

Entropy and local uncertainty of data from sensory neurons

R. Steuer* and W. Ebeling†

Institute of Physics, Humboldt-University Berlin, Invalidenstrasse 110, 10115 Berlin, Germany

D. F. Russell,‡ S. Bahar,§ A. Neiman,|| and F. Moss¶

Center for Neurodynamics, University of Missouri, 8001 Natural Bridge Road, St. Louis, Missouri 63121

(Received 21 June 2001; revised manuscript received 4 September 2001; published 27 November 2001)

We present an empirical comparison between neural interspike interval sequences obtained from two different kinds of sensory receptors. Both differ in their internal structure as well as in the strength of correlations and the degree of predictability found in the respective spike trains. As a further tool in this context, we suggest the local uncertainty, assigning a well-defined predictability to individual spikes. The local uncertainty is demonstrated to reveal significant patterns within the interspike interval sequences, even when its overall structure is (almost) random. Our approach is based on the concept of symbolic dynamics and information theory.

DOI: 10.1103/PhysRevE.64.061911

PACS number(s): 87.19.La, 87.19.Bb, 87.80.Tq, 05.45.Tp

I. INTRODUCTION

In recent years, several authors have investigated the encoding of information in neural spike trains using the concepts of entropy and information theory [1–7]. Of particular interest has been the correlation between the ability of an animal to detect weak sensory signals and the existence of extended memory in the respective spike train [3,8]. Weak deviations, caused by external stimuli, must be detected against the ongoing background of the spontaneous activity. Intuitively, this task should be easier when the spontaneous activity is rather regular and predictable [3]. On the other hand, predictability and correlations are known to reduce entropy, and thus might be expected to degrade information transmission. While it is presently not clear what effect correlations will have on the signal detection performance of different animals, we may safely assume that nature has developed a variety of different mechanisms to provide an effective (well-adapted) representation of outside stimuli.

Here we present an empirical comparison between the spontaneous activity obtained from two different sensory receptors, the paddlefish electroreceptor (ER) and the crayfish mechanoreceptors. Both are sensory receptors, for perception of the outside world, still they differ in their internal structure and the complexity of the generated interspike interval sequences. We will study the entropy and predictability of these sequences based on the concept of symbolic dynamics.

The paper is organized as follows. First, we will give a brief description of the experimental setup. After that, we will shortly review the concepts of symbolic dynamics and Shannon entropy [9]. In the last section, we will summarize the results obtained from the empirical interspike interval

sequences and discuss the biological implications of the findings.

II. PADDLEFISH AND CRAYFISH

Our first experimental object is the paddlefish *Polyodon spathula*, an ancient freshwater species native to the Mississippi River drainage, having evolved 65 million years ago [10]. The paddlefish is characterized by a unique, rostral extension in front of the head, covered with tens of thousands of electrosensory receptors, morphologically similar to the ampullae of Lorenzini of sharks and rays [11–13]. The electrosense is passive, and is used to detect weak electrical signals from planktonic prey [14].

Significant for this study, each electroreceptor is a rather complex system, consisting of a cluster of 1–35 skin pores, each leading into a short canal, which ends in a sensory epithelium containing approximately 400 hair cells. The hair cells of each pore, which are considered electrosensitive, synaptically excite the terminals of a primary afferent axon projecting to the brain [11,14]. The synapse from each hair cell, together with the spiking properties of the primary afferent endings, convert the analog signal from the hair cells into spike trains, coding the electrosensory information as a time series (the intervals between spikes).

We recorded single-unit spikes of electroreceptor afferents, *in vivo*, using tungsten microelectrodes. For a detailed description of the apparatus and methods see [14–16]. Each recording contains 10 000 to 50 000 interspike intervals. All sets of interspike intervals had coefficients of variation $CV \leq 0.3$, defined as $CV := \sigma_\tau / \langle \tau \rangle$ with $\langle \tau \rangle$ being the mean interspike interval and $\sigma_\tau = \sqrt{\langle \tau^2 \rangle - \langle \tau \rangle^2}$. The interspike interval histogram (ISIH) of all files resembled a log-normal distribution and showed a well-expressed unimodal peak.

