



Universidade de São Paulo

Biblioteca Digital da Produção Intelectual - BDPI

Departamento de Física e Ciência Interdisciplinar - IFSC/FCI

Artigos e Materiais de Revistas Científicas - IFSC/FCI

2008-02

A complex biological system: the fly's visual module

Philosophical Transactions of the Royal Society A, London, v. 366, n. 1864, p. 345-357, 2008
<http://www.producao.usp.br/handle/BDPI/49103>

Downloaded from: Biblioteca Digital da Produção Intelectual - BDPI, Universidade de São Paulo

A complex biological system: the fly's visual module

BY MURILO S. BAPTISTA^{1,†}, LIRIO O. B. DE ALMEIDA²,
JAN F. W. SLAETS², ROLAND KÖBERLE^{2,*} AND CELSO GREBOGI³

¹*Institut für Physik Am Neuen Palais 10, Universität Potsdam,
14469 Potsdam, Germany*

²*Instituto de Física, Universidade de São Paulo, C.P. 369,
3560-970 São Carlos, São Paulo, Brazil*

³*Centre for Applied Dynamics Research, King's College,
University of Aberdeen, Aberdeen AB24 3UE, UK*

Is the characterization of biological systems as *complex systems* in the mathematical sense a fruitful assertion? In this paper we argue in the affirmative, although obviously we do not attempt to confront all the issues raised by this question. We use the fly's visual system as an example and analyse our experimental results of one particular neuron in the fly's visual system from this point of view. We find that the motion-sensitive 'H1' neuron, which converts incoming signals into a sequence of identical pulses or 'spikes', encodes the information contained in the stimulus into an alphabet composed of a few letters. This encoding occurs on multilayered sets, one of the features attributed to *complex systems*. The conversion of intervals between consecutive occurrences of spikes into an alphabet requires us to construct a *generating partition*. This entails a one-to-one correspondence between sequences of spike intervals and words written in the alphabet. The *alphabet dynamics* is multifractal both with and without stimulus, though the multifractality increases with the stimulus entropy. This is in sharp contrast to models generating independent spike intervals, such as models using Poisson statistics, whose dynamics is monofractal. We embed the support of the probability measure, which describes the distribution of words written in this alphabet, in a two-dimensional space, whose topology can be reproduced by an M-shaped map. This map has positive Lyapunov exponents, indicating a chaotic-like encoding.

Keywords: complex systems; neuron dynamics; multifractality; dynamical systems

1. The fly's optical tract as a complex system

The neural system of any living creature, the fly in particular, is certainly a complex system. In fact, it consists of individual neurons¹ joined into modules, which interact with themselves and with the external world. This multiple-component

* Author for correspondence (rk@if.sc.usp.br).

† Present address: Max-Planck-Institut für Physik Komplexer Systeme, Nöthnitzerstrasse 38, 01187 Dresden, Germany.

¹ Even a neuron itself may be considered as such a system.

One contribution of 15 to a Theme Issue 'Experimental chaos I'.

system evolves and adapts as a consequence of these internal and external dynamical interactions. The interacting components lead to hierarchical structures with different causations at different levels. In an experiment we usually only observe causations in a few levels, in our case at the ‘last’ neuron. There are various mathematical characterizations of a *complex system* (Grassberger 1986; Poon & Grebogi 1995; Bialek et al. 2001). We do not attempt to satisfy all the criteria spelt out there, but suggest that the encoding of external stimuli in a part of the fly’s optical tract takes place in *multilayered sets* (Baptista et al. 2006).

Roughly, in the optical system of the fly, the incoming information is processed first by the photoreceptors and then by various modules—the lamina, medulla, lobula and lobula plate. From here the information goes to the sensory system. We record from one neuron, the ‘H1’, located in the lobula plate. H1 encodes the incoming signals into a sequence of identical pulses, termed action potentials or spikes. If such a spiking neuron has to encode *relevant* features of the stimulus, it has to fire precisely timed spikes depending on the presynaptic input generated by the instantaneous stimuli the organism receives. Our experimental results strongly suggest that this encoding takes place on multilayered sets, a characteristic of the complex nature of this system. These sets are defined in terms of symbolic sequences of letters, selected from an alphabet according to the size of spike time intervals. Furthermore, we present strong evidence that the underlying dynamics (UD) on each layer is multifractal, thus allowing for a chaotic-like type of encoding. Already in the no-stimulus regime, owing to the multifractal dynamics, the UD is highly flexible, ready to adapt its dynamics to the statistical properties of the stimulus (Fairhall et al. 2001) to be encoded. Then, in the presence of the stimulus, finely tuned spike times ride on a set whose UD has now an increased multifractality, shaped by the properties of the stimulus. We also show that a two-letter alphabet is required for the UD to represent the no-stimulus dynamics of spontaneous firing, while a four-letter alphabet is necessary for the UD to encode the information contained in the stimulus.

