THE UNIVERSITY OF QUEENSLAND

AUSTRALIA

Probing the sub-thalamic nucleus: development of bio-markers from very Local Field Potentials

By

John Varghese

B. Eng. (Hons), B. Sci, MBBS

A thesis submitted for the degree of Doctor of Philosophy at The University of Queensland in 2016

School of Mechanical and Mining Engineering

Abstract

Deep Brain Stimulation (DBS) surgery for neuro-psychiatric disorders involves the insertion of Micro Electrode Recording (MER) targeting probes into a specific location in the patient's brain to confirm the precise location of the candidate nucleus for the stimulation therapy. The unique geometry and electrical properties of these MERs probes, when inserted into neuron dense tissue such as the Sub-thalamic Nucleus (STN) results in the acquisition of signals which typically contain contributions from the spiking behaviour of multiple nearby neurons in addition to a low frequency component similar to the Local Field Potentials (LFP) generated by more distant neurons (over length scales smaller than typical LFPs). We refer to these signals acquired from the MER probes, which contain both some of the nearby (resolvable) spikes and these smaller scale LFPs as very Local Field Potentials (vLFPs).

The unique signal contributions to the vLFPs raise the immediate question what contribution of the signal is best used in order to characterise (in terms of identifying the underlying physiology or detecting changes of) the state of the STN. In this thesis we develop methodologies to analyse vLFPs using both model based and model free analysis of the entire (nearby spiking, distant spiking and non-spiking contributions) vLFP and only the nearby spiking neurons.

We apply concepts from Mori-Zwanzig non-equillibrium kinetic theory to develop model-free estimates of the entire vLFP. With this approach we show that the Non-Markov Parameter (NMP) can be used to identify statistically significant changes in the electrical behaviour of the STN when presented with different neuro-linguistic stimuli. We show that these changes are due to variations in the low frequency bands associated with the power spectrum of the vLFP.

We then develop a model-based analysis of the entire vLFP using the renewal theory of stochastic processes. We show that when the ensemble of neural processes forming the vLFP satisfy the assumptions of an independent renewal process model, given a measurement of

the power spectrum, the common probability distribution driving the spiking statistics of the individual neurons can be identified. We show with simulation that, when the assumptions of the model are satisfied, this approach can outperform state of the art spike sorting algorithms in the challenging situation of identifying the spiking statistics when an unknown number of neurons with near identical spike shapes contribute to the vLFP.

Finally we develop spike-only analysis by constructing spike sorting strategies using convex optimisation and clustering theory to identify the precise timing and shapes of the spikes associated with individual neurons nearby to the MER probe. With this approach the spike detection and clustering is performed using a Basis Pursuit De-Noising (BPDN) strategy which is a subset of ℓ_1 regularised least squares techniques. We show that this method outperforms state of the art spike sorting algorithms for a range of signal to noise ratios. We use this method to identify Poisson counting statistics with average firing rates between 20-56 Hz for the STN neurons in patients with Parkinson's Disease during DBS surgery. These results are consistent with previous studies.

The results of these methods show that information contained in the different scales of the vLFP can successfully be used to assist in characterising the state of the STN in patients with Parkinson's Disease. We conclude that both the BPDN spike sorting methodology and the NMP approach offer robust performance and require minimal a priori assumptions. The BPDN approach provides excellent physiological insight (identifying nearby spiking shapes and timing) whereas the NMP provides a well-tested discriminator of vLFP electrical state.

We recommend that future work explore whether these metrics can be used as biomarkers which correlate to the degree of disease state in Parkinson's disease. The development and identification of these biomarkers which provide continual and real time analysis of the patient state will provide a crucial component for the development of future adaptive DBS systems which can minimise battery expenditure, offer superior symptom amelioration and provide treatment to patients who have thus far been considered refractory to pharmacological or neuromodulation therapy.

Declaration by author

This thesis is composed of my original work, and contains no material previously published or written by another person except where due reference has been made in the text. I have clearly stated the contribution by others to jointly-authored works that I have included in my thesis.

I have clearly stated the contribution of others to my thesis as a whole, including statistical assistance, survey design, data analysis, significant technical procedures, professional editorial advice, and any other original research work used or reported in my thesis. The content of my thesis is the result of work I have carried out since the commencement of my research higher degree candidature and does not include a substantial part of work that has been submitted to qualify for the award of any other degree or diploma in any university or other tertiary institution. I have clearly stated which parts of my thesis, if any, have been submitted to qualify for another award.

I acknowledge that an electronic copy of my thesis must be lodged with the University Library and, subject to the policy and procedures of The University of Queensland, the thesis be made available for research and study in accordance with the Copyright Act 1968 unless a period of embargo has been approved by the Dean of the Graduate School.

I acknowledge that copyright of all material contained in my thesis resides with the copyright holder(s) of that material. Where appropriate I have obtained copyright permission from the copyright holder to reproduce material in this thesis.

Publications during candidature

- J.J. Varghese, P.A. Bellette, K.J. Weegink, A.P. Bradley P.A. Meehan, "*Analysis of the non-Markov parameter in continuous-time signal processing*, Phys. Rev. E., vol 89(**2**), p. 022109, 2015.

- J. Varghese, K. Weegink, P. Bellette, T. Coyne, P. Silburn, and P. Meehan, "*Theoretical & Experimental Analysis of the Non Markov Parameter to Detect Low Frequency Synchronisation in Time Series Analysis*" in Proceedings of the 33rd Annual International Conference of the IEEE EMBS, Boston, Massachusetts USA, August 30-September 3,2011, 2011, pp. 1500-1505.

- Paul A. Meehan, Paul Bellette, Andrew P. Bradley, Joanna E. Castner, Helen J. Chenery, David A. Copland, John Varghese, Terry Coyne and Peter Silburn (2011). *"Investigation of the non-markovity spectrum as a cognitive processing measure of deep brain microelectrode recordings"*. In: F. Babiloni, A. Fred, J. Filipe and H. Gamboa, Proceedings of the International Conference on Bio-Inspired Systems and Signal Processing. BIOSIGNALS 2011 - 4th International Conference on Bio-Inspired Systems and Signal Processing, Rome, Italy, (144-150). 26-29 January 2011.

- Varghese, J.J., Weegink, K.J., Bellette, P.A. and Bradley, A.P. (2015). *"Spectral properties of neuronal pulse interval modulation.* In: 2015 IEEE International Conference on Acoustics, Speech, and Signal Processing: Proceedings. International Conference on Acoustics, Speech, and Signal Processing (ICASSP) Conference, Brisbane, QLD, Australia, (1007-1011). 19-24 April 2015. doi:10.1109/ICASSP.2015.7178121.

- Kristian J. Weegink, Paul A. Bellette, John J. Varghese, Peter A. Silburn, Paul A. Meehan and Andrew P. Bradley, *"A Parametric Simulation of Neuronal Noise from Microelectrode Recordings"* Transactions on Neural Systems and Rehabilitation Engineering, IEEE Transactions on, 2015.

- Kristian James Weegink, Varghese, John Joseph and Bradley, Andrew P. (2015). *"Spikes from compound action potentials in simulated microelectrode recordings"*. In: 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2015 International Conference on Acoustics, Speech and Signal Processing, South Brisbane, QLD, Australia, (813-816). 19-24 April, 2015. doi:10.1109/ICASSP.2015.7178082.

- Weegink, K. J., Varghese, J. J., Bellette, P. A., Coyne, T., Silburn, P. A. and Meehan, P. A. (2012). *"An efficient stochastic based model for simulating microelectrode recordings of the deep brain"*. In: Proceedings of Biosignals 2012, International Conference on Bio-Inspired Systems and Signal Processing. 5th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC), Vilamoura, Portugal, (76-84). 1-4 February 2012.

- Weegink, K. J., Varghese, J, Bellette, P., Coyne, T., Silburn, P. and Meehan, P. (2012). *"Simulating micro electrode recordings from deep brain surgery for applications in a clinical framework"*. In: Abstracts of WFN XIX World Congress on Parkinson's Disease and Related Disorders. WFN XIX World Congress on Parkinson's Disease and Related Disorders, Shanghai, China, (S225-S225). 1114 December 2011.

Publications included in this thesis

1. J.J. Varghese, P.A. Bellette, K.J. Weegink, A.P. Bradley P.A. Meehan, "*Analysis of the non-Markov parameter in continuous-time signal processing*, Phys. Rev. E., vol 89(**2**), p. 022109, 2015. *Incorporated as part of the theory & experiment sections of chapter 2*

Contributor	statement of contribution
J.J. Varghese (Candidate)	Development of Concept (50%) Implementation (70%) Analysis of Results (90%) Writing of Paper (100%) Reviewing of Paper (70%)
P.A. Bellette	Development of Concept (30%) Implementation (10%)
K.J. Weegink	Implementation (20%) Analysis of Results (10%) Analysis of Results (90%)
A.P. Bradley	Reviewing of Paper (20%)
P.A. Meehan	Development of Concept (20%) Reviewing of Paper (10%)

2. J. Varghese, K. Weegink, P. Bellette, T. Coyne, P. Silburn, and P. Meehan, "*Theoretical & Experimental Analysis of the Non Markov Parameter to Detect Low Frequency Synchronisation in Time Series Analysis*" in Proceedings of the 33rd Annual International Conference of the IEEE EMBS, Boston, Massachusetts USA, August 30-September 3,2011, 2011, pp. 1500-1505. *Incorporated as part of the theory & experiment sections of chapter 2*

Contributor	statement of contribution
J. Varghese (Candidate)	Analysis of Data (70%) Writing of Paper (70%) Reviewing of Paper (60%)
K. Weegink	Analysis of Data (10%) Writing of Paper (10%)
P. Bellette	Analysis of Data (20%) Writing of Paper (20%) Reviewing of Paper (20%)
T. Coyne	Acquisition of Data (80%)
P. Silburn	Design of Experiment (100%) Acquisition of Data (20%)
P. Meehan	Reviewing of Paper (20%)

3. Varghese, J.J., Weegink, K.J., Bellette, P.A. and Bradley, A.P. (2015). *"Spectral properties of neuronal pulse interval modulation.* In: 2015 IEEE International Conference on Acoustics, Speech, and Signal Processing: Proceedings. International Conference on Acoustics, Speech, and Signal Processing (ICASSP) Conference, Brisbane, QLD, Australia, (1007-1011). 19-24 April 2015. doi:10.1109/ICASSP.2015.7178121. *Incorporated into Appendix 7*

Contributor	statement of contribution
J.J Varghese (Candidate)	Development of Concept (50%) Implementation (70%) Analysis of Results (70%) Writing of Paper (100%) Reviewing of Paper (50%)
K.J. Weegink	Implementation (10%) Analysis of Results (10%)
P.A. Bellette	Reviewing of Paper (10%)
A.P. Bradley	Development of Concept (50%) Implementation (20%) Analysis of Results (20%) Reviewing of Paper (40%)

Contributions by others to this thesis

None

Statement of parts of the thesis submitted to qualify for the award of another degree

None

Acknowledgements

Completing this thesis would not of been possible without the love, support and friendship of many who assisted me practically, financially and emotionally over one of the most difficult journeys I have ever undertaken. In many ways the success of this thesis is less a reflection of my technical abilities or determinedness and more a consequence of the professional and personal relationships I have been lucky enough to forge. The long list of names should not be considered a dilution of my appreciation to any individual, the gratitude I hold for each person mentioned below cannot adequately be put into words.

Paul Meehan: for providing me with such an interesting project, funding to conduct this research and a much longer leash than is afforded to most PhD students to tackle the problem, shape the questions and make my own mistakes attempting to answer them. It has been an amazing learning experience.

Peter Silburn: the progenitor of this research project whose infectious enthusiasm sparked my interest in the clinical side of medicine and continues to provide an excellent role model as a clinician-researcher. Thank you also for all your help and encouragement in getting me to Oxford in order to learn from giants in the field of Deep Brain Stimulation like Prof. Tipu Aziz.

Andrew Bradley: who had to juggle his time both with left field technical questions and providing continual moral support whenever I would (frequently!) come to him as an apostle of doom proclaiming the project was impossible, infeasible and destined for failure. I am incredibly grateful for the combination of technical, moral and financial support you gave me well outside of office hours, across multiple continents and coffee shops across Brisbane. Without your help I would not have been able to complete this thesis.

Paul Bellette: the greatest supervisor I never (officially) had. No single person has been so pivotal in helping me complete this thesis. I can never adequately express the profound

gratitude I have for the level of technical and moral support I have received from you over the last 6 years. Even when you were no longer part of the research group you still made time for me at ungodly hours despite juggling full time work and raising small children. If it was not for your help I can unequivocally say that I would not have been able to complete this thesis.

Thank you to my thesis committee chairmen: Prof. Martin Veidt and Dr. Bill Daniel for your help in guiding the development of this thesis. Thank you to both the Australian federal government and Medtronics for providing me the funding to undertake this PhD. Thank you to the school of Mechanical and Mining Engineering, especially Brianne Mackinnon, Kristin Greer and Katie Gollschewski for providing a welcoming, enthusiastic and encouraging place to undertake my research. I would also like to thank my school, my supervisors and my thesis review committees for providing the best bulwark they could against a sadly emerging bureaucratic university culture which seems to prize meeting milestones and arbitrary deadlines over actually conducting interesting, difficult and often time-consuming research.

Nathan Walk: an expert in quantum communications by choice, and an expert in signals processing in neuroscience due to my persistent badgering. I dont think there has been a single journal article or major mathematical theorem in my thesis which we didnt discuss in your office (usually unannounced), over beers in pubs I didnt know existed around Brisbane or inconveniently located gourmet food trucks.

Kristian Weegink: my brother in arms on this project and the only other person who truly understood what an odyssey this PhD was. I couldnt even begin to calculate the hours we spent together trying to shape our theses and get a conceptual grasp on what seemed like one insurmountable technical problem after the next.

To the boys and girls at Ronin BJJ especially my mentors Jarek Janczewski, Anthony Jacobs, Paul (Old Paul) Navruk and my coach John Walken: for giving me an outlet for much needed stress relief and giving me something else to obsessively channel my thought processes into. I feel like less of a fraud about being called ‘Doctor John’ now.

To my dorogaya Uliana Systerova thank you for your unconditional love, affection and support, for your understanding in how much time I had to put into this thesis and for tolerating such a neurotic, stressed out and moody boyfriend for the last 2 years.

To my family: my wonderful parents, my sister and my grandparents, who I dedicate this thesis to, thank you for all your emotional and financial support especially while I began to lose hope in completing this PhD during the demanding clinical years of my medical degree. I could not of completed this thesis without your help.

Keywords

Micro Electrode Recording, neural signal processing, spectral analysis, deep brain stimulation, Parkinson's Disease, spike sorting, ℓ_1 minimisation.

Australian and New Zealand Standard Research Classifications (ANZSRC)

ANZSRC code: 090303 , Biomedical Instrumentation 60%

ANZSRC code: 090609 , Signal Processing, 40%

Fields of Research (FoR) Classification

FoR code: 0903, Biomedical Engineering, 60%

FoR code: 0102, Applied Mathematics, 40%

Contents

List of Abbreviations

AR	Auto Regressive
ARMA	Auto Regressive Moving Average
ASRDF	Asymptotically Shifted Renewal Density Function
AUC	Area and the Curve
BPDN	Basis Pursuit De-Noising
CBP	Continuous Basis Pursuit
CLDBS	Closed Loop Deep Brain Stimulation
CoV	Coefficient of Variation
CWT	Continuous Wavelet Transform
DALM	Dual Augmented Lagrangian Multiplier
DBS	Deep Brain Stimulation
DC	Direct Current
DPIM	Digital Pulse Interval Modulation
ECAP	Evoked Compound Action Potential
EEG	Electroencephalogram
FFT	Fast Fourier Transform
FT	Fourier Transform
GIC	Generalised Information Criteria
GLE	Generalised Langevin Equation
GPi	Globus Pallidus internal
i.i.d	independent identically distributed
ISI	Inter Spike Interval
KDE	Kernel Density Estimate

KKT	Karush Kuhn Tucker
LARS	Least Absolute Regression
LASSO	Least Absolute Selection and Shrinkage Operator
LFP	Local Field Potential
LIF	Leaky Integrate and Fire
LMM	Linear Mixed Model
LT	Laplace Transform
LZC	Lempel-Ziv Complexity
MA	Moving Average
MAD	Median Absolute Deviation
MDS-UPDRS	Movement Disorder Society Unified Parkinson's Disease Rating Scale
MER	Micro Electrode Recording
MFR	Mean Firing Rate
MUA	Multiple Unit Activity
NMP	Non-Markov Parameter
NSR	Noise to Signal Ratio
OLS	Ordinary Least Squares
PCA	Principal Component Analysis
PCG	Preconditioned Conjugate Gradient
PD	Parkinson's Disease
PDF	Probability Density Function
PSD	Power Spectral Density
RDF	Renewal Density Function
RIP	Restricted Isometry Property
ROC	Receiver Operating Curve
S-ASRDF	symmetric Asymptotically Shifted Renewal Density Function

SDE	Spectral Density Estimator
SHO	Simple Harmonic Oscillator
SPC	Super Paramagnetic Clustering
S-RDF	symmetric Renewal Density Function
STN	Sub Thalamic Nucleus
TN, FN	True Negative, False Negative
TNIP	Truncated Newton Interior Point
TP, FP	True Positive, False Positive
TPR, FPR	True Positive Rate, False Positive Rate
vLFP	very Local Field Potential
$\mathbf{ZF - NMP_1}$	Zero Frequency of first Non Markov Parameter

List of Figures

List of Tables

1

Introduction

"When we talk mathematics, we may be discussing a secondary language built on the primary language of the nervous system."

– John von Neumann

Therapy for treatment resistant neuro-psychiatric disorders, such as Parkinson's Disease, Tourettes & Obsessive Compulsive Disorder involves a procedure known as Deep Brain Stimulation (DBS) [2], [3], [4]. In the United States DBS was approved by the Food and Drug Administration (FDA) for essential tremor in 1997, Parkinson's Disease in 2002 and was granted humanitarian device exemption for dystonia in 2003 and treatment resistant Obsessive Compulsive Disorder in 2009 [5]. In DBS the skull and meninges are breached in a keyhole neurosurgical operation and electrodes are placed into a specific deep brain nucleus and a high frequency square wave current (typically $\geq$ 130 Hz [6]) is consistently

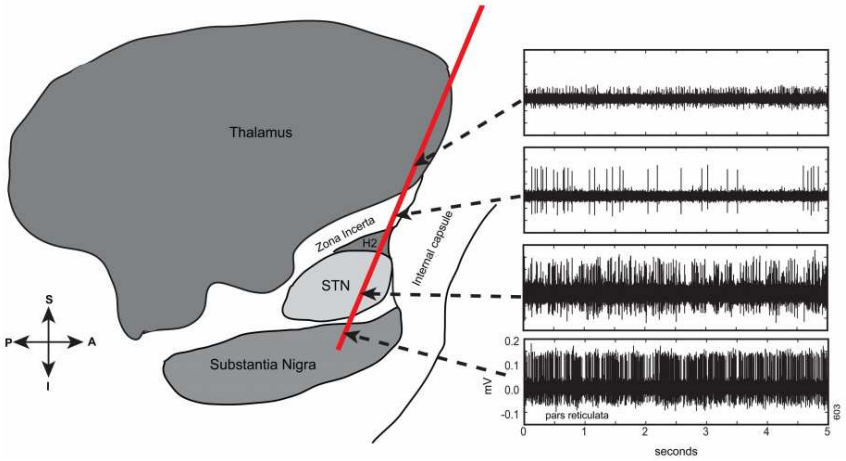


Figure 1.1: Micro Electrode Recordings of different nuclear structures on an approach of the STN. In this image the electrode is passed from the thalamus through the Zona Incerta, through the second field of Forel (H2) to the STN and then finally out through to the Substantia Nigra. Notice that as the STN is approached the background noise increases, and the firing pattern of the neurons becomes more irregular and 'burst' like. Identification of this pattern can be used by the neurosurgeon to confirm the location of the Sub Thalamic Nucleus. Image from [1].

applied. In many cases this undeniably radical treatment modality been shown to provide the best symptomatic relief to these debilitating disorders. For example it has been shown that DBS in conjunction with best medical therapy is superior to best medical therapy alone for alleviating the symptoms of Parkinson's Disease [7] [8],[9].

The need for treatment approaches such as DBS is further underscored by the rising incidence and costs associated with these neuro-psychiatric disorders treatable with stimulation therapies. In Australia alone it is estimated that as of 2014 there were over $65,000$ people suffering from Parkinson's Disease. This is an increase in incidence of $5,100$ more patients than in 2011 and $14,500$ more than in 2005. Recent studies [10] have suggested that 50% of Parkinson's Disease patients (excluding those presenting with atypical varieties of the illness) are good candidates for DBS. It is estimated that worldwide over $140,000$ people have benefitted from DBS [11]. The health care (due to aged care facilities, pharmaceuticals,

hospital, pathology and imaging) costs in Australia in 2014 were estimated at over $567 million and total productivity costs (due to disease impact on both patients and their carers) were estimated at over $182 million [12]. This is in addition to the incalculable impact of pain, suffering, psychological distress and reduction in quality of life of those suffering from Parkinson's Disease, their families and their carers. When we further consider that this is only one disease which can be treated with DBS, the therapeutic power of this treatment modality and the motivation for researching methods to further optimise symptom control becomes evident.

The nuclei which are targeted with DBS, and whether the targeting is bilateral or unilateral, is determined by the underlying pathology and the clinical judgement of the neurosurgical team. Typical sites for stimulation for Parkinson's Disease are the Sub Thalamic Nucleus (STN) or the Globus Pallidus Internus (GPi) [2]. Sites stimulated for Tourette's Syndrome are the centromedian-parafascicular (CM-Pfc) and ventralis oralis complex of the thalamus [3]. For Obsessive Compulsive Disorder the nucleus accumbens [4] is stimulated. These nuclei are remarkably small. For example, the size of the STN is approximately 6.5 x 7.8 x 9 (W x L x D) mm [13] which is similar to the size of a pea. Targeting these nuclei is a difficult neurosurgical operation because these small structures deep within the brain must be targeted without damaging vital adjacent neuro-vascular structures which can lead to strokes, speech and motor deficits or death.

There are multiple ways to target the candidate nucleus but a commonly used method [14] is a two stage approach of pre-operative imaging using fused X-Ray Computerized Tomography (CT) and Fluid Attenuated Inversion Recovery (FLAIR) Magnetic Resonance Imaging (MRI) scans for the neurosurgeon to plan the surgical approach and then peri-operative analysis using Micro Electrode Recordings (MER) to confirm the target nucleus. This use of electrical recordings is employed because certain nuclei generate characteristic electrical signals which the surgeon can use to confirm the location of the target nucleus to within sub millimetre accuracy [1] [15], [16]. For example, when targeting the STN, insertion into the STN is associated with an increase in background noise and the appearance

of irregular burst firing neurons [17],[16]. This change in the behaviour of the electrical signal is shown in Figure 1.1. In the case of DBS surgery for Parkinson's Disease, once the location of the candidate nucleus (described for the STN in [14]) is identified the recording tip is withdrawn and the distal sheath provides high frequency ($\geq$ 130 Hz) stimulation to this targeted nucleus. The awake patient [1] is then examined for improvement of the clinical signs of bradykinesia, tremor and rigidity and an absence of adverse (typically speech and language) effects. These MERs are then withdrawn and the permanent, larger ($\sim$ 1 mm) stimulating electrodes are inserted.

This thesis is concerned with analysing the one dimensional voltage signals obtained from the single channel MER probes which are used to confirm the target nucleus. These electrodes are relatively low impedance (0.5 $M\Omega$ @ 1kHz) tungsten probes but have a relatively large 500 μm diameter conducting tip. The size of this tip is much larger and the impedance much smaller than the electrodes used for recording the precise activity of single neurons ($\sim 1-40\mu m$, $1M\Omega$ [15]). Conversely these MERs also have tips with comparable surface areas but higher resistance to the so called 'macro electrodes' used to measure Local Field Potentials (LFPs) which have diameters $\sim 50-1270\mu m$ [18] and impedances of $\sim 1-100k\Omega$ @ 1kHz [15]. For the remainder of this thesis we will refer to the smaller high impedance probes which measure single unit activity as *single unit probes*, the larger low impedance probes as *macro electrodes* and the larger radius but lower impedance probes used to confirm the candidate nucleus as *micro electrodes*.

The electrodynamics of the interaction between the neural field and the probe tip is highly non-trivial, but we can utilise two basic principles of circuit theory to guide our understanding of what biological processes we expect to identify with these different recording probes:

1. The first principle is that the larger the surface area of the probe, the greater the charge it can accommodate on the surface and thus the capacitance will be greater. This larger capacitance will generate a low-pass filtering effect on the recording probe

[1] The brain does not contain nociceptors (pain receptors), so local anaesthetic can be applied to the scalp and the patient can be woken up mid operation with minimal distress!

and therefore will preferentially record the lower frequency aggregate activity of the neuronal populations. This component of the signal is the LFP. There are a multitude of neuro-biological processes which contribute to the LFP. Indeed this is an active area of research as the precise biophysical origins of the LFP are not completely understood [19]. See [20] for an excellent review of these processes contributing to the LFP. Essentially, the LFP consists of both spiking and non-spiking events. Non-spiking contributions include synaptic currents [21],[20], and voltage-dependent membrane oscillations [22]. The spiking contribution to LFPs includes afterhyperpolarization of nearby neurons [19], [23] and the spiking patterns of more distant neurons (the electric fields of which have been heavily attenuated due to the low-pass filtering properties of the extra-cellar fluid [24],[25] and the geometry of the dendrites [20]). Thus the LFP records information over a large spatial distribution [26].

Therefore, we expect that the macro electrodes which have a large surface area and low impedance to primarily record these LFPs. We also expect that the micro electrodes which have large surface area but (relatively) high impedance to also record a contribution (over a smaller spatial distribution) from the LFPs. These results are indeed seen experimentally in [15],[27] and specifically for the STN in [28].

2. The second principle is that the larger the impedance, the lower the current source the probe can record for a given voltage. This is easily seen as a consequence of Ohm's law. Thus, we expect that the higher the impedance of the probe - the lower the number of electrical sources which will contribute to the recording. Therefore, we expect that the single unit probes having high impedance and low surface areas will, as their name suggests, primarily record the contribution from a minimal, usually single number of neurons nearest to the probe. We also expect that the micro electrodes with their higher impedance than the macro electrodes, but larger surface area than the single unit recordings, will record the contribution from a potentially larger group of nearby, resolvable neurons in addition to recording a contribution of the LFPs. This result is also seen experimentally in [15], [28],[29].

For all intents and purposes, these different electrodes can be understood as follows. The single unit probes are excellent at recording the precise electrical signal of a very localised space of brain tissue which, with appropriate operator skill may be single neuron behaviour. We refer to these as *single unit recordings*. The larger macro electrodes record an aggregate of neuronal activity (the *LFP*) over a much larger spatial scale which is too complex to identify the behaviour (e.g. a single neuron) of the contributing processes. The signals, analysed in this thesis, acquired from the MER probes to confirm the location of the candidate nucleus lie in between these two extremes. They will record the behaviour of a few neurons ($\sim 10 - 100$) in a small area ($\sim 10 - 100\mu m$) [15] in addition to the background neural processes. We refer to these signals measured by the MERs which contain both resolvable single unit recordings and significant contributions to the LFP as *very* Local Field Potentials (vLFPs). It is important to realise that there has been much work showing that correlations exist between the aggregated LFP activity and the spiking activity [19],[30],[27],[31] and as stated in [20]: "*spike 'contamination' of the LFP should be regarded as good news, in that high-frequency LFP power can provide a 'proxy' for the assessment of neuronal outputs*". In a very broad sense, vLFPs therefore may contain similar information over the different scales of the single unit recordings and the component of the LFPs.

It is important to take a step back and appreciate both the breadth and depth of the complexity of the 'system' we are analysing compared to the vLFP measurements we are obtaining with the MERs. We are obtaining *in-vivo* recording from the human STN, a structure which has a size on the order of millimetres [13] with a single channel probe which measures the electrical activity on the order of hundreds of micrometres [15]. Thus, we are not even simultaneously recording the entire STN at any given time. The neurons in the STN will have an exquisitely complicated set of synaptic dendritic connections to each other which we cannot hope (especially in the in-vivo situation) to identify [32]. The collection of neurons comprising the STN will have extensive connections to the other collections of neurons comprising other nuclei in the Basal Ganglia [33]. The Basal Ganglia itself will have a multitude of connections to other nuclei, collections of nuclei and cortical structures [34]. The dynamical evolution of each of these neurons is *approximated* by the Hodgkin-Huxley

equations [35] which are four coupled non-linear differential equations which are capable of exhibiting nonlinear behaviour such as bifurcations [36], mode locking [37] and chaotic oscillations [37],[38]. The kinetics of the ion channels which regulate the flow of ions and drive the membrane voltages are intrinsically stochastic [39]. Therefore, there is complexity in the breadth of how the neurons being measured are connected to each other (both neurons being measured and not explicitly measured) and complexity in the depth of how the electrical behaviour of a single one of these neurons evolves.

Given these recordings, a fundamental question is: what can we expect to identify from these Micro Electrode Recordings ? This is more a question of philosophy than of engineering science. Clearly we cannot expect from these single channel recordings, to completely characterise the evolution of all the 10^{11} neurons in the measured brain. The problem is intrinsically ill posed. Can we completely identify the firing patterns of a single neuron close to the recording probe? Can we identify the firing pattern of a collection of neurons near the probe? Can we characterise the behaviour of at least a portion of the STN with the LFP signal? Which part of the measured signal should be analysed? For the small and large stimulating recording probes the component of the signal which should be analysed is dictated by the electrodynamics of the measured process. The single neuron recordings (if correctly placed) can only provide information about the precise firing patterns of the single identified neuron. For the larger macro electrode probes, the averaged activity of the behaviour of the lower frequency LFP is collected and patterns can be attempted to be identified. For the MERs considered in this thesis, the vLFPs will record a combination of the multiple (and in theory individually resolvable) spiking patterns of individual neurons and the spatially averaged activity (although over a smaller length scale compared to the stimulating probes) of the lower frequency LFP behaviour.

Given the complexity of the problem at hand, another question is: why should we attempt to identify information about the signal over different scales? Why bother attempting to isolate and cluster the firing times of individual neurons? Why should we try to characterise the patterns observed from the LFPs? These are more direct questions of engineering science.

We argue that this multi-scale analysis of the vLFPs can be used to help construct Closed Loop DBS (CLDBS) stimulation devices to improve the outcomes of patients suffering from the wide range of neuro-psychiatric disorders we have discussed. We discuss this motivation further in the next section.

1.1 Thesis Motivation

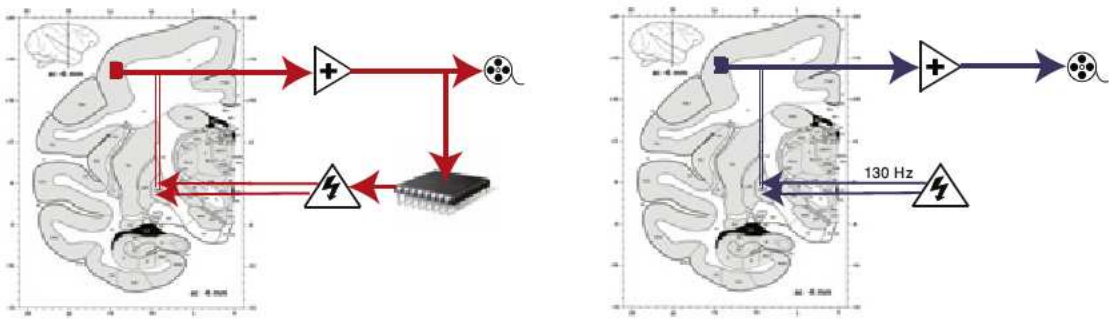


Figure 1.2: Schematic difference between closed loop (left) and open loop (right) control of DBS pacemakers. In the open loop configuration (current DBS scheme) the stimulation parameters remain the same independent of the MER signal. In closed loop control the stimulation parameters would alter depending on the MER signal. Image from [40].

The motivation for this work extends beyond exploring the academically interesting problem of attempting to develop classification schemes for a network of an unknown number of neurons, with unknown but dense and widely distributed connections which individually display non-linear dynamics and evolve together with firing patterns best described as stochastic processes. The analysis of the vLFPs associated with these MERs can be used to construct biomarkers in CLDBS strategies. For the remainder of this thesis we restrict our consideration of biomarkers and CLDBS for the treatment of Parkinson's Disease. Nonetheless the general principles discussed can potentially be applied to the development of CLDBS for any neuro-psychiatric disorder. In this section we do the following:

- Introduce CLDBS and describe its theoretical superiority to current open loop DBS therapy.

- Discuss the need for the identification of biomarkers for CLDBS algorithms.
- Argue that analysis of the vLFPs provides a natural, robust and practical framework for the construction of these biomarkers.

DBS for Parkinson's Disease is currently implemented as an open loop strategy where chronic stimulation is provided independent of the recorded state [2]. The problem with this approach, compared to an adaptive DBS stimulating device which could intelligently vary (or turn off) the delivered stimulation is the sizeable energy expenditure (resulting in an increased rate of required surgical battery replacements), the onset of more side effects due (in part) to the large volume of tissue stimulated and stimulation therapy which becomes less effective (due to static stimulation parameters) as the disease progresses. Since Parkinson's Disease is a progressive illness [41] there is a need for the stimulation parameters (stimulation frequency, pulse width and amplitude) to be varied by an expert (guided by clinical experience) at 3-12 monthly intervals [40]. It is perhaps not surprising that there is improved patient outcomes with more frequent consultations [42]. It is attempting to solve these problems and utilise the advantages of frequent stimulation parameter variation which has motivated researchers to explore CLDBS protocols.

In CLDBS the stimulation parameters (and indeed whether stimulation is provided at all) are determined in real time by some controller based on the physiological state of the measured system. In a sense, the CLDBS represents the parameter changes by the clinician in the limit of updates in differential steps of time, with the parameter choice dictated by *signal processing* not by clinical acumen and the parameter variation performed by a microprocessor not a medical specialist. The difference between an open and closed loop DBS system is illustrated schematically in Figure 1.2. An ideal CLDBS controller would use information about the state of the patient to only provide stimulation when clinically necessary, provide the minimal stimulating current required and constantly update the stimulation parameters to provide optimal symptomatic relief. The advantages of CLDBS would therefore include:

[2] There are some newer devices which allow the user the option to turn the stimulation on and off and vary the intensity of injected current within a range preset by the treating clinician.

constant adaptation to deal with habituation and tolerance to the parameter settings [43], minimisation of side effects with lower amplitude current stimulation, maximisation of battery life (and subsequent delay in operations to replace the battery), reducing the burden on the health care system by reducing required neurology consultation visits and introducing treatment strategies which work for people who have thus far been refractory to open loop DBS.

Central to the development of these CLDBS systems is the need to characterise the physiological state of the patient in order to determine *when* the stimulation should be provided and with *what* stimulation settings. In classical control engineering we would refer to this as identifying the state estimator of the system [44]. These physiological characterisations of the patient are referred to as *biomarkers* [5], [43], [45]. There is a very large class of biomarkers which could be employed to 'measure' the state of the patient for a CLDBS algorithm and they can be broadly classified as *internal* or *external* biomarkers [43]. The internal biomarkers are based on direct measurements of the neuro-dynamics and candidates include the LFP, individual spiking patterns, electroencephalograms, magnetoencephalography, near infrared spectroscopy and functional magnetic resonance imaging [43] and even chemical analysis of neurotransmitter metabolites [46]. The external biomarkers are based on measurements of the clinical sequelae Parkinson's Disease such as accelerometers [47] or even video based monitoring systems [48] [43] which observe the motor pathology.

For a thorough review of candidate biomarkers for closed loop DBS see [43], [45]. In effect, all the candidate biomarkers have advantages and disadvantages. Nonetheless, we argue on engineering grounds that the biomarkers for CLDBS based on MERs (either the LFP or the individual neuronal spikes) have the most practical advantages for the reasons outlined below:

- The vLFPs measured with the MERs provide real time recordings with high temporal resolution.

- The MERs provide continuous recordings which are readily available. This is in contrast to other considered biomarkers such as fMRIs, infrared spectroscopy, accelerometers, EEGs or video based monitoring systems which are only feasible to record for discrete periods of time in the hospital or laboratory setting.

- The MERs which measure the vLFP provide robust and relatively cheap sensors compared with the chemical sensors, such as the silicon electrodes with microdialysis tubes considered in [49], which as discussed in [45] may be susceptible to tube blockage and difficulty in miniaturizing the chemical analysis hardware for implementation in permanently implanted devices. This is in addition to the concerns regarding the temporal resolution of these chemical sensors [45].

- There is a large body of work which has explored whether patterns in the signals measured by MERs can detect Parkinson's Disease pathology. There has been considerable interest exploring whether the emergence of β band (10-35Hz) synchrony in LFPs is linked to the motor disorders of Parkinson's Disease [50], [51], [52], [53], [54], [55], [56].

- Prototype CLDBS systems using biomarkers derived from components of the vLFP have previously been constructed using the power in the β band of the LFP as the biomarker [56] and using identified single unit events in deep brain and cortical nuclei as the biomarker [40].

- Since the MER probes are also capable of providing stimulation we argue that a CLDBS system which can provide both stimulation and recording (although not necessarily simultaneously) from the same location would provide a simplified design and minimise neuro-surgical incisions. This will decrease surgical operation times, decrease costs associated with the operations and minimise adverse outcomes such as breaching vital neuro-vascular structures. In addition a single stimulating and recording site is likely to be more robust to errors introduced by effects such as electrode drift [57].

Past construction of biomarkers based on MERs for CLDBS systems have followed two distinct paths, utilising different components of the vLFPs which we discuss below:

Low frequency LFP content: The first approach is based on the previously mentioned approach of removing β $(10 - 35$ Hz) band peaks from the power spectrum of the measured LFPs.

In [56] a CLDBS system which used the amplitude in the β band (measured in real time from the stimulating electrode) as the biomarker to determine when to deliver stimulation to the STN of eight human patients with advanced Parkinson's Disease was developed. The details are provided in [56], but we briefly discuss the methodology of this CLDBS system and its clinical outcomes. For each patient the frequency within the β band with the greatest amplitude was determined, and the LFPs measured from the stimulating electrode were filtered in real time using a Moving Average around this frequency. Whenever the energy at this specific frequency in the β band went above a predetermined threshold, a DBS pulse (at 130 Hz) of predetermined amplitude was applied until this energy fell below the threshold. When the energy in this specific frequency of the β band was below the threshold value, the stimulator remained quiescent. This CLDBS system was compared to the standard open loop DBS system, and a DBS system which randomly applied stimulation with the same parameters as the CLDBS system. This random stimulation DBS system was developed such that, on average, it was in the stimulating and quiescent phases for the same lengths of time as the CLDBS system. This CLDBS showed marked improvement in the tremor, rigidity and bradykinesia of patients compared to the open loop and randomly stimulating DBS systems, as judged by both blinded and unblinded clinicians using the Unified Parkinson's Disease Rating Scale (UPDRS), and (unsurprisingly) less energy usage than the open loop DBS system. Interestingly, as the CLDBS stimulations progressed the emergence of β bursts above threshold became less frequent.

In [58] a theoretical closed loop controller was constructed where the input-output relationship between the stimulation current and the resulting LFPs was determined

using an Auto Regressive Exogenous (ARX) model where the parameters were determined by minimising the mean squared prediction error using a Recursive Least Squares (RLS) algorithm. A generalised minimum variance control law was implemented which attempted to determine the current amplitude that would provide minimum mean squared error between the predicted next value of the LFP (using the ARX model) and a predetermined reference LFP (indicative of no tremor) defined using an Auto Regressive (AR) model. The parameters of the ARX model were updated online using the RLS algorithm. The success of the controller was judged on how similar the power spectra from the LFPs resulting from the neurons exhibiting tremor stimulated with the closed loop DBS protocol were to the spectra from the LFP of the 'tremor free' neurons.

The advantage of working with the LFP is that it requires minimal signals processing effort to obtain properties such as the peaks in different spectral bands. The disadvantage of this approach is that due to the complexity of LFPs there are few 'simple models' of their behaviour and thus it is difficult to predict how the measured LFP will evolve when stimulated with a specific current. For example, in [58] statstical ARX models without recourse to the underlying biology were employed.

High frequency individual spike content: The second approach is based on identifying the firing patterns of individual neurons near the recording probe. This involves removing the lower frequency LFP content and employing signals processing intensive spike sorting algorithms to identify action potentials and grouping them to the correct neuron.

In [40] a physical closed loop DBS controller was tested on 'Parkinsonised' green rhesus monkeys where a pulse train (of seven or one pulses) of stimulation was applied to the GPi following detection of a spike at an electrode in the primary motor cortex or the GPi. This closed loop scheme was shown to provide superior improvement of the akinesia symptoms, reduce GPi neuron discharge rates, alter the discharge pattern

and remove or attenuate oscillatory activity (as measured with wavelet spectrograms) of neurons at the motor cortex and GPi.

The work of [40] raised interesting questions about *where* biomarkers based on MERs should be measured from. It was shown that superior amelioration of akinesia was demonstrated when the stimulus was applied to the GPi but the signal was measured from the motor cortex rather than simultaneously recording and stimulating the GPi. Furthermore, recording from a site distant to the location of stimulation has advantages in terms of minimising stimulus artefact [43],[45] but signals processing techniques can be applied to reduce this [59]. We argue that the benefits of recording from a secondary location must be balanced against the increased surgical risk of introducing a second incision site in an area of such neuro-vascular sensitivity. We also mention that biomarkers based on estimation the LFP are largely robust to the presence of stimulation [43] because the low frequency content of the LFP [20] has minimal overlap with the high frequency stimulation.

More recently [5] have used advanced multi electrode hardware (originally used for spinal cord stimulation) to record the response of neural tissue in the area of stimulation with minimal latency ($\leq$ 0.5 milliseconds) after stimulus onset. This work showed that the morphology of the late ($\sim$ 3 milliseconds after stimulation) response of Evoked Compound Action Potentials (ECAPs) correlated well with therapeutic relief from rigidity symptoms, and it was suggested that this shape could be used as a CLDBS biomarker.

The advantage of working with the individual neuron firing patterns is that the observed behaviour is well understood in terms of the underlying biology and relatively simple phase oscillator (typically leaky integrate and fire) models of neurons can be used to predict how the firing patterns of the measured neurons will respond to input stimulus current. The disadvantage of this approach is that it requires more intensive

signals processing in order to identify the spikes and group them to their appropriate neuron.

In this section we have introduced CLDBS and justified why it would provide superior therapeutic treatment than the currently available chronically applied open loop DBS. We have suggested that the construction of a CLDBS system will require a biomarker to act as an estimator of the state to determine the specific stimulation parameters when the CLDBS is activated. We have argued that one of the most practical, robust and time tested biomarkers for a CLDBS system is based on the vLFPs obtained from MERs. We have surveyed the literature and shown that previous biomarkers based off MERs have used information across different length scales: using the low frequency contribution from the LFP to identify β band oscillations and using the high frequency contribution from the individual spike trains of nearby neurons.

1.1.1 Thesis Motivation In A Nutshell

In effect, the analysis of the STN using transformations of the vLFP over different scales (individual neuron behaviour or spectral bands of the LFPs) may generate biomarkers which can assist in the development of *future* closed loop DBS algorithms. These adaptive DBS systems offer the exciting prospect of simple (from a surgical perspective) devices which record and stimulate from the same deep brain region which provide superior clinical amelioration of Parkinson's Disease symptoms, minimal side effects and have longer lasting battery life.

1.2 Thesis Hypothesis

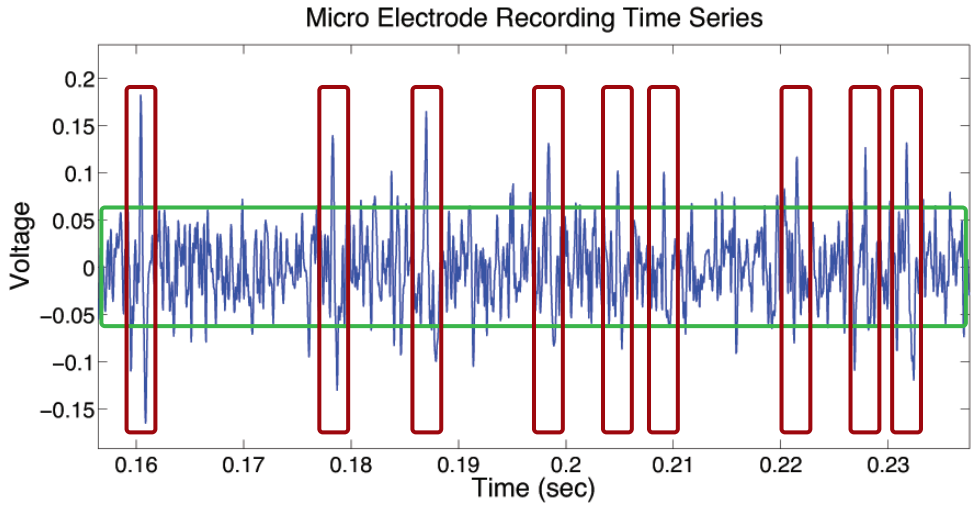


Figure 1.3: Micro Electrode Recording of a human Sub Thalamic Nucleus from a patient with Parkinson's Disease. Notice that the signal detected by the MER probe consists of both individual, resolvable spikes of nearby neurons (red boxes) and the Local Field Potential (green box) which is the localized spatially averaged aggregate of distant spiking neurons and non spiking neuro-biological processes. We refer to this signal as the Very Local Field Potential (vLFP).

The research question addressed with this thesis is to explore whether *vLFPs measured from single channel targeting MERs during DBS can be used to characterise the state of the human Sub Thalamic Nucleus.* The concept of 'characterising the state of the STN' refers to the development of transformation operations applied to the MER which can either detect changes in the electrical activity of the STN or reveal information about the underlying activity of some subset of the constituent neurons in the STN. The recordings which we analyse in this thesis were obtained from patients with Parkinson's Disease. We envisage that these identified transformations used to 'characterise the state of the STN' could be considered in future work as potential biomarkers for CLDBS strategies. The previously discussed electrodynamic nature of the MER is such that the vLFP is unique (compared to macro electrodes and single unit recordings) insomuch that it contains contributions from *both* the *LFP* and (potentially multiple and resolvable) spike trains of *single neurons* near

to the recording probe. This contribution over multiple scales is shown in Figure 1.3. Thus in order to explore this problem we analyse the vLFP on multiple different scales:

1. Model-free analysis of the entire vLFP (spikes + LFP) using complexity metrics.

2. Model-based analysis of the entire vLFP (spikes + LFP) assuming it consists of an ensemble of similar, independent, stochastic processes.

3. Spike-only scale of identifying the precise firing statistics of the neurons closest to the recording probe.

More specifically, the entire vLFP is analysed in a model-free framework using the non-Markov family of parameters and a model-based approach where the measured system is modelled as an ensemble of independent renewal processes. The spike-only scale is analysed using a spike sorting approach based on applying Basis Pursuit De-Noising to separate and sort the spiking behaviour of the closest neurons from the entire measured signal.

1.3 Thesis Approach

In this thesis we aim to analyse the vLFP signals measured from the MER over the different scales below:

- Model-free entire vLFP (LFP + neuronal spikes) analysis using the Non-Markov Parameter. This approach requires minimal a priori assumptions but provides little insight into the underlying neuro-physiology of the recorded process.

- Semi-parametric Model-Based entire vLFP (LFP + neuronal spikes) analysis using transformation operators in the spectral domain. In this approach the measured neurons are modelled as an ensemble of independent stochastic renewal oscillators. This approach introduces a series of a priori assumptions but uses all the information in the entire signal and provides excellent insight into the underlying spiking statistics of the contributing neurons (if the assumptions are satisfied).

- Spike-only analysis using Basis Pursuit De-Noising: This approach introduces minimal a priori assumptions, and provides deep neuro-physiological insight by identifying the spike times and shapes of the observed resolvable neurons. The disadvantage of this approach is that only the behaviour of the neurons nearest to the probe are considered while the LFP component of the signal is not analysed.

The scales we choose to analyse the vLFP over, from the broadest (model-free entire vLFP) approach to the most localised (spike-only) analysis span the range of scales which were considered for past biomarkers based on the MERs for CLDBS systems in section 1.1. At each of these scales we characterise the signal using a transformation operator derived from solving for the inverse of different linear response problems. The methodology we use to analyse the vLFPs over the different scales is described in more detail in the following section. We delay more technical descriptions of these approaches until section 1.4.

It is important to stress that this analysis over multiple scales of the vLFP is performed from MER signals acquired from the (previously discussed) temporarily inserted targeting electrodes, not the permanently inserted chronically stimulating electrodes.

1.3.1 Entire vLFP (LFP + Neuronal Spikes) Model-Free Analysis

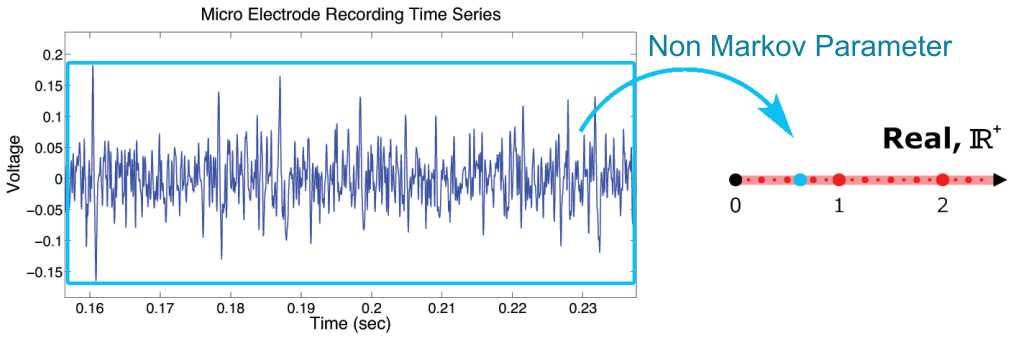


Figure 1.4: Analysis of the entire vLFP (spikes + LFP) using the Non-Markov Parameter. The entire one dimensional vLFP consisting of *both* the LFP and the neuronal spikes is mapped to a single positive number on the positive real number line.

Starting with the *broadest* model-free entire vLFP scale - we use *all* of the signal available to us: the individually resolvable spikes and the spatially averaged LFP, without recourse to relying on models for these components which reduce the validity of this approach in the general case. This approach of analysing the entire signal in a model-free framework certainly contains the most information available about the neural state, but the price for this is the difficulty of unravelling what this information is telling us about the underlying neuro-biology. It is unlikely that we can identify the precise, *individual* behaviour of any of the constituent neural processes contributing to this signal with this approach, but we can hope to detect patterns in the signal. That is detect, either directly or by transformations of the signal, the presence of some emergent phenomena in the bulk behaviour (at the level of the MER probe tip) of these strongly interacting neurons. This is in a sense the philosophy of other clinical techniques such as Electroencephalograms (EEGs) where repeatable electrical rhythms can be identified from the ensemble behaviour of billions of neurons measured by only dozens of electrodes at the level of the scalp [60], or the Electrocardiogram (ECG) where cutaneous electrical measurements of millions of myocytes and electrical pacemaker cells generate predictable complexes, with changes that can be identified with the onset of specific pathologies. [61]

In previous model-free approaches to analyse vLFPs, the LFP is separated from the spiking activity detected in the MER by low-pass filtering. The LFP is then characterised by some transformation operator. Typically this operator attempts to characterise the low-pass filtered one dimensional signal by a single number. The most popular linear transformation approach has been to assume stationarity of the signals and analyse the power spectrum to observe how the energy of the signal is distributed across frequencies. This approach has yielded insight into the association of peaks in the β (13-30Hz) band of the power spectrum of the LFP and associated movement disorders in Parkinson's Disease [50],[51],[52],[53]. More specifically, there is evidence that peaks in the β band are associated with akinesia/bradykinesia [51],[54], whereas suppression of these peaks may be associated with tremor [55]. As with all LFP analysis it is difficult to relate these identified changes of the β band to the underlying biology. The current consensus opinion is that the rise in β band activity is related to synchronous oscillatory activity in the cortico-basal ganglia circuits [62].

More sophisticated approaches based on nonlinear transformation operators have also been applied to low-pass filtered MER signals. Bispectral analysis of the STN LFPs of patients with Parkinson's Disease was introduced in [63], which identified a correlation between low β (13-20 Hz) and high β (20-35 Hz) bands which was attenuated with the medical treatment of Levodopa. Lempel Ziv Complexity (LZC) analysis of the STN LFPs in the β (13-35 Hz) band of Parkinson's Disease patients was performed in [64], which showed that there was a strong negative correlation between the LZC and the degree of akinesia-rigidity as measured by the Unified Parkinson's Disease Rating Score (UPDRS).

The approach we undertake to analyse the system on this entire vLFP (spikes+LFP) scale is to perform a nonlinear transformation using the Non-Markov Parameter, which is developed from the Mori-Zwanzig framework [65] of non-equilibrium statistical mechanics. This approach belongs to the subset of *statistical complexity metrics*, which is a measure of the correlation structure of an interacting system and its subsets [66]. The motivation for using the NMP is based on its previous success of differentiating states of complex biological systems, such as healthy and post-myocardial infarction hearts based on ECG recordings

[67] and, more saliently for this thesis, attempting to differentiate Parkinsonian tremor under different treatment modalities (Levodopa and DBS) based on signals obtained from velocity transducing laser recordings [68].

Using the NMP, the one dimensional vLFP signal is mapped to a single number on the positive real line, $\mathbb{R}^+$, as shown in Figure 1.4. It is shown that the transformation applied by this mapping requires solving for the inverse of the linear integro-differential Langevin equation. We show that this metric can be used to differentiate the state of the STN in patients when presented with different neuro-linguistic stimuli during lexical decision making tasks. We also show with aid of the Weiner-Khinchtine theorem that this NMP complexity metric has an explicit link to the transformation operators based on feature selection of the power spectrum.

The primary advantages of this model-free analysis of the entire vLFP is two-fold. Firstly this approach uses the entire signal, both the spiking structure of nearby spikes and the neural aggregate LFP behaviour to characterise the state of the STN. In a general sense, using the entire signal guarantees that the contribution of the entire neural ensemble being measured is used to characterise the state of the system, not just the neurons that happen by chance to be closest to the probe. Secondly,this approach is model-free and thus is reliant on a minimal set of assumptions. The disadvantage of this approach is also the model-free nature, as it is not possible to relate any changes identified to the behaviour of the underlying neurons.

1.3.2 Entire vLFP (LFP + Neuronal Spikes) Model-Based Analysis

This approach is based on trying to incorporate information from all of the vLFP (spikes + LFP) *and* relating the observed signal to the underlying physiology. In order to perform this step we must introduce a model of the portion of the STN analysed and relate this to either the measured signal or the signal under an appropriate transformation. The model we introduce is to approximate that the rich behaviour of the entire (LFP and neuronal spikes)

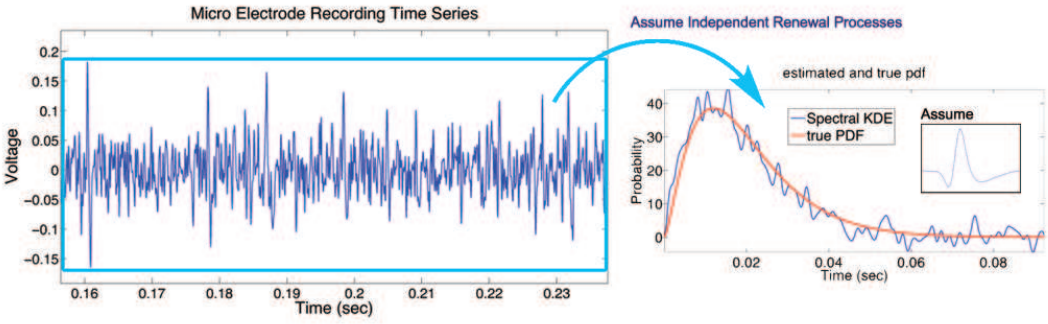


Figure 1.5: Analysis of the entire vLFP (spikes + LFP) by modelling the measured system as a stochastic ensemble of independent filtered renewal processes with similar spike shapes. Using spectral analysis the one dimensional time series is mapped to the probability distribution driving the renewal process.

strongly interacting neuronal system being measured is dominated by the spiking patterns of the nearby neurons (the neuronal spikes) and the more distant neurons (the filtered neuronal spikes contributing to the LFP). More specifically, we model the portion of the STN measured by the probe tip as an ensemble of independent stochastic processes following renewal statistics with similar (but linearly filtered) spike shapes.

The use of renewal theory in neuroscience is widespread in both experiment [69], [70], [71] and theory [72], [73], [74]. There is evidence that the discharge patterns of neurons in certain anatomical zones of lower order mammals such as the spike trains from the retinal ganglion cells to the Lateral Geniculate Nucleus (LGN) of the thalamus in cats [75] and the responses of neurons in the Antero-Ventral Cochlear Nucleus [76] follow renewal statistics. More recently, [77] has shown that the dopaminergic neurons of the Substantia Nigra Pars Compacta in rhesus macaques follow different classes of renewal statistics while performing saccadic eye movements in response to different visual stimuli. The popularity of renewal theory models, where the gap between firing times are statistically independent, is consistent with the Hodgkin-Huxley Na^+,K^+ channel model, where the equations governing these variables is reset after each spike [72]. Obviously, this model is not always valid and there are many examples in the literature where non-renewal statistics have been identified. See

[78] for a detailed review of these examples and proposed mechanisms for the generation of non-renewal statistics.

There has been an interest recently in modelling the behaviour of neurons in models of sparse neural networks using renewal theory [79],[80],[81]. In these works the statistical behaviour of a single neuron in the network is analysed with the other neurons connected in the network altering its voltage by a sequence of synaptic bombardments which follow either renewal [79] or pooled renewal [80] statistics. In these works the analysed neuron is modelled as a Leaky Integrate and Fire (LIF) neuron. It is important to realise in this model that the non-spiking effects (i.e. the non-spiking component of the LFP) are ignored. The model we incorporate for the neural field-probe tip interaction is similar to this described work, but instead of considering the voltage evolution of a single neuron in the network we consider the voltage evolution of the probe tip embedded in the network. Also, instead of considering the inputs to a modelled neuron being a barrage of synapses from connected neurons which follow pooled renewal statistics, we consider the inputs to the modelled probe tip as the voltages generated from the barrage of electric fields, propagating through the extra-cellular fluid, associated with the action potentials of the nearby neurons.

The methodology we employ to develop this model-based transformation operator, termed the *Spectral Density Estimator*, is described in chapter 3. Briefly, we mention that we transform the measured signal to its power spectrum and then use the Weiner-Khincthine theorem to identify the systems' correlation structure. The assumptions of the ensemble of renewal oscillators are then introduced and specific mathematical properties of renewal processes are used to construct the probability distribution describing the timing between the spikes from the power spectrum.

Using these approximations we introduce a series of (inverse) transformation operators to map the observed signal to the probability distribution driving the duration of the gaps between spikes for a single spike train, referred to as the Inter-Spike-Intervals (ISI). Notice that this effectively maps the one dimensional measured electrical signal to the one dimensional

Probability Distribution Function (PDF) describing the ISI. This methodology is illustrated in Figure 1.5. This PDF can then be used to characterise the measured system by construction of the appropriate statistical moments of the ISI such as the mean firing rate, the coefficient of variation or the information entropy associated with the spike times (assuming they exist). Notice that similar to the model-free approach using the NMP this approach maps the one dimensional measured electrical signal to a (series of possible) single number(s).

This model-based entire vLFP approach can be considered the 'middle ground' between the entire vLFP model-free approach using the NMP, which uses the entire measured signal and the spike-only approach which, as the name suggests, only uses information from the spikes of the closest neurons. In this approach the entire vLFP is used, but it is assumed that the dominant effect is the spiking activity (both of the nearest neurons and the spiking component of the LFP). The advantage of this approach is being able to relate the entire measured signal to the underlying physiology of the PDF driving the spiking patterns of the neurons without explicitly detecting the spikes and measuring their ISI. The disadvantage is that unlike the other two methods, there are a series of a priori assumptions introduced in this model-based approach which will not always be necessarily valid. Indeed, in section 3.5.4 we identify degenerate cases where this method does not work at all.

1.3.3 Spike-Only analysis of the vLFP

In this approach we consider the spike-only scale analysis, characterising the vLFP by the identified spike times of the neurons closest to the recording probe. One of the primary objectives of neuroscience is attempting to understanding how neurons encode information [82],[83], [84], [85], [86], [87]. Due to the *assumed* stereotyped shape of the action potential associated with a neuron, it is frequently assumed that the information is encoded in the patterns of the spike trains [88], [89], [85], [90]. This philosophy of the importance of timing is all pervasive through neuroscience: ranging from the debate of rate [88] or temporal [89] codes for information transmission of a stimulus to Hebbian learning models where neural

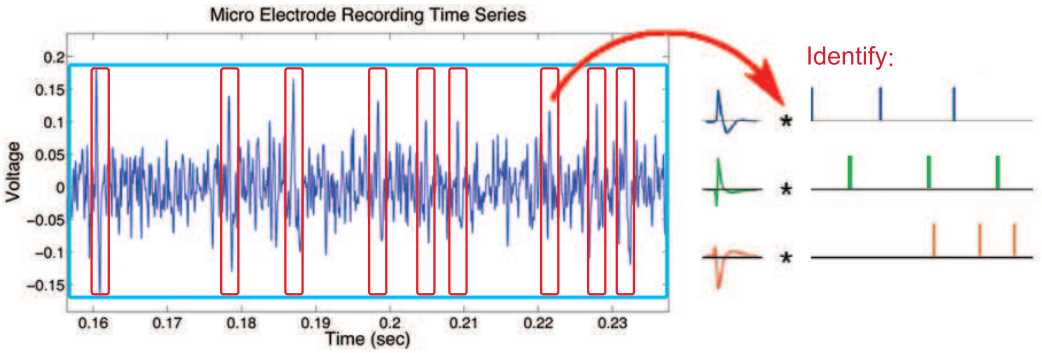


Figure 1.6: Analysis of the subsection of the vLFP (neuronal spikes only) using spike sorting. Using the spike sorting algorithm the one dimensional time series is mapped to a sequence of identified spike times and associated spike shapes.

connections strengthen based on increased firing rates [91] to Spike Timing Dependent Plasticity [92] where connections between neurons are strengthened or weakened based on the order of pre and post-synaptic firing times within small timing windows.

In order to identify the spike times associated with the neurons around the probe tip we use spike sorting algorithms. These algorithms strip all of the information from measured signal away except for the times and shapes of the different action potentials identified which are then appropriately clustered into groups associated with individual neurons. This approach is illustrated in Figure 1.6. The problem of spike sorting in extracellular recordings (such as the recordings considered in this thesis) is a highly non-trivial problem [93]. There are multiple spike sorting algorithms available for single channel recordings (many with extensions to multi channel recordings) which all have limitations with regard to what spikes they can identify amongst noise [94], how well they can be correctly clustered into their correct groups [95] and how many of the different neurons they can actually detect [96]. We review multiple different spike sorting algorithms and how they perform the task of spike sorting in chapter 4.

The problem with spike sorting is that it is almost always easiest to detect and separate spikes from neurons which have larger amplitudes and higher firing rates (as better estimates of the shape can be constructed). Unfortunately, sometimes the neuron conveying the most

information about a process or a stimulus is a quiescent or sparsely firing neuron [93] which may have a reduced amplitude because it is further away from the recording probe than other neurons. Thus, a constant problem in spike sorting is developing more accurate spike sorters which can detect more than just the neurons associated with the loudest and largest action potentials.

The methodology we employ to generate a more accurate spike sorting algorithm is to introduce techniques from Basis Pursuit De-Noising (BPDN), which are a subset of ℓ_1 minimization techniques, that have shown great success in constructing accurate and *sparse* estimates to linear problems in signals processing [97],[98] statistics [99] and compressed sensing [100],[101]. The motivation for applying these BPDN methods is that if we assume a linear model for the measured spike trains and certain conditions described by the Restricted Isometry Property (see section 4.4) are satisfied, this method is guaranteed to give the optimally sparse accurate estimation of the different neuron spiking times.

The advantage of this spike-only approach is that we are able to understand the underlying biological processes we are measuring by accurately identifying the firing patterns of the nearby individual neurons in a (nearly) model-free fashion with minimal a priori requirements. The disadvantage of this approach is that the majority of the information in the vLFP (the LFP portion) is discarded and only the behaviour of the nearest neurons (which happen to be nearest to the probe purely by chance) are recorded. We mention however that, as discussed in [96], these spike sorting methodologies which are developed for single channel recordings can easily be applied to multiple recording probe systems [102], where different regions in a similar area (e.g. different regions of the Sub Thalamic Nucleus) are recorded concurrently. Applying these accurate spike sorting algorithms to all the individual probes in the system could provide spike timing information for potentially hundreds of different neurons in the system. This would allow for a microscopic classification (in terms of spike times and shapes) of neural structures over *meso-scopic* length scales using minimal a priori assumptions about the measured system.

1.4 Thesis Structure Overview

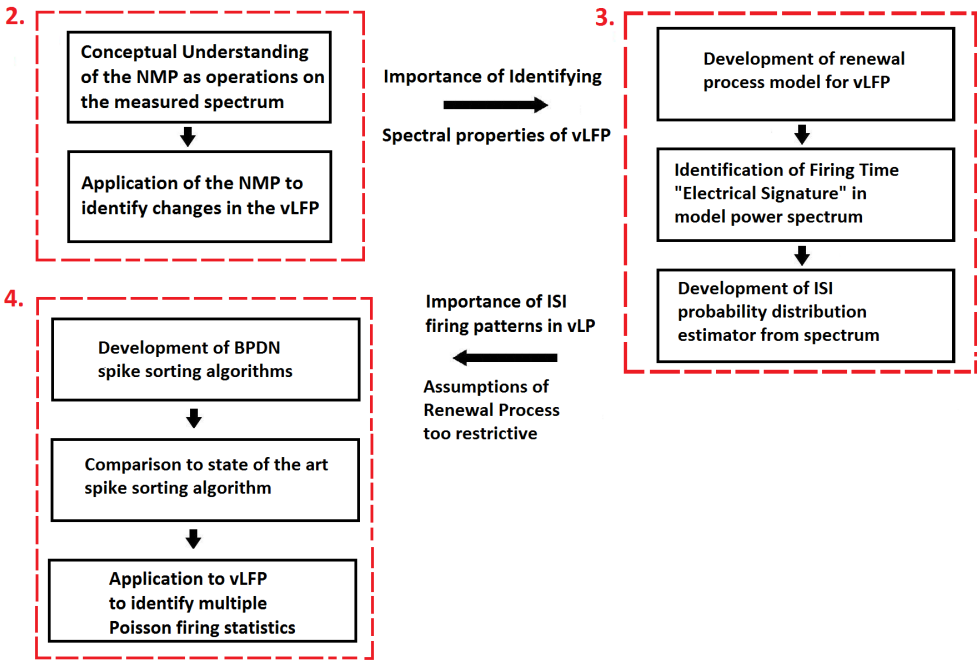


Figure 1.7: Breakdown of the thesis chapter content and their inter-relationships.

The organisation of the subsequent chapters and their relationship to each other is shown in Figure 1.7. The content of these chapters is summarised below:

Chapter 2 develops the entire vLFP model-free approach discussed in Section 1.3.1. Specifically, the following is performed:

- The family of *Non-Markov Parameters* (NMP) from the Mori-Zwanzig theory of non-equilibrium statistical mechanics is introduced. This family of parameters is analysed in a signals processing framework, showing with the Weiner-Khinchtine theorem that the set of NMP for a system can be constructed as nonlinear functions of the measured power spectrum. This is achieved in part by solving an inverse problem based on the integro-differential Langevin equation. It is then argued that only the first NMP contains useful information about the measured system.

- Closed form expressions for the zero frequency value of the first NMP is determined for a series of instructive physical systems in order to identify parametric sensitivity.
- The NMP in different frequency bands (zero frequency, β (13-35Hz) and fast (80-200Hz) bands) is applied to vLFPs obtained from MERs of the STN obtained from patients with Parkinson's Disease while undergoing neuro-linguistic testing. It is shown, with two separate experimental data sets, that the NMP is able to identify different neural states under the different experimental conditions.

Chapter 3 develops the entire vLFP model-based approach discussed in Section 1.3.2. Specifically, the following is performed:

- Renewal stochastic processes are introduced and their application to neuroscience is discussed.
- The spectral properties of renewal processes and pooled independent renewal processes is discussed. It is the spectral properties of these pooled processes which forms the basis for the model-based analysis of this chapter.
- The *Spectral Density Estimator* (SDE) which allows estimation of the ISI PDF of an ensemble of independent renewal oscillators given an estimate of the power spectrum is developed and discussed. The construction of this estimator requires the serial solution of two linear inverse integral equations of renewal theory.
- The SDE is compared to the spike sorting algorithm *Osort* on simulated datasets of vLFPs. It is shown through extensive Monte-Carlo simulation that the SDE is superior to the classical spike sorting algorithm Osort when the individual spike trains are renewal processes and the action potentials of different neurons are sufficiently similar.
- In the Appendix it is shown that the renewal theory model of a neuronal spike train is equivalent to the Digital Pulse Interval Modulation (DPIM) coding scheme (frequently used in fibre-optics and communications theory) in the limit of the allowed pulse times occurring on the continuous time line.

Chapter 4 develops the spike-only scale discussed in Section 1.3.3. Specifically the following is performed:

- An overview of spike sorting, its canonical sub-problems (spike detection, spike estimation, feature selection and clustering) and state of the art spike sorting algorithms are discussed.
- The Basis Pursuit De-Noising (BPDN) framework for sparse estimation of spiking times from MERs is motivated.
- An overview of BPDN is provided and the three BPDN algorithms (homotopy, InCrowd with Truncated Newton Interior Point and Dual Augmented Lagrangian Method) which are subsequently analysed are discussed.
- The dictionary terms necessary to drive the BPDN algorithm using a sequential three step process of Continuous Wavelet Transformation, Diffusion Mapping and Mean Shift clustering are introduced.
- The developed BPDN spike sorting algorithm is shown by extensive Monte-Carlo simulation with Receiver Operating Characteristic (ROC) plots and Chi squared (χ^2) analysis to be superior to the state of the art spike sorter *wav-clus* when analysed on simulated MER data.
- The developed BPDN spike sorting algorithm is then applied to the vLFPs obtained from MERs inserted into the STN of patients with Parkinson's Disease prior to DBS surgery. It is shown using this algorithm that the analysed patients have Poisson firing statistics with average firing rates between 20-56 Hz, which is consistent with previous analysis.

Chapter 5 develops the conclusions to the findings of the previous three content chapters and identifies limitations, recommendations and extensions for future work.

1.5 Thesis Scope

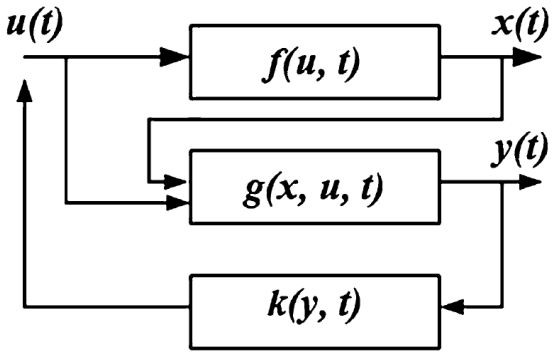


Figure 1.8: Structure of a CLDBS system described in [43]. $u(t)$ is the input current stimulation. $x(t)$ is the state of the neural system. $y(t)$ is the biomarker of the neural system under the transformation $g(x, u, t)$. $k(y, t)$ is the feedback signal based on the biomarker $y(t)$. This thesis is *only* concerned with identifying *potential* biomarkers $y(t)$ under more specific transformations $g^{\star}(x, t)$.

This thesis has been motivated within a frame-work of developing biomarkers for potential use in CLDBS systems. It is important to realise that the construction of a CLDBS system is a highly non-trivial problem, in terms of the development of requisite underlying theory, experimental validation and practical implementation, which is well beyond the scope of a single thesis. In this thesis we restrict our analysis to characterising the state of the STN (both in terms of identifying changes in the electrical activity or describing the behaviour of the individual constituent neurons) by developing transformations of measured vLFPs obtained from MERs. We *do not* consider:

- How the DBS stimulating current affects the neural state ($f(u, t)$ in Figure 1.8), nor the structure of the feeback controller based on the biomarker ($k(y, t)$ in Figure 1.8).
- The effects of stimulating current in our transformation models (therefore the transformations $g(x, u, t)$ considered in 1.8 are restricted to the sub-class $g^{\star}(x, t)$). Nonetheless, with sufficiently sophisticated hardware recording devices (such as introduced in [5]) the biomarkers developed in this thesis should be valid either during or immediately after (< 0.5 milliseconds) the stimulating current is provided.

- Correlations between our identified transformations of the measured signal and clinical pathological state of patients with Parkinson's Disease or the success of specific stimulation settings.

Nonetheless, the metrics and transformations identified and evaluated within this thesis can be considered as providing some of the ground work for the development of potential biomarkers for future CLDBS strategies.

1.6 Contribution of Thesis

This thesis provides a systematic analysis of vLFPs measured with MERs in the STN over different length scales, introducing new methods to analyse these recordings and applying methods which had previously not been applied to MERs. More specifically:

- For the model-free entire vLFP approach (chapter 2), the family of NMPs is applied to deep brain recordings for the first time and the previous complex descriptions of these metrics using the complex language of non-equilibrium statistical mechanics is condensed to simpler explanations based on nonlinear spectral analysis. The NMP is then used to identify signal changes in the human STN when presented with different neuro-linguistic stimuli.

- For the model-based entire vLFP approach (chapter 3), a new method of analysis based on *Spectral Density Estimator* is introduced. It is shown that given the assumptions of this model are satisfied, this estimator can outperform state of the art spike sorting algorithms when the spikes associated with different neurons are sufficiently similar. It is also mathematically shown that if the information transmission of a neuron is modelled as a Digital Pulse Interval Modulation encoder, where the information is encoded in the time *between* the spikes, that the power spectrum of this process in the limit of the spikes being allowed to occur (as biologically expected) at any time on the continuous timeline is equivalent to a renewal process.

- For the spike-only approach (chapter 4), a new spike sorting algorithm based on *Basis Pursuit De-Noising* is developed. It is shown that this algorithm can outperform state

of the art spike sorting algorithms in a range of signal to noise ratios. It is then shown that this algorithm can successfully be applied to the *in-vivo* MER to identify the spike shapes and times associated with different neurons.

2

Model-Free Entire vLFP Analysis

"With four parameters I can fit an elephant, and with five I can make him wiggle his trunk."

– John von Neumann

2.1 Chapter Summary

This chapter introduces the model-free approach to characterising the entire vLFP (spikes + LFP) of a single channel MER. This method is based on the family of Non-Markov Parameters (NMP) developed from non-equillibrium statistical mechanics. The solution of the family of NMP values requires solving an infinite sequence of deconvolution problems based on the estimated auto-correlation function. This will be a consistent theme throughout this thesis of characterising the vLFP in terms of a transformation which requires the solution

of inverse problems.

We show that this family of parameters can be constructed in terms of a modular set of operations on the measured power spectrum of the studied system. We suggest by theory and example that only the first NMP has any useful value in complexity analysis, eliminating the need to solve the infinite set of NMPs. We show that the first NMP is a function of both the spread of the power spectrum and the ratio of the spectrum to its amplitude envelope. We show that the zero frequency component of this parameter, frequently used as a metric of discrimination between states of complex systems is solely a function of the spread of the measured power spectrum and the DC offset. Thus, we show that using this parameter is effectively equivalent to performing spectral feature selection.

We apply this NMP metric to the entire vLFP from MERs of the Sub Thalamic Nucleus of patients suffering from Parkinson's Disease. We identify changes in the low frequency components ($\leq$ 200Hz) of the NMP during different states of neuro-linguistic testing. Using the theoretical analysis of the NMP, we identify that these changes are indicative of variations in the measured power spectrum of the vLFP. The results of this chapter suggest that the low frequency spectral behaviour of the Sub Thalamic Nucleus changes with different neuro-linguistic stimuli, and that the NMP can be used to detect these changes.

