

# Markov Analysis of Neural Interspike Intervals: Experiment and Model

Marcel J. van der Heyden<sup>123</sup> and Andre Longtin<sup>4</sup> Jacob DeGoede<sup>1</sup>

<sup>1</sup> Department of Physiology, Leiden University, P.O. Box 9604, 2300 RC Leiden, the Netherlands, Tel: + 31 71 527 67 51, Fax: + 31 71 527 67 82

<sup>2</sup> Graduate School of Neurosciences Amsterdam

<sup>3</sup> heijden@rullf2.leidenuniv.nl

<sup>4</sup> Département de Physique, Université d'Ottawa, 150 Louis Pasteur, Ottawa, Ontario, Canada K1N 6N5

**Abstract.** Statistical tests for the null-hypothesis that a symbolic sequence has  $n^{\text{th}}$  Markov order, based on Monte-Carlo resampling methods, are outlined. One such test is applied to binary and heptary representation of interspike interval sequences of an electroreceptor of the weakly electric eel *Apteronotus leptorhynchus* and a generic model of this sensory neuron.

## 1 Introduction

The analysis of a symbolic (discrete) representation of an otherwise continuous data set may have several advantages such as decreased statistical fluctuations in estimation and inference problems. In general the selection of a partition for coarse-graining a continuous data set is not trivial in the sense that the properties of the resulting symbolic sequence depend on the partition chosen. However, in certain cases one may be able to select meaningful partitions on the basis of physi(ologi)cal arguments that highlight some features of the data set while effectively assuming that other structures are due to, in a particular context, uninteresting fluctuations. Furthermore, one could also use discrete representations for model validation for instance.

A subject of interest in the neurosciences is the correlation structure and information content of sequences of neural interspike intervals (ISI) [2, 6, 8]. When treating ISI sequences as discrete Markov chains, the order provides important complementary information on the correlation structure in a data set and may thus shed light on information coding in (sensory) neural spike trains.

## 2 Markov Chains

Assume that a symbolic sequence  $\{x_i\}_{i=1}^N$  where  $x_i \in \{A_1, \dots, A_\lambda\}$  (the alphabet) is a discrete Markov chain. We will call a sequence composed of independent identically distributed (iid) elements to be of  $0^{\text{th}}$  Markov order. One can further define  $n^{\text{th}}$  order ( $n \geq 1$ ) Markov sequences through the  $n^{\text{th}}$  order (conditional)

transition probabilities which have the property that

$$p(a_i|a_{i-m}, \dots, a_{i-n}, \dots, a_{i-1}) = p(a_i|a_{i-n}, \dots, a_{i-1}) \quad (1)$$

for  $m \geq n$  and satisfy

$$\sum_{i=1}^{\lambda} p(a_i|a_{i-n}, \dots, a_{i-1}) = 1. \quad (2)$$

The transition probabilities can be conveniently assembled in the transition probability matrix  $\mathbf{P}_n$ .

The maximum-likelihood sample estimates of the joint- and conditional probabilities are estimated using the  $n^{\text{th}}$  order counts  $n_{a_{i-n}, \dots, a_i}$  which denote the number of overlapping tuples  $a_{i-n}, \dots, a_i$  present in the (circular) data set:

$$\hat{p}(a_{i-n}, \dots, a_i) = \frac{n_{a_{i-n}, \dots, a_i}}{N} \quad (3)$$

$$\hat{p}(a_i|a_{i-n}, \dots, a_{i-1}) = \frac{n_{a_{i-n}, \dots, a_{i-1}, a_i}}{n_{a_{i-n}, \dots, a_{i-1}}} \quad (4)$$

For an  $n^{\text{th}}$  order Markov discrete Markov process the  $n^{\text{th}}$  order counts are a *sufficient statistic* for the transition matrix defining such process.

### 3 Symbolic Surrogate Data

The surrogate data procedure for testing the Markov order consists first in specifying a null hypothesis  $H_0$  ("the data are of  $n^{\text{th}}$  Markov order") which defines the model sought to be rejected as a candidate for the data. Further, a number of realizations  $B$  of the  $H_0$  process is generated (the surrogate data) which mimic the  $H_0$  properties of the data as closely as possible so as to optimize the test performance. A test statistic  $S$  is selected which is sensitive to properties of higher Markov order than that of the  $H_0$  and thus can be used to detect differences between  $S^d$  of the data set and the ensemble of  $S_i^s$  of the surrogates. In this study the  $(n+1)^{\text{st}}$  order conditional entropy  $h_{n+1}$  [10] is used as the test statistic which allows for one-sided testing since  $h_{n+1} \leq h_n$  [8].  $S^d$  is ranked with respect to the ensemble  $\{S_i^s\}_{i=1}^B$  and the rank  $r$  determines the  $p$ -value  $\frac{r}{B+1}$  at which the  $H_0$  can be rejected.