2. Multifractality

We study this underlying dynamical regime in a prominent example of a spiking neuron: the H1 neuron, located in the lobula plate of the fly, *Chrysomya megacephala*. This area is several layers back from each compound eye of the fly. It includes large motion-detector neurons with wide receptive fields and pronounced direction selectivity. The two H1 neurons, one for each lobula plate, are sensitive to image motion around a vertical axis (Hausen 1981). Their response is excited by horizontal back-to-front motions, generated by rotations around a vertical axis and suppressed by motions in the opposite direction. One of these neurons was stimulated by a computer-controlled random, vertical bar pattern with horizontal velocity $v(t)$. Here we discretize time in bins of 2 ms, which is roughly the refractory period of the H1 neuron. The fly therefore saw a new frame on the monitor every $\delta t = 2 \text{ ms}$,² whose change in position δx was given by $\delta x(t) = v(t)\delta t$.

²The experiments were performed in *DipteraLab*, Instituto de Física, São Carlos, Brazil. Flies were immobilized with wax and viewed an image displayed on a Tektronix 608 monitor from a distance of 12 cm, the setup being similar to that of Fairhall et al. (2001). The light intensity corresponds roughly to that seen by a fly at dusk.

In order to be minimally representative, we selected four types of stimuli $v(t)$: S_0 , no stimulus; S_1 , constant velocity; S_2 , a stimulus generated by an Ornstein–Uhlenbeck (Uhlenbeck & Ornstein 1930) process with correlation time $\tau_c = 20$ ms;³ S_3 , an uncorrelated ($\tau_c = 0$) Gaussian stimulus.

In each experiment, such a spiking neuron generates a sequence of spike times t_i , $i = 1, 2, 3, \dots, N_s$, where $N_s \sim 10^5$ in our experiments. All the information received is compressed into the sequence of intervals $\Delta t_i = t_{i+1} - t_i$. In order to extract the UD, we classify all the possible intervals into a discrete set of cardinality N , depending on their size: $\Delta t_i \leq d_1$, $d_1 < \Delta t_i \leq d_2$, $d_2 < \Delta t_i \leq d_3$, etc., where d_j , $j = 1, 2, \dots, N-1$, are a set of *dividers*. Evidently, if we make this set large enough, we recover the original intervals, when $d_j - d_{j-1}$ approaches 2 ms. The question is: can we choose a reasonably small set of dividers without compromising the UD of the original intervals Δt_i ? In other words, *what is the size of the alphabet the fly's H1 neuron uses to speak postsynaptically?*

We choose our dividers so as to maximize Shannon's entropy (Strong *et al.* 1998). Equivalent results are obtained, although with larger errors, if we minimize the information the spike train provides about the stimulus. For a given set of $N-1$ dividers, we study the dynamics of our sequence of spike intervals Δt_i converted into a sequence of words of length L containing N letters. Note that an L -letter word may comprise a very long time interval. Now count up all the different sequences showing up in an experiment, get their probabilities P_k and compute the entropy per letter of a word of length L with N possible letters of the experimental sequence

$$H(L, N) = -\frac{1}{L} \sum_k P_k \log_2(P_k). \quad (2.1)$$

Figure 1a shows the entropy for $L=10$ with only one divider for the four different datasets. The entropy shows a maximum at

$$[d_1(S_0), d_1(S_1), d_1(S_2), d_1(S_3)] = [23, 6, 5, 5]$$

bins. We now construct a uniquely defined *generating* (Plumecoq & Lefranc 2000a,b) partition of our time intervals. *Generating* means that our alphabet has to be able to allow a one-to-one encoding of all possible sequences of intervals. It will turn out to be the partition with $N = 2 \times 2^l = 2, 4, 8, \dots$. This is a highly non-trivial endeavour. Even for such well-studied systems as the Henon map (Grassberger & Kantz 1985), this question has not been settled. In order to search for such a generating partition, we require the following consistency requirement to be satisfied: *in order to be generating, all the dividers of a particular layer J must already be contained in the previous layers $j = 1, 2, \dots, J-1$* . If this were not true, conflicting information would result as we go from layer $J-1$ to layer J . Dividers are extracted from the data using always $L=1$. We will see that, within errors, only one partition satisfies this requirement. We now construct table 1 in two steps. Firstly, maximizing the entropy $H_2(N = 2 \times 2^l)$ for $N = 2, 4, 8, \dots$,⁴ we get the respective dividers in

³The fly's reaction time is of the order of 20 ms, and this is thus a behaviourally relevant time scale.

⁴The subscript '2' in $H_2(N)$ indicates that we start our search for dividers using a sequence beginning with this prime number.

