Course Module

Analysis of spike correlation in multi-channel recordings

Prof. Sonja Grün
Institute of Neuroscience and Medicine (INM-6)
Computational and Systems Neuroscience
Research Center Juelich
Juelich, Germany
and
Theoretical Systems Neurobiology
RTWH Aachen University
Aachen, Germany
s.gruen@fz-juelich.de
www.csn.fz-juelich.de

Synopsis

The present course material is designed to introduce students to the detection and significance evaluation of spike correlation between simultaneous spike trains. Two different analysis tools for spike correlation will be implemented and explored in their properties: For 1) the cross-correlation analysis and 2) the Unitary Event (UE) analysis in a basic form. In a first step, we will study the cross-correlation analysis and also how the properties of the individual spike trains map into the shape of the cross-correlation histogram (CCH), and we explore in this context methods to estimate the significance of the correlation. In a next step, we will implement the UE analysis for a stationary piece of data and study its properties, before we move on to implement the time-resolved UE analysis to extract time-dependent modulation of excess spike synchrony. By implementing the UE analysis for the analysis of not only pairs of neurons, but also for three parallel spike trains we will get insight into the increase of complexity when the various pattern constellations need to be considered if the order of correlation is of interest. For both correlation methods we get experience with estimating significance of correlation, learn how various features of the spike trains may complicate this and will study the application of surrogate methods for significance evaluation.

This course material is designed to provide hands-on experience with data analysis. We provide sample data sets of simulated data and extracellular in vivo recordings from monkey motor cortex. This course addresses graduate students with either experimental or theoretical background, and with limited experience in statistical data analysis. However, ample programming skills in Matlab are required.

Supplemental Material

Together with this course module we provide an Introductory Lecture 'Analysis of spike correlation in multi-channel recordings'. Supplement data files are provided for practical analysis, as well as matlab functions to accelerate the exercises.

Requirements

Practical work in this course module is designed for Matlab Version 7 or higher. Supplemental data files are required for data analysis.

Introductory and further reading

For an introduction into the methods and their assumptions relevant for this practical course module refer to the supplement introductory lecture.

The Unitary Events analysis studied in this course is summarized in

- Grün S (2009) Data-driven significance estimation of precise spike correlation. *J Neurophysiology*, 101: 1126-1140 (review)
- Grün S, Diesmann M, Aertsen A. Unitary Event Analysis. In: Analysis of parallel spike trains. Grün & Rotter (eds). Springer Series Computational Neuroscience (2010)

The original literature on these issues:

- Grün S, Diesmann M, Grammont F, Riehle A, and Aertsen A (1999) Detecting unitary events without discretization of time. *Journal of Neuroscience Methods*94:67-79
- Grün S, Diesmann M, and Aertsen A (2002) `Unitary Events' in multiple single-neuron activity. II. Non-Stationary data. *Neural Computation*14(1): 81-119,
- Grün S, Diesmann M, and Aertsen A (2002) `Unitary Events' in multiple single-neuron activity. I. Detection and significance. *Neural Computation* 14(1): 43-80
- Grün S, Riehle A, and Diesmann M (2003) Effect of cross-trial nonstationarity on joint-spike events. *Biological Cybernetics* 88(5): 335-351

For introduction into the cross-correlogram analysis we refer to:

- Aertsen, Gerstein, Habib, Palm (1989) Dynamics of Neuronal Firing Correlation: Modulation of 'Effective Connectivity'. J Neurophysiol 61(5): 900—917
- Tetzlaff T and Diesmann M. Dependence of Spike-Count Correlations on Spike-Train Statistics and Observation Time Scale. In: Analysis of parallel spike trains. Grün & Rotter (eds). Springer Series Computational Neuroscience (2010)
- Eggermont J. Pair-Correlation in the Time and Frequency Domain. In: Analysis of parallel spike trains. Grün & Rotter (eds). Springer Series Computational Neuroscience (2010)
- Eggermont J (1990) The Correlative Brain, Springer-Verlag, Studies of Brain Function, Berlin, ISBN 3-540-52326-X
- Perkel, Gerstein, Moore (1967) Neuronal spike trains and stochastic point processes. I. + II. Biophysical Journal 7(4): 391--418; 419--440

As an Introductory literature on numerical methods for significance evaluation of spike correlation is:

- Louis S, Borgelt C, Grün S. Generation and selection of surrogate methods for correlation analysis. In: Analysis of parallel spike trains. Grün & Rotter (eds). Springer Series Computational Neuroscience (2010)

General Instructions

- The exercises of Task 1 relate to the cross-correlation analysis, whereas Task 2 deals with the Unitary Event analysis. Although both can be treated independently, I recommend to have both tools available when analyzing the data files, since the methods can provide complimentary information.

- The two Task packages are composed of a) programming tasks, and b) analysis of data and its interpretation. The latter should motivate you to 'play' with the data and the analysis parameters.
- Subtasks indicated as 'Extra' are not required for the basic understanding, but are add ons.
- It is strongly suggested to not solely apply the correlation analysis, but also to analyze the individual neurons for their statistical properties to correctly interprete the result of the correlation analysis. Looking at the dot display and the PSTH are the minimal approach, better is to also observe the ISI distribution, and possibly the auto-correlation and the spike count distribution. However, since the focus of the exercises today is not the analysis of single neuron activity, we provide matlab functions which may supply you with the mentioned plots.
- **Create and use functions – do not copy-paste your code to perform the analysis on the various data sets (this is a well-known error source!). The first indication to do so is when one tends to copy code!**
- The exercises are organized such, that they can be partly performed with provided functions, but we would recommend to use your own programs yourself to enhance the learning effect.
- Comment the code while scripting that you later will later be able to rehearse the code for presenting the results. You may also comment within your code questions and answers that arise during discussion with the tutors. Include for each script a header with date and name of author.
- When producing figures showing the results make sure that the axes contain appropriate axes labels incl the correct units. Print all result figures to files for presentation and documentation. Choose an appropriate image format that you can use with the presentation software of your choice.
- Alternatively, one could use the matlab function **publish('script')** that runs the M-file script named script in the base workspace one cell at a time, and saves the code, comments, and results to an output file (formats: 'doc','html' (default), 'latex', 'ppt', 'xml'). The output file is stored along with other supporting output files (see the documentation for details).