2.2 Chapter Overview

In this chapter we introduce the Non-Markov Parameter which has previously been used to analyse many-body systems where the underlying dynamics driving the system are too complex to be predicted. The use of this family of metrics requires solving an inverse problem associated with the integro-differential equations which describe the Mori-Zwanzig kinetic equations. This approach of using the NMP to characterise the vLFP will represent the first of three approaches based on solution of an inverse problem to characterise the neural signal acquired from a single channel recording.

In this chapter we perform the following:

1. Provide a brief overview of the applications of statistical complexity metrics (which the NMP belongs) to dynamical systems.

2. Introduce the Mori-Zwanzig kinetic equations which underpin the NMPs.

3. Show that the family of NMPs can be understood as a set of closed-form expressions that only depend on a nonlinear set of integral transformations on the measured signal's power spectrum. We then show that the first NMP contains the most useful information about the measured system, and that the higher order NMPs veil the underlying correlation structure of the measured system.

4. Show that the zero frequency value of the first NMP, ZF-NMP_1, is solely a function of the DC offset and spread of the measured power spectrum.

5. Apply the ZF-NMP_1 to four instructive systems. The first three systems are analytical models of a Simple Harmonic Oscillator driven by white noise, band limited white noise and the output of white noise into an idealised all-pole low-pass filter. Analytical calculation of the NMP identifies a primary sensitivity to the decay rate of the tail of the power spectrum. The ZF-NMP_1 is then numerically calculated from the time series of a second order Auto-Regressive process driven by white noise. We show that the numerically determined NMP is in agreement with the theoretically expected

value. This example provides a 'real world' application, similar to our neural signals problem, of estimating the non-Markovity spectra where the underlying continuous power spectra are unknown and only discrete samples of the measured time series, corrupted with noise, are available.

6. Lastly we use the ZF-NMP$_1$ to identify changes of *in-vivo* vLFPs obtained from MER probes from the left and right STN of patients with Parkinson's Disease when they are presented with neuro-linguistic stimuli. We show that this metric, applied to the datasets of two separate neuro-linguistic experiments, is able to differentiate the signals at different time of stimulus presentation, the side of the brain recorded and the structure of the neuro-linguistic stimuli.

Submitted Work

The theory portion of this chapter (sections 2.3-2.8) is based on the following journal article:

- J.J. Varghese, P.A. Bellette, K.J. Weegink, A.P. Bradley P.A. Meehan, "*Analysis of the non-Markov parameter in continuous-time signal processing*, Phys. Rev. E., vol 89(**2**), 022109, 2015.

The experiment portion of this chapter (sections 2.9-2.10) is based on the following conference articles:

- J. Varghese, K. Weegink, P. Bellette, T. Coyne, P. Silburn, and P. Meehan, "*Theoretical & Experimental Analysis of the Non Markov Parameter to Detect Low Frequency Synchronisation in Time Series Analysis*" in Proceedings of the 33rd Annual International Conference of the IEEE EMBS, Boston, Massachusetts USA, August 30-September 3,2011, 2011, pp. 1500-1505.

- P. Meehan, P. Bellette, A. Bradley, J. Castner, H. Chenery, D. Copland, J. Varghese, T. Coyne, and P. Silburn, "*Investigation of the Non-Markovity Spectrum as a Cognitive Processing Measure of Deep Brain Microelectrode Recordings*,International Conference on Bio-inspired Systems and Signal Processing, Rome, Italy, pp. 144-151, 2011.

Contribution

The contribution of this chapter is the development of expressions for the NMP in a signals processing framework which shows that the NMP is simply a series of nonlinear functions of the measured power spectrum. We then go on to apply different spectral bands of the NMP to the vLFPs from MER obtained from in vivo human STN, showing that different states can be identified using the entire signal (not just a subset set of the closest identified spikes) using a novel parameter free methodology.

2.3 Introduction

Time series analysis is often employed to characterise systems where the generating physics is either too complex or involves too many degrees of freedom to be predicted. These systems frequently arise in financial [103],[104] and biological [64],[105],[67],[106] systems. One of the earliest metrics used to characterise a system is the Shannon information entropy which ascribes a numerical score to the system based on the randomness of the underlying statistics driving the process [107]. This methodology, originally used in communications theory, was adapted to nonlinear deterministic dynamical systems with the use of Kolmogorov-Sinai entropy [108]. While successful for the analysis of chaotic systems this approach can fail to detect the statistical simplicity of random behaviour [66]. This has motivated the development of "statistical complexity metrics" which measure the correlation structure of an interacting system and its subsets [109], allowing for the analysis of multi degree of freedom probabilistic systems.

This methodology has generated a plethora of statistical complexity metrics, often with ambiguous relationships to each other, which often do not provide a clear interpretation of what the metric is actually measuring [109]. The application of these complexity metrics to time series analysis has thus seen the emergence of an interesting phenomenon where signals from complex systems can be successfully differentiated, but there is little insight into the nature of these differences.

An excellent example of this problem has been the application of the generalised Non-Markov Parameter (NMP), which is effectively a complexity metric developed from the Mori-Zwanzig theory of non equillibrium statistical physics [110]. The NMP has been used in a diverse range of fields such as geology [111], astrophysics [112], cardiology [67] and neurophysiology [106]. In these complex systems the NMP has been developed as an informational tool to analyse the degree of randomness or "Markovity" of the system. Particular attention has been paid to the Zero Frequency value of the first order NMP (ZF-NMP$_1$) [67],[111],[113],[114] which

in a similar sense to the Shannon information entropy, maps the Markovity of a system interacting with its environment to a scale from unity for non Markov processes (the transition to the next state is history dependent) to infinity for purely Markov processes (the transition to the next state is history independent) [110]. In specific applications the quantification of this randomness has proved useful as a metric of discrimination between different states in complex systems. In section 2.9 we use the ZF-NMP$_1$ to differentiate states of vLFPs from MERs of the STN of patients with Parkinson's disease during neuro-linguistic processing tasks [115], [116].

Previous papers have derived the NMP for measured systems from discrete time equations to define chaos or a non Markov correlation structure between a system and its environment [110]. This chapter will show that for stationary processes the NMP can be expressed in closed form in terms of operations on the power spectrum of the measured system. It is then shown that with the additional constraint of a smooth autocorrelation function (specifically belonging to the C^1 or higher differentiability class) the first NMP has a particularly simple structure depending solely on the spread and amplitude envelope of the measured power spectrum. These results provide a more conventional signal processing perspective from which to understand the NMP in terms of the power spectrum of the measured system. This result is entirely complementary to the original Mori-Zwanzig framework of complex interactions between the measured system and its environment. In essence these results allow the NMP to be expressed simply without a detailed understanding of Mori-Zwanzig theory.

We then go on to show that closed form expressions for the higher order NMP can be constructed in a modular fashion by a set of nonlinear operations on the measured power spectrum. We then suggest, but do not prove, that these operations remove correlation structure from the spectrum and thus there is limited information about the system in the higher order NMP. This analysis is consistent with [117], which argued that the memory kernels can 'veil' the properties of the physical system. The ZF-NMP$_1$ is then analytically calculated for three instructive systems: simple harmonic oscillator (SHO) driven by white

noise, band limited white noise and white noise passed through an ideal all pole filter. We show that the dominant feature of ZF-NMP$_1$ is the slope of the tail of the measured power spectrum. We lastly show with numerical simulation that these expressions are also valid for noisy sampled systems where the power spectrum is not known a priori.

The work performed in this chapter may be considered an extension to that in [118]. The work in [118] was concerned with developing the zero frequency NMP from time correlation functions generated from time propagation operators in dynamical systems theory. These time correlation functions are constructed such that they are only non-zero for positive time. This greatly simplifies the spectral analysis in [118] because the Fourier Transforms (FT) can be represented as Laplace transforms rotated by 90° in the complex plane. Note that this is distinct from the autocorrelation functions considered in time series analysis which are symmetric functions defined for both positive and negative time. In this chapter we derive analytical expressions for the generalised NMP spectra for measured systems with autocorrelation functions defined for both positive and negative time. Lastly this work is concerned with observing how the NMP varies with the measured power spectrum whereas [118] uses successively higher order zero frequency values of the non-Markov parameters to explicitly explore the Markovity of specific causal systems.

The theory portion (sections 2.4-2.8) of this chapter are divided into four sections. The first section gives an overview of the Zwanzig Mori kinetic equations from which the NMP is derived. The second section develops closed form expressions for the hierarchy of generalised NMPs and shows how the zero frequency value of the NMPs can be simplified. The third section uses these simplifications to derive analytical expressions for the ZF-NMP$_1$ for three stochastic processes. The analysis of the SHO driven by white noise provides a conceptual bridge between the analysis of the Markovity of the physical system and the signal processing interpretation of the ZF-NMP$_1$ introduced in this chapter. The analysis of the band limited white noise and the ideal all pole filter provide an explicit understanding of how this parameter varies with spectral properties of corner frequencies and stop band decay rates. The fourth section determines the generalised NMP from the discrete time series data of a model

of the SHO driven by white noise. This highlights that the closed form expressions for the generalised NMP are applicable to 'real world' problems where only noisy sampled realisations of a process are available. This methodology will then be used to distinguish different experimental conditions of single channel MER signals placed into human Sub Thalamic Nuclei in section 2.9.

2.4 Mori-Zwanzig Kinetic Equations

Consider a system of interacting objects with defined observables which completely describe the phase space of the system of interest. Often we are only concerned with the evolution of a subset of all the objects in the system. For example the voltage contribution to an electrode of the closest neuron in a highly connected neural network.

The number of observables of interest can be extended to an arbitrarily high number but for simplification we will consider the evolution of one observable $\mathcal{G}(t)$. The evolution of the observable of interest is described by a generalised Langevin equation (GLE) [119]:

$$\dot{\mathcal{G}}(t) = \lambda_0 \mathcal{G}(t) - \Lambda_0 \int_0^t m_1(t-t')\mathcal{G}(t')dt' + \mathcal{S}(t) \quad t \geq 0, \tag{2.1}$$

where $\mathcal{G}(t)$ and $\dot{\mathcal{G}(t)}$ are the observable of interest and its time derivative respectively, t' is a dummy variable of integration, $m_1(t)$ is the first memory kernel which introduces history dependence, $\mathcal{S}(t)$ is a stochastic forcing function and λ_0 & Λ_0 are the zeroth order relaxation parameters.

The contribution of the neglected variables is accounted for in the memory (convolution) function and stochastic forcing terms. The convolution term is a consequence of a general result that when the evolution of a multi degree of freedom dynamic system, which is history independent, is described by a reduced number of degrees of freedom it is transformed to a history dependent dynamical system [65]. The presence of the stochastic forcing term is a consequence of the state vector which describes the observables of interest. This resides in

a subspace of the Hilbert space of the full dynamic system at time zero, rotating as time increases outside of this subspace into the full Hilbert space. The evolution of this state vector outside this subspace is modelled as stochastic forces randomly rotating the state vector [65].

The difficulty with the GLE (2.1) is that the presence of the noise term makes the system a *stochastic* integro-differential equation which is mathematically difficult to analyze. The equation can be reduced to a standard integro-differential by projecting the observable at time zero $\mathcal{G}(0)$ onto the evolution equation and performing an ensemble average $\langle \cdots \rangle$. The noise terms are constructed such that for all time they stay orthogonal to the observable at time zero [120] and thus the noise term is removed. Thus:

$$\langle \mathcal{S}(t)\mathcal{G}(0) \rangle = 0 \tag{2.2}$$

$$\langle \mathcal{G}(t)\mathcal{G}(0) \rangle = m_0(t). \tag{2.3}$$

Where $m_0(t)$ is the autocorrelation function of the observable $\mathcal{G}(t)$. Applying these operations to (2.1) yields an integro-differential equation for the evolution of the autocorrelation function, $m_0(t)$, of our single observable of interest [65]:

$$\frac{dm_0(t)}{dt} = \lambda_0 m_0(t) - \Lambda_0 \int_0^t m_1(t-t')m_0(t')dt' \quad t \geq 0. \tag{2.4}$$

There is a subtlety regarding the evolution of the autocorrelation function which must be highlighted. The autocorrelation function by definition is a symmetric function defined for both negative and positive time, however the GLE (2.1) is only defined for positive time and thus the evolution of the autocorrelation function in (2.4) is only defined for positive time. This point will be important when considering the FT of the memory kernel in this section.

The first memory kernel can be shown by the second fluctuation dissipation theorem to be the autocorrelation (and thus symmetric) function of the stochastic forcing function [120]:

$$m_1(t) = \frac{\langle \mathcal{S}(t)\mathcal{S}(0) \rangle}{\langle \mathcal{G}(0)\mathcal{G}(0) \rangle}. \tag{2.5}$$

Arbitrarily higher order equations can be constructed by interchanging the positions of the autocorrelation function $m_{n-1}(t)$ with the memory kernel $m_n(t)$ and introducing a new

memory kernel $m_{n+1}(t)$ to replace $m_n(t)$ into the convolution term. This is known as the Mori-Zwanzig chain, with the memory kernel autocorrelation functions acting as the links of the chain.

$$\frac{dm_n(t)}{dt} = \lambda_n m_n(t) - \Lambda_n \int_0^t m_{n+1}(t-t')m_n(t')dt' \quad t \geq 0. \tag{2.6}$$

By convention all of the considered autocorrelation functions are normalised such that $m_n(0) = 1$ [110]. For the first autocorrelation function ($m_0(t)$) this can be achieved by dividing every term in (2.4) by an appropriate normalising factor. For the remaining memory autocorrelation functions the normalisation is ensured by the form of the Λ_n relaxation factors.

Recall that in this thesis we use the following FT convention:

$$F(\omega) = \mathcal{F}\left[f(t)\right](\omega) = \int_{-\infty}^{\infty} f(t)e^{-i\omega t}\,dt \tag{2.7}$$

$$f(t) = \mathcal{F}^{-1}\left[F(\omega)\right](t) = \frac{1}{(2\pi)}\int_{-\infty}^{\infty} F(\omega)e^{i\omega t}\,d\omega. \tag{2.8}$$

Notice that an equivalent normalisation of unity requirement on the FT of the n^{th} memory autocorrelation function which we refer to as the n^{th} order memory power spectrum, $M_n(\omega)$, can be constructed using the Wiener-Khinchin theorem [121]:

$$\int_{-\infty}^{+\infty} M_n(\omega)d\omega = 2\pi. \tag{2.9}$$

The form of the λ_n relaxation parameters can be determined from manipulation of (2.4) as:

$$\lambda_n = \lim_{t\to 0^+} \frac{dm_n(t)}{dt} = 0 \quad \forall \quad m_n(t) \in C^1. \tag{2.10}$$

Notice that because the time correlation function in (2.4) is only defined for positive time, the limit is taken from above. Due to the symmetry of autocorrelation functions, as long as there is no breakdown in smoothness of the derivative (that is it belongs to the C^1 or higher set of functions) at the origin, this parameter must be zero in continuous time. An

equivalent requirement can be constructed in the frequency domain:

$$\lambda_n = \lim_{h\to 0^+} \frac{1}{2\pi} \frac{\int_{-\infty}^{\infty} M_n(\omega)\left(e^{i\omega h}-1\right) d\omega}{h}. \tag{2.11}$$

By the dominated convergence theorem [122], if $M_n(\omega)$ decays $O(\omega^{-2})$ or faster, the limit can be brought inside the integral and the integrand evaluated to yield the indeterminate function 0/0. Applying L'Hôpital's rule, if $M_n(\omega)$ decays $O(\omega^{-3})$ or faster, then the limit can again be brought inside the integral and shown to be zero. Since the power spectrum must be symmetric, this $O(\omega^{-3})$ decay is not possible. Thus if $M_n(\omega)$ decays $O(\omega^{-4})$ or faster, λ is zero. As a counter example, an autocorrelation function with a $c(t) = e^{-a|t|}$ and $M_n(\omega) \equiv O(\omega^{-2})$ structure, which has a non zero λ value, was considered in [115]. A full derivation of the functional form of this λ_n relaxation parameters is provided in section A.1.

The Λ_n relaxation parameter cannot be defined from (2.4) without the previously stated constraint that $M_{n+1}(0) = 1$. Taking the derivative of (2.4), applying the Leibniz rule for differentiating the convolution term and taking the limit as time goes to zero yields:

$$\Lambda_n M_{n+1}(0) = \lambda_n \lim_{t\to 0^+} \frac{m_n(t)}{dt} - \lim_{t\to 0^+} \frac{d^2 m_n(t)}{dt^2}. \tag{2.12}$$

As discussed previously, if $M_n(\omega)$ decays $O(\omega^{-4})$ or faster, λ_n will be zero. Enforcing the condition that the memory kernel $M_{n+1}(t)$ must be unity at time zero gives the following expression:

$$\begin{aligned} \Lambda_n &= -\lim_{t\to 0^+} \frac{d^2 m_n(t)}{dt^2}, \quad \forall \quad m_n(t) \in C^1 \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \omega^2 M_n(\omega) d\omega. \end{aligned} \tag{2.13}$$

The second expression has been generated by application of the Weiner-Khinchin theorem and is in agreement with that obtained in [123] (which considered a similar Mori-Zwanzig kinetic equation with λ set to zero) and [118]. Notice that this expression shows that Λ_n is a measure of the spread (second central moment) of the $M_n(\omega)$ memory power spectrum. A full derivation of the functional form of the Λ_n relaxation parameters is provided in section A.2. It is important to note that despite this description of Λ_n as the second moment of the memory power spectrum, $M_n(\omega)$ should not be interpreted as a probability distribution.

2.5 Analysis of the Non-Markov Parameters

The Non-Markov parameters $\epsilon_n(\omega)$ are defined as the square root of the ratio of FT of the preceding memory kernel ($M_{n-1}(\omega)$) and the memory kernel $M_n(\omega)$ [67],[110],[112]:

$$\epsilon_n(\omega) = \sqrt{\frac{M_{n-1}(\omega)}{M_n(\omega)}} \quad n \geq 1. \tag{2.14}$$

Using the second fluctuation-dissipation theorem, (2.5), we can write the NMP as the ratio of the FT of the autocorrelation functions of the observable variables and the interacting variables:

$$\epsilon_n(\omega) = \frac{1}{\sqrt{\langle \mathcal{G}(0)\mathcal{G}(0)\rangle}} \cdot \sqrt{\frac{\mathcal{F}\left[\langle \mathcal{S}(t)\mathcal{S}(0)\rangle\right](\omega)}{\mathcal{F}\left[\langle \mathcal{G}(t)\mathcal{G}(0)\rangle\right](\omega)}} \quad n \geq 1. \tag{2.15}$$

Notice that the zero frequency value of the NMP is the ratio of the correlation times, τ_c, of the observable and interacting variables:

$$\epsilon_n(0) = \frac{1}{\sqrt{\langle \mathcal{G}(0)\mathcal{G}(0)\rangle}} \cdot \sqrt{\frac{\int_0^{+\infty} \langle \mathcal{S}(t)\mathcal{S}(0)\rangle dt}{\int_0^{+\infty} \langle \mathcal{G}(t)\mathcal{G}(0)\rangle dt}} \propto \sqrt{\frac{\tau_{c,\text{observe}}}{\tau_{c,\text{interact}}}} \tag{2.16}$$

When the correlation time of the system is high,$\tau_{c,\text{observe}} >> \tau_{c,\text{interact}}$, the observable dynamics are highly Markovian. When the correlation time of the observables is comparable to the correlation time of the interacting variables, $\tau_{c,\text{observe}} \approx \tau_{c,\text{interact}}$, the observable dynamics are Non Markovian.

By the Wiener-Khinchin theorem the $M_0(\omega)$ term in the first NMP is immediately recognised as the measured power spectrum of the signal. The higher order memory power spectrum can be evaluated by taking the FT of both sides of (2.4), with a Heaviside distribution, $\theta(t)$, included in the Fourier kernel. The resultant equation in Fourier space can be algebraically re-arranged to yield an expression for the FT of the memory kernel. The inclusion of the Heaviside distribution is necessary as the evolution of the autocorrelation function in (2.4) is only valid for positive time, whereas the FT is defined for all positive and negative time. This function was evaluated previously in [115]:

$$\mathcal{F}\left[\frac{dm_n}{dt}\theta(t)\right] = \lambda_n \mathcal{F}\left[m_n(t)\theta(t)\right] - \Lambda_n \mathcal{F}\left[\int_0^t m_{n+1}(t-t')m_n(t')\theta(t)dt'\right]. \tag{2.17}$$

These Fourier transforms will be considered individually:

$$\mathcal{F}\left[\frac{dm_n(t)}{dt}\theta(t)\right](\omega) \quad = \quad i\omega\mathcal{F}\left[m_n(t)\theta(t)\right] - 1, \tag{2.18}$$

Where integration by parts has been used and it is recognised that the definite integral component vanishes, and then the sifting property of the Dirac delta distribution has been applied.

The FT of the convolution term will be determined in the following steps. First we write the expression as a double integral:

$$\mathcal{F}\left[\Lambda_n \int_0^t m_{n+1}(t-t')m_n(t')dt'\theta(t)\right](\omega) = \Lambda_n \int_0^\infty \int_0^t m_{n+1}(t-t')m_n(t')e^{-i\omega t}dt'dt \tag{2.19}$$

Note that the Heaviside distribution is removed because the bounds of the FT have been changed from $(-\infty, +\infty)$ to $(0, +\infty)$. We evaluate this integral by applying Fubini's theorem to switch the order of integration and introduce the variable substitution: $p = t - t'$, $dp = dt$ to transform this into two separable integrals:

$$\int_0^\infty \int_0^t m_{n+1}(t-t')m_n(t')e^{-i\omega t}dt'dt = \int_0^\infty m_{n+1}(p)e^{-i\omega p}dp \int_0^\infty m_n(t')e^{-i\omega t'}dt' \tag{2.20}$$

We can write this expression on the right hand side of (2.20) as the product of two separable FTs by re-introducing a Heaviside distribution into the Fourier kernels:

$$\mathcal{F}\left[\Lambda_n \int_0^t m_{n+1}(t-t')m_n(t')dt'\theta(t)\right](\omega) \quad = \quad \Lambda_n\mathcal{F}\left[m_{n+1}(t)\theta(t)\right] \cdot \mathcal{F}\left[m_n(t)\theta(t)\right] \tag{2.21}$$

Using (2.18) & (2.21) we may algebraically re-arrange (2.17) as:

$$\mathcal{F}\left[m_{n+1}(t)\theta(t)\right] = \frac{1}{\Lambda_n}\left(\lambda_n - i\omega + \frac{1}{\mathcal{F}\left[m_n(t)\theta(t)\right]}\right) \tag{2.22}$$

This expression can be simplified using the following identity [121]:

$$\mathcal{F}\left[m_{n+1}(t)\theta(t)\right] = \frac{1}{2}\left(M_{n+1}(\omega) - j\mathcal{H}\left[M_{n+1}(\omega)\right](\omega)\right) \tag{2.23}$$

Where $\mathcal{H}\left[\cdots\right] =$ is the Hilbert transform integral, which is defined in terms of the Cauchy principal value (p.v.):

$$\mathcal{H}\left[F(\omega)\right](\omega) = \frac{1}{\pi}\,\text{p.v.}\int_{-\infty}^{+\infty}\frac{F(\omega')}{\omega - \omega'}\,d\omega'. \tag{2.24}$$

This identity (2.23) can easily be derived by applying the convolution theorem and the definition of the Fourier Transform of the Heaviside distribution. Applying this identity to (2.22) yields:

$$\frac{1}{2}\left(M_{n+1}(\omega) - j\mathcal{H}\left[M_{n+1}(\omega)\right](\omega)\right) = \frac{1}{\Lambda_n}\left(\lambda_n - i\omega + 2\cdot\frac{M_n(\omega) + j\mathcal{H}\left[M_{n+1}(\omega)\right](\omega)}{M_n(\omega)^2 + \mathcal{H}\left[M_{n+1}(\omega)\right]^2(\omega)}\right) \quad (2.25)$$

Taking the real part of (2.25) and re-arranging yields the recursive expression for the FT of $(n+1)^{th}$ memory kernel as a function of the FT of the previous n^{th} memory kernel:

$$\begin{aligned} M_{n+1}(\omega) &= \frac{2\lambda_n}{\Lambda_n} + \frac{4M_n(\omega)}{M_n(\omega)^2 + \mathcal{H}\left[M_{n+1}(\omega)\right]^2(\omega)} \\ &= \frac{2\lambda_n}{\Lambda_n} + \frac{4}{\Lambda_n}\left(\frac{M_n(\omega)}{|V_n(\omega)|^2}\right). \end{aligned} \quad (2.26)$$

Note the expression in the denominator of (2.26), $|V_n(\omega)|^2 = M_n(\omega)^2 + \mathcal{H}\left[M_n(\omega)\right]^2$ is known in communications theory as the square of the Amplitude Envelope [107] of $M_n(\omega)$. Thus we see that the FT of the memory kernel is an algebraic function of the previous memory kernel and its complex envelope.

The recursive nature of (2.26) allows the n^{th} memory power spectrum $M_n(\omega)$ to be determined from the previous $M_{n-1}(\omega)$ memory power spectrum. The n^{th} memory power spectrum can then be constructed in a modular fashion from the measured power spectrum $M_0(\omega)$ using (2.10),(2.13) & (2.26). For signals with memory kernel spectrums with decay rates $O\left(\omega^{-4}\right)$ or greater, or equivalently with memory autocorrelation functions belonging to the C^1 or greater set this has a particularly simple form:

$$M_n(\omega) = \frac{4^n M_0(\omega)}{\prod_{i=0}^{n-1}\Lambda_i\left(M_i(\omega)^2 + \mathcal{H}\left[M_i(\omega)\right]^2(\omega)\right)^2}. = \frac{4^n M_0(\omega)}{\prod_{i=0}^{n-1}\Lambda_i|V_i(\omega)|^2} \quad (2.27)$$

Thus the FT of the memory kernels are defined in terms of a product series of nonlinear integral transforms of the measured power spectrum. We will pay particular attention to the FT of the first memory kernel:

$$M_1(\omega) = \frac{4M_0(\omega)}{\Lambda_0\left(M_0(\omega)^2 + \mathcal{H}\left[M_0(\omega)\right]^2(\omega)\right)} = \frac{4}{\Lambda_0}\cdot\frac{M_0(\omega)}{|V(\omega)|_0^2}. \quad (2.28)$$

The generalised NMP, when (2.10) is satisfied, is given in closed form using (2.14) & (2.27) :

$$\begin{aligned} \epsilon_n(\omega) &= \frac{\sqrt{\Lambda_{n-1}}}{2} \cdot \sqrt{M_{n-1}(\omega)^2 + \mathcal{H}\left[M_{n-1}(\omega)\right]^2} \\ &= \frac{\sqrt{\Lambda_{n-1}}\,|V_{n-1}(\omega)|}{2}. \end{aligned} \tag{2.29}$$

Thus, when the conditions of (2.10) are satisfied, the first NMP is solely a function of the spread and amplitude envelope of the power spectrum. The description of the memory power spectrum in (2.26) as a ratio of the previous memory power spectrum and its amplitude envelope suggests two interesting properties of these parameters. Firstly, the successive memory power spectra will decay at slower rates than the previous memory power spectra. Secondly, since the action of the amplitude envelope is to smooth out the underlying function, the higher memory power spectra will become flatter and flatter over the support of the original power spectrum. We will show this behaviour of the higher order memory power spectra in the systems we analyse in sections 2.6 and 2.7.

This raises questions about the validity of using the higher order NMP to analyse a measured system. Firstly in the noise free case the successive amplitude envelopes will "smear" out the measured spectrum, thereby losing the correlation structure. Thus the higher order NMP may not be measuring anything 'interesting' about the system. This interpretation may explain why two of the physical systems considered in [118] (ideal gas and an ideal gas with linear interaction perturbations) had distinct ZF-NMP$_1$ but identical higher order zero frequency NMP values. Secondly in any signal acquisition process, noise will certainly be present. It can be seen from (2.4), with $\lambda = 0$, that extracting the memory kernel is a deconvolution of a Volterra integral equation of the first kind. These convolution equations are often ill posed [124]. Thus it is possible that in any measured system, interesting structure seen in the higher order NMP are actually the manifestation of numerical errors and or noise.

Notice that by the symmetry property of the power spectrum ($M_n(\omega') = M_n(-\omega')$) the

Hilbert transform of the power spectrum is zero at zero frequency:

$$\mathcal{H}\left[M_n(\omega)\right](0) = \frac{-1}{\pi}\ \text{p.v.}\int_{-\infty}^{+\infty}\frac{M_n(\omega')}{\omega'}\,d\omega' = 0. \tag{2.30}$$

Using (2.14) (2.27) and (2.30), closed form expressions for the zero frequency values of the generalised NMP that solely depend on the spread and the DC offset of the n^{th} memory power spectrum can be obtained. These expressions are in agreement with the calculations in [118]:

$$\epsilon_n(0) = \frac{\sqrt{\Lambda_n}}{2}M_n(0). \tag{2.31}$$

We pay particular attention to the zero frequency value of the first NMP (ZF-NMP$_1$):

$$\epsilon_1(0) = \frac{M_0(0)}{2\sqrt{2\pi}}\sqrt{\int_{-\infty}^{\infty}\omega^2 M_0(\omega)d\omega}. \tag{2.32}$$

The complicated structure of the Mori-Zwanzig chain (2.4) obfuscates the fact that the ZF-NMP$_1$ (and indeed the generalised NMP) is solely a function of the measured power spectrum. Indeed the Mori-Zwanzig equations and associated memory kernels do not even need to explicitly be considered. These results suggest that signal metrics to differentiate complex systems can be developed by analysing different properties of the measured power spectrum. The previous success of the NMP [67],[106],[111],[112],[115] helps elucidate specifically what properties (spectral spread and DC offset) should be explored, but does not necessarily require Mori-Zwanzig theory to interpret the results.

The dependence of the ZF-NMP$_1$ on the DC offset value of the spectrum raises an important digital signal processing issue regarding the use of this parameter for "real-world" measured systems. Spectrum values determined from nonparametric estimation methods (e.g. Welch's method) will be random variables (typically Chi-Squared distributed [125]). Thus if the ZF-NMP$_1$ is to be used as a signal metric, it would be advisable to use a large number of signal samples and appropriate statistical tests or parametric spectrum estimation techniques (e.g. Burg's method) to reduce the variance associated with this metric [125].

It is interesting to note that the rich class of behaviour generated by Markovian dynamics

will all have a common ZF-NMP_1 value of infinity [110]. Thus in this framework different purely Markov systems cannot be differentiated, representing a degeneracy. Methodologies have been developed to explore and analyse the behaviour of these Markov processes given only measurements of the system. An example of this is modelling the unknown Markov system as a set of coupled, Langevein equations (representing the restricted case of (2.1) with the memory kernel term set to zero) and performing an Eigenanalysis on the diffusion matrix of the corresponding Fokker-Planck equation [126]. This approach was used to successfully model the power response curves of wind farm turbines [127].

2.6 Explicit Calculation of ZF-NMP$_1$ for physical systems

In this section we derive the ZF-NMP$_1$ for three instructive stochastic processes: simple harmonic oscillation driven by white noise, band limited white noise and the output of white noise passed through an idealised all pole filter. We perform this analysis to understand how sensitive this parameter is to specific variation in the measured power spectra. This provides insight into what changes, in terms of spectral properties, the ZF-NMP$_1$ was detecting in the measured complex systems analysed in [67],[111],[113],[114].

2.6.1 Simple Harmonic Oscillation Driven by White Noise

The SHO driven by white noise provides an excellent bridge between understanding the (ZF-NMP$_1$) in the original framework of the Markovity of the system and the signal processing framework of the structure of the spectrum. This system also provides one of the few systems where the higher order memory power spectra can be analytically calculated. The difficulty in the general case is due to the evaluation of the Hilbert Transforms. The dynamics of the SHO driven by white noise are given by:

$$\frac{d^2x(t)}{dt^2} + 2\zeta\omega_0\frac{dx(t)}{dt} + \omega_0^2 x(t) \quad = \quad W(t), \tag{2.33}$$

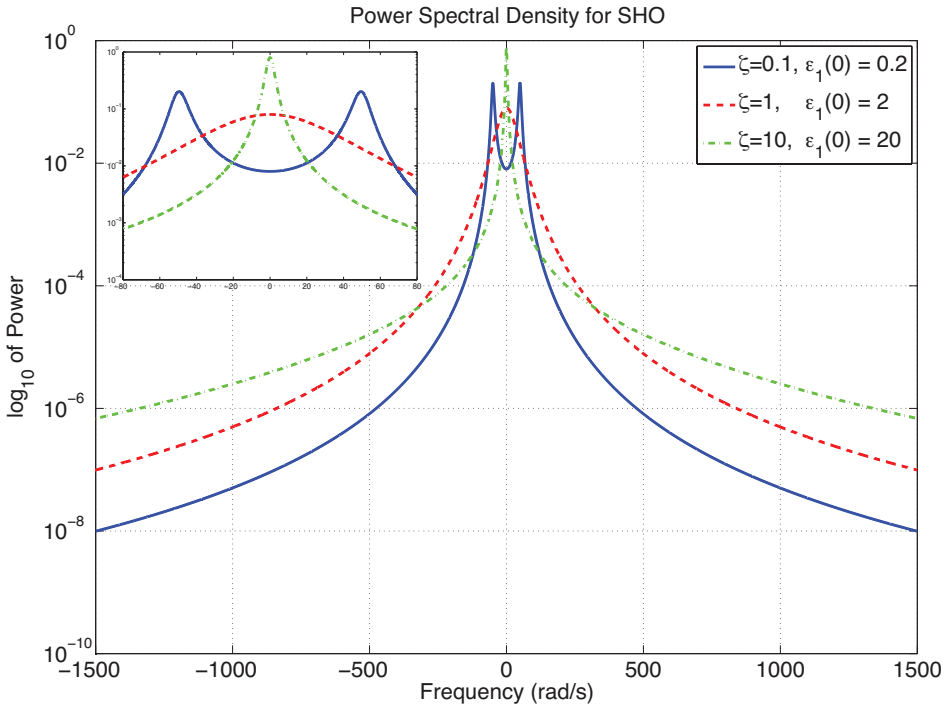


Figure 2.1: Power Spectral Density ($M_0(\omega)$) for the SHO for different damping regimes of under ($\zeta = 0.1$, blue solid line), over ($\zeta = 10$, green dash-dottted line) and critically damped ($\zeta = 1$, red dashed line). Notice that as the damping ratio is increased the spectrum becomes more spread, the zero frequency power increases and the NMP increases. Inset is the low frequency behaviour of the 3 oscillators.

where ω_0 is the angular natural frequency, ζ is the damping ratio and W(t) is the white noise stochastic process with constant amplitude power spectrum.

The normalised power spectrum of this process is given by [128]:

$$M_0(\omega) = \frac{4\zeta\omega_0^3}{\left(\omega_0^2 - \omega^2\right)^2 + 4\zeta^2\omega_0^2\omega^2}. \tag{2.34}$$

Notice that $M_0(\omega)$ decays $O(\omega^{-4})$ and thus $\lambda_0 = 0$. The contour integrals required to determine the Λ_0 from the spectral form of (2.13) are relatively difficult for arbitrary ζ, ω_0 parameters. Instead the Λ_0 relaxation parameter will be determined from the autocorrelation function form for the different damping regimes: over, under and critically damped.

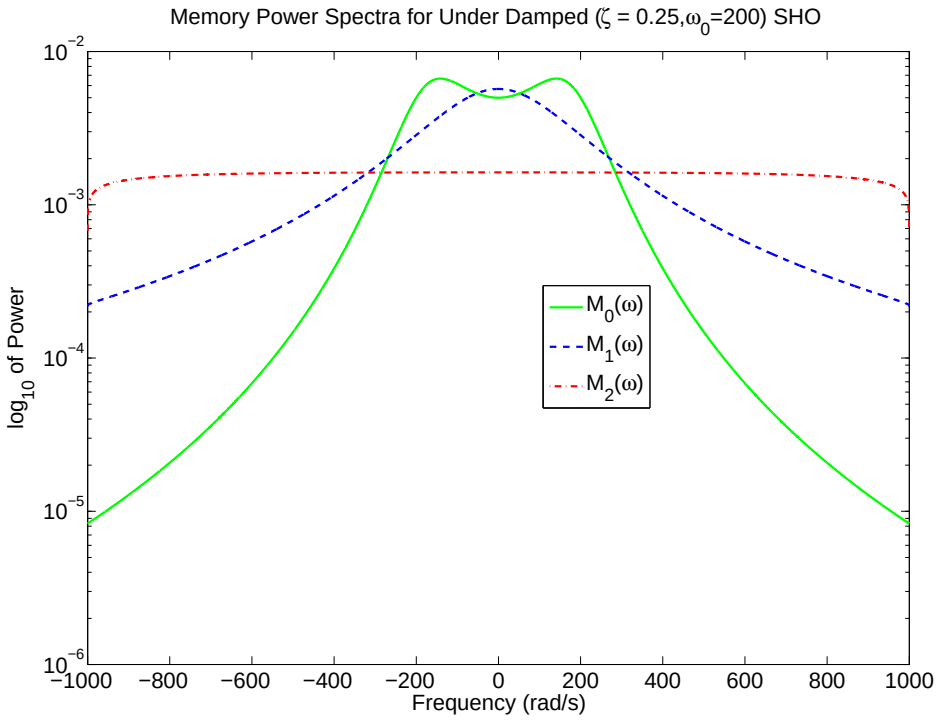


Figure 2.2: Numerically determined first three memory power spectra for the underdamped SHO. Notice that for the successively higher memory power spectra the correlation structure is lost, the spectra become more flat and the third order memory power spectra $M_2(\omega)$ appears to be approaching a white noise solution. The edge effects are numerical issues associated with the numerical estimate of the Hilbert Transform.

The normalised autocorrelation structure of the underdamped (UD) SHO is given by [129]:

$$m_0(t)_{\text{UD}} = e^{-\zeta\omega_0|t|}\left(cos(\omega_1 t) + \frac{\zeta}{\sqrt{1-\zeta^2}}sin(\omega_1|t|)\right), \tag{2.35}$$

where $\omega_1^2 = \omega_0^2(1-\zeta^2)$ is the damped natural frequency. The critically damped case is determined in the limit of $\omega_1 \to 0$. The overdamped case is determined by setting $\omega_1 \to i\omega_1$, which transforms the trigonometric functions in (2.35) to hyperbolic trigonometric functions.

The Λ_0 relaxation parameter is given for all three damping regimes by:

$$\Lambda_0 = \omega_0^2. \tag{2.36}$$

The ZF-NMP$_1$ of the SHO driven by white noise can now be written using (2.31),(2.34) & (2.36) as:

$$\epsilon_1(0) = 2\zeta. \tag{2.37}$$

Notice that the ZF-NMP$_1$ depends solely on the damping ratio which is a measure of the amount of energy dissipation in the system. Inspection of Figure 2.1 shows that as the damping ratio is increased the measured spectrum $M_0(\omega)$ becomes more spread out and the DC offset increases. The sole dependence of the ZF-NMP$_1$ on this ratio is particularly interesting, because it explicitly states that the "Markovity" of this system (as measured by the ZF-NMP$_1$) is directly related to how quickly the energy is dissipated. If the system is under-damped then the deterministic free response will dominate the stochastic forcing by the white noise process. If the system is over-damped then the predictable free response will quickly die out and the response will be dominated by the stochastic white noise process. Thus the damping is a measure of the memory the system. This analysis is entirely consistent with the original description of the ZF-NMP$_1$ [106] in terms of the Markovity of a system with respect to its environment.

Notice that in the limit of an infinitely large damping ratio the NMP approaches infinity. Analysing the SHO dynamics in the limit of the β term approaching infinity and recognising that the white noise process can informally be written as the time derivative of a Wiener process: $W(t) = dw(t)/dt$ the dynamics of this system in this limit can be written as:

$$dx(t) = -\frac{\omega_0}{2\zeta}x(t)dt + \frac{1}{2m\zeta\omega_0}dw(t). \tag{2.38}$$

This is the stochastic differential equation which describes the Ornstein-Uhlenbeck process. It is interesting to note that this is a process that satisfies the conditions of being stationary, Markov and Gaussian [130]. Thus the NMP of the Ornstein-Uhlenbeck process is infinite, which is in agreement with the original definition of this parameter.

The higher order memory power spectrum and kernel for the critically damped SHO (the

Hilbert Transform being too difficult for arbitrary damping ratios) can be analytically calculated :

$$\begin{aligned} |V_0(\omega)|^2 &= \frac{4(\omega^2+4\omega_0^2)}{(\omega^2+\omega_0^2)^2} \\ M_1(\omega) &= \frac{4\omega_0}{\omega^2+4\omega_0^2} \qquad (2.39) \\ m_1(t) &= e^{-2\omega_0 t}. \qquad (2.40) \end{aligned}$$

Notice that $M_1(\omega)$ decays $O(\omega^{-2})$ and thus λ will be non zero. Using (2.10), (2.12) & (2.40):

$$\begin{aligned} |V_1(\omega)|^2 &= \frac{4}{\omega^2+4\omega_0^2} \\ \frac{M_1(\omega)}{|V_1(\omega)|^2} &= \omega_0 \qquad (2.41) \\ \lambda_1 &= -2\omega_0 \qquad (2.42) \\ \Lambda_1 &= 0. \qquad (2.43) \end{aligned}$$

Thus the $M_2(\omega)$ memory power spectrum will be infinite due to the scaling by the inverse of Λ_1. This shows that the second order zero frequency NMP will be zero. Analysis of (2.26) shows the higher order memory kernels and generalised NMP will give pathological divide by zero solutions. This result can be explained as follows: The exponential first order memory function satisfies the differential equation:

$$\frac{dm_1(t)}{dt} = \lambda_1 m_1(t). \qquad (2.44)$$

This is exactly the Zwanzig-Mori equation (2.6) for the first memory kernel $m_1(t)$ with the convolution term (and thus Λ_1) equal to zero. It is interesting to identify the unscaled second memory power spectrum $M_2(\omega)$ given by (2.41) & (2.42) (that is not obeying the constraint in (2.9)) is white noise. This indicates the second order memory kernel $m_2(t)$ will be a Dirac delta distribution centred at time zero. This distribution cannot be scaled such that it satisfies the requirements of (2.9), which also helps explain why Λ_1 is zero.

Figure 2.2 shows the first three numerically determined memory kernels for the underdamped

harmonic oscillator ($\zeta = 0.25$, $\omega_0 = 200$) using (2.13) and (2.26) and assuming (2.10) is satisfied. Notice that the $M_2(\omega)$ term is constructed assuming $\lambda_2 = 0$ (thus Λ_2 is defined from (2.13) rather than (2.12)) so that the flat white noise structure can be identified. Notice that this flattening of the higher memory kernels is exactly as was postulated previously in section 2.4.

2.6.2 Band Limited White Noise

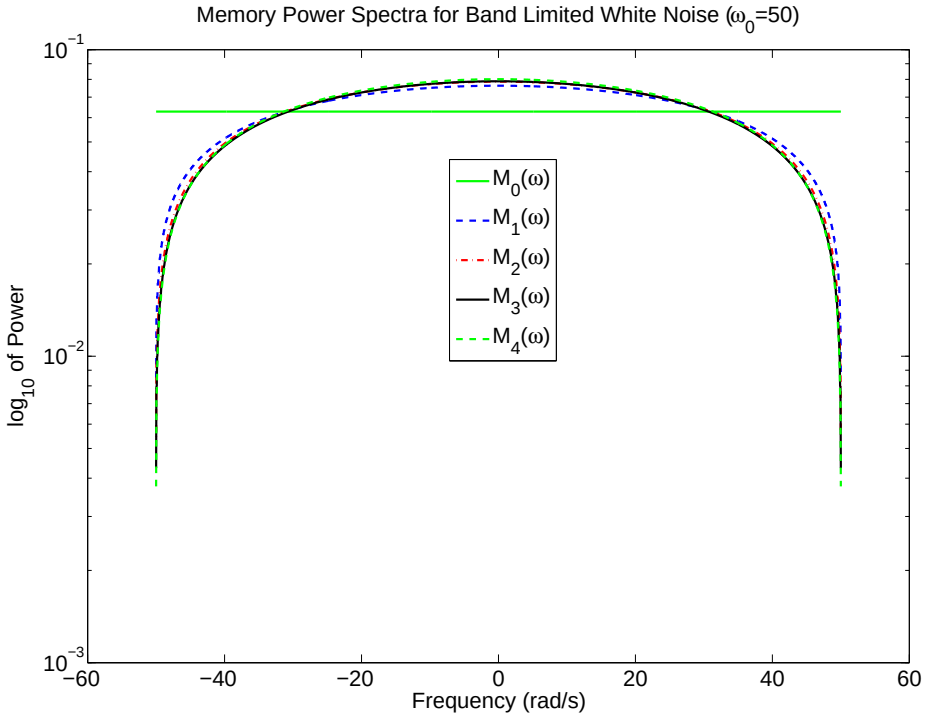


Figure 2.3: Numerically determined first five memory power spectra for the band limited white noise case ($\omega_0 = 50$). Notice that the higher order memory power spectra after the zeroth $M_0(\omega)$ term appear to converge to a common spectral structure.

The normalised power spectrum for the white noise process banded between $(-w_0, +w_0)$ is given by:

$$M_0(\omega) \quad = \quad \frac{\pi}{\omega_0} \left[\theta(\omega + \omega_0) - \theta(\omega - \omega_0)\right]. \tag{2.45}$$

Where $\theta(\cdots)$ is the Heaviside distribution. The Λ_0 relaxation parameter is given by the

spectral form of (2.13) and (2.45):

$$\Lambda_0 = \frac{1}{2\omega_0}\int_{-\omega_0}^{+\omega_0}\omega^2 d\omega = \frac{\omega_0^2}{3}. \tag{2.46}$$

The first memory power spectrum is given by:

$$M_1(\omega) = \frac{12\pi\left[\theta(\omega+\omega_0)-\theta(\omega-\omega_0)\right]}{\omega_0\left(\pi^2 + ln\left(\left|\frac{\omega+\omega_0}{\omega-\omega_0}\right|\right)^2\right)}. \tag{2.47}$$

The ZF-NMP$_1$ is given by (2.31),(2.45) and (2.46):

$$\epsilon_1(0) = \frac{\pi}{2\sqrt{3}}. \tag{2.48}$$

This is interesting because it shows that the ZF-NMP$_1$ is independent of the bandwidth of the band limited white noise. Mathematically this can be understood to be due to the normalisation requirement, as the noise process occupies more bandwidth and the spread increases, the amplitude of this noise decreases and so does the DC offset. These two parameters must change such that the ZF-NMP$_1$ remains constant. We cannot extend our analysis to the infinite bandwidth white noise process because the correlation structure is a Dirac delta distribution for which it is not possible to normalise to unity, nor define its derivatives in the limit of zero time as is required for the relaxation parameters.

Figure 2.3 plots the first five numerically determined memory power spectra of the band limited white noise ($\omega_0 = 50$) system using (2.13),(2.27) and (2.45). Notice that all the memory power spectra (except for the measured power spectrum $M_0(\omega)$) converge to a common spectral structure, indicating limited utility of the higher order NMP. The higher order Λ_n relaxation parameters will not equal zero because the compact support of the band limited white noise prevent the memory power spectra from becoming white noise solutions, or equivalently the autocorrelation function from ever becoming a Dirac delta distribution.

2.6.3 Ideal All Pole Filter

The ideal all pole filter refers to a piecewise continuous spectrum that consists in log-log (base e) space of a straight line of height h which goes from 0 to the corner frequency ω_c and

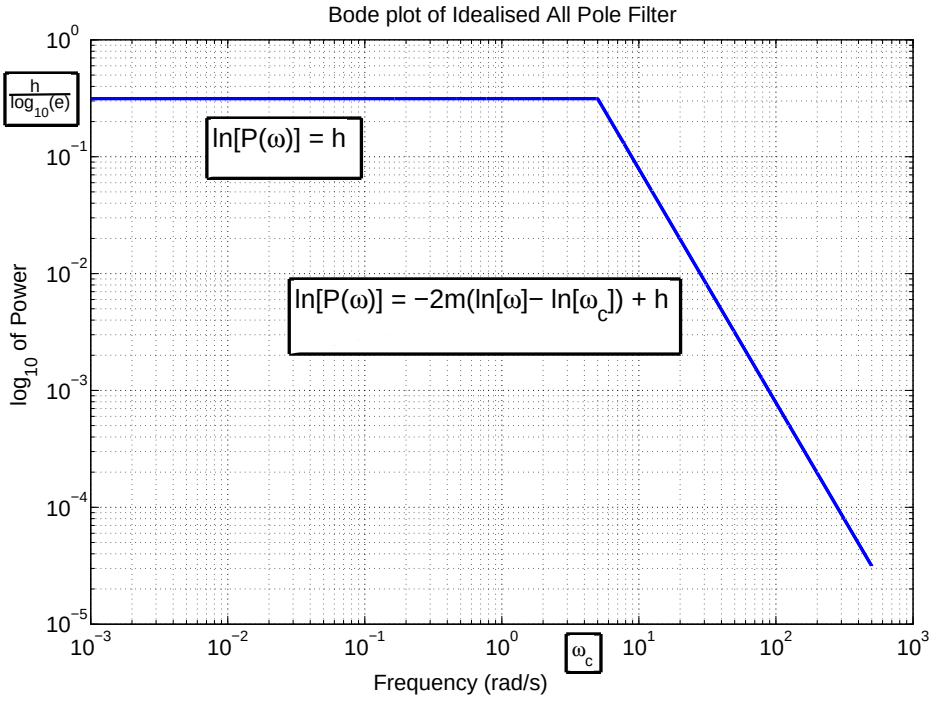


Figure 2.4: Bode plot of the ideal all pole filter spectrum with corner frequency ω_c and filter order m.

then another straight line which has a negative slope proportional to the order of m, which goes from the corner frequency ω_c to infinity. The Bode plot of this spectrum is shown in Figure 2.4. The actual value of the height h is determined by the normalization requirement of (2.9). This is an idealised filter because the cusp (breakdown of the first derivative) at the corner frequency creates an unphysical 'infinite-power' requirement on the filter [107]. Intuitively we expect the NMP to depend on both the order and corner frequency of the filter, but we show that it depends solely on the slope of the tail (i.e. m) of the power spectrum.

We can mathematically represent the power spectrum of our idealised all pole filter in log-log (base e) frequency space as:

$$ln\left[M_0(\omega)\right] = \begin{cases} h & 0 \geq ln\left[\omega\right] \geq ln\left[\omega_c\right] \\ -2m\left(ln\left[\frac{\omega}{\omega_c}\right]\right) + h & ln\left[\omega_c\right] \geq ln\left[\omega\right] \geq \infty. \end{cases} \tag{2.49}$$

We can map this to frequency space by taking the antilog (in base e) of both sides of

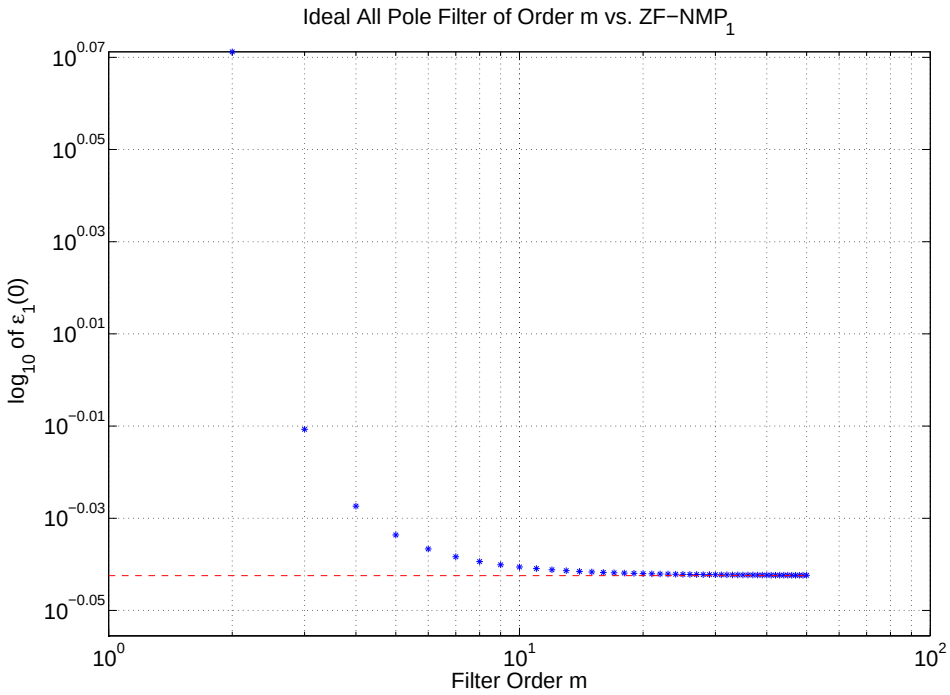


Figure 2.5: Plot of ZF-NMP$_1$ of the output of white noise fed into the ideal all pole filter vs slope order (m). Notice that for sufficiently large slope order the NMP converges to the band limited white noise solution.)

(2.49). There are two things to notice: Firstly the power spectrum is symmetric about the origin whereas the log-log (in base e) power spectrum is one sided. Thus we make the solution obtained in frequency space symmetric about the zero frequency origin. The log-log (in base e) power spectrum when mapped to the power spectrum will start at $\omega = 1$ (because $e^0 = 1$ in the bounds for the constant straight line). We simply extend the bounds of the power spectrum to $\omega = 0$ in frequency space.

The constant $\alpha = e^h$ will be determined such that the power spectrum has the appropriate normalisation required by (2.9):

$$\alpha = \frac{\pi\,(2m-1)}{2m\omega_c}. \tag{2.50}$$

The power spectrum is given by:

$$M_0(\omega) = \begin{cases} \dfrac{\pi\,(2m-1)}{2m\omega_c} & 0 \geq |\omega| \geq \omega_c \\ \dfrac{\pi\,(2m-1)}{2m\omega_c}\left(\dfrac{\omega}{\omega_c}\right)^{-2m} & \omega_c \geq |\omega| \geq \infty. \end{cases} \tag{2.51}$$

In order to simplify analysis, we will only consider ideal all pole filters of order m of 2 or higher.

The Λ_0 relaxation parameter can be determined from the spectral form of (2.13) and (2.51):

$$\Lambda_0 = \frac{(2m-1)\,\omega_c^2}{3\,(2m-3)}. \tag{2.52}$$

Using (2.31),(2.51) and (2.52) the ZF-NMP$_1$ is given by :

$$\epsilon_1(0) = \frac{\pi}{m}\sqrt{\frac{(2m-1)^3}{48\,(2m-3)}}. \tag{2.53}$$

The most interesting result from (2.53) is that, similar to the band limited white noise case, the ZF-NMP$_1$ for the ideal all pass filter is independent of the corner frequency. The ZF-NMP$_1$ is sensitive to the order of the filter. Figure 2.5 shows an inverse relationship between the decay rate of the tail of the power spectrum and the ZF-NMP$_1$ value. Similar to the SHO, the more spread out the power spectrum, the larger the ZF-NMP$_1$. An obvious difference between these two systems is that the ZF-NMP$_1$ for the SHO is unbounded, whereas (at least for $m \geq 2$) the ZF-NMP$_1$ of the ideal all pole filter is approximately bounded between 1.17 (m=2) and 0.9 ($m \to \infty$). Notice that while the ZF-NMP$_1$ does depend on the tail of this spectrum, it is not particularly sensitive to it.

It can be seen that the ZF-NMP$_1$ value for the ideal all pole filter converges to the band limited white noise case for sufficiently large slope order (m $\approx$ 10 or higher). This result is to be expected, because as the slope of the ideal all pole filter increases the spectrum will converge to the band limited white noise spectrum. It is trivial to formally show that the NMP of this idealised all pole filter converges to the band limited white noise process in the

limit of infinite filter order:

$$\lim_{m\to\infty} \epsilon_1(0) = \lim_{m\to\infty} \pi\sqrt{\frac{(2m)^3}{48m^2\,(2m)}} = \frac{\pi}{2\sqrt{3}}. \tag{2.54}$$

Figure 2.6 plots the first five numerically determined memory power spectra of the ideal all pole filter ($m = 4$, $\omega_0 = 50$) using (2.13),(2.27) and (2.51). Notice that (similar to the critically damped simple harmonic oscillator shown in Figure 2.2) the higher order memory power spectra are flatter and lose the correlation structure present in the measured $M_0(\omega)$ spectrum. This further validates the flattening of the higher memory kernels as postulated previously in section 2.4. Again, this raises questions about the utility of the higher order NMP as system metrics.

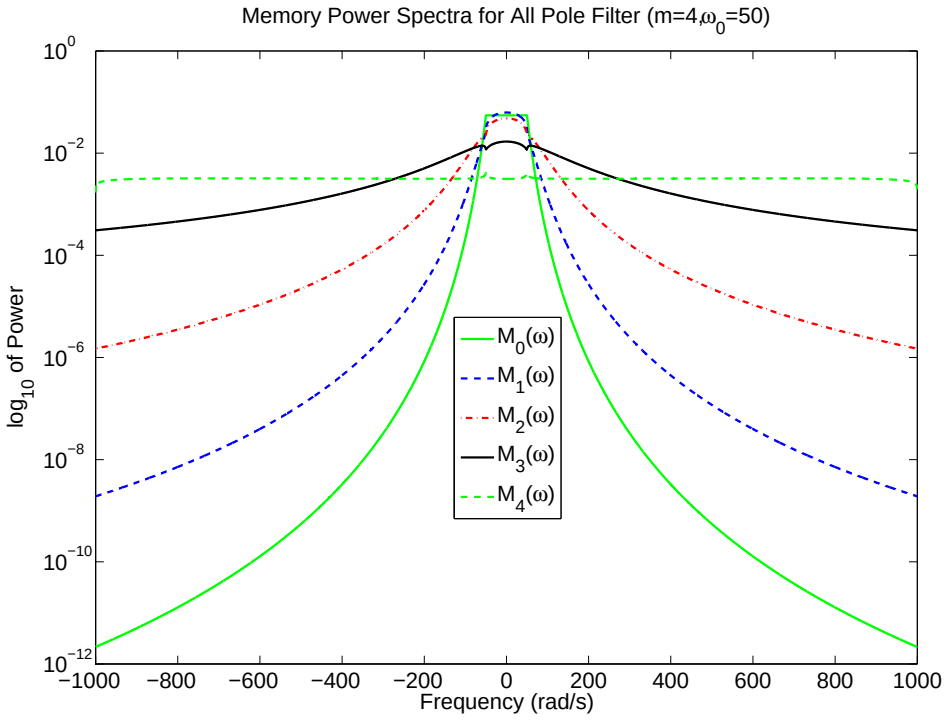


Figure 2.6: Numerically determined memory power spectra for the ideal all pole filter ($m = 4$, $\omega_0 = 50$). Notice that the higher order memory power spectra get considerably flatter.

2.7 Numerical Determination of NMP from Sampled Time Series

We finish by determining the memory power spectra and ZF-NMP$_1$, using the closed form expressions (2.27)-(2.31), from what can be considered as a model of the sampled time series of the displacement of the SHO driven by white noise. We model the system as a second order Auto Regressive (AR(2)) process:

$$\begin{aligned} x[n] &= \phi_1 x[n-1] + \phi_2 x[n-2] + \epsilon_n \qquad (2.55) \\ \text{where:} \qquad & (\phi_1 = 0.9, \phi_2 = -0.8), \quad \epsilon_n \sim \mathcal{N}(0,1). \end{aligned}$$

The choice of AR coefficients in (2.55) can be considered related to the under damped SHO driven by white noise [131]. We simulate $120,000$ data points of this process to generate a discrete time series. This number of data points is equivalent to 5 seconds of a signal sampled at 24 kHz, which is the sampling rate of the measured vLFPs of the STN which we analyse in section 2.9. We determine the memory power spectra and ZF-NMP$_1$ using solely this time series with no a priori information about the AR coefficients or innovations, ϵ_n, which define the process. A realisation of this process (and thus time series to be analysed) is shown in Figure 2.7.

The power spectrum of this process is determined from the time series using the *ARMAsel* parametric spectrum estimator. *ARMAsel* seperately fits the data to an optimal order (p) AR , an order (q) Moving Average (MA) and an optimal order ($r, r-1$) Auto Regressive Moving Average (ARMA) model. Each 'optimal order' is determined by a specified information criteria. The Prediction Error (PE) associated with these three models is then generated. The model with the smallest PE is used to generate the parametric power spectrum in the standard fashion using Fast Fourier Transforms [132]. The technical details of this estimator are provided in [133] but we provide more specific information about how these three models are constructed and compared below.

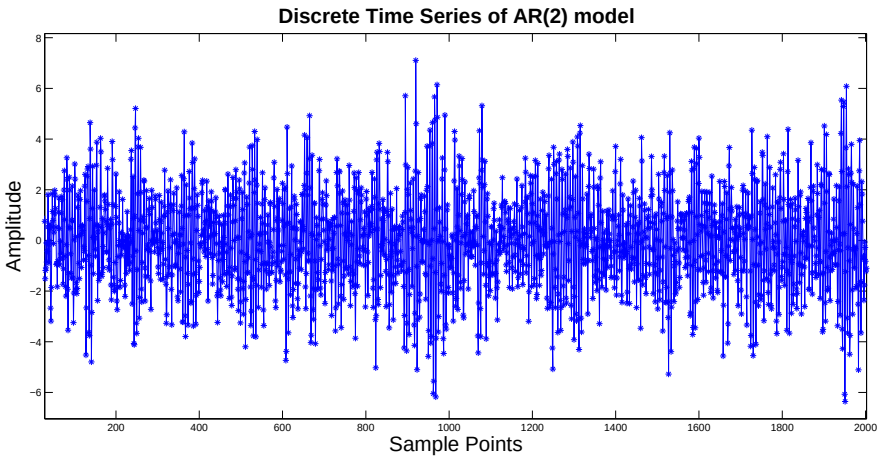


Figure 2.7: Sample path realisation of the AR(2) process ($\phi_1 = 0.9, \phi_2 = -0.8, \epsilon_n \sim \mathcal{N}(0,1)$) generated from (2.55). This discrete time series is used to generate the memory power spectra.

The AR model is solved for a sequentially higher orders using Burg's method [132] and the optimal model order p is selected such that the minimum Generalised Information Criteria (GIC) with a finite order correction factor (referred to in [133] as the CIC) is obtained. Similar to most information criteria the CIC provides a balance between the variance of the residuals and the order of the model. The specific difference of the CIC is that the order penalty term is a compromise between the standard asymptotic penalty term (used in GIC) and finite sample effects. The MA model is solved for sequentially higher orders using Durbin's method [134]. The optimal model order q is chosen as the minimum of the standard asymptotic GIC criteria [133]. The ARMA $(r, r-1)$ model is solved for sequentially higher orders using Durbin's second method [135] and the optimal order is selected by minimising the same GIC criteria as for the MA model (although the CIC used for selecting the AR model can also be used). Once the optimal order p AR, order q MA and order $(r, r-1)$ ARMA models are constructed, the prediction error associated with these models is then computed. The prediction error for the MA and ARMA models can be computed using Akaike's first order selection criterion [136] whereas a separate formula, derived using finite sampling theory [133]. is used to develop the AR prediction error.

An estimate of the ZF-NMP$_1$ of this AR process can be obtained from the time series by estimating the power spectrum with the ARMAsel algorithm and using (2.13) and (2.31). In Figure 2.8 we generate a beeswarm plot from estimating the ZF-NMP over 50 simulations of the AR(2) process ($\phi_1 = 0.9$, $\phi_2 = -0.8$, $\sigma = 1$) and compare these values with the analytical value of the ZF-NMP. It can be seen that the estimated ZF-NMP value provides an accurate and consistent estimator of the true ZF-NMP.

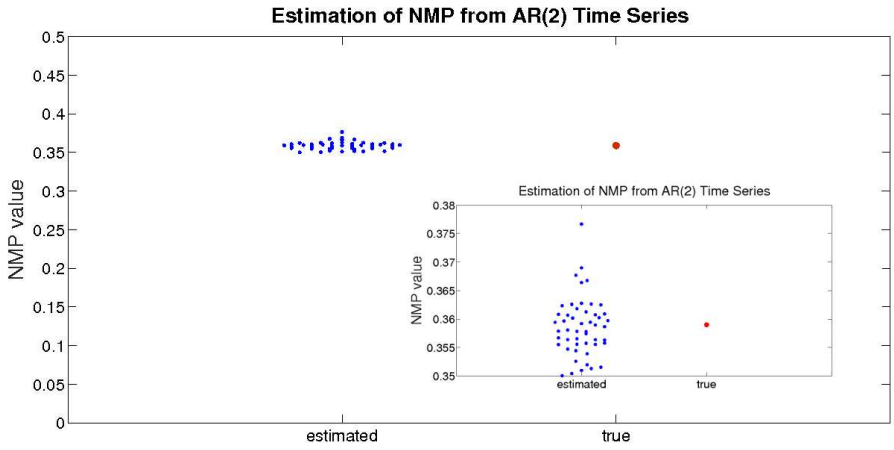


Figure 2.8: Beeswarm plot of the 50 estimations of the NMP and the true NMP for the AR(2) process with ($\phi_1 = 0.9, \phi_2 = -0.8.$) generated from (2.55). Notice that the values are consistently distributed around the true NMP value. Inset: Zoomed in version of distribution of estimates around the true value of the ZF-NMP.

The mean and standard error of the mean from these 50 realisations is:

$$\hat{\epsilon}_1(0) = 0.358 \pm 0.001 \tag{2.56}$$

This solution can be compared to the true ZF-NMP$_1$. The unnormalised power spectrum of this process is given by [137]:

$$M_0(\omega) = \frac{\sigma^2}{1 + \phi_1^2 + \phi_2^2 - 2\phi_1(1 - \phi_2)cos(\omega) - 2\phi_2 cos(2\omega)}, \tag{2.57}$$

where σ^2 is the variance of the innovations. Using (2.13),(2.31) and (2.57) with ($\phi_1 = 0.9$, $\phi_2 = -0.8$, $\sigma = 1$) the ZF-NMP$_1$ of this process can be calculated by numerical integration to be:

$$\epsilon_1(0) = 0.359 \tag{2.58}$$

Therefore it can be seen that with appropriate statistical averages (which are necessary given finite time recordings of stochastic processes) the ZF-NMP$_1$ can be accurately estimated from the time series data alone.

With an estimate of the power spectrum using the *ARMAsel* algorithm the higher order memory power spectra and generalised NMP can be determined using (2.13),(2.27) and (2.29). The estimated memory power spectra are shown in Figure 2.9. Notice that similar to systems with a continuous power spectrum analysed in section 2.6, the higher order memory power spectra appear to smear out the correlation structure observable in the power spectrum.

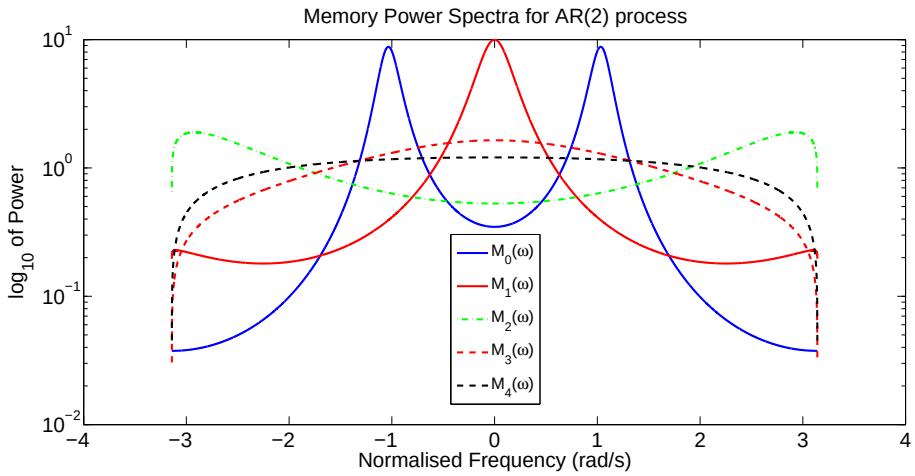


Figure 2.9: First five numerical memory power spectra estimates of the AR(2) process defined in (2.55). Notice that for the successively higher order memory power spectra the correlation structure observable in the power spectral density is smeared out. Also notice the spectrum is defined over the normalised frequency range of $(-\pi, \pi)$.

This example highlights several key points which are not immediately clear in the analysis of the physically motivated systems considered in section 2.6. Firstly this time series could be acquired without any knowledge of the underlying physics driving this system. This would make the Mori-Zwanzig interpretation of the Non Markovity (which requires partitioning the dynamical system into subsets of observables of interest and an interacting environment) extremely difficult to perform. This is in contrast to the signal processing approach introduced in this paper which interprets the Non Markovity spectrum in the concrete terms of operations on the measured power spectra. Secondly this example provides a 'real world' application of estimating the Non Markovity spectra where the underlying continuous power spectra is unknown and only discrete samples of the measured time series corrupted with noise are available. Problems of this nature are commonplace in signals analysis and thus it is important to identify that the closed form expressions (2.27)-(2.29) for the memory power spectra and generalised NMP are applicable to this class of problem. We refer the reader to [117] for a detailed description regarding the issues associated with interpreting the underlying continuous memory kernel from its discrete time estimate.

2.8 Discussion of NMP

The generalised Non-Markov Parameters have been used to successfully differentiate states (as defined by the degree of chaosity (sic) [138] and randomness) of complex interacting systems. In the preceding part of this chapter we have shown these parameters can be understood as a set of closed form expressions which only depend on a nonlinear set of integral transform operations on the measured signal's power spectrum. We have argued that the operations yielding the higher order memory power spectra and generalised NMP veil the underlying correlation structure of the measured system in agreement with [117]. We have supported this argument with numerical simulation of four instructive stochastic processes: a SHO driven by white noise, band limited white noise, the output of white noise fed into an ideal all pole filter and an AR(2) process with Gaussian innovations. These results suggest a sensitivity of the ZF-NMP_1 to the decay rate of the tail of the spectrum.

Lastly we have shown that under the appropriate condition of C^1 or higher smoothness of the autocorrelation (or equivalently $O(\omega^{-4})$ or faster decay rates of the tail of the spectral function), the closed form expression for the ZF-NMP_1 can be reduced to depending solely on the spread and DC offset of the measured power spectrum. These results provide an alternative interpretation of the generalised NMP which only depends on the measured signal and does not require knowledge of Mori-Zwanzig theory, nor interpretation of a complex relationship between a measured system and its environment. We have shown that these equations for the memory power spectra and generalised NMP can readily be applied to systems where only noisy discrete time samples are available. These simplified expressions of the NMP in light of its previous success in the analysis of complex systems provides insight into what properties of the spectrum could be used in future signal analysis of complex systems.

2.9 NMP analysis of vLFPs

In this section we apply the 1^{st} NMP, $\epsilon_1(\omega)$, at zero frequency (ZF-NMP_1) and biologically significant (β: 10-30Hz and fast: 80-200Hz) band frequencies to the vLFPs taken from MER probes in the STN of patients undergoing DBS surgery. These frequency bands were selected because of their extensive use in EEG analysis [60]. These biologically significant NMP bands are calculated after developing the first NMP spectrum, $\epsilon_1(\omega)$, and then restricting this spectrum to the desired frequency ranges. This methodology is discussed in more detail in Section 2.9.2.

As discussed in section 2.3 the application of the NMP to complex datasets is not new. Indeed the NMP has specifically been used for discriminating states of biological systems such as the R-R interval of Electro Cardio Grams (ECG) for healthy patients and patients post-myocardial infarction [67], EEG readings to predict the onset of seizures for epileptic patients [106] and finger tremor data to differentiate patients with and without Parkinson's

Disease [114]. Nonetheless this experimental analysis represents the first use of the NMP for deep brain recordings.

The application of complexity based metrics to the unfiltered neural signals is relatively unique. Most extracellular analysis is based on spike sorting followed by ad hoc clustering and analysis of resultant inter-arrival times [139], [140]. For example, mean firing rates and a burst index (calculated by dividing the burst firing rate by the mean firing rate) were considered in [141]. The template matching and clustering, if not automated are highly dependent upon the skills of the person performing the analysis [139] and accurate estimates of the firing statistics often require reasonably long sequences of data. The signal analysis in this experiment simply required the unfiltered vLFP signals which were fed into the NMP signal processing metric.

There have been limited attempts previously to analyse MER signals with complexity metrics. Typically these studies have been clinically motivated, with the score of the metric being correlated to a marker of disease state. For example The Lempel-Ziv Complexity (LZC) estimator was applied to MER data (filtered in the $13-35$ Hz range and less than 13 Hz range) taken from the STN of Parkinson's Disease patients. The details of the general Lempel-Zev method can be found in [142]. As an oversimplification the estimator can be understood as follows. The signal is broken up into discrete values and assigned a 0 if the signal is below the mean and a 1 if it is at the mean or above. The Lempel Ziv algorithm determines how many distinct ways you can have sequences of up and down crossings (the size of the dictionary in standard information theory) N_{DS}. In order to make the LZC independent of the string length, N_{DS} is scaled by the string length, N. The LZC is given by: $C = \frac{N_{DS}}{(N/log_2(N))}$. In essence this is a measure of the phasic structure of the signal. The results of this paper showed a statistically significant correlation ($\rho = -0.542, p = 0.008$) between hemibody UPDRS (Unified Parkinson's Disease Rating Scale) score and the LZC score when applied to the $13-35$ Hz filtered data. No statistically significant correlations were observed for either the less than 13 Hz data and the combined akinesia and rigidity UPDRS score.

A second issue which [64] analyses is the LZC score for simulated data with an oscillatory signal in the beta frequency range (generated by an auto-regressive model) overlaid with fractional noise with a Hurst exponent of 0.5. As the signal to noise ratio increases the LZC estimator decreases. This result is somewhat intuitive because as the stochastic noise term becomes stronger the "complexity" of the signal will increase and it would be expected that the LZC will increase. This result is interesting because a similar result would be expected for the NMP. Recall that in the statistical mechanics framework the NMP is a measure of the 'randomness' of the signal [143].

The primary advantage of applying the NMP given the theoretical analysis performed in the previous section is that any changes identified can be linked to variations in the measured power spectrum without recourse to a model of the underlying dynamics of the neural ensemble. The advantage of this may not be immediately obvious, but consider the original Mori-Zwanzig statistical mechanical framework of partitioning a dynamical system into observed and interacting subspaces. For the MER recording it is not immediately clear what either the observable or the interacting variables are. Is a single neuron the observable? Or a collection of neurons? Is some set of neurons near the probe the observables and the set further away the interacting variables? If so, where geometrically does this partition lie? This is to say nothing of the lack of a tractable model for the interacting multitude of neurons contributing to the signal. In effect without a priori knowledge of how the space is partitioned we are forced to rely on vague interpretations of the NMP as a measure of the 'chaosity (sic) [138] and randomness' of the neural ensemble. The previous analysis allows us to identify that the changes appear as variations in the measured power spectrum.

2.9.1 Experimental Methodology

Two separate datasets with slightly different experimental methodology were analysed using the first NMP across different biologically relevant frequency bands and the ZF-NMP_1. We will briefly discuss both experimental methodologies, their results and implications. We note

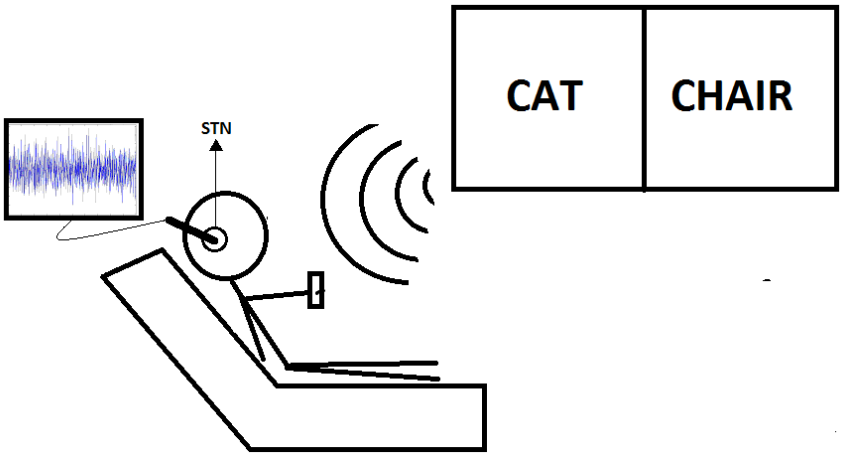


Figure 2.10: Rough cartoon of the neuro-linguistic testing set up. The patient is awake with a MER probe inserted into their Sub Thalamic Nucleus (STN) which provides the vLFP. The patient is shown two words which are either semantically similar or different (in this diagram the words are semantically different). The patient determines if they are the same or different using a clicker in the hand ipsilateral to the STN MER probe.

that interpreting the significance of these results in a speech pathology/linguistic framework is beyond the scope of this thesis and restrict our observations to concepts of signal processing.