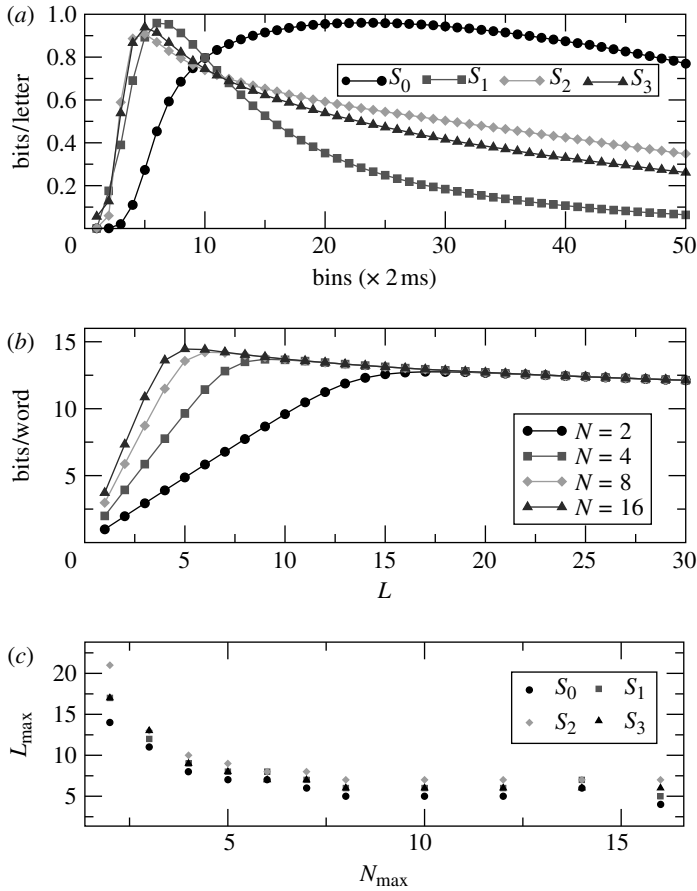


Figure 1. (a) Entropy versus divider for a two-letter code, for the four datasets: S_0 , S_1 , S_2 and S_3 . The entropy maximum gives the best divider d_1 . (b) Word entropy ($H_{\max}(L, N) \times L$) versus the length of words L for the data S_1 . Note that for each N there is a value of L for which the entropy reaches its maximum. We refer to these N and L values as N_{\max} and L_{\max} . Each curve corresponds to a different number of letters. (c) Scaling law that relates N_{\max} with L_{\max} for which the maximum of the entropy is achieved for all the datasets.

table 1. Thus, for experiment S_0 , we found the dividers: 23; 9, 23 and 49; and 6, 14, 23, 34, 49 and 71. Secondly, we turn to the remaining values of N , searching for dividers satisfying our consistency condition. This is indeed possible, if we fix, for example, the bold-faced dividers and perform a constrained maximization to obtain the remaining ones. We obtain the new S_0 dividers: 9, 71, 34 and 14. Note that the divider 9, obtained using two letters, arose in the sequence starting with $N=2$ only when we used four letters. The corresponding entropy per letter $H_2(N=3, 5, 6, 7, \dots)$ equals to within 5% the entropy obtained with an unconstrained maximization, which yields completely different dividers. We conclude that $H_2(N) = 2 \times 2^l = 2, 4, 8, \dots$ does indeed satisfy our consistency condition to be a generating partition. Is it unique? If we perform the same procedure for other partitions, e.g. $N = 3 \times 2^l = 3, 6, 12, \dots$, the first step yields the same value for the unconstrained entropy maxima $H_3(N = 3, 6, 12, \dots)$. In the second step, to obtain the remaining dividers, the

Table 1. The best time dividers (rows 3, 4, ...) in units of bins, which generate the entropies H_2 for all datasets, considering a different number of letters N (second row). (All dividers are calculated using $L=1$.)

	S_0								S_1								S_2								S_3							
	2	3	4	5	6	7	8	9	2	3	4	5	6	7	8	9	2	3	4	5	6	7	8	9	2	3	4	5	6	7	8	9
N	2	3	4	5	6	7	8	9	2	3	4	5	6	7	8	9	2	3	4	5	6	7	8	9	2	3	4	5	6	7	8	9
d_1	23	9	9	9	9	9	6	6	6	6	4	4	4	4	4	3	3	5	5	5	3	3	3	3	3	3	3	3	3	3	3	3
d_2	—	23	23	23	23	14	9	—	—	10	6	6	6	5	4	—	—	5	5	4	4	4	4	3	2	5	4	5	4	3	—	5
d_3	—	—	49	49	34	23	14	—	—	10	10	8	6	5	—	—	—	11	5	5	5	5	4	—	—	—	9	—	—	—	—	
d_4	—	—	—	71	49	34	23	—	—	—	—	15	10	8	6	—	—	—	11	11	7	7	5	—	—	—	—	—	—	—	—	
d_5	—	—	—	—	—	71	49	34	—	—	—	—	15	10	8	—	—	—	—	34	11	7	7	—	—	—	—	—	—	—	—	
d_6	—	—	—	—	—	—	71	49	—	—	—	—	—	15	10	—	—	—	—	—	34	11	7	—	—	—	—	—	—	—	—	
d_7	—	—	—	—	—	—	—	71	—	—	—	—	—	—	15	—	—	—	—	—	—	—	34	—	—	—	—	—	—	—	—	

