Spike Train Analysis IV:

SPADE and other methods for higher-order spike pattern detection

Dr. Junji Ito (j.ito@fz-juelich.de)

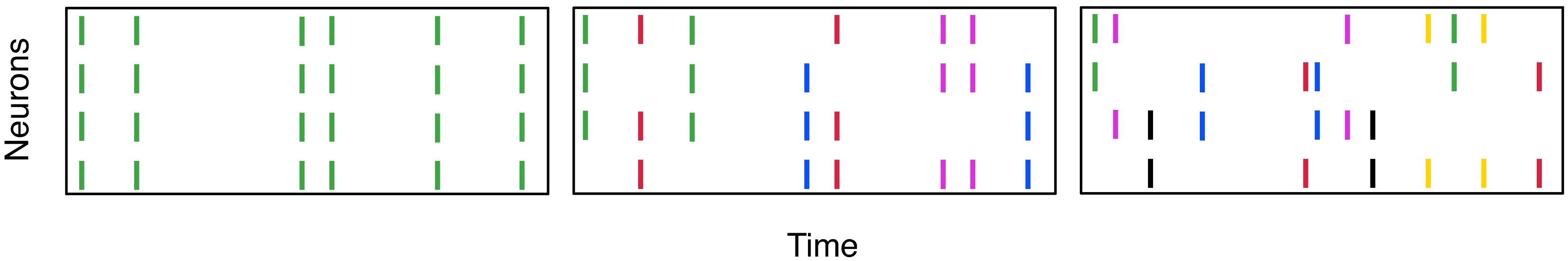
Institute of Neuroscience and Medicine (INM-6) and Institute for Advanced Simulation (IAS-6), Jülich Research Centre, Jülich, Germany

9th Latin American School on Computational Neuroscience (LASCON 2024)

NeuroMat, University of Sao Paulo, Sao Paulo, Brazil | January 25, 2024

Recap | higher-order correlation

- Same number of neurons, same number of spikes, but different order of correlation



- Analysis of only pairwise correlations cannot fully capture **higher-order correlations** (HOCs).

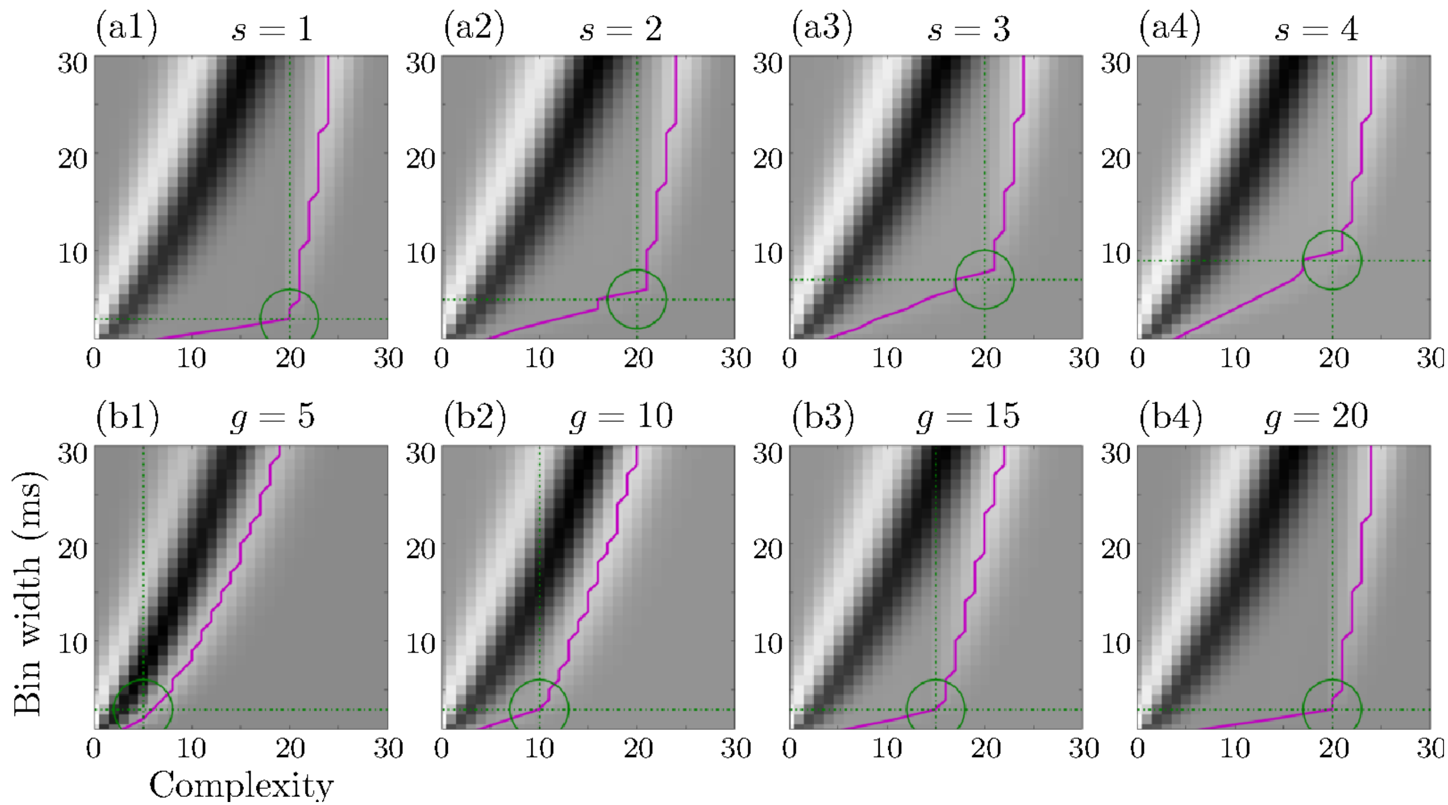
  → need for methods for detecting and analysing HOCs in massively parallel spike train data

Recap | challenges in HOC analysis

- UE analysis is, in principle, capable of detecting excess spike synchrony among more than 2 units
- However, in practice, scaling the UE analysis up to a large number N of neurons leads to…
  - **combinatorial explosion** of the number of patterns to be considered
    - e.g., for $N = 100$, $2^{100} \sim 10^{30}$ patterns
  - massive **multiple testing problem**
    - 10^{30} significant tests at a 5% significance level produce $\sim 5 \times 10^{28}$ false positives
    - Bonferroni correction (divide the p-value by the number of tests) makes the test too conservative

Recap | HOC in complexity distribution

- The order and temporal jitter of HOCs in massively parallel spike trains can be inferred from the complexity distribution
  → identity of neurons involved in the HOCs?



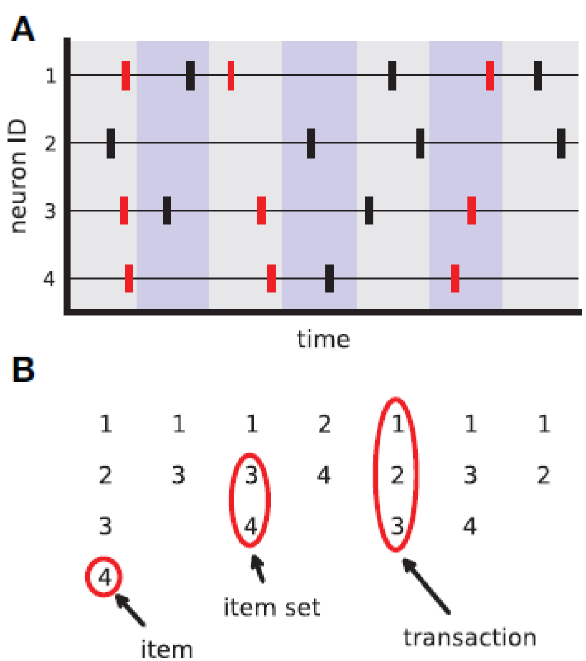
Louis et al. (2010) *Neural Networks*

Outline

- **SPADE for synchronous pattern identification**
  - methods
  - application
- **Extension to spatio-temporal pattern identification**
  - 3d-SPADE
  - ASSET

Methods | counting synchronous patterns

- **Frequent item-set mining** (FIM): an efficient counting algorithm, imported from market basket analysis

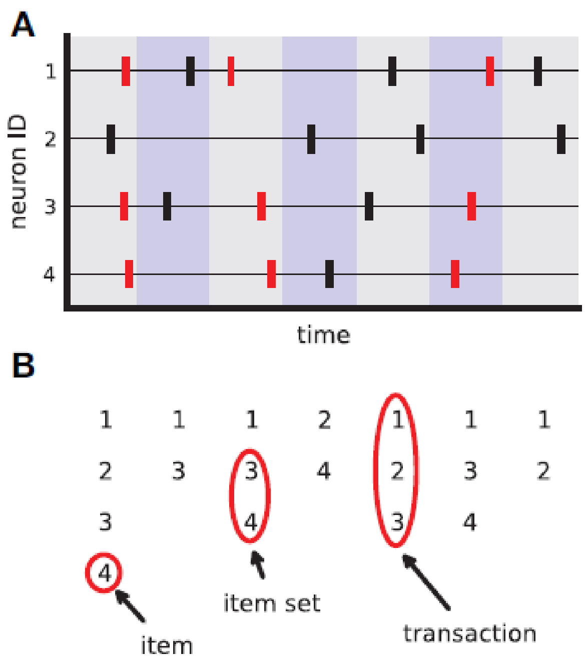


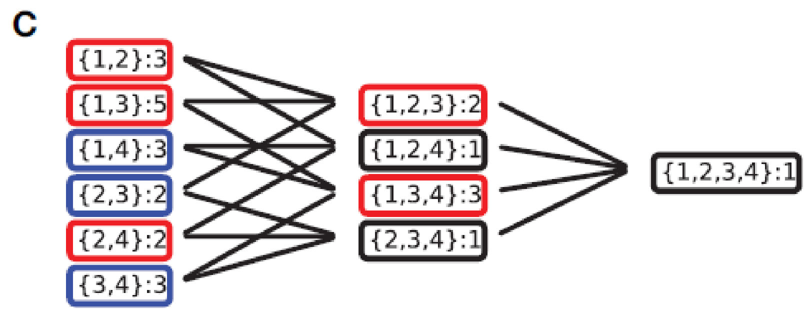
Mathematics	Market basket analysis	Spike train analysis
Item	Product	Neuron
Item base	Set of products	Set of neurons
Transaction *id*	Customer	Time bin
Transaction	Set of products bought by a customer	Set of neurons firing in a time bin
Frequent item set	Set of products frequently bought together	Set of neurons frequently firing together

