

A stochastic model for the joint evaluation of burstiness and regularity in oscillatory spike trains

Dissertation
zur Erlangung des Doktorgrades der Naturwissenschaften

vorgelegt beim Fachbereich Informatik und Mathematik
der Johann Wolfgang Goethe-Universität
in Frankfurt am Main

von
Markus Bingmer
geboren in Hanau



Frankfurt 2012
(D30)

vom Fachbereich Informatik und Mathematik der

Johann Wolfgang Goethe - Universität als Dissertation angenommen.

Dekan:

Prof. Dr. Thorsten Theobald

Gutachter:

Dr. Gaby Schneider

Prof. Dr. Anton Wakolbinger

Datum der Disputation:

15. April 2013

Summary

The thesis provides a stochastic model to quantify and classify neuronal firing patterns of oscillatory spike trains. A spike train is a finite sequence of time points $t_1 < t_2 < \dots < t_n$ at which a neuron has an electric discharge (spike) which is recorded over a finite time interval $[0, \mathcal{T}]$. In this work, these spike times are analyzed regarding special firing patterns like the presence or absence of oscillatory activity and clusters (so called bursts). These bursts do not have a clear and unique definition in the literature. They are often fired in response to behaviorally relevant stimuli, e.g., an unexpected reward or a novel stimulus, but may also appear spontaneously. Oscillatory activity has been found to be related to complex information processing such as feature binding or figure ground segregation in the visual cortex. Thus, in the context of neurophysiology, it is important to quantify and classify these firing patterns and their change under certain experimental conditions like pharmacological treatment or genetical manipulation. In neuroscientific practice, the classification is often done by visual inspection criteria without giving reproducible results. Furthermore, descriptive methods are used for the quantification of spike trains without relating the extracted measures to properties of the underlying processes.

For that reason, a doubly stochastic point process model is proposed and termed 'Gaussian Locking to a free Oscillator' - GLO. The model has been developed on the basis of empirical observations in dopaminergic neurons and in cooperation with neurophysiologists. The GLO model uses as a first stage an unobservable oscillatory background rhythm \mathbb{B} which is represented by a stationary random walk $(B_i)_{i \in \mathbb{Z}}$ whose increments are normally distributed with mean μ and variance σ_1^2 . Two different model types are used to describe single spike firing ($m = 0$) or clusters of spikes ($m = 1$). Thus, m is a parameter and does not change over time. For both model types, the distribution of the random number of spikes P_i per beat B_i has mean γ , but different probability distributions (Bernoulli in the single spike case or Poisson in the cluster case). In the second stage, the random spike times are placed around their birth beat according to a normal distribution with variance σ_2^2 . These spike times represent the observed point process Φ which has the five easily interpretable parameters $\mu, \sigma_1, \sigma_2, \gamma, m$ to describe the regularity and the burstiness of the firing patterns.

It turns out that the point process Φ is stationary, simple and ergodic. Φ can be characterized as a cluster process and for the bursty firing mode ($m = 1$) as a Cox process. Furthermore, the distribution of the waiting times between spikes can be derived for some parameter combinations and for $\sigma_2 \rightarrow \infty$ the processes $\Phi = \Phi(\sigma_2)$ converge weakly to a homogenous Poisson process with intensity γ/μ . The conditional intensity function of the point process is derived which is also called autocorrelation function (ACF) in the neuroscience literature. This function arises by conditioning on a spike at time zero and measures the intensity of spikes l time units later. The autocorrelation histogram (ACH) is an estimate for the ACF. The parameters of the GLO are estimated by fitting the ACF to the ACH with a nonlinear least squares algorithm. This is a common procedure in neuroscientific practice and has the advantage that the GLO ACF can be computed for all parameter combinations and that its properties are closely related to the burstiness and regularity of the process. The precision of estimation is investigated for different scenarios using Monte-Carlo simulations and bootstrap methods.

The GLO provides the neuroscientist with objective and reproducible classification rules for the firing patterns on the basis of the model ACF. These rules are inspired by visual inspection criteria often used in neuroscientific practice and thus support and complement usual analysis of empirical spike trains. When applied to a sample data set, the model is able to detect significant changes in the regularity and burst behavior of the cells and provides confidence intervals for the parameter estimates.

Acknowledgements

This work was supported by the BMBF Project "Bernstein Fokus: Neurotechnologie Frankfurt, FKZ 01GQ0841" and I would like to express my gratitude for the financial support.

I would like to thank my advisor Dr. Gaby Schneider for her support, ideas, enthusiasm and her guidance into neuroscience.

I am grateful to Prof. Dr. Anton Wakolbinger and Dr. Brooks Ferebee for the many fruitful discussions and ideas.

I thank the colleagues from the institute of neurophysiology, Prof. Dr. Jochen Roeper and Julia Schiemann, for the good collaboration, for providing me with the data and the neurophysiological motivation of this work.

Furthermore I thank my colleague and room mate Michael Messer for the uncountable helpful comments and discussions.

In the end I would like to thank my wife Katharina, my brother Michael and my parents Karlfred and Christine, who encouraged me to finish this thesis.

Contents

| | |
|--|------------|
| Summary | i |
| Acknowledgements | iii |
| Contents | v |
| 1. Introduction | 1 |
| 1.1. Motivation | 3 |
| 1.2. Point processes | 7 |
| 2. The spike train model | 19 |
| 2.1. The GLO | 19 |
| 2.2. Basic properties | 24 |
| 2.3. Spike train simulation | 37 |
| 3. Interspike interval distributions | 47 |
| 3.1. Introductory remarks | 47 |
| 3.2. Non-bursty spike trains in the regular case | 53 |
| 3.3. Bursty spike trains in the regular case | 56 |
| 3.4. Spike inversions and the irregular case | 62 |
| 4. Autocorrelograms | 65 |
| 4.1. The autocorrelation function | 66 |
| 4.2. The autocorrelation histogram | 70 |
| 4.3. Describing burstiness and regularity | 72 |
| 5. Parameter estimation | 79 |
| 5.1. Overview of fitting procedure | 80 |
| 5.2. Estimation precision | 89 |
| 5.3. Remark: Model diagnostics | 99 |
| 6. Variability estimation of estimated parameters | 103 |
| 6.1. Methods for estimating the standard error | 103 |
| 6.2. Bootstrap confidence intervals | 112 |

| | |
|--|------------|
| 7. Data analysis | 121 |
| 7.1. Goodness of fit | 121 |
| 7.2. Classification of spike trains | 123 |
| 7.3. Comparison of different experimental conditions | 127 |
| 8. Conclusion | 129 |
| A. Additional information | 133 |
| A.1. Bursts and regularity | 133 |
| A.2. The sample data set | 134 |
| A.3. R Code | 136 |
| German Summary | 145 |
| List of Figures | 153 |
| List of Tables | 155 |
| List of R codes | 157 |
| List of Notations and Abbreviations | 159 |
| Bibliography | 165 |

Chapter 1.

Introduction

In neuroscience it is commonly agreed that information processing in the brain is based on electric discharges (*spikes*) of special nerve cells — the neurons. Mathematical description concentrates on the times $0 < t_1 < t_2 < \dots < t_n < \mathcal{T} < \infty$ (*spike times*) at which the spikes occur within the recording time interval $[0, \mathcal{T}]$. A *spike train* $\mathcal{S} = (t_1, t_2, \dots, t_n)$ is the sequence of spike times and is identified with the point configuration

$$\varphi = \sum_i \delta_{t_i}, \quad (1.1)$$

where δ_{t_i} denotes the Dirac measure, indexing the occurrence times t_i . Randomness comes in by regarding the time points t_i as \mathbb{R} -valued random variables T_i . Then, the spike trains are often modeled as a point process

$$\Phi = \sum_i \delta_{T_i}. \quad (1.2)$$

Spike train analysis (Johnson, 1996; Awiszus, 1997; Gabbiani and Koch, 1998; Brown et al., 2004) is the attempt to find patterns in spike trains which reflect some aspect of neural functioning. Famous spike train models are pseudo-Markov models (Ekholm and Hyvärinen, 1970), Poisson processes (Abeles, 1982), hidden Markov models (Camproux et al., 1996), inverse gaussian probability models (Iyengar and Liao, 1997) and Gamma processes (Barbieri et al., 2001; Shimokawa and Shinomoto, 2009).

In dopaminergic neurons, prominent empirical firing patterns (see Figure 1.1) include highly regular pacemakers (A), processes with repetitive bursts (B), irregular bursty spike trains (C) or irregular non-bursty firing patterns (D). The terms *regular* (or *irregular*) and *bursty* (or *non-bursty*) are fundamental for this work and have to be explained in some more detail and in a neuroscientific context. Both of them lack a clear unique definition and thus constitute a subjective description of special firing patterns: Regularity is reflected by a continuum of different firing patterns, ranging from very regular,

almost clock like, neurons (A) to irregular Poisson process like cells (D) (Maimon and Assad, 2009). Regular spike trains are often associated with a small variability of the waiting times between spikes. Here, regular (or irregular) is associated with the presence (or absence) of rhythmic activity and oscillation. So in contrast to mathematical intuition a homogenous Poisson process is regarded as irregular, because there is no dominant oscillation.

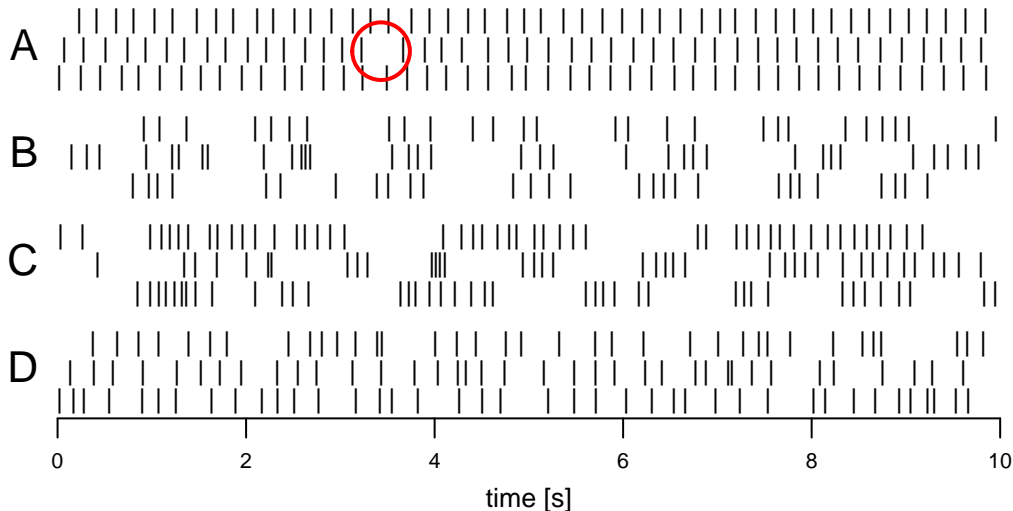


Figure 1.1: Rasterplots of prominent empirical firing patterns (every bar represents the time position of one spike): (A) highly regular pacemaker, (B) repetitive oscillatory bursts, (C) irregular bursty and (D) irregular non-bursty spike train. Red circle shows a 'hole' in the regular spike train.

Furthermore, bursts are periods of enhanced firing frequency (Cocatre-Zilgien and Delcomyn, 1992) and considered as an important unit of neuronal information (Izhikevich et al., 2003). Unfortunately, a burst also lacks a clear definition (Gourevitch and Eggermont, 2007), so in this work it will be simply regarded as a cluster of spikes. The clusters may have a different width and may contain a different number of spikes. A bursty spike train contains many bursts (B & C), in contrast to non-bursty cells (A & D). Of course, a bursty neuron can be regular (B) as well as irregular (C), whereas regular means that the burst packages follow some periodic behavior.

In the context of neurophysiology it is important to quantify and classify the firing patterns of Figure 1.1 under certain experimental conditions like pharmacological treatment (Bingmer et al., 2011; Schiemann et al., 2012; Schiemann, 2012), genetical manipulation, stimulation (Gray and Singer, 1989; Berger et al., 1990; König et al., 1995) or for different brain regions (Shinomoto et al., 2009). For neuronal populations, synchronous

oscillatory activity has been found to be related to complex information processing such as feature binding or figure ground segregation in the visual cortex (Gray and Singer, 1989; Engel et al., 1991). Furthermore, burst signals are fired in response to behaviorally relevant stimuli, e.g. an unexpected reward or a novel stimulus (Schultz, 2007; Bromberg-Martin et al., 2010). For more background information about bursts and regularity/oscillation see appendix A.1.

Overview of the thesis: The aim of this thesis is to give the experimenter an easily understandable model for the quantification and classification of firing patterns in neuronal spike trains. In the following section, the motivation is described in more detail and in Section 1.2 a short introduction is given to the theory of point processes which is the mathematical basis of spike train analysis. In Chapter 2, the spike train model is introduced which intends to describe the mentioned firing patterns. Furthermore, the basic properties of the model will be investigated, as well as a procedure explained for spike train simulation. The model is called 'Gaussian Locking to a free Oscillator' - GLO and has been published by Bingmer et al. (2011). In Chapter 3, the distribution of the waiting times between spikes get a special focus until finally in Chapter 4 the autocorrelograms are developed. These are conditional intensities of the point process and are used for classification of firing patterns, as well as parameter estimation in Chapter 5. The estimation procedure will be described in detail and the precision of estimation investigated for different conditions. In Chapter 6, different bootstrap methods are introduced for resampling and confidence interval construction with the aim of visualizing the variability of the estimated parameters of the GLO model. Finally, the model will be applied to a real data set (Chapter 7) which is part of the data set described in Bingmer et al. (2011); Schiemann et al. (2012) and Schiemann (2012). At the end, the properties of the GLO, its usability for neuroscientists and open problems will be discussed (Chapter 8). Additional details about the sample data set, burstiness and regularity, as well as R Code of the proposed procedures can be found in the appendix. This thesis contains results of the article by Bingmer et al. (2011) but are presented in some more detail here.

1.1. Motivation

The following short example represents a typically situation for a neuroscientist after an experiment: In Figure 1.1, spike train B represents the activity of a cell before the injection of a pharmacological substance and spike train C shows the same neuron after application. How should one quantify the change in the firing patterns? Is it an experimentally induced change or does it happen by chance? The following additional

main questions have arisen in our cooperation with neurophysiologists:

- How to classify a spike train into a regular or irregular and bursty or non-bursty firing pattern in an objective and consistent way?
- Does the oscillatory behavior change? And in which way?
- Is there a change in the burst behavior? What happens to the average number of spikes per burst? Does the burst width increase or decrease?

Existing spike train models and statistical approaches usually assume renewal processes as null models (Moore et al., 1966; Perkel et al., 1967; Johnson, 1996), which is insufficient when it comes to the description of bursts and oscillation. So far, bursts and oscillation are considered jointly (Kaneoke and Vitek, 1996) and investigated on the descriptive level, without relating the extracted measures to properties of the underlying processes. As a consequence, the connection between bursting and regularity/oscillation remains unspecified because no explicit spike train model is used as a basis for the detection of the features.

Thus, motivated by observations in empirical data (which will be described in Section 1.1.1), a parametric spike train model (Chapter 2) is introduced and termed 'Gaussian Locking to a free Oscillator' - GLO. The GLO (Bingmer et al., 2011) is a doubly stochastic point process model and similar to the ELO (Schneider, 2008) which stands for 'Exponential Locking to a free Oscillator'. While the ELO has been developed for the description of phase offsets in multi unit recordings (Schneider and Nikolic, 2006), the GLO parameterizes bursts and regularity in five easily understandable parameters and captures the diversity of empirically observed firing patterns. With these parameters, a spike train can be classified in a consistent and objective way which is very similar to visual inspection criteria used in empirical practice (Wilson et al., 1977; Paladini et al., 2003). Furthermore, the model assists the neuroscientist by providing an objective tool to investigate functional changes in neuronal spiking behavior in response to external conditions (Bingmer et al., 2011; Schiemann et al., 2012; Schiemann, 2012).

1.1.1. Observations in empirical data

Modeling a spike train $\mathcal{S} := (t_1, t_2, \dots, t_n)$ should be based on the knowledge of neuronal information processing and be motivated by observations in empirical data. By using exploratory techniques, simple data summaries and graphical displays one can get a first understanding of the underlying structure and its features. Keeping the scientific

issue in mind, on this basis a first probability model can be developed. Of course, this represents a process of assessment of goodness of fit and model refinement, until finally statistical inference can be made (Kass et al., 2005).

The sample data set (Schiemann et al., 2012; Schiemann, 2012) consists of 146 extra-cellular single-unit recordings of dopamine (DA) neurons of anesthetized mice in vivo (for more details see appendix A.2). It is analyzed using rasterplots, interspike interval distributions and autocorrelation histograms which aim at giving a first overview. The graphics will be discussed in the following paragraphs with examples and observations representative for the entire data set.

Rasterplots: These plots are used to represent the raw data. Every spike is drawn as a bar against the time axis. Figure 1.1 shows parts of rasterplots of four empirical spike trains of this data set. The following characteristics can be summarized: Rhythmic activity is found in bursty as well as non-bursty spike trains (A & B). There are moments of clearly increased firing intensity and spikes seem to arrange themselves in clusters (B). The cluster sizes vary in the number of spikes and the width of the cluster (B & C). The cell may stop producing spikes (A, red circle) and there is different variability between spikes (A - D).

Interspike interval distribution: A further useful graphic for the description of the underlying process is a visualization of the empirical waiting time distribution, i.e. the distribution of the length of the interval from one point to the next. In spike train analysis this is called the *interspike interval (ISI) distribution* which will be regarded in detail in Chapter 3. Figure 1.2 shows the ISI distribution of the four cells of Figure 1.1. It seems as if regular spike trains exhibit a gaussian shaped ISI distribution (A). Oscillatory bursty cells show two peaks where one peak represents intervals in a cluster and the other one intervals between clusters (B). In irregular burst processes the second peak decreases or disappears (C). Furthermore, irregular spike trains tend to have a skewed ISI distribution (D).

Autocorrelation histogram: Finally, an *autocorrelation histogram* (ACH) is used to quantify the oscillatory activity in a spike train. In neuroscience, ACHs are special histograms which show the distances $t_j - t_i$ of all pairs of spike times (t_i, t_j) with $i < j$. Thus, the ACH visualizes periodic behavior and dependence and is basically an estimate for the conditional intensity of the point process. The ACH will be discussed in detail in Chapter 4. Figure 1.3 shows the four corresponding ACHs of the empirical spike trains of Figure 1.1: For a regular neuron, one can see oscillatory activity in non-bursty (A) and bursty (B) spike trains. In the bursty case, there is a longer period and an initial peak close to zero. Because of this initial peak some spike trains can be detected as bursty even in the irregular case (C). Very irregular spike trains have a flat ACH

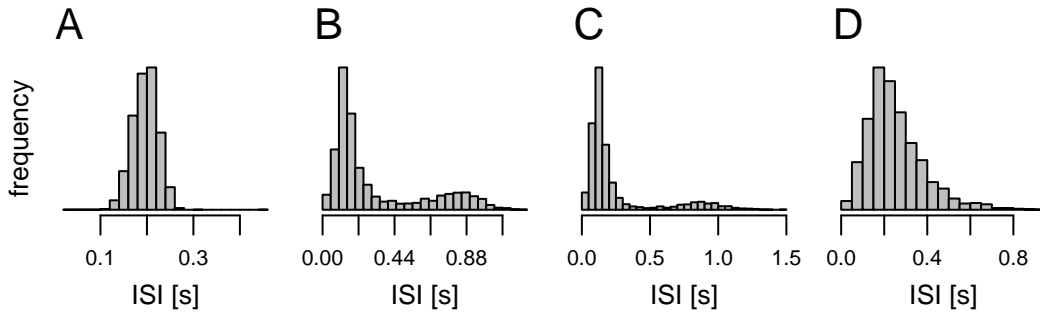


Figure 1.2: Distribution of waiting times between events in empirical spike trains seen in Figure 1.1. There are the following characteristics: (A) gaussian shaped, (B) two peaks, (C) two peaks, but only a flat second one and (D) a skewed distribution.

(D). Furthermore, regular bursty cells (B) tend to have lower frequency oscillations than regular non-bursty spike trains (A).

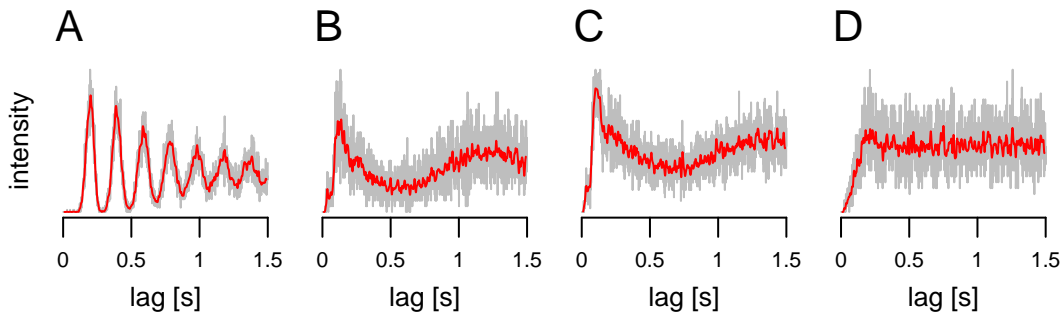


Figure 1.3: Autocorrelation histograms (ACHs) of the four empirical spike trains presented in Figure 1.1. Grey lines represent raw ACH and red lines a smoothed version resulting of a gaussian filter. The main characteristics of these ACHs are: strong oscillatory activity represented by multiple peaks in regular spike trains (A, B), peaks near zero lags for bursty cells (B, C) and flat ACH for irregular neurons (D). Regular bursty cells (B) tend to have lower frequency oscillations than regular non-bursty spike trains (A).

In this work, the proposed model is motivated by the basis of these observations and tries to account for the relevant features to quantify regularity and bursting behavior of empirical processes. Before introducing the GLO model in Chapter 2.1 the mathematical basics of point process theory as well as some important notations and definitions are summarized in the next section.

1.2. Point processes

In the following section, some basic definitions and results of the theory of point processes are summarized which is the mathematical basis of spike train analysis (Perkel et al., 1967; Johnson, 1996). We will mainly follow the books of Daley and Vere-Jones (1988) and König and Schmidt (1992), but do not give a full description of the theory. Further references for the theory of point processes are Srinivasan (1974); Snyder (1975); Cox and Isham (1980); Diggle (1983) and Thompson (1988). More mathematical requirements are needed for Matthes et al. (1978); Bremaud (1981); Franken et al. (1981); Kallenberg (1986); Karr (1986); Stoyan et al. (1987) and Reiss (1993). For an introduction to measure theory we refer to Kingman and Taylor (1966) and Bauer (1990).

1.2.1. Basic definitions

Definition 1.1 (Sets): We define $\mathbb{N} := \{1, 2, \dots\}$ to be the set of natural numbers, \mathbb{Z} as the set of integers, \mathbb{R} the set of real numbers and \emptyset as the empty set. Furthermore, we define $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$, $\mathbb{Z}^* := \mathbb{Z} \setminus \{0\}$, $\mathbb{R}_+ := [0, \infty)$, $\mathbb{R}_\infty := \mathbb{R} \cup \{-\infty, \infty\}$ and $[n] := \{1, 2, \dots, n\}$.

Definition 1.2 (Measures and measure spaces): Let E be a *complete separable metric space* (c.s.m.s.), that means a metric space which has a countable dense subset and where every Cauchy sequence in E converges in E . If \mathcal{A} is a σ -algebra on E , then the tuple (E, \mathcal{A}) is called a *measurable space*. The smallest σ -algebra which contains the open sets of E is called the *Borel σ -algebra* $\mathcal{B}(E)$ on E and elements of $\mathcal{B}(E)$ are called *Borel sets*. A *Borel measure* φ is any measure defined on the Borel sets. It is *boundedly finite* if $\varphi(A) < \infty$ for every bounded Borel set A . The space M_E consists of all boundedly finite Borel measures φ on E . The set of counting measures M_c consists of all boundedly finite, integer-valued measures φ defined on the Borel subsets $\mathcal{B}(E)$. M_c contains the subset of *simple counting measures*

$$M_s := \{\varphi \in M_c : \varphi(\{x\}) \leq 1 \forall x \in E\}. \quad (1.3)$$

Thus, $M_s \subset M_c \subset M_E$. For $\varphi \in M_E$, the triple $(E, \mathcal{A}, \varphi)$ is called a *measure space*. Furthermore, we define an *atom* of the measure φ as a $x \in E$ with $\varphi(\{x\}) > 0$ and the *support* of the measure φ as the set $S_\varphi := \{x \in E : \varphi(\{x\}) > 0\}$.

Definition 1.3 (Dirac measure): For every $x \in E$ the *Dirac measure* δ_x is given as

$$\delta_x(A) = \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases}, \quad A \in \mathcal{B}. \quad (1.4)$$

The Dirac measures are the building blocks for other counting measures, as the following proposition shows.

Proposition 1.4 (see Daley and Vere-Jones, 1988, p. 198): An element of M_E belongs to M_c if and only if it is expressible as

$$\varphi = \sum_i k_i \delta_{t_i}, \quad (1.5)$$

where each k_i is a positive integer and the distinct points $\{t_i\}$ indexing the atoms of the measure form a countable set with at most finitely many t_i in any bounded Borel set. Each φ in M_c defines via (1.5) its support counting measure φ^* by

$$\varphi^* = \sum_i \delta_{t_i}. \quad (1.6)$$

Thus φ belongs to M_s if and only if in (1.5) $k_i = 1 \forall i$.

Remark 1.5 (see Daley and Vere-Jones, 1988, p. 627): Taking a fixed origin $x_0 \in E$, let $S(r) = S(r, x_0)$ for $0 < r < \infty$ (sphere with radius r and center x_0), and introduce a distance function \hat{d} on M_E by setting

$$\hat{d}(\mu, \nu) = \int_0^\infty e^{-r} d_r(\mu^{(r)}, \nu^{(r)}) [1 + d_r(\mu^{(r)}, \nu^{(r)})]^{-1} dr, \quad (1.7)$$

where $\mu^{(r)}, \nu^{(r)}$ are the totally finite restrictions of μ, ν to $S(r)$, and d_r is the Prohorov distance between the restrictions. We call the metric topology generated by \hat{d} the \hat{w} -topology and write $\mu_k \rightarrow_{\hat{w}} \mu$ for convergence with respect to this topology.

Proposition 1.6 (see Daley and Vere-Jones, 1988, p. 628): Let $\{\mu_k : k = 1, 2, \dots\}$ and μ be measures in M_E ; then the following conditions are equivalent:

1. $\mu_k \rightarrow_{\hat{w}} \mu$;
2. $\int_E f(x) \mu_k(dx) \rightarrow \int_E f(x) \mu(dx)$ for all bounded continuous functions $f(\cdot)$ on E vanishing outside a bounded set.

Proposition 1.7 (see Daley and Vere-Jones, 1988, p. 629): M_E with the \hat{w} -topology is a c.s.m.s.

Proposition 1.8 (see Daley and Vere-Jones, 1988, p. 199): M_c is a closed subset of M_E , and M_s is a measurable subset of M_c (and hence of M_E).

Corollary 1.9 (see Daley and Vere-Jones, 1988, p. 199): As a metric space in its own right, with the \hat{w} -topology, M_c is a c.s.m.s.

Corollary 1.10 (see Daley and Vere-Jones, 1988, p. 199): $\mathcal{B}(M_c)$ is the smallest σ -algebra with respect to which the mappings $\varphi \mapsto \varphi(A)$ are measurable for each Borel set $A \in \mathcal{B}(\Omega)$.

Definition 1.11 (Point process): A *point process* Φ is a measurable mapping from a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ into the measurable space $(M_c, \mathcal{B}(M_c))$. The point process is *simple* when

$$\mathbb{P}(\Phi \in M_s) = 1. \quad (1.8)$$

Remark 1.12 (Distribution of point process): Given a point process Φ defined on the probability space $(\Omega, \mathcal{A}, \mathbb{P})$, the *distribution* of Φ is the probability measure given by $\mathbb{P}_\Phi = \mathbb{P} \circ \Phi^{-1}$ on $(M_c, \mathcal{B}(M_c))$. Point processes Φ_1 and Φ_2 are said to be *identically distributed* if $\mathbb{P}_{\Phi_1} = \mathbb{P}_{\Phi_2}$. In this case we write $\Phi_1 \stackrel{d}{=} \Phi_2$ and denote $\stackrel{d}{=}$ as equality in distribution.

For a non-negative integer-valued random variable, the probability generating function completely describes its distribution. For a point process Φ the probability generating function can be generalized to the so-called probability generating functional (Westcott, 1972).

Definition 1.13 (Probability generating functional): $\mathcal{V}(E)$ denotes the class of all \mathbb{R} -valued Borel functions f (Borel measurable functions) defined on the c.s.m.s. E with $1 - f$ vanishing outside some bounded set and satisfying $0 \leq f(x) \leq 1$ ($\forall x \in E$). The *probability generating functional (p.g.fl.)* of a point process Φ on the c.s.m.s. E is defined by

$$G[f] = G_\Phi[f] = \mathbb{E} \left[\exp \left(\int_E \log f(x) \Phi(dx) \right) \right], \quad f \in \mathcal{V}(E). \quad (1.9)$$

Because a spike train is typically defined on $E = \mathbb{R}$, it can be regarded as a sequence of random time points and be characterized by a sequence of waiting times between events. This is made more precise in the next section.

1.2.2. Characterization on the line

In spike train analysis, a spike train $\mathcal{S} = (t_1, t_2, \dots, t_n)$ is a sequence of occurrence times of undistinguishable events which cannot happen at the same time and are recorded over some finite time interval $[0, \mathcal{T}]$ with *recording time* $\mathcal{T} \in (0, \infty)$. Thus, a spike train can be regarded as a part of a realization of a point process Φ on the real line and we make the following assumptions on Φ which will be assumed for the rest of the thesis.

Assumption 1.14 (Main assumptions): The r.v. $\Phi: (\Omega, \mathcal{A}, \mathbb{P}) \mapsto (M_c, \mathcal{B}(M_c))$ is a point process with

1. $E = \mathbb{R}$ with $\mathcal{B} = \mathcal{B}(\mathbb{R})$,
2. $\mathbb{P}(\Phi(A) < \infty) = 1, \forall A \in \mathcal{B}$ bounded,
3. $\mathbb{P}(\Phi(\mathbb{R}) = \infty) = 1$,
4. $\mathbb{P}(\Phi \in M_s) = 1$.

Thus, we consider only processes with finitely many points in every bounded set, but with infinitely many points on the real line and focus on simple point processes. So we have a sequence of random variables $\dots < T_{-1} < T_0 < 0 \leq T_1 < T_2 < \dots$ representing the random occurrence times of the point process Φ and a sequence of deterministic points $\dots < t_{-1} < t_0 < 0 \leq t_1 < t_2 < \dots$ representing a realization φ of Φ .

In the following, we want to make the description of a point process as a random sequence of points more precise:

Definition 1.15: For every $\varphi \in M_s$ we number the elements of the support as $t_n(\varphi)$, $n \in \mathbb{Z}$, according to

$$t_n(\varphi) := \begin{cases} \min\{t : t \geq 0, \varphi(\{t\}) > 0\} & , n = 1 \\ \min\{t : t > t_{n-1}(\varphi), \varphi(\{t\}) > 0\} & , n > 1 \\ \max\{t : t < t_{n+1}(\varphi), \varphi(\{t\}) > 0\} & , n < 1. \end{cases} \quad (1.10)$$

Proposition 1.16 (see König and Schmidt, 1992, p. 34): For every integer $n \in \mathbb{Z}$, $t_n(\varphi): \varphi \rightarrow t_n(\varphi)$ is a measurable mapping from $(M_c, \mathcal{B}(M_c))$ into $(\mathbb{R}_\infty, \mathcal{B}(\mathbb{R}_\infty))$.

Remark 1.17: We denote the random occurrence times of the point process Φ by

$$T_n := t_n(\Phi) \quad (1.11)$$

and the occurrence times of a realization φ of Φ by

$$t_n := t_n(\varphi). \quad (1.12)$$

Thus a simple point process Φ can be described by $\{T_i : i \in \mathbb{Z}\}$ and a realization φ by the set $\{t_i : i \in \mathbb{Z}\}$.

Definition 1.18 (Spike train): In this setting, a *spike train* \mathcal{S} is the restriction of the realization φ of a simple point process Φ to the recording interval $[0, \mathcal{T}]$. Thus

$$\mathcal{S} := (t_1, t_2, \dots, t_n) = \{t_i \in S_\varphi : i \in \mathbb{Z}\} \cap [0, \mathcal{T}] \quad (1.13)$$

with $n = \#\{\{t_i \in S_\varphi : i \in \mathbb{Z}\} \cap [0, \mathcal{T}]\}$ and $t_1 < t_2 < \dots < t_n$.

The equivalence between counting measures and random occurrence times will be discussed in the next proposition:

Proposition 1.19 (Specification of point processes on the line): The following descriptions of a simple point process Φ on \mathbb{R} are equivalent:

1. Counting measure:

$$\Phi(A) = \#\{i \in \mathbb{Z} : T_i \in A\} = \sum_{i \in \mathbb{Z}} \delta_{T_i}(A), \quad \forall A \in \mathcal{B}(\mathbb{R}). \quad (1.14)$$

2. Nondecreasing integer-valued step functions:

$$N(t) := \begin{cases} \Phi((0, t]) & , t > 0 \\ 0 & , t = 0 \\ -\Phi((t, 0]) & , t < 0. \end{cases} \quad (1.15)$$

3. Sequence of points $\{T_i : i \in \mathbb{Z}\}$, with $T_i := \inf\{t \in \mathbb{R} : N(t) \geq i\}$, $i \in \mathbb{Z}$.
4. Sequence of intervals together with the first occurrence time $\{W_i : i \in \mathbb{Z}\} \cup \{T_1\}$, where $W_i := T_{i+1} - T_i$, $\forall i \in \mathbb{Z}$.

Proof. Obviously $\Phi(A)$ is a random counting measure and thus defines a point process on \mathbb{R} . $N(t)$ determines $\Phi(A)$ for all Borel sets A and vice versa. Because of the fundamental relation

$$\{T_i \leq t\} = \{N(t) \geq i\} \quad (1.16)$$

specifying the sequence of points $\{T_i\}$ is equivalent to specifying the function $N(t)$. From the sequence of points the sequence of intervals can be derived and together with some given occurrence time (typically T_1) the converse is true. \square

Remark 1.20: In this setup the functions $\Phi(A), N(t), T_i, W_i$ are well-defined random variables and Figure 1.4 illustrates their relationship.



Figure 1.4: Point process on the real line with illustration of the relationship between the nondecreasing inter-valued step function $N(t)$, the sequence of time points $\{T_i\}$ and the sequence of intervals $\{W_i\}$.

1.2.3. Stationarity

In the following, the concept of stationarity is discussed in the context of random point processes. Stationarity basically means that the structure of the process does not change under translation of the time axis. Still, there are different types of stationarity to consider.

Definition 1.21 (Stationary point process): A point process Φ is *stationary* when for every $r \in \mathbb{N}$ and all bounded Borel subsets A_1, \dots, A_r of \mathbb{R} the joint distribution of $\{\Phi(A_1 + t), \dots, \Phi(A_r + t)\}$ does not depend on t ($-\infty < t < \infty$). This is equivalent to

$$\sum_i k_i \delta_{T_i} \stackrel{d}{=} \sum_i k_i \delta_{T_i+t}, \quad \forall t \in \mathbb{R}. \quad (1.17)$$

Because the full restriction of stationarity is seldom required, one also considers weaker forms of stationarity.

Definition 1.22 (Crudely stationary): A point process Φ is *crudely stationary* when

$$p_k(x) := \mathbb{P}(\Phi((t, t+x]) = k), \quad x > 0, k = 0, 1, \dots \quad (1.18)$$

depends on the length x but not the location $t \in \mathbb{R}$.

Definition 1.23 (Interval stationary): A point process Φ is *interval stationary* if for every $r \in \mathbb{N}$ and all integers i_1, \dots, i_r the joint distribution of $\{W_{i_1+k}, \dots, W_{i_r+k}\}$ does not depend on $k \in \mathbb{Z}$.

One should note that interval stationarity does not necessarily imply stationarity. A necessary and sufficient condition for a point process to be stationary is given by the following proposition:

Proposition 1.24 (see Daley and Vere-Jones, 1988, p. 318): Let Φ be a point process on the state space $E = \mathbb{R}^d$. The following condition is necessary and sufficient for Φ to be stationary: For each $u \in \mathbb{R}^d$ and $f \in \mathcal{V}(\mathbb{R}^d)$, the p.g.f. G of Φ satisfies

$$G[f(\cdot)] = G[f(\cdot - u)]. \quad (1.19)$$

Definition 1.25 (k th-order stationary): The point process Φ is *k th-order stationary*, if Φ is stationary and its k th moment measure exists, which means that

$$\mathbb{E}[\Phi(A_1)\Phi(A_2) \dots \Phi(A_k)] < \infty \quad (1.20)$$

for all bounded Borel sets A_1, \dots, A_k .

Finally, the density of points in a stationary point process is measured in a natural way:

Definition 1.26 (Mean density): The *mean density* of a stationary point process Φ on \mathbb{R} is given as

$$\underline{m} := \mathbb{E}[\Phi(0, 1]]. \quad (1.21)$$

Proposition 1.27 (Khinchin's Existence Theorem, see Daley and Vere-Jones, 1988, p. 44): For a stationary (or even crudely stationary) point process, the limit

$$\lambda := \lim_{h \downarrow 0} \frac{\mathbb{P}(\Phi(0, h] > 0)}{h} \quad (1.22)$$

exists, though it may be infinite. The parameter λ is called the *intensity* of the point process.

These definitions coincide in the following theorem.

Proposition 1.28 (Korolyuk's Theorem, see Daley and Vere-Jones, 1988, p. 45): For a crudely stationary simple point process,

$$\lambda = \underline{m}, \quad (\text{finite or infinite}). \quad (1.23)$$

Remark 1.29: In the context of neuroscience and a crudely stationary simple point process, λ (or \underline{m}) is called the *mean firing rate*.

Proposition 1.30 (see Daley and Vere-Jones, 1988, p. 58): For a stationary point process with finite mean density $\underline{m} = \mathbb{E}[\Phi(0, 1]]$, $\xi \equiv \lim_{x \rightarrow \infty} \Phi(0, x]/x$ exists a.s. and is a random variable with $\mathbb{E}[\xi] = \underline{m}$.

Definition 1.31 (Ergodic point process): A stationary point process with finite mean density \underline{m} is *ergodic* if

$$\mathbb{P}(\Phi((0, x])/x \rightarrow \underline{m} \quad (x \rightarrow \infty)) = 1. \quad (1.24)$$

1.2.4. Important examples

Now three special classes of point processes on \mathbb{R} are introduced which are related to each other: Poisson processes, random walks and renewal processes. For a more detailed introduction of the theory see Spitzer (1976) for random walks, Cox and Isham (1980) and Kingman (1993) for Poisson processes, Cox (1962) for renewal processes or Feller (1966) for all of them. For a detailed discussion of the connection between random walks and point processes see Daley and Oakes (1974) and for random walks regarded as a general renewal process see Gut (2009).

Definition 1.32 (Poisson process): Let $\lambda(t)$ be a non-negative real valued measurable function and for $a < b$, $\Lambda(a, b) = \int_a^b \lambda(t) dt$. Then a point process Φ is called an *inhomogeneous Poisson process*, if

$$\mathbb{P}(\Phi((a_i, b_i]) = n_i, i = 1, \dots, k) = \prod_{i=1}^k \frac{[\Lambda(a_i, b_i)]^{n_i}}{n_i!} e^{-\Lambda(a_i, b_i)} \quad (1.25)$$

for $a_i < b_i \leq a_{i+1}$. If $\lambda(t)$ is constant, we call Φ a *homogeneous Poisson process*.

Remark 1.33 (Properties of Poisson process): It follows that the number of points in each finite interval is Poisson distributed, all disjoint intervals are independent and the process is in general not stationary. If Φ is a homogeneous Poisson process, it is stationary. In more general spaces (but also in \mathbb{R}), $\Lambda(a, b)$ can be replaced by a boundedly finite Borel measure $\Lambda(\cdot)$ which is called the *parameter measure* or *intensity measure*.

Proposition 1.34 (Superposition Theorem, Kingman 1993, p. 16): Let Φ_1, Φ_2, \dots be a countable collection of independent Poisson processes on E and let Φ_n have intensity measure Λ_n for each n . Then their superposition

$$\Phi = \bigcup_{n=1}^{\infty} \Phi_n \quad (1.26)$$

is a Poisson process with intensity measure $\Lambda = \sum_{n=1}^{\infty} \Lambda_n$.

Definition 1.35 (Random walk and renewal process): Let X_1, X_2, \dots be independent identically distributed (i.i.d.) random variables on \mathbb{R} and $S_n, n \geq 0$, be the partial sums with $S_0 = 0$ and $S_n = X_1 + \dots + X_n$. The sequence $(S_n)_{n \in \mathbb{N}_0}$ is called a *random walk* on \mathbb{R} . If X_1 is nonnegative, it is called a *renewal process*. The distribution of X_1 is denoted by \mathbb{P}_{X_1} , and we suppose that the mean and variance of \mathbb{P}_{X_1} exist and are denoted by ν and σ^2 .

Definition 1.36 (Renewal measure): The expected number of visits in the interval or set $I \in \mathcal{B}(\mathbb{R})$ is defined as

$$U(I) := \sum_{n=0}^{\infty} \mathbb{P}(S_n \in I) \quad (1.27)$$

and defines a measure on the Borel sets of \mathbb{R} , called the *renewal measure*.

Definition 1.37 (Arithmetic distribution): A distribution \mathbb{P}_{X_1} on the Borel sets of $(-\infty, \infty)$ is said to be *arithmetic* if and only if there is a $d > 0$ such that the distribution is concentrated on $\{0, \pm d, \pm 2d, \dots\}$.

Proposition 1.38 (Renewal theorem, see Woodroffe, 1982, p. 16): If \mathbb{P}_{X_1} has mean ν with $0 < \nu < \infty$, then $U(I) < \infty$ for every finite interval I . If \mathbb{P}_{X_1} is non-arithmetic, then

$$\lim U(a + I) = \begin{cases} \nu^{-1}h & , a \rightarrow \infty \\ 0 & , a \rightarrow -\infty \end{cases} \quad (1.28)$$

for every finite interval I , where h denotes the length of the interval.

Definition 1.39 (Excess of random walk): Let $(S_n)_{n \in \mathbb{N}_0}$ be a random walk started at 0 and for $a \geq 0$, let

$$\tau_a := \inf\{n \in \mathbb{N} : S_n > a\} \quad (1.29)$$

be the time at which the random walk first reaches the height a , or ∞ . Given the event $\{\tau_a < \infty\}$,

$$R_a := S_{\tau_a} - a \tag{1.30}$$

is the *excess* over the boundary a at the time which the random walk first crosses a .

Remark 1.40: R_a is also called the *overshoot*. Alternatively in the case of a renewal process, if S_n , $n \geq 1$, represents the times at which equipment is renewed, then R_a is the *residual waiting time* until the next renewal after time a . If $0 < \nu < \infty$, it follows that $\mathbb{P}(\tau_a < \infty) = 1$ and R_a is defined.

Proposition 1.41 (see Woodroffe, 1982, p. 18): Suppose that $0 < \nu < \infty$. If \mathbb{P}_{X_1} is non-arithmetic, then R_a has a limiting distribution H_∞ as $a \rightarrow \infty$, where

$$H_\infty(dx) = \frac{1}{\mathbb{E}[S_{\tau_0}]} \mathbb{P}(S_{\tau_0} > x) dx, \quad x > 0. \tag{1.31}$$

Corollary 1.42 (Asymptotic residual waiting time): Suppose that $0 < \nu < \infty$ and \mathbb{P}_{X_1} is non-arithmetic, then

$$R_\infty := \lim_{a \rightarrow \infty} R_a \tag{1.32}$$

exists and R_∞ is called the *asymptotic residual waiting time*.

1.2.4.1. The connection between point processes and random walks

Let $(S_n)_{n \in \mathbb{N}_0}$ be a random walk with $\mathbb{E}[S_1] = \nu \in (0, \infty)$. The random counting measure Φ can be defined by setting

$$\Phi(A) := \#\{n \in \mathbb{N}_0 : S_n \in A\} \tag{1.33}$$

for all bounded Borel subsets A of \mathbb{R} . Because of the renewal theorem (Proposition 1.38) $\Phi(A)$ is finite with probability one. So, every renewal process is a random walk and every random walk is a point process with random time points S_0, S_1, S_2, \dots which are not necessarily ordered. Figure 1.5 shows an example of a random walk represented as a point process, where the points $T_0 < T_1 < \dots < T_6$ are the ordered values of S_0, S_1, \dots, S_6 .

Chapter 1 ends with a remark, how to generalize a random walk to a stationary point process which is used for the construction of the GLO spike train model in the next chapter.

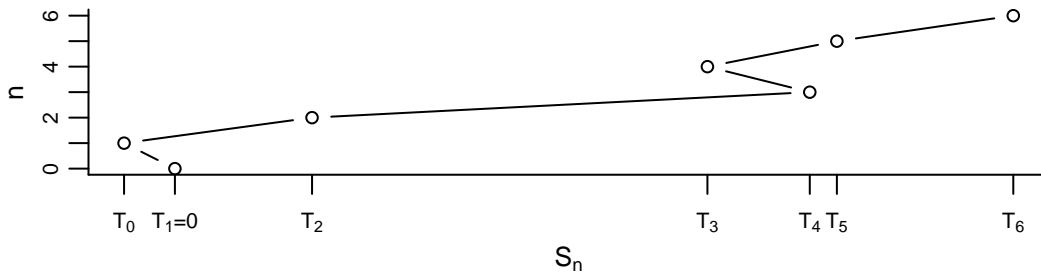


Figure 1.5: A random walk regarded as a point process.

Remark 1.43 (Stationary random walk): Let $X_i, i \in \mathbb{Z}$, be i.i.d. with expectation $\nu \in (0, \infty)$ and \mathbb{P}_{X_1} non-arithmetic. For $n \in \mathbb{N}$, the steps of the random walk are written as $S_n := S_0 + X_1 + \dots + X_n$ and as $S_{-n} := S_0 - (X_0 + X_{-1} + \dots + X_{-n+1})$ with $S_0 := 0$. Thus, the random walk is started at zero and evolves to the left and to the right and represents a point process Φ on \mathbb{R} by setting

$$\Phi(A) := \#\{n \in \mathbb{Z} : S_n \in A\}. \quad (1.34)$$

Because $0 < \nu < \infty$ and \mathbb{P}_{X_1} is non-arithmetic, assumption 1.14 holds. Thus Φ can be described by the sequence $\{T_i : i \in \mathbb{Z}\}$ which are the ordered elements of $(S_n)_{n \in \mathbb{Z}}$ with $T_1 = 0$. Furthermore, we get the sequence of intervals $\{W_i : i \in \mathbb{Z}\}$ which is interval stationary (Daley and Oakes, 1974), but the intervals are not necessarily independent. To make Φ a stationary point process, we have to set the origin of the time axis at random. This is done by taking a large boundary a (which represents the new origin) and by renaming the indices of the random walk

$$\dots, \quad S'_{-1} := S_{\tau_a-2}, \quad S'_0 := S_{\tau_a-1}, \quad S'_1 := S_{\tau_a}, \quad S'_2 := S_{\tau_a+1}, \quad \dots$$

In the same way as (1.32) converges, this holds true for all neighbor points. Thus, the steps of the shifted random walk are given as

$$S'_n := \lim_{a \rightarrow \infty} S_{n+\tau_a-1} - a, \quad n \in \mathbb{Z}. \quad (1.35)$$

Then the point process $\Phi(A) := \#\{i \in \mathbb{Z} : S'_i \in A\}$ is stationary in the sense of point processes and $(S'_i)_{i \in \mathbb{Z}}$ is called a *stationary random walk*. Furthermore, we have $S'_1 = R_\infty$ in distribution. Figure 1.6 visualizes this construction.

Remark 1.44 (Independence of steps in a stationary random walk): When started at zero, the increments of the random walk $S_{n+1} - S_n$ are independent of each other ($\forall n \in \mathbb{Z}$). But when regarding the stationary random walk (using the construction mentioned in Remark 1.43 for setting an arbitrary time origin) this is not true anymore.

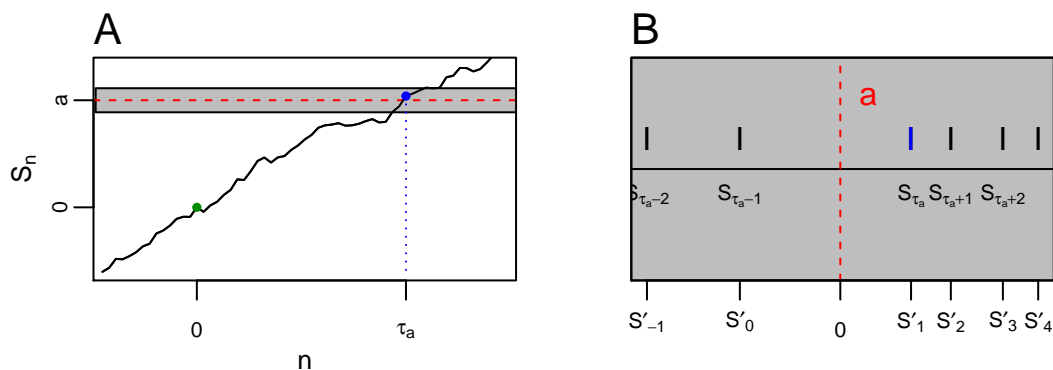


Figure 1.6: Construction of a stationary random walk. (A) A random walk started at zero (green point) and going to the left and to the right. Red dotted line represents the boundary a and the blue point the height S_{τ_a} . (B) shows the grey area of plot (A). The threshold a becomes the new origin and the heights of S_n are renamed to S'_n according to equation (1.35).

In this case, only all increments after the time τ_a are independent of each other. This means that $S_{n+1} - S_n$ are independent for all $n \geq \tau_a$ or in terms of the stationary random walk that $S'_{n+1} - S'_n$ are independent for all $n \in \mathbb{N}$. All other increments are not independent and not normally distributed anymore. The reason is that because of the construction of the stationary random walk, τ_a is defined to be the time at which the height a is reached for the 'first' time. Thus, when going backwards from S_{τ_a} (i.e. regarding all S_n with $n < \tau_a$), we have to condition on the event that the random walk does not reach a . Anyway, the exact distribution and dependence structure for these increments is not of importance, because we will only focus on increments after time τ_a .

In the next chapter, the GLO model (Gaussian Locking to a free Oscillator) is introduced which uses this kind of stationary random walk to model a random background process. This background process intends to describe oscillatory activity of the spike train with the mean and variance of the normal distribution and is assumed unobservable.

Chapter 2.

The spike train model

The following model (Bingmer et al., 2011) is able to quantify regularity and burstiness of single neurons and to reproduce a lot of different firing patterns. It is motivated by observations in empirical data which have been discussed in Section 1.1.1. The model contains a characterization of two primary components: single events or clusters of events and interevent intervals, which is similar to other characterizations (Smith and Smith, 1965; Bair et al., 1994; Kaneoke and Vitek, 1996). Because an oscillatory background rhythm is used with normally distributed intervals, errors and normally distributed burst packages, the model is termed 'Gaussian Locking to a free Oscillator' - GLO. It is very similar to the ELO which stands for 'Exponential Locking to a free Oscillator' (Schneider, 2008).

Mathematically, a doubly stochastic model is used to describe a spike train. The starting point is given by a stationary point process where each point is independently copied or thinned and randomly shifted. After the formal introduction of the GLO (Section 2.1), some basic properties are derived (Section 2.2). The chapter ends with an algorithm for spike train simulation (Section 2.3) which will be important for later chapters.

2.1. The GLO

The model assumptions are motivated by observations in empirical spike trains and cover important features of them (compare Section 1.1.1). For the rest of this work, we denote an underlying probability space by $(\Omega, \mathcal{A}, \mathbb{P})$. The GLO makes extensive use of the normal distribution. A random variable X is said to be normally distributed with mean $\mu \in \mathbb{R}$ and variance $\sigma^2 \in (0, \infty)$, which we write as $X \sim \mathcal{N}(\mu, \sigma^2)$, if its cumulative

distribution function (c.d.f.) is given as

$$F_{\{\mu, \sigma^2\}}(x) := \mathbb{P}(X \leq x) = \int_{-\infty}^x \varphi_{\{\mu, \sigma^2\}}(y) dy, \quad (2.1)$$

with $\varphi_{\{\mu, \sigma^2\}}(y)$ denoting the density function of the normal distribution which is given as

$$\varphi_{\{\mu, \sigma^2\}}(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y-\mu)^2}{2\sigma^2}}. \quad (2.2)$$

The GLO is constructed in the following way:

- 1. The background rhythm:** It is assumed that there is some kind of background rhythm (BR) consisting of random background beats B_i , $i \in \mathbb{Z}$, to which the spikes lock. For this purpose, let $X_0, X_{\pm 1}, X_{\pm 2}, \dots$ be i.i.d. according to a normal distribution with mean $\mu \in (0, \infty)$ and variance $\sigma_1^2 \in (0, \infty)$, so the distribution \mathbb{P}_{X_1} is non-arithmetic. The random walk $(S_n)_{n \in \mathbb{Z}}$ is constructed and made a stationary process $(S'_n)_{n \in \mathbb{Z}}$ as described in Remark 1.43 on page 17. Setting $B_i := S'_i$, $\forall i \in \mathbb{Z}$, the BR is described as a stationary random walk

$$(B_i)_{i \in \mathbb{Z}}. \quad (2.3)$$

Because of the Renewal theorem (see Proposition 1.38 on page 15), there are finitely many points in each bounded Borel set and because the stationary random walk can be regarded as a random counting measure by setting $\mathbb{B}(A) := \#\{i \in \mathbb{Z} : B_i \in A\}$, it can be written as a stationary point process

$$\mathbb{B} = \sum_{i \in \mathbb{Z}} \delta_{B_i}. \quad (2.4)$$

The points B_i are called the *beats* of the BR and are assumed unobservable.

