

Sensory Coding in Natural Environments

Lessons From the Grasshopper Auditory System

DISSERTATION

zur Erlangung des akademischen Grades

doctor rerum naturalium

(Dr. rer. nat.)

im Fach Biophysik

eingereicht an der

Mathematisch-Naturwissenschaftlichen Fakultät I

der Humboldt-Universität zu Berlin

von

Herrn Dipl.-Phys. Christian Machens

geboren am 19. November 1970 in Münster/Westfalen

Präsident der Humboldt-Universität zu Berlin:

Prof. Dr. Jürgen Mlynek

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät I:

Prof. Dr. Bernhard Ronacher

Gutachter:

1. Prof. Dr. Andreas Herz
2. Prof. Dr. Klaus Obermayer
3. Prof. Dr. Bernhard Ronacher

Tag der mündlichen Prüfung: 23. Januar 2002

Mind: “By convention colour, by convention cold, by convention bitter:
in reality atoms and void.”

Senses: “Poor mind, do you take your evidence from us and then try
to overthrow us? Our overthrow is your fall.”

Democritus (460-370 BC)

Contents

1	Introduction	1
1.1	Randomness and Structure	1
1.2	Sensory Systems and Natural Stimuli	2
1.3	Grasshoppers and Acoustic Communication	3
1.4	Overview	4
2	A Physicist's View	6
2.1	Stages of the Research Process	6
2.1.1	Goals of Investigation	6
2.1.2	Definitions and Assumptions	7
2.1.3	The System and Its Environment	9
2.1.4	Model Building	9
2.1.5	Testing a Sensory System	10
2.1.6	Fitting Parameters	11
2.2	Models of the Sensory Environment	12
2.2.1	Low-Order Statistics	13
2.2.2	Invariances, Symmetries, and Higher-Order Statistics	13
2.3	Models of Encoding	13
2.3.1	Integral Series	14
2.3.2	Tuning Curves	15
2.4	Methods of Decoding	16
2.4.1	Stimulus Reconstruction Methods	17
2.4.2	Quantisation of Responses and Stimuli	18
2.5	Assessing the Information Transfer	19
2.6	Fundamental Problems of Input-Output Analysis	20
2.6.1	Dimensionality Reduction	20
2.6.2	Unknown Inputs	21
2.6.3	Long-Term Memory	22
2.6.4	Experimental Design	22
3	Acoustic Communication of Grasshoppers	24
3.1	Behaviour	24
3.1.1	Mate Detection and Selection	25
3.1.2	Sexual Selection	26

3.2	The Sender	27
3.2.1	Carrier Frequencies	27
3.2.2	Amplitude Modulations	29
3.3	The Communication Channel	31
3.4	The Receiver	32
3.4.1	Auditory Receptor Neurons	32
4	Decoding Natural Stimuli	35
4.1	Introduction	35
4.1.1	Questions	35
4.1.2	Stimulus Design	36
4.1.3	Discrimination Analysis	38
4.2	Discrimination Performance	39
4.2.1	Resolution Parameter	39
4.2.2	Time Window	41
4.2.3	Comparison of Stimulus Paradigms	41
4.2.4	Summary of Experiments	43
4.3	Relation of Song Features and Body Symmetry	46
4.3.1	Song Features and Male Impairments	46
4.3.2	Male Impairment and Fluctuating Asymmetry	48
5	Decoding Artificial Stimuli	49
5.1	Introduction	49
5.1.1	Questions	49
5.1.2	Stimulus Design	50
5.1.3	Stimulus Preprocessing and Reconstruction	52
5.2	Single Auditory Receptors	53
5.2.1	A Sample Reconstruction	53
5.2.2	The Importance of Gaps	55
5.2.3	Breakdown at Higher Modulation Frequencies	58
5.3	Population Codes	58
5.3.1	Redundancy	60
5.3.2	Population Reconstruction	60
5.3.3	Gap Detection	64
5.3.4	Reconstruction of a Natural Song	65
5.4	Discussion	68
5.4.1	Implications for Behaviour	68
5.4.2	Discrimination Analysis and Stimulus Reconstruction	68
5.4.3	Use of Naturalistic Stimuli in the Literature	69
6	Adaptive Sampling	71
6.1	Experimental Design	71
6.1.1	The Optimal Stimulus Ensemble	72
6.2	The Algorithm	73

6.2.1	Modelling the Input Distribution	74
6.2.2	The Iterative Algorithm	74
6.3	Applications	76
6.3.1	Spike Count	76
6.3.2	Spike Timing	80
6.4	Neurophysiological Interpretations of the Best Ensemble	85
6.5	Problems and Refinements	86
6.5.1	Estimating the Input-Output Relation	86
6.5.2	Quantising Input and Output	86
6.5.3	Systematic Expansions	87
6.5.4	Incorporation of Constraints	87
7	Conclusions	89
A	Stimulus Design	91
A.1	Rescaling of Calling Songs	91
A.2	Design of Artificial Stimuli	91
B	Electrophysiology	93
C	Quantisation in Response Space	94
C.1	Spike Train Distances	94
C.1.1	Victor-Purpura Distance	94
C.1.2	Van Rossum Distance	95
C.1.3	Limitations	97
C.2	Clustering Algorithm	97
D	Stimulus Reconstruction	98
D.1	Stimulus Preprocessing and Calibration	98
D.2	Linear Reconstruction and Filter Calculation	99
D.3	Signal-to-Noise Ratio	100
D.4	Information Transfer	101
D.5	Coding Efficiency	101
D.6	Redundancy	102

Acknowledgments

It is impossible to remember all those people whose thoughts have influenced me and who have helped, in one or another way, to shape this thesis. First and foremost, however, I want to thank Andreas Herz for his whole-hearted support, his unparalleled optimism and enthusiasm, and his guidance throughout these years. He always encouraged me to pursue my own research directions. I had the opportunity to see the lab grow from the very start; along this way, the grasshopper research effort has constantly expanded and I had the pleasure to closely collaborate with Jan Benda, Hartmut Schütze, and Tim Gollisch. There was always something to learn and discuss and there was always something to laugh about. Furthermore I want to thank Martin Stemmler for his constant curiosity, his many ideas, and his many critical comments (whose frequency and vehemency had a slight tendency to increase when it came towards submitting a manuscript).

I am also grateful to Bernd Ronacher for his firm interest in the doings of the theorists and his patience with their ignorance in biological matters. The contact with his group has strongly shaped my scientific path from physics to biology. Within this group, Petra Prinz performed the experiments for Chapter 5.

During these years, I have also been inspired by those whose research I had the opportunity to accompany and advise, namely Franziska Matthäus, Susanne Schreiber, and Fredrik Edin. Many more people have contributed to the unique atmosphere of the Innovationskolleg Theoretische Biologie and provided a great working environment.

I am greatly indebted to Björn Reineking for many fruitful discussions on fundamental issues of data analysis and modelling. Furthermore, the Graduiertenkolleg “Signalketten in lebenden Systemen” enabled me to get a glimpse of the full spectrum of neuroscience. Therefore, I was always given the chance of looking beyond my own research projects and re-viewing them from a different perspective.

The Deutsche Forschungsgemeinschaft funded my research during the last three years with a scholarship from the Graduiertenkolleg “Signalketten in lebenden Systemen”.

Last not least, I want to thank my parents for their support throughout all these years.

Christian Machens
Berlin, October 2001

Chapter 1

Introduction

1.1 Randomness and Structure

Three monkeys are confined to a cage where they hit the keys of three typewriters. In a booth above them sits Dr. Rosenbaum, a scientist from Columbia University, and surveys the cage. Dr. Rosenbaum intends to demonstrate, once and for all, that three monkeys typing into infinity will sooner or later produce Shakespeare's *Hamlet*. In David Ives' comical play *Words, words, words*, the three monkeys are well aware of their task although they have no clue what *Hamlet* is. In between typing, they scorn at Dr. Rosenbaum for giving them such an uncreative role.

One does not need to be a scientist to understand that Dr. Rosenbaum's project is not bound to be successful.¹ First of all, there is no need to test experimentally that a random combination of infinitely many letters contains the drama *Hamlet*: this is simply a mathematical fact. More importantly, however, the likelihood of producing a meaningful sequence of letters is so incredibly small that Dr. Rosenbaum will die long before the first line of *Hamlet* is written down. Most of the pages produced by the monkeys will just contain a nonsensical combination of letters.

For very much the same reasons, it is practically unfeasible to create a meaningful image such as the one shown in Figure 1.1 (left) by a random mixture of grey shades. We can reformulate this observation in terms of a high-dimensional "image" space: the array of numbers that represents the grey shades of an image corresponds to a single point in this space. Vice versa, every point in this space corresponds to an image. However, almost all of these points, when translated into visual images will rarely resemble anything that we would perceive in nature; just randomly picking a point from this huge space, we end up with an image as the one shown in Figure 1.1 (right). The set of all meaningful visual images constitutes an unbelievably tiny, low-dimensional subset in this huge space and the image of a tree, Figure 1.1 (left), is a point from this subset.

A similar amount of structure can be found for other sensory modalities as well. For instance, a sound pressure wave of a certain length can be discretised in time and

¹In fact, one might rather wonder where Dr. Rosenbaum raised the funding for his project.



Figure 1.1: A natural image and a random image.

the resulting sequence of numbers once more represents a point in a high-dimensional “sound” space. Most points within this space, when played back acoustically, simply sound like noise. The set of points that represents sounds as we hear them in our everyday life is again restricted to a low-dimensional subset.²

1.2 Sensory Systems and Natural Stimuli

The goal of this thesis is not to study the visual or acoustic stimuli of our natural environment. Rather, the introduction should set the stage for a problem that we encounter when studying the computational properties of sensory systems. To test a specific sensory system experimentally, one often presents a set of stimuli while measuring some property of the system. Conventionally, one chooses stimuli with a simple, well-controlled structure that allow to find some simple relation between the stimulus and the property measured. This approach has been highly successful; most of what we know today about information processing in sensory systems can be ascribed to experiments using such simple stimuli. One famous example is the discovery and subsequent analysis of orientation selectivity in the visual cortex (Hubel and Wiesel, 1962). However, within the space of all possible stimuli, these traditional experimental paradigms sample only specific stimulus areas whose relation to the set of behaviourally relevant stimuli is often obscure. Since sensory systems are designed to process natural stimuli, the information gained by tests with simple stimuli is limited. On the other hand, these limitations might be overcome by the use of natural stimuli or artificial stimuli that

²For language, the space of all possible images or sounds corresponds to Jorge Luis Borges’ famous *Library of Babel*, i.e., the library of all possible books of a certain length. In this story, the librarians have explored the library for centuries, yet book after book is just a weird combination of letters. As the library contains all possible books, there are zillions of useful books; alas, the library is too large to find them.

mimic some of the properties of natural stimuli. Let us illustrate in what respect these stimuli can further our understanding of sensory systems.

I. Natural stimuli and behaviour: Sensory systems are built for specific purposes and the stimuli processed by a particular sensory system are of high relevance to the animal's survival or reproductive success. A sensory system will usually extract certain features of natural stimuli and disregard others. Yet why certain features are extracted and others disregarded, can only be understood in the behavioural context. Consequently, only natural stimuli and their relevance allow us to understand why a sensory system computes in a certain way. Moreover, understanding how a sensory system functions in an animal's natural environment opens the path towards an understanding of the animal's behaviour.

II. Natural stimuli and efficient sampling: It might be argued, though, that to understand purely mechanical aspects of a sensory system, experiments using simple artificial stimuli should suffice. Yet there remains the technical problem of the high dimensionality of the input space. It is impossible to test all possible stimuli during the finite duration of an experiment, and any set of artificial stimuli bears the risk of missing many stimuli that might be highly informative about the system. One might, therefore, ask what set of artificial stimuli is least likely to miss any of the essential stimuli. One successful approach has recently been to use artificial stimuli that incorporate some of the statistical properties of natural stimuli (Rieke et al., 1995).

Within this thesis, we will discuss both topics and show how the more traditional use of simple stimuli can be complemented by investigations using natural and naturalistic stimuli. Natural stimuli will allow us to study a sensory system in terms of its behavioural relevance. Naturalistic stimuli promise to capture a large range of relevant stimuli and allow us to discover previously unexplored aspects of a sensory system. Finally, we will show how one can efficiently sample the complete range of relevant stimuli. For concreteness, we will focus on the auditory system of grasshoppers.

1.3 Grasshoppers and Acoustic Communication

The complexity and size of sensory systems varies greatly, from the auditory system of a noctuid moth consisting of just a few neurons (Roeder, 1967), to the primary visual system consisting of ~ 250 million neurons (Hubel, 1988). This size corresponds to the difficulty of the tasks that a sensory system has to fulfil. Essentially, the auditory system of the noctuid moth only needs to detect bat cries while the visual system of a primate has to analyse and interpret a whole variety of complex visual scenes. The amount of computations needed to perform these tasks differs correspondingly.

Somewhere within this range of complexity lies the grasshopper auditory system with a few hundred neurons (for a review, see Pollack, 1998). This system has two advantages: (1) The set of natural stimuli is well known, most importantly the communication signals that are employed by grasshoppers in the mate finding process (Elsner,

1974). These signals exhibit enough structure to be interesting for our analysis and are at the same time simple enough to allow for some mathematical characterisation. (2) Despite their simplicity, some grasshoppers show highly evolved behavioural patterns which are accessible to systematic investigations (e.g. Helversen, 1972; Helversen and Helversen, 1994). Indeed, several studies have been performed with respect to acoustic communication, providing us with the great advantage of interpreting neurophysiological findings within a behavioural context.

1.4 Overview

In Chapter 2, we will give an overview of some approaches common to the field of systems neuroscience. The chapter serves to present the methods and techniques used throughout this thesis from a unifying perspective. Examining how sensory systems have been tested and modelled shall also provide a first hint in which way tests of the system could be made more effective.

In Chapter 3, we will take a brief look at acoustic communication in grasshoppers. We will study some of the behaviourally relevant stimuli, namely the advertisement calls of male grasshoppers of the species *Chorthippus biguttulus*. These calls exhibit several characteristic features that distinguish different individuals.

In Chapter 4, we will show that individual variations between these mating calls are preserved by the auditory periphery of females, allowing females not only to detect the mating calls of conspecific males, but also to choose among them. A discrimination analysis based on eight representative calls indicates that the information conveyed by single auditory receptor neurons suffices to distinguish the calls even if they differ in minor aspects only. This investigation provides a first step towards understanding the specific design of the auditory periphery in terms of sexual selection.

In Chapter 5, we broaden our view and investigate which region of stimulus space is optimally encoded by auditory receptor neurons. We do this by presenting different stimulus ensembles that correspond to different regions in stimulus space, and we relate the stimulus ensemble that best explores the dynamic range of the system with certain aspects of the advertisement calls of male grasshoppers. Both on the single cell and the population level, it turns out that the most “interesting” aspects of the calls are closely related to the aspects that allow a successful discrimination on the basis of single auditory receptors.

Chapters 4 and 5 show that the properties of auditory receptors are closely matched to the behaviourally relevant aspects of the communication signals. This close match reveals itself most conspicuously in the high rates of information conveyed about the relevant stimuli. Hence, even if the relevant stimuli are a priori unknown, one might be able to find them by a systematic search for stimulus ensembles that maximise the information transfer. Additionally, such stimulus ensembles provide an efficient sampling of the respective sensory system. In Chapter 6, we present an algorithm that automatically approaches this stimulus ensemble in an iterative manner during the course of an experiment. This algorithm is applicable to any input-output system and therefore extends

the scope of the present thesis beyond both natural and artificial sensory systems.

Chapter 7 provides a short summary of the results and their implications for the study of sensory systems. The Appendices contain technical details that were excluded from the main text for the purpose of readability.

Chapter 2

A Physicist's View of Sensory Neurophysiology

When scanning experimental studies of sensory systems on the systems level, one finds that problems are often tackled with seemingly different methods and approaches. Various sorts of stimuli are used to investigate a specific problem and the data are subsequently analysed with a multitude of tools from data analysis. The diversity of approaches found in the literature sometimes obscures the common structure of the underlying problem: the analysis of input-output systems. A similar difficulty holds for this thesis where we will employ different methods to investigate the same system. Therefore, the present chapter endeavours to derive the relevant approaches used from first principles, thus revealing otherwise hidden assumptions and providing the theoretical foundations for the rest of the thesis. As we are also interested in issues of experimental design, we will place a special emphasis on the selection of stimulus ensembles.

We will assume that the reader is familiar with the basic concepts and terms of systems neuroscience and with the fundamentals of probability theory. Background literature on neuroscience can be found in Kandel et al. (1991); an introduction to probability theory is provided in Papoulis (1991).

2.1 Stages of the Research Process

2.1.1 Goals of Investigation

Several experimental methods exist to investigate the activity in a sensory system. These range from intracellular recordings of single neurons to monitoring the complete system at reduced spatial and temporal resolution. For instance, one might stimulate the auditory system of grasshoppers with sound and simultaneously record the voltage response of single receptor neurons with electrodes that penetrate the cell membrane (Römer, 1976). Investigations of the mammalian visual system might rely on multi-electrode arrays that record local field potentials in the lateral geniculate nucleus while the animal views a movie on a monitor screen (Stanley et al., 1999). The large-scale

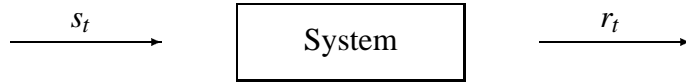


Figure 2.1: At time t , the system receives a multi-variate input s_t while we measure the multi-variate output r_t .

activity within a restricted area of a sensory system, e.g., the antennal lobe in the olfactory system of bees, might be measured using calcium imaging techniques (Galizia et al., 1999).

In all of these examples, the first goal is usually to find some accurate description of the relation between the stimuli presented and the measured response. Such a description may be purely phenomenological or it may rely, to varying degree, on a mechanistic background. A description of the stimulus-response relation will usually help to understand part of what the sensory system does, e.g., what information it processes. In the literature, the problem of describing the stimulus-response relation is often referred to as the *encoding* problem (Rieke et al., 1997; Dayan and Abbott, 2001).

Closely related is the *decoding* problem, concerned with the question of what an organism can infer about its environment from the responses of its neurons. To yield meaningful answers, a decoding procedure additionally requires some knowledge about the structure of the organism’s sensory environment (Rieke et al., 1997; Dayan and Abbott, 2001). Given such knowledge, one can reconstruct the perception of an animal in its natural surroundings.

Taking the “organism’s point of view” is, therefore, an important step towards studying how the sensory system’s function relates to the behaviour of the animal. Building on this fundament, one might ultimately seek to put the system and its function in the context of evolution (Mayr, 1997).

2.1.2 Definitions and Assumptions

Within this chapter, we will focus on the encoding and decoding problems. Therefore, we will be concerned with the analysis of input-output systems, i.e., systems that receive some input and produce an output which we can measure (cf. Figure 2.1). We take for granted that we have complete freedom to choose the inputs at our own will. We will use the terms input and stimulus interchangeably; the same holds for output and response.

First, we need to describe input and output in mathematical terms. We will assume that both can be modelled as discrete time series of multi-variate variables s_t and r_t , respectively. At a particular time instant t , we will write $\mathbf{s}_t = (s_t, s_{t-1}, \dots)$ for the complete, infinite history of the stimulus and $\mathbf{r}_t = (r_t, r_{t-1}, \dots)$ for that of the response. As any physical quantity is bandwidth-limited in time, no loss of generality results from the assumption of discrete time steps if these steps are small enough. Furthermore, physical quantities are always bounded in magnitude and the effective resolution of input

and output is limited by noise in the system. We therefore also assume that both input and output are discrete and finite, i.e., both can be modelled as a sequence of discrete numbers. Whenever we choose to describe input or output as vectors or continuous functions, we take this description merely as a convenient approximation of the discrete reality. A finite segment of a time series will be denoted by $\mathbf{x}_{t,t-m} = (x_t, \dots, x_{t-m})$.

The most general description of the stimulus-response relation assigns a conditional probability to every possible stimulus-response pair. In particular, if the stimulus is given by \mathbf{s}_t , then the conditional probability for a response \mathbf{r}_t will be denoted by $p(\mathbf{r}_t|\mathbf{s}_t;t)$. The explicit dependence on the time instant t leaves the possibility that the stimulus-response relation of the system changes with time: the system might age, for instance. However, such changes might play a negligible role during the finite duration of an experiment. Within the analysis of input-output systems, one therefore makes some reasonable assumptions that simplify the problem (Wiener, 1958; Marmarelis and Marmarelis, 1978):

I. Stationarity: The intrinsic properties of the system do not change over time. In this case, the input-output relation at any time t is given by the same conditional probability distribution,

$$p(\mathbf{r}_t|\mathbf{s}_t) = p(\mathbf{r}_t|\mathbf{s}_t;t) \quad \text{for all } t. \quad (2.1)$$

This time-invariance assumption rules out the existence of unknown inputs or input-independent dynamics, unless these are stationary,¹ i.e., time-invariant themselves.

II. Finite memory: Inputs before a time instant $t - m$ do not influence the output at time t where $m > 0$ is called the temporal memory of the system. Summing over the history of all outputs that are possible for a given input \mathbf{s}_t , we obtain the conditional probability $p(r_t|\mathbf{s}_t) = \sum_{\mathbf{r}_{t-1}} p(\mathbf{r}_t|\mathbf{s}_t)$.² Systems with a finite memory then fulfil the requirement

$$p(r_t|\mathbf{s}_t) = p(r_t|\mathbf{s}_{t,t-m}) \quad (2.2)$$

for the output r_t at time t .

III. Causality: Any physical system is causal, i.e., inputs after a time instant t do not influence the output at or before that time instant t ,

$$p(\mathbf{r}_t|\mathbf{s}_{t+u}) = p(\mathbf{r}_t|\mathbf{s}_t) \quad \text{for all } u > 0. \quad (2.3)$$

While the first two assumptions constrain the physics of the input-output system, causality is a characteristic of every real input-output system; the respective assumption

¹Our use of the term stationarity is sometimes referred to as “strict stationarity” (Brockwell and Davis, 1991), in contrast to “weak stationarity” which only requires the time-invariance of the first and second-order moments of a time series and its underlying probability distribution.

²Using the time series notation, we obtain $p(\mathbf{r}_t|\mathbf{s}_t) = p(r_t, \mathbf{r}_{t-1}|\mathbf{s}_t) = p(\mathbf{r}_{t-1}|r_t, \mathbf{s}_t)p(r_t|\mathbf{s}_t)$. Hence, $\sum_{\mathbf{r}_{t-1}} p(\mathbf{r}_t|\mathbf{s}_t) = \sum_{\mathbf{r}_{t-1}} p(\mathbf{r}_{t-1}|r_t, \mathbf{s}_t)p(r_t|\mathbf{s}_t) = p(r_t|\mathbf{s}_t) \sum_{\mathbf{r}_{t-1}} p(\mathbf{r}_{t-1}|r_t, \mathbf{s}_t) = p(r_t|\mathbf{s}_t)$. Note that the summation over \mathbf{r}_{t-1} means summing over all possible values that can be adopted by \mathbf{r}_{t-1} .

is, therefore, of a more formal nature in that it restricts the set of possible probability distributions. The conditions of stationarity and finite memory will only approximately be fulfilled by biological systems. We postpone the discussion of possible violations and their consequences to the end of this chapter.

2.1.3 The System and Its Environment

The conditional probability distribution $p(\mathbf{r}_t|\mathbf{s}_t)$ describes the response of the system to any stimulus. Therefore, it provides the foundation for a study of the system's functionality. However, to tackle the decoding problem, we rather need the probability distribution $p(\mathbf{s}_t|\mathbf{r}_t)$, which denotes the probability that the stimulus was \mathbf{s}_t for a given response \mathbf{r}_t . The Formula of Bayes allows us to formulate the mathematical relation between these two distributions,

$$p(\mathbf{s}_t|\mathbf{r}_t) = \frac{p(\mathbf{r}_t|\mathbf{s}_t)p(\mathbf{s}_t)}{p(\mathbf{r}_t)} . \quad (2.4)$$

where the prior knowledge of the outside world is summarised in the probability distribution $p(\mathbf{s}_t)$. The unconditional probability of a certain response can be obtained via $p(\mathbf{r}_t) = \sum_{\mathbf{s}_t} p(\mathbf{r}_t|\mathbf{s}_t)p(\mathbf{s}_t)$. Bayes' Formula shows that the decoding problem is unvariably linked to the prior probability of a stimulus \mathbf{s}_t (Rieke et al., 1997). A proper understanding of the natural environment and its statistical structure is, therefore, a prerequisite to any interpretation regarding the perception of a particular animal in nature.

2.1.4 Model Building

We will hardly ever know the "real" input-output relation $p(\mathbf{r}_t|\mathbf{s}_t)$ or the "real" distribution of natural stimuli $p(\mathbf{s}_t)$. The goal is therefore to approximate these distributions or some of their aspects by a model. Here we define model as any, possibly stochastic, relation between variables. Hence, the most general form of a model is again a probability distribution. For instance, if we model the system's input-output relation by a stochastic difference equation, then the solutions to this equation and their respective probabilities yield the probability distribution of the model.³ We will refer to these modelling distributions as $p_M(\mathbf{r}_t|\mathbf{s}_t)$ or $p_M(\mathbf{s}_t)$ where the subscript M stands for model.

The model building process usually starts with a detailed specification of input and output.⁴ Possible adjustments of input and output might reflect experimental limits as well as prior notions of what is necessary and relevant. For instance, the experimental setup might necessitate a relatively coarse discretisation of input or output. If the output is given by the electrical activity within a sensory system, then one will usually neglect

³Usually, one rather uses stochastic differential equations. As noted before, however, this can be seen as a continuous approximation to the discrete reality.

⁴By no means a trivial task: For instance, the output of a single neuron consists of many dynamic variables. Voltage, although of utmost significance for rapid signaling via action potentials, is only one of them; if information is "hidden" in other dynamical variables, we might miss an important feature of the system (Eguia et al., 2000).

details of the voltage trace and only retain the precise times of the action potentials. In any case, both stimulus and response space are usually very high-dimensional. This fact notably aggravates the modelling process and is often referred to as the “curse of dimensionality” (Bellman, 1961; Bishop, 1995).

Given a proper definition of input and output and given that some ideas exist about the system under study, one can develop a model $p_M(\mathbf{r}_t|\mathbf{s}_t)$ for the input-output relation or some aspect of it. Depending on one's goals and prior knowledge, there are two types of models one can build: mechanistic or phenomenological models. Mechanistic models incorporate the biophysical structure of the system. The scientific benefit of such models is that they are able to predict the behaviour of the system from its components, thus bridging levels of description. A famous example is the explanation of action potential generation in terms of the dynamics of ionic currents (Hodgkin and Huxley, 1952). In contrast, phenomenological models are often less detailed but nevertheless capture the behaviour of the system, while their parameters and variables are not interpretable in biological or physical terms. In this category, we include “non-parametric” models of the input-output relation, i.e., models in which the probabilities $p(\mathbf{r}_t|\mathbf{s}_t)$ are directly measured. Note, though, that many models are located in between these opposing poles, thus having some mechanistic and some phenomenological components. Examples include the integrate-and-fire neuron model (see, e.g., Hoppenstaedt, 1997).

To investigate the decoding problem, we additionally need a model $p_M(\mathbf{s}_t)$ of the sensory environment as follows from Bayes' Formula, equation (2.4). Such models seek to describe the subset of natural stimuli, as discussed in the introduction, and are thus of a phenomenological nature. The same holds for any approaches that model the distribution $p(\mathbf{s}_t|\mathbf{r}_t)$ or aspects of it as the parameters of these models are only interpretable via $p(\mathbf{r}_t|\mathbf{s}_t)$.

Within this thesis, we will neglect mechanistic models and exclusively focus on phenomenological models.

2.1.5 Testing a Sensory System

A crucial stage within the research process is the design of a specific experimental protocol. Due to the huge dimensionality of the input space and the limited experimental time, only a small subset of the possible stimuli can be tested on the system. This fact puts a severe constraint on the number of parameters that one can fit to a model. Furthermore, inputs that lead to a broad range of outputs will often constitute only a tiny subset of all possible inputs. If too many inputs lie outside this particular region (e.g., if the inputs do not elicit any action potentials) the data might not be particularly instructive, and the respective testing time will be lost. Consequently, care has to be taken to design sophisticated tests of the system. There are two lines of thought that guide the design of test stimuli for sensory systems.

I. Deductive approach: If there exists already a hypothesis or, more generally, a model $p_M(\mathbf{r}_t|\mathbf{s}_t)$, then one usually designs test stimuli that enable one to fit the model

parameters while minimising the experimental time. Ultimately, one can choose every new sample point to minimise the remaining uncertainty about the parameters (MacKay, 1992). Note, however, that such strategies seek to acquire data that is highly informative about a particular model; if the model does not provide a good fit, there is the risk that the collected data is not particularly informative about the system itself.

An additional criterion in the choice of test stimuli has often been whether they allow a technically simple fit of the model parameters. This is the case for the well-known white-noise analysis which has been developed by Wiener (1958) and Lee and Schetzen (1965). However, such considerations become less important with the current increase in computational capabilities.

II. Inductive approach: Sometimes, there exist no prior notion about the system’s functionality and no model can be constructed beforehand. Nevertheless, one might know the stimuli that the system has to process in its natural environment. Such knowledge can be used to design stimuli that capture the relevant characteristics of natural stimuli. If we collect data with such stimuli, we can expect the data to be highly informative about the system under study. These data can then guide us in developing a model $p_M(\mathbf{r}_t|\mathbf{s}_t)$. This approach has often been the method of choice within neuroethology. The auditory system of grasshoppers, for instance, has often been investigated with stimuli that were abstractions of natural stimuli (Ronacher and Römer, 1985; Ronacher and Stumpner, 1988).

The question of how to select stimuli to test a given input-output system will accompany us throughout this thesis. In any case, one wants to collect data that is highly informative about the given system. Moreover, it is often desirable to obtain data that is “robust” against a specific question or model, i.e., that can be used to investigate several aspects of a system. A priori, it is unclear how one should design experiments that collect such data. We will take a closer look at this issue in Chapter 6.

2.1.6 Fitting Parameters

Equipped with both a model $p_M(\mathbf{r}_t|\mathbf{s}_t)$ and appropriate data, one can fit the model parameters by standard statistical methods such as the maximum likelihood method (cf., e.g., Ripley 1996). Ideally, the validity of a model is tested on new data sets. To see whether the model is a proper global description of the system, the new data set should be obtained for a different set of stimuli. For instance, one can probe the system with different stimulus ensembles $p(\mathbf{s}_t)$; if the parameter values of the model change with the stimulus ensembles, then the model is only locally valid. Note, however, that a local fit often suffices if it is sufficiently good for the ensemble of natural stimuli, thus covering the biologically relevant range.

Model building is essentially an iterative procedure. One builds a model, designs test stimuli, measures the system’s responses, fits the model parameters, and then refines the model in the next iteration. Therefore, the stages of the research process

presented here are meant to be schematic while in practice they will be much more interwoven.

The application of input-output analysis within systems neuroscience is certainly not restricted to sensory systems. However, as this thesis focuses primarily on sensory systems, we will neglect any other applications for the rest of the chapter and focus on experimental setups where the output is given by the action potentials (spikes) of one or more neurons. In the following, we will discuss some approaches to the modelling task which we will employ later on.

2.2 Models of the Sensory Environment

It is useful to start any investigation of sensory systems with an examination of the stimuli the system processes. Knowing the ensemble of natural stimuli will aid in the process of building a model $p_M(\mathbf{r}_t|\mathbf{s}_t)$, it will allow to design efficient and informative tests of the system, and it will enable us to infer what an animal perceives in nature.

As noted in the introduction, animals confront only a very specific set of stimuli in their natural environment. We can assign a certain probability to every one of these stimuli, depending on their probability of occurrence. The sensory environment of the system is then completely characterised by the probability distribution $p(\mathbf{s}_t)$. Assuming that temporal correlations practically vanish for large enough interval size n , it is sufficient to model $p(\mathbf{s}_{t,t-n})$.

As the stimulus space is very high-dimensional, one might wonder whether $p(\mathbf{s}_t)$ is at all well-defined. For instance, visual images might very well be unique so that one cannot assign a probability to them. The distribution $p(\mathbf{s}_t)$ is, therefore, of a more hypothetical nature. In practice, one will only model aspects of this distribution, such as the sample statistics.

Another problem arises from non-stationary changes of the sample statistics in different environments. For example, the statistics of natural images is different in a desert and a forest. Likewise, the statistics of natural sounds highly depends on the environment in which the sounds are recorded. We therefore need to model several probability distributions $p_E(\mathbf{s}_t)$ where the subscript E denotes the specific environment. Most noteworthy, some sensory systems can adapt their response characteristics to changes in the statistics of the sensory environment (Brenner et al., 2000). This fact serves to show that sensory systems are often highly specialised towards a particular set of stimuli, thus emphasising the importance of the statistics of natural stimuli.

2.2.1 Low-Order Statistics

Perhaps the simplest quantities available to characterise $p(\mathbf{s}_t)$ are low-order statistics such as mean and (co)variance. The low-order statistics of static visual images have been extensively studied for a long time (for a review, see Ruderman, 1994). As objects can appear everywhere in a visual scene, the probability distribution of images exhibits a translation invariance so that its covariance matrix is completely described by the

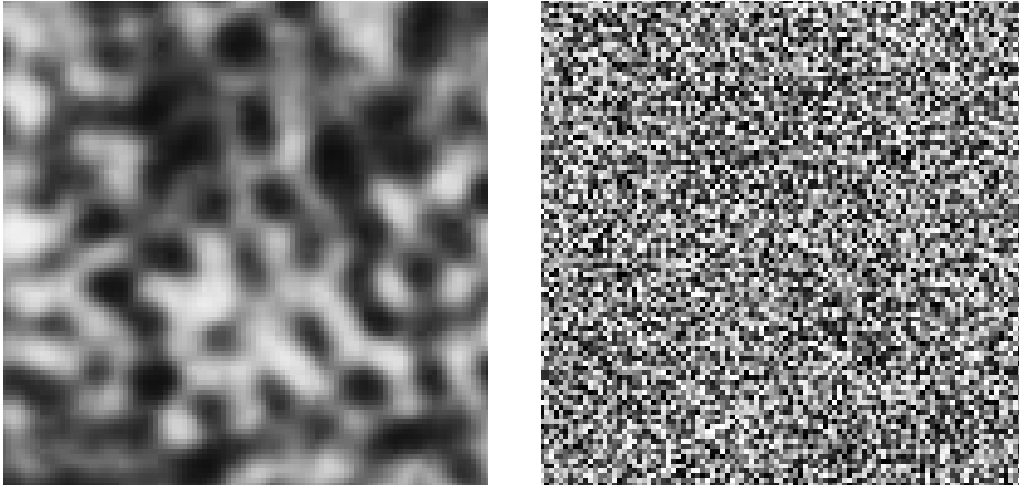


Figure 2.2: Left: A random image whose power spectrum behaves like $S(k) \propto k^{-1.8}$ for higher spatial frequencies. Consequently, the image features long-range correlations. Right: A completely random image, same image as in Figure 1.1.

power spectrum of spatial frequencies. One finds that the power spectrum of images usually obeys a power law scaling, $S(k) \propto k^{-2+\eta}$ where k denotes spatial frequency and $\eta \ll 1$ (Field, 1987). Such knowledge can be used to obtain a first approximation $p_M(\mathbf{s}_t)$. When drawing random images from this distribution, the stimuli do not at all resemble natural stimuli, cf. Figure 2.2. However, they do capture some of the characteristic features of natural stimuli such as long-range correlations.

The low-order statistics of different natural sound ensembles exhibit a similar power law scaling with respect to modulation frequencies within different frequency bands (Attias and Schreiner, 1997). For the sounds produced by grasshoppers, there is no real need to study frequency bands as differences therein are of negligible behavioural relevance. Here we encounter a great advantage of the auditory system of grasshoppers: the behavioural relevance of the communication signals provides a valuable guidance for studying the statistical properties of acoustic communication signals. Moreover, many of the important properties of these signals are already captured by low-order statistics as we will see in Chapter 3.

2.2.2 Invariances, Symmetries, and Higher-Order Statistics

Within this thesis, it will not be necessary to move beyond low-order statistics. However, for higher areas of sensory systems, it is often essential to do so. Estimating higher-order statistics such as the three- and four-point correlation functions is made very difficult by the large amounts of data needed. Additionally, these correlation functions are often hard to interpret. Therefore, other methods have been used to capture properties of the distributions beyond the low-order statistics. These include attempts to uncover hidden symmetries and invariances, e.g., size invariance of visual stimuli (Ruderman, 1994). Furthermore, researchers have sought to factorise the ensemble of

natural stimuli $p(\mathbf{s}_t)$ by transforming \mathbf{s}_t into statistically independent parameters (Olshausen and Field, 1996).

2.3 Models of Encoding

In many cases, it is impossible to model the complete stimulus-response relation $p(\mathbf{r}_t|\mathbf{s}_t)$ at once. However, one can try to break up the modelling into manageable pieces by neglecting certain aspects of the system or by restricting a model $p_M(\mathbf{r}_t|\mathbf{s}_t)$ to particular regions in input and output space. Here we present two common approaches that exemplarily show how input-output systems can be modelled and tested.

2.3.1 Integral Series

The first approach goes under the name of Volterra-Wiener system identification and seeks to model the stimulus-response relation on the complete stimulus space (Wiener, 1958; Marmarelis and Marmarelis, 1978). For that purpose, one first determines the relation between the stimulus and the mean of the system's conditional probability distribution. For the average response at time t , we write

$$\bar{r}(\mathbf{s}_t) = \sum_{\mathbf{r}_t} r_t p(\mathbf{r}_t|\mathbf{s}_t) \quad (2.5)$$

$$= \sum_{r_t} r_t p(r_t|\mathbf{s}_t) \quad (2.6)$$

$$\stackrel{(2.2)}{=} \sum_{r_t} r_t p(r_t|\mathbf{s}_{t,t-m}) \quad (2.7)$$

which in our case is given by the time-varying rate of action potential generation. This “firing rate” is then assumed to represent the relevant output for the transformation performed by the system. In practice, the firing rate can be estimated by averaging over repeated presentations of the same stimulus, the resulting average is usually called the peristimulus time histogram (PSTH). Note that this procedure averages over both measurement errors and the intrinsic stochasticity of the system.

The Volterra integral series presents a systematic way to model the mapping between the time series $\bar{\mathbf{r}}_t$ and \mathbf{s}_t . In the limit of continuous time, $\bar{\mathbf{r}}_t \rightarrow \bar{r}(t)$ and $\mathbf{s} \rightarrow s(t)$, and for single-variate responses and stimuli, the Volterra integral series is given by

$$\bar{r}(t) = k_0 + \int_0^\infty d\tau k_1(\tau) s(t-\tau) + \int_0^\infty \int_0^\infty d\tau_1 d\tau_2 k_2(\tau_1, \tau_2) s(t-\tau_1) s(t-\tau_2) + \dots$$

where k_0 and the integral kernels $k_1(\tau)$ and $k_2(\tau_1, \tau_2)$ provide the parameters of the model. To enforce the finite temporal memory of the system, the integral kernel must vanish for any $\tau > m$. For all practical purposes, the Volterra series has to be terminated after two or at most three orders as the number of parameters that need to be fitted grows exponentially with every order included.

To fit the model parameters to a specific input-output system, we need to design an ensemble of test stimuli. The integral kernels can then be estimated by minimising the mean-square error between the measured and the modelled output.⁵ In the most general case, the mean-square error has to be computed by averaging over the complete ensemble of test stimuli. The averaging operation can be tremendously simplified by using stationary stimuli which allow to replace the ensemble average by a time average. It then suffices to test a single stationary stimulus of sufficient length. As the parameters, i.e., the integral kernels, enter only linearly in the Volterra integral equation, the mean-square minimisation results in a set of linear equations which, given modern computational capabilities, can nowadays be solved numerically.

To complete the modelling process, the time-varying firing rate $\bar{r}(t)$ has to be complemented by a statistical model of the spiking process. The simplest and most widely used model of this kind is the Poisson process which assumes that spikes are generated independently of each other and that the probability of occurrence of a spike is directly proportional to the firing rate at that particular time instant. For the Poisson model, the conditional probability distribution $p(\mathbf{r}_t|\mathbf{s}_t)$ decomposes in the following way,

$$p(\mathbf{r}_t|\mathbf{s}_t) = \prod_{n=0}^{\infty} p(r_{t-n}|\mathbf{s}_t) \stackrel{(2.3)}{=} \prod_{n=0}^{\infty} p(r_{t-n}|\mathbf{s}_{t-n}) \quad , \quad (2.8)$$

greatly simplifying the overall problem.