It was recently discovered that each electroreceptor contains two coupled oscillators [17]: one oscillator resides in the population of hair cells and the other oscillator is associated with the afferent neuron. Stochastic biperiodic oscillations of the electroreceptor system are reflected in an auto-

*Email address: steuer@physik.hu-berlin.de

†Email address: ebeling@physik.hu-berlin.de

‡Email address: drussell@admiral.umsl.edu

§Email address: bahar@neurodyn.umsl.edu

||Email address: neiman@neurodyn.umsl.edu

¶Email address: mossf@umsl.edu

correlation function of interspike intervals that displays alternations (anticorrelations) extending to up to 20 to 40 interspike intervals.

The spike trains obtained from the paddlefish electroreceptor will be set against those obtained from mechanoreceptor afferents in the sensory nerve roots of the crayfish. Compared to the paddlefish, the crayfish has a much longer evolutionary history, having ancestors that arose more than 600 million years ago [18]. The mechanoreceptor system is thought to be primarily a predator avoidance system, sensitive to hydrodynamic motions. Details of the apparatus and methods used for the crayfish experiments are given in [19–21]. Each mechanoreceptor may be viewed as having two elements: a stiff hair, which compresses a stretch-sensitive region of the sensory neuron and the neuron itself. There are no synapses. The extracellular recordings were made with a suction electrode attached to the afferent sensory neuron. For this, the crayfish tailfan and part of the abdominal nerve cord were excised intact and immersed in an appropriate saline solution. Each recording yielded 10 000 to 220 000 interspike intervals under relatively constant experimental conditions.

The firing sequences of mechanoreceptors are characterized by high variability: CV 's are in the range from 0.7 to 1.2. The ISIH's of mechanoreceptors do not possess any expressed maxima and can be well approximated by an exponential distribution. This indicates the lack of oscillations. The autocorrelation function is characterized by a very fast drop that is less than five interspike intervals. In the subsequent sections, we will compare the spike trains from both receptors. Before that, we shall shortly review the relevant results for Shannon entropy.

III. SHANNON ENTROPY AND PREDICTABILITY

The starting point is a symbolic sequence S , consisting of successive symbols (letters) drawn from a finite alphabet \mathcal{A} . Substrings $x^{(n)} = A_1 \cdots A_n$, $A_i \in \mathcal{A}$ of length n are termed n -words or n -blocks and are supposed to appear with a well-defined probability $p(x^{(n)})$ within the infinite sequence S . Following Shannon's approach [9], the n -block entropies H_n are given by

$$H_n := - \sum p(x^{(n)}) \log p(x^{(n)}). \quad (1)$$

The summation is carried out over all n -blocks with nonzero probability $p(x^{(n)}) > 0$. The n -block entropies H_n are a measure of *uncertainty* and give the average amount of information contained in a word of length n . Consequently, the conditional entropies h_n

$$h_n := H_{n+1} - H_n, \quad h_0 := H_1, \quad (2)$$

give the average amount of information required to predict the $(n+1)$ th symbol, when the preceding n symbols are known [22]. Note that h_n is monotonically decreasing: $h_{n+1} \leq h_n$.

Of particular interest is the *entropy of the source* or *limit entropy* h , defined as the limit of the conditional entropy for large n . The convergence of the conditional entropy h_n to its

limit h may be taken as a measure of correlations. If no correlations exist beyond a finite-range m (Markov chain property) the asymptotic value is reached for $n = m$. However, for most systems, the Markov property applies to an approximate description of the system only. The generic case yields an exponential decay of the conditional entropy h_n to its limit [23]. Sequences showing a subexponential decay of the h_n are related to long-range correlations [24,22].

Particular attention must be paid to the estimation of the described quantities for finite sequences of length N . In this case, the n -block entropy H_n becomes systematically underestimated [25,26].

$$H_n^{\text{observed}} = H_n - \frac{M_n - 1}{2N} + O\left(\frac{1}{N^2}\right). \quad (3)$$

Here, l is the size of the alphabet \mathcal{A} and $M_n \leq l^n$ denotes the number of different n -blocks within the sequence.

