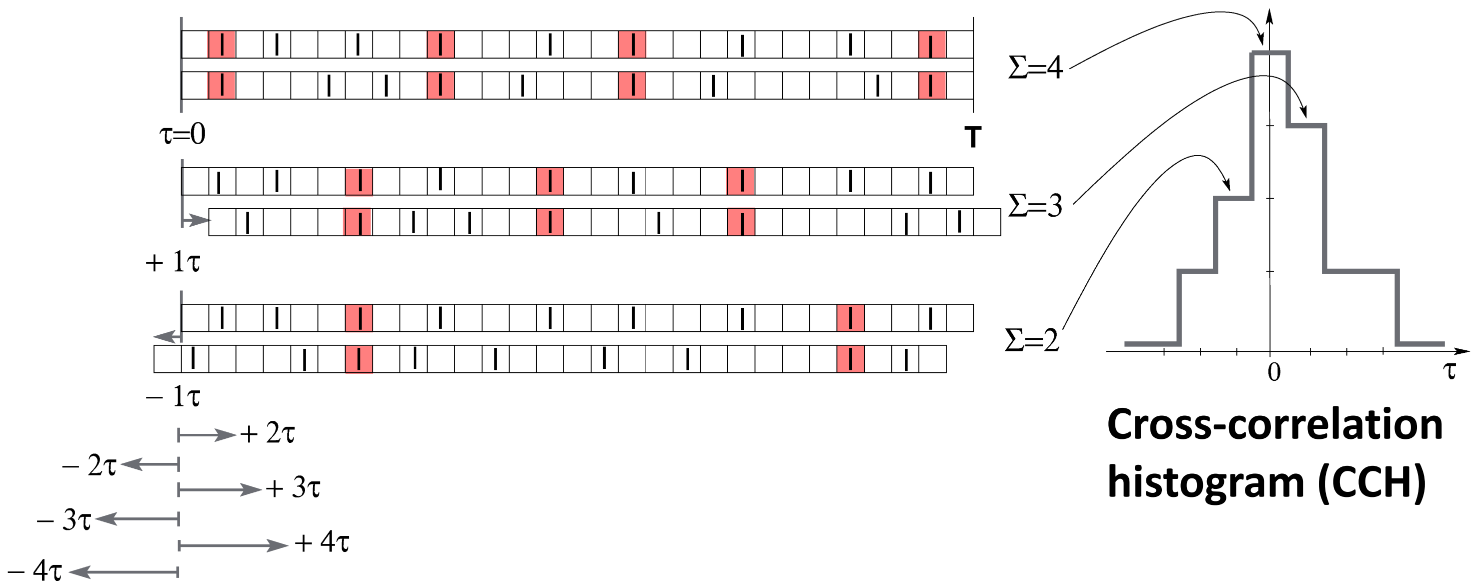
# Spike Train Analysis III:
## unitary events and higher-order correlations

**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

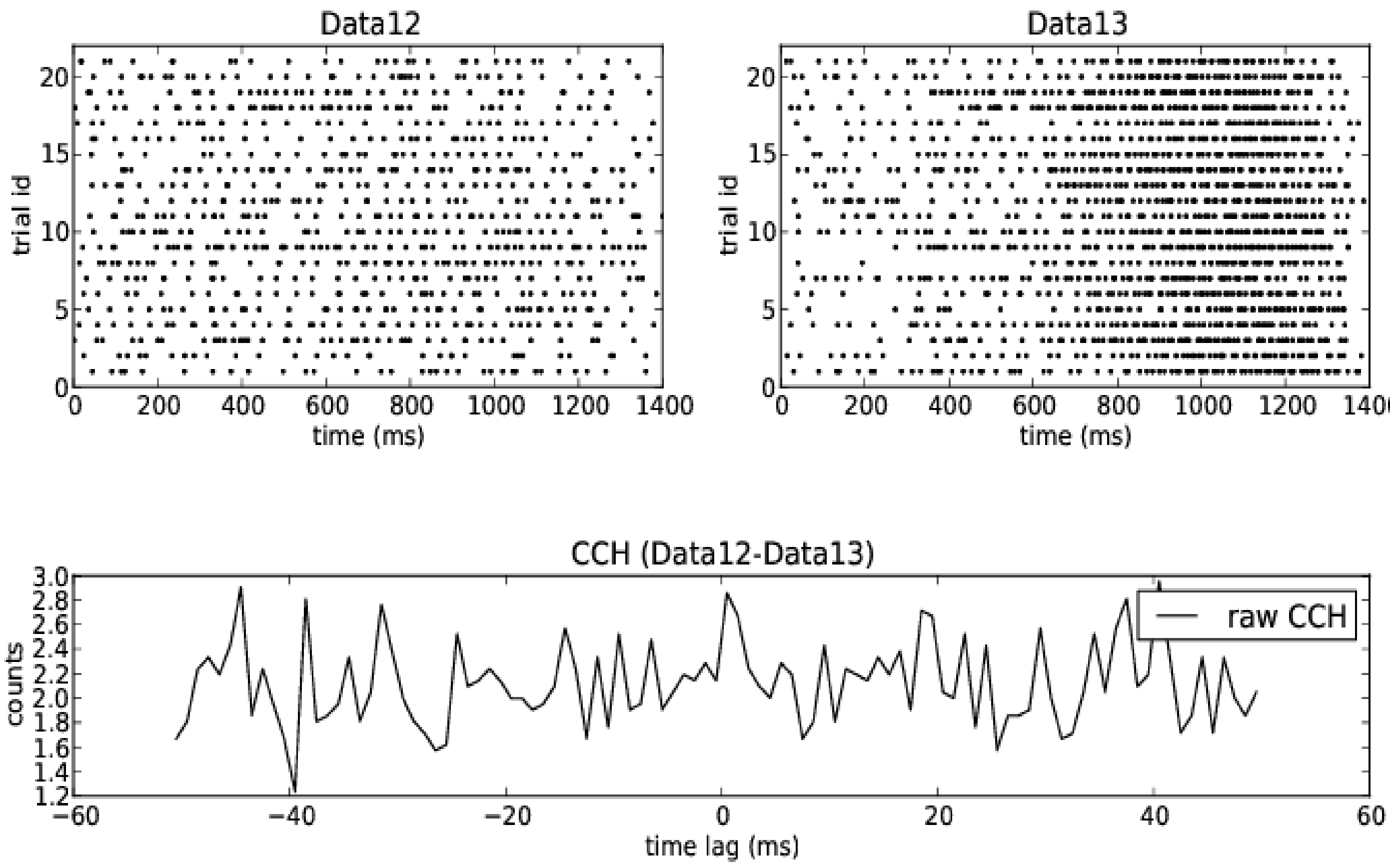
9th Latin American School on Computational Neuroscience (LASCON 2024)
NeuroMat, University of Sao Paulo, Sao Paulo, Brazil | January 24, 2024

# Recap | correlation analysis




**Cross-correlation histogram (CCH)**

# **Recap** | correlation analysis



# **Motivation** | drawbacks of correlation analysis

- Correlations identified by the CCH is “averaged over time”.
- But, neurons may change dynamically their correlation in time.
- Need for methods for time-resolved correlation analysis
  - Joint peri-stimulus time histogram (JPSTH)
    Aertsen et al. (1989) *J Neurophysiol* **61**(5):900–917
    Vaadia et al. (1995) *Nature* **373**:515--518
  - **Unitary event analysis**
    Riehle et al. (1997) *Science* **278**:1950-1953
    Kilavic et al. (2009) *J Neurosci* 28(40):12653-63

# Outline

- **Unitary event (UE) analysis**
  - methods
  - application
- **Higher-order correlation**
  - massively parallel spike trains
  - pairwise correlation-based approach
  - complexity distribution

# Methods | basic idea

- Bin the time axis, to represent parallel spike trains as binary processes


- Select a joint-spike 0-1 pattern across neurons (e.g., 0101000)