Based on tests with Gaussian white-noise stimuli, the Volterra-Wiener system identification has been extensively used to study neural responses in the early visual or auditory system (Marmarelis and Marmarelis, 1978; Palm and Pöpel, 1985; Eggermont, 1993). In this context, we mention two difficulties and their solutions: (1) Many of the nonlinearities encountered in neural systems are caused by saturations whose sigmoid shape is not well-fitted by quadratic or cubic terms. Hence, the integral series is usually terminated after the linear term and nonlinearities are accounted for by static transformations of stimulus or response (Hunter and Korenberg, 1986; Korenberg and Hunter, 1986). Nonlinearities caused by adaptation processes are usually neglected. (2) As the Poisson process does not incorporate a refractory period, it gives only a rough approximation of a neuron's spiking process. More realistic models include correlations between spikes (Gabbiani and Koch, 1998; Barbieri et al., 2001; Keat et al., 2001) and correlations of the spiking behaviour to the average firing rate.

In principle, Gaussian white-noise stimuli sample the complete stimulus space and will therefore cover all relevant ranges. In practice, however, white noise-stimuli sam-

⁵Note that the integral kernels computed by mean-square minimisation are in general not the Volterra integral kernels. The term Volterra integral kernels is reserved for the expansion coefficients of the system's functional around zero input. For the case of discrete time, the Volterra integral kernels therefore model the expansion coefficients of a Taylor series. Hence, the integral kernels computed by minimising the mean-square error only correspond to the Volterra integral kernels if the terminated integral series completely describes the system, i.e., if the mean-square error between measured and modelled response is zero (Palm and Pöpel, 1985).

ple the stimulus space too coarsely when it comes to higher sensory areas; they fail, for instance, to provide sufficient data to characterise the properties of auditory interneurons in grasshoppers (Sippel and Breckow, 1983). Higher sensory areas often react only to a very specific set of stimuli and Gaussian white-noise stimuli are too unlikely to ever sample this subset; this is exactly the problem mentioned in the introduction.

2.3.2 Tuning Curves

When using white noise or other stationary stimuli, many stimulus parameters are varied at once. A simpler, classical method characterises the stimulus-response relation by varying only a single stimulus parameter which corresponds to modelling the input-output relation on a very specific subset of stimuli only. The resulting relations are often called tuning curves when they have a bell-like or sigmoid shape, suggesting that the system under study is “tuned” to encode the stimulus parameter varied (Dayan and Abbott, 2001). Details of the responses’ precise spike pattern are disregarded and it is assumed that all relevant aspects of the response are contained in the average firing rate. Accordingly, we further simplify equation (2.7) by averaging over n response time bins,

$$\bar{r}(\mathbf{s}) = \frac{1}{n} \sum_{u=t-n+1}^t \bar{r}(\mathbf{s}_{u,u-m}) \quad . \quad (2.9)$$

Due to the finite memory of the system, the effective stimulus space with respect to time is only $(m+n)$ - dimensional. Every point in this space can be assigned a single number: the average response \bar{r} . To measure a tuning curve, one designs stimuli that can be described by a single parameter α . The set of parametrised stimuli $\mathbf{s}_{t,t-n-m}^{(\alpha)}$ then forms a one-dimensional curve through stimulus space and the respective tuning curve is given by

$$f(\alpha) = \bar{r}(\mathbf{s}_{t,t-n-m}^{(\alpha)}) \quad . \quad (2.10)$$

Deviations from $f(\alpha)$ are often assumed to be Gaussian and it suffices to estimate their standard deviation. The term tuning curve is usually reserved for stimulus parameters that drive the respective system. However, it is also of interest to consider stimulus parameters that move along isoresponse-manifolds in stimulus space, as these describe the invariances of the system.

As the input-output relation is only one-dimensional, the system can be tested by simply stepping through all relevant parameter values, repeating every stimulus several times, and finally plotting the curve $f(\alpha)$. With a judicious choice of parametrised stimuli, such a curve can provide an excellent first description of the system’s functionality. Hence, the success of this approach depends on the choice of appropriate stimulus parameters, i.e., directions and regions in stimulus space that are informative about the system. This is by no means a trivial task: for instance, it has taken quite some time to determine the stimuli that neurons in the visual cortex V1 respond to; orientation selectivity of cells in V1 was discovered with a bit of fortune (Hubel, 1988). Note, though, that a simple parametrisation of the stimulus ensemble always risks to miss stimuli that

might be highly informative about other aspects of the system. In this respect, tuning curves provide only a first step towards characterising a system's input-output relation.

2.4 Methods of Decoding

Eventually, the organism has to base its decisions on the representation of information achieved by the neurons within the sensory system. The animal's potential knowledge of the world is given by the probability distribution $p(\mathbf{s}_t|\mathbf{r}_t)$. As external observers, we can use this distribution to investigate what an animal can infer about its environment from the responses of its sensory neurons. To decode the response, every response needs to be assigned to a specific set of stimuli. A decoding procedure is always connected to a certain loss function that specifies how costly a wrong assignment is.⁶ Depending on the nature of the loss one is willing to admit, different decoding procedures can be formulated.

If we have a successful model of the encoding process, $p(\mathbf{r}_t|\mathbf{s}_t)$, and a model of the environment $p(\mathbf{s}_t)$, we can use Bayes' Formula, equation (2.4), to construct the distribution $p(\mathbf{s}_t|\mathbf{r}_t)$. However, any modelling effort carries some errors with it; when modelling the distribution $p(\mathbf{r}_t|\mathbf{s}_t)$, for instance, parameters are often fitted so as to minimise the prediction error. They are, therefore, not optimised to minimise the loss function of a possible decoding procedure. Hence, one usually seeks to model the distribution $p(\mathbf{s}_t|\mathbf{r}_t)$ such that a given loss function is minimised (Ripley, 1996; Dayan and Abbott, 2001). Here, we will focus on methods that we will use later on in Chapters 4–6.

2.4.1 Stimulus Reconstruction Methods

Within stimulus reconstruction methods, the goal is to assign every response to a single stimulus only. To quantify the deviations between the actual stimulus and the reconstructed stimulus, one commonly uses the mean-square error, $e = \langle (\mathbf{s}_t - \hat{\mathbf{s}}_t)^2 \rangle$ as a loss function. Here \mathbf{s}_t denotes the actual stimulus, $\hat{\mathbf{s}}_t$ the reconstructed stimulus, and the angular brackets averaging over the joint stimulus ensemble $p(\mathbf{r}_t, \mathbf{s}_t) = p(\mathbf{r}_t|\mathbf{s}_t)p(\mathbf{s}_t)$ or, in the case of stationary stimulus ensembles, over time. The stimulus estimate $\hat{\mathbf{s}}_t$ that minimises the mean-square error is given by the mean of the conditional probability distribution $p(\mathbf{s}_t|\mathbf{r}_t)$ (see, e.g., Poor, 1994; Rieke et al., 1997).

Observing a response r_t at time t will supply us with information about the complete past of the stimulus. Due to the finite memory of the system, the information will reach

⁶A trivial decoding procedure would assign the complete stimulus space to every response. Needless to say, such an assignment wastes all the information about the stimulus that was present in the original probability distribution $p(\mathbf{s}_t|\mathbf{r}_t)$.

back to time $t - m$ only. Therefore, the conditional mean at time $t - m$ is given by

$$\bar{s}_{-m}(\mathbf{r}_t) = \sum_{\mathbf{s}_t} s_{t-m} p(\mathbf{s}_t | \mathbf{r}_t) \quad (2.11)$$

$$= \sum_{s_{t-m}} s_{t-m} p(s_{t-m} | \mathbf{r}_t) \quad (2.12)$$

$$\stackrel{(2.3)}{=} \sum_{s_{t-m}} s_{t-m} p(s_{t-m} | \mathbf{r}_{t,t-m}) \quad (2.13)$$

Similarly to Section 2.3.1, we will focus on the single-variate case and take the continuous limit of the discrete time series, i.e., $\bar{s}_t \rightarrow \bar{s}(t)$ and $\mathbf{r} \rightarrow r(t)$. Again, we can expand the conditional mean into a Volterra integral series. Shifting our focus to time t , we find that the decoded stimulus now depends on the next m time steps of the response. With the expansion restricted to first order, the integral series is therefore given by

$$\bar{s}(t) = h_0 + \int_{-\infty}^0 d\tau h_1(\tau) r(t - \tau), \quad (2.14)$$

where h_0 and the linear reconstruction kernel $h_1(\tau)$ are the model parameters. The kernel will vanish for any $\tau < -m$, enforcing the finite temporal memory of the decoding procedure.

The stimulus reconstruction method is only sensible if the reconstructed stimulus was actually likely to have caused the response. For instance, if $p(\mathbf{s}_t | \mathbf{r}_t)$ is a bimodal distribution, then the conditional mean can well lie right in between the modal hills of the distribution. In this case, it might be completely improbable that the reconstructed stimulus, when presented to the system, would ever elicit the measured response (Rieke et al., 1997). This problem can usually be avoided by an appropriate transformation of the stimulus space. In fact, stimulus preprocessing is often essential to a successful reconstruction. As reconstruction amounts to a stimulus-dependent rate estimate (cf. Appendix D), preprocessing essentially has to mimic the forward transformation. Reconstruction is therefore a method that meets the forward approaches half way.

The parameters h_0 and $h_1(\tau)$ are estimated by minimising the mean square error. Hence, the technical procedure is equivalent to that of the Volterra-Wiener analysis⁷ and the same holds for the selection of stimulus ensembles. However, due to the nature of decoding, manifested in Bayes' Formula, the reconstruction kernel depends on the stimulus ensemble $p(\mathbf{s}_t)$. This must be taken into account when testing the system with stationary time series: the reconstruction algorithm always reflects properties of both the system and the stimuli.⁸ Therefore, the choice of the ‘‘correct’’ stimulus ensemble is essential to a potential success of the stimulus reconstruction methods.

We will employ stimulus reconstruction methods in Chapter 5. Technical details can be found in Appendix D.

⁷Since equation (2.14) is a linear integral equation, an analytical solution for the linear reconstruction kernel can be found. If the stimulus has zero mean, the Fourier transform of the reconstruction kernel is given by $h_1(f) = S_{sr}(-f)/S_{rr}(f)$ where $S_{sr}(f)$ denotes the cross spectrum between stimulus and spike train and $S_{rr}(f)$ the power spectrum of the spike train (Gabbiani and Koch, 1998).

⁸Unlike fitting parameters to $p(\mathbf{r}_t | \mathbf{s}_t)$, cf. Section 2.1.6, changes of the model parameters with different stimulus ensembles do here not necessarily imply that the model is only locally valid.

2.4.2 Quantisation of Responses and Stimuli

A different approach seeks to obtain a direct estimate of the probabilities that govern the stimulus-response relation by reducing the effective dimensionality of the problem.

One possibility of reducing the size of input and output space is to group either responses or stimuli into a small number of classes. Here we will concentrate on forming such classes in the output space. One often uses clustering algorithms which group responses that are “close” to each other. Hence, these clustering algorithms require a distance measure on the set of responses. For spike trains, several distance measures have been proposed in the literature (Victor and Purpura, 1997; van Rossum, 2001; see also Appendix C). Using one of these distances, one can group the responses into N clusters R_i , with $i = 1 \dots N$. If N is not too large, one might be able to estimate the conditional probability $p(R_i|\mathbf{s}_t)$. Often, this is done on a selected set $\{S_j : j = 1 \dots M\}$ of stimuli only, yielding an estimate $p(R_i|S_j)$, see for instance Wiener and Richmond, 1999.

Ideally, grouping responses that are “close” to each other will form clusters that correspond to the different stimuli. A decoding procedure based on the assignment of the spike trains from one response cluster to one of the stimuli can therefore be used to estimate a lower bound of the system’s discrimination abilities. We will investigate the discrimination performance of auditory receptor neurons based on this method in Chapter 4.

An approach that we will use in Chapter 6 reduces the number of temporal dimensions by restricting the length of the stimuli. For instance, one can use “chunks” of stimuli, $\mathbf{s}_{t,t-n}$ and measure the output in the corresponding time window, $\mathbf{r}_{t,t-n}$. The input-output relation is then given by integrating over all possible histories such that

$$p(\mathbf{r}_{t,t-n}|\mathbf{s}_{t,t-n}) = \frac{1}{p(\mathbf{s}_{t,t-n})} \sum_{\mathbf{r}_{t-n-1}} \sum_{\mathbf{s}_{t-n-1}} p(\mathbf{r}_t|\mathbf{s}_t)p(\mathbf{s}_t), \quad (2.15)$$

where the summation variables \mathbf{r}_{t-n-1} and \mathbf{s}_{t-n-1} imply summing over the corresponding sets of responses and stimuli. In turn, Bayes’ Formula allows to estimate the response-conditional ensemble $p(\mathbf{s}_{t,t-n}|\mathbf{r}_{t,t-n})$. Note though, that this non-parametric approach does not seek a reduction of stimulus and output space that is in any sense optimised. Yet if one corrects for latencies between stimulus and response, and if the influences of the system’s memory are not too strong, then the reduced probability distribution can give a good approximation of the system’s input-output behaviour. Furthermore, when using stationary input traces, the probabilities $p(\mathbf{r}_{t,t-n}|\mathbf{s}_{t,t-n})$ can be estimated in overlapping time windows of size n (Strong et al., 1998).

These non-parametric methods do not impose any restrictions on the stimuli used to test the system and are therefore well suited to directly investigate the encoding of natural stimuli.

2.5 Assessing the Information Transfer

To estimate how much information about sensory stimuli the organism can possibly obtain from the responses of its neurons, one needs to compute the mutual information (Shannon and Weaver, 1949; Cover and Thomas, 1991). The mutual information I between any two random variables x and y can be calculated if both $p(x)$ and $p(y|x)$ are known,

$$I = \sum_x p(x) \sum_y p(y|x) \log_2 \frac{p(y|x)}{p(y)} \quad (2.16)$$

with $p(y) = \sum_x p(y|x)p(x)$. The mutual information is symmetric in the variables x and y which can be seen when rewriting it in terms of the joint probability distribution $p(x,y) = p(y|x)p(x)$,

$$I = \sum_x \sum_y p(x,y) \log_2 \frac{p(x,y)}{p(x)p(y)}. \quad (2.17)$$

The mutual information quantifies how many bits of information about the sensory stimulus are carried by the response. An information of 1 bit suffices to halve the uncertainty about the stimulus. If both $p(x)$ and $p(y|x)$ are Gaussian, one can simplify the formula in terms of covariance matrices; usually, however, the mutual information has to be directly estimated from the underlying probability distributions. Using one of the quantisation procedures described above, such direct estimates are quite feasible and can sometimes yield a good approximation of the true mutual information (Strong et al., 1998).

The mutual information depends not only on the transmission properties of the system under study but also on the input distribution $p(x)$. A high mutual information can be achieved only if the input distribution is well matched to the properties of the system. The maximum mutual information achievable by any distribution $p(x)$ defines the information capacity of the system. Hence, the mutual information allows to measure how well a particular stimulus ensemble is matched to a given system. Fortunately, any local maximum of the mutual information with respect to $p(x)$ is a global maximum (Cover and Thomas, 1991). We will exploit this fact in Chapter 6 when searching for an optimal stimulus ensemble $p(x)$.

As we assume the number of possible inputs and outputs to be finite, the mutual information is always bounded from above. For instance, if there are N possible output symbols, then I cannot exceed $\log_2 N$. A small value of the mutual information relative to this maximum can have three causes: (1) The stimulus ensemble used is not sufficiently matched to the properties of the input-output system. (2) The system's output is strongly influenced by unknown inputs. (3) The system is very noisy. Note that the difference between noise and unknown inputs is one of interpretation. We will come back to this issue in Section 2.6.2.

2.6 Fundamental Problems of Input-Output Analysis

Although the input-output analysis as introduced in this chapter provides a mathematical framework for the investigation of sensory systems, it does not make these investigations self-evident. Some of the most fundamental problems are the following: (1) Often, the probability distributions $p(\mathbf{s}_t)$ and $p(\mathbf{r}_t|\mathbf{s}_t)$ have a very complicated structure. (2) Biological systems are “open” systems, i.e., systems that are usually influenced in many ways; the input \mathbf{s}_t might be only one of them. Such unknown influences make it difficult to find a systematic relation between stimulus and response. (3) Feedback-loops can lead to systems that have a very long memory. (4) The amount of data that can be accumulated to fit a particular model is limited. This fact constrains the parametric size of the models that one can investigate.

2.6.1 Dimensionality Reduction

Essentially, one needs to find models that capture the complicated structure of the input-output relation without relying on innumerable (and, therefore, unmeasurable) parameters. Many problems that are seemingly hard to solve can be solved quite easily once they are transformed to a specific set of coordinates. Famous examples of the physical sciences include the two-body problem in classical mechanics and the hydrogen atom in quantum mechanics (Goldstein, 1980; Sakurai and Tuan, 1993). Likewise, modelling of the input-output relation of a sensory system can often be made much more simple by appropriate transformations of either the stimulus or the responses.

Transformations that reduce the dimensionality of stimulus or response space are of specific interest. Furthermore, such transformations might imitate what sensory systems do to solve their problems. This observation gives rise to the hope that by successively investigating a sensory system stage by stage, one might be able to expose the corresponding transformations and use these transformations in turn to investigate higher stages of the system. In other words, one has to find representations of the stimulus space that simplify the analysis of higher stages. For instance, if the orientation of a bar in a visual image is described by an angle, i.e., a single parameter, the input-output relation between an image and the response of a simple cell in the visual cortex is quite straightforward (Hubel, 1988). An equivalent but more complicated description characterises the stimulus by a vector whose components are the pixels of the visual image.

Although we will use simple transformations of sound pressure waves in Chapter 5 to simplify the stimulus-response relation, transformation and dimensionality reduction will not be an integral part of this thesis.

2.6.2 Unknown Inputs

At higher levels of the mammalian visual system, the neural responses depend not only on the visual image but also on the attentional state and the expectation of the organism (e.g., Moran and Desimone, 1985). These phenomena can be viewed as further external

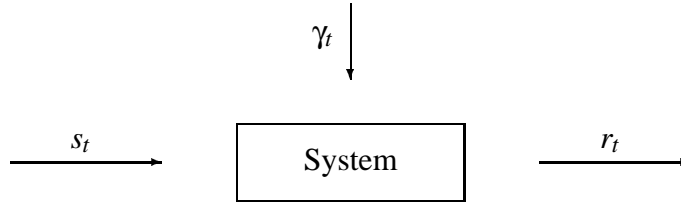


Figure 2.3: At time t , the input-output system now receives additional, unknown inputs γ_t .

inputs that act on our local system (cf. Figure 2.3). If these inputs are stationary, they do not violate our first condition. Nevertheless, it is useful to analyse their influence. For that purpose, we summarise their history by the variable $\Gamma_t = (\gamma_t, \gamma_{t-1}, \dots)$ where γ_t denotes the multi-variate input at time t . The input-output relation of our system is then given by summing over the unknown inputs,

$$p(\mathbf{r}_t | \mathbf{s}_t) = \sum_{\Gamma_t} p(\mathbf{r}_t | \mathbf{s}_t; \Gamma_t) p(\Gamma_t). \quad (2.18)$$

These perturbations become even more severe if the unknown inputs are not stationary. Under these circumstances, the assumption of stationarity is violated and the input-output relation of the system becomes explicitly time-dependent,

$$p(\mathbf{r}_t | \mathbf{s}_t; t) = \sum_{\Gamma_t} p(\mathbf{r}_t | \mathbf{s}_t; \Gamma_t) p(\Gamma_t | t) \quad (2.19)$$

which will result in an even less predictable response. However, if the unknown inputs change in a non-stationarity way on rather long time scales, then the stationarity assumption is approximately fulfilled and the input-output analysis will still give good results.

The effect of the unknown inputs will be to obscure the systematic relation between stimulus and response.⁹ In other words, the more unknown inputs there are, the smaller we expect the mutual information to be. In practice, these inputs can work in two opposing ways: (1) They can lead to a narrowing of $p(\mathbf{r}_t | \mathbf{s}_t)$, e.g., by inhibiting the system so that it does not respond at all. (2) They can lead to a broadening of $p(\mathbf{r}_t | \mathbf{s}_t)$, e.g., by "randomly" exciting the system which could result in a Poisson-like spontaneous activity. Note, that a small information transfer between input and output does not imply the existence of unknown inputs or internal noise. As noted above, it could simply mean that we are using the wrong stimulus ensemble. The auditory receptor neurons which we will investigate come comparably close to the maximum mutual information; we will take this as an indication that there are no unknown influences.

The problem of unknown influences is of course a generic problem of science (especially the life sciences) not restricted to the analysis of sensory systems. As long as the influences remain unknown, the only solution is to model the system with latent variable models such as Hidden-Markov Models (Gat et al., 1997; Sahani, 1999).

⁹We implicitly assume that the unknown inputs Γ_t are statistically independent of the input \mathbf{s}_t , i.e., $p(\mathbf{s}_t, \Gamma_t) = p(\mathbf{s}_t)p(\Gamma_t)$.

2.6.3 Long-Term Memory

For all practical purposes, systems that have a long memory pose an additional problem. It has been proposed, for instance, that horizontal and vertical motion detectors in the fly visual system form a network which has different stable states which in turn are modulated by the sensory input (Borst, 2001). Indeed, one finds that the horizontal motion detector H1 not only adapts to changes in contrast of the stimuli, but that the time scale of this contrast adaptation also adapts depending on how fast changes in contrast occur (de Ruyter van Steveninck, personal communication). In such cases, the system has a memory that extends on time scales of hours rather than minutes or seconds and the phenomenological models described in this chapter break down.

Sometimes, the long time scales are separable from the short time scales; for instance, one can often keep a system in a certain state of adaptation. This will be the case for auditory receptor neurons; hence we will neglect the influences of long-term memory.

2.6.4 Experimental Design

One issue of this thesis is the selection of stimulus ensembles to test a given sensory system. To make better use of the limited experimental time, the selection of stimuli can be updated during the run of the experiment. For that purpose, one can design experimental setups in which the data that have been gathered up to a certain time instant are analysed online and the results are then used to determine the subsequent stimulation of the system. Such experiments can help to minimise the time needed to fit a particular model or collect data for a particular question. Common approaches in the neurosciences have been the Alopex and Simplex algorithms that search for a stimulus that maximally drives the cell (Harth and Tzanakou, 1974; Nelken et al., 1994). An algorithm that searches for an ensemble of stimuli that maximises the information transfer between input and output will be presented in Chapter 6.

SUMMARY AND OUTLOOK:

The stimulus-response relation of a given system can be described by a probability distribution which determines the probability of a certain response for a given stimulus. As both stimulus space and response space are high-dimensional, one usually seeks to model only certain aspects of the system; within the Volterra-Wiener analysis, for instance, one models the relation between stimulus and average response. Furthermore, one often quantifies the stimulus-response relation on a very specific subset of stimuli only; a tuning curve describes the response as a function of a single stimulus parameter.

To understand the perception of an animal in its natural environment, one additionally needs to know the ensemble of natural stimuli. One can then attempt to decode a stimulus from the observed response. Depending on how one quantifies the success of decoding, one can formulate different decoding procedures. Stimulus reconstruction

techniques minimise the mean-square error between the actual and the reconstructed stimulus. Quantisation procedures seek to group spike trains so that, ideally, spike trains produced by the same stimulus fall into the same group. The mutual information quantifies how much information one can gain from a response about the stimulus. Given this theoretical background, we will now introduce the grasshopper auditory system.

Chapter 3

Acoustic Communication of Grasshoppers

Within this chapter, we will give a short introduction to acoustic communication of grasshoppers and analyse some of the statistical properties of grasshopper sounds. This shall aid us in understanding the specific task the grasshopper auditory system has to solve, both from a behavioural perspective and from the point of view of signal processing.¹

3.1 Behaviour

The hearing organs of grasshoppers have several functions (for review, see Hoy et al., 1999; Stumpner and Helversen, 2001). First of all, they serve to detect and localise potential mates. Additionally, they can help the animal to evade potential predators. In some species, acoustic communication also serves territorial defence. Here we will exclusively consider aspects of mate detection and selection.

The gomphocerine grasshoppers are remarkable for their bidirectional communication (Faber, 1953; von Helversen and von Helversen, 1994): while in many vertebrates and invertebrates only one sex sings to attract a partner, in many gomphocerine grasshoppers both sexes participate acoustically in the mate finding process. When searching for a female, a male repeatedly produces a species-specific sound, the “calling song”. Upon hearing this song, a nearby female, if ready to mate, will usually answer with a “response song”. In turn, the response allows the male to localise the female and move towards her. This communication process is iterated; as the localisation is not exact, the male approaches the stationary female in a zig-zag way.

According to Shannon (Shannon and Weaver, 1949), any communication process can be subdivided into a sender, a communication channel, and a receiver. As sound propagates through the environment, the communication channel is given by a grasshop-

¹The calling songs of *Ch. biguttulus* males analysed throughout this chapter were kindly provided by Dagmar and Otto von Helversen. I am grateful to Jan Benda and Hartmut Schütze for assistance in recording calling songs from *Ch. brunneus* and *Ch. mollis*.

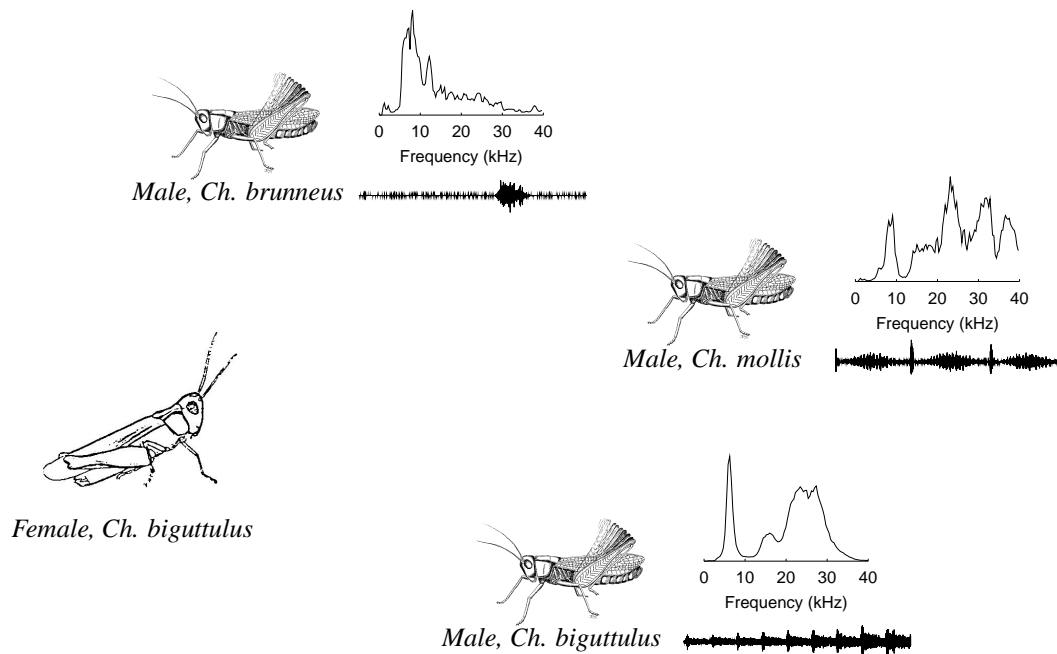


Figure 3.1: Mate detection and selection. To avoid cross-fertilisation, a female grasshopper (left) needs to identify conspecific males and should ignore heterospecific males. The identification relies almost exclusively on the calling songs of the male grasshoppers. For three exemplary males, we show a sound pressure wave of the calling song and the respective frequency content of the carrier waves. The carrier exhibit minor differences in their spectral content which generally ranges from 5–40 kHz. The behavioural relevant features of the songs, however, are the distinct rhythmic modulation patterns which are visible in the sound pressure waves. (Grasshopper drawings adapted from Helversen and Helversen, 1994; Benda, 2001)

per's natural habitat. As both male and female can be sender and receiver, their roles alternate. Subsequently, we will exclusively consider that part of the communication process where the male is the sender and the female the receiver.

3.1.1 Mate Detection and Selection

Often, many grasshopper species live in the same habitat; quite common within central Europe are the closely related species *Chorthippus biguttulus*, *Chorthippus mollis*, and *Chorthippus brunneus* (Bellmann, 1985). These grasshoppers are so similar to each other that they have long been thought to constitute a single species until differences in their calling songs were noticed. Also for female grasshoppers, the calling songs provide the primary clue to identify a conspecific male (von Helversen, 1972).

When ready to mate, a female will usually listen to the calling songs of different males from several species. This situation is displayed in Figure 3.1 for a female grasshopper of the species *Ch. biguttulus*. To avoid potential cross-fertilisation, the female should respond only to the songs of *Ch. biguttulus* males. A female that mates

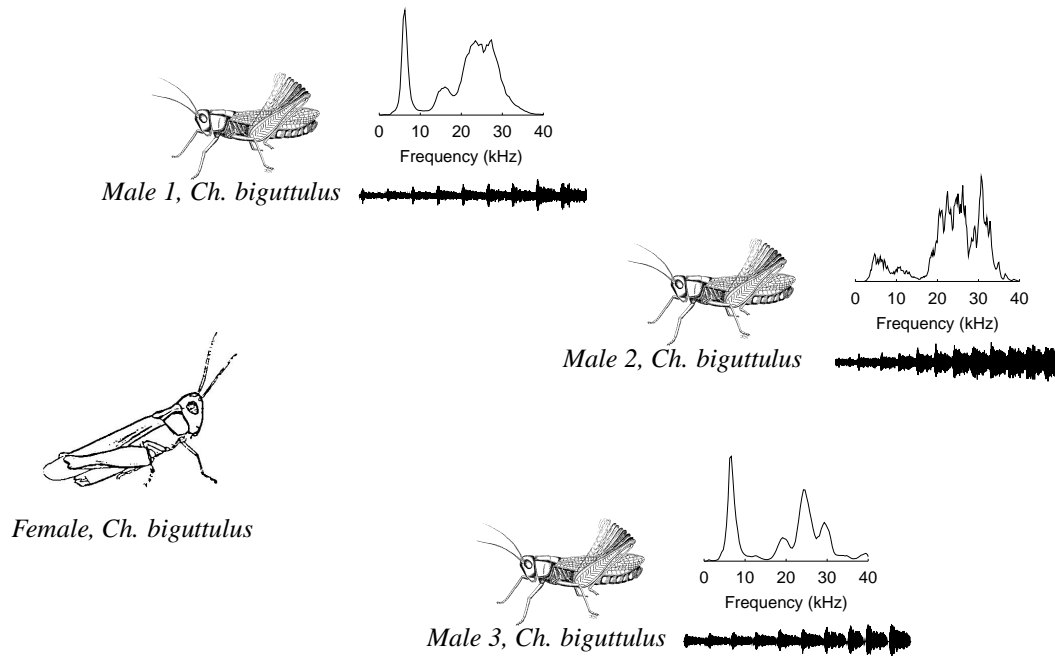


Figure 3.2: Sexual selection. To insure its reproductive success, a female grasshopper (left) will seek to mate only males with a high genetic quality. Such a decision will most likely be based on differences between the calling songs of conspecific males. The sound pressure waves and the frequency content of the respective carrier waves for three exemplary males are shown on the right.

with a male from a different species produces hybrid offspring which is either infertile or not at all viable. As female grasshoppers lay only about ten eggs every 2–3 days, cross-fertilisation significantly decreases the reproductive success of the female. The most important task for the female is, therefore, to identify the songs of conspecific males and reject all other songs.

3.1.2 Sexual Selection

Male grasshoppers invest far less in their gametes than females. Moreover, males are able to fertilise several females during a single day. The asymmetry in gamete production leads to an asymmetry in mating strategies between the sexes. To increase their reproductive success, males will attempt to mate as many females as possible while females will seek to mate only males with a high genetic quality. The asymmetry in mating strategies, however, is not restricted to grasshoppers. Throughout the animate world, the sex that invests more into the offspring is also the one that is more selective (Anderson, 1994). This principle, called sexual selection, has been recognised by Darwin to be the second force of evolution besides natural selection (Darwin, 1871).

The difference in mating strategies leads to a divergence of interests between the two sexes. To avoid mating a less desirable male, a female will try to detect any clues in the signal that reveal the male's true state. In turn, to increase the frequency of

populations, a male grasshopper will try to disguise any signal traits that might betray his physical defects. Communication systems, therefore, tend to evolve towards signals that are bluff-resistant, i.e., that cannot easily be feigned by less fit males (Anderson, 1994). On the same grounds, these signal traits need not necessarily be obvious to external observers; instead, to us they could turn out to be only minor differences.

As acoustic communication in grasshoppers serves the mate finding process, the evolution of the acoustic communication systems is largely driven by these selective pressures (von Helversen and von Helversen, 1994; Hauser, 1996). A female's auditory system should therefore be able to recognise even minute deviations in the signals of conspecific males. Behavioural experiments have shown that female grasshoppers of the species *Ch. biguttulus* are indeed able to do so: they do usually not answer the songs of male grasshoppers that have autotomised one of their hindlegs (von Helversen, 1972) and thus prevent mating with such males (Kriegbaum, 1989).

3.2 The Sender

Grasshoppers generate calling songs by rasping their hindlegs across their forewings in a species-specific rhythm. The inner surface of the hindleg's upper segment (femur) is furnished with a row of small teeth; if these are scrubbed against the nervures of the forewing, the latter vibrates with frequencies in the range of 4–40 kHz. The amplitude of these broad-band vibrations grows with the force of the stridulation. The vibrations cease almost immediately if the movement stops.

In *Ch. biguttulus* males, the rhythmic structure of the sound arises in the following way (Elsner, 1974): during the basic movement (syllable), each hindleg moves up and down three to four times. After a short rest (pause), the cycle starts from the beginning and is repeated up to 50 times. Depending on the individual animal and the ambient temperature, the combination of syllable and pause lasts for about 60–140 ms.

Usually, the movement patterns of both hindlegs differ a little; moreover, the movement is not synchronous but slightly out of phase. These phase shifts obviate the cessation of sound that results if one of the hindlegs reverses its movement. Male grasshoppers that have lost or damaged a hindleg or forewing can no longer avoid these instants of silence: their songs exhibit short gaps (1–2 ms) at the inevitable movement reversal point of the remaining hindleg.

The sound pressure wave that results from the stridulation pattern constitutes the calling song. The song, therefore, consists of broad-band carrier frequencies whose amplitude is rhythmically modulated. Three such sound pressure waves are shown in Figure 3.2, the corresponding spectrograms are displayed in Figure 3.3. After a short examination of the carrier frequencies, we will focus on statistical properties of the amplitude modulation signal (AM signal).

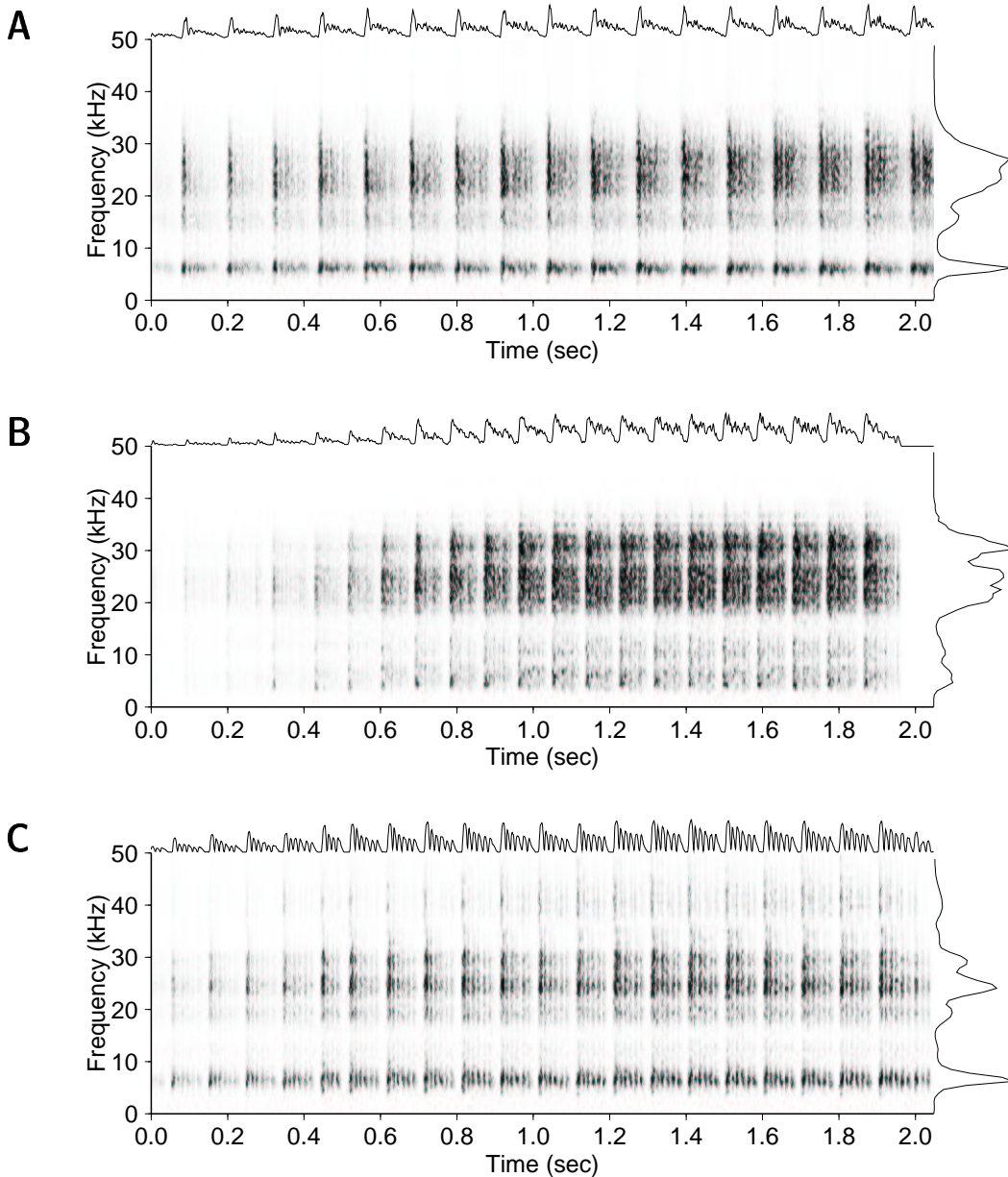


Figure 3.3: Spectrograms of the three calling songs of *Ch. biguttulus* shown in Figure 3.2. **A** The song of this intact male grasshopper features a syllable-pause length of about 120 ms and a broad-band carrier that is separated into a low-frequency component around 4–8 kHz and a high-frequency component around 10–35 kHz. On top of each spectrogram, the song’s amplitude modulation (AM signal) is shown; to the right, the song’s power spectral density, compare Figure 3.2. **B** This panel depicts the song of another intact male. The song has a shorter syllable-pause length of 90 ms and more power in the high-frequency band, here 15–35 kHz. **C** The third panel shows the song of a grasshopper that has lost one hindleg. The stridulation of the remaining hindleg lays bare the many short gaps within each syllable that originate at the hindlegs’ movement reversal points.

3.2.1 Carrier Frequencies

For *Ch. biguttulus* males, we will distinguish two frequency regimes: a low-frequency regime (4–8 kHz) and a high frequency regime (10–40 kHz). Both regimes capture roughly the main frequencies created by the wing vibrations, as can be seen in Figure 3.3. Other species have generally a slightly different composition of frequencies (Meyer and Elsner, 1996) although there are large overlaps, cf. Figure 3.1. Whether the precise spectral content of a song has any behavioural relevance remains unknown. While the carrier frequencies do not aid in species discrimination, they could still relate to certain properties of the wings and thus potentially be informative about male fitness.

In most songs, the spectral content of the carrier remains roughly constant over the full duration of the song. The spectrograms show that the songs have no significant frequency modulation. We will neglect slight differences between the temporal structure of the low- and high-frequency bands. However, the spectral content of the carrier does differ between different male individuals of the same species: while the song displayed in Figure 3.3B has most of its power in the high-frequency band, a significant portion of the power of the songs in Figure 3.3A and 3.3C lies in the low-frequency regime.