Picado-Muiño et al (2013) *Front Neuroinform*
Torre et al (2013) *Front Comput Neurosci*

Methods | counting synchronous patterns

- **Frequent item-set mining** (FIM): an efficient counting algorithm, imported from market basket analysis
  - Consider only the patterns with size $z \geq Z$ and occurrence count $c \geq C$



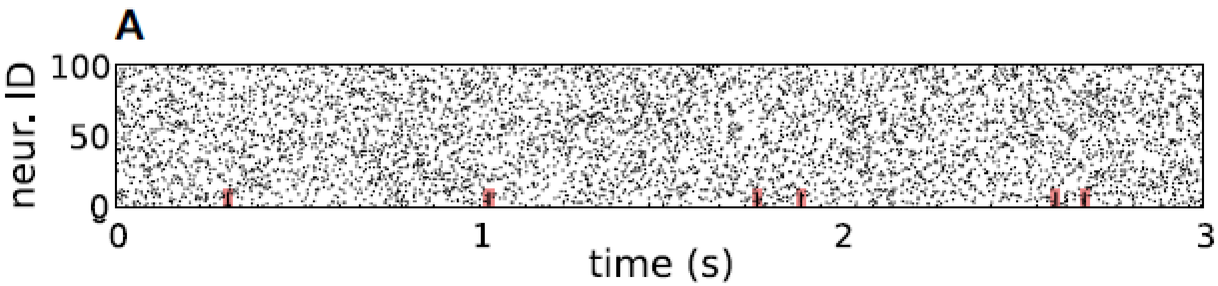


- **Non-frequent item-set**: occurring less than C times
- **Non-closed frequent item-set**: occurring only as many times as its superset
- **Closed frequent item-set** (CFIS): all the others; the output of the FIM algorithm

Picado-Muiño et al (2013) *Front Neuroinform*
Torre et al (2013) *Front Comput Neurosci*

Methods | testing significance of patterns

- A test data with a size-10 pattern (50 CFISs found by FIM)
- Characterize CFISs by **signatures**: size z and occurrence c

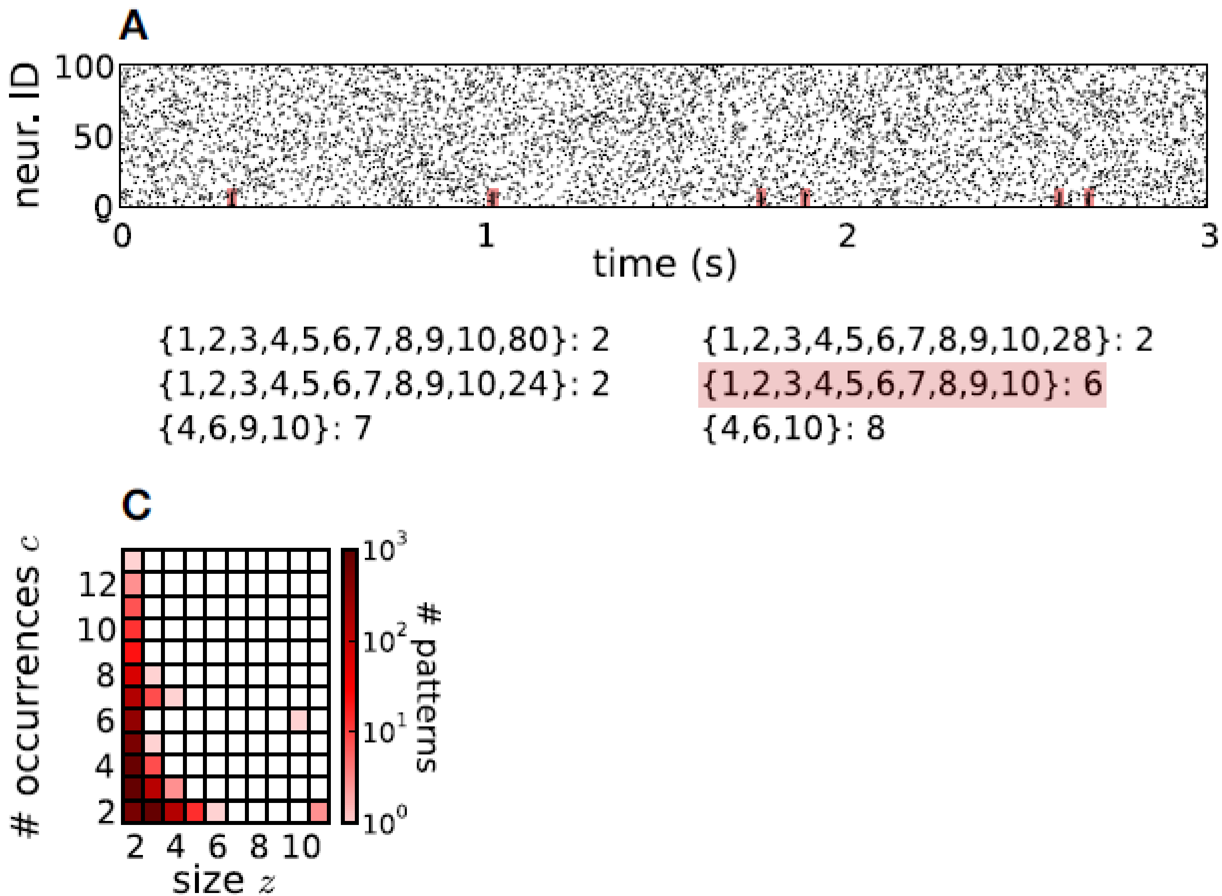


{1,2,3,4,5,6,7,8,9,10,80}: 2
{1,2,3,4,5,6,7,8,9,10,24}: 2
{4,6,9,10}: 7
{1,2,3,4,5,6,7,8,9,10,28}: 2
{1,2,3,4,5,6,7,8,9,10}: 6
{4,6,10}: 8

Torre et al (2013) *Front Comput Neurosci*

Methods | testing significance of patterns

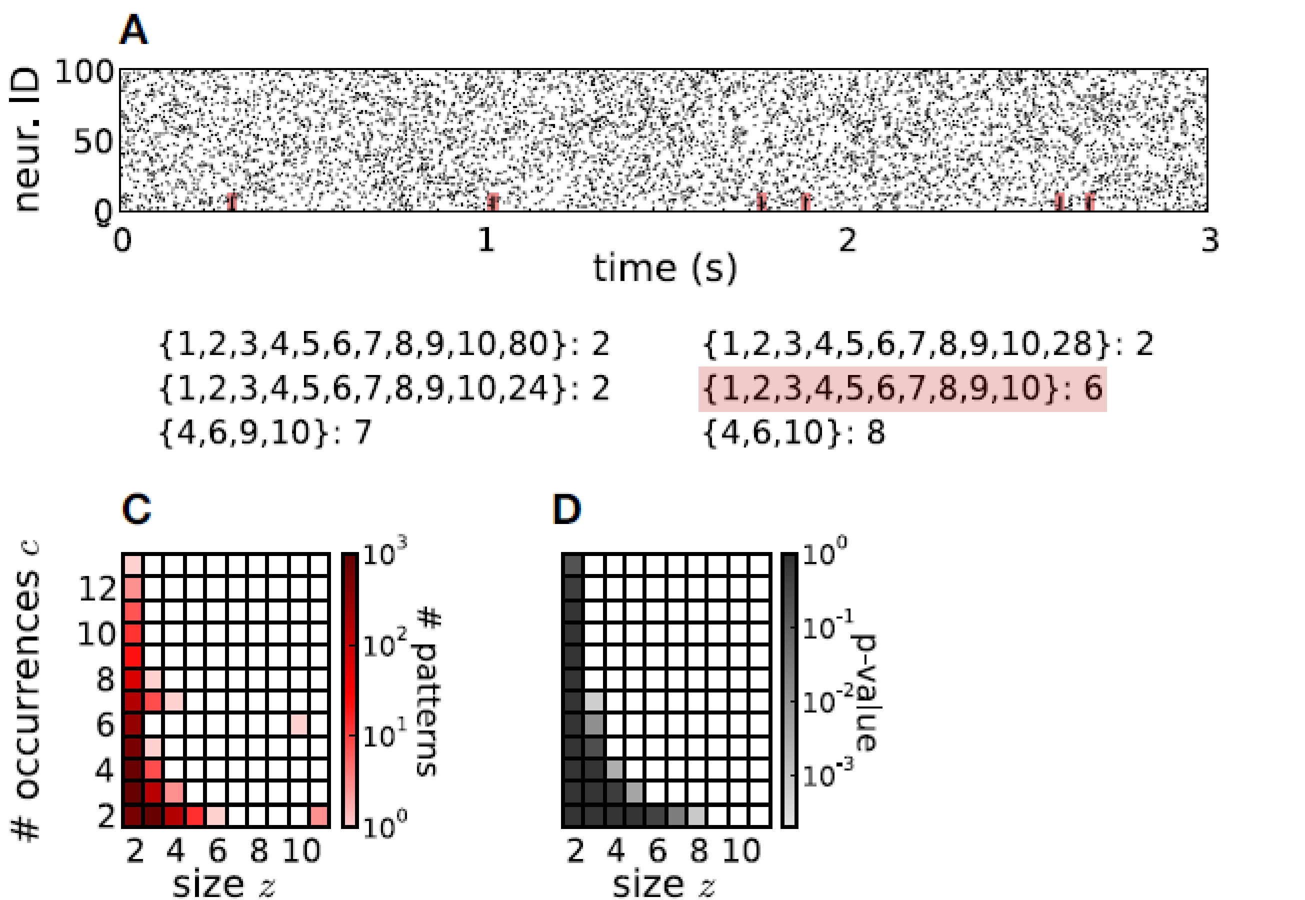
- **Pattern spectrum**: 2-dimensional histogram of the size z and the occurrence c of CFISs



