Recognizing and classifying temporal patterns in neuronal spike trains

Institute of Neuroinformatics

Winterthurerstrasse 190, CH-8057 Zurich, Tel: (+41) (0)1 635 30 52, http:// www.ini.ethz.ch

Markus Christen (markus@ini.phys.ethz.ch), Albert Kern, Thomas Ott and Ruedi Stoop

The temporal structure of single spike trains, expressed as patterns in series of interspike intervals (ISIs), is an important topic of the analysis of neuronal information processing [1]. We recently proposed a unbiased method for the detection of spike patterns based on the correlation integral [2]. Below, the features and the capability of the method to classify neurons according to their firing behavior is investigated.

1. Correlation integral based pattern recognition

The correlation integral was originally introduced for the determination of the correlation dimension [3]. It is calculated as

$$C_N^{(m)}(\epsilon) = \frac{1}{N(N-1)} \sum_{i \neq j} \theta(\epsilon - \|\xi_i^{(m)} - \xi_j^{(m)}\|),$$

where $\xi_k^{(m)}$ are data points embedded in dimension m, $\theta(x)$ is the Heavyside function ($\theta(x) = 0$ for $x \leq 0$ and $\theta(x) = 1$ for $x > 0$) and N is the number of embedded points. The correlation integral averages the probability of measuring a distance smaller than ϵ between two randomly chosen points $\xi_i^{(m)}$ and $\xi_j^{(m)}$. Therefore, it allows the detection of clusters, which are formed by the embedded points. To demonstrate this, we constructed a series as a repetition of the sequence $\{1,2,4\}$. The embedding of this series for $m = 2$ leads to three clusters, represented by the points $P_1 = \{1,2\}$, $P_2 = \{2,4\}$ and $P_3 = \{4,1\}$. Calculating the correlation integral and plotting $\log C_N^{(m)}(\epsilon)$ against $\log \epsilon$ leads to a clean-cut staircase structure in the log-log plot (Fig. 1a).

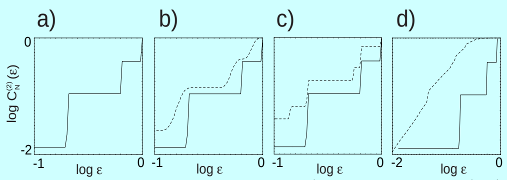

Figure 1: Log-log steps for different classes of data ($m = 2$, Euclidean norm). a) Virtually noise free (noise ±1%, solid line in all four plots), b) noisy (noise ±10%, dashed line), c) unstable periodic (dashed line), d) random background (dashed line).

In natural systems, three influences contribute towards a blurring of the steps, if only one pattern is present. First, the ISI series could be affected by noise. Second, the system which generates the ISI series could be of a chaotic orbit nature. Third, the sequence could be included in a noisy background. To demonstrate the effect of these influences on the log-log steps, we generated corresponding model data. The results are in agreement with the theoretical predictions (Fig. 1b-d, details in [4]).

2. Pattern length estimation

The pattern length is defined as the number of ISI involved. An indicator for the length can be obtained from the following reasoning: A pattern will emerge in the embedded ISI series in its most genuine form (neither cut into pieces, nor spoilt by points that do not belong to the pattern), if the pattern length equals the chosen embedding dimension ($m = n$). To investigate the potential of this criterion, we performed a number of experiments.

First, we include the sequences $\{5, 25, 10, 2\}$ and $\{5, 25, 10, 2, 17, 33\}$, respectively, with probability $p = 0.06$ into a noisy background. The background was provided by a homogeneous Poisson spike generator with refractory period. Additionally, the Poisson distribution is chosen so as to produce a mean interspike interval identical with the one generated by the patterns alone. Consistent with our expectations, the clearest steps emerge at the embedding dimensions 4 and 6 (Fig. 2a,b).