An important property of tests is the *power*: the probability of correctly rejecting the  $H_0$  at a certain significance level  $p$ . When the  $H_0$  is incorrectly not rejected a Type II error occurs. The prespecified probability of rejecting the  $H_0$  when the data actually comply to the  $H_0$  (a Type I error) is called the *nominal size* of the test. In practice both the observed power and true size are dependent on the data set that is tested and the way the surrogate data are generated.

In statistical inference it is often attempted to eliminate the influence of nuisance parameters by conditioning on a sufficient statistic under the  $H_0$ . Conditioning often results in tests that have increased power and a true size that better approximates the nominal size  $p$  (see [12] for an example and [9] for an

review on conditioning in inference). We outline two methods for testing the  $H_0$  of  $n^{\text{th}}$  Markov order called the *typical realizations* surrogates (TRS) and the *constrained probabilities* surrogates (CPS) method. These methods are described in more detail in [5].

**Typical Realization Surrogates** Typical  $n^{\text{th}}$  order surrogate data are obtained by fitting the "best"  $n^{\text{th}}$  order model (in the maximum-likelihood sense)  $\hat{\mathbf{P}}_{\mathbf{n}}^{\mathbf{d}}$  to the data and then use this model to generate surrogate realizations. Note, however, that the  $\hat{\mathbf{P}}_{\mathbf{n}}^{\mathbf{s}}$  of a surrogate sequence is in general not identical to  $\hat{\mathbf{P}}_{\mathbf{n}}^{\mathbf{d}}$  of the original data due to statistical fluctuations. This means we have not conditioned on a sufficient statistic under the  $H_0$ .

**Constrained Surrogate Probabilities** The constrained surrogate probabilities method yields a set of counts of order  $n + 1$ , used to estimate  $\hat{\mathbf{P}}_{\mathbf{n}+1}^{\mathbf{s}}$  of the surrogate data, which has *identical*  $n^{\text{th}}$  order sample Markov properties as the original data.

We construct the CPS from the original data set by first extracting all  $N$   $(n + 1)^{\text{st}}$  order tuples  $(a_i, \dots, a_{i+n+2})$  of length  $n + 2$ . These tuples contain information of at most Markov order  $n + 1$ . With this set of tuples one can estimate all sample probabilities of order less than or equal to  $n + 1$  by selecting, at the same location in all tuples, subsets of the appropriate length.

We now construct a surrogate set of tuples by first splitting the set of  $n + 2$ -tuples  $\{a_i, \dots, a_{i+n+2}\}$  into 2 new sets of  $n + 1$ -tuples:  $\{a_i, \dots, a_{i+n+1}\}$  and  $\{a_{i+1}, \dots, a_{i+n+2}\}$ . Next, the surrogate set of  $n + 2$ -tuples is constructed by taking (without replacement) a  $n + 1$ -tuple  $(a_i, \dots, a_{i+n+1})$  from the first set of  $n + 1$  tuples. Then a second tuple  $(a_{i+1}, \dots, a_{i+n+2})$  is randomly drawn (without replacement) from the second set, where the elements  $a_{i+1}, \dots, a_{i+n+1}$  of the second tuple match the corresponding ones of the first tuple. In other words, the last  $n$  elements in the first  $n + 1$ -tuple are identical to the first  $n$  in the second  $n + 1$ -tuple. Thus we have conditioned on a sufficient statistic under the  $H_0$ .

A similar method had been described in [8] which does not however condition *exactly* on lower Markov order sample properties because the element  $a_{i+n+2}$  is drawn with replacement from the  $0^{\text{th}}$  order sample distribution of  $\{x_i\}_{i=1}^N$ .