In both datasets the general experimental protocol is largely similar. vLFPs were obtained using MERs inserted into the STN of patients with Parkinson's Disease undergoing Deep Brain Stimulation surgery. The studies were approved by The University of Queensland Medical Research Ethics Committee and UnitingCare Health Human Research Ethics Committee. All participants gave informed written consent. The electrical properties of these probes, combined with the high neuron density of the STN suggests that multiple neurons will contribute to the measured signal. The signals were recorded from the neural tissue at a sample rate of 24 (experiment 1) or 22 (experiment 2) kHz. Three Butterworth filters, as recommended by the manufacturer were applied to the signal (high-pass: 500 Hz first order, low-pass: 5k Hz first order and anti-aliasing: 5 kHz fourth order). The 500 Hz high-pass filter is included to minimise the signal contamination from muscle artifact, background EEG

activity and 50 Hz mains power interference [139]. Recordings were obtained at baseline (epoch 1), when presented with the stimulus (epoch 2) and when the patient must categorise the stimulus using a motor task (epoch 3). Note that the motor task is performed using a manual clicking device in the hand *ipsilateral* to the side of the brain that the STN is being recorded from. This is to minimise the influence of the motor task on the electrical activity of the STN [144].

Experiments of this form allow for an examination of whether the Basal Ganglia may be involved in semantic processing and decision making in addition to its well known motor modulation functions [145]. One of the unique features of this approach is the use of highly localised MER recordings in contrast to the indirect methods of subcortical activity such as functional Magnetic Resonance Imaging [146]. It is important to note that [147] used MER of the Basal Ganglia, showing that the STN activity was not modulated with the changes in linguistic processes in contrast to the results we present here which suggest a statistically significant change when using the NMP metric [148] [115].

2.9.2 Experiment 1: Semantically Similar & Different Stimuli

The complete description of the experimental set up and methodology used in this research is provided in [148]. In essence the experiments performed were based on seeing whether the electrical behaviour of the STN, as determined by MER probes, varied when a patient was presented with pairs of words which were either semantically similar (n=14) or different (n=14). The words forming the pairs were drawn from either household items or animals. Thus 'cat' and 'dog' would be considered semantically similar word pairs whereas 'cat' and 'chair' would be considered semantically different word pairs. The auditory recordings of the word pairs were presented to the patient, who responded manually using his ipsilateral index finger to indicate which category the word pair belonged to. The electrical behaviour of the STN was analysed under the additional permutations of the two brain hemispheres (the left and right sides were considered in separate trials) and the three time epochs. The

first epoch was before the word pair was given for baseline activity, the second epoch was immediately after the word pair was given during the cognitive processing of the stimuli and the third epoch was during the motor response to the stimuli. The MER signals were taken from the STN of 7 patients (all male, non senile and right handed) comprising 666 individual trials sampled at 24 kHz which were analysed off line with no additional post-processing.

We analyse the NMP value in three frequency bands: zero frequency (ZF-NMP_1) and the biologically relevant (in EEG studies): β (10-30Hz) and fast (80-200Hz) bands. The NMP spectrum, $\epsilon(\omega)$, is calculated by taking the raw vLFPs (which have been band-pass filtered between 500-5000 Hz as discussed previously in Section 2.9.1) estimating the power spectrum using the WOSA method and then using (2.13) and (2.29). The zero frequency NMP (ZF-NMP_1) is calculated by taking the zero frequency value of this spectrum. The β and fast band NMPs are calculated by taking the maximal value of this calculated NMP spectrum within the respective frequency bands of $(10-30)$ Hz and (80-200) Hz:

$$\epsilon_{[\omega_1 \cdots \omega_2]}(\omega) = \text{Max}\left\{\epsilon(\omega_1 \cdots \omega_2)\right\} \tag{2.59}$$

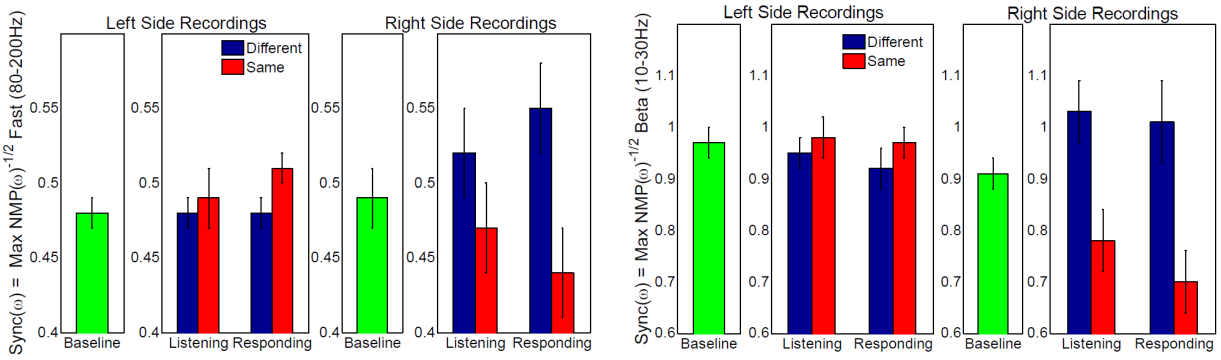
It is important to realise that these NMP metrics are calculated using frequency bands of the measured vLFPs which are filtered during the signal acquisition process. With regards to the 5 kHz low-pass filters we do not expect there to be any relevant biological contribution to the signal at frequencies beyond 5kHz. With regards to the 1^{st} order 500 Hz high-pass Butterworth filter, due to the low order, the signal attenuation (especially for the NMP fast band) will be negligible. Secondly this low-pass filter is consistently applied (with the same parameter values) to every signal acquired and in the subsequent analysis we only compare the *relative* values of these NMP metrics under the different experimental conditions.

For the statistical analysis a rank preserving power law (Box-Cox) transform [149] was applied to the NMP values to yield what we refer to as the 'Synch' parameter:

$$\text{Synch} = -2\left(\frac{1}{\sqrt{\text{NMP}}} - 1\right). \tag{2.60}$$

The normality of the Synch parameter for this data set was confirmed by Kolmogorov-Smirnov tests. The synch metric was applied to the MER signals and the resulting data was analysed using a Linear Mixed Model (LMM) to determine if there were any statistically significant interaction effects between the three fixed factors of brain side (left or right), semantic condition (same or different) and time epoch (before, processing and responding to stimulus). The LMM was set up such that brain side, semantic condition and time epoch were modeled as fixed effects whereas the patients were modeled as random effects and significance of interaction was set at the ($\alpha = 0.05$) level.

2.9.3 Experiment 1: Results



(a) fast ($80 - 200$ Hz) band. (b) β ($10 - 30$ Hz) band.

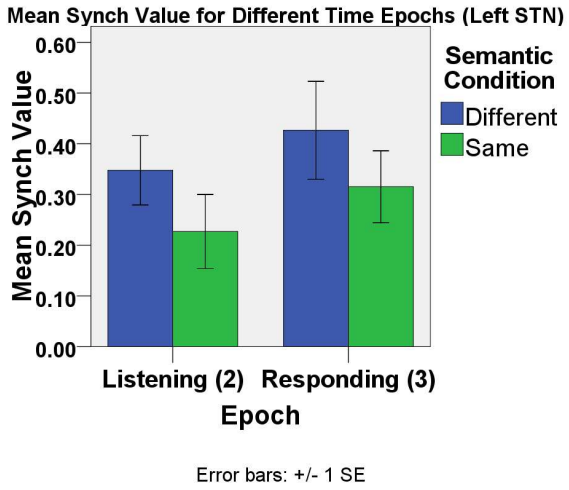
Figure 2.11: Mean Synch values for fast (2.11a) and β (2.11b) Bands

Statistically significant three way interactions (fast band: $p = 0.004$, β band: $p = 0.001$) were observed between the three variables in all 3 frequency bands using the LMM. Interaction effects of interest were then explored with planned contrasts. Since the left hemisphere is typically associated with linguistic processing we focus on the left-side recordings initially. It can be seen that for the fast band (Figure 2.11a) for the left side recording, an increase in the synch metric was observed for the same semantic (word-meaning) category compared to the different semantic (word-meaning) category during both the listening and responding phases. It is interesting to note for the right side recordings these findings were reversed, with the synch measure decreasing for the same semantic (word-meaning) category during

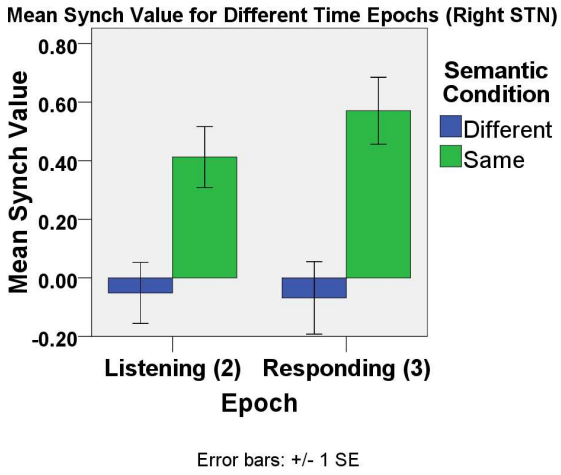
both the listening and responding phases. This may indicate important left-right hemispherical interactions are occurring in the STN during semantic processing.

Figure 2.11b shows the behaviour of the synch metric in the β (10-30Hz) band. Notice that there are similar trends to the fast frequency band results of Figure 2.11a. Interestingly, these results indicate a substantial difference in the β band NMP between the two semantic conditions which is more pronounced for the right brain side recordings. Similar to the fast band NMP, the findings for the right brain side β band NMP are the reverse of the left side with the different semantic condition increasing and the same semantic condition decreasing relative to baseline. It is important to recall that this experimental task included a motor activity i.e. button push (in the responding phase). In addition, it should be highlighted that the MER were taken from PD patients and so are not necessarily indicative of the general population. In particular, recent research [51], [54], [64] has highlighted enhanced beta band synchrony associated with STN local field potential (LFP) recordings from PD patients using power spectra and complexity-based analyses of Parkinson's disease patients.

Figure 2.12 shows post-hoc contrasts for the zero frequency synch (Box-Cox transformed ZF-NMP_1) value comparing the mean synch values with standard error as uncertainties between the listening and responding epochs on both brain hemispheres. The change in mean synch value was shown not to be statistically significant between the listening and responding epochs for the left brain, but statistically significant ($p < 0.01$) for the right brain by unpaired t-tests. It is interesting to note that, similar to the fast and β bands, the difference in the synch metric between the listening and responding phases for both the same and different word conditions is more pronounced for the right sided recordings than the left sided recordings. It is especially interesting to note however that for the left sided recordings during the listening and responding phases the synch measure for the fast and β bands has larger value for the same word condition than for the different word condition, whereas for the zero frequency value the different word condition has larger synch measure. This consistent difference is also present for the right sided recordings.



(a) Left STN. No significant difference



(b) Right STN. significant difference

Figure 2.12: Mean Synch values at zero frequency (Box-Cox transformed ZF-NMP_1) *during* and *after* (Epoch 2 and 3) *same* and *different* word pair associations from the left and right STN MER recordings. Notice that there is only statistically significant (at the $\alpha = 0.05$ level) differences between the semantic conditions for the right brain.

2.9.4 Experiment 2: Word/NonWord Stimuli

The complete description of the experimental set up & methodology for this data set is provided in section A.3. This experiment is similar to the previously described and analysed experiment of identifying whether the electrical behaviour of the STN as measured by the same MER probes varied when a patient was presented with two different categories of neuro-linguistic stimuli. In this experiment the stimuli was the presentation of a *word* (n=30, for each hemisphere) or a *non word* (n=30, for each hemisphere). The non words were taken from the Australian Research Council non word database [150] and were orthographically legal, pronounceable & not homophonic with respect to English words.

The experiments began with a fixation cross presented at the center of a screen the patient was looking at directly. Either a word or a non word was then displayed on the screen. The patient then had to respond with a motor task in their ipsilateral hand whether the stimulus was a word or not a word. The motor task was using two-button clicker: one button

corresponding to a word, one button corresponding to a non word. Note that similar to the previous experiment the motor task is performed on the hand ipsilateral to the side of the STN being recorded in order to minimise the influence of the motor task on the electrical activity of the STN [144]. Similar to the previous experiment the signal was recorded at baseline prior to presentation of the stimulus (epoch 1), during presentation of the stimulus (epoch 2) and during the response to the stimulus (epoch 3).

934 MER time series segments, sampled at 22 kHz were recorded from six male patients with idiopathic Parkinson's Disease without dementia. The signals were then pre-amplified and filtered with a bandwidth of $500 - 5000$ Hz using the previously mentioned filter setting recommended by the manufacturer: Medtronic. The vLFPs for this experiment were only analysed using the zero frequency value of the first NMP: ZF-NMP_1.

2.9.5 Experiment 2 Results

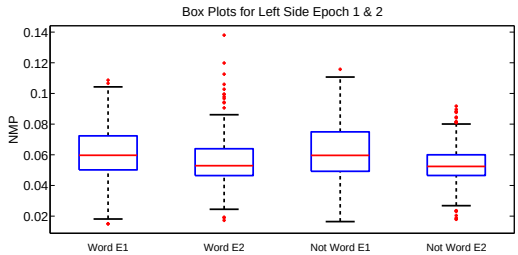
Despite applying an optimal Box-Cox transformation, the assumption of normality could not be established using the Kolmogorov-Smirnov test statistic with an $\alpha = 0.05$ significance level. Since this data set did not satisfy normality the use of Linear Mixed Models was not applicable and the data was therefore analysed in a non-parametric framework using the Friedmann 2-way analysis of ranks by variance with repeated measures. Since the Friedmann test cannot incorporate 3 way interactions the data was partitioned into two separate sub data sets of left and right STN data.

For each sub data set the permutations of patient ID and semantic condition were considered as subjects. For example patient 5 when presented with *is word* stimuli was considered a single variable, whereas patient 5 when presented with *non-word* stimuli was considered another variable. The recordings at the different epochs of baseline (epoch 1), during presentation of stimuli (epoch 2) and motor response to categorise (epoch 3) constituted the repeated measures for the data. Note that for each subject (permutation of patient ID and stimuli

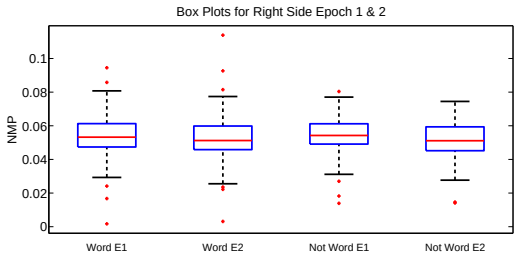
presented) at each epoch there were multiple recordings. The methodology to generalise the Friedmann test to this situation of multiple treatments per condition was developed in [151].

For both the left and the right data subsets, at the $\alpha = 0.05$ level of significance, the modified Friedmann test detected differences between the subjects. The differences between conditions were further explored with Wilcoxon rank sum tests. The application of the Wilcoxon rank sum test for both the left and right data subsets, for the presentation of both word and non word stimuli showed statistical significance between the epoch 1 and epoch 2 & 3. These results are shown in Table 2.1. Interestingly for either side or stimuli there is no statistically significant difference between epochs 2 & 3. It is also important to note that for both epoch 2 & 3 on both the left & right sub datasets that there was no statistically significant difference between the word and non word condition. These results are shown in Table 2.2. Lastly we note the consistency of the results by identifying that no statistically significant difference was identified for either the left or the right data subsets at epoch 1 between the is word and non word conditions. This result is important because epoch 1 occurs prior to the presentation of the stimuli and so we would not expect an identifiable difference in the signal between these two conditions. These results are shown in Table 2.3.

The interquartile range of the subsets of data compared using the Wilcoxon rank sum tests are displayed in box plots in Figures 2.13-2.15.



(a) Left Side Epoch 1 & 2 for word & non-word semantic condition



(b) Right Side Epoch 1 vs 2 for word & non-word semantic condition

Figure 2.13: Interquartile range box plots of ZF-NMP_1 for epoch 1 & 2.

Left Side		
	Is Word	Not Word
Epoch		
E1 vs E2	7.4×10^{-4} *	6.9×10^{-4} *
E1 vs E3	3.7×10^{-4} *	1.8×10^{-4} *
E2 vs E3	0.053	0.11
Right Side		
	Is Word	Not Word
Epoch		
E1 vs E2	0.034 *	0.001 *
E1 vs E3	0.002 *	7.5×10^{-4} *
E2 vs E3	0.3	0.85

Table 2.1: Wilcoxon Rank Sum tests comparing epochs for the left and right recordings. The $*$ signifies significance at the $\alpha = 0.05$ level. Notice that for both left and right sides statistically significant difference is identified between epochs 1 & 2 and 1 & 3.

2.10 Experiment Conclusion

The NMP/Synch metric has been applied to two separate datasets to distinguish between different states of the STN acquired using vLFPs obtained from MERs. This approach is interesting, because unlike classical spike sorting approaches, which analyse the signal using a subset of the data (the estimated spikes) the NMP is constructed in a model-free framework using the entire signal. Using the theoretical analysis performed in this chapter it is shown that the NMP essentially performs non-linear feature selection on the vLFP power spectrum. The discrepancy between these results and those discussed in [147] is most likely a consequence of using the synch/NMP metric. The MER signals in [147] were analysed in a linear signal processing framework by comparing whether the mean peak voltage and latency time observed between semantically different word pair experiments was significant ($p \leq 0.05$ level). Furthermore these results suggest that if STN is involved in lexical processing,

Is Word vs Not Word		
	Left	Right
Epoch		
E2	0.42	0.67
E3	0.41	0.75

Table 2.2: Wilcoxon Rank Sum tests comparing semantic condition. Notice that no significantly significant differences are identified at the $\alpha = 0.05$ level.

Epoch 1		
	Left	Right
Word vs Non Word	0.91	0.41

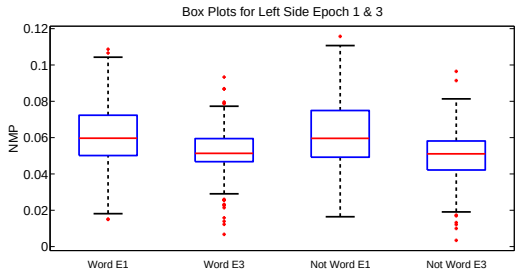
Table 2.3: Wilcoxon Rank Sum tests comparing epoch 1 for the two semantic conditions. Notice that because the stimuli is presented after epoch 1 we would expect, and indeed observe, that there is no statistically significant difference between these two states.

in order to see its effects (at least for short time recordings) we must analyse the *entire* signal. These results, where changes between states have been identified by analysing the entire signal, rather than just the spikes suggest that the 'neural noise' may in fact have valuable information content. This implies that the neurons both close and distant to the probe are modulated by experimental condition.

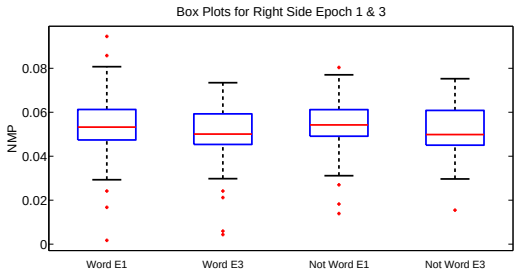
2.11 Conclusion & Thesis Contribution

There are two major contributions of this chapter which can be understood separately in a theoretical and experimental framework.

Theoretical: The family of Non-Markov Parameters is analysed within a signal processing framework. This approach allows the NMP to be understood in the concrete terms of operations of the power spectrum in contrast to the more typical framework of non-equilibrium statistical mechanics. In broad terms this analysis makes the application

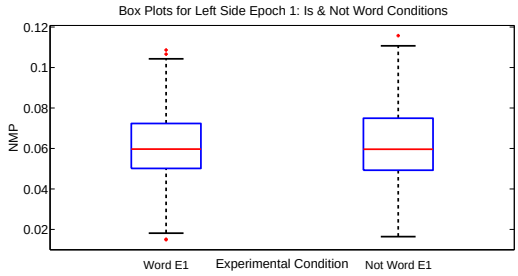


(a) Left Side Epoch 1 & 3 for word & non-word semantic condition

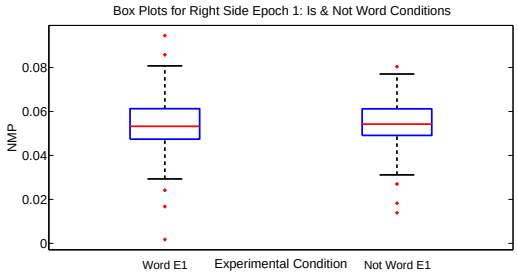


(b) Right Side Epoch 1 & 3 for word & non-word semantic condition

Figure 2.14: Interquartile range box plots of ZF-NMP_1 for epoch 1 & 3.



(a) Left Side Epoch 1: Is Word vs Not Word



(b) Right Side Epoch 1: Is Word vs Not Word

Figure 2.15: Interquartile range box plots of epoch 1. Notice for both the left and right sided recordings the Wilcoxon rank sum test found no statistically significant difference at the $\alpha = 0.05$ level between the same and different conditions.

of the NMP more accessible to the engineering community which is more comfortable with the concepts of spectral analysis.

More specifically we have shown that the usefulness of the (sequentially more complicated) infinite family of Non-Markov Parameters is actually limited to the first NMP. We have shown that the ZF-NMP_1 can be understood in terms of simple operations of the measured spectrum, rather than the abstract definitions of defining 'chaosity (sic) [138] & randomness'.

Experimental: We have shown that the Non-Markov Parameter can be used to successfully differentiate the time series of the Sub Thalamic Nucleus Micro Electrode Recordings

under different neuro-linguistic stimuli. The key novelty of this approach is that in calculating the NMP the entire vLFP (spikes + LFP) is used and no underlying model for the measured neural processes is assumed. This approach is perhaps particularly suitable for MER probes recording from neuron dense Sub Thalamic Nuclei which will pick up the contribution of many more neurons than can be expected to be detected with classical spike sorting approaches [96]. In theory the NMP will include the contribution of all neurons contributing to the MER probe, including sub threshold oscillations and lower frequency synaptic-dendritic connections. This is in contrast to the most common alternative of spike sorting strategies which characterise the measured system in terms of the behaviour of a small subset of neurons closest to the recording probe. The additional advantages of using the NMP for neural recordings are:

1. Applying the NMP to a time series represents a parameter free signals analysis methodology. This is in contrast to spike sorting approaches where multiple thresholding parameters must be chosen prior to analysis to set the detection and clustering of the identified spikes.

2. The assumptions required to apply the NMP are minimal. Using the NMP does not pre-suppose any underlying model for the data. The only requirement is that the measured signal is covariance (wide-sense) stationary so that the autocorrelation function is non-stationary and the power spectrum is defined.

The disadvantage of using the NMP to characterise neural signals is that fundamentally, the NMP provides a sensitive means to identify changes in a measured power spectrum. We cannot link these identified changes to the underlying biological behaviour of the neurons being measured, beyond broad statements about the presence of peaks (such as that seen in β band synchronisation) representing the recruitment of synchronised clusters.

This deficiency of the NMP will motivate the approach of the third chapter which is to model the constituent spike trains as an ensemble of filtered renewal processes. This approach will allow us to link changes in the measured spectrum to changes in the statistics

of the contributing spike trains, which is consistent with information theoretic approaches to characterise neuron firing patterns. This benefits of this approach will be at the expense of a much more exhaustive list of assumptions of the underlying signal.

3

Model-Based Entire vLFP Analysis

"All models are wrong, but some are useful. "

– George E. P. Box

3.1 Chapter Summary

In the previous chapter we used a model-free approach to analyse the entire vLFP (spikes + LFP) using the NMP. This methodology was successful in identifying changes in the state of the STN during the presentation of experimental stimuli, but these changes could only be described in terms of the measured spectrum, not the underlying biology of the system.

In this chapter we consider a model-based approach to analysing the entire vLFP. The model used is to assume that the vLFP which is recorded is generated from an ensemble of

independent neurons firing with renewal process statistics. This approach can be considered to provide the conceptual bridge between the changes in the identified power spectrum and changes in the underlying spike train statistics.

Using this methodology we will show, for non-Poisson firing statistics, that for an ensemble of independent neurons the common probability density defining the firing statistics of the constituent processes can be identified given an estimate of the power spectrum of the ensemble. Identifying this probability density given the measured power spectrum continues the strategy employed throughout this thesis of characterising the neural time series by solving inverse problems given statistical averages of the signal (in this case the measured power spectrum). In this chapter the inverse problem is fundamentally based on developing deconvolution strategies to the renewal equation.

3.2 Chapter Overview

In this chapter we consider a renewal model-based approach to analyse the entire vLFP (spikes + LFP). We seek to link the identified changes in the measured power spectrum of the vLFP to changes in the firing patterns of the contributing neurons modelled as renewal processes. We will then generate an algorithm, termed the *Spectral Density Estimator*, to solve the associated inverse problem: given the power spectrum of an ensemble of independent neurons, identify the probability distribution generating the renewal statistics.

The steps of this chapter will be as follows:

1. Provide an overview of the the properties of renewal processes necessary to understand the remainder of the chapter.

2. Develop the Spectral Density Estimation algorithm which, via Fourier inversion techniques, allows the identification of the probability distribution generating the firing statistics associated with an ensemble of neurons. This estimate requires the measured power spectrum associated with the ensemble. We show by extensive numerical simulation that this methodology can resolve the underlying statistics associated with an ensemble of identical non-Poisson neural processes.

3. Generate a model of an ensemble of neurons with identical independent renewal statistics which are variable distances from the recording probes. The distance is accounted for with a simplified model of dispersion and amplitude attenuation (to model the modifications to the electric fields of the neurons as they propagate from the neurons to the electrical probe) which is a function of distance from the probe. We compare how the Spectral Density Estimator and a widely used spike sorting algorithm 'Osort' estimate the underlying PDF driving the renewal statistics as the relative distance of the neurons to the recording probe is varied. For the case where there is no relative distance between neurons, the action potentials from all neurons are identical. This represents a pathological problem which most spike sorting algorithms (based on identifying differences in the identified action potential shapes) cannot solve. We show that for similar action potential shapes the Spectral Density Estimator is consistently

more accurate, but the superiority of the estimates decreases as the relative distance increases and the action potentials associated with the different neurons become more distinct.

4. We suggest that this Spectral Density Estimator may find utility in the situation of analysing vLFPs obtained from MERs in neuron dense structures like the Sub Thalamic Nucleus where multiple similarly orientated, near equi-distant neurons are contributing to the measured time series.

5. Separate to the development, simulation and analysis of the Spectral Density Estimator we show that the model of a neuron firing as a renewal process is equivalent to the widely used communications protocol of Digital Pulse Interval Modulation (DPIM) when the allowed firing times are allowed to occur along a continuum (as would be expected for a biological system) instead of only occurring at discrete times. We then explore how robust the associated spectrum of different neural spiking statistics are to the presence of firing time jitter. This analysis is provided in section A.7.1.

Submitted Work

The development, analysis and benchmarking of the Spectral Density Estimator is based on the following submitted, but as yet not published, journal article:

- J. Varghese, K. Weegink, P. Bellette, and A. Bradley, *"Spectral techniques to estimate renewal spiking statistics with near identical spike shapes*" Phys. Rev. E, 2017.

The analysis of the renewal theory model of a spike train as the continuum limit of a DPIM encoder (section A.7.1) and the discussion of the Bartlett spectrum (section 3.3.2) is based on the following conference article:

- J.J. Varghese, K.J. Weegink, P.A. Bellette, and A.P. Bradley, "*Spectral properties of neuronal pulse interval modulation,"* , in Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on, April 2015, pp. 1007-1011

Contribution

The novelty of this chapter is introducing a method which, given a set of a priori assumptions is able to estimate the PDF of an ensemble of neurons given just an estimate of the PSD. This method uses the entire signal (not just a subset of identified spikes), is parameter free and we show is capable of solving the pathological problem of resolving the firing time distribution of a renewal process embedded with a collection of identical renewal processes. Secondly we show that the power spectrum of a neuron firing as a Digital Pulse Interval Modulation encoder, in the continuum limit of allowed firing times is identical to that of the renewal processes frequently used to model individual neurons. Given that one of the primary roles of the neural system is to convey information through the spike trains, the link between such a biologically accepted model and a currently used communications protocol in fields such as fibre optics where the information is also encoded in the pattern of the pulse times [152], [153] provides an interesting link between computational neuroscience and communications engineering.

3.3 Renewal Processes Models of Neural Time Series

Individual neurons are complicated systems that respond to synaptic input currents with a multitude of different firing patterns. Indeed the dopaminergic neurons in the basal ganglia have been shown to generate periodic, random and even chaotic firing patterns [154]. Despite the complexity of the firing patterns, to good approximation, the shape of the action potential does not change. As a result it is believed that information is encoded by neurons in the timing of spikes [88], [89], [82], [83], [155]. Replicating these firing patterns by pumping noise into dynamical models inspired by the neuron physiology such as the Hodgkin-Huxley or Fitzhugh-Nagumo models is possible, but in all but a few cases is mathematically intractable and one has to rely on numerical approximations [90]. A viable alternative is to do away with the physiological underpinnings of action potential dynamics and treat the generated spikes in the train as an abstract collection of points in a set, with a probabilistic relation between the points. Models of this form are known as 'point-process models'. An

excellent introduction to the theory of point processes in neuroscience is provided in [72]. The interval between action potentials is of special significance in neuroscience and is termed the *Inter Spike Interval* (ISI). One of the most important subset of point processes in neuroscience are called *renewal processes* and are defined by the condition that the probability distribution describing the ISI time is *stationary* and independent of prior ISI times.

The importance of the renewal process stems from the drastic simplification in mathematical analysis the assumptions allow for, while still providing a reasonably accurate model of neural spike trains [69]. Indeed this seemingly simple approximation allows for the generation of a remarkably diverse set of spike trains [90]. The key mathematical simplification introduced by renewal processes is that all statistical properties (e.g. probability of the n^{th} action potential firing, the expectation and variance of spike count, correlation structure etc ...) are derivable solely from knowledge of the ISI PDF [156].

The use of renewal theory in neuroscience is widespread in both experiment [69], [70], [71] and theory [72],[73],[74]. There is evidence that the discharge patterns of neurons in certain anatomical zones of lower order mammals such as the spike trains from the retinal ganglion cells to the Lateral Geniculate Nucleus (LGN) of the thalamus in cats [75] and the responses of neurons in the Antero-Ventral Cochlear Nucleus [76] follow renewal statistics. More recently [77] has shown that the dopaminergic neurons of the Substantia Nigra Pars Compacta in rhesus macaques follow different classes of renewal statistics while performing saccadic eye movements in response to different visual stimuli. The popularity of renewal theory models, where the gap between firing times are statistically independent, is consistent with the Hodgkin-Huxley Na^+,K^+ channel model, where the equations governing these variables is reset after each spike [72]. Obviously this model is not always valid and their are many examples in the literature where non-renewal statistics have been identified. See [78] for a detailed review of these examples and proposed mechanisms for the generation of non-renewal statistics.

Similar to our analysis in the frequency domain, [157] showed that measurement of the power spectra of single neuron spike trains could be used to identify the associated firing rates and refractory times of a discharging neuron. These renewal theory approaches have also been used to model collections of neurons in neural networks. The power spectra of an ensemble of neurons following renewal statistics was developed in [158]. More recently there has been an interest in modelling the behaviour of neurons in sparse neural networks using renewal theory [79],[80],[81].

The work of these papers can broadly be considered extensions to Steins model of a neuron in a network [159]. In Stein's model the behaviour of a single neuron is described by a leaky integrate and fire neuron which is has its membrane voltage altered by synaptic barrages at the pre-synaptic terminal, the arrival times of which are described by Poisson processes [160]. Although undoubtedly a powerful model, as discussed in [158] it generates an inconsistency where the statistics of the input arrival times (Poisson) are inconsistent with the output firing statistics (non-Poisson) of the neuron which would be input to other neurons in the network. In [79] an analysis, including renewal statistic inputs, was performed to identify inputs which are consistent with their output. Similar to the approach we follow in this chapter statistical independence of the neurons in the network was assumed. In [80] a generalisation of Stein's model is introduced where the Poisson input is generalised to two (one excitatory and one inhibitory) pooled renewal processes. In [161] it was shown that non-Poisson input statistics could greatly affect the output statistics of a feedforward network. In [81] expressions for output train statistics of single integrate and fire neuron given a general time modulated renewal input processes were developed.

Renewal theory has also been used to construct spike sorting programs that detect and cluster spikes from individual neurons given noisy extracellular recordings which contain the contribution of multiple neurons embedded in noise. In [162] a model which assumes that the spike trains of the individual neurons follow log-normal renewal statistics and the spike amplitude is modulated with firing rate was constructed. This method used Bayes rule and

Markov Chain Monte Carlo techniques to develop estimates of the probability that identified spikes belonged to a specific cluster (referred to as 'soft clustering'). By incorporating information about both the spiking pattern and the spike amplitude this approach was able to successfully deal with the non-stationarity of spike shape problem present with burst firing where the spike shape of a single neuron varies when a sequence of spikes occur in rapid ($<$ 100 milliseconds) [163] [164] progression. This spike shape non-stationarity is a problem which is well known to introduce errors into standard spike sorting programs [57] [165].

3.3.1 Introduction

In this section we provide a brief overview of renewal processes in time series analysis required to understand the subsequent chapter. For excellent introductions to and further discussion of renewal processes see [156],[166]. Consider a sequence of *non negative* random variables $\{\Delta T_i\}$ which are independent and identically distributed (i.i.d.), drawn from an arbitrary Probability Density Function (PDF) with positive support:

$$\Delta T_i \sim p(t), \quad t \in [0, \infty) \tag{3.1}$$

The random variable can be as abstract as necessary, but in the context of this thesis it is best to consider a spike train, with sequence $\{\Delta T_i\}$ representing the time *between* spikes arriving at the synapse of a neuron [69],[78],[157]. The density, $p(t)$, determines the statistics of the spacing of these spikes.

Following [156] we define the combination of renewal events, S_n, which is itself a random variable representing the time to the n^{th} renewal (spike) event by:

$$S_0 = 0, \quad S_n = \Delta T_1 + \Delta T_2 + \cdots + \Delta T_n \tag{3.2}$$

Let N(t) denote the number of renewals (spikes in our situation) by time t, $t \geq 0$. In order for $N(t) = n$, that is n renewals (spikes) to occur at time t, the n^{th} renewal, S_n must occur *at latest* at time t and the $(n+1)^{th}$ renewal (spike) must occur *after* time t. That is:

$$\{N(t) = n\} \Leftrightarrow \{S_n \leq t, S_{n+1} > t\} \tag{3.3}$$

Stated more formally:

$$N(t) = \sup\{n : S_n \leq t\}, \quad t \geq 0 \tag{3.4}$$

It should be immediately obvious that the counting variable, $N(t)$, is a random variable. The expectation of the number of spikes at time t plays a pivotal role in renewal theory and is known as the *renewal function*: $M(t) = \mathbb{E}[N(t)]$. For this chapter we will be more interested in the derivative of this function, which is referred to as the *Renewal Density Function* (RDF):

$$m(t) = \frac{dM(t)}{dt} = \frac{d\mathbb{E}[N(t)]}{dt} \tag{3.5}$$

The RDF, $m(t)$, represents the rate of change of the expected number of events at time t. We will now show a dual interpretation of the RDF as a marker for how the renewal events are distributed in time. The RDF is related to the underlying PDF, $p(t)$, by the following Volterra convolution integral of the second kind with corresponding form in Laplace space [156]:

$$\begin{aligned} m(t) &= p(t) + \int_0^t m(t-t')p(t')dt' \qquad t \geq 0 && (3.6) \\ &\Updownarrow \\ M(s) &= \frac{P(s)}{1-P(s)} && (3.7) \end{aligned}$$

The Laplace transform shows that the RDF completely describes the PDF: $p(t)$. Historically there has been more concern with determining the RDF given a known PDF, whereas here we are interested in determining the PDF from an estimated RDF. We can write the RDF in a form more appropriate for understanding its relation to the distribution of renewal events by applying the geometric sum formula to (3.7) (recall $P(s) < 0$, except at $s = 0$) and applying the inverse Laplace transform:

$$\begin{aligned} m(t) &= \mathcal{L}^{-1}\{P(s)\}(s) + \mathcal{L}^{-1}\left\{P(s)^2\right\}(s) + \mathcal{L}^{-1}\left\{P(s)^3\right\}(s) + \cdots && (3.8) \\ &= \underbrace{p(t)}_{\text{spike 1}} + \underbrace{\int_0^\infty p(t')p(t-t')dt'}_{\text{spike 2}} + \underbrace{\int_0^\infty \int_0^\infty p(t')p(t'')p(t-t'-t'')dt'dt''}_{\text{spike 3}} + \cdots && (3.9) \end{aligned}$$

Thus the RDF can be represented as an infinite sum of n-fold convolutions of the PDF. Recall that the PDF associated with the sum of two independent random variables is the convolution of their respective PDFs. Therefore the first term in (3.9) represents the probability of the first spike occurring at time t, the second term represents the probability of the second spike occurring at time t and so forth. Thus we see that the RDF can be interpreted as the probability of seeing *any* spike at time t. Note that as we show below, the RDF does not integrate to unity and is therefore not a true probability density function.

We will outline some additional properties of the RDF useful for our analysis.

1. *Poisson Behaviour*: The RDF for the Poisson process, which has exponential PDF: $p(t)_{\text{Poisson},\nu} = \nu e^{-\nu t}\theta(t)$, is given by the constant value: $m(t)_{\text{Poisson},\nu} = \nu\theta(t)$, [167] where $\nu = \left(\int_0^\infty tp(t)dt\right)^{-1}$ is the Mean Firing Rate (MFR) and $\theta(t)$ is the Heaviside distribution. Therefore for a Poisson counting process the probability of seeing a spike is constant.

   The Poisson counting process is the only continuous distribution which is *memoryless*. That is to say that if we have waited time period t without a spike, ΔT, occuring $P(\Delta T \geq t)$, the probability of a spike, ΔT, occurring during an additional δt time period $P(\Delta T \leq t + \delta t | \Delta T \geq t)$ is the same the probability of observing a spike in any arbitrary δt time period. We can mathematically express this as:

   $$Pr(\Delta T \leq t + \delta t | \Delta T \geq t) = Pr(0 \leq \Delta T \leq \delta t) \tag{3.10}$$

2. *Asymptotic Behaviour*: By the Erdos-Feller-Pollard theorem [168] the RDF asymptotes to the constant value of the mean firing rate ν: $\lim_{t\to\infty} m(t) = \nu$. In the less general case of sufficiently smooth PDFs we can see this result by applying the Final Value Theorem to the Laplace Transform of the RDF given by (3.7). We introduce the square integrable *Asymptotically Shifted Renewal Density Function* (ASRDF) $m_{\text{AS}}(t)$:

   $$m_{\text{AS}}(t) = m(t) - \nu\theta(t) \tag{3.11}$$

3. *Non-Negativity*: Inspection of (3.9) shows that the RDF is an infinite series of n-fold convolutions of strictly positive PDFs. Since the convolution of two strictly positive functions is positive, the RDF must be strictly positive.

4. *Positive Support:* Inspection of (3.9) also shows that the RDF is an infinite series of n-fold convolutions of PDFs defined for positive support. Thus the RDF, $m(t)$, is also only defined for positive support. We define symmetric functions: s-RDF and s-AS-RDF which are defined over the entire (positive and negative) $\mathbb{R}^1$ number line:

$$\text{s-RDF} \equiv m(t) + m(-t) \tag{3.12}$$

$$\text{s-ASRDF} \equiv m(t) + m(-t) - \nu \tag{3.13}$$

Note that by the properties of Fourier Transforms of symmetric functions [121], the FT of these symmetric forms of the RDF will be purely real functions.

The renewal process can be represented as a time series where, at the time of each event, a Dirac delta distribution is embedded into the continuous number line $\mathbb{R}^1$. The spaces between the Dirac delta distributions are the random variables ΔT_i. The times series is thus: $x(t) = \sum_{i=1}^{\infty} \delta(t - t_i)$, where $(t_i - t_{i-1}) \equiv \Delta T_i$. For modeling many real world phenomena, when the event occurs a characteristic shape is generated. For example the voltage at the synapse of a neuron will generate a characteristic action potential shape when the spike occurs. We can incorporate this effect into the time series by convolving the renewal process (represented as the Dirac comb) with an appropriate filter function g(t). This is referred to as a filtered renewal process and has the following time series form:

$$y(t) = g(t) * \sum_{i=1}^{\infty} \delta(t - t_i) \tag{3.14}$$

The auto-correlation structure, $R_y(\tau)$, for this time series is given by the convolution of the pulse shape, $g(t)$ with the correlation structure associated with the pulse times [157], [167]:

$$R_y(t) = \nu \int_{-\infty}^{\infty} g(t - t') \left[\delta(t') + m(t') + m(-t')\right] dt' \tag{3.15}$$

Where $\nu = \left(\int_0^\infty tp(t)dt\right)^{-1}$ is the Mean Firing Rate (MFR). Recall that the RDF, $m(t)$ is only defined for $t \geq 0$. Thus the autocorrelation is a function of the s-RDF.

A closed form expression for the power spectrum of the filtered renewal process can be developed, using the Weiner-Khinchtine theorem, by taking the Fourier Transform (FT) of the autocorrelation function (3.15) of this process [157], [167], [169]. In Appendix A.5 we develop this expression along a different route, by considering the infinite limit of the periodogram of the time series. The power spectrum, $P(\omega)$, is given by:

$$P(\omega) = \nu G(\omega)\left[2\pi\nu\delta(\omega) + 1 + \underbrace{2Re\left\{\frac{H(\omega)}{1-H(\omega)}\right\}}_{\text{Correlation Spectrum: } 2\cdot\Phi(\omega)}\right] \tag{3.16}$$

Where $G(\omega) = |\mathcal{F}[g(t)]|^2$ is the energy spectrum associated with the filter shape g(t), and $H(\omega)$ is the characteristic function of the $\{\Delta T\}$ random variables. Note that if ΔT admits a distribution then the characteristic function is given by the FT of the PDF: $H(\omega) = \mathcal{F}\{p(t)\}(\omega)$. The expression in the square parentheses $[\cdots]$ is termed the Barltett spectrum and, by the Weiner-Khincthine theorem, represents the FT of the auto-correlation structure of the pulse train [157]. The non constant component of the Bartlett Spectrum, $\Phi(\omega) = Re\left\{\frac{H(\omega)}{1-H(\omega)}\right\}$, which we will refer to as the *correlation spectrum* will play a central role in our subsequent analysis.

It is important to note that (3.16) is the analytic expression for the power spectrum of a filtered renewal process. Specifically (3.16) *does not* represent the periodogram used to estimate the power spectrum given a single realisation (i.e. the time series) of a renewal process. Thus the $G(\omega)$ term represents the contribution to the analytical power spectrum due to the stereotypical shape of the action potential generated at every firing time. It *does not* represent a window or data taper used in the estimation of the power spectrum of the measured time series.

3.3.2 Analysis Of the Bartlett Spectrum

The Spectral Density Estimator discussed in section 3.5 is based heavily on the estimation and subsequent integral transformation of the correlation spectrum $\Phi(\omega)$. In section 3.3.3 we will show the unintuitive result that given the correlation spectrum, $Re\left\{\frac{H(\omega)}{1-H(\omega)}\right\}$, which is a nonlinear function of $H(\omega)$, we can recover $H(\omega)$ and thus the density function $p(t)$.

Before we begin the formal analysis of the correlation spectrum,$\Phi(\omega)$, we provide some general properties of both the Bartlett and correlation spectrum (which are related by a constant, unity, offset). Notice that by inspection of (3.16) due to the non-negativity requirements of the power spectrum the Bartlett spectrum must be non-negative. There is no such requirement of the correlation spectrum.

The Bartlett spectrum represents the correlation structure of the spiking times in frequency space. A problem is that despite the simple expression for the Bartlett spectrum, for most probability distributions it is remarkably resistant to closed form analysis. It is rarely possible to develop closed form expressions for the conditions for the presence of peaks, their locations or width or indeed how rapidly the spectrum decays to zero. Instead we are forced to rely on heuristics to understand the Bartlett spectrum. For example if a neuron is firing perfectly periodically we would expect to see peaks in the power spectrum at the firing rate and its harmonics. Purely periodic firing represents one of the few cases we can develop an analytic expression for the Bartlett spectrum:

$$\begin{aligned} H(\omega)_{\text{periodic}} &= \int_0^\infty \delta(t-\frac{1}{\lambda})e^{i\omega t} = e^{(i\omega)/\lambda} \\ \sum_{n=1}^{\infty} H(\omega) + \sum_{n=1}^{\infty} H^*(\omega) &= \sum_{n=-\infty}^{\infty} \delta(\omega - n\lambda) \\ \nu\left[1+\Phi(\omega)\right]_{\text{Periodic}} &= \nu + \nu \sum_{n=-\infty}^{\infty} \delta(\omega - n2\pi\lambda) \end{aligned}$$

As the variance associated with the timing of these spikes increases we would expect to see the peaks get wider, and the higher order harmonics begin to disappear. As the variance is increased such that the firing pattern becomes maximally random (that is, it becomes a *Poisson* counting process) there would be no statistically expected deviation from uniformly distributed firing times and thus we would expect the Bartlett spectrum to be a constant. This is another case where an analytic expression for the Bartlett spectrum can be developed, and it can be shown [167], in agreement with Carson's theorem the Bartlett spectrum is a constant:

$$\begin{aligned} H(\omega)_{\text{Poisson}} &= \frac{\lambda}{\lambda - i\omega} \longrightarrow 2\text{Re}\left(\frac{H(\omega)}{1 - H(\omega)}\right) = 0, \\ \nu\left[1 + \Phi(\omega)\right]_{\text{Poisson}} &= \nu + \nu^2\delta(f) \end{aligned}$$

It is convenient to identify that these two firing patterns represent extremes of the Weibull family of distributions for the ISI random variable:

$$p(t)_{\text{ISI}} = \left(\frac{k}{\lambda}\right)\left(\frac{t}{\lambda}\right)^{k-1} e^{-(t/\lambda)^k}, \qquad t, \lambda, k \geq 0 \tag{3.17}$$

Where λ and k are termed the scale & shape parameter respectively. When k is unity the maximally random Poisson statistics are recovered, whereas periodic firing occurs in the limit of the shape parameter, k, approaching infinity. The variance of the Weibull distribution is given by: $\sigma^2 = \lambda^2\Gamma(1+2/k) - \mu^2$. Notice that as k increases the variance decreases, and the Poisson solution has maximal variance. Between the two extremes of periodic and Poisson count firing statistics we can observe the Bartlett spectrum by parametrically fitting the ISI density function to the Weibull distributions and tuning the shape parameter. Examples of this process are shown in Figure 3.1 for $k = 1$, 5, 10 and validate this heuristic approach. Inspection of Figure 3.1 shows that the peaks do not occur precisely at the harmonics of the fundamental (mean firing rate, ν) frequency. This is because these are not spectra of purely periodic processes .

There are two important points to observe from Figure 3.1. Firstly for the different Weibull ($k = 1$, $k = 5$ and $k = 10$) ISI distributions considered, with mean firing rate of 30 Hz, there is minimal structure (beyond constant behavior) in the Bartlett spectrum

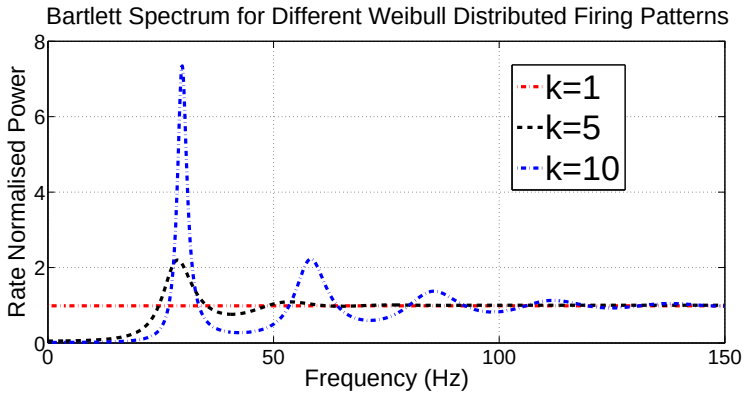


Figure 3.1: Bartlett spectrum of Weibull distributions with different shape parameters k=1,5,10 for a neuron with mean firing rate of 30 Hz. Notice that as the firing becomes less periodic (k decreases) the higher order peaks will disappear, the remaining peaks get smaller and wider until eventually with Poisson counting statistics (k=1) the Bartlett spectrum shows no features. Notice that as these are not purely periodic processes the peaks do not occur precisely at the harmonics of the fundamental (30 Hz) frequency.

above roughly 150 Hz. This intuitively makes sense, as the probability of the time between firing events for these Weibull distributions being five times smaller than the mean time to fire is exceedingly low. Therefore we would expect minimal correlations at firing rate rates that are unlikely to occur. To compare these simulation values with real STN neurons, in [170] the firing rates from 351 MER signals from 65 patients with Parkinsons Disease were measured, on average, to be 40 ± 20.3 Hz. This is a similar mean firing rate to the 30 Hz we have considered for these simulations. It is important to realise that while neurons can *temporarily* fire at much faster rates (for example during burst firing events) this represents an *instantaneous* increase in the firing rate, and not an increase in the *mean* (i.e. statistical expectation) of the firing rate. For example it can be shown [171] that high frequency burst firing patterns can be generated by stretched exponential ($k < 1$) Weibull ISI distributions with standard physiological *mean* ($\nu \approx 30$ Hz) firing rates. The second important point to note is that there is no structure in the Bartlett spectrum for a Poisson counting process beyond constant amplitude equal to the mean firing rate.

3.3.3 Analysis of the Correlation Spectrum

In the proceeding section we will show the unintuitive result that given the correlation spectrum, $Re\left\{\frac{H(\omega)}{1-H(\omega)}\right\}$, which is a nonlinear function of $H(\omega)$, we can recover $H(\omega)$ and thus the density function $p(t)$. The key to this result is the positivity and non-negativity properties of the RDF. Inspection of (3.15) and (3.16) shows that the correlation spectrum $\Phi(\omega)$ and the symmetric form of the Renewal Density Function (s-RDF) are a FT pair [157]:

$$\mathcal{F}\left\{m(t)+m(-t)\right\}(\omega) = 2\Phi(\omega) + 2\pi\nu\delta(\omega) \tag{3.18}$$

The Dirac delta term $\delta(\omega)$ is present because the RDF asymptotes to ν and thus is not square integrable. Because the support of the renewal density function is $\in [0,\infty)$ there is no 'overlap' between the causal, $m(t)$, and anti-causal, $m(-t)$, components. We can develop the FT of the (one sided) renewal density function by inserting a Heaviside distribution, $\theta(t)$, into the Fourier kernel and use the following identity developed from the convolution theorem [121]:

$$\mathcal{F}\left[f(t)\theta(t)\right] = \frac{1}{2}\left(F(\omega) - j\mathcal{H}\left[F(\omega)\right](\omega)\right) \tag{3.19}$$

Recall that $\mathcal{H}\left\{f(t)\right\}(\omega)$ is the Hilbert transform defined (in chapter 2) as $\mathcal{H}\left\{f(t)\right\}(\omega) = \frac{1}{\pi}\text{p.v.}f(t) * \frac{1}{t}$, where p.v. is the Cauchy principle value. Inserting the Heaviside distribution into the Fourier Kernel of (3.18) and using (3.19) yields:

$$\begin{aligned}
\mathcal{F}\left\{m(t)\right\}(\omega) &= \mathcal{F}\left\{[m(t)+m(-t)]\theta(t)\right\}(\omega) \\
&= \frac{1}{2}\left(2\Phi(\omega) + 2\pi\nu\delta(\omega)\right) - \frac{i}{2}\left(\mathcal{H}\left\{2\Phi(\omega) + 2\Phi(\omega)\right\}(\omega)\right) \\
&= \Phi(\omega) - i\mathcal{H}\left[\Phi(\omega)\right] + \pi\nu\delta(\omega) - \frac{i\nu}{\omega}
\end{aligned} \tag{3.20}$$

Thus given an estimate of the correlation function, $\Phi(\omega)$, and the mean firing rate, ν, the RDF can be recovered. This result can also be developed using the Kramers-Kronig relationship, recognising that the FT of a causal function will be complex, with the real and imaginary components of this transform related by a Hilbert transformation in the frequency

domain. For practical purposes, from (3.18) we can obtain the AS-RDF, $m_{\text{AS}}(t)$, by Fourier inversion of $\Phi(\omega)$ multiplied by the Heaviside distribution:

$$m_{\text{AS}}(t) \equiv m(t) - \nu\theta(t) = \mathcal{F}^{-1}\left\{\Phi(\omega)\right\}(t)\theta(t) \tag{3.21}$$

Note that the inversion of $\Phi(\omega)$ recovers the AS-RDF: $\hat{m}_{AS}(t)$. We must separately develop an estimate of ν in order to estimate the RDF: $m(t)$. We specifically discuss this issue in Section 3.5.3.

Uniqueness of Correlation Function

Notice that we can express $\Phi(\omega)$ as the FT of the *difference* of the s-RDF ,$m(t) + m(-t)$, of a given PDF with mean firing rate ν and the s-RDF of a Poisson process with rate ν.

$$\mathcal{F}[(\underbrace{m(t) + m(-t)}_{\text{s-RDF: } \int_0^\infty tp(t)dt = \nu}) - \underbrace{(\nu\theta(t) + \nu\theta(t))}_{\text{Poisson s-RDF},\nu}] = \Phi(\omega) \tag{3.22}$$

Thus $\Phi(\omega)$ is the FT of the difference of two uniquely defined distributions. We argue that, excluding Poisson processes, $\Phi(\omega)$ completely describes $p(t)$. The Poisson process represents a degeneracy point where $\Phi(\omega)$ is zero [167] for all mean firing rates ν. Thus for the Poisson process, $\Phi(\omega)$ does not completely describe the PDF $p_{\text{poisson}}(t)$.

3.4 Super-Position of Renewal Processes

We now consider the super-position of many independent renewal processes. In Section 3.7 we will use a modification of this super-position of renewal processes to model an extracellular recording probe interacting with the electrical fields of multiple neurons. We consider below the time series, $y(t)_\Sigma$, associated with the restrictive case when the processes all have the same filter shape but variable amplitude, a, uncorrelated with the firing times ($\langle a\Delta T\rangle = \langle a\rangle\langle\Delta T\rangle$). This is referred to as spike shape stationarity. Each renewal process is independent with the spacing between pulses ΔT for each process being drawn from the

same, but separate, PDF p(t):

$$y(t)_{\Sigma} = \sum_{j=1}^{N} a_j g(t) * \left(\sum_{i=1}^{\infty} \delta_j \left(t - \Delta T_{i,j} \right) \right) \quad \forall j \quad \Delta T_i \sim p(t) \tag{3.23}$$

Where $\Delta T_{i,j}$ refers to the i^{th} pulse gap associated with the j^{th} renewal process. In general this super-position of independent renewal processes results in a pooled process which is *not* a renewal process, but a subset of Markov-renewal processes [172]. The key reason for this is, due to the summation, the independence of spacing between the pooled spike stream is violated. The properties of the infinite limit of summed renewal process was well studied in [158]. It was shown that the general properties of a pooled renewal process are that the random variable associated with the spacing, the Inter-Spike-Interval (ISI), is drawn from an exponential distribution (similar to a Poisson counting process) but with correlations between these random variables. This is *not* a property of renewal processes. Most interestingly the power spectrum of this super-position is a scaled form of the individual processes contributing to the pool. We can trivially extend the analysis in [158] to the summation of filtered renewal processes with variable amplitude:

$$P_{\Sigma}(\omega) = N\nu\langle a^2 \rangle G(\omega) \left[1 + 2Re \left\{ \frac{H(\omega)}{1 - H(\omega)} \right\} \right] \tag{3.24}$$

The derivation of this result is provided in section A.6. Note that $\langle a^2 \rangle$ is the square of the mean of the filter shape amplitudes. In effect the super-position of renewal processes creates a non-renewal process, but with a renewal process power spectrum. As mentioned in [158] this spectral result can be understood as a consequence of the variance of the summation of independent random variables being the sum of the variance of those variables. The property we will exploit is that the pooled renewal process, although having ISI different to the individual contributing processes has (up to a scaling factor) the same correlation structure, $\Phi(\omega)$, as the individual contributing processes. That is (3.24) is scaled version of (3.16)

Note that it is frequently misinterpreted that the *Palm-Khincthine theorem* guarantees that the pooled renewal process is a Poisson renewal process. As is emphasized in [158], this theorem only guarantees that for the summation of low intensity renewal processes, on

a time scale smaller than the renewal times of the individual processes, that the resulting process will appear as a Poisson process.

3.5 Spectral Density Algorithm

Our algorithm to estimate the PDF, shown in Figure 3.2, consists of three iterated steps: Step 1: estimate the correlation spectrum $\Phi(\omega)$ and obtain the AS-RDF $m_{\text{AS}}(t)$. Step 2: estimate the mean firing rate $\hat{\nu}$ to obtain an estimate of the RDF, $\hat{m}(t) = \hat{m}_{\text{AS}}(t) + \hat{\nu}\theta(t)$. Step 3: estimate the density function by solving (3.6) for $p(t)$. In order to estimate ν we will use an iterative procedure which will require multiple cycling of steps 2-3.

In explaining the following steps we consider the estimation of $\Phi(\omega)$ of the super-position of an unknown number of renewal processes with unknown amplitudes in the presence of zero mean white noise with variance σ^2 which is uncorrelated with the renewal processes:

$$y(t) = y(t)_{\Sigma} + w(t), \quad w(t) \sim \mathcal{N}(0, \sigma^2) \tag{3.25}$$

3.5.1 Step 1: Correlation Spectrum Estimation

The first requirement for an accurate estimate of $\Phi(\omega)$ is obviously an accurate estimate of the power spectral density. As we are generating a data driven estimate of the PDF, we wish the estimates to be as general as possible and thus restrict the spectral estimation to non-parametric methods. In the proceeding sections we use Welches Overlapping Segment Averaging (WOSA) [173]. In certain circumstances, for instance if only short time recordings are available, alternative methods such as Thompson's multitaper method, (which, instead of partitioning the time series, reduces variance in the estimate with a series of orthogonal data tapers applied to the entire time series) [174],[125] may be more appropriate.

There are alternative approaches to estimating the power spectrum of more general stationary point processes (of which renewal processes are a special case) based on knowledge of the precise timing of the spikes observed over the length of the measured time series

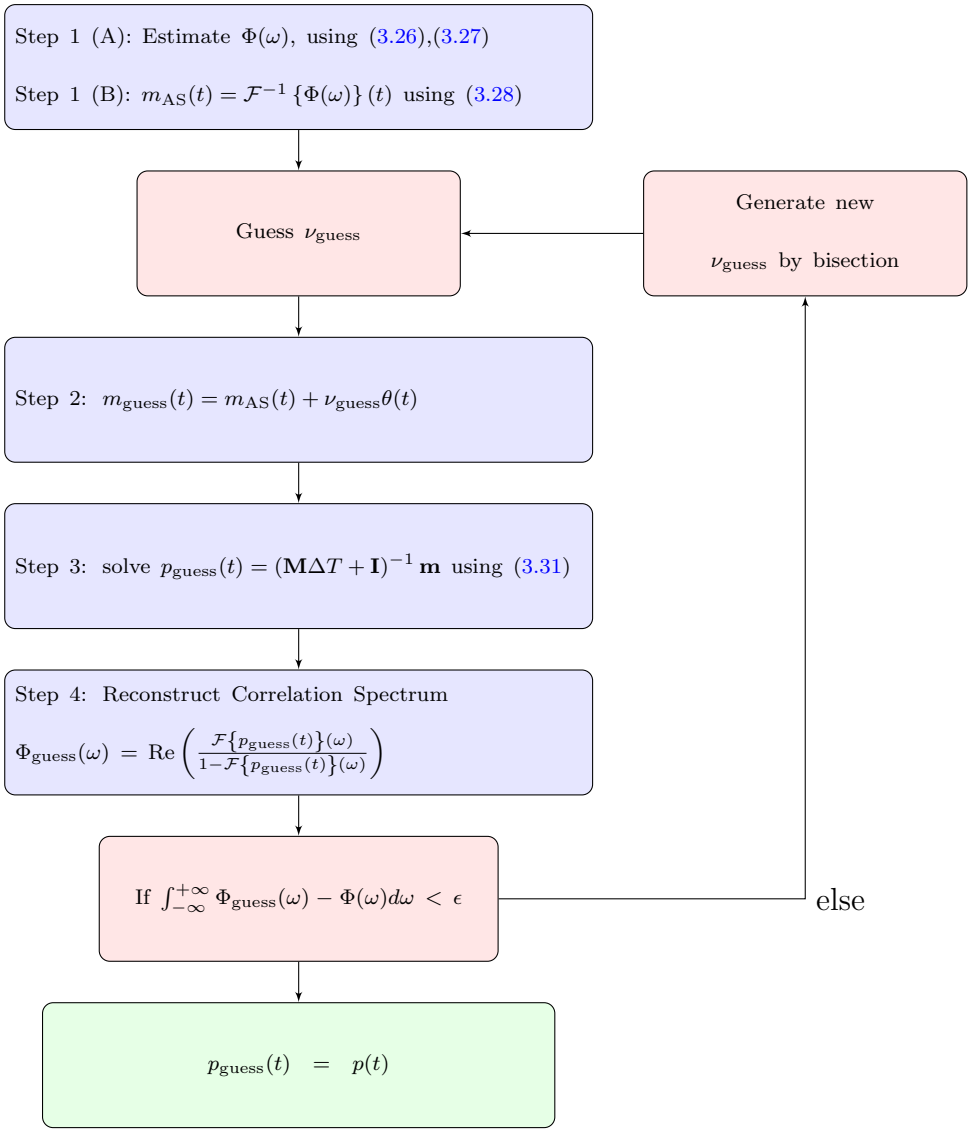


Figure 3.2: Spectral Density Estimator algorithm flow chart.

[175],[176]. Thus, instead of developing the power spectrum from a sequence of windows of the time evolution of the stochastic process (i.e. the standard periodogram approach) the spectrum is estimated from the discrete set of firing times which occur in these windows of time. These approaches which require detailed knowledge of the individual neuron firing times *cannot* be applied (without significant post-processing) to the vLFP problem because we do not know the precise firing times of the constituent neurons contributing to the noisy single channel MER signal.

The ability to estimate $\Phi(\omega)$ is determined by the signal to noise ratio of the energy spectrum $G(\omega)$ to the background white noise in the region of support of $\Phi(\omega)$. This support is dependent on the physical problem considered. For example with extra-cellular neural recordings, the firing rates of individual neurons are roughly in the 1-150 Hz range. Thus we would expect the support of $\Phi(\omega)$ for the neural system to be roughly (0-150 Hz). We show the Bartlett spectra associated with different gamma distributed renewal processes with physiological neural mean firing rate of 50 Hz as well as the typical energy spectrum for the spike shape of an STN neuron, $G(\omega)$, in Figure 3.4.

We estimate $\Phi(\omega)$ in the following steps: Given estimates of the PSD, $\hat{P}(\omega)$, and energy spectrum, $G(\omega)$, either by a priori knowledge or estimated from the signal, perform the blind deconvolution: $\chi(\omega) = \hat{P}(\omega)/G(\omega)$. Using (3.24)-(3.25) this can be shown to be equivalent to:

$$\chi(\omega) = N\nu\langle a^2\rangle\,[1 + \Phi(\omega)] + \frac{\sigma^2}{G(\omega)} \tag{3.26}$$

By the Riemann-Lebesgue lemma [177], at sufficiently high frequencies $\Phi(\omega)$ will decay to zero. As discussed previously on heuristic grounds we expect that this will decay to zero around the value of the firing rate. We also expect, excluding the unphysical situation $g(t) = \delta(t)$, that for sufficiently high frequencies $\lim_{\omega\to\infty} G(\omega) = 0$. Thus at high frequencies the noise component of $\chi(\omega)$ (where the correlation spectrum will most likely be zero) may diverge. If we choose a range $[\omega_L \cdots \omega_H]$ where we expect $\Phi(\omega)$ to be approximately zero but $G(\omega)$ to be sufficiently large we can develop an approximation for the product of unknown quantities $N\nu\langle a^2\rangle$:

$$\alpha = N\nu\langle a^2\rangle = \frac{1}{M}\sum_{\omega=\omega_L}^{\omega=\omega_H}\chi(\omega), \qquad M = \frac{\omega_H - \omega_L}{\Delta\omega} \tag{3.27}$$

Using (3.26),(3.27) we can develop our estimate of the correlation spectrum as: $\hat{\Phi}(\omega) \approx \chi(\omega)/\alpha - 1$. Because of the noise term of $\chi(\omega)$ diverging at higher frequencies we can low-pass filter our solution. We use a *shifted* complementary Gaussian error function: $\mathrm{erfc}(\omega - \omega_0) = \frac{2}{\sqrt{\pi}}\int_{\omega}^{\infty} e^{-(\omega'-\omega_0)^2}\,\mathrm{d}\omega'$. as a low-pass filter, with ω_0, the cut off frequency, set at a value higher

than the highest expected firing rate. The AS-RDF, $m_{\text{AS}}(t)$, is recovered using (3.21):

$$\hat{m}(t)_{\text{AS}} = \mathcal{F}^{-1}\left\{\text{erfc}(\omega - \omega_0) \cdot \left[\frac{\chi(\omega)}{\alpha} - 1\right]\right\}(t)\theta(t) \tag{3.28}$$

Notice that the algebraic manipulation required to obtain $\Phi(\omega)$ from $P(\omega)$, from (3.24), is equivalent to a deconvolution of the filter shape and the s-ASRDF, in the time domain. Thus an alternative approach based on Tikhonov regularisation of the auto-correlation function in the time domain could also be employed [178].

3.5.2 Step 2: Solving the Volterra Integral Equation

We will discretise the RDF: $\mathbf{m} \rightarrow \mathbf{m}[t]$ and the desired PDF $p(t) \rightarrow \mathbf{p}[t]$ and use numerical quadrature techniques to perform the deconvolution required to solve (3.6) for $p(t)$. Note that if we discretise these continuous quantities at a sampling rate, F_s, the numerical time step is given by: $\Delta T = 1/F_s$. We can write the convolution term as a matrix-vector product:

$$\int_0^t m(t-t')p(t')dt' \approx \mathbf{M}\mathbf{p}\Delta T \tag{3.29}$$

$$\mathbf{M}\mathbf{p} = \begin{bmatrix} m[1] & 0 & 0 & \cdots & 0 \\ m[2] & m[1] & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ m[n] & m[n-1] & m[n-2] & \cdots & m[1] \end{bmatrix} \begin{bmatrix} p[1] \\ p[2] \\ \vdots \\ p[n] \end{bmatrix} \tag{3.30}$$

We can solve this for $\mathbf{p}[t]$ as:

$$\mathbf{p}[t] = (\mathbf{M}\Delta T + \mathbf{I})^{-1}\,\mathbf{m}[t] \tag{3.31}$$

Since the RDF does not asymptote to zero we cannot use Fourier Transform techniques to estimate the discretised distribution $\mathbf{p}[t]$. This problem can be solved using Ordinary Least Squares (OLS), or Tikhonov regularisation with a penalty associated with either the ℓ_2 norm of the vector $\mathbf{p}[t]$ or its first order difference (to promote smooth solutions). In the proceeding sections we solve (3.31) using OLS.

A problem now is that the methodology in Section 3.5.1 determines an estimate of AS-RDF: $m_{\text{AS}}(t)$. From (3.21) we require an estimate of the MFR, ν, to construct an estimate of the RDF. We describe an iterative process to achieve this in the next section.

3.5.3 Step 3: Mean Estimation

It is important to notice that the RDF with different offsets: $m'(t) = m(t) + \nu'\theta(t)$ yields a structurally different integral equation to the standard Volterra integral equation of the 2^{nd} kind (3.6). Re-writing the RDF in (3.6) as: $m(t) = m_{\text{AS}}(t) + \nu_{\text{guess}}\theta(t)$ yields the expression:

$$m_{\text{AS}}(t) + \nu_{\text{guess}}\theta(t) \quad = \quad p_{\text{guess}}(t) + \int_0^t m_{\text{AS}}(t-t')p_{\text{guess}}(t')dt' + \nu_{\text{guess}}\int_0^t p_{\text{guess}}(t)dt \tag{3.32}$$

There is no guarantee that the solution, $p_{\text{guess}}(t)$ for all ν_{guess}, of this modified renewal equation will satisfy the properties of a PDF of integrating to unity or being strictly *non-negative*. We can identify that in the limit of $\nu_{\text{guess}} >> m(t) \quad \forall t$, (3.32) reduces to the standard renewal equation for a Poisson process with mean firing rate ν_{guess}.

This suggests that we can estimate the mean firing rate ν for a non-Poisson process using the following iterative procedure:

1. Choose a value v_{guess} on the interval $[0,+v_{\max})$, where $v_{\max}$ is determined by an estimate of the maximum possible firing rate. The lower bound is because the mean firing rate cannot be negative.

2. Given the estimate $\hat{m}_{AS}(t)$ and choice v_{guess}, generate $\hat{m}_{guess}(t)$ and solve (3.32) for $p_{\text{guess}}(t)$.

3. Form the associated correlation spectrum: $\Phi_{\text{guess}}(\omega) = 2\text{Re}\left\{\dfrac{\mathcal{F}\left\{p_{\text{guess}}(t)\right\}(\omega)}{1-\mathcal{F}\left\{p_{\text{guess}}(t)\right\}(\omega)}\right\}$.

4. Determine the sum of squared error: $\int_{-\infty}^{+\infty}\left(\Phi(\omega) - \Phi_{\text{guess}}(\omega)\right)d\omega$

5. Use a bisection method to determine the $v'_{\min}$ which minimises the sum of squared error. $\hat{v} \to v_{\text{guess}}$ (in the proceeding section we use the golden search method). Return to step 2.

3.5.4 Degeneracy of Poisson Processes

Notice that for Poisson processes $Re\left\{\frac{H(\omega)}{1-H(\omega)}\right\} = 0$ [167] and thus $P_{\text{Poisson}}(\omega) = N\nu$. There is no way without a-priori knowledge to tell from the spectrum whether the system consists of a single Poisson process firing with rate $= N\nu$, N oscillators firing with rate $= \nu$ or m oscillators firing with rate $= \nu/m$. This is a consequence of $\Phi(\omega) = 0$ and thus it represents a degeneracy point. This limits the utility of this method to the super-position of Poisson firing oscillators without a priori knowledge of the mean firing rate. We note however that if ν can be known a priori (e.g. by biological or technical constraints of the problem), the SDE will successfully recover the PDF. We also note that although it is certainly less useful, inspection of the power spectrum associated with the superposition of renewal processes will inform us that the underlying processes are following Poisson statistics, but with an unknown rate.

3.6 Validation of Methodology

In this section we validate the methodology described by estimating the PDF associated with super-positions of filtered renewal processes. We consider spikes with *gamma* firing statistics. The gamma distribution is given by $p(t) = \dfrac{\lambda^k t^{k-1} e^{-\lambda t}}{\Gamma(k)}$. Where k and λ are referred to as the shape and scale parameters respectively. For the gamma distribution the MFR, ν_γ, is related to these parameters by the relation: $\nu_\gamma = \lambda/k$.

We choose the gamma distribution for two reasons. Firstly Poisson firing is a special case of the gamma distribution, with $k = 1$. This allows us to explore how the accuracy of this method varies as the firing statistics become more correlated (less Poisson-like). Recall in Section 3.5.4 we identified that the iterative process used to estimate the mean firing rate,

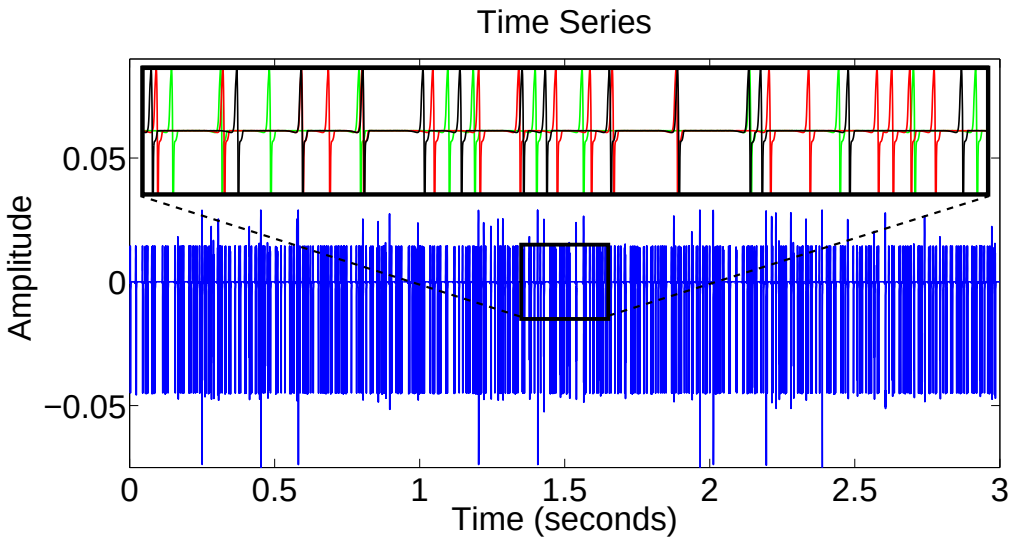


Figure 3.3: Sample time series of the superposition of three identical filtered renewal processes and a zoomed in version of the individual constituent (different colours) processes.

ν, (and subsequently the PDF) breaks down for the case of Poisson statistics. By varying the shape parameter, k, we can quantify the accuracy of the estimates $\hat{\nu}$, the PDF and subsequent statistical moments (which are calculated using this estimate of the PDF) as the firing patterns become more correlated. The second reason is that, as mentioned previously, in Section 3.7 we apply this method to simulated datasets of neurons interacting with an extra-cellular probe. Gamma statistics are frequently used to model neuron firing patterns [179],[180],[181].

We consider the problem of attempting to estimate the PDF for a range of different shape parameters from $k = 1.8$ (near-Poisson) to $k = 4$ (strongly correlated). We also consider whether the accuracy of the estimates vary with the MFR over the range $\nu = 20$Hz to $\nu = 200$Hz. We consider a filter shape, g(t), which is the action potential generated by solving the Hodgkin-Huxley equations with the appropriate parameters for an STN neuron [182]. We consider this filter function as *known apriori*, although we note that it would be possible to estimate this filter function, albeit less accurately, from the times series.

We generate the resulting time series as the superposition of three filtered gamma renewal processes, but do not have a priori knowledge about this number of processes. A sample realisation of the time series is shown in Figure 3.3. We use a transition frequency $\omega_0 = 500$Hz for the low-pass Filter described in Section 3.5.1. A plot of the Bartlett spectra for different shape parameters, with mean firing rate $\nu = 50$Hz, the energy spectrum $|G(\omega)|^2$ and the low-pass Filter are shown in Figure 3.4. Notice the support of the non unity component of the Bartlett Spectra, the correlation spectrum $\Phi(\omega)$, overlaps with the support of $|G(\omega)|^2$ which is vital for the blind deconvolution described in Section 3.5.1 to be successful. Also note that for almost all shape parameters considered the non-unity component of the Bartlett spectrum is within roughly 5% of its final value by 150 Hz and continues to decrease at higher frequencies. The majority (but not all) of the support of the correlation spectrum being restricted to the low frequency (< 150 Hz) region will be important in Section 3.7 when we consider the low-pass filter effects of the extracellular fluid which causes frequency selective attenuation of the measured vLFPs. Recall that frequency selective attenuation of the measured vLFP will alter the estimate of power spectrum and in turn the correlation spectrum which provides the foundation for estimating the underlying probability distribution using the SDE algorithm.