- Count the empirical number of occurrences: $n_{\mathrm{emp}}$

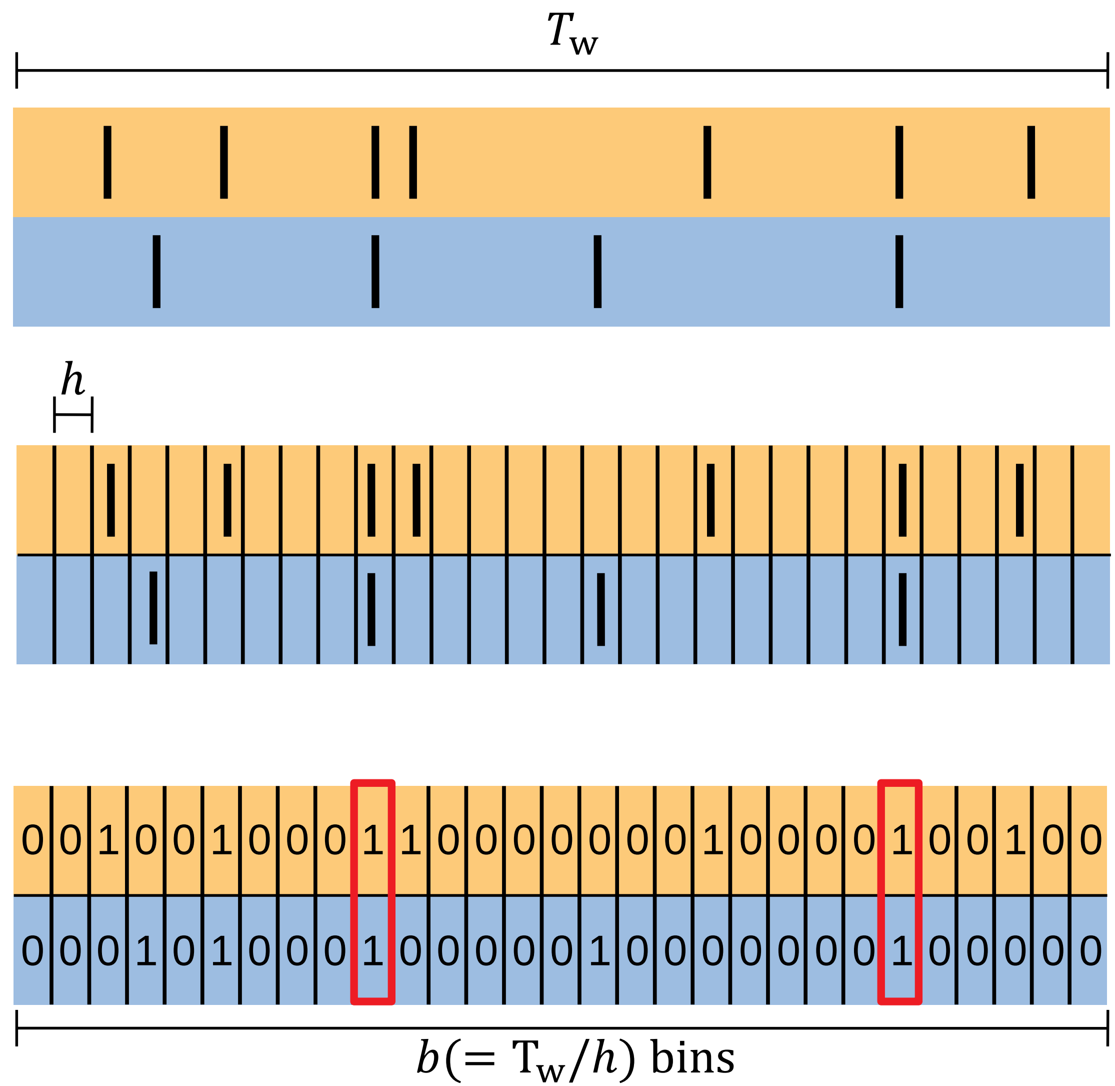
- Compute an expected number of occurrences: $n_{\mathrm{exp}}$

- Derive a null-distribution of occurrence count (conditioned by $n_{\mathrm{exp}}$), to test the significance of $n_{\mathrm{emp}}$

- If significant: **unitary events**

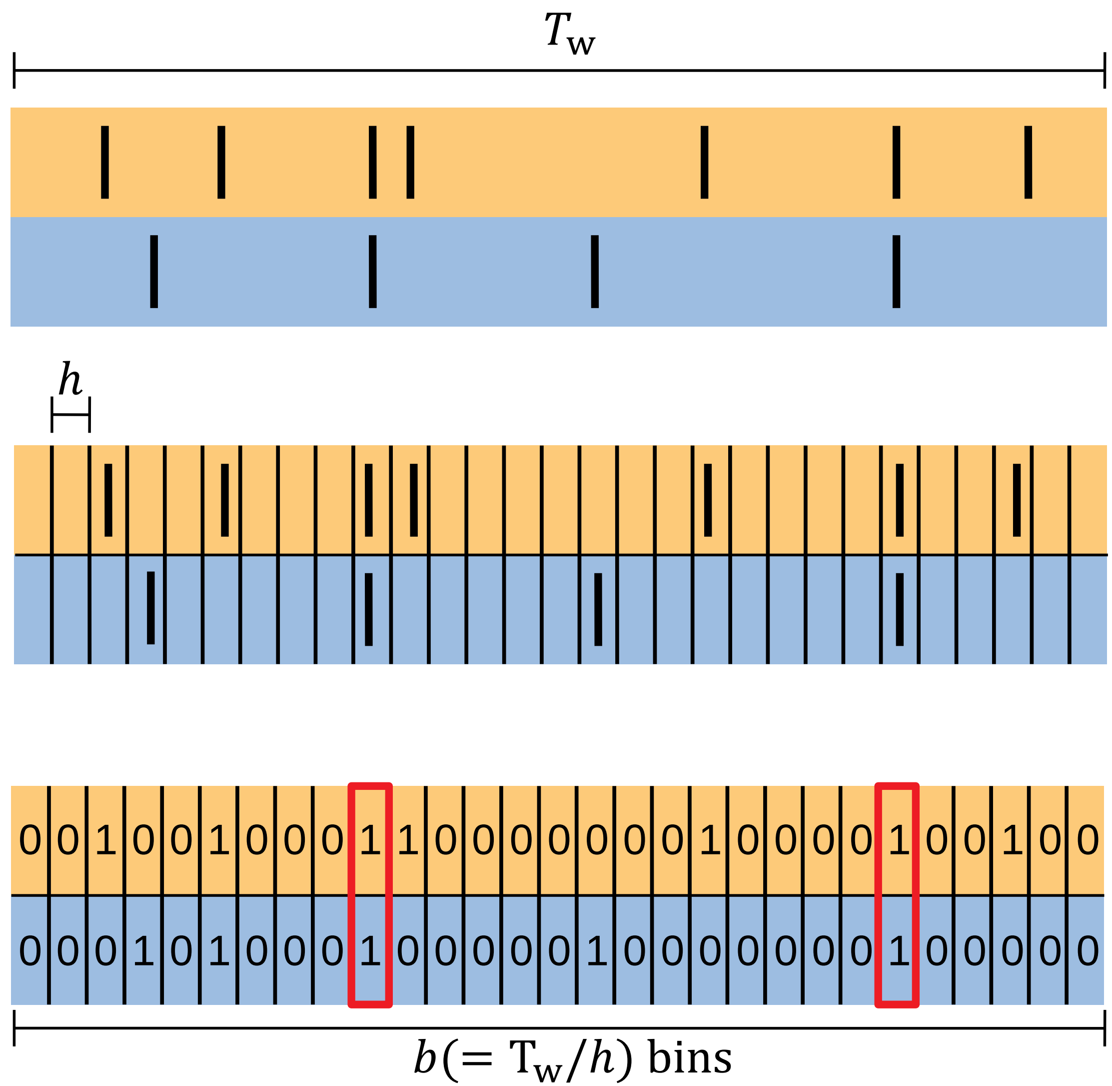
# **Methods** | formulation in a pairwise case

- Simultaneously recorded spike trains of unit 1 and unit 2
- Focus on the activity within a time window $T_\mathrm{w}$
- Discretise the time window $T_w$ into $b$ bins of width $h$: $b = T_w/h$
- Count the number $n_\mathrm{emp}$ of coincident spikes, i.e., the number of bins occupied by spikes of both unit 1 and 2