3.2.2 Amplitude Modulations

A song usually lasts 2–4 seconds and consists of several repetitions of syllable-pause combinations each of which is produced by up- and down-strokes of the hindlegs (Elsner, 1974). This stridulation pattern results in a distinct amplitude modulation (AM signal) of the broad-band carrier frequencies. To extract this AM signal from the original sound pressure wave, one can employ the Hilbert transform (Haykin, 1994). The resulting estimates of the AM signal, smoothed with a 0–200 Hz low-pass filter, are shown as traces above the spectrograms (cf. Figure 3.3). The syllables themselves exhibit an individually characteristic amplitude modulation resulting from the up and down movements of the hindlegs. This difference of the intra-syllable structure is visible in the left column of Figure 3.4.

To analyse the statistical properties of the AM signal, we estimated the power spectral density of the AM signal, as displayed in the right column of Figure 3.4. While the first peak in each spectrum corresponds to the length of a syllable-pause duration, the other peaks, also called harmonics, describe the spectral content of the AM signal within the syllables. A song of a male that has lost a hindleg is shown in Figure 3.3C and 3.4C; whenever the remaining hindleg reverses its movement, a short yet pronounced gap (1–2 milliseconds) appears within the syllable. In turn, each combination of a gap and subsequent sound leads to a strong modulatory component of around 70 Hz in the power spectrum (cf. Figure 3.4C, right panel). This component can sometimes be seen in intact males if one of the hindlegs is weaker as visible in Figure 3.4B. In all of the twelve songs analysed, no significant power of the modulation frequencies beyond 100 Hz could be found.

The pauses within the songs manifest themselves in a highly non-Gaussian distribution of modulation amplitudes: neglecting the slow rising phase of each song, one

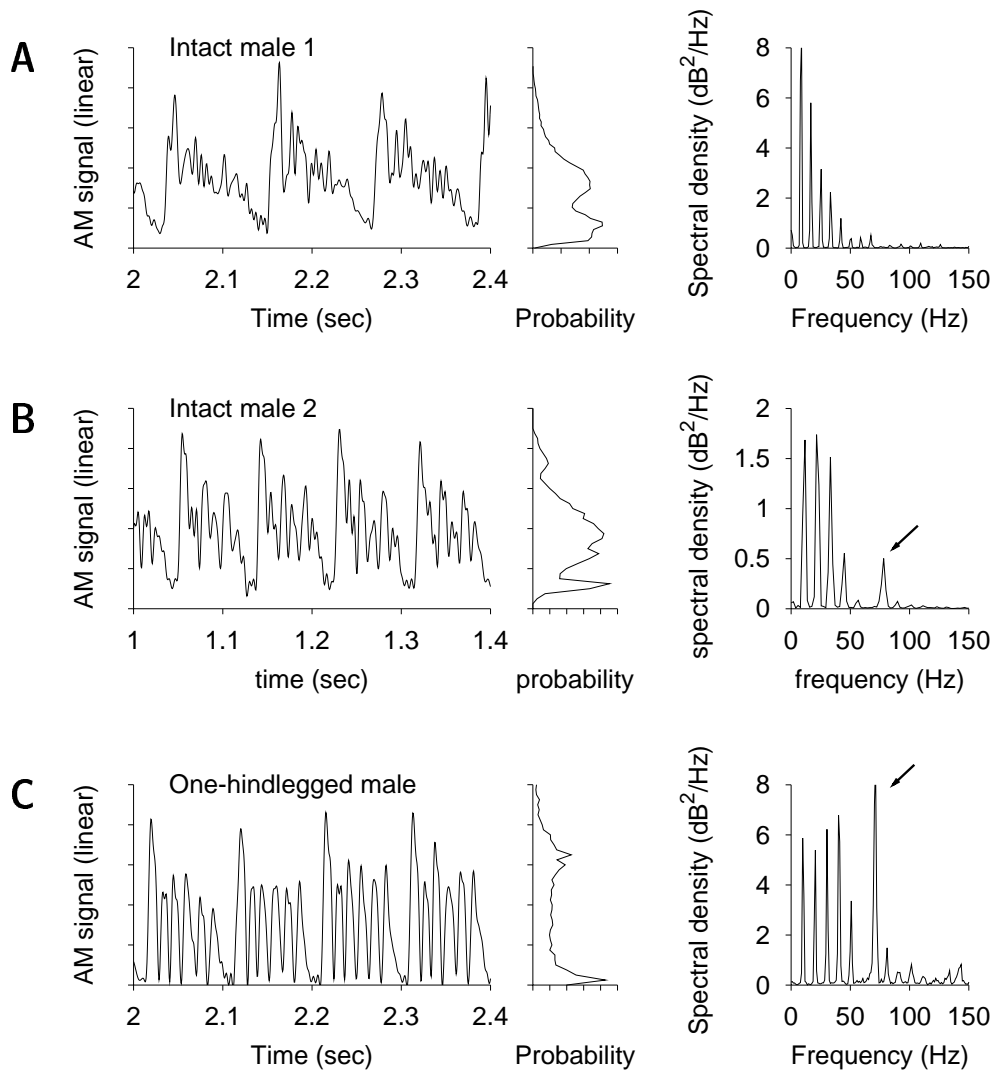


Figure 3.4: Statistical properties of the calling songs. Shown are the AM signal of a short song section (left), the distribution of modulation amplitudes (middle), and the power spectral density of the AM signal (right). Because of their syllable-pause structure, all calling songs display a double-peak distribution of modulation amplitudes. The duration of syllable and pause becomes evident in the first peak of the power spectral density. **A** An intact male with two hindlegs in good order. **B** An intact male with two hindlegs one of which is weaker. **C** A male that has lost one hindleg. This loss leads to an additional peak around 70–80 Hz in the power spectrum (arrow), indicating a strong modulation within each syllable. The strong modulation of the AM signal can also be seen on a less pronounced basis if one of the hindlegs is weaker as in B. (A and C adapted from Machens et al., 2001c)

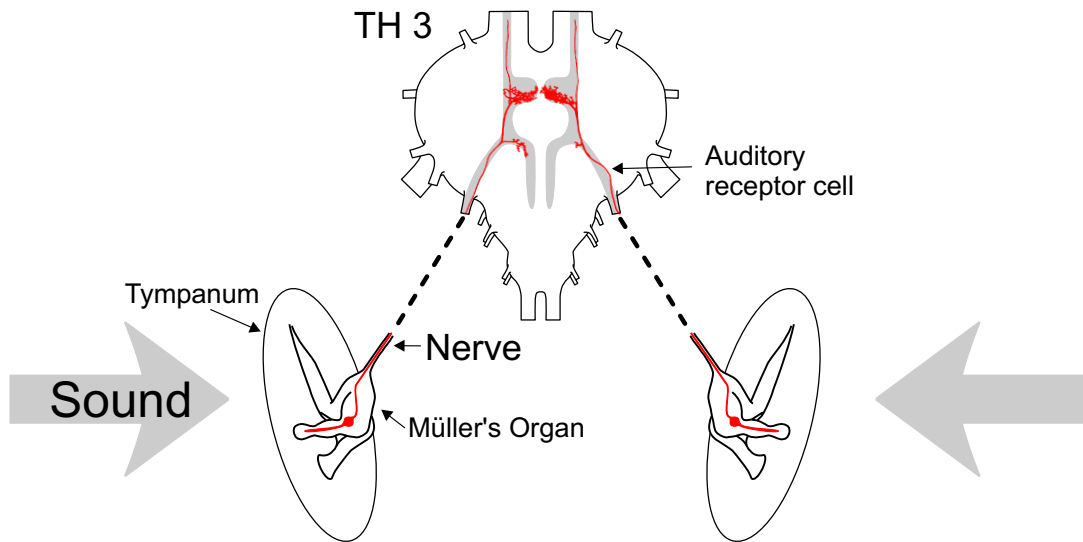


Figure 3.5: The auditory system of grasshoppers and locusts. Sound falls on the two tympana where the receptor neurons translate the sound into electrical activity which is forwarded into the metathoracic ganglion (TH 3). (Figure designed by Hartmut Schütze, Humboldt-University Berlin)

finds that the amplitude distribution always displays a double-peak structure with one peak at low amplitudes for the pauses between distinct syllables and one peak at higher amplitudes for the syllable segments (Figure 3.4, centre column). The low amplitude peak is centred away from zero because the pauses do not consist of silence but of relative quiet. Apart from the characteristic composition of modulation frequencies, the rhythmic structure of the calling songs thus also causes a large depth in the amplitude modulations.

3.3 The Communication Channel

The calling songs are propagated through the air before they reach a receiver. During the propagation, the signal is disturbed through physical properties of the natural environment. These disturbances include reverberations, scattering of sounds, and frequency-dependent excess attenuation (Römer, 1998). Excess attenuation can lead to losses of up to 18 dB/m close to the ground, thus severely limiting the communication distances to at most a few meters (Lang, 1999). As high-frequency components are affected more strongly by excess attenuation, one might hypothesise that male songs with weak low-frequency components (cf. Figure 3.3B) are less successful in attracting a female response.

The propagation of a signal is additionally disturbed by the songs of other grasshoppers. Altogether, these disturbances tend to fill the pauses in the song of a specific grasshopper, resulting in an effective decrease of the modulation depth of the songs. We will investigate the importance of this phenomenon in greater detail in Chapter 5.

3.4 The Receiver

On the receiver side, the songs are encoded by roughly 40–50 auditory receptor neurons per ear (Pauls et al., 2001). On each side, the receptor cells are located within the Müller's Organ (Gray, 1960) that is attached to the tympanum of the respective ear; their axons extend through the tympanal nerves to the metathoracic ganglion where auditory information is processed by local interneurons and then sent to the brain via ascending neurons, cf. Figure 3.5. The activity of the receptor neurons is usually recorded in the tympanal nerve.

3.4.1 Auditory Receptor Neurons

Comparative studies have shown that the auditory periphery of different acridid grasshopper species, including locusts, is strikingly similar in both anatomy and physiology. The physical properties of the ears and sound intensity threshold curves are highly conserved (Meyer and Elsner, 1996) and responses of receptor cells of locusts and *Ch. biguttulus* are almost indistinguishable (Stumpner et al., 1990; Ronacher and Krahe, 2000). Remarkably, the conservation of the systems seems to hold even at the interneuron level (Ronacher and Stumpner, 1988). As it is better suited for demanding electrophysiological experiments, we will therefore adapt *Locusta migratoria* as a model system for *Ch. biguttulus*.

There are three types of receptor neurons, attached to different parts of the membrane and sensitive to different frequency ranges. Two of these types, called low-frequency receptors, are most sensitive for frequencies in the range 4–8 kHz, the third (high-frequency receptor) in the range 10–25 kHz (Römer, 1976; Jacobs et al., 1999). The frequency sensitivity of the receptor cells is therefore well matched to the frequency content of the songs.

The basic properties of receptor neurons have been demonstrated by testing them with sine wave stimuli (Römer, 1976; Stumpner and Ronacher, 1991; Benda, 2001). At a given intensity, this corresponds to tracing the system's input-output relation on a one-dimensional manifold in stimulus space, as described in Section 2.3.2. Neglecting influences of the system's memory (such as adaptation) and using sine waves of 50 ms length separated by longer pauses (≈ 500 ms) yields tuning curves as shown in Figure 3.6A for three different intensities. The shape of the curves indicates that the cell is tuned to a frequency of around 4 kHz, also termed best frequency. In a second step, we fix the frequencies and change the intensity, thus cutting another one-dimensional manifold through stimulus space. The resulting curves are shown in Figure 3.6B and exhibit a typical sigmoid shape. Combining the data from both tests allows us to plot isoresponse curves on the two-dimensional manifold spanned by frequency and intensity. Figure 3.6 shows the respective curve for the onset response, yielding the threshold curve of the cell.

To obtain a fuller picture of the system's input-output relation, one needs to enlarge the view of stimulus space. A model that applies to the complete stimulus space while keeping a one-dimensional output, the average firing rate, has been successfully devel-

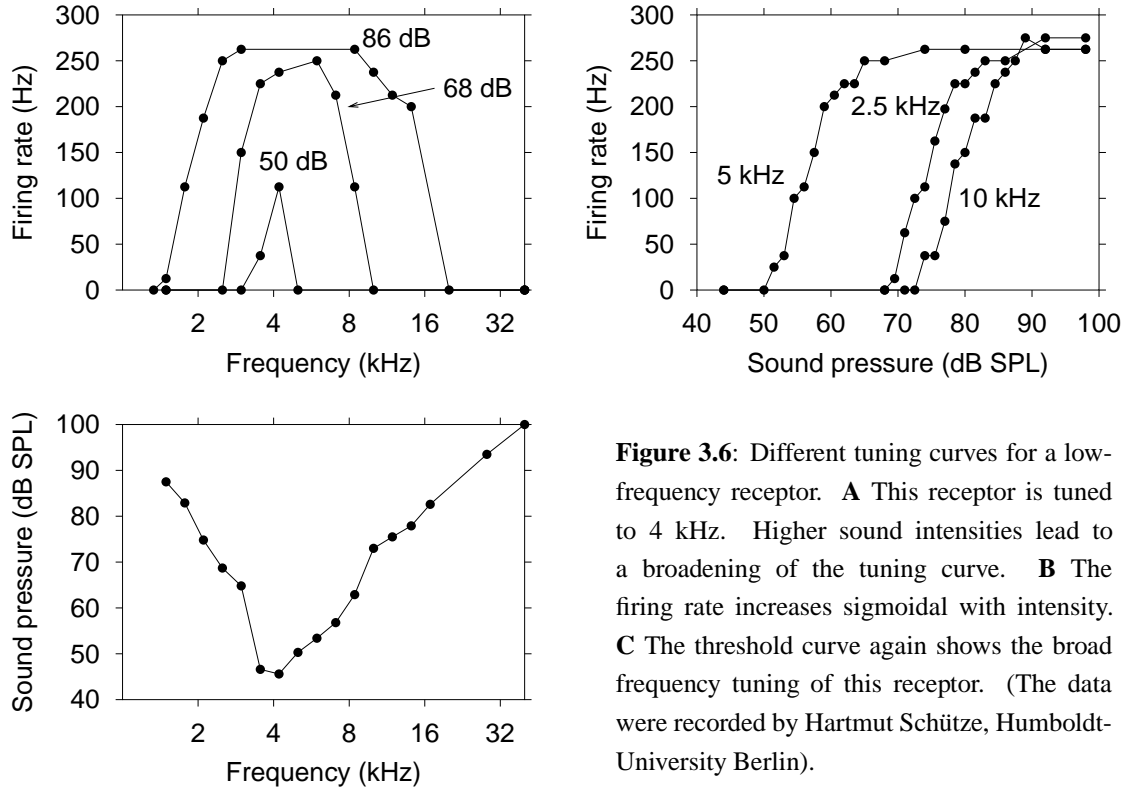


Figure 3.6: Different tuning curves for a low-frequency receptor. **A** This receptor is tuned to 4 kHz. Higher sound intensities lead to a broadening of the tuning curve. **B** The firing rate increases sigmoidal with intensity. **C** The threshold curve again shows the broad frequency tuning of this receptor. (The data were recorded by Hartmut Schütze, Humboldt-University Berlin).

oped by Gollisch et al. (2001). According to this model, the receptor neurons perform a linear filtering of a sound stimulus $s(t)$,

$$u(t) = \int d\tau T(\tau) s(t - \tau) \quad (3.1)$$

where the integral kernel $T(\tau)$ denotes the system's transfer function and $u(t)$ corresponds to the filtered sound pressure wave. The power of the frequency-filtered stimulus is then averaged over time, $I = \overline{|u(t)|^2}$, and the average firing rate is obtained from I via some sigmoid function, e.g.,

$$\bar{r} = r_{\max} [1 + \tanh(I - I_{\theta})] \quad (3.2)$$

where r_{\max} denotes the maximum response and I_{θ} the threshold of the system.

These models do not yet take into account stimuli that contain time-varying amplitude modulations. They also neglect the stochasticity of the spiking process. The results suggest, however, that receptor neurons use their firing rates to encode the amplitude modulation (or power) within a certain frequency band. We will use this as a working assumption in Chapter 5 and therefore neglect the details of the frequency composition of the carrier.

SUMMARY AND OUTLOOK:

Mate finding of grasshoppers primarily relies on acoustic communication. Female

grasshoppers identify the calling songs of conspecific males and choose among these males based on specific features of the songs. Most likely, this is done to avoid genetically less desirable males. The calling songs consist of broad-band carrier frequencies modulated into a repetitive series of basic patterns, termed syllables. The modulation pattern within a syllable conveys behaviourally important information down to the millisecond time scale. The amplitude modulation of the songs is encoded into a firing rate by auditory receptor neurons, the first stage of the grasshopper's auditory system. In the following two chapters we will investigate what and how much information a female can draw from its receptor neurons.

Chapter 4

Decoding Natural Stimuli

In this chapter, we start our investigation of the auditory periphery using a small set of natural stimuli, the calling songs of eight individual males of the species *Ch. biguttulus*. In particular, we investigate to what extent variations between these songs can be faithfully encoded by single auditory receptor neurons. The resolution of sound patterns achieved by the receptor neurons limits a female's ability to select between different male individuals and is, therefore, highly relevant to behaviour.¹

4.1 Introduction

4.1.1 Questions

Female grasshoppers have a particular interest in mating with conspecific males that have a high genetic quality (cf. Section 3.1.2). Behavioural data suggest that the choice of a specific male is mostly based on his calling song (von Helversen, 1972; Kriegbaum, 1989). In order to choose a particular male, a female must recognise the deviations in the signals of conspecific males. The detection of these differences is the task of the female's auditory system. However, the stochastic nature of neural responses tends to mask small signal differences. The ability of a female grasshopper to discern differences in the male songs thus depends both on the properties of these songs and the limits imposed by the response variability of auditory neurons.

The information contained in a calling song has to be represented at the level of the auditory periphery (cf. Section 3.4.1). To investigate some of the neural limits to sexual selection, we therefore explored whether single auditory receptors supply a female grasshopper with sufficient information to consistently rate mating songs. In particular, we asked whether the information available at the level of receptor neurons allows one

¹This chapter is in part based on Machens et al. (2001a, 2001b). I am especially grateful to Hartmut Schütze (Humboldt-University at Berlin) who performed the electrophysiological experiments and to Martin Stemmler (Humboldt-University at Berlin) who developed the procedures necessary for the stimulus design. The grasshopper songs were kindly provided by Dagmar and Otto von Helversen (University of Erlangen, Germany).

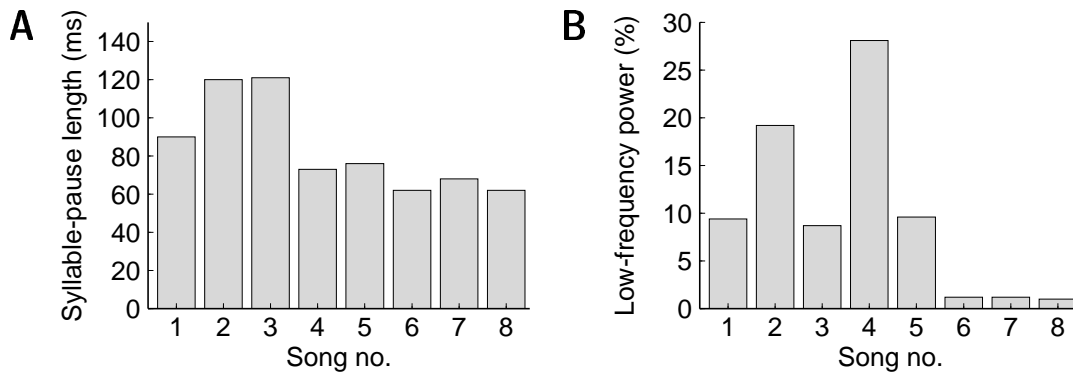


Figure 4.1: Song features. Altogether eight songs were employed in the discrimination analysis. **A** The syllable-pause durations of the eight songs vary between 60–120 ms. **B** While all songs have roughly the same overall power, different fractions of that power fall below 10 kHz. In particular, songs no. 6–8 contain only about 1% of their power in the low-frequency regime.

to discriminate the songs. If this were the case, then females would definitely have the capability to evaluate individual songs.

4.1.2 Stimulus Design

To investigate the discrimination performance, we took a random sample of eight songs, recorded from eight male individuals of the species *Ch. biguttulus*. All males were still endowed with both of their hindlegs. The signal differences apparent in the songs of one-hindlegged grasshoppers, cf. Figures 3.3C and 3.4C are therefore excluded, making the discrimination task more difficult.

As noted in Section 3.2, the songs differ in several aspects. One aspect concerns the length of each syllable-pause combination. As can be seen in Figure 4.1A, these durations vary between 60–120 ms for the songs chosen. However, the syllable-pause duration is unlikely to convey any information about male viability, as syllable-pause duration is greatly affected by body temperature. In fact, this duration can vary more than two-fold when the male moves from sun to shadow (von Helversen and von Helversen, 1987). Hence, a song evaluation based on syllable-pause duration alone is probably of limited use under natural conditions. These considerations suggest that female grasshoppers should be able to choose among males even if their songs have the same syllable-pause duration. To test this possibility, we generated a second set of songs by rescaling the original songs to a common syllable-pause length of 100 ms while maintaining the original spectra, cf. Figure 4.2A–C. The technical procedure for the rescaling operation is described in Appendix A.

A second aspect in which the songs differ are the carrier frequencies. In particular, the relative power of the high frequency band (10–40 kHz) versus the low-frequency band (4–10 kHz) varies considerably between the eight songs as can be inferred from Figure 4.1B. As described in Section 3.4.1, a single receptor neuron encodes either

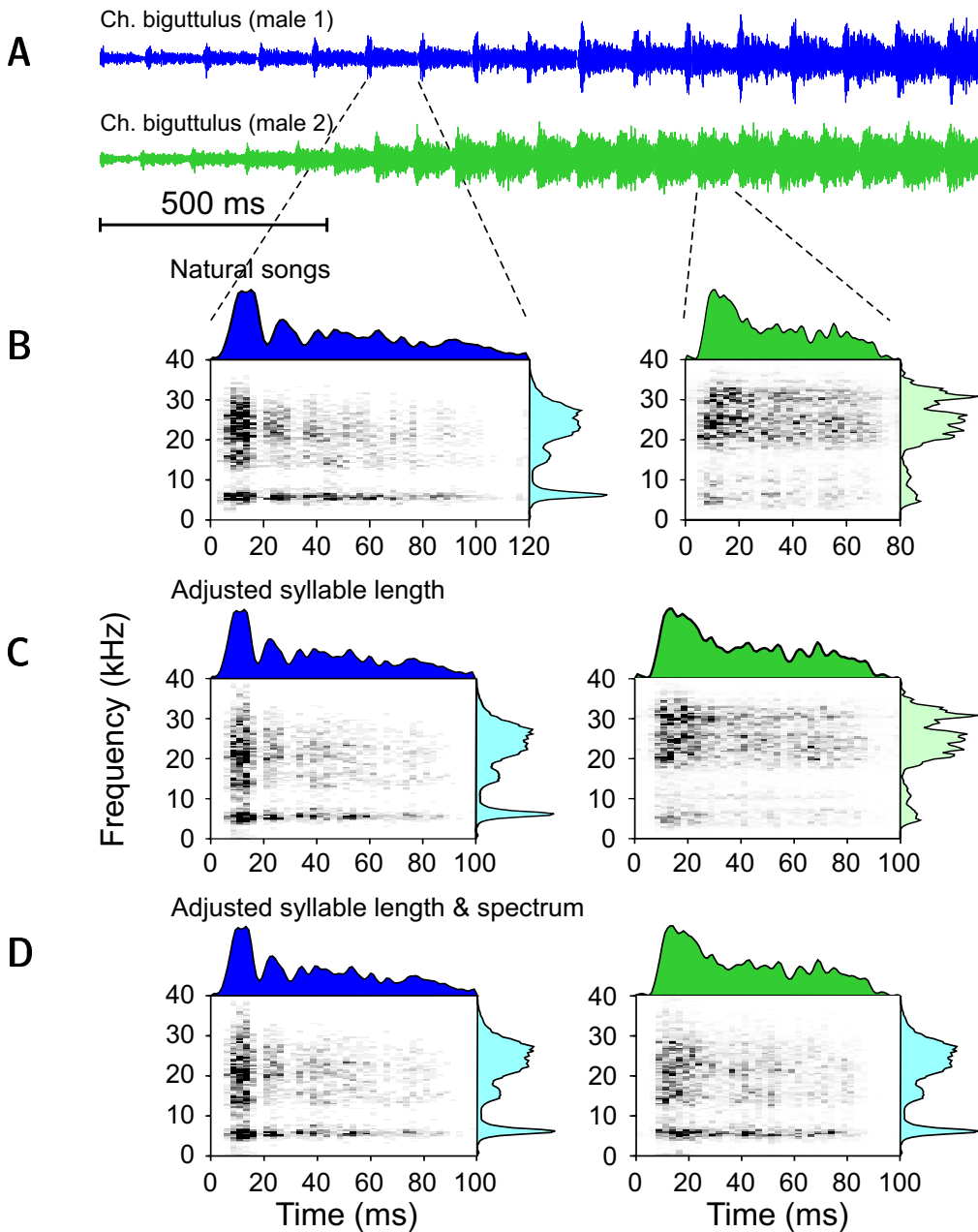


Figure 4.2: Stimulus design. **A** Sound pressure waves of two sample songs. **B** The differences between the songs are illustrated in the two spectrograms that highlight a single syllable. The spectrogram shows both the amplitude modulation and the spectral content of the syllable. **C** A second set of songs rescales the syllable-pause duration to 100 ms while keeping the spectral content of each song. **D** A third set of songs additionally equalises the spectral content of all songs. During the last second of the songs, the remaining differences in the mean intensities are as little as 0.1 dB. (Figure adapted from Machens et al., 2001b)

the low-frequency or the high-frequency band. The differences in the power of the frequency bands therefore translate directly into differences of the firing rates of a receptor. To preclude a discrimination based on differences in the average firing rate, we generated a third set of songs by additionally equalising the carrier frequencies (Fig. 4.2D). With equal syllable-pause duration and equal carrier spectra, the eight songs in this set retain differences only in the precise amplitude modulation of each syllable.

Altogether, we thus tested the system with three sets of eight songs each. The songs were repeatedly presented to a female while recording the responses of single auditory receptors; details of the experimental protocol can be found in Appendix B.

4.1.3 Discrimination Analysis

Within the space of all possible acoustic stimuli of a certain length and discretisation, the eight songs correspond to eight points only. However, these eight points are highly relevant as they belong to the small subset of natural stimuli. To obtain a useful measure of the discrimination ability, we need to quantise the response set, i.e., the set of all possible spike trains, into eight corresponding subsets (cf. Section 2.4.2). Ideally, spike trains evoked by the same stimulus should always fall into the same subset; in spite of response variability, one would then be able to perfectly discriminate the stimuli given the spike trains.

Usually, spike trains elicited from repeated presentations of the same song will be more similar to each other than spike trains elicited by different songs. This observation suggests a quantisation procedure based on spike train similarity. Such a similarity can be measured by defining a distance on the set of all spike trains. Two such distance measures which account for differences in spike count as well as in spike timing are described in detail in Appendix C. The Victor-Purpura distance seeks to align two spike trains by a set of elementary steps that are associated with a cost: inserting or deleting a spike costs one unit; shifting a single spike by Δt seconds cost $q\Delta t$ units. The parameter q weights the temporal resolution on which the spike train conveys information. The “cheapest” composition of steps defines the distance (Victor and Purpura, 1997). Within the van Rossum Distance, each spike is replaced with an EPSP-like response function whose width τ represents the temporal resolution of the spike train; the root-mean-square difference between the resulting EPSP-filtered traces provides the distance (van Rossum, 2001). Appendix C shows that both distances are closely related and that the resolution parameter τ approximately corresponds to the reciprocal of the cost parameter q .

Once a distance on the set of all possible spike trains is defined, there exist many methods that allow to quantise the response set for a given pool of data. Here we divide the set of all spike trains into eight clusters by randomly picking a template spike train from each song; each spike train then belongs to the cluster with the nearest template (cf. Appendix C). This procedure completes the quantisation of the response space and allows to estimate the conditional probability distribution $p(R_i|S_j)$ where S_j , with $j = 1 \dots 8$, denotes the songs, corresponding to points in the stimulus space,

and R_i , with $i = 1 \dots 8$, denotes the response clusters, corresponding to subsets of the response space. Hence, $p(R_i|S_j)$ is the probability that a spike train elicited from the song number j falls into the response cluster of song number i . Since the eight songs form a representative sample of natural songs, we will (implicitly) use a uniform prior distribution, $p(S_j) = 1/8$.

4.2 Discrimination Performance

The length of the original songs varies between 1.9 and 3.7 seconds. However, it might not be necessary to use all of this time for the discrimination task. To analyse how the discrimination capability unfolds in time, we therefore took only the first T milliseconds of each spike train into account. In the following, we will focus on the van Rossum Distance unless mentioned otherwise. Hence, we are left with two parameters: the parameter τ , weighting the temporal resolution of the information conveyed by the spikes; and the parameter T , giving the window length of the spike trains.

To gain some understanding of the discrimination procedure, we will first focus on a sample experiment. In this experiment, the original eight songs were played 15 times to a female grasshopper while recording the spike trains of single auditory receptor neurons. The average firing rates varied between 95 Hz for song no. 6 and 175 Hz for song no. 4.

4.2.1 Resolution Parameter

The role of the parameter τ can be best understood by first fixing the window length. Here we chose to use only the first $T = 500$ ms of every spike train. The distributions $p(R_i|S_j)$, resulting from the discrimination analysis, are shown in matrix form in Figure 4.3A for six values of τ . Columns and rows of these matrices denote the consecutively numbered songs and the respective response clusters. The diagonal elements correspond to the probability that a spike train elicited by song number i falls into the response cluster of song number i . The probability of correct discrimination can therefore be read directly from the diagonal elements:² the percentage of correctly assigned spike trains is given by averaging over the diagonal elements, $\gamma = \sum_{i=1}^8 p(R_i|S_i)p(S_i) = 1/8 \sum_{i=1}^8 p(R_i|S_i)$.

When largely neglecting temporal information ($\tau \rightarrow \infty$), the songs cannot be discriminated as indicated from the many off-diagonal entries in the corresponding matrix. The matrix approximately decomposes into two blocks, signifying that the songs fall into two groups: a group consisting of the songs no. 1–5 and a group consisting of the songs no. 6–8. For very large τ , the spike train distance measures only the spike count difference (see Appendix C). The average firing rate and, therefore, the spike count,

²Alternatively, one can measure the discrimination performance by the mutual information between original and assigned songs, cf. Section 2.5. This is the measure that was used in Machens et al. (2001a). However, both mutual information and percentage correct yield qualitatively the same results; for simplicity, we use the latter measure here.

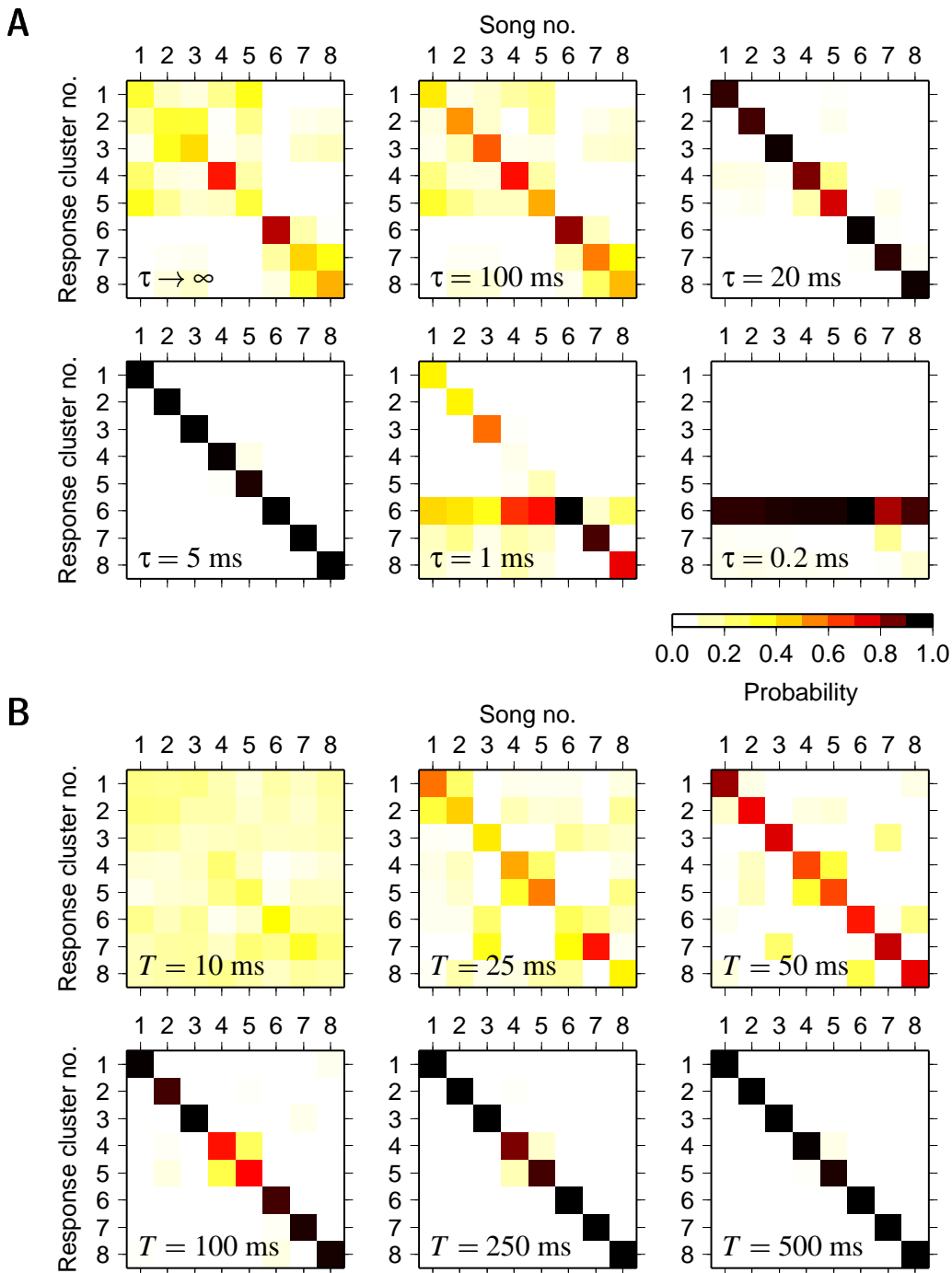


Figure 4.3: The conditional probabilities $p(R_i|S_j)$ shown as classification matrices. The horizontal axis denotes the song no. j , the vertical axis the response cluster of song no. i . The diagonal elements of the matrices correspond to correctly classified, the off-diagonal elements to misclassified spike trains. **A** Taking only the first $T = 500$ ms of every spike train into account, these matrices were computed for different values of the parameter τ which measures the time scale on which information relevant to the discrimination task is conveyed. **B** Matrices computed at a resolution $\tau = 5$ ms for different window length T . After about 250 ms, the probability of correct classification exceeds 95%.

depend on the power of the low-frequency component of a stimulus. Hence, songs with large differences in the power, and thus the firing rate, can be discriminated while songs with small differences in power tend to be confused. Consequently, the grouping reflects the differences in the power of the low-frequency component already shown in Figure 4.1B.

For intermediate resolutions, $\tau = 5$ ms, all spike trains are correctly assigned to the songs that evoked them. Information in the spike trains varying on a scale of about 5 ms seems to be well suited for song discrimination. This time scale reflects both the unique qualities of each song and the encoding mechanism of the auditory receptor neurons.

When insisting on a very high temporal resolution ($\tau = 0.2$ ms), the discrimination ability decreases again. In this case, the distance between two spike trains is given by the sum of the spike counts minus two times the coincident spikes (see Appendix C). If the number of coincident spikes is negligible, then every spike train is closest to the spike train with the fewest spikes. For this experiment, song no. 6 elicits the spike trains with the smallest firing rates. Consequently, the response cluster of song no. 6 becomes so large that it effectively incorporates all spike trains, leading to the probability distribution $p(R_i|S_6) = 1$ for all i and $p(R_i|S_j) = 0$ for all i and $j \neq 6$, as shown in the lower right matrix of Figure 4.3A.

4.2.2 Time Window

In the next step, we fix the temporal resolution parameter to $\tau = 5$ ms, and vary the length T of every spike train window taken into account. Parameter values $T = 10$ – 500 ms allow to investigate how much time is needed for a successful discrimination. The resulting probability distributions are shown in Figure 4.3B as classification matrices. After 10 ms, all spike trains are classified completely at random and the distribution $P(R_i|S_j)$ is almost uniform. However, already after 20 ms, corresponding to an average of 3–4 spikes per spike train, the discrimination performance is significantly above chance level. With the inclusion of ever longer time windows T , the ability to discriminate the eight songs grows until, for $T = 500$ ms, the songs can be discriminated with almost perfect score.

4.2.3 Comparison of Stimulus Paradigms

We can obtain a view of the discrimination performance for every (τ, T) -pair by calculating the respective percentage of correctly classified spike trains. For the original songs, this procedure yields the colour plot shown in Figure 4.4A. It takes about 200–300 ms until the probability of correct classification exceeds 95% for a resolution of $\tau = 5$ ms. Unexpectedly, longer samples of the spike trains improve the discrimination ability only marginally for other values of the parameter τ , as can be seen by the nearly horizontal contours surrounding the area of 95% correct classification, cf. Figure 4.4A. This effect can be attributed to the repetitive structure of these songs. After a certain amount of time, additional syllables do not significantly aid in the song discrimination.

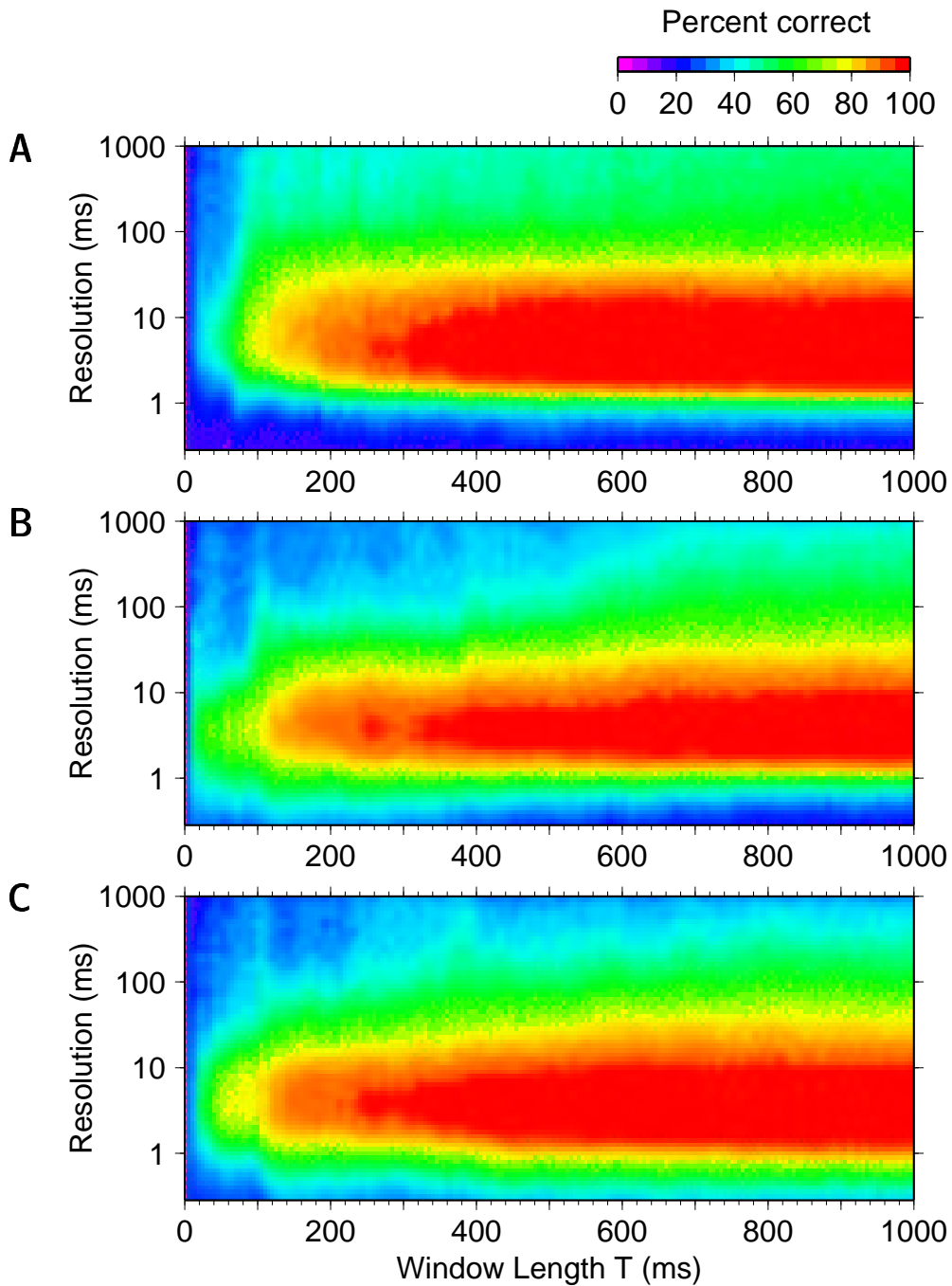


Figure 4.4: Percent of correctly classified spike trains as a function of both τ and T . The results here are again based on the van Rossum Distance. **A** For the original set of eight songs, the discrimination performance is best for τ -values around 5 ms. About 200–300 ms have to be taken into account to achieve a classification performance of at least 95%. **B** Surprisingly, the discrimination performance is approximately preserved for the set of songs with rescaled syllable-pause length. **C** Even additionally equalising the carrier frequencies does not diminish the discrimination performance.