A. The local predictability

In addition to the conditional entropies h_n , which characterize the average uncertainty, we use the *local predictability* to quantify structure within the interspike interval sequences. At each time step, the local conditional entropy $h_n(x_{\text{prev}})$ gives the uncertainty of the next symbol A_{n+1} , based on the preceding n -block x_{prev} [33] [22,27].

$$h_n(x_{\text{prev}}) := - \sum_{\{A_{n+1}\}} p(A_{n+1}|x_{\text{prev}}) \log p(A_{n+1}|x_{\text{prev}}). \quad (4)$$

Note that the average over all possible prehistories is the previously defined conditional entropy h_n

$$h_n = \langle h_n(x_{\text{prev}}) \rangle = \sum_{\{x_{\text{prev}}\}} p(x_{\text{prev}}) h_n(x_{\text{prev}}). \quad (5)$$

The local predictability or local redundancy $r_n(x_{\text{prev}})$ is given by

$$r_n(x_{\text{prev}}) := 1 - h_n(x_{\text{prev}}). \quad (6)$$

For that we assume that $h_n(x_{\text{prev}})$ is normalized to the size of the alphabet: All logarithms are taken to base $l = 2$. In the following, we refer to all entropy measures as dimensionless quantities in the interval $[0,1]$.

In the next section, we will demonstrate that local measures provide an appropriate tool to identify significant patterns within a sequence, even when its overall structure is close to random.

B. Symbolic dynamics

The application of the aforementioned concepts to the interspike interval sequences requires a coarse-grained representation of the signal, and the results will partly depend on how this coarse graining is performed. There have been several approaches to the discretization of data, the most common being to divide the time series into small bins, assigning the value “1” or “0” to each bin according to the occurrence

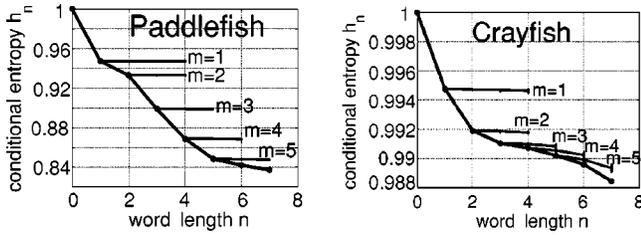


FIG. 1. The conditional entropy h_n as a function of word length n for an interspike interval sequence of the paddlefish (left) and the crayfish (right). The straight lines denote the average conditional entropy obtained from a m th-order Markov process. Both sequences have memory at least up to order $m=5$. Note that the y axis is different in both plots.

or nonoccurrence of a spike [6]. Here, we will use a different approach. Each interspike interval τ_n will be labeled either long (“1”) or short (“0”), depending on whether it is longer or shorter than a given threshold c .

$$\tau_n > c \Rightarrow S_n = 1, \quad \tau_n \leq c \Rightarrow S_n = 0. \quad (7)$$

Naturally, the question arises whether such a crude representation of the data is sufficient to capture the essential structure within the time series.

Previous investigations revealed that one may give heuristic criteria for choosing an optimal threshold-crossing partition [5,28,29], which ensure that the coarse-grained sequence is a reasonable approximation of the dataset. It is understood that the situation here is more complex. Still, we argue that while some microscopic detail of the dynamics may be lost on the symbolic level, most temporal correlations are embedded in the structure of the n -word distributions. Here, this view is mainly supported by the results obtained.

Another crucial point is the timescale on which our analysis is applied. Based on behavioral experiments, an upper bound for the reaction time of the paddlefish is given by $T \approx 125-200$ msec [16]. Given that the mean interspike interval for the paddlefish is $\langle \tau_n \rangle \approx 10-30$ msec, we may speculate that the relevant information is encoded in no more than a few spikes. For the crayfish, reflex reactions may take place in a fraction of this time (though as with all animals, slow adaptive responses occur over an extended period of time) [30,31]. Given these considerations, we restrict ourselves to a study of 3–10 interspike intervals in spontaneous recordings from both crayfish and paddlefish. Our way of partitioning the data allows us to investigate this timescale with high precision and without the need to make further assumptions about the scaling of entropies for finite data. In the subsequent analysis, Eq. (3) was used to monitor the expected deviations.

IV. APPLICATION ON DATA FROM SENSORY NEURONS: RESULTS

A total of 29 paddlefish and 29 crayfish recordings were analyzed. The data were partitioned according to Eq. (7), with the partition threshold optimized to produce a maximal

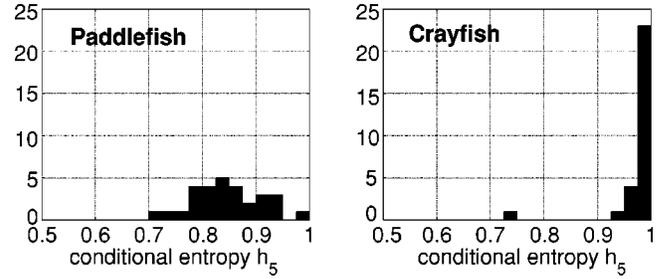


FIG. 2. The histogram of the conditional entropies h_n for a word length $n=5$. Left, paddlefish, right, crayfish.

higher-order conditional entropy $h_{n_{\max}}(c)$. Since each file had a different length (number of spikes) the order n_{\max} up to which the conditional entropy h_n could be estimated differs from file to file. All files allowed for at least order $n=7$.