# **Methods** | derivation of $n_{\text{exp}}$

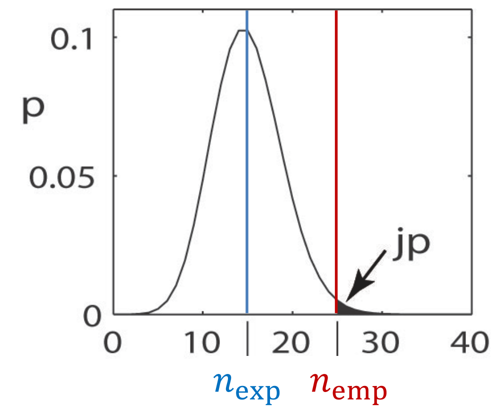
- Count the number $n_1$ of spikes of unit 1
- Probability $p_1$ of unit 1 firing in a bin: $\mathrm{p}_1 = n_1/b$ (Poissonity assumption)
  - Also, $p_2 = n_2/b$
- Probability $p_{12}$ of unit 1 and 2 firing together in a bin: $p_{12} = p_1 p_2$ (independency assumption)
- Expected number $n_{\text{exp}}$ of coincident spikes: $n_{\text{exp}} = p_{12} b = n_1 n_2 h/T_w$



# **Methods** | derivation of null-distribution

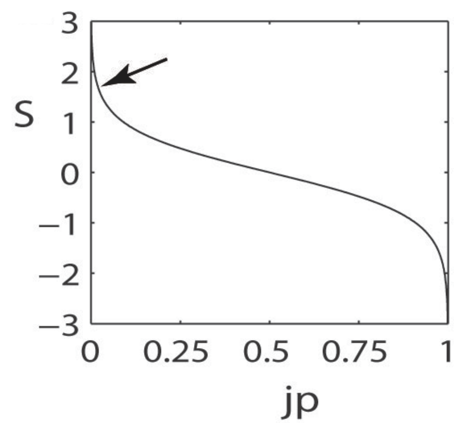
- Null-distribution of coincidence count
  - If the spike trains are independent Poisson processes, the spike coincidences should also be a Poisson process
    → null-distribution: **Poisson distribution** with mean $n_{\text{exp}}$

- Derive the p-value jp of $n_{emp}$ from this null-distribution




- For visualization: express the p-value jp as **surprise** $S$:

$$S(\text{jp}) = \log_{10}\frac{1-\text{jp}}{\text{jp}}$$



# **Methods** | time-resolved analysis

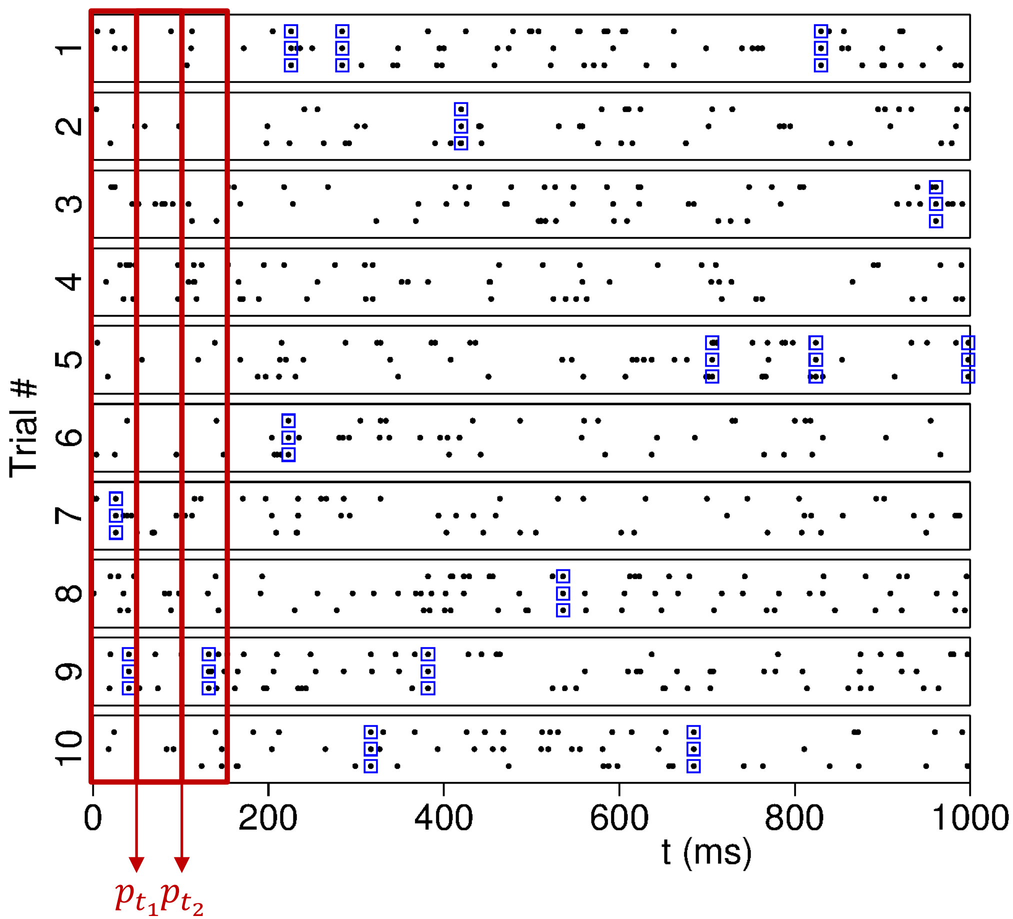
- Experimental data are typically recordings across multiple trials.

- General interest: how activity changes over time within a trial (non-stationarity in trial-time), whereas assuming stationarity across trials

- Approach
  - accumulate statistics (such as $n_{\mathrm{emp}}$ and $n_{\mathrm{exp}}$) across trials, within a certain (short) time interval defined on trial-time: **analysis time window**
  - systematically shift the position of the time window along trial-time: **sliding time window analysis**

# **Methods** | time-resolved analysis

- Consider a time window of width $T_{\mathrm{w}}$, starting at the beginning of trial
- For each trial, compute $n_{\mathrm{emp}}$ and $n_{\mathrm{exp}}$ from the data within the window, and sum them up over the trials
- Derive the p-value from the summed $n_{\mathrm{emp}}$ and $n_{\mathrm{exp}}$
- Shift the window position by a fixed time amount (**window step** $\leq T_{\mathrm{w}}$), and perform the same analysis
- Repeat this until the window reaches the end of trial



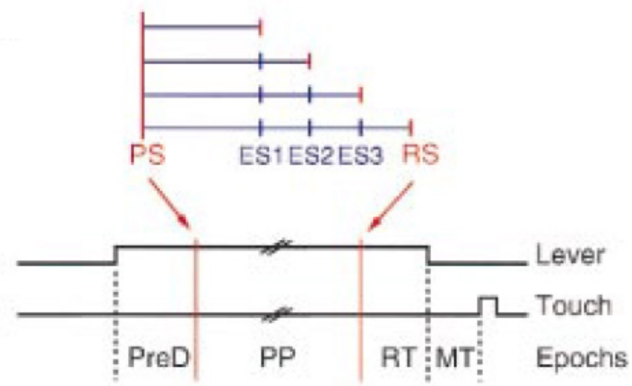
# Outline

- **Unitary event (UE) analysis**
  - methods
  - application
- **Higher-order correlation**
  - massively parallel spike trains
  - pairwise correlation-based approach
  - complexity distribution

# **Application** | expectation-related synchrony

- Two macaque monkeys were trained for a delayed reaching task
  - The duration of the delay was selected from four possible durations: 600, 900, 1200, and 1500 ms, and varied randomly from trial to trial.

**Task protocol**




Riehle et al. (1997) *Science*

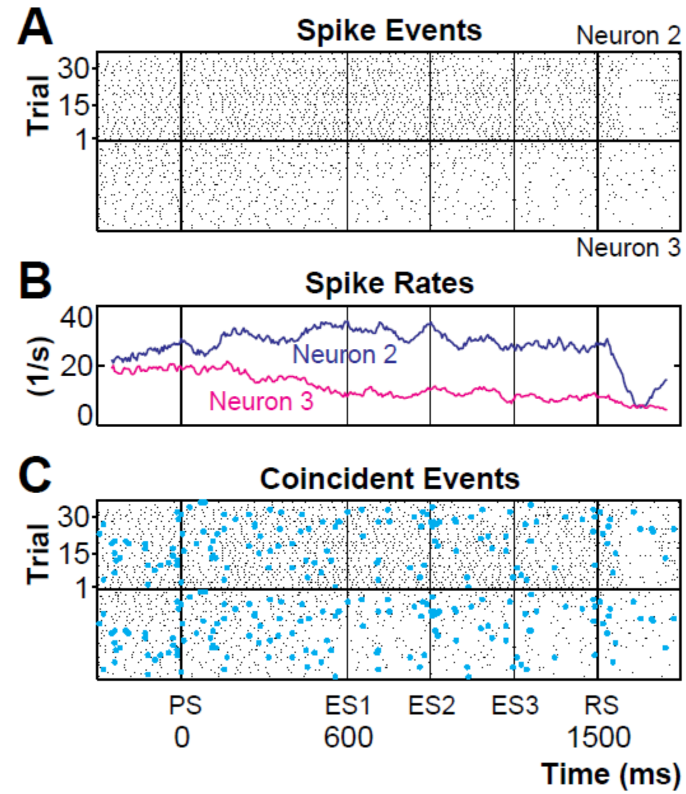
- Spiking activity of the primary motor cortex was recorded during the performance of the task.