So far we have investigated the discrimination performance on the original eight songs that feature differences in both syllable-pause duration and frequency content. It could therefore be possible, for instance, that the discrimination works mainly because the songs differ strongly in their syllable-pause length (60–120 ms). Due to these differences, the pauses in between syllables appear at different times in spike trains of different songs. If the discrimination were primarily based on these obvious differences, then the discrimination performance should decrease significantly for the second set of songs where syllables and pauses are rescaled to a common length. Given the resulting alignment of syllables and pauses, a successful discrimination of the rescaled songs must be based on the remaining differences in the amplitude modulations and frequency content of the syllable substructures. As Figure 4.4B show, both the total time and temporal resolution required for successful discrimination remain approximately the same.

We have seen that differences in the low-frequency content of the songs lead to different firing rates. These differences in firing rates also aid the discrimination; for instance, they suffice to distinguish songs no. 1–5 from 6–8. Within the third set of songs, this clue is also deleted by equalising the low-frequency content. Surprisingly, the ability to differentiate still does not decrease (cf. Figure 4.4C). A few hundred milliseconds suffice to distinguish all eight stimuli ($> 95\%$). Hence, the intrasyllable amplitude modulation is completely sufficient to discriminate the songs given our discrimination measure. As the optimal resolution resides around $\tau = 5$ ms, the intrasyllable structure should be the feature of choice for the discrimination task.

Appendix C shows that the van Rossum Distance and the Victor-Purpura Distance are closely related. Since this relation is of special interest, we recomputed the discrimination performance for the Victor-Purpura Distance, cf. Figure 4.5. As predicted by the heuristic considerations in Appendix C, the result is essentially the same, underlining the robustness of the distance measures.

4.2.4 Summary of Experiments

These observations hold for all the experiments performed. While $n = 22$ receptor neurons were recorded altogether, the complete stimulus protocol, consisting of all three sets of songs, was tested on only $n = 7$ neurons which were all low-frequency receptors. A summary of the discrimination performance of these receptors is displayed in Figure 4.6A–C. For all three sets of songs, the percentage of correctly classified spike trains surpasses 95% if the resolution is $\tau = 5$ ms and the time incorporated $T = 200$ –300 ms, corresponding to about 30 spikes. Note that the chance level, i.e., the level of random assignment of songs is given by 12.5%. Even for $\tau < 1$ ms, the discriminability remains above chance level, indicating that the spike timing precision extends well into the submillisecond regime.

Although the songs in the third set have approximately the same power in the low-frequency band, a significant fraction ($\approx 40\%$) of the spike trains is correctly classified when the discrimination is based on the average firing rate alone, cf. Figure 4.6C. This

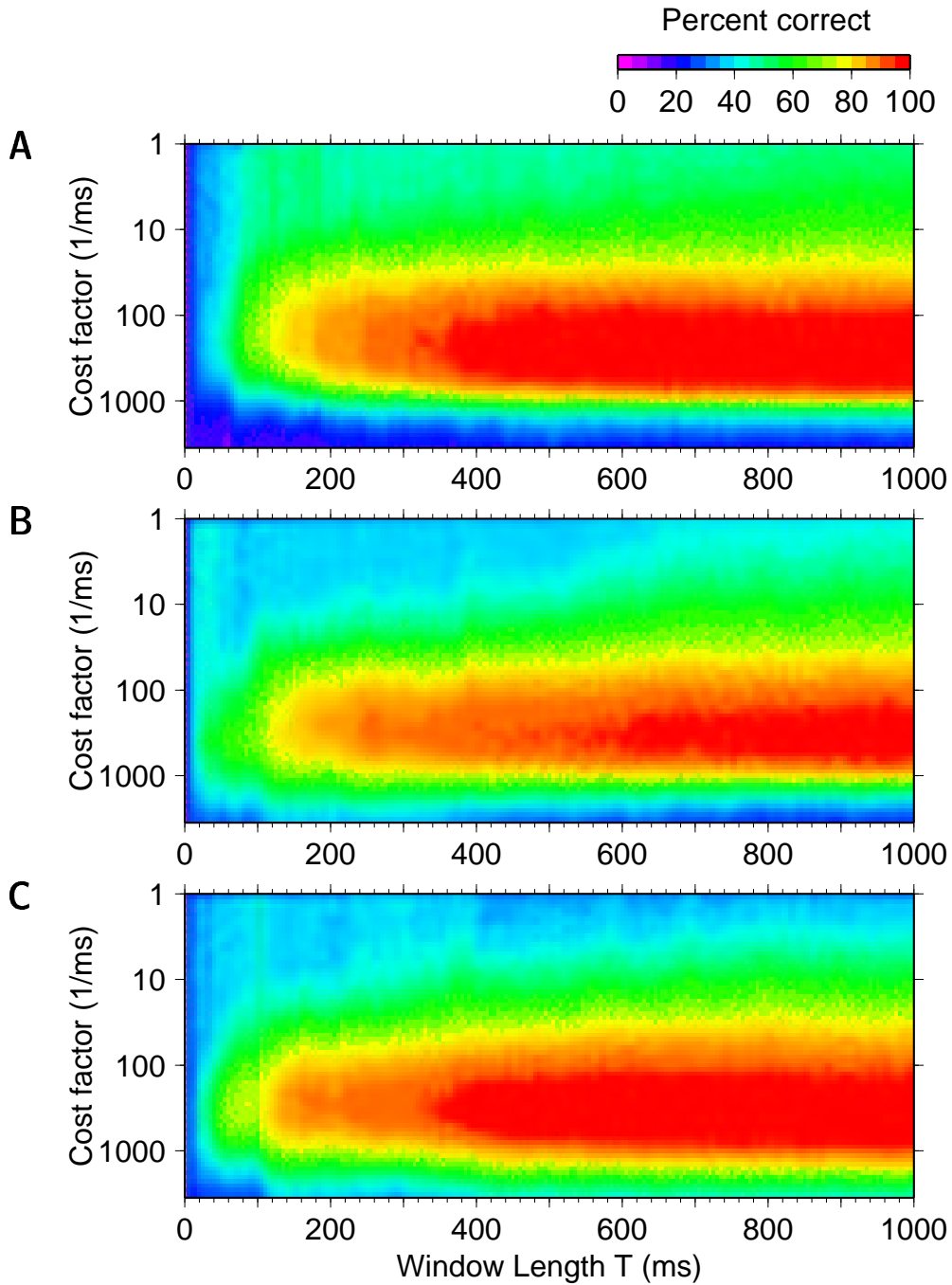


Figure 4.5: The Figure is equivalent to Figure 4.4, however the results here are based on the Victor-Purpura Distance. Shown are the percent of correctly classified spike trains as a function of the cost factor q and the window length T . Since the cost factor q corresponds to the reciprocal of the resolution parameter τ , the Victor-Purpura Distance essentially yields the same results.

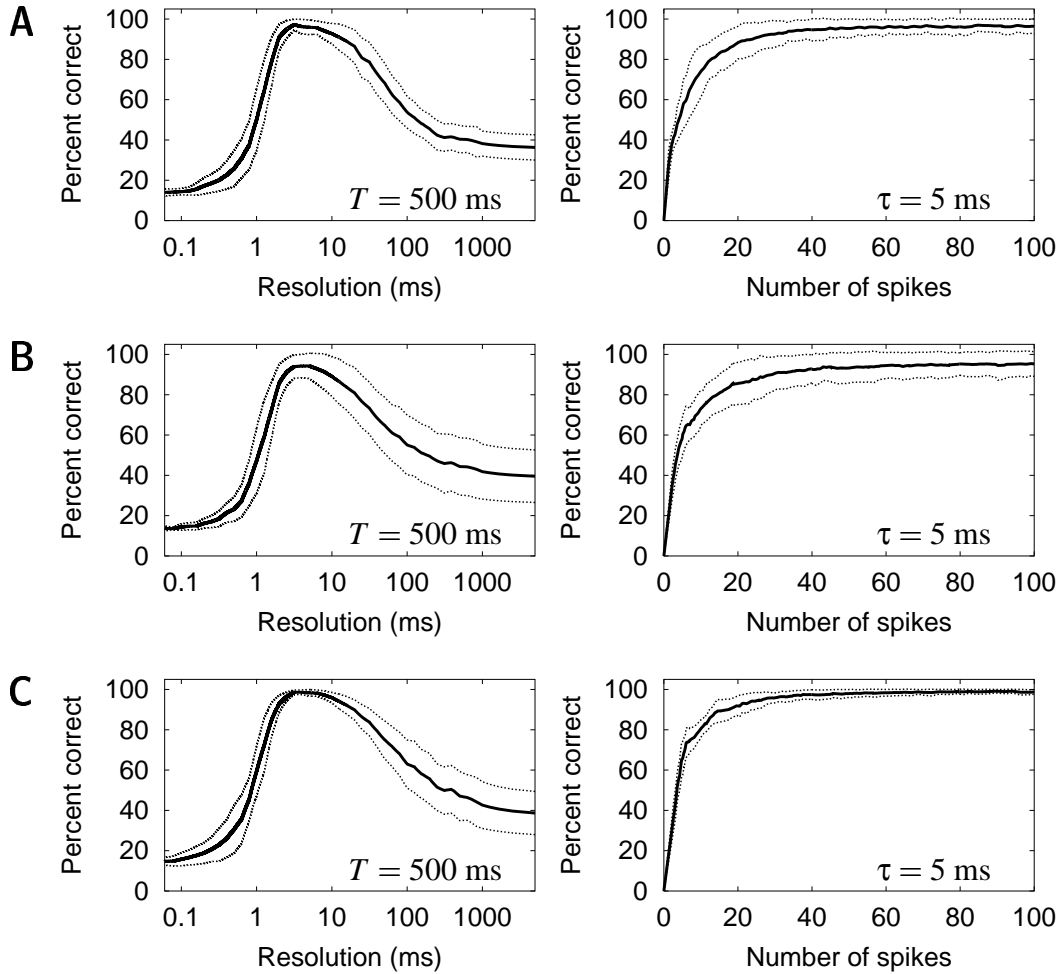


Figure 4.6: Summary of the experiments. Thick lines indicate the mean, the dotted lines the standard deviation. **A** Original songs. The left panel shows the percentage of correct classification for a window length $T = 500$ ms and different values of the resolution parameter τ . In contrast, the plot on the right fixes the resolution to $\tau = 5$ ms while varying the window length T . The abscissa shows the average number of spikes present for a certain window length T . **B** Songs with rescaled syllable-pause duration. **C** Songs with rescaled syllable-pause duration and equalised frequency content.

phenomenon can be attributed to the slow rising phase of the songs: as visible in Figure 3.2 and Figure 3.3, the amplitude of the songs grows slowly with time during the first 500–1000 ms. The length of this phase differs slightly between songs, a fact that leads to minor differences in the overall firing rates during the beginning of each song.

To avoid the rising phase, we therefore also performed the discrimination procedure on the last second of the songs, cf. Figure 4.7. In three out of seven experiments this led to a decrease of the discrimination ability. However, this is the result of an artifact as the corresponding receptors were driven into the saturation regime for some of the songs; a problem that was independent of the stimulus paradigm. In the remaining four experiments the discrimination ability did not decrease, although almost no information can be extracted from the average firing rates, cf. Figure 4.7C, left.

It is quite noteworthy that during the last second, the remaining mean sound intensity differences between songs can be as little as 0.1 dB. Still, the discrimination performance works amazingly well; indeed, two (instead of eight) songs can often be discriminated after a few tens of milliseconds, or three to four spikes. On the other hand, more than one hundred spikes must accumulate before one can differentiate between two constant sine tones with a comparable intensity difference (data not shown). This coincides with the observation that constant or slowly varying stimuli lead to lower spike timing precision (Machens et al., 2001c; see also Chapter 5), as is also observed in neocortical neurons (Mainen and Sejnowski, 1995).

For all experiments, the best temporal resolution is always around 2–10 ms. Information conveyed by the spike trains on such a resolution seems to be best suited for the discrimination task. Therefore, the question arises, whether there are features of the songs, varying on such time scales, that might be informative about the condition of the respective male singer.

4.3 Relation of Song Features and Body Symmetry

With a population of 40–50 receptors per ear, even more information is available about the precise amplitude modulation pattern. To discriminate conspecific from heterospecific signals, however, such precision is not necessary. One might therefore suspect that the acquired information is used to evaluate the male singer.

4.3.1 Song Features and Male Impairments

As mentioned in Chapter 3, grasshoppers produce their songs by a rhythmic movement of the hindlegs against the forewings. During a syllable, each hindleg moves up and down three to four times. Each of these strokes lasts for about 7–12 ms and is succeeded by a short gap (2 ms) at the movement reversal point. In a healthy male these brief gaps are camouflaged by systematic phase shifts between the movement of both legs, leading to a fairly smooth total amplitude modulation pattern within a syllable. However, if one of the hindlegs is weaker, the sound produced by the other side will dominate. This can happen in young immature males, as well as in males that have a damaged hindleg

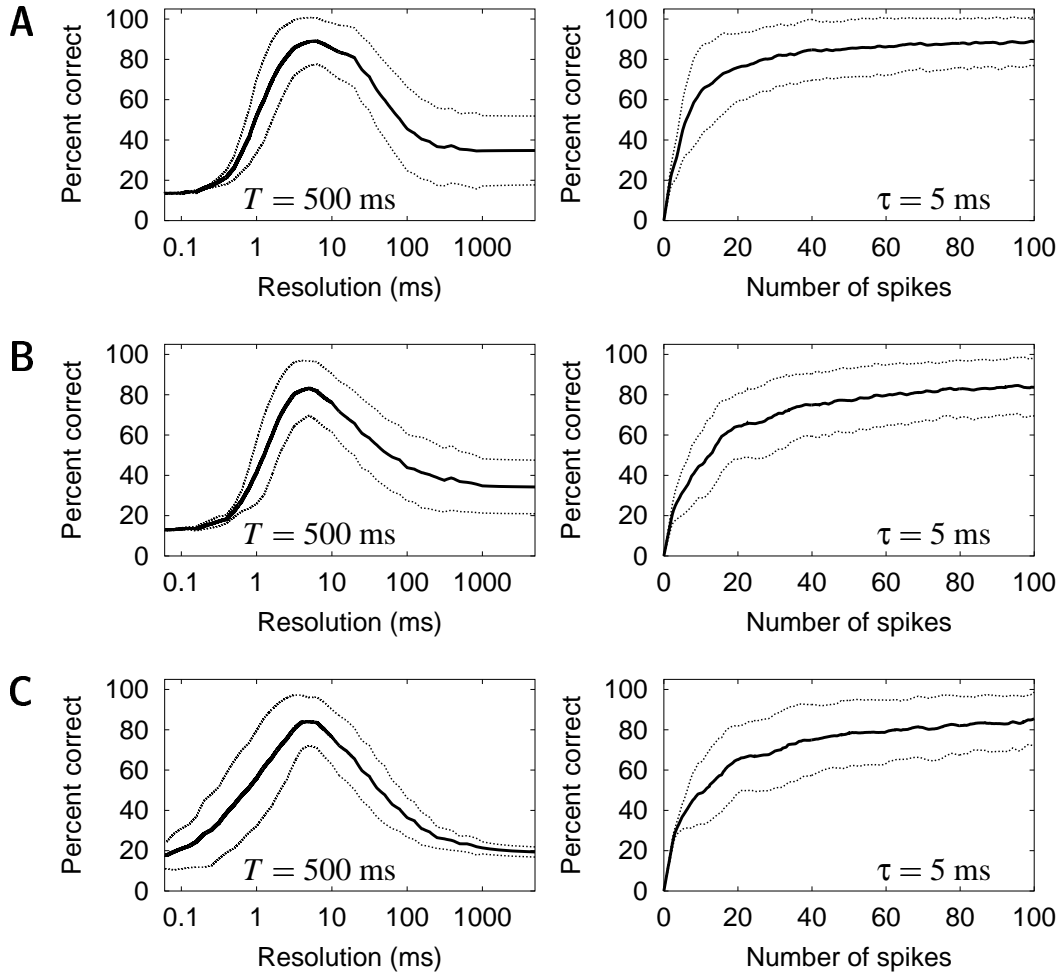


Figure 4.7: Summary of the experiments. This Figure is organised as Figure 4.6; it shows the discrimination performance for the last second of each song. **A** Original songs. **B** Songs with rescaled syllable-pause duration. **C** Songs with rescaled syllable-pause duration and equalised frequency content. As the firing rates on the last second are approximately the same for all songs, the discrimination performance drops to less than 20% for resolutions $\tau \rightarrow \infty$, showing that only negligible information can be gained from the average firing rate, compare Figure 4.6C.

or crippled wing due to a moulting accident; in the extreme case of loss of a hindleg (for instance, due to a predator), the songs exhibit pronounced 2 ms gaps within each syllable as shown in Figure 3.3C and 3.4C. Behavioural studies have shown that females refuse to mate with such impaired singers (Kriegbaum, 1989; von Helversen and von Helversen, 1997). A deviant amplitude modulation pattern can also be caused by an inadequate coordination between the two hemisegmental central pattern generators (CPGs) that are responsible for the stridulation movement (Hedwig, 1986; Ronacher, 1989). In a healthy male, either hindleg can take the leading part to achieve the phase shift. Hence, both CPGs must be highly symmetrical.

4.3.2 Male Impairment and Fluctuating Asymmetry

The amplitude modulations of a calling song are therefore likely to provide detailed clues to the state of the respective male, in particular to the symmetry of its nervous system, muscles, legs and wings. It is thus advantageous for a female grasshopper to carefully “inspect” these amplitude modulations when evaluating the song of a potential mate. The data analysed in this Chapter demonstrate that the auditory periphery preserves even the most minute differences between the songs of conspecific grasshoppers.

It is interesting to note that the features of the songs that allow a discrimination are related to the symmetry of a male. Especially with respect to visual ornaments, it has previously been proposed that body symmetry might be an indicator of genetic quality, in particular, to the ability of an organism to develop harmonically despite environmental disturbances (Møller, 1992; Watson and Thornhill, 1994). In this context, one might therefore suspect that information about the body symmetry of male grasshoppers is conveyed in the acoustic domain. Previous studies in crickets have suggested a similar role for the spectral properties of cricket calling songs (Simmons and Ritchie, 1996).

SUMMARY AND OUTLOOK:

The calling songs of male grasshoppers constitute the most important set of stimuli that a female’s auditory system has to process. We have shown that single auditory receptor neurons encode even slight variations in the amplitude modulations of the songs with a very high precision, endowing the female with the ability to easily distinguish between songs. Such a discrimination works best when comparing the songs on time scales of 2–10 milliseconds. Song features varying on these time scales are highly informative about the body symmetry, and presumably, the genetic quality of a male singer.

The time scale reflects properties of both the songs and the receptor neurons. To investigate characteristics of the receptor neurons only, one must test the system with more general ensembles of stimuli. We will do this in the next chapter in order to reveal the stimulus properties that the receptor neurons encode best.

Chapter 5

Decoding Artificial Stimuli

Neurons often convey information about various aspects of a stimulus, e.g., frequency and intensity. To uncover the ensemble of stimuli that auditory receptor neurons convey most information about, we will compare several ensembles that differ with respect to their similarity to natural sounds. Using stimulus reconstruction methods (explained in full detail in Appendix D), we can estimate the information transfer on both the single cell and population level.¹

5.1 Introduction

5.1.1 Questions

We have seen in Chapter 3 that auditory receptor neurons encode the amplitude modulation of a song within a specific frequency band and a limited range of intensities, cf. Figure 3.6. Certainly, these investigations provide us with important information about the stimuli that receptor neurons encode. However, they do not yet take into account stimuli with a time-varying amplitude modulation. As we can tell from studying natural songs in Chapter 4, dynamic amplitude modulations are an indispensable condition of acoustic communication in grasshoppers.

We will therefore broaden our view and scan large regions of the complete stimulus space. With this in mind, we are interested in two particular questions: Do receptor neurons encode a large range of acoustic stimuli or are they specifically tuned to behaviourally relevant features, such as the temporal structure of a grasshopper calling song? Is the information carried by the spike train of a single auditory receptor sufficient to identify a given stimulus or are several neurons required to do so?

¹The following has been adapted from Machens et al. (2001c). I am grateful to Petra Prinz (Humboldt-University Berlin) who performed the experiments. I also want to thank Martin Stemmler (Humboldt-University Berlin) who generated the LMD stimuli. Some of the data (concerning the SMD stimuli) has previously been analysed in a preliminary fashion in Machens (1998).

5.1.2 Stimulus Design

To investigate these questions, we designed artificial stimuli that vary the most salient statistical properties of grasshopper sounds. Two aspects of the behaviourally relevant amplitude modulation (the AM signal) have been discussed in Chapter 3: (1) The songs alternate between noise bursts and pauses, leading to a characteristic double-peak distribution of sound amplitudes (Figure 5.1A, centre). (2) The AM signals of natural songs vary on time scales from one millisecond to several seconds. To investigate the importance of these structural aspects, two different classes of stimuli were generated:

- (1) The first class consists of random stimuli that have the same distribution as a typical grasshopper song and thus imitate the gap-infiltrated structure of these songs. The modulation depth of these stimuli, defined as the range covered by the central 95% of the amplitude distribution, is approximately 24 dB. The stimuli are therefore called large-modulation-depth stimuli (LMD stimuli).

To analyse the impact of spectral properties of the AM signal, one subtype of stimulus was chosen to exhibit the power spectral density prescribed by the natural song shown in Figure 5.1A. Having both the large modulation depth of a grasshopper song and a song-like spectrum (SLS), this artificial stimulus (Figure 5.1B) comes closest to the properties of a natural song. The remaining stimuli contain a uniform or “white” mix of all modulation frequencies, from zero up to a cut-off frequency f_c of either 25, 50, 100, 200, 400, or 800 Hz (Figure 5.1C). Technical details of the stimulus generation are provided in Appendix A.

- (2) Within the second class, stimuli have a Gaussian amplitude distribution (Figure 5.1D) with a modulation depth of 10 dB and are called small-modulation-depth stimuli (SMD stimuli). These random stimuli mimic a situation where pauses and gaps of individual songs are blurred by other sounds as might occur when many grasshoppers sing simultaneously. Because the acoustic waves sum linearly and different individuals do not synchronise their song patterns, the amplitude modulations of the summed sound pressure waves will have a nearly Gaussian distribution. Assuming that 5–10 grasshoppers sing at the same time, the modulation depth of their summed sound pressure waves is roughly 10 dB. To restrict the amplitude values within a finite range, the tails of the Gaussian distributions were cut off at 3.5 standard deviations.

As Gaussian white noise stimuli have been frequently used in previous stimulus-reconstruction studies, the SMD stimuli facilitate the comparison with results obtained in other sensory systems (Bialek et al., 1991; Rieke et al., 1995; Wessel et al., 1996; Theunissen et al., 1996).

In all experiments, the amplitude distribution for each stimulus was kept constant by fixing the integrated AM signal power. A larger bandwidth, therefore, corresponds to a lower power spectral density. The final AM signal was used to modulate the amplitude of a 5 kHz sine tone, the carrier frequency preferred by low-frequency receptors

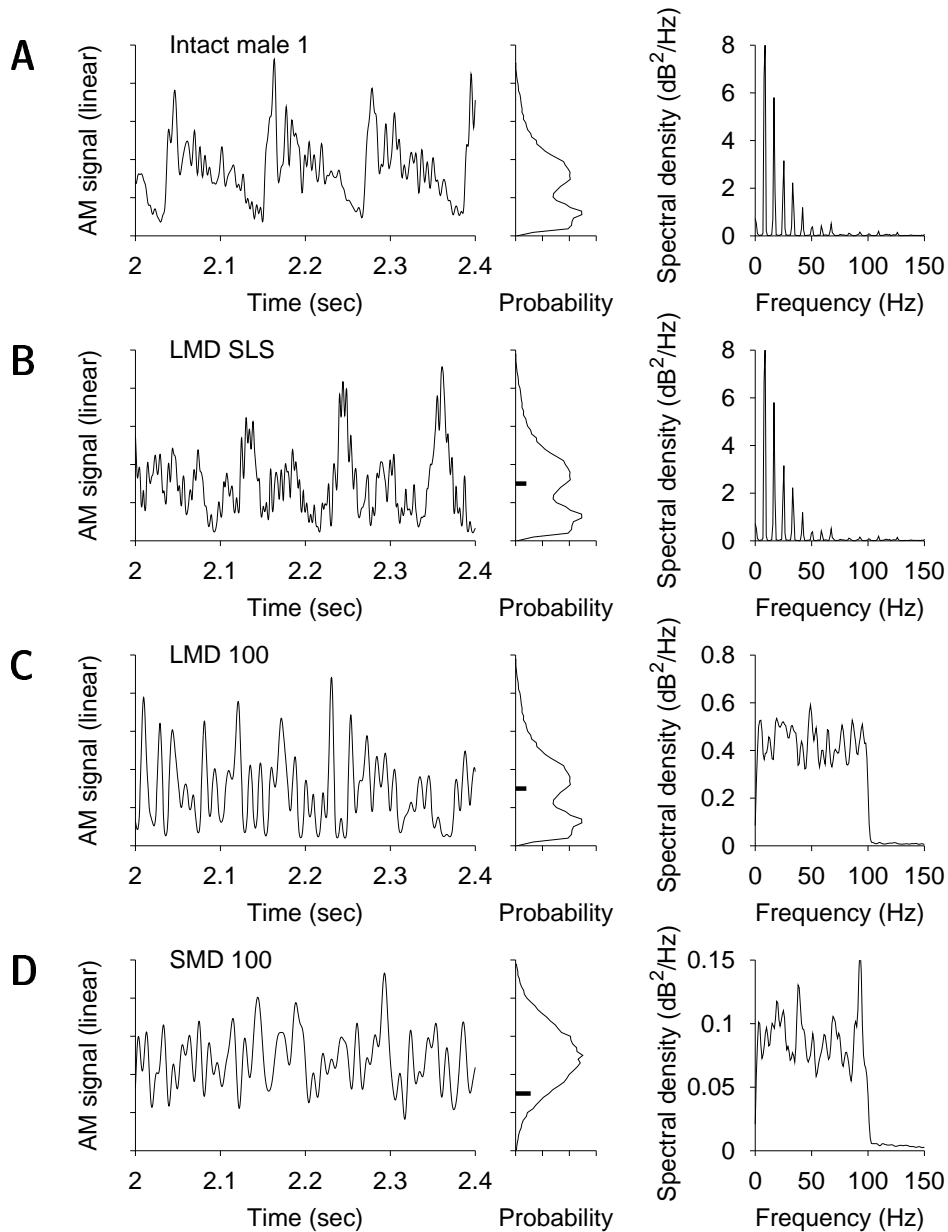


Figure 5.1: Stimulus design. **A** (*left*) The AM signal of a short song section (three syllables), cf. Figure 3.4A. (*middle*) Distribution of modulation amplitudes. (*right*) Power spectral density of the AM signal. The first peak, at about 8 Hz, corresponds to the mean duration of a syllable (~ 125 msec). **B** Design of artificial stimuli with the same amplitude distribution and the same spectrum as the natural song in A. **C** Design of artificial stimuli with the same amplitude distribution as in A and a spectrum that is “white”, i.e. flat, up to a certain cut-off frequency, here 100 Hz (*right*). Deviations from the ideal, flat spectrum result from the finite signal length. Because of their large modulation depth (24 dB) such stimuli are called LMD stimuli. **D** Design of artificial stimuli with a Gaussian amplitude distribution. These stimuli have a modulation depth of 10 dB and are referred to as small-modulation-depth stimuli (SMD stimuli). The short horizontal bars in B–D denote a threshold at which the supra-threshold stimulus power is equal for all artificial stimuli. Adapted from Machens et al. (2001c).

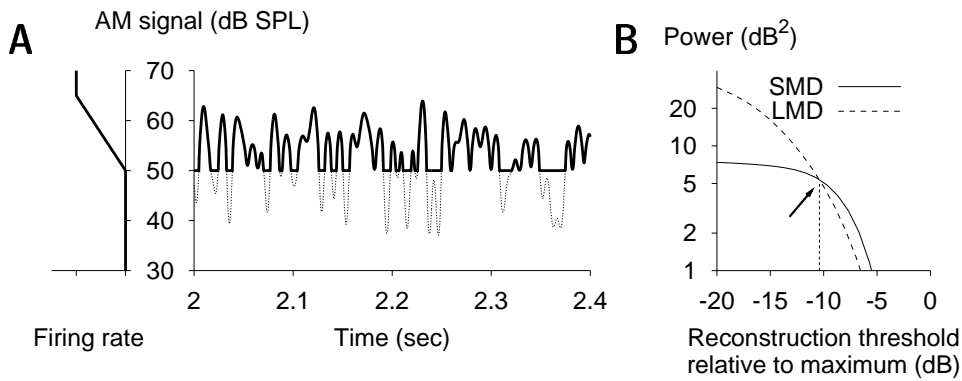


Figure 5.2: Stimulus preprocessing. **A** Transformation of the AM signal. Using a logarithmic scale for the AM signal yields a piecewise linear curve of firing rate versus sound intensity, as shown schematically on the left. The resulting preprocessed LMD stimulus is depicted as thick line (*centre*) and exhibits a short pause whenever the original amplitude modulation meanders sub-threshold. **B** Calibration of SMD and LMD stimuli. Shown is the supra-threshold power of the respective AM signal. To allow for a fair comparison of SMD and LMD stimuli, both must have the same supra-threshold AM signal power in the experiments. This point is indicated by the vertical line and by the short horizontal line in Figure 5.1 B–D. Adapted from Machens et al. (2001c).

(cf. Section 3.4.1). Stimuli were digitised to a resolution of 20 kHz, and each stimulus lasted for 10 seconds.

This set of artificial AM signals with well-defined amplitude distributions and spectral characteristics allowed us to test whether auditory receptors can encode arbitrary stimulus features down to the millisecond time scale and whether the pauses and gaps in individual songs are of any importance for the encoding procedure. Most importantly, as each stimulus is a stationary time series, it is representative for the respective stimulus ensemble, and ensemble averages can be replaced by time averages (cf. Section 2.3.1, 2.4.1). Moreover, estimating the amount of information that receptors transmit becomes a straightforward task (see Appendix D).

5.1.3 Stimulus Preprocessing and Reconstruction

Above their threshold and below saturation, the firing rate of the receptors increases approximately linearly with sound pressure level if the latter is measured on a logarithmic scale (Römer, 1976; Stumpner and Ronacher, 1991), as shown in Figure 3.6. A representation of the signal appropriate for stimulus reconstruction rather consists of a thresholded AM signal in decibels² (cf. Figure 5.2).

²Incidentally, all attempts to reconstruct the original sound pressure wave failed (data not shown). In fact, the reconstruction of the sound pressure wave approximates the mean sound pressure level, suggesting that the response-conditional ensemble is bimodal as briefly discussed in Section 2.4.1: Since receptor neurons are not sensitive to the phase of the sound pressure wave $w(t)$, both $w(t)$ and $-w(t)$ elicit the same response; the same problem has been discussed in Rieke et al., (1997) with respect to auditory afferents of frogs.

The total power of the half-wave-rectified signal depends on the applied threshold (see Figure 5.2). Both SMD and LMD stimuli have the same total power in the amplitude modulations, if the peak signal intensity is chosen to be 10 dB above the receptor's threshold. The threshold applied in this case is also indicated by small black bars in the centre panels of Figure 5.1B–D. Given such a calibration of the threshold, the stimulus ensembles retain differences in two primary traits: containing pauses (LMD) or lacking them (SMD), and having a periodic or white spectrum (all others).

To reconstruct the preprocessed AM signal from a spike train, each spike is replaced by a bandwidth-limited filter, resulting in a smooth time-varying function. After minimisation of the mean-square distance between this function and the original AM signal, the (optimal) linear reconstruction of the signal is obtained (see Appendix D).

In all, responses were recorded from $n = 27$ receptor neurons, and each stimulus class was presented to up to 10 neurons. Stimuli were presented with peak intensities ranging from 3 dB to 21 dB above the threshold of individual receptors, entailing firing rates from 40 Hz up to 160 Hz.

5.2 Single Auditory Receptors

5.2.1 A Sample Reconstruction

As illustrated in Figure 5.3, amplitude modulations of sound pressure waves can be reconstructed from the spike trains of individual auditory receptors. Figure 5.3A displays the amplitude modulation of a 5 kHz tone, the resulting spike train from an auditory receptor, and the estimated amplitude modulation as reconstructed from this spike train. The stimulus is an LMD stimulus whose spectrum of amplitude modulations is flat up to a cut-off frequency f_c of 50 Hz. The original AM signal was thresholded at 34 dB corresponding to the experimentally determined threshold of the investigated cell. The signal estimate closely follows the thresholded signal with deviations of at most a few decibels. As quantified in Figure 5.3B, the distribution of reconstruction errors has a standard deviation of about 1.4 dB.

The linear reconstruction filter is shown in Figure 5.3C and represents the contribution of a single spike, situated at time zero, to the reconstruction of the AM signal. The central peak of the filter is shifted by 7–8 msec from zero, reflecting the intrinsic delay between the stimulus presentation and the response of the auditory receptor. Results obtained for cut-off frequencies from 25 to 400 Hz and intermediate firing rates indicate that the full width at half maximum, w , of the optimal reconstruction filter is roughly given by $w \approx (2f_c)^{-1}$. The qualitative shape of the filter remains the same as that seen in Figure 5.3C (data not shown).

The signal-to-noise ratio (SNR) quantifies the reconstruction success by the ratio of AM signal variance to the variance of the effective noise as referred to the input (see Appendix D). In the example illustrated in Figure 5.3, the signal-to-noise ratio was 6:1, demonstrating that even a single auditory receptor can accurately represent the time course of the AM signal. Without proper thresholding of the AM signal, the SNR would

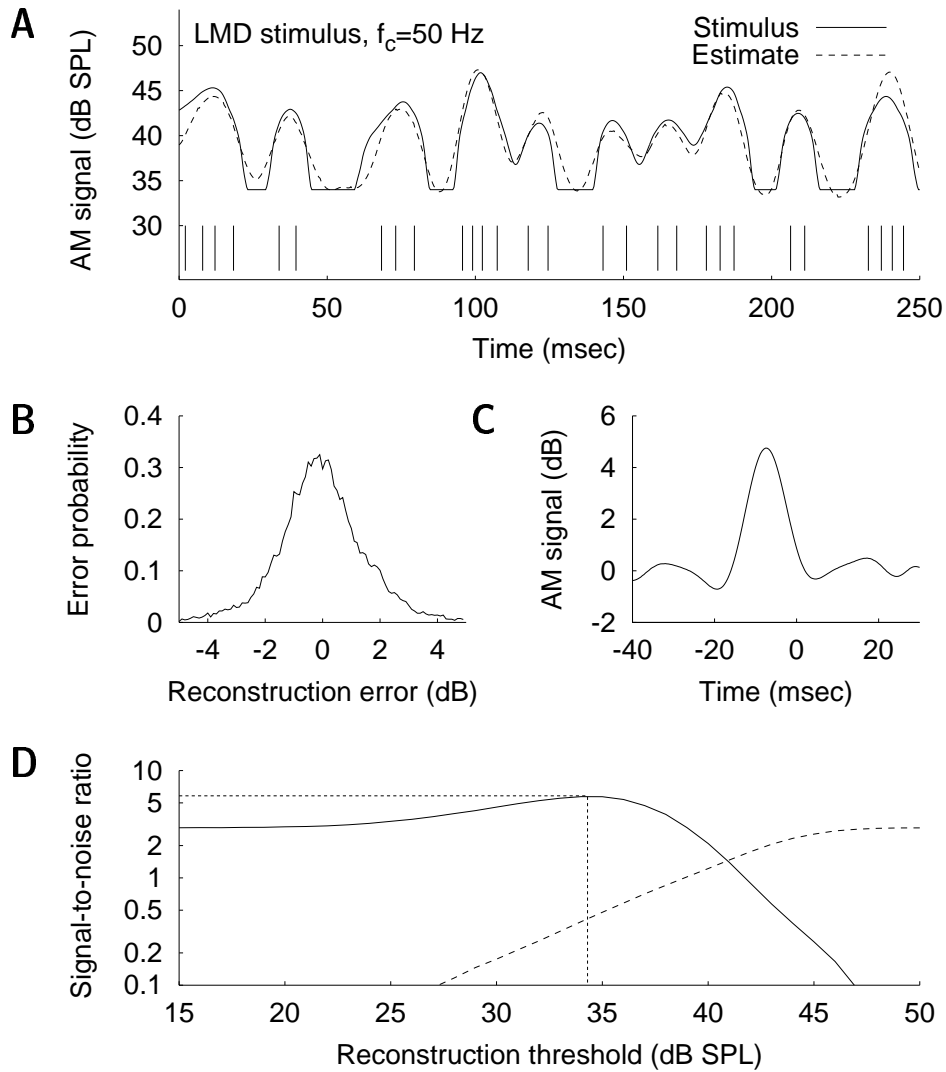


Figure 5.3: Reconstruction of a LMD signal with 50 Hz cut-off frequency from the responses of a single receptor. **A** Preprocessed stimulus, spike train, and stimulus estimate. The stimulus was thresholded at 34 dB, the threshold of this particular neuron. **B** Distribution of reconstruction errors. **C** Linear reconstruction filter. **D** Dependence of the reconstruction quality on the stimulus representation. The AM signal was split into a supra-threshold and a sub-threshold signal each of which was independently reconstructed from the spike train. Shown are the signal-to-noise ratios of the supra-threshold signal (solid line) and the sub-threshold-signal (dashed line) for different values of the applied threshold. Adapted from Machens et al. (2001c).

