

DISSERTATION

submitted to the
Combined Faculties for the Natural Sciences and for Mathematics
of the Ruperto-Carola University of Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

presented by
Diploma-Biomathematician Olga Kornienko
born in Kiew, Ukraine
Oral-examination: 16.11.2015

Neural Representations and Decoding with Optimized Kernel Density Estimates

Referees:

Prof. Dr. Daniel Durstewitz

Prof. Dr. Christoph Schuster

Statement of Originality

Declarations according to §8 (3) b) and c) of the doctoral degree regulations: a) I hereby declare that I have written the submitted dissertation myself and in this process have used no other sources or materials than those expressly indicated, b) I hereby declare that I have not applied to be examined at any other institution, nor have I used the dissertation in this or any other form at any other institution as an examination paper, nor submitted it to any other faculty as a dissertation.

Abstract

In in-vivo neurophysiology, firing rates from single neurons are traditionally presented in the form of spike counts or peri-stimulus time histograms which are accumulated and averaged across many presumably identical trials. These histograms may on the one hand provide either only noisy representations of the true underlying spiking activity, or on the other hand do not enable single trial resolution. Kernel density estimates (KDE), a weighted moving average with Gaussian kernels centered around spike times, act as a low-pass filters averaging out rapid changes in the firing frequency. Optimized KDEs with the width of the Gaussians (bandwidth) determined through cross-validation or bootstrapping reflect more accurately the underlying spiking activity and also allow for single trial resolution.

We found that optimized bandwidth estimates obtained through unbiased cross-validation (UCV) are an information rich measure, which is applicable to more problems than firing rate estimation, by analyzing both simulations and multiple single-unit recordings from the prefrontal cortex (PFC) of behaving rats. Optimized bandwidth estimates provide a characteristic value for the temporal spiking structure of single units and can be modeled as a function of the temporal precision within spiking patterns accounting for the signal-to-noise ratio in simulated data. The distribution of optimized bandwidth estimates of PFC units and their joint distribution with further spike train metrics allows to segregate groups of cells with distinct spiking properties. Additionally, optimized KDEs obtained with UCV-based bandwidths perform reliable or superior compared to non-optimized KDEs when decoding behavioral events during the task. Moreover, when applied to analyze mechanisms of encoding and internal processing during self-paced cognitive tasks, optimized KDEs facilitate across-trial comparisons of firing activity during trials varying in length, enable to identify neuronal ensembles encoding for task-related events and can unfold population dynamics displaying the underlying neural process.

Zusammenfassung

In der In vivo-Neurophysiologie werden Feuerraten von einzelnen Neuronen traditionell mittels der Zählung von Aktionspotentialen (AP) innerhalb eines Zeitfensters oder mithilfe des Peri-Stimulus Zeitistogramms dargestellt, in welchem möglichst viele identische Versuche akkumuliert und gemittelt worden sind. Diese Histogramme koennen einerseits entweder nur verrauschte Repraesentationen der tatsaechlich zugrunde liegenden neuronalen Aktivitaeten liefern oder ermoeglichen es andererseits nicht, single-trial Aufloesung zu erhalten. Kerndichteschaetzer (KDS), ein gewichteter gleitender Mittelwert, bei welchen Gauss'sche Dichtefunktionen um Spike-Zeiten zentriert werden, fungieren als low-pass Filter, welche hochfrequente Fluktuationen in der Feuerrate herausmitteln. Optimierte KDS, bei welchen die Weite oder Varianz der Gauss'schen Dichtefunktion (die Bandweite) durch Kreuzvalidierungsverfahren oder Bootstrapping bestimmt wird, reflektieren akkurater die zugrunde liegenden AP-Aktivitaeten und ermoeglichen eine single-trial Aufloesung. Wir haben herausgefunden, dass optimierte Bandweiteschaetzer, welche mittels "unbiased cross-validation" (UCV) ermittelt wurden, ein informationsreiches Mass darstellen, welches anwendbar auf viele weitere Probleme als nur die Bestimmung der Feuerrate ist, sowohl in simulierten Daten als auch in Tetroden-basierten elektrophysiologischen Ableitungen des Praefrontalen Cortex (PFC) von Ratten waehrend Verhaltensexperimenten. Ermittelte optimierte Bandweiten liefern einen charakteristischen Wert fuer die zeitliche Aktivitaet von einzelnen Neuronen und koennen in simulierten Daten als Funktion in Abhaengigkeit von der zeitlichen Praezision und des Signal-Rausch-Verhaeltnisses innerhalb der Spike-Muster modelliert werden. Die Verteilung der optimierten Bandweiten-Werte fuer PFC-Zellen sowie ihre gemeinsame Verteilung mit weiteren Massen fuer Regularitaet von AP erlauben es, Klassen von Zellen mit bestimmten Feuerverhalten zu unterscheiden. Ausserdem verhalten sich optimierte KDS, welche mittels UCV-basierten Bandweiten geschaezt wurden, zuverlaessiger und liefern bessere Ergebnisse im Vergleich zu nicht optimierten KDS, wenn diese eingesetzt werden, um Verhalten des Tieres waehrend eines Experiments zu dekodieren. Darueber hinaus erlauben es optimierte KDS wenn sie zur Analyse von Mechanismen der Neuronalen Kodierung waehrend Durchfuehrung von komplexen kognitiven Aufgaben eingesetzt werden Einzel-Trial Vergleiche von zeitlich variablen Versuchsdurchgaengen. Sie ermoeglichen die Identifizierung von neuronalen Ensembles, die aufgabenbezogene Ereignisse kodieren und koennen Populationsdynamiken, welche einen zugrunde liegenden neuronalen Prozess abbilden, aufdecken.

Contents

1	Introduction	1
1.1	The coding problem	2
1.1.1	Encoding schemes of the brain	2
1.1.2	Decoding and stimulus reconstruction - methods and algorithms for retrieving information conveyed by spiking activity	4
1.2	Representations of neuronal spiking activity	6
1.2.1	Model-free and model-based approaches in firing rate estimation	6
1.2.2	Model complexity, parameter estimation and optimization in probabilistic and non-probabilistic approaches	8
1.3	Motivation & goal of the thesis	9
2	Methods	12
2.1	Spike train regularity measures	12
2.2	Kernel density estimation	13
2.3	Selecting the optimal smoothness of kernel density estimates by unbiased cross-validation	14
2.4	Decoding behavior from spike train data: the classification procedure	16
2.5	Evaluating decoding accuracy by m-fold cross-validation	18
2.6	Experimental procedures and electrophysiological recordings	20
3	General properties of UCV-based single unit bandwidth estimates	22
3.1	Motivation: inferring temporal structure of spike trains	22
3.2	Surrogate data: dependency of optimized bandwidth estimated on temporal spiking patterns	23
3.2.1	Generation of temporally precise spiking patterns	23
3.2.2	Relationship between temporal structure of spike trains and the bandwidth outcomes	25
3.3	In-vivo recordings: UCV-based bandwidth estimates of rat mPFC single-units	26
3.3.1	Distribution and properties of bandwidth estimates	27
3.3.2	Relationship between bandwidth estimates and other measures	29
3.4	Summary & Discussion	32
4	Validity of optimized KDEs in neuronal decoding	35
4.1	Validity measures	36
4.1.1	Evaluation of UCV-based bandwidth estimates in decoding	37
4.1.2	Comparing decoding with optimized on non-optimized bandwidths	39

4.2	Surrogate data: validity of optimized bandwidth estimates dependent on distinct firing states	40
4.2.1	Generation of spiking activity with multiple firing states	40
4.2.2	UCV-bandwidth validity for simulated data sets	43
4.3	Decoding spiking activity of in-vivo recordings from the rat mPFC	44
4.3.1	Bandwidth validity for experimental data	44
4.3.2	Decoding performance compared to non-optimized KDEs	45
4.4	Summary & Discussion	46
5	Outlook and possible extensions	48
5.1	Single-trial analysis with optimized single neuron representations	49
5.2	Identifying cell ensembles encoding for task-related information	51
5.3	Reconstructing population dynamics during sequence processing	52
5.4	Concluding remarks	55

List of Figures

1.1	Spiking activity extracted from extracellular recordings	1
1.2	Schematic representation of neuronal codes	2
1.3	Illustration of Bayesian classification and the neuronal coding problem	5
1.4	Comparison of firing activity obtained with PSTH and by Gaussian kernel smoothing	6
1.5	Setting the kernel bandwidth determines temporal precision and smoothness of the firing rate estimate	7
1.6	Error criterion and the optimal bandwidth	8
1.7	Illustration of the two-stage approach for evaluating the optimized bandwidth estimates in decoding	10
2.1	JISI scatter plots	13
2.2	Illustration of Gaussian smoothing	14
2.3	UCV estimate as a function of the bandwidth h	16
2.4	Simulation illustrating Fisher's LDA with two classes and two spiking units	17
2.5	Simplified illustration of the m-fold cross-validation scheme	19
3.1	Illustration of precise spike patterns in surrogate spike trains	24
3.2	KDE of precise spike patterns in surrogate spike trains	24
3.3	Functional dependency between UCV-optimal bandwidth estimates and the temporal structure of spike trains	25
3.4	Distribution single unit bandwidth estimates	27
3.5	Properties of single neuron spiking activity: JISI scatterplots, JISI densities and relationship to task-related events	28
3.6	Correlation to other measures reflecting precise temporal spiking patterns	29
3.7	Relationship to other spike train irregularity measures	30
3.8	Relationship of optimal bandwidth estimates to mean spiking activity	31
4.1	Concept: How to choose the smoothness? Illustration of the two-stage approach for evaluating the validity of optimized bandwidth estimates in decoding	36
4.2	Illustration of bandwidth validity measures and decoding performance as a function of the scaling parameter λ	37
4.3	Decoding performance as function of the bandwidth scaling λ or smoothing degree h	39
4.4	Steps of the HMM simulation	40
4.5	Experimental and surrogate firing rate profile distributions of the Markov sequence states	42
4.6	Bandwidth validity for simulated data sets	43
4.7	Validity measures and decoding performance as a function of the bandwidth estimate	44

4.8	Comparison in decoding performance of optimized KDEs and most predictive non-optimized KDEs	45
5.1	Illustration of time-warped KDEs	49
5.2	Optimized instantaneous firing rates of encoding single units	50
5.3	Identifying encoding units via stepwise feature selection procedures	51
5.4	Reconstructing population dynamics from optimal KDEs	53
5.5	Sequence switch task apparatus and stages of pre-training sessions on the maze	I

Symbols and Abbreviations

Abbreviation	Description
CV	cross-validation
CVE	cross-validation error
HMM	hidden Markov model
ifr	instantaneous firing rate
ISE	integrated squared error
ISI	interspike interval
JISI	joint interspike interval
KDE	kernel density estimate
LDA	Fisher's linear discriminant analysis
MDS	multi dimensional scaling
MISE	mean integrated squared error
mPFC	medial prefrontal cortex
pdf	probability density function
PFC	prefrontal cortex
PSTH	peri-stimulus time histogram
QDA	quadratic discriminant analysis
SD	standard deviation
SEM	standard error of mean
UCV	unbiased cross-validation

Symbol	Description
δ_{Err}	error percentage in deviation from optimal value
λ	bandwidth scaling parameter
λ_{opt}	optimal bandwidth scaling parameter in decoding
μ	mean of a variable
ν	instantaneous firing rate
σ_j	inverse of the temporal precision
σ^2	variance of a variable
h	bandwidth of the Gaussian kernel
h_{opt}	optimal bandwidth in decoding for a fixed set of units
h_{ucv}	optimal unbiased cross-validation bandwidth estimate
C_v	coefficient of variation
Err	averaged cross-validated test set prediction error
L_v	local variation
$N(\mu, \sigma^2)$	normal distribution with mean μ and variance σ^2
r	Pearson correlation coefficient
r^2	coefficient of determination

1 Introduction

How does the brain encode, transmit, store or retrieve information? Recently developed large-scale techniques measuring neuronal activity in the brain on a microscopic scale, such as multiple single unit tetrode recordings make it possible to acquire data from a vast number of simultaneously recorded cells at a high temporal resolution. This allows for more detailed insight into the spatio-temporal structure and functional organization of neuronal activity.

A vertebrate brain contains billions of electrically excitable cells called neurons which possess elaborate branching structure, connected via synapses, through which each cell receives thousands of inputs from other neurons. Physiological features of neurons such as cell membrane-spanning ion channels, control the net flow of sodium, potassium, calcium, and chloride ions in response to internal and external signals. Under resting conditions, a neuron has a net negative charge of about -70 mV termed a negative membrane potential, which is defined as the difference in electrical charge between the interior and the surrounding extracellular medium. At this point, the cell is said to be polarized due to different ion concentrations on either side of the cell membrane. Ion channel gating and the capability to vary the membrane potential plays a key role in the ability of a neuron to generate and propagate electrical signals (neuronal signaling).

A change in the voltage gradient across the membrane which leads to a sufficient depolarization and exceeds a certain threshold level, causes the generation of a short electrical pulse of few milliseconds, termed action potential or spike, which in the case of a temporal sequence of action potentials or multiple spikes is also referred to as a spike train (fig. 1.1). The action potential sequence can be then characterized by a list of successive spike arrival times. For n spikes, we denote these times by $\{t_i\} = \{t_1, \dots, t_n\}$ with $t_1 < \dots < t_n$.

Transmission of these spike trains is believed to allow neurons to communicate with each other and to assemble building blocks which encode information in their activity.

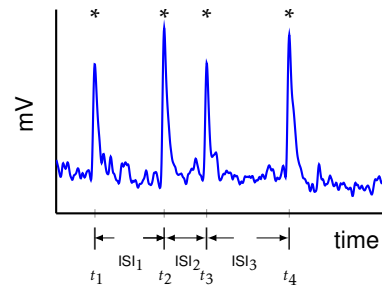


Figure 1.1: Spiking activity extracted from extracellular recordings. Voltage trace containing four spike events at times t_1, \dots, t_4 indicated by stars. After the voltage signal from an electrode is amplified and band-pass filtered, firing of neurons in the vicinity appears as action potentials on top of background activity. Spikes are then detected using an amplitude threshold and assigned to single units according to spike sorting algorithms.

1.1 The coding problem

Deciphering how the brain processes information and the impact on subsequent behavior is one of the most intriguing and debated questions in neuroscience. Progress in this field requires a better understanding of the diverse neural coding strategies used by different brain areas and, in the specific mechanisms by which neurons represent or code for different entities. One ongoing debate in the field is identifying what characteristics of neuronal spike trains serve as the coding signals that carry information (Rieke et al., 1997; Shadlen and Newsome, 1995, 1998; Softky, 1995; Tovee et al., 1993; Fujii et al., 1996; DeCharms and Zador, 2000). While decoding represents the process of reconstructing or extracting information about a stimulus from a given neural response, for instance by predicting the most probable stimulus that could have elicited an observed spike train, encoding can be regarded as the generation of specific activity patterns that serve as representation for these stimuli or behavioral events.

1.1.1 Encoding schemes of the brain

Several candidate codes have been proposed in order to explain how neurons represent information. In the following we will briefly review main ideas and concepts underlying different more fundamental coding strategies, for a more detailed overview see e.g. DeCharms and Zador (2000); Rieke et al. (1997); Feng (2003).

Perhaps the most widely debated question in neural coding is whether information is conveyed in the precise spike timing of action potentials, *temporal code* (fig. 1.2 a), or rather in their frequency, *rate code* (fig. 1.2 b). As the timing of successive action potentials is highly irregular (Softky and Koch, 1993), the interpretation of this irregularity has led to two divergent views of cortical organization which involve two distinct, commonly considered as mutually exclusive encoding schemes that support the propagation of either asynchronous (*rate code*) or synchronous (*temporal code*) spiking activity.

The rate coding hypothesis, initially proposed by Adrian and Zotterman (1926), claims that if irregularity arises from stochastic fluctuations the irregular interspike interval (ISI) reflects a random process.

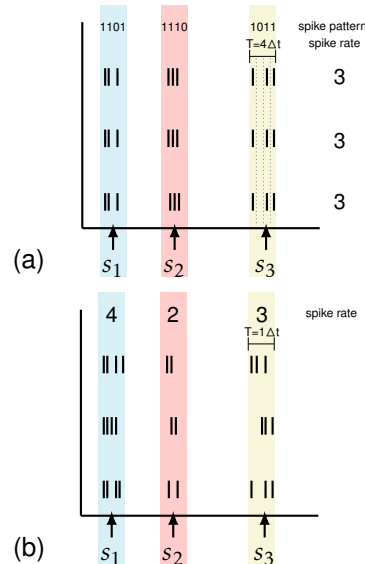


Figure 1.2: Schematic representation of temporal and rate codes. The x-axis represents the time axis, short black vertical lines denote spikes times and rows correspond to different trials. **a.:** Temporal code. Stimuli are encoded by the relative timing of spikes. From the response in a given window of length T to three distinct stimuli s_1, s_2, s_3 indicated by colored bars, one can extract the spike pattern (here a binary four-digit number). Time windows of same rate can contain distinct spike patterns which are determined by the choice of the temporal precision or the size of the binning parameter Δt . **b.:** Rate code. Distinct stimuli are encoded by the number of spikes within the encoding window.

This implies that it would take the pooled responses of many individual neurons to elicit an instantaneous spike rate and requires that the spiking activities of neurons in a population are mostly uncorrelated (Mazurek and Shadlen, 2002). Accordingly, temporal pattern of spikes would convey little information. By contrast, in the *the temporal coding hypothesis* the irregular ISIs may result from precise coincidences of presynaptic events. Consequently, synchrony must represent a necessary feature of spiking activity and information is conveyed by the exact timing of spikes, their intervals and patterns.

From a statistical point of view, the distinction between the temporal and the rate code is more a question of the timescale and can be circumvent by selecting an appropriate temporal precision when estimating the spike rate (Rieke et al., 1997; DeCharms and Zador, 2000). Given that the firing rate is defined as the number of spikes over some time interval Δt (e.g. Tovee et al., 1993), the firing rate can be estimated reliably from a single spike train employing time bins sufficiently longer than the time between the spikes so that enough spikes occur in each bin (fig. 1.2 b). However, when using very small bins with a high temporal precision, each bin will contain only one or zero spikes (fig. 1.2 a), e.g. when the firing rate changes faster than average ISI and the time bins required to capture these changes must be very small, one is effectively measuring the position of individual spikes, being more a measure of spike timing than spike rate. For that reason, the distinction between rate and temporal encoding when analyzing individual spike trains is principled but in that sense rather arbitrary and is based upon the selected temporal precision or time interval chosen for counting the spikes. Not surprisingly, some studies argue that both encoding mechanisms might represent two extreme modes of a continuum and both can be integrated into a single extended encoding concept (Kumar et al., 2010; Ainsworth et al., 2012).

So far we discussed coding properties of single cells, but apart from different spiking patterns or the rate of action potentials (Georgopoulos et al., 1986; Shadlen and Newsome, 1998) much of the attention has focused on whether information carried by single neurons, or by neuronal populations. While the sparse or selective coding strategy postulates that an individual neuron, sometimes also referred to as "grandmother cell", may encode for only one item (e.g. Barlow, 1972; Quiroga et al., 2005), in a fully distributed, ensemble or population coding scheme, each stimulus is coded by a pattern of activity across a larger number of cells. Sparse distributed coding falls between these two schemes, where the simultaneous activation of a small proportion of neurons encodes one item, and each neuron contributes to the representation of only a few stimuli. As neurons in the cortex are densely interconnected, locally and distally, despite evidence for a clear role of individual neurons, involvement of multiple cells during sensory processing or motor activity clearly implicates that population coding must play an important role too (Sakurai, 1996). The question arises of how many neurons are needed or, what size of a functional group of neurons is required to generate a valid representation. For example, the encoding of information can emerge from a more complex temporal organization in networks of spiking cells, such as functional groups of neurons which act as temporally transient cell assemblies (Hebb, 1949; Gerstein et al., 1989; Nicolelis et al., 1995; Harris, 2005; Riehle et al., 1997). Moreover, it is of importance whether interactions between the spiking of different neurons provide additional significant information about a stimulus that cannot be obtained by considering all of their firing patterns individually (*correlation coding*), or, whether individual neurons fire spikes independently of each other.

Such correlations might play a role for even more elaborate encoding strategies which suggest that different electrophysiological signals can be integrated and that *spike phase-timing*, i.e. the relative timing of action potentials with respect to an ongoing oscillation, may serve as an encoding mechanism (Hopfield, 1995; Kayser et al., 2009; Panzeri et al., 2010).

1.1.2 Decoding and stimulus reconstruction - methods and algorithms for retrieving information conveyed by spiking activity

Information conveyed by spiking activity can be retrieved either using the classical approach based upon the tuning curve or by decoding or predicting the stimulus or behavior that elicits a particular neuronal response. Thereby, the response can represent any kind of neuronal activity, either a single neuron or a population response, a (trial-averaged) spike count, a spike sequence or an instantaneous firing rate.

In the tuning curve approach, which has been widely applied for analyses of various sensory areas, the neural response is characterized as function of one particular stimulus attribute. Well-known studies employing this method have revealed evidence for the tuning of primary visual cortex neurons to the spatial location, orientation, direction of motion of visual stimuli (Hubel and Wiesel, 1962), sound frequency and intensity in the primary auditory cortex (Schreiner et al., 2000), and to planned direction of reaching movements in the primary motor cortex (Georgopoulos et al., 1986). Rather than analyzing what neural activity a particular stimulus leads to, more recent reconstruction or classification methods attempt to predict what stimulus produces a particular neuronal response (Bialek et al., 1991; Rieke et al., 1997). This paradigm, is analogous to the task a downstream neuron might perform when reading out its input spike trains. A great variety of machine learning and pattern recognition techniques has been employed for reconstructing stimuli from neuronal signals. A thorough discussion of decoding algorithms can be found in e.g. Oram et al. (1998); Dayan and Abbott (2005); Pouget et al. (2000).

A descriptive example can be provided by Bayesian classification (see fig. 1.3): if $P(s_i)$ denotes the prior probability of observing a stimulus s_i belonging to the set of stimuli $S = \{s_1, s_2, \dots, s_N\}$ and $P(r|s_i)$ is the conditional probability to obtain the neuronal response r when stimulus s_i was presented, then, by employing Bayes' theorem, we obtain:

$$P(s_i|r) = \frac{P(r|s_i)P(s_i)}{P(r)} \quad (1.1)$$

with

$$P(r) = \sum_{j=1}^N P(r|s_j)P(s_j) \quad (1.2)$$

Equation 1.1 gives the posterior probability for a stimulus s_i being present given the observed single-trial response r . We can then predict the most probable stimulus \hat{s} that could have elicited the response r by assigning it to its most likely stimulus class label $\hat{s} = \underset{s_i}{\arg \max} P(s_i|r)$, after having obtained posterior probabilities for all stimuli $s_i \in S$ (fig. 1.3).

Besides Bayesian decoding further examples of classifiers include nearest-neighbor algorithms which assign a given neural response to the class of its nearest neighbor. Fisher linear discriminant classifiers and support vector machines project the original data to a space that optimally separates the samples

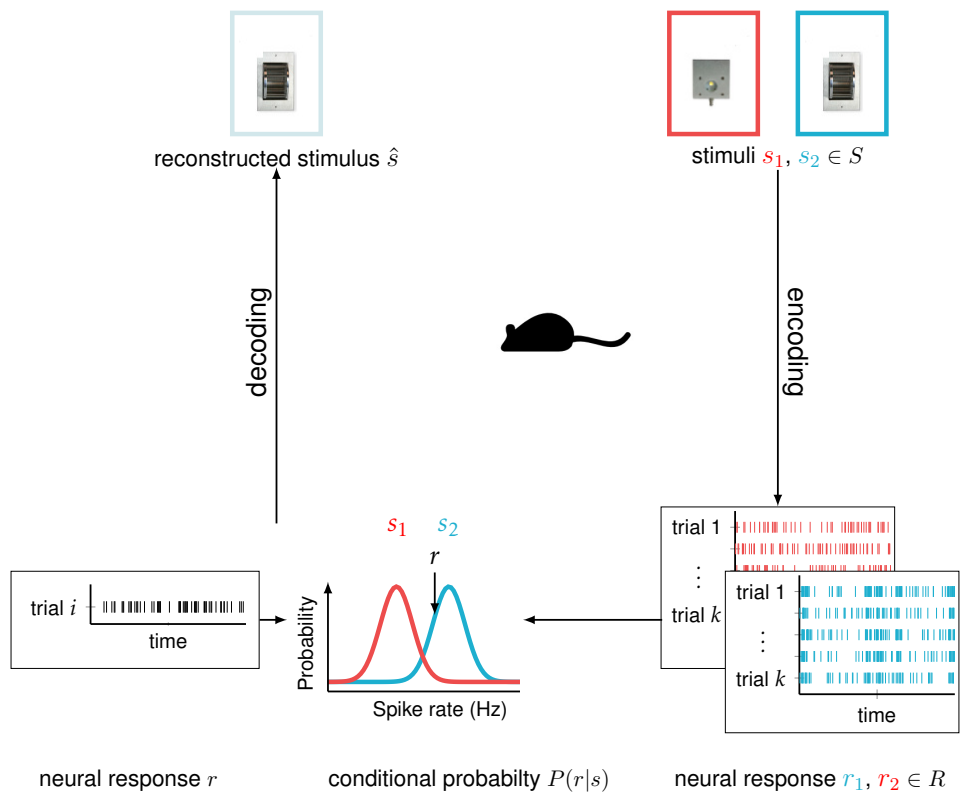


Figure 1.3: Illustration of Bayesian classification and the neuronal coding problem. While encoding can be regarded as the generation of specific activity patterns that serve as representation for these stimuli or behavioral events, decoding represents the process of reconstructing/extracting information about a stimulus from a given neural response, for instance through Bayesian decoding by predicting the most probable stimulus that could have elicited an observed spike train. First, spike trains from multiple trials of a simulated single neuron are converted to spike rates that represent the neuronal response. Then, the response distributions $P(r|s)$ corresponding to two different stimuli, s_1, s_2 , here: a wheel-turn and a lever-press, are estimated. When the neuron fires a single-trial response r just about the average spike rate to s_2 , stimulus s_2 will be decoded.

of each stimulus class (Balaguer-Ballester et al., 2011). Additionally, decoding can be performed by training an artificial neural network (Nicolelis et al., 1998; Wessberg et al., 2000).

So far we discussed principles of decoding. However, before neural activity can be decoded, first the sequence of successive spike arrival times $t_1 < \dots < t_N$ needs to be transformed to a more feasible representation of spiking activity, e.g. by converting a spike train to a spike count rate as in the Bayesian decoding example (fig. 1.3).

