

The Neuronal Code

Development of tools and hypotheses
for understanding the role of
synchronization of neuronal activity

vorgelegt von
Dipl.-Phys. Gordon Pipa
aus Frankfurt am Main

von der Fakultät IV - Elektrotechnik und Informatik
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften
- Dr. rer. nat. -

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender: Herr Professor Dr. Ehrig
Gutachter: Herr Professor Dr. Obermayer
Gutachter: Herr Professor Dr. Singer (MPI für Hirnforschung/ Ffm.)
Gutachterin: Frau Dr. Grün (FU Berlin)

Tag der wissenschaftlichen Aussprache: 14. August 2006

Berlin 2006
D 83

Contents

Acknowledgments	vii
Eidesstattliche Versicherung	viii
Deutsche Zusammenfassung der Arbeit	ix
I. Introduction	1
1. The Neuronal Code	3
1.1. Complex Systems and Evolution	3
1.2. Hypotheses About the Neuronal Code	4
1.3. Testing the Assembly Hypothesis	6
2. Neuronal Signals and Electrophysiological Recordings	7
2.1. Spiking Activity	8
2.1.1. Nature and Source of Spiking Activity	8
2.1.2. Recording of Neuronal Activity	9
2.1.3. Spike Detection	9
2.1.4. Spike Sorting	9
2.1.5. Event Based Representation of Spiking Activity	11
2.2. Local Field Potential (LFP)	12
2.2.1. Source of Local Field Potentials	12
3. Neuronal Assemblies and Coordinated Neuronal Activity	13
3.1. Temporal Pattern of Spiking Activity	13
3.1.1. Synchronous Events (Joint Spike Activity)	13
3.1.2. Spatiotemporal Pattern	13
3.1.3. Variability of Spike Timing (Jitter)	14
3.1.4. Pattern Complexity and Correlation Order	14
3.2. Local Field Potential and Pattern of Neuronal Activity	16

3.2.1.	Rhythmicity of the Local Field Potential	16
3.2.2.	Frequency-Locking, Coherent Oscillations and Phase-Locking	17
3.3.	Coupling of Complex Systems	17
3.3.1.	n:m Frequency and Phase-Locking in Complex Systems	17
3.3.2.	Synchronization by Coupling: Arnold Tongues	17
4.	Tools to Detect Coordinated Neuronal Activity	19
4.1.	Tools for Analyzing the Local Field Potential	19
4.1.1.	Definition of Coherence	19
4.1.2.	Definition of Phase-Locking	20
4.1.3.	Comparison of Coherence and Phase-Locking	21
4.1.4.	New Approaches to Identify Coordinated Neuronal Activity	22
4.2.	Tools for Spiking Data	22
4.2.1.	Variability, Rareness, Non-Stationarity and Auto-Structure	23
4.2.2.	Unitary Event Analysis	23
4.2.3.	Limitations of the Unitary Event Method	26
4.2.4.	New Approaches to Identify Coordinated Spiking Activity	28
II.	New Tools for Analyzing Coordinated Neuronal Activity	29
5.	New Resampling Tools for Uni-, Bi-, and Multivariate Data	31
5.1.	Uni-Variate Data: Assessment of Stability and Reliability	31
5.1.1.	Stability and Reliability of Phase-Locking and Coherence	32
5.1.2.	Stability and Reliability of Significant Joint Spike Patterns	33
5.2.	Bi-Variate Data: Assessment of Effects	35
5.3.	Permutation Test for Bivariate Data	36
5.3.1.	Number of Permutation Derived by Additional False Positives	38
5.4.	Multi-Variate Data: Assessment of Effects	38
6.	NeuroXidence	41
6.1.	Method	41
6.1.1.	Detection of Joint Spike Events	44
6.1.2.	Definition of the Hypothesis Test	45

6.1.3.	Generation of Surrogate Data	46
6.1.4.	Shuffled Spike Trains Used as Surrogate Data	47
6.1.5.	Implementation of the Statistical Test	48
6.1.6.	Rare Events versus Significant	49
6.1.7.	Parametric t -Test versus Non-Parametric Wilcoxon-Rank Test	49
6.1.8.	Parameters of the NeuroXidence Significance Estimation	50
6.1.9.	False Positive Rates	52
6.2.	Results	53
6.2.1.	Joint Spike Pattern Detection	53
6.2.2.	False Positives for Stationary Processes	54
6.2.3.	Test Power for Stationary Processes	55
6.2.4.	Comparison of the Test Power of NeuroXidence and of the UE Method	58
6.2.5.	False Positives and Test Power (Non-Stationary Process)	61
6.2.6.	Test Power of NeuroXidence for Oscillatory Processes	65
6.2.7.	Computational Complexity	66
6.3.	Discussion	67
6.3.1.	Discussion of the NeuroXidence Method	67
6.3.2.	Oscillatory Processes	68
6.3.3.	Rare and Spurious Events versus Reliably Reoccurring Events	69
6.3.4.	Analysis Window Length	69
6.3.5.	NeuroXidence versus UE Method and Shuffle-Corrected Cross-Correlogram	70
 III. New Tools Applied to Data		71
 7. High Complexity Joint Spike Activity in Cat Visual Area 17		73
7.1.	NeuroXidence Based Analysis of Single-Unit Activity from 48 Cells	73
7.2.	Discussion: Significant Joint Spike Activity in Real Data Evaluated by NeuroXidence	75
 8. Short-Term Memory Related LFP Oscillation in the Prefrontal Cortex		77
8.1.	Introduction	77
8.2.	Dynamics of Local Field Potentials	78
8.3.	Behavioral Performance	79
8.3.1.	Performance Effects	80

8.3.2.	Performance Effects in Grand Average Power and Phase-Locking	82
8.3.3.	λ -Maps	82
8.3.4.	Modulation of λ -Maps	84
8.3.5.	Task-Specific Changes	85
8.3.6.	Comparisons of the Modulation of the Grand Average and λ -Maps	85
8.3.7.	Phase-Locking - Spatial Extent of Synchronization	86
8.3.8.	Controlling for Induced Power by Eye-Movements	87
8.4.	Stimulus Specific LFP Power Modulation	88
8.5.	Discussion	90
9.	Short-Term Memory-Related High Complexity Joint Spike Activity	93
9.1.	Significant Joint Spike Pattern in Prefrontal Multi-Unit Activity	93
9.2.	Rate of Significant Joint Spike Patterns	95
9.2.1.	Average Rate of Significant Joint Spike Pattern ($\bar{\lambda}(c, t)$)	95
9.2.2.	Task-Related Modulations of $\bar{\lambda}(c, t)$	96
9.2.3.	Results	96
9.3.	Performance-Related Modulation of the Strength of Synchronization ($\lambda_c(c)$ and $\lambda_i(c)$)	97
9.3.1.	Performance Dependent Formation of Neuronal Assemblies ($\psi(c)_c$ and $\psi(i)_c$)	100
9.3.2.	Results	101
9.4.	Motor-Related Modulations of the Strength of Synchronization	104
9.4.1.	Results	104
9.5.	Confirmation of Fine-Temporal Cross-Structure by Jittering of Spikes	106
9.6.	Discussion	107
IV.	Discussion and Conclusions	109
10.	Discussion	111
10.1.	Discussion: New Tools to Investigate Joint Spike Activity	111
10.1.1.	Bootstrap Unitary Event Method	112
10.1.2.	NeuroXidence	113
10.2.	Discussion: Results Revealed by New Tools for Analyzing JS Activity	115
10.2.1.	High-Complexity JS Patterns Revealed by NeuroXidence	115
10.2.2.	Task-, Performance-, and Motor-Related Neuronal Assembly Formation	116

10.3. Discussion: Short-Term Memory-Related LFP Oscillation in the Prefrontal Cortex	117
10.3.1. Oscillations and Synchronization Involved in Short-Term Memory	118
10.3.2. Spatial Extend of Neuronal Oscillation and Synchronization	118
10.3.3. Performance-Related and Memorandum-Specific Neuronal Oscillation	119
10.3.4. Model of Processes Underlying Short-Term Memory	120
10.4. Concurrence of Changes of the LFP and JS Activity during Short-Term Memory	121
11. Final Conclusions	123
11.1. Oscillation and Synchronization Underlying Short-Term Memory	123
11.2. Task- and Performance-Related Formation of Neuronal Assemblies	123
11.3. NeuroXidence	124
11.4. Conclusive Evidence for Joint Spike Activity Supporting the Assembly Hypothesis	124
 V. Appendix	 125
A. Appendix: NeuroXidence Methods	127
A.1. Preprocessing	127
A.2. Standard Parameters for NeuroXidence	129
A.3. Analytical Description of the Test Power for Coherent Oscillatory Rate Changes	129
B. Appendix: Short-Term Memory-Related LFP Oscillation in the Prefrontal Cortex	131
B.1. Inter-Subject Comparison of λ -Maps	131
B.2. Detailed Discussion of the Coordination of Memory Processes by an n:m Locking	132
C. Appendix: High Complexity JS Activity in Cat Visual Area 17	135
C.1. Preparation	135
C.2. Recording	135
C.3. Visual Stimulation	135
D. Appendix: Hypothesis Tests	139
D.1. The Concept of Hypothesis Testing	139
D.1.1. Null Hypothesis (H_0)	139
D.1.2. Alternative Hypothesis (H_1)	139
D.1.3. Type I and Type II Errors	139

D.1.4. Critical Region	140
D.1.5. One Sided Test	141
D.1.6. Two Sided Test	141
D.1.7. p -Value and Test Power	141
D.2. Parametric Tests	142
D.2.1. Unpaired t-Test	143
D.2.2. Paired t-Test	143
D.2.3. One-Way Analysis of Variance (ANOVA)	143
D.3. Non-Parametric Tests	144
D.3.1. Sign Test for One-Sample or Paired Data	145
D.3.2. Wilcoxon Signed-Rank Test for One-Sample or Paired Data	145
D.3.3. Wilcoxon Rank Sum (Mann-Whitney) for Two Unpaired Samples	146
D.3.4. Kruskal-Wallis Non-Parametric Equivalent of ANOVA	146
D.4. Parametric versus Non-Parametric	146
D.4.1. Advantages of Non-Parametric Methods	146
D.4.2. Disadvantages of Non-Parametric Methods	146
E. Appendix: Bootstrapping and Resampling	147
E.1. Introduction to Resampling	147
E.2. Formal Definition of Bootstrapping	149
E.3. Monte Carlo Estimation	151
E.4. Bootstrap Standard Deviation of an Arbitrary Parameter	151
E.5. Hypothesis Testing by Bootstrapping	152
F. Appendix: Parallel and Load Balanced Computation with ClusterMatlab	153
G. Appendix: Supplementary Material and Figures	155
G.1. Spike Sorting	155
G.2. Supplementary Figures	157
 List of Variables	 162
List of Figures	164
Bibliography	169

Acknowledgments

There are many people to whom I owe a debt of thanks for their support over the last years.

- First, on a personal note, I am grateful to my wife, Gabriela Pipa, and my family, especially my parents, who gave me constant support, in every sense of the word. In addition, the ways in which Gabi helped me to finish this thesis are truly indescribable.
- I would like to sincerely acknowledge my supervisor, Prof. Wolf Singer. He inspired me and formed my scientific character with his great ideas and his ability to keep seeing the whole while integrating details. His personality as well as his inspiring way of doing and describing science are just fascinating.
- I especially would like to thank Dr. Sonja Grün. Sonja accompanied and supported my scientific development since my diploma thesis. I really enjoy her rigorous way of doing science and developing ideas. Without her support, most of the presented work in this thesis wouldn't have been possible.
- Dr. Matthias H. J. Munk, Ellen Städler, Dr. Eugenio Rodriguez-Balboa and Dr. Danko Nikolić were great colleagues, with whom working together was a great experience for me. In particular, I would like to thank all four for their supplied data that they had either recorded or pre-processed.
- I would like to recognize Prof. Dr. Klaus Obermayer, Prof. MD, Ph.D. Emery N. Brown, Dr. Diek Wheeler, Prof. Jochen Triesch, Dr. Raul Mureşan and Dr. Lars Schwabe for our stimulating discussions.
- Lastly, I would like to sincerely acknowledge the proof reading of this thesis by Gabriela Pipa, Dr. Diek Wheeler, Jürgen Leipner, and Dr. Matthias H. J. Munk.

Eidesstattliche Versicherung:

Hiermit erkläre ich an Eides statt, dass ich die an der Fakultät VI - Elektrotechnik und Informatik der Technischen Universität zu Berlin eingereichte Dissertation mit dem Titel "The neuronal code Development of tools and hypotheses for understanding the role of synchronization of neuronal activity" von mir selbst ohne unzulässige Hilfe Dritter verfasst wurde, auch in Teilen keine Kopie anderer Arbeiten darstellt und die benutzten Hilfsmittel sowie die Literatur vollständig angegeben sind.

Gordon Pipa, Berlin, den 28.März 2006.

Deutsche Zusammenfassung der Doktorarbeit

Wichtige, in den Neurowissenschaften lebhaft diskutierte und häufig untersuchte Funktionen des Gehirns sind die Informationsverarbeitung und das Gedächtnis, denen der neuronale Code zugrunde liegt. Seine Entschlüsselung würde uns ermöglichen, Krankheiten wie Schizophrenie, Epilepsie, Morbus Parkinson und Morbus Alzheimer, die auf Fehlfunktionen dieses Codes beruhen, zu verstehen und unter Umständen Methoden zu eröffnen, diese Krankheiten heilen zu können. Zudem würde ein Verständnis des neuronalen Codes mit Sicherheit unsere technischen Konzepte der Informationsverarbeitung weitestgehend in Frage stellen und uns zu neuen Lösungen inspirieren.

Es stehen sich zwei Hypothesen über den neuronalen Code gegenüber. Die erste Hypothese postuliert Informationsverarbeitung auf der Basis einzelner Neurone und wird als Großmutterzellen-Hypothese bezeichnet (engl: "grand mother cell hypothesis" oder "labeled line code", siehe auch (Barlow, 1972; Bialek et al., 1991; Bialek and Rieke, 1992; Shadlen and Newsome, 1998)). Dieser Annahme steht die Hypothese gegenüber, dass Information durch Gruppen von Zellen verarbeitet wird (Hebb, 1949). Bei dieser Hypothese unterscheidet man weiter zwischen der Assemblytheorie, der Gruppenratenkodierung und der Spike Rankkodierung, die jeweils Informationsverarbeitung durch Gruppen von Zellen annehmen. Unterschiede zwischen der Assemblytheorie (von der Malsburg, 1981; Singer, 1999), der Gruppenratenkodierung (Georgopoulos et al., 1988) und der Spike Rankkodierung (Fabre-Thorpe et al., 1998) sind die angenommenen Zeitskalen, die für die Informationsverarbeitung genutzt werden. Die Assemblytheorie basiert auf der Annahme, dass eine millisekundenpräzise Koordinierung der Spikeaktivität genutzt wird, um Relationen zwischen den individuellen Repräsentationen von Information einzelner Neurone zu enkodieren. Die Spike Rankkodierung basiert alleine auf der Reihenfolge der Spikes einer Population, wobei die Hypothese der Gruppenratenkodierung annimmt, dass Information alleine durch die Häufigkeit von Spikes enkodiert wird.

In den letzten Jahren wurden sowohl die Assemblyhypothese als auch die Ratenhypothese kontrovers und leidenschaftlich diskutiert. Beide Hypothesen fanden durch experimentelle Befunde Unterstützung, welches letztendlich zu einer starken Polarisierung und zu Zweifeln an der jeweiligen anderen Hypothese, den dazugehörigen Ergebnissen und Methoden führte. Es ist deshalb für weiterführende Untersuchungen und Experimente nötig, klare und möglichst einfache Arbeitshypothesen aufzustellen und geeignete Methoden zu benutzen, die eindeutige Antworten und überzeugende Argumente liefern können.

In dem ersten Teil dieser Doktorarbeit widmeten wir uns deshalb der Entwicklung vier neuer Methoden, die es erlauben, eine klare Trennungslinie zwischen Experimenten und Daten, die die Raten- oder Assemblyhypothese unterstützen, zu ziehen. Jede der entwickelten Methoden basiert auf nichtpara-

metrischen Signifikanzschätzungen, die auf der Generierung von surrogaten Daten, Bootstrapping sowie Permutationstests beruhen, so dass jede der neuen Methoden eine hohe statistische Robustheit aufweist. Eine dieser neuen Methoden ist NeuroXidence, die zur Untersuchung von zeitlich präzise koordinierter Spikeaktivität mehrerer Neurone von uns entwickelt wurde. Unseres Wissens nach ist NeuroXidence die erste Methode, die es erlaubt, unter Berücksichtigung der kompletten Autostruktur der Spikeaktivität eines Neurons, millisekundenpräzise Aktivitätsmuster von augenblicklich bis zu 100 Neuronen zu finden und statistisch robust und zugleich sehr sensitiv zu evaluieren.

Im zweiten Teil dieser Doktorarbeit wenden wir diese neuen Methoden auf zwei verschiedene Arten elektrophysiologischer Daten an: 1) simultan aufgenommene Spikeaktivität von Neuronen, 2) simultan aufgenommene lokale Feldpotentiale ('LFP'), die jeweils die synaptische Aktivität mehrerer tausend Neurone darstellen. Alle in dieser Doktorarbeit verwendeten Daten wurden entweder im präfrontalen Kortex im wachen Affen oder im visuellen Areal 17 in anästhesierten Katzen aufgenommen.

Unsere Ergebnisse demonstrieren, dass präzise koordinierte Spikeaktivität in den aufgenommenen Daten häufiger auftritt als sie per Zufall erwartet würde. Zudem konnten wir zeigen, dass diese erhöhte Häufigkeit koordinierter Spikeaktivität durch intrinsische Mechanismen des neuronalen Netzwerkes von Neuronen generiert und durch das Verhalten des Versuchstiers oder durch einen Stimulus moduliert wurde. Durch die Analyse von Spike- und LFP-Aktivität, die in einem wachen Affen, der eine Kurzzeitgedächtnisaufgabe ausführte, aufgenommen wurde, konnten wir die Rolle von periodischer Aktivität des LFP sowie die Rolle von synchroner Spikeaktivität weiter erklären. Unsere Ergebnisse zeigen, dass sowohl oszillatorische Komponenten des LFP-Signals als auch Spikesynchronisation mit Millisekundenpräzision eng mit den kognitiven Prozessen der Kurzzeitgedächtnisaufgabe wie Enkodieren, Speichern und Wiederabrufen der Information sowie motorischer Aktivität korrelieren. Zudem konnten wir zeigen, dass die Stärke von Oszillationen des LFPs, die Synchronisation von Oszillationen des LFPs verschiedener Elektroden sowie die millisekundengenaue Synchronisation von Spikes verschiedener Neurone von der erfolgreichen Ausführung der Kurzzeitgedächtnisaufgabe abhängt.

Abschließend können wir festhalten, dass die während unserer Forschung entwickelten Methoden eindeutige Antworten und überzeugende Argumente über die grundlegenden Eigenschaften des neuronalen Codes liefern können. Auf physiologische Daten angewendet, konnten wir mit Hilfe unserer entwickelten Methoden eindeutig zeigen, dass Synchronisation neuronaler Aktivität vorhanden ist und durch kognitive Prozesse des Versuchstieres moduliert wird.

Part I.

Introduction

1. The Neuronal Code

Neuroscience and brain research, formerly being mainly descriptive, are now undergoing a similar process of change as physics did at the beginning of the 20th century. Neuroscience is more and more becoming a research field that is interdisciplinary and driven by the huge and unmanageable amount of as well as diversity of results, to search for the principle mechanisms in the brain. Likewise, as it was in physics more than 100 years ago, the success for identifying these principle mechanisms is crucially dependent on a fruitful and inspiring interaction between experimental and theoretical sciences.

Nevertheless, the research principles in physics at that time and in neuroscience nowadays are fundamentally different. The success of physics was possible because it was focused on the microscopic description of basic elements of the universe, e.g. the forces and particles. This enabled experimentalists and theoreticians to work with simple and isolated systems as well as well-defined hypotheses. In contrast, neuroscience deals with a complex system, whose properties are likely emerging from its own complexity. This prohibits investigating the underlying principles of the brain based on a microscopic view that is focused only on the basic elements. As in most complex systems, the basic elements of the brain are the substrate of the system, and therefore, likely not sufficient to describe the system's properties.

1.1. Complex Systems and Evolution

Even though complex systems are complex, they are not unclassifiable and uncharacterizable. Quite the contrary, the underlying principles of complex systems are often much more simple than suggested by their behavior and dynamics. Powerful examples that illustrate this are the fractal dimension that characterizes fractals (Mandelbrot, 1967; Mandelbrot, 1973), the universal scaling behavior of the frequency of bifurcations that distinguishes chaotic systems (Feigenbaum, 1978; Feigenbaum, 1979), as well as the type of connectivity described by the graph theory of networks (e.g. small world versus scale free networks) (Barabási and Albert, 1999; Barabási et al., 1999; Barabási, 2005; Albert et al., 1999; Albert et al., 2000). As these examples demonstrate, the complex systems that are well understood are those that can be tackled analytically, which requires identifying the crucial elements that characterize the system. The latter requires that the system be reducible.

Reducibility, in particular, is the critical point of the neuroscience field today. Like any system that was designed by evolution, the brain is a system that is composed of various subsystems and units. These

1. The Neuronal Code

parts were changed and adapted by evolution to optimize the whole system by following implicit, and to us mostly unknown, rules. The spatial scales of the subsystems of the brain range from molecular reactions (about 10^{-9} m) to whole brain areas (about 10^{-1} m). Furthermore, the intrinsic temporal scales of the subsystems range from nearly instantaneous electric-field interactions and neuronal spiking (about one millisecond) to slow modifications of the neuronal system on timescales of weeks, months, and years. This illustrates the difficulty, or even the impossibility, to reduce the whole system to one core that enables us to study the elementary features of the brain.

Nevertheless, to reduce the complexity and to approach reducibility, one can study parts or aspects of the system. This allows one to ignore features, interactions, or elements of the system that either do not match the temporal scale of interest or are assumed to be not important. Without any doubt, the neuronal code, which is underlying the information processing in the brain, has been the most attended aspect in recent years, since understanding the neuronal code could serve the understanding of brain diseases and could induce a revolution in technical systems.

1.2. Hypotheses About the Neuronal Code

It is generally accepted that neurons are the basic elements that process information in the brain and form the substrate for the neuronal code. The human brain consists of up to 10^{12} neurons, each of them connected with approximately a thousand other neurons, that are the basis for an extremely interconnected and complex network (Braitenberg and Schüz, 1998). Within the neuronal network, information is transmitted from one neuron to another through action potentials, the so called spikes (Adrian, 1928; Hodgkin and Huxley, 1939).

Thus, focusing on the neuronal code reduces the complexity of the system that has to be investigated by delineating an upper and lower bound to the important timescales that define the system. The upper bound is given by the amount of time the neuronal system needs to process information, which is on the order of a few hundred milliseconds. The lower bound is given by the intrinsic timescale of spiking, since spiking is the only form of communication between neurons that is faster than the upper bound. These constraints allow one to reduce the investigated system to a network of neurons whose communication is based on spikes. With spiking being the only form of coupling across the neuronal network, intrinsic dynamics of each neuron that are faster than the dynamics of the network can be considered to be independent of the other neurons of the network.

Nevertheless, the way information is encoded and processed by the neuronal network is still the subject of research. The first experiments investigated the principles of neuronal information coding by recording spiking activity from single neurons. These recordings showed that neurons react to stimulation by changing the frequency of spikes, commonly referred to as the spiking rate. These results have led to the hypothesis that each neuron encodes only one certain kind of information, which is clearly associated to

itself ('Grandmother neuron,' 'single cell,' or 'labeled line code'), as well as to the hypothesis that the transportation of information between neurons is enabled by changes in the spiking rate ('rate coding') (Baker and Lemon, 2000; Barlow, 1972; Bialek et al., 1991; Bialek and Rieke, 1992; Brody, 1999; Oram et al., 1999; Shadlen and Newsome, 1994; Shadlen and Newsome, 1998). Especially the second assumption of rate coding seemed to stand up to many neurophysiologic experiments, in which the observed neurons were predictably changing the spiking rate when reacting to a specific stimulus. Moreover, it was shown that simplified models of a nerve cell ('Integrate-and-Fire-Neurons') are well suited to receive rate-coded information. Without any doubt, the most appealing characteristic of the single-cell coding hypothesis is its simplicity.

In 1949 Donald Hebb (Hebb, 1949) formulated an alternative hypothesis that assumes that neurons form functional units ('cell assemblies') that code information in a synergistic manner. This hypothesis was latter extended by von der Malsburg and Singer (Gray and Singer, 1987; von der Malsburg, 1981; Singer et al., 1988; Singer, 1993; Singer, 1999), who introduced the concept that the formation of cell assemblies, each defined by synchronized spiking of the neurons belonging to the same assembly, serves the binding of information. Thus, in comparison to the grandmother neuron theory, the assembly hypothesis is more complex since it assumes that information is processed by groups of neurons based on the temporal relation of their spikes on a millisecond timescale. The most appealing points of the assembly hypothesis are, first, that it contains the basic idea that complex systems are shaped by self organization, such as by formation of groups and synchronization (Neda et al., 2000; Pikovsky et al., 2001), and second, that it overcomes several limitations of the single-cell coding hypothesis, such as very limited coding potential, no flexible grouping and binding of information and classes, and the lack of robustness (Singer, 1999).

Besides the assembly hypothesis, there are other hypotheses that are based on the concept of synergistic population coding, for example population rate coding (Georgopoulos et al., 1988) and population rank coding (Fabre-Thorpe et al., 1998). Contrary to the cell assembly hypothesis, the population rate code does not assume that relations between neurons are based on the fine-temporal structure of spiking activity on a millisecond scale but rather by slower modulations of the spike rate, defined on a timescale of more than 10 ms. In contrast, the rank coding hypothesis from Fabre-Thorpe postulates a latency code that assumes that the spike timing of the first spike of each neuron in relation to an event, such as a stimulus, contains the information. Thus, when discussing the characteristics of these hypotheses, one has to differentiate between both the single-cell and population coding as well as the timescale that is assumed to be involved in information processing.

A different class of hypotheses about the neuronal code are the concepts of echo-states (Jaeger, 2002) and liquid state machines (Haeusler and Maass, 2006; Kaske and Maass, 2005; Maass et al., 2002; Maass et al., 2004; Melamed et al., 2004; Natschlager and Maass, 2005). The basic idea underlying both concepts is that time delays expand the dimensionality of time-continuous signals to infinity. Thus, the concepts of

1. *The Neuronal Code*

liquid and echo-states are similar to the idea of kernels used by support vector machines (Schölkopf and Smola, 2002). The kernel serves as an expansion of the feature space due to a non-linear mapping, with the consequence that linearly non-separable data in the original feature space eventually becomes linearly separable in the additional dimensions introduced by the kernel. The same is the case for echo-states and liquid state machines in the time domain. The expansion of the dimensionality of the data caused by delays in the system likely maps features of spiking activity onto a linearly separable space. The most appealing properties of these machines are that they have been shown to have universal computational power, while they are at the same time biologically plausible, since they incorporate memory and are not engineered. Nevertheless, the disadvantages are four-fold. First, any useful computation based on a liquid state machine is crucially dependent on the classification by the readout which has to be trained (Maass et al., 2002). Second, biologically plausible implementations of readout training is not yet well understood (Legenstein et al., 2005). Third, the importance and impact of properties that modify and shape neuronal activity, such as synaptic plasticity, are so far pretty unclear (Izhikevich et al., 2004; Lazar et al., 2006). Fourth, different readouts may be required for different epochs in time to allow for stable feature classification across time.

1.3. Testing the Assembly Hypothesis

There are many polarized and controversial discussions in the fields of experimental and theoretical neuroscience about the neuronal code. Their resolution requires tools and techniques that allow for conclusive tests to support or falsify hypotheses. What is needed is first, precisely defined and, if possible, simple working hypotheses and second, analysis tools that are well focused and free from assumptions. This motivated us to develop new tools and to apply them to data to investigate the concept of cooperative and synergistic coding in the brain. In the first part of this thesis, we present four non-parametric tools for the analysis of oscillations and synchronization of neuronal activity, which allow for the required conclusive tests. In the second part, we apply the new tools to data recorded simultaneously with multiple electrodes in awake monkeys and anaesthetized cats. The results demonstrate the existence of task and behavior related neuronal synchronization that indicate cooperative neuronal activity and strongly supports the assembly hypothesis.

2. Neuronal Signals and Electrophysiological Recordings

Extracellular electrophysiology is currently the best technique for monitoring the activity of small populations of neurons in an awake animal. In dependence of the spatial extent of the recorded neuronal population there are two components of the signal that have to be distinguished (Fig. 2.2): First, the *spiking activity* of the very next cells of the electrode, and second, the *Local Field Potential* ('LFP') caused by the synaptic currents from larger groups of cells. As a mechanistic interpretation spiking is often understood as the output because it is generated by individual cells based on the integration of their synaptic inputs, whereas the LFP is understood as the average input to a group of cells since it mostly represents the average synaptic activity of the neurons in the recorded area.

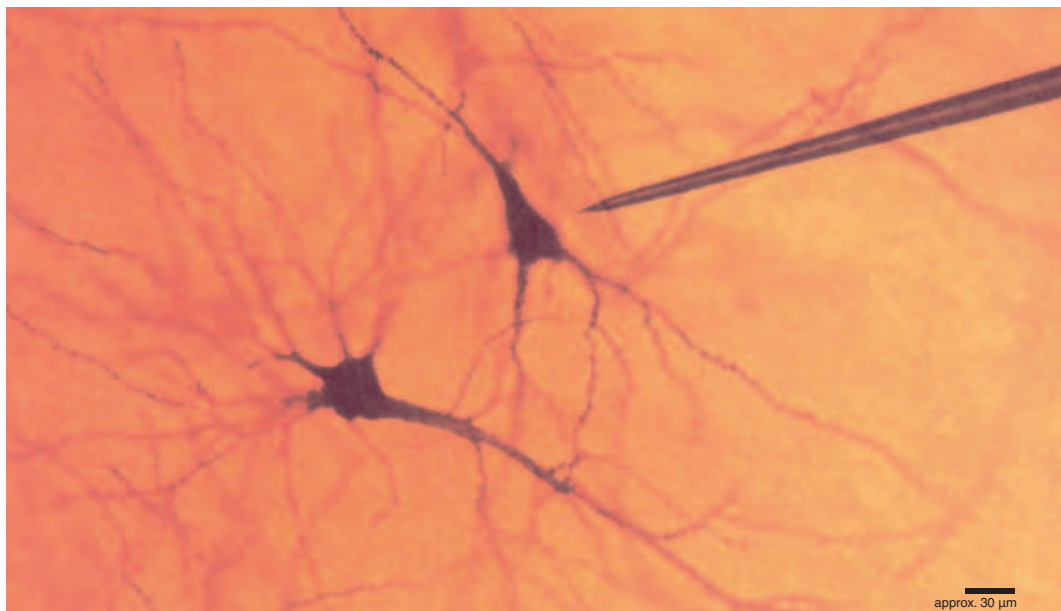


Figure 2.1.: **Schematic picture of a recording electrode in tissue.** Schematic picture of an extra cellular recording. The recording electrode measures electrical signals originated by a few neurons surrounding it. The scale bar is indicating a scale of 30 μm .

2.1. Spiking Activity

The *spiking activity* of a cell corresponds to the activity that is transmitted via the axon to other neurons.

2.1.1. Nature and Source of Spiking Activity

Measuring the electric activity of neurons is possible because neurons open transiently sodium channels in order to fire an action potential. This allows positively charged sodium ions to rush down the voltage gradient into the cell which creates a negative change in potential in the immediately surrounding area. Consequently, the latter leads to a transient change in voltage between the extracellular recording electrode and the distant reference electrode. The signal of a typical extra cellular recording is constituted by the activity of multiple neurons surrounding the electrode (Fig. 2.1).

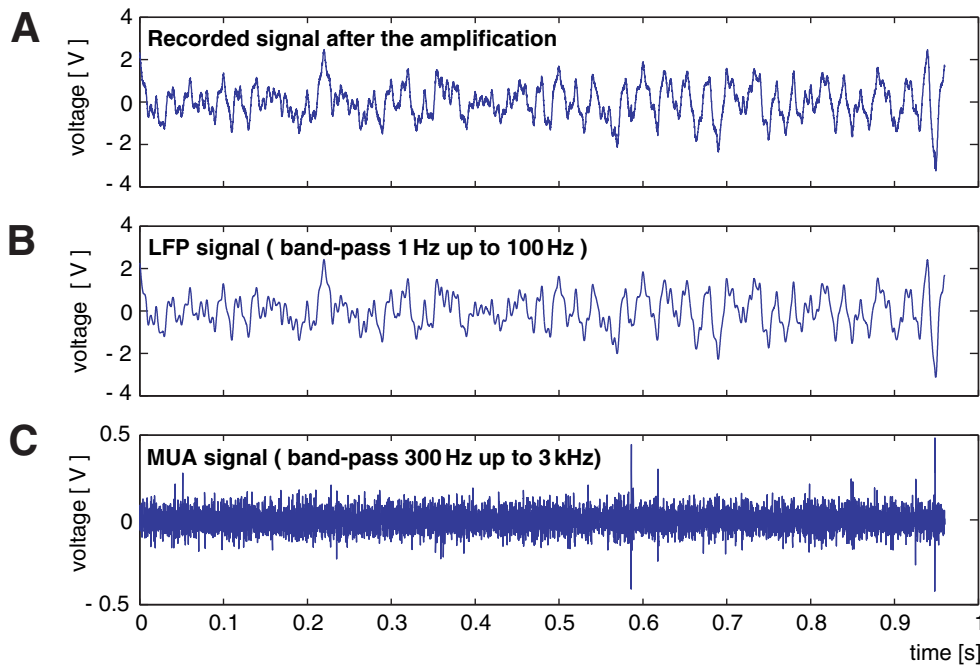


Figure 2.2.: **Multi-unit activity (MUA) and local field potential (LFP) signal.** An analog signal recorded in the visual cortex of a cat with a sampling frequency of 10.04 ks/s. **(A)** Amplified original signal of the recording electrode. Low frequency components dominate in the signal. **(B)** LFP: Low frequency components of the original signal shown in (A) filtered with a band-pass filter with cut off frequencies at 1 Hz and 100 Hz. **(C)** Spiking MUA signal: High frequency components of the original signal shown in (A) but filtered with a band-pass filter with cut off frequencies at 300 Hz and 3 kHz. The high-frequency components of signal (A) become visible in (C) only after filtering, since their amplitude is comparably small to the amplitude of the low frequency component.

2.1.2. Recording of Neuronal Activity

The spiking activity is a very fast signal on a timescale of one millisecond and signal components that range up to a few kHz, while the LFP is a comparably slow signal with signal components between 1 and 300 Hz. Thus, utilizing two band-pass filters, each either optimized for the frequency components of the spiking or the LFP signal, allows to record both signals at the same time from only one recording electrode (Figure 2.3). A typical filter used to record a spike signal is a band-pass between 300 Hz and 5 kHz (-3 dB), while the LFP component is extracted based on a band-pass with a bandwidth of 1 Hz to 300 Hz (-3 dB). Typical electrodes used for recording have a diameter of 20-200 μm and an impedance in the range of 0.1-10 $M\Omega$ at 1 kHz (e.g. Wire-electrodes (\emptyset : approx. 25 μm) or Tunkston-electrodes (\emptyset : approx. 100 μm), the length of the conical tip approx. 10-15 μm (Harris et al., 2000)). Despite this quite small size they are still thick in relation to a nerve cell (soma 10-30 μm). Therefore, most recorded spike signals are composed from signals of several cells (Fig. 2.1). Spiking activity of several neurons is referred to as *Multi-Unit Activity* ('MUA').

2.1.3. Spike Detection

Spikes are primarily detected by a window discriminator that detects signals that have amplitudes that are in between of a lower and an upper threshold. Signals which are rejected are those, that have a smaller amplitude than a lower bound and are therefore considered as noise, as well as those signals, that exceed an upper threshold (Fig. 2.4) and therefore are likely to be induced artificially by other sources than neurons (e.g. electrical equipments like computers or monitors, etc.).

2.1.4. Spike Sorting

To identify activity from single-neurons one has to identify the different sources of the MUA signal. Since the shape of the spikes for each neuron is very stereotypical, the spike waveform can be used to identify different neurons as sources of the MUA signal. Even though the spike waveforms from different neurons, but from the same class, are very similar, their separability is increased by modifications of the signals on their way from the cell to the electrode. Modifications of the signal are likely to be different for different neurons because of different distances of the neurons to the electrode and because of the inhomogeneous consistence of the tissue in respect to its electric properties. As a first order approximation the modifications of the signal can be described by a low-pass filter. Thus, spikes from neurons which are further away from the electrode are damped and contain less high frequency components leading to less transient slopes. Hence, the latter modifications can be utilized to sort the MUA signal based on the typical spike waveforms, that are slightly different for each recorded neuron, however (Gray et al., 1995; Harris et al., 2000; Lewicki, 1998; Vollgraf et al., 2005). Sorted data is referred to as *Single-Unit Activity*

2. Neuronal Signals and Electrophysiological Recordings

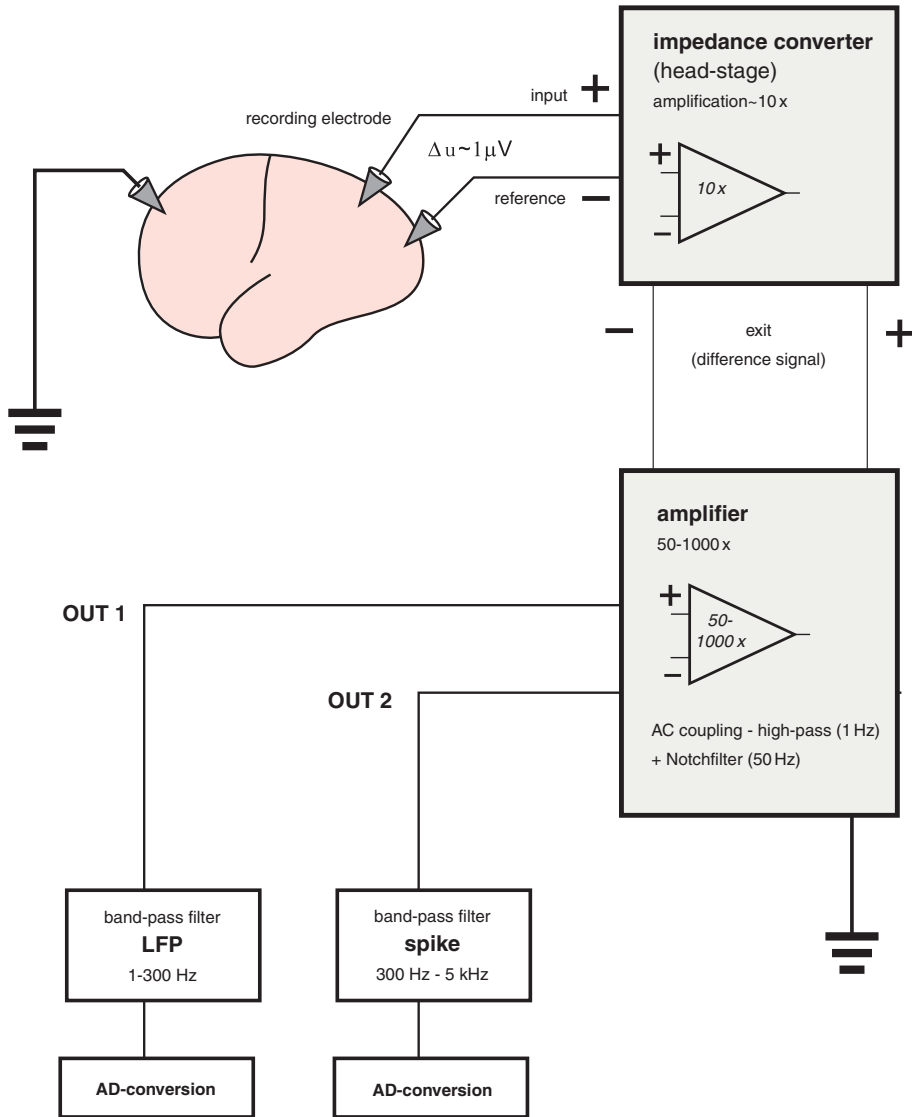


Figure 2.3.: **Recording setup.** Typical amplifier-system for recording neuronal activity used in electrophysiology. The amplified signal is a difference signal between a reference electrode and a recording electrode. The signals go first through an impedance converter ('head-stage') with a minor amplification. In order to minimize the impact of potential noise sources this head-stage is positioned as close as possible to the electrodes. Next, after the head-stage, the main amplifier increases the signal amplitude with a gain of about 50-1000. To prevent saturation of the amplification caused by small DC-offsets between the reference and the recording electrode, both, the head-stage and the main amplifier are AC-coupled. After the amplification, the very transient spike signal and the low frequency LFP become separated by two band-pass filters.

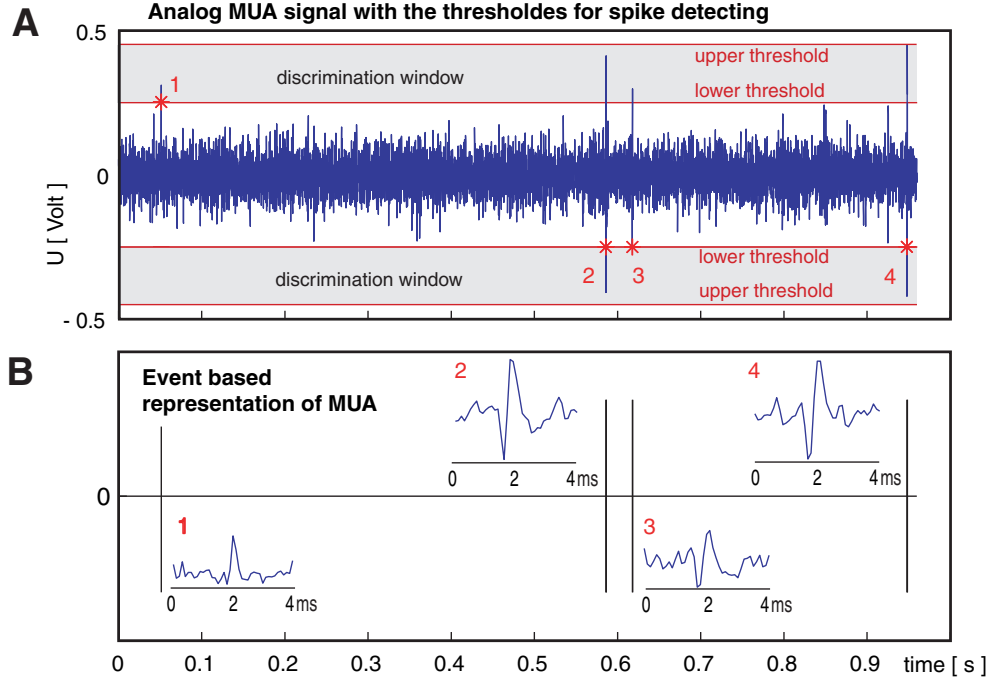


Figure 2.4.: **Analog signal and detection.** (A) Window discrimination used to detect multi-unit activity ('MUA') in high-pass filtered signal. The window discrimination triggers on events whose absolute value of the voltage U is in between of a lower and an upper bound. (B) Analog signal shown in (A) is mapped to an event based representation of the MUA signal. Red lines indicate events on the time axis. In addition, the spike waveforms of each of the four spikes detected in (A) are visualized. Single-unit activity ('SUA') that has been identified based on clustering of features of the spike waveforms. Four detected spikes can be assigned to three different cells (spike 1 = cell 1 / spike 2, 4 = cell 2 / spike 3 = cell 3).

('SUA'). In Appendix G.1 spike sorting of data recorded in an awake monkey is more technically addressed (Software package: MClust by A David Redish, University of Minnesota).

2.1.5. Event Based Representation of Spiking Activity

The waveform of a spike is stereotypical for each neuron. Thus it is assumed that the shape of a spike itself does not transmit information. As a consequence, the information contained in spiking activity can be reduced on the spike timing. In case the timing is sampled with a sampling frequency f_s it is mapped onto a time series $v_i(t)$, where v is the total number of spikes per bin. To maintain the whole information contained in the spike times of the neuron, the sampling frequency $f_s = 1/\Delta_s$ has to be chosen high enough to resolve the intrinsic time scales of the neuronal spiking activity. The lower limit of this is assumed to be in a range of 0.1 to 1 ms. Therefore it is reasonable to reduce the sampling frequency $f_s = 1/\Delta_s$ to $f_b = 1/\Delta_b$ by binning with a bin width of Δ_b . ($\Delta_b = 0.1, \dots, 1 \text{ ms}$)

$$v_i(t) = \text{number of spikes in } [t, t + \Delta_b) \quad \text{with} \quad t = 0, 1\Delta_b, 2\Delta_b, \dots, (B-1)\Delta_b \quad (2.1)$$

2. Neuronal Signals and Electrophysiological Recordings

If the bin width Δ_b of the time series $v_i(t)$ is smaller than the expected minimal inter-event interval ($\sim 1\text{ ms}$), that is corresponding to the absolute refractory period, the time series $v_i(t)$ is expected to be binary in case of a SUA signal. In case of a MUA signal it might happen that more than one spike is falling into one bin, since the inter-event period between spikes originated by different neurons can be arbitrarily small.

2.2. Local Field Potential (LFP)

In comparison to the spiking activity the *Local Field Potential* ('LFP') is a slow signal composed by the activity from a very large group of cells whose size ranges between a few hundred to thousands of cells. The spatial extent ranges between a few hundred micrometers to a few millimeters.

2.2.1. Source of Local Field Potentials

A cell responses on a received excitatory synaptic input by opening ion channels that allow currents to flow into the cell. Likewise, an inhibitory synaptic input often results in a current flow out of the cell. Thus, each presynaptic event induces changes of the concentration of ions in the very local proximity of the synapses. The latter results in individual electric fields. The superposition of each individual electric field is described by the LFP. Thus, the LFP recorded at any given site at any given time reflects the sum of fields generated by current sources (e.g. 'EPSPs: **E**xcitatory **P**ost-**S**ynaptic **P**otentials') and current sinks (e.g. 'IPSPs: **I**nhibitory **P**ost-**S**ynaptic **P**otentials') in the close proximity of the recording electrode. The recorded fluctuations thus can reflect the synchronized synaptic activity of the population of neurons in the local area surrounding the recording electrode. Fast sodium action potentials do not significantly contribute to this field recording because of the low-pass (capacitative) filtering properties of the extracellular environment that separates the recording electrode and the cell (Nadasdy et al., 1998).

3. Neuronal Assemblies and Coordinated Neuronal Activity

The assembly hypothesis assumes that groups of neurons ('assemblies') are defined by temporally coordinated spiking activity. Thus, to study the assembly hypothesis, *temporal patterns of neuronal activity* have to be studied. This chapter discusses different kinds of temporal patterns of neuronal activity. To this end, it first introduces *spike patterns* and discusses related concepts. Second, patterns based on mass activity of large groups of neurons recorded by the *Local Field Potential* ('LFP') are introduced and discussed.

3.1. Temporal Pattern of Spiking Activity

Temporal pattern of spiking activity has been defined in at least two different ways (see the following Subsections 3.1.1 and 3.1.2).

3.1.1. Synchronous Events (Joint Spike Activity)

The first definition is based on synchronous spiking of at least two neurons with a millisecond precision (Fig. 3.1A). These coordinated events have been referred to as a synchronous firing (Gray et al., 1989; König, 1994), a temporal pattern (Vaadia and Abeles, 1987), or a joint spike event (Grün et al., 2002a; Grün et al., 2002b). In this thesis the term *Joint Spike activity* ('JS activity') will be used to refer to a pattern arising from coordinated firing. Each incidence of a coordinated firing event is referred to as a *Joint Spike Event* ('JSE'), while the identity of a JSE is defined by the set of neurons that have exhibited the coordinated firing. This set of neurons is referred to as a *Joint Spike pattern* ('JS pattern').

3.1.2. Spatiotemporal Pattern

The second type of temporal pattern of spiking activity is defined as two or more neurons spiking in sequence over a finite period of time (Fig. 3.1B). Such a sequence has been referred to as a *spatiotemporal pattern* (Abeles and Gerstein, 1988) or, more simply, a pattern (Ikegaya et al., 2004). Spatiotemporal patterns can arise from intrinsic properties of the recurrent network of neurons or from synchronous

3. Neuronal Assemblies and Coordinated Neuronal Activity

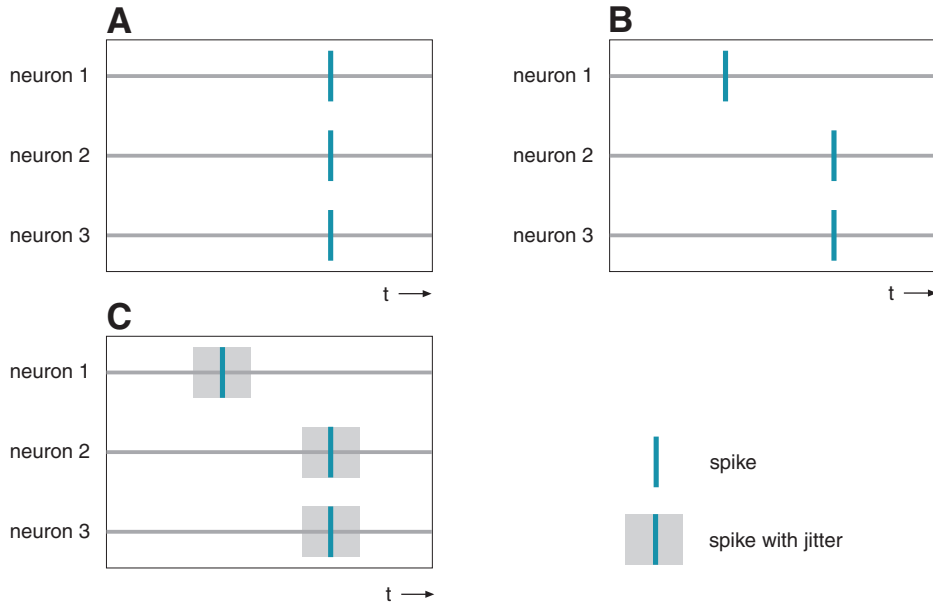


Figure 3.1.: **Different temporal patterns of three neurons.** (A) Joint spike activity of three neurons. (B) Spatiotemporal pattern of three neurons. In case of a spatiotemporal pattern neurons are spiking in a fixed sequence. (C) Statistical variability ('jitter') of the spike timing indicated by grey shadowed areas.

events which are delayed by different transmission delays (Salinas and Sejnowski, 2001; Vogels et al., 2000).

3.1.3. Variability of Spike Timing (Jitter)

Beside the systematical shifting of the spike timing in the case of spatiotemporal pattern, the spike timing of each individual spike can vary due to statistical variability (Abeles, 1991; Feng and Brown, 1998; Gur et al., 1997; Ikegaya et al., 2004; Lestienne, 1995; Marsalek et al., 1997; Salinas and Sejnowski, 2001). The latter variability can either arise by random fluctuations or by changes of the system. This variation is referred to as '*jitter*' (Fig. 3.1C). The upper limits of a jitter is assumed to be in a range of 1 to 10 ms (Aertsen et al., 2001; Bi and Poo, 1998; Grün et al., 1999; Hopfield and Brody, 2000; Hopfield and Brody, 2001; Koch, 1999; Mainen and Sejnowski, 1995; Markram et al., 1997; Sjostrom et al., 2001).

3.1.4. Pattern Complexity and Correlation Order

The complexity of a JSE is defined by the number of neurons participating in the event. It is important to note that the definition of a JSE allows each neuron to participate in any JSE only once, so that the

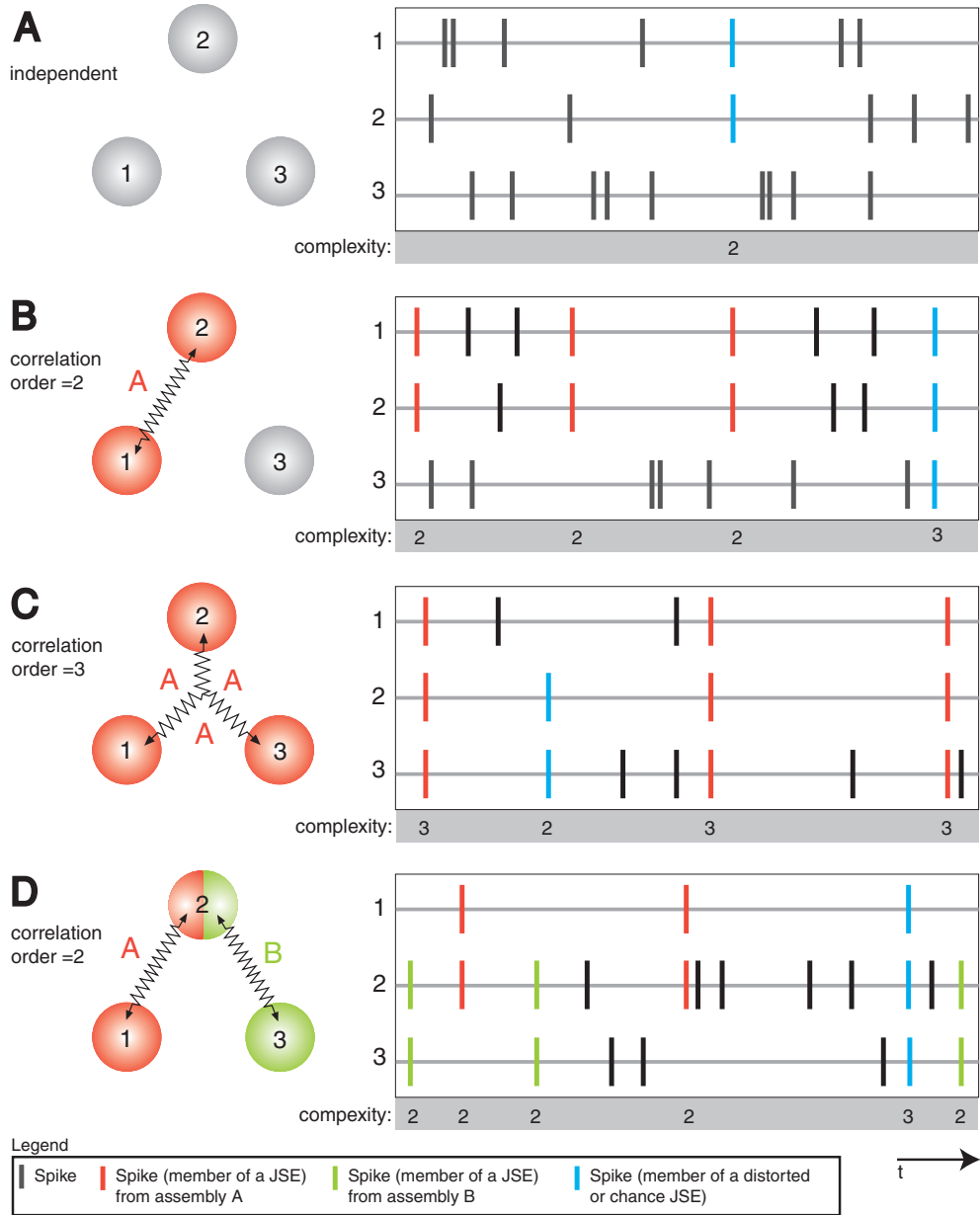


Figure 3.2.: **Synchronization of spikes by neuronal coupling.** (A-D) show three units, indicated by 1-3, representing either single-neurons or three neuronal populations that are mutually independent. The right sub-panel shows examples of the resulting set of spike trains. Under each set the complexity of occurring joint spike events (JSEs) is indicated on top of a grey shadowed bar. In the case that the units represent single-neurons, each spike train contains single-unit activity (SUA), while each spike train is composed of multi-unit activity (MUA) if the units represent small populations of neurons. (A) Since the units are not coupled, they are mutually independent and exhibit spike trains that do not share synchronized spikes beyond chance level. JSEs that are occurring just by chance are blue. In (B), (C) and (D) units are coupled. JSEs that are induced by underlying coupling of cells are red or green. Induced JSEs, that are induced and perturbed, are also blue. Spike trains produced by coupled units share synchronized spikes (JSEs) (indicated by red color). (B) A second order correlation between units 1 and 2 in a JSE of complexity 3. (C) A third order correlation between units 1, 2, and 3. (D) Two second order correlations A and B between units 1, 2, and 3. The first second order correlation A couples 1 and 2, while the second, B, couples 2 and 3.

3. Neuronal Assemblies and Coordinated Neuronal Activity

number of participating neurons is equal to the total number of spikes per event. The assembly hypothesis assumes that JSEs are caused by correlated neuronal firing of a group of cells. Thus, the concept of the complexity (Baker and Lemon, 2000) of a JSE has to be separated from the concept of the order of a correlation (Fig. 3.2). The order of a correlation is defined by the number of neurons that are directly coupled, which is not necessarily the same as the total number of neurons participating in the JSE due to one or more spurious events (Fig. 3.2). Therefore, the correlation is the source of coordinated firing. Since the observed processes are most likely mixtures of different correlation structures and noise, JSEs are likely to be distorted. Thus, the order of a neuronal correlation is, in most cases, not equal to the complexity of the observed JS patterns.

3.2. Local Field Potential and Pattern of Neuronal Activity

In comparison to the spiking activity, LFP is a slow signal composed by the activity from a very large group of cells whose size can range between a few hundred to thousands of cells. The spatial extent ranges between a few hundred micrometers to a few millimeters (Nadasdy et al., 1998). Both, the spatial extent and the size of the recorded population are determining the nature of the signal. Given such a large number of sources one would expect that the average population activity, that is built up by linear superpositions of the fields of the individual sources, has very tiny fluctuations only (central limit theory). That implies, that significant modulations of the LFP cannot be explained by independent activity of the neurons belonging to the recorded population. Thus, already the analysis of the LFP from a single electrode allows to study the synchronization of neurons in local populations with a spatial extent that ranges between a few hundred micrometers to a few millimeters.

3.2.1. Rhythmicity of the Local Field Potential

When neurons in a network simultaneously receive many concerted inputs, this population events can be recorded as local voltage fluctuations in the extracellular field. If these population events occur regularly, the resulting voltage fluctuations can be seen as rhythmical oscillations. Consequently, frequency decomposition allows to study the individual oscillating components that contribute to the LFP. Tools that are frequently applied to transform the signal into the frequency domain are the Fourier (Percival and Walden, 1993), the wavelet transform (Lachaux et al., 1999), filtering and hybrid methods like the multitaper method (Percival and Walden, 1993) that combines the idea of the Fourier decomposition with special types of preprocessing, like tapering, that allows an optimal concentration of the signal power in a certain frequency band.

3.2.2. Frequency-Locking, Coherent Oscillations and Phase-Locking

Beyond the analysis of the LFP signal of one electrode, the analysis of large sets of simultaneously recorded LFPs is of interest. The latter allows to study the interaction between spatially distinct sites by studying their coherence, frequency-locking or phase-locking.

3.3. Coupling of Complex Systems

Coupled systems are understandable if they are simple like two pendulums coupled via a spring. For more complex systems the interaction becomes pretty soon hard or impossible to be described analytically. Nevertheless, one can study coupled systems even if they are complex by analyzing the time courses of state variables from the systems on oscillations and synchronization.

3.3.1. $n:m$ Frequency and Phase-Locking in Complex Systems

Oscillatory processes at different frequencies can interact based on coupling that leads to synchronization. Synchronized oscillations can either be frequency synchronized, which implies that the frequencies (ω_0 and ω_1) of the two synchronized oscillations match an $n : m$ ratio ($n * \omega_0 = m * \omega_1$), or phase-synchronized, which implies that in addition to the frequency synchronization the phases of both oscillations are correlated.

3.3.2. Synchronization by Coupling: Arnold Tongues

In contrast to synchronized oscillations, independent oscillations can have any ratio of frequencies. Thus, the process of synchronization implies tuning of the coupled processes to adapt their frequencies to match an $n : m$ relation. The force necessary for synchronizing can be dependent on the type of the system, the absolute frequencies (ω_0, ω_1), the necessary amount of detuning as well as the ratio $n : m$ of the frequencies. *Arnold tongues* describe the force (ε) required for coupling a first system (1) to a second system (0), oscillating at a given frequency ω_0 . Each tongue describes the required force in relation to the amount of necessary detuning. With increasing tightness of Arnold tongues the force necessary for the same amount of detuning increases. Typically higher order $n : m$ synchronization requires stronger forces than lower order synchronization like 1:1. For further explanation, we cite from a recent book "Synchronization" (Cambridge University Press, 2001) by Pikovsky, Rosenblum and Kurths.