3.6.1 Variation of Statistics

We generate a Monte-Carlo simulation where for each permutation of shape parameter, k, and MFR, ν, considered, 20 different realisations of a 100 second time recording was generated and the mean firing rate, $\hat{\nu}$, the PDF, $\hat{p}(t)$, and the variance, $\hat{\sigma}^2$, estimates were compared with the true values and relative errors were calculated. Note that the MFR estimate $\hat{\nu}$ was constructed using the iterative process described in Section 3.5.3. The PDF was then constructed using the methodology described in Section 3.5 using the MFR estimate $\hat{\nu}$. The variance estimate was then constructed using *both* $\hat{\nu}$ and $\hat{p(t)}$ with the formula: $\hat{\sigma}^2 = \int_0^{+\infty}(t - 1/\hat{\nu})^2\hat{p}(t)dt$. The choice of these three parameters, where the estimation relies on the estimates of the previous parameter(s) allows us to observe how the errors in the estimate propagate.

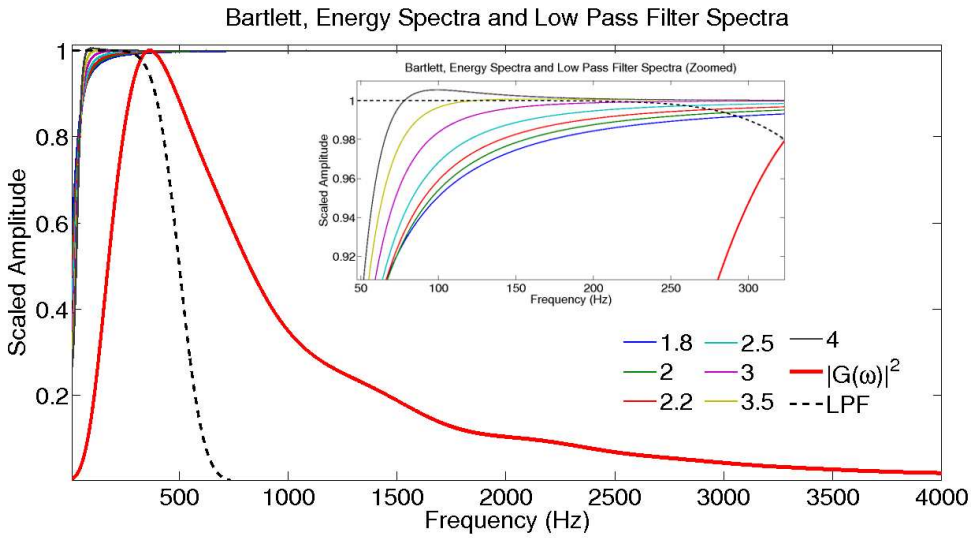


Figure 3.4: Bartlett Spectrum, Energy Spectrum $|G(\omega)|^2$ and low-pass filter @ $\nu = 50Hz$. *Inset*: Zoomed in version of Bartlett spectra of gamma distribution for different shape parameters k.

Figure 3.5 shows the error associated with the estimates of the mean firing rate, $\hat{\nu}$, PDF $\hat{p}(t)$ and variance $\hat{\sigma}^2$ respectively for each of the 20 trials for the different shape parameters and true MFR. Analysis of these parameter estimates shows the following:

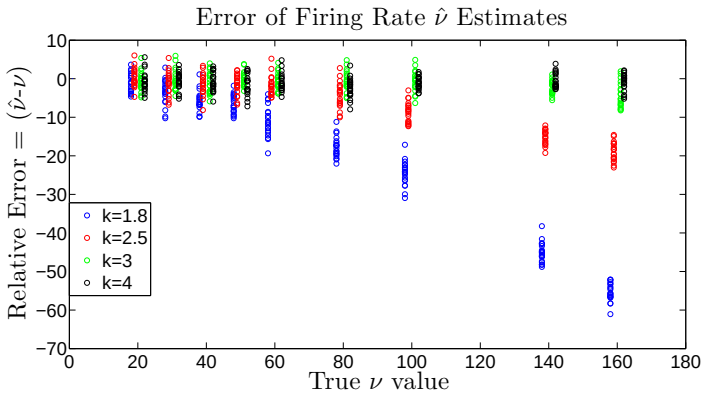
1. *MFR:* At smaller true MFR the estimates $\hat{\nu}$ for all firing patterns, from near-Poisson (k=1.8) to strongly correlated (k=4) are accurate. As the true MFR increases, the estimates closer to Poisson firing become less accurate, whereas estimates associated with the more correlated firing patterns stay accurate. It is interesting to note that for the near-Poisson case at high frequency the methodology appears to be biased towards under-estimating the MFR.

2. *PDF:* The accuracy of the PDF estimate is largely determined by the accuracy of the MFR estimate. The pattern of errors for both the NFR and PDF estimates follow the same trend: The near-Poisson ($k = 1.8$) errors are larger for higher MFR. In general, the higher MFR cases are associated with a larger biased error values. This is especially pronounced for the less correlated (smaller k values) firing statistics. This

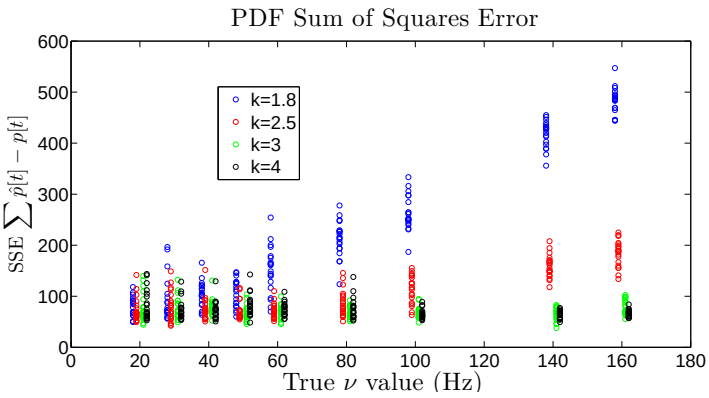
can be understood because if the average firing rate is higher, for a given time window there will be more pulses and a more accurate estimate of the correlation structure can be obtained.

3. *Variance:* The distribution of errors roughly follows the same pattern as that of the PDF and MFR errors. The more accurate the PDF estimate the smaller the relative error in the variance estimate. It is interesting to note that for near-Poisson (k=1.8) firing, even though the MFR estimate is biased to under-estimation, the variance estimates do not appear to be biased. This is especially interesting given that the variance estimate depends on MFR both explicitly, and through its effect on the construction of the PDF estimate.

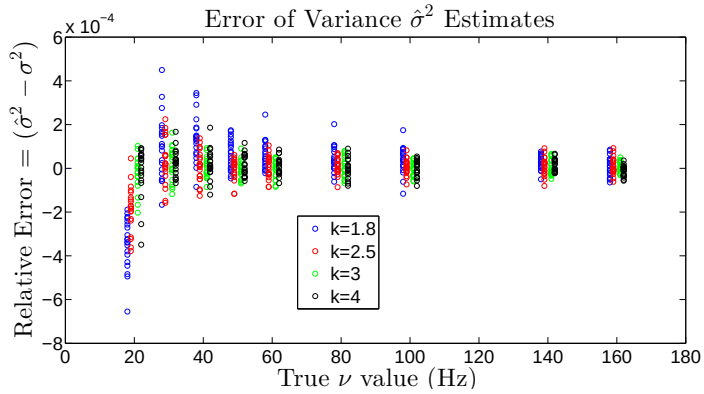
We have shown with simulation that the theory describing the solution to the inverse problem can indeed be used to estimate the firing time distribution given the superposition of an unknown number of identical renewal processes with similar filter shape. We have shown, as a general rule, that this method is more accurate for less 'Poisson-like' firing statistics. In the proceeding section we consider the application of this methodology to attempting to estimate the firing statistics of an ensemble of neurons which are equally close to an extra-cellular recording probe and therefore difficult to distinguish from one another.



(a) Relative error of NFR estimates $\hat{\nu}$



(b) Sum of Squared Error of $\hat{p(t)}$



(c) Relative error of variance estimates $\hat{\sigma}^2$

Figure 3.5: Error plots of MFR, $\hat{\nu}$, density, $\hat{p}(t)$ and variance $\hat{\sigma}^2$ estimates for different NFR and shape parameters. Each point represents the result from a single trial.

3.7 Application to Extra-Cellular Recordings

We consider an application of this Spectral Density Estimator to the vLFP estimation problem. When extra-cellular recording probes are used they often record the electrical activity of multiple nearby neurons [19]. This is especially true for the vLFPs recorded from MERs when they are placed in neuron-dense nuclei such as the STN targeted in Deep Brain Stimulation surgery for neuro-psychiatric disorders [14]. However in computational neuroscience we frequently wish to identify the spiking times associated with a single neuron, so that we may characterise that neuron by its firing properties [90],[89],[155]. These properties can either be the distribution that the firing times are drawn from or the statistical moments such as the mean firing rate, the coefficient of variation or the information entropy associated with this distribution (assuming they exist).

In order to isolate single neuron statistics when an ensemble of neural firing times are inadvertently measured spike sorting is used. Spike sorting consists of four main steps [165]: The first step is spike detection which estimates when a spike occurs in the time series. The second step is spike extraction where the shapes of the detected spikes are identified from the noisy time series. The third step is feature selection, where certain properties of the candidate spikes are selected and extracted. The fourth step is spike clustering, which attempts to partition the spikes into similar groups based on the similarity of the selected features. Once the set of spike times in a group have been identified the statistical moments $(\hat{\nu}, \hat{\sigma}, ...)$ can be estimated and the underlying distribution can be obtained either by using histograms or Kernel Density Estimation (KDE).

Recall that clustering in spike sorting is based on grouping some subset of the spikes together because of the similarity of some feature(s) such as the Euclidean distance of their shape to one another [57] or their principle components [183]. The problem is that when the spike shapes associated with different neurons are sufficiently similar, the clustering processes will not be able to differentiate the spikes. This will result in the clustering process placing spikes from different neurons into the same cluster. This situation could

occur when (excluding different spatial orientations) multiple neurons are equidistant from an extra-cellular recording probe. Conversely the Spectral Density Estimator solves the inverse problem assuming that the filter shapes, $g(t)$, are similar. The only difference which is accounted for in the super-position of renewal processes model is the variation in amplitude due to the scale factor a in (3.23). Therefore the more similar the action potential shapes are, the more accurate we expect the Spectral Density Estimates to be.

3.7.1 Simplified vLFP Model

We consider a simplified model of the vLFP associated with the MER receiving signals from multiple neurons as the superposition of multiple independent filtered renewal processes. This model is described by (3.23). It is important to note that this model introduces many simplifications to the vLFP recording. We discuss the implications of the two most important approximations below:

Renewal firing spike trains: This model assumes that the individual firing times are renewal processes which by construction assume that the time *between spikes* are i.i.d. There is a subtlety regarding this point: the property of i.i.d. times between spikes does not imply that the firing times are uncorrelated. This can be understood as follows. Inspection of (3.15) shows that the autocorrelation function of the spike times is solely a function of the renewal density function and the mean firing rate. With the exception of the Poisson process (a subset of renewal processes) different ISI probability distributions will generate non-constant renewal density functions and therefore will generate correlated firing patterns.

Independence of neurons: This model also assumes that each neuron will fire independently of their surrounding neurons. This is indeed a simplification to the true behaviour measured by the vLFP signal. Previous in-vitro studies measuring the spiking patterns of pairs of STN neurons in the Basal Ganglia tissue of rats has showed through cross correlation analysis that STN neurons often fire in synchronous bursts [184]. It is therefore unlikely that the STN neurons truly fire independently. Nonetheless in [171]

it was shown that a model of independent STN neurons firing with renewal statistics could successfully reproduce the structure of the measured vLFP power spectrum. Thus this assumption may assume an accurate power spectrum and therefore the estimated ISI PDF may also be accurate. In effect, [171] was interested in solving the 'forward problem' of choosing a specific ISI distribution such that the simulated power spectrum matched the measured one, whereas this chapter is concerned with solving the 'inverse problem' of estimating the ISI distribution given measurements of the vLFP power spectrum. The primary motivation for incorporating this independence assumption is to make the solution of the inverse problem (estimating the ISI PDF from the measured PSD) tractable. In future work we will consider whether this inverse approach can be extended to incorporate simple correlation models between the individual neurons.

In addition to the model given by (3.23) we include an amplitude scaling factor and low-pass filtering of the action potential shapes, g(t), to incorporate the effect of attenuation and dispersion of the electric fields as they travel through the extra-cellular fluid [185] to the probe tip. It is important to note that the amplitude attenuation is modeled by the a factor in (3.23), whereas the dispersion by low-pass filtering is not. Thus we expect that the presence of low-pass filtering will cause the Power Spectrum, $P(\omega)$, to deviate from the form expected from (3.23) and therefore will reduce the accuracy of the SDE.

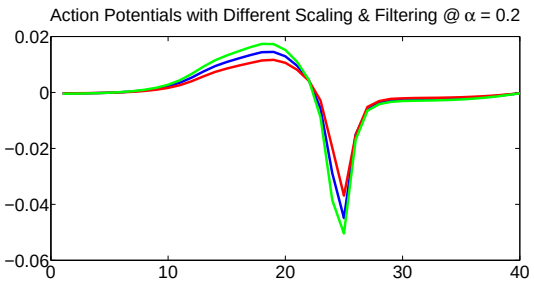
For the simulations we consider the situation of three neurons of variable distance from the MER. We wish to identify how both the spectral density estimator and classical spike sorting compare in identifying the PDF defining the neural statistics as the action potentials of the neurons become more similar. This similarity is a model for the neurons becoming more equidistant. We ignore the effects of white noise (due to signal acquisition errors) or neural noise (due to the contribution of more distant neurons) to minimise the number of variables when we compare these two methods. We note that if the neural noise is associated with the same renewal statistics as the three neurons of interest this will boost the signal to noise ratio for the SDE, but reduce accuracy for the classical spike sorting methodology.

The time series for this model is given by.

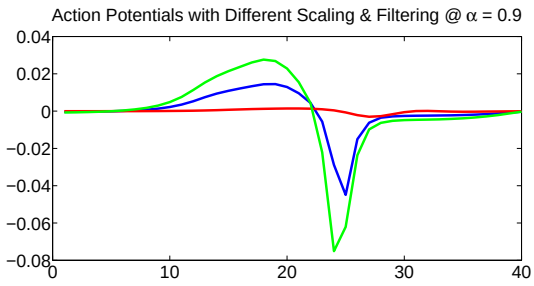
$$y(t) = [g(t) * h_1(t)] * \sum_{i=1}^{\infty} \delta(t - t_i) + (1 + \alpha) [g(t) * h_2(t)] * \sum_{j=1}^{\infty} \delta(t - t_j) \quad (3.33)$$

$$+ (1 - \alpha) [g(t) * h_3(t)] * \sum_{k=1}^{\infty} \delta(t - t_k) \quad (3.34)$$

Where $\{h(t)_1, h(t)_2, h(t)_3\}$, are Butterworth low-pass filters of order 2 with a transition frequencies given by: $\omega_c = \{0.12 f_s/2, (0.12 + \alpha/10) f_s/2, (0.12 - \alpha/10) f_s/2\}$ respectively and f_s is the sampling frequency. These low-pass filters act as the simple model for dispersion due to the extra-cellular fluid. In this simple model the α scaling factor controls the similarity of the three action potentials. In the case of the three neurons being similarly orientated and equidistant from the extra-cellular probe the action potential shapes and amplitudes would be identical, which would be modeled by $\alpha = 0$. As the relative distance to the probe between the neuron varies the action potential shapes vary and α will increase. When $\alpha = 1$, the third neuron vanishes completely and the shape difference (in terms of dispersion) of the first and second neurons action potentials is greatest. Figure 3.6 illustrates this effect, showing the three different action potentials for different α values.



(a) $\alpha = 0.2$.



(b) $\alpha = 0.9$

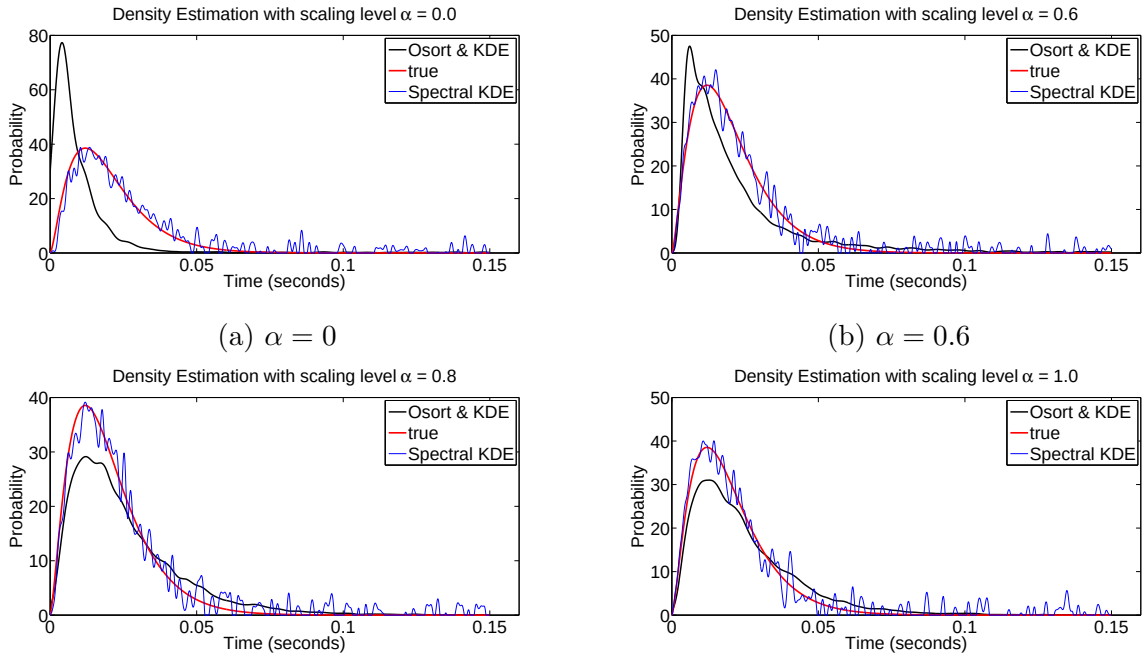
Figure 3.6: Action potentials with different levels of dispersion and attenuation determined by the α scaling factor. The green and red action potentials have the least and greatest dispersion respectively. Notice that the greater the value of α (i.e. the greater the dispersion and attenuation) the more distinct the waveforms.

We vary this scaling factor α over the range $[0, 1]$ and compare the PDF estimates obtained by the spectral method to classical spike sorting. We run all simulations with gamma firing ($k = 2.5$) statistics with mean firing rate, ν, of 50 Hz and generate the subsequent time series at sampling rate of $f_s = 10kHz$. Similar to the previous section the action potential shape is generated by solving the Hodgkin-Huxley equation with appropriate parameters for the neuron in the Sub-thalamic nucleus (STN) [182].

We compare the accuracy of PDF reconstructions generated by our methodology with a classical spike sorting approach used by the highly popular program Osort [57]. Osort performs the spike detection step using a multi-scale continuous wavelet denoising algorithm. For a full description of this process see [94]. The spike clustering in Osort is performed using a template matching scheme which groups the (up sampled) waveforms identified at the estimated spike times into different clusters based on the Euclidean distance from each other. For a full description of this clustering process see [57]. Once we have identified the spiking times associated with each cluster we generate estimates of the ISI times for each cluster by subtracting the identified firing time from the previous identified firing time: $\Delta T_i = t_i - t_{i-1}$. We then merge the ISI times for each cluster into a single grand cluster and generate an estimate of the PDF using an adaptive Kernel Density Estimator based on the diffusion process [186].

3.7.2 Results

Inspection of Figures 3.7 and 3.8 show that as the α scaling value gets closer to zero (modeling the neurons becoming more equidistant from the recording probe) the accuracy of the classical spike sorting estimates decrease. As α decreases, the clustering (incorrectly) places more and more neurons into the same cluster. This results in estimates of PDF with much smaller mean time, μ, between events with a shape that looks similar to the exponential distribution. This was identified in [158] where the spike times associated with a superposition of renewal processes was observed to have shorter mean ISI times and more exponential



(a) $\alpha = 0$ (b) $\alpha = 0.6$

(c) $\alpha = 0.8$ (d) $\alpha = 1.0$

Figure 3.7: Sample realisations of the Spectral Density Estimator and Osort PDF estimates, $\hat{p}(t)$, for different α scaling values.

looking distribution estimates as more renewal processes contributed to the super-position. As the α value gets larger (modeling the AP shapes becoming more distinct) the classical spike sorting result gets more accurate, because the template matching correctly clusters neurons. It is interesting to note that despite the general trend towards better estimates for the classical spike sorting, at $\alpha = 0.8$ there is a local increase in the SSE. The cause of this is not obviously clear. It is likely due to the re-emergence of an Osort detected similarity between two separate action potentials, which disappears for even larger α values in the region where the amplitude of the third neuron is approximately zero. Note that the spectral density estimator is largely insensitive to the α value.

3.7.3 Analysis

The results shown in Figure 3.8 and discussed in the previous section confirm the intuition discussed in Section 3.7 that the more similar the action potential shapes the worse the

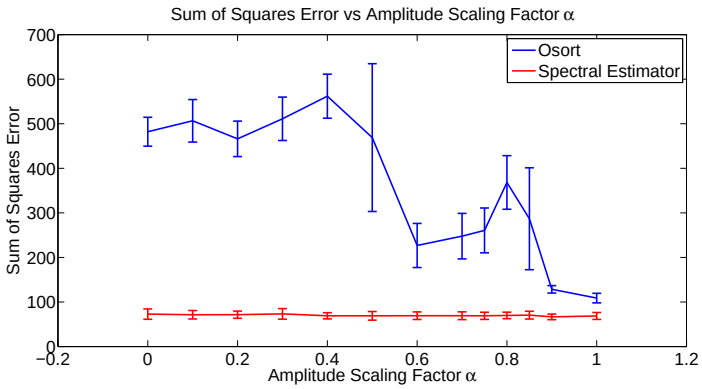


Figure 3.8: Mean and SEM of the SSE averaged over 25 trials for $k = 2.5$ gamma firing neurons ($\nu = 50$Hz) for various α scaling levels. Notice the general trend that as the α scaling factor increases the classical spike sorting is more successful, whereas the spectral KDE is largely independent of the relative amplitudes.

Osort and better the Spectral Density Estimator function respectively. This suggests that the Spectral Density Estimator may have utility over classical spike sorting approaches when the renewal firing statistics of individual neurons are required but unable to be separated from other nearby renewal firing neurons which are also being recorded.

It is very interesting that in Figure 3.8 the accuracy of the spectral density estimator is largely unaffected by the dispersion (low-pass filtering) effect. Recall that the dispersion changes the filter function, g(t), shape. Also recall that the Spectral Density Estimator solves the inverse problem associated with the superposition of filtered renewal processes with (up to a scaling factor) identical filter functions. Intuitively we would expect this dispersion effect to violate the assumptions of the model, alter the power spectrum, $P(\omega)$, and in turn alter the estimate of the correlation spectrum $\Phi(\omega)$ (described in Section 3.5.1) and therefore the estimate of distribution: $p(t)$. This does not occur because for almost all transition frequencies, ω_c, considered the support of the correlation spectrum, $\Phi(\omega)$, overlaps the constant passband of the Butterworth low-pass Filter. For example, the smallest transition frequency considered in these simulations was (at $\alpha = 1$ for the third neuron in (3.34)) $\omega_c = 100$Hz. Inspection of Figure 3.4 shows that for a wide range of gamma shape parameters with mean firing rate, ν, of 50Hz (including $k = 2.5$) the correlation spectrum

is close to zero by approximately $150Hz$. Thus filtering distorts the measured spectra in a region that contains minimal information about the correlation spectrum and therefore the density estimate is largely unaffected.

The results of these simulations are also interesting because they suggest that given the power spectrum of highly filtered and attenuated neural time series (firing with renewal statistics) it is possible to reconstruct the distribution that these firing statistics are drawn from. A situation where this occurs is the analysis of *Local Field Potentials* (LFPs). The contribution to LFPs is still debated [19], but in addition to synaptic/dendritic contributions it also must include the effect of multiple far field highly dispersed neurons. This is especially plausible when low impedance, high capacitance MERs (such as those used in Deep Brain Stimulation surgery) are employed. Results from [19] have suggested that in some instances, with appropriate machine learning techniques, features of LFPs can be used to predict with relatively high accuracy when a particular neuron will generate an action potential. It is particularly interesting that this methodology was most effective when the feature considered was the power of the LFPs in the gamma band ($40-90$Hz), which, as shown in Figure 3.4, is precisely the region where the correlation spectrum for a physiologically firing neuron will be non-zero. We also note that there is rarely a clear demarcation between single unit recordings and LFPs for extra-cellular recordings [20]. This is especially true for the vLFPs from MERs which, even when placed near a dominant neuron will still pick up the electrical contribution from (heavily) attenuated and dispersed nearby neurons.

As discussed in the previous section we have generated a highly simplified model of neurons interacting with an extra-cellular MER. Multiple effects such as the spatial orientation of the neurons (which will further alter the action potential shape) and firing rate dependent spike shapes have not been included. Additional phenomena such as sub-threshold oscillations [187] and correlation between the neurons [188] are not modeled. In addition the filtering effect of the extra-cellular medium is more complex than a second order Butterworth low-pass filter [185]. Nonetheless this model incorporates many of the core features and can be considered a proof of concept that there are circumstances (of multiple renewal

firing neurons) when estimates of the power spectrum can yield accurate reconstructions of the underlying density function which outperform classical spike sorting approaches.

3.8 Numerical Simulation Conclusion

We have developed a methodology to estimate the PDF associated with the superposition of filtered renewal processes given an estimate of the power spectrum. We have shown through extensive Monte-Carlo simulation that this method generates more accurate estimates when the firing patterns are more strongly correlated. We have suggested that this spectral approach to density estimation may have applications in analysing vLFPs from MERs, outperforming classical spike sorting techniques when the shapes of the action potentials associated with different neurons being recorded are sufficiently similar. These results suggest that this method could be most appropriate when attempting to estimate the firing statistics associated with multiple neurons when they cannot be separated using classical spike sorting techniques.

3.9 Conclusion & Contributions

In this chapter we have attempted to develop a model-based approach to analyse vLFPs from MERs in neuron dense structures such as the STN. We have developed a transformation operator, termed the Spectral Density Estimator which maps the measured power spectrum to the underlying common probability distribution driving the firing statistics of the neural ensemble. The methodology implemented in this chapter has demonstrated the remarkable result that given the measured power spectrum of a neural ensemble, conditional on a series of assumptions being satisfied, the common probability density function driving the constituent spike trains can be identified. This methodology continues along the common theme introduced in this thesis, where the vLFPs are characterised by solution of an inverse problem. In this case there were two serial inverse problems: decoupling the

Bartlett spectrum from the action potential energy spectrum associated and the deconvolution operation on the renewal density equation to identify the probability density function.

This spectral approach has allowed us to solve the pathological mathematical problem of identifying the firing statistics of an individual renewal process when it is embedded with identical processes which are impossible to distinguish using conventional techniques. We have argued that this pathological situation can arise when there are multiple neurons of the same type, some with similar orientation, equidistant from a single channel recording probe. This situation is highly likely to occur when recording probes are placed in neuron-dense structures like the Sub Thalamic Nucleus. Therefore, the methodology of the Spectral Density Estimator introduced in this chapter may be useful to resolve the spiking statistics if this situation is expected and classical spike sorting approaches fail to yield accurate or biologically plausible solutions.

We have also (in Appendix A.7) developed a model of a neuron encoding information in the space between firing times using a Digital Pulse Interval Modulation (DPIM) scheme popular in telecommunications engineering. We have shown (unlike classical DPIM strategies used in communication systems) that if the firing times are allowed to occur any time in the continuous timeline (as we would expect for a biological process) that the power spectrum of this encoder will be exactly the same as that of an idealised renewal process.

Similar to the NMP, an additional advantage of this Spectral Density Estimator approach is that the estimates are generated using the entire signal, rather than small subset of spikes in the case of spike sorting. Also similar to the NMP approach this methodology is parameter free (with the exception of parameters introduced in the identification of the spectrum step to boost the accuracy of the estimate).

The limitation of this approach is that in order for this Spectral Density Estimator

approach to be valid there is a series of stringent signal requirements.

- The signal must not only be wide-sense (covariance) stationary for the power spectrum to be defined, but it must be strictly stationary for the probability density being estimated to not vary over the duration of the recordings.
- The firing patterns of all the individual neurons must satisfy renewal statistics drawn from a common distribution.
- The neural ensemble is modelled as a statistically independent set of renewal processes.

In the next chapter we introduce the final approach to characterise the MER signal based on Basis Pursuit Denoising and clustering using Diffusion Mapping and the Mean Shift algorithm. Similar to chapters 2 & 3 this will involve solving an inverse problem, but unlike the previous approaches, where the inversion is performed on the time averaged ensembles of the autocorrelation function or the power spectrum, the inversion is associated with the actual time series representing stochastic realisations of the vLFP. This approach will drastically reduce the a priori assumptions, principally no assumption of stationarity of the firing times or independence of the constituent neurons. The disadvantage of this approach is that fundamentally it is a variant of a spike sorting algorithm and so only uses information from a small subset (the identified spikes) of the signal. In addition, it requires multiple parameters which must be determined apriori.

4

Spike-Only vLFP Analysis

"It is pointless to do with more what can be done with fewer."

– William of Ockham, in *Summa Totius Logicae*

4.1 Chapter Summary

In the previous chapter we showed that the vLFPs could be analysed in a model-based framework by assuming the contributing neurons could be modelled as an ensemble of independent renewal processes. Using this model, conditional on multiple assumptions being satisfied, the common probability distribution driving the spike times could be estimated with a series of inverse operations applied to the measured power spectrum. The problem with this approach was that the assumptions were highly restrictive. In addition the probability distribution

could not be determined if the neurons were firing with maximally random Poisson statistics.

In this chapter we introduce a method to identify the precise firing times of a subset of the individual neurons in the ensemble. This method places no assumptions on the individual firing patterns of, or the interactions between the constituent neurons and can deal with Poisson firing statistics. In this approach we strip away all information from the vLFP except for the firing times and action potential shapes of the nearest neurons interacting with the MER. The major assumption, also invoked in the previous chapter, is that these spike shapes and firing times for the individual neurons are considered decoupled variables.

The method introduced in this chapter is a spike sorting algorithm constructed using Basis Pursuit Denoising (BPDN), which is a subset of the sparse ℓ_1 minimisations schemes used widely in signals processing, statistics and compressed sensing. Similar to the methods introduced in the previous chapters the BPDN scheme requires the solution of an inverse problem, although unlike the previous chapters this inversion is solved in the time domain.

We show through extensive Monte Carlo simulation that the sensitivity, specificity and χ^2 estimates of the spike sorting using this BPDN strategy are superior to state of the art spike sorting techniques. We then apply this method to vLFPs from MER in the STN of human patients with Parkinon's Disease undergoing DBS. The firing patterns we identify ($\sim 20 - 56$ Hz, Poisson statistics) are consistent with previous experimental analysis.

4.2 Chapter Overview

In this chapter we develop a spike sorting algorithm to characterise the firing patterns observed in a vLFP from the contribution of the neurons (with action potentials which are resolvable) which are closest to the (single channel) MER. Unlike the previous chapters which were based primarily on frequency domain operations, this algorithm is based on operations in the time domain.

With this approach we only use the information from the time and shapes associated with spikes of the neurons nearest to the MER. This philosophy is markedly different from the approach considered previously in chapters 2 & 3, where the time series was characterised by operations on the entire vLFP. The principal advantage of this spike sorting approach is the minimal a priori assumptions required to use this method: the firing times of the constituent neurons may be highly nonstationary, with complex correlations between the constituent neurons which may have arbitrarily complex individual firing patterns. The only requirement of this model is that the action potential shapes associated with individual neurons does not vary with time. This is referred to as spike shape stationarity.

The outline of this chapter is as follows:

1. Provide an introduction to spike sorting and the four canonical sub-problems of spike detection, spike identification, feature selection and clustering. We then provide an overview of the different methodologies introduced to solve these sub-problems.

2. Introduce the sparse least squares algorithm of Basis Pursuit De-Noising (BPDN) and explain why it provides a natural framework for spike detection and clustering problems. We then provide an overview & analysis of three popular approaches to solving BPDN: *homotopy*, the *InCrowd* algorithm with *Truncated Newton Interior Point* (TNIP) and the *Dual Augmented Lagrange Multiplier* (DALM) method.

3. Introduce the solution to the problem of identifying the spike shapes (the dictionaries) required to use BPDN for spike sorting. We identify these spike shapes using a three

step method of *Continuous Wavelet Transforms* to develop an initial estimate of the firing times, *Diffusion Mapping* to reduce the dimensionality of the identified waveforms and the *Mean Shift* algorithm to automatically identify the average of the individual neuron's spike shapes.

4. Show the results of this spike sorting algorithm when applied to synthetic vLFP time series. The accuracy of the different spike sorting algorithms are displayed with Receiver Operating Characteristic (ROC) scatter plots which provides a graphical representation of how many true positives (identifying the correctly clustered spike) and false positives (either the detection of a spike that was not present or incorrectly clustering a spike that was present). The following numerical simulations are generated and analysed:

   - The accuracy of the three different BPDN algorithms: homotopy, InCrowd with TNIP and DALM when the true action potential shapes are known a priori.
   - The accuracy of the BPDN clustering algorithm compared with the 'preliminary dictionary obtaining' (wavelet + diffusion mapping + mean shift clustering) method. We show that while the BPDN algorithm requires accurate estimates from this 'preliminary dictionary obtaining' method, the BPDN approach consistently outperforms this 'preliminary dictionary obtaining' method for all noise levels considered.
   - The accuracy of the BPDN clustering algorithm compared against the state of the art spike sorting software *Wav-Clus*. We show that at high signal to noise levels both of these algorithms give comparable results, but at the more challenging lower signal to noise levels the BPDN spike sorting algorithm outperforms *wav-clus*.

5. Apply the BPDN spike sorting algorithm to vLFPs obtained from MERs inserted in human STN. We show that these neurons fire, on average, with the physiologically plausible pattern of $\sim 20 - 56$ Hz with Poisson firing statistics.

Contribution

The contribution of this chapter is the introduction of a non-parametric spike sorting algorithm based on Basis Pursuit De-Noising which is able to outperform the state of the art spike sorting software *Wav-Clus* at low signal to noise ratios. This algorithm is then applied to *in-vivo* recordings of human sub thalamic nuclei, showing that at rest these neurons fire with Poisson statistics in the physiological range of $20 - 56$ Hz. The results on both benchmarked synthetic and real extracellular data suggest that this BPDN spike sorting algorithm can be successfully used for analysis and spike sorting of signals obtained from *in-vivo* extracellular micro electrode recorders in low signal to noise environments.

4.3 Introduction

One of the core objectives of neuroscience is *unravelling the neural code*. That is, given a neuron which generates a series of action potentials, understanding how this stream of spikes transmit information? It is near universally agreed that the information is encoded in the timing of spikes [82], [83], [84], [85], [86], [87], [88], [89]. But *How* this information is encoded is a contentious question: in some cases it is believed the information encoded by the lower order statistical moments (e.g. the firing rate [88] or coefficient of variation [90]), in other cases it is the precise timing between the spikes [89]. The answer to this question is further obsfucated by the rich diversity of observed firing patterns. For example dopaminergic neurons of the Basal Ganglia have been identified with firing patterns ranging from near periodic, to maximally (Poisson) random to chaotic [154]. Irrespective of the neural coding mechanism, in order to characterise a firing pattern, techniques must be developed to identify these spikes contributing to the pattern [93].

Once the firing times of the individual neurons are known they can be compared to the timing of the presentation or some property of the external stimulus. Alternatively, assuming the firing patterns on a short enough time scale are stationary, the neurons can be described by the statistical moments (firing rate, coefficient of variation or information entropy), if they exist, of the spacing between their spikes. Recall that this spacing between the spikes is referred to as the Inter Spike Interval (ISI). In intracellular experiments, or with well controlled *in-vitro* probes, a skilled physiologist can record the firing times of individual neurons. For *in-vivo* experiments recorded with extracellular probes, identifying individual neuron firing times is is manifestly more difficult [189],[96],[57],[165]. This is because the signal will be composed of the contribution referred to as the Multiple Unit spiking Activity (MUA) [19] of the multiple local neurons [190] hidden within the so called 'neural noise' consisting of the heavily filtered [24] spiking activity of far field neurons and the multitude of low frequency (< 300 Hz) neural events which contribute (synaptic currents [20], spike after potentials [19], voltage-dependent membrane oscillations [22]) to the Local Field Potential (LFP) [19]. This field of research, of not only identifying the MUA spikes hidden within the

noise, but clustering them together into groups associated with individual neurons is referred to as *Spike sorting*.

Spike sorting effectively involves solving two coupled problems: firstly identifying all the firing times (action potentials) in a signal and then secondly grouping those identified spikes into appropriate clusters. It is important to note that there is an implicit assumption of most spike sorting programs [165], [191], [192], [57] that each neuron generates a stereotyped action potential which is unique for each of the measureable neurons. The exact shape and amplitude of these spikes will depend on the distance and orientation to the recording probe and the morphology of the neuron [185]. The issues associated with this assumption, particularly in the case of burst firing neurons, will be discussed further in section 5.3.

4.3.1 Overview of Spike Sorting

The two sub problems of spike timing: identifying firing times and then clustering them are usually solved in 4 steps [165]. We will outline these steps and discuss the most commonly used approaches utelized by automated spike sorting software.

The first step is **spike detection** which involves detecting all the candidate spikes, irrespective of which cluster they belong to. This spike detection step can be solved with a multitude of techniques. The simplest approach is based on amplitude thresholding [165], where a datapoint with amplitude greater than some value or factor of the standard deviation in the measured times series is considered a spike. The performance of this approach, while simple to implement, degrades rapidly in situations of low signal to noise [94] because the high variance noise events greater than threshold are considered spikes. Similar, but more sophisticated approaches based on identifying localised increases in energy using the multi-resolution Teager Energy Operator (mTEO) [193] have also been considered. Another widely used approach is to use multi-scale wavelet decompositions [94] with mother wavelets carefully chosen to match the expected shape of action potentials (for example the *bior1.5*

family [194]). Briefly this multi-scale resolution allows the signal to be decomposed independently into different frequency bands so that different transients (spikes) at different time scales can be identified robustly in the presence of noise. We will discuss this approach in more detail in section 4.5.1.

The second step, once the candidate firing times have been identified is **spike extraction**. This involves collecting a windowed time-series centred around the estimated firing times. This process is consistent among different spike sorting strategies, with the only variable being how much of the time series is recorded before and after the estimated spike. This decision is usually selected based on an assumed time constant of the expected action potentials associated with the neurons being examined.

The third step, once all the candidate firing times and their associated spike shapes have been detected is **feature extraction**. The purpose of this step is to condense the high-dimensional action potential shapes from the traces of the time series into low-dimensional structures such that similar shaped action potentials (presumed to come from the same neuron) will reside in similar locations in this subspace, but be separated from action potentials of different shapes belonging to a different group. The selection of the dimensionality of the subspace is a non-trivial problem. The reason for this is that differences between spikes are most clearly seen in high dimensions (in the extreme limit this would be the original action potential shapes). However the success of algorithms used to cluster action potentials in the reduced subspace markedly degrades in higher dimensions. This is referred to as the "curse of dimensionality" [195].

Multiple different features of spike shapes have been used to construct this reduced dimension subspace. One of the earliest approaches was to perform Principal Component Analysis (PCA) on the identified spike shapes [183]. In [165] the ten Haar wavelet coefficients with the greatest degree of multi-modality, as indicated by the Kolmogorov-Smirinov test, (as an indication of being different for different neurons) was used. More recently [196] argued that diffusion maps developed in [197] provides a superior approach for clustering

than the Haar wavelets with the greatest multi-modality. This feature extraction problem is an open area of research that we will discuss in more detail in section 4.5.2.

Once the shapes of candidate spikes have been mapped to a reduced dimension subspace they must be clustered into appropriate groups. This fourth step is called **spike clustering**. The simplest approach to clustering is for the process to be user driven, but even in the hands of a skilled neurophysiologist this approach is time consuming and is subject to bias [165] with wide variance between users [198]. This has motived the development of automated clustering algorithms. Automated clustering is a very large field and the development of new clustering algorithms is an area of active research. In this section we will only discuss a subset of the available clustering methods, focussing on density based and graph based approaches. For an excellent overview and introduction of automated clustering see [199, Chapter 5].

One of the simplest and robust approaches is *K-means* clustering [200], which groups elements of a dataset by the shortest Euclidean distance to a number of centroids. One of the problems with the K-means approach is it requires a priori knowledge of the number of centroids (in the spike sorting case this is the number of neurons contributing to the recording). With extra-cellular recordings the number of contributing neurons is seldom known. Extensions to K-means, which do not require a priori knowledge of the number of clusters, have been developed including using the gap statistic [201], silhouette statistics [202] and information criterion approaches such as the Bayesian Information Criterion (BIC) (used in the X-means algorithm) [203] . In the context of spike sorting these extensions to K-means using the gap statistic and silhoutte statistics were applied to features based on both the most multi-modal Haar wavelet coefficients and diffusion maps in [196].

Another successful clustering approach is the Super-Paramagnetic Culstering (SPC) algorithm which was applied to spike sorting in [165]. This approach is based on simulated interactions between data points in the reduced subspace and its K nearest neighbours. The name SPC is based on simulations of a statistical mechanical model of spin states in a

crystalline lattice called the Pott's model. For a detailed explanation see [165], but briefly, points in the reduced subspace are randomly assigned to one of a fixed number of states. The probability that other points will also be assigned to this state is an exponentially decreasing function of the Euclidean distance (in the reduced subspace) of these points to the original point. A free parameter, referred to as the *temperature* (due to the algorithms statistical physics origins) scales the Euclidean distance function. For sufficiently high temperatures this coupling function drops off rapidly and each data point in the subspace will be considered as an independent cluster, at low temperatures the strength of the coupling factor is large and all data points belong to the same cluster. In between these two extremes there will be a distribution of clusters, similar to the paramagnetic state of a spin glass, where a small subset of clusters will form. The success of this method is largely driven by judicious choice of the temperature parameter.

A simpler alternative to clustering on the subspace is to use the mean shift algorithm. The mean shift algorithm is a non-parametric method for identifying peaks (modes) in the Kernel Density Estimate (KDE) of a set of datapoints without actually constructing the KDE. If the data points in the reduced subspace are assumed to be realisations of a multi-dimensional random variable the mean shift algorithm can be applied to identify which peak each datapoint in the reduced subspace belongs to. Applying this method to each datapoint partitions them into unique clusters associated with a mode. Similar to the SPC algorithm the success of this method relies on the choice of the bandwidth parameter associated with the KDE. We will provide further analysis of this method in Section 4.5.3. This mean shift algorithm method was applied to both the most multi-modal Haar wavelet coefficients and diffusion maps of a priori known spike shapes in [196].

An alternative, much simpler clustering strategy based on the Euclidean distance between candidate spike shapes themselves, rather than the distance in some transformed subspace was considered in [57]. In this approach spike shapes with a Euclidean distance smaller than some pre-defined threshold are considered to belong to the same cluster. This approach has the advantage of being computationally efficient enough to be performed online, and does

not require *apriori* information about the number of clusters, but has also been found to be less accurate than other clustering methods [204].

It is important to stress that we have only scratched the surface of clustering algorithms. There are other powerful density based clustering algorithms such as DBSCAN [205] (used in a spike clustering approach in [206]), it extensions OPTICS [207], DENCLUE [208] (similar to the mean shift approach we discuss in Section 4.5.3) and DENCLUE 2.0 (which incorporates sparse sampling and can be shown to be a subset of Expectation-Maximisation [209]). There are additional graph based approaches such as clustering on the Normalised Graph Laplacian (applied to spike sorting in [210]) and more general spectral clustering [211],[212] strategies. There are also Bayesian, model-based, clustering strategies based on Gaussian Mixture Models (GMM) [213], mixtures of Student t-distributions [214], infinite Gaussian mixture models [215] or greedy 'binary pursuit strategies' [191] which maximise the a posteriori (MAP) distributions of spike times and shapes. There are also Dynamic Hidden Markov Model approaches to spike sorting [216],[217] which attempt to incorporate as much 'biological' information (such as the refractory times of neurons and the variation of spike amplitude with firing rate in bursting regimes). This is by no means an exhaustive list and indicates the breadth of the field of automated clustering and its application to the spike sorting problem.

These four steps (spike detection, spike estimation, feature selection and clustering) are largely modular with different algorithms for the four steps being effectively interchangeable. For example the spike sorting package *Osort* uses the multiscale wavelet decomposition to identify the firing times and then clusters based on the Euclidean distance of the spike shapes to each other. The *wav-clus* algorithm, which we use to bench mark our spike sorting method against in section 4.6.5 uses amplitude thresholding to identify spike times, Haar wavelet coefficients for feature selection and SPC for feature clustering. The *SpikeOMatic* algorithm uses amplitude thresholding or template matching to identify spikes and then uses the Gaussian Model of Mixtures or the Dynamic Hidden Markov Model to cluster the data. As discussed previously the majority of algorithms solve these sub problems sequentially. One

notable exception is [192] which uses an iterated co-ordinate descent approach of holding the firing times constant then solving the sub-problem to identify the optimal spike shapes, and then holding these newly determined spike shapes constant while solving the sub-problem to identify the optimal firing times. This process is then repeated in an iterative fashion until a stopping condition is satisfied.

4.3.2 Model Development

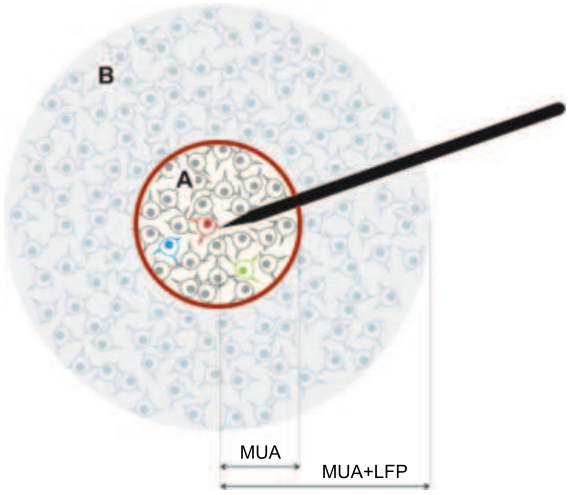


Figure 4.1: Simplified Model of the MER interacting with the electric field of nearby neurons. We consider a boundary (the red circle) which separates our neural sources into 2 components: the MUA $\mathcal{S}(t)$ (inside the red circle), and the neural noise $\eta(t)$ consisting of far field neurons and the LFP contribution (outside the red circle). The division is arbitrary, and it is not immediately clear how this demarcation should be made. Image adapted from [96].

In this section we develop a linear response model of the vLFP measured by a single channel MER embedded in neural tissue. This basic model formulation were also developed in [191] & [192]. The probe records a time dependent voltage, y(t), which consists of two components. The first component is the MUA, $\mathcal{S}(t)$, associated with the spiking of the nearby resolvable neurons (the neurons within the red circle of Figure 4.1) and the second component is the neural noise term, $\eta(t)$, associated with the spiking activity of more distant, non-resolvable neurons and the slower non spiking contributions to the signal.

The firing times of each of the neurons in the MUA set, $\mathcal{S}(t)$, can be represented without loss of generality as an arbitrary point process. That is, the spike times are isolated stochastic elements embedded in the continuous timeline. We can express these firing times as a function of time as a summation of Dirac delta distributions $\sum_{k=1}^{N} \delta(t - t_k)$, where $\{t_k\}$ form a set of random variables. Note that there is no restrictions on this random variable

beyond non-negativity. The firing times need not be i.i.d or even drawn from a stationary distribution. Recall that this is distinct from the modelling in chapter 3 where a renewal model for the spacing between the firing times was considered.

As discussed in section 4.3 the action potential shape for each neuron is considered stereotyped with a specific structure that is a function of the neurons morphology and the distance and orientation from the recording probe. The shapes associated with the MUA set of neurons is *not* considered known a priori and must be estimated from the recorded time series. The spike trains, $y_i(t)$, of the individual neurons which belong to the MUA, may be written as the convolution of the spike shape, $g_i(t)$, and the Dirac delta distribution of firing times: $y_i(t) = g_i(t) * \sum_{i=k}^{N} \delta(t - t_k)$. It is important to note that the assumption that the action potential of a single neuron is stereotyped is equivalent to assuming a linear framework for the spike trains such that the spike shape and firing time variables are decoupled from each other. This linearity heuristic breaks down in certain scenarios such as bursting [163], where the spike shape for a single neuron varies along the bursting event.

The signal generated by the set of neurons comprising the MUA is given by their superposition:

$$\mathcal{S}(t) = \sum_{i=0}^{N} \underbrace{g_i(t) * \left(\sum_{k=1}^{\infty} \delta(t - t_{k,i}) \right)}_{i^{th}\text{Neuron}} \tag{4.1}$$

The neural noise term, $\eta(t)$, consists of every contribution to the recording probe which is not the spiking activity of the resolvable neurons which contribute to the MUA. As discussed in section 4.3 this term will include, but not be limited to, the heavily filtered spiking activity of far field neurons and the multitude of low frequency (< 300 Hz) neural events which contribute (synaptic currents, spike after potentials, voltage-dependent membrane oscillations) to the Local Field Potential (LFP) [19]. Note that the filtering of the far field neurons occurs as the associated electric fields travel through the extra-cellular medium to the recording probe [24],[185]. It is important to note that this neural-noise term $\eta(t)$ will be coloured and correlated with both itself and the neural source term $\mathcal{S}(t)$.

The model for the single channel extra-cellular MER times series is given by the superposition of the MUA term, $\mathcal{S}(t)$ (4.1), and the neural noise term, $\eta(t)$:

$$y(t) = \sum_{i=0}^{N} g_i(t) * \left(\sum_{k=1}^{\infty} \delta(t - t_{k,i}) \right) + \eta(t) \tag{4.2}$$

It is important to appreciate that in order to model the vLFP we have (similar to others [57],[192] [191]) introduced the heuristic that the spike train dynamics and their interactions are linear. More specifically we have assumed linearity for the following:

- **spike shape stationarity:** As discussed previously the spike shape (for a single neuron) does not vary and is not influenced by firing times.
- **MUA linearity:** the behaviour of the MUA is given by the linear superposition of the individual spike trains of the constituent neurons.
- **signal + noise linearity:** the overall time series is given by the linear summation MUA signal, $\mathcal{S}(t)$, and the neural noise term $\eta(t)$. Note that there is no restriction on the correlation between the signal and the noise.

It is important to identify that these assumptions are not unique to our approach. Indeed these assumptions are implicit in other state of the art spike sorters [57][1],[165],[191],[192]. The approach we take (similar to [192]) given this linear framework is to explicitly develop estimates of the firing times (that are subsequently clustered) which are theoretically guaranteed, under certain conditions dictated by the Restricted Isometry Property (see section 4.4) to be optimally sparse.

A notable exception to the spike shape stationarity assumption is the Dynamic Hidden Markov Model approach used in the SpikeOMatic [162] spike sorting algorithm. In this methodology the firing times of the neurons are modelled as renewal processes and the amplitude of a spike is dependent on the time elapsed since the last spike. This approach has

[1] In the Osort algorithm the threshold for belonging to the same cluster is often set sufficiently low such that action potentials which are modified during burst sequences are still clustered into the same group. The trade-off for this is less accurate spike separation in the general case.

the advantage of easily incorporating the burst phenomenon, where the amplitude of a spike decreases if a sequence of spikes occur in a rapid succession. The disadvantage of this model is that it requires a priori information about the number of neurons contributing to the signal.

If this model (4.2) is discretised as is necessary for digital acquisition and subsequent numerical operations, the convolution of the i^{th} action potential shape $g_i(t)$ with the variable Dirac comb representing the firing times $\{t_k\}_i$ can be written in matrix-vector notation as:

$$\begin{aligned} g_i(t) * \left(\sum_{k=1}^{\infty} \delta(t - t_{k,i}) \right) &= \begin{bmatrix} g_0 & 0 & 0 & \cdots & 0 \\ g_1 & g_0 & 0 & \cdots & 0 \\ g_2 & g_1 & g_0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \\ g_M & g_{M-1} & g_{M-2} & & g_0 \end{bmatrix} \begin{bmatrix} x_0^i \\ x_1^i \\ x_2^i \\ \vdots \\ x_{M-1}^i \end{bmatrix} \\ &= \mathbf{A}_i \mathbf{x}_i \end{aligned} \tag{4.3}$$

Note that $\mathbf{A}_i \in \mathbb{R}^{M \times M}$ is a convolution matrix and $\mathbf{x}_i \in \mathbb{R}^{M \times 1}$ is a vector associated with the i^{th} spike train. The vector $\mathbf{x_i}$ is the discretised form of the Dirac delta comb associated with the firing times, with the non-zero elements representing the time and amplitude of the spike $g_i(t)$. For example the time series for the i^{th} spike train with sampling rate of F_s Hz with non-zero values in the 412^{th} and 756^{th} elements of magnitude 1 and 1.5 respectively indicates that spikes with amplitudes of magnitude (1) & (1.5) occur at $[412/F_s]$ & $[746/F_s]$ seconds respectively.

We can write the entire problem in matrix-vector notation as:

$$\mathbf{y} = [\mathbf{A}_1, \mathbf{A}_2, \cdots \mathbf{A}_N]\, \mathbf{x} + \boldsymbol{\eta} \tag{4.4}$$

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \boldsymbol{\eta} \tag{4.5}$$

Where $\mathbf{A} \in \mathbb{R}^{M \times (N \times M)}$ is the block form of the N convolution matrices associated with the set of spike shapes $\{g(t)\}$. $\mathbf{x} \in \mathbb{R}^{(N \times M) \times 1}$ is the vertical concatenation of the firing time vectors, $\{\mathbf{x}_i\}$, associated with the N neurons. We can write this using the transpose operator as:

$$\mathbf{x} = \left[\mathbf{x}_1^T, \mathbf{x}_2^T, \cdots, \mathbf{x}_M^T\right]^T \tag{4.6}$$

Since the zero elements of $\mathbf{x}$ represent no spike, for a sufficiently high sampling rate, F_s, we expect $\mathbf{x}$ to be a sparse vector. For example the time series we analyse in section 4.6.6 is sampled at $F_s = 24$kHz. If we consider 5 neurons all firing with Poisson statistics at a physiologically fast rate of 100Hz we expect only one in every 48 elements of the solution vector to be non-zero. This requirement of sparsity plays a central role in our motivation for the development of a Basis Pursuit De-Noising (BPDN) approach to estimating the firing times.

4.4 Basis Pursuit De-Noising Approaches to Spike Time Detection

Given the dictionary set $\{g_i[t]\}$ the problem of estimating the set of firing times$\{t_k\}_i$ is an inverse linear problem of the deconvolution class. The simplest approach to solve this problem would be to use Ordinary Least Squares, which by the Gauss-Markov theorem is guaranteed to provide the Best Linear Unbiased Estimator (BLUE) in the presence of errors which are uncorrelated, with equal variance and an expectation value of zero [218]. Recall that for this problem (4.5) the noise sources are expected to be highly correlated, with no guarantee of constant variance and thus the Gauss-Markov theorem does not hold. Furthermore, as discussed in section 4.3.2 because neuronal firing rates are relatively low (maximally ~ 100 Hz) and signal sampling rates are relatively high ($\sim 24kHz$) we wish to bias the estimated firing time vector, $\hat{\mathbf{x}}$, to be sparse.

We can incorporate this sparsity by writing (4.5) as a constrained least squares problem with the constraint that the number of non-zero elements in the solution vector $\mathbf{x}$ is less than some value δ. We can mathematically represent this condition by requiring that the ℓ_0 pseudo-norm (which measures the number of non-zero elements of a vector) of $\mathbf{x}$ satisfy: $|\mathbf{x}|_0 \leq \delta$. Thus we seek the solution of the following mathematical problem:

$$\min_{\mathbf{x}} \quad |\mathbf{A}\mathbf{x} - \mathbf{y}|_2^2 \quad \text{s.t.} \quad |\mathbf{x}|_0 \leq \delta, \tag{4.7}$$

The problem with this strategy is that ℓ_0 minimisation is an NP-hard [219] combinatorial search problem [220]. Notice that to find the minimum least squared error for a fixed ℓ_0 pseudo-norm of δ we must search over $\binom{(N\times M)}{\delta}$ combinations. Since the least squares problem is constrained such that the ℓ_0 pseudo-norm can take any value from 0 to δ we must perform the search over $\sum_{i=1}^{\delta}\binom{(N\times M)}{i}$ combinations. Clearly this is not a feasible approach to develop sparse solutions. Multiple strategies have been developed to speedily deal with this problem, broadly these can be categorised into greedy approaches and convex relaxation approaches.

Greedy approaches have the advantage of being computationally efficient but have no theoretical guarantees regarding existence or stability criteria [221]. They typically work in a step-wise fashion, sequentially turning on elements according to some heuristic rule. Examples include Orthogonal Matching Pursuit (OMP) [222] which develops a linear combination of the active elements by identifying variables associated with the column of the system matrix which has the greatest inner product with the current residual. Another widely used alternative is Least Angle Regression (LARS) which sequentially adds elements which maximally reduce the residual correlation of the system. Greedy approaches specialised to the spike detection problem have also been constructed such as *Binary Pursuit* [191] which attempt to selectively turn elements on and off such that the decrease in a log-likelihood function is maximised.

The alternative solution is to use convex relaxation, where the ℓ_0 pseudo-norm in (4.7) is replaced by the ℓ_1 norm. This, in effect, replaces the ℓ_0 norm with its convex hull, converting the minimisation problem from a combinatorial one to a convex (but non-differentiable) one. The conditions under which the ℓ_0 pseudo-norm and ℓ_1 norm problems are equivalent is described by the Restricted Isometry Property (RIP) [223],[224],[225]. The RIP places conditions on both the system matrix $\mathbf{A}$ (and all permutations of its sub-matrices) and the sparsity of the solution vector $\mathbf{x}$. Unfortunately showing a matrix has bounded restricted isometry properties is also an NP-hard problem [226] and in practice ℓ_1 minimisation schemes are utilized without verifying whether the data satisfies the RIP. The ℓ_1 minimised least squares

problem may be written in an equivalent form as an OLS problem with an ℓ_1 regularisation term:

$$\min_{\mathbf{x}} \qquad |\mathbf{A}\mathbf{x} - \mathbf{y}|_2^2 + \lambda|\mathbf{x}|_1 \tag{4.8}$$

In the signals processing community this problem (4.8) is known as *Basis Pursuit De-Noising* [98] and is used for detecting signals, assumed to be sparse, from highly over complete dictionaries. In the statistics community (an equivalent form) of (4.8) is known as *Least Absolute Selection and Shrinkage Operator* (LASSO) [99] and is used for parsimonious model selection of high-dimensional multi variable regression problems. The use of these ℓ_1 minimisation problems is used in problems as diverse as magnetic resonance image processing [97], [227],[228], portfolio optimisation [229], (one of its original applications) reflection seismology [230] and more recently in the field of compressed sensing [100],[101].

The number of algorithms developed to solve (4.8) is legion. These methods include gradient methods [231], Iterative Shrinkage Thresholding (IST) [232], an accelated variant (FISTA) [233], and a variant with an intelligent step size [234] and Fixed Point Continuation (FPC) [235]. Alternative methods using Bregman iteration [236], shown to be equivalent to Augmented Lagrange Multipler (ALM) methods [237], Split Bregman Iteration (SBI) [238], Dual Augmented Lagrangian Methods (DALM) [239], interior point methods [227] and homotopy methods [240], [241] are also used. This list is by no means exhaustive, with different solution methodologies providing advantages and disadvantages depending on the sparsity of the signal to be estimated, the measurement noise level, the structure of the system matrix and computational efficiency.

There have been some studies which have attempted to catalogue the differences between these approaches. For example [242] compared five algorithms: gradient projection, homotopy, IST, Proximal Gradient, and ALM on both simulated datasets and on image processing problem of facial recognition. The results of [242] found that for the synthetic data no algorithm overall provided a superior combination of speed and accuracy for different sparsity and noise levels. Interestingly the results pertaining to the different algorithms obtained

from the simulated data were markedly different when applied to the real data. For the real world data the authors concluded that the homotopy method and ALM provided the best balance of performance and computational efficiency.

For the remainder of this chapter we will consider the following non-negative ℓ^1 minimisation problem:

$$\begin{aligned} \min_{\mathbf{x}} \quad & |\mathbf{A}\hat{\mathbf{x}} - \mathbf{y}|_2^2 + \lambda|\mathbf{x}|_1 \\ \text{s.t.} \quad & \mathbf{x} \succeq \mathbf{0} \qquad \text{(Non-negativity)} \end{aligned} \tag{4.9}$$

Where the non-negativity has been introduced to incorporate the principle that the Dirac delta trains representing the spiking times in $\mathbf{x}$ must, by definition, be positive. Note that the $\succeq$ symbol indicates element-wise greater than or less than value.

A variant of BPDN, referred to as *Continuous Basis Pursuit* was applied to spike sorting in [192]. The full description of CBP is provided in [243] but briefly, it can be considered an extension to BPDN which attempts to remedy errors associated with attempting to identify the spike times on a discretised lattice when the signal being sampled has the spikes embedded in continuous time. To improve on these errors, local shifts in the dictionary terms (identified spike shapes), of the order less than the sampling period: $\Delta T = 1/F_s$, are approximated using interpolation functions. In [243] linear Taylor series & trigonometric splines were considered and their associated minimisation problems were explicitly constructed.

In [192] this CBP algorithm was applied to extra-cellular recordings and the firing time estimates were further refined using the process of Iterative Reweighted ℓ_1 Minimisation. Reweighted ℓ_1 minimisation was introduced in [244] as a means of estimating sparser solutions with comparable reconstruction error to standard ℓ_1 minimisation. In Reweighted ℓ_1 minimisation the ℓ_1 norm: $|\mathbf{x}|_1$ is replaced by the weighted norm: $|\mathbf{W}\mathbf{x}|_1$. In [192] the elements of the diagonal weighting matrix were (similar to [244]) chosen proportional to an estimate of the elements of the solution vector: $W_{ii} \propto 1/x_i$. This approach penalises solution elements closer to zero. The Iterative Reweighted ℓ_1 Minimisation employed in [192] solves

the CBP initially with a weighting matrix equal to the identity matrix: $\mathbf{W} = \mathbf{I}$ and then iteratively solves the weighted ℓ_1 minimisation problem with the weighting matrix updated from the previous estimate until a convergence criterion is met. The individual weighted ℓ_1 minimisation problems are solved using the convex optimisation solver package *CVX* [245],[246].

The methodology introduced in [192] of using CBP & Iterative Reweighted ℓ^1 Minimisation forms the conceptual scaffolding for this chapter, nonetheless we argue that they introduce some unnecessary complications to the spike sorting problem. We argue that the advantages of CBP over BPDN are negligible for sufficiently over sampled neural systems. For example, although digital sampling of the extracellular recordings involves the discretisation of a fundamentally continuous process, in section 4.6.6 the experimental data is sampled at 24 kHz, whereas the physiological firing rates of the measured neurons is roughly 100Hz. Secondly, although Iterative Reweighted ℓ_1 Minimisation will undoubtedly improve the sparsity of the firing time estimates, these iterated ℓ_1 minimisations must be performed serially and thus the computational time will be multiplied by the number of iterations. Thirdly the interpolation functions associated with the CBP process generates a more complicated constrained convex optimisation problem to be solved, as well as additional variables which must be estimated. Lastly the cvx package used to solve the weighted ℓ_1 minimisation problem is self-described as not being recommended for large-scale problems [245].

For the remainder of this chapter we only consider the application of non-negative BPDN (4.9) to the spike sorting problem. We consider three algorithms: homotopy, Incrowd with Truncated-Newton Interior Point (TNIP) and Dual Augmented Lagrangian. We will discuss these algorithms individually in more depth in sections 4.4.1-4.4.3 but we briefly mention the motivation for choosing these three methods out of the multitude of previously discussed BPDN algorithms.

The homotopy approach is first considered because of the program's solution approach that single spikes are identified (i.e. non-zero elements of the solution vector are individually added or removed) in each iteration of the solver [240]. This spike by spike

approach seems intuitively appropriate for detecting sparse firing neurons.

The InCrowd approach is then considered because, in many ways, it can be considered an extension of the homotopy algorithm because it can identify groups of spikes (i.e. multiple non-zero elements of the solution vector are added or removed) in each iteration [220]. This 'groups of spikes' approach seems intuitively appropriate for situations where there is a high density of spikes contributing to the measured signal (i.e. if the firing rates are higher or there are more neurons contributing to the vLFP).

The Dual Augmented Lagrangian approach is finally considered as an alternative to these two previous methods because the program solves the BPDN problem using an entirely different approach in the Fenchel dual space [247].

we will develop these algorithms and in section 4.6.3 we will compare the accuracy of these algorithms for spike sorting simulated extracellular vLFP datasets.

4.4.1 Positive Homotopy Algorithm

The homotopy algorithm was first introduced in [240] for solution to the LASSO problem for over-determined systems. It was extended in [248] to deal with arbitrary measurement matrices. An extensive analysis of the theory and freely available matlab code was provided in [241], which as a special case, included the positive homotopy algorithm to the more general problem of applying the homotopy algorithm which could incorporate a warm-start (initial guess) solution vector.

KKT Conditions for Positive BPDN

We apply the Karush-Kuhn-Tucker (KKT) conditions to problem (4.9). Our objective function to minimise is $f(\mathbf{x}) = \frac{1}{2}|\mathbf{A}\hat{\mathbf{x}} - \mathbf{y}|_2^2 + \lambda \sum_{i=1}^{n} \mathbf{x}$. Notice that because of the non-negativity constraint we can replace the ℓ_1 norm with the summation operator over the individual $\mathbf{x}$ elements. We incorporate these non-negativity requirements, which are essentially *inequality*

constraints in the standard KKT framework as $g_i(x_i) = -x_i \quad \forall i = 1 \cdots N$. Notice that we have introduced the constant prefactor of $(1/2)$ in order to simplify calculations.

The Lagrangian associated with this problem is given by:

$$\begin{aligned} \mathcal{L}(\mathbf{x}, \mu) &= f(\mathbf{x}) + \sum_{i=1}^{n} \mu_i g_i(\mathbf{x}) \\ &= \frac{1}{2}|\mathbf{A}\mathbf{x} - \mathbf{y}|_2^2 + \lambda \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \mu_i x_i \end{aligned} \tag{4.10}$$

Where μ is the KKT multiplier associated with the inequality constraints. The KKT conditions to ensure the optimal solution $\mathbf{x}$ are:

$$\mathbf{A}^T(\mathbf{A}\mathbf{x} - \mathbf{y}) + \lambda \mathbf{1} - \boldsymbol{\mu} = 0 \quad \text{stationarity} \tag{4.11}$$

$$\forall i = 1 \cdots N: \quad \mu_{\mathbf{i}} \mathbf{x}_{\mathbf{i}} = 0 \quad \text{Complimentary Slackness} \tag{4.12}$$

$$\mu_i \geq 0 \quad \text{dual feasibility} \tag{4.13}$$

$$\mathbf{x} \succeq 0 \quad \text{primal feasibility} \tag{4.14}$$

Where $\mathbf{1}$ is the $N \times 1$ column vector of unity values and $\boldsymbol{\mu}$ is a column vector of the KKT multipliers such that $\boldsymbol{\mu} = [\mu_1, \mu_2 \cdots \mu_N]^T$. We partition the solution vector $\mathbf{x}$ into an active set $\mathbf{x}_I$ (s.t. $x_i \neq 0, \quad \forall i \in I$) and inactive set $\mathbf{x}_J$ (s.t. $x_j \neq 0, \quad \forall j \in J$) and $\mathbf{x} = \mathbf{x}_I \bigcup \mathbf{x}_J$. We define $\#(I)$ and $\#(J)$ as the number of elements (i.e. the ℓ_0 pseudo-norm) in the active and inactive sets respectively. We similarly partition the columns, $\mathbf{a}_k = \mathbf{A}_{k,:}$ of the system matrix $\mathbf{A}$, associated with elements of the solution vector $\mathbf{x}$ into two sub-matrices:

$$\mathbf{A}_I = (\mathbf{a}_i)_{i \in I} \in \mathbb{R}^{M \times \#(I)} \tag{4.15}$$

$$\mathbf{A}_J = (\mathbf{a}_j)_{j \in J} \in \mathbb{R}^{M \times \#(J)} \tag{4.16}$$

Where $\mathbf{A}_I$ and $\mathbf{A}_J$ represent the columns of matrix $\mathbf{A}$ associated with the active and in-active variables respectively. We now define the *residual correlations* consistent with [221],[241]:

$$\mathbf{c} = \mathbf{A}^T(y - \mathbf{A}\mathbf{x}) \tag{4.17}$$

Where $\mathbf{c}(I)$ & $\mathbf{c}(J)$ are the residual correlations of the elements of the active and in-active sets respectively. We will now develop the optimisation conditions separately for elements

associated with the active (non-zero) $\mathbf{x}_I$ and in-active (zero) $\mathbf{x}_J$ elements. On the active set, by the complimentary-slackness condition (4.12) the KKT multipliers must be zero and thus the stationarity condition (4.11) for active elements can be simplified to:

$$\mathbf{c}(I) = \mathbf{A}_I^{\mathbf{T}} (y - \mathbf{A}\mathbf{x}) = \lambda \mathbf{1}_I \tag{4.18}$$

On the inactive set by the dual feasibility condition guarantees $(\lambda - \mu_j) < \lambda$. We can therefore remove the μ_j variables, but include their effect by replacing the equality in the stationarity condition (4.11) equation with an inequality:

$$\mathbf{c}(J) = \mathbf{A}_J^{\mathbf{T}} (y - \mathbf{A}\mathbf{x}) \leq \lambda \mathbf{1}_J \tag{4.19}$$

Positive Homotopy Algorithm

The name homotopy comes from the fact that initially (4.9) is solved for the trivial problem of very large regularisation parameter λ such that the solution vector is the inactive set: $\mathbf{x} \rightarrow \mathbf{x}_J \equiv \mathbf{0}$. The regularisation parameter λ is subsequently reduced until it reaches the required value. For each value of the regularisation parameter the optimality conditions (4.18)-(4.19) must be obeyed, which results in elements being added (joining the active set) or being removed (joining the inactive set) for specific values of λ. Between the addition and removal of elements from the active set the solution vector follows the affine transformation given by the re-arrangement of active-set optimisation condition (4.18):

$$\mathbf{x}_{I,\lambda} \quad = \quad (\mathbf{A}_I^T \mathbf{A}_I)^{-1} \left(\mathbf{A}_I^T \mathbf{y} - \lambda \mathbf{1}_I \right) \tag{4.20}$$

Thus the change in the solution vector between these breakpoints can be described by replacing: $\lambda \rightarrow (\lambda - \gamma)$:

$$\mathbf{x}_{I,\lambda-\gamma} = \mathbf{x}_{I,\lambda} + \underbrace{(\mathbf{A}_I^T \mathbf{A}_I)^{-1} \mathbf{1}_I}_{\mathbf{u}} \gamma \tag{4.21}$$

Thus the update direction of the solution vector $\mathbf{u}$ is simply the inverse of the Gram matrix associated with the columns of the (currently) active set multiplied by a column vector of ones. Rather than continuously reducing the regularisation parameter, λ, we can reduce it in discrete jumps of γ associated with the addition or removal of an element from

the active set. An element will move from the inactive set to the active set when its residual correlation changes from *less* to equality with the current λ value.

Saturating the optimality condition for the *inactive* set (4.19) by *equality* of the left hand side with the regularisation parameter (the boundary of violation of the optimality condition for the inactive set) as the regularisation parameter is varied from $\lambda \rightarrow (\lambda - \gamma_{j,\min,\text{add}})$ yields:

$$\begin{aligned} \mathbf{a_j}^T \left(\mathbf{y} - \mathbf{A}_I \mathbf{x}_{I,\lambda-\gamma_{j,\min,\text{add}}}\right) &= \lambda - \gamma_{j,\min,\text{add}} \\ \underbrace{\mathbf{a_j}^T \left(\mathbf{y} - \mathbf{A}_I \mathbf{x}_{I,\lambda}\right)}_{\mathbf{c}_j} - \mathbf{a}_j^{\mathbf{T}} \underbrace{\mathbf{A}_I \mathbf{u}}_{\mathbf{v}} \gamma_{j,\min,\text{add}} &= \lambda - \gamma_{j,\min,\text{add}} \end{aligned} \tag{4.22}$$

Substituting in (4.21):

$$\gamma_{j,\min,\text{add}} = \frac{\lambda - c_j}{1 - \mathbf{a}_j^{\mathbf{T}} \mathbf{v}} \tag{4.23}$$

Scanning over all j elements of the inactive set we find the minimum $\gamma_{j,\min,1}$ which would result in an inactive set element becoming an active set element $x_j \rightarrow x_i$:

$$\gamma_{\min,\text{add}} = \min\left(\gamma_{j,\min,\text{add}}\right) \quad \forall j \in J \quad \text{s.t.} \quad \gamma_{\min,\text{add}} > 0 \tag{4.24}$$

The other alternative for the affine transformation to be broken is that an element of the active set, $x_i \rightarrow \mathbf{x}_J = 0$. The value for this to occur for each element of the active set is given using (4.21):

$$\gamma_{i,\min,\text{remove}} = -\frac{x_i}{u_i} \tag{4.25}$$

Scanning over all i elements of the active set we find the minimum $\gamma_{i,\min,\text{remove}}$ which would result in an active set element becoming an inactive set element $x_i \rightarrow x_j$:

$$\gamma_{\min,\text{remove}} = \min\left(\gamma_{i,\min,\text{remove}}\right) \quad \forall i \in I \quad \text{s.t.} \quad \gamma_{\min,\text{remove}} > 0 \tag{4.26}$$

The distance γ is the smallest distance which will either cause an element to be added or removed from the active set:

$$\gamma_{\min} = \min\left(\gamma_{\min,\text{add}}, \gamma_{\min,\text{remove}}\right) \tag{4.27}$$

The value of λ is sequentially reduced in these determined amounts $\gamma_{\min}$, and the solution vector $\mathbf{x}$ is transformed in the direction $\mathbf{u}$. Because the ℓ_1 minimisation problem is convex, by always obeying the optimality conditions we are guaranteed to head down to the global minimum associated with the problem. In the noise free case we keep decreasing λ until $\lambda \to 0^+$. At this point the $x_{\lambda \to 0}$ solution has the minimum $|x|_1$ which satisfies $A^T(Ax - y) = 0$ (that is has zero residual) [227]. In our situation where noise is present we bring λ down to the level of the ℓ_2 norm of the noise $|w|_2^2$. In practice this noise level is unknown, so the stopping criteria becomes a free parameter which must be tuned. It has been shown [221] that if there are k non-zero elements of a sparse solution vector, the homotopy path should have k vertices.

There is a very similar *greedy* algorithm, Least Angle Regression (LARS) which can also be applied to data where the solution vector $\mathbf{x}$ is restricted to non-negative values. LARS starts with a zero solution vector, and then sequentially turns on the elements with the greatest residual correlation. When one of the inactive variables has residual correlation equal to the active set elements, this element joins the active set and the solution vector updates in direction $\mathbf{u}$ which is the bisection between the active variables. The key difference between these two algorithms is that in homotopy elements can be removed from the active set, whereas in LARS they cannot. See [249] for further information on LARS and [221] for a discussion of when solution paths of homotopy and LARS converge.

Efficient Homotopy

The most computationally expensive steps of computing the homotopy path is calculating the inverse of the Grammian matrix $\left(\mathbf{A}_I^T \mathbf{A}_I\right)^{-1}$ for the update direction $\mathbf{u}$ in (4.21) [221] [241] and determining the residual correlations which are subsequently partitioned into $\mathbf{c}(I)$ & $\mathbf{c}(J)$. The residual correlations can be computed rapidly by recognising that the convolution operations can be quickly constructed using the Fast Fourier Transform (FFT). We describe how to calculate the residual correlation using the FFT in section A.8. The optimisation of the update direction, $\mathbf{u}$, calculation is more difficult and is described below.