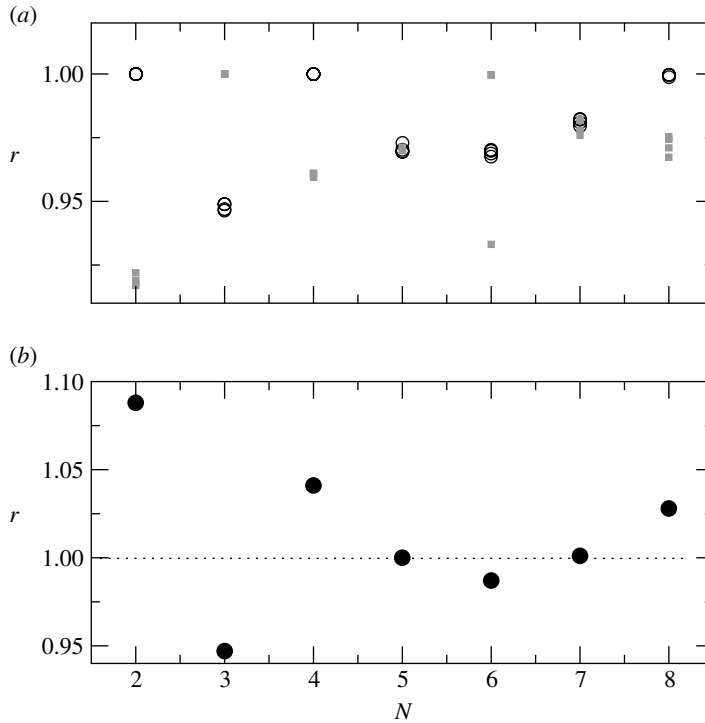


Figure 2. (a) Entropy per letter normalized by $\log_2(N)$ for all datasets. Circles represent the entropy $H_2(N)$ (bits per letter)/ $\log_2(N)$, and squares, the entropy $H_3(N)$ (bits per letter)/ $\log_2(N)$ for all datasets. (b) Ratio of $r = \langle H_2 \rangle / \langle H_3 \rangle$, where $\langle \rangle$ indicates the average over S_0, \dots, S_3 .

constrained maximizations yield smaller entropies $H_3(N = 2, 4, 5, 7, 8, \dots)$ by as much as 10% as compared with the unconstrained ones (as shown in figure 2). Here, $H_3(N=2)$ means: (i) maximize $H_3(N=3)$ to get $[d_0, d_1]$ and (ii) maximize $H_3(N=2)$, using either d_0 or d_1 , whichever gives the larger entropy. The only exception here is one point corresponding to $N=3$. The same is also true for other partitions, starting with prime numbers 5 and 7, thus indicating that within errors the binary partition is the only generating one.

Figure 1c shows that the number of letters needed to maximize the entropy decreases with increasing word length L , due to correlations manifesting themselves and undersampling effects. Since these put a limit to the size of the alphabet usable, we address the UD using just a few letters. We note, however, that even for N smaller than N_{\max} , e.g. seven dividers for experiments S_2 and S_3 , we obtain essentially the information the original spike trains convey about the stimulus (Strong et al. 1998). Figure 3 shows the information per letter I_A for S_2 , divided by the information the original spike trains convey about the stimulus, using the dividers of table 1. In other words, for $N = 2, 3, 4, 5 \dots$, we use the dividers $[5]$, $[3, 5]$, $[3, 5, 11]$, $[3, 4, 5, 11]$, The error bars shown are due to systematic errors in the computation of the *noise entropy*, i.e. the entropy of the spike train, given the stimulus. In the conversion of spike times into an alphabet, the precise timing relation of the stimulus to a particular letter is partially lost and this fact is responsible for the error bars.

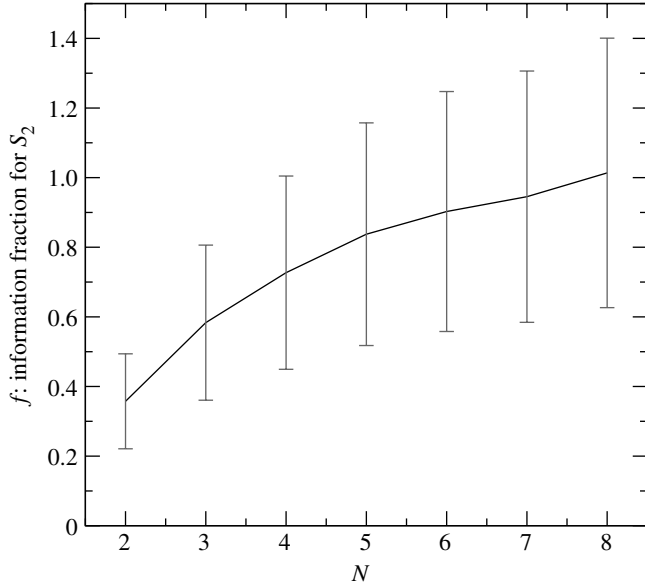


Figure 3. Information per letter I_A the alphabet conveys about the stimulus for S_2 , divided by the average information per spike I_S : $f = I_A/I_S$. The computation of the noise entropy H_{noise} , where $I_A = H - H_{\text{noise}}$, introduces the large error bars. Our data are extrapolations to $L \rightarrow \infty$ (Strong *et al.* 1998). Experiment S_3 shows the same trend.

Since we are going to extrapolate to L large, we focus on sequences of words of length L containing four letters, using the best dividers of table 1. Only fine details of our results depend on this choice.

We now encode each of these L -letter words in a one-to-one map into a real number W_i , $0 \leq W_i \leq 1$,⁵ whose frequencies we count in order to compute their probabilities p_i . The structure of the space of sequences $\mathcal{C}(W)$ for a given N may now be uncovered by computing the generalized dimensions⁶ D_q (Renyi 1971; Grassberger & Procaccia 1983b; Hentschel & Procaccia 1983). These are logarithmic ratios between the probabilities p_i and their physical occupation ϵ , which in our space \mathcal{C} is given by $\epsilon = N^{-L}$. The index q can be thought of as a filter: a larger q enhances this ratio for large probabilities, whereas a negative one emphasizes the smaller probabilities,

$$D_q = \lim_{L \rightarrow \infty} \frac{\ln(\sum_i p_i^q)}{(1-q)L \ln(N)}. \tag{2.2}$$

An equivalent quantity is f_α , the Legendre transform of $(1-q)D_q$. The index α measures the possible local fractal dimensions (Grassberger & Procaccia 1983a; Halsey *et al.* 1986) in our space \mathcal{C} , occurring with singularity strength f_α . This is

⁵ Our results are independent of the particular encoding used, as long as it is one-to-one.