- 2. The firing mode and number of spikes:** It is assumed that a spike train has either a non-bursty ($m = 0$, single spiking) or a bursty ($m = 1$, cluster of spikes) firing mode, where m is deterministic and does not change over time. According to the firing mode m , every beat B_i of the background rhythm is assumed to give rise to a random number P_i of spikes, where all P_i are i.i.d. according to the following rules.

In the non-bursty mode ($m = 0$), the random variable P_i is Bernoulli-distributed with parameter $\gamma \in (0, 1]$, which we write as $P_i \sim \text{Bernoulli}(\gamma)$, and has probability weights

$$\mathbb{P}(P_i = 1) = \gamma \quad \text{and} \quad \mathbb{P}(P_i = 0) = 1 - \gamma. \quad (2.5)$$

This allows to model holes in the spike train (compare red circle in Figure 1.1).

In the bursty mode ($m = 1$), the number of spikes P_i at beat B_i is assumed to be Poisson distributed with parameter $\gamma \in (0, \infty)$ which is used to model clusters or bursts of spikes. In this case, P_i has probability weights

$$\mathbb{P}(P_i = k) = \frac{\gamma^k}{k!} e^{-\gamma}, \quad k \in \mathbb{N}_0. \quad (2.6)$$

and is written as $P_i \sim \text{Pois}(\gamma)$.

Thus in summary

$$P_i^{(m)} \sim \begin{cases} \text{Bernoulli}(\gamma), & m = 0 \\ \text{Pois}(\gamma), & m = 1 \end{cases} \quad (2.7)$$

$\forall i \in \mathbb{Z}$. So γ represents the mean number of spikes per beat. For simplicity, we will omit the index (m) in the future if m follows from the context.

3. Spike variation: It is assumed that the spikes vary randomly around their birth beats which is modeled with a family of random variables

$$(Z_{i,j}), \quad \text{where } Z_{i,j} \text{ are i.i.d. } \mathcal{N}(0, \sigma_2^2), \quad (2.8)$$

with $\sigma_2 \in (0, \infty)$, $j \in \mathbb{N}$ and $i \in \mathbb{Z}$. For example, if there are $P_i = 2$ spikes ($m = 1$) at beat B_i , they will be placed according to $Z_{i,1}$ and $Z_{i,2}$ such that the spike times are given as $B_i + Z_{i,1}$ and $B_i + Z_{i,2}$.

4. Independence of steps 1-3: All random variables Z_{i_1, i_2}, P_{i_3} are independent $\forall i_2 \in \mathbb{N}$ and $\forall i_1, i_3 \in \mathbb{Z}$ and independent of the stationary random walk $(B_i)_{i \in \mathbb{Z}}$.

Figure 2.1 shows a visualization of the GLO assumptions. In the first instance, the background rhythm is generated. Then, dependent on the firing mode, the random number of spikes is drawn for each beat and all spikes are placed around their birth beat. The model can be summarized in a more compact form using its representation as a random counting measure.

Definition 2.1 (GLO process): Let $\psi = (\mu, \sigma_1, \sigma_2, \gamma, m)$ be a parameter vector with $m \in \{0, 1\}$, $\mu, \sigma_1, \sigma_2 \in (0, \infty)$ and $\gamma \in (0, 1]$ if $m = 0$ or $\gamma \in (0, \infty)$ if $m = 1$. Then a point process Φ is a GLO process with parameter vector ψ , if the counting measure representation of Φ is of the form

$$\Phi = \sum_{i \in \mathbb{Z}} \sum_{k=1}^{P_i^{(m)}} \delta_{B_i + Z_{i,k}}, \quad \text{with} \quad (2.9)$$

1. a stationary random walk $(B_i)_{i \in \mathbb{Z}}$, with $B_{i+1} - B_i$ i.i.d. $\mathcal{N}(\mu, \sigma_1^2) \forall i \in \mathbb{N}$,
2. number of points $P_i^{(m)} \sim \begin{cases} \text{Bernoulli}(\gamma), & m = 0 \\ \text{Pois}(\gamma), & m = 1 \end{cases}, \forall i \in \mathbb{Z}$,
3. random variations $(Z_{i,k})$, where $Z_{i,k}$ are i.i.d. $\mathcal{N}(0, \sigma_2^2), \forall i \in \mathbb{Z}$ and $\forall k \in \mathbb{N}$,
4. all variables Z_{i_1, i_2}, P_{i_3} independent $\forall i_2 \in \mathbb{N}, \forall i_1, i_3 \in \mathbb{Z}$ and independent of $(B_i)_{i \in \mathbb{Z}}$.

We write $\Phi \sim \text{GLO}(\mu, \sigma_1, \sigma_2, \gamma, m)$ or in short form $\Phi \sim \text{GLO}(\psi)$ or $\Phi \sim \text{GLO}(m)$ to focus on a special firing mode.

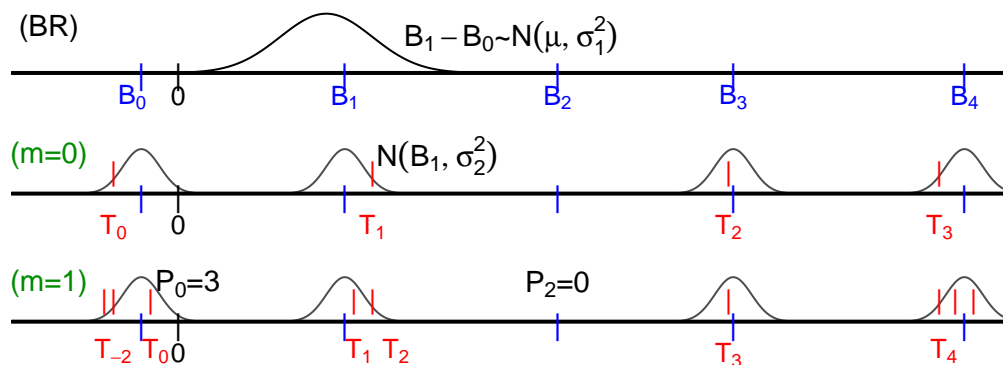


Figure 2.1: Construction of the GLO spike train model. The background rhythm (BR) of the GLO has independent normally distributed increments $B_{i+1} - B_i \sim \mathcal{N}(\mu, \sigma_1^2)$. The number of spikes P_i at each beat B_i (blue) depends on the firing mode m (green). In the non-bursty case ($m = 0$) P_i is Bernoulli(γ)-distributed and Poiss(γ) in bursty mode ($m = 1$). Every spike T_k (red) is placed around its birth beat B_i according to $\mathcal{N}(B_i, \sigma_2^2)$.

2.1.1. Interpretation of parameters and assumptions

One of the aims of this construction is to give the experimenter an easily understandable and interpretable model for the overall quantification of burstiness and regularity in single spike trains. The background rhythm evokes oscillatory activity with period μ and intervals with variance σ_1^2 . So μ is basically a scale parameter and a change of σ_1 has a direct impact on the regularity of the process. The firing mode m tells if the process is bursty ($m = 1$) oder non-bursty ($m = 0$). The parameter γ denotes the expected number of spikes per beat of the background rhythm and is thus connected with the burstiness. In addition, the parameter σ_2 represents the width of the bursts

for $m = 1$, and in the non-bursty mode, σ_2 may be necessary to describe variations and correlations in the distribution of waiting times even when the background rhythm is regular. Since the model intends to describe regular oscillations with distinguishable bursts, σ_1 and σ_2 should usually be small relative to μ .

The assumptions are chosen to reproduce important characteristics of the data set (compare Section 1.1.1). As mentioned in the previous paragraph, the model parameters have a natural interpretation, but some motivations regarding the distributions have to be clarified. While the choice of a Bernoulli and a Poisson distribution for the number of spikes seems to be obvious, the choice of a normal distribution for the increments of the BR is not.

Because the normal distribution is defined on \mathbb{R} , there is always a positive probability to have increments $B_{i+1} - B_i < 0$. However, this probability is small for $\sigma_1/\mu < 1/3$ (< 0.0014), but the model is not restricted to such parameter combinations. A negative increment only implies that the corresponding beats occur in different order than they were originally generated in the model. So, the distribution of ordered interbeat intervals (IBIs) differs from the normal distribution because it takes only positive values. Figure 2.2 shows that the resulting distribution of IBIs changes from a normal density towards a density that decreases monotonically similar to an exponentially distribution (if the IBI distribution is exponential, the process will be a Poisson cluster process or more specifically a Neyman-Scott process (Neyman and Scott, 1972)).

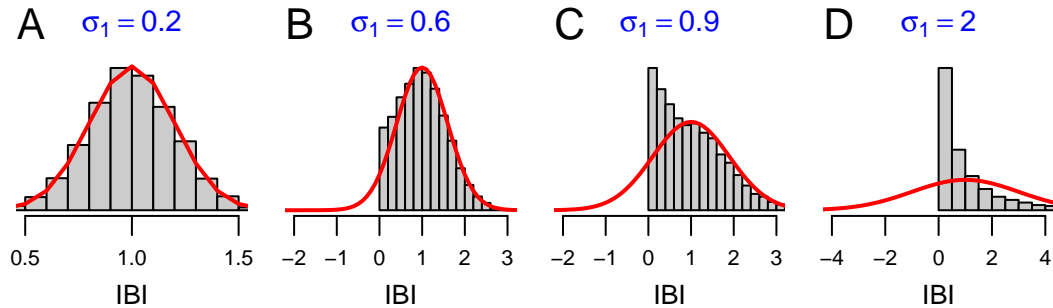


Figure 2.2: Interbeat interval (IBI) distribution visualized by histograms, generated from a normal distribution with mean $\mu = 1$ and variance σ_1^2 (red lines). Although the normal distribution has values in \mathbb{R} , the IBI distributions are non-negative and tend to an exponentially looking shape for increasing σ_1 .

Therefore, the model is also able to describe activity that is not locked to a regularly oscillating backbone process. But the two parameters of the normal distribution are considered a simple and sufficiently general description in order to grasp the period and

the precision of the background random walk. In addition, the simple assumption of normal distribution leads to easily interpretable parameters and in contrast to other distributions allows straightforward calculations.

2.2. Basic properties

This section shows a few basic properties of the GLO process Φ . First of all, some details of the BR are investigated (Section 2.2.1). It is shown that \mathbb{B} is a simple point process, the distribution of B_1 is derived with help of the asymptotic residual waiting time R_∞ and possible inversions of beats are discussed. Secondly, because the final process Φ results of independent thinning or copying and translations of the points in \mathbb{B} , it turns out that Φ is also stationary. Furthermore, it is shown that Φ is a simple point process and the mean firing rate is derived (Section 2.2.2). Thirdly, a short characterization of the GLO process is given in terms of a Cluster process and in the special case of $m = 1$, there is a characterization as a Cox process (Section 2.2.3). Finally the convergence of Φ to a homogenous Poisson process is proved for increasing σ_2 (Section 2.2.4). The waiting times between spikes get a special focus in Chapter 3.

2.2.1. Background rhythm

Let \mathbb{B} be as defined in equation (2.4). The following three lemmas follow immediately from the construction of the BR (compare Remark 1.43 on page 17).

Lemma 2.2: \mathbb{B} is a simple point process.

Proof. The condition for \mathbb{B} being simple is given as $\mathbb{P}(\mathbb{B} \in M_s) = 1$ (compare equation (1.3)) which is the same as the condition that all beats B_i are different from each other with probability one. We have

$$\mathbb{P}(\{B_i \neq B_j, \forall i, j \in \mathbb{Z}, i \neq j\}) = 1 - \mathbb{P}\left(\bigcup_{\substack{i, j \in \mathbb{Z} \\ i \neq j}} \{B_i = B_j\}\right) \quad (2.10)$$

$$\geq 1 - \sum_{\substack{i, j \in \mathbb{Z} \\ i \neq j}} \mathbb{P}(\{B_i = B_j\}) = 1 \quad (2.11)$$

The last step follows, because the beats B_i are \mathbb{R} -valued r.v.s without atoms. \square

Lemma 2.3: \mathbb{B} is a stationary point process.

Proof. Because of the construction of \mathbb{B} in Remark 1.43 on page 17, it follows that $\sum_{i \in \mathbb{Z}} \delta_{B_i} \stackrel{d}{=} \sum_{i \in \mathbb{Z}} \delta_{B_i+x}, \forall x \in \mathbb{R}$. \square

Lemma 2.4: The mean density of \mathbb{B} is $\mathbb{B}((0, 1]) = 1/\mu$.

Proof. This follows directly of the Renewal theorem (Proposition 1.38 on page 15). \square

Proposition 2.5: \mathbb{B} is an ergodic point process.

Proof. According to definition 1.31 on page 14, \mathbb{B} is ergodic if

$$\lim_{x \rightarrow \infty} \frac{\mathbb{B}((0, x])}{x} \rightarrow \frac{1}{\mu} \quad a.s. \quad (2.12)$$

We know that \mathbb{B} is given by the stationary random walk $(B_i)_{i \in \mathbb{Z}}$. So it should be enough to show this property for a usual random walk started at 0, because the extension of the random walk to both directions as well as the random shifting as described in Remark 1.43 should not change the behavior of the process in the limit. Thus, we regard $S_0 := 0$ and $S_n := \sum_{i=1}^n X_i$ with X_1, X_2, \dots i.i.d. $\mathcal{N}(\mu, \sigma_1^2)$, $\mu, \sigma_1 \in (0, \infty)$. Define the counting function $N(t) := \#\{n \in \mathbb{N} : S_n \in (0, t]\}$, then we have to show that

$$\lim_{t \rightarrow \infty} N(t)/t = 1/\mu \quad a.s. \quad (2.13)$$

Because of the following relationship

$$S_{\tau_t-1} \leq t \leq S_{\tau_t} \quad (2.14)$$

we have

$$\frac{\tau_t - 1}{N(t)} \frac{S_{\tau_t-1}}{\tau_t - 1} \leq \frac{t}{N(t)} \leq \frac{S_{\tau_t}}{N(t)} \leq \frac{S_{\tau_t-1}}{\tau_t - 1} + \frac{X_{\tau_t}}{\tau_t - 1} \quad (2.15)$$

Because of the strong law of large numbers and $\lim_{t \rightarrow \infty} \tau_t = \infty$ a.s., $S_{\tau_t-1}/(\tau_t - 1)$ converges almost surely against μ . Finally, because $(\tau_t - 1)/N(t) \rightarrow 1$ a.s. (for $t \rightarrow \infty$) and $X_{\tau_t}/(\tau_t - 1) \rightarrow 0$ a.s. (for $t \rightarrow \infty$), we get $\mu \leq t/N(t) \leq \mu$. \square

2.2.1.1. Distribution of B_1

Given the stationary BR \mathbb{B} , the distribution of B_1 is determined. Because \mathbb{B} is basically a stationary random walk $(S'_n)_{n \in \mathbb{Z}}$ and S'_1 is given by (1.35), there is the relation

$$B_1 \stackrel{d}{=} S'_1 \stackrel{d}{=} R_\infty. \quad (2.16)$$

Thus, the distribution of B_1 equals the distribution of the asymptotic residual waiting time which is given by Proposition 1.41. According to equation (1.31), the ladder height distribution S_{τ_0} (Feller, 1966) of the underlying random walk is needed to compute the distribution H_∞ .

An alternative way is to compute H_∞ via its Laplace transform \mathcal{H}_∞ which is given by

$$\mathcal{H}_\infty(\alpha) = \int_0^\infty e^{-\alpha x} H_\infty(dx), \quad \alpha \geq 0. \quad (2.17)$$

Still in the setting of Proposition 1.41 on page 16, Woodroffe (1982) gives the following expression.

Corollary 2.6 (Woodroffe, 1982, p.24):

$$\mathcal{H}_\infty(\alpha) = \frac{1}{\alpha\nu} \exp[-B(\alpha)], \quad \alpha > 0, \quad (2.18)$$

where

$$B(\alpha) = \sum_{k=1}^\infty \frac{1}{k} \mathbb{E} \left[e^{-\alpha S_k^+} \right], \quad \alpha > 0, \quad (2.19)$$

with $S_k^+ := \max(\{0, S_k\}) \in [0, \infty)$.

The connection of Corollary 2.6 to the GLO model is given by assuming $\mathcal{N}(\mu, \sigma_1^2)$ as the increment distribution of the random walk, then the expression $\mathbb{E} \left[e^{-\alpha S_k^+} \right]$ can be derived.

Lemma 2.7: Let $S_k = \sum_{i=1}^k X_i$, with X_i i.i.d. $\mathcal{N}(\mu, \sigma_1^2)$, $\mu > 0$ and $\sigma_1^2 > 0$, then

$$\mathbb{E} \left[e^{-\alpha S_k^+} \right] = F_{\{0,1\}} \left(-\frac{\mu}{\sigma_1} \sqrt{k} \right) + \exp \left(\left[\frac{\sigma_1^2 \alpha^2}{2} - \mu \alpha \right] k \right) F_{\{0,1\}} \left(- \left[\frac{\alpha \sigma_1^2 - \mu}{\sigma_1} \right] \sqrt{k} \right). \quad (2.20)$$

Proof. We have $S_k \sim \mathcal{N}(k\mu, k\sigma_1^2)$ with c.d.f. $F_{\{k\mu, k\sigma_1^2\}}(x)$. For $t \geq 0$,

$$F_{S_k^+}(t) := \mathbb{P}(S_k^+ \leq t) = \mathbb{P}(S_k^+ = 0) + \mathbb{P}(S_k^+ \in (0, t]) \quad (2.21)$$

$$= \mathbb{P}(S_k \leq 0) + \mathbb{P}(S_k \in (0, t]) \quad (2.22)$$

$$= F_{\{0,1\}} \left(\frac{t}{\sigma_1 \sqrt{k}} - \sqrt{k} \frac{\mu}{\sigma_1} \right) \quad (2.23)$$

and $\mathbb{P}(S_k^+ \leq t) = 0$, for $t < 0$. Then we have

$$\mathbb{E} \left[e^{-\alpha S_k^+} \right] = \int_{\mathbb{R}} e^{-\alpha x} dF_{S_k^+}(x) \quad (2.24)$$

$$= F_{S_k^+}(0) + \int_{(0, \infty)} e^{-\alpha x} dF_{S_k^+}(x) \quad (2.25)$$

$$= F_{\{0,1\}} \left(-\sqrt{k} \frac{\mu}{\sigma_1} \right) + \int_0^\infty e^{-\alpha x} \frac{1}{\sigma_1 \sqrt{k}} \varphi_{\{0,1\}} \left(\frac{x}{\sigma_1 \sqrt{k}} - \sqrt{k} \frac{\mu}{\sigma_1} \right) dx. \quad (2.26)$$

It is

$$\exp(-\alpha x) \exp \left(\frac{1}{2} \left(\frac{x}{\sigma_1 \sqrt{k}} - \sqrt{k} \frac{\mu}{\sigma_1} \right)^2 \right) = \exp \left(-\frac{x}{\sigma_1 \sqrt{k}} + \left(\frac{\mu}{\sigma_1^2} - \alpha \right)^2 k \sigma_1^2 - \frac{k \mu^2}{\sigma_1^2} \right) \quad (2.27)$$

and

$$c := \int_0^\infty e^{-\alpha x} \frac{1}{\sigma_1 \sqrt{k}} \varphi_{\{0,1\}} \left(\frac{x}{\sigma_1 \sqrt{k}} - \sqrt{k} \frac{\mu}{\sigma_1} \right) dx. \quad (2.28)$$

So that

$$c = \frac{1}{\sigma_1 \sqrt{k}} \frac{1}{\sqrt{2\pi}} \exp \left(\left(\left(\frac{\mu}{\sigma_1^2} - \alpha \right)^2 k \sigma_1^2 - \frac{k \mu^2}{\sigma_1^2} \right) / 2 \right) \quad (2.29)$$

$$\cdot \int_0^\infty \exp \left(-\left(\frac{x}{\sigma_1 \sqrt{k}} - \left(\frac{\mu}{\sigma_1^2} - \alpha \right) \sqrt{k} \sigma_1 \right)^2 / 2 \right) dx \quad (2.30)$$

$$= \exp \left(\left[\frac{\sigma_1^2 \alpha^2}{2} - \mu \alpha \right] k \right) \int_0^\infty \frac{1}{\sigma_1 \sqrt{k}} \varphi_{\{0,1\}} \left(\frac{x}{\sigma_1 \sqrt{k}} - \left(\frac{\mu}{\sigma_1^2} - \alpha \right) \sqrt{k} \sigma_1 \right) dx \quad (2.31)$$

$$= \exp \left(\left[\frac{\sigma_1^2 \alpha^2}{2} - \mu \alpha \right] k \right) F_{\{0,1\}} \left(-\left(\frac{\alpha \sigma_1^2 - \mu}{\sigma_1} \right) \sqrt{k} \right). \quad (2.32)$$

Finally, putting (2.32) and (2.28) together into (2.26) ends the proof. \square

The distribution of B_1 can be derived by inverting the Laplace transform $\mathcal{H}_\infty(\alpha)$. Obviously this procedure contains some effort, but fortunately the distribution of B_1 is not really needed to describe the firing patterns of a spike train in experimental practice. Furthermore, simulations can be used to approximate the distribution (see Section 2.3.4).

2.2.1.2. Inversion probability of BR beats

Let $(B_i)_{i \in \mathbb{Z}}$ be a stationary random walk as defined in Section 2.1. Due to the fact that the normal distribution has values on \mathbb{R} , steps to the left may occur (or downward steps – depending on the point of view). This we call an inversion of beats and it is an event of the kind $\{B_j < B_i\}$ for integers $j > i$ (compare Figure 2.3). In the following, we focus on beats B_i with $i \in \mathbb{N}$, because in this case, the increments $B_{i+1} - B_i$ are independent. Furthermore, because of the stationarity, the events $\{B_{j+1} < B_1\}$ and $\{B_{j+1+k} < B_{1+k}\}$ have the same probability for all $j, k \in \mathbb{N}$. Thus, we focus on events of the kind $\{B_j < B_1\}$ for $j > 1$. The next lemmas represent some simple results for the inversion probability.

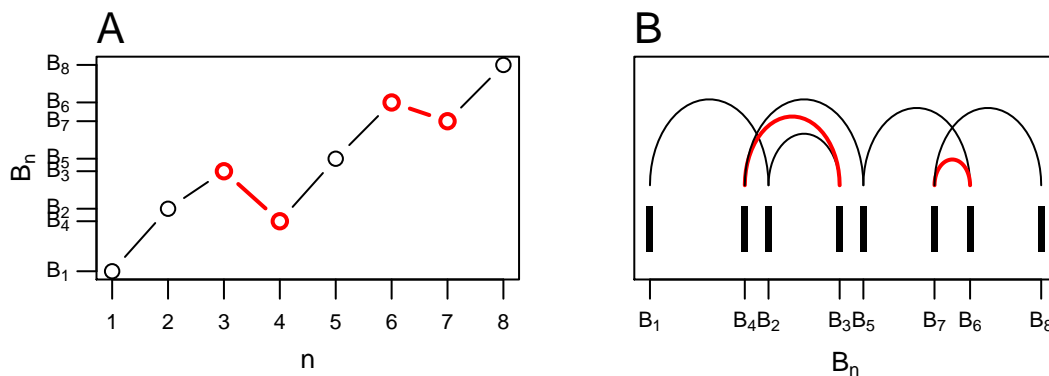


Figure 2.3: Visualization of an inversion of beats, in the context of a random walk (A) and for the induced point process (B). Red lines indicate events of the kind $\{B_j < B_i\}$ for integers $j > i$.

Lemma 2.8 (Inversion probability for one pair of beats): Let $k \in \mathbb{N}$, then

$$\mathbb{P}(B_{1+k} < B_1) = F_{\{0,1\}}(-\mu\sqrt{k}/\sigma_1). \quad (2.33)$$

Proof. Because of $B_{1+k} - B_1 \sim \mathcal{N}(k\mu, k\sigma_1^2)$, the rest follows immediately. \square

Remark 2.9: For $k \in \mathbb{N}$, we have that

$$\lim_{\sigma_1 \rightarrow \infty} \mathbb{P}(B_{1+k} < B_1) = F_{\{0,1\}}(0) = 0.5. \quad (2.34)$$

Lemma 2.10: If $j, k \in \mathbb{N}$ with $j < k$,

$$\mathbb{P}(B_{1+j} < B_1) > \mathbb{P}(B_{1+k} < B_1). \quad (2.35)$$

Proof. $\mathbb{P}(B_{1+j} < B_1) > \mathbb{P}(B_{1+k} < B_1) \Leftrightarrow F_{\{0,1\}}(-\mu\sqrt{j}/\sigma_1) > F_{\{0,1\}}(-\mu\sqrt{k}/\sigma_1) \Leftrightarrow \sqrt{j} < \sqrt{k}$. \square

Thus, pairs of neighbor beats (concerning the index) have the largest inversion probability. Equations (2.35) and (2.33) consider only one pair of beats. For taking all possible pairs into account there is the following upper boundary.

Lemma 2.11 (Upper bound for the inversion probability of beats):

$$\mathbb{P}(\{\exists j \in \mathbb{N} : B_{1+j} < B_1\}) \leq \sum_{k=0}^{\infty} F_{\{0,1\}}(-\mu\sqrt{k}/\sigma_1). \quad (2.36)$$

Proof. $\mathbb{P}(\{\exists j \in \mathbb{N} : B_{1+j} < B_1\}) = \mathbb{P}(\cup_{j=1}^{\infty} \{B_{1+j} < B_1\}) \leq \sum_{j=1}^{\infty} \mathbb{P}(B_{1+j} < B_1)$. \square

Lemma 2.12 (Relationship between inversion events): For $k \in \mathbb{N}$ and the event $\{B_{1+k} < B_1\}$ there is a $l \in \mathbb{N}$ with $l \leq k$, such that

$$\{B_{l+1} < B_l\} \subseteq \{B_{1+k} < B_1\}. \quad (2.37)$$

Furthermore, we have

$$\{B_{1+k} < B_1\} \subset \{\exists j \in \mathbb{N} : B_{1+j} < B_1\}. \quad (2.38)$$

Proof. The second inclusion follows because

$$\{B_{1+k} < B_1\} \subset \cup_{j=1}^{\infty} \{B_{1+j} < B_1\} = \{\exists j \in \mathbb{N} : B_{1+j} < B_1\}. \quad (2.39)$$

The equality in the first inclusion is given if $k = 1$ and because of

$$\{B_{1+k} < B_1\} \implies \cup_{i=1}^k \{B_{1+i} < B_i\} \quad (2.40)$$

the left inclusion follows. \square

The probability of the event $\{\exists j \in \mathbb{N} : B_{1+j} < B_1\}$ can be computed for some parameter constellations. For this purpose, let us assume a random walk $(S_n)_{n \in \mathbb{N}}$ with $S_0 := 0$ and normally distributed increments with mean $-\beta < 0$ and variance 1. Let $\zeta(z)$ denote the Riemann Zeta function, then Janssen and van Leeuwaarden (2007) give the following proposition (which is a theorem of Chang and Peres (1997)) for the probability that the maximum M of this gaussian random walk is zero ($M := \max\{S_n : n \in \mathbb{N}_0\}$).

Proposition 2.13 (see Chang and Peres (1997)): The probability that the maximum of the Gaussian random walk is zero satisfies

$$\mathbb{P}(M = 0) = \sqrt{2}\beta \exp \left\{ \frac{\beta}{\sqrt{2\pi}} \sum_{r=0}^{\infty} \frac{\zeta(1/2 - r)}{r!(2r + 1)} \left(\frac{-\beta^2}{2} \right)^r \right\} \quad (2.41)$$

for $0 < \beta < 2\sqrt{\pi}$.

Corollary 2.14 (Exact inversion probability of beats): The inversion probability is given by $\mathbb{P}(\{\exists j \in \mathbb{N} : B_{1+j} < B_1\}) = 1 - \mathbb{P}(M = 0)$ with β in equation (2.41) given by $\beta = \mu/\sigma_1$ and $0 < \mu/\sigma_1 < 2\sqrt{\pi}$.

Proof. The event to get an inversion of beats is the same as the event that the underlying random walk with mean increment μ is started at zero and walks below zero, i.e.

$$\{\exists j \in \mathbb{N} : B_{1+j} < B_1\} = \{S_n < 0 : S_0 = 0, n > 0\}. \quad (2.42)$$

But this is the same as regarding the event that the random walk is started at zero, has a mean increment of $-\mu$ and walks above zero. Because this is equal to the following events

$$\{M > 0\} = \{M \leq 0\}^C = \{M = 0\}^C \quad (2.43)$$

Proposition 2.13 can be applied by rescaling the random walk to have an increment variance of 1. In this case, also the mean has to be scaled to μ/σ_1 . \square

Example 2.15 (Inversion probabilities): Table 2.1 shows the inversion probabilities for different values of μ/σ_1 . They are computed for neighbor beats, for the upper boundary (2.36) and by using $1 - \mathbb{P}(M = 0)$ of equation (2.41). Obviously the inversion probability will decrease, if the fraction μ/σ_1 is increased and vice versa. So under the assumption that $\mu \gg \sigma_1$, the possibility of an inversion is ruled out. Furthermore the use of $\mathbb{P}(B_2 < B_1)$ as a simplified inversion probability seems to be adequate for ratios $\mu/\sigma_1 \geq 2$, because in these cases they are close to the upper boundary. But for $\mu/\sigma_1 < 2$ the boundary gets imprecise, as one can see by comparison with the true inversion probability given by $1 - \mathbb{P}(M = 0)$. However, we want to describe parameter constellations for which the inversion probability is smaller than a (small) ε and for which it is larger. So the use of $\mathbb{P}(B_2 < B_1)$ seems to be appropriate.

| μ/σ_1 | 0.1 | 1 | 2 | 3 | 5 | 7 | 10 |
|-------------------------|-------|-------|-------|-------|---------------------|----------------------|----------------------|
| $\mathbb{P}(B_2 < B_1)$ | 0.460 | 0.159 | 0.023 | 0.001 | $2.9 \cdot 10^{-7}$ | $1.3 \cdot 10^{-12}$ | $7.6 \cdot 10^{-24}$ |
| $1 - \mathbb{P}(M = 0)$ | 0.867 | 0.199 | 0.024 | 0.001 | - | - | - |
| bound (2.36) | 49.76 | 0.331 | 0.025 | 0.001 | $2.9 \cdot 10^{-7}$ | $1.3 \cdot 10^{-12}$ | $7.6 \cdot 10^{-24}$ |

Table 2.1: Inversion probabilities of beats for different ratios μ/σ_1 . For increasing μ/σ_1 the inversion probability is decreasing and for $\mu/\sigma_1 \geq 2$, the probability $\mathbb{P}(B_2 < B_1)$ is very close to the upper boundary.

2.2.2. Stationarity, simpleness and mean density

The GLO process Φ builds upon the stationary point process \mathbb{B} whose points are independently copied, thinned and shifted. Thus, Φ will also be stationary and simple.

Proposition 2.16: $\Phi \sim \text{GLO}(\psi)$ is a stationary point process.

Proof. For each B_i there is an independent P_i which either deletes the point B_i (if $P_i = 0$) or replaces it by P_i copies. The resulting process after the independent copying and thinning of points is given by

$$\mathbb{B}_P = \sum_{i \in \mathbb{Z}} P_i \delta_{B_i}. \quad (2.44)$$

The stationarity of the background rhythm \mathbb{B} means

$$\sum_{i \in \mathbb{Z}} \delta_{B_i} \stackrel{d}{=} \sum_{i \in \mathbb{Z}} \delta_{B_i+x} \quad \forall x \in \mathbb{R}. \quad (2.45)$$

It follows that \mathbb{B}_P is also stationary, because

$$\mathbb{B}_P = \sum_{i \in \mathbb{Z}} P_i \delta_{B_i} \stackrel{d}{=} \sum_{i \in \mathbb{Z}} P_i \delta_{B_i+x} \quad \forall x \in \mathbb{R}. \quad (2.46)$$

Now, let the point $B_{i,j(i)}$ represent the j -th copy of the point B_i , such that the indices are given by $i \in \{k \in \mathbb{Z} : P_k > 0\}$ and $j(i) = 1, \dots, P_i$. Assume that each point $B_{i,j(i)}$ is independently and randomly shifted through a random distance $Z_{i,j}$, where all random variables $\{Z_{i,j} : i \in \mathbb{Z}, j \in \mathbb{N}\}$ are i.i.d. $\mathcal{N}(0, \sigma_2^2)$ with common distribution function $F_{\{0, \sigma_2^2\}}$. The resulting process is Φ as defined in (2.9). The respective p.g.f.s $G_{\mathbb{B}_P}$ and G_{Φ} of the shifted process are related by

$$G_{\Phi}[h(\cdot)] = G_{\mathbb{B}_P} \left[\int_{\mathbb{R}} h(y) F_{\{0, \sigma_2^2\}}(dy - \cdot) \right]. \quad (2.47)$$

Because \mathbb{B}_P is itself stationary, we have $G_{\mathbb{B}_P}[h(\cdot)] = G_{\mathbb{B}_P}[h(\cdot - u)]$ for all $u \in \mathbb{R}$ and $h \in \mathcal{V}(\mathbb{R})$. The right-hand side of the expression for $G_{\Phi}[\cdot]$ then equals

$$G_{\mathbb{B}_P} \left[\int_{\mathbb{R}} h(y - u) F_{\{0, \sigma_2^2\}}(dy - \cdot) \right] = G_{\Phi}[h(\cdot - u)]. \quad (2.48)$$

So that according to Proposition 1.24, the transformed process Φ (the final GLO process) is again stationary. \square

Proposition 2.17: $\Phi \sim \text{GLO}(\psi)$ is a simple point process.

Proof. By definition, the point process Φ is simple if $\Phi(\{x\}) \leq 1, \forall x \in \mathbb{R}$ with probability one. Because of (2.9) this is fulfilled if

$$\mathbb{P}(B_{i_1} + Z_{i_1, k_1} \neq B_{i_2} + Z_{i_2, k_2}, \forall (i_1, i_2, k_1, k_2) \in A) = 1 \quad (2.49)$$

with the set A given as

$$A := \{(i_1, i_2, k_1, k_2) : i_1, i_2 \in \mathbb{Z}, k_1 \in \{1, \dots, P_{i_1}\}, k_2 \in \{1, \dots, P_{i_2}\}, i_1 \neq i_2 \wedge k_1 \neq k_2\}.$$

Again because of (2.9), we have countably many random variables in (2.49). These random variables are the random spike times $\{T_0, T_{\pm 1}, \dots\}$ so (2.49) can be rewritten as

$$\mathbb{P}(\{T_i \neq T_j, \forall i, j \in \mathbb{Z}, i \neq j\}) = 1 - \mathbb{P}\left(\bigcup_{\substack{i, j \in \mathbb{Z} \\ i \neq j}} \{T_i = T_j\}\right) \quad (2.50)$$

$$\geq 1 - \sum_{\substack{i, j \in \mathbb{Z} \\ i \neq j}} \mathbb{P}(\{T_i = T_j\}) \quad (2.51)$$

$$= 1 - 0 = 1 \quad (2.52)$$

The last step follows, because $B_{i_1} + Z_{i_1, k_1}$ are \mathbb{R} -valued without atoms. So Φ is simple. \square

The mean density of the process is given in the following proposition.

Proposition 2.18: If $\Phi \sim \text{GLO}(\psi)$, then $\lambda := \mathbb{E}[\Phi(0, 1]] = \gamma/\mu$.

Proof. The mean density λ equals the intensity of the stationary point process. Because Φ is simple, it follows from Korolyuk's theorem (Proposition 1.28) that $\lambda = \mathbb{E}[\Phi(0, 1]]$. We have

$$\lambda = \mathbb{E}[\Phi(0, 1]] = \mathbb{E}[\mathbb{B}_P(0, 1]] = \mathbb{E}[P_i] \mathbb{E}[\mathbb{B}(0, 1]] = \frac{\gamma}{\mu}. \quad (2.53)$$

The last step follows of Lemma 2.4. \square

2.2.3. Process characterization as a Cluster and Cox process

Definition 2.19 (Cluster process): Φ is a *cluster process* on the c.s.m.s. E_1 , with *centre process* Φ_c on the c.s.m.s. E_2 and *component processes* which is the measurable

family of point processes $\{\Phi(\cdot | y) : y \in E_2\}$, when for every bounded $A \in \mathcal{B}(E_1)$,

$$\Phi(A) = \int_{E_2} \Phi(A | y) \Phi_c(dy) = \sum_{y_i \in \Phi_c(\cdot)} \Phi(A | y_i) < \infty \quad a.s. \quad (2.54)$$

If $\{\Phi(\cdot | y) : y \in E_2\}$ is an independent family of component processes, Φ is called an *independent cluster process*.

Lemma 2.20 (see Daley and Vere-Jones, 1988): The stationary cluster process Φ exists if the centre process Φ_c is stationary, the distribution of the cluster members depends only on their positions relative to the cluster centre and the mean cluster size is finite.

Proof. Follows directly of definition 2.19. □

Lemma 2.21: $\Phi \sim \text{GLO}(\psi)$ is a stationary cluster process.

Proof. In the GLO case, $E_1 = E_2 = \mathbb{R}$, $\Phi_c = \mathbb{B}$ (stationary) and the component processes are given by

$$\Phi(\cdot | y_i) = \sum_{j=1}^{P_i^{(m)}} \delta_{y_i + Z_{i,j}} \quad (2.55)$$

with $y_i \in \mathbb{B}(\cdot)$. It is

$$\Phi(A) = \sum_{y_i \in \mathbb{B}(\cdot)} \Phi(A | y_i) < \infty \quad a.s. \quad (2.56)$$

Thus, according to Lemma 2.20 and definition 2.19 the GLO process Φ is a stationary cluster process. □

Proposition 2.22 (see Daley and Vere-Jones, 1988, p. 347): A stationary cluster process is ergodic whenever the cluster centre process has the same property.

Corollary 2.23: $\Phi \sim \text{GLO}(\psi)$ is ergodic.

Proof. Because of Lemma 2.21, we know that Φ is a stationary Cluster process, and because the cluster centre process \mathbb{B} is ergodic (Proposition 2.5), Φ has the same property (Proposition 2.22). □

Furthermore, if $\Phi \sim \text{GLO}(1)$ (bursty process, $m = 1$), it will turn out that Φ is a Cox process. In the GLO case, this means that conditional on \mathbb{B} which determines the random firing intensity of the process, Φ can be considered a time-inhomogeneous Poisson process (see definition 1.32) or more specifically, a generalized shot noise Cox process (Møller and Torrisi, 2005).

Definition 2.24 (Cox process): Let Λ be a random measure on E . A point process Φ on E is a *Cox process* directed by Λ if, conditional on Λ , Φ is a Poisson process $\Phi(\cdot | \Lambda)$ on E with intensity measure Λ (see definition 1.32 and Remark 1.33).

Proposition 2.25: $\Phi \sim \text{GLO}(\mu, \sigma_1, \sigma_2, \gamma, 1)$ is a Cox process with random intensity

$$\rho_{\mathbb{B}}(t) = \gamma \sum_{j \in \mathbb{Z}} \varphi_{\{B_j, \sigma_2^2\}}(t) \quad (2.57)$$

and random intensity measure

$$\Lambda(A) = \int_A \rho_{\mathbb{B}}(t) dt, \quad A \in \mathcal{B}. \quad (2.58)$$

Proof. The GLO process Φ is decomposed in subprocesses

$$\Phi_j = \sum_{k=1}^{P_j^{(1)}} \delta_{B_j + Z_{j,k}} \quad (2.59)$$

where Φ_j is the point process consisting of spikes only generated of beat B_j . Thus,

$$\Phi = \sum_{j \in \mathbb{Z}} \Phi_j. \quad (2.60)$$

Consequently, conditioned on the background rhythm \mathbb{B} , the conditioned process can be written as

$$\Phi | \mathbb{B} = \sum_{j \in \mathbb{Z}} \Phi_j | \mathbb{B}. \quad (2.61)$$

Now, we show that $\Phi_j | \mathbb{B}$ is an inhomogeneous Poisson process. For this purpose, let $(a_1, b_1), (a_2, b_2), \dots, (a_k, b_k)$ be disjoint intervals and $\Lambda_i = \int_{a_i}^{b_i} \varphi_{\{B_j, \sigma_2^2\}}(t) dt$ equal the probability that one spike of beat B_j falls into the interval (a_i, b_i) and $n_i \in \mathbb{N}_0$ the number of spikes in that interval. Then

$$\begin{aligned} & \mathbb{P}(\Phi_j(a_i, b_i) = n_i, i = 1, \dots, k | \mathbb{B}) \\ &= \sum_{n=0}^{\infty} \mathbb{P}(\Phi_j(a_i, b_i) = n_i, i = 1, \dots, k | P_j = n, \mathbb{B}) \mathbb{P}(P_j = n | \mathbb{B}). \end{aligned} \quad (2.62)$$

Let $A := \mathbb{R} \setminus \bigcup_{i=1}^k (a_i, b_i)$, then

$$\begin{aligned} & \mathbb{P}(\Phi_j(a_i, b_i) = n_i, i = 1, \dots, k | P_j = n, \mathbb{B}) \\ &= \mathbb{P}\left(\Phi_j(a_1, b_1) = n_1, \dots, \Phi_j(a_k, b_k) = n_k, \Phi_j(A) = n - \sum_{i=1}^k n_i \mid P_j = n, \mathbb{B}\right). \end{aligned} \quad (2.63)$$

Obviously, (2.63) is zero for $\sum_{i=1}^k n_i > n$, so we concentrate on the case $\sum_{i=1}^k n_i \leq n$. Using the multinomial distribution, equation (2.62) equals

$$\sum_{n=0}^{\infty} \binom{n}{n_1, \dots, n_k, n - \sum_{i=1}^k n_i} e^{-\gamma} \frac{\gamma^n}{n!} \left(1 - \sum_{i=1}^k \Lambda_i\right)^{n - \sum_{i=1}^k n_i} \prod_{i=1}^k \Lambda_i^{n_i} \quad (2.64)$$

$$= e^{-\gamma} \prod_{i=1}^k \left(\frac{\Lambda_i^{n_i}}{n_i!}\right) \sum_{j=0}^{\infty} \frac{\gamma^{j + \sum_{i=1}^k n_i}}{j!} \left(1 - \sum_{i=1}^k \Lambda_i\right)^j \quad (2.65)$$

$$= \prod_{i=1}^k e^{-\gamma \Lambda_i} \frac{(\gamma \Lambda_i)^{n_i}}{n_i!}. \quad (2.66)$$

Thus $\Phi_j|\mathbb{B}$ is an inhomogeneous Poisson process with intensity function $\gamma \varphi_{\{B_j, \sigma_2^2\}}(t)$.

Kingman (1993) describes in its Superposition Theorem that the superposition of countably many independent Poisson processes is again a Poisson process (see Proposition 1.34). Furthermore, it is known that the superposition of independent inhomogeneous Poisson processes is again an inhomogeneous Poisson process. This directly applies to $\Phi|\mathbb{B}$, because all subprocesses $\Phi_j|\mathbb{B}$ are independent of each other. Thus Φ is a Cox process with random intensity (2.57) and random intensity measure (2.58). \square

However, Proposition 2.25 does not hold for the non-bursty mode ($m = 0$) in which the number of spikes is not Poisson distributed. But as we will see in the next subsection, there is a convergence of Φ to a homogenous Poisson process for both firing modes ($m = 0$ and $m = 1$) and increasing σ_2 .

2.2.4. Convergence of GLO process

When σ_2 increases and is getting large, the spikes of a beat are spread on \mathbb{R} . Furthermore, the random spike intensity, which is given at time point t by

$$f(t) = \frac{\gamma}{\sigma_2 \sqrt{2\pi}} \sum_{i \in \mathbb{Z}} \exp\left\{-\frac{(t - B_i)^2}{2\sigma_2^2}\right\}, \quad (2.67)$$

is turning into a constant, because

$$\lim_{\sigma_2 \rightarrow \infty} \frac{df(t)}{dt} = \lim_{\sigma_2 \rightarrow \infty} \frac{\gamma}{\sigma_2 \sqrt{2\pi}} \sum_{i \in \mathbb{Z}} -\frac{(t - B_i)}{\sigma_2^2} \exp\left\{-\frac{(t - B_i)^2}{2\sigma_2^2}\right\} = 0. \quad (2.68)$$

Thus, it is natural to expect that Φ turns for increasing σ_2 , regardless of the firing mode m , into a homogenous Poisson process. This can be proved by using a theorem of Daley and Vere-Jones (1988) which is a result of Dobrushin (1956) about the convergence of point processes after repeated independent translations of the points and has been further investigated by Stone (1968).

Definition 2.26 (Invariant σ -algebra): Let $(\Omega, \mathcal{A}, \varphi)$ be a measure space and \mathcal{T} a measure-preserving operator on this space; that is $\varphi(\mathcal{T}^{-1}A) = \varphi(A)$ for $A \in \mathcal{A}$. The σ -algebra \mathcal{A}_I of invariant events under \mathcal{T} , contains those sets $A \in \mathcal{A}$ for which $\varphi(\mathcal{T}^{-1}A\Delta A) = 0$, where $A\Delta B := (A \setminus B) \cup (B \setminus A)$.

Proposition 2.27 (see Daley and Vere-Jones, 1988, p. 342): A stationary random measure or point process is ergodic if and only if it is metrically transitive; that is the invariant σ -algebra \mathcal{A}_I is trivial.

Proposition 2.28 (see Daley and Vere-Jones, 1988, p. 305): Let Φ_0 be a second-order stationary point process on $E = \mathbb{R}^d$ and ν a distribution on \mathbb{R}^d that is nonlattice. Then the sequence of point processes $\{\Phi_n\}$, derived from Φ_0 by successive random translations according to ν , converges weakly to the stationary mixed Poisson process with p.g.fl.

$$G[h] = \mathbb{E} \left[\exp \left(-Y \int_E [1 - h(x)] dx \right) \right], \quad (2.69)$$

where Y is given by

$$Y = \mathbb{E} \left[\Phi_0(\mathbb{U}^d) \mid \mathcal{A}_I \right], \quad (2.70)$$

with \mathbb{U}^d the unit cube in \mathbb{R}^d and \mathcal{A}_I the invariant σ -algebra.

Remark 2.29: In Proposition 2.28, the invariant σ -algebra \mathcal{A}_I is given under the operator \mathcal{T}_u , $u \in \mathbb{R}$, where \mathcal{T}_u is the shift operator given by

$$(\mathcal{T}_u \varphi)(A) := \varphi(\mathcal{T}_u A) = \varphi(\{x + u : x \in A\}). \quad (2.71)$$

In the case of a stationary process Φ , \mathcal{T}_u is a measure-preserving operator.

Proposition 2.30: Let $\Phi \sim \text{GLO}(\mu, \sigma_1, \sigma_2, \gamma, m)$. If σ_2 tends to infinity, Φ converges weakly to a homogenous Poisson process with rate $\lambda = \gamma/\mu$.

Proof. This is simply an application of Proposition 2.28 with

$$\Phi_0 := \Phi = \sum_{i \in \mathbb{Z}} \sum_{k=1}^{P_i^{(m)}} \delta_{B_i + Z_{i,k}}. \quad (2.72)$$

First of all, we show that the GLO process Φ_0 is second order stationary, which by definition 1.25 means that Φ_0 is stationary and $\mathbb{E}[\Phi_0(A_1)\Phi_0(A_2)] < \infty$ for all bounded Borel sets A_1 and A_2 . The stationarity of the GLO process has been shown in Proposition 2.16. Because we are having a countable number of spikes in (2.9) we can rewrite the GLO process in terms of spike times $\Phi_0 = \sum_{i \in \mathbb{Z}} \delta_{T_i}$. It is

$$\mathbb{E}[\Phi_0(A_1)\Phi_0(A_2)] = \mathbb{E} \left[\sum_{i \in \mathbb{Z}} \delta_{T_i}(A_1) \sum_{j \in \mathbb{Z}} \delta_{T_j}(A_2) \right] \quad (2.73)$$

$$= \sum_{i, j \in \mathbb{Z}} \mathbb{E}[\delta_{T_i}(A_1)\delta_{T_j}(A_2)] \quad (2.74)$$

$$= \sum_{i, j \in \mathbb{Z}} \mathbb{P}(T_i \in A_1, T_j \in A_2) < \infty. \quad (2.75)$$

The last step follows, because the probability to fall into the sets A_1 and A_2 gets exponential small for increasing i and j .

Furthermore, let $\nu = \varphi_{\{0, \sigma_2^2\}}$ which is nonlattice. Then n successive translations of the points are given by randomly shifting the original points according to the distribution $\nu^{*n} = \varphi_{\{0, n\sigma_2^2\}}$. Thus the process after n translations can be written as

$$\Phi_n = \sum_{i \in \mathbb{Z}} \sum_{k=1}^{P_i^{(m)}} \delta_{B_i + Z_{i,k}^{[n]}} \quad \text{with} \quad Z_{i,k}^{[n]} \sim \mathcal{N}(0, (n+1)\sigma_2^2). \quad (2.76)$$

In this setting, Proposition 2.28 can be applied and so for increasing n (which represents an increasing σ_2) Φ_n converges weakly to the stationary mixed Poisson process with p.g.f. (2.69). Because Φ_n is ergodic, the invariant σ -algebra \mathcal{A}_I is trivial (Proposition 2.27). Thus, we have

$$Y = \mathbb{E} [\Phi_0(\mathbb{U}^d) \mid \mathcal{A}_I] = \mathbb{E} [\Phi_0(\mathbb{U}^d)] = \mathbb{E} [\Phi_0((0, 1])] = \gamma/\mu \quad (2.77)$$

and thus Φ_n converges weakly to the homogeneous Poisson process with rate γ/μ . \square

2.3. Spike train simulation

With increasing computer power, simulation procedures for complicated models become more and more popular. The term simulation is very general and in the literature there are different distinctions for stochastic simulation, simulation, Monte Carlo method and Monte Carlo simulation (Ripley, 1987; Ross, 1991; Gamerman, 1997). But these

distinctions are not always clear cut. For that reason, we just use the term 'simulation' meaning that we are using random numbers to generate one or more realization of the GLO process which can be used for quantifying parameter estimation (Chapter 5.2) or for statistical inference (Chapter 6.2).

The GLO model allows the efficient simulation of spike trains with parameter vector ψ and recording time \mathcal{T} . In the progress of this work, we make repeated use of simulated spike trains according to GLO assumptions. For that reason, the procedure is described in detail (Section 2.3.1) and two different implementations in the free statistic software R Chambers (2008) are compared with each other (Section 2.3.2). Finally, a rule of thumb is given for \mathcal{T} to have at least n spikes in the interval $[0, \mathcal{T}]$ (Section 2.3.3) and the chapter is ended with some exemplary simulations (Section 2.3.4).

2.3.1. Simulation procedure

We are interested in the simulation of a stationary GLO spike train $\mathcal{S} = (t_1, \dots, t_n)$ with parameters $\psi = (\mu, \sigma_1, \sigma_2, \gamma, m)$ for the interval $[0, \mathcal{T}]$, where $\mathcal{T} > 0$ is determined by the investigator. For this purpose, we make use of a Markov Chain Monte Carlo procedure in order to generate a realization φ of the stationary point process Φ . In the following, the procedure is described in detail:

1. In a first step, the BR is modeled with a random random walk $(S_n)_{n \in \mathbb{N}_0}$ by setting $S_0 := -K_0\mu$ with $K_0 > 0$ and $S_{n+1} - S_n \sim \mathcal{N}(\mu, \sigma_1^2)$ for $n \in \mathbb{N}$. The random walk is ended when it reaches the height $K_2 \geq \mathcal{T}$ for the first time, i.e. after τ_{K_2} steps. The points S_n represent the beats of the BR and are renamed to

$$\begin{aligned} B_{-\tau_0+1} &:= S_0, & B_{-\tau_0} &:= S_1, & \dots, & B_{-1} &:= S_{\tau_0-2}, & B_0 &:= S_{\tau_0-1}, \\ B_1 &:= S_{\tau_0}, & B_2 &:= S_{\tau_0+1}, & \dots, & B_{\tau_{K_2}-\tau_0} &:= S_{\tau_{K_2}-1}. \end{aligned}$$

For large K_0 and K_2 the induced point process is approximately stationary on the interval $[0, \mathcal{T}]$.

2. In the next step, the number of spikes at each beat are generated and placed according to a normal distribution with variance σ_2^2 . One should note, that although we are only interested in the interval $[0, \mathcal{T}]$, there can be beats outside this interval which can place spikes into $[0, \mathcal{T}]$. Thus, it is important to take account of all beats in some interval $[-K_1, \mathcal{T} + K_1]$ with constant $K_1 > 0$ (or all beats in $[-K_0\mu, K_2]$, but this is an unnecessary large interval). Furthermore, the end of the process K_2 should be even larger than $\mathcal{T} + K_1$ because the random walk $(S_n)_{n \in \mathbb{N}_0}$ may

return below $\mathcal{T} + K_1$, but this is only relevant for very irregular GLO spike trains. Figure 2.4 shows a visualization of this procedure.