The result for each file was tested against Markov sequences of given order $m=0,1,\dots,6$. These surrogate data sets were produced as follows: For each shuffle step, two symbols were randomly chosen within the sequence and exchanged only if the m -nearest neighbors of both symbols matched each other. For $m=0$, this corresponds to an unconstrained random shuffle. With $m=1$, two symbols are exchanged only if the left neighbor of the first symbol is identical to the left neighbor of the second symbol and the same also holds for the right neighbors.

By construction, multiple iterations of shuffle steps will preserve the two-block frequency, hence, the conditional entropy h_1 . For higher m , the shuffling preserves the conditional entropy h_n up to order $n=m$, corresponding to an m th-order Markov process. Figure 1 shows a representative result for a paddlefish electroreceptor (left) and a crayfish mechanoreceptor (right). In both cases, the conditional entropy h_n decreases with increasing word length, indicating correlations and long memory in the respective spike trains. However, the magnitude of the decrease is different (Note that in Fig. 1, the y axes of both plots are not identical). The first observation is thus, that the paddlefish sequences pos-

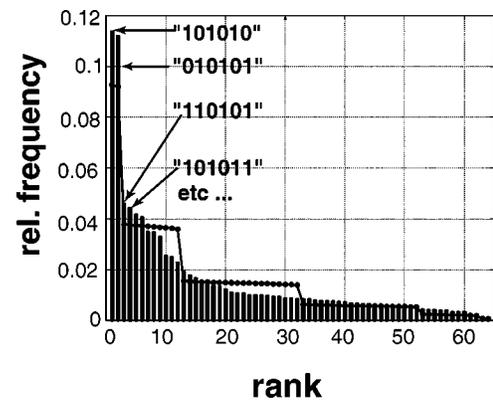


FIG. 3. The Zipf-ordered frequency distribution for six blocks, calculated from a paddlefish sequence. The structure is dominated by the pattern “101010” and “010101.” The bold line denotes the average for surrogates with preserved conditional probabilities $p(A_{n+1}|A_n)$.

TABLE I. Some examples for typical n -blocks found within a (partitioned) paddlefish sequence. For all word length n , the sequence was dominated by patterns with alternating “1”s and “0”s. The first row contains the conditional probabilities, which define a first-order Markov process. The relative frequencies of occurrence of the n -blocks were compared to the respective product of conditional probabilities.

$p(0) \approx$	0.501	n -block	Relative frequency	Expected
$p(1) \approx$	0.499	“101”	0.286	0.269
$p(0 0) \approx$	0.267	“011”	0.089	0.098
$p(1 0) \approx$	0.733	“1010”	0.230	0.197
$p(0 1) \approx$	0.735	“0110”	0.059	0.072
$p(1 1) \approx$	0.265	“1100”	0.033	0.026

sess a larger amount of structure, while the data from the crayfish is more close to random. Figure 2 shows a histogram of the estimated conditional entropies h_5 for all sequences. The clustering of the crayfish data sets at large entropies is clearly seen.

To be more specific about which substructures within the spike trains contribute to this difference, we examine the respective n -word distributions. For the paddlefish ER, the decrease of conditional entropy is mostly due to long-short anticorrelations, already reported for this kind of sensory receptor [32]. Long interspike intervals are often followed by short ones, and vice versa. Consequently, for 20 of 29 sequences the n -block distribution was dominated by the patterns of alternating series of ones and zeros, e.g., “101010” and “010101” for $n=6$.

However, in no case could this finding be explained by a first-order Markov process, incorporating the long-short anticorrelation (see Fig. 3 and Table I). For four sequences, the n -block distribution was dominated by more complicated patterns, e.g., such as “010010” or “011001” for $n=6$. The remaining five sequences have to be treated separately. Here, the structure was due to blocks of repeating ones or zeros, indicating nonstationarity of the time series. This will be discussed in more detail for the crayfish data.