Citation 1, page 65: '*Thus, an oscillator with frequency ω_0 can be entrained by a force having a frequency close (but not necessarily equal!) to $\omega_0/2$, and synchronization then appears as the onset of the following relation between the frequencies: $2 * \omega_0 = \Omega$.*', here $\Omega = \omega_1$ '*This regime is called synchronization of order 2:1. Obviously, entrainment by every third pulse can be achieved as well, although it would*

3. Neuronal Assemblies and Coordinated Neuronal Activity

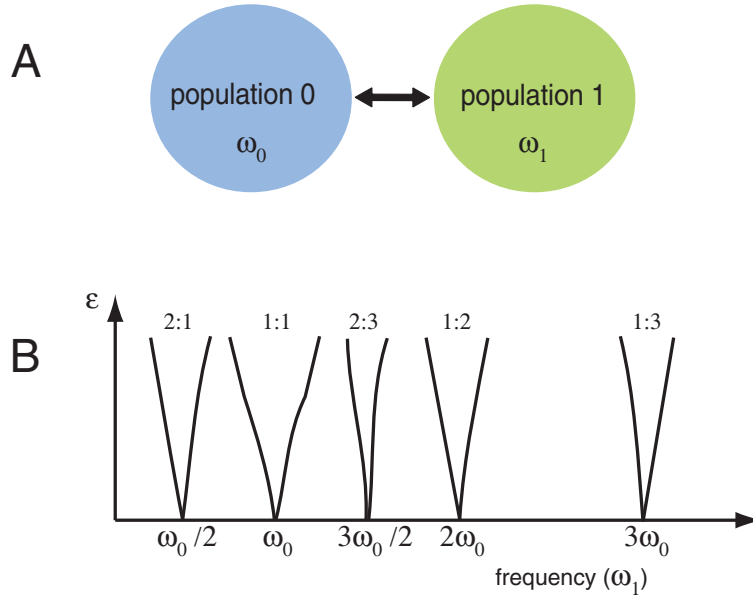


Figure 3.3.: $n : m$ **Synchronization and Arnold tongues.** (A) Coupling of population 0 oscillating at frequency ω_0 , and population 1 oscillating at the frequency ω_1 . (B) Schematic representation of Arnold tongues, or regions of $n : m$ synchronization. The number at the top of each Arnold-tongue indicates the order of locking. Y-axis (ϵ) gives the driving force necessary for detuning.

require an even higher amplitude of the pulses for the same detuning. Generally, the synchronous regimes of arbitrary order $n : m$ (n pulses with m oscillatory cycles) can be observed, and the whole family of synchronization regions can be plotted.' (Fig. 3.3) 'The regions are now commonly called Arnold tongues. It is important to mention that higher-order tongues are typically very narrow so that it is very difficult (if not impossible) to observe them experimentally. We can see that, for the same values of detuning synchronization of order 2:1 requires an essentially larger amplitude of pulses. On the contrary, if an amplitude is fixed, then resetting by, say, every second pulse can compensate a smaller detuning than resetting by every pulse, meaning exactly that the region of 2:1 frequency locking is narrower than the region of 1:1 locking.'

Citation 2 from 'Synchronization' (Cambridge University Press, 2001) by Pikovsky, Rosenblum and Kurths, page 67: 'The synchronization properties we have described are general for weakly coupled forced oscillators, and independent for the features of the particular system, i.e. whether it is a quasilinear or a relaxation oscillator. They are also independent of the form of the periodic forcing, whether it is harmonic, rectangular, or pulse-like. Generally, synchronization of order $n : m$ can be observed, with Arnold tongues touching the ω -axis, this means that synchronization can be achieved by an arbitrary small force' (Fig. 3.3).

4. Tools to Detect Coordinated Neuronal Activity

In this chapter we are going to present tools to detect coordinated neuronal activity. We will discuss frequency domain approaches to analyze local field potentials ('LFPs') and spiking activity on temporal structure as well as synchronization of the signal phase across pairs. We will also present the basic concept of the *Unitary Event method* ('UE method') (Grün et al., 2002a; Grün et al., 2002b) that was developed to analyze spiking activity from simultaneously recorded neurons on joint spike ('JS') activity. Subsequently we are going to compare tools and discuss their constraints and limitations.

4.1. Tools for Analyzing the Local Field Potential

In following section we are going to introduce the basic ideas of two widely used concepts *coherence* and *phase-locking*, (Lachaux et al., 1999; Percival and Walden, 1993) for assessing synchronization of the LFP.

4.1.1. Definition of Coherence

Coherence is a measurement of covariation of power and phase between two signals $x(t)$ and $y(t)$. It is defined by c_{xy} based on the power spectral density P_{xx} of x and the power spectral density P_{yy} of y as well as the cross power spectral density P_{xy} of x and y :

$$c_{xy} = \frac{|P_{xy}|^2}{P_{xx}P_{yy}} \quad (4.1)$$

The power spectral densities P_{xx} and P_{yy} of x and y and the cross power spectral density P_{xy} is derived based on the Fourier transform $X(f)$ and $Y(f)$ from $x(t)$ and $y(t)$. The latter is defined given a sampling frequency of f_s by

$$X(f) = \sum_{m=-\infty}^{+\infty} x(m) e^{-2\pi i m \frac{f}{f_s}} \quad (4.2)$$

Then P_{xx} , P_{yy} and P_{xy} are defined by:

$$P_{xx} = X(f) * X(f)^* \quad (4.3)$$

4. Tools to Detect Coordinated Neuronal Activity

$$P_{yy} = Y(f) * Y(f)^* \quad (4.4)$$

$$P_{xy} = X(f) * Y(f)^* \quad \text{and} \quad |P_{xy}|^2 = P_{xy} * P_{xy}^* \quad (4.5)$$

4.1.2. Definition of Phase-Locking

Phase-locking detects synchrony in a precise frequency range between two recording sites. It uses responses to a repeated stimulus and looks for latencies at which the phase difference between the signals varies little across trials (phase-locking). Given two series of signals $x(t)$ and $y(t)$ and a frequency of interest f_{oi} , the procedure computes for each latency a measure of phase-locking (the measure itself is referred to as *Phase-Locking Value* or 'PLV') between the components of $x(t)$ and $y(t)$ at the frequency of interest f_{oi} . The procedure follows three steps (Lachaux et al., 1999):

Step 1. Band-pass filtering of $x(t)$ and $y(t)$ to make both signals $x'(t)$ and $y'(t)$ narrow banded with frequency components centered by the frequency of interest. In principle, this step is technically not necessary since phase-locking could be computed for unfiltered signals. But, nevertheless, only filtered signals allow a meaningful interpretation of the phase-locking value (Lachaux et al., 1999).

Step 2. Convolution of both signals $x'(t)$ and $y'(t)$ with a complex Gabor wavelet $G(t, f)$ centered at frequency f_{oi} .

$$G(t, f_{oi}) = e^{\left(\frac{-t^2}{2\sigma_t^2}\right)} * e^{(2i\pi f_{oi}t)} \quad \text{with } \sigma_t = \kappa/f_{oi} \quad \text{with } \kappa \text{ usually } 7 \quad (4.6)$$

The latter step extracts the instantaneous phase of $x(t)$ and $y(t)$.

$$X'(t, f_{oi}) = \int x'(t-t')G(t', f)dt' \quad (4.7)$$

As an alternative approach to the wavelet decomposition, one can use a standard Fourier transform to extract the phase information of the filtered signals $x'(t)$ or $y'(t)$.

Step 3. The phases $\varphi(x'_m(t))$ of $x'(t)$ and $\varphi(y'_m(t))$ of $y'(t)$ is derived for each trial m ($m = [1, \dots, M]$). Next, the pairwise phase difference $\theta_m(t)$ between $\varphi(x'_m(t))$ and $\varphi(y'_m(t))$ is computed for each trial m .

$$\theta_m(t) = \varphi(x'_m(t)) - \varphi(y'_m(t)) \quad (4.8)$$

The PLV is then defined at time t as the average value of $\theta_m(t)$ across trials.

$$PLV(t) = \frac{1}{N} \sum_{n=1}^M e^{i\theta_m(t)} \quad (4.9)$$

The measurement PLV is bounded between 0 and 1. A PLV of 1 corresponds to perfect phase-locking of signal x and y across all trials while a PLV close to zero indicates uncorrelated phases of x and y .

The difference between the wavelet and the Fourier approach is that the wavelet analysis is a scale analysis, that means that the wavelet length σ_t is adapted to each frequency of interest, in such a way that always κ cycles are covered by the wavelet. In contrast, the sliding window length of the Fourier transform is constant across all frequencies. This implies that the wavelet is more precise in localizing high frequency components in the time domain, but increases the frequency uncertainty, since the windows are comparably shorter than in the Fourier analysis. A disadvantage of this adaptation of the window length concerning the statistical evaluation of the PLV is, that the amount of sampling points entering the estimation for different frequencies of interest is different. Therefore we have used the wavelet transform to screen our data based on time frequency plots of power and phase-locking, but we also have used Fourier based approaches that ensure the same number of samples across frequency bands for statistical issues.

4.1.3. Comparison of Coherence and Phase-Locking

Coherence does not specifically quantify phase relationships since coherence also increases with amplitude covariance. Thus, the relative importance of amplitude and phase covariance in the coherence value is not clear. Since phase-locking is not including any information about the amplitude of the signal, it is, by definition, not influenced by the amplitude covariation.

Nevertheless, the advantage of the phase-locking method does not hold for stochastic signals, since the signal-to-noise ratio might change. In this scenario an increase of the power contained in a certain frequency band, of course, influences the reliability of the phase-detection. Thus we expect a covariation of the PLV with the modulation of the signal-to-noise ratio if the signals are phase-locked. And indeed, instances in which the power exhibits covariation with the PLV are often observed in LFP and EEG studies. Still, although this is a very important point for electrophysiological signals, since they are partly stochastic and the signal-to-noise might be modulated, it has not been acknowledged so far as a serious drawback of the phase-locking analysis. Nevertheless, in case of a high signal-to-noise ratio, the phase-locking analysis is able to detect changes in the phase-locking if they are modulated stronger by the internal coupling and decoupling than by task induced changes of the signal-to-noise ratio. Thus, only if one can demonstrate on the recorded data that modulations of the power and of the PLV are occurring at different times and at different frequencies, one can consider this phase-locking modulation as independent of modulations of power.

4.1.4. New Approaches to Identify Coordinated Neuronal Activity

Motivated by the fact that both the coherence and phase-locking methods do not allow to estimate a confidence interval analytically, we developed new tools that utilize bootstrapping of the phase-locking and coherence to estimate confidence limits as well as to allow hypothesis tests in case of bivariate and multivariate datasets (see Chapters 5, 8, and Appendix E). Previous approaches (Multi-taper analysis and Chronux Software toolbox for Matlab, see also (Mitra and Pesaran, 1999)) used the delete-one-jackknife (drop one of m) to derive the confidence limits of the coherence in terms of a standard deviation of the coherence across M trials. The disadvantage of the standard delete-one-jackknife approach is, that it is rather inflexible and does not always derive correct estimates (Efron and Tibishirani, 1993). Thus, the method we have developed uses the n -jackknife that drops n elements instead of one, in combination with the concept of bootstrapping based on random resampling. Both, the n -jackknife and the bootstrap have been proofed in contrast to the normal jackknife, to be universal methods that allow to derive the variability and confidence limits of an estimator robustly (Efron and Tibishirani, 1993).

Beyond the estimation of confidence intervals it is important to be able to compare the influence of different experimental conditions on the strength of coherence and phase-locking (see Chapters 5, 8, and Appendix E). To this end we have developed two tools that utilize a combination of n -jackknife and bootstrapping with a permutation test to estimate the statistical significance of observed differences in coherence, phase-locking or UE p -values across different experimental conditions.

4.2. Tools for Spiking Data

The assembly hypothesis assumes that coordinated neuronal firing plays an important role in the encoding and processing of information in the brain (Fetz, 1997; Gerstein et al., 1989; Hebb, 1949; Singer, 1999; von der Malsburg, 1981). Multiple methods have been developed to detect coordinated spiking events and to investigate whether these events are correlated either to information processing or to states in the neuronal system (Abeles and Goldstein, 1977; Abeles and Gerstein, 1988; Aertsen and Gerstein, 1985; Baker and Gerstein, 2000; Barbieri et al., 2004; Barbieri et al., 2005; Brown et al., 1998; Brown et al., 2004; Czanner et al., 2005; Gerstein and Perkel, 1969; Gerstein and Perkel, 1972; Gray et al., 1989; Grün et al., 1999; Grün et al., 2002a; Grün et al., 2002b; Ikegaya et al., 2004; Kass et al., 2005; König, 1994; Martignon et al., 2000; Nakahara and Amari, 2002; Okatan et al., 2005; Pipa and Grün, 2003; Radons et al., 1994; Samonds and Bonds, 2004; Sharpee et al., 2004; Tetko and Villa, 2001). The methods differ in the definitions of coordinated firing patterns, the techniques to detect these patterns, and the approaches to analyze the resulting data (descriptive, statistical hypothesis testing, maximum likelihood, and Bayesian approaches).

4.2.1. Variability, Rareness, Non-Stationarity and Auto-Structure

Even though the assembly hypothesis formulates precisely what constitutes a coordinated spiking event, it turns out to be a non-trivial problem to design a method that detects the existence of such events and investigates their information content, without being confounded by other properties of the data (Baker and Gerstein, 2001; Brody, 1999; Grün et al., 2003; Oram et al., 1999; Roy et al., 2000).

Four main properties of neuronal spike trains make the analysis of coordinated spiking events difficult:

1. **Variability in time and across trials.** Neuronal recordings show a high degree of variability, which is partially caused by changes in the properties of responses that reflect the dynamics of the information processing in the brain. Another source of variability is the intensity with which cells respond to the presentation of the same stimulus. Any analysis of coordinated firing events has to consider both kinds of variability.
2. **Changing properties in time on short time scales (e.g. rate).** Evidence indicates that neuronal states linked to the processing of information can last for only a short period of time (e.g., a few tens of milliseconds; (Oram and Perrett, 1992; Thorpe et al., 1996)). This implies that analysis methods need to operate on short time series containing a very limited number of samples.
3. **History dependencies and auto-structures.** Neuronal activity might have strong history dependencies, where the likelihood that a spike will occur at a certain point of time depends on the times at which previous spikes have occurred. Therefore, the auto-structures of the spike trains need to be considered during the analysis.
4. **Rareness of events.** Past investigations of coordinated firing events indicated that these events are rare in respect to time and to space (Abeles and Gerstein, 1988; Ikegaya et al., 2004). Thus, detection of coordinated firing events might be difficult, since even highly parallel recordings from up to a hundred simultaneous recording sites might still grossly under-sample the neurons that participate in such events.

4.2.2. Unitary Event Analysis

The *Unitary Event analysis* ('UE analysis', see (Grün et al., 2002a; Grün et al., 2002b)) was one of the first approaches that allowed the analysis of simultaneously recorded spiking activity. Since both new approaches that are presented in this thesis (see Section 5.1.2 and Chapters 9 and 6) are either based on the UE analysis or inspired by it, we are going to introduce it in detail. The UE analysis had been developed to detect coordinated spiking activity based on a statistical comparison of an observed number of JSEs and an expected number of JSEs. The method can be applied to short windows of the data ranging down to a window length of a few tens of milliseconds. Thus, the UE analysis addresses the property of fast changes in the data. (see previous Section 4.2.1)

4. Tools to Detect Coordinated Neuronal Activity

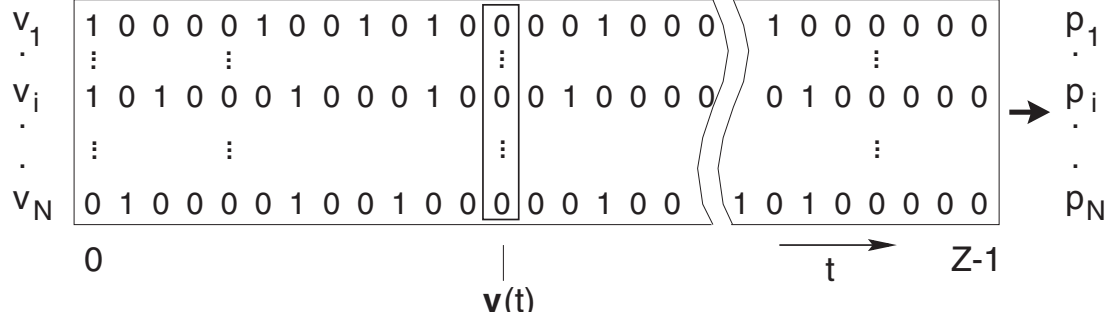


Figure 4.1.: **Binned representation of N spike trains.** Each spike train $v_i(t)$ consists of Z bins, each element of one or zero. Based on the number of spikes represented by ones across Z bins, the probability of spiking p_i can be estimated based on the assumption that each spike train is a stationary Bernoulli process.

The basic idea behind the UE analysis is to infer if a certain number of observed JSEs can be explained by chance by employing a hypothesis test. To this end the observed number of JSEs is compared with an expected number. The estimation of the expected number is based on the assumption that simultaneously recorded spike trains are originated by independent neurons. To this end, activity of each neuron is represented as a spike train with exclusive bins (Fig. 4.1 and Eq. 2.1). The bin width Δ_b reflects the expected maximal jitter of individual spikes (usually $\Delta_b \sim 5 \text{ ms}$). Since the analytical description used by the UE method requires a binary process, clipping is used, that modifies the data and reduces any number of events per bin that is larger than one to one. Thus, the mathematical description of a spike train that was given in Eq. 2.1 has to be modified to match the description of a binary process and to be applicable for the analytical description used by the UE method.

$$v_i(t) = \begin{cases} 1, & \text{one or more spikes in } [t, t + \Delta_b) \\ 0, & \text{else} \end{cases} \quad (4.10)$$

$$\text{with } t = 0, 1\Delta_b, 2\Delta_b, \dots, (Z-1)\Delta_b \quad (4.11)$$

For N simultaneously recorded neurons given a binary representation holds (Fig. 4.1):

$$v(t) = \begin{bmatrix} v_1(t) \\ v_2(t) \\ \vdots \\ v_N(t) \end{bmatrix} \quad \text{with } v_i \in \{0, 1\} \quad \text{and } i = 1, \dots, N \quad (4.12)$$

For each time step there are $k = 2^N$ potentially existing configurations v^k . In case of the null hypothesis (H_0) that assumes that each neuron (i) is independent of any other neuron (j), the probability of any configuration v^k is given by:

$$H_0 : P_k = \prod_{i=1}^N P(v_i^k), \text{ with } P(v_i^k) = \begin{cases} P(v_i = 1), & \text{if } v_i^k = 1 \\ 1 - P(v_i = 1), & \text{if } v_i^k = 0 \end{cases} \quad (4.13)$$

Consequently, the probability of occurrences for an arbitrary configuration k in M trials and Z timesteps is given by:

$$n_k^{pred} = P_k \cdot Z \cdot M \quad (4.14)$$

In the following steps the probability of a certain configuration under H_0 , given a set of simultaneously recorded spike trains, will be derived. To this end the UE analysis assumes that the observed spike trains are originated by a Bernoulli process (Cox and Isham, 1980). A Bernoulli process is a binary and memory free process. Consequently, in case the spike frequency is assumed to be constant throughout a segment containing Z bins, the probability of spiking of an individual neuron i can be derived by:

$$p_i = \frac{c_i}{Z} \quad \text{with } c_i \text{ total number of spikes from neuron } i \text{ in } Z \text{ timesteps} \quad (4.15)$$

and

$$H_0 : P_k = \prod_{i=1}^N P(v_i^k), \text{ with } P(v_i^k) = \begin{cases} p_i(v_i = 1), & \text{if } v_i^k = 1 \\ 1 - p_i(v_i = 1), & \text{if } v_i^k = 0 \end{cases} \quad (4.16)$$

The combination of H_0 , that assumes that all spike trains are originated by mutually independent neurons, and the assumption of a stationary Bernoulli process leads to mutually independent bins of the set of N spike trains across time and across neurons (Fig. 4.1). Thus, the set of spike trains can be described by an N -dimensional Bernoulli trial (Feller, 1968; Grün et al., 2002a). Therefore, the probability ψ that one configuration v^k occurs exactly n_k times in Z timesteps is given by:

$$\psi(n_1, n_2, \dots, n_m; P_1, P_2, \dots, P_m; Z) = \frac{Z!}{\prod_{k=1}^m n_k!} \cdot \prod_{k=1}^m P_k^{n_k} \quad (4.17)$$

Equation 4.17 describes, given the following normalization, a multinomial distribution.

$$\sum_{k=1}^m P_k = 1 \quad (4.18)$$

$$\sum_{k=1}^m n_k = Z \quad (4.19)$$

Consequently, the probability distribution for one configuration is given by Eq. 4.20, that is equivalent to a binomial distribution parameterized by P_k .

$$\psi(n_k; P_k; Z) = \frac{Z!}{n_k! \cdot (Z - n_k)!} \cdot P_k^{n_k} \cdot (1 - P_k)^{Z - n_k}, \quad k = 1, \dots, m \quad (4.20)$$

4. Tools to Detect Coordinated Neuronal Activity

The binomial distribution can be approximated by a Poisson distribution if the number of timesteps Z is large and the product $P_k \cdot Z$ at the same time stays small :

$$\psi(n_k; P_k; Z) = \frac{(P_k \cdot Z)^{n_k}}{n_k!} \cdot \exp(-P_k \cdot Z), \quad k = 1, \dots, m \quad (4.21)$$

Given that the expected number $n_k^{pred} = P_k \cdot Z$ is Poisson distributed and parameterized by n_k^{pred} (Feller, 1968; Grün, 1996), then the Eq. 4.21 can be used to compute the probability that an observed frequency of JSEs, referred to as empirical frequency, can be explained by H_0 that assumes independent Bernoulli processes.

4.2.3. Limitations of the Unitary Event Method

Based on the previous description of the UE analysis we are going to discuss limitations of the approach next.

Robustness of Findings Revealed by the UE Method

In case it is unlikely that the null hypothesis (H_0) can explain an empirical frequency of one JS pattern observed in an experiment, H_0 has to be rejected according to the idea of hypothesis testing. Since H_0 is formulated on the basis of two assumptions, first, stationary Bernoulli processes and second, mutually independent processes, any violation of one of the two might lead to a rejection of H_0 . Hence, any violation of H_0 due to a violation of the assumed stationary Bernoulli process, is equivalent to a false positive rejection of the hypothesis that neurons are mutually independent. Thus, a rejection of H_0 does not necessarily indicate evidence for the assembly hypothesis. This emphasizes that the UE method is crucially dependent on the capability to demonstrate that a violation of H_0 is based on fine-temporal cross-structure. This implies that various other properties that violate the assumed stationary Bernoulli process, like changing spike rates in time and across trial, rate covariation of sets of neurons, or auto-structure in the spike trains (bursting, regular spiking) have to be excluded.

To demonstrate that fine-temporal structure caused the rejection of H_0 , two alternative approaches can be used. The first checks, if any of the discussed properties, like changing spike rates that might lead to a false positive rejection, are expressed in the data. This approach is very problematic, because, first, one can never prove that the analyzed feature is not expressed, since a negative result might just reflect an insensitive method, and second, the variety of features of neuronal activity that is not described by a stationary Bernoulli process is just too large to be tested for everything. The alternative approach (see Chapter 9) inverts the argumentation. Instead of demonstrating the unimportance of other features in the data it rather demonstrates the importance of fine-temporal structure for the rejection of H_0 . To this end the UE analysis is applied two times to the same data, but in the second run, the fine-temporal

cross-structure is destroyed by jittering (dithering of individual spikes). The appealing idea behind this is that a difference in the results between before and after jittering can only be explained by fine-temporal cross-structure.

Another limitation of the UE method is that rare events can lead to false positive rejections of H_0 . The limitation arises by the fact that the UE method does not consider trial by trial variability of the empirical frequency of JSEs (Roy et al., 2000). Since the estimation of the statistical significance is based on the total frequency, rare events that are occurring only once or a few times, cannot be taken apart from events that are occurring reliably across trials. (For further discussion see Sections 6.1.6, 6.2.5 and 6.3.5).

Thus, since in recorded data, first, spike rates are likely changing in time and across trials, second, spike trains are likely not described by a Bernoulli process, and third, JSEs might be pretty rare, conclusions which one might draw based on the rejection of H_0 based on the UE method, are pretty fragile and have only a limited value for argumentations.

Exclusive Binning

The intention of the coarsening of the sampling frequency is to allow for detection of JS activity that is jittered, but less apart in time than Δ_b (see also Section 4.2.2 and Eq. 4.10). However, exclusive binning does not consider the exact difference between the spike times of the neurons, rather it reduces the precision of the sampling to Δ_b . This poses the problem that spikes, which are less than Δ_b apart and should be considered as a JSE, can be detected as non-coincident if they happen to fall across separate bins ((Grün et al., 2002a), Fig. 6.2G). For that reason, exclusive binning isn't likely to detect all JSEs. This problem is further aggravated by increases in both, the complexity of the JSEs and the bin size Δ_b (Grün et al., 2002a).

One can overcome the problem of exclusive binning with the multiple-shift method (Grün et al., 1999), which also utilizes binning, but the bin-size is rather short ($b \sim 1$ ms, see Section 2.1.5). To detect JSEs that are scattered across larger intervals than b , the multiple-shift method shifts the entire spike train of each neuron in relation to one another, step by step. The maximal shift corresponds to the maximal allowed jitter. However, this method can be applied only to datasets with a small number of neurons because the number of shift combinations and the resulting computational complexity increases exponentially with the number of neurons included in the analysis. Thus, in order to analyze JSEs across a larger number of neurons, a computationally more efficient method is needed. (For further discussion see Section 6.1.1).

4.2.4. New Approaches to Identify Coordinated Spiking Activity

Motivated by these limitations we developed two new approaches for analyzing spiking activity from multiple neurons that we are going to present in this thesis.

Since currently available methods fail to extract the full correlation structure from massive parallel recordings (activity of up to hundreds of neurons recorded simultaneously), both new methods do not try to infer the correlation structure underlying JSEs (Martignon et al., 2000; Nakahara and Amari, 2002; Okatan et al., 2005), but utilize the underlying idea of the UE analysis and test, whether a certain type of JS pattern occurs more or less often than expected by chance. Thus, both methods do not try to derive a model of the correlation structure, but rather they test whether coordinated firing is random or likely to be induced by, as well as correlated to information processing or neuronal states.

The first new approach, the *Bootstrap Unitary Event method* ('BUE method'), that we are going to present in Chapter 5 (for results see Chapter 9) is an extension of the UE method (see Section 4.2.2) that addresses the limitations of the standard UE analysis that arise from rare events and the variability of the empirical and expected frequency of a certain JSE across trials. Thus this new method is able to deal with variability across trials, changes of the data on short time scales and rareness of events. (see Section 4.2: features one, two, and four).

The second new approach, called *NeuroXidence* (see Chapter 6, for results see Chapter 7), is based on a concept that is quite different to existing approaches including the UE method. It utilizes a new concept of statistical evaluation of the difference between the expected and empirical frequency and uses surrogate data to formulate H_0 . Thus the combination of these two concepts allows NeuroXidence to address all four features (see Section 4.2) like variability across trials, changes of the data on short time scales, rareness of events, history dependencies and auto-structures of recorded spike trains, that have confounded the data analysis so far.

Part II.

New Tools for Analyzing Coordinated Neuronal Activity

5. New Resampling Tools for Uni-, Bi-, and Multivariate Data

Motivated by the fact that neither for coherence and phase-locking in case of the analysis of the local field potential ('LFP') nor for the p -value of joint spike ('JS') patterns derived by the unitary event ('UE') analysis, a robust method for the estimation of a confidence interval is existing, we developed tools based on a combination of the n -jackknife and bootstrapping as well as a permutation test, that estimate confidence intervals and allow to compare the influence of two or n different conditions on the test statistics by means of a hypothesis test. The tools are in analogy with the two-sample Mann-Whitney U test (see Appendix D.3.3) or an ANOVA test (Kruskal-Wallis, see Appendix D.3.4).

5.1. Uni-Variate Data: Assessment of Stability and Reliability

Given typical electrophysiological recordings we were concerned about two frequent observations. First, properties or statistics that are used to describe the data might have large variabilities and second, the experimental conditions or the state of the investigated system might change during the experiment, since the experiment can last between a couple of minutes to a couple of hours, and thus, these changes can either be induced by manipulations from outside (e.g. changed level of anaesthesia, drugs etc.), or might be intrinsic, caused by changes in the investigated system. In an awake animal, like the awake monkeys from whom we have analyzed data (see Chapters 8 and 9), these intrinsic changes might occur due to different levels of alertness, tiredness, motivation, or other more basic changes, or just episodic changes without any obvious reason. Thus, one has to consider these changes when data is analyzed. Parameters like the performance or the eye-movements of the experimental animal are usually used by the experimentalist to control for stability. And indeed, performance is a very good criteria on stability for experiments with awake animals, since it can be assumed that the probability to identify characteristics in the data induced by processes necessary to perform the task are not changing over time if the performance stays constant throughout the whole experiment. Still, in addition to these attempts that try to stabilize and to test the stability of the experimental conditions, one would like to use the observed properties themselves to judge how stable they were expressed during the experiment. For standard statistics like the mean value these issues can easily be handled by, first, the standard deviation that allows to estimate

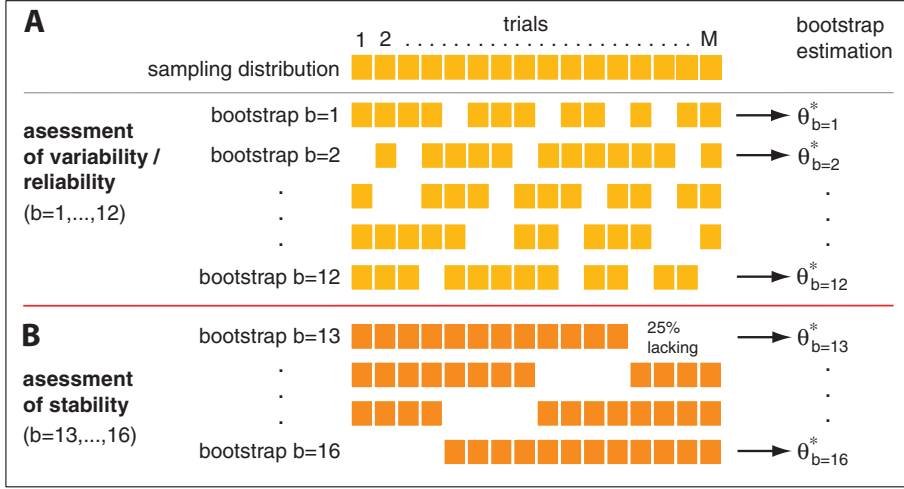


Figure 5.1.: **Bootstrap and n -jackknife of one univariate sample.** (A) Bootstrapping of M trials leads to bootstrap samples each containing 75% of the trials. Bootstrapping is performed based on random sampling with replacement (putting back). (B) The n -jackknife is resampling of the same M trials. Each n -jackknife sample drops one quarter of the data so that each element exists only once. Based on each bootstrap sample the estimation of Θ by Θ_b^* is derived. In total $b = 1 \dots B$ with $B=16$ bootstrap samples are used.

variability, and second, the assessment of trends to estimate stability based on linear models. But, in case of statistics for that these estimators haven't been developed so far (see for further discussion Sections 4.1.4, 4.2.4, 4.2.3), like the p -value of the UE method or the coherence and phase-locking value ('PLV'), one has to develop new tools. Thus, we next present a new tool that estimates the reliability and variability including trends in univariate data.

The basic idea is to employ two different resampling techniques. First, we use bootstrapping that randomly resamples the data set \mathbf{x} to estimate variability and reliability, and second, the n -jackknife that considers trends in order to estimate stability. To this end, the method utilizes ξ_{boot} bootstrap samples each consisting of random 75% of the samples drawn with replacement from \mathbf{x} and 4 n -jackknife samples that are lacking systematically either the first, second, third or fourth quarter of the data. The 4 n -jackknife samples are used to estimate the stability of the experiment (Note, the elements of \mathbf{x} have to be ordered accordingly to the experimental time). Each of the bootstrap and n -jackknife samples is denoted by \mathbf{x}_b^* with $b = 1 \dots B = \xi_{boot} + 4$. Based on each bootstrap sample \mathbf{x}_b^* the test statistic of interest $\Theta^*(\mathbf{x}_b^*)$ is estimated (see Appendix E). Thus $\Theta(\mathbf{x}_b^*)$ represents the coherence or PLV in case of LFP data, or the p -value in case of the UE method.

5.1.1. Stability and Reliability of Phase-Locking and Coherence

For the purpose of simplification we denote both, the PLV and the coherence, with the same estimator Θ . Consequently the empirical distribution H^* of the bootstrap estimation Θ^* of Θ is derived by all B

bootstrap samples. Thus, $H^*(\Theta^*)$ can be utilized to derive the confidence interval of Θ and the PLV or coherence. To this end the standard deviation $\hat{\sigma}_{\Theta}^*$ based on B bootstrap samples of Θ^* is computed (see also Appendix E.4):

$$\hat{\sigma}_{\Theta}^* = \sqrt{\sum_{b=1}^B \frac{[\hat{\Theta}^*(\mathbf{x}_b^*) - \langle \hat{\Theta}^*(\mathbf{x}_b^*) \rangle]^2}{B}} \quad (5.1)$$

In comparison to previous approaches (Multi-taper analysis and Chronux Software toolbox for Matlab, see also (Mittra and Pesaran, 1999)) the advantage of this new approach is that it uses a combination of bootstrapping and n -jackknife instead of the delete-one-jackknife. This increases on the one hand the robustness, since both, bootstrapping and n -jackknife, but not the delete-one-jackknife, had been shown to be universal methods that derive correct estimates, and, on the other hand, it allows to incorporate the stability analysis that drops systematically one quarter of the data.

5.1.2. Stability and Reliability of Significant Joint Spike Patterns

In contrast to the stability analysis of the phase-locking or coherence, the new method, called *Bootstrap Unitary Event method* ('BUE method'), that we are going to present next and that is based on the UE analysis (see Section 4.2.2, and (Grün et al., 2002a; Grün et al., 2002b)), estimates the reliability and stability of JS patterns but does not rely on confidence intervals.

In case of the UE method we were concerned about three properties of spiking data and JSEs. The first two, that are variability and stability, were discussed in Section 5.1. Thus, since the estimation of the p -value is based on the empirical and total number of JSEs across M trials, both estimators n_{emp} and n_{exp} have to be included into an estimation of stability and variability. The third property we were concerned about are low spike rates and the rareness of JSEs, since it had been shown, that low rates lead to an increase in the false error rate of the UE method ((Roy et al., 2000), and see Chapter 6). The reason for this increase is that the UE analysis compares the total number of empirical JSEs with the total expected number (both across M trials), instead of comparing the difference of the distributions of the single trial estimations of both. Thus, in case of a very few empirical JSEs and a low spike rate, the Poisson distribution, which models the probability of the empirical number under H_0 , fails to model an appropriate H_0 distribution ((Roy et al., 2000), for further discussion see Sections 6.1.5, 6.1.6 and 6.3.1). The stability, rareness and variability estimation of the joint- p -value of the UE method utilizes the concept that is introduced in Section 5.1 (Fig. 5.1). The UE method is applied to each of B bootstrap samples \mathbf{x}_b^* and computes the test statistics $\Theta^*(\mathbf{x}_b^*)$ corresponding to the bootstrap joint- p -value:

$$\Theta^*(\mathbf{x}_b^*) = jp_b^* \quad \text{with } jp_b^* : \text{bootstrap estimation of the joint-}p\text{-value} \quad (5.2)$$

The joint- p -value (jp_b^*) describes how likely it is to observe an empirical number of JSEs by chance.

5. New Resampling Tools for Uni-, Bi-, and Multivariate Data

A joint- p -value of 0.5 indicates that JSEs are at chance level, while smaller values indicate an excess of JSEs and larger values a deficiency of JSEs. As a significance criterium for the new bootstrap version of the UE method it is required that all jp_b^* are smaller or equal than the test level (α).

$$\text{if } jp_b^* < \alpha \text{ for } b = 1 \dots B \text{ than significant excess of JSEs} \quad (5.3)$$

and

$$\text{if } jp_b^* > 1 - \alpha \text{ for } b = 1 \dots B \text{ than significant deficiency of JSEs} \quad (5.4)$$

To simplify later definitions we define $joint - p - value_{boot}$ by:

$$joint - p - value_{boot} = \max(jp_1^*, jp_2^*, \dots, jp_B^*) \quad (5.5)$$

Thus the condition $joint - p - value_{boot} < \alpha$ is equivalent to Equation 5.3.

Since we are using multiple estimates and require that all have to be significant, the effective test level (α_{eff}) is different from the applied test level (α) leading to a different expected number of false positives. Next, we derive a lower and an upper bound of the effective test level based on $joint - p - value_{boot}$. In case we assume that each of the bootstrap samples is mutually independent, the probability of a false positive event is given by:

$$\alpha_{eff}^{indep} = \alpha^B \quad (5.6)$$

But, if we assume the other extreme that each bootstrap sample is the same, the effective level is of course identical to the applied level (α), since all jp_b^* are the same. Thus, the effective test level is bounded by α_{eff}^{indep} and α .

$$\alpha_{eff}^{indep} < \alpha_{eff} < \alpha \quad (5.7)$$

Because only 75% of the samples are used in each bootstrap sample, the effective level α_{eff} must be smaller than the applied level, which makes the hypothesis test more conservative. The new bootstrap UE method approach addresses the rareness of spiking and JSEs, and requires for a significant pattern, that the p -value for each bootstrap sample is significant. That implies, first, that the number of JSEs has to be significantly different from chance level in any of the tested bootstrap samples, and second, that the total frequency of observed JSEs has to be at least two. Latter is the case, since the n -jackknife procedure splits the data in pieces of 3/4 of the original length. Thus there must be at least two JSEs distributed across two disjunct quarters to have at least one in any 3/4 long bootstrap samples. Since the p -value has to be significant for each bootstrap sample, the stability constraints on the empirical and expected number of JSEs get more strict for p -values close to the significance threshold. That implies that the difference between the observed and the empirical frequency has to be quite stable. Note, in case of the UE analysis the difference is not evaluated pairwise between the empirical and expected number of the same trial rather than between the total number of both estimators across all M trials. This implies that differences between the total empirical and total expected number of JSEs can still be originated

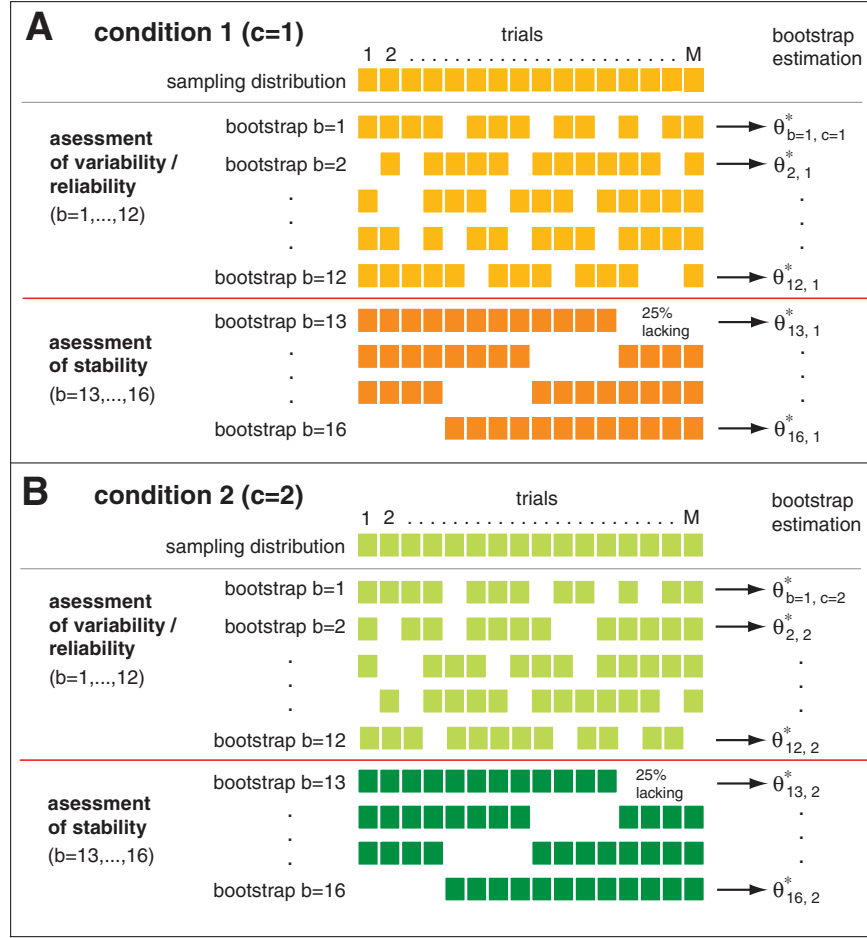


Figure 5.2.: **Bootstrap and n -jackknife of bivariate data.** Bootstrapping and n -jackknife of two-samples (**A**: sample 1, **B**: sample 2) from two conditions ($c = 1, 2$). Each sample is resampled by 16 bootstrap samples. Each contains 75% of the trials. Bootstrapping is performed based on random sampling with replacement (putting back), while the n -jackknife drops one quarter of the data (each element exists only once). Based on each bootstrap sample, $\Theta_{b,c}^*$ is derived. In total, two times $B=16$ bootstrap samples are used.

in different trials. In conclusion, the new approach checks for stability and variability and prevents that significant events may be originated by differences between the expected and the empirical number of JSEs that are clustered in a few trials or in a certain period of the experiment.

5.2. Bi-Variate Data: Assessment of Effects

Next we describe a new tool that allows to estimate if two different conditions modulate the strength of phase-locking, the coherence or the joint- p -value of the UE analysis. With the same motivation as for the one-sample design (see Section 5.1) we incorrupted the estimation of stability and variability. The basic concept of the method is to estimate first, the variability of the test statistics Θ by a combination of

5. New Resampling Tools for Uni-, Bi-, and Multivariate Data

bootstrapping and n -jackknife, and second, to derive a H_0 distribution of the expected differences of the test statistics based on a permutation test. The latter has the disadvantage that it is computationally demanding but the strong advantage that it is the statistically safest and most robust and precise way to construct a distribution of H_0 .

To support this arguments we cite the introductory paragraph of the permutation test chapter from 'An Introduction to the Bootstrap' from Efron and Tibshirani, page 202 (1993, Chapman and Hall/CRC):

'Permutation tests are a computer-intensive statistical technique that predates computers. The idea was introduced by R.A. Fisher in the 1939's, more as a theoretical argument supporting Student's t-test than as a useful statistical method in its own right. Modern computational power makes permutation tests practical to use on a routine basis. The basic idea is attractively simple and free of mathematical assumptions.[...] The greatest virtue of permutation testing is its accuracy.'

Due to the accuracy of permutation tests the empirical percentage of false positives is, (if H_0 is true) for any given test statistics, almost exactly the test level itself (see page 210 in 'An Introduction to the Bootstrap' from Efron and Tibshirani). As illustrated in Figure 5.2 for each condition ($c = 1, 2$) indexed with c , B bootstrap samples are used to derive estimates of the test statistic $\Theta_{b,c}^*$. The first $B-4$ are used to estimate variability and reliability based on bootstrapping while the last 4 are used to estimate stability. To compare both sets of $\Theta_{b,c}^*$ for condition $c=1$ and $c=2$ we tested if the median of $\Theta_{b=1\dots B,1}^*$ is significantly different from the median of the second condition $\Theta_{b=1\dots B,2}^*$.

The latter allows a robust estimation that is not influenced by outliers or bimodal-distributions, as it would be the case if we tested on average differences based on the mean values of the samples one and two (see Appendix D and D.4). Thus the difference is by analogy with the Mann-Whitney U test (see Appendix D.3.3) derived by the difference in the sum of the ranks of $\Theta_{b=1\dots B,1}^*$ and $\Theta_{b=1\dots B,2}^*$. To this end, first, ranks of the elements of a joint set of both individual sets are derived, and second, the sum of the ranks k^* of the samples of condition one $k^*(c = 1)$ and condition two $k^*(c = 2)$ are derived. The latter are used to derive the difference Δ^*k^* of $k^*(c = 1)$ and $k^*(c = 2)$ that is used to describe the difference of the coherence, phase-locking or p -value between condition one and two.

5.3. Permutation Test for Bivariate Data

To judge if the difference Δ^*k^* is significant or not, a permutation test is utilized to derive the probability distribution $H^*(\Delta^*k^{*H_0})$ of Δ^*k^* under H_0 (H_0 formulates that the difference between condition one and two is by chance). To this end the trials from condition one and two are randomly permuted between the conditions. This implies that even an identical copy or a mirror copy, for which each trial

is exchanged between both conditions, might occur. The probability of these extreme cases, of course, drops with increasing number of trials. The latter illustrates the tight link between the number of samples and the degree of freedom that are used in the t-statistics. To estimate the distribution of Δk under H_0 denoted by $\hat{H}^*(\Delta^*k^{*H_0})$ we use $B_{perm} = 160$ permutations (see Monte Carlo estimation in Appendix E.2 and Section 5.3.1). Based on the concept of additional false positive events (those false positives that are caused by the Monte Carlo approximation of the permutation test) it had been shown (Pipa, 2001) that the $B_{perm} = 160$ gives enough confidence in the empirical distribution $\hat{H}^*(\Delta^*k^{*H_0})$ to be used for a hypothesis test based on a testlevel of one percent.

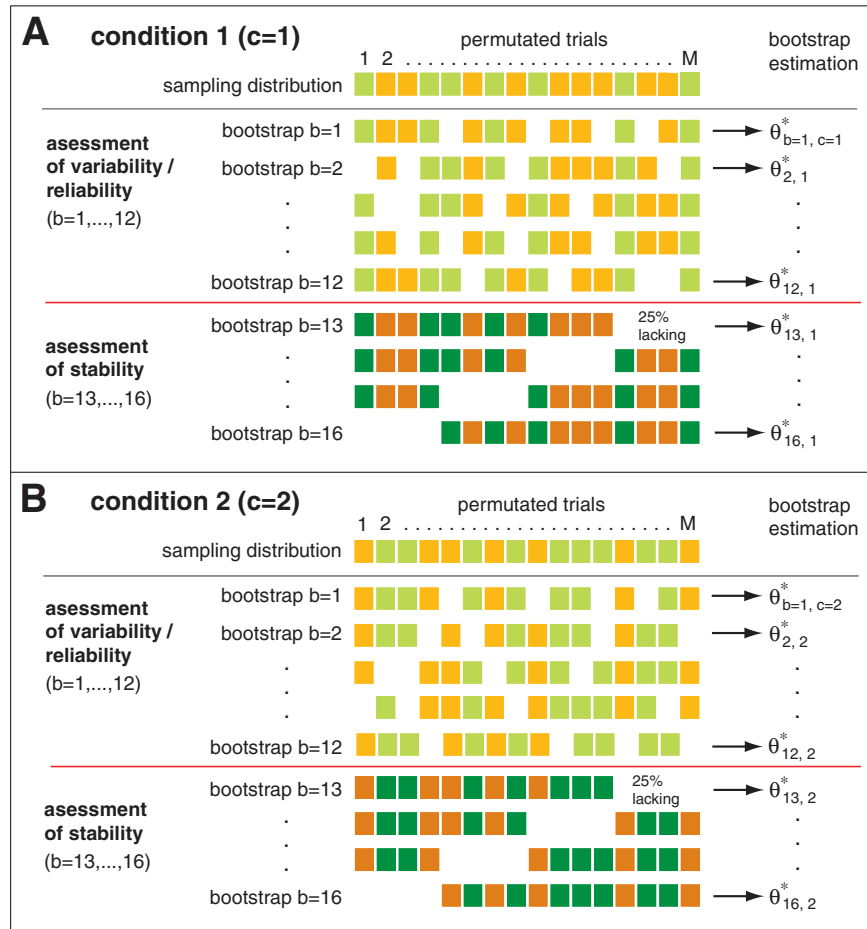


Figure 5.3.: **Permutation, bootstrap and n -jackknife of bivariate data.** Example that illustrates the concept of a permutation test. Bootstrapping and n -jackknife of two-samples (**A**: sample 1, **B**: sample 2) from two conditions ($c = 1, 2$). Before each sample is resampled by 16 bootstrap samples, trials from condition one and two are randomly permuted across both conditions. Each bootstrap sample contains 75% of the trials. Bootstrapping is performed based on random sampling with replacement (putting back), while the n -jackknife drops one quarter of the data (each element exists only once). Based on each bootstrap sample $\Theta_{b,c}^*$ is derived. In total two times $B=16$ bootstrap samples are used.

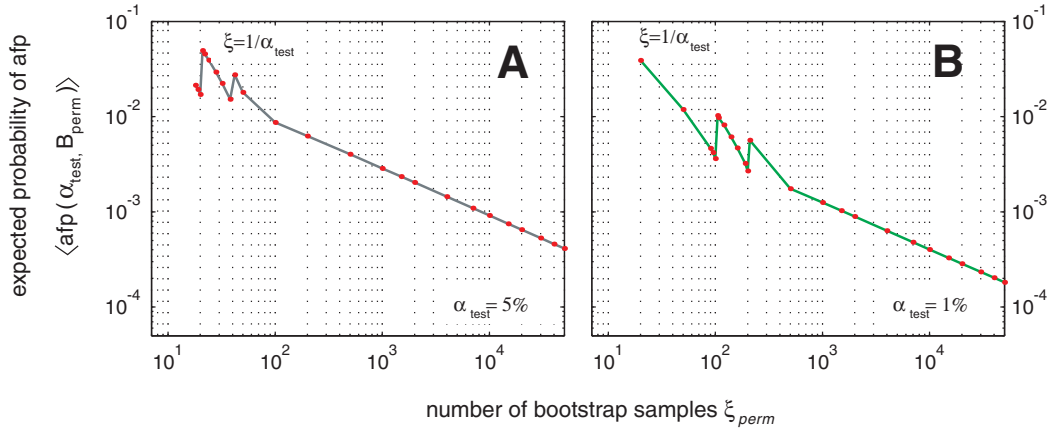


Figure 5.4.: (A,B) **Expected probability of additional false positive events.** Expected probability of additional false positive events ('afp') in dependence of the number of permutations in case of a 5% significance criteria in **A** and a 1% criteria in **B**

5.3.1. Number of Permutation Derived by Additional False Positives

The expected resulting percentage of false positive events $\langle \alpha \rangle$ is bounded by the sum of the test level α_{test} and the expected percentage of additional false positive events $\langle afp(\alpha_{test}, B_{perm}) \rangle$, (Pipa, 2001).

$$\langle \alpha \rangle \leq \langle afp(\alpha_{test}, B_{perm}) \rangle + \alpha_{test} \quad (5.8)$$

The expected percentage of the additional false positives $\langle afp(\alpha_{test}, B_{perm}) \rangle$ can be derived analytically (Pipa, 2001) and amounts to 0.4% in case of $B_{perm} = 160$ and $\alpha_{test} = 1\%$. Thus, in case of 160 permutations and a test level of 1% the effective test level is smaller than 1.4% (Fig. 5.4).

5.4. Multi-Variate Data: Assessment of Effects

Next we describe a new tool that allows to estimate, by analogy with the ANOVA, if any from N different conditions modulate the strength of phase-locking, the coherence or the joint- p -value of the UE analysis. We again used the same design as for the one- and two-sample design (see Section 5.1) that estimates stability and variability.

As for the two-sample design the basic concept of the method is first, to estimate the variability of the statistics by a combination of bootstrapping and n -jackknife, and second, to derive a H_0 distribution of the expected differences of the statistics based on a permutation test (H_0 assumes that no difference exists). As illustrated in Figure 5.5 for each condition ($c = 1, 2, \dots, N$) B bootstrap samples are used to derive estimates of $\Theta_{b,c}^*$. The first $b - 4$ are used to estimate variability and reliability based on bootstrapping, the last 4 are used to estimate stability. We again use a median test on the analogy of the Kruskal-Wallis test (see also Appendix D.3.4) and test, if at least one median of $\Theta_{c,b=1\dots B,1}^*$ of one condition is

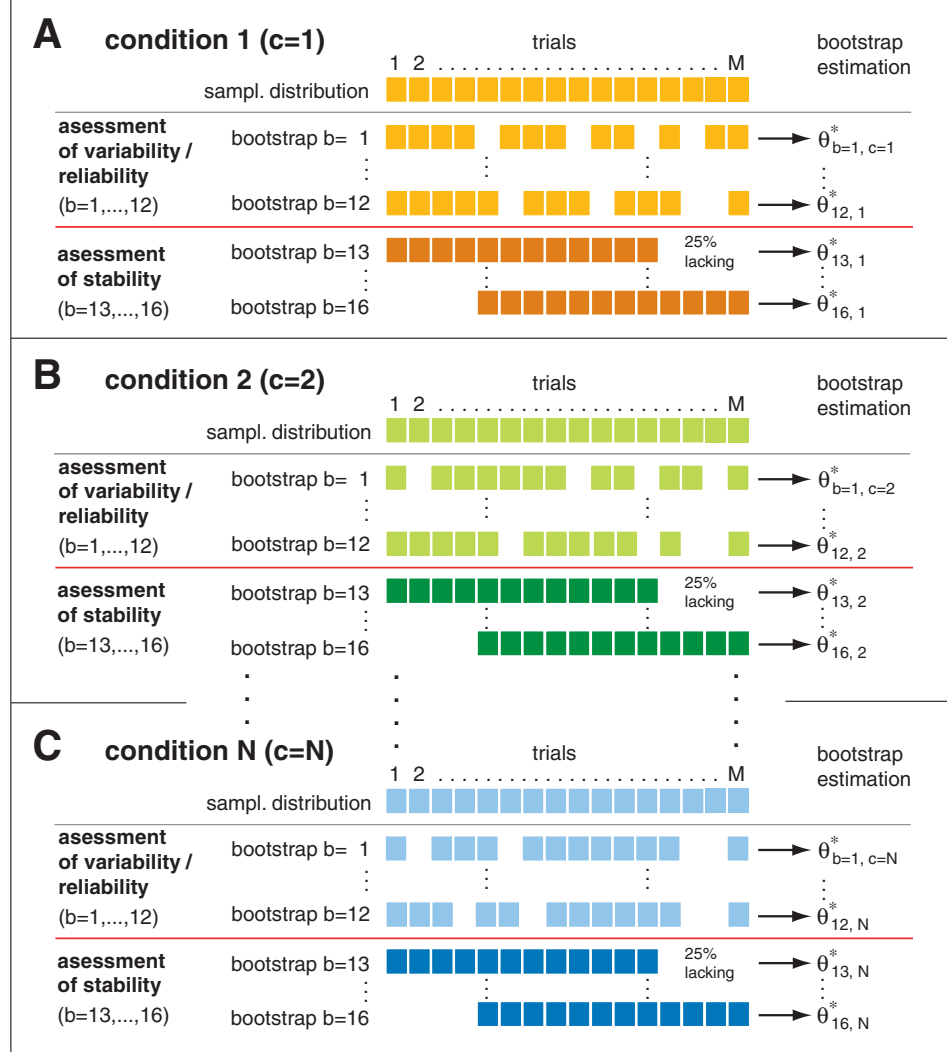


Figure 5.5.: **Bootstrap and n -jackknife of multivariate data.** Bootstrapping and n -jackknife of N samples (**A**: sample 1, **B**: sample 2, **C**: sample N) from N different experimental conditions ($C = 1, 2, \dots, N$). Each sample is resampled by 16 bootstrap samples. Each contains 75% of the trials. Bootstrapping is performed based on random sampling with replacement (putting back), while the n -jackknife drops one quarter of the data (each element exists only once). Based on each bootstrap sample $\theta_{b,c}^*$ is derived. In total N times $B=16$ bootstrap samples are used.

5. New Resampling Tools for Uni-, Bi-, and Multivariate Data

significantly different from any median of the other conditions. As discussed for the two-sample test, the median test allows a robust estimation that is not influenced by outliers or bimodal-distributions, as it would be the case if we tested the average differences based on the mean value (see Appendix D and D.4). Since we use a multivariate test we must use a different concept of how to compare the medians across the different conditions than the concept used by the Mann-Whitney test in the bivariate case (see Section 5.2). For this multivariate test the difference is derived by analogy with the Kruskal-Wallis test (see also Appendix D.3.4). Thus, we use the variance of the ranks and compare the group variance of the corresponding ranks of each of the N conditions with the variances of the ranks of all N conditions times 16 bootstrap samples. To this end, first all elements $\Theta_{b,c}^*$ in the joint set of all N conditions are ranked, and second, the sum and variance of the ranks per condition is computed. Like in the standard ANOVA (see Appendix D.2.3) the difference between the groups is expressed by \hat{F} :

$$\hat{F} = \hat{\sigma}_{pop}^2 / \hat{\sigma}_{error}^2 \quad (5.9)$$

To derive the significance of the value \hat{F} , the empirical probability distribution of \hat{F} under H_0 denoted by \hat{F}^{*H_0} is derived (H_0 : the medians from all conditions are the same) by a permutation test as it has been used for the two-sample design in Section 5.3.

6. NeuroXidence

In the last chapter we introduced the bootstrap unitary event ('BUE') method that was based on the unitary event ('UE') method, but was designed to account for variability in the date and rareness of joint spike events ('JSEs'). But the problem of auto-structure in the spike trains wasn't addressed by the BUE method. This motivated us to develop a different strategy than the UE method to detect excess or deficiency of JSEs. The new method that will be presented in detail in the following chapter is called *NeuroXidence*.

NeuroXidence addresses the problems of variability and rareness, utilizes a non-parametric approach, and simplifies the process of detecting JSEs (see Section 4.2.1). NeuroXidence analyzes coordinated neuronal firing utilizing a novel and computationally-efficient algorithm for identifying and counting JSEs that are based on joint spike ('JS') activity that is not precisely synchronous.