Torre et al (2013) *Front Comput Neurosci*

Methods | testing significance of patterns

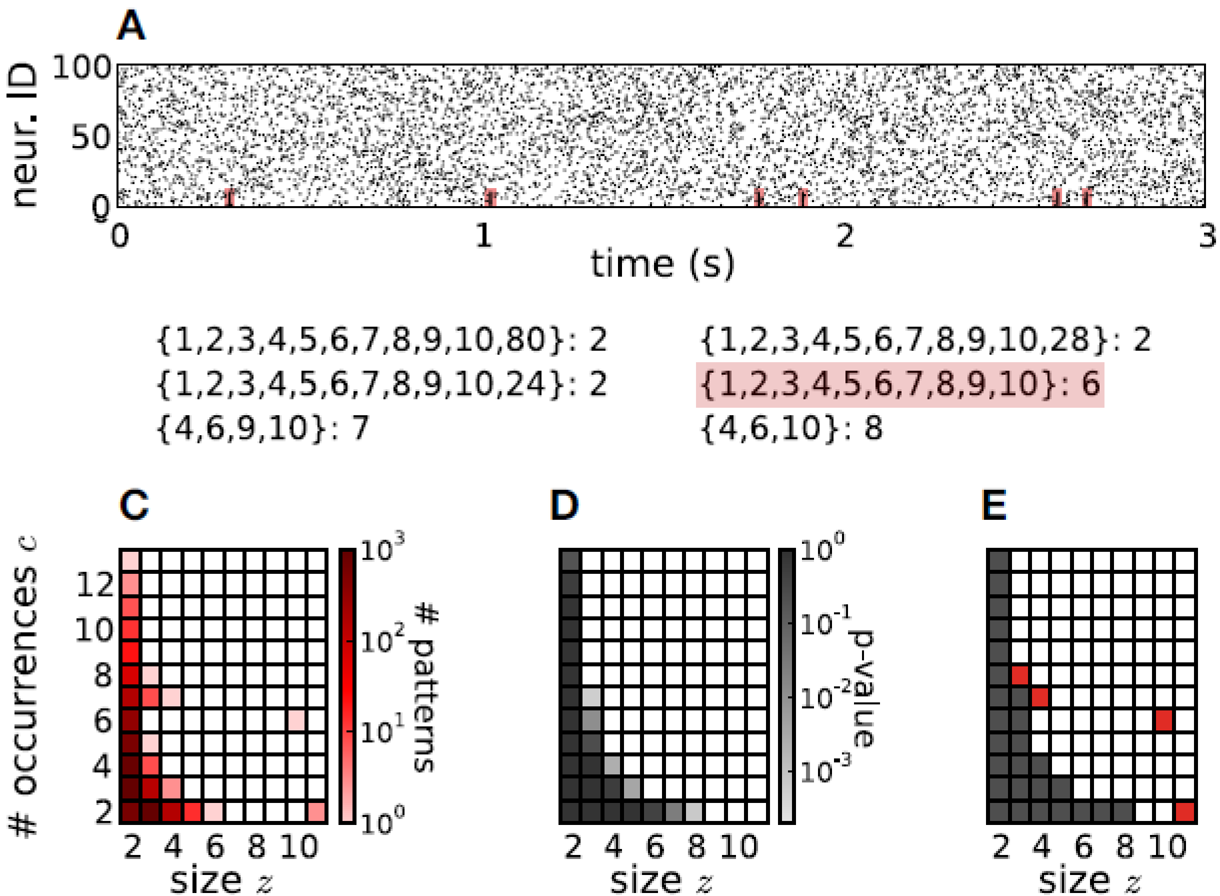
- **P-value spectrum**: 2-dimensional matrix of the occurrence probability of CFISs with size ≥ z and occurrence ≥ c



Torre et al (2013) *Front Comput Neurosci*

Methods | testing significance of patterns

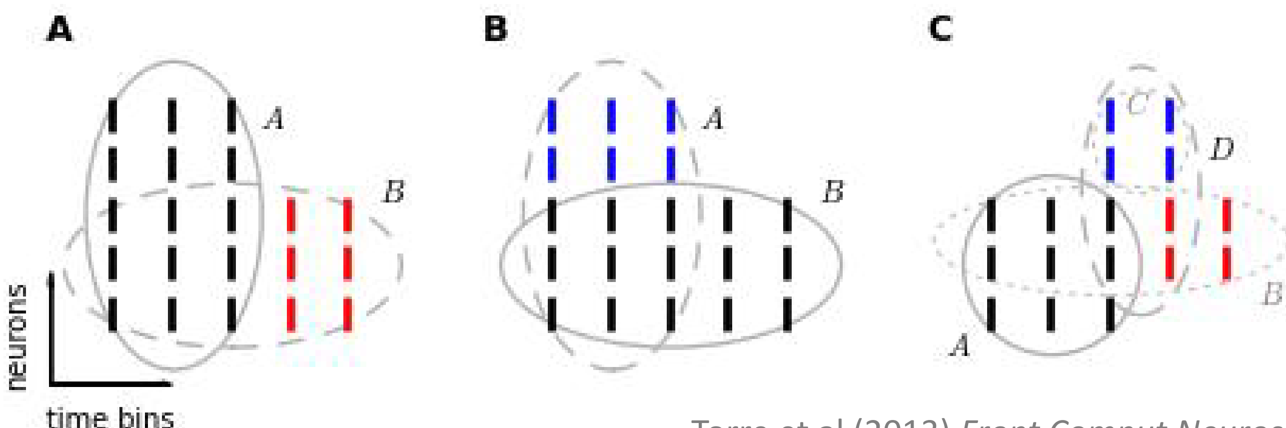
- **Significance spectrum**: 2-dimensional binary matrix indicating the significance of CFISs with signature (z, c)



Torre et al (2013) *Front Comput Neurosci*

Methods | rejecting false positives

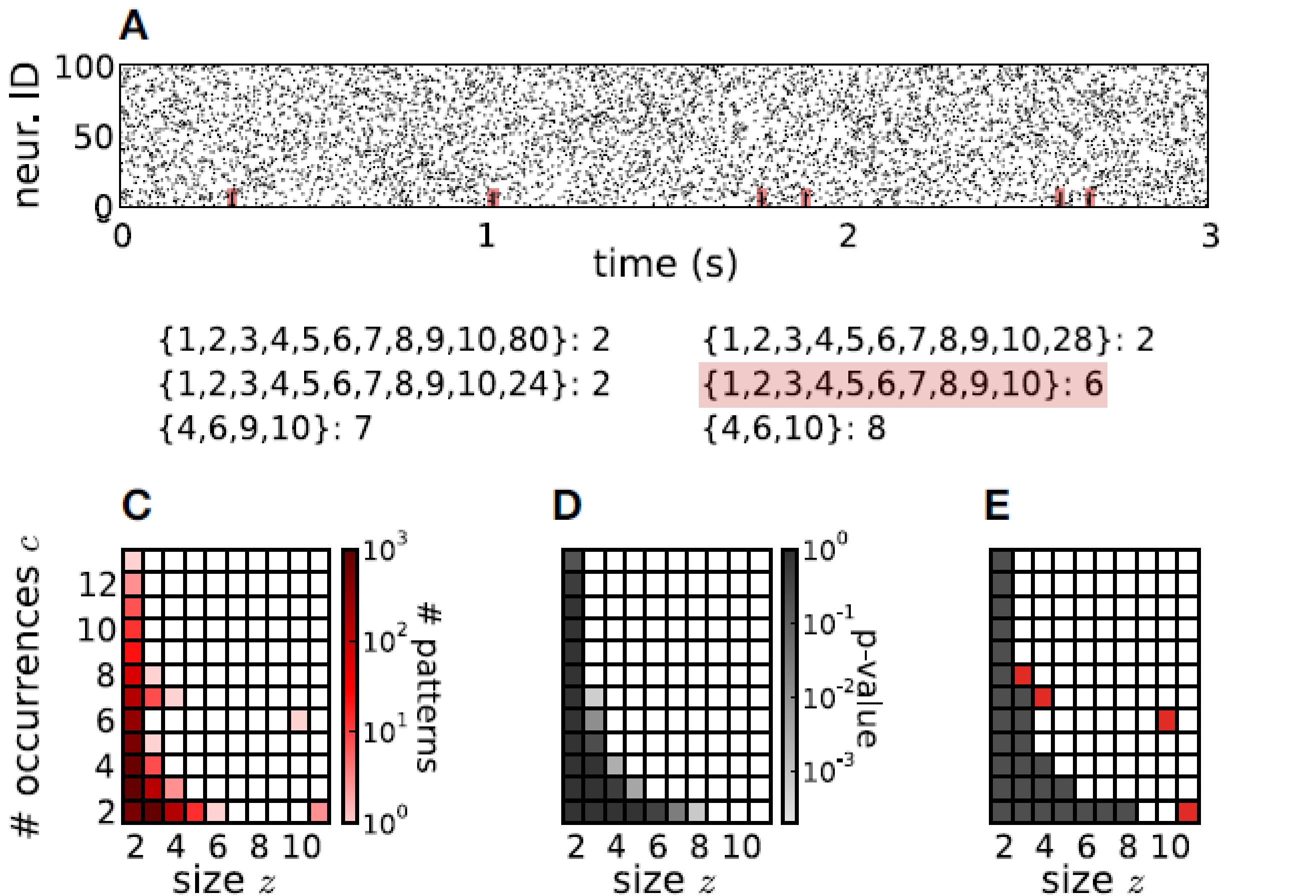
- Typically, sub- and super-patterns of a pattern from the actual HOC are detected as false positives
- They can be filtered out by additional **pattern set reduction**
  - *Subset and superset filtering*: reject a sub- or super-patterns if the extra synchrony is non-significant
  - *Covered-spikes criterion*: for any overlapping patterns, reject all patterns but the one that covers the most spikes



Torre et al (2013) *Front Comput Neurosci*

Methods | rejecting false positives

- After the pattern set reduction, only the pattern from the embedded HOC remains.