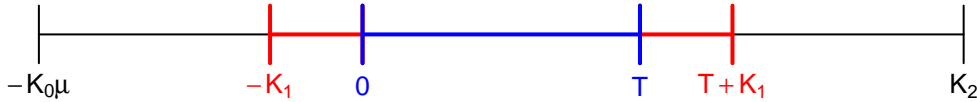


Figure 2.4: Visualization of the GLO simulation procedure. The GLO process is started at $-K_0\mu$ and ended at K_2 . All beats in the interval $[-K_1, \mathcal{T} + K_1]$ (red) have a noticeable probability to place spikes into the recording interval $[0, \mathcal{T}]$ (blue).

Remark 2.31 (Simulation parameters): Here are some guidelines for the choice of the simulation parameters K_0 , K_1 and K_2 .

The choice of K_0 : Larger values of K_0 improve the approximation of the stationary distribution on $[0, \mathcal{T}]$. In our case, we propose a value of

$$K_0 := \max\{1000, 5\sigma_1/\mu\}. \quad (2.78)$$

This should ensure that the process becomes virtually stationary beyond zero.

The choice of K_1 : Let us suppose that there is a beat at $\mathcal{T} + K_1$. How has K_1 to be chosen that a spike of this beat is only with small probability $\varepsilon_1 \in (0, 1)$ below \mathcal{T} ? Of course, this means that beats which are even larger than $\mathcal{T} + K_1$ have a much smaller probability for putting a spike below \mathcal{T} . So if $X \sim \mathcal{N}(\mathcal{T} + K_1, \sigma_2^2)$ is the distribution of this random spike, we want to have $\mathbb{P}(X \leq \mathcal{T}) \leq \varepsilon_1$. It follows that

$$K_1 \geq -q_{\varepsilon_1}\sigma_2, \quad (2.79)$$

where q_{ε_1} is the ε_1 -quantile of the standard normal distribution.

The choice of K_2 : In the context of ruin probabilities of random walks, there is the following lemma:

Lemma 2.32 (Ruin probability): Let X_1, X_2, \dots be i.i.d. and $b > 0$ with $\mathbb{E}[e^{-bX_1}] \leq 1$. Then for $a > 0$, we have $\mathbb{P}(a < \sum_{i=1}^n X_i \text{ for some } i \in \mathbb{N}) \leq e^{-ba}$.

This can be translated in our setting. If the random walk is started at $K_2 > \mathcal{T}$, how large should be the difference $a := K_2 - \mathcal{T}$, that the random walk jumps only with small probability ε below \mathcal{T} . This is the same situation as having a random walk started at 0 with mean increment $\mu < 0$ and looking for an $a > 0$, such that

the probability to cross a is small. Thus applying Lemma 2.32, we have X_1, X_2, \dots i.i.d. $\mathcal{N}(\mu, \sigma_1^2)$. Let $b > 0$ with

$$\mathbb{E}[e^{bX_1}] \leq 1 \tag{2.80}$$

$$\Leftrightarrow e^{\mu b + \sigma_1^2 b^2 / 2} \leq 1 \tag{2.81}$$

$$\Leftrightarrow \mu b + \sigma_1^2 b^2 / 2 \leq 0 \tag{2.82}$$

$$\Leftrightarrow b \leq -2\mu / \sigma_1^2. \tag{2.83}$$

Thus with $b = -2\mu / \sigma_1^2$, we have $\mathbb{P}(a < \sum_{i=1}^n X_i \text{ for some } i \in \mathbb{N}) \leq e^{2a\mu / \sigma_1^2}$. It follows that

$$e^{2a\mu / \sigma_1^2} \leq \varepsilon \tag{2.84}$$

$$\Leftrightarrow a \geq \frac{\sigma_1^2}{2\mu} \log \varepsilon \tag{2.85}$$

We set $K_2' := \mathcal{T} + \frac{\sigma_1^2}{2\mu} \log \varepsilon$ and

$$K_2 := \max\{K_2', \mathcal{T} + K_1\}. \tag{2.86}$$

Algorithm 1 summarizes the simulation procedure in a pseudo code and a R function called SIM can be found in appendix A.3.

2.3.2. Efficient implementation in R

The actual implementation of the simulation procedure depends on the programming language. In the statistical package R (Chambers, 2008), loops should be avoided. Thus, in this section, a more efficient version of the simulation procedure is constructed for the R environment.

In algorithm 1, there is a 'while' and a 'for' loop which can be avoided in the implementation. The idea is to use the already existing functions of R and to generate a large number of beats and spikes by just two function calls and to remove all values which are not needed. In this procedure, one has to ensure that enough random variables are simulated or the spike train will be too short. But if too many spikes and beats are generated, this procedure may waste resources and be worse than using loops. For this reason, the numbers of beats needed to generate the spike train are estimated.

Number of beats to cross K : We have a random walk $(S_k)_{k \in \mathbb{N}_0}$, with $S_0 = 0$ and $S_k \sim \mathcal{N}(k\mu, k\sigma_1^2)$ with $0 < \mu < \infty$ and $0 < \sigma_1 < \infty$. $k \in \mathbb{N}$ should be the smallest


```

input : parameters  $\mu, \sigma_1, \sigma_2, \gamma, m$ 
         interval length  $\mathcal{T}$ 
         error  $\varepsilon$  and  $K_0$ 
output: spike times  $t_1, \dots, t_n$ 

 $B := -K_0\mu$  ;
 $k := 0$  ;
 $K_2 := \max\{\mathcal{T} - q_\varepsilon\sigma_2, \mathcal{T} + \frac{\sigma_1^2}{2\mu} \log \varepsilon\}$ ;
while  $B < K_2$  do
    if  $m = 0$  then
        |  $N \sim \text{Bernoulli}(\gamma)$  ;
    else
        |  $N \sim \text{Pois}(\gamma)$ ;
    end
    if  $N > 0$  then
        |  $Z_1 \dots, Z_N \sim \mathcal{N}(0, \sigma_2^2)$  ;
        for  $i \leftarrow 1$  to  $N$  do
            |  $t_{k+i} := B + Z_i$ 
        end
        |  $k := k + N$ 
    end
     $X \sim \mathcal{N}(\mu, \sigma_1^2)$ ;
     $B := B + X$  ;
end
 $\{t_1, \dots, t_n\} := \{t_i : t_i \geq 0, t_i \leq \mathcal{T}, i = 1, \dots, k\}$  ;
return  $(t_{(1)}, t_{(2)}, \dots, t_{(n)})$ 

```

Algorithm 1: Pseudo-code for generating a spike train according to the GLO assumptions with parameters $\mu, \sigma_1, \sigma_2, \gamma, m$ and time length \mathcal{T} . Optional arguments are K_0 which controls the convergence to a stationary process and the error ε which is used for determining K_1 and K_2 .

index value, such that the probability for the random walk to be smaller than a constant $K > 0$ is smaller or equal than $\varepsilon > 0$, i.e.

$$k = \inf\{j \in \mathbb{N} : \mathbb{P}(S_j < K) \leq \varepsilon\}. \quad (2.87)$$

Let $Z \sim \mathcal{N}(0, 1)$, then

$$\mathbb{P}(S_k < K) \leq \varepsilon \quad (2.88)$$

$$\Leftrightarrow \mathbb{P}\left(Z < \frac{K - k\mu}{\sigma_1\sqrt{k}}\right) \leq \varepsilon \quad (2.89)$$

$$\Leftrightarrow \frac{K - k\mu}{\sigma_1\sqrt{k}} \leq q_\varepsilon \quad (2.90)$$

where q_ε denotes the ε -quantile of the standard normal distribution. Typically ε should be small and close to zero, thus $q_\varepsilon < 0$ and the lefthand side of (2.90) is smaller than zero. Thus, squaring both sides would change the direction of the sign and then (2.90) is equivalent to

$$k^2 - k \frac{(2K\mu + q_\varepsilon^2\sigma_1^2)}{\mu^2} + \frac{K^2}{\mu^2} \geq 0. \quad (2.91)$$

This quadratic function can be solved in the usual way. Resulting in

$$k \geq \frac{2K\mu + q_\varepsilon^2\sigma_1^2}{2\mu^2} + \frac{1}{2\mu} \sqrt{q_\varepsilon^2\sigma_1^2(4K + q_\varepsilon^2\sigma_1^2)}. \quad (2.92)$$

Remark 2.33 (Efficient R implementation): The R function `SIM` in appendix A.3 can be modified to account for vector valued programming of R by using the following procedure:

1. Let $K = K_0\mu + K_2$ and determine k according to equation (2.92).
2. Generate the beats of the BR as k steps of the random walk S_0, S_1, \dots, S_k with $S_0 = -K_0\mu$, $S_i = S_{i-1} + X_i$ and $X_1, \dots, X_k \sim \mathcal{N}(\mu, \sigma_1^2)$.
3. Delete all beats outside the interval $[-K_1, \mathcal{T} + K_1]$.
4. For each of the remaining beats draw the number of spikes and place them around their birth beat according to a normal distribution with variance σ_2^2 .
5. Delete all spikes outside the interval $[0, \mathcal{T}]$ return the ordered spike times.

A R code of this procedure called `GL0` can also be found in appendix A.3.

Remark 2.34 (Comparison of two different R implementations): The real advantage of this vector valued programming is the reduced computational time for simulation. Figure 2.5 shows the speed gain by using vectorized programming with function `GLO` in contrast to loops in the function `SIM`. The simulations are performed on a MAC Pro (OS X 10.6.8) with a Quad-Core Intel Xeon 2,93 GHz processor with 8 Gb RAM and R version 2.13.1.

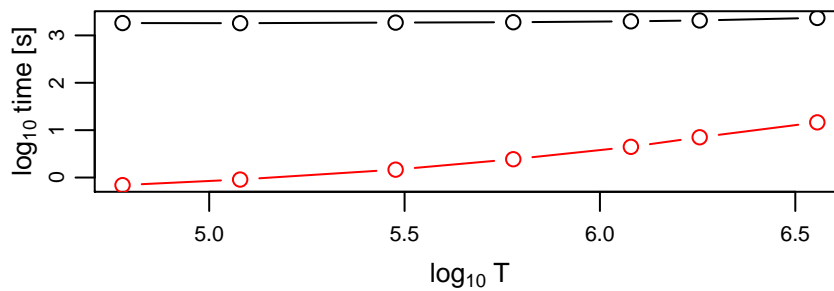


Figure 2.5: Running time analysis. For different recording times \mathcal{T} the time for simulating 1000 GLO spike trains is measured for a vectorized R Code (`GLO`, red) and a R Code using loops (`SIM`, black). Details of the R functions can be found in appendix A.3.

2.3.3. Predefined number of spikes

The code in algorithm 1 and its implementation in the R code can be slightly modified to return a predefined number of spikes n . In general, it is still necessary to simulate over an interval and not stopping after n spikes, because in the case of a very irregular parameter combination ($\sigma_1 \gg \mu$ or $\sigma_2 \gg \mu$) future spikes can be placed before actual spikes. So, as a short workaround to solve this problem, one can simply increase \mathcal{T} and take only the first n spikes of the spike train. But how large has \mathcal{T} to be taken? In a first step we crudely estimate the number of beats needed to generate n spikes. In the second step, we estimate the time \mathcal{T} to generate that many beats.

1. Number of beats: Let $X := \sum_{i=1}^k P_i^{(m)}$ denote the number of spikes in the interval $[0, \mathcal{T}]$ and k the number of beats needed to generate these spikes. One has

$$X \sim \begin{cases} \text{Bin}(k, \gamma) & , m = 0 \\ \text{Pois}(k\gamma) & , m = 1, \end{cases} \quad (2.93)$$

where $\text{Bin}(k, \gamma)$ denotes the Binomial distribution with k trials and success probability γ . Furthermore, we want to ensure to have n spikes with high probability,

that is

$$\mathbb{P}(X < n) < \varepsilon \quad (2.94)$$

for some given (small) ε . Because typically n is large, we can use the normal approximation, such that

$$\frac{n - \mathbb{E}[X]}{\sqrt{\text{Var}(X)}} < q_\varepsilon \quad (2.95)$$

is equivalent to (2.94) with q_ε as the ε -quantile of the standard normal distribution. Solving (2.95) for k , we get

$$k > \frac{2n + q_\varepsilon^2(1 - \gamma)^{1-m}}{2\gamma} + \frac{1}{2\gamma} \sqrt{4nq_\varepsilon^2(1 - \gamma)^{1-m} + (q_\varepsilon^2(1 - \gamma)^{1-m})^2}. \quad (2.96)$$

- 2. Size of \mathcal{T} :** Given the number of beats k , we assume $B_1 = 0$ and further require $\mathbb{P}(B_k \geq \mathcal{T}) < \varepsilon$. Because $B_k \sim \mathcal{N}(k\mu, k\sigma_1^2)$, we choose \mathcal{T} to be the $1 - \varepsilon$ -quantile of this distribution.

This choice of \mathcal{T} ensures with high probability (depending on ε), that there are at least n spikes in the interval $[0, \mathcal{T}]$.

2.3.4. Examples of spike train simulation

In this section, two examples are shown to illustrate possible applications of the simulation procedure. In the first case (Example 2.35), we make use of simulations to approximate the distribution of the first beat B_1 and in the second case (Example 2.36), different firing patterns are generated for different GLO parameters.

Example 2.35 (Distribution of B_1): For approximating the distribution of B_1 , 100.000 GLO spike trains are simulated and the value of B_1 saved each time (\mathcal{T} is not important in that case). Figure 2.6 shows the distribution of B_1 for different parameter values.

Example 2.36 (Firing patterns): Figure 2.7 shows four examples of simulated spike trains representing a regular oscillating (A), regular bursty (B), irregular bursty (C) and irregular firing pattern (D). The change in irregularity is achieved by an increase of σ_1 and σ_2 and the burst behavior by switching the firing mode m and parameter γ .

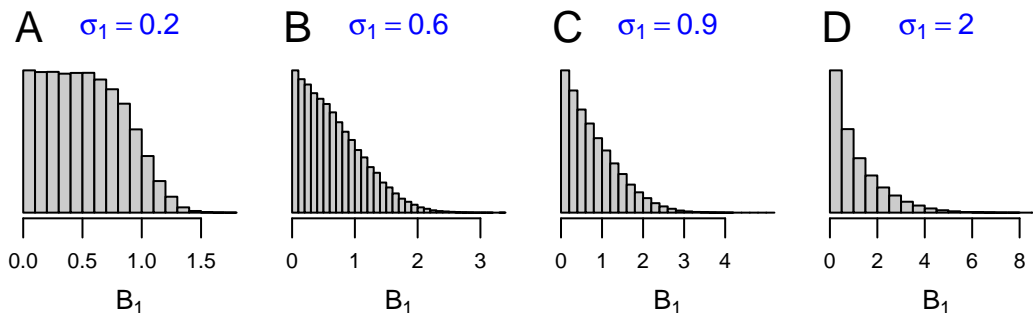


Figure 2.6: Simulated distribution of B_1 for $\mu = 1$ and different values of σ_1 . For large σ_1 , the distribution has an exponentially looking shape.

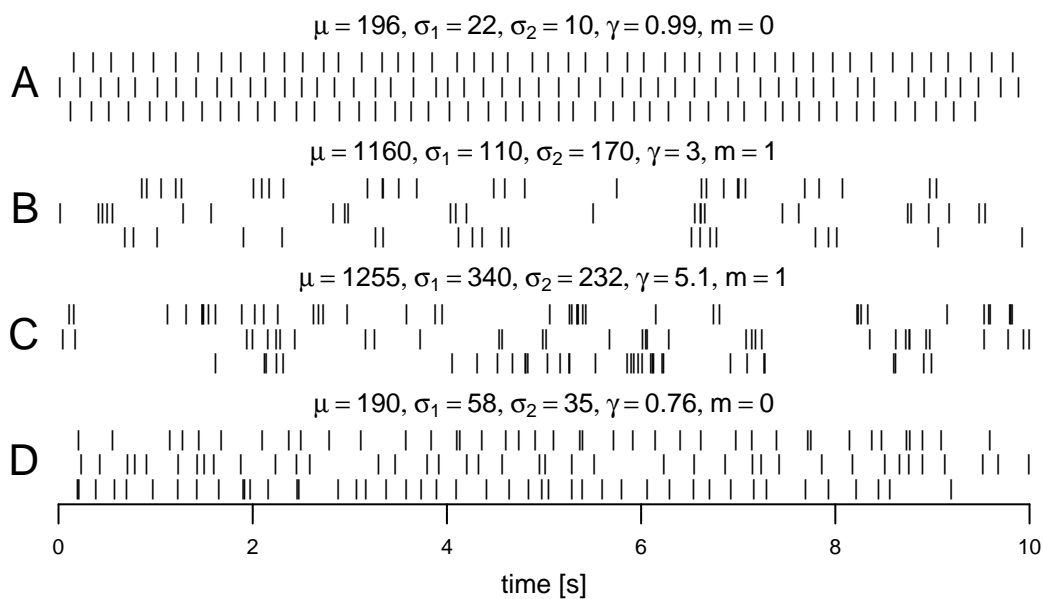


Figure 2.7: Rasterplots of simulated GLO spike trains with different parameter constellations. Different firing patterns are produced: non-bursty (A & D), bursty (B & C), regular (A & B) and irregular (C & D).

Chapter 3.

Interspike interval distributions

In many theoretical spike train models such as pseudo-markov models (Ekholm and Hyvärinen, 1970), Poisson processes (Abeles, 1982), hidden markov models (Camproux et al., 1996), inverse gaussian probability models (Iyengar and Liao, 1997) or Gamma processes (Barbieri et al., 2001; Shimokawa and Shinomoto, 2009), the distribution of the waiting times W_k between spikes, the so called *interspike intervals* (ISIs), determines crucially the nature of the processes and is often used for parameter estimation. The present spike train model is not a renewal process and due to its doubly stochastic nature, the ISI distributions can only be approximated for some parameter combinations and thus provides only limited information.

In Section 3.1, we start with the ordinary definition of spike times and interspike intervals. In the context of the GLO, we additionally introduce the concept of beat-ordered spike and waiting times. Furthermore, we define intuitively the event of a spike inversion which is used to describe 'regular' GLO spike trains. Then the ISI distributions for regular non-bursty (Section 3.2) and regular bursty GLO spike trains (Section 3.3) are approximated by the temporal ordered waiting times. Finally the waiting time distribution is regarded for the irregular case and in the context of spike inversions (Section 3.4).

3.1. Introductory remarks

Although, we know the distribution of the difference $B_{i+1} - B_i$ for $i \in \mathbb{N}$, we have seen that the distribution of B_1 is difficult to derive (Section 2.2.1). So for simplification, we assume that $B_1 = 0$ for the rest of this chapter. Although both processes (Φ and \mathbb{B}) lose their stationarity property, they will be interval stationary. Because the GLO process stays a simple process, we can introduce the ordinary spike and waiting times in the following way.

Definition 3.1 (Spike times): The random *spike times* $(T_k)_{k \in \mathbb{Z}}$ of the GLO process Φ are defined as

$$T_k := \begin{cases} \inf\{t \geq 0 : \Phi([0, t]) = k\} & , k > 0 \\ \sup\{t < 0 : \Phi([t, 0]) = |k| + 1\} & , k \leq 0. \end{cases} \quad (3.1)$$

This coincides with equation (1.11) and definition 1.15 on page 10.

Definition 3.2 (Interspike intervals): We define an *interspike interval (ISI)* of j th-order as

$$W_k^{[j]} := T_{k+j} - T_k, \quad k \in \mathbb{Z}, \quad j \in \mathbb{N}. \quad (3.2)$$

So the spikes are temporally ordered and we have a sequence of spike times $\{T_k : k \in \mathbb{Z}\}$ with $\dots < T_{-1} < T_0 < 0 \leq T_1 < T_2 < \dots$ and a sequence of waiting times $\{W_k^{[j]} : k \in \mathbb{Z}\}$ with $W_k^{[j]} \geq 0, \forall k \in \mathbb{Z}, \forall j \in \mathbb{N}$. Figure 3.1 visualises ISIs of first and second order. Furthermore, we define $f_{W^{[j]}}$ as the density function and set $W_k := W_k^{[1]}$.

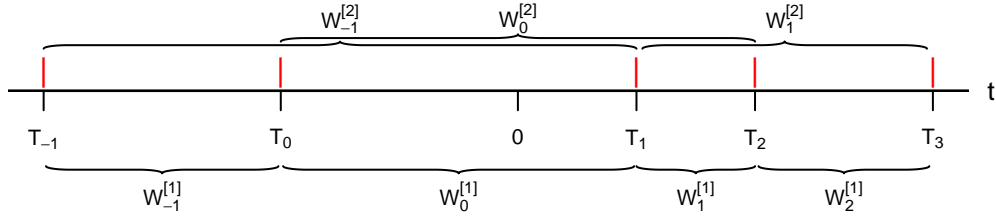


Figure 3.1: ISIs of first and second order between spikes (red).

While definition (3.1) sorts the spikes along the line, there is also the possibility to order them in the sequence of how the spikes are generated according to the random walk $(B_k)_{k \in \mathbb{Z}}$. This we call *beat-ordered* and is used for simplifying some calculations and is not connected with neurophysiological assumptions.

Definition 3.3 (Beat-ordered spike times): Let $\Phi \sim \text{GLO}(\psi)$ be as in equation (2.9). We define the k -th *beat-ordered spike time* ($k \in \mathbb{Z}$) as

$$\tilde{T}_k := B_{i_0(k)} + Z_{i_0(k), j_0(k)} \quad (3.3)$$

with the indices $i_0(k)$ and $j_0(k)$ depending on k and given as follows: Let

$$l_0 := \begin{cases} \#\{i \in [P_1] : Z_{1,i} < 0\} & , P_1 > 0 \\ 0 & , P_1 = 0 \end{cases} \quad (3.4)$$

be the number of spikes of beat B_1 which fall into the negative axis. Then the k -th spike is generated in beat $B_{i_0(k)}$ with

$$i_0(k) := \begin{cases} \inf\{i \in \mathbb{Z} : \sum_{j=1}^i P_j \geq k + l_0\} & , k > 0 \\ \sup\{i \in \mathbb{Z} : i \leq 1, l_0 + \sum_{j=i}^0 P_j \geq |k| + 1\} & , k \leq 0. \end{cases} \quad (3.5)$$

Furthermore, the position of the k -th spike in beat $B_{i_0(k)}$ is

$$k_0(k) := \begin{cases} k + l_0 - \sum_{l=1}^{i_0(k)-1} P_l & , k > 0 \\ l_0 - |k| - 1 + \sum_{l=i_0(k)+1}^0 P_l & , k \leq 0 \end{cases} \quad (3.6)$$

which gives

$$j_0(k) := \begin{cases} \{i \in [P_{i_0(k)}] : \#\{j : Z_{i_0(k),j} \leq Z_{i_0(k),i}\} = k_0(k)\} & , k > 0 \\ \{i \in [P_{i_0(k)}] : \#\{j : Z_{i_0(k),j} \geq Z_{i_0(k),i}\} = k_0(k)\} & , k \leq 0. \end{cases} \quad (3.7)$$

In words: For \tilde{T}_1 , we look for the first beat $B_{i_0(1)}$ which gives birth to a spike on the positive half axis. This spike is defined to be \tilde{T}_1 . Supposing there are five spikes at beat $B_{i_0(1)}$. Spikes of the same beat which are larger than \tilde{T}_1 (e.g., there would be two), are written as \tilde{T}_2 and \tilde{T}_3 and the rest as \tilde{T}_0 and \tilde{T}_{-1} . Then going to the next beat $B_{i_0(1)+1}$, which maybe has two spikes and the smaller one is defined to be \tilde{T}_4 and so on.

Example 3.4: Examples of classical and beat-ordered spike times can be seen in Figure 3.2. In the case of a regular GLO process ($\sigma_1 \ll \mu$ and $\sigma_2 \ll \mu$), the spike times are equal, but in the irregular case they are not identical.

Definition 3.5 (Beat-ordered interspike intervals): Analogously to equation (3.2) we define the *beat-ordered interspike interval* as

$$\tilde{W}_k^{[j]} := \tilde{T}_{k+j} - \tilde{T}_k, \quad k \in \mathbb{Z}, \quad j \in \mathbb{N} \quad (3.8)$$

and write $f_{\tilde{W}^{[j]}}$ for the density function and $\tilde{W}_k := \tilde{W}_k^{[1]}$.

Remark 3.6: In this setting, it is possible that $\tilde{T}_k < \tilde{T}_j$ for $k > j$ and that $\tilde{W}_k^{[j]}$ can take negative values.

Remark 3.7 (Spike inversions): The definition of the spike times in equation (3.1) accounts for the possibility of placing spikes of beat B_k after spikes of beat B_{k+j} ($j \in \mathbb{N}$), e.g. Figure 3.2. This is called a *spike inversion* and is a similar problem to the inversion of beats (compare Section 2.2.1). Typically a spike inversion happens when σ_1 and/or

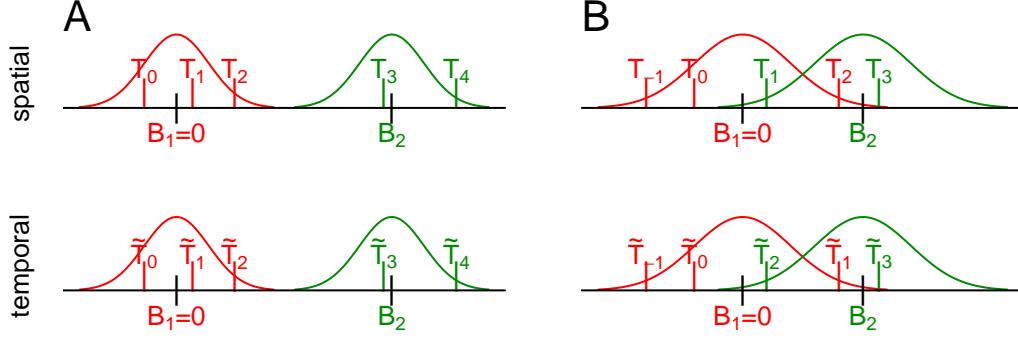


Figure 3.2: Classical spike times and beat-ordered spike times. Given $B_1 = 0$, $P_1 = 3$ and $P_2 = 2$, there are shown two different situations: (A) the classical and beat-ordered spike times are identical and (B) the spike times are not identical.

σ_2 are large relative to μ and is thus a characteristic of irregular firing patterns. Because the normal distribution is defined on \mathbb{R} , there is always a positive probability to have an inversion - but dependent on the parameters, it can be exponentially small.

Although spike inversions may appear between beats which are far away from each other, it is more likely to happen between beats which are close together (see Lemma 2.10). So for simplicity, we focus on spikes of two adjacent beats, i.e. B_k and B_{k+1} , and exactly one spike per beat. Furthermore, because of the stationarity of Φ , we can set $k = 1$.

Definition 3.8 (Spike inversion): Let $\Phi \sim \text{GLO}(\psi)$ be given as in equation (2.9), then the event

$$\eta := \{B_1 + Z_{1,1} > B_2 + Z_{2,1}\} \quad (3.9)$$

is called a *spike inversion* (a more general definition will be introduced in in Section 3.4).

Remark 3.9: In a regular GLO spike train, a spike inversion will only happen with a small probability. The restriction on the parameters to satisfy this condition is given as the following relationship

$$\varepsilon > \mathbb{P}(\eta) = \mathbb{P}(B_2 - B_1 + Z_{2,1} - Z_{1,1} < 0) = F_{\{0,1\}} \left(-\mu / \sqrt{\sigma_1^2 + 2\sigma_2^2} \right). \quad (3.10)$$

For the parameters, it follows that

$$\mu > -q_\varepsilon \sqrt{\sigma_1^2 + 2\sigma_2^2}, \quad (3.11)$$

where q_ε denotes the ε -quantile of the standard normal distribution. Because $F_{\{0,1\}}(-z) \leq \exp(-z^2/2)/(z\sqrt{2\pi})$, the probability of an inversion gets exponentially small with increasing z . Figure 3.3 visualizes the inversion probability for different parameter constellations.

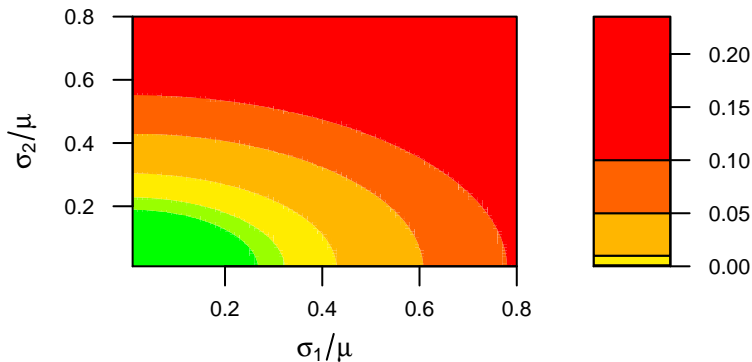


Figure 3.3: Inversion probability for different parameter constellations. Green combinations have an inversion probability less than 0.001.

Remark 3.10 (The regular case): As mentioned previously, a high inversion probability is connected with irregular spike trains. So the 'regular case' of a GLO spike train is defined as all parameter combinations $(\mu, \sigma_1, \sigma_2)$ with $\sigma_1 \ll \mu$ and $\sigma_2 \ll \mu$, such that $\mathbb{P}(\eta) < \varepsilon$. Of course the size of ε depends on the investigator. In this regular case, $\mathbb{P}(\tilde{T}_k \neq T_k) < \varepsilon$, so the ISI distribution $W_k^{[j]}$ can be approximated by $\tilde{W}_k^{[j]}$ for the non-bursty ($m = 0$, Section 3.2) and bursty firing mode ($m = 1$, Section 3.3). So, we will identify $\tilde{W}_k^{[j]}$ with $W_k^{[j]}$ and call both ISIs.

3.1.1. Coefficient of variation of ISIs

Definition 3.11: The *coefficient of variation (CV)* of a random variable X is defined to be

$$c_v(X) := \frac{\sqrt{\text{Var}(X)}}{\mathbb{E}[X]}. \quad (3.12)$$

It is a measure of dispersion of the distribution \mathbb{P}_X . Given a realization $\mathbf{x} = (x_1, \dots, x_n)$

and $\bar{x} = \sum_{i=1}^n x_i$, an estimator for the CV is

$$\hat{c}_v(\mathbf{x}) := \frac{\sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}}{\bar{x}}. \quad (3.13)$$

Remark 3.12 (Classification of firing patterns according to the CV): In computational neuroscience, the CV is typically derived for the theoretical ISI distribution represented by W or for the empirical observations $\mathbf{w} = (w_1, \dots, w_n)$ of the ISIs. The CV is used to quantify the regularity of stationary spike trains (Holt et al., 1996; Davies et al., 2006).

- A regular spike train has small variability between spikes, so that the nominator in equation (3.12) is small in comparison to the denominator. Thus the CV will be close to zero.
- In contrast, a Poisson process is regarded as a typical example of an irregular spike train. In this case, the ISI distribution is exponential, $W \sim \text{Exp}(\lambda)$, so that $\mathbb{E}[W] = 1/\lambda$ and $\text{Var}(W) = 1/\lambda^2$. Thus the CV will be equal to one.
- Finally for bursty neurons, because there are clusters with short ISIs, separated by larger intervals, the nominator will be typically larger than the denominator and the CV will be larger than one.

Because the main disadvantage of the CV is its high sensitivity to rate changes and outliers in ISIs, this considerations do not hold true for non-stationary spike trains. Table 3.1 shows the \hat{c}_v values of the spike trains shown in Figure 1.1 and Figure 1.2. Although spike train (B) is regarded as regular bursty by the investigators, its CV is less than one.

| spike train | A | B | C | D |
|-------------|------|------|------|------|
| \hat{c}_v | 0,14 | 0,91 | 1,12 | 0,52 |

Table 3.1: The estimated coefficient of variation \hat{c}_v for the empirical spike trains shown in Figure 1.1 and Figure 1.2. The regular spike train (A) has the lowest CV and the irregular one (D) an increased CV between 0 and 1. Although (B) and (C) are regarded as bursty, this cannot be clearly detected by the CV, but because the CV of (C) is bigger than one, there is an indication for burstiness in this case.

3.2. Non-bursty spike trains in the regular case

Assumption 3.13 (The regular non-bursty case): We assume $B_1 = 0$, $m = 0$ and $\sigma_1 \ll \mu$, $\sigma_2 \ll \mu$ according to Remark 3.10. Furthermore, we are only interested in $\tilde{W}_k^{[j]}$ with $k \in \mathbb{N}$.

Lemma 3.14: Given assumption 3.13 and $\gamma = 1$, $\tilde{W}_k^{[j]} \sim \mathcal{N}(j\mu, j\sigma_1^2 + 2\sigma_2^2)$ for $k \in \mathbb{N}$.

Proof. The spike times can be written as

$$\tilde{T}_k = \begin{cases} B_k + Z_{k,1} & , Z_{1,1} \geq 0 \\ B_{k+1} + Z_{k+1,1} & , Z_{1,1} < 0. \end{cases} \quad (3.14)$$

Both cases are identical, besides of the shift of the indices, so we focus on the first one. Because for $k \in \mathbb{N}$, the increments $B_{k+1} - B_k$ are independent, we will only consider this case.

$$\begin{aligned} \tilde{W}_k^{[j]} &= \tilde{T}_{k+j} - \tilde{T}_k \\ &= (B_{k+j} + Z_{k+j,1}) - (B_k + Z_{k,1}) \\ &= (B_{k+j} - B_k) + Z_{k+j,1} - Z_{k,1} \\ &= (B_{k+1} - B_k) + \dots + (B_{k+j} - B_{k+j-1}) + Z_{k+j,1} - Z_{k,1} \end{aligned}$$

Thus, $\tilde{W}_k^{[j]}$ has distribution $\mathcal{N}(j\mu, j\sigma_1^2 + 2\sigma_2^2)$. \square

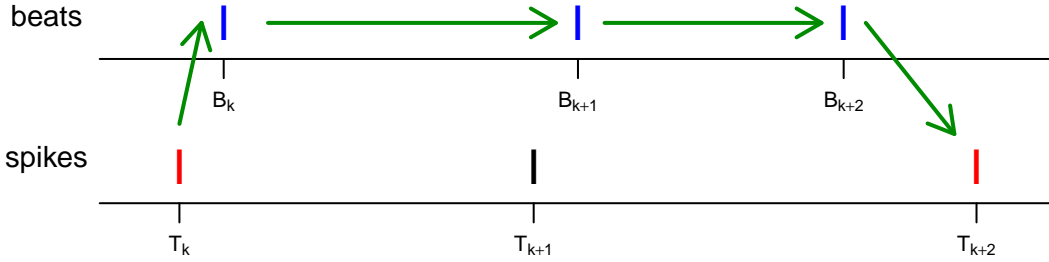


Figure 3.4: Schematic computation of a beat-ordered waiting time distribution. For deriving $\tilde{W}_k^{[2]} = \tilde{T}_{k+2} - \tilde{T}_k$ in the case of $m = 0$ and $\gamma = 1$, there are the following independent steps: From spikes to beat according to $\mathcal{N}(0, \sigma_2^2)$, from beat to beat according to $\mathcal{N}(\mu, \sigma_1^2)$ and from beat to spike according to $\mathcal{N}(0, \sigma_2^2)$. Thus, $\tilde{W}_k^{[2]} \sim \mathcal{N}(2\mu, 2\sigma_2^2)$.

Figure 3.4 visualizes this result which is an application of backward techniques (Liemant et al., 1988) which also arise in the context of backward and forward trees for cluster

fields (Kallenberg, 1977). However, for $0 < \gamma < 1$, an ISI has this distribution only with probability γ^j , because a beat has a spike only with probability γ . Given a spike at beat B_k , the probability that the j -th next spike comes from beat B_{k+n} is given by the weights of a negative binomial distributed random variable G (we write $G \sim \text{NB}(j, \gamma)$). Here, $\gamma \in (0, 1)$ represents the success probability and $j \in \mathbb{N}$ the number of successes. The probability weights are given as

$$\mathbb{P}(G = n) = \binom{n-1}{j-1} \gamma^j (1-\gamma)^{n-j}, \quad n \in \mathbb{N}. \quad (3.15)$$

Furthermore, the expectation and variance of a negative binomial distribution are well known and given as

$$\mathbb{E}[G] = \frac{j}{\gamma} \quad \text{and} \quad \text{Var}(G) = \frac{j(1-\gamma)}{\gamma^2}. \quad (3.16)$$

Now, this can be used to compute the beat-ordered waiting time distribution for $m = 0$ and $0 < \gamma < 1$.

Proposition 3.15: Let $B_1 = 0$, $m = 0$ and $0 < \gamma < 1$. For $k \in \mathbb{N}$, the ISI $\tilde{W}_k^{[j]}$ has density

$$f_{\tilde{W}^{[j]}|m=0}(x) = \sum_{k=1}^{\infty} \mathbb{P}(G = k) \cdot \varphi_{\{k\mu, k\sigma_1^2 + 2\sigma_2^2\}}(x), \quad (3.17)$$

Proof. Obviously, the ISI $\tilde{W}_k^{[j]}$ should have the same distribution as $\tilde{T}_{k+G} - \tilde{T}_k$ and we write

$$\tilde{W}_G^{[j]} := \tilde{T}_{k+G} - \tilde{T}_k, \quad (3.18)$$

with

$$\tilde{W}_G^{[j]}|G \sim \mathcal{N}(G \cdot \mu, G \cdot \sigma_1^2 + 2\sigma_2^2). \quad (3.19)$$

Using the law of total probability, the unconditioned density of $\tilde{W}_k^{[j]}$ for $m = 0$ can thus be described by

$$f_{\tilde{W}^{[j]}|m=0}(x) = \sum_{k=1}^{\infty} \mathbb{P}(G = k) \cdot \varphi_{\{k\mu, k\sigma_1^2 + 2\sigma_2^2\}}(x), \quad (3.20)$$

where $\varphi_{\{\mu, \sigma^2\}}(x)$ denotes the density of the normal distribution with mean μ and variance σ^2 at x . \square

Example 3.16: To investigate the approximation of W_k by \tilde{W}_k , a GLO spike train $\mathcal{S} = (t_1, \dots, t_n)$ with $n = 100.000$ spikes is simulated with parameter vector ψ (compare Section 2.3). Figure 3.5 shows three simulated ISI W_k distributions and the corresponding theoretical density function of \tilde{W}_k . If $\sigma_1 \ll \mu$ and $\sigma_2 \ll \mu$, the approximation seems to be adequate (A and B). But if the variances increase, there will be a larger probability of spike inversions, which has the consequence that $f_{\tilde{W}^{[j]}|m=0}(x)$ has too much mass on the negative axis (C).

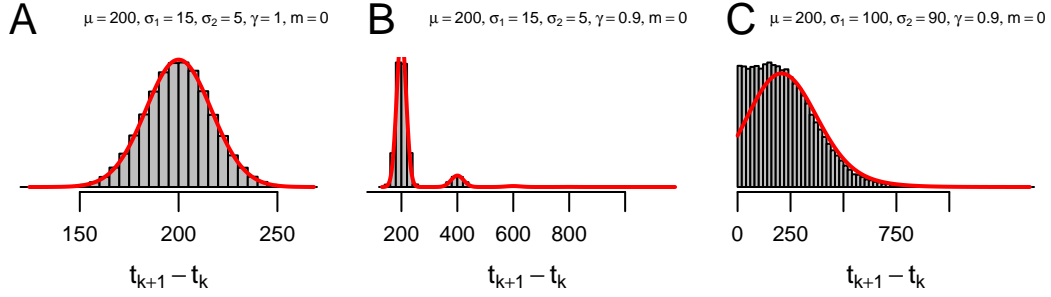


Figure 3.5: Simulated ISI distributions ($n = 100.000$, grey histograms) and probability density function according to (3.17) (red lines) for different parameter settings in the non-bursty case.

Remark 3.17 (Expectation, variance, CV and covariance of ISIs): The expectation of $\tilde{W}_k^{[j]}$ is given by

$$\mathbb{E} \left[\tilde{W}_k^{[j]} \right] = \mathbb{E} \left[\tilde{W}_G^{[j]} \right] = \mathbb{E} \left[\mathbb{E} \left[\tilde{W}_G^{[j]} \mid G \right] \right] = j \frac{\mu}{\gamma}. \quad (3.21)$$

The variance can be calculated as follows:

$$\text{Var} \left(\tilde{W}_k^{[j]} \right) = \mathbb{E} \left[\text{Var} \left(\tilde{W}_G^{[j]} \mid G \right) \right] + \text{Var} \left(\mathbb{E} \left[\tilde{W}_G^{[j]} \mid G \right] \right) \quad (3.22)$$

$$= j \frac{\sigma_1^2}{\gamma} + \left(\frac{\mu}{\gamma} \right)^2 (1 - \gamma) + 2\sigma_2^2. \quad (3.23)$$

The coefficient of variation (see equation 3.12) is given as

$$c_v \left(\tilde{W}_k^{[j]} \right) = \frac{\sqrt{\text{Var} \left(\tilde{W}_k^{[j]} \right)}}{\mathbb{E} \left[\tilde{W}_k^{[j]} \right]}. \quad (3.24)$$

If G is deterministic (i.e., $\gamma = 1$), the formulas (3.21) and (3.23) reduce to the mean and variance of eq. (3.15). With decreasing firing probability γ , the variance of the interval

increases due to the increased probability to skip a beat of the background rhythm, the same holds true for the mean.

Because of the fundamental relation between ISIs of j th-order and ISIs of first order

$$\tilde{W}_k^{[j]} = \tilde{W}_k^{[1]} + \tilde{W}_{k+1}^{[1]} + \dots + \tilde{W}_{k+j-1}^{[1]} \quad (3.25)$$

we get for the covariance of two successive ISIs of first order as

$$\text{Cov}(\tilde{W}_k^{[1]}, \tilde{W}_{k+1}^{[1]}) = 1/2 \left[\text{Var}(\tilde{W}_k^{[1]} + \tilde{W}_{k+1}^{[1]}) - \text{Var}(\tilde{W}_k^{[1]}) - \text{Var}(\tilde{W}_{k+1}^{[1]}) \right] \quad (3.26)$$

$$= 1/2 \cdot \text{Var}(\tilde{W}_k^{[2]}) - \text{Var}(\tilde{W}_k^{[1]}) \quad (3.27)$$

$$= -\sigma_2^2. \quad (3.28)$$

Because $\tilde{W}_k^{[1]}$ is independent of $\tilde{W}_{k+i}^{[1]}$ for $|i| > 1$, we have $\text{Cov}(\tilde{W}_k^{[1]}, \tilde{W}_{k+i}^{[1]}) = 0$ for $|i| > 1$. If $I_1 = \{k_1, k_1 + 1, \dots, k_1 + j_1 - 1\}$ and $I_2 = \{k_2, k_2 + 1, \dots, k_2 + j_2 - 1\}$ with $j_1, j_2 \in \mathbb{N}$, we get the covariance of two arbitrary ISIs $\tilde{W}_{k_1}^{[j_1]}$ and $\tilde{W}_{k_2}^{[j_2]}$ as

$$\begin{aligned} \text{Cov}(\tilde{W}_{k_1}^{[j_1]}, \tilde{W}_{k_2}^{[j_2]}) &= \sum_{i_1 \in I_1} \sum_{i_2 \in I_2} \text{Cov}(\tilde{W}_{i_1}^{[1]}, \tilde{W}_{i_2}^{[1]}) \\ &= -\#\{(i_1, i_2) \in I_1 \times I_2 : |i_1 - i_2| = 1\} \sigma_2^2 \\ &\quad + \#\{(i_1, i_2) \in I_1 \times I_2 : i_1 = i_2\} \text{Var}(\tilde{W}_k^{[1]}). \end{aligned} \quad (3.29)$$

3.3. Bursty spike trains in the regular case

Assumption 3.18 (The regular bursty case): We assume $B_1 = 0$, $m = 1$ and $\sigma_1 \ll \mu$, $\sigma_2 \ll \mu$ according to Remark 3.10 on page 51. Furthermore, we are only interested in $\tilde{W}_k^{[j]}$ with $k \in \mathbb{N}$ and throughout this section we focus on ISIs of first order. So $j = 1$ and we will omit the index and write \tilde{W}_k .

Remark 3.19 (Decomposition in between- and in-beat intervals): For bursty firing patterns we can split \tilde{W}_k in two kinds of intervals: intervals between spikes of the same beat which we call 'in-beat (ib) intervals' \tilde{W}_k^{ib} and intervals representing the time between two spikes of different beats which we call 'between-beats (bb) intervals' \tilde{W}_k^{bb} (see Figure 3.6). For both types, order statistics (David, 1970) are used to determine the conditioned interspike interval distribution (Section 3.3.1), until finally the unconditioned distribution is derived in Section 3.3.2.

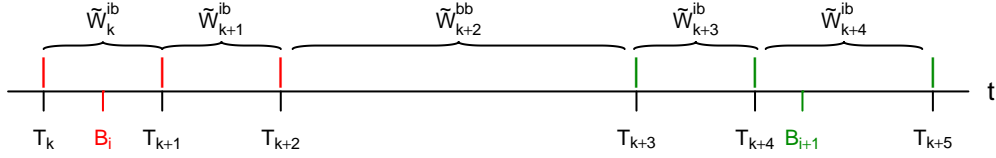


Figure 3.6: In-beat and between-beat intervals in a bursty GLO spike train. The first three spikes (red) belong to the same beat, as well as the last three spikes (green). In-beat (ib) ISIs result of spikes of the same beat and between-beat (bb) ISIs represent the intervals between spikes of different beats.

3.3.1. Conditioned ISIs with order statistics

We start with a basic introduction of definitions and results of order statistics (David, 1970) and translate them in the GLO context.

3.3.1.1. Order statistics

Let Z_1, Z_2, \dots, Z_n be n independent variates, each with cumulative distribution function (c.d.f.) $F(x)$ and probability density function (p.d.f.) $f(x)$. If the independent identically distributed random variables are rearranged in ascending order of magnitude and written as

$$Z_{(1:n)} \leq Z_{(2:n)} \leq \dots \leq Z_{(n:n)}, \quad (3.30)$$

then $Z_{(k:n)}$ is called the k -th order statistic ($k = 1, \dots, n$) and it has p.d.f.

$$f_{(k:n)}(x) = \frac{n!}{(k-1)!(n-k)!} F^{k-1}(x) [1 - F(x)]^{n-k} f(x). \quad (3.31)$$

The joint density function of $Z_{(r:n)}$ and $Z_{(s:n)}$ ($1 \leq r < s \leq n$) is denoted by $f_{(r,s:n)}(x, y)$. For $x \leq y$, it is given by

$$f_{(r,s:n)}(x, y) = \frac{n!}{(r-1)!(s-r-1)!(n-s)!} F^{r-1}(x) f(x) \cdot [F(y) - F(x)]^{s-r-1} f(y) [1 - F(y)]^{n-s}. \quad (3.32)$$

and for $x > y$ it equals 0.

3.3.1.2. In-beat (ib) intervals

\tilde{W}_k^{ib} are ISIs of first order and result from spikes of the same beat (see Figure 3.6). Given an arbitrary beat B_i of the background rhythm and conditioned on $\{P_i = n\}$ (with $n > 1$, because otherwise there are not any in-beat intervals), every spike of this beat can be written as $B_i + Z_{i,j}$, $j \in \{1, \dots, n\}$. Because only the difference between beats is important, we can set $B_i = 0$. Omitting the index i and relative to the beat, the spike positions Z_1, \dots, Z_n are i.i.d. $\mathcal{N}(0, \sigma_2^2)$. The c.d.f. of this normal distribution is denoted by $F_0(x)$ and the p.d.f. as $\varphi_0(x)$. Using the notation of the previous paragraph, $Z_{(k:n)}$ denotes the k -th ordered spike position ($1 \leq k \leq n$).

The ISIs in such a cluster are defined as the differences

$$\tilde{W}_{(k:n)}^{ib} := Z_{(k+1:n)} - Z_{(k:n)} \quad (3.33)$$

for $k = 1, \dots, n-1$. Because of (3.32) the p.d.f. of $(Z_{(k:n)}, Z_{(k+1:n)})$ is known and given as

$$f_{(k,k+1:n)}(x, y) = \frac{n!}{(k-1)!(n-(k+1))!} F_0^{k-1}(x) \varphi_0(x) \varphi_0(y) [1 - F_0(y)]^{n-(k+1)}. \quad (3.34)$$

So $\tilde{W}_{(k:n)}^{ib}$ has density

$$f_{\tilde{W}_{(k:n)}^{ib}}(z) = \int_{\mathbb{R}} f_{(k,k+1:n)}(x, z+x) dx. \quad (3.35)$$

3.3.1.3. Between-beat (bb) intervals

Still given a beat B_i , but now with $P_i = n_1$ ($n_1 > 0$), a between-beat interval \tilde{W}_k^{bb} is the difference between the last relative spike position $Z_{(n_1:n_1)}$ of beat B_i and the first spike of the next firing beat

$$Z_{(1:n_2)[j]} := B_{i+j} - B_i + Z_{(1:n_2)} \quad (3.36)$$

(so beat B_{i+j} gives rise to n_2 spikes according to $\mathcal{N}(j\mu, j\sigma_1^2 + \sigma_2^2)$). Here j denotes the number of trials till the next beat fires, i.e. $j = \inf\{k \in \mathbb{N} : P_{i+k} > 0\}$. Let F_j and φ_j be the c.d.f. and p.d.f. of $\mathcal{N}(j\mu, j\sigma_1^2 + \sigma_2^2)$. Because of (3.31) the p.d.f. of $Z_{(n_1:n_1)}$ is given by

$$f_{(n_1:n_1)}(x) = n_1 F_0(x)^{n_1-1} \varphi_0(x) \quad (3.37)$$

and the p.d.f. of $Z_{(1:n_2)[j]}$ by

$$f_{(1:n_2)[j]}(y) = n_2(1 - F_j(x))^{n_2-1}\varphi_j(x). \quad (3.38)$$

A between-beat interval is defined as the difference

$$\tilde{W}_{(n_1,n_2)[j]}^{bb} := Z_{(1:n_2)[j]} - Z_{(n_1:n_1)}. \quad (3.39)$$

Because two different beats are independent, the spike positions are also independent and it follows that $(Z_{(n_1:n_1)}, Z_{(1:n_2)[j]})$ has density

$$f_{(n_1,n_2)[j]}(x, y) = f_{(n_1:n_1)}(x) \cdot f_{(1:n_2)[j]}(y). \quad (3.40)$$

So $\tilde{W}_{(n_1,n_2)[j]}^{bb}$ has density

$$f_{\tilde{W}_{(n_1,n_2)[j]}^{bb}}(z) = \int_{\mathbb{R}} f_{(n_1,n_2)[j]}(x, z+x) dx. \quad (3.41)$$

3.3.2. Unconditioned ISIs

Proposition 3.20: Let $B_1 = 0$ and $m = 1$. For $k \in \mathbb{N}$, the ISI \tilde{W}_k has density

$$f_{\tilde{W}|m=1}(x) := f_1(x) := \sum_{k=1}^{\infty} \frac{p_k}{\gamma} \left(\sum_{u=1}^{k-1} f_{\tilde{W}_{(u:k)}^{ib}}(x) + \sum_{n=1}^{\infty} \sum_{g=1}^{\infty} p_0^{g-1} p_n f_{\tilde{W}_{(k,n)[g]}^{bb}}(x) \right), \quad (3.42)$$

with $p_j := \mathbb{P}(P_i = j)$, $j \in \mathbb{N}_0$.

Example 3.21: Figure 3.7 shows the empirical ISI distributions of three simulated GLO spike trains with $n = 100.000$ spikes and different parameters. The corresponding theoretical densities (red) indicate that the ISI distributions are well approximated by equation (3.42) (at least for $\sigma_1 \ll \mu$ and $\sigma_2 \ll \mu$).

3.3.2.1. Proof of Proposition 3.20

The p.d.f. $f_1(x) := f_{\tilde{W}|m=1}(x)$ is derived by conditioning and using the law of total probability. The procedure is shown step by step:

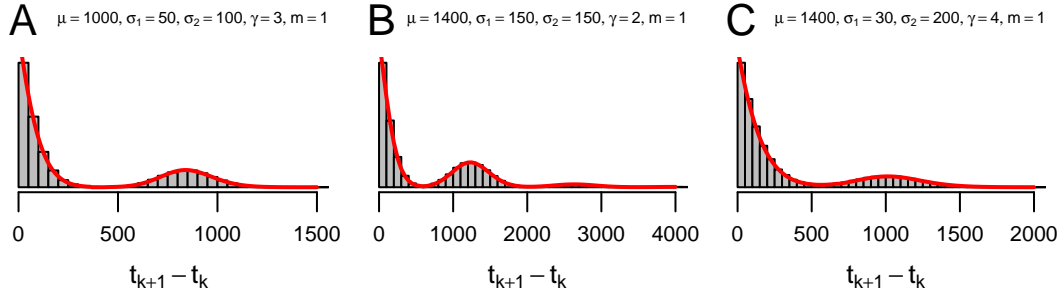


Figure 3.7: Simulated ISI distributions ($n = 100.000$, grey histograms) and probability density function according to (3.42) (red lines) for different parameter settings in the bursty case.

1.) Distribution of the number of spikes: We know that the number of spikes P_i at beat B_i is poisson distributed with

$$p_k := \mathbb{P}(P_i = k), \quad k \in \mathbb{N}_0. \quad (3.43)$$

But if we draw a random point of the stationary point process and look at the distribution of the number of spikes in the same beat (we call this random variable \hat{P}_0 and the belonging beat \hat{B}_0), this number will not be poisson distributed. In fact, we have an above average chance of drawing a spike of a beat with a large number of spikes. Furthermore, by drawing a random spike, we implicitly condition on the event that we have at least one spike in the particular beat. So beats with no spikes are not relevant anymore and we set $I := \{i \in \mathbb{Z} : P_i > 0\}$. Furthermore, we define a \mathbb{N} -valued random variable \tilde{P}_i as

$$\tilde{p}_n := \mathbb{P}(\tilde{P}_i = n) := \mathbb{P}(P_i = n \mid P_i > 0) = \frac{p_n}{1 - p_0}, \quad i \in I, \quad (3.44)$$

so that we get a sequence $(\tilde{P}_i)_{i \in I}$.