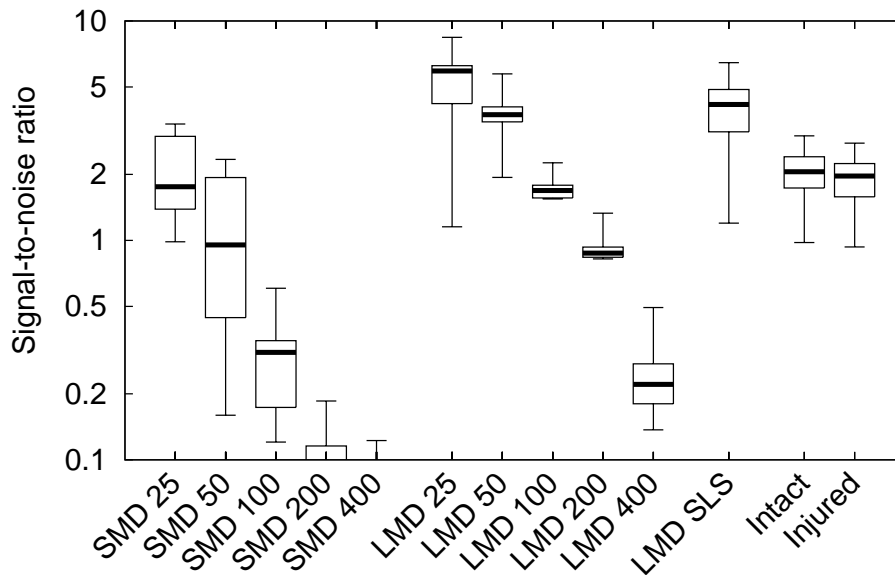


Figure 5.4: Summary of all experiments with equal supra-threshold AM signal power. Altogether 27 cells from 14 animals were analysed, with mean firing rates ranging from 40 to 100 Hz. Shown are the signal-to-noise ratios for natural and artificial stimuli. For equal cut-off frequency, LMD stimuli are always reconstructed better than SMD stimuli. The signal-to-noise ratio for stimuli with a song-like spectrum (SLS) is comparable to that of LMD stimuli with $f_c = 50$ Hz. *Thick lines* indicate the median, *boxes* the quartiles, and *bars* the maximum and minimum observed values. Adapted from Machens et al. (2001c).

decrease significantly, to values of around 3:1, as depicted in Figure 5.3D. By contrast, applying an *upper* threshold and trying to reconstruct only the *sub-threshold* portion of the AM signal leads to considerably worse results (dashed line in Figure 5.3D). At the threshold indicated by the vertical line, the SNR in this case is less than 0.5:1. Hence, little, if any, information can be recovered from the sub-threshold regime of the stimulus. These results underline the importance of carefully adjusting stimulus-reconstruction techniques to the specific properties of the neural system under study.

5.2.2 The Importance of Gaps

To allow for a direct comparison of stimuli with large and small modulation depth, a set of experiments were performed in which the AM signal power *above* the receptor threshold was made to be equal for both classes of stimuli. For this purpose, the AM signals were thresholded, and the remaining variance in the AM signals was computed. As shown in Figure 5.2, the above-threshold AM signal power is the same for LMD and SMD stimuli if the peak stimulus intensity is adjusted to be 10 dB above the receptor threshold.

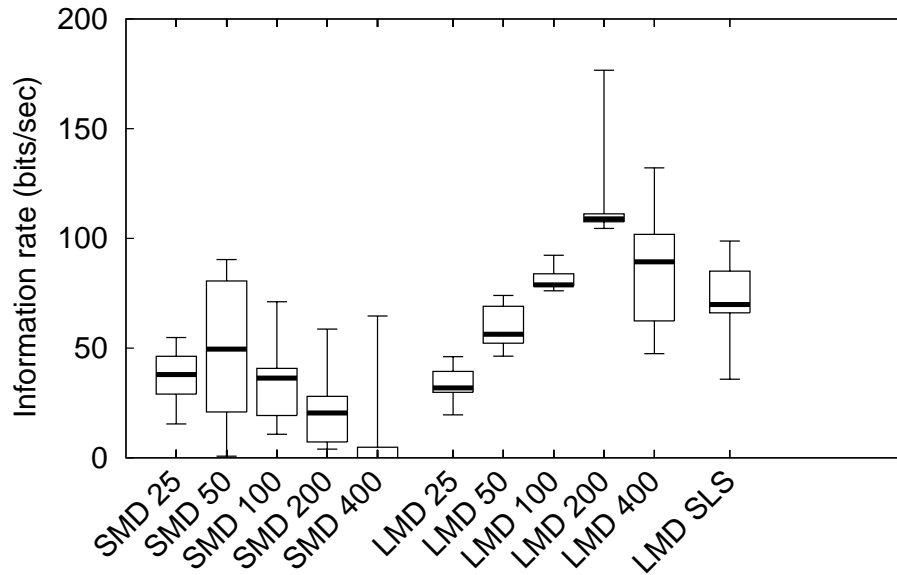


Figure 5.5: Summary of all experiments with equal supra-threshold AM signal power. The mutual information rate R_{info} quantifies the information carried by the spike train about the AM signal. LMD stimuli with a cut-off frequency of 200 Hz result in the highest values for R_{info} , suggesting that single receptor neurons are optimised for stimuli with such statistics. Since the probability distribution of natural songs is unknown, R_{info} cannot be estimated for the songs. Adapted from Machens et al. (2001c).

To compare the reconstructions obtained for different stimuli, we applied three measures: (1) Signal-to-noise ratios, to determine how accurately a given reconstruction follows an AM signal, shown in Figure 5.4. (2) Information rates, to assess how much information is conveyed by each spike train about the stimulus, shown in Figure 5.5. (3) Coding efficiencies, to measure how efficient receptor neurons employ their resources to transmit that information, shown in Figure 5.6 (see also Appendix D).

With most of the original power in the LMD stimuli well below receptor threshold, there exists no *a priori* reason to expect any difference in the signal reconstruction success for the two calibrated signals. Nonetheless, for all three measures, LMD stimuli clearly outperform SMD stimuli. No matter for which cut-off frequency, LMD stimuli can always be reconstructed more accurately than SMD stimuli. In addition, spike trains always convey more information about LMD stimuli than SMD stimuli, and receptor neurons use their resources more efficiently when presented with LMD stimuli. Hence, receptor neurons are much better suited to convey information about stimuli that feature a natural amplitude distribution and, therefore, gaps.

On the other hand, matching the spectral properties of the songs, such as the one depicted in the right panels of Figure 5.1B and D, has almost no effect: While the songs' sharp spectral peaks confer strong rhythmicity to the natural grasshopper calls,

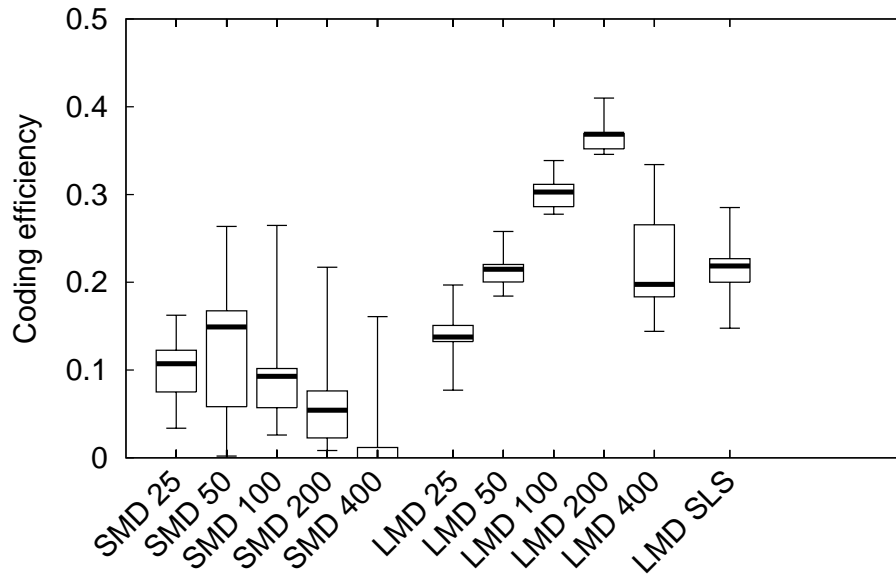


Figure 5.6: Summary of all experiments with equal supra-threshold AM signal power. The coding efficiency ϵ measures the fraction of the maximum possible information transfer. The highest efficiency is reached when using LMD stimuli with a cut-off frequency of 200 Hz: for such stimuli, the receptor neurons use their resources optimally. Adapted from Machens et al. (2001c).

rhythmicity neither impairs nor aids the quality of the reconstructions. Given that the predominant portion of the natural spectrum falls between zero and 50 Hz, the signal-to-noise ratios, information rates, and coding efficiencies of the spectrally matched stimulus (LMD SLS) are comparable to that with a 50 Hz cut-off frequency (LMD 50). Thus, filling in the valleys between the peaks in the spectrum to create the “white” LMD stimulus has almost no consequences.

The highest information rates (up to 180 bits per second) and coding efficiencies (up to 40%) are reached for the LMD stimulus with a cut-off frequency of 200 Hz. This stimulus has a natural amplitude distribution and varies randomly on time scales down to 2.5 msec. Among all stimuli tested, this stimulus best exploits the operating regime of a receptor neuron; in other words, single auditory receptors are tuned to “look” into the region of stimulus space defined by the statistical properties of this stimulus ensemble.

This finding should be compared with behavioural studies in which various artificial auditory stimuli were presented that were generated by filtering the Fourier components of model songs with regular or irregular syllable composition (von Helversen and von Helversen, 1998). These studies demonstrate that depending on the original syllable structure, Fourier components between 150 and 300 Hz are required by *Ch. biguttulus* females to reliably detect gap signals. Apparently, the response properties of single receptor neurons are optimised for features of the acoustic environment that are of

prime importance for behavioural decisions.

For all stimuli, signal-to-noise ratios decrease with increasing cut-off frequency f_c . For a linear system, this is to be expected as the power density of the AM signal decreases with increasing cut-off frequency, a consequence of keeping the total power of the AM signals constant.

5.2.3 Breakdown at Higher Modulation Frequencies

To explicitly analyse the frequency-response properties of the system under investigation, signal-to-noise ratios were resolved in the frequency domain (Figure 5.7). Our data show that for stimuli with high cut-off frequencies, slow temporal variations are represented better than fast variations. The decrease of the measured SNR's with frequency suggests that the auditory receptors act as band-pass filters for the AM signal, instead of being tuned to any particular modulation frequency. High signal-to-noise ratios can be achieved even for artificial broad-band stimuli by matching the amplitude distribution of the natural grasshopper calling songs, as in the LMD stimuli.

Signal-to-noise ratios decrease for high cut-off frequencies, as seen in Figure 5.7D–F. Faster variations can only be reconstructed for LMD stimuli, in which the AM signal repeatedly falls below threshold. But even then no significant amount of information about the stimulus is retrievable for frequencies above 400 Hz (Figure 5.7F). On the basis of the spike train of a locust auditory receptor, decoding arbitrary signal features at time scales of 1.5 milliseconds is nearly impossible. This drop off at higher frequencies does not depend on the mean firing rate of the receptor (data not shown).

We have shown that receptor neurons convey most information about LMD stimuli with a cut-off frequency of 200 Hz. However, as revealed by their signal-to-noise ratios, the stimuli cannot be accurately reconstructed. This seemingly paradox situation can be explained as follows: The receptor neurons are tuned to the region in stimulus space described by the statistical properties of the LMD 200 stimuli, as the information transfer is maximal in this case. However, the region is not yet resolved with a precision that allows an accurate reconstruction of stimuli; on the level of single auditory receptors, the resolution of stimulus space is simply too coarse.

5.3 Population Codes

Distributed representations based on the activity patterns of a population of many auditory receptors might admit an improved resolution of stimulus features in both time and intensity. Additionally, the population serves to increase the overall range of sound intensities covered, as the 40–60 low-frequency receptor neurons on each body side of the animal have thresholds spreading over more than forty decibels SPL (Römer, 1976; Ronacher and Krahe, 2000; Jacobs et al., 1999). The question then poses itself: can the population encode arbitrary stimulus features of one to two milliseconds duration, the minimal time needed to detect the gap in the song of a “one-hindlegged” male?

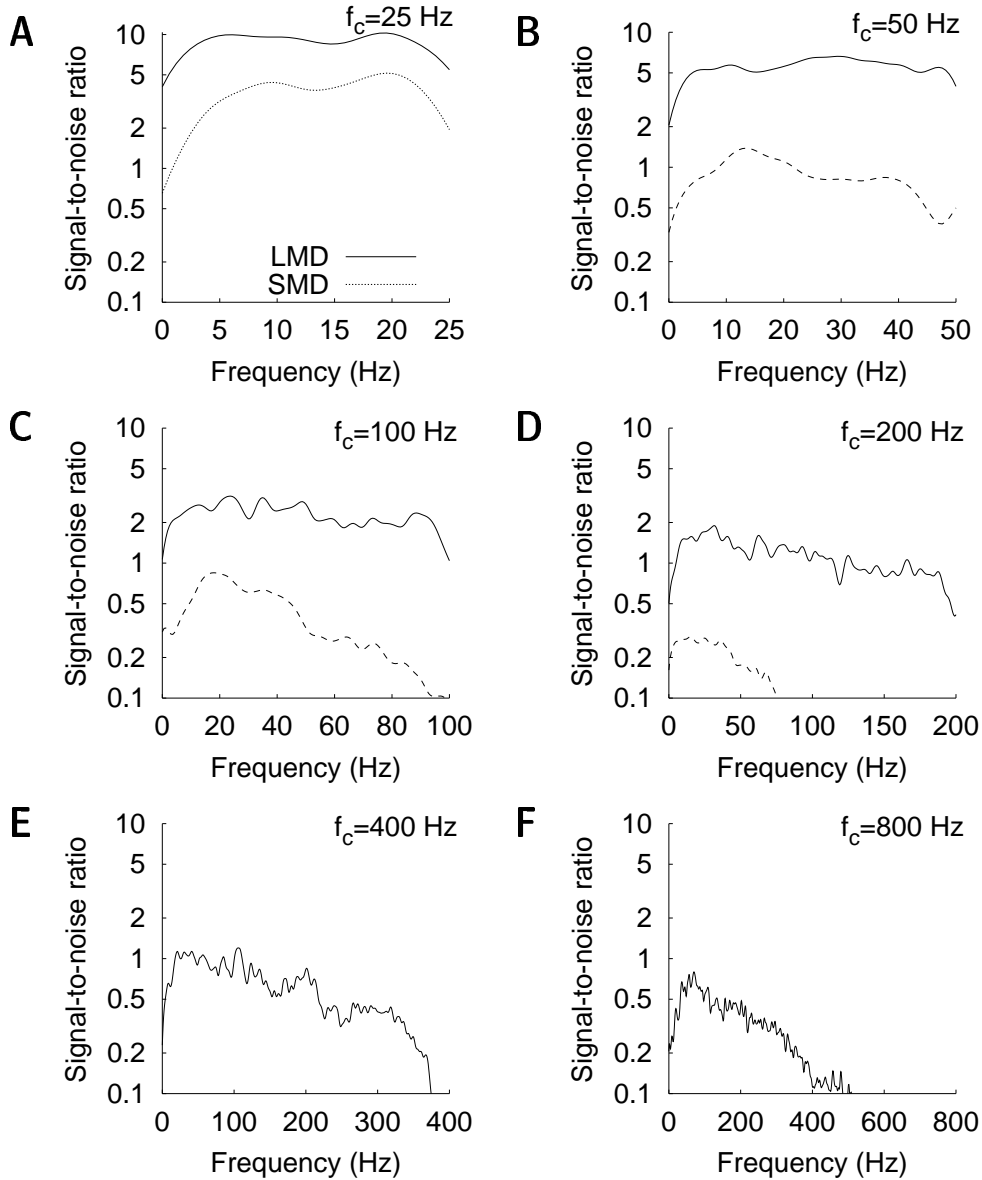


Figure 5.7: Signal-to-noise ratio for the artificial stimuli, shown in the frequency domain. Again, sound intensities were chosen such that above the receptor thresholds, LMD and SMD stimuli had the same AM signal power (4.8 dB^2). Firing rates were 40–60 Hz for the LMD stimuli (solid lines) and 60–80 Hz for the SMD stimuli (dashed lines). Although LMD stimuli are encoded by fewer spikes, their longer sub-threshold periods and steeper onsets result in more accurate reconstructions as shown by the higher signal-to-noise ratios. The fine structure of the curves is not statistically significant since the signal-to-noise values have relative errors of 17.5%. Adapted from Machens et al. (2001c).

No evidence has been found to date for any physiological coupling between auditory receptors in acridid species. On the assumption that the responses of different receptors are independent and that the receptor properties do not change during the experiment, sequential recordings from different cells may be pooled to estimate the information carried by a population of auditory receptors.

5.3.1 Redundancy

To quantify the information that can be gained by pooling responses, we computed the information redundancy of two cells as a function of the stimulus type presented. Both cells had the same threshold and encoded the same stimulus range. As shown in Figure 5.8A (left) repeated LMD stimuli with low cut-off frequencies (25 Hz, 50 Hz, and 100 Hz) result in highly redundant spike trains. For these stimuli, then, pooling responses yields almost no additional information. By contrast, SMD stimuli with higher cut-off frequencies (100 Hz and more) elicit spike trains that carry largely independent information.

The information redundancy calculated over repeated stimulations of the same cell (Figure 5.8A right) is almost the same as that from different cells (Figure 5.8A left). This indicates that the encoding procedure of cells sensitive to the same stimulus range is almost identical. It thus matters little whether spike trains from the same or different neurons are being compared, as long as the neurons have the same sound-intensity threshold.

The information gained by additional spike trains must come from the variability between these spike trains, as identical spike trains are necessarily completely redundant. The trial-to-trial variability for some of the tested stimuli can be inferred from Figure 5.8B and C. Spike train variability, however, does not always correlate with non-redundant information. For instance, although spike responses elicited from an LMD stimulus with $f_c = 200$ Hz (Figure 5.8C, right panel) are more reliable than those elicited from an SMD stimulus with $f_c = 25$ Hz (Figure 5.8B, left panel), the latter has a higher redundancy.

If two receptor neurons respond to distinct, non-overlapping stimulus intensity ranges—in other words, if each receptor “hears” a different part of the stimulus—the information redundancy in the spike trains decreases correspondingly (data not shown).

The high cell-to-cell redundancy in the response to stimuli with large modulation depth suggests that synchronous response patterns are solely due to stimulus locking and do not carry additional information based on the relative timing of different spike trains. Synchronicity may nevertheless serve as an important signal for down-stream neurons and improve their capability to reliably detect rapid stimulus onsets. In addition, a high cell-to-cell redundancy endows the system with fault tolerance with respect to failure of individual receptors.

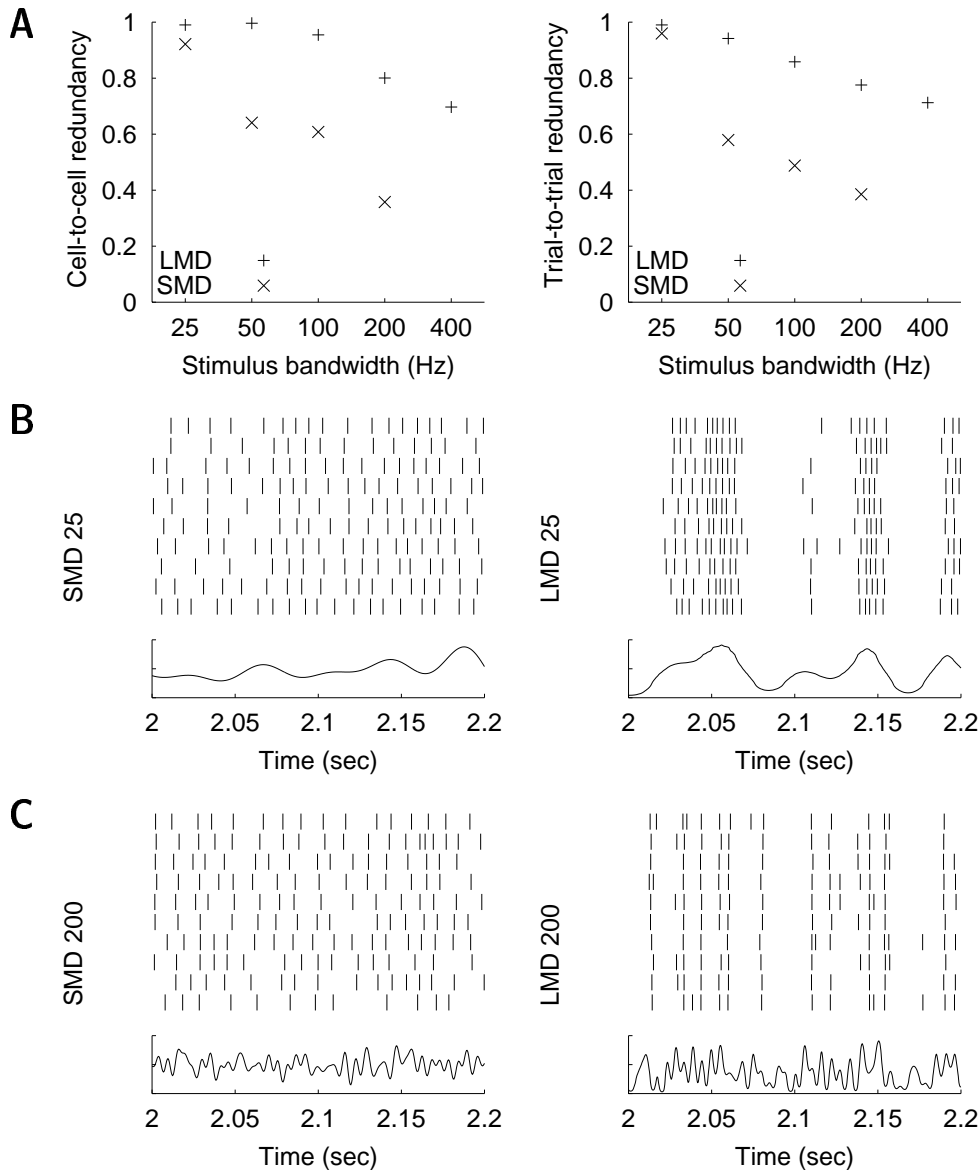


Figure 5.8: Redundancy and reliability of neural responses. **A** (*left*) Redundancy of spike trains from two different cells as a function of stimulus type and cut-off frequency. The peak stimulus intensity was 60 dB and both cells had a threshold of 50 dB. Neurons convey identical information about a stimulus if the redundancy equals one, while if the redundancy is zero, they convey independent information. (*right*) The redundancy of spike trains from one cell responding to repeated presentations of the same stimulus. LMD stimuli (+) cause larger redundancy than SMD stimuli (x) and in both cases, cell-to-cell redundancy is comparable to trial-to-trial redundancy. **BC** Spike raster plots for one of the cells in response to the stimuli denoted on the ordinate. In terms of spike timing precision, the LMD 200 stimulus (**C**, *right*) clearly stands out but other stimuli result in higher trial-to-trial redundancies (**A**, *right*). Adapted from Machens et al. (2001c).

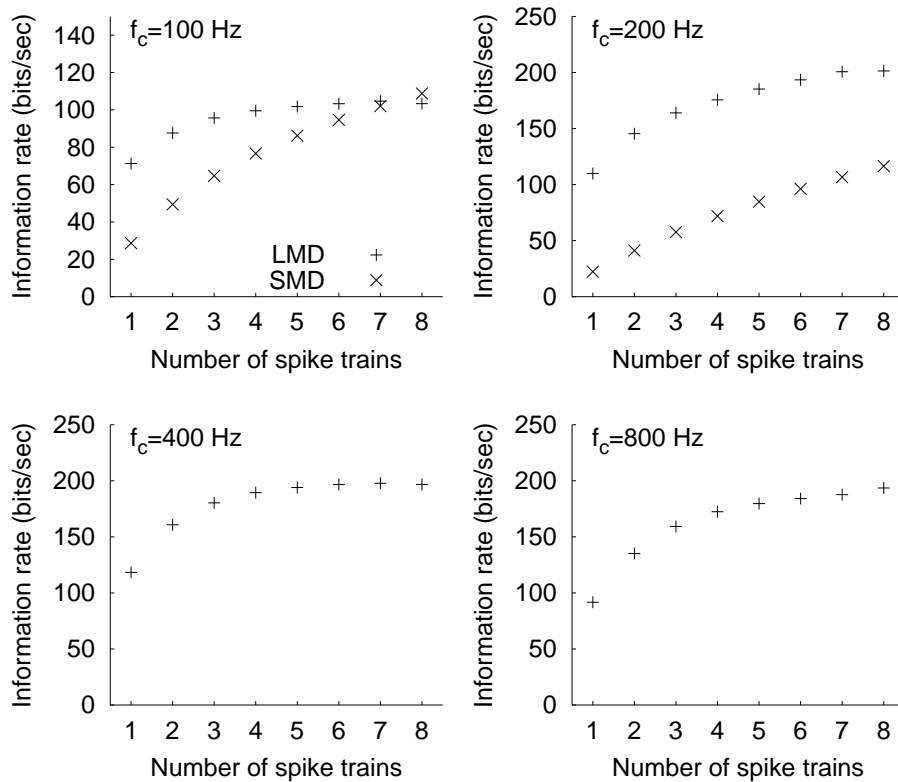


Figure 5.9: Reconstruction from multiple spike trains. Depicted is the mutual information as a function of the number of spike trains used in the reconstruction for stimuli with cut-off frequency f_c . Here a pool of eight spike trains was used that were taken from two cells having the same threshold (50 dB). The peak stimulus intensity was 60 dB. As the information redundancy shown in Figure 5.8 already suggests, information rates improve particularly strongly for SMD stimuli and stimuli with high cut-off frequency. Adapted from Machens et al. (2001c).

5.3.2 Population Reconstruction

Given the reduced redundancy of information for stimuli with higher bandwidth and/or small modulation depth, pooling neuronal responses should aid most in uncovering information about short-time and small-intensity features of the AM signal. This is indeed true. For instance, Figure 5.9, upper left panel, shows that the information rate for the SMD stimulus with a cut-off frequency of 100 Hz increases much faster with the number of spike trains than the information rate for the LMD stimulus which begins to saturate already when four or five spike trains are used for the reconstruction.

Signal-to-noise ratios based on a pool of eight spike trains (Figure 5.10) are significantly improved as compared to SNR's in the reconstruction from a single spike train, especially at high frequencies (cf. Figure 5.7D–F). Pooling spike trains is similarly efficient and important for stimuli with larger bandwidth. A pool of eight spike trains evoked by the SMD stimulus with a cut-off frequency of 200 Hz provides more than four times as much information as a single spike train (Figure 5.9). Even more dra-

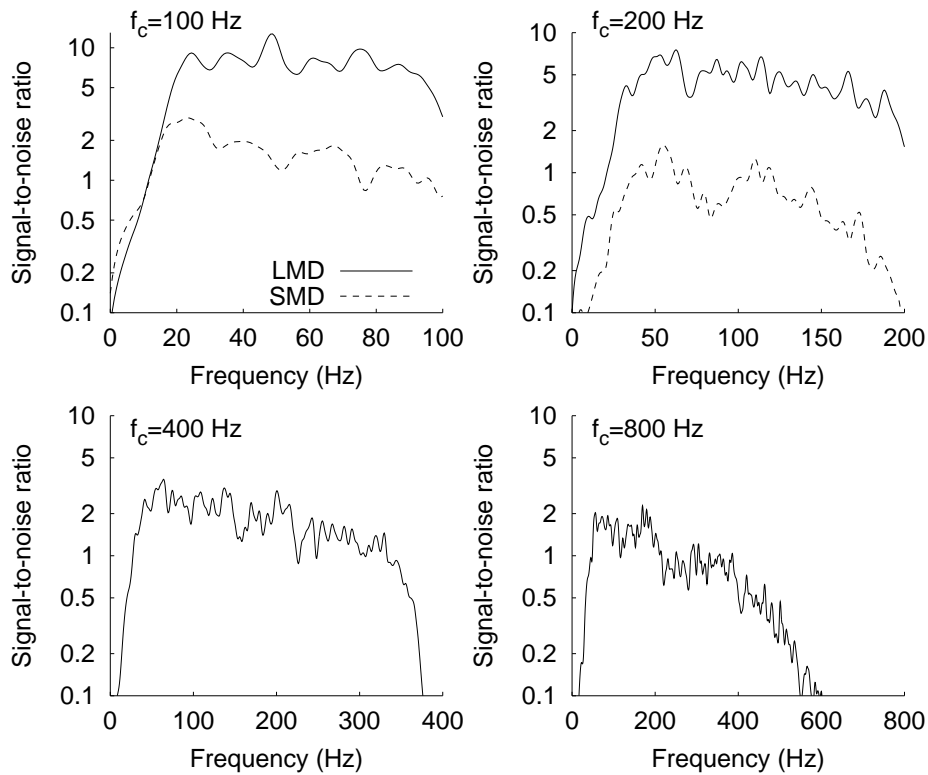


Figure 5.10: Signal-to-noise ratio when all eight spike trains are combined for the reconstruction. SMD stimuli with cut-off frequencies of 400 Hz and 800 Hz are not shown since they led to insignificant information transfer and vanishing SNR values. Signal-to-noise ratios obtained from the full pool of spike trains are significantly larger than those calculated for single cells, but decrease rapidly at high frequencies. Signal components above 400 Hz are thus still poorly represented in the stimulus reconstruction. Adapted from Machens et al. (2001c).

matic results could be expected for SMD stimuli with higher cut-off frequency but the measured values for R_{info} are too small and too variable to give statistically significant results after cross-validation. Finally, for cut-off frequencies $f_c = 200$ Hz, 400 Hz or 800 Hz, the information rate for LMD stimuli increases by roughly 100 bits/sec when pooling eight spike trains (Figure 5.9), as compared to reading from a single spike train.

Signal-to-noise ratios decay quickly for very high modulation frequencies but are still above or around 1:1 at 300 Hz, corresponding to a time resolution that samples stimulus features every 1.5 to 2 milliseconds (Figure 5.10, $f_c = 400$ Hz). Stimulus features are thus faithfully represented at all behaviourally relevant time scales. This increase in the coding performance is possible because cell-to-cell variations in the response patterns of individual receptors may be exploited on the population level.

Distributed codes involving many receptor neurons thus help to represent the acoustic environment in greater detail, especially improving the resolution for stimuli that cannot be reconstructed well from single spike trains (Figure 5.9). Combining receptor neurons that cover the same sound intensity range helps in two important ways: (1) The bandwidth of modulation frequencies that can be faithfully encoded is increased which

leads to a greater detectability of very short gaps (down to 1.5–2 milliseconds). In fact, the resolution achieved by the receptor neurons matches the resolution limits that have been found in behavioural experiments. (2) The representation of stimuli with a small modulation depth is enhanced. Information that is gained about such stimuli might help grasshoppers to detect acoustic communication signals in a noisy environment.

Strong correlations in the response patterns across neurons limit the information gained by considering multiple spike trains; the net information rate saturates at 5 to 8 combined spike trains. This number should be compared with the number of receptor cells that have a linear firing-rate characteristic in a specified intensity range. Given the threshold distribution measured by Römer (1976), a maximum of five receptors covering the same intensity range appears to be a realistic estimate. Interestingly, similar numbers for information saturation have been found for peripheral neurons in other sensory systems (Warland et al., 1997; Stanley et al., 1999).

On the level of a population of receptor neurons with the same threshold, we conclude that stimuli with a cut-off frequency of 200–400 Hz stimulate the system best and can be faithfully reconstructed at the same time. However, the population of spike trains is less efficient in the Shannon-information sense than a single spike train. The representation of an LMD stimulus recoverable from a single spike train captures up to 40% of the theoretical maximum information about the stimulus that the spike train could possibly convey. A set of eight receptor spike trains yields only up to 8% of the corresponding theoretical maximum, as receptor neurons encoding the same intensity range do not spike independently of each other.

It is natural to ask whether the receptors are optimally adapted to the environment at the complete population level. The distribution of receptor thresholds is species-dependent; hence, the thresholds could reflect the adaptation of a given grasshopper species to a specific environment. However, a thorough evaluation awaits further study.

5.3.3 Gap Detection

Although high signal-to-noise ratios indicate that the reconstruction accurately captures the time course of the AM signal, it is useful to focus on brief gaps in the AM signal, as these are of specific behavioural relevance.

A gap is a brief, silent interruption of an acoustic stimulus. Sampled at twice the AM signal's cut-off frequency, a stimulus was defined to exhibit a gap whenever the AM signal remained below the neuron's threshold for exactly one sampling point. The average length of a gap is, therefore, simply given by the inverse of the corresponding sampling frequency. Note that in this definition a gap is not a completely silent part of the stimulus, but rather a part that appears to be silent as perceived by the investigated neuron. When stimuli are presented with a peak amplitude of 10–20 dB above threshold, LMD stimuli contain a significant number of gaps, while SMD stimuli comprise at most a few gaps during the whole course of the stimulus.

A gap was called detected if the reconstructed stimulus at that time instant was smaller than a detection threshold (Figure 5.11A, “correct detection”). Varying the de-

tection threshold balances the tradeoff between the two types of error that might occur: a miss when the stimulus exhibits a gap that was not detected, or a false alarm when the stimulus exhibits no gap, but the reconstruction falsely indicates a gap (Figure 5.11A, “miss” and “false alarm”).

The tradeoff between miss and false alarm can be quantified by the receiver operating characteristics (Poor, 1994) in which the probability of correct detection is plotted against the probability of false alarm, both being parametrised by the detection threshold. This measure differs from the previous measures in that it focuses solely on the reliability of gap detection and does not take into account how accurately the stimulus is encoded between two gaps.

Figure 5.11B shows the receiver operating characteristics (ROC) for different LMD stimuli (SMD stimuli will only rarely fall below threshold and were, therefore, excluded from this analysis). Two conclusions can be drawn from this figure: (1) Using eight spike trains instead of a single one significantly enhances the detection of gaps. (2) In all cases, the ROC curve is asymmetric regarding the balance between correct detection and false alarm; even for detection thresholds that lead to few false alarms, almost all gaps can be detected.

The spurious recognition of gaps can be attributed to the fact that stimulus parts that come very close to the neuron’s threshold from above without crossing it might easily be mistaken for gaps. The nature of the stimulus as a whole may, however, help to assess whether coming close to a particular neuron’s threshold truly constitutes a gap. Later processing stages, taking into account larger sections of the stimulus or multiple spike trains, can, therefore, refine the decision of what is a gap and what is not.

5.3.4 Reconstruction of a Natural Song

Based on the results obtained for artificial stimuli, a population of receptor neurons seems to be capable of representing stimulus features on time scales down to 1.5 to 2 milliseconds. Therefore, we predicted that also the natural song of an injured, i.e., “one-hindlegged” *Ch. biguttulus* male can be reconstructed from a small group of spike trains. This is indeed the case, as shown in Figure 5.12. Here, the AM signal of the song is depicted together with pooled reconstructions from four spike trains of two receptors that had roughly the same intensity thresholds. If the sound-pressure level of the song is just above threshold, only the onset of each syllable is encoded. With increasing sound intensity the full syllable structure is recovered, including the gaps within each syllable. Interestingly, at the highest sound intensity, 21 dB above threshold, the maximum amplitude of each reconstructed syllable is fairly invariant from syllable to syllable (Figure 5.12B, bottom row), despite the slowly rising overall intensity of the recorded auditory input (Figure 5.12A). This phenomenon coincides from the adaptation of the firing rate as the sound intensity increases.

Comparing Figure 5.12 with Figure 5.4, these adaptation effects also explain why the signal-to-noise ratios from the natural *Ch. biguttulus* songs do not reach those obtained from LMD stimuli with a natural spectrum (LMD SLS) or cut-off frequencies of

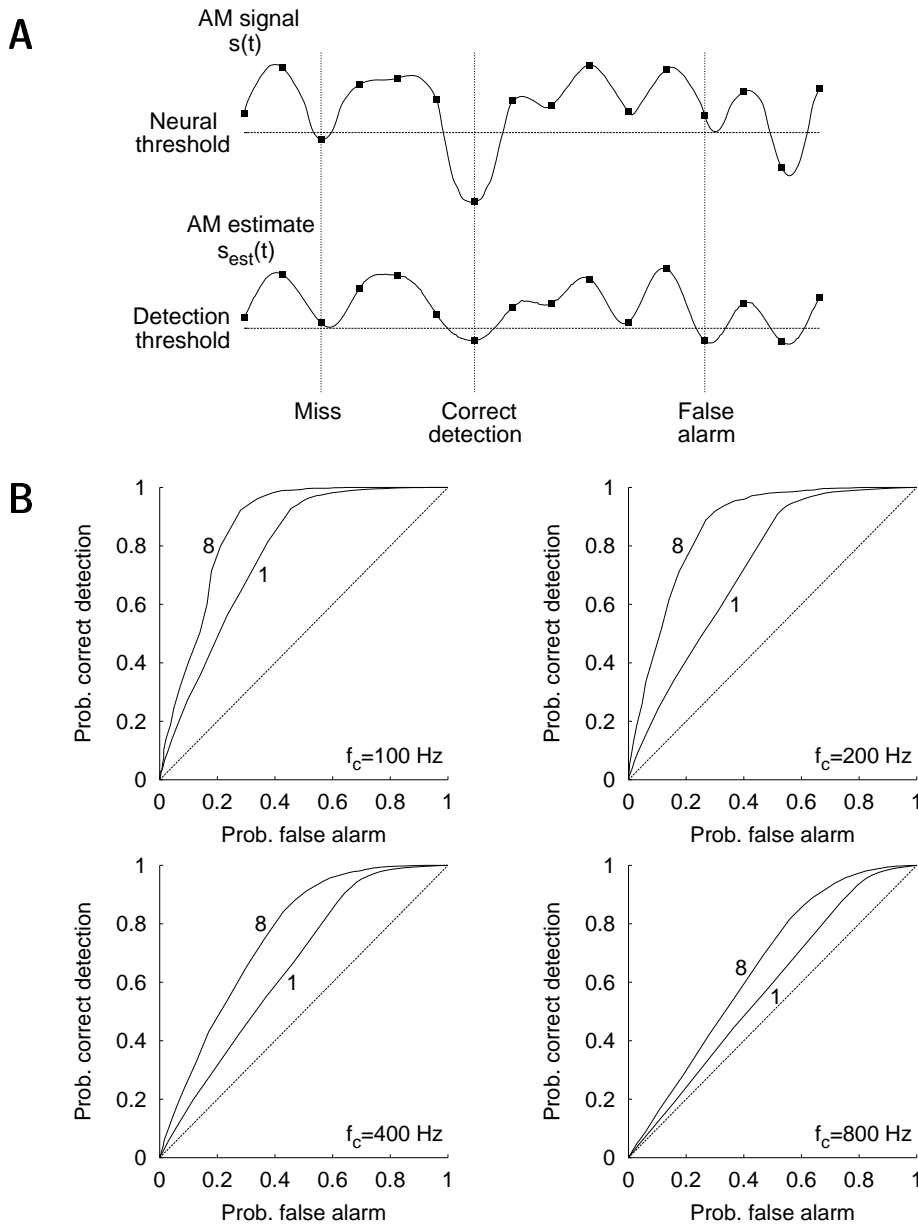


Figure 5.11: Gap detection. **A** After sampling a stimulus at twice its cut-off frequency (filled squares), gaps were defined to be all those parts of the stimulus that fell below threshold for exactly one sampling point (upper trace). A gap was classified as correctly detected if the reconstructed stimulus (lower trace) was smaller than a given detection threshold; otherwise the gap was classified as missed. A false alarm occurs when the reconstructed stimulus falls below threshold, but the signal contains no gap. **B** Receiver operating characteristics for gap detection as a function of the number of spike trains (1 versus 8) used for the reconstruction. Shown are the probability of detecting a gap (correct detection) vs. the probability of falsely predicting a gap (false alarm). The dashed identity curve marks chance level. Shown are LMD stimuli only as SMD stimuli contain almost no gaps. Using all spike trains significantly increases the gap detection performance. Adapted from Machens et al. (2001c).

