Basic issues in neural data analyses: Finding patterns

Jean-Marc Fellous

University of Arizona
1/6/2010

Summer of Spikes
University of Queensland

Neural Data Analyses

Neural Data Analysis: Incremental process

The 'master plan'

Experiment(s) -> Data

Analysis 1
Analysis 2
Analysis 3
…
Analysis N

- Typical, classical

- 'Conclusion' driven
- Result

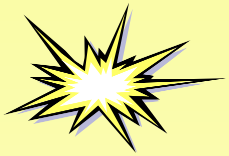
Analysis

Qualitative (phenomenological)

Quantitative
- Statistical (group data, trends…)
- Analytical

Neural Data Analyses

- Analysis results can suggest new analyses: **Combinatorial explosion**



Analysis 1
Analysis 2 → Analysis 1'
Analysis 3 Analysis 2'
… Analysis 3'
Analysis N …
Analysis N'

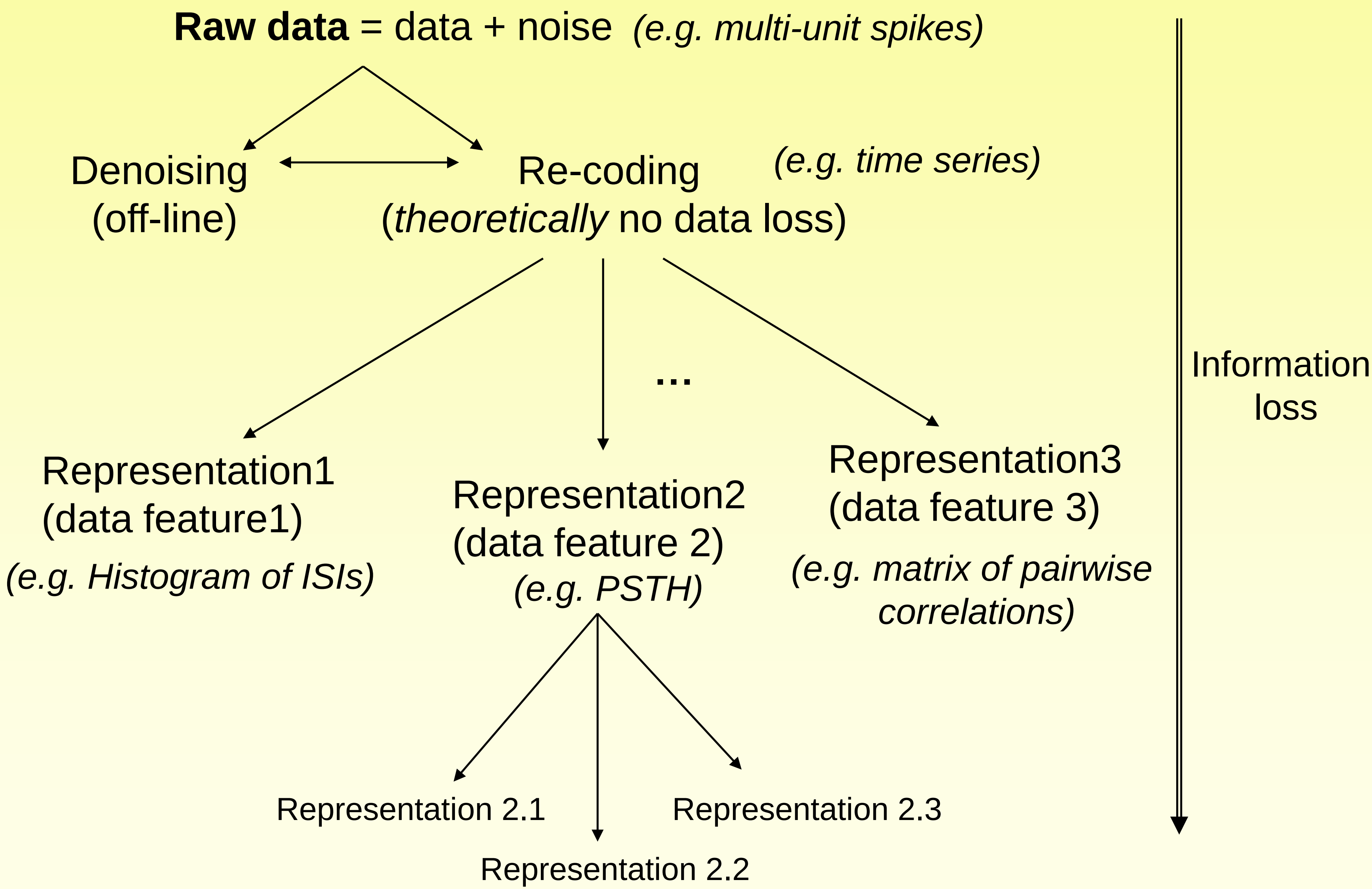
Compromise between depth/breadth first

- Analysis results can suggest new experiments: **Long time scales**

Analysis 1
Analysis 2 → Experiment1 -> data1 → Analysis 1.1, Analysis 2.1, Analysis 3.1, …, Analysis P.1
Analysis 3
…
Analysis N

Careful planning/design of the initial experiment(s)
(controls, alternative hypotheses…)

Data Representations



Data Representations

Analysis level / type *j*

Representation *i*
(data feature *i*)

Archiving (database)

<u>Backups</u>

- Raw data (permanent, multiple copies)
- Various representations (depends on amount of processing of raw data)
- Code (permanent, multiple copies)

Showing Data Analyses: The typical progression

1 - Analysis method. Use surrogate dataset, simulation data set, cartoon.

2 - Show typical single cell examples (raw data): voltage traces, rasterplots.