6.1. Method

As pointed out already in Section 4.2, *NeuroXidence* tests whether a certain type of JS pattern occurs more or less often than expected by chance (Aertsen and Gerstein, 1985; Abeles and Gerstein, 1988; Baker and Lemon, 2000; Gerstein and Perkel, 1969; Gerstein and Perkel, 1972; Gray et al., 1989; Grün et al., 1999; Grün et al., 2002a; Grün et al., 2002b; Ikegaya et al., 2004; Kass et al., 2005; König, 1994; Samonds and Bonds, 2004) and whether such an excess or deficiency of JSEs is correlated with information processing or neuronal states (Castelo-Branco et al., 2000; Gray et al., 1989; Riehle et al., 1997) and does not try to infer the correlation structure underlying JSEs (Martignon et al., 2000; Nakahara and Amari, 2002; Okatan et al., 2005). Thus, NeuroXidence does not try to derive a model of the correlation structure, but rather it tests whether coordinated firing is random or likely to be induced by, as well as correlated to, information processing or neuronal states. The central feature of this method is the establishment of the timescale τ_c , which defines the required temporal precision of fine-temporal cross-structure that is typically assumed to be in the range between 1 and 10 ms (Aertsen et al., 2001; Bi and Poo, 1998; Feng and Brown, 1998; Grün et al., 1999; Hopfield and Brody, 2000; Hopfield and Brody, 2001; Markram et al., 1997; Lestienne, 1995; Sjostrom et al., 2001). NeuroXidence utilizes τ_c as the maximum temporal difference between the first and the last spike of a JSE (Fig. 6.1D). By using τ_c in this way, NeuroXidence avoids the problems of previous approaches that utilized τ_c for exclusive binning ((Grün et al., 2002a;

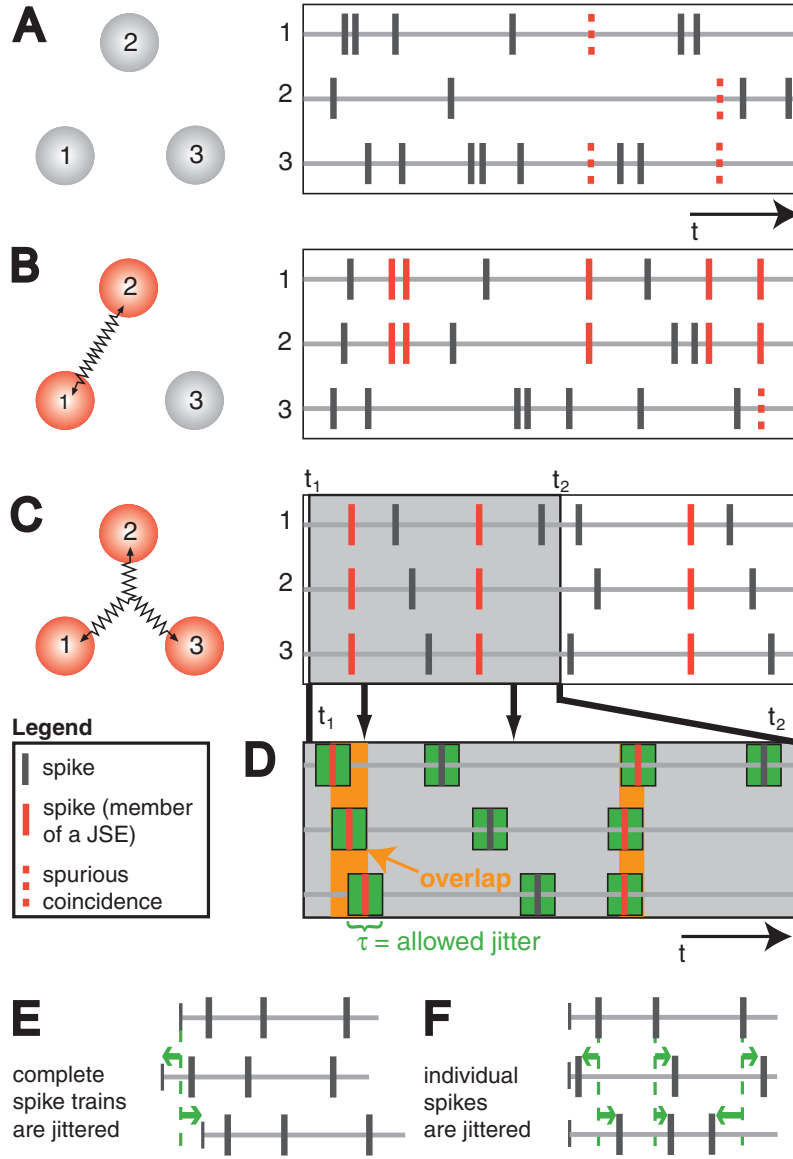


Figure 6.1.: **Synchronization of spikes by neuronal coupling.** (A-C) Show three units, indicated by 1-3, representing either single-neurons or three neuronal populations that are mutually independent. The right sub-panel shows the resulting spike trains. In the case that the units represent single-neurons, each spike train contains single-unit activity (SUA), while each spike train is composed of multi-unit activity (MUA) if the units represent small populations of neurons. (A) Since the units are not coupled, they are mutually independent and exhibit spike trains that do not share synchronized spikes beyond chance level. In (B) and (C) units are coupled. Spike trains produced by coupled units share synchronized spikes (indicated by red colour), referred to as joint spike events (JSEs). (B) A second order correlation between units 1 and 2 in a JSE of complexity 3. (C) A third order correlation between units 1, 2, and 3. The definition of a JSE is sketched in (D). Spikes that belong to a JSE share overlapping regions of the maximally allowed jitter. Regions of allowed jitter represent either a maximally-allowed offset, a maximal random jitter, or both together. (E,F) Two different ways of jittering spikes. (E) In NeuroXidence, each complete spike train is jittered on a timescale given by τ_r . This jittering destroys coordinated spiking on all shorter timescales, while preserving the auto-structure of each spike train on all timescales. (F) Each spike of each spike train is jittered individually. Therefore, the auto-structure, as well as the cross-structure, is modified.

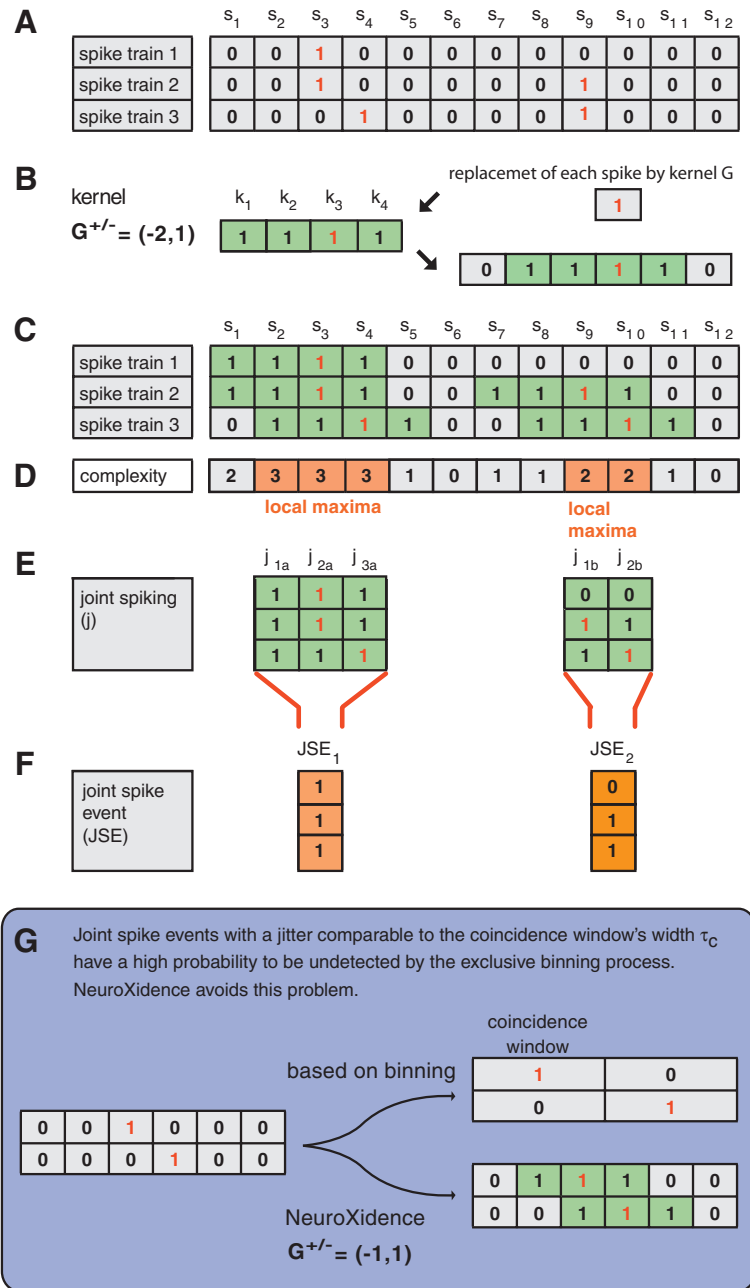


Figure 6.2.: **Detection of joint spike events ('JSEs')**. (A) A spike train is defined as a binary process (1 = spike, 0 = no spike) with a bin width b small enough to resolve an individual spike. (B) Each spike is replaced by the kernel $G^{+/-}$ that defines the maximal allowed jitter. The red ones ('1') indicate the position of real spikes, while the black ones ('1') indicate the maximally-allowed jitter given by the kernel $G^{+/-}$. (C) The results of the replacements are JSEs, which are defined by overlapping regions of allowed jitter (green) across different spike trains. (D) To determine the complexity of JSEs, the number of spikes (ones ('1'), black and red) per time bin is computed. The local maxima of the number of synchronous spikes determines the complexity of joint spiking. (E,F) Joint spikes belonging to one period, defined by a local maximum of the complexity, are mapped to one JSE. (G) The difference between the exclusive binning used by the unitary event method (UE method) and the centered expansion of spikes with a kernel $G^{+/-}$ used by NeuroXidence. Since the UE method is based on exclusive binning, it is likely that JSEs are not detected.

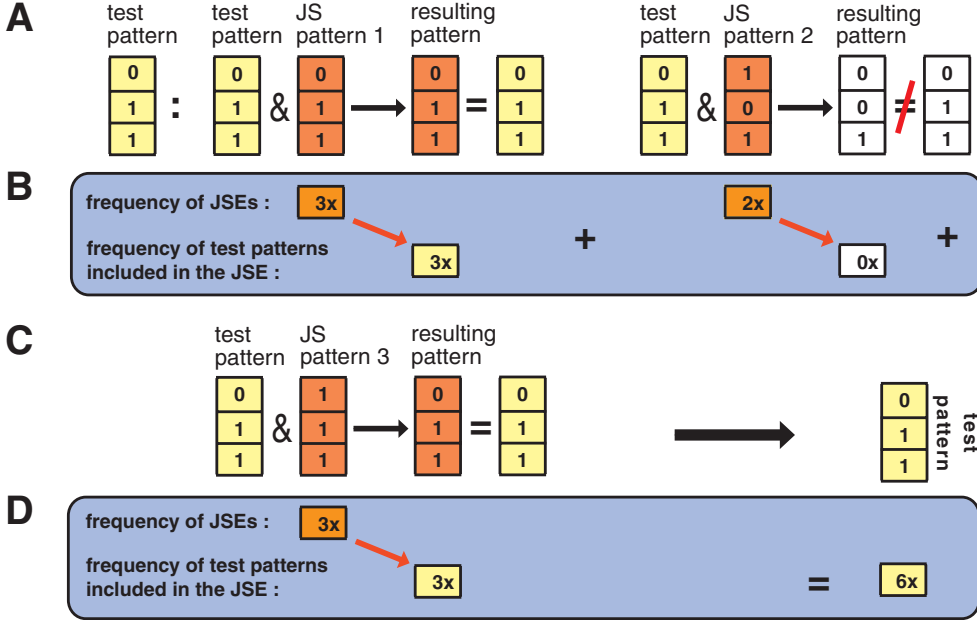


Figure 6.3.: **Total frequency of a test pattern.** Deriving the frequency of occurrence of a test pattern requires accounting for the frequency of occurrence of sub-patterns included in other joint spike events ('JSEs'). (**A,C**) To determine if a JSE is a sub-pattern of the test pattern, an AND operation between the test pattern and each JSE is applied. (**B,D**) The frequency of occurrence of the resulting pattern is considered only if it is identical to the test pattern. The total frequency of occurrence of the test pattern is given by the sum of all frequencies of all qualifying resulting patterns

Grün et al., 2002b; Pipa and Grün, 2003), Fig. 6.2G).

6.1.1. Detection of Joint Spike Events

NeuroXidence is more computationally efficient than previously proposed algorithms (Grün et al., 1999) because it takes advantage of the sparseness of spiking activity to reduce the computational complexity of finding JSEs. The algorithm also returns an exact solution and allows for variations in the timing of each individual spike that are less than the defined amount of jitter τ_c . This is accomplished by the use of binning to represent the spiking data as a binary process, where the bin-size b is considerably shorter than τ_c (e.g., $b = 1$ ms). Thus, this method preserves the precision of the spike times in the original recording when detecting JSEs (Fig. 6.1D).

There are three key steps in the detection of JSEs: the preprocessing of spike trains (see Appendix A.1: Preprocessing, and Fig. A.1), the detection of jittered JSEs, and the identification of JS patterns and their complexity. The preprocessing step ensures that the minimal inter-spike interval of any unit is larger than τ_c , corresponding to the kernel $G^{+/-}$ defined in units of the bin-size b . This preprocessing prevents temporally-overlapping JSEs, while preserving each individual JSE and its exact frequency of

occurrence. In the second step, NeuroXidence detects jittered JSEs by extending the temporal extent of each individual spike to the number of bins defined by $G^{+/-}$. For that purpose, each spike and its neighborhood is replaced by a kernel which contains $G^{+/-}$ bins of ones (Fig. 6.2A-C). The third step in JSE detection consists of identifying JS patterns and determining their complexity. This is accomplished by evaluating the local complexity (i.e. the number of synchronous spikes) per time bin (Fig. 6.2D). Each local maxima of the local complexity corresponds to one instance of joint spiking (Fig. 6.2E). Thus, zones of maximal local complexity are used to derive the corresponding JSE within an identified JS pattern (Fig. 6.2F).

The determination of the total frequency of occurrences of one particular JS pattern in the dataset involves the sum of two JS pattern frequencies. One is the frequency of JSEs that are identical to the JS pattern of interest. The other is the frequency of all supra-patterns in which the JS pattern is included as a sub-pattern. NeuroXidence identifies these supra-patterns using an AND operation between the JS pattern of interest and any other detected JS pattern. Only if the resulting JS pattern is equal to the JS pattern of interest, is the frequency of the supra-pattern considered (Fig. 6.3A-D).

6.1.2. Definition of the Hypothesis Test

To test whether JSEs of one type of JS pattern occur more or less often than by chance, a statistical hypothesis test is used. The *null hypothesis* (H_0) is formulated based on the assumption that neurons are firing independently from each other (Fig. 6.1). The *alternative hypothesis* (H_1) assumes that there are more (or less) coordinated firing events than expected by chance since spiking events are interdependent across neurons. Consequently, the observed frequency of JSEs has to be compared to the expected frequency of JSEs in the case where the neurons are assumed to be independent. Therefore, the core of the statistical hypothesis test is the estimation of the expected frequency of JSEs.

This estimation can be based on either a parametric or a non-parametric approach. Parametric approaches hypothesize an explicit model of the data, while non-parametric methods do not pose such assumptions. An example of a parametric method is the analysis of UE method, (Grün et al., 2002a; Grün et al., 2002b), which hypothesizes that spike trains can be described by a stationary Bernoulli process (Cox and Isham, 1980), which is a binary and discrete process with constant probability of spike generation across time as well as with no history dependency. Most parametric methods have a higher test power (the probability to reject H_0 if H_1 is true), if the assumptions underlying the method are satisfied. However, if these assumptions are violated, parametric methods might increase the rate of false rejections of H_0 , thereby increasing the number of type I errors in statistical inference. In the case of the UE method, several factors have been discussed that might induce false rejections of H_0 . These include low rates (Roy et al., 2000), fast modulations of the rates (Grün et al., 2002b), auto-structure in the spike trains (Baker and Lemon, 2000), and trial-by-trial variability in fir-

ing rates and in latencies of responses (Azouz and Gray, 1999; Baker and Gerstein, 2001; Brody, 1999; Grün et al., 2003). Therefore, a non-parametric method might provide considerable advantages in detecting JSEs, as such a method would not assume any model of the spiking activity.

6.1.3. Generation of Surrogate Data

NeuroXidence allows for two different non-parametric approaches to estimate the chance frequency of JSEs (H_0). The first estimates the chance frequency of JSEs (H_0) based on surrogate data, a technique that has been widely used (Schreiber and Schmitz, 2000; Baker and Gerstein, 2001). The ideal surrogate dataset is derived from the original dataset, such that it maintains all of the original properties of the data but does not contain any coordinated firing. Therefore, the art of developing a non-parametric method is to produce surrogate data that differs from the original dataset in only one property. This property defines the alternative hypothesis H_1 . A surrogate dataset should preserve the auto-structure of each individual spike train, trial-by-trial variability, rate modulations, and all history dependencies (see also Section 4.2.3).

NeuroXidence generates surrogate data by discriminating two disparate time scales. The first timescale τ_c corresponds to fine-temporal cross-structures that is corresponding to JSEs. This range is typically between 1 and 10 ms (Aertsen et al., 2001; Bi and Poo, 1998; Grün et al., 1999; Hopfield and Brody, 2000; Hopfield and Brody, 2001; Markram et al., 1997; Sjöström et al., 2001). The second timescale τ_r is defined to be distinguishably slower than τ_c , so a factor η is introduced, which is the ratio between the two timescales:

$$\tau_r = \eta \cdot \tau_c \quad (6.1)$$

In the present analysis η is chosen to be in the range of 2-5, meaning that τ_r is about 2 to 5 times slower than τ_c . As we will show, this difference in timescales is sufficient to test reliably for the presence of fine-temporal cross-structure defined by JSEs, even in data in which the auto-structure of the spike trains changes on a timescale that is comparable to τ_c .

The surrogate dataset is derived from the original spike trains by jittering spikes by an amount given by the slow timescale τ_r (Fig. 6.1E, F). In contrast to random jittering, also referred to as '*dithering*' of each individual spike (Fig. 6.1F), which destroys the fine-temporal auto-structure, NeuroXidence jitters each spike of each individual spike train by the same amount (Fig. 6.1E). Thus, the jittering is equivalent to random shifts of an entire spike train, which destroys fine-temporal cross-structure on timescales faster than τ_r , but preserves auto-structures on all timescales. Hence, random shifting of entire spike trains is an optimal strategy to produce surrogate data, to test on excess and deficiency of JSEs.

6.1.4. Shuffled Spike Trains Used as Surrogate Data

The second non-parametric approach of NeuroXidence formulates H_0 based on shuffling of spike trains. Shuffling of spike trains is defined by a random and across neurons independent permutation of trials in the dataset ((Brody, 1999; Gray et al., 1989; König, 1994), Fig. G.5). The permutation is intended to destroy patterns that are induced by the underlying coupling of neurons. The basic assumption underlying the idea of spike train shuffling is that the spike train generating process is composed by two constituents.

The first constituent is assumed to be a process that might be changing in time, but has to stay the same across trials (ergodic). The dynamics of this first process are not constraint in any sense.

The second constituent of the spike train generating process is assumed to represent activity that is based on coupling of neurons and therefore is the origin of JS patterns that might be observed in the dataset. The assumed nature of this process is that it is inducing fine-temporal cross-structure on the time-scale τ_c , but the precision of the reoccurrence in time and across trials is much more sloppy than the time scale τ_c of the fine-temporal cross-structure itself. In other words, it is assumed that patterns that might be constituted by synchronous spikes on a millisecond timescale are reoccurring across trials less precisely than the precision of the pattern itself. Thus the precision, with which patterns are reoccurring across trials, might be of the order of tens of milliseconds or more.

This split between the two processes is motivated by the idea that processes in the recorded neuronal activity might be evoked by fast changes from outside the system like visual stimulation (referred to as '*evoked activity*'). In this case it is assumed that the process evoked by the stimulation is precisely time locked to the stimulus, while the dynamics of internal activity, that is generated by the system itself, are not tightly locked to the stimulus, but still might include fine-temporal cross-structure. The latter is referred to as '*induced activity*'. If this scenario is true, shuffling of spike trains destroys the JS patterns that are based on induced activity. Since the induced activity is the one that is assumed to reflect the internal activity of the system as well as the information processing, shuffling of spike trains destroys only the activity that is assumed to be representative for the assembly hypothesis. Therefore, the difference between the shuffled and the original, simultaneously recorded data is the difference that is assumed to be based on coupling of neurons.

Still, the basic idea of comparing shuffled and original spike trains requires that spike train generating processes stay the same across trials. In physical terms that means that the processes must be ergodic. Thus, only the respective stochastic realization of the same processes is expected to be different from trial to trial. Nevertheless, since data recorded in real neuronal systems often show a high degree of variability and strong systematic changes, the main assumption of the shuffling procedure is not fulfilled. Thus,

in case processes are changing over time, shuffling of trials recombines different processes that might lead to more changes in the data than the destruction of induced fine-temporal structure (Brody, 1999; Grün et al., 2003). To correct for these additional changes tools were developed that allow to use trial shuffling even in case of latency covariation, or changing rates across trials (Ventura et al., 2005; Grün et al., 2003). Still, the correction requires the estimation of parameters of the processes like the spike rates. In this case the correction itself requires the assumption a model of the spike train generating process to estimate its rate.

As a consequence, the idea to separate induced and evoked activity is appealing but technically and statistically very difficult to realize. As for any method that requires crucial assumptions about the data that cannot be completely fulfilled, trial shuffling and results revealed by it remains a question at issue. Since the alternative described in Section 6.1.3 based on the separation of timescales does not require the assumption of ergodicity it is superior and more robust. Still, the alternative that is based on the separation of the two timescales τ_c and τ_r cannot differentiate between evoked and induced activity if the evoked activity is locked to the stimulus with a precision that is compatible with the precision of JSEs. Thus trial shuffling can be seen as a complementary procedure for analyzing spiking activity on JS patterns if it is required to distinguish between evoked and induced activity on a timescale of a few milliseconds. Since such a tight millisecond locking is only the case for a very few classes of stimuli (e.g. flashed visual stimulation) we are using in this thesis exclusively the idea that is based on the separation of the timescales τ_c and τ_r .

6.1.5. Implementation of the Statistical Test

To assess a deficiency or an excess in the frequency of JSEs within a particular JS pattern, we compare the frequency of their occurrence in the original dataset to their frequency in the surrogate dataset or the shuffled dataset. Thus, for a particular JS pattern k and trial t , the difference Δf_t^k between the two frequencies is computed. The variability of Δf_t^k is then assessed on a trial-by-trial basis,

$$\Delta f_t^k = f_t^k(org) - f_t^k(sur) \quad (6.2)$$

where $f_t^k(org) \in N_0$, for trial $t = 1 \dots T$ and JS pattern $k = 1 \dots K$.

To improve the sampling of the chance frequency of JSEs occurring in surrogate data (H_0), NeuroXidence allows the use of multiple (S) surrogates. In case of method one (see Section 6.1.3) that employs random shift each surrogate sample is based on S random and independent shifts to derive the average chance frequency of JSEs per trial. Consequently, in case of the shuffling procedure (Section 6.1.4) each of the surrogates is based on S independent permutations of trials. Thus, for $S > 1$, the difference Δf_t^k becomes the average difference between the original data and the S surrogates $\bar{\Delta} f_t^k$,

$$\bar{\Delta} f_t^k = \frac{1}{S} \sum_{s=1}^S \Delta f_{t,s}^k = f_t^k(org) - \frac{1}{S} \sum_{s=1}^S f_{t,s}^k(sur) \quad (6.3)$$

$$\text{with } \bar{\Delta}f_t^k \in Q \wedge f_t^k(\text{org}) \in N_0 \quad (6.4)$$

A value of $\bar{\Delta}f_t^k$ larger than zero indicates an excess of JSEs within one particular JS pattern k in the original dataset. Similarly, a negative value of $\bar{\Delta}f_t^k$ indicates a deficiency of JSEs in the original dataset. The differences $\bar{\Delta}f_t^k$ for one particular JS pattern k computed for all available trials form the set $\bar{\Delta}F^k$.

$$\bar{\Delta}F^k = \{\bar{\Delta}f_1^k, \bar{\Delta}f_2^k, \dots, \bar{\Delta}f_T^k\} \quad (6.5)$$

NeuroXidence uses the set $\bar{\Delta}F^k$ to test whether the excess or deficiency of JSEs is significant and consistent across trials. To this end, NeuroXidence uses either the parametric t -test, to investigate whether the mean value of $\bar{\Delta}F^k$ is significantly different from zero, or the non-parametric Wilcoxon-rank test, to determine whether the median of $\bar{\Delta}F^k$ is significantly different from zero. Note, since the statistical evaluation is based on the trial-based difference $\bar{\Delta}f_t^k$, both tests are paired tests that keep the trial identity of the frequency of JSE in the original and surrogate data. The statistical tests are always used as single-sided tests but are applied to both sides of the distributions to test for both excesses and deficiencies of JSEs. Note, if a t -test is used to derive the statistical significance of $\bar{\Delta}F^k$, the distribution of the mean value of $\bar{\Delta}F^k$ is assumed to be Gaussian. In this instance, NeuroXidence uses a combination approach that utilizes non-parametric surrogate data to derive $\bar{\Delta}F^k$, while the statistical evaluation of $\bar{\Delta}F^k$ is parametric. In contrast, if the statistical inference is based on a Wilcoxon-test, then all stages of the analysis are non-parametric.

6.1.6. Rare Events versus Significant

Estimation of the statistical significance utilized by NeuroXidence is derived from the trial-based evaluation of $\bar{\Delta}f_t^k \neq 0$. Therefore, NeuroXidence requires that a significant increase or decrease of the frequency of JSEs be based on a consistent trend across trials. This allows for the discrimination of significant events from those that are rare and spurious. This is one of the main advantages of NeuroXidence, as it contrasts with other methods that consider only the total number of JSEs across trials and do not take trial-by-trial variability into consideration (Abeles and Gerstein, 1988; Grün et al., 1999; Grün et al., 2002a; Grün et al., 2002b; Ikegaya et al., 2004).

6.1.7. Parametric t-Test versus Non-Parametric Wilcoxon-Rank Test

In practical applications, the Wilcoxon-rank test (Wilcoxon-test) is more likely to be appropriate for hypothesis testing than Student's t -test. The t -test assumes that the mean value of $\bar{\Delta}F^k$ is normally distributed, while the Wilcoxon-test does not make any assumptions about the distribution of the test statistics. This difference between the two tests is important since JSEs are expected to be rare when spike rates are low and JSE complexities are high (i.e. large number of neurons are recorded in parallel). With such rare events, the probability distribution of the total number of JSEs in the original dataset is

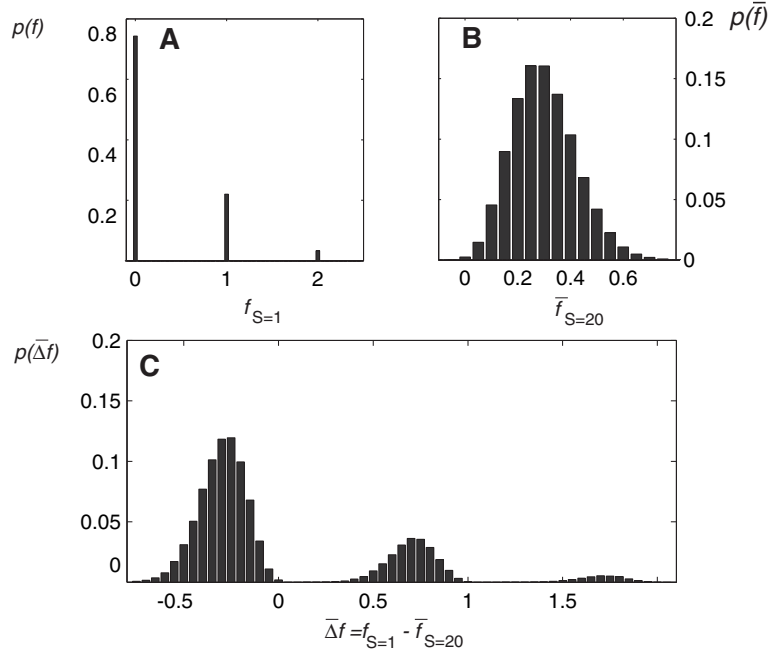


Figure 6.4.: **Impact of multiple surrogates S .** Probability distributions of f and 20 surrogates of \bar{f} . (A) Estimated distribution (10^6 samples) of the frequency of f of JSEs per trial, given a Poisson distribution with an expected value of 0.3 JSEs per trial. (B) Estimated distribution \bar{f} based on averaging $S = 20$ surrogates of the distribution in (A). (C) Probability distribution of the difference $\bar{\Delta}f = f - \bar{f}$.

expected to be better approximated by a Poisson than by a normal distribution. Thus, the distribution of JSEs is likely to violate the assumptions of the t -test. The same holds for the probability distribution of the total number of JSEs that occur in the surrogate dataset. This distribution is expected to be less skewed than the one in the original dataset (Fig. 6.4), since the number of averaged samples is larger by the factor S (surrogate samples per trial). As a consequence of the skewness and non-normality of both probability distributions, the distribution of the average difference, corresponding to the mean value of $\bar{\Delta}F^k$, is also not expected to be normal.

Thus, if JSEs are rare, the assumption of the t -test that the mean value of $\bar{\Delta}F^k$ is normally distributed, is likely to be violated. Although the t -test is to some degree robust against skewed distributions (Boneau, 1960), the Wilcoxon-test may, in many cases, be a much better choice, especially when spike rates are low and complexities of JSEs are greater than two.

6.1.8. Parameters of the NeuroXidence Significance Estimation

We next investigate the properties of the parameters η and S utilized by NeuroXidence to estimate the statistical significance of JSEs (see Appendix A.2 for recommended standard parameters). To this end, we discuss how both affect the false positive rate and the test power of the method.

A: The Scaling Factor η

As previously mentioned, NeuroXidence differentiates between two timescales at which neural processes operate. The first, τ_c , defines the timescale of JSEs. The second timescale τ_r is slower and is relevant for the other processes considered by the significance evaluation. The ratio between the two timescales is given by $\eta = \tau_r/\tau_c$ (see Eq. 6.1).

The selection of η is based on a trade-off between the information maintained in the surrogate data and the information destroyed by jittering the spike trains. On the one hand, τ_r should be as fast as possible to allow for rapid changes in the cross-structure, such as co-variations in firing rates (Baker and Gerstein, 2001; Brody, 1999). On the other hand, a large value of τ_r (i.e. large η) ensures that jittering destroys most of the coordinated firing in the surrogate data, which is necessary for a high test power (probability to reject H_0 if H_1 is true). The probability that jittering does not destroy JSEs, decreases exponentially with the complexity of JSEs, since the number of existing different jitter configurations is exponentially growing with the complexity of the JS pattern. Therefore, large η improves mainly the test power for JS patterns whose complexities do not exceed two or three (Fig.6.6). In conclusion, η should be chosen to be in the range of 2 to 5, which provides a good balance between the test power and the dynamics of the cross-structure.

B: Number of Surrogates S

Using $S > 1$ random surrogates per trial improves the sampling of the distribution of the frequency $f_t^k(sur)$ of JSEs for one particular JS pattern (k). This improves the estimate of the expected number of JSEs under H_0 , which increases the reliability of the estimated p -value. In addition, multiple surrogates map the frequency of JSEs from an integer value, when $S = 1$, onto the average frequency, which is an element of Q , when $S > 1$. The distribution of the average difference between the frequency of JSEs in the surrogate and original data can be derived by a convolution of the respective distributions of the frequency:

$$p_{\bar{\Delta}f}(f) = \int df' p_{\bar{f}_{sur}}(f') \cdot p_{f_{org}}(f + f') \quad (6.6)$$

In the case of rare events, for which the probability of zero JSEs per trial is the largest, the median of $\bar{\Delta}f_t^k$ is expected to be smaller than zero (Fig. 6.4). This has important implications for the Wilcoxon-test, as its purpose is to test whether the median of $\bar{\Delta}f_t^k$ is significantly different from zero. It follows from Eq. 6.6 that $S > 1$ biases the median of $\bar{\Delta}f_t^k$ towards negative values. This bias makes NeuroXidence more conservative when testing for an excess of JSEs and more liberal when testing for a deficiency of JSEs (i.e. increased number of false positive events). Therefore, it is recommended to use larger values of S only for testing for an excess of JSEs, while one needs to be cautious when testing for a deficiency of JSEs, in which case $S = 1$ is a more proper choice to prevent type 1 errors. Since the t -test is based

on the mean value as an estimator of the expected value, which is not affected by increasing S , the t -test can be applied in the case of $S > 1$ for the estimation of excess as well as of lacking JSEs.

6.1.9. False Positive Rates

False positive events are defined as false rejections of the null hypothesis (H_0). If a statistical hypothesis test is applied to a recording of large sets of neurons, one has to differentiate between three different scenarios leading to different definitions of false error rates:

Scenario One: The test is applied to only one JS pattern of interest. In this case, the rate of false positives is defined by the percentage of p -values lower than the test level (e.g., $p < 0.05$). This percentage of false positives is referred to as the *individual-false-positive-rate*, since it is defined for a single JS pattern. The inclusion of more than one JS pattern requires a multiple-comparison correction (e.g. Bonferroni) if at least one p -value is smaller than test level.

Scenario Two: The test is applied to a set of JS patterns that is fixed in advance (e.g. all JS patterns potentially existing in the dataset). In this case, one can evaluate the grand average percentage of the false positives across all tested JS patterns. This makes it likely that the test is performed even on JS patterns that do not exist in the data. This average false error rate is referred to as the *population-false-positive-rate*, which corresponds to the percentage of JS patterns that are falsely evaluated as significant. Thus, scenario two is an extension of scenario one with multiple patterns. To ensure that a conservative population-false-positive-rate is smaller than the test level, the average of the individual-false-positive-rates estimated across all tested JS patterns has to be smaller than the test level.

Scenario Three: The test is applied only to those JS patterns that occurred in the dataset at least once. In this case, the grand average false error rate is referred to as the *actual-false-positive-rate*. In contrast to scenario two, a test that ensures a conservative actual-false-positive-rate has to ensure that the average individual-false-positive-rates of all occurring JS patterns are below the test level. This implies that the statistical significance must be independent of the detection of the JS patterns, which is achieved by evaluating the significance on the basis of multiple trials. The consideration of the trial-by-trial variability of the frequencies of occurrences prevents patterns, which are by chance occurring in a few or even only one trial, from being considered to be significant. Thus, significant patterns are required to occur reliably across trials.

In the context of simultaneous recordings of large sets of neurons, only scenarios two and three are of practical usage, since for scenario one the JS pattern has to be picked randomly, or a multiple comparison correction has to be performed (e.g Bonferroni). The latter corrects the test level for the number of JS patterns that were tested (K). Since the number of JS patterns can be very large ($K > 10^6$ for more than 20 simultaneously-recorded neurons), the Bonferroni correction would lead to a corrected test level that is extremely small (a test level of 1% would correspond to a corrected test level smaller than 10^{-8} for 20 simultaneously-recorded neurons). In this case, a statistical significance test of any of the JS patterns has to have a precision and reliability comparable to the corrected test level to allow for a reliable evaluation of its statistical significance. Given the size of typical datasets (up to a few hundreds of seconds of recording time) and the total frequency of JSEs per JS pattern (up to a few hundreds or thousands of occurrences in the whole recording), it is obvious that the required precision and reliability of the estimation of the statistical significance, in the case of large sets of neurons, can not be fulfilled. A multi-comparison correction does not have to be used in scenarios two and three because the statistical hypothesis test is used to derive the percentage of JS patterns for which the H_0 was rejected. In fact, since the expected population-false-positive-rate in scenario two and the actual-false-positive-rate in scenario three are both required to be below the test level, any excess beyond the test level indicates support for H_1 .

To decide which scenario is relevant, one has to differentiate between statistical hypotheses tests that allow for a conservative actual-false-positive-rate and those that permit a conservative population-false-positive-rate. The former test is the most relevant to NeuroXidence, as it allows for a computationally-effective analysis of a dataset. In contrast, the latter test only ensures that a conservative population-false-error-rate is applied to each JS pattern of interest, independently of its existence.

There are many reasons why statistical hypothesis tests might fail to meet the actual-false-positive-rate criteria. Some are not independent of the existence of JS patterns. They are based either on the frequency of JSEs in only one trial or on the total frequency across trials (Abeles and Gerstein, 1988; Grün et al., 1999; Grün et al., 2002a; Grün et al., 2002b; Ikegaya et al., 2004). These tests also do not consider trial-by-trial variability and might evaluate a small number of frequencies or even only one JSE as a significant excess, if the expected chance frequency of JSEs under H_0 is small enough.

6.2. Results

6.2.1. Joint Spike Pattern Detection

To prove that the NeuroXidence JS pattern detection algorithm identifies all JSEs and derives their exact frequency of occurrence, toy-data were analyzed. The toy-dataset contained different JSEs with given frequencies of occurrence. The dataset contained JSEs that were isolated as well as overlapping. The complexity of the induced JS patterns varied between 2 and 10, based on a set of neurons varying between

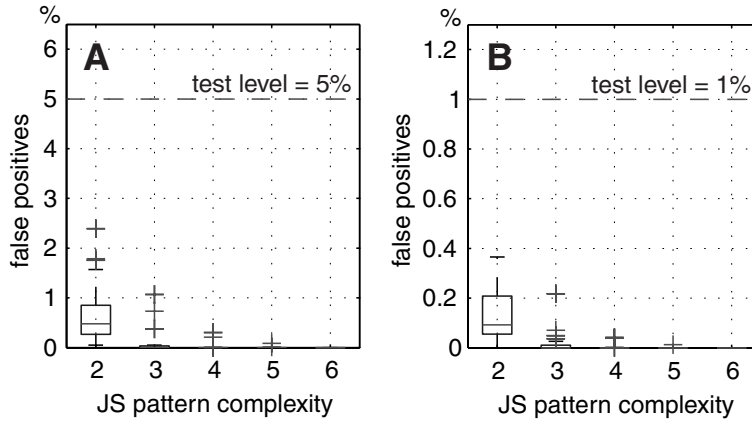


Figure 6.5.: **Percentage of false positives.** Percentile plots of the percentage of false positives estimated by NeuroXidence used to detect JS patterns of complexities 2 to 6. Independent and homogenous Poisson processes were used to generate toy-data that were analyzed by NeuroXidence to estimate the false positives rate based on 100 independent realizations of each toy-data model. Each individual percentile plot represents 15 parameter-sets describing 15 toy-data models, which differed, compared to the standard data model (50 trials, spike rate 15 ap/s, 20 surrogates, and $\eta = 3$), by variation of the mean spike rate ($r = 7, 10, 30, 60, 90$ ap/s), the number of trials ($T = 20, 100, 200$), the number of surrogates ($S = 1, 50, 250$), and $\eta = 2, 5, 7$. (A) The percentage of false positives for a test level of 5% and (B) for a test level of 1%.

2 and 8. NeuroXidence detected in each toy-dataset all induced JS patterns, as well as their sub-patterns, and derived the frequencies of occurrence correctly.

6.2.2. False Positives for Stationary Processes

Toy-data, generated by independent and stationary Poisson processes, were used to assess the individual-false-positive-rate of NeuroXidence for one individual JS pattern of interest. The individual-false-positive-rate for each individual JSE was derived from 100 independent hypothesis tests based on independent realizations of the same toy-data model.

The standard set of parameters used to generate the toy-data was defined by 50 trials (T), a mean spike rate of 15 ap/s (r), 20 surrogate samples (S), and η equal to 3. From this standard set, 15 different combinations of parameters were derived by varying the number of trials ($T = 20, 50, 100, 200$), the number of surrogates ($S = 1, 50, 250$), the mean spiking rate ($r = 7, 10, 30, 60, 90$ ap/s), and the scaling factor η ($\eta = 2, 5, 7, 9$). NeuroXidence was applied to each toy-dataset using a sliding window with duration of $l = 200, 400$ and 800 ms. In total, 48 different toy-data models were used to derive the individual-false-positive-rate for 5 JS patterns of complexities 2, 3, 4, 5 and 6. None of the individual-false-positive-rates for any of the parameter sets was above chance level, for either test level of 5% or 1% (Fig. 6.5), demonstrating that NeuroXidence is a conservative statistical hypothesis test. Furthermore, NeuroXidence becomes more conservative with increasing complexity.

6.2.3. Test Power for Stationary Processes

Test power for individual JS patterns was derived from correlated Poisson processes (Cox and Isham, 1980), generated by a single interaction process ('SIP') (Kuhn et al., 2003). Thus, correlated spike trains were characterized by a background rate, corresponding to the independent spiking of neurons, and by a JSE rate, defining the expected frequency of the JS pattern of interest beyond chance level. The test level used by NeuroXidence was 1%. To demonstrate that NeuroXidence is capable of detecting JSEs that are jittered less than the allowed jitter τ_c , we produced two sets of toy-data with $\tau_c = 5$ ms. The first toy-dataset contained exact JSEs, while the second was deduced from the first by random jittering of individual spikes by τ_c . The agreement of the test power for both datasets across different complexities, sliding-window lengths, and frequencies of JSEs demonstrates that NeuroXidence detects jittered JSEs and precise JSEs equally well (Fig. 6.6).

To illustrate the impact of η , the ratio of the timescales τ_c and τ_r ($\eta = \tau_r/\tau_c$) on test power, a third dataset was generated from the first by jittering individual spikes by $\tau_r = 3\tau_c = 15$ ms. This amount of jittering also corresponds to that used by NeuroXidence to destroy JSEs in the surrogate dataset. The result is an increased test power above chance level (1%) when JSE rates are high for complexity two and three JS patterns (Fig. 6.6A1-2), which indicates that some JSEs in the jittered third toy-dataset survived jittering. Test power could be increased further by increasing η , which lowers the probability of survival. For higher complexity JS patterns (Fig. 6.6A3-4, B3-4, C3-4), the large number of potential jitter configurations makes survival of JSEs after jittering unlikely. Therefore, $\eta = 3$ is high enough to destroy all JS patterns in the third dataset, leading to a test power close to or equal to zero. Sixteen toy-datasets were used to study how the test power of NeuroXidence is influenced by the length of the sliding window, the number of trials, the background spike rate, and the number of surrogates. Dataset 1 contained 50 trials, each 800 ms long. The background spike rate was 15 ap/s. The maximum allowed jitter was 5 ms and η was 3. Each of the remaining 15 sets was derived from the first set by variations of either the sliding window length ($l = 0.2, 0.4, 0.8$ s, Fig. 6.7A1-4), the number of trials ($T = 20, 50, 100, 200$, Fig. 6.7B1-4), the background spike rate ($r = 10, 15, 30, 60, 90$ ap/s, Fig. 6.7C1-4), or the number of surrogate datasets ($S = 1, 20, 50, 250$, Fig. 6.7D1-4).

Based on these 16 toy-datasets, the test power was derived for five JS patterns of complexity 2, 3, 4, 5 and 6. Increasing the length of the analysis window, as well as increasing the number of trials, led to increased test power. Given the same frequency of excess JSEs, the test power increases along with JSE complexities. This was expected since the probability of a chance occurrence of a JSE exponentially decreases with its complexity. The number of surrogates used for the estimation of the average frequency $\bar{\Delta f}_t^k$ changes the test power only very subtly in comparison to the changes observed when varying the window length, the number of trials, and the spike rate. In particular, the differences in test power for $S > 20$ were negligible, indicating that going beyond 20 surrogates does not substantially improve the

6. NeuroXidence

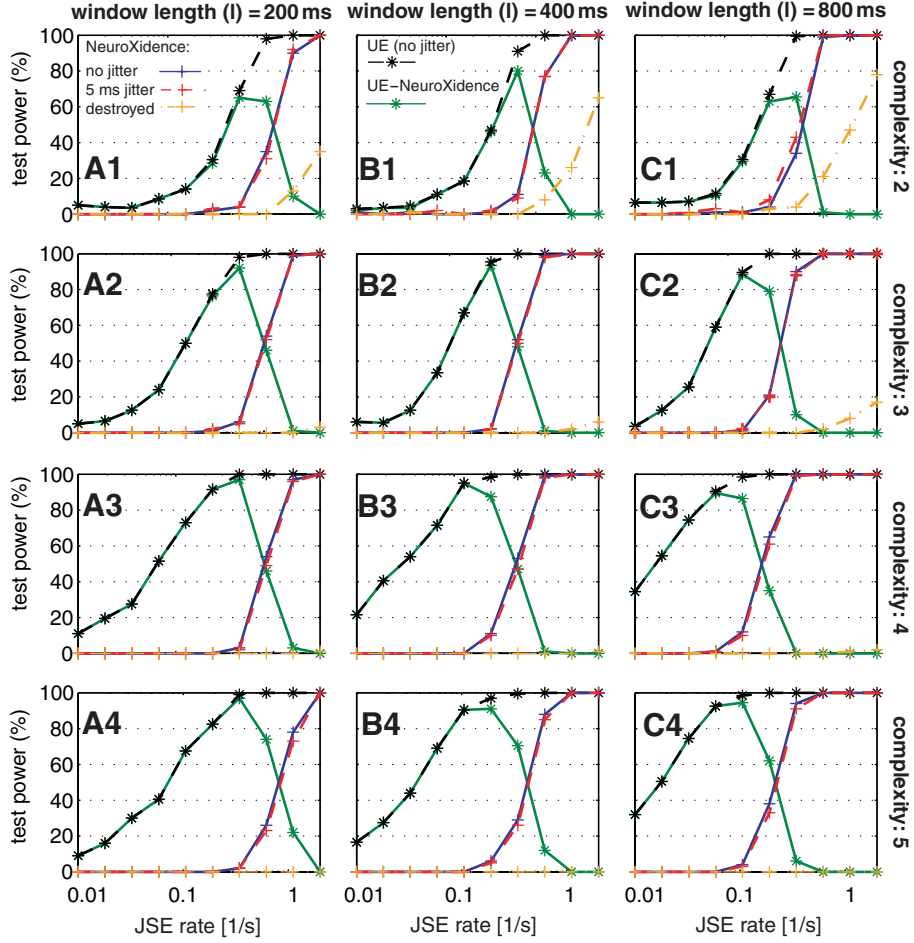


Figure 6.6.: Comparisons of test power in relation to the frequency of excess joint spike events (JSEs). Subfigures show the test power (y-axis) of NeuroXidence (blue, dashed red, dashed orange) and of the unitary event method (UE method) (dashed black). The toy-data model consisted of 50 trials, each with 5 'simultaneous' spike trains. Synchronization of spike trains was modeled by an SIP process. The excess rate of JSEs beyond the chance level is given on the x-axis. The spike rate was 10 ap/s. The blue curve indicates the test power of NeuroXidence, the dashed black is the test power of the UE method, and in both cases, the JSEs were absolutely synchronous (dataset 1). The green curve shows the difference between the test power of the UE method (dashed black) and of NeuroXidence (blue). The dashed red curve shows the test power of NeuroXidence for the same data as used before, but each spike was jittered randomly by an allowed maximum of 5 ms (dataset 2). Thus, the jitter of spikes in the test pattern was the same as the maximal imprecision considered by NeuroXidence. The yellow curve indicates a test pattern with a jitter of 15 ms, which was three times larger ($\eta = 3$) than the maximal jitter considered by NeuroXidence (dataset 3). The latter test pattern was used to describe H_0 . The complexities of the test patterns are changing down the columns (1-4), and the lengths of the analysis windows are changing across the rows (1-3).

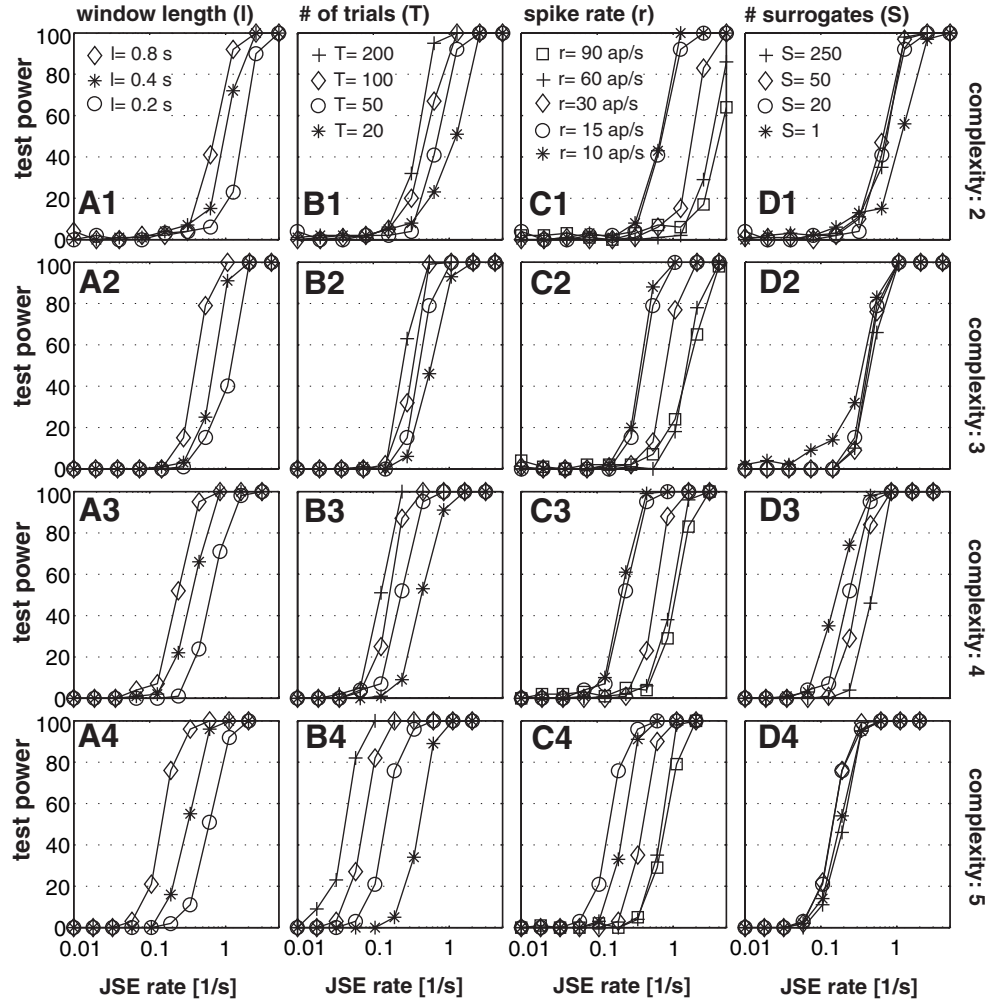


Figure 6.7.: **Test power.** Test power of NeuroXidence in relation to the window length (l), number of trials (T), spike rate (r), and number of surrogates (S). Subfigures plot the test power of NeuroXidence as a function of the frequency of joint spike events (JSEs) beyond chance level. Spike trains were analyzed by NeuroXidence and modeled as a single interaction process (SIP) based on Poisson processes. Rows (1-4) show the test power dependencies on the complexities of the analyzed joint spike patterns ranging from 2 to 5. In columns (A-D) variations in (A) the length of the analysis window l , (B) the number of trials T , (C) the spike rate r , and (D) the number of surrogates S from the standard parameter set ($T = 50$, $r = 15$ ap/s, $S = 20$, $\eta = 3$) are presented.

reliability of the estimates (Fig. 6.7D1-4).

To study the test power of sub-patterns and supra-patterns of individual JS patterns induced by an SIP process (Fig. 6.8), we performed a statistical evaluation of the frequency of occurrence of any potentially existing JS pattern in 4 toy-datasets. Each dataset comprised 50 trials of 16 simultaneous spike trains. The 4 toy-datasets differed in the complexities (2-5) of the mother-patterns, which are the JS patterns induced by the SIP process ('mother-process'; Fig. 6.8). Any JS pattern of complexity higher than 2 includes sub-patterns that are expected to be detected at least as frequently as the mother-pattern. On the other hand, the test power of each of these sub-patterns is expected to be lower than the test power of the mother-pattern, since the lower complexity leads to a higher chance frequency under H_0 . Thus, when sub-patterns are only induced by one mother-process and not by additional correlations with orders equal to or smaller than the complexity of the sub-patterns, the test power is expected to decrease with the decreasing complexity of the sub-pattern.

Supra-patterns are composed of the mother-pattern itself and spikes from additional neurons, which are by chance coinciding with the mother-pattern. Therefore, supra-patterns are of higher complexity than the induced mother-pattern. The maximal frequency of a supra-pattern is bound by the maximal frequency of any sub-pattern, including the induced mother-pattern. Furthermore, since the additional spikes are coinciding by chance, the frequency of the supra-pattern occurring is expected to be smaller than that for the mother-pattern. Thus, as long the excess frequency of the mother-pattern is not so high that the supra-pattern still occurs reliably across trials, the test power of the supra-pattern is expected to be substantially reduced in relation to the mother-pattern.

6.2.4. Comparison of the Test Power of NeuroXidence and of the UE Method

To compare the test power of NeuroXidence with that of the UE method, we applied both methods to two toy-datasets, which were characterized by the same set of parameters. Toy-dataset one contained JSEs that were absolutely synchronous, while toy-dataset two contained JSEs that were deduced from toy-dataset one by jittering each individual spike by an amount smaller than the allowed jitter τ_c (Fig. 6.6, 6.9). Positive values indicate a higher test power for the UE method. We used absolutely precise JS patterns to derive the test power of the UE method, since it is based on exclusive binning, which limits its capability to detect jittered JS patterns of complexities higher than 2 (Grün et al., 1999). The combination of parameters, given by 50 trials, a spike rate of the background process of 15 ap/s, 20 surrogates, and $\eta = 3$, served as a standard parameter set, from which 11 different toy-datasets were derived by changing either the window length ($l = 0.2, 0.4$ or 0.8 s), the number of trials ($T = 20, 50, 100, 200$), or the spike rate ($r = 10, 15, 30, 60, 90$ ap/s).

First, we studied the difference in the test power as a function of analysis window length. Initially, we generated toy-datasets where the NeuroXidence analysis window had the same length as that of the UE

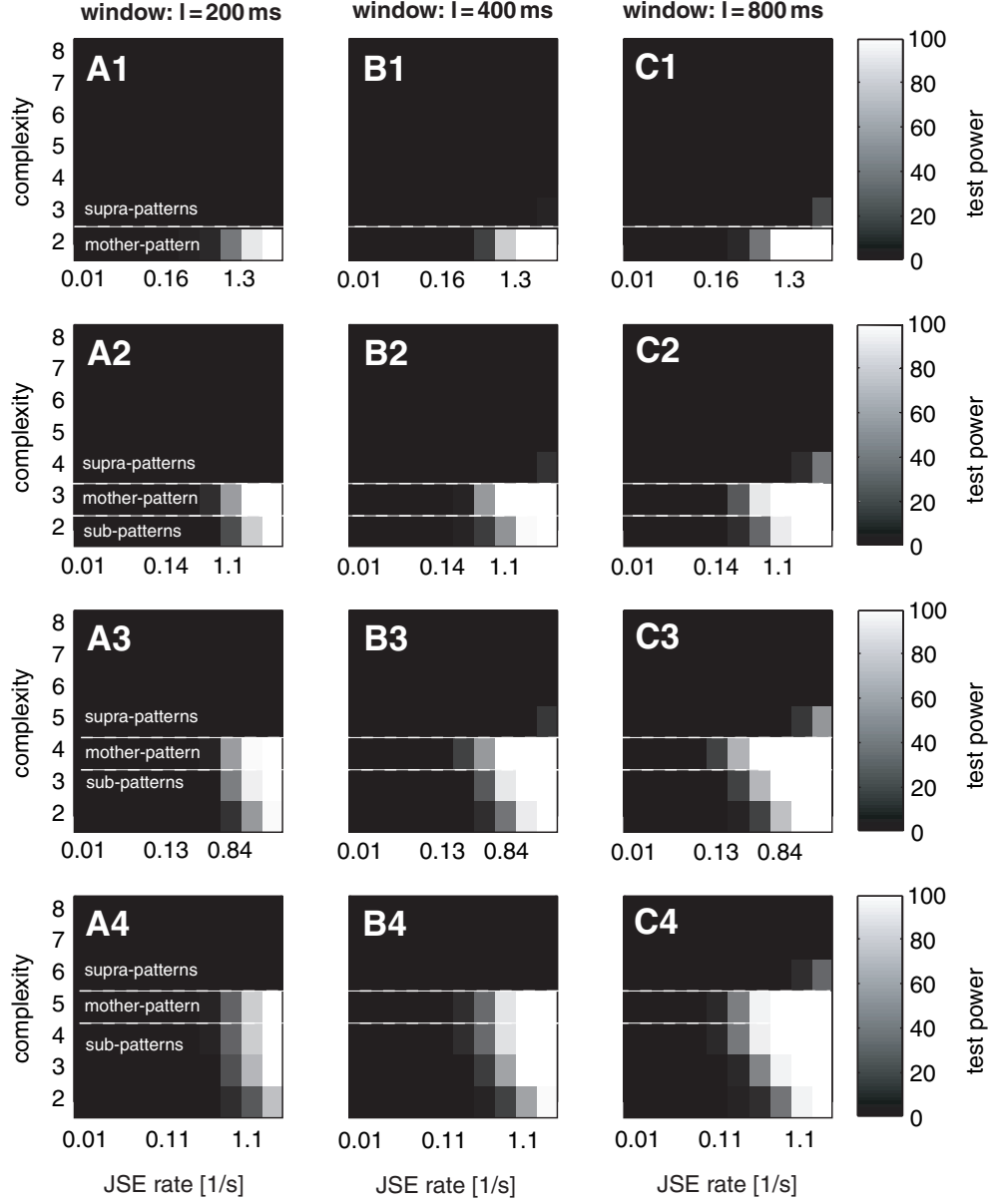


Figure 6.8.: **Test power of NeuroXidence for an induced mother-pattern and its supra-patterns and sub-patterns.** Each subfigure shows the gray-coded test power of a certain mother-pattern, all sub-patterns of lower complexities, and all supra-patterns of higher complexities. The excess rate of the JSEs beyond the chance level, which corresponds to the mother-pattern, is given on the x-axis. In row (1) the mother-pattern was of complexity 2, in row (2) of complexity 3, in row (3) of complexity 4, and in row (4) of complexity 5. The data used to derive the test power consisted of 50 trials of 16 spike trains, modeled as a single interaction process (SIP) that is based on homogeneous Poisson processes. A spike rate of 15 ap/s, $S = 20$ surrogates, and $\eta = 3$ were used by NeuroXidence for deriving the statistical significance. In columns (A-C) variations in the length of the analysis window l (A: $l = 200$ ms, B: $l = 400$ ms, C: $l = 800$ ms) are presented.

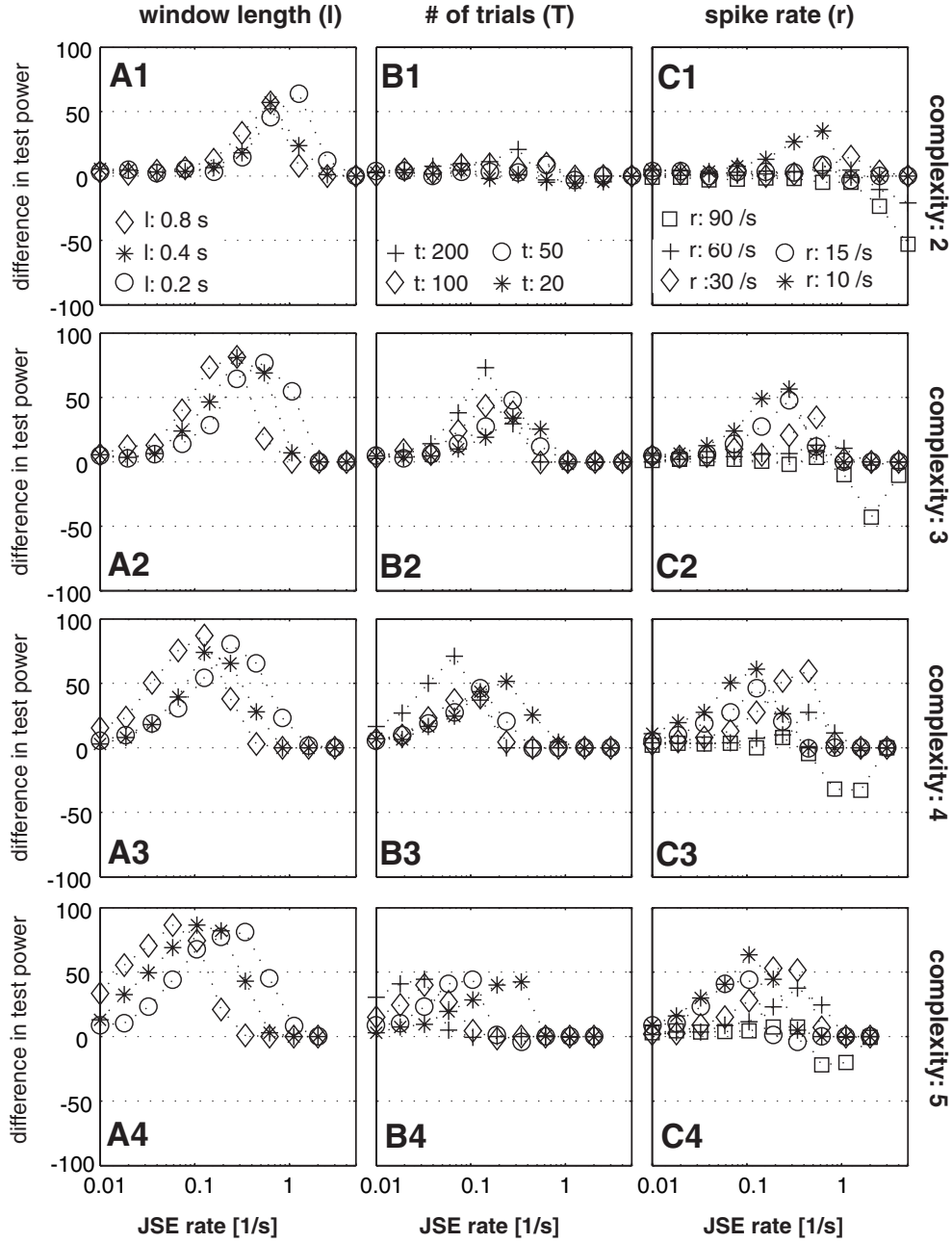


Figure 6.9.: **Test power NeuroXidence versus Unitary Event ('UE') method.** Difference in the test power between NeuroXidence and the UE method. (A-C) Variation of length l (A) of the analysis window, (B) the number of trials T , and (C) the spike rate r from the stand parameter set ($T = 50$, $r = 15$ ap/s, $S = 20$, $\eta = 3$) are presented. Spike trains were modeled as a single interaction process (SIP) based on homogenous Poisson processes. The frequency of excess joint spike events (JSEs) beyond the chance level is given by the x-axis. In column (A) the length of the analysis window used for both, NeuroXidence and the UE method, varied between 0.2 s, 0.4 s and 0.8 s. In (B) and (C) the NeuroXidence analysis window was 0.8 s while the UE window was 0.2 s long. Rows (1-4) show the difference in test power for test patterns of complexity 2-5.

method (Fig. 6.9A1-4). In that case the test power of the UE method was for all tested parameters higher than for the NeuroXidence method and the difference increased with increasing JS pattern complexity. However, since the length of the NeuroXidence analysis window is not constrained by the dynamics in the auto-structure, we increased the window length of the NeuroXidence window up to 0.8 s, while we kept the length of the UE analysis window at 0.2 s (Fig. 6.9B1-4, 6.9C1-4). The latter was selected to be compatible with the window length used to analyze real data and reflects the constraint that data inside a window have to be approximately described by a stationary Bernoulli process. The longer NeuroXidence window reduced the test power advantage of the UE method. For JS patterns of complexities higher than two, the difference between the test power of both methods was substantially reduced and dropped for complexity two JS patterns below 15% (Fig. 6.9B1).