We can efficiently speed up the calculation of the direction vector $\mathbf{u}$ using the following

two ideas. Firstly the homotopy algorithm sequentially adds or removes elements (corresponding to single columns of the system matrix $\mathbf{A}_I$) and thus are rank-1 updates/downdates of $\mathbf{A}_I$. The second idea is that rather than explicitly calculating the Grammian inverse, we can generate the new direction vector $\mathbf{u}_{\text{new}}$ based on an update to the previous direction vector $\mathbf{u}_{\text{old}}$. We will consider the generation of the new direction vector $\mathbf{u}_{\text{new}}$ for element addition and deletion seperately below.

In all calculations we represent the system matrix using the QR factorised form such that $\mathbf{A}_I = \mathbf{QR}$. We choose QR factorisation because it provides an excellent balance between numerical stability and computational speed. Note that if there are k active elements such that $\mathbf{A}_I \in \mathbb{R}^{M\times k}$ the factored matrices will be of size: $\mathbf{Q} \in \mathbb{R}^{M\times k}$ & $\mathbf{R} \in \mathbb{R}^{k\times k}$. These approaches were implemented, but not discussed in the code of [241].

Adding An Element to the Active Set

When an element is added to the active set, an additional column is appended to the end of the sub matrix $[\mathbf{A_I}, \mathbf{a}_i]$. The $\mathbf{R}$ matrix associated with the QR decomposition can be rapidly constructed using the Modified Gram-Schmidt algorithm with 'twice is enough' re-orthogonalisation [250]. The new update direction $\mathbf{u}_{\text{new}}$ can then be calculated from this $\mathbf{R}$ matrix and the previous update direction $\mathbf{u}_{\text{old}}$ as follows: Firstly calculate the update direction $\Delta\mathbf{u}$

$$\mathbf{r} = \mathbf{R}_{\text{new}}^{-1}[0_1, 0_2 \cdots 0_{\#I}, 1]^T \quad \text{(Backwards substitution)} \tag{4.28}$$

$$\mathbf{\Delta u} = \mathbf{r}^T \mathbf{1}_I \mathbf{r} \tag{4.29}$$

$$\mathbf{u}_{\text{new}} = [\mathbf{u}_{\text{old}}, 0]^T + \mathbf{\Delta u} \tag{4.30}$$

Removing an Element from the Active Set

In the case of removing a column, the efficient update is slightly more complicated. The QR re-factorisation can be calculated using Givens rotation matrices $\mathbf{G}$ [250]. Similar to adding a column we can iteratively determine the update direction as follows: Assume we are removing the i^{th} column from the system matrix: $\mathbf{A_I} \rightarrow [\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_{i-1}, \mathbf{a}_{i+1}, \cdots, \mathbf{a}_N]$.

We first apply a permutation matrix $\mathbf{P}$ which brings the i^{th} column of the current $\mathbf{R}$ matrix $\mathbf{r}_i$ to the position of the final column of $\mathbf{R}_{\text{perm}}$: $\mathbf{R}_{\text{perm}} \rightarrow \mathbf{PR}$.

$\mathbf{R}_{\text{perm}}$ does not have the required upper triangular structure for a QR decomposition. To restore this we apply a series of Given rotatation matrices, $\mathbf{G}$ to subsections of the permuted $\mathbf{R}$ matrix to restore $\mathbf{R}_{\text{perm}}$ (excluding the last column) to upper triangular form. The first Givens matrix $\mathbf{G}$ will be applied to the i^{th} and $(i+1)^{th}$ rows, which due to the zeros in the preceeding columns of $\mathbf{R}_{\text{perm}}$, will only effect the proceeding columns $\mathbf{r}_i \cdots \mathbf{r}_N$:

$$\mathbf{R}_{\text{new,perm}}(i : i+1, i : N) \quad \rightarrow \quad \mathbf{G}^{\mathbf{T}}_{\mathbf{i,i+1}} \mathbf{R}_{\text{perm}}(i : i+1, i : N) \tag{4.31}$$

We continue this process starting at column i and continuing until column $(N-1)$. We thus re-form the upper triangular matrix structure (excluding the final permuted column) with a series of these Givens matrices:

$$\mathbf{R}_{\text{new,perm}} \quad \rightarrow \quad \mathbf{G}^{\mathbf{T}}_{\mathbf{N-1,N}} \cdots \mathbf{G}^{\mathbf{T}}_{\mathbf{i+1,i+2}} \mathbf{G}^{\mathbf{T}}_{\mathbf{i,i+1}} \mathbf{R}_{\text{perm}} \tag{4.32}$$

Where similar to equation (4.31) only a subset of the entire $\mathbf{R}_{\text{new,perm}}$ needs to be recalculated with each pre-multiplication by the Givens rotator. See [250] for a more detailed description of this process and [251] for a more in depth discussion about applying the Givens matrices to a sub matrix of the $\mathbf{R}_{\text{perm}}$ matrices.

Sub matrices of the Q matrix can be similarly updated:

$$\mathbf{Q}_{\text{new,perm}}(1 : M, i : i+1) = \mathbf{Q}_{\text{perm}}(1 : M, i : i+1)\mathbf{G}_{i,i+1} \tag{4.33}$$

We use a series of these:

$$\mathbf{Q}_{\text{new,perm}} = \mathbf{Q}(1 : M, i : i+1)\mathbf{G}_{i,i+1}\mathbf{G}_{i+1,i+2} \cdots \mathbf{G}_{N-1,N} \tag{4.34}$$

The effect of this column which is to be removed (which is now in the final column) can be similarly calculated with:

$$\mathbf{r} \quad = \quad \mathbf{R}^{-1}_{\text{perm,new}} \left[0_1, 0_2 \cdots 0_{\#I-1}, 1\right]^T \tag{4.35}$$

The *change* in the update direction is given by:

$$\Delta \mathbf{u} = \mathbf{r}^T \mathbf{1}_I \mathbf{r} \tag{4.36}$$

The update direction can then be iteratively calculated from the previous direction using:

$$\mathbf{u}_{\text{new}} = \mathbf{P}\mathbf{u}_{\text{old}} - \Delta \mathbf{u} \tag{4.37}$$

$$\mathbf{u}_{\text{new}} \rightarrow \mathbf{u}_{\text{new}(1\cdots I-1)} \tag{4.38}$$

Because we have permuted the $\mathbf{R}$ matrix, we must similarly permute the update direction. Because we are removing an element, our update direction must involve one less direction and thus we must remove the final element of the calculated update direction (because we permuted the element to be removed to the end of the system matrix)

Positive Homotopy Pseudo-Code

1. Initialise solution vector $\mathbf{x} = 0$. Determine initial residual correlation $\mathbf{C}_0 = \mathbf{A}^T\mathbf{y}$.
2. Starting homotopy parameter value: $\lambda = \max(\mathbf{C}_0)$. Initial active element x_i is the element with maximal residual correlation.
3. Determine initial direction vector $\mathbf{u}$ using (4.21).

   while $\lambda < \tau$

4. Calculate the γ_{min} (4.27) using (4.24) & (4.26).
5. Update active elements: $\mathbf{x}_I = \mathbf{x}_I + \gamma \mathbf{u}$.
6. Decrease $\lambda \rightarrow \lambda - \gamma$.
7. Add or remove active element depending on outcome of (4.27).
8. Calculate correlated residuals over the new inactive set $\mathbf{c_J}$ using (A.99).
9. Determine new direction vector $\mathbf{u}$ using (4.30) when an element is added or (4.38) when an element is removed.

4.4.2 InCrowd Algorithm with Truncated Newton Interior Point

In the preceding sections (4.4.1) we discussed the homotopy algorithm which solves the BPDN problem by starting with a zero vector and then sequentially adding and removing elements from the active solution set as the regularisation parameter is decreased. We now introduce the *InCrowd algorithm* [220] which can be thought of as an extension to homotopy where instead of adding/removing individual elements, at each step we add and remove a *collection* of active elements. The advantage of this approach is that if we have multiple neurons contributing spikes to the signal we may be able to identify potential sets of these spikes in a single pass, compared to the step by step approach of identifying spikes which occurs in the homotopy method.

The InCrowd algorithm works, similar to homotopy, by iteratively updating a solution vector $\mathbf{x}$ which is partitioned into active $\mathbf{x}_I$ and inactive $\mathbf{x}_J$ subsets. Also similar to homotopy the initialisation of the solution vector is the zero vector: $\mathbf{x} = \mathbf{0} \equiv \mathbf{x}_J$.

At each step of the InCrowd algorithm the residual correlations are calculated over the inactive set. The L_{ic} variables with residual correlations $\mathbf{c_J}$ greater than some threshold τ_{ic} are added to the active set as candidate active set elements. The BPDN algorithm (4.9) is then solved over the active subset using any desired method from convex optimisation. Thus at each step this process can result in candidate active elements either becoming inactive or active and elements activated in previous steps either staying active or becoming inactive. This process is repeated until no remaining elements of the inactive set have a residual correlation greater than τ_{ic}. The principle advantage of this method is that, for $L_{ic} \ll (N \times M)$ the computationally intensive part of BPDN is solved for a significantly smaller problem. In the code provided in [220] the default value is $L_{ic} = 20$ Note that there are similar algorithms where instead of BPDN an OLS step is performed, such as *subspace pursuit* [252] and *compressed orthogonal matching pursuit* [253] which also have similar theoretical guarantees if the RIP is satisfied. Similar to [220] we have found these methods to provide solutions with higher error rates and slower speeds of convergence compared to the BPDN approaches.

Similar to the homotopy method we can speed up the calculation of the residual correlations using FFTs as described in section A.8. Note that in order to enforce the non-negativity solution condition we only consider candidate active elements which have positive residual correlations. This is due to the theorem, proven in [254] that a solution vector element, x_i, must have the same sign as its residual correlation: $\text{sgn}(x_i) = \text{sgn}(c_i)$. Note that in addition the algorithm used to solve the BPDN problem over the identified active subset must incorporate the non-negativity condition.

The following properties of the InCrowd algorithm are discussed in [220]. The InCrowd method is guaranteed to solve the BPDN and converge to the global minimum. Under certain conditions on the columns of the system matrix $\mathbf{A}$ ($|\mathbf{a}_i|_2^2 = 1$), conditions are developed for the error term to initially decrease exponentially. It is also shown that this method cannot retrace its own path. Despite this, similar to homotopy, individual elements can be re-added and removed over multiple cycles (described as *model-churning* in [221]). Nonetheless the combination of elements comprising the active set at each step is guaranteed to be unique.

InCrowd Inner Loop: Truncated Newton Interior Point Method:

As discussed in the previous section, The BPDN solution step for the InCrowd algorithm over the active set can be performed with any BPDN solver. We solve the BPDN step using the Truncated Newton Interior Point (TNIP) method, which with our simulations has been fast, acurate and allows for easy incorporation of the non negativity condition. The TNIP was described in [227] and provided as the Matlab package *ls-l1*. We provide an overview of the TNIP method for BPDN following [227], but refer the reader to this article for further technical details.

The TNIP algorithm is an interior point method which finds the solution in the interior point using an iterative central path method with a series of log-barrier functions with

increasing penalties. The central path framework is set up as follows:

$$\phi_t(\mathbf{x}) = |\mathbf{Ax} - \mathbf{y}|_2^2 + \lambda \sum_{i=1}^{N} x_i - \left(\frac{1}{t}\right) \sum_{i=1}^{N} log(x_i) \tag{4.39}$$

The last term in (4.39) is termed the logarithmic barrier function, which allows us to incorporate the non-negativity condition as an unconstrained optimisation problem. Notice that this central path problem (4.39) is convex, and due to the domain of the logarithmic functions is restricted to the interior of the feasible points $x_i \leq 0 \quad \forall i$ [247]. The set of points, $\mathbf{x}_t^\star$ which minimise (4.39) for varying values of t is known as the *central path*. The value of t controls the penalty for being in an infeasible region ($x_i < 0$). The larger the parameter t, the more accurately the infeasable region is modelled, but obtaining the solution becomes numerically more difficult because the Hessian varies rapidly near the boundary of the feasible set [247]. The approach taken is to iteratively generate the central path $\{\mathbf{x}_t^\star\}$ by optimising equation (4.39) for small values of t, then using this solution for the starting point for generating the central point for a larger value of t.

It can be shown that a point on the central path $\mathbf{x}_t^\star$ satisfy all the KKT conditions, except complimentary slackness, which satisfies $\mu_i x_i = 1/t$ [247]. Note that in the limit of t approaching infinity the KKT conditions will be satisfied and for t sufficiently large the KKT conditions are 'almost' satisfied. It is also important to note the bounds on the sub-optimality of the solution given by: $\mathbf{x}_t^\star$ [247], [227] :

$$f_0(\mathbf{x_t^\star}) - p^\star \geq \frac{2N}{t} \tag{4.40}$$

Where $p^* = f_0(x^*)$ is the unknown optimal primal solution and N is the number of constraints, in this case the length of the solution vector $\mathbf{x}_t^\star$. It is interesting to compare the central path approach to solving (4.9) to the homotopy approach discussed in section 4.4.1. The homotopy method starts by solving an easier problem (associated with a large regularisation parameter and subsequently a single active element) and then slowly deforms the problem, satisfying the KKT conditions at each step until the required solution (corresponding to a specific regularisation parameter) is obtained. The central path approach does not vary the regularisation parameter, but rather constructs a sequence of simpler problems

which at each step only approximate the KKT conditions. For each step in this sequence of simpler problems the approximation to the KKT conditions become more valid.

A given value of the parameter t in (4.39) is solved for the minimal value of $\mathbf{x}_t$ using Newtons method to determine the search direction (Δx). The search direction is the solution of:

$$\nabla^2 \phi_t(\mathbf{x}) \Delta x + \nabla \phi_t = 0 \tag{4.41}$$

Where:

$$\nabla^2 \phi_t(\mathbf{x}) = 2t\mathbf{A}^T\mathbf{A} + \text{diag}\left[\frac{1}{x_1^2}, \frac{1}{x_2^2} \cdots \frac{1}{x_N^2}\right] = 2t\mathbf{A}^T\mathbf{A} + \mathbf{D} \tag{4.42}$$

$$\nabla \phi_t(\mathbf{x}) = 2t\mathbf{A}^T(\mathbf{A}\mathbf{x} - \mathbf{y}) + \left[t\lambda - \frac{1}{x_1}, t\lambda - \frac{1}{x_2}, \cdots t\lambda - \frac{1}{x_N}\right]^T \tag{4.43}$$

Note that the Hessian is positive and symmetric and thus the search direction is solved approximately using the Pre-Conditioned Conjugate Gradient (PCGC) method. The pre-conditioner which is recommended in [227] and used in *l1-ls* is the *Jacobi* preconditioner, $\mathbf{P}$, based on the diagonals of the Hessian:

$$\mathbf{P} = \text{diag}\left(2t\mathbf{A}^\mathbf{T}\mathbf{A}\right) + \mathbf{D} \tag{4.44}$$

The $\mathbf{x}_{t,i}$ value is then updated to:

$$\mathbf{x}_{t,i+1} = \mathbf{x}_{t,i} + s\Delta x \tag{4.45}$$

The value of the step size s is calculated using a back-tracking line search. The method described so far can be continued indefinitely, subsequently increasing the penalty pre-factor t resulting in slower and slower convergence to a solution. Note that since the problem satisfies Slaters condition [227] the duality gap Δ_{pd} between the optimal primal and dual loss functions is zero:

$$\Delta_{p^\star, d^\star} = p^\star - d^\star = 0 \tag{4.46}$$

Thus a stopping criteria can be constructed based on the duality gap being sufficiently small. In order to identify the duality gap the Lagrangian dual problem is constructed and a subsequent dual feasible point is identified.

Calculation of Duality Gap

The ℓ^1 regularised least squares problem (4.9) without the non-negativity constraint can be re-written as:

$$\text{minimise} \quad f(x,z) = \mathbf{z}^t\mathbf{z} + \lambda|\mathbf{x}|_1 \tag{4.47}$$

$$\text{such that} \quad z = \mathbf{A}\mathbf{x} - \mathbf{y} \tag{4.48}$$

Thus the Lagrangian dual problem to (4.48) can be written as:

$$\begin{aligned} \text{maximise} \quad G(\mathbf{v}) &= -(\frac{1}{4}\mathbf{v}^T\mathbf{v} + \mathbf{v}^T\mathbf{y}) \\ \text{subject to} \quad \mathbf{a}_i^T\mathbf{v} &\leq \lambda \quad \forall i = 1 \cdots N \end{aligned} \tag{4.49}$$

We can construct a dual feasible point which satisfies (4.49):

$$\tilde{\mathbf{v}} = \frac{\lambda\,(\mathbf{A}\mathbf{x} - \mathbf{y})}{||\mathbf{A}^T\,(\mathbf{A}\mathbf{x} - \mathbf{y})\,||_\infty} \tag{4.50}$$

With this dual feasible point $\tilde{\mathbf{v}}$, we can use equations (4.47), (4.49) & (4.50) to develop the duality gap $\Delta_{p,d}$:

$$\begin{aligned} \Delta_{p,d} &= f(\mathbf{x},\mathbf{z}) - G(\tilde{\mathbf{v}}) \\ &= |\mathbf{A}\mathbf{x} - \mathbf{y}|_2^2 + \lambda|\mathbf{x}|_1 + (\frac{1}{4}\tilde{\mathbf{v}}^T\tilde{\mathbf{v}} + \tilde{\mathbf{v}}^T\mathbf{y}) \end{aligned} \tag{4.51}$$

The TNIP algorithm is terminated when the *relative* duality gap, $\Delta_{p,d}/f(\mathbf{x},\mathbf{z})$, is less than the specified tolerance: τ_{gap}.

InCrowd Pseudo-Code

1. Start with the active set being the null set $I \in \emptyset$, and the solution vector being the zero vector , $\mathbf{x} = \mathbf{0}$.

2. Calculate the correlated residual over the in-active set $\mathbf{c}_J$ using (A.99) and indexing the residual correlation terms associated with the inactive set.. In [220] this is referred to as the 'usefulness' of the inactive set.

3. Determine if any of the residual correlations, $\mathbf{c_J}$ are above the regularisation threshold $(\mathbf{c}_J) < \tau_{ic}$. If no residual correlations are above this threshold then end the algorithm with the current solution $\mathbf{x}$.

4. If residual correlations exist greater than τ, select the elements with the L_{ic} largest residual correlations.

5. Solve the BPDN problem exactly over the subspace of the active set variables $\mathbf{x}_I$. Note that this will be a *dense* problem, albeit in a much smaller dimensional space than the full scale problem:

   (a) set: $t = v1/\lambda \quad \mathbf{x}_I = 0$.
   While $\Delta_{p*,d*}/f(\mathbf{x}, \mathbf{z}) \leq \tau_{\text{gap}}$.

   (b) Compute direction $\Delta\mathbf{x}$ using Newton's method to solve (4.41) with PGC.

   (c) Compute step size γ in direction Δx using a back-tracking line search.

   (d) $\mathbf{x}_t = \mathbf{x}_t + s\Delta x_t$.

   (e) Construct dual feasible point $\mathbf{v}$ from (4.50).

   (f) calculate relative duality gap (4.51).

   (g) update t.

6. Prune the solution components from the active set which were set to zero $\mathbf{x}_i = 0, I \rightarrow J$ in constrained minimisation problem (step 5). Update the solution vector on the active set, $\mathbf{x}_I$, with the values determined in step 5.

7. Update the correlated residuals over the inactive set, $\mathbf{c}_J$, using (A.99) and indexing the residual correlation terms associated with the *new* inactive set.

8. Return to step 3.

4.4.3 Positive Dual Augmented Lagrangian Method

In this section we describe an alternative method to solving the non-negative BPDN problem, based on solution in the dual space described in [239]. This approach is based on developing an alternative dual, the *Fenchel dual* and then attempting to maximise the Lagrangian associated with this dual problem. The Lagrangian is iteratively maximised in a step-wise fashion using the *Augmented Lagrangian* (with quadratic penalty) method, with the Lagrange multipliers updated in each iteration. The advantage of this approach is that, similar to the homotopy method, the computational complexity of this algorithm is proportional to the sparsity of the system being estimated. This is in contrast to the TNIP method used with the InCrowd method where the computational complexity is related to the structure of the system matrix $\mathbf{A}$. We provide a brief overview of this method with enough detail to grasp the underlying intuition and implement it numerically. The interested reader is referred to [239] for the full technical details of the convergence guarantees and [255] for an extension of this method to a more general family of BPDN approaches. The authors of both these papers have made code available based on this method, which we use in simulations, as part of the *DAL* program.

The key to understanding this method is that the Lagrange multiplier associated with the Lagrangian of the Fenchel dual is also the primal (spike firing time) variable $\mathbf{x}$. In the Augmented Lagrangian framework *both* the variables to be optimised and the Lagrange multiplers are updated in an iterative fashion towards their optimal value. Thus this process of updating the Lagrange multipliers associated with the Fenchel dual problem provides the primal solution variable $\mathbf{x}$. Similar to the TNIP method described in section 4.4.2, the solution is considered optimal once the duality gap is sufficiently small.

Similarly to Section 4.4.1 we will rewrite the positive BPDN problem (4.9) into a slightly different, but equivalent form more amenable for the positive DAL algorithm:

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2}|\mathbf{A}\hat{\mathbf{x}} - \mathbf{y}|_2^2 + \lambda \sum_{i=1}^{P} x_i \\ \text{s.t.} \quad & \mathbf{x} \succeq \mathbf{0} \qquad \text{(Non-negativity)} \end{aligned} \tag{4.52}$$

Since we are constrained to positive elements of the solution vector $\mathbf{x}$, the ℓ_1 norm of $\mathbf{x}$ has been replaced with the summation over the elements of $\mathbf{x}$. We have also pre-multiplied the OLS term by $1/2$ in order to make some of the subsequent calculations simpler. Central to the DAL algorithm is the Fenchel-Rockafeller Duality theorem, which provides the relationship between the infimum of a cost function and the supremum of its dual problem constructed from Fenchel-Legendre transformations. We first introduce these Fenchel-Legendre transformations:

$$f^*(\boldsymbol{\alpha}) = \sup_{\mathbf{x}\in\mathbb{R}^n} \quad \langle \mathbf{x}, \boldsymbol{\alpha} \rangle - f(\mathbf{x}) \tag{4.53}$$

The Fenchel-Legendre transformation, also referred to as the *convex conjugate*, can be understood as the convex hull associated with the epigraph of function $f(\mathbf{x})$ in terms of its supporting hyperplanes. Note that by construction the Fenchel-Legendre transformation is a convex function.

The Fenchel-Rockafeller duality theorem [256],[257] states:

$$\inf_{x\in\mathbb{R}^m} f(\mathbf{x}) + g(\mathbf{A}\mathbf{x}) = \sup_{\boldsymbol{\alpha}\in\mathbb{R}^n} -f^*(-\mathbf{A}^{\mathbf{T}}\boldsymbol{\alpha}) - g^*(-\alpha) \tag{4.54}$$

Where $\mathbf{A}$ is a matrix which maps from $\mathbb{R}^m$ to $\mathbb{R}^n$. The Fenchel-Rockafeller duality theorem (4.54) states that the minimum in the primal space can be determined by an equivalent minimum problem in terms of the Fenchel-Legendre conjugates. Notice also that the point wise supremum of convex functions is also necessarily convex. For the BPDN problem (4.9) The individual Legendre-Fenchel transforms are given below.

The Fenchel-Legendre transformation of the regularisation (summation) term in (4.52) is derived below:

$$f^*(\boldsymbol{\alpha}) \quad = \quad \sup_{\mathbf{x}\in\mathbb{R}_P^+} \quad \langle \boldsymbol{\alpha}, \mathbf{x} \rangle - \lambda \sum_{i=1}^{P} x_i = \sum_{i=1}^{P} (\alpha_i - \lambda)\, x_i \tag{4.55}$$

Notice that because of the non-negativity constraint on $\mathbf{x}$ we are constrained to consider only positive values of $\mathbf{x}$. Also notice that if any element $\alpha_i > \lambda$ the maximum value that

$f^*(\boldsymbol{\alpha})$ may obtain is infinity, in the limit of x_i approaching infinity. For $\alpha_i \leq \lambda$ the maximum value that $f^*(\boldsymbol{\alpha})$ may obtain is zero. Thus:

$$f^*(\boldsymbol{\alpha}) = \ell_\lambda^{\infty^+}(\boldsymbol{\alpha}) = \begin{cases} 0 & \text{if} \quad \alpha_i \leq \lambda \\ +\infty & \text{if} \quad \alpha_i > \lambda \end{cases} \quad \forall i \tag{4.56}$$

Where we define $\ell_\lambda^{\infty^+}(\alpha)$ as the subset of the ℓ_∞ ball of radius λ for only *positive* values. The Fenchel-Legendre transform of the regularisation term is therefore:

$$f(\mathbf{x}) = \lambda \sum_{i=1}^{P} x_i \quad \Rightarrow \qquad f^*(\boldsymbol{\alpha}) = \ell_\lambda^{\infty^+}(\boldsymbol{\alpha}) \tag{4.57}$$

The Legendre-Fenchel transformation of the sum of squares term is:

$$g^*(\boldsymbol{\alpha}) = \sup_{\mathbf{x} \in \mathbb{R}_P^+} \quad \langle \boldsymbol{\alpha} \mathbf{x} \rangle - \frac{1}{2} ||\mathbf{x} - \mathbf{y}||^2 \tag{4.58}$$

Since all the expressions in (4.58) are smooth, standard calculus techniques can be applied to identify that the maximal value of $\mathbf{x}$ is given by: $\mathbf{x} = \mathbf{y} + \boldsymbol{\alpha}$. Substituting this value into (4.58) and taking the *negative argument* of the convex conjugate, as is required for the Fenchel-Rockafeller (4.54) theorem, yields:

$$g^*(-\boldsymbol{\alpha}) = \frac{1}{2} ||\boldsymbol{\alpha} - \mathbf{y}||^2 - \frac{1}{2} ||\mathbf{y}||^2 \tag{4.59}$$

The Fenchel-Rockafeller dual to BPDN problem (4.9) can be written using (4.54), (4.57) & (4.59) as:

$$\sup_{\boldsymbol{\alpha} \in \mathbb{R}^m} \underbrace{-\delta_\lambda^{\infty^+}\left(\mathbf{A}^\mathbf{T} \boldsymbol{\alpha}\right) - \frac{1}{2} ||\boldsymbol{\alpha} - \mathbf{y}||^2 + \frac{1}{2} ||\mathbf{y}||^2}_{d(\boldsymbol{\alpha})} \tag{4.60}$$

One of the principal problems with (4.60) is the coupling of the $\boldsymbol{\alpha}$ variables due to the system matrix $\mathbf{A}$. We can re-write the above coupled optimisation problem, as a decoupled optimisation problem, with the coupling included as a constraint which must be satisfied:

$$\sup_{\boldsymbol{\alpha}, \mathbf{v}} \qquad \underbrace{-\delta_\lambda^{\infty^+}(\mathbf{v}) - \frac{1}{2} ||\boldsymbol{\alpha} - \mathbf{y}||^2 + \frac{1}{2} ||\mathbf{y}||^2}_{d(\boldsymbol{\alpha}, \mathbf{v})} \tag{4.61}$$

$$\text{subject to} \qquad \mathbf{v} = \mathbf{A}^\mathbf{T} \boldsymbol{\alpha} \tag{4.62}$$

We can now write a Lagrangian associated with this constrained dual problem. We will use the *Augmented Lagrangian* function. In this approach our constraint is incorporated with a Lagrange multiplier and in addition we have a quadratic penalty term for not satisfying the constraint:

$$\begin{aligned}\mathcal{L}_\eta(\boldsymbol{\alpha}, \mathbf{v}, \mathbf{w}) &= \mathbf{d}(\boldsymbol{\alpha}, \mathbf{v}) - \mathbf{w}^\mathbf{T}\left(\mathbf{A}^\mathbf{T}\boldsymbol{\alpha} - \mathbf{v}\right) - \eta||\mathbf{A}^\mathbf{T}\boldsymbol{\alpha} - \mathbf{v}||^2 \\ &= -\delta_\lambda^{\infty^+}\left(\mathbf{A}^\mathbf{T}\boldsymbol{\alpha}\right) - \frac{1}{2}||\boldsymbol{\alpha} - \mathbf{y}||^2 + \frac{1}{2}||\mathbf{y}||^2 - \mathbf{w}^\mathbf{T}\left(\mathbf{A}^\mathbf{T}\boldsymbol{\alpha} - \mathbf{v}\right) - \frac{\eta}{2}||\mathbf{A}^\mathbf{T}\boldsymbol{\alpha} - \mathbf{v}||^2\end{aligned} \tag{4.63}$$

Where $\mathbf{w}$ are the Lagrange multipliers associated with the constraint (4.62). Notice that because this is a convex problem, and we are taking the Lagrangian associated with the dual problem, these Lagrange multipliers are also the primal solution $\mathbf{x}$ variable we seek. In the augmented Lagrangian framework the Lagrange multipliers are updated in an iterative fashion according to the equation:

$$\mathbf{w}_{k+1} = \mathbf{w}_k + \eta_k\left(\mathbf{A}^\mathbf{T}\boldsymbol{\alpha_k} - \mathbf{v}_k\right) \tag{4.64}$$

The dual problem [247] associated with the Lagrangian of our Fenchel dual problem is:

$$f_\eta(\mathbf{w}) = \sup_{\boldsymbol{\alpha}\in\mathbb{R}^m, \mathbf{v}\in\mathbf{R}^n} \mathcal{L}_\eta(\boldsymbol{\alpha}, \mathbf{v}, \mathbf{w}) \tag{4.65}$$

The maximisation in (4.65) with respect to $\mathbf{v}$ can be performed analytically, reducing the problem to a single variable optimisation:

$$\begin{aligned}\sup_{\mathbf{v}\in\mathbf{R}^n} \mathcal{L}_\eta(\boldsymbol{\alpha}, \mathbf{v}, \mathbf{w}) &= -\frac{1}{2}\left|\left|\boldsymbol{\alpha} - \mathbf{y}\right|\right|^2 - \min_{\mathbf{v}\in\mathbb{R}^n}\left(\delta_\lambda^{\infty^+}(\mathbf{v}) + \frac{\eta}{2}||\nu - \mathbf{A}^T\boldsymbol{\alpha} - \frac{\mathbf{w}}{\eta}||^2\right) \\ &\quad -c(\mathbf{w}, \eta)\end{aligned} \tag{4.66}$$

Notice that we have split finding the supremum of the Lagrangian with respect to the dual variable $\mathbf{v}$, (4.66) into a component which depends on the variables to be optimised $(\boldsymbol{\alpha}, \mathbf{v})$ and a component which depends on the variables which are not to be optimised $(\mathbf{w},\eta)$. Notice that this $c(\mathbf{w}, \eta)$ function will not be required for the optimisation procedure because the function $f_\eta(\omega)$ only requires the additional optimisation with respect to the $\boldsymbol{\alpha}$ variable and the Lagrange multiplier (which is also the primal solution vector $\mathbf{x}$) $\mathbf{w}$ are updated based

only on these $\boldsymbol{\alpha}, \mathbf{v}$ optimised variables.

Notice that in the above equation all the terms containing the $\mathbf{v}$ variable are ℓ_2 norm values pre-multiplied by negative one and therefore must be negative. Thus in order to maximise the Lagrangian the $\mathbf{v}$ terms must all be minimal. The presence of the ℓ_{∞^+} term and the minimisation condition constrains us to the projection onto the positive subset of the ℓ_∞ ball of radius λ: $P_\lambda^{\infty^+}(\cdots)$. Therefore:

$$\sup_{\mathbf{v}\in\mathbf{R}^n} \mathcal{L}_\eta(\boldsymbol{\alpha}, \mathbf{v}, \mathbf{w}) = -\frac{1}{2}\left|\left|\alpha - \mathbf{y}\right|\right|^2 - \frac{\eta}{2}\left|\left|\mathbf{A}^T\boldsymbol{\alpha} + \frac{\mathbf{w}}{\eta} - P_\lambda^{\infty^+}\left(\mathbf{A}^T\boldsymbol{\alpha} + \frac{\mathbf{w}}{\eta}\right)\right|\right|^2 \tag{4.67}$$

Following [239] the projection onto the ℓ_∞ ball of radius λ can be related to the soft thresholding function ($\mathrm{ST}_\lambda = \mathbf{w} - P_\lambda^\infty(\mathbf{w})$). Since our problem projects only onto the positive component of the ℓ_∞ ball we consider the positive soft thresholding function:

$$\begin{aligned} \mathrm{ST}_\lambda^+(\mathbf{w}) &= \mathbf{w} - P_\lambda^{\infty^+}(\mathbf{w}) \\ &= \left(\max(w_j - \lambda, 0)\right)_j \qquad \forall j = 1\cdots n \end{aligned} \tag{4.68}$$

The supremum of the Lagrangian with respect to the $\mathbf{v}$ dual variable, is given by substituing (4.68) into (4.67). Therefore $f_\eta(\mathbf{w})$ can now be written in terms of the maximisation of the single dual variable α:

$$f_\eta(\mathbf{w}) = \max_{\boldsymbol{\alpha}\in\mathbf{R}^m} \quad -\left|\left|\boldsymbol{\alpha} - \mathbf{y}\right|\right|^2 - \frac{\eta}{2}\left|\left|\mathrm{ST}_\lambda^+\left(\mathbf{A}^T\boldsymbol{\alpha} + \frac{\mathbf{w}}{\eta}\right)\right|\right|^2 \tag{4.69}$$

The expression for the update of the Lagrange multiplier $\mathbf{w}_{k+1}$ (4.64) can be written in terms of the projection operator on the $\mathbf{v}_k$ dual variable: $\mathbf{v}_k = \Gamma_\lambda^{\infty^+}\left(\mathbf{A}^T\boldsymbol{\alpha} + \omega/\eta\right)$:

$$\begin{aligned} \mathbf{w}_{k+1} &= \mathbf{w}_k + \eta_k\mathbf{A}^T\boldsymbol{\alpha}_k - \eta_k\Gamma_\lambda^{\infty^+}\left(\frac{\mathbf{w}_k + \eta_k\mathbf{A}^T\boldsymbol{\alpha_k}}{\eta_k}\right) \\ &= \mathrm{ST}_{\lambda\eta_k}^+\left(\mathbf{w}_k + \eta_k\mathbf{A}^\mathbf{T}\boldsymbol{\alpha}\right) \end{aligned} \tag{4.70}$$

Where we have used the identity for the projection operation: $\Gamma_{\lambda\eta_k}^{\infty^+}(\mathbf{w}) = \Gamma_\lambda^{\infty^+}(\eta_k\mathbf{w})$.

We now turn our attention to maximising $f_\eta(\mathbf{w})$ (4.69) over the variable $\boldsymbol{\alpha}$. Notice that similar to before, the objective to be maximised consists of ℓ_2 norm of functions which are

pre-multiplied by negative one. Thus the maximisation of $f_\eta(\mathbf{w})$ is equivalent to the following minimisation over the $\boldsymbol{\alpha}$ dual variable:

$$\boldsymbol{\alpha_k} = \arg\min_{\boldsymbol{\alpha}_k \in \mathbb{R}^m} \left\{ \frac{1}{2} ||\boldsymbol{\alpha} - \mathbf{y}||^2 + \frac{\eta_k}{2} ||\mathrm{ST}^+_\lambda \left(\mathbf{A^T}\boldsymbol{\alpha} + \frac{\mathbf{w}_k}{\eta} \right) ||^2 \right\} \tag{4.71}$$

The $\boldsymbol{\alpha}$ which minimises (4.71) can be solved using interior point methods. Similar to the TNIP method used with the InCrowd algorithm described in section 4.4.2, [239] solves (4.71) using Newton's method. The gradient,$\nabla\psi$, and Hessian Matrices, $\nabla^2\psi$, with respect to the $\boldsymbol{\alpha}$ variable are given by [239]:

$$\nabla_{\boldsymbol{\alpha}}\psi = \boldsymbol{\alpha} - \mathbf{y} + \mathbf{A}\mathrm{ST}^+_{\lambda\eta_k}\left(\mathbf{w}_k + \eta_k \mathbf{A}^T\boldsymbol{\alpha}\right) \tag{4.72}$$

$$\nabla^2_{\boldsymbol{\alpha}}\psi = \mathbf{I}_m + \eta_k \mathbf{A}_I \mathbf{A}_I^T \tag{4.73}$$

Where $\mathbf{I}_m$ is the identity matrix of size $m \times m$ and $\mathbf{A}_I$ is the submatrix of the system matrix $\mathbf{A}$ consisting of the columns associated with the non-zero elements of $\mathbf{w}$. Notice that the Hessian and gradient only depend on the active elements. Therefore the complexity of the Newton step is proportional to the sparsity of the solution vector $\mathbf{x}$, not the density of the system matrix $\mathbf{A}$. In this sense, this approach is similar to the computational cost associated with the homotopy method and is unlike the TNIP, which depends on the complexity of the system matrix.

Similar to TNIP [227], it is recommended that a Preconditioned Conjugate Gradient (PCG) method should be used to solve Newton system [239]. The preconditioner used is the Jacobi preconditioner which is constructed using the diagonal elements of the Hessian matrix $\nabla^2_{\boldsymbol{\alpha}}\psi$.

The methodology to solve for the firing time vector $\mathbf{x}$ using the DAL approach can now be explained. Choosing a sequence of increasing barrier terms $\{\eta_k\}$: $\eta_k < \eta_{k+1}$ (this guarantees supra-linear convergence [239]) the $\boldsymbol{\alpha}_k$ which minimises (4.71) is solved using the Newton method and a backtracking line search. The solution to the Newton method is defined when the magnitude of the gradient function $\nabla_{\boldsymbol{\alpha}}\psi$ is less than some threshold ϵ_k The Lagrange multiplier (primal variable) $\mathbf{w}_{k+1}$ is updated using (4.70). The barrier term is increased: $\eta_k \to \eta_{k+1}$ and the Newton step solution threshold is decreased $\epsilon_k \to \epsilon_{k+1}$. This process is

repeated until the stopping criteria is reached. The stopping criterion, the same as for the TNIP described in section 4.4.2, is defined as when the relative duality gap $\Delta_{p,d}/f(\mathbf{x})$ is less than the specified tolerace: τ_{gap}.

The dual feasible point, satisfying (4.62), is chosen in [239] using exactly the same methodology as the TNIP described in section 4.4.2:

$$\hat{\alpha} = \frac{\lambda\left(\mathbf{A}\mathbf{w} - \mathbf{b}\right)}{||\mathbf{A}^T\left(\mathbf{A}\mathbf{w} - \mathbf{b}\right)||_\infty} \tag{4.74}$$

Notice that this feasible point satisfies, by construction, the requirement that $\mathbf{A}^T\tilde{\alpha} \leq \lambda$. The duality gap can be developed using (4.52) & (4.60) given by:

$$\begin{aligned} \Delta_{p,d} &= f(\mathbf{x}) - d(\hat{\boldsymbol{\alpha}}, \mathbf{A}^T\hat{\boldsymbol{\alpha}}) \\ &= \frac{1}{2}|\mathbf{A}\hat{\mathbf{x}} - \mathbf{y}|_2^2 + \lambda\sum_{i=1}^{P} x_i + \frac{1}{2}||\tilde{\boldsymbol{\alpha}} - \mathbf{y}||^2 - \frac{1}{2}\,||\mathbf{y}||^2 \end{aligned} \tag{4.75}$$

DAL pseudo-code

choose sequences $\{\eta_k\}$ such that $\eta_{k+1} > \eta_k$ and $\{\epsilon_k\}$ such that $\epsilon_{k+1} < \epsilon_k$

while $\Delta_{p,d}/f(\mathbf{x}) < \tau_{\text{gap}}$

1. set η_k.
2. solve $\Delta\boldsymbol{\alpha}_{k+1}$ for the direction to minimise (4.71) using Newton's method to tolerance: $|\nabla\psi| \leq \epsilon_{k+1}$.
3. Choose step size γ_{k+1} which maximises the decrease in (4.71) using a backtracking line search.
4. Update $\boldsymbol{\alpha_{k+1}} = \boldsymbol{\alpha_k} + \gamma_{k+1}\Delta\boldsymbol{\alpha}_{k+1}$
5. Update $\mathbf{w}_{k+1}$ using (4.70).
6. Update the gradient $\nabla\psi$ (4.72) & the Hessian matrix $\nabla^2_{\boldsymbol{\alpha}}\psi$ (4.73).
7. calculate $\Delta_{p,d}$.

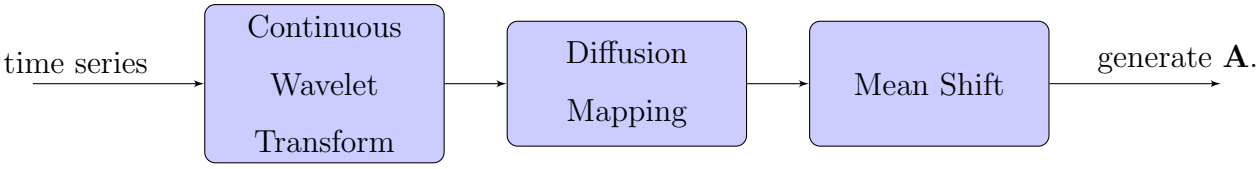
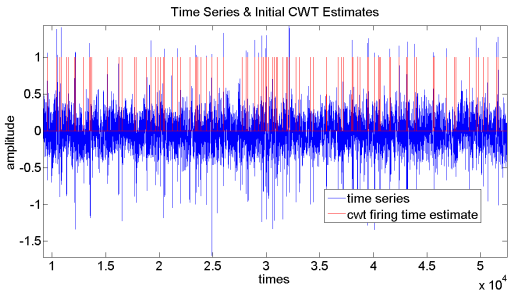


Figure 4.2: The three step process to identify dictionaries and construct the system matrix $\mathbf{A}$ for BPDN. The candidate spike times are estimated using *CWT*, the reduced dimensional feature space is constructed using *Diffusion Maps* which is then clustered using the *Mean Shift* algorithm.

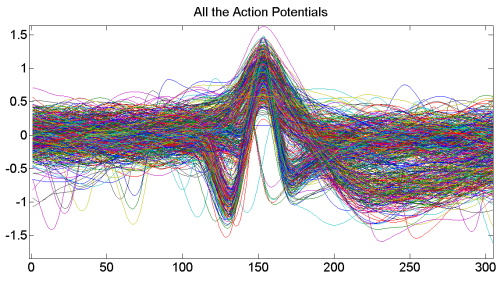
4.5 Developing The Dictionaries

In the previous section we have shown that given an estimate of the set of spike shapes, $\{g_i\}$ we can form the system matrix $\mathbf{A}$ and use BPDN to identify the firing times $\mathbf{x}$. Since the firing time vector $\mathbf{x}$ is a concatenation of the firing time vectors associated with each dictionary term (spike), the BPDN algorithm not only performs accurate spike detection but also accurate and automatic spike clustering. The problem with this approach is that given an extra-cellular recording we *do not know* the spike shapes apriori and therefore do not have the structure of the system matrix $\mathbf{A}$ necessary to perform BPDN. In the proceeding section we will develop a methodology to identify the spike shapes, which will be used as the dictionary set (and thus construction of the system matrix $\mathbf{A}$) for the BPDN algorithm. This approach will involve 3 steps: initial spike time detection using multi-scale continuous wavelet decomposition, dimensionality reduction using Diffusion Mapping and then clustering using the Mean Shift algorithm. This process is shown conceptually in Figure 4.2. The centroids associated with these clusters are then used to estimate the spike shapes $\{g_i\}$. This process is shown on simulated data in Figure 4.3.

We argue that this sequential approach of estimating the spike shapes to generate the dictionary terms (and thus the system matrix $\mathbf{A}$) for use in the BPDN algorithm (4.9) is superior to the sequential approach introduced in [192] of initial spike detection using amplitude thresholding, feature selection using Principal Component Analysis, then clustering using the K-means algorithm and using the centroids of these clusters to drive the CBP

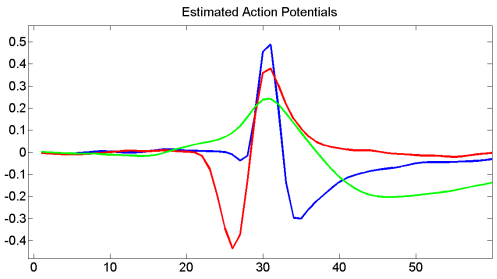


(a) CWT estimate of firing times



(b) Identified action potentials

(c) clustering by diffusion mapping and means shift clustering



(d) estimated action potential shapes (dictionary terms)

Figure 4.3: The three step process of getting the dictionaries. The firing times are estimated by CWT. The candidate spike shapes are extracted from the voltage trace. The dimensionality of the dataframe is reduced using diffusion mapping (3 dimensions are shown). The reduced data is clustered using mean shift and the centroid of the waveforms amongst the same members are used as the dictionaries.

algorithm.

The main problem with this approach is the requirement of the K-means clustering algorithm to know the number of clusters a priori. Any least squares solver (BPDN, CBP or otherwise) cannot be accurate if its dictionary terms (and thus system matrix $\mathbf{A}$) is incorrect. For real extra-cellular recordings the number of clusters is not known apriori and thus frequently the incorrect number of clusters can be specified. If the incorrect number of clusters is specified the K-means clustering will produce incorrect dictionary terms and the spike detection and subsequent clustering will be poor. In effect, the basis pursuit algorithms (BPDN or CBP) are all highly reliant on their preliminary clustering algorithms.

This problem with the K-means clustering algorithm does not occur with the mean shift algorithm (described in section 4.5.3) which can successfully cluster data without a priori knowledge of the number of clusters.

As discussed in section 4.3.1 automated clustering is an incredibly broad area of active research. The K-means clustering (without modification) is a reasonably simple clustering algorithm. Indeed in the family of density based estimators (which mean shift belongs to) there are several viable candidate clustering algorithms which could be considered such as DBSCAN [205], OPTICS [207], DENCLUE [208] and DENCLUE2.0 [209]. This is to say nothing of the other discussed clustering approaches including, but not limited to, Bayesian approaches of Gaussian Mixed Models [213], Student Distribution Mixed Models [214], Infinite Mixed Models [215] or spectral clustering [212], [211].

Secondary problems with the approach of [192] is the use of amplitude thresholding and feature selection based on the principal components. It was shown by extensive simulation in [94] that wavelet analysis was superior to amplitude thresholding techniques to estimate firing times in neural data. Likewise it was shown in [165],[95] that wavelet coefficients and diffusion maps were superior to Principle Component Analysis (PCA) for feature selection in neural data. It was argued in [165] that while the eigenvectors associated with PCA account for the largest variance of the identified waveforms, this does not necessarily provide the best basis for separation of the spike clusters.

It is important to note that this method of developing the dictionaries is itself a rudimentary spike sorting algorithm. The subsequent application of the BPDN procedure boosts the sensitivity and specificity of the spike sorting. In section 4.6.4 we explicitly demonstrate that the additional BPDN step does indeed improve spike sorting estimates over this dictionary building procedure of CWT, diffusion mapping and then mean shift clustering.

In the proceeding section we describe the sub-steps employed to obtain the dictionary terms in detail:

1. The continuous wavelet transform used to identify the spike times.

2. The diffusion mapping procedure used to reduce the dimensionality of the identified spike shapes and perform feature selection.

3. The mean shift algorithm used to group the reduced dimensional spike shapes into their appropriate clusters.

4.5.1 Multi-Scale Continuous Wavelet Transform

The Continuous Wavelet Transform (CWT) is a commonly used method for spike detection [94] and is implemented in the spike sorting package *Osort* [57]. The general idea is as follows. Recall that with the Fourier Transform we represent an arbitary function (subject to certain conditions) as a linear super-position of basis functions of sines and cosines. The idea of the CWT is to decompose an arbitrary function into a linear super position of basis functions called *wavelets*. The basis functions, termed *daughter wavelets* are all generated by dilations and translations of a *mother wavelet* $\Psi(t)$.

$$\Psi_{a,b} = \frac{1}{\sqrt{a}}\Psi\left(\frac{t-b}{a}\right), \qquad \text{a,b} \in \mathbb{R} \tag{4.76}$$

We refer to terms a & b as the dilating & translating parameters respectively. The wavelets must satisfy certain conditions of compact support, zero mean: $\int_{-\infty}^{+\infty}\Psi(t) = 0$ and specific orthogonality conditions [194]. The CWT projets the arbitrary function ($y(t)$) onto the daughter functions ($\Psi_{a,b}$):

$$c(a,b) = \int_{-\infty}^{+\infty} y(t)\Psi\left(\frac{t-b}{a}\right)dt \tag{4.77}$$

The CWT maps from the one dimensional space of the arbitary function y(t), to the two dimensional space of wavelet coefficients $c(a,b)$. The CWT is, for all intents and purposes a measure of the correlation between the function y(t) and the daughter wavelet with dilation and translating parameters a,b chosen from the set $\{a\}$,$\{b\}$. Notice that if we choose a specific dilation parameter value a, then the CWT represents the *cross-correlation* function between $y(t)$ and the wavelet shape at this specific scale. If a mother wavelet is chosen to resemble a spike shape, then at scale a the wavelet coefficients $c(a,b)$ over the set $\{b\}$

identify where the signal correlates with the spike shape with time constant $\approx 1/a$. Thus for a fixed dilation parameter the CWT, over the set of translation parameters, acts like a matched filter [258]. By generating the wavelet transform for multiple dilation and shifting parameters, the cross correlation of the signal, y(t), with wavelets with the same basic shape but different time constants can be constructed. This is referred to in the literature as *multi-resolution analysis.* [194]

The full CWT spike sorting method is described in [94], but an overview of this process is described below. The mother wavelet $\Psi(t)$ is chosen with a shape as similar to the action potentials we are attempting to detect. This is often a wavelet from the bior1.5 family. A set of scale parameters $\{a_i\}$ are selected which are consistent with the expected width of the action potentials. This set may subtly vary with anatomical location or physiological conditions, but is typically in the 0.2-1.5 milliseconds range. The set of translation parameters, $\{b_k\}$ is selected to correspond to every discretised element of the time series. That is, if the time series is T_s seconds long and sampled at $F_s = 1/\Delta_T$ the set of translation parameters will be $[\Delta_T, 2\Delta_T, \cdots T_s]$. Notice that this set of translation parameters is different from the *dyadic* set often used in wavelet analysis [194].

For each dilation parameter, a_i, the wavelet transform over the set of translation parameters is a cross correlation between the measured signal and the daughter wavelet. This set of cross-correlation signals are used to identify the firing times using the following methodology. For each cross correlation function, at each translation point a hypothesis test is performed to estimate whether that point represents a spike (signal + noise) or no spike (just noise). The details are provided in [94], but the hypothesis test is:

$$|c(a,b)| \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \frac{\hat{\mu}_i}{2} + \frac{\hat{\sigma_i^2}}{\hat{\mu}_i} ln\left(L_M \cdot L + ln\left[\frac{P(\mathcal{H}_0)_i}{P(\mathcal{H}_1)_i}\right]\right) \qquad \forall \quad b \in \{b\} \tag{4.78}$$

Where μ_i is the sample mean of the absolute value of the wavelet coefficients at scale a_i under the hypothesis $\mathcal{H}_1$ (that a spike is present) and σ_i is the standard deviation of the wavelet coefficients at scale a_i. The term in the parenthesis $(\cdots)$ sets the costs of false

negatives vs. false positives at the wavelet scale a_i. The terms $P(\mathcal{H}_0)_i/P(\mathcal{H}_1)_i$ is the ratio of the prior probability distributions associated with spike and no spike respectively at scale a_i. This ratio is considered equivalent of the ℓ^0 pseudo-norm of the number of spikes and non spikes respectively. $L_M = 36.7368$ is chosen as the maximum ratio of the cost of false positives to false negatives which does not cause arithmetic overflow under a double-precision floating point representation. The L variable represents a free parameter which controls the cost of false positives and false negatives.

Note that if the statistics of the signal are known, the linearity of the wavelet transform guarantees that the statistics of the wavelets at the different scales can be calculated. The problem is that the statistics of the signal are *not* known, and in the case of the μ_i & $P(\mathcal{H}_0)_i/P(\mathcal{H}_1)_i$ terms we do not know which set of translation coefficients correspond to spikes.

The mean, μ_i, is estimated by partitioning the wavelet coefficients at scale a_i into a spike containing set and a non spike containing set based on a hard thresholding rule (based on a nonlinear wavelet denoising approach used in [259]) such that coefficients greater than the threshold are considered spikes and coefficients less are considered to be noise. The threshold, given by $T_i = \sqrt{2N\hat{\sigma}_i{}^2}$, is based on the estimate of the variance and the number of time samples, N. Similarly the ratio $P(\mathcal{H}_0)/P(\mathcal{H}_1)$ is calculated at scale a_i by the ratio of the number of coefficients greater and less than this threshold respectively. There is a subtly discussed in [94] which explains the different options of how to proceed if no wavelet coefficients are detected above the hard threshold value. Briefly the options are to consider no spikes at this scale a_i (termed the conservative estimate) or to consider that a single wavelet at this scale exceeds this threshold and set $\mu_i = T_i$ (termed the liberal estimate).We explain how the variance at each scale a_i is estimated below.

As previously explained, at sufficiently high sampling rates the spikes are sparsely distributed in the time series. Thus, with respect to estimation of the variance σ, of the background noise, the spikes can be considered outliers. The variance at each wavelet scale

is estimated by assuming that the action of the wavelet transform is to make the signal statistics more white and uncorrelated and then applying the Median Absolute Deviation (MAD) estimator to the set of wavelet coefficients for each dilation parameter value a. Recall that for Gaussian random variables the MAD estimator provides accurate estimates of the variance in the presence of outliers (spikes).

For a given scale parameter at each point where the alternative hypothesis is true, neighbouring points will often also satisfy this hypothesis. This is referred to as *temporal contiguity* [94]. For each contiguous series of points at scale a_i we identify whether this point of an estimated spike is also identified at other scales. This is referred to as *scale contiguity* [94]. If the alternate hypothesis is true in a contiguous region over multiple scales the spike time in that contiguous region is identified as the arithmetic mean of the maximum wavelet coefficient at each scale value. This approach provides an estimate which attempts to deal with the noise sources on the different scales independently jittering the location of the maximal coefficient. If the contiguous region is only present at one scale, a_i, the spike time is estimated at the time of the maximum wavelet coefficient over that contiguous region. This process is repeated for every contiguous region to identify all the spike times. Notice that, as identified in [94], the process of identifying a single spike in each contiguous region prevents the detection of sufficiently close overlapping spikes. This is in contrast to the BPDN approach which can deal with this overlapping spike situation.

CWT spike detection pseudo-code

Choose a range of wavelet scale coefficients $[a_1 \cdots a_n]$ and detection parameter L:

1. Calculate the CWT over the set of all scale coefficients $[a_1 \cdots a_n]$ and translation coefficients $b = [\Delta_T, 2\Delta_T \cdots T_s]$
2. **at each scale a**
   (a) Estimate the standard deviation σ_i using the MAD estimator.

(b) Determine the hard threshold T_i

(c) Calculate the mean μ_j associated with the wavelet coefficients with absolute value greater than the threshold $|c(a,b)| > T_i$.

(d) Calculate the ratio $P(\mathcal{H}_0)_i/P(\mathcal{H}_1)_i$ at scale a_i as the ratio of the number of wavelet coefficients from the set $\{b\}$ with absolute value greater and less than T_j respectively.

(e) Identify the wavelet coefficients which satisy the alternative hypothesis $\mathcal{H}_1$ using (4.78).

3. Identify wavelet coefficients which satisfy the alternative hypothesis over different scales $[a_1 \cdots a_n]$ (satisfy scale contiguity).

4. Estimate the spike time in the temporally contiguous regions as the arithmetic average of the time associated with the maximum wavelet coefficient in this region across the valid scales.

Notice that this CWT approach can often successfully perform the first phase of spike sorting: detecting spikes amongst noise, but it cannot perform the second phase which is assigning the identified spikes into their appropriate clusters. In the next sections we describe how given these estimates of the spiking times, diffusion mapping (section 4.5.2) and the mean shift algorithm (section 4.5.3) can be used to successfully cluster the spikes.

4.5.2 Diffusion Mapping

Diffusion mapping is a non-linear dimensionality reduction technique (closely linked to spectral clustering) which was introduced in [197]. This method was applied to extra-cellular neural recordings where the firing times were known apriori in [95]. We describe the diffusion mapping process below, but refer the reader to [197] for the original (and more mathematically detailed) formulation of the methodology. Given our series of L dimensional datapoints: $\mathbf{x}_i = [x_1, \cdots, x_L]^{\mathbf{T}}$ which describe an identified spike shape, we form a *data frame* $\mathbf{X}$ by stacking these p datapoints: $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \cdots \mathbf{x}_p]$. We view the data points associated with the different spike shapes as nodes of a weighted undirected graph $G = (V, E, W)$ (vertex, edge

and weight). We define the weights according to the heat (Gaussian) kernel and form the associated $p \times p$ dimension Gram matrix as:

$$W_{i,j} = \exp\left(-\frac{||\mathbf{x}_i - \mathbf{x}_i||^2}{\sigma_{DM}^2}\right) \tag{4.79}$$

The σ_{DM} function represents the strength of the scaling between data points in the P dimensional space and is a free parameter which must be tuned by the operator. The use of this kernel shows that we are considering a *fully connected* graph. We see that points which are close in the Euclidean sense are weighted more highly. Notice that this Gaussian kernel is isotropic. Reduced dimensional clustering based on a normalised graph Laplacian of this isotropic kernel was considered in [210]. Indeed spike clustering using this approach was considered in [260]. In [197] a family of anisotropic diffusion processes (parameterised by a factor α) are introduced which specify the amount of influence of the density of the datapoints in the (infinitesimal) transitions of the diffusion process:

$$\mathbf{W}^\alpha = \mathbf{D}^{-\alpha}\mathbf{W}\mathbf{D}^{-\alpha} \tag{4.80}$$

Where $\mathbf{D}$ is the diagonal matrix with diagonal elements: $D_{ii} = \sum_j W_{ij}$. Following [197] we consider $\alpha = 1$, which corresponds to diffusion described by the Beltrami-Laplace operator, and removes the influence of the density of the distribution of the datapoints and recovers the underlying (Riemannian) geometry of the dataset.

From our anisotropically weighted matrix we define the Markov probability transition matrix: $\mathbf{P} = (\mathbf{D}^\alpha_{\text{norm}})^{-1}\mathbf{W}^\alpha$. The diagonal matrix $\mathbf{D}^\alpha_{\text{norm}}$ is defined as: $\mathbf{D}^\alpha_{\text{norm},ii} = \sum_j \mathbf{W}^\alpha_{i,j}$. Thus the matrix $\mathbf{P}$ is a row normalised form of W^α such that each row sums to unity. This $\mathbf{P}$ matrix is the transition matrix of the Markov chain associated with a random walk on the data points. It is shown in [197] that for the choice of $\alpha = 1$ the Markov chain converges to Brownian motion on the sub manifold of $\mathbb{R}^l$. The behaviour of the random walk tells us which points are close (the random walker is likely to transition to) and which points are further apart (the random walker is less likely to transition to). This provides a natural framework for clustering over the dataset.

We determine the eigenvalues & eigenvectors associated with this Markov transition matrix: $\mathbf{P}\psi = \lambda\psi$. The first eigenvalue λ_1 will be unity (to see this consider that for non-negative matrix elements a_{ij}, which are row-normalised to unity, $\mathbf{P}\psi = 1\psi$ will always be a solution) and the magnitude of the remaining eigenvalues will be less than unity: $\lambda_1 = 1 \geq |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_p|$. The set of eigenvectors, scaled by its eigenvalue provides a basis set for the data points $\{\lambda\psi\}$. Because the ordered eigenvalues are all less than unity and have reducing magnitude we may truncate all but the q largest eigenvalues (excluding the unity one). This corresponds to the mapping: $\mathbf{\Psi} : \mathbb{R}^L \to \mathbb{R}^q$.

$$\mathbf{\Psi}(\mathbf{x}_i) \to \begin{bmatrix} \lambda_1\psi_1(i) \\ \lambda_2\psi_2(i) \\ \vdots \\ \lambda_q\psi_q(i) \end{bmatrix} = \begin{bmatrix} \mathbf{q}_1(i) \\ \mathbf{q}_2(i) \\ \vdots \\ \mathbf{q}_q(i) \end{bmatrix} \tag{4.81}$$

Immediately we can see that if $q \leq L$ this method can be used for dimensionality reduction. We will briefly discuss the intuitive advantage of this method. Consider the Diffusion distance which is related to summations over our Markov transition matrix $\mathbf{P}$:

$$\begin{aligned} D_t^2(\mathbf{x}_i, \mathbf{x}_j) &= \sum_z \left(\mathbf{P}_{iz}^t - \mathbf{P}_{jz}^t\right)^2 \cdot d_z^{-1} \\ &= \sum_{\mathbf{z}} \Big(p(\mathbf{z}, t|\mathbf{x}_i) - p(\mathbf{z}, t|\mathbf{x}_j)\, w(\mathbf{z}) \Big) \end{aligned} \tag{4.82}$$

Note that $p(\mathbf{z}, t|\mathbf{x}_i)$ is the probability that the random walker starts at point $\mathbf{z}$ and arrives at point $\mathbf{x}_i$ after t steps. The diffusion distance $D_t(\mathbf{x}_i, \mathbf{x}_j)$ defines the ℓ_2 distance between the probability clouds $\mathbf{x}_i$, $\mathbf{x}_j$ with weight w. As identified in [197] as this distance sums over *all* paths of length t connecting $\mathbf{x}_i$ & $\mathbf{x}_j$ it is robust to noise. The diffusion distance can also be written in terms of the Euclidean distance of the eigenvectors weighted by the eigenvalues. Using the truncated set of eigenvalues and eigenvectors we may write:

$$D_t^2(\mathbf{x}_i, \mathbf{x}_j) = ||\mathbf{\Psi}(\mathbf{x}_i) - \mathbf{\Psi}(\mathbf{x}_j)||^2 + \mathcal{O}(t, q) \tag{4.83}$$

Thus by using the mapping provided by (4.81) we can see that we achieve a dimensional reduction from $\mathbb{R}^l \to \mathbb{R}^q$ which attempts to preserve the diffusion distance between data

points in the original higher dimensional space. It tries to preserve the structure between points which are readily connected by a random walk on the data. Thus at lower dimensions it identifies which points are 'connected', which provides a natural representation for clusters.

Diffusion Mapping Pseudo-Code

1. Form the $\mathbb{R}^{(L\times p)}$ data-frame: $\mathbf{X}$.
2. Construct the isotropic kernel $\mathbf{W}$ using (4.79).
3. Construct the anisotropic kernel $\mathbf{W}^{\alpha}$ (with $\alpha = 1$) using (4.80).
4. Construct the Normalisation Matrix $\mathbf{D}_{\text{norm}}$.
5. Form the Markov transition matrix $\mathbf{P} = (\mathbf{D}^{\alpha}_{\text{norm}})^{-1}\mathbf{W}^{\alpha}$ (with $\alpha = 1$).
6. Generate the first $q+1$ eigenvalues and eigenvectors.
7. Generate the q dimensional Diffusion Mapping using (4.81) and the 2^{nd} to $(q+1)^{th}$ eigenvalues and eigenvectors.

Note that in subsequent section we used $q = 3$ dimensional Diffusion Maps.

4.5.3 Mean Shift

We now describe the non-parametric algorithm used to automatically cluster the waveforms in the reduced dimensional space provided by the diffusion mapping. The mean shift algorithm is a non-parametric clustering technique which identifies modes in the underlying probability distribution describing a dataset. The mean shift is based on Kernel Density Estimation (KDE), but the key to this algorithm is that in order to find these modes the KDE does not need to be explicitly constructed. If we assume that the data points are realisations of a random variable drawn from the multi-dimensional distribution, we can use the mean shift algorithm to identify the cluster shape (mode) that the data points belong to. We will briefly explain this algorithm following the approach of [261]. The KDE, $\hat{f}(\mathbf{x})$

associated with the multivariate data is given by:

$$\hat{f}(\mathbf{x}) = \frac{1}{nh^d}\sum_{i=1}^{n} K\left(\frac{\mathbf{x}-\mathbf{x}_i}{h}\right) \tag{4.84}$$

Where h is referred to as the *bandwidth* of the kernel. Notice the similarity between this bandwidth term and the sigma function σ_{DM} used in the diffusion mapping procedure (4.5.2). Similar to σ_{DM}, this bandwidth represents a free parameter for the spike sorting algorithm which scales the strength of the contribution of nearby data points. Multiple different kernels can be considered (Uniform, Epanechikov, Gaussian) but they must satisfy certain conditions including positive definiteness, zero mean, integrating to unity and vanishing behaviour in the appropriate limits [261]. In the mean shift procedure we only consider isotropic, radially symmetric kernels: $K(\mathbf{x}) = k(||\mathbf{x}||^2)$. This function $k(\cdots)$ is referred to as the *profile* of the kernel. This simplifies the KDE $\hat{f}(\mathbf{x})$ (4.84) to:

$$\hat{f}(\mathbf{x}) = \frac{c_{k,d}}{nh^d}\sum_{i=1}^{n} k\left(\frac{||\mathbf{x}-\mathbf{x}_i||^2}{h}\right) \tag{4.85}$$

Where $c_{k,d}$ is a normalisation constant introduced to guarantee that the PDF integrates to unity. The modes of the PDF, by definition have gradients of zero. Taking the gradient of (4.85), introducing the new profile associated with the derivative of the previous profile: $g(||\mathbf{x}||) = -\frac{dk(||\mathbf{x}||)}{d\mathbf{x}}$ and defining a new kernel associated with this profile $G(\mathbf{x}) = g(||\mathbf{x}||^2)$ yields:

$$\nabla f_{h,K}(\mathbf{x}) = \underbrace{\frac{2c_{k,d}}{nh^{d+2}}\sum_{i=1}^{n} g\left(||\frac{\mathbf{x}-\mathbf{x}_i}{h}||^2\right)}_{\hat{f}_{G,h(\mathbf{x})}} \times \underbrace{\left[\frac{\sum_{i=1}^{n}\mathbf{x}_i g\left(||\frac{\mathbf{x}-\mathbf{x}_i}{h}||^2\right)}{\sum_{i=1}^{n} g\left(||\frac{\mathbf{x}-\mathbf{x}_i}{h}||^2\right)} - \mathbf{x}\right]}_{\mathbf{m}_{h,G}} \tag{4.86}$$

As explained in [261] the first term, $\mathbf{f}_{G,h}(\mathbf{x})$ represents the KDE of $f(\mathbf{x})$ with kernel $G(\mathbf{x})$. The second term, $\mathbf{m}_{h,G}(\mathbf{x})$, is referred to as the *mean shift* vector and represents the difference between the point $\mathbf{x}$ and the weighted mean (as scaled by the kernel $G(\mathbf{x})$). Algebraically re-arranging (4.86) yields:

$$\mathbf{m}_{h,G}(\mathbf{x}) = \frac{h^2 c}{2}\frac{\nabla f_{h,K}(\mathbf{x})}{\hat{f}_{G,h(\mathbf{x})}} \tag{4.87}$$

Therefore we see from (4.87), due to the normalisation by the PDF estimate $\hat{f}_{G,h(\mathbf{x})}$, that the mean shift vector always points in the direction of *increasing* probability density. This implies that if $\mathbf{m}_{h,G}(\mathbf{x})$ equals zero, then the point $\mathbf{x}$ must represent a mode point.

Note that (4.87) shows that the mean shift vector can be used to identify the modes and (4.86) shows how to identify these modes without explicitly constructing the KDE. Given a point $\mathbf{x}$ we may identify the nearby mode in an iterative process by determining the mean shift vector $\mathbf{m}_{h,G}(\mathbf{x})$ associated with this point, then determining the mean shift vector associated with this new point. We continue this process until the magnitude of the mean shift vector is sufficiently small:

$$\begin{aligned} &\text{while} \qquad |\mathbf{m}_{\mathbf{h},\mathbf{G}}(\mathbf{x}^k)|^2 \leq \epsilon \\ &\mathbf{x}^{k+1} \quad = \mathbf{m}_{\mathbf{h},\mathbf{G}}(\mathbf{x}^k) = \frac{\sum_{i=1}^{n} \mathbf{x^k}_i g\left(||\frac{\mathbf{x}^k - \mathbf{x}_i}{h}||^2\right)}{\sum_{i=1}^{n} g\left(||\frac{\mathbf{x^k} - \mathbf{x}_i}{h}||^2\right)} - \mathbf{x}^k \\ &\text{end} \end{aligned} \tag{4.88}$$

Mean Shift Algorithm Pseudo-Code

Choose parameters ϵ, ϵ_{merge} and bandwidth h.

1. For each data point $\mathbf{x}_i$ (which corresponds to a diffusion mapped spike estimate) apply algorithm (4.88). This will (to some high degree of tolerance) map each data point (spike) to the centroid associated with that cluster.

2. Once each point has been clustered, all modes within some specified radius $|\mathbf{x}_i - \mathbf{x}_j|^2 \leq \epsilon_{\text{merge}}$ are merged.

3. Identify the spike shapes associated with the different clusters by taking the arithmetic average (centroid) of the spike shapes (in the original L dimensional space) in each cluster.

We briefly mention three improvements which could be implemented to improve the clustering ability. Notice that the mean shift algorithm is constructed assuming that the bandwidth of the kernel at each data point is constant. The use of *variable* bandwidth estimators can improve estimates when the characteristics of the feature space vary across the space of the data [262]. A procedure for choosing a unique bandwidth for each point, proportional to the density estimate, is described in [262]. The problem with this approach is that it requires an initial estimate of the distribution (termed the *pilot*) using the standard KDE. Unfortunately this defeats one of the primary advantages of using the mean shift algorithm, that it does not require an explicit construction of the KDE. The second problem with using the standard mean shift algorithm for clustering is that it is computationally expensive. It is possible to improve the runtime of the algorithm by approximating the mean shift vector in (4.86) by a form which only considers the data points close to the current point of interest [263]. This strategy is often viable for Gaussian kernels because the function rapidly decays away from point of interest. A third improvement, suggested in [263] is to improve run time by estimating the kernel for the mean shift using random sampling.

Attractive alternatives to mean shift clustering are the previously mentioned DENCLUE 2.0 [209] and the *Gaussian Blurred Mean Shift* [264] algorithms. The Gaussian Blurred Mean Shift simultaneously shifts every single data point towards its cluster centroid. The problem with this approach however is that if the algorithm is run for too many iterations it converges all data points to a single cluster. In [264] a stopping criterion based on information entropy estimates are developed to stop iterating once it has successfully clustered the data sets, but before the cluster structure is lost.

4.6 Results

In this section we apply the BPDN algorithm to both simulated and real data. The outline of these results will be as follows:

section 4.6.1 describes the simulation methodology used to generate the artificial vLFP. This data is used in the subsequent sections to benchmark the accuracy of the BPDN spike sorting algorithm and compare it to other spike sorting approaches.

section 4.6.2 describes how the ROC plots associated with the spike sorting algorithms are constructed. These ROC plots and χ^2 statistics of the estimates are used to quantify the success of the analysed spike sorting algorithms.

section 4.6.3 provides comparisons in ROC space of the three BPDN algorithms described in section 4.4 to see which provides spike sorting estimates with the greatest sensitivity and specificity.

section 4.6.4 compares the sensitivity and specificity of the BPDN spike sorting algorithm to the simpler method based on the preliminary estimate of the firing times using the CWT approach, which is used to develop the dictionary terms (and subsequently the system matrix $\mathbf{A}$) for the BPDN approach.

section 4.6.5 compares the sensitivity, specificity and associated χ^2 estimates of spike sorting using BPDN algorithm and the well established state of the art *wav-clus* spike sorting algorithm.

section 4.6.6 applies the BPDN spike sorting algorithm to real vLFPs collected from MERs inserted into the STN of human patients prior to undergoing deep brain stimulation for Parkinson's Disease.

These steps can be understood in a pedagogical fashion. Firstly if we are going to employ a BPDN approach, we need to identify which BPDN algorithm (homotopy, InCrowd or DAL) should be used. Once we have identified the specific BPDN solver to implement we must show that the spike sorting estimates identified using the dictionary building approach

(CWT, diffusion mapping and the mean shift algorithm) are improved with the additional BPDN step. We then need to show that this BPDN approach provides spike sorting estimates which are superior (or at least comparable) to other spike sorting programs (we specifically compare against *wav-clus*). Once this BPDN approach has been validated and compared on synthetic data (where we know the ground truth of when spikes occur) we can more confidently apply it to real extra-cellular recordings where the spike times are not known a priori.

4.6.1 Simulation Data

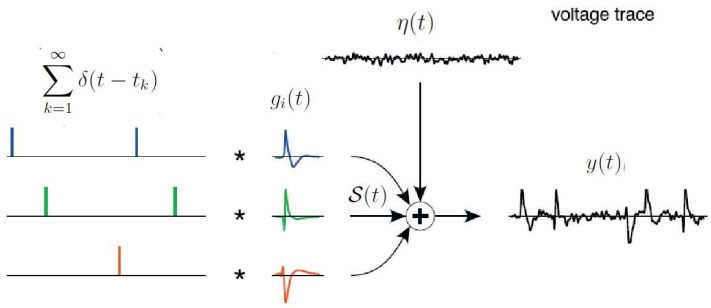


Figure 4.4: Schematic representation of the generation of the simulated extra-cellular vLFP. The MUA term $\mathcal{S}(t)$ is generated by superposition of convolving the spike times $\{t_k\}$ with the spike shapes $g(t)$. The voltage trace is then given by the superposition of the MUA and neural noise term: $y(t) = \mathcal{S}(t) + \eta(t)$. Figure adapted from [192].

In this section we describe the methodology used to generate the simulated extra-cellular vLFPs which are used to asses the accuracy of the spike sorting algorithms. The use of simulated data is necessary because, with the exception of rare circumstances where concurrent intra-cellular and extra-cellular recordings are obtained [163], there is no "ground truth" for the spiking times of the real data.

The basic idea of the simulated data is to generate a time series which consists of the super-position of two processes. This concept of two components of the signal: the resolvable spikes of nearby neurons making up the MUA, $\mathcal{S}(t)$, and the non-resolvable neural noise, $\eta(t)$, was discussed in section 4.3.2. The first process, the MUA, is the set of spike trains of

the neurons with distinct spike shapes which are closer to the recording probe. We consider this part of the signal as containing the detectable and separable spikes. The second process is the neural noise which consists of every biological process not associated with the spiking of the MUA neurons. As discussed in section 4.3.2 these neural noise terms include (but are not limited to) the spiking behaviour of the more distant neurons, synaptic currents, subthreshold oscillations and after potentials. We consider this part of the signal as containing no detectable spikes, but impeding our ability to identify and cluster the detectable spikes making up the MUA. In principle a perfect spike sorting algorithm will ignore the contribution of the neural noise, identify all the detectable spikes contributing to the MUA and correctly cluster them into their appropriate groups.

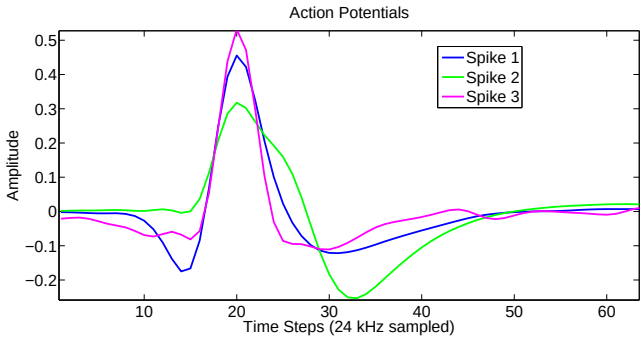


Figure 4.5: The three detectable spike shapes used in the simulations.