- Spike trains of simultaneously recorded single units were analysed for excess synchrony beyond chance coincidence, using the UE analysis.

# **Application** | expectation-related synchrony

- Spike trains of a simultaneously recorded pair of single units, for 36 trials with the longest delay

- No strong modulations of firing rates in relation to the expectation



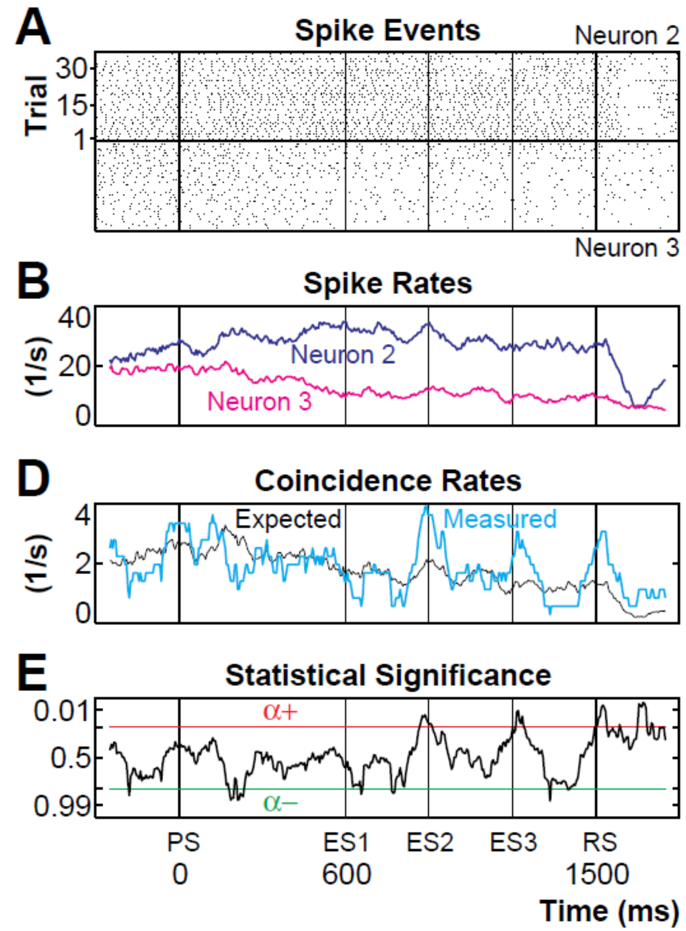

Riehle et al. (1997) *Science*

# **Application** | expectation-related synchrony

- Spike trains of a simultaneously recorded pair of single units, for 36 trials with the longest delay

- No strong modulations of firing rates in relation to the expectation

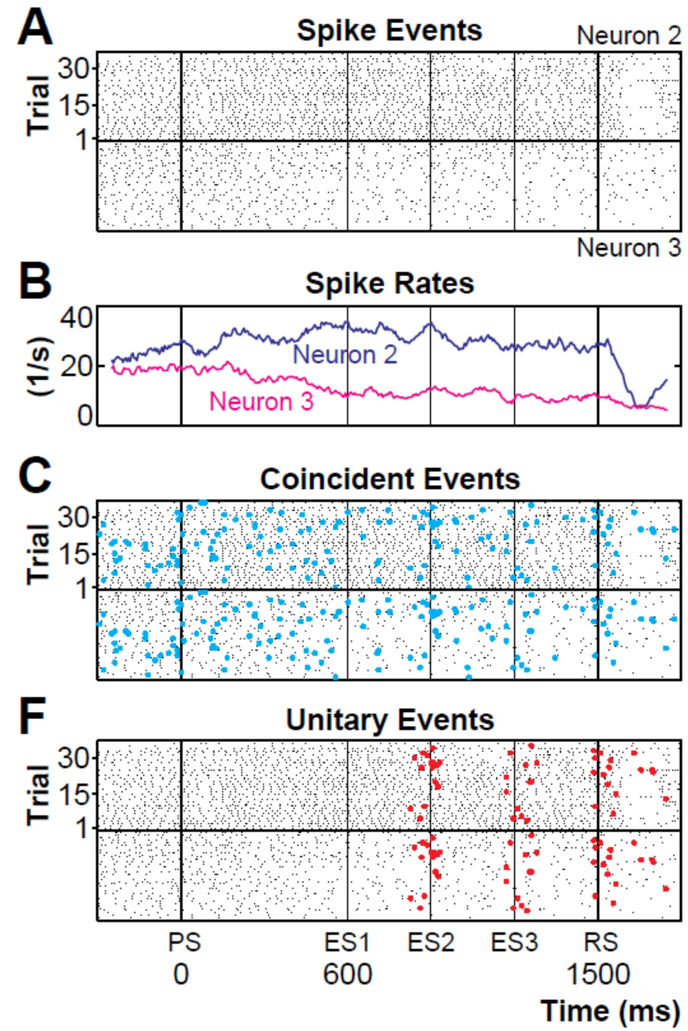
- Marked rises in the coincidence rate at the timings of the expectation




Riehle et al. (1997) *Science*

# **Application** | expectation-related synchrony

- Spike trains of a simultaneously recorded pair of single units, for 36 trials with the longest delay
- No strong modulations of firing rates in relation to the expectation
- Marked rises in the coincidence rate at the timings of the expectation
- Thus, spike synchronization and rate modulation are differentially involved in motor cortical function.



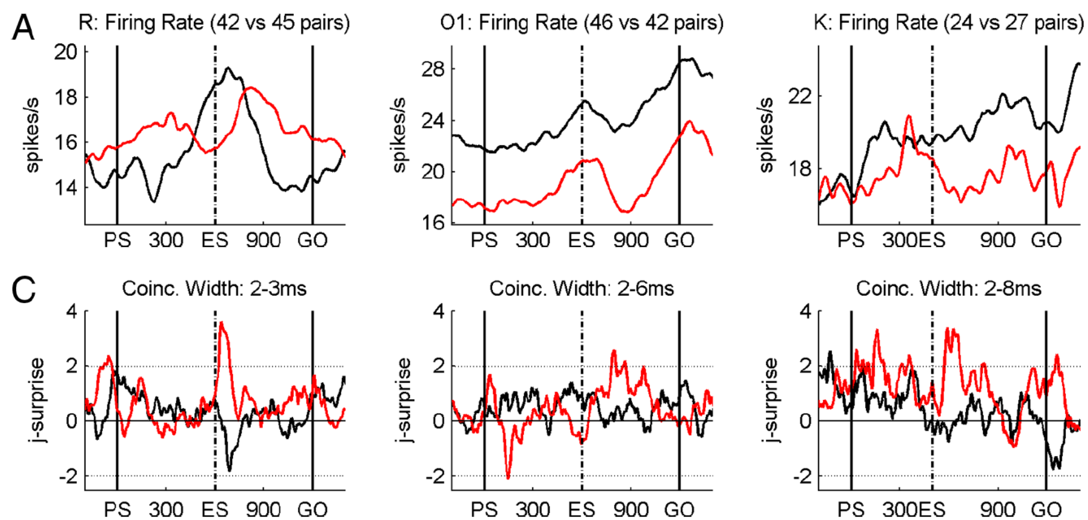

Riehle et al. (1997) *Science*

# **Application** | expectation-related synchrony

- Comparison between two populations of unit pairs recorded in the **1st** and **2nd** half of consecutive recording sessions
  - Monkeys were more trained for the task in the **2nd** half.
- Here $n_{\mathrm{emp}}$ and $n_{\mathrm{exp}}$ were summed over pairs of units, and the surprise was derived from those summed counts (**population UE analysis**)

- Generally lower firing rate in the **2nd** half

- Stronger expectation-related spike synchrony in the **2nd** half




Kilavik et al. (2009) *J NeuroSci*

# Unitary event analysis | summary

- Unitary event analysis allows for time-resolved investigation of event-related excess spike synchrony beyond chance level.
  - Important parameters
    - **bin size** $h$: allowed coincidence width
    - **analysis window size** $T_\mathrm{w}$: should be small enough such that *the Poissonity assumption* is valid, but also large enough to contain enough samples for the statistics

- Application to spike trains from the primary motor cortex revealed expectation-related excess spike synchrony.
  - The spike synchrony is modulated differently from the firing rates of single units.
  - Temporal locking of the synchrony to expected events is build over training.