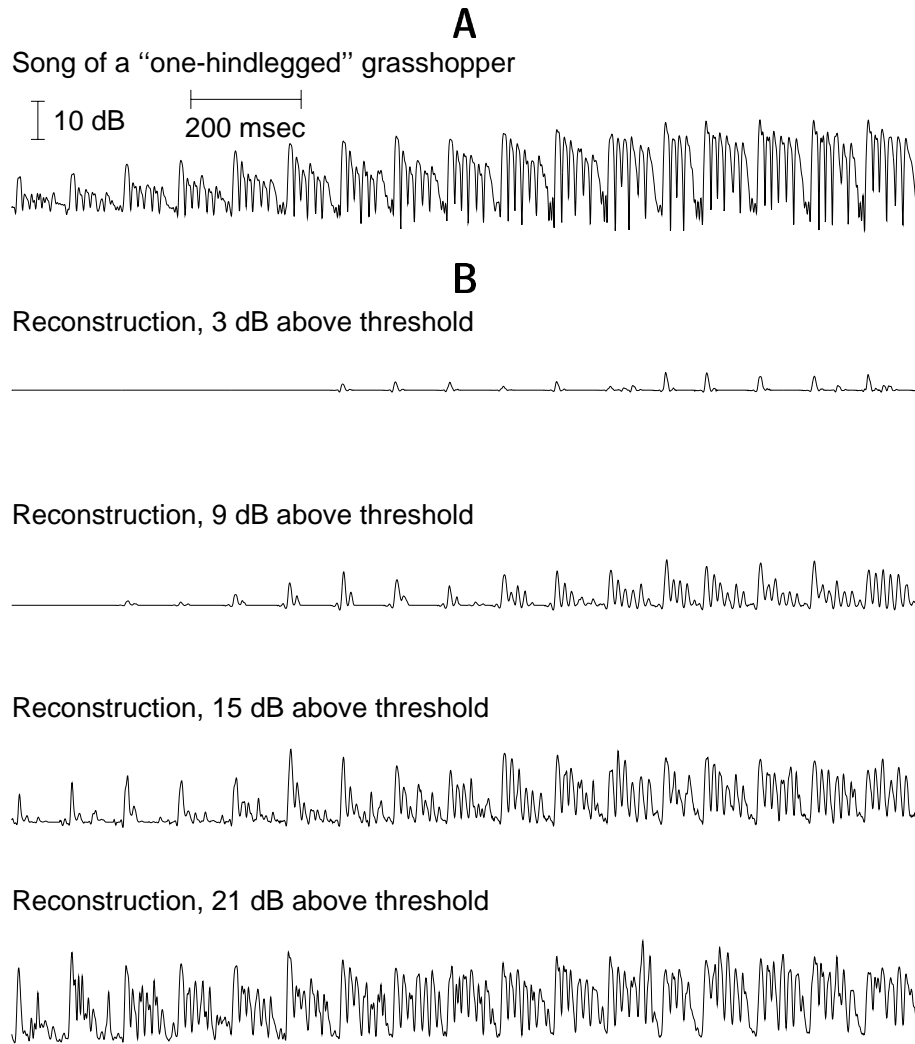


Figure 5.12: Reconstruction of the calling song of a “one-hindlegged” grasshopper. **A** AM signal. This signal has not yet been thresholded and is displayed on a decibel scale. **B** Reconstruction of the signal, thresholded at 55 dB, for different stimulus intensities (peaks at 58, 64, 70, and 76 dB). Four spike trains from two cells with thresholds of 55 dB were pooled. At sound-pressure levels just exceeding the firing threshold, only the onset of each syllable is encoded in the spike train. With increasing sound intensity, more and more details of the song appear in the reconstruction until at 21 dB above threshold even the short gaps of 2 msec length are almost perfectly preserved. Note that in the last reconstruction, the maximum amplitude of each reconstructed syllable remains approximately the same throughout the entire song. This demonstrates that adaptation effects balance the rising overall intensity of the song. Down-stream processing stages therefore receive a fairly invariant representation of each syllable onset. Adapted from Machens et al. (2001c).

25 Hz or 50 Hz (LMD 25 and LMD 50) although the songs contain their major spectral power in this frequency regime. In reconstructions using artificial stimuli, the first second of the ten-second response pattern was discarded to avoid adaptation effects. When reconstructions of the respective stimuli are based on the first 2–4 seconds, as in the reconstructions of natural songs, the obtained signal-to-noise ratios decrease and reach values similar to those obtained from the songs (data not shown).

Retrieving the AM signal from the entire set of auditory receptors characterised by staggered intensity thresholds must emphasise the start of each syllable, as many spikes in different trains coincide at the syllable upstroke (Adam, 1977; Ronacher and Römer, 1985). Stimulus reconstructions from the full receptor population will, therefore, inevitably display systematic deviations from the original AM signal structure, emphasising certain features, in particular rapid increases of the sound intensity following a pause or gap, and downplaying others.

5.4 Discussion

5.4.1 Implications for Behaviour

Against background noise generated by competing grasshoppers or other sound sources, and due to multiple sound reverberations in the habitat, the effective modulation depth of an individual song, as evaluated by a female, decreases rapidly with the distance of the male singer (Lang, 1999). With increasing distance, the song, therefore, no longer resembles an LMD stimulus and is expected to become more and more similar to an SMD stimulus in its modulation depth. This implies that the precise shape of the song's amplitude modulations can no longer be reconstructed faithfully since SMD stimuli lead to much lower signal-to-noise ratios than LMD stimuli (Figure 5.4). Heard at a distance in the field, there is thus no large difference between the reconstructed song of an intact and an injured male grasshopper. We thus predict that females only discriminate against “one-hindlegged” males if they are nearby. In fact, *Ch. biguttulus* females avoid mating with such males (von Helversen, 1972; Kriegbaum, 1989). At close distances, the high signal-to-noise ratios for LMD stimuli (Figure 5.4) should also allow the detection of much finer details in a song, which might provide the female with additional information on the male's fitness.

5.4.2 Discrimination Analysis and Stimulus Reconstruction

With appropriate preprocessing of the stimulus, the linear stimulus reconstruction from the response of a single neuron can be interpreted as an estimate of that neuron's firing rate. The high values obtained for signal-to-noise ratios, information rates and coding efficiencies (Figure 5.4, 5.5, 5.6) thus indicate that a firing-rate code suffices to recover large amounts of information.

The van Rossum Distance introduced in Chapter 4 is based on the convolutions of spike trains with a linear filter such as an alpha function. Because this procedure also

amounts to estimating the time-varying firing rate of the spike train, the distance between spike trains is simply the root-mean-square distance between the corresponding rate estimates. Interestingly, the kernels obtained from the reconstruction of natural songs roughly have a unimodal shape with a full width at half maximum of about 10 milliseconds. On the other hand, the temporal resolution parameter τ of Chapter 4, corresponding to the full width at half maximum of an alpha function, yields the best discrimination for values of $\tau = 2\text{--}10$ ms. To obtain reliable estimates of the time-varying firing rate, the width of a kernel is far more important than its precise shape (Nawrot et al., 1999). Therefore, we can re-interpret the distances used in Chapter 4: they are roughly equivalent to the mean square distances between the amplitude modulations of the original songs.

5.4.3 Use of Naturalistic Stimuli in the Literature

The application of systems-analysis methods based on white-noise signals (Marmarelis and Marmarelis, 1978) to the auditory system has a long history, see Eggermont (1993) for an overview. Studying auditory receptor neurons of *L. migratoria*, Sippel and Breckow (1983) pointed out that the response properties depend strongly on the test stimuli used, which in their case were sinusoidal and Gaussian white-noise signals. Motivated by findings of this type, various recent studies have specifically addressed the role of natural-scene statistics for the processing of acoustic stimuli. For example, Rieke et al. (1995) showed that auditory afferents in bullfrogs transmit most information about artificial stimuli whose spectrum matches that of conspecific calls. For auditory receptor neurons of grasshoppers, it is not the spectrum but the amplitude distribution of the song that is important. The importance of amplitude distributions has also been demonstrated by Attias and Schreiner (1998) in the inferior colliculus of cats showing that highest information rates are reached for stimuli whose amplitude distribution resembles that of natural sounds.

A close match of the response properties of a sensory system with the statistics of the most relevant natural stimuli has long been identified as a functionally important evolutionary design principle. The relevance of this principle for auditory systems has mainly been discussed with respect to the frequency domain (Suga, 1989). The studies of Rieke et al. (1995), Attias and Schreiner (1998) and Theunissen et al. (2000) suggest that also in the temporal domain, auditory coding has evolved to optimally exploit the regularities of behaviourally significant signals, most importantly conspecific communication signals. Our results on the neural processing of grasshopper calling songs demonstrate that the same principle holds even in an insect system. Together, these studies imply that any quantitative analysis of computations performed by auditory systems requires a proper understanding of the statistics of natural stimuli.

SUMMARY AND OUTLOOK:

In this chapter, reconstruction methods and information theory were used to analyse

the representation of different stimulus ensembles on the level of single and multiple receptor neurons. The highest rates of information transmission (180 bits/sec) and the highest coding efficiencies (40%) are obtained for stimuli that capture both the time scales and amplitude distributions of natural songs. While stimuli with typical time scales of > 40 ms can be read from single spike trains with high accuracy, faster stimulus variations can only be faithfully reconstructed when using multiple spike trains. Apparently, auditory receptor neurons are optimised to extract both the time scales and the amplitude distribution of natural songs.

It might seem inefficient to test so many stimulus ensembles on the system and compare their information rates afterwards. A simple method to speed up this procedure would be to evaluate the information transfer online, i.e., during the experiment and then change the stimulus ensembles with the aim of maximising the mutual information. This idea leads us back to a general issue already touched in Chapter 2: how can one acquire the most informative data about an input-output system?

Chapter 6

Adaptive Sampling by Information Maximisation

In this chapter, we will return to the problem of data acquisition for input-output systems as first described in Chapter 2 and re-address some of the questions raised there. In particular, we will introduce a method that allows one to collect data that is highly informative about a given system.¹

6.1 Experimental Design

Consider an arbitrary input-output system. The system could be a real physical system such as an auditory receptor neuron in which case the input would be a sound pressure wave and the output a spike train. However, we can also think of a complex model that requires time-consuming numerical simulations. The input could then be the parameters of a differential equation and the output some property of the solution such as the value of a fix point. We will be interested in analysing how changes in the input systematically influence the output. We assume that the input-output relation is not or only partially known and that we can test the system by presenting specific inputs at will and measuring the resulting outputs.

When designing a specific experimental test of the system, one will usually have some question in mind, some hypothesis about the system, or even a model. The acquired data will then allow to answer the question, test the hypothesis, or fit the model. If the research aim is very specific, the data will in general be very specific, too. Hence, data that has been acquired for one question might be of no use to investigate another question; instead, new questions usually require new data and one has to design a different test of the system.

It might therefore be useful to acquire data that depends less on a specific question while being highly informative about the input-output system. Such data could subsequently be used for several purposes. As we have seen in Chapter 2, conventional sampling strategies either sample the stimulus space too coarsely (White-noise method,

¹This chapter is adapted from Machens (2001).

Section 2.3.1) or too narrowly (Tuning curve method, Section 2.3.2). We propose that informative sampling means to sample the system in the “correct range”. To illustrate one possibility to define this range, we will look at a simple model neuron.

6.1.1 The Optimal Stimulus Ensemble

Consider a deterministic model neuron with a single-valued input x and an average firing rate y as output. We sample the system with inputs that are drawn at random from a distribution $p(x)$. Figure 6.1A shows the relation of x and y for the neuron, corresponding to a typical sigmoid tuning curve. The inputs x that elicit different firing rates y are confined to the region where the sigmoid curve has a steep slope. Useful information is encoded only about stimuli that lie within this range as no downstream mechanism will be able to distinguish stimuli falling below the threshold of the neuron or into its saturation regime. Only within the intermediate regime do changes of the input systematically influence the output.

To optimally sample the system, we propose that one should use a distribution $p_{\text{opt}}(x)$ that maximises the mutual information between input x and output y . For a deterministic system, this input distribution will lead to a uniform distribution of the outputs y .² In the continuous limit, one can analytically calculate $p_{\text{opt}}(x)$ if the system’s input-output relation is given by a monotonically increasing or decreasing function $y = f(x)$. The distributions of input and output are then related via (Papoulis, 1991)

$$p(x) = \left| \frac{dy}{dx} \right| p(y) \quad (6.1)$$

For the sigmoid tuning curve $y = r_{\text{max}}[1 + \tanh(x)]$, the optimal distribution is given by

$$p_{\text{opt}}(x) = \frac{1}{Z} \left| \frac{dy}{dx} \right| \quad (6.2)$$

$$= \frac{r_{\text{max}}}{Z \cosh^2(x)} \quad (6.3)$$

and shown as a dashed line in Figure 6.1A. The parameter Z ensures the normalisation of $p_{\text{opt}}(x)$. The system is sampled efficiently when many test inputs lie in regions where the function $\tanh(x)$ changes the most and only a few sampling points in regions where the input-output relation does not change at all. The bell-like shape of the distribution $p_{\text{opt}}(x)$ has long been recognised to be optimal for neurons with a sigmoid firing characteristic (Laughlin, 1981; Stemmler and Koch, 1999).

Although a bit more complicated, the story remains essentially the same when adding noise to the system’s output (cf. Figure 6.1B). Again we define the system’s optimal input ensemble as that ensemble of stimuli which maximises the information

²The mutual information is given by $I(x, y) = H(y) - H(y|x)$. For a deterministic system, one has a conditional entropy $H(y|x) = 0$. Hence, the maximum of the mutual information is achieved if the entropy $H(y)$ is maximal, which is uniquely given by a uniform distribution for $p(y)$ (Cover and Thomas, 1991).

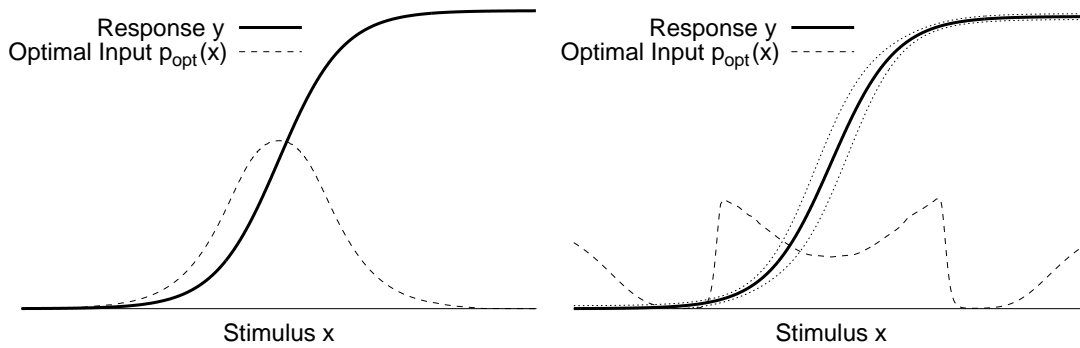


Figure 6.1: **A** Tuning curve (solid line) and optimal distribution of test inputs (dashed line) for a deterministic neuron. **B** The same for a noisy neuron. Here, the dotted lines denote the standard deviation of the tuning curve. In this case, the optimal distribution avoids stimuli that lead to very noisy responses.

transfer between input and output. An analytical solution can be found for a few special cases only; however, the optimal input ensemble can be computed using the Blahut-Arimoto algorithm which we will describe below. The ensemble obtained from this algorithm avoids stimuli that are drowned in noise as these are hardly recoverable from the output; the ensemble prefers stimuli that lead to a reliable and distinctive output.

We generalise these simple observations to the multi-variate case. The optimal sampling procedure should always prefer inputs that lead to reliable and distinctive outputs; it should downgrade the importance of inputs that share a common output. This is exactly what is achieved by maximising the mutual information. Note, that the proposed sampling strategy is not per se optimal; it simply reflects our specific interest in the system.

6.2 The Algorithm

Unfortunately, the optimal stimulus ensemble can only be computed when the full stimulus-response relation of the system is known. As the relation is usually not available beforehand, all we can do is to approach the optimal ensemble in an iterative fashion. Here, we present an algorithm that solves this task.

The basic idea is the following: define a parametric model for the input distribution $p(x)$. Draw a set of stimuli from this distribution and test them on the system. Given the system's response, determine those stimuli that contribute most to the information transfer. Finally, re-estimate the parameters of $p(x)$ to move the stimulus ensemble into those regions where the “interesting” stimuli are. The algorithm, therefore, seeks to increase the “usefulness” of the data that is acquired during an experiment.

6.2.1 Modelling the Input Distribution

Let us assume that we have a model for the input distribution $p(x)$ which depends on K parameters ϕ_i . We will write $p(x|\phi)$ with $\phi = (\phi_1, \dots, \phi_K)$ to emphasise this dependence. To guarantee the convergence of the algorithm, we require that $p(x|\phi)$ is always non-zero, i.e., $p(x|\phi) > 0$ for all x and ϕ . The model distribution expresses our prior assumptions about the structure of the optimal stimulus ensemble. A simple, unimodal model would be a Gaussian where the parameters ϕ_i are given by mean and (co)variance. Whenever sufficient amounts of data have been collected, one can refine the model and add more parameters to allow for additional structure.

6.2.2 The Iterative Algorithm

I. Initialising the stimulus ensemble: At the beginning, we need to choose initial values ϕ_i for the parameters. These values reflect a first guess of the optimal stimulus ensemble. The statistical properties of the ensemble of natural stimuli might guide us in our choice. In case of doubt, one should pick values leading to a broader distribution: such a choice avoids being too specific at the beginning. Given the values ϕ_i , we can construct the corresponding probability distribution $p(x|\phi)$.

II. Testing the system: During every iteration of the algorithm, we draw L stimuli x_i , with $i = 1 \dots L$, from the distribution $p(x|\phi)$. Each of these stimuli is presented several times to the system and the respective responses are measured. Combining these new data with the data from previous iterations, we take that altogether N stimuli have been presented and the total amount of recorded responses is given by y_{ij} with $i = 1 \dots N$ and $j = 1 \dots M_i$. Here M_i signifies that different stimuli might have been repeated a different number of times. Some of the response values might have appeared more than once. Let us assume that altogether K different response values have been recorded up to now. For the set of all different responses we write $\{y_k : k = 1 \dots K\}$.

III. Evaluating the stimuli: In the next step, we want to evaluate which of the stimuli were “good” and informative sampling points and which were “bad”. For that purpose, we define a probability distribution $q(x_i)$ on all of the tested stimuli x_i with $i = 1 \dots N$. This distribution shall weight the importance of the stimuli x_i . Initially we assume that all stimuli x_i were equally informative, setting $q_1(x_i) = 1/N$. To evaluate the stimuli, we need some estimate of the stimulus-response relation. Given the data, we can make a non-parametric estimate of the conditional probability that a response y_k was obtained from the stimulus x_i ,

$$q(y_k|x_i) = \frac{1}{M_i} \sum_{j=1}^{M_i} k(y_{ij} - y_k) \quad (6.4)$$

where the function $k(y) = 1$ for $y = 0$ and $k(y) = 0$ otherwise. Given this estimate of the stimulus-response relation, $q(y_k|x_i)$, we now seek to redistribute the probabilities $q(x_i)$ of the stimuli x_i such that the mutual information is maximised. Given the initial

distribution $q_1(x_i)$, we can use the Blahut-Arimoto algorithm (Arimoto, 1972; Blahut, 1972), which iterates the formula

$$q_{n+1}(x_i) = \frac{1}{Z} q_n(x_i) \exp \left(\sum_{k=1}^K q(y_k|x_i) \log \frac{q(y_k|x_i)}{q_n(y_k)} \right) \quad (6.5)$$

where $q_n(y_k) = \sum_{i=1}^N q(y_k|x_i) q_n(x_i)$ and Z is again a normalisation constant which here assures that $\sum_{i=1}^N q_{n+1}(x_i) = 1$. In every iteration, the Blahut-Arimoto algorithm will decrease the probability of inputs x_i whose conditional response distribution $q(y_k|x_i)$ is very similar to the total response distribution $q_n(y_k)$. The algorithm will increase the probabilities of stimuli x_i for which these distributions are very different.³ The convergence of the algorithm requires that $|1 - q_{n+1}(x_i)/q_n(x_i)| < \epsilon_1$ for all i . Upon convergence we set $q_{\text{opt}}(x_i) = q_{n+1}(x_i)$. This optimal distribution weights the importance of the stimuli tested so far.

IV. Moving the stimulus ensemble: Using the optimal distribution $q_{\text{opt}}(x_i)$, we can re-estimate the parameters ϕ_i of our input distribution $p(x|\phi)$. The new parameters should be chosen so that $p(x|\phi)$ captures the structure of $q_{\text{opt}}(x_i)$. For that purpose, we use the maximum-likelihood estimator (see, e.g., Ripley, 1996). Here, the parameters ϕ_i are determined at the maximum of the log-likelihood function,

$$\log L(x_1, \dots, x_N|\phi) = \sum_{i=1}^N q_{\text{opt}}(x_i) \log p(x_i|\phi) \quad (6.6)$$

where the probabilities $q_{\text{opt}}(x_i)$ provide the appropriate weights. For some model types, e.g. Gaussians, the maximum can be found analytically. In general, one has to compute the new parameter values ϕ_i numerically.

V. Assessing the model fit: Naturally, the quality of the fit depends on the possible distributions allowed for by our model. However, the magnitude of the deviations are not decisive in this case, but rather whether $p(x|\phi)$ still provides a stimulus ensemble that permits to obtain informative data. For the distribution $q_{\text{opt}}(x_i)$, the information transfer is given by (cf. equation 2.16)

$$I_{\text{R}} = \sum_{i=1}^N q_{\text{opt}}(x_i) \sum_{k=1}^K q(y_k|x_i) \log_2 \frac{q(y_k|x_i)}{q_{\text{opt}}(y_k)} \quad (6.7)$$

where $q_{\text{opt}}(y_k) = \sum_{i=1}^N q(y_k|x_i) q_{\text{opt}}(x_i)$. The information transfer achieved by $q_{\text{opt}}(x_i)$ should be compared to the information transfer of the model. For that purpose, we design a distribution $p_q(x_i)$, defined on the set of all tested stimuli x_i , $i = 1 \dots N$, so that $p_q(x_i)$ has the same relative values as $p(x|\phi)$,

$$p_q(x_i) = \frac{p(x_i|\phi)}{\sum_{j=1}^N p(x_j|\phi)}. \quad (6.8)$$

³This can be seen from the fact that the exponent is the Kullback-Leibler distance between the distributions $q_n(y_k)$ and $q(y_k|x_i)$ (Cover and Thomas, 1991).

Hence, the information transfer achieved by the model is given by

$$I_M = \sum_{i=1}^N p_q(x_i) \sum_{k=1}^K q(y_k|x_i) \log_2 \frac{q(y_k|x_i)}{p_q(y_k)} \quad (6.9)$$

where the distribution $q_{\text{opt}}(x_i)$ has been replaced by $p_q(x_i)$. The output distribution can be computed via $p_q(y_k) = \sum_{i=1}^N q(y_k|x_i) p_q(x_i)$. The fraction γ of the mutual information captured by $p(x|\phi)$ is then defined as

$$\gamma = \frac{I_M}{I_R} \quad (6.10)$$

and provides a measure for the quality of the model.

Steps II–V are iterated until enough data have been accumulated. Note, that due to the high dimensionality of real input spaces, one can never be sure that the iterative procedure has converged to the optimal input distribution unless all stimuli have been tested. The goal of the algorithm is, therefore, rather to increase the usefulness of the data acquisition with every iteration.

6.3 Applications

We illustrate the method using a Hodgkin-Huxley-type, class-I neuron model (Wang and Buzsáki, 1996). The voltage output V of this neuron is modelled by the following equation

$$C \frac{dV}{dt} = g_{\text{Na}} n^4 (V - E_{\text{Na}}) + g_{\text{K}} h^3 m (V - E_{\text{K}}) + g_{\text{L}} (V - E_{\text{L}}) + I + \eta \quad (6.11)$$

where g_{Na} , g_{K} , and g_{L} denote the conductances of the different ion channels; E_{Na} , E_{K} , and E_{L} the reversal potentials of the ion channels; and h , m , and n the time-dependent (in)activation variables. The function and precise values of all quantities have been given in Wang and Buzsáki (1996).

To obtain a stochastic model, we add white noise η with cut-off frequency $f_c = 1000$ Hz and a total power of $\sigma_{\eta}^2 = 16 \mu\text{A}^2/\text{cm}^4$. The input to this system is given by the injected current I , the output by the voltage V . We keep the input currents within a physiological range by constraining I to the range $-12 \mu\text{A}/\text{cm}^2$ to $28 \mu\text{A}/\text{cm}^2$.

6.3.1 Spike Count

In our first example, we use a one-dimensional parametrisation of stimulus and response. This allows us to fully comprehend the functioning of the algorithm. We employ stimuli consisting of 100 millisecond long, discretised current steps ($\Delta I = 0.5 \mu\text{A}/\text{cm}^2$) that lie in the range $I = -12 \dots 28 \mu\text{A}/\text{cm}^2$. The responses are given by the spike count C during the corresponding time window, cf. Figure 6.2. To emphasise

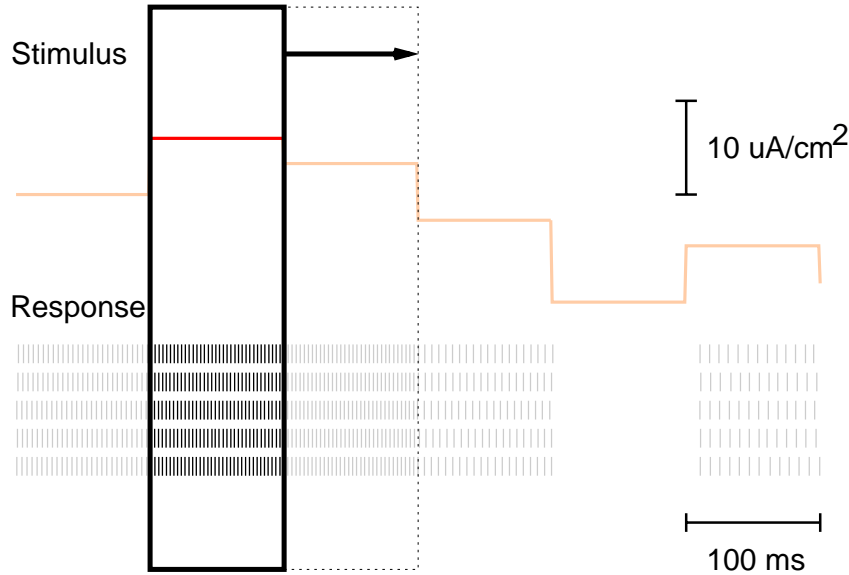


Figure 6.2: For the first example, we use a one-dimensional parametrisation of input and output. Here, the stimulus is given by constant current steps of 100 ms length, discretised to a resolution of $\Delta t = 0.5 \mu\text{A}/\text{cm}^2$ and the response is given by the total spike count during the corresponding time window. The figure shows six of these current steps and the corresponding spike trains, obtained from five repetitions of the stimulus. An exemplary “window” is highlighted (black frame).

the meaning of input and output, we use the notation I for the input current and C for the spike count, instead of the general terms x and y used in the preceding section.

For this simple problem, we can numerically calculate the stimulus-response relation $p(C|I)$. The resulting relation of spike count versus current is displayed in Figure 6.3A. The model neuron starts spiking for inputs $I > 0 \mu\text{A}/\text{cm}^2$. With increasing input, the spike size decreases until, above $I > 25 \mu\text{A}/\text{cm}^2$, the spikes have shrunk to a size that makes it impossible to detect them in the noisy voltage output. The resulting uncertainties in the spike count estimate result in a broad distribution of output values.

Given the full distribution $p(C|I)$, we can compute the optimal input distribution $p_{\text{opt}}(I)$ using the Blahut-Arimoto algorithm. The result is depicted by the vertical bars in Figure 6.3B. The shape of the distribution reflects the properties of the stimulus-response relation. Similar to the deterministic case, equation (6.2), the optimal input distribution roughly corresponds to the slope of the stimulus-response relation, cf. Figure 6.3A. Note, that there is a slight increase in the probabilities of stimuli far below threshold ($I \leq -10 \mu\text{A}/\text{cm}^2$). These stimuli will almost certainly result in a zero spike count response. At the same time, stimuli closer to threshold ($I \approx -9 \dots -1 \mu\text{A}/\text{cm}^2$) will less reliably produce no response. As the optimal input distribution prefers inputs that are more reliable, the stimuli closer to threshold are neglected.

Knowing the optimal distribution, we can now test the iterative algorithm to see how it performs. For that purpose, we model the input distribution by a Gaussian,

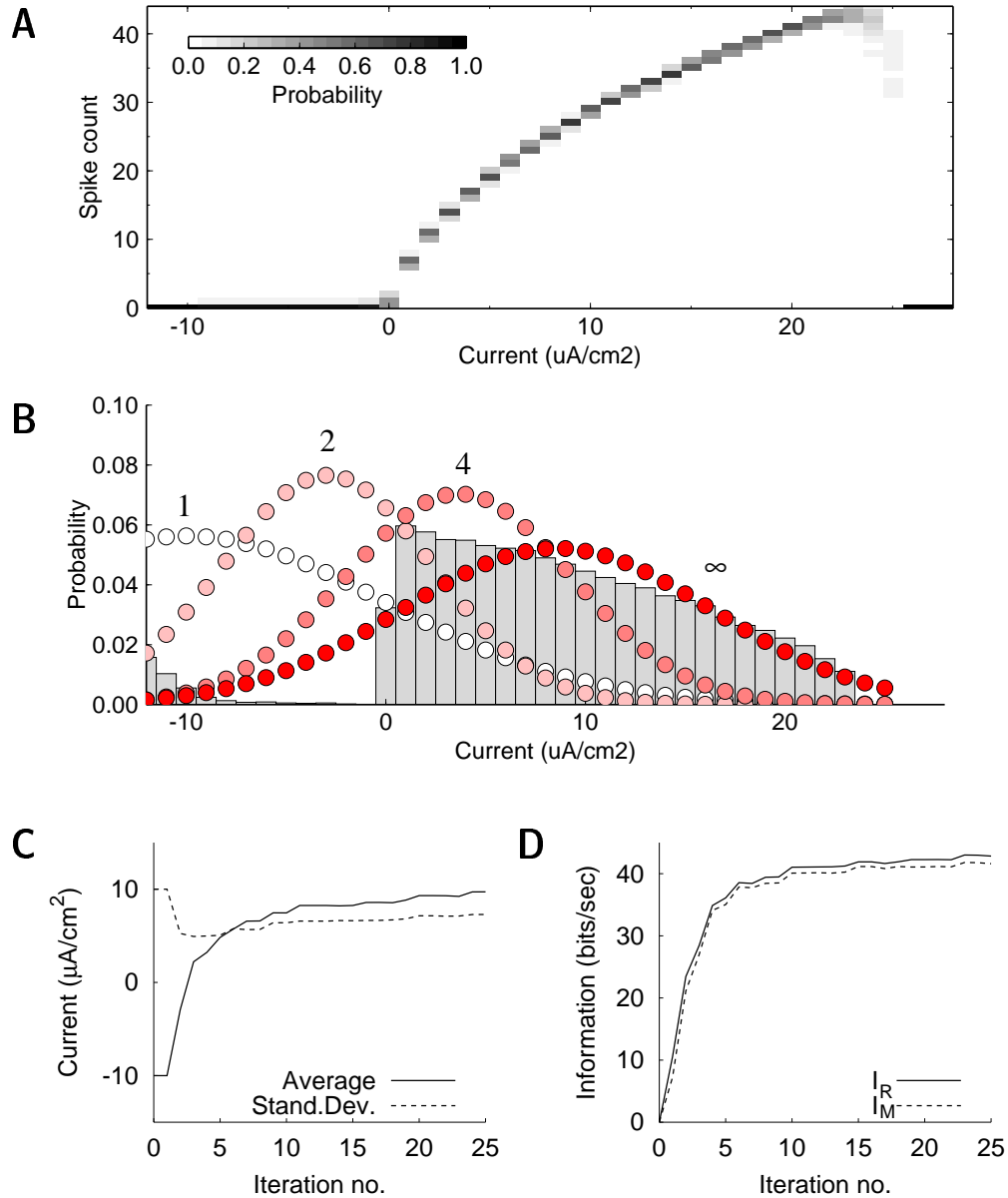


Figure 6.3: Search for the optimal ensemble for a system with one-dimensional input and output. **A** Colour plot of the conditional probability distribution $p(C|I)$ with spike count C and input current I . **B** Approaching the optimal distribution. Shown are the initial stimulus distribution (1), the distributions of the iterations (2) and (4), as well as the final distribution (∞) which is reached after about 20 iterations. The optimal distribution is shown as grey bars; the final Gaussian (∞) has the same mean and variance as the optimal distribution. **C** Change of the model parameters, mean and standard deviation, during the iterations. **D** Increase of information rate for the simulation data (I_R) and the Gaussian model (I_M). Note that the information rate of the simulation data converges to the maximum information rate.

parametrised by the mean ϕ_1 and the standard deviation ϕ_2 . As initial values, we choose $\phi_1 = -10 \mu\text{A}/\text{cm}^2$ for the mean and $\phi_2 = 10 \mu\text{A}/\text{cm}^2$ for the standard deviation. The Gaussian distribution given by these parameter values is shown in Figure 6.3B. The distribution does extend into the correct range, however, it is inclined heavily towards negative values.

At each step of the iteration, we randomly draw ten stimulus values I_i from the input distribution. To keep these input currents within the physiological range, we will map inputs that fall out of this range onto the margin values $I_{\min} = -12 \mu\text{A}/\text{cm}^2$ and $I_{\max} = 28 \mu\text{A}/\text{cm}^2$. Each input is presented five times to the system, obtaining the responses C_{ij} . The data collected suffice to give a rough histogram estimate of the distribution $q(C_k|I_i)$ describing the stimulus-response relation, cf. equation (6.4). In turn, we can re-evaluate the stimuli I_i by determining the probabilities $q_{\text{opt}}(I_i)$ which allow us to re-estimate the model parameters by maximising the log-likelihood function, equation (6.6). For a Gaussian model distribution, the parameter values at the maximum are given by

$$\begin{aligned}\phi_1 &= \sum_{i=1}^N I_i q_{\text{opt}}(I_i) \\ \phi_2 &= \left[\sum_{i=1}^N (I_i - \phi_1)^2 q_{\text{opt}}(I_i) \right]^{\frac{1}{2}},\end{aligned}\quad (6.12)$$

where N denotes the total number of stimuli tested so far. The equations yield the average and standard deviation for the input distribution in the next iteration.

The Gaussian model distributions are displayed in Figure 6.3B for the first few iterations. The input distributions move relatively fast towards the correct range; after about 20 iterations, both average and standard deviation have approximately converged to their final values as shown in Figure 6.3C. The large changes in the input statistics right after the start are reflected in the fast growth of the information rate during the first 5–10 iterations, cf. Figure 6.3D. After about 10 iterations, changes in the inputs statistics lead to only minor increases of the information rate. Although the Gaussian distribution obtained after convergence of the algorithm does not resemble the optimal input distribution $p_{\text{opt}}(I_i)$, given by the vertical bars in Figure 6.3B, both distributions lead to approximately the same information transfer (Figure 6.3D). Hence, the landscape of the mutual information with respect to the input distribution seems to be relatively flat around the maximum; the precise shape of the final input distribution is not that important as long as the distribution covers the correct range of inputs ($I = 0 \dots 25 \mu\text{A}/\text{cm}^2$).

In general there will be other Gaussian distributions which might achieve a slightly higher information transfer than the final Gaussian distribution. Instead of computing the Gaussian distribution that leads to the highest information transfer, we have obtained a Gaussian that provides the best fit to the optimal distribution. In particular, the final Gaussian distribution has the same mean and variance as the optimal distribution. This property of the final distribution stems from the maximum-likelihood estimation, equation (6.6) and is well-known in statistics (Ripley, 1996).

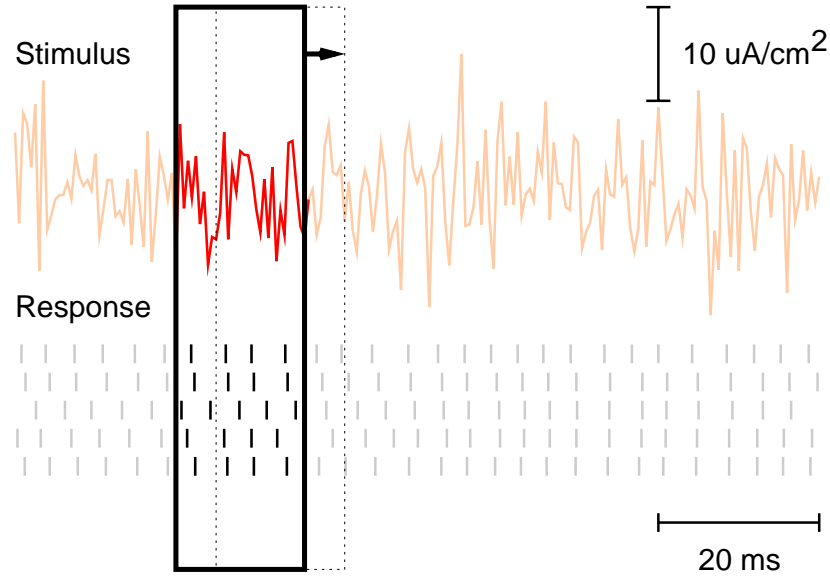


Figure 6.4: The second example uses a multi-dimensional input and output. Here, the stimulus is discretised in time steps of 0.5 ms and the response in time steps of 1 ms. The input is given by the current values within overlapping time windows of 16 ms length and the output is given by the spike times during the corresponding window.

6.3.2 Spike Timing

The second example uses time-varying, stationary stimuli, discretised in time steps of 0.5 ms and one second long. To keep the dimensionality of the stimulus space within a manageable range, we slide a 16 ms-long window across the stimulus trace and use the 32 current values within each window as input vector $I_i = (I_{i,1}, \dots, I_{i,32})$, cf. Figure 6.4. Moving the window in steps of 1 ms, each trace provides us with $1000 - 16 = 984$ vector-valued inputs I_i . For each of these inputs I_i , the output C_{ij} is assembled by the spike times, discretised in time steps of 1 ms, during the corresponding window. Hence, each input consists of 32 real-valued numbers and each output of 16 numbers whose values are either zero for no spike or one for a spike. Note, that we do not explicitly discretise the stimulus space; we will instead assume that every input I_i is unique. As before, we will map input values falling out of the physiological range onto the bordering values $I_{\min} = -12 \mu\text{A}/\text{cm}^2$ and $I_{\max} = 28 \mu\text{A}/\text{cm}^2$.

The large size of the input space prevents any computation of the real optimal input distribution. However, we can again seek to model this distribution by a Gaussian. As the stimulus is stationary, its power spectrum completely describes the covariance of the Gaussian input distribution (see, for instance, Rieke et al., 1997). Given the 32-dimensional input space, we therefore need 16 parameters ϕ_i to model the power spectrum and the average. Given the estimate $q(C_k|I_i)$ of the stimulus-response relation, we obtain the optimal weighting of inputs, $q_{\text{opt}}(I_i)$. Maximising the log-likelihood

function then results in the following equations for re-estimating average and power spectrum,

$$\begin{aligned}\phi_1 &= \sum_{i=1}^N \bar{I}_i q_{\text{opt}}(I_i) \\ \phi_l &= \sum_{i=1}^N A_{il}^2 q_{\text{opt}}(I_i) \quad \text{for } l = 2 \dots 16\end{aligned}\quad (6.13)$$

where \bar{I}_i denotes the average of the i -th input and A_{il} the amplitude of the l -th Fourier coefficient. The re-estimated parameters of the new distribution can in turn be used to create a new one-second long stimulus trace. The stimulus creation follows the procedure outlined in Appendix D.

For the numerical simulations, each one-second long stimulus was repeated 20 times. The number of responses per stimulus provides a sufficiently good estimate of the conditional probabilities $q(C_k|I_i)$. The initial distribution has a flat power spectrum with a cut-off frequency of $f_c = 1000$ Hz which, according to the Nyquist theorem, is the upper limit for a discretisation of 0.5 ms (Press et al., 1992). However, both the overall power $\sigma^2 = \sum_{i=2}^{16} \phi_i f_c / 16$ and the average of the initial distribution were varied to analyse how the algorithm depends on the initial conditions. Subsequently, we will analyse three different initial conditions that capture more or less of the correct input regime.

The first initial condition (Figure 6.5AB, solid line) has an average of $\phi_1 = 0 \mu\text{A}/\text{cm}^2$ and a standard deviation of $\sigma = 10 \mu\text{A}/\text{cm}^2$. Since the average is equivalent to the neuron's threshold, about 50% of the current values drawn from this distribution fall below threshold. Moreover, with a standard deviation of $10 \mu\text{A}/\text{cm}^2$, the stimulus will rarely reach the current values needed for very high firing rates, cf. Figure 6.3. Consequently, to cover the relevant input range, the algorithm immediately shifts the input distribution towards higher current values by increasing the average, cf. Figure 6.5A. This fast growth in average is accompanied by a slower growth in the standard deviation which levels off at around $\sigma \approx 15 \mu\text{A}/\text{cm}^2$ as shown in Figure 6.5B. While the average, therefore, reaches roughly the same value as in the previous case with a spike count output, $\phi_1 \approx 10 \mu\text{A}/\text{cm}^2$, the standard deviation is much higher.

