Comparison of surrogate techniques for evaluation of spatio-temporal patterns in massively parallel spike trains

Alessandra Stella[1,3], Peter Bouss[1,3], Günther Palm[2,1], Sonja Grün[1,3]

[1] Institute of Neuroscience and Medicine (INM-6, INM-10), Institute for Advanced Simulation (IAS-6) and Jara Brain Institute I (INM-10), Jülich Research Centre

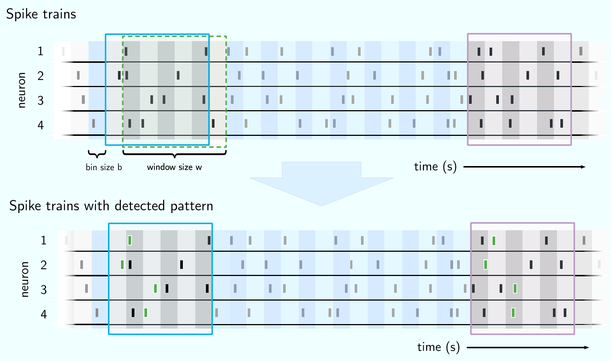
[2] Institute of Neural Information Processing, University of Ulm, Ulm, Germany

[3] Theoretical Systems Neurobiology, RWTH Aachen University, Germany

Contact: a.stella@fz-juelich.de

Introduction

Correlated activity between neurons is considered as a signature of the activation of a cell assembly [1,2]. To identify active cell assemblies we developed a method to detect significant **spatio-temporal spike patterns** (STPs). **SPADE** [3,4,5], identifies repeating ms-precise spike patterns across neurons.



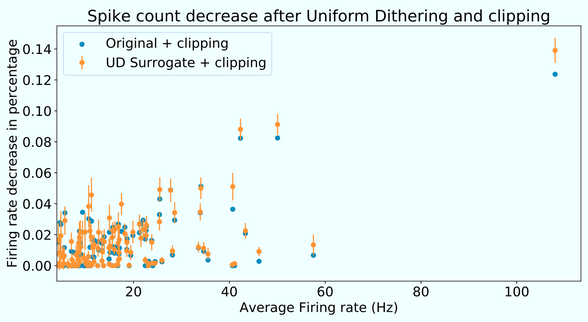
Schematized representation of four parallel spike trains, before and after the identification of a spatio-temporal pattern. Adapted from [5].

Candidate pattern are detected by Frequent Itemset Mining [6] after spike train discretization, or *clipping*. The method then employs **surrogate** generation in order to construct the null hypothesis of independence between spike trains. Candidate patterns are evaluated for significance based on the occurrence distributions in the surrogates.

Spike loss due to Uniform Dithering

Surrogate data implement the null-hypothesis of independence across neurons, and a classical choice is to apply **uniform dithering** (UD), i.e. independent, uniformly distributed local displacement of each spike. Important **issues** of UD are:

- it does not maintain the absolute refractory period
- destroys potentially existing ISI regularity
- leads to spike count reduction after clipping of the spike train, especially for high firing rates. (Figure)



Spike train reduction in an experimental session before and after UD. Each dot represents a neuron.

Surrogate techniques

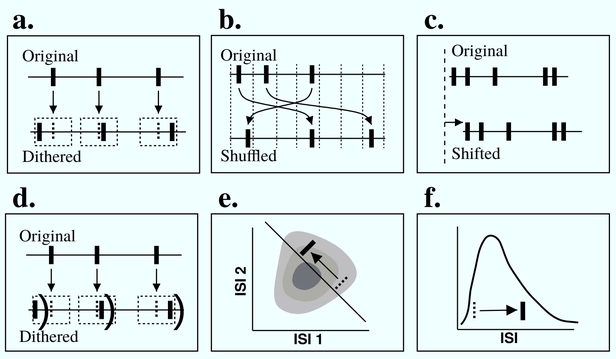


Illustration of the surrogate methods examined. Adapted from [8]

a) Uniform Dithering (UD) [7, 8]
b) Bin Shuffling (Bin-Shuff)
c) Spike Train Shifting (ST-Shift) [7,8]
d) Uniform dithering with Refractory Period (UD-RP)
e) Joint ISI-Dithering (JISI-D) [9]
f) ISI-Dithering (ISI-D)