Torre et al (2013) *Front Comput Neurosci*

Methods | summary

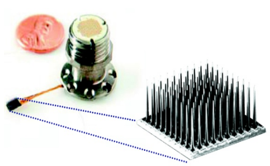
- **SPADE** (Spike PAttern Detection and Evaluation) method consists of the following steps:
  1. **Frequent item-set mining**:
     extract closed frequent item-sets (CFISs) from the given spike train data
  2. **Pattern spectrum filtering**:
     reduce CFISs to their signatures, and evaluate their significance based on surrogates
  3. **Pattern set reduction**:
     reject false positive patterns due to extra chance coincidences

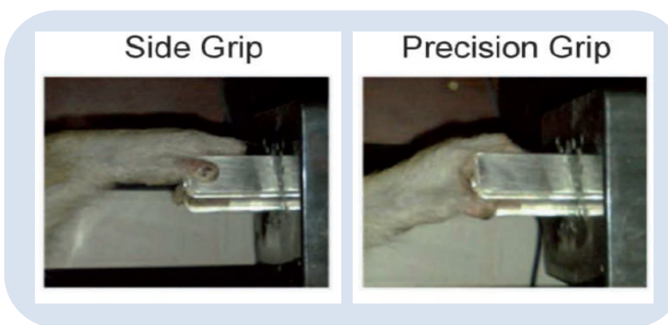
Outline

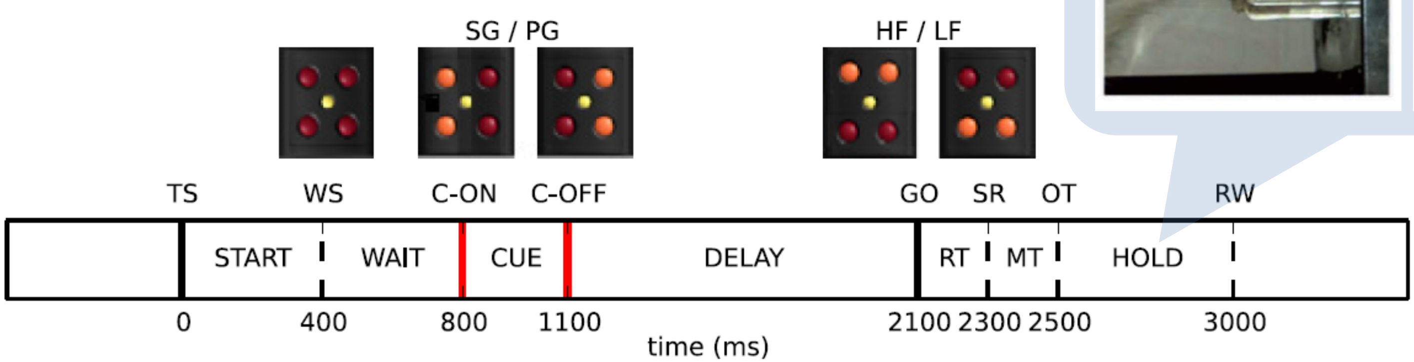
- **SPADE for synchronous pattern identification**
  - methods
  - application
- **Extension to spatio-temporal pattern identification**
  - 3d-SPADE
  - ASSET

Application | reach-to-grasp experiment

- Two macaques, trained for a delayed reaching task to hold an object with side/precision grip (SG/PG) using high/low force (HF/LF) → **4 trial types** by the combination
- Spiking activity was recoded with a 10x10 Utah array covering the primary motor and premotor cortices
  → ~100 simultaneously recorded single units per monkey



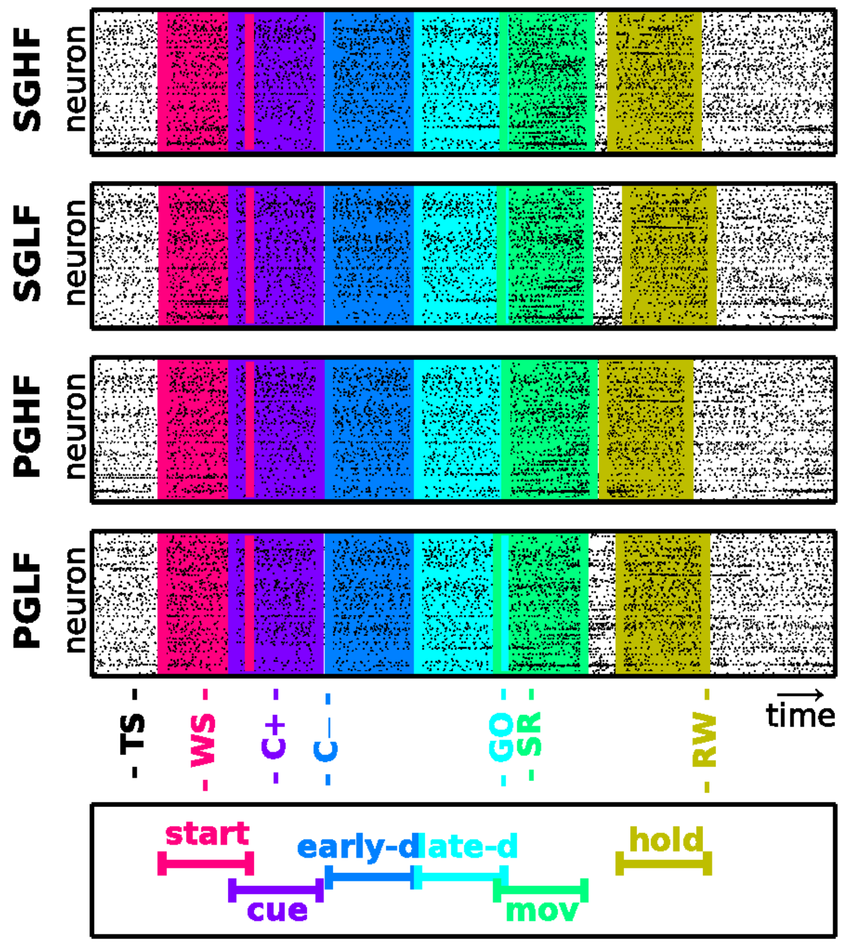


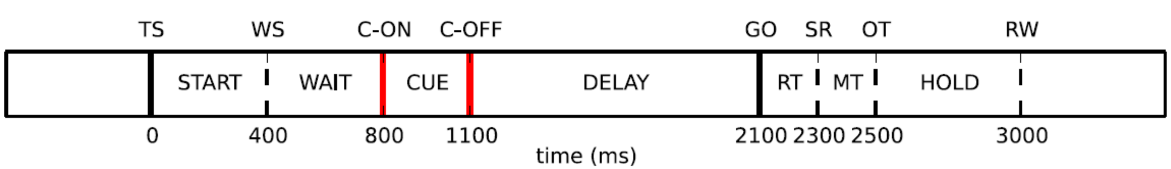


Riehle et al. (2013) *Front Neural Circuits*
Torre et al. (2016) *J Neurosci*

Application | data structure

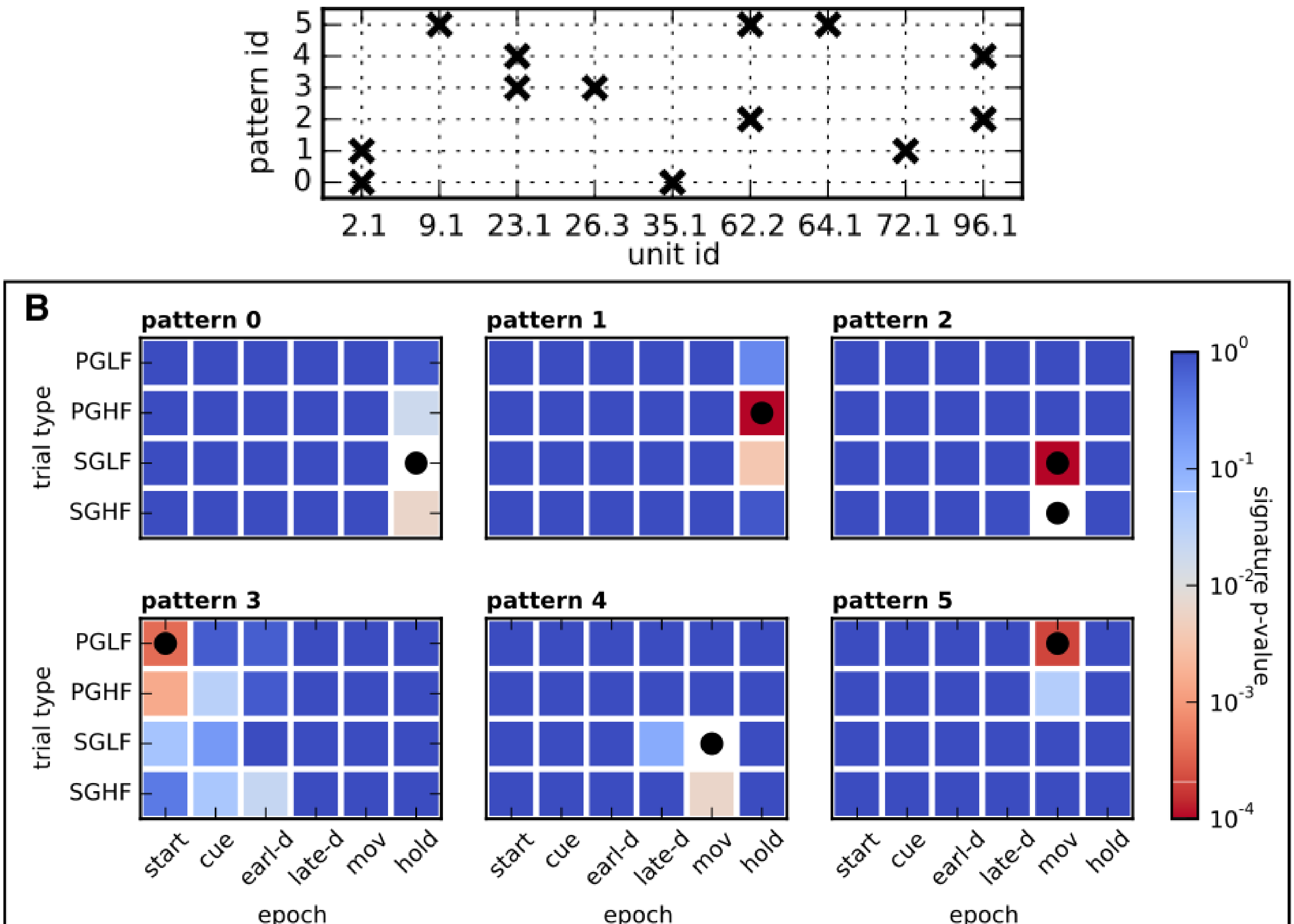
- 10 recording sessions per monkey
  - ~30 trials for each of the 4 trial types
  - Each trial is cut into 6 epochs of 500 ms
- SPADE analysis is applied separately to the data for each of the 24 combinations
- **Question**: *do patterns occur specifically for a particular trial type or epoch?*