1.2 Representations of neuronal spiking activity

The complex temporal structure and dynamics of univariate single neuron and multivariate spiking activity pose the need for advanced methods when analyzing precise temporal spiking of single cells, correlations across multiple neurons or population dynamics (Brown et al., 2004; Churchland et al., 2007; Kass et al., 2005). Thus, prior to being able to access the information which is conveyed by spiking activity e.g. by decoding, spike train data (successive spike arrival times) needs to be transformed first to generate a denoised, more interpretable representation of neuronal spiking activity. Discrete spike trains can be transformed in various ways to representations of the underlying neuronal activity. These Representations can comprise either a single neuron or a population response, a binary spike sequence, a (trial-averaged) spike count or an instantaneous firing rate.

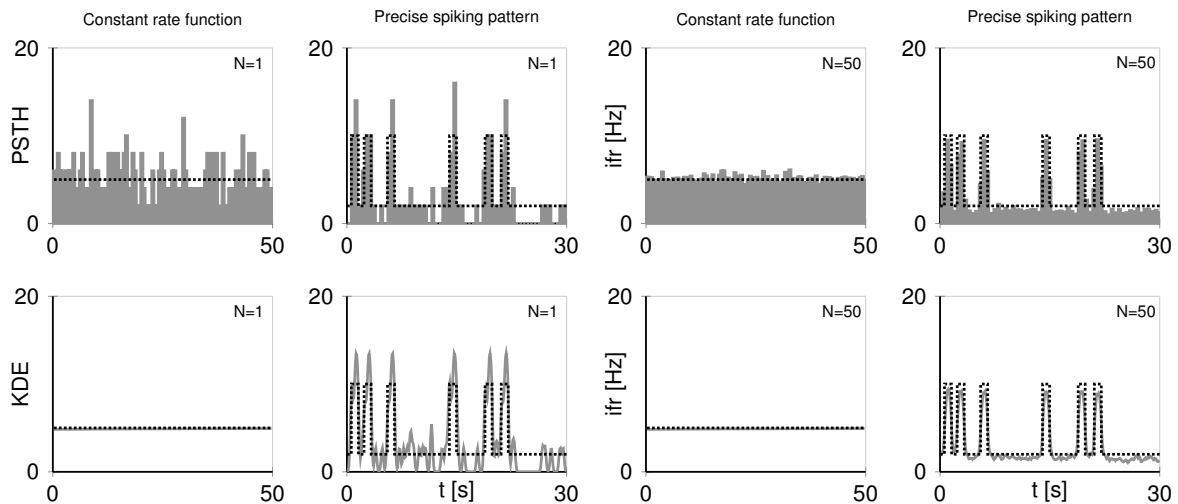


Figure 1.4: Comparison of firing activity obtained with PSTH and by Gaussian kernel smoothing for constant spiking activity and temporally precise spiking patterns when averaged over $N=50$ simulated trials and single trial estimates, dashed line indicates the true underlying rate. While PSTH only converges to the true rate when averaged over a sufficiently high number of trials for both precise pattern and constant activity (top panel), instantaneous firing rates obtained by kernel smoothing can approximate the underlying spiking activity on a single trial basis (lower panel).

1.2.1 Model-free and model-based approaches in firing rate estimation

The irregular spiking activity of a neuron is considered mathematically as a stochastic point-process or probability function (Johnson, 1996; Tuckwell, 1988; Perkel et al., 1967; Cox and Isham, 1980) and can be generally described by the observed number of spikes per time interval, the spike count or spike rate. Due to the probabilistic nature of spike trains, the same underlying input rate will not produce an identical spiking pattern. Therefore, in the most commonly employed approach, spike counts are averaged across many presumably equal trials in form of a peri-stimulus time histogram (PSTH; Gerstein and Kiang, 1960), which in case of a sufficiently large number of trials will approximate the true underlying rate. Figure 1.4, top panel shows PSTHs (gray bars) for simulated constant spiking activity of 5 Hz and precise spike patterns (short pulses of milliseconds of elevated firing of 10 Hz on top of lower back-

ground spiking rates of 3 Hz). When averaging over 50 trials (fig. 1.4, top right) the noise level is clearly reduced and estimated rates approximate true underlying rates (dashed lines) compared to non-averaged, single-trial PSTHs (fig. 1.4, top left). Although noise may be reduced by averaging across multiple experimental trials to produce a smooth firing rate estimate, averaging is sometimes not desirable and can even be misleading. For instance, if neural responses reflect internal processing over external stimulus drive, time courses of the neural responses differ or, trials are self-paced (in the case of freely moving animal), single-trial analysis should be preferred. This is particularly important in behavioral tasks involving motor planning, decision making, rule learning and perception (Nawrot et al., 1999; Horwitz and Newsome, 2001; Czanner et al., 2008; Mante et al., 2013; Churchland et al., 2010).

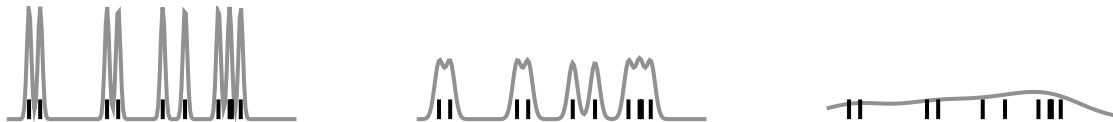


Figure 1.5: Setting the kernel bandwidth determines temporal precision and smoothness of the firing rate estimate (gray lines). For KDE we employed a normal (Gaussian) $N(\mu, \sigma^2)$ probability density function as kernel. μ is centered at spike times (denoted by black vertical lines) and the standard deviation σ controls the width of the kernel. When the bandwidth is small (left) compared to average ISI length one is effectively measuring the position of individual spikes, making it more a measure of spike timing while the estimated rate is more noisy. By setting reasonable bandwidth (middle) one may capture trends and precise spike patterns, whereas a too large bandwidth (right) leads to oversmoothing and gives an estimate of the average spike rate.

The instantaneous firing rate obtained by kernel smoothing can resolve this problem by giving denoised single-trial estimates without averaging and at the same time capture precise spike timing if the temporal resolution is sufficiently high. Figure 1.4, bottom panel, shows trial-averaged and single-trial firing rates obtained with KDE. In this setting, single-trial KDEs already approximate the true underlying rate, in contrast to the PSTH. Kernel smoothing, originally known as the Parzen window method (Parzen, 1962) or, also referred to as kernel density estimation (KDE; Shimazaki and Shinomoto, 2010; Nawrot et al., 1999; Bowman and Azzalini, 1997; Silverman, 1986; Wand and Jones, 1994) is closely related to the PSTH but provides a smooth and continuous firing rate estimate by after setting time interval for counting the number of spikes, replacing each spike by a kernel, i.e. probability density function. Each point (firing rate value) in time is then set to a weighted average of that point's neighborhood. The convolution of the spike train with a kernel acts as low-pass filter averaging out rapid changes in the firing frequency. Similarly to the bin width of the PSTH, varying the kernel bandwidth, i.e. the width of probability density function, will change the temporal precision of resulting firing rate estimates, as shown in figure 1.5. Alternatively, recently employed techniques for obtaining denoised, smoothed estimates of single-neuron instantaneous firing rates on a trial-by-trial basis which are more tailored to neural data include Bayesian binning (BB; Endres et al., 2008), Bayesian adaptive regression splines (BARS; Dimatteo et al., 2001; Olson et al., 2000; Kaufman et al., 2005; Kass et al., 2005), Gaussian process firing rates (GPFR; Cunningham et al., 2008) or generalized linear models (GLMs; Truccolo et al., 2005; Barbieri, 2001; Czanner et al., 2008).

1.2.2 Model complexity, parameter estimation and optimization in probabilistic and non-probabilistic approaches

Many sophisticated methods are probabilistic (BB, BARS, GPFR, GLM), i.e. they require assumptions of the true firing rate distribution (based on a parametric form, such as a homogeneous Poisson distribution, Daley and Vere-Jones, 2003) to estimate the most likely firing rate function via maximum likelihood or Bayesian inference. These methods perform poorly if the parametric form deviates from the empirical firing rate. Additionally, in case of many parameters to be estimated they suffer from technical complexity which leads to increasing computational run-time (Cunningham et al., 2009; Kass et al., 2005). Particularly for exploratory data analysis, where the underlying spiking structure is not well known and first has to be investigated informally (Tukey, 1977), KDEs provide a good trade-off between rather inaccurate multiple-trial firing rate estimators (PSTH) and more complex probabilistic models, which are more efficient, but require greater computational cost and perform poorly when the assumptions of the parametric model are not met.

The most obvious advantages of kernel density methods are that they are simple to implement and non-parametric (i.e. requiring no assumptions about the underlying firing rate distribution).

Equivalently to choosing the time interval for counting the number of spikes PSTH, the temporal precision of the KDE is set by a single parameter (the kernel width) which also determines the smoothness of the instantaneous firing rate (fig. 1.5). The optimal bandwidth can be inferred in multiple ways: informally, e.g. by visual inspection, by rule of thumb (as a multiple of the average ISI) or, according to some formal optimization measures (Bowman, 1984; Rudemo, 1982; Shimazaki and Shinomoto, 2010). These usually quantify the discrepancy between the estimate and the underlying rate by some error criterion. The optimized bandwidth will then be the bandwidth value that minimizes the error measured by the error criterion (fig. 1.6).



Figure 1.6: Error criterion for the bandwidth parameter and the optimal bandwidth. Left: True underlying rate (black dashed line), noisy firing rate estimate with a too narrow bandwidth choice (light gray), oversmoothed firing rate estimate with a too large bandwidth (dark gray). Right: mean integrated squared error (MISE) quantifies deviation from the underlying rate (on the left) as a function of the bandwidth, MISE values for the two non-optimal firing rate estimates shown in the left plot are indicated as marks on the x-axis.

Throughout the thesis we will make use of optimized KDEs with bandwidths estimated through unbiased cross-validation (UCV) for obtaining instantaneous firing rates which serve as a representation of the underlying neuronal spiking activity. Both approaches will be discussed in detail in sections 2.3 and 2.2.

1.3 Motivation & goal of the thesis

The Link between neuronal representations and decoding

Decoding and finding a representation of spiking activity are tightly interlinked processes. The estimated firing rate will impact subsequent classification and inference about encoding mechanisms. Conversely, encoding mechanisms influence the temporal structure of single neuron and population spiking activity that will be reflected in the value of optimized single neuron bandwidth estimates.

The main goal of the thesis is to outline this relationship by focusing on three major properties of optimized bandwidth estimates: First, their role as characteristic value which implicates information on temporal spiking structure and single neuron coding properties as well as their contribution to segregate groups of cells with distinct spiking structure; Second, their contribution to improved decoding performance by optimal noise reduction in instantaneous firing rate estimates; And third, the conclusions we can draw about neuronal coding mechanisms when applying optimized bandwidths and KDEs to analyze single-trial representations, to identify ensembles of cells encoding for task-events and to unfold neuronal population dynamics.

Although bandwidth optimization is not a novel approach, its application in this context is unique: Despite an abundance of general-purpose statistical bandwidth optimization routines (Scott and Terrell, 1987; Silverman, 1986; Simonoff, 1996; Sain, 2002; Sain and Scott, 2002; Hazelton, 2003; Antoniadis et al., 2009; Bouezmarni and Rombouts, 2010; Tran, 2010; Zougab et al., 2014), surprisingly few of these approaches have been applied to neural data. Instead, increasing attention is directed towards model-based techniques which are specifically tailored to analyze spike trains. However, these are derived from strict assumptions about the underlying properties of data, e.g. homogeneous Poisson spiking, as the method proposed by Shimazaki and Shinomoto (2010), which often do not hold. Therefore, there is a need to exploit these pre-existing model-free bandwidth optimization routines for spike train analysis. Furthermore, bandwidth optimizer have been employed to spike train data primarily for facilitating estimation of firing rates and subsequent decoding, in a two-stage approach. Most these studies focus on the general performance compared to other methods for firing rate estimation (Cunningham et al., 2009; Kass et al., 2005). However, to date validity and unbiasedness of the bandwidth optimization routine outcome in decoding have not been analyzed systematically. Additionally, exploration of many highly interesting features, such as temporal structure of spike trains, which can be highlighted by optimized bandwidth estimates, have been neglected.

This work aims to address this gap in the research by analyzing data from multiple simultaneously recorded prefrontal cortex neurons of freely behaving rodents during the performance of complex stimulus operant-based self-paced tasks.

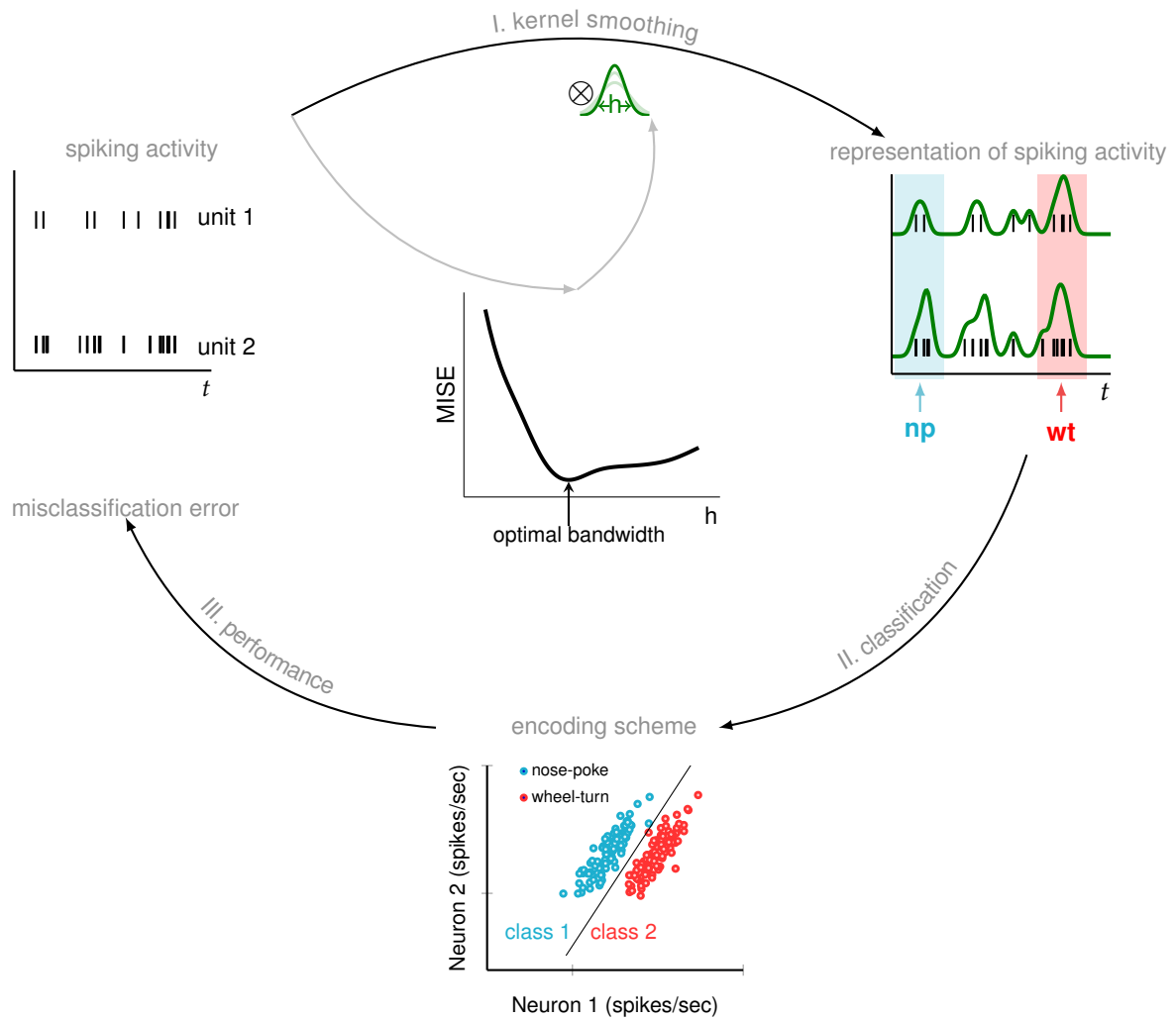


Figure 1.7: Synthesis and link between the coding problem and finding a valid representation of neuronal spiking activity: illustration of the two-stage approach for the simulated spiking activity of two hypothetical neurons in response to two stimuli: wheel-turn (wt) and nose-poke (np). In a first step optimized bandwidth estimates are obtained from both units respectively. These are plugged-in as parameters for subsequent KDE. The estimated multivariate firing rates (ifr neuron one vs. ifr neuron two) are then used to uncover the encoding mechanism, here: correlation- and rate-coding. Decoding performance can then be quantified by the misclassification error, i.e. the relative number of hits or matches.

Thesis organization:

After the general introduction, an overview over the employed methods, and a description of the behavioral task and electrophysiological recordings is given. The first part will cover the basic properties of optimized bandwidth estimates which are usually employed to infer the instantaneous firing rate by kernel density estimation. We focus on general features of the single unit UCV-based bandwidth outcomes by examining their distribution and correlation with further spike train statistics in experimental data. We investigate what information on temporal structure of spike trains bandwidth estimates can reflect by analyzing simulated data with temporally structured spiking patterns.

In the second part, we combine firing rate estimation methods with classification techniques to decode animal behavior during the task based on spiking activity from in-vivo recordings and simulated data sets. As part of this two-stage approach (kernel density estimation in combination with decoding, see fig. 1.7), we address the advantages, disadvantages of the method and its contribution to the improvement in decoding performance by comparing the outcome of optimized with non-optimized KDEs. This is especially important given, in optimization procedures bandwidth estimates are inferred for each recorded neuron individually while non-optimized bandwidths are equal across all cells. Additionally, we examine validity and unbiasedness of UCV-based bandwidth estimates by systematically scaling their outcomes to ascertain if a more optimal scaled value for the decoding procedure exists.

Lastly, we will give an outlook which inferences can be made about encoding mechanisms when applying optimized KDEs to study the ensemble activity of simultaneously recorded neurons from the rat mPFC during a cognitive higher-order sequence-processing task.

The second and third part have already been published as conference abstracts in the context of several conference poster presentations:

Kornienko O, Ma L, Hyman JM, Seamans JK, Durstewitz D. (2013). Analysis of population coding in the rat medial prefrontal cortex. (Poster presentation). *Society for Neuroscience 43rd Annual Meeting*, San Diego, CA.

Kornienko et al.: Neuronal coding in the rodent prefrontal cortex. *BMC Neuroscience* 2013 14(Suppl 1): P117. (Poster presentation). *22nd Annual Computational Neuroscience Meeting*, Paris, France.

Kornienko O, Ma L, Hyman JM, Seamans JK, Durstewitz D. (2012). Reconstructing neural population dynamics during sequence processing in rat prefrontal cortex. (Poster presentation). *Society for Neuroscience 42nd Annual Meeting*, New Orleans, LA.

2 Methods

In the following chapter we will introduce methods statistical concepts for spike train data analysis and the experimental data which will be used throughout this study. First, we will make reference to more general spike train statistics measures and regularity metrics. Then, we will outline the representation of the spike train by its instantaneous firing rate and the procedure how to determine the optimal bandwidth for instantaneous firing rate estimation. And, finally the last sections of this chapter are dedicated to neuronal decoding with the aid of linear discrimination analysis based on instantaneous firing rates, the validation of the decoding results and the in-vivo data sets we applied our methods to.

All procedures and analyses described throughout the thesis are implemented and run using custom written code (MathWorks, Inc., Natick, Massachusetts).

2.1 Spike train regularity measures

A standard metric for measuring the variability or irregularity of a spike train is the coefficient of variation C_v , which defined as

$$C_v = \frac{\sigma_{\text{ISI}}}{\mu_{\text{ISI}}} \quad (2.1)$$

where μ_{ISI} and σ_{ISI} denote the mean and the standard deviation of the ISI distribution. A C_v value close to zero implies that ISIs are almost constant and that the spike train is highly regular, while values close to unity are expected for a Poisson process (Softky and Koch, 1993; Holt et al., 1996; Ponce-Alvarez et al., 2010). It is worth noting that the C_v will not give an accurate estimate of the ISI variability if μ_{ISI} and σ_{ISI} are not stationary and the firing rate changes over time (Softky and Koch, 1993; Holt et al., 1996). Thus, the C_v represents rather a global measure of regularity of spikes emitted by a neuron that is sensitive to fluctuations in the neurons firing rate. Accordingly, a very high C_v value might reflect high firing rate fluctuations (Wohrer et al., 2013).

To decrease the effect of such fluctuation and allow for local non-stationarities metrics which measure the local regularity have been developed. Proposed by (Shinomoto et al., 2003, 2002) the local variation L_v compares only adjacent inter-spike intervals

$$L_v = \frac{3}{N-1} \sum \frac{(\text{ISI}_n - \text{ISI}_{n+1})^2}{(\text{ISI}_n + \text{ISI}_{n+1})^2} \quad (2.2)$$

where $\text{ISI}_n = t_n - t_{n-1}$ and $n=2, \dots, N$. N is the number of emitted spikes. The L_v is equal to zero for distributions reflecting constant ISIs and expected to be near one for Poisson processes.

A further method to characterize the serial dependence of adjacent ISIs is by graphical examination of joint ISI (JISI) scatter plots. JISI plots are a widely used technique in detecting nonlinear dynamics, which are also referred to as Poincare, return or recurrence maps (Abarbanel et al., 1996; Segundo et al., 1998; Fitzurka and Tam, 1999; Szucs et al., 2003). In such plots each point corresponds to a value pair of consecutive ISIs (ISI_n, ISI_{n+1}) among three adjacent spikes. Points falling into the right-lower part indicate spike triplets with a short ISI following a longer one and points in the left-upper half correspond to triplets with a longer ISI following a shorter one (fig. 2.1, top panel).

This plots can highlight points concentrating in increased local density which indicate precisely replicating spike triplet patterns (Szucs et al., 2003).

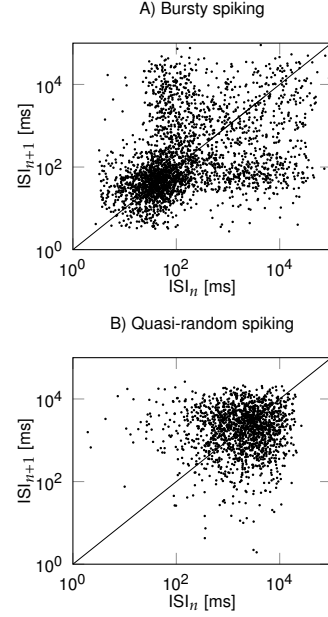


Figure 2.1: JISI scatter plots of spiking activity of two recorded units from the rat mPFC, taken from sequence-switch task (described in sec. 5.4 and 2.6)

2.2 Kernel density estimation

The classical exploratory approach, representation of the data in histogram or PSTH form, which conveys visual information of frequency and the relative frequencies of observations, which is the essence of any density function, comes with several drawbacks. As displayed in bin counts which are constructed on the set of equal-sized non-overlapping intervals, the PSTH suffers from discontinuities (see fig. 1.4, top panel), i.e. sharp steps between the bins and variability when the placement of bins is differed. To overcome both shortcomings one can apply a moving average, which calculates the local average value in a window centered around each data point t_i and which can be viewed mathematically as convolving the original data with a uniform rectangular window

$$\hat{v}_h(t) = \frac{\#\{t_i \in (t-h, t+h)\}}{2nh} \quad (2.3)$$

Equation 2.3 gives the basic kernel density estimator which was introduced as the Parzen-Rosenblatt window method (Rosenblatt, 1956; Parzen, 1962) and can be rewritten as

$$\hat{v}_h(t) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{t-t_i}{h}\right) \quad (2.4)$$

where h is the bandwidth and $K(u)$ is the kernel, which is the uniform probability density function $K(u) = \frac{1}{2}$, if $-1 < u \leq 1$ and zero otherwise, n is the total number of spikes and $t_1 < \dots < t_n$ are successive single unit spike arrival times.

that more accurate measures are needed. More formal, mathematical approaches usually quantify the discrepancy between the estimate $\hat{\nu}_h$ and the underlying rate ν by some error criterion. A commonly employed measure to quantify the accuracy of the firing rate estimate is the mean integrated square error (MISE; Simonoff, 1996; Silverman, 1986). MISE calculates the distance between the estimated rate $\hat{\nu}_h$ and the actual rate ν as a function of the kernel width h .

$$MISE(h) = E[ISE(h)] = E \int (\hat{\nu}_h - \nu)^2(t) dt \quad (2.6)$$

The optimal bandwidth will then be identified as the bandwidth value that minimizes the error function (as illustrated in fig. 1.6). However, since the underlying rate is usually unknown, MISE itself has to be approximated asymptotically. Several studies suggested bandwidth selection with approximate MISE by cross-validating the data (Scott and Terrell, 1987; Silverman, 1986; Rudemo, 1982; Bowman, 1984; Loader, 1999), i.e. calculating MISE for partitions of the data where one part is treated as the actual and the other as the estimated firing rate.

In the following we will make use of the bandwidth selection method (fully described in Rudemo, 1982; Bowman, 1984) referred to as least-squares cross-validation or unbiased cross-validation (UCV) which provides an explicit solution for the error function. The idea is to consider the expansion of the integrated square error (ISE) in the following way

$$\begin{aligned} ISE(h) &= \int (\hat{\nu}_h - \nu)^2(t) dt \\ &= \int \hat{\nu}_h^2(t) dt - 2 \int \nu(t) \hat{\nu}_h(t) dt + \int \nu^2(t) dt \end{aligned} \quad (2.7)$$

The last term in eq. (2.7) is constant as it does not depend on $\hat{\nu}_h$ or h , so that only the first two terms need to be considered. The ideal choice of bandwidth is the one which minimizes

$$ISE(h) - \int \nu^2(t) dt = \int \hat{\nu}_h^2(t) dt - 2 \int \nu(t) \hat{\nu}_h(t) dt \quad (2.8)$$

The CV approach suggests removing one observation at a time t_j , $j=1, \dots, n$, and calculating the usual kernel estimator based on the remaining $n-1$ data points

$$\hat{\nu}_{h,-j}(t) = \frac{1}{(n-1)h} \sum_{i \neq j} K\left(\frac{t-t_i}{h}\right), \quad j = 1, \dots, n \quad (2.9)$$

Accordingly, the last integral in eq. (2.8) can be then replaced by $\int (\hat{\nu}_h \nu)(t) dt = E[\hat{\nu}_{h,n-1}(t)]$ where the expectation is computed using the sample mean of the leave-one-out kernel density estimator in (2.9)

$$\begin{aligned} E[\hat{\nu}_{h,n-1}(t)] &= \frac{1}{n} \sum_{j=1}^n \hat{\nu}_{h,-j}(t_j) \\ &= \frac{1}{n(n-1)h} \sum_{j=1}^n \sum_{i \neq j} K\left(\frac{t_j - t_i}{h}\right) \end{aligned} \quad (2.10)$$

Using the right hand side of (2.8) and replacing $\int (\hat{\nu}_h \nu)(t) dt$ by $E[\hat{\nu}_{h,n-1}(t)]$ results in the UCV function

$$UCV(h) = \int \hat{\nu}_h^2(t) dt - \frac{2}{n} \sum_{j=1}^n \hat{\nu}_{h,-j}(t_j) \quad (2.11)$$

After substituting the Gaussian KDE (2.5) in (2.8) we obtain the following expression (Taylor, 1989),

$$UCV(h) = \frac{1}{2n^2h\sqrt{2\pi}} \left[\sqrt{2} \sum_{i,j} \exp\left(\frac{-(t_j - t_i)^2}{4h^2}\right) - \frac{4n}{n-1} \sum_{i,j} \exp\left(\frac{-(t_j - t_i)^2}{2h^2}\right) + \frac{4n^2}{n-1} \right] \quad (2.12)$$

The optimal UCV-based bandwidth h_{ucv} is then the bandwidth for which the value of the UCV function has its minimum. We minimized $UCV(h)$ with respect to h numerically by means of the Nelder-Mead algorithm (a direct unconstrained nonlinear optimization method provided by the built-in Matlab function 'fminsearch', for more details see Lagarias et al., 1998). The mean interspike interval was used throughout the thesis as an initial estimate for the fminsearch fitting procedure.