Only in the case of high rates was the NeuroXidence test power higher than the test power of the UE method. The reason is that the UE method assumes that data can be approximately described by a binary Bernoulli process. Using this assumption, the method detects JSEs based on exclusive and binary binning. To ensure a binary process, the UE method has to utilize clipping, which only counts one spike, no matter how many spikes actually fell in the bin. Thus, the detection scheme of the UE method changes the data structure and does not detect all existing JS patterns. Since the toy-data were based on Poisson spike trains, increasing the spike rates led to an increased probability of more than one spike per bin. This led to an increasing number of JS patterns that went undetected by the UE method. It also explains the relative drop in the test power compared to NeuroXidence, which, in contrast, detects all existing JS patterns due to the preprocessing (see Appendix A.1) that uncovers overlapping JSEs.

6.2.5. False Positives and Test Power (Non-Stationary Process)

We compared NeuroXidence (Fig. 6.10C,D) with the widely-used shuffle-corrected cross-correlogram (Fig. 6.11A) and the UE method (Fig. 6.11B-D) using data generated by a non-stationary process. The toy-data comprised 18 simultaneous units and 50 trials, which were based on 15 periods, each 2 seconds long. The toy-dataset also contained features which are frequently observed in real datasets (Fig. 6.10A,B), such as low rates (periods 5, 6), rate modulations (periods 7, 8, 9, 10, 11, 12), latency co-variation of rate responses across neurons (periods 10, 12), trial-by-trial variability of rate levels (period 13), three different types of spike train generating processes (Poisson processes in periods 1, 5, 7, 8, 10-15, bursty γ -processes with a shape factor $\gamma=0.7$ in period 2 and $\gamma=0.3$ in period 3, as well as regular γ -processes with a shape factor $\gamma=7$ in periods 4, 6, 9). Only periods 14 and 15 contained JSEs above chance level that were based on correlations modeled by a multiple interaction process 'MIP' ((Kuhn et al., 2003), period 14: $r_{MIP} = 0.12$, period 15: $r_{MIP} = 0.3$).

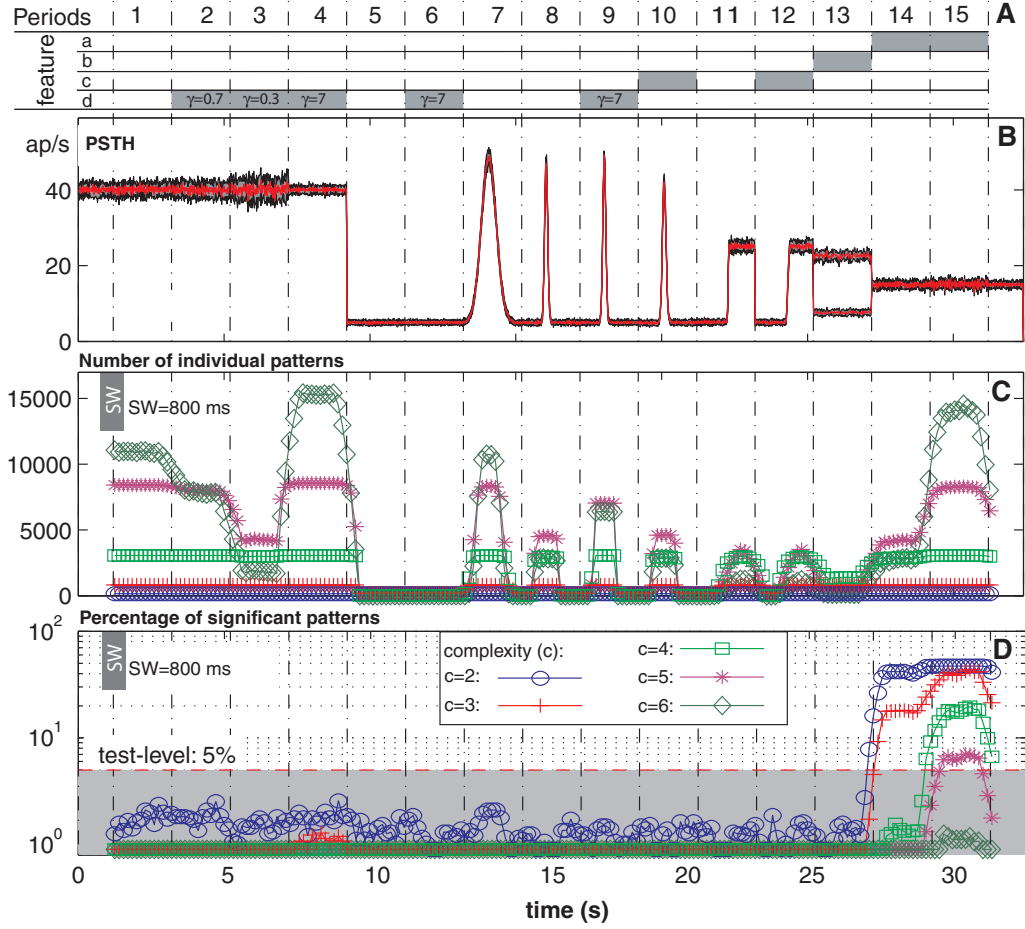


Figure 6.10.: **Test power in non-stationary data.** Statistical significance evaluated by NeuroXidence: NeuroXidence was used to detect a statistical excess of JSEs. Generated toy-data consisted of 50 trials of 18 ‘simultaneous’ spike trains. The toy-data contained 15 periods each 2 s long. Each period was described by one set of features used to describe the toy-data model and to generate the spike trains. An inhomogeneous and independent Poisson process served as a standard model. Four additional features, tabulated in panel (A), were used as modifications of this standard model. Only periods 14 and 15, indicated by feature **a** spike trains, were inter-correlated and thus exhibited a statistical excess of JSEs (MIP-process, period 12 correlation $r_{MIP} = 0.12$, period 13 $r_{MIP} = 0.3$). Feature **b**, which was used for period 13, indicates changing rates across trials and across neurons. Neurons 1-9 were modeled by a homogenous Poisson process with a rate = 15 ap/s, while the rate of neurons 10-18 changed from trial to trial between 15-30 ap/s. Feature **c** (periods 10, 12) represents latency covariations. From trial to trial, the latency for all 18 neurons varied randomly by the same amount between 0-100 ms. During the periods characterized by feature **d**, inhomogeneous gamma processes (shape factor $\gamma = 7$ for periods 4, 6 and 9 (regular), $\gamma = 0.7$ (bursty) for period 2 and $\gamma = 0.3$ (bursty) for period 3) were used instead of inhomogeneous Poisson processes. (B) PSTH displays the rate profile of the used inhomogeneous processes. During period 7, rate had been modulated between 5 and 50 ap/s with a Gaussian shape with $\sigma_t = 250$ ms, while during periods 8, 9, and 10, $\sigma_t = 50$ ms. The rates in periods 11 and 12 were modulated between 5 and 30 ap/s by a step function. (C) The number of individual and unique JS patterns of complexities 2 to 6 that were detected in each sliding window ($\tau_c = 5$ ms, ‘SW’ = sliding window: 800 ms). (D) The percentage of JS patterns that could not be explained by chance (test level 5%). In the case of independent processes (periods 1-13), this percentage corresponds to the percentage of false positives, while it corresponds to the test power during periods 14 and 15.

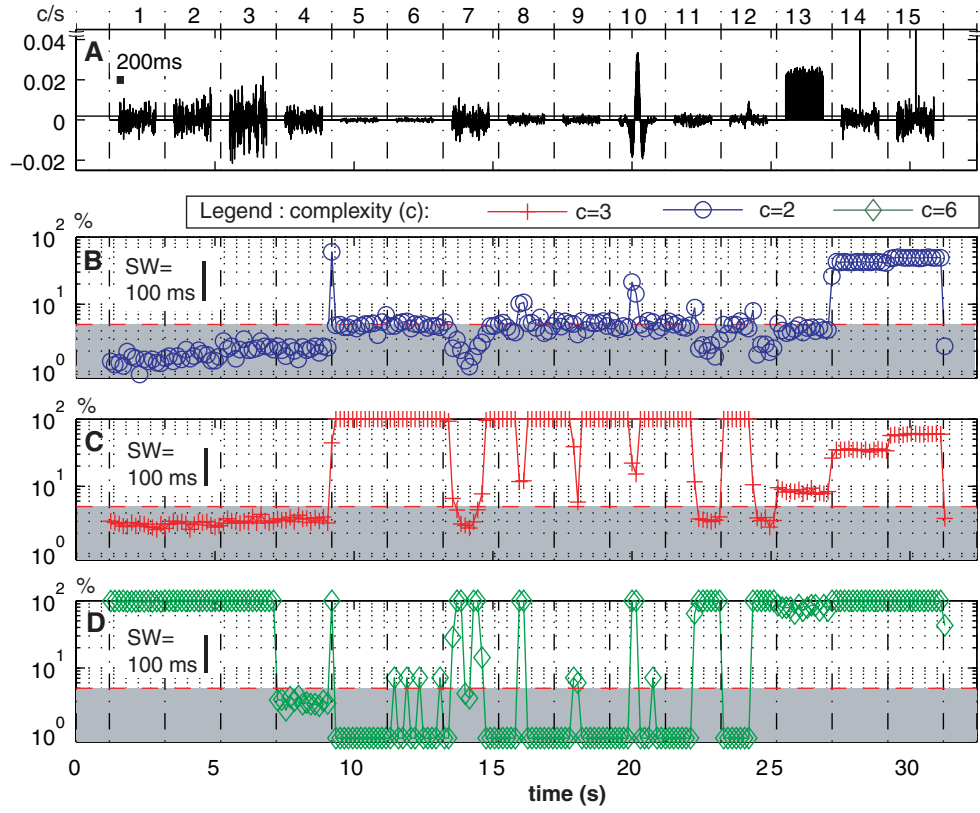


Figure 6.11.: **Comparison of cross-correlogram and UE method for detecting statistical excess of JSEs.** The comparison was based on the same toy-dataset as in Figure 6.10. (A) Cross-correlograms for the respective periods (the scale bar in period 1 indicates 200 ms). (B,C,D) The percentage of JS patterns, with complexities ranging from 2 to 6, that couldn't be explained by chance (5%). In the case of independent processes (periods 1-13), this percentage corresponds to the percentage of false positives, while it corresponds to the test power during periods 14 and 15. (binning $b = 5$ ms, 'SW' = sliding window with length 100 ms)

NeuroXidence was applied to the data in a sliding window of length 800 ms, and it performed significance tests on excesses of JSEs with complexity 2-6. The statistical significance was evaluated for each occurring JS pattern. We derived the percentage of JS patterns that occurred significantly more frequently than expected by chance (Fig. 6.10D). To make comparisons across complexities, the number of significant JS patterns per complexity was normalized by the total number of identified JS patterns per complexity (Fig. 6.10C). Spike trains during periods 1-13 were independent, which implied that H_0 should not be rejected more often than the test level (5%), if the actual-false-positive-rate was conservative. The percentage of false rejections of H_0 corresponding to the actual-false-positive-rate is clearly below 5% for all complexities and throughout all sliding windows during periods 1-13. Therefore, NeuroXidence is a hypothesis test with a conservative actual-false-positive-rate that is effected neither by low rates, rate modulations, latency variability, and cross-trial rate chances, nor by different model processes (γ -process, Poisson) used to generate the spike train. High test power during periods 14 and 15, which

6. NeuroXidence

was characterized by increased frequencies of JSEs based on an MIP correlation with $r_{MIP} = 0.12$ (period 14) and $r_{MIP} = 0.3$ (period 15), demonstrates the applicability of NeuroXidence for the detection of excesses of JSEs.

Unlike NeuroXidence, the shuffle-corrected cross-correlogram (bin-width $b = 20$ ms, data segments $= 800$ ms) indicates the existence of cross-structure due to its modulations in periods 10 and 12 (Fig. 6.11A). The reason for these modulations is the shuffle-correction (Baker and Gerstein, 2001; Brody, 1999). Periods 10 and 12 are characterized by latency co-variations across neurons that are uniformly distributed between 0 and 100 ms across trials. Thus, the cross-correlogram of the shuffled data is broader, since trials with different latencies from different trials are combined. These differences in the widths and the slopes between the cross-correlogram and its shuffled version induce modulations in the shuffle-corrected cross-correlogram, as it is the same for the difference of two Gaussians with different variances. This indicates that modulation in the shuffle-corrected cross-correlogram does not necessarily indicate an excess of fine-temporal cross-structure. Instead, modulations can be induced by fast, and across neurons coherent rate changes which are combined with latency co-variations across trials, which lead to a false correction by the shuffled data (Baker and Gerstein, 2001; Brody, 1999) and to broad peaks in the cross-correlogram. In contrast, fine-temporal cross-structure, such as that in periods 14 and 15, induces very tight peaks with widths equal to the precision of the induced JSEs.

We also applied the UE method to the same dataset. Since the method assumes that data in every sliding window can be described by a stationary Bernoulli process, the sliding window was chosen to be rather short (100 ms) compared to the 800 ms long NeuroXidence window. This window length of 100 ms justifies the assumption of a stationary Bernoulli process for a slower modulation of the rate profiles. Therefore, 100 ms is a good trade-off between the danger of false positive events induced by non-stationarity and reduced test power due to a small amount of data used by each window.

For the UE method, we derived the actual-false-positive-rate per JS pattern complexity, which ranged from 2 to 6 during periods 1-13. As with NeuroXidence, the actual-false-positive-rate was defined as the rate of false rejections of H_0 normalized by the number of occurred JS patterns.

As expected, rate changes on a time scale slower than the sliding window (100 ms) were fully corrected for complexity 2 JSEs (Fig. 6.11B), while fast changing rates lead to an increased actual-false-positive-rate (transition from period 4 to 5, and during periods 8 and 10). For high complexity JS patterns fast and relatively slow rate changes lead to an increased actual-false-positive-rate (transition from period 4 to 5, periods 7, 8, 9, 10, 11, and 12 (Fig. 6.11D)). The second source for increased actual-false-positive-rates was a low spike rate. During periods 5 and 6 and all parts of other periods in which the rate was low (5 ap/s), the actual-false-positive-rate increased dramatically for complexity 2 JS patterns and up to 100% for complexity 3 patterns (Fig. 6.11C). In contrast, the actual-false-positive-rate for complexity 6 JSEs was close to zero. The reason for the decrease from 100% actual-false-positive-rate for complexity 3 to

an actual-false-positive-rate of 0% for complexity 6 JS patterns was either that complexity 6 JS patterns did not exist due to the low expectation of high complexity JSEs, or that they were not detected due to the exclusive binning and the rather short analysis window used by the UE method. In the instance that zero JS patterns of one certain complexity were detected, the corresponding actual-false-positive-rate per complexity was defined to be zero. The reason for the dramatic increase in the actual-false-positive-rate is that the UE method does not consider any trial-based variation in the frequency of occurrences, rather it tests if the total number of a particular JSE across all trials is significantly different from the expected total number based on the assumption that the latter is Poisson distributed. Since the significance estimation of the UE method is only based on the variability of the expected total number and not on the variability of frequency of occurrences across trials, a few or even only one JSE might be evaluated as a significant excess if the expected total number is low enough. The latter is the case if spike rates are low or the complexity is large. In conclusion, the comparison of the NeuroXidence method with the widely used shuffle-corrected cross-correlogram and the UE method illustrates the performance and robustness of the NeuroXidence method and motivates its usage for reliable detection of periods with excesses or deficiencies of JSEs.

6.2.6. Test Power of NeuroXidence for Oscillatory Processes

Since oscillatory modulations of the spike rate are frequently observed in experimental data (Fries et al., 2002; Fries et al., 2001; Fries et al., 2001a; Fries et al., 1997; Gray et al., 1989; Lachaux et al., 2000; Rodriguez et al., 1999; Tallon-Baudry et al., 1999; Tallon-Baudry et al., 2004), we studied if such modulations lead to rejection of H_0 and to support of H_1 . To this end we assumed as a model a continuous instantaneous rate function $r_i(t)$ that is coherently modulated across neurons by a sinusoidal oscillation with frequency ν .

$$r_i(t) = \frac{A}{2} * [1 + \sin(\nu \cdot t)] \text{ with } i = 1 \dots N \quad (6.7)$$

We used the instantaneous rate function ($A = 10$ ap/s) to model spike trains by inhomogeneous Poisson processes and estimated numerically the test power of complexity 2 JS patterns. NeuroXidence, based on 50 trials of 0.8 s long analysis windows, was applied to 100 independent realizations of the data. The oscillation frequency ranged between 1 and 80 Hz, while the jitter-kernel width was kept fixed ($\tau_r = 11$ ms in Fig. 6.12B and $\tau_r = 21$ ms in Fig. 6.12A). For oscillation frequencies below 6 Hz in the case of $\tau_r = 21$ ms and 9 Hz for $\tau_r = 11$ ms, the percentage of significant events stayed below the test level of 5%. With increasing oscillation frequencies the test power increased and reached up to 100% at 50 Hz for $\tau_r = 21$ ms and at 80 Hz for $\tau_r = 11$ ms.

In addition to this numerical estimation of the NeuroXidence test power, we used an analytical description of the expected difference Δf between the frequency of JSEs in the original $f(org)$ and the jittered

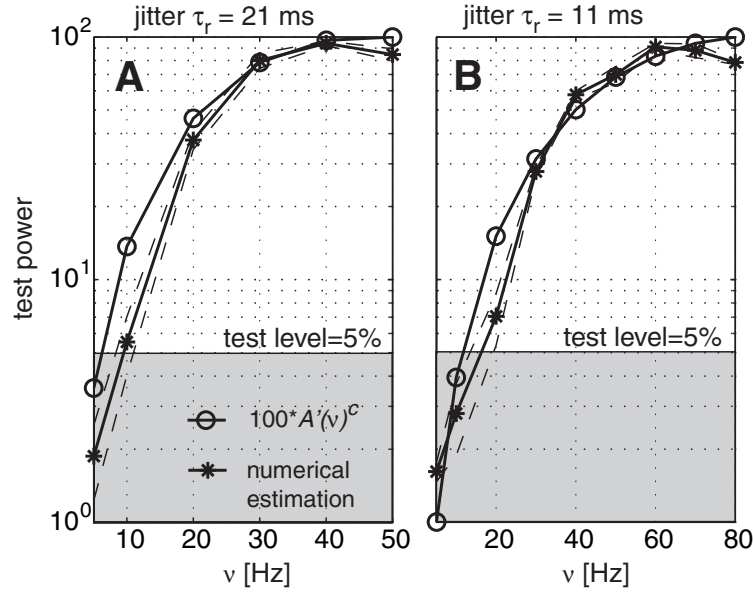


Figure 6.12.: **Test power in case of oscillatory rate modulations.** Test power of NeuroXidence based on data modeled as independent and inhomogenous Poisson processes. The instantaneous rate functions were modulated by a sinusoid with frequency ν and a peak rate of 20 ap/s. Twenty surrogates and 21 ms of jitter were used in (A), while in (B) 11 ms of jitter were used for deriving statistical significance. The black solid curve (*) shows the mean test power (dashed black: confidence interval of $1 * \sigma$) based on 100 independent realizations of the same data model, while the black curve using circles (o) as symbols shows the theoretical prediction given by $A'(\nu)^c$.

$f(\text{sur})$ datasets to approximate the theoretically predicted test power $A'(\nu)^c$ (see Appendix A.3). The estimate Δf_t^k of the difference Δf assessed in multiple trials ($t = 1 \dots T$) enables NeuroXidence to derive the statistical significance of a JS pattern k . Thus, the analytical description of Δf_t^k allowed us to study to which degree a frequency-dependent modulation of Δf_t^k explained a modulation of the test power of NeuroXidence. Since NeuroXidence is using Δf_t^k as an estimate for $\langle f \rangle_c$, we use $A'(\nu)^c$ (see Appendix A.3, Equation A.5) to predict the modulation of the test power, which depends on the frequency (ν) of the coherent oscillatory rate modulation of $r_i(t)$. The agreement between the theoretical prediction of the test power, in percent given by $100 * A'^c$, and the numerical estimation based on toy-data, both for $c = 2$, indicates that the modulation of the test power is well described by the analytical solution. Therefore, $A'(\nu)^c$ can be used to discuss the impact of coherent oscillations on the probability of a rejection of H_0 .

6.2.7. Computational Complexity

To estimate the computational complexity of NeuroXidence as a function of the number of neurons and the amount of allowed jitter τ_c , we generated toy data containing 20 trials, of length 0.8 s, based on a spike train model given by a homogeneous gamma-process (with γ -shape factor of 10) and a spike rate of 10 ap/s. The number of neurons that were contained in the dataset changed between 2 and 48.

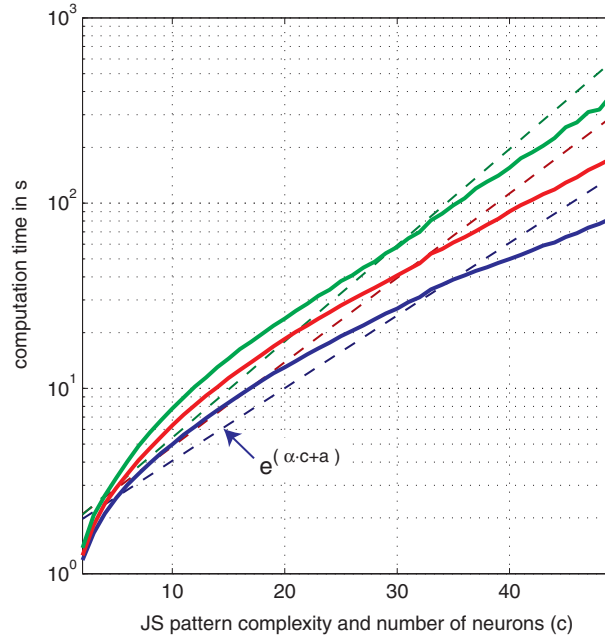


Figure 6.13.: **Computational complexity of NeuroXidence.** Computational complexity of NeuroXidence as a function of the number of neurons that were analyzed. The y-axis gives the time needed to analyze the data on one JS pattern with a complexity equal to the number of neurons. The data consisted of one $l=800$ ms sliding window and 20 trials. Spike trains were modeled by an homogeneous gamma process (shape factor $\gamma=10$) with a spike rate of 10 ap/s. The x-axis gives the complexity of the analyzed JS pattern, which was equivalent to the number of neurons in the data-set. The color of the individual curves corresponds to the maximally allowed jitter τ_c in units of the bin length (blue = 3, red = 5, green = 7). Each curve represents the average computational complexity measured for 15 independent realization of the same data model. The dashed curves give the expected computational complexity if complexity were an exponential function of number of neurons $y = e^{\alpha c + a}$.

The allowed jitter varied between 3, 5, and 7 bins which is equivalent, in the case of a bin length of 1 ms, to $\tau_c = 3, 5$, and 7 ms. For each dataset all existing JS patterns were detected but the statistical significance was derived for only one JS pattern with the maximal complexity. For each parameter set, the computational complexity of NeuroXidence grew less than exponentially as the number of neurons increased (Fig. 6.13). This makes NeuroXidence, in contrast to the multiple shift method (Grün et al., 1999), applicable to data-sets with large number of neurons.

6.3. Discussion

6.3.1. Discussion of the NeuroXidence Method

The rate and assembly hypothesis about the neuronal coding strategy in the cortex have been very controversially discussed in the last years (Aertsen et al., 1989; Aertsen et al., 2001; Baker and Lemon, 2000; Brody, 1999; Fetzer, 1997; Gerstein et al., 1989; Grün et al., 2003; Ikegaya et al., 2004; von der Malsburg,

6. NeuroXidence

1981; Riehle et al., 1997; Shadlen and Newsome, 1994). The rate hypothesis assumes that information is encoded in the spiking of individual cells, while the assembly hypothesis assumes that coordinated firing of cell assemblies with a millisecond precision serves the coding and binding of information. To dissociate the assembly hypothesis from the rate hypothesis, the H_0 has to consider all features of the data that might be causally related to the rate hypothesis. Only then is a rejection of H_0 equivalent to support for the assembly hypothesis and evidence for cooperative coding based on JS activity.

NeuroXidence implements a statistical hypothesis test that allows for a dissociation of both hypotheses. In contrast to other methods (Abeles and Gerstein, 1988; Grün et al., 1999; Grün et al., 2002a; Grün et al., 2002b; König, 1994), NeuroXidence is a non-parametric statistical hypothesis test and does not assume any model about the spike train generating process. The estimation of the statistical significance of JS patterns is based on a comparison of the frequency of JSEs between the recorded dataset and a surrogate dataset, which is derived from the recorded dataset but lacking any coordinated-firing. Random shifting of spike trains was utilized to destroy fine temporal cross-structure while keeping the complete auto-structure in the surrogate dataset intact. Therefore, NeuroXidence estimates the chance frequency of JSEs under H_0 by allowing for the consideration of rate modulations or for history dependences, which might exist in each individual spike train.

To this end, NeuroXidence defines two time-scales τ_r and τ_c . The fast time-scale τ_c corresponds to fine-temporal joint activity across neurons and is believed to be in the millisecond range. The second time-scale τ_r is a lower bound of the timescales of rate modulations that are coherent across neurons and that should be considered as rate, not as fine-temporal joint activity. The scale between τ_r and τ_c is given by η . As we have shown already, small amounts of shifting with $\eta = 3$ lead to a considerable test power of the NeuroXidence method. Thus, even in cases of very small differences between the timescale τ_r and τ_c (e.g. $\tau_c = 15$ ms for $\tau_r = 5$ ms and $\eta = 3$), NeuroXidence can detect a significant excess or deficiency of JSEs and therefore it allows for a clear distinction between the rate and the assembly hypothesis.

6.3.2. Oscillatory Processes

Since oscillatory processes are frequently observed in neuronal data, we studied the impact of oscillatory and, across neurons, coherent rate modulations by using toy-data based on inhomogeneous Poisson processes whose rates were modulated sinusoidally. Based on this toy-data, we estimated the test power numerically and also derived an approximate analytical solution of the test power (see Section 6.2.6 and Appendix A.3). Both, the analytical approximation and the numerical estimation were in agreement and increased with increasing frequency of the coherent rate modulations. This behavior is expected since NeuroXidence discriminates between fast processes, that are defined on a timescale τ_c corresponding to the precision of JSEs, and processes that are slower than a timescale τ_r . Therefore, we have to compare the timescale of the coherent oscillatory rate modulation with both timescales τ_c and τ_r .

Since the slope of the sinusoid is constantly changing, there is not one clearly defined timescale for which the rate does not change, as it would be the case for a series of square pulses. However, we can define an upper bound of the fast timescale by using the interval between the two steepest slopes in one period of a sinusoid, which correspond to the zero crossings ($\sin(x) = 0$) at $\varphi = 0$ and $\varphi = \pi$. Since the steepness of the slopes is highest at these two points, the strongest modulation of the rate is bounded within an interval of length $T/2$. For example, assuming an oscillation frequency of $\nu = 50$ Hz, equivalent to a period $T = 20$ ms, $\tau_c = 5$ ms and $\tau_r = 20$ ms. Thus, an oscillatory rate modulation of $\nu = 50$ Hz, coherent across neurons $i = 1 \dots N$, induces a fast cross-structure on a time-scale of 10 ms or faster, which is slower than the fast timescale τ_c and faster than the slow timescale τ_r . Therefore, we expect that jittering of individual spikes destroys fine-temporal cross-structure, as it is predicted by the analytical approximation and observed in the toy-data.

To summarize, if coherent rate-modulations are on a time-scale that is comparable with τ_c , NeuroXidence is expected to reject H_0 and to support H_1 . This illustrates that the definition of coordinated firing and fine-temporal cross-structure includes explicitly fast and, across neurons, coherent changes of the instantaneous spike rate, which occur on a timescale comparable to τ_c .

6.3.3. Rare and Spurious Events versus Reliably Reoccurring Events

NeuroXidence differentiates between rare and spurious events and events that are reliably reoccurring. Only if events are reliably reoccurring across trials might they be classified as JS patterns with a significant excess or deficiency of JSEs. This allows NeuroXidence to estimate the statistical significance reliably even for JS patterns of complexities higher than 2 that are expected to be rare. In contrast, other methods, which are based only on the total frequency across trials or the frequency in one single trial, might reject H_0 if the expected frequency is low enough that a few occurrences or even one single JSE is enough to be evaluated as a significant excess. Since the latter gets more likely for higher complexity JSEs, for which the expectancy is exponentially decreasing, methods based on single trials or on the total number in all trials are likely to fail to differentiate between rare and spurious events and events based on an underlying correlation.

6.3.4. Analysis Window Length

NeuroXidence can be applied to short time windows to track transient neuronal states. In contrast to other methods, the NeuroXidence window length is not restricted by any stationary assumption on the data, since the full auto-structure including any rate modulation is considered by the statistical hypothesis test. Thus, the length of the NeuroXidence analysis window can be freely chosen by the operator to match the assumed duration of neuronal states of interest.

6.3.5. NeuroXidence versus UE Method and Shuffle-Corrected Cross-Correlogram

To compare NeuroXidence, the unitary event ('UE') method, and the shuffle-corrected cross-correlogram, we applied all three methods to the same toy-dataset based on 18 simultaneous Poisson- or Gamma-processes. The toy-data contained epochs with constant rates between 5 and 40 ap/s, rate modulations between 5 and 45 ap/s, latency variability, and trial-by-trial variability of rate-levels. Only NeuroXidence did not exhibit an increased false positives rate for any of the epochs of the toy-data indicating that the consideration of the auto-structure, as well as a trial-based evaluation, is crucial for dissociating the rate and the assembly hypothesis.

However, if NeuroXidence and the UE method use the same length of analysis window, the test power of the UE method is higher. This has two reasons. First, the UE method is a parametric approach and assumes a model which in general improves the test power. The disadvantage is that a model-based approach is only valid as long as the model is describing approximately well the real observed data. Second, NeuroXidence incorporates the trial-by-trial variability of the frequency of occurrence. Therefore, it requires that the frequency of JSEs across all of the trials is reliable and for most of the trials is higher in the real than in the surrogate dataset. In contrast, the UE method is only based on the total frequency across trials. If the expected frequency is very low, a few occurrences or even one single JSE across all trials might lead to a rejection of H_0 . This of course increases the sensitivity, as reflected by the higher test power, but on the other hand, this also increases the probability of false rejections in the case of rare events. The danger of false rejections of H_0 by the UE method, in combination with the decreasing difference between the test power of both methods, if NeuroXidence is based on longer windows, strongly indicates the necessity of using NeuroXidence instead of the UE method.

In conclusion, we have presented the new analysis tool NeuroXidence for detecting coordinated firing events in spike trains. We have demonstrated its performance, reliability, and applicability compared to the capabilities of other popular and currently-used methods. The results from analyzing real data with NeuroXidence give strong evidence for cooperative coding and the assembly hypothesis.

Part III.

New Tools Applied to Data

7. High Complexity Joint Spike Activity in Cat Visual Area 17

We applied NeuroXidence to data recorded in visual area 17 of an anaesthetized cat. Two Michigan probes (each 16 channels) were used to record neuronal activity in response to 20 repetitions of a drifting sinusoidal grating covering the cells' receptive fields (supplementary Fig. C.1). The contacts at each probe were arranged into a 4x4 matrix with 200 μm spacing between the nearest contacts, allowing us to record from different layers and columns simultaneously. Gratings were presented for 3.5 seconds and were oriented optimally for eliciting strong responses in most of the cells (see Appendix E for details on experimental methods). By using off-line spike sorting techniques, 48 single-units were isolated and were stable throughout the entire recording period of about 30 minutes.

7.1. NeuroXidence Based Analysis of Single-Unit Activity from 48 Cells

We computed post-stimulus time histograms (PSTHs) (supplementary Figure C.2) as well as cross-correlation histograms (Aertsen and Gerstein, 1985; König, 1994) for all pairs of neurons and for the period during which neurons showed sustained responses to the presented stimuli. After the subtraction of shift-predictors (König, 1994), cross-correlograms indicated that a number of spikes are synchronized with a precision of several milliseconds (supplementary Fig. C.3), raising the question whether higher complexity JSEs might exist in this dataset.

The NeuroXidence algorithm was applied to the entire duration of the recording, including the 1 s of pre-stimulus period as well as the 1 s post-stimulus activity (Fig. 7.1). The average computation time per sliding window amounted to 35 min on a 1.4 GHz CPU (Pentium M) with 1 GB RAM. We investigated JSEs of complexities 2 to 48 and used 1% as a test level ($\tau_c = 5$ ms, $\tau_r = 20$ ms, $S = 50$, bin-size = 1 ms, Wilcoxon-test). Up to 18,000 different and individual JS patterns per sliding window (length 0.5 s, window shift 0.16 s) were detected (Fig. 7.1A). JS patterns occurred significantly more often than expected by chance only during the period of stimulation and their complexities ranged between 2 and 6. No significant JS patterns were found of complexity 7 or higher (Fig. 7.1B).

7. High Complexity Joint Spike Activity in Cat Visual Area 17

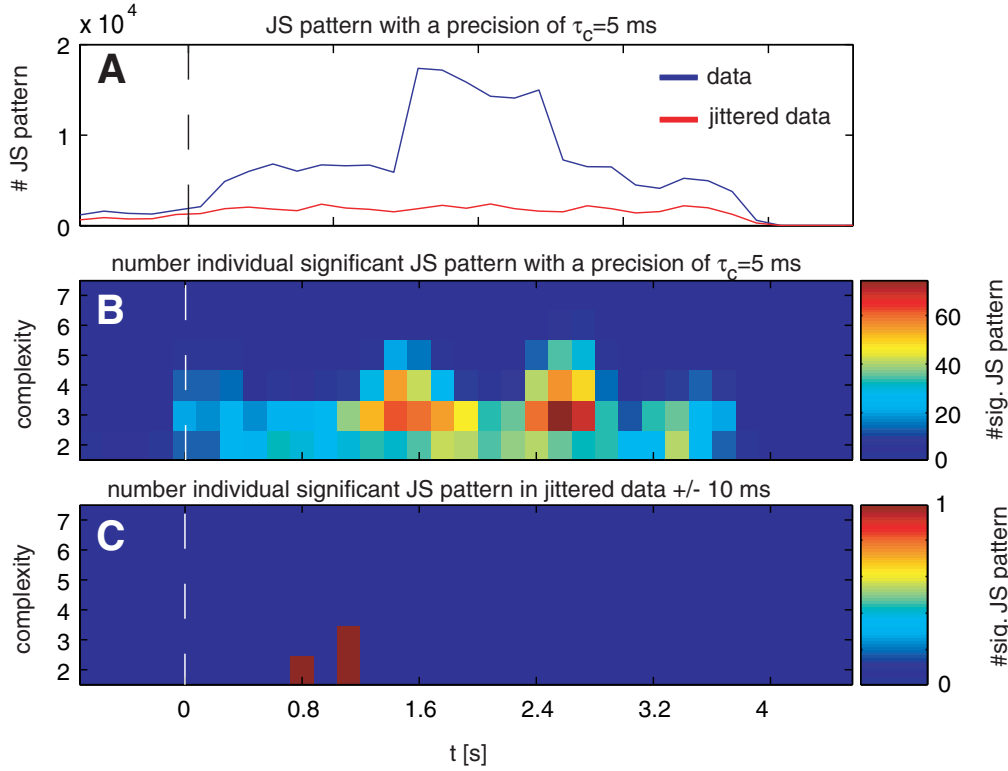


Figure 7.1.: **Analysis of simultaneously recorded single-unit activity from 48 neurons from an anaesthetized cat.** The recordings were made in visual area 17 during visual stimulation with moving gratings, which were presented for 3.5 seconds. Stimulus onset occurs at 0 s. Data comprise 20 trials. NeuroXidence was applied to sliding windows (0.5 s) that were shifted by 0.16 s. **(A)** The number of unique JS patterns detected per sliding window (no significance criteria applied). The blue curve corresponds to the original recorded data, and the red curve is based on data in which each spike was jittered by ± 10 ms to destroy fine-temporal cross-structure. **(B,C)** The frequencies of significant JS patterns (test level 1 %) for each sliding window (x-axis) are color coded and sorted according to their complexities (y-axis). Frequencies shown in **(B)** are based on the original data set, while in **(C)** the spikes in the original data were jittered by ± 10 ms to destroy fine-temporal cross-structure.

Interestingly, the time-course of the total number of JS patterns detected (Fig. 7.1 A) was considerably different from the time-course of the number of JS patterns that exceeded the chance level. The time-course of significant number of patterns partially followed the temporal rhythm of the stimulation, the number of patterns recurring with the period of 0.82 s with which the stripes of the grating crossed the receptive fields (compare to PSTHs in supplementary Fig. C.2). In contrast, the time course of the total number of JSEs indicated the largest number of patterns in the middle of the stimulation period at 1.7 s. Also remarkable is the steep increase and decrease at the beginning and the end of the period of stimulation.

To validate the results given by the number of JS patterns that exhibited a significant increase of JSEs, we also applied NeuroXidence to a dataset derived by jittering from the original dataset. In the

new dataset we destroyed any coordinated firing by jittering of spikes with ± 10 ms (Fig. 7.1A,C). This jittering of the data is the same as that used by NeuroXidence to derive the difference between the real data and H_0 ($j = \pm 10$ ms, $\eta = 4$). Thus, the application of NeuroXidence to the jittered dataset allows the derivation of the chance frequency of significant JS patterns (H_0 against H_0). The number of significant JS patterns that occurred in the jittered version of the original dataset amounted to 0.2% of the number observed in the original dataset. Thus, the error rate was much smaller than the nominal alpha value of 0.01, indicating that the test is conservative.

The positive side is that 99.8% of the significant JS patterns cannot be explained by chance (Fig. 7.1C). These results demonstrate that NeuroXidence allows for an evaluation of JS patterns independent of rate modulations, auto-structure, and trial-by-trial variability. Thus, the dynamics of cells that exhibit significant JS activity across different complexities can be separated from the various rate modulations observed for different cells in the data. These rate modulations concerned either the strength of the responses or the relative phase shift of the responses, because of the different times the stimulus was crossing the cells' receptive fields. This highlights the strength of NeuroXidence for investigating the complex interactions of neurons in terms of fine-temporal structure that might coexist with rate modulations of individual neurons. These results suggest that the present method is applicable to the analysis of neuronal activity in visual cortex and that it can be used to investigate changes in the number and structure of JSEs, as a function of different experimental variables such as the stimulus properties, behavior, or attention.

7.2. Discussion: Significant Joint Spike Activity in Real Data Evaluated by NeuroXidence

Our results demonstrate that JSEs with a precision of 5 ms occurred significantly more often than expected by chance for complexity 2 to 7 JS patterns. Since the frequency of JS patterns with excess JSEs is increased only during the period of stimulus, our results indicate that stimulus encoding involves fine-temporal cross-structure induced by neuronal correlations. We validated our result by analyzing a jittered version of the original data. This validation was equivalent to a comparison of H_0 against H_0 and served as the estimation of the chance frequency of JS patterns with excess JSEs. Since only 0.2% of JS patterns with excess JSEs remained significant after jittering, we conclude that the increase of the frequency of JS patterns with excess of JSEs is highly significant for complexities ranging between 2 to 7. The validation also revealed that fine-temporal cross-structure on a timescale of 5 ms is present in simultaneous recordings from cat visual area 17. This is, to our knowledge, the first proof that considers the complete auto-structure and trial-by-trial variability of spike trains when calculating JSE frequencies in real data.

7. *High Complexity Joint Spike Activity in Cat Visual Area 17*

8. Short-Term Memory Related LFP Oscillation in the Prefrontal Cortex

The maintenance of information during *Short-Term Memory* (STM) has originally been associated with sustained firing of cells in prefrontal cortex (Fuster and Alexander, 1971; Goldman-Rakic, 1995; Miller et al., 1996; Rainer and Miller, 2000). It has been proposed that persistent neuronal activity is sustained by cellular bistability (Lisman et al., 1998; Seamans et al., 2001) or by reverberating neuronal activity (Durstewitz et al., 2000; Machens et al., 2005). Reverberating neuronal activity that could serve the maintenance of stimulus information has been proposed to express itself in common rate modulations (Golomb et al., 1990) or irregular (Compte et al., 2003) or oscillatory activity patterns (Compte et al., 2000). If oscillatory activity is involved in the organization of STM, the question arises whether oscillations participate in the coding of a specific content (Lisman and Idiart, 1995; Singer, 1999), or reflect unspecific arousal due to behavioral demands, or provide a temporal reference signal for the coordination of distributed representations (Lisman and Idiart, 1995; Sommer and Wennekers, 2001).

In the context of distributed perceptual processing in visual cortex, synchronized beta- and gamma-frequency oscillations have been shown to be involved in attention dependent stimulus selection (Bichot et al., 2005; Fries et al., 2001b; Taylor et al., 2005). Oscillations provide an effective mechanism for the synchronization (Pikovsky et al., 2001) of neuronal activity both at global and local scales (Buzsáki and Draguhn, 2004) and, thus, could play a role in the coordination of STM sub-processes serving the encoding, retention, and retrieval of information. Physiological evidence for neuronal oscillations participating in STM has been obtained in parietal (Pesaran et al., 2002) and extrastriate temporal cortex (Tallon-Baudry et al., 2004), but not for prefrontal cortex.

8.1. Introduction

In this study we investigated first, whether task-related oscillations occur in local field potentials (LFPs) of ventral prefrontal cortex and second, whether power and phase-locking of oscillations in various frequency bands is correlated with behavioral performance and identifies of the presented stimuli. To this end we use a new tool that has been developed during this thesis (Section 4.1.4 and Chapter 5).

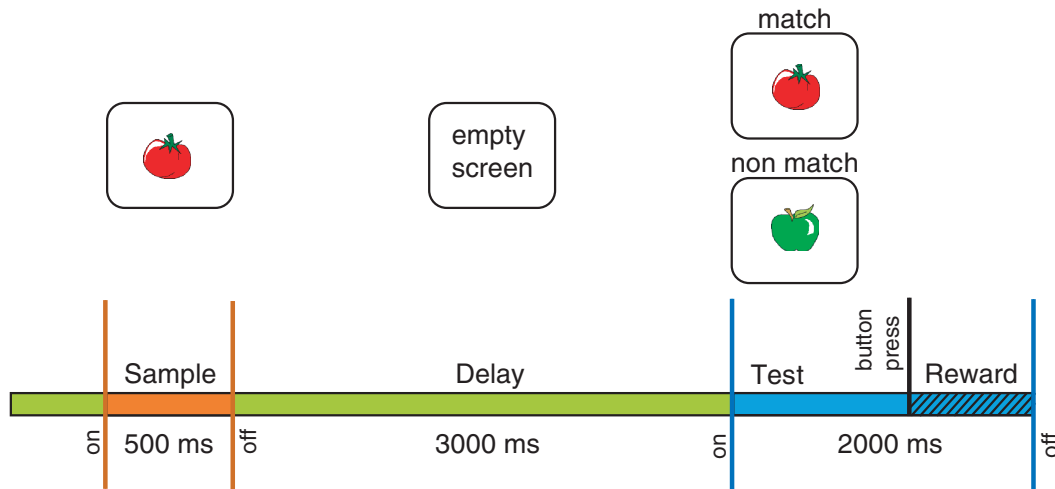


Figure 8.1.: **Time course of the visual short-term memory task.** After a 0.5-1 s baseline, a sample stimulus was presented for 500 ms which was followed by a 3 s delay. Thereafter a test stimulus was presented for 2 s during which the monkey had to respond by a differential button press. In case of a matching test stimulus, the left of two buttons in front of the monkey had to be pressed and in case of a non-matching test stimulus the right button. Thus, this visual memory task combines the classical matching-to-sample and non-matching-to-sample tasks so that the monkey has to decide in each trial whether a matching or non-matching test stimulus was presented. The reward was delivered after the monkey released the button which happened on average 200-300 ms later.

We trained two female monkeys (*Macaca mulatta*) to perform a visual memory task (0.5 s sample, 3 s delay, 2 s test presentation; Fig. 8.1). The percentage of correct responses made by the monkeys ranged between 71% and 87% for a given session. Anatomical MRI scans (T1-flash) were performed in a 1.5 T MRI scanner and used to guide implantation of recording chambers and to reconstruct recording positions (Fig. 8.2). Simultaneous recording of unit and field potential activity was performed with up to 16 Pt-W-in-quartz fiber microelectrodes from ventral prefrontal cortex. Signals were filtered (0.5-5 kHz (MUA) and 5-150 Hz (LFP) 3dB/octave) and digitized at 1 kHz, preprocessed by rejecting artifacts (movements, licking) and removing line noise at 50 ± 0.5 Hz. Spectral analyses were performed for trialsets constructed from the stimulus and behavioral protocol using the NeuroMeter software package (J. Munk). Stimulus selectivity of encoding and maintenance-related multi-unit activity, estimated by ANOVA (5% level), amounted to 25% stimulus-selective sites (supplementary Fig. G.3), which is compatible with published data for single-unit recordings (Miller et al., 1996; Rainer and Miller, 2000).

8.2. Dynamics of Local Field Potentials

The dynamics of LFPs (Fig. 8.3B) were analyzed for evoked responses (supplementary Fig. G.4), spontaneous and induced oscillations (Fig. 8.3C). Evoked responses were observed in 78% of the recording sites during the presentation of sample and test stimuli with a mean latency of 112 ms (range 58-176

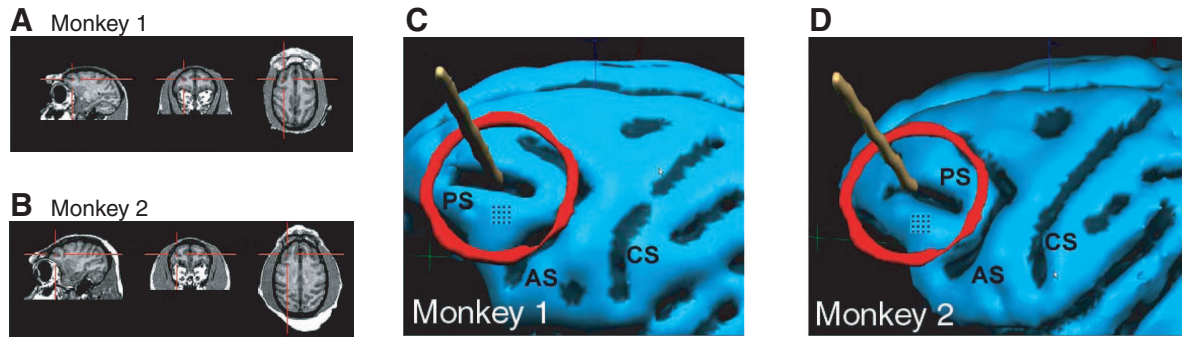


Figure 8.2.: **MRI-based reconstruction of recording positions.** (A,B) 3D-MRI datasets for Monkey 1 and 2 which were used to guide implantation of recording chambers. The red crosses point to the principle sulcus of the left hemisphere very close to the actual positions of the chamber centers. The coordinates refer to the anterior commissure. (C,D) Lateral view on surface reconstructions of the frontal cortex of both monkeys. The labels denote principle sulcus (PS), arcuate sulcus (AS) and central sulcus (CS). The 3D-positions of the chambers were measured in a stereotaxic frame relative to the ear bars and the bone above the center of the left eye. The brown rods represent the axes of the actual chamber positions which came to be at $x=-17$ $y=7$ $z=16$ for Monkey 1 and $x=-17$ $y=11$ $z=13$ for Monkey 2. These coordinates indicate displacements of maximally 2-3 mm in the anterior-posterior and dorso-ventral directions relative to the planned target positions. The red circles represent the walls of the recording cylinders which were placed into the skull at 45 in the frontal plain and 10 in the transversal plain. The 4x4 dot matrices illustrate the electrode-grids over ventral prefrontal cortex at the level of insertion through the surface of the cortex which was always in the ventral half of the chambers.

ms). The amplitude of evoked field potential responses to sample and test stimuli was larger in trials with correct behavioral responses (mean difference 10.4%, $p<0.01$). Object selectivity (ANOVA, 5% level) was observed for 14% of the sites. Induced and spontaneous oscillations in the frequency range between 12 and 95 Hz occurred during all epochs of the trial. Although the mean amplitude of oscillations was about three times higher in the range of 12-35 Hz than in the range of 35-95 Hz (Fig. 8.3C, Fig. 8.4), oscillatory activity in the high frequency range occurred reliably across sessions (Fig. 8.4). Grand-average time-frequency ('TF') spectra of all field potential recordings revealed that oscillatory activity (Fig. 8.3) was stable throughout all recording sessions. Except immediately after the evoked responses to sample and test stimuli and during early delay, oscillations dominated the LFP all the way through the trial. Synchronization of oscillations was assessed by computing PLVs ((Lachaux et al., 1999) see also Section 4.1.2). In the grand-average, the time course of phase-locking resembled that of power across all frequencies (Fig. 8.3).

8.3. Behavioral Performance

First we investigated performance effects. On average, the monkeys gave correct responses in 80% of the trials, ranging from 71 to 87% across sessions ($n=13$; 4124 trials). To test whether behavioral performance or the memorization (see Results in Sections 8.3.1, 8.3.2, 8.3.5) of specific stimulus information (see

8. Short-Term Memory Related LFP Oscillation in the Prefrontal Cortex

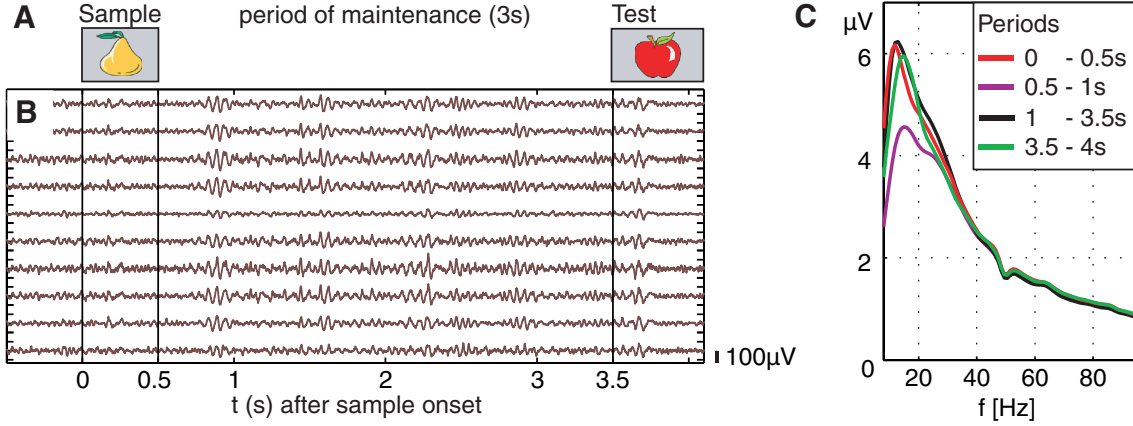


Figure 8.3.: **Oscillations and time-frequency responses in prefrontal cortex during a visual STM task.** (A) Visual memory task in case of a non-match condition. After sample stimulus presentation for 0.5 s, the information of the sample has to be maintained for 3 s and compared to the test stimulus. (B) Ten simultaneous field potential recordings from microelectrodes in ventral prefrontal cortex, scale bar indicates 100µV. (C) Average amplitude of oscillatory components during 4 different periods of the task.

Results in Sections 8.4, 8.5), or both (see Results in Section 8.5), were correlated with the strength and synchronization of prefrontal oscillations, we grouped our data into balanced sub-samples. This enabled us first, to compare neuronal activity recorded during trials in which the monkey gave a wrong response ('error trials') with activity recorded during trials with a correct response ('correct trials'), and second, to search in correct trials for stimulus-selective activity patterns related to the 16 or 20 different visual stimuli, used within individual sessions. To allow for unbiased estimation, we selected sets of trials (subsequently dubbed as trialsets) balanced in number for correct and incorrect performance. Performance amounted on average to 85%.

8.3.1. Performance Effects

In order to assess dynamical changes of oscillatory activity in LFP recordings, we estimated signal power (P) based on a sliding window FFT for all recording sites, experiments and monkeys individually corresponding to correct and incorrect behavior (200 ms length, 10 ms shift, signal tapered by a hamming window). Pairwise phase relations between recording sites were assessed by computing the PLV, which describes the stability of phase differences across trials among pairs of sites ((Lachaux et al., 1999), see also 4.1.2 and 4.1.3). A constant phase difference is described by $PLV=1$, while random phase differences express as $PLV=0$.

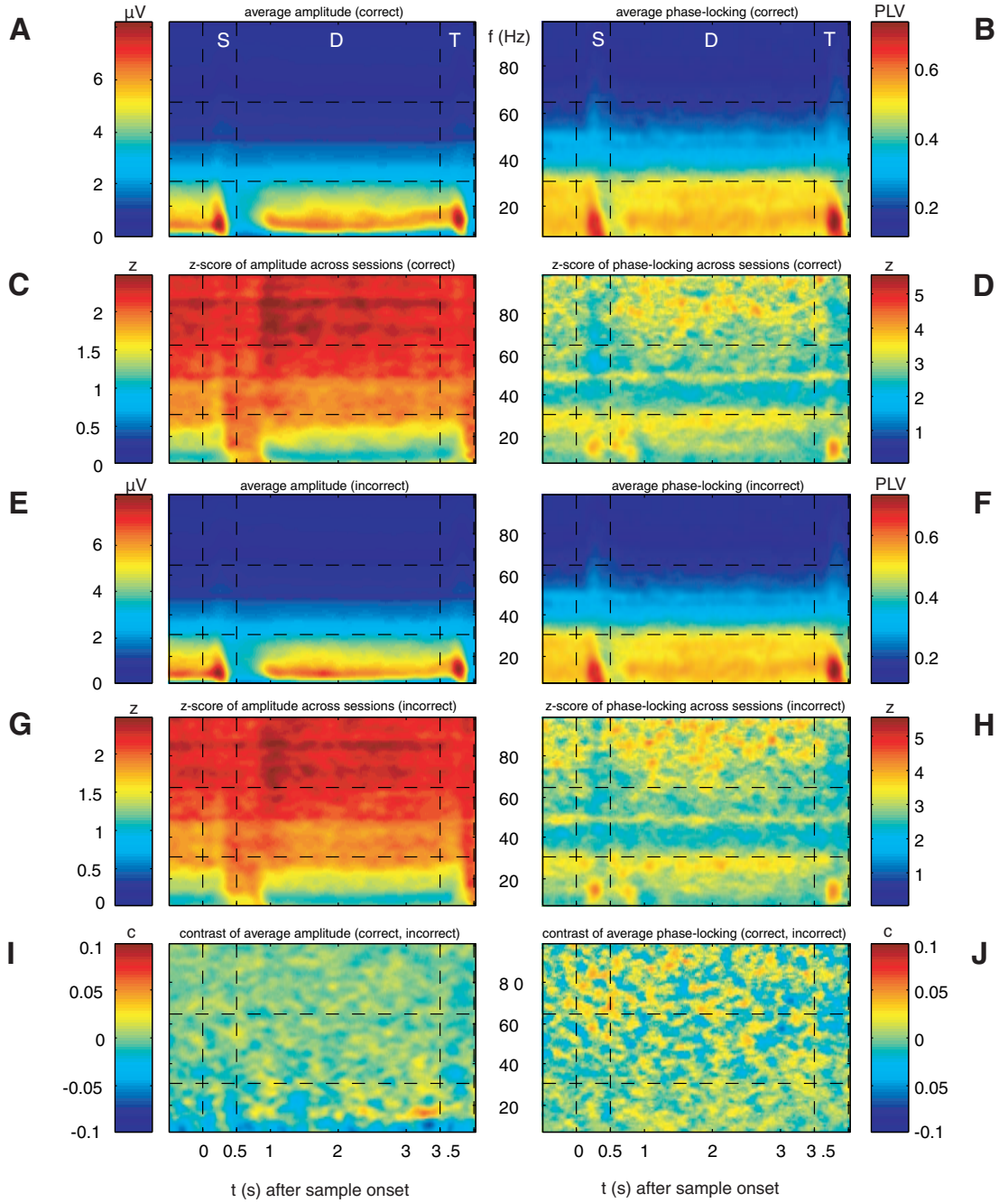


Figure 8.4.: **Grand average time-frequency ('TF') plots for power and phase-locking.** (A-J) TF plots of grand average power and phase-locking of all simultaneously recorded signals from two monkeys recorded in 12 sessions with a total of 4124 trials, from 146 sites and 921 pairs. Different task epochs are labeled with S (sample), D (delay), and T (test). (A,E) Grand average power for trials with correct and incorrect responses, respectively (C,G) z-score for grand average power in relation to variance (standard deviation = S.E.M.) across experiments for trials with correct and incorrect responses, respectively. (B,F) Grand average phase-locking for trials with correct and false responses, respectively (D,H) z-score for grand average phase-locking in relation to variance (S.E.M.) across experiments for trials with correct and incorrect responses, respectively. (I,J) Contrast= $((A-E)/(A+E))$ of grand average power and phase-locking as TF map. The maximum contrast of $c=0.08$ indicates a relative difference between trials with correct and false responses of 0.5%.

8.3.2. Performance Effects in Grand Average Power and Phase-Locking

We analyzed differences in the average power and phase-locking TF maps corresponding to correct and incorrect behavior. To this end we computed for each session the average power and phase-locking map for correct and incorrect performance. We used these maps to compute the average maps and the variability across sessions. Next we used both derived for each TF bin as a statistical measure of change by computing the z -score=(mean/std). Neither the average nor the stability across sessions revealed any performance-related differences. The maximum relative modulation across all frequency bands amounted on =0.5% (Fig. 8.4). We suspected that the variability across sites/pairs and sessions masked differences in power and phase-locking that might have occurred at individual sites and pairs. Therefore, we developed a method that permits assessment of differences for individual sites and pairs and than can be extended to the population of analyzed sites/pairs for each TF bin (see Section 5.2).