Conclusions

We analyze experimental data and ad-hoc artificially generated indepedent data with six different surrogate techniques, in order to evaluate their statistical performances when looking for spatio-temporal spike patterns. We find:

- all surrogate techniques despite UD keep the spike counts at least approximately identical to the original data after clipping
- the firing rate reduction after clipping of the surrogate spike train is the primary reason for FPs
- loosing the ISI/J-ISI property of the original spike train may create only a few FPs
- we observe a low number of FPs (below significance threshold) across surrogate techniques besides UD
- we detect the same patterns in experimental data across different surrogate techniques, evidencing the robustness of our findings.

Results: Application of surrogate techniques on artificial and experimental data

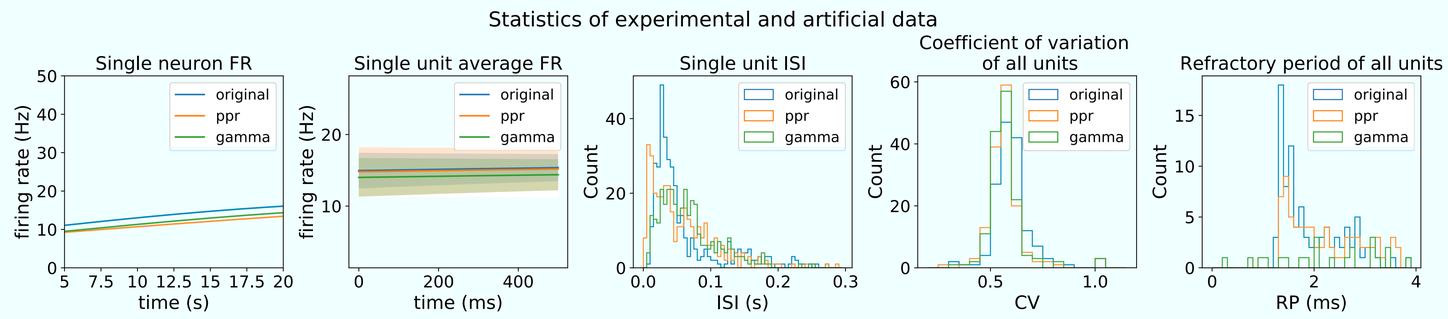
Our purpose is to verify and evaluate the statistical properties of the surrogates generated via the six methods presented. We apply the six different surrogate techniques to artificial data modeled on two sessions of experimental data (Motor cortex of a macaque monkey [10]), and compare the STPs resulting from the SPADE analysis. Finally, we apply the same analysis on the two experimental sessions.

Characteristics of the artificial data:

- Same number of neurons as in experimental data
- Spike trains modeled by independent inhomogeneous point processes, with the same firing rate profile of the original data
- Spike trains are modeled by Poisson process with refractory period (PPR) [11] and Gamma process
- Refractory process for PPR estimated from the data for each unit
- Shape factor of Gamma estimated from the average CV of each unit.

Method/Feature	UD	Bin-Shuff	ST-Shift	UD-RP	J-ISI-D	ISI-D
Firing rate modulation	approx.	approx.	yes	approx.	approx.	approx.
Spike count	no	yes	yes	approx.	approx.	approx.
ISI	no	no	yes	no	approx.	approx.
J-ISI	no	no	yes	no	approx.	no
Refractory period	no	yes	yes	yes	no	no

Table summarizing the statistical properties conserved/not-conserved/approximately conserved by the six surrogate techniques.