Torre et al. (2016) *J Neurosci*

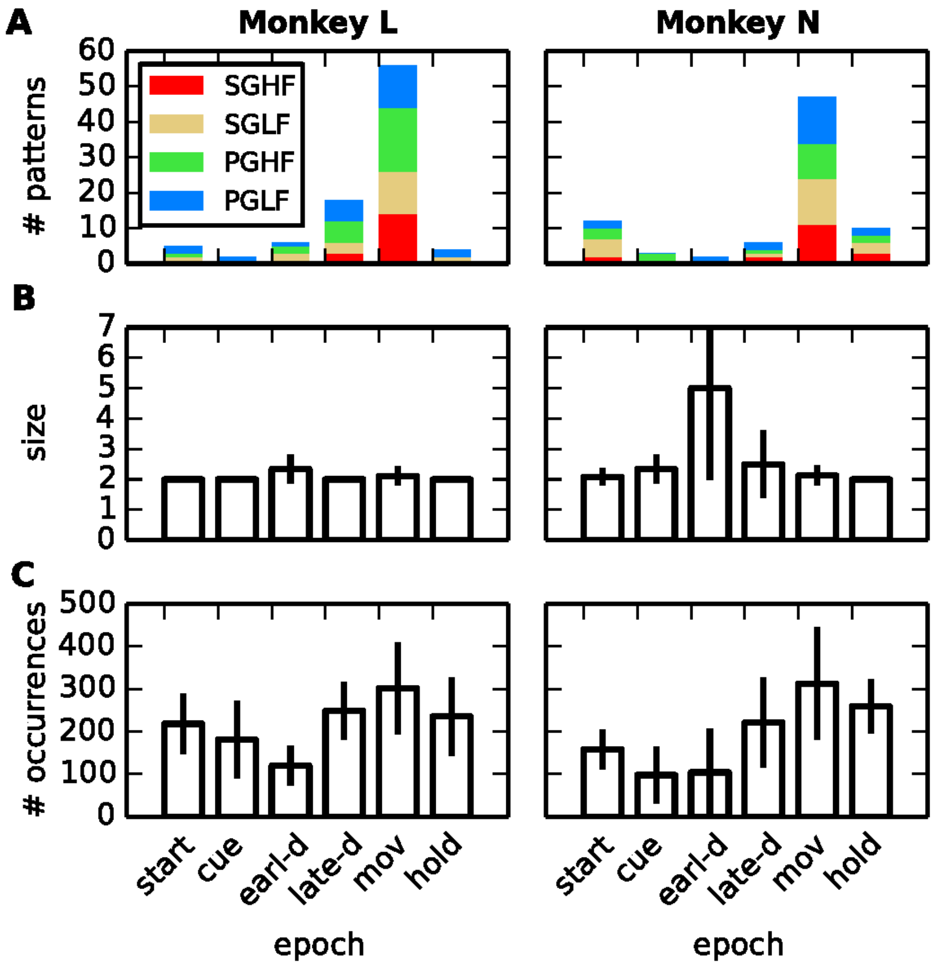
Application | example patterns in one session



Torre et al. (2016) *J Neurosci*

Application | population statistics

- Patterns occur primarily during late delay or during movement
- Pattern size on average of about 2
- Individual patterns occur > 100 times



Torre et al. (2016) *J Neurosci*

Application | are patterns specific?

- **Question**: *do patterns occur specifically for a particular trial type or epoch?*

- Let $\mathcal{P}_{\alpha,\mu}$ be the set of significant patterns found for the combination of trial type α and epoch μ.
- Specificity S_μ for **epoch** μ is defined as:

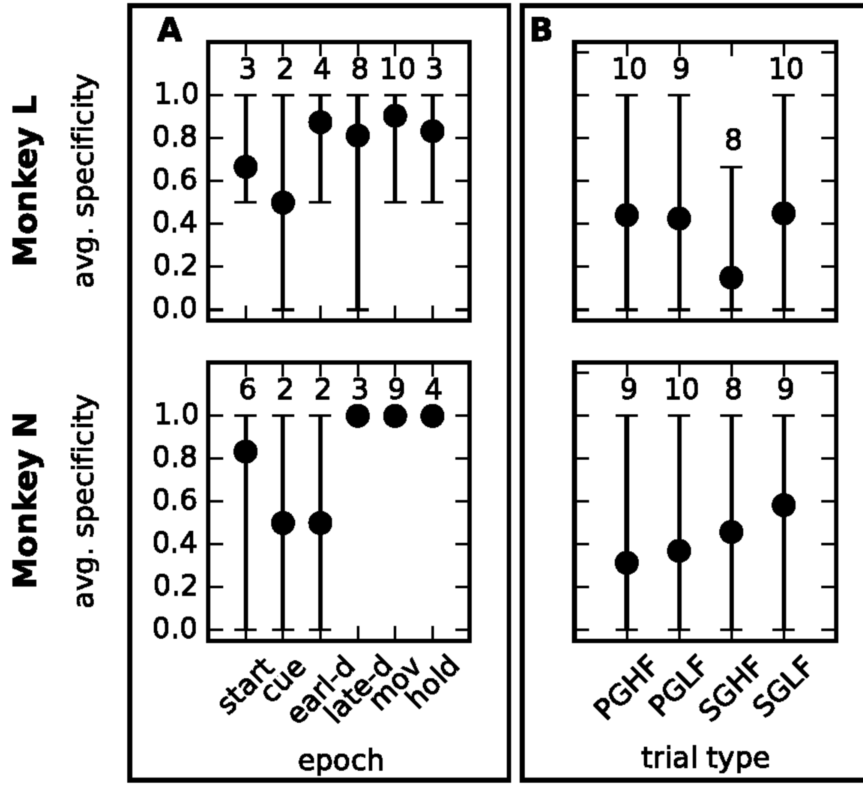
$$S_\mu = \frac{\left|\bigcup_\alpha \mathcal{P}_{\alpha,\mu} \setminus \bigcup_{\alpha,\nu\neq\mu} \mathcal{P}_{\alpha,\nu}\right|}{\left|\bigcup_\alpha \mathcal{P}_{\alpha,\mu}\right|}$$

  - $S_\mu = 1 \rightarrow$ all the patterns that occur in epoch μ never occur in the other epochs.
  - $S_\mu = 0 \rightarrow$ any pattern that occurs in epoch μ also occurs in at least one other trial type.
- Specificity S_α for **trial type** α is defined in the same manner.

Application | are patterns specific?

- **Question**: *do patterns occur specifically for a particular trial type or epoch?*

- Average specificity values over sessions per monkey
- Generally high specificity to epoch
- Medium to low specificity to trial type

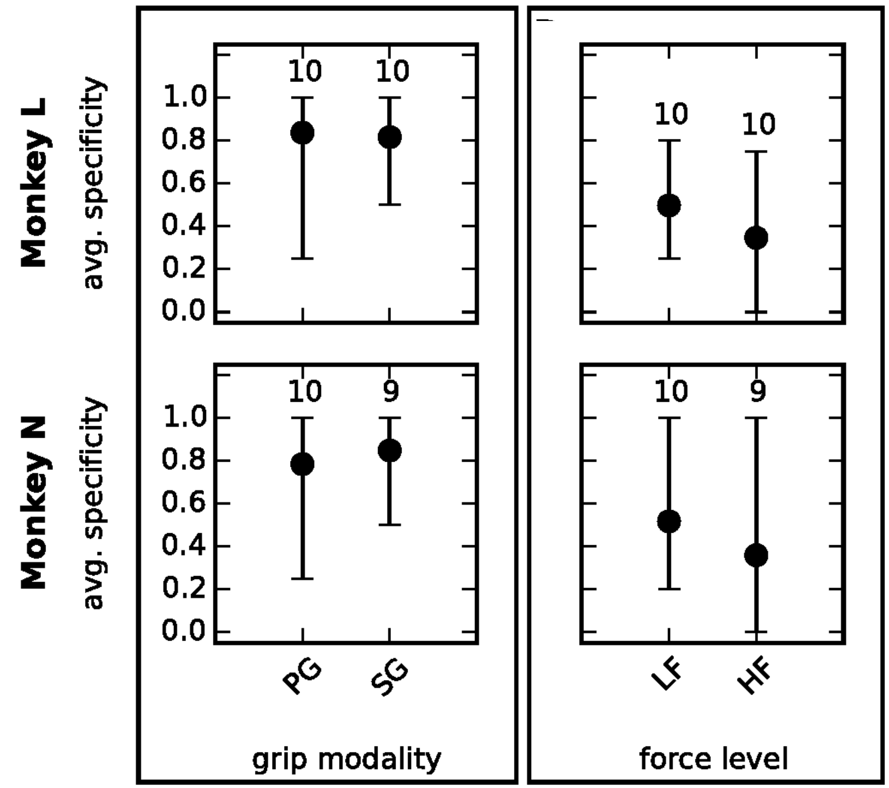


Torre et al. (2016) *J Neurosci*

Application | are patterns specific?