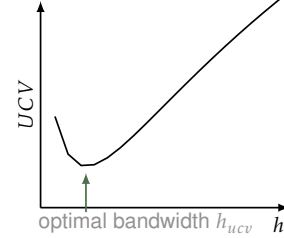


Figure 2.3: UCV estimate as a function of the bandwidth h for spiking activity of a simulated unit (as shown in fig. 1.4, second column).

2.4 Decoding behavior from spike train data: the classification procedure

In order to decode animal behavior from multiple single unit spiking activity we employed Fisher's linear discriminant analysis (LDA) in combination with cross-validation (e.g. Hastie et al., 2009).

First, Kernel density estimates were obtained from all recorded neurons by convolving spike trains with Gaussian kernels binned at a temporal resolution of 100 ms (as described in sec. 2.2). Depending on the goal of classification the smoothing degree, i.e. the width of the Gaussian kernel, was either selected arbitrarily or determined by the UCV optimization routine (eq. 2.12). Single unit instantaneous firing rates were then combined to p -dimensional population vectors $\nu(t) = \{\nu_1(t), \dots, \nu_p(t)\}$ as a function of time bin t .

As LDA is a supervised learning algorithm where labeled data is necessary to build the classification rule we constructed time-dependent vectors $y(t)$ containing class labels corresponding to the different events. Based on event time stamps which were acquired during a recording session, each sampling point in time t_i centered around the time window of 500 ms preceding and following task-relevant behavior of the animal was assigned to $k = 3$ distinct classes c_k corresponding to three different actions performed on task and summarized in a response-class vector $y(t) \in \{1, 2, 3\}$ as a function of time (fig. 2.4, a). Similarly we extracted connected blocks corresponding response-class specific population vectors of spiking activity $\nu(t|c_k)$ discarding all points lying outside the time window of task-relevant events. Also, we analyzed systematically the classification outcome applying different time windows. However, since we did not observe any qualitatively significant change, the time window of 1 s was kept fixed for all classification routines. All labeled points were divided into segments according to distinct trials.

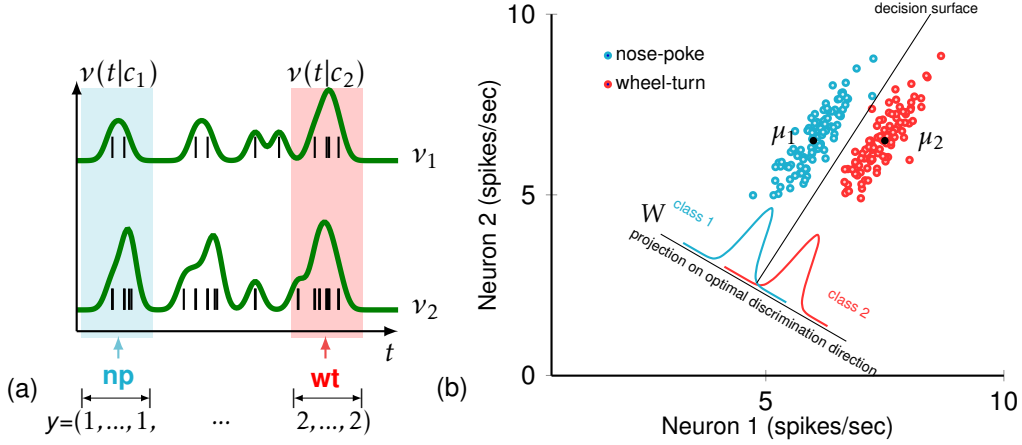


Figure 2.4: Simulation illustrating Fisher's LDA with two classes and two spiking units. a: Obtained firing rates (green) and extracted consecutive blocks of spiking activity $v(t|c_k)$ centered around ± 500 ms task-relevant behavior of the animal, here: nose-poke (np) and a wheel-turn (wt). Arrows on the x-axis indicate event time-stamps. Similarly, the corresponding response-class vector $y(t) \in \{1, 2\}$ is constructed by assigning each point in time to one of the two classes. **b:** Firing rates of two units during the two different responses shown in a) are plotted in a 2-dimensional coordinate system and color-coded according to their class-membership. The underlying firing rates of the two neurons are correlated for both responses, clearly discernible by the covariance ellipses. LDA finds the direction in the two dimensional space of greatest separation between the two classes.

Class-specific population vectors of spiking activity $v(t|c_k)$ and the response-class vector $y(t)$ were then employed to build a LDA classifier and to assess decoding accuracy by means of cross-validation as discussed in the next section.

Fisher's discriminant analysis works by maximizing the difference between class means μ_k and thus the between-class scatter S_b while minimizing within-class covariances S_w , in other words, by finding the direction W of the high-dimensional space along which the overlap between class distributions is minimized (this is illustrated in fig 2.4 b using two neurons) and μ_k , S_b , S_w are defined as follows

$$\mu_k = \frac{1}{N_k} \sum_{t_i \in c_k} v(t_i), \quad \mu = \frac{1}{N} \sum_{k=1}^K \mu_k \quad (2.13)$$

$$S_w = \sum_{k=1}^K \sum_{t_i \in c_k} (v(t_i) - \mu_k)(v(t_i) - \mu_k)^\top, \quad S_b = \sum_{k=1}^K N_k (\mu_k - \mu)(\mu_k - \mu)^\top \quad (2.14)$$

The projections $y(t_i)$ of high-dimensional vectors $v(t_i)$ onto the optimally discriminating direction W are obtained by

$$y(t_i) = W^\top v(t_i) \quad (2.15)$$

The optimally discriminating direction or the so-called weight values or the projection matrix W are determined by those eigenvectors of $S_w^{-1} S_b$ that correspond to the K largest eigenvalues, by maximizing the following equation

$$W^* = \operatorname{argmax} \frac{|W^\top S_b W|}{|W^\top S_w W|} \quad (2.16)$$

This corresponds to maximizing between-class while minimizing within-class scatter as illustrated in figure 2.4 b for the two dimensional case with two neurons. Similarly, projecting the firing rate vector of multiple single units $\nu(t_i)$ for one point in time t_i is equivalent to assigning the same vector to the class with minimum Mahalanobis distance δ_k

$$\delta_k = \sqrt{(\nu(t_i) - \mu_k)\Sigma^{-1}(\nu(t_i) - \mu_k)^\top} \quad (2.17)$$

where Σ is the pooled covariance matrix so that the predicted class labels \hat{y} are

$$\hat{y}(\nu(t_i)) = \underset{k}{\operatorname{arg\,min}} \delta_k(\nu(t_i)) \quad (2.18)$$

Performance of LDA is limited by the fact that the model assumes the data has a Gaussian mixture distribution with varying class means but the same covariance matrix for each class. Since that is often not the case we also employed a generalized model, quadratic discriminant analysis (QDA), where both means and covariances of each class vary and decision boundaries are not linear.

2.5 Evaluating decoding accuracy by m-fold cross-validation

To evaluate the prediction power of a model or in our case smoothed multiple single unit spike train data, and the applied classifier, it is of great importance to make use of objective measures that assess the agreement between the prediction and the underlying observations. One of the most widely used methods for estimating the decoding performance is cross-validation, which estimates well the expected prediction error (e.g. Efron, 2004; Hastie et al., 2009; Bishop, 2006; Krzanowski, 2000). The idea behind cross-validation is that it avoids over-fitting, in terms of the bias-variance trade-off (Hastie et al., 2009). As, by fitting all available data at once by one single model, one tends to fit rather the variance or noise than the signal. Instead, we employ multiple subsets of the sample for both estimation of model parameters, in our case classifier weights W , and assessment of classification performance. This procedure provides a more generalized and accurate measure of the decoding performance (Efron, 2004; Hastie et al., 2009).

In m-fold cross-validation a subset of the available data is used to build the classification rule, and a disjoint, hold out set to test it. In more detail, the data is split into m roughly equal-sized parts, each containing n points. We fit the model to and build the classifier on $m - 1$ parts of the data, the training set, and calculate the out-of-sample prediction error of the fitted model by predicting the left out, m th part of the data, the test set. Thus, the test set prediction error Err_m is denoted as

$$Err_m = \frac{1}{n} \sum_{i=1}^n I[\hat{y}^{-i}(\nu(t_i)) = y(\nu(t_i))] \quad (2.19)$$

where I is the indicator function, the superscript $-i$ indicates that data points $\nu(t_i)$ were left out for building the classifier, \hat{y} denotes predicted class labels and y is the true class membership of $\nu(t_i)$.

This procedure is repeated for all $i = 1, \dots, m$ parts. Averaging over the m resulting estimates yields the m-fold cross-validated test set classification error (CVE) or more generally the expected prediction error, which for the sake of simplicity will be referred to in the following chapters of the thesis as

misclassification error or Err , given by

$$Err = \frac{1}{m} \sum_{j=1}^m Err_j \quad (2.20)$$

After obtaining response-class vectors $y(t)$ and class-specific population firing rates $v(t|c_k)$ for each data set all points were further assigned to l distinct trials. We preserved the temporal structure of trials in the data. Thus, any sample could be split maximally into $m \leq l$ parts each containing at least $n/3$ points of each response-class. If the number of folds was smaller than the amount of trials per recording session $m < l$ there is a multinomial number of ways of partitioning l trials into m distinct parts, with n_i objects in each part i resulting in $\frac{l!}{n_1! \dots n_m!}$ possible combinations to divide trials in training and test sets. For instance for a data set comprised of $l = 30$ trials and $m = 3$ folds this gives a total of $\frac{30!}{10!10!10!} = 5.551 \cdot 10^{13}$ different possible combinations to choose training and test sets. To account for the vast number of possible sample partitions the CV procedure was repeated multiple times with shuffled trials.



Figure 2.5: Simplified illustration of the m-fold cross-validation scheme with $l=2$ trials, $k=3$ response-classes, class labels are color-coded

The cross-validation procedure consisted of the following consecutive steps summarized in the pseudo-code:

1. **if** $m < l$
 - . permute blocks of trials randomly N times
 - . **else** $N = 1$
 - . **for** $j = 1$ **to** N
 2. partition $v(t|c_k)$ and $y(t)$ into m roughly equal-sized distinct parts preserving the trial structure
 - . **for** $i = 1$ **to** m
 3. build classifier on the training set consisting of $m - 1$ parts of the data
 4. test classifier on the m th part and calculate prediction error Err_i
 5. Average error obtained for training parts i and their hold out part $-i$
 - . $Err_j = 1/m \sum_m Err_i$
 - 6. Average over all permutations N to obtain the cross-validation test set error
 - . $Err = 1/N \sum_N Err_j$
- m-fold cross-validation pseudo-code.**

Based on our findings in chapter 4 two criteria had a substantial impact on the CV outcome: first the number of partitions m and second the auto-correlation which is present in each consecutive block of response class-specific firing rates $\nu(t|c_k)$. E.g. taking $m = 3$ folds resulted in a CVE converging to zero. However, the test sets contained many similar (autocorrelated) samples which made the interpretation of the CV result difficult. To overcome this issue, we averaged consecutive blocks of the same class k for each trial instead of taking multiple sample points from the entire time window centered around responses. Decoding with temporally averaged class-specific firing rates yielded a poorer performance compared to non-averaged $\nu(t|c_k)$, but also implied qualitatively different information which will be discussed in chapter 4. We employed two different cross-validation schemes.

1. Leave-one-third-trial-out CV, where $m=3$ and for each trial we take block-wise averaged response-class firing rates. This method gives a large pool of temporally uncorrelated training and test samples which provides a more robust estimate of the error.
2. Leave-one-trial-out CV, where $m = l$ and folds correspond the number of trials per data set. Taking the non-averaged time-series we preserve auto-correlation and temporal structure of the original spike train series.

We varied these depending on the purpose of the classification and the issue of interest: e.g. evolution of single-trial dynamics (chap. 5) was analyzed with the latter and robustness and general inference about prediction accuracy (chap. 3 and 4) was examined with the former CV scheme. Summarizing, at first instance we did not aim to achieve a misclassification error close to zero, taking CVE as prediction accuracy measure, which was the case for particular cross-validation schemes. More importantly, our primary goal was to determine a suitable measure which allowed for comparisons among certain criteria of interest, e.g. choice of the bandwidth for kernel smoothing or selection of set of units to build the classifier. In general, the aim was to identify the model which is superior compared to others with the aid of the CVE.

2.6 Experimental procedures and electrophysiological recordings

All of our experimental analyses were performed on data from behavioral experiments and extracellular in-vivo recordings from the rat anterior cingulate cortex (ACC) which were conducted and generously provided by Dr. Liya Ma and Dr. James Hyman, Brain Research Centre and Department of Psychiatry, University of British Columbia, Vancouver, Canada.

In-vivo recordings were partly taken from the study published in Hyman et al. (2013), and partly from the study conducted by Ma (personal communication). Both were involving two distinct operant-based behavioral tasks ("sequence-switch" and "alternation task") with multiple stimuli, task events and responses.

Sequence switch task: In the study conducted by Ma, animals had to perform a fixed sequence of three actions on a maze before reward, which was reversed after a given number of trials. The first sequence consisted of the following actions: wheel-turn, lever-press nose-poke (see fig. 5.5). After at least 20 trials or 20 min animals were removed and placed back in the maze which indicated the switch of the sequential order.

Alternation task: In the Hyman et al., 2013 study rats performed a continuous or delayed alternation task where following a nose-poke one of two levers had to be pressed in an alternating fashion prior to delivery of a reward. In the delayed version, a delay of 10 sec was introduced between each lever-press from a preceding trial and the nose-poke initiating the next trial.

For a more detailed description and procedures of the alternation task the reader is referred to the publication by Hyman et al. (2013). Behavioral and electrophysiological details of the sequence-switch task are best described by the citation provided by Dr. Ma in section 5.4.

We ran our analysis procedures on spike-sorted single units plus time stamps for relevant events, which were held in the format of MatLab matrices. Overall data from 7 animals, consisting of 19 recording sessions with 40 trials on average and up to 74 isolated units per recording session, containing 905 units in total was analyzed.

For decoding analyses we extracted event-time stamps related to the 4s delay period preceding nose poke, nose-poke and lever press from the alternation task, and to a nose-poke, wheel-turn and lever-press from the sequence-switch task. Only correct trials were included.

3 General properties of UCV-based single unit bandwidth estimates

3.1 Motivation: inferring temporal structure of spike trains

Precise spike train patterns in in-vivo recordings, i.e. exact timing of single neuron spikes in behaving animals during repeated stimulus presentation have been reported in multiple cortical areas (Bair and Koch, 1996; Buračas et al., 1998; Reinagel and Reid, 2002; Fellous et al., 2004; Brown et al., 2005; Raman et al., 2010). Particular precision of spike timing can be characteristic for both specific stimuli and single neurons (Mainen and Sejnowski, 1995; Reinagel and Reid, 2002). As pointed out in a previous study by Song et al. (2009) a bandwidth estimate "essentially determines the optimal temporal resolution used in comparing the predicted spike train with the actual spike train where large bandwidth values indicate low temporal resolution whereas small values imply high temporal resolution".

The main aim of the following chapter is twofold: first, to examine if the UCV-based bandwidth estimate reflects information about temporal structure of the spike train as well as what sort of information it implicates (experimental data) and second, conversely, if the temporal structure, i.e. the precision of spikes and their signal-to-noise ratio, directly influence the bandwidth outcome (surrogate data).

To answer the second question we generated simulated spike trains with known underlying ground truth and examined the relationship between bandwidth outcome and precision and signal-to-noise ratio (sec. 3.2). To answer the first question we looked at measures in experimental in-vivo recordings which are associated with precise spike timing and examined the relationship to the h_{ucv} outcome (sec. 3.3). We analyzed the three following single unit spike train properties which are related to the temporal structure of spiking activity: First, single unit discrimination power of task-related event occurrences; second, recurrence of spike triplet patterns, i.e. distribution of joint interspike intervals of successive ISIs pairs and regions densely clustered/spread; third, general spike train statistics and irregularity measures, i.e. the coefficient of variation, the local variation and the average firing activity).

In the following, we will refer to the temporal resolution of spike times interchangeably as precision or jitter where the precision (units: 1/s) is defined to be the inverse of the jitter (units: s).

3.2 Surrogate data: dependency of optimized bandwidth estimated on temporal spiking patterns

To examine how temporal structure of spike trains impacts optimized bandwidth estimates, we generated surrogate spike trains with known underlying structure and analyzed the UCV-based bandwidth outcome as function of the parameters which determine the spiking pattern.

3.2.1 Generation of temporally precise spiking patterns

We conducted following consecutive steps:

A homogeneous Poisson point process was generated with specified constant firing rate $\nu = N/T$ of duration $T = 600$ s with N number of spikes corresponding initially the background spiking activity or noise $\nu = \nu_{noise}$.

We introduced n_{signal} short pulses or number of non-overlapping spike train patterns of short duration T_{signal} with mean firing rate $\nu_{signal} = 10$ Hz corresponding the signal.

To preserve the mean firing rate of the entire spike train ν , ν_{noise} was decreased by removing the number of introduced signal spikes N_{signal} from the background spike train at random $N' = N - N_{signal}$ so that $\nu = N'/T = \nu_{signal} + \nu_{noise}$.

To reduce temporal precision of spike patterns which is equivalent to introducing spike-timing variability or temporal jitter σ_j , signal spike times were jittered by adding independent normally distributed random variables with zero mean and standard deviation specified by the amount of jitter $\sim \mathcal{N}(0, \sigma_j^2)$

While keeping $\nu_{signal} = 10$ Hz and duration of spike train $T = 600$ s constant, we examined the impact of two parameters on the bandwidth estimate, which essentially determine the temporal structure of the generated spike patterns:

Precision: Adding the amount of the jitter $\sigma_j \in (0, 32]$ s to the data changed temporal precision of spikes.

Signal-to-noise ratio (SNR): We varied number of spiking patterns/short pulses corresponding the signal $n_{signal} \in \{5, 10, 25, 50, 75, 100, 200\}$, signal duration $T_{signal} \in \{0.5, 1\}$ s and the average firing activity of the entire spike train $\nu \in \{1, 2.5, 5\}$ Hz.

SNR of precise patterns to background activity can be then computed directly applying the following equations

$$\text{SNR} = \frac{N_{signal}}{N_{noise}} \quad (3.1)$$

$$= \frac{\nu_{signal} \cdot T_{signal} \cdot \frac{n_{signal}}{T}}{\nu - \nu_{signal} \cdot T_{signal} \cdot \frac{n_{signal}}{T}} \quad (3.2)$$

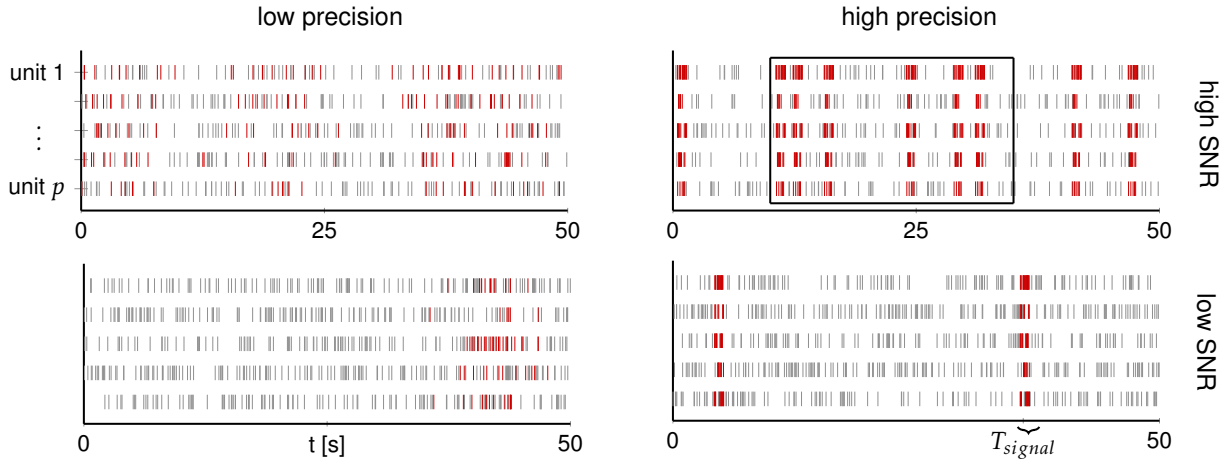


Figure 3.1: Illustration of precise spike patterns in surrogate spike trains. Rasterplots show spiking activity of five exemplary simulated units (y-axis) for different SNRs and temporal precisions. Gray-colored horizontal bars correspond to spike time occurrences of the background activity and red bars to (precise) spike patterns, $T_{signal} = 1$ s, $\nu = 2.5$ Hz, high SNR=2, low SNR=0.06, high precision $\sigma_j=0$ s, low precision $\sigma_j=2$ s.

Figure 3.1 shows examples of simulated spike trains with different temporal precisions $\sigma_j \in \{0, 2\}$ of the short pulses and different SNRs $\in \{0.06, 2\}$. Figure 3.2 shows the framed box in top right panel for high precision and high SNR (SNR=2, $\sigma_j=0$) as instantaneous firing rate. The dashed black line displays the underlying rate and the gray solid line the firing rate estimate obtained by KDE as average over $n=50$ exemplified surrogate spike trains. Precise patterns can be identified as short pulses with increased activity by looking at optimally smoothed instantaneous firing rates.

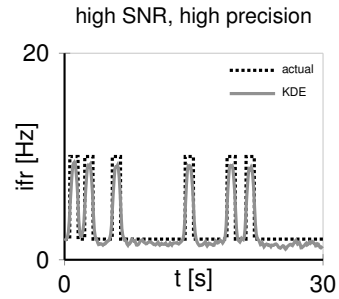


Figure 3.2: KDE of precise spike patterns in surrogate spike trains Instantaneous firing rates obtained by optimized kernel density estimation as average over 50 exemplified simulated spike trains.

We calculated bandwidth estimates for all simulated spike trains with different SNRs, precisions and average firing rates.

3.2.2 Relationship between temporal structure of spike trains and the bandwidth outcomes

Analysis of UCV-based optimal bandwidth estimates h_{ucv} as a function of temporal structure in simulated spike trains reveals that lower bandwidth estimates are associated with more precise temporal patterns of the spike trains, a higher signal-to-noise ratio and higher average spiking activity.

To illustrate the dependence between bandwidth estimates and the level of jitter we first averaged values over $n=5000$ simulations and then base-10 logarithm transformed both axes. From figure 3.3, b it follows that h_{ucv} values can be characterized by a sigmoid: the UCV-based bandwidth increases initially approximately exponentially with decreasing precision of the precise patterns. Then, depending on SNR levels, after reaching the inflection point (jitter>5s) the growth slows and h_{ucv} attains its maximal value. Following this capacity limit, after precise patterns are completely smeared out, h_{ucv} values remain constant.

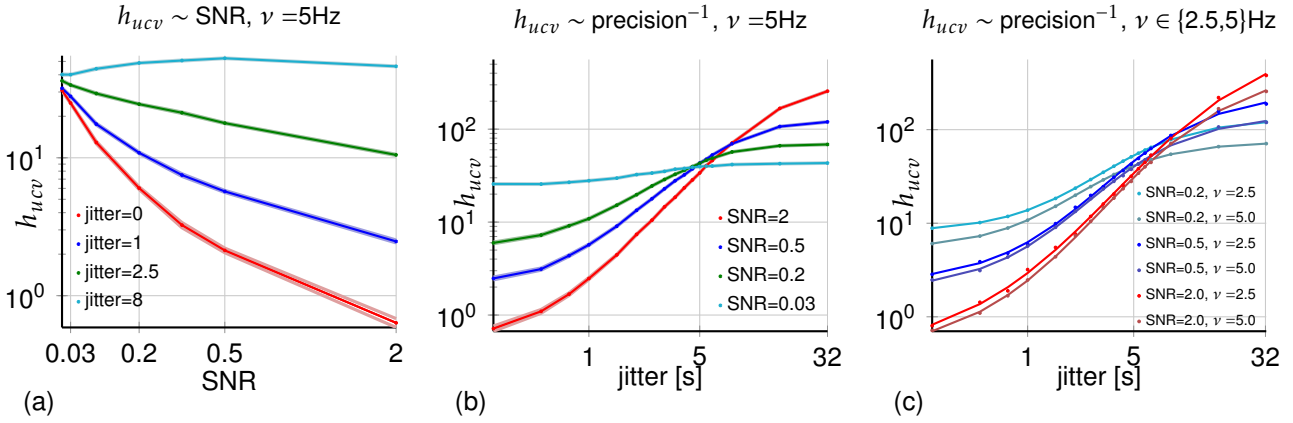


Figure 3.3: Functional dependency between UCV-optimal bandwidth estimates and the temporal structure of spike trains. **a:** Optimized bandwidth as a function of the signal-to-noise ratio for different precisions $\sigma_j \in \{0,1,2,5,8\}$. **b:** Optimized bandwidth as a function of the temporal precision for different SNRs $\in \{0.03,0.2,0.5,2\}$, $\nu=5\text{Hz}$. Lines correspond to the mean over $n=5000$ bandwidth estimates obtained from simulated spike trains. Shaded area indicates the SEM. **c:** Generalized logistic functions fitted to h_{ucv} values dependent on the temporal precision for different average spiking activity $\nu \in \{2.5,5\}$ Hz, SNR $\in \{0.2,0.5,2\}$. Dots correspond actual values and lines to fitted values according to eq. 3.3. SEM in the right panel has been left out for clarity.