⁶ Note that, since no metric characterization of the space $\mathcal{C}(W)$ is performed, it is irrelevant to the way we encode words into real numbers to calculate the generalized dimensions.

the global dimension of the set of points, which locally scales with singularity strength α . They are given by $\alpha = d[(q-1)D_q]/dq$, $f_\alpha = q\alpha - (q-1)D_q$.⁷

In figure 4a,b, we show the spectra of the symbolic sequences in $\mathcal{C}(W)$ with their error,⁷ which allow us to draw the following conclusions. The H1 neuron has a *multifractal character*, exhibiting the existence of an infinite number of dimensions α with densities f_α . This is in sharp contrast to a memoryless, uncorrelated spike train with a Poissonian or similar probability distribution. Any distribution with independent increments yields a non-trivial f_α for suboptimal dividers, but is really monofractal for the optimal ones with $L \ln(N)D_q = \log_2(N)$, independent of q . In fact, for $N=2$, let p_0/p_1 be the probabilities for an interval being smaller/greater than d_1 . Then $p_1(\text{Poisson}) = e^{-\lambda d_1}$, where λ is the spike rate. Then $L \ln(N)D_q = \log_2(p_0^q + p_1^q)/(1-q)$. Maximizing the entropy gives $p_0 = p_1 = 1/2$ and $e^{-d_1(\text{Poisson})\lambda} = 1/2$, yielding $L \ln(N)D_q = \log_2(2)$.

Real spike trains by contrast are multifractal even for the best dividers for all types of stimuli. The fractal dimensions are roughly the same for all the datasets, since $D_0 = f_1 = 1$. This means that the neuron's dynamics has continuous support on \mathcal{C} , the probability measure being distributed without 'holes'. The spectra's shape— $f_{\alpha_{\min}} = 0$ and $f_{\alpha_{\max}} > 0$ (except for S_3)—indicates that \mathcal{C} has at least two scales and two main components: high-probability sets, hot spots localized in small portions of the symbolic space \mathcal{C} with density $f_{\alpha_{\min}}$; and another low-probability set, cold spots spread out all over \mathcal{C} with density $f_{\alpha_{\max}}$. In the set associated with S_0 , the number of cold spots very much dominates— $f_{\alpha_{\max}} \sim \alpha_{\max} \sim 1$ —the hot ones, implying one dominant scale. For the other datasets, the scales are comparable. Therefore, as the dynamics of the stimulus becomes faster as we go from $\tau_c = \infty$ to $\tau_c = 0$, the suppressed scale emerges.

Given the f_α spectra of our data, what is the simplest set with this spectrum? To address this question, we construct a probabilistic two-scale set with the following rule: an interval of length unity is divided into $b+1$ equal pieces, such that one piece receives p_0 of the original measure and the remaining b receive p_1 each, with $p_0 + bp_1 = 1$. Iterating this process self-similarly yields a set of dimension

$$D_q = \frac{\ln(p_0^q + bp_1^q)}{(1-q)\ln(1+b)}, \quad (2.3)$$

with $p_0 = (1+b)^{-\alpha_{\min}}$ and $p_1 = (1-p_0)/b$, where b is adjusted to produce the correct value for $f_{\alpha_{\max}}$ in figure 4b. The resulting f_α spectra are shown in figure 4c,d.

Note that b jumps to lower values, once a stimulus is turned on. This means that for no stimulus the dynamics distributes the measure rather uniformly into 31 equal intervals, where 30 of them receive p_1 of the measure and one receives p_0 with $p_0 > p_1$. At the r th iteration of the set, 30 intervals will contain p_1^r of the measure (cold spots), one interval will contain p_0^r of the measure (hot spots), i.e. virtually all of the measure, and the rest will receive combinations proportional to $p_0^{r-k} p_1^k$, $0 \leq k \leq r$. When stimulated, the number of pieces drops dramatically to approximately one, the hot spots approximately balancing the cold ones. Stimuli

⁷ D_q can be reliably computed, if there are linear regimes, when plotting the numerator versus the denominator of equation (2.2). The q values were selected so as to produce D_q values with a variance of no larger than 5% relative to the absolute value (this is obtained from the residual of the linear fitting) and which produce an error in f_α of no more than 10% relative to the absolute value, except for some low-probability sets, which have larger errors.