# “Recipe” for unitary event analysis

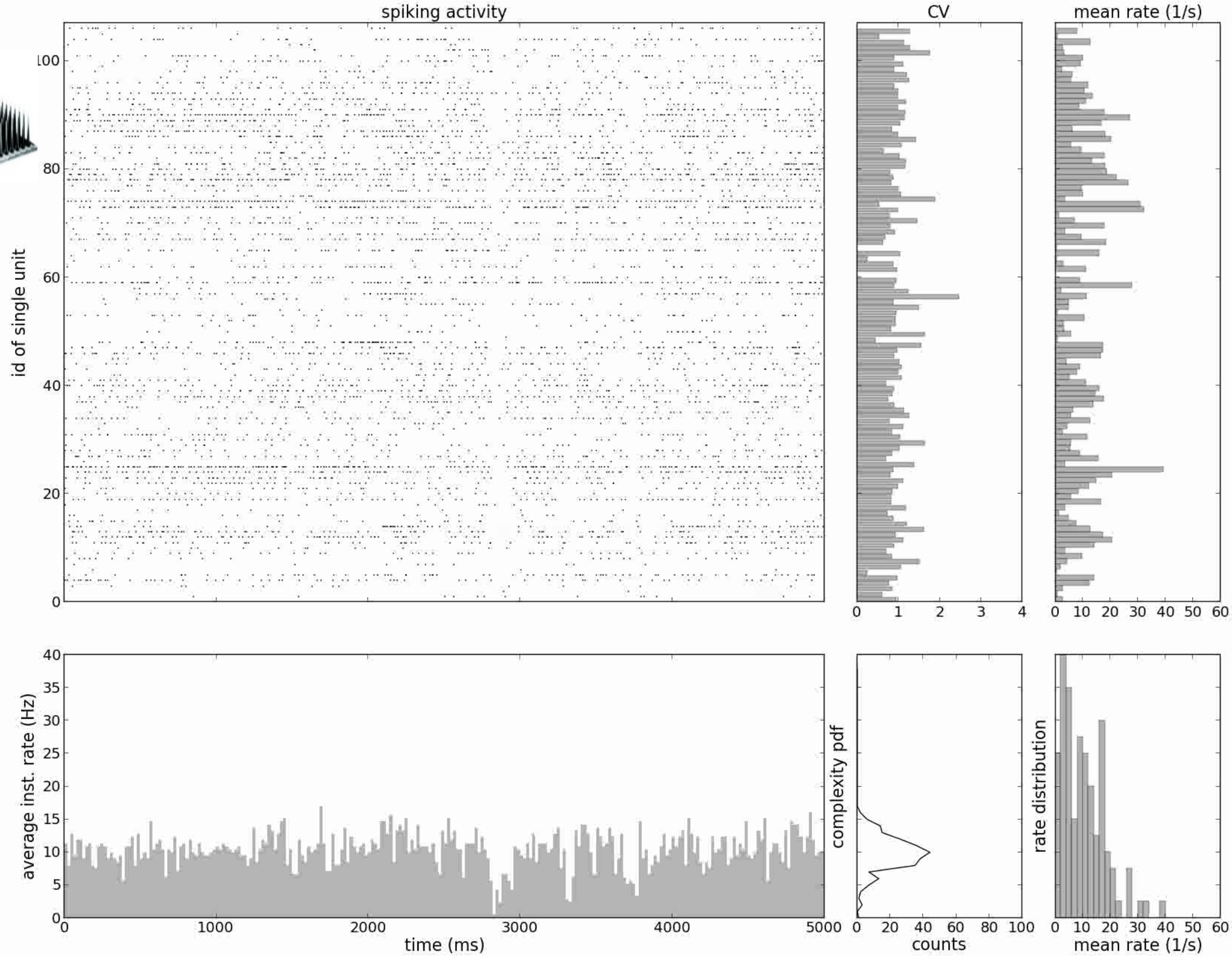
**Unitary Event computation**

1. Align trials, decide on width of analysis window.
2. Decide on allowed coincidence width.
3. Perform a sliding window analysis. In each window:
    a. Detect and count coincidences.
    b. Calculate expected number of coincidences.
    c. Evaluate significance of detected coincidences.
    d. If significant, the window contains Unitary Events.
4. Explore behavioral relevance of UE epochs.

# Outline

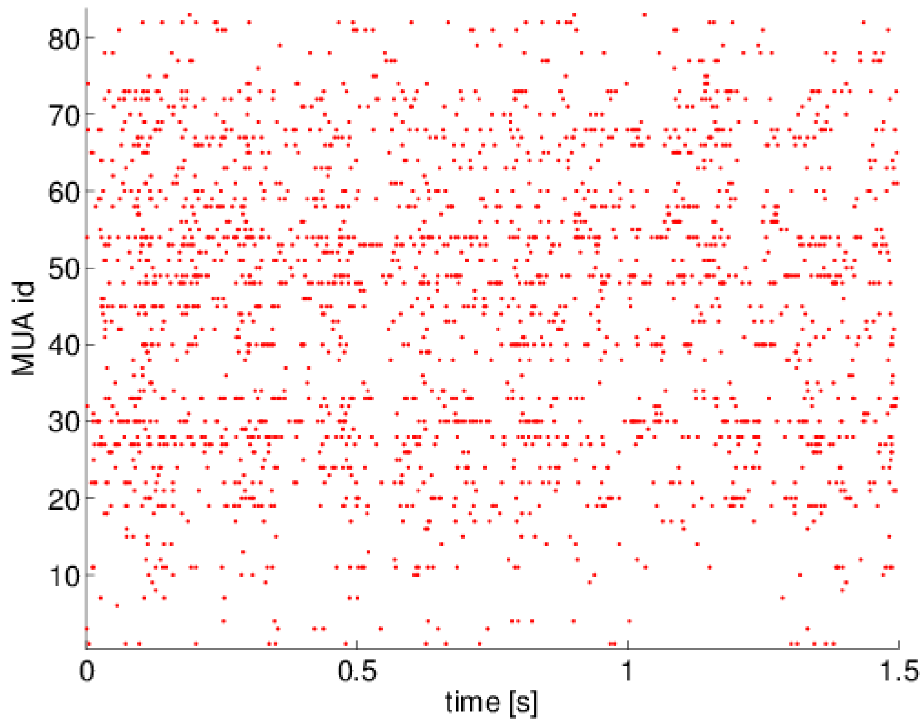
- **Unitary event (UE) analysis**
  - methods
  - application
- **Higher-order correlation**
  - massively parallel spike trains
  - pairwise correlation-based approach
  - complexity distribution