The transformed bandwidth values dependent on the jitter σ_j can then be fitted by a generalized logistic function of the following form for different SNR and average firing activity ν levels (fig. 3.3, c)

$$h_{ucv}(\sigma_j) = A + \frac{K - A}{1 + e^{-(a+b \cdot \sigma_j)}} \quad (3.3)$$

Parameters A , K , a , b denote the lower and upper asymptote, growth rate, jitter at maximum slope and were determined by maximizing the log-likelihood with the MATLAB built-in function "fminsearch". From figure 3.3 c it also follows that the UCV-based bandwidth outcome always increases with lower average firing rates if SNR levels are kept constant. Moreover, lower bandwidth estimates are associated with higher SNRs within the temporal structure of simulated spike trains.

SNR/ ν [Hz]	A	K	a	b	r^2
0.2/ 2.5	1.292 (1.263, 1.321)	2.366 (2.332, 2.400)	-1.614 (-1.779, -1.450)	3.09 (2.809, 3.371)	$1-1.2 \cdot 10^{-3}$
0.2/ 5.0	0.907 (0.873, 0.941)	2.118 (2.085, 2.151)	-1.438 (-1.594, -1.282)	3.234 (2.958, 3.510)	$1-1.2 \cdot 10^{-3}$
0.5/ 2.5	0.731 (0.705, 0.757)	1.870 (1.852, 1.889)	-1.021 (-1.122, -0.920)	3.373 (3.176, 3.570)	$1-6 \cdot 10^{-4}$
0.5/ 5.0	0.358 (0.314, 0.402)	2.398 (2.354, 2.442)	-1.294 (-1.397, -1.191)	2.788 (2.611, 2.965)	$1-5 \cdot 10^{-4}$
2.0/ 2.5	0.287 (0.250, 0.323)	2.158 (2.129, 2.188)	-1.105 (-1.190, -1.020)	2.946 (2.794, 3.097)	$1-4 \cdot 10^{-4}$
2.0/ 5.0	-0.405 (-0.53, -0.28)	3.040 (2.893, 3.186)	-1.080 (-1.199, -0.961)	1.991 (1.787, 2.195)	$1-7 \cdot 10^{-4}$

Table 3.1: Parameter solutions and coefficients of determination for equation 3.3

Estimated model parameters and their respective 95% confidence intervals for the functions in fig. 3.3 c are shown in the table 3.1.

In general, the SEM was low (fig. 3.3, a and b) ranging from 10^{-2} to below 10^0 and the proportion of variance explained by the model in equation 3.3 r^2 is almost unity. The estimated sigmoid functions $h_{ucv}(\sigma_j|\text{SNR}, \nu)$ shown in figure 3.3 c and parameter solutions for equation 3.3 given in table 3.1 are unique in the probed SNR and ν range.

3.3 In-vivo recordings: UCV-based bandwidth estimates of rat mPFC single-units

Besides the distribution of the bandwidth estimates, we examined the three following single unit spike train properties which are related to the temporal structure of spiking activity: First, single unit discrimination power of task-related event occurrences; second, the recurrence of spike triplet pattern: distribution of joint interspike intervals of successive ISIs pairs and how densely those cluster; and, third, general spike train statistics and global and local irregularity measures such as the mean firing rate, the coefficient of variation and the local variation.

To examine the single unit discrimination power we first measured the decoding performance as prediction error of a linear discriminant classifier built on temporally averaged response-class specific optimized firing rates of single cells (discussed in detail in sec. 2.4). This measure reflects how well spiking activity of a single unit discriminates between distinct task events. The misclassification rate will be low if the single unit discrimination power is high. Consequently, a misclassification rate below chance implies that spiking activity and spike timing must be time-locked to task-related events. Thus, low misclassification rate can be associated with the occurrence of temporally precise spiking patterns.

In order to determine the clustering or spread of recurrent spike triplet pattern we looked at the local autocorrelation structure of spike trains by examining the relationship of adjacent ISIs. Besides being related to temporal correlation of ISIs JISI plots can also reflect the degree of regularity or periodicity in the underlying activity and more importantly the precision of recurrent triplet spiking patterns (Faure et al., 2000). We also examined the correlation between optimized bandwidth estimates and the coefficient of variation C_ν which reflects the global regularity of a spike train, the local variation L_ν and the average firing rate $\langle \nu \rangle$ (as introduced in sec. 2.1).

3.3.1 Distribution and properties of bandwidth estimates

We obtained bandwidth estimates of all recorded units $N=905$ (units with spiking activity below 0.1Hz were excluded). From figure 3.4 it follows that UCV-based bandwidth estimates of single units are distributed mixed log-normally with heavy tails and values ranging from 0.05-356.7. The majority of single units has an UCV-optimal bandwidth at $h_{ucv} \approx 0.5$ (visible as pronounced mode in fig. 3.4 a). This value corresponds to the optimal temporal resolution, i.e. variance of the Gaussian kernel, for computation of the instantaneous firing rate. Logarithm-transforms of bandwidth estimates reveal that values follow a log-normal mixture distribution with at least two (alternation task, fig. 3.4 b) or three distinct modes (sequence-switch task, fig. 3.4 c). We further investigated if distinct modes represent different subgroups of units which share common properties in their temporal spiking structure.

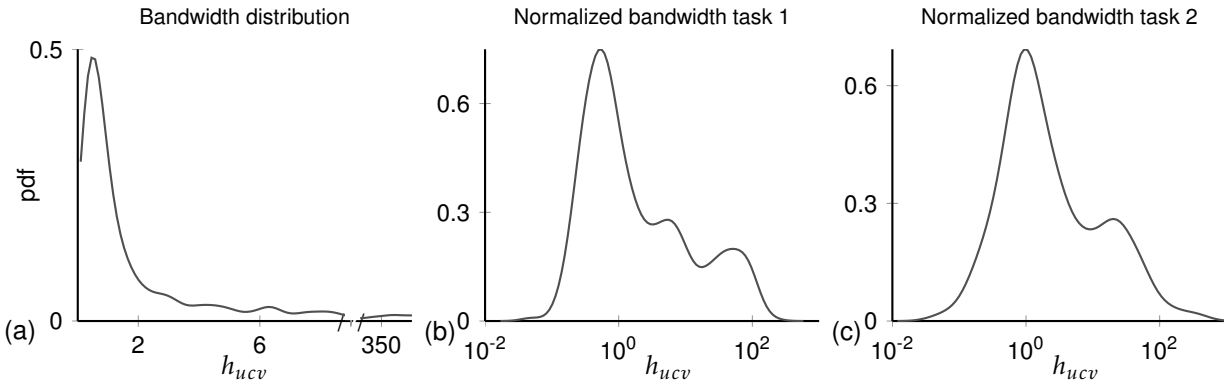


Figure 3.4: **a:** Distribution of $N = 905$ single unit bandwidth estimates. **b:** Base-10 logarithm transformed bandwidth estimates on task 1 (sequence-switch task) and **c:** on task 2 (alternation task).

To illustrate what sort of information on the temporal structure of single units can be detected by the bandwidth estimate we selected three exemplary units with $h_{ucv} = \{0.15, 0.26, 81.73\}$ from tails of the log-transformed h_{ucv} distribution. We first visualized their spiking activity with the aid of return maps (fig. 3.4 top and middle panel) and second displayed their spike patterns in relation to onsets of task-related events (fig. 3.4 lower panel).

The JISI plots we obtained showed the presence of distinct clusters (fig 3.5, top). The cluster size and the spread of the points reflects the serial dependence of successive ISIs Dodla and Wilson (2010). To compute how likely joint ISIs are going to recur we calculated the joint ISI pair density using a bivariate Gaussian kernel similar to the univariate KDE in section 2.2. These densities are visualized for log-transformed ISIs as color-coded contour maps (fig. 3.5, middle). Hues changing from white to magenta indicate increasing probability of recurrent spike triplet patterns, i.e., when pairs of ISIs are more likely to appear. Points which scatter in clusters indicate recurrent spike triplet patterns.

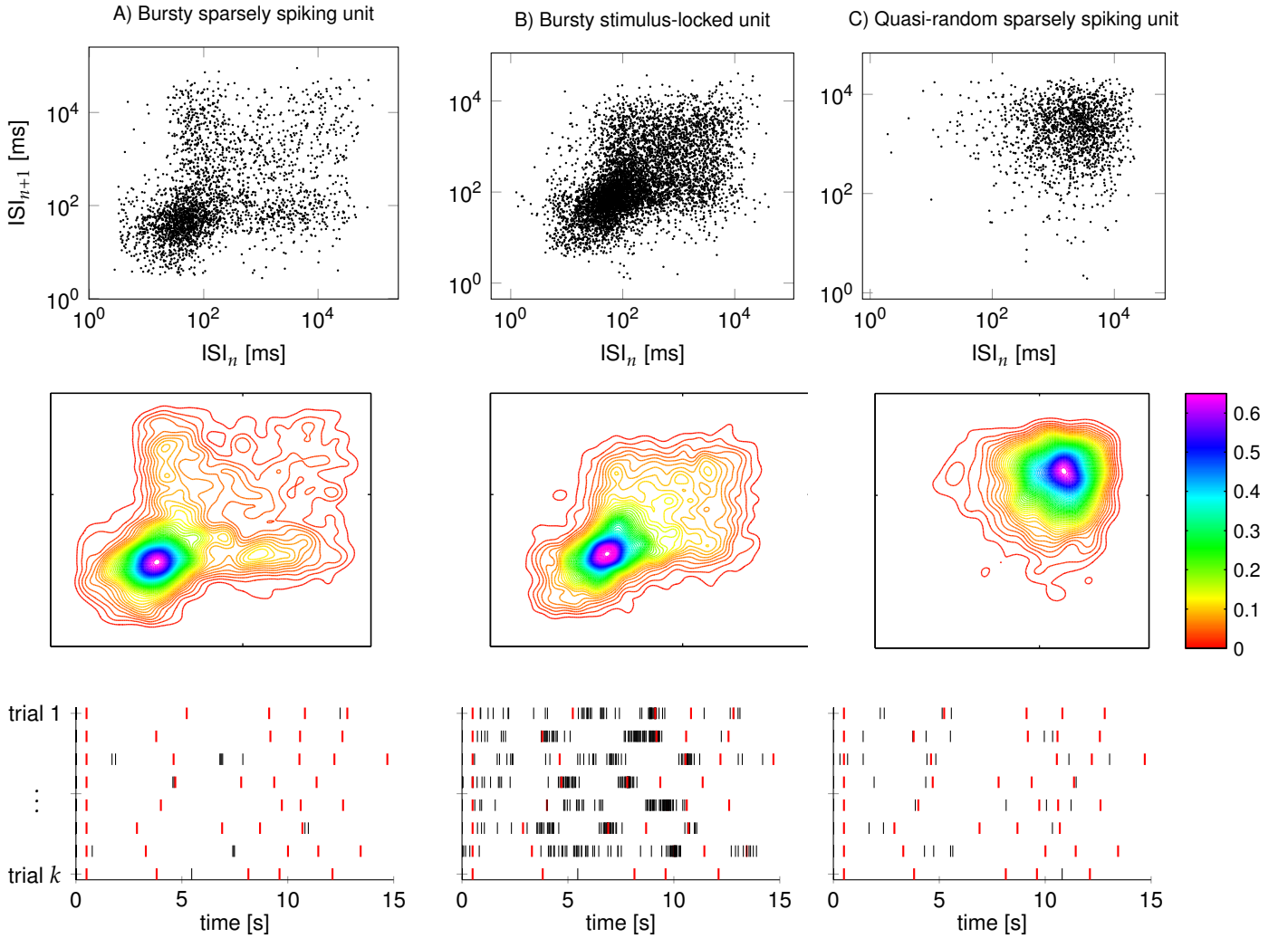


Figure 3.5: Properties of single neuron spiking activity: JISI scatterplots, JISI densities and relationship to task-related events. Columns are arranged from left to right in order of increasing h_{ucv} values for three selected units (chosen from the same data set) with $h_{ucv} = \{0.15, 0.26, 81.73\}$. **Top:** Scatter plots of base-10 logarithm transformed JISIs. **Middle:** JISI density contour plots, regions with increased densities of successive ISIs are color-coded. ISI pairs were logarithm transformed for visualization purposes and provided a more compact overview compared to non-transformed heavy-tailed ISI distributions. **Bottom:** Spiking activity (black horizontal bars) and time-stamps of successive task-related events (red horizontal bars) for 8 depicted trials of the same units. Spike trains were aligned at the beginning of each trial. Units are taken from the sequence-switch task, five task-related events per trial comprise: wheel-turn, nose-poke, lever-press, approaching the reward and reward consumption respectively.

Exemplary units with small bandwidth estimates ($h_{ucv} < 1$) exhibit irregular spike sequences with sparse and bursty firing behavior (fig. 3.5, lower left and middle panel). JISI scatter and JISI density plots of the unit having the lowest estimate ($h_{ucv} = 0.15$) show three discernible clusters (fig. 3.5 left, top and middle panel). The two arms parallel to the axes correspond to interburst intervals, where the mode to the right displays ISIs associated with initiation and the upper mode completion of bursts. The unit with a slightly higher h_{ucv} value of 0.26 (figure 3.5 middle panel) is clearly time-locked to some of the task-related events (fig. 3.5 lower, middle panel), while the third unit with the highest $h_{ucv} = 81.73$ is spiking quasi-random and is not time-locked to any events of the task-related events.

3.3.2 Relationship between bandwidth estimates and other measures

After we have characterized distribution of the bandwidth estimates $p(h_{ucv})$ separately, we examine the covariation of bandwidth estimates and other spike train statistics. In order to quantify this covariation statistically, we first examined if there is a linear correlation between h_{ucv} and the other measures by estimating Pearson's correlation coefficient r and fitting a linear regression model with h_{ucv} as the independent variable. If the correlation was non-linear, we attempt to fit a non-linear model. Otherwise, if there was no direct relationship which can be characterized by an explicit function, i.e. bandwidth estimates and values of other measures exhibit multiple distinct clusters in their joint scatter plots, we estimated their joint distribution, denoted as $p(h_{ucv}, \cdot)$, by fitting a Gaussian mixture model. We then characterized each component of the joint distribution separately by determining the cluster centers and the respective covariances within the clusters.

Single unit discrimination power. Bandwidth estimates and decoding performance of single units were significantly correlated ($r=0.36$, $p\text{-value} < 10^{-29}$, Pearson correlation coefficient). By fitting a linear model of the form $Err = 3.2 \cdot h_{ucv} + 61.4$ 13% of the variation can be explained ($r^2=0.13$, F: 138, $p\text{-value} = 1.01e-29$, F-statistic vs. constant model). Most of the cells were encoding above chance. The mode of the error distribution is centered at $\approx 66\%$ misclassification error (fig. 3.6 b, dotted horizontal line and distribution right panel). Decoding three actions this corresponds to a chance level of 1/3 correct classification rate.

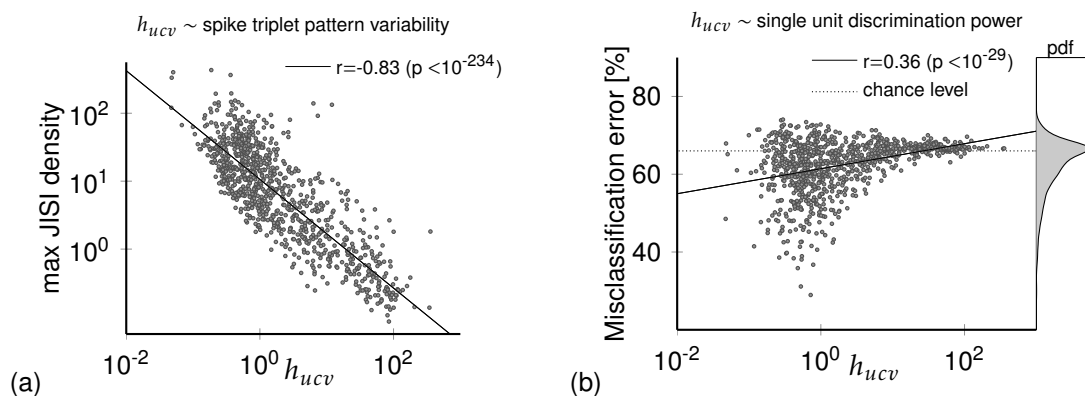


Figure 3.6: Correlation with other measures reflecting precise temporal spiking patterns. Optimal bandwidth estimates and **a**: base-10 logarithm transformed maximum JISI density values, and **b**: single unit misclassification error.

Probability of recurrent spike triplet patterns. We analyzed the clustering of ISI pairs by examining the maximal probability values ISI pairs $p(\text{ISI}_n, \text{ISI}_{n+1})$ across all single units (fig. 3.6 a). This quantity is proportional to the local number of points in the JISI scatter plot (Dodla and Wilson, 2010). A density of unity corresponds to the case where all ISI pairs fall into the same region in the JISI plot. This would imply that there exist only one exact spike triplet pattern. In contrast, smaller probability values indicate that ISIs are more spread in time, i.e. occupy larger regions in return maps and there is more variability of recurrent spike triplet patterns. The maximal probability of recurrent spike triplet patterns was significantly correlated with lower bandwidth estimates ($r=-0.83, p\text{-value}<1.35e-234$, Pearson correlation coefficient).

Irregularity. We computed two further measures which are associated with global and local irregularity of the underlying spiking activity: the coefficient of variation C_v , the local variation L_v . We then quantified their relationship to UCV-based bandwidth estimates. Coefficients of variation and bandwidth estimates were negatively correlated. C_v values exponentially decrease with increasing bandwidth estimates (fig. 3.7 a). Single units spiking close to a Poisson-like or random, i.e. $C_v \approx 1$ or regular regime ($C_v < 1$) display increasing bandwidth estimates. We fitted an exponential function of the form $C_v = 0.24 + e^{-0.68 \cdot h_{ucv}}$ which could explain $r^2=0.34$ variation contained in the data.

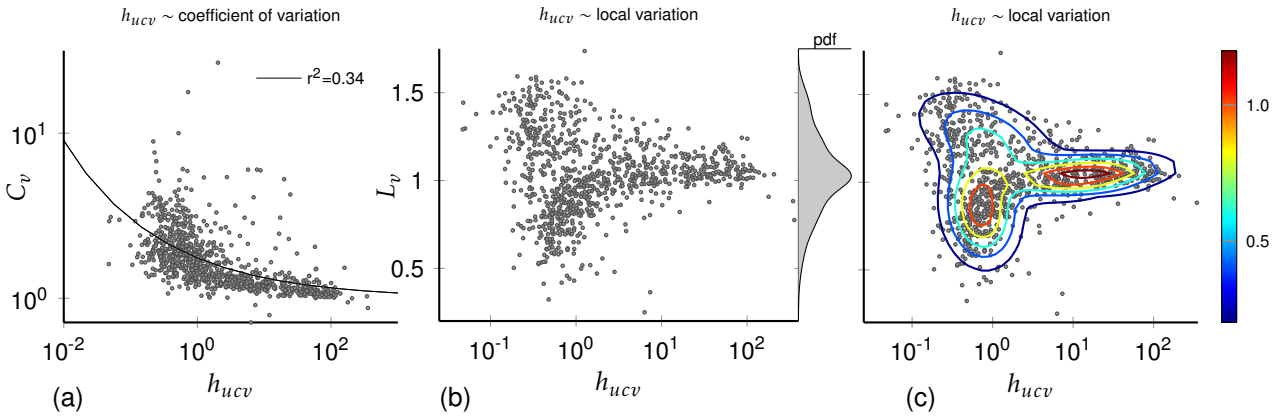


Figure 3.7: Correlation between optimal bandwidth estimates and **a:** the coefficient of variation and **b:** the local variation for $N=905$ units. **c:** Joint distribution $p(h_{ucv}, L_v)$ by fitting a three-component Gaussian mixture model

Although we did not observe explicit correlation between the local variation and bandwidth estimates, scatter plots showed at least two or three discernible clusters. We next fitted their joint distribution $p(h_{ucv}, L_v)$ by a three-component Gaussian mixture model (fig. 3.7 c). One cluster was centered around $L_v \approx 1$ and scattered along a line parallel to the h_{ucv} -axis and two more clusters at $h_{ucv} = 10^{-0.07}$ and $h_{ucv} = 10^{-0.15}$ scattered parallel to the L_v -axis. According to the definition of the L_v (as introduced in sec. 2.1) points falling within the first cluster centered at $(10^{1.19}, 1.04)$ represent Poisson-spiking units ($L_v \approx 1$) with high bandwidth estimates $> 10^1$. More regularly spiking cells ($L_v < 1$) fall in the second cluster centered at $(10^{-0.15}, 0.84)$, while units firing irregularly ($L_v > 1$) lie within the third cluster centered at $(10^{-0.07}, 1.25)$.

Average spiking activity. Figure 3.8 shows the relationship between base-10 logarithm transformed average firing rates, the local variation L_v and optimized bandwidth estimates. To establish a connection to the previous findings (fig. 3.7), we computed the correlation between average firing activity of cells and their local irregularity measure L_v (fig. 3.8 a). Both are negatively correlated with $r = -0.72$ (p-value = $1.35e-142$, Pearson correlation coefficient) and could be fitted by a linear model $L_v = -0.31 \cdot \langle \nu \rangle + 1.02$ (F: 945, p-value = $1.35e-142$, F-statistic vs. constant model) which explains $r^2 = 0.51$ of the variation in the data. It is worth noting that this relationship implies that cells spiking more irregularly with L_v above unity exhibit low average firing rates $\langle \nu \rangle < 1$ Hz (fig. 3.8 a, left upper quadrant), while units with higher average firing activity $\langle \nu \rangle > 1$ Hz are spiking more regularly $L_v < 1$ (fig. 3.8 a, right lower quadrant).

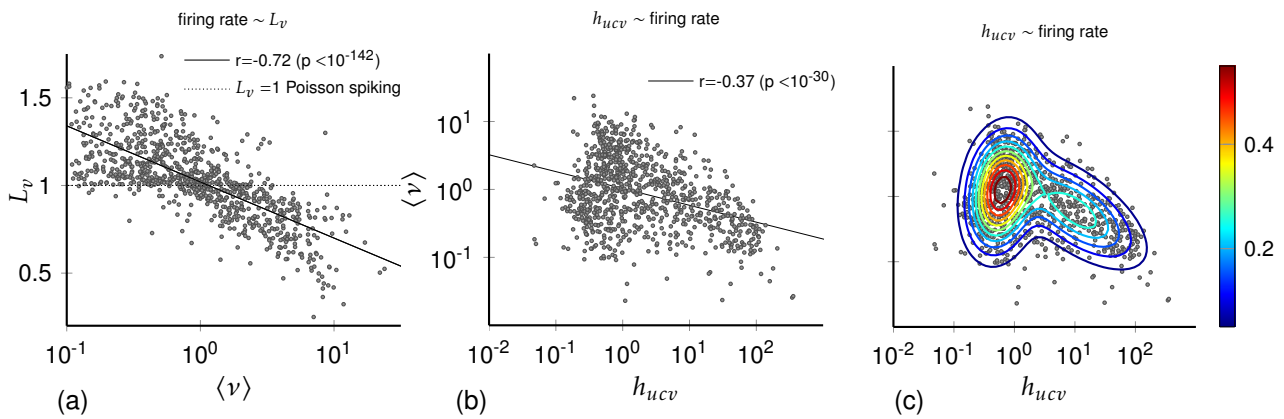


Figure 3.8: Correlation between average firing activity, optimal bandwidth estimates and the local variation. The log-transformed mean firing rate $\langle \nu \rangle = 1/\langle \text{ISI} \rangle$ and **a**: local variation L_v and **b**: log-transformed bandwidth estimates. **c**: Obtained joint distribution $p(h_{ucv}, \langle \nu \rangle)$ by fitting a two-component Gaussian mixture model.

In general, bandwidth estimates and average firing rate of single units were significantly correlated ($r = -0.37$, p-value = $1.63 \cdot 10^{-30}$, Pearson correlation coefficient). We quantified the relationship by fitting a linear model of the form $\nu = -0.25 \cdot h_{ucv} + 0.014$, which could explain 14% of the variation ($r^2 = 0.14$, F: 142, p-value < $1.63e-30$, F-statistic vs. constant model). Furthermore, scatter plots displayed distinct clusters. We fitted a two-component Gaussian mixture model to characterize the joint distribution of single unit UCV-based bandwidth estimates and average firing rates $p(h_{ucv}, \langle \nu \rangle)$. Both components are visualized as two ellipses in figure 3.8 b. The first cluster was centered at $(10^{-0.23}, 10^{0.06})$ with a mixing proportion of 0.58 and a second cluster at $(10^{1.0}, 10^{-0.22})$ with a mixing proportion of 0.42. By extracting correlation coefficients of the Gaussian covariances we found that for points falling within the first cluster bandwidth estimates and average firing activity are positively correlated with $r = 0.18$, while for points falling within the second component h_{ucv} and $\langle \nu \rangle$ are negatively correlated with $r = -0.64$.

3.4 Summary & Discussion

Neither bandwidth optimization nor the application of Kernel density estimates to neurophysiological data is a novel approach. However, the sort of information which UCV-based bandwidth estimates of single units alone or in combination with other measures enable to retrieve, e.g. on the temporal structure, global or local regularity of a spike train or its encoding properties, is unique and can be employed to answer qualitatively different questions. The main findings of this chapter can be summarized as follows:

First, bandwidth estimates are tuned to the temporal structure of spike trains and there exists an explicit functional dependence between UCV-based the bandwidth outcome and the precision of spiking patterns which also accounts for signal-to-noise ratio and average firing activity of a spike train. Second, by providing a characteristic value for the temporal spiking structure, bandwidth estimates can highlight single units encoding for task-related events. Third, the distribution of bandwidth estimates reveals distinct subgroups of cells with common firing properties.

Bandwidth estimates are tuned to the temporal structure of surrogate spike trains

The simulations we conducted show that smaller bandwidth estimates are associated with first, higher precisions of the spiking patterns, second, higher signal-to-noise ratio within the structure of spiking activity and third, higher average firing rates of spike trains. Moreover, the relationship between the temporal structure of the simulated spike trains and optimized bandwidth estimates can be fully described by a sigmoid function given in eq. 3.3 (fig. 3.3 c) which provides unique parameter solutions for different temporal structure settings: SNR, ν .

The first two findings, relationship between bandwidth estimates and the precision and SNR of spike train patterns are in line with general properties of KDEs and reflect that the lower variability, i.e. the more densely points cluster within a restricted region the smaller bandwidths are needed to represent the underlying distribution (Scott and Terrell, 1987; Silverman, 1986; Rudemo, 1982; Bowman, 1984; Simonoff, 1996). Similarly, higher spike rates have lower mean interspike intervals and variances (Gerstner and Kistler, 2002; Dayan and Abbott, 2005), i.e. lower variability and thus higher average firing rates implicate lower bandwidth estimates.

