Université d'Ottawa · University of Ottawa

PERMISSION DE REPRODUIRE ET DE DISTRIBUER LA THÈSE

PERMISSION TO REPRODUCE AND DISTRIBUTE THE THESIS

NOM DE L'AUTEUR / NAME OF AUTHOR:	CHACRON, Maurice Jacques
ADRESSE POSTALE / MAILING ADDRESS:	610-200 Terrace Rideau Ottawa, ON K1M 2C8
GRADE / DEGREE:	**ANNÉE D'OBTENTION / YEAR GRANTED**
Ph.D.(Physics)	2003
TITRE DE LA THÈSE / TITLE OF THESIS:	
Neural Dynamics Leading to Optimized Information ~~Processing~~ Transfer	

L'auteur permet, par la présente, la consultation et le prêt de cette thèse en conformité avec les règlements établis par le bibliothécaire en chef de l'Université d'Ottawa. L'auteur autorise aussi l'Université d'Ottawa, ses successeurs et cessionnaires, à reproduire cet exemplaire par photograhie ou photocopie pour fins de prêt ou de vente au prix coûtant aux bibliothèques ou aux chercheurs qui en feront la demande.

Les droits de publication par tout autre moyen et pour vente au public demeureront la propriété de l'auteur de la thèse sous réserve des règlements de l'Université d'Ottawa en matière de publication de thèses.

The author hereby permits the consultation and the lending of this thesis pursuant to the regulations established by the Chief Librarian of the University of Ottawa. The author also authorizes the University of Ottawa, its successors and assignees, to make reproductions of this copy by photographic means or by photocopying and to lend or sell such reproductions at cost to libraries and to scholars requesting them.

The right to publish the thesis by other means and to sell it to the public is reserved to the author, subject to the regulations of the University of Ottawa governing the publication of theses.

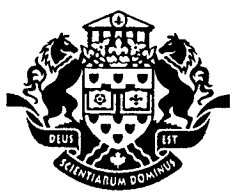
N.B. LE MASCULIN COMPREND ÉGALEMENT LE FÉMININ

25/04/03

DATE

Maurice J. Chacron

(AUTEUR) SIGNATURE (AUTHOR)



Université d'Ottawa · University of Ottawa

Université d'Ottawa · University of Ottawa

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTDOCTORALES

FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

CHACRON, Maurice Jacques

AUTEUR DE LA THÈSE - AUTHOR OF THESIS

Ph. D. (Physics)

GRADE - DEGREE

Physics

FACULTÉ, ÉCOLE, DÉPARTEMENT - FACULTY, SCHOOL, DEPARTMENT

TITRE DE LA THÈSE - TITLE OF THE THESIS

Neural Dynamics Leading to Optimized Information Transfer

A. Longtin

DIRECTEUR DE LA THÈSE - THESIS SUPERVISOR

EXAMINATEURS DE LA THÈSE - THESIS EXAMINERS

F. Gabbiani

P. Johns

I. L'Heureux

R. Munger

J.-M. De Koninck, Ph.D.

LE DOYEN DE LA FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTDOCTORALES

SIGNATURE

DEAN OF THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

Neural Dynamics Leading to Optimized Information Transfer

Maurice J. Chacron

Department of Physics

University of Ottawa, Ottawa

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of PhD in Physics.

SUMMARY

Neural information processing by trains of action potentials is studied in the context of weakly electric fish electroreceptor neurons. A simple but accurate dynamical model for the firing activity of these neurons is presented and compared with experimental results. Dynamical analysis of the model reveals the mechanism by which it reproduces features present in experimental data, such as relative refractoriness and bursting behaviour. Approximations necessary for application of information theory to neural spike trains are presented and different measures are compared. Finally, the consequences of spike patterning caused by relative refractoriness and bursting on information transfer are investigated. It is found that relative refractoriness can increase information transfer while bursting provides a non-linear mechanism for encoding information that might be more efficient than firing of isolated spikes.

SOMMAIRE

Nous étudions le processus par lequel les neurones des poissons électriques traitent l'information sensorielle à l'aide de potentiels d'action. Notre étude propose d'abord un modèle mathématique simple qui reproduit les résultats expérimentaux obtenus sur des neurones sensibles aux champs électriques émis par ces poissons. Une analyse de la dynamique du modèle nous révèle le mécanisme par lequel le modèle reproduit l'importante période réfractaire relative ainsi que la tendance de ces neurones d'émettre des groupes de potentiels d'actions, i.e. des bouffées. Après avoir présenté quelques approximations nécessaires afin de pouvoir quantifier le transfert d'information, nous montrons qu'une période réfractaire relative augmente le transfert d'information, tandis que les bouffées apparaissent comme un méchanisme non-linéaire permettant de transmettre l'information d'une manière plus efficace que les potentiels d'action seuls.

Acknowledgements

I would first like to thank my advisor Dr. André Longtin for his support and advice throughout the course of this work.

Special thanks go to Dr. Leonard Maler for taking the time to teach me about neuroscience and weakly electric fish. His advice was invaluable in order to give behavioural consequences to some results presented here. Discussions with Dr. Joseph Bastian were also insightful for some of this work. I also wish to thank Dr. Khashayar Pakdaman for discussions and help pertaining to non-linear dynamical analysis and Dr. Mark Nelson for sharing his data on electroreceptor neurons with me.

I would also like to thank all my co-workers for insightful comments as well as invaluable discussions: Martin St-Hilaire, Brent Doiron, Anne-Marie Oswald, Jason Middleton, and Drs. Carlo Laing, John Lewis, Jan Benda, and Benjamin Lindner.

Finally, I would like to thank my parents, Maurice and Bernadette Chacron, and my love, Renée-Xavière Larouche, for patience and support, without which none of this would have been possible.

LIST OF ACRONYMS

AM: amplitude modulation

B: bursting

BLIFDT: Bursting leaky integrate-and-fire with dynamic threshold

CV: coefficient of variation

EOD: electric organ discharge

FPT: first passage time

ISI: Interspike interval

ISIH: Interspike interval histogram

LIF: Leaky integrate-and-fire

LIFDT: Leaky integrate-and-fire with dynamic threshold

NB: non-bursting

OU: Ornstein-Uhlenbeck

pdf: probability density function

PND: pulse number distribution

PSD: Power spectral density

RAM: random amplitude modulation

ROC: Receiver operating characteristic

SAM: sinusoidal amplitude modulation

SCC: serial correlation coefficient

SNR: signal-to-noise ratio

SSR: suprathreshold stochastic resonance.

LIST OF SYMBOLS

α	memory parameter
β	constant
ΔI_b	parameter in BLIFDT model
$\Delta\tau$	bin width
γ	constant
Γ	Lyapunov exponent
δ	Dirac delta function
ξ_1	uncorrelated Gaussian random variable
ξ_2	uncorrelated Gaussian random variable
κ	burst feature detection measure
λ_1	Ornstein-Uhlenbeck process
λ_2	Ornstein-Uhlenbeck process
μ	current
ρ_j	serial correlation coefficient at lag j
σ	standard deviation
σ_{SAM}	SAM contrast
τ_1	time constant of λ_1
τ_2	time constant of λ_2
τ_a	time constant
τ_b	time constant

τ_{I_b}	time constant in BLIFDT model
τ_v	voltage decay time constant
τ_s	threshold decay time constant
Θ	Heaviside function
A(t)	amplitude modulation of the EOD
A_0	baseline EOD amplitude
c	constant
C_1	fitting constant
C_2	fitting constant
d	time delay in BLIFDT model
d'	discriminability
D_1	intensity of λ_1
D_2	intensity of λ_2
dt	integration time step
F	Fano factor
f	frequency
f_c	stimulus cutoff frequency
f_{SAM}	SAM frequency
f_{EOD}	EOD frequency
G_a	gain term
G_b	gain term
G_c	gain term
h	entropy

H	entropy rate
H_{gaus}	entropy rate estimated over the Gaussian stimulus ensemble
H_{noise}	entropy rate estimated with external noise
H_{spon}	entropy rate estimated from spontaneous activity
H_{stim}	entropy rate estimated with a long stimulus
i	mutual information
I	mutual information rate
I_b	current in BLIFDT model
I_{gaus}	mutual information rate estimated over the Gaussian stimulus ensemble
I_{noise}	mutual information rate estimated with external noise
I_{spon}	mutual information rate estimated from spontaneous activity
I_{stim}	mutual information rate estimated with a long stimulus
I_c	average information transmitted during the correlation time of the stimulus.
I_i	i^{th} ISI
I_{syn}	synaptic current in LIFDT model
$<I>$	mean of the ISI distribution
L	word length
m	number of subprocesses used in Nelson's model
P	probability distribution
P_D	probability of correct detection

P_{FA}	probability of false alarm
PND	pulse number distribution
r	time dependent firing rate in Nelson's model
r_{base}	baseline firing rate in Nelson's model
s	threshold in LIFDT model
s_0	constant by which the threshold is incremented
s_r	threshold equilibrium value
T	counting time window length
T_r	absolute refractory period duration in LIFDT
v	membrane voltage in the LIFDT model
VS	vector strength
W	stimulus ensemble
w	angular frequency
X	filtered stimulus
X_a	filter variable
X_b	filter variable

LIST OF FIGURES

Table of contents

1 Introduction

As you are reading this text, neurons in your eyes encode the visual scene into trains of brief electric pulses known as action potentials or spikes (abbreviated spk throughout). These trains travel through the optic nerve into the visual cortex at the back of your head and are decoded there, producing vision. One of the aims of neuroscience is the unravelling of the principles by which neurons encode, transmit and decode information (i.e. the neural code). It is a vast field where electrophysiologists, psychologists, biologists, medical practitioners, geneticists, engineers, physicists, and mathematicians interact in order to gain deeper understanding into brain functions. Amongst the possible applications of this understanding are computer vision, artificial intelligence, as well as the development of neural prosthetics.

An action potential is a brief yet large change in the potential across the cellular membrane. As the action potential shape is very stereotyped, it is thought that information is carried in their timing and/or frequency of occurrence (see Gabbiani and Koch 1998 and references therein). One of the aims of Computational Neuroscience is to understand the neural code by constructing quantitative measures of information transfer that can be applied to experimental data. It also uses more or less detailed neuron models in order to explain complex biological phenomena in terms of simple physical concepts. Theoretical predictions can then be verified experimentally, and these in turn are used to validate and refine these models.

The human brain is composed of 10^{12} neurons and has an extremely complex architecture (Felleman and Van Essen 1991). For this reason, many scientists believe that studying simpler organisms can eventually give greater understanding of how the brain functions. In fact, many of the great discoveries in neuroscience did not come from human or even mammalian studies: Hodgkin and Huxley (1952) won the Nobel Prize for building a detailed ionic model of action potential generation in the squid giant axon. Furthermore, most early studies on cellular excitability were done on amphibian nerve muscle (e.g. Lucas 1906). Neuron shapes seem to be conserved across species: many fish neurons and monkey neurons, for example, have similar morphological characteristics. Thus, many investigators have concentrated on studying lower organisms.

Weakly electric fish use amplitude and phase modulations of a self-generated electric field in order to gather information about their environment (Bastian 1981), a process called electrolocation. An electric organ located in their tail generates this electric field, known as the electric organ discharge (EOD). Much is known about weakly electric fish. For example, the neural circuitry responsible for processing amplitude-modulated signals is well known (Maler et al. 1991). Moreover, this is one of the few systems where animal behaviour has been linked to specific neural circuitry (see Heiligenberg 1991 for a review).

The work done here concerns information transfer in the primary sensory neurons for these fish and its relation to animal behavior. Thus, we begin by describing the electric sense that these fish possess as well as the associated sensory neurons.

1.1 Electrolocation

We will concentrate on the species *Apteronotus Leptorhynchus* (brown Ghost knife fish). The process of electrolocation is illustrated in Figure 1-1. Objects with a conductivity different than that of the surrounding water will alter field line trajectories. This results in amplitude and phase modulations of the fish's own sinusoidally time-varying EOD. Electroreceptor neurons on the fish's skin encode these modulations into trains of action potentials that are then decoded in the higher brain. There are two types of electroreceptor neurons: T-types encode phase modulations while P-types encode amplitude modulations (Scheich et al. 1973; Bastian 1981; Zakon 1986).

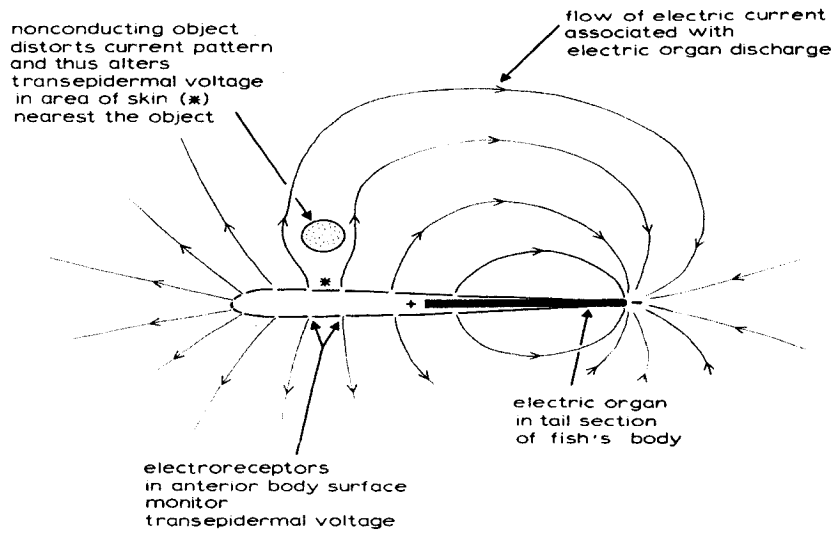


Figure 1-1 : The process of electrolocation. Adapted from Heiligenberg 1981

1.1.1 P-type electroreceptors

Much is known experimentally about electroreceptor neurons. They display spiking activity in the presence of an unmodulated EOD. Furthermore, their spike trains are phase-locked to the EOD since action potential generation tends to occur around peaks of the EOD cycle (Scheich et al. 1973; Bastian 1981; Wessel et al. 1996; Xu et al. 1996). Electroreceptor afferents tend to skip a seemingly random number of EOD cycles between consecutive firings. They are composed of 25-40 receptor cells and a nerve fiber making synaptic contact unto upwards of 16 active neurotransmitter release sites per receptor cell (Bennett et al. 1989) Although intracellular recordings are not yet possible, there is much indirect evidence that the EOD amplitude affects individual receptor potentials, which govern the rate of release of neurotransmitter onto the nerve. The probability P that the nerve fires in an EOD cycle is proportional to the amount of transmitter that binds on target channels in the nerve membrane. The probability of firing increases smoothly with the EOD amplitude (Wessel et al. 1996). Thus, electroreceptors will modulate their firing rate in response to an EOD amplitude modulation.

It is important to note that the actual "meaningful" stimulus for this fish consists of the amplitude modulation only, and that the EOD is merely a carrier wave. The present work aims

at understanding how P-type electroreceptors transmit information about amplitude modulations caused by stimuli and moreover aims at linking specific spike train features to increased information transfer.

1.1.2 Experimental data

Numerous aspects of P-type electroreceptor baseline firing statistics are presented in Xu et al. (1996) as well as Ratnam and Nelson (2000). There is considerable trial-to-trial variability amongst individual units in both their firing rates (100-500 Hz) as well as other firing statistics. Some units are considered "bursty" (i.e. pausing after firing on a few consecutive EOD cycles) while others are considered "non-bursty". These issues will be addressed in detail in chapter 3.

1.2 Behaviour of weakly electric fish

Weakly electric fish display stereotyped responses during prey detection (Nelson and MacIver 1999) as well as social interactions (Zupanc and Maler 1993). They offer a good system in which to study neuroethology, i.e. the biological approach to the study of the neural basis of behaviour. Behavioural constraints are important for modeling as they tell us what is relevant. These fish have a few stereotyped behavioural responses that we will now review.

1.2.1 Prey capture

Weakly electric fish prey on small aquatic invertebrates such as Daphnia (1-2 mm diameter) that emit faint electric signals. Nelson and MacIver (1999) studied prey capture in isolation. The animal was placed in an otherwise empty water tank in complete darkness. Single Daphnia were then introduced in the tank and the prey capture scenario was observed through infrared cameras. These movies give remarkable insight into the movements needed for prey capture and are available online at: http://soma.npa.uiuc.edu/labs/nelson/public_resources.html.

The following is a simplified description of prey capture. The animal takes about 250 ms to pass by the Daphnia. After this, the animal has knowledge of the Daphnia's existence and spatial location as it executes a series of movements that will bring it next to its mouth and eat it. The animal thus has 250 ms in order to make a decision on whether or not a Daphnia is present. We will review theories that help us deal with such problems in chapters 2 and 6. Furthermore, the amplitude modulations caused by a single Daphnia at the time of detection

are very faint (<1 μV). These fish are remarkable in the fact that they can detect electric signals down to 0.1 μV (Knudsen 1974). Prey stimuli are slow and contain temporal frequencies typically under 10 Hz (Nelson and MacIver 1999).

1.2.2 Social communication

Weakly electric fish interact socially during mating season. Males and females engage in courtship behaviour while males become more aggressive amongst themselves. Males emit communication calls or "chirps" that are characterized by a brief increase in EOD frequency and a decrease in EOD amplitude. Chirps typically contain temporal frequencies above 50 Hz. (Zupanc and Maler 1993) and other fish can detect and respond to them. There is a nice segregation between prey and communication signals in these fish since the former typically contain low frequencies while the latter typically contain high frequencies. An understanding of P-units must build on existing knowledge about the signal integration and firing properties of neurons. This is the realm of neural modeling to which we now give a brief introduction.

1.3 Introduction to neuronal modeling

Perhaps the simplest neuron model is the Leaky Integrate-and-fire (LIF) neuron, reviewed in Tuckwell (1988). The model is shown to reproduce the all-or-none characteristic of action potential generation and is given by the following equation:

$$Cv = -\frac{v}{R} + \mu \tag{1.1}$$

Equation (1.1) is simply the current balance equation for an RC circuit with applied current I. The voltage v will relax exponentially to its equilibrium value Rμ. However, the following firing rule is included in the dynamics: when the voltage reaches a threshold value s, it is reset to a value v_0 and we say that an action potential has occurred. Equation (1.1) with this firing rule constitutes a nonlinear system. To see this, consider the response to a current pulse illustrated in Figure 1-2. If the current value μ is less than s/R, then v will always be less than threshold and will settle to μR. We are then in the subthreshold regime where the response is linear (since it is described by Ohm's law) as shown in Figure 1-2a. However, when $\mu > s/R$,

one or more action potentials can be elicited as shown in Figure 1-2b: this is a strongly nonlinear response due to the firing rule based on exceeding a fixed threshold. Although extremely simplified, the LIF model can reproduce the all-or-none nature of action potential generation (Tuckwell 1988). The LIF neuron description is adequate if one is not concerned with the action potential shape or the particular mechanism of action potential generation.

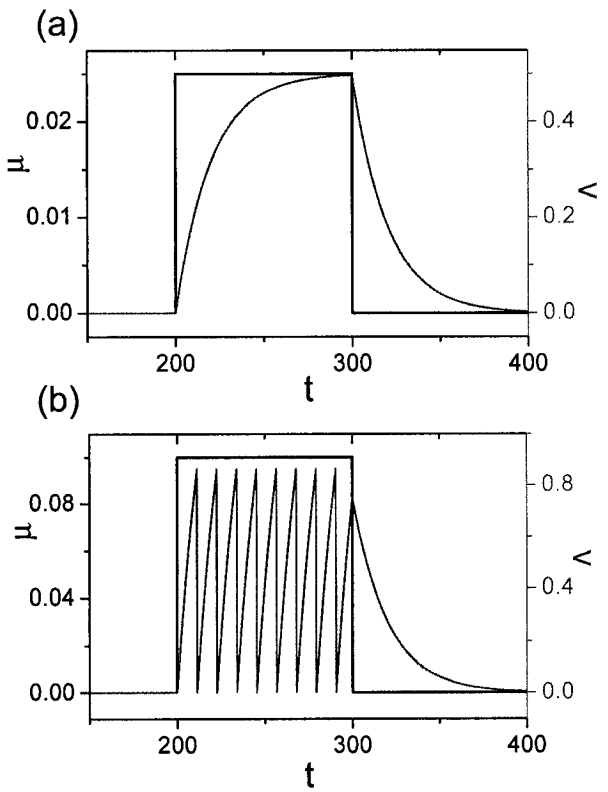


Figure 1-2: (a) Voltage v and current I as a function of time t in the subthreshold regime. (b) Same as (a), but in the suprathreshold regime. Parameter values used were C=1, R=20, s =0.865, v_0=0. Note the different vertical scale for both panels.

There are more detailed models such as the Hodgkin-Huxley (HH) model based on various ionic channels and their kinetics that govern both the subthreshold and suprathreshold behavior (Hodgkin and Huxley 1952). However, the LIF neuron retains enough simplicity to be analytically tractable. In contrast, the HH model has four dynamical variables and is analytically intractable for the most part. HH models are appropriate for situations in which

much is known about the ionic channels such as their kinetics as is the case for e.g. (Doiron et al. 2001). However, little is known about the ionic channel kinetics inside an electroreceptor neuron. Thus, we feel that building the simplest phenomenological model will give the most insight into the internal dynamics of an electroreceptor neuron by making the least amount of assumptions.

1.4 Scope and organization of thesis

Our working hypothesis is that it is possible to answer behavioural questions from analysis of neuron dynamics. This discipline is called neuroethology and has had profound influence on research in weakly electric fish (Heiligenberg 1991). Furthermore, we believe that analysis of simple phenomenological models reproducing key features seen in experimental data will lead to understanding of their functionality in the animal. Our goal is to investigate the consequences of electroreceptor neuron firing dynamics on their information transfer capabilities. Specifically, we address the following questions:

- What constitutes a good model for an electroreceptor neuron? Are electroreceptor neurons optimized for the coding of naturalistic stimuli? If so, how do their dynamics optimize information transfer?

- Why are electroreceptor neurons so noisy? Why is there enormous firing statistics variability across a population of receptors? Are there any computational advantages in having this?

We will provide some answers to these questions by constructing a simple model of electroreceptor neuron spiking activity. In chapter 2, we review some elementary concepts of spike train analysis. The model is presented in chapter 3 and compared with experimental data. Varying the model's parameters allows us to reproduce experimental data from a variety of electroreceptor neurons. A detailed non-linear dynamical analysis of the response to simple input is presented in chapter 4. Chapter 5 reviews and presents some new results on the application of information theory to neural spike trains. Chapter 6 will concentrate on the role of a particular spike train feature on optimizing information transfer for naturalistic stimuli. Finally, in chapter 7, we provide some answers regarding the role of the variability of

experimentally observed electroreceptor neuron spike train statistics by comparing the information transfer properties of two electroreceptor neurons having different firing dynamics. Chapter 8 summarizes the results obtained and provides a general discussion of our results and possible future directions.

1.5 Statement of Contribution

This thesis results from extensive discussions with my advisor Dr. André Longtin as well as with Dr. Leonard Maler, an expert on weakly electric fish. Moreover, Dr. Khashayar Pakdaman and I have worked together in deriving the Lyapunov exponent expressions in chapter 4. All the numerical work presented here was done by the author. Parts of chapter 3 have been published in Chacron et al. (2000, 2001a, 2001b) and Longtin et al. (2003). The contributions from St-Hilaire in Chacron et al. (2000), as well as those of Longtin and Laing in Longtin et al. (2003), are however not presented in this thesis. Parts of chapter 4 have been published in Chacron et al. (2003). Most of chapter 6 has been published in Chacron et al. (2001a) and reviewed in part in Longtin et al. (2003).

2 Methods

In this chapter, we describe different measures used in spike train analysis that will be used throughout this thesis. Since there always is a stochastic component to neural activity, all quantities discussed here are stochastic variables.

2.1 Spike Trains

As mentioned in chapter 1, it is believed that many neurons communicate by means of action potentials. Since all action potentials have the same shape, it is futher believed that information is conveyed by the times at which these action potentials occur. This allows us to represent a sequence of action potentials as a sum of delta functions:

$$x(t) = \sum_{i=1}^{N} \delta(t - t_i) \tag{2.1}$$

where x(t) is the spike train, N is the number of spikes in the spike train, and the t_i are the times at which individual action potentials occur. In practice, the action potential has a finite width, so the firing time is defined by the onset time.

2.1.1 Firing rate

The time varying firing rate of a neuron is defined by:

$$r(t) = < x(t) > \tag{2.2}$$

where the " <>" denotes an average over realisations (i.e. an ensemble of spike trains obtained under similar experimental conditions).

If the spike train is ergodic, then the firing rate does not depend on time and we can replace the average over realizations by an average over time. We thus obtain:

$$r = \lim_{T\to\infty} \frac{1}{T} \int_0^T x(t)\,dt \tag{2.3}$$

r is then the time independent mean firing rate of the neuron. In practice, all spike trains have finite duration. Hence the limit in equation (2.3) is estimated with T finite but large.

2.1.2 The pulse number

Neurophysiologists often are interested in the number of action potentials that occur during a time window of length T from t to t+T. If the spike train is stationary, this quantity does not depend on t and is called the pulse number n(T). It is given by:

$$n(T) = \int_0^T x(t)\,dt \tag{2.4}$$

In practice, n(T) is a random variable that must be described by a probability distribution P(n,T) that is called the pulse number distribution (PND).

2.1.3 Fano factor

The Fano factor F(T) (Fano 1948) is defined as:

$$F(T) = \frac{\sigma^2(T)}{< n(T) >} \tag{2.5}$$

where $\langle n(T) \rangle$, $\sigma^2(T)$ are the mean and variance of P(n,T).

2.2 *Interspike Intervals*

If the spike train is stationary, one can define from the spike times t_i the interspike interval (ISI) sequence I_i by:

$$I_i \equiv t_{i+1} - t_i \tag{2.6}$$

2.2.1 ISI distribution and coefficient of variation

The ISI distribution P(I) is usually obtained experimentally or in numerical experiments by binning an ISI sequence into a histogram, yielding an ISI histogram. The coefficient of variation (CV) of P(I) is defined as the standard deviation-to-mean ratio:

$$CV = \frac{\sigma_I}{\mu_I} \tag{2.7}$$

where μ_I is the mean of P(I) and σ_I is the standard deviation.

2.2.2 ISI correlation coefficients

The ISI serial correlation coefficients (SCC) are a measure of linear correlation between successive ISI's. They are defined by:

$$\rho_L = \frac{\langle I_i I_{i+L} \rangle_i - \mu_I^2}{\sigma_I^2} \tag{2.8}$$

where $\langle\rangle_i$ denotes an average over the ISI sequence. ρ_L is the correlation coefficient at lag L. A negative correlation coefficient implies the following: if an ISI is (on average) shorter/longer than average, then the L^{th} ISI after it will (on average) be longer/shorter than average. A positive correlation coefficient implies that, if an ISI is shorter/longer than average, then the L^{th} ISI after it will be shorter/longer than average.

Note that $\rho_0=1$ by definition. If all other ISI correlation coefficients are zero, then we have a renewal process. Otherwise the process is non-renewal. Also, the asymptotic value of the Fano factor is related to the ISI correlation coefficients as well as the coefficient of variation (Cox and Lewis 1966):

$$\lim_{T\to\infty} F(T) = CV^2\left(1+\sum_{i=1}^{\infty}\rho_i\right) \tag{2.9}$$

2.2.3 The Power Spectral Density

The power spectral density (PSD) of the ISI sequence is the discrete Fourier transform of the SCC's (Cox and Lewis 1966) and is defined for positive frequencies f by:

$$PSD(f) = \frac{1}{\pi}\left(1+2\sum_{j=1}^{\infty}\rho_j\cos(2\pi jf)\right) \tag{2.10}$$

Note that the frequency f is dimensionless and can be thought to be measured in units of the mean firing rate. This formula can be inverted using the inverse Fourier transform to yield an expression for ρ_j (Cox and Lewis 1966):

$$\rho_j = 2\pi\int_0^{1/2} PSD(f)\cos(2\pi jf)\,df \tag{2.11}$$

Note that the PSD is always positive. Moreover, specifying the SCC sequence allows us to uniquely determine the PSD and vice-versa. The two quantities are thus completely equivalent (Cox and Lewis 1966).

2.3 Signal Detection Theory

Signal detection theory was originally used for radar applications. It has many applications in psychophysics (Green and Swets 1966) as well as engineering (Poor 1994) and can be also used in the analysis of neural spike trains (Gabbiani and Koch 1998). We will only be interested in one of its simplest applications here. Let us say that stimulus 1 gives rise to a pulse number $n_1(T)$ and that stimulus 2 gives rise to a pulse number $n_2(T)$. The goal is to distinguish between stimulus 1 and stimulus 2 based on the random variables $n_1(T)$ and $n_2(T)$. Clearly, if these variables have zero variance (i.e. they are deterministic), then this is trivial. However, the task becomes much more difficult if $n_1(T)$, $n_2(T)$ have significant variance. Hence, the ability of an observer to distinguish between the two stimuli will depend on the amount of overlap between the respective probability distributions $P_1(n,T)$, $P_2(n,T)$ of $n_1(T)$, $n_2(T)$, respectively. These concepts are illustrated in Figure 2-1.

This allows us to define the probabilities of correct detection P_D (i.e. of identifying stimulus 2 when stimulus 2 was given) and false alarm P_{FA} (i.e. of identifying stimulus 2 when stimulus 1 was given) P_{FA}. We have (Gabbiani and Koch 1998):

$$\begin{aligned} P_D &= \sum_{n \geq k}^{\infty} P_2(n,T) \\ P_{FA} &= \sum_{n \geq k}^{\infty} P_1(n,T) \end{aligned} \tag{2.12}$$

where k is a positive integer called a threshold. The receiver operating characteristic (ROC) curve is obtained by varying k between 0 and infinity and plotting P_D as a function of P_{FA}.

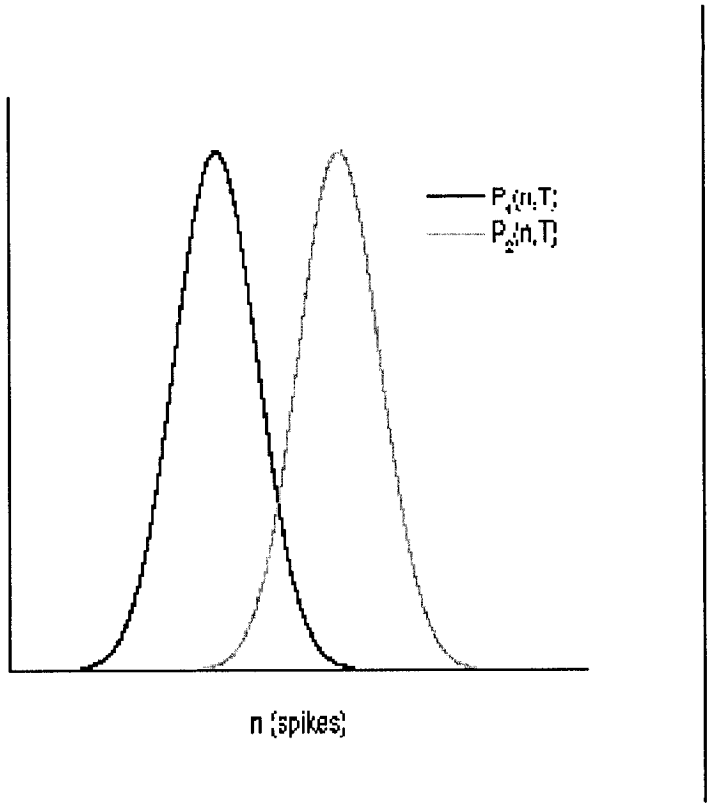


Figure 2-1: The probability distributions $P_1(n,T)$ and $P_2(n,T)$

For a perfect detector, we have $P_D=1$ independently of P_{FA}, whereas for a chance detector we have $P_D=P_{FA}$. All detectors lie somewhere in between these two extremes and this is illustrated in Figure 2-2. Note that there is 100% overlap between $P_1(n,T)$ and $P_2(n,T)$ for chance detection and that there is no overlap for perfect detection. One can thus define a quantity measuring the amount of overlap, the signal-to-noise ratio (SNR):

$$\mathrm{SNR} = \frac{|\mu_2 - \mu_1|}{\sqrt{\sigma_1^2 + \sigma_2^2}} \tag{2.13}$$

The SNR is sometimes denoted by d' (Green and Swets 1966). If there is no overlap, then the SNR is infinite whereas it is zero if the two distributions have equal means or if one of them has infinite variance. It is usually agreed that an SNR of 3-5 corresponds to effectively perfect detection.

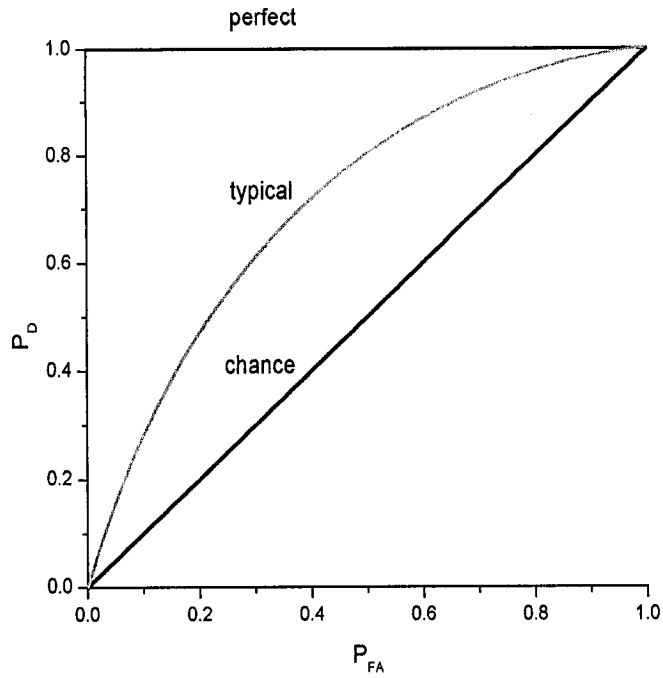


Figure 2-2: The ROC curve. The "chance detection" is obtained when P_D equals P_{FA}. In contrast, perfect detection is obtained when $P_D = 1$ independently of P_{FA}. The ROC curve for a typical detector lies between these extremes.

2.4 *Information Theory*

Information theory has been developed by Shannon (1948) in the context of communication channels. It has been successfully applied in many engineering fields (e.g. fibreoptics) and has been applied to neural spike trains (Holden 1976; Rieke et al. 1997). As we will discuss the applicability of this theory to spike trains in detail in chapter 5, we will only give a brief introduction here. Throughout, we assume that the entropies and mutual information exist for the random variables we consider.

2.4.1 The entropy of a discrete variable

Let X be a discrete random variable with probability distribution P(X). The entropy of X is defined as (Cover and Thomas 1991):

$$h(X) = -\sum_{x} P(x) \log_2 P(x) \tag{2.14}$$

and is expressed in bits. Note that this is the same definition used in statistical mechanics under the micro-canonical ensemble (see e.g. Pathria 1996). The entropy h(X) is always positive since $P(X) \leq 1$. As in statistical mechanics, the entropy of a random variable represents the uncertainty in that variable. Thus, a random variable with a uniform probability distribution has maximal entropy whereas a random variable with a Kronecker-delta distribution (i.e. $P(x)=1$ if $x=x_0$ and $P(x)=0$ otherwise) has zero entropy.

2.4.2 Mutual Information

Let X and Y be random variables with probability distributions P(X), P(Y). The mutual information between X and Y is expressed as (Cover and Thomas 1991):

$$i(X,Y) = h(X) - h(X/Y) \tag{2.15}$$

where h(X/Y) is the entropy of X given Y and is given by:

$$h(X/Y) = -\sum_{x,y} P(x,y) \log_2 P(x/y) \tag{2.16}$$

Here, P(X,Y) is the joint probability distribution of X and Y and P(X/Y) is the probability distribution of X conditioned on Y. Clearly, if X and Y are independent, we have $P(X,Y)=P(X)P(Y)$ and $P(X/Y)=P(X)$ giving $i(X,Y)=0$. The mutual information is thus a measure of correlation between the random variables X and Y. In fact, it measures the reduction in the entropy of X brought by the knowledge of Y. We will compute entropy and mutual information in chapters 5,6 and 7.

2.5 Fitting Experimental Data with a Model

In this section, we discuss the general problem of building a model with the goal of fitting experimental data. Usually, one would like to perform fitting and statistical tests such as a t-test or a χ^2 test in order to ensure that the data is properly fitted. We define the χ^2 between two sets of binned data A_i and B_i of length N as:

$$\chi^2 = \frac{1}{\nu}\sum_{i=1}^{N}\frac{(A_i - B_i)^2}{A_i} \tag{2.17}$$

where ν is the number of degrees of freedom. If the χ^2 is on the order of one, then we can say the data sets are similar and most likely come from the same parent distribution.

We will see in chapter 3 that experimental data from electroreceptor neurons displays considerable variability across the population. Different electroreceptor neurons will in general have very different spiking statistics. However, experimental data from a single electroreceptor neuron is very reproducible in the sense that the statistics are weakly stationary. With this in mind, it is the author's belief that it is not useful to build a model that will exactly reproduce spike train statistics seen experimentally for one electroreceptor neuron. Rather, the underlying philosophy of our modelling is to obtain a minimal generic model that can reproduce basic statistics of single cell firing as well as the variability across the population. As there is a trade-off between model simplicity and closeness to experimental data, it is expected that spike train statistics from the model and the data will have different but similar parent distributions. This philosophy is in line with current standards of modelling neuronal data (Koch, 1999).

A comparable situation to this approach would be the modelling of snowflake shape. It would be useless to produce a model that would perfectly fit a particular snowflake since no two snowflakes are alike. Rather, one would like to know some of the underlying principles by which snowflakes are formed, like the population parent distribution from which they are generated. Our approach to modeling electroreceptor neurons is similar: we build a simple model whose spiking statistics are similar to those seen experimentally. One might then ask the

question: by varying the model parameters, can one obtain firing statistics that are similar to those seen in experimental data from another electroreceptor? The answer is yes if the model captures some of the underlying principles that govern firing statistics of electroreceptor neurons.

3 Models for electroreceptor neurons

In this chapter, we look in greater detail at the electroreceptors neurons used by weakly electric fish to probe their environment. After a review of their characteristics and a previous mathematical model proposed for these electroreceptors, we present a simple biophysically justified model for these neurons. This model is then extended to include bursting dynamics displayed by some electroreceptor neurons. Some of these results have been presented in Chacron et al. 2000; 2001a; 2001b.

3.1 Baseline Activity

3.1.1 Experimental data

It is possible to record action potentials produced by electroreceptor neurons. Bastian (1981) has characterized their spontaneous activity. The probability P of firing in an EOD cycle is a smooth increasing function of the instantaneous EOD amplitude (Wessel et al. 1996; Nelson et al. 1997). P-units can thus encode amplitude modulations (AM's) of this EOD or "carrier", arising from environmental stimuli, into the frequency of "randomly" phase locked firings.

The baseline activity from a typical electroreceptor is presented in Figure 3-1. Panel (a) shows an ISI raster plot that is characterized by bands centered on integer multiples of the EOD period. The ISI return map shown in panel (b) displays clusters with a negative sloped elongation. Panel (c) shows that the serial correlation coefficient (SCC) ρ_1 defined in chapter 2 is negative: this implies that short ISIs are preferentially followed by long ones and vice-versa. Panel (d) shows the ISI histogram: it is multi-modal with Gaussian-shaped modes centered at multiples of the EOD period, as expected from the raster plot on panel (a). Furthermore, the

envelope is approximately Gaussian with mean around five EOD cycles. P-units thus skip a random number of EOD cycles between successive firings.

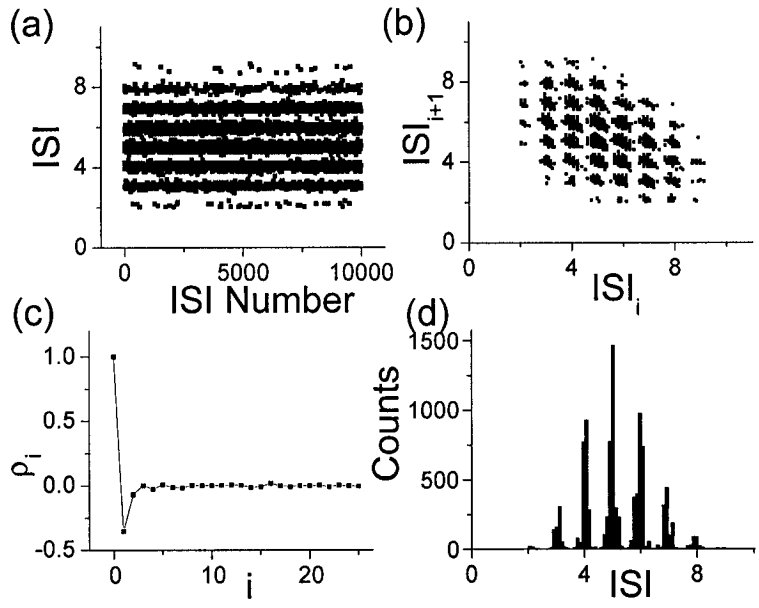


Figure 3-1: Analysis of 10,000 consecutive interspike intervals from a P-unit of the weakly electric fish A. Leptorhynchus (data courtesy of Mark Nelson, Beckmann Institute, Illinois, USA). Time is in EOD cycles; the EOD frequency is 755 Hz. The firing rate is 145 Hz which corresponds to P=0.192. (a) Raster plot of ISI duration versus ISI number. (b) Return map. (c) Serial correlation. (d) ISI Histogram. Note that the error bars on the ISI SCC's are too small to be seen and are thus not included.

3.1.2 Interspike interval correlation analysis

We assume that the i-th firing time is given by $t_i=n_iT+\xi_i$ where $T=1/f_{eod}$ is the EOD period, n_i is a positive integer that increases with i by integer jumps, and the ξ_i are uncorrelated identically distributed random values. We furthermore assume that n_i and ξ_I are uncorrelated. They represent jitter in the firing times around multiples of T. Defining the integer random variable $m_i \equiv n_{i+1}-n_i$, we can write for the SCC at lag one:

$$\rho_1^{est} \equiv \frac{\left\langle (I_{i+1} - \langle I \rangle)(I_i - \langle I \rangle) \right\rangle}{\left\langle (I_i - \langle I \rangle)^2 \right\rangle} = \frac{\left(\langle m_{i+1} m_i \rangle - \langle m_i \rangle^2 \right) T^2 - \langle \xi_i^2 \rangle}{\left(\langle m_i^2 \rangle - \langle m_i \rangle^2 \right) T^2 + 2 \langle \xi_i^2 \rangle} \qquad (3.1)$$

These correlations can be estimated from the data in Figure 3-1 (the EOD period T is normalized to one), since m_i represents the nearest integer to a given ISI: $\langle \xi_i^2 \rangle = 0.00461$, $\langle m_{i+1} m_i \rangle = 24.86$, $\langle m_i^2 \rangle = 26.78$, $\langle m_i \rangle^2 = 25.39$, yielding an SCC of **-0.381**. This is close to the actual value $\rho_1 = -0.351$, which justifies our assumption about the firing times t_i. Another justification comes from inspection of the ISIH; an ISI can be written as: $I_i = m_i T + \eta_i$. The η_i have zero mean and govern the ISIH peak widths, and satisfy $\langle \eta_{i+1} \eta_i \rangle / \langle \eta_i^2 \rangle = -0.5$. This second serial correlation, which the data exhibit (not shown) and which explains the negatively sloped elongation of each cluster in the ISI return map (Figure 3-1b), results from phase locking: when a firing leads (lags) the mean EOD phase at which firings occur, the next firing will lag (lead) on average, so that firings keep in step with the EOD.

3.2 A previous model of electroreceptor activity

We present the model developed by Mark Nelson and collaborators and described in Nelson et al. (1997).

3.2.1 Description

It is known from experiments that various filtering mechanisms are at work inside a P-unit (Hopkins 1976; Bastian 1981; Wessel et al. 1996; Xu et al. 1996; Nelson et al. 1997). Nelson et al. (1997) measured the gain and phase response characteristics of P-units to sinusoidal AM's of frequencies in the range 0.1-200 Hz, leading to the following set of differential equations:

$$
\begin{aligned}
\dot{X}_a &= -\frac{X_a}{\tau_a} + \frac{G_a}{\tau_a} A(t) \\
\dot{X}_b &= -\frac{X_b}{\tau_b} + \frac{G_b}{\tau_b} A(t) \\
X(t) &= -X_a - X_b + (G_a + G_b + G_c)\ A(t)
\end{aligned}
\tag{3.2}
$$

where the dot denotes differentiation with respect to time, A(t) is the stimulus (i.e. the time varying EOD amplitude minus its baseline value), and X(t) is the filtered stimulus. The G's are gains in units of spikes/s/mV, and the τ's, time constants in units of seconds. A baseline firing rate r_{base} is then added to X(t) and the sum Z(t) is then passed through a clipping non-linearity to account for saturation effects (Nelson et al. 1997):

$$
r(t) = \begin{cases} 0 & Z(t) < 0 \\ f_{EOD} & Z(t) > f_{EOD} \\ Z(t) & \text{otherwise.} \end{cases}
\tag{3.3}
$$

This filter will be used again in chapters 5, 6, and 7. The probability p(t) of firing per EOD cycle is thus $r(t)/f_{EOD}$ where f_{EOD} is the EOD frequency. At each maximum of the EOD, the P-unit has probability p(t) of firing. If the unit fires, jitter is added to the spike time in the form of Gaussian white noise of zero mean and standard deviation 0.04 EOD cycles. Throughout the paper, we take f_{EOD}=1000 Hz, hence an EOD cycle corresponds to one millisecond. Furthermore, to reduce the coefficient of variation of the ISIH, it is possible to implement m independent random subprocesses each with an event rate equal to the spike rate r(t) (Nelson et al. 1997). Each subprocess is simulated as described above and output spikes are generated at the time of occurrence of every m^{th} subprocess event. The model was constructed to give the correct responses to sinusoidal AM's (Nelson et al. 1997) and was used to give the firing dynamics in response to changes in transdermal potential caused by a prey (Nelson and MacIver 1999).