Before interpreting the results, we want to analyse a second initial condition, cf. Figure 6.5AB (dotted line). Here, we start with a very narrow distribution, featuring a standard deviation $\sigma = 1 \mu\text{A}/\text{cm}^2$ and with an average $\phi_1 = 20 \mu\text{A}/\text{cm}^2$ located in an input regime that leads to very high firing rates. The small standard deviation of the initial stimulus ensemble does not leave much maneuvering space for the parameter re-estimation as this initial distribution explores only a relatively small area in input space. Consequently, the algorithm discovers only slowly, but steadily, that input values outside this area are quite valuable, which leads to an increase in the standard deviation. The average moves even more slowly towards smaller values; the reason is that initially not much information can be gained by moving the average if the distribution is very narrow.

A third case starts with an initial condition of $\phi_1 = -6 \mu\text{A}/\text{cm}^2$ and $\sigma = 10 \mu\text{A}/\text{cm}^2$ which is well outside the correct range, cf. Figure 6.5AB (dashed line) and Figure 6.6A.

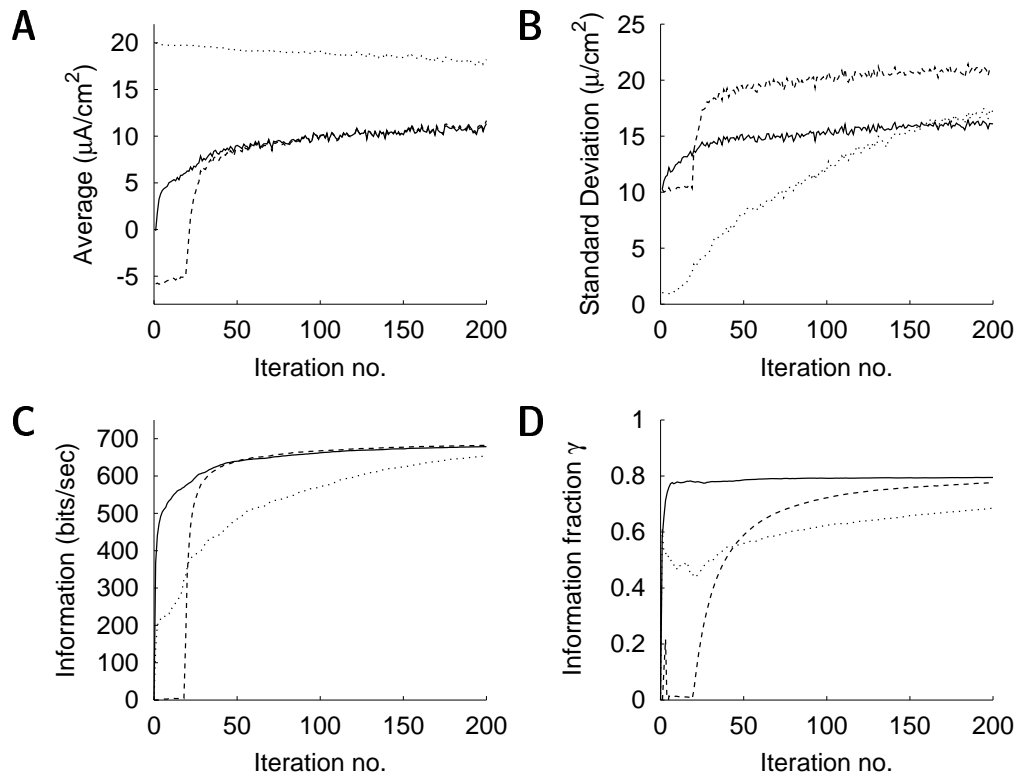


Figure 6.5: Three different initial conditions. **AB** Evolution of average ϕ_1 and standard deviation σ . The parameters move the stimulus ensembles into the correct regime although there seems to be no clear convergence point. **C** In all cases, the information rate saturates around 670 bits/sec. **D** The fraction γ of the information transfer captured by the model $p(x|\phi)$ reaches about 80%. Note that an initially “bad” model can get better if more data is available.

Although the initial distribution does reach slightly above threshold, the high cut-off frequency prevents spiking: a single current value above threshold during 0.5 ms does not suffice to elicit a spike. During the first five iterations, the stimulus does not cause any spikes. This situation corresponds to a worst-case scenario because, with only one output value (“no spikes”), the Blahut-Arimoto algorithm cannot decide which stimuli have been “better” than others. Hence, up to random fluctuations, the parameters ϕ_i do not change from one iteration to the next. Fortunately, a single spike appears during some of the stimulus traces after about five iterations. Given the shifting time window, there are now altogether 17 possible output values: “no spike” and 16 versions of “one spike”. Subsequently, the statistics of the model distribution moves into the direction of the statistics of the inputs I_i that caused the spike. Still, the impact of that movement is not large. It takes about 15 more iterations until the algorithm has really tracked the correct input range; a downright explosion of the information rate follows.

We can look at the input distribution after 50 iterations to see which stimuli have been most important so far. For that purpose, we sum over the probabilities $q_{\text{opt}}(I_i)$ of all stimuli that compose one stimulus trace. Figure 6.6B shows that the stimuli that

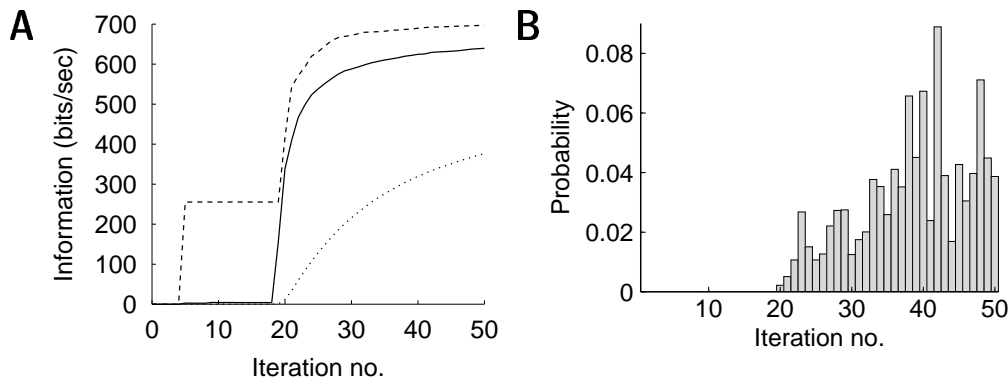


Figure 6.6: Bad initial conditions, $\phi_1 = -6 \mu\text{A}/\text{cm}^2$ and $\sigma = 10 \mu\text{A}/\text{cm}^2$. **A** No response is obtained during the first 4 iterations. At the fifth iteration a spike occurs in some of the stimulus repetitions. Due to the shifting time windows, there are now $16+1$ possible output values, such that $I_{\max} = \log_2 17 / (16 \text{ ms}) \approx 255 \text{ bits/s}$ as shown by the dashed line which denotes the upper bound on the mutual information, $\log_2 K$. In turn, the stimuli that lead to the spike are emphasised, resulting in an increase in stimulus power. When the correct stimulus range has been found, an explosion in information transfer I_R follows (solid line). The dotted line shows the information rate I_M of the model. **B** The data obtained during the first 20 iterations no longer influence the shape of the input distribution after 50 iterations.

have been used during the first 20 iterations are completely neglected; they do no longer influence the properties of the input distribution.

In all cases, the information rate I_R saturates around 670 bits/s as can be seen in Figure 6.5C. Depending on the initial condition, this value can be reached after about 25 iterations.⁴ Here, the simulation with the very narrow initial distribution ($\sigma = 1 \mu\text{A}/\text{cm}^2$) fares much worse because a lot of iterations are needed before the standard deviation covers the relevant range within input space. The fraction of the mutual information captured by the model reaches 80%, cf. Figure 6.5D. Note, though, that a model that is initially bad might well become better when more data is available.

A good choice of the initial conditions is therefore a necessity for the algorithm. If the initial distribution samples a system so ineffective that one does not obtain a non-zero response, then the algorithm has no possibility to move into any direction.

Although the mutual information reaches approximately the same value independent of the initial conditions, average and standard deviation have not converged to the same values during the 200 iterations. This suggests that the landscape around the maximum is very flat. Whether there exists a single best Gaussian approximation to the optimal ensemble remains therefore unknown. However, this need not worry us as all of the final ensembles are equally good in achieving a high rate of information transfer.

Figure 6.7A shows the power spectra of the final stimulus ensembles for the three initial conditions. The increase in standard deviation can be attributed to a growth of the power of frequencies below 500 Hz. This cut-off approximately coincides with the

⁴This corresponds to about 12 minutes of data acquisition, a time span well available for electrophysiological experiments.

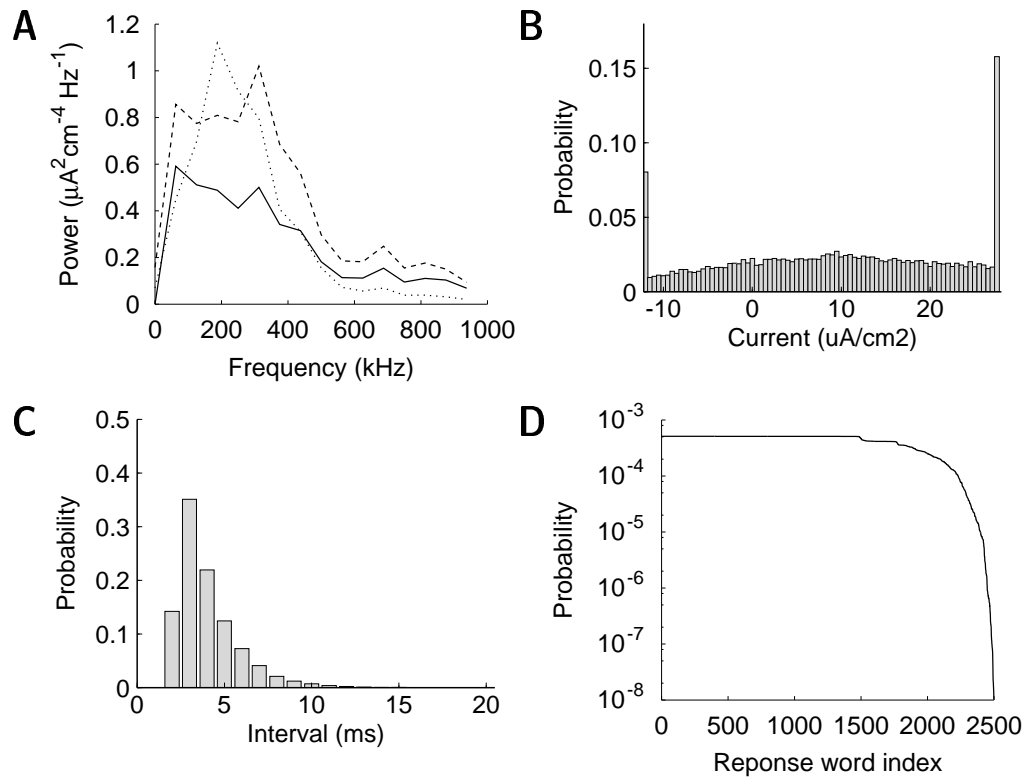


Figure 6.7: Properties of the final stimulus ensembles. **A** The final power spectra for the three initial conditions. Due to the low-pass properties of the neuronal membrane, the final stimulus ensembles neglect frequencies beyond 500 Hz. **B** The large standard deviation of the final stimulus ensemble leads to a two-tailed distribution of currents as current values falling out of the physiological range are mapped onto the boundary values. This distribution of currents seeks to drive the neuron maximally. **C** The refractory period of the model neuron prevents a Poisson-like spike rate. Nevertheless, the interspike interval histogram shows an exponential tail that is characteristic for Poisson spike trains. **D** Moreover, about 60% of the 16-bit responses are uniformly distributed.

maximum firing rate of the neuron.⁵ The corresponding dimensions in stimulus space are degenerate and the statistical parameters ϕ_i that represent the power along these dimensions can assume any value without influencing the information transfer.

The large values of the standard deviation σ lead to an effective distribution of stimuli as shown in Figure 6.7B. The stimulation seeks to drive the neuron maximally by accumulating many inputs at the bordering values. Such a distribution of inputs overrides the additive noise η of the model neuron, and stimuli drawn from this distribution produce very reliable responses.⁶

⁵This result was also independent of the specific discretisation of stimulus and response.

⁶Note, that the effective distribution of inputs is therefore highly non-Gaussian: The high standard deviations drive the tails of the Gaussian distributions outside the boundary values $I_{\min} = -12 \mu\text{A}/\text{cm}^2$ and $I_{\max} = 28 \mu\text{A}/\text{cm}^2$. As noted before, input values dropped out of this range are mapped onto the boundary values, causing the non-Gaussian shape of the effective input distribution.

The high information rates reached are reflected in the interspike-interval distribution shown in Figure 6.7C. After the refractory period of 1–2 ms, the distribution falls off exponentially like a Poisson distribution. Correspondingly, the 16-digit binary outputs are uniformly distributed over a wide range, cf. Figure 6.7D.

6.4 Neurophysiological Interpretations of the Best Ensemble

The original goal of this chapter has been to find an ensemble of inputs that optimally explores the relevant regions of the input-output system under study. For the investigation of sensory systems, we can re-interpret this ensemble. A particular neuron within a sensory system, acting as a feature detector, encodes information about specific aspects of a stimulus. Usually, only a very specific subset of all possible stimuli will share these aspects. Since the optimal stimulus ensemble covers all stimuli that the neuron conveys information about, the optimal ensemble is focused to this specific subset. Hence, the optimal stimulus ensemble can be interpreted as representing, in a probabilistic manner, that region of stimulus space that a particular neuron is “looking” at. In contrast to previous online algorithms such as Alopex or Simplex (Harth and Tzanakou, 1974; Nelken et al., 1994), we are not looking for a single optimal stimulus but for an optimal stimulus ensemble that covers the complete range of relevant inputs.

Hence, if a neuron conveys information in its average firing rate, as suggested in Section 6.3.1, then it best encodes slow-varying current values in the range $I = 0 \dots 23 \mu\text{A}/\text{cm}^2$. The synaptic input of a neuron should therefore drive the neuron in the corresponding range. If a neuron encodes its information in the precise timing of spikes, then the synaptic input should be of a more binary nature, i.e., either fully exciting the neuron or fully inhibiting it.

Note that we do not assume that a particular neuron indeed receives input from the optimal stimulus ensemble. Instead, we simply state that the neuron is “looking” into that particular direction. Indeed, both experimental and theoretical work has suggested that neurons are able to redirect their “gaze” (Stemmler and Koch, 1999; Fairhall et al., 2001) in order to track the ensemble of stimuli present in their natural environment if this ensemble is “moving” within stimulus space.

Although the optimal stimulus ensemble provides a first hint, it does not necessarily uncover which aspects of a stimulus are encoded by a sensory system. For instance, an invariance of the system towards a specific stimulus aspect means that the system’s response does not change when the given stimulus aspect is varied. From the point of view of the optimal ensemble, the respective stimulus dimensions along which this stimulus aspect varies, are degenerate: the total probability assigned to the respective stimulus subset can be distributed in an arbitrary way on the subset.

Once the optimal stimulus ensemble has been found, it might therefore be of specific interest to explore these degenerate stimulus subsets by fixing a specific response. Applying methods from the machine learning theory (Tenenbaum et al., 2000; Roweis

and Saul, 2000), these iso-response stimuli might subsequently be used to deduce the invariance properties of the system under study, see also Gollisch et al. (2001).

6.5 Problems and Refinements

6.5.1 Estimating the Input-Output Relation

There are two crucial issues to the algorithm. One issue concerns the choice of an appropriate model for the input distribution; such a choice is always problem-specific and will not be discussed here. The other issue concerns the estimation of the input-output relation as described by the distribution $q(y_k|x_i)$. This estimation is closely related to the measurement process. To obtain a reliable, non-parametric estimate of this distribution for a given stimulus x_i , cf. equation (6.4), the stimulus must be repeated often enough. In particular, it should be repeated so frequently that several of the measured outputs y_{ij} are equivalent.

Alternatively, one might seek to model $q(y_k|x_i)$ by a parametric distribution. This strategy amounts to assuming additional structure on the output space, for instance, about the similarity of different outputs. With a judicious choice of a model, one might reduce the number of stimulus repetitions necessary to estimate $q(y_k|x_i)$ for a specific stimulus x_i . However, if the model is not a good choice, one risks to mis-estimate the stimulus-response relation which, in turn, leads to a mis-estimation of the parameters of the input distribution.

6.5.2 Quantising Input and Output

The algorithm will only work if it gets a good representation of the possible outputs of the system during the iterative process. For instance, if the responses measured are all the same, e.g., no spikes as in one of the examples above, then the input distribution cannot be changed. Often, the reason for this problem is that the set of relevant inputs is very small compared to the set of all possible inputs. Unless the input distribution extends into the correct regime from the outset on, without being too broad, the inputs drawn will usually miss the correct regime.

An equivalent problem is posed if the set of possible outputs is very large. In this case, every presentation of an input could elicit a new, different output. Without any overlap on the response side, the algorithm will not be able to change the input distribution. Since the original non-parametric estimation of the stimulus-response relation does not assume any relation between the output values, we do not have any method to grade or rank the output values.

Both are generic problems of input-output analysis. One possible solution is to reduce the size of input or output space. For instance, one can quantise either input or output space by one of the methods described in Section 2.4.2. Ideally, the quantisation should not decrease the potential mutual information between input and output. In

this respect, future developments might seek to systematically incorporate the clustering methods of Tishby et al. (1999) and Dimitrov and Miller (2001) into the online algorithm. Given a specific size of input or output space, these methods choose a quantisation that preserves as much of the mutual information as possible.

One might also cope with the size of the output space by using a parametric model for the stimulus-response relation, as described above. However, while this might help to overcome the data acquisition problem, one is still prone to be quickly limited by the abilities of present-day computers.

6.5.3 Systematic Expansions

Essentially, the goal of the algorithm is to increase the value of every new data point by using all the information available. At some point, the number of stimuli that have been tested will greatly surpass the number of statistical parameters used to characterise the optimal distribution. Unless the distribution $p(x)$ is in excellent agreement with the restricted distribution $q(x_i)$, i.e., $\gamma \approx 1$, one should refine the model by increasing the number of statistical parameters to allow for additional structure. In turn, this could lead to a better focusing of the input distribution onto the relevant range.

In this respect, it would be highly desirable to do this in a systematic way. This might be achieved by using expansions of probability distributions, e.g., expansions based on orthonormal functions (Ripley, 1996).

6.5.4 Incorporation of Constraints

In many cases, there might be specific constraints on the input. For instance, the power of an acoustic stimulus is always limited by the properties of a specific setup. Various investigations have suggested that neural systems try to decrease the amount of energy expenditure (for instance, Balasubramanian et al., 2001; Schreiber et al., 2001). Since every spike costs a certain amount of energy, decreasing energy expenditure means avoiding high spiking rates whenever possible. Hence, one might also want to consider the possibility of introducing energetic constraints for the system.

Such constraints can be directly incorporated into the Blahut-Arimoto algorithm by appropriately modifying equation (6.5), see Blahut (1972).

SUMMARY AND OUTLOOK:

To optimally sample the most interesting aspects of an input-output system, one might search for the ensemble of stimuli that maximises the information transfer between input and output. Within this chapter, we have proposed an algorithm that automatically adjusts a stimulus ensemble during the course of an experiment to focus on the area of stimulus space that is most relevant for a given system. For that purpose, the algorithm needs a model of the stimulus ensemble. In an iterative procedure, the algorithm uses whatever data is available to estimate the input-output relation and then uses

this estimate to guess the model parameters of the optimal stimulus ensemble. From a neurophysiological perspective, the optimal stimulus ensemble represents the region in stimulus space that a particular sensory system encodes.

Chapter 7

Conclusions

Within this thesis, we were mainly interested in two questions: (1) How can one employ natural stimuli to relate the function of a sensory system to an animal's behaviour? (2) Which stimuli are best suited to sample a particular sensory system? For that purpose, we have investigated receptor neurons in the grasshopper auditory system. To see what we have learnt, we will summarise the non-introductory Chapters 4–6.

In Chapter 4, we investigated the encoding of a small set of natural stimuli, the calling songs of male grasshoppers. These songs are of eminent behavioural relevance, as female grasshoppers use the information within these songs to choose among males. We could show that auditory receptor neurons of females achieve an amazing resolution of the amplitude modulation patterns of the male songs. While other properties of the songs seem to be of minor behavioural relevance, the amplitude modulation pattern conveys important information about the state of a particular male. If a male grasshopper exhibits some degree of asymmetry, for instance due to loss of a hindleg, the amplitude modulation pattern produced significantly deviates from the one of a healthy male. Hence, we were able to show that the information available to a female allows her to choose among males and that the information represents important properties of the males.

In Chapter 5, the aim was to broaden the experimental test of the auditory receptor neurons by designing stimulus ensembles that sample large regions of the stimulus space. These stimulus ensembles differed in the extent to which they incorporated features of the natural songs. We discovered that one of these stimulus ensembles best explored the coding regime of a given receptor. Within this ensemble, the amplitude modulation varies on time scales of 2.5 ms and exhibits large fluctuations. We could show that this finding holds on both the single-cell and the population level. On the population level, the differences between the stimulus ensembles are less pronounced because the population helps to convey information about stimulus ensembles which are not well encoded by single cells. In particular high-frequency components and shallower amplitude modulations are resolved much better on the population level; mainly, because the population helps to average out the stochasticity of the neural responses. The large depth of the amplitude modulations featured by the optimal stimulus ensemble is a characteristic of all grasshopper songs, caused by the repetitions of sound and

pauses. The time scales of the optimal ensemble resolve all relevant features up to the gaps in the songs of one-hindlegged grasshoppers, suggesting that the receptor neurons are matched to a regime in stimulus space that is highly informative about the state of the male singer.

In Chapter 6, we point out how the optimal stimulus ensemble can be found in a systematic and automated way. Here the idea was to let the system itself show the path towards the optimal stimulus ensemble. Given a first guess of the optimal ensemble, we developed an online algorithm that alternately uses stimuli from this ensemble to record data and estimate the stimulus-response relation; and then uses the estimate of the stimulus-response relation to update the guess of the optimal ensemble. We tested this method on a stochastic model neuron and showed that the algorithm is well suited to approximate the optimal ensemble, thus improving the performance of experiments such as those in Chapter 5 within the limited time of an electrophysiological recording session.

We have shown that the properties of auditory receptor neurons are well matched to the behaviourally relevant aspects of natural grasshopper songs. This was only possible by moving beyond simple sine-wave and noise stimuli and testing the system with complex natural or naturalistic stimuli. Most likely, a close match between behaviourally important aspects of natural stimuli and the properties of sensory systems is an evolutionary design principle. Investigations that make explicit use of the properties of natural stimuli can therefore help to further our understanding of sensory systems. Yet even if the properties of the behaviourally relevant stimuli are not known, one can potentially extract them from a given system using our novel online algorithm. The increase of computational power permits to re-evaluate the stimulus presentation in neurophysiology. In the future, automated methods might complement more traditional approaches and, hopefully, lead to new and exciting insights.

Appendix A

Stimulus Design

Here we explain the technical procedures used for the design of stimuli in Chapters 4–6. The stimuli used in Chapters 4 and 5 were generated by Martin Stemmler (Humboldt-University at Berlin).¹

A.1 Rescaling of Calling Songs

The calling songs of male grasshoppers were rescaled to a common syllable-pause duration of 100 ms in the following way: The amplitude modulation (AM signal) of the songs was computed using the Hilbert transform (Haykin, 1994) and used to demodulate the acoustic signal. A uniform syllable-pause duration was then obtained by rescaling the AM signal in time. The spectrum was locally adjusted by combining the rescaled AM signal with a Gaussian signal carrier whose time-resolved spectrum was matched to that of the demodulated original signal. Hence, this procedure also allowed to fill the AM signal of one song with the time-resolved spectrum of another song.

A.2 Design of Artificial Stimuli

Within Chapters 5 and 6, we designed random stimuli that obey a certain frequency composition. These stimuli were created by choosing Fourier components with the specified spectral amplitudes but random phases. Using the inverse Fourier transform, a Gaussian signal with the prescribed spectral properties is generated.

To obtain stimuli that additionally feature a specific non-Gaussian signal distribution (such as the LMD stimuli in Chapter 5), signal values drawn from the non-Gaussian distribution were sorted into increasing order; the same number of random, Gaussian variables were also sorted; finally, the Gaussian set of signal values was mapped in a one-to-one fashion onto the non-Gaussian set. Without corrective measures, this procedure could generate nontrivial higher-order correlations among the phases and distort the spectrum (Li and Hammond, 1975); our correction scheme involved an iterative

¹This appendix is adapted from Machens et al. (2001b, 2001c).

procedure, alternately shaping the spectrum and then mapping the cumulative distribution of the artificial variables onto that of the target distribution. In the investigated cases, however, the distortion of the spectrum was negligible, so that the simple one-to-one mapping nonlinearity sufficed for the transformation. This also allowed us to use the reverse transformation back into Gaussian stimuli for the purpose of calculating information-theoretic quantities, cf. Appendix D.

Appendix B

Electrophysiology

For completeness, we here present the electrophysiological methods. The experiments were performed by Hartmut Schütze and Petra Prinz (both at Humboldt-University at Berlin).¹ The experimental protocol differed in minor respects for the investigations of Chapter 4 and 5. In general, however, the experiments followed these lines:

All experiments were done on adult *Locusta migratoria*. Anatomical and physiological properties of their auditory periphery closely match those of *Chorthippus biguttulus* (Ronacher and Krahe, 2000). Unlike the latter, however, *L. migratoria* is well suited to the demanding experimental protocol needed for the investigations. After removing of the legs, the wings and the dorsal part of the thorax, the metathoracic ganglion and auditory nerve were exposed. Auditory receptors were recorded intracellularly in the auditory nerve with standard glass micro electrodes (borosilicate, 50–110 M Ω , filled with 1 M KCl solution).

Neural responses were amplified (NPI BRAMP-01) and recorded by a data acquisition board (National Instruments PCI-MIO-16E-1) with a sampling rate of 10 kHz. Spikes were detected automatically by a custom made software. Acoustic stimuli were generated by the same software and played by the same data acquisition board, with sampling rates between 200 and 250 kHz. The experiments were performed in a Faraday cage lined with sound attenuating foam to reduce echoes. The preparation was placed between two speakers (D28-2, Dynaudio, Skanderborg, Denmark) which were oriented towards the animal's ears at a distance of 35 cm. The temperature during the measurements was 30°C (Chapter 4) or 33–35°C (Chapter 5).

For Chapter 4, responses from $n = 22$ receptor neurons were recorded; for $n = 7$ neurons stable responses to 8–16 repetitions of all songs and artificial song variants were obtained. Only these were included in the analysis, although the other (partial) recordings confirmed the results.

For Chapter 5, responses were recorded from $n = 27$ receptor neurons, and each stimulus class was presented to up to 10 neurons. Stimuli were presented with peak intensities ranging from 3 dB to 21 dB above the threshold of individual receptors, entailing firing rates from 40 Hz up to 160 Hz.

¹This appendix is adapted from Machens et al. (2001b).

Appendix C

Quantisation in Response Space

C.1 Spike Train Distances

Distance measures allow one to quantify the similarity between spike trains. In the following, we will review the two types of distances used in Chapter 4, show their similarity, and discuss their limitations.

C.1.1 Victor-Purpura Distance

Distance measures based on the explicit alignment, insertion and deletion of spikes have been proposed by Victor and Purpura (1996, 1997). To compute the distance between two spike trains, one transforms one spike train into the other by a series of steps each of which carries a certain “cost” (cf. Fig. C.1): insertion and deletion of spikes cost one unit each, while a shift in time of length Δt costs $q|\Delta t|$ units where q is a free parameter that weights how important the precise timing of spikes is. The distance $D_{VP}(q)$ between two spike trains is defined as the minimal cost of transforming one spike train into the other.

For a cost factor $q = 0$, spikes can be shifted without any costs. After aligning as many spikes as possible by shifting, the distance between two spike trains is simply given by inserting or deleting the remaining spikes. If n_1 and n_2 denote the number of spikes of the two spike trains, we obtain

$$D_{VP}(q = 0) = |n_1 - n_2|. \quad (\text{C.1})$$

In the case $q \rightarrow \infty$, shifting spikes becomes so expensive that the cheapest transformation is to simply delete all spikes of the first spike train and insert all spikes of the second spike train at the appropriate places. Only coincident spikes do not have to be deleted and inserted again. Therefore, the minimal cost is given by the total number of non-coincident spikes,

$$D_{VP}(q \rightarrow \infty) = n_1 + n_2 - 2c_{12}, \quad (\text{C.2})$$

where c_{12} is the number of coincident spikes. Hence, while $D_{VP}(q = 0)$ ignores the precise timing of spikes, $D_{VP}(q = \infty)$ considers any difference in time of occurrence to

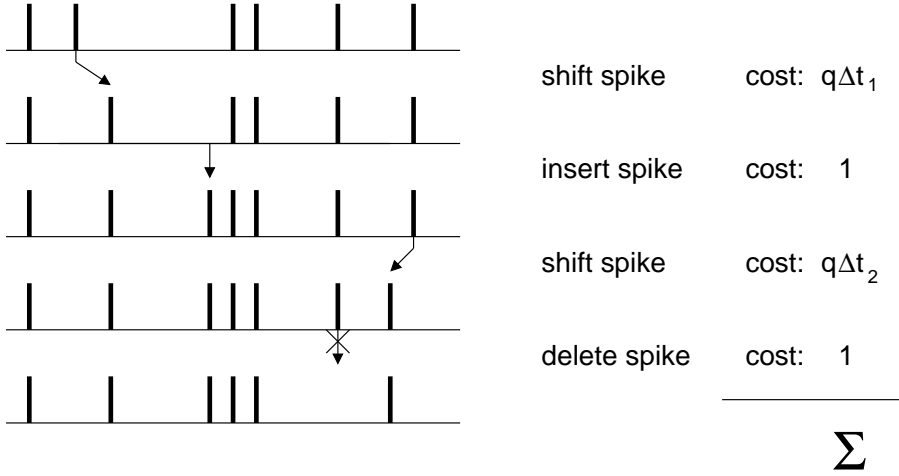


Figure C.1: The Victor-Purpura Distance D_{VP} . The distance is computed by a stepwise transformation of one spike train into another. Each step (inserting a spike, deleting a spike, shifting a spike) is associated with a cost; the distance is given by the total cost of the “cheapest” transformation possible.

constitute a different spike.

In general, the distance can be calculated by an iterative equation. For a fixed q , let $G_{i,j}$ be the distance between the first i spikes of the first spike train and the first j spikes of the second spike train. It follows that $G_{0,j} = j$ and $G_{i,0} = i$. The distance $G_{i,j}$ is then given by

$$G_{i,j} = \min(G_{i-1,j} + 1, G_{i,j-1} + 1, G_{i-1,j-1} + q|a_i - b_j|) \quad (\text{C.3})$$

where a_i and b_j denote the time of the i th and j th spike of the first and second spike train, respectively. Using the initial conditions for $G_{i,0}$ and $G_{0,j}$, the equation can be iterated up to the point where both spike trains are completely included,

$$D_{VP}(q) = G_{n_1, n_2} . \quad (\text{C.4})$$

The proof that $D_{VP}(q)$ fulfils all the requirements of a distance has been given by Victor and Purpura (1997).

C.1.2 Van Rossum Distance

Distances based on the convolutions of spike trains with linear filters were first proposed by van Rossum (2001). Here it is assumed that each spike train is convolved with a linear filter, e.g., an exponential function,

$$h(t) = \theta(t) \exp^{-t/\tau} \quad (\text{C.5})$$

where the θ -function is zero for all $t < 0$ and one otherwise. The free parameter τ weights the temporal influence of a single spike. Essentially, this function roughly mimics an EPSP of a certain width, cf. Figure C.2. The distance itself is given by the

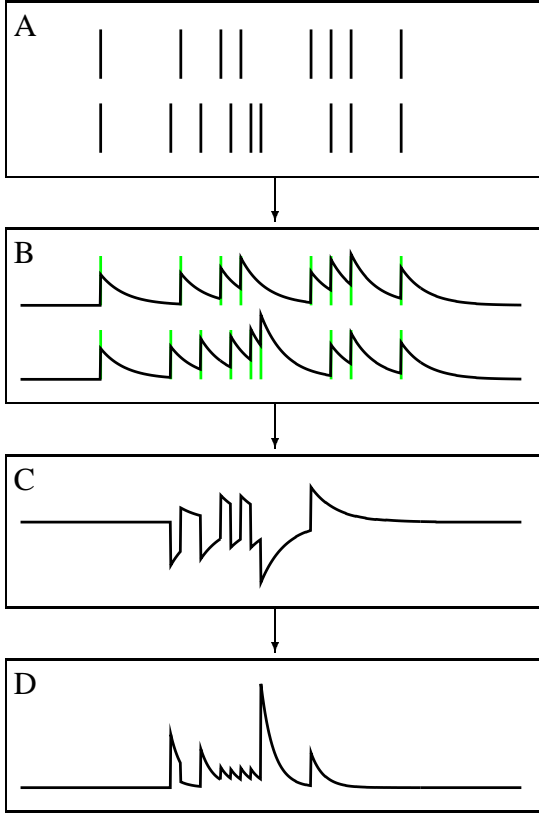


Figure C.2: The van Rossum Distance D_R . Each spike of the original spike trains (A) is replaced by a caricature EPSP (B). To compute the root-mean square difference, the resulting traces are first subtracted (C) and then squared (D). The distance corresponds to the square root of the area under the final trace (D).

root mean-square distance between the convolved traces of the two spike trains,

$$D_R(\tau) = \left[\frac{1}{\tau} \int_0^\infty dt [(r_2(t) - r_1(t)) * h(t)]^2 \right]^{\frac{1}{2}}. \quad (\text{C.6})$$

For all practical purposes, this distance is comparable to the Victor-Purpura Distance explained above when replacing the time constant τ by the reciprocal of the cost factor q . For the limiting cases we find,

$$D_R^2(\tau \rightarrow 0) = \frac{1}{2}(n_1 + n_2 - 2c_{12}) \quad (\text{C.7})$$

$$= \frac{1}{2}D_{VP}(q \rightarrow \infty) \quad (\text{C.8})$$

$$D_R^2(\tau \rightarrow \infty) = \frac{1}{2}(n_1 - n_2)^2 \quad (\text{C.9})$$

$$= \frac{1}{2}D_{VP}^2(q = 0). \quad (\text{C.10})$$

A heuristic argument shows the similarity for the cases of intermediate τ : inserting one additional spike into one of two identical spike trains leads to a distance $D_R^2(\tau) = 1/2$ between the two spike trains (van Rossum, 2001). The same answer is obtained for the removal of one spike. When shifting a spike in one of two identical spike trains by

an amount of Δt , the resulting distance is $D_R^2(\tau) = 1 - \exp(-|\Delta t|/\tau)$ so that $D_R^2(\tau) \approx |\Delta t|/\tau$ for small time shifts. For a given τ or q , the van Rossum Distance is therefore approximately a one-to-one mapping of the Victor-Purpura Distance.

For the data analysis of Chapter 4, we used an alpha function,

$$h(t) = \theta(t) \frac{t}{\tau'} \exp(-t/\tau') \quad (\text{C.11})$$

and defined the temporal resolution parameter to represent the full-width at half maximum of the alpha function. A numerical evaluation yields

$$\tau \approx \frac{\tau'}{2.45}. \quad (\text{C.12})$$

C.1.3 Limitations

Both distance measures are affected by overall time shifts of spike trains. Two identical spike trains that are shifted against each other do have a distance different from zero. To avoid a trivial discrimination based on time shifts, the spike trains were always aligned at the first syllable onset that elicited a spike in at least 90% of the trials.

Likewise, the distance measures do not take into account time dilatation or contraction. Two spike trains that are simply stretched versions of each other will have a finite distance. This fact facilitates the discrimination of calling songs with different syllable-pause length: even if two songs were simply stretched versions of each other, the respective spike trains could be well distinguished by their distances. To avoid a discrimination based on time stretch, two of the stimulus sets of Chapter 4 were designed to have the same syllable-pause duration.

C.2 Clustering Algorithm

Using one of the distance measures introduced above, spike trains elicited from the same stimulus should ideally have much smaller distances than spike trains elicited by different stimuli, resulting in a clustering of spike trains in the respective space. To compare this clustering with the clustering we would expect from the stimulus, we apply a supervised clustering algorithm.

For each song S_i , $i = 1 \dots 8$, one spike train is arbitrarily chosen to be the template for that song. This template provides the centre for the response cluster R_i of song S_i . A spike train is then always assigned to the response cluster of the closest template spike train. After assigning all the recorded spike trains, one can estimate the probability distribution $p(R_i|S_j)$ which denotes the probability that a spike train from the j -th song was assigned to the response cluster of the i -th song.

To minimise statistical fluctuations due to the choice of a particular set of templates, the entire procedure was repeated by permuting all possible template choices. The probability distributions $p(R_i|S_j)$ used through Chapter 4 are averages over the distributions calculated for each set of templates.

Appendix D

Stimulus Reconstruction

Stimulus reconstruction methods seek to reconstruct a stimulus from the spike train that it produced. In turn, such a reconstruction allows one to estimate the mutual information contained in the spike trains of sensory neurons about an external stimulus (Bialek et al., 1991; Rieke et al., 1997). In this appendix, we explain some technical details of stimulus reconstruction and introduce the measures used throughout Chapter 5.¹

The use of stimulus reconstruction methods does not imply that we assume that auditory receptor neurons try to map acoustic stimuli in a one-to-one fashion on their spike trains. To the contrary, by comparing the reconstruction quality for different stimulus ensembles, we seek to find out which characteristics of acoustic signals are encoded faithfully and which features are discarded.

D.1 Stimulus Preprocessing and Calibration

Stimulus reconstruction methods meet the forward approaches halfway. In particular, the analysis is considerably facilitated if one uses *a priori* knowledge about what qualitative aspects of a stimulus a neuron encodes. Auditory receptor neurons of grasshoppers are sensitive to amplitude modulations of broad-band sound-pressure waves that exceed a certain response threshold. Below this threshold, the cells remain silent, cf. Section 3.4.1. Therefore, the appropriate preprocessed stimulus $s(t)$ for applying reconstruction techniques (Fig. D.1) is *not* the original sound-pressure wave $w(t)$ but that part of the AM signal that lies in the sound intensity range covered by the particular receptor. Within the stimulus reconstruction algorithm, therefore, the AM signal was first half-wave rectified at the threshold of each cell (Fig. 5.2A) and then used for the stimulus reconstruction algorithm. The thresholded AM signal will from now on simply be referred to as “signal”.

¹This appendix is adapted from Machens et al. (2001c).

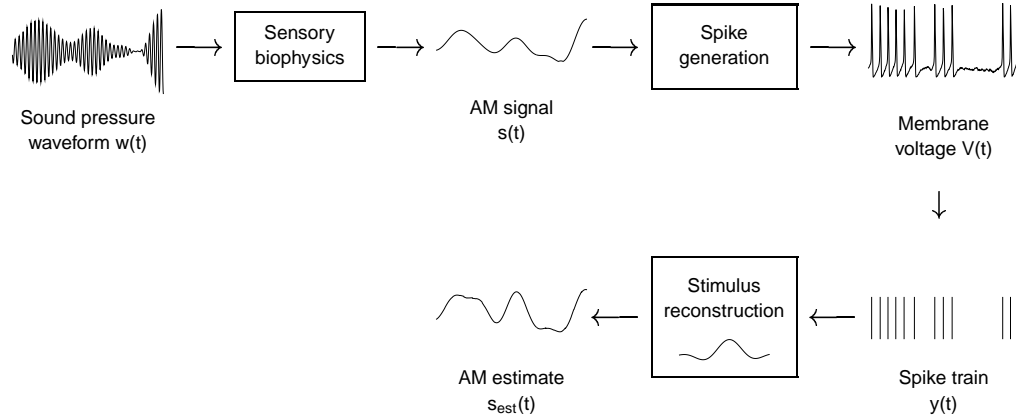


Figure D.1: Stimulus preprocessing and reconstruction. The mechanics of the receiver’s ear extract the slow amplitude modulation $s(t)$ of a rapidly oscillating sound-pressure wave $w(t)$. Auditory receptor neurons then encode $s(t)$ into the membrane voltage $V(t)$. As a first step of the stimulus reconstruction, the spike train $y(t)$ is extracted from the voltage trace. Within linear reconstruction, each spike is then replaced by an optimal filter function to yield $s_{\text{est}}(t)$, the estimate of $s(t)$. As shown by this example, stimulus reconstruction does not aim at recovering the original, complete physical stimulus $w(t)$ but instead requires the identification of a representation of the stimulus that is relevant for the animal, in the present case the AM signal $s(t)$.