The sigmoid functional dependence between bandwidth estimates and the precision of spiking patterns which also takes into account signal-to-noise ratio and average firing activity of the spike train $h_{ucv}(\sigma_j|\text{SNR}, \nu)$ implies that there most likely exists an analytical solution for h_{ucv} which can then be expressed as a function of the temporal structure. Consequently, if this is the case, bandwidth estimates and temporal structure of spike trains are not just correlated, but estimates must to some extent reflect exact information on particular aspects of the spiking patterns.

Bandwidth estimates can highlight single units encoding for task-related events

We found significant correlation between higher decoding performance when predicting task-related events with LDA and small single unit bandwidth estimates (fig. 3.6 b). This finding implies that spike trains of encoding units must convey temporally precise spiking patterns which are time-locked to these events and is in agreement with results for surrogate data (in sec. 3.2) that UCV-based bandwidth estimates highlight the temporal structure in spike train data. Units which had low bandwidth estimates but high prediction error values might fall into the following categories: units which are not discriminating events in their activity but have a high average firing rate (supported by fig. 3.8 b) or bursty spiking activity. Also encoding units could have high prediction error values for instance if they exhibit a high trial-by-trial variability, are locked to all stimuli, but do not discriminate in their firing rate, or reflect internal processing (e.g. delay activity or reward).

Distribution of bandwidth estimates reveals distinct subgroups of cells

To examine in more detail what sort of temporal information can be revealed with UCV-based bandwidth estimates we analyzed mPFC in-vivo multiple single unit recordings of behaving rats. In-vivo recordings show that the distribution of bandwidth estimates exhibits at least two distinct modes. Exemplary units depicted from the distinct modes display different types of spiking activity structure with respect to recurrence of spike triplet patterns or locking to task-related events. Similar to findings for surrogate data summarized in the previous paragraph, temporally more structured spiking activity seems to be correlated with lower optimized bandwidth outcomes (fig. 3.5).

More detailed analysis across all units in combination with other spike train metrics and regularity measures demonstrates either significant correlation or the presence of multiple clusters. We found that first, probabilities of recurrent spike triplet patterns (fig. 3.6 a) and global spike train irregularity, i.e. coefficients of variation (fig. 3.7 a) are significantly correlated with lower bandwidth estimates. Second, examining the joint distribution with the average firing rate shows the presence of two and with the local irregularity measure L_v the presence of three clusters (fig. 3.8 and 3.7 c). One cluster is centered at L_v close to unity and higher bandwidth values ($>10^1$), and two more clusters centered below and above $L_v = 1$, and lower h_{ucv} values $<10^0$. These clusters represent cells which are spiking random or Poisson ($L_v \approx 1$, first cluster), locally regular ($L_v < 1$, second cluster) or locally irregular ($L_v > 1$, third cluster) (Shinomoto et al., 2002). It is worth noting, that by examining the L_v distribution alone it is not possible to discriminate these three separate components as it has only one mode centered around unity (fig. 3.7 b, right panel). This is in agreement with the L_v distribution for PFC neurons which has been reported by Shinomoto et al. (2003). Also the clear negative correlation we found for average firing rates $\langle \nu \rangle$ and local irregularity values of spike trains, i.e. high $\langle \nu \rangle$ are associated with low L_v , has been reported in other studies which show either a power-law relationship for awake rat PFC (Peyrache et al., 2009) or a linear relationship for anesthetized cat V2 neurons (Blanche et al., 2005).

The components of the joint $p(h_{ucv}, L_v)$ distribution might be categorized in the following way: units with low bandwidth estimates, locally regular spiking activity and thus high firing rates which fall in the second cluster might represent putative interneurons, while units with low bandwidth estimates, locally irregular spiking activity and thus low firing rates which fall in the third cluster might represent puta-

tive pyramidal cells. This categorization is supported by other studies which used the L_ν distribution either to classify cells from distinct functional brain areas or to distinguish distinct cell classes from the same brain region (Shinomoto et al., 2002, 2009, 2005). Also, by combining several spike train metrics, $\langle \nu \rangle$, C_ν , L_ν and two more spike waveform measures, Ardid et al. (2015) were able to segregate several functional classes of broad spiking putative pyramidal cells and narrow spiking putative inhibitory cells in the PFC of macaques engaged in an attention task. Authors reported that putative inhibitory cells were associated with lower L_ν and higher $\langle \nu \rangle$ and vice versa for putative pyramidal cells which is in agreement with our categorization.

Summarizing, qualitative information conveyed by UCV-based bandwidth estimates clearly differs from other spike train irregularity measures, as L_ν , C_ν and also from general spike train statistics as the average spike rate $\langle \nu \rangle$. Therefore, combining these measures might represent a powerful tool for discrimination of different cells types as it was already shown for the local variation (Shinomoto et al., 2002, 2009, 2005) or multiple measures in combination with clustering analyses Ardid et al. (2015). Concluding, apart from their application in firing rate estimation in the first place, optimized bandwidths can be used to solve different problems: give a characteristic value of temporal precision within spiking patterns, discriminate task-related units and possibly even distinguish different cell types.

4 Validity of optimized KDEs in neuronal decoding

In the previous chapter we have studied general features and characteristics of UCV-based optimized bandwidths obtained for in-vivo spike train data from the rat mPFC of behaving animals. We have seen that optimized bandwidth provide an information rich measure reflecting temporal structure or encoding properties of single units. In this chapter we will examine properties of optimized KDEs in decoding, i.e. when KDEs obtained with optimized bandwidths are applied in subsequent decoding analyses.

The goal of this chapter is two-fold: First, we aim to assess if UCV method gives a reliable bandwidth estimate when optimized KDEs are applied subsequent classification and second, compare the decoding performance achieved with optimized KDEs to non-optimized KDEs.

This chapter is structured in the following way: First, we will describe the used measures to assess the decoding validity of optimized KDEs, next will apply the validity measures to surrogate data and to mPFC in-vivo recordings, then we will compare decoding performance of optimized to non-optimized KDEs in-vivo data and finally we will draw which conclusions can be made based on the outcome.

4.1 Validity measures

We have described the procedure to obtain optimized KDEs in section 2.2, how to decode task-related events in section 2.4 and how compute a generalized, expected prediction error of a classifier by cross-validation in section 2.5. In order to examine how reliable the obtained optimized bandwidth estimate are, we analyzed how the classification performance changes, depending on to which extent we scale single unit bandwidth estimates. As the scaling will effect smoothness of instantaneous firing rates and thus also the decoding performance.

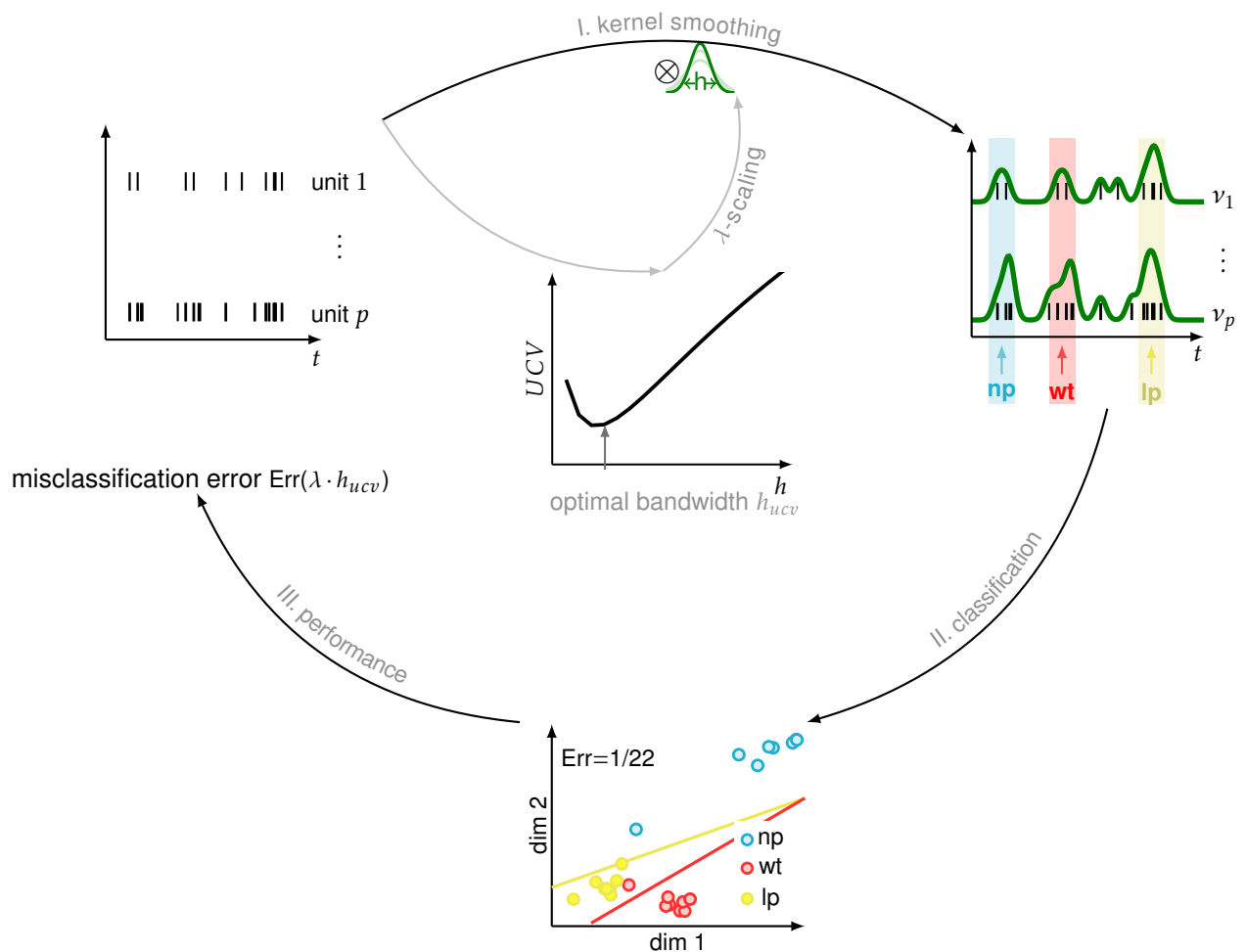


Figure 4.1: Illustration of the two-stage approach for evaluating the validity of optimized bandwidth estimates in decoding. Before obtaining firing rates, UCV-based bandwidth estimates are scaled by a parameter λ and after subsequent decoding the validity of the UCV method is evaluated by measuring the misclassification error as a function of λ .

4.1.1 Evaluation of UCV-based bandwidth estimates in decoding

We implemented the LDA in combination with the 3-fold cross-validation scheme (as introduced in sec. 2.4 and 2.5). Experimental data: only correct trials were included. For the decoding procedure we took class labels from alternation task: 4s delay period preceding nose poke, nose-poke and lever press. Sequence-switch task: nose-poke, wheel-turn and lever-press. Instead of using optimized KDEs $\nu(t|h_{ucv})$ to calculate the prediction error $Err(h_{ucv})$, population vectors of firing activity were obtained employing a scaled version of the bandwidth estimate $\nu(t|\lambda \cdot h_{ucv})$, $\lambda \in (0, 20]$, and then used for subsequent stimulus class prediction (fig. 4.1). To test if the true bandwidth estimate gives reliable unbiased classification results, i.e. if the scaling does not significantly improve prediction, we determined two measures:

First, the optimal scaling which gives the best prediction and second, the percentage deviation in decoding performance of the true compared to the best scaled bandwidth estimate.

The optimal scaling parameter λ_{opt} is then defined as the value for which $Err(\lambda \cdot h_{ucv})$ achieves best prediction (eq. 4.1) and the percentage difference in prediction error represents the deviation when using the true bandwidth estimate relative to the best scaled estimate (eq. 4.2).

Both measures λ_{opt} and δ_{Err} indicate if a scaling λ exists which allows more accurately to discriminate between distinct stimulus classes and neuronal activity states, or, whether the UCV method provides a sufficiently valid and unbiased estimate (when $\lambda_{opt} \approx 1$ and $\delta_{Err} \approx 0$) so that decoding performance cannot be improved by bandwidth scaling.

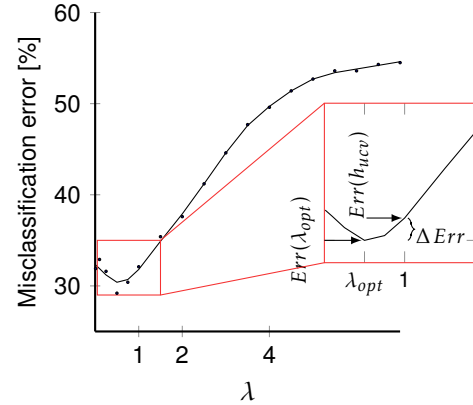


Figure 4.2: Illustration of bandwidth validity measures and decoding performance as a function of the scaling parameter λ , dataset 15 alternation task. Smoothed (solid black line) and raw (left plot, dots) misclassification error values as a function of the scaling parameter λ follow a v-shaped curve with a pronounced minimum at λ_{opt} . The discrepancy between true error $Err(h_{ucv})$ at $\lambda = 1$ and the best error $Err(\lambda_{opt})$ can be measured as absolute ($\Delta Error$) or relative value (δ_{Err}).

$$\lambda_{opt} = \underset{\lambda}{\operatorname{arg\,min}} \operatorname{Err}(\lambda \cdot h_{ucv}) \quad (4.1)$$

$$\delta_{Err} = \frac{\operatorname{Err}(h_{ucv}) - \min \operatorname{Err}(\lambda \cdot h_{ucv})}{\max \operatorname{Err}(\lambda \cdot h_{ucv}) - \min \operatorname{Err}(\lambda \cdot h_{ucv})} \times 100 \quad (4.2)$$

We computed δ_{Err} and λ_{opt} applying the following procedure to each data set:

```

.for units  $i = 1$  to  $p$ 
.   1. compute optimal UCV-based bandwidth estimates  $h_{ucv}$ 
2. define a set of  $n$  scaling parameters  $\lambda = \{\lambda_1, \dots, \lambda_n\} \in (0, 20]$ 
.for  $j = 1$  to  $n$ 
.   3. scale the true estimates  $\lambda_j \cdot h_{ucv}$ 
.   4. obtain KDEs by plugging in scaled bandwidth estimates  $\nu(t | \lambda_j \cdot h_{ucv})$ 
.   5. estimate decoding performance  $Err(\lambda_j \cdot h_{ucv})$  based on 3-fold-CVE scheme (sec. 2.5)
6. smooth resulting  $Err(\lambda)$ -functions with the MATLAB built-in cubic spline function 'csaps'
7. determine scaling factor  $\lambda_{opt}$  which gives best decoding performance
.    $\lambda_{opt} = \underset{j}{\operatorname{arg\,min}} Err(\lambda_j \cdot h_{ucv})$ 
8. determine the relative deviation of the true error  $Err(h_{ucv})$  from the best error  $Err(\lambda_{opt})$ 
.    $\delta_{Err} = [Err(h_{ucv}) - Err(\lambda_{opt}) / (\underset{j}{\operatorname{arg\,max}} Err(\lambda_j \cdot h_{ucv}) - Err(\lambda_{opt}))] \times 100$ 

```

Bandwidth validity pseudo-code.

4.1.2 Comparing decoding with optimized on non-optimized bandwidths

Since in experimental settings most commonly used bandwidths are chosen arbitrarily which makes them equal or fixed across units for a given data set, we examined how the classification performance changes, depending on what bandwidth selection method is used for prior kernel smoothing.

We computed the decoding performance as a function of the fixed bandwidth $h \in (0, 20]$ and identified the best non-optimized bandwidth as the value which gives the lowest prediction error $h_{opt} = \arg_h \min Err(h)$. We employed Wilcoxon one-sided signed rank test for paired samples to compare $Err(h_{ucv})$ (optimized) and $Err(h_{opt})$ (non-optimized) decoding outcomes.

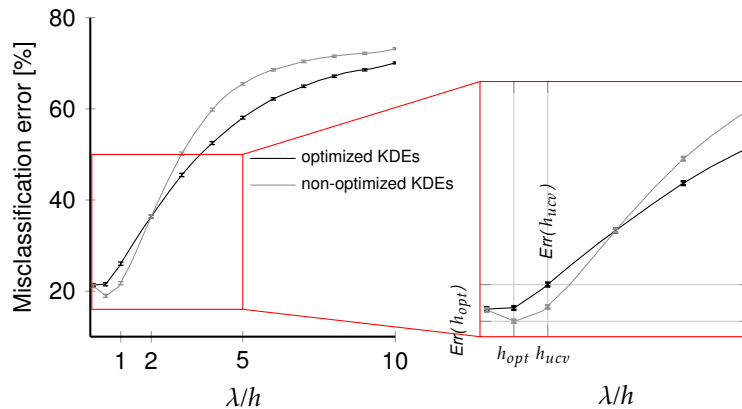


Figure 4.3: Decoding performance with scaled optimized and non-optimized KDEs. The misclassification error is shown as a function of the bandwidth scaling λ for optimized KDEs (black lines) and dependent on a fixed bandwidth h for non-optimized KDEs (gray lines) which is equal across all units for a given data set.

Selecting different cross-validation error schemes and subsets units from experimental data.

Both bandwidth selection methods were probed using different subsets of units for a given data set. Employing optimized KDEs and also spike counts over 500 ms, we identified for each data set the most predictive subset by means of sequential unit elimination starting with the complete set and successively removing units which most improve the misclassification error (for the complete description of the method see appendix 5.4). So that for each of the 19 experimental data sets used for our analyses we identified and compared the performance using three different sets:

First, the complete set, i.e. all units of one recording session; second, the most predictive set of cells when the firing rate was computed as spike count over 500 ms; third, the most predictive set of cells when the firing rate was obtained by Kernel density estimation with UCV-based bandwidths h_{ucv} .

Deploying these sets we also compared the decoding performance of optimized to non-optimized KDEs when the misclassification error was calculated using different cross-validation schemes. We applied 3-fold and m-fold CV schemes, both are discussed in section 2.5.

4.2 Surrogate data: validity of optimized bandwidth estimates dependent on distinct firing states

4.2.1 Generation of spiking activity with multiple firing states

To analyze the reliability of the optimal bandwidth estimates in more detail we generated surrogate spike trains of known underlying structure which reflect response-class specific firing activity $\nu(t|c_k)$ during performance of task-related events and applied the previously introduced validity measures.

In order to simulate animal behavior and to generate surrogate spike train data we implemented a traditional hidden Markov model (HMM) framework (Rabiner, 1989) which has been widely applied and extended to spike train data (Abeles et al., 1995; Seidemann et al., 1996; Jones et al., 2007; Yu et al., 2006; Escola et al., 2011). In this context, neuronal activity can be characterized by a HMM which states that responses of different neurons reflect a common dynamical process in the network, a network state. These states use to drive spiking activity of the neurons. Transitions from one state to another can undergo abruptly, at variable times as described in models of network dynamics (Hopfield, 1982; Amit, 1992).

The implementation of the HMM is based on the following assumptions

1. Neuronal activity can be characterized by a small number of hidden states s which correspond to average firing rates $\nu_i(s)$ of cells i . At every moment in time, the system is in one of these states.
2. In each state the neurons fire according to an approximately stationary homogeneous Poisson process with a constant firing rate. whereas the precise spike timing is random.
3. Hidden states change as a time-homogeneous Markov chain, i.e. the probability of a transition from one state i to another state j from one time bin to the next, denoted as $p_{i,j}$, is constant over time and can be summarized by transition probability matrix T . For states to be persistent for much longer than the time bin of 1 ms, the diagonal elements of T are almost unity, $p_{i,i} \sim 1$.

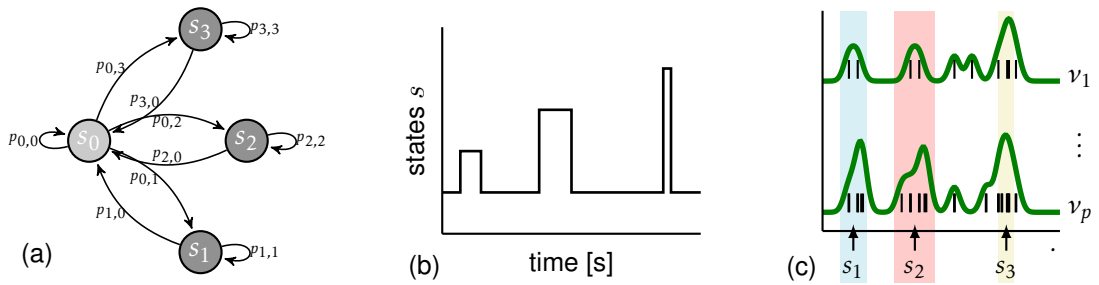


Figure 4.4: Steps of the HMM simulation. **a:** Markov chain with nodes representing states and edges possible state transitions and assigned transition probabilities. **b:** Markov state sequence corresponding animal behavior for one simulated trial. **c:** Markov states drive spiking activity and give rise to different ifr profiles of several units.

We first simulated animal behavior by a Markov state sequence with $S = 4$ distinct states corresponding a resting state s_0 and three events/behaviors s_1, s_2, s_3 performed on task based on the transition proba-

bilities shown as edges of the Markov chain shown in fig 4.4 a and summarized in the transition matrix T .

$$T = \begin{bmatrix} p_{0,0} & p_{0,1} & p_{0,2} & p_{0,3} \\ p_{1,0} & p_{1,1} & p_{1,2} & p_{1,3} \\ p_{2,0} & p_{2,1} & p_{2,2} & p_{2,3} \\ p_{3,0} & p_{3,1} & p_{3,2} & p_{3,3} \end{bmatrix} = \begin{bmatrix} 1 - 3 \cdot 10^{-5} & 10^{-5} & 10^{-5} & 10^{-5} \\ 10^{-4} & 1 - 1 \cdot 10^{-4} & 0 & 0 \\ 10^{-4} & 0 & 1 - 1 \cdot 10^{-4} & 0 \\ 10^{-4} & 0 & 0 & 1 - 1 \cdot 10^{-4} \end{bmatrix}$$

Apart from recurrence to the same state, transitions were possible from the resting state to any other state, however, vice versa from a non-resting state only to s_0 (e.g. fig. 4.4 a). Probability of remaining in s_0 within the next time bin of 1 ms was set to $1 - 3 \cdot 10^{-5}$ and to $1 - 1 \cdot 10^{-4}$ for any other state, transition from s_0 to any other state 10^{-4} , and back to s_0 10^{-5} . Minimum duration of persistence in a state was set to 500 ms, transitions were then arranged in sequential order so that each trail consists of the following state sequence $(s_0, s_1, s_0, s_2, s_0, s_3, s_0)$ (fig. 4.4 b) as in the experiment. The states were then used to drive spiking activity of the model neurons and gave rise to different firing rate profiles ν_i for each state s (fig. 4.4 c). The firing rate of the resting state is set to $\nu_i(s_0) = 0.1$ Hz.

To be as close to experimental conditions as possible we analyzed empirical firing rate profiles for each behavior (non-resting state) prior to surrogate data generation. The distribution of average firing rates within 1 sec time windows centered around behavioral events of all units and data sets is shown in figure 4.5 a (black line), the mean and variance of firing rate profiles across all task-related events and cells for each data set $\langle \nu \rangle_{s,i}$ is shown in figure 4.5 b. From figure 4.5 a we concluded that average state-specific single-unit firing rates $\nu_i(s)$ are log-normally distributed with mean μ_s and variance σ_s^2 which can be fitted by a log-normal probability density function $\ln \mathcal{N}(\mu_s, \sigma_s^2)$ with $\mu_s = 1.75$ and $\sigma_s^2 = 8$ (fig. 4.5 a, gray line).

Additionally, as reported by Abeles et al. (1995) state transitions are associated not only with a change in the firing rate profile of several units, but also pairwise correlations between the units vary between the different states. To account for these pairwise correlations of surrogate spike trains multivariate binary patterns were generated at a millisecond resolution from a dichotomized Gaussian distribution (Cox and Wermuth, 2002; Macke et al., 2009) with population mean and correlation structure specified by the Markov sequence state and converted to spike times (MATLAB implementation of the binary pattern generation algorithm adapted from Macke et al., 2009). We double checked that assumption (2) was met and for each state ISIs were exponentially distributed, i.e. model neurons fired according to an approximately stationary homogeneous Poisson process with constant firing rate.

Thus, to analyze optimized bandwidth validity systematically, with known underlying ground truth, but at the same time to remain as close to experimental conditions as possible, each HMM comprised $p = 10$ units, $s = 4$ states and the following state-dependent, multivariate spiking activity parameters: pairwise correlation coefficients $\rho \in \mathbb{R}^{p \times s}$, 3 non-resting states with log-normally distributed average firing activity $\nu(s) \sim \ln \mathcal{N}(\mu_s, \sigma_s^2)$ with mean $\mu_s \in \{1, 3, 5\}$ and variance $\sigma_s^2 \in \{10, 15, 20\}$ and resting state s_0 with $\nu(s_0) = 0.1$ Hz, $\nu(s) \in \mathbb{R}^{p \times s}$.

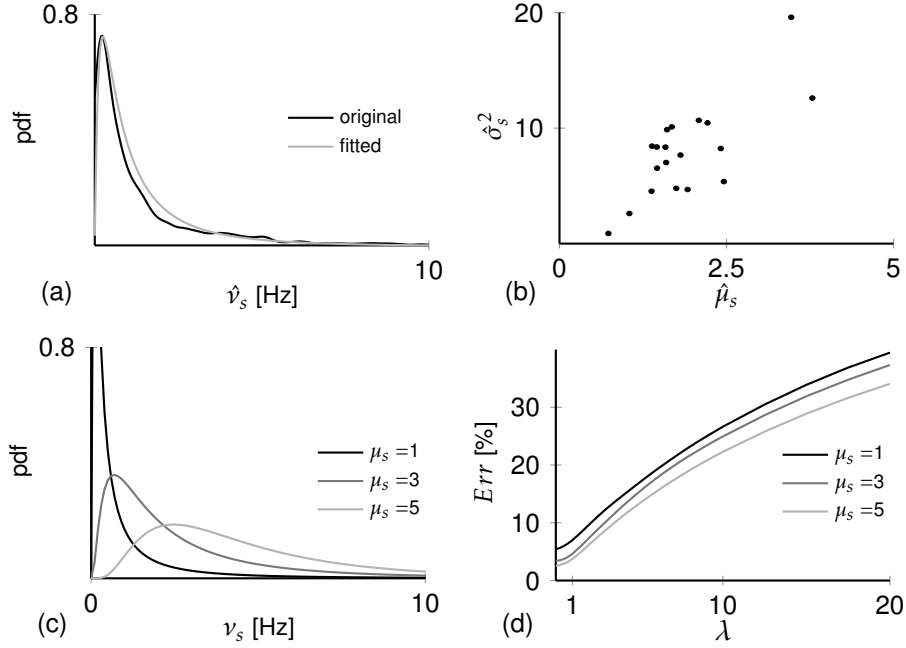


Figure 4.5: Experimental and surrogate firing rate profile distributions of the Markov sequence states. **a:** Empirical distribution of firing rates during different behaviors on the task (black) and log-normal fit with $\mu_s = 1.75$ and $\sigma_s^2 = 8$ (gray). **b:** Mean and variance of firing rates across different behaviors and cells per data set. **c:** Surrogate state firing rates are drawn from a log-normal distribution e.g. here: $\sigma_s^2 = 20$ **d:** Corresponding misclassification error curves as a function of the bandwidth scaling averaged over $N=100$ models for state firing rates drawn from the distribution shown in c.