3.2.2 Results

Figure 3-2 shows the results obtained with Nelson's model. The ISI raster plots (a) , ISI return maps (b), and ISI histograms (d) are similar to those obtained experimentally (compare with Figure 3-1). However, the model fails to display the prominent negative SCC at lag one displayed by the data (c).

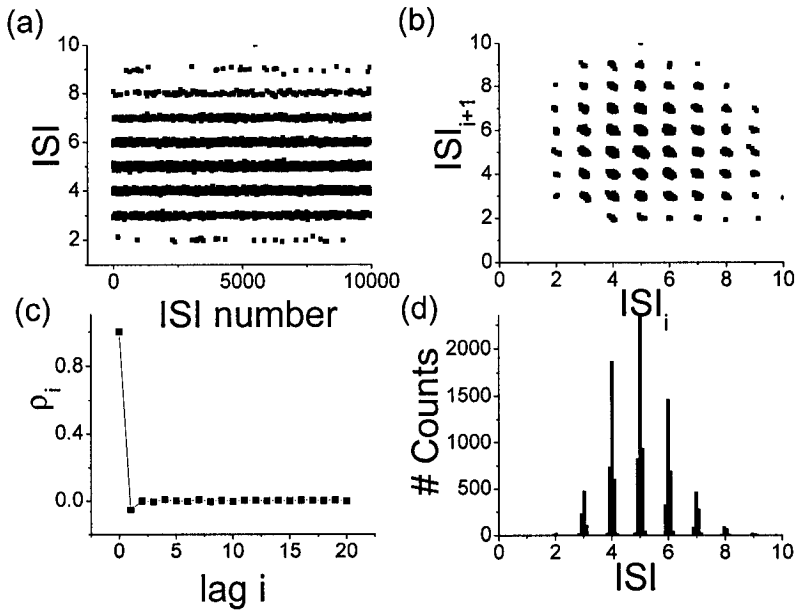


Figure 3-2: Results obtained using Nelson's model to generate 10000 ISIs with A(t)=0. (a) ISI raster plot, (b) ISI return map, (c) SCC sequence, (d) ISI histogram. Time units are in EOD cycles. Parameter values used were: r_{base}=200 Hz; m=18; G_a =14100 spk/s/mV; G_b=470 spk/s/mV; G_c=670 spk/s/mV; τ_a=2.6 ms; τ_b=210 ms.

The small value obtained at lag one can be explained from the interspike interval analysis carried out in the previous section. A similar calculation yields: $\langle \xi_i^2 \rangle$= 0.00309, $\langle m_{i+1} m_i \rangle$= 24.83164, $\langle m_i^2 \rangle$=26.15862, $\langle m_i \rangle^2$=24.913075, yielding an SCC of -0.0628 using equation (3.1). This value is close to the actual value of -0.065. The dominant serial correlation once again comes from the term $\langle m_{i+1} m_i \rangle$.

3.3 The LIFDT model

We first provide a biophysical justification of our model before introducing the equations. As mentioned above, we expect relative refractoriness (i.e. the fact that it is harder to elicit an action potential just after one occurred) to be important at such high firing rates (Stein 1965). It leads to negative SCC's for the ISI's at low lags (Geisler and Goldberg 1966; Chacron et al. 2000). Since there is currently no biophysical characterization of the ionic conductances inside the P-unit, we can only speculate based on our knowledge of other systems about the possible mechanisms responsible for this experimentally observed relative refractoriness. The physiological mechanisms responsible for these correlations could be presynaptic in origin: for example, long term depression at the synapses connecting the receptor cells to the afferent nerve (Bennett et al. 1989, Hausser and Roth 1997) would lead to relative refractoriness. However, the recovery time constant of the neurotransmitter at typical synapses is usually in the thousands of milliseconds range (von Gernsdorff et al. 1997), which is much too long for the phenomenon at work here since serial correlations are significant only up to lag two (see Figure 3-1d). We are thus looking at a time scale of about ten EOD cycles (twice the mean ISI which is 5 EOD cycles long for the unit considered here). A likely candidate would be a postsynaptic spike-activated potassium channel that slowly deactivates and thus summates to produce a negative adaptation current. The KV3.1 channel has the right activation and deactivation kinetics (Wang et al. 1998). Members of the KV3 family are richly expressed in the electrosensory system (Asim J. Rashid, personal communication), but it remains to be shown whether similar channels are present in P-units.

From the foregoing discussion, we feel that the best approach for studying the effects of correlations on information transfer and signal detection is to use a simple yet biophysically plausible model that reproduces the essential features of P-unit baseline and evoked discharge.

3.3.1 Description

We will first give an expression for the synaptic current at the spike initiation zone in the axon, and then describe the spiking mechanism. We use the simple model in Chacron et al. (2000)

that has been proposed to model the baseline firing dynamics of P-type electroreceptors and extend it to get the proper responses to time varying AM's of the EOD amplitude. We write the transdermal potential on the fish's skin as $(A_0 + A(t))\sin(2\pi f_{EOD} t)$, where A_0 is the constant EOD amplitude corresponding to baseline firing dynamics (this is similar to r_{base} in Nelson's model) and A(t) is the AM. The stimulus A(t) is filtered using equation (3.2). As many receptors rectify a periodic input (French et al. 1972; Gabbiani 1996), we take the total synaptic current to be:

$$I_{syn} = (\beta X(t) + \gamma A_0)\Theta(\beta X(t) + \gamma A_0)\sin(2\pi f_{EOD} t) \times \Theta(\sin(2\pi f_{EOD} t))(1 + \lambda_1) + \lambda_2 \tag{3.4}$$

where β,γ are constants to make units match, λ_1,λ_2 are noise terms, and $\Theta(.)$ is the Heaviside function ($\Theta(x)=1$ if $x \geq 0$ and $\Theta(x)=0$ if $x<0$) to account for rectification (i.e. for the fact that synaptic current is positive). Thus, the deterministic component of the synaptic current is zero whenever $\sin(2\pi f_{EOD} t)$ is negative or $\beta X(t)+\gamma A_0$ is negative. This is to ensure that firings will always occur near the maxima of the EOD sine wave, as experimentally observed (Sheich et al. 1973). Because our model is phenomenological, we take the synaptic current to be dimensionless.

Since neurons are stochastic in nature, we include noise sources in the model. These sources could model conductance and synaptic fluctuations and are represented respectively by two Ornstein-Uhlenbeck (Gardiner, 1985) processes λ_1 and λ_2 given by:

$$\lambda_1 = -\frac{\lambda_1}{\tau_1} + \sqrt{D_1}\,\xi_1$$
$$\lambda_2 = -\frac{\lambda_2}{\tau_2} + \sqrt{D_2}\,\xi_2 \tag{3.5}$$

where ξ_1,ξ_2 are two independent Gaussian random variables of zero mean and $\langle\xi(t)\xi(s)\rangle = \delta(t-s)$, D_1,D_2 are constants proportional to the intensities of λ_1,λ_2 and τ_1,τ_2 are time constants. Figure 3-3 shows a time series for λ_1 and λ_2. It can be shown (Gardiner 1985) that

λ_1 and λ_2 are stationary Gaussian random variables with zero mean and respective variances $D_1 \tau_1/2$ and $D_2 \tau_2/2$. However, unlike Gaussian white noise, λ_1 and λ_2 are correlated in time and their correlation functions decay exponentially with respective time constants τ_1 and τ_2. We take τ_1 to be much less than an EOD cycle, hence λ_1 can be thought of as "fast" compared to the model dynamics (it is almost white noise in fact, see Figure 3-3). It could thus model fluctuations that occur on time scales much faster than the EOD cycle (e.g. membrane noise caused by channel flicker low-pass filtered by the membrane capacitance – Manwani and Koch 1999). In this case, D_1 would be related to the strength of the membrane noise.

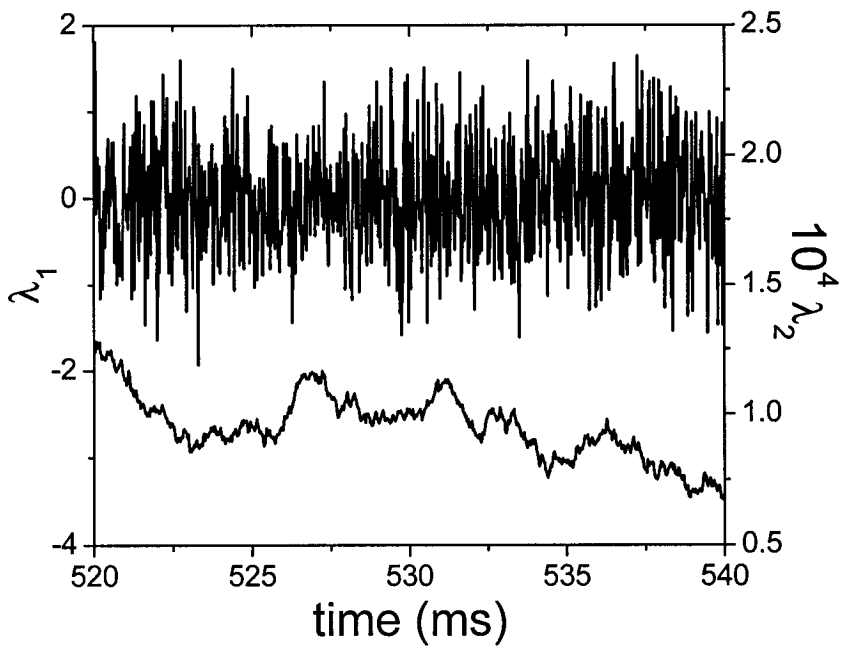


Figure 3-3: Noise terms λ_1 (upper curve) and λ_2 (lower curve) as a function of time. λ_2 varies more slowly than λ_1 and is four orders of magnitude smaller. Parameter values used were: τ_1=0.025 ms; τ_2=50000 ms; D_1=8 ms^{-1}; D_2=0.000009 ms^{-1}.

In contrast, we take τ_2 to be much greater than the EOD cycle. Hence this noise term can be thought of as "slow" compared to the model dynamics (it is almost constant for time scales much smaller than τ_2, see Figure 3-3). It thus models fluctuations on slow time scales (e.g. fluctuations in synaptic vesicular release rate, see Lowen et al. 1997 and discussion below). This term is needed to accurately reproduce the Fano factor curve as shown in the results section.

The spiking mechanism is a simple extension to the leaky integrate-and-fire (LIF) model in which a spike is said to have occurred when the membrane potential V reaches a constant threshold s. Immediately after, the voltage is reset to its resting value (usually taken to be 0). LIF models are memoryless in the sense that consecutive ISI's are not correlated (ρ_i=0 for all i>0). To include refractory effects, we make s also a dynamical variable (see e.g. Geisler and Goldberg 1966). But instead of making it random (Gestri et al. 1980; Gabbiani and Koch 1996; Gabbiani and Koch 1998), we let it carry the memory by the following firing rule (Chacron et al. 2000): when voltage equals threshold, it is reset to zero as in the LIF model while threshold is incremented by a constant amount s_0 and kept constant for the duration of the absolute refractory period T_r after which it relaxes exponentially towards its equilibrium value θ_0 until the next spiking time. The equations for voltage and threshold between times of occurrence of action potentials and after the absolute refractory period are thus:

$$\begin{aligned} v &= -\frac{v}{\tau_v} + \frac{I_{syn}}{\tau_v} \\ s &= (s_r - s)/\tau_s \end{aligned} \qquad (3.6)$$

Like the synaptic current, v and θ are dimensionless. A stretch of simulation showing v and θ is shown in Figure 3-4b. The filter given by equation (3.2) gives the linear transfer properties of the afferent while our spiking mechanism gives the correct baseline dynamics (i.e. when A(t)=0). We thus combine the two to get the correct responses to AM's. Kreiman et al. (2000) had a similar approach to model the P-type electroreceptors of another species of weakly electric fish (*Eigenmania*): they used a high-pass filter to give proper AM response characteristics and fed the output to an LIF model. However, as their LIF model had a random threshold (i.e. uncorrelated from firing to firing), it did not take into account the relative refractory effects that could exist in this species.

3.3.2 Results

Figure 3-4 shows a voltage and threshold trace obtained with the model when A(t)=0 (i.e. baseline activity). Suppose an ISI shorter than $\langle I \rangle$ just occurred, then the threshold (having had less time to decay) will typically be high after the spike and will thus take a long time to decay. Consequently, the next ISI will (on average) be longer than $\langle I \rangle$. We expect that this will give rise to a negative SCC at lag one. This is also the case when an ISI longer than $\langle I \rangle$ occurs: as a result, the threshold will now be lower and the next ISI will be shorter than $\langle I \rangle$.

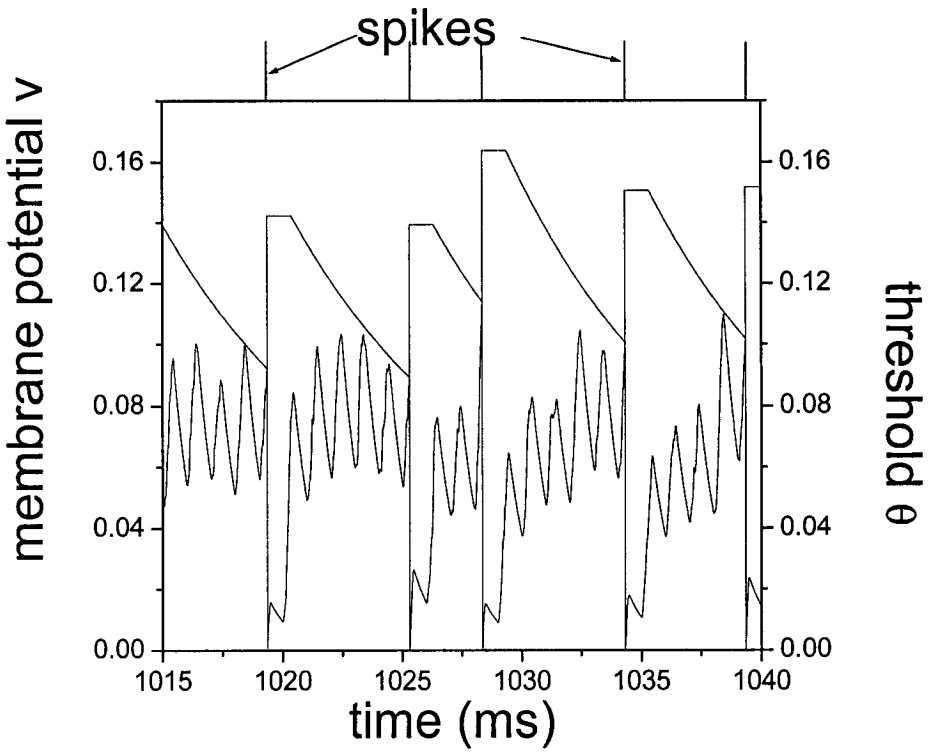


Figure 3-4: Voltage (lower curve) and threshold (upper curve) trace obtained with the LIFDT model for baseline activity (no AM's, thus A(t)=0) showing the firing rule. When voltage equals threshold, a spike is said to have occurred and voltage is reset to zero while threshold is incremented by a constant s_0. The threshold is kept constant to simulate the absolute refractory period T_r (equal to one EOD cycle) and then decays exponentially with time constant τ_s to its equilibrium value s_r. Parameter values are T_r=1 ms; f_{eod}=1000 Hz; γ=0.3266 mV^{-1}; β=1 $(spk/ms)^{-1}$; G_a =14100 spk/s/mV; G_b=470

spk/s/mV; G_c=670 spk/s/mV; τ_a=2.6 ms; τ_b=210 ms; s_r=0.03; s_0=0.05; τ_v=1 ms; τ_s=7.75 ms; A_0=0.8 mV; τ_1=0.025 ms; τ_2=50000 ms; D_1=8 ms^{-1}; D_2=0.000009 ms^{-1}.

This is verified in Figure 3-5 where it is seen that the ISI raster plot (a), the ISI return map (b), the SCC sequence (c), and the ISI histogram (d) are similar to the data (compare with Figure 3-1). The experimentally obtained SCC at lag one was -0.35 while it is -0.391 for the model. We computed χ^2 between the data and model ISI histograms and obtained 50.25. As $\chi^2 \gg 1$, one might conclude that the fit is bad and that the model should be rejected. However, the model qualitatively reproduces the phase-locking, the ISI return maps and ISI SCC's seen experimentally. Moreover, the ISI mean, variance, and SCC's are well within the experimental range observed by Ratnam and Nelson (2000). Chapter 4 will deal with a more detailed analysis of why the model gives rise to a negative SCC at lag one.

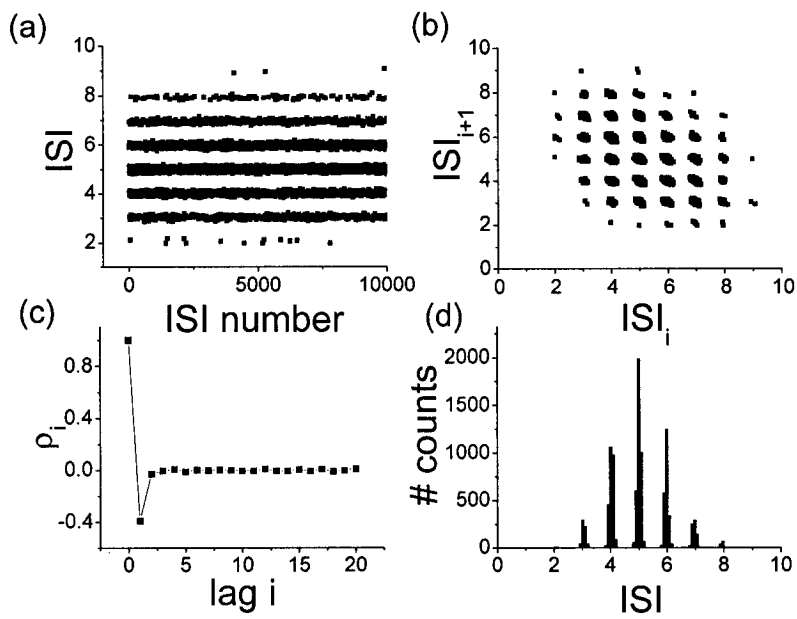


Figure 3-5: Analysis of 10000 ISIs generated by the LIFDT model. (a) ISI raster plot. (b) ISI return map. (c) SCC sequence. (d) ISI histogram. Parameter values are the same as in Figure 3-4 and were chosen to give a good eyeball fit to the experimental data.

3.3.3 Effects of varying parameters on the output spike train

Studying the effects of varying every parameter in the LIFDT model is beyond the scope of this thesis. However, we provide some intuition about the role of the different parameters in the model. As the ISI CV increases with the mean ISI, it is sufficient to look at how different parameters affect the mean ISI and the ISI SCC at lag one.

A_0: The baseline EOD amplitude controls the strength of the carrier wave. Increasing A_0 will lead to a lower mean ISI since the ISI distribution shifts to the left. This will lead to smaller SCC's in magnitude.

D_1, τ_1: The noise intensity D_1 will increase the ISI CV and decrease the mean ISI for high values. It will also decrease the SCC's in magnitude. The effects of varying τ_1 are seen by looking at the effects of noise source λ_2.

D_2, τ_2: Increasing D_2 will have the same effects as increasing D_1. The effects of varying τ_2 are discussed in chapter 6.

τ_v, τ_θ. These decay time constants control the mean ISI. Increasing τ_θ will increase the mean ISI and make SCC's at higher lags non-zero. Increasing τ_v will increase the mean ISI.

s_r: Increasing the steady-state threshold value will increase the mean ISI.

s_0: This parameter controls the appearance of negative SCC's at lags one and higher. The SCC's will increase in magnitude with s_0, as will the mean ISI.

3.4 Bursting electroreceptors

The LIFDT model as presented cannot model the bursting dynamics seen experimentally in some electroreceptors (Bastian 1981; Xu et al. 1996; Nelson et al. 1997). We will thus develop a model for this activity.

3.4.1 Experimental data

Figure 3-6 shows experimental data obtained from a bursting electroreceptor afferent. The ISI raster plot in panel (a) shows a banding pattern just like a non-bursting electroreceptor afferent (compare with Figure 3-1a). Also, the ISI return map in panel (b) shows clusters elongated in the negative direction just like a non-bursting electroreceptor. The SCC sequence in panel (c) shows a prominent negative SCC at lag one. However, the SCC's at higher lags are also non-zero and alternate in sign. Finally, the ISI histogram shown in panel (d) is multi-modal with Gaussian shaped modes centered at integer multiples of the EOD period. However, the envelope is not Gaussian in shape but instead displays two peaks, one around 1 EOD cycle and one around 5 EOD cycles. This electroreceptor tends to fire packets of action potentials called "bursts" separated by periods of quiescence. In the following, we say that an ISI is part of a burst if its value is less than 1.5 EOD cycles.

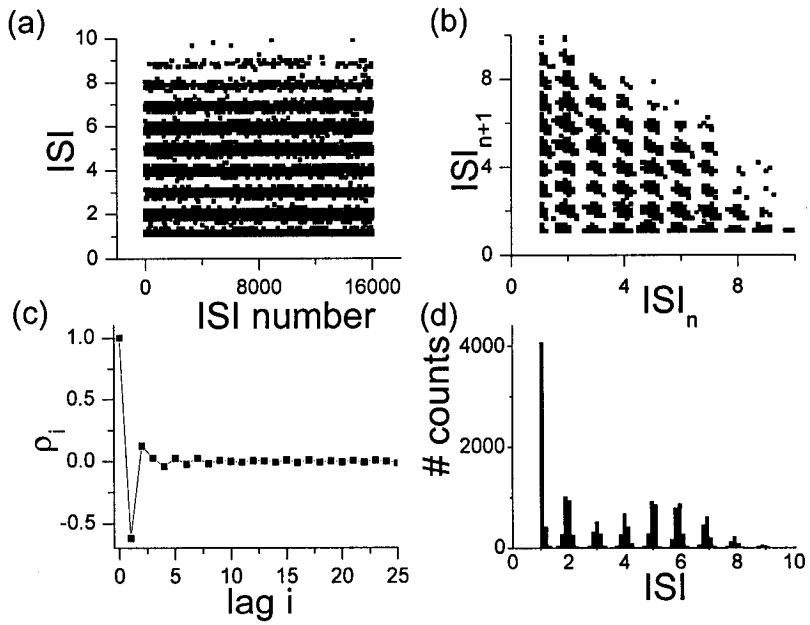


Figure 3-6: Analysis of 16000 ISIs from a bursting electroreceptor afferent (data courtesy of Mark Nelson, Beckman Institute, Ill, USA). Time is in EOD cycles, the EOD frequency is 850 Hz. We have $\langle I \rangle = 3.557$ and $\langle I^2 \rangle - \langle I \rangle^2 = 4.880$. (a) ISI raster plot. (b) ISI return map. (c) SCC sequence. (d) ISI histogram.

3.4.2 An extension of LIFDT model: the BLIFDT model

We extend the LIFDT model by including a current I_b that works as follows: I_b activates (i.e. it is incremented by a constant ΔI_b) after a delay d after an action potential has occurred. This activation is assumed to be instantaneous and such currents are biophysically plausible (Treves, 1996). The current then inactivates exponentially with time constant τ_b. The current I_b is thus governed by the equation:

$$I_b = -\frac{I_b}{\tau_{I_b}} + \Delta I_b \delta(t - t_{fire} - d) \tag{3.7}$$

The term "I_{syn}" is thus replaced by "$I_{syn}+I_b$" in equation (3.6). We call this extended model the Bursting LIFDT (BLIFDT). Note that the LIFDT model is just a special case of the BLIFDT model with $\Delta I_b=0$. Figure 3-7 shows a voltage and threshold trace obtained with the model.

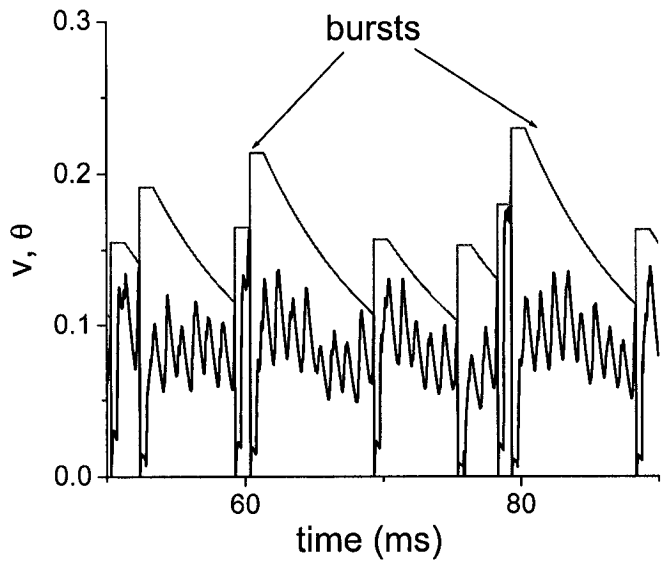


Figure 3-7: Voltage (black curve) and threshold (grey curve) trace obtained with the BLIFDT model. Parameter values are the same as in Figure 3-5 except: τ_s=3.35 ms; s_0=0.1; D_1=39.0625 ms^{-1}. Furthermore, we have: τ_{Ib}=0.25 ms; d=1 ms; ΔI_b=1.4.

One can see that the model produces firings on consecutive EOD cycles (bursts). These tend to be initiated when the threshold s is low. After the first action potential, the threshold is raised and I_b can bring the voltage v to threshold on the next EOD cycle. However, threshold summation will eventually terminate the burst. The bursting seen is thus a competition between the I_b current and the dynamic threshold s. Results obtained with the model are shown in Figure 3-8. The ISI raster plot shown in panel (a) is similar to the data. Furthermore, the ISI return map shown in panel (b) is also similar to the data. The SCC's at lags one and two are quantitatively similar to the data (ρ_1=-0.620 for the data vs. -0.595 for the model and ρ_2=0.125 for the data vs 0.135 for the model) as seen by comparing panels (c) of Figure 3-6 and Figure 3-8. The ISI histogram is similar to the one from the data (Figure 3-6d): they are both multi-modal and have bimodal envelopes, although we got χ^2=101.54 which might indicate a poor fit.

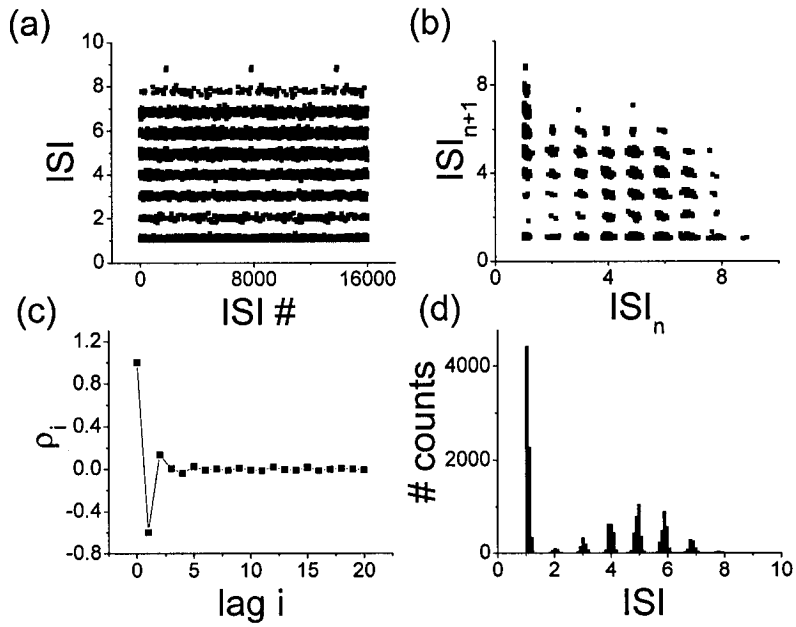


Figure 3-8: Analysis of 16000 ISIs generated by the BLIFDT model. Time is in EOD cycles. We have $\langle I \rangle$=3.248 and $\langle I^2 \rangle - \langle I \rangle^2$=4.505. (a) Raster plot. (b) ISI return map. (c) SCC sequence. (d) ISI histogram. Parameter values are the same as in Figure 3-7 and were adjusted to give a good fit to experimental data.

3.5 *Modeling the trial-to-trial variability obtained in experiments across different electroreceptor neurons*

As mentioned in chapter 2, it is known from experiments that there is significant variability across the electroreceptor afferent population. We try to partially answer the question as to whether the models presented can display similar spiking statistics to other electroreceptor neurons. Xu et al. (1996) have shown that the probability of firing per EOD cycle varied between 0.1 and 0.6. Furthermore, about 30% of electroreceptor afferents are bursty whereas the rest are non-bursty (Nelson et al. 1997). It is thus of interest to see whether the BLIFDT can be used to model the activity recorded from other electroreceptor neurons recorded from.

3.5.1 Comparisons

We performed fits to two different electroreceptor neurons with the BLIFDT model. The results from one such fit are presented in Figure 3-9. Panels (a) and (b) show the ISI histograms obtained with the data and model respectively. The mean ISI $<I>$ was 4.861 and 4.691 for the data and model respectively, while the ISI variance was 4.393 and 3.4229. Panels (c) and (d) show the SCC sequences obtained for the data and model respectively. The SCC ρ_1 was -0.4096 and -0.414 for the data and model, respectively.

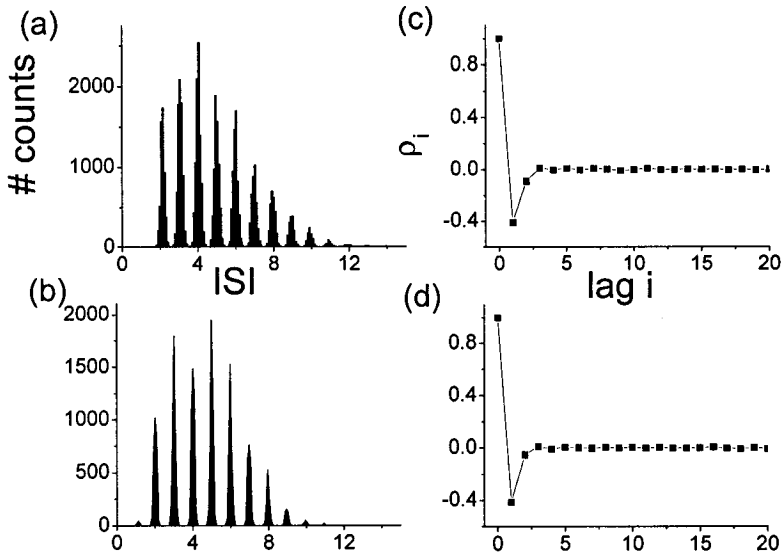


Figure 3-9: Comparison between 21309 ISIs obtained experimentally (data courtesy of L. Maler, University of Ottawa, Canada) and with the BLIFDT model. Time is in EOD cycles and the EOD frequency was 885 Hz. (a) ISI histogram obtained experimentally. (b) ISI histogram obtained with the model. (c) SCC sequence obtained experimentally. (d) SCC sequence obtained with the model. Parameters have the same value as in Figure 3-7 except: γ=0.45 mV^{-1}; s_0=0.05; τ_θ=11 ms; ΔI_b=0.08; τ_{Ib}=0.54 ms.

Figure 3-10 shows another fit. Panels (a) and (b) show the ISI histograms obtained with the data and model, respectively. The mean ISI $<I>$ was 4.394 and 4.181 for the data and model respectively, while the ISI variance was 5.067 and 5.9803. Panels (c) and (d) show the SCC sequences obtained for the data and model, respectively. The SCC sequence was (1, -0.80038, 0.51321, -0.34704, 0.24003, -0.16638, 0.11522, -0.07584, 0.04579, -0.02344, 0.01058) for the data and (1, -0.79873, 0.53405, -0.37551, 0.25781, -0.17395, 0.1128, -0.0732, 0.04873, -0.03748, 0.03279). This particular unit had anti-correlations that decayed up to lag 10 and the BLIFDT model can easily reproduce them.

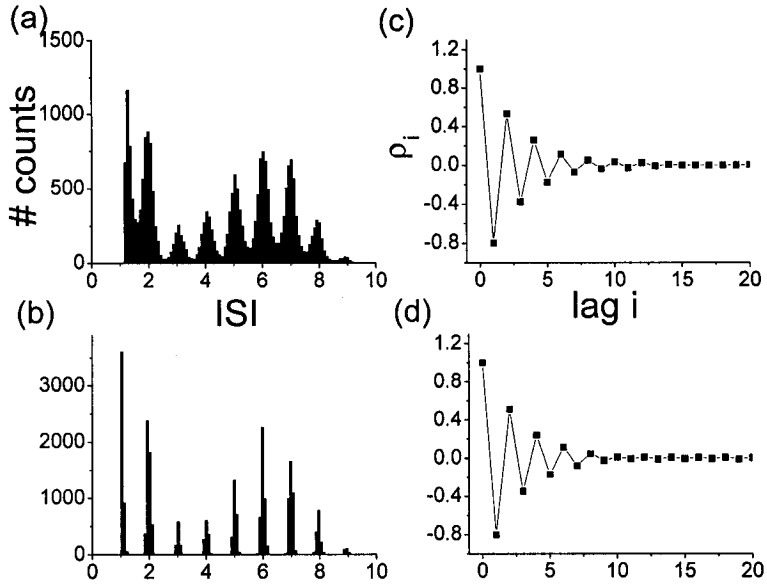


Figure 3-10: Comparison between model and data for 23924 ISIs (data courtesy of L. Maler, University of Ottawa, Canada). Time is in EOD cycles and the fish's EOD frequency was 905 Hz. (a) ISI histogram obtained from the data. (b) ISI histogram obtained from the model. (c) SCC sequence obtained from the data. (d) SCC sequence obtained from the model. Parameter values are the same as in Figure 3-7 except: τ_s=7.75 ms; D_1=4.913 ms^{-1}; ΔI_b=0.26; d=0.4 ms; τ_{Ib}=0.58 ms.

3.6 *Discussion*

3.6.1 Chapter summary

In this chapter, we have introduced electroreceptor neurons that weakly electric fish use to sense amplitude modulations of their self-generated electric field. We have described a previous model of electroreceptor activity and showed that it failed to reproduce the ISI correlations during baseline activity. We have presented a biophysically justified simple model that could reproduce these correlations. We have further extended the model to include the bursting dynamics seen in some electroreceptor neurons. Furthermore, we have shown that

different parameter sets could quantitatively reproduce experimental data obtained from different electroreceptor neurons.

3.6.2 ISI histograms

We have seen that the BLIFDT model could reproduce ISI histograms obtained from both bursting and non-bursting electroreceptors. However, the width of individual modes in the ISI histogram was always higher for the data than for the model. We now address this discrepancy. It should be noted that the mode width can be increased by increasing the parameter D_1 (data not shown). However, this is at the expense of the SCC's that are then lower in magnitude than the data. It was impossible to find a parameter set for which the ISI mode width was as large for the data as for model and for which the SCC's were approximately the same. This is most likely because the model is too simple and thus cannot reproduce this feature of the data.

3.6.3 SCC sequences

It was seen that the model could reproduce the SCC sequences observed experimentally for both bursting and non-bursting electroreceptor neurons. Non-bursting electroreceptors tend to display a single negative SCC at lag one with all other coefficients zero at higher lags while bursting electroreceptors tend to display anti-correlations that alternate in sign and decay over a few lags (L. Maler, personal communication). The source for this difference is not known at present. However, our model shows that the bursting current I_b together with the dynamic threshold θ can account for this type of firing activity.

3.6.4 Modeling electroreceptor neuron variability

It is known experimentally that there is significant variability across the electroreceptor neuron population (Xu et al. 1996; Nelson et al. 1997; Ratnam and Nelson 2000). The four cases presented in this chapter can attest to this fact. The functional consequences of this variability

on information transfer are at present unknown experimentally. Chapter 7 will however provide some answers by comparing the information transfer properties of both a bursting neuron and a non-bursting one.

4 Non-linear dynamical analysis of the LIFDT model

In this chapter, we will be concerned with the analysis of the "core" LIFDT model. The goal is to provide a more complete explanation of results presented in previous chapters. In particular, we will be interested in the properties of the model that will give rise to ISI correlations as well as phase-locking to a sinusoidal input. We present a simplified version of this model that is analytically tractable. Furthermore, we will be interested in the non-linear dynamics displayed by the model under sinusoidal forcing. Most of these results have been published in Chacron et al. (2003).

4.1 A simplified LIFDT model

We will study the model given by the following equations:

$$\begin{aligned} v &= -\frac{v}{\tau_v} + \mu(t) \\ s &= \frac{s_r - s}{\tau_s} \\ v(t^+) &= v_0 \qquad\qquad \text{if } v(t) = s(t) \\ s(t^+) &= s_0 + W[s(t), \alpha] \quad \text{if } v(t) = s(t) \end{aligned} \tag{4.1}$$

where v is the voltage, s is the threshold, $\mu(t)$ is the stimulation current, τ_v and τ_s are the time constants for voltage and threshold respectively, and s_r is the value at which the threshold stabilizes in the absence of firing. Firing occurs when the voltage reaches the threshold. Following this, the voltage is reset to v_0 and the threshold is set to $s(t^+)=s_0+W[s(t),\alpha]$ where s_0 is a parameter and W is a monotically increasing function of s and α with $W(s,0)=0$. α is a positive parameter controlling the memory in the model. We will mostly look at the cases

$W(s,\alpha)=W_1(s,\alpha)=\alpha s$ as well as $W(s,\alpha)=W_2(s,\alpha)=\exp(\alpha s)-1$ (non-linear behavior). Throughout this chapter, we assume that $v_0 \leq 0 < s_r \leq s_0$. The model dynamics are graphically illustrated in Figure 4-1.

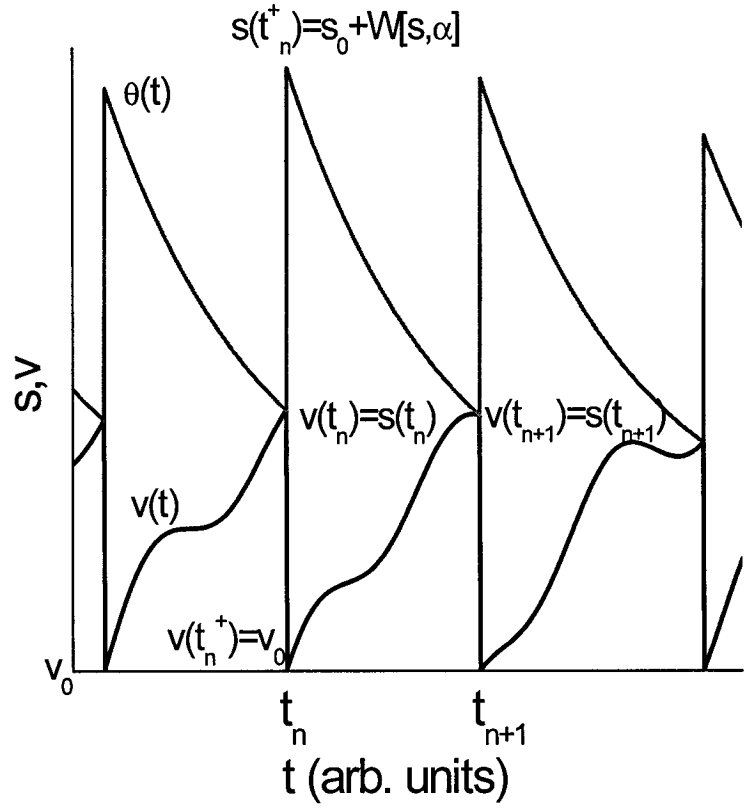


Figure 4-1: Voltage (black solid line) and threshold (grey solid line) time series obtained with the model. An action potential occurs when voltage and threshold are equal. The firing times t_n thus satisfy $v(t_n)=s(t_n)$. Immediately after an action potential, the voltage is reset to v_0 while the threshold is set to a value $s_0+W[s(t),\alpha]$.

The function W_1 represents threshold fatigue with α being the memory parameter. Indeed, when $\alpha=0$, the threshold value immediately after a discharge is independent of the threshold value at that discharge, so that future discharges bear no memory of the firing history. The case

where $\alpha=0$ and $s_0=s_r$ corresponds to the standard LIF with constant threshold. Conversely, when $W=W_1$ and $\alpha=1$ one recovers the threshold fatigue implemented in previous studies, i.e. after each discharge the threshold is raised by a fixed amount s_0 from its value at firing (Segundo et al. 1968; Vibert et al. 1994; Pakdaman and Vibert 1995; Pakdaman et al. 1996; Chacron et al. 2000; 2001a; 2001b; Liu and Wang 2001). Note that this is a generalization of the firing mechanism presented in chapter 3. Other models used a conductance that was decremented by a fixed amount immediately after an action potential to achieve similar effects (Wehmeier et al. 1989; Treves 1996; Liu and Wang 2001). Liu and Wang (2001) have shown that the latter approach gave qualitatively similar results to our model with $W=W_1$ and $\alpha=1$. Previous studies were thus limited to the case $\alpha=1$.

Generally, increasing α from zero corresponds to increasing the degree of fatigue in the model and consequently the dependence of the dynamics on the past firing history. The following sections examine how this parameter affects the correlation between consecutive ISIs and the response of the model to various stimuli. The stimuli of interest are: constant current stimulation, noise, and sinusoidal input. As the realistic situation in electroreceptor afferents is a combination of all these currents, it is important to understand their individual effects.

4.2 Constant current stimulation

In this section, we assume $W=W_1$ and consider constant current $\mu(t)=\mu$. When $\mu\tau_v<s_r$, the stimulus is subthreshold, the membrane potential v(t) stabilizes at its steady-state value $\mu\tau_v$ and no firing occurs. Conversely, when $\mu\tau_v>s_r$, the model generates sustained firing. The remainder of this section analyzes the corresponding discharge pattern.

4.2.1 A 1-D map of ISIs

We construct a map relating consecutive post-discharge threshold values. Let v(t) be the voltage solution with $v(0)=v_0$. From now on, we will consider the threshold as a function of

two variables: the time since the last firing and the initial threshold value right after the last firing S. These are given by:

$$\begin{aligned} v(t) &= (v_0 - \mu\tau_v)\exp\left(-\frac{t}{\tau_v}\right) + \mu\tau_v \\ s(t,S) &= (S - s_r)\exp\left(-\frac{t}{\tau_s}\right) + s_r \end{aligned} \tag{4.2}$$

Let us assume that a discharge occurred at a time t_n and resulted in the postdischarge threshold S_n^+. The next firing takes place after a time interval $I_n \equiv t_{n+1} - t_n$ such that:

$$s(I_n, S_n^+) = v(I_n) \tag{4.3}$$

where $s(t, S_n^+) > v(t)$ for all $0 \leq t < I_n$. This firing yields the postdischarge threshold S_{n+1}^+ at the firing time t_{n+1}^+:

$$S_{n+1}^+ = s_0 + \alpha\left[(S_n^+ - s_r)\exp\left(-\frac{I_n}{\tau_s}\right) + s_r\right] \tag{4.4}$$

From equation (4.2), one can derive a relation between I_n and S_n^+, substituting this in equation(4.4) gives us a relation between the nth and n+1th ISIs I_n and I_{n+1}:

$$I_n = -\tau_v \ln\left\{\frac{\left[(v_0 - \mu\tau_v)e^{-I_{n+1}/\tau_v} + \mu\tau_v - s_r\right]e^{I_{n+1}/\tau_s} + s_r - s_0 - \alpha\mu\tau_v}{\alpha(v_0 - \mu\tau_v)}\right\} \tag{4.5}$$

4.2.2 Stability of the map

While equation (4.5) can be used to discuss the relation between successive ISIs, we prefer to focus on the relation between successive values of the postdischarge threshold. Indeed for constant stimulation, from the ISI I_n one can unambiguously derive the nth postdischarge

threshold S_n^+, and conversely, given S_n^+ there is a unique ISI I_n. Thus, using either the ISIs or the postdischarge thresholds yields the same information about the dynamics of the model. However, for the sinusoidal stimuli, this is not necessarily so: different values of S_n may yield the same I_n. This in turn makes it more difficult to obtain the relation between I_n and I_{n+1}. For this reason, the following paragraphs describe the relation between S_n^+ and S_{n+1}^+.

Equation (4.3) uniquely defines $I_n=I_n(S_n^+)$ as a function of S_n^+, so that we can rewrite equation (4.4) as:

$$S_{n+1}^+ = s_0 + \alpha s[I_n(S_n^+), S_n^+] \equiv F(S_n^+) \tag{4.6}$$

This relation implies that the postdischarge threshold S_n^+ after the nth firing entirely determines the next postdischarge firing S_{n+1}^+. In other words, the dynamics of the LIF with

threshold fatigue receiving a constant stimulation is determined by the iterates of the map F defined by equation (4.6). This map is defined on the interval $[s_0+\alpha s_r, +\infty)$. In general, I(S) and consequently F(S) cannot be explicitly written in closed form because equation (4.3) cannot be solved analytically for arbitrary τ_s. However, the geometrical properties of F that determine the behaviour of the sequence S_n^+ can be determined in the absence of such knowledge.

For fixed t, the function s(t,S) is monotonic increasing, so that if $S>S'$, then s(t,S) cannot intersect the voltage prior to s(t,S'). In other words, $S>S'$ implies that $I(S)>I(S')$, which means that the larger the postdischarge threshold is, the longer the following interspike interval. Given that $v_0\leq 0<\mu\tau_v$ (the postdischarge potential is reset below the resting voltage) and that v(t) is monotonic increasing, we have that $I(S)>I(S')$ implies $v[\Delta(S)]>v[\Delta(S')]$. This relation together with equation (4.3) and $\alpha>0$ yields that $F(S)>F(S')$ whenever $S>S'$, that is, F is monotonic increasing. This, combined with the observation that $F(s_0+\alpha s_r)>s_0+\alpha s_r$, and

$F(S)\to s_0+\alpha\mu\tau_v$ as $S\to+\infty$, ensures that for any $S_1^+ \in [s_0+\alpha s_r, +\infty)$, the sequence $\{S_n^+\}$ is either increasing or decreasing and converges to some value in $[s_0+\alpha s_r, s_0+\alpha\mu\tau_v]$. Finally, all sequences converge to the same value, i.e. F has a unique fixed point, because F is concave.