3 - Show a single cell analysis: Extract interesting feature(s) from step 2.

4 - Show population results: statistical analyses, population features, controls.

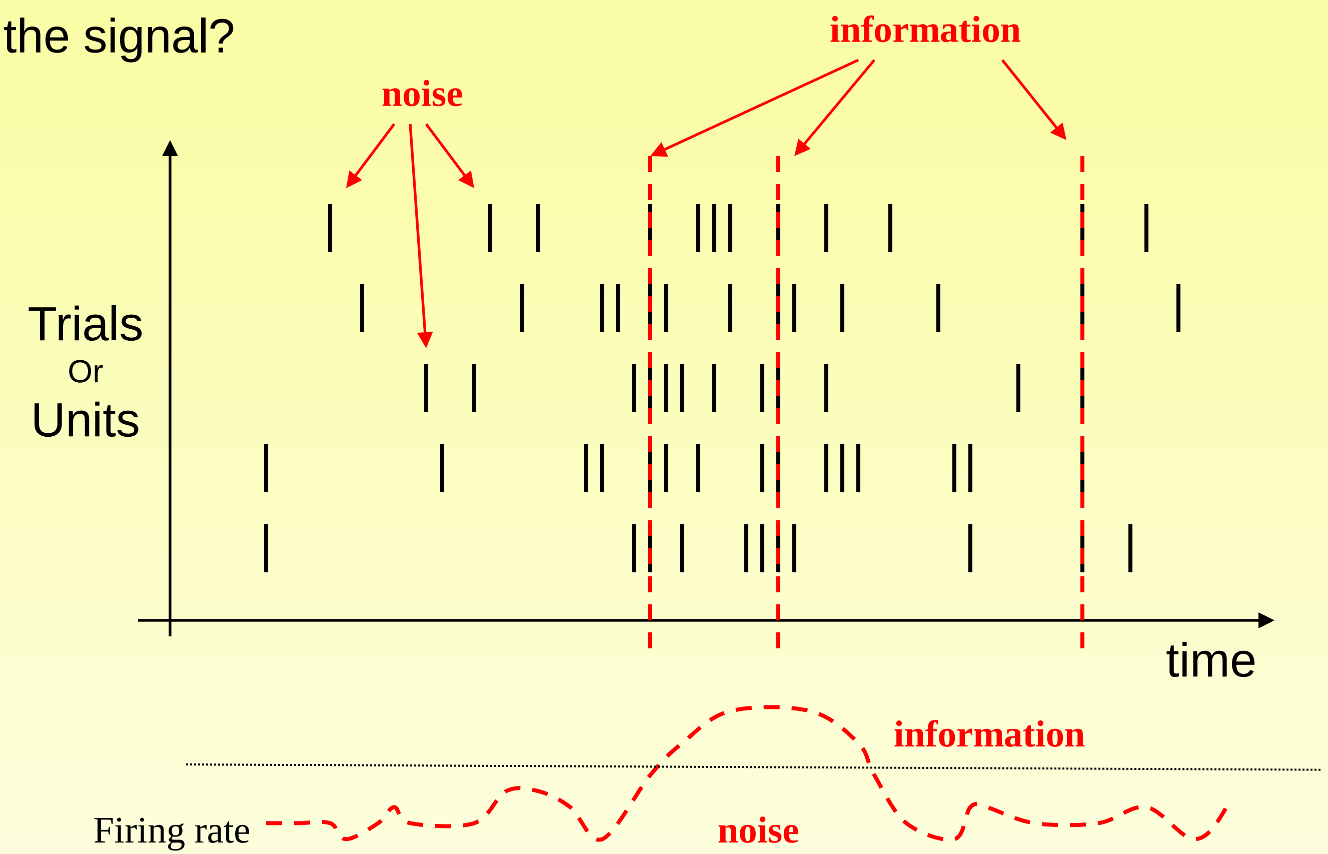
5 - Propose an interpretation (explanation), generate prediction(s):
Use a (conceptual or computational) model.

Good examples:

Reinagel and Reid, J Neuroscience, 2002.
Usrey, Sceniak Chapman, J Neurophys, 2003.

Neural Data Analyses

Where is the signal?



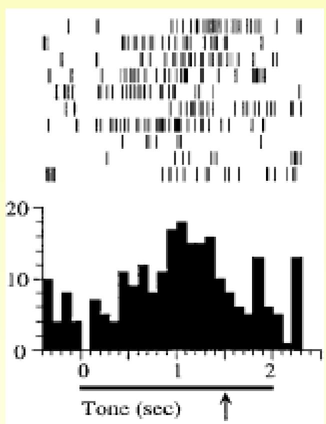
Spike patterns
- Single neurons, single train
- Single neurons, multiple trains
- Multiple neurons single trains
- Multiple neurons, multiple trains

Neural Data Analyses

Single neuron:

- Relative to a stimulus: e.g. peristimulus event histogram, Information.
- Relative to another cell: e.g. cross correlogram.
- Relative to itself: e.g. power spectrum, Inter Spike Interval return map.

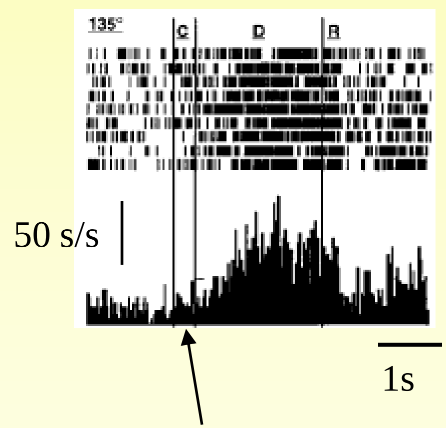
Stimulus ‘driven’



Bin size
(100 ms)

Stimulus

Stimulus ‘triggered’



Bin size
(25 ms)

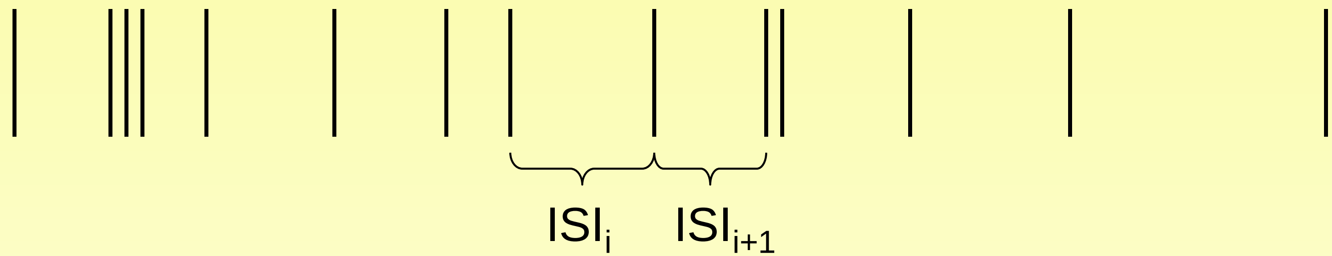
Stimulus

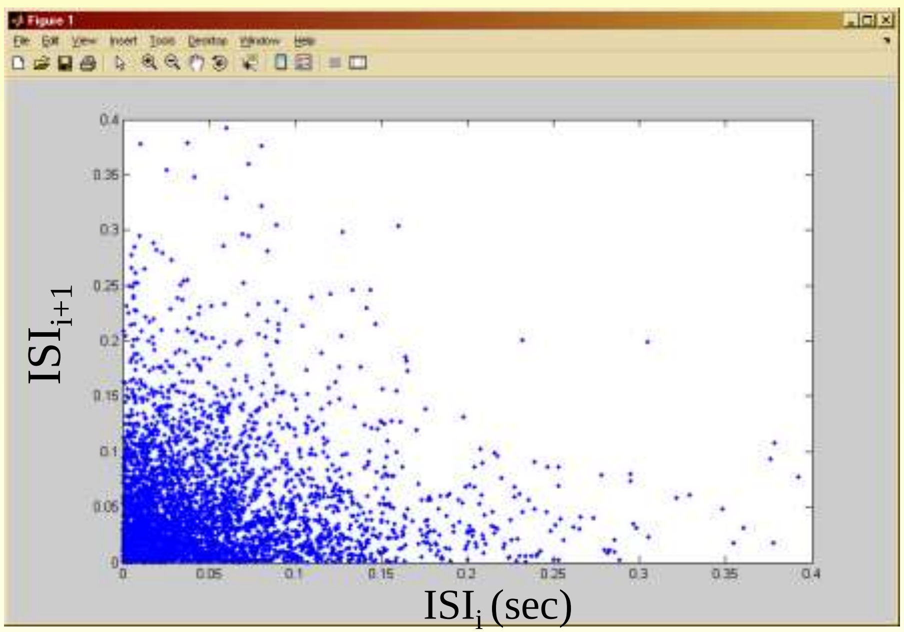
Next:

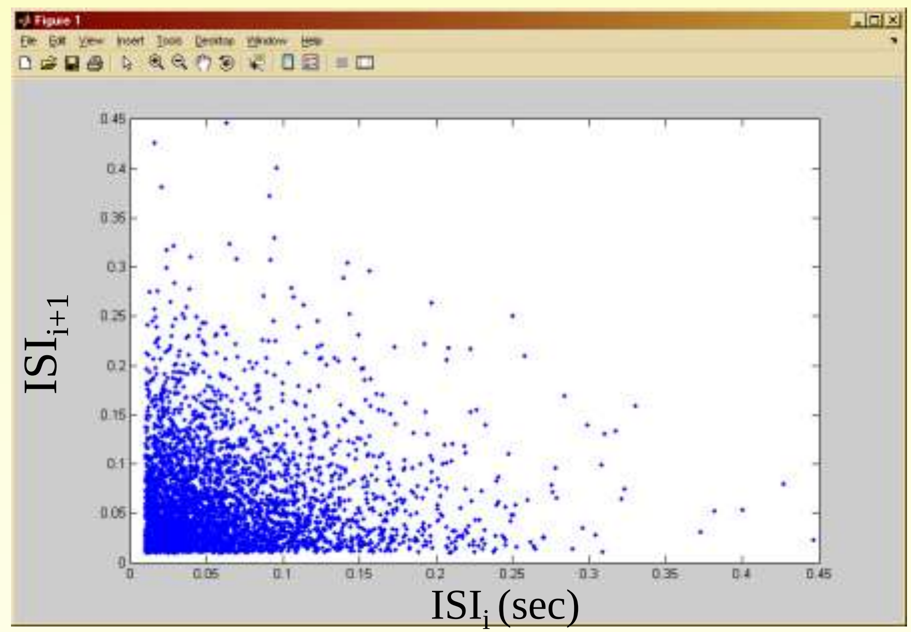
- Gaining intuition from good data *displays* (ISI return maps)
- Some data analyses can give a false ‘picture’ of the results (explained variance)

Inter **S**pike **I**nterval return map

- Goal: Detecting ‘temporal structures’ in spike trains.
- Poincaré map – ISI return map.

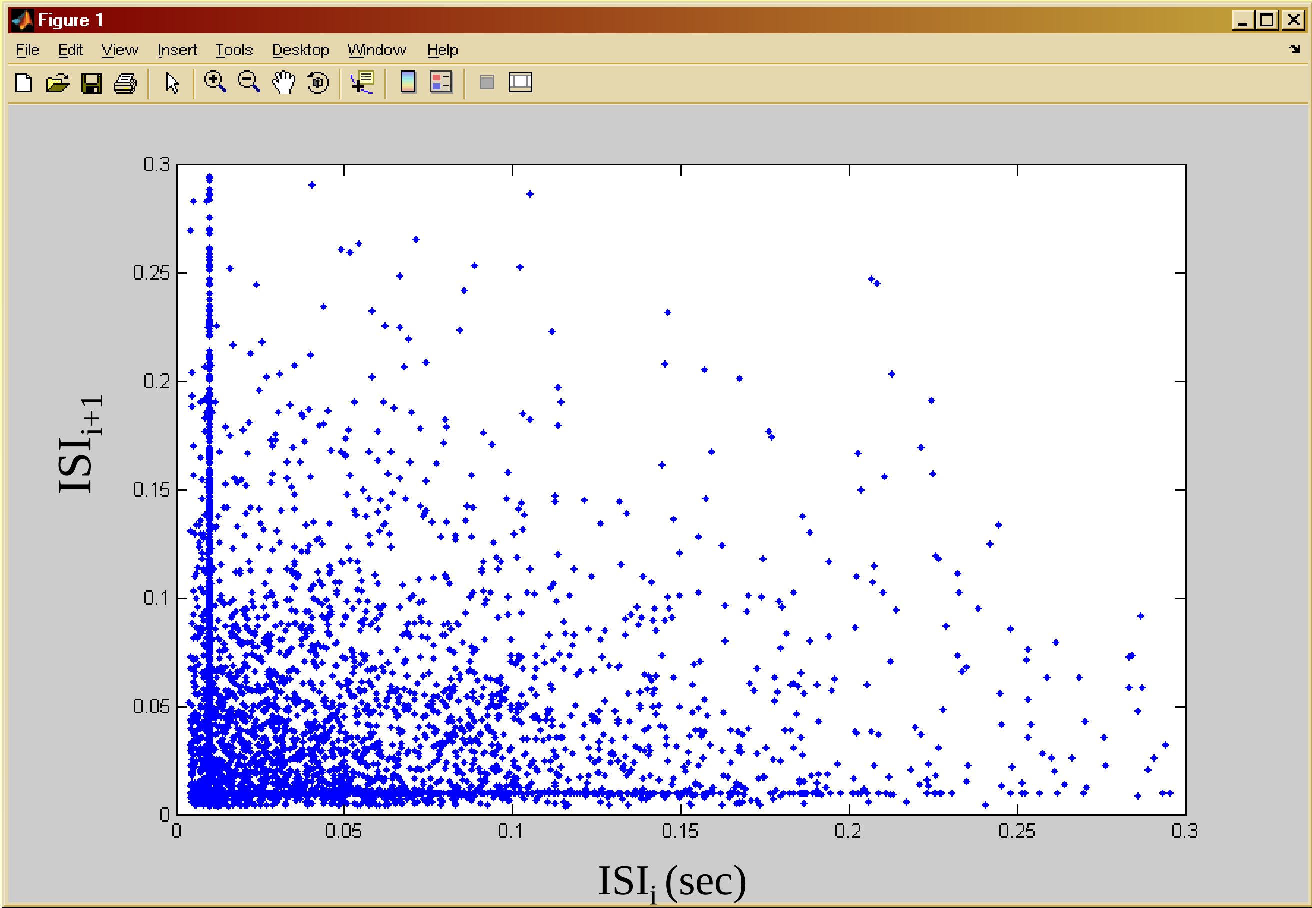


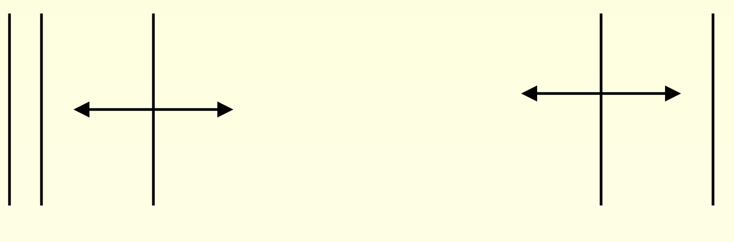




(20 Hz Poisson train)

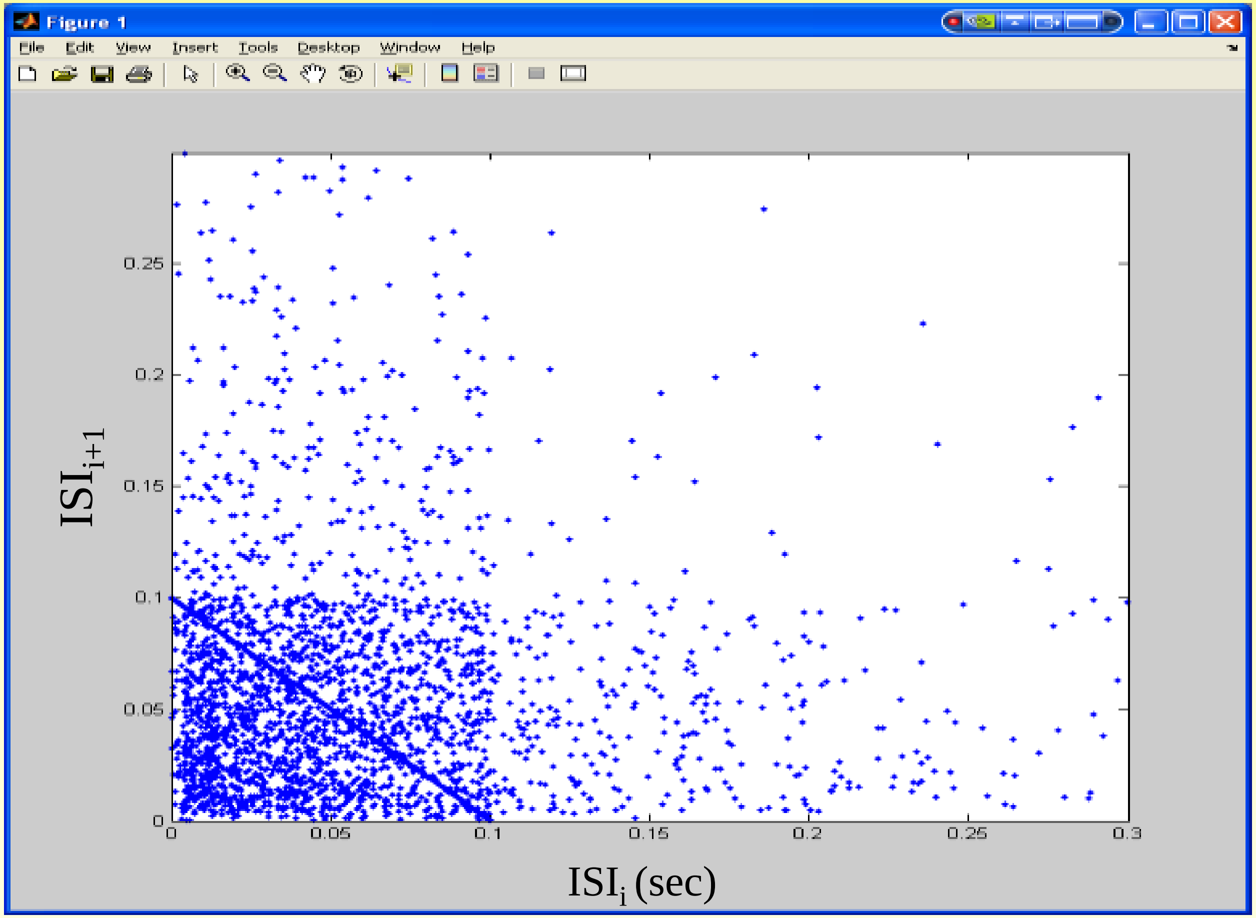
ISI return maps

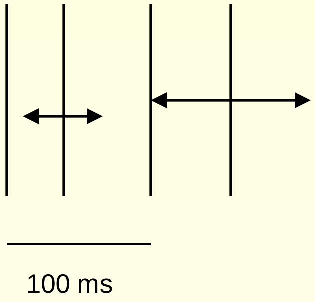




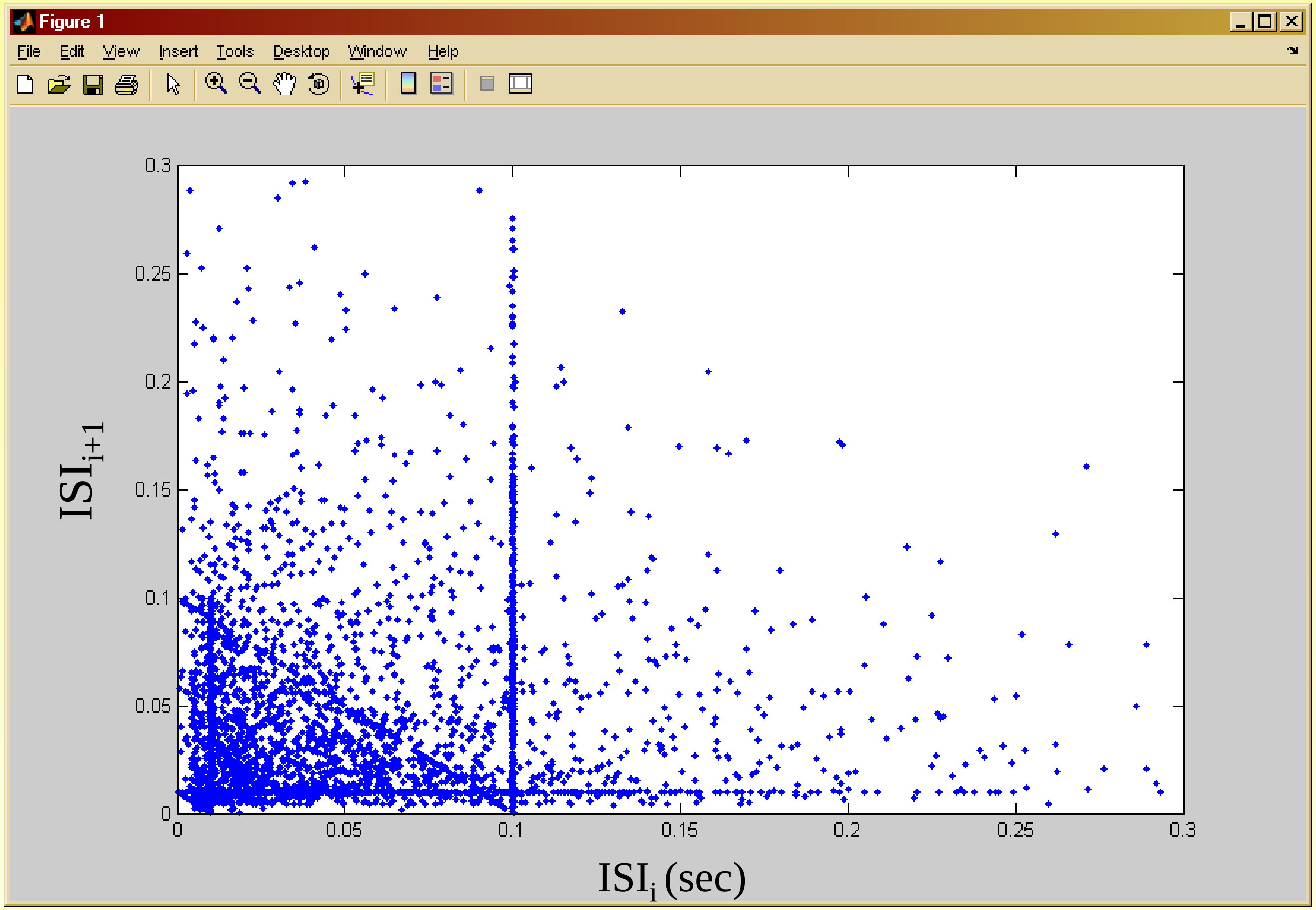
10 ms

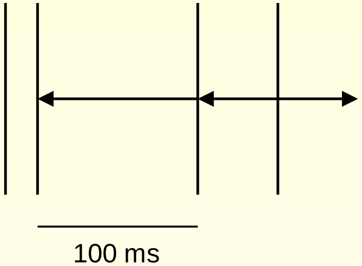
ISI return maps





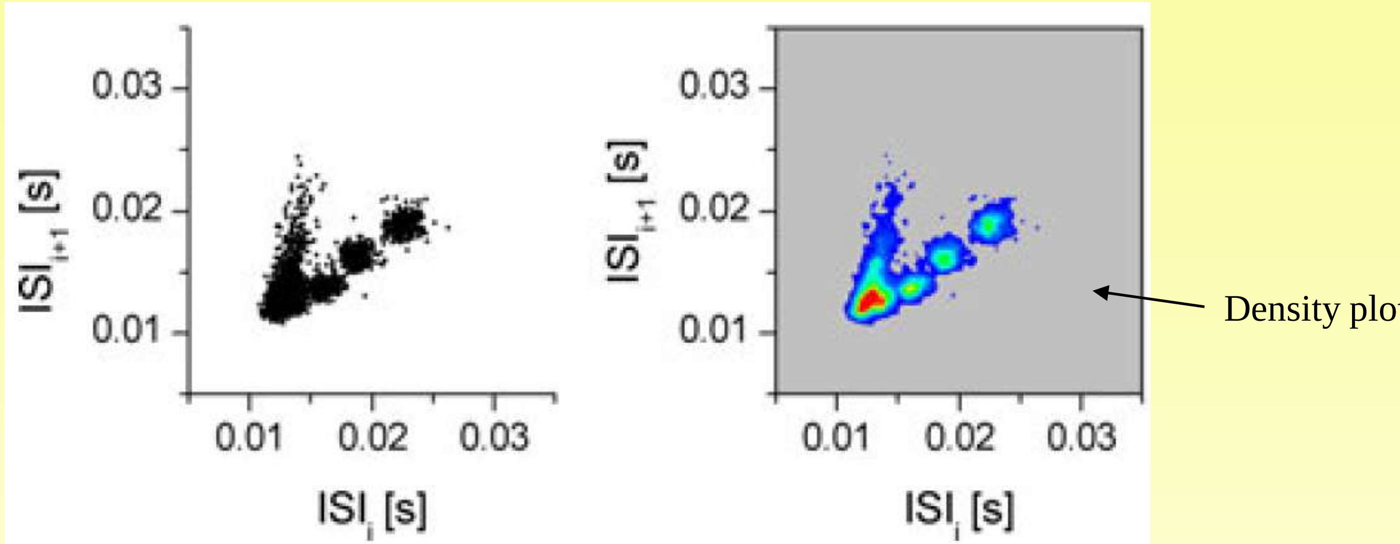
ISI return maps



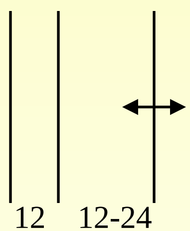


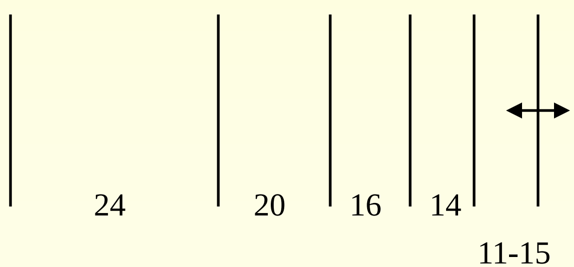
ISI return maps

Pyloric neuron lobster STG

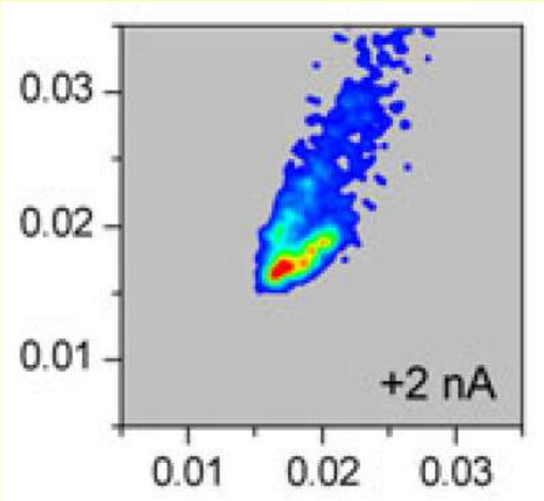


(Szucs et. al. 2005)

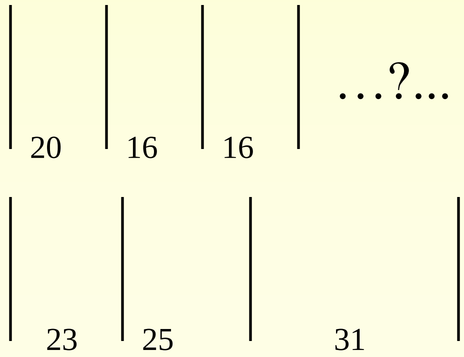




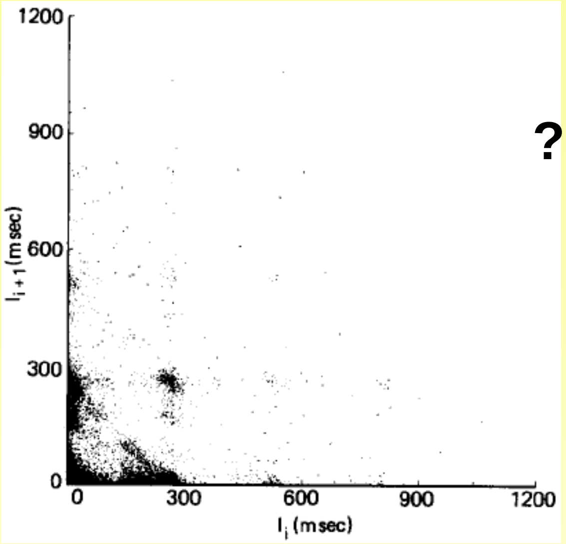
ISI return map

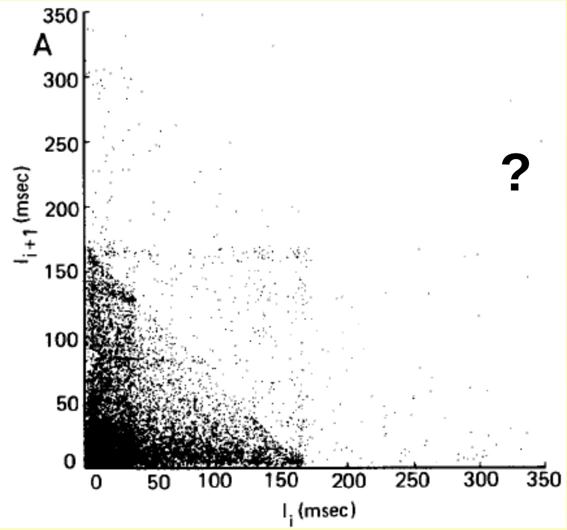


(Szucs et. al. 2005)

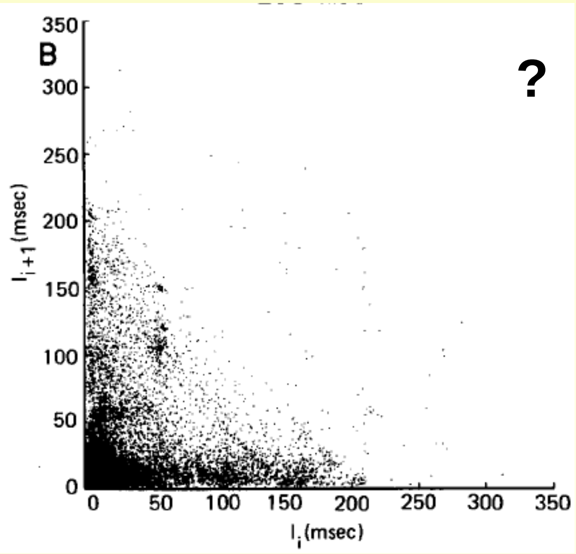


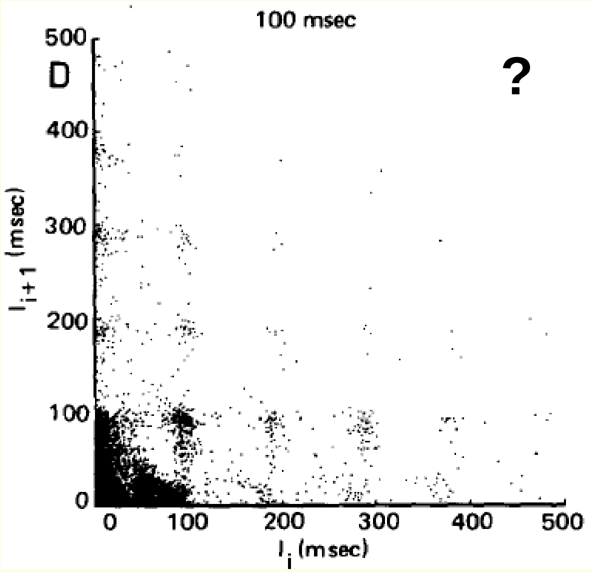
Other real ISI return maps





(Siegel, 1990)





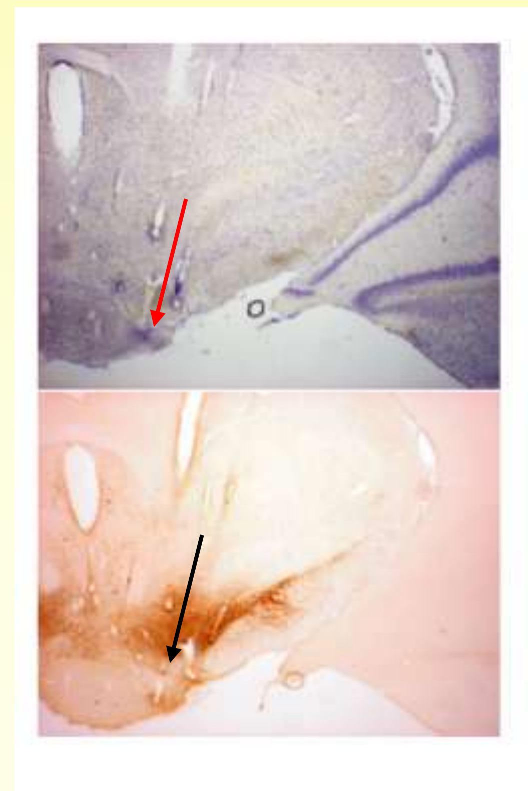
Patterns: Multiple neurons, single trains

Question: Does a rat consolidate memory for specific rewarding events?

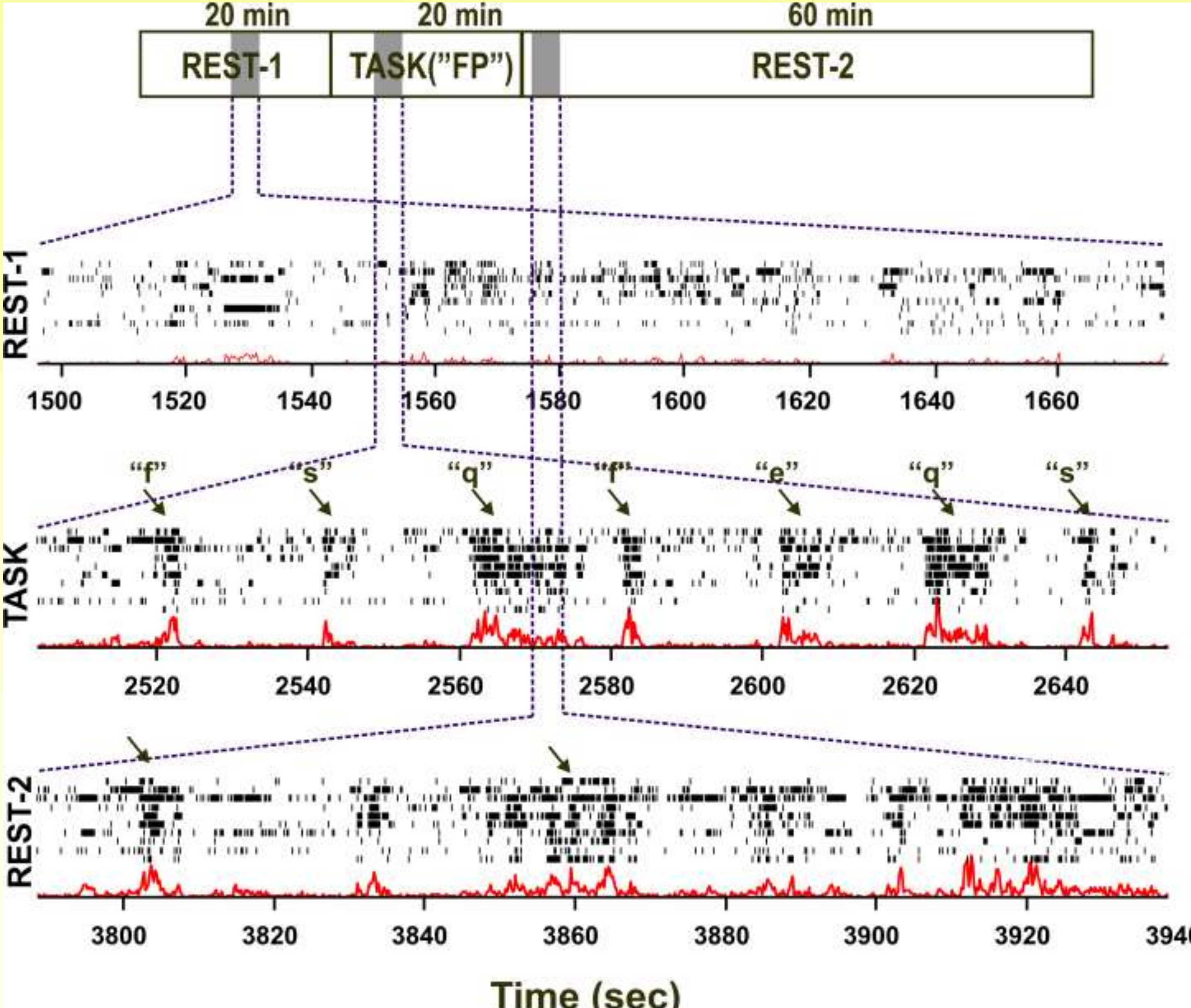
Task: - Non-spatial task (learn probability of reward delivery)
- Different kinds of rewards (regular food, sugar pellets, quinine, ‘empty’)

Technique: Multi-unit chronic tetrode recordings from the Ventral Tegmental Area.

Rest1 – task – **Rest2**



Patterns: Multiple neurons, single trains



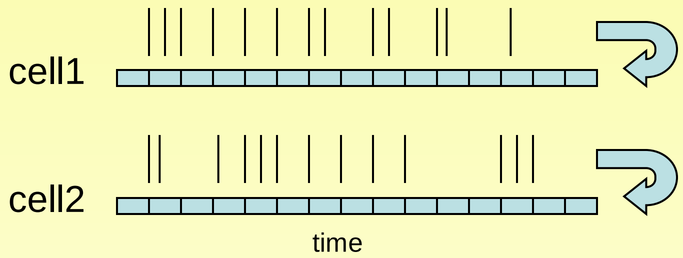
f = food
s = sugar
q = quinine
e = empty

'Reactivation'

‘Classic’ EV method: High Level Description