We conducted the following procedure to compute δ_{Err} and λ_{opt} for simulated data sets:

```

for  $\sigma_s^2 = \{10, 15, 20\}$ 
.   for  $\mu_s = \{1, 3, 5\}$ 
.     generate  $N = 100$  models a  $p = 10$  units and  $s = 4$  states
.     for  $j = 1$  to  $N$ 
.       1. set state-dependent HMM parameters  $\rho, \nu(s) \in \mathbb{R}^{p \times s}$ 
.         .draw equally distributed  $\rho_s \in (0, 0.02]$  and  $\nu(s_i) \sim \ln \mathcal{N}(\mu_s, \sigma_s^2), i = 1, \dots, 3$ 
.       2. generate a state sequence of 60 min based on transition probabilities defined in  $T$ 
.       3. produce Poisson spike train output from parameters in (1) and (2)
.       4. obtain optimal UCV-based bandwidth estimates  $h_{ucv}$ 
.       5. compute the error function dependent on the scaling parameter  $Err_j(\lambda \cdot h_{ucv})$ 
.       6. average over  $N=100$  models to obtain the expected prediction error
.          $\langle Err(\lambda \cdot h_{ucv} | \mu_s, \sigma_s^2) \rangle = 1/N \sum_j Err_j(\lambda \cdot h_{ucv})$ 
.       7. from averaged error curves  $\langle Err(\dots) \rangle$  estimate bandwidth validity  $\lambda_{opt}(\mu_s, \sigma_s^2)$  and  $\delta_{Err}(\mu_s, \sigma_s^2)$ 

```

Model generation and validation procedure.

4.2.2 UCV-bandwidth validity for simulated data sets

Deviation from the best prediction performance of actual compared to scaled optimized bandwidths is very low ($\delta_{Err} < 5\%$) when units have higher state-specific firing rates compared to background activity. Employing bandwidth validity measures to averaged cross-validated misclassification error curves (as shown in fig. 4.5 d for $\sigma_s^2 = 20$, $\mu_s \in \{1, 3, 5\}$) yields results showing that with high signal-to-noise ratio, i.e. increasing simulated state firing rates ν_s the deviation from the best prediction performance of the true compared to scaled optimized bandwidth estimate converges to zero (fig. 4.6 a,b).

This essentially means that the higher state-dependent firing rates in surrogate data and firing rates in response to stimulus presentation compared to the background firing activity in in-vivo recordings the more reliable the application of optimized KDEs in decoding will be. However, the optimal scaled bandwidth values which gave the best prediction do not match exactly the actual ones (fig. 4.6 c,d).

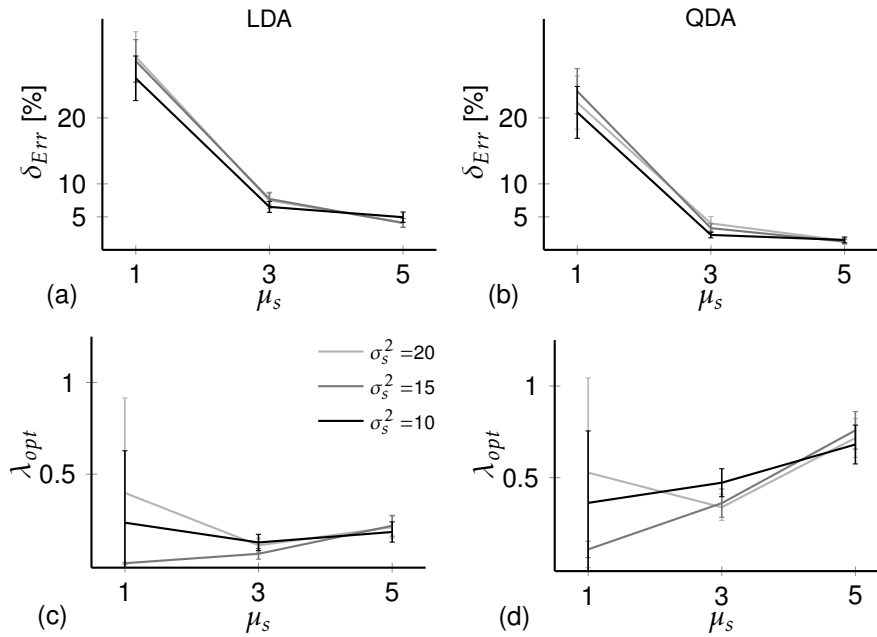


Figure 4.6: Bandwidth validity for simulated data sets, N=10 units. a, b: Percentage deviation in misclassification error when using the true bandwidth estimate relative to the best scaled bandwidth. **c, d:** Optimal scaling which gives the best prediction.

Applying QDA which takes into account that covariance matrices of different activity states vary in contrast to LDA which assumes their identity, yielded qualitatively similar results, however outperformed LDA significantly by a lower deviation in prediction error (fig. 4.6 b) and an optimal scaling closer to the true estimate (fig. 4.6 d).

4.3 Decoding spiking activity of in-vivo recordings from the rat mPFC

4.3.1 Bandwidth validity for experimental data

Decoding accuracy as a function of the scaling λ of UCV-based bandwidth estimates follows a v-shaped curve with a pronounced minimum centered around one (fig. 4.7 a, b). This results implicate that the best prediction performance is achieved with the unscaled, actual estimate h_{ucv} . The summary across all experimental data sets shows that when using optimized bandwidths for decoding for 80% of the data (15 out of 19 sets) the misclassification error differs from the minimum by less than 10% (fig. 4.7 c, d: $\lambda_{opt} \approx 1$ and $\delta_{Err} < 10$).

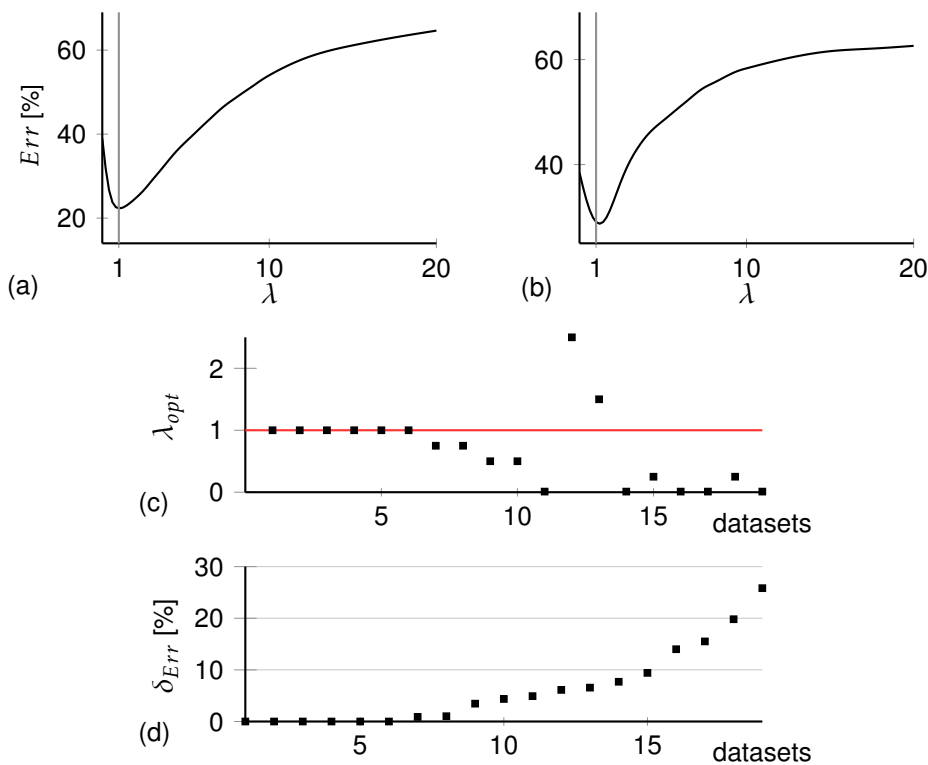


Figure 4.7: Validity measures and decoding performance as a function of the bandwidth estimate. Illustration of the prediction error as a function of the scaling parameter λ for two exemplary data sets from **a**: task two (alternation task, data set 13) and **b**: task one (sequence-task, data set 12). Confidence intervals were $\leq 10^{-1}$ and left out for clarity reasons. **c**: Optimal scaling which gives the best prediction. **d**: Percentage deviation in misclassification error when using the true bandwidth estimate relative to the best scaled optimized bandwidth. Both measures λ and δ_{Err} are sorted in order of increasing δ_{Err} values.

4.3.2 Decoding performance compared to non-optimized KDEs

Decoding performance with optimized and non-optimized KDEs was probed on different sets of units for a given data set and the misclassification error was calculated using two different cross-validation schemes. To estimate the prediction error we employed 3-fold and m -fold CV. For 3-fold CV the classifier was trained on two third and tested one third of the trials. KDEs of the same response-class were averaged block-wise per trial respectively. For m -fold CV we employed non-averaged KDEs and the classifier was trained on $m-1$ and tested on the left out trial respectively (see sec. 2.5). The different sets of units comprised the complete set of units for a given recording session, the most predictive set when decoding was performed with spike counts and the most predictive set when using optimized KDEs (sec. 4.1.2). The resulting prediction errors for the two different bandwidth selection methods were then compared using the one-sided Wilcoxon signed rank test for paired samples to determine which method performed superior.

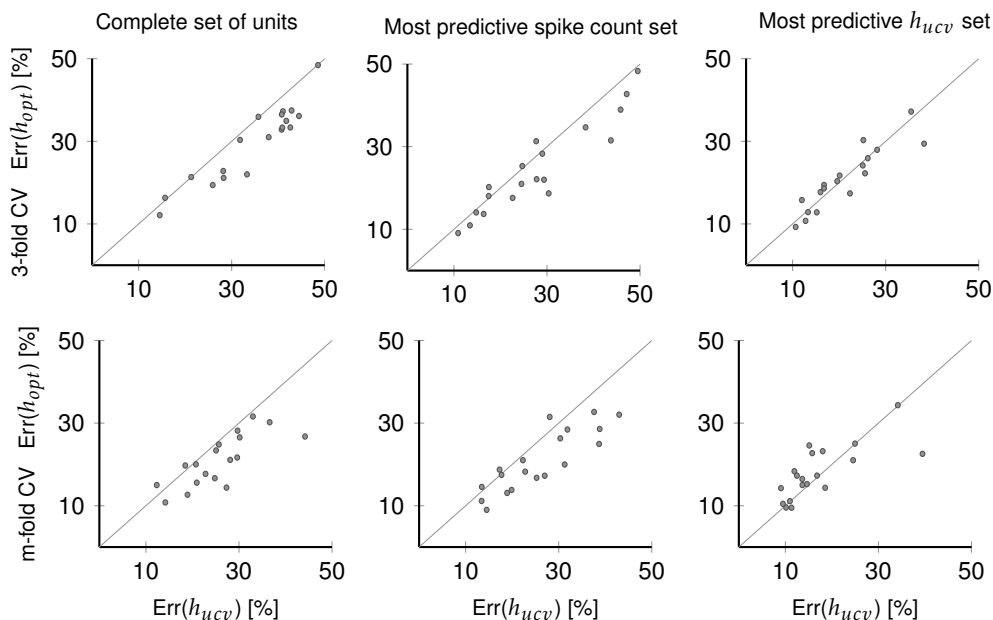


Figure 4.8: Comparison in decoding performance of optimized KDEs and most predictive non-optimized KDEs with different sets of units (columns arranged from left to right), 3-fold and leave-one-trial-out cross-validation (upper and lower panel). Each point corresponds to misclassification errors for one data set obtained with optimized (x-axis) and non-optimized KDEs (y-axis). The gray bisector indicates the positions at which values of both methods are identical.

For the different conditions we tested (2 CV schemes, 3 unit subsets) the overall best decoding performance across all the data sets on average was achieved with optimized KDEs of the most predictive subset of units when employing m -fold CV ($Err=14.8\pm 1.1\%$ and $Err=17.1\pm 1.9\%$ with and without excluding the outlier at $Err=34.2$ and 39.5%). This decoding performance was significantly better compared to the best achieved decoding with non-optimized KDEs for the same condition ($p<0.018$, $z=2.09$, fig. 4.8, lower right panel). And also when comparing the same optimized with non-optimized KDE results across the different sets of units but the same CV scheme optimized KDEs yielded sig-

nificantly better performance (complete sets: $p < 9.94e-04$, $z = 3.09$; most predictive set based on spike counts: $p < 0.0018$, $z = 2.91$). Comparing both methods for m-fold CV and the same sets of units directly yielded significantly poorer performance for optimized KDEs for all the other cases ($p < 3.36e-04$, $z = 3.40$, $p < 5.20e-04$, $z = 3.28$, fig. 4.8, lower left and middle). Employing 3-fold CV resulted in a similar outcome: optimized KDEs performed poorer for same conditions (complete sets: $p < 2.4e-04$, $z = 3.48$; most predictive set based on spike counts: $p < 0.0023$, $z = 2.84$, fig. 4.8, top left and middle panel) or equally with most predictive optimized KDEs ($p = 0.71$, $z = 0.36$, fig. 4.8, top right).

Summarizing, optimized KDEs yielded the best overall decoding performance which is significantly better or equally good compared the best performance of non-optimized KDE for the analyzed conditions (fig. 4.8 top and lower right panel). However, the classification accuracy highly depends on the subset of units used for firing rate estimation and subsequent classification: while employing the most predictive subset of optimized KDEs outperforms non-optimized KDEs, for decoding with other sets of units the outcome is significantly poorer. Decoding accuracy also depends on the validation method when estimating the prediction error.

4.4 Summary & Discussion

In a two-stage approach we first assessed how reliable we can predict task-relevant behavior of the animal when applying instantaneous firing rates obtained with optimized bandwidths, using both surrogates and in vivo recordings from mPFC multiple single-unit spike times of behaving rats, and second, how well we can decode experimental data with optimized compared to non-optimized KDEs.

Our findings show that first, in experimental data decoding outcome with optimized KDEs highly depends on the sets of units used for classification and also on the cross-validation scheme for estimation of the prediction error. Second, the UCV-method provides a reliable bandwidth estimate for 80% of the analyzed in-vivo data, i.e. the UCV-based differs from the best misclassification error when scaling optimized bandwidths by less than 10% when employing robust validation sets (3-fold CV). Third, this finding is in agreement with the results obtained for surrogate data which shows that prediction performance with actual KDEs deviates less than 5% from best performance achieved with scaled estimates for sufficiently high non-resting state activity (fig. 4.6). This implies that decoding outcome of optimized KDEs will be more reliable the higher state-dependent firing rates compared to the background activity in surrogate data and the higher firing rates in response to stimulus presentation in in-vivo recordings are.

In the following we will briefly review possible reasons why the decoding outcome for optimized KDEs varies depending on the set of single units and the CV choice when applied to mPFC in-vivo recordings. One reason for dependency on single units is that neuronal activity under same conditions evolves over time. Single-units can be responding to stimuli in some trials but being silent in other trials. This is in line with previous studies which report inter-trial (Nawrot et al., 2008), e.g. responding non-stationary when presented with the same stimulus, and long-term variability meaning that the state of neuronal mPFC ensembles systematically exposed to the same conditions shifts with time (Hyman et al., 2012;

Balaguer-Ballester et al., 2011). It is also in agreement with the general concept that neural ensembles encoding for different task events or stimuli are context-dependent (Balaguer-Ballester et al., 2011; Lapish et al., 2008). This may account for the flexibility needed during the performance of higher cognitive tasks such spatial and temporal context-encoding (Hyman et al., 2012), rule learning (Durstewitz et al., 2010), or decision making (Balaguer-Ballester et al., 2011). In non-optimized KDEs spike trains of all stationary and non-stationary units within a data set are convolved by a Gaussian kernel of the same width resulting in instantaneous firing rates which are smoothed to the same extent. Instead, in optimized KDEs individual spike trains are smoothed based on their individual bandwidth outcomes. A non-stationary unit encoding task-related events but which activity shifts over time will have a lower UCV-based bandwidth estimate (as pointed out in sec. 2.1 that firing rate fluctuations are associated with higher C_v and high C_v with low h_{ucv} values, sec. 3.3, fig. 3.7 a), give a more peaked instantaneous firing rate and thus add more variability to a given data set. Consequently, non-stationary units will contribute to a higher prediction error when employing optimized KDEs for classification. By removing units by sequential feature selection we can presumably identify the set of cells with firing rates which are stationary over time for same task events. Thus, to overcome this issue one can either remove units with high trial-by-trial variability, or alternatively, employ adaptive KDEs with variable bandwidths which adjust to non-stationary regions of spike trains (Sain, 1994; Hazelton, 2003; Terrell and Scott, 1992; Shimazaki and Shinomoto, 2010). Also multivariate Kernel density estimates or optimized KDEs developed for analysis of time-series might provide a more accurate alternative (Antoniadis et al., 2009; Bouezmarni and Rombouts, 2010; Tran, 2010; Zougab et al., 2014; Balaguer-Ballester et al., 2014).

The reasons why decoding results with optimized KDEs depend on the CV scheme are based on general statistical properties of the CV procedure: as pointed out by Hastie et al. (2009) the leave-one-out (here referred to as m-fold) CV will have low bias, but high variance since training sets are similar to each other. In 3-fold CV the test set comprises a time series with many auto-correlated samples. Thus, we computed mean KDEs by averaging consecutive blocks of the same event or class label for each trial. Decoding with temporally averaged class-specific firing rates results in a poorer performance (higher bias) compared to non-averaged KDEs, however lower variance. It also conveys qualitatively different information: when temporal patterns which are highlighted by optimized KDEs are averaged out, performance of optimized and non-optimized KDEs does not differ significantly (figure 4.8, upper right panel).

To conclude, although UCV-optimal bandwidth selection is an unsupervised method, which does not include any prior information on the timing of task-relevant events, it highlights temporal structure in spike train data and improves subsequent decoding performance. Thus, the UCV method presents a helpful tool for automated bandwidth selection and instantaneous firing rate estimation. The limitations of validity assessment of optimized KDEs in decoding are that the model we used to generate spiking activity with multiple firing states does not incorporate trial-by-trial (Nawrot et al., 2008), long-term variability (Hyman et al., 2012) or precise temporal spiking patterns present in in-vivo recordings (Mainen and Sejnowski, 1995; Bair and Koch, 1996; Buračas et al., 1998; Reinagel and Reid, 2002; Fellous et al., 2004; Brown et al., 2005; Raman et al., 2010) which poses the need for more analyses including parameters which specify temporal precision, inter-trial and long-term variability of spiking activity.

5 Outlook and possible extensions

UCV-based optimized bandwidths provide a information rich measure reflecting spiking activity of single cells and enhance classifier performance when applied to the recorded neuronal population (as seen in the previous chapters 4 and 3). However, employing optimized KDEs we can tackle many more recent challenging problems of the neuroscience community. Such might present single trial analysis of neural time courses, detection of encoding neuronal ensembles or the unfolding population dynamics (Brown et al., 2004; Baeg et al., 2007). Although there exist certainly many more possible considerable extensions related to optimized bandwidths and KDEs, this outlook will be devoted to show the above stated applications of optimized KDE to in-vivo recordings of the rat mPFC and, which general inferences can be made about neuronal coding mechanisms based on the outcome.

Apart from being involved in working memory the rat mPFC has been implicated in spatial and temporal context-encoding (Hyman et al., 2012) and rule deduction (Durstewitz et al., 2010). Furthermore it has been reported that the context-dependent organization of neuronal ensembles which encode for different task events or stages (Lapish et al., 2008; Balaguer-Ballester et al., 2011; Ma et al., 2014) may account for the great flexibility required during the performance of higher cognitive tasks.

Employing optimized KDEs to the activity of simultaneously recorded neurons from the rat mPFC during the sequence-switch task (sec. 5.4) we will first examine single-trial neural representation of the different sequences and actions in single unit activity. To enable across-trial comparisons of activity during self-paced events, we will use 'time-warped' optimized KDEs. Second, we will identify neuronal ensembles, i.e. subsets of units, encoding for task-related events by keeping those which maximally decrease the misclassification rate. And third, we will unfold population dynamics during the different sequences by reconstructing the neural trajectories in a 3-dimensional space by multidimensional scaling.

5.1 Single-trial analysis with optimized single neuron representations

We can examine how units represent task events and sequence information by comparing the time courses of their neuronal activity across trials. To enable comparisons of spiking activity across trials varying in length (when animals were freely moving), we aligned KDEs at time-stamps of task-related behavior by "time-warping" or temporal scaling (for a detailed description see sec. 5.4).

Representation of sequence information by single units: Figure 5.1 illustrates the transformation from the original (left) to time-warped (middle and right) single-trial KDEs. Across-trial comparisons of time-warped optimized KDEs in figure 5.2 confirm that single units significantly differentiate between distinct events in time courses of their activity. We then examined whether the firing of single cells can represent combinations of independent task-related attributes encoding for both actions and context information. Therefore, we grouped single-trials according to the task stage or order of performed actions.

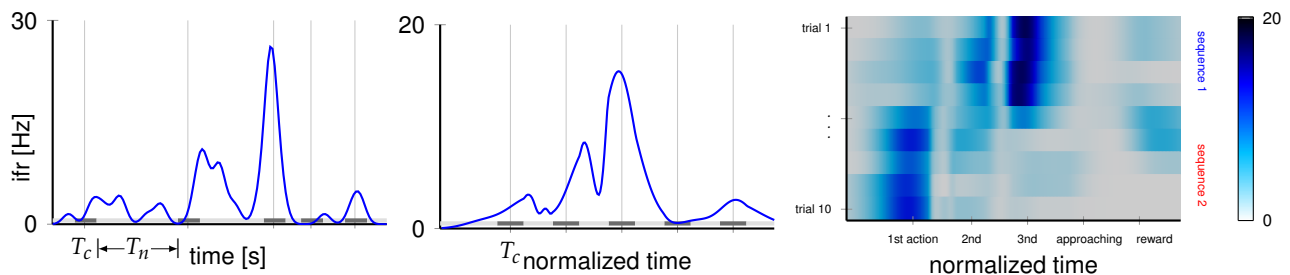


Figure 5.1: Illustration of single-trial KDEs varying in length and time-warped KDEs. **Left:** single unit optimized KDE estimated at equidistant time points. Event time-stamps are denoted by vertical lines. **Middle:** same single trial spiking activity as 'time-warped' representation aligned at task-related behaviors. Gray-colored sections on the x-axis indicate response (T_c , dark gray) and non-response-specific time-windows (T_n , light gray). Sections of the same color in left span the same time periods as in the middle plot. **Right:** contour plot of ten 'time-warped' single-trial KDEs for the two different task stages: actions performed in clock-wise (sequence 1, trials 1-5) and anti-clock-wise order (sequence 2, trials 6-10). Dark blue regions indicate high firing activity and light blue regions activity close to zero Hz.

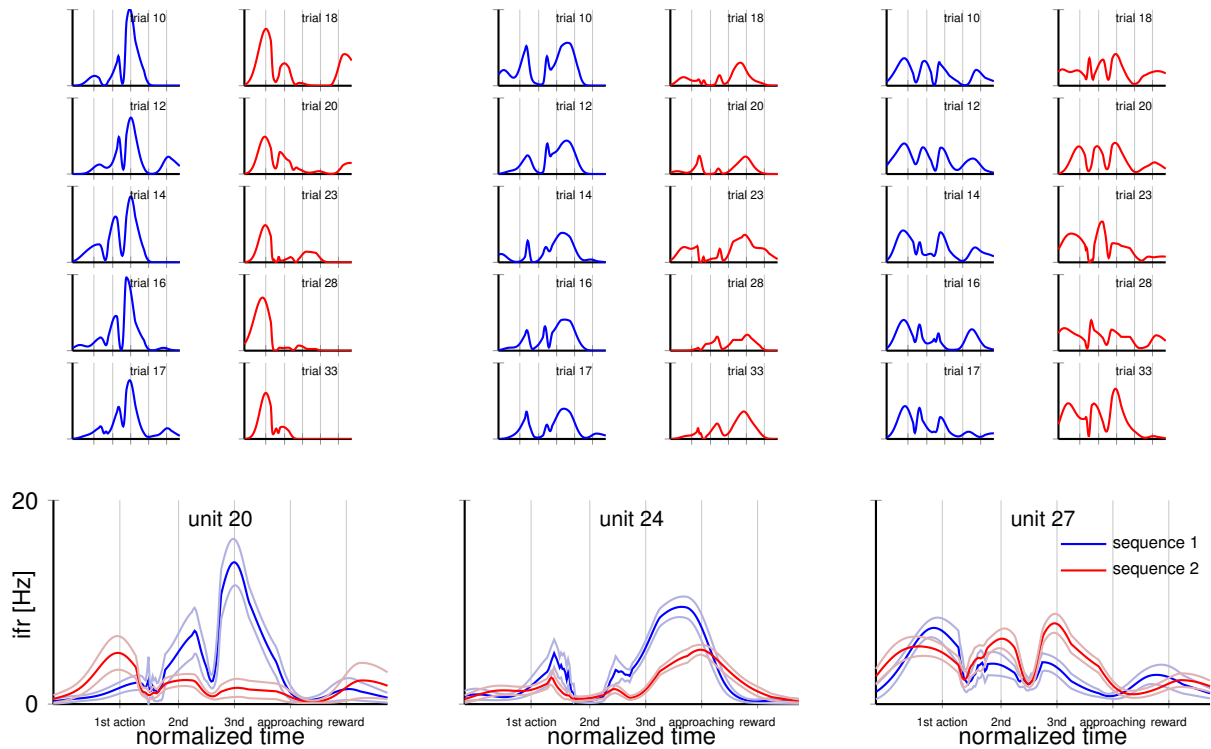


Figure 5.2: Time courses of time-warped optimized single trial instantaneous firing rates of single units which convey maximal information about task-related responses, grouped according to the task-stages. Data set 12 taken from the sequence-switch task. Upper panels: Ten single-trial ifrs of three selected units during the two different task stages (color-coded sequences one and two). Lower panel: trial-averaged activity across the two sequences, light blue and red lines indicate 95% confidence regions. The left column displays the same busy, stimulus-locked unit as shown in chap. 3, figure 3.5, middle panel. Time-warped single trial KDEs are aligned at time-stamps corresponding to actions performed on the task denoted by vertical lines (level-press, nose-poke, wheel-turn, approaching the reward and reward consumption).