This property comes from the fact that on the one hand $S \to s(t,S)$ is contracting, i.e. $|s(t,S)-s(t,S')| = \exp(-t/\tau_s)|S-S'| < |S-S'|$ for $t>0$, and $v(t)$ (given by equation(4.2)) is concave, so that F', the derivative of F, is monotonic decreasing.

In summary, we have established that the dynamics of the modified LIF with constant current is governed by:

$$S_{n+1}^{+} = F(S_n^{+}) \tag{4.7}$$

where F, which maps $[s_0+\alpha\, s_r\,, +\infty)$ onto $[s_0+\alpha\, s_r\,, s_0+\alpha\, \mu\, \tau_v)$, is a concave monotonic increasing function with a unique fixed point that we denote by S*. Thus for all $S_1^* < S^*$, the sequence $\{S_n^+\}$ increases and converges to S*, and, conversely, for all $S_1^* > S^*$ the sequence S_n^+ decreases towards S*. The first consequence of the above result is that in response to a constant stimulation the modified LIF stabilizes at a periodic firing with constant interspike intervals I* given by:

$$I^* = \tau_v \ln\left[\frac{\alpha(v_0 - \mu\tau_v)}{S^* - s_0 - \alpha\mu\tau_v}\right] \tag{4.8}$$

We note that, since $I_n(S_n^+)$ is independent of α, the map F defined in equation (4.6) is an increasing function of α. A consequence of this observation is that the equilibrium value of the postdischarge threshold S* increases with α. Given that I* satisfies $I^*=I(S^*)$, the firing period also increases with α. In other words, the stationary firing slows down as the factor representing memory or threshold fatigue increases.

4.2.3 Dependence of I* on μ

We now study how the equilibrium ISI I* varies with the input current μ. Although it is not possible in general to solve equation (4.5) to find I* as a function of μ, it is possible to solve for μ as a function of I*.

We obtain:

$$\mu(I^*) = \frac{-s_0 + s_r - s_r e^{I^*/\tau_s} + v_0 e^{I^*/\tau_s} e^{-I^*/\tau_v} - \alpha v_0 e^{-I^*/\tau_v}}{\tau_v \left(\alpha - e^{I^*/\tau_s} + e^{I^*/\tau_s} e^{-I^*/\tau_v} - \alpha e^{-I^*/\tau_v}\right)} \tag{4.9}$$

The derivative of μ with respect to I^* is always negative under the assumptions $v_0 \leq 0 < s_r \leq s_0$, $\alpha > 0$, and $\Delta^* \geq \tau_s \ln(\alpha)$. We first show that the ISI I^* can never be less than $\tau_s \ln(\alpha)$.

If $\alpha \leq 1$ this condition is trivially satisfied as $\tau_s \ln(\alpha) \leq 0$ and $I^* > 0$. We thus concentrate on the case $\alpha > 1$. The denominator of equation (4.9) is zero when $I^* = \tau_s \ln(\alpha)$, so the function $\mu(\Delta^*)$ has a pole at this point. Since it is monotonically decreasing and continuously differentiable on the range $(\tau_s \ln(\alpha), +\infty)$ and $\lim_{\Delta^* \to +\infty} \mu(I^*) = s_r/\tau_v > 0$, we have that $\lim_{\Delta^* \to \tau_s \ln(\alpha)} \mu(I^*) = +\infty$. Thus, as we increase μ from s_r/τ_v (the minimal value for sustained firing) to $+\infty$, the ISI I^* decreases from $+\infty$ to $\tau_s \ln(\alpha)$. The ISI I^* can thus never be less than $\tau_s \ln(\alpha)$ if $\alpha > 1$. This has important implications for the model dynamics. In particular, it implies that the stationary ISI I^* remains greater than $\tau_s \ln(\alpha)$ as we increase μ to arbitrarily large values.

Moreover, the fact that $\mu(I^*)$ is a decreasing function of I^* implies that the firing frequency $1/I^*$ is always a monotonically increasing function of μ. However, if $\alpha > 1$, the firing frequency will saturate to a finite value given by $[\tau_s \ln(\alpha)]^{-1}$. This situation is illustrated in Figure 4-2

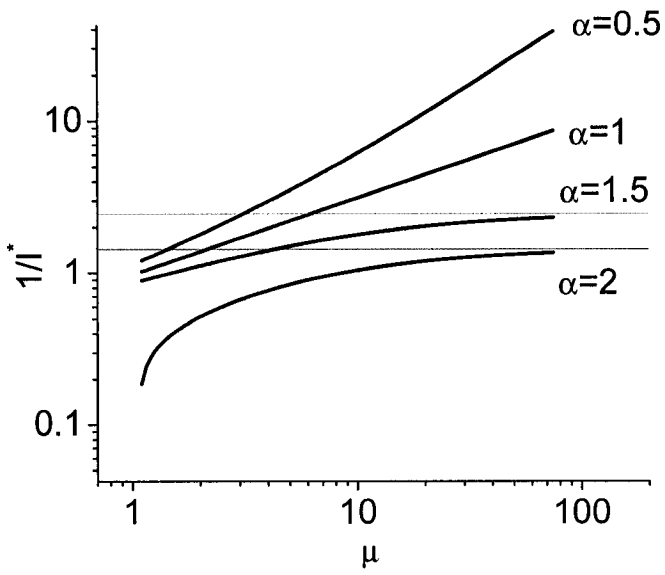


Figure 4-2: The firing frequency 1/I* as a function of μ for different values of α. A saturation is seen for α>1 (the two grey lines represent the theoretical values for saturation). Parameter values used were τ_s=2, τ_v=1, α=1, s_r=0.2, s_0=0.1.

As mentioned above, this feature is absent from the standard LIF model ($\alpha=0, s_r=s_0$) where the firing rate diverges as a function of the input current. A saturation in the firing frequency is commonly seen in experimental recordings and is usually associated with absolute refractoriness (i.e. a dead time immediately after an action potential has occurred during which no action potential can be generated). Our analysis shows that the saturation in firing rate does not imply that an absolute refractory period is present in the model. Indeed, the model can give rise to an ISI smaller than $\tau_s \ln(\alpha)$ following a sufficiently high current pulse or step. Thus, we show that even in the absence of an absolute refractory period in the classical sense, threshold fatigue can give rise to a saturation in the firing frequency under constant current input. This has important implications: it may be that the saturation in firing rate seen in experimental recordings under constant depolarizing current is due to threshold fatigue rather than the absolute refractory period of the neuron.

4.2.4 Response to step currents

A modification $\Delta\mu$ of the value of μ leads to a transient modification of the discharge rate of the model. Notably, when $\Delta\mu > 0$, the sequence S_n^+ increases from the previous value of S^* towards the new value, resulting in a progressive lengthening of intervals to the new stationary ISI. Note that this ISI will be shorter than the old one. If initially $\mu\tau_v < s_r$ is subthreshold and the new value is suprathreshold, the transient regime in which the ISIs are shorter than in the stationary regime corresponds to neuronal adaptation. In this case, this adaptation consists of a transient frequency increase at the onset of stimulation as illustrated in Figure 4-3.

The above results hold for all $\alpha > 0$. However, some distinctions exist between the different values of α which clarify the role of this parameter. One point is that for $\alpha \leq 1$, the map F is contracting, that is, $F'(S) < 1$ for all S. This is seen from the following argument: consider two initial values S and S' and their corresponding ISIs I and I'. Without loss of generality, we assume $S > S'$. Then $I' > I$ and $F(S')-F(S) < \alpha(S'-S) \exp(-I'/\tau_s) < S'-S$ if $\alpha \leq 1$. However, this is not necessarily so for $\alpha > 1$, where the map may be expanding in some intervals of S. This means that the effects such as firing rate adaptation tend to be more pronounced in maps with large α than in those with small α. Figure 4-3A shows the response of the model to a step current going from the subthreshold regime to the suprathreshold regime. Note that the ISIs progressively increase to the new equilibrium value as is seen in experimental recordings (see e.g. Adrian and Zotterman 1926). Figure 4-3B shows the response of the model when the value of the current before the step was set to be in the suprathreshold regime with all other parameters unchanged. Figure 4-3C shows the response of the model to the same step as in Figure 4-3B but with α increased. Note the faster rate of adaptation. The rate at which the model adapts to a step current can thus be varied by changing the parameter α.

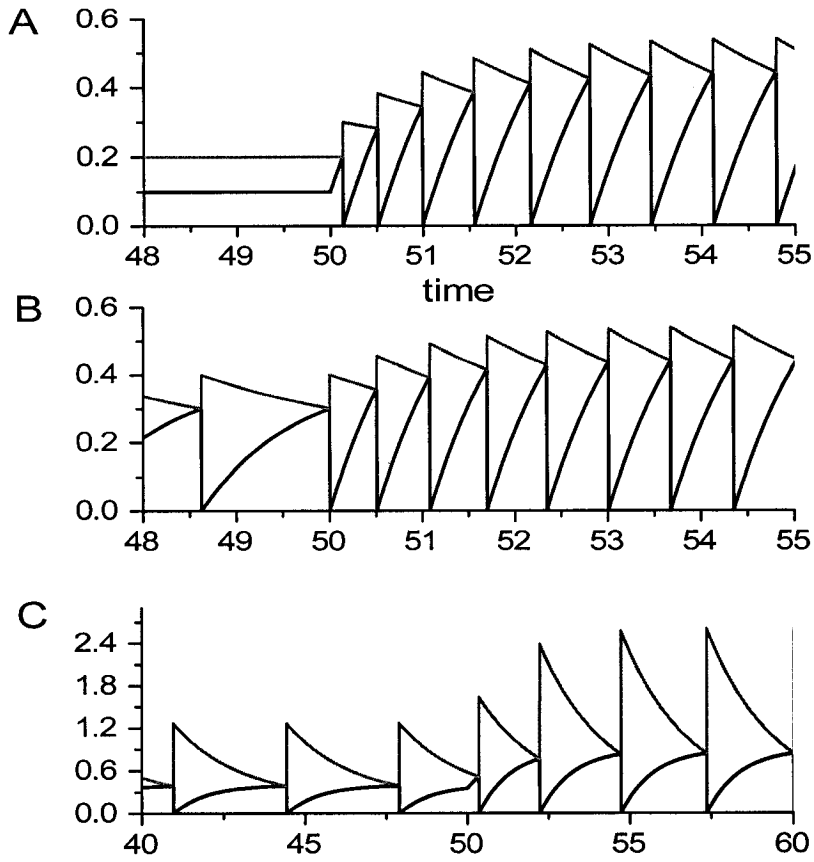


Figure 4-3: Response of the model to a step increase in current. A: the current goes from subthreshold (μ=0.1) to suprathreshold (μ=0.9) at t=50. Note the adaptation in the threshold (upper grey curve) and the progressive lengthening of ISI's to the new equilibrium value. The voltage is the lower black curve. Parameter values used were τ_s=2, τ_v=1, α=1, s_r=0.2, s_0=0.1. B: the current goes from μ=0.4 to μ=0.9 at t=50 with all other parameters unchanged. C: illustration of the effects of increased α. μ goes from 0.4 to 0.9 at t=50, but α=10 with all other parameters unchanged. Note the increased rate of adaptation.

4.3 *Noisy input current*

Throughout this section, we assume $W=W_1$ and $\mu(t)=\mu+\sigma\,\xi(t)$ where μ is constant and $\xi(t)$ is white Gaussian noise with unit intensity. In this situation, the voltage and threshold of the

model are stochastic processes that depend on the particular noise realization. We will first consider the first passage time distribution of the stochastic process described by our model.

4.3.1 The first passage time distribution

The interspike intervals are defined as the first passage times (FPTs) of the voltage through the threshold. When $\alpha=0$, that is, in the absence of threshold fatigue, the ISIs are independent identically distributed random variables. We denote by $g(t|s_0)$ the probability density function (pdf) of these variables, that is, the function $g(t|s_0)$ is the FPT pdf of the Ornstein-Uhlenbeck process η defined by:

$$\frac{d\eta}{dt} = -\frac{\eta}{\tau_v} + \mu + \sigma\xi(t) \quad \text{with } \eta(0) = \eta_0 \tag{4.10}$$

through the threshold $s(t,s_0) = (s_0 - s_r) \exp(-t/\tau_s) + s_r$. The FPT pdf $g(t|s_0)$ contains all information concerning the point process formed by the discharge times of the model.

When $\alpha>0$, the description of the discharge times needs to take into account the variations in the postdischarge time threshold. In the same way as for constant stimulation, the key point for the description of the behavior of the model is in establishing the relation between consecutive postdischarge thresholds. In this case, given that the stimulation is noise, the description involves the construction of a Markov chain rather than a map. This is detailed in the following.

The conditional pdf of S_{n+1}^+ given S_n^+ can be written as:

$$\begin{aligned} \prod_1(S_{n+1}^+ \mid S_n^+) &= \left|\frac{dI_n}{dS_{n+1}^+}\right| g(I_n \mid S_n^+) \\ &= \frac{\tau_s}{S_{n+1}^+ - s_0 - \alpha s_r} g\left[\tau_s \ln \frac{\alpha(S_n^+ - s_r)}{S_{n+1}^+ - s_0 - \alpha s_r} \mid S_n^+\right] \end{aligned} \tag{4.11}$$

where we have used equation (4.4) and g represents the FPT pdf of the Ornstein-Uhlenbeck process through the threshold.

The pdf Π_1 defines the first order transition probabilities of a Markov chain. We denote by h*(S) the stationary distribution of this chain, in other words, in the long run, the probability for the postdischarge threshold to take a value within (S, S+dS) is given by h*(S)dS. The characteristics of the point process formed by the discharge times of the model are then defined in terms of this stationary distribution together with g and Π_1. For instance, the pdf of ISI distribution g*(t) is given by:

$$g^*(t) = \int g(t \mid S) h^*(S)\, dS \tag{4.12}$$

4.3.2 The ISI SCC

We are interested in the serial correlation coefficient of ISI's defined by:

$$\rho_p = \frac{< I_n I_{n+p} > - < I_n >^2}{< I_n^2 > - < I_n >^2} \tag{4.13}$$

where $<I_n I_{n+p}>$, $<I_n^2>$ and $<I_n>$ are the expectations of $I_n I_{n+p}$, I_n^2 and I_n, respectively.

For n large, the last two quantities are respectively determined by the following two expressions: $<I_n^2> = \int t^2 g^*(t)\, dt$ and $<I_n> = \int t\, g^*(t)\, dt$. For the first term, i.e. the expectation of the product of two intervals, we first consider the case where n is large and p=1. Given S_n^+ the postdischarge threshold after the nth discharge, the ISI between the nth and n+1th firings, i.e. I_n, is distributed according to $g(t|S_n^+)$. The knowledge of S_n^+ and I_n uniquely determine S_{n+1}^+ as:

$$S_{n+1}^+ = f(S_n^+, I_n) \tag{4.14}$$

Where f is given by equation (4.4). This knowledge in turn yields that the ISI between the n+1th and n+2th firings, i.e. I_{n+1} has pdf $g(t|S_{n+1}^+)=g[t|\ f(S_n^+,I_n)]$. Combining this knowledge with the fact that for n large, S_n^+ is distributed according to h* yields:

$$< I_n I_{n+1} >= \int_S \int_I \int_{I'} I I' g[I'|f(S,I)]g[I|S]h^*(S)\,dI'\,dI\,dS \qquad (4.15)$$

In a similar way, we derive:

$$\begin{aligned} < I_n I_{n+p} >= & \int_S \int_I \int_{S'} \int_{I'} I I' g[I'|S'] \times \\ & \prod_{p-1}[S'|f(S,I)]g[I|S]h^*(S)\,dI'\,dS'\,dI\,dS \end{aligned} \qquad (4.16)$$

where Π_{p-1} is the p-1th transition pdf, i.e. $\Pi_0(u|S)=\delta(u-S)$ the Dirac function at S, and $\Pi_k(u|S)=\int_{S'} \Pi_1(u|S)\ \Pi_k(S'|S)\ dS'$. In the absence of threshold fatigue, i.e. when $\alpha=0$, the postdischarge threshold is a fixed value s_0, so that $\Pi_1(u|S)= \delta(u-s_0)$, $h^*(S)= \delta(S-s_0)$ and $f(S,I)=s_0$. Substituting these quantities into equation (4.16) yields that, for $p\leq 1$, $<I_n I_{n+p}>=<I_n><I_{n+p}>$ and consequently that $\rho_p=0$.

The analysis of the previous paragraphs establishes that the response to Gaussian white noise of the modified LIF with threshold fatigue is not described by a renewal process when $\alpha>0$. This is in contrast with the standard LIF whose response to white Gaussian noise can be described by a renewal process. It also reveals that the modified model with noisy forcing was characterized by a Markov chain relating consecutive postdischarge thresholds. Finally, it showed that the dependence of a postdischarge threshold on the previous value induces correlation between ISIs.

Once it is theoretically established that ISIs of the modified model are correlated with one another, we examine the nature of this dependence. For this purpose, rather than using the integral expressions obtained previously, it is appropriate to use numerical simulations of the model. Indeed, the integrals require the computation of the FPT pdf of the Ornstein-Uhlenbeck process through an exponential boundary. Given that no general analytical expression is available for this quantity (such an expression is not available even for an OU

process through a constant boundary), derivation of the correlation from the integrals can be computationally more demanding than estimating the same quantities from simulations.

Figure 4-4a shows the ISI distribution $g^*(t)$ obtained when the system given by equation (4.1) is driven by Gaussian white noise for $\alpha=1$ while Figure 4-4b shows the ISI correlation coefficients ρ_p.

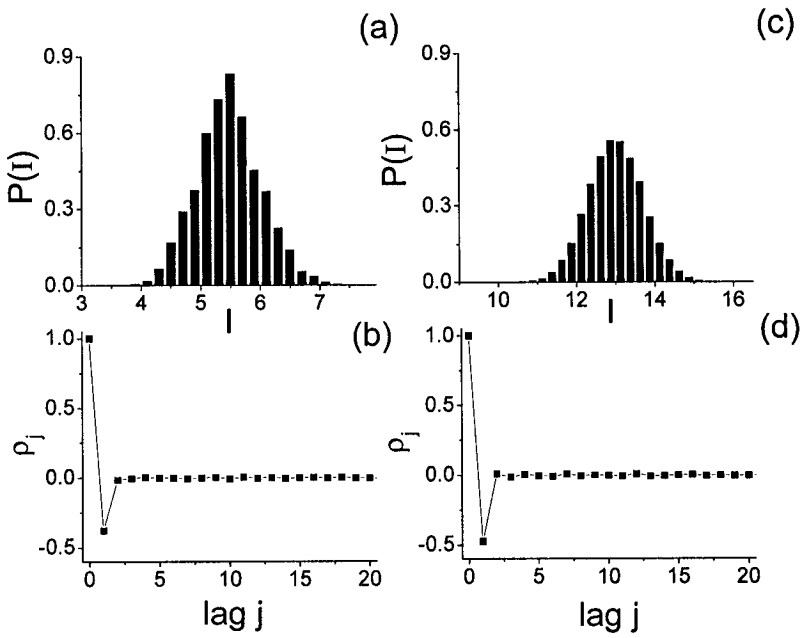


Figure 4-4: A: ISI distribution obtained for $W=W_1$ and $\alpha=1$ in the presence of Gaussian white noise of standard deviation 0.1. B: Correlation coefficients ρ_j as a function of lag. Note that only $\rho_1=-0.38$ is negative and that all coefficients are zero for higher lags. C: ISI distribution obtained for $\alpha=4$. D: Correlation coefficients. Note that $\rho_1=-0.48$ is lower than for $\alpha=1$. Other parameters values were: $\tau_s=8$, $\tau_v=1$, $\mu=1$, $s_r=0$, $s_0=1$.

The estimated ρ_1 is negative. Serial correlation coefficients at higher lags are close to zero illustrating the rapid return of the system to equilibrium. The fact that ρ_1 is negative and that it is the only coefficient substantially different from zero implies that ISI's shorter (longer) than average will be followed by ISI's longer (shorter) than average. Figure 4-4c shows the ISI

distribution obtained for α=4. We can see that increasing α increases the mean ISI, as predicted earlier since this will increase the mean threshold value. Also, the magnitude of the coefficient ρ_1 increases with α, which is consistent with the predictions that more memory will give rise to more pronounced ISI correlations.

The magnitude of the serial correlation coefficient at lag one depends on the slope of the map $S \rightarrow F(S)$ and on the strength of the noise. For high noise intensity, it is expected that the noise will "wash out" the correlations induced by the map. We thus expect ρ_1 to increase as a function of noise intensity from a negative value at zero mean. This is shown in Figure 4-5 where noise "expresses" the deterministic properties of the map (i.e. it produces sufficient perturbations from the map's fixed point to see transient ISIs and their correlations). The value of ρ_1 has been linked to the detectability of weak signals in neurons as well as their information transfer capabilities (Chacron et al. 2001a; see chapter 6).

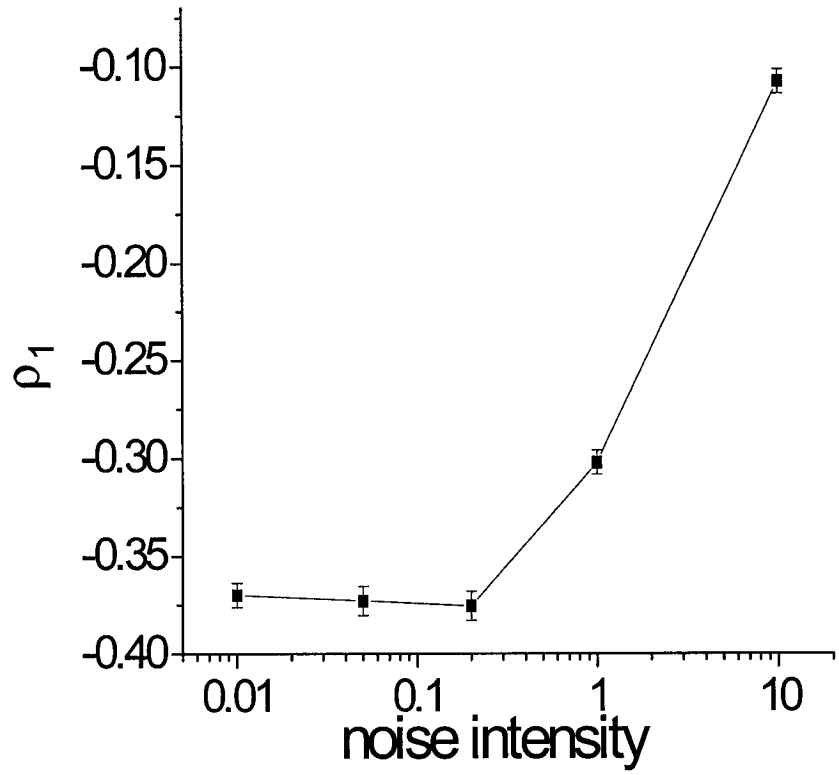


Figure 4-5 ρ_1 as a function of the noise standard deviation. ρ_1 increases as a function of noise intensity. Parameter values were the same as in Figure 4-4 with α=1. 20000 ISIs were used in each case.

4.3.3 Response to perturbations

To understand how the deterministic properties of the map can give rise to ISI correlations in the presence of noise, we study the response of the model to perturbations without noise (i.e. σ=0).

We assume that the stationary regime has been reached and that the model fires periodically with ISI I*, and that a single sufficiently large pulsatile stimulation is delivered a time lapse $I < I^*$ after a discharge, to make the neuron fire ahead of time. This results in a shortened ISI with length I, and consequently the postdischarge threshold S_I^+ is larger than S*. Following the perturbation, the sequence S_n progressively decreases towards S*. Given that the interspike

interval I(S) is a monotonic increasing function of S, the sudden increase in S at the time of the perturbation and its progressive decrease indicate that the shortened interval is followed by one that is longer than the steady state period, and that subsequent intervals progressively decrease towards this steady state period. If the system returns quasi-instantaneously to equilibrium, then only ρ_1 will be significantly negative. We concentrate on the regime in which the voltage has almost reached its asymptotic value $\mu\tau_v$ before a firing; this occurs when $\tau_s \gg \tau_v$. We start from equation (4.3) in which we set $v(\Delta_n)=\mu\,\tau_v$:

$$s(I_n, S_n^+) = \mu\tau_v \tag{4.17}$$

The stationary threshold value is then $S^*=s_0+\alpha\mu\tau_v$ and equation (4.8) gives us the stationary ISI I*:

$$I^* = \tau_s \ln\left(\frac{s_0 + \alpha\mu\tau_v - s_r}{\mu\tau_v - s_r}\right) \tag{4.18}$$

Let us suppose that the voltage perturbation advances or delays an action potential, thus producing an ISI I which can be smaller or greater than I*. Then the new value of the threshold immediately after the action potential is given by:

$$S_I^+ = s_0 + \alpha\left[s_r + (s_0+\alpha\mu\tau_v - s_r)\exp\left(-\frac{I}{\tau_s}\right)\right] \tag{4.19}$$

The next ISI, in the absence of further perturbations, will then be given by:

$$I_{next} = \tau_s \ln\left\{\frac{S_I^+ - s_r}{\mu\tau_v - s_r}\right\} \tag{4.20}$$

Equations (4.19) and (4.20) give us the relation between I_{next} and I:

$$I_{next} = \tau_s \ln\left\{\frac{s_0 - s_r + \alpha\left[s_r + (s_0 + \alpha\mu\tau_v - s_r)\exp\left(-\frac{I}{\tau_s}\right)\right]}{\mu\tau_v - s_r}\right\} \tag{4.21}$$

The derivative of the map $I_{next}=U(I)$ is given by:

$$U'(I) = -\frac{\alpha(s_0 + \alpha\mu\tau_v - s_r)\exp(-I/\tau_s)}{s_0 - s_r + \alpha\left[s_r + (s_0 + \alpha\mu\tau_v - s_r)\exp(-I/\tau_s)\right]} \tag{4.22}$$

and is negative for $s_0 > s_r > 0$ and $\alpha\mu\tau_v > 0$. We now consider the effects of α on the ISI I_{next}.

If $\alpha=0$, we have:

$$I_{next} = I^* = \tau_s \ln\left\{\frac{s_0 - s_r}{\mu\tau_v - s_r}\right\} \tag{4.23}$$

i.e. I_{next} is independent of I since we have $U'(I)=0$. This is because the system returns instantaneously to equilibrium and there is no memory extending beyond one ISI. We now consider the effects of $\alpha>0$. I_{next} increases with decreasing I. Thus, a perturbation that causes I larger (smaller) than I* will make I_{next} smaller (larger) than I*. Furthermore, increasing α increases I_{next}. We illustrate this point with numerical examples.

Figure 5 shows a time series in which a perturbation in the voltage was applied at t=1099.

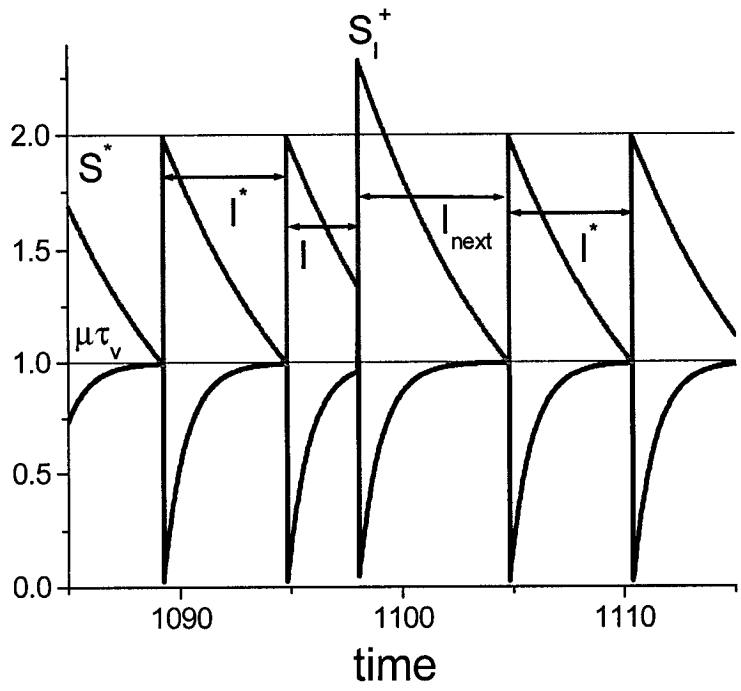


Figure 4-6: Voltage (black solid line) and threshold (grey solid line) time series. A perturbation in voltage was applied at t=1099. The perturbation caused an action potential to occur earlier than expected. As such, the ISI I was shorter than I* and the value of the threshold immediately after that action potential, S_I^+, was higher than the threshold equilibrium value S*. Consequently, the threshold took a longer time to decay to $\mu\tau_v$ and the next ISI I_{next} was longer than I*. Parameters values were the same as in Figure 4-5.

The perturbation increased voltage and caused an action potential to occur prematurely. The ISI I is thus shorter than I* and consequently S_I^+ (equation (4.19)) is greater than S*. Since the threshold starts from a higher value, it will take a longer time to decay to the value $\mu\tau_v$. As such, the next ISI I_{next} will be greater than I*. Similarly, had Δ been longer than I*, S_I^+ would have been smaller than S* and consequently I_{next} would have been smaller than I*. The actual map $I_{next}=U(I)$ is shown in Figure 4-7a and has a negative slope as expected. Thus, the deterministic properties of the map lead to a perturbed ISI longer (shorter) than average being followed by an ISI shorter (longer) than average.

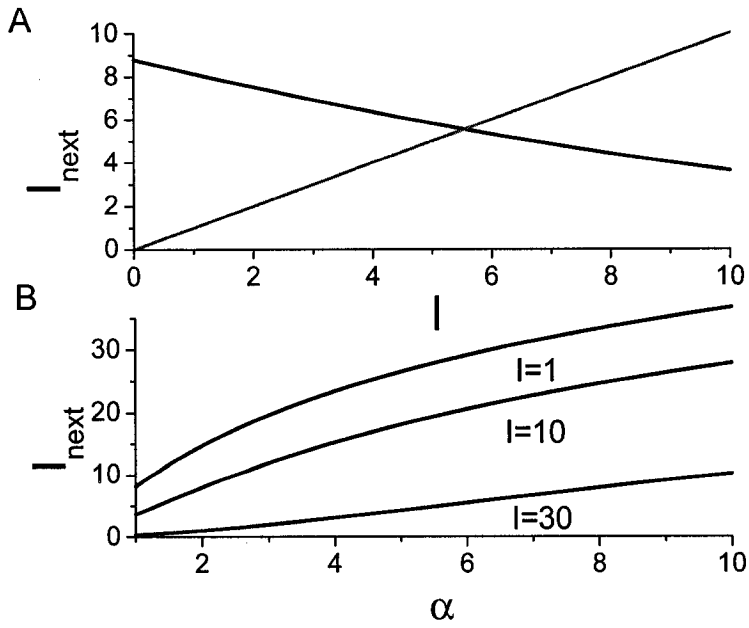


Figure 4-7: A: ISI I_{next} as a function of I (solid black line). The diagonal with unit slope is also represented in grey. The slope of the map is negative as expected (see text). B: ISI I_{next} as a function of α for different values of I. Increasing α will increase I_{next}. Parameters have the same value as in Figure 4-6 except α.

Figure 4-7b illustrates the dependence of I_{next} on α for different values of I. As discussed above, I_{next} increases with α. The rate at which I_{next} increases is however greater for low I than for higher I. This shows how the deterministic properties of the map in response to perturbations can give rise to ISI correlations in the presence of noise as demonstrated in the previous section.

4.4 Sinusoidal stimulation

In this chapter we study the model's response under sinusoidal stimulation. This has particular relevance for weakly electric fish, as the EOD is sinusoidal and drives the electroreceptors. We thus take $\mu(t) = \mu + a\sin(w\,t)$ with $a > 0$. We will also assume $W(s,\alpha) = W_2(s-s_r,\alpha) = \exp(\alpha(s-s_r))-1$.

4.4.1 An annulus map

The voltage and threshold solutions to equations (4.1) are given by:

$$\begin{aligned} v(t,t_n) &= \mu\tau_v\left(1-\exp\left[(t_n-t)/\tau_v\right]\right)+ \\ &\frac{a\tau_v}{1+w^2\tau_v^2}\left[\sin(wt)-w\tau_v\cos(wt)\right]- \\ &\frac{a\tau_v\exp\left[(t_n-t)/\tau_v\right]}{1+w^2\tau_v^2}\left[\sin(wt_n)-w\tau_v\cos(wt_n)\right] \\ s(t,t_n,S_n^+) &= s_r+\left(S_n^+-s_r\right)\exp\left(\frac{t_n-t}{\tau_s}\right) \end{aligned} \tag{4.24}$$

with $v(0,t_n)=0$ and $s(0,t_n,S^+_n)=S^+_n$ and provided that $t \geq t_n$ with no firing on the time interval $[t_n,t]$. We note that we now need three arguments in the threshold solution s as the system is no longer time translation invariant, this is in contrast to the previous sections. A necessary and sufficient condition for sustained firing is given by:

$$\mu\tau_v + a\tau_v/\sqrt{1+w^2\tau_v^2} > s_r \tag{4.25}$$

and we will assume that this condition holds throughout. The firing time t_{n+1} is given by:

$$t_{n+1} = H_1(t_n,S_n^+) \equiv \inf\left\{t : t > t_n, v(t,t_n) = s(t,t_n,S_n^+)\right\} \tag{4.26}$$

and the threshold $s(t^+_{n+1},t_n,S^+_n)\equiv S^+_{n+1}$ immediately after the firing is given by:

$$S_{n+1}^+ = H_2(t_n,S_n^+) \equiv s_0 + W_2\left[s\left(t_{n+1},t_n,S_n^+\right),\alpha\right] \tag{4.27}$$

A two-dimensional map H can thus be defined as: $(t_n,S^+_n)\rightarrow(t_{n+1}=H_1(t_n,S^+_n),S^+_{n+1}=H_2(t_n,S^+_n))$ on the interval $[0,\infty)\times[s_0+W(s_r,\alpha),s_0+W(V_{max},\alpha)]$, where $V_{max}=\mu\tau_v+a\tau_v/(1+w^2\tau_v^2)^{1/2}$ is the maximum value of the asymptotic voltage solution.

It is then convenient to introduce the firing phase as in Coombes (1999):

$$\phi_n = 2\pi\left\{\frac{t_n}{T} - \text{int}\left[\frac{t_n}{T}\right]\right\} \tag{4.28}$$

where the int[.] denote the integer part and $T\equiv 1/f$ is the forcing period. We can thus associate to H a map $h=(h_1,h_2)$ on the annulus $[0,2\pi)\times[s_0+W(s_r,\alpha),s_0+W(V_{max},\alpha)]$:

$$\begin{aligned} h_1(\phi_n, S_n^+) &= \frac{2\pi}{T} H_1\left(\frac{T}{2\pi}\phi_n, S_n^+\right) \quad \text{mod}(T) \\ h_2(\phi_n, S_n^+) &= H_2\left(\frac{T}{2\pi}\phi_n, S_n^+\right) \end{aligned} \tag{4.29}$$

4.4.2 Lyapunov exponents of the map

We now derive the Lyapunov exponents for the model given by equation (4.1) with periodic forcing. Lyapunov exponents measure the rate of convergence or divergence of solutions that start close to one another at t=0. Thus, for a chaotic system, at least one Lyapunov exponent is positive and thus nearby solutions diverge from each other (see e.g. Strogatz 1994). We proceed similarly to Coombes (1999) for the standard LIF model (i.e. without threshold dynamics) using ideas developed for the study of impact oscillators (Müller 1995). For the standard LIF model, it is only necessary to consider a voltage solution v(t) and a perturbed solution $v_p(t)$ with $\delta v(t)\equiv v_p(t)-v(t)$. It is then possible to derive a recurrence relation between $\delta v(t_{n+1})$ and $\delta v(t_n)$ keeping only first order terms and thus to define a Lyapunov exponent (Coombes 1999). In our case, we must also consider an unperturbed threshold solution s(t) (we denote $s(t,t_n,S_n^+)$ by s(t)) and a perturbed threshold solution $s_p(t)$ with $\delta s(t)\equiv s_p(t)-s(t)$. The unperturbed solutions v(t) and s(t) intersect at times t_n and t_{n+1} while the perturbed solutions $v_p(t)$ and $s_p(t)$ will intersect at times $t_n+\delta_n$ and $t_{n+1}+\delta_{n+1}$. The situation is illustrated in Figure 4-8.

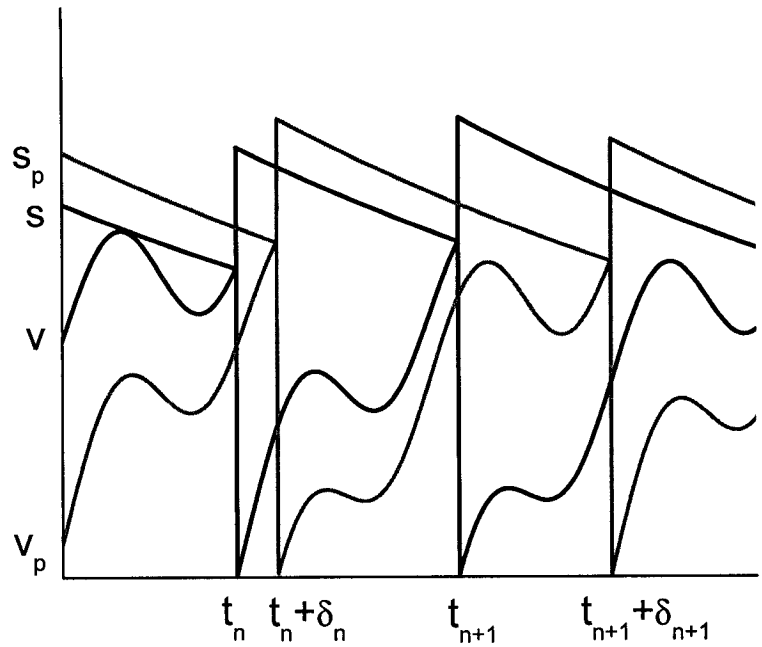


Figure 4-8: Illustration of the situation to compute Lyapunov exponents. The unperturbed solutions v(t) and s(t) intersect at times $\{t_n\}$ while the perturbed solutions intersect at times $\{t_n+\delta_n\}$.

The goal is to derive a relation between $(\delta v(t_n), \delta s(t_n))$ and $(\delta v(t_{n+1}), \delta s(t_{n+1}))$ in order to look at the time evolution of perturbations keeping only first order terms. At time $t_n+\delta_n$, we have:

$$\begin{aligned} 0 &= s_p(t_n+\delta_n) - v_p(t_n+\delta_n) \\ 0 &\approx s_p(t_n) + \delta_n s_p(t_n) - v_p(t_n) - \delta_n v_p(t_n) \\ 0 &\approx \delta s(t_n) - \delta v(t_n) + \delta_n [s(t_n) - v(t_n)] \end{aligned} \tag{4.30}$$

Thus the perturbation in the firing time δ_n is approximately given by:

$$\delta_n \approx \frac{\delta s(t_n) - \delta v(t_n)}{v(t_n) - s(t_n)} \tag{4.31}$$

where we have assumed that the denominator is not zero. At $t=(t_n+\delta_n)^+$, we have:

$$
\begin{aligned}
\delta s\left[(t_n+\delta_n)^+\right] &= s_p\left[(t_n+\delta_n)^+\right]-s\left[(t_n+\delta_n)^+\right] \\
&\approx s_0+W_2\left[s_p(t_n)+\delta_n s_p(t_n)-s_r,\alpha\right]- \\
&\quad s(t_n^+)-\delta_n s(t_n^+) \\
&\approx s_0+W_2\left[s_p(t_n)-s_r,\alpha\right]+ \\
&\quad \delta_n s_p(t_n)W_2'\left[s_p(t_n)-s_r,\alpha\right]- \\
&\quad s(t_n^+)-\delta_n s(t_n^+) \\
&\approx \delta s(t_n)W_2'\left[s(t_n)-s_r,\alpha\right]+ \\
&\quad \delta_n\left\{s(t_n)W_2'\left[s(t_n)-s_r,\alpha\right]-s(t_n^+)\right\}
\end{aligned}
\tag{4.32}
$$

where $W'[s,\alpha]\equiv\partial W/\partial s$. Furthermore, we have:

$$
\begin{aligned}
\delta v\left[(t_n+\delta_n)^+\right] &= v_p\left[(t_n+\delta_n)^+\right]-v\left[(t_n+\delta_n)^+\right] \\
&\approx -\delta_n v(t_n^+)
\end{aligned}
\tag{4.33}
$$

Finally, we have:

$$
\begin{aligned}
\delta v(t_{n+1}) &= \delta v\left[(t_n+\delta_n)^+\right]\exp\left[(t_n+\delta_n-t_{n+1})/\tau_v\right] \\
&\approx \delta v\left[(t_n+\delta_n)^+\right]\exp\left[(t_n-t_{n+1})/\tau_v\right] \\
\delta s(t_{n+1}) &= \delta s\left[(t_n+\delta_n)^+\right]\exp\left[(t_n+\delta_n-t_{n+1})/\tau_s\right] \\
&\approx \delta s\left[(t_n+\delta_n)^+\right]\exp\left[(t_n-t_{n+1})/\tau_s\right]
\end{aligned}
\tag{4.34}
$$

Using eqs (4.30) to (4.34) in conjunction with equation (4.1), we get that:

$$
\begin{aligned}
\delta v(t_{n+1}) &= \left\{a_n\delta v(t_n)-a_n\delta s(t_n)\right\} \\
\delta s(t_{n+1}) &= \left\{-b_n\delta v(t_n)+(b_n+c_n)\,\delta s(t_n)\right\}
\end{aligned}
\tag{4.35}
$$

with $\Delta_n\equiv t_{n+1}-t_n$ and a_n, b_n, c_n given by:

$$
\begin{aligned}
a_n &= \frac{-v_0\tau_v^{-1} + \mu + a\sin(wt_n)}{s(t_n)[\tau_s^{-1} - \tau_v^{-1}] - s_r\tau_s^{-1} + \mu + a\sin(wt_n)} e^{-I_n/\tau_v} \\
b_n &= \frac{[s_r - s(t_n)]W_2'[s(t_n) - s_r, \alpha] + s_0 + W_2[s(t_n) - s_r, \alpha] - s_r}{s(t_n)[1 - \tau_s/\tau_v] - s_r + \mu\tau_s + a\tau_s\sin(wt_n)} e^{-I_n/\tau_v} \\
c_n &= W_2'[s(t_n) - s_r, \alpha] e^{-I_n/\tau_s}
\end{aligned}
\tag{4.36}
$$

The system (4.35) is equivalent to the following matrix equation:

$$
[\delta_v(t_{n+1}), \delta_s(t_{n+1})]^T = M_n [\delta_v(t_n), \delta_s(t_n)]^T \tag{4.37}
$$

where the superscript T denotes the transpose and the matrix M_n is given by:

$$
M_n = \begin{pmatrix} a_n & -a_n \\ -b_n & b_n + c_n \end{pmatrix} \tag{4.38}
$$

This allows us to define the Lyapunov exponent of the system as:

$$
\begin{aligned}
\Gamma_1 &= \lim_{n\to\infty} \frac{1}{t_n - t_1} \ln|\Lambda_1(t_n)| \\
\Gamma_2 &= \lim_{n\to\infty} \frac{1}{t_n - t_1} \ln|\Lambda_2(t_n)|
\end{aligned}
\tag{4.39}
$$

where $\Lambda_1(t_n)$, $\Lambda_2(t_n)$ are the eigenvalues of the matrix R_n given by:

$$
R_n = \prod_{i=1}^{n} M_i \tag{4.40}
$$

The stability of solutions can also be considered from the kernels κ_1 and κ_2 defined by:

$$\kappa_1 = \lim_{n\to\infty} \Lambda_1(t_n)^{\frac{1}{t_n - t_1}}$$
$$\kappa_2 = \lim_{n\to\infty} \Lambda_2(t_n)^{\frac{1}{t_n - t_1}} \tag{4.41}$$

The stability of a solution thus depends on $|\Re(\kappa_1)|$ and $|\Re(\kappa_2)|$ where the $\Re$ denotes the real part. If both are less than one, then the solution is stable (note that this corresponds to both Lyapunov exponents being negative). Otherwise, the solution is unstable. Bifurcations are thus expected to occur when $|\Re(\kappa_1)|$ and $|\Re(\kappa_2)|$ are equal to one. Tangent bifurcations occur when $\text{Max}[\Re(\kappa_1), \Re(\kappa_2)]=1$ and period-doubling ones occur when $\text{Min}[\Re(\kappa_1), \Re(\kappa_2)]=-1$.

4.4.3 Dynamics of the map

Figure 4-9 shows an example of the rich dynamics displayed by our LIFDT model without noise. Period doubling tangent bifurcations occur as well as chaotic dynamics. This is interesting since the LIF model with sinusoidal forcing cannot display chaos (Coombes 1999). However, before studying chaotic dynamics in more detail, we will study the stability of periodic solutions.