The comparison with higher-order surrogates revealed that except for two, no sequence could be approximated by a Markov process up to order $m=6$ (see Table II).

For the crayfish, the situation is different. Here, 1 of 29 sequences matched a first-order Markov process, seven sequences were compatible with Markov processes up to order $m=6$ (Table II). For 17 of 29 sequences, the *small* deviations from random sequences were mostly due to blocks of repeating ones or zeros. Caution requires us to interpret this as a sign of nonstationarity, almost unavoidable in recordings from living animals [34].

TABLE II. Each column denotes the number of files for which the hypothesis of a m th-order Markov sequence was accepted.

$m =$	0	1	2	3	4	5	6	$m \geq 7$
Paddlefish	0	0	0	1	0	0	1	27
Crayfish	0	1	1	3	1	1	0	22

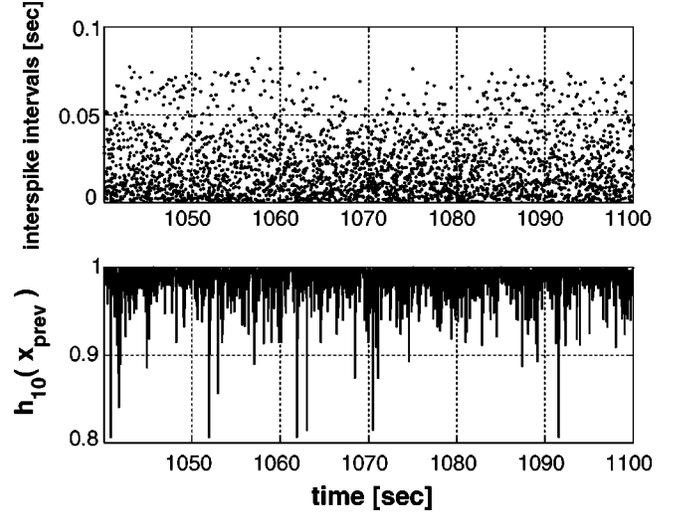


FIG. 4. The local uncertainty $h_{10}(x_{\text{prev}})$ based on the previous ten-block for a recorded interspike interval sequence of the crayfish mechanoreceptor. Upper plot, interspike intervals τ_n ; lower plot, the local uncertainty $h_{10}(x_{\text{prev}})$. For example, the peak with the low uncertainty at time $t \approx 1092$ sec was preceded by the pattern “1110100000.”

Can the crayfish recognize certain words? At least for some crayfish recordings, the deviations from a random sequence must be attributed to nontrivial structure in the n -block distribution.

To emphasize this, we take into account the local uncertainty $h_n(x_{\text{prev}})$, as defined in Eq. (4). As an example, Fig. 4 shows the local uncertainty $h_{10}(x_{\text{prev}})$ of crayfish interspike intervals based on the previous ten-block. Although the overall structure of the spike train is almost completely random, rare events still stand out: we observe a few particular individual patterns, after which the sequence is significantly

TABLE III. The local uncertainty $h_{10}(x_{\text{prev}})$ for a crayfish sequence. Also included are slight variations of the first pattern (flipped symbols are depicted in bold). The results for the pattern with the lowest local uncertainty stay reasonably stable when restricting the sequence length to \tilde{N} .

x_{prev}	$p(1 x_{\text{prev}})$	$p(0 x_{\text{prev}})$	$h_n(x_{\text{prev}})$
“1110100000”	0.75	0.25	0.806
“1011100000”	0.75	0.25	0.815
“0000011111”	0.27	0.73	0.840
“0001100000”	0.72	0.28	0.856
“0110000000”	0.71	0.29	0.863
\vdots	\vdots	\vdots	\vdots
“0110100000”	0.59	0.41	0.979
“1110110000”	0.57	0.43	0.986
“1110100001”	0.46	0.54	0.996
“1110100000”	-	-	-
$\tilde{N} = \frac{2}{3}N$	0.75	0.25	0.811
$\tilde{N} = \frac{1}{2}N$	0.72	0.28	0.852

more predictable than on average (see Table III). To ensure that the occurrences of these patterns are not simply fluctuations, as expected in a random process, we have to test the results against their counterparts in shuffled sequences. Since we deliberately choose the patterns with the lowest uncertainty, we must expect them to lie on the edge of their respective distributions. Most conventional tests would therefore assign to them a spurious significance. Still, our results show that the local uncertainties given in Table III lie well outside their distributions, obtained from an ensemble of first-order Markov surrogates.