Single-trial optimized KDEs of single units in figure 5.2, lower panel reveal that some cells exhibit temporal activity patterns that clearly differentiate between sequences, even when actions within a sequence are re-ordered to match both sequences. This finding implies that information on actions and on sequence order is encoded in some of the single unit activity.

5.2 Identifying cell ensembles encoding for task-related information

In order to make inferences about population coding mechanisms we examined if mPFC neurons follow rather a distributed or sparse coding scheme by identifying units encoding for distinct task-related events and determining the size of these neuronal ensembles proportional to the recorded population.

When coding is fully distributed information will be conveyed by the activity of many cells and accordingly including all units to build the classifier will yield a higher decoding performance, otherwise, when encoding is sparse, classification accuracy will be improved when activity of only few predictive cells is considered.

We can address the problem from a statistical analytical view-point by using a procedure termed feature or model selection. In this context, each feature corresponds to the optimized kernel density or instantaneous firing rate estimate v_i of a single unit, $i=1, \dots, p$ and a model represents the set of features the classifier is built on. Units which most improve the misclassification error are successively added or removed from the model (for a more detailed description see appendix 5.4).

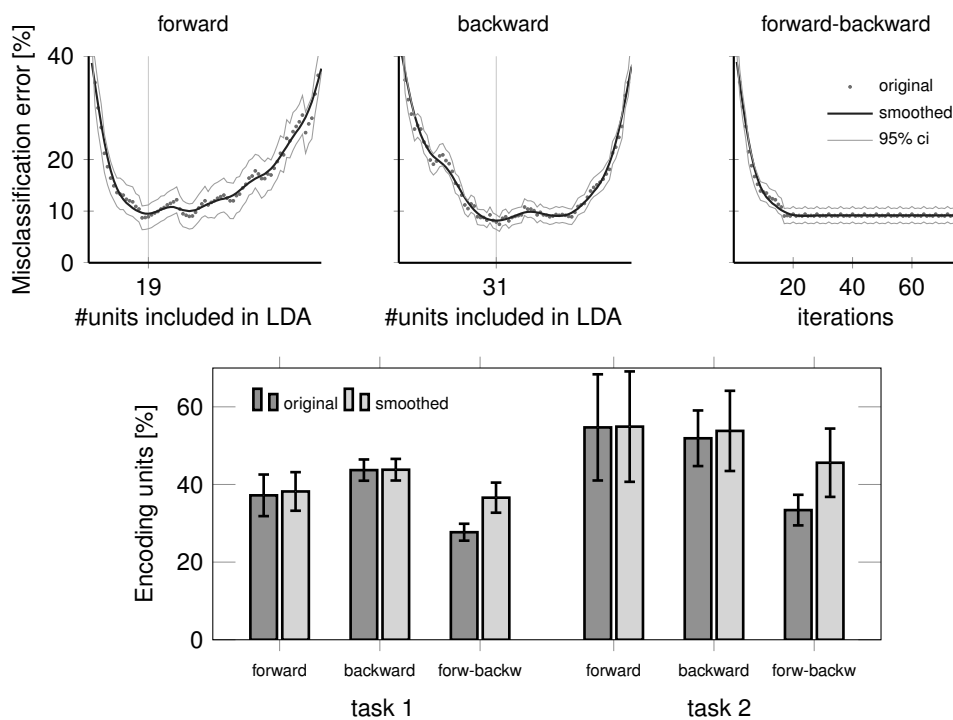


Figure 5.3: Identifying encoding units via stepwise feature selection procedures. Top: decoding performance as a function of the subset size used for LDA classifier construction, dataset 12, sequence-switch task. The optimal feature set size is indicated by vertical lines. Lower panel: average size of neuronal ensembles conveying maximal task-related information proportional to the recorded population identified by the three different feature selection algorithms (mean value \pm 95% confidence intervals).

Impact of the ensemble size on the decoding accuracy: Figure 5.3, top panel illustrates misclassification error curves as a function of the subset size of neurons which was employed to construct the LDA classifier for three different variable selection methods. Results obtained with forward-backward selection are not considered for further analyses due to lack of convergence to a global minimum. All prediction error curves are typically u-shaped due to the bias-variance trade-off. Starting with one unit the misclassification rate decreases as the activity of more units are used for building the classifier until a global minimum is reached, which defines the optimal size of the most predictive neuronal ensemble, indicated by gray horizontal lines in figure 5.3, top panel. Prediction accuracy then again deteriorates as more units are added to the model. Decoding accuracy significantly improves when taking into account exclusively the activity of the most predictive optimized KDEs of cells which are identified by feature selection. For instance, in backward selection the misclassification error is reduced from about $41 \pm 2.5\%$, when including the entire set of recorded units, to below $7.9 \pm 1.43\%$ for the most predictive ensemble. Furthermore, from figure 5.3, top middle column it follows that the complete set has almost the same decoding performance as the best single unit $46.9 \pm 2.6\%$. This means that discrimination of task-related event can significantly be improved with by selection the subset of encoding cells.

Sizes of encoding neuronal ensembles: The application of optimized KDEs in combination with feature selection suggests that around 40% of the cells are encoding for task-related events on the sequence-switch task (task one: 38.2 ± 7.5 , forward and 43.8 ± 3.2 backward selection) and above 50% in the alternation task (task two: 54.9 ± 13.0 , forward and 53.8 ± 10.6 backward selection), figure 5.3, lower panel. On average more cells were recorded per session in task one (~ 50) compared to task two, (~ 40). Generally, most predictive feature sets identified by backward selection slightly differ from those found by forward selection, but not to a significant extent.

5.3 Reconstructing population dynamics during sequence processing

By analyzing high-dimensional population dynamics and low-dimensional neural trajectories in the PFC it has already been possible to uncover interesting phenomena during rule learning, context perception or decision making either on a single-trial or on a multiple-trial basis (Mante et al., 2013; Stokes et al., 2013; Durstewitz et al., 2010). Here, we show first, that also optimized KDEs enable to extract low-dimensional single-trial time courses which allow to monitor the simultaneous activity of a neuronal population and second, which conclusions on mechanisms of sequence encoding can be inferred from trajectories based on optimized KDEs.

The idea behind this approach is that each neuron is considered as a noisy sensor which reflects an underlying neural process (Brown et al., 2005; Carrillo-Reid et al., 2008; Mazor and Laurent, 2005; Yuan and Niranjana, 2010; Stopfer et al., 2003; Yu et al., 2006; Balaguer-Ballester et al., 2011). This process can be uncovered by extracting a low-dimensional neural trajectory from the recorded high-dimensional population activity. In behavioral experiments high trial-by-trial variability of single neurons can be related to internal processing, e.g. in behavioral tasks involving motor planning, decision making, rule learning and perception (Nawrot et al., 1999; Horwitz and Newsome, 2001; Czanner et al., 2008; Mante et al., 2013; Churchland et al., 2010) and will reflect the state of the network and thus will be shared among many cells. The neural trajectory facilitates the visualization of the underlying neuronal process by providing a reduced representation of the shared high-dimensional population activity.

We applied multidimensional scaling (MDS) to reduce the dimensionality of the population vector of optimized instantaneous firing rates. Figure 5.4, middle and top panel shows the trajectories obtained with optimized and non-optimized KDEs for an exemplary data set from the sequence-switch task. Each dot represents the state of the entire recorded ensemble in one 100-ms bin. All points corresponding to different 100-ms bins in the epochs of the same behavior are shown in the same color.

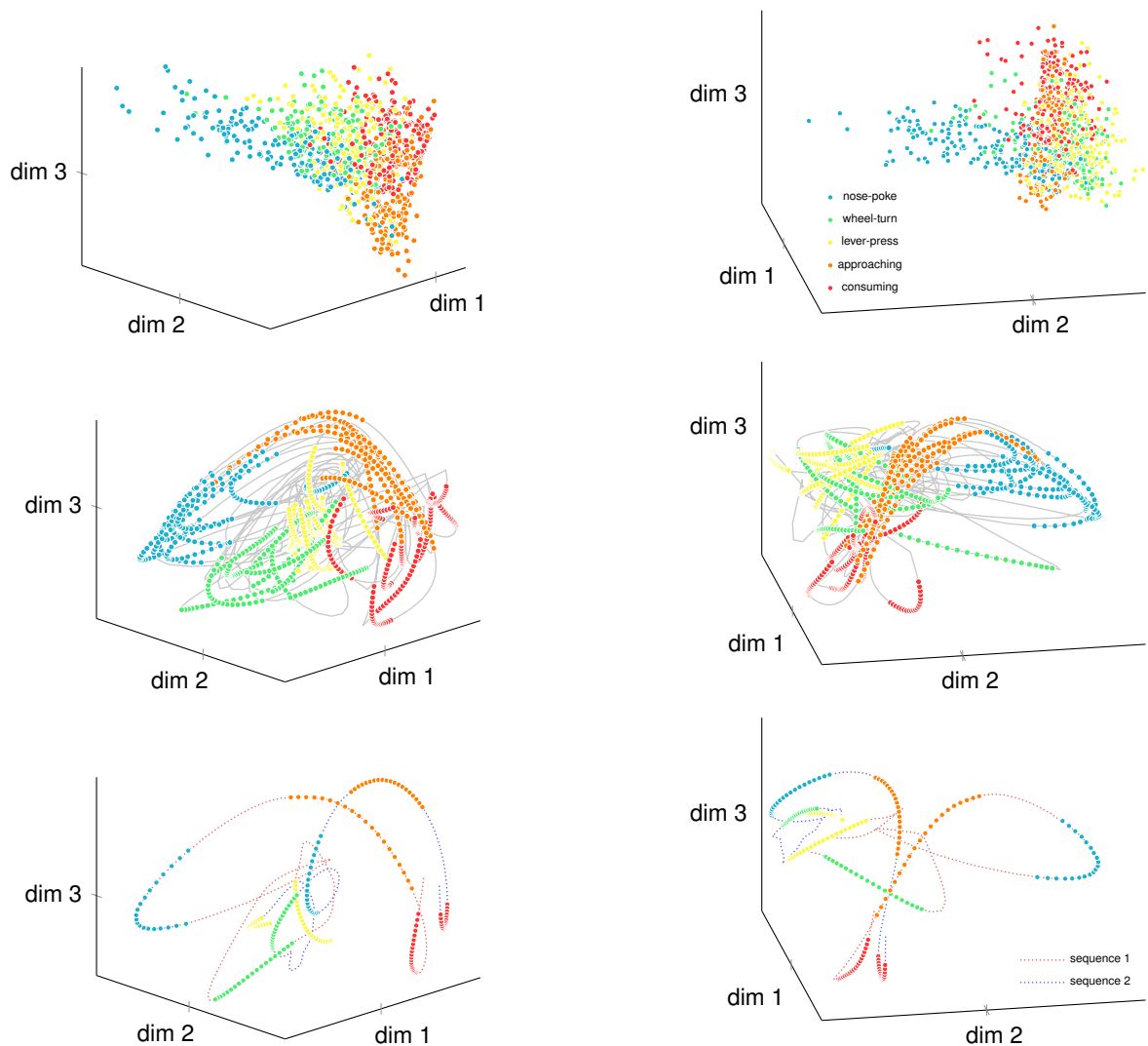


Figure 5.4: Reconstructing population dynamics from optimal KDEs. Data set 12 taken from the sequence-switch task. Population vector of optimized instantaneous firing rates mapped to a 3-dimensional space with multi-dimensional scaling. **a-b:** neuronal single trial trajectories based on non-optimized KDEs. **c-d:** neuronal single trial trajectories based on optimized KDEs, both showing 10 consecutive trials. **e-f:** averaged trajectories over the two different task stages: sequence 1 and sequence 2.

Representation of sequence information by the neuronal population: Our findings demonstrate that first, optimized KDEs with automated bandwidth selection enable to unfold population dynamics and show smooth single-trial trajectories (fig.5.4, middle panel). This is clearly not possible with non-optimized KDEs where the bandwidth is chosen at random (fig.5.4, top panel). Second, when averaging over the sequences in the reduced space, task events which appear in the same order in both sequences (reward-approaching, reward-consumption) occupy similar regions of the space. Instead, same actions but with reserved order can be discriminated along some dimensions of the reduced space while aligned along other dimensions (fig.5.4, lower panel). The conclusions on mechanisms of sequence encoding we can draw from analyzing trajectories based on optimized KDEs are that information encoding for both actions as well as sequence order (context) seems to be represented simultaneously in the population activity. These findings support the view that the mPFC is involved in higher-level contextual encoding (Hyman et al., 2012; Ma et al., 2014; Rigotti et al., 2013).

5.4 Concluding remarks

To date, studies which seek to determine more accurately the underlying spike density function by bandwidth optimization of KDEs are mostly employed to analyze representations of neuronal activity in form of instantaneous firing rates (Nawrot et al., 1999; Lehky, 2010; Shimazaki and Shinomoto, 2010).

However, in this study we have seen that optimized bandwidths provide an information rich measure which helps to address and study more challenging questions and can provide valuable insight into the understanding of multiple different mechanisms: By examining optimized bandwidths and their distributions, we can discriminate cells with specific temporal spiking structure and moreover, segregate cell groups with distinct spiking characteristics. By employing optimized KDEs obtained with automated bandwidth selection we can enhance classification performance and unfold population dynamics reflecting the underlying neuronal process and thereby draw conclusions on encoding mechanisms.

Theoretical research on kernel density estimation is far ahead of its practical application to neurophysiological data for a long time: adaptive KDEs with variable bandwidths which adjust to non-stationary regions (Sain, 1994; Hazelton, 2003; Terrell and Scott, 1992; Tran, 2010; Antoniadis et al., 2009) could be employed to overcome trial-by-trial variability in spike trains. Also, multivariate KDEs (Sain, 2002; Bouezmarni and Rombouts, 2010; Zougab et al., 2014) might provide a more accurate method to estimate the underlying spike density function of multiple-single unit recordings. Yet, the PSTH is widely applied and the most commonly used method in in-vivo neurophysiology for the representation of spiking activity.

This poses the need for future studies to integrate these already present more advanced methods on automated bandwidth selection and adaptive and multivariate KDE and will be crucial for a better understanding of functional organization of neuronal activity. Also bandwidth estimates obtained by optimization should be studied more systematically in a variety of cortical areas. We hope that the application of bandwidth estimates and KDEs determined by optimization will in future shed light on the neuronal coding problem and contribute to a better understanding of the functional mechanisms of internal processing during complex cognitive tasks in higher order areas as the PFC.

I would like to conclude with the remark that I hope that this work will fill at least a tiny piece of the gap between the field of Kernel density estimation and the complex but fascinating world of how the brain works which may inspire and enrich future research.

Acknowledgements

I would like to thank all my colleagues and friends who contributed with valuable and thoughtful discussions to the work described in this thesis.

First and foremost I offer my sincerest gratitude to my supervisor, Professor Daniel Durstewitz, supporting me throughout my thesis with his patience and knowledge whilst allowing me the room to work in my own way and demanding a high quality of work in all my endeavours.

Additionally, I would like to thank my dissertation committee members Professors Ursula Kummer, Christoph Schuster, and Dr. Kevin Allen for their interest in my work and their important feedback.

I am very grateful to Liya Ma, James Hyman, and Jeremy Seamans for providing the vivo data sets used to test our methods and analysis.

Every result described in this thesis was accomplished with the help and support of fellow labmates and collaborators at the Central Institute for Mental Health, Mannheim and the Bernstein Center for Computational Neuroscience, Heidelberg and Mannheim. In my daily work I have been blessed over the years with a friendly and cheerful group of fellow students. Loreen Hertaeg, Charmaine Demanuele, Emili Balaguer-Ballester, Grant Sutcliffe, Claudio Sebastian Quiroga-Lombard, Tatiana Golovko, Joachim Hass, Thomas Fucke, Sven Berberich, Thomas Hahn, Elonora Russo, Georgia Koppe, Carla Filosa, Sadjad Sadeghi, and Hazem Toutounji made my time here at the Central Institute a lot more fun.

Finally, I would like to acknowledge friends and family who supported me during my time here. First and foremost I would like to thank Mom, Dad, Igor, Jessica, Oliver for their constant love and support.

Appendix

Methods

Supplemental Information on the electrophysiological recordings & sequence-switch task

The following description was provided by Dr Liya Ma (personal communication).

Apparatus. The maze consisted of 4 platforms connected by 4 two-foot long passages that formed a diamond shape (fig. 5.5), with platforms 1 and 3 at the sharp tips of the diamond. The individual platforms differed in size, shape, floor texture and wall patterns. Platforms 1 to 3 contained unique manipulandae: a nose-poke port in platform 1 to the right; a lever in platform 2 in the middle; and a response wheel in platform 3 to the left. A 3W signal light was located above each of these manipulandae. During initial task shaping, a food-cup was inserted above each manipulanda for food-pellet delivery, which was always accompanied by a 0.5s pure tone at 1.5 KHz. Once trained, these food cup dispensers were removed and food was only delivered on platform 4 in association with the same tone. All cue lights, tone-generator, manipulandae and pellet dispensers were operated by a MedPC IV system (Med Associates, Georgia, VT). Doors located at the start of each passage could be controlled from outside of the maze.

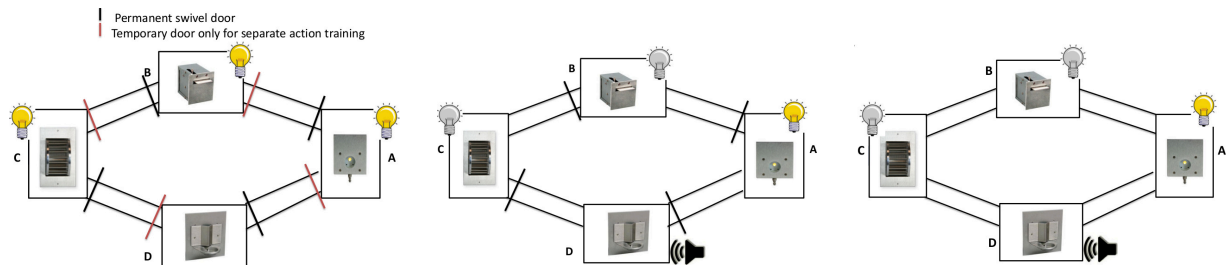


Figure 5.5: Sequence switch task apparatus and stages of pre-training sessions on the maze. Figure and description taken from Ma (personal communication).

Behavior. Pre-training on the Maze: Training on the maze started with single-action instrumental conditioning. The animals were restricted for fixed periods to each of the nose poke (np)(right), lever press (lp) (middle) and wheel turn (wt) (left) platforms, thus ensuring daily training on each of the 3 instrumental actions (fig. 5.5). They progressed through FR1, fixed-interval 10s (FI10s) and random-interval 15s (RI15s) schedules with a performance criterion of 30pellets/20min. Ten daily sessions of sequence shaping started the day after they attained criterion performance on all three individual operant actions. The animals were placed in the reward platform at the beginning of the session and the only open passage led to Platform 1, where the light above the np port was illuminated (fig. 5.5 middle). Once the animals reached this platform, two doors blocked both exits and the light stayed on until a np response was emitted. At that point a door opened to allow access to Platform 2 and the light above the lever was illuminated. The animals could only leave the Platform 2 after they performed a lp. At that point the light above the lever was extinguished and the door was opened and the light above the wheel on Platform 3 was illuminated. Once the animal reached the wt platform and turned on the wheel for a full circle, the light above the wheel was extinguished and the last door opened to allow entrance to Platform where 4 food pellets were delivered accompanied by a 0.5-s pure tone at 1.5 KHz. After 4 seconds, another door opened so they could move to Platform 1 and start the next trial. These sessions lasted for no less than 45min and ended either when the animal stopped responding for at least 3min or 60min had elapsed.

Maze sequence task: After 10 daily shaping sessions, all doors were removed. Rats were still required to perform the 3 actions in the aforementioned order, but could commit out-of-sequence errors such as running in the wrong direction around the maze. Performance was evaluated by the number of out-of-sequence errors per trial. Since the intention was to examine the plasticity of action representations rather than sequence learning per se, it was desirable to minimize these types of errors. Therefore the rats continued to be guided through the sequence by cue lights that were activated by the manipulanda and that illuminated the next platform in the sequence. Repeated responses on the manipulanda in the 4s after the initial correct response were not considered as errors. A session typically lasted for 60min, but extra time (up to 10min) was occasionally given if the animal was in the middle of the sequence at 60min. The animals continued to receive this self-paced sequence training for a total of 23 days.

Sequence switch task: After training on the original sequence (sequence A), animals were trained to perform the action-sequence in the reversed order (i.e., wt lp np reward: sequence B). On the maze, rats were first guided through sequence B with doors, then without, until they reached the same level of efficiency as exhibited previously on sequence A. At that point the sequence-switch sessions commenced on the following days, when the animals were required to complete at least 20 trials on sequence B within 20min at which point they were removed from the maze for a minute. They were then placed back in the apparatus and the light above the np was illuminated instead of the light above the wt (as was the case for sequence B). This prompted them to perform sequence A. Only the trials free of out-of-sequence errors were used in the analyses of neural data. In the results, we refer to the first sequence (i.e. sequence B) as "sequence 1" and the 2nd as "sequence 2".

Subjects & surgery. Male Long-Evans rats (450-550g) were housed in a facility with 12hr light-dark cycle, with all training and recording taking place during the light cycle. For the duration of the behavioral experiments, the rats were food-restricted to just below 90% of their free-feeding weights. Feeding took place in the home cage after their daily training/recording sessions, and water was available ad libitum

in the cages at all times. All procedures were carried out in accordance with the Canadian Council of Animal Care and the Animal Care Committee at the University of British Columbia. Stereotaxic surgeries were performed on naive rats with sterilized-tip procedures. NSAIDs analgesic, antibiotic, and a local anesthetic, were given before incision. An elliptical-shaped craniotomy was made, centered at: AP: +3.2mm, ML: ± 0.5 mm. Once the dura mater was retracted, the bottoms of the two bundles of 8, 30-gauge tubes, containing a total of 16 tetrodes were placed bilaterally immediately beside the central sinus, touching the cortical surface. Each bundle had a cylindrical shape with bottom radius ~ 0.4 mm, and were angled by $3.5\sim 5$ degrees. The implants were fixed with bone screws and dental acrylic. All tetrodes were extended ~ 0.7 mm into the brain at the end of the surgery. After 10d of recovery, the tetrodes were advanced ventrally into the ACC. Once all tetrodes were placed into the dorsal ACC according to lowering records and atlas coordinates, small adjustments were made with hyperdrives to maximize the number of neurons recorded.

Acquisition of electrophysiological data. For data acquisition, an EIB-36TT board (Neuralynx Inc., Bozeman, MT, USA), connected to the extracellular electrodes were plugged into HS-36 headstages and tether cables (Neuralynx Inc., Bozeman, MT). Signals were converted by a Digital Lynx 64 channel system (Neuralynx Inc., Bozeman, MT) and sent to a PC workstation, where electrophysiological and behavioral data were read into Cheetah 5.0 software (Neuralynx Inc., Bozeman, MT). Files were then read into Offline Sorter (Plexon Inc., Dallas, TX) for spike sorting, based on visually dissociable clusters in 3D projections along multiple axes for each electrode of a tetrode (peak and valley amplitudes, peak-to-valley ratio, principal components and area). Sorting was confirmed by examining auto- and cross-correlations, and ANOVAs were conducted from the 2D and 3D projections. Spike timestamps were then read into MATLAB (Mathworks Inc., Natick, MA) for all further analysis. At the end of the studies, the animals were deeply anesthetized using urethane i.p. injection, and a $100\mu\text{A}$ current was passed through the electrodes for 30s. Animals were then perfused with a solution containing 250ml 10% buffered formalin, 10ml glacial acetic acid, and 10 g of potassium ferrocyanide. This solution causes a Prussian blue reaction, which marks with blue the location of the iron particles deposited by passing current through the electrodes. The brains were then removed and stored in a 10% buffered formalin/20% sucrose solution for at least 1 week, before being sliced and mounted to determine precise electrode tracks. Since multiple sessions were recorded from individual animals the precise recording locations could not be derived from electrode lesions, but all electrode tracks were inferred between the entrance point and the dyed spot. All tracks ended within the medial frontal cortex with the vast majority of tracks limited to the ACC and a minority extending into superficial layers of the prelimbic region.

Single-trial time warping

In order to align single trial KDEs, instead of using a grid of usually equidistant points at which the instantaneous firing rate has to be estimated (fig. 5.1, left), we divided each trial into an equal number of non-overlapping scaled time bins, which length was adapted to the timing of task events (fig. 5.1, middle and right panel). We obtained the temporal binning for each trial in the following way:

Based on recorded time-stamps corresponding to actions c_k , $k = 1, \dots, 5$, performed on the task (level-press, nose-poke, wheel-turn, approaching the reward and reward consumption, see sec. 5.4), we identified periods of 500 ms preceding and following all task-related behaviors as time windows T_c of response-specific activity (fig. 5.1, dark gray horizontal lines). Time intervals T_n lying in-between were labeled as non-response-specific (fig. 5.1, light gray horizontal lines). Then, the temporal resolution of KDEs, i.e. grid at which the instantaneous firing rate was estimated, was adjusted depending on the window type. Response-specific time-windows of length $T_c=1$ s were divided into $L=20$ equally-spaced bins, so that the resulting response-specific instantaneous firing rates $\nu(t|c_k)$ had a temporal resolution of $T_c/L = \Delta t = 50$ ms. Non-response-specific periods T_n were partitioned into the same number of bins respectively. However, as the timing of self-paced behavior varied, the temporal resolution was adjusted to the length of time intervals between consecutive responses $\Delta t = T_n/L$.

Feature selection

Before feature selection we partition the input and output dataset of optimized KDEs ν and stimulus-vectors y into two subsets based on the leave-one-trial-out CV principle (sec 2.5). The training set is used for classifier construction and the test set is retained for subsequent validation.

We employed LDA in combination with forward, backward and forward-backward feature selection algorithms, which are best described by the following pseudo-code:

<ol style="list-style-type: none"> 1. Initialize feature set <ul style="list-style-type: none"> · $S_0 = \langle \emptyset \rangle; m = 0$ 2. Select the next best feature <ul style="list-style-type: none"> · $\nu^* = \underset{\nu_i \in S_m}{\arg \min} \text{Err}(S_m \cup \nu_i)$ 3. Update feature set <ul style="list-style-type: none"> · $S_{m+1} = S_m \cup \nu^*$ 4. While $m < p$ <ul style="list-style-type: none"> · $m = m + 1$ · Go to Step 2 <p>Sequential forward selection pseudo-code.</p>	<ol style="list-style-type: none"> 1. Initialize feature set <ul style="list-style-type: none"> · $S_0 = \nu; m = 0$ 2. Remove the worst feature <ul style="list-style-type: none"> · $\nu^* = \underset{\nu_i \in S_m}{\arg \min} \text{Err}(S_m \setminus \nu_i)$ 3. Update feature set <ul style="list-style-type: none"> · $S_{m+1} = S_m \setminus \nu^*$ 4. While $m > p$ <ul style="list-style-type: none"> · $m = m + 1$ · Go to Step 2 <p>Sequential backward selection pseudo-code.</p>
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

The backward feature selection procedure starts with the full set of features or single neuron firing rate estimates $S_p = \nu = \{\nu_1, \dots, \nu_p\}$ and then sequentially eliminates features $\nu_i \in \nu$ which most improve the misclassification error $\text{Err}(S_{p-1})$, $S_{p-1} = S_p \setminus \nu_i$, while in forward selection units are added to the current set. Forward-backward selection combines both procedures so that units are added or deleted from the subset S_m as long as the prediction error is still decreasing $\text{Err}(S_{m+1}) \leq \text{Err}(S_m)$.