Drawing a spike at random and looking for the distribution of \hat{P}_0 is the same as regarding $(\tilde{P}_i)_{i \in I}$ as the life times of a discrete renewal chain and looking for the distribution of the interval containing the origin. Because

$$\mathbb{E}[\tilde{P}_i] = \frac{\gamma}{1 - p_0}, \quad (3.45)$$

it follows from renewal theory that

$$\mathbb{P}(\hat{P}_0 = k) = \frac{k \mathbb{P}(\tilde{P}_i = k)}{\mathbb{E}[\tilde{P}_i]} = \frac{k}{\gamma} p_k. \quad (3.46)$$

So if we draw a spike at random, we have a length biased distribution for the number of spikes in the same beat.

2.) By conditioning on the number of spikes in the initial beat of the interspike interval, we can decompose the p.d.f. $f_1(x)$ of the ISI distribution:

$$f_1(x) = \sum_{k=1}^{\infty} f_1(x | \hat{P}_0 = k) \mathbb{P}(\hat{P}_0 = k). \quad (3.47)$$

3.) Decomposition of intervals: As mentioned in the previous section, there are two kinds of ISIs — in-beat and between-beat intervals. An in-beat (ib) interval is one that contains spikes of the same beat and a between-beat (bb) interval is defined by spikes which come from different beats. We introduce a random variable L which denotes the type of the ISI and takes values ib or bb .

If we draw the last spike of a beat, we automatically get a between-beat interval and an in-beat one in the other case. It is $\mathbb{P}(L = ib | \hat{P}_0 = k) = \frac{k-1}{k}$ and $\mathbb{P}(L = bb | \hat{P}_0 = k) = \frac{1}{k}$, so that we get

$$f_1(x | \hat{P}_0 = k) = f_1(x | \hat{P}_0 = k, L = ib) \frac{k-1}{k} + f_1(x | \hat{P}_0 = k, L = bb) \frac{1}{k}. \quad (3.48)$$

4.) In-beat intervals: If $\hat{P}_0 = k$ there are $k-1$ possible in-beat intervals. Let U be uniform distributed on the set $\{1, \dots, k-1\}$ and denote the number of the U -th ordered spike, then

$$f_1(x | \hat{P}_0 = k, L = ib) = \sum_{u=1}^{k-1} f_1(x | \hat{P}_0 = k, L = ib, U = u) \frac{1}{k-1} \quad (3.49)$$

where

$$f_1(x | \hat{P}_0 = k, L = ib, U = u) = f_{W_{(u,k)}^{ib}}(x) \quad (3.50)$$

is known by (3.35).

5.) Between-beat intervals (number of beats fallen out): Let G be a negative binomial distributed random variable with values in \mathbb{N} which denotes the number of trials till the next beat fires, i.e. $G = \inf\{n \in \mathbb{N} : P_{i+n} > 0\}$. So G is geometric distributed with parameter $1 - p_0$ and it follows that

$$f_1(x | \hat{P}_0 = k, L = bb) = \sum_{g=1}^{\infty} f_1(x | \hat{P}_0 = k, L = bb, G = g) p_0^{g-1} (1 - p_0) \quad (3.51)$$

6.) Between-beat intervals (number of spikes in the second beat): If \tilde{P}_{i+1} denotes the number of spikes in the next firing beat, then

$$f_1 \left(x \mid \hat{P}_0 = k, L = bb, G = g \right) = \sum_{n=1}^{\infty} f_1 \left(x \mid \hat{P}_0 = k, L = bb, G = g, \tilde{P}_{i+1} = n \right) \frac{p_n}{1 - p_0} \quad (3.52)$$

with

$$f_1 \left(x \mid \hat{P}_0 = k, L = bb, G = g, \tilde{P}_{i+1} = n \right) = f_{W_{(k,n)[g]}^{bb}}(x). \quad (3.53)$$

7.) Final result: Putting (3.47) – (3.53) together, one finally gets the result (3.42).

3.4. Spike inversions and the irregular case

In the beginning of this chapter, there has been a first intuitive definition of a spike inversion by regarding two adjacent beats and assuming exactly one spike per beat (compare equation (3.9)). Now this should be made more general:

Definition 3.22 (Spike inversion - updated): Conditioned on the event that at least one spike occurs at a specific backbone beat B_k , $k \in \mathbb{N}$, a spike inversion is then defined as the event that at least one spike from the next firing beat B_{k+G} ($G := \inf\{i > 0 : P_{k+i} > 0\}$) is placed before at least one spike belonging to B_k . This equals the event that the smallest spike of B_{k+G} is smaller than the largest spike of B_k . A spike inversion is given as the event

$$\eta_0 := \left\{ B_k + Z_{(P_k:P_k)} \geq B_{k+G} + Z_{(1:P_{k+G})} \mid P_k > 0, P_{k+G} > 0 \right\} \quad (3.54)$$

$$= \left\{ B_k + Z_{(\tilde{P}_k:\tilde{P}_k)} \geq B_{k+G} + Z_{(1:\tilde{P}_{k+G})} \right\} \quad (3.55)$$

Obviously $\eta \subseteq \eta_0$ and thus $\mathbb{P}(\eta) \leq \mathbb{P}(\eta_0)$. The probability of an inversion can be computed, using the information of the previous sections.

Remark 3.23 (Probability of a spike inversion): For $m = 0$ and an arbitrary parameter combination $(\mu, \sigma_1, \sigma_2, \gamma)$, $f_{W|m=0}^{[1]}(x)$ (compare equation (3.17)) is the density of the difference between the spike of beat B_k and of the spike of the next firing beat B_{k+G} . This equation is true even in the irregular case (i.e. when σ_1 or σ_2 is large

relative to μ), although its interpretation as the waiting time is not true anymore. So the probability of an inversion equals the probability of observing a negative difference.

For $m = 1$, we are interested in the difference D between the largest spike of beat B_k which is $Z_{(\tilde{P}_1, \tilde{P}_1)}$ and the smallest spike of next firing beat B_{k+G} which is $Z_{(1, \tilde{P}_{k+1})[G]}$. The density of this difference (which we denote as $f_D(x)$) can be decomposed by conditioning in the number of spikes of beat B_k , the number of trails G till the next beat fires at least one spike and in the number of spikes of beat B_{k+G} . So it is given as

$$f_D(x) = \frac{1}{1 - p_0} \sum_{i \in \mathbb{N}} p_i f_D(x | \tilde{P}_k = i) \quad (3.56)$$

$$= \frac{1}{1 - p_0} \sum_{i, g \in \mathbb{N}} p_i p_0^{g-1} (1 - p_0) f_D(x | \tilde{P}_k = i, G = g) \quad (3.57)$$

$$= \frac{1}{1 - p_0} \sum_{i, j, g \in \mathbb{N}} p_i p_j p_0^{g-1} f_D(x | \tilde{P}_k = i, G = g, \tilde{P}_{k+G} = j) \quad (3.58)$$

$$= \frac{1}{1 - p_0} \sum_{i, j, g \in \mathbb{N}} p_i p_j p_0^{g-1} f_{\tilde{W}_{(i, j)[g]}}(x). \quad (3.59)$$

Because of this explanations the probability of having a general spike inversion is given as

$$\mathbb{P}(\eta_0) = \int_{-\infty}^0 f_{W|m=0}(x) dx, \quad \text{for } m = 0 \quad (3.60)$$

and

$$\mathbb{P}(\eta_0) = \int_{-\infty}^0 f_D(x) dx, \quad \text{for } m = 1. \quad (3.61)$$

Remark 3.24 (The inversion probability as a measure of irregularity): One could assume that $\mathbb{P}(\eta_0)$ would be possible measure for irregularity in spike trains, because regular firing patterns should have a small inversion probability and irregular spike trains a large one. Unfortunately, if the parameter γ is decreasing, the inversion probability is decreasing too, but at least in the non-bursty case we would expect an increase of irregularity. For that reason $\mathbb{P}(\eta_0)$ is not used as a measure of irregularity.

Remark 3.25 (ISI distribution in the irregular case): When $\mathbb{P}(\eta) > \varepsilon$ the occurrence of spike inversions cannot be ignored and the distribution of \tilde{W}_k will not be a good approximation of the distribution of W_k . Of course one could try to derive the ISI distribution in the same way as for the regular bursty spike train by successive conditioning,

but this would be very complex, because the number of cases dramatically increases. So due to spike inversions, we propose to use simulations of large spike trains to investigate the shape of the ISI distributions in the irregular case. Figure 3.8 shows realizations with $n = 100.000$ spikes. With increasing σ_2 , the ISI distribution tends towards an exponential distribution which is the result already mentioned in Section 2.2.4.

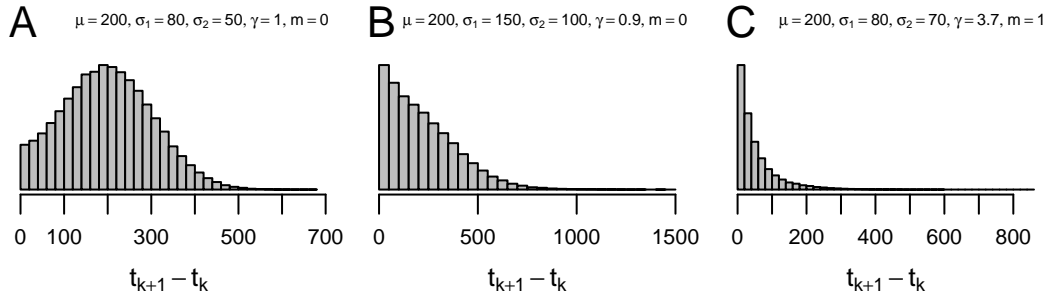


Figure 3.8: Simulated ISI distributions ($n = 100.000$, grey histograms) for different parameter settings in the irregular case.

Because the ISI distribution cannot be handled analytically in the general GLO case (i.e. for all possible parameter vectors ψ), it cannot be used for parameter estimation. For that reason, we introduce the autocorrelation function of a spike train in the next chapter.

Chapter 4.

Autocorrelograms

In the classification of dopaminergic neurons into bursty, irregular and pacemaker cells, a number of studies rely on the shape of the autocorrelation histogram (ACH; Moore et al. (1966); Perkel et al. (1967)), which contains information about burstiness and regularity. For example a peak or the absence of a peak around zero, represents a bursty or a non-bursty spike train and the number of side peaks indicates the regularity (Wilson et al., 1977; Gray et al., 1992; König, 1994; Paladini et al., 2003). The empirical ACH is an estimate for the autocorrelation function (ACF; (Gerstein and Kiang, 1960)) of the process which conditioned on a spike to happen at time x measures the intensities of spikes l time units later.

However, experimental classifications mostly rely on visual inspection or on fitting plausible functions to the observed ACH (Engel et al., 1992; König, 1994; Celada et al., 1999; Hyland et al., 2002; Paladini et al., 2003; Schneider and Nikolic, 2006). While the ISI distributions can only be approximated for some parameter combinations, the ACF of the GLO model provides an explicit analytical solution and can be quickly computed, regardless of the firing mode m and other parameter values. For that reason the ACF and ACH have a central role in this work (see the classification of spike trains and parameter estimation in Chapter 5).

We start with a definition of the ACF, regard its relation to the Palm distribution of the process and derive the ACF of the GLO model (Section 4.1). In the next step the estimate of the ACF, the ACH, is defined (Section 4.2). Finally, we introduce typical visual inspection criteria of the ACF to classify the firing pattern of a spike train and discuss the possibility of using the GLO ACF to classify the firing patterns in an objective way (Section 4.3).

4.1. The autocorrelation function

The autocorrelation function (ACF) represents a second-order property of the counts of the stochastic point process, i.e. refers to the second moments. It is basically an intensity which measures the occurrence of spikes per time unit, conditioned on a spike at a particular time point. In the stationary case which we focus on, this amounts to condition on a spike at 0.

Definition 4.1 (Autocorrelation function): Let Φ be a stationary point process with the properties given in Assumption 1.14 on page 10. The autocorrelation function (ACF) of Φ is defined for lags $l > 0$ as

$$f(l) := \lim_{\delta_1, \delta_2 \rightarrow 0^+} \frac{\mathbb{E}[\Phi(l, l + \delta_1) \mid \Phi(-\delta_2, 0] > 0]}{\delta_1}. \quad (4.1)$$

Lemma 4.2: The limiting value of $f(l)$ (as defined in Definition 4.1) exists.

Proof. This follows directly from Assumption 1.14 and Proposition 1.27. □

In the described GLO, the ACF for lag $l > 0$ can be derived by regarding the conditional probability of a spike in the small interval $I_l := [l, l + \delta_1)$ given a spike in the small interval $(-\delta_2, 0]$. Thus, the ACF can be written down because most intervals result from sums of normal distributions.

Proposition 4.3 (GLO ACF): Let $\Phi \sim \text{GLO}(\psi)$, the ACF of Φ is

$$f_m(l) = \gamma \sum_{i \in K(m)} \varphi_{\{i\mu, |i|\sigma_1^2 + 2\sigma_2^2\}}(l), \quad \text{with } K(m) := \begin{cases} \mathbb{Z}^* & , m = 0 \\ \mathbb{Z} & , m = 1. \end{cases} \quad (4.2)$$

Example 4.4 (ACF construction): Figure 4.1 shows a visualization of the construction of the ACF for both firing modes m . As mentioned they differ only by the summation sets \mathbb{Z} and \mathbb{Z}^* . For computational implementations, one would sum over some finite set. R Code for the construction of the ACF can be found in the appendix under listing A.3 (function `ACF`). Furthermore in Figure 4.1, it is illustrated that $f_m(l)$ converges to the mean rate $\lambda = \gamma/\mu$ for increasing l .

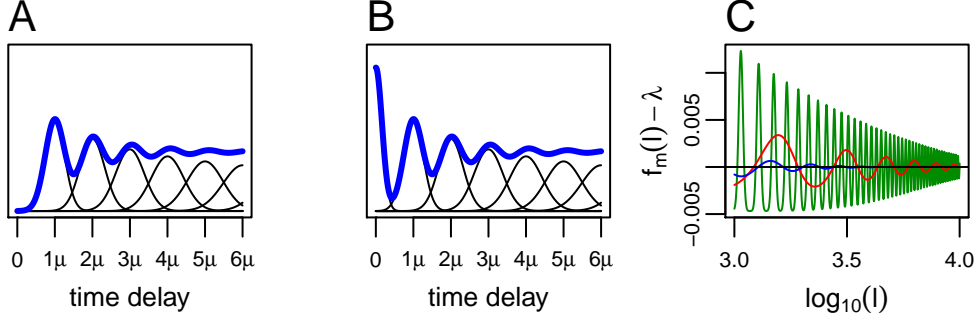


Figure 4.1: Construction of the GLO autocorrelation function (ACF) as the sum of densities from normal distributions for the non-bursty (A, $m = 0$) and bursty firing mode (B, $m = 1$). The two firing modes differ only by the summation sets \mathbb{Z} and \mathbb{Z}^* . Furthermore $f_m(l)$ converges to the mean rate $\lambda = \gamma/\mu$ for increasing l (C).

4.1.1. Proof of Proposition 4.3

According to Definition 4.1 and the random counting measure representation (2.9) on page 21, the ACF of the GLO is defined for $l > 0$ as

$$f(l) = \lim_{\delta_1, \delta_2 \rightarrow 0+} \frac{\mathbb{E}[\Phi(l, l + \delta_1) \mid \Phi(-\delta_2, 0] > 0]}{\delta_1} \quad (4.3)$$

$$= \lim_{\delta_1, \delta_2 \rightarrow 0+} \frac{\mathbb{E}[\sum_{i \in \mathbb{Z}} \sum_{j=1}^{P_i} \delta_{B_i + Z_{i,j}}(l, l + \delta_1) \mid \Phi(-\delta_2, 0] > 0]}{\delta_1}. \quad (4.4)$$

Conditioning on the event $\{\Phi(-\delta_2, 0] > 0\}$ with $\delta_2 \rightarrow 0+$ basically means to condition on the event, that there is a spike at time zero. We further assume that this spike comes from beat B_0 (otherwise we can simply rename the beats) which implies $P_0 > 0$. So P_0 has another distribution than the other P_i (we write \hat{P}_0 instead of P_0), but is still independent of them. For that reason, we split (4.4) into

$$f(l) = \lim_{\delta_1, \delta_2 \rightarrow 0+} \frac{\mathbb{E}[\sum_{j=1}^{\hat{P}_0} \delta_{B_0 + Z_{0,j}}(l, l + \delta_1) \mid \Phi(-\delta_2, 0] > 0]}{\delta_1} + \lim_{\delta_1, \delta_2 \rightarrow 0+} \frac{\mathbb{E}[\sum_{i \in \mathbb{Z}^*} \sum_{j=1}^{P_i} \delta_{B_i + Z_{i,j}}(l, l + \delta_1) \mid \Phi(-\delta_2, 0] > 0]}{\delta_1}. \quad (4.5)$$

We focus on the second summand of (4.5). Because all P_i are independent and independent of all B_j and all Z_{k_1, k_2} ($\forall i, j, k_1 \in \mathbb{Z}$ and $\forall k_2 \in \mathbb{N}$) and because of the linearity of

the conditional expectation, we can write the summand as

$$\lim_{\delta_1, \delta_2 \rightarrow 0+} \frac{\mathbb{E}[P_i] \mathbb{E}[\sum_{i \in \mathbb{Z}^*} \delta_{B_i + Z_{i,j}}(l, l + \delta_1) \mid \Phi(-\delta_2, 0] > 0]}{\delta_1} \quad (4.6)$$

$$= \lim_{\delta_1, \delta_2 \rightarrow 0+} \frac{\gamma \sum_{i \in \mathbb{Z}^*} \mathbb{E}[\delta_{B_i + Z_{i,j}}(l, l + \delta_1) \mid \Phi(-\delta_2, 0] > 0]}{\delta_1} \quad (4.7)$$

$$= \lim_{\delta_1, \delta_2 \rightarrow 0+} \gamma \sum_{i \in \mathbb{Z}^*} \frac{\mathbb{P}(B_i + Z_{i,j} \in (l, l + \delta_1) \mid \Phi(-\delta_2, 0] > 0)}{\delta_1}. \quad (4.8)$$

The conditional probability in (4.8) can be derived. Basically it is the probability that given a spike at time zero which comes from beat B_0 , one spike of beat B_i falls into the interval $(l, l + \delta_1)$. Thus, we are interested in

$$\lim_{\delta_2 \rightarrow 0+} \mathbb{P}(B_i + Z_{i,j} \in (l, l + \delta_1) \mid B_0 + Z_{0,1} \in (-\delta_2, 0]). \quad (4.9)$$

This probability is known and has been computed in the context of the ISI distributions in section 3.2. Furthermore, it is an application of formula 1.9.1 of Liemant et al. (1988) which derives the Palm distribution (see subsection 4.1.2) of a point process resulting of an ancestor process. In our GLO case, the step between spikes and beats is normally distributed with mean 0 and variance σ_2^2 . From beat B_0 to B_i we have a sum of i normal distributions with mean μ and variance σ_1^2 which gives $\mathcal{N}(i\mu, |i|\sigma_1^2)$. And then again a step between beat and spike (compare Figure 3.4 on page 53). Thus we get $\mathcal{N}(i\mu, |i|\sigma_1^2 + 2\sigma_2^2)$. Then, the probability (4.9) can be written with the c.d.f. of the normal distribution as

$$F_{\{i\mu, |i|\sigma_1^2 + 2\sigma_2^2\}}(l + \delta_1) - F_{\{i\mu, |i|\sigma_1^2 + 2\sigma_2^2\}}(l). \quad (4.10)$$

Putting (4.10) into (4.8), we get

$$\lim_{\delta_1 \rightarrow 0+} \gamma \sum_{i \in \mathbb{Z}^*} \frac{(F_{\{i\mu, |i|\sigma_1^2 + 2\sigma_2^2\}}(l + \delta_1) - F_{\{i\mu, |i|\sigma_1^2 + 2\sigma_2^2\}}(l))}{\delta_1} \quad (4.11)$$

$$= \gamma \sum_{i \in \mathbb{Z}^*} F'_{\{i\mu, |i|\sigma_1^2 + 2\sigma_2^2\}}(l) \quad (4.12)$$

$$= \gamma \sum_{i \in \mathbb{Z}^*} \varphi_{\{i\mu, |i|\sigma_1^2 + 2\sigma_2^2\}}(l). \quad (4.13)$$

Finally, there is still the first summand in (4.5). In the non-bursty case ($m = 0$), the expectation is zero, because given a spike at beat B_0 which is placed at time zero, there cannot be a second spike from this beat. In the bursty case ($m = 1$), the length biased distribution \hat{P}_0 of P_0 has to be considered. As mentioned in the proof of Proposition 3.20 in subsection 3.3.2.1, the length biased distribution is given by

$$\mathbb{P}(\hat{P}_0 = k) = \frac{k}{\gamma} \mathbb{P}(P_0 = k), \quad \text{for } k \in \mathbb{N}. \quad (4.14)$$

Thus the expectation is $\mathbb{E}[\hat{P}_0] = 1 + \gamma$. With one spike conditioned at time zero, there are in the mean γ spikes which are independent of that spike. The further procedure is the same as before, thus we can rewrite the first summand in (4.5) as

$$\lim_{\delta_1, \delta_2 \rightarrow 0+} \gamma \frac{\mathbb{P}(B_0 + Z_{0,j} \in (l, l + \delta_1) \mid \Phi(-\delta_2, 0] > 0)}{\delta_1} = \gamma \varphi_{\{0, 2\sigma_2^2\}}(l). \quad (4.15)$$

The last step follows, because given a spike at time zero we have a normally distributed step with mean 0 and variance σ_2^2 back to beat B_0 and again a step to a spike with the same distribution. Taking both summands in (4.5) together, the ACF of the GLO can be written in the form as given in (4.2).

4.1.2. Connection to the Palm distribution

Because Φ is a simple point process and as the interval $(l, l + \delta_1)$ gets sufficiently small, the probability to have more than one spike in this interval becomes also small. Thus, there should only be 0 or 1 spike in the interval $(l, l + \delta_1)$, so that we have

$$\mathbb{E}[\Phi(l, l + \delta_1)] \approx \mathbb{P}(\Phi(l, l + \delta_1) = 1) \approx \mathbb{P}(\Phi(l, l + \delta_1) \geq 1). \quad (4.16)$$

Thus we can write

$$f(l) \approx \lim_{\delta_1, \delta_2 \rightarrow 0+} \frac{\mathbb{P}(\Phi(l, l + \delta_1) \geq 1 \mid \Phi(-\delta_2, 0] > 0)}{\delta_1} \quad (4.17)$$

or equivalently

$$f(l) \approx \lim_{\delta_1, \delta_2 \rightarrow 0+} \frac{\mathbb{P}(\Phi(l, l + \delta_1) = 1 \mid \Phi(-\delta_2, 0] > 0)}{\delta_1}. \quad (4.18)$$

The numerator of the fraction in (4.17) is called a Palm probability of the stationary point process (König and Schmidt, 1992).

Definition 4.5 (Palm distribution): The Palm distribution of a stationary point process is defined as

$$P^0(A) = \frac{1}{\lambda} \int \int_{M_c [0,1]} 1_A(\mathcal{T}_x \varphi) \varphi(dx) \mathbb{P}(d\varphi), \quad A \in \mathcal{B}M_c, \quad (4.19)$$

where $1_A(x)$ is an indicator function and \mathcal{T}_x is the shift operator which for every $x \in \mathbb{R}$ is a measurable mapping with $\mathcal{T}_x: M_c \rightarrow M_c$, such that every counting measure φ with $\varphi = \sum_{t \in S_\varphi} \varphi(\{t\}) \delta_t$ is related to the counting measure $\mathcal{T}_x \varphi = \sum_{t \in S_\varphi} \varphi(\{t\}) \delta_{t-x}$.

There is the following local characterization of the palm distribution:

Proposition 4.6 (König and Schmidt (1992), p.135): For every $k \in \mathbb{N}$ and every sequence $\{(a_i, b_i] \mid i = 1, \dots, k\}$ of pairwise disjoint intervals with the property

$$P^0(\varphi: \varphi(\{a_i\}) > 0 \text{ or } \varphi(\{b_i\}) > 0) = 0 \quad \text{for } i = 1, \dots, k$$

and every sequence of intervals $\{I_n\}$ with $0 \in I_n$ and $\lim_{n \rightarrow \infty} \nu(I_n) = 0$ and every k -tuple (j_1, \dots, j_k) of positive integers, there is a convergence of

$$P^0(A) = \lim_{n \rightarrow \infty} \mathbb{P}(A \mid \{\varphi: \varphi(I_n) > 0\}) \quad (4.20)$$

for the set

$$A = \bigcap_{i=1}^k \{\varphi: \varphi((a_i, b_i]) = j_i\}. \quad (4.21)$$

Thus, Proposition 4.6 shows the connection of (4.18) to the Palm distribution. It is

$$P^0(\Phi(l, l + \delta_1) = 1) = \lim_{\delta_2 \rightarrow 0^+} \mathbb{P}(\Phi(l, l + \delta_1) = 1 \mid \Phi(-\delta_2, 0] > 0). \quad (4.22)$$

4.2. The autocorrelation histogram

The autocorrelation histogram (ACH) is an empirical estimate of the ACF and thus of the intensity function, conditioned on a point at zero. The ACH uses a discrete binning for the time axis and basically counts the number of spike pairs having a particular difference according to this binning. The calculation of the ACH is straightforward, but obviously depends on the chosen discretization. So usually, the first step is to define an analysis window which should be large enough to capture the main dependencies.

Definition of the analysis window: Let t_1, t_2, \dots, t_n denote the empirical spike times, l_b the beginning and l_e the end of the analysis window of the ACH, with $l_e > l_b \geq 0$. Let K be a positive integer representing the number of intervals to partition the interval $(l_b, l_e]$, then the bin size is given as $\delta := (l_e - l_b)/K$. But in computational practice, δ is sometimes predefined and l_e chosen according to δ and K . A partition of $(l_b, l_e]$ is given by the set of intervals

$$L_{l_b, l_e}^\delta := \{L_1, L_2, \dots, L_K\}, \quad (4.23)$$

where L_k is an interval of length δ , i.e.

$$L_k := (l_b + (k - 1)\delta, l_b + k\delta] \quad (4.24)$$

with $k = 1, \dots, K$. Furthermore, each interval has a centre

$$l_k := l_b + k\delta - \delta/2 \quad (4.25)$$

which can be collected in the set

$$\mathcal{L}_{l_b, l_e}^\delta := \{l_1, l_2, \dots, l_K\}. \quad (4.26)$$

Obviously both sets have the same number of elements - we have K intervals and K interval centers. $\mathcal{L}_{l_b, l_e}^\delta$ is called the analysis window of the ACH. In the following, we will drop the indices and write \mathcal{L} and L for short.

Definition of the ACH: Given an analysis window \mathcal{L} , the counts \tilde{h} of the ACH are defined as a function of the interval centers given by \mathcal{L} , according to

$$\tilde{h}(l_k) = \tilde{h}_L(l_k) := \#\{(t_i, t_j) \mid t_j - t_i \in L_k, i, j \in [n]\}, \quad \forall l_k \in \mathcal{L}. \quad (4.27)$$

Thus for every interval, there is an interval center and a value for \tilde{h} . In order to make the ACH and the ACF comparable, we need to standardize the ACH by the recording time \mathcal{T} of the spike train, the number of spikes n and the bin size δ used in the ACH:

$$h(l) := \frac{\mathcal{T} \cdot \tilde{h}(l)}{n \cdot \delta \left(\mathcal{T} - l\delta - \frac{1}{2}\delta\right)}, \quad \forall l \in \mathcal{L}. \quad (4.28)$$

$h(l)$ can then be used as an estimate of the intensity function $f_m(l)$ at lag l (Cox, 1965). R Code for the construction of the ACH can be found in appendix A.3. The ACH is typically visualized by plotting the pairs $(l_k, h(l_k))$, $k = 1, \dots, K$, as a histogram or the points connected by lines. Figure 1.3 shows four ACHs from the empirical spike trains shown in Figure 1.1. The analysis window uses the following options: $l_b = 0$, $l_e = 1500$ ms and $\delta = 10$ ms. Oscillatory activity in the spike train, is represented by the presence of one or more side peaks. An initial peak near 0, reflects many small intervals which may indicate bursty spike trains.

Remark 4.7 (Observations regarding the ACH): Definition (4.28) uses the empirical observed spike times to generate the ACH. If we switch to random variables T_i and T_j , the ACH will become a random variable and we write $H(l)$ instead of $h(l)$. In the context of the GLO, we want to mention the following observations for these random ACHs: The variance of the ACH at lag l is proportional to the value of $f_m(l)$, i.e.

$$\text{Var}(H(l)) \approx f_m(l)/n, \quad \forall l \in \mathcal{L}. \quad (4.29)$$

Furthermore, with increasing number of spikes n , the bin size δ can be reduced and the difference between $H(l)$ and $f_m(l)$ is approximately zero, i.e.

$$\lim_{n \rightarrow \infty} \lim_{\delta \rightarrow 0} |H(l) - f_m(l)| \approx 0. \quad (4.30)$$

Although these observations are not proven, they are illustrated in Figure 4.2 by simulation.

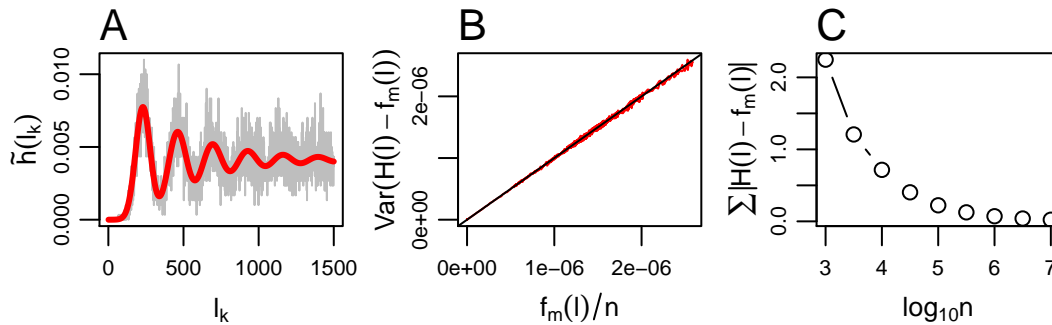


Figure 4.2: Properties of the autocorrelation histogram (ACH). (A) The ACH (grey) and ACF (red) of a simulated GLO spike train with parameters $\mu = 230$, $\sigma_1 = 40$, $\sigma_2 = 20$, $\gamma = 0.95$, $m = 0$ and $n = 3.000$ spikes. (B) The variance of $h(l) - f_m(l)$ is plotted against $f_m(l)/n$. (C) The sum of absolute error over the analysis window decreases for increasing n .

4.3. Describing burstiness and regularity

As well as the shape of the ISI distribution, the shape of the ACF and ACH contain information about burstiness and regularity. For example a peak or the absence of a peak around zero, represents a bursty or a non-bursty spike train and the number of side peaks indicates the regularity (Wilson et al., 1977; Gray et al., 1992; König, 1994; Paladini et al., 2003). However, experimental classifications mostly rely on visual inspection or on fitting plausible functions to the observed ACH (Engel et al., 1992; König, 1994; Celada et al., 1999; Hyland et al., 2002; Paladini et al., 2003; Schneider and Nikolic, 2006).

In this section, we start with a short description of the usual visual inspection criteria of the ACH for classification of a spike train (Section 4.3.1). In Section 4.3.2 we describe a number of possibilities that the ACF of a GLO process offers to grasp burstiness and regularity. These measures that are directly related to the shape of the ACF are thus

closely related to classification measures based on visual inspection of the ACF (Wilson et al., 1977; Paladini et al., 2003).

4.3.1. Visual inspection

The visual classification of firing patterns into bursty or single spike processes and into regular/oscillatory or irregular patterns focuses on two criteria (Wilson et al., 1977; Paladini et al., 2003):

1. A spike train is called bursty, if there is a clear narrow peak near zero in the ACH. It is called non-bursty, if there is not a narrow peak. Inspection of burst discharges in the raw spike train can complement this ACH-based classification. A peak that occurs directly after the refractory period in the ACH is termed 'first' or 'central peak'.
2. A non-bursty process with single spikes is called oscillatory or regular if there are at least 3 'visible' peaks in the ACH and irregular if there are less than 3 peaks. A bursty process is termed oscillatory or regular bursty, if the first peak is followed by a clear trough and repetitive peaks, otherwise it is called irregular bursty.

Table 4.1 summarizes these classification criteria.

Example 4.8: Figure 1.3 shows four ACHs of empirical spike trains. The first one (A) would be clearly classified as regular non-bursty because it contains many side peaks. The second ACH (B) has a clear central peak and trough with side peak, which classifies it as a regular bursty neuron. For the third one (C) it is less clear cut to decide for regular or irregular, but for sure it is more irregular than (B). In the last graphic the ACH is almost a line, so the spike train is very irregular, but it is not sure if there is an initial peak or not.

Obviously, these method for classification is a very subjective criteria, because sometimes it is not clear cut if there is a peak or not and different observers may have a different classification for the same ACH. For that reason, we want to introduce objective and reproducible criteria to classify spike trains into bursty or non-bursty and regular or irregular. These criteria will turn out to be highly connected with the criteria of visual inspection.

4.3.2. GLO criteria

The GLO classification criteria aims at describing the visual inspection criteria in an objective way and is based on the model ACF which depends on the parameters μ , σ_1 , σ_2 , γ and m .

4.3.2.1. Burstiness

The question, if there is a central peak in the ACH, is answered by the parameter m . If $m = 0$, there is no central peak, for $m = 1$ there is a central peak (compare equation (4.2)). So per definition of the GLO, the spike train is termed bursty if $m = 1$ and non-bursty if $m = 0$. If the spike train is bursty, the burstiness is given by γ .

4.3.2.2. Regularity and Irregularity

In order to measure the regularity of a neuron, one should note that in a regular spike train the background rhythm with period μ should be easily detectable. Thus, a regular spike train should have clear side peaks. The regularity is identified by the height of the first side peak in the ACH in relation to the baseline level γ/μ of the ACH. If the peak gets higher relative to the baseline level, the spike train will be more regular. Because in the GLO the first side peak is located at μ , the height of the first side peak is given by $f_m(\mu)$. Because $\varphi_{\{\mu, \sigma_1^2 + 2\sigma_2^2\}}(\mu) \geq \varphi_{\{k\mu, |k|\sigma_1^2 + 2\sigma_2^2\}}(\mu) \forall k \in \mathbb{Z}^*$, the first side peak is typically dominated by the summand $\varphi_{\{\mu, \sigma_1^2 + 2\sigma_2^2\}}(\mu)$. So reducing the height to this main contribution and removing constants, the height is given by $\gamma/\sqrt{\sigma_1^2 + 2\sigma_2^2}$. So relative to the baseline we get

$$\frac{\mu}{\sqrt{\sigma_1^2 + 2\sigma_2^2}} \quad (4.31)$$

as a simplified measure for the regularity of a spike train (compare fig. 4.3 A). This has the intuitive interpretation that the regularity decreases if the variances increases. The irregularity (reciprocal of the regularity) is given by

$$\beta_1 := \frac{\sqrt{\sigma_1^2 + 2\sigma_2^2}}{\mu}. \quad (4.32)$$

Figure 4.3 B shows a 3d plot of σ_1/μ , σ_2/μ and β_1 for different parameter combinations.

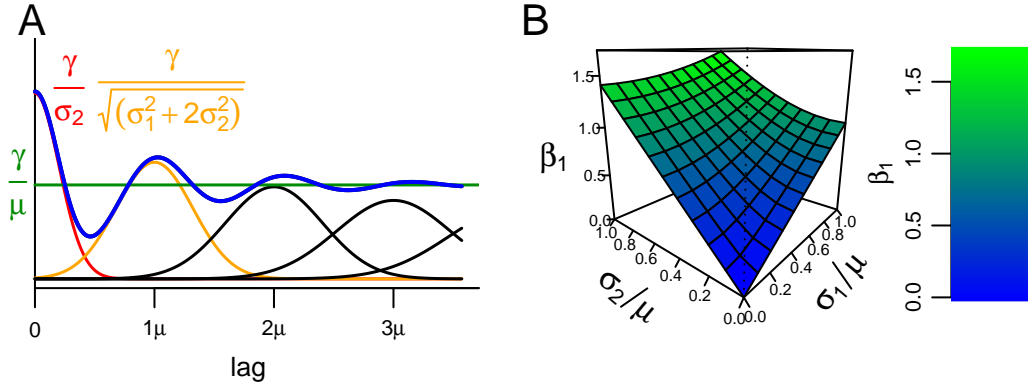


Figure 4.3: (A) Visualization of regularity and burstiness in the ACF (green). Burst identifiability is given by the height of the central peak, which is approximated by $\varphi_{\{0,2\sigma_2^2\}}(0)$ (red), relative to baseline γ/μ (green). Regularity is represented by the height of the first side peak, which is approximated by $\varphi_{\{\mu,\sigma_1^2+2\sigma_2^2\}}(\mu)$ (orange), relative to baseline. (B) The value of β_1 for different ratios of σ_1/μ and σ_2/μ .

4.3.2.3. Burst identifiability

Similar to the description of irregularity by the shape of the first side peak, the bursty firing mode ($m = 1$) can be identified easily if the central peak is high relative to baseline in the ACF (compare Figure 4.3). The height of the central peak is given by $\lim_{l \rightarrow 0^+} f_m(l)$ and dominated by the term $\varphi_{\{0,2\sigma_2^2\}}(0)$. Although the ACF is defined for $l > 0$, we use $l = 0$ as an approximation for the central peak. Disregarding constants, the burst identification thus basically depends on

$$\beta_2 = \frac{\mu}{\sigma_2}, \quad (4.33)$$

which is the reciprocal of the relative burst width (with increasing relative burst width, bursts are less detectable, because the spikes are placed over a wider range). So the parameter β_2 is useful to separate extremely irregular bursters from the group of highly and moderately regular bursters.

4.3.2.4. Classification

The parameters m , β_1 and β_2 are used to classify firing patterns as bursty or non-bursty and regular or irregular. While m is discrete with values 0 or 1, β_1 and β_2 have values in $(0, \infty)$. Although we know that there is a continuous change for regularity and burst

identifiability, there is also a practical need for a clear classification. So for β_1 and β_2 thresholds have to be used to classify spike trains. There are the following rules

- If β_2 is too small, this means that the variance σ_2^2 is too high in comparison with μ . It follows that regardless of m , spikes of different beats cannot be distinguished from each other which contradicts the idea of a burst as clusters of spikes. So we will regard all spike trains with $\beta_2 \leq c_2$ as non-bursty and irregular. Where c_2 is a predefined threshold and depends on the data set, brain region and the understanding of irregularity of the investigator.
- If $\beta_2 > c_2$, it depends on m , if the spike train is called bursty ($m = 1$) or non-bursty ($m = 0$).
- if $\beta_2 > c_2$, it depends on β_1 , if the spike train is classified as regular or irregular. Here again thresholds are used which we call c_1 for $m = 0$ and \tilde{c}_1 for $m = 1$. Firing patterns with a β_1 larger than these thresholds are called irregular. If β_1 is smaller than the threshold, they are called regular. The reason why there are two thresholds for β_1 is that bursty spike trains will typically have larger β_2 values than non-bursty spike trains. The final choice of the thresholds also depends on the dataset and the investigator.

These classification rules are an improvement of the usual visual inspection criteria, because they can be implemented in an automatic way and give the same result for the same spike train. The GLO classification rules are published by Bingmer et al. (2011). A summary can be found in table 4.1.

| Classification | Visual inspection of ACH | GLO parameters |
|-------------------------|--|--|
| regular non-bursty | no peak at zero time lag ≥ 3 side peaks | $m = 0, \beta_1 \leq c_1, (\beta_2 > c_2)$ |
| irregular non-bursty | no peak at zero time lag < 3 side peaks | $m = 0, \beta_1 > c_1$ (OR $\beta_2 \leq c_2$) |
| regular bursty | peak at zero time lag followed by trough and side peaks | $m = 1, \beta_1 \leq \tilde{c}_1, (\beta_2 > c_2)$ |
| irregular bursty | peak at zero time lag without trough and side peaks | $m = 1, \beta_1 > \tilde{c}_1, (\beta_2 > c_2)$ |

Table 4.1: Summary of classification rules for burstiness and regularity based on visual inspection and on the GLO parameters. Thresholds for c_1 and \tilde{c}_1 should be chosen differently for bursty and single spike processes.

Example 4.9: In Figure 4.4, the ACHs and corresponding ACFs are illustrated for

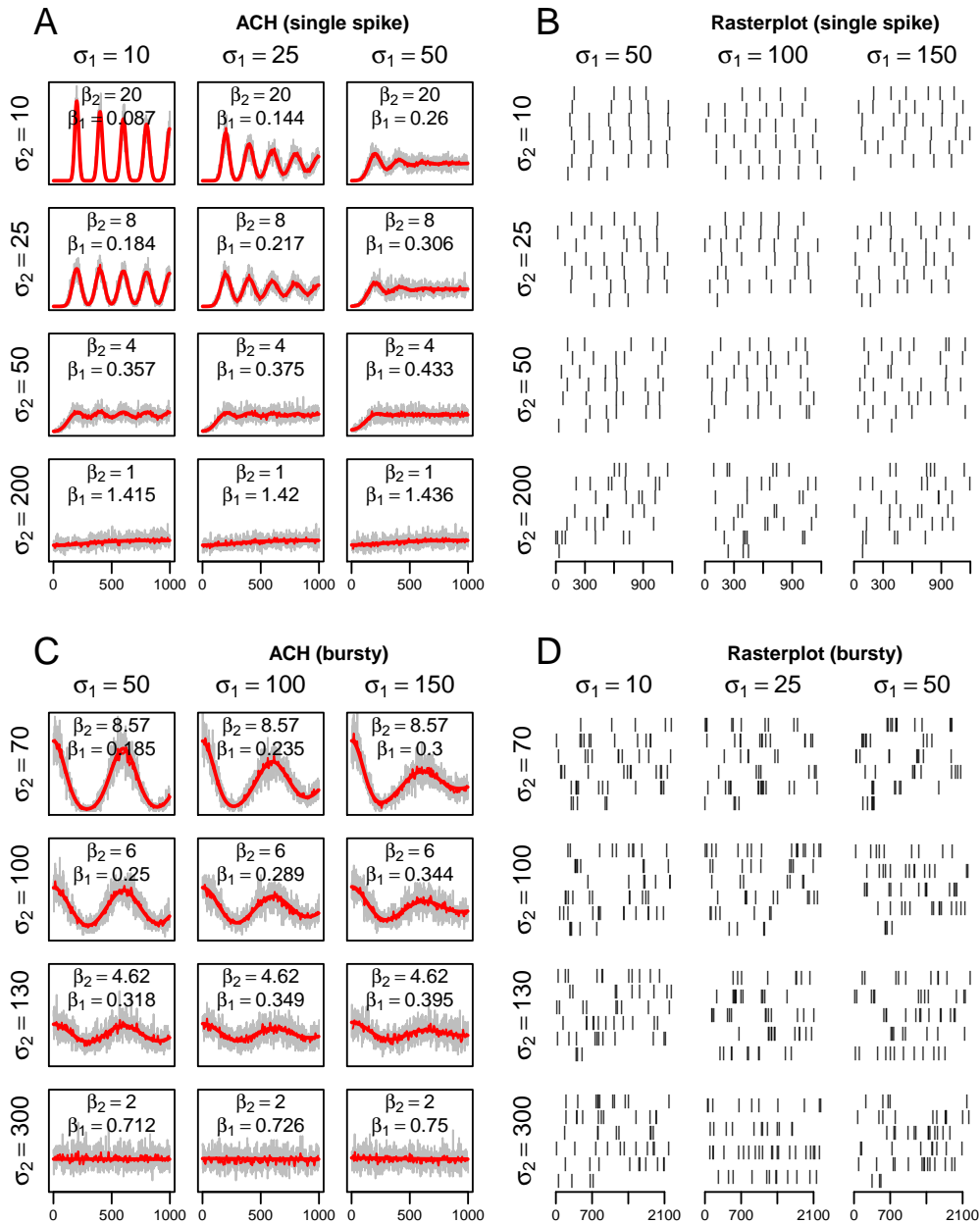


Figure 4.4: Illustration of β_1 and β_2 in combination with ACFs and rasterplots. GLO spike trains simulated with $n = 3,000$ spikes and parameters $\mu = 200$, $\gamma = 0.95$, $m = 0$ (A,B) and $\mu = 600$, $\gamma = 3$, $m = 1$ (C,D). Different values of σ_1 and σ_2 are used to illustrate β_1 and β_2 and the corresponding firing patterns. (A, C) Raw ACHs (gray) and theoretical ACFs (black). (B, D) Rasterplots of the spike trains.

different parameter combinations of the GLO. As described, the only difference between the bursty and non-bursty mode is the peak around zero. For $l \rightarrow \infty$, the ACFs converge to the mean firing rate γ/μ . Increasing σ_1 results in a decrease in oscillation strength of the background rhythm. If σ_2 increases, successive bursts are less separated and thus, all ACH peaks decrease towards the background level γ/μ . Furthermore β_1 and β_2 are given to illustrate the effect of a change of parameters. If the variances increase, β_1 increases as well. In contrast, β_2 only increases if σ_2 increases.

Chapter 5.

Parameter estimation

In this chapter, a technique is presented that uses the empirical autocorrelation histogram (ACH; Section 4.2) to estimate the parameters μ , σ_1 , σ_2 , γ and m of the GLO model. Basically, the GLO autocorrelation function (ACF; Section 4.1) is fitted to the ACH by a nonlinear least squares algorithm.

There are several advantages by choosing such a procedure: In contrast to the ISI distributions, the ACF can be handled mathematically and be quickly computed – even in the irregular case. So inversions of beats (Section 2.2.1) or spikes (Section 3.1) are not a problem. Furthermore, the ACF contains information about burstiness and regularity (Section 4.3) and the concept of fitting functions to the ACH is very well known in neuroscientific practice (Engel et al., 1992; König, 1994; Celada et al., 1999; Hyland et al., 2002; Paladini et al., 2003; Schneider and Nikolic, 2006; Schneider, 2008). But these articles usually use arbitrary functions (e.g. cosine) which are not based on an underlying point process model (except Schneider (2008)).

In the beginning of the chapter (Section 5.1), there will be a short overview of the fitting procedure and every step of it will be described in detail (analysis window, nonlinear regression, weights and starting values). Then (Section 5.2), the precision of estimation is regarded in context of different bin sizes, recording times and parameter combinations. At the end of the chapter (Section 5.3), for four exemplary parameter combinations, the residuals of the nonlinear regression model are investigated regarding the usual assumptions (homogeneous variances, normal distributed and independence of residuals).

5.1. Overview of fitting procedure

Given a spike train $\mathcal{S} = (t_1, \dots, t_n)$, the goal is to get estimators for the parameters μ , σ_1 , σ_2 , γ and m . This is done by fitting the model ACF to the empirical ACH with the following procedure: In a first step, one has to determine the analysis window, i.e. the values l_b , l_e and δ for \mathcal{L} , to compute the empirical normalized ACH h (see equation (4.28)). Then conditioned on the firing mode m , the theoretical ACF f_m of the GLO model is fitted to h by minimizing the weighted residual sum of squares (WRSS)

$$\sum_{l \in \mathcal{L}} v_l (h(l) - f_m(l))^2. \quad (5.1)$$

The weights v_l are introduced according to the observation that the error variance in the ACH at lag l seems to equal $f_m(l)$ up to a constant, see Remark 4.7. Therefore, we propose to use an estimate $\hat{f}_m(l)$ of $f_m(l)$ obtained by smoothing the ACH with a Gaussian filter. Different starting values $\tilde{\psi} = (\tilde{\mu}, \tilde{\sigma}_1, \tilde{\sigma}_2, \tilde{\gamma}, \tilde{m})$ are used to compute (5.1). The parameter combination which minimizes WRSS is used to fit the theoretical ACF to the empirical ACH by an iterative nonlinear least squares procedure based on a Gauss Newton algorithm implemented in the statistical package R (Chambers, 2008), i.e. using the R function `nls()`. Finally this procedure yields the estimates $\hat{\mu}$, $\hat{\sigma}_1$, $\hat{\sigma}_2$, $\hat{\gamma}$ and \hat{m} .

Remark 5.1: Summary of parameter estimation procedure:

1. Determine analysis window \mathcal{L} (Section 5.1.1).
2. Compute the weights v_l by smoothing the ACH with a Gaussian filter (Section 5.1.3).
3. Derive a set of starting values from the spike train (Section 5.1.4).
4. The starting value combination which minimizes equation (5.1) is used for the nonlinear weighted least squares algorithm to compute the estimates $\hat{\mu}$, $\hat{\sigma}_1$, $\hat{\sigma}_2$, $\hat{\gamma}$ and \hat{m} (Section 5.1.2).

Code of the estimation procedure (function `estimation`) can be found in appendix A.3. Different fitted ACHs can be found in Figure 5.1 B and C, Figure 5.9 and Figure 5.5, where the empirical ACH is given in gray and the fitted ACF of the GLO in blue.

5.1.1. Analysis window

The choice of the analysis window for the ACH, i.e. the set \mathcal{L} , is crucial for the fitting procedure. A poorly chosen \mathcal{L} eventually would not let the algorithm converge and there are a few difficulties to consider for choosing l_b , l_e and δ :

- l_b : Neurons have a short period after firing at which they are not able to fire a spike again, called the refractory time or period. In the GLO, due to its construction, there is not any refractory time and spikes can be arbitrarily close together. This issue is important for bursty GLO spike trains and can be seen by comparing the lags in the ACH close to zero for bursty GLO spike trains (Figure 4.4 C) and for empirical observed ones (Figure 1.3 B and C). While the empirical ACHs tend to zero for small lags, the model ACHs tend to a strictly positive constant. So setting $l_b = 0$ would result in large quadratic deviations for lags close to zero.
- l_e : It should be large enough to account for the main dependencies (in particular the dominant oscillations in the ACH), but not too large, because the asymptotic behavior of the ACH would dominate these dependencies.
- δ : If δ is chosen too small, the ACH will be noisy and the estimation procedure will be slowed down. On the other hand, a bin size chosen too large may smooth out important dependencies. Typically the choice of δ depends on the number of spikes in a spike train and how the data is recorded.

So, because the number of visible peaks depends on the size of μ , as well as the refractory period is different for each cell, the analysis window \mathcal{L} for the ACH is chosen as a function of the individual spike train. The following guidelines work well with the present data set, but it may be necessary to modify them for neurons of other brain regions.

The choice of the analysis window: The bin size is set to $\delta = 10$ milliseconds (ms). Given the spike train $\mathcal{S} = (t_1, \dots, t_n)$, the ISI intervals $\mathbf{w} = (w_1, \dots, w_{n-1})$ are derived. Because the coefficient of variation of the ISIs is roughly able to discriminate between bursty and non-bursty neurons (see Section 3.1.1), the minimal time lag l_b is chosen depending on the CV:

- For $\hat{c}_v(\mathbf{w}) \leq 0.4$, the ISIs are very regular and we use

$$l_{b_1} := \min\{100, q_{0.05}(\mathbf{w})\}, \quad (5.2)$$

with $q_{0.05}(\mathbf{w})$ denoting 5 per cent quantile of the empirical ISI distribution.

- For $\hat{c}_v(\mathbf{w}) > 1.2$, the position of the maximum of a kernel density estimation is used. For the empirical waiting time distribution, a gaussian kernel density estimate (Silverman, 1986; Wand and Jones, 1995; Simonoff, 1996) is computed according to

$$\hat{f}_W(x) := \frac{1}{(n-1)b} \sum_{i=1}^{n-1} K\left(\frac{x-w_i}{b}\right) \quad \text{with} \quad K(\cdot) = \varphi_{\{0,1\}}(\cdot). \quad (5.3)$$

b is called the bandwidth and is estimated by silverman's rule of thumb with

$$\hat{b} := 0.9 \min(\hat{\sigma}, R/1.34)(n-1)^{-1/5} \quad (5.4)$$

where $\hat{\sigma}$ is the estimated standard deviation and R the estimated interquartile range of w_1, \dots, w_{n-1} . The kernel density estimate $\hat{f}_K(x)$ is derived for pre specified values $x \in I$, with

$$I := \{w_{(1)} - 3b, w_{(1)} - 3b + q, w_{(1)} - 3b + 2q, \dots, w_{(n-1)} + 3b\} \quad (5.5)$$

and $q = (w_{(n-1)} - w_{(1)} + 6b)/(511)$. Thus l_{b_2} is given as

$$l_{b_2} := \arg \max_{x \in I} \hat{f}_W(x). \quad (5.6)$$

- For $\hat{c}_v(\mathbf{w}) \in (0.4, 1.2)$, a linear interpolation between l_{b_1} and l_{b_2} is used, i.e.,

$$l_{b_3} := l_{b_1} + \frac{\hat{c}_v(\mathbf{w}) - 0.4}{0.8}(l_{b_2} - l_{b_1}). \quad (5.7)$$

Thus the beginning of the analysis window is defined as

$$l_b := \begin{cases} l_{b_1}, & \hat{c}_v(\mathbf{w}) \leq 0.4 \\ l_{b_2}, & \hat{c}_v(\mathbf{w}) > 1.2 \\ l_{b_3}, & \text{otherwise.} \end{cases} \quad (5.8)$$

The maximum time lag l_e is set to be

$$l_e := l_b + \left\lceil \frac{l_{\bar{e}} - l_b}{\delta} \right\rceil \delta, \quad (5.9)$$

where $l_{\bar{e}}$ is set to be the minimum of 5 times the 90%-quantile of the ISIs and 6000 ms.