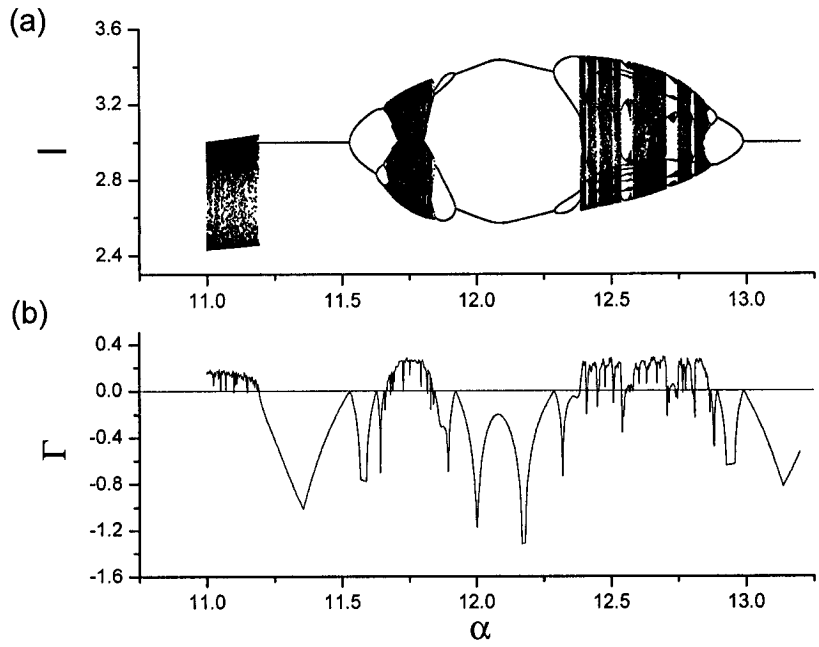


Figure 4-9: A: Bifurcation diagram showing the ISI I as a function of α. B: Largest Lyapunov exponent $\Gamma=\max(\Gamma_1, \Gamma_2)$ of the system. A positive Γ implies chaotic dynamics. Parameter values used were: $\tau_s=0.8$, $\tau_v=1$, $a=0.1732$, $w=2\pi$, $\mu=0$, $s_0=0.01$, $s_r=-0.1732$.

4.4.4 Reduction to a one-dimensional circle map

We now consider that the voltage has reached its asymptotic time varying form when an action potential occurs. Effectively, this corresponds to taking the limit $\tau_v \to 0$ while keeping the quantities $\mu\tau_v$, $a\tau_v$, and $w\tau_v$ constant. Thus, we have:

$$v(t) = \mu\tau_v + \frac{a\tau_v}{1+w^2\tau_v^2}\left[\sin(wt) - w\tau_v\cos(wt)\right]$$
$$s(t, t_n, S_n^+) = s_r + \left(S_n^+ - s_r\right)\exp\left(\frac{t_n - t}{\tau_s}\right) \qquad (4.42)$$

To simplify, we take $\tau_v=1$ in the quantities $\mu\tau_v$, $a\tau_v$, and $w\tau_v$ for the remainder of this section. The system (4.42) is equivalent to a one-dimensional map with the firing time t_{n+1} implicitly given by:

$$F(t_{n+1}) = \left(S_n^+ - s_r\right)\exp\left(\frac{t_n}{\tau_s}\right) \tag{4.43}$$

with F(t) given by:

$$F(t) = \left\{\mu - s_r + \frac{a}{1+w^2}\left[\sin(wt) - w\cos(wt)\right]\right\}\exp\left(\frac{t}{\tau_s}\right) \tag{4.44}$$

If F is invertible, we can write:

$$t_{n+1} = f(t_n) = F^{-1}\left[\left(S_n^+ - s_r\right)\exp\left(\frac{t_n}{\tau_s}\right)\right] \tag{4.45}$$

We have that $F'(t) > 0$ if:

$$\mu - s_r > \frac{a}{\sqrt{1+w^2}} \tag{4.46}$$

Hence F is invertible provided equation (4.46) holds by the implicit function theorem. Note that this is equivalent to equation (4.25) with $\tau_v=1$. Thus, the condition for sustained firing is the same as the condition for invertibility for this map. It is furthermore easily verified that F[f(t+T)]=F(f(t)+T). Hence, the map defined by equation (4.43) is equivalent to a circle map (Coddington and Levinsen, 1955; Keener et al, 1981). We thus have:

$$\phi_{n+1} = f(\phi_n) \mod(T) \tag{4.47}$$

We now derive the Lyapunov exponent of the map.

4.4.4.1 *Lyapunov exponents of the circle map*

We consider a threshold solution s(t) as well as a perturbed solution $s_p(t)$. A calculation similar to that performed in section 4.4.2 gives:

$$\delta s(t_{n+1}) = d_n \delta s(t_n) \tag{4.48}$$

with d_n given by:

$$\frac{W_2'[s(t_n)-s_r,\alpha]e^{-I_n/\tau_s} + W_2'[s(t_n)-s_r,\alpha]\{s_r - s(t_n)\} + s_0 - s_r + W_2[s(t_n)-s_r,\alpha]}{\tau_s\{-v(t_n)+a\sin(wt_n)\}+s(t_n)-s_r} e^{-I_n/\tau_s} \tag{4.49}$$

The unique Lyapunov exponent is then given by:

$$\Gamma_c = \lim_{n\to\infty} \frac{1}{t_n - t_1} \sum_{i=1}^{n} \ln|d_n| \tag{4.50}$$

We now illustrate the dynamics of this map with numerical simulations.

4.4.4.2 *Dynamics of the circle map*

Figure 4-10 shows the dynamics of the map for the same parameter set as for Figure 4-9. It is seen that the bifurcation diagrams are quite similar. In particular, the rich dynamical behavior obtained for the annulus map is also obtained for the circle map. The positivity of the Lyapunov exponent Γ_c for certain values of α indicates that the circle map can support chaotic solutions. We will now explain why our circle map can display chaotic solutions while the circle map obtained for the LIF model cannot.

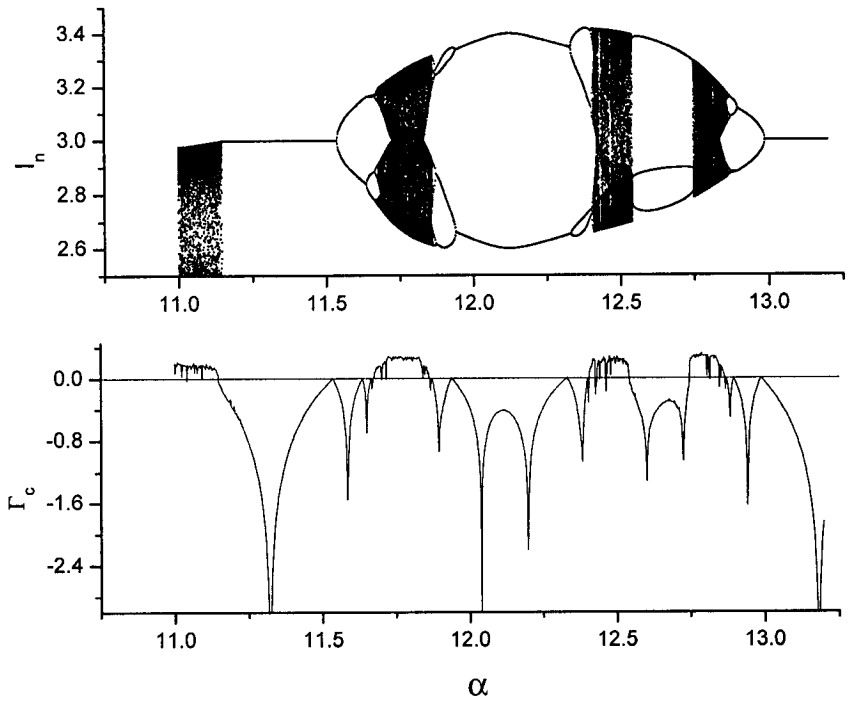


Figure 4-10: A: bifurcation diagram of the circle map obtained by varying α. It is seen that the circle map displays similar dynamics to the annulus map B: Lyapunov exponent of the map. Parameter values are the same as in Figure 4-9.

4.4.4.3 Origins of chaotic dynamics

It is known that the circle map associated to the LIF model cannot display chaotic dynamics (Keener et al. 1981; Coombes 1999). This is because the circle map $\phi_{n+1}=f_{LIF}(\phi_n)$ associated to the LIF neuron is orientation preserving when restricted to its range (Pakdaman 2001). However, our map is not necessarily orientation preserving since the function:

$$F[f(t)]=\{s_0-s_r+W[s(t)-s_r,\alpha]\}\exp\left(\frac{t}{\tau_s}\right) \qquad (4.51)$$

is not necessarily monotonic in t. Indeed, s(t) is always monotonic decreasing while W[s,α] is always monotonic increasing in s. The function s_0-s_r+W[s(t)-s_r,α] is thus monotonic decreasing in t. However, the function exp(t/τ_s) is monotonic increasing in t. For α=0, we have that F[f(t)] is always monotonic increasing, hence the map can never display chaotic dynamics as the LIF model. However, as α increases, chaotic dynamics can occur since the map is not necessarily orientation preserving.

4.5 Noisy Sinusoidal Stimulation

In this section, we will briefly consider the effects of combining noisy stimulation with sinusoidal stimulation. Thus, we take μ(t)=μ+a sin(wt)+σ ξ(t). This is relevant to results presented in chapter 6. We construct tuning curves for the LIFDT model under this input. We quantify the degree of phase locking by the mean vector strength (see Mardia and Jupp 1999 and references therein) defined by:

$$VS = < \cos(\phi_i) >_i^2 + < \sin(\phi_i) >_i^2 \qquad (4.52)$$

where ϕ_i is the phase defined by equation (4.28). Figure 4-11 shows the mean vector strength VS as a function of w. A maximum is seen for $w \approx 0.81 \approx 2\pi\tau_s^{-1}$. This can be explained intuitively as follows. The threshold will have time to recover considerably from cycle to cycle when the cycle period is on the order of the threshold decay time constant. On the other hand, the threshold will effectively decay to its equilibrium value during the stimulus cycle if w is too low, thus cancelling the beneficial effects brought by its dynamics. Moreover, the threshold will not have the time to decay for w too high, again cancelling the beneficial effects.

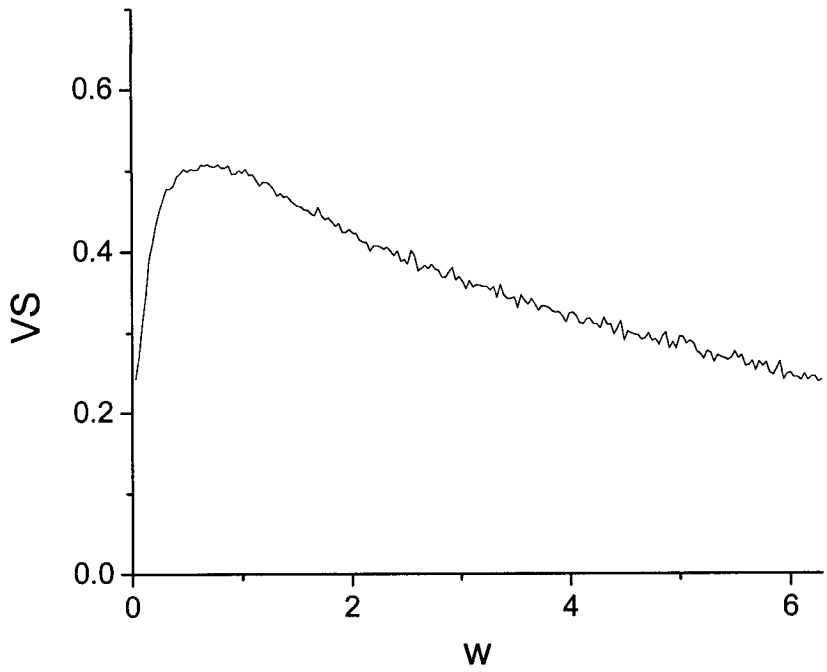


Figure 4-11: Mean vector strength VS obtained as a function of w. Parameter values were τ_s=7.75, τ_v=1, μ=0, s_r=0.03, s_0=0.05, a=0.2613, σ=0.18, α=1.

4.6 Discussion

4.6.1 Adaptation and firing rate saturation

In this study, we have shown that a simple single neuron model that accounts for the memory seen in several classes of neurons has surprisingly rich dynamical behavior. Previous studies involved either α=0 (Geisler and Goldberg; 1966) or α=1 (Geisler and Goldberg 1966; Segundo et al. 1976; Vibert et al. 1994; Pakdaman and Vibert 1995; Pakdaman, et al. 1996; Chacron et al. 2000; 2001a; 2001b; Liu and Wang 2001). We instead varied α continuously and studied its effects.

Different values for the parameter α (measuring the amount of memory in the system) gave rise to qualitatively different regimes. The rate and degree of adaptation to a step in input current were shown to vary with α. Adaptation is commonly seen in biological neurons. We have found that our model could reproduce this interesting feature and that the rate of adaptation depended on the amount of memory. Another common feature of biological neurons is the saturation in firing rate as the input current is increased. Our model reproduces this feature when $\alpha>1$, in contrast with the LIF neuron for which the firing rate diverges as a function of the input current. In experimental recordings, this saturation is usually thought to be associated with an absolute refractory period or network effects. Our results show that an absolute refractory period is not necessary to obtain such an effect. In fact, this prediction could be verified experimentally in vitro if the firing rate saturation due to adaptation is felt before the effects of the absolute refractory period and is furthermore lesser than the firing rate saturation due to the absolute refractory period.

4.6.2 Origins of ISI correlations

Under the influence of perturbations and noise, we have shown that the model can give rise to negative ISI interval correlations. The value and rate of decay of these correlations depends on the parameter α and on the noise strength and can be varied to give rise to very different regimes. Expressions were derived for the ISI correlation coefficients. The model was also shown to give rise to ISI correlations similar to those seen in experimental data (Longtin and Racicot 1998; Chacron et al. 2000; see chapter 3). Further, the value of the correlation coefficient at lag one was shown to increase as a function of noise intensity.

ISI correlations have been observed in experimental data taken from neurons in different sensory systems (Schäfer et al. 1995; Bahar et al. 2001). The advantages of a negative correlation coefficient at lag one for stimulus detection through long-term spike train regularization as well as stimulus encoding have been outlined in another study (Chacron et al. 2001a; see chapter 6). Moreover, a similar firing mechanism has already been used to successfully model electroreceptors of weakly electric fish which display this negative serial correlation coefficient at lag one (Longtin and Racicot 1998; Chacron et al., 2000; 2001a;

2001b; see chapter 3) and could potentially be used to model other neurons which also display negative ISI correlations (Kuffler et al. 1957; Goldberg et al. 1964; Geisler and Goldberg 1966). In fact, models similar to our own have already been used to model several classes of neurons in the visual system (Keat et al. 2001), cortical neurons (Liu and Wang 2001) and electroreceptor afferents (Brandman and Nelson 2002).

Results similar to our own can be obtained by having adaptation in the reset value of the membrane potential. Geisler and Goldberg (1966) have shown that a adapting reset could give rise to ISI correlations. However, as mentioned earlier, there is experimental evidence for the firing threshold being dependent on the spiking history (Azouz and Grey 1999) while there is, to our present knowledge, no such evidence for the reset value. We thus believe a dynamic threshold to be more physiologically realistic than a dynamic reset.

4.6.3 The annulus map

The response of the model to perturbations and noise will carry over in the presence of sinusoidal forcing (Chacron et al. 2000). However, it is then impossible to obtain analytical explicit expressions for the ISI under voltage perturbations. Furthermore, under sinusoidal forcing, we have shown that our model could be described by a map defined on an annulus. While this may seem a simple extension of the many studies which have described the dynamics of sinusoidally forced standard LIFs by circle maps (Rescigno et al. 1970; Keener et al. 1981; Coombes and Bressloff 1999; Pakdaman 2001), it has major implications in terms of the dynamics of the system. Indeed, the circle map associated with the periodically forced standard LIF is one-to-one and orientation preserving, so that it produces only one of three following regimes, phase locked, quasi-periodic or strange non-chaotic behavior (Keener 1980). For fixed parameters, it can neither produce a mix of these nor chaos.

In contrast, the annulus map associated with the LIF model with threshold fatigue is not constrained by such limitations and may very well lead to more complex dynamics (Le Calvez 2000), in agreement with our numerical investigations. Also, introducing threshold fatigue as in our model yields an adapting sequence of ISI's in the presence of periodic forcing. In contrast, for a step increase in bias current, models in which periodic forcing is taken into account by

sinusoidally modulating the threshold or the voltage reset value see an abrupt change in ISI. It should be noted that the annulus map is dissipative in general. This system is thus different from the so-called "standard map" studied by various authors (see e.g. Meiss 1992). Our results suggest that neurons with strong adaptation are more prone to displaying complex dynamics than those that either do not adapt or adapt more slowly.

We have furthermore reduced the annulus map to a circle map that preserved most of the rich dynamics. That circle map was not necessarily orientation preserving and thus chaotic dynamics could occur, as we confirmed by numerical simulations. Such sensitivity to initial conditions may confer some advantages, as chaotic systems have the potential to transmit much more information about time-varying stimuli than non-chaotic ones, depending on the Lyapunov exponents of the system (Abarbanel 1996). Further studies are needed to investigate such possible functional roles of chaos in signal processing in nervous systems.

4.6.4 Physiological origins of a dynamic threshold

In this chapter, we have considered how the addition of threshold fatigue affected the spike train of a single LIF neuron. However, it is clear that adaptation and ISI correlations seen in neurons could also be due to network and/or synaptic dynamics. To expand on this point, we now discuss the possible biological mechanisms that the threshold fatigue could model. Cumulative inactivation of sodium channels (Mickus et al. 1999) could for example give rise to neural adaptation and ISI correlations. However, a fast spike-activated slowly inactivating negative current could also give rise to similar effects. A likely candidate would be members of the KV family of potassium currents (Wang et al. 1998). Another candidate could be a calcium-activated potassium current such as I_{ahp}. This current was shown to be present in cortical neurons (Madison and Nicoll 1984). Liu and Wang (2001) have shown that a model with an I_{ahp} current was qualitatively equivalent to our threshold fatigue model with $W=W_1$ (i.e. a linear threshold reset) and $\alpha=1$. It is further known that most ionic currents have voltage-dependent conductances. The parameter s_0 represents the amount of reset that is voltage independent. The parameter α in our model could then represent the degree of dependence

upon the past firing time in the ionic channel associated with the dynamic threshold. Increasing α would thus lead to a more history-dependent conductance.

Long term depression at a synapse could also give rise to relative refractoriness (Hausser and Roth 1997). In fact, a depressing synapse produces negative spike train correlations (Goldman et al. 2002). Our model thus bears some similarity with models of short-term plasticity (Fuhrman et al., 2002; Goldman et al., 2002). Furthermore, it is known that many synapses are voltage-dependent. The parameter α in our model could then represent the voltage dependence of synaptic depression. It should however be noted that the recovery time constant of the neurotransmitter at typical synapses is usually in the hundreds if not thousands of milliseconds range (von Gernsdoff et al. 1997). Although neural adaptation usually occurs on much shorter time scales, our model could in principle still be used to account for such phenomena if the threshold time constant is made sufficiently long. Moreover, it has been observed that neural adaptation was in many ways similar to recurrent inhibition via a cellular network (Ermentrout et al. 2001). Our single neuron model could therefore describe network effects.

It has thus been shown that threshold fatigue can be similar to recurrent inhibition, synaptic depression, or an intrinsic I_{ahp} current responsible for adaptation.

4.6.5 Conclusion/Outlook

We have shown that the LIFDT model could display adaptation and firing rate saturation. Experiments have been suggested to test whether this actually occurs in real neurons. It will be exciting to have experimental verification of this theoretical prediction. We have looked at the effects of noise on the ISI correlation coefficient ρ_1 and gotten a basic understanding of the mechanism by which the model displays negative ISI correlations and thus how negative ISI correlations arise in the real system.

Furthermore, we have shown that chaotic dynamics and phase-locking could occur under sinusoidal stimulation. A method for calculating Lyapunov exponents for the model was presented. It is not yet clear how noise will affect the phase-locking patterns and chaotic

dynamics seen deterministically in the model; further studies are needed to investigate this. It is also unknown at present whether receptor afferents display chaotic dynamics. However, many neurons *in vitro* can display chaotic dynamics under sinusoidal current injection. Our results suggest that any combination of intrinsic ionic conductances, recurrent inhibition, and synaptic depression could induce chaotic dynamics. It would be interesting to block some of these mechanisms and to look at their effects on chaotic dynamics. Furthermore, we have looked at the effects of noise on frequency tuning. A maximum in vector strength was found when the cycle period was on the order of the threshold time constant. Further studies are needed in order to fully understand this effect.

5 Application of Information theory to neural spike trains

In this chapter, we will be interested in the application of information theory introduced in chapter 2 to neural spike trains. Information theory (Shannon 1948; Cover and Thomas 1991) is an increasingly popular tool to analyze input-output properties of neural systems and characterize neural coding (Borst and Theunissen 1999). The mutual information rate (i.e. the mutual information per unit time), I(R,S), between the neuronal response R and a stimulus S is defined as the reduction in the entropy rate of the neural response R brought about by the stimulus S. Thus, according to Shannon (1948), we have I(R,S)=H(R)-H(R/S) where H(R) is the entropy rate of the response averaged over the stimulus ensemble and H(R/S) is the entropy rate of the response to the stimulus S. H(R/S) measures the trial-to-trial variability in the neural response that cannot be accounted for by the stimulus (Strong et al. 1998).

In an optimal situation, the trial-to-trial variability is minimized, thus minimizing H(R/S), and maximizing mutual information. On the other hand, H(R) measures the uncertainty in the spike train brought about by a stimulus ensemble: it is also the maximum value that the mutual information rate I(R,S) can attain. The mutual information rate is usually expressed in bits/s and expresses the capacity of a neuron to discriminate different stimuli that are part of the stimulus ensemble per unit time. Thus, if the mutual information rate is equal to 10 bits/s, this means that the neuron is able to discriminate between 2^{10} different stimuli from 1 second of neuronal spiking activity.

There are different techniques for estimating bounds to the mutual information rate of a neuron (see Borst and Theunissen 1999 for a review). Indirect methods of estimating information transfer such as linear reconstruction have been used to estimate I(R,S) (Gabbiani 1996; Gabbiani et al. 1996; Rieke et al. 1997; Gabbiani and Koch 1998). On the other hand, the direct method of estimating the mutual information rate makes no assumptions on the nature of the neural code (Strong et al. 1998). The chapter is organized as follows. We will first

review existing results on the application of the indirect method to neural spike trains. Then we review existing estimates of the direct method and present new measures that are appropriate in different situations. These estimates are then compared numerically.

5.1 The indirect method of estimating information transfer

We will review the indirect method of estimating information transfer as described in Gabbiani (1996) and Rieke et al. (1997). The interested reader should thus consult these references for more information.

We consider the following situation. We give a neuron a stimulus S and look at its response R. We wish to estimate the mutual information rate I(R,S) between the stimulus and the response. We start from the definition of the mutual information from chapter 2:

$$\begin{aligned} i(R,S) &= h(R) - h(R/S) \\ &= -\sum_{r} P(r)\log_2[P(r)] + \sum_{r,s} P(r,s)\log_2[P(r/s)] \\ &= \sum_{r,s} P(r,s)\log_2\left[\frac{P(r,s)}{P(r)P(s)}\right] \\ &= \sum_{r,s}\{-P(r,s)\log_2[P(s)] + P(r,s)\log_2[P(s/r)]\} \\ &= -\sum_{s} P(s)\log_2[P(s)] + \sum_{r,s} P(r,s)\log_2[P(s/r)] \\ &= h(S) - h(S/R) \end{aligned} \tag{5.1}$$

Thus, the mutual information between the neural response R and a stimulus S can be viewed as the reduction in the entropy of the response caused a stimulus S, or equivalently as the reduction in the entropy of the stimulus caused by the response. The indirect method uses the second point of view.

5.1.1 Assumptions

We will assume that the stimulus S is drawn from a Gaussian distribution with mean zero and variance σ^2 and is time invariant:

$$P(S) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{S^2}{2\sigma^2}\right] \tag{5.2}$$

We further assume that the response R is given by $R=gS+\eta$ (i.e. we assume that the output R is proportional to the input S with gain g and added noise η). We further assume that η is uncorrelated with S and is drawn from a Gaussian distribution:

$$P(\eta) = \frac{1}{\sqrt{2\pi\langle\eta^2\rangle}} \exp\left[-\frac{\eta^2}{2\langle\eta^2\rangle}\right] \tag{5.3}$$

where $\langle\eta^2\rangle$ denotes the variance of η. We can then calculate the probability distribution P(R/S) as:

$$\begin{aligned} P(R/S) &= P(\eta = R - gS) \\ &= \frac{1}{\sqrt{2\pi\langle\eta^2\rangle}} \exp\left[-\frac{(R-gS)^2}{2\langle\eta^2\rangle}\right] \end{aligned} \tag{5.4}$$

Furthermore, we can calculate the probability distribution P(R) as:

$$P(R) = \int dS\, P(R/S)\, P(S) = \frac{1}{\sqrt{2\pi < R^2 >}} \exp\left[-\frac{R^2}{2\langle R^2\rangle}\right] \tag{5.5}$$

with $\langle R^2\rangle = g^2\langle S^2\rangle + \langle\eta^2\rangle$. Using equation (5.1), we have:

$$i = \int_{r,s} dr\, ds\, P(r/s)P(s)\log_2\left[\frac{P(r/s)}{P(r)}\right]$$
$$= \frac{1}{2}\log_2\left[1+\frac{\langle s^2\rangle}{\langle \eta^2\rangle / g^2}\right] \tag{5.6}$$

where we have used equations (5.2), (5.4), and (5.5). Defining $\eta_{eff}=\eta/g$, we have:

$$i = \frac{1}{2}\log_2\left[1+SNR\right] \tag{5.7}$$

where we defined the signal-to-noise ratio (SNR) as $\langle S^2\rangle / \langle \eta^2_{eff}\rangle$.

5.1.2 Extension to time varying stimuli

We now consider the case where S is time varying. The extension of equation (5.7) does not pose any particular problems since all Fourier components of the signal S(t) are independent of one another, this comes from the assumption that S(t) has a Gaussian distribution and is stationary. Thus, we can simply add all contributions from each Fourier component (Rieke et al. 1997) to get:

$$i = \frac{T}{2}\int_{-\infty}^{+\infty} df\, \log_2\left[1+SNR(f)\right] \tag{5.8}$$

where T is the total duration of the stimulus and SNR(f) is given by:

$$SNR(f) = \frac{P_{ss}(f)}{N_{eff}(f)} \tag{5.9}$$

and where $P_{ss}(f)$ is the power spectrum of s(t) and $N_{eff}(f)$ is the power spectrum of $\eta_{eff}(t)$. Thus, the mutual information rate I is given by:

$$I = \lim_{T\to\infty} \frac{i}{T} = \frac{1}{2}\int_{-\infty}^{\infty} df \log_2 [1 + SNR(f)] \tag{5.10}$$

5.1.3 The signal-to-noise ratio

We now focus on estimating the signal-to-noise ratio SNR(f) for an arbitrary system. Suppose that we construct some arbitrary estimator that takes the response R(t) and returns some estimate of S(t) which we shall denote by $S_{est}(t)$. The power spectrum of the effective noise $N'_{eff}(f)$ of the error in this estimate will always be greater or equal to $N_{eff}(f)$. Thus, we have that:

$$I \geq \frac{1}{2}\int_{-\infty}^{+\infty} df \log_2 \left[1 + \frac{P_{ss}(f)}{N'_{eff}(f)}\right] \tag{5.11}$$

Let us defined the difference between the original stimulus S(t) and the estimated stimulus by:

$$\varepsilon(t) = S(t) - S_{est}(t) \tag{5.12}$$

We will now describe a procedure for estimating the stimulus that gave rise to a spike train.

5.1.4 Stimulus reconstruction

We now review the theory for stimulus estimation (Rieke et al. 1997). We assume that the stimulus S(t) can be written as:

$$\begin{aligned} S(t) = &\int d\tau K_1(\tau) r(t-\tau) + \\ &\frac{1}{2}\int d\tau d\tau' K_2(\tau,\tau') r(t-\tau) r(t-\tau') + \end{aligned} \tag{5.13}$$

equation (5.13) gives the first two terms in the Volterra (1930) series expansion of S(t). This series is analogous to a Taylor series but works in a functional space. For simplicity, we will only consider the first term in equation (5.13). Thus, we have:

$$s_{est}(t) = \int d\tau K(\tau) r(t-\tau) \tag{5.14}$$

we wish to choose the Kernel $K(\tau)$ so as to minimize the time integral of the mean square error. Thus we must solve:

$$\frac{\delta \varepsilon^2}{\delta K} = 0 \tag{5.15}$$

where δ denotes the functional derivative and with ε^2 given by:

$$\varepsilon^2 = \frac{1}{T}\int_0^T dt\left[s(t) - s_{est}(t)\right]^2 \tag{5.16}$$

The calculus of variations provides us with tools for solving equation (5.15). To solve it, note that:

$$\varepsilon^2 = \lim_{\Delta\tau\to 0} \frac{\Delta\tau}{T} \sum_{i=0}^{T/\Delta\tau} \left(\Delta\tau \sum_{k=-\infty}^{\infty} K_k r_{i-k} - s_i \right)^2 \tag{5.17}$$

where $K_k \equiv K(k\,\Delta\tau)$ and $r_i \equiv r(i\,\Delta\tau)$. For clarity, we drop the limit $\Delta\tau \to 0$ in the following derivation. ε^2 will be minimized by setting the derivative of equation (5.17) with respect to K_j to zero for all values of j:

$$0 = \frac{2\Delta\tau}{T} \sum_{i=0}^{T/\Delta\tau} \left(\Delta\tau \sum_{k=-\infty}^{\infty} K_k r_{i-k} - s_i \right) r_{i-j} \Delta\tau \tag{5.18}$$

Rearranging equation (5.18) gives:

$$\Delta\tau \sum_{k=-\infty}^{\infty} K_k \left(\frac{\Delta\tau}{T} \sum_{i=0}^{T/\Delta\tau} r_{i-j} r_{i-k} \right) = \frac{\Delta\tau}{T} \sum_{i=0}^{T/\Delta\tau} s_i r_{i-j} \tag{5.19}$$

Taking the limit $\Delta\tau \to 0$ in equation (5.19), we get:

$$\int_{-\infty}^{\infty} d\tau' K(\tau') \left(\frac{1}{T} \int_0^T dt\, r(t-\tau') r(t-\tau) \right) = \frac{1}{T} \int_0^T dt\, s(t) r(t-\tau) \tag{5.20}$$

Using the definitions for the auto and cross-correlation functions, we get:

$$\int_{-\infty}^{+\infty} d\tau' C_{rr}(\tau-\tau') K(\tau') = C_{rs}(-\tau) \tag{5.21}$$

where $C_{ss}(t)$ and $C_{rs}(t)$ are defined by:

$$\begin{aligned} C_{rr}(t) &= \frac{1}{T} \int_0^T d\tau\, r(\tau) r(t+\tau) \\ C_{rs}(t) &= \frac{1}{T} \int_0^T d\tau\, s(\tau) r(t+\tau) \end{aligned} \tag{5.22}$$

Taking the Fourier transform of equation (5.21) gives:

$$K(f) = \frac{P_{rs}(-f)}{P_{rr}(f)} \tag{5.23}$$

where $P_{rr}(f)$ and $P_{ss}(f)$ are the respective Fourier transforms of $C_{rr}(t)$ and $C_{rs}(t)$. Thus, the Fourier transform of the kernel $K(t)$ is given by the cross-spectrum $P_{rs}(-f)$ divided by the power spectrum of the spike train $P_{rr}(f)$.

The quality of the reconstruction can be evaluated by computing the coding fraction γ (Gabbiani, 1996) given by:

$$\gamma = 1 - \varepsilon / \sigma \tag{5.24}$$

where σ is the standard deviation of S. We have $\gamma=1$ when the spike train is perfectly linearly correlated with the stimulus and $\gamma=0$ when the spike train is linearly uncorrelated with the stimulus.

5.1.5 A lower bound for the mutual information rate

We can define an effective noise as:

$$s_{est}(f) = g(f)[s(f) + n_{eff}(f)] \tag{5.25}$$

where ñ(f) is the Fourier transform of n(t). Taking the modulus square of equation (5.25) and averaging over realizations gives:

$$\left|K(f)\right|^2 P_{rr}(f) = \left|g(f)\right|^2 [P_{ss}(f) + N'_{eff}(f)] \tag{5.26}$$

$N'_{eff}(f)$ can be estimated in the following way (Rieke et al. 1997). One subdivides both the estimated stimulus and the stimulus into segments and computes their power spectra over each segment. For each frequency f, one then plots the estimated stimulus power versus the stimulus power. The gain $|g(f)|^2$ is then the slope of the best fit line while the scatter around this line gives the effective noise power $N'_{eff}(f)$. A lower bound for the mutual information rate can then be computed as:

$$I_{LB} = \frac{1}{2}\int_{-\infty}^{+\infty} df \log_2\left[1 + \frac{P_{ss}(f)}{N'_{eff}(f)}\right] \tag{5.27}$$

Computing the coding fraction and mutual information rates from neural spike trains using the aforementioned method poses no particular difficulty. We will defer numerical examples to chapter 7 where we will compare the mutual information rates obtained with the direct and indirect methods. We now describe the direct method.

5.2 *The direct method for applying information theory to neural spike trains*

Unlike the indirect method, the direct method makes no assumptions on the nature of the neural code. Thus, there is no need to assume that the stimulus is drawn from a Gaussian distribution. We will first introduce the concept of a stimulus ensemble. We will review existing measures as well as introduce new ones.

5.2.1 The stimulus ensemble

The total stimulus ensemble is denoted by W and consists of all stimuli S that an animal can encounter in nature. This is illustrated in Figure 5-1. Each member S of the stimulus ensemble W has an associated probability of occurrence P(S). In some cases, these probabilities have been experimentally measured such as for light contrast in the fly visual system (Laughlin 1981). However, they are presently unknown for weakly electric fish. The stimulus ensemble W is in general multi-dimensional and is hard to fully describe. For these reasons, researchers often concentrate on subsets that are easier to characterize.

5.2.1.1 Gaussian stimulus ensemble

One such subset is the ensemble of filtered Gaussian white noise stimuli that has been widely used in characterizing the information transfer properties of neurons (Gabbiani et al. 1996; Roddey and Jacobs 1996; Wessel et al. 1996; Clague et al. 1997; Warland et al. 1997). We denote this ensemble by W_{gaus}. Each member of W_{gaus} is characterized by two parameters: its contrast σ and its cutoff frequency f_c. The contrast is defined as the standard deviation of the stimulus (note that since we use a baseline EOD amplitude of 1mV, the contrastσ has the same numerical value as the standard deviation-to-mean ratio). All Gaussian stimuli used have a power-spectrum σ^2/f_c for frequencies between zero and the f_c and zero otherwise (Wessel et al. 1996). Since every member of W_{gaus} is uniquely characterized by the parameters σ and f_c, the

set W_{gaus} is two-dimensional. It is thus easy to fully characterize, but natural stimuli seldom display Gaussian statistics (Simoncelli and Olshaussen, 2001).

5.2.1.2 The sinusoidal amplitude modulation ensemble

Keeping in mind that we will eventually apply these measures to models of electroreceptor neurons in weakly electric fish, we now describe a natural stimulus ensemble that is particular to these animals.

When two weakly electric fish with different EOD frequencies are in proximity of one another, there is interference between the quasi-sinusoidal electric fields generated by each fish, producing a beating phenomenon. This interference will cause a sinusoidal amplitude modulation (SAM) in each fish's own electric field (Bastian 1981; Heiligenberg et al. 1991). The frequency of this modulation depends on the frequency difference between the two fish and can range from a few Hz to 400 Hz. We thus used SAM's within the physiological range as natural stimuli; other natural electrocommunication signals such as chirps (Zupanc and Maler 1993) are not considered here. We denote by W_{SAM} the set of all such stimuli. Each member of this set can be described by two parameters: the SAM contrast σ_{SAM} and frequency f_{SAM} of the sinusoid. However, to fully characterize the set W_{SAM}, we would need to know the probability of natural occurrence of a SAM with contrast σ_{SAM} and frequency f_{SAM} in nature for the fish (this would be an experiment similar to the one performed by Laughlin (1981) for the fly). Although such data can be acquired, this has to our knowledge not yet been accomplished. Note that W_{SAM} is but a small subset of W as the latter contains other natural stimuli such as the chirps mentioned earlier as well as stimuli caused by prey (Nelson and MacIver 1999). This example illustrates that the contents of the stimulus ensemble are very dependent on the particular system one wishes to study. For example, the stimulus ensemble for weakly electric fish is very different from the visual stimulus ensemble of primate vision.

5.2.1.3 The "no-stimulus" ensemble

Finally, we discuss the notion of a "no-stimulus" ensemble. We denote this set by W_0 because this set contains only one member: a stimulus whose value is identically zero (i.e. the EOD is presented alone without amplitude modulations). It is easily seen that this set contains all SAM and Gaussian stimuli with zero contrast. As such, W_0 intersects both W_{gaus} and W_{SAM} (figure 1). We will see that the no-stimulus ensemble proposed here is useful when dealing with spontaneously active neurons such as electroreceptors.

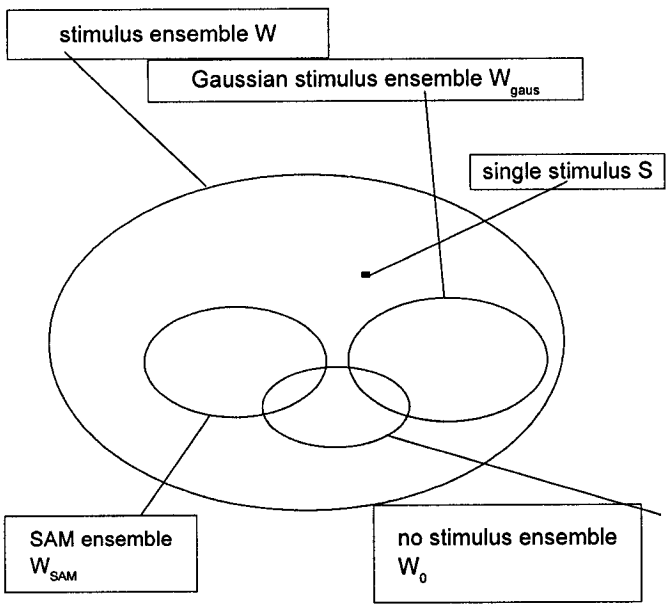


Figure 5-1: The stimulus ensemble W. Shown are the subset W_{gaus} of zero-mean Gaussian stimuli of contrast σ and cutoff frequency f_c used in this study, the subset W_{SAM} of all sinusoidal stimuli with contrast σ_{SAM} and f_{SAM}. The "no-stimulus" ensemble W_0 is also shown; it intersects W_{SAM} and W_{gaus} since the cases $\sigma=\sigma_{SAM}=0$ correspond to no stimulus.

5.2.2 Estimating the mutual information rate

The calculation of the mutual information rate I(R,S) requires the calculation of both the response entropy rate averaged over the stimulus ensemble H(R) and the response entropy

rate given the stimulus H(R/S). The estimation of H(R/S) is standard (Strong et al. 1998) and poses no conceptual difficulty. We will describe it first.

5.2.2.1 Estimation of H(R/S)

We use the method described by Strong et al. (1998). We assume that the same stimulus S was repeated many times under identical conditions and that each trial resulted in a particular spike train representing the neural response R to a particular trial. As earlier mentioned, the response entropy rate given the stimulus, H(R/S), measures the response trial-to-trial variability to a stimulus.

To estimate H(R/S), we partition the spike train into bins of length $\Delta\tau$, with $\Delta\tau$ smaller than the minimum ISI so that each bin contains at most one spike. A bin is then assigned the value 0 if no spikes occurred during that bin and 1 otherwise. This procedure converts a spike train into a binary string and is illustrated in figure 2a. The alphabet for this code thus consists of two symbols (no spike: 0 and 1: spike) and words consist of binary strings of various lengths (e.g. '010','00101').

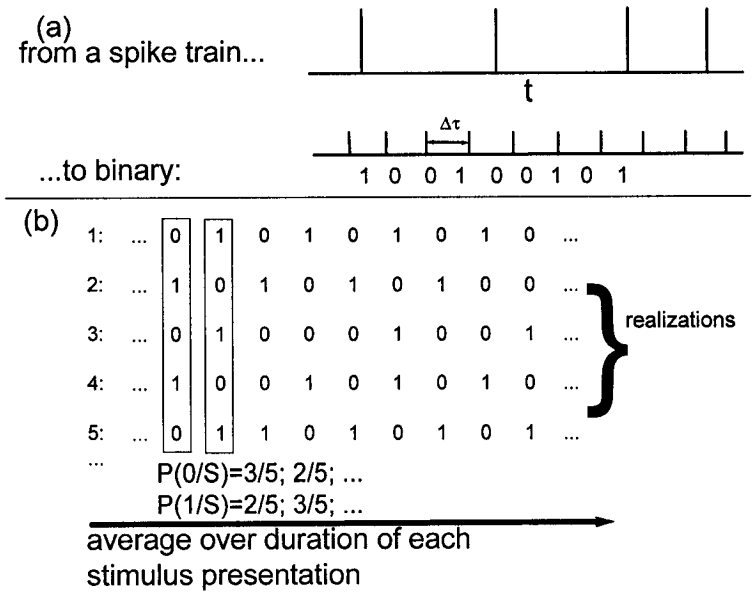


Figure 5-2: Calculating the entropy rate of a spike train. (a) Outline of the procedure for turning a spike train into a binary string consisting of 0's and 1's. The spikes are represented by vertical bars while time is represented horizontally. The time axis is then divided into bins of length $\Delta\tau$ and each bin is assigned the value 1 if a spike occurred and zero otherwise. The binary string obtained will of course depend on the bin width $\Delta\tau$. (b) Outline of the procedure used to calculate the probabilities of occurrence of words w when the length of these words is one (i.e. 0 and 1). Each binary sequence represents one realization. Multiple stimulus presentations will give rise to multiple spike trains that can be converted into binary data using the procedure outlined above. The probabilities of occurrence of words of length 1 are calculated for each bin from all trials. The final probabilities of occurrence P(w/S) are calculated by averaging the results obtained for each bin. This procedure is then repeated for words of length two, etc... Note that the windows used for calculating the probabilities are non-overlapping.

The choice of $\Delta\tau$ is critical and depends on various considerations such as the spike timing jitter of the particular neuron (i.e. the degree to which the precise timing of action potentials is important for information transfer). This spike timing jitter can usually be measured by adding "jitter" to the spike times and looking at the resulting loss in information (Kreiman et al. 2000). Another important factor is the temporal discrimination of postsynaptic neurons (Bender et al. 2001): this usually depends on the properties of the synapse(s) and possible inter-neuron(s) connecting the neurons in question as well as on the temporal integration properties of the

post-synaptic neurons themselves. In our case, the EOD cycle is a natural time scale since P-receptor afferents can fire at most one spike per EOD cycle (Bastian 1981). It is also known that electroreceptor input onto basilar pyramidal cells results in excitatory postsynaptic potentials of duration barely greater than an EOD cycle (Berman and Maler 1998). We thus take $\Delta\tau$ equal to one EOD cycle. Since we assume an EOD frequency equal to 1000 Hz, one EOD cycle corresponds to 1 millisecond.

We mark a particular time t with respect to the stimulus and compute the probabilities of occurrence across trials of each word w of length L that begin at t. These probabilities are denoted by P(w/S,t). The entropy of words of length L is then calculated as (Strong et al. 1998):

$$h(L/S,t) = \sum_{w \in \Gamma(L)} -P(w|S,t)\ \log_2 P(w|S,t) \tag{5.28}$$

where $\Gamma(L)$ is the set of words of length L (e.g. $\Gamma(2)=\{01,10,00,11\}$). The entropy h(L/S,t) thus denotes the local entropy of words of length L in the time window (t,t+L). These entropies h(L/S,t) are then averaged across all such windows during the entire time course of the stimulus. We thus have:

$$h(L/S) = \langle h(L/S,t) \rangle_t \tag{5.29}$$

where the angle brackets denote the average over the stimulus time course. This procedure is illustrated graphically in figure 2b. The entropy rate of the spike train given the stimulus is then obtained by (Strong et al. 1998):

$$H(R/S) = \lim_{L\to\infty} \frac{h(L/S)}{L\Delta\tau} \tag{5.30}$$

In practice, the entropies h(L/S) calculated from equation (5.29) underestimate the real value for large word length L because of undersampling due to the finiteness of the spike train data.

We hence used an extrapolation procedure inspired from Strong et al. (1998) to get an estimate of the entropy rate (Chacron et al. 2001a):

$$\frac{H(L/S)}{L\Delta\tau} = H(R/S) + \frac{C_1}{L} + \frac{C_2}{L^2} + \ldots \qquad (5.31)$$

where C_1, C_2 are constants. Such expansions are used to correct for finite size effects. To estimate the entropy rate H(R/S), we plot the quantities $H(L)/(L\,\Delta\tau)$ as a function of 1/L and perform a quadratic least-squares fit. The entropy rate H(R/S) is then estimated from the value of the fitted parabola at 1/L=0. An example is given in Figure 5-3. The accuracy was evaluated by computing the quantities $H(L)/(L\,\Delta\tau)$ for a binomial distribution having the same statistics for L=1. We have $H(L)/(L\,\Delta\tau) = H(1)/\Delta\tau$ for all L for such a process. However, Figure 5-3 shows that undersampling errors are significant for $1/L < 0.2$.

We estimated the entropy rate H(R/S) from 1000 trials each lasting 10 seconds. Note that such long simulation times are necessary here due to the high firing rates of the electroreceptors we study. These usually range from 100 to 600 Hz (Ratnam and Nelson 2000). This as well as other considerations (see above) forces us to take a small bin size $\Delta\tau$. However, for other systems, a larger $\Delta\tau$ can sometimes be taken if the firing rate is lower, and as few as 100 repetitions can sometimes lead to a good estimate of the response entropy rate given the stimulus (Strong et al. 1998; Reinagel and Reid 2000).