The simulated vLFP data is generated using a modification to the publicly available simulation datasets of single channel extra-cellular recordings taken from the Sub Thalamic Nuclei of human patients undergoing Deep Brain Stimulation developed in [204]. The spike trains associated with each of the MUA, $\mathcal{S}(t)$, was constructed using the following procedure:

1. The spike shapes were manually obtained from the extra-cellular recordings from different portions of the STN of patients undergoing DBS. These spike templates were scaled such that the maximal amplitude was unity. The three spike shapes considered are shown in Figure 4.5. Notice that the structure of the depolarisation-polarisation phase of the action potential is very similar between spikes one & three. Therefore we expect that separating noisy samples of these two spikes to be more difficult than spike

two, which has a much broader characteristic depolarisation-polarisation phase.

2. The ISI times, $t_k - t_{k-1}$, for the spike train of each neuron was a random sample of a Weibull distribution with randomly selected rate and shape parameters. Specifically the shape parameter, $\tilde{k}$, was drawn from the bounded continuous uniform distribution: $\tilde{k} \sim \mathcal{U}(1,4)$ giving firing patterns that range from Poisson (k=1) to pseudo-periodic (k=4). The mean firing rate $\tilde{\nu}$ was also drawn from the bounded continuous uniform distributions $\tilde{\nu} \sim \mathcal{U}(10,35)$. Thus the firing rates of the spike trains ranged from 10 to 35 Hz. The scale parameter $\tilde{\lambda}$ was chosen such that the randomly sampled shape and firing rates could both be satisfied: $\tilde{\lambda} = 1/(\tilde{\nu}\Gamma(1+1/\tilde{k}))$. Thus the ISI time are drawn from the distribution $(t_k - t_{k-1}) \sim p(t)_{\text{ISI}}$ such that:

$$p(t)_{\text{ISI}} = \tilde{k}t^{\tilde{k}-1}\left(\tilde{\nu}\Gamma\left(1+\frac{1}{\tilde{k}}\right)\right)^{\tilde{k}} e^{-t\tilde{\nu}\Gamma\left(1+\frac{1}{\tilde{k}}\right)^{\tilde{k}}} \tag{4.89}$$

3. This process was then repeated for all of the detectable neurons and the resulting spike trains were linearly superposed. The resulting time series were then scaled such that the power spectrum integrated to unity: $\mathcal{S}(t) = \mathcal{S}(t)/\sqrt{\int_{-\infty}^{\infty} P_{\mathcal{S}(t)}(\omega)d\omega}$, where $P_{\mathcal{S}(t)}(\omega)$ is the power spectrum of the unscaled $\mathcal{S}(t)$ process.

The neural noise, $\eta(t)$, part of the signal was generated from the real STN recordings in regions where no spikes were observed to occur. This is to mimic both the behaviour of the more distant neurons and the previously mentioned non-spiking neural processes (such as synaptic currents, subthreshold oscillations and after potentials) which will contribute to the recording probe time series. This time series is then scaled to an appropriate Noise to Signal Ratio (NSR) as follows: $\eta(t) = \eta(t) \cdot \sqrt{\frac{\text{NSR}}{\int_{-\infty}^{\infty} P_{\eta(t)}(\omega)d\omega}}$, where $P_{\eta(t)}(\omega)$ is the power spectrum of the unscaled neural noise, $\eta(t)$, process. Obviously the greater this neural noise level the more difficult it is to detect and sort the spikes associated with the detectable set. In the proceeding sections the NSR parameter is varied to identify how the different spike sorting algorithms operate in different noise conditions. Three NSR levels of 0.05, 0.2, 0.45 corresponding to low, medium and high neural-noise respectively are considered in simulations in Sections 4.6.3-4.6.5. The power spectra of a representative sample of the MUA ,$\mathcal{S}(t)$, and

the neural noise, $\eta(t)$, at these three NSR levels is shown in Figure 4.6.

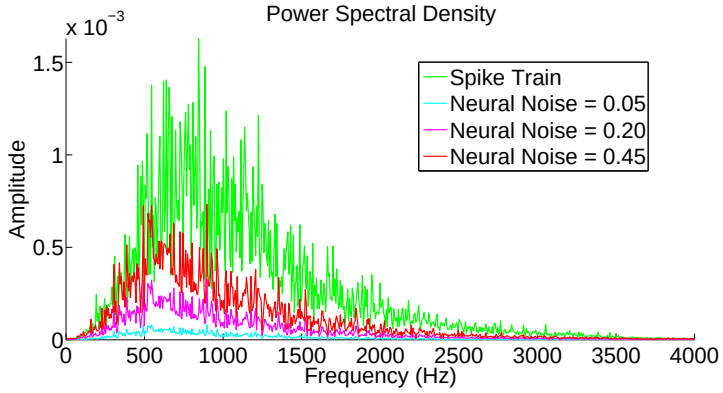


Figure 4.6: Power Spectral Density of the detectable spike train and the non-detectable neural noise for the three NSR levels of 0.05, 0.2 & 0.45. Note the neural noise component is a coloured noise process with a similar spectrum to the detectable spike train.

The MUA, $\mathcal{S}(t)$, (detectable) and neural noise,$\eta(t)$, (non-detectable) parts of the signal are linearly superposed and white Gaussian noise is added to the signal to model thermal and signal aquisition effects. For all simulations this white noise process is scaled similarly to the neural noise, with the constant value of $\text{NSR}_{WN} = 0.1$. This process is shown in Figure 4.4. Note that in the subsequent sections we limit the performance analysis of the spike sorting algorithms to a system with the three detectable spike shapes (neurons). This process of analysing the ability and accuracy of the spike sorting algorithm to identify and cluster only three distinct action potential shapes was also performed in [192] and [165] to characterise the performance of *CBP* and the *wav-clus* algorithms respectively.

4.6.2 ROC Curve Criteria

In this section we describe how Receiver Operating Characteristic (ROC) plots are used to quantify the success of the different spike sorting algorithms. The ROC plot provides a convenient means of measuring and displaying the ability of a procedure to perform binary classification tasks. This approach was originally developed in radar theory [265] but has since found extensive use in the medical field [266]. The ROC scatter-plot displays the False

Positive Rate (FPR) on the abscissa and the True Positive Rate (TPR) on the ordinate. The success of the binary classifier can be assessed by identifying where in the ROC space it resides. We would ideally like to construct classifiers which have high TPR and low FPR estimates. These quantities of FPR and TPR are defined below and can be related to other commonly used statistical metrics:

$$TPR = \frac{TP}{C_P} = \text{Sensitivity} = 1 - \beta \tag{4.90}$$

$$FPR = \frac{FP}{C_N} = 1 - \text{Specificity} = \alpha \tag{4.91}$$

Where C_P and C_N refer to the number of positive events (spikes) and negative events (no spike). TP refers to the number of true positives where the detector has classified the state as positive when the state truly is positive (i.e. identified a spike when there is a spike) and FP refers to the number of false positives where the detector has classified the state as positive when the state truly is negative (i.e. identifying a spike when there is no spike). Notice that by construction both TPR and FPR values are restricted to between zero and unity and therefore can be considered the *probabilities* of true positive and false positive classification respectively. α is the level of significance of a statistical test and represents the probability of a false positive. $(1 - \beta)$ is referred to as the power of a statistical test and represents the probability of a true positive. Notice that for a binary classifier we may also define false negatives (FN) where an event is classed as negative when the state is truly positive (i.e. missing a spike when there is a spike) and true negatives (TN) where an event is classified as a negative when the state is truly negative (i.e. identifying no spike when there is no spike).

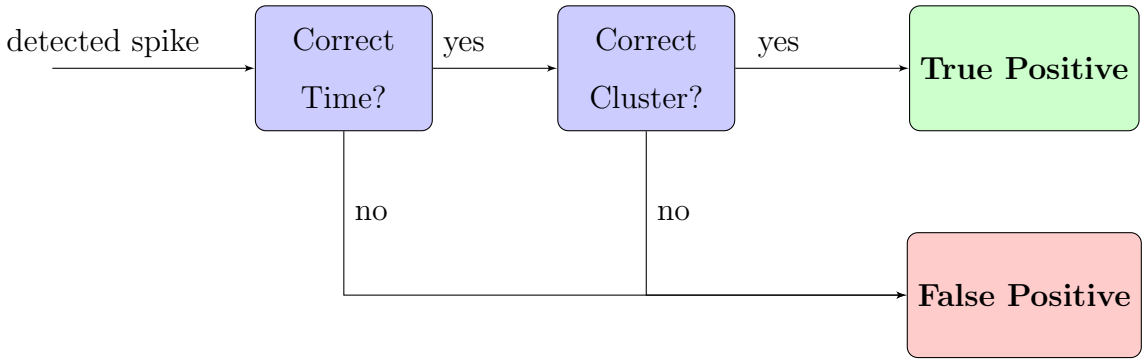


Figure 4.7: To define a true positive the correct *time* and *cluster* must be identified.

Note that the diagonal line extending from (0,0) to (1,1) is the locus of points where $TPR = FPR$ (or equivalently $\alpha = 1 - \beta$). This line defines the set of points associated with a random classification scheme which doesn't use any of the input (i.e. signal) information. The point (0,0) corresponds to a point where the probability of defining a random negative state is unity, $P_n = 1$, (each state is considered a negative so no FPs will be identified, but equally no TP will be identified). The point (1,1) corresponds to a point where the probability of defining a random positive state is unity, $P_p = 1$ (each state is considered a positive so all TPs will be identified, but equally every possible FP will also occur). The points between these extremes represent the variation in the probabilities of randomly labelling positive and negative states such that $P_n + P_p = 1$. Obviously any successful classifier which uses information about the measured system to classify states should reside in a point above this random classification ($\alpha = 1 - \beta$) line.

Quantifying the success of a spike sorting operation is difficult for a multitude of reasons. The primary reason is that spike sorting is *not* a binary classification problem. When spikes are detected they are assigned (or classified) into one of many clusters. One approach to deal with this would be to have a separate ROC plot for each of the clusters but this would be impractical for a number of reasons. Firstly the number of identified clusters can change between recordings and secondly spikes can vary with which cluster they are assigned to as the threshold parameter is varied. Lastly this is an impractical method to compare different spike sorting algorithms. For example it is unclear how you would quantify the success of two separate spike sorters when one algorithm classifies one cluster accurately and another cluster poorly whereas the second algorithm classifies the first cluster poorly but the second cluster accurately. Another approach to dealing with the multi classifier nature of spike sorting would be to develop a multi-class ROC hypersurface with two axes, TPR and FPR, per cluster. The main problem with this approach is the difficulty of displaying the data in such high dimensions.

The approach we introduce is to consider a true positive classification of the spike sorter

when the program not only correctly detects the time of a spike but also correctly classifies the spike into the appropriate cluster. This is effectively a method of mapping the multi-class ROC hypersurface to a binary ROC plot with only positive and negative classification. Thus in this framework if the spike is correctly detected but incorrectly clustered it is considered a false positive. This requirement for a true positive is shown in Figure 4.7. In constructing these ROC plots there are subtleties with both identifying whether the spike sorter has detected the spike at the correct time and assigned it to the correct cluster which we discuss below.

The timing of the spike is defined as occurring in a particular time bin of the discretised lattice. If the spike time is detected in a bin (or bins) adjacent to the defined spike time there is an issue of whether this should be considered a correct or incorrect detection. We define a correct spike detection if it is within $\pm N$ bins of the true spike time. For subsequent simulations we set $N = 10$. Note that for time series sampled at 24 kHz this corresponds to spike detection which is accurate to within 0.4 milliseconds. For comparison, the main depolarisation-polarisation lobes of the spikes considered (shown in figure 4.5) have widths of roughly 0.625 milliseconds. See [267] for an excellent review of the action potential shapes of neurons in different regions of the mammalian central nervous system.

The second issue is how we define whether the identified spike is classified into the correct or incorrect cluster. Associated with this is the issue of how we link the clusters identified from the spike sorting program to the true clusters defined in the synthetic data. We introduce the following two definitions: the identified clusters are the collections of sets of spike times for the clusters *identified* by the spike sorting program, the true clusters are the collections of sets of spike times for the clusters *defined* from the simulation data. Note that the number of positive, C_P, and negative, C_N, events in (4.90) & (4.91) are defined from the number of spike and non spike events of the simulated data respectively.

We link the identified and true clusters together by pairing the identified and true clusters that have the greatest agreement in their set of spike times. Specifically for each permutation

of identified and true clusters we calculate the TPR between their sets of spike times. We then pair the clusters in an iterative greedy approach: linking the identified and true pair with the highest TPR, removing these clusters from the linking process and then finding the next pair (of the remaining clusters) with the greatest TPR. Notice that if there is more identified clusters than true clusters, the spike times associated with these unmatched clusters in the subsequent ROC analysis are automatically considered false positives. Likewise if there are less identified clusters than true clusters, all of the spike times at these unidentified true clusters are considered false negatives.

It is important to note that this process of detecting spikes and then classifying them is distinct from the classical ROC approach of linearly varying a threshold parameter and considering elements above this value an event and elements below this value not an event. In the case of the BPDN algorithm, variation of the regularisation parameter λ has a nonlinear effect [227] on the number of spikes detected and their location. The amplitude thresholding, A_{thresh}, for spike detection in *wav-clus* and the L factor (re-parameterisation of the ratio of false positives to false negatives) in (4.78) for the CWT are similar to the classical thresholding in ROC analysis, but the subsequent clustering procedures will introduce nonlinear variation. In effect as you vary λ or A_{thresh} you construct a new classifier which can result in either the TPR and FPR increasing or decreasing as these threshold parameters are varied. Therefore the variation of these parameter will not necessarily trace out a convex hull in the ROC space. Thus there is no simple way, by variation of a single parameter to vary the misclassification cost between positive and negative classification. Nonetheless the closest surrogates to a single parameter are the regularisation parameter, λ for BPDN and the amplitude threshold level, A_{thresh} for *wav-clus*.

A second important consideration is that the collapse of the multi-class classifier (i.e classifying to multiple different clusters) to a single positive/negative detection condition will restrict the points of any spike sorting algorithm to a subset of the ROC space. For example since any clustering algorithm will only assign each spike to a single cluster it is highly unlikely that any choice of parameter for any spike sorting algorithm will result in

a point in the ROC space of 100 percent true positive rate (every spike not only detected but correctly clustered). Due to these conditions we display the data in scatter form in the ROC space rather than with the traditional (and convex) ROC curves. Specifically the ROC scatter plots show the operating points of the different spike sorters (our BPDN and the wav-clus SPC algorithm) as the regularisation λ and amplitude thresholding parameters A_n respectively are varied.

χ^2 Estimates of the Classifiers

The ROC plots quantify the behaviour of the classifier by the two quantities of TPR and FPR (or equivalently sensitivity and specificity) as a parameter is varied. These two quantities provide complimentary but independent descriptions of how successful the candidate classification scheme is. In order to explicitly compare the classification ability between the different spike sorting algorithms we must characterise this classifier by a single value. Unfortunately any attempt to reduce the information from the collection of points on the ROC space to a single value will lose information about the pattern of positive and negative classifications. The most popular summary statistic is the Area Under the Curve (AUC) [266]. Due to the points in the ROC space not defining a convex set and the different spike sorting algorithms being restricted to different subsets of this space, this statistic is not appropriate.

The single value metric we use to characterise the success of the different spike sorting algorithm is the *chi squared statistic*: χ^2. The χ^2 measures the Euclidean distance between the observed values in the TP, TN, FP and FN classes and and the expected values of these classes based on the random classification scheme:

$$\chi^2 = \sum_c \frac{(O-E)^2}{E} \tag{4.92}$$

The expression for the χ^2 value for this binary classification system was developed in [268]:

$$\chi^2 = \left(\mathbb{E}\left(T_P\right) - T_P\right)^2 \left(\frac{1}{\mathbb{E}\left(T_P\right)} + \frac{1}{\mathbb{E}\left(F_P\right)} + \frac{1}{\mathbb{E}\left(T_N\right)} + \frac{1}{\mathbb{E}\left(F_N\right)}\right) \tag{4.93}$$

Where:

$$\mathbb{E}(T_P) = \frac{C_P[C_P(1-\beta)+C_N\alpha]}{C_N+C_P} \tag{4.94}$$

$$\mathbb{E}(F_P) = \frac{C_N[C_P(1-\beta)+C_N\alpha]}{C_N+C_P} \tag{4.95}$$

$$\mathbb{E}(F_N) = \frac{C_P[C_N(1-\alpha)+C_P\beta]}{C_N+C_P} \tag{4.96}$$

$$\mathbb{E}(T_N) = \frac{C_N[C_N(1-\alpha)+C_P\beta]}{C_N+C_P} \tag{4.97}$$

Thus the χ^2 statistic can be generated for each spike sorting estimate by using (4.90)-(4.91) to convert the TPR and FPR values into equivalent α and β scores which are inserted into (4.94)-(4.97) to generate the expectation values which, along with the TP count are substituted into (4.93). It is interesting to note that, as identified in [268] the χ^2 value is not only a function of the distance from the random classification (TPR=FPR) line, but also where in the (TPR,FPR) space the classifier resides.

Following [268] we consider the χ^2 value as a measure of how much 'work' the classifier is doing, i.e. exactly how different it is to a random classification system. A primary advantage of using the χ^2 statistic is that this metric is robust to wide discrepancies of the number of positive and negative states. For example it is shown in [268] that if the *accuracy* metric is determined for a system with $C_p = 5$ and $C_n = 55$, always labelling a state as negative can lead to accuracy estimates as high as 92% whereas the χ^2 value remains low. For the MER recordings we are considering the time series are sampled at $F_s = 24kHz$ whereas the spike rates of the individual neurons are less than $100Hz$, thus their will be a wide discrepancy between the number of spike and non spike states.

For each ROC plot considered in the subsequent sections we overlay χ^2 contour values in order to describe the success of the different spike sorters in terms of a single value. In figure 4.8 we demonstrate how to interpret which classifiers do more or less 'work' based on the relative value of the χ^2 contour they lie on. Note from (4.94)-(4.97) that the χ^2 estimate depends on the number of true, C_P, and negative, C_N states. Since this will vary between simulations, for each ROC plot we generate the χ^2 contours based on the average of the C_P

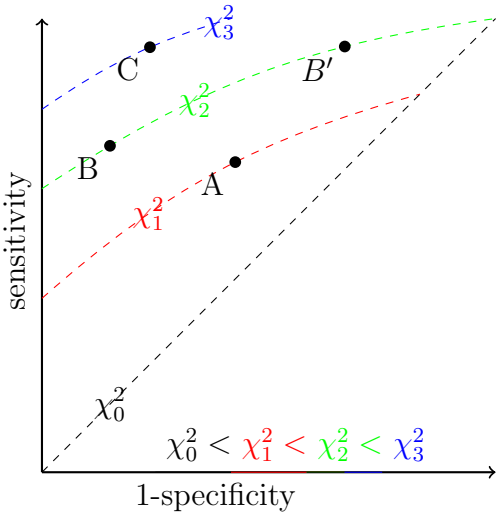


Figure 4.8: Interpreting the 'work' done by a classifier based on its χ^2 value. The χ_0^2 contour represents the classifier based purely on chance ($\alpha = 1 - \beta$). The classifier associated with point C does the most 'work'. The classifier associated with point A does the least 'work'. The classifiers associated with points B and B' do the same amount of 'work' because even though they lie on different points in ROC space they are on the same χ_2^2 contour. In this χ^2 framework we would refer to classifier C as the best classifier, point A as the worst classifier and points B & B' equally good classifiers.

and C_N values over the simulations used to generate the ROC plot.

4.6.3 Comparison of BPDN strategies

In this section we apply the BPDN approach for the three methods: homotopy, InCrowd and DAL discussed in Sections 4.4.1, 4.4.2 and 4.4.3 respectively. This provides a means of deciding which BDPN algorithm to implement for subsequent comparison against other spike sorting solvers (the *wav-clus* algorithm) and application to real data. In this section, in order to minimise the number of variables we consider the dictionary terms known a priori. Thus the system matrix $\mathbf{A}$ is already constructed. This identifies the limits of discriminability we can expect using the BPDN algorithm because we expect this algorithm to be most accurate when the dictionary terms (and system matrix, $\mathbf{A}$) are perfectly known a priori. Figures 4.9-4.11 show the ROC plots with overlaying χ^2 contours for the low (NSR=0.05), medium (NSR=0.2) and high (NSR=0.45) neural noise levels. For each of these three noise levels, we generate ten Monte Carlo simulations and solve the BPDN problem for each of the three algorithms over a range of regularisation parameters, λ.

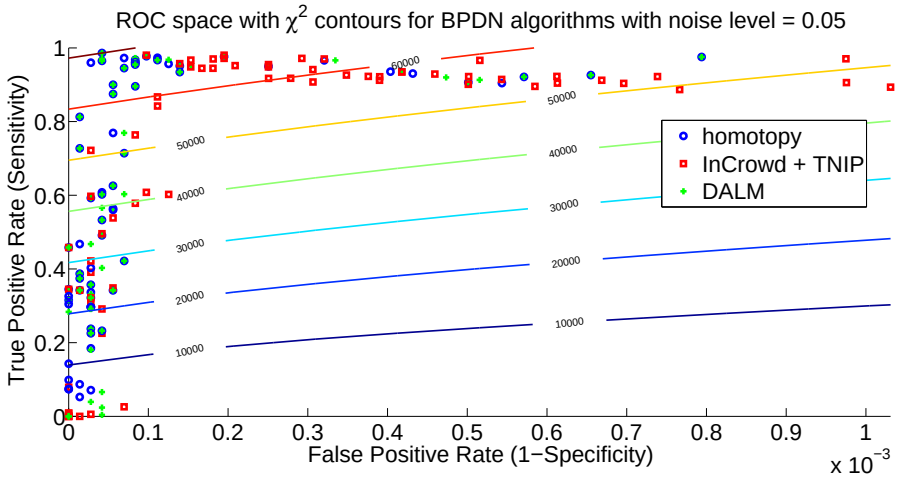


Figure 4.9: ROC plots with χ^2 contours for different BPDN solvers: homotopy (blue circles), InCrowd with TNIP (red squares) and DALM (green crosses) with neural noise = 0.05

It can be seen for Figures 4.9-4.11 that for all noise levels considered, both the homotopy and DALM methods provide estimates with equal or higher sensitivity and specificity (and also χ^2 estimates) than the InCrowd method with Truncated Newton Interior Point. At the low noise (neural noise =0.05) level it can be seen in Figure 4.9 that the three methods provide comparable results, with similar numbers of estimates of all BPDN strategies within

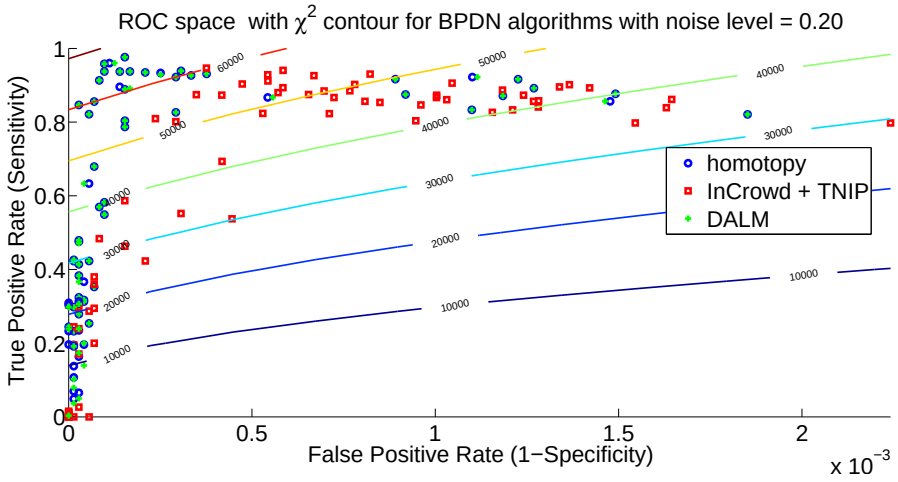


Figure 4.10: ROC plots with χ^2 contours for different BPDN solvers: homotopy (blue circles), InCrowd with TNIP (red squares) and DALM (green crosses) with neural noise = 0.2

the highest χ^2 value contour. At the higher noise levels (neural noise = 0.20 and 0.45) it can be seen in Figures 4.10-4.11 that the homotopy and DALM provide superior estimates, with higher sensitivity, specificity and markedly more estimates within the highest χ^2 value contours. It is unclear whether the decreased performance of the InCrowd+TNIP method at higher noise levels is due to the InCrowd algorithm or the TNIP step of the algorithm. The deficiency is unlikely to be due to the TNIP step, as the DALM method uses this same step in the dual space. Nonetheless it may be of interest to explore the success of the InCrowd algorithm with other ℓ_1 solvers.

For the noise levels considered, due to the similarity of the results, it is difficult to assess whether the homotopy or DALM algorithm is superior. There is much overlap between the results for both these methods. For the remainder of this chapter (sections 4.6.4-4.6.6) the BPDN problem will be solved using the DALM method and the system matrix,$\mathbf{A}$ will *not* be known apriori.

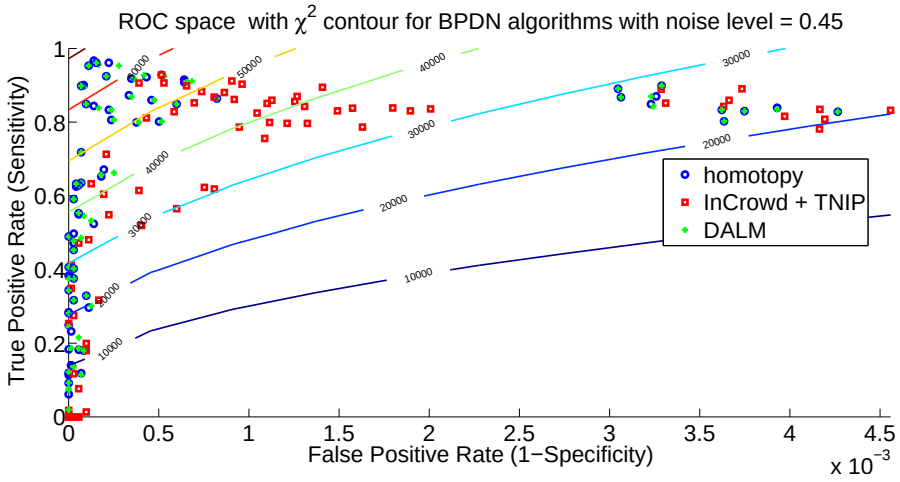


Figure 4.11: ROC plots with χ^2 contours for different BPDN solvers: homotopy (blue circles), InCrowd with TNIP (red squares) and DALM (green crosses) with neural noise = 0.45

4.6.4 Comparison of CWT and BPDN

Recall that the dictionary terms necessary to construct the system matrix $\mathbf{A}$ for the BPDN solution requires preliminary estimates of the spike times using the CWT and then the spike shapes associated with these times are clustered using the Diffusion Mapping and Mean Shift procedure. This approach was described in Section 4.5. A simple spike sorting algorithm can be constructed by partitioning the firing times estimated with the CWT into the appropriate clusters determined using the diffusion mapping and means shift procedure. We refer to this algorithm as *wavelet clustering*.

The additional step of using these estimated dictionary terms and solving the spike sorting problem using the BPDN algorithm introduces computational complexity. Thus in order to justify this complexity we must show that the inclusion of this BPDN step actually improves the spike sorting estimates. In this section we explore and quantify this improvement by comparing the sensitivity and specificity (and χ^2 estimates) of both the wavelet and BPDN clustering algorithms under the low, medium and high (NSR = 0.05,0.2,0.45) neural noise conditions. These ROC plots with associated χ^2 contours are shown in Figures 4.13-4.15.

For both algorithms ten Monte-Carlo simulations using the method described in Section

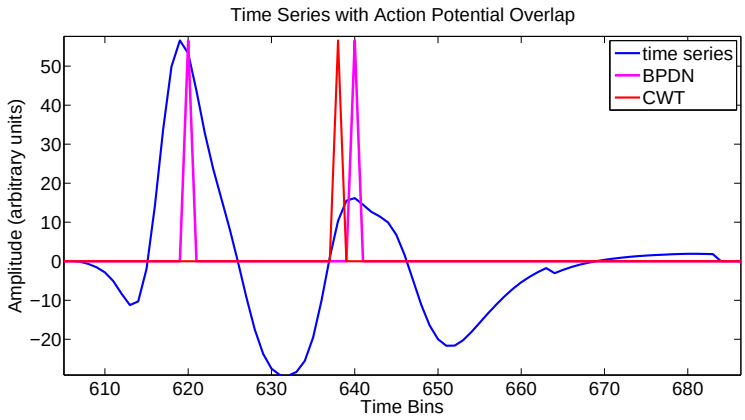


Figure 4.12: Spike detection using BPDN and wavelet clustering (CWT in legend). There are two overlapping spikes centred at time bins 620 and 640. The CWT detects a single spike at bin 638 whereas the BPDN detects the seperate spikes at bins 620 and 640. Time Series sampled at 24kHz.

4.6.1 were generated and subsequently clustered and the estimates were plotted in the ROC space. The CWT clustering was solved with the free parameter L given in (4.78) set to the default value of zero. The BPDN problem is solved (with the DALM algorithm) with the system matrix constructed from the wavelet clustering step. The BPDN problem is solved for regularisation values, λ, of $\{5, 10, 12.5, 15, 17.5, 20, 22.5\}$.

Inspection of Figures 4.13-4.15 show that the additional step of BPDN clustering does indeed boost the sensitivity and specificity compared with the CWT clustered estimates. Furthermore it can be seen that the BPDN estimates are associated with consistently higher χ^2 values. One obvious reason for the superiority of the BPDN is the ability to resolve spiking times of overlapping spikes, which the CWT approach cannot perform. Recall from Section 4.5.1 that it was discussed that due to the temporal and scale contiguity of wavelet transforms, the regions with nonzero wavelet coefficients across time and different scales are collapsed to a single point representing a spike time. Thus if two different spikes occur near simultaneously they will have maximum amplitude coefficients at different scales but likely over the same contiguous region of translation coefficients (time). Therefore these two spikes will be likely be detected but merged by the CWT as a single spike occurring at a time given

by the average of the maximum wavelet coefficient at each observed scale. This effect of the CWT algorithm observing a single spike but the BPDN algorithm correctly identifying the two overlapping spikes for simulated data is shown in Figure 4.12.

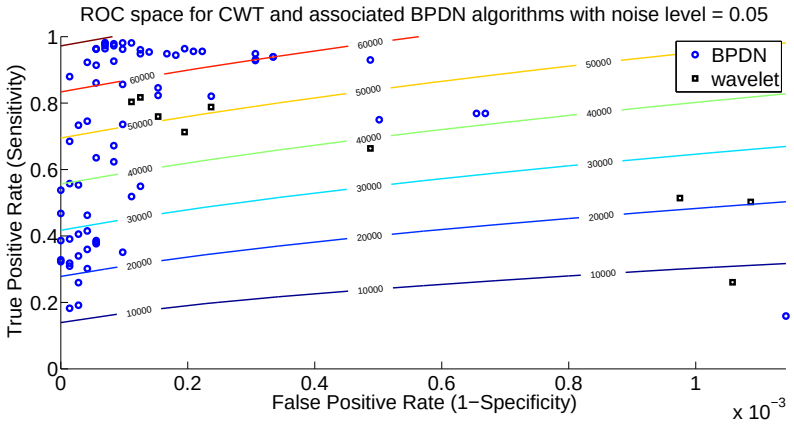


Figure 4.13: ROC plots for BPDN (blue circles) and CWT (black square) spike sorting estimates for SNR = 0.05

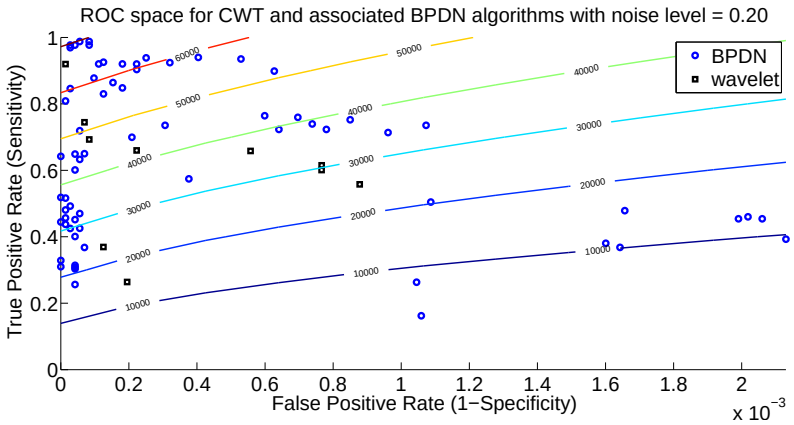


Figure 4.14: ROC plots for BPDN (blue circles) and CWT (black square) spike sorting estimates for SNR = 0.2

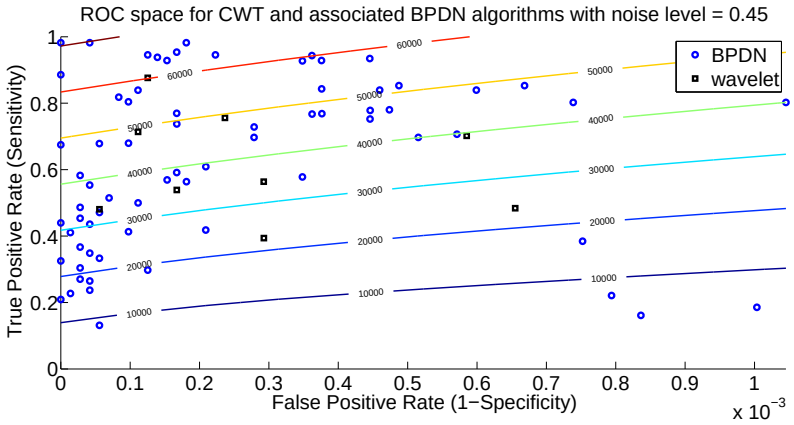


Figure 4.15: ROC plots for BPDN (blue circles) and CWT (black square) spike sorting estimates for SNR = 0.45

4.6.5 Comparison of Integrated Approach against Super Paramagnetic Clustering

In this section we compare the accuracy of the BPDN spike sorting algorithm to the state of the art method *wav-clus*. We compare our BPDN algorithm against the *wav-clus* algorithm for two main reasons. Firstly *wav-clus* is widely employed in scientific studies involving the simulation of extra-cellular recordings and subsequent analysis using spike sorting techniques [96],[269]. Secondly the accuracy of *wav-clus* has consistently been benchmarked against other spike sorting programs [204], [192]. Indeed in [204] explicit comparisons showed that the *wav-clus* solver was more accurate than both the alternative state of the art spike sorting algorithms *Osort* and *KlustaKwik*.

The BPDN component of the BPDN spike sorting algorithm is solved using the DALM method. The *wav-clus* algorithm was discussed in section 4.3.1 but we review how this method works here. Similar to the BPDN approach *wav-clus* serially solves the problems of spike detection, feature selection and clustering. The spikes are detected by considering any point in the time series which is greater than the threshold, $A_{\text{thresh}} \cdot \sigma_n$, where A_{thresh} is any number greater than zero and σ_n is an estimate of the standard deviation of the background

noise which is constructed using the Median Absolute Deviation estimator discussed in section 4.5.1. Obviously the larger values of A_{thresh} will detect very few spikes but very few false positives, whereas very small values will detect many true spikes but also many false positives. The feature selection is based on performing a Haar wavelet decomposition on the identified spike shapes and then selecting the ten wavelet coefficients with the greatest degree of multi-modality (to best separate the spike shapes in this reduced dimensional space) determined by a Kolmogorov-Smirnov test. The wavelet coefficients are then clustered using the Super-Paramagnetic Clustering algorithm discussed in section 4.3.1. We briefly mention that this clustering algorithm depends on a free parameter referred to as the temperature: T, similar to the bandwidth selection parameter for the Mean Shift algorithm, h. When this temperature value is very small all the data is partitioned into a single cluster, whereas when this temperature value is very large each data point belongs to its own group. Between this range is a 'super-paramagnetic regime' where the appropriate number of clusters should be generated.

Figures 4.16-4.18 present the results of the BPDN and the *wav-clus* spike sorting algorithms in ROC space for ten Monte Carlo simulations at each neural noise levels of 0.05, 0.2 and 0.45. For the BPDN algorithm only the regularisation parameter λ was varied. For the *wav-clus* algorithm only the extraction threshold parameter A_{thresh} was varied. Note that the clustering temperature of the SPC algorithm in *wav-clus* is determined automatically by varying the temperature in a range (T=0 to T=0.301 in steps of 0.01) and identifying the highest temperature in this range where a large membership (> 20 elements) was created. The temperature for the subsequent clustering operations is set as this identified temperature. The rationale for this approach is that as the temperature is varied in this range the clustering will sweep through the paramagnetic then super-paramagnetic then ferromagnetic regimes. In both the paramagnetic and ferromagnetic phases clusters with minimal membership will develop, whereas in the super-paramagnetic phase large membership clusters will form [165]. Table 4.1 shows the parameter values used for the BPDN and *wav-clus* algorithms during these simulations.

BPDN		*wav-clus*	
wavelet type:	bior 1.5	wavelet type:	Haar
wavelet L:	0	temperature:	automatic detection
kernel scaling σ:	1	threshold A_{thresh}:	{ 3,3.5,4,4.5,5,5.5,6 }
bandwidth h:	$1 * 10^{-4}$		
regularisation λ:	{ 5,10,12.5,15,17.5,20,22.5 }		

Table 4.1: Parameter values for BPDN and *wav-clus* analysis of simulated data.

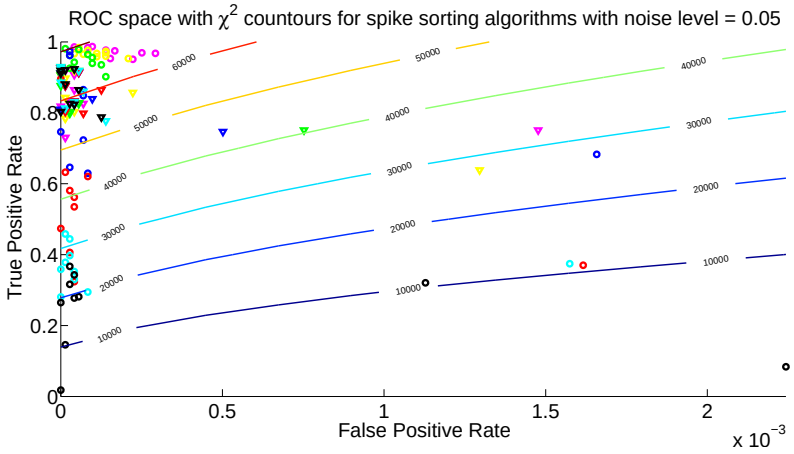
Inspection of Figures 4.16-4.17 shows that for the three NSR regimes (neural noise = 0.05, 0.2 and 0.45) the BPDN spike sorting algorithm consistently performs superior spike classification than the *wav-clus* algorithm. This superiority of the BPDN spike sorter is most pronounced for the high NSR signals. We discuss the behavior of both these algorithms over the three noise regimes below.

It can be seen from Figure 4.16 that for the low NSR (neural noise = 0.05) that the χ^2 values of the estimates of the BPDN algorithm are superior to the *wav-clus* algorithms. It can be seen from Figure 4.16c that across the range of regularisation and amplitude threshold values (for BPDN and *wav-clus* respectively) that the BPDN estimates appear to have greater sensitivity whereas the *wav-clus* estimates have slightly greater specificity. Analysis of figures 4.16a-4.16b shows that for this (neural noise = 0.05) noise environment the optimal (in terms of the highest χ^2 values) regularisation and amplitude threshold parameter values are $\lambda = 12$ (green circles) and $A_{\text{thresh}} = 4.5$ (blue triangles). The optimal A_{thresh} values are difficult to appreciate in these ROC plots because of the close overlap of the different A_{thresh} points.

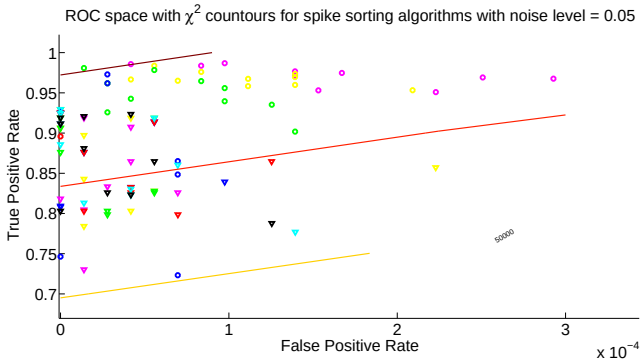
Analysis of Figure 4.17 shows that for the medium NSR (neural noise = 0.20) level more of the BPDN spike sorting classifications have higher χ^2 values than the *wav-clus* estimates.

Inspection of Figures 4.17c clearly shows that, for the range of regularisation and amplitude threshold parameters considered, the BPDN algorithm generates a 'hull' of higher sensitivity and specificity (and therefore higher χ^2 values) points around the *wav-clus* estimates. Analysis of Figures 4.17a-4.17b shows that for this (neural noise = 0.20) noise environment the optimal (in terms of the highest χ^2 values) regularisation and amplitude threshold parameter values are $\lambda = 12$ (green circles) and $A_{\text{thresh}} = 6$ (black triangles).

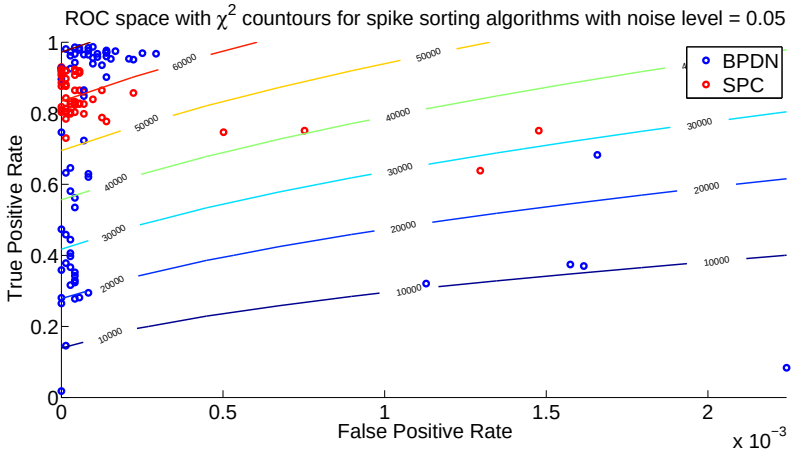
Inspection of Figure 4.18 shows that for the high NSR (neural noise = 0.45) environment the BPDN spike sorting classifications consistently have higher sensitivity, specificity and χ^2 values than the *wav-clus* estimates. The superiority of the BPDN solver over the *wav-clus* algorithm at this high NSR is particularly evident from observation of Figure 4.18b. Analysis of Figures 4.18a shows that for this (neural noise = 0.45) noise environment the optimal (in terms of the highest χ^2 values) regularisation and amplitude threshold parameter values are $\lambda = 12$ (green circles) and $A_{\text{thresh}} = 5$ (red triangles).



(a) ROC plot

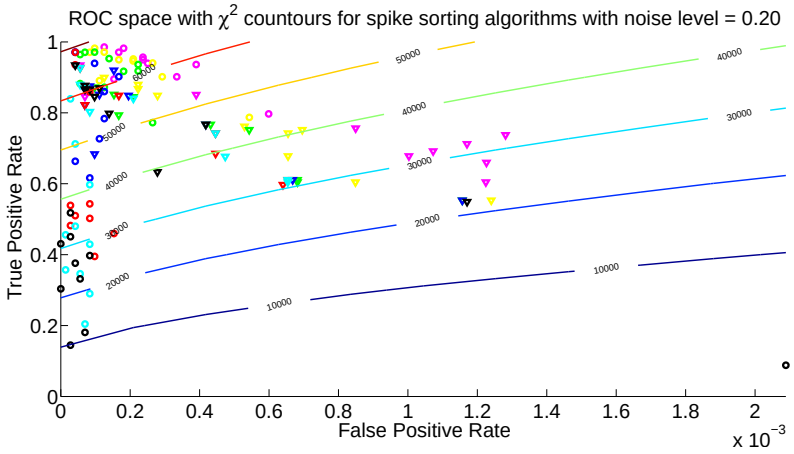


(b) ROC plot zoomed

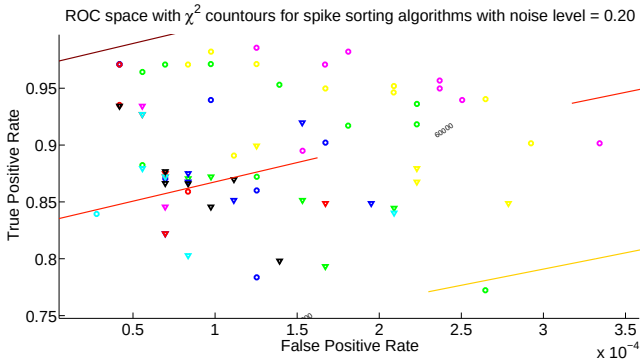


(c) single colour ROC plot

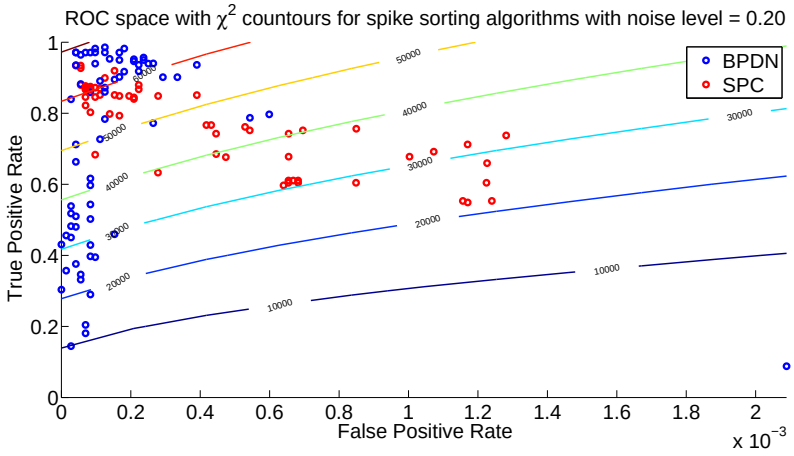
Figure 4.16: ROC plots with χ^2 contours for BPDN (circles) and SPC (triangle) spike sorting estimates for neural noise = 0.05. The colour ordering is { magenta, yellow, green, blue, red, cyan, black } for regularisation parameters, λ, of { 5,10,12.5,15,17.5,20,22.5 } and amplitude thresholding values of { 3,3.5,4,4.5,5,5.5,6 } respectively. Note that there is some overlap of data points corresponding to different tuning parameters (λ or A_{thresh}) for the same algorithm (BPDN or SPC) which makes some data points difficult to identify.



(a) ROC plot

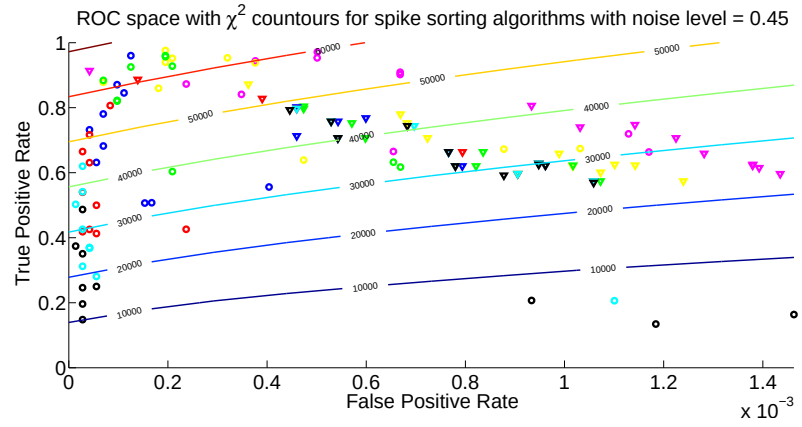


(b) ROC plot zoomed

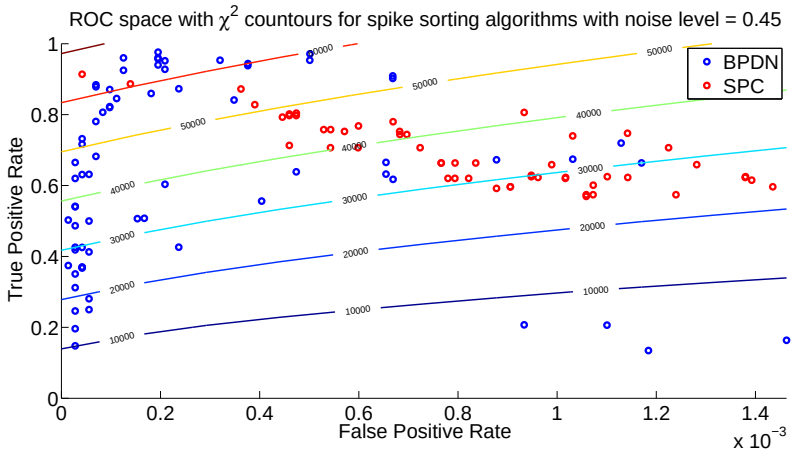


(c) single colour ROC plot

Figure 4.17: ROC plots with χ^2 contours for BPDN (circles) and SPC (triangle) spike sorting estimates for neural noise = 0.20. The colour ordering is { magenta, yellow, green, blue, red, cyan, black } for regularisation parameters, λ, of { 5,10,12.5,15,17.5,20,22.5 } and amplitude thresholding values of { 3,3.5,4,4.5,5,5.5,6 } respectively. Note that there is some overlap of data points corresponding to different tuning parameters (λ or A_{thresh}) for the same algorithm (BPDN or SPC) which makes some data points difficult to identify.



(a) ROC plot

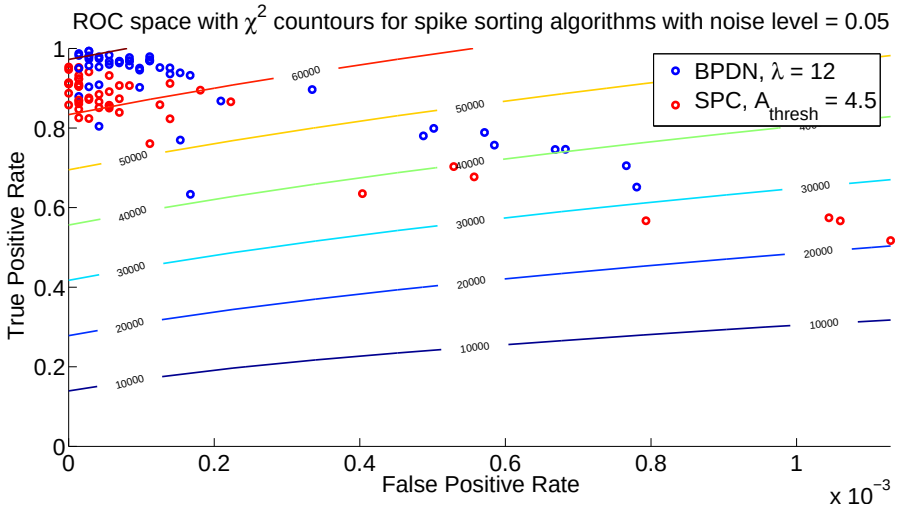


(b) single colour ROC plot

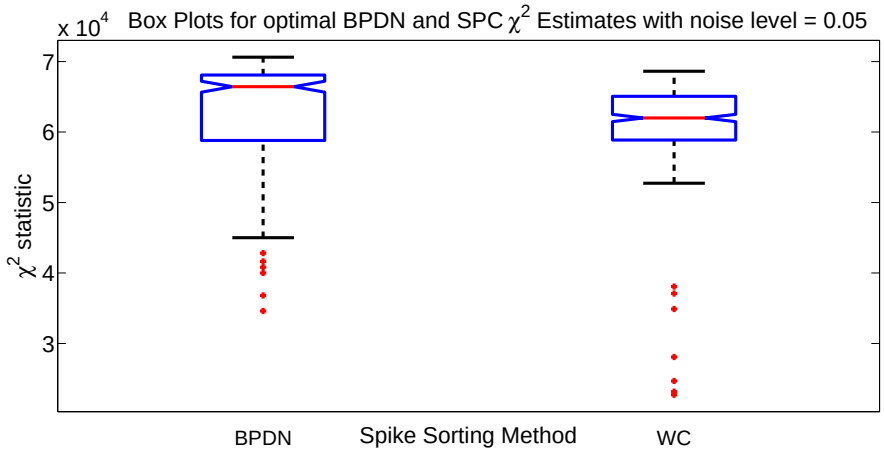
Figure 4.18: ROC plots with χ^2 contours for BPDN (circles) and SPC (triangle) spike sorting estimates for neural noise = 0.45. The colour ordering is { magenta, yellow, green, blue, red, cyan, black } for regularisation parameters, λ, of { 5,10,12.5,15,17.5,20,22.5 } and amplitude thresholding values of { 3,3.5,4,4.5,5,5.5,6 } respectively. Note that there is some overlap of data points corresponding to different tuning parameters (λ or A_{thresh}) for the same algorithm (BPDN or SPC) which makes some data points difficult to identify.

Comparison of Integrated Approach against Super Paramagnetic Clustering with Optimal Tuning Parameters

In order to further compare the spike sorting ability of the BPDN solver and *wav-clus* for each noise level (NSR= 0.05,0.2,0.45) we compare the regularisation parameter, λ, and amplitude threshold A_{thresh} which yield the estimates with the greatest chi-squared values identified in the previous section. For each of the three noise levels considered we generate fifty simulations with the optimal regularisation and amplitude thresholding parameters. The χ^2 statistic associate with each estimate of the BPDN and the *wav-clus* is then displayed on a box plot. These results are shown below in Figures 4.19-4.21. For each noise level the Wilcoxon Rank Sum test is performed to identify that the population means between these two spike sorting algorithms are indeed different. For each of the three noise levels considered, at the $\alpha = 0.01$ level of significance, statistically significant differences in the means of the χ^2 values of the BPDN and *wav-clus* algorithms was identified, with the BPDN mean consistently higher. Based on these results we can conclude that the BPDN algorithm does provides superior spike sorting estimates to the *wav-clus* algorithm across a dynamic range of noise environments.

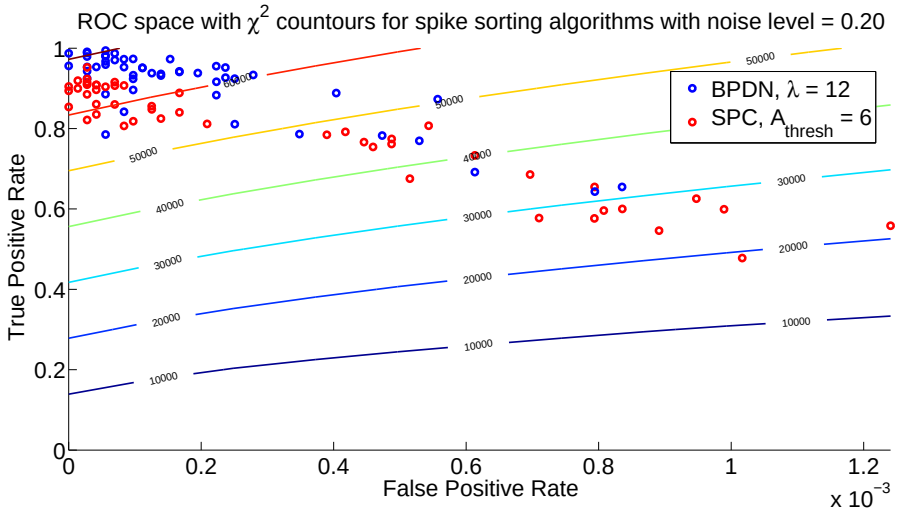


(a) ROC plot

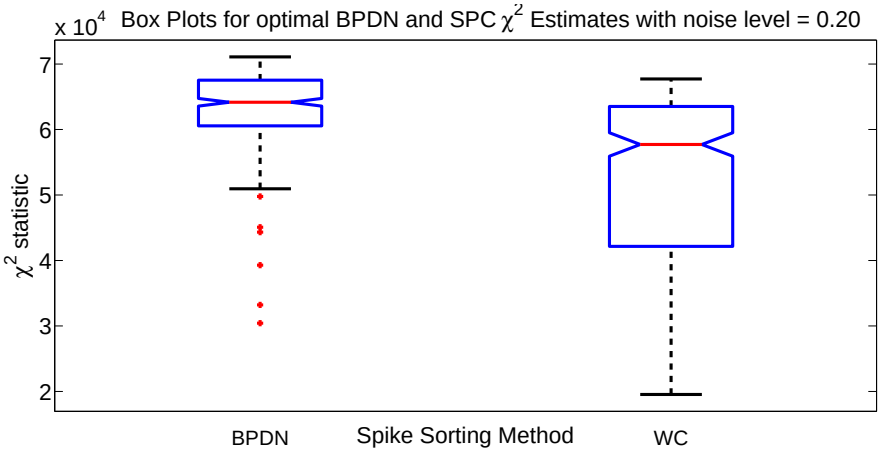


(b) ROC plot

Figure 4.19: Box plots for χ^2 statistic for BPDN $\lambda = 12$ and SPC $A_{\text{thresh}} = 4.5$ at noise level $= 0.05$. Wilcoxon Rank Sum test p-value between populations: < 0.01.

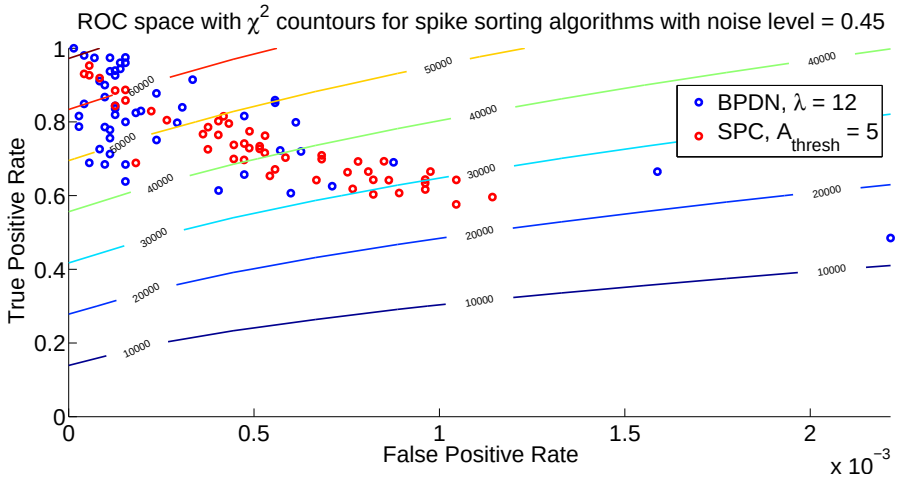


(a) ROC plot

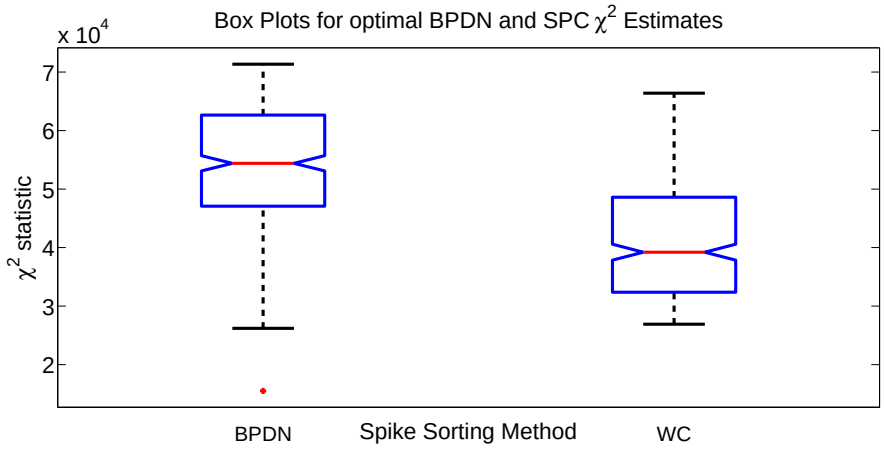


(b) ROC plot

Figure 4.20: Box plots for χ^2 statistic for BPDN $\lambda = 12$ and SPC $A_{\text{thresh}} = 6$ at noise level = 0.2. Wilcoxon Rank Sum test p-value between populations: < 0.01.



(a) ROC plot



(b) ROC plot

Figure 4.21: Box plots for χ^2 statistic for BPDN $\lambda = 12$ and SPC $A_{\text{thresh}} = 5$ at noise level = 0.45. Wilcoxon Rank Sum test p-value between populations: < 0.01.

4.6.6 Application of BPDN to Real Data

In this section we apply the BPDN algorithm to vLFPs of extracellular MERs from the left and right STN of nine human patients with Parkinson's Disease (PD) undergoing DBS surgery. The study was approved by The University of Queensland Medical Research Ethics Committee and UnitingCare Health Human Research Ethics Committee. The signals were recorded from the neural tissue at a sample rate of 24 kHz. Three Butterworth filters, as recommended by the manufacturer were applied to the signal (high-pass: 500 Hz first order, low-pass: 5k Hz first order and anti-aliasing: 5 kHz fourth order). For each recording the spike times were identified and clustered and then the mean firing rate, ν, and Coefficient of Variation (CoV) $C_v = \sigma\nu$ were calculated for each cluster. These estimated mean firing rates and CoV were then averaged into *grand* mean and CoV estimates. Thus for each recording a mean firing rate and CoV was estimated. The parameters of the BPDN algorithm are shown in Table 4.2. In Table 4.4 the mean firing rate and CoV (both sides, left and right side) are shown for each individual patient analysed.

Figures 4.22 and 4.23 show the Kernel Density Estimates (using the KDE method described in [186]) and box plots of the mean firing rates and CoV of 218 recordings across all nine patients for both the left and right side STN recordings. Notice that inspection of Figures 4.22a-4.23a shows that the firing rates are distributed around a mode of approximately 20 Hz and the CoV is centred around a mode of approximately unity. This suggests that *on average* the neurons in the STN of PD patients fire with Poisson statistics at roughly 20Hz. Analysis of the box plot of firing rates displayed in Figure 4.22b shows that the mean firing rate is 44.84 Hz and the median firing rate is 36.22 Hz ($19.65 - 56.36$ Hz, 25^{th}-75^{th} percentile). Similarly analysis of the box plot of CoV displayed in Figure 4.23b shows that the mean CoV is 1.47 and the median CoV is 1.14 ($0.94 - 1.5$, 25^{th}-75^{th} percentile). We discuss the validity of these results in the context of previous studies below. It is highly likely that these estimates of mean firing rate and CoV are sensitive to the higher valued outliers whereas the median, being a more robust statistic [270], is less biased.

BPDN parameter values	
wavelet type:	bior 1.5
wavelet L:	0
kernel scaling σ:	3.5
bandwidth h:	1×10^{-3}
regularisation λ:	0.5

Table 4.2: Parameter values for BPDN analysis of real data.

These results are consistent with the results of previous studies of the firing patterns of STN in humans with PD. For example in [170] the firing patterns in the STN from 351 recordings was averaged over bursting and non-bursting neurons to yield a mean firing rate 40.5 ± 20.3 Hz with a CoV 1.327 ± 0.52. Similarly [28] investigated the firing patterns of the STN in human patients with early and advanced PD. The results showed that in the advanced group (77 recordings) the mean firing rate was 36.3 Hz (25.8-48.5 Hz, 25^{th}-75^{th} percentile) with CoV 1.1 (0.9-1.4, 25^{th}-75^{th} percentile). Similarly the results showed that in the early group (113 recordings) the mean firing rate was 28.7 Hz (19.7-38.7 Hz, 25^{th}-75^{th} percentile) with CoV 1.2 (1.0-1.5, 25^{th}-75^{th} percentile). We display the comparisons in firing rates and CoV between the advanced and early PD patients studied in [28] in Table 4.3. Note that the patients analysed in our study would most likely be more similar to the advanced PD patients in this study as they were undergoing DBS surgery. It is especially interesting to note in Table 4.3 the similarity in these firing rates and CoV between the advanced PD patients studied in [28] and our results.

To further investigate these results we now split the data of all the patients into sub sets associated with the left and right hand side recordings. This partitioning of the data serves two purposes. Firstly it allows us to explore whether there are differences between the firing patterns on the left and right hand side STN of the patients analysed. Secondly this splitting of the data ensures that the previously identified physiological firing rates and

	Early PD patients [28]	**Advanced PD [28]**	**BPDN measure-ments**
mean FR (Hz)	28.7	36.3	44.84
median FR (Hz)	-	-	36.22
$(25^{th} - 75^{th})$ percentile FR (Hz)	(19.7-38.7)	(25.8-48.5)	(19.65-56.36)
mean CoV	1.2	1.1	1.47
median CoV	-	-	1.14
$(25^{th} - 75^{th})$ percentile CoV	(1-1.5)	(0.9-1.4)	(0.94-1.5)

Table 4.3: Comparison of firing rate and CoV for BPDN and previous studies [28] of early and advanced PD patients.

pattern are not a consequence of averaging over a large number of both very large and very small un-physiological firing rates and patterns. Figures 4.24 and 4.25 show that both the mean firing rates and the CoV are similar for the left and right hand side recordings, with both sides firing with approximately Poisson ($C_v = 1$) statistics. It is interesting to note in Figure 4.24 that the KDE of the left hand side recordings may suggest a bimodal distribution with one peak at 20 Hz (similar to the right hand side recordings) and then a higher peak at roughly 50 Hz which is not reproduced in the right hand side recordings.

4.6.7 Summary of Results

In this section we have applied the BPDN algorithm to both simulated vLFP data (where the ground truth of spike times is known) and experimentally obtained vLFPs from human STN of Parkinson's disease patients during DBS surgery (where the ground truth of spike times is unknown).

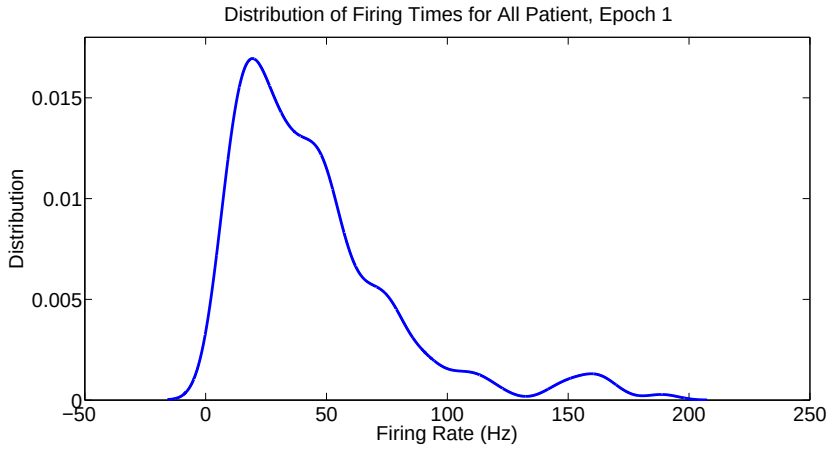
	Both Sides		Left		Right	
Patient	Rate (Hz)	CoV	Rate (Hz)	CoV	Rate (Hz)	CoV
1	35.76	1.05	22.34	1.11	53.01	0.98
2	32.31	1.13	26.64	1.20	39.71	1.05
3	83.34	2.46	105.00	2.79	37.00	1.75
4	59.42	2.47	68.07	3.34	43.69	0.90
5	15.01	1.06	12.33	1.14	15.91	1.03
6	55.38	1.15	55.04	1.10	55.84	1.23
7	31.70	1.30	32.54	1.31	11.40	1.21
8	33.90	1.18	18.93	1.16	59.56	1.23
9	38.21	1.12	48.30	1.06	15.16	1.26

Table 4.4: Mean firing rates & Coefficient of Variation for individual patients.

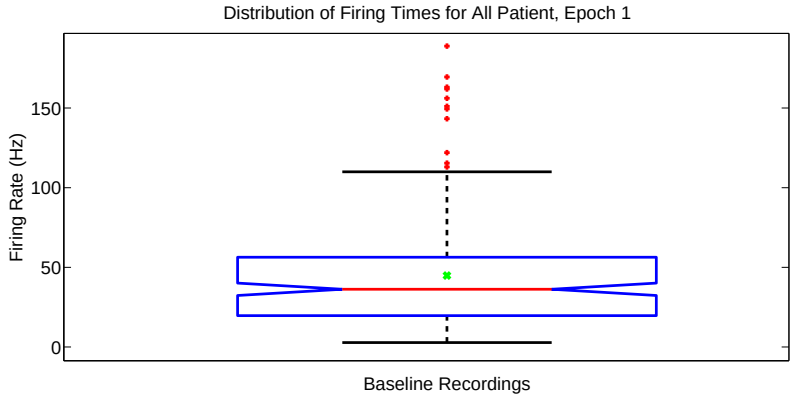
For the simulated data we have quantified the success of the spike sorting with the aid of ROC scatter plots and identifying the χ^2 statistic of the estimators compared to a binary classifier based purely on chance. We have shown that for the BPDN component of the BPDN algorithm the homotopy and DALM methods give comparative, superior results whereas the InCrowd method (with the ℓ_1 minimisation step solved with TNIP) gave worse results. We have shown that including the additional step of BPDN to the preliminary spike sorter consisting of Continuous Wavelet Transform, Diffusion Mapping and Mean Shift clustering (the method to determine the dictionaries described in section 4.5) significantly boosts the sensitivity and specificity (and the χ^2 statistic) of the spike sorting estimates. We suggested that this is due to the ability of the BPDN algorithm to correctly identify highly overlapped spikes. We then showed that this BPDN algorithm provides superior spike sorting estimates to the state of the art spike sorter *wav-clus* which is based on amplitude thresholding and Super Paramagnetic Clustering.

Lastly we have applied the BPDN spike sorter to recordings taken from human Sub

Thalamic Nuclei. The BPDN algorithm identified that, on average, the detected neurons fire with physiologically plausible Poisson statistics with firing rates between 20-56 Hz. These estimates are in agreement with previous studies which have identified the firing patterns of neurons in the STN of patients with Parkinson's Disease.

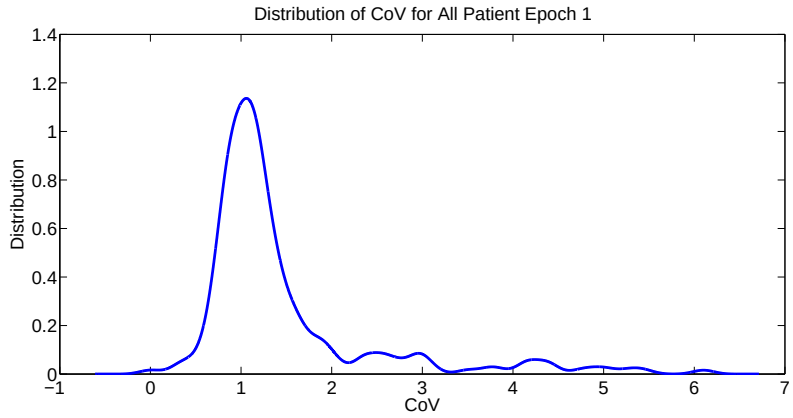


(a) Kernel Density Estimate

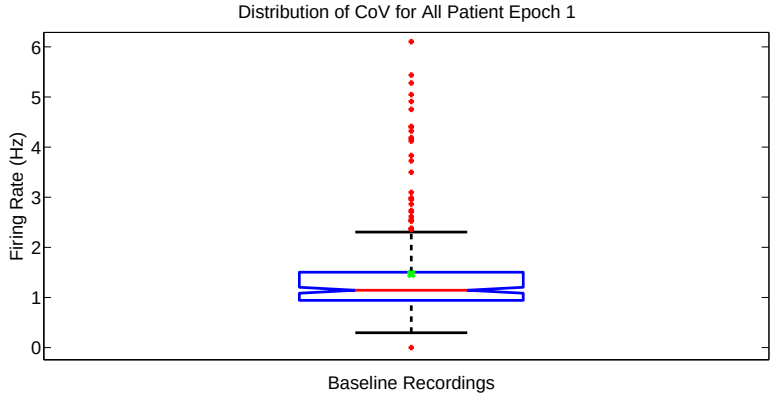


(b) Box Plot

Figure 4.22: Kernel Density Estimates and Box Plots of the firings rates determined by BPDN algorithm to all patients. The mean firing rate is 44.84 Hz and the median firing rate is 36.22 (19.65 – 56.36 Hz, 25^{th}-75^{th} percentile) Hz. It is likely that the mean value is sensitive to the higher frequency outliers.



(a) Kernel Density Estimate



(b) Box Plot

Figure 4.23: Kernel Density Estimates and Box Plots of the Coefficient of Variation determined by BPDN algorithm applied to all patients. Notice that the mode of the distribution is approximately unity, corresponding to Poisson firing statistics. The mean CoV is 1.47 Hz and the median CoV is 1.14 ($0.94 - 1.5$, 25^{th}-75^{th} percentile) Hz. It is likely that the mean value is sensitive to the higher CoV outliers.

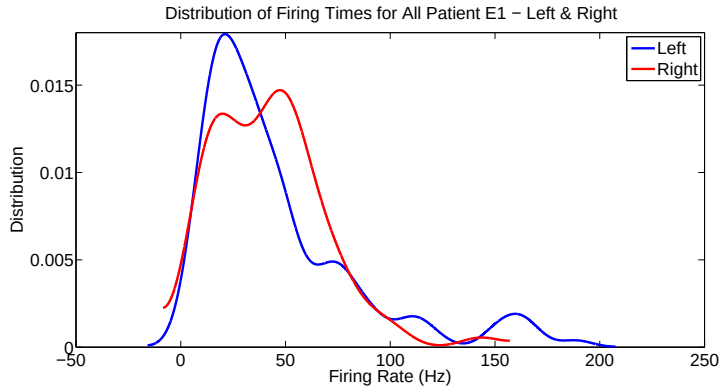


Figure 4.24: Kernel density estimates of the firing rate determined by BPDN algorithm applied to all patients in epoch 1 with left (blue) and right hand (red) data sets considered seperately.

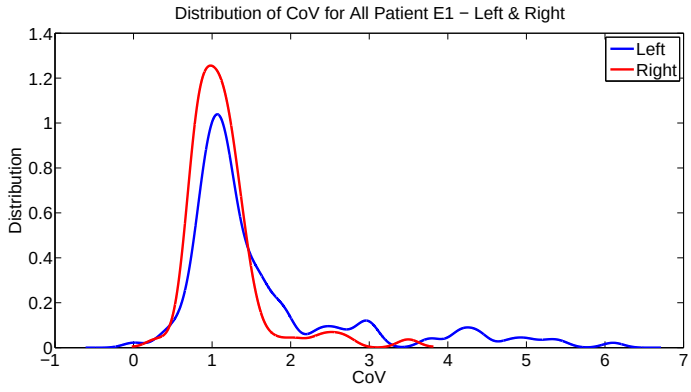


Figure 4.25: Kernel density estimates of the Coefficient of Variation determined by BPDN algorithm applied to all patients in epoch 1 with left (blue) and right (red) hand data sets considered seperately.

4.7 Conclusion & Contributions

We have proposed a BPDN spike sorting algorithm which we have shown, through extensive simulation, is capable of outperforming other state of the art spike sorting algorithms such as *wav-clus* in high noise environments. We have then applied this algorithm to data obtained from extra-cellular, single channel micro electrode recordings of the STN of patients with Parkinson's Disease undergoing DBS. We have identified that the neurons in the STN of these patients fire, on average with Poisson statistics at $(\sim 20 - 56)$Hz, consistent with previous experimental analysis.

The a priori assumptions required for this algorithm are minimal. The primary heuristic introduced is that the spiking patterns of the individual neurons are linear dynamical systems such that the spike timing and spike shape are decoupled. Consistent with previous spike sorting algorithms we have suggested this heuristic is valid for a large parameter space of extracellular recordings, but the validity will break down in situations such as *burst firing* where the spike shapes and firing rates are coupled variables.