In appendix A.3, there is a R function called `Awindow` which derives the analysis window for the function `estimation`.

Example 5.2: Figure 5.1 shows the impact of the analysis window on the fitting procedure. In panel A is the ACH of an empirical spike train and analysis window given by $l_b = 5$, $l_e = 3005$ and $\delta = 10$. Plot B shows a not appropriate fit. Because there is no refractory period in the GLO, it is important to ignore the first lags in the ACH. In plot C, $l_b = 105$ (green line) which results in a better fit of the ACH.

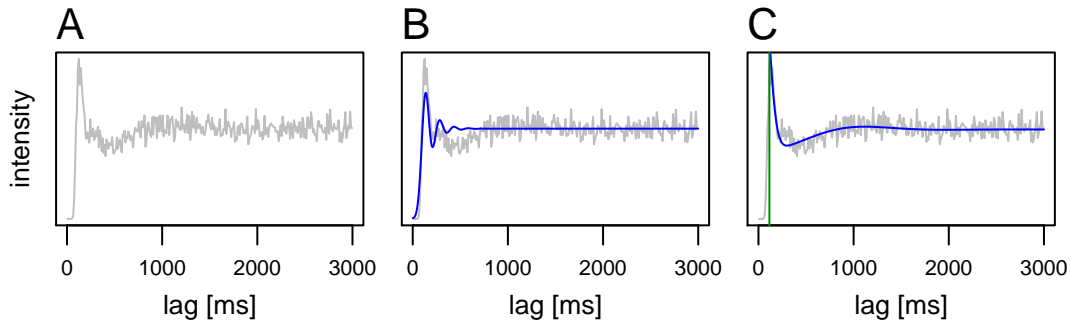


Figure 5.1: The choice of the analysis window and its consequences. (A) Raw ACH of an empirical spike train and analysis window given by $l_b = 5$, $l_e = 3005$ and $\delta = 10$. (B) Fit of the ACH. (C) Again a fit of the same spike train, but with $l_b = 105$ (position of the green line).

5.1.2. Nonlinear weighted regression

We give a short summary of nonlinear weighted regression. More details on nonlinear regression theory can be found in the books by Bates and Watts (1988) and Seber and Wild (1989). Application of these techniques in combination with the statistical software package R is described by Ritz and Streibig (2008).

Summary of the nonlinear regression model: The nonlinear regression model is given by

$$Y_i = f(\theta, \mathbf{x}_i) + \epsilon_i \quad (5.10)$$

where Y_i is a random variable, representing the univariate response of a nonlinear function f of a vector of parameters $\theta = (\theta_1, \dots, \theta_p)$ and a vector of deterministic predictor variables $\mathbf{x}_i = (x_{i1}, \dots, x_{iq})$ for the i -th observation of n . ϵ_i is a random error, assumed to be normally distributed, independently of the errors for other observations, with expectation 0 and constant variance σ^2 .

If Y_1, \dots, Y_n vary around the mean function $f(\theta, \mathbf{x}_i)$ with unequal variance, this can be

considered with the weights $v_i > 0$ and the new model

$$Y_i = f(\theta, \mathbf{x}_i) + \epsilon_i/\sqrt{v_i}. \quad (5.11)$$

The parameter vector θ is typically estimated by minimizing the weighted residual sum of squares (WRSS) which is

$$WRSS(\theta) = \sum_{i=1}^n v_i (y_i - f(\theta, \mathbf{x}_i))^2 \quad (5.12)$$

and equivalent to maximizing the likelihood function

$$L(\theta, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp \left\{ -\frac{\sum_{i=1}^n v_i [y_i - f(\theta, \mathbf{x}_i)]^2}{2\sigma^2} \right\}. \quad (5.13)$$

The resulting estimates $\hat{\theta}$ are called the least-squares parameter estimates.

Example 5.3: (Application to the GLO) The analysis window for the fitting procedure is given as $\mathcal{L} = \{l_1, \dots, l_n\}$, θ represents the parameters $\psi = (\mu, \sigma_1, \sigma_2, \gamma)$, \mathbf{x}_i is given by the lag l_i and thus $f(\theta, \mathbf{x}_i)$ represents the ACF $f_m(l_i | \mu, \sigma_1, \sigma_2, \gamma)$. Note that m has to be predefined. Because y_i represents $h(l_i)$, (5.12) is equivalent to (5.1).

Remark 5.4: We summarize some important comments of Ritz and Streibig (2008): Due to the nonlinearity of f , the minimization of WRSS will in general be a nonlinear problem, so numerical optimization methods are used. These iterative procedures ideally approach the optimal parameter values in a stepwise manner. The most famous algorithm is the Gauss-Newton method, which relies on linear approximations to the nonlinear mean function at each step. Two common complications arise when using numerical optimization:

- How to choose the starting parameter values?
- How to ensure to reach the global minimum rather than a local minimum?

The procedure will usually get closer to the optimal parameter value within a few steps, i.e. the algorithm is said to converge, if the initial parameter values are sufficiently close to the optimal parameter values. So, it is very important to provide sensible starting parameter values. In contrast, poorly chosen starting values will often lead the procedures astray, so no useful model fit is obtained (see Figure 5.2).

The solutions of nonlinear regression problems differ as a consequence of different algorithms, different implementations of the same algorithm (for example, different criteria

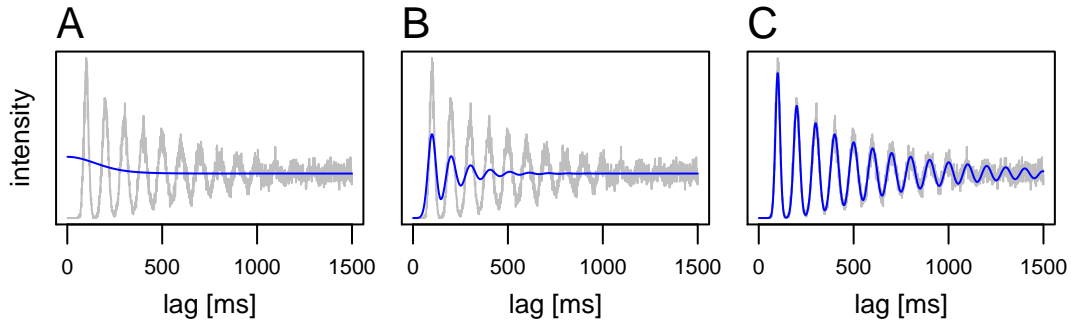


Figure 5.2: Fits resulting of different starting values. Poorly chosen starting values can prevent the nonlinear least squares algorithm to converge or let it converge to a local minimum (plot A and B). On the other hand, starting values chosen near the optimal values, typically will converge to them (C).

for declaring convergence or whether or not first derivatives are computed numerically or explicit expressions are provided), different parameterizations, or different starting values. However, the resulting parameter estimates often will not differ much.

5.1.3. Weights

The values of $h(l)$ do not vary with constant variance around the mean function $f_m(l)$ (compare Section 4.2) and for that reason the weighted nonlinear regression model (5.11) with weights v_l is used to stabilize the variance. Because it seemed that the variance of $h(l) - f_m(l)$ equals $f_m(l)$ up to a constant, it would be naturally to set $v_l := 1/f_m(l)$. Unfortunately $f_m(l)$ is not known in the first instance, so that we have to use an estimate of it. One possibility is to use a smoothed or filtered version of $h(l)$, $l \in \mathcal{L}$. The following procedure can be understood as a suggestion.

Smoothed ACH and weights: So let $\mathcal{L} = \{l_1, \dots, l_K\}$ be the analysis window as derived in Subsection 5.1.1 with $l_1 < l_2 < \dots < l_K$ and $K \in \{7, 8, \dots\}$. To derive $\hat{f}_m(l)$, a weighted moving average with weights according to a normal distribution is applied which is defined as

$$\hat{f}_m(l_i) = \sum_{j=-3}^3 a_j h(l_{i+j}), \quad i \in \{4, \dots, K-3\} \quad (5.14)$$

with

$$a_j = \frac{1}{\sqrt{2\pi}} e^{-\frac{j^2}{2}}. \quad (5.15)$$

Because $\hat{f}_m(l_i)$ cannot be derived for values close to the boundary of \mathcal{L} , we simply set

$$\hat{f}_m(l_i) = h(l_i), \quad i \in \{1, 2, 3, K-2, K-1, K\}. \quad (5.16)$$

Finally, before setting the weights $v_l = 1/\hat{f}_m(l)$, we remove all values of $\hat{f}_m(l)$ which are zero and replace them with the minimum of $\{\hat{f}_m(l), \dots, \hat{f}_m(l)\} \setminus \{0\}$ (otherwise we would have weights equal to infinity).

A R function called `Weights` can be found in appendix A.3.

5.1.4. Starting values

Carefully chosen starting values are crucial for the nonlinear least squares algorithm. In the worst case, the procedure will not converge or will stop in a local minimum (see Figure 5.2). For that reason different parameter combinations are tried as starting values. The more combinations are tried, the higher is the chance that the algorithm converges to the right minimum. Unfortunately more combinations mean longer computational time which is relevant for practical purposes. To save time, we compute the WRSS for all parameter combinations without applying the nonlinear least squares algorithm. It is just started once with the starting value combination which has minimal WRSS. Of course, if time would not play a role, one should use all parameter combinations with the algorithm.

Set of starting values: In the following, some guidelines and rules are presented for the automatic construction of a set of starting values which worked well for the current data set which is described in appendix A.2. Let again $\mathcal{S} = (t_1, \dots, t_n)$ be the empirical observed spike train, $\mathbf{w} = (w_1, \dots, w_{n-1})$ the set of waiting times between spikes, \mathcal{T} the recording time and $\hat{c}_v(\mathbf{w})$ the empirical coefficient of variation of \mathbf{w} . The starting values are given as follows:

- **The initial starting value for μ :** It is derived by two procedures:
 1. In the non-bursty regular case ($m = 0$) μ should be located at the peak of the ISI distribution (compare Figure 5.3). So for the empirical waiting time distribution, a gaussian kernel density estimate is used (compare Section 5.1.1) and thus the starting value for μ is given as

$$\tilde{\mu}_0 := \arg \max_{x \in I} \hat{f}_W(x). \quad (5.17)$$

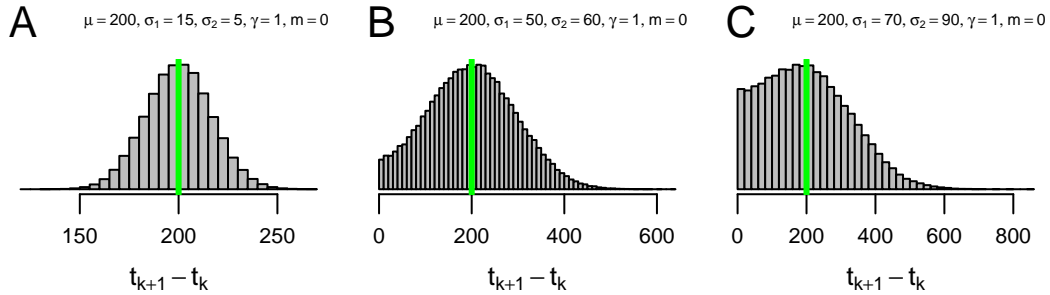


Figure 5.3: Starting value for μ in the non-bursty case. μ (green) is located at the peak of the ISI distribution. These holds true for moderate irregular cases (B and C).

- For $m = 1$, μ is estimated as the dominant period in a spectral density estimate denoted by $\tilde{\mu}_1$ (compare Figure 5.4). For this purpose, we set

$$\mathcal{T}_0 := 2^{\lceil \log_2(\lceil \mathcal{T} \rceil) \rceil}, \quad (5.18)$$

$\mathcal{S}_0 := \mathcal{S} \cap [0, \mathcal{T}_0]$ and define

$$x_i := |\{\mathcal{S}_0 \cap (i-1, i]\}|, \quad i = 1, 2, \dots, \mathcal{T}_0. \quad (5.19)$$

For the sequence $(x_1, x_2, \dots, x_{\mathcal{T}_0})$ the dominant frequency is estimated according to the Discrete Fourier Transform (DFT, see Chu (2008)), i.e. we regard

$$d(j/n) := \frac{1}{\sqrt{n}} \sum_{t=1}^n x_t \exp(-2\pi i t j/n), \quad j = 1, 2, \dots, n/2. \quad (5.20)$$

Because very small values of μ are very unlikely in empirical data (as well as very large values), we expect μ to be in the interval $[a, b]$ with $a, b \in \mathbb{R}$ and $a < b$. This means the connected frequency should be in the interval $[1/b, 1/a]$. Thus, when

$$I := \{j/n : j = 1, 2, \dots, n/2, j/n \in [1/b, 1/a]\} \quad (5.21)$$

the starting value for μ is given as

$$\tilde{\mu}_1 := 1/\arg \max_{j/n \in I} |d(j/n)|^2. \quad (5.22)$$

For the current data set, we use $a = 200$ and $b = 3000$. A R function called `spektral` can be found in appendix A.3 which uses the R function `spectrum` to estimate the dominant frequency.

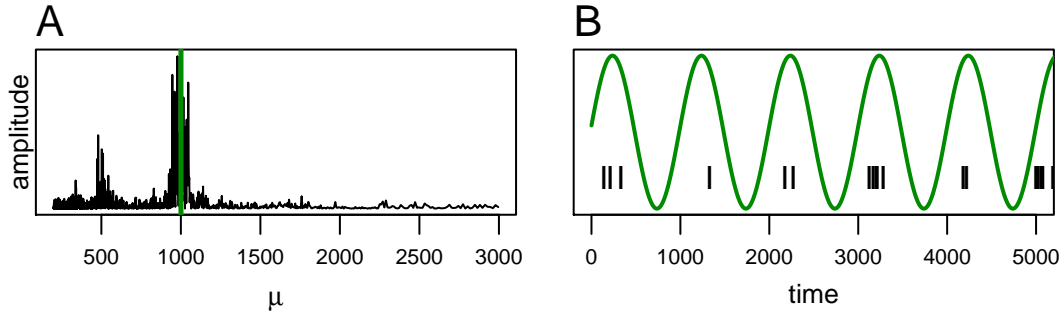


Figure 5.4: Starting value for μ in the bursty case. μ (green) should be the dominant period in the spike train (A and B).

- **The starting value for γ :** We make use of the fact that the mean firing rate of a GLO spike train is given as $\lambda = \gamma/\mu$, see equation (2.53). Given the starting value $\tilde{\mu}_m$ for μ , the initial value for γ is calculated from an estimate of the mean firing rate

$$\tilde{\gamma}(\tilde{\mu}_m) = \hat{\lambda} \cdot \tilde{\mu}_m. \quad (5.23)$$

Because the data may be non-stationary, the obvious estimator $\hat{\lambda} := n/T$ may give wrong results. Instead we use the property that the ACH should converge to the mean firing rate and use the following estimator

$$\hat{\lambda} = \frac{1}{1001} \sum_{l \in \mathcal{L}} h(l) \quad \text{with} \quad \mathcal{L} = \{5000, 5001, \dots, 6000\}. \quad (5.24)$$

- **Starting values for the variances σ_1 and σ_2 :** For σ_1 and σ_2 , a grid of starting values is used. For this purpose we define

$$C_1(x) := \{cx : c \in \{0.01, 0.03, 0.05, 0.1, 0.15, \dots, 0.3\}\} \quad (5.25)$$

$$C_2(x) := \{cx : c \in \{0.01, 0.05, 0.1, 0.15, \dots, 0.75\}\} \quad (5.26)$$

$$C_3(x) := \{cx : c \in \{0.01, 0.04, \dots, 0.25, 0.3, 0.35, \dots, 0.75\}\}. \quad (5.27)$$

The values for $\tilde{\sigma}_1$ and $\tilde{\sigma}_2$ are relative to the estimate $\tilde{\mu}_m$. So we take all values $\tilde{\sigma}_1, \tilde{\sigma}_2 \in C_k(\tilde{\mu}_m)$, where k depends on the CV.

The set of starting values is given as

$$SV := \begin{cases} \tilde{C}_1 & , \hat{c}_v(\mathbf{w}) < 0.4 \\ \tilde{C}_2 & , \hat{c}_v(\mathbf{w}) > 1.2 \\ \tilde{C}_3 & , \text{otherwise.} \end{cases} \quad (5.28)$$

The sets $\tilde{C}_1, \tilde{C}_2, \tilde{C}_3$ represents different sets of starting values and are defined as

$$\tilde{C}_1 := \{(\tilde{\mu}_0, \tilde{\sigma}_1, \tilde{\sigma}_2, \tilde{\gamma}(\tilde{\mu}_0), 0) : \tilde{\sigma}_1, \tilde{\sigma}_2 \in C_1(\tilde{\mu}_0)\} \quad (5.29)$$

$$\tilde{C}_2 := \{(\tilde{\mu}_1, \tilde{\sigma}_1, \tilde{\sigma}_2, \tilde{\gamma}(\tilde{\mu}_1), 1) : \tilde{\sigma}_1, \tilde{\sigma}_2 \in C_2(\tilde{\mu}_1)\} \quad (5.30)$$

$$\begin{aligned} \tilde{C}_3 := & \{(\tilde{\mu}_0, \tilde{\sigma}_1, \tilde{\sigma}_2, \tilde{\gamma}(\tilde{\mu}_0), 0) : \tilde{\sigma}_1, \tilde{\sigma}_2 \in C_3(\tilde{\mu}_0)\} \\ & \cup \{(\tilde{\mu}_1, \tilde{\sigma}_1, \tilde{\sigma}_2, \tilde{\gamma}(\tilde{\mu}_1), 1) : \tilde{\sigma}_1, \tilde{\sigma}_2 \in C_3(\tilde{\mu}_1)\} \\ & \cup \{(\tilde{\mu}_1, \tilde{\sigma}_1, \tilde{\sigma}_2, \tilde{\gamma}(\tilde{\mu}_1), 0) : \tilde{\sigma}_1, \tilde{\sigma}_2 \in C_3(\tilde{\mu}_1)\}. \end{aligned} \quad (5.31)$$

As final starting value combination for the nonlinear least squares algorithm, we take the combination which minimizes WRSS, i.e.

$$\tilde{\psi} := \arg \min_{\psi \in SV} \sum_{l \in \mathcal{L}} v_l (h(l) - f_m(l|\psi))^2. \quad (5.32)$$

In appendix A.3, there is a R function called `Svalues` which can be used to generate starting values.

In the following section, we discuss the behavior of this fitting procedure for different parameter constellations.

5.2. Estimation precision

We start with introducing four representative parameter combinations (Section 5.2.1) which will be used for different kind of simulations. After defining some basic notations, error measures and robust statistics (Section 5.2.2), different simulations are performed to investigate the impact of the recording time \mathcal{T} and the bin size δ of the ACH on the fitting procedure (Section 5.2.3). Furthermore, it is investigated, if the standard errors, which are given as the output of the nonlinear regression procedure, can be used for quantifying the precision of the estimated model parameters (Section 5.2.4). Finally, the fitting procedure is quantified for different representative parameter constellations (Section 5.2.5).

5.2.1. Representative parameter combinations

We take four typical parameter constellations which are found in table 5.1 and represent the four typical firing patterns described by the GLO. These four parameter combinations will be used as representative of their category (non-bursty, bursty, regular, irregular)

and used for further simulations in the current and the next chapter. Their visualizations in rasterplots, ACFs and ACHs can be seen in Figure 5.5.

| combination | μ | σ_1 | σ_2 | γ | m | β_1 | β_2 | type |
|-------------|--------|------------|------------|----------|-----|-----------|-----------|----------------------|
| A | 195.97 | 21.7 | 9.6 | 0.99 | 0 | 0.13 | 20.4 | regular non-bursty |
| B | 1160 | 110 | 170 | 3 | 1 | 0.23 | 6.8 | regular bursty |
| C | 1255 | 500 | 252 | 5.1 | 1 | 0.49 | 5 | irregular bursty |
| D | 190 | 58 | 40 | 0.76 | 0 | 0.43 | 4.75 | irregular non-bursty |

Table 5.1: Exemplary parameter constellations used for simulation studies. They represent the typical firing patterns described by the GLO model. Figure 5.5 shows rasterplots, ACHs and ACFs.

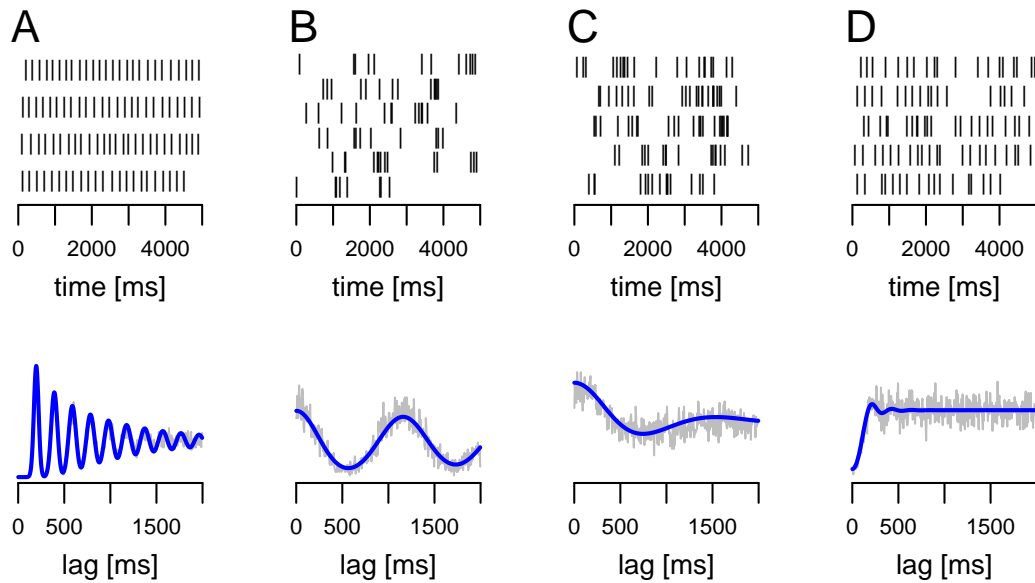


Figure 5.5: Rasterplots, ACH and ACFs of exemplary parameter combinations given in table 5.1. They show the four typical firing patterns: (A) regular non-bursty, (B) regular bursty, (C) irregular bursty and (D) irregular non-bursty.

5.2.2. Basic notations, error measures and robust statistics

The goal of this section is to investigate the estimation precision with simulations under different conditions or settings. A setting/condition is a special parameter combination $\psi = (\mu, \sigma_1, \sigma_2, \gamma, m)$, with a special recording time \mathcal{T} and bin size δ or some other condition. Typically, we have $J \in \mathbb{N}$ different settings and for each setting $n \in \mathbb{N}$ spike

trains are simulated (with the R function `GL0`, see appendix A.3). For each of these spike trains the parameters are estimated again (with R function `estimation`), resulting in

$$\hat{\psi}_{ij} := (\hat{\mu}_{ij}, \hat{\sigma}_{ij,1}, \hat{\sigma}_{ij,2}, \hat{\gamma}_{ij}, \hat{m}_{ij}), \quad i = 1, \dots, n, \quad j = 1, \dots, J. \quad (5.33)$$

Because $\psi_j = (\mu_j, \sigma_{j,1}, \sigma_{j,2}, \gamma_j, m_j)$ represent the true values of setting j , the values $\hat{\psi}_{ij}$ ($i = 1, \dots, n$) can be used to quantify the precision of the estimation procedure. For the quantification different kinds of error measures are used.

5.2.2.1. Error measures

Let $\hat{x} \in \mathbb{R}$ be an estimator for a parameter $x \in \mathbb{R}$ (we assume $\mathbb{E}[|\hat{x}^2|] < \infty$) and \hat{x}_i ($i = 1, \dots, n$) be several realizations of \hat{x} . The *mean squared error (MSE)* is defined as

$$\text{MSE}(\hat{x}) := \mathbb{E}[(\hat{x} - x)^2] = \text{Var}(\hat{x}) + \mathbb{E}[\hat{x} - x]^2. \quad (5.34)$$

It can be estimated by $1/n \sum_{i=1}^n (\hat{x}_i - x)^2$. Unfortunately, it is difficult to compare the MSE of different parameters, e.g. $\text{MSE}(\hat{\mu})$ and $\text{MSE}(\hat{\gamma})$. For that reason, the following additional error measures are introduced: The *absolute error (AE)* is introduced as

$$\text{AE}(\hat{x}_i) := \text{AE}_x(\hat{x}_i) := |\hat{x}_i - x| \quad (5.35)$$

and the *absolute relative error (RE)* as

$$\text{RE}(\hat{x}_i) := \text{RE}_x(\hat{x}_i) := \frac{\text{AE}(\hat{x}_i)}{x}. \quad (5.36)$$

For each setting and each parameter, the *mean absolute relative error (MRE)* is derived, i.e. for parameter μ_j of the j -th setting,

$$\overline{\text{RE}}(\mu_j) := \frac{1}{n} \sum_{i=1}^n \text{RE}_{\mu_j}(\hat{\mu}_{ij}). \quad (5.37)$$

The *mean absolute error (MAE)* is defined analogously and denoted by $\overline{\text{AE}}(\mu_j)$. Furthermore, because the MRE is dimensionless and the RE of different parameters is comparable, we can define the *overall relative error (ORE)* of setting j as

$$\text{ORE}_j := \overline{\text{RE}}(\mu_j) + \overline{\text{RE}}(\sigma_{j,1}) + \overline{\text{RE}}(\sigma_{j,2}) + \overline{\text{RE}}(\gamma_j). \quad (5.38)$$

These error measures are not used for the binary parameter m . Instead the relative frequency of having a wrong m is computed as

$$\tilde{m}_f := \sum_{i=1}^n \frac{|\hat{m}_{ij} - m_j|}{n}. \quad (5.39)$$

5.2.2.2. Robust statistics

When doing 10.000 or more simulations, it can happen that the fitting procedure gives extreme or even wrong results. For example, the algorithm does not converge or stops in a local minimum instead of the global one. For that reason we make use of some robust statistics (Hoaglin et al., 1983; Staudte and Sheather, 1990), instead of using the usual sample mean and sample standard deviation.

Definition 5.5: For data values $x_1, \dots, x_n \in \mathbb{R}$ with order statistics $x_{(1)} \leq \dots \leq x_{(n)}$, the α -trimmed mean ($\alpha \in [0, 1)$) is defined as

$$\bar{x}_t := \frac{1}{n - 2\lfloor \alpha n \rfloor} \sum_{i=1+\lfloor \alpha n \rfloor}^{n-\lfloor \alpha n \rfloor} x_{(i)}, \quad (5.40)$$

where $\lfloor y \rfloor$ denotes the whole-number part of y .

Definition 5.6: The *median absolute deviation (MAD)* is defined as

$$\text{MAD}(x_1, \dots, x_n) = \text{median}(\{|x_i - \tilde{x}| : i = 1, \dots, n\}), \quad (5.41)$$

where \tilde{x} is the median of x_1, \dots, x_n .

Remark 5.7: In order to use the MAD as a consistent estimator for the estimation of the standard deviation σ , one takes

$$\hat{\sigma} = K \cdot \text{MAD}(x_1, \dots, x_n). \quad (5.42)$$

For normally distributed data $K=1.4826$ (Staudte and Sheather, 1990) which we will use in the following sections.

5.2.3. Precision for different recording times and bin sizes

The impact of the bin size δ and the recording time \mathcal{T} is investigated regarding the precision of the fitting procedure. For this purpose, we take the four typical parameter constellations given in table 5.1 (A, B, C and D) and select $\delta = 1, 5, 10, 20, 40$ (milliseconds) and $\mathcal{T} = 300, 600, 900$ (seconds). Thus we have $4 \cdot 5 \cdot 3 = 60$ possible settings – four parameter combinations, five different δ and three \mathcal{T} (see table 5.2).

For each setting j of table 5.2, we simulate $n = 1000$ new spike trains with recording time \mathcal{T}_j and estimate the parameters again. For each setting, the corresponding δ_j is

| j | δ_j | \mathcal{T}_j | parameters | μ_j | ... | m_j |
|----------|------------|-----------------|------------|----------|----------|----------|
| 1 | 1 | 300 | A | 195.97 | ... | 0 |
| 2 | 1 | 600 | A | 195.97 | ... | 0 |
| 3 | 1 | 900 | A | 195.97 | ... | 0 |
| 4 | 5 | 300 | A | 195.97 | ... | 0 |
| \vdots | \vdots | \vdots | \vdots | \vdots | \vdots | \vdots |
| 60 | 40 | 900 | D | 190 | ... | 0 |

Table 5.2: Settings to investigate the impact of δ and \mathcal{T} on the fitting procedure.

used as the bin size for the function estimation. So, 60 sets of estimated parameters are generated, $\hat{\psi}_{ij}$, $i = 1, \dots, n$, $j = 1, \dots, 60$ and for each setting the ORE is derived, as well as the MRE for each parameter.

Figure 5.6 shows the result of the simulations. In only 2 of 60000 simulations the fitting procedure did not converge. For the rest it turns out, that the MRE and ORE decreases for increasing recording time \mathcal{T} as would be expected, because an increase of \mathcal{T} represents an increase of the number of spikes n . Furthermore, it turns out that δ should not be chosen too small, because in this case too much noise is represented in the ACH. Of course, if \mathcal{T} gets extremely large or the spike train has a large firing rate, small values of δ are preferred, because otherwise important dependencies will be averaged out (see Havenith et al. (2009) which recommend a bin size of 1 ms in the context of phase offsets and large firing rates). This can be seen in particular for parameter combination A which has $\mu \approx 196$. Finally, it turns out that a bin size of $\delta = 10$ gives good results for almost all combinations. For that reason, we will take $\delta = 10$ for further simulation analysis.

The relative frequency of wrong m fittings \tilde{m}_f has been zero, for all settings except the ones with parameter constellation D . These \tilde{m}_f values can be found in table 5.3. While there are only slight changes between different δ values, the change induced by different \mathcal{T} values is more apparent. However, these results do not surprise, because in the irregular non-bursty case it is difficult to decide if some clusters appear at random or not. But with an increasing recording time, the true m can be better detected.

5.2.4. Standard errors of nonlinear regression

The nonlinear regression procedures typically return the estimated variance-covariance matrix of the estimators and thus their estimated standard errors. Unfortunately, they cannot be used to construct confidence intervals, because the estimated standard errors

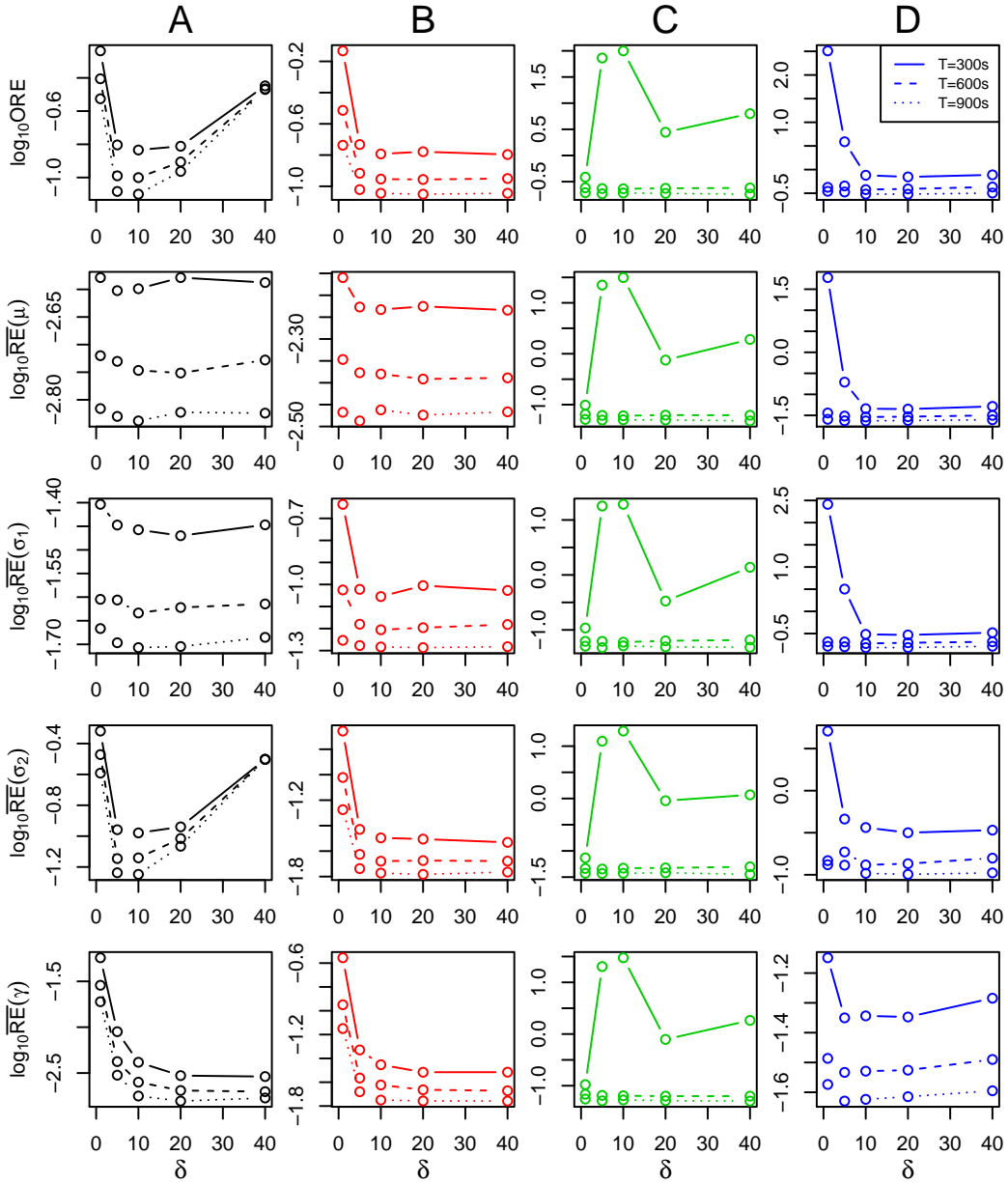


Figure 5.6: Comparison of estimated standard errors for different T and δ . $n = 1000$ simulations have been performed for each parameter combination given in table 5.1. Each column in the graphic represents one parameter constellation. Each plot shows the logarithm of the mean absolute relative error (RE) versus δ for different T values (lines). The first row represents the overall relative error (ORE). The other rows show the logarithm of the mean RE for the single parameters.

| \mathcal{T} [s] | δ | | | | |
|-------------------|----------|-------|-------|-------|-------|
| | 1 | 5 | 10 | 20 | 40 |
| 300 | 0.089 | 0.094 | 0.086 | 0.097 | 0.119 |
| 600 | 0.047 | 0.047 | 0.050 | 0.060 | 0.079 |
| 900 | 0.033 | 0.038 | 0.033 | 0.046 | 0.047 |

Table 5.3: Relative frequency \tilde{m}_f of wrong estimations of m for parameter constellation D of Figure 5.6. All other parameter constellations had $\tilde{m}_f = 0$.

are underestimating the true variability most of the time. This may be because of violations of the nonlinear regression assumptions (independence and normal distribution of errors).

To illustrate this problem, we simulate $n = 10.000$ spike trains with the parameter combinations given in table 5.1 for different recording times $\mathcal{T} = 5, 10, 15, 20, 25, 30$ (minutes). Thus we have $4 \cdot 6 = 24$ settings. For every simulated spike train the parameters are estimated with the function `estimation` and a bin size of $\delta = 10$ ms. As a result of the nonlinear regression, we get estimators $\hat{\psi}_{ij}$, as well as their estimated standard errors

$$\hat{s}e_R(\hat{\psi}_{ij}) := (\hat{s}e_R(\hat{\mu}_{ij}), \hat{s}e_R(\hat{\sigma}_{ij,1}), \hat{s}e_R(\hat{\sigma}_{ij,2}), \hat{s}e_R(\hat{\gamma}_{ij})), \quad (5.43)$$

where $\hat{s}e_R(\hat{\mu}_{ij})$ denotes the estimated standard deviation of $\hat{\mu}_{ij}$. For every recording time \mathcal{T} and every parameter (i.e. μ), we derive the 0.01-trimmed mean of the parameter estimates, i.e.

$$\bar{\hat{\mu}}_j := \frac{1}{n-200} \sum_{i=1+100}^{n-100} \hat{\mu}_{(i)j}, \quad (5.44)$$

as well as the 0.01-trimmed mean of all standard errors given by the nonlinear regression procedure

$$\hat{s}e_{\text{NLS}}(\hat{\mu}_j) := \frac{1}{n-200} \sum_{i=1+100}^{n-100} \hat{s}e_R(\hat{\mu}_{(i)j}), \quad (5.45)$$

(which we call the mean of the NLS SD estimators) and the median absolute deviation of all estimated parameters, i.e.

$$\hat{s}e(\hat{\mu}_j) = 1.4826 \cdot \text{MAD}(\hat{\mu}_{1j}, \dots, \hat{\mu}_{nj}). \quad (5.46)$$

Although $\hat{s}e(\hat{\mu}_j)$ should converge to the true standard deviation of $\hat{\mu}_j$ for increasing n , it cannot be used to estimate the standard deviation, because the true parameter value

(which is needed to simulate the spike trains) is not known in general. Furthermore, it is not clear what happens to $\hat{s}e_{\text{NLS}}$ for increasing n . For that reason, the simulations are used to investigate this question.

Figure 5.7 shows the results. In the mean (red line) the estimators $\hat{\psi}_{ij}$ vary around their true value (black line). For increasing \mathcal{T} the variability of the estimator decreases. But most of the time, the mean of the NLS SD estimators (green) is smaller than the real variability (blue) represented by the MAD. For that reason, it is not recommended to use the estimated standard errors of the NLS procedure. Alternative methods and ideas will be presented in Chapter 6.

5.2.5. Quantification of different parameter constellations

For the further quantification of the precision of the described fitting procedure some more predefined parameter combinations are needed. To focus on combinations which are relevant for experimental practice, we use the estimated parameter constellations of the 146 spike trains given in the sample data set (see appendix A.2). \mathcal{T} and δ are constant, so we have 146 settings. For each parameter combination $\psi_j = (\mu_j, \sigma_{j,1}, \sigma_{j,2}, \gamma_j, m_j)$, $j = 1, \dots, 146$, 10.000 new spike trains were simulated with $\mathcal{T} = 720$ seconds which represents a typical recording time. The parameter estimates $\hat{\psi}_{ij}$ were obtained with $\delta = 10$ ms.

Figure 5.8 shows the 80%-quantiles of the AE and RE of each parameter set. The AE is chosen for the non-bursty case, because the parameter values are relatively small (e.g. $\sigma_2 < 1$), such that even very small AEs appear as large REs. Conversely bursty spike trains have larger values in the parameters, therefore the RE is shown. Some parameter combinations, such as γ/μ , can be estimated rather precisely. For 104 out of the 146 parameter constellations (green lines), the errors were small in the following sense: If $m = 1$, the 80%-quantile of RE was less than 0.2 for all parameters, and if $m = 0$, the 80%-quantile of AE was less than 20 milliseconds for all parameters. In the remaining 42 constellations, errors were large as would be expected from large values of σ_2/μ or very low firing rates (Figure 5.8C).

Only 3 of 146 constellations had more than 1% aborts in the simulations, 99 had no aborts and for 44 constellations the abort rate has been between 0 and 1 percent. For 79 of 146 constellations there has been no wrong estimation of m , 56 had $0 < \tilde{m}_f < 0.1$ and for 11 constellations $\tilde{m}_f \geq 0.1$. The parameter constellations with high abort rate and high \tilde{m}_f are found in Figure 5.8 E and F.

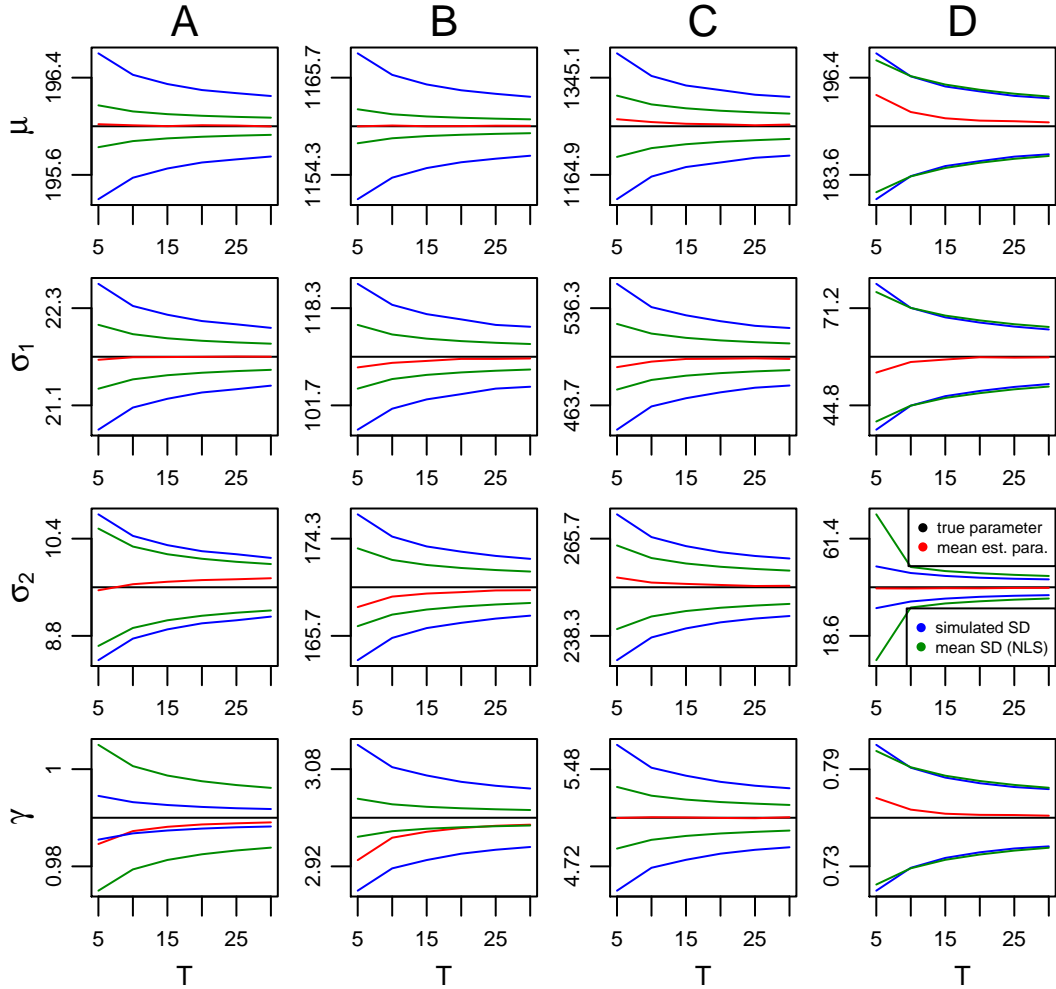


Figure 5.7: Comparison of estimated NLS standard errors. For every recording time \mathcal{T} , $n = 10.000$ new spike trains are generated with the parameter combinations given in table 5.1 and the parameters are estimated with a bin size of $\delta = 10$ ms. For every parameter (except m), the mean of the parameter estimates (red), the mean of all standard errors given by the nonlinear regression procedure (green) and the standard deviation of all estimated parameters (blue) is derived. The estimators vary around their true value (black line). For increasing \mathcal{T} the variability of the estimators decreases. But in general the mean of the NLS SD estimators is smaller than the real variability.

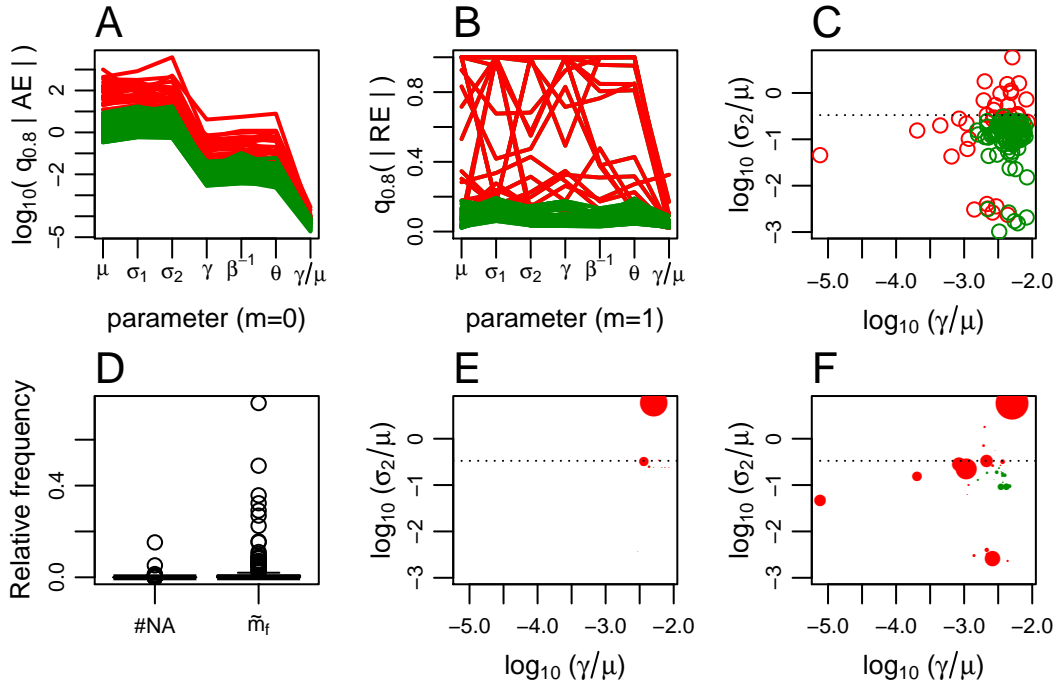


Figure 5.8: Precision of parameter estimates obtained from 10.000 simulations for each of 146 parameter constellations. (A) $\log(80\%)$ -quantile of the absolute error (AE) for each parameter and (B) 80%-quantile of the relative error (RE). (C) shows a scatterplot of $\log(\frac{\gamma}{\mu})$ and $\log(\frac{\sigma_2}{\mu})$ for the parameter estimates with a horizontal line at $\sigma_2/\mu = 1/3$. Green lines and points indicate constellations with either small AE or RE. (D) Boxplots of the relative frequency of simulations with an abort of the fitting procedure due to error (#NA) and the relative number \tilde{m}_f of wrong m estimates. (E) Parameter constellations with high #NA and (F) high \tilde{m}_f indicated by the size of the points.

5.3. Remark: Model diagnostics

After using the fitting procedure which results in the estimates $\hat{\mu}$, $\hat{\sigma}_1$, $\hat{\sigma}_2$, $\hat{\gamma}$ and \hat{n} , one should investigate the goodness of the fit. For further analysis (classification of firing patterns and comparison of parameter estimates), one has to decide to trust the model or not. We recommend to look on the empirical ACH together with the fitted ACF of the GLO (compare Figure 5.1 on page 83 and Figure 5.2 on page 85). If it looks like the algorithm has converged to a local minimum, one should not trust the estimated parameters. If the fit seems to be appropriate, one can use the estimated parameters for further analysis. Of course a formal or more systematic goodness-of-fit test for the GLO would be preferable, but there is none in preparation yet. In contrast to visual inspection of the fit, a formal test could lead to conservative decisions if the GLO should be applied. However, if the GLO is not the correct model, it still can be a suitable approximation to the data.

Another possibility for model diagnostic is (as it is common for regression analysis) to investigate the assumptions on the residuals: normal distributed, heteroscedasticity and independent of each other. Although these are important mainly for statistical inference, it shows if the model is appropriate or if there are structural deviations.

Example 5.8: Figure 5.9 shows four types of diagnostic plots for the four spike trains given in Figure 5.5 on page 90 which result of the GLO parameters given in Table 5.1. Each column represents one spike train (A-D) and each row one diagnostic plot. The fits in the first row seem to be adequate, as well as the quantile-quantile-plots for investigating the normal distribution of residuals. Furthermore, the other two graphics indicate that there are no abnormalities in the variance and correlation of residuals. Although, there are some correlations outside of the red confidence intervals in the fourth row for spike train A, this can have happened by chance.

Remark 5.9: (Possible problems and solutions) If the fit does not look appropriate, there is no convergence or the assumptions on the residuals are not fulfilled, there are the following possible reasons and solutions:

1. The model is not appropriate. This can be because of non-stationarities in the spike train or another structure underlying the firing pattern. In this case it depends on the underlying structure if maybe another model should be used.
2. The analysis window or the starting values are poorly chosen. Then, the set of possible starting values should be increased or values should be chosen by hand.

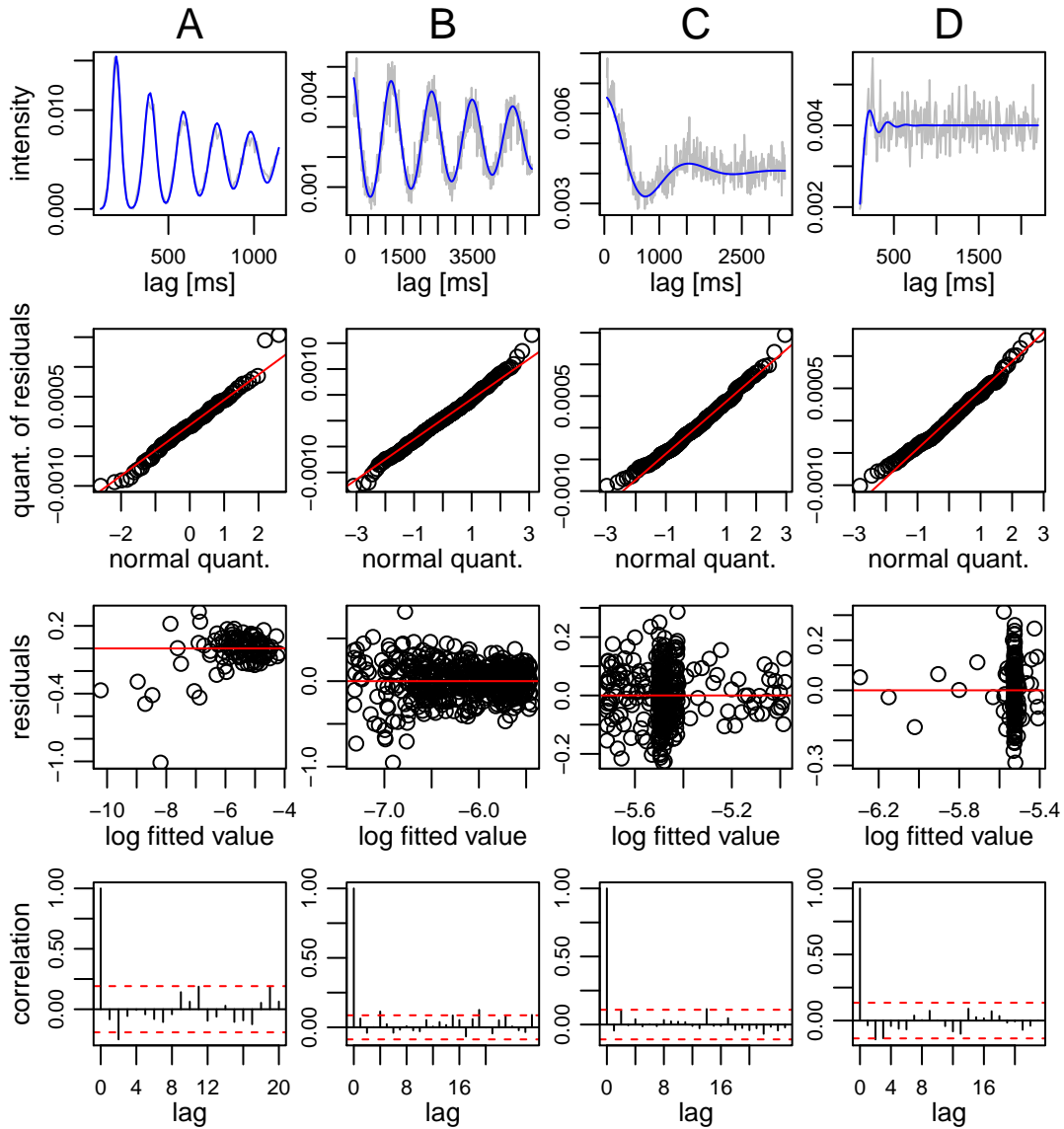


Figure 5.9: Four diagnostic plots are shown for the four exemplary spike trains given in Figure 5.5 which result of the GLO parameters given in table 5.1. In the first row are the raw ACHs (grey) and the fitted ACF curve (blue). The second row shows a quantile-quantile plot of the residuals and a normal distribution. In the third row the residuals are plotted against the logarithm of the fitted ACF values (red lines indicate zero) and in the last row the correlations between different lags of residuals are visualized (red lines indicate typical variations of white noise).

The same holds true for the analysis window.

3. There are not enough spikes in the spike train. The recording time \mathcal{T} should be increased (which is not always possible because the neuron can get instable). A possible solution can be to increase the bin size δ of the analysis window.

Chapter 6.

Variability estimation of estimated parameters

In Section 5.2, it has been shown that the applied estimation method yields relatively small errors for most parameter combinations, when a spike train fully complies with the GLO model assumptions. In practice, these strict assumptions are seldomly fulfilled, in particular regarding the stationarity of the process. But even if they are fulfilled, the preceding chapter (Section 5.2.4) showed that the estimated standard errors of the nonlinear regression procedure underestimated the true variability. Furthermore, in general it is not clear if the least squares algorithm converges correctly and resulted in the right parameter estimates.