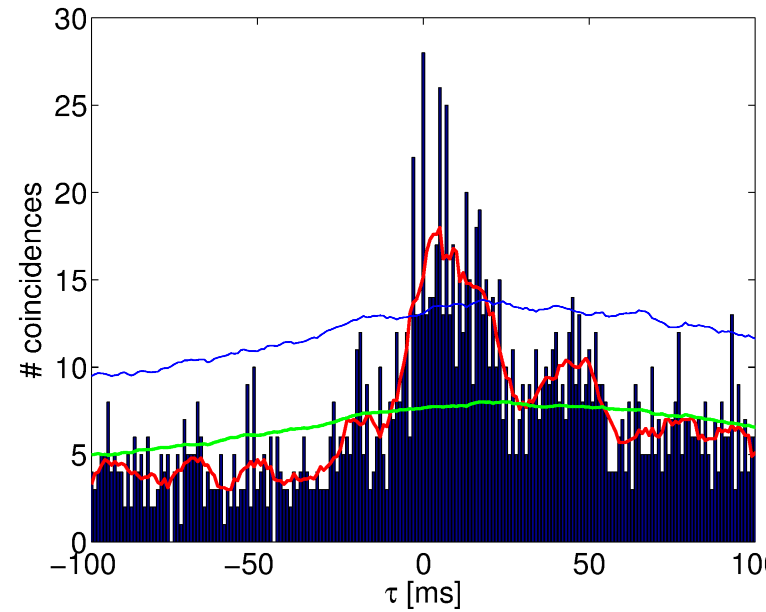
# Massively parallel spike trains




Collaboration with A. Riehle and T. Brochier, INT, CNRS-AMU, Marseille

# CCH-based approach | methods



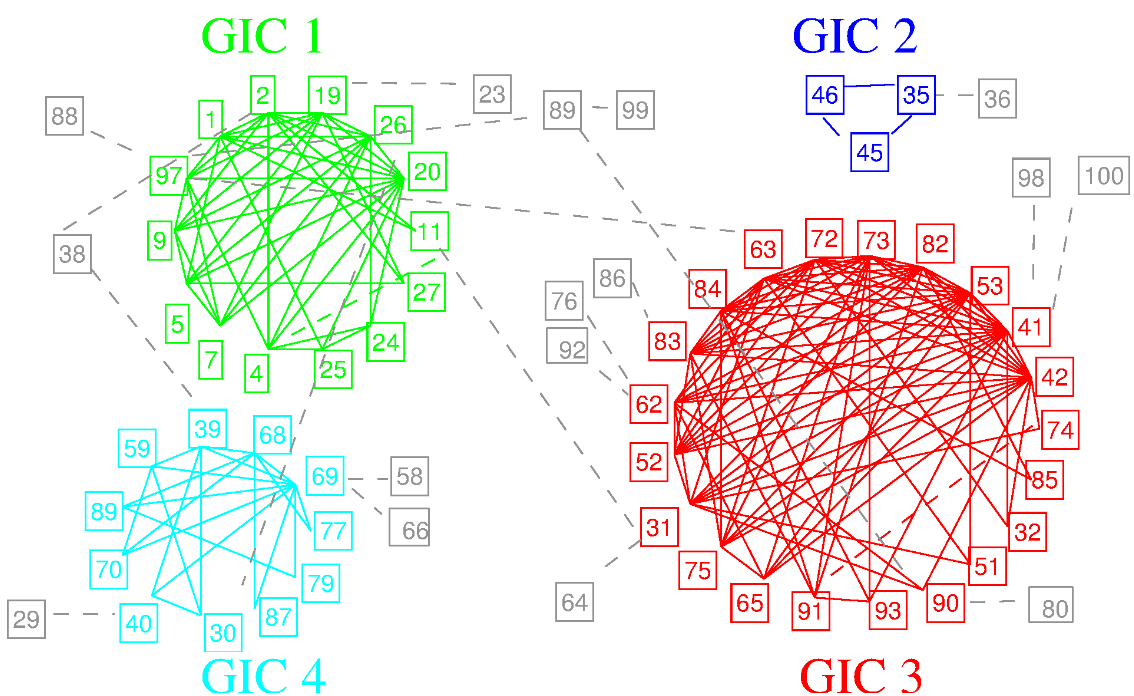




Berger et al. (2007) Neural Computing

- Utah array recording from macaque motor cortex
- Spike train of the multi unit activity (MUA) on each electrode
- Cross-correlation histogram for all pairs of MUAs
- Identify significant pairs

# **CCH-based approach** | clique grouping

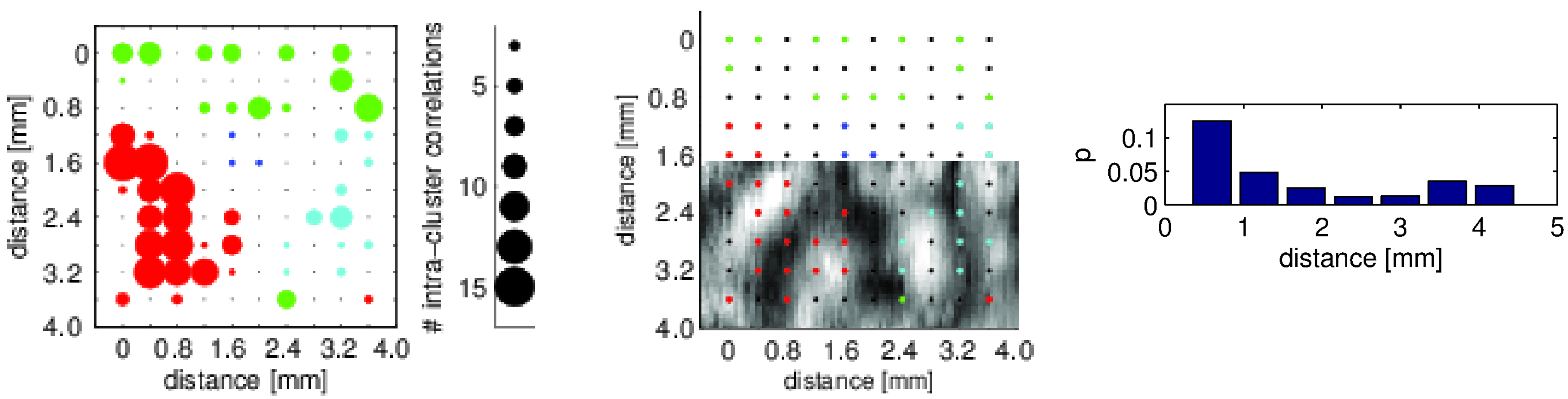
- **Number**: electrode ID
- **Line**: significant correlation between the respective MUAs
- Electrodes form a lot of cliques (fully-connected subsets)
- Cluster the cliques (size > 2) that overlap with at least 1 MUA
  → four distinct groups of mutually correlated MUAs




Berger et al. (2007) *Neural Comput*

# **CCH-based approach** | clique grouping

- Members of a cluster also cluster in cortical space.
- Spatial scale of the clusters are very similar to stimulus orientation domains in the visual cortex.
- Correlated pairs decrease with cortical distance, but increase again at an intermediate distance, also similar to the “patchy” connectivity in the visual cortex.



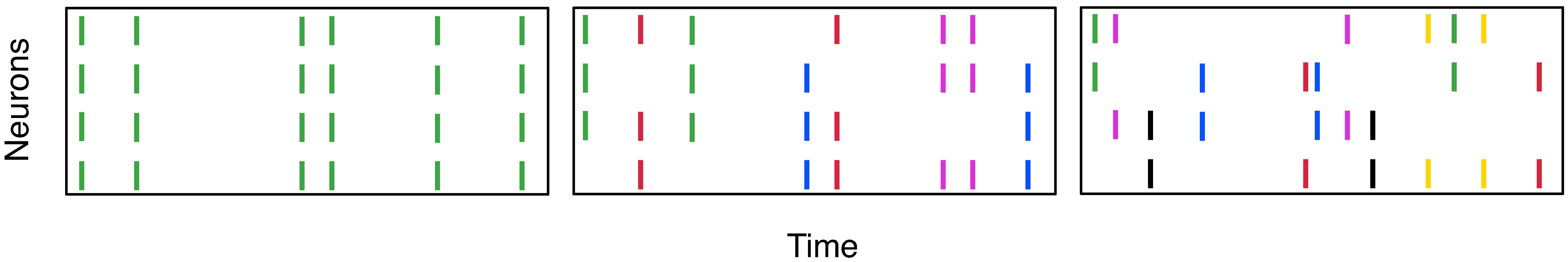

Berger et al. (2007) *Neural Comput*

# Outline

- Unitary event (UE) analysis
  - methods
  - application
- **Higher-order correlation**
  - massively parallel spike trains
  - pairwise correlation-based approach
  - complexity distribution