8.3.3. λ -Maps

To allow for unbiased estimations of performance-related differences we matched for each recording session correct and incorrect trials in terms of number of trials and temporal proximity, leading to subsets containing in total 2402 trials (7 Session/86 sites/507 pairs) for Monkey 1 and 1722 for Monkey 2 (5 Sessions/66 sites/414 pairs). To assess significant differences between trials with incorrect and correct responses, we used first, bootstrapping to estimate variability and stability, and second, permutation of trials between conditions to construct H_0 which predicts no performance-dependent difference (see Sections 5.2, 5.3). The bootstrapping procedure comprised a total of 16 samples. 12 of the 16 bootstrap samples, each constructed by a random selection of 75% of all trials within one condition (correct/incorrect), were used for assessing variability. The four remaining bootstrap samples were used for detection of slow trends throughout a session. They were each systematically lacking the first, second, third or last quarter of trials. This procedure will only retain performance-related changes that are reliable and stable over time. Differences between conditions are assessed in analogy to the Mann-Whitney U-test (see Section D.3.3) by a difference of rank sums (k) between the two sets of 16 bootstrap samples. The same bootstrapping approach was used on each of a set of 160 permutations of the trials of both conditions leading to a distribution of differences in rank sums k_0 under H_0 (H_0 : No difference between conditions). The p -value is estimated by computing $p_l(k \leq k_0)$ for a left sided and $p_r(k \geq k_0)$ for a right sided test. Based on this p -value (test level 1%) for each frequency bin and sliding window, we computed the percentage of sites and pairs of sites per session which showed an increase in power and PLV for correct ($f_c(P)$, $f_c(PL)$) and incorrect ($f_i(P)$, $f_i(PL)$) performance, respectively. To estimate the expected probability of sites/pairs with significant modulations in power/phase-locking, f_c and f_i are averaged across sessions. To allow for variability in the timing and frequencies of states or processes across sessions and subjects, TF maps of f_c and f_i were smoothed with a Gaussian kernel ($\sigma_t=150$ ms / $\sigma_f=5$ Hz). The parameter σ_t was selected to

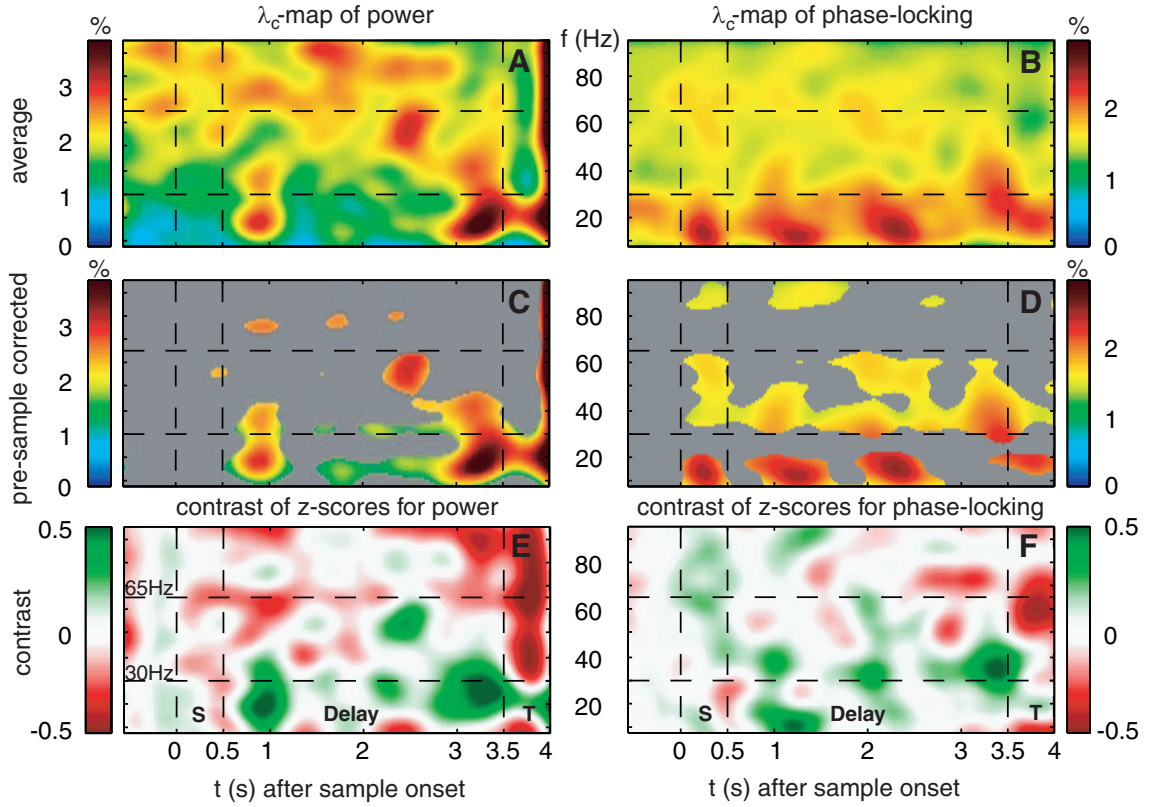


Figure 8.5.: **Performance-dependent increases of β - and γ -oscillations and their phase-locking in prefrontal field potentials.** (A,B) Time-frequency λ_c -maps provide the percentage of sites/pairs with significant increases of power/phase-locking in trials with correct responses. Abscissa and ordinate represent peri-stimulus time aligned to the onset of sample stimuli and signal frequency, respectively. Color-codes provide the percentage of sites/pairs with significant changes ('S', 'T': sample and test stimulus presentation, 'Delay': during delay the information of the sample has to be maintained). The left column provides maps of power and the right column maps of phase-locking. (A) Average of λ_c -map across both monkeys of power and (B) of phase-locking (based on 12 sessions, 4124 trials, 152 sites and 921 pairs of sites). (C,D) show increase of average data presented in A and B but corrected for the pre-sample period (-0.5 to 0 s) for each individual frequency, test level 1%, Bonferroni corrected for 450 sliding windows during sample, delay, and test. (E,F) Contrast of the z -scores of λ_c (correct) and λ_i (incorrect responses) with respect to a pre-sample period of 500 ms. Green indicates that λ_c exhibits larger increases in respect to the pre-sample epoch than λ_i , while red indicates the reverse. A contrast of $c=0.5$ indicates a difference in modulation by a factor of 3.

match approximately the observed latency variation of evoked responses during sample and test (range 58-176 ms), while σ_f is matching the uncertainty of the frequency revealed by a Fourier transform based on a 200 ms long rectangular sliding window. Smoothed TF maps of f are referred to as λ -maps. λ_c describes the percentage of sites with a significant increase in power for correct trials and λ_i an increase in incorrect trials. Note that an increase in λ_c does not imply a decrease in λ_i for incorrect performance. Two different subpopulations could still behave in an opposite fashion. One subpopulation could exhibit increased power for correct and the other for incorrect performance.

8.3.4. Modulation of λ -Maps

By referring to a λ -map one refers to the number of sites (pairs) with a significant increase or decrease in power (phase-locking) across sessions and monkeys. Thus, λ -maps reveal performance-dependent population dynamics in the TF domain. In the following, we describe increases of λ_c which identify TF zones reflecting increases of power (phase-locking) correlated to correct responses (Fig. 8.5A, inter-subject comparison in supplementary Fig. B.1). Intertrial intervals were fixed at 3 s in 11 of 13 recording sessions, and varied by less than half a second in the remaining two, so that the onset of the next sample stimulus was predictable for the monkeys with sub-second precision. Therefore and motivated by previous studies postulating or showing attention-related oscillatory cortical activity (Fries et al., 2001b; Niebur et al., 2002; Steinmetz et al., 2000), we investigated whether oscillatory patterns in the pre-sample period (-0.5 to 0 s) were correlated with performance. During the pre-sample period, $\lambda_c(P)$ exceeded the chance level by a factor of 2 at low (35-60 Hz), and a factor of 3 at high gamma-frequencies (60-70 and 80-95 Hz). Thus, enhanced oscillatory activity in the gamma-frequency range prior to trial onset was positively correlated to correct performance, most likely reflecting enhanced attention. In contrast, during sample presentation $\lambda_c(P)$ was increased by less than a factor of 2 for frequencies > 50 Hz. During the pre-stimulus interval and the sample response, $\lambda_c(P)$ for frequencies below 30 Hz did not exceed chance level. The response to test stimuli was associated with oscillations in the beta- and low-gamma frequency range (15-30 Hz). During the period of maintenance, the $\lambda_c(P)$ -map exhibits four prominent zones with an increased number of sites expressing enhanced power for correct responses. Two circumscribed zones simultaneously cover beta (14-27 Hz) and low gamma frequencies (30-45 Hz) in the early delay (0.6-1.2 s). A very similar pattern occurs directly before test stimulus onset (2.8-3.5 s). Note that the monkeys could anticipate the time of test stimulus presentation. Both zones extended over 500-650 ms. The third zone included mid gamma-frequencies (45-70 Hz) during the late delay (2.1-2.7 s) and preceded the second zone. In contrast to the previous three zones, the fourth zone covered the entire first 2.5 s of the delay at frequencies in the high gamma band (65-95 Hz).

8.3.5. Task-Specific Changes

In order to differentiate task-specific changes of power $\lambda_c(P)$ and phase-locking for the pairs of $\lambda_c(PL)$ on the population level from anticipatory attentional effects that seem to occur during the pre-sample-stimulus epoch, we performed a baseline-correction of $\lambda_c(P)$ and $\lambda_c(PL)$ for each frequency. To this end we computed baseline-corrected z -scores given by the modulation of $\lambda_c(P)$ with respect to its average and standard deviation during the 500 ms of the presample-sample-stimulus epoch. Then we applied a significance criterion of 1%, Bonferroni corrected for 450 comparisons (all sliding windows from 0 to 4.5 s). Frequencies are treated separately to consider frequency-band specific pre-sample differences like attention-related increases of certain oscillation frequencies (Fig. 8.5C,D). Despite the increase in $\lambda_c(P)$ for high frequencies during the pre-sample-stimulus epoch, three zones at 75-80 Hz still reached the significance threshold during early and middle delay. Note that there was no significant modulation of $\lambda_c(P)$ following the sample presentation. Otherwise, all prominent features of the $\lambda_c(P)$ -map survived the baseline correction and, that, cannot be explained by the preceding attentional state.

To compare the modulation λ_c and λ_i in relation to the baseline we contrasted the z -score for correct and incorrect performance, describing a percentage difference in z for correct with respect to incorrect trials (Fig. 8.5E,F). The contrast was for z -scores of λ_c , and λ_i with respect to the pre-stimulus baseline: $contrast = (z_c - z_i)/(z_c + z_i)$. A positive contrast indicates that the relative increase of λ_c with respect to the baseline is stronger than the relative increase of λ_i . Hence, a contrast of 0.5, observed in TF zones between 10 Hz and 65 Hz during the delay, shows that λ_c is modulated three times stronger than λ_i . Thus more sites (pairs) exhibiting significant differences among conditions expressed dynamical changes in the same direction.

8.3.6. Comparisons of the Modulation of the Grand Average and λ -Maps

Both, the performance dependent modulation of grand-average power and phase-locking, as well as the modulation of λ_c and λ_i -maps, are assessed by the contrast between the two conditions correct (c) versus incorrect (i) responses:

$$contrast = \frac{c - i}{c + i}$$

The contrast is ranging between -1 and 1. The contrast is a nonlinear transformation and the relative difference between c and f increases supra-linearly with increasing contrast. If we assume a modulation factor α scaling c and f with $c = \alpha * i$, the contrast can be expressed only by α :

$$contrast = \frac{\alpha - 1}{\alpha + 1}$$

And we can express α by the contrast:

$$\alpha = -\frac{\text{contrast} + 1}{\text{contrast} - 1}$$

Thus, the maximum contrast observed for grand-average power (phase-locking) and the maximum contrast of λ -maps, both describing the performance dependent modulation, can also be expressed by a modulation factor α .

For example: An $z=0.5$ leads to a modulation factor $\alpha=3$, meaning that the compared value (e.g. power, phase-locking or λ) had been 3 times, or 300% higher for correct than for incorrect responses. The maximum observed contrast in case of the grand-average power and phase-locking amounted to 0.07, which is equal to a modulation factor of 1.005. Note that the performance related modulation of λ -maps amounted to 0.5 which is equal to a modulation factor of 3. Thus, the performance-related modulation of λ -maps exceeds the modulation of the grand-average power and phase-locking by a factor of 298. This illustrates the advantage of the combination of bootstrapping and permutation tests, applied to each individual site (pairs) and TF bin, to identify performance-related or stimulus-selective modulations in power and phase-locking.

8.3.7. Phase-Locking - Spatial Extent of Synchronization

If oscillatory activity provides a temporal reference (Lisman and Idiart, 1995; Sommer and Wennekers, 2001) for distributed processes or drives subsequent circuits (Salinas and Sejnowski, 2001), it should be synchronized. We studied the precision and spatial extent of synchronization among recording sites by assessing phase-locking and its performance-related modulation. This measure describes the stability of phase differences in a given TF bin across trials. For frequencies below 35 Hz, grand-average-phase-locking was particularly strong (> 0.5), while for frequencies above 50 Hz, it was rather weak (< 0.2). This indicates that neuronal oscillations in the beta and low-gamma band are synchronized over a larger spatial extent than the average electrode spacing (400-900 μm), while neuronal oscillations above 50 Hz were much less well synchronized across recording sites. This suggests that faster oscillations are restricted to single or at most a few cortical columns (Pucak et al., 1996) while slower oscillations involve many columns and probably other cortical areas. This is supported by studies showing that beta- but not gamma-oscillations synchronize across remote cortical areas (Brovelli et al., 2004; Roelfsema et al., 1997). Performance modulated increases of phase-locking during trials with correct responses occurred in three TF zones with increased $\lambda_c(PL)$ (Fig. 8.5). First, during sample presentation a transient increase occurred in the beta band, which is remarkable, because the $\lambda_c(P)$ -map did not reveal a related change. Second, $\lambda_c(PL)$ exhibits a prolonged increase in the beta band with peaks at 900 ms and 1900 ms after the beginning of the delay, and third, as for power, there is a zone of enhanced phase-locking at beta and low gamma frequencies (15-50 Hz) that starts before and merges into the test response. Baseline correction of the $\lambda_c(PL)$ -map (Fig. 8.5B,D,F) uncovered in addition to the beta oscillations described

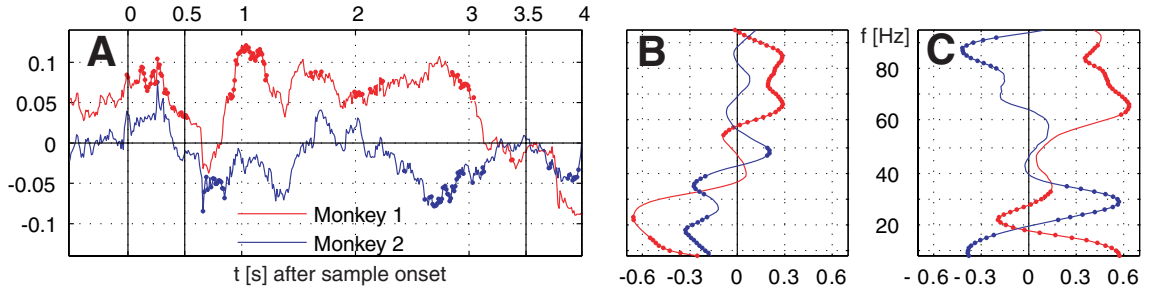


Figure 8.6.: **Control for induced power by eye-movements.** Time course of saccade probability and correlation analysis with main effects. (A) difference in the probability (correct performance - incorrect performance) of saccades per sliding window for Monkey 1 (red) and Monkey 2 (blue). Dotted segments identify periods which are significantly different (1% level) to zero. (B,C) Pearson correlation coefficients for Monkey 1 (red) and Monkey 2 (blue) between the probability difference of saccades per sliding window displayed in (A) and the percentage of sites (B) with increased power or pairs of sites with increased phase-locking (C) for correct performance.

above, a strong and long-lasting increase of phase-locking in the lower gamma band (30-65 Hz).

8.3.8. Controlling for Induced Power by Eye-Movements

To control for the influence of saccades that could have induced oscillatory activity, we computed for both monkeys the difference of the probabilities of saccades in correct and incorrect trials during each sliding window. This resulted in a time dependent function for each monkey which we could correlate (Pearson correlation coefficient, test level 1%) with λ_c -maps of power and phase-locking for each individual frequency (Fig.8.5). The Pearson correlation coefficient indicates if changes in the probability of saccades for correct and incorrect trials co-vary with the time course of λ -maps for individual frequencies. Because positive differences in the probability of saccades indicate an excess for correct performance, a positive Pearson coefficient indicates that the difference in saccades and the values of the lambda map are both either strengthened or weakened for correct responses. In contrast, a negative Pearson correlation coefficient indicates that less power is correlated with an increasing number of saccades and vice versa. The dynamics of the differences between trials with correct and false responses show that both monkeys exhibited an increased saccade probability during and after sample stimulus presentation in correct trials. While the traces for both monkeys differ significantly from each other during the delay, the performance-related difference is close to zero for both monkeys directly before test onset (Fig.8.6). Since correlations between the time-courses of the λ -maps and the difference in the probability of a saccade in correct and incorrect trials were either not significant, or not consistent across monkeys, an influence of performance-related differences in the saccade probability on power and phase-locking of individual sites is very unlikely. Thus performance-related changes in λ -maps of power and phase-locking cannot be explained by performance-related differences in the probability of saccades.

8.4. Stimulus Specific LFP Power Modulation

The results described so far do not differentiate among performance- and stimulus-related processes. We therefore assessed differences in power and related these to the presented stimuli (16 or 20 different memoranda). To this end bootstrapping and permutation tests were used in analogy to an ANOVA ($H_0: \mu_i = \mu_j$, see Section 5.4 and Appendix D.2.3) to determine the percentage $\lambda_s(P)$ of sites per session and TF bin which exhibited stimulus-selective changes in power. To assess stimulus selective differences in the power of oscillations (TF resolved), we balanced the number of trials for all presented stimuli per session by taking random subsets of the same size. To correct for sequence effects, which could lead to stimulus selective pre-sample activity, we rejected trials in which the sample stimulus was identical with the test-stimulus of the preceding trial. Stimulus sets consisting of 16 or 20 different stimuli had been used. Sessions which had been considered for the analysis of performance-related changes and had less than 20 trials per stimulus were discarded, leading to 4 sessions for Monkey 1 (in total 2280 trials, 56 sites) and 3 sessions for Monkey 2 (in total 1760 trials, 38 sites). Since phase-locking statistics require more trials than the estimation of average power to reach the same reliability of the estimation, the number of trials for individual objects had not been large enough to check for selective differences in phase-locking. As for the analysis of performance-related changes we used a bootstrapping procedure that yielded 16 bootstrap samples for each object to compute the mean power per frequency bin and sliding window. In analogy to an ANOVA we computed on these samples the F -value (see also Section 5.4 and Appendix D.2.3).

$$F = \hat{\sigma}_{pop}^2 / \hat{\sigma}_{error}^2 \quad (8.1)$$

The same bootstrapping approach was used on each set of 100 permutations of the trials of all conditions, leading to a distribution of F -values (F_0) under H_0 (all $\mu_i = \mu_j$). We performed a right sided significance test (test level 1%) for each frequency bin and sliding window and estimated the percentage $f_s(P)$ of sites per session with stimulus selective responses in power. To estimate the probability of stimulus selective modulations in power at individual sites $\lambda_s(P)$ we averaged $f_s(P)$ across sessions and smoothed the average with a Gaussian kernel ($\sigma_t=150$ ms, $\sigma_f=5$ Hz as used for performance-related modulation). Smoothed TF maps of f_s are referred to as λ_s -maps. $\lambda_s(P)$ allows for variability in the timing and frequencies of selective states or processing across sessions and subjects. An increase to baseline (-0.5 s to 0 s) was measured by the baseline corrected z -score, where average baseline selectivity and its standard error was estimated across all frequencies. This takes into consideration that stimulus-selective changes during baseline must be false positives irrespective of their frequency.

The resulting ' $\lambda_s(P)$ -maps' are shown in Figure 8.7. The power modulations observed in the pre-stimulus period, which only reached chance level, were used as baseline. Stimulus-selective increases in power as revealed by $\lambda_s(P)$ -maps were restricted to the delay period (Fig. 8.7B), although event-related

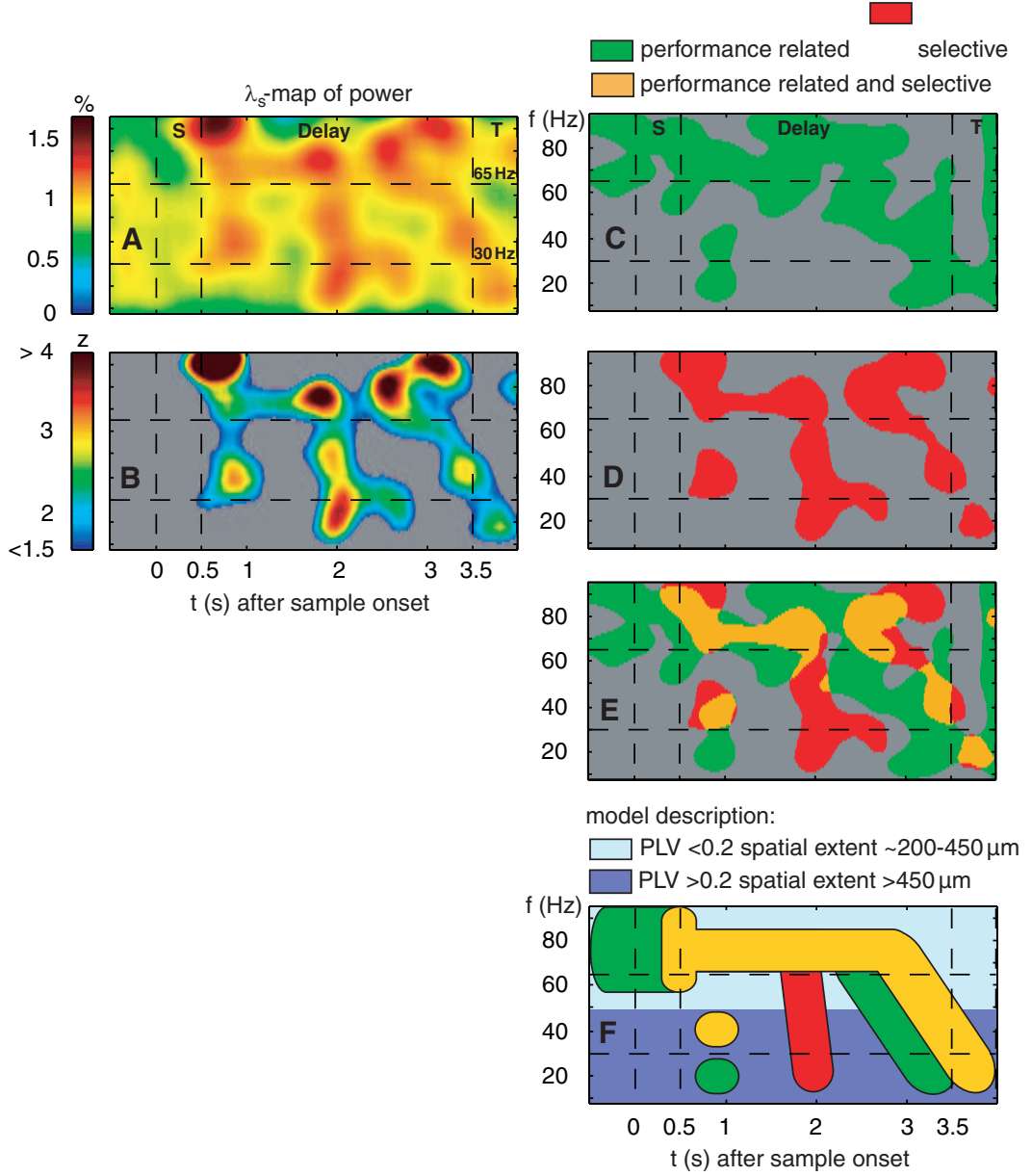


Figure 8.7.: **Stimulus-selective and performance-dependent increases of β - and γ -oscillations in prefrontal field potentials.** (A) Time-frequency ('TF') λ_s -map averaged across data from two monkeys (7 sessions, 4040 trials, and 94 sites). Abscissa and ordinate represent peri-stimulus time aligned to the onset of sample stimuli and signal frequency, respectively. Color codes provide the percentage of sites with significant differences in power across different stimuli ('S', 'T': sample and test stimulus presentation). (B) Modulations of the $\lambda_s(P)$ -map expressed as z -score in respect to baseline (-0.5 to 0 s). (C) Performance-related (λ_c -maps, green) and (D) stimulus-selective (λ_s -maps, red). The threshold used in (C,D,E) reflects 30% of the modulation between chance level and maximum λ_s during delay. (E) Direct comparison of performance-related (green) and stimulus-selective (red) modulation in power. Zones marked in orange represent periods during which significant increases of both performance and stimulus-related oscillations are present in the same frequency range. (F) Model description of TF zones with performance-related increase in power (green), object-selective power increase (red) and overlapping zones that exhibit both effects (orange). Light blue indicates grand average phase-locking value ('PLV') lower than 0.2, reflecting oscillatory processes confined to the average spacing between two electrodes. In contrast, dark blue indicates PLV higher than 0.2, which is compatible with spatially more extended processes that span distances over 450 μm .

potentials ('ERPs') revealed selectivity for the memorized stimuli already at much shorter latencies of 58-153 ms. This suggests that oscillatory activity in prefrontal cortex, which outlasts the initial encoding, contains stimulus-selective information.

8.5. Discussion

To distinguish between performance- and stimulus-related power changes, we compared $\lambda_c(P)$ and $\lambda_s(P)$ -maps. Because the dynamical range of $\lambda_c(P)$ and $\lambda_s(P)$ -maps was different, we considered 30% of the respective relative modulations of each map relative to chance level (Fig. 8.7C,D,E) as threshold for the comparison, which in case of $\lambda_s(P)$ -maps (Fig. 8.7D) corresponds to an increase of $z > 2$ with respect to baseline.

The low-gamma components (30-45 Hz) of the two transient TF zones that occurred during early and late delay were modulated both in a performance-related and stimulus-selective way, while the beta components showed only performance-related modulations. The sustained increases in the high-gamma band (65-95 Hz) during the first 2.5 s of the delay were stimulus-selective as well as performance-related. Among the memory-related signals described above, only the performance- and stimulus-related increase of oscillations in the high gamma band lasted throughout the entire memory delay. Because these high-frequency oscillations exhibited also stimulus-selectivity, the sustained modulation of high-frequency gamma oscillations might constitute an important part of the maintenance process. The low average phase-locking (< 0.2) of these high-gamma oscillations implies that the spatial extent of the processes based on high-frequency oscillations was small. However, the information about the memorized stimuli must be distributed across a large number of such local oscillating networks, because otherwise our arbitrary sampling of a few sites from a small patch of cortex would not have yielded that many sites with memory-related activity. The high phase-locking (> 0.5) of beta- and low-gamma oscillations across the extension of the electrode-grid, its persistence throughout the entire trial duration and the performance-dependence of this locking suggests that slower oscillations might provide a global temporal reference signal for the coordination of the distributed local processes that contain information about the stimuli. The power decrease that occurred across a broad range of frequencies shortly after sample presentation (Fig. 8.4A, E) appears incompatible with the possibility that oscillations are involved in stimulus encoding. However, a comparison between correct and incorrect trials revealed a reliable performance-related power increase at distinct sites. The transient episodes of beta- and low-gamma oscillations occurring at the beginning and the end of the maintenance interval were performance-dependent but only the gamma and not the beta-oscillations were stimulus-selective during these epochs. This might suggest that the transient beta/low-gamma-oscillations serve the organization of the memory process during transitions from encoding to maintenance and from maintenance to retrieval. Alternatively, these episodes with beta/low-gamma-oscillations could directly reflect encoding and retrieval processes.

We prefer the first interpretation, because these low-frequency oscillations were not stimulus-selective. A more likely correlate of encoding and retrieval processes are the high-gamma oscillations because they had shorter latencies and were stimulus-selective. As suggested by strong phase-locking of low-frequency oscillations across the whole electrode grid these oscillations appear to comprise larger networks. Because these persist during maintenance while stimulus-selective oscillations decrease their frequency in mid delay, these slower oscillations could reflect a rehearsal or recruitment process that involves representations in other cortical areas, such as visual area V4, for which locking of spikes to slow oscillations has been shown to occur during STM (Lee et al., 2005).

As summarized schematically in Fig.8.7F, performance-related TF zones overlap with stimulus-selective TF zones. This could reflect interactions of different sub-processes that need to be coordinated during the memory process. In particular, during early delay, two transient zones around 20 Hz and 40 Hz coincided with a sustained TF zone around 80 Hz, all reflecting increases of power in correct trials. The 80 Hz increase started earlier and coincided with stimulus encoding. Since power around 60 Hz was not modulated, the interaction of the 20 Hz and 80 Hz processes could be mediated by a 40 Hz process, as a harmonic of the 20 Hz and a sub-harmonic of the 80 Hz process. This suggests that the 40 Hz process could mediate a handshake between a global performance-dependent 20 Hz- and a local stimulus-selective 80 Hz process by means of a 1:2 and 2:1 synchronization. Such synchronization among harmonic frequencies was recently reported for human MEG (Palva et al., 2005) (See Fig. 3.3 for Arnold tongues, and Section 3.3 for $n : m$ synchronization). During the proposed period of handshaking, the bandwidth of the local high-frequency processes is sharpened in relation to the pre-sample and sample periods. Therefore it is conceivable that the power increase of the global and performance-dependent 20 Hz and 40 Hz oscillations serves to tune and coordinate the local stimulus-selective high-frequency processes. The persistence of the performance-dependent increase in phase-locking of the 20 Hz oscillations during the entire delay supports the idea that the global 20 Hz oscillations function as a temporal reference signal for the maintenance of the 1:4 synchronization with the local 80 Hz oscillations.

In conclusion, the organization of STM seems to involve networks of different sizes within and most probably beyond prefrontal cortex that oscillate at different frequencies. Small groups of neurons probably comprising a cortical column participate in gamma-oscillations around 80 Hz and their activity carries partial information about the memorized stimulus, while larger networks in and most likely beyond prefrontal cortex appear to be coordinated by coherent low-gamma- and beta-oscillations (14-45 Hz) that are correlated with performance but not the stimulus content. These global oscillations could coordinate the various stages of the memory process, provide a link to representations in other cortical areas containing more detailed information, and eventually organize executive functions. Global cortical circuits that organize behavior might employ short episodes of beta- and low-gamma-oscillations to embed local processes into a distributed maintenance network. Oscillatory brain activity and its synchronization

8. *Short-Term Memory Related LFP Oscillation in the Prefrontal Cortex*

across space and frequencies might serve as a mechanism for the integration of distributed signals over different temporal and spatial scales.

9. Short-Term Memory-Related High Complexity Joint Spike Activity

Recent findings, reporting that oscillatory components of the EEG and LFP are related to short-term memory ('STM') processes (see Section 8, and (Compte et al., 2000; Compte et al., 2003; Lisman and Idiart, 1995; Pesaran et al., 2002; Tallon-Baudry et al., 2004)), motivated us to analyze spiking data that was recorded in the prefrontal cortex of an awake monkey, performing a visual STM task on spike synchronization across groups of cells. In this chapter we are going to present results that raise the hypothesis, that the formation of neuronal assemblies is crucial for sub-processes of the STM like encoding, maintenance and decision making and/or motor planing. For further information about the task and findings concerning the oscillatory activity correlated to STM see Chapter 8 and the Figures 8.1 and 8.2.

9.1. Significant Joint Spike Pattern in Prefrontal Multi-Unit Activity

To investigate if joint spike events ('JSEs') are occurring by chance or more frequently than expected by chance in multi-unit activity ('MUA') that was recorded in the prefrontal cortex of an awake monkey performing an STM paradigm, we used the bootstrap unitary event ('BUE') method, a combined approach of the unitary event ('UE') analysis, bootstrapping and the n -jackknife. Details of this test are given in Sections 5.1 and 5.1.2. We used 10 bootstrap samples and 4 n -jackknife sub-samples, each lacking one quarter of the data. JSEs are evaluated as significant only if the $joint - p - value_{boot}$ (see Section 5.1.2), corresponding to the maximum of 14 joint- p -values derived from 14 bootstrap samples, was below the test level. The new test is a conservative significance test that detects joint spike ('JS') patterns that occurred reliably across the investigated trials (see for further discussion Section 5.1.2).

We applied the analysis to the data of 4 sessions from 2 monkeys (in total 5109 trials and 40 recording channels). For each session we chose all trials with correct responses (correct response of the monkey are successful trials with correct classification of the match or non-match condition, see Fig. 8.1) and derived for each JS pattern that occurred in any of the sliding windows (100 ms length) the expected and empirical number of JSEs as well as the corresponding joint- p -value. From Monkey 1 we analyzed 2 sessions: one with 11 recording channels and 1368 correct trials, and one with 10 recording channels and 1479 correct trials. The other two sessions from Monkey 2 contained 11 recording channels and 1179

9. Short-Term Memory-Related High Complexity Joint Spike Activity

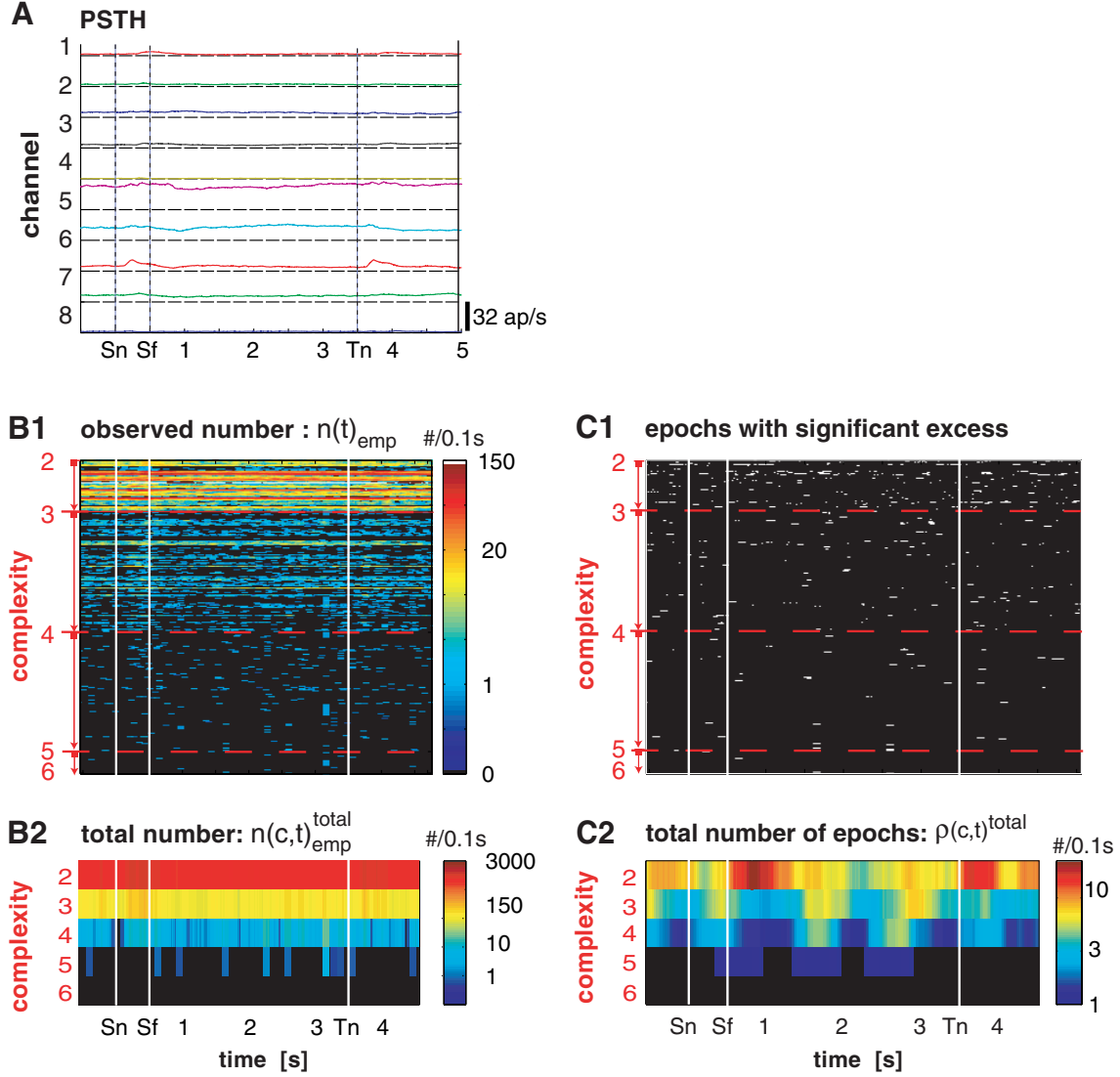


Figure 9.1.: **Introductory figure for the analysis of MUA during a visual short-term memory task.** (A) PSTH of 8 simultaneously recorded channels of multi-unit activity (MUA) during an STM task ('Sn'=sample onset, 'Sf'=sample off, 'Tn'=test onset). The scale bar indicates 32 spikes per second. (B1) Observed empirical frequency $n(t)_{emp}$ per sliding window and JS pattern. JS patterns are ordered along the y-axis according to their complexity. The frequency is color coded and given in units of number per 100 ms. (B2) Total frequency $n(c,t)_{emp}^{total}$ derived for all JS patterns of the same complexity c . (C1) Epochs with increased frequency of JSEs are marked with white squares. (C2) Frequency $p(c,t)$ of JS patterns per complexity c that have significantly more JSEs than expected by chance (in units of number per 100 ms).

correct trials as well as 9 recording channels and 1083 correct trials. Since the PSTHs from all 4 analyzed experiments revealed fast modulations of the spike rate after the onset of the sample- and test-stimulus, the UE analysis was applied in short windows of 100 ms length (Fig. 9.1A, and supplementary Fig. G.3). For each of the occurred JS patterns we derived the observed frequency n_{emp} as a function of time that is displayed color coded and ordered by the pattern complexity in Figure 9.1B1. Based on the $joint - p - value_{boot}$ (for more details see Section 5.1.2) we identified epoches for which JSEs occurred significantly more often than by chance and marked them with white squares ordered by their complexity (Figure 9.1C1).

9.2. Rate of Significant Joint Spike Patterns

Next we derived the total number of occurred JSEs per complexity and sliding window $n_{emp}^{total}(c, t)$ given by the sum of JSEs that occurred in all JS patterns of complexity c . Since $n_{emp}^{total}(c, t)$ is evaluated per sliding window, it is given in units of number per 100 ms ($[1/0.1s]$).

$$n_{emp}^{total}(c, t) = \sum_k n(t)_{emp}^{v^k} \text{ for all } v^k \text{ that have complexity } c_k = c \text{ with } c_k = \sum_{i=1}^N v_i^k \quad (9.1)$$

By analogy with the total frequency of occurred pattern per complexity $n_{emp}^{total}(c, t)$ we derived the total frequency of JS patterns with complexity c that have a significant excess of JSEs $\rho(c, t)^{total}$.

9.2.1. Average Rate of Significant Joint Spike Pattern ($\bar{\lambda}(c, t)$)

We derived $\rho(c, t)^{total}$ for each of the four sessions from Monkey 1 and 2. To compare $\rho(c, t)^{total}$ across sessions that included different numbers of recorded MUA channels we normalized $\rho(c, t)^{total}$ from each session by the number of potentially existing patterns per complexity (with N =number of channels, c =complexity).

$$\lambda(c, t) = \frac{\rho(c, t)^{total}}{\Omega} \text{ with } \Omega = \binom{N}{c} \quad (9.2)$$

E.g. the normalization for session one of Monkey 1 that included 11 channels would be 55 for complexity 2, 165 for complexity 3 and 330 for complexity 4. The advantage of the normalized frequency $\lambda(c, t)$ is that it is independent of the number of recording channels. Thus we can average $\lambda(c, t)$ across the four sessions of the two monkeys and derive the average $\bar{\lambda}(c, t)$ rate of significant JS patterns per complexity c and as a function of time (Fig. 9.2A).

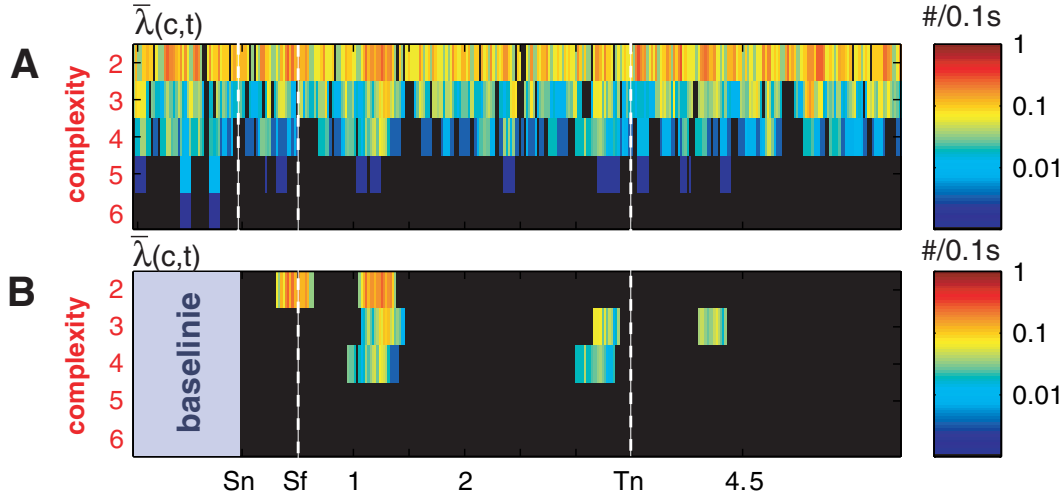


Figure 9.2.: **Modulations and task-related modulations of $\bar{\lambda}(c, t)$ of JS patterns during a short-term memory task.** (A) frequency $\bar{\lambda}(c, t)$ of JS pattern per complexity c that have significantly more JSEs than expected by chance ('Sn'=sample onset, 'Sf'= sample off, 'Tn'=test onset).(B) Significant modulations of the frequency $\bar{\lambda}(c, t)$ in respect to the pre-sample period starting 1 s before and ending at sample onset (Sn). Non-significant modulation of $\bar{\lambda}(c, t)$ are masked (test level 5%). The frequency $\bar{\lambda}(c, t)$ is given in units of number per 100 ms.

9.2.2. Task-Related Modulations of $\bar{\lambda}(c, t)$

To identify task-related modulations we baseline corrected $\bar{\lambda}(c, t)$. To this end we computed the mean and the standard deviation of $\bar{\lambda}(c, t)$ across sliding windows in the interval between sample onset and sample onset minus one second, to derive a z -score that describes the strength of the modulation of $\bar{\lambda}(c, t)$ during the trial in respect to the average pre-sample value $\bar{\lambda}(c, t)$ and its variability.

$$z(c, t)_{\bar{\lambda}} = \frac{\bar{\lambda}(c, t) - \text{mean}(\bar{\lambda}(c, t))}{\text{std}(\bar{\lambda}(c, t))} \quad (9.3)$$

We used $z(c, t)_{\bar{\lambda}}$ to identify significant increases (test level 5%) of $\bar{\lambda}(c, t)$ in relation to baseline and highlighted those by applying a threshold in order to display only significant modulations (Fig. 9.2B).

9.2.3. Results

We observed JS patterns with a significant increase of the frequency of JSEs for complexities ranging from 2 to 6. To identify task-related modulations we used baseline corrected values of $\bar{\lambda}(c, t)$. As illustrated in Figure 9.2B the frequency of JS patterns with significantly increased frequencies of JSEs $\bar{\lambda}(c, t)$ is task modulated for complexities two, three and four. Remarkably, these task-related increases occur at crucial periods of the STM task:

1. **Late sample presentation:** During the late sample presentation, from 250 ms after sample onset until 150 ms after sample off, the frequency $\bar{\lambda}(c, t)$ was significantly modulated in respect to

9.3. Performance-Related Modulation of the Strength of Synchronization ($\lambda_c(c)$ and $\lambda_i(c)$)

baseline for complexity two JS patterns. This is remarkable, since the fast increases in the spike rates occurred earlier, already at 85-120 ms after sample onset.

2. **Early delay:** During the early delay, starting at 420 ms and ending at 900 ms after sample off (corresponding to 920 ms till 1400 ms after sample onset), the frequency $\bar{\lambda}(c, t)$ was significantly modulated in respect to baseline for complexity two, three and four JS patterns.
3. **Late delay:** Directly before test stimulus onset, from 3 to 3.5 s after sample on, the frequency $\bar{\lambda}(c, t)$ was significantly modulated in respect to baseline for complexity three and four JS patterns.
4. **Late test presentation:** During a period that started 800 ms after the test onset, the frequency $\bar{\lambda}(c, t)$ was significantly modulated in respect to baseline for complexity three JS patterns.

These findings allow us to draw four conclusions: First, the BUE method (100 ms sliding window length) is able to correct for the fast changes of the rate during the sample period, since during the period, that starts with the raising phase of the rate at 85-120 ms and ends at the onset of the increase of $\bar{\lambda}(c, t)$ at 250 ms after sample onset, no significant modulation of $\bar{\lambda}(c, t)$ was observed. Second, the dynamics of the JS patterns seem not to be correlated to the dynamics of the rate. Third, the epoch during and directly after sample presentation, that might be linked to encoding of stimulus information, is characterized by task-related modulations of $\bar{\lambda}(c, t)$ with lower complexity of the JS patterns ($c = 2$) than during the delay periods that no stimulus was presented for. This is remarkable since the activity in the prefrontal cortex is likely to be more driven during stimulus presentation than during the delay period. Thus, this finding could indicate that the non-driven and intrinsic states during the delay are more complex in terms of the size of formed neuronal assemblies. Fourth, the second and third increase during the delay are tightly linked to the beginning and the end of the delay. This indicates that JS activity could be involved in the encoding and rehearsal of information in pre-frontal cortex.

9.3. Performance-Related Modulation of the Strength of Synchronization ($\lambda_c(c)$ and $\lambda_i(c)$)

So far we have described task-related modulations of the frequency of JS patterns each having a significant excess of the corresponding JSEs by $\bar{\lambda}(c, t)$ and $z(c, t)_{\bar{\lambda}}$. Nevertheless, both $\bar{\lambda}(c, t)$ and $z(c, t)_{\bar{\lambda}}$ cannot clarify if the observed JS patterns and the underlying processes have any importance for the monkey's task or if they are just induced by processes that are either intrinsic or induced by the stimulus but both not relevant for the STM task. Thus we developed a test based on the bivariate resampling presented in Chapter 5.2 to test, if the synchronization of JS patterns is modulated by the monkeys performance.

This allows us to study if correlations between modulations of the strength of synchronization and the processes that are necessary to encode, maintain and recall information in an STM paradigm (see

9. Short-Term Memory-Related High Complexity Joint Spike Activity

Section 9.1). The bivariate test uses bootstrap estimates of the joint- p -value to estimate its variance. Since the joint- p -value describes the likelihood to observe an empirical number of JSEs that occurred in the data, given the expected number that is estimated based on the rates of the respective neurons, the joint- p -value can be interpreted as a rate normalized measurement of the synchronization of the respective neuron. Thus, for the purpose of simplification we refer to the joint- p -value as normalized synchrony, and used it as an estimator of the strength of synchronization. A joint- p -value equivalent to the normalized synchrony of 0.5 indicated that the normalized synchrony is just at chance level, while small normalized synchrony values indicate a strong synchronization. If the normalized synchrony is smaller in condition one than in condition two, that means that the strength of the synchronization in condition one is larger than in condition two. Thus, the bi-variate analysis of the joint- p -values allows us to compare trials with correct and incorrect performance of the monkey, in order to infer if the strength of the synchronization underlying the individual JS pattern is correlated to performance (success versus failure of the monkey).

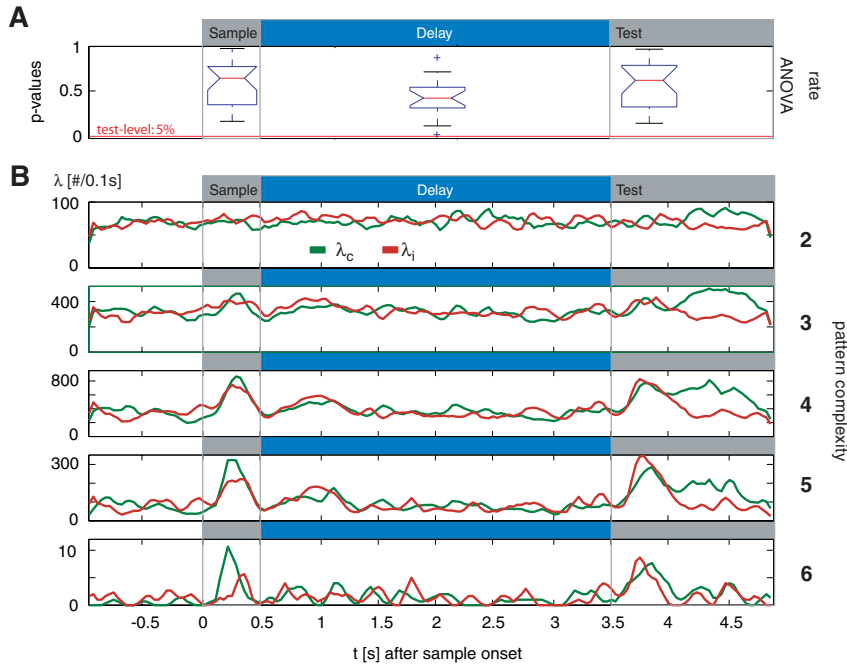


Figure 9.3.: **Task and performance-related modulations of $\lambda_c(c)$ and $\lambda_i(c)$ of JS patterns during a short-term memory task.** (A) Resulting p -values of an ANOVA test on performance-related modulations of trial based spike counts during the Sample, Delay, and Test period. Any of the p -values of any of the 27 simultaneously recorded SUA yielded significant performance-related modulations of spike counts on a 5% significance level. (B) Frequency of JS patterns with significant and performance-related modulations of the synchronization derived from the same dataset as shown in (A). Green lines represent the frequency $\lambda_c(c)$ of JS patterns per complexity that had a significantly higher normalized synchrony for correct trials than for incorrect trials, while red represents the frequency $\lambda_i(c)$ of JS patterns per complexity that had a significantly lower normalized synchrony for correct trials than for incorrect trials. In case $\lambda_c(c)$ is larger than $\lambda_i(c)$, the frequency $\lambda_c(c)$ of JS patterns, with a higher normalized synchrony for correct behavior, is higher than the frequency $\lambda_i(c)$ of JS patterns, with a higher normalized synchrony for incorrect behavior.

9.3. Performance-Related Modulation of the Strength of Synchronization ($\lambda_c(c)$ and $\lambda_i(c)$)

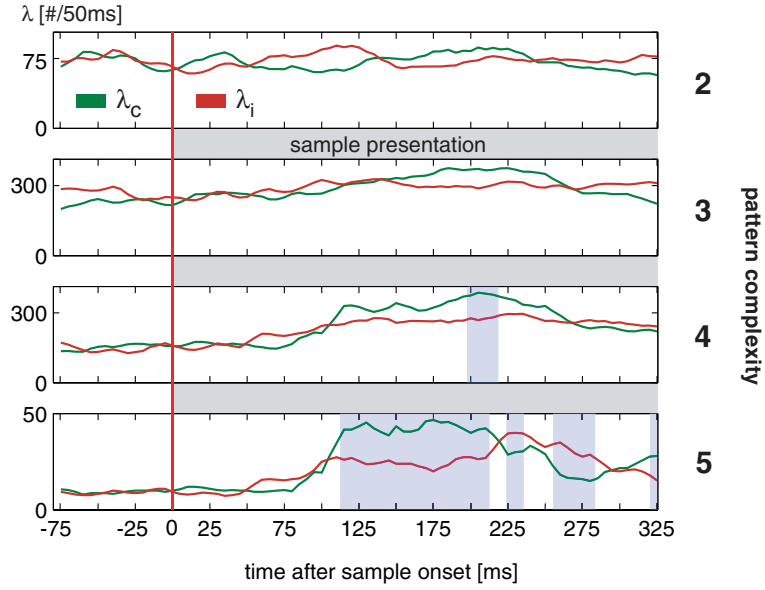


Figure 9.4.: **Task and performance-related modulations of $\lambda_c(c)$ and $\lambda_i(c)$ of JS patterns during the sample presentation of a short-term memory task.** The same as in Figure 9.3 but enlarged for the period around sample presentation (from -75 ms before and 325 ms after sample onset) and based on shorter sliding windows of 50 ms instead of 110 ms as used in Figure 9.3. Blue shadowed areas indicate a significant difference between $\lambda_c(c)$ and $\lambda_i(c)$ in respect to the baseline difference and its standard deviation (test level 5% and Bonferroni corrected for 325 multiple comparisons)

Since the joint- p -value is derived by a nonlinear transformation of the empirical and expected number of JSEs, it is important not to use the mean value, but rather the median to compare the bootstrap estimates $\Theta_{b,c}^*$ across the conditions $c = 1, 2$ (see for further discussion Section 5.2). Thus, first we use the median of $\Theta_{b,c}^*$ to describe the difference in the normalized synchrony between condition one and two, and second, a permutation test to assign a statistical significance to the difference for each individual JS pattern (H_0 : no performance-related difference in the normalized synchrony - joint- p -value, see Section 5.1.2). Since the test is applied to each individual JS pattern, we would have to apply a multiple comparison correction in case we were interested in the statistical significance from individual JS patterns (see for further discussion in 'False Positive Rates', Section 6.1.9).

To avoid the multiple comparison correction and to describe population dynamics we derive the frequency of JS patterns that have an increased normalized synchrony for correct performance ($\lambda_c(c)$: c ='correct'). As for $\lambda_c(c)$ we compute the frequency ($\lambda_i(c)$: i ='incorrect') of the JS patterns that have an increased normalized synchrony if the monkey fails to perform the task correctly. Note, that a decreased normalized synchrony in failure trials is the same as an increased one in successful trials, but the frequency of distinct JS patterns $\lambda_c(c)$ and $\lambda_i(c)$, that express either strengthening or weakening of the normalized synchrony for correct trials, can be different. To identify significant modulations of $\lambda_c(c)$ and $\lambda_i(c)$ we baseline corrected both and identified periods of $\lambda_c(c)$ and $\lambda_i(c)$ that were significantly increased in relation to baseline (baseline corresponds to the pre-sample period from -1 s to 0 s before

9. Short-Term Memory-Related High Complexity Joint Spike Activity

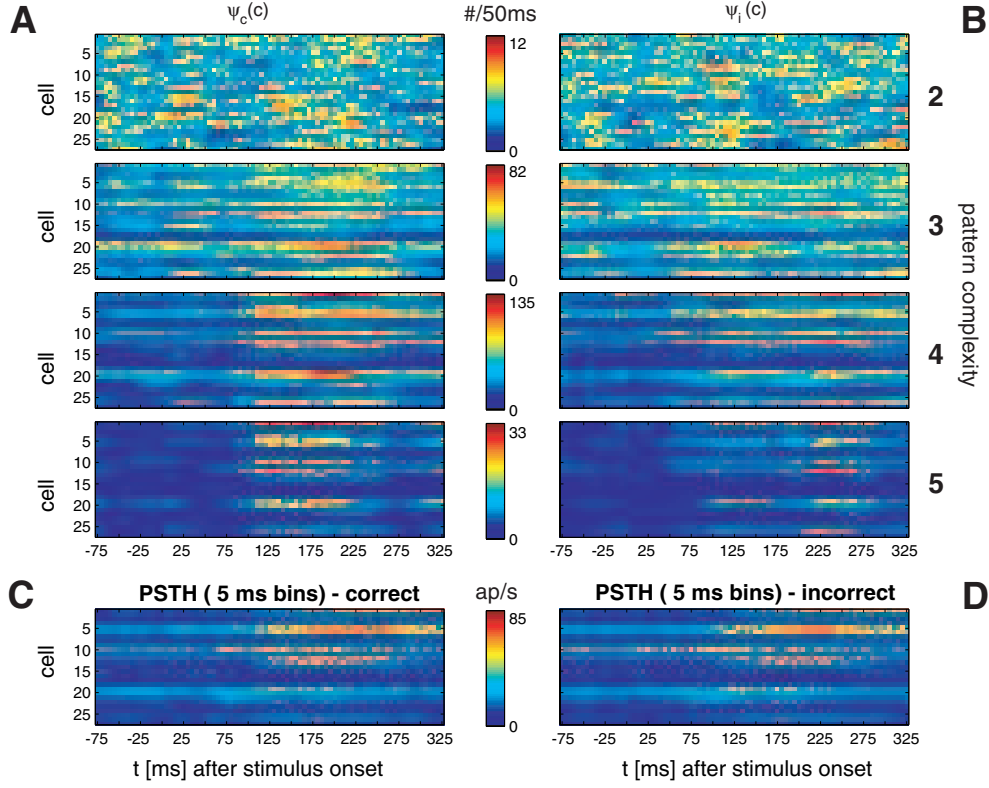


Figure 9.5.: **Neuronal groups with performance dependent modulations of synchrony ($\psi(c)_c$ and $\psi(i)_c$).** (A,B) Frequency of participation of cell $i = 1 \dots 27$ in JS patterns of complexity c that expressed significant and performance-related modulations of the normalized synchrony. (A) Frequency $\psi(c)_c$ of participation in JS patterns with increased normalized synchrony for correct trials. (B) Frequency $\psi(c)_i$ of participation in JS patterns with increased normalized synchrony for incorrect trials. (C,D) PSTH color coded for cell $i = 1 \dots 27$ for trials with correct performance (A) and incorrect performance (B).

sample onset, test level 5%, for detailed procedure see previous Section 9.2.2). Latter allows to identify performance-related modulations that are also task-related.

9.3.1. Performance Dependent Formation of Neuronal Assemblies ($\psi(c)_c$ and $\psi(i)_c$)

So far we have lost the identities of the neurons that participated in JS patterns that expressed performance-related modulations of the normalized synchrony. Thus we next identify the individual cells and how often they participated in JSEs of a certain complexity that expressed performance-related modulations of the normalized synchrony. To this end we just counted how often each cell was involved in all patterns that contributed to $\lambda_c(c)$ and $\lambda_i(c)$ of complexity c , leading to $\psi(c)_c$ and $\psi(i)_c$.

To identify significant modulations of $\psi(c)_c$ and $\psi(i)_c$ in respect to the pre-sample period we baseline corrected both and identified periods of $\psi(c)_c$ and $\psi(i)_c$ that were significantly increased (5%, for detailed procedure see previous Section 9.2.2). As described in the previous section, latter allows to identify

9.3. Performance-Related Modulation of the Strength of Synchronization ($\lambda_c(c)$ and $\lambda_i(c)$)

performance-related modulations that are also task-related, but not related to pre-sample effects like attentional modulations.

9.3.2. Results

We applied the analysis described in the previous section to derive $\lambda_c(c)$, $\lambda_i(c)$, $\psi_c(c)$ and $\psi_i(c)$ to one session from Monkey 2 (session C002), for that we have sorted the spiking activity and selected 27 single-units that were stable and well classified throughout the whole experiment (for further information about the sorting see Appendix G.1). As described in Section 5.2 we balanced the datasets for the two conditions (correct versus incorrect) in terms of number of trials and temporal proximity, leading to subsets each containing 226 trials. First we detected all JS patterns that occurred throughout the whole trial based on a sliding window UE analysis (100 ms window length). We detected in total 14314 distinct JS patterns in correct trials and 13771 in incorrect trials. To ensure that we did not bias latter statistics of $\lambda_c(c)$, $\lambda_i(c)$, $\psi_c(c)$ and $\psi_i(c)$, we rejected all JS patterns that didn't occur in both conditions, leading to 10024 JS patterns ('common') that occurred in correct as well as in incorrect trials. To each of the 10024 common JS patterns we applied the bivariate analysis based on a combination of bootstrapping, n -jackknife and permutation test to derive $\lambda_c(c)$, $\lambda_i(c)$, $\psi_c(c)$ and $\psi_i(c)$ per complexity. In Figure 9.3 $\lambda_c(c)$ and $\lambda_i(c)$ are illustrated for the whole trial that lasts 6 s (the trial starts at -1 s before sample onset and ends 2 s after test stimulus onset).

It is remarkable, that differences between $\lambda_c(c)$ and $\lambda_i(c)$ occurred only during the sample and the test-period. Especially during the delay period we would have expected modulations in case, that assemblies would have been involved in processes that are necessary to perform the task. Nevertheless, that does not necessarily indicate that assemblies are not involved in theses processes underlying the STM of the monkey. One problem of the analysis that might have made it impossible to detect modulations in $\lambda_c(c)$ and $\lambda_i(c)$ is the short window of the UE analysis. On the one hand we need this short window to be able to correct for rate modulations, on the other hand we require very precise reoccurrence of JSEs in time and across trials since JS patterns must occur in the same window to be statistically detectable.