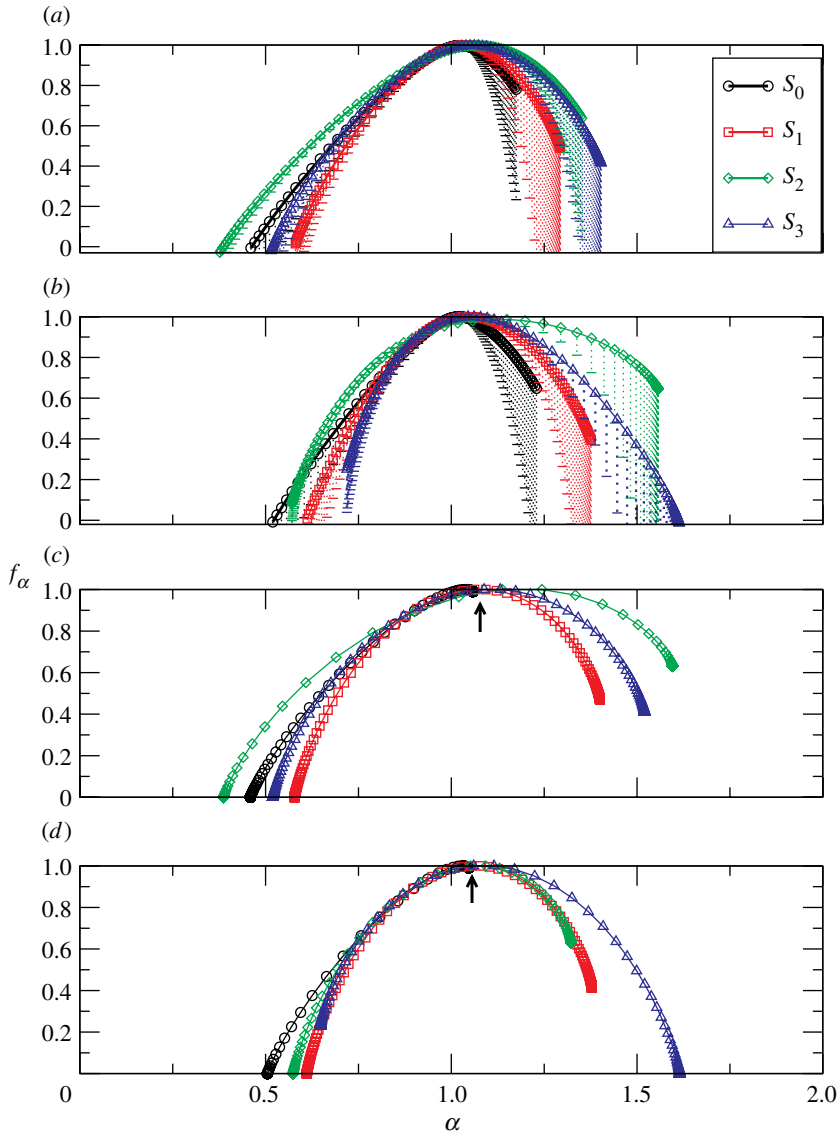


Figure 4. The f_α spectra using table 1: (a) two letters and (b) four letters. To avoid cluttering, we show only one-sided error bars as dashed lines. Spectra analogous to (a,b), using equation (2.3): (c) for two letters with $b_{S_0;S_1;S_2;S_3} = [30; 1.55; 2.0; 1.44]$ and (d) for four letters with $b_{S_0;S_1;S_2;S_3} = [30; 1.45; 2.0; 1.2]$. The arrows indicate the endpoint α_{\max} of the curve for S_0 .

thus reshape the probability landscape, which becomes more and more structured as the stimulus entropy increases.

3. The M map and a two-dimensional symbolic space

The uncovering of a neural code, the rule by which a stimulus is encoded into a sequence of spike times, is one of the central issues in neuroscience. Here we address a much simpler, though preliminary, issue: finding a *pseudo-code*, which

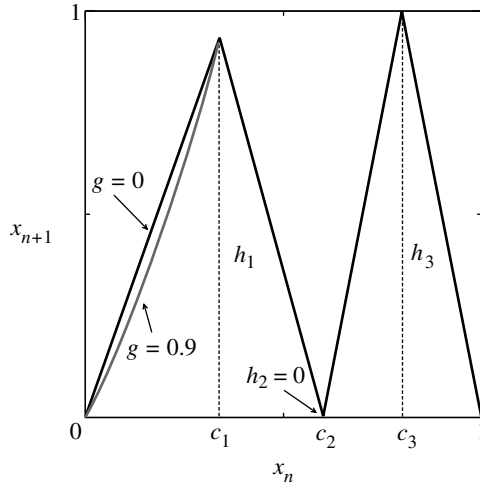


Figure 5. The M mapping. A piecewise *almost* linear map with four distinct domains. For $x_{n+1} > 1$ take $x_{n+1} = 2 - x_{n+1}$. Each domain is associated with a symbol.

generates sequences with four letters, having the same probability distributions and the same f_α spectra as the data.

We propose a piecewise almost linear map of the form shown in figure 5.⁸ It was selected out of several dozens of multiparameter maps, which did not work. The parameters (c_1 , c_2 and c_3) of the M map were adjusted, starting from initial divider intervals set proportional to the respective letter probabilities. It models the most probable words, meaning that, for stimuli $[S_0, S_1, S_2, S_3]$, we only consider words appearing more than $[4, 20, 20, 25]$ times, discarding words with frequency approximately less than 0.012%. Note that cold spots are sensitive to low-probability words. The nonlinearity for $0 \leq x_n \leq c_1$ models the prominent singularity due to words mostly composed of ‘0’s in dataset S_0 and is given by

$$x_{n+1} = x_n(h_1/c_1 + (1/c_1)^g x_n^g - 1), \quad 0 \leq x_n \leq c_1, \tag{3.1}$$

with $g=0.9$. The jump in b (see figure 5) is reflected in this nonlinearity. If $g=0$, the map is piecewise linear.