5

Conclusion

"I predict that within 10 years, computers will be twice as powerful, ten thousand times larger, and so expensive that only the five richest kings of Europe will own them"

– Professor Frink, *The Simpsons*

This thesis contains novel theoretical and experimental analysis of vLFPs from in vivo single channel MERs obtained from human STN. We have developed methods to analyse these vLFPs over multiple scales (using the contribution from the localised LFP and/or the nearby spikes) using techniques from non-equilibrium statistical mechanics, stochastic processes, signals processing, convex optimisation and clustering theory. For each of the three methodologies considered we have introduced a transformation operator which maps the one dimensional vLFP to either a number on the positive real ($\mathbb{R}^+$) line (chapter 2), the

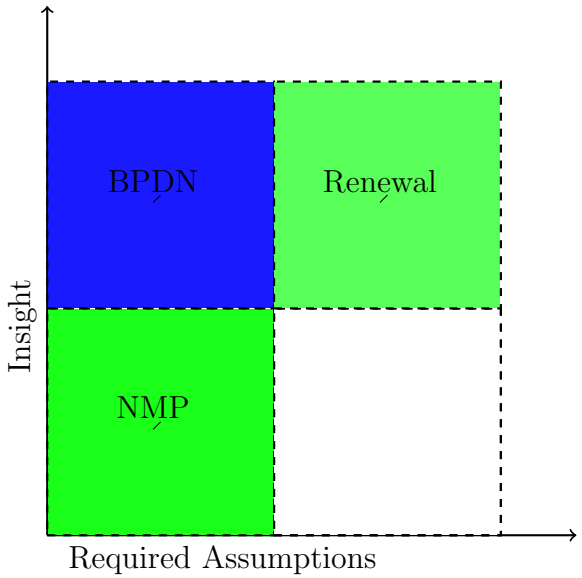


Figure 5.1: The three separate transformations considered to analyse the vLFP and their associated required assumptions (abscissa) and insight into the behaviour of the STN (ordinate).

probability distribution associated with the firing times (chapter 3) or the precise times and shapes of action potentials from the nearby neurons (chapter 4).

We have used these techniques to either identify changes in the electrical activity of the STN under different experimental conditions or understand the behaviour of the neurons nearest to the probe contributing to the measured signal. The results of using these methods highlights that both the measured power spectrum of the entire signal (LFP + spikes) and the precise firing patterns of the neurons nearest to the recording probe can be used to characterise the state of the STN. Thus, there is information about the STN in the vLFP over the multiple scales analysed in this thesis. These methods may have future utility as bio-markers in CLDBS algorithms. In the proceeding section we sum up the different methods used, the results obtained, their required assumptions, and their advantages and disadvantages.

5.1 Summary of Methods

Model-Free: NMP This methodology was the most general non-parametric approach considered in this thesis. In this method the first NMP is the transformation operator which maps the entire vLFP (LFP + spikes) to a single number on the positive real, $\mathbb{R}^+$, line. This methodology was used to identify statistically significant changes in the low frequency (β: 10-30 Hz, *fast*: 80-200 Hz and *zero frequency*: 0 Hz) bands of the NMP while patients were presented with different stimuli from two separate neuro-linguistic experiments. This work was published in the following conference articles:

- J. Varghese, K. Weegink, P. Bellette, T. Coyne, P. Silburn, and P. Meehan, "*Theoretical & Experimental Analysis of the Non Markov Parameter to Detect Low Frequency Synchronisation in Time Series Analysis*" in Proceedings of the 33rd Annual International Conference of the IEEE EMBS, Boston, Massachusetts USA, August 30-September 3,2011, 2011, pp. 1500-1505.
- P. Meehan, P. Bellette, A. Bradley, J. Castner, H. Chenery, D. Copland, J. Varghese, T. Coyne, and P. Silburn, "*Investigation of the Non-Markovity Spectrum as a Cognitive Processing Measure of Deep Brain Microelectrode Recordings*, International Conference on Bio-inspired Systems and Signal Processing, Rome, Italy, pp. 144-151, 2011.

The family of NMPs is constructed by solving a series of inverse problems associated with the stochastic integro-differential (Langevin) equations which describe the Mori-Zwanzig chains of non-equilibrium statistical mechanics. We showed that the first NMP contained the most useful information about the measured system, with higher order NMP veiling the systems correlation structure. We also showed that the first NMP could be understood in a signals processing framework as a series of nonlinear operations on the measured power spectrum. Thus, we suggested that when the NMP was used during the neuro-linguistic experiments, it was detecting subtle but statistically significant changes in the vLFP power spectrum. This work was published in the following journal article:

- J.J. Varghese, P.A. Bellette, K.J. Weegink, A.P. Bradley P.A. Meehan, "*Analysis of the non-Markov parameter in continuous-time signal processing*, Phys. Rev. E., vol 89(**2**), p. 022109, 2015.

The NMP method of analysis places effectively no requirements on the measured signal except for covariance-stationarity (so that the WienerKhinchin theorem guarantees that the power spectrum is the Fourier Transform of the autocorrelation function). The disadvantage of this method is that it does not provide much insight into the underlying physics driving the measured system, and provides even less link to concepts typically used to quantify measurements in neuroscience based on the statistics of the underlying spike trains. The best conceptual understanding we can develop from applying the first NMP to the neuro-linguistic experiments is that there are subtle changes in the distribution of the vLFP power spectrum. Nonetheless, this method is capable of resolving differences in the in vivo signals in a consistent fashion, using the entire signal (rather than a subset of the closest spikes) without introducing free parameters. Certainly, from a machine learning perspective, any metric (such as the NMP) which can distinguish different behavioural or functional states in the STN has the potential to be used as a biomarker. Considerations of biophysical accuracy are, for all intents and purposes, secondary for such applications.

Summary: The NMP approach requires minimal a priori assumptions of the vLFP, is capable of identifying changes in the in vivo STN vLFP recordings but provides minimal insight into the measured process beyond information about how energy is distributed in different frequency bands.

Renewal Model-Based: This methodology introduced the most stringent model restrictions considered in this thesis. In this method the measured signal is modelled as an ensemble of independent filtered renewal processes. The transformation operator maps the entire vLFP (LFP + spikes) to The PDF associated with the ISI times of the individual spike trains. This PDF is constructed by the serial solution of two

inverse problems. Firstly the Bartlett spectrum (the contribution to the power spectrum due to correlation structure of the individual spike trains) is identified from the measured power spectrum using inverse techniques in the frequency domain. Secondly the Bartlett spectrum is then transformed to the renewal density function by integral transforms and then finally the PDF is constructed by solving another inverse problem associated with the renewal density equation which is a Volterra integral equation of the second kind. The analysis of the relationship between the firing statistics and the power spectrum and the consequences for neural signals processing were explored in the following conference paper:

- J.J. Varghese, K.J. Weegink, P.A. Bellette, and A.P. Bradley, "*Spectral properties of neuronal pulse interval modulation,"* , in Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on, April 2015, pp. 1007-1011

In addition this paper also showed that the spectrum of a neuron encoding information using a Digital Pulse Interval Modulation (DPIM) scheme converges to a renewal process in the limit of continuous time. Building on this work, we developed the *Spectral Density Estimator* which, when the assumptions of the model are satisfied, is able to successfully identify the firing statistics (without explicitly identifying the spiking times) in the pathological situation of simultaneously recording multiple neurons with similar action potential shapes. We consider that this situation may arise when there are voltage contributions from multiple neurons which are a similar distance and orientation from a recording probe. The development of this estimator was submitted (but not yet accepted) in the following journal article:

- J. Varghese, K. Weegink, P. Bellette, and A. Bradley, *"Spectral techniques to estimate renewal spiking statistics with near identical spike shapes*" Phys. Rev. E, 2017.

Unfortunately the assumptions of this model were too stringent to successfully apply to the in vivo vLFPs. The assumptions require strict stationarity, that the individual

neurons can be treated as renewal processes and that the neurons contributing to the MER time series form a statistically independent ensemble. There is also an implicit assumption that the LFP contribution to the vLFP is dominated by the highly filtered distant spiking neurons. The failure of this methodology suggests that the assumptions of this method are not valid for the vLFPs measured. It is unclear which assumptions are invalid. Nonetheless, the advantage of this approach is in stark contrast to most classical spike sorting approaches, which rely on identifying differences between action potential shapes in order to separate individual neurons and develop estimates of their firing statistics.

Summary: the Spectral Density Estimator requires very restrictive a priori assumptions about the vLFP and was too restrictive to apply to the in vivo STN recordings but provides maximal insight into the firing statistics in pathological situations (when the assumptions are valid) where classical spike sorting techniques fail.

BPDN Spike Sorting: The characterisation of only the nearby spiking neurons using BPDN techniques provides an excellent balance of both minimal a priori assumptions while providing insight into the behaviour of the nearby neurons contributing to the vLFP. In this approach the transformation operator is the spike sorter, which maps the signal to the spiking shapes and times of the nearby neurons. This spike sorter is constructed in a two step process. The candidate spike shapes are estimated using a combined wavelet transform, diffusion mapping and mean-shift clustering approach. The spike times are then estimated by assuming the vLFP is a linear dynamical system, using convex relaxation, and subsequently employing BPDN strategies to develop accurate and sparse estimates of the firing times.

We demonstrated with simulated vLFPs that this BPDN spike sorter was superior (based on sensitivity, specificity and χ^2 estimates) to the state of the art spike sorting

program *wav-clus* (which is based on amplitude thresholding and Super Paramagnetic Clustering) in challenging low signal to noise ratio environments. We then applied this BPDN spike sorter to the vLFPs obtained from MERs inserted into human STNs of patients with Parkinson's Disease undergoing DBS. We identified Poisson statistic firing patterns with firing rates in the range of $20-56$ Hz. These results are consistent with previous studies of the behavior of the STN in patients with Parkinson's Disease [28],[170].

This BPDN spike sorting approach places no restrictions on the firing patterns of the individual neurons or their interaction with other neurons. The primary assumption introduced by this approach is that the spike trains evolve as linear dynamical systems and therefore the spike shape of a neuron is independent of the firing pattern. In this approach the BPDN solution is constructed by solving the inverse problem associated with the ℓ_1 regularised least squares problem. This method provides the precise firing times of a number of neurons close to the MER. The disadvantage of this approach is that it only characterises the signal by the behaviour of the neurons which happen, by chance, to be nearest to the recording probe and thus information from non nearby spiking processes is lost.

Summary: the BPDN spike sorter requires minimal a priori assumptions and provides excellent insight into the behaviour (in terms of elucidating the firing patterns) of the nearby contributing neurons, but only uses a very small subset of the entire vLFP.

5.1.1 Conclusion of Analysis

Throughout this thesis we have emphasised the differences of the three methods in terms of the scale of the vLFP considered (the localised LFP and/or the nearby spikes) and whether the methodology is parametric or non-parametric. With the results of these methods now provided and compared we are now in a position to summarise the *validity* of the application and the *information* provided by these three approaches. Specifically, we consider the

differences of these approaches in terms of how many *a priori requirements* must be satisfied and how much insight they provide into the *underlying biology* of the measured system. We see that the NMP approach (chapter 2) requires the least assumptions, but provides the least insight into the underlying biology of the measurements. The renewal approach (chapter 3) requires the most stringent assumptions, but also provides excellent insight into the underlying biology. The BPDN approach (chapter 4) provides a balance between minimal required assumptions and good insight into the underlying biology. This summary is provided graphically in Figure 5.1.

The key result from this thesis is that the vLFPs obtained from MERs contain useful information about the STN over the *multiple scales* (localised LFP, nearest neuron spiking patterns and the combination of both these effects) of analysis. This result is consistent with the assessment in [20], (also stated in the introduction chapter) that "*spike 'contamination' of the LFP should be regarded as good news, in that high-frequency LFP power can provide a 'proxy' for the assessment of neuronal outputs*". The next stage of research should ideally focus on exploring whether the bio-markers developed in this thesis can be correlated with clinical features of disease state. This could be achieved with the aid of scoring systems such as the Movement Disorder Society sponsored revision of Unified Parkinson's Disease Rating Scale (MDS-UPDRS) for Parkinon's Disease [271]. This can begin by using within-subject studies that may start with a small number of participants. Depending on how successful these bio-markers are at detecting pathological states, much longer term research goals would ideally be to consider implementing these bio-markers in future CLDBS systems.

It is not immediately clear which bio-marker future research should focus on. On one hand, the NMP analysed in chapter 2 requires minimal a priori assumptions of the vLFP, is capable of identifying changes in the in-vivo STN vLFP recordings but provides minimal insight into the measured process. On the other hand, the BPDN spike sorter developed in chapter 4 provides physiological insight, robust performance and minimal required a priori assumptions but only uses a subset of the entire vLFP. Another advantage of the BPDN approach is that this method could easily be scaled up to the more modern multi-electrode

systems which provide multiple simultaneous channels of in vivo recording to provide much larger 'meso-scale' information about the constituent neural processes in the targeted neural tissue.

5.2 Contributions

The following are believed to be novel contributions to the literature:

- Development of closed form expressions for the NMP - showing that it is a nonlinear series of operations on the power spectrum.
- Identifying that within the family of NMPs, the first NMP contains the most information, with subsequent NMPs losing information.
- Showing the NMP can be applied to vLFP recordings from MERs obtained from human STN to differentiate electrical states during neuro-linguistic testing.
- Showing that (excluding Poisson processes) the underlying PDF which drives the individual ISI times for a super-position of renewal processes can be accurately estimated using the measured power spectrum.
- Showing that this estimation of the ISI PDF from the measured power spectrum can, in certain circumstances (when there are multiple neurons equidistant and with similar orientation to a recording probe) provide superior estimates of the ISI PDF than classical spike sorting algorithms.
- Showing that the power spectrum of neuron encoding information using a DPIM coding scheme will have the power spectrum of a renewal process in the limit of continuum time.
- Developing the BPDN spike sorting algorithm which we show is superior to state of the art spike sorting algorithms in the challenging situation of low signal to noise ratios.
- Applying the BPDN spike sorter to the vLFP recordings from MERs obtained from human STN and identifying that the firing patterns of the individual neurons have, in

agreement with the literature [28],[170], physiologically plausible Poisson firing statistics with firing rates between $\sim 20 - 56$ Hz.

5.3 Limitations

In this section we identify limitations associated with the methods we have developed to analyse the vLFPs:

Accuracy of the Spectral Estimates

Recall that the NMP (chaper 2) and Spectral Density Estimator (chapter 3) approaches are constructed from spectral estimates of the vLFPs. Therefore, the accuracy of these metrics are fundamentally limited by how accurately the power spectrum can be estimated. It is a highly non-trivial problem to develop accurate spectral estimates from short time recordings. This problem is further compounded by the fact that the NMP bands considered in this thesis (β: 10-30 Hz, *fast*: 80-200 Hz and *zero frequency*: 0 Hz) and the relevant spectral information for the Spectral Density Estimator is in the low frequency content. Recall from section 3.3.2 of chapter 3 that we suggested, but did not prove, that the Bartlett spectrum which represents the contribution from the statistics of the ISI to the power spectrum is centred about the mean firing rate ν. In most circumstances physiological *mean* firing rates are between 1-100 Hz.

Throughout this thesis we have used the non-parametric Welch's Overlapping Segment Averaging (WOSA) method with fifty percent overlap which further reduces the length of the time series and thus degrades the quality of the low frequency contribution. The development of more sophisticated spectral estimation techniques is an active area of research. As discussed in section 3.5.1 of chapter 3, methods such as Thompson's multitaper method which, instead of partitioning the time series into smaller segments, reduces variance in the estimate with a series of orthogonal data tapers applied to the entire time series [174],[125] may be more appropriate for these short time recordings.

An alternative approach is to use parametric spectral estimators which are specifically accurate in low frequency regimes. One method proposed in [272] uses frequency selective parametric (Auto Regressive Moving Average model) spectral analysis in defined sub-bands. This estimator works by estimating the entire spectrum using a very high order Auto Regressive (AR) model. The spectrum in the sub band of interest is then transformed to an auto-covariance function and then an ARMA model is fitted to this auto-covariance function. The spectrum is then described using the coefficients of the ARMA model. The order of the ARMA model can be automatically selected using Generalised Information Criteria [273]. This approach is referred to as *sub-band remodelling*. This sub-band remodelling may find particular utility with the Spectral Density Estimator approach where the sampling rates of most MER acquisition systems is of the order 20 kHz, whereas the useful information content is only up to the order 100 Hz.

Dealing With Nonlinearity

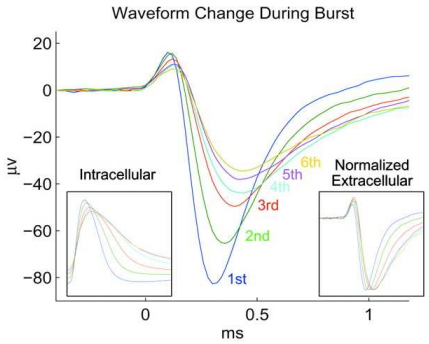


Figure 5.2: Bursting is a nonlinear phenomenon, introducing a coupling between the spike shape and firing times. Notice that the amplitude of the spikes decreases with the burst sequence. Figure taken from [163].

For both the Spectral Density Estimator (chapter 3) and the BPDN spike sorter (chapter 4) approaches, multiple linear heuristics have been introduced to deal with the complexity of the vLFP. As discussed in Section 4.3.2 of chapter 4, these linear assumptions are:

1. **spike shape stationarity:** The spike shape (for a single neuron) does not vary and is not influenced by firing times.

2. **nearby neuron linearity:** The behaviour of the nearby neurons is given by the linear superposition of the individual spike trains of the constituent neurons.

3. **LFP + nearby spikes linearity:** the overall vLFP is given by the linear summation of the nearby neurons and the LFP.

For all intents and purposes the linearity of the voltage contribution from the nearby neurons (assumption 2) is likely to be valid. It is unclear how to identify whether the linearity of the nearby spikes and the LFP (assumption 3) is valid, indeed it is even less clear how this validity could be explored. It is important to note that while we have assumed that the nearby spikes and the LFP can be linearly summed, with the Spectral Density estimator (chapter 3) we have also assumed that these two terms are uncorrelated, whereas with the BPDN spike sorter (chapter 4) we *have not* assumed that these two terms are uncorrelated. We discuss the implications of the spike shape stationarity (assumption 1) to the validity and accuracy of the BPDN spike sorter below:

The application of the BPDN approach requires that the firing patterns of the individual neurons can be modelled as linear dynamical systems. Stated another way, the BPDN spike sorting methodology requires the introduction of a heuristic that the spike trains can be modelled in a linear framework where the action potential shapes, $g(t)$, and firing times, $\{t_k\}$ are decoupled. This can mathematically be represented by: $f(g(t), \{t_k\}) = g(t) * \sum_{k=1}^{\infty} \delta(t - t_k)$. For a very large parameter space of extracellular recordings this assumption is likely valid. This is evidenced by almost every spike sorting algorithm [165], [57],[192] operating under the assumption that the spike shape does not vary over the time course of the recording.

One important area where this assumption of linearity is known to break down is in the presence of burst firing. When a neuron fires in a burst sequence (with spikes typically less than 100 milliseconds apart [163], [164]) the action potential shape is continuously deformed, usually with the amplitude sequentially decreasing. This is an intracellular phenomenon [57], which is shown in Figure 5.2. Applying the linearity heuristic to bursting data will likely result in the different amplitude spikes being (incorrectly) partitioned into different clusters.

We note that the *SpikeOMatic* program, which can solve the spike sorting program using a Dynamic Hidden Markov Model [216], combines information about both the determined firing times *and* amplitudes with a simple model for variation of the spike amplitude in the presence of burst firing to often successfully cluster spikes in the presence of bursting.

An open area of research is determining how to extend the BPDN spike sorting algorithm to this fundamentally nonlinear problem. An interesting extension of the BPDN spike sorter would be to use an underlying model, similar to [216], which combines information about the firing times *and* the spike amplitude to incorporate the variation of amplitude when spikes from the same neuron occur in rapid succession.

We discuss additional limitations of the methods introduced in this thesis below:

- The NMP and renewal chapters assume that the measured signals are stationary over the time scales examined.

- With a single channel recording it is not possible to capture any of the spatial dependence of the signal. This is a fundamental limitation of a single channel system.

- The methods we have used to analyse the signal do not consider the synaptic weightings between the different neurons. It may be possible to argue that the evolution of the firing rates (in a Hebbian learning framework) or the specific timing (in a Spike Timing Dependent Plasticity framework) between specific neurons may provide are a 'proxy marker' for the synaptic weighting between these identified neurons.

- We have not attempted to correlate these transformations of the signal to either the clinical state of patients with Parkinson's Disease or the success of different stimulation protocols. If we are to consider any of the methodologies introduced in this thesis as potential bio-markers for future CLDBS strategies, larger scale clinical correlation and animal studies will be required. We discuss this further in section 5.4.3

5.4 Extensions and Future Work

In this section we identify extensions to the methods developed in this thesis and future work which may be performed. The recommendations are broken up into three broad sections: combining information about the vLFP from multiple scales, extending current methods and future clinical work.

5.4.1 Combining Information From Multiple Scales

In this section we describe how combining information from the methods developed in this thesis may provide more insight into the measurements obtained of the STN and improve the robustness of these algorithms. We specifically consider two extensions: firstly, identifying how the zero frequency value of the NMP varies with firing pattern if the STN is modelled as an ensemble of independent neurons (combining information from the chapters 2 and 3 approach) and secondly, suggesting how the spectral properties of a renewal process may improve the accuracy of the BPDN spike sorter when the action potential associated with two or more neurons is sufficiently similar (combining information from the chapters 3 and 4 approach).

Zero Frequency NMP for Poisson and Periodic Renewal Processes

In section A.9 the zero frequency value of the first NMP of a renewal process is calculated as:

$$\epsilon_{1,\Sigma}(0) = \pi G(0) C_v^2 \cdot \sqrt{\frac{\int_{-\infty}^{+\infty} \omega^2 G(\omega) \left[1 + 2Re\left\{\frac{H(\omega)}{1-H(\omega)}\right\}\right] d\omega}{\left(\int_{-\infty}^{+\infty} G(\omega) \left[1 + 2Re\left\{\frac{H(\omega)}{1-H(\omega)}\right\}\right] d\omega\right)^2}} \tag{5.1}$$

Where $C_v = \nu\sigma$ is the coefficient of variation of the ISI distribution. We now use this expressions (5.1) to develop explicit expressions of the $\epsilon_1(0)$ value for the two different extremes of renewal statistics: maximally random Poisson and periodic firing statistics.

Recall from section 3.3.2 of chapter 3 that for Poisson processes $2Re\left\{\frac{H(\omega)}{1-H(\omega)}\right\} = 0$ and $C_v = 1$. Therefore the NMP of a Poisson process is given by:

$$\epsilon_{1,\text{Poisson}}(0) = \pi G(0) \cdot \sqrt{\int_{-\infty}^{+\infty} \omega^2 d\omega} = \infty \tag{5.2}$$

This result is consistent with the analysis in section 2.5 of chapter 2 where we showed that Markov processes have maximal (infinite) zero frequency NMP values.

Recall that by definition the variance of a purely periodic processes is zero and thus $C_v = \sigma = 0$. Using the expression for the Bartlett spectrum of a periodic process given in section 3.3.2 of chapter 3 the NMP of a periodic renewal process is given by:

$$\epsilon_{1,\text{Periodic}}(0) = \pi G(0) \cdot 0 \cdot \sqrt{\frac{\int_{-\infty}^{+\infty} \omega^2 \left[1 + \sum_{n=-\infty}^{n=+\infty} \delta(\omega - n\lambda)\right] d\omega}{\left(\int_{-\infty}^{+\infty} G(\omega) \left[1 + \sum_{n=-\infty}^{n=+\infty} \delta(\omega - n\lambda)\right] d\omega\right)^2}} = 0 \tag{5.3}$$

This result is consistent with the analysis in section 2.5 of chapter 2 where we showed that Non-Markov processes have very small zero frequency NMP values.

Therefore, it appears that the NMP might be useful in distinguishing different renewal firing patterns. This also suggests that the changes of the electrical behaviour of the STN using the zero frequency component of the NMP in section 2.9 of chapter 2 may reflect changes in the firing patterns of the measured neurons. However, it is important to remember that since the NMP was applied to the entire vLFP, many non spiking low frequency (localised LFP) contributions to the LFP would of also affected the NMP value.

The following represents extensions to the work performed in this thesis which may provide further insight into the NMP and the behaviour of the STN under the experimental conditions considered in sections 2.9:

- As discussed in section 3.3.2 of chapter 3 the Poisson and periodic firing patterns represent the two different extremes of ISI time drawn from the very general Weibull

distribution. Can simplifications of the power spectrum of the Weibull distribution be introduced to develop simplified expressions for the NMP as the firing pattern is varied?

- We have identified that the two different extremes of firing patterns, Poisson and periodic, are mapped by the NMP to the extremes of the number line (approaching infinity and zero respectively). How sensitive is the NMP to variation of the firing statistics between these extremes?

- If the models of the vLFP used in chapter 4 are considered (individual, resolveable spike trains embedded in highly correlated non-white background noise from the LFP), how robust is the NMP to measuring changes in the renewal statistics of the resolvable spikes as the amplitude of the LFP background noise increases?

- Similarly, is the value of the NMP dominated by high frequency spiking behaviour of the nearest neurons or the non-spiking, low frequency LFP processes?

Applying renewal theory to the BPDN spike sorter

In chapter 4 we have shown with extensive simulation that the BPDN spike sorting algorithm is capable of outperforming (as judged by sensitivity, specificity and χ^2 estimates) state of the art spike sorters in challenging low signal to noise ratio environments. Nonetheless, as we identified in section 3.7.2 of chapter 3, *any* spike sorting algorithm which performs clustering based on the amplitude and or shape of the identified action potentials will fail when multiple neurons produce identical spike shapes. Therefore, the BPDN spike sorter will also provide inaccurate results for the pathological situation when one or more neurons produces sufficiently similar action potential shapes. This situation may indeed occur in neuron dense environments such as the STN when neurons are equidistant and similarly aligned relative to the MER probe.

In section 3.3.3 of chapter 3 we argued that (with the exception of Poisson processes) each firing pattern of a renewal process has a unique Bartlett spectrum. In a very broad sense

we considered this unique Bartlett spectrum a 'spectral fingerprint' in the low frequency component of the measured spectrum. We can use this property to introduce an 'error-checking' step to the BPDN spike sorter. Consider the following situation: the MER probe is inserted into neural tissue such that there are two neurons: neuron A and neuron B which are similarly orientated and an equal distance from the MER probe. If the spike shapes of neurons A and B are sufficiently similar, the BPDN spike sorter will either merge these two distinct firing clusters into a single cluster or attribute a large number of spikes from neuron A to neuron B, or vice versa. If the firing patterns of neurons A and B are, or can be approximated, as *independent* renewal processes we can "error check" the set of firing times identified for each neuron using classical spike sorting as follows:

1. Apply the BPDN (or any spike sorting) algorithm.

2. For each cluster, generate the KDE of the ISI probability distribution associated with the identified firing times. *Calculate* the Bartlett spectrum associated this KDE of the ISI PDF of this cluster using (3.16).

3. *Measure* the Bartlett spectrum of this cluster directly from the set of identified spike times (see [175],[176]) in this cluster.

4. Compare these measured and calculated Bartlett spectra. If there is a large discrepancy in values there has likely been an error with how the spike sorting algorithm has clustered the detected spike times.

Note that this potential method would work on the principle elucidated in [158] and discussed in section 3.4 of chapter 3, that the super-position of independent renewal processes generates a non-renewal process with a spectrum which is a scaled form of the constituent renewal processes. If the spike sorter incorrectly generates a cluster which consists of the spikes from neuron A in addition to multiple spikes from other neurons there will be a discrepancy between the calculated Bartlett spectrum and the measured Bartlett spectrum. This is because the calculated spectrum will be calculated using the incorrect KDE of the ISI distribution whereas the measured Bartlett spectrum (if the independent renewal assumptions are valid) will be the linear combination of the individual Bartlett spectra associated

with the ISI distribution. Therefore, if the measured and calculated Bartlett spectra for each of the clusters are consistent then the BPDN (or any alternative) spike sorter has likely identified most of the spikes associated with a single cluster.

The principle problem with this extension is that while it may identify if the estimated firing times are consistent with the measured power spectrum, it does not suggest how to *improve* the estimates if these values are inconsistent.

5.4.2 Development of potential bio-markers from BPDN

In chapter 4 we developed the BPDN algorithm and showed that it was capable of resolving the firing times and shapes of individual neurons contributing to a vLFP with high accuracy. The outputs from this algorithm are the individual spike shapes and sets of the discrete firing times associated with individual neurons. In the context of this thesis an immediate question is: given this output, what metrics should be constructed to characterise the state of the STN? An obvious approach, similar to chapter 3, would be to assume that the spike trains are renewal processes, and develop metrics from the statistical moments (firing rate, the coefficient of variation or the information entropy) of the ISI times. An alternative approach would be to model the data as a point process and develop the Bartlett spectrum (the power spectrum of the firing time correlation structure discussed in chapter 3) using the discrete sets of firing times (see [175],[176]). Given an accurate estimation of this Bartlett spectrum, previously developed metrics such as the peak energy in the β band ($10 - 35$ Hz) [56] or even the NMP (considered in chapter 2) could be used to characterise the state of the STN and potentially be used as future CLDBS bio-markers.

5.4.3 Efficacy of developed metrics as CLDBS biomarkers

The methods (NMP, Spectral Density Estimator and the BPDN spike sorter) developed in this thesis to analyse the vLFPs obtained from MERs have demonstrated that they can be used to characterise the state of the STN and identify the firing patterns of a subset of the

contributing neurons. Recall that the motivation for this research was largely driven by attempting to develop potential biomarkers for use in future CLDBS systems for the treatment of Parkinson's Disease. The first required step is to identify whether these metrics can be correlated to the degree of severity of Parkinson's Disease pathology. This is required in order to confirm a 'numerical measure' of the patient's disease state (which can be measured near instantaneously by the MER probes) for use in any future adaptive CLDBS system. There is a long history in Parkinson's Disease research of exploring whether metrics can characterise the state of the disease. Multiple studies have suggested that β band synchronisation may play a role in the motor symptoms of Parkinson's [50], [51], [52], [53], [54], [55]. More recently, complexity metrics such as the Lempel-Ziv Complexity (LZC) metric have been applied to vLFPs band pass filtered in the β (13-35 Hz) range and correlated with the UPDRS score for akinesia/rigidity.

The first recommended step in exploring the efficacy of these metrics as potential biomarkers would be to explore whether correlations exist between the patient's clinical state (assessed using scoring tools such as the MDS-UPDRS system) and the metrics (NMP, statistical moments or the spectra of the identified spike trains) developed in this thesis. One approach would be to set up the designated metric(s) and the MDS-UPDRS scores in a regression framework with a relatively large collection of patients and, similar to [99], use a LASSO (ℓ_1 minimisation) approach to determine which *specific* MDS-UPDRS scores correlate most strongly to the developed metrics. Due to the comorbidities of patients, the variable state of their Parkinson's Disease and the subjective nature of the MDS-UPDRS scoring systems there is likely to be a high degree of variability associated with the detected correlations. Due to this intrinsic variability, it is instead recommended that these metrics be analysed using within-subject studies that compare their values with and without DBS or on and off medication. For practical reasons these studies could begin with a relatively small cohort, with an extension to larger sample sizes pending the outcome of the smaller studies.

If these bio-markers are successfully able to identify pathological disease states, much further future work may be able to explore whether they could be used in future CLDBS

systems. This style of approach was considered in [5] where it was suggested that certain characteristic shapes of late onset evoked action potentials near the site of DBS in the STN are more responsive to stimulation than others. A simpler approach could be to follow the methodology in [56] where a stimulating current with an a priori selected voltage and frequency (perhaps the standard open loop values) are applied when the biomarker gets above a threshold value and stays quiescent when the biomarker is below the threshold value. These approaches could first be analysed theoretically using computer simulation studies, similar to [58], but will more than likely require animal model studies, similar to the approach of [40].

Exploring the efficacy of these bio-markers as markers of disease states will be a difficult and time consuming process. Furthermore, the subsequent analysis of whether these metrics could be considered as bio-markers for CLDBS will be a highly non-trivial future task. Nonetheless, the success of this approach may provide the exciting prospect of drastically reducing the impact of a cruel, debilitating and currently incurable disease which affects more than 10 million people worldwide.

A

Appendices

A.1 Deriving λ Relaxation Parameter

In this appendix we will determine two equivalent expressions for the set of first relaxation parameters $\{\lambda_n\}$ in terms of the autocorrelation, $m_n(t)$, and the power spectrum: $M_n(\omega)$. We will show that for C^1 smooth autocorrelation functions or equivalently power spectrums which decay faster than $\mathcal{O}(\omega)$ this parameter is always equal to zero.

The λ_n relaxation parameter is very easy to understand in terms of the autocorrelation function. Applying the limit as time goes to zero from the positive side to (2.6) yields:

$$\lim_{t\to 0^+} \frac{dm_n(t)}{dt} = \lambda_n \lim_{t\to 0^+} m_n(t) - \Lambda_n \lim_{t\to 0^+} \int_0^t m_{n+1}(t-t')m_n(t')dt' \quad t \geq 0. \tag{A.1}$$

Notice the following simplifications:

1. Because the autocorrelation function is normalised: $\lim_{t\to 0^+} m_n(t) = m_n(0) = 1$.

2. The convolution integral in the limit of zero time is zero:

$$\lim_{t\to 0^+} \int_0^t m_{n+1}(t-t')m_n(t')dt' = \int_0^0 m_{n+1}(t-t')m_n(t')dt' = 0 \tag{A.2}$$

Thus:

$$\lambda_n = \lim_{t\to 0^+} \frac{dm_n(t)}{dt} \tag{A.3}$$

Recall that autocorrelation functions are by definition symmetric. Also recall that the first derivative around the origin of a symmetric function is always zero. Thus the λ_n relaxation parameter must be zero unless there is a breakdown in smoothness at the origin of the $m_n(t)$ autocorrelation function. An example of a function with a breakdown in smoothness at the origin is $m_n(t) = e^{-a|t|}$. Figures A.1 & A.2 show the functions and derivatives respectively of a smooth function $m_n(t) = sinc(\omega t)$ and a non-smooth function $m_n(t) = e^{-a|t|}$ (with breakdown at the origin).

We will now provide the analysis to understand the requirements on the power spectrum for a non zero first relaxation parameter λ_n:

Notice that the first derivative at the origin can be written as:

$$\lim_{t\to 0^+} \frac{dm_n(t)}{dt} = \lim_{h\to 0^+} \frac{m_n(h) - m_n(0)}{h} \tag{A.4}$$

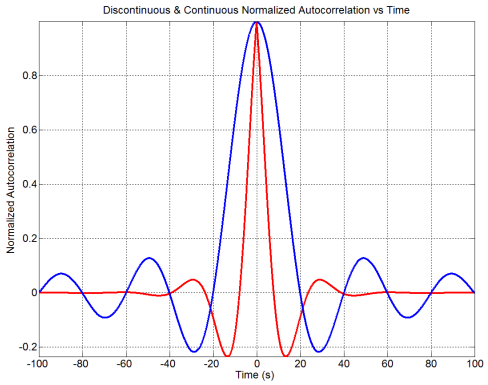


Figure A.1: Example of smooth $m_n(t) = sinc(\omega t)$ and non-smooth $m_n(t) = e^{-a|t|}$ memory autocorrelation functions. Notice the breakdown in smoothness at the origin.

Now $m_n(h)$ (which is the autocorrelation function where we have represented the time variable by h) can be written in terms of the inverse Fourier transform of the power spectral density by the Wiener-Khinchine theorem:

$$\begin{aligned} m_n(h) &= \mathcal{F}^{-1}\left[M_n(\omega)\right] \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} M_n(\omega)e^{i\omega h}d\omega \end{aligned} \tag{A.5}$$

Now $m_n(0)$ can similarly be written:

$$m_n(0) = \frac{1}{2\pi}\int_{-\infty}^{\infty} M_n(\omega)e^{i\omega(t=0)}d\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty} M_n(\omega)d\omega \tag{A.6}$$

Thus:

$$\begin{aligned} \lim_{h\to 0^+}\frac{m_n(h)-m_n(0)}{h} &= \lim_{h\to 0^+}\left\{\frac{1}{2\pi}\int_{-\infty}^{\infty} M_n(\omega)e^{i\omega h}d\omega - \frac{1}{2\pi}\int_{-\infty}^{\infty} M_n(\omega)d\omega\right\} \\ &= \lim_{h\to 0^+}\frac{1}{2\pi}\int_{-\infty}^{\infty} M_n(\omega)\left(e^{i\omega h}-1\right)d\omega\Big/h \end{aligned} \tag{A.7}$$

Therefore:

$$\lambda_n = \lim_{h\to 0^+}\frac{1}{2\pi}\int_{-\infty}^{\infty} M_n(\omega)\left(e^{i\omega h}-1\right)d\omega\Big/h \tag{A.8}$$

By the theorem of *Dominated Convergence* we can see that if $M_n(\omega)$ decays faster than $1/\omega$ then $M_n(\omega)$ is integrable and its behaviour will dominate the integration and the limit can be brought inside the integral. If the limit is brought inside the integral we have:

$$\lim_{h\to 0^+}\left(e^{i\omega h}-1\right) = 0 \tag{A.9}$$

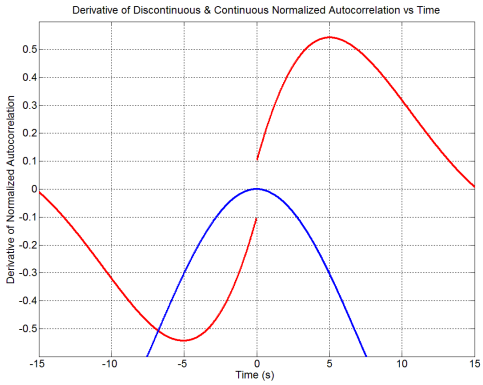


Figure A.2: Derivative of the smooth $m_n(t) = sinc(\omega t)$ and non-smooth $m_n(t) = e^{-a|t|}$ memory autocorrelation functions. Notice that the derivative is zero at the origin of the smooth function, but is non zero (and the limits from above and below zero are different values) for the non-smooth function.

In this case that the power spectrum decays faster than $1/\omega$ both the numerator and the denominator will be zero:

$$\lambda_n = \frac{0}{0} \tag{A.10}$$

Using L'Hospitals Rule:

$$\lambda_n = \lim_{h\to 0^+} \frac{1}{2\pi} \cdot \frac{\dfrac{\partial}{\partial h}\int_{-\infty}^{\infty} \left(M_n(\omega)e^{i\omega h} - 1\right) d\omega}{\dfrac{\partial h}{\partial h}} \tag{A.11}$$

$$= \lim_{h\to 0^+} \frac{i}{2\pi} \int_{-\infty}^{\infty} \omega M_n(\omega) e^{i\omega h} d\omega \tag{A.12}$$

Now if the Power Spectrum decays faster than $O(1/\omega^2)$ then by the theorem of Dominated Convergence then the limit can be brought inside the integral:

$$\lambda_n = \frac{i}{2\pi} \int_{-\infty}^{\infty} \omega M_n(\omega) d\omega \qquad \text{For} \qquad M_n(\omega) > O\left(\frac{1}{\omega^2}\right) \tag{A.13}$$

The power spectrum is by definition an *even* function whereas ω is clearly and odd function. The product of an even and odd function is an odd function. The integral of an odd function over the entire real line must necessarily be zero. Thus:

$$\lambda_n = 0 \qquad \text{For} \qquad M_n(\omega) > O\left(\frac{1}{\omega^2}\right) \tag{A.14}$$

A.2 Deriving Λ Relaxation Parameter

In this appendix we will develop expressions for the set of $\{\Lambda_n\}$ relaxation parameters. We show that the relaxation parameters Λ_n can be understood in terms of the limit as time goes to zero of the second derivative of the autocorrelation functions $m_n(t)$. We will then show that if certain conditions are met that this can be alternatively analysed in terms of the 'spread' of the power spectral density $M_n(\omega)$.

Taking the first derivative of (2.6) yields:

$$\frac{d^2 m_n(t)}{dt^2} = \lambda_n \frac{dm_n(t)}{dt} - \Lambda_n \frac{d}{dt} \int_0^t m_{n+1}(t-t')m(t')dt' \tag{A.15}$$

The only difficulty with the above equation is determining the derivative of the convolution term. Notice that this is not a convolution defined with integration bounds from -∞ to +∞ and thus the derivative operation cannot be brought inside the integral and applied to either of the convolution products. Also notice that the *fundamental theorem of calculus* cannot be used because the integral bounds depend on the time variable which is a function of the integrand. We use the generalised Leibniz rule for differentiation of an integral:

$$\text{Let} \qquad \phi(\alpha) = \int_{a(\alpha)}^{b(\alpha)} f(t',\alpha)dt' \tag{A.16}$$

$$\frac{\partial \phi}{\partial \alpha} = \int_{a(\alpha)}^{b(\alpha)} \frac{\partial}{\partial \alpha} f(t',\alpha)dt' + f(b,\alpha)\frac{\partial b}{\partial \alpha} - f(a,\alpha)\frac{\partial a}{\partial \alpha}. \tag{A.17}$$

For the convolution term in (A.15) our definitions translate as:

$$\alpha \equiv t \tag{A.18}$$

$$b(\alpha) \equiv t \tag{A.19}$$

$$a(\alpha) \equiv 0 \tag{A.20}$$

$$f(t',\alpha) \equiv m_{n+1}(t-t')c(t')dt' \tag{A.21}$$

For our ZM chain the following quantities need to be calculated:

$$f(b,\alpha) = f(t'=t,t) = m_{n+1}(0)m_n(t) \tag{A.22}$$

$$f(a,\alpha) = f(t'=0,t) = m_{n+1}(t)m_n(t) \tag{A.23}$$

$$\frac{\partial b}{\partial \alpha} = \frac{\partial t}{\partial t} = 1 \tag{A.24}$$

$$\frac{\partial a}{\partial \alpha} = \frac{\partial 0}{\partial t} = 0 \tag{A.25}$$

So:

$$\begin{aligned}\frac{d}{dt}\int_0^t m_{n+1}(t-t')m_n(t')dt' &= \int_0^t \frac{\partial m_{n+1}(t-t')}{\partial t}m_n(t')dt' + m_{n+1}(0)m_n(t) - m_{n+1}(t)m_n(t)\cdot 0 \\ &= \int_0^t \frac{\partial m_{n+1}(t-t')}{\partial t}m_n(t')dt' + m_{n+1}(0)m_n(t)\end{aligned} \tag{A.26}$$

Substituting this result back into (A.15) yields:

$$\frac{d^2m_n(t)}{dt^2} = \lambda_n\frac{dm_n(t)}{dt} - \Lambda_n\left(\int_0^t \frac{\partial m_{n+1}(t-t')}{\partial t}m_n(t')dt' + m_{n+1}(0)m_n(t)\right) \tag{A.27}$$

If we take the limit as time goes to zero of this equation:

$$\lim_{t\to 0^+}\frac{d^2m_n(t)}{dt^2} = \lambda_n\lim_{t\to 0^+}\frac{dm_n(t)}{dt} - \Lambda_n\left(\lim_{t\to 0^+}\int_0^t \frac{\partial m_{n+1}(t-t')}{\partial t}m_n(t')dt' + m_{n+1}(0)\lim_{t\to 0^+}m_n(t)\right) \tag{A.28}$$

The convolution term in the limit will be zero (see section A.1). By definition the limit: $\lim_{t\to 0^+} m_n(t)$ is unity. Thus this expression can be simplified:

$$\lim_{t\to 0^+}\frac{d^2m_n(t)}{dt^2} = \lambda_n^2 - \Lambda_n m_{n+1}(0) \tag{A.29}$$

Re-arranging we can get a form for the Λ_n relaxation parameter:

$$\Lambda_n m_{n+1}(0) = \lambda_n^2 - \lim_{t\to 0^+}\frac{d^2m_n(t)}{dt^2} \equiv \lim_{t\to 0^+}\frac{dm_n(t)}{dt}^2 - \lim_{t\to 0^+}\frac{d^2m_{n+1}(t)}{dt^2} \tag{A.30}$$

If we limit our analysis to 'sensible' signals ($M_n(\omega) \geq O\left(1/\omega^2\right)$) where the λ_n relaxation parameter is zero we can simplify the Λ_n relaxation parameter equation to:

$$\Lambda_n m_{n+1}(0) = -\lim_{t\to 0^+}\frac{d^2m_n(t)}{dt^2} \tag{A.31}$$

Imposing the condition that the memory function at time zero $m_{n+1}(0)$ must be unity leads to the expression for 'sensible' signals:

$$\Lambda_n = -\lim_{t\to 0^+} \frac{d^2 m_n(t)}{dt^2} \tag{A.32}$$

We will now show the Λ_n relaxation parameter in terms of the power spectral density.

Recall from the Wiener-Khinchtine formula that:

$$m_n(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} M_n(\omega) e^{i\omega t} d\omega \tag{A.33}$$

Taking the second derivative of this equation with respect to time yields:

$$\frac{d^2 m_n(t)}{dt^2} = \frac{-1}{2\pi}\int_{-\infty}^{+\infty} \omega^2 M_n(\omega) e^{i\omega t} d\omega \tag{A.34}$$

Taking the limit of this equation as time goes to zero yields:

$$\lim_{t\to 0^+} \frac{d^2 m_n(t)}{dt^2} = \frac{-1}{2\pi}\int_{-\infty}^{+\infty} \omega^2 M_n(\omega) d\omega \tag{A.35}$$

Substituting this result into the equation for the Λ_n relaxation parameter in terms of the limit of the second derivative of the autocorrelation function (A.32) yields:

$$\Lambda_n = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \omega^2 M_n(\omega) d\omega \tag{A.36}$$

A.3 Neuro-Linguistic Experiment 2 Procedure

In this appendix we describe the experimental setup and signals acquisition protocol used to obtain the MERs from the second experiment which were subsequently analysed using the NMP as described in Section 2.9.4. This description is based heavily on the (as yet) unpublished article *Microelectrode recordings of subthalamic nucleus activity during language processing in people with Parkinson's disease* with authors: *Helen J Cheney, Anna D Macdonald, Andrew P Bradley, Peter A Silburn, Andrew Smith, John Varghese, Terry Coyne, David A Copland* which used the same data, but analysed the signals differently using the spike sorting package *Osort*.

Participants

Ten participants were initially recruited to the study but the data from four of these participants were discarded due to participant drowsiness during the intra-operative testing procedure or technical difficulties with data acquisition. Subsequently, six participants (five male) diagnosed with idiopathic PD without dementia who were undergoing bilateral DBS surgery for treatment of their symptoms participated in the study (mean age 66.77 $\pm$ 6.52 years [range 59-78]; mean years of education, 15.5 $\pm$ 4.13 years [range 10-21]; mean disease duration from time of diagnosis, 8.5 $\pm$ 4.70 years [range 4-17]). A summary of the participant characteristics is shown in Table A.1.

All participants were native English speakers except for participant 10 who reported being multi-lingual (Greek and Afrikaans) with English as a first language. The data from this participant was retained as preliminary inspection of the data showed no significant deviation from the group. The study was approved by The University of Queensland Medical Research Ethics Committee and UnitingCare Health Human Research Ethics Committee. All participants gave informed written consent.

Participant	Age	Years of Education	Duration Since Diagnosis	Average LD Dose (mg/day)	EHI score	GDS-SF	SMMSE	PD-CRS
2	64	10	8	1300	R (100)	3	27	100
10	78	21	5	400*	R (78)	3	30	89
11	62	18	10	800	Ambi (30)	3	30	105
14	65	13	17	850	R (79)	1	30	82
16	59	13	7	750	R (79)	0	30	na
18	69	18	4	750	R (71)	1	28	na

Table A.1: Patient Characteristics for Second Neuro-Linguistic experiment. Acronyms: EHI: *Edinburgh Handedness Inventory*, GDS-SF: *Geriatric Depression Scale Short Form*, SMMSE: *Standardized Mini-Mental State Examination*, PD-CRS: *Parkinson's disease Cognitive Rating Scale*. * Patient 10 was also taking 150mg/day desvenlafaxine.

Surgical Procedure

The DBS leads were implanted under local anaesthesia after overnight withdrawal of all antiparkinsonian medication. The neurosurgical team targeted the dorsolateral aspect of the STN using a Cosman-Roberts-Wells stereotactic frame with patients set on a Medtronic Stealth surgical planning station. The stereotactic coordinates were determined via direct visualisation of the target after merging of preoperative thin slice (1mm) CT with 3-T MRI T1 and fluid-attenuated inversion recovery (FLAIR) sequences.

During surgery, a 500 micron tungsten microelectrode (Fred Hayer Corporation) was passed via a microdrive attached to the head frame to 5mm above the subthalamic nucleus. The procedure is described fully in [14], but briefly, microelectrode recordings were obtained in a selected trajectory to verify the characteristic STN firing patterns that assist in optimal target identification. The participant was examined for effect on clinical signs (dyskinesia, tremor, rigidity, bradykinesia) and absence of adverse effects. At this stage, the experimental testing was carried out (full details are provided below). After testing was completed, the microelectrode was withdrawn and the permanent DBS lead (Medtronic Lead 3387) was placed through the same guidetube.

The procedure was then repeated on the other brainside. As a general guide,the neurosurgical team implanted the first lead in the STN on the contralateral side to the patient's worst motor symptoms, which for five of the six participants was the left STN.

Language Stimuli and Procedure

Two stimulus sets (one for each hemisphere) each consisting of 30 real words and 30 nonwords were created and presented in a counterbalanced fashion across participants. Real words were nouns and were matched across sets (F values for all tests < 1.14; p values for all tests $> .289$) for mean concreteness (498.25 ± 107.65), imageability (532.72 ± 72.66), Kucera-Francis written frequency (173.65 ± 140.51), number of letters (5.18 ± 1.32) [274] and mean lexical decision time (594.4 ± 45.75) and accuracy ($.98 \pm .02$) [275] and mean total CELEX frequency (169.05 ± 145.92) [276]. Nonwords were taken from the ARC Nonword Database [150], and were orthographically legal, pronounceable and not homophonic with real English words. The mean number of letters for nonwords (5.07 ± 1.3) was matched to real words within a list (List 1, $F = 0.038$, $p = 0.846$; List 2, $F = 0.239$, $p = 0.627$). Nonwords were also matched across sets for number of letters ($F = 0$; $p = 1$) and neighbours ($F = 0.08$; $p = 0.779$).

The presentation of the lexical decision task and collection of behavioural responses (accuracy and response times) was controlled and timed by E-Prime Version 1.1 running on a

DELL laptop with a modified USB two-button mouse attached. An external monitor was positioned over the patient on the operating table via an articulating monitor arm. Experimental stimuli were pseudorandomly ordered such that there were no more than three consecutive word or nonword trials. Each experimental trial began with a fixation cross (+) presented in the centre of the screen for 1500 ms followed by the stimulus which remained on the screen until the participant responded or a maximum duration of 3 seconds elapsed. Participants were instructed to indicate via button press whether the stimulus was a real word or not (i.e., a nonword). The hand that was ipsilateral to the hemisphere of the STN MER acquisition executed the button press. For each hand, the index finger was always used to indicate Yes (real word) and the middle finger for No (nonword). The duration of the inter-trial interval was set at a minimum of 500 ms and, dependent on the response time for each stimulus, varied to make each trial length 5 seconds in total. Participants completed a short practice task with stimuli not used in experimental sessions 1-2 days prior to their surgery.

MER Recordings and Analysis

In this study, we acquired 1,440 baseline and task micro-electrode recordings (MERs) using two hardware systems: a Leadpoint system (Medtronic) with sampling rates of 22 kHz and a FHC Guideline LP+ system, sampling frequency 48 kHz. We removed individual MERs that were linked to participant errors on the lexical decision task or where there were too few spikes detected in an epoch leaving 934 of the total 1,440 MERs (or 64.9%) for final analysis. The MERs were pre-amplified and filtered to a bandwidth of 500-5000 Hz prior to sampling. The filter settings used were those recommended by Medtronic to reduce the effect of muscle artefact, 50Hz mains interference and background electroencephalographic activity.

Time Synchronisation:

Each MER recording was synchronized (time aligned) with the experimental timings using the E-prime software to insert detectable 'trigger' pulses into the MER recordings at the end of both the fixation and inter-trial epochs for each trial (see Figure A.3). The trigger was inserted using an opto-isolated connection between the parallel port on the laptop running the E-Prime experiment and the Leadpoint/FHC bio-amplifier input. These trigger pulses were then combined with the timing information from E-prime so as to enable each MER recording to be broken up into individual trials and within these trials, individual epochs. In cases where the timing difference between the E-prime system (accurate to within ± 1 ms) and the recorded MER triggers could not be resolved to within ± 3ms for an epoch in a trial, the data from the epoch in that trial was not subsequently analysed. This accounted for 340 MERs or 23.6 % of the data. These events were attributed to data loss in the Leadpoint/FHC system as a result of saving the MER data to disk.

For each participant/hemisphere combination we created a single amalgamated MER consisting of each individual MER from every trial (baseline and stimulus) that were then concatenated (i.e., joined) together. While this technique of amalgamating the epochs from different trials could potentially introduce discontinuities where each MER is concatenated, it was judged that background noise statistics could be more reliably estimated using the longer amalgamated MERs.

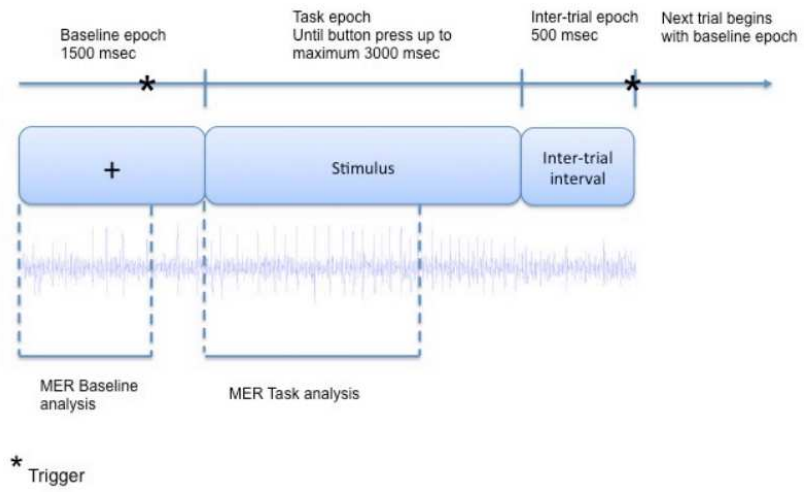


Figure A.3: Testing and recording sequence of the neuro-linguistic experiment.

A.4 Asymptotic Behaviour of RDF:

In this appendix we prove the result that the asymptotic value of the renewal density function, $m(t)$, is equal to the mean firing rate of the underlying probability distribution $p(t)$. As stated in section 3.3.1 this can be considered a special case of the more general result of the Erdos-Feller-Pollard theorem [168] when the function is sufficiently smooth such that the Laplace transform of the function and its derivative exists.

Using the Final Value Theorem for Laplace Transforms [44]:

$$\lim_{t\to\infty} m(t) = \lim_{s\to 0^+} sM(s) \tag{A.37}$$

Substituting in our expression for the Laplace transform of the renewal function (3.7) yields:

$$\lim_{t\to\infty} m(t) = \lim_{s\to 0^+} \frac{sP(s)}{1-P(s)} = \frac{0}{0} \tag{A.38}$$

Since:

$$P(s=0) = \int_0^\infty p(t)dt = 1 \tag{A.39}$$

Applying L'Hospitals rule to this limit yields:

$$\begin{aligned}
\lim_{t\to\infty} m(t) &= \lim_{s\to 0^+} \frac{\dfrac{d}{ds}[sP(s)]}{\dfrac{d}{ds}[1-P(s)]} \\
&= \lim_{s\to 0^+} \frac{P(s)}{-P'(s)} + \lim_{s\to 0^+} \frac{sP'(s)}{-P'(s)} \\
&= \lim_{s\to 0^+} \frac{P(s)}{-P'(s)} \qquad \text{(A.40)}
\end{aligned}$$

Using the following identity:

$$\begin{aligned}
\lim_{s\to 0^+} \frac{dP(s)}{ds} &= \lim_{s\to 0^+} \frac{d}{ds} \int_0^\infty p(t)e^{-st}dt \\
&= \int_0^\infty \lim_{s\to 0^+} \frac{d}{ds} p(t)e^{-st}dt \\
&= -\int_0^\infty tp(t)dt \\
&= -\mu \qquad \text{(A.41)}
\end{aligned}$$

Thus the limit of the renewal density function becomes:

$$\lim_{t\to\infty} m(t) = \frac{1}{\mu} = \nu \qquad \text{(A.42)}$$

A.5 Derivation of Power Spectrum of a Single Filtered Renewal Process

In this appendix we derive the power spectrum of a filtered renewal process. It should be made explicitly clear that this result has been known for over 50 years (see [277],[167]). Indeed a derivation using the Wiener-Khinchtine theorem was developed in [157] specifically for neural discharge spectra, although that paper only considered the case where all the action potentials have the same amplitude. The problem with these previous derivations is they are too mathematically loquacious to be accessible to (most) engineers and lose track of the physics behind the problem. This derivation is based on a periodogram approach and is an extension to Carson's theorem for shot noise processes, which was restricted to Poisson counting (i.e. memoryless) events.

Assumptions

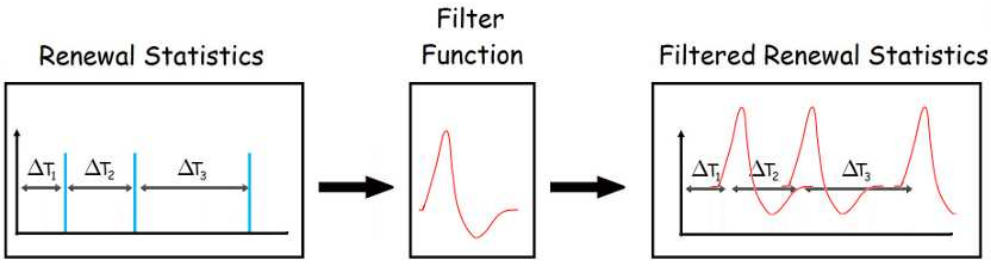


Figure A.4: Schematic of the filtered renewal process. The forcing function is a Dirac comb of impulses which follow the statistics of the renewal process. This forcing function is fed into a response filter which yields an output which is the convolution of the Dirac comb forcing function with the response shape function (for our analysis the action potential shape).

This is a *highly simplified* model of the expected neuron dynamics and in all certainty is not a truly accurate reflection of the true biological process. The following assumptions are introduced

1. The *individual neurons* follow renewal firing statistics. See Figure A.4. Recall that for renewal processes the Inter-Spike Interval (ISI) times ΔT_i form a set of *independent,*

identically distributed (i.i.d.) random variables drawn from probability distribution p(t). Note that this *does not* imply that the firing times t_i form an i.i.d. process.

2. All the STN action potentials have exactly the same shape, with the exception of a free amplitude scaling factor which is drawn from an arbitrary probability distribution. This makes our model a *marked* filtered renewal process.

3. The amplitude of an arbitrary action potential is independent of the firing times or the ISI time. This is referred to as spike shape stationarity. Note that this is a consequence of analysing the spike trains in a linear dynamical systems framework. For most circumstances this assumption will hold true, but it has been found that sometimes neurons can exhibit *rate dependent action potential shapes* during bursting events [163] for example.

4. The neurons, individually modelled as renewal processes are independent of each other. Certainly this assumption appears the most unphysical. Neurons are highly nonlinear, densely connected, possibly chaotic, deterministic oscillators. We can consider that modelling the firing times as stochastic processes is an approximation to these nonlinear coupled dynamics. Indeed in other physical systems we have seen the success of mean field theories were highly correlated systems can be treated as an ensemble of independent statistical objects. This process is frequently used in condensed matter physics with the Druid model of how a highly correlated systems of electrons and positive ions in a crystal lattice can be modelled with high accuracy as gas of independent electrons.

We model the voltage time series at the MER probe as:

$$y(t) = \sum_{m=1}^{N} a_m g(t - t_m) \tag{A.43}$$

Where the set $\{t_m\}$ are the firing times. Recall that the ISI times: $(t_m - t_{m-1})$, are i.i.d. and are drawn from a probability density function p(t).

Because we are dealing with a stochastic process which is not square integrable we have

to be careful how we define our Fourier Transforms. We formally define the Fourier Transforms in the following limiting sense:

$$\mathcal{F}\left\{y(t)\right\}(\omega) = \lim_{T\to\infty}\int_{T/2}^{T/2} y(t)e^{-i\omega t}dt \tag{A.44}$$

The Fourier Transform of (A.43) can be determined using the *shifting property* of Fourier Transforms.

$$\mathcal{F}\left\{y(t)\right\}(\omega) = Y(i\omega) = G(i\omega)\sum_{m=1}^{N} a_m e^{-i\omega t_m} \tag{A.45}$$

Where we have taken the Fourier Transform of the pulse shape $G(i\omega)$ outside of the summation because of our assumption that it is independent of the firing times.

The complex conjugate of the Fourier Transform is given by:

$$\mathcal{F}\left\{y(t)\right\}(\omega)^* = Y(i\omega)^* = G(i\omega)^*\sum_{n=1}^{N} a_n e^{+i\omega t_n} \tag{A.46}$$

The power spectral density of a stochastic process is defined as:

$$P(\omega) \;=\; E\left(\lim_{T\to\infty}\frac{1}{T}Y(i\omega)Y^*(i\omega)\right) = \lim_{T\to\infty}\frac{1}{T}E\Big(Y(i\omega)Y^*(i\omega)\Big) \tag{A.47}$$

There are subtleties associated with exchanging the limit and expectation operations which are guaranteed by the Weiner-Khinchtine theorem. The power spectrum for our renewal process generated voltage time history is thus given by:

$$P(\omega) \;=\; \lim_{T\to\infty}\frac{1}{T}E\left(G(i\omega)\sum_{m=1}^{N} a_m e^{-i\omega t_m}G(i\omega)^*\sum_{n=1}^{N} a_n e^{+i\omega t_n}\right) \tag{A.48}$$

$$\;=\; \lim_{T\to\infty}\frac{1}{T}|G(i\omega)|^2\sum_{m=1}^{N}\sum_{n=1}^{N}\langle a_m a_n e^{-i\omega(t_m-t_n)}\rangle \tag{A.49}$$

Where $\langle\cdots\rangle$ represents taking the ensemble average. Using assumption 3 that the action potential amplitudes are independent of the firing times:

$$P(\omega) = \lim_{T\to\infty}\frac{1}{T}|G(i\omega)|^2\sum_{m=1}^{N}\sum_{n=1}^{N}\langle a_m a_n\rangle\langle e^{-i\omega(t_m-t_n)}\rangle \tag{A.50}$$

We break this double summation up into three components: one where $m = n$ (which will have N terms), one where $m > n$ and one where $m < n$ (which will have combined $(N^2 - N)$ terms).

$$P(\omega) = \lim_{T\to\infty} \frac{1}{T}|G(i\omega)|^2 \left(\underbrace{\sum_{m=1}^{N}}_{m=n} + \underbrace{\sum_{m=1}^{N}\sum_{n=1}^{N}}_{m>n} + \underbrace{\sum_{m=1}^{N}\sum_{n=1}^{N}}_{m<n} \left[\langle a_m a_m\rangle \langle e^{-i\omega(t_m - t_n)}\rangle\right] \right) \tag{A.51}$$

We will tackle the $m = n$ term of the summation first:

$$P(\omega)_{m=n} = \lim_{T\to\infty} \frac{1}{T}|G(i\omega)|^2 \underbrace{\sum_{m=1}^{N}}_{m=n} \langle a_m a_m\rangle \langle e^{-i\omega(t_m - t_m)}\rangle \tag{A.52}$$

$$= \lim_{T\to\infty} \frac{1}{T}|G(i\omega)|^2 \langle a^2\rangle \cdot N \tag{A.53}$$

Where in the last equation we have assumed that all the action potential amplitudes are drawn from a common distribution. Note that in the limit of time, T, approaching infinity the number of pulses, N, will also approach infinity. At this point consider the limits of $N \to \infty$ and $T \to \infty$ such that there ratio converges to the mean firing rate ν:

$$\lim_{\substack{N\to 0 \\ T\to 0}} \frac{N}{T} = \nu \tag{A.54}$$

The limit of $N \to \infty$ arises because when we consider the process over an infinitely long time period (i.e. $T \to \infty$) there will be an infinite number of pulses. Therefore we have for the $m = n$ terms of the power spectra:

$$P(\omega)_{m=n} = \nu|G(i\omega)|^2\langle a^2\rangle \tag{A.55}$$

Will now tackle the $m > n$ terms of the double summation:

$$P(\omega)_{m\neq n} = \lim_{T\to\infty} \frac{1}{T}|G(i\omega)|^2 \underbrace{\sum_{m=1}^{N}\sum_{n=1}^{N}}_{m>n} \langle a_m a_n\rangle \langle e^{-i\omega(t_m - t_n)}\rangle \tag{A.56}$$

The key simplification introduced by assuming renewal firing times is that the difference of firing times can be written as: $t_m - t_n = (m - n)\tau$ where τ is the i.i.d. random variable

representing the ISI firing times. Thus:

$$P(\omega)_{m>n} = \lim_{T\to\infty} \frac{1}{T}|G(i\omega)|^2 \underbrace{\sum_{m=1}^{N}\sum_{n=1}^{N}}_{m>n} \langle a_m a_n\rangle \langle e^{-i\omega(m-n)\tau}\rangle \tag{A.57}$$

Notice that the constraint that: $m > n$ guarantees that $m - n > 1$ and that this double sum involving the *difference* of two counting variables $m - n$ can be written as a single sum involving a single counting variable 'k' that ranges from 1 to N:

$$\underbrace{\sum_{m=1}^{N}\sum_{n=1}^{N}}_{m>n} \langle a_m a_n\rangle \langle e^{-i\omega(m-n)\tau}\rangle \;=\; \sum_{k=1}^{N} \langle a_k\rangle\langle a_k\rangle\langle C(k)e^{-i\omega k\tau}\rangle = \langle a\rangle^2 \sum_{k=1}^{N} C(k)\langle e^{-i\omega k\tau}\rangle \tag{A.58}$$

There are *two really important* concepts to notice here:

1. We have assumed that the amplitudes associated with the old counting variables m & n are drawn from the same probability distribution, and that they are independent of each other. This is a reasonable assumption given that firing times might be correlated, the amplitudes are unlikely to be.

2. We have included a function $C(k)$ in the summation. The is the combinatoric function which accounts for the fact there are multiple copies of the $e^{-i\omega k\tau}$ term. The constraint that $m > n$ and the fact that both m and n run from 0 to N leads to the conclusion that the form of the combinatoric function is: $C(k) = N - k$.

Therefore:

$$\underbrace{\sum_{m=1}^{N}\sum_{n=1}^{N}}_{m>n} \langle a_m a_n\rangle \langle e^{-i\omega(m-n)\tau}\rangle = \langle a\rangle^2 \sum_{k=1}^{N} (N-k)\langle e^{-i\omega k\tau}\rangle \tag{A.59}$$

Now notice that the characteristic function for the Inter-Spike-Interval (ISI) distribution is given by:

$$H(\omega) = \langle e^{i\omega\tau}\rangle \quad , \quad H(\omega)^* = \langle e^{-i\omega\tau}\rangle \tag{A.60}$$

Also recall that for *independent* random variables the following property:

$$\langle e^{i\omega(\tau_1+\tau_2\ldots+\tau_k)}\rangle = H(\omega)^k \tag{A.61}$$

Now recall that one of the defining properties of a renewal process is that the ISI times are i.i.d. Thus we can write:

$$e^{-i\omega k\tau} = e^{-i\omega(\tau_1+\tau_2+\cdots\tau_k)} \tag{A.62}$$

Thus we can write out our sum in (A.59) as:

$$\underbrace{\sum_{m=1}^{N}\sum_{n=1}^{N}}_{m>n}\langle a_m a_n\rangle\langle e^{-i\omega(m-n)\tau}\rangle = \langle a\rangle^2\sum_{k=1}^{N}(N-k)\left(H(\omega)^*\right)^k \tag{A.63}$$

Algebraically re-arranging this expression:

$$\underbrace{\sum_{m=1}^{N}\sum_{n=1}^{N}}_{m>n}\langle a_m a_n\rangle\langle e^{-i\omega(m-n)\tau}\rangle = \langle a\rangle^2 N\sum_{k=1}^{N}(1-\frac{k}{N})\left(H(\omega)^*\right)^k \tag{A.64}$$

The exact same process can be performed with the $m < n$ summation term in (A.51), with the substitution that $i \to -i$ in the exponent:

$$\underbrace{\sum_{m=1}^{N}\sum_{n=1}^{N}}_{m<n}\langle a_m a_n\rangle\langle e^{-i\omega(m-n)\tau}\rangle = \langle a\rangle^2 N\sum_{k=1}^{N}(1-\frac{k}{N})\langle e^{+i\omega k\tau}\rangle \tag{A.65}$$

$$= \langle a\rangle^2 N\sum_{k=1}^{N}(1-\frac{k}{N})H(\omega)^k \tag{A.66}$$

Now we can write the power spectrum contribution from the $m \neq n$ terms of the summation. Substituting (A.64) & (A.66) into (A.51) yields:

$$\begin{aligned} P(\omega)_{m\neq n} &= \lim_{T\to\infty}\frac{1}{T}|G(i\omega)|^2\left(\underbrace{\sum_{m=1}^{N}\sum_{n=1}^{N}}_{m>n}+\underbrace{\sum_{m=1}^{N}\sum_{n=1}^{N}}_{m<n}\left[\langle a_m a_m\rangle\langle e^{-i\omega(t_m-t_n)}\rangle\right]\right) \\ &= |G(i\omega)|^2\langle a\rangle^2\lim_{T\to\infty}\frac{N}{T}\left(\sum_{k=1}^{N}(1-\frac{k}{N})\left(H(\omega)^*\right)^k+\sum_{k=1}^{N}(1-\frac{k}{N})H(\omega)^k\right) \end{aligned} \tag{A.67}$$

At this point, similar to the $m = n$ case we take the limit of N going to infinity and T going to infinity such that:

$$\lim_{T\to\infty}\lim_{N\to\infty}\frac{N}{T} = \nu \tag{A.68}$$

Notice that in this limit:

$$\lim_{N\to\infty}\sum_{k=1}^{N}(1-\frac{k}{N})H(\omega)^k \approx \sum_{k=1}^{\infty}H(\omega)^k = \frac{H(\omega)}{1-H(\omega)} \tag{A.69}$$

Where we have used the geometric series formula in the last equation above. Therefore the power spectrum contribution from the $m \neq n$ terms of the summation can be written as:

$$\begin{aligned} P(\omega)_{m\neq n} &= \nu|G(i\omega)|^2\langle a\rangle^2\left[\left(\frac{H(\omega)}{1-H(\omega)}\right)+\left(\frac{H(\omega)}{1-H(\omega)}\right)^*\right] &\text{(A.70)}\\ &= 2\nu|G(i\omega)|^2\langle a\rangle^2 Re\left\{\frac{H(\omega)}{1-H(\omega)}\right\} &\text{(A.71)}\end{aligned}$$

Thus the power spectrum of the filtered renewal process can be obtained by adding the contribution from the $m = n$ (A.55) & the $m \neq n$ (A.71) terms of the power spectrum:

$$\begin{aligned} P(\omega) &= P(\omega)_{m=n} + P(\omega)_{m\neq n}\\ &= \nu|G(i\omega)|^2\left(\langle a^2\rangle + 2\langle a\rangle^2 Re\left\{\frac{H(\omega)}{1-H(\omega)}\right\}\right) &\text{(A.72)}\end{aligned}$$

If we now restrict the amplitudes of the individual neuron to be a constant value: $\langle a\rangle^2 = \langle a^2\rangle = 1$. Also using the notation that: $|G(i\omega)|^2 \equiv G(\omega)$ the power spectrum for the individual renewal process is given by:

$$P(\omega) = \nu G(\omega)\left(1 + 2Re\left\{\frac{H(\omega)}{1-H(\omega)}\right\}\right) \tag{A.73}$$

This is the power spectral density given by (3.16) in Section 3.3.1.

A.6 Spectrum of Ensemble of i.i.d Renewal Processes is Scaled form of Individual Process

In this appendix we derive the power spectrum of a sum of independent filtered renewal processes which each have a distinct, but constant, amplitude. As discussed in section 3.4, this derivation and the implication that the resulting PSD of the superposition of identical renewal processes is a scaled form of PSD of the constituent processes was developed in [158]. In this section we perform the trivial extension of developing the spectrum when the constituent renewal processes have distinct, but constant amplitudes.

Let the time series voltage at the MER be a linear combination of the voltages from the individual neurons with variable amplitude as given by (3.23):

$$y(t)_\Sigma = \sum_{n=1}^{N} \alpha_n y_n(t) \tag{A.74}$$

The autocorrelation function of the MER time series is given by:

$$\begin{aligned} R_{y_\Sigma}(t) &= \langle y_\Sigma(t) y_\Sigma(t-t') \rangle - \langle y_\Sigma(t) \rangle \langle y_\Sigma(t-t') \rangle \\ &= \langle \sum_{\mathrm{n,l}} \alpha_n \alpha_l y_n(t) y_l(t-t') \rangle - \sum_{\mathrm{n,l}} \langle \alpha_n y_n(t) \rangle \langle \alpha_l y_l(t) \rangle \end{aligned} \tag{A.75}$$

We separate the double summation in both terms into N identical terms where the n and l indices are identical and another where the indices are distinct. We then invoke both the statistical independence of the firing times of different renewal processes and the independence of the amplitude and firing times of a single renewal process.

$$\begin{aligned} R_{y_\Sigma}(t) &= N\langle \alpha^2 \rangle \left(\langle y(t) y(t-t') \rangle - \langle y(t) \rangle \langle y(t-t') \rangle \right) \\ &+ \sum_{\substack{n,l \\ n \neq l}} \langle \alpha_n y_n(t) \rangle \langle \alpha_l y_l(t) \rangle - \sum_{\substack{n,l \\ n \neq l}} \langle \alpha_n y_n(t) \rangle \langle \alpha_l y_l(t) \rangle \\ &= N\langle \alpha^2 \rangle R_y(t) \end{aligned} \tag{A.76}$$

Where $\langle \alpha^2 \rangle$ is the second moment of the amplitude distribution which is independent of the firing times. Using the Weiner-Khinchtine theorem the power spectrum at the MER due

to the superposition of iio ensemble renewal oscillators is given by:

$$P_{\Sigma}(\omega) = \langle \alpha^2 \rangle N P(\omega) \tag{A.77}$$

A.7 Spectral Equivalence of DPIM and renewal theory

In this Appendix we develop the power spectrum for a neuron that encodes information in the gaps between its spikes as a Digital Pulse Interval Modulation encoder. We consider the situation which (unlike telecommunications protocols) the pulses can occur at any time on the continuous timeline. We show that the resulting spectrum in this limit is identical to that of the renewal processes considered in chapter 3.

Digital Pulse Interval Modulation (DPIM) is a communications scheme where information content is transmitted in the variable number of idle slot times between pulses [152]. The DPIM information transmission methodology is ideally suited for optical fibre communications where classical amplitude or frequency modulation techniques are non-trivial. The wide bandwidth offered by optical fibre motivates the use of particularly narrow pulses with low duty cycles which allows for low average, but high peak power. This property provides accurate signal detection at the receiver end during transmission over a noisy channel [153]. Carrying the information between the pulses also eliminates the need for the pulse times to be defined relative to a central clock.

These properties of robust transmission over noise, lack of a central clock and minimal energy expenditure are also ideal for information transmission of neurons. Indeed the concept of neurons transmitting information in the space between firing events has been considered as far back as [278] where it was shown that a DPIM coding scheme offers a far greater channel capacity than a binary on/off keying (OOK) coding scheme. Fundamentally the debate about rate vs time dependent coding is about *how* neurons encode information in the timing between spikes. Nevertheless there is a key difference between neuron information exchange and DPIM on communications channels. DPIM processes are discrete time processes embedded in continuous time. Although the signals being sent and received are occurring in continuous time, the pulse intervals are constrained to be separated by a discrete number of packets, termed slots or chips. Neurons have no such embedding and the firing times can occur at any point along the continuous time line. This issue is philosophically complicated

by the fact that at some fine enough level of temporal resolution a neuron must consider a spike arriving at two different times to be the same spike.

The time intervals between pulses in DPIM are driven by some encoding strategy and thus is strictly speaking deterministic. Nonetheless from the receiver's perspective, who does not know the message a priori, the series of gaps between the pulses is a random process. Neurons are most likely similar in the sense that the action potential timings are generated by highly nonlinear processes, with exceptionally rich and complex synaptic/dendritic connections, such that from the 'perspective' of a neuron the arriving pulse times are random variables. This is the basis for mathematical modeling of neuron firing times as stochastic point processes.