For that reason, three bootstrap methods are described to estimate the variability of the estimated parameters (Section 6.1) and different types of confidence intervals are introduced to quantify this variability for the user (Section 6.2).

6.1. Methods for estimating the standard error

After giving a short introduction to the bootstrap (Subsection 6.1.1), three methods are described for estimating the standard errors of the estimated parameters: A parametric bootstrap procedure (Subsection 6.1.2), a nonparametric random block bootstrap (Subsection 6.1.3) and a marked point process bootstrap in combination with random blocks (Subsection 6.1.4). Finally, all three methods are investigated and compared with each other using simulations (Section 6.1.5).

6.1.1. Introduction to bootstrap methods

Bootstrap procedures are computer based methods to assess the accuracy of estimators. The basic idea is to construct new realizations of the observed data and to compute the quantity of interest several times. Thus an approximating distribution is generated from which bias, standard errors or confidence intervals can be derived. Originally these procedures were developed for simple univariate data, but the idea has been generalized to more complicated models, like linear regression, time series or point processes. The big advantage of bootstrap procedures is their simpleness and intuitive approach, but their consistence is not always ensured – in particular in more complicated models. For a detailed description of bootstrap methods the books of Efron and Tibshirani (1994) and Davison and Hinkley (2009) are recommended.

Basic idea: The bootstrap is explained by Efron (2003) in the following way: We have an unknown probability model P , e.g. a linear regression model, which depends on some unknown parameters and gives an observed data vector \mathbf{x} . From \mathbf{x} a statistic $\hat{\theta} = s(\mathbf{x})$ is calculated to estimate some parameter or quantity $\theta = t(P)$ of interest, e.g. the slope of the linear regression line. Typically we are interested in the accuracy of estimating θ by $\hat{\theta}$ in terms of bias, variance and confidence intervals. A point estimate \hat{P} of P is used to generate b bootstrap samples \mathbf{x}_i^* , $i = 1, \dots, b$. In the nonparametric case when P is totally unknown, \hat{P} is typically estimated by the empirical c.d.f.. In both cases, one gets bootstrap estimates $\hat{\theta}_i^* = s(\mathbf{x}_i^*)$, $i = 1, \dots, b$. The number of bootstrap samples b is only limited by time and computational power, but usually, a larger b means more accurate results. In the next step, the empirical distribution of the $\hat{\theta}_1^*, \dots, \hat{\theta}_b^*$ is used to assess the accuracy of $\hat{\theta}$. Figure 6.1 visualizes the idea of the bootstrap.

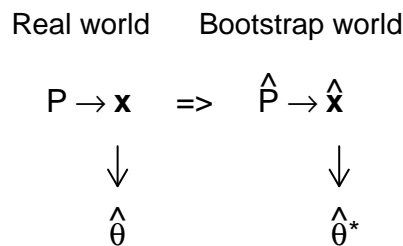


Figure 6.1: Bootstrap diagram (taken from Efron (2003) FIG. 1). The model P generates the data vector \mathbf{x} from which an estimate $\hat{\theta}$ of some quantity θ is computed. \mathbf{x} is also used to get an estimate \hat{P} of the model P , which can be used to construct a bootstrap sample $\hat{\mathbf{x}}$ from which the bootstrap estimate $\hat{\theta}^*$ can be derived.

Definition 6.1: The *bootstrap estimate of standard error* of $\hat{\theta}$ is defined analogously to the empirical standard deviation as

$$s\hat{e}_b(\hat{\theta}) = \left(\sum_{i=1}^b (\hat{\theta}_i^* - \bar{\theta}^*)^2 / (b-1) \right)^{1/2} \quad (6.1)$$

with $\bar{\theta}^* = 1/b \sum_{i=1}^b \hat{\theta}_i^*$ being the mean of the bootstrap estimates.

Example 6.2: In the simplest situation, we have realizations $\mathbf{x} = (x_1, \dots, x_n)$ of a random variable X with unknown c.d.f. F . In this univariate nonparametric case, the probability model P is represented by F and an estimator of interest may be the expectation $\theta = t(F)$ which is typically estimated by $s(\mathbf{x}) = 1/n \sum_{i=1}^n x_i$. Because F is unknown, \mathbf{x}_i^* represents a random sample of size n of \mathbf{x} (with replacement) and the bootstrap estimates are given as $s(\mathbf{x}_i^*)$, $i = 1, \dots, b$. If we do not know the standard error of $\hat{\theta}$, we can estimate it by the bootstrap estimate of standard error (6.1).

Example 6.3: In our case, the probability model P is given by $\Phi \sim \text{GLO}(\psi)$, from which we have a realization $\mathcal{S} = (t_1, \dots, t_n)$ which corresponds to \mathbf{x} . The quantity of interest are the parameters $\psi = (\mu, \sigma_1, \sigma_2, \gamma, m)$ and combinations of them. The statistic $s(\cdot)$ is given by the fitting procedure. After generating new bootstrap samples (spike trains) $\mathcal{S}_1^*, \dots, \mathcal{S}_b^*$, they are used to get the bootstrap estimates $\psi_i^* = s(\mathcal{S}_i^*)$, $i = 1, \dots, b$.

Of course, one can use the bootstrap estimate of standard error to quantify the precision of the estimates in example 2. But it can happen that for some of the b bootstrap spike trains, the fitting procedure gives wrong fits which may result in extreme values (outliers) which affect (6.1). For this reason, it can be necessary to take a robust estimate of variability. So, we define the bootstrap MAD estimate (compare equation (5.41)) as

$$\widehat{\text{MAD}}(\hat{\theta}) := 1.4826 \cdot \text{MAD}(\hat{\theta}_1^*, \dots, \hat{\theta}_b^*). \quad (6.2)$$

Finally, the procedure of generating new spike train replications has not been discussed yet, although it is crucial and very important for the bootstrap. This will be regarded in detail in the following subsections where three different methods are introduced: The first (Section 6.1.2) and second (Section 6.1.3.1) method generate bootstrap replications of the empirical spike train and the third method (Section 6.1.4) constructs a bootstrap replication of the ACH.

6.1.2. Parametric Simulations

The idea is rather simple: Let us suppose that the GLO is an appropriate model for the underlying spike train with recording time \mathcal{T} and the parameter vector ψ is estimated using the described fitting procedure (R function `estimation`). Given these estimated parameters $\hat{\psi}$, n new spike trains are simulated with recording time \mathcal{T} (R function `GLO`) and the parameters are estimated again, resulting in estimates $\hat{\psi}_i$, $i = 1, \dots, n$. This is the same procedure which has been used for the quantification of the fitting procedure in the previous chapter. Finally, the standard error of $\hat{\psi}$ can be estimated using equation (6.1) or (6.2). The R function `boot1` (see appendix A.3) uses $\hat{\psi}$ to simulate n new spike trains and gives the resulting bootstrap estimates $\hat{\psi}_i$, $i = 1, \dots, n$, as output.

In practice, one should be cautious to directly apply such parametric simulations for the derivation of the precision of the obtained estimates. The reason is that the simulation generates data with the GLO mechanism and does not consider slight deviations from the assumptions, e.g. non-stationarities in the data, which would increase the variability of the estimators (see Figure 6.2). Furthermore, the obtained parameter estimates may be wrong, e.g. the algorithm converged only in a local minimum, so that one should not use them for further simulations.

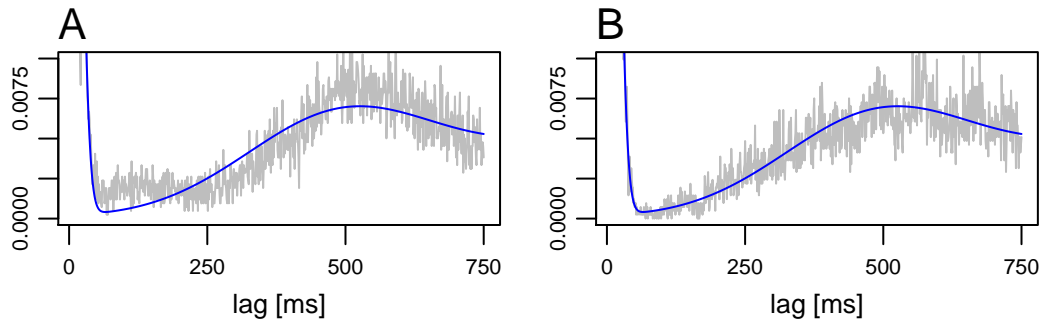


Figure 6.2: Comparison of a fit between empirical and simulated spike train. (A) ACH of an empirical spike train of the sample data set. The fitted GLO ACF with parameters $\hat{\psi}$ is shown as a blue line. (B) ACH of a simulated GLO spike train with parameters $\hat{\psi}$. Obviously, the fit for the simulated spike train looks better.

6.1.3. Nonparametric random blocks bootstrap

Because of the problems of the parametric simulation, we can use a nonparametric approach which does not contain the GLO assumptions. The idea is that all relevant

information about the true process is already given in the empirical spike train. To generate a bootstrap replication, r random pieces or blocks are sampled at random with replacement from the spike train and pasted together again. If the spike train is a stationary GLO process, this method should give similar results as the parametric simulations (at least if r is small). But if the spike train is non-stationary or has another structure, the bootstrap replications of this method should have variability closer to the true variability than the parametric simulation whose variability would be too small.

Method: A new realization of the spike train is constructed in the following way: From the given spike train $\mathcal{S} = (t_1, \dots, t_n)$ which has been recorded over the time interval $[0, \mathcal{T}]$, we draw r random blocks of length \mathcal{T}/r , allowing for overlap of different blocks. So we have

$$X_1, \dots, X_r \sim \text{Unif}(0, \mathcal{T} - \mathcal{T}/r) \quad (6.3)$$

and set the blocks as the intervals

$$A_i := [X_i, X_i + \mathcal{T}/r), \quad i = 1, \dots, r. \quad (6.4)$$

The spikes in these blocks are given by the sets

$$A_i^* := \mathcal{S} \cap A_i, \quad i = 1, \dots, r, \quad (6.5)$$

which are pasted together to the new realization

$$\mathcal{S}^* := \bigcup_{i=1}^r (A_i^* - X_i + (j-1)\mathcal{T}/r). \quad (6.6)$$

Thus the spike times $\tilde{t}_1, \dots, \tilde{t}_{n_0}$ ($n_0 = |\mathcal{S}^*|$) are the ordered elements of \mathcal{S}^* which can be used to estimate the parameters in the usual way. The R function `boot2` (see appendix A.3) gives the resulting bootstrap estimates $\hat{\psi}_i$, $i = 1, \dots, n$, as output.

6.1.3.1. Number of blocks

How to choose the number of blocks r ? On the one hand, r has to be large enough to construct adequately many new bootstrap spike trains. On the other hand, \mathcal{T}/r has to be large enough to capture all important dependencies. To investigate which r should be taken to estimate the unknown standard deviation of the parameter estimates, we again make use of simulations.

We take the four exemplary parameter combinations given in table 5.1, three different recording times $\mathcal{T} = 300, 600, 900$ (seconds) and seven different numbers of blocks with

$r = 15, 30, 45, 60, 75, 90, 105$. Thus there are 84 different settings. In practice, we only have one spike train from which the bootstrap samples are generated, but because there is a possibility for extreme observations, it is better to use several spike train realizations for the simulation study. For that reason, for every setting i , we simulate $n = 20$ new spike trains \mathcal{S}_{ik} , $i = 1, \dots, 84$ and $k = 1, \dots, n$ (thus simulating 1680 spike trains). For each of these spike trains $b = 500$ bootstrap samples are constructed according to the nonparametric random blocks bootstrap method and the GLO parameters are estimated with the function `estimation` and a bin size of $\delta = 10$ milliseconds. The function `boot2` performs the task and gives the estimated parameters $\hat{\psi}_{ijk}$, $j = 1, \dots, b$ as output. Then, for each of the 1680 spike trains the bootstrap estimate of standard error is computed over all b bootstrap replications of each parameter estimate, e.g. for the parameter μ

$$\widehat{\text{se}}_b(\hat{\mu}_{ik}) = \left(\sum_{j=1}^b (\hat{\mu}_{ijk}^* - \bar{\mu}_{ik}^*)^2 / (b-1) \right)^{1/2} \quad (6.7)$$

with $\bar{\mu}_{ik}^* = 1/b \sum_{j=1}^b \hat{\mu}_{ijk}^*$. Figure 6.3 visualizes the idea of this procedure.

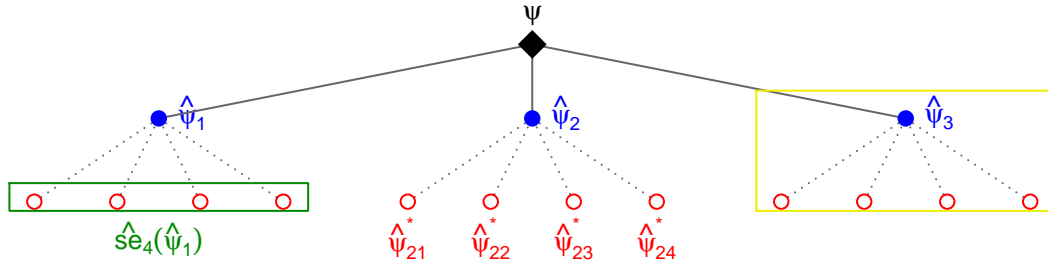


Figure 6.3: Layout for the bootstrap simulation: According to the true parameter combination ψ (black), $n = 3$ new spike trains are simulated and the parameters estimated, resulting in $\hat{\psi}_i$ (blue), $i = 1, 2, 3$. From each of these spike trains $b = 4$ bootstrap replications are generated and their parameters are estimated by $\hat{\psi}_{ij}^*$ (red). These are used to estimate the standard error of $\hat{\psi}_i$ with $\widehat{\text{se}}_4(\hat{\psi}_i)$ (green). In practice, when the true parameters are unknown only the yellow part is observed.

Finally for each setting i , the 0.1-trimmed mean of the n bootstrap estimates of standard error

$$\bar{\text{se}}_b(\hat{\mu}_i) = \frac{1}{n - 2[0.1n]} \sum_{k=1+[0.1n]}^{n-[0.1n]} \widehat{\text{se}}_b(\hat{\mu}_{ik}) \quad (6.8)$$

can be compared with the true variability. The true variability is approximated by the standard deviation of the parameter estimates of the simulated spike trains for a given

setting i and is denoted by $\bar{s}e(\hat{\mu}_i)$. The results of the simulation in Section 5.2.4 can be used for this purpose. Then the error is quantified with $RE_{\bar{s}e(\hat{\mu}_i)}(\bar{s}e_b(\hat{\mu}_i))$ (compare equation (5.36)).

Figure 6.4 shows the results of the simulation regarding the number of blocks for the bootstrap method. The difference of estimated standard error and true variability given by $\bar{s}e(\hat{\mu}_i)$ is quantified with the relative absolute error. In the regular case (A & B) the error increases if the number of drawn blocks r increases. This has been expected, because if the spike train is really a regular GLO process, then additional variability is introduced by increasing r . Furthermore, the irregular case (C & D) seems to be independent of r . For that reason, $r = 15$ is suggested for further analysis.

6.1.4. Marked point process bootstrap with random blocks

We propose a marked point process bootstrap procedure (Braun and Kulperger, 1998) in order to retain the dependence structure of the underlying process. The method has been used to estimate second order properties of spatial point processes (Loh, 2010) and is inspired by the block resampling algorithm used in time series bootstrapping suggested by Künsch (1989).

The method is as follows: Given a spike train of length \mathcal{T} , with spikes t_1, \dots, t_n and analysis window $\mathcal{L} = \{l_1, \dots, l_K\}$. The ACH $\tilde{h}(l)$ at lag l can be decomposed in

$$\tilde{h}(l) = \sum_{i=1}^n \tilde{h}_i(l), \quad (6.9)$$

where

$$\tilde{h}_i(l) := \#\{j \in [n] : l - \delta/2 < t_j - t_i \leq l + \delta/2\}. \quad (6.10)$$

This means that we fix spike time t_i and count the number of spikes t_j which have a time difference $t_j - t_i \in (l - \delta/2, l + \delta/2]$. Thus the unnormalized ACH $\tilde{h}(l)$ is split into n parts $\tilde{h}_i(l)$ and the sum of all parts is the original ACH. Here, we call $\tilde{h}_i(l)$ the 'contribution' of spike i to the counts at lag l . The same holds true, if the ACH is regarded for all lags as a vector

$$\tilde{\mathbf{h}} = (\tilde{h}(l_1), \dots, \tilde{h}(l_K))' \quad (6.11)$$

and

$$\tilde{\mathbf{h}}_i = (\tilde{h}_i(l_1), \dots, \tilde{h}_i(l_K))'. \quad (6.12)$$

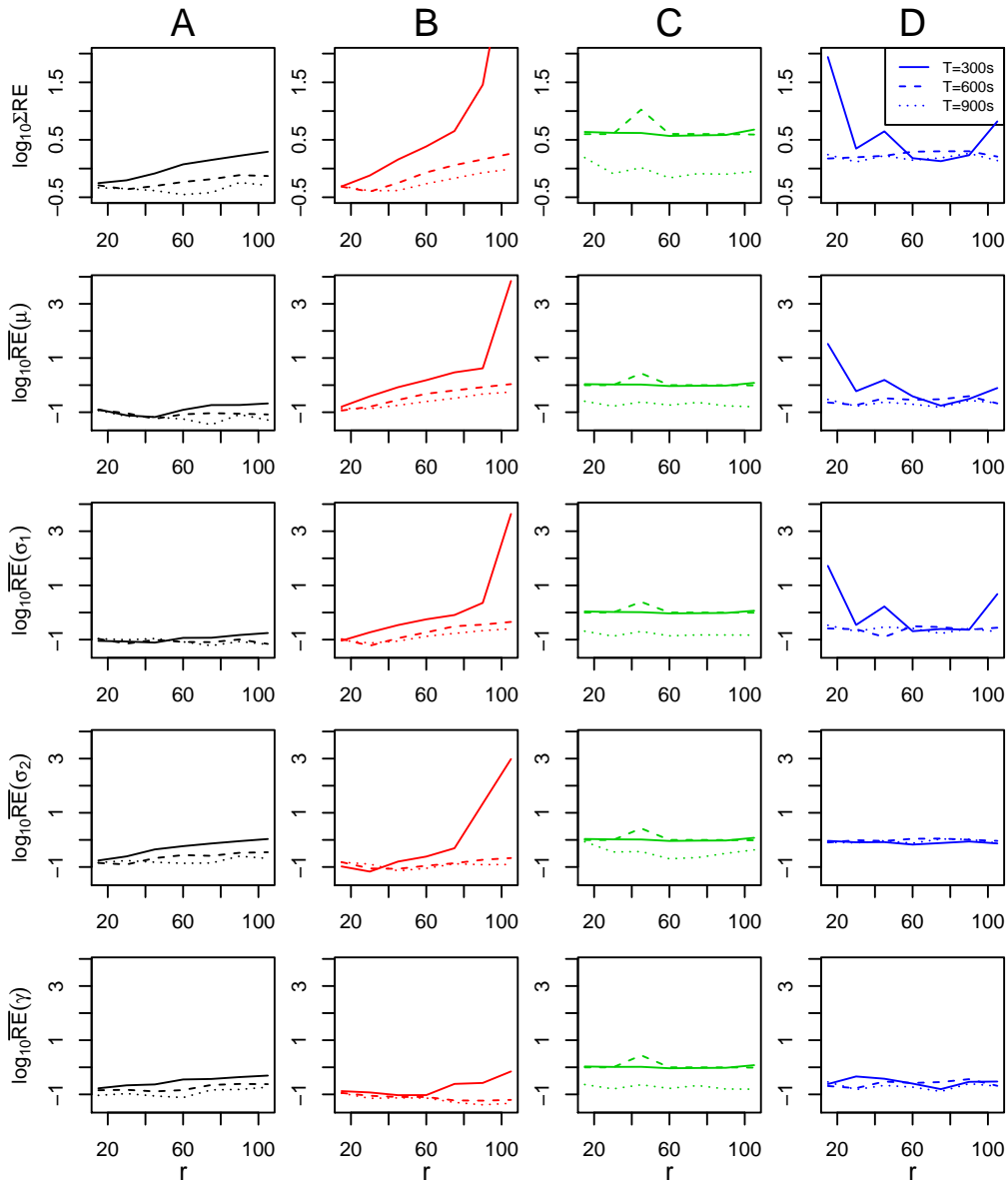


Figure 6.4: Number of blocks for block resampling. For the four parameter combinations given in table 5.1, three different recording times $\mathcal{T} = 300, 600, 900$ (seconds) and seven different numbers of blocks $r = 15, 30, 45, 60, 75, 90, 105$ (84 settings), $n = 20$ spike trains have been simulated. For every spike train $b = 500$ bootstrap replications have been generated to compute the bootstrap estimate of standard error. This is averaged over all n values and compared with the approximate true variability of the setting with the relative error.

So, we also call $\tilde{\mathbf{h}}_i$ the 'contribution' of spike i to the ACH. Thus again, we can write $\tilde{\mathbf{h}} := \sum_{i=1}^n \tilde{\mathbf{h}}_i$. Furthermore we regard $\tilde{\mathbf{h}}_i$ as the mark of spike t_i and write $\tilde{t}_i = (t_i, \tilde{\mathbf{h}}_i)$ as an element of the marked point process $(\tilde{t}_i)_{i=1, \dots, n}$.

To this marked point process we apply a block resampling method which is almost the same as for the random blocks bootstrap: From the given spike train we draw r random blocks of length \mathcal{T}/r , allowing for overlap of different blocks. So we have

$$X_1, \dots, X_r \sim \text{Unif}(0, \mathcal{T} - \mathcal{T}/r) \quad (6.13)$$

and set the blocks as

$$A_i := [X_i, X_i + \mathcal{T}/r), \quad i = 1, \dots, r. \quad (6.14)$$

The marks of the spikes in these blocks are then used to built up a bootstrap version of the ACH

$$\tilde{\mathbf{h}}^* := \sum_{i=1}^n \sum_{j=1}^r 1_{\{t_i \in A_j\}} \tilde{\mathbf{h}}_i. \quad (6.15)$$

The normalized ACH is then computed analog to (4.28) by using

$$n^* := \sum_{i=1}^n \sum_{j=1}^r 1_{\{t_i \in A_j\}} \quad (6.16)$$

instead of n . The estimation procedure described in Chapter 5 is used with the starting values of the original spike train and results in the estimated parameters $(\hat{\mu}, \hat{\sigma}_1, \hat{\sigma}_2, \hat{\gamma}, \hat{m})$ which we call bootstrap estimates. This procedure is repeated b times, yielding the estimates $(\hat{\mu}_i, \hat{\sigma}_{1,i}, \hat{\sigma}_{2,i}, \hat{\gamma}_i, \hat{m}_i)$, $i = 1, \dots, b$. The procedure is implemented in the R function `boot3` in appendix A.3.

6.1.5. Comparison of Methods

All three bootstrap methods have advantages and disadvantages. If the spike train fulfills the GLO assumptions, the parametric method is preferable, otherwise one of the other methods can be taken. The marked point process bootstrap tries to avoid the cut points of the block resampling method (the region where two blocks are pasted together) by directly going to the ACH. Of course, there are also other possibilities for resampling. For example, one can permute the ISI intervals, but this would destroy the underlying dependency structure. For that reason, we focus on the three described methods and investigate their behavior regarding the relative absolute error.

For this purpose, we take again the four exemplary parameter combinations given in table 5.1, three different recording times $\mathcal{T} = 300, 600, 900$ (seconds) and choose $r = 15$ as the number of drawn blocks for both nonparametric methods. Furthermore we have three different resampling methods (`boot1`, `boot2`, `boot3`) and thus there are 36 different settings.

The same simulation approach is used as in Section 6.1.3.1: For every setting i , we simulate $n = 20$ new spike trains \mathcal{S}_{ik} , with $i = 1, \dots, 36$ and $k = 1 \dots, n$. For each of these spike trains $b = 500$ bootstrap samples are constructed and parameters estimated accordingly. The estimated parameters $\hat{\psi}_{ijk}$, $j = 1, \dots, b$ are given as output. Then, for each of the 720 spike trains \mathcal{S}_{ik} the bootstrap estimate of standard error is computed over all b bootstrap replications of each parameter estimate (see equation (6.7) for parameter μ). Figure 6.3 again visualizes the idea of this procedure.

Finally for each setting i , the mean of the n bootstrap estimates of standard error (see equation 6.8) is compared with the true variability $\bar{s}e(\hat{\mu}_i)$ which is estimated with the mean absolute deviation. The results of the simulation in Section 5.2.4 can be used again and the error is quantified with $RE_{\bar{s}e(\hat{\mu}_i)}(\bar{s}e_b(\hat{\mu}_i))$.

The results can be found in Figure 6.5. It turns out that the parametric method has the most accurate estimates, because the simulated spike trains fully complies with the GLO model assumptions. Furthermore, one can see that the error is higher for irregular spike trains, especially for the irregular non-bursty case (combination D) and that the error is decreasing for increasing \mathcal{T} . Between the two nonparametric methods, there seems to be no large difference. Because the computational complexity of the random blocks bootstrap given by the R function `boot2` is less than the complexity of the marked point process bootstrap represented by `boot3` and because arbitrary empirical spike train does not need to fulfill the GLO assumptions, it is recommended to use `boot2` for further analysis.

6.2. Bootstrap confidence intervals

In the previous section, different methods have been introduced and investigated to generate bootstrap estimates of the GLO parameters. Typically, the bootstrap estimate of standard deviation is used to quantify the precision of the parameter estimates, but confidence intervals are usually preferred in experimental practice. Thus in the current section, different types of confidence intervals will be introduced (Section 6.2.1) and compared with each other (Section 6.2.2). Finally, a method is shown how confidence inter-

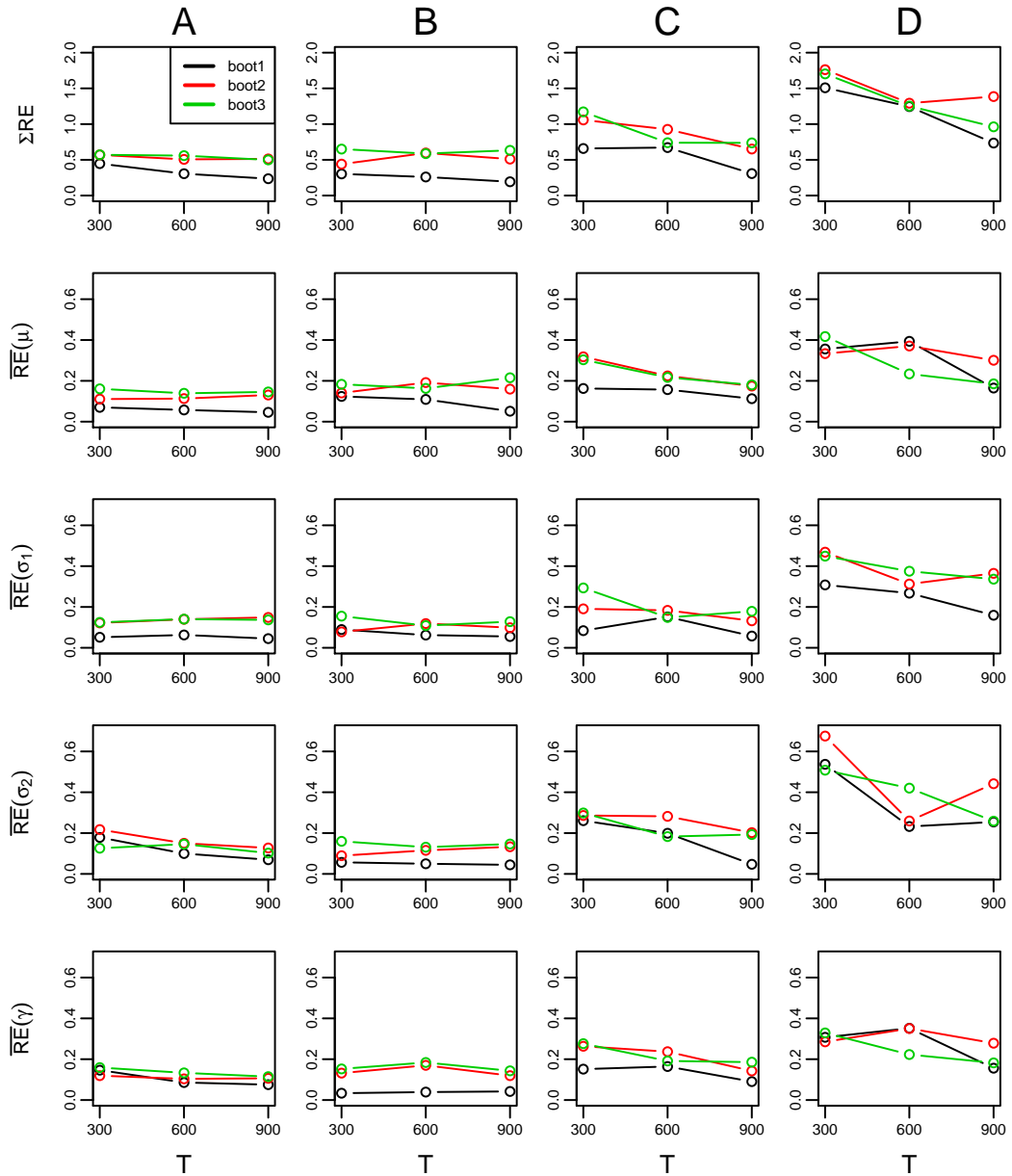


Figure 6.5: Comparison of bootstrap methods. For the four parameter combinations given in table 5.1, three different recording times $T = 300, 600, 900$ (seconds) and the three resampling methods (*boot1*, *boot2*, *boot3*) (36 settings), $n = 20$ spike trains have been simulated. For every spike train $b = 500$ bootstrap replications have been generated to compute the bootstrap estimate of standard error. This is averaged over all n values and compared with the approximate true variability of the setting with the relative error.

vals can be used for hypothesis testing regarding the GLO parameters (Section 6.2.3).

6.2.1. Different types of confidence intervals

Definition 6.4: Given a random sample $\mathbf{X} = (X_1, \dots, X_n)'$ of some probability distribution with parameter θ , a *confidence interval (CI)* for θ with confidence level $\alpha \in (0, 1)$ is a random interval $C_\alpha(\mathbf{X}) := [u_\alpha(\mathbf{X}), v_\alpha(\mathbf{X})]$ on \mathbb{R} with

$$\mathbb{P}(\theta \in C_\alpha(\mathbf{X})) \geq 1 - \alpha. \quad (6.17)$$

$u_\alpha(\mathbf{X})$ and $v_\alpha(\mathbf{X})$ are the real valued random end points of the interval.

The probability $\mathbb{P}(\theta \in C_\alpha(\mathbf{X}))$ is called the *coverage probability* of the confidence interval. In some definitions the inequality in (6.17) is replaced by equality. $C_\alpha(\mathbf{X})$ is random, because it only depends on the random sample \mathbf{X} . In practice, a confidence interval is constructed of a realizations \mathbf{x} of \mathbf{X} and we write $C_\alpha(\mathbf{x})$. Furthermore, let $\hat{\theta}$ be an estimator of θ . Then, $\hat{\theta}$ is called a point estimate and $C_\alpha(\mathbf{x})$ is called an interval estimate which is more informative and represents the precision of $\hat{\theta}$.

Sometimes, the CI has the desired coverage probability only asymptotically, this means that

$$\lim_{n \rightarrow \infty} \mathbb{P}(\theta \in C_\alpha((x_1, \dots, x_n)')) = 1 - \alpha. \quad (6.18)$$

But in complex models (e.g. in the GLO), it can be even difficult to derive asymptotic confidence intervals. In this case, bootstrap methods can be used to construct asymptotic CIs. Of course this depends on the data generating process, the resampling method and the construction of the confidence interval and it is not assured to get the desired coverage probability (at least in complex models).

Let $\mathbf{y} = (\hat{\theta}_1^*, \dots, \hat{\theta}_b^*)'$ be the bootstrap estimates of θ , derived with one of the resampling methods. Then there are these different famous types of bootstrap confidence intervals DiCiccio and Efron (1996):

The standard interval: Often it is appropriate to assume that $\hat{\theta}$ is approximately normal distributed. In this case and if $\hat{s}e_b(\hat{\theta})$ denotes the bootstrap estimate of standard error, the confidence interval is given by

$$C_\alpha^{(1)}(\mathbf{y}) := [\hat{\theta} - z^{(1-\alpha/2)} \hat{s}e_b(\hat{\theta}), \hat{\theta} - z^{(\alpha/2)} \hat{s}e_b(\hat{\theta})], \quad (6.19)$$

where $z^{(\alpha)}$ denotes the α -quantile of the standard normal distribution.

The bootstrap-t interval improves the standard interval which is only an approximation to finite samples. If there are b bootstrap samples $\mathbf{x}_1^*, \dots, \mathbf{x}_b^*$, we compute

$$Z_i^* = \frac{\hat{\theta}_i^* - \hat{\theta}}{\hat{s}\hat{e}_b(\hat{\theta}_i^*)}, \quad i = 1, \dots, b, \quad (6.20)$$

where $\hat{\theta}_i^* = s(\mathbf{x}_i^*)$ and $\hat{s}\hat{e}_b(\hat{\theta}_i^*)$ is the estimated standard error of $\hat{\theta}_i^*$ for the i -th bootstrap sample. Thus one has to use the bootstrap twice. First of all the empirical spike train is bootstrapped and in a second step, each bootstrap sample is bootstrapped to estimate $\hat{s}\hat{e}_b(\hat{\theta}_i^*)$.

Let $\hat{t}^{(\alpha)}$ be defined as the α -quantile of Z_1^*, \dots, Z_b^* , then the bootstrap-t interval is

$$C_\alpha^{(2)}(\mathbf{y}) := [\hat{\theta} - \hat{t}^{(1-\alpha/2)}\hat{s}\hat{e}_b(\hat{\theta}), \hat{\theta} - \hat{t}^{(\alpha/2)}\hat{s}\hat{e}_b(\hat{\theta})]. \quad (6.21)$$

Basic bootstrap interval: Let $\hat{\theta}^{*(\alpha)}$ denote the α -quantile of $\hat{\theta}_1^*, \dots, \hat{\theta}_b^*$. The basic bootstrap interval is

$$C_\alpha^{(3)}(\mathbf{y}) := [2\hat{\theta} - \hat{\theta}^{*(1-\alpha/2)}, 2\hat{\theta} - \hat{\theta}^{*(\alpha/2)}]. \quad (6.22)$$

The percentile interval also uses the quantiles of $\hat{\theta}_1^*, \dots, \hat{\theta}_b^*$. It is given as

$$C_\alpha^{(4)}(\mathbf{y}) := [\hat{\theta}^{*(\alpha/2)}, \hat{\theta}^{*(1-\alpha/2)}]. \quad (6.23)$$

6.2.2. Coverage probabilities

In this section, we investigate the coverage probabilities for the different types of confidence intervals described in Section 6.2.1. As mentioned, they depend on the data generating process, the resampling method and the confidence interval. Of course we would like to use a CI $C_\alpha(\mathbf{x})$ which fulfills (6.17) and is as small as possible. But in practice, the bootstrap intervals have this coverage probability only asymptotically, i.e.

$$\lim_{b \rightarrow \infty} \mathbb{P}(\theta \in C_\alpha((\hat{\theta}_1^*, \dots, \hat{\theta}_b^*)')) = 1 - \alpha. \quad (6.24)$$

In general, computational simulation time and the number of bootstrap samples b is limited. But if b is large enough, we can assume

$$\mathbb{P}(\theta \in C_\alpha((\hat{\theta}_1^*, \dots, \hat{\theta}_b^*)')) \approx 1 - \alpha. \quad (6.25)$$

In the GLO context, it is not clear how the different CI types will behave. Furthermore, we introduce two modifications of the CIs:

1. In $C_\alpha^{(1)}(\mathbf{y})$ and $C_\alpha^{(2)}(\mathbf{y})$, we replace $\hat{s}_b(\hat{\theta})$ by the MAD estimator

$$K \cdot MAD(\hat{\theta}_1^*, \dots, \hat{\theta}_b^*) \quad (6.26)$$

and K as defined in (5.42). This ensures that possible outliers produced by the fitting procedure will be of no consequence.

2. For $C_\alpha^{(2)}(\mathbf{y})$, $\hat{s}_b(\hat{\theta}_i^*)$ in (6.20) is replaced by $\hat{s}_b(\hat{\theta}^*)$ which saves a lot of computational time.

The modified CIs will be denoted as $C_\alpha^{(1')}(\mathbf{y})$ and $C_\alpha^{(2')}(\mathbf{y})$.

To investigate the coverage probabilities of $C_\alpha^{(1')}(\mathbf{y})$, $C_\alpha^{(2')}(\mathbf{y})$, $C_\alpha^{(3)}(\mathbf{y})$ and $C_\alpha^{(4)}(\mathbf{y})$, we take again the four exemplary parameter combinations given in table 5.1 and set $\mathcal{T} = 900$ (seconds) which will represent a typical empirical recording time. For every ψ of the four parameter combinations, $n_1 = 400$ new spike trains \mathcal{S}_i are simulated and the parameters estimated with the R function `estimation` and $\delta = 10$ milliseconds, resulting in $\hat{\psi}_i$ ($i = 1, \dots, n_1$). Each of these spike trains is resampled $b = 1000$ times with the random block bootstrap described in Section 6.1.3 and with $r = 15$ as the number of drawn blocks. This results in the bootstrap samples \mathcal{S}_{ij}^* , $i = 1, \dots, n_1$ and $j = 1, \dots, b$. The parameters are estimated in the same way as before, resulting in the bootstrap estimates $\hat{\psi}_{ij}^*$ (using R function `boot2`). The confidence intervals are constructed as described for every parameter of interest and for every interval type with $\alpha = 0.8$. For example, for parameter μ , the i -th spike train \mathcal{S}_i and the percentile interval, the confidence interval is given by $C_{i,\alpha}^{(4)}((\hat{\mu}_{i1}^*, \dots, \hat{\mu}_{ib}^*)')$. Then its coverage probability for the true parameter μ is estimated by

$$\frac{1}{n_1} \sum_{i=1}^{n_1} 1_{\{\mu \in C_{i,\alpha}^{(4)}((\hat{\mu}_{i1}^*, \dots, \hat{\mu}_{ib}^*)')\}} \quad (6.27)$$

Of course, it would be preferred to simulate new spike trains for each CI type, but this would increase the already large number of simulations ($4 \cdot n_1 \cdot b = 1.600.000$) by a factor of four.

Figure 6.6 shows these coverage probabilities. Each column represents one of the four exemplary parameter combinations given in table 5.1 and each row one parameter or a combination of parameters. In every plot, there are the coverage probabilities of the four confidence interval types, discussed in the previous section. The estimated coverage probability is represented by a point estimate, surrounded by an 95%-confidence interval which represents the unsureness by doing 'only' $n_1 = 400$ confidence intervals. The attempted coverage of 80% is visualized by a dotted line. Most of the time, all four

CI types are close to the coverage of 80%. Only in some cases the coverage probability is even smaller than 0.7. Of course combinations of parameters are influenced by the raw parameters, i.e. if the coverage probability for σ_1 is bad, the same holds true for σ_1/μ . Surprisingly, the modified standard interval has an adequate coverage probability in almost all plots. The two CIs $C_\alpha^{(3)}(\mathbf{y})$ and $C_\alpha^{(4)}(\mathbf{y})$ have a disadvantage, because $b = 1000$ is not large enough (typically in the literature 10.000 bootstrap samples are recommend, in particular when using $\alpha = 0.05$).

The decision which interval type to take, depends on the time needed for bootstrapping and on the underlying assumptions. If one has doubts that the distribution of the bootstrap estimates is bell shaped, the quantile methods should be preferred, although they need b to be large enough. If the distribution is assumed bell shaped and high accuracy is needed, the bootstrap-t interval should be taken. Nevertheless, high accuracy is seldom required in experimental practice and all four intervals fulfill the task to visualize the variability of the estimators.

6.2.3. Hypothesis Testing

Although confidence intervals are preferable to quantify the variability of parameter estimates, there can be a need or wish for getting p -values of a hypothesis test. This may be important for the comparison of differences between two parameter estimates, e.g. γ_1 and γ_2 , derived from two different empirical spike trains \mathcal{S}_1 and \mathcal{S}_2 (for example for pre/post comparisons in the context of pharmacological treatment). Of course any other parameter or parameter combination can be taken instead. For the testing problem

$$\begin{aligned} H_0: \gamma_1 &= \gamma_2 && \text{(null hypothesis)} \\ H_1: \gamma_1 &\neq \gamma_2 && \text{(alternative hypothesis),} \end{aligned}$$

one could regard the test statistic

$$\frac{\hat{\gamma}_1 - \hat{\gamma}_2}{\sqrt{\hat{s}_b(\hat{\gamma}_1)^2 + \hat{s}_b(\hat{\gamma}_2)^2}}$$

which is approximately standard normal distributed – assuming that

$$\begin{aligned} \hat{\gamma}_1 &\sim \mathcal{N}(\gamma_1, \hat{s}_b(\hat{\gamma}_1)) \\ \hat{\gamma}_2 &\sim \mathcal{N}(\gamma_2, \hat{s}_b(\hat{\gamma}_2)) \end{aligned}$$

and both spike trains to be independent. Of course this is seldom fulfilled in practice and furthermore it is not clear if the variability estimates $\hat{s}_b(\hat{\gamma}_1)$ and $\hat{s}_b(\hat{\gamma}_2)$ are precise.

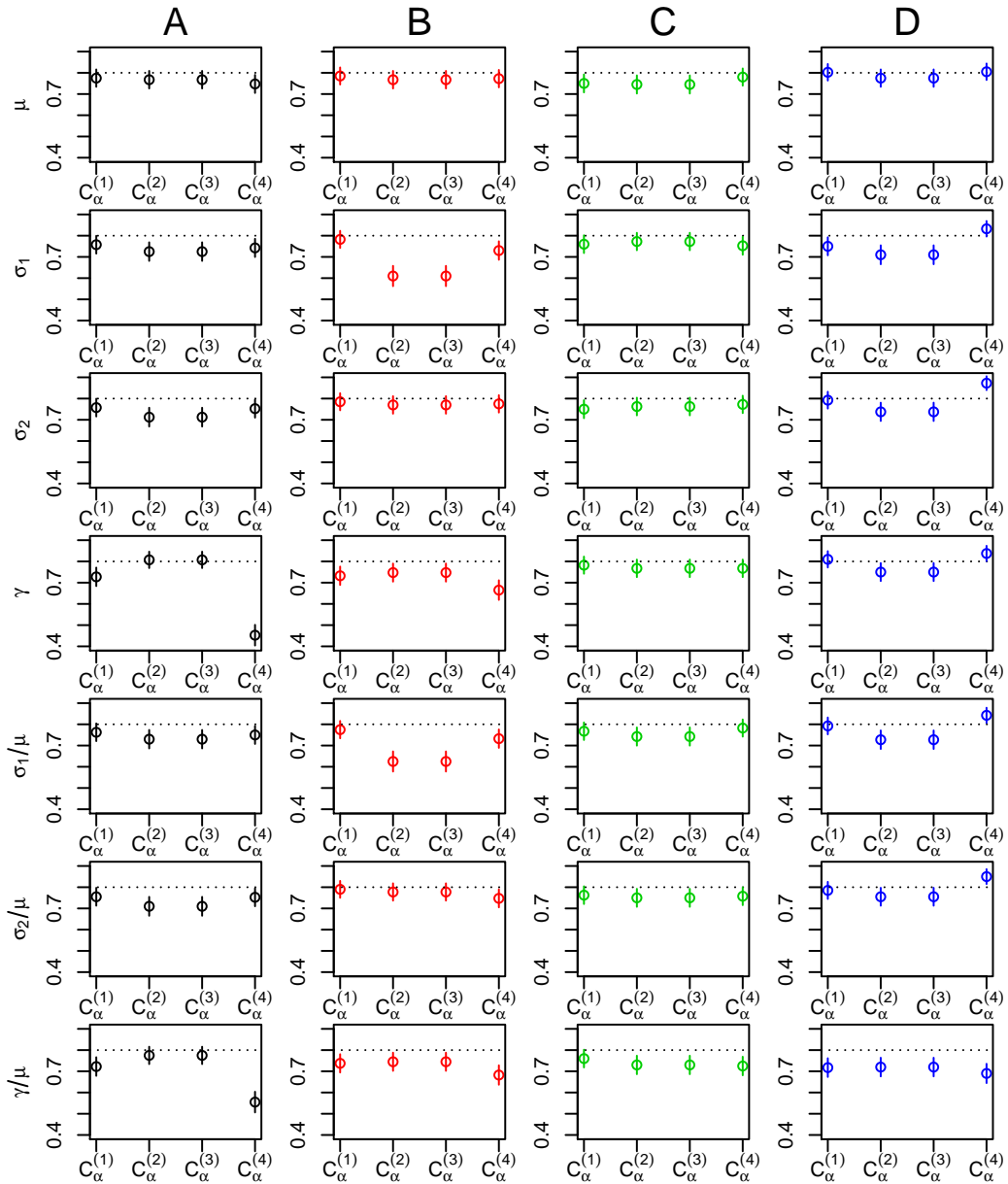


Figure 6.6: Comparison of different bootstrap confidence intervals. For each of the four parameter combinations given in table 5.1 and $\mathcal{T} = 900$ (seconds), $n_1 = 400$ new spike trains are simulated and resampled $b = 1000$ times with the random block bootstrap and with $r = 15$ as the number of drawn blocks. Each column represents one parameter combination and each row a parameter. In every plot, there are coverage probabilities of the four confidence interval types together with a 95% interval for the coverage probability.

Thus, a general procedure is introduced to test $H_0: \gamma_1 = \gamma_2$ and generate approximate p -values. It is motivated by the idea that a confidence interval is equivalent to a hypothesis test as it is mentioned by Wonnacott and Wonnacott (1969) and Efron and Tibshirani (1994). But this procedure should be regarded as a workaround and not as a formal hypothesis test and will not be investigated in further detail.

Testing procedure: Let us suppose that we have an estimated parameter or parameter combination $\hat{\theta}_1$ of a spike train \mathcal{S}_1 and $\hat{\theta}_2$ of a spike train \mathcal{S}_2 . Typically we observe $\hat{d} := \hat{\theta}_1 - \hat{\theta}_2 \neq 0$, thus there seems to be a difference. For testing the null hypothesis $H_0: d := \theta_1 - \theta_2 = 0$ against the alternative $H_1: d \neq 0$, the duality between confidence intervals and statistical tests is used. That is, the null hypothesis $H_0: d = 0$ can be rejected at a significance level α if the corresponding $1 - \alpha$ confidence interval around \hat{d} does not contain 0, and vice versa. To this end, we use the b bootstrap differences $\hat{d}_i := \hat{\theta}_{1,i} - \hat{\theta}_{2,i}$, $i = 1, \dots, b$ derived with some predefined resampling scheme and can construct confidence intervals for d . Let $C_\alpha(\mathbf{d})$ denote a $1 - \alpha$ confidence interval for d depending on the bootstrap estimates $\mathbf{d} = (\hat{d}_1, \dots, \hat{d}_b)'$. The largest confidence interval $C_\alpha(\mathbf{d})$ which does not contain 0 then yields the minimal significance level α at which the corresponding test would be statistically significant and thus, the p -value of the respective test. So the p -value is given as

$$p := \max\{\alpha \in [0, 1] : 0 \notin C_\alpha(\mathbf{d})\}. \quad (6.28)$$

This procedure works with all resampling methods described in Chapter 6.1 and with all types of confidence intervals.

Chapter 7.

Data analysis

In this chapter, the GLO model is applied to the sample data set. It consists of 146 extracellular single-unit recordings of dopamine (DA) neurons of mice in vivo (details are given in appendix A.2). First of all, the parameter estimates are derived by fitting the theoretical ACF to the empirical ACH using the R function `estimation` (as described in Section 5.1). Then, the goodness of fit will be judged by visual inspection of the fitted ACHs and with diagnostic plots (Section 7.1). In the next step, all spike trains are classified into bursty and non-bursty/pacemaker and irregular or regular firing patterns (Section 7.2). The variability of these classifications is visualized and the distribution of the classified firing patterns can be used to compare different groups of neurons (e.g., pre / post or KO/WT). Finally, the differences in the underlying parameters are investigated in order to study neurophysiologically and pharmacologically relevant differences or changes in the firing patterns (Section 7.3).

7.1. Goodness of fit

One of the first important steps is to judge the goodness of fit of the GLO model for data analysis. There is no formal test at the moment, so one has to rely on visual inspection of the fitted ACH. Of course, there is the possibility of checking the assumption of stationarity with a graphic (e.g., see Figure A.1) or a formal test, but it is not recommended, because in practice, only a fraction of spike trains can be regarded as stationary. Furthermore, although there may be non-stationarities in a spike train, the GLO model may still be a good approximation to get an overall summary of the firing patterns.

Figure 7.1 shows diagnostic plots of the GLO ACH fits for 16 different spike trains of the sample data set. Most of the fits look appropriate. Only for spike train L, there

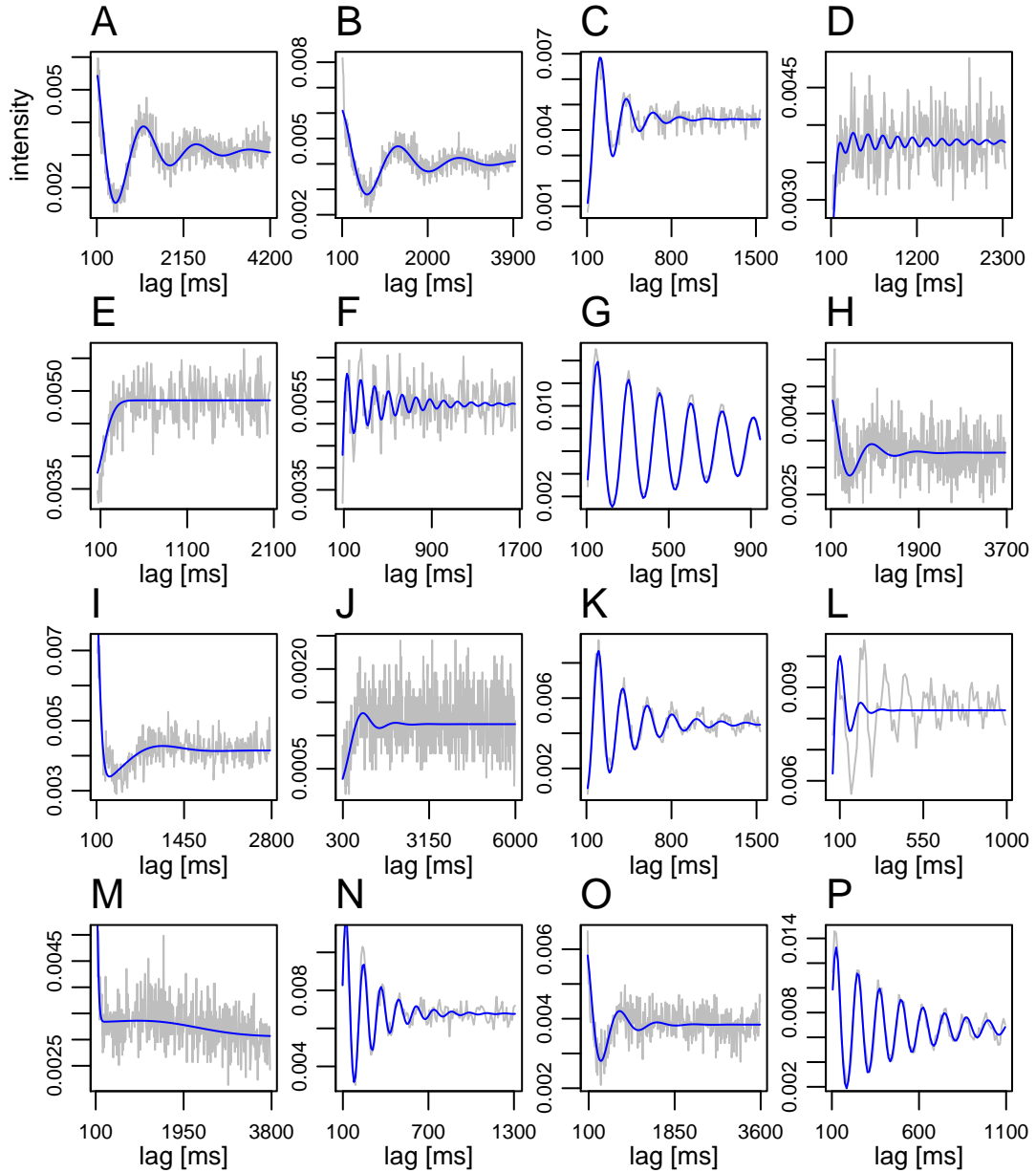


Figure 7.1: Overview of the GLO ACH fits in the sample data set.

could have happened a convergence to a local minimum. In fact, 141 of 146 spike trains have a proper fit. Only one spike train of the sample data set, could not be fitted by the function `estimation`. A possible solution would be to modify the fitting procedure or to add some more starting values. For the remaining four spike trains, one has to decide if they should be excluded of the further analysis or not. Of course, this depends on the individual spike train. Here, it has been decided to remove them.