- **Question**: *do patterns occur specifically for a particular trial type or epoch?*

- Average specificity values over sessions per monkey
- Generally high specificity to epoch
- Medium to low specificity to trial type
  - If focused only on grip type or force level, specificity is high to grip type but low to force level



Torre et al. (2016) *J Neurosci*

SPADE analysis | summary

- The SPADE method enables to identify the groups of units that fire synchronously above the chance level
  - The chance level is estimated based on surrogates
    → can be adapted to various assumptions on the data

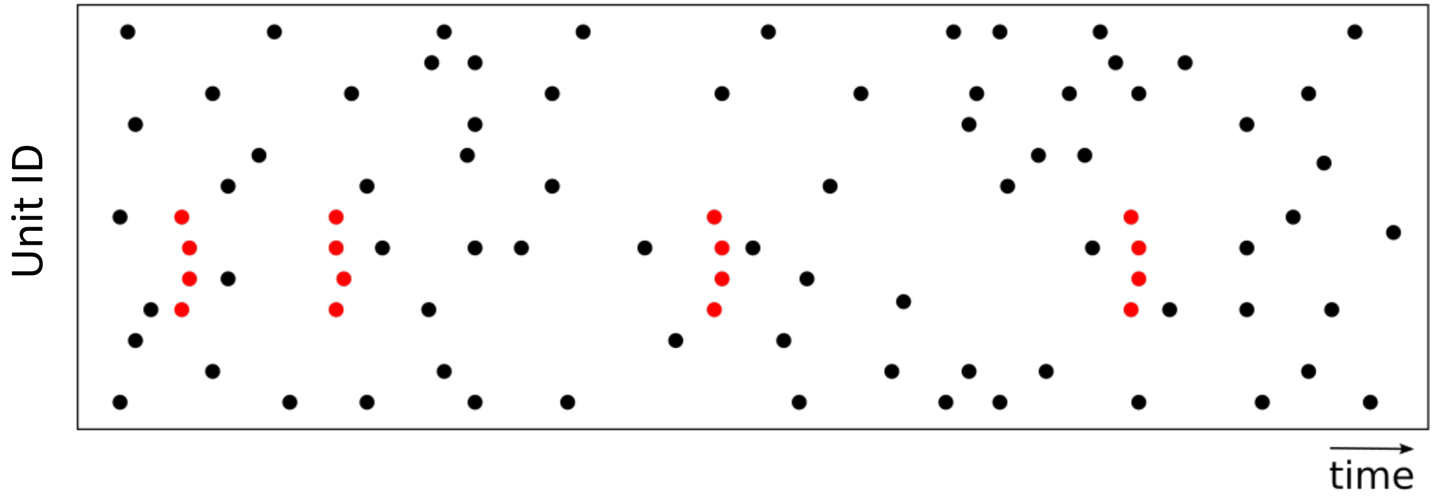
- The SPADE analysis on the reach-to-grasp data revealed behaviour-specific synchronous spike patterns
  - High specificity to epoch
  - Higher specificity to grip type than to force level
- Not clear whether these patterns represent information independent from- and/or complementary to the information represented by firing rates

Outline

- SPADE for synchronous pattern identification
  - methods
  - application
- **Extension to spatio-temporal pattern identification**
  - Beyond synchronous spike patterns
  - 3d-SPADE
  - ASSET

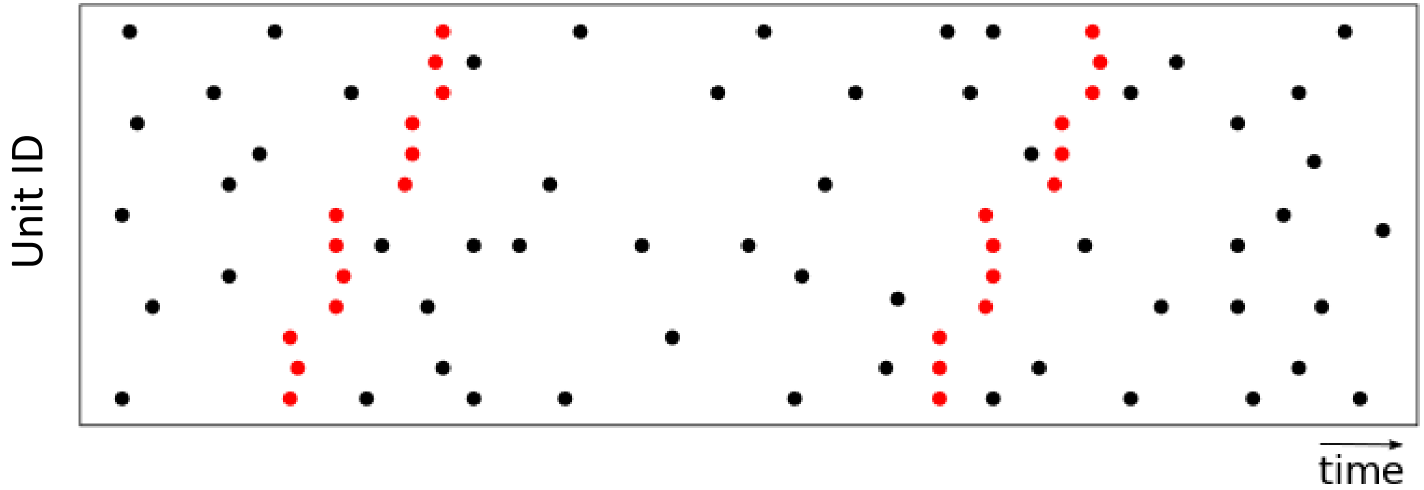
Beyond synchronous spike patterns

- Synchronous spikes can effectively activate post-synaptic neurons
- Those post-synaptic neurons also fire synchronously?
  → stable transmission of synchronous volleys of spikes
  → **synfire chain** (Abeles, 1982; 1991)



Beyond synchronous spike patterns

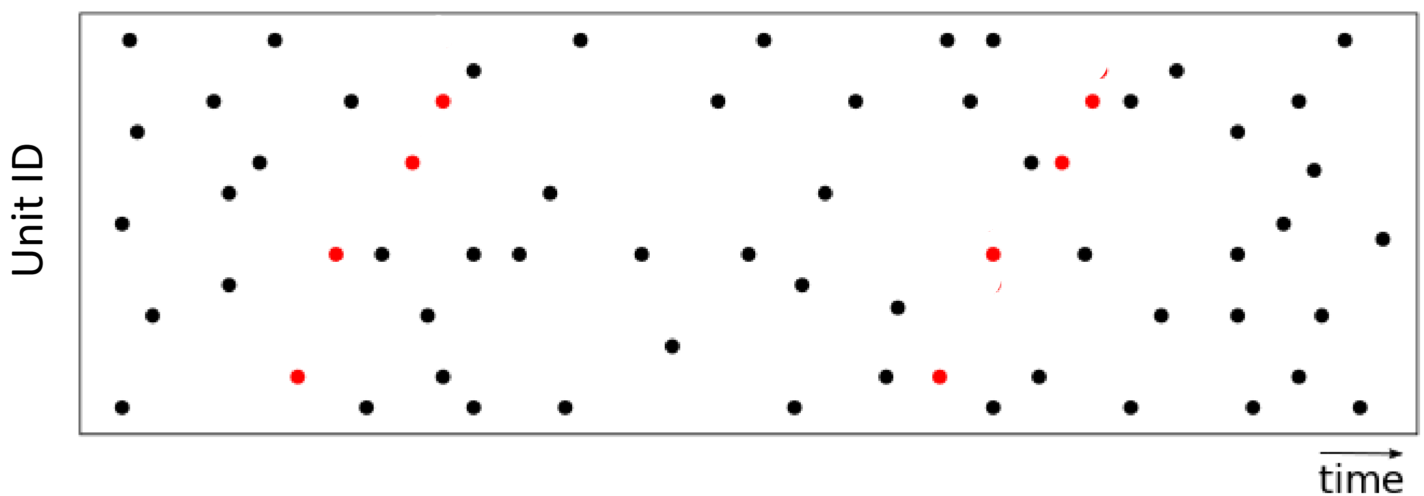
- If such a transmission occurs within the recorded neurons, multiple synchronous spike patterns with fixed time delays would be observed



Beyond synchronous spike patterns

- If such a transmission occurs within the recorded neurons, multiple synchronous spike patterns with fixed time delays would be observed
- More realistically, at most one neuron per synchronous group (or another model: synfire braid (Bienenstock, 1996); polychrony (Izhikevich, 2006))