Provided matlab functions:

The following matlab functions are provided to accelerate the exercises:

- **LoadData_Gnode.m:** load provided data sets
- **convert2bin.m:** convert a spike train given by a list of spike times into a binned spike train
- **jointSjointP.m:** calculates the joint-p and joint-surprise value needed in UE analysis

The following set of functions are provided to accelerate the exercises if needed:

- **MakeAnalysisSingleNeuron.m:** analysis of a single spike train (dot display, PSTH, spike counts per trial, ISI distribution, ACH) – calls a number of further matlab functions

In case of emergency – if you run out of time or got stuck:

- **MakeCCH.m:** cross-correlation analysis
- **UE_analysis_nonstat.m:** very basic version of the UE analysis for two neurons only. Please refer to the UE program package for serious data analysis !

Data sets:

The data files contain data of individual neurons and can be read by the function **LoadData_Gnode.m** providing you with 4 variables: a list (column) of spike times (1ms resoution), a list of corresponding trial ids, the number of trials in the file, and the number of time steps (1ms width) within a trial.
This list provides the files names of the neurons recorded simultaneously:

- Data5, Data6 (monkey data)
- Data7, Data8 (simulated)
- Data12, Data13 (monkey data)
- Data14, Data15 (monkey data)
- Data23, Data24 (simulated data)
- Data20,Data21,Data22 (simulated data)

Task 1: Cross-correlation analysis

a) Implement a function that computes the raw cross-correlation histogram. The function should get as inputs the data sets of two neurons (or alternatively, if used as a auto-correlation function twice the data of the same neuron), the maximal time lag of the CCH. The function should provide as output all the relevant information to produce a plot (x-axis, histogram). Make a plot. Make sure that also the bins at large delay contain the same number of data (or correct accordingly). Why is the latter relevant?
b) Are features in the CCH a consequence of the properties of the individual spike trains? Therefore also look at the dot displays and the PSTH of the neurons, and if required other measures.
c) Also include in the CCH the predicted CCH based on the average firing rate. What information do you gain?
d) Extra: Express the CCH in terms of the correlation coefficient computed bin by bin. Are the neuron pairs of the various data sets correlated?
e) Extra (but fun!): Implement the trial shuffle surrogate, and test if the neurons are significantly correlated. How many drawings of the surrogates have to be performed to apply a significance test level of 1%? Plot the original CCH and the mean of the surrogates +/- 2 (5% level) or 3 (1% level) standard deviations per bin. Are the resulting significantly correlated pairs of neurons different from the former analyses? If yes, why? Would there be an alternative way of calculating the predictor considering the non-stationarities of the firing rates without trial shuffling?

Analyze all the provided sets of pairs of neurons and discuss, if the neurons are correlated and why you come to this conclusion. Also note additional observations.

Task 2: Unitary Event analysis

a) Implement the Unitary Event analysis for two neurons for an assumed stationary data set. To do so, calculate trial-by-trial (why?) the expected number of coincidences, then sum their numbers to compare to the empirial numbers found in all trials. Use the provided function **jointSjointP.m** to calculate the joint-p or the joint-surprise value since here are a few tricks involved. A p-value of 5% corresponds to a joint-surprise value of 1.2788, a p-value of 1% to S=2.
b) Implement now the time-resolved UE analysis by performing the analysis implemented in a) in a sliding window fashion. Thus, extract in windows of e.g. 100ms width the data from each trial and analyze these as done in a). Then move the window by one bin and do the

analysis again. Plot as a result: the dot displays of the two neurons, their firing rates as estimated in the sliding window fashion, the time course of the expected and the empirical number of coincidences and the joint-suprise. Make sure that the time axis is properly aligned! Mark in the dot displays all spikes involved in a coincidence by a cyan circle, mark in the dot displays the spikes involved in Unitary Events by red diamonds.
c) Implement a way to vary the allowed coincidence width. The simplest way of doing this is to bin the data into larger bins (and clip, the UE analysis requires 0-1 entries) before feeding the data into the UE analysis. You may use the provided function **convert2bin.m** for the binning.
d) Extra: Implement the spike train shuffle surrogate to estimate the significance instead assuming a Poisson coincidence distribution for significance estimation. Are the results different compared to the analysis done in b) ?
e) Extra: Expand the Unitary Event analysis to 3 neurons (or better to N neurons). Calculate for each spike pattern constellation across the neurons the time dependent measures.

Analyze all the provided sets of pairs of neurons now with the UE analysis. Can you any thing conclude from the time averaged UE analysis (a)? Why is a time resolved UE analysis relevant? What do you gain compared to the CCH analysis and the average analysis? How does the binning of the data or the choice of the width of the sliding window influence the results?

f) Related to d): Are there tripple correlations contained in the data? Could this been observed/concluded from the pairwise analysis as well?

Acknowledgements
I am grateful to Alexa Riehle, CNRS Marseille, France who agreed to make her data available for this course. This G-Node Winter Courses on Neural Data Analysis are funded by the German Ministry of Education and Research.

This course was first tought at the 4^{th} G-node data analysis course in Munich. Part of the teaching material was developed and used in the NWG-BCCN course on 'Analysis and Models in Neurophysiology', held at the BCF at the Univ of Freiburg.