A more faithful indicator of significance may be applied to the pattern with the lowest local uncertainty ‘‘1110100000.’’ Within an ensemble of 499 first-order Markov surrogates, no pattern with less or equal $h_n(x_{\text{prev}})$ was found. Another crucial point is the question of nonstationarity. As could be observed, all patterns in Table III share the common feature of a final block of repeating ones or zeros. As a first indication that the high predictability is not solely due to the block of repeating zeros, Table III also contains the results for slight variations of the pattern ‘‘1110100000.’’ Moreover, the results stay reasonably stable when restricting the analysis to a part of the original recording.

Similar results were obtained for other crayfish recordings. Examples of a pattern with the lowest local uncertainty for *different* files are: ‘‘1111011000,’’ ‘‘1000001010,’’ ‘‘0000010011,’’ and ‘‘0000001010.’’ However, due to the finite length of the datasets and nonstationarity, the number of examples is limited. Thus, the claim that *all* files have this property would not be justified by the data.

While it is beyond the scope of this paper to argue whether these patterns of low uncertainty are of biological relevance, we will speculate on how they might influence the balance between order and disorder. Moreover, each such specific pattern corresponds to a precise time in the interspike interval sequence, making experimental tests about their relevance possible.

V. DISCUSSION

In the first section, we have compared the overall structure of spike trains from the paddlefish electroreceptor with those from the crayfish mechanoreceptor. Both yield a different conditional entropy, the former showing a high degree of structure, while the latter is, apart from nonstationarity, almost random. Up to here, our findings support earlier reports on correlations and memory found in different kinds of spike trains [3,4]. In particular, for the paddlefish ER, short-term anticorrelations were reported in [17].

Regarding the comparison of two different kinds of sen-

sory receptors, we want to emphasize that our analysis does not aim at measuring the information transfer in neuronal signals. Our study was restricted to spontaneous activity, containing no information about external stimuli at all. Still, (anti-)correlations in spontaneous activity do have an influence on the detectability of weak changes in the interspike interval probability distribution [3]: A weak signal is potentially difficult to detect, given a relatively high variability and the lack of correlations in the spontaneous activity. Correlations will decrease the possibility that small dynamical changes in the spike activity are obscured by statistical fluctuations, but maybe at the cost of reducing the signals capacity to allow for an effective encoding of different stimuli.

While the correlation between the existence of extended memory in the afferent spike train and the sensitivity of the receptive systems is presently unknown, we may note that the juvenile paddlefish feeds on individual *Daphnia*, which produce oscillations at about 5 – 10 Hz. The electrosensory system is tuned to this frequency (shown by both the behavioral experiments and electrophysiology) [14–17].

On the other hand, the crayfish mechanoreceptor array lacks a primary specialization. The sensitivity of single mechanoreceptors may vary considerably over a broad range from 1 Hz to 40 Hz [21].

It is understood that general conclusions may not be drawn from a study of peripheral systems of only two animals. Our comparison focuses on characterizing the variability of the spontaneous activity of both sensory systems. In particular, we want to emphasize that entropy measures, such as the local uncertainty, are also applicable to (almost) random spike trains and to situations in which the animal must make a decision on the basis of a finite number of interspike intervals.

As demonstrated, even in almost random sequences, rare events can still stand out. Thus, the crayfish may indeed be able to recognize certain ‘‘words’’ even though they appear rarely in its neural sequences. To this end, we have proposed the local uncertainty, which quantifies the predictability of *individual* spikes, as an additional tool.

ACKNOWLEDGMENTS

We are grateful to Lon Wilkens for an edifying discussion. This work was supported by the Office of Naval Research, Physics Division; the Deutsche Forschungsgemeinschaft (Sfb555); and the Fetzer Institute (A.N.). A.N. and D.R. acknowledge support from the University of Missouri Research Board. F.M. acknowledges support from the Alexander von Humboldt Foundation.

-
- [1] S.P. Strong, R. Koberle, R.R. de Ruyter van Steveninck, and W. Bialek, *Phys. Rev. Lett.* **80**(1), 197 (1998).
 [2] M.C. Eguia, M.I. Rabinovich, and H.D.I. Abarbanel, *Phys. Rev. E* **62**(5), 7111 (2000).
 [3] R. Ratnam and M.E. Nelson, *J. Neurosci.* **20**(17), 6672 (2000).
 [4] M.J. van der Heyden, C.G.C. Diks, B.P.T. Hoekstra, and J.