→ **spatio-temporal pattern** (STP) of spikes

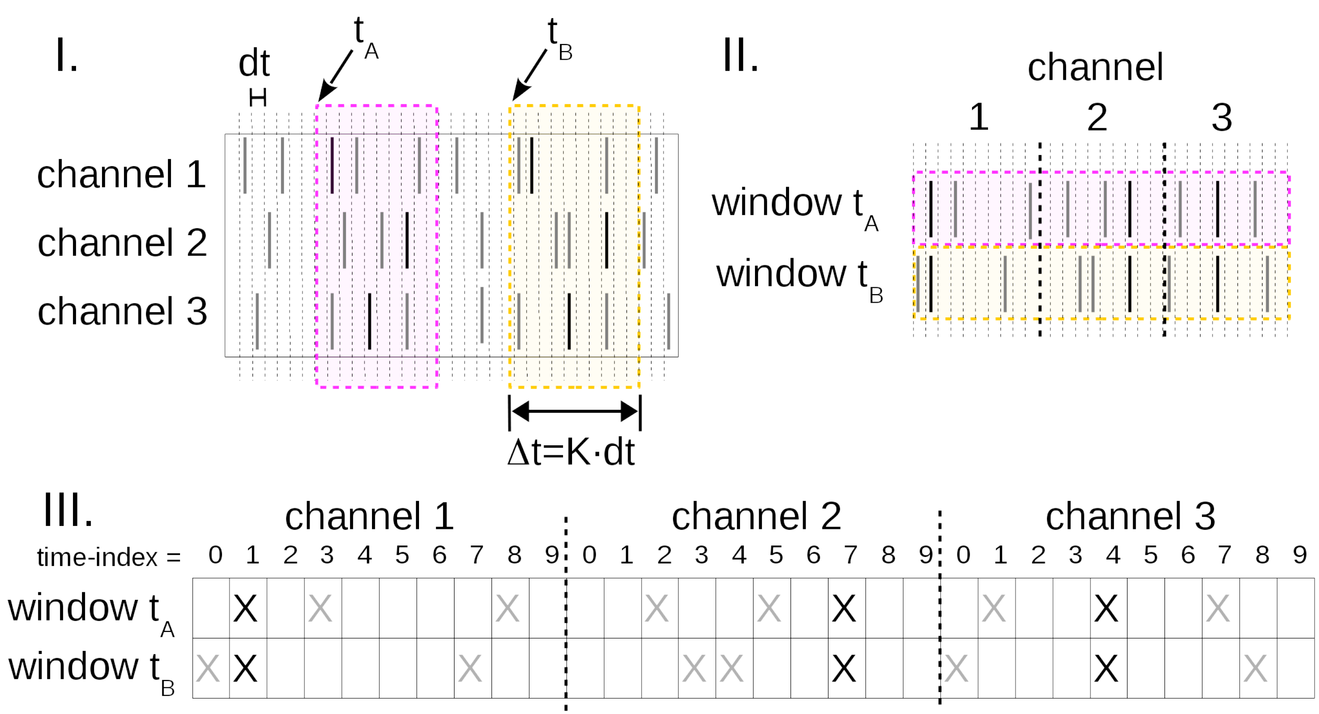


Outline

- **SPADE for synchronous pattern identification**
  - methods
  - application
- **Extension to spatio-temporal pattern identification**
  - Beyond synchronous spike patterns
  - 3d-SPADE
  - ASSET

3d-SPADE

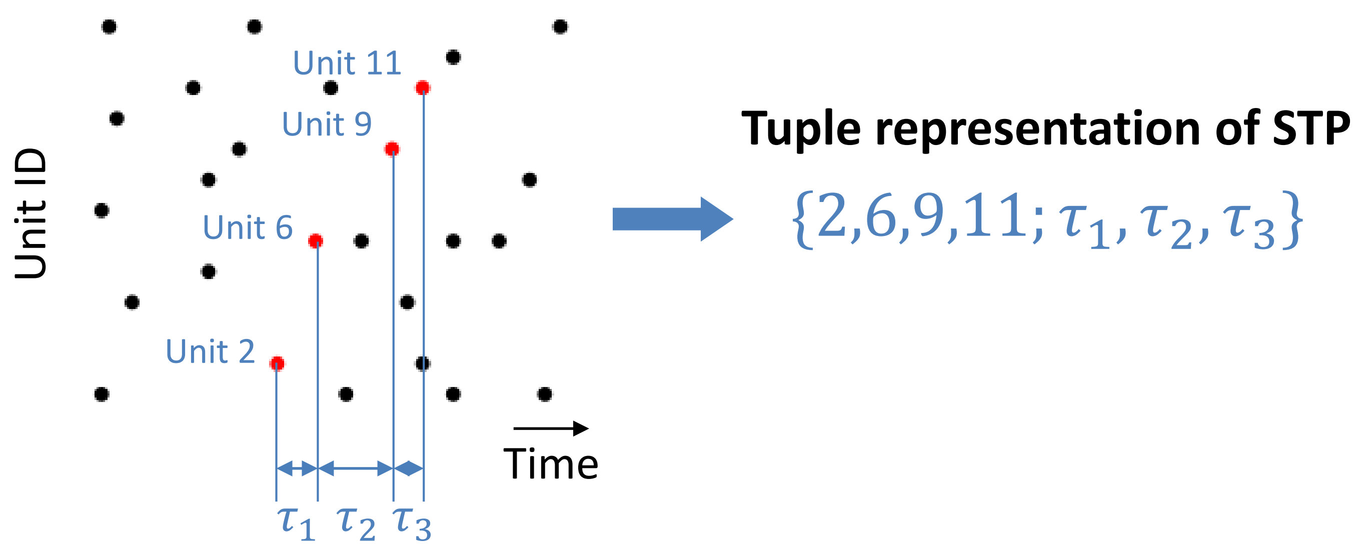
- Extension of the SPADE method for detection of synchronous spike patterns to detection of SPTs
- **Main idea**: reformat the spike trains to represent STPs at each time bin



Yegenoglu et al. (2016) *Lect Note Comput Sci*; Quaglio et al. (2017) *Front Comp Neurosci*; Quaglio et al (2017) *Biol Cybern*; Stella et al. (2019) *Biosystems*

3d-SPADE | problem

- There are more longer patterns than shorter patterns per signature
  - Example: patterns of unit 1, 2 and 3
    - Duration: 1 bin → $\{1,2,3;0,0\}$
    - Duration: 2 bins → $\{1,2,3;0,1\},\{1,2,3;1,0\},\{1,3,2;0,1\}$ $\{2,1,3;1,0\},\{2,3,1;0,1\},\{3,1,2;1,0\}$
    - Duration: 3 bins → even more…



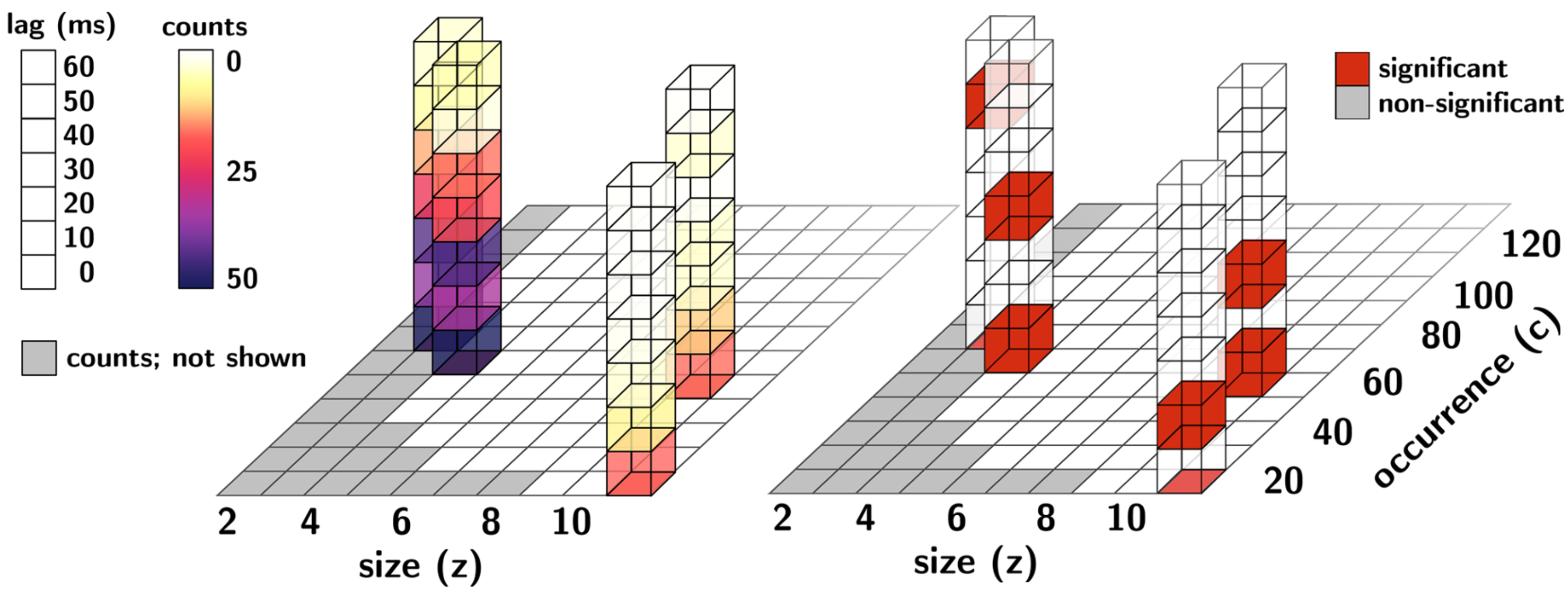
3d-SPADE | problem

- There are more longer patterns than shorter patterns per signature
  - Example: patterns of unit 1, 2 and 3
    - Duration: 1 bin → $(1,2,3;0,0)$
    - Duration: 2 bins → $(1,2,3;0,1), (1,2,3;1,0), (1,3,2;0,1)$
      $(2,1,3;1,0), (2,3,1;0,1), (3,1,2;1,0)$
    - Duration: 3 bins → even more…
  - A signature $(3, c)$ is occupied more by longer patterns

- **Solution**
  - extend the signature to 3-dimensional as (size z, occurrence c, duration d)
  - Accordingly, extend the pattern spectrum to 3-dimensional

3d-SPADE | 3d spectra

- **Pattern spectrum**: 3-dimensional histogram of size z, occurrence c, and duration z of CFISs
- **Significance spectrum**: 3-dimensional binary matrix indicating the significance of CFISs with signature (z, c, d)