7.2. Classification of spike trains

Given the estimated parameters $\hat{\psi}$, every spike train is classified into one of the following firing patterns:

- regular non-bursty
- irregular non-bursty
- regular bursty
- irregular bursty

Thresholds: The classification is done according to the GLO classification criteria as described in Section 4.3.2 and summarized in table 4.1. We set the threshold of parameter β_2 as $c_2 = 3$ and term all spike trains with $\beta_2 < 3$ 'irregular single spike' regardless of the remaining parameters. For all other spike trains the parameter m can be used to distinguish between bursty and non-bursty neurons. The threshold c_1 of β_1 was set to 0.4 for bursty and \tilde{c}_1 to 0.35 for non-bursty processes (compare table 4.1) because the latter showed a lower level of β_1 in general (see Figure 7.2 A). Figure 7.2 B shows all estimated values of β_1 and β_2 and table 7.1 summarizes the classification of the GLO: 45 cells were classified as oscillatory single spike, 45 as irregular single spike, 22 as oscillatory bursty and 29 as irregular bursty.

Comparison with visual inspection: In table 7.1, there is also the classification of the 141 spike trains with visual inspection criteria (compare table 4.1). It turns out, that in 104 cases both classification methods give the same results. 24 of the 37 differences come from a different regular/irregular classification. In 10 cases there was a difference between bursty and non-bursty and in only 3 classifications there were complete opposite results.

Variability of classification: While some spike trains can be classified clearly into bursty or non-bursty and regular or irregular, the classification is not always clear cut. The transition between regular and irregular firing patterns is continuous which should be also true for bursty and non-bursty, although this is not directly taken into account

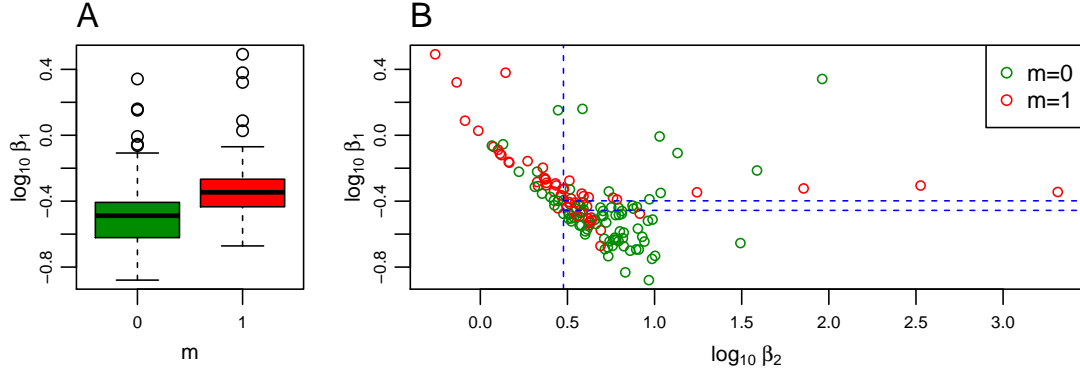


Figure 7.2: Thresholds for GLO classification criteria. (A) In general, the parameter β_1 is larger for bursty neurons ($m = 1$) than for non-bursty cells ($m = 0$). (B) Scatterplot of parameter combinations β_1 and β_2 which are used for classification of the 141 spike trains (see table 4.1). The threshold for β_2 is set to $c_2 = 3$ (vertical blue line). The threshold for c_1 of β_1 is 0.4 for bursty and $\tilde{c}_1 = 0.35$ for non-bursty processes (horizontal blue lines).

| classification | GLO | | | | Σ |
|--------------------------------|-----------------------|-------------------------|-------------------|---------------------|----------|
| | regular non-bursty | irregular non-bursty | regular bursty | irregular bursty | |
| regular non-bursty | 40 | 10 | 0 | 1 | 51 |
| VI irregular non-bursty | 5 | 27 | 0 | 4 | 36 |
| regular bursty | 0 | 2 | 16 | 3 | 21 |
| irregular bursty | 0 | 6 | 6 | 21 | 33 |
| Σ | 45 | 45 | 22 | 29 | 141 |

Table 7.1: Classification of the sample data set by GLO and visual inspection (VI) criteria as described in Section 4.3.

in the GLO model, because m takes only values 0 and 1. Of course, γ or β_2 can be used to measure the degree of burstiness.

However, as a consequence of this classification into categories, some spike trains tend in the simulations to be classified sometimes as bursty and sometimes as non-bursty and the same holds true for regular and irregular firing patterns. This uncertainty in the classification reflects the fact that the threshold is set arbitrarily, and that due to random variation, some processes tend to be classified sometimes left and sometimes right of the boundary. In a similar way, visual inspection will lead to results that are not always reproducible in ambiguous cases.

For quantifying the variability of the classification, we generate $b = 1000$ bootstrap samples \mathcal{S}_{ij}^* of the empirical spike trains \mathcal{S}_i , $i = 1, \dots, 141$ and $j = 1, \dots, b$. For this purpose, the random block bootstrap method is used with $r = 15$ blocks (see Section 6.1.3). The R function `boot2` gives the bootstrap estimates $\hat{\psi}_{ij}^*$, which are used again according to the GLO classification criteria. Thus, for every spike train \mathcal{S}_i , one gets a discrete distribution of the classified firing patterns which represents the uncertainty in the classification. If there is only small variability, one category will have large frequency (in the ideal case close to one).

Figure 7.3 shows the distributions of the classified firing patterns of the bootstrap samples for every spike train \mathcal{S}_i (represented by a vertical bar). Most of the spike trains are clearly dominated by one category. Furthermore, one sees that some cells are sometimes classified as regular and sometimes as irregular. The same holds true for bursty and non-bursty in the irregular case. Finally, one has to mention that although a spike train is classified with a special firing pattern, the bootstrap sample can be classified as something different (e.g. see the last bar in the regular non-bursty category).

Distribution of firing patterns: Finally, the classifications can be used to compare the distribution for different populations of neurons and under different experimental conditions. Typically, each condition or population is represented by a distribution. Thus, if one wants to know if the distribution is independent of the condition, one can perform a chi-square-test for independence. If the assumptions on the cell counts are not fulfilled, a permutation test can be performed instead.

Table 7.2 shows an example for WT and KO neurons in the substantia nigra (SN) with spontaneous activity. In this case, the permutation test gives a p -value of $p = 0.2993$ and thus the null hypothesis of independence cannot be rejected. In contrast, table 7.3 shows a significant ($p = 0.002$) difference between WT and DN cells which are viral treated SN neurons. Of course also other populations and conditions can be compared (e.g. pre/post or VTA/SN).

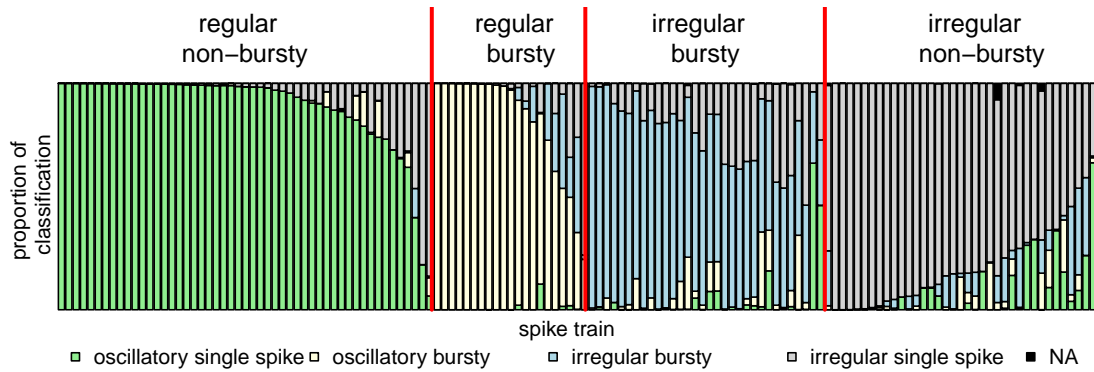


Figure 7.3: Variability of the GLO classification criteria. Every spike train is represented by one vertical bar which represents the distribution of the firing patterns generated by the bootstrap estimates. Red lines separate the four categories of the initial classification of the empirical spike trains.

| SN spon. | regular non-bursty | irregular non-bursty | regular bursty | irregular bursty | Σ |
|----------|--------------------|----------------------|----------------|------------------|----------|
| KO | 19 | 3 | 9 | 7 | 38 |
| WT | 10 | 7 | 8 | 6 | 31 |
| Σ | 29 | 10 | 17 | 13 | 69 |

Table 7.2: Distribution of classified KO and WT spike trains (SN spon.). The permutation test with 5000 simulations results in $p = 0.2993$.

| SN VIRUS | oscillatory single spike | irregular single spike | oscillatory bursty | irregular bursty | Σ |
|----------|--------------------------|------------------------|--------------------|------------------|----------|
| DN | 11 | 1 | 1 | 2 | 15 |
| WT | 1 | 3 | 4 | 6 | 14 |
| Σ | 12 | 4 | 5 | 8 | 29 |

Table 7.3: Distribution of classified DN and WT spike trains (SN viral). The permutation test with 5000 simulations results in $p = 0.002$.

7.3. Comparison of different experimental conditions

In this section, the GLO is used to compare spike trains in different experimental conditions such as pharmacological treatment or genetic manipulation. According to the direct interpretation of the model parameters, one can identify changes in the burst width, the regularity of the background rhythm or other features of the processes. As a small example, the described sample data set contains a subset of six dopaminergic neurons in the substantia nigra whose firing patterns were measured pre and post injection of a dopamine antagonist (see appendix A.2 or Bingmer et al. (2011)). The parameters of the GLO are estimated as described by fitting the theoretical ACFs to the empirical ACHs of the spike train with R function `estimation` (see Section 5.1). $b = 10.000$ bootstrap samples are generated with the random blocks bootstrap method with $r = 15$ blocks (Section 6.1.3). These are used to construct confidence intervals and for statistical hypothesis testing (as described in Section 6.2), whereas the modified standard interval $C_\alpha^{(1')}$ is taken.

The main results are summarized in Figure 7.4 and have been discussed in Bingmer et al. (2011). For every neuron, the change in the parameter estimates is illustrated by a red arrow, beginning at the pre injection estimate and pointing towards the post injection estimate. In the pre injection case, one spike train (a) is classified as regular non-bursty, two spike trains (b & c) are classified as irregular, and three spike trains (d, e & f) are classified as regular bursty (panel A). Only cells d and e change their classification from pre to post treatment, being classified as irregular post treatment. In addition, the cells a, b and c show an increase in their irregularity. All bursty cells show a significant increase in the number of spikes per burst, γ , from pre to post treatment, while the regular non-bursty cell shows a significant decrease in γ . Panel B shows that all cells have an increased mean firing rate after treatment. For the bursters, this is due to the increase in γ , and for the pacemaker it results from a significant decrease in μ (panel C). Finally, the higher number of spikes per burst also results in an increased burst width, σ_2 (panel D). The corresponding p -values of the tests for differences between parameter estimates pre and post injection are shown in table 7.4.

Thus, the proposed GLO and the presented methods for confidence interval construction and statistical testing can be used to identify meaningful changes in the firing behavior of oscillatory spike trains. In the present case, this preliminary analysis suggests that the applied dopamine antagonists might change the mean firing rate of the cells as well as increase the number of spikes per burst, the burst width and the irregularity of the oscillation.

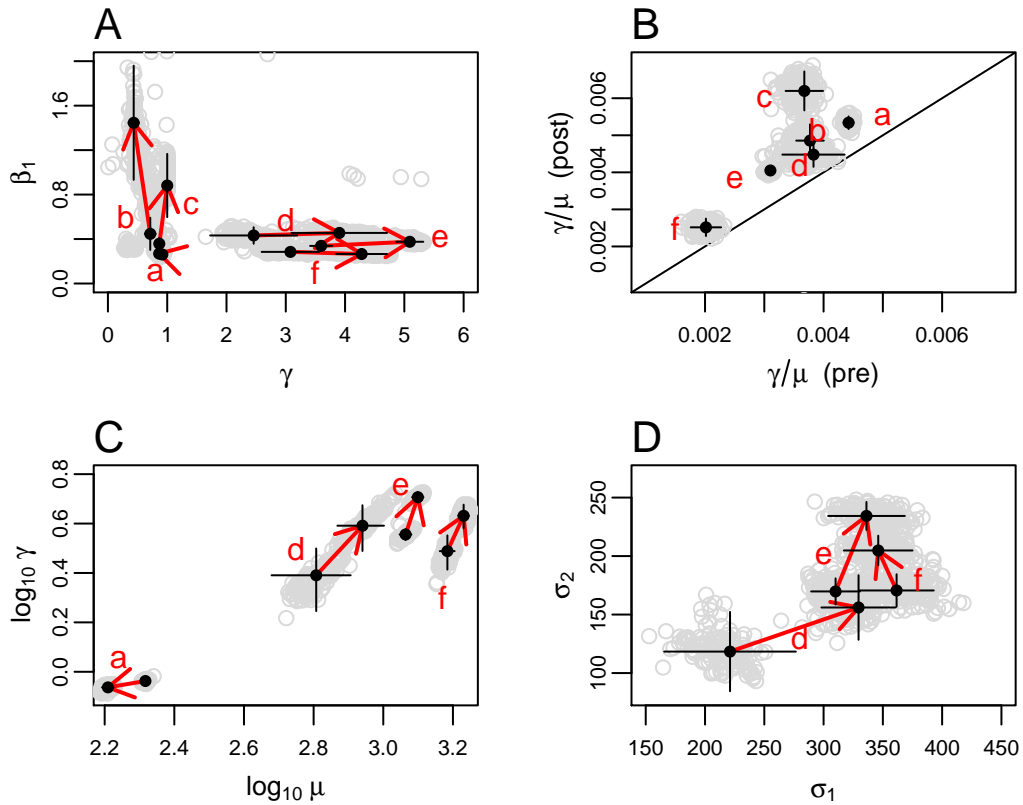


Figure 7.4: Changes in the parameter estimates of six substantia nigra neurons (denoted by a - f, respectively) before and after treatment (begin and end of red arrows). The black lines represent 95% confidence intervals. A) The bursty neurons fire more spikes per burst after treatment. In most cells, irregularity tends to increase. B) The firing rate increases for all cells. C) This increase can be explained by the increase of γ for the bursters. D) For all bursters, the burst width increases.

| | μ | σ_1 | σ_2 | γ | γ/μ | σ_1/μ | σ_2/μ | β_1 |
|---|--------------|--------------|--------------|--------------|--------------|----------------|----------------|--------------|
| a | 0.001 | 0.003 | 0.890 | 0.001 | 0.001 | 0.824 | 0.678 | 0.621 |
| b | 0.001 | 0.893 | 0.039 | 0.002 | 0.001 | 0.570 | 0.004 | 0.027 |
| c | 0.001 | 0.042 | 0.445 | 0.006 | 0.001 | 0.001 | 0.235 | 0.001 |
| d | 0.064 | 0.004 | 0.148 | 0.049 | 0.046 | 0.511 | 0.824 | 0.639 |
| e | 0.010 | 0.249 | 0.001 | 0.001 | 0.001 | 0.996 | 0.001 | 0.031 |
| f | 0.001 | 0.526 | 0.001 | 0.001 | 0.005 | 0.050 | 0.114 | 0.144 |

Table 7.4: p-values of the test between pre and post measurements as described in Section 6.2.3. Bold numbers indicate significant tests on the 5% level.

Chapter 8.

Conclusion

This work proposed a doubly stochastic spike train model to describe regularity and burstiness of neuronal firing patterns within a stationary point process framework. The point process Φ builds up on the stationary background point process $\mathbb{B} = \sum_{i \in \mathbb{Z}} \delta_{B_i}$ which is modeled by a stationary random walk $(B_i)_{i \in \mathbb{Z}}$ with independent normal distributed increments $B_{i+1} - B_i$. The distribution of the number of spikes per beat B_i depends on the firing mode. They are Bernoulli distributed in the non-bursty mode and Poisson distributed in the bursty case. Finally, the spikes are placed around their birth beats according to a normal distribution. Thus, the beats B_i of the BR are independently copied, thinned and then shifted, resulting in the final spike process Φ . It turned out that the GLO process Φ can be regarded as a general cluster process which in case of the bursty firing mode can be further specified as a Cox process. The GLO model uses only five parameters $\psi = (\mu, \sigma_1, \sigma_2, \gamma, m)$ which can be related directly to the underlying properties of the point process Φ : The two parameters μ and σ_1 describe the period and variability of the background rhythm. The parameter m , classifies the process into a bursty or non-bursty firing pattern. In both cases, the parameter γ denotes the expected number of spikes per background beat. Finally, σ_2 represents the variability of spikes around their background beat ($m = 0$) or the width of the bursts in the bursty firing mode ($m = 1$).

The assumptions are chosen to reproduce important observed characteristics of the empirical data set and to have easily interpretable parameters. But because the normal distribution is defined on \mathbb{R} , there is always a positive probability to have negative increments $B_{i+1} - B_i < 0$ in the background rhythm. A negative increment only implies that the corresponding beats occur in different order than they were originally generated in the model. So, the distribution of ordered interbeat intervals differs from the normal distribution because it takes only positive values. Therefore, the model is also able to describe activity that is not locked to a regularly oscillating backbone process. Anyway, the two parameters of the normal distribution are considered a simple and sufficiently

general description in order to grasp the period and the precision of the background random walk.

Although these assumptions make it difficult to derive the ISI distribution in the general case, the intensity of spikes at time t can be computed conditioned on a spike at 0. In a neuroscientific context, this is called the autocorrelation function (ACF; Gerstein and Kiang (1960)) and is basically an intensity function for the points of the process given a point at 0. It is similar to the autocorrelation function of time series analysis (Brockwell and Davis, 2010), but should not be confounded with it. The autocorrelation histogram (ACH; Moore et al. (1966); Perkel et al. (1967)) is an estimate of the ACF and often used to classify a spike train with visual inspection criteria into regular or irregular and bursty or non-bursty (Wilson et al., 1977; Gray et al., 1992; König, 1994; Paladini et al., 2003). According to the ACF of a GLO process, classification rules has been stated in terms of the model parameters which are similar to these criteria. For getting estimates of the parameters, the ACF f_m of a GLO process can be fitted to the ACH h of an observed process. This is done with a nonlinear least squares algorithm by minimizing the weighted residual sum of squares

$$\sum_{l \in \mathcal{L}} v_l (h(l) - f_m(l))^2.$$

For this purpose, the ACH h is smoothed with a gaussian filter to get a smoother estimate \hat{f}_m for f_m , because the variance of h is approximately given by f_m . Thus, the weights v_l are given by the reciprocal of \hat{f}_m . Furthermore, a large set of starting values is used for the fitting procedure. The precision of this procedure has been investigated for several experimental conditions. In addition, different methods for resampling and confidence interval construction have been tried, to represent the variability of these parameter estimates.

Fitting procedures of ACFs are often used in neuroscientific practice (Engel et al., 1992; König, 1994; Celada et al., 1999; Hyland et al., 2002; Paladini et al., 2003; Schneider and Nikolic, 2006). For the GLO model, the current implementation tries to estimate the parameters in an almost automatic way and should work well with other empirical spike trains, although it can be improved. For example, the bin size δ should be chosen in an automatic way and be spike train dependent or the algorithm be implemented more efficiently. Furthermore, the choice of the analysis window \mathcal{L} can be analyzed in a more systematic way and optimized. While at the moment, the time needed for the parameter estimation and simulation is important for empirical practice, this will be negligible as computers get faster in the future. In this case, it will be preferred to start the fitting procedure with all starting values, instead of just taking one. However, the analyses have shown, that the current fitting procedure gives accurate and reproducible results in most of the cases. The parameter estimates that result from such a fit indi-

cate spike properties on a continuous scale. This allows an objective quantification of regularity and burstiness that are usually, in an experimental context, described and classified on the basis of visual inspection of the ACHs (Wilson et al., 1977; Paladini et al., 2003). Thus, the present model provides a classification tool that can support usual classification tools in an objective way (Bingmer et al., 2011; Schiemann et al., 2012; Schiemann, 2012). In addition to the classification of a spike train into bursty or single spike and into regular/oscillatory or irregular, this classification and parameter estimation is associated with confidence intervals and thus, with a measure of uncertainty of classification. Furthermore, the parameter σ_2 can be used for firing rate kernel estimation. Typically the bandwidth of the kernel is unknown, but by choosing a normal distribution as the kernel, σ_2 is an appropriate candidate for the bandwidth, because it represents the random variation of a spike.

By applying the GLO to a sample data set, it has been shown that its minimal set of assumptions can be very useful in reproducing empirically obtained spike trains and thus, describe and quantify these processes to a large extent, in particular for dopaminergic neurons (Bingmer et al., 2011; Schiemann et al., 2012; Schiemann, 2012). The reproducible firing patterns include approximate normal ISI distributions in the single spike case, the classical bi- or trimodal ISI distributions for the bursty mode and even Poisson process-like patterns under high irregularity. The objective classification mirrors the visual inspection used in experimental contexts. Moreover, the model allows to analyze the effect on burstiness and regularity for different experimental conditions like pharmacological treatment or genetical manipulation.

The measure of burstiness that we propose here is, in contrast to other methods (Legendy and Salcman, 1985; Gourevitch and Eggermont, 2007), not based on an algorithm that detects the bursts. In the present approach, burst detection is not necessary because burstiness is grasped by the model parameters, which can be estimated directly from the ACH. For example, the parameter γ directly estimates the average number of spikes per burst and σ_2 is connected with the burst width. As a consequence, the present approach avoids several difficulties that may arise in classical burst detection algorithms, including subjective variability from visual inspection, rate dependence from fixed criteria or manual parameter adjustment. While algorithmic approaches usually suffer from the lack of a burst definition, the proposed spike train model conceptually provides a definition of what constitutes a burst, namely all spikes that originate in the same background beat. But because empirically it is not observable, which spikes belong to which unobservable background beat, direct burst detection is difficult. However, the model offers the possibility to use surrogate data in order to develop and to validate an algorithm that can identify the bursts.

Some properties of experimental spike trains cannot be represented with the present

model. For example, the refractory period between successive spikes is not taken into account, which needs to be considered in the fitting procedure by ignoring small lags in the ACH. In addition, the present model clearly focusses on oscillatory processes by assuming a normal distribution in the backbone random walk and small values of σ_1/μ . The model thus intends to describe either regular pacemakers or oscillatory bursters, and the less regular a process becomes, the less precisely the parameters can be estimated. However, the model is also able to represent higher irregularity in the backbone rhythm by increasing σ_1/μ . One should note that the present model assumes stationarity of the firing rate and of all other parameters. Although it describes additional patterns of rhythmic and bursting properties, these properties are assumed constant in time. However, these assumptions can be loosened to some extent: A change of the number of spikes per burst γ throughout the process does not affect the fitting procedure or the estimates as long as all other parameters remain constant. In this case, the estimate of γ still describes the mean number of spikes per backbone beat and thus, is robust against non-stationarity.

In conclusion, the present model provides a tool for the description and quantification of regular oscillatory activity in bursty and non-bursty processes. Because of its high similarity to subjective procedures in experimental practice, it can provide an objective measure for burstiness and regularity that corresponds closely to visual inspection criteria. Finally, it can be helpful in the identification of changes in the parameters of the underlying processes and thus, provide a measure to investigate functional changes in neuronal spiking behavior in response to external conditions.

Appendix A.

Additional information

A.1. Bursts and regularity

Bursting was first described for slow-wave sleep and drowsiness (Livingstone and Hubel, 1981; Steriade et al., 1993), but recent data on cats and monkeys (Guido and Weyand, 1995; Edeline et al., 2000; Ramcharan et al., 2000) indicate it as a relay mode in the awake state (Sherman, 1996). Bursts have also been observed for thalamic neurons in awake humans under different pathological conditions (Lenz et al., 1998; Radhakrishnan et al., 1999; Magnin et al., 2000). Although the physiological correlates of bursting remain unclear (Izhikevich et al., 2003), it is considered as an important unit of neuronal information (Lisman, 1997).

While a burst lacks a clear definition (Gourevitch and Eggermont, 2007), it is commonly referred to periods during which spike frequency is relatively high, separated by periods during which frequency is relatively low (Cocatre-Zilgien and Delcomyn, 1992). Because of the flexible definition, there is a variety of methods to detect and measure bursts: power spectrum analysis (Bair et al., 1994), comparison of the variance of the sum of two interspike intervals (ISIs) and the sum of variance of ISIs (van Elburg and van Ooyen, 2004), detection of a critical interval value in the ISI distribution (Cocatre-Zilgien and Delcomyn, 1992), a minimum number of ISIs all shorter than a given value (Harris et al., 2001; Kepecs and Lisman, 2004; Chiappalone et al., 2005), empirical criteria (Grace and Bunney, 1984), the Poisson surprise method (Legendy and Salzman, 1985) and the rank surprise method (Gourevitch and Eggermont, 2007).

Spiking neurons also vary in the regularity of interspike intervals (Grigoryan et al., 2007) which is represented by a continuum of firing patterns (Maimon and Assad, 2009). On one side are Poisson like cells which arise most often in visual cortex, have a variance-to-mean ratio of spike counts equal to unity (Buracas et al., 1998) and are posited as

the fundamental unit of cortical communication (Geisler et al., 2005). At the other end are clock like neurons that are common in the peripheral nervous system (Werner and Mountcastle, 1965).

The regularity of a spike train is analyzed by different approaches: The coefficient of variation (CV) of ISIs is used to characterize the temporal structure of neurons. Other methods are the squared CV (Nawrot et al., 2008), comparison of adjacent ISIs (Holt et al., 1996), shape parameter of ISIs modelled by a gamma distribution (Shimokawa and Shinomoto, 2009; Maimon and Assad, 2009), the Fano factor, i.e. the variance divided by the mean of the spike count (Buracas et al., 1998), different irregularity metrics (Davies et al., 2006) and measures of local variations (Shinomoto et al., 2005).

A.2. The sample data set

The data set consists of 146 extracellular single-unit recordings of dopamine (DA) neurons of mice in vivo and has been collected by Schiemann (2012). The measurements were performed on anesthetized adult (10 - 16 weeks) male C57bl6 mice, a common inbred strain of laboratory mice, we denote as wildtype (WT). Furthermore, gene manipulated ATP sensitive potassium (K-ATP) channel knockout (KO) mice were used for in vivo recordings. These general KO-mice lack the pore-forming Kir6.2 subunit of K-ATP channels. The regions of interest in the brain are the substantia nigra pars compacta (SN) and the ventral tegmental area (VTA) which are both located in the midbrain, have a high density of DA neurons and which, together with their projections to multiple brain areas, compose the dopaminergic system.

For most neurons, the spontaneous activity (spon.) has been recorded. For some cells a dopamine D2-receptor antagonist (eticlopride) was injected during recording to obtain measurements of basal activity (pre) and after pharmacological D2-receptor inhibition (post). Furthermore, for silencing of K-ATP channel function selectively in SN DA neurons, a viral approach was chosen (virus). rAAV2-mediated (recombinant adeno-associated virus, serotype 2) gene transfer was established for dopamine cell-selective expression of dominant-negative (DN) pore-mutant Kir6.2 subunits. Electrophysiological single-unit activity of these virally transduced DA neurons were also recorded in vivo, whereas cells expressing native Kir6.2 subunits (WT) were used as control.

Table A.1 shows a summary of experimental conditions, the number of cells, the range of the recording time and the mean and standard deviation of the number of spikes. Figure A.1 shows the variation of the firing rates for different regions and conditions.

| Region | Type | Condition | #Cells | recording time | #Spikes (mean \pm SD) |
|--------|------|-----------|--------|----------------|-------------------------|
| SN | WT | spon. | 32 | 540s - 900s | 3016 \pm 1294 |
| SN | KO | spon. | 40 | 480s - 900s | 3158 \pm 1178 |
| VTA | WT | spon. | 14 | 600s - 900s | 3671 \pm 1844 |
| VTA | KO | spon. | 18 | 660s - 720s | 2318 \pm 1545 |
| SN | WT | virus | 14 | 600s - 780s | 2941 \pm 905 |
| SN | DN | virus | 16 | 720s - 780s | 3423 \pm 1465 |
| SN | WT | pre | 6 | 720s - 840s | 2475 \pm 596 |
| SN | WT | post | 6 | 1036s - 1260s | 5279 \pm 1634 |

Table A.1: Summary of the empirical data set of Schieman (2012). The spike trains are recorded in the substantia nigra (SN) or ventral tegmental area (VTA) for wildtype (WT), *K-ATP* channel knockout (KO) mice or dominant-negative (DN) pore-mutant *Kir6.2* subunits. For most cells only the spontaneous activity (spon.) has been recorded. For some cells measurements of basal activity (pre) and after injection of a D2-receptor antagonist (post) are available and some cells got a viral transduction (virus).

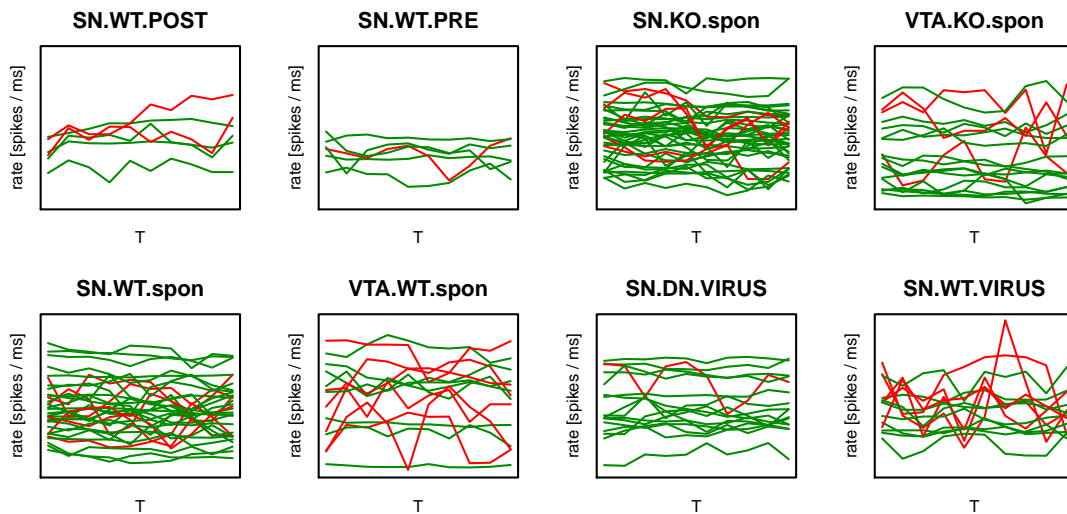


Figure A.1: Empirical firing rates for different brain regions and experimental conditions (see table A.1). Most of the spike trains have a firing rate approximately constant. Every spike train is divided in ten parts and for every part the mean firing rate is estimated. Then, for all spike trains the variability of the firing rate is computed and all spike trains having a variability larger than the 20%-quantile are shown in red.

A.3. R Code

The current section contains a collection of R Code which has been developed for the underlying study. This allows a reproduction of the results in this thesis. The background details of these R functions are discussed in the connected chapters. The computations have been performed on R version 2.13.1. For a general introduction to R, there are the books of Chambers (2008) and Kabacoff (2010). Ritz and Streibig (2008) gives details about nonlinear regression with R and Robert and Casella (2010) about Monte Carlo methods.

```
SIM <- function(psi,T,K=1000,eps=1*10^(-6)){

  mu <- psi[1]
  s1 <- psi[2]
  s2 <- psi[3]
  gam <- psi[4]
  m <- psi[5]

  q <- qnorm(eps)
  k <- 0
  B <- -K*mu
  K2 <- max(c(T-q*s2, T+s1^2/(2*mu)*log(eps)))

  t0 <- rep(NA,10^7)

  while (B < K2) {
    if (m==0) {N <- rbinom(1,1,gam)}
      else {N <- rpois(1,gam)}
    if (N>0) {
      Z <- rnorm(N,0,s2)
      for (i in 1:N) {t0[k+i] <- B+Z[i]}; k <- k+N}
    B <- B+rnorm(1,mu,s1)
  }
  t0 <- na.omit(t0)

  return(sort(t0[t0>=0 & t0 <=T]))
}
```

Listing A.1: INPUT: GLO parameters $\psi = (\mu, \sigma_1, \sigma_2, \gamma, m)$ and recording time \mathcal{T} . OUTPUT: Spike train $\mathcal{S} = (t_1, \dots, t_n)$. REMARKS: For details see Algorithm 1 on page 41 in Section 2.3.1.


```
GLO <- function(psi,T,K1=1000,eps=1*10^(-6)){
  mu <- psi[1]; s1 <- psi[2]; s2 <- psi[3]
  gam <- psi[4]; m <- psi[5]; q <- qnorm(eps)

  K2 <- max(c(T-q*s2, T+s1^2/(2*mu)*log(eps)))
  K <- K2 + K1*mu; k <- (2*K*mu+(q*s1)^2)/(2*mu^2)
  k <- ceiling(k+sqrt((q*s1)^2*(4*K+(q*s1)^2))/(2*mu))

  B <- cumsum(rnorm(k,mu,s1))-K1*mu
  B <- B[B >= q*s2 & B <= T -q*s2]
  N <- length(B)
  if (m==0){ P <- rbinom(N,1,gam)
    B <- B[P > 0]}
  else { P <- rpois(N,gam); P0 <- P > 0
    B <- rep(B[P0],P[P0])}
  B <- sort(B + rnorm(length(B),0,s2))
  return(B[B >= 0 & B <= T ])}

```

Listing A.2: INPUT: GLO parameters $\psi = (\mu, \sigma_1, \sigma_2, \gamma, m)$ and recording time \mathcal{T} . OUTPUT: Spike train $\mathcal{S} = (t_1, \dots, t_n)$. REMARKS: More efficient than listing A.1 by using vector-valued programming (see Section 2.3.2).

```
ACF <- function(psi,L,eps=10^(-300),B=100){
  R <- c(0,max(L))
  K <- rep(NA,2)

  for (i in 1:2) {
    u <- ceiling(R[i]/psi[1]); k <- u:(u+B*(-1)^i)
    b <-dnorm(R,k*psi[1],sqrt(abs(k)*psi[2]^2+2*psi[3]^2))
    if (all(b>= eps)) { K[i] <- k[length(k)]}
    else {K[i] <- k[which(b<eps)[1]]}
  }

  K <- K[1]:K[2]; if (psi[5]==0) {K <- K[K!=0]}
  sdK <- sqrt(abs(K)*psi[2]^2 + 2*psi[3]^2)
  fkt <- function(x){ sum(dnorm(x,mean=K*psi[1],sd=sdK))}
  tmp <- apply(as.matrix(L),1,fkt)*psi[4]
  return(tmp)}

```

Listing A.3: INPUT: Parameters $\psi = (\mu, \sigma_1, \sigma_2, \gamma, m)$ and set of lags \mathcal{L} . OUTPUT: Values $f_m(l)$ for all $l \in \mathcal{L}$. REMARKS: For theoretical details see Section 4.1.

```

ACH <- function(x,L=1:1500,norm=TRUE,plot=TRUE){
  minD <- 0; k <- 1; n <- length(x); T0 <- max(x)
  delta <- L[2]-L[1]; K <- length(L)
  L0 <- c(L-delta/2,L[length(L)]+delta/2)
  l_b <- min(L0); l_e <- max(L0)
  anz <- rep(0,length(L))

  while (minD <= l_e & k < n){
    D <- diff(x,lag=k)
    minD <- min(D)
    D <- D[D >= l_b & D <= l_e]
    anz <- anz + hist(D,breaks=L0,plot=FALSE)$counts
    k <- k+1}
  if (norm) {
    K1 <- round((l_b-delta/2)/delta)
    y <- (n*delta*(T0-(K1:(K1+K-1))*delta-delta/2))
    anz <- anz*T0/y}
  if (plot) {plot(L,anz,type="l")}
  return(anz)
}

```

Listing A.4: INPUT: Spike train $\mathcal{S} = (t_1, t_2, \dots, t_n)$ and set of lags \mathcal{L} . OUTPUT: ACH values $h(l)$ for all $l \in \mathcal{L}$. REMARKS: See Section 4.2 for details.

```

Awindow <- function(x,delta=10) {
  isi <- diff(x); cv <- sd(isi)/mean(isi)
  d <- density(isi)
  lb1 <- min(c(quantile(isi,0.05),100))
  lb2 <- d$x[which.max(d$y)]
  lb3 <- lb1 +(cv-0.4)/0.8 *(lb2-lb1)
  if (cv<=0.4) { lb<-lb1 }
  if (cv>1.2) { lb<-lb2 }
  if (cv >0.4 & cv <=1.2) { lb<-lb3 }
  l_e0 <- min(c(6000,quantile(isi,0.9))*5)
  le <- lb + floor((l_e0-lb)/delta)*delta
  L <- seq(lb+delta/2,le-delta/2,delta)
  return(L)
}

```

Listing A.5: INPUT: Spike train $\mathcal{S} = (t_1, t_2, \dots, t_n)$ and bin size δ . OUTPUT: Analysis window \mathcal{L} . REMARKS: Compare Section 5.1.1.

```
Weights <- function(achy) {
  V <- filter(achy, filter=dnorm(-3:3))

  V[c(1:3, (length(achy)-2):length(achy))] <-
    c(achy[1:3], achy[(length(achy)-2):length(achy)])

  V[V==0] <- min(V[V!=0])
  V      <- as.vector(1/V)

  return(V)
}
```

Listing A.6: INPUT: ACH values $h(l)$. OUTPUT: The weights v_l for weighted non-linear regression. REMARKS: See Section 5.1.3 for details.

```
spektral <- function(x) {
  maxT <- 2^floor(log2(ceiling(max(x))))
  x0 <- x[x<=maxT]
  h <- hist(x0, breaks=0:maxT, plot=FALSE)

  ss <- spectrum(h$counts, plot=FALSE)
  freq <- ss$freq
  auswahl <- 1/freq>200 & 1/freq<3000

  mu <- 1/ss$freq[auswahl][which.max(ss$spec[auswahl])]

  return(mu)
}
```

Listing A.7: INPUT: Spike train $\mathcal{S} = (t_1, t_2, \dots, t_n)$. OUTPUT: The starting value $\tilde{\mu}_1$ for the bursty case. REMARKS: See Section 5.1.4 for details.

```
GA <- function(gam) {
  gam <- abs(gam)
  y <- trunc(gam)

  if (y%%2==0) {return(gam-y)}
  else {return(1-(gam-y))}
}
```

Listing A.8: INPUT: Parameter γ . OUTPUT: Transformed γ . REMARKS: Ensures that γ is between 0 and 1 in the non-bursty case.

```

Svalues <- function(x) {
  isi <- diff(x)
  cv <- sd(isi)/mean(isi)
  d <- density(isi)
  lam0 <- mean(ACH(x,L=5000:6000,plot=FALSE))

  if (cv<=0.4) {
    mu00 <- max(c(d$x[which.max(d$y)],90))
    ga00 <- mu00 * lam0
    s22 <- s11 <- c(c(1,3)/100,seq(5,30,5)/100,1)
    mat <- as.matrix(expand.grid(s11,s22))*mu00
    n <- dim(mat)[1]
    sv <- cbind(rep(mu00,n),mat,rep(ga00,n),rep(0,n))}

  if (cv>1.2) {
    mu00 <- spektral(x)
    ga00 <- mu00 * lam0
    s22 <- s11<-c(1/100,seq(5,35,5)/100,seq(40,75,5)/100,1)
    mat <- as.matrix(expand.grid(s11,s22))*mu00
    n <- dim(mat)[1]
    sv <- cbind(rep(mu00,n),mat,rep(ga00,n),rep(1,n))}

  if (cv >0.4 & cv <=1.2) {
    mu00 <- max(c(d$x[which.max(d$y)],90))
    ga00 <- mu00 * lam0
    s22 <- s11<-c(1/100,seq(4,22,3)/100,seq(25,75,5)/100,1)
    mat <- as.matrix(expand.grid(s11,s22))
    n <- dim(mat)[1]
    sv1 <- cbind(rep(mu00,n),mat*mu00,rep(ga00,n),rep(0,n))

    mu00 <- spektral(x)
    ga00 <- mu00 * lam0
    sv2 <- cbind(rep(mu00,n),mat*mu00,rep(ga00,n),rep(1,n))
    sv3 <- cbind(rep(mu00,n),mat*mu00,rep(ga00,n),rep(0,n))
    sv <- rbind(sv1,sv2,sv3)}
  return(sv)
}

```

Listing A.9: INPUT: Spike train $\mathcal{S} = (t_1, t_2, \dots, t_n)$. OUTPUT: Starting values for weighted nonlinear regression. REMARKS: See description in Section 5.1.4.

```

estimation <- function(x,delta=10) {
  L      <- Awindow(x,delta=delta)
  achy   <- ACH(x,L=L)
  V      <- Weights(achy)
  sv     <- Svalues(x)

  aus    <- function(x) {return(sum((ACF(x,L=L)-achy)^2))}
  i0     <- which.min(apply(sv,1,aus)); sv<-as.numeric(sv[i0,])

  nl0    <- try(nls(achy~ACF(c(mu,s1,s2,gam,sv[5]),L=L),
    start=list(mu=sv[1],s1=sv[2],s2=sv[3],gam=sv[4]),
    weights=V,control=list(minFactor=1/2048/2,
    warnOnly=TRUE)),TRUE)

  if (class(nl0)=="try-error")
  {list(parameter=rep(NA,5),nls=NA)} else
  {para <- c(abs(coefficients(nl0)),m=sv[5])
  if (sv[5]==0) {para[4]<-GA(para[4])}
  list(parameter=para,nls=nl0)}
}

```

Listing A.10: INPUT: Spike train $\mathcal{S} = (t_1, t_2, \dots, t_n)$ and bin size δ . OUTPUT: The estimated GLO parameters $\hat{\psi}$ and details of the nonlinear regression procedure. REMARKS: Compare Section 5.1.

```

boot1<- function(para,b=1000,T0=600000) {
  mat <- matrix(NA,b,5,dimnames=list(
    c(),c("mu","s1","s2","gam","m")))

  if (all(!is.na(para))) {
    for (j in 1:b){
      spike <- GLO(para,T=T0)
      mat[j,] <- estimation(spike,delta=10)$parameter}
  }
  return(mat)
}

```

Listing A.11: INPUT: Parameter vector $\psi = (\mu, \sigma_1, \sigma_2, \gamma, m)$, number of simulations b and recording time \mathcal{T} . OUTPUT: Estimated parameters $\hat{\psi}_i, i = 1, \dots, b$. REMARKS: Parametric bootstrap and GLO parameter estimation as described in Section 6.1.4.

```

boot2 <- function(x,b=1000,r=15,T0=600000,delta=10) {
  t0 <- T0/r
  mat <- matrix(NA,b,5,dimnames=list(
    c(),c("mu","s1","s2","gam","m")))
  for (j in 1:b) {
    U <- runif(r,0,T0-t0)
    spike <- rep(NA,length(x)*10)
    k <- 0
    for (i in 1:r) {
      auswahl <- x>=U[i] & x <= U[i]+ t0
      if (any(auswahl)) {
        y <- x[auswahl] - U[i]
        ni <- length(y)
        spike[(k+1):(k+ni)] <- y+t0*(i-1)
        k <- k+ni}
      }
    spike <- as.vector(na.omit(spike))
    mat[j,] <- estimation(spike,delta=delta)$parameter}
  return(mat)}

```

Listing A.12: INPUT: Spike train $\mathcal{S} = (t_1, t_2, \dots, t_n)$, number of simulations b , number of blocks r , recording time \mathcal{T} and bin size δ . OUTPUT: Estimated parameters $\hat{\psi}_i$, $i = 1, \dots, b$. REMARKS: Moving blocks bootstrap of spike trains and GLO parameter estimation as described in Section 6.1.3.

```

mark <- function(x,delta=10) {
  br1 <- Awindow(x,delta=delta)
  br1 <- c(br1-delta/2,max(br1)+delta/2)

  count <- function(i) {
    y <- x-x[i]
    y <- y[y>=min(br1) & y<=max(br1)]
    tmp <- hist(y,breaks=br1,plot=F)$counts
    return(tmp)}

  return(sapply(1:length(x),count))
}

```

Listing A.13: INPUT: Spike train \mathcal{S} and bin size δ . OUTPUT: Marks $h_i(l)$. REMARKS: Construction of the marked point process as described in Section 6.1.4.

```

boot3<-function(x,b=2000,delta=10,r=20,T0=max(x)){

  mat <- matrix(NA,b,5,dimnames=list(
    c(),c("mu","s1","s2","gam","m")))
  MARK <- mark(x,delta=delta)
  L <- Awindow(x,delta=delta)
  sv <- Svalues(x)

  for (j in 1:b){
    U <- runif(r,0,T0-T0/r)
    index <- unlist(apply(t(U),2,
      function(y){which(y<x&y+T0/r>x)}))
    COUNT <- MARK[,index]
    achy <- apply(COUNT,1,sum)/dim(COUNT)[2]/delta
    aus <- function(x){return(sum((ACF(x,L=L)-achy)^2))}
    i0 <- which.min(apply(sv,1,aus))
    V <- Weights(achy)

    nl0 <- try(nls(achy~ACF(c(mu,s1,s2,gam,sv[i0,5]),L=L),
      start=list(mu=sv[i0,1],s1=sv[i0,2],
        s2=sv[i0,3],gam=sv[i0,4]),weights=V,
      control=list(minFactor=1/2048/2,warnOnly=TRUE,
        maxiter=100)),TRUE)

    if (class(nl0)=="try-error") {mat[j,] <- rep(NA,5)}
    else {
      para <- c(abs(coefficients(nl0)),m=sv[i0,5])
      if (sv[i0,5]==0) {para[4] <- GA(para[4])}
      mat[j,] <- para}
  }

  return(mat)
}

```

Listing A.14: INPUT: Spike train $\mathcal{S} = (t_1, t_2, \dots, t_n)$, number of simulations b , bin size δ , number of blocks r and recording time \mathcal{T} . OUTPUT: Estimated parameters $\hat{\psi}_i$, $i = 1, \dots, b$. REMARKS: Bootstrap of spike train and GLO parameter estimation as described in Section 6.1.4.

German Summary

Diese Arbeit schlägt ein stochastisches Modell zur Quantifizierung und Klassifikation von Feuermustern in Spike Trains einzelner dopaminerger Neuronen vor. *Kapitel 1* gibt eine kurze Einführung in die Motivation der Arbeit und der Spike Train Analyse. Ein 'Spike' ist die elektrische Entladung eines Neurons und die Basis der Informationsverarbeitung im Gehirn. Die mathematische Beschreibung der Spikes konzentriert sich in der Regel auf die Zeitpunkte $0 \leq t_1 < t_2 < \dots < t_n \leq \mathcal{T}$ zu denen die Spikes im betrachteten Zeitintervall $[0, \mathcal{T}]$, $\mathcal{T} \in (0, \infty)$, auftauchen. Ein 'Spike Train' $\mathcal{S} = (t_1, t_2, \dots, t_n)$ ist die endliche Folge der Spike Zeitpunkte und wird mit der Punktconfiguration

$$\varphi = \sum_i \delta_{t_i} \quad (0.1)$$

identifiziert, wobei δ_{t_i} das Dirac Maß ist. Der Zufall wird dadurch eingeführt, dass man die Zeitpunkte t_i als \mathbb{R} -wertige Zufallsvariablen T_i auffasst. In diesem Fall wird ein Spike Train als zufälliger Punktprozess

$$\Phi = \sum_i \delta_{T_i} \quad (0.2)$$

modelliert. Übliche Spike Train Modelle sind pseudo-Markov Modelle (Ekholm und Hyvärinen, 1970), Poisson Prozesse (Abeles, 1982), hidden Markov Modelle (Camproux et al., 1996) und Gamma Prozesse (Barbieri et al., 2001; Shimokawa und Shinomoto, 2009).

Diese Modelle werden in der Spike Train Analyse (Johnson, 1996; Awiszus, 1997; Gabbiani und Koch, 1998; Brown et al., 2004) für die Suche nach Mustern in Spike Trains genutzt, die die neuronale Funktionsweise im Gehirn reflektieren. Diese Feuermuster (vgl. Abbildung 0.2) sind unter anderem sehr regelmäßig (A), können aber auch Cluster von Spikes (sogenannte Bursts) beinhalten (B & C) oder sehr unregelmäßig sein (D). Dabei wird die Regelmäßigkeit als oszillatorische Aktivität verstanden und kann auch in Kombination mit Bursts auftreten (B). Beide Konzepte (Bursts und Regelmäßigkeit) haben jedoch keine eindeutige Definition in der Literatur (Gourevitch und Eggermont, 2007) und repräsentieren ein Kontinuum von Feuermustern (Maimon und Assad, 2009).

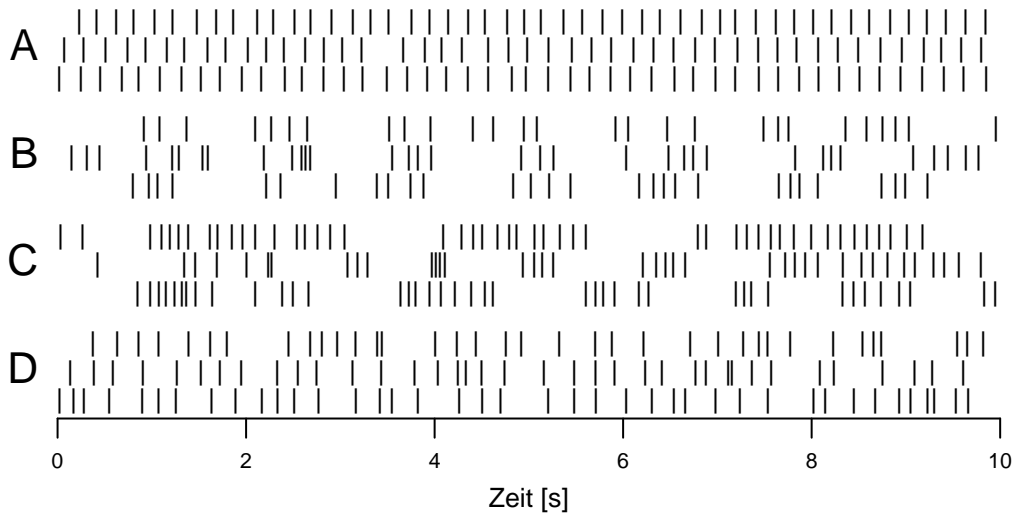


Abbildung 0.2: *Typische Spike Train Feuermuster. Jeder Strich repräsentiert einen Spike Zeitpunkt.*

Bursts werden oft als Antwort zu verhaltensrelevanten Stimuli abgegeben, wie z.B. eine unerwartete Belohnung oder ein neuer Reiz (Schultz, 2007; Bromberg-Martin et al., 2010), können aber auch spontan auftreten. Die oszillatorische Aktivität wurde als relevant für die komplexe Informationsverarbeitung im visuellen Kortex entdeckt, wozu z.B. Bindungsprozesse oder Objektunterscheidung (Gray und Singer, 1989; Engel et al., 1991) gehören. Deswegen ist es im Bereich der Neurophysiologie von Interesse die Feuermuster unter verschiedenen experimentellen Bedingungen zu untersuchen und zu klassifizieren. Dazu gehören pharmazeutische Behandlungen (Bingmer et al., 2011; Schiemann et al., 2012), genetische Veränderungen (Schiemann, 2012), Stimulation (Gray und Singer, 1989; Berger et al., 1990; König et al., 1995) sowie die Betrachtung verschiedener Hirnregionen (Shinomoto et al., 2009). Dabei gibt es die folgenden wichtigen Fragen: Wie unterscheiden sich die Feuermuster in diesen Fällen? Werden Sie regelmäßiger oder unregelmäßiger? Gibt es längere Bursts und sind mehr Spikes in einem Burst enthalten? Sind potenzielle Unterschiede statistisch signifikant? Wie kann man die Feuermuster objektiv und reproduzierbar klassifizieren? Oftmals werden visuelle Kriterien zur Klassifikation benutzt, die keine reproduzierbaren Ergebnisse liefern. Außerdem werden deskriptive Statistiken zum Quantifizieren der Spike Trains verwendet, ohne sie auf die Eigenschaften des darunterliegenden Prozesses zu beziehen.

Deswegen wird nach einer kurzen Einführung der wichtigsten Punktprozess Begriffe in *Kapitel 1.2*, welches hauptsächlich an dem Buch von Daley und Vere-Jones (1988) angelehnt ist, das Spike Train Modell in *Kapitel 2.1* vorgestellt. Das Modell wurde auf

Basis empirischer Beobachtungen in dopaminergen Neuronen und in Zusammenarbeit mit Neurophysiologen entwickelt und wird 'Gaussian Locking to a free Oscillator' (GLO) genannt. Es ist verwandt mit dem ELO (Schneider, 2008), welches für 'Exponential Locking to a free Oscillator' steht und zur Modellierung von Phasenverschiebungen bei parallel aufgenommenen Neuronen verwendet wird.

Das GLO ist doppelt stochastisch und benutzt als ersten Schritt einen oszillatorischen Hintergrundrhythmus \mathbb{B} der als stationäre Irrfahrt $(B_i)_{i \in \mathbb{Z}}$ mit einer Zuwachsverteilung $\mathcal{N}(\mu, \sigma_1^2)$ repräsentiert wird. Es werden zwei Teilmodelle eingeführt die nicht burstiges ($m = 0$) und burstiges ($m = 1$) Feuerverhalten beschreiben. Dabei ist m ein Modellparameter der konstant über die Zeit ist. In Abhängigkeit des Feuermodus m wird für jeden Punkt B_i eine zufällige Anzahl P_i von Spikes ausgewürfelt. Für beide Teilmodelle ist die mittlere Anzahl der Spikes pro Erzeugerpunkt B_i durch den Parameter γ gegeben, jedoch wird für $m = 0$ die Bernoulli-Verteilung und für $m = 1$ die Poisson-Verteilung verwendet. Im zweiten Schritt werden die zufälligen Spike Zeitpunkte um ihren Erzeugerpunkt B_i mit einer zentrierten Normalverteilung mit Varianz σ_2^2 verteilt. Diese so erzeugten Spike Zeitpunkte repräsentieren den beobachteten Punktprozess Φ (vgl. dazu Abbildung 0.3), welcher von den fünf leicht zu interpretierenden Parametern $\mu, \sigma_1, \sigma_2, \gamma, m$ abhängt. Wir schreiben $\Phi \sim \text{GLO}(\mu, \sigma_1, \sigma_2, \gamma, m)$.