**Power and Size** Power and size of the TRS and CPS tests were evaluated using binary and heptary (number of elements in the alphabet  $\lambda = 7$ ) test data of known order and length  $N = 512$ . Test data were generated using randomly constructed transition matrices. Size and power were estimate using 1000 different models of known order. The order of the test data for evaluating the power was always 1 larger than the order of the  $H_0$ . The  $H_0$  was rejected at a significance value of 0.05 using  $B = 19$  surrogate data sets for each test. The estimated power and size can thus be considered averages over a set of all Markov models of a given order. We took measures to avoid generating test data with too strong

correlations since our simulations showed that highly correlated data (with correlation lengths of the order of the length of the test data) yielded many false rejections of the  $H_0$  (Type I errors).

*Typical Realization Surrogates*

**Table 1.** Estimation of the power and size of the typical realization surrogate test for the binary and heptary siddogate data. Test parameters are given in the text.

$H_0$	0	1	2	3	4	5
$\lambda = 2$	0.46	0.53	0.55	0.50	0.32	0.08
<b>Power</b>	0.00	0.00	0.00	0.00	0.00	0.00
<b>Size</b>						
$\lambda = 7$						
<b>Power</b>	0.85	0.13	1.00	1.00	1.00	
<b>Size</b>	0.00	0.00	0.00	0.00	0.00	

The size of the TRS test is always much smaller than the the nominal size arguably to a lack of conditioning on the nuisance parameters [12] and resulting in a lower power compared to the CPS test (see further). Initially the power of the test increases as the Markov order of the  $H_0$  increases followed by a decrease in power due to finite sample effects which can be reduced using sufficiently long sequences. For higher values of  $\lambda$  (or smaller data sets) the size and the power can attain anomalous values close to 1 due to the occurrence of extremely sparse transition matrices  $\hat{\mathbf{P}}_{\mathbf{n}}$ .

*Constrained Probability Surrogates* The results of applying the CPS to the test data indicate that the CPS is an accurate test: the size of the CPS test is always approximately equal to the nominal value of 0.05, even for very short data sets or heptary data. The power of this test is consistently larger than that of the TRS method, and the CPS method does not suffer from finite sample effects to the extend the TRS method does. Therefor, we will use the constrained probabilities method for estimating the Markov order of the neural ISI sequences in the remainder of this paper.

**Table 2.** Estimation of the power and size of the constrained probabilities surrogate test for the binary and heptary suddogate data. Test parameters are given in the text.

$H_0$	<b>0</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
$\lambda = 2$	0.72	0.86	0.91	0.93	0.93	0.84
<b>Power</b>	0.04	0.05	0.06	0.05	0.06	0.05
<b>Size</b>						
$\lambda = 7$						
<b>Power</b>	0.96	0.60	0.21	0.07	0.05	
<b>Size</b>	0.05	0.06	0.05	0.04	0.06	

## 4 Application to Experimental and Model ISI Data

### 4.1 Experimental Data

We apply the CPS test for Markov order to symbolic representations of ISI sequences which have previously been analyzed in detail in [6]. The experimental ISI sequence was recorded in-vivo from a bursting P-type electroreceptor [13] of a weakly electric eel *Apteronotus leptorhynchus* [4] in the absence of moving environmental disturbances. This fish uses its electrical organ to generate a quasi-sinusoidal (660 Hz) electrical field: the electric organ discharge (EOD). The (reflected) EOD field is picked up by electroreceptors covering the body of the fish and used for electrolocation and electrocommunication. The electroreceptors have a high probability of firing at a preferred phase of the EOD and the ISIs are approximately equal to the EOD period. Often one or more periods are skipped, as can be seen from the banded structure of the plot of the raw data in Fig. 1.

### 4.2 Model

A generic model of this system can be constructed using a periodically driven noisy Fitzhugh-Nagumo excitable system [7]:

$$\begin{aligned} \epsilon \frac{dv}{dt} &= v(v - a)(1 - v) - w + [b + r \sin(\beta t) + \mu(t)] \\ \frac{dw}{dt} &= v - dw \\ \frac{d\mu}{dt} &= -l\mu + \xi(t). \end{aligned}$$

where the periodic forcing  $r \sin(\beta t)$  is applied to the voltage variable  $v$  and  $\xi(t)$  is Gaussian iid noise such that  $\mu(t)$  is an Ornstein-Uhlenbeck process with an

autocorrelation determined by the value of  $l$ . The exact values of the parameters are given in [7] and were adapted so as to give maximal similarity to the experimental data on visual inspection. It is of interest to investigate what aspects of the electroreceptor behavior can be reproduced by the model.