We create a symbolic sequence from a trajectory of the M map, associating letters $\beta_1, \beta_2, \beta_3, \beta_4$, if $x_n \in [0, c_1[, [c_1, c_2[, [c_2, c_3[, [c_3, 1]$, respectively, to real numbers.

The extraction of the second layer’s equivalent dynamics is one of the basic steps in analysing this system within the framework of *dynamical systems theory*. We therefore convert the dynamics of our sequence of spike intervals Δt_i into a sequence of words of length $L=6$ containing $N=4$ letters and embed our sequences into a Cartesian two-dimensional space⁹ $\mathcal{G} = W_i \otimes W_{i+1}$ for visualization. The symbolic spaces for the sequences generated by the M model with the parameters of table 2 are shown in figure 6e–h. The symbolic

⁸Notice that h_3 may be greater than 1.

⁹The false nearest neighbour technique (Kantz & Schreiber 1997) indicates an embedding dimension $d_e \sim 3-6$.

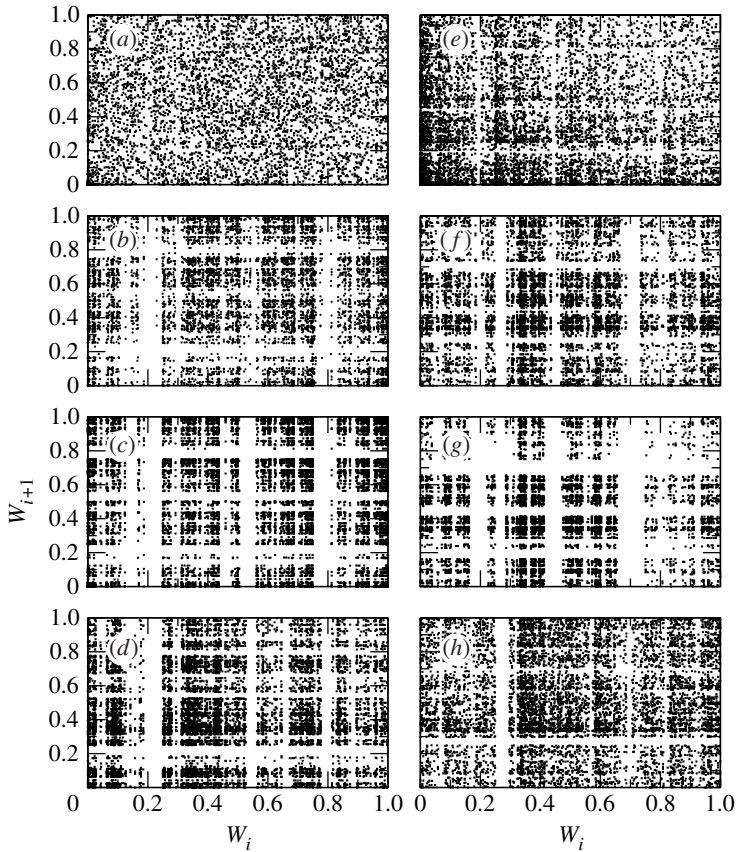


Figure 6. Symbolic spaces \mathcal{G} for $N=4$ letters and $L=5$: (a) S_0 , (b) S_1 , (c) S_2 and (d) S_3 for the original data; (e) S_0 , (f) S_1 , (g) S_2 and (h) S_3 for the M mapping with parameters from table 2.

Table 2. Parameters of the M mapping (figure 5), whose sequences generate the symbolic spaces of figure 6. (Parameters adjusted in order to reproduce both the spectra in figure 4a,b and \mathcal{G} in figure 6a-d.)

dataset	c_1	c_2	c_3	h_1	h_2	h_3	g
S_0	0.337	0.6	0.8	0.93	0	1.0	0.9
S_1	0.3	0.6	0.8	0.9	0.25	1.1	0
S_2	0.3	0.6	0.8	0.85	0.25	1.0	0
S_3	0.3	0.6	0.8	1.0	0.25	1.2	0

sequences extracted from the original data using four letters, embedded in \mathcal{G} ,¹⁰ again for the most probable words, are shown in figure 6a-d. They exhibit the same qualitative features as the M model-generated ones.

As the fly is subjected to different stimuli, the H1 dynamics is rather robust (as can be gleaned from table 2, which shows only a small change in parameter values).

¹⁰The position of the stripes depends on how we code for W_i , but quantities like D_q , f_α and Lyapunov exponents are coordinate independent.

Table 3. Lyapunov exponents λ for the four datasets extracted from M map sequences.

dataset	S_0	S_1	S_2	S_3
λ	1.85	1.81	2.08	2.14

We can also infer the *chaoticity* of the experimental data, measuring the Lyapunov exponents of the corresponding M map for the four datasets (see table 3).

As expected, the exponents increase as the stimuli's entropy increases. Note though that, even for no or constant stimuli, the exponents are positive, indicating a genuine chaotic UD.