# 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).

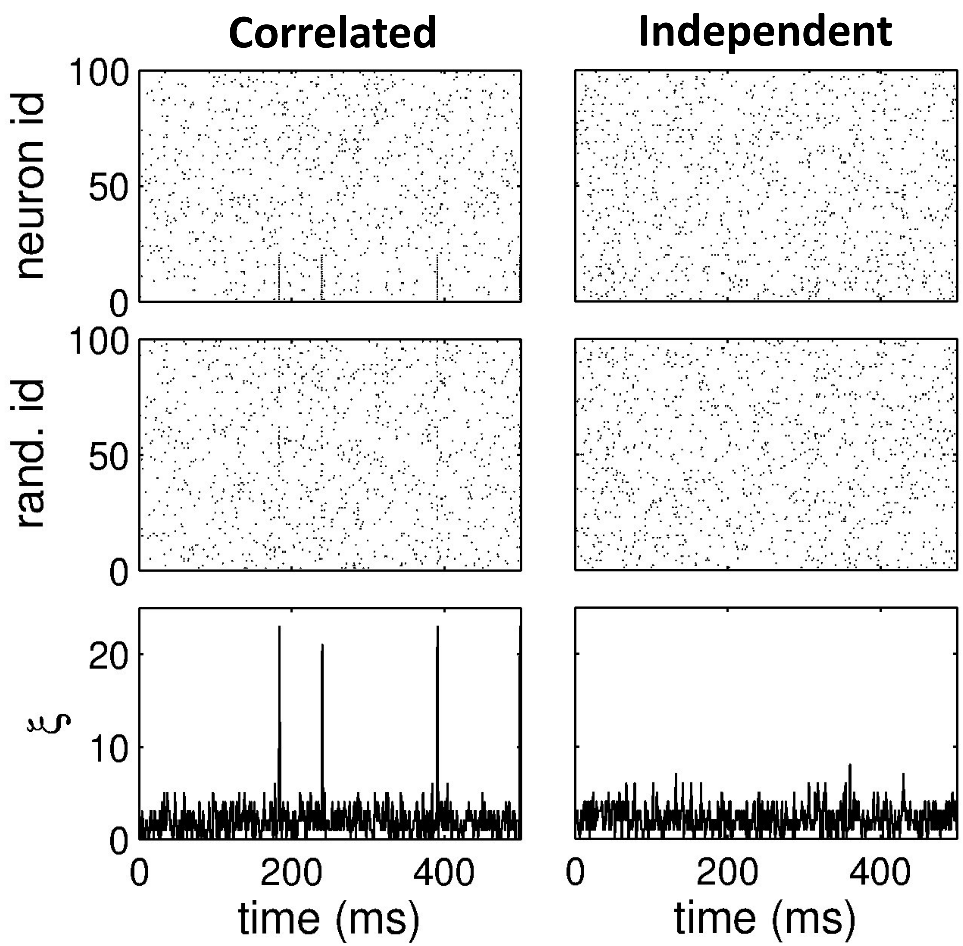
  → need for methods for detecting and analysing HOCs in massively parallel spike train data

# Higher-order correlation | challenges

- 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

# Population histogram

- Simplest approach: population spike time histogram, or **population histogram**

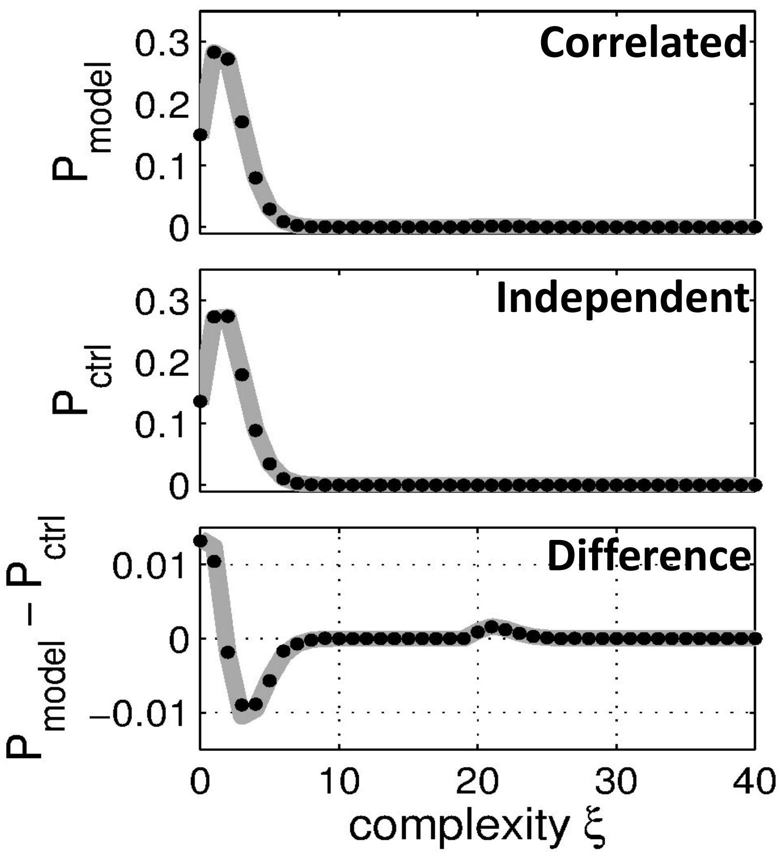



- $N$ Poisson spike trains with synchrony introduced only to $n$ units (correlated data)
- Independent control data generated by spike randomization surrogate
- Population histogram can easily detect this type of HOC

Grün et al. (2008) *Lecture Notes in Computer Science*
Louis et al. (2010) *Neural Networks*

# Complexity distribution

- **Complexity**: the number of units firing in a bin (i.e., the height of each entry in the population histogram)



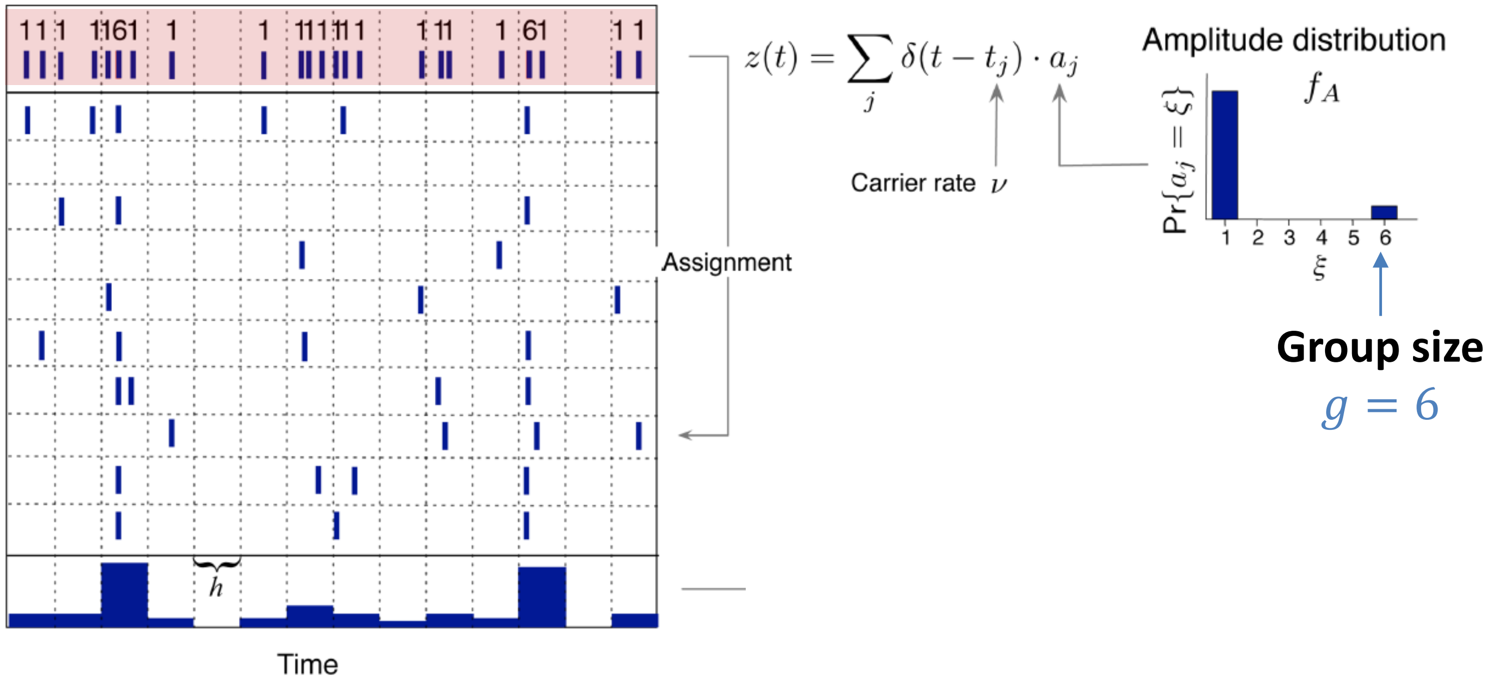

- Distribution of the complexity values $\xi$ obtained from the correlated data does not exhibit a pronounced peak at $n$.

- But the difference from the complexity distribution of the independent data clearly shows a "bump" at around $n$.