Since we find modulations of $\lambda_c(c)$ and $\lambda_i(c)$ during the sample and test presentation, it might be that the external stimulus triggers processes that modulate the synchronization of neurons and makes occurrences of JS patterns locked to the stimulus with a precision higher than 100 ms and thus makes induced JSEs detectable by relatively short sliding windows. In contrast, the absence of modulations of $\lambda_c(c)$ and $\lambda_i(c)$ could either indicate that task-related strengthening or weakening of synchronization is indeed not existing during the delay, or, that the intrinsic activity is less precisely locked to task or behavior than 100 ms, and therefore not statistically detectable with short sliding windows. This unresolved question motivates to utilize the NeuroXidence method, since this method does not require short windows to correct for rates, rather it can be adapted to the underlying timescale of the processes

9. Short-Term Memory-Related High Complexity Joint Spike Activity

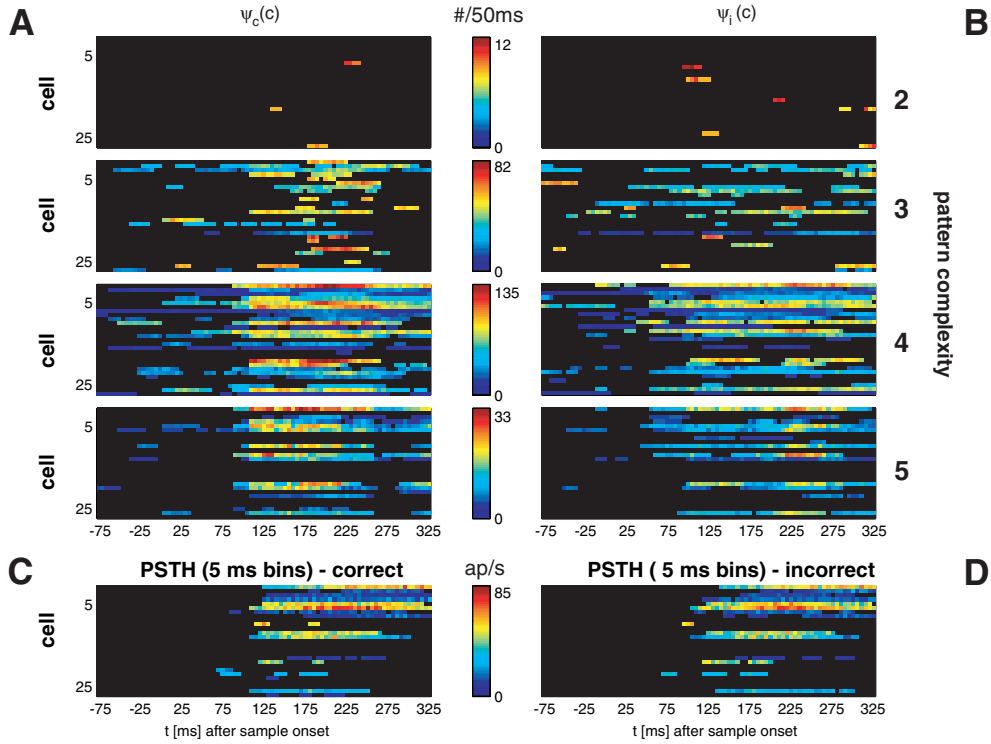


Figure 9.6.: **Neuronal groups with performance dependent and task-related modulations of synchronization** ($\psi(c)_c$ and $\psi(i)_c$). (A,B,C,D) Same as Figure 9.5A,B but modulations that are not significantly modulated in relation to baseline are masked (5% test level).

that might be in between of 0.1 to 1 s. Sliding windows of this length would allow JS patterns to be loosely locked to the task or the monkey's behavior.

Modulations that occur during the sample period are illustrated in Sections 9.4, 9.5 and 9.6 while Figures 9.7 and 9.8 show results for the test presentation period. Note that Figures 9.4, 9.5 and 9.6 present results for the period of the sample presentation that are based on shorter sliding windows (50 ms) than used so far. Motivated by results that indicated modulations of $\lambda_c(c)$ and $\lambda_i(c)$, but were based on 100 ms long windows, we shortened the windows in order to enable an even better rate correction and thus to reduce the chance of false positive modulations of the normalized synchrony. To identify significant differences between $\lambda_c(c)$ and $\lambda_i(c)$ we derived the difference $\Delta\lambda(c) = \lambda_c(c) - \lambda_i(c)$. Next we computed the mean value and the standard deviation of $\Delta\lambda(c)$ during the baseline (-1s till sample stimulus onset) and baseline corrected $\Delta\lambda(c)$ and expressed $\Delta\lambda(c)$ in units of standard deviation leading to the z -score $Z_{\Delta\lambda(c)}$. Based on a 5% significance level that was Bonferroni corrected for 325 multiple comparisons ($\alpha_{corrected} = \alpha_{test}/325$, because 325 windows were tested from 0 s till 325 ms after sample onset) only complexity four and five expressed significant differences of $\lambda_c(c)$ and $\lambda_i(c)$ during the sample presentation (Fig. 9.4 periods with significant differences in the modulation of $\lambda_c(c)$ and $\lambda_i(c)$ are indicated by blue shadowed areas). Differences of complexity three and two JS pattern were at chance

9.3. Performance-Related Modulation of the Strength of Synchronization ($\lambda_c(c)$ and $\lambda_i(c)$)

level. It is remarkable that significant differences between $\lambda_c(c)$ and $\lambda_i(c)$ occurred at about the same time as the latency of rate modulations after sample onset increases were observed. Nevertheless the spike counts were not performance modulated during the same period from 120 to 275 ms (see Fig. 9.3A for statistical evaluation and Fig. 9.5C,D for PSTH, and Fig. 9.6C,D for baseline corrected PSTH). Still, even though we used a very short sliding window of 50 ms, we have to validate that the presented findings were based on fine-temporal cross-structure rather than on rate modulation that remained undetected because of an insufficient sensitivity of the spike count significance test. In section 9.5 we are going to present this validation based on a jittering approach that is used to destroy effects that were based on fine-temporal cross-structure intentionally.

To identify the members of the JS patterns that participated in these significant and performance-related modulations of the normalized synchrony, we derived $\psi_c(c)$ and $\psi_i(c)$. $\psi_c(c)$ is the frequency with which the respective neuron participated in JS patterns that had a significantly increased normalized synchrony in case of correct compared to incorrect trials. Consequently, $\psi_i(c)$ is the frequency with which the respective neuron participated in JS patterns that had a significantly increased normalized synchrony during incorrect compared to correct trials. Fig. 9.5A,B illustrates $\psi_c(c)$ and $\psi_i(c)$ for individual JS patterns with complexities ranging from 2 to 5 while Fig. 9.6A,B is baseline corrected by the same procedure as in the previous sections (baseline -1 s to sample onset; test level 1%, but no Bonferroni correction for multiple comparisons). While the PSTHs for correct and incorrect trials are nearly identical (Fig. 9.5C,D, Fig. 9.6C,D) the time courses for $\psi_c(c)$ and $\psi_i(c)$ are remarkably different for higher complexities of 4 and 5. While $\psi_c(c)$ increases abruptly at about 90-100 ms after sample onset for complexity 4 and 5, this change isn't present in $\psi_i(c)$. Thus, this abrupt increase of $\psi_c(c)$, that seems to occur slightly ahead of the spike rate increase, is equivalent to an increased probability of neurons to participate in JS patterns that have a stronger normalized synchrony in correct than in incorrect trials. This indicates that the correct behavior of the monkey is tightly linked to the formation of groups of neurons that are synchronized. Since the spiking data was recorded during an STM task in the prefrontal cortex, and the reported observation occurred during the sample presentation, it is likely that this process that seems to be based on the formation of assemblies, is either encoding or transmitting information that was presented. Nevertheless, only if we show in addition that these performance-related modulations of the normalized synchrony are also informative about the nature of the stimulus would we have strong evidence that the assembly formation is representing information rather than it is involved in processes that are necessary to perform the task but are not informative about the stimulus material.

So far we haven't been able to perform this required multivariate test since the trial number per stimulus was in the order of 15 to 80. This number, even though it seems to be large, is not sufficient to analyze JSEs of higher order since spike rates in data from prefrontal cortex are of the order of a few spikes per second and thus JSEs are pretty rare. Thus, a new experiment with reduced number of stimuli

9. Short-Term Memory-Related High Complexity Joint Spike Activity

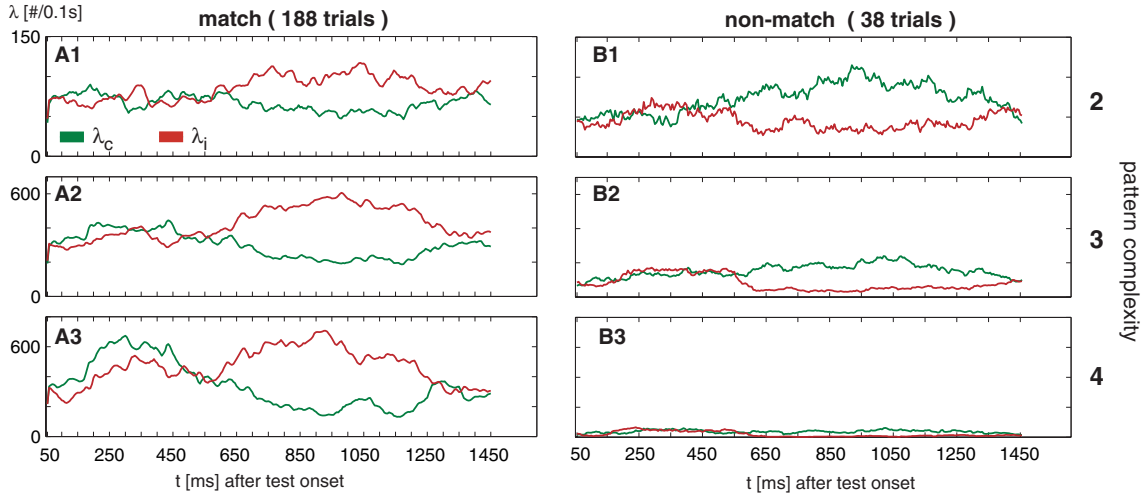


Figure 9.7.: **Task and performance-related modulations of $\lambda_c(c)$ and $\lambda_i(c)$ of JS patterns during test-stimulus presentation of a short-term memory task.** (B) Frequency $\lambda_c(c)$ and $\lambda_i(c)$. The same as in Figure 9.3 but for the period during test-stimulus presentation.

(e.g. 4 different stimuli instead of 16 to 20 per session) would have to be conducted to clarify the role of the identified assemblies during the stimulus presentation period.

9.4. Motor-Related Modulations of the Strength of Synchronization

The second period that contained significant modulations of $\lambda_c(c)$ and $\lambda_i(c)$ was the period during the test stimulus presentation. Since the task consists of a match and non-match condition we have to differentiate between correct and incorrect behavior in case of match and non-match. Thus we have two pairs of two different conditions. Therefore, we computed beside the performance-related modulations described by $\lambda_c(c)$ and $\lambda_i(c)$, modulations in respect to the motor response of the monkey $\lambda_r(c)$ (right hand movement) and $\lambda_l(c)$ (left hand movement). We balanced the number of trials, and used 188 correct and 188 incorrect trials for the match condition and 38 correct and 38 incorrect trials for the non-match condition.

9.4.1. Results

Remarkably, during the first 250 ms after test stimulus onset more JS patterns had higher normalized synchrony for correct than incorrect performances across all three complexities 2, 3 and 4 and across the two conditions match and non-match. Since the match and non-match analysis were performed on independent trial sets this is a strong cross-validation of the results. After this first period the development of the time courses of $\lambda_c(c)$ and $\lambda_i(c)$ become mirrored for the match and non-match condition. While $\lambda_i(c)$ increased and $\lambda_c(c)$ decreased in the match condition, it was the opposite in the non-match condition.

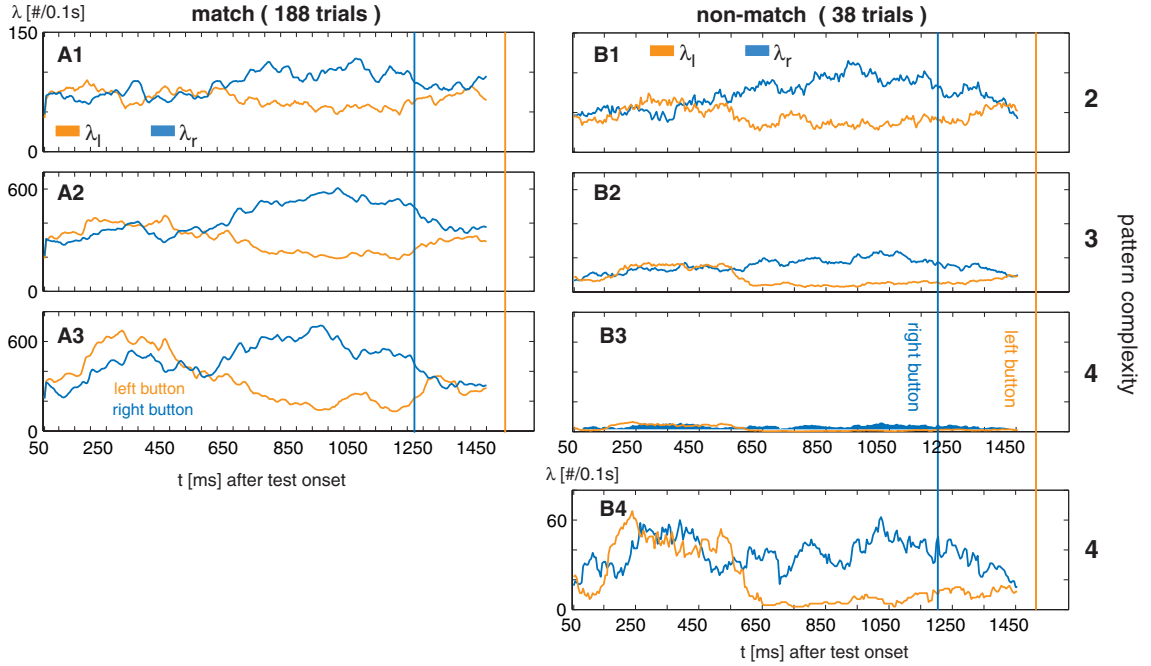


Figure 9.8.: **Task and motor-related modulations of $\lambda_r(c)$ and $\lambda_l(c)$ of JS patterns during test-stimulus presentation of a short-term memory task.** (A,B) Frequency per complexity of JS patterns with significant and motor-related modulations of the number of JSEs derived for the same dataset as shown in (A). Orange represents the frequency $\lambda_l(c)$ of JS patterns per complexity that had a significantly higher normalized synchrony for trials with left hand movements to press the left button than for right hand movements to press the right button. Blue represents the frequency $\lambda_r(c)$ of JS patterns per complexity that had a significantly higher normalized synchrony for trials with right hand movements to press the right button than for left hand movements to press the left button. In case $\lambda_r(c)$ is larger than $\lambda_l(c)$, more patterns occurred that were more synchronous in case of right hand movements than with left hand movements. (B4) The same as (B3) but a different scale on the y-axis to amplify the modulations that are smaller than in (A3). Vertical orange and blue lines indicate the mean response time for left (orange) and right (blue) button press.

This and the fact that this difference increased strongly 600 ms after the test stimulus onset motivated us to interpret the results from the perspective of the motor response.

Figure 9.8 shows $\lambda_r(c)$ and $\lambda_l(c)$ corresponding to the number of JS patterns with an increased normalized synchrony for right ($\lambda_r(c)$) and left handed ($\lambda_l(c)$) movements. Even though the absolute range and the sign of the modulation of $\lambda_l(c)$ and $\lambda_r(c)$ were different they were modulated consistently for both conditions, match and non-match, and complexities 2, 3 and 4. Remarkably, the difference between the modulations of $\lambda_l(c)$ and $\lambda_r(c)$ disappeared at about the same time as the average response time of the monkey. In contrast to task-related modulations that were observed during the sample period for higher complexities 4 and 5, these motor-related modulations occurred across low complexity 2 and higher complexity 3 and 4 JS patterns.

Thus, the combination of the motor-related modulations of $\lambda_l(c)$ and $\lambda_r(c)$, the cross validation based on two independent trial sets and the tight locking of the increase to the motor response from 600 ms

after test onset till the average response time, indicate that the formation of assemblies is involved in the motor activity. Since it takes, first, about 600 ms after test stimulus onset and about 400 ms after the first performance-related response to establish this strong motor-related modulation, and second, the recordings were performed in the prefrontal cortex that has been associated with memory and decision making so far, it is likely that the observed modulations of $\lambda_l(c)$ and $\lambda_r(c)$ are signaling motor commands based on the evaluation of the maintained stimulus material during delay and the stimulus material presented during the test period.

9.5. Confirmation of Fine-Temporal Cross-Structure by Jittering of Spikes

To support the findings from the previous sections, we tested, if fine-temporal cross-structures were the property that have led to the presented results in order to exclude other kinds of violation of the null hypothesis of the UE analysis (H_0 : cells are independent and can be described by stationary Bernoulli processes, for further discussion see Section 4.2.3 and Chapter 6). To this end we destroyed fine-temporal cross-structure by dithering of spikes with plus minus 6 ms. Since dithering destroys most of the fine-temporal cross-structures but rate modulations stay rather constant (see Discussion in Section 6.1.3), it is a robust procedure to demonstrate the importance of fine-temporal cross-structure on the presented results. Thus, we expect the results to break down after dithering, if the observed effects would be based on fine-temporal cross-structure on a time scale of less than 12 ms. And indeed, after dithering the same datasets as used for the investigation of performance-related differences (see Section 9.3.2), the majority of the results broke down. To demonstrate this we again analyzed 228 correct and incorrect trials for the same 27 single-units and derived $\lambda_c(c)$ and $\lambda_i(c)$ based on 100 ms long sliding windows (same length as used for the analysis of delay and test period, but twice as long as used for the analysis of the sample period). The increased sliding window from 50 to 100 ms length for the analysis of the sample period increased the danger of false positives due to violations of the stationarity assumption of the UE analysis. It is remarkable that the results based on the 100 ms long window are qualitatively pretty close to the 50 ms version, even though the total number of JS patterns that expressed significant modulations is higher in case of the 50 ms window. After applying the dithering to the spikes of plus minus 6 ms, the total number of JS patterns that expressed significant modulations of the normalized synchrony, went down to 20% for complexity 2, 12% for complexity 3, 5% for complexity 4, 2% for complexity 5 and 0.3% for complexity 6. That means that 99.7 percent of the JS patterns that had significant modulations of the normalized synchrony before dithering, became non-modulated after dithering for complexity 6 JS patterns. This finding strongly indicates that, indeed, the fine-temporal cross-structure on a time scale of less than 12 ms has caused the above reported effects and results.

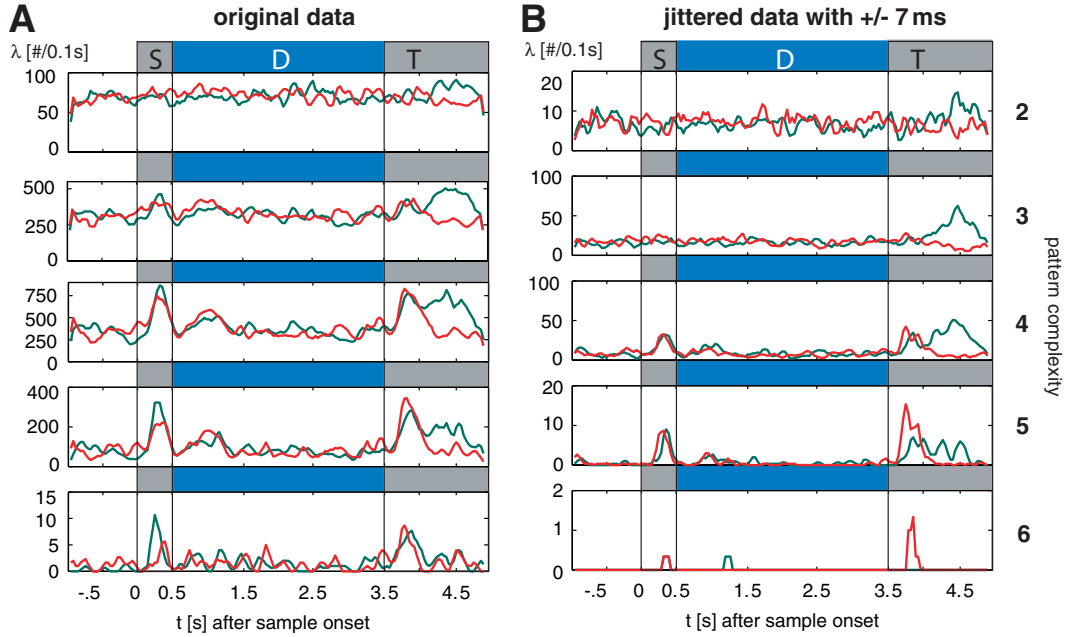


Figure 9.9.: **Task and performance-related modulations of $\lambda_c(c)$ and $\lambda_i(c)$ of JS patterns during a short-term memory task.** (A,B) Frequency of JS patterns with significant and performance-related modulations of the synchronization derived for the same dataset as shown in (A). (A) Results based on the original data and (A,B) results based on data that was derived from the original but individual spikes were jittered (dithered) with ± 7 ms. Green lines represent the frequency $\lambda_c(c)$ of JS pattern per complexity that had a significantly higher normalized synchrony for correct trials than for incorrect trials, while red represents the frequency $\lambda_i(c)$ of JS pattern per complexity that had a significantly lower normalized synchrony for correct trials than for incorrect trials. In case $\lambda_c(c)$ is larger $\lambda_i(c)$, the frequency $\lambda_c(c)$ of JS patterns, with a higher normalized synchrony for correct behavior, is higher than the frequency $\lambda_i(c)$ of JS pattern, with a higher normalized synchrony for incorrect behavior

9.6. Discussion

We applied three different analysis steps to spiking data that were simultaneously recorded from up to 27 single-units in the prefrontal cortex of an awake monkey, performing an STM paradigm. The results of the first analysis, based on the frequency of JS patterns with significantly increased number of JSEs, indicate the existence of task-related formations of neuronal assemblies. Especially during the late sample period as well as in the early and the late delay, an increase of the frequency of significant JS patterns was observed. Nevertheless, the presented findings, even though they are strong, seem to be plausible and are based on about 5000 trials, cannot be generalized, since the statistics are based on only 4 sessions from 2 monkeys. To generalize these findings we would have to have a couple of more sessions with similar results. Nevertheless, we demonstrated that the strategy chosen to analyze the data and the developed tools are performing well.

In a second step we tested the importance of JS activity for the monkeys performance. We selected one session and sorted the data to get single-unit activities from 27 cells. This large number of cells

9. *Short-Term Memory-Related High Complexity Joint Spike Activity*

allowed us to gain reliable statistics for even higher complexity JS patterns. This is because the expected probability is exponentially falling with increasing complexity, but at the same time the number of potentially existing number of JS patterns of the same complexity is factorially increasing. Thus a large number of simultaneously recorded neurons increases the robustness of statistics of higher complexity JS patterns, since the high number of different patterns of the same complexity compensates partially for the low expected probability for individual JS patterns. Based on the bivariate analysis of the activity of these 27 cells we have found JS activity that was performance-related and indicated stronger synchronization if the monkey performed the task correctly. The tight linking of these performance-related modulations of the synchronization to the monkey's task indicates that neuronal assemblies are formed task-related and are necessary for the monkey for a successful classification of the stimuli as match or non-match. If these assemblies are carrying information about the stimulus that is presented, or, if they are engaged in other processes necessary to encode and maintain information in the STM, is unclear so far, and should be tested with a test design proposed in Section 5.4.

In a third step we tested, based on the same 27 single-units, to which degree neuronal assemblies are related to the motor actions of the monkey (see Figure 9.4). During the first period of the test presentation (150 to 550 ms after test onset) modulations in the number of JS patterns with significant modulations of the normalized synchrony seem to be performance-related rather than motor-related. After this first period, during which the maintained stimulus from the sample has to be compared to the test-stimulus to enable a correct response of the monkey, the modulations in the number of JS patterns with significant modulations of the normalized synchrony become motor-related. This motivates the hypothesis that JS activity in the prefrontal cortex is related to decision making rather than to stimulus maintenance.

Still, since we used only one experiment, the value of argumentation might be rather limited since findings cannot be generalized due to few independent samples. Nevertheless we can raise the hypothesis that the formation of neuronal assemblies and the strengthening or weakening of synchronization is correlated to the periods of sample presentation, the early and late delay, the test stimulus presentation and the motor action as well as to the monkey's performance.

Part IV.

Discussion and Conclusions

10. Discussion

The neuronal coding strategy in the cortex has been very controversially discussed in the last years. The discussion is polarized by two extremes. The first is the 'grand mother' or single-cell coding hypothesis that assumes that neuronal information is encoded, processed, or maintained by single-neurons. The second, the assembly hypothesis assumes a synergistic and cooperative neuronal code formed by large groups of cells (*single-cell coding*: (Bialek et al., 1991; Bialek and Rieke, 1992; Baker and Lemon, 2000; Brody, 1999; Oram et al., 1999; Shadlen and Newsome, 1994; Shadlen and Newsome, 1998) *cooperative neuronal coding*: (Abeles et al., 1993; Aertsen et al., 1989; Arieli et al., 1995; Castelo-Branco et al., 2000; Eckhorn et al., 1988; Engel et al., 1991; Fries et al., 1997; Gray and Singer, 1989; Kreiter and Singer, 1996; von der Malsburg, 1981; Markram and Tsodyks, 1996; Prut et al., 1998; Riehle et al., 1997; Singer, 1999; Volgushev et al., 1998) *discussions of important aspects concerning both hypotheses*: (Bernander et al., 1994; Bohte et al., 2000; Fabre-Thorpe et al., 1998; Fetzer, 1997; Gerstner et al., 1997; Hopfield, 1995; Lestienne and Tuckwell, 1998; MacLeod et al., 1998; Softky and Koch, 1993; Softky, 1994)). Remarkably, evidence for each of the two hypotheses has been accumulating rapidly in the last years and decades. This alone was reason enough to cause disbelief in studies and tools that were used to demonstrate the validity and importance of either of the two hypotheses.

Thus, the strong polarization in the neuroscience field requires tools and techniques that allow for conclusive tests to draw a line of demarcation. This requires, first, precisely defined and, if possible, simple working hypotheses, and second, analysis tools that are well focused and free from assumptions. This motivated us to devote one part of this thesis to the development of non-parametric tools for the analysis of neuronal oscillations, neuronal synchronization, and Joint Spike ('JS') activity, while allowing for the required conclusive tests. In the second part, we demonstrate the existence of task and behavior-related synchronization of neuronal activity that give strong evidence for the assembly hypothesis by applying the new tools to data recorded simultaneously with multiple electrodes in awake monkeys and anaesthetized cats.

10.1. Discussion: New Tools to Investigate Joint Spike Activity

The single-cell or rate hypothesis assumes that information is encoded in the spiking of individual cells. In contrast, the assembly hypothesis assumes that coordinated firing of cell assemblies with millisecond

10. Discussion

precision serves the coding and binding of information (see Chapter 3). Thus, the working hypotheses for a statistical test that is intended to test the assembly hypothesis are the following:

Working hypothesis 1: *Fine-temporal cross-structure exists on a millisecond timescale, which is represented by Joint-Spike Events ('JSE') in simultaneously-recorded spiking activity from multiple neurons, and its frequency cannot be explained by chance.*

Working hypothesis 2: *Synchronization of spiking activity across groups of cells is correlated to the underlying neuronal processes.*

The statistical hypothesis test is crucially dependent on the statistical formulation of the chance level of JSEs under the null hypothesis (H_0 : rate hypothesis). To allow for a line of demarcation and to enable a clear dissociation of the assembly hypothesis from the rate hypothesis, the H_0 has to consider all features of the data that might be causally related to the single cell or rate hypothesis. Only then is a rejection of H_0 equivalent to support for the assembly hypothesis and evidence for cooperative coding based on JS activity (see Sections 4.2.2 and 6.1.2). We developed two new methods to approach this goal. The first method ('Bootstrap Unitary Event', see Section 5.1.2) is an improved version of the Unitary Event ('UE') method (Grün et al., 2002a; Grün et al., 2002b) and overcomes limitations of the standard UE method that arise due to a lacking assessment of the variability of the expected and observed number of JSE across trials. The second method, 'NeuroXidence' (see Chapter 6), is a new approach that, in order to formulate H_0 , allows for the consideration of all features of the data that might be causally related to the rate hypothesis. Thus, NeuroXidence is a tool that is well focused, free from assumptions, and tests the precisely-defined and simple working hypothesis that JSEs occur more or less often than expected by chance.

10.1.1. Bootstrap Unitary Event Method

The *Bootstrap Unitary Event* ('BUE') method, (see Section 5.1.2) was designed to tackle two limitations of the standard UE method (Grün et al., 2002a; Grün et al., 2002b): an increased rate of false positives in the case of low rates (Roy et al., 2000) and the lacking stability and reliability check of JS patterns. To this end, the BUE method estimates the reliability, variability, and stability of the difference between the expected and occurred frequency of JSEs. It does so by employing two different resampling techniques. The first is bootstrapping to estimate variability and reliability, and the second is the n -jackknife to estimate stability. The new BUE method is a conservative hypothesis test that accounts for the rareness of spiking and JSEs. In addition, it checks on the stability and variability, and it prevents significant events originating from differences between the expected and the empirical number of JSEs, which are clustered in a few trials or in a certain period of the experiment. Therefore, the number of JSEs has to be significantly different from chance level across all bootstrap samples.

Based on this principle idea of the BUE method, we developed an extension that allows one to estimate if two or N different conditions modulate the joint- p -value of the UE analysis (bi-variate, multi-variate, see Sections 5.2 and 5.4). The first basic concept of the extension is to estimate by a combination of bootstrapping and an n -jackknife the variability and stability of the difference between the expected and occurred frequency of JSEs. The second is to derive an H_0 distribution of the expected differences of the test statistics based on a permutation test. The latter has the disadvantage that it is computationally demanding, but the strong advantage is that it is the most robust way to construct a distribution of H_0 (Efron and Tibishirani, 1993). This new bi-variate and multi-variate BUE method allows one, for the first time, to analyze if, how, and to what degree different experimental conditions modulate the synchronization of spiking activity, which is stable and reliable across trials for more than two neurons.

Still, since the BUE method is based on the principal concept of the standard UE method that formulates that H_0 be based on the assumption of independent and stationary Bernoulli processes, the susceptibility to false rejections based on the violation of the Bernoulli assumption, rather than on the fine-temporal cross structure, persisted. This motivated us to develop a new tool, named NeuroXidence, that uses a different concept for defining H_0 .

10.1.2. NeuroXidence

NeuroXidence implements a statistical hypothesis test that allows for a clear dissociation of the single-cell rate and the assembly hypothesis (see Chapter 6). In contrast to other methods (Abeles and Gerstein, 1988; Grün et al., 1999; Grün et al., 2002a; Grün et al., 2002b; König, 1994), NeuroXidence is a non-parametric statistical hypothesis test and does not assume any model about the spike-train generating process. It is suitable for the analysis of spiking activity from large groups with about a hundred simultaneously-recorded neurons. New concepts for the detection of JS activity as well as the statistical evaluation make NeuroXidence computationally very efficient. The estimation of the statistical significance of JS patterns is based on a comparison of the frequency of JSEs between the recorded dataset and a surrogate dataset, which is derived from the recorded dataset but lacking any coordinated firing. NeuroXidence uses two concepts to generate surrogate data.

The first is based on trial shuffling and is intended to separate induced from evoked activity (see Section 6.1.4). Even though this concept is appealing, it is difficult to realize in a statically-robust way, since it assumes that the process underlying the observed spiking activity is ergodic. The latter implies the assumption that the underlying processes are identical across trials. Since data recorded in real neuronal systems often show a high degree of variability and strong systematic changes, the main assumption of ergodicity is not fulfilled. Although one can partly correct for these changes (Ventura et al., 2005) in the case of certain test statistics, such as the UE analysis (Grün et al., 2003), the correction itself requires again the estimation of parameters, such as the spike rates. Consequently, the correction itself makes

10. Discussion

assumptions, e.g. of the model of the spike train generating process, to estimate rate. As for any method that requires crucial assumptions about the data that cannot be completely fulfilled, trial shuffling and results revealed by it remain questions at issue.

The second concept, which generates surrogate data to implement H_0 , utilizes random shifting of spike trains to destroy fine temporal cross-structure, while keeping the complete auto-structure in the surrogate dataset intact (see Section 6.1.3). Therefore, using this concept to generate surrogate data, NeuroXidence estimates the chance frequency of JSEs under H_0 , by allowing for the consideration of rate modulations or for history dependencies, which might exist in each individual spike train. In comparison to trial shuffling to generate surrogate data, the random shifting of spike trains to destroy fine temporal cross-structure is the more general and robust method. It also does not need any assumption about the data, which allows the operator to choose the length of the analysis window freely. However, since it implements a comparison of two timescales, it is limited in its capabilities, like the UE method, to dissociate evoked from induced activity, if both have about the same millisecond precision. Since the latter is only the case for very few classes of stimulation, such as flashed visual stimuli, the limitation is rather more theoretical than of practical relevance. In conclusion, since the concept to generate surrogate data based on spike train shifting is statistically more robust than the alternative based on trial shuffling, the latter should be used as a complementary procedure to the first to disentangle induced and evoked JSE only if the precision of the stimulus locking of the evoked activity is of the order of the precision of JSE.

The new statistical evaluation of NeuroXidence relies on the trial-based difference between the expected and occurred number of JSEs in combination with a non-parametric statistical hypothesis test (see Section 6.1.5). This allows NeuroXidence to differentiate between rare and spurious events and events that are reliably reoccurring. Only if events are reliably reoccurring across trials, might they be classified as JS patterns with a significant excess or deficiency of JSEs. Thus, NeuroXidence is able to estimate the statistical significance reliably, even for JS patterns of complexities higher than 2, which are expected to be rare.

NeuroXidence can be applied to short time windows to track transient neuronal states. In contrast to other methods, the NeuroXidence window length is not restricted by any stationary assumption of the data, since the full auto-structure, including any rate modulations, is considered by the statistical hypothesis test. Thus, the length of the NeuroXidence analysis window can be freely chosen by the operator to match the assumed duration of neuronal states of interest. This makes NeuroXidence predestined for the analysis of spontaneous or induced activity, in which reoccurring JS patterns might be only loosely locked and, therefore, might require large analysis windows to be detected.

10.2. Discussion: Results Revealed by New Tools for Analyzing JS Activity

We applied NeuroXidence, the BUE method (see Section 5.1.2), the bootstrap power, and the coherence estimation (see Sections 5.1.1, 5.2, 5.4) to data recorded in awake monkeys and an anaesthetized cat.

10.2.1. High-Complexity JS Patterns Revealed by NeuroXidence

We applied NeuroXidence to real data recorded from 48 single units in visual area 17 in an anaesthetized cat. Our results demonstrate that JSEs with a precision of 5 ms occurred significantly more often than expected by chance, for complexity 2 to 7 JS patterns (see Chapter 7). Since the frequency of JS patterns with excess JSEs was increased only during the period of stimulus, our results indicate that stimulus encoding involves fine-temporal cross-structure induced by neuronal correlations.

We validated our result by analyzing a jittered version of the original data (see Section 7.1). This validation was equivalent to a comparison of H_0 against H_0 and served as the estimation of the chance frequency of JS patterns with excess JSEs. Since only 0.2% of JS patterns with excess JSEs remained significant after jittering, we conclude that the increase of the frequency of JS patterns with excess of JSEs is highly significant for complexities ranging between 2 to 7. The validation also revealed that fine-temporal cross-structure on a timescale of 5 ms is present in simultaneous recordings from cat visual area 17.

The dynamics of cells that exhibit significant joint activity across different complexities can be separated from the various rate modulations observed for different cells in the data. These rate modulations concerned either the strength of the responses or the relative phase shift of the responses because of the different times the stimulus was crossing the cells' receptive fields. This highlights the strength of NeuroXidence for investigating, in terms of fine-temporal structure, the complex interactions of neurons that might coexist with rate modulations of individual neurons. In addition, the time-course of the total number of JS patterns detected (See Fig. 7.1A in Chapter 6) was considerably different from the time-course of the number of JS patterns that exceeded the chance level. Both together strongly indicate that the dynamics of the neuronal coupling, which is reflected by the number of groups of cells that exhibited a significant increase of JSEs, are mostly independent of the drive given by the external stimulus.

Since NeuroXidence was based on surrogate data generated by spike train shifting rather than trial shuffling, we have to discuss if the increased frequency of significant JS patterns is due to induced or evoked activity. The temporal frequency of the drifting sinusoid was 1.2 s and fits pretty well the interval between the two mean peaks of the number of significant patterns per complexity, which were centered at 1.5 s and 2.7 s. This could imply that the observed increase is due to stimulus-evoked JS activity. However, the latter is very unlikely because of four arguments. First, the validation of the results

10. Discussion

demonstrated that the observed fine-temporal structure is on a timescale of a few milliseconds. Thus, if the observed increase of JS activity is due to stimulus-evoked activity, the latter would have to have the required precision of a few milliseconds. Nevertheless, the latter is very unlikely, since the stimulus wasn't changing very fast, as would have been the case for a flashed stimulus. Instead, it was a slow sinusoidal modulation with a temporal frequency of 1.2 s. Thus, the intrinsic timescale of the grating was more than 200 times slower than the required millisecond precision of the observed JS pattern. The second argument against evoked JS patterns is that the number of significant JS patterns is increased during the whole stimulus presentation. Third, the effect of the flashed stimulus onset at 0 s is much weaker than the increase at 1.5 s and 2.7 s. Since the flashed onset is a much more optimal stimulus than a sinusoidal grating to evoke fine-temporal cross structure, this is a strong indication that the observed increases of the frequency of significant JS patterns are induced. The fourth argument is the strongest against evoked activity. Since each neuron had a slightly displaced receptive field relative to the others, the stimulus didn't arrive at the same time at each of them. This would imply that evoked activity with millisecond precision would also be mapped onto the latencies with millisecond precision. However, since the latencies range up to a few hundred milliseconds, the observed fine-temporal cross-structure on a millisecond scale cannot be evoked.

This is, to our knowledge, the first proof of induced and task-related fine-temporal cross-structure on a millisecond timescale, which indicates cooperative coding of large groups of cells on a millisecond timescale and which considers the complete auto-structure and trial-by-trial variability of spike trains.

10.2.2. Task-, Performance-, and Motor-Related Neuronal Assembly Formation

We applied three different analyses to spiking data that were recorded simultaneously in the prefrontal cortex of an awake monkey performing a short-term memory ('STM') paradigm (see Section 8.1).

The results of the first analysis (see Section 9.1), based on the frequency of JS patterns with significantly increased number of JSEs, indicate the existence of the task-related formation of neuronal assemblies. Especially during the late sample period and the early and late parts of the delay period, an increase of the frequency of significant JS patterns of complexities 2 to 4 were observed. Since the presented results were based on only 4 sessions from two monkeys, we cannot generalize the findings so far. Nevertheless, we demonstrated that the strategy chosen to analyze the data and the developed tools are performing well.

In a second step (see Section 9.3.2), we tested the importance of JS activity in the monkey's performance, based on a session in which single-unit activity from 27 cells was recorded. The large number of simultaneously recorded neurons increased the robustness of the group statistics of higher-complexity JS patterns of complexities 5 and 6, since the high number of different JS patterns of the same complexity compensates partially for the low expected probability of individual JS patterns. Based on the bi-variate

10.3. Discussion: Short-Term Memory-Related LFP Oscillation in the Prefrontal Cortex

(two-sample) UE analysis of the activity of these 27 cells, we have found JS activity that was performance-related and indicated by stronger synchronization if the monkey performed the task correctly. The tight linking of these performance-related modulations of the synchronization to the monkey's task indicates that neuronal assemblies are formed task-related and are necessary for the monkey to classify successfully the stimuli as match or non-match.

In a third step (see Section 9.4), we tested, based on the same 27 single-units, to which degree neuronal assemblies are related to the motor actions of the monkey (see Section 9.4). During the first period of the test presentation (150 to 550 ms after test onset), modulations in the number of JS patterns of complexities 2 and 5, with significant modulations of the normalized synchrony, seem to be performance-related rather than motor-related. After this first period, during which the maintained stimulus from the sample has to be compared to the test-stimulus to allow for a correct response by the monkey, the modulations in the number of JS patterns with significant modulations of the normalized synchrony become motor-related. This motivates the hypothesis that activity in the prefrontal cortex is related to decision making rather than to stimulus maintenance.

To validate these findings we have dithered spikes in the original dataset with ± 6 ms (see Section 9.5). After applying the dithering to the spikes, the total number of JS patterns that expressed significant modulations of the normalized synchrony went down to 20% for complexity 2, 12% for complexity 3, 5% for complexity 4, 2% for complexity 5, and 0.3% for complexity 6. That means in case of complexity 6 JS patterns that 99.7% of the JS patterns, which had significant modulations of the normalized synchrony before dithering, became non-modulated after dithering. This finding strongly indicates that, indeed, the fine-temporal cross-structure on a time scale of less than 12 ms has caused the above reported effects and results.

Despite these results, since we used only one experiment, we cannot generalize these latter findings so far. Nevertheless, the robustness and the strength of the observed modulations motivate the hypothesis that the formation of neuronal assemblies and the strengthening or weakening of synchronization across large groups of neurons is correlated to the periods of sample presentation, early and late delay, the test stimulus presentation, and the motor action, as well as correlated to the monkey's performance.

10.3. Discussion: Short-Term Memory-Related LFP Oscillation in the Prefrontal Cortex

In this study, we investigated first, whether task-related oscillations occur in local field potentials (LFPs) of ventral prefrontal cortex of an awake monkey performing an STM task and second, whether power and phase-locking of oscillations in various frequency bands is correlated with behavioral performance and identifies the presented stimuli (see Chapter 8). To this end, we applied the bi-variate and multi-variate

10. Discussion

bootstrapping tool that has been developed during this thesis (Section 4.1.4 and Chapter 5) to analyze the data. We geared the developed tools to test 3 working hypotheses:

Working hypothesis 1: *Neuronal oscillations and their synchronization are involved in processes underlying the STM in the pre-frontal cortex.*

Working hypothesis 2: *Oscillation of and coupling across neuronal cell assemblies are modulated in their amplitude and strength by the task and behavior of the monkey.*

Working hypothesis 3: *Oscillation of neuronal cell assemblies is carrying information about the memoranda and, therefore, is likely to be involved in the maintenance of information.*

10.3.1. Oscillations and Synchronization Involved in Short-Term Memory

Given typical electrophysiological recordings, we were concerned about two frequent observations. First, properties or statistics that are used to describe the data might have large variabilities, and second, the experimental conditions or the state of the investigated system might change during the experiment, since the latter can last anywhere between a couple of minutes to a couple of hours. These changes can either be induced by manipulations from outside (e.g. changed level of anaesthesia, drugs, etc.) or might be intrinsic, caused by changes in the experimental animal. In an awake animal, such as the awake monkeys from whom we have analyzed data, these intrinsic changes might occur due to different levels of alertness, tiredness or motivation. In addition episodic changes without any obvious reason might occur. We developed tools that account for these changes and that enabled us first, to perform bi-variate and multi-variate testing of the power and phase-locking in the case of two or N conditions, respectively, and second, to test if detected effects were significant, stable and reliable across the experiment (Section 4.1.4 and Chapter 5).

To test our three working hypotheses, we performed four analyses. First, we studied the precision and spatial extent of synchronization among recording sites, by assessing phase-locking and its performance-related modulations. Second, we identified performance-dependent modulations (bi-variate design) of the oscillation amplitude and the phase-locking. Third, we performed, in analogy with the bivariate test, a multi-variate test to identify stimulus-specific oscillations. Fourth, we compared and distinguished between performance- and stimulus-related power changes to allow for a mechanistic model of processes that are underlying the STM.

10.3.2. Spatial Extend of Neuronal Oscillation and Synchronization

We studied the precision and spatial extent of synchronization among recording sites by assessing phase-locking and its performance-related modulations (see Section 8.3.7). The results indicate that neuronal

oscillations in the beta and low-gamma bands between 10 Hz and 45 Hz are synchronized over a larger spatial extent than the average electrode spacing (400-900 μm), while neuronal oscillations above 50 Hz were not synchronized. This suggests that faster oscillations are restricted to single, or at most a few, cortical columns (Pucak et al., 1996) while slower oscillations involve many columns and probably other cortical areas. This is supported by studies showing that beta- but not gamma-oscillations synchronize across remote cortical areas (Brovelli et al., 2004; Roelfsema et al., 1997).

10.3.3. Performance-Related and Memorandum-Specific Neuronal Oscillation

Low-gamma components (30-45 Hz), which occurred during the early and late parts of the delay period, were both performance-related and stimulus-selective, while beta components were only performance-related (see Section 8.3.5). In analogy with the mechanistic idea that information is maintained by sustained activation similar to a labeled line code (Durstewitz et al., 2000; Fuster and Alexander, 1971; Goldman-Rakic, 1995; Gutkin et al., 2001; Miller et al., 1993), we were interested in signal modulations that lasted throughout the entire delay period. We found that among the memory-related signals, only the performance- and stimulus-related increases of oscillations in the high-gamma-band (65-90 Hz) lasted throughout the entire memory delay.

Because these high-frequency oscillations exhibited also stimulus-selectivity, the sustained modulations of high-frequency gamma-oscillations might constitute an important part of the maintenance process (see Section 8.4). Nevertheless, the mechanistic idea that information is maintained by sustained activity and, therefore, bound to localized units might be an over-simplification. Rather, information may be maintained through the trajectories of changing network activity that are an identical representation of the memorandum (Maass et al., 2002; Maass et al., 2004). The latter implies, first, that not the states themselves, but rather the trajectories are maintaining the information and second, that states that maintain information do not have to be sustained or stable in time.

Our measurements based on λ -maps account for this, since λ -maps represent the percentage of sites or pairs that expressed performance- or stimulus-related modulations (see Section 8.3.3). Therefore, it is not required that modulations of λ be due to the same sites and pairs throughout the whole delay. Indeed, our results of individual sites or pairs indicate that the performance- and stimulus-related increases of oscillations in the high gamma band, which lasted throughout the entire memory, are based on groups of sites and pairs that change over time.

Thus, we found evidence that the information about the memorized stimuli must be distributed across a large number of such local oscillating networks. Otherwise our arbitrary sampling of a few sites from a small patch of cortex would not have yielded that many sites with memory-related activity. Also, we found that units that are involved in the maintenance of information are changing during the delay.

The low average phase-locking (< 0.2) of these high-gamma oscillations (> 50 Hz) implies that the spa-

10. Discussion

tial extent of the processes based on high-frequency oscillations was small (see Section 8.3.7). This raises the question of how local processes that maintain information can be integrated into a larger network. The high phase-locking (> 0.5) of beta- and low-gamma oscillations across the extension of the electrode-grid, its persistence throughout the entire trial duration, and the performance-dependence of this locking suggest that slower oscillations might provide a global temporal reference signal for the coordination of the distributed local processes, which contain information about the stimuli. The transient episodes of beta- and low-gamma oscillations occurring at the beginning and the end of the maintenance interval were performance-dependent, but only the gamma- and not the beta-oscillations were stimulus-selective during these epochs. This might suggest that the transient beta/low-gamma oscillations serve the organization of the memory process during transitions from encoding to maintenance and from maintenance to retrieval. Alternatively, these episodes with beta/low-gamma oscillations could directly reflect encoding and retrieval processes. We prefer the first interpretation, because these low-frequency oscillations were not stimulus-selective.

A more likely correlate of encoding and retrieval processes are the high-gamma oscillations, because they had shorter latencies and were stimulus-selective. As suggested by strong phase-locking of low-frequency oscillations across the whole electrode grid, these oscillations appear to comprise larger networks. Because these persist during the maintenance interval, while stimulus-selective oscillations decrease their frequency in mid delay, these slower oscillations could reflect a rehearsal or recruitment process that involves representations in other cortical areas, such as visual area V4, for which locking of spikes to slow oscillations has been shown to occur during STM (Lee et al., 2005).

10.3.4. Model of Processes Underlying Short-Term Memory

We compared performance-related with stimulus-selective time frequency zones and identified overlaps between both. These overlaps could reflect interactions of different sub-processes that need to be coordinated during the memory process (see Section 8.5). In particular, during early delay, two transient zones around 20 Hz and 40 Hz coincided with a sustained time-frequency-zone ('TF-zone') around 80 Hz, all reflecting increases of power in correct trials. Since power around 60 Hz was not modulated, the interaction of the 20 Hz and 80 Hz processes could be mediated by a 40 Hz process, as a harmonic of the 20 Hz and a sub-harmonic of the 80 Hz process. This suggests that the 40 Hz process could mediate a handshake between a global performance-dependent 20 Hz and a local stimulus-selective 80 Hz process by means of a 1:2 and 2:1 synchronization (see Section 3.3). Such a synchronization among harmonic frequencies was recently reported for human MEG (Palva et al., 2005). During the proposed period of handshaking, the bandwidth of the local high-frequency processes is sharpened in relation to the pre-sample and sample periods. Therefore, it is conceivable that the power increase of the global and performance-dependent 20 Hz and 40 Hz oscillations serve to tune and coordinate the local stimulus-selective high-frequency

processes. The persistence of the performance-dependent increase in phase-locking of the 20 Hz oscillations during the entire delay supports the idea that the global 20 Hz oscillations function as a temporal reference signal for the maintenance of the 1:4 synchronization with the local 80 Hz oscillations.

10.4. Concurrence of Changes of the LFP and JS Activity during Short-Term Memory

Among the performance-related and stimulus-specific modulations of the LFP oscillations and synchronization, the performance-related strengthening of the phase-locking in the β -band during the sample period was the only signal modulation that appears to be correlated to an increased number of significant JS patterns (see Chapters 8 and 9). Since both phase-locking and JS activity assesses the mutual correlation of neuronal activity across electrodes, a concurrent performance-related strengthening for correct performance is, first, plausible and second, indicative that performance-related strengthening involves activity of much larger groups than the number of simultaneously recorded neurons. Also, the synchronization of spiking activity across neurons seems to be part of a larger network that is synchronized in the beta band.

Since we haven't performed an analysis of the spike field coherence so far, the latter can only be deduced from the concurrence, although, remarkably, the sample period was the only period that exhibited this concurrence. We suspect that this is the case because of lacking test power for effects of the JS activity during the delay period, rather than due to lacking effects. As we discussed in Chapter 9, we suspect that the BUE method wasn't able to detect relevant JS activity during the delay since we had to use a rather short sliding window (100 ms) to correct for changing rates, even though we expected that relevant JS activity during the non-stimulus driven delay period was scattered across times larger than the length of the window. If this were the case, JS activity might not have been detected in the same sliding window and was, therefore, statistically not detectable. To test this hypothesis, we have to use NeuroXidence, which allows us to make the sliding windows rather long. Since the LFP study revealed epochs of about 600 ms in length for which the power and synchronization of the LFP were performance-related and stimulus-specific, we propose to use a sliding window length of 600 ms to reanalyze the same data with NeuroXidence.

10. Discussion

11. Final Conclusions

11.1. Oscillation and Synchronization Underlying Short-Term Memory

The organization of short-term memory ('STM') seems to involve networks of different sizes within and, most probably, beyond prefrontal cortex, which oscillate at different frequencies. Small groups of neurons, probably comprising a cortical column, participate in gamma oscillations that carry partial information about the memorized stimulus. Larger networks in and, most likely, beyond prefrontal cortex appear to be coordinated by coherent low-gamma and beta oscillations, which are correlated with performance but not stimulus content. These global oscillations could coordinate the various stages of the memory process, provide a link to representations in other cortical areas containing more detailed information, and eventually organize executive functions. Global cortical circuits that organize behavior might employ short episodes of beta- and low-gamma oscillations to embed local processes into a distributed maintenance network. Oscillatory brain activity and its synchronization across space and frequencies might serve as a mechanism for the integration of distributed signals over different temporal and spatial scales.

11.2. Task- and Performance-Related Formation of Neuronal Assemblies

Based on the BUE method, we demonstrated task- and performance-related formations of neuronal assemblies during an STM task in the prefrontal cortex of a monkey. We have shown that JS activity during the sample period discriminates for behavioral performance and motor activity. The latter motivated the hypothesis that encoding during sample presentation, as well as motor preparation in the prefrontal cortex, relies on the formation of neuronal assemblies characterized by precise and reliable synchronization of spiking activity on a millisecond time scale.

11.3. NeuroXidence

The newly-developed analysis tool, NeuroXidence, detects coordinated firing events in spike trains. We have demonstrated its performance, reliability, and applicability compared to the capabilities of other popular and currently-used methods. NeuroXidence allows for a conclusive test of JS activity and for a clear dissociation of the assembly hypothesis from the rate hypothesis, since it considers all features of the data that might be causally related to the rate hypothesis in order to formulate H_0 (e.g. rate modulations or history dependencies, which might exist in each individual spike train).

11.4. Conclusive Evidence for Joint Spike Activity Supporting the Assembly Hypothesis

To our knowledge, we demonstrated for the first time the existence of induced and task-related JS activity of higher complexities ranging from 2 to 6 based on a hypothesis test that considers the complete auto-structure and trial-by-trial variability of spike trains. Thus, since NeuroXidence allows for a clear dissociation of the assembly hypothesis from the rate hypothesis, this finding can be seen as clear and conclusive evidence for the assembly hypothesis and for the cooperative coding of large groups of cells.

Part V.

Appendix

A. Appendix: NeuroXidence Methods

A.1. Preprocessing

NeuroXidence preprocesses coordinated neuronal firing patterns before the detection of all unique JSEs can be performed. This preprocessing prevents temporally overlapping JSEs, while preserving each individual JSE and its exact frequency of occurrence. Afterwards, the minimal interval between JSEs occurring within any single spike train is larger than the amount of allowed jitter.

Preprocessing a dataset requires two steps. First, an array of bins that contains zeros is padded to both ends of the spike trains to prevent border problems at later processing steps (Fig. A.1(A)). Second, NeuroXidence applies a recursive algorithm sequentially to each individual spike train to isolate all JSEs that are included in overlapping JSEs and to represent them as individual and isolated JSEs in the dataset for later processing (Fig. A.1(B)).

The resolving of overlapping JSEs might multiply sub-patterns that are included in the resulting isolated JSEs. The recursive algorithm has to correct this multiplication to allow for an exact derivation of the frequency of occurrence of each JSE. The algorithm first identifies, one spike train at a time, all sets of spikes that have an interspike interval smaller than $G^{+/-}$, a kernel defining the amount of allowed jitter. Operational fields (Fig. A.1(B)) describe the temporal windows spanning such sets. A field includes the G^- bins preceding the first and the G^+ bins following the last spikes of a set. To separate the JSEs localized by the overlapping operational fields (Fig. A.1(B), dark green), a modified copy of the overlapping JSEs is appended to the end of the dataset. Note, that only the data covered by operational fields is replicated and not the entire dataset. The copy is missing the second spike in the overlapping field, leaving an intact and isolated JSE. Likewise, in the original of the overlapping JSEs the first spike is deleted, producing the second isolated JSE. Next, the sub-patterns that were doubled by the copying process, i.e. those spikes that did not have overlapping operational fields, are copied into a d-dataset. Since the algorithm is applied sequentially in time, there is no difficulty in separating three or more spikes that occur in an interval smaller than $G^{+/-}$. For uncovering all overlapping JSEs in the original dataset, as well as in the d-dataset, the recursive algorithm is applied sequentially to all units in the dataset and in the d-dataset. In this way, while the algorithm processes unit n in the original dataset, it is processing unit $n+1$ in the d-dataset. This ensures that JSEs that are copied in the d-dataset no longer contain

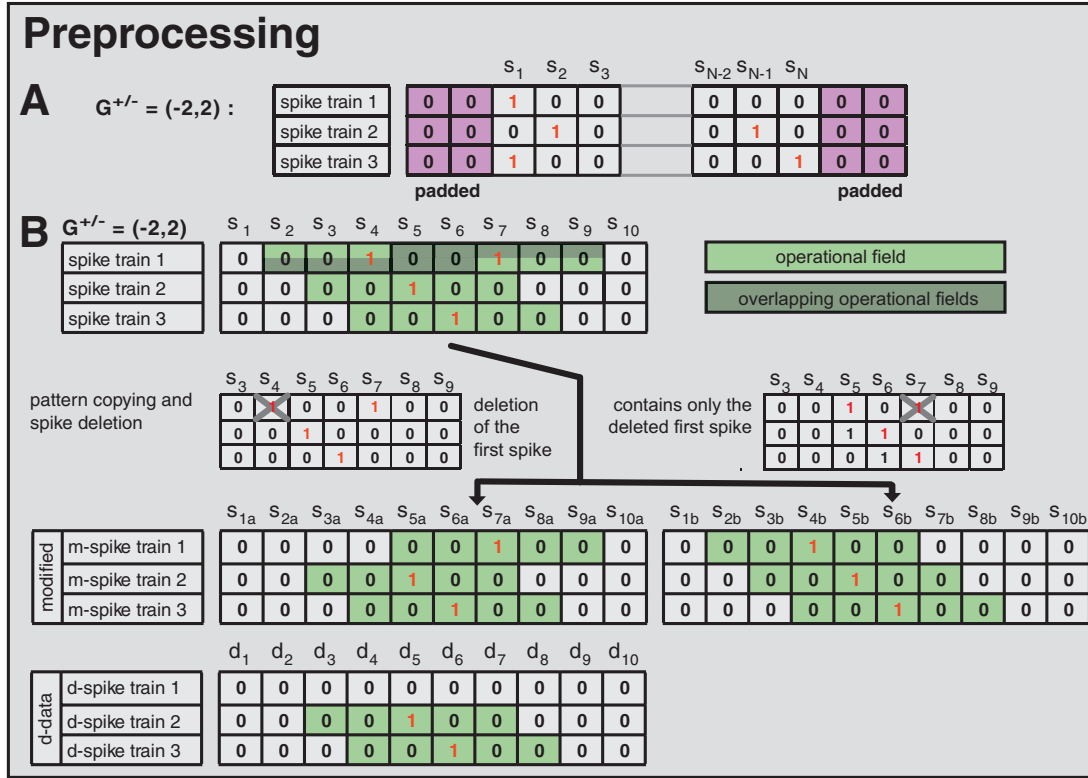


Figure A.1.: **Two preprocessing steps of NeuroXidence.** (A,B) Two preprocessing steps, which have to be applied to the spike trains in advance of the detection of joint spike events ('JSEs'). (A) Padding zeros to the beginning and the end of the spike trains ensures that the replacement of spikes by the kernel $G^{+/-}$ is possible. (B) Overlapping JSEs are isolated. Overlapping regions of allowed jitter (green) are identified, which correspond to overlapping operational fields. The algorithm uncovers overlapping patterns in two steps: First, the copying of patterns, and second, the deletion of single spikes. After revealing the overlapping patterns, the 'modified' dataset contains all isolated patterns, while the 'd-data' dataset contains all patterns that have been multiplied. The total frequency of occurrence of a test-pattern in the original data is given by the frequency in the 'modified' dataset, minus the frequency in the 'd-data'

any overlapping operational fields for units 1 to n . To accurately count the total number of isolated JSEs, the number of individual JSEs in the d-dataset has to be subtracted from the number of the same JS patterns in the original data, while the number of JSEs that were doubled in the d-dataset have to be added to the total. The computational complexity of resolving overlapping JSEs increases with the number of spikes that are less than $G^{+/-}$ apart. Thus, a higher computational effort is required if spike trains contain a lot of short intervals, as in the case of bursting cells or Poissonian firing. However, since the computational effort does not exponentially increase with the number of neurons, NeuroXidence can be effectively applied to large sets of simultaneously recoded neurons.

A.2. Standard Parameters for NeuroXidence

The standard choice for analysis parameters is $\tau_c = 5$ ms and $\tau_r = 20$ ms equivalent to $\eta = 4$. The duration of the analysis window l should be longer than 100 ms. The number of surrogates should be $S = 20$ if NeuroXidence is used to identify JS patterns with an excess of JSEs, while S should be 1 if NeuroXidence is used to identify a deficiency of JSEs. To allow for a reasonable test power, not less than 20 trials should be used. The bin-length corresponding to the timing resolution b in the spike train should be 1 ms. The statistical evaluation should be based on the Wilcoxon-test.