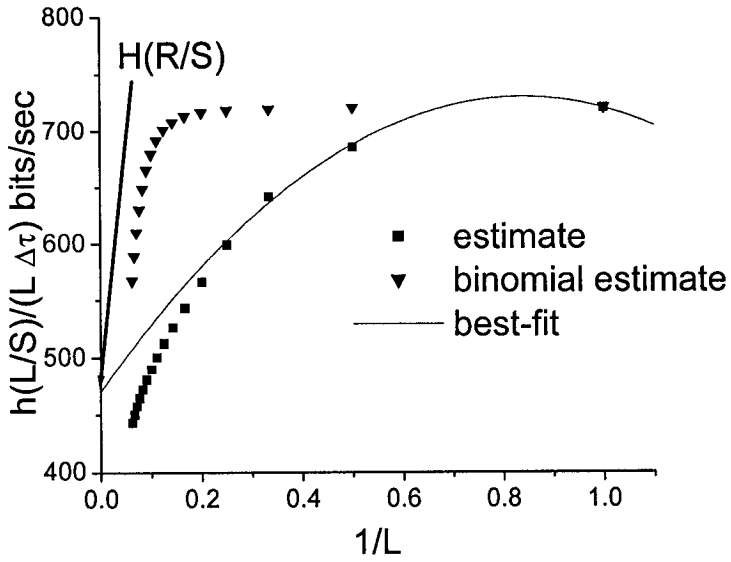


Figure 5-3: The entropy rates h(L/S)/(L Δτ) as a function of 1/L. The best fit parabola is then obtained and the entropy rate H(R/S) is obtained from value of the best fit parabola at the origin. Also shown are the values obtained from a binomial distribution showing that only the values obtained for small L are accurate. Parameter values for the LIFDT model are the same as in Figure 3-4.

5.2.2.2 Estimation of H(R)

Having discussed our estimation procedure for the response entropy rate H(R/S), we can now discuss estimation procedures for the entropy rate of the response averaged over the stimulus ensemble. According to probability theory, we have that:

$$P(w,t) = \sum_{S \in W(S)} P(w \mid S,t) \;\; P(S) \tag{5.32}$$

where P(S) is the probability of occurrence of stimulus S within the stimulus ensemble W and P(w/S,t) is, as before, the probability of w occurring at time t with respect to the stimulus present at time t. The entropy rate H(R) could then in theory be estimated by using equations (5.28) through (5.31) with P(w,t) instead of P(w/S,t) in equation (5.28) and H(R) instead of

H(R/S) in equations (5.29) through (5.31). However, as mentioned above, the stimulus ensemble W cannot be characterized fully in general and the probabilities of occurrence of various stimuli P(S) are not known from nature. It thus becomes necessary to make simplifying assumptions at this point and we now discuss some of them.

5.2.2.2.1 H(R) averaged over the Gaussian stimulus ensemble

It is possible to restrict oneself to an easily characterizable subset of the stimulus ensemble W where the calculation described in the previous paragraph can be undertaken. Such a subset is the Gaussian stimulus ensemble W_{gaus}. We shall denote the estimate obtained by $H_{gaus}(R)$ to emphasize its dependence on the subset W_{gaus}.

As such, we estimated the entropy rate $H_{gaus}(R)$ in the following way. Gaussian stimuli of contrasts σ and cutoff frequencies f_c each lasting 10 seconds were each given 1000 times. The probabilities P(w/S,t) were estimated for each Gaussian stimulus S as for the calculation of the response entropy rate given the stimulus S. These probabilities were then averaged using equation (5.32) with W_{gaus} instead of W to obtain the probabilities P(w,t). It was assumed that all stimuli were equiprobable. Equations (5.28) through (5.31) with P(w,t) instead of P(w/S,t) in equation (5.28) and $H_G(R)$ instead of H(R/S) in equations (5.29) through (5.31) were then used to compute $H_{gaus}(R)$. The mutual information rate can thus be estimated as $I_{gaus}(R,S) = I_{gaus} = H_{gaus}(R) - H(R/S)$.

5.2.2.2.2 Calculation from an unrepeated stimulus

Even averaging over the Gaussian stimulus ensemble can require much more data than is available. Thus, Strong et al. (1998) have proposed that the response entropy rate H(R) could be estimated from the response entropy rate of the spike train in the presence of an unrepeated stimulus S. This is based on the assumption that a single member S of the stimulus ensemble W would provide a good sample of the full stimulus ensemble W. The spike train entropy rate estimated in this way will of course depend on the particular stimulus S used; to

emphasize this dependence we shall denote it by $H_{stim}(R)$. The corresponding mutual information rate is thus $I_{stim}(R,S) \equiv I_{stim} = H_{stim}(R) - H(R/S)$.

We presented a long (10000 seconds duration) stimulus and obtained a spike train that was converted into a binary sequence as described above. We estimated the probabilities of occurrence $P_{stim}(w)$ of words w within the sequence in the following way. The binary sequence was divided into non-overlapping windows of length L. The probabilities of occurrence of words of length L were then estimated across these time windows. This allowed us to compute the entropies $h_{stim}(L)$ as:

$$h_{stim}(L) = -\sum_{w \in \Gamma(L)} P_{stim}(w) \log_2 P_{stim}(w) \qquad (5.33)$$

The entropy rate $H_{stim}(R)$ was then estimated from:

$$\frac{h_{stim}(L)}{L\Delta\tau} = H_{stim}(R) + \frac{C_1}{L} + \frac{C_2}{L^2} + \ldots \qquad (5.34)$$

The mutual information rate can then be estimated as $I_{stim}(R,S) = H_{stim}(R) - H(R/S)$. However, this estimate assumes that the stimulus used to compute $H_{stim}(R)$ provides a good sample of the stimulus ensemble W.

5.2.2.2.3 Estimation from the no-stimulus ensemble

We now present an information estimate that does not depend on the stimulus ensemble per se. It is known that many neurons are spontaneously active and that, in most cases, this spontaneous activity *in vivo* is irregular. We propose an information measure based on the entropy rate of the spontaneous activity of such neurons $H_{spon}(R)$. To estimate it, we take a long record (10000 seconds in our case) of spontaneous activity which is then converted into a binary sequence. We then used the same estimation procedure as for $H_{stim}(R)$. The difference between the spontaneous response entropy rate $H_{spon}(R)$ and the response entropy rate given the stimulus H(R/S) is a measure of information. We shall denote this quantity by

$I_{spon}(R) \equiv I_{spon} = H_{spon}(R) - H(R/S)$. However, this quantity is not a measure of information in the "classical" sense since it does not take into account the stimulus ensemble W. Instead, this quantity measures the reduction of the spontaneous response entropy brought about by a stimulus S. This is based on the assumption that spontaneously active neurons transmit information about stimuli through a change from their spontaneous activity (Ratnam and Nelson 2000; Chacron et al. 2001a). We now discuss the relation between I_{spon} and the Shannon mutual information rate I. We have:

$$\begin{aligned} I(R,S) &= H(R) - H(R/S) \\ &= H(R) - H_{spon}(R) + H_{spon}(R) - H(R/S) \\ &= H(R) - H_{spon}(R) + I_{spon} \\ &= C + I_{spon} \end{aligned} \tag{5.35}$$

Thus, the mutual information rate calculated from the spontaneous activity of the neuron I_{spon} differs by only a constant C from Shannon's mutual information rate I(R,S). This constant C depends on the stimulus ensemble W. Let us assume that no stimulus is given to the neuron, and we then have I(R,S) =C. C thus represents the information rate that a spontaneously active neuron transmits about the fact that no stimulus is present in the animal's environment. It depends only on the stimulus ensemble and on the neuron's intrinsic properties. It is thus expected that I_{spon} and I(R,S) will behave in the same manner when the stimulus S is varied.

5.2.2.2.4 Estimation in the presence of background noise

Note that all the previous measures are based on the assumption of a noise-free environment in which stimulus encoding and decoding is performed. We consider the effects of background noise on information transfer. Common sources of background noise in our system include environmental factors and conspecifics. For example, lightning is common in these fish's natural habitat and generates electrical pulses and oscillations that propagate over hundreds of kilometers through tropical waters (Hopkins 1973). The lightning pulses interfere with weakly electric fish's abilities to detect amplitude modulations (i.e. stimuli). Moreover, the fish must detect prey (invertebrates) that may be hiding in root masses (Crampton 1998). The electric

image caused by these root masses can be considered noise when the fish is trying to detect its prey. Finally, two fish might be able to detect each other even when other fish are close by (Parthridge and Heiligenberg 1980): this is the weakly electric fish equivalent of the cocktail-party effect where one entertains a conversation amidst many other (Cherry, 1953). From the above examples, it is clear that background noise will be very system-dependent. In our study, we assume background noise to be Gaussian with contrast σ=0.01 mV and cutoff frequency f_c=50 Hz. These values where chosen arbitrarily for lack of experimental data on the matter. These parameters are fixed throughout. Furthermore, we will assume that the background noise is additive (i.e. the noise is added to the stimulus).

Often, sensory neurons must thus encode and decode stimuli that are perturbed by variable amounts of noise such as in the cocktail party effect discussed above. We now extend the measure I_{spon} to include the effects of background noise. To do so, we must consider the effects of background noise on both the spontaneous response entropy rate $H_{spon}(R)$ and on the response entropy rate given the stimulus H(R/S).

We computed the response entropy rate $H_{noise}(R)$ in the same manner as H_{stim} (i.e. using one long realization from the ensemble) except that background noise was given to the neuron instead of a stimulus. We computed the response entropy rate given the stimulus in the presence of background noise in the same manner as the response entropy rate given the stimulus H(R/S). The only difference is that background noise is added to the stimulus before each trial. This background noise changes across trials. We will denote the estimate obtained by $H_{noise}(R/S)$. The mutual information rate can then be estimated as $I_{noise}(R,S)=I_{noise}=H_{noise}(R)-H_{noise}(R/S)$.

5.2.3 Summary

In summary, we have presented different ways to estimate the entropy rate H(R) and to compute the corresponding mutual information rate. We have restricted the calculation of H(R) over the Gaussian stimulus ensemble to calculate the information rate I_{gaus}. We have then reviewed a measure proposed by Strong et al. (1998) denoted by I_{stim} that estimates H(R) from a long stimulus. We have also calculated the information rate from the spontaneous activity of

a neuron I_{spon} and showed that it differed by only a constant from the "true" mutual information rate I. Finally, we have extended I_{spon} to the case where background noise is present in the environment and obtained I_{noise}. Table 5.1 summarizes the different information rate estimates we have presented, their assumptions, and a brief description on their estimation procedure. We will now compare results obtained using these different estimates.

5.2.4 Numerical results

We first present results calculations of the mutual information rate measures I_{stim}, I_{gaus}, I_{spon}, and I_{noise} using Gaussian stimuli (i.e. the stimuli are part of W_{gaus}) of various physiologically relevant contrasts σ and cutoff frequencies f_c. The estimates I_{stim}, I_{spon}, and I_{noise} are then compared for naturalistic SAM's of various frequencies. Throughout, parameter values for the LIFDT model are the same as in Figure 3-4.

Information measure:	Description:	Stimulus ensemble used:	Assumptions:
I_{gaus}	Mutual information rate calculated over the Gaussian stimulus ensemble W_{gaus}.	The Gaussian stimulus ensemble W_{gaus}.	The Gaussian stimulus ensemble provides a good sample of the total stimulus ensemble W.
I_{stim}	Mutual information rate calculated from a single member S of the stimulus ensemble.	W	The stimulus ensemble is well sampled by the single member S.
I_{spon}	Mutual information rate calculated from the spontaneous activity of a neuron	The "no-stimulus" ensemble W_0.	The neuron must have sufficiently irregular spontaneously activity.
I_{noise}	Mutual information rate calculated in the presence of background noise	The "no-stimulus" ensemble in the presence of background noise.	The background noise must elicit sufficiently irregular activity from the neuron.
I	In theory: Mutual information rate calculated from the stimulus ensemble W	The full stimulus ensemble W.	In theory: none except that the stimulus ensemble W is properly characterized.

Table 5.1: Summary of the different information measures.

5.2.4.1 *Results obtained with Gaussian stimuli*

We wish to study the effects of varying the stimulus contrast σ. We thus first keep the cutoff frequency constant at100 Hz and compare the different mutual information rate estimates obtained by varying the stimulus contrast σ. The results are shown in Figure 5-4a. We see that the estimates I_{gaus} (the mutual information rate estimated over the Gaussian stimulus ensemble W_{gaus}) and I_{stim} (the mutual information rate estimated from a single stimulus of W_{gaus}) are in good agreement for high contrast σ. However, I_{stim} is less than I_{gaus} as σ tends to zero. To understand this, let us assume that the stimulus has $\sigma=0$. We then have that the entropy rate H(R/S) is equal to the response entropy rate $H_{stim}(R)$, which in turn is equal the spontaneous response entropy rate $H_{spon}(R)$. As a consequence, the estimate $I_{stim}=H_{stim}(R)-H(R/S)$ is equal to zero. However, the response entropy rate averaged over the Gaussian stimulus ensemble $H_{gaus}(R)$ is not equal to the response entropy rate H(R/S) when the stimulus S has zero contrast. Hence, we have $I_{gaus}=H_{gaus}(R)-H_{spon}(R)>0$ when $\sigma=0$.

Our results thus show that approximating the Gaussian stimulus ensemble-averaged response entropy rate $H_{gaus}(R)$ by the response entropy rate obtained from a single member of the stimulus ensemble $H_{stim}(R)$ is justified for mid to high contrasts σ at this particular cutoff frequency f_c. This is due to the particular nature of the stimulus ensemble W_{gaus}. Note that a stimulus S with contrast σ and cutoff frequency f_c contains all Gaussian stimuli with lower contrasts and cutoff frequencies. Thus, a member with a high contrast and cutoff frequency provides a good sample of the stimulus ensemble W_{gaus} and thus provides a good estimate of the spike train entropy rate calculated over this subset W_{gaus} of the stimulus ensemble W as we will show below in Figure 5-4b.

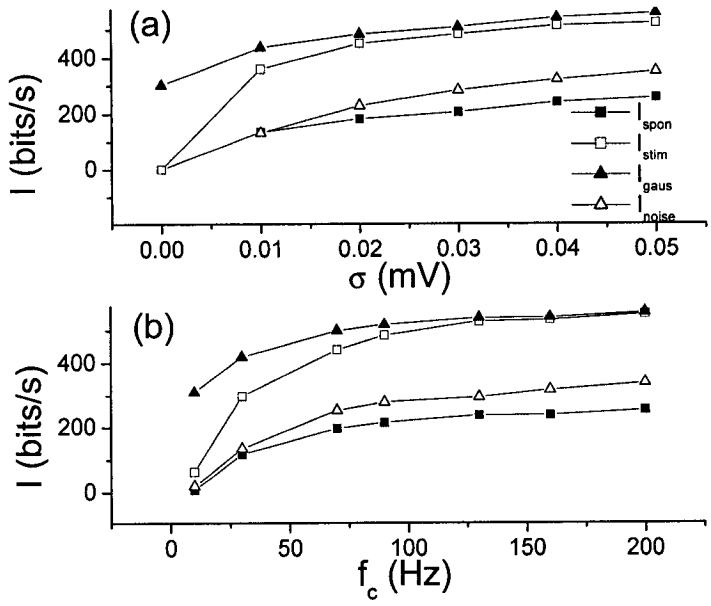


Figure 5-4: Effects of varying the RAM contrast and cutoff frequency. (a) Information measures as a function of stimulus contrast for f_c=100 Hz. Note that both I_{gaus} and I_{stim} are much higher than I_{spon} and that $I_{gaus} \approx I_{stim}$ for high contrasts. (b) Information rate measures as a function of cutoff frequency for σ=0.03 mV. Note again that $I_{gaus} \approx I_{stim}$ for high cutoff frequencies. Note that we have that $\sigma^2 = 2 f_c \varepsilon$ where ε is the constant value of the power spectrum, we can thus vary σ and f_c independently.

It is seen that the mutual information rate estimated over the Gaussian stimulus ensemble I_{gaus} and the mutual information rate estimated over the "no-stimulus" ensemble I_{spon} differ by a constant. This is expected from the previous discussion since $I_{gaus} = I_{spon} + C'$ with $C' = H_{gaus}(R) - H_{spon}(R)$. The constant C' measures how well the neuron can discriminate the absence of a stimulus from other members of the stimulus ensemble W_{gaus}. This extra information is absent from the estimate I_{stim} when the contrast σ is zero because the entropy rate $H_{stim}(R)$ is estimated from the same stimulus ensemble member S (the "no-stimulus" case) that is used to estimate the noise entropy rate H(R/S).

Note that the mutual information rate estimated from the spontaneous activity I_{spon} is equal to 0 when the stimulus contrast σ=0 (i.e. no stimulus is present). Hence this measure implies that no information is transmitted by the neuron when no stimulus is present, whereas the measure

I_{gaus} implies that information is being transmitted when no stimulus is present (i.e. about the absence of stimulus). We will return to this point in the discussion.

Finally, we note that the mutual information rate calculated with background noise I_{noise} is slightly greater than I_{spon} for high contrasts but is still lower than I_{stim}. To understand this, note that the spike train entropy rate in the presence of background noise $H_{noise}(R)$ is always greater than the spike train entropy rate in the absence of background noise $H_{spon}(R)$ (i.e. with spontaneous activity only) because the extra randomness due to background noise increases the entropy rate. For the same reason, the response entropy rate given the stimulus in the presence of background noise $H_{noise}(R/S)$ is always greater than the noise entropy rate in the absence of background noise H(R/S). The signal-to-noise ratio (i.e. the ratio of the stimulus contrast to the noise contrast) increases with stimulus contrast and consequently the noise becomes less and less important. The noise entropy rate in the presence of background noise $H_{noise}(R/S)$ tends towards the noise entropy rate obtained when no background noise is present H(R/S), hence making the mutual information rate in the presence of background noise I_{noise} higher than I_{spon}.

Note also that I_{noise} is lesser or equal to the mutual information calculated from a single member of the stimulus ensemble I_{stim}. This is because the noise entropy rate in the presence of background noise $H_{noise}(R/S)$ is higher than the noise entropy rate obtained in the absence of background noise H(R/S). Moreover, the spike train entropy rate obtained from a single member of the stimulus ensemble $H_{stim}(R)$ is higher than the spike train entropy rate $H_{noise}(R)$ if the stimulus S has higher contrast σ or cutoff frequency f_c than the background noise N (note that the two are equal if the stimulus and background noise have the same contrast σ and cutoff frequency f_c).

We now keep the contrast σ constant at 0.03 mV and vary the cutoff frequency f_c of the stimulus. Results are shown in Figure 5-4b. We note that all four measures increase with f_c. This increase is due to the high-pass filtering characteristics of the electroreceptors we study (Xu et al. 1996) and which are incorporated in the model we use (Chacron et al. 2001a; see chapter 3). The mutual information rate calculated from the Gaussian stimulus ensemble I_{gaus} and the mutual information rate calculated from a single member of the stimulus ensemble I_{stim}

agree for high cutoff frequencies f_c. This is again due to a particular property of the Gaussian stimulus subset W_{gaus}. Because the power spectrum of a stimulus with cutoff frequency f_c contains all frequencies up to f_c, the stimulus essentially contains all stimuli with the same contrast σ with lower cutoff frequencies f_c. For this reason, a stimulus with a high cutoff frequency and a reasonably high contrast σ provides a good sample of the entire stimulus subset W_{gaus}. However, our results show that a single Gaussian stimulus with low frequency content will not provide a good sample of the stimulus ensemble subset W_{gaus}; consequently, the mutual information rate will be underestimated if one uses I_{stim}. Note that the mutual information rate calculated over the Gaussian stimulus ensemble I_{gaus} again tends towards the constant $H_{gaus}(R)$-$H_{spon}(R)$ as f_c goes to zero.

The mutual information rate calculated from the spontaneous activity I_{spon} is again lower than the mutual information rate calculated over the Gaussian stimulus ensemble I_{gaus} for the same reason as above: the measure I_{spon} only gives us the reduction in the spike train entropy rate based on the "no-stimulus" ensemble (i.e. spontaneous activity) rather than the reduction in entropy rate from the Gaussian stimulus ensemble W_{gaus}. Further, the mutual information rate obtained in the presence of background noise I_{noise} is lower than the mutual information rate I_{gaus} calculated over the stimulus ensemble subset W_{gaus}. This is again for the same reasons as above. However, the mutual information rate in the presence of background noise I_{noise} is again higher than the mutual information rate in the absence of background noise I_{spon}, since the spike train entropy rate in the presence of background noise $H_{noise}(R)$ is always greater than or equal to the spike train entropy rate in the absence of background noise $H_{spon}(R)$; equality occurs only in the limit where the background noise goes to zero. Due to the aforementioned high-pass filtering characteristics of the neurons we study, increasing the cutoff frequency of the stimulus f_c increases the signal-to-noise ratio, and thus the noise entropy rate in the presence of background noise $H_{noise}(R/S)$ then tends towards the noise entropy rate in the absence of background noise H(R/S) because the noise becomes less and less important.

5.2.4.2 Results obtained using naturalistic stimuli

The mutual information rates calculated from the spontaneous activity of the spike train in the absence (I_{spon}) and presence (I_{noise}) of background noise easily extend to naturalistic stimuli as

they do not require a priori knowledge of the neuron's full stimulus ensemble W. Figure 5-5a shows the measures I_{spon} and I_{noise} obtained by varying the SAM contrast σ_{SAM} and for constant SAM frequency f_{SAM}=100 Hz. We see that in the absence of background noise, the mutual information rate I_{spon} increases with SAM contrast. This is similar to what was obtained for Gaussian stimuli. The mutual information rate I_{noise} also increases with stimulus contrast. However, we have I_{noise} almost equal to I_{spon} for low contrasts and I_{noise} greater than I_{spon} for higher contrasts σ_{SAM}. This occurs for the same reason as described above with Gaussian stimuli: the SNR increases with contrast σ_{SAM} and the noise entropy rate in the presence of background noise $H_{noise}(R/S)$ decreases towards the noise entropy rate obtained in the absence of background noise H(R/S).

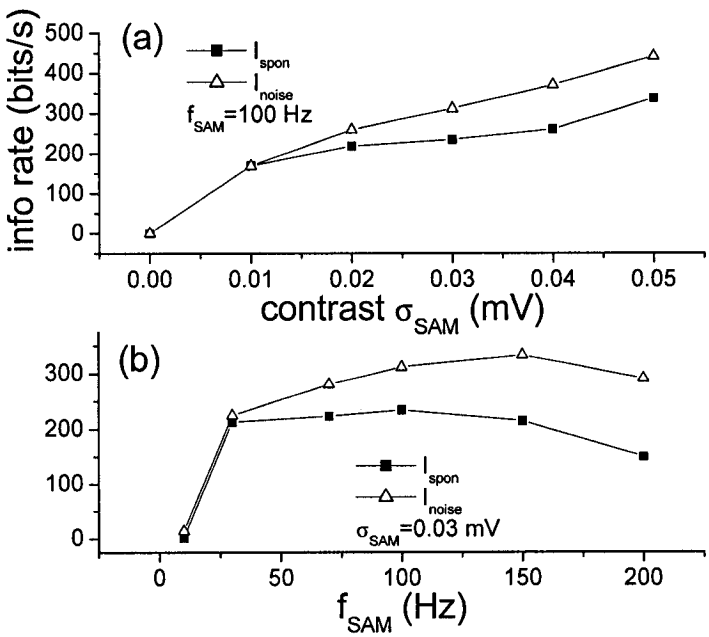


Figure 5-5: Effects of varying the SAM contrast and cutoff frequency. (a) Information measures I_{spon} and I_{noise} as a function of SAM contrast for a constant SAM frequency of 100 Hz. I_{noise} is greater than I_{spon} for high contrasts (see text for explanation). (b) Information measures I_{spon} and I_{noise} as a function of SAM frequency for a constant SAM contrast of 0.03 mV. Note that both information rates exhibit a maximum as a function of SAM frequency.

Figure 5-5b shows the measures I_{spon} and I_{noise} obtained by varying the SAM frequency f_{SAM} and keeping the SAM contrast σ_{SAM} constant. We note that the mutual information rate in the presence of background noise I_{noise} and in the absence of background noise I_{spon} are almost equal for low SAM frequencies f_{SAM} but that the difference $(I_{noise}-I_{spon})$ increases with f_{SAM}. This is due to the high-pass filtering properties of the electroreceptors reproduced by our model which amplifies the SAM's at higher frequencyes and thus produces a relatively stronger stimulus S. Moreover, both measures exhibit a maximum as a function of f_{SAM}. This can be explained as follows: neurons typically exhibit low-pass filtering characteristics due to their membrane time constant. However, electroreceptor neurons also exhibit high-pass filtering characteristics. Putting these two effects together, the neurons might exhibit band-pass filtering characteristics. A resonance in the mutual information rate as a function of stimulus frequency f_{SAM} can then be expected (Hutcheon and Yarom 2000).

Note however that this resonance was not seen for Gaussian stimuli. This is partly because one merely adds higher frequency components to a Gaussian signal by increasing its cutoff frequency f_c. On the other hand, increasing the SAM frequency f_{SAM} shifts the frequency of the unique Fourier component of the signal, thereby probing the neural response to different frequencies. Figure 5-5b can thus be thought of as an "information tuning curve" for the neuron. Note that this is different than the looking at the variation in information transfer of Gaussian stimuli by varying the EOD (i.e. carrier) frequency itself (Longtin and St-Hilaire 2000). This provides another example where the neuronal response to naturalistic stimuli can be very different from the response to low-passed filtered Gaussian white noise stimuli (Rieke et al. 1995; Lewen et al. 2001).

5.3 *Discussion*

5.3.1 The indirect method of estimating information transfer

We have reviewed how the indirect method of estimation is derived. It was assumed that the stimulus had a Gaussian probability distribution. This method furthermore only gives a lower bound on information transfer. However, it has the advantage of not requiring a lot of data.

We will compare the direct and indirect methods of computing information transfer in chapter 7.

5.3.2 The direct method of estimating information transfer

We have reviewed how the mutual information rate can be computed directly from a spike train. It was seen that the entropy rate of the spike train H(R) could not be computed exactly in practice and that approximations have to be made.

We have compared different information measures for different stimulus ensemble subsets as well as different contrasts, frequencies, and bandwidths in the case where the neuron is spontaneously active. Our results show that different information measures can give us quantitatively and in some cases qualitatively different results. Thus, results obtained with a particular measure must be interpreted carefully. We now discuss all four measures and their implications.

The mutual information rate calculated from a single member S of the stimulus ensemble I_{stim} has been used to quantify the information transfer by neurons in the visual system (Strong et al. 1998; Reinagel and Reid 2000; Fairhall et al. 2001; Lewen et al. 2001; Nirenberg et al. 2001). We have shown that this generally agrees well with the measure I_{gaus} obtained by averaging over a subset of the stimulus ensemble when the subset was the set W_{gaus} of zero-mean Gaussian stimuli of contrast σ and cutoff frequency f_c. However, some of these studies (Lewen et al. 2001; Nirenberg et al. 2001) used some naturalistic stimuli and it remains to be shown whether the measure I_{stim} would agree with the mutual information rate computed over the entire stimulus ensemble W.

A mutual information rate estimate about a stimulus S is only valid when the stimulus ensemble W is well sampled (Borst and Theunissen 1999). It remains to be shown whether the subset of all Gaussian stimuli W_{gaus} studied here provides a good sample of the total stimulus ensemble W and if so, in which cases. Thus, it is unclear whether the mutual information rate I_{gaus} calculated over the subset W_{gaus} would agree with the mutual information rate I calculated over the stimulus ensemble W. There is evidence that this might not be the case in sensory neurons as they must be most efficient at coding natural stimuli relevant for the animal's

survival. Indeed, natural stimuli have been shown to elicit higher information rates than low-pass filtered Gaussian white noise (Rieke et al. 1995; Attias and Schreiner 1998).

The total stimulus ensemble W is in general multi-dimensional and is not easy to characterize like W_{gaus} or W_{SAM}. It is therefore useful to have information measures like I_{spon} and I_{noise} that do not depend on the stimulus ensemble used, but rather depend on the properties of the neurons under study and on the particular stimulus used only through the noise entropy rate H(R/S). Accordingly, we have compared the mutual information rate calculated from the spontaneous activity of the neuron I_{spon} to the mutual information rate I_{gaus} calculated over the Gaussian stimulus ensemble W_{gaus} and showed that the two had the same dependence on stimulus contrast and cutoff frequency. The mutual information rate calculated from spontaneous activity I_{spon} was shown to be lower than I_{gaus}. This is because the measure I_{spon} only measures the reduction in the entropy rate of the spike train from spontaneous activity $H_{spon}(R)$ brought about by a stimulus. This difference is perhaps most clearly seen when no stimulus is present. We then have that $I_{spon}=0$ and $I_{gaus}=H_{gaus}(R)-H_{spon}(R)>0$. As mentioned earlier, the quantity $H_{gaus}(R)-H_{spon}(R)$ represents the information transmitted by the neuron about the fact that no stimulus is present in the environment since the "no-stimulus" case is just another "stimulus" in the stimulus ensemble W_{gaus}.

The two information rates I_{gaus} and I_{spon} were shown to differ by the quantity $H_{gaus}(R)-H_{spon}(R)$ which has no dependence on the particular stimulus S used (it only depends of W_{gaus}). Furthermore, we note that $H(R)-H(R/S)= H(R)-H_{spon}(R)+H_{spon}(R)-H(R/S)$. Thus, we have that $I= H(R)-H_{spon}(R)+I_{spon}$ using the definitions of I and I_{spon}. Thus, the information rate calculated from the spontaneous activity of the neuron I_{spon} differs by only a constant from the mutual information rate calculated over the full stimulus ensemble I. The two measures will thus behave in the same manner. The measure I_{spon} is zero when there is no stimulus and has the advantage of having no dependence on the stimulus ensemble.

It has been argued that a spontaneously active neuron must change one or more properties of its spike train in order to signal the presence of a stimulus (see Ratnam and Nelson 2000; Chacron et al. 2001a and references therein). Thus, a change from spontaneous activity might signal the presence of a stimulus for these neurons. One possible way to look at this change is to calculate the difference between the entropy rates of the spontaneous activity and the

activity in the presence of a stimulus: this is exactly the information rate I_{spon}. Such a measure might be applicable for example to a situation in which an animal must detect a stimulus starting from the absence of stimuli: this occurs for example when the electric fish must detect a prey in midstream (Nelson and MacIver 1999) or when an animal detects a sound in an otherwise silent environment using a spontaneously active auditory afferent.

We now discuss the conditions of applicability of the measure I_{spon}. Clearly, I_{spon} cannot be applied to a neuron that is not spontaneously active. Moreover, it cannot be applied to a neuron with a very regular spike train (e.g. a stretch receptor). Thus, there needs to be some irregularity in the spontaneous activity of the neuron. However, most spontaneously active neurons have irregular patterns of activity, and the trial-to-trial variability decreases when driven by a stimulus.

In nature, stimuli are often contaminated by unwanted noise. Thus, we have introduced a simple extension of the measure I_{spon} that accounts for background noise: I_{noise}. This measure could for be applied in the context of the aforementioned cocktail party effect (Cherry, 1953). Our results show that, when the stimulus contrast was low, I_{noise} and I_{spon} were approximately equal. However, as stimulus contrast increased, I_{noise} became higher than I_{spon}. This is because the spike train entropy rate in the presence of background noise $H_{noise}(R)$ was higher than $H_{spon}(R)$ and that the noise entropy rate in the presence of background noise $H_{noise}(R/S)$ tended towards $H(R/S)$ as the signal-to-noise ratio increased. This result might seem surprising at first. However, we are only calculating the effect of background noise on the spontaneous activity of a neuron. It has been shown experimentally that an animal could take advantage of background noise for neural coding and prey detection (Russell et al. 1999). The measure I_{noise} might underlie the means of neural computation used for such a task. In the electrosensory system, I_{noise} might be the appropriate measure for the problem the fish faces when detecting prey in background clutter (e.g. root masses). In the auditory system, this measure might be appropriate for quantifying the information transmitted by a spontaneously active neuron about a stimulus in the presence of background noise.

It is important to note that we only considered the case where the background noise was additive (i.e. the noise was added to the stimulus) and independent of the stimulus. This may not be the case in every situation (such as noise caused by other fish's EOD's for example).

Furthermore, we assumed that the background noise had a flat power spectrum up to 50 Hz and had a contrast of 0.01 mV. The validity of this assumption has yet to be tested and this will certainly not be true in general. The strength and frequency content of the background noise will most certainly vary depending on the particular situation the animal is in.

It is also known that there are internal noise sources such as channel and synaptic noise that contribute to the variability seen in experimental data. Stochastic processes (Tuckwell, 1988) account for these and other possible noise sources in our model. Such noise sources are responsible for the irregularity of the spontaneous activity of our model neuron and are present in all our simulations. As mentioned earlier, some neurons have little or no spontaneous activity. While I_{spon} could not be applied to such neurons, I_{noise} could in principle be applied if the background noise is sufficiently high so that it might elicit a sufficiently irregular response from that neuron. The background noise would in this case act as internal noise sources. Note that this is different from the measure I_{stim} as the background noise is not "frozen" across trials.

We now discuss the differences between natural and Gaussian stimuli. Unrepeated presentations of the latter will always elicit a highly variable neural response due to their high entropy since the Gaussian distribution is the distribution with the highest entropy when the variance of the distribution is constrained (Cover and Thomas 1991; Rieke et al. 1997). However, natural stimuli seldom obey Gaussian statistics as mentioned ealier. For example, the SAM's used in our study are not Gaussian. These stimuli have zero entropy as they are deterministic in nature and contain only a single Fourier component. Thus, it is not practical to estimate the mutual information rate with I_{stim} since the entropy rate of the spike train when stimulated by an unrepeated SAM is actually equal to the noise entropy rate obtained with that same SAM (i.e. $I_{stim}=0$). The stimulus presented must thus have some irregularity in order to apply the measure I_{stim}. The ensemble of SAM's is characterized by the SAM contrast σ_{SAM} and SAM frequency f_{SAM} as well as the probability of occurrence of each SAM. Although it is much easier to work with the SAM ensemble than the full stimulus ensemble W, the characterizing of the SAM ensemble would require much work as the probability of occurrence of each SAM in nature would have to be determined experimentally. However, the measures I_{noise} and I_{spon} can easily be calculated as they do not depend on the stimulus ensemble.

By comparing the measures I_{spon} and I_{noise} obtained for both low-passed filtered Gaussian stimuli and the SAM's, we found that the model neuron exhibited a resonance in information rate when the SAM frequency varied. This resonance can be expected from the combination of high-pass and low-pass filtering characteristics of the electroreceptor neurons (Hutcheon and Yarom 2000). Furthermore, the resonance in tuning seen in chapter 4 might interact with this resonance, further studies are needed to look into the nature of such interactions. Note that, in this reference, only responses to sinusoidal signals were used and no information theoretic calculations were carried out. It was however only revealed once we used naturalistic SAM stimuli rather than low-passed filtered Gaussian noise. The reason that low-passed Gaussian noise did not elicit this resonance is that we are merely adding higher frequency components to the signal as we increase the cutoff frequency f_c. Note however that the resonance might be revealed using higher cutoff frequencies than the ones used in this study. The responses of a neuron to natural stimuli can be very different than to artificial stimuli (Rieke et al. 1995; Attias and Schreiner 1998).

In conclusion, we have shown that different information measures are designed to answer different questions and thus give different results. Hence, one must thus carefully consider the measure one uses based on the physiological context. Although the direct method of calculating mutual information rate makes no assumptions about the neural code (Borst and Theunissen 1999), the mutual information rate can only be calculated in practice for very restricted subsets of the total stimulus ensemble W. Assumptions must thus be made either in terms of the stimulus ensemble itself or in terms of calculating the information based on the no-stimulus ensemble alone. For neurons that exhibit irregular spontaneous activity, I_{spon} provides a natural information measure that does not require a complicated stimulus ensemble. Further, I_{spon} differs from the Shannon mutual information rate by only a constant: the two information measures thus behave similarly. The mutual information rate in the presence of background noise I_{noise} could also be applied to neurons that are not spontaneously active.

6 Effects of Interspike Interval Correlation on information transfer

In this chapter, we look at the effects of interspike interval correlations (see chapter 2) on information transfer by neurons. These results have been published in Chacron et al. (2001b).

6.1 Setting up the problem

6.1.1 Introduction

It is known that electroreceptors display negative interspike interval correlations (see chapter 2). We wish to quantify the effects of these correlations on information transfer by comparing the LIFDT that was shown to display these correlations and the Nelson model that was shown to not display these correlations (see chapter 3). Throughout, the EOD frequency was chosen to be 1000 Hz. An EOD cycle thus corresponds to one millisecond.

6.1.2 Comparison

To ensure a fair comparison between the two models, parameters were chosen such that the ISIH's obtained under baseline activity were nearly identical. Results are shown in Figure 6-1, panels (a) and (c). The experimentally obtained SCC at lag one for the P-receptor data is -0.35 (Chacron et al. 2000) while it is -0.385 for the model. In contrast, the one obtained for the Nelson model is -0.05 and is negligible.

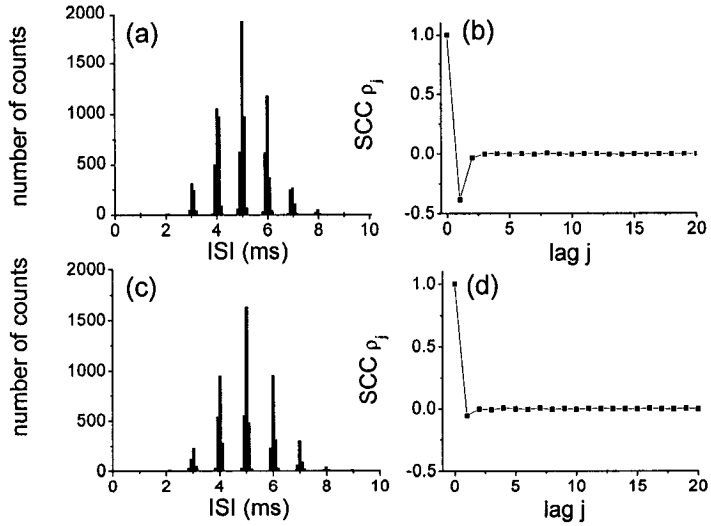


Figure 6-1 Comparison of the two models used in our study. (a) ISIH obtained from the analysis of 10000 consecutive ISI's from the LIFDT model ($\langle I \rangle$=4.9912 ms, VAR(I)= 1.1449 ms^2, CV=0.2143). (b) SCC's obtained with the model. (c) ISIH obtained from 10000 consecutive ISI's with Nelson's model ($\langle I \rangle$=4.9982 ms, VAR(I)= 1.1003 ms^2, CV=0.2098) . (d) SCC's obtained. Note that both models have the same distribution of ISI's but that the Nelson model does not exhibit any significant ISI correlations. Error bars are too small to be visible. Parameter values for the LIFDT model are the same as in Figure 3-4. Parameter values for the Nelson model are the same as in Figure 3-2.

To further ensure a good comparison between both models in the presence of time varying stimuli, we constructed the gain and phase response curves of both models to sinusoidal stimuli with frequency ranging from 0.1 Hz to 200 Hz. This was done by computing a cycle histogram from the data and fitting the function: G(f) sin($2\pi f\ t+\phi(f)$). The gain and phase values are given respectively by G and ϕ. This is illustrated in Figure 6-2.

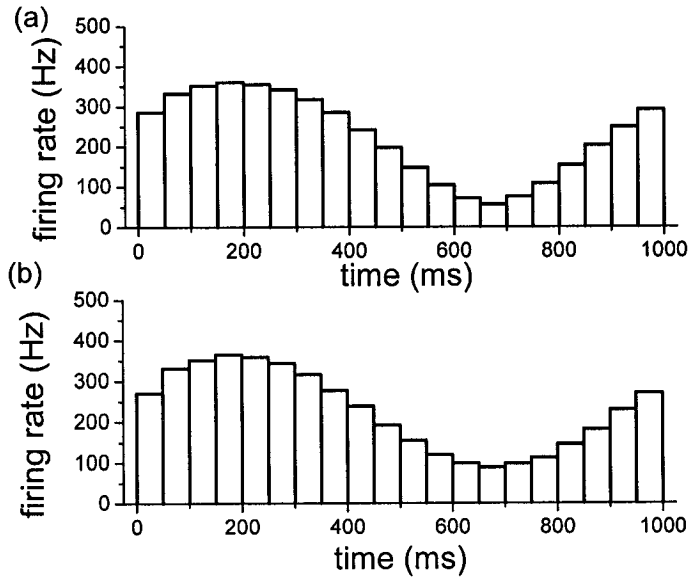


Figure 6-2 Cycle histograms. (a) Cycle histogram obtained with Nelson's model using a 1Hz sinewave. The amplitude was 0.05 mV. (b) Cycle histogram obtained with the LIFDT model with the same sinewave.

The gain and phase response curves are shown in Figure 6-3 and are nearly identical. Now that we have a good calibration, we can compare results obtained using both models to quantify the effects of negative ISI correlations on spike train variability as well as coding.

6.2 *Results*

6.2.1 Effects of correlations on spike train variability

The Fano factor curve (see chapter 2) obtained for the LIFDT model with $D_2=0$ is plotted in Figure 6-4 (triangles). We see that the spike train is more regular at all time scales than a Poisson process since $F(T) < 1$ (Cox and Lewis 1966). F(T) decreases for T in the 1-5000 ms range and has an asymptotic value of 0.00685 (n=5 line). We had previously stated an expression relating the asymptotic value of the Fano factor to the ISI correlation coefficients (see equation (2.9)):

$$\lim_{T\to\infty} F(T) = CV^2\left(1+2\sum_{i=1}^{\infty}\rho_i\right) \tag{6.1}$$

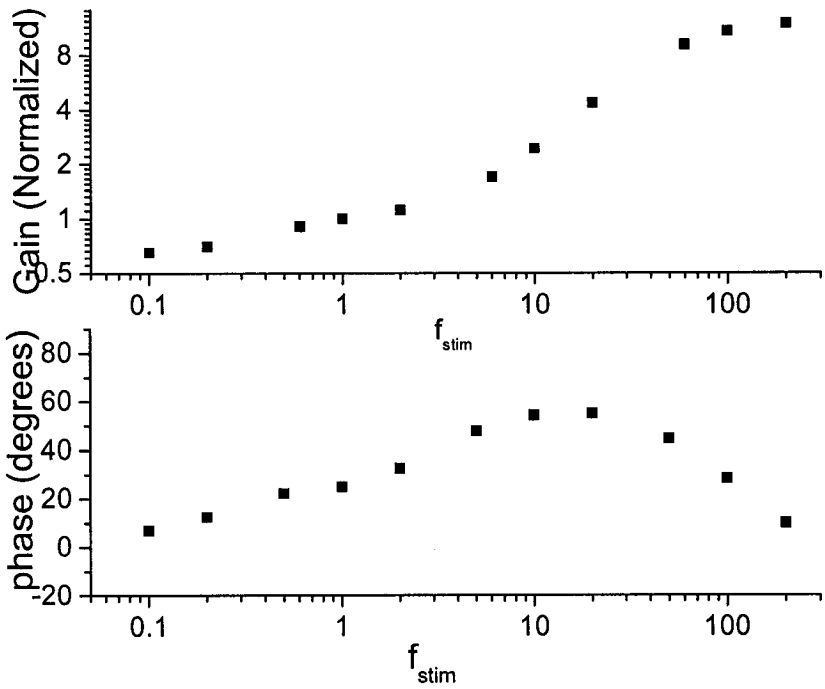


Figure 6-3: Gain (a) and phase (b) response curves obtained with both models for sinusoidal AM's of various frequencies. The curves lie on top of each other. The root mean squared baseline transdermal potential is $A_0/\sqrt{2}$=0.566 mV which is in the physiological range (Xu et al. 1996; Nelson et al. 1997; Nelson and MacIver 1999). As in Nelson et al. (1997), we consider that a sinusoidal AM has 0dB intensity when it produces a 1 mV change (RMS) in transdermal potential. SAM's of various frequencies were then presented to the model with the same intensities used in Nelson et al. (1997) to construct the phase and gain curves. The gains have been normalized by the value 1060 spikes/s/mV (this value is in the physiological range) obtained for f_{stim}=1 Hz.

If we take only the SCC's to be non-zero up to lag 5, then we get F_∞=0.00681 from equation (6.1) which is very close to the observed asymptotic value of 0.00685. For comparison, the Fano factor time curve obtained by random shuffle of the ISI sequence is also plotted in Figure 6-4 (diamonds). Since all ISI correlations have been eliminated by this operation, we

now have a renewal process (Cox and Lewis 1966) for which F(T) tends towards CV^2 from equation (6.1) (CV^2 line in Figure 6-4). Note that since $CV^2 \approx 0.0436 < 1$ in our case, we have $F_\infty < 1$ even in the absence of ISI correlations. The two curves (triangles and diamonds) are almost on top of one another for short counting times (<10 ms), implying that correlations do not play a significant role over this range from the Fano factor curve perspective; however, as we will see below, this is not the case from an information theoretic perspective. However, they become different for longer times: the Fano factor curve without correlations tends towards CV^2 while the one with correlations has a lower asymptotic value.

We plot the Fano factor curve obtained with the Nelson model in Figure 6-4 (squares). We see that it matches the one obtained for randomly shuffled ISI sequences from the LIFDT model. This match is not surprising, since the two models have for all practical purposes identical ISI distributions and thus identical CV's (see Figure 6-1). Furthermore, it demonstrates that there are no significant ISI correlations in the Nelson model. Since the SCC's are effectively negligible beyond lag 5 in the model, the Fano factor tends towards a constant for long counting times given by equation (6.1); this is not what is observed experimentally.