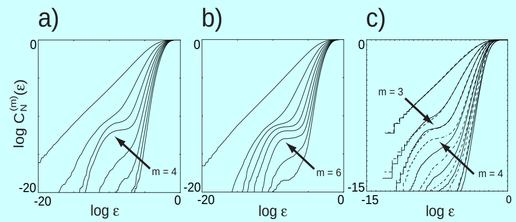

Figure 2: Log-log plots for varying embedding dimensions ($m = 1, \ldots, 8$, maximum norm) indicate the pattern size. a) Sequence of length 4 included in a Poisson spike train: Most pronounced step for $m = 4$. b) Sequence of length 6: Most pronounced step for $m = 6$. c) Sequences of length 3 and 4 included at different ratios (solid line: ratio 3:1, dashed line: ratio 1:3): Most pronounced step is where m equals the length of the dominating sequence.

Second, we refined this investigation by varying the individual pattern inclusion probabilities (using the sequences $\{4, 17, 12\}$ and $\{5, 25, 10, 2\}$). For the generation of the first series, the first pattern was chosen with $p = 0.12$ and the second with $p = 0.04$. For the generation of the second series, the probabilities were exchanged. The obtained results imply that also in this setting, the clearest steps emerge for m being equal to the pattern length n, but the influence of a particular pattern is weighted by its probability of occurrence (Fig. 2c).

3. Application to neuronal data

The method has been applied to spike data obtained from the striate cortex (V1) and the lateral geniculate nucleus (LGN) of cats. We analyzed data from extracellular field potential measurements of neurons from V1 (17 time series from 4 neurons) and LGN (17 time series from 6 neurons) of anesthetized cats (for details see [5,6]). Figure 3a shows the results of three cells (V1), displaying bimodal ISI histogram distributions in all cases, while the corresponding log-log plots indicate clear differences in the firing behavior. Earlier investigations of V1 data [5] suggested the existence of three classes of neurons: I) the class of randomly firing neurons, II) the class of neurons where simple patterns are included into a random or incompatible background, and III) the class of neurons that preferentially fire in patterns.

To investigate the pattern-length of class II spike trains, we calculated the ratio between the slopes of the flat and of the steep part of the steps in each embedding dimension. The most pronounced steps emerge for $m = 2$ and $m = 3$ (Fig. 3b and table below), indicating patterns of length 2 and 3. The results obtained for the LGN data are compatible with the above classification (Fig. 3c). Generally, the classification of the neuronal response is quite robust against the different stimulation paradigms used.

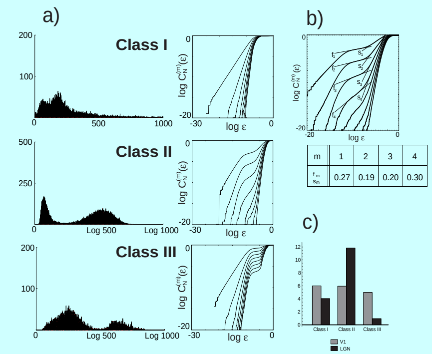

Figure 3: Cat data (V1, $m = 1, \ldots, 8$ in log-log plots). a): Neurons with bimodal ISI distributions display distinct pattern properties, leading to three classes (see text). b) Log-log plot of a cat V1 neuron of class II: The most pronounced steps appear at $m = 2$ and $m = 3$, indicating patterns of length 2 and 3. c) Histograms for cat LGN and V1 data.

4. Main Features

- Our method allows unbiased testing for pattern occurrence.
- The presence of patterns can be detected against a random environment.
- Robust indicators for the lengths of patterns are provided.

5. References

[1] Lestienne, Tuckwell, Neuroscience **82**(2), 315 (1998). [2] Christen, Kern, Stoop in: *Proceedings of the IEEE Conference on Nonlinear Dynamics of Electronic Systems* (2003), pp. 49-53. [3] Grassberger, Procaccia, Physica D **13**, 34 (1984). [4] Christen, Kern, Nikitchenko, Steeb, Stoop, Phys. Rev. E, submitted. [5] Stoop, Blank, Kern, van der Vyver, Christen, Lecchini, Wagner, Cog. Brain Res. **13**, 293 (2002). [6] Freeman, Durand, Kiper, Carandini, Neuron **35**, 759 (2002). The authors thank V. Mante, M. Carandini and K. Martin for providing the data used in this analysis.