Long-tailed interspike interval distributions from cortical neurons

Ruedi Stoop† and Logan Kleinwaks‡

†*Institut für Neuroinformatik, ETHZ/UNIZH, Zürich 8006, Switzerland*
ruedi@ini.phys.ethz.ch
‡*Princeton University, Princeton, NJ 08544, USA*
kleinwks@princeton.edu

***Abstract*—**
We investigate the distributions of interspike intervals, both from the experimental and the theoretical point of view. We show experimental evidence for the fact that these distributions are long-tailed, and we supply theoretical arguments for how such tails may arise.

I. INTRODUCTION

Neocortical circuits are formed of recurrently connected neurons. These neurons are of two basic types, inhibitory and excitatory, and are reciprocally coupled in monosynaptic or polysynaptic arcs. Their possible roles have been the subject of many experimental and theoretical analyses. Less attention has been given to the effect of such recurrent coupling on the global patterns of activity generated in extended recurrent circuits of spiking neurons. Preliminary evidence from combined optical and single unit recordings in the primate visual cortex indicate that single unit responses occur within complex global patterns of activity [1]. However, the nature of this activity in large populations of neurons in not well understood. These interactions are mostly studied by looking at neuron interspike interval distributions. Common belief is that these distributions are Gaussian or Poissonian, or, at least, approximately so. Our point, however, is that models of long-tailed distributions are more appropriate. Our reasons to believe this originate from information theoretic arguments, from measured interspike distributions of *in vivo* experiments, and from models based on *in vitro* experiments. We begin by showing how α-stable interspike interval distributions can arise from information considerations of a biologically suitable network theory.

II. NETWORK THEORY

We consider a network consisting of the elements E_1, E_2, ..., E_k, which receive a common signal S, and, depending on this signal, generate outputs O_i^S, which are all read by another element E (see Figure 1). It may be that the outputs O_i^S only depend on certain characteristics of the signal S; to clarify this notion we introduce the formal definition of *signal characteristics*. We say that a set $\mathcal{F}$ of functionals defined on the space of signals is a *set of signal characteristics* of an element E_i if (i) $f(S_1) \stackrel{\mathrm{d}}{=} f(S_2) \forall f \in \mathcal{F} \Rightarrow O_i^{S_1} \stackrel{\mathrm{d}}{=} O_i^{S_2}$ and if (ii) $\exists f \in \mathcal{F}$ s.t. $f(S_1) \stackrel{\mathrm{d}}{\neq} f(S_2) \Rightarrow O_i^{S_1} \stackrel{\mathrm{d}}{\neq} O_i^{S_2}$. Here, we have used the notation that "d" above a relation symbol means that the relation holds for the distributions of the two objects. This allows for randomness in the signal and in the response of the element. Henceforth, we assume that the E_i have a common set of signal characteristics, which we denote by $\mathcal{F}$.

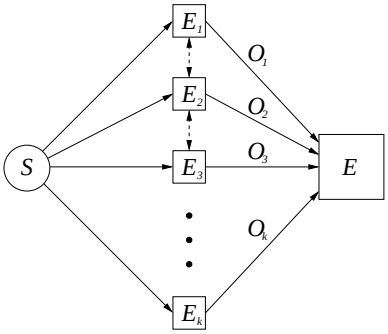

Figure 1: The network geometry. Dashed arrows are referred to in Section III.

We say that a signal S is *stable in the signal characteristic* $f \in \mathcal{F}$ if $f(S)$ is time-independent. An *absolutely stable* signal is stable in every signal characteristic in $\mathcal{F}$. Extreme examples are the cases where the outputs do not depend on the signal, so that $\mathcal{F}$ may consist only of a functional which maps every signal to the same value, in which case every signal is absolutely stable, and the case where everything about the signals matters, so that $\mathcal{F}$ may contain a single invertible functional. We suspect that biological neural networks fall between these extreme cases, and that perhaps functionals which measure local signal frequency are signal characteristics. However, the exact nature of the signal characteristics and absolutely stable signals is, at the present time, speculative, and all that will matter for us is that absolutely stable signals exist.

Now we let S be an absolutely stable signal and consider the element E, which receives and processes all of the outputs O_i^S. The importance of the absolutely stable signals we have just introduced is that they are the "constant" signals, since all of the information-carrying features of signals are constant in them. We would like E to be able to identify absolutely stable signals; that is, to determine that the information in S does not change. If E can do this, then the information that it derives from the outputs O_i^S must converge in some sense over time. Furthermore, in the biological system, this convergence should take place quickly. To state this precisely, we must assume something about how E derives information from the outputs.