To look at the effects of the weak correlated noise λ_2 (see chapter 3), we now plot the Fano factor curve obtained with $D_2 = 9 \times 10^{-6}$ ms^{-1} in the LIFDT model. This curve is on top of the other ones for short counting times. In particular, the noise λ_2 is weak and has negligible effect on the ISIH and the SCC's at low lags (see below). However, the Fano factor curve differs from the others by increasing in a power law fashion for long counting times before saturating. The behavior can be understood from a plot of the mean and variance of the PND (Figure 6-5). The mean increases linearly with counting time; this is because the mean number of spikes expected in a time window of length T is equal to the length of that window multiplied by the mean firing rate. However, the variance is almost constant for short counting times, hence F(T) decreases. The variance then increases at a greater rate than the mean, hence F(T) increases. Finally, the variance and mean both increase with the same rate and the Fano factor is constant.

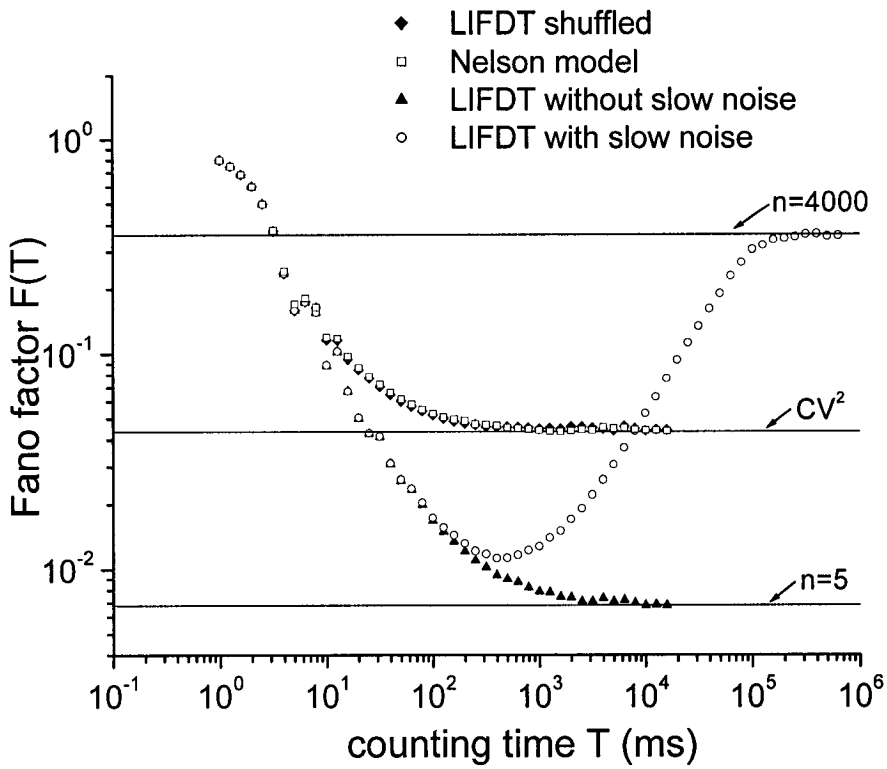


Figure 6-4: The Fano factor curves obtained with the LIFDT model with slow noise (open circles) and without the slow noise (filled triagngles), shuffled ISI sequence (filled diamonds), and Nelson model (open squares).

We now show a possible mechanism for the increase of the Fano factor at long counting times. It relies on the presence of weak positive SCC's that extend out to long lags. These positive correlations are extremely small and cannot be seen from a plot of the SCC's ρ_j as a function of j. However, their presence is revealed by the increase of the ISI PSD at low frequencies (Cox and Lewis 1966) (Figure 6-6). This can be seen in the following simple example, in which the form for the SCC's at lags greater than zero is assumed to be:

$$\rho_j = -0.385\ \delta_{1j} + 0.00225\ e^{-\frac{j}{4000}} \qquad (6.2)$$

where δ_{ij} is the Kronecker delta function ($\delta_{ij}=1$ if $i=j$ and $\delta_{ij}=0$ if $i\neq j$). Parameters were chosen to fit the expression for the PSD (see chapter 2) that is plotted in Figure 6-6. We have only retained the negative SCC at lag one as it is dominant. We see that the PSD corresponding to the SCC's given by eq. (6.2) is very similar to the one obtained with our model for $D_2=9\times10^{-6}$ ms^{-1}. Since the PSD and the SCC sequence are completely equivalent (see Cox and Lewis 1966 and chapter 2), this justifies our assumptions for equation (6.2). These positive correlations at lags >1 sum up according to equation (6.1) to give the increase in the Fano factor. Thus, adding a slow noise to the membrane voltage increases spike train variability at long counting times.

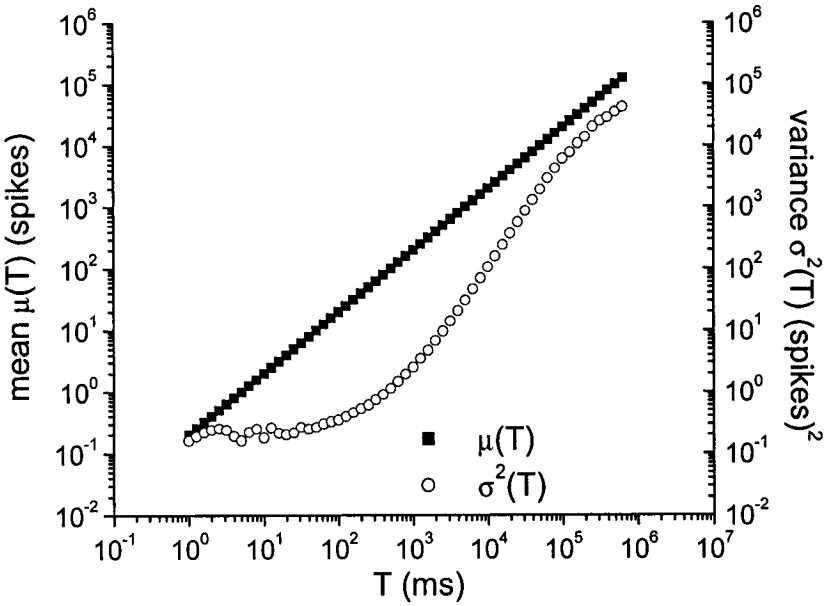


Figure 6-5: Mean and variance of the PND as a function of counting time T for the LIFDT model with slow noise of intensity $D_2=9\times10^{-6}$ ms^{-1}. The mean increases linearly with counting time. The variance is at first almost constant, which leads to a decrease in F(T); it then increases faster than the mean (F(T) increases). At long counting times, both increase at the same rate, hence F(T) is constant.

ISI sequences obtained from other neural systems have been shown in some cases to display negative ISI correlations at short lags and positive ISI correlations at long lags (Lowen and Teich 1992). Furthermore, an increase in the Fano factor curve has been observed in many preparations (Teich 1992; Lowen and Teich 1992; Teich et al. 1996; Lowen and Teich 1996;

Turcott and Teich 1996; Teich et al. 1997) and has been modeled by driving the rate of a Poisson spike generator with coloured noise (Teich 1992; Teich et al. 1996; Teich et al. 1997). We have gone beyond this standard description by showing how this increase arises in a more biophysically plausible model.

It is known experimentally that the Fano factor curves obtained for different P-units have the same qualitative shape, although their respective minimum occurs between 40 and 1000 ms (Ratnam and Nelson 2000). Our model produces Fano factor curves quantitatively similar to those obtained experimentally. Further, the location of the minimum is found to be dependent mainly upon D_2; in fact, by a suitable choice of D_2, we can obtain Fano factor curves matching the full experimental range (data not shown). For example, if we take $D_2=10^{-4}$ ms^{-1}, then the minimum of the Fano factor curve is at T=40 ms. On the other hand, if we take $D_2=10^{-6}$ ms^{-1}, then the minimum is at T=1000 ms. The correlation time τ_2 of λ_2 also influences the position of the minimum.

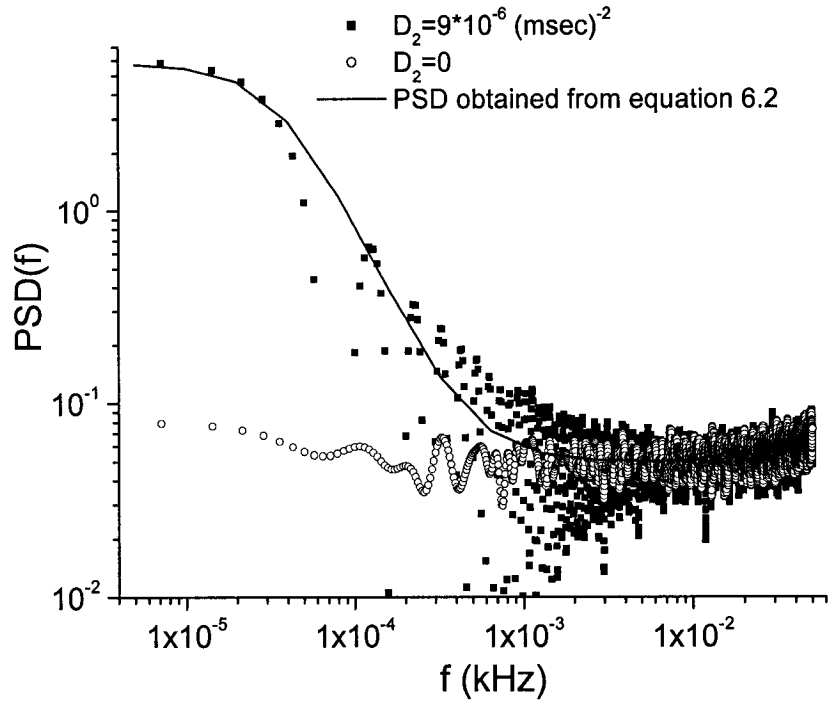


Figure 6-6: PSD of the ISI sequence obtained with the LIFDT model with D_2=0 and D_2=9x10^{-6} ms^{-1}. The PSD obtained by assuming a form for the SCC's (see text) is also shown (solid line) and matches the curve obtained for D_2=9x10^{-6} ms^{-1}.

These results imply that the remarkable regularity of P-unit spike trains at counting times of roughly 250 ms with the parameters chosen (within the experimentally observed range - Ratnam and Nelson 2000) can be entirely explained by negative ISI correlations present experimentally and in the LIFDT model at low lags. These arise from cumulative relative refractoriness exhibited by our dynamic threshold. We have shown here that negative SCC's contribute to the reduction of the Fano factor and that positive SCC's increase the Fano factor according to equation (6.1). Ratnam and Nelson (2000) have shown that modeling the ISI sequence by a first-order Markov chain gave the correct SCC at lag one. However, their SCC's at longer lags had higher absolute values than observed experimentally and alternated in sign. As a consequence, the SCC sum for their model was greater (less negative) than the one calculated from the experimental data. Thus, the Fano factor calculated with their Markov chain model did not decrease as much as the Fano factor from their experimental data. In

contrast, our model reproduces the descending part of the Fano factor curve seen experimentally (compare Figure 6-4 with figure 11F of Ratnam and Nelson 2000).

We now discuss the power law increase of the Fano factor curve in more detail. We first note that since we are using Ornstein-Uhlenbeck noise, which has a finite correlation time τ_2, the Fano factor will eventually saturate to a finite value. This value is equal to 0.36188 (figure 4 n=4000 line) and is obtained approximately by taking the SCC's up to lag 4000 in equation (6.1) (hence $\rho_j=0$ for $j>4000$). This implies that the 4000th ISI is still correlated with the first one. Ratnam and Nelson (2000) have found that, for many P-units, the ISI sequence cannot be modeled by a Markov process of order less than ten. Our results suggest that a Markov chain of ISI's of order at least 4000 would be required to correctly reproduce this feature of the Fano factor curve. In contrast, our simple dynamical model accurately accounts for the full behavior of the Fano factor curve. One plausible origin of the slow Ornstein-Uhlenbeck noise might be fluctuations in synaptic neurotransmitter secretion rates that exhibit long term correlations (Lowen et al. 1997). However, it could also be due to slow drifts in EOD amplitude or frequency (Moortgat et al. 1998).

6.2.2 Effects of ISI correlations on the detection of weak signals

We show in Figure 6-7 the PND's obtained in the presence and absence of ISI correlations for four different counting times for the LIFDT model. We see that ISI correlations have minimal effects at short counting times such as 20 ms (a). The effect of negative ISI correlations increases with counting time (b). It is very important around e.g. 255 ms (c), where the variance of the PND is reduced while the mean is left unchanged. This effect diminishes for longer counting times (d) where positive ISI correlations contribute to the broadening of the PND.

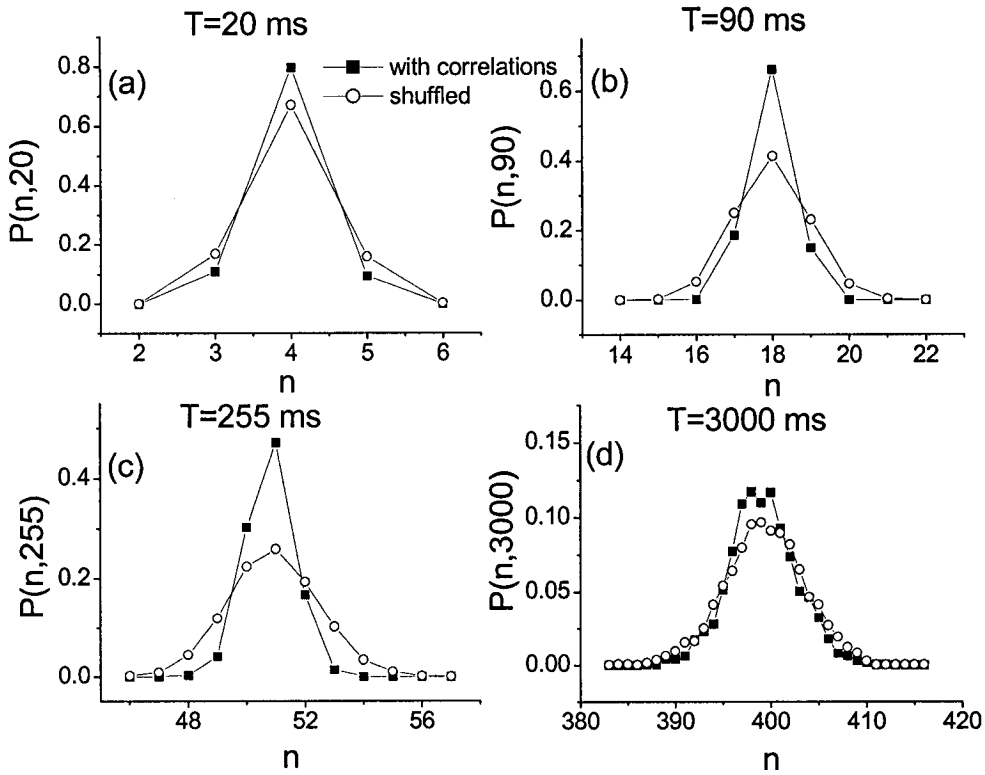


Figure 6-7: PND's obtained for both models for various counting times. (a) 20, (b) 90, (c) 255, (d) 3000 ms. ISI correlations reduce the variance of the PND while keeping the mean unchanged. This effect is maximal at counting times where the Fano factor is minimal.

This has implications for weak signal detection using the ideal observer paradigm (Green and Swets 1966). Intuitively, if $P_0(n,T)$ and $P_1(n,T)$ do not overlap much, then we have a very good detector (Gabbiani and Koch 1998). As mentioned in chapter 2, a good measure for discriminability is the discriminant measure d' or SNR (Green and Swets 1966; Nachimas 1972; Snippe and Koenderink 1992; see chapter 2) defined by:

$$\mathrm{SNR} = \frac{|\mu_1 - \mu_0|}{\sqrt{\sigma_1^2 + \sigma_0^2}} \tag{6.3}$$

As mentioned in chapter 2, we have assumed in equation (6.3) that the $P_i(n,T)$ are Gaussian (i.e. we have neglected their third and higher moments). This is not too restrictive as they approach Gaussian distributions for high T by the central limit theorem. Furthermore, the PND's obtained with the models are bell-shaped (Figure 6-7) and the Gaussian approximation is reasonable. Optimal discrimination, and hence detection, requires SNR to be higher than some threshold SNR_{crit}. Let f_0 be the baseline steady state firing rate of the electroreceptor and further suppose that the stimulus varies slowly with time and leads to a new steady state firing rate f_1. Here, we do not consider transients in P-unit firing rate that can occur after a change in EOD amplitude (Xu et al. 1996, Nelson et al. 1997), but rather only the new steady state firing rate. Furthermore suppose that the signal is weak, and that as a consequence, the variances of the PND's are approximately equal. Hence $\mu_1(T)=(f_1/f_0)\ \mu_0(T)$ and $\sigma_1^2(T) \cong \sigma_0^2(T)$ and using equation (6.3) the inequality $SNR \geq SNR_{crit}$ becomes:

$$F(T) \leq \left| \frac{f_1}{f_0} - 1 \right| \frac{\sigma_0(T)}{\sqrt{2}\ SNR_{crit}} \qquad (6.4)$$

since $F(T) = \sigma_0^2(T)/\ \mu_0(T)$. Furthermore, $\sigma_0(T)$ does not vary much if we consider low counting times (Figure 6-5), hence it can be considered constant to a first approximation.

From equation (6.4), the lower the Fano factor, the better the detector. A good value for SNR_{crit} that gives almost no overlap between the $P_0(n,T)$ and $P_1(n,T)$ distributions is 3-5 (see chapter 2). Using this value, one can find a lower bound for $|(f_1/f_0)-1|$ from equation (6.4). From the data for T=255 ms, we have: $\sigma_0(255)=0.78$, and $F(255)=0.012$, hence a single P-unit can near perfectly discriminate steady state response to slow amplitude modulations that cause a steady state firing rate as low as 6.5% higher than baseline firing rate. We tested this by giving a step increase in EOD amplitude to the model. The negative ISI correlations at low lags lead to a significant improvement in the ROC curve (Figure 6-8). These fish are very good at detecting prey using their electrosensory system (Nelson and MacIver 1999). As there are relatively few numbers of false strikes towards the prey, P_{FA} must be low (Ratnam and Nelson 2000). Our results show that the improvement in the detection probability P_D due to ISI

correlations is in fact greatest for low P_{FA}. Thus the animal significantly improves its chances at detecting prey by having negatively correlated ISI's.

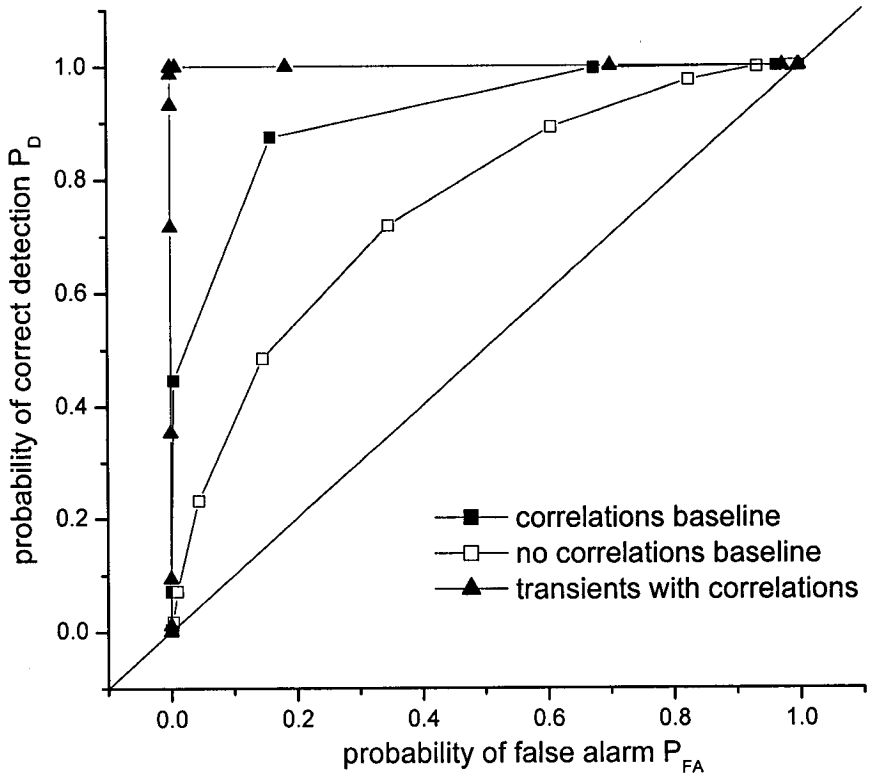


Figure 6-8: ROC curves obtained for both models. Correlations improve the ROC curve by decreasing the variance of the PND's, leading to a better discriminability between the distributions. The stimulus was a 4% step increase in EOD amplitude (the new amplitude value is equal to 1.04 times the old value). Including the transients leads to a further increase of the mean of the PND with stimulus, which further increases discriminability and hence further improves the ROC curve.

In the above analysis, we did not consider the effect on signal detection of transients in firing rates that are associated with changes in EOD amplitude (Xu et al. 1996; Nelson et al. 1997). Transients will help the animal in detecting weak signals by increasing/decreasing the firing rate, thus shifting the mean of the PND with stimulus away from the baseline PND. This is confirmed in Figure 6-8 where the transients resulting from a 4% step increase in EOD amplitude lead to a near perfect detection of the signal. These transients are present in both models since both incorporate the Nelson filter that is responsible for response to changes. Thus, there are two phenomena at work here: transients resulting from a change in EOD

amplitude will tend to shift the mean of the PND while negative correlations will reduce the variance. The two effects work in tandem to achieve better discriminability between the PND's, hence a better ROC curve and a lower stimulus contrast threshold for signal detection.

Usually, for renewal point processes, fluctuations due to noise average out over time and the ROC curve improves with counting time (Gabbiani and Koch 1998). In our case, the ROC curve becomes worse with respect to the one obtained with shuffled ISIs at longer counting times as the positive ISI correlations increase variability (Figure 6-4). Thus there is a time window where signal detectability is optimal for the animal and it corresponds to the counting time at which the Fano factor is minimal. This optimal counting time has been shown to vary between 40 and 1000 ms (Ratnam and Nelson 2000) within the P-unit population, a range that is easily reproduced by our model.

6.2.3 Effects of correlations on information transfer

All the previous analysis assumed a rate code, and thus spike timing was considered unimportant. A recent study on the electroreceptors of a similar electric fish (*Eigenmannia*, see Kreiman et al. 2000) showed that significant jitter could be added to the spike train without affecting the quality of encoding using the stimulus reconstruction technique (Rieke et al. 1997, Gabbiani and Koch 1998). This study however was carried out when the stimulus cutoff frequency was low (<20 Hz). This suggests that a rate code might be more relevant for the encoding of low frequency stimuli. However, electrocommunication signals contain much higher frequency components (Metzner and Heiligenberg 1991; Zupanc and Maler 1993; Dulka et al. 1995). In fact, the animal can detect AM frequencies greater than 200Hz (Bastian et al. 2001). We use information theoretic measures to assess the quality of encoding time varying stimuli by P-units at such frequencies. In particular, we will assess the role of ISI correlations by comparing the information rates obtained with and without these correlations.

To assess the P-unit's ability to encode different frequency stimuli, we used low-passed filtered Gaussian white noise with a variable cutoff frequency f_c (Wessel al. 1996, see chapter 5). The same stimulus was presented to both models and the resulting baseline and noise

entropies were calculated as detailed in chapter 4. We thus used the measure I_{spon}. We plot the baseline entropies obtained for each model in Figure 6-9. Note that the baseline entropy values for our model are always lower than for Nelson's as correlations reduce entropy (Shannon 1948; Cover and Thomas 1991). As ISI correlations will also reduce the noise entropy, they might therefore be expected to either reduce or increase the mutual information. We will show that noise entropies will in fact lead to an increase in mutual information.

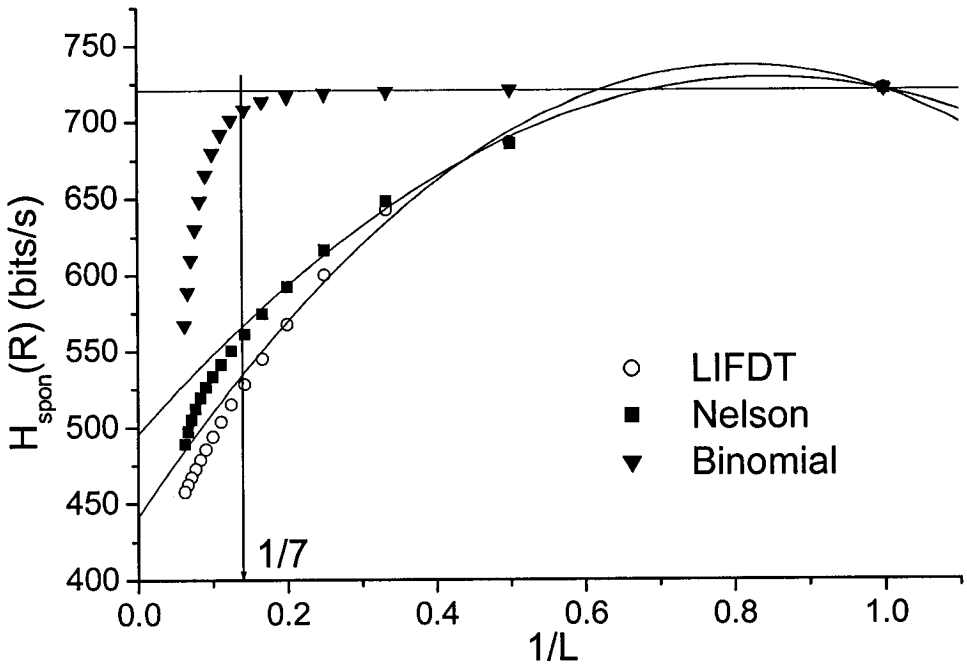


Figure 6-9: Baseline entropies of words of length n calculated from 1000 realizations of duration 10000 EOD cycles each for both models. Comparison with a binomial process for which these entropies should be constant is shown to verify the accuracy of the results. Finite sampling errors occur for words of length greater than 7 and lead to results that are lower than the true value (as in Strong et al. 1998). Clear deviation from linearity is still seen for the first few points in both cases. Entropies are lower for the LIFDT model due to correlations. Also shown are the best fits obtained for the data.

6.2.3.1 Varying the stimulus contrast

We calculated the information rates at different stimulus contrasts σ_{stim} for both models. The baseline and noise entropy rates for different contrasts were calculated at a fixed cutoff frequency f_c=100 Hz to assess the P-unit's capacity to encode high-frequency stimuli (as discussed above). We plot in Figure 6-10 the information rate for both models. As expected, it increases with stimulus contrast for both models; however the information for the LIFDT model is always higher than the one obtained for Nelson's model. For example, for stimulus contrasts between 0.004 mV and 0.005 mV, the gain in information rate is approximately 40 bits/second (30%). For both models, the correlation between successive spikes in the presence of a stimulus increases with σ_{stim} resulting in a decrease in noise entropy rate. This leads to an increase in the information rate. However, this correlation is higher for the LIFDT model due to the dynamic threshold, which reduces the noise entropy even more and leads to a higher information rate.

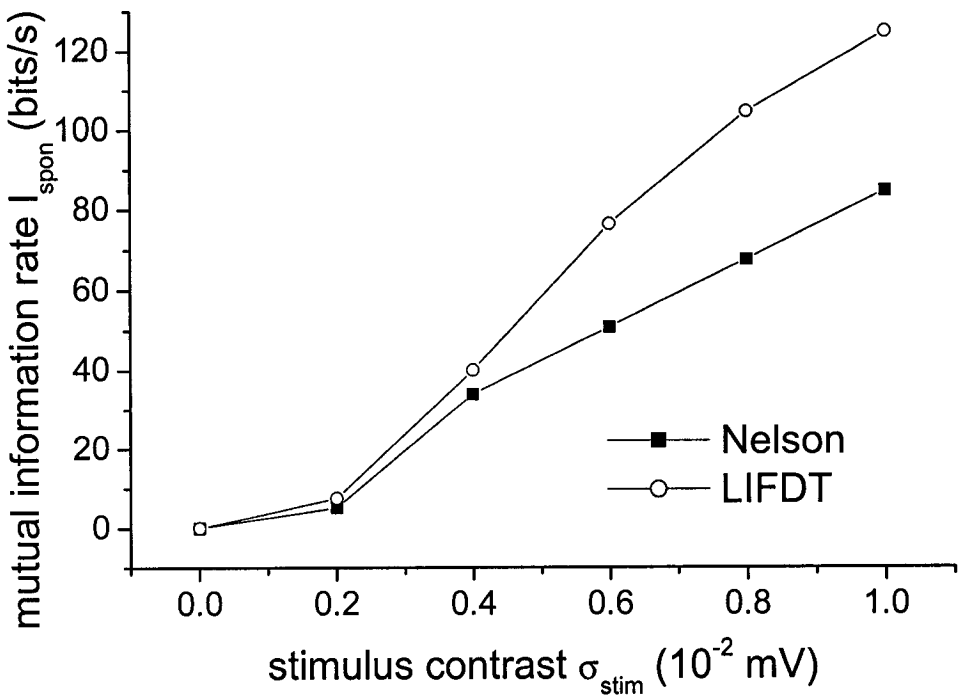


Figure 6-10: Information rate as a function of stimulus contrast for both models. The stimulus is low-pass filtered Gaussian white noise with cutoff frequency 100 Hz and has zero mean and standard deviation σ_{stim}. Information rate increases with contrast for both models but it is higher in the LIFDT model than in the Nelson model.

6.2.3.2 *Varying the cutoff frequency*

We now look at the dependence of information rate on stimulus cutoff frequency f_c. We thus construct an "information tuning curve" for the two models (see Figure 6-11a), i.e. the dependence of information rate on the stimulus' frequency content, quantified here by its bandwidth $(0,f_c)$. For both models, information rates are small at low cutoff frequencies and increase for higher cutoff frequencies, such as those near the mean baseline firing rate of the neuron (200 Hz in our case). The general shape is due to the fact that the noise entropy decreases as a function of f_c (over the range of interest). This decrease may be due in part to the high-pass filter characteristics of both models over the range of frequencies considered, since this filtering enhances the amplitude of the Fourier components at mid to high frequencies.

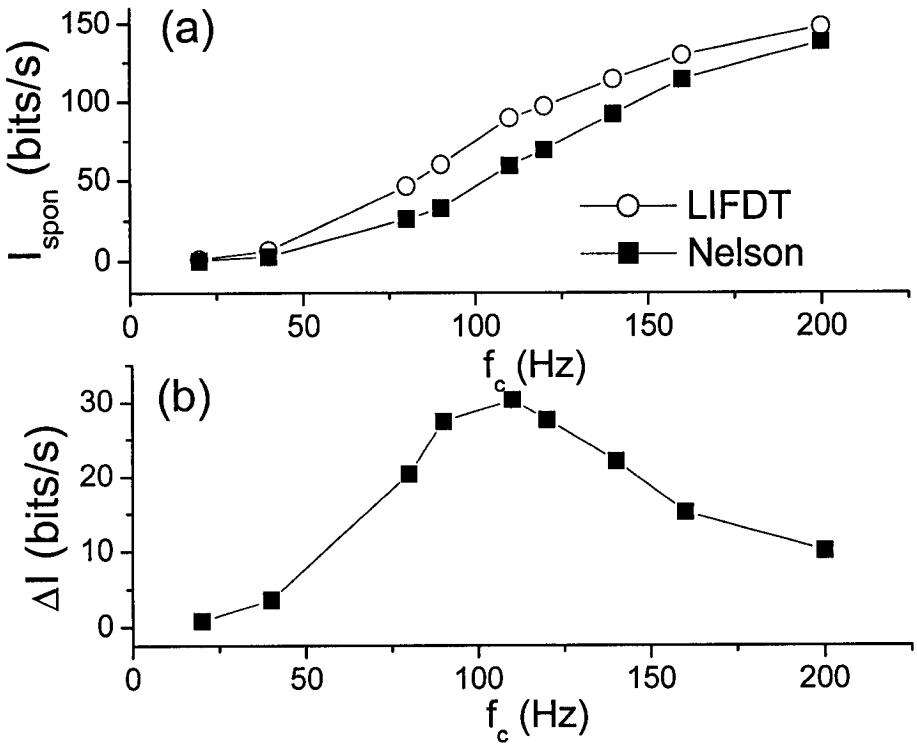


Figure 6-11: Effects of varying the cutoff frequency. (a) Mutual information rate as a function of stimulus cutoff frequency. The stimulus contrast was σ_{stim}=0.006 mV. The information rate increases with frequency for both models. However, information rate is higher for the LIFDT model than for the Nelson model. (b) Difference between the information rate for the LIFDT model and the Nelson model ΔI as a function of f_c. This quantity displays a resonance for frequencies on the order of the inverse of the decay time constant of the dynamic threshold τ_θ.

Information rates were higher for the LIFDT model compared to the Nelson model over the full range of cutoff frequencies f_c considered. A surprising result is the fact that the gain in information in the LIFDT model due to ISI correlations exhibits a maximum at frequencies around 110 Hz (Figure 6-11b). Hence our dynamic threshold not only enhances the information tuning curve, but this effect is maximal for stimuli with a specific cutoff frequency f_c. Interestingly, this frequency is on the order of the inverse of the decay time constant of the dynamic threshold τ_θ.

This effect can be understood intuitively as follows. Cumulative relative refractoriness leads to a less variable neural response to repeated stimuli because it causes negative ISI correlations. If the unit fires e.g. on the rising phase of the stimulus, then it has time to recover and fire on the next rising phase of the stimulus, which occurs on average at least after one correlation time f_c^{-1}. One thus expects this effect to be maximal when f_c^{-1} is on the order of the time constant of the cumulative relative refractoriness (modeled here by our dynamic threshold). If f_c is too low, the stimulus fluctuates on a longer time scale than the dynamic threshold and there is no enhancement of information transmission. Hence the information rates obtained with both models are approximately the same (see Figure 6-11a, below f_c=50Hz). For high f_c, the dynamic threshold cannot follow the fast stimulus variations and the two information rates are again the same. There is thus a resonance in the gain in information rate at approximately 100 Hz due to cumulative relative refractoriness. The maximum gain is about 30 bits/s (about 30%). This resonance in information transfer also comes from the fact that the LIFDT model exhibits a resonance in phase locking for frequencies equal to the inverse of the threshold correlation time (see chapter 4).

Behavioral experiments have demonstrated that the animal can reliably detect AM's in the range 100-200 Hz frequencies and that this might be relevant for courtship behavior (Bastian et al. 2001). We hypothesize that the resonance in the P-unit information tuning curve caused by our dynamic threshold may be responsible for this observed sensitivity.

Note that if the stimulus correlation time f_c^{-1} were high, then one might expect that the information rate over an EOD cycle would be lower than the one obtained if f_c^{-1} were low. This would be due to over-sampling. Note also that for a stimulus cutoff frequency f_c, the electrosensory system might integrate the input using a time window proportional to f_c^{-1} rather than counting over an optimal counting window. These possible effects can be examined by dividing the information rate by f_c. This results in an estimate (I_c) of the average information transmitted during a time window equal to the correlation time f_c^{-1} of the stimulus. This quantity is shown to increase as a function of f_c for both models (Figure 6-12a) but saturates for high f_c. Thus, more information is transmitted about stimuli with a larger bandwidth even when the integration times are normalized. This quantity saturates and decreases for high f_c due to under-sampling by the P-unit's baseline firing rate (Nyquist theorem). However, I_c is

still higher for the LIFDT model. The gain also has a resonance with the maximum around 100 Hz (Figure 6-12b).

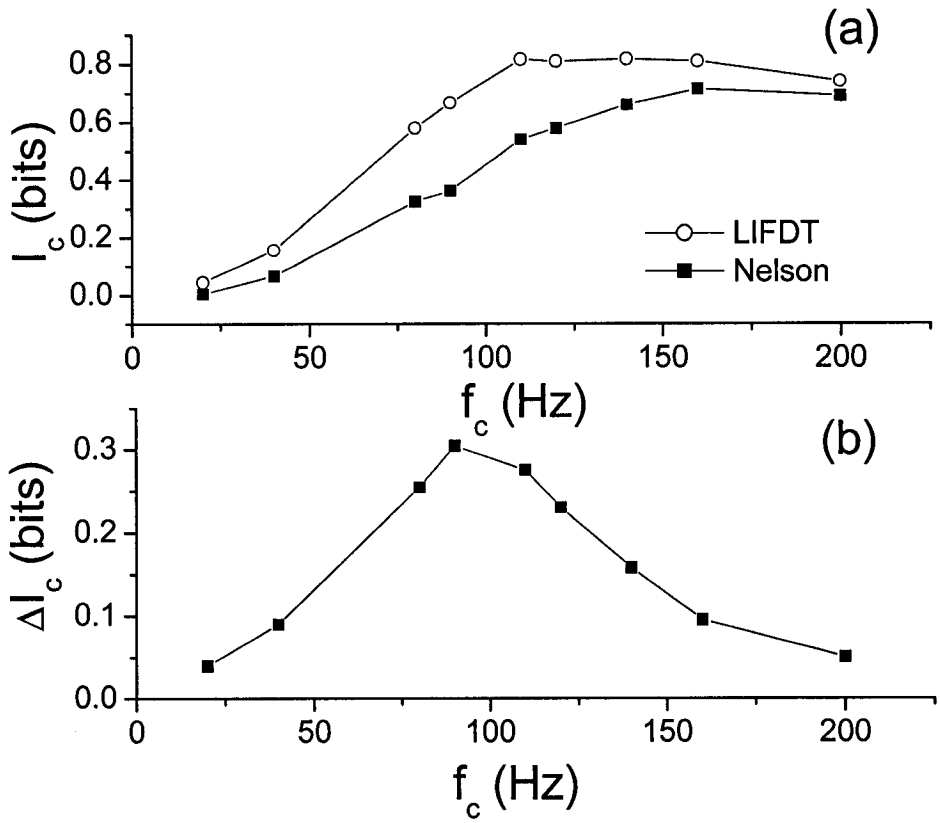


Figure 6-12: Effects of varying the cutoff frequency on the average information transmitted during the reciprocal of the cutoff frequency. (a) Average information transmitted I_c during a time window whose length is equal to the stimulus correlation time f_c^{-1} plotted as a function of cutoff frequency f_c. This quantity increases with f_c (see text for explanation). (b) Difference in I_c between the LIFDT and Nelson models as a function of f_c.

6.2.3.3 Effects of varying the bin width $\Delta\tau$

Finally, we show that the importance of spike timing increases with stimulus cutoff frequency. We introduce "blurring" in the spike train by making the bin width $\Delta\tau$ greater and studying the incurred loss in information for different cutoff frequencies of the stimulus. We see that this loss is greater for higher cutoff frequencies (Figure 6-13). This suggests that spike timing has minimal contributions in encoding information about slowly varying stimuli. Furthermore, this

finding agrees with the fact that significant jitter does not affect the quality of the reconstructed stimulus for low stimulus cutoff frequencies (Kreiman et al. 2000). However, spike timing is important for high-frequency time-varying stimuli and our results show in fact that electroreceptors can encode stimuli that vary on time scales at least as fast as 5-10 ms. This agrees with the results from a previous study that found that spike timing jitter in electroreceptors was on the order of 1-2 ms (Kreiman et al. 2000). This also justifies our use of PND's for low frequency stimuli in section 6.2.2.

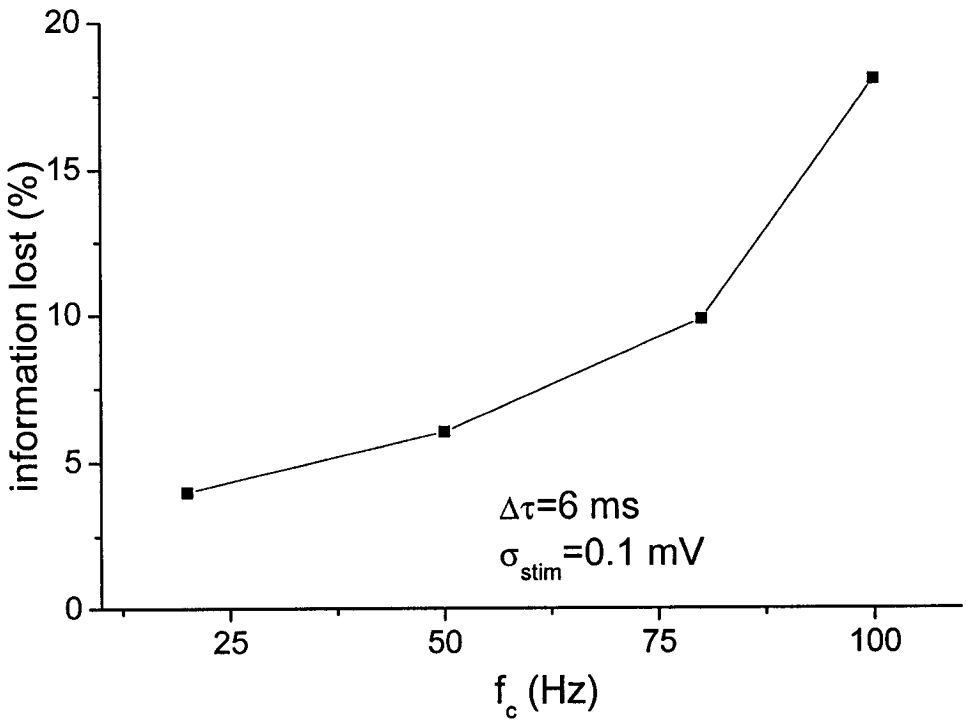


Figure 6-13: Information loss caused by "blurring" the spike train. We counted the number of spikes that occurred in successive time intervals of 6 ms in length and computed the information rate for that spike train. We can see that fractional loss of information increases with f_c: this quantity is calculated as $100\times[I_{spon}(\Delta\tau=1\text{ ms})-I_{spon}(\Delta\tau=6\text{ ms})]/I_{spon}(\Delta\tau=1\text{ ms})$.

6.2.4 Contrasting signal detection and information theory

As the fish swims by a prey, it will experience a small change in transdermal potential in a time window of roughly 200 ms (Nelson and MacIver 1999). However, information rates are almost zero for such low cutoff frequencies (this corresponds to frequencies around 5 Hz, the information rate is almost zero for $f_c<50$ Hz for the particular P-unit cell modeled here: see Figure 6-11a). The fact that the animal can readily detect these signals (Nelson and MacIver 1999) suggests that the mutual information rate calculated using baseline entropy is poorly suited for coding of these low frequency stimuli. This is also the case after the correlation time of the stimulus has been taken into account (Figure 6-12a). However, we have shown that measures based on signal detection theory were apt at quantifying the electroreceptor's ability to transmit information about constant stimuli (Figure 6-7 and Figure 6-8). It is thus more natural to analyse slow time varying stimuli by looking at the trial-to-trial variability of the PND.

We use a weak slow time varying stimulus (σ_{stim}=0.01 mV, f_c=1.96 Hz) and look at the PND in a time window of 255 ms. The portion of stimulus used is shown in Figure 6-14(a). The PND's obtained with both models are shown in Figure 6-14(b). Note that for each model, the PND's with and without stimulus have almost the same variances. Thus the main factor for discriminating between the two distributions is the difference in their means. This is not captured by entropy measures as they do not depend on the mean of the distribution used (Shannon 1948).

However, as shown previously for a 4% step increase in EOD amplitude (Figure 6-7), the variance of the PND's obtained with the LIFDT model are smaller than those obtained with the Nelson model for T=255 ms. Note that the means are separated by as few as 3 spikes in both cases (as low as 6% difference). The lesser overlap for the LIFDT model results in a dramatic improvement in discriminability. For example, the probability of obtaining 53 spikes for the LIFDT model with stimulus is 0.3, and is 0.05 without stimulus corresponding to a ratio of 6. For the Nelson model, this ratio is 2.5. Our results thus show that a single P-unit could discriminate as few as 2 extra spikes in a time window of 255 ms as already suggested by Ratnam and Nelson (2000). Their first-order Markov process was not able to reproduce the experimentally observed probabilities of correct detection. This is because the Fano factor for their model was higher than the one for the experimental data. The key factor in this

remarkable sensitivity is the negative ISI correlations observed experimentally and which result from cumulative relative refractoriness exhibited by our simple model with dynamic threshold.

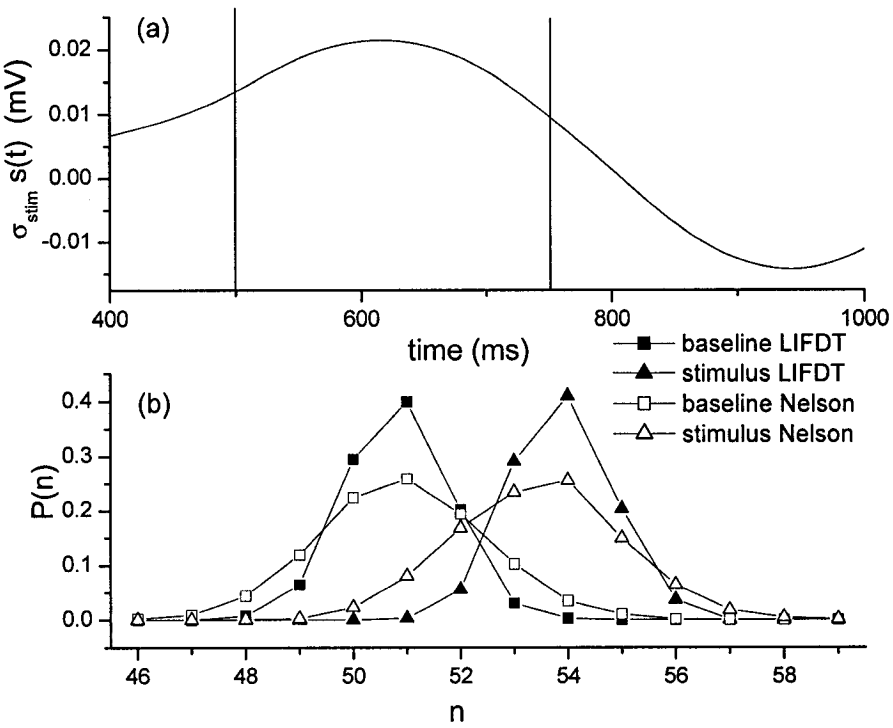


Figure 6-14: Effects of ISI correlations on signal detection. (a) A realization of low-pass filtered Gaussian white noise with cutoff frequency f_c=1.96 Hz and contrast σ_{stim}=0.01 mV. The counting time window is between the horizontal lines (from 500 ms to 755 ms). Note that this stimulus does not have zero mean. (b) PND obtained with this portion of stimulus as compared to the one obtained from baseline dynamics. The PND's variances are nearly identical and low while the means are very different, which leads to a good discriminability d'. The difference between the means is not captured by entropy measures, since they do not depend on the mean and also assume infinite spike trains. As there is no averaging over the stimulus in this case, signal detection yields better results than information theory.

6.3 Discussion

Using two models, one with baseline ISI correlations and one without, we have shown that negative ISI correlations present experimentally may play an important role in the animal's ability to detect both slowly and rapidly time-varying stimuli. A first analysis based on signal detection theory, which assumes rate coding, revealed that the ISI correlations dramatically enhanced the detectability of low frequency weak signals. We then used information theory to quantify the electroreceptor's ability to encode time-varying stimuli with various cutoff frequencies. We found that ISI correlations also helped in this case and that the effect was maximal for cutoff frequencies on the order of the inverse of the decay time constant of our dynamic threshold used to model cumulative relative refractoriness. By comparing both approaches, our study suggests that a rate code can be assumed for low frequency stimuli while spike timing is important for high frequency signals.

6.3.1 Comparison of models

To quantify the effects of correlations, we used the LIFDT model as well as Nelson's model. To ensure a fair comparison, parameter values were chosen such that both models had identical first order ISI statistics and responses to sinusoidal AM's. However, the Nelson model did not display negative ISI correlations while the LIFDT model did.

6.3.2 Spike Train variability and signal detection

Spike train variability as measured by the Fano Factor was computed for both models at various time scales. Although able to reproduce first order ISI statistics as well as responses to AM's, the Nelson model did not reproduce experimentally observed spike train variability as measured by the Fano factor at counting times greater than 10 ms (see Figure 6-4 and Ratnam and Nelson 2000). The asymptotic value of the Fano factor for the Nelson model was shown to equal the square of the coefficient of variation of the ISIH as expected from eq. 1 in the absence of ISI correlations. The Fano factor obtained from our LIFDT model also tends towards this value when the ISI sequence was randomly shuffled to eliminate ISI correlations, thus supporting the fact that the Nelson model did not display any significant ISI correlations.

The negative ISI correlations obtained with our simple LIFDT model with only one noise source bring the Fano factor down (Cox and Lewis 1966) to experimentally observed values for longer counting times. A discrepancy was however observed for counting times greater than 150 ms as the Fano factor decreased monotonically while the experimentally observed one increased. To get this increase in variability at long counting times it was necessary to add a weak noise with a long correlation time to the model dynamics. This noise had no effect at low counting times as explained above. However, it led to positive ISI correlations extending to long lags. These were also reflected in the increase of the PSD at low frequencies (Figure 6-6). The positive ISI correlations were weak and decayed exponentially up to long lags and summed up according to equation (6.1) to give the increase in the Fano factor. These ISI correlations make modeling of electroreceptors by Markov chains of ISI's difficult as ISI correlations can extend out to lags in the thousands. However, the addition of an extra noise term in our simple biophysical model gave quantitatively accurate results and can also be used to incorporate these effects into other neuron models.

We used the ideal observer paradigm (Green and Swets 1966) to optimally discriminate between pulse number distributions with and without stimulus (Koch and Gabbiani 1998). It was shown that negative ISI correlations reduced the variance of these pulse number distributions without changing their means significantly, hence increasing discriminability. This led to an improvement in the receiver operating characteristic curve that was maximal at counting times where the Fano factor was minimal. A criterion for near perfect discrimination was derived and it was shown that signals leading to a 6.5% increase in firing rate could be discriminated in this manner based on steady state dynamics alone. Transients caused by the filter further lowered this detection threshold by increasing the firing rate computed over the duration of the stimulus thus leading to a further increase in discriminability. The assumptions made for the signal detection analysis in this chapter (i.e. the animal is trying to look at spike count over a certain time window) are not appropriate for zero-mean high-frequency signals (e.g. beat frequencies generated by fish with high EOD frequency differences) as these stimuli might not cause a significant change in firing rate over the counting time window. Signal detection is however appropriate for low frequency signals such as those caused by prey that will cause a change in firing rate.

6.3.3. Using information theory to quantify the coding of time varying stimuli

The trial-to-trial variability of the neural response to a repeated stimulus was characterized by the noise entropy rate. The entropy rate of the baseline spike train was estimated and the difference between baseline and noise entropy rates was used to quantify the information transfer rate about the stimulus. This definition is natural in our case as we have baseline firing activity and signal detection must be based on a change from this baseline activity (Ratnam and Nelson 2000, see chapter 5). We compared the information rates obtained with our model and Nelson's. It was shown that the information rate increased in both cases with stimulus contrast. This agrees with the fact that the electroreceptor is better at encoding stronger stimuli (Wessel et al. 1996). However, negative ISI correlations could help the electroreceptor in the coding of fast time-varying stimuli, as information rates computed with our model were higher. This is due to the fact that correlations reduce the noise entropy even more than the baseline entropy. We note that our results are consistent with previous studies that have demonstrated that a refractory period can improve the linear correlation between the stimulus and the instantaneous firing rate (Chialvo et al. 1997) as well as the neural precision (and hence the mutual information) (Berry and Meister 1998). However, the models used in these studies did not incorporate ISI correlations. Furthermore, in the latter paper, the increase in refractory period led to a decrease in mean firing rate; thus the causes of the increase in mutual information were not clear. In our analysis, we examined the effect of ISI correlations on mutual information and signal detection without concomitant changes in the mean firing rate.

Information rates as a function of stimulus cutoff frequency were also computed. Our results show that the information rate increased with cutoff frequency as expected from the high-pass characteristics of P-afferents (Xu et al. 1996). Information rates were higher for the LIFDT model and the gain due to correlations exhibited a resonance with a maximum around 100 Hz. This frequency corresponds to the inverse of the decay time constant of our dynamic threshold. It also corresponds to the maximum in tuning seen in chapter 4 using the mean vector strength. These are potentially related. However, further studies are needed to look at this. Remarkably, there is evidence that these fish respond preferentially to electrocommunication signals with a frequency content of around 100 Hz (Bastian et al. 2001).

6.3.4 Comparing signal detection and information theory

Finally, we have shown that the loss of information about the stimulus incurred by decreasing the spike timing resolution increases with cutoff frequency (Figure 6-13). This suggests that spike timing is not as important for low as it is for high frequency stimuli. Moreover, since entropy measures assume a very long (ideally infinite) spike train, they are not applicable to situations where the animal must make a decision based on a short time window (e.g. when trying to decide whether a prey is present or not). However, signal detection theory assumes a rate code and can be applied in cases where spike timing is not important. Our results show that an optimal detector receiving a single P-unit spike train with baseline negative ISI correlations can detect the presence of stimuli that would give rise to as few as two extra spikes over a counting time of approximately 250 ms as suggested by Ratnam and Nelson (2000); this means two extra spikes out of approximately 50 during that time. For low-frequency stimuli, the assumptions made in the signal detection analysis are appropriate because the change in mean firing rate (computed over an appropriate time window) without a concomitant change in firing rate variance can signal the presence/absence of the stimulus. However, for zero-mean high frequency stimuli, there may be almost no change in mean firing rate computed over a long time window, hence the signal detection analysis is not appropriate. It is then more appropriate to use information theory. Due to the high-pass filtering characteristics, high-frequency stimuli will elicit almost no trial-to-trial variability, hence the noise entropy rate will be low. The mutual information rate will then be high.

The ideas presented above are compatible with a recent analysis by Salinas and Sejnowksi (2000) in which neurons can be driven either by the mean excitatory level (mean firing rate as assessed by the PND in our case) or by fluctuations around this mean (as assessed by mutual information in our case). Hence the high variability of P-afferent spike trains gives a high mutual information rate when looking at high frequency stimuli. On the other hand, negative ISI correlations at short lags improves signal detectability for low-frequency stimuli.

6.3.5 Conclusion/Outlook

We have shown that negative ISI correlations seen experimentally can improve a neuron's ability to code **both** slow and fast time-varying stimuli. There is still theoretical work to be done in order to look at the mechanism by which ISI correlations increase information transfer. This involves looking at information transfer in simplified models (i.e. without an EOD or the Nelson filter). Some of this work has been carried out (Longtin et al. 2003). Moreover, problems are being worked out in our group.

The P-units we have studied are known to converge onto basilar pyramidal cells of the electrosensory lateral line lobe (ELL, Bastian 1981). Population averaging is thus expected and might explain the extreme behavioural sensitivity to AM's, down to 0.1% of baseline EOD (Knudsen 1976, Nelson and MacIver 1999). The position of the minimum of the Fano factor is highly variable in experiments (40-1000 ms, Ratnam and Nelson, 2000). Moreover, it is possible that different P-units (probability of firing per EOD cycle ranges from 0.1 to 0.5: Nelson et al. 1997) will have different cumulative relative refractoriness decay time constants. It will thus be very interesting to study ELL decoding of slowly vs. rapidly time-varying input processed by this heterogeneous electroreceptor population.

7 To burst or not to burst?

In this chapter, we are interested in the consequences of neural bursting on information transfer. An important question is why this weakly electric fish might have two kinds of electroreceptors: ones that do not burst and ones that do. Thus, we are addressing the possible advantages of having a heterogeneous electroreceptor population.

7.1 Introduction: function of bursting in sensory processing

This problem is not only related to weakly electric fish, but also to higher order primates. Indeed, it has been shown that thalamo-cortical neurons can fire tonically or burst (Sherman 2001). It was further proposed that each mode of firing would correspond to a different function. An interesting study of electric fish pyramidal neurons (to which electroreceptors provide input) shows that they are incapable of correctly estimating a time-varying stimulus yet can extract certain features through bursts of action potentials (Gabbiani et al. 1996). However, results on Lateral Geniculate Nucleus neurons (Reinagel et al., 1999) show that bursts encode less information per spike than single spike firing. Thus, how is information transferred by a bursting neuron? In order to answer this question: we compare the information transfer properties of a bursting neuron (B) to those of a non-bursting neuron (NB).

7.2 Calibration

We use the BLIFDT model described in chapter 3 under two different parameter regimes. We use the following calibration.

7.2.1 Calibration for baseline activity

ISI histograms for baseline activity (EOD only) are shown in Figure 7-1. Both show modes around multiples of the EOD period characteristic of phase-locking in these units and are similar to representative units of both classes (Xu et al. 1996). NB (Figure 7-1a) tends to fire single spikes but skip a variable number of EOD cycles between each firing. On the other hand, B has a high peak at one EOD cycle and has a lower peak at 6 EOD cycles (Figure 7-1b): this implies that the unit tends to fire action potentials on consecutive EOD cycles (bursts) followed by a variable period of quiescence (Xu et al. 1996; see also chapter 3). Parameters for the model were chosen such that both units have essentially the same firing rate for baseline activity. However, the inter-spike interval histogram for B has a higher coefficient of variation (the standard deviation to mean ratio) than the ISI histogram for NB (Figure 7-1). The inter-spike interval serial correlation coefficients are shown in Figure 7-1c and Figure 7-1d for NB and B respectively. Non-bursting P-units typically have a negative serial correlation coefficient at lag one. On the other hand, bursting P-units show correlation coefficients that alternate in sign and decay over a few lags.

7.2.2 Calibration for time-varying input

To further ensure a fair comparison for time-varying input, the gain and phase response curves (see chapter 6, Figure 6-3) were calculated for both models for sinusoidal amplitude modulations of the EOD with frequencies ranging from 0.1 Hz to 200 Hz (Figure 7-2a and Figure 7-2b). The values obtained for N and NB agree well and further lie in the physiological range (Nelson et al. 1997). The frequency of firing is approximately linearly related to the EOD amplitude (Figure 7-2c) over the entire range used in this study and again the curves obtained for each model neuron agree well.

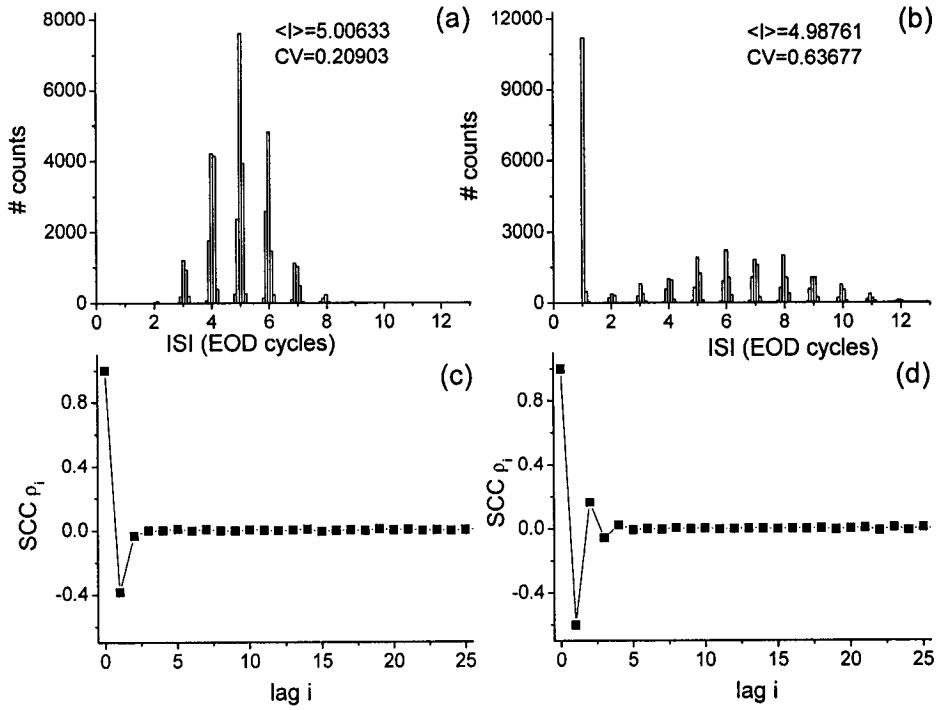


Figure 7-1: Baseline activity obtained for both model neurons. (a) Unimodal ISI histogram obtained for NB with mean ISI $\langle I \rangle$ and coefficient of variation CV defined as the standard deviation-to-mean ratio. (b) Bimodal ISI histogram obtained for B with $\langle I \rangle$ and CV. Note that $\langle I \rangle$ is approximately the same for NB and B. B is however more variable than NB as indicated by the higher CV value. (c) ISI correlation coefficients as a function of lag for NB. The coefficient at lag one is negative as is typically seen experimentally in these units. (d) ISI correlation coefficients as a function of lag for B. These coefficients decay over a few lags and alternate in sign as is seen experimentally. Parameter values for NB are the same as in Figure 3-4 with $\Delta I_b=0$. Parameter values for B are the same except for: $\tau_\theta=9.2$ ms, $D_1=4$ ms^{-1}, $D_2=1.6\times10^{-5}$ ms^{-1}, $\Delta I_b=1.5$, $\tau_b=0.09$ ms, $d=0.4$ ms.

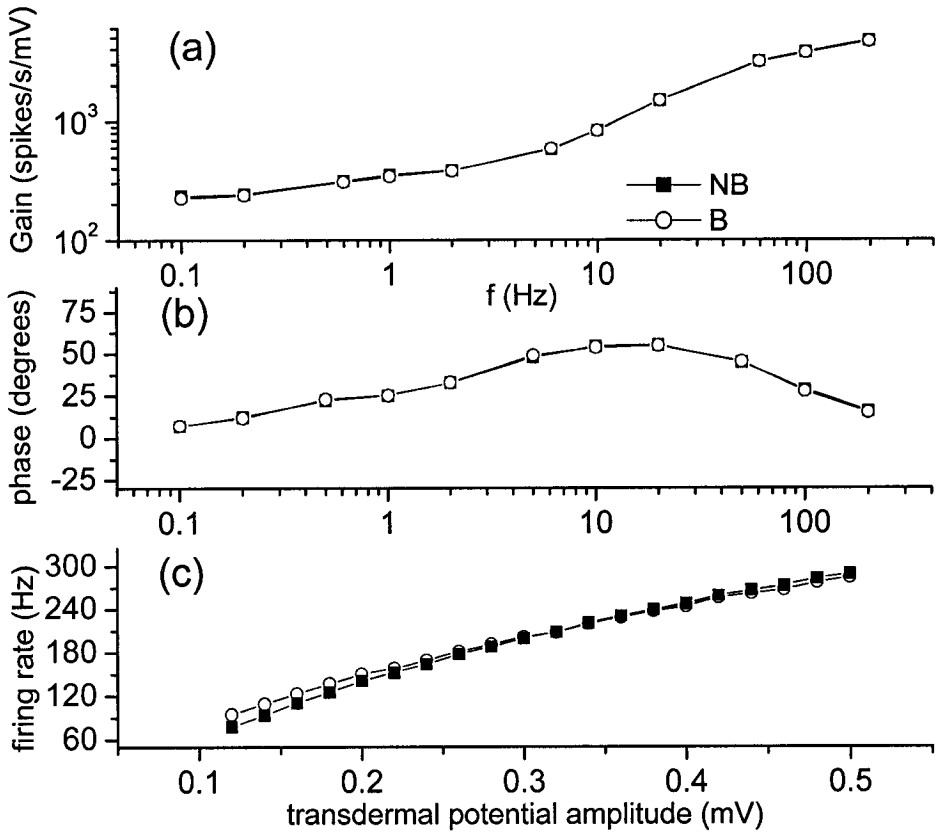


Figure 7-2: Calibration for both models as a function of stimulus frequency and contrast. (a) Gain as a function of the frequency of the sinusoidal amplitude modulation for both units. The curves are similar to those obtained experimentally and the values are in the physiological range. (b) Phase difference between the stimulus and the response as a function of frequency. The curve is again quantitatively similar to those obtained experimentally. (c) Firing rate as a function of transdermal potential amplitude for both units. The curves are approximately linear and equal over the entire range of amplitude modulations used in this study.

7.3 Results

We use signal detection theory (described in chapter 2), indirect methods of estimating information transfer (described in chapter 5), and direct methods of estimating information transfer (also described in chapter 5) as well as feature detection measures in order to compare both models.

7.3.1 Signal detection theory

We have shown in chapter 6 that spike train variability (as measured by the Fano factor F(T) defined in chapter 2 was minimal at a counting time varying between 40-1000 ms. It was further shown that this decreased variability was a consequence of negative ISI correlation coefficients as the reduction depends on the sum of the correlation coefficients. Negative correlation coefficients thus reduce spike train variability while positive correlation coefficients increase it. The Fano factor curves for both model neurons are shown in Figure 7-3. Parameters were chosen such that both curves have their minima for a counting time of 250 ms. However, the minimum value is less for NB than for B. This can be explained as follows.

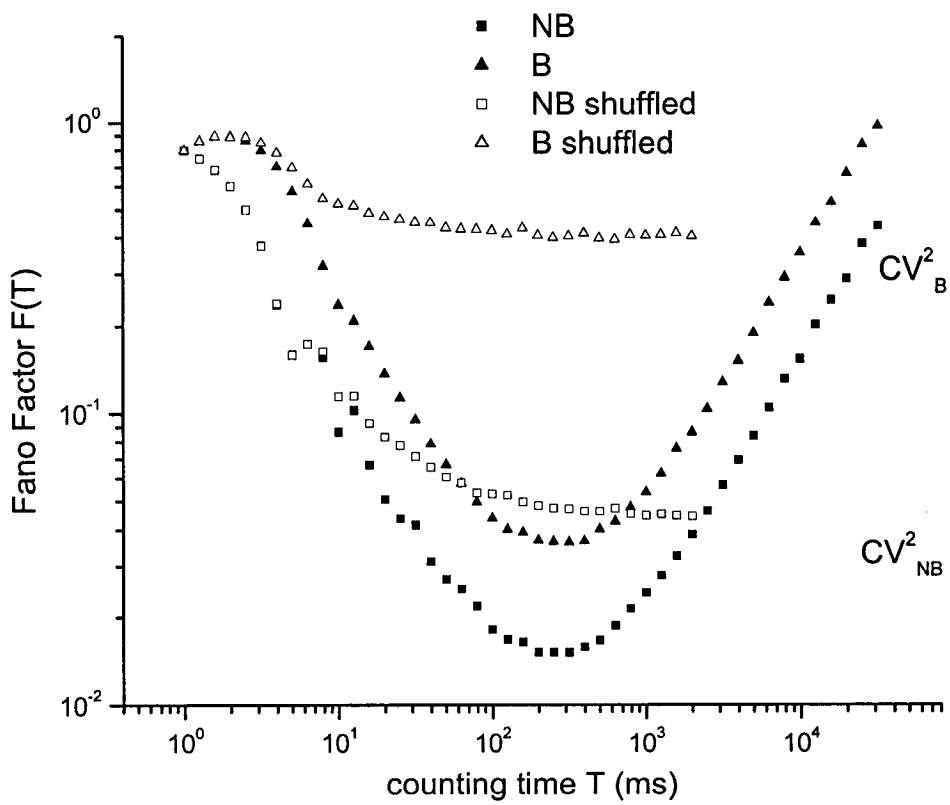


Figure 7-3: Fano factor time curves obtained from baseline activity for both neurons in the presence and absence of ISI correlations. In the absence of ISI correlations, the Fano factor for B is higher than for NB because of the higher coefficient of variation.

The negative ISI correlations reduce the Fano factor for both units. However, the Fano factor minimum is still greater for B than for NB.

It is possible to eliminate ISI correlations by randomly shuffling the ISI sequence. The Fano factor curve obtained for shuffled ISI sequences for both model neurons during baseline activity are also shown in Figure 7-3. As we now have a renewal process in each case, the Fano factor now tends asymptotically towards the square of the coefficient of variation for both models (Cox and Lewis, 1966). However, since the coefficient of variation is higher for B than for NB (Figure 7-1), the Fano factor is thus consequently higher for B in the absence of correlations.

For NB, the negative ISI correlation coefficient decreases the Fano factor for low counting times (Figure 7-3). However, the effects of negative correlation coefficients will be partly cancelled by positive correlation coefficients for B, hence the Fano factor will not decrease by as much. The higher coefficient of variation (calculated from first order ISI statistics) and the positive correlation coefficients (second order ISI statistics) thus contribute to make spike train variability higher for B, hence the minimum value of the Fano factor is greater.

To test both units, we used a portion of a weak (15 μV) low cutoff frequency (2 Hz) stimulus that is similar to that perceived by P-afferents when the fish passes by an isolated prey in mid-water (Nelson and MacIver 1999). The portion used in shown in Figure 7-4a. Detection of signals must in this case presumably be based on a change from baseline activity. We thus calculated the spike count distributions obtained in the absence and presence of this stimulus. Figure 7-4 shows the results: the variances of the spike count distributions for NB are lower than the variances obtained for B. As explained above, this is due to the fact that the Fano factor is lower for NB. Consequently, the signal-to-noise ratio (SNR, see chapter 2) is higher for NB than for B (Figure 7-4b and Figure 7-4c). Our results thus show that NB is more apt at discriminating stimuli than B.

We next use information theory (described in chapter 5) to extend our analysis to more complex stimuli and to look at the effects of spike timing.

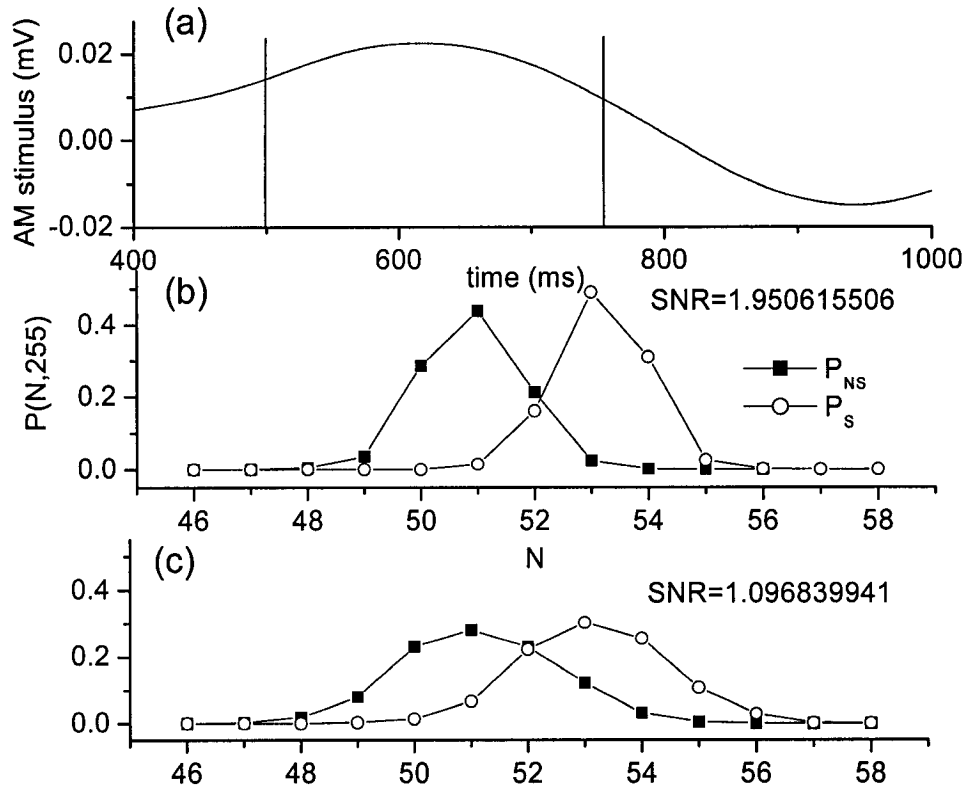


Figure 7-4: Signal detection performance of both models. (a) Stimulus (i.e. the transdermal potential amplitude minus its baseline value). The portion of stimulus used to calculate the discriminability d is between the two vertical lines. (b) Spike count distributions obtained in the absence (P_{NS}) and presence (P_S) of stimulus for NB. The distributions have a low variance and thus a small amount of overlap. The SNR is thus high. (c) Spike count distributions obtained for B. The distributions have higher variances and thus d is reduced due to a greater amount of overlap between the two distributions. Note that the distributions obtained for B and NB in the absence of stimulus have the same mean. This is also the case for the distributions obtained in the presence of stimulus.

7.3.2 Information theory

We will compare results obtained using both the direct and indirect methods of estimation of information transfer. However, we will first compare results obtained using stimulus estimation techniques (as described in chapter 5).

7.3.2.1 *Stimulus estimation (through the indirect method)*

The coding fractions obtained for both units are shown in Figure 7-5. The coding fraction is seen to increase with stimulus contrast for constant cutoff frequency. Stronger stimuli reduce the intrinsic variability of the neuron and thus allow for a better estimate of the stimulus. Our results are consistent with those of Wessel et al. (1996) who found the same trend in an experimental study of a related electric fish. The coding fraction for the bursting unit was lower than for NB. The difference between the two is maximal for low contrast. This result is consistent with the results obtained using signal detection theory. At a constant contrast, the coding fraction estimates obtained for B and NB decrease as a function of cutoff frequency (Figure 7-5c and Figure 7-5d). Again our results are consistent with those obtained experimentally by Wessel et al. (1996). Thus, the extra variability of the bursting unit is detrimental to stimulus reconstruction over the entire frequency and contrast range.

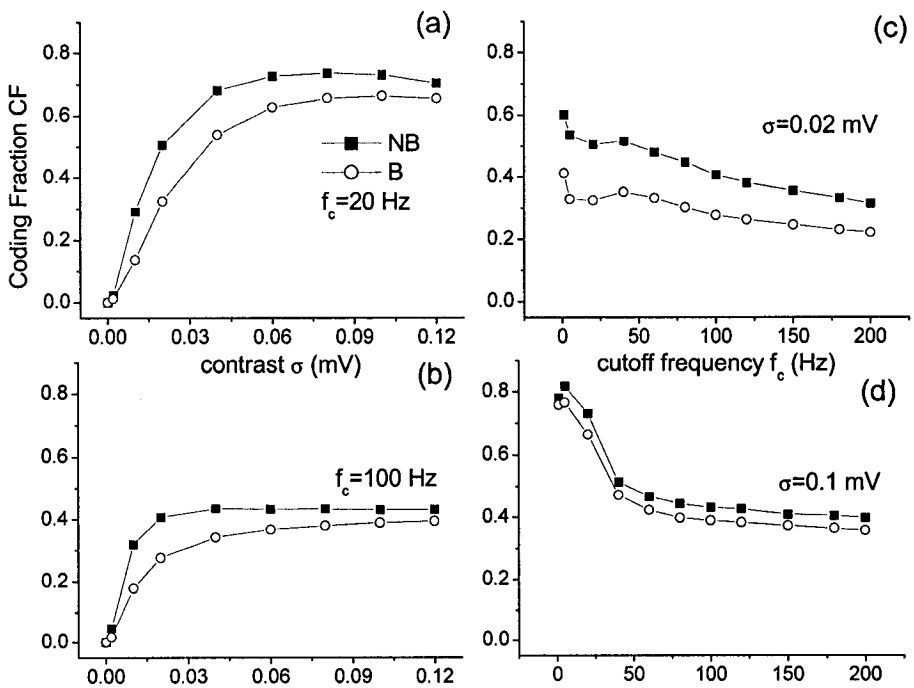


Figure 7-5: Effects of stimulus contrast and cutoff frequency on coding fraction. Coding fraction as a function of stimulus contrast for (a) a 20 Hz cutoff frequency and

(b) for a 100 Hz cutoff frequency. Coding fraction as a function of cutoff frequency for (c) a 0.02 mV contrast and (d) for a 0.1 mV contrast. Note that the coding fraction for NB is always higher than for B. Coding fraction increased with contrast and decreased with cutoff frequency.

7.3.2.2 Comparison between the direct method and the indirect method

We first use the measure I_{stim} obtained by the direct method. We show results obtained as a function of contrast for two different cutoff frequencies in Figure 7-6a and Figure 7-6b respectively. For low cutoff frequencies, NB transmits more information than B over the contrast range studied. These results are consistent with the ones obtained using stimulus reconstruction and signal detection theory. However, for a high cutoff frequency, B transmits as much information as NB for high contrasts, in contrast to what is predicted from the stimulus reconstruction technique.

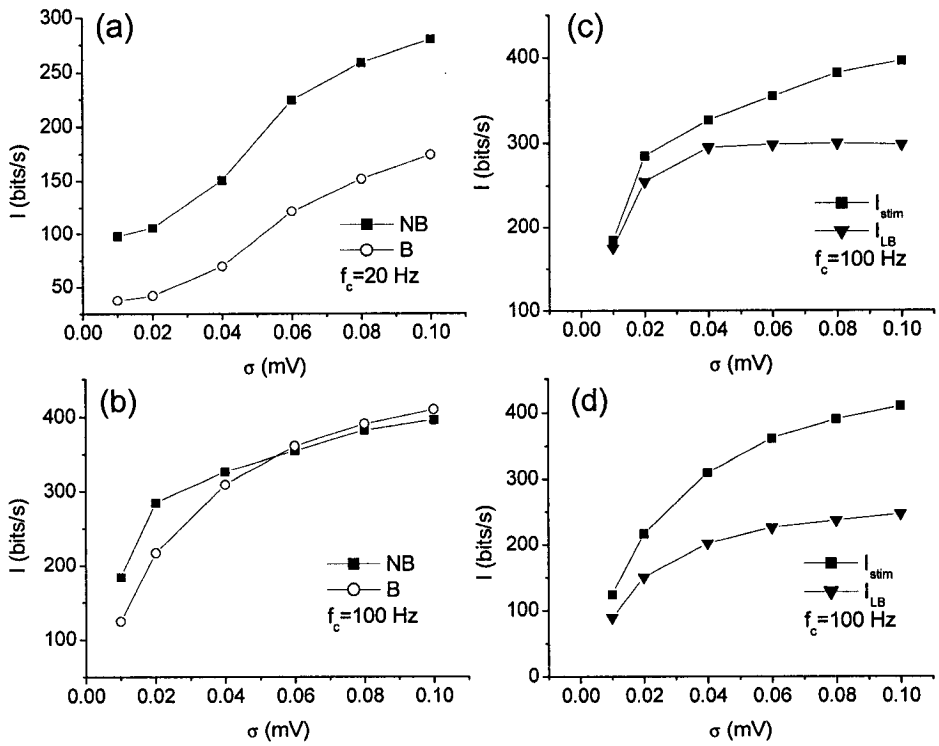


Figure 7-6: Mutual information rate as a function of contrast. (a) Mutual information rate estimate I_{stim} as a function of contrast for a 20 Hz cutoff frequency. Note that the estimate is always higher for NB than for B. (b) Mutual information rate estimate I_{stim} as a function of contrast for a 100 Hz cutoff frequency. In this case, the two information rates have similar values for high contrasts. (c) Comparison for NB of the estimates I_{stim} and I_{LB} obtained as a function of contrast and for a 100 Hz cutoff frequency. The estimate I_{LB} is at least equal to 75% of the estimate I_{stim}. (d) comparison for B of the estimates I_{stim} and I_{LB} as a function of contrast and for a 100 Hz cutoff frequency. Note that the estimate I_{LB} is only equal to at least 50% of the estimate I_{stim}.

In order to explore this discrepancy, the information rate estimates I_{LB} (described in chapter 5) and I_{stim} (also described in chapter 5) are shown for NB (Figure 7-6c) and B (Figure 7-6d). By comparing the two figures, one can see that the estimates I_{low} obtained for NB are always higher than the respective estimates obtained for B, as expected from the results obtained using stimulus estimation and signal detection theory. A comparison with the direct method estimate I_{stim} shows that the estimate I_{LB} for NB captures at least 75% of the total information rate. In contrast, the estimate I_{LB} for B captures only at least 50% of the available information for high cutoff frequencies. Our results are consistent with those of Buracas et al. (1998) that

show a difference between the direct and indirect methods of estimating information rates. This difference comes from the fact that the assumptions of the indirect method are less valid for a bursting neuron than for a non-bursting one. This implies that bursting neurons encode different aspects of the stimulus than non-bursting neurons.

Information rate estimates calculated using the direct method as a function of cutoff frequency are shown in Figure 7-7a and Figure 7-7b respectively for a low contrast and a high contrast. As expected, the estimate obtained for NB is higher than the estimate obtained for B for a low stimulus contrast (Figure 7-7a). However, both estimates are nearly equal with B sometimes transmitting more information than NB for a high contrast and in the high cutoff frequency range (Figure 7-7b). This is again contrary to what is expected from stimulus reconstruction. As before, we compare the estimate using the direct method I_{stim} to the estimate I_{LB} obtained from the indirect method for both NB and B for a low contrast respectively in Figure 7-7c and Figure 7-7d. For NB, we see that the estimate I_{LB} is very close to I_{stim} over the entire frequency range. This permits us to infer that NB encodes aspects of the stimulus that are captured by linear stimulus reconstruction for low contrasts. It is known that P-receptor afferents modulate their firing rate with the EOD amplitude; the receptor thus modulates its mean firing rate with the time-varying AM. Its low variability allows a good estimate of the mean firing rate over a relatively short time window, hence allowing a good reconstruction. However, the situation for B (Figure 7-7d) is different with I_{LB} lower than I_{stim} by as much as a factor of two. The bursting unit thus encodes information about aspects of the stimulus captured by linear stimulus reconstruction. We have thus shown that a bursting neuron performs poorly at stimulus reconstruction as compared to a non-bursting neuron. This is because of the intrinsic bursting dynamics.

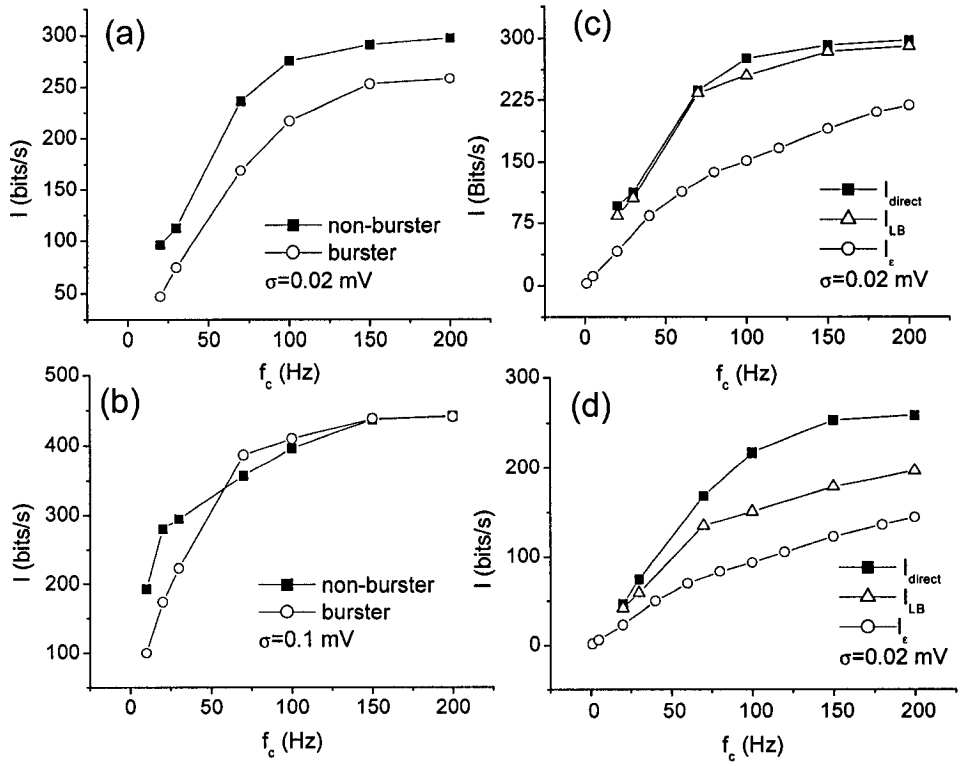


Figure 7-7: Mutual information rate as a function of cutoff frequency. (a) Mutual information rate estimate I_{stim} as a function of cutoff frequency for a 0.02 mV contrast. Note that the estimate obtained for NB is higher than the one obtained for B. (b) Mutual information rate estimate I_{stim} as a function of cutoff frequency for a 0.1 mV contrast. Note that both estimates have similar values for high cutoff frequencies. (c) Comparison for NB of the estimates I_{stim} and I_{LB} obtained as a function of cutoff frequency and for a 0.02 mV contrast. Note that I_{LB} and I_{stim} are approximately equal over the entire frequency range. (d) Comparison for B of the estimates I_{stim} and I_{LB} obtained as a function of cutoff frequency and for a 0.02 mV contrast. Note that the estimate I_{stim} is always significantly larger than I_{LB}.

7.3.2.3 Feature detection

In order to understand what are the differences in the encoding strategies of bursting and non-bursting neurons, we look at feature extraction by both B and NB.

7.3.2.3.1 Definition of a feature

By a feature, we mean a lapse of time during which the time derivative of the stimulus was positive (i.e. the rising phase). It is known that electroreceptor afferents preferentially emit action potentials during these phases (Bastian, 1981; Nelson et al., 1997). Furthermore, the pyramidal neurons to which these electroreceptors synapse respond preferentially to these rising phases.

7.3.2.3.2 Definition of the "κ" measure

We wish to use a simple measure of feature detection. We define κ_1 as the fraction of bursts that occur on a rising phase (i.e. the number of bursts that occurred during an rising phase divided by the total number of bursts). κ_1 is a measure of correlation between bursts and rising phases. However, we must also address the issue of reliability (i.e. whether most rising phases are associated to at least one burst). We hence define κ_2 as the fraction of rising phases during which at least one burst occurred (i.e. the number of rising phases during which at least one burst occurred divided by the total number of rising phases). We define κ as a measure of feature detection and encoding by bursts by:

$$\kappa = \kappa_1 \kappa_2 \tag{6.5}$$

7.3.2.3.3 Results

We calculated κ for both units as a function of stimulus cutoff frequency and contrast. The results are shown in Figure 7-8. For a low contrast (Figure 7-8a), B significantly outperforms NB over the entire frequency range. For a higher contrast (Figure 7-8b), B still outperforms NB but the difference is lessened. Looking at the measure κ_1 for B, we see that it is equal to 0.5 for low cutoff frequencies (Figure 7-8a and Figure 7-8b). It however increases and becomes nearly one for high cutoff frequencies. There is thus a gradual transition from

disorder (the bursts occur randomly) to order (the bursts occur almost exclusively on the rising phase of the stimulus). This is due to a competition between the stimulus and the intrinsic burst mechanism of B. As shown in Figure 7-8a and Figure 7-8b, reliability κ_2 is not an issue for B except at high cutoff frequencies where the measures κ and κ_1 are different.

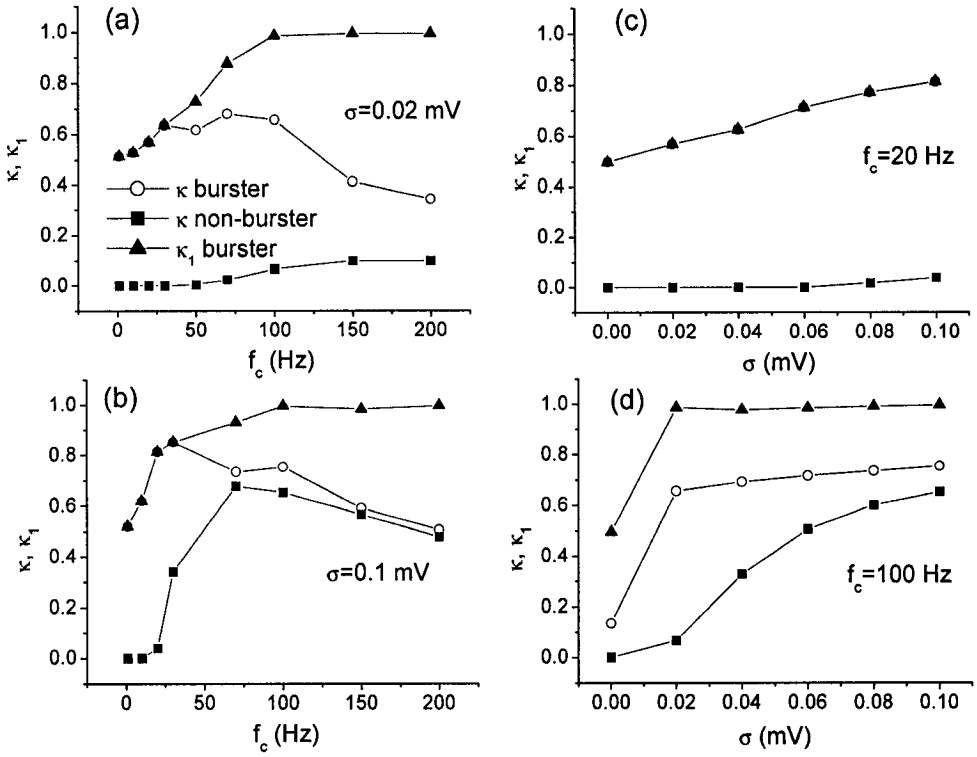


Figure 7-8: Feature detection performance of both models. (a) κ and κ_1 as a function of cutoff frequency for a 0.02 mV contrast. (b) κ and κ_1 as a function of cutoff frequency for a 0.1 mV contrast. (c) κ and κ_1 as a function of contrast for a 20 Hz cutoff frequency. (d) κ and κ_1 as a function of contrast for a 100 Hz cutoff frequency. Note the transition from a disordered state where bursts occur randomly on either the rising or falling phase of the stimulus to an ordered state where bursts occur preferentially on the rising phase in B.

The situation is very different for NB. Bursting can occur but only for sufficiently high contrast and cutoff frequency values: there is thus a "burst threshold" below which there is no bursting. These bursts are purely driven by the stimulus and are thus perfectly correlated to it.

The measure κ_1 is thus always equal to one (not shown) and thus $\kappa=\kappa_2$ is a measure of reliability. Reliability is very much an issue for NB as indicated by the low κ values obtained over the entire frequency range for low contrast (Figure 7-8a). The reliability increases with a higher contrast (Figure 7-8b) and approaches that of B for high cutoff frequencies. We are then in a regime where the stimulus dominates: both neurons thus perform about the same.

This also explains the results obtained for a low cutoff frequency (Figure 7-8c) where B again clearly outperforms NB over the entire contrast range. For a higher cutoff frequency (Figure 7-8d), B still outperforms NB but the difference is lessened as we are again in a regime were the stimulus dominates the dynamics.

Thus, B is always better at encoding features of the stimulus than NB, even in regimes where it is possible to get evoked bursting from NB. This is because the reliability of B is always greater or equal than that of NB and because a transition occurs in B in which the bursts occur preferentially of the rising phase of the stimulus.

7.3.2.4 Comparison of different measures obtained using the direct method

Finally, we compare the measure I_{stim} obtained from one long presentation of the stimulus to the measure I_{spon} based on spontaneous activity presented in chapter 5. Results are presented in Figure 7-9. The information rate I_{spon} is seen to increase as a function of contrast for both low (Figure 7-9a) and high (Figure 7-9b) cutoff frequencies and higher for NB at low cutoff frequencies. However, it is higher for B at higher cutoff frequencies. As a function of cutoff frequency and for a low contrast (Figure 7-9c), we see that the information I_{spon} is higher for NB for low frequencies and higher for B at high frequencies. The results for a higher contrast are qualitatively similar (Figure 7-9d).

These results are very different than those obtained using the direct method with I_{stim} where the information rate calculated for NB always exceeded the information rate calculated for B. This is due to the fact that spontaneous activity for B is more variable than baseline activity for NB. Hence $H_{spon}(R)$ for B is higher than $H_{spon}(R)$ for NB. Consequently, H(R/S) for B is always higher than for NB. However, H(R/S) is reduced as contrast and/or cutoff frequency

increases. For high contrasts and/or cutoff frequencies, both units have comparable noise entropies. However, since $H_{spon}(R)$ is much higher for B than for NB, the mutual information is higher for B.

Thus, our results suggest that non-bursting P-receptor afferents are involved in the coding of low frequency stimuli which correspond e.g. to signals from prey and from interactions between members of the same sex. Bursting P-receptor afferents are maybe involved in the coding of high frequency stimuli that correspond to interactions between members of the opposite sex and electrocommunication signals such as chirps. We thus hypothesize that the animal must have two kinds of electroreceptors in order to deal with both prey and communication stimuli.

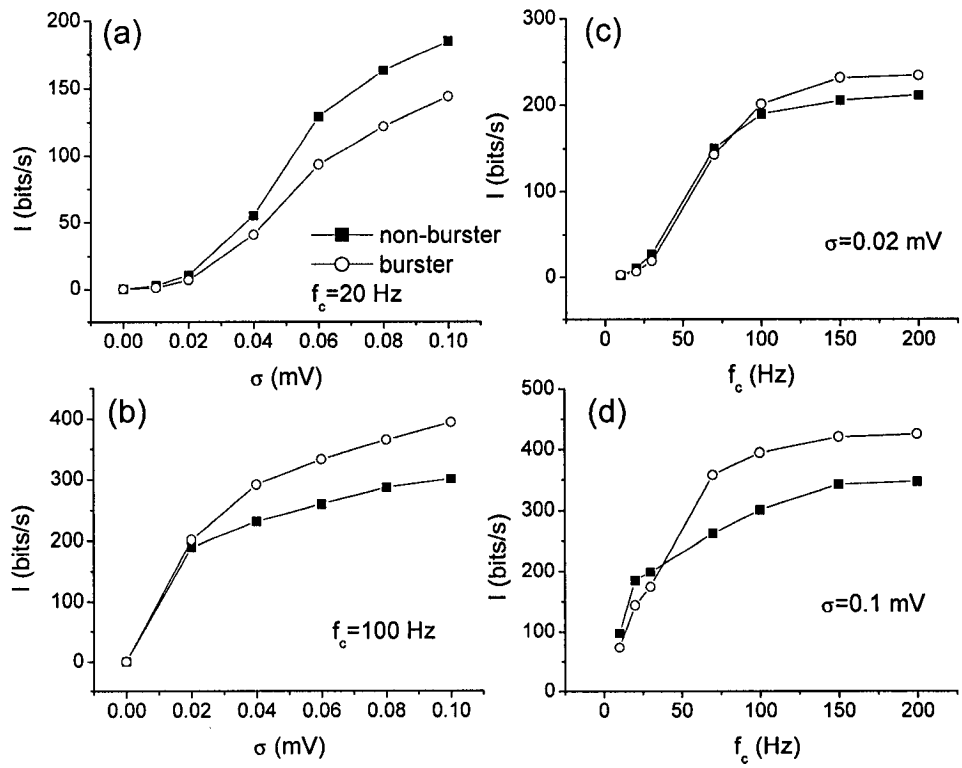


Figure 7-9: Mutual information rate I_{spon}. (a) I as a function of contrast for a 20 Hz cutoff frequency. The estimate obtained for NB is higher than the estimate obtained for B. (b) I_{spon} as a function of contrast for a 100 Hz cutoff frequency. Note that it is now the estimate obtained for B that is higher than the estimate obtained for NB. (c)

I_{spon} as a function of cutoff frequency for a 0.02 mV contrast. (d) I_{spon} as a function of cutoff frequency for a 0.1 mV contrast. For panels (c) and (d), for low cutoff frequencies, the estimate obtained for NB is higher than the estimate obtained for B. For high cutoff frequencies, the opposite occurs.

7.4 Discussion

7.4.1 Calibration

We have compared results obtained from two models of P-receptor afferents: one was bursting intrinsically while the other did not. Parameter values were chosen such that the two units had the same firing rates. Moreover, the gain and phase values for sinusoidal modulations of the EOD amplitude over the range 0.1-200 Hz were essentially the same for each model. This helped ensure that the mean firing rate response was the same for both models over the frequency range of stimuli used in this study. Hence the only difference between the two models was the fact that one tended to fire packets of action potentials (bursts) and thus had a higher variability than the other.

7.4.2 Signal detection theory

Using signal detection theory and using a portion of stimulus akin to what the animal might get from passing by a prey, we have shown that the extra variability due to bursting was detrimental to detection of these weak slowly-varying signals. This was due to the higher coefficient of variation for the bursting unit but also to ISI correlation coefficients that alternated in sign. These alternating ISI correlations have been observed in other periphery neurons (Schäfer et al. 1995; Bahar et al. 2001). As the positive coefficients somewhat cancelled the variability reducing effects of negative ISI correlation coefficients, the variance of the spike count distributions obtained in the presence and absence of stimulus for the bursting unit was thus higher. This led to a greater overlap between the two distributions and thus a lower SNR. For the two units considered here, the SNR was two times greater for NB than for B. Our results thus suggest that the function of these bursting neurons is not to detect signals based on their firing rate. On the contrary, non-bursting units typically only have negative

correlation coefficients that reduce spike train variability and thus increase discriminability as shown in chapter 6. These results have implications for animal behavior in that we can make a prediction that non-bursting electroreceptors are primarily used for prey detection by weakly electric fish.

7.4.3 Information theory

7.4.3.1 Stimulus estimation (indirect method)

To look at the issue of encoding a stimulus rather than detecting it, we used the stimulus reconstruction technique. The coding fraction increased with stimulus contrast and decreased with cutoff frequency for both units. Our results thus agreed with those reported in other studies (Wessel et al. 1996; Kreiman et al. 2000). However, the coding fraction obtained for NB was higher than the one obtained for the bursting unit over the entire frequency and contrast range used for this study. This agreed with our results using signal detection theory and other results on burst vs. tonic modes in thalamocortical neurons (Reinagel et al. 1999; Sherman 2001). This further agreed with the fact that bursting neurons are poor stimulus estimators (Gabbiani et al. 1996) due to their higher variability.

7.4.3.2 Information theory: direct vs indirect method

The indirect method gave us the estimate I_{LB} of the information rate transmitted by the neuron about the stimulus. We found that I_{LB} was always greater for NB than for B over the entire frequency and contrast range, once again validating our results up to date. It is however important to gauge the goodness of the lower bound estimate I_{LB} by comparing it with the estimate I_{stim} obtained using the direct method (Borst and Theunissen 1999). For NB, we found that the two estimates agreed for low stimulus contrast. This implies that linear stimulus reconstruction captures all the aspects of the stimulus that are encoded by the neuron at these low contrasts. For higher contrasts however, the estimate I_{stim} was greater than I_{LB} indicating that other aspects of the stimulus are being encoded as well. The estimate I_{LB} was at least equal

to 75% of the estimate I_{stim}, indicating that linear stimulus reconstruction captured most of the information. This implies that the features of the stimulus encoded by the non-bursting P-afferent are pretty well captured by linear stimulus reconstruction. Linear stimulus reconstruction has been shown to be valid for neurons in other sensory systems (Rieke et al. 1997).

For the bursting unit however, the estimate I_{low} was always significantly lower than I_{stim}, indicating that other aspects of the stimulus are being encoded. This difference was shown to increase with stimulus contrast and cutoff frequency where the estimate I_{stim} was twice greater than I_{LB}. This implies that the bursting unit encodes aspects of the stimulus not captured by linear stimulus reconstruction. Such a systematic comparison of information rates obtained using both the direct and indirect methods as a function of stimulus contrast and cutoff frequency has, to our knowledge, never been made.

By comparing the estimates I_{stim} obtained for both B and NB, we found that NB transmitted more information at low frequencies than B. This agreed with previous results. However, the two estimates were approximately equal for high frequency stimuli. This is in contradiction with results obtained using the indirect method and shows that one should always compare estimates using both methods to ensure validity. Such a comparison has been done in visual neurons (Buracas et al., 1998) and it was found that the estimate I_{stim} was significantly greater than I_{low}. Thus, bursting P-receptor afferents transmit just as much information as non-bursting ones for high frequency stimuli. They however use a different coding strategy.

7.4.3.3 Feature detection

To explore the strategy, we used a measure of efficiency in signaling particular features of the stimulus. Our results showed that the bursting P-receptors afferent had a significantly greater efficiency than NB. In the absence of stimulus, bursts occur through the intrinsic burst mechanism of B only. This mechanism is not correlated to the stimulus and as thus bursts do not occur preferentially on the rising phase of the stimulus. However, as stimulus contrast and/or cutoff frequency increases, the stimulus's influence on the probability of bursting increases and bursts become correlated to the stimulus (i.e. they tend to occur on the rising

phases of the stimulus where the synaptic current is greatest). There is thus a transition from disorder to order in the presence of a stimulus. The probability of bursting is high and reliability is not an issue except for very high frequencies. It is possible to get evoked bursting from the intrinsically non-bursting P-receptor afferent only if the stimulus is strong enough. Such bursting is completely driven by the stimulus and is completely correlated to it. However, the probability of bursting is not as high for the intrinsically bursting P-afferent. B was clearly better at reliable feature extraction than NB.

7.4.3.4 Direct Method: I_{stim} vs I_{spon}

We calculated the mutual information I_{spon} for N and NB. We found that NB performed better for low frequency stimuli and B better for high frequency stimuli. This has implications for animal behaviour, as the fish must detect stimuli from a wide frequency range. High frequency stimuli are global (they come from other fish) and do not require precise spatial location. We thus hypothesize that the bursting P-receptor afferents would be mostly located on the fish's trunk as they would only be used to detect high frequency global stimuli. A previous study (Xu et al. 1996) reported that 31% of P-receptor afferents could be classified as bursters. That study however considered afferents from the trunk region. Such a study remains to be undertaken for the head region which has an overall higher density of electroreceptor neurons.

7.4.4 Conclusion/ Outlook

In conclusion, we have compared results from two phenomenological accurate models of P-receptor afferents of weakly electric fish. We have shown that extra variability due to bursting was detrimental to signal detection, stimulus reconstruction, and information transfer of low-frequency weak signals. The information estimates obtained by the indirect method underestimated the information transmitted by bursters. This information transmission is partly due to feature extraction. Finally, using a measure of information more appropriate in the case of spontaneous activity, we have shown that non-bursters encoded more information about low-frequency stimuli and bursters encoded more information about high-frequency

stimuli. These results provide an example in which a bursting neuron can perform just as well as a non-bursting neuron and even out-performs it in some cases.

It will be interesting to experimentally verify the predictions made in this chapter. If they are correct, they would show that the animal has evolved highly developed neural circuitry in order to interact with its environment.

8 Conclusion

8.1 Summary of Results

8.1.1 Models of electroreceptor neurons

After reviewing previous models of electroreceptor neurons, we presented a new model that was capable of reproducing first and second order ISI statistics from experimental data: the LIFDT model. This model provided a simple extension to the LIF model by making the spike threshold dynamic. The spike threshold allowed memory to propagate from one firing to the next, hereby giving rise to correlations (second order ISI statistics) as seen in experimental data. An extension of the LIFDT model, the BLIFDT model, was furthermore shown to reproduce the bursting dynamics seen in some electroreceptor neurons. The BLIFDT model could thus reproduce a range of ISI statistics caused by heterogeneities present in electroreceptor neurons.

8.1.2 Non-linear dynamics exhibited by the LIFDT model

We looked at a simplified LIFDT model in order to gain deeper understanding of the dynamics. We introduced a parameter α that controlled the amount of memory in the system. We used simple inputs: constant current, step currents, noisy current, and sinusoidal current. Under constant current, the model displayed stable periodic firing like the LIF model. Furthermore, the firing frequency was shown to remain finite for $\alpha>1$ even for large input

currents as is seen experimentally. For $\alpha \leq 1$, the firing frequency diverged as a function of input current as in the LIF model.

The model displayed adaptation (i.e. a gradual decrease in firing rate over time) in response to step currents. This behaviour is not seen in the LIF model but is observed experimentally. The rate of adaptation was further shown to depend on the memory parameter α.

Under noisy current input, the model displayed negative ISI correlations. Through a perturbation analysis, it was shown that these correlations arise from the intrinsic deterministic dynamics of the model. The noise perturbs the system away from equilibrium and thus produces negative ISI correlations. Analytical expressions for the ISI correlation coefficients were obtained.

Under sinusoidal input, the model was shown to be equivalent to an annulus map. Expressions for the Lyapunov exponents of the system were derived. Numerical estimation of these exponents revealed that the model was capable of chaotic dynamics, unlike the LIF model. Phase-locking behaviour was also observed. Under some approximations, we reduced the annulus map to a circle map. This circle map was however shown to be non-orientation preserving and thus could support chaotic behaviour: this was confirmed through numerical simulations. A different circle map can also describe the standard LIF model under sinusoidal input. However, others (Pakdaman, 2001) have shown that this circle map was orientation preserving and thus could not support chaotic dynamics.

8.1.3 Application of information theory to neural spike trains

We have reviewed the application of the indirect method to neural spike trains. The indirect method can only be applied if the stimulus is Gaussian. Moreover, it only provides a lower bound to the rate of information transfer.

The direct method in theory provides an exact value of the mutual information rate. However, the estimation requires a huge amount of data as well as a priori knowledge of the stimulus ensemble. For this reason, approximations are necessary. We reviewed previous estimates and

introduced other measures. One was based on the Gaussian stimulus ensemble; the other used the fact that some neurons are spontaneously active while others are not. The measure based on spontaneous activity was shown to differ by only a constant from the exact mutual information rate. Another measure taking into account background noise in the environment was also introduced.

Comparison between the different measures revealed that the amount by which they agree or disagree depends on the stimulus as well as the stimulus ensemble considered.

8.1.4 Effect of ISI correlations on information transfer

We studied the effects of the negative ISI correlations displayed by electroreceptor neurons on information transfer by comparing the LIFDT model to the Nelson model. In order to ensure a fair comparison, steps were taken to ensure that the gain and phase responses of each model were the same. Furthermore, each model had nearly identical ISI distributions.

It was found that the LIFDT model did not display correct spike count statistics for large counting times. This was corrected by adding a weak coloured noise to the dynamics. This weak noise had negligible effects on the ISI distribution as well as the dynamics at short time scales. It also produced weak slowly decaying positive ISI correlations. These positive ISI correlations interacted with the strong negative ones produced by the LIFDT model to give a minimum in spike train variability as seen in the experimental data. This was because negative correlations reduce spike train variability while positive ones will increase it. In contrast, the Nelson model produced monotonically decaying spike train variability as a function of counting time that is very close to the one produced by the LIFDT model after random shuffling of the ISI sequence.

Negative ISI correlations improved signal detectability for a step increase as measured by the ROC curve. Furthermore, the LIFDT model provided greater signal detectability for a stimulus similar to that caused by prey for a counting time on the order of 250 ms. This is due to the reduction in spike count variance for both baseline and stimulated activity, causing a lesser overlap between the two distributions and thus leading to greater signal detectability.

Application of information theory to both models confirmed this result. The LIFDT model showed greater information transfer than the Nelson model. The gain in information transfer due to negative ISI correlations was shown to be greatest when the stimulus' cutoff frequency was on the order of the reciprocal of the time constant of the dynamic threshold (~100 Hz). Combined with our results on signal detection, we can assert that negative ISI correlations can thus improve information transfer for high frequency stimuli as well as low frequency stimuli. Results suggested that a rate code would be more appropriate for low frequency stimuli while a timing code would be more appropriate for high-frequency stimuli. It will be interesting to verify these theoretical results experimentally.

8.1.5 To burst or not to burst?

In order to gain deeper understanding why some electroreceptor neurons display non-bursting dynamics and some do, we compared information transfer from the BLIFDT model in non-bursting mode (NB) to the BLIFDT in bursting mode (B). Steps were again taken to ensure that both models displayed similar gain and phase responses as well as nearly equal firing rates.

The first comparison was done using signal detection theory. It was found that NB displayed better signal detectability than B. This was due to the greater spike train variability displayed by B, due to a higher ISI variance, as well as to positive ISI correlations at even lags that somewhat cancelled the beneficial effects of negative ISI correlations at odd lags.

The next comparison used stimulus estimation techniques. It was found that NB performed much better than B over a wide range of stimulus contrast and frequency content. The results agreed well with those obtained using signal detection theory.

Information theory was then used. We compared results obtained using both the direct and indirect methods. It was found that the indirect method captured about 75% of the available information for NB while it only captured 50% for B. However, mutual information rates obtained for NB and B using the direct method were similar for the most part. It was thus seen that B encoded information in a more non-linear fashion than NB.

In order to get some insight into how B encoded information, we constructed a feature detection measure and compared the performance of B and NB. It was found that B performed overall as a much better feature detection than NB when the feature was an upstroke in the stimulus.

Finally, we compared information rates based on the direct method in conjunction with the spontaneous activity for both models (rather than an unrepeated stimulus). It was found that NB encoded low-frequency stimuli better than B, in complete agreement with previous results. However, it was found that B encoded high frequency stimuli better than NB due to its more variable baseline activity. Our results suggested that NB and B might serve different functions: namely the differential encoding of low and high frequency stimuli, respectively.

8.2 *Future directions*

8.2.1 Modeling a population of electroreceptor neurons

The model presented here is extremely simple and provides accurate spike train statistics that closely match those obtained experimentally. If one has experimental data from many different electroreceptor neurons, one could fit a BLIFDT model to each and thus build a realistic heterogeneous population of receptor afferents. It is expected that the heterogeneous population will show capabilities of encoding a wider range of stimuli than each individual unit.

Further studies are also needed to understand the role of noise in each electroreceptor neuron. Recently, Stocks and Mannella (2001) have discussed the phenomenon of suprathreshold stochastic resonance (SSR). This phenomenon is seen in homogeneous neural populations. Small to moderate amounts of intrinsic noise in each agent allow population de-synchronization and greater population information transfer. Of course, too much noise will degrade information transfer. There is thus a value of noise for which information transfer is optimized. For a heterogeneous population, the noise value might be unit dependent. It would be interesting to build such a model population and vary the amount of intrinsic noise in each unit. If the amount of noise needed for optimal information transfer was near or equal to the

amount of noise present physiologically, this would give a reason for having noisy neurons rather than deterministic ones.

It would also be interesting to connect this population to the neuron actually receiving the input: the pyramidal cell (Maler 1979). A detailed compartmental model of pyramidal cell activity has recently been developed (Doiron et al. 2001). Further modeling work is necessary in order to understand how information coming from the electroreceptor neuron population is processed at the pyramidal cell level. It was found experimentally that electroreceptor neurons perform good stimulus estimation (Wessel et al. 1996). However, pyramidal cells perform comparatively poorer stimulus estimation (Gabbiani et al. 1996; Bastian et al. 2002) than single electroreceptor neurons, a surprising result since they receive input from a population of them. It has been proposed that pyramidal cells might selectively encode a few features of the stimulus (Gabbiani et al. 1996), thus discarding some information. It remains to be seen how this is actually done. Perhaps a realistic model of receptor afferent input unto a pyramidal cell would give some insight into the problem.

8.2.2 Effects of coloured noise on spike count statistics

It was found that the addition of weak coloured noise to the LIFDT model had a significant impact on spike count statistics at large counting times. Further theoretical work is required in order to fully understand this interesting phenomenon that occurs in experimental data. Jason Middleton, Benjamin Lindner and myself are currently carrying some of this work out. However, this will be presented in Jason's thesis.

8.2.3 Effects of negative ISI correlations on information transfer

We have shown, through numerical simulations, that negative ISI correlations increase information transfer. Further theoretical work is required to gain an understanding of the mechanism by which this occurs. This theoretical work should focus on simpler models of spiking activity; such as on the model proposed by Brandman and Nelson (2000). This would give us greater understanding on how spike patterning would affect information transfer.

Schultz and Panzeri (2001) have studied the effects of spike train correlations on information transfer. As mentioned before, many other neurons also display negative ISI correlations. Such a study might uncover how nature optimizes information transfer in sensory systems. Benjamin Lindner and myself are now working on such theoretical work.

8.2.4 Bursting vs non-bursting dynamics?

It is known that many neurons display bursting dynamics. However, many neurons also display tonic firing. Thalamocortical neurons can in fact display both non-bursting as well as bursting dynamics (see Sherman 2001 for a review). It has further been proposed that non-bursting dynamics would lead to stimulus estimation while bursting dynamics would lead to feature detection. Our results from chapter 7 support this view. However, there are different intrinsic burst mechanisms (see Izhikevich 2000 for a summary). It would be interesting to see how different burst mechanisms will perform at feature detection.

Our results suggest that bursting neurons may be more apt at encoding high frequency stimuli while non-bursting ones would be more apt at encoding low frequency stimuli. This has important behavioural consequences for weakly electric fish. Non-bursting electroreceptor neurons would then preferentially be used for the encoding of prey stimuli while bursting electroreceptor neurons would preferentially be used for the encoding of communication stimuli. This would provide an interesting possible function for each class. Further experimental and behavioural work is needed to verify this prediction.

Further work is also required on how pyramidal cells actually decode the information coming from each class of receptor.

8.3 Conclusion

This thesis provides some insights into the role of neural dynamics on information transfer capabilities through the use of non-linear dynamics, information theory, as well as modeling and numerical simulations.

Bibliography

[1] Abarbanel, H.D.I. (1996). *Analysis of observed chaotic data*. New York: Springer.

[2] Adrian, E.D., and Zotterman, Y. (1926) The impulse produced by sensory nerve endings. Part 2: the response of a single end-organ. *J. Physiol. (lond.)*, **61**:151-171.

[3] Attias H., and Schreiner, C.E. Coding of naturalistic stimuli by auditory midbrain neurons. In: *Advances in Neural Information Processing Systems*, edited by Jordan M, Kearns M and Solla S. Cambridge: MIT press, 1998, p. 103-109.

[4] Azouz, R., and Gray, C.M. (1999). Cellular Mechanisms contributing to response variability of cortical neurons in vivo. *J. Neurosci.*, **19**:2209-2223.

[5] Bahar, S. et al. (2001). Long-range temporal anti-correlations in paddlefish electroreceptors. *Eur. Phys. Letters*, **56**:454-460.

[6] Barlow, H.B. and Levick, W.R. (1969a). Three factors limiting the reliable detection of light by the retinal ganglion cells of the cat. *J. Physiol. (Lond.)*, **200**:1-24.

[7] Barlow, H.B. and Levick, W.R. (1969b). Changes in the maintained discharge with adaptation level in the cat retina. *J. Physiol. (Lond.)*, **202**:699-718.

[8] Bastian, J. (1981). Electrolocation I. How the electroreceptors of Apteronotus Albifrons code for moving objects and other electrical stimuli. *J. Comp. Physiol. A*, **144**:465-479.

[9] Bastian, J., Schneiderjen, S., and Nguyenkim, J. (2001). Arginine vasotocin modulates a sexually dimorphic communication behavior in the weakly electric fish, *Apteronotus leptorhynchus*. *J. Exp. Biol.*, **204**:1909-1923.

[10] Bastian, J., Chacron, M.J., and Maler, L. (2002). Receptive field Organization Determines Pyramidal Cell Stimulus-Encoding Capability and Spatial Stimulus Selectivity. *J. Neurosci.*, **22**: 4577-4590.

[11] Bender, J.A., Dimitrov, A.G., and Miller, J.P. (2001). Biophysical Constraints on the Precision of Neural Coding. *Soc. Neurosci. Abstr.*, **27**.

[12] Bennett, M.V.L., Sandri, C., and Akert, K. (1989). Fine Structure of the tuberous electroreceptor of the high-frequency electric fish, *Sternachus albifrons* (gymnotiformes). *J. Neurocytol.*, **18**:265-283.

[13] Berman, N.J., and Maler, L. (1998). Inhibition evoked from primary afferents in the electrosensory lateral line lobe of the weakly electric fish (*Apteronotus leptorhynchus*). *J. Neurophysiol.* **80**:3173-3196.

[14] Berry, M.J. and Meister, M. (1998). Refractoriness and Neural Precision, *J. Neurosci.*, **18**:2200-2211.

[15] Borst, A. and Theunissen, F. (1999). Information theory and neural coding. *Nature Neuroscience*, **2**:947-957.

[16] Borst, A. and Haag, J. (2001). Effects of Mean Firing on Neural Information Rate. *J. Comp. Neurosci.* **10**:213-221.

[17] Brandman, R. and Nelson, M.E. (2002). A Simple Model of Long-term Spike Train Regularization, *Neural. Comp.*, **14**:1507-1544.

[18] Buracas, G.T., Zador, A.M., DeWeese, M.R., and Albright, T.D. (1998). Efficient discrimination of temporal patterns by motion-sensitive neurons in primate visual cortex. *Neuron*, **20**:959-969.

[19] Chacron, M.J., Longtin, A., St-Hilaire, M., and Maler, L. (2000). Suprathreshold Stochastic Firing Dynamics with Memory in P-Type Electroreceptors. *Phys. Rev Lett.*, **85**:1576-1579.

[20] Chacron, M.J., Longtin, A., and Maler, L. (2001a). Negative Interspike Interval Correlations increase the neuronal capacity for encoding time-dependent stimuli, *J. Neurosci.*, **21**:5328-5343.

[21] Chacron, M.J., Longtin, A., and Maler, L. (2001b). Simple models of bursting and non-bursting P-type electroreceptors. *Neurocomputing*, **38**:129-139.

[22] Chacron, M.J., Pakdaman, K., and Longtin, A. (2003). Interspike Interval Correlations, Memory, Adaptation, and refractoriness in a Leaky Integrate-and-Fire model with threshold fatigue. *Neural Comp.* **15**:253-278.

[23] Cherry, E.C. (1953). Some experiments on the recognition of speech, with one and two ears. *J. Acoust. Soc. Am* **25**:975-979.

[24] Chialvo, D.R., Longtin, A. and Muller-Gerking, J. (1997). Stochastic resonance in neuronal ensembles, *Phys. Rev E.*, **55**:1798-1808.

[25] Clague, H., Theunissen, F., and Miller, J.P. (1997). Effects of adaptation on neural coding by primary sensory interneurons in the cricket cercal system. *J. Neurophysiol.*, **77**:207-220.

[26] Coombes, S. (1999). Liapunov exponents and mode-locked solutions for integrate-and-fire dynamical systems. *Phys. Lett. A.*, **255**:49-57.

[27] Coombes, S., and Bressloff, P.C. (1999). Mode locking and Arnold tongues in integrate-and-fire neural oscillators. *Phys. Rev E.*, **60**:2086-96.

[28] Cover, T. and Thomas, J. (1991). *Elements of Information Theory*. New York: Wiley.

[29] Cox, D.R. and Lewis, P.A.W. (1966). *The Statistical Analysis of Series of Events*. London: Methuen.

[30] Crampton, W.G.R. (1998). Electric signal design and habitat preferences in a species rich assembly of gymnotiform fishes from the upper Amazon basis. *Anais da Acad. Brasiliera de Ciencias* **70**:805-847.

[31] Doiron, B., Longtin, L., Turner, R.W., Maler, L. (2001). Model of gamma frequency burst discharge generated by conditional backpropagation. *J. Neurophysiol.*, **86**:1523-1545.

[32] Dulka, J.G., Maler, L., and Ellis, W. (1995). Androgen-Induced Changes in Electrocummunicatory Behavior Are Correlated with Changes in Substance P-like Immunoreactivity in the Brain of the Electric Fish *Apteronotus leptorhynchus*. *J. Neurosci.*, **15**:1879-1890.

[33] Ermentrout, B., Pascal, M., and Gutkin, B. (2001). The effects of Spike Frequency Adaptation and Negative Feedback on the Synchronization of Neural Oscillators. *Neural Comp.*, **13**:1285-1310.

[34] 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**:787-792.

[35] Fano, U. (1947). Ionization yield of radiations. II. The fluctuations of the number of ions. *Phys. Rev*, **72**:26-29.

[36] Felleman, D.J., Van Essen, D.C. (1991). Distributed hierarchical processing in the primate visual cortex. *Cereb. Cort.*, 1:1-47.

[37] French, A.S., Holden, A.V., and Stein, R.B. (1972). The estimation of the frequency response function of a mechanoreceptor, *Kybernetik*, **11**:15-23.

[38] Fuhrmann, G., Segev, I., Markram, H., and Tsodyks, M. (2002). Coding of temporal information by activity-dependent synapses. *J. Neurophysiol.*, **87**:140-148.

[39] Gabbiani, F. (1996). Coding of time varying signals in spike trains of linear and half-wave rectifying neurons. *Network*, **7**:61-85.

[40] Gabbiani, F. and Koch, C. (1996). Coding of Time-Varying Signals in Spike Trains of Integrate-and-Fire Neurons with Random Threshold. *Neural Comp.*, **8**:44-66.

[41] Gabbiani, F., Metzner, W., Wessel, R. and Koch, C. (1996). From stimulus encoding to feature extraction in weakly electric fish. *Nature*, **384**:564-567.

[42] Gabbiani, F. and Koch, C. (1998). Principles of spike train analysis. In Koch, C. and Segev, I., editors, *Methods in Neuronal Modeling: From Ions to Networks*, pages 313-360. Cambridge: MIT Press.

[43] Gardiner, C.W. (1985). *Handbook of stochastic methods*. Berlin: Springer-Verlag.

[44] Geisler, C.D. and Goldberg, J.M. (1966). A stochastic model of the repetitive Activity of Neurons, *Biophys. J.*, **6**:53-69.

[45] Gestri, G., Masterbroek, H.A.K., and Zaagman, W.H. (1980). Stochastic constancy, variability and adaptation of spike generation: Performance of a giant neuron in the visual system of the fly. *Biol. Cybern.*, **38**:31-40.

[46] Goldberg, J.M., Adrian, H.O., and Smith, F.D. (1964). Response of Neurons of the Superior Olivary Complex of the Cat to Acoustic Stimuli of Long Duration. *J. Neurophysiol.*, **27**:706-749.

[47] Goldman, M.S., Maldonado, P., and Abbott, L.F. (2002). Redundancy reduction and sustained firing with stochastic depressing synapses (2002). *J. Neurosci.*, **22**:584-591.

[48] Green, D. and Swets, J. (1966). *Signal Detection Theory and Psychophysics*. New York: John Wiley & Sons.

[49] Hausser, M. and Roth, A. (1997). Dendritic and somatic glutamate receptor channels in rat cerebellar Purkinje cells, *J. Physiol. (Lond.)*, **501**:77-95.

[49] Heiligenberg, W. (1991). *Neural Nets in Electric Fish*. Cambridge (MA): MIT press.

[50] Heiligenberg, W., Keller, C.H., Metzner, W., and Kawasaki, M. (1991). Structure and function of neurons in the complex of the nucleus electrosensorius of the gymnotiform fish *Eigenmannia*: Detection and processing of electric signals in social communication. *J. Comp. Physiol. A*, **169**:151-164.

[51] 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. (lond.)*, **117**:500-544.

[52] Holden, A.V. (1976). Models of the Stochastic Activity of Neurones. New York: Springer Verlag.

[53] Hopfield, J.J. (1995). Pattern recognition computation using action potential timing for stimulus representation. *Nature*, **376**:33-36.

[54] Hopkins. C.D. (1973). Lightning as background noise for communication among electric fish. *Nature*, **242**:268-270.

[55] Hopkins, C.D. (1976). Stimulus filtering and electroreception: tuberous electroreceptors in three species of gymnotid fish. *J. Comp. Physiol. A*, **111**:171-207.

[56] Hutcheon, B. and Yarom, Y. (2000). Resonance, oscillation and the intrinsic frequency preferences of neurons. *TINS*, **23**:216-222.

[57] Izhikevich, E.M. (2000). Neural Excitability, spiking, and bursting. *Int. J. Bif. Chaos Appl. Sci. Eng.*, 10:1171-1266.

[58] Jaeger, D., and Bauer, J.M. (1994). Prolonged responses in rat cerebellar Purkinje cells following activation of the granule cell layer: an intracellular in vitro and in vivo investigation. *Exp. Brain Res.*, 100:200-214.

[59] 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.

[60] Keener, J.P. (1980) Chaotic behavior in piecewise continuous difference equations. *Trans. Amer. Math. Soc.*, **261**:589-604.

[61] Keener, J.P., Hoppensteadt, F.C., and Rinzel, J. (1981). Integrate-and-fire models of nerve membrane response to oscillatory input. *SIAM J. of Appl. Math.*, 41:816-823.

[62] Koch, C. (1999). *Biophysics of Computation*. New York: Oxford University Press.

[63] Kreiman, G., Krahe, R., Metzner, W., Koch, C., and Gabbiani, F. (2000). Robustness and Variability of Neuronal Coding by Amplitude Sensitive Afferents in the Weakly Electric Fish *Eigenmania*. *J. Neurophysiol.*, 84:189-224.

[64] Kuffler, S.W., Fitzhugh, R., and Barlow, H.B. (1957). Maintained activity in the cat's retina in light and darkness. *J. Gen. Physiol.*, 40:683-702.

[65] Laughlin, S. (1981). A Simple Coding Procedure Enhances A Neuron's Information Capacity. *Z. Naturforsch.* 36:910-912.

[66] Le Calvez, P. (2000) *Dynamical properties of diffeomorphisms of the annulus and of the torus*. SMF/AMS Texts and Monographs, American Mathematical Society, Providence, RI; Société Mathématique de France, Paris, 2000.

[67] Lewen, G.D., Bialek, W., and de Ruyter van Steveninck, R.R. (2001). Neural coding of naturalistic motion stimuli. *Network*, 12:317-329.

[68] Liu, Y.H., and Wang, X.J. (2001). Spike frequency adaptation of a generalized Leaky Integrate-and-fire neuron. *J. Comput. Neurosci.*, 10:25-45.

[69] Longtin, A. and Racicot, D.M. (1997). Spike train patterning and forecastability. *BioSystems*, 40:111-118.

[70] Longtin, A., and St-Hilaire, M. (2000). Encoding carrier amplitude modulations via stochastic phase synchronization. *Int. J. Bifur. Chaos*, 10:1-16.

[71] Longtin, A., Laing, C., and Chacron, M.J. (2003). Correlations and memory in neurodynamical systems. In: *Long-Range-Dependent Stochastic Processes: Theory and Applications*. Rangarajan. G. and Ding, M., editors. Springer-Verlag: Berlin.

[72] Lowen, S.B. and Teich, M.C. (1992). Auditory nerve action potentials form a nonrenewal process over short as well as long time scales. *J. Acoust. Soc. Am*, 92:803-806.

[73] Lowen, S.B. and Teich, M.C. (1996). The periodogram and Allan variance reveal fractal exponents greater than unity in auditory-nerve spike trains. *J. Acoust. Soc. Am*, 99:3585-3591.

[74] Lowen, S.B., Cash, S.S., Poo, M., and Teich, M.C. (1997). Quantal Neurotransmitter Secretion Rate Exhibits Fractal Behavior. *J. Neurosci.*, **17**:5666-5677.

[75] Lucas, K. (1906). The Analysis of complex exciable tissues by their response to electric currents of short duration. *J. Physiol.*, **35**:310-331.

[76] Madison, D.V., and Nicoll, R.A. (1984). Control of the repetitive discharge of rat CA 1 pyramidal neurones in vitro. *J. Physiol.*, **354**:319-331.

[77] Maler, L., Sas, E., Johnston, S., Ellis, W. (1991). An atlas of the brain of the weakly electric fish *Apteronotus Leptorhynchus*. J. Chem. Neuro., **4**:1-38.

[78] Manwani, A., and Koch, C. (1999). Detecting and Estimating Signals in Noisy Cable Structures, I: Neuronal noise Sources. *Neural. Comp.*, **11**:1797-1829.

[79] Mardia, K.V., and Jupp, P.E. (1999). *Directional Statistics*. New York: Wiley.

[80] Meiss, J.D. (1992). Symplectic Maps, Variational Principles, and Transport. *Rev Mod. Phys.*, **64**:795-848.

[81] Mickus, T., Jung, H.Y., and Spruston, N. (1999). Properties of slow cumulative sodium channel inactivation in rat hippocampal CA1 pyramidal neurons. *Biophys. J.*, **76**:846-860.

[82] Metzner, W. and Heiligenberg, W. (1991). The coding of signals in the electric communication of the gymnotiform fish *Eigenmannia*: From electroreceptors to neurons in the torus semicircularis of the midbrain. *J. Comp. Physiol. A*, **169**:135-150.

[83] Moortgat, K.T., Keller, C.H., Bullock, T.H., and Sejnowski, T.J. (1998). Submicrosecond pacemaker precision is behaviourally modulated: the gymnotiform electromotor pathway. PNAS, **95**:4684-4689.

[84] Nachimas, J. (1972). Signal Detection Theory and its Application to Problems in Vision. In Jameson, D. and Hurvitch, L.M., editors, *Visual Psychophysics* VII/4 pages 56-77. Berlin: Springer-Verlag.

[85] Nelson, M.E, Xu, Z., and Payne, J.R. (1997). Characterization and modeling of P-type electrosensory afferent responses to amplitude modulations in a wave-type electric fish. *J. Comp. Physiol. A*, **181**:532-544.

[86] Nelson, M.E. and MacIver, M.A. (1999). Prey capture in the weakly electric fish apteronotus albifrons: sensory acquisition strategies and electrosensory consequences, *J. Exp. Biol.*, **202**:1195-1203.

[87] Nirenberg. S., Carcieri, S.M., Jacobs, A.L., and Latham, P.E. (2001). Retinal ganglion cells act largely as independent encoders. *Nature*, **411**:698-701.

[88] Pakdaman, K. (2001). The periodically forced leaky integrate-and-fire model. *Phys. Rev E.*, **63**:1907-1912.

[89] Pakdaman, K., Tanabe, S., and Shimokawa, T. (2001) Coherence resonance and discharge time reliability in neurons and neuronal models. *Neural Networks*, **14**:895-905.

[90] Pakdaman, K., and Vibert, J.F. (1995) Modeling excitatory networks. *Journal of Sice*, **34**:788-793.

[91] Pakdaman, K., Vibert, J.F., Boussard, E., and Azmy, N.(1996) Single neuron with recurrent excitation: Effect of the transmission delay. *Neural Networks*, **9**:797-818.

[92] Panzeri, S. and Schultz, S.R. (2001). A Unified Approach to the Study of Temporal, Correlational, and Rate Coding. *Neural Comp.* **13**, 1311-1349.

[93] Pathria, R.K. (1996). *Statistical Mechanics*. Oxford: Butterworth-Heineman.

[94] Parthridge, B.L., and Heiligenberg, W. (1980). Three's a crowd? Predicting Eigenmania's response to multiple jamming. *J. Comp. Physiol.*, **136**:153-164.

[95] Poor, H.V. (1994) *An introduction to signal detection and estimation*. New York: Springer.

[96] Ratnam, R. and Nelson, M.E. (2000). Non-renewal statistics of electrosensory afferent spike trains: implications for the detection of weak sensory signals. *J. Neurosci.*, **20**:6672-6683.

[97] Reinagel, P., Godwin, D., Sherman, S.M. and Koch, C. (1999). Encoding of visual information by LGN bursts. *J. Neurophysiol.* **81**, 2558-2569.

[98] Reinagel, P. and Reid, R.C. (2000). Temporal Coding of Visual Information in the Thalamus. *J. Neurosci.*, 20:5392-5400.

[99] Rescigno, A., Stein, R.B., Purple, R.L., and Popele, R.E. (1970). A neuronal model for the discharge patterns produced by cyclic inputs. *Bull. Math. Biophys.*, 32:337-353.

[100] Rieke. F., Bodnar, D.A., and Bialek, W. (1995). Naturalistic stimuli increase the rate and efficiency of information transmission by primary auditory afferents. *Proceedings of the Royal Society of London - Series B: Biological Sciences* 262:259-265.

[101] Rieke, F., Warland, D., de Ruyter van Steveninck, R., and Bialek ,W. (1997). *Spikes: exploring the neural code.* Cambridge, M.A.: MIT.

[102] Roddey, J.C., and Jacobs, G.A. (1996). Information Theoretic analysis of dynamical encoding by filiform mechanoreceptors in the cricket cercal system. *J. Neurophysiol.* 75:1365-1376.

[103] Russell, D.F., Wilkens, L.A., and Moss, F. (1999). Use of behavioural stochastic resonance by paddle fish for feeding. *Nature*, 402:291-294.

[104] Salinas, E. and Sejnowski, T.J. (2000). Impact of Correlated Synaptic Input on Output Firing Rate and Variability in Simple Neuronal Models. *J. Neurosci.*, 20:6193-6209.

[105] Schäfer, K., Braun, H.A., Peters, C., and Bretschneider, F. (1995). Periodic firing pattern in afferent discharges from electroreceptor organs of catfish. *Eur. J. Physiol.* **429**, 378-385.

[106] Scheich, H., Bullock, T.H., and Hamstra, R.H. (1973). Coding properties of two classes of afferent nerve fibers: high frequency electroreceptors in the electric fish, *eigenmania. J. Neurophysiol.*, 36:39-60.

[107] Segundo, J.P., Perkel, D.H., Wyman, H., Hegstad, H., and Moore, G.P. (1968) Input output relations in computer simulated nerve cells. *Kybernetik*, 4:157-175.

[108] Shannon, C.E. (1948). The mathematical theory of communication. *Bell. Syst. Tech. J.*, **27**:379-423,623-656.

[109] Sherman, S.M. (2001). Tonic and burst firing: dual modes of thalamocortical relay. *TINS*, **24**: 122-126.

[110] Simoncelli, E.P., and Olshausen, B.A. (2001). Natural Image Statistics and Neural Representation. *Annu. Rev Neurosci.* **24**:1193-1216.

[111] Snippe, H.P. and Koenderink, J.J. (1992). Discrimination thresholds for channel-coded systems. *Biol. Cybern.*, **66**:543-551.

[112] Stein, R.B. (1965). A theoretical analysis of neuronal variability, *Biophys. J.*, **5**:173-194.

[113] Steriade, M. (1978). Cortical long-axoned cells and putative interneurons during the sleep-waking cycle. *Behav Brain Res.* **3**:465-514.

[114] Stocks, N.G., Mannella, R. (2001). Generic noise-enhanced coding in neuronal arrays: suprathreshold stochastic resonance. *Phys. Rev E*, **63**:041114/1-9.

[115] Strogatz, S.H. (1994). *Nonlinear dynamics and chaos*, Reading: Perseus Books.

[116] 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**:197-200.

[117] Teich, M.C. and Khanna, S.M. (1985). Pulse-number distributions for the neural spike train in the cat's auditory nerve. *J. Acoust. Soc. Am*, **77**:1110-1128.

[118] Teich, M.C. (1992). Fractal Neuronal Firing Patterns. In McKenna, T., Davis, J. and Zornetzer, S.F., editors, *Single Neuron Computation*, pages 589-622. San Diego: Academic Press.

[119] Teich, M.C., Turcott, R.G., and Siegel, R.M. (1996). Temporal correlations in cat striate-cortex neural spike trains. *IEEE Eng Med Biol. Mag.*, **15**:79-87.

[120] Teich, M.C., Heneghan, C., Lowen, S.B., Ozaki, T., and Kaplan, E. (1997). Fractal Character of the neural spike train in the visual system of the cat. *J. Opt. Soc. Am. A*, 14:529-546.

[121] Treves, A. (1996). Mean-field analysis of neuronal spike dynamics. *Network*, 4:259-284.

[122] Tuckwell, H.C. (1988). *Introduction to Theoretical Neurobiology: I.*. Cambridge: Cambridge University Press.

[123] Turcott, R.G. and Teich, M.C. (1996). Fractal Character of the electrocardiogram: distinguishing heart failure and normal patients. *Ann. Biomed. Eng.*, 24:269-293.

[124] Vibert, J.F., Pakdaman, K., and Azmy, N. (1994). Inter-neural delay modification synchronizes biologically plausible neural networks. *Neural Networks*, 7:589-607.

[125] Volterra, V. (1930). *Theory of Functional and of Integral and Integro-differential Equations*. London: Blackwell Scientific.

[126] Von Gernsdoff, H., Schneggenburger, R., Weis, S., and Neher, E. (1997). Presynaptic depression at a calyx synapse: the small contribution of metabotropic glutamate receptors. *J. Neurosci.*, 17:8137-8146.

[127] Wang, L.Y., Gan, L., Forsythe, I.D., and Kaczmarek, L.K. (1998). Contribution of the Kv3.1 potassium channel to high-frequency firing in mouse auditory neurones. *J. Physiol. (Lond.)*, 509:183-194.

[128] Warland, D.K., Reinagel, P., and Meister, M. (1997). Decoding visual information from a population of retinal ganglion cells. *J. Neurophysiol.* 78:2336-2350.

[129] Wehmeier, W., Dong, D., Koch, C., and van Essen, D. (1998) Modeling the mamalian visual system. In C. Koch and I. Segev (Eds.). *Methods in Neuronal Modeling*. Cambridge, MA: MIT press.

[130] Wessel, R., Koch, C., and Gabbiani, F. (1996). Coding of time varying electric field amplitude modulations in a wave-type electric fish. *J. Neurophysiol.*, 75:2280-2293.

[131] Xu, Z., Payne, J.R., and Nelson, M.E. (1996). Logarithmic Time Course of Sensory Adaptation in Electrosensory Afferent Nerve fibers in a weakly electric fish. *J. Neurophysiol.*, **96**:2020-2032.

[132] Zakon, H.H. (1986). The electroreceptive periphery. In Bullock, T.H. and Heiligenberg, W., editors, *Electroreception*, pages 103-156. New York: John Wiley and Sons.

[133] Zupanc, GKH and Maler, L. (1993). Evoked chirping in the weakly electric fish *Apteronotus Leptorhynchus*: a quantitative biophysical analysis. Can. J. Zool. **71**:2301-2310.