4. Conclusion

We show that methods from *dynamical systems* from the point of view of *complex systems*, when applied to a biological system, are able to extract new and relevant features. In our case, we extract the UD of the H1 neuron by a finite size alphabet with about four letters. Analysing sequences written in this alphabet allows us to exhibit the multifractal character of the sequence space. We did model our data sequences by a piecewise almost linear map, whose dynamics compares well with the original data in a two-dimensional state space. The stimulus shapes the landscape of the sequence space from mainly uniform to highly structured, as the stimuli become increasingly dynamically variable. But this reshaping is played out on different chaotically unstable layers. These are dynamically linked—a behaviour typical of complex systems.

The existence of multifractal sets strongly suggests a chaotic dynamics—also supported by the positive Lyapunov exponents that we compute using the M map. This leads us to think about strange attractors, which exhibit a natural way to generate hot spots: the existence of folds of an attractor with bounded measure (Grebogi *et al.* 1988). There the system loses its hyperbolicity, since the stable and unstable manifolds are tangent. Owing to ergodicity or recurrence, these folds map densely into the attractor, generating high-probability spots.

The tools developed here for the analysis of our data are of general applicability. The fly's optical processing system reveals its fascinating complexity, even in the basic aspects of the spike generating dynamics. Aspects of this spike generating dynamics provide a means for its manipulation and control. It remains to be seen whether the layered structure uncovered here has a counterpart in the fly's optical information processing system.

Thanks go to the Fundação de Amparo a Pesquisa do Estado de São Paulo (FAPESP; M.S.B., R.K., C.G.) and the Max-Planck-Institute (M.S.B.) for financial support. It is R.K.'s pleasure to thank W. Bialek and R. R. van Steveninck for their infinite patience, when sharing their insights. We also thank I. Zuccoloto for her help in fly preparation.

References

- Baptista, M. S., Grebogi, C. & Köberle, R. 2006 Dynamically multilayered visual system of the multifractal fly. *Phys. Rev. Lett.* **97**, 178 102–178 105. (doi:10.1103/PhysRevLett.97.178102)
- Bialek, W., Nemenman I. & Tishby, N. 2001 Complexity through nonextensivity. arXiv: physics/0103076v1.

- Fairhall, A. L., Lewen, G. D., Bialek, W. & de Ruyter van Steveninck, R. R. 2001 Efficiency and ambiguity in an adaptive neural code. *Nature* **412**, 787–792. (doi:10.1038/35090500)
- Grassberger, P. 1986 Toward a quantitative theory of self-generated complexity. *Int. J. Theor. Phys.* **25**, 907–938. (doi:10.1007/BF00668821)
- Grassberger, P. & Kantz, H. 1985 Generating partitions for the dissipative Hénon. *Phys. Lett. A* **113**, 235–238. (doi:10.1016/0375-9601(85)90016-7)
- Grassberger, P. & Procaccia, I. 1983a Measuring the strangeness of strange attractors. *Physica D* **9**, 189–208. (doi:10.1016/0167-2789(83)90298-1)
- Grassberger, P. & Procaccia, I. 1983b Characterization of strange attractors. *Phys. Rev. Lett.* **50**, 346–349. (doi:10.1103/PhysRevLett.50.346)
- Grebogi, C., Ott, E. & Yorke, J. A. 1988 Unstable periodic orbits and the dimensions of multifractal chaotic attractors. *Phys. Rev. A* **37**, 1711–1724. (doi:10.1103/PhysRevA.37.1711)
- Halsey, T. C., Jensen, M. H., Kadanoff, L. P., Procaccia, I. & Shraiman, B. I. 1986 Fractal measures and their singularities: the characterization of strange sets. *Phys. Rev. A* **33**, 1141–1151. (doi:10.1103/PhysRevA.33.1141)
- Hausen, K. 1981 Monocular and binocular computation of motion in the lobula plate of the fly. *Verh. Dtsch. Zool. Ges.* **74**, 49–70.
- Hentschel, H. G. E. & Procaccia, I. 1983 The infinite number of generalized dimensions of fractals and strange attractors. *Physica D* **8**, 435–444. (doi:10.1016/0167-2789(83)90235-X)
- Kantz, H. & Schreiber, T. 1997 *Nonlinear time series analysis*. Cambridge, UK: Cambridge University Press.
- Plumecoq, J. & Lefranc, M. 2000a From template analysis to generating partitions: I: periodic orbits, knots and symbolic encodings. *Physica D* **144**, 231–258. (doi:10.1016/S0167-2789(00)00082-8)
- Plumecoq, J. & Lefranc, M. 2000b From template analysis to generating partitions II: characterization of the symbolic encodings. *Physica D* **144**, 259–278. (doi:10.1016/S0167-2789(00)00083-X)
- Poon, L. & Grebogi, C. 1995 Controlling complexity. *Phys. Rev. Lett.* **75**, 4023–4026. (doi:10.1103/PhysRevLett.75.4023)
- Renyi, A. 1971 *Probability theory*. Amsterdam, The Netherlands: North-Holland.
- Strong, S. P., Köberle, R., de Ruyter van Steveninck, R. R. & Bialek, W. 1998 Entropy and information in neural spike trains. *Phys. Rev. Lett.* **80**, 197–200. (doi:10.1103/PhysRevLett.80.197)
- Uhlenbeck, G. E. & Ornstein, L. S. 1930 On the theory of the Brownian motion. *Phys. Rev.* **36**, 823–841. (Reprinted: Wax, W. 1945 *Selected papers on noise and stochastic processes*. New York, NY: Dover Publications.) (doi:10.1103/PhysRev.36.823)