The time history of DPIM, $v(t)$, can be represented as follows [152]:

$$v(t) = \sum_{m=-\infty}^{m=+\infty} g(t - \tau_m T) \tag{A.78}$$

Where $g(t)$ is the pulse shape, characteristically a rectangle with a given duty cycle, T is the slot length and τ_m is a set of random integers indicating when a pulse occurs in terms of numbers of slots.

The power spectrum of this process contains both a continuous and discrete part which are both functions of the energy spectrum, $G(f)$, of the pulse shape. They are also both dependent on the average pulse rate, ν_b, which is the inverse of the expected length (in number of slots) between pulses. The continuous component, $R_c(fT)$, termed the *Bartlett spectrum* in point process literature [279], is best understood from the Weiner-Khinchtine theorem as the Fourier transform of the autocorrelation structure of the pulse times. The discrete component forms a Dirac comb with Dirac delta distributions $\delta(\cdots)$ spaced $f_m = m/T, (m \in Z)$ apart. This discrete component arises because the expectation of the signal is non zero ($\mathbb{E}[v(t)] \neq 0$) [280]. Intuitively the repetition of the Dirac delta pulses can be understood because the discrete process is embedded in the continuous time, creating an effective Nyquist frequency (the inverse of the slot time) for the frequency structure to be periodic about. The

spectrum of the DPIM process is determined in [152] as:

$$\begin{aligned} S_{\text{DPIM}}(f) &= \underbrace{\frac{1}{T}\left|G(f)\right|^2 R_{c,\text{DPIM}}(fT)}_{\text{Continuous}} \\ &+ \underbrace{\sum_{m=-\infty}^{+\infty} \frac{1}{T^2}\left|G(f)\right|^2 \nu_b^2 \delta(f - f_m)}_{\text{Discrete}} \end{aligned} \tag{A.79}$$

The *Bartlett spectrum* is given in [152] by the following evaluation of the Z transform on the unit circle:

$$\underbrace{R_{c,\text{DPIM}}(u)}_{\text{Bartlett Spectrum}} = \nu_b \left[\nu_b - 1 + 2\text{Re}\left(X(z)\Big|_{z=e^{2\pi j u}}\right)\right] \tag{A.80}$$

Where $X(z)$ can be understood to be the Z transform of the cumulative probabilities associated with the spacing (in number of slots) between two arbitrary pulses. We use an alternate definition of $X(z)$ defined in [281]:

$$X(z) = 1 + X(z)\sum_{\lambda=0}^{\infty} p[\lambda] z^{-\lambda} - \nu_b \sum_{\lambda=0}^{\infty} p[\lambda] \sum_{k=0}^{\lambda-1} z^{-k} \tag{A.81}$$

Where $p[\lambda]$ represents the probability of the pulse interval being λ slots long. We algebraically re-arrange (A.81) and evaluate the z transform on the unit circle as defined in (A.80) to show that this term depends on the characteristic function associated with the pulse arrival time random variables:

$$\begin{aligned} X(e^{j2\pi u}) &= \frac{1}{1 - \sum_{\lambda=0}^{\infty} p[\lambda] e^{-2\pi j u \lambda}} \\ &\quad -\nu_b \cdot \underbrace{\left(\frac{\sum_{\lambda=0}^{\infty} p[\lambda] \sum_{k=0}^{\lambda-1} e^{-2\pi j u k}}{1 - \sum_{\lambda=0}^{\infty} p[\lambda] e^{-2\pi j u \lambda}}\right)}_{\text{simplify}} \end{aligned} \tag{A.82}$$

We will now develop a simplified expression for the bracketed term in (A.82) (labelled simplify) by recognizing that the nested sum in the numerator can be re-written *except at*

zero frequency using the geometric series formula and identifying that the probability mass function, $p[\lambda]$, must sum to unity:

$$\begin{aligned}\sum_{\lambda=0}^{\infty} p[\lambda] \sum_{k=0}^{\lambda-1} e^{-2\pi juk} &= \sum_{\lambda=0}^{\infty} p[\lambda] \left(\frac{1 - e^{-2\pi ju\lambda}}{1 - e^{-2\pi ju}} \right) \\ &= \frac{\left(1 - \sum_{\lambda=0}^{\infty} p[\lambda] e^{-2\pi ju\lambda} \right)}{1 - e^{-2\pi ju}} \end{aligned} \tag{A.83}$$

Substituting (A.83) into the numerator of the bracketed part of (A.82) (labelled simplify) and cancelling the common denominator & numerator term yields the following simpler expression:

$$\left(\frac{\sum_{\lambda=0}^{\infty} p[\lambda] \sum_{k=0}^{\lambda-1} e^{-2\pi juk}}{1 - \sum_{\lambda=0}^{\infty} p[\lambda] e^{-2\pi ju\lambda}} \right) = \frac{1}{1 - e^{-2\pi ju}} \tag{A.84}$$

Therefore:

$$X(e^{j2\pi u}) = \frac{1}{1 - \sum_{\lambda=0}^{\infty} p[\lambda] e^{-2\pi ju\lambda}} - \nu_b \cdot \frac{1}{1 - e^{-2\pi ju}} \tag{A.85}$$

Note the identity:

$$2Re \left(\frac{1}{1 - e^{-2\pi ju}} \right) = 1 \tag{A.86}$$

Thus taking the real part of (A.85) and using (A.86) yields:

$$2\mathrm{Re}\left[X(e^{2\pi ju}) \right] = 2\mathrm{Re} \left(\frac{1}{1 - \sum_{\lambda=0}^{\infty} p[\lambda] e^{-2\pi ju\lambda}} \right) - \nu_b \tag{A.87}$$

Substituting (A.87) into (A.80), adding and subtracting unity yields:

$$R_{c,\mathrm{DPIM}}(u) = \nu_b \left[1 + 2\mathrm{Re} \left(\frac{\sum_{\lambda=0}^{\infty} p[\lambda] e^{-2\pi ju\lambda}}{1 - \sum_{\lambda=0}^{\infty} p[\lambda] e^{-2\pi ju\lambda}} \right) \right] \tag{A.88}$$

Thus we can write the power spectrum for the DPIM scheme in [152] as:

$$\begin{aligned} S_{\text{DPIM}}(f) &= \frac{\nu_b}{T}|G(f)|^2\Bigg[\sum_{m=-\infty}^{+\infty}\frac{\nu_b}{T}\delta(f-\frac{m}{T}) \\ &+ 1+2\text{Re}\left(\frac{\sum_{\lambda=0}^{\infty}p[\lambda]e^{-2\pi jf\lambda T}}{1-\sum_{\lambda=0}^{\infty}p[\lambda]e^{-2\pi jf\lambda T}}\right)\Bigg] \end{aligned} \tag{A.89}$$

Note that if the pulse slot time approach zero, the expected gap (in units of the number of slots) between pulses will approach infinity. Thus the average rate of pulses (in units of the inverse of the number of slots) ν_b, will also approach zero. These limits will both converge to zero such that their ratio remains the statistical average rate of the number of pulses, $N(t)$, in units of time:

$$\lim_{\substack{T\to 0\\ \nu_b\to 0}}\frac{\nu_b}{T}=\nu_\cdot=\lim_{t\to\infty}\mathbb{E}\left(\frac{N(t)}{t}\right) \tag{A.90}$$

Note that as the pulse slot time approaches zero the discrete component of the spectrum will change from a train of Dirac delta pulses (spaced 1/T apart) to a single Dirac delta pulse (the $m = 0$ solution) centred at zero frequency:

$$\lim_{\substack{T\to 0\\ \nu_b\to 0}}\sum_{m=-\infty}^{+\infty}\frac{\nu_b^2}{T^2}\delta(f-\frac{m}{T})\longrightarrow\nu^2\delta(f) \tag{A.91}$$

Also note that reducing the pulse slot length to zero allows the pulse times to occur at any point on the continuous time line. Thus the characteristic function will transform from a periodic DTFT for a discrete random variable to a non-periodic continuous Fourier transform for a continuous random variable:

$$\sum_{\lambda=0}^{\infty}p[\lambda]e^{-2\pi ju\lambda}\longrightarrow\int_0^{+\infty}p(t)e^{-2\pi ft}dt\equiv H(\omega) \tag{A.92}$$

Where $H(\omega)$ is the characteristic function of the Inter Spike Interval (ISI) of firing times. Thus the DPIM spectrum in the continuum reduces to:

$$\lim_{\substack{T\to 0\\ \nu_b\to 0}} S_{\mathrm{DPIM}}(f) \;=\; \nu|G(f)|^2\Bigg[\underbrace{1+2\mathrm{Re}\left(\frac{H(\omega)}{1-H(\omega)}\right)}_{R_{c,DPIM}}+\nu\delta(f)\Bigg] \tag{A.93}$$

A.7.1 Spectral Effect of Jittering the Neural DPIM Spectra

In this section we explore how robust the spectral features of a neuron following renewal statistics are to noise. The noise sources on a neuron in a network are exceptionally complex. We model noise effects to the PIM encoding scheme in the simplest possible fashion of Gaussian displacements of the firing times. In the most general sense we can consider that noise will introduce variation to the membrane voltage, which in turn will introduce variation to when the neuron reaches threshold and fires, thus jittering the firing times. These ideas were formalised in [282] with a leaky integrate and fire model of a neuron subject to Gaussian white noise input. We show that the effect of the jittering is to attenuate, but not distort, the non-Poisson features of the Bartlett spectrum.

We consider the effect of the jittering on the spectral properties of the Bartlett spectrum rather than the full time series power spectrum. This is equivalent to considering the action potentials to be Dirac delta pulses ($g(t)=\delta(t)$, $|G(f)|^2=1$). This spectrum can be determined with aid from the Fundamental Isometry Theorem. Briefly this theorem allows for the determination of the spectrum of a *marked* point process when the spectrum of the original process and the statistics of the *marks* are known. It can be shown [277] that the jittering is easily accounted for as follows:

$$R_{\mathrm{c,DPIM}}(\omega)_{\mathrm{jittered}} = |\phi_z(2\pi f)|^2 R_{\mathrm{c,DPIM}}(\omega) + \nu(1-|\phi_z(2\pi f)|^2) \tag{A.94}$$

We consider that the firing times are independently jittered by a Gaussian random variable with mean μ_{J} and variance σ_{J}^2. Thus the absolute value squared of the characteristic function is $|\phi_z(2\pi f)|^2 = e^{-\sigma_j^2 f^2}$. Note that the effect of jittering is to randomise firing patterns, remove correlation structure and thus make the observed firing times more uniform.

Using (A.94) the Gaussian jittered Bartlett spectrum is given by:

$$R_{\text{c,DPIM}}(\omega)_{\text{jittered}} = \nu \left[1 + 2e^{-\sigma_j^2\omega^2}\text{Re}\left(\frac{H(\omega)}{1-H(\omega)}\right)\right] \tag{A.95}$$

There are two points to notice about (A.95). Firstly jittering removes the non-Poisson structure, which represents deviations from uniformly distributed patterns, from the Bartlett spectrum at a rate exponentially proportional to the jittering variance. Secondly jittering has no effect on the Bartlett spectrum of a Poisson process, which already has uniform distributed firing times.

Figure A.5 shows the jittered Bartlett spectrum of Weibull processes with shape parameters of k=5 (pseudo-periodic) & k=10 (strongly periodic) for different strengths of jittering ($\sigma_J = 0.01, 0.025, 0.05$ seconds). As a comparison the time constant of a leaky integrate and fire model of a Sub Thalamic Nucleus is 0.01 seconds [182]. Notice that, as expected, the larger the variance of jitter the more the features of the Bartlett spectra are reduced but not distorted. Notice also that the spectra of the more periodic Weibull process is more robust to jittering. This intuitively makes sense, as we expect it would require jitter of a higher variance to transform the more 'strongly' periodic processes to a maximally random Poisson process.

This result shows that the spectra of both the maximum (Poisson) and minimum (purely periodic) information entropy firing distributions are highly resistant to firing time jitter, whereas patterns in between these extremes are sensitive. This is especially true for nearly Poisson ($k \approx 1$) firing patterns. This is interesting for neural spectral feature selection because neurons can have highly variable firing patterns under different anatomical and physiological conditions. For example cortical neurons alone can display firing statistics ranging from Poisson to weakly periodic depending on anatomical area, anaesthetic state and behavioural task [283].

A.8 Simplification of Residual Correlations

In this section we show how the calculation of the residual correlations required for the homotopy algorithm (section 4.4.1) and the InCrowd algorithm (section 4.4.2) can be efficiently computed.

We can speed up the calculation of the residual correlations by exploiting the fact that our system matrix, $\mathbf{A}$, is a concatenation of convolution matrices. Notice that for a single convolution matrix, $\mathbf{A}_i$ with a single solution vector $\mathbf{x}_i$ we can write the matrix vector product using Fourier Transforms as:

$$\mathbf{A}_i\mathbf{x}_i = \mathcal{F}^{-1}\left\{\mathcal{F}\left\{\mathbf{g}_i\right\}(\omega)\cdot\mathcal{F}\left\{\mathbf{x}_i\right\}(\omega)\right\}(t) \tag{A.96}$$

With the use of the Fast Fourier Transform this reduces our operation count from roughly $2M^2$ to $Mlog(M)$. Similarly we can write the correlation of the entire system with an estimate of the firing time vector $\hat{\mathbf{x}}$ as:

$$(\mathbf{y}-\mathbf{A}\hat{\mathbf{x}}) = \underbrace{\mathbf{y}-\sum_{i=1}^{M}\left(\mathcal{F}^{-1}\left\{\mathcal{F}\left\{\mathbf{g}_i\right\}(\omega)\cdot\mathcal{F}\left\{\hat{\mathbf{x}}_i\right\}(\omega)\right\}(t)\right)}_{\mathbf{z}} \tag{A.97}$$

We now work out the residual correlation. Due to the re-ordering introduced by the transpose of the system matrix, we must re-order elements. Introducing the transformation operator: $\hat{\mathbf{T}}\left[(x_0,x_1,x_2,\cdots,x_N)\right] = (x_N,x_{N-1},x_{N-2},\cdots,x_0)$, we can write the residual correlation associated with $\mathbf{A}_i$ as:

$$\mathbf{c}_i = \hat{\mathbf{T}}\left[\mathcal{F}^{-1}\left\{\mathcal{F}\left\{g_i\right\}(\omega)\cdot\hat{\mathbf{T}}\left[\mathcal{F}\left\{\mathbf{z}\right\}(\omega)\right]\right\}(t)\right] \tag{A.98}$$

The residual correlation can now be written as:

$$\mathbf{C} = \mathbf{A}^{\mathbf{T}}(\mathbf{A}\mathbf{x}-\mathbf{y}) = \left[\mathbf{c}_1^T,\mathbf{c}_2^T,\cdots\mathbf{c}_M\right]^{\mathbf{T}} \tag{A.99}$$

A.9 Deriving NMP for a Superposition of Renewal Processes

In this Appendix we develop closed form expressions for the full spectrum and zero frequency value of the first NMP, $\epsilon_1(\omega)$, for a superposition of renewal processes.

Recall from (2.13) & (2.29) that the first generalised NMP is given by:

$$\epsilon_1(\omega) = \frac{\sqrt{\Lambda_0}}{2}\sqrt{P(\omega)^2 + \mathcal{H}\left\{P(\omega)\right\}(\omega)^2} \tag{A.100}$$

$$\text{Where:} \quad \Lambda_0 = \frac{1}{2\pi}\int_{-\infty}^{\infty} \omega^2 P(\omega) d\omega \tag{A.101}$$

Recall from (3.24) that the resulting power spectrum of a super-position of renewal processes, $P_\Sigma(\omega)$, is given by:

$$P_\Sigma(\omega) = N\nu\langle a^2\rangle G(\omega)\left[1 + 2Re\left\{\frac{H(\omega)}{1-H(\omega)}\right\}\right]$$

Recall from section 2.4 that the NMP requires a power spectrum normalised such that $\int_{-\infty}^{+\infty} P(\omega)d\omega = 2\pi$. Therefore the normalised spectrum is given by:

$$P_{\Sigma,\text{norm}}(\omega) = \frac{2\pi G(\omega)\left[1 + 2Re\left\{\frac{H(\omega)}{1-H(\omega)}\right\}\right]}{\int_{-\infty}^{\infty} G(\omega)\left[1 + 2Re\left\{\frac{H(\omega)}{1-H(\omega)}\right\}\right]d\omega} \tag{A.102}$$

The second relaxation parameter for a super-position of renewal processes is therefore given by combining (A.101) & (A.102):

$$\Lambda_0 = \frac{\int_{-\infty}^{+\infty} \omega^2 G(\omega)\left[1 + 2Re\left\{\frac{H(\omega)}{1-H(\omega)}\right\}\right]d\omega}{\int_{-\infty}^{+\infty} G(\omega)\left[1 + 2Re\left\{\frac{H(\omega)}{1-H(\omega)}\right\}\right]d\omega} \tag{A.103}$$

Therefore we may calculate the generalised NMP associated with a super-position of

renewal processes by combining (A.100),(A.103) & (A.102):

$$\epsilon_{1,\Sigma}(\omega) = \pi \cdot \sqrt{\frac{\int_{-\infty}^{+\infty} \omega^2 G(\omega)\left[1+2Re\left\{\frac{H(\omega)}{1-H(\omega)}\right\}\right]d\omega}{\left(\int_{-\infty}^{+\infty} G(\omega)\left[1+2Re\left\{\frac{H(\omega)}{1-H(\omega)}\right\}\right]d\omega\right)^2}} \times \sqrt{\left(G(\omega)\left[1+2Re\left\{\frac{H(\omega)}{1-H(\omega)}\right\}\right]\right)^2 + \mathcal{H}\left\{G(\omega)\left[1+2Re\left\{\frac{H(\omega)}{1-H(\omega)}\right\}\right]\right\}(\omega)^2} \tag{A.104}$$

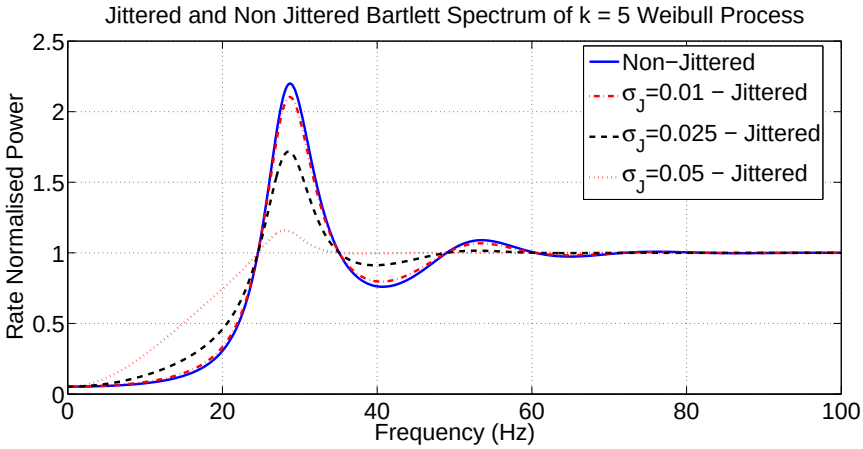
This expression (A.104) is far too complicated to allow simple insight into how the firing statistics alter the NMP spectrum, beyond identifying that this spectrum is explicitly a function of these firing statistics. We consider the simpler case of the Zero Frequency component of the first NMP ($\epsilon_{1,\Sigma}(0)$), which we used to compare states of the STN in sections 2.9.3-2.9.4.

By taking the Maclaurin series approximation (to second order) of the correlation function it can be shown [157]:

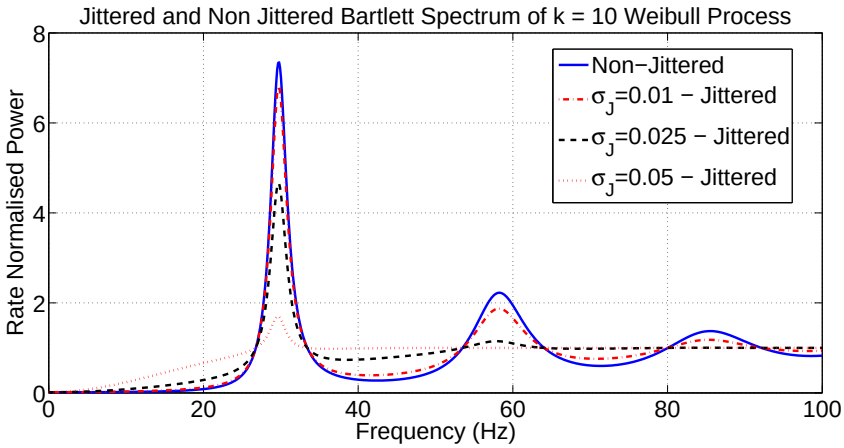
$$\lim_{\omega\to 0}\left(1+2Re\left\{\frac{H(\omega)}{1-H(\omega)}\right\}\right) = \nu^2\sigma^2 = C_v^2, \tag{A.105}$$

where, ν is the mean firing rate, σ^2 is the variance of the ISI and C_v is the coefficient of variation of the ISI. Using (A.104) and (A.105) the zero frequency component, $\epsilon_{1,\Sigma}(0)$, of the first NMP is given by:

$$\epsilon_{1,\Sigma}(0) = \pi G(0) C_v^2 \cdot \sqrt{\frac{\int_{-\infty}^{+\infty} \omega^2 G(\omega)\left[1+2Re\left\{\frac{H(\omega)}{1-H(\omega)}\right\}\right]d\omega}{\left(\int_{-\infty}^{+\infty} G(\omega)\left[1+2Re\left\{\frac{H(\omega)}{1-H(\omega)}\right\}\right]d\omega\right)^2}} \tag{A.106}$$



(a) $k = 5$.



(b) $k = 10$

Figure A.5: Rate normalised jittered Bartlett Spectrum of a Weibull process with shape parameters $(k) = 5$ & 10, mean firing rate of 30Hz for different strengths of jittering (σ_J).

References

[1] C. R. Camalier, P. E. Konrad, C. E. Gill, C. Kao, M. R. Remple, H. M. Nasr, T. L. Davis, P. Hedera, F. T. Phibbs, A. L. Molinari, *et al. Methods for surgical targeting of the stn in early-stage parkinson's disease.* Front. Neurol **5**, 25 (2014).

[2] N. R. Williams, K. D. Foote, and M. S. Okun. *Subthalamic nucleus versus globus pallidus internus deep brain stimulation: Translating the rematch into clinical practice.* Movement Disorders Clinical Practice **1**(1), 24 (2014).

[3] D. Servello, M. Porta, M. Sassi, A. Brambilla, and M. M. Robertson. *Deep brain stimulation in 18 patients with severe gilles de la tourette syndrome refractory to treatment: the surgery and stimulation.* Journal of Neurology, Neurosurgery & Psychiatry **79**(2), 136 (2008).

[4] V. Sturm, D. Lenartz, A. Koulousakis, H. Treuer, K. Herholz, J. C. Klein, and J. Klosterkötter. *The nucleus accumbens: a target for deep brain stimulation in obsessive–compulsive-and anxiety-disorders.* Journal of chemical neuroanatomy **26**(4), 293 (2003).

[5] G. E. Gmel, T. J. Hamilton, M. Obradovic, R. B. Gorman, P. S. Single, H. J. Chenery, T. Coyne, P. A. Silburn, and J. L. Parker. *A new biomarker for subthalamic deep brain stimulation for patients with advanced parkinson's disease - a pilot study.* Journal of Neural Engineering **12**(6), 1 (2015).

[6] M. Okun. *Deep-brain stimulation for parkinson's disease.* The New England journal of medicine **367**(16), 1529 (2012).

[7] G. Deuschl, C. Schade-Brittinger, P. Krack, J. Volkmann, H. Schäfer, K. Bötzel, C. Daniels, A. Deutschländer, U. Dillmann, W. Eisner, *et al. A randomized trial of deep-brain stimulation for parkinson's disease.* New England Journal of Medicine **355**(9), 896 (2006).

[8] A. Williams, S. Gill, T. Varma, C. Jenkinson, N. Quinn, R. Mitchell, R. Scott, N. Ives, C. Rick, J. Daniels, *et al. Deep brain stimulation plus best medical therapy versus best medical therapy alone for advanced parkinson's disease (pd surg trial): a randomised, open-label trial.* The Lancet Neurology **9**(6), 581 (2010).

[9] F. M. Weaver, K. Follett, M. Stern, K. Hur, C. Harris, W. J. Marks, J. Rothlind, O. Sagher, D. Reda, C. S. Moy, *et al. Bilateral deep brain stimulation vs best medical therapy for patients with advanced parkinson disease: a randomized controlled trial.* Jama **301**(1), 63 (2009).

[10] M. Katz, C. Kilbane, J. Rosengard, R. L. Alterman, and M. Tagliati. *Referring patients for deep brain stimulation: an improving practice.* Archives of Neurology **68**(8), 1027 (2011).

[11] L. Berry. Personal Communication using Medtronic Data on File (2015).

[12] D. A. Economics. *Living with Parkinson's Disease: An updated economic analysis.* Tech. rep., Deloitte Access Economics (2014).

[13] I. Mavridis, E. Boviatsis, and S. Anagnostopoulou. *Anatomy of the human subthalamic nucleus: A combined morphometric study.* Anatomy research international (2013).

[14] T. Coyne, P. Silburn, R. Cook, P. Silberstein, G. Mellick, F. Sinclair, G. Fracchia, D. Wasson, and P. Stanwell. *Rapid subthalamic nucleus deep brain stimulation lead placement utilising ct/mri fusion, microelectrode recording and test stimulation.* Acta Neurochir Suppl. **99**, 49 (2006).

[15] J. L. Shils, M. Tagliati, and R. L. Alterman. *Chapter 16 - neurophysiological monitoring during neurosurgery for movement disorders.* In V. D. L. Shils, ed., *Neurophysiology in Neurosurgery*, pp. 405 – 448 (Academic Press, San Diego, 2002).

[16] A. Benazzouz, S. Breit, A. Koudsie, P. Pollak, P. Krack, and A.-L. Benabid. *Intraoperative microrecordings of the subthalamic nucleus in parkinson's disease.* Movement disorders **17**(S3), S145 (2002).

[17] W. J. Marks Jr. *Deep brain stimulation management* (Cambridge University Press, 2015).

[18] A. R. Kent and W. M. Grill. *Analysis of deep brain stimulation electrode characteristics for neural recording.* Journal of neural engineering **11**(4), 046010 (2014).

[19] M. J. Rasch, A. Gretton, Y. Murayama, W. Maass, and N. K. Logothetis. *Inferring spike trains from local field potentials.* Journal of Neurophysiology **99**(3), 1461 (2008).

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

[21] P. L. Nunez and R. Srinivasan. *Electric fields of the brain: the neurophysics of EEG* (Oxford University Press, USA, 2006).

[22] A. Kamondi, L. Acsady, X. Wang, and G. Buzsaki. *Theta oscillations in somata and dendrites of hippocampal pyramidal cells in vivo: Activity-dependent phase-precession of action potentials.* Hippocampus **8**(3), 244 (1998).

[23] G. Buzsaki, R. G. Bickford, G. Ponomareff, L. Thal, R. Mandel, and F. H. Gage. *Nucleus basalis and thalamic control of neocortical activity in the freely moving rat.* The Journal of neuroscience **8**(11), 4007 (1988).

[24] C. Bedard, H. Kroger, and A. Destexhe. *Modeling Extracellular Field Potentials and the Frequency-Filtering Properties of Extracellular Space.* Biophysical Journal **86**(3), 1829 (2004).

[25] M. Bazhenov, P. Lonjers, S. Skorheim, C. Bedard, and A. Destexhe. *Non-homogeneous extracellular resistivity affects the current-source density profiles of up–down state oscillations.* Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences **369**(1952), 3802 (2011).

[26] C. C. Chen, A. Pogosyan, L. U. Zrinzo, S. Tisch, P. Limousin, K. Ashkan, T. Yousry, M. I. Hariz, and P. Brown. *Intra-operative recordings of local field potentials can help localize the subthalamic nucleus in parkinson's disease surgery.* Experimental neurology **198**(1), 214 (2006).

[27] J. S. Winestone, A. Zaidel, H. Bergman, and Z. Israel. *The use of macroelectrodes in recording cellular spiking activity.* Journal of neuroscience methods **206**(1), 34 (2012).

[28] M. Remple, C. Bradenham, C. Kao, P. Charles, J. Neimat, and P. Konrad. *Subthalamic nucleus neuronal firing rate increases with parkinson's disease progression. movement disorders?* official journal of the Movement Disorder Society **26**(9), 1657 (2011).

[29] M. Magnin, A. Morel, and D. Jeanmonod. *Single-unit analysis of the pallidum, thalamus and subthalamic nucleus in parkinsonian patients.* Neuroscience **96**(3), 549 (2000).

[30] N. E. Crone, A. Korzeniewska, and P. J. Franaszczuk. *Cortical gamma responses: searching high and low.* International Journal of Psychophysiology **79**(1), 9 (2011).

[31] K. J. Miller, E. C. Leuthardt, G. Schalk, R. P. Rao, N. R. Anderson, D. W. Moran, J. W. Miller, and J. G. Ojemann. *Spectral changes in cortical surface potentials during motor movement.* The Journal of neuroscience **27**(9), 2424 (2007).

[32] C. Lambert, L. Zrinzo, Z. Nagy, A. Lutti, M. Hariz, T. Foltynie, B. Draganski, J. Ashburner, and R. Frackowiak. *Confirmation of functional zones within the human subthalamic nucleus: patterns of connectivity and sub-parcellation using diffusion weighted imaging.* Neuroimage **60**(1), 83 (2012).

[33] E. Marani, T. Heida, E. Lakke, and K. Usunoff. *The Subthalamic Nucleus Part I: Development, Cytology, Topography and Connections* (Springer Berlin Heidelberg, 2008).

[34] J. H. Martin. *Neuroanatomy Text and Atlas* (McGraw Hill, 2012).

[35] A. L. Hodgkin and A. F. Huxley. *A quantitative description of membrane current and its application to conduction and excitation in nerve.* The Journal of physiology **117**(4), 500 (1952).

[36] K. Aihara and G. Matsumoto. *Chaotic oscillations and bifurcations in squid giant axons.* In H. A. V., ed., *Chaos*, pp. 257–269 (Manchester University Press, Manchester, 1986).

[37] G. Matsumoto, K. Aihara, Y. Hanyu, N. Takahashi, S. Yoshizawa, and J.-i. Nagumo. *Chaos and phase locking in normal squid axons.* Physics Letters A **123**(4), 162 (1987).

[38] K. Aihara, G. Matsumoto, and M. Ichikawa. *An alternating periodic-chaotic sequence observed in neural oscillators.* Physics Letters A **111**(5), 251 (1985).

[39] C. Dangerfield, D. Kay, and K. Burrage. *Stochastic models and simulation of ion channel dynamics.* Procedia Computer Science **1**(1), 1587 (2010).

[40] B. Rosin, M. Slovik, R. Mitelman, M. Rivlin-Etzion, S. Haber, I. Z., E. Vaadia, and H. Bergman. *Closed-loop deep brain stimulation is superior in ameliorating parkinsonism.* Neuron **722**, 370 (2011).

[41] W. Schuepbach, J. Rau, K. Knudsen, J. Volkmann, P. Krack, L. Timmermann, T. Hälbig, H. Hesekamp, S. Navarro, N. Meier, *et al. Neurostimulation for parkinson's disease with early motor complications.* New England Journal of Medicine **368**(7), 610 (2013).

[42] E. Moro, Y. Y. Poon, A. M. Lozano, J. A. Saint-Cyr, and A. E. Lang. *Subthalamic nucleus stimulation: improvements in outcome with reprogramming.* Archives of neurology **63**(9), 1266 (2006).

[43] M. K. Hosain, A. Kouzani, and S. Tye. *Closed loop deep brain stimulation: an evolving technology.* Australasian Physical & Engineering Sciences in Medicine **37**(4), 619 (2014).

[44] N. S. Nise. *Control Systems Engineering* (John Wiley & Sons, Inc., New York, NY, USA, 2000), 3^{rd} ed.

[45] S. Little and P. Brown. *What brain signals are suitable for feedback control of deep brain stimulation in parkinson's disease?* Annals of the New York Academy of Sciences **1265**(1), 9 (2012).

[46] M. Kilpatrick, E. Church, S. Danish, M. Stiefel, J. Jaggi, C. Halpern, M. Kerr, E. Maloney, M. Robinson, I. Lucki, *et al. Intracerebral microdialysis during deep brain stimulation surgery.* Journal of neuroscience methods **190**(1), 106 (2010).

[47] I. Basu, T. D., D. Graupe, and K. Slavin. *Adaptive control of deep brain stimulator for essential tremor: Entropy-based tremor prediction using surface emg.* Proceedings of the 33rd Annual International Conference of the IEEE EMBS, Boston, Massachusetts USA, August 30-September 3,2011 **1**, 7711 (2011).

[48] J. Pham, S. M. Cabrera, C. Sanchis-Segura, and M. A. Wood. *Automated scoring of fear-related behavior using ethovision software.* Journal of neuroscience methods **178**(2), 323 (2009).

[49] A. B. Schwartz, X. T. Cui, D. J. Weber, and D. W. Moran. *Brain-controlled interfaces: movement restoration with neural prosthetics.* Neuron **52**(1), 205 (2006).

[50] P. Brown, A. Oliviero, P. Mazzone, A. Insola, P. Tonali, and V. Di Lazzaro. *Dopamine dependency of oscillations between subthalamic nucleus and pallidum in parkinson's disease.* The Journal of neuroscience **21**(3), 1033 (2001).

[51] A. Eusebio and P. Brown. *Synchronisation in the beta frequency band - the bad boy of parkinsonism or an innocent bystander ?* Experimental Neurology **217**, 1 (2009).

[52] A. Kuhn, A. Tsui, T. Aziz, N. Ray, C. Brucke, A. Kupsch, G. Schneider, and P. Brown. *Pathological synchronisation in the subthalamic nucleus of patients with parkinson's disease relates to both bradykinesia and rigidity.* Experimental Neurology **215**, 380 (2009).

[53] C. Hammond, H. Bergman, and P. Brown. *Pathological synchronization in parkinson's disease: networks, models and treatments.* Trends in neurosciences **30**(7), 357 (2007).

[54] M. Weinberger, N. Mahant, W. D. Hutchison, A. M. Lozano, E. Moro, M. Hodaie, A. E. Lang, and J. O. Dostrovsky. *Beta oscillatory activity in the subthalamic nucleus and its relation to dopaminergic response in parkinson's disease.* Journal of Neurophysiology **96**, 3248 (2006).

[55] S. Wang, Y. Chen, M. Ding, J. Feng, J. Stein, T. Aziz, and X. Liu. *Revealing the dynamic causal interdependence between neural and muscular signals in parkinsonian tremor.* Journal of the Franklin Institute **344**(3), 180 (2007).

[56] S. Little, A. Pogosyan, S. Neal, B. Zavala, L. Zrinzo, M. Hariz, T. Foltynie, P. Limousin, K. Ashkan, J. FitzGerald, *et al. Adaptive deep brain stimulation in advanced parkinson disease.* Annals of neurology **74**(3), 449 (2013).

[57] U. Rutishauser, E. Schuman, and A. Mamelak. *Online detection and sorting of extracellularly recorded action potentials in human medial temporal lobe recordings, in vivo.* Journal of Neuroscience Methods. **154**, 204 (2006).

[58] S. Santaniello, G. Fiengo, L. Glielmo, and W. Grill. *Closed-loop control of deep brain stimulation: A simulation study.* IEEE Transactions on neural systems and rehabilitation engineering **191**, 15 (2011).

[59] L. Rossi, G. Foffani, S. Marceglia, F. Bracchi, S. Barbieri, and A. Priori. *An electronic device for artefact suppression in human local field potential recordings during deep brain stimulation.* Journal of neural engineering **4**(2), 96 (2007).

[60] J. M. Stern. *Atlas of EEG patterns* (Lippincott Williams & Wilki, 2013), 2^{nd} ed.

[61] M. S. Thaler. *The Only EKG Book You'll Ever Need* (Lippincott Williams & Wilki, 2006), 5^{th} ed.

[62] A. Oswal, M. Beudel, L. Zrinzo, P. Limousin, M. Hariz, T. Foltynie, V. Litvak, and P. Brown. *Deep brain stimulation modulates synchrony within spatially and spectrally distinct resting state networks in parkinson's disease.* Brain **139**(5) (2016).

[63] S. Marceglia, G. Foffani, A. Bianchi, G. Baselli, F. Tamma, M. Egidi, and A. Priori. *Dopamine-dependent non-linear correlation between subthalamic rhythms in parkinson's disease.* The Journal of physiology **571**(3), 579 (2006).

[64] C. Chen, Y. Hsu, H. Chan, S. Chiou, P. Tu, S. Lee, C. Tsai, C. Lu, and P. Brown. *Complexity of subthalamic 13-35hz oscillatory activity directly correlates with clinical impairement in patients with parkinson's disease.* Experimental Neurology **224**(1), 234 (2010).

[65] R. Zwanzig. *Nonequilibrium Statistical Mechanics* (Oxford University Press, 2001).

[66] J. Crutchfield and K. Young. *Inferring statistical complexity.* Phys. Rev. Lett. **63**, 105 (1989).

[67] R. Yulmetyev, P. Hänggi, and F. Gafarov. *Quantification of heart rate variability by discrete nonstationary non-markov stochastic processes.* Phys. Rev. E **65**(4), 046107 (2002).

[68] R. Yulmetyev, S. Demin, O. Y. Panischev, P. Hnggi, S. Timashev, and G. Vstovsky. *Regular and stochastic behavior of parkinsonian pathological tremor signals.* Physica A: Statistical Mechanics and its Applications **369**(2), 655 (2006).

[69] G. Moore, D. Perkel, and J. Segundo. *Statistical analysis and funtional interpretation of neuronal spike data.* Annual Review of Physiology **28**, 493 (1966).

[70] G. Maimon and J. A. Assad. *Beyond poisson: increased spike-time regularity across primate parietal cortex.* Neuron **62**(3), 426 (2009).

[71] J. Bastian and J. Nguyenkim. *Dendritic modulation of burst-like firing in sensory neurons.* Journal of Neurophysiology **85**(1), 10 (2001).

[72] D. H. Johnson. *Point process models of single-neuron discharges.* Journal of computational neuroscience **3**(4), 275 (1996).

[73] Z. Pawlas and P. Lansky. *Distribution of interspike intervals estimated from multiple spike trains observed in a short time window.* Physical Review E **83**, 259 (2011).

[74] W. J. Wilbur and J. Rinzel. *A theoretical basis for large coefficient of variation and bimodality in neuronal interspike interval distributions.* Journal of Theoretical Biology **105**(2), 345 (1983).

[75] J. Troy and B. Lee. *Steady discharges of macaque retinal ganglion cells.* Visual neuroscience **11**(01), 111 (1994).

[76] I. M. Winter and A. R. Palmer. *Responses of single units in the anteroventral cochlear nucleus of the guinea pig.* Hearing research **44**(2-3), 161 (1990).

[77] H. M. Bayer, B. Lau, and P. W. Glimcher. *Statistics of midbrain dopamine neuron spike trains in the awake primate.* Journal of Neurophysiology **98**(3), 1428 (2007).

[78] O. Avila-Ackerberg and M. Chacron. *Nonrenewal spike train statistics: causes and functional consequences on neural coding.* Experimental Brain Research **210**, 353 (2011).

[79] B. Dummer, S. Wieland, and B. Lindner. *Self-consistent determination of the spike-train power spectrum in a neural network with sparse connectivity.* Frontiers in Computational Neuroscience **8**(104) (2014).

[80] K. Rajdl and P. Lansky. *Stein's neuronal model with pooled renewal input.* Biological cybernetics **109**(3), 389 (2015).

[81] C. Ly and D. Tranchina. *Spike train statistics and dynamics with synaptic input from any renewal process: a population density approach.* Neural Computation **21**(2), 360 (2009).

[82] W. Bialek, F. Rieke, R. de Ruyter van Steveninck, and D. Warland. *Reading a neural code.* Science **252**(5014), 1854 (1991).

[83] W. Bialek and F. Rieke. *Reliability and information transmission in spiking neurons.* Trends in neurosciences **15**(11), 428 (1992).

[84] R. Romo and E. Salinas. *Flutter discrimination: neural codes, perception, memory and decision making.* Nature Reviews Neuroscience **4**(3), 203 (2003).

[85] R. VanRullen, R. Guyonneau, and S. Thorpe. *Spike times make sense.* Trends in Neurosciences **28**(1), 1 (2005).

[86] D. A. Butts, C. Weng, J. Jin, C. I. Yeh, N. A. Lesica, J. M. Alonso, and G. B. Stanley. *Temporal precision in the neural code and the timescales of natural vision.* Nature **449**(7158), 92 (2007).

[87] C. Stevens and A. Zador. *Information through a spiking neuron.* Advances in neural information processing systems pp. 75–81 (1996).

[88] E. Adrian and Y. Zotternman. *The impulses produced by sensory nerve endings. part 3. impulses set up by touch and pressure.* Journal of Physiology **61**, 456 (1926).

[89] R. Stein, E. Gossen, and K. Jones. *Neuronal variability: Noise or part of the signal ?* Nature Review Neuroscience **6(5)**, 389 (2005).

[90] L. Kostal, P. Lansky, and J. Rospars. *Neuronal coding and spiking randomness.* European Journal of Neuroscience **26**, 2693 (2007).

[91] D. O. Hebb. *The organization of behavior: A neuropsychological theory* (Psychology Press, 2005).

[92] N. Caporale and Y. Dan. *Spike timing-dependent plasticity: a hebbian learning rule.* Annu. Rev. Neurosci. **31**, 25 (2008).

[93] R. Quian Quiroga and S. Panzeri, eds. *Principles of neural coding* (Taylor & Francis/CRC Press, Boca Raton, 2013).

[94] Z. Nenadic and J. Burdick. *Spike detection using the continuous wavelet transform.* Biomedical Engineering, IEEE Transactions on **52**(1), 74 (2005).

[95] T. Nguyen, A. Bhatti, A. Khosravi, S. Haggag, D. Creighton, and S. Nahavandi. *Automatic spike sorting by unsupervised clustering with diffusion maps and silhouettes.* Neurocomputing **153**(4), 199 (2015).

[96] C. Pedreira, J. Martinez, J. I. Matias, and R. Q. Quiroga. *How many neurons can we see with current spike sorting algorithms?* Journal of Neuroscience Methods **211**, 58 (2012).

[97] M. Lustig, J. H. Lee, D. L. Donoho, and J. M. Pauly. *Faster Imaging with Randomly Perturbed, Undersampled Spirals and L1 Reconstruction.* In *Proceedings of ISMRM* (2005).

[98] S. S. Chen, D. L. Donoho, and M. A. Saunders. *Atomic decomposition by basis pursuit.* SIAM Journal On Scientific Computing **20**, 33 (1998).

[99] R. Tibshirani. *Regression shrinkage and selection via the lasso.* Journal of the Royal Statistical Society, Series B **58**, 267 (1994).

[100] D. L. Donoho. *Compressed sensing.* IEEE Trans. Inform. Theory **52**, 1289 (2006).

[101] Y. Tsaig and D. L. Donoho. *Extensions of compressed sensing.* Signal Process. **86**(3), 549 (2006).

[102] E. W. Keefer, B. R. Botterman, M. I. Romero, A. F. Rossi, and G. W. Gross. *Carbon nanotube coating improves neuronal recordings.* Nature nanotechnology **3**(7), 434 (2008).

[103] K. Bassler, G. Gunaratne, and J. McCauley. *Markov processes, hurst exponents, and nonlinear diffusion equations: With application to finance.* Physica A-Statistical Mechanics and its Applications. **369(2)**, 343 (2006).

[104] D. Schmitt and M. Schultz. *Analyzing memory effects of complex systems from time series.* Phys. Rev. E **73(5)**, 056204 (2006).

[105] A. Knezevic and M. Martinis. *(multi) fractality of physiological time-series.* International Journal of Bifurcation and Chaos. **16**(7), 2103 (2006).

[106] R. Yulmetyev, P. Hanggi, and F. Gafarov. *Stochastic processes of demarkovization and markovization in chaotic signals of human brain electric activity from eegs during epilepsys.* Journal of Experimental and Theoretical Physics **96(3)**, 572 (2003).

[107] A. Carlson and P. Crilly. *Communication Systems: An introduction to signals and noise in electrical communication* (McGraw Hill, 2010).

[108] M. Baptista, E. Ngamga, P. Pinto, B. Margarida, and J. Kurths. *Kolmogorov-sinai entropy from recurrence times.* Physics Letters A. **374**, 1135 (2010).

[109] D. Feldman and J. Crutchfield. *Measures of statistical complexity: Why?* Physics Letters A **238(4-5)**, 244 (1998).

[110] R. Yulmetyev, P. Hänggi, and F. Gafarov. *Stochastic dynamics of time correlation in complex systems with discrete time.* Phys. Rev. E **62**, 6178 (2000).

[111] R. Yulmetyev, F. Gafarov, P. Hänggi, R. Nigmatullin, and S. Kayumov. *Possibility between earthquake and explosion seismogram differentiation by discrete stochastic non-markov processes and local hurst exponent analysis.* Phys. Rev. E **64**(6), 066132 (2001).

[112] R. Yulmetyev, S. Demin, R. Khusnutdinov, O. Panischev, and P. Hanggi. *Non-markov statistical effects of x-ray emission intensity of the microquasar grs 1915+105.* Nonlinear Phenomena in Complex Systems **9**, 313 (2007).

[113] R. Yulmetyev, S. Demin, R. Khusnutdinov, O. Panischev, and P. Hänggi. *Non-markov statistical effects of x-ray emission intensity of the microquasar grs 1915+105.* Nonlinear Phenomena In Complex Systems **9**(4), 313 (2006).

[114] R. Yulmetyev, S. Demin, O. Y. Panischev, P. Hänggi, S. Timashev, and G. Vstovsky. *Regular and stochastic behavior of parkinsonian pathological tremor signals.* Physica A **369**, 655 (2006).

[115] J. Varghese, K. Weegink, P. Bellette, T. Coyne, P. Silburn, and P. Meehan. *Theoretical & experimental analysis of the non markov parameter to detect low frequency synchronisation in time series analysis.* In *Proceedings of the 33rd Annual International*

Conference of the IEEE EMBS, Boston, Massachusetts USA, August 30-September 3,2011, pp. 1500–1505 (2011).

[116] K. Weegink, J. Varghese, P. Bellette, T. Coyne, P. Silburn, and P. Meehan. *Investigation of the non-markovity spectrum as a cognitive processing measure of deep brain microelectrode recordings.* In *International Conference on Bio-inspired Systems and Signal Processing, Rome, Italy* (2011).

[117] M. Niemann, T. Laubrich, E. Olbrich, and H. Kantz. *Usage of the mori-zwanzig method in time series analysis.* Phys. Rev. E **77(1)**, 011117 (2008).

[118] R. Yulmetyev and N. Khushnutdinov. *Statistical spectrum of the non-markovity parameter for simple model systems.* Journal of Physics A **27**, 5363 (1994).

[119] P. Hänggi. *Correlation functions and masterequations of generalized (non-markovian) langevin equations.* Zeitschrift fr Physik B Condensed Matter **314**, 407 (1978).

[120] G. Mazenko. *Nonequilibrium Statistical Mechanics* (Wiley, Weinhelm, Germany, 2006).

[121] R. N. Bracewell. *The Fourier transform and its applications* (Mcgraw-Hill, 2000).

[122] R. Bartle. *The Elements of Integration and Lebesgue Measure.* Wiley Classics Library (Wiley, 2011).

[123] R. Yulmetyev, R. Galeev, and V. Shurygin. *Investigation of non-markovian kinetics of microscopic vortices in liquids.* Physics Letters A **4**, 258 (1996).

[124] P. Hansen. *Deconvolution and regularization with toeplitz matrices.* Numerical Algorithms **29**, 323 (2002).

[125] D. Percival and A. Walden. *Spectral Analysis for Physical Applications: Multitaper and Conventional Univariate Techniques* (Cambridge University Press, 1993).

[126] V. Vasconcelos, F. Raischel, M. Haase, J. Peinke, M. Wächter, P. G. Lind, and D. Kleinhans. *Principal axes for stochastic dynamics.* Phys. Rev. E. **84**, 031103 (2011).

[127] F. Raischel, T. Scholz, V. Lopes, and P. Lind. *Uncovering wind turbine properties through two-dimensional stochastic modeling of wind dynamics.* Phys. Rev. E. **88**, 042146 (2013).

[128] W. Coffey, Y. Kalmykov, and J. Waldron. *The Langevin Equation With Applications in Physics, Chemistry and Electrical Engineering* (World Scientific, Singapore, 1996).

[129] M. Wang and G. Uhlenbeck. *On the theory of the brownian motion ii.* Reviews of Modern Physics **17**((2&3)), 323 (1945).

[130] J. Doob. *The brownian movement and stochastic equations.* Annals of Mathematics **43**, 351 (1942).

[131] L. Lankhorst. *A self-contained identification scheme for eddies in drifter and float trajectories.* J. Atmos. Oceanic Technol. **23**, 1583 (2006).

[132] J. G. Proakis and D. G. Manolakis. *Digital Signal Processing: Principles, Algorithms, and Applications* (Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1996).

[133] P. Broersen. *Facts and fiction in spectral analysis.* IEEE Transactions on Instrumentation and Measurement **49**(4), 766 (2000).

[134] J. Durbin. *Efficient estimation of parameters in moving-average models.* Biometrika **46**, 306 (1959).

[135] J. Durbin. *The fitting of time-series models.* Revue de l'Institut International de Statistique pp. 233–244 (1960).

[136] H. Akaike. *Statistical predictor identification.* Annals of the Institute of Statistical Mathematics **22**(1), 203 (1970).

[137] D. Wilks. *Statistical Methods in the Atmospheric Sciences*, vol. **100**(3) (Elsevier, 2011).

[138] R. M. Yulmetyev, S. A. Demin, and P. Hänggi. *The Logistic Map and the Route to Chaos: From The Beginnings to Modern Applications*, chap. Manifestation of Chaos

in Real Complex Systems: Case of Parkinson's Disease, pp. 175–196 (Springer Berlin Heidelberg, 2006).

[139] Z. Israel and K. J. Burchiel. *Microelectrode Recordings in Movement Disorder Surgery* (Thieme, 2004).

[140] E. Zelniker, A. Bradley, J. Castner, H. Chenery, D. Copland, and P. Silburn. *Estimation of neuronal firing rates with the three-state biological point process model.* J Neurosci Methods **174**, 281 (2008).

[141] W. D. Hutchison, R. J. Allan, H. Opitz, R. L. amd J.O. Dostrovsky, A. Lang, and A. Lozano. *Neurophysiological identification of the subthalamic nucleus in surgery for parkinsons disease.* Ann. Neurol. **44**, 622 (1998).

[142] S. Haykin. *Communication Systems* (John Wiley & Sons, 2001).

[143] R. M. Yulmetyev, D. Yulmetyeva, and F. F. Gafarov. *How chaosity and randomness control human health.* Physica A **354**, 404 (2005).

[144] A. Crossman and D. Neary. *Neuroanatomy: An Illustrated Colour Text.* Illustrated colour text series (Churchill Livingstone, 1995).

[145] F. A. Middleton and P. Strick. *Basal ganglia output and cognition: Evidence from anatomical, behavioral, and clinical studies.* Brain Cogn **42**, 183 (2000).

[146] J. Rissman, J. Eliassen, and S. E. Blumstein. *An event-related fmri investigation of implicit semantic priming.* J. Cogn. Neurosci **15**, 1160 (2003).

[147] M. Wahl, F. Marzinzik, A. Friederici, A. Hahne, A. Kupsch, G. Schneider, D. Saddy, G. Curio, and F. Klostermann. *The human thalamus processes syntactic and semantic language violations.* Neuron **59**, 695 (2008).

[148] P. Meehan, P. Bellette, A. Bradley, J. Castner, H. Chenery, D. Copland, J. Varghese, T. Coyne, and P. Silburn. *Investigation of the non-markovity spectrum as a cognitive processing measure of deep brain microelectrode recordings.* International Conference on Bio-inspired Systems and Signal Processing, Rome, Italy, pp. 144–151 (2011).

[149] G. E. Box and D. R. Cox. *An analysis of transformations.* Journal of the Royal Statistical Society. Series B (Methodological) pp. 211–252 (1964).

[150] K. Rastle, J. Harrington, and M. Coltheart. *358,534 nonwords: The arc nonword database.* The Quarterly Journal of Experimental Psychology Section A **55**(4), 1339 (2002).

[151] J. de Kroon and P. van der Laan. *A generalization of friedman's rank statistic.* Statistica Neerlandica **37**(1), 1 (1983).

[152] G. Cariolaro, T. Erseghe, and L. Vangelista. *Exact spectral evaluation of the family of digital pulse interval modulated signals.* IEEE Transactions on Information Theory **477**, 2983 (2001).

[153] T. Wysocki, H. Razavi, and B. Honary. *Digital signal processing for communication systems* (Kluwer Academic Publishers, 1997).

[154] G. di Giovanni, V. Di Matteo, and E. Esposito. *Birth, Life and Death of Dopaminergic Neurons in the Substantia Nigra*, vol. 73 of *Journal of Neural Transmission. Supplementa* (Springer Vienna, 2009).

[155] F. Rieke, D. Warland, R. van Steveninck, and W. Bialek. *Spikes: Exploring the neural code* (MIT press, 1997).

[156] D. Heyman and M. Sobel. *Stochastic Models in Operations Research: Stochastic Processes and Operating Characteristics* (Dover Courier Publications, 2004).

[157] J. Franklin and W. Bair. *The effect of a refactory period on the power spectrum of neuronal discharge.* SIAM Journal of Applied Maths **55**, 1074 (1995).

[158] B. Lindner. *Superposition of many independent spike trains is generally not a renewal process.* Physical Review E. **73**, 353 (2006).

[159] R. B. Stein. *A theoretical analysis of neuronal variability.* Biophysical Journal **5**(2), 173 (1965).

[160] W. J. Wilbur and J. Rinzel. *An analysis of stein's model for stochastic neuronal excitation.* Biological Cybernetics **45**(2), 107 (1982).

[161] H. Câteau and A. D. Reyes. *Relation between single neuron and population spiking statistics and effects on network activity.* Physical review letters **96**(5), 058101 (2006).

[162] C. Pouzat, M. Delescluse, P. Viot, and J. Diebolt. *Improved spike-sorting by modeling firing statistics and burst-dependent spike amplitude attenuation: a markov chain monte carlo approach.* Journal of neurophysiology **91**(6), 2910 (2004).

[163] K. D. Harris, D. A. Henze, J. Csicsvari, H. Hirase, and G. Buzsaki. *Accuracy of tetrode spike separation as determined by simultaneous intracellular and extracellular measurements.* Journal of Neurophysiology **84**, 401 (2000).

[164] M. S. Fee, P. P. Mitra, and D. Kleinfeld. *Variability of extracellular spike waveforms of cortical neurons.* Journal of Neurophysiology **76**(6), 3823 (1996).

[165] R. Quiroga, Z. Nadasdy, and Y. Ben-Shaul. *Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering.* Neural Computation **16**(8), 1661 (2004).

[166] S. Resnick. *Adventures in stochastic processes* (Birkhäuser, 2002).

[167] E. Banta. *A note on the correlation function of nonindependent, overlapping pulse trains.* IEEE Transactions on Information Theory **10**, 160 (1964).

[168] P. Erdos, W. Feller, and H. Pollard. *A property of power series with positive coefficients.* Bull. Amer. Math. Soc. **55**, 201 (1949).

[169] D. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes. vol I* (Springer-Verlag, 2002).

[170] F. Steigerwald, M. Pötter, J. Herzog, M. Pinsker, F. Kopper, H. Mehdorn, G. Deuschl, and J. Volkmann. *Neuronal activity of the human subthalamic nucleus in the parkinsonian and nonparkinsonian state.* Journal of Neurophysiology **100**(5), 2515 (2008).

[171] K. Weegink, P. Bellette, J. Varghese, P. Silburn, P. Meehan, and A. Bradley. *A parametric simulation of neuronal noise from microelectrode recordings*. IEEE Transactions on Neural Systems and Rehabilitation Engineering **PP**(99), 1 (2016).

[172] A. J. Lawrence. *Dependency of intervals between events in superposition processes*. Journal of the Royal Statistical Society. Series B (Methodological) **35**(2), pp. 306 (1973).

[173] P. D. Welch. *The use of fast fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms*. IEEE Transactions on audio and electroacoustics **15**(2), 70 (1967).

[174] D. J. Thomson. *Spectrum Estimation and Harmonic Analysis*. IEEE Proceedings **70**, 1055 (1982).

[175] M. Bartlett. *The spectral analysis of point processes*. Journal of the Royal Statistical Society. Series B (Methodological) pp. 264–296 (1963).

[176] M. S. Bartlett. *Statistical estimation of density functions*. Sankhyā: The Indian Journal of Statistics, Series A pp. 245–254 (1963).

[177] C. M. Bender and S. A. Orszag. *Advanced mathematical methods for scientists and engineers*. International series in pure and applied mathematics (McGraw-Hill Book Company, Auckland, Bogota, Paris, 1978).

[178] P. Hansen. *Discrete Inverse Problems* (Society for Industrial and Applied Mathematics, 2010).

[179] K. Fujiwara and K. Aihara. *Time-varying irregularities in multiple trial spike data*. The European Physical Journal B **68**(2), 283 (2009).

[180] D. E. McKeegan. *Spontaneous and odour evoked activity in single avian olfactory bulb neurones*. Brain Research **929**(1), 48 (2002).

[181] M. Badoual, M. Rudolph, Z. Piwkowska, A. Destexhe, and T. Bal. *High discharge variability in neurons driven by current noise*. Neurocomputing **65-66**, 493 (2005).

[182] D. Terman, J. E. Rubin, A. C. Yew, and C. J. Wilson. *Activity patterns in a model for the subthalamopallidal network of the basal ganglia*. Journal of Neuroscience pp. 2963–2976 (2002).

[183] E. Glaser and W. Marks. *On-line separation of interleaved neuronal pulse sequences*. Data Acquisition and Processing in Biology and Medicine **5**, 137 (1968).

[184] D. Plenz and S. T. Kital. *A basal ganglia pacemaker formed by the subthalamic nucleus and external globus pallidus*. Nature **400**(6745), 677 (1999).

[185] C. Gold, D. A. Henze, C. Koch, and G. Buzsáki. *On the origin of the extracellular action potential waveform: A modeling study*. Journal of Neurophysiology **95**(5), 3113 (2006).

[186] Z. I. Botev, J. F. Grotowski, and D. P. Kroese. *Kernel density estimation via diffusion*. Ann. Statist. **38**(5), 2916 (2010).

[187] T. Verechtchaguina, L. Schimansky-Geier, and I. Sokolov. *Spectra and waiting-time densities in firing resonant and nonresonant neurons*. Physical Review E **70**, 1829 (2004).

[188] K. Weegink, J. Varghese, and A. Bradley. *Spikes from compound action potentials in simulated microelectrode recordings*. In *Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on*, pp. 813–816 (2015).

[189] M. S. Lewicki. *A review of methods for spike sorting: the detection and classification of neural action potentials*. Network: Computation in Neural Systems **9**(4), 53 (1998).

[190] G. Buzsaki. *Large-scale recording of neuronal ensembles*. Nature Neuroscience **7**(5), 446 (2004).

[191] J. Pillow, S. Jonathon, J. Chichilnisky, and E. Simoncelli. *A model-based spike sorting algorithm for removing correlation artifacts in multi-neuron recordings*. PLOS one **8**(5) (2013).

[192] C. Ekanadham, D. Tranchina, and E. P. Simoncelli. *A unified framework and method for automatic neural spike identification.* Journal of Neuroscience Methods **222**, 47 (2014).

[193] J. H. Choi, H. K. Jung, and T. Kim. *A new action potential detector using the mteo and its effects on spike sorting systems at low signal-to-noise ratios.* Biomedical Engineering, IEEE Transactions on **53**(4), 738 (2006).

[194] S. Mallat. *A Wavelet Tour of Signal Processing: The Sparse Way* (Academic Press, 2008), 3^{rd} ed.

[195] B. Georgescu, I. Shimshoni, and P. Meer. *Mean shift based clustering in high dimensions: a texture classification example.* In *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*, pp. 456–463 vol.1 (2003).

[196] T. Nguyen, A. Bhatti, A. Khosravi, S. Haggag, D. Creighton, and S. Nahavandi. *Automatic spike sorting by unsupervised clustering with diffusion maps and silhouettes.* Neurocomputing **153**, 199 (2015).

[197] R. R. Coifman and S. Lafon. *Diffusion maps.* Applied and Computational Harmonic Analysis **21**(1), 5 (2006).

[198] F. Wood, M. J. Black, C. Vargas-Irwin, M. Fellows, and J. P. Donoghue. *On the variability of manual spike sorting.* IEEE Transactions on Biomedical Engineering **51**(6), 912 (2004).

[199] C. C. Aggarwal. *Data Mining - The Textbook* (Springer, 2015).

[200] J. Macqueen. *Some methods for classification and analysis of multivariate observations.* In 5^{th} *Berkeley Symposium on Mathematical Statistics and Probability*, pp. 281–297 (1967).

[201] R. Tibshirani, G. Walther, and T. Hastie. *Estimating the number of clusters in a data set via the gap statistic.* Journal of the Royal Statistical Society: Series B (Statistical Methodology) **63**(2), 411 (2001).

[202] P. J. Rousseeuw. *Silhouettes: A graphical aid to the interpretation and validation of cluster analysis*. Journal of Computational and Applied Mathematics **20**, 53 (1987).

[203] D. Pelleg, A. W. Moore, *et al. X-means: Extending k-means with efficient estimation of the number of clusters.* In *ICML*, vol. 1 (2000).

[204] J. Wild, Z. Prekopcsak, T. Sieger, D. Novak, and R. Jech. *Performance comparison of extracellular spike sorting algorithms for single-channel recordings*. Journal of Neuroscience Methods **203**(2), 369 (2012).

[205] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu. *A density-based algorithm for discovering clusters in large spatial databases with noise.* In *2nd International Conference on Knowledge Discovery and Data Mining*, vol. 96(34), pp. 226–231 (1996).

[206] C. Braune, C. Borgelt, and R. Kruse. *Behavioral clustering for point processes*. In *Advances in Intelligent Data Analysis XII*, pp. 127–137 (Springer, 2013).

[207] M. Ankerst, M. M. Breunig, H.-P. Kriegel, and J. Sander. *Optics: ordering points to identify the clustering structure*. In *ACM Sigmod Record*, vol. 28(2), pp. 49–60 (ACM, 1999).

[208] A. Hinneburg and D. A. Keim. *An efficient approach to clustering in large multimedia databases with noise*. In *KDD*, vol. 98, pp. 58–65 (1998).

[209] A. Hinneburg and H.-H. Gabriel. *Denclue 2.0: Fast clustering based on kernel density estimation*. In *Advances in Intelligent Data Analysis VII*, pp. 70–80 (Springer, 2007).

[210] M. Belkin and P. Niyogi. *Laplacian eigenmaps for dimensionality reduction and data representation*. Neural Computation **15**, 1373 (2002).

[211] A. Y. Ng, M. I. Jordan, Y. Weiss, *et al. On spectral clustering: Analysis and an algorithm*. Advances in neural information processing systems **2**, 849 (2002).

[212] U. Von Luxburg. *A tutorial on spectral clustering*. Statistics and computing **17**(4), 395 (2007).

[213] M. S. Lewicki. *Bayesian modeling and classification of neural signals.* Neural computation **6**(5), 1005 (1994).

[214] S. Shoham, M. R. Fellows, and R. A. Normann. *Robust, automatic spike sorting using mixtures of multivariate t-distributions.* Journal of neuroscience methods **127**(2), 111 (2003).

[215] F. Wood, S. Goldwater, and M. J. Black. *A non-parametric bayesian approach to spike sorting.* In *Engineering in Medicine and Biology Society, 2006. 28th Annual International Conference of the IEEE*, pp. 1165–1168 (2006).

[216] M. Delescluse and C. Pouzat. *Efficient spike-sorting of multi-state neurons using inter-spike intervals information.* Journal of neuroscience methods **150**(1), 16 (2006).

[217] J. A. Herbst, S. Gammeter, D. Ferrero, and R. H. Hahnloser. *Spike sorting with hidden markov models.* Journal of neuroscience methods **174**(1), 126 (2008).

[218] H. Barreto and F. Howland. *Introductory Econometrics: Using Monte Carlo Simulation with Microsoft Excel* (Cambridge University Press, 2006).

[219] D. Ge, X. Jiang, and Y. Ye. *A note on the complexity of l_p minimization.* Mathematical Programming **129**(2), 285 (2011).

[220] P. R. Gill, A. Wang, and A. Molnar. *The in-crowd algorithm for fast basis pursuit denoising.* Trans. Sig. Proc. **59**(10), 4595 (2011).

[221] D. Donoho and Y. Tsaig. *Fast solution of ℓ_1 -norm minimization problems when the solution may be sparse.* Information Theory, IEEE Transactions on **54**(11), 4789 (2008).

[222] G. Davis, S. Mallat, and M. Avellaneda. *Adaptive greedy approximations.* Constructive Approximation **13**(1), 57 (1997).

[223] E. J. Candes and T. Tao. *Decoding by linear programming.* IEEE Transactions on Information Theory **51**(12), 4203 (2005).

[224] E. Candes. *The restricted isometry property and its implications for compressed sensing.* Comptes Rendus Mathematique **346**(9-10), 589 (2008).

[225] E. Candes and M. Wakin. *An introduction to compressive sampling.* IEEE Signal Processing Magazine **25**(2), 21 (2008).

[226] A. Bandeira, E. Dobriban, D. Mixon, and W. Sawin. *Certifying the restricted isometry property is hard.* Information Theory, IEEE Transactions on **59**(6), 3448 (2013).

[227] S.-J. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky. *An interior-point method for large-scale l1-regularized least squares.* Selected Topics in Signal Processing, IEEE Journal of **1**(4), 606 (2007).

[228] D. Kim, J. Trzasko, M. Smelyanskiy, C. Haider, P. Dubey, and A. Manduca. *High-performance 3d compressive sensing mri reconstruction using many-core architectures.* Journal of Biomedical Imaging **2**, 1 (2011).

[229] M. S. Lobo, M. Fazel, and S. Boyd. *Portfolio optimization with linear and fixed transaction costs.* Annals of Operations Research **152**(1), 341 (2006).

[230] H. L. Taylor, S. C. Banks, and J. F. McCoy. *Deconvolution with the ℓ^1 norm.* Geophysics **44(1)**, 39 (1979).

[231] Y. Nesterov. *Gradient methods for minimizing composite functions.* Mathematical Programming **140**(1), 125 (2012).

[232] I. Daubechies, M. Defrise, and C. De Mol. *An iterative thresholding algorithm for linear inverse problems with a sparsity constraint.* Comm. Pure Appl. Math. **57**(11), 1413 (2004).

[233] A. Beck and M. Teboulle. *A fast iterative shrinkage-thresholding algorithm for linear inverse problems.* SIAM J. Img. Sci. **2**(1), 183 (2009).

[234] S. J. Wright, R. D. Nowak, and M. A. T. Figueiredo. *Sparse reconstruction by separable approximation.* IEEE Transactions on Signal Processing **57**(7), 2479 (2009).

[235] E. T. Hale, W. Yin, and Y. Zhang. *Fixed-point continuation for ℓ_1-minimization: Methodology and convergence*. SIAM Journal on Optimization **19**(3), 1107 (2008).

[236] W. Yin, S. Osher, D. Goldfarb, and J. Darbon. *Bregman iterative algorithms for l1-minimization with applications to compressed sensing*. SIAM J. Imaging Sci **1**(1), 143 (2008).

[237] J. Yang and Y. Zhang. *Alternating direction algorithms for ℓ_1-problems in compressive sensing*. SIAM Journal on Scientific Computing **33**(1), 250 (2011).

[238] T. Goldstein and S. Osher. *The split bregman method for l1-regularized problems*. SIAM Journal on Imaging Sciences **2**(2), 323 (2009).

[239] R. Tomioka and M. Sugiyama. *Dual-augmented lagrangian method for efficient sparse reconstruction*. Signal Processing Letters, IEEE **16**(12), 1067 (2009).

[240] M. R. Osborne, B. Presnell, and B. A. Turlach. *On the lasso and its dual*. Journal of Computational and Graphical Statistics **9**, 319 (1999).

[241] S. Asif. *Dynamic Compressive Sensing: Sparse recovery algorithms for streaming signals and videos*. Ph.D. thesis, Georgia Institute of Technology (2013).

[242] A. Y. Yang, Z. Zhou, A. G. Balasubramanian, S. S. Sastry, and Y. Ma. *Fast ℓ_1 -minimization algorithms for robust face recognition*. IEEE Transactions on Image Processing **22**(8), 3234 (2013).

[243] C. Ekanadham, D. Tranchina, and E. P. Simoncelli. *Recovery of sparse translation-invariant signals with continuous basis pursuit*. IEEE Transactions on Signal Processing **59**(10), 4735 (2011).

[244] E. J. Candes, M. B. Wakin, and S. P. Boyd. *Enhancing sparsity by reweighted ℓ_1 minimization*. Journal of Fourier Analysis and Applications **14**(5), 877 (2008).

[245] M. Grant and S. Boyd. *CVX: Matlab software for disciplined convex programming, version 2.1* (2014).

[246] M. Grant and S. Boyd. *Graph implementations for nonsmooth convex programs.* In V. Blondel, S. Boyd, and H. Kimura, eds., *Recent Advances in Learning and Control*, Lecture Notes in Control and Information Sciences, pp. 95–110 (Springer-Verlag Limited, 2008).

[247] S. Boyd and L. Vandenberghe. *Convex Optimization* (Cambridge University Press, USA, 2004).

[248] D. Malioutov, M. Cetin, and A. Willsky. *Homotopy continuation for sparse signal representation.* In *Acoustics, Speech, and Signal Processing, 2005. Proceedings. (ICASSP '05). IEEE International Conference on*, vol. 5, pp. 733–736 (2005).

[249] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani. *Least angle regression.* Annals of Statistics **32**, 407 (2004).

[250] D. S. Watkins. *Fundamentals of Matrix Computations* (John Wiley & Sons, Inc., New York, USA, 1991).

[251] S. Hammarling and C. Lucas. *Updating the qr factorization and the least squares problem.* Manchester Institute for Mathematical Sciences (2008).

[252] W. Dai and O. Milenkovic. *Subspace pursuit for compressive sensing signal reconstruction.* Information Theory, IEEE Transactions on **55**(5), 2230 (2009).

[253] D. Needell and J. Tropp. *Cosamp: Iterative signal recovery from incomplete and inaccurate samples.* Applied and Computational Harmonic Analysis **26**(3), 301 (2009).

[254] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani. *Least angle regression.* Annals of Statistics **32**, 407 (2004).

[255] R. Tomioka, T. Suzuki, and M. Sugiyama. *Super-linear convergence of dual augmented lagrangian algorithm for sparsity regularized estimation.* J. Mach. Learn. Res. **12**, 1537 (2011).

[256] A. Dhara and J. Dutta. *Optimality Conditions in Convex Optimization: A Finite-Dimensional View* (Taylor & Francis, 2011).

[257] J. M. Borwein and A. S. Lewis. *Convex analysis and nonlinear optimization, theory and examples* (2000).

[258] G. Turin. *An introduction to matched filters*. IRE Transactions on Information Theory **6**(3), 311 (1960).

[259] D. Donoho, I. Johnstone, and I. Johnstone. *Ideal spatial adaptation by wavelet shrinkage*. Biometrika **81**, 425 (1993).

[260] Y. Ghanbari, P. Papamichalis, and L. Spence. *Graph-spectrum-based neural spike features for stereotrodes and tetrodes*. pp. 598–601 (2010).

[261] D. Comaniciu and P. Meer. *Mean shift: a robust approach toward feature space analysis*. Pattern Analysis and Machine Intelligence, IEEE Transactions on **24**(5), 603 (2002).

[262] D. Comaniciu, V. Ramesh, and P. Meer. *The variable bandwidth mean shift and data-driven scale selection*. In *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, vol. 1, pp. 438–445 vol.1 (2001).

[263] S. Hader and F. A. Hamprecht. *Between Data Science and Applied Data Analysis: Proceedings of the 26th Annual Conference of the Gesellschaft für Klassifikation e.V., University of Mannheim, July 22–24, 2002*, chap. Efficient Density Clustering Using Basin Spanning Trees, pp. 39–48 (Springer Berlin Heidelberg, 2003).

[264] M. A. Carreira-Perpiñán. *Fast nonparametric clustering with gaussian blurring mean-shift*. In *Proceedings of the 23rd International Conference on Machine Learning*, ICML '06, pp. 153–160 (ACM, New York, NY, USA, 2006).

[265] J. Fan, S. Upadhye, and A. Worster. *Understanding receiver operating characteristic (roc) curves*. Canadian Journal of Emergency Medicine **8**(01), 19 (2006).

[266] A. P. Bradley. *The use of the area under the roc curve in the evaluation of machine learning algorithms*. Pattern recognition **30**(7), 1145 (1997).

[267] B. P. Bean. *The action potential in mammalian central neurons*. Nature Reviews Neuroscience **8**(6), 451 (2007).

[268] A. P. Bradley. *Roc curves and the χ^2 test.* Pattern Recognition Letters **17**(3), 287 (1996).

[269] R. Q. Quiroga. *What is the real shape of extracellular spikes?* Journal of neuroscience methods **177**(1), 194 (2009).

[270] S. Janke and F. Tinsley. *Introduction to Linear Models and Statistical Inference* (Wiley, 2005).

[271] C. G. Goetz, S. Fahn, P. Martinez-Martin, W. Poewe, C. Sampaio, G. T. Stebbins, M. B. Stern, B. C. Tilley, R. Dodel, B. Dubois, *et al. Movement disorder society-sponsored revision of the unified parkinson's disease rating scale (mds-updrs): Process, format, and clinimetric testing plan.* Movement Disorders **22**(1), 41 (2007).

[272] P. M. Broersen and S. De Waele. *Time series analysis in a frequency subband.* Instrumentation and Measurement, IEEE Transactions on **52**(4), 1054 (2003).

[273] P. M. Broersen. *Automatic spectral analysis with time series models.* Instrumentation and Measurement, IEEE Transactions on **51**(2), 211 (2002).

[274] M. Wilson. *Mrc psycholinguistic database: Machine-usable dictionary, version 2.00.* Behavior Research Methods, Instruments, & Computers **20**(1), 6 (1988).

[275] D. A. Balota, M. J. Yap, K. A. Hutchison, M. J. Cortese, B. Kessler, B. Loftis, J. H. Neely, D. L. Nelson, G. B. Simpson, and R. Treiman. *The english lexicon project.* Behavior research methods **39**(3), 445 (2007).

[276] C. J. Davis. *N-watch: A program for deriving neighborhood size and other psycholinguistic statistics.* Behavior research methods **37**(1), 65 (2005).

[277] A. Ridolfi and M. Win. *Ultrawide bandwidth signals as shot noise: A unifying approach.* IEEE Journal on selected areas in communications **244**, 899 (2006).

[278] D. MacKay and W. McCulloch. *The limiting information capacity of a neuronal link.* The bulletin of mathematical biophysics **142**, 127 (1952).

[279] G. Lindgren. *Stationary Stochastic Processes* (CRC Press, 2013).

[280] M. Simon and S. Million. *Power spectrum of unbalanced nrz and biphase signals in the presence of data asymmetry.* The Telecommunications and Data Acquisition Progress Report pp. 42–126 (1996).

[281] G. Cariolaro and G. Pierobon. *Stationary symbol sequences from variable-length word sequences.* IEEE Transactions on Information Theory **232**, 243 (1977).

[282] W. Gerstner, W. Kistler, R. Naud, and L. Paninski. *Neuronal Dynamics: From Single Neurons to Networks and Models of Cognition* (Cambridge University Press, 2014).

[283] P. DiLorenzo and J. Victor. *Spike Timing: Mechanisms and Function* (CRC Press, 2013).