### 4.3 Coarse-Graining

The question as to the properties of the skipping behavior may be addressed by coarse-graining the sample-space into symbols corresponding to the integer number of the EOD periods in an interval. Thus each interval is mapped onto a symbol representing the number of skipped periods in an interval and for this data set we obtain a heptary symbol sequence by applying the partition indicated by the dotted lines in Fig. 1.

Properties of fluctuations *within* a band may be investigated by making a binary partition within each band with respect to the mean interval length *within* a band (dashed lines in Fig. 1). The resulting binary symbol sequences thus contains information on in-band firing properties and may be seen as binary representations of the “interval deviates” in [6].

### 4.4 Results

For both the continuous and the symbolic sequences we display the (linear) autocorrelation time as the lag beyond which the autocorrelation function is not significantly different from 0 anymore (see also [6]). The results of the analysis of the symbolic representations of the ISI data are displayed in Table 3. All analyses were performed using sequences of length 9165.

**Autocorrelations** For the *experimental* ISI data the autocorrelation between neighboring intervals and symbols is negative for a lag of 1 and slightly positively correlated at lag 2 for the binary case only. The heptary and continuous data show no significant autocorrelations beyond a lag of 1. The continuous and heptary *model* data are not at all autocorrelated, whereas the binary model data show only a small negative autocorrelation at lag 1. Thus we can conclude that the model presented here does not capture all linear correlations present in the experimental ISI sequence.

**Markov Order** When using a significance level of 0.01 for rejecting the  $H_0$  and  $B = 999$  surrogate data in testing we estimate the Markov order of the experimental data to be at least 4 for both the binary and heptary sequences. The discrete model ISI sequences have a Markov order of 2 and 0 for the binary and heptary case respectively. We remark that we do not consider the rejection of the  $H_0$  of  $4^{th}$  Markov order for the heptary model sequence as valid since we were unable to reject all smaller order. We suspect that this is either an uncommon statistical fluctuation or the result of unknown properties of the data set. In fact there are no solid statistical arguments to decide whether the rejection of  $4^{th}$  Markov order is “valid” or not.

**Nonlinearity** The large Markov order compared to the relatively short (linear) autocorrelations, especially for the experimental ISI sequence, are not necessarily in conflict with each other. Markov models incorporate not only linear structure, and high order Markov models may be constructed with a short autocorrelation time. Furthermore, it is consistent with the hypothesis that nonlinear structure is present in the sequences (see also Longtin [6, 7]).

**Table 3.** Summary of the results of the analysis of the experimental and model ISI sequences. The autocorrelation time (ACT), defined here as the lag beyond which the autocorrelation function was not significantly different from 0, was estimated for the discrete, as well as the continuous sequences. It is expressed in the number of samples. The Markov order of both discrete representations was estimated. Displayed are the  $p$ -values obtained.

type	$\lambda$	ACT	0	1	2	3	4	5
cell	2	2	< 0.01	< 0.01	< 0.01	< 0.01	0.51	0.74
	7	1	< 0.01	< 0.01	< 0.01	< 0.01	0.40	-
	cont.	1						
model	2	1	< 0.01	< 0.01	0.05	0.21	0.94	0.75
	7	0	0.12	0.53	0.13	0.17	< 0.01	-
	cont.	0						

## 5 Discussion

In this study we outlined two symbolic surrogate tests for estimating the Markov order of discrete sequences. The application of the symbolic surrogate data tests is straightforward and especially suited to analyze short data sets. Conditioning on a sufficient statistic under the  $H_0$  in this problem results in increased power and a more accurate test compared to the unconditional case, as well as better finite sample properties. We furthermore analyzed ISI sequences from a sensory neuron as well as a generic model of this cell. The coarse-graining of the sample-space allows us to focus on specific features of the data.

All experimental ISI data as well as the binary representation of the model data displayed negative linear correlation between subsequent intervals. Thus, long intervals are predominately followed by short ones and vice versa or in the case of the binary representations: firing prior to the mean preferred phase was often followed by firing after the mean preferred phase of EOD. The negative correlations between neighboring symbols in the binary representation of the experimental ISI data are simply the result of the phase-locking of the receptor activity to the EOD. These linear correlation properties are to some extent reproduced by the model, as can be expected.

The skipping behavior of the electroreceptor (studied using the heptary sequences) displayed only short linear anticorrelations at lag 1 probably due to membrane currents or intracellular processes with a relatively long time constant such as currents underlying the refractory period. In contrast, the heptary

model data showed no autocorrelations and was indistinguishable from an iid heptary sequence. Thus, the model failed to capture salient autocorrelation features of the experimental data.