Stella et al. (2019) *Biosystems*

3d-SPADE | application

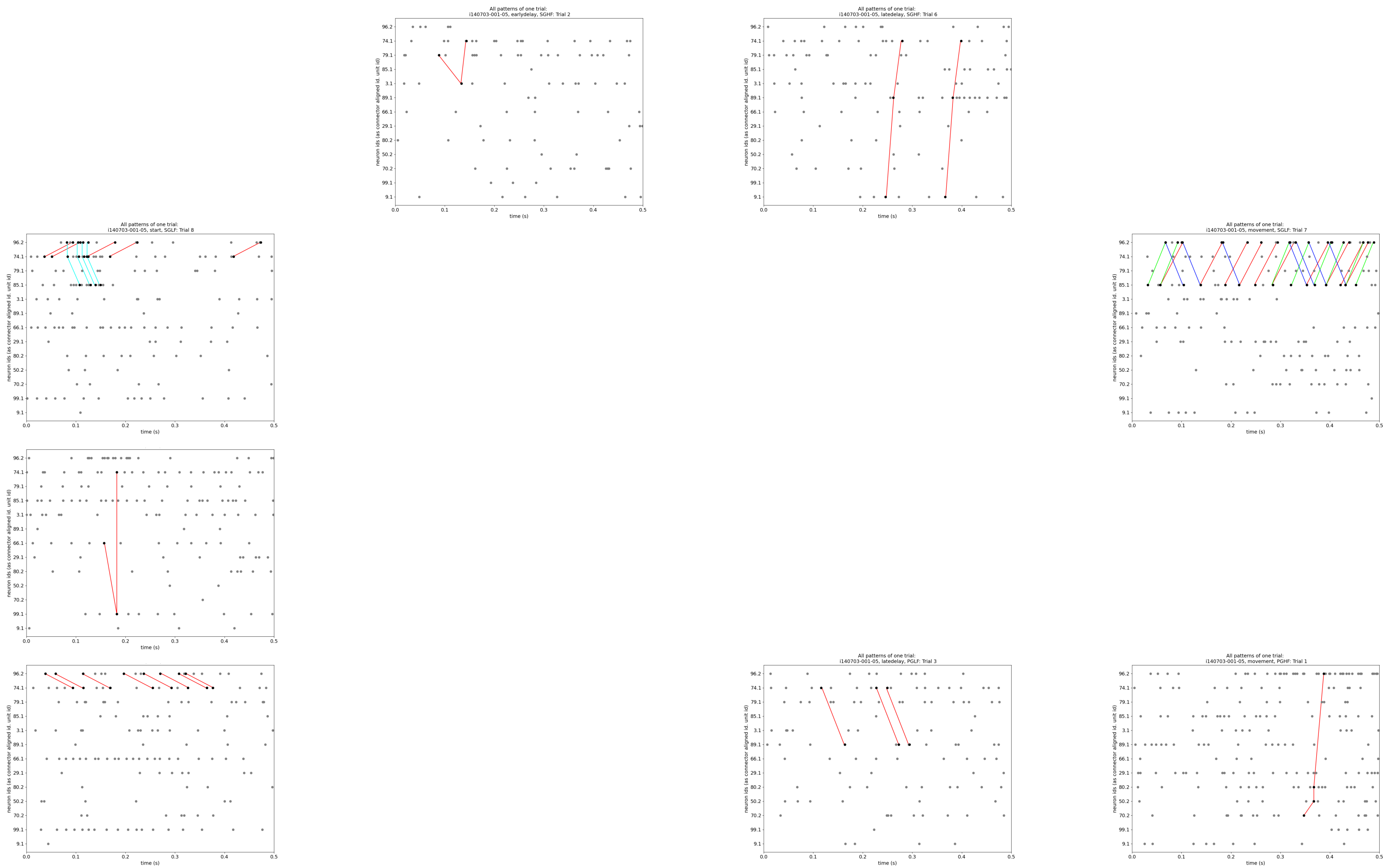
start | early-d | late-d | mov

SGHF

SGLF

PGLF

PGHF



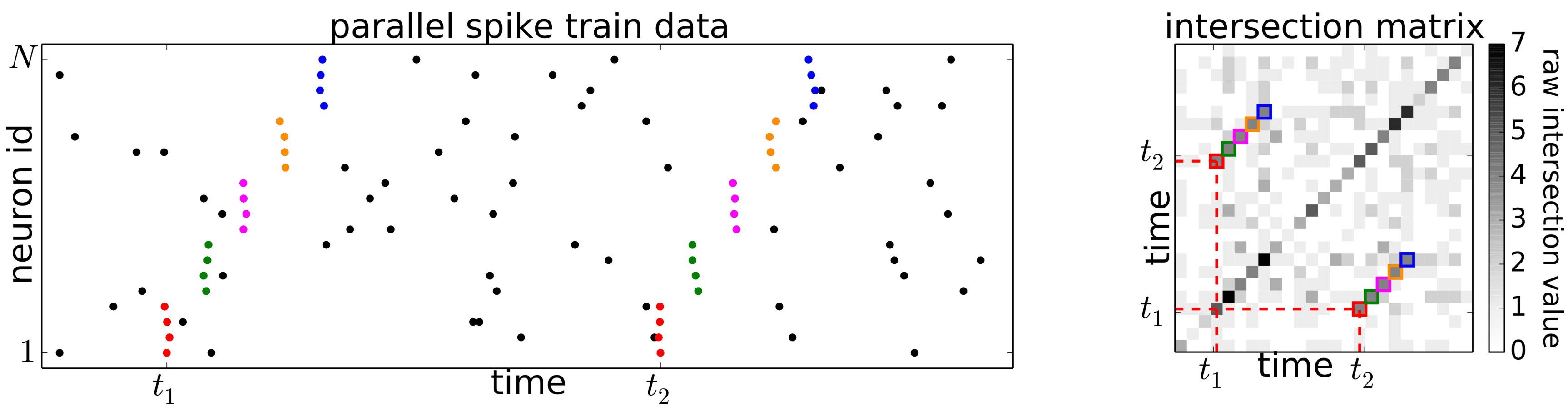
Stella et al. *in prep.*

Outline

- SPADE for synchronous pattern identification
  - methods
  - application
- **Extension to spatio-temporal pattern identification**
  - Beyond synchronous spike patterns
  - 3d-SPADE
  - ASSET

ASSET

- Short for “Analysis of Sequences of Synchronous EvenTs”
- Designed specifically for detecting sequences of synchronous spike patterns
- **Intersection matrix** I: degree of overlap between unit membership in synchronous spike patterns at two time points
- Test the significance of diagonal patterns in I



Schrader et al. (2008) *J Neurophysiol*; Gerstein et al. (2012) *J Neurosci Meth*; Torre et al. (2016) *PLoS CB*

Summary

- Various methods for analysis of HOCs
  - Complexity distribution, SPADE, 3d-SPADE, ASSET, CuBIC (Staude, Grün and Rotter, 2010; Staude, Rotter and Grün, 2010), and so on…
  - Assume different models of HOC, and design statistical tests for detecting expected patterns
- Common problems: combinatorial explosion of patterns and multiple testing problem
- Surrogate is a versatile and flexible method for estimating p-values.

Literature

- Overview of spike pattern detection:
  Quaglio P., Rostami V., Torre E., Grün S. (2018)
  Methods for identification of spike patterns in massively parallel spike trains
  Biological Cybernetics **112**:57-80 DOI:10.1007/s00422-018-0755-0

… and more

- Grün S, Abeles M, and Diesmann M. (2008) Impact of higher-order correlations on coincidence distributions of massively parallel data. Lecture Notes in Computer Science, 'Dynamic Brain - from Neural Spikes to Behaviors' 5286:96-114. DOI:10.1007/978-3-540-88853-6
- Louis S, Borgelt C, Grün S (2010) Complexity distribution as a measure for assembly size and temporal precision Neural Networks 23:705-712. DOI:10.1016/j.neunet.2010.05.004
- Staude B, Rotter S, Grün S (2010) CuBIC: cumulant based inference of higher-order correlations in massively parallel spike trains. J Comput Neurosci 29(1-2):327-350. DOI: 10.1007/s10827-009-0195-x
- Staude B, Grün S, Rotter S. (2010) Higher-order correlations in non-stationary parallel spike trains: statistical modeling and inference. Front. Comput. Neurosci. 4:16. DOI:10.3389/fncom.2010.00016
- Staude B, Grün S, Rotter S (2010) Higher order correlations and cumulants. In: Analysis of parallel spike trains. eds. Grün & Rotter. Springer Series in Computational Neuroscience DOI:10.1007/978-1-4419-5675-0_12
- Picado-Muiño D., Borgelt C., Berger D., Gerstein G., Grün S. (2013). Finding neural assemblies with frequent item set mining. Frontiers in Neuroinformatics 7:9. DOI: 10.3389/fninf.2013.00009.
- Torre E., Picado-Muiño D., Denker M., Borgelt C., Grün S. (2013). Statistical evaluation of synchronous spike patterns extracted by frequent item set mining. Frontiers in Computational Neuroscience 7:132. DOI: 10.3389/fncom.2013.00132.
- Torre E., Canova C., Denker M., Gerstein G., Helias M., Grün S. (2016b). ASSET: analysis of sequences of synchronous events in massively parallel spike trains. PLOS Computational Biology 12:e1004939. DOI: 10.1371/journal.pcbi.1004939.
- Torre E., Quaglio P., Denker M., Brochier T., Riehle A., Grün S. (2016a). Synchronous spike patterns in macaque motor cortex during an instructed-delay reach-to-grasp task. Journal of Neuroscience 36:8329–8340. DOI: 10.1523/JNEUROSCI.4375-15.2016.
- Yegenoglu A, Quaglio P, Torre E, Grün S, Enders D. (2016) Exploring the Usefulness of Formal Concept Analysis for Robust Detection of Spatio-Temporal Spike Patterns in Massively Parallel Spike Trains. In: Graph-Based Representation and Reasoning 22nd International Conference on Conceptual Structures, ICCS 2016, Annecy, France. pp 3-16. DOI:10.1007/978-3-319-40985-6_1 ISBN: 978-3-319-40984-9
- Quaglio P., Yegenoglu A., Torre E., Endres DM., Grün S. (2017). Detection and evaluation of spatio-temporal spike patterns in massively parallel spike train data with spade. Frontiers in Computational Neuroscience 11:41. DOI: 10.3389/fncom.2017.00041.