A.3. Analytical Description of the Test Power for Coherent Oscillatory Rate Changes

To derive an analytical description of Δf , we analytically formulated the impact of jittering. Jittering of spikes is equivalent to a convolution of the instantaneous rate with a jitter-kernel that is defined by the corresponding probability distribution of the jitter values. We assumed a rectangular jitter-kernel $\rho_{rect}(j)$, corresponding to a uniform distribution of random jitter values bounded between 0 and τ_c . The Fourier transform of $\rho_{rect}(j)$ is $P_{rect}(\nu')$.

$$\rho_{rect}(j) = \begin{cases} 1 & \text{for } 0 \leq j \leq \tau_c \\ 0 & \text{else} \end{cases} \Leftrightarrow P_{rect}(\nu') = \frac{\sin(\pi\tau_c\nu')}{\pi\tau_c\nu'} \quad (\text{A.1})$$

Therefore, the convolution of $r_i(t)$ with $\rho_{rect}(j)$ in the time domain is equivalent to a modulation of the Fourier spectrum of the instantaneous rate function $R_i(\nu)$ by $P_{rect}(\nu')$ in the frequency domain. In the case that the rate function is continuous, has infinite length, and is a sinusoid with frequency ν , the Fourier spectrum of $r_i(t)$ (see Equation 6.7) is a delta-peak at ν . Thus, the jittering of spikes maps the instantaneous rate $r_i(t)$ on $r_i^{jit}(t)$:

$$r_i^{jit}(t) = \frac{A'(\nu)}{2} * \left[\sin\left(2\pi\frac{t}{f}\right) \right] \text{ with } A'(\nu) = A \frac{\sin(\pi\tau_c\nu)}{\pi\tau_c\nu} \quad (\text{A.2})$$

Based on the instantaneous rates $r_i(t)$ and $r_i^{jit}(t)$, the instantaneous joint probabilities of a certain kind of JS pattern of complexity c is defined by $jp_c(t)$ for the original and $jp_c^{jit}(t)$ for the jittered datasets:

$$jp_c(t) = \prod_{i=1}^{i=c} r_i(t) \text{ and } jp_c^{jit}(t) = \prod_{i=1}^{i=c} r_i^{jit}(t) = A'(\nu)^c \prod_{i=1}^{i=c} r_i^{jit}(t) \quad (\text{A.3})$$

Thus, the expected frequency $\langle f \rangle_c$ of a JSE of complexity c in the time-interval between 0 and l is given by:

$$\langle f(org) \rangle_c = \int_0^l jp_c(t) dt \text{ and } \langle f(sur) \rangle_c = A'(\nu)^c \langle f(org) \rangle \quad (\text{A.4})$$

A. Appendix: NeuroXidence Methods

Accordingly, the expected difference between the expected frequency of JSEs in the original and the jittered datasets is given by:

$$\langle \Delta f \rangle_c = \langle f(org) \rangle_c - \langle f(sur) \rangle_c = (1 - A'(\nu)^c) \cdot \langle f(org) \rangle_c \quad (\text{A.5})$$

Note that the modulation is independent of the duration of the interval (l) used to derive the expected frequency of JSEs.

B. Appendix: Short-Term Memory-Related LFP Oscillation in the Prefrontal Cortex

B.1. Inter-Subject Comparison of λ -Maps

Memory-related activity during correct behavioral responses is characterized by several prominent modulations in λ_c -maps of power $\lambda_c(P)$ (Fig. 8.5(A, C) and supplementary Fig. B.1(A, C, E)) and phase-locking $\lambda_c(PL)$ (Fig. 8.5(B, D) and supplementary Fig. B.1(B, D, F)) in both monkeys (Monkey 1: 2402 trials, 7 sessions, 86 sites, 507 pairs; Monkey 2: 1722 trials, 5 sessions, 66 sites, 414 pairs). The percentage of sites and pairs exhibiting a significant increase in power or phase-locking ranged from 0.7 to 6%, 1% representing chance level. Power λ_c -maps of each of the two monkeys (Fig. B.1) show an increase above chance level by a factor of 3 to 6: first, during early delay (~ 0.35 s after sample off) and second, shortly before the onset of the test stimulus at beta- and low gamma-frequencies (10-45 Hz). This means that 3-6 times more sites than expected by chance generated significantly more power for correct responses in the respective time-frequency zones. Both periods lasted for about one second. At halftime of the memory delay (~ 1.2 - 2.7 s), both monkeys expressed a 2- to 4-fold increase of λ_c at low gamma-frequencies (45-70 Hz). Although weaker but much more extended in time, λ_c was elevated by a factor of 1.5-3 for 2.5 s during the delay at high gamma-frequencies (65-95 Hz). Because λ -maps of the two monkeys shared a substantial number of features, we assume that the underlying processes are tightly coupled to the task and are comparable across subjects. Therefore we used averaged λ -maps across all sessions of both monkeys for further analysis.

As described for $\lambda_c(P)$ -maps, $\lambda_c(PL)$ -maps of each of the two monkeys (Fig.B.1) exhibited several common features. First, a 2- to 3-fold increase of pairs with significantly increased phase-locking of beta- and low gamma-frequencies (7-35 Hz) during sample presentation in trials with correct responses. This increase, however, was more tightly locked to the presentation of sample stimuli for one of the monkeys. Second, during the delay, $\lambda_c(PL)$ was increased during several epochs by a factor of 2-3 in the range of 7-35 Hz. Third, similar to $\lambda_c(P)$, $\lambda_c(PL)$ was increased shortly before the onset of the test stimulus, albeit at higher frequencies. Fourth, increases of $\lambda_c(PL)$ in the high gamma band throughout the delay were less consistent across subjects and, like for power, their dominant frequency was lower in Monkey

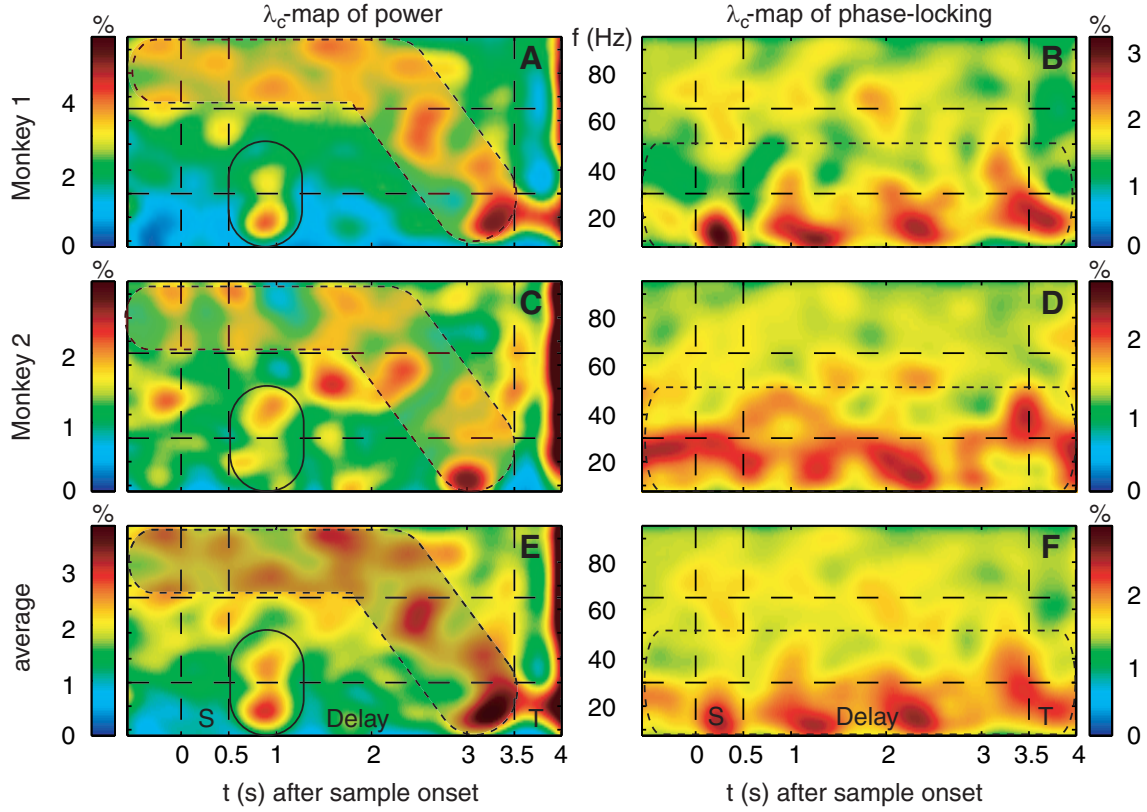


Figure B.1.: **Inter-subject comparison of λ_c -maps of correct responses (λ_c -maps).** Time-frequency λ_c -maps provide the percentage of sites/pairs with significant increases of power/phase-locking in trials with correct responses. Abscissa and ordinate represent peri-stimulus time aligned to the onset of sample stimuli and signal frequency, respectively changes (sample presentation 0-0.5 s /delay 0.5-3.5 s / test onset at 3.5 s). Color-codes provide the percentage of sites/pairs with significant differences. Maps of power are provided on the left side while maps of phase-locking on the right side. (A,B) provide λ_c -maps of Monkey 1 (2402 trials, 7 sessions, 86 sites, 507 pairs). (C,D) provide λ_c -maps of Monkey 2 (1722 trials, 5 sessions, 66 sites, 414 pairs). (E,F) provide the average of λ_c -maps of all sessions from both monkeys. (2402+1722=4124 trials, 7+5=12 sessions, 86+66=152 sites, 507+414=921 pairs)

2.

B.2. Detailed Discussion of the Coordination of Memory Processes by an $n:m$ Locking

Detailed handshaking by $n:m$ synchronization requires tuning of the frequencies of the individual processes. The driving force required for tuning is given by Arnold tongues (see Section 3.3). Higher order synchronization like a 1:4 locking, necessary for direct handshaking between 20 Hz and 80 Hz, likely requires stronger driving forces or more precisely adjusted frequencies of the individual processes than for a 1:2 and 2:1 locking (see Figure 8.7). Thus, a mediating process at 40 Hz would allow for a reliable and fast

B.2. Detailed Discussion of the Coordination of Memory Processes by an n:m Locking

handshaking by synchronization of the global performance-dependent 20 Hz- and the local coding-specific 80 Hz processes, even in case of weak driving forces and relatively broad frequency tunings. During the period of handshaking, the bandwidths of the local high frequency processes are sharpened in relation to the pre-sample and sample period. Thus, it is likely that the performance dependent transient increases in power of the global 20 Hz and 40 Hz oscillations reflect the tuning and coordination of the local stimulus-selective high frequency processes. In addition, the performance dependent strengthening of phase-locking at 20 Hz during the delay supports the hypothesis, that the global 20 Hz oscillations serve as a time reference signal by means of a 1:4 synchronization with tuned local stimulus-selective high frequency processes.

B. Appendix: Short-Term Memory-Related LFP Oscillation in the Prefrontal Cortex

C. Appendix: High Complexity JS Activity in Cat Visual Area 17

C.1. Preparation

Anaesthesia was induced with ketamine and maintained with a mixture of 70% N_2O and 30% O_2 supplemented with halothane (0.4-0.6%). The animal was paralysed with pancuronium bromide (Pancuronium, Organon, 0.15 mg $kg^{-1}h^{-1}$). All the experiments were conducted according to the guidelines of the American Physiological Society and the German Law for the Protection of Animals, approved by the local government's ethical committee and overseen by a veterinarian.

C.2. Recording

Multi-unit activity ('MUA') was recorded from a region of area 17 corresponding to the central part of the visual field. We used two SI-based multi-electrode probes (16-channels per electrode) supplied by the Center for Neural Communication Technology at the University of Michigan (Michigan probes) with an inter-contact distance of 200 μm (0.3-0.5 $M\Omega$ impedance at 1000 Hz). Signals were filtered between 500 Hz and 3.5 kHz for extracting MUA and stored in computer memory. The probes were inserted in the cortex approximately perpendicular to the surface, which allowed simultaneous recording from neurons at different cortical depths and along an axis tangential to the cortical surface. The receptive fields of neurons overlapped to a high degree, resulting in a cluster of receptive fields all activated simultaneously by the stimulus (see supplementary Fig. C.1). All MUA signals that showed good responses to visual stimuli and orientation selectivity were used for the analysis.

C.3. Visual Stimulation

Stimuli were presented binocularly on a 21" computer screen (HITACHI CM813ET) with 100 Hz refresh rate. To obtain binocular fusion, the optical axes of the two eyes were first determined by mapping the borders of the respective receptive fields and then aligned on the computer screen with an adjustable

C. Appendix: High Complexity JS Activity in Cat Visual Area 17

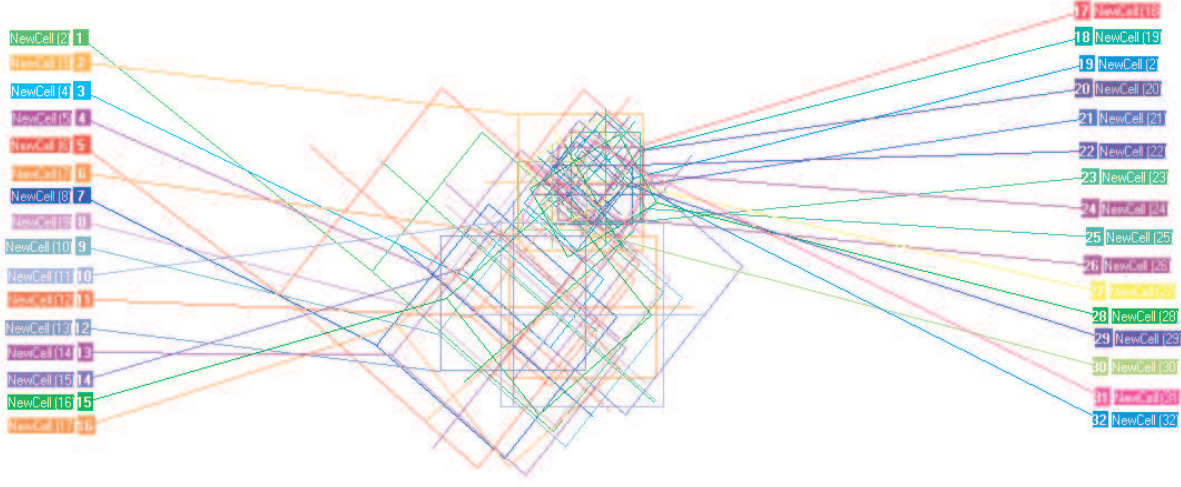


Figure C.1.: **Orientation tuning of recorded channels.** Squares indicate the size of the receptive fields for each recorded channel. Mid-lines represent orientation preferences

prism placed in front of one eye. The software for visual stimulation was a commercially available stimulation tool, ActiveSTIM (www.ActiveSTIM.com). The stimulus consisted of a drifting sinusoidal grating, spanning 15° of visual angle (spatial frequency: $3^\circ/\text{cycle}$; temporal frequency of the drift: $3.6^\circ/\text{second}$), which was sufficient to cover the receptive fields of all the recorded cells simultaneously and to stimulate also the regions surrounding the receptive fields. The orientation and the direction of the drift of the stimulus was optimal to produce strong rate responses of the cells. At each of the 20 stimulus presentations (trials), one second of spontaneous activity was recorded, which was followed by a 3.5 second presentation of the stimulus.

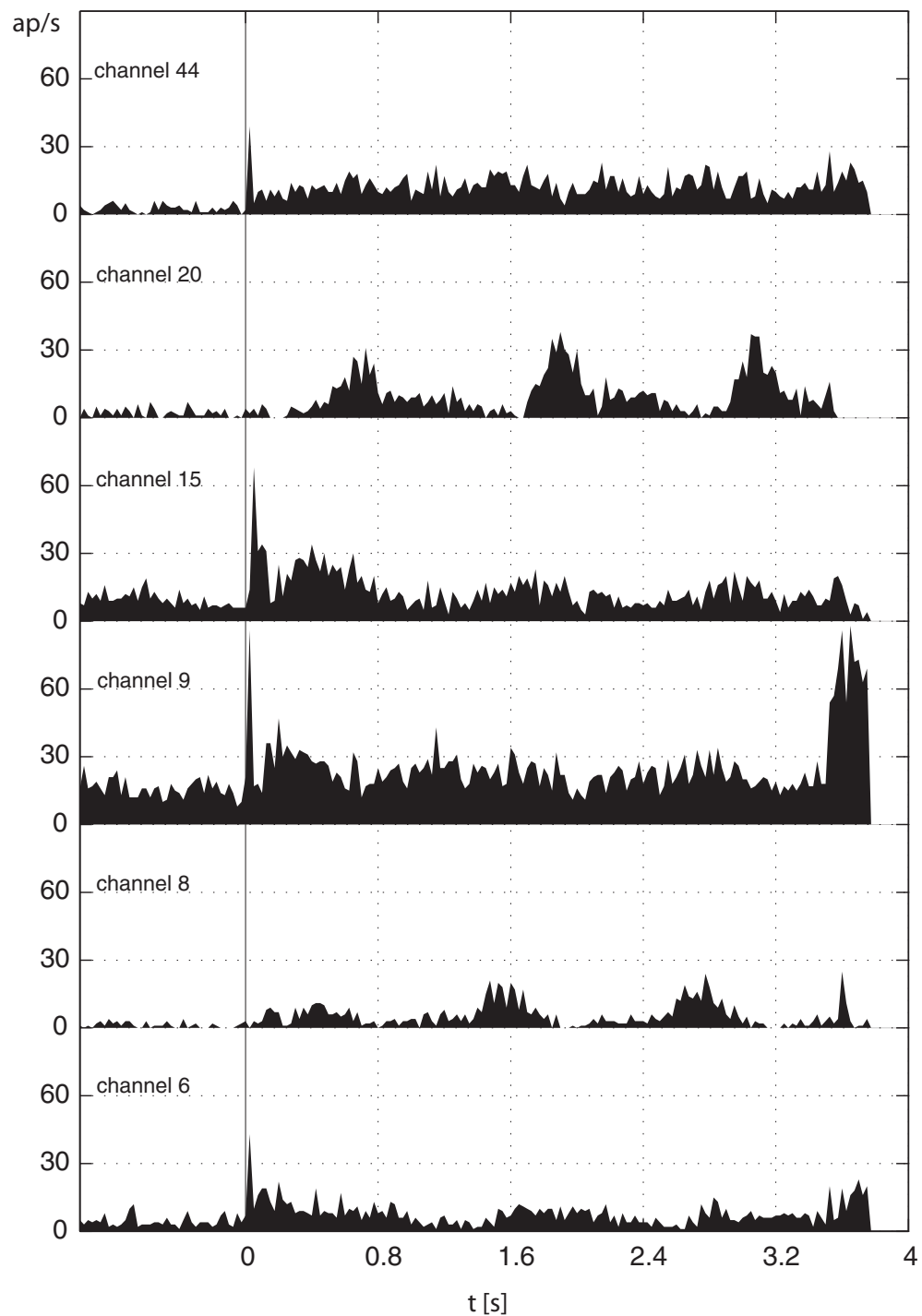


Figure C.2.: **Peri-stimulus-histogram (PSTH) recorded in 17.** PSTH of 6 single units in 20 trials recorded in cat visual area 17. The PSTH was derived based on a bin width of 25 ms. Channels were selected to be representative for 48 simultaneously-recorded channels.

C. Appendix: High Complexity JS Activity in Cat Visual Area 17

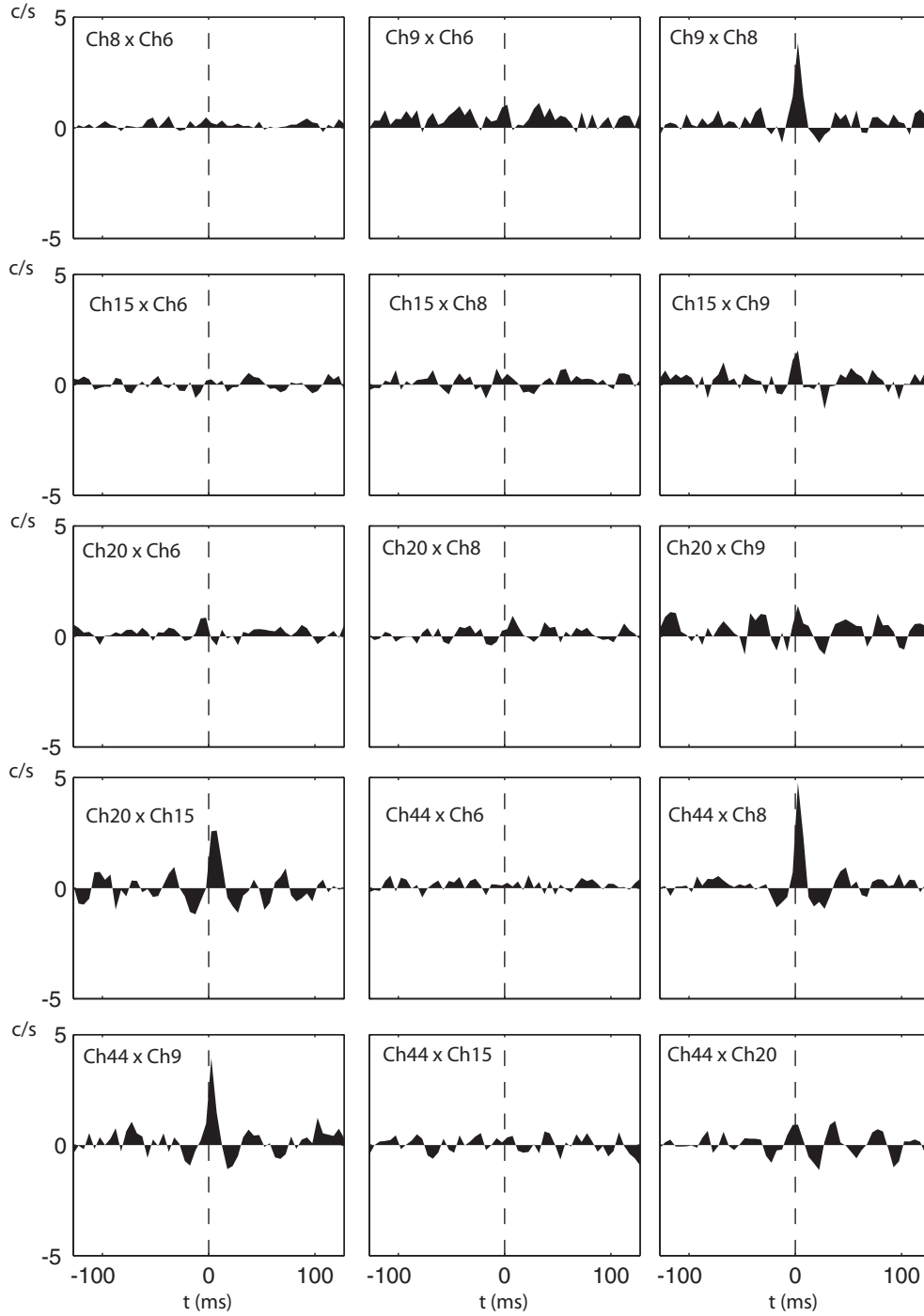


Figure C.3.: **Shift-corrected cross-correlograms.** All 15 shift-corrected cross-correlograms from the 6 single-units in Fig. C.2. The cross-correlograms were derived for the period between 0 and 3.5 s with 5 ms wide bins. Each subplot represents the shift-corrected cross-correlograms for one combination of two channels. The y-axis gives the number of excess coincidences per second and trial. The x-axis gives the time delay in milliseconds.

D. Appendix: Hypothesis Tests

D.1. The Concept of Hypothesis Testing

A statistical hypothesis test provides a mechanism for making quantitative decisions. It can be used to make a decision based on comparisons made between a single observed value versus some hypothesized quantity. The intent is to determine whether there is enough evidence to "reject" a conjecture or hypothesis. The conjecture is called the null hypothesis.

D.1.1. Null Hypothesis (H_0)

The null hypothesis (' H_0 ') represents a theory that has been put forward, either because it is believed to be true or because it is to be used as a basis for argument, but has not been proved.

D.1.2. Alternative Hypothesis (H_1)

The alternative hypothesis (' H_1 ') is the hypothesis that we consider as an alternative explanation of the data in case the test gives evidence that H_0 might not be true. The final conclusion, once the test has been carried out, is always given in terms of H_0 . We either 'reject H_0 in favor of H_1 ' or 'do not reject H_0 '; we never conclude 'reject H_1 ', or even 'accept H_1 '. If we conclude 'do not reject H_0 ', this does not necessarily mean that H_0 is true, it only suggests that there is not sufficient evidence against H_0 in favor of H_1 ; rejecting H_0 then, suggests that H_1 may be true.

D.1.3. Type I and Type II Errors

Type I Error (α -error, false positive):

In a hypothesis test, a type I error (also referred to as α -error or false positive) occurs when H_0 is rejected although it is in fact true; thus, H_0 is wrongly rejected. A type I error is considered to be more serious, and therefore more important to avoid, than a type II error (see following section). The hypothesis test procedure is therefore adjusted so that there is a guaranteed 'low' probability of rejecting H_0 wrongly. (Fig. D.3)

D. Appendix: Hypothesis Tests

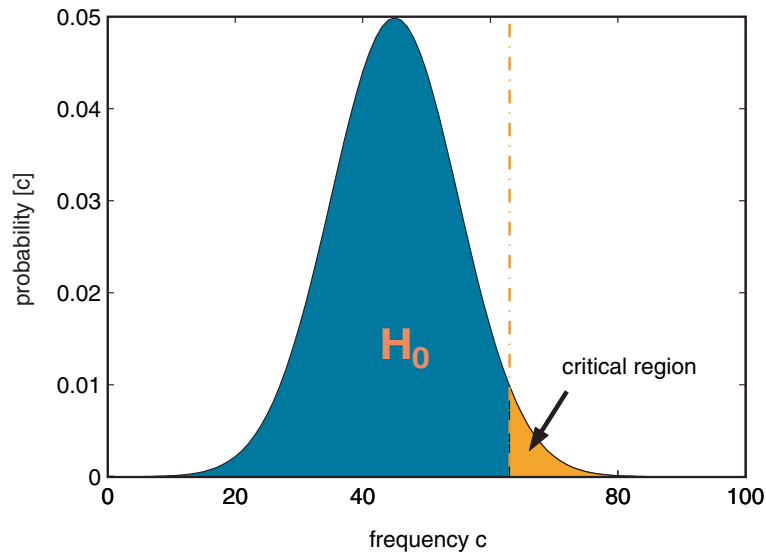


Figure D.1.: **Single sided hypothesis test:** A single sided hypothesis test tests if a value c is significantly different to the expectation formulated based on the null hypothesis (H_0). The null hypothesis is assumed to be a potential explanation if c is not in the critical region. In case c is in the critical region that is defined by the testlevel, H_0 must be rejected.

Type II Error (β -error, false negative):

In a hypothesis test, a type II error (also referred to as β -error or false negative) occurs, when H_0 is not rejected, when it is in fact false. If we do not reject H_0 , it may still be false (a type II error) as the sample may not be big enough to identify the falseness of H_0 (especially if the truth is very close to the hypothesis). For any given set of data, type I and type II errors are inversely related; the smaller the risk of one, the higher the risk of the other. (Fig. D.3)

	reject H_0	don't reject H_0
truth: H_0	type I error	right decision
truth: H_1	right decision	type II error

D.1.4. Critical Region

The critical region is a set of values of the test statistic for which H_0 is rejected in a hypothesis test; that is, the sample space for the test statistic is partitioned into two regions; one region (the critical region) will lead us to 'reject H_0 ', the other not. So, if the observed value of the test statistic is a member of the critical region, we conclude 'reject H_0 '; if it is not a member of the critical region then we conclude 'do

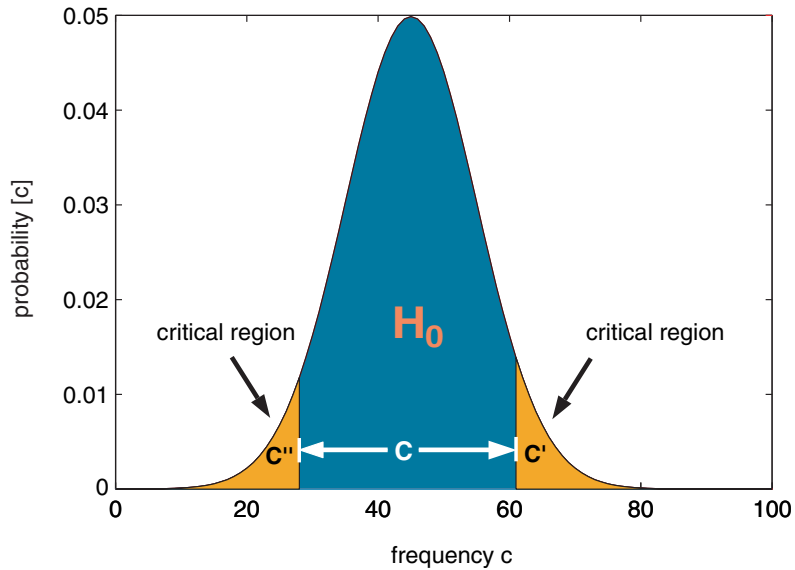


Figure D.2.: **Two sided hypothesis test:** A two sided hypothesis test tests if a value c is significantly different to the expectation formulated based on the null hypothesis H_0 . The null hypothesis is assumed to be a potential explanation if c is not in one of the two critical regions. In case c is in one of the two critical regions (either $c \leq c''$ or $c \geq c'$), H_0 must be rejected.

not reject H_0 '. (Fig. D.1 and Fig. D.2)

D.1.5. One Sided Test

In case of a one sided test the test statistic is compared with a one critical value to test if the test statistic is comparably larger or smaller. Since a one sided test either tests if the test statistic is larger or smaller than a critical value, it is equivalent to either a right sided or left sided test. In case of a right sided test H_0 is rejected if the test statistic is larger than this critical value, while, in case of a left sided test, H_0 is rejected if the test statistic is smaller than this critical value. (Fig. D.1)

D.1.6. Two Sided Test

A two sided test tests if a test statistic is either smaller than a lower critical value or larger than an upper critical value. (Fig. D.2)

D.1.7. p-Value and Test Power

There is a wide range of statistical tests available, depending on the nature of the investigation. However, the end result of any statistical test is a p -value. The (p) stands for probability, and measures how likely

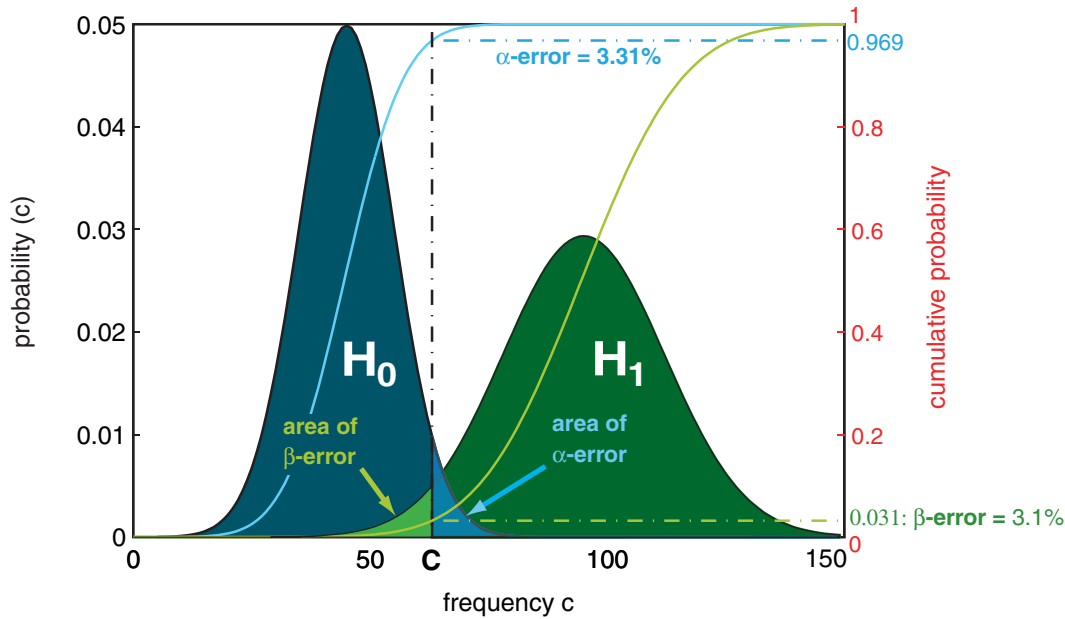


Figure D.3.: **False positives, false negatives and test power:** H_0 is the null hypothesis, H_1 the alternative hypothesis. Light blue indicates the area of α -errors equivalent to false positives, while light green indicates the area of the beta errors equivalent to false negatives. The testpower (β) of the test is given by $\beta = 1 - p(\text{false negatives})$

it is that any observed difference between groups is due to chance. In other words, the p -value is the probability of seeing the observed difference, or greater, just by chance if H_0 is true. Being a probability, p can take any value between 0 and 1. Values close to 0 indicate that the observed difference is unlikely to be due to chance, whereas a p -value close to 1 suggests that there is no other difference between groups than that due to random variation.

The power of a statistical hypothesis test measures the test's ability to reject H_0 when it is actually false. In other words, the power of a hypothesis test is the probability of not committing a type II error. It is calculated by subtracting the probability of a type II error from 1, usually expressed as: power = 1 - $p(\text{type II error})$. The maximum power a test can have is 1, the minimum is 0. Ideally we want a test to have high power, close to 1.

D.2. Parametric Tests

Methods such as the t -test are known as 'parametric' because they require an estimation of the parameters that define the assumed underlying distribution of the data; in the case of the t -test, for instance, these parameters are the mean and standard deviation that define the normal distribution.

D.2.1. Unpaired t-Test

The t -test performs a hypothesis test that tests if a sample with N elements $\vec{x} = \{x_1, \dots, x_N\}$ comes from a normal distribution with an expected value of zero and unknown variance.

$$T = \sqrt{N} \frac{\bar{x} - m}{std(x)} \quad (D.1)$$

Where $std(x)$ is, is sample standard deviation. The p -value is derived given the t -statistic and the degree of freedom.

D.2.2. Paired t-Test

The paired t -test is an extension of the unpaired t -test. The only difference is that the samples that enter the test $\vec{x} = \{x_1, \dots, x_N\}$ are paired differences $x_i = x_i^a - x_i^b$ of two samples a and b .

D.2.3. One-Way Analysis of Variance (ANOVA)

The purpose of the one-way ANOVA is to find out whether samples $\vec{x}^i = \{x_1^i, \dots, x_M^i\}$ from several groups (i), with $i = 1 \dots I$ come from the same sampled distribution F and thus have a common expected value (μ). That is, to determine whether the groups are actually different in the measured characteristic. Thus H_0 and H_1 of the ANOVA test is:

$$H_0 : \mu_i = \mu_j \text{ with } \mu_i \text{ expected value of } F_i \quad (D.2)$$

$$H_1 : \text{for at least for one pairs holds } \mu_i \neq \mu_j \quad (D.3)$$

Thus a rejection of H_0 gives evidence that at least one of the $\vec{x}^i = \{x_1^i, \dots, x_M^i\}$ is not originated from the same distribution as the others. Nevertheless, H_0 is also rejected if one half of the groups is different to another half. Hence, ANOVA does not allow to identify which sample is different from whom. Only a posthoc test that is applied after H_0 had been rejected can identify differences between pairs of samples, but ANOVA is suitable to screen for differences between different samples.

The basic concept that is used by the ANOVA test is to estimate to which degree the variance across elements across groups can be predicted by the variance in groups. In case that the expected value of at least one group is different from the other the variance across groups is increased. To estimate the difference of the mean value across groups, an analysis of variance is performed. To this end three error terms are defined, the total error QS_{tot} , the error due to hidden sources or noise QS_{error} and the treatment error due to differences across the groups QS_{treat} . The latter describes the effect caused by differences across groups and has to be compared to the random variation in each group QS_{error} .

For $i = 1 \dots I$ groups with each $m = 1 \dots M$ samples holds:

D. Appendix: Hypothesis Tests

$$QS_{tot} = QS_{error} + QS_{treat} \quad (D.4)$$

and

$$\hat{\sigma}_{tot}^2 = \frac{QS_{tot}}{df_{tot}} \quad \wedge \quad \hat{\sigma}_{error}^2 = \frac{QS_{error}}{df_{error}} \quad \wedge \quad \hat{\sigma}_{treat}^2 = \frac{QS_{treat}}{df_{treat}} \quad (D.5)$$

and for the degrees of freedom

$$df_{tot} = df_{error} + df_{treat} \quad (D.6)$$

$$df_{tot} = IM - 1 \quad \wedge \quad df_{error} = I(M - 1) \quad \wedge \quad df_{treat} = M - 1 \quad (D.7)$$

For the individual error terms holds:

$$QS_{tot} = \sum_{i=1}^I \sum_{m=1}^M (x_{i,m} - \bar{G})^2 \quad \text{with} \quad \bar{G} = \sum_{i=1}^I \sum_{m=1}^M x_{i,m} \quad \text{the total mean} \quad (D.8)$$

$$QS_{error} = \sum_{i=1}^I \sum_{m=1}^M (x_{i,m} - \bar{A}_i)^2 \quad \text{with} \quad \bar{A}_i = \sum_{m=1}^M x_{i,m} \quad \text{the group mean} \quad (D.9)$$

$$QS_{treat} = M \sum_{i=1}^I (\bar{G} - \bar{A}_i)^2 \quad (D.10)$$

As a test statistics the quotient of $\hat{\sigma}_{treat}^2$ and $\hat{\sigma}_{error}^2$ is used:

$$F = \hat{\sigma}_{treat}^2 / \hat{\sigma}_{error}^2 \quad (D.11)$$

D.3. Non-Parametric Tests

Parametric methods require assumptions to be made about the sampled distribution. For example, the paired t -test requires the distribution of the differences to be approximately normal, while the unpaired t -test requires an assumption of normality to hold separately for both sets of samples. Fortunately, these assumptions are often true, and where they are not true in the raw data, it is often possible to apply a suitable transformation. There are situations in which even transformed data may not satisfy the assumptions, however, in these cases it may be inappropriate to use traditional (parametric) methods of analysis. Thus, one of the major concerns about parametric approaches is that a rejection of H_0 may be originated by a violation of the assumed model of the data like normality rather than by an effect like a difference of the mean value in case of the t -test. Non-parametric methods provide an alternative series of statistical methods that require no or very limited assumptions to be made about the data. There is a wide range of methods that can be used in different circumstances. But there is of course also a tradeoff, that is, that non-parametric methods, particularly when used in small samples, have rather less power (i.e. less chance of detecting a true effect where one exists) than their parametric equivalents.

D.3.1. Sign Test for One-Sample or Paired Data

The sign test is the simplest of all the non-parametric methods. It is used to compare a single sample with some hypothesized value, and it is therefore of use in those situations in which the one-sample or paired t -test might traditionally be applied. The sign test is called so because it allocates a sign, either positive (+) or negative (-), to each observation according to whether it is greater or less than some hypothesized value. Mathematically explained, the sign test tests if the median of the sample is significantly different from the assumed median. In case of an exact solution of the sign test, the p -values are based on the binomial distribution.

D.3.2. Wilcoxon Signed-Rank Test for One-Sample or Paired Data

The sign test is intuitive and extremely simple to perform. However, one immediately obvious disadvantage is that it simply allocates a sign to each observation, according to whether it lies above or below some hypothesized value, and does not take the magnitude of the observations into account. Omitting information about the magnitude of the observations is rather inefficient and may reduce the statistical power of the test. An alternative test that does account for the magnitude of the observations is the Wilcoxon signed-rank test. Since this test is extensively used in this thesis we are going to introduce the technical details next:

Procedure of a Wilcoxon signed-rank test: Like the t -test for correlated samples, the Wilcoxon signed-rank test applies to a two-sample design involving repeated measures, matched pairs, or "before" and "after" measures. Given two sets A and B of M paired values of X_i^a and X_i^b with $i = 1 \dots M$, first the absolute value of the paired difference d_i between X_i^a and X_i^b is derived.

$$d_i = |X_i^a - X_i^b| \quad (\text{D.12})$$

The set \vec{d} includes all d_i that are not zero. Thus, all d_i that are zero ($d_i = 0$) are omitted since they do not show an effect. Next, the values in \vec{d} are ranked from smallest to largest employing tied ranks where appropriate. Ties in the ranks are cases for which two or more elements ($d_i = 0$) have the same value. In such a case the subset of \vec{d} , that contains all d_i with the same value, is selected. Then the average rank across the subset is assigned to each of the elements of the subset. In a last step the sum $W+$ of the ranks of those elements that have a difference d_i larger than zero, and the sum $W-$ of the ranks of all those elements d_i that are smaller zero is computed. Based on the difference of $W+$ and $W-$ the Wilcoxon signed-rank decides if the median of all d_i is significantly different from zero in case of a two tailed test. Consequently, in case of a one tailed test the Wilcoxon signed-rank decides based on the difference of the rank sums ($W+$ and $W-$) if the median of all d_i is significantly larger (right handed test) or significantly smaller (left handed test) than zero.

D.3.3. Wilcoxon Rank Sum (Mann-Whitney) for Two Unpaired Samples

The sign test and Wilcoxon signed-rank test are useful non-parametric alternatives to the one-sample and paired t -tests. A non-parametric alternative to the unpaired t -test for two samples is given by the Wilcoxon rank sum test, which is also known as the Mann-Whitney test. This is used when comparison is made between two independent groups.

D.3.4. Kruskal-Wallis Non-Parametric Equivalent of ANOVA

The Kruskal-Wallis test is a non-parametric version of the one-way analysis of variance (ANOVA, see Appendix D.2.3). The assumption behind this test is that the measurements come from a continuous distribution, but not necessarily a normal distribution. The test is based on an analysis of variance using the ranks of the data values, not the data values themselves.

D.4. Parametric versus Non-Parametric

Inevitably there are advantages and disadvantages of non-parametric versus parametric methods, and the decision regarding which method is most appropriate depends very much on individual circumstances. As a general guide, the following (not exhaustive) guidelines are provided.

D.4.1. Advantages of Non-Parametric Methods

Non-parametric methods require no or very limited assumptions to be made about the format of the data, and they may therefore be preferable when the assumptions required for parametric methods are not valid. Since most non-parametric methods rely on order statistics rather than on interval statistics, they can deal with unexpected, outlying observations that might be problematic with a parametric approach. Most non-parametric methods are intuitive and are simple to carry out by hand, for small samples at least. Non-parametric methods are often useful in the analysis of ordered categorical data in which assignation of scores to individual categories may be inappropriate.

D.4.2. Disadvantages of Non-Parametric Methods

Non-parametric methods may lack power compared to more traditional approaches. This is a particular concern if the sample size is small or if the assumptions for the corresponding parametric method (e.g. normality of the data) hold. Non-parametric methods are geared to hypothesis testing rather than estimation of effects. Tied values can be problematic when these are common, and adjustments to the test statistic may be necessary.

E. Appendix: Bootstrapping and Resampling

There are a lot of valuable statistical methods that are practically guaranteed to work well if the data is approximately normally distributed and we are mainly concerned with linear functions of random variables (see Appendix D, D.2). An example for such a statistic is the mean value of a sample that is proved by the central limit theorem to converge on normality as the sample size increases. However, in neuroscience we are often forced to use small samples or we deal with estimates for which we do not have an estimation of the variability of the estimator itself. An example for which an estimator of the variability does not exist is the coherence or the phase-locking value ('PLV'). However, the estimation of the variability of an estimator is essential for the statistical assessment of reliability or for hypothesis testing.

A relatively new development in statistical methodology offers a way out of this dilemma. The technique is called "bootstrapping", which, according to Efron and Tibishirani (Efron and Tibishirani, 1993) was named from the phrase "to pull oneself up by one's bootstraps", i.e., to accomplish a physical impossibility. Efron and Tibishirani ((Efron and Tibishirani, 1993) on page 56) note that the bootstrap was introduced by Efron in 1979, making it quite a recent development in contrast to many other statistical techniques. It was preceded by "jackknifing" which was originated by Quenouille (1956) as a way to study bias in estimators, but named by John Tukey (1958). One of the nice things about bootstrapping is that it is simple to apply, so long as one has access to a computer.

E.1. Introduction to Resampling

The probability distribution F of a random variable x is any complete description of the probabilistic behavior of x . In the following example we are interested in the empirical $\hat{H}(y)$ and theoretical $H(y)$ probability distribution of the total number of spots y of 8 dice.

$$y = g(x) = \sum_{i=1}^8 x_i \quad (\text{E.1})$$

In rolling a dice, there are 6 outcomes, each having $p(x) = 1/6$. Based on the theoretical value that is equivalent to the expected value of the probability $p(x)$ one can derive the expected probability distribution of the total number y of 8 replications of independent tosses.

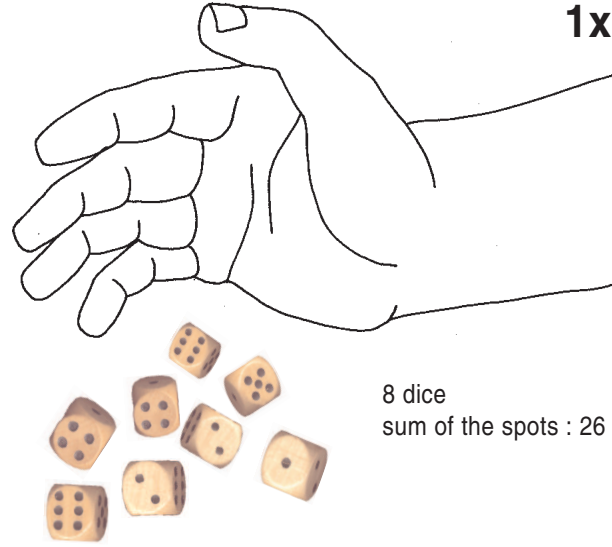


Figure E.1.: **One toss of eight dice:** One toss with eight dice is utilized to estimate the empirical probability distribution $\hat{F}(x)$

In case the distribution $F(x)$ is unknown one has to conduct an experiment that estimates $F(x)$ by $\hat{F}(x)$. The example in Fig. E.1 illustrates this situation. $\hat{F}(x)$ is estimated based on 8 independent tosses with one dice. Thus, the empirical probability distribution $\hat{F}(x)$ is estimated by the empirical probability of each of the occurred number of spots,

$$\hat{F}(x) = \hat{p}(x = k) = \frac{1}{8} * [1, 2, 0, 2, 1, 2] \text{ for } k = (1 \dots 6) \quad (\text{E.2})$$

Note, the experiment does not ensure that all potentially existing 6 configurations are occurring nor it estimates the empirical probability very accurately.

Based on the 'plug-in principle' (Efron and Tibishirani, 1993) resampling of the empirical distribution $\hat{F}(x)$ can be utilized to derive the empirical probability distribution $\hat{H}(y)$. To this end the original dataset containing n items (here $n = 8$) is randomly sampled with replacement B times, with each

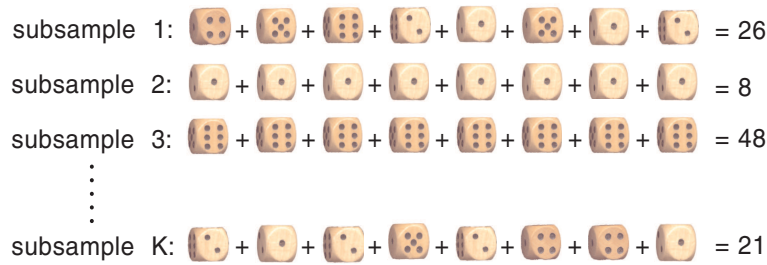


Figure E.2.: **Resampling of $\hat{F}(x)$:** Four bootstrap samples built by resampling of $\hat{F}(x)$

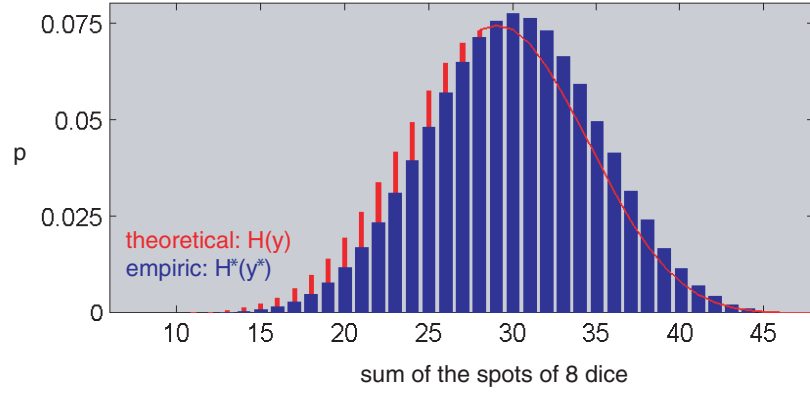


Figure E.3.: **Theoretical and bootstrapped distribution of y :** Theoretical (red) and bootstrapped (blue) distribution derived by resampling of \hat{F}

sample containing exactly n items (four of these B samples are shown in Fig. E.2). Note that an individual value from the original data set may appear repeatedly in a bootstrap sample. For each of the B bootstrap samples the total number y^* is derived. The latter is a bootstrap replication. We then use these B replicate values to derive the empirical distribution $\hat{H}^*(y^*)$ of y^* corresponding to the total sum of spots of eight dice (Fig. E.3). Even though the number of samples used to derive $\hat{F}(x)$, that is the only information that entered into the bootstrap estimation of $H(y)$, was pretty low, the estimation of $\hat{H}^*(y^*)$ comes pretty close to the theoretical distribution $H(y)$.

E.2. Formal Definition of Bootstrapping

Based on the distribution $F(x)$ of a random variable x , a parameter $\Theta(F)$ of F can be defined.

$$\Theta = g(F) \quad (\text{E.3})$$

An example of such a parameter might be the expected value of the variance of the distribution F . In case the distribution $F(x)$ is unknown, one has to sample $F(x)$ by using an experimental approach. Thus, sampling leads to a limited number (n) of samples x_i distributed according to F , which can be described as a set \vec{x} .

$$F \longrightarrow \vec{x} = (x_1, x_2, \dots, x_n) \quad (\text{E.4})$$

Hence, the empirical distribution \hat{F} is given by the relative frequency of all elements that occur in the sample \vec{x} .

$$\hat{F} = \hat{F}(x_1, x_2, \dots, x_n) \quad (\text{E.5})$$

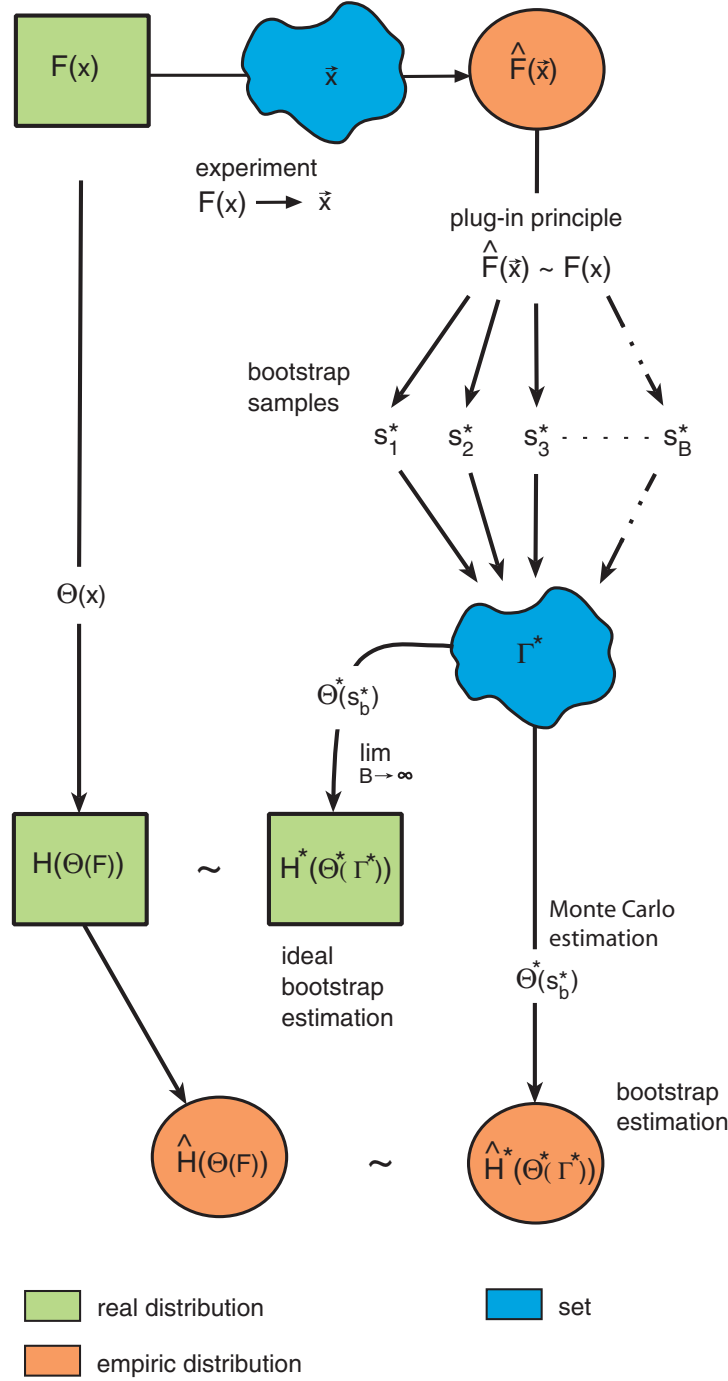


Figure E.4.: **Basic principals of the bootstrapping idea.** The basic idea is that F is sampled by \vec{x} . To estimate the distribution $H(\theta(F))$, bootstrapping is resampling \hat{F} . Based on the set of all potentially existing bootstrap samples Γ $H(\theta(F))$ can be approximated by and ideal bootstrap estimation ($H^*(\theta^*(\Gamma^*))$) or estimated by a Monte Carlo estimation ($\hat{H}^*(\theta^*(\Gamma^*))$)

In case the distribution F is unknown one has to use \hat{F} to estimate $\Theta(F)$ by $\hat{\Theta}(\hat{F})$. An example for this estimation is the estimation of the expected value μ of F by the mean value \bar{m} of \hat{F} .

So far we haven't employed any bootstrapping. Bootstrapping becomes of interest if one is interested in the distribution $H(\Theta(F))$ of the parameter $(\Theta(F))$ without knowing F itself. In this case the distribution H has to be estimated based on $\hat{F}(x)$. To this end bootstrapping utilizes the plug-in principles. It assumes that sampling \hat{F} is a good approximation of sampling F . Consequently $H(\Theta(F))$ can be estimated by $\hat{H}(\Theta(\hat{F}))$. To this end bootstrapping resamples $\hat{F}(x)$ and creates multiple (B) bootstrap samples \mathbf{s}_b^* that are used to derive B estimation of Θ^* .

The additional $*$ indicates that estimations are based on bootstrap replications.

$$\vec{x} \longrightarrow \mathbf{s}_b^* = (x_1^*, x_2^*, \dots, x_n^*) \quad \text{with} \quad \mathbf{s}_b^* \quad \text{bootstrap sample } b \quad (\text{E.6})$$

Given the B bootstrap estimations of Θ^* , the distribution $H(\Theta)$ is estimated by the empirical distribution H^* of $\Theta^*(\hat{F})$.

E.3. Monte Carlo Estimation

Given the set \vec{x} used to estimate \hat{F} the set Γ of all potentially existing bootstrap samples \mathbf{s}_b^* is defined by:

$$\Gamma^* = \{\mathbf{s}_b^* \mid \mathbf{s}_b^* = (x_1^*, x_2^*, \dots, x_n^*) \wedge b = 1 \dots B\} \quad (\text{E.7})$$

In case all B potentially existing bootstrap samples \mathbf{s}_b^* are used for the estimation H^* all the information included in \hat{F} is considered. Thus, in case the complete set Γ^* is used, one refers to an ideal bootstrap estimation. However, since B can be very large one might consider to approximate the estimation of H^* by a random sampling of Γ^* . In the latter it is referred to as a Monte Carlo estimation of H^* by \hat{H}^* .

E.4. Bootstrap Standard Deviation of an Arbitrary Parameter

In this section the standard deviation of an arbitrary parameter $\Theta(F)$ is derived (Efron and Tibishirani, 1993). The distribution F is approximated by \hat{F} . Based on \hat{F} the distribution \hat{F}^* is derived by resampling \hat{F} . Consequently the standard deviation of $\Theta(F)$ is given by:

$$\hat{\sigma}_{\Theta(\hat{F})}^* = \sqrt{\frac{\sum_{b=1}^B [\hat{\Theta}^*(\mathbf{s}_b^*) - \langle \hat{\Theta}^*(\mathbf{s}_b^*) \rangle]^2}{B}} \quad (\text{E.8})$$

with

$$\langle \hat{\Theta}^*(\mathbf{s}_b^*) \rangle = \frac{1}{B} \sum_{b=1}^B \hat{\Theta}^*(\mathbf{s}_b^*) \quad (\text{E.9})$$

E. Appendix: Bootstrapping and Resampling

If all elements of Γ are used, the estimation is an ideal estimation (Efron and Tibishirani, 1993):

$$\hat{\sigma}_{\Theta(\hat{F})}^* = \sigma_{\hat{F}}^* = \sigma_{\Theta(\hat{F})} \quad (\text{E.10})$$

In case Γ is randomly sampled by a Monte Carlo procedure for $\hat{\sigma}_{\hat{F}}$ holds (Efron and Tibishirani, 1993):

$$\lim_{B \rightarrow \infty} \hat{\sigma}_{\Theta(\hat{F})}^* = \sigma_{\hat{F}}^* = \sigma_{\Theta(\hat{F})} \quad (\text{E.11})$$

E.5. Hypothesis Testing by Bootstrapping

The distribution H^* of an estimator Θ , that is derived by bootstrapping the empirical distribution \hat{F} can be used to derive confidence intervals of Θ , by using percentiles of H^* . Thus, since the estimation of confidence intervals is tightly related to the issue of hypothesis testing, H^* can also be used for a statistical hypothesis test that is testing if an assumed value κ is significantly different from the estimator Θ that is derived by \hat{F} . The usage of H^* to derive percentiles is called the standard percentile method. The standard percentile method simply uses the $\alpha/2$ and $1 - \alpha/2$ percentiles of the bootstrap distribution H^* to define the confidence interval of Θ for a given α -level. This method performs well for quantiles and for statistics that are unbiased and have a symmetric sampling distribution. For a statistic that is biased, the percentile method unfortunately amplifies the bias. Thus, one has to be careful if H^* should be employed for a hypothesis test, since a bias might lead to a false rejection of H_0 . In Chapter 5 we will present a new alternative approach that we have developed to overcome the latter limitation.

F. Appendix: Parallel and Load Balanced Computation with ClusterMatlab

All the presented methods and projects in the preceding chapters wouldn't have been possible without an immense amount of computational power. At the same time the development of the new tools had to be as efficient as possible in terms of man power. Therefore we used Matlab, because of the high algorithmic complexity of each of the analysis tools. To make the computation efficient we first wrote a Matlab application and identified, in a second step, the computationally most demanding procedures that were then substituted by more efficient C-implementations. Still, the computing time, required for the results presented in this thesis, amounted to approx. 10^5 hours (on a conventional computer). To make this possible despite of this enormous computing time, we utilized a Linux cluster with 16 CPU (P4 3.2 GHz, 2 GB RAM) and wrote a toolkit 'ClusterMatlab' (copyright by Gordon Pipa and Na Hui Chen, Fig. F.1) for Matlab that allowed parallel and load balanced computation of Matlab functions. The basic idea of the toolkit is that statistical evaluation is often based on batch like processes of different data segments evaluated by the same function. This offered us the possibility to parallelize the presented statistical tools based on independent computations on different parts of the data. This can either be based on different sliding windows or time-frequency bins. To hide from the user all the technical details necessary for parallel computation, we wrote a Matlab toolkit that requires only 4 definitions to be made by the user:

1. **FileToExecute:** Name of your Matlab function that you want to run on your cluster with ClusterMatlab (e.g. FileToExecute= 'DoComputation')
2. **ParameterfileID:** ID of the parameter file. Based on different IDs you can run different jobs with the same framework (e.g. ParameterfileID= 'ID001')
3. **wishedNrCPUs:** Wished Nr of CPUs: if wishedNrCPUs=-1 maximal available number is used. Otherwise the wished number of CPUs is used (e.g. wishedNrCPUs=13)
4. **Nicelevel:** Optional, allows to give a certain priority (1=maximal priority, 19=minimal priority) to your jobs (e.g. Nicelevel=19)

F. Appendix: Parallel and Load Balanced Computation with ClusterMatlab

Prerequisite of the toolbox is the Matlab compiler that is utilized to make an executable from the Matlab *m*-files and to get independent from Matlab licences on the computing nodes.

```
%+++++
%   Definitions and parameters that have to be modified by the operator
%-----
File_to_execute   = 'Do_computation_M'           % Name of your Matlab function that you want to run
                                                    % on your Cluster with Cluster Matlab

Parameterfile_ID  = 'ID_001'                     % ID of the parameterfile: based on different IDs you
                                                    % can run different jobs with the same framework

wished_Nr_CPUs    = 5;                           % Wished Nr of CPUs: if -1 maximal available number
                                                    % otherwise the wished number of CPUs is included in
                                                    % the MPI Cluster

Nice_level        = 19;                           % Nice level of your jobs

%#####
%   Code - Don't touch the code from here on
%-----
Filename_executable      =[File_to_execute '_EXE'];
[Nr_CPUs Nr_Nodes CPU_per_node] = Check_MPI_Cluster;

%-----
%   Computation control file is generated
%-----
[Parameterfile,file_name_PMatlabGP_Controll] = ...
    make_execution_control_file(Computations_to_do,Filename_executable,Parameterfile_ID,Nice_level);

%-----
%   MPI run script file is generated %
%-----
[Used_CPUs Script_filename] = ...
    make_MPIRUN_scriptfile(Filename_executable,Parameterfile_ID,Nr_CPUs,wished_Nr_CPUs)

%-----
%   Matlab function which has to be executed is compiled
%-----
Compile_matlab_code_genEXE(File_to_execute)

%-----
%   Start computation
%-----
unix_command = ['mpirun ' Script_filename ];
[t Shell_output]= system(unix_command);

%*****
%   Computation finished on all nodes *****
%*****
```

Figure F.1.: **Start script of ClusterMatlab.** Main script of the toolbox ClusterMatlab. To make *m*-files run on a LINUX cluster based on ClusterMatlab the user has to adjust 4 parameters of the script file. 1) FileToExecute 2)ParameterfileID 3)WishedNrOfCPU 4)NiceLevel

G. Appendix: Supplementary Material and Figures

G.1. Spike Sorting

Spike sorting of the recording performed in the prefrontal cortex of wake monkeys (see Chapter 9) was performed with a modified MClust version that we developed to run a LINUX cluster (software package: MClust by A. David Redish, University of Minnesota).