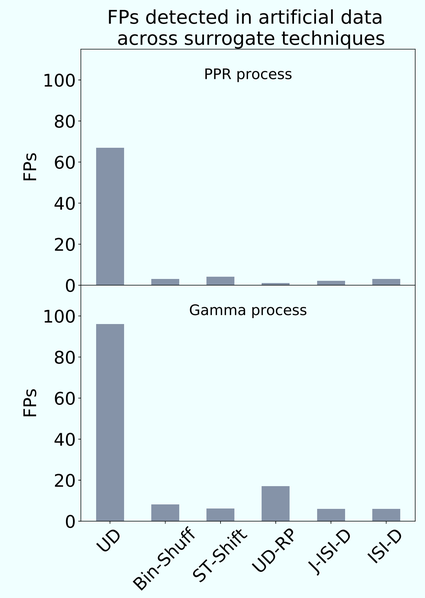
Comparison of statistics of the artificial data to the experimental data. In blue, orange and green, statistics of the original data, experimental data modeled by a Gamma and PPR process, respectively. Description of the panels from left to right. First panel: modulation of firing rate across one concatenated spike train for a single neuron. Second panel: time resolved average firing rate of a single neuron across trials of $500ms$. Third panel: ISI distribution of a single unit. Fourth panel: Average coefficient of variation across trials for all units. Fifth panel: Average refractory period across trials for all units.

In artificial data:

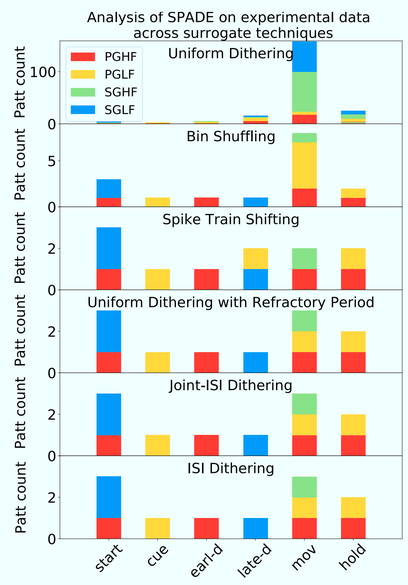
- In case of non-Poisson data, UD leads to a strong reduction in spike count after clipping, causing a large number of false positives (FPs)
- The other methods show a small number of FPs
- More patterns are detected when their intrinsic regularity is included in the artificial data (Gamma process). Fewer FPs for PPR process
- UDRP shows a higher number of FPs in the case of Gamma process (for neurons with high firing rate).

In experimental data:

- We detect most patterns (ca. 200 across two sessions) by using UD (likely to be FPs)
- Higher amount of patterns are found during the start and the movement period of the trials for all surrogate techniques besides UD
- Same patterns detected across ST-Shift, UDRP, J-ISI-D, ISI-D. Patternn sizes (number of spikes) ranging from 2 to 5. Pattern occurrences ranging from 11 to 386 (depending on size).



Number of False Positives detected by SPADE in artificial data across surrogate techniques. Top/bottom panel: Artificial data modeled by PPR/gamma process respectively.



Histogram of number of patterns detected per epoch of the experiment. Different colors represent different behavioral conditions (PG/SG: Precision/Side Grip; HF/LF: High/Low force)

References:

[1] Hebb D. (1949), Wiley and Sons
[2] Harris K. (2005), Nature Reviews Neuroscience
[3] Torre E., Picado-Muino D., Denker M., Borgelt C., Grün S. (2013), Frontiers in Computational Neuroscience
[4] Quaglio P., Yegenoglu A., Torre E., Endres D.M., Grün S. (2017), Frontiers in Computational Neuroscience
[5] Stella A., Quaglio P., Torre E., Grün S. (2019), Biosystems
[6] Louis S., Gerstein G., Grün S., Diesmann M. (2010), Frontiers in Computational Neuroscience
[7] Han J., Pei J., Yin Y. (2000), Data Mining and Knowledge Discovery
[8] Grün S., Rotter S. (2010), Springer
[9] Gerstein G. (2004), Acta Neurobiologiae Experimentalis
[10] Brochier T., Zehl L., Hao Y., Dure M., Sprenger J., Denker M., Grün S., Riehle A. (2018), Scientific data
[11] Deger M., Helias M., Boucsein C., Rotter S. (2012), Journal of Computational Neuroscience

Acknowledgments: The project is funded by the Helmholtz Association Initiative and Networking Fund (ZT-I-0003), by Human Brain Project HBP Grant No. 785907 (SGA2 and SGA3), and by RTG2416 MultiSenses-MultiScales (DFG).