Diese Parameter beschreiben die Regelmäßigkeit und das Burstverhalten der Feuermuster. Der Hintergrundrhythmus $(B_i)_{i \in \mathbb{Z}}$ hat eine oszillatorische Aktivität mit Periode μ und Varianz σ_1^2 . Der Parameter μ ist somit ein Skalenparameter und eine Änderung von σ_1 hat einen direkten Einfluss auf die Regelmäßigkeit des Prozesses. Der Feuermodus m besagt, ob der Prozess Bursts enthält ($m = 1$) oder nicht ($m = 0$). Der Parameter γ bezeichnet die erwartete Anzahl an Spikes an jedem Punkt B_i des Hintergrundrhythmus und ist somit mit dem Burstverhalten verbunden. Im Falle von $m = 1$ repräsentiert σ_2 die Burstbreite und für $m = 0$ wird es benutzt um die Korrelationen zwischen benachbarten Spike Wartezeiten zu beschreiben.

Um das GLO Modell besser zu verstehen, befasst sich *Kapitel 2.2* mit einigen mathematischen Eigenschaften. Der den Hintergrundrhythmus beschreibende Punktprozess $\mathbb{B} = \sum_{i \in \mathbb{Z}} \delta_{B_i}$ ist einfach (Lemma 2.2), stationär (Lemma 2.3) und ergodisch (Satz 2.5). Diese Eigenschaften übertragen sich auf den GLO Prozess Φ (Satz 2.16, Satz 2.17 und Korollar 2.23). Des Weiteren hat Φ eine mittlere Intensität von γ/μ (Satz 2.18) und repräsentiert einen Cluster Prozess (Lemma 2.21). Im Falle von $m = 1$ kann Φ als ein Cox Prozess mit zufälliger Intensität

$$\rho_{\mathbb{B}}(t) = \gamma \sum_{j \in \mathbb{Z}} \varphi_{\{B_j, \sigma_2^2\}}(t) \tag{0.3}$$

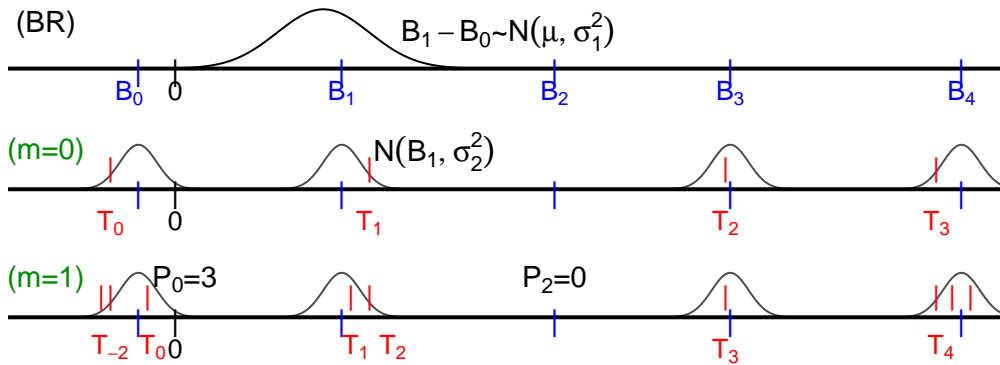


Abbildung 0.3: Konstruktion des GLO Punktprozesses Φ (rot).

und zufälligem Intensitätsmaß

$$\Lambda(A) = \int_A \rho_{\mathbb{B}}(t) dt, \quad A \in \mathcal{B} \quad (0.4)$$

aufgefasst werden (Satz 2.25), wobei $\varphi_{\{B_j, \sigma_2^2\}}$ die Dichte der Normalverteilung mit Mittelwert B_j und Varianz σ_2^2 darstellt und \mathcal{B} die Borel σ -algebra auf \mathbb{R} ist. Ausgehend von Satz 9.4.II. aus Daley und Vere-Jones (1988), welches auf einem Ergebnis von Stone (1968) und Dobrushin (1956) über die Konvergenz eines Punktprozesses nach wiederholter unabhängiger Translation der Punkte beruht, ergibt sich der folgende Satz.

Satz 0.1: Sei $\Phi \sim \text{GLO}(\mu, \sigma_1, \sigma_2, \gamma, m)$. Für $\sigma_2 \rightarrow \infty$ konvergiert Φ schwach gegen einen Poisson Prozess mit Rate γ/μ .

Die Interspike Intervalle (ISI) genannten Wartezeiten zwischen den einzelnen Spike Zeitpunkten werden in Kapitel 3 ausführlich untersucht. Die ISIs j -ter Ordnung ($j \in \mathbb{N}$) sind definiert als $W_k^{[j]} := T_{k+j} - T_k$, $k \in \mathbb{Z}$, wobei die reellwertigen Zufallsvariablen $\dots < T_{-1} < T_0 < 0 \leq T_1 < T_2 < \dots$ die zufälligen Spike Zeitpunkte des Punktprozesses Φ repräsentieren. Um die ISI Verteilung zu approximieren, werden 'beat-sortierte' Spike Zeitpunkte $\{\tilde{T}_k : k \in \mathbb{Z}\}$ eingeführt (siehe Definition 3.3). Diese sind nicht mehr anhand ihrer Reihenfolge auf der reellen Achse nummeriert, sondern anhand der Reihenfolge ihres Auftauchens im Hintergrundrhythmus $(B_i)_{i \in \mathbb{Z}}$. Die zugehörigen Intervalle sind gegeben als $\tilde{W}_k^{[j]} := \tilde{T}_{k+j} - \tilde{T}_k$ und können im Gegensatz zu den ISIs auch negative Werte annehmen. Im Falle von $\sigma_1 \ll \mu$ und $\sigma_2 \ll \mu$, lässt sich die Verteilung von W_i durch die von \tilde{W}_i approximieren und die dazugehörigen Dichten bestimmen (Satz 3.15 für den Fall $m = 0$ und Satz 3.20 für $m = 1$). Dabei wird im burstigen Fall auf Ordnungsstatistiken (David, 1970) zurückgegriffen.

Ursprünglich erhoffte man sich über die ISI Verteilung die Parameter des GLO Modells zu schätzen. Da dies aber im Allgemeinen nicht ohne Weiteres möglich ist, wird auf die in *Kapitel 4* vorgestellten Autokorrelogramme zurückgegriffen. Dazu gehört die autocorrelation function (ACF; Gerstein und Kiang, 1960) und das autocorrelation histogram (ACH; Moore et al., 1966; Perkel et al., 1967). Für die Klassifikation der Feuermuster wird üblicherweise die Form des ACH benutzt. Um dieses zu definieren, hat man eine Partition $L := \{L_1, \dots, L_K\}$ eines Analysebereichs $(l_b, l_e]$, $0 < l_b < l_e < \infty$, bestehend aus $K \in \mathbb{N}$ disjunkten benachbarten Intervallen $L_k := (a_k, a_k + \delta]$ mit Länge δ und die Menge der Intervallmittelwerte $\mathcal{L} := \{l_1, \dots, l_K\}$. Dann ist das ACH zum Wert $l_k \in \mathcal{L}$ definiert als

$$\tilde{h}(l_k) = \tilde{h}_L(l_k) := \# \{(t_i, t_j) \mid t_j - t_i \in L_k, i, j \in \{1, 2, \dots, n\}\}, \quad \forall l_k \in \mathcal{L} \quad (0.5)$$

und zählt somit die Anzahl der Spike Paare die eine Differenz haben, die im Intervall L_k liegt. Das empirische ACH wird normiert (Cox, 1965) um als Schätzer für die ACF zu dienen. Diese bedingt auf einen Spike zur Zeit 0 und bestimmt die Intensität der Spikes l Zeiteinheiten später:

$$f(l) := \lim_{\delta_1, \delta_2 \rightarrow 0^+} \frac{\mathbb{E}[\Phi(l, l + \delta_1) \mid \Phi(-\delta_2, 0] > 0]}{\delta_1}. \quad (0.6)$$

Da im GLO Fall die Differenzen zwischen den Spike Zeitpunkten T_j und ihren Erzeugerpunkten B_i normalverteilt sind, genauso wie die Intervalle zwischen den Erzeugerpunkten, lässt sich das ACF leicht bestimmen.

Satz 0.2: Sei $\Psi \sim \text{GLO}(\psi)$, dann ist die ACF für $l > 0$ gegeben als

$$f_m(l) = \gamma \sum_{j \in K(m)} \varphi_{\{j\mu, |j|\sigma_1^2 + 2\sigma_2^2\}}(l), \quad \text{wobei} \quad K(m) := \begin{cases} \mathbb{Z}^* & , m = 0 \\ \mathbb{Z} & , m = 1. \end{cases} \quad (0.7)$$

Der Aufbau der ACF für $m = 0$ ist in Abbildung 0.4 A angedeutet.

Experimentelle Klassifikation benutzt visuelle Inspektion des ACH oder versucht plausible Funktionen daran anzupassen (Engel et al., 1992; König, 1994; Celada et al., 1999; Hyland et al., 2002; Paladini et al., 2003; Schneider und Nikolic, 2006). Anhand der GLO ACF können aber objektive Kriterien für die Klassifikation angegeben werden, die inspiriert von der visuellen Klassifikation (Wilson et al., 1977; Paladini et al., 2003) und reproduzierbar sind (vgl. dazu *Kapitel 4.3*). Zu diesen Kriterien gehören die Parameterkombinationen μ/σ_2 und $\mu/\sqrt{\sigma_1^2 + 2\sigma_2^2}$, welche die ersten zwei lokalen Maxima der ACF in Relation zur durchschnittlichen Intensität γ/μ darstellen (vgl. Abbildung 0.4 B).

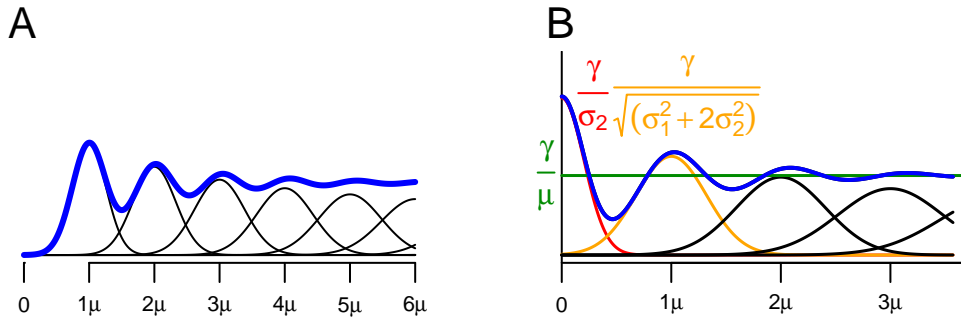


Abbildung 0.4: Zusammensetzung der ACF (A) sowie Parameterkombinationen für die Klassifikation der Feuermuster (B).

Kapitel 5 beschreibt die Parameterschätzung des GLO Modells. Die Parameter werden mittels Anpassung der ACF f_m an das empirische normierte ACH h geschätzt. Dabei werden die gewichteten quadratischen Abweichungen

$$\sum_{l \in \mathcal{L}} v_l (h(l) - f_m(l))^2 \quad (0.8)$$

mit einem nichtlinearen Kleinste Quadrate Algorithmus (Gauss-Newton) minimiert. Die Wahl des Analysefensters \mathcal{L} , der Gewichte v_l sowie der Startwerte der Prozedur ist von Bedeutung und wird im *Kapitel 5.1* ausführlich beschrieben. Diese Vorgehensweise liefert Schätzer $\hat{\mu}, \hat{\sigma}_1, \hat{\sigma}_2, \hat{\gamma}, \hat{m}$.

Die Genauigkeit der Prozedur wird in *Kapitel 5.2* für unterschiedliche Szenarien untersucht und es stellt sich heraus, dass die Güte der Parameterschätzung von verschiedenen Faktoren abhängt. Dazu gehört die Anzahl der Spikes n im Beobachtungsintervall $[0, \mathcal{T}]$, die Parameter des GLO Modells sowie die Bin Größe δ , welche die Auflösung des Analyse Fensters \mathcal{L} vorgibt. Je mehr Spikes (oder je länger das Intervall $[0, \mathcal{T}]$) desto besser die Schätzung. Dabei hängt die Wahl der Bin Größe δ auch von n ab. Je größer n desto kleiner kann δ gewählt werden. Des Weiteren lassen sich die Parameter von GLO Prozessen mit oszillatorischem Verhalten ($\sigma_1 \ll \mu$ und $\sigma_2 \ll \mu$) besser schätzen als solche die diese Eigenschaft nicht aufweisen.

Kapitel 6 beschäftigt sich mit der Schätzung der Variabilität der Parameterschätzer. Dazu werden verschiedene Bootstrap Verfahren (Efron and Tibshirani, 1994) zum Resampeln der Spike Trains vorgestellt und mittels Monte-Carlo Simulation untersucht. Das Kapitel endet mit einer Betrachtung verschiedener Bootstrap Konfidenzintervalle.

In *Kapitel 7* wird das GLO Modell auf den Datensatz angewendet anhand dessen es entwickelt wurde. Der Datensatz besteht aus 146 Spike Trains von dopaminergen Neuronen

aus der Substantia Nigra von Mäusen. Es stellt sich heraus, dass die Klassifikation der Feuermuster durch das GLO in ca. 75% der Fälle mit der Klassifikation durch visuelle Kriterien übereinstimmt und nur in ca. 2% der Fälle gegensätzlich war. Das GLO Modell scheint also gut verträglich mit den visuellen Verfahren. Darüber hinaus bietet es die Möglichkeit die Variabilität der Klassifikation zu visualisieren. Das Quantifizieren und Vergleichen der Feuermuster auf der Basis des GLO Modells wird anhand eines kleinen Teildatensatzes exemplarisch durchgeführt. Die Daten wurden aus 6 Mäusen erhoben, deren Gehirnaktivität vor und nach der Injektion eines Dopamin Rezeptor Blockers gemessen wurde. Nach der Injektion gab es signifikante Unterschiede in den Feuermustern der Zelle. Dies spiegelte sich insbesondere in der Erhöhung der Feuerrate, sowie in einer größeren Anzahl von Spikes pro Burst und einer größeren Burstbreite wider.

Im *Kapitel 8* endet die Arbeit mit einer kurzen Zusammenfassung und einem Fazit. Das GLO gibt dem Anwender objektive und reproduzierbare Klassifikationskriterien und unterstützt ihn beim Quantifizieren von experimentell gewonnen Spike Trains.

Im Anhang werden zusätzliche Informationen und Literatur zum neurophysiologischen Hintergrund, eine kurze Beschreibung des Datensatzes sowie der verwendete R Code zur Verfügung gestellt.

Veröffentlichte Inhalte: Einige Abschnitte und Ergebnisse dieser Arbeit sind im Artikel von Bingmer et al. (2011) veröffentlicht. Die Daten sind bei Schieman (2012) ausführlich beschrieben und zusätzlich unter Verwendung des GLO in Schieman et al. (2012) veröffentlicht.

List of Figures

| | | |
|------|--|----|
| 1.1. | Rasterplots of prominent empirical firing patterns | 2 |
| 1.2. | Distribution of waiting times between events in empirical spike trains . . . | 6 |
| 1.3. | Autocorrelation histograms of the four empirical spike trains | 6 |
| 1.4. | Point process on the real line | 12 |
| 1.5. | A random walk regarded as a point process | 17 |
| 1.6. | Construction of a stationary random walk | 18 |
| 2.1. | Construction of the GLO spike train model | 22 |
| 2.2. | Interbeat interval (IBI) distribution | 23 |
| 2.3. | Visualization of an inversion of beats | 28 |
| 2.4. | Visualization of the GLO simulation procedure | 39 |
| 2.5. | Running time analysis | 43 |
| 2.6. | Simulated distribution of B_1 for $\mu = 1$ and different values of σ_1 | 45 |
| 2.7. | Rasterplots of simulated GLO spike trains | 45 |
| 3.1. | ISIs of first and second order | 48 |
| 3.2. | Classical spike times and beat-ordered spike times | 50 |
| 3.3. | Inversion probability for different parameter constellations | 51 |
| 3.4. | Schematic computation of a beat-ordered waiting time distribution | 53 |
| 3.5. | Simulated ISI distributions in the non-bursty case | 55 |
| 3.6. | In-beat and between-beat intervals in a bursty GLO spike train | 57 |
| 3.7. | Simulated ISI distributions in the bursty case | 60 |
| 3.8. | Simulated ISI distribution in the irregular case | 64 |
| 4.1. | Construction of the GLO autocorrelation function (ACF) | 67 |
| 4.2. | Properties of the autocorrelation histogram (ACH) | 72 |
| 4.3. | Visualization of regularity and burstiness in the ACF | 75 |
| 4.4. | Illustration of β_1 and β_2 in combination with ACFs and rasterplots | 77 |
| 5.1. | The choice of the analysis window and its consequences | 83 |
| 5.2. | Fits resulting of different starting values | 85 |
| 5.3. | Starting value for μ in the non-bursty case | 87 |
| 5.4. | Starting value for μ in the bursty case | 88 |

List of Figures

| | |
|---|-----|
| 5.5. Rasterplots, ACH and ACFs of exemplary parameter combinations | 90 |
| 5.6. Comparison of estimated standard errors for different \mathcal{T} and δ | 94 |
| 5.7. Comparison of estimated NLS standard errors | 97 |
| 5.8. Precision of parameter estimates | 98 |
| 5.9. Four diagnostic plots are shown for the four exemplary spike trains | 100 |
| 6.1. Bootstrap diagram | 104 |
| 6.2. Comparison of a fit between empirical and simulated spike train | 106 |
| 6.3. Layout for the bootstrap simulations | 108 |
| 6.4. Number of blocks for block resampling | 110 |
| 6.5. Comparison of bootstrap methods | 113 |
| 6.6. Comparison of different bootstrap confidence intervals | 118 |
| 7.1. Overview of the GLO ACH fits in the sample data set | 122 |
| 7.2. Thresholds for GLO classification criteria | 124 |
| 7.3. Variability of the GLO classification criteria | 126 |
| 7.4. Changes in the parameter estimates of six substantia nigra neurons | 128 |
| A.1. Empirical firing rates for different brain regions | 135 |

List of Tables

| | |
|--|-----|
| 2.1. Inversion probability of beats for different ratios μ/σ_1 | 30 |
| 3.1. The estimated coefficient of variation \hat{c}_v for the empirical spike trains . . | 52 |
| 4.1. Summary of classification rules for burstiness and regularity | 76 |
| 5.1. Exemplary parameter constellations used for simulation studies | 90 |
| 5.2. Settings to investigate the impact of δ and \mathcal{T} on the fitting procedure . . | 93 |
| 5.3. Relative frequency \tilde{m}_f of wrong estimations of m | 95 |
| 7.1. Classification of the sample data set by GLO and VI criteria | 124 |
| 7.2. Distribution of classified KO and WT spike trains (SN spon.) | 126 |
| 7.3. Distribution of classified DN and WT spike trains (SN viral) | 126 |
| 7.4. p -values of tests between pre and post measurements | 128 |
| A.1. Summary of the empirical data set | 135 |

List of R codes

| | | | |
|-------|-------------------------|---|-----|
| A.1. | <code>SIM</code> | (Simulation of a GLO spike train) | 136 |
| A.2. | <code>GLO</code> | (Efficient simulation of a GLO spike train) | 137 |
| A.3. | <code>ACF</code> | (Autocorrelation function of the GLO process) | 137 |
| A.4. | <code>ACH</code> | (Computation of the ACH for an arbitrary spike train) . . . | 138 |
| A.5. | <code>Awindow</code> | (Computation of the analysis window) | 138 |
| A.6. | <code>Weights</code> | (Weights for weighted nonlinear regression) | 139 |
| A.7. | <code>spektral</code> | (Possible starting value for μ) | 139 |
| A.8. | <code>GA1</code> | (Restriction for γ in the non-bursty case) | 139 |
| A.9. | <code>Svalues</code> | (Starting values for weighted nonlinear regression) | 140 |
| A.10. | <code>estimation</code> | (GLO estimation procedure) | 141 |
| A.11. | <code>boot1</code> | (Parametric bootstrap and GLO estimation) | 141 |
| A.12. | <code>boot2</code> | (Moving blocks bootstrap and GLO estimation) | 142 |
| A.13. | <code>mark</code> | (Construction of a marked point process) | 142 |
| A.14. | <code>boot3</code> | (Marked point process bootstrap and GLO estimation) . . . | 143 |

List of Notations and Abbreviations

| | | |
|---|---|-----|
| \emptyset | empty set | 7 |
| $\mathcal{A}, \mathcal{A}_1, \mathcal{A}_2$ | σ -algebra | 7 |
| $\overline{\text{AE}}(\mu)$ | mean absolute error of the estimator $\hat{\mu}$ | 91 |
| $\text{AE}(\hat{x})$ | absolute error of an estimator \hat{x} | 91 |
| ACF | autocorrelation function | 66 |
| ACH | autocorrelation histogram | 70 |
| AE | absolute error | 91 |
| β_1 | irregularity in the GLO model | 74 |
| β_2 | burst identifiability in the GLO model | 75 |
| $(B_i)_{i \in \mathbb{Z}}$ | stationary random walk modeling the BR \mathbb{B} | 20 |
| \mathbb{B} | stationary point process representing the BR | 20 |
| \mathbb{B}_P | copied and thinned points of \mathbb{B} | 31 |
| $\mathcal{B}(M_c)$ | Borel σ -algebra of counting measures | 8 |
| $\mathcal{B}(E)$ | Borel σ -algebra on E | 7 |
| $\mathcal{B}, \mathcal{B}(\mathbb{R})$ | Borel subsets of \mathbb{R} | 9 |
| $B_0, B_{\pm 1}, B_{\pm 2}, \dots$ | points of the stationary background rhythm \mathbb{B} | 20 |
| Bernoulli(γ) | Bernoulli distribution with parameter γ | 20 |
| BR | background rhythm | 20 |
| $C_\alpha(\mathbf{x})$ | confidence interval constructed of \mathbf{x} | 114 |

List of Notations and Abbreviations

| | | |
|--------------------------------------|--|-----|
| c.d.f. | cumulative distribution function..... | 20 |
| c.s.m.s | complete separable metric space..... | 7 |
| CI | confidence interval..... | 114 |
| CV | coefficient of variation..... | 51 |
| δ | bin size for the ACH..... | 70 |
| δ_x | Dirac measure..... | 7 |
| DA | dopamine..... | 4 |
| E | complete separable metric space (c.s.m.s.)..... | 7 |
| ELO | the 'Exponential Locking to a free Oscillator' model..... | 4 |
| $f_{(k:n)}$ | density of k -th order statistic..... | 57 |
| $f_m(l)$ | ACF of the GLO model..... | 66 |
| $f_{(r,s:n)}$ | joint density of r -th and s -th order statistic..... | 57 |
| $f_{\tilde{W}_{(k:n)}^{ib}}$ | density of k -th order in-beat interval..... | 58 |
| $f_{\tilde{W}_{(n_1, n_2)[j]}^{bb}}$ | conditioned density of between-beat interval..... | 59 |
| $F_{\{\mu, \sigma^2\}}(x)$ | cumulative distribution function of $\mathcal{N}(\mu, \sigma^2)$ | 20 |
| γ | mean number of spikes per beat in the GLO model..... | 21 |
| $G[f], G_{\Phi}[f]$ | probability generating functional (p.g.fl.) of Φ | 9 |
| GLO | the 'Gaussian Locking to a free Oscillator' model..... | 19 |
| $\tilde{h}(l)$ | ACH..... | 71 |
| $\tilde{\mathbf{h}}$ | vector $(\tilde{h}(l_1), \dots, \tilde{h}(l_K))'$ | 109 |
| $\tilde{\mathbf{h}}_i$ | vector $(\tilde{h}_i(l_1), \dots, \tilde{h}_i(l_K))'$ | 109 |
| $h(l)$ | standardized ACH..... | 71 |
| H_{∞} | distribution of asymptotic residual waiting time R_{∞} | 15 |
| $h_i(l)$ | contribution of spike i to the ACH..... | 111 |

List of Notations and Abbreviations

| | | |
|--|--|-----|
| i.i.d. | independent identically distributed | 14 |
| IBI | interbeat interval | 23 |
| ISI | interspike interval | 47 |
| KO | knockout | 134 |
| λ | intensity of a stationary point process | 13 |
| Λ | intensity measure of a Poisson point process | 14 |
| $\mathcal{L}, \mathcal{L}_{l_b, l_e}^\delta$ | set of lags for the ACH | 71 |
| l_b | beginning lag of analysis window | 70 |
| l_e | end lag of analysis window | 70 |
| M_E | space of boundedly finite Borel measures on E | 7 |
| μ | mean increment of the random walk $(B_i)_{i \in \mathbb{Z}}$ | 20 |
| $\text{MAD}(x_1, \dots, x_n)$ | median absolute deviation of x_1, \dots, x_n | 92 |
| $\text{MSE}(\hat{x})$ | mean squared error of an estimator \hat{x} | 91 |
| \tilde{m}_f | relative frequency of wrong m classifications | 91 |
| \underline{m} | mean density of a stationary point process | 13 |
| m | firing mode of the GLO process | 20 |
| M_c | set of counting measures defined on Borel subsets $\mathcal{B}(E)$ | 7 |
| M_s | set of simple counting measures | 7 |
| MAD | median absolute deviation | 92 |
| MAE | mean absolute error | 91 |
| MRE | mean absolute relative error | 91 |
| MSE | mean squared error | 91 |
| \mathbb{N} | set of natural numbers | 7 |
| \mathbb{N}_0 | $\mathbb{N} \cup \{0\}$ | 7 |

List of Notations and Abbreviations

| | | |
|------------------------------------|--|-----|
| $\mathcal{N}(\mu, \sigma^2)$ | Normal distribution with mean μ and variance σ^2 | 19 |
| $\text{NB}(j, \gamma)$ | negative binomial distribution | 54 |
| $N(t)$ | nondecreasing integer-valued step function | 11 |
| ORE | overall relative error | 91 |
| \mathbb{P}_Φ | distribution of a point process Φ | 8 |
| $\Phi, \Phi_1, \Phi_2, \dots$ | point processes | 8 |
| ψ | vector of GLO parameters $(\mu, \sigma_1, \sigma_2, \gamma, m)$ | 21 |
| ψ^* | bootstrap estimate of GLO parameter vector ψ | 105 |
| $P_0, P_{\pm 1}, P_{\pm 2}, \dots$ | random number of spikes at beats $B_0, B_{\pm 1}, B_{\pm 2}, \dots$ | 20 |
| p.g.fl. | probability generating functional | 9 |
| $\text{Pois}(\gamma)$ | Poisson distribution with parameter γ | 21 |
| \mathbb{R} | set of real numbers | 7 |
| \mathbb{R}_+ | $[0, \infty)$ | 7 |
| \mathbb{R}_∞ | $\mathbb{R} \cup \{-\infty, \infty\}$ | 7 |
| $\overline{\text{RE}}(\mu)$ | mean absolute relative error of the estimator $\hat{\mu}$ | 91 |
| $\text{RE}(\hat{x})$ | absolute relative error of an estimator \hat{x} | 91 |
| R_∞ | asymptotic residual waiting time | 15 |
| R_a | excess of random walk $(S_n)_{n \in \mathbb{N}_0}$ over a boundary a | 15 |
| RE | absolute relative error | 91 |
| $\varphi_{\{\mu, \sigma^2\}}(x)$ | density function of $\mathcal{N}(\mu, \sigma^2)$ | 20 |
| $(S'_i)_{i \in \mathbb{Z}}$ | stationary random walk | 17 |
| $(S_n)_{n \in \mathbb{N}_0}$ | random walk or renewal process | 14 |
| $s\hat{e}_b(\hat{\theta})$ | bootstrap estimate of standard error | 105 |
| \mathcal{S} | spike train | 9 |

List of Notations and Abbreviations

| | | |
|--|--|-----|
| \mathcal{S}^* | bootstrapped spike train | 105 |
| σ_2 | spike variability in the GLO model | 21 |
| σ_1 | variance of the increments of the random walk \mathbb{B} | 20 |
| φ | counting measure | 7 |
| φ^* | support counting measure of φ | 7 |
| $S_0, S_1, \dots,$ | partial sums of the random walk $(S_n)_{n \in \mathbb{N}_0}$ | 14 |
| S_φ | support of measure φ | 7 |
| $\hat{\theta}^*$ | bootstrap estimate of parameter θ | 104 |
| \mathcal{T} | recording time | 9 |
| \mathcal{T} | measure-preserving operator | 36 |
| τ_a | time at which random walk first reaches a , or ∞ | 15 |
| θ | parameter | 104 |
| $\tilde{T}_0, \tilde{T}_{\pm 1}, \tilde{T}_{\pm 2}, \dots$ | beat-ordered spike times | 48 |
| $T_0, T_{\pm 1}, T_{\pm 2}, \dots$ | random occurrence times of a point process | 10 |
| $t_0, t_{\pm 1}, t_{\pm 2}, \dots$ | realization of occurrence times of a point process | 10 |
| t_1, t_2, \dots, t_n | empirical spike times of a spike train | 9 |
| U | renewal measure | 14 |
| $\mathcal{V}(E)$ | class of all \mathbb{R} -valued Borel functions defined on E | 9 |
| $\tilde{W}_0, \tilde{W}_{\pm 1}, \tilde{W}_{\pm 2}, \dots$ | beat-ordered ISIs of first order | 49 |
| $\tilde{W}_0^{[j]}, \tilde{W}_{\pm 1}^{[j]}, \tilde{W}_{\pm 2}^{[j]}, \dots$ | ISIs of j th-order of beat-ordered spike times | 49 |
| $W_0, W_{\pm 1}, W_{\pm 2}, \dots$ | waiting times between spikes | 11 |
| $W_0^{[j]}, W_{\pm 1}^{[j]}, W_{\pm 2}^{[j]}, \dots$ | interspike intervals of j th-order | 48 |
| $W_{(k:n)}^{ib}$ | k -th order in-beat interval | 58 |
| $\tilde{W}_{(n_1, n_2)[j]}^{bb}$ | between-beat interval | 59 |

List of Notations and Abbreviations

| | | |
|---------------------------|---|-----|
| WT | wildtype..... | 134 |
| X, X_1, X_2, \dots | random variables..... | 14 |
| \mathbb{Z} | set of integers..... | 7 |
| \mathbb{Z}^* | $\mathbb{Z} \setminus \{0\}$ | 7 |
| $Z_{(k:n)}$ | k -th order statistic of Z_1, Z_2, \dots, Z_n | 57 |
| $Z_{i,1}, Z_{i,2}, \dots$ | spike variations in the GLO model..... | 21 |
| $[n]$ | set of numbers $1, 2, \dots, n$ | 7 |
| $\#A$ | number of elements of the set A | 10 |
| \sim | distributed as..... | 19 |
| $\stackrel{d}{=}$ | equality in distribution..... | 9 |

Bibliography

- Abeles, M. Quantification, smoothing, and confidence limits for single-units' histograms. *Journal of Neuroscience Methods*, 5(4):317–325, 1982.
- Awiszus, F. Spike train analysis. *Journal of Neuroscience Methods*, 74(2):155–166, 1997.
- Bair, W., Koch, C., Newsome, W., and Britten, K. Power Spectrum Analysis of Bursting Cells in Area MT in the Behaving Monkey. *The Journal of Neuroscience*, 14(5):2870–2892, 1994.
- Barbieri, R., Quirk, M. C., Frank, L. M., Wilson, M. A., and Brown, E. N. Construction and analysis of non-poisson stimulus-response models of neural spiking activity. *Journal of Neuroscience Methods*, 105(1):25–37, 2001.
- Bates, D. M. and Watts, D. G. *Nonlinear Regression Analysis and Its Applications*. John Wiley and Sons, New York, 1988.
- Bauer, H. *Maß- und Integrationstheorie*. De Gruyter-Verlag, Berlin/New York, 1990.
- Berger, D., Pribram, K., Wild, H., and Bridges, C. An analysis of neural spike-train distributions: determinants of the response of visual cortex neurons to changes in orientation and spatial frequency. *Experimental Brain Research*, 80(1):129–134, 1990.
- Bingmer, M., Schiemann, J., Roeper, J., and Schneider, G. Measuring burstiness and regularity in oscillatory spike trains. *Journal of Neuroscience Methods*, 201(2):426–437, 2011.
- Braun, W. J. and Kulperger, R. J. A bootstrap for point processes. *Journal of Statistical Computing and Simulation*, 60:129–155, 1998.
- Bremaud, P. *Point Processes and Queues*. Springer, New York/Heidelberg/Berlin, 1981.

- Brockwell, P. and Davis, R. *Introduction to Time Series and Forecasting*. Springer, New York, 2010.
- Bromberg-Martin, E. S., Matsumoto, M., and Hikosaka, O. Dopamine in Motivational Control: Rewarding, Aversive, and Alerting. *Neuron*, 68(5):815–834, 2010.
- Brown, E. N., Kass, R. E., and Mitra, P. P. Multiple neural spike train data analysis: state-of-the-art and future challenges. *Nature Neuroscience*, 7(5):456–461, 2004.
- Buracas, G. T., Zador, A. M., DeWeese, M. R., and Albright, T. D. Efficient discrimination of temporal patterns by motion-sensitive neurons in primate visual cortex. *Neuron*, 20(5):959–969, 1998.
- Camproux, A. C., Saunier, F., Chouvet, G., Thalabard, J. C., and Thomas, G. A Hidden Markov Model Approach to Neuron Firing Patterns. *Biophysical Journal*, 71(5):2404–2412, 1996.
- Celada, P., Paladini, C. A., and Tepper, J. M. Gabaergic control of rat substantia nigra dopaminergic neurons: role of globus pallidus and substantia nigra pars reticulata. *Neuroscience*, 89(3):813–825, 1999.
- Chambers, J. M. *Software for Data Analysis: Programming with R*. Springer, New York, 2008.
- Chang, J. T. and Peres, Y. Ladder heights, gaussian random walks and the riemann zeta function. *The Annals of Probability*, 25(2):787–802, 1997.
- Chiappalone, M., Novellino, A., Vajda, I., Vato, A., Martinoia, S., and Vanpelt, J. Burst detection algorithms for the analysis of spatio-temporal patterns in cortical networks of neurons. *Neurocomputing*, 65–66:653–662, 2005.
- Chu, E. *Discrete and Continuous Fourier Transforms*. CRC Press, London, New York, 2008.
- Cocatre-Zilgien, J. H. and Delcomyn, F. Identification of bursts in spike trains. *Journal of Neuroscience Methods*, 41(1):19–30, 1992.
- Cox, D. R. *Renewal Theory*. Methuen’s monographs on applied probability and statistics. Methuen & CO LTD, London, 1962.
- Cox, D. R. On the Estimation of the Intensity Function of a Stationary Point Process.

Bibliography

- Journal of the Royal Statistical Society. Series B (Methodological)*, 27(2):332–337, 1965.
- Cox, D. R. and Isham, V. *Point Processes*. CRC Monographs on Statistics & Applied Probability. Chapman & Hall/CRC, 1980.
- Daley, D. J. and Oakes, D. Random walk point processes. *Probability Theory and Related Fields*, 30(1):1–16, 1974.
- Daley, D. J. and Vere-Jones, D. *An Introduction to the Theory of Point Processes*. Springer Berlin, 1988.
- David, H. A. *Order Statistics*. Wiley Series in Probability and Statistics. Wiley-Interscience, 1970.
- Davies, R. M., Gerstein, G. L., and Baker, S. N. Measurement of Time-Dependent Changes in the Irregularity of Neural Spiking. *Journal of Neurophysiology*, 96(2): 906–918, 2006.
- Davison, A. C. and Hinkley, D. V. *Bootstrap Methods and their Applications*. Cambridge University Press, 2009.
- DiCiccio, T. J. and Efron, B. Bootstrap confidence intervals. *Statistical Science*, 11(3): 189–212, 1996.
- Diggle, P. J. *Statistical Analysis of Spatial Point Patterns*. Academic Press, London, 1983.
- Dobrushin, R. L. On the poisson law of distribution of particles in a space. *Ukrainskii Matematicheskii Zhurnal*, 8:127–138, 1956.
- Edeline, J. M., Manunta, Y., and Hennevin, E. Auditory thalamus neurons during sleep: changes in frequency selectivity, threshold and receptive field size. *Journal of Neurophysiology*, 84(2):934–952, 2000.
- Efron, B. Second thoughts on the bootstrap. *Statistical Science*, 18(2):135–140, 2003.
- Efron, B. and Tibshirani, R. J. *An Introduction to the Bootstrap*. Chapman & Hall/CRC, 1994.

- Ekholm, A. and Hyvärinen, J. A pseudo-markov model for series of neuronal spike events. *Biophysical Journal*, 10(8):773–796, 1970.
- Engel, A. K., König, P., Kreiter, A. K., Schillen, T. B., and Singer, W. Temporal coding in the visual cortex: New vistas on integration in the nervous system. *Trends in Neurosciences*, 15(6):218–226, 1992.
- Engel, A. K., König, P., and Singer, W. Direct Physiological Evidence for Scene Segmentation by Temporal Coding. *Proceedings of the National Academy of Sciences of the United States of America*, 88(20):9136–9140, 1991.
- Feller, W. *An Introduction to Probability Theory and Its Applications, Volume II*. John Wiley & Sons, Inc, 1966.
- Franken, P., König, D., Arndt, U., and Schmidt, V. *Queues and Point Processes*. John Wiley & Sons, Chichester, 1981.
- Gabbiani, F. and Koch, C. *Methods in Neuronal Modeling: From Synapses to Networks*, chapter Principles of spike train analysis, pages 313–360. MIT Press, Cambridge, MA, 1998.
- Gamerman, D. *Markov Chain Monte Carlo*. Chapman & Hall, London, New York, 1997.
- Geisler, C., Brunel, N., and Wang, X. J. Contributions of intrinsic membrane dynamics to fast network oscillations with irregular neuronal discharges. *Journal of Neurophysiology*, 94(6):4344–4361, 2005.
- Gerstein, G. L. and Kiang, N. Y. An approach to the quantitative analysis of electrophysiological data from single neurons. *Biophysical Journal*, 1:15–28, 1960.
- Gourevitch, B. and Eggermont, J. J. A nonparametric approach for detection of bursts in spike trains. *Journal of Neuroscience Methods*, 160(2):349–358, 2007.
- Grace, A. A. and Bunney, B. S. The control of firing pattern in nigral dopamine neurons: burst firing. *The Journal of Neuroscience*, 4(11):2877–2890, 1984.
- Gray, C. M., Engel, A. K., König, P., and Singer, W. Synchronization of oscillatory neuronal responses in cat striate cortex: Temporal properties. *Visual Neuroscience*, 8(4):337–347, 1992.
- Gray, C. M. and Singer, W. Stimulus-specific neuronal oscillations in orientation columns

- of cat visual cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 86(5):1698–1702, 1989.
- Grigoryan, G., Minasyan, S. M., and Saakyan, S. G. Background activity of neurons of the supraoptic hypothalamic nucleus in rats under conditions of vibrational stimulation and electromagnetic extrahigh-frequency irradiation. *Neurophysiology*, 39(6):433–442, 2007.
- Guido, W. and Weyand, T. Burst responses in thalamic relay cells of the awake behaving cat. *Journal of Neurophysiology*, 74(4):1782–1786, 1995.
- Gut, A. *Stopped Random Walks*. Springer Series in operations research and financial engineering. Springer, 2009.
- Harris, K. D., Hirase, H., Leinekugel, X., Henze, D. A., and Buzsáki, G. Temporal interaction between single spikes and complex spike bursts in hippocampal pyramidal cells. *Neuron*, 32(1):141–149, 2001.
- Havenith, M. N., Zemmar, A., Yu, S., Baudrexel, S. M., Singer, W., and Nikolic, D. Measuring sub-millisecond delays in spiking activity with millisecond time-bins. *Neuroscience Letters*, 450(6):296–300, 2009.
- Hoaglin, D. C., Mosteller, F., and Tukey, J. W. *Understanding Robust and Exploratory Data Analysis*. John Wiley & Sons, New York, 1983.
- Holt, G. R., Softky, W. R., Koch, C., and Douglas, R. J. Comparison of discharge variability in vitro and in vivo in cat visual cortex neurons. *Journal of Neurophysiology*, 75(5):1806–18014, 1996.
- Hyland, B. I., Reynolds, J. N., Hay, J., Perk, C. G., and Miller, R. Firing modes of midbrain dopamine cells in the freely moving rat. *Neuroscience*, 114(2):475–492, 2002.
- Iyengar, S. and Liao, Q. Modeling neural activity using the generalized inverse gaussian distribution. *Biological Cybernetics*, 77(4):289–295, 1997.
- Izhikevich, E. M., Desai, N. S., Walcott, E. C., and Hoppensteadt, F. C. Bursts as a unit of neural information: selective communication via resonance. *Trends in Neurosciences*, 26(3):161–167, 2003.
- Janssen, A. J. E. M. and van Leeuwaarden, J. S. H. On lerch’s transcendent and the gaussian random walk. *The Annals of Applied Probability*, 17(2):421–439, 2007.

- Johnson, D. H. Point process models of single-neuron discharges. *Journal of Computational Neuroscience*, 3(4):275–299, 1996.
- Kabacoff, R. *R in Action*. Manning, 2010.
- Kallenberg, O. Stability of critical cluster fields. *Mathematische Nachrichten*, 77(1):7–43, 1977.
- Kallenberg, O. *Random Measures*. Academic Press, London, 1986.
- Kaneoke, Y. and Vitek, J. L. Burst and oscillation as disparate neuronal properties. *Journal of Neuroscience Methods*, 68(2):211–223, 1996.
- Karr, A. F. *Point processes and their statistical inference*. Probability: Pure and applied. Marcel Dekker, Inc., 1986.
- Kass, R. E., Ventura, V., and Brown, E. N. Statistical issues in the analysis of neuronal data. *Journal of Neurophysiology*, 94(1):8–25, 2005.
- Kepecs, A. and Lisman, J. How to read a burst duration code. *Neurocomputing*, 58–60:1–6, 2004.
- Kingman, J. F. C. *Poisson Processes*. Oxford University Press, USA, 1993.
- Kingman, J. F. C. and Taylor, S. J. *Introduction to Measure and Probability*. Cambridge University Press, Cambridge, 1966.
- König, D. and Schmidt, V. *Zufällige Punktprozesse*. Teubner, Stuttgart, 1992.
- König, P. A method for the quantification of synchrony and oscillatory properties of neuronal activity. *Journal of Neuroscience Methods*, 54(1):31–37, 1994.
- König, P., Engel, A. K., and Singer, W. Relation between oscillatory activity and long-range synchronization in cat visual cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 92(1):290–294, 1995.
- Künsch, H. R. The jackknife and the bootstrap for general stationary observations. *The Annals of Statistics*, 17:1217–1241, 1989.
- Legendy, C. R. and Salzman, M. Bursts and Recurrences of Bursts in the Spike Trains

- of Spontaneously Active Striate Cortex Neurons. *Journal of Neurophysiology*, 53(4): 926–939, 1985.
- Lenz, F. A., Garonzik, I. M., Zirh, T. A., and Dougherty, P. M. Neuronal activity in the region of the thalamic principal sensory nucleus (ventralis caudalis) in patients with pain following amputations. *Neuroscience*, 86(4):1065–1081, 1998.
- Liemant, A., Matthes, K., and Wakolbinger, A. *Equilibrium distributions of branching processes*, volume 34 of *Mathematics and its Applications (East European Series)*. Kluwer Academic Publishers Group, Dordrecht, 1988.
- Lisman, J. E. Bursts as a unit of neural information: making unreliable synapses reliable. *Trends in Neurosciences*, 20(1):38–43, 1997.
- Livingstone, M. S. and Hubel, D. H. Effects of sleep and arousal on the processing of visual information in the cat. *Nature*, 291(5816):554–561, 1981.
- Loh, J. M. Bootstrapping an inhomogeneous point process. *Journal of Statistical Planning and Inference*, 140(3):734–749, 2010.
- Magnin, M., Morel, A., and Jeanmmod, D. Single-unit analysis of the pallidum, thalamus and subthalamic nucleus in parkinsonian patients. *Neuroscience*, 96(3):549–564, 2000.
- Maimon, G. and Assad, J. A. Beyond Poisson: Increased Spike-Time Regularity across Primate Parietal Cortex. *Neuron*, 62(3):426–440, 2009.
- Matthes, K., Kerstan, J., and Mecke, J. *Infinitely Divisible Point Processes*. John Wiley & Sons, New York, 1978.
- Møller, J. and Torrisi, G. L. Generalised shot noise Cox processes. *Advances in Applied Probability*, 37(1):48–74, 2005.
- Moore, G. P., Perkel, D. H., and Segundo, J. P. Statistical analysis and functional interpretation of neuronal spike data. *Annual Review of Physiology*, 28:493–522, 1966.
- Nawrot, M. P., Boucsein, C., Rodriguez Molina, V., Riehle, A., Aertsen, A., and Rotter, S. Measurement of variability dynamics in cortical spike trains. *Journal of Neuroscience Methods*, 169(2):374–390, 2008.
- Neyman, J. and Scott, E. L. *Stochastic Point Processes*, chapter Processes of clustering and applications, pages 646–681. Wiley, 1972.

- Paladini, C. A., Robinson, S., Morikawa, H., Williams, J. T., and Palmiter, R. D. Dopamine controls the firing pattern of dopamine neurons via a network feedback mechanism. *Proceedings of the National Academy of Sciences of the United States of America*, 100(5):2866–2871, 2003.
- Perkel, D. H., Gerstein, G. L., and Moore, G. P. Neuronal Spike Trains and Stochastic Point Processes I: The Single Spike Train. *Biophysical Journal*, 7(4):391–418, 1967.
- Radhakrishnan, V., Tsoukatos, J., Davis, K. D., Tasker, R. R., Lozano, A. M., and Dostrovsky, J. O. A comparison of the burst activity of lateral thalamic neurons in chronic pain and non-pain patients. *Pain*, 80(3):567–575, 1999.
- Ramcharan, E. J., Gnadt, J. W., and Sherman, S. M. Burst and tonic firing in thalamic cells of unanesthetized, behaving monkeys. *Visual Neuroscience*, 17(1):55–62, 2000.
- Reiss, R. D. *A Course on Point Processes*. Springer, New York, 1993.
- Ripley, B. D. *Stochastic Simulation*. John Wiley & Sons, New York, Chichester, 1987.
- Ritz, C. and Streibig, J. C. *Nonlinear Regression with R*. Springer, New York, 2008.
- Robert, C. and Casella, G. *Introducing Monte Carlo Methods with R*. Springer, 2010.
- Ross, S. M. *A course in simulation*. Macmillan Publishing Company, New York, 1991.
- Schiemann, J. *An exciting in vivo function of ATP-sensitive potassium channels in substantia nigra dopamine neurons - Implications for burst firing and novelty coding*. PhD thesis, Goethe University Frankfurt, 2012.
- Schiemann, J., Klose, V., Schlaudraff, F., Bingmer, M., Seino, S., Magill, P. J., Zaghoul, K. A., Schneider, G., Liss, B., and Roeper, J. K-ATP channels in dopamine substantia nigra neurons control bursting and novelty-induced exploration. *Nature Neuroscience*, accepted, 2012.
- Schneider, G. Messages of oscillatory correlograms - a spike train model. *Neural Computation*, 20(5):1211–1238, 2008.
- Schneider, G. and Nikolic, D. Detection and assessment of near-zero delays in neuronal spiking activity. *Journal of Neuroscience Methods*, 152(1-2):97–106, 2006.

- Schultz, W. Multiple Dopamine Functions at Different Time Courses. *Annual Review of Neuroscience*, 30(1):259–288, 2007.
- Seber, G. A. F. and Wild, C. J. *Nonlinear Regression*. John Wiley & Sons, New York, 1989.
- Sherman, S. M. Dual response modes in lateral geniculate neurons: mechanisms and functions. *Visual Neuroscience*, 13(2):205–213, 1996.
- Shimokawa, T. and Shinomoto, S. Estimating Instantaneous Irregularity of Neuronal Firing. *Neural Computation*, 21(7):1931–1951, 2009.
- 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. Relating neuronal firing patterns to functional differentiation of cerebral cortex. *PLoS Computational Biology*, 5(7):e1000433+, 2009.
- Shinomoto, S., Miyazaki, Y., Tamura, H., and Fujita, I. Regional and laminar differences in in vivo firing patterns of primate cortical neurons. *Journal of Neurophysiology*, 94(1):567–575, 2005.
- Silverman, B. W. *Density Estimation for Statistics and Data Analysis*. Chapman & Hall, London, 1986.
- Simonoff, J. S. *Smoothing Methods in Statistics*. Springer-Verlag, New York, 1996.
- Smith, D. R. and Smith, G. K. A statistical analysis of the continual activity of single cortical neurones in the cat unanaesthetized isolated forebrain. *Biophysical Journal*, 5(1):47–74, 1965.
- Snyder, D. L. *Random Point Processes*. Wiley, New York, 1975.
- Spitzer, F. *Principles of Random Walk*. Graduate Texts in Mathematics. Springer-Verlag, New York, Heidelberg, Berlin, 1976.
- Srinivasan, S. K. *Stochastic Point Processes and Their Applications*. Griffin, London, 1974.
- Staudte, R. G. and Sheather, S. J. *Robust Estimation and Testing*. John Wiley & Sons Ltd., New York, 1990.

Bibliography

- Steriade, M., McCormick, D. A., and Sejnowski, T. J. Thalamocortical oscillations in the sleeping and aroused brain. *Science*, 262(5134):679–685, 1993.
- Stone, C. On a theorem by dobrushin. *The Annals of Mathematical Statistics*, 39(5): 1391–1401, 1968.
- Stoyan, D., Kendall, W. S., and Mecke, J. *Stochastic geometry and its applications*. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley & Sons Ltd., Chichester, 1987.
- Thompson, W. A. *Point Process Models with Applications to Safety and Reliability*. Chapman & Hall, London/New York, 1988.
- van Elburg, R. A. and van Ooyen, A. A new measure for bursting. *Neurocomputing*, 58–60:497–502, 2004.
- Wand, M. P. and Jones, M. C. *Kernel Smoothing*. Chapman & Hall/CRC, London, 1995.
- Werner, G. and Mountcastle, V. B. Neural activity in mechanoreceptive cutaneous afferents: stimulus-response relations, weber functions, and information transmission. *Journal of Neurophysiology*, 28:359–397, 1965.
- Westcott, M. The probability generating functional. *Journal of the Australian Mathematical Society*, 14:448–466, 1972.
- Wilson, C. J., Young, S. J., and Groves, P. M. Statistical properties of neuronal spike trains in the substantia nigra: Cell types and their interactions. *Brain Research*, 136 (2):243–260, 1977.
- Wonnacott, T. H. and Wonnacott, R. J. *Introductory Statistics*. John Wiley & Sons, Inc., New York, 1969.
- Woodroffe, M. *Nonlinear Renewal Theory in Sequential Analysis*. Society for Industrial and Applied Mathematics, Philadelphia, 1982.

Curriculum Vitae

Personal Information

Name: Markus Bingmer
Adress: Taunusstr. 39, 63067 Offenbach
Date of Birth: June 20, 1983
Place of Birth: Hanau, Germany
Marital Status: married
Citizenship: german



Education

- since 10/2008 **Graduate studies in mathematics**, *Goethe University*, Frankfurt am Main.
Dissertation-Supervisor: Dr. Gaby Schneider
- 2003–2008 **Undergraduate studies in mathematics**, *Goethe University*, Frankfurt am Main.
Minor field of study: economics
Thesis-Supervisor: Dr. Gaby Schneider
Thesis-Title: Eine anschauliche Methode zur mehrdimensionalen Ausreißerererkennung in der Kieferorthopädie.
Degree: Diplom (average grade: 'Sehr gut')
- Academic teachers: Prof. Dr. J. Bliedtner, Prof. Dr. P. Kloeden, Prof. Dr. R. Neining, Prof. Dr. A. Wakolbinger, Prof. Dr. J. Wolfart, Dr. B. Ferebee, Dr. G. Schneider
- 1999–2002 **High School Diploma**, *Secondary school*, Bruchköbel.
- 1993–1999 **Certificate of Secondary Education**, *comprehensive school*, Nidderau.
- 1989–1993 **Elementary school**, *Büdesheim*.

Positions

- since 10/2008 **Research assistant**, *Goethe University*, Frankfurt am Main.
At the Institute of Mathematics and in a research project of the Berstein Focus Neurotechnology Frankfurt.
- 04/2007 - 09/2008 **Student assistant**, *Goethe University*, Frankfurt am Main.
At the Institute of Mathematics.

Publications

- [1] M. Bingmer, J. Schiemann, J. Roeper, and G. Schneider. Measuring burstiness and regularity in oscillatory spike trains. *Journal of Neuroscience Methods*, 201(2):426–437, 2011.
- [2] M. Bingmer, V. Özkan, J.-m. Jo, K.-J. Lee, H.-S. Baik, and G. Schneider. A new concept for the cephalometric evaluation of craniofacial patterns (multiharmony). *European Journal of Orthodontics*, 32(6):645–654, 2010.