We computed the prediction error Err by means of leave-one-trial-out CV (see sec. 2.5). After smoothing the resulting $Err(S_m)$ -functions with the MATLAB built-in cubic spline function "csaps", the most predictive subset of neurons which conveys maximal stimulus-related information is then identified as the set for which the prediction error attains its global minimum $S_{opt} = \underset{m}{arg\ min} Err(S_m)$.

Bibliography

- Abarbanel, H. D., Huerta, R., Rabinovich, M. I., Rulkov, N. F., Rowat, P. F., and Selverston, I. (1996). Synchronized action of synaptically coupled chaotic model neurons. *Neural computation*, 8:1567–1602.
- Abeles, M., Bergman, H., Gat, I., Meilijson, I., Seidemann, E., Tishby, N., and Vaadia, E. (1995). Cortical activity flips among quasi-stationary states. *Proceedings of the National Academy of Sciences of the United States of America*, 92(September):8616–8620.
- Adrian, E. D. and Zotterman, Y. (1926). The impulses produced by sensory nerve-endings: Part II. The response of a Single End-Organ. *The Journal of physiology*, 61(2):151–171.
- Ainsworth, M., Lee, S., Cunningham, M. O., Traub, R. D., Kopell, N. J., and Whittington, M. A. (2012). Rates and Rhythms: A Synergistic View of Frequency and Temporal Coding in Neuronal Networks.
- Amit, D. J. (1992). *Modeling Brain Function: The World of Attractor Neural Networks*. Addison-Wesley.
- Antoniadis, A., Paparoditis, E., and Sapatinas, T. (2009). Bandwidth selection for functional time series prediction. *Statistics and Probability Letters*, 79(6):733–740.
- Ardid, S., Vinck, M., Kaping, D., Marquez, S., Everling, S., and Womelsdorf, T. (2015). Mapping of Functionally Characterized Cell Classes onto Canonical Circuit Operations in Primate Prefrontal Cortex. *Journal of Neuroscience*, 35(7):2975–2991.
- Baeg, E. H., Kim, Y. B., Kim, J., Ghim, J.-W., Kim, J. J., and Jung, M. W. (2007). Learning-induced enduring changes in functional connectivity among prefrontal cortical neurons. *The Journal of neuroscience*, 27(4):909–918.
- Bair, W. and Koch, C. (1996). Temporal precision of spike trains in extrastriate cortex of the behaving macaque monkey. *Neural computation*, 8:1185–1202.
- Balaguer-Ballester, E., Lapish, C. C., Seamans, J. K., and Durstewitz, D. (2011). Attracting dynamics of frontal cortex ensembles during memory-guided decision-making. *PLoS Comput. Biol.*, 7(5):e1002057.
- Balaguer-Ballester, E., Tabas-Diaz, A., and Budka, M. (2014). Can we identify non-stationary dynamics of trial-to-trial variability? *PLoS ONE*, 9(4).
- Barbieri, R. (2001). Construction and analysis of non-Poisson stimulus-response models of neural spiking activity. *Journal of Neuroscience Methods*, 105(1):25–37.

- Barlow, H. B. (1972). Single units and sensation: a neuron doctrine for perceptual psychology? *Perception*, 1(4):371–394.
- Bialek, W., Rieke, F., de Ruyter van Steveninck, R. R., and Warland, D. (1991). Reading a neural code. *Science*, 252(5014):1854–1857.
- Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*, volume 4. Springer.
- Blanche, T. J., Spacek, M. A., Hetke, J. F., Swindale, N. V., Timothy, J., Spacek, M. A., Hetke, J. F., and Swindale, N. V. (2005). Polytrodes : High-Density Silicon Electrode Arrays for Large-Scale Multiunit Recording. *Journal of Neurophysiology*, 93:2987–3000.
- Bouezmarni, T. and Rombouts, J. V. (2010). Nonparametric density estimation for positive time series. *Computational Statistics & Data Analysis*, 54(2):245–261.
- Bowman, A. W. (1984). An alternative method of cross-validation for the smoothing of density estimates. *Biometrika*, 71(2):353–360.
- Bowman, A. W. and Azzalini, A. (1997). *Applied Smoothing Techniques for Data Analysis: The Kernel Approach with S-Plus Illustrations*. Oxford Statistical Science Series. Oxford University Press, USA.
- Brown, E. N., Kass, R. E., and Mitra, P. P. (2004). Multiple neural spike train data analysis: state-of-the-art and future challenges. *Nature Neuroscience*, 7(5):456–461.
- Brown, S. L., Joseph, J., and Stopfer, M. (2005). Encoding a temporally structured stimulus with a temporally structured neural representation. *Nature Neuroscience*, 8(11):1568–1576.
- Buračas, G. T., Zador, A. M., DeWeese, M. R., and Albright, T. D. (1998). Efficient discrimination of temporal patterns by motion-sensitive neurons in primate visual cortex. *Neuron*, 20:959–969.
- Carrillo-Reid, L., Tecuapetla, F., Tapia, D., Hernández-Cruz, A., Galarraga, E., Drucker-Colin, R., and Bargas, J. (2008). Encoding network states by striatal cell assemblies. *Journal of neurophysiology*, 99(January 2008):1435–1450.
- Churchland, M. M., Yu, B. M., Cunningham, J. P., Sugrue, L. P., Cohen, M. R., Corrado, G. S., Newsome, W. T., Clark, A. M., Hosseini, P., Scott, B. B., Bradley, D. C., Smith, M. A., Kohn, A., Movshon, J. A., Armstrong, K. M., Moore, T., Chang, S. W., Snyder, L. H., Lisberger, S. G., Priebe, N. J., Finn, I. M., Ferster, D., Ryu, S. I., Santhanam, G., Sahani, M., and Shenoy, K. V. (2010). Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nature neuroscience*, 13(3):369–378.
- Churchland, M. M., Yu, B. M., Sahani, M., and Shenoy, K. V. (2007). Techniques for extracting single-trial activity patterns from large-scale neural recordings. *Current opinion in neurobiology*, 17(5):609–618.
- Cox, D. R. and Isham, V. (1980). *Point Processes*. CRC Monographs on Statistics & Applied Probability. Chapman & Hall/CRC.
- Cox, D. R. and Wermuth, N. (2002). On some models for multivariate binary variables parallel in complexity with the multivariate Gaussian distribution. *Biometrika*, 89(2):462–469.

- Cunningham, J. P., Gilja, V., Ryu, S. I., and Shenoy, K. V. (2009). Methods for estimating neural firing rates, and their application to brain-machine interfaces. *Neural Networks*, 22(9):1235–1246.
- Cunningham, J. P., Shenoy, K. V., and Sahani, M. (2008). Fast Gaussian process methods for point process intensity estimation. In *Proceedings of the 25th international conference on Machine learning, ICML '08*, pages 192–199, New York, NY, USA. ACM.
- Czanner, G., Eden, U. T., Wirth, S., Yanike, M., Suzuki, W. A., and Brown, E. N. (2008). Analysis of between-trial and within-trial neural spiking dynamics. *Journal of neurophysiology*, 99(5):2672–2693.
- Daley, D. J. and Vere-Jones, D. (2003). *An Introduction to the Theory of Point Processes, Volume 1 (2nd ed.)*. Springer, New York.
- Dayan, P. and Abbott, L. F. (2005). *Theoretical neuroscience: computational and mathematical modeling of neural systems*. MIT Press.
- DeCharms, R. C. and Zador, A. (2000). Neural Representation and the Cortical Code. *Annual Review of Neuroscience*, 23:613–647.
- Dimatteo, I., Genovese, C. R., and Kass, R. E. (2001). Bayesian curve-fitting with free-knot splines. *Biometrika*, 88(4):1055–1071.
- Dodla, R. and Wilson, C. J. (2010). Quantification of clustering in joint interspike interval scattergrams of spike trains. *Biophysical Journal*, 98(11):2535–2543.
- Durstewitz, D., Vittoz, N. M., Floresco, S. B., and Seamans, J. K. (2010). Abrupt transitions between prefrontal neural ensemble states accompany behavioral transitions during rule learning. *Neuron*, 66(3):438–448.
- Efron, B. (2004). The Estimation of Prediction Error. *Journal of the American Statistical Association*, 99(467):619–632.
- Endres, D., Oram, M., Schindelin, J., and Foldiak, P. (2008). Bayesian binning beats approximate alternatives: estimating peri-stimulus time histograms. In Platt, J. C., Koller, D., Singer, Y., and Roweis, S., editors, *Advances in Neural Information Processing Systems 20*. MIT Press.
- Escola, S., Fontanini, A., Katz, D., and Paninski, L. (2011). Hidden markov models for the stimulus-response relationships of multistate neural systems. *Neural computation*, 23(2006):1071–1132.
- Faure, P., Kaplan, D., and Korn, H. (2000). Synaptic efficacy and the transmission of complex firing patterns between neurons. *Journal of neurophysiology*, 84(6):3010–3025.
- Fellous, J.-M., Tiesinga, P. H. E., Thomas, P. J., and Sejnowski, T. J. (2004). Discovering spike patterns in neuronal responses. *The Journal of neuroscience*, 24(12):2989–3001.
- Feng, J. (2003). *Computational Neuroscience: A Comprehensive Approach*. CRC Press.

- Fitzurka, M. a. and Tam, D. C. (1999). A joint interspike interval difference stochastic spike train analysis: detecting local trends in the temporal firing patterns of single neurons. *Biological cybernetics*, 80:309–326.
- Fujii, H., Ito, H., Aihara, K., Ichinose, N., and Tsukada, M. (1996). Dynamical cell assembly hypothesis - Theoretical possibility of spatio-temporal coding in the cortex. *Neural Networks*, 9(8):1303–1350.
- Georgopoulos, A. P., Kettner, R. E., and Schwartz, A. B. (1986). Neuronal population coding of movement direction. *Science*, 233:1416–1419.
- Gerstein, G. L., Bedenbaugh, P., and Aertsen, A. M. H. J. (1989). Neuronal assemblies. *IEEE Transactions on Biomedical Engineering*, 36(1):4–14.
- Gerstein, G. L. and Kiang, N. Y. (1960). An Approach to the Quantitative Analysis of Electrophysiological Data from Single Neurons. *Biophys. J.*, 1(1):15–28.
- Gerstner, W. and Kistler, W. K. (2002). *Spiking Neuron Models*. Cambridge University Press.
- Harris, K. D. (2005). Neural signatures of cell assembly organization. *Nature reviews. Neuroscience*, 6(May):399–407.
- Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning*, volume 27. Springer.
- Hazelton, M. L. (2003). Variable kernel density estimation. *Australian & New Zealand Journal of Statistics*, 45(3):271–284.
- Hebb, D. O. (1949). *The Organization of Behavior: A Neuropsychological Theory*. Wiley, New York, new ed edition.
- Holt, G. R., Softky, W. R., Koch, C., and Douglas, R. J. (1996). Comparison of discharge variability in vitro and in vivo in cat visual cortex neurons. *Journal of neurophysiology*, 75(5):1806–1814.
- Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences of the United States of America*, 79(April):2554–2558.
- Hopfield, J. J. (1995). Pattern recognition computation using action potential timing for stimulus representation. *Nature*, 376:33–36.
- Horwitz, G. D. and Newsome, W. T. (2001). Target selection for saccadic eye movements: prelude activity in the superior colliculus during a direction-discrimination task. *Journal of neurophysiology*, 86(5):2543–2558.
- Hubel, D. H. and Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of physiology*, 160:106–154.
- Hyman, J. M., Ma, L., Balaguer-Ballester, E., Durstewitz, D., and Seamans, J. K. (2012). Contextual encoding by ensembles of medial prefrontal cortex neurons. *Proc. Natl. Acad. Sci. U.S.A.*, 109(13):5086–5091.

- Hyman, J. M., Whitman, J., Emberly, E., Woodward, T. S., and Seamans, J. K. (2013). Action and outcome activity state patterns in the anterior cingulate cortex. *Cerebral Cortex*, 23(June):1257–1268.
- Johnson, D. H. (1996). Point process models of single-neuron discharges. *Journal of computational neuroscience*, 3(4):275–299.
- Jones, L. M., Fontanini, A., Sadacca, B. F., Miller, P., and Katz, D. B. (2007). Natural stimuli evoke dynamic sequences of states in sensory cortical ensembles. *Proceedings of the National Academy of Sciences of the United States of America*, 104:18772–18777.
- Kass, R. E., Ventura, V., and Brown, E. N. (2005). Statistical issues in the analysis of neuronal data. *Journal of neurophysiology*, 94(1):8–25.
- Kaufman, C. G., Ventura, V., and Kass, R. E. (2005). Spline-based non-parametric regression for periodic functions and its application to directional tuning of neurons. *Statistics in Medicine*, 24(14):2255–2265.
- Kayser, C., Montemurro, M. A., Logothetis, N. K., and Panzeri, S. (2009). Spike-Phase Coding Boosts and Stabilizes Information Carried by Spatial and Temporal Spike Patterns. *Neuron*, 61:597–608.
- Krzanowski, W. J. (2000). *Principles of multivariate analysis*. Oxford University Press.
- Kumar, A., Rotter, S., and Aertsen, A. (2010). Spiking activity propagation in neuronal networks: reconciling different perspectives on neural coding. *Nature reviews. Neuroscience*, 11:615–627.
- Lagarias, J. C., Reeds, J. A., Wright, M. H., and Wright, P. E. (1998). Convergence properties of the nelder-mead simplex method in low dimensions. *SIAM Journal of Optimization*, 9:112–147.
- Lapish, C. C., Durstewitz, D., Chandler, L. J., and Seamans, J. K. (2008). Successful choice behavior is associated with distinct and coherent network states in anterior cingulate cortex. *Proc. Natl. Acad. Sci. U.S.A.*, 105(33):11963–11968.
- Lehky, S. R. (2010). Decoding Poisson spike trains by Gaussian filtering. *Neural computation*, 22(5):1245–1271.
- Loader, C. R. (1999). Bandwidth selection: classical or plug-in? *The Annals of Statistics*, 27(2):415–438.
- Ma, L., Hyman, J. M., Lindsay, A. J., Phillips, A. G., and Seamans, J. K. (2014). Differences in the emergent coding properties of cortical and striatal ensembles. *Nature neuroscience*, 17(June):1100–1106.
- Macke, J. H., Berens, P., Ecker, A. S., Tolias, A. S., and Bethge, M. (2009). Generating spike trains with specified correlation coefficients. *Neural Comput.*, 21(2):397–423.
- Mainen, Z. F. and Sejnowski, T. J. (1995). Reliability of Spike Timing in Neocortical Neurons. *Science*, 268(5216):1503–1506.
- Mante, V., Sussillo, D., Shenoy, K. V., and Newsome, W. T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature*, 503(7474):78–84.

- Mazor, O. and Laurent, G. (2005). Transient dynamics versus fixed points in odor representations by locust antennal lobe projection neurons. *Neuron*, 48:661–673.
- Mazurek, M. E. and Shadlen, M. N. (2002). Limits to the temporal fidelity of cortical spike rate signals. *Nature neuroscience*, 5:463–471.
- Nawrot, M., Aertsen, A., and Rotter, S. (1999). Single-trial estimation of neuronal firing rates: from single-neuron spike trains to population activity. *J Neurosci Methods*, 94(1):81–92.
- Nawrot, M. P., Boucsein, C., Rodriguez Molina, V., Riehle, A., Aertsen, A., and Rotter, S. (2008). Measurement of variability dynamics in cortical spike trains. *Journal of Neuroscience Methods*, 169(2):374–390.
- Nicolelis, M., Baccala, L., Lin, R. C., and Chapin, J. K. (1995). Sensorimotor encoding by synchronous neural ensemble activity at multiple levels of the somatosensory system. *Science*, 268(June):1353–1358.
- Nicolelis, M. A., Ghazanfar, A. A., Stambaugh, C. R., Oliveira, L. M., Laubach, M., Chapin, J. K., Nelson, R. J., and Kaas, J. H. (1998). Simultaneous encoding of tactile information by three primate cortical areas. *Nature neuroscience*, 1:621–630.
- Olson, C. R., Gettner, S. N., and Kass, R. E. (2000). Neuronal activity in macaque supplementary eye field during planning of saccades in response to pattern and spatial cues. *Journal of Neurophysiology*, 84:1369.
- Oram, M. W., Földiák, P., Perrett, D. I., and Sengpiel, F. (1998). The 'ideal homunculus': Decoding neural population signals. *Trends in Neurosciences*, 21:259–265.
- Panzeri, S., Brunel, N., Logothetis, N. K., and Kayser, C. (2010). Sensory neural codes using multiplexed temporal scales. *Trends in Neurosciences*, 33(3):111–120.
- Parzen, E. (1962). On estimation of a probability density function and mode. *The Annals of Mathematical Statistics*, 33(3):1065–1076.
- Perkel, D. H., Gerstein, G. L., and Moore, G. P. (1967). Neuronal Spike Trains and Stochastic Point Processes. *Biophysical Journal*, 7(4):391–418.
- Peyrache, A., Khamassi, M., Benchenane, K., Wiener, S. I., and Battaglia, F. P. (2009). Replay of rule-learning related neural patterns in the prefrontal cortex during sleep. *Nature Publishing Group*, 12(7):919–926.
- Ponce-Alvarez, A., Kilavik, B. E., and Riehle, A. (2010). Comparison of local measures of spike time irregularity and relating variability to firing rate in motor cortical neurons. *Journal of Computational Neuroscience*, 29:351–365.
- Pouget, A., Dayan, P., Zemel, R., and House, A. (2000). Information Processing with Population Codes. *Nature Reviews Neuroscience*, 1(2):125–132.

- Quiroga, R. Q., Reddy, L., Kreiman, G., Koch, C., and Fried, I. (2005). Invariant visual representation by single neurons in the human brain. *Nature*, 435(7045):1102–1107.
- Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition.
- Raman, B., Joseph, J., Tang, J., and Stopfer, M. (2010). Temporally Diverse Firing Patterns in Olfactory Receptor Neurons Underlie Spatiotemporal Neural Codes for Odors. *The Journal of Neuroscience*, 30(6):1994–2006.
- Reinagel, P. and Reid, R. C. (2002). Precise firing events are conserved across neurons. *The Journal of neuroscience*, 22(16):6837–6841.
- Riehle, A., Grün, S., Diesmann, M., and Aertsen, A. (1997). Spike synchronization and rate modulation differentially involved in motor cortical function. *Science*, 278:1950–1953.
- Rieke, F., Warland, D., de Ruyter van Steveninck, R. R., and Bialek, W. (1997). *Spikes: Exploring the Neural Code*. MIT Press, Cambridge, MA, 1st edition.
- Rigotti, M., Barak, O., Warden, M. R., Wang, X.-J., Daw, N. D., Miller, E. K., and Fusi, S. (2013). The importance of mixed selectivity in complex cognitive tasks. *Nature*, advance online publication(7451):585–590.
- Rosenblatt, M. (1956). Remarks on Some Nonparametric Estimates of a Density Function. *The Annals of Mathematical Statistics*, 27:832–837.
- Rudemo, M. (1982). Empirical choice of histograms and kernel density estimators. *Scandinavian Journal of Statistics*, 9(2):65–78.
- Sain, S. (1994). Adaptive Kernel Density Estimation.
- Sain, S. R. (2002). Multivariate locally adaptive density estimation. *Computational Statistics and Data Analysis*, 39:165–186.
- Sain, S. R. and Scott, D. W. (2002). Zero-Bias Locally Adaptive Density Estimators. *Scandinavian Journal of Statistics*, 29:441–460.
- Sakurai, Y. (1996). Population coding by cell assemblies - what it really is in the brain. *Neuroscience Research*, 26:1 – 16.
- Schreiner, C. E., Read, H. L., and Sutter, M. L. (2000). Modular Organization of Frequency Integration in Primary Auditory Cortex. *Annual Review of Neuroscience*, 23:501–529.
- Scott, D. W. and Terrell, G. R. (1987). Biased and Unbiased Cross-Validation in Density Estimation. *Journal of the American Statistical Association*, 82(400):1131–1146.
- Segundo, J. P., Sugihara, G., Dixon, P., Stiber, M., and Bersier, L. F. (1998). The spike trains of inhibited pacemaker neurons seen through the magnifying glass of nonlinear analyses. *Neuroscience*, 87(4):741–766.

- Seidemann, E., Meilijson, I., Abeles, M., Bergman, H., and Vaadia, E. (1996). Simultaneously recorded single units in the frontal cortex go through sequences of discrete and stable states in monkeys performing a delayed localization task. *The Journal of neuroscience*, 16(2):752–768.
- Shadlen, M. N. and Newsome, W. T. (1995). Is there a signal in the noise? *Current opinion in neurobiology*, 5:248–250.
- Shadlen, M. N. and Newsome, W. T. (1998). The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *The Journal of neuroscience*, 18(10):3870–3896.
- Sheather, S. J. and Jones, M. C. (1991). A Reliable Data-Based Bandwidth Selection Method for Kernel Density Estimation. *Journal of the Royal Statistical Society. Series B (Methodological)*, 53(3):683–690.
- Shimazaki, H. and Shinomoto, S. (2010). Kernel bandwidth optimization in spike rate estimation. *J Comput Neurosci*, 29(1-2):171–182.
- Shinomoto, S., Kim, H., Shimokawa, T., Matsuno, N., Funahashi, S., Shima, K., Fujita, I., Tamura, H., Doi, T., Kawano, K., Inaba, N., Fukushima, K., Kurkin, S., Kurata, K., Taira, M., Tsutsui, K. I., Komatsu, H., Ogawa, T., Koida, K., Tanji, J., and Toyama, K. (2009). Relating neuronal firing patterns to functional differentiation of cerebral cortex. *PLoS Computational Biology*, 5(7).
- Shinomoto, S., Miyazaki, Y., Tamura, H., and Fujita, I. (2005). Regional and laminar differences in in vivo firing patterns of primate cortical neurons. *Journal of neurophysiology*, 94(March 2005):567–575.
- Shinomoto, S., Shima, K., and Tanji, J. (2002). New classification scheme of cortical sites with the neuronal spiking characteristics. *Neural Networks*, 15:1165–1169.
- Shinomoto, S., Shima, K., and Tanji, J. (2003). Differences in spiking patterns among cortical neurons. *Neural computation*, 15:2823–2842.
- Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman and Hall/CRC Monographs on Statistics and Applied Probability. Chapman and Hall/CRC, first edition.
- Simonoff, J. (1996). *Smoothing methods in statistics*. Springer, New York.
- Softky, W. R. (1995). Simple codes versus efficient codes. *Current Opinion in Neurobiology*, 5:239–250.
- Softky, W. R. and Koch, C. (1993). The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. *The Journal of Neuroscience*, 13:334–350.
- Song, D., Chan, R. H. M., Marmarelis, V. Z., Hampson, R. E., Deadwyler, S. a., and Berger, T. W. (2009). Nonlinear modeling of neural population dynamics for hippocampal prostheses. *Neural Networks*, 22(9):1340–1351.
- Stokes, M. G., Kusunoki, M., Sigala, N., Nili, H., Gaffan, D., and Duncan, J. (2013). Dynamic Coding for Cognitive Control in Prefrontal Cortex. *Neuron*, 78(2):364–375.

- Stopfer, M., Jayaraman, V., and Laurent, G. (2003). Intensity versus identity coding in an olfactory system. *Neuron*, 39:991–1004.
- Szucs, A., Pinto, R. D., Rabinovich, M. I., Abarbanel, H. D. I., and Selverston, A. I. (2003). Synaptic modulation of the interspike interval signatures of bursting pyloric neurons. *Journal of neurophysiology*, 89:1363–1377.
- Taylor, C. C. (1989). Bootstrap choice of the smoothing parameter in Kernel density estimation. *Biometrika*, 76(4):705–712.
- Terrell, G. R. and Scott, D. W. (1992). Variable Kernel Density Estimation. *The Annals of Statistics*, 20(3):1236–1265.
- Tovee, M. J., Rolls, E. T., Treves, A., and Bellis, R. P. (1993). Information encoding and the responses of single neurons in the primate temporal visual cortex. *Journal of neurophysiology*, 70(2):640–654.
- Tran, L. T. (2010). Variable-Kernel Density Estimates for Time Series. *The Canadian Journal of Statistics*, 19(4):371–387.
- Truccolo, W., Eden, U. T., Fellows, M. R., Donoghue, J. P., and Brown, E. N. (2005). A point process framework for relating neural spiking activity to spiking history, neural ensemble, and extrinsic covariate effects. *Journal of neurophysiology*, 93(2):1074–1089.
- Tuckwell, H. C. (1988). *Introduction to theoretical neurobiology*. Cambridge studies in mathematical biology, 8. Cambridge University Press.
- Tukey, J. W. (1977). *Exploratory Data Analysis*. Addison-Wesley.
- Wand, M. P. and Jones, M. C. (1994). *Kernel smoothing*, volume 60. Crc Press.
- Wessberg, J., Stambaugh, C. R., Kralik, J. D., Beck, P. D., Laubach, M., Chapin, J. K., Kim, J., Biggs, S. J., Srinivasan, M. A., and Nicolelis, M. A. (2000). Real-time prediction of hand trajectory by ensembles of cortical neurons in primates. *Nature*, 408(1):361–365.
- Wohrer, A., Humphries, M. D., and Machens, C. K. (2013). Population-wide distributions of neural activity during perceptual decision-making. *Progress in Neurobiology*, 103:156–193.
- Yu, B. M., Afshar, A., Santhanam, G., Ryu, S., Shenoy, K., and Sahani, M. (2006). Extracting dynamical structure embedded in neural activity. *Advances in Neural Information Processing Systems*, 18:1545–1552.
- Yuan, K. and Niranjan, M. (2010). Estimating a state-space model from point process observations: a note on convergence. *Neural computation*, 22:1993–2001.
- Zougab, N., Adjabi, S., and Kokonendji, C. C. (2014). Bayesian estimation of adaptive bandwidth matrices in multivariate kernel density estimation. *Computational Statistics & Data Analysis*, 75:28–38.