We used the following features to describe each recorded spike wave:

1. **Energy** the energy contained within the waveform of the spike. Also known as the L2 norm.
2. **Peak** the maximum height of the waveform of the spike.
3. **Valley** the maximum depth of the waveform of the spike.
4. **Spike width** the width of each spike.
5. **Wave FFT** returns a value based on the fast Fourier transform of the spike waveform.
6. **Wave PC1, PC2** returns for each waveform the contribution to the waveform that is due to the first (wavePC1) or second (wavePC2) principal component.

Based on the extracted features we performed an automatic pre-clustering of the data using the EM cluster algorithm KlustaKwik. KlustaKwik performs an expectation-maximization fit of n Gaussians to the data. This creates a set of putative clusters. It may happen that KlustaKwik splits a single cells data into multiple clusters and may sometimes merge two clusters that are in fact separate cells. We decided to bias the clustering towards high number of clusters (~ 20) and used the results as a pre-classification rather than a classification. This allowed us to deal with banana like shapes of the dot-clouds that often arise by smooth changes during the experiment. Thus, after we pre-classified the data we applied a user based clustering. For this step we also considered the time stamps of each spike that allowed us to check

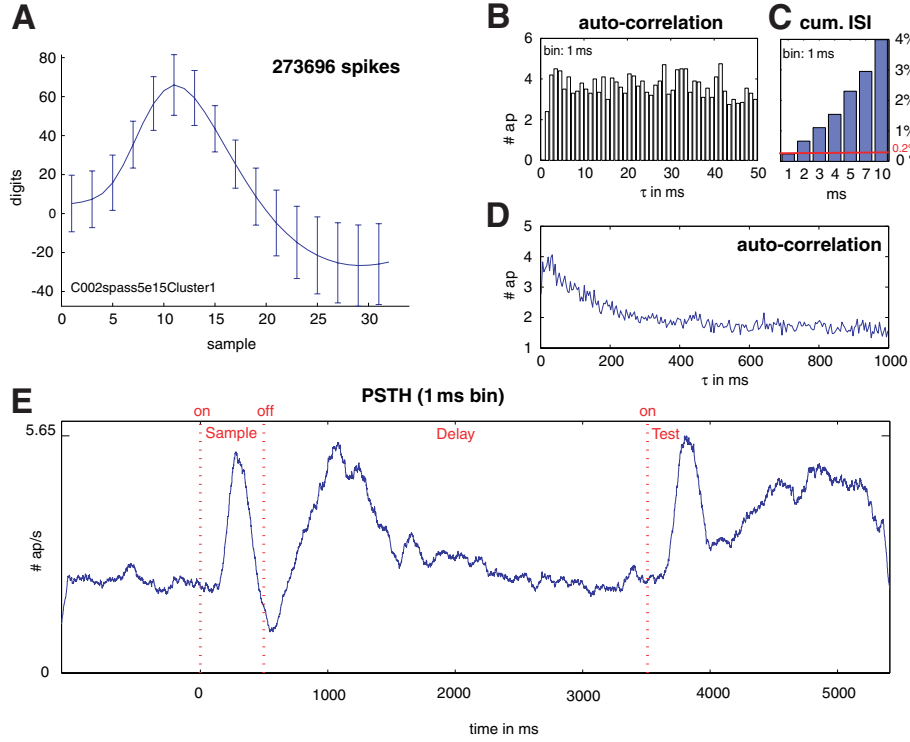


Figure G.1.: **Short-term memory related high complexity JS activity:** (A) Average spike waveform. (B,D) Auto-correlogram. (C) Cumulative interspike interval (ISI) for the first 10 ms. (E) Peri-stimulus-histogram (PSTH)

if changes in features are smooth in time. In case changes are smooth the Gaussian pre-classification gives several clusters distributed along the banana shape which can be jointed by the user of MClust.

To characterize the quality of spike sorting we first used the separability of the data and distinguished between three classes:

1. **Single-unit:** single-units that are clearly separated from other spike waves in the feature space. In addition it was required for single-unit activity that less than 0.5% of all spikes from each individual cluster had interspike intervals of less than 2 ms.
2. **Likely single-unit:** clearly defined cluster, but larger variability between spike waves than for the class of single-units. As for single-unit activity it is required for this class that less than 0.5% of all spikes from each individual cluster had interspike intervals of less than 2 ms.
3. **Multi-unit:** not sortable activity from more than one unit.

We also distinguished between 3 forms of stability:

1. **Stable throughout the session:** clearly defined cluster throughout the whole session.
2. **Stable:** clearly defined cluster throughout most of the session, but the cluster might not exist throughout the whole session. This class was mostly assigned to cells that died or were lost during the recording session.
3. **Unstable:** clearly defined cluster throughout most of the session, but the cluster might cross another cluster at one point in time.

For the analysis we used the classes **single-unit** or **likely single-unit** that were **stable throughout the session**.

G.2. Supplementary Figures

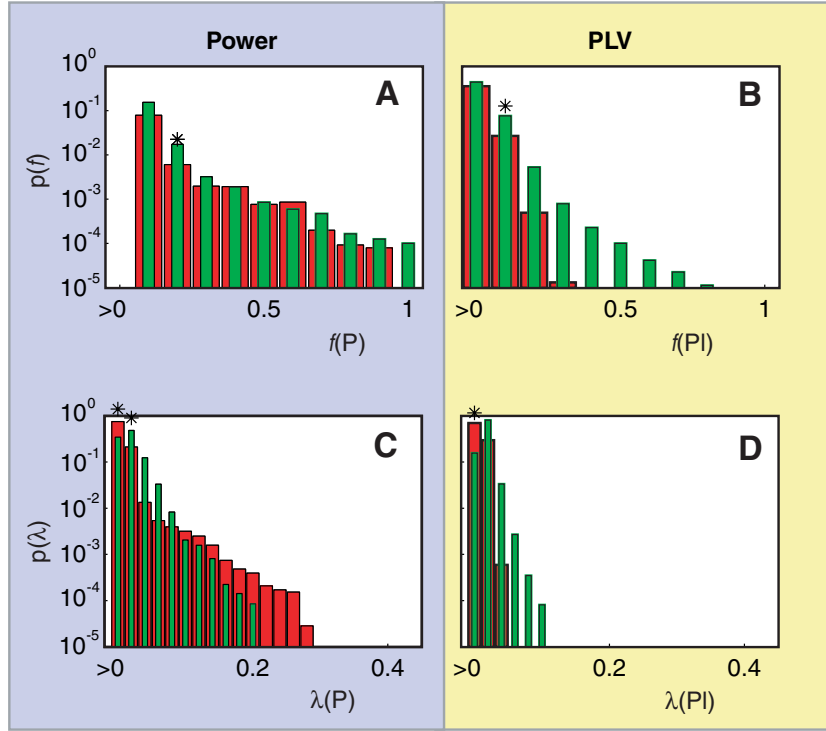


Figure G.2.: **Distributions of unsmoothed f -values and (smoothed) λ -values for power and phase-locking.** Number of channels with significant performance-related differences in power and phase-locking per time-frequency (TF) bin. $p < 0.01$, red and green are data from trials with correct and false responses, respectively. Y-axis is the probability of observing a given f - or λ -value across all TF bins. X-axis represents the f - and λ -values.

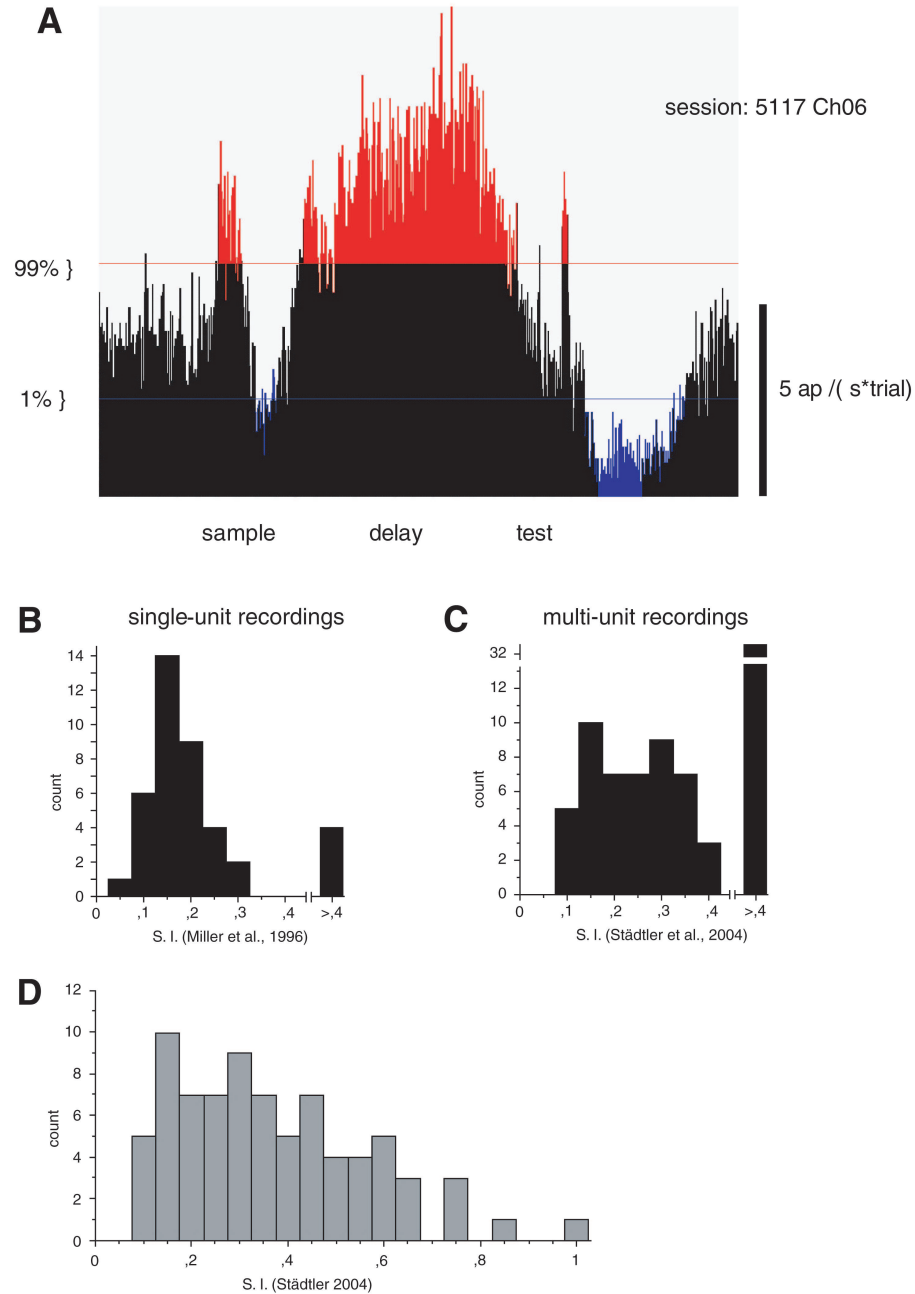


Figure G.3.: **Selectivity of multi-unit responses.** (A) Example of a peri-stimulus time histogram computed from multi-unit activity. Bin width 20 ms, amplitude is normalized to spikes per second and trial. The red and blue line represent the 99th and 1st percentile of the spike count distribution during the baseline (here first second). (B) Distribution of Selectivity Indices (S.I.) for single-unit delay firing as published by Miller et al., 1996.. (C) Distribution of S.I. for multi-unit delay firing recorded in our study and analyzed in the same way as by Miller et al. (ANOVA 5%). (D) The same data shown with the full range of values.

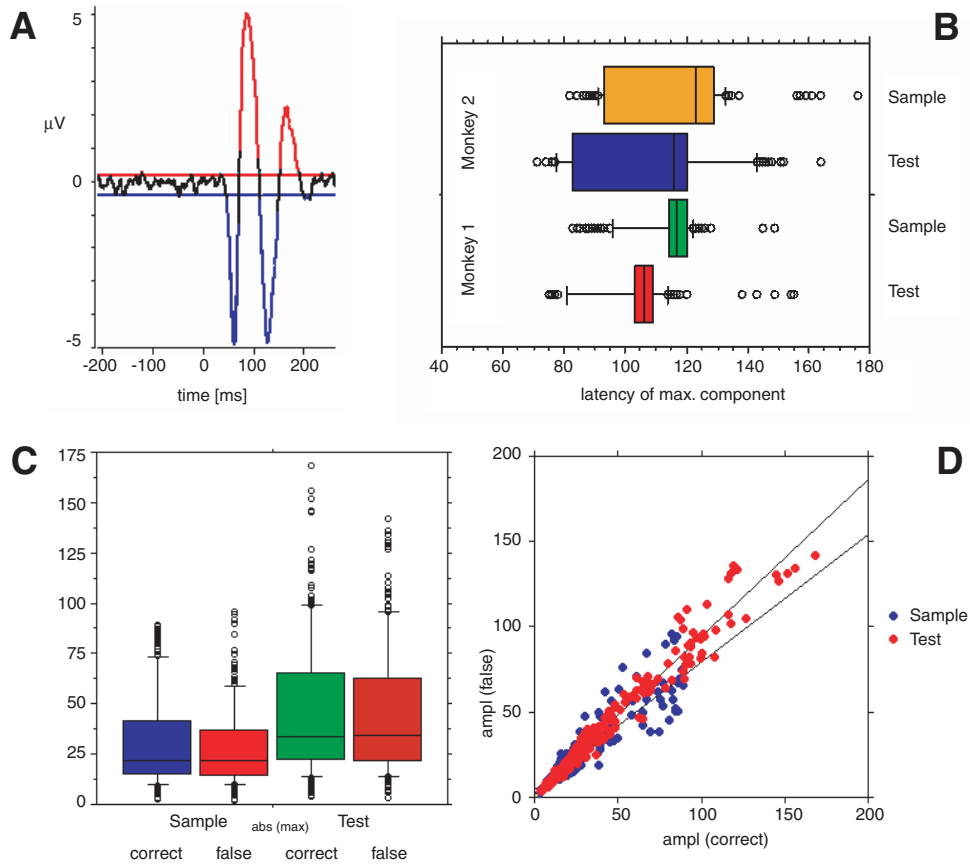


Figure G.4.: **Response strength and timing of evoked responses of LFPs.** (A) Example of a visually evoked field potential from prefrontal cortex. Evoked responses detected by estimating the distribution of the signal amplitude in a pre-stimulus epoch 500 ms before stimulus onset. Responses were considered as significant if 5 or more successive bins exceeded a 1% threshold. (B). Response latencies of the maximal peak of the evoked potentials, grouped for the two monkeys and sample versus test stimulus epochs. (C) Amplitude distributions of the evoked potential responses in trials with correct and incorrect responses, grouped for sample versus test stimulus epochs. (D) Scatter plot for peak evoked amplitudes in trials with correct and incorrect responses. The response amplitudes are reliably correlated across conditions and their regression lines differ significantly from the slope 1 indicating that there is a weak amplitude difference in favor of evoked responses in trials with correct behavioral responses. (Regression Test: $amplitude_{false} = 21,3 + 0,92 amplitude_{correct}$ with $R^2 = 0,958$ / Sample: $amplitude_{false} = 46,1 + 0,74 amplitude_{correct}$ with $R^2 = 0,731$)

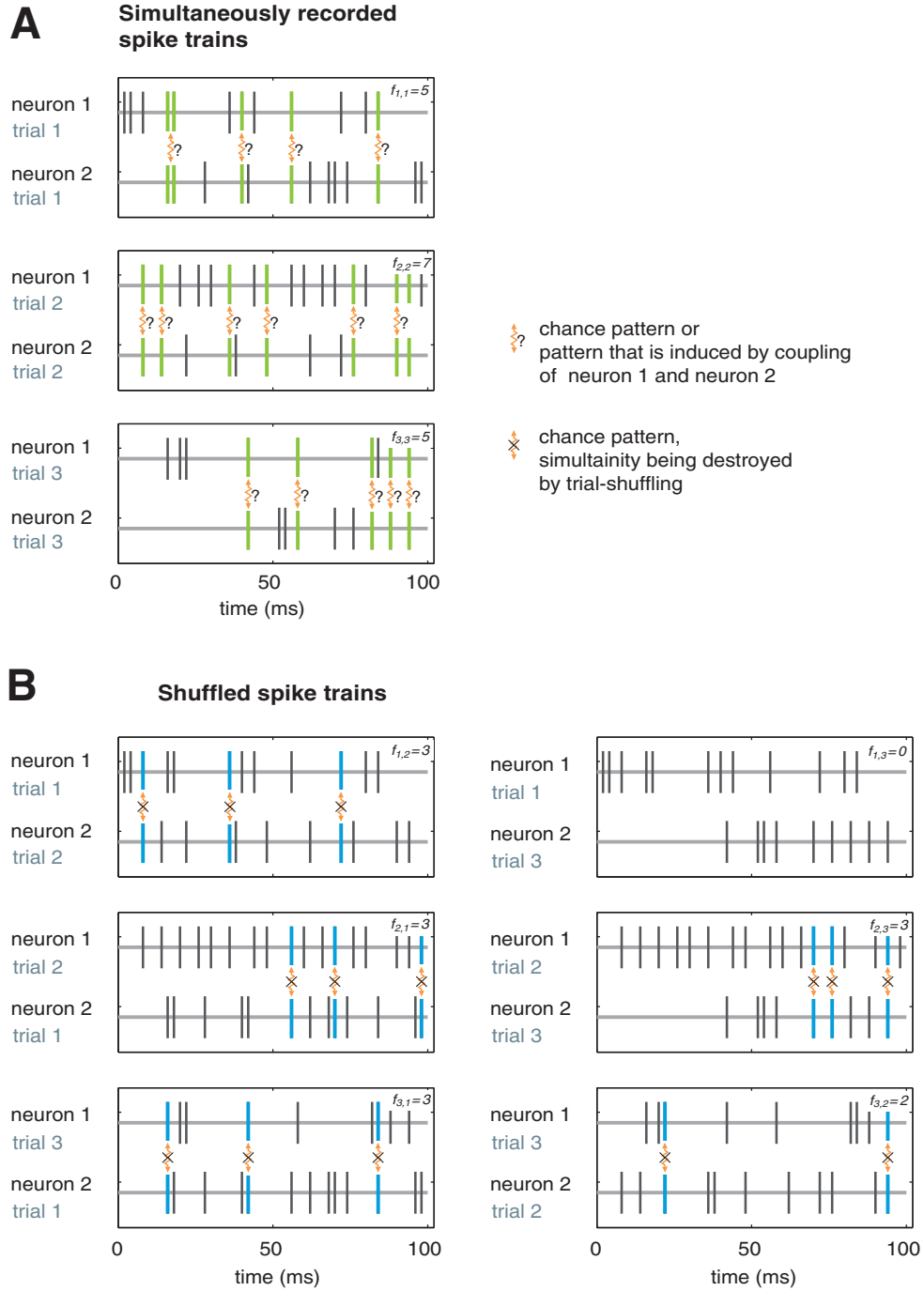


Figure G.5.: **Shuffling of spike trains.** (A) Simultaneously recorded pairs of spike trains. Synchronous spiking is indicated with green spikes. (B) Shuffled spike trains that are built by permutation of spike trains across trials. Synchronous spiking, that only can occur at chance level, is indicated with blue spikes. $f_{i,j}$ indicates the number of the synchronous spiking events per pair of spike trains. The trials are denoted by the indices i and j .

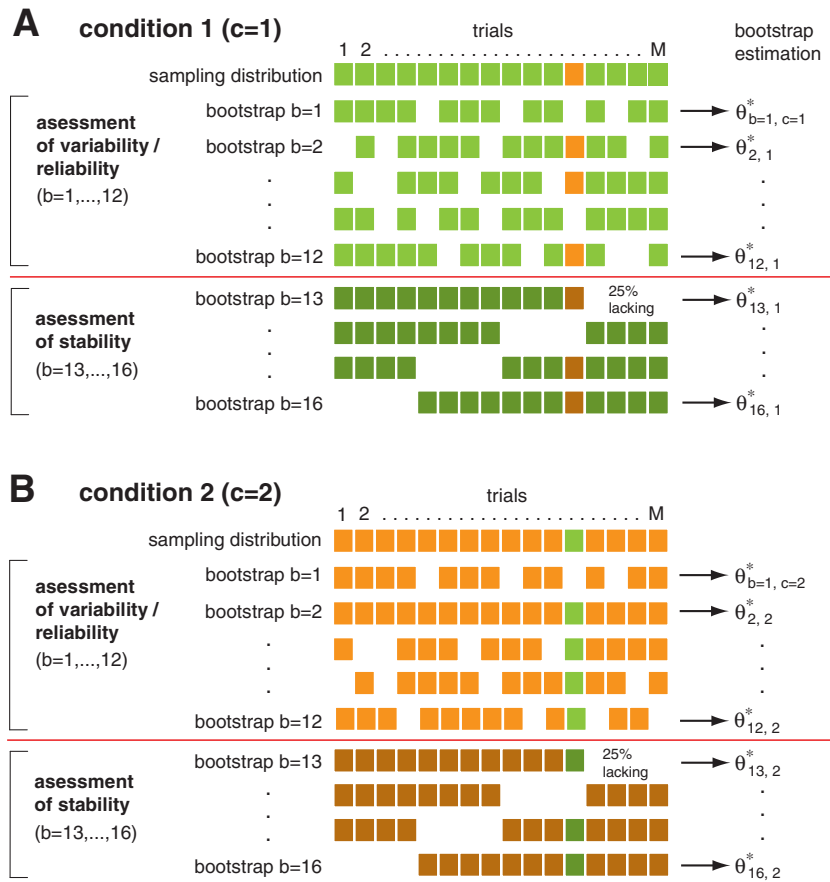


Figure G.6.: **Permutation, bootstrap and n -jackknife of bivariate data.** Second example that illustrates the concept of a permutation test. It shows an extreme case for which only one trial was exchanged by chance for condition 1 and 2. For details see Fig. 5.3 in Chapter 5

List of Variables

$v_i(t)$	number of spikes per bin
b	bin length in ms
Δ_s	sampling period with sampling frequency $f_s = 1/\Delta_s$
Δ_b	sampling period with sampling frequency $f_b = 1/\Delta_b$ and bin width $b = \Delta_b$
$G^{+/-}$	kernel defining the amount of allowed jitter of spike times in a JS pattern
Z	number of bins per analysis window
l	length of analysis window in units of seconds
T	number of trials
N	number of neurons
f	frequency of a JSE
f_{org}	frequency of a JSE in the ordinal data
f_{sur}	frequency of a JSE in the surrogate data
S	number of surrogates
Δf_t^k	difference between f_{org} and f_{sur} for trial t and JS pattern k
$\bar{\Delta} f_t^k$	mean difference between f_{org} and S estimates of f_{sur}
$\bar{\Delta} F^k$	set of $\bar{\Delta} f_t^k$ for M trials
k	index of JS patterns with $k = 1 \dots K$
K	number of distinct JS patterns for each sliding window
η	scaling factor between τ_c and τ_r
τ_c	temporal time scale of JSEs
τ_r	temporal time scale of cross-structure that is considered as rate covariation
r	spike rate in ap/s
r_{MIP}	correlation factor of an MIP process
γ	shape factor of a γ -process ($\gamma < 1$: bursty $\gamma \leq 1$: Poisson $\gamma > 1$: regular)
F	probability distribution
\hat{F}	empirical estimation of F
H	probability distribution of an estimator based on F
H^*	ideal bootstrap estimation of H
\hat{H}^*	bootstrap estimation of H based on the empirical distribution \hat{F}
B	number of bootstrap samples

Θ	test statistics
α	statistical significance value
α_{test}	test level
β	test power
$A'(\nu)^c$	theoretical test power of NeuroXidence for oscillatory rate modulations
ν	frequency of oscillatory rate modulations
c	JS pattern complexity
$\langle f \rangle_c$	expected frequency of JSEs
\mathbf{v}	temporal pattern
H_0	null hypothesis
H_1	alternative hypothesis
k	rank sum
z	z-value = mean / std
$\Theta_m(f, t)$	phase difference between two signals at frequency f in trial m
$c_{xy}(f, t)$	coherence at frequency f between $x(t)$ and $y(t)$
PLV	phase-locking value
$\lambda_s(P)$	percentage of tf-bins with stimulus selective power modulation
$\lambda_c(P)$	percentage of tf-bins with increased power for correct
$\lambda_c(PL)$	percentage of tf-bins with increased PLV for correct
$\lambda_i(P)$	percentage of tf-bins with increased power for incorrect
$\lambda_i(PL)$	percentage of tf-bins with increased PLV for incorrect
$n_{emp}^{total}(c, t)$	total number of occurred JSEs per complexity and sliding window
$\bar{\lambda}(c, t)$	average rate of significant JS pattern
$\lambda_c(c)$	frequency JS patterns with stronger synchrony in correct trials
$\lambda_i(c)$	frequency JS patterns with stronger synchrony in incorrect trials
$\Delta\lambda(c)$	difference of $\lambda_c(c)$ and $\lambda_i(c)$
$\psi(c)_c$	frequency of participation in JS pattern that contributed to $\lambda_c(c)$
$\psi(i)_c$	frequency of participation in JS pattern that contributed to $\lambda_i(c)$
$z(c, t)_{\bar{\lambda}}$	baseline corrected z -score of $\lambda_c(c)$

List of Figures

2.1. Chapter Neuronal Signals and Electrophysiological Recordings: Schematic picture of a recording electrode in tissue	7
2.2. Chapter Neuronal Signals and Electrophysiological Recordings: Multi-unit activity (MUA) and local field potential (LFP) signal	8
2.3. Chapter Neuronal Signals and Electrophysiological Recordings: Recording setup	10
2.4. Chapter Neuronal Signals and Electrophysiological Recordings: Analog signal and detection	11
3.1. Chapter Neuronal Assemblies and Coordinated Neuronal Activity: Different temporal patterns of three neurons.	14
3.2. Chapter Neuronal Assemblies and Coordinated Neuronal Activity: Synchronization of spikes by neuronal coupling	15
3.3. Chapter Neuronal Assemblies and Coordinated Neuronal Activity: $n : m$ Synchronization and Arnold tongues	18
4.1. Chapter Tools to Detect Coordinated Neuronal Activity: Binned representation of N spike trains	24
5.1. Chapter New Resampling Tools for Uni-, Bi- and Multivariate Data: Bootstrap and n -jackknife of one univariate sample	32
5.2. Chapter New Resampling Tools for Uni-, Bi- and Multivariate Data: Bootstrap and n -jackknife of bivariate data	35
5.3. Chapter New Resampling Tools for Uni-, Bi- and Multivariate Data: Permutation, bootstrap and n -jackknife of bivariate data	37
5.4. Expected probability of additional false positive events	38
5.5. Chapter New Resampling Tools for Uni-, Bi- and Multivariate Data: Bootstrap and n -jackknife of multivariate data	39
6.1. Chapter NeuroXidence: Synchronization of spikes by neuronal coupling	42
6.2. Chapter NeuroXidence: Detection of joint spike events ('JSEs')	43

6.3. Chapter NeuroXidence: Total frequency of a test pattern	44
6.4. Chapter NeuroXidence: Impact of multiple surrogates S	50
6.5. Chapter NeuroXidence: Percentage of false positives	54
6.6. Chapter NeuroXidence: Comparisons of test power in relation to the frequency of excess joint spike events (JSEs)	56
6.7. Chapter NeuroXidence: Test power	57
6.8. Chapter NeuroXidence: Test power of NeuroXidence for an induced mother-pattern and its supra-patterns and sub-patterns	59
6.9. Chapter NeuroXidence: Test power NeuroXidence versus Unitary Event ('UE') method	60
6.10. Chapter NeuroXidence: Test power in non-stationary data	62
6.11. Chapter NeuroXidence: Comparison of cross-correlogram and UE method for detecting statistical excess of JSEs	63
6.12. Chapter NeuroXidence: Test power in case of oscillatory rate modulations	66
6.13. Chapter NeuroXidence: Computational complexity of NeuroXidence	67
7.1. Chapter High Complexity Joint Spike Activity in Cat Visual Area 17: Analysis of simultaneously recorded single-unit activity from 48 neurons from an anaesthetized cat	74
8.1. Chapter Short-Term Memory Related LFP Oscillation in the Prefrontal Cor- tex: Time course of the visual short-term memory task	78
8.2. Chapter Short-Term Memory Related LFP Oscillation in the Prefrontal Cor- tex: MRI-based reconstruction of recording positions	79
8.3. Chapter Short-Term Memory Related LFP Oscillation in the Prefrontal Cor- tex: Oscillations and time-frequency responses in prefrontal cortex during a visual STM task	80
8.4. Chapter Short-Term Memory Related LFP Oscillation in the Prefrontal Cor- tex: Grand average time-frequency ('TF') plots for power and phase-locking	81
8.5. Chapter Short-Term Memory Related LFP Oscillation in the Prefrontal Cor- tex: Performance-dependent increases of β - and γ -oscillations and their phase-locking in prefrontal field potentials	83
8.6. Chapter Short-Term Memory Related LFP Oscillation in the Prefrontal Cor- tex: Control for induced power by eye-movements	87
8.7. Chapter Short-Term Memory Related LFP Oscillation in the Prefrontal Cor- tex: Stimulus-selective and performance-dependent increases of β - and γ -oscillations in prefrontal field potentials	89

9.1. Chapter Short-Term Memory-Related High Complexity Joint Spike Activity:	
Introductory figure for the analysis of MUA during a visual short-term memory task . . .	94
9.2. Chapter Short-Term Memory-Related High Complexity Joint Spike Activity:	
Modulations and task-related modulations of $\bar{\lambda}(c, t)$ of JS patterns during a short-term memory task	96
9.3. Chapter Short-Term Memory-Related High Complexity Joint Spike Activity:	
Task and performance-related modulations of $\lambda_c(c)$ and $\lambda_i(c)$ of JS patterns during a short-term memory task	98
9.4. Chapter Short-Term Memory-Related High Complexity Joint Spike Activity:	
Task and performance-related modulations of $\lambda_c(c)$ and $\lambda_i(c)$ of JS patterns during the sample presentation of a short-term memory task	99
9.5. Chapter Short-Term Memory-Related High Complexity Joint Spike Activity:	
Neuronal groups with performance dependent modulations of synchrony ($\psi(c)_c$ and $\psi(i)_c$)	100
9.6. Chapter Short-Term Memory-Related High Complexity Joint Spike Activity:	
Neuronal groups with performance dependent and task-related modulations of synchronization ($\psi(c)_c$ and $\psi(i)_c$)	102
9.7. Chapter Short-Term Memory-Related High Complexity Joint Spike Activity:	
Task and performance-related modulations of $\lambda_c(c)$ and $\lambda_i(c)$ of JS patterns during test-stimulus presentation of a short-term memory task	104
9.8. Chapter Short-Term Memory-Related High Complexity Joint Spike Activity:	
Task and motor-related modulations of $\lambda_r(c)$ and $\lambda_l(c)$ of JS patterns during test-stimulus presentation of a short-term memory task	105
9.9. Chapter Short-Term Memory-Related High Complexity Joint Spike Activity:	
Task and performance-related modulations of $\lambda_c(c)$ and $\lambda_i(c)$ of JS patterns during a short-term memory task	107
A.1. Appendix - NeuroXidence Methods: Two preprocessing steps of NeuroXidence . . .	128
B.1. Appendix - Short-Term Memory Related LFP Oscillation in the Prefrontal Cortex: Inter-subject comparison of λ -maps of correct responses (λ_c -maps)	132
C.1. Appendix - High Complexity JS Activity in Cat Visual Area 17: Orientation tuning of recorded channels	136
C.2. Appendix - High Complexity JS Activity in Cat Visual Area 17: Peri-stimulus-histogram (PSTH) recorded in 17	137
C.3. Appendix - High Complexity JS Activity in Cat Visual Area 17: Shift-corrected cross-correlograms	138

D.1. Appendix - Hypothesis Tests: Single sided hypothesis test	140
D.2. Appendix - Hypothesis Tests: Two sided hypothesis test	141
D.3. Appendix - Hypothesis Tests: False positives, false negatives and test power	142
E.1. Appendix - Bootstrapping and Resampling: One toss of eight dice	148
E.2. Appendix - Bootstrapping and Resampling: Resampling of $\hat{F}(x)$	148
E.3. Appendix - Bootstrapping and Resampling: Theoretical and bootstrapped distribution of y	149
E.4. Appendix - Bootstrapping and Resampling: Basic principals of the bootstrapping idea	150
F.1. Appendix - Parallel and Load Balanced Computation with ClusterMatlab: Start script of ClusterMatlab	154
G.1. Appendix - Supplementary Material and Figures: Short-term memory related high complexity JS activity	156
G.2. Appendix - Supplementary Material and Figures: Distributions of unsmoothed f-values and (smoothed) λ -values for power and phase-locking	157
G.3. Appendix - Supplementary Material and Figures: Selectivity of multi-unit responses	158
G.4. Appendix - Supplementary Material and Figures: Response strength and timing of evoked responses of LFPs	159
G.5. Appendix - Supplementary Material and Figures: Shuffling of spike trains	160
G.6. Appendix - Supplementary Material and Figures: Permutation, bootstrap and n -jackknife of bivariate data	161

Bibliography

- Abeles M, Bergman H, Margalit E, Vaadia E (1993) Spatiotemporal firing patterns in the frontal cortex of behaving monkeys. *J. Neurophysiol.* 70:1629–1638.
- Abeles M, Gerstein GL (1988) Detecting spatiotemporal firing patterns among simultaneously recorded single neurons. *J. Neurophysiol.* 60:909–924.
- Abeles M (1991) *Corticonics: Neural Circuits of the Cerebral Cortex* Cambridge University Press, Cambridge, 1 edition.
- Abeles M, Goldstein MH (1977) Multispike train analysis. *Proc. IEEE* 65:762–773.
- Adrian E (1928) The basis of sensation: The action of the sense organ. *London: Christophers* .
- Aertsen A, Diesmann M, Gewaltig MO, Grün S, Rotter S (2001) Neural dynamics in cortical networks—precision of joint-spiking events. *Novartis. Found. Symp.* 239:193–204.
- Aertsen AM, Gerstein GL, Habib MK, Palm G (1989) Dynamics of neuronal firing correlation: modulation of "effective connectivity". *J. Neurophysiol.* 61:900–917.
- Aertsen A, Gerstein GL (1985) Evaluation of neuronal connectivity: Sensitivity of cross-correlation. *Brain Research* 340:341–354.
- Albert R, Jeong H, Barabási AL (1999) Diameter of the world wide web. *Nature* 401:130–131.
- Albert R, Jeong H, Barabási AL (2000) Error and attack tolerance of complex networks. *Nature* 406:378–482.
- Arieli A, Shoham D, Hildesheim R, Grinvald A (1995) Coherent spatiotemporal patterns of ongoing activity revealed by real-time optical imaging coupled with single-unit recording in the cat visual cortex. *J. Neurophysiol.* 73:2072–2093.
- Azouz R, Gray CM (1999) Cellular mechanisms contributing to response variability of cortical neurons in vivo. *J. Neurosci.* 19:2209–2223.
- Baker SN, Gerstein GL (2001) Determination of response latency and its application to normalization of cross-correlation measures. *Neural Comput.* 13:1351–1377.

- Baker S, Lemon R (2000) Precise spatiotemporal repeating patterns in monkey primary and supplementary motor areas occur at chance level. *J. Physiol. (Lond)* 84:1770–1780.
- Baker SN, Gerstein GL (2000) Improvements to the sensitivity of gravitational clustering for multiple neuron recordings. *Neural Comp.* 12:2597–2620.
- Barabási AL (2005) Taming complexity. *Nature Physics* 1:68–70.
- Barabási AL, Albert R (1999) Emergence of scaling in random networks. *Science* 286:509–512.
- Barabási AL, Albert R, Jeong H (1999) Mean-field theory for scale-free random networks. *Physica A* 272:173–187.
- Barbieri R, Frank LM, Nguyen DP, Quirk MC, Solo V, Wilson MA, Brown EN (2004) Dynamic analyses of information encoding in neural ensembles. *Neural Comput.* 16:277–307.
- Barbieri R, Wilson MA, Frank LM, Brown EN (2005) An analysis of hippocampal spatio-temporal representations using a bayesian algorithm for neural spike train decoding. *IEEE Trans. Neural Syst. Rehabil. Eng* 13:131–136.
- Barlow HB (1972) Single units and sensation: a neuron doctrine for perceptual psychology? *Perception* 1:371–394.
- Bernander O, Koch C, Usher M (1994) The effect of synchronized inputs at the single neuron level. *Neural Comp.* 6:622–641.
- Bi GQ, Poo MM (1998) Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. *J. Neurosci.* 18:10464–10472.
- Bialek W, Rieke F (1992) Reliability and information transmission in spiking neurons. *TINS* 15:428–434.
- Bialek W, Rieke F, De Ruyter van Steveninck RR, Warland D (1991) Reading a neural code. *Science* 252:1854–1857.
- Bichot NP, Rossi AF, Desimone R (2005) Parallel and serial neural mechanisms for visual search in macaque area v4. *Science* 308:529–534.
- Bohte SM, Spekreijse H, Roelfsema PR (2000) The effects of pair-wise and higher-order correlations on the firing rate of a postsynaptic neuron. *Neural Comp.* 12:153–179.
- Boneau CA (1960) The effects of violations of assumptions underlying the t-test. *Psychological Bulletin* 57:49–64.
- Braitenberg V, Schüz A (1998) *Cortex: Statistics and Geometry of Neuronal Connectivity* Springer-Verlag, Berlin, 2nd edition.

- Brody CD (1999) Correlations without synchrony. *Neural Comp.* 11:1537–1551.
- Brovelli A, Ding M, Ledberg A, Chen Y, Nakamura R, Bressler SL (2004) Beta oscillations in a large-scale sensorimotor cortical network: directional influences revealed by granger causality. *Proc. Natl. Acad. Sci. U. S. A* 101:9849–9854.
- Brown EN, Frank LM, Tang D, Quirk MC, Wilson MA (1998) A statistical paradigm for neural spike train decoding applied to position prediction from ensemble firing patterns of rat hippocampal place cells. *J. Neurosci.* 18:7411–7425.
- Brown EN, Kass RE, Mitra PP (2004) Multiple neural spike train data analysis: state-of-the-art and future challenges. *Nat Neurosci.* 7:456–461.
- Buzsáki G, Draguhn A (2004) Neuronal oscillations in cortical networks. *Science* 304:1926–1929.
- Castelo-Branco M, Goebel R, Neuenschwander S, Singer W (2000) Neural synchrony correlates with surface segregation rules. *Nature* 8:685–689.
- Compte A, Brunel N, Goldman-Rakic PS, Wang XJ (2000) Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cereb. Cortex* 10:910–923.
- Compte A, Constantinidis C, Tegner J, Raghavachari S, Chafee MV, Goldman-Rakic PS, Wang XJ (2003) Temporally irregular mnemonic persistent activity in prefrontal neurons of monkeys during a delayed response task. *J. Neurophysiol.* 90:3441–3454.
- Cox DR, Isham V (1980) *Point Processes* Monographs on Applied Probability and Statistics. Chapman and Hall.
- Czanner G, Grün S, Iyengar S (2005) Theory of the snowflake plot and its relations to higher-order analysis methods. *Neural Comput.* 17:1456–1479.
- Durstewitz D, Seamans JK, Sejnowski TJ (2000) Neurocomputational models of working memory. *Nat. Neurosci.* 3 Suppl:1184–1191.
- Eckhorn R, Bauer R, Jordan W, Brosch M, Kruse W, Munk M, Reitböck HJ (1988) Coherent oscillations: A mechanism of feature linking in the visual cortex? *Biol. Cyber.* 60:121–130.
- Efron B, Tibshirani RJ (1993) *An Introduction to the Bootstrap* Chapman and Hall/CRC, London, 1 edition.
- Engel A, König P, Kreiter A, Singer W (1991) Interhemispheric synchronization of oscillatory responses in cat visual cortex. *Science* 252:1177–1179.

- Fabre-Thorpe M, Richard G, Thorpe SJ (1998) Rapid categorization of natural images by rhesus monkeys. *Neuroreport* 9:303–308.
- Feigenbaum MJ (1978) Quantitative universality for a class of non-linear transformations. *J. Stat. Phys.* 19:25–52.
- Feigenbaum MJ (1979) The universal metric properties of nonlinear transformations. *J. Stat. Phys.* 21:669–706.
- Feller W (1968) *An Introduction to Probability Theory and Its Applications*, Vol. 1 John Wiley & Sons, New York, 3 edition.
- Feng J, Brown D (1998) Spike output jitter, mean firing time and coefficient of variation. *J. Phys. A* 31:1239–1252.
- Fetz EE (1997) Temporal coding in neural populations? *Science* 278:1901–1902.
- Fries P, Neuenschwander S, Engel AK, Goebel R, Singer W (2001a) Rapid feature selective neuronal synchronization through correlated latency shifting. *Nat Neurosci.* 4:194–200.
- Fries P, Reynolds JH, Rorie AE, Desimone R (2001b) Modulation of oscillatory neuronal synchronization by selective visual attention. *Science* 291:1560–1563.
- Fries P, Roelfsema PR, Engel AK, König P, Singer W (1997) Synchronization of oscillatory responses in visual cortex correlates with perception in interocular rivalry. *Proc. Natl. Acad. Sci. U. S. A* 94:12699–12704.
- Fries P, Schröder JH, Roelfsema PR, Singer W, Engel AK (2002) Oscillatory neuronal synchronization in primary visual cortex as a correlate of stimulus selection. *J. Neurosci.* 22:3739–3754.
- Fries P, Schröder JH, Singer W, Engel AK (2001) Conditions of perceptual selection and suppression during interocular rivalry in strabismic and normal cats. *Vision Res.* 41:771–783.
- Fuster JM, Alexander GE (1971) Neuron activity related to short-term memory. *Science* 173:652–654.
- Georgopoulos AP, Kettner RE, Schwartz AB (1988) Primate motor cortex and free arm movements to visual targets in three-dimensional space. II. coding of the direction of the movement by a neuronal population. *J. Neurosci.* 8:2928–2937.
- Gerstein GL, Bedenbaugh P, Aertsen MH (1989) Neuronal assemblies. *IEEE Trans. Biomed. Eng* 36:4–14.
- Gerstein GL, Perkel DH (1969) Simultaneously recorded trains of action potentials: analysis and functional interpretation. *Science* 164:828–830.

- Gerstein GL, Perkel DH (1972) Mutual temporal relationships among neuronal spike trains. statistical techniques for display and analysis. *Biophys. J.* 12:453–473.
- Gerstner W, Kreiter AK, Markram H, Herz VM (1997) Neural codes: Firing rates and beyond. *Proc. Nat. Acad. Sci. USA* 94:12740–12741.
- Goldman-Rakic PS (1995) Cellular basis of working memory. *Neuron* 14:477–485.
- Golomb D, Rubin N, Sompolinsky H (1990) Willshaw model - associative memory with sparse coding and low firing rates. *Physical Review A* 41:1843–1854.
- Gray CM, König P, Engel AK, Singer W (1989) Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature* 338:334–337.
- Gray CM, Singer W (1987) Stimulus-specific neuronal oscillations in the cat visual cortex: A cortical functional. *Soc. Neurosci. Abstr.* 13 404.3.
- Gray CM, Singer W (1989) Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *Proc. Nat. Acad. Sci. USA* 86:1698–1702.
- Gray C, Maldonado P, Wilson M, , McNaughton B (1995) Tetrodes markedly improve the reliability and yield of multiple single-unit isolation from multi-unit recordings in cat striate cortex. *Journal of Neuroscience Methods* 63:43–54.
- Grün S, Diesmann M, Aertsen A (2002a) Unitary events in multiple single-neuron spiking activity: I. detection and significance. *Neural Comput.* 14:43–80.
- Grün S, Diesmann M, Aertsen A (2002b) Unitary events in multiple single-neuron spiking activity: II. nonstationary data. *Neural Comput.* 14:81–119.
- Grün S, Diesmann M, Grammont F, Riehle A, Aertsen A (1999) Detecting unitary events without discretization of time. *J. Neurosci. Methods* 94:67–79.
- Grün S, Riehle A, Diesmann M (2003) Effect of cross-trial nonstationarity on joint-spike events. *Biol. Cybern.* 88:335–351.
- Grün S (1996) *Unitary Joint-Events in Multiple-Neuron Spiking Activity: Detection, Significance, and Interpretation* Reihe Physik, Band 60. Verlag Harri Deutsch, Thun, Frankfurt/Main.
- Gur M, Beylin A, Snodderly M (1997) Response variability of neurons in primary visual cortex (V1) of alert monkeys. *J. Neurosci.* 17:2914–2920.

- Gutkin BS, Laing CR, Colby CL, Chow CC, Ermentrout GB (2001) Turning on and off with excitation: the role of spike-timing asynchrony and synchrony in sustained neural activity. *J. Comput. Neurosci.* 11:121–134.
- Haeusler S, Maass W (2006) A statistical analysis of information-processing properties of lamina-specific cortical microcircuit models. *Cereb. Cortex* .
- Harris H C, Hirase, Buzsaki (2000) Accuracy of tetrode spike sorting as determined by simultaneous intracellular measurements. *J. Physiol. (Lond)* 84:401–414.
- Harris K, Henze D, Csicsvari J, Hirase H, , Buzsaki G (2000) Accuracy of tetrode spike separation as determined by simultaneous intracellular and extracellular measurements. *J Neurophysiol* 84:401–414.
- Hebb DO (1949) *Organization of behavior. A neurophysiological theory* John Wiley & Sons, New York.
- Hodgkin AL, Huxley AF (1939) Action potentials recorded from inside a nerve fibre. *Nature* 144:710–711.
- Hopfield JJ (1995) Pattern recognition computation using action potential timing for stimulus representation. *Nature* 376:33–36.
- Hopfield JJ, Brody CD (2000) What is a moment? ”cortical” sensory integration over a brief interval. *Proc. Natl. Acad. Sci. U. S. A* 97:13919–13924.
- Hopfield JJ, Brody CD (2001) What is a moment? transient synchrony as a collective mechanism for spatiotemporal integration. *Proc. Natl. Acad. Sci. U. S. A* 98:1282–1287.
- Ikegaya Y, Aaron G, Cossart R, Aronov D, Lampl I, Ferster D, Yuste R (2004) Synfire chains and cortical songs: temporal modules of cortical activity. *Science* 304:559–564.
- Izhikevich EM, Gally JA, Edelman GM (2004) Spike-timing dynamics of neuronal groups. *Cereb. Cortex* 14:933–944.
- Jaeger H (2002) Adaptive nonlinear system identification with echo state networks Proc. NIPS 02.
- Kaske A, Maass W (2005) A model for the interaction of oscillations and pattern generation with real-time computing in generic neural microcircuit models. *Neural Netw.* .
- Kass RE, Ventura V, Brown EN (2005) Statistical issues in the analysis of neuronal data. *J. Neurophysiol.* 94:8–25.
- Koch C (1999) *Biophysics of computation: information processing in single neurons* Oxford University Press.
- König P (1994) A method for the quantification of synchrony and oscillatory properties of neuronal activity. *J. Neurosci. Methods* 54:31–37.

- Kreiter A, Singer W (1996) Stimulus-dependent synchronization of neuronal responses in the visual cortex of awake macaque monkey. *J. Neurosci.* 16:2381–2396.
- Kuhn A, Aertsen A, Rotter S (2003) Higher-order statistics of input ensembles and the response of simple model neurons. *Neural Comput.* 15:67–101.
- Lachaux JP, Rodriguez E, Martinerie J, Adam C, Hasboun D, Varela FJ (2000) A quantitative study of gamma-band activity in human intracranial recordings triggered by visual stimuli. *Eur. J. Neurosci.* 12:2608–2622.
- Lachaux JP, Rodriguez E, Martinerie J, Varela FJ (1999) Measuring phase synchrony in brain signals. *Hum. Brain Mapp.* 8:194–208.
- Lazar A, Pipa G, Triesch J (2006) The combination of stdp and intrinsic plasticity yields complex dynamics in recurrent spiking networks.
- Lee H, Simpson GV, Logothetis NK, Rainer G (2005) Phase locking of single neuron activity to theta oscillations during working memory in monkey extrastriate visual cortex. *Neuron* 45:147–156.
- Legenstein R, Naeger C, Maass W (2005) What can a neuron learn with spike-timing-dependent plasticity? *Neural Comput.* 17:2337–2382.
- Lestienne R, Tuckwell H (1998) The significance of precisely replicating patterns in mammalian cns spike trains. *Neuroscience* 82:315–336 single units!
- Lestienne R (1995) Determination of the precision of spike timing in the visual cortex of anaesthetised cats. *Biol. Cyber.* 74:55–61 single units!!
- Lewicki M (1998) A review of methods for spike sorting: the detection and classification of neural action potentials. *Network: Computation in Neural Systems* 9:R53–R78.
- Lisman JE, Fellous JM, Wang XJ (1998) A role for nmda-receptor channels in working memory. *Nat. Neurosci.* 1:273–275.
- Lisman JE, Idiart MA (1995) Storage of 7 ± 2 short-term memories in oscillatory subcycles. *Science* 267:1512–1515.
- Maass W, Natschlager T, Markram H (2002) Real-time computing without stable states: a new framework for neural computation based on perturbations. *Neural Comput.* 14:2531–2560.
- Maass W, Natschlager T, Markram H (2004) Fading memory and kernel properties of generic cortical microcircuit models. *J. Physiol Paris* 98:315–330.

- Machens CK, Romo R, Brody CD (2005) Flexible control of mutual inhibition: a neural model of two-interval discrimination. *Science* 307:1121–1124.
- MacLeod K, Bäcker A, Laurent G (1998) Who reads temporal information contained across synchronized and oscillatory spike trains. *Nature* 395:693–698.
- Mainen ZF, Sejnowski TJ (1995) Reliability of spike timing in neocortical neurons. *Science* 268:1503–1506.
- Mandelbrot B (1967) How long is the coast of great britain, statistical self similarity and fractional dimension. *Science* 155:636–638.
- Mandelbrot B (1973) *The Fractal Geometry of Nature* W. H. Freeman and Co., New York.
- Markram H, Lubke J, Frotscher M, Sakmann B (1997) Regulation of synaptic efficacy by coincidence of postsynaptic aps and epsps. *Science* 275:213–215.
- Markram H, Tsodyks M (1996) Redistribution of synaptic efficacy between neocortical pyramidal neurons. *Nature* 382:807–10.
- Marsalek P, Koch C, Maunsell J (1997) On the relationship between synaptic input and spike output jitter in individual neurons. *Proc. Nat. Acad. Sci. USA* 94:735–740.
- Martignon L, Deco G, Laskey K, Diamond M, Freiwald W, Vaadia E (2000) Neural coding: higher-order temporal patterns in the neurostatistics of cell assemblies. *Neural Comput.* 12:2621–2653.
- Melamed O, Gerstner W, Maass W, Tsodyks M, Markram H (2004) Coding and learning of behavioral sequences. *Trends Neurosci.* 27:11–14.
- Miller EK, Erickson CA, Desimone R (1996) Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *Journal of Neuroscience* 16:5154–5167.
- Miller EK, Li L, Desimone R (1993) Activity of neurons in anterior inferior temporal cortex during a short-term memory task. *Journal of Neuroscience* 13:1460–1478.
- Mitra PP, Pesaran B (1999) Analysis of dynamic brain imaging data. *Biophys. J.* 76:691–708.
- Nadasdy Z, Csicsvari J, Penttonen M, Hetke J, Wise K, Buzsáki G (1998) Extracellular recording and analysis of neuronal activity: from single cells to ensembles In *Strategies for Recording and Processing: Neuronal Ensembles*, pp. 17–55. Wiley press.
- Nakahara H, Amari S (2002) Information-geometric measure for neural spikes. *Neural Comput.* 14:2269–2316.

- Natschläger T, Maass W (2005) Dynamics of information and emergent computation in generic neural microcircuit models. *Neural Netw.* 18:1301–1308.
- Neda Z, Ravasz E, Brechet Y, Vicsek T, Barabási AL (2000) Self-organizing processes: The sound of many hands clapping. *Nature* 403:849–850.
- Niebur E, Hsiao SS, Johnson KO (2002) Synchrony: a neuronal mechanism for attentional selection? *Curr. Opin. Neurobiol.* 12:190–194.
- Okatan M, Wilson MA, Brown EN (2005) Analyzing functional connectivity using a network likelihood model of ensemble neural spiking activity. *Neural Comput.* 17:1927–1961.
- Oram MW, Perrett DI (1992) Time course of neural responses discriminating different views of the face and head. *J. Neurophysiol.* 68:70–84.
- Oram MW, Wiener MC, Lestienne R, Richmond BJ (1999) Stochastic nature of precisely timed spike patterns in visual system neuronal responses. *J. Neurophysiol.* 81:3021–3033.
- Palva JM, Palva S, Kaila K (2005) Phase synchrony among neuronal oscillations in the human cortex. *J. Neurosci.* 25:3962–3972.
- Percival DB, Walden AT (1993) *Spectral Analysis for Physical Applications: Multitaper and Conventional Univariate Techniques* Cambridge University Press.
- Pesaran B, Pezaris JS, Sahani M, Mitra PP, Andersen RA (2002) Temporal structure in neuronal activity during working memory in macaque parietal cortex. *Nat. Neurosci.* 5:805–811.
- Pikovsky A, Rosenblum M, Kurths J (2001) *Synchronization* Cambridge University Press.
- Pipa G (2001) Diplomarebit: Entwicklung und Untersuchung einer nicht-parametrischen Methode zur Schätzung der Signifikanz zeitlich koordinierter Spike-Aktivität Thesis/dissertation, Max-Planck Institute for Brainresearch, Neurophysiology, Frankfurt/M, Germany; J.W.Goethe-Universität, Frankfurt/M, Germany.
- Pipa G, Grün S (2003) Non-parametric significance estimation of joint-spike events by shuffling and resampling. *Neurocomputing* 52-54:31–37.
- Prut Y, Vaadia E, Bergman H, Haalman I, Hamutal S, Abeles M (1998) Spatiotemporal structure of cortical activity: Properties and behavioral relevance. *J. Neurophysiol.* 79:2857–2874.
- Pucak ML, Levitt JB, Lund JS, Lewis DA (1996) Patterns of intrinsic and associational circuitry in monkey prefrontal cortex. *J. Comp Neurol.* 376:614–630.

- Radons G, Becker DJ, Dülfer B, Krüger J (1994) Analysis, classification and coding of multielectrode spike trains with hidden markov models. *Biol. Cyber.* 71:359–373.
- Rainer G, Miller EK (2000) Effects of visual experience on the representation of objects in the prefrontal cortex. *Neuron* 27:179–189.
- Riehle A, Grün S, Diesmann M, Aertsen A (1997) Spike synchronization and rate modulation differentially involved in motor cortical function. *Science* 278:1950–1953.
- Rodriguez E, George N, Lachaux JP, Martinerie J, Renault B, Varela FJ (1999) Perception’s shadow: long-distance synchronization of human brain activity. *Nature* 397:430–433.
- Roelfsema PR, Engel AK, König P, Singer W (1997) Visuomotor integration is associated with zero time-lag synchronization among cortical areas. *Nature* 385:157–161.
- Roy A, Steinmetz PN, Niebur E (2000) Rate limitations of unitary event analysis. *Neural Comput.* 12:2063–2082.
- Salinas E, Sejnowski TJ (2001) Correlated neuronal activity and the flow of neuronal information. *Nat. Neurosci.* 2:539–550.
- Samonds JM, Bonds AB (2004) Real-time visualization of neural synchrony for identifying coordinated cell assemblies. *Journal of Neuroscience Methods* 139:51–60.
- Schölkopf B, Smola A (2002) *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond (Adaptive Computation and Machine Learning)* MIT Press, Cambridge, MA,.
- Schreiber T, Schmitz A (2000) Surrogate time series. *Physica D-Nonlinear Phenomena* 142:346–382.
- Seamans JK, Durstewitz D, Christie BR, Stevens CF, Sejnowski TJ (2001) Dopamine d1/d5 receptor modulation of excitatory synaptic inputs to layer v prefrontal cortex neurons. *Proc. Natl. Acad. Sci. U. S. A* 98:301–306.
- Shadlen MN, Newsome WT (1994) Noise, neural codes and cortical organization. *Curr. Opin. Neurobiol.* 4:569–579.
- Shadlen MN, Newsome WT (1998) The variable discharge of cortical neurons: Implications for connectivity, computation, and information coding. *J. Neurosci.* 18:3870–3896.
- Sharpee T, Rust NC, Bialek W (2004) Analyzing neural responses to natural signals: maximally informative dimensions. *Neural Comput.* 16:223–250.
- Singer W (1999) Neuronal synchrony: a versatile code for the definition of relations? *Neuron* 24:49–25.

- Singer W, Gray C, Engel A, König P (1988) Spatio-temporal distribution of stimulus-specific oscillations in the cat visual cortex ii: Global interactions. *Soc. Neurosci. Abstr.* 14:899.
- Singer W (1993) Synchronization of cortical activity and its putative role in information processing and learning. *Annu. Rev. Physiol.* 55:349–374.
- Sjostrom PJ, Turrigiano GG, Nelson SB (2001) Rate, timing, and cooperativity jointly determine cortical synaptic plasticity. *Neuron* 32:1149–1164.
- Softky WR (1994) Sub-millisecond coincidence detection in active dendritic trees. *Neuroscience* 58:13–41.
- Softky WR, Koch C (1993) The highly irregular firing of cortical cells is inconsistent with temporal integration of random epsps. *J. Neurosci.* 13:334–350.
- Sommer FT, Wennekers T (2001) Associative memory in networks of spiking neurons. *Neural Netw.* 14:825–834.
- Steinmetz PN, Roy A, Fitzgerald PJ, Hsiao SS, Johnson KO, Niebur E (2000) Attention modulates synchronized neuronal firing in primate somatosensory cortex. *Nature* 404:187–190.
- Tallon-Baudry C, Kreiter A, Bertrand O (1999) Sustained and transient oscillatory responses in the gamma and beta bands in a visual short-term memory task in humans. *Vis. Neurosci.* 16:449–459.
- Tallon-Baudry C, Mandon S, Freiwald WA, Kreiter AK (2004) Oscillatory synchrony in the monkey temporal lobe correlates with performance in a visual short-term memory task. *Cereb. Cortex* 14:713–720.
- Taylor K, Mandon S, Freiwald WA, Kreiter AK (2005) Coherent oscillatory activity in monkey area v4 predicts successful allocation of attention. *Cereb. Cortex* .
- Tetko IV, Villa AE (2001) A pattern grouping algorithm for analysis of spatiotemporal patterns in neuronal spike trains. 2. application to simultaneous single unit recordings. *J. Neurosci. Methods* 105:15–24.
- Thorpe S, Fize D, Marlot C (1996) Speed of processing in the human visual system. *Nature* 381:520–522.
- Vaadia E, Abeles M (1987) Temporal firing patterns of single units, pairs and triplets of units in the auditory cortex. *Isr. J. Med. Sci.* 23:75–83.
- Ventura V, Cai C, Kass RE (2005) Statistical assessment of time-varying dependency between two neurons. *J. Neurophysiol.* 94:2940–2947.
- Vogels TP, Rajan K, Abbott L (2000) Neural network dynamics. *Annu. Rev. Neurosci.* 28:357–376.
- Volgushev M, Christakova M, W. S (1998) Modification of discharge patterns of neocortical neurons by induced oscillations of the membrane potential. *Neuroscience* 83:15–25.

Vollgraf R, Munk M, Obermayer K (2005) Optimal filtering for spike sorting of multi-site electrode recordings. *Network*. 16:85–113.

von der Malsburg C (1981) The correlation theory of brain function Internal report 81-2, Max-Planck-Institute for Biophysical Chemistry, Göttingen, FRG.