- DeGoede, *Physica D* **117**, 299 (1998).
 [5] R. Steuer, L. Molgedey, W. Ebeling, and M.A. Jimenez-Montaño, *Eur. Phys. J. B* **19**, 265 (2001).
 [6] F. Rieke, D. Warland, R. R. de Ruyter van Steveninck, and W. Bialek, *Spikes* (Bradford Books, MIT Press, Cambridge, 1997).

- [7] M.A. Jiménez-Montaño, H. Penagos, A. Hernandez Torres, and O. Diez-Martínez, *Biosystems* **58**, 117 (2000).
- [8] M.J. Chacron, A. Longtin, and L. Maler, *J. Neurosci.* **21**, 25 (2001).
- [9] C.E. Shannon, *Bell Syst. Tech. J.* **27**, 379 (1948); *ibid.* **27**, 623 (1948).
- [10] L. Grande and W.E. Bmis, *J. Vertebrate Paleontology* **11**, 1 (1991).
- [11] J.M. Jorgenson, A. Flock, and J. Wersall, *Z. Zellforsch Mikrosk Anat.* **130**, 362 (1972).
- [12] T.H. Bullock, *Annu. Rev. Neurosci.* **5**, 121 (1982).
- [13] H.A. Braun, H. Wissing, K. Schäfer, and M.C. Hirsch, *Nature (London)* **367**, 270 (1994).
- [14] L.A. Wilkens, D.F. Russell, X. Pei, and C. Gurgens, *Proc. R. Soc. London, Ser. B* **264**, 1723 (1997).
- [15] A. Neiman, D. F. Russell, X. Pei, W. Wojtenek, J. Twitty, E. Simonotto, B. Wattering, E. Wagner, L. Wilkens, and F. Moss, *Int. J. Bifurcation Chaos Appl. Sci. Eng.* **10(11)**, 2499 (2000).
- [16] D. F. Russell, A. Tucker, B. A. Wetrting, A. Neiman, L. Wilkens, F. Moss, *Fluct. Noise Lett.* (2001) (to be published).
- [17] A. Neiman and D. F. Russell, *Phys. Rev. Lett.* **86**, 3443 (2001).
- [18] R. C. Brusca and C. J. Brusca, *Invertebrates* (Sinauer Press, Sunderland, 1990).
- [19] L. Wilkens and J. Douglass, *J. Exp. Biol.* **189**, (1994).
- [20] J.K. Douglass, L. Wilkens, E. Pantazelou, and F. Moss, *Nature (London)* **365**, 337 (1993).
- [21] E. Pantzelou, C. Dames, F. Moss, J. Douglass, and L. Wilkens, *Int. J. Bifurcation Chaos Appl. Sci. Eng.* **5**, 101 (1995).
- [22] W. Ebeling, *Physica D* **109**, 42 (1997).
- [23] W. Ebeling and M.A. Jiménez Montaño, *Math. Biosci.* **52**, 53 (1980).
- [24] T. Pöschel, W. Ebeling, and H. Rosé, *J. Stat. Phys.* **80(5/6)**, 1443 (1995).
- [25] H. Herzel, A.O. Schmitt, and W. Ebeling, *Chaos, Solitons Fractals* **4(1)**, 97 (1994).
- [26] P. Grassberger, *Phys. Lett. A* **128**, 369 (1988).
- [27] L. Molgedey and W. Ebeling, *Physica A* (2000).
- [28] J.P. Crutchfield and N.H. Packard, *Physica D* **7**, 201 (1983).
- [29] W. Ebeling, R. Steuer, and M. Titchener, *Stochastics and Dynamics* **1**, 45 (2001).
- [30] F. Krasne and S.C. Lee, *J. Neurosci.* **8(10)**, 3703 (1988).
- [31] D.H. Edwards, W.J. Heitler, and F.B. Krasne, *Trends Neurosci.* **22(4)**, 153 (1999).
- [32] P. Teunis, F. Bretschneider, J. Bedaux, and R. Peters, *Neuroscience* **41(2-3)**, 809 (1991).
- [33] Here, the index n is omitted. The prehistory x_{prev} always refers to the previous n symbols.
- [34] With the term “nonstationarity,” we refer to a slight systematic increase of the interspike intervals. In our analysis, this would lead to an overrepresentation of “short” intervals in the first part of the sequence.