D.2 Linear Reconstruction and Filter Calculation

To reconstruct the signal $s(t)$ each spike in the recorded spike train $r(t)$, a series of Dirac impulse functions, is replaced by a linear reconstruction filter $h_1(\tau)$, resulting in the signal estimate $s_{\text{est}}(t)$, cf. equation (2.14),

$$s_{\text{est}}(t) = h_0 + \int_{-T}^0 d\tau h_1(\tau) r(t - \tau) \quad (\text{D.1})$$

where h_0 is the mean signal level in the absence of spiking. The parameters h_0 and $h_1(\tau)$ are determined by minimising the mean-square error $\langle n_{\text{mse}}(t)^2 \rangle$ where the angular brackets $\langle \dots \rangle$ denote a time average over the section of the experiment used for parameter estimation and $n_{\text{mse}}(t)$ is the time-dependent reconstruction error $n_{\text{mse}}(t) = s(t) - s_{\text{est}}(t)$.

To analyse the activity of a population of N receptor neurons, different reconstruction filters $h_{1,i}(\tau)$, $i = 1, \dots, N$ are allowed for each spike train $r_i(t)$, i.e.,

$$s_{\text{est}}(t) = h_0 + \frac{1}{N} \sum_{i=1}^N \int_{-T}^0 d\tau h_{1,i}(\tau) r_i(t - \tau). \quad (\text{D.2})$$

As before, h_0 and $h_{1,i}(t)$ are obtained by minimising the mean-square reconstruction error.

To restrict the number of parameters to be estimated, each reconstruction filter was expanded into an orthonormal series of Hermite functions (Arfken, 1985) of order up to 20. The mean-square-error minimisation then results in a linear system of equations for the expansion parameters which can be solved numerically (Rieke et al., 1997, appendix A.8.2; Press et al., 1992).

By design, reconstruction filters perform above average on the data section used to estimate the filter parameters. To avoid this bias, filters were always estimated on 9 out of 10 segments of a recording and then evaluated on the remaining segment. Sampling errors were reduced by taking averages over repeated permutations of this procedure.

A nonlinear relationship between the AM signal and the firing rate could require higher-order reconstruction filters for adequate signal reconstruction. Such filters seem to suggest relational codes, i.e., coding schemes that involve higher orders of the spike-train statistics, as in interspike-interval-based codes (Theunissen and Miller, 1995). Since the firing-rate responses of auditory receptors of grasshoppers are approximately threshold-linear if amplitude modulations are measured on a logarithmic scale (Römer, 1976; Stumpner and Ronacher, 1991; Ronacher and Krahe, 2000), this potentially misleading interpretation of higher-order kernels was obviated by transforming $s(t)$ and $s_{\text{est}}(t)$ into the decibel scale.

In reconstructions from multiple spike trains, nonlinearities arise if the intensity ranges of the different neurons do not fully overlap. This problem was avoided in reconstructions from multiple traces by only including neurons that had approximately the same threshold.

Firing-rate adaptation can in principle be described by higher-order or time-dependent reconstruction filters but their estimation requires enormous amounts of data. In the studies involving artificial stimuli, these complications were circumvented by discarding the first second of the ten-second neural response patterns.

D.3 Signal-to-Noise Ratio

The reconstruction error $n_{\text{mse}}(t)$ can be separated into random and systematic components. Systematic errors occur if one attempts to reconstruct a signal $s(t)$ that is incompatible with the signal the neuron actually encodes. For instance, if only a low-pass-filtered version of the signal is encoded, any attempts to reconstruct higher frequencies have to fail. Systematic errors can be corrected for by introducing a frequency-dependent gain $g(f)$ such that $s_{\text{est}}(f) = g(f) [s(f) + n_{\text{eff}}(f)]$ where $n_{\text{eff}}(f)$ denotes the random errors or “effective noise”, as referred to the input (Theunissen et al., 1996; Rieke et al., 1997).

Given the effective noise $n_{\text{eff}}(f)$, the success of a stimulus reconstruction for each frequency can be measured by the *signal-to-noise ratio* (SNR),

$$\text{SNR}(f) = \frac{S(f)}{N_{\text{eff}}(f)} \quad (\text{D.3})$$

where $S(f)$ and $N_{\text{eff}}(f)$ are the power spectral densities of the signal and the effective

noise, respectively. In the linear filter case, the gain $g(f)$ is related to the signal-to-noise ratio by $g(f) = \text{SNR}(f)/(1 + \text{SNR}(f))$. Using this relation, one can also calculate the signal-to-noise ratio based on the power spectral densities of estimate and reconstruction error, $\text{SNR}(f) = S_{\text{est}}(f)/N_{\text{mse}}(f)$.

A high SNR indicates an accurate reconstruction, while a SNR of zero implies chance level. The SNR allows one to assess which frequency components are best decoded by signal reconstruction. Reconstruction of signals with high bandwidth serves to estimate the cut-off frequency of the system; this cut-off will be unveiled as the frequency where the signal-to-noise ratio approaches zero.

A measure for the overall success of a reconstruction can be defined by using the total power of signal and noise,

$$\text{SNR} = \frac{\int_0^\infty df S(f)}{\int_0^\infty df N_{\text{eff}}(f)}. \quad (\text{D.4})$$

D.4 Information Transfer

The mutual information, already introduced in section 2.5, quantifies how many bits of information about the signal $s(t)$ are carried by a spike train. Measuring the mutual information per time unit, we obtain the mutual information rate R_{info} . If $s(t)$ is a Gaussian random signal, a lower bound on R_{info} (Rieke et al., 1997) is given by

$$R_{\text{info}} \geq \int_0^\infty df \log_2[1 + \text{SNR}(f)]. \quad (\text{D.5})$$

The LMD and SMD signals of Chapter 5 were generated from Gaussian distributions by nonlinear but invertible transformations. Such transformations conserve the information carried by the signal. After performing the corresponding inverse transformation on both the AM signal and its reconstruction, the lower bound on R_{info} is, therefore, still given by equation (D.5). Accordingly, the signal-to-noise ratio in equation (D.5) was calculated using the original Gaussian AM signal and the inversely transformed reconstruction. Furthermore, the artificial signals were reconstructed without applying a threshold to the preprocessed signal, since the reconstruction errors thus obtained followed a Gaussian distribution more closely, making it more likely that equation (D.5) is a tight lower bound. As the receptor neurons do not encode details of the sub-threshold signal (see also Fig. 5.3D), the information rates thus computed must almost exclusively be conveyed by the supra-threshold signal.

Note that the mutual information rate can be large even if the signal is only poorly reconstructed as might occur for stimuli with high bandwidth.

D.5 Coding Efficiency

Given a time resolution Δt , the efficiency of a neuron to transmit information can be measured by comparing the estimated mutual information rate $R_{\text{info}}(\Delta t)$ with the

information-theoretic limit $R_{\max}(\Delta t)$ which is reached if the spike train is maximally disordered, i.e., Poisson (see Rieke et al. 1995 for details). The *coding efficiency* $\varepsilon(\Delta t)$ is then defined as

$$\varepsilon(\Delta t) = \frac{R_{\text{info}}(\Delta t)}{R_{\max}(\Delta t)}, \quad (\text{D.6})$$

and takes values between zero and one. While $R_{\max}(\Delta t)$ tends to infinity for $\Delta t \rightarrow 0$, this is not the case for $R_{\text{info}}(\Delta t)$ which will instead achieve the value R_{info} given in equation (D.5). To yield non-trivial results, the coding efficiency, therefore, has to be evaluated at a finite time resolution which reflects spike-timing variability due to intrinsic noise sources. This time resolution was estimated by cross-correlation analysis. The full width at half maximum of the cross-correlation peak was calculated for each two spike trains recorded from the same stimulus. The smallest width obtained, $\Delta t \approx 1$ msec, was then taken to be the approximate time resolution of the system. Values of Δt ranging from 0.5 to 2 msec yielded comparable results and underline the robustness of the method.

D.6 Redundancy

Let $R_{\text{info}}(1)$, $R_{\text{info}}(2)$ and $R_{\text{info}}(1, 2)$ denote the mutual information rates obtained from reconstructions based on two individual spike trains and from the corresponding population reconstruction, respectively, with all quantities calculated at a time resolution of $\Delta t = 0.1$ msec. A measure of *redundancy* ρ of these two cells was defined as

$$\rho = \frac{R_{\text{info}}(1) + R_{\text{info}}(2) - R_{\text{info}}(1, 2)}{\min\{R_{\text{info}}(1), R_{\text{info}}(2)\}} \quad (\text{D.7})$$

where $\min\{x, y\}$ denotes the smaller of the two variables x and y . A value of $\rho = 1$ implies complete redundancy and $\rho = 0$ corresponds to complete independence. Negative values of ρ occur if the two cells are synergistic.

Since identical spike trains carry the same information, their redundancy is one. The reverse, however, is not true: a redundancy of one does not imply that the spike trains were identical, as even two different spike trains might carry identical information. Therefore, redundancy is not simply a measure of spike-train variability, but a measure of the information-theoretic consequences of that variability.

Bibliography

- Adam, L.-J. (1977). The oscillating summed action potential of an insect's auditory nerve (*Locusta migratoria*, Acrididae). II. Underlying spike pattern and causes of spike synchronization. *Biol. Cybern.*, 28:109–119.
- Anderson, M. (1994). *Sexual selection*. Princeton University Press.
- Arfken, G. (1985). *Mathematical methods for physicists*. Academic Press.
- Arimoto, S. (1972). An algorithm for computing the capacity of arbitrary discrete memoryless channels. *IEEE Trans. Inform. Theory*, IT-18(1):14–20.
- Attias, H. and Schreiner, C. E. (1997). Temporal low-order statistics of natural sounds. In M. C. Mozer et al. , editor, *Advances in Neural Information Processing Systems 9*, pages 27–33. MIT Press.
- Attias, H. and Schreiner, C. E. (1998). Coding of naturalistic stimuli by auditory mid-brain neurons. In M. I. Jordan et al. , editor, *Advances in Neural Information Processing Systems 10*, pages 103–109. MIT Press.
- Balasubramanian, V., Kimber, D., and Berry, M. J. I. (2001). Metabolically efficient information processing. *Neural Comp.*, 13(4):799–815.
- Barbieri, R., Quirk, M. C., Frank, L. M., Wilson, M. A., and Brown, E. N. (2001). Construction and analysis of non-Poisson stimulus-response models of neural spiking activity. *J. Neurosci. Methods*, 105:25–37.
- Bellman, R. (1961). *Adaptive control processes: A guided tour*. Princeton University Press.
- Bellmann, H. (1985). *Heuschrecken*. Verlag J. Neumann-Neudamm.
- Benda, J. (2001). *Single neuron dynamics—models linking theory and experiment*. PhD thesis, Humboldt-University Berlin.
- Bialek, W., Rieke, F., de Ruyter van Steveninck, R. R., and Warland, D. (1991). Reading a neural code. *Science*, 252:1854–1857.
- Bishop, C. M. (1995). *Neural networks for pattern recognition*. Oxford University Press.

- Blahut, R. E. (1972). Computation of channel capacity and rate-distortion functions. *IEEE Trans. Inform. Theory*, IT-18(4):460–473.
- Borges, J. L. (2000). *The library of Babel*. David R. Godine Publisher.
- Borst, A. (2001). Neural computation of visual motion information in the fly. *Proc. 28th Goettingen Neurobiology Conference*, pages 117–132.
- Brenner, N., Bialek, W., and de Ruyter van Steveninck, R. R. (2000). Adaptive rescaling maximizes information transmission. *Neuron*, 26:695–702.
- Brockwell, P. J. and Davis, R. A. (1991). *Time series: theory and methods*. Springer-Verlag.
- Cover, T. M. and Thomas, J. A. (1991). *Elements of information theory*. Wiley.
- Darwin, C. (1871). *The descent of man and selection in relation to sex*. John Murray, London, reprinted by Princeton University Press (1992).
- Dayan, P. and Abbott, L. F. (2001). *Theoretical Neuroscience*. MIT Press.
- Dimitrov, A. G. and Miller, J. P. (2001). Neural coding and decoding: communication channels and quantization. *Network: Comput. Neural Syst.*, 12(4):441–472.
- Eggermont, J. J. (1993). Wiener and Volterra analysis applied to the auditory system. *Hearing Research*, 66:177–201.
- Eguia, M. C., Rabinovich, M. I., and Abarbanel, H. D. (2000). Information transmission and recovery in neural communication channels. *Phys. Rev. E*, 62:7111–7122.
- Elsner, N. (1974). Neuroethology of sound production in gomphocerine grasshoppers (Orthoptera: Acrididae). *J. Comp. Physiol.*, 88:67–1.
- Faber, A. (1953). *Laut- und Gebärdensprache bei Insekten, Orthoptera I. Vergleichende Darstellung von Ausdrucksformen als Zeitgestalten und ihre Funktion*. Stuttgart.
- Fairhall, A. L., Lewen, G. D., Bialek, W., and de Ruyter van Steveninck, R. R. (2001). Efficiency and ambiguity in an adaptive neural code. *Nature*, 412(22):787–792.
- Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *J. Opt. Soc. Am. A*, 4(12):2379–2394.
- Gabbiani, F. and Koch, C. (1998). Principles of spike train analysis. In Koch, C. and Segev, I., editors, *Methods of neuronal modelling: From ions to networks*. MIT Press, Cambridge, MA.
- Galizia, C. G., Sachse, S., Rappert, A., and Menzel, R. (1999). The glomerular code for odor representation is species specific in the honeybee *Apis mellifera*. *Nature Neurosci.*, 2:473–478.

- Gat, I., Tishby, N., and Abeles, M. (1997). Hidden Markov modeling of simultaneously recorded cells in the associative cortex of behaving monkeys. *Network: Comput. in Neural Syst.*, 8(3):297–322.
- Goldstein, H. (1980). *Classical Mechanics*. Addison Wesley.
- Gollisch, T., Benda, J., Schütze, H., and Herz, A. V. M. (2001). Phenomenological model for sound-intensity coding in auditory receptor neurons of locusts. *Proc. 28th Goettingen Neurobiology Conference*, page 360.
- Gray, E. G. (1960). The fine structure of the insect ear. *Phil. Trans. R. Soc. Lond. B*, 243:75–94.
- Harth, E. and Tzanakou, E. (1974). Alopex: a stochastic method for determining visual receptive fields. *Vision Res.*, 14:1475–1482.
- Hauser, M. D. (1996). *The evolution of communication*. MIT Press.
- Haykin, S. (1994). *Communication systems*. Wiley.
- Hedwig, B. (1986). On the role in stridulation of plurisegmental interneurons of the acridid grasshopper *Omocestus viridulus* L. *J. Comp. Physiol. A*, 158:413–427.
- Hodgkin, A. L. and Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol.*, 117:500–544.
- Hoppenstaedt, F. C. (1997). *An Introduction to the mathematics of neurons*. Cambridge University Press.
- Hoy, R. R., Popper, A. N., and Fay, R. R., editors (1999). *Comparative Hearing: Insects*. Springer-Verlag.
- Hubel, D. H. (1988). *Eye, brain, and vision*. Scientific American Library.
- Hubel, D. H. and Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.*, 160:106–154.
- Hunter, I. W. and Korenberg, M. J. (1986). The identification of nonlinear biological systems: Wiener and Hammerstein cascade models. *Biol. Cybern.*, 55:135–144.
- Ives, D. (1995). Words, words, words. In *All in the timing. Fourteen plays*. Vintage Books.
- Jacobs, K., Otte, B., and Lakes-Harlan, R. (1999). Tympanal receptor cells of *Schistocerca gregaria*: Correlation of soma positions and dendrite attachment sites, central projections and physiologies. *J. Exp. Zool.*, 283:270–285.

- Kandel, E. R., Schwartz, J. H., and Jessell, T. M. (2000). *Principles of Neural Science*. McGraw-Hill.
- Keat, J., Reinagel, P., Reid, R. C., and Meister, M. (2001). Predicting every spike: a model for the response of visual neurons. *Neuron*, 30:803–817.
- Korenberg, M. J. and Hunter, I. W. (1986). The identification of nonlinear biological systems: LNL cascade models. *Biol. Cybern.*, 55:125–134.
- Kriegbaum, H. (1989). Female choice in the grasshopper *Chorthippus biguttulus*. *Naturwissenschaften*, 76:81–82.
- Lang, F. (1999). Acoustic communication distances of a gomphocerine grasshopper. *Bioacoustics*, 10:233–258.
- Laughlin, S. (1981). A simple coding procedure enhances a neuron's information capacity. *Z. Naturforsch.*, 36:910–912.
- Lee, Y. L. and Schetzen, M. (1965). Measurement of the Wiener kernels of a non-linear system by cross-correlation. *Internat. J. Control*, 2:237–254.
- Li, S. T. and Hammond, J. J. (1975). Generation of pseudorandom numbers with specified univariate distributions and correlation coefficients. *IEEE Trans. Systems, Man, and Cybernetics*, 4:557–561.
- Machens, C. K. (1998). Statistical analysis of spike trains. Master's thesis, Humboldt-University Berlin.
- Machens, C. K. (2001). Adaptive sampling by information maximization. *physics/0112070*.
- Machens, C. K., Prinz, P., Stemmler, M. B., Ronacher, B., and Herz, A. V. M. (2001a). Discrimination of behaviorally relevant signals by auditory receptor neurons. *Neurocomputing*, 38-40:263–268.
- Machens, C. K., Schütze, H., Stemmler, M. B., Ronacher, B., and Herz, A. V. M. (2001b). Auditory receptor neurons preserve the individuality of acoustic communication signals. *submitted*.
- Machens, C. K., Stemmler, M. B., Prinz, P., Krahe, R., Ronacher, B., and Herz, A. V. M. (2001c). Representation of acoustic communication signals by insect auditory receptor neurons. *J. Neurosci.*, 21:3215–3227.
- MacKay, D. J. C. (1992). Information-based objective functions for active data selection. *Neural Comp.*, 4:590–604.
- Mainen, Z. F. and Sejnowski, T. J. (1995). Reliability of spike timing in neocortical neurons. *Science*, 268:1503–1506.

- Marmarelis, P. Z. and Marmarelis, V. Z. (1978). *Analysis of physiological systems. The white-noise approach*. Plenum Press.
- Mayr, E. (1997). *This is biology*. Harvard University Press.
- Meyer, J. and Elsner, N. (1996). How well are frequency sensitivities of grasshopper ears tuned to species-specific song spectra? *J. Exp. Biol.*, 199:1631–1642.
- Møller, A. P. (1992). Female swallow preference for symmetrical male sexual ornaments. *Nature*, 357:238–240.
- Moran, J. and Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science*, 229:782–784.
- Nawrot, M., Aertsen, A., and Rotter, S. (1999). Single-trial estimation of neuronal firing rates - from single neuron spike trains to population activity. *J. Neurosci. Meth.*, 94:81–92.
- Nelken, I., Prut, Y., Vaadia, E., and Abeles, M. (1994). In search for the best stimulus: an optimization procedure for finding efficient stimuli in the cat auditory cortex. *Hearing Res.*, 72:237–253.
- Olshausen, B. A. and Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381:607–609.
- Palm, G. and Pöpel, B. (1985). Volterra representation and Wiener-like identification of nonlinear systems: scope and limitations. *Quart. Rev. Biophys.*, 18(2):135–164.
- Papoulis, A. (1991). *Probability, random variables, and stochastic processes*. McGraw-Hill.
- Pauls, S., Pauls, M., Elsner, N., and Lakes-Harlan, R. (2001). Composition of the tympanal nerves of different grasshopper species. *Proc. 28th Goettingen Neurobiology Conference*, page 372.
- Pollack, G. S. (1998). Neural processing of acoustic signals. In Hoy, R. R., Popper, A. N., and Fay, R. R., editors, *Comparative hearing: insects*, pages 139–196. Springer-Verlag.
- Poor, H. V. (1994). *An introduction to signal detection and estimation*. Springer-Verlag.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (1992). *Numerical recipes in C*. Cambridge University Press.
- Rieke, F., Bodnar, D. A., and Bialek, W. (1995). Naturalistic stimuli increase the rate and efficiency of information transmission by primary auditory afferents. *Proc. R. Soc. Lond. B*, 262:259–265.

- Rieke, F., Warland, D., de Ruyter van Steveninck, R. R., and Bialek, W. (1997). *Spikes - exploring the neural code*. MIT Press.
- Ripley, B. D. (1996). *Pattern Recognition and Neural Networks*. Cambridge University Press.
- Roeder, K. D. (1967). *Nerve cells and insect behavior*. Harvard University Press.
- Römer, H. (1976). Die Informationsverarbeitung tympanaler Rezeptorelemente von *Locusta migratoria* (Acrididae, Orthoptera). *J. comp. Physiol.*, 109:101–122.
- Römer, H. (1998). The sensory ecology of acoustic communication in insects. In Hoy, R. R., Popper, A. N., and Fay, R. R., editors, *Comparative hearing: insects*, pages 63–96. Springer-Verlag.
- Ronacher, B. (1989). Stridulation of acridid grasshoppers after hemisection of thoracic ganglia: evidence for hemiganglionic oscillators. *J. Comp. Physiol. A*, 164:723–736.
- Ronacher, B. and Krahe, R. (2000). Temporal integration versus parallel processing: coping with the variability of neuronal messages in directional hearing of insects. *Eur. J. Neurosci.*, 12(6):2147–2156.
- Ronacher, B. and Römer, H. (1985). Spike synchronization of tympanic receptor fibres in a grasshopper (*Chorthippus biguttulus* L., Acrididae). *J. Comp. Physiol. A*, 157:631–642.
- Ronacher, B. and Stumpner, A. (1988). Filtering of behaviourally relevant temporal parameters of a grasshopper's song by an auditory interneuron. *J. Comp. Physiol. A*, 163:517–523.
- Roweis, S. T. and Saul, L. K. (2000). Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290:2323–2326.
- Ruderman, D. L. (1994). The statistics of natural images. *Network: Comput. in Neural Systems*, 5:517–548.
- Sahani, M. (1999). *Latent variable models for neural data analysis*. PhD thesis, California Institute of Technology.
- Sakurai, J. J. and Tuan, S. F. (1993). *Modern Quantum Mechanics*. Addison Wesley.
- Schreiber, S., Machens, C. K., Herz, A. V. M., and Laughlin, S. (2002). Energy efficient coding with discrete stochastic events. *Neural Comp.*, in press.
- Shannon, C. E. and Weaver, W. (1949). *The mathematical theory of communication*. University of Illinois Press.

- Simmons, L. W. and Ritchie, M. G. (1996). Symmetry in the songs of crickets. *Proc. R. Soc. Lond. B*, 263:305–311.
- Sippel, M. and Breckow, J. (1983). Non-linear analysis of the transmission of signals in the auditory system of the migratory locust *Locusta migratoria*. *Biol. Cybern.*, 46:197–205.
- Stanley, G. B., Li, F. F., and Dan, Y. (1999). Reconstruction of natural scenes from ensemble response in the lateral geniculate nucleus. *J. Neurosci.*, 19(18):8036–8042.
- Stemmler, M. B. and Koch, C. (1999). How voltage-dependent conductances can adapt to maximize the information encoded by neuronal firing rate. *Nat. Neurosci.*, 2(6):521–527.
- Strong, S. P., Koberle, R., de Ruyter van Steveninck, R. R., and Bialek, W. (1998). Entropy and information in neural spike trains. *Phys. Rev. Lett.*, 80(1):197–200.
- Stumpner, A. and Ronacher, B. (1991). Auditory interneurons in the metathoracic ganglion of the grasshopper *Chorthippus biguttulus*. I. Morphological and physiological characterization. *J. exp. Biol.*, 158:391–410.
- Stumpner, A., Ronacher, B., and von Helversen, O. (1990). Transmission and filtering of behaviourally relevant song parameters by auditory neurons in a grasshopper. In Gribakin, F. G., Wiese, K., and Popov, A. V., editors, *Sensory systems and communication in arthropods*. Birkhäuser.
- Stumpner, A. and von Helversen, D. (2001). Evolution and function of auditory systems in insects. *Naturwissenschaften*, 88:159–170.
- Suga, N. (1989). Principles of auditory information-processing derived from neuroethology. *J. exp. Biol.*, 146:277–286.
- Tenenbaum, J. B., de Silva, V., and Langford, J. C. (2000). A global geometric framework for nonlinear dimensionality reduction. *Science*, 290:2319–2323.
- Theunissen, F. and Miller, J. (1995). Temporal encoding in nervous system: a rigorous definition. *J. Comput. Neurosci.*, 2:149–162.
- Theunissen, F., Roddey, J. C., Stufflebeam, S., Clague, H., and Miller, J. P. (1996). Information theoretic analysis of dynamical encoding by four identified primary sensory interneurons in the cricket cercal system. *J. Neurophys.*, 75(4):1345–1364.
- Theunissen, F. E., Sen, K., and Doupe, A. J. (2000). Spectral-temporal receptive fields of nonlinear auditory neurons obtained by using natural sounds. *J. Neurosci.*, 20(6):2315–2331.

- Tishby, N., Pereira, F. C., and Bialek, W. (1999). The information bottleneck method. *physics/0004057*.
- van Rossum, M. (2001). A novel spike distance. *Neural Comp.*, 13:751–763.
- Victor, J. D. and Purpura, K. P. (1996). Nature and precision of temporal coding in visual cortex: a metric-space analysis. *J. Neurophysiol.*, 76(2):1310–1326.
- Victor, J. D. and Purpura, K. P. (1997). Metric-space analysis of spike trains: theory, algorithms and application. *Network: Comput. in Neural Syst.*, 8:127–164.
- von Helversen, D. (1972). Gesang des Männchens und Lautschema des Weibchens bei der Feldheuschrecke *Chorthippus biguttulus*. *J. Comp. Physiol.*, 81:381–422.
- von Helversen, D. and von Helversen, O. (1987). Innate receiver mechanisms in the acoustic communication of orthopteran insects. In Guthrie, D. M., editor, *Aims and methods in neuroethology*. Manchester University Press.
- von Helversen, D. and von Helversen, O. (1994). Forces driving coevolution of song and song recognition in grasshoppers. *Prog. Zool.*, 39:253–284.
- von Helversen, D. and von Helversen, O. (1998). Acoustic pattern recognition in a grasshopper: Processing in the time or frequency domain? *Biol. Cybern.*, 79:467–476.
- von Helversen, O. and von Helversen, D. (1997). Recognition of sex in the acoustic communication of the grasshopper *Chorthippus biguttulus*. *J. Comp. Physiol. A*, 180:373–386.
- Wang, X.-J. and Buzsáki, G. (1996). Gamma oscillations by synaptic inhibition in a hippocampal interneuronal network model. *J. Neurosci.*, 16(20):6402–6413.
- Warland, D. K., Reinagel, P., and Meister, M. (1997). Decoding visual information from a population of retinal ganglion cells. *J. Neurophysiol.*, 78:2336–2350.
- Watson, P. J. and Thornhill, R. (1994). Fluctuating asymmetry and sexual selection. *TREE*, 9:21–25.
- Wessel, R., Koch, C., and Gabbiani, F. (1996). Coding of time-varying electric field amplitude modulations in a wave-type electric fish. *J. Neurophysiol.*, 75(6):2280–2293.
- Wiener, M. C. and Richmond, B. J. (1999). Using response models to estimate channel capacity for neuronal classification of stationary visual stimuli using temporal coding. *J. Neurophysiol.*, 82:2961–2875.
- Wiener, N. (1958). *Nonlinear Problems in Random Theory*. MIT Press.

Deutsche Zusammenfassung

Unsere Sinnessysteme erfassen und verarbeiten ständig die vielfältigen und komplexen Reize unserer Umwelt. Um die funktionellen Eigenschaften eines solchen Systems zu untersuchen, verwendet man jedoch oft relative einfache, abstrakte Reize. So werden Neuronen im visuellen System z.B. oft durch die Präsentation geometrischer Muster untersucht, Neuronen im auditorischen System durch Sinustöne unterschiedlicher Frequenz und Intensität. Obwohl diese Experimente große Erfolge für das Verständnis sensorischer Verarbeitung gezeitigt haben, sind die verwendeten Stimuli oft weit von der Komplexität und dem Reichtum natürlicher Stimuli entfernt.

In dieser Arbeit untersuchen wir anhand des auditorischen Systems von Heuschrecken, welche zusätzlichen Erkenntnisse sich durch die Verwendung natürlicher Stimuli oder künstlicher Stimuli, deren Eigenschaften denen natürlicher Stimuli ähneln, erzielen lassen. Die Verwendung solcher Reize bietet zwei Vorteile: Zum einen erlauben sie, die jeweiligen neuronalen Befunde im Kontext des Verhaltens zu interpretieren. Zum anderen tasten künstliche Stimuli, die natürliche Eigenschaften beinhalten, die wesentlichen Bereiche des Stimulusraums ab. Dadurch ermöglichen diese Reize eine umfassendere Charakterisierung des jeweiligen sensorischen Systems als dies durch einfache, abstrakte Reize erreicht wird. Im günstigsten Fall werden dadurch neue Eigenschaften des zu untersuchenden Systems ans Tageslicht gefördert.

Um das auditorische System der Heuschrecke in seinen verhaltensrelevanten Rahmen zu stellen, behandeln wir in Kapitel 3 die akustische Kommunikation von Heuschrecken. Gegenüber vielen anderen natürlichen Stimuli haben die Gesänge von Heuschrecken den großen Vorteil, dass die Verhaltensrelevanz der Signalstruktur genau untersucht ist. Die Verwendung akustischer Signale dient bei Heuschrecken der Partnerfindung und -selektion. Hierbei singen die Männchen und warten auf die akustische Antwort eines paarungsbereiten Weibchens. Die Antwort eines Weibchens kann dabei hoch selektiv ausfallen, da Weibchen speziell daran interessiert sein müssen, unter verschiedenen männlichen Individuen diejenigen mit der höchsten genetischen Qualität auszuwählen. Der Gesang wird durch ein rhythmisches Reiben der Hinterbeine an den Vorderflügeln erzeugt. Die Amplituden des so generierten Schalls werden durch die rhythmischen Bewegungen stark moduliert. Diese bis zu 2 ms schnelle Amplitudenmodulation ist das verhaltensrelevante Signal. Auf der Empfängerseite wird das Schallsignal durch Rezeptorneuronen in eine Reihe von elektrischen Pulsen (Aktionspotentialen) umgewandelt.

Um überhaupt zwischen verschiedenen Männchen wählen zu können, müssen

Weibchen in der Lage sein, Unterschiede zwischen den Gesängen verschiedener Männchen zu detektieren. Kapitel 4 untersucht, ob die entsprechende Information tatsächlich im auditorischen System von Weibchen vorliegt. Das Problem ist hierbei, dass die Unterschiede zwischen den Gesängen leicht in der neuronalen Variabilität untergehen könnten. Die auf acht Männchengesängen basierende Untersuchung zeigt jedoch, dass die Weibchen eine sehr feine Auflösung der Struktur dieser Gesänge erreichen. In der Tat genügen oft bereits 200–300 ms des Gesangs, um alle acht Männchen voneinander zu unterscheiden. Dies gilt selbst dann noch, wenn man die Längen der repetitiven Untereinheiten der Gesänge normiert. Selbst die völlige Auslöschung jeglicher Information über die Trägerfrequenzen der Gesänge verringert die Diskriminationsleistung nicht. Die zur Diskrimination verbleibenden Unterschiede der Gesänge bestehen aus Fluktuationen der Amplitudenmodulation auf Zeitskalen von 2–10 ms. Diese Unterschiede geben Auskunft, wie genau das übliche Muster der Hinterbeinbewegung eingehalten wird. Bei vielen Männchen kommt es durch Häutungsfehler, Kontakt mit natürlichen Feinden oder Asymmetrien der neuronalen Mustergeneratoren zu Störungen dieses Stridulationsmusters. Die gute Auflösung der feinen Gesangsunterschiede erklärt sich daher durch das Interesse der Weibchen, die entsprechenden Defizite der Männchen zu entdecken.

In Kapitel 5 wird die Untersuchung der auditorischen Rezeptoren auf weitere Teilbereiche des Stimulusraums ausgedehnt. Damit soll gezeigt werden, wie spezifisch die Rezeptoren auf Eigenschaften der Heuschreckengesänge angepasst sind. Um dies zu untersuchen, testen wir das System mit verschiedenen künstlichen Stimuli, die in unterschiedlichem Grade die Eigenschaften natürlicher Heuschreckengesänge erfassen. Die eine Gruppe von Stimuli besitzt die gleiche bimodale Verteilung von Schallamplituden wie ein natürlicher Heuschreckengesang. In einer anderen Gruppe sind alle Stimuli Gauß-verteilt. Die Untersuchung selbst basiert auf den Stimulusrekonstruktionsmethoden, die erstmals von Bialek et al. (1991) eingeführt wurden. Diese Methoden erlauben, einen gegebenen Stimulus aus der Antwort eines Neurons annähernd zu rekonstruieren. Die Güte dieser Rekonstruktion gestattet wiederum, die Informationsübertragung abzuschätzen. Es zeigt sich dabei, dass die Stimuli, die den vom Neuron am besten codierten Bereich des Stimulusraums optimal abtasten, sowohl die bimodale Verteilung natürlicher Gesänge besitzen als auch auf Zeitskalen von 2.5 ms fluktuieren. Diese Eigenschaften der besten Stimuli entsprechen gerade den verhaltensrelevanten Eigenschaften der natürlichen Gesänge.

Kapitel 4 und 5 zeigen also, dass die wichtigen Eigenschaften natürlicher Stimuli sehr gut durch auditorische Rezeptoren abgebildet werden. Umgekehrt zeigt Kapitel 5 ebenfalls, dass die Stimuli, die gut abgebildet werden, auch die wichtigen sind. Falls eine solche Abstimmung ein allgemeines Prinzip sensorischer Systeme ist, so kann man die verhaltensrelevanten Stimuli oder Stimuluseigenschaften ohne vorheriges Wissen aus dem System selbst extrahieren. Diese Idee führt in Kapitel 6 zu einem Algorithmus, der ein Ensemble von Stimuli sucht, das die Informationsübertragung (und somit die Güte und Zuverlässigkeit der Abbildung) maximiert. Während ältere Algorithmen üblicherweise einen einzigen Stimulus suchen, der die Feuerrate eines

Neurons maximiert, suchen wir hier also ein ganzes Ensemble von Stimuli, welches die Informationsübertragung maximiert. Ausgehend von einem parametrischen Modell eines Stimulusensembles wird eine Auswahl von Stimuli zufällig gezogen und dem System präsentiert. Die gewonnenen Daten werden anschließend benutzt, um eine erste Abschätzung der Stimulus-Antwort-Charakteristik des Systems zu gewinnen. Im Hinblick auf die Maximierung der Informationsübertragung lässt sich nun die "Wichtigkeit" der präsentierten Stimuli bewerten. Die entsprechend gewichteten Stimuli dienen wiederum einer Neuabschätzung der Parameter des Stimulusensembles. Wenn dieses Verfahren iteriert wird, "wandert" das Stimulusensemble in den Bereich des Stimulusraums, der von dem jeweiligen System gut abgebildet wird. Beispielhafte Berechnungen an einem Modellneuron verdeutlichen, dass die Methode in der Lage ist, das optimale Stimulusensemble in der Zeit zu finden, die ein durchschnittliches elektrophysiologisches Experiment dauert.

Insgesamt haben wir also gezeigt, dass die Eigenschaften der auditorischen Rezeptoren sehr gut auf die natürlichen Gesänge abgestimmt sind. Dies war nur möglich, indem wir über Untersuchungen mit einfachen Reizen hinausgegangen sind und die Eigenschaften eines sensorischen Systems im Hinblick auf natürliche Stimuli untersucht haben. Falls sensorische Systeme generell gut auf die jeweilig verhaltensrelevanten Stimuli abgestimmt sind, so erlaubt der vorgestellte iterative Algorithmus, diese Stimuli automatisch und ohne vorherige Kenntnis zu finden.

Veröffentlichungen

Wissenschaftliche Publikationen

- Machens CK (2001). Adaptive sampling by information maximization. physics/0112070.
- Machens CK, Stemmler MB, Prinz P, Krahe R, Ronacher B, Herz AVM (2001a). Representation of acoustic communication signals in insect auditory receptors. *J Neurosci*, 21(9):3215-3227.
- Machens CK, Prinz P, Stemmler MB, Ronacher B, Herz AVM (2001b). Discrimination of behaviorally relevant signals by auditory receptor neurons. *Neurocomputing*, 38-40: 263-268.
- Schreiber S, Machens CK, Herz AVM, Laughlin SB (2001). Energy efficient coding with discrete stochastic events. *Neural Comp.*, in press.

Tagungsbeiträge

- Machens CK, Schütze H, Stemmler MB, Ronacher B, and Herz AVM (2001). Discrimination of individual mating songs as a first step towards sexual selection. *Neural Information and Coding Workshop*, Big Sky, Montana.
- Machens CK (2001). Finding stimulus ensembles that maximize information transfer in single neurons. *Proc of the 28th Göttingen Neurobiology Conference I:242*.
- Machens CK, Stemmler MB, Prinz P, Krahe R, Ronacher B, Herz AVM (2001). Representation of acoustic communication signals by insect auditory receptors. *Proc of the 28th Göttingen Neurobiology Conference I:237*.
- Schreiber S, Machens CK, Herz AVM, Laughlin (2001). Energy efficient coding with discrete stochastic events. *Proc of the 28th Göttingen Neurobiology Conference I:238*.

- Edin F, Machens CK, Schütze H, Benda J, Herz AVM (2001). Iterative stimulus reconstruction in closed-loop experiments. Proc of the 28th Göttingen Neurobiology Conference I:240.
- Schütze H, Machens CK, Stemmler MB, Ronacher B, Herz AVM (2001). Discrimination of behaviorally relevant signals by auditory receptor neurons. Proc of the 28th Göttingen Neurobiology Conference II:365.
- Machens CK, Stemmler MB, Prinz P, Ritz R, Ronacher B, and Herz AVM (2000). Representation of acoustic communication signals in the grasshopper auditory system. Soc Neurosci Abs 26(1):174.
- Machens CK, Prinz P, Stemmler MB, Ronacher B, and Herz AVM (2000). Discrimination of behaviorally relevant signals by auditory receptor neurons. Computational Neuroscience Meeting. Neurocomputing, 38–40:263–268.
- Prinz P, Machens CK, Ronacher B (1999). Encoding of natural songs by auditory receptor cells of grasshoppers. Proc of the 27th Göttingen Neurobiology Conference, Band I, Seite 66.
- Machens CK, Stemmler MB, Prinz P, Krahe R, Ronacher B, Herz AVM (1998). Coding in the locust auditory system as revealed by stimulus reconstruction methods. Soc Neurosci Abs 24(1):2113.
- Machens CK, Stemmler MB, Herz AVM (1998). Decoding spike trains of model neurons: role of nonlinearity and feedback. Eur J Neurosci 10, 10:404.

Lebenslauf

- 19.11.1970 Geboren in Münster/Westfalen als erster Sohn von
Maria-Elisabeth Machens und Prof. Dr. Dietrich Machens
- 1977-1981 Grundschule Walluf/Rheingau
- 1981-1982 Gymnasium Eltville
- 1982-1990 Altes Kurfürstliches Gymnasium Bensheim
- 1987-1988 Smith Cotton High School Sedalia, MO, USA.
(Nacel Austauschschüler)
- 1990 Abitur
- 1990-1991 Zivildienst bei der Arbeiterwohlfahrt
Kreis Bergstraße
- 1991-1992 Studium generale am Leibniz Kolleg in Tübingen
- 1992-1995 Physikstudium, Universität Tübingen
- 1994 Vordiplom
- 1995-1996 Physikstudium, State University of New York at Stony Brook
(Austauschprogramm der Universität Tübingen)
- 1996-1998 Physikstudium, Humboldt-Universität zu Berlin
- 1998 Diplom in Physik
- 8/1998 Sommerschule “Methods in Computational Neuroscience”
Marine Biological Laboratory, Woods Hole, MA
- seit 9/1998 Promotion bei Prof. Dr. Andreas Herz

Erklärung

Hiermit erkläre ich, dass ich die Dissertation selbstständig verfasst und keine anderen als die angegebenen Hilfsmittel benutzt habe.

Berlin, den 5. Oktober 2001