The Markov orders of the symbolic representations of the experimental as well as the model ISI sequences are all larger than the corresponding autocorrelation times suggesting nonlinear structure in the ISI generating mechanisms. They may be the results of correlations already present in the EOD [6]<sup>5</sup> as well as processes on cellular level with a time constant much longer than the EOD period such as intracellular messenger dynamics.

More detailed interpretations of the nature of the higher order structure in the coarse-grained data remains speculative but could perhaps be related to recent reports on adaptive processes in electrosensory systems of *Apteronotus leptorhynchus* on time scales of minutes [1] and the absence of a clear time scale of the adaptation of the spike rate in response to step stimuli of the P-type electroreceptor [13].

We conclude that the symbolic surrogate method is a useful tool for statistical inference on discrete data sets in the natural science such as DNA or languages. Although many experimental data are not discrete and the choice of a partition determines to some extent the properties of the resulting symbolic sequence, one can make a meaningful partition of the sample space to highlight certain features of the data set. The validation of biophysical models using methods such as proposed here is promising and may provide more insight into the questions of coding in neural systems and information content of (sensory) spike trains.

**Acknowledgments** The constrained surrogate probabilities were suggested by F. Takens. The interspike interval data were provided by J. Bastian. The authors gratefully acknowledge helpful conversations with C. Diks.

## References

1. J. Bastian, Plasticity in an electrosensory system. I. General features of a dynamics sensory filter. *J. Neurophysiol.* **76**(4) pp. 2483–2496, 1996.
2. D.R. Brillinger, Nerve cell spike train data analysis: a progression of technique. *J. Amer. Statist. Assoc.* **87** pp. 270–271, 1992.
3. B. Efron and R.J. Tibshirani, *An introduction to the bootstrap*. Monographs on Statistics and Applied Probability **57**, 1993, Chapman and Hall, New York.
4. W. Heiligenberg and J. Bastian, The electric sense of weakly electric fish. *Ann. Rev. Physiol.* **46**, pp. 561–583, 1984.
5. M.J. van der Heyden, C. Diks, B.P.T. Hoekstra and J. DeGoede, Testing the order of discrete Markov chains using surrogate data. *submitted Physica D*.
6. A. Longtin and D.M. Racicot, Assessment of linear and nonlinear correlations between neural firing events. *Fields Inst. Comm.* **11**, pp. 223–239, 1997.
7. A. Longtin and D.M. Racicot, Spike train patterning and forecastability, *Biosystems* (in press) , 1997.

---

<sup>5</sup> We found the correlations in the EOD interval series to be not decaying with increasing lag and to be of the order of the length of the series itself.

8. H. Nakahama, M. Yamamoto, N. Ishi, H. Fujii and K. Aya, Dependency as a measure to estimate the order and the values of Markov processes. *Biol. Cybern.* **25**, pp. 209–226, 1977.
9. N. Reid, The roles of conditioning in inference. *Statist. Sci.* **10**(2), pp. 138-199, 1995.
10. C.E. Shannon and W. Weaver, *The mathematical Theory of Communication*, University of Illinois Press, 1963.
11. J. Theiler, S. Eubank, A. Longtin, B. Galdrikan and J.D. Farmer, Testing for nonlinearity in time series: the method of surrogate data. *Physica D* **58** pp. 77–94., 1992.
12. J. Theiler and D. Prichard, Constrained-realization Monte-Carlo method for hypothesis testing. *Physica D* **94**, pp. 221–235, 1996.
13. Z. Xu, J.R. Payne and M.E. Nelson, Logarithmic time course of sensory adaptation in electrosensory afferent nerve fibers in a weakly electric fish. *J. Neurophysiol.* **76**(3) pp. 2020–2032, 1996.

**Fig. 1.** Raw data set of interspike intervals of the electroreceptor of the electric eel (top) and of the model (bottom). Plotted are only the first 2048 data points (out of a total of 9165). The horizontal lines indicate the partitioning of the data into binary (dashed lines) and heptary data (dotted lines). The range of EOD periods (see text) skipped in the model data is larger than for the electroreceptor. The few ISI's that are larger than 7 EOD periods in the model data are all assigned the same symbol corresponding to a 7-EOD-period ISI.