Grün et al. (2008) *Lecture Notes in Computer Science*
Louis et al. (2010) *Neural Networks*

# Compound Poisson process (CPP)

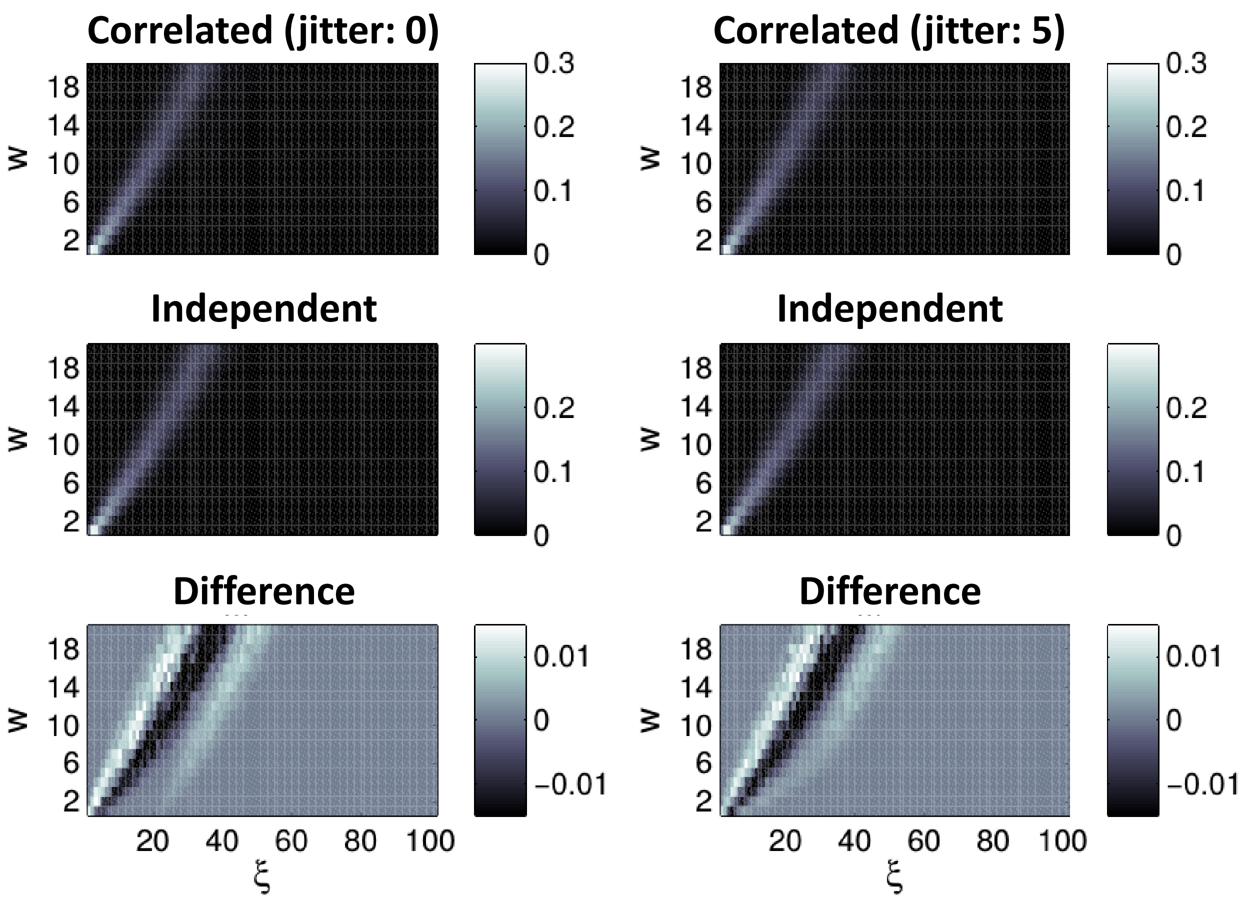
- Carrier process: a Poisson process with an "amplitude" $a_j$ at each event time $t_j$, drawn from an **amplitude distribution** $f_\mathrm{A}$
  - Carrier event at $t_j$ is copied as spikes of $a_j$ units
  - **Temporal jitter** of $\pm s$ ms is introduced to the $a_j$ spikes




Staude, Grün and Rotter (2010) *Front Comput Neurosci*

# Complexity distribution | application to CPP

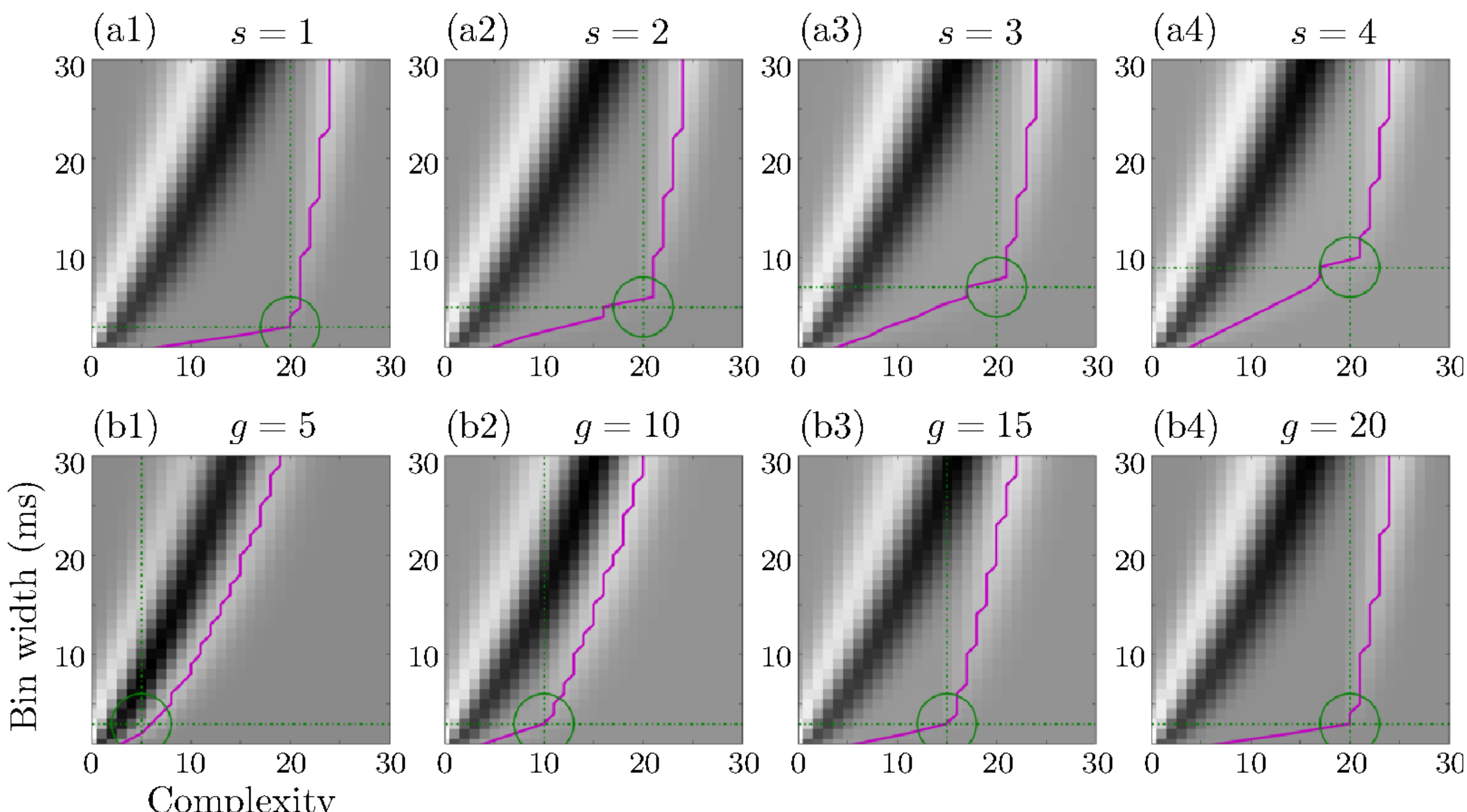
- Complexity distributions computed with varied bin sizes $w$




Louis et al. (2010) *Neural Networks*

# Complexity distribution | jitter and group size

- Trace of the “bump” shows a kink at $(g, 2s)$




Louis et al. (2010) *Neural Networks*

JÜLICH FORSCHUNGSZENTRUM

# Complexity distribution | summary

- Simple measure, easy to compute
- Use surrogates for generating control data sets
- Take the difference from the control data to highlight the excess synchrony by HOCs as a “bump”.
- Examine the distributions with various analysis bin size to estimate the group size and the temporal jitter of coincidences
- Drawback: the identity of the units participating in the group cannot be uniquely determined.