In a biological neural network, the elements are neurons and the outputs O_i^S are membrane potentials. We make two critical assumptions about how E derives information from the membrane potentials. First, we assume that only the intervals between spikes in the membrane potentials affect the output of E (this partly specifies the signal characteristics of E). We denote the time between the n^{th} and $(n+1)^{th}$ spikes in the output O_i^S by X_{in}. We assume that the X_{in} are realizations of random variables with a common distribution F, which is the interspike interval distribution, and let $\mathrm{S}_{in} = \sum_{j=1}^{n} \mathrm{X}_{ij}$. We let $\mathrm{S}_n = \sum_{j=1}^{n} \mathrm{X}_j$, where X_j are random variables which also have the distribution F. E can only approximate the distribution of S_n by its empirical distribution, determined from the S_{in}. Second, we assume that the output of E at a time t depends only on the empirical distributions of $a_n\mathrm{S}_n - b_n$ that have accumulated up to the time t, where $a_n > 0$ and b_n are constants that encapsulate the processing E does to the interspike intervals. If the network size k is large, as we expect it is in the biological case (k may be on the order of 10^4), then these empirical distributions should approximate the true distributions well.

Now we can state exactly what we mean by the convergence of the information derived from the outputs O_i^S. If E can identify S as being absolutely stable, then the empirical distribution of $a_n\mathrm{S}_n - b_n$ must converge to a distribution G as $n \to \infty$. There are three factors which affect the rate of this convergence: the network size k, the tolerance of E, and the signal S itself. We assume that k is large enough and the tolerance is strict enough so that convergence of the empirical distribution implies that the distribution must also converge to G. The problem of relating F, G, a_n, and b_n provided that this convergence takes place is the study of domains of attraction in mathematics, and we state here only the results which concern us (proofs can be found in [2]).

By its definition, F must be concentrated on $(0, \infty)$, and this implies that if G is a proper distribution not concentrated at a point, then

(1) $\exists$ a function L which is slowly varying at ∞ and a constant α, $0 < \alpha < 1$ such that

$$1 - F(x) \sim \frac{x^{-\alpha}L(x)}{\Gamma(1-\alpha)},\ x \to \infty$$

(2) the a_n must satisfy $na_n^{-\alpha}L(a_n) \to 1$, and

(3) G must be a stable distribution with characteristic exponent α

(1) implies that the tail of the density of the interspike interval distribution F must decay as $x^{-(\alpha+1)}$, (2) relates the a_n to this decay, and (3) specifies the exact form of G, up to its location parameters.

In a biological neural network, we know that this convergence must occur quickly. Certainly, if the interspike interval distribution is itself α-stable, then we have the quickest convergence possible for a given network size k and tolerance of E. However, we also expect that for interspike interval distributions which satisfy (1)-(3) and are close to being α-stable, convergence should take place quickly. Since the interspike interval distribution must be supported on $(0, \infty)$, the only way it can be close to an α-stable distribution is to have a large skewness. This imples that the interspike interval density should fall off rapidly towards zero. The possible case where G is concentrated at a point, however, only admits fast convergence if the interspike interval distribution is a narrow and quickly decaying peak.

Thus, we have shown that for absolutely stable distributions to be identified quickly in a biological neural network, the densities of the interspike interval distributions should have power-law decaying tails and rapid decay towards zero, or they should be sharp peaks. Furthermore, we have established a direct relationship between the processing done by the neurons on the interspike intervals and the shape of the interspike interval distributions. We note that in the much debated model of simple averaging of the intervals, corresponding to $a_n = n^{-1}$, absolutely stable signals can only be identified quickly if the interspike interval distributions are sharp peaks. If we require the a_n to be fixed in E, then E can only distinguish absolutely stable signals by the three location parameters of G.

III. A MODEL OF NOISE-DRIVEN NEURON INTERACTIONS

We now outline our model of interaction among noise-driven quasistatic neurons [3], and show that it also generates long-tailed interspike interval distributions. This model is based on experimentally measured neuron behavior. Neurons in the cortex receive input from other neurons and, when the firing threshold is

reached, document this by firing an action potential (the "spike"). Although synaptic output is released in quantals, neurons receive inputs of different orders of magnitude:

- Small-scale noisy input (e.g., from remote synapses) drives the neuron towards regular spiking with well-defined periodicity. The noise may be considered as without structure; therefore, it may be represented by a constant driving current.
- Strong input by next neighbors (neurons or a group of synchronized neurons) arrives at the neuron as a simple, ideally periodic, structure in time.
- Medium-size interactions are neglected in this picture. In a refined approach, they may be treated by a coupling mechanism, for example on a coupled map lattice.

Starting from this picture, we first mention that upon constant current driving, the neuron indeed responds with regular spiking. In experiments with real neurons, slices of rat neocortex are prepared for *in vitro* recording (for details of the preparation see [4]). The neuron then is supplied with a constant current [5] which drives the membrane potential towards the spiking threshold. Under these conditions, the cell starts to fire regularly, on the basis of small membrane fluctuations that have a drive towards the spiking threshold. In the mathematical abstraction, the unperturbed regular spiking behavior corresponds to a limit cycle solution of the associated oscillator equations of the cable model of the neuron [6]. Strong inhibitory or strong excitatory inputs correspond to a perturbation of this solution. This concept will lead us to a simplification of the description of neuron spiking, which can be directly based on experimental measurements.

In addition to the small-scale noise, information from other neurons arrives in our model at the neuron in the form of substantial packages of spikes, received within a certain small time interval. (This type of interaction is represented in Figure 1 by the dashed arrows between the symbols E_i, where the signal S is the driving noise.) Experimentally, this perturbation is performed by the stimulation of a synaptic connection to the neuron. To investigate the perturbed limit cycle, we applied the techniques originally put forward by Glass and Mackey [7]. We are interested in the typical response of an intrinsically regularly spiking neuron to synaptic perturbation by an also regularly spiking neuron. The response of the targetted neuron has strongly nonlinear characteristics. At fixed perturbation strength, the effect of the perturbation depends on the *phase* ϕ (with respect to the neuron's own regular spiking) at which the perturbation is applied. This property is revealed by the *phase response function* $g(\phi)$, which returns the quotient between the perturbed interspike interval length to the intrinsic (i.e., unperturbed) interval length as a function of ϕ. The phase response and phase return function are related through

$$f_\Omega : \phi_2 = \phi_1 + \Omega - g(\phi) \pmod 1,$$

where the parameter Ω is the quotient of the intrinsic interspike time T_0 of the targetted neuron divided by the interspike time T_s of the targetting neuron (for details, see [4, 8]). Taking into account how the phase response function depends on the perturbation strength K, the above equation can be seen as defining a circle map [9]. The functional dependence on K has the form

$$g_{\Omega,K}(\phi) = g_{\Omega,1}(\phi - 1)K + 1.$$

Investigation of the returned periodicities as a function of Ω, K results in typical Arnol'd tongue structures [3], see Figure 2. For each periodicity p, there are different Arnol'd tongues which comprise areas in the Ω, K parameter space which have solutions of the same periodicity p. For the different areas, the stability properties of the solutions, which can be expressed by the Lyapunov exponent $\lambda_{\Omega,K}$ [10], are of interest. Zooming in on the Arnol'd tongue plots reveals that for inhibition, chaotic behavior is possible ($\lambda_{\Omega,K} > 0$). However, large input strengths are needed to generate a chaotic response. Analytic investigations in [3] reveal that this occurs on a nonzero set in the relevant parameter space. Excitatory stimulations always yield invertible phase return maps on the biologically meaningful parameter space.

We now consider a given neuron and start, for convenience, with an intrinsic interspike interval of length one. In order to mimic the interaction with not only the strongest, but with every possible interaction, we perform a perturbation average over the Arnol'd tongue structure, with respect to the Lebesgue measure. We start, for convenience, with a unit interval that is perturbed by a random number of randomly excitatory or inhibitory perturbations of different perturbation strengths. From this process, as in experiments, an interspike interval distribution is generated. A typical result is shown in the left of Figure 3. As can be inferred from this figure, almost perfect long-tail behavior is displayed, along with the expected type of behavior towards zero. The histogram even shows Lévy-type behavior, with an exponent of $\alpha \approx 1.8$. The histogram was obtained by restricting excitabilities to a range of strengths $K \in [0.005, 0.255]$, measured in units of the maximally applicable perturbation strength. These working conditions are biologically reasonable and significant for close-to-equilibrium states of the brain.

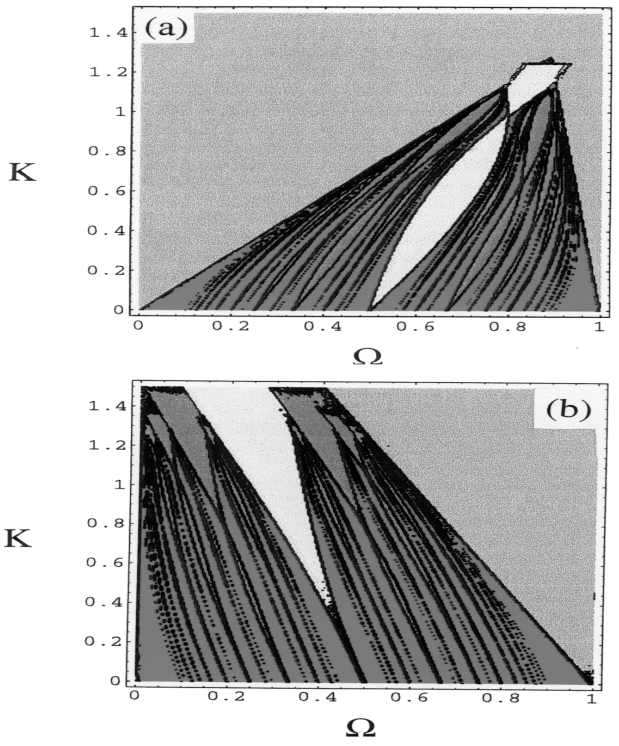

Figure 2: Arnol'd tongues obtained from inhibitory interaction (a) and excitatory interaction (b).

Choosing higher ranges of excitabilities yields less evident Lévy behavior. This is partly due to the appearance of a phenomenon which is widely known to the experimentalists as the "second frequency." This phenomenon consists of an additional wiggle in the interspike interval density, which obstructs the formation of a clean power-law decay. In our approach, these wiggles emerge in a systematic way, and a large (in principle, infinite) number of such high order frequencies should be observable, given a sufficient resolution of the data.

IV. DISCUSSION

Our information theoretic analysis suggests that it may be possible to determine features of neuronal information processing through measured interspike interval distributions. Our model of noise-driven neuron interactions also points to an as yet unrealized usefulness of interspike interval distributions, namely that they may contain traces of extreme excitability. However, to make such inferences from experimental data, there must be enough data so that the tail statistics are good, and confounding factors, such as adaptation, must be accounted for.

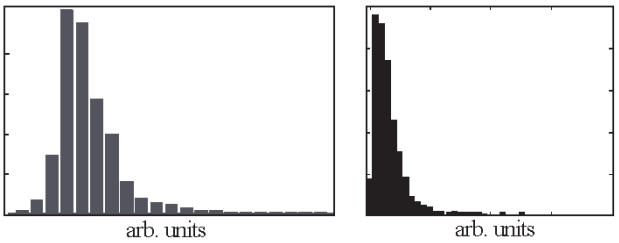

Figure 3: Long-tailed histograms based on experimental data. Left: Interspike interval distribution from the model of Section III. Right: ISI histogram recorded from a complex-type neuron in cat striate cortex. The histogram contains the neuron's combined responses to five, 4 sec. presentations of a spatiotemporally optimized sine wave grating at 40% contrast.

V. ACKNOWLEDGEMENTS

It is a pleasure for us to thank K.A.C. Martin for enthusiasm towards this project, and John Allison for providing the data from cat striate cortex. L.K. is especially grateful to R. Douglas and J. Fröhlich for the opportunity to work at the INI, and to the Nicholas Green Scholarship Foundation for financial support.

References

[1] A. Arieli, A. Sterkin, A. Grinvald, and A. Aertsen, *Science*, vol.273, pp.1868–1871, 1996.

[2] W. Feller, "An Introduction to Probability Theory and Its Applications," vol.2, John Wiley: New York, 1971.

[3] R. Stoop, K. Schindler, and L.A. Bunimovich, submitted.

[4] K. Schindler, C. Bernasconi, R. Stoop, P. Goodman, and R. Douglas, *Z. Naturforsch.*, vol.52a, pp.509–512, 1997.

[5] M. Abeles, "Local cortical circuits," Springer: Berlin, 1982.

[6] M. Hines, *Int. J. Biomed. Comp.*, vol.24, pp.55–68, 1989.

[7] L. Glass and M. Mackey, "From Clocks to Chaos," Princeton University Press: Princeton, 1988.

[8] R. Stoop and L.A. Bunimovich, *Proc. NOLTA'97*, vol.1, pp.169–172, 1997.

[9] E.P. Cornfeld, S.V. Fomin, and Ya.G. Sinai, "Ergodic Theory," Springer: Berlin, 1982.

[10] J. Argyris, G. Faust, and M. Haase, E.N. "Die Erforschung des Chaos," Vieweg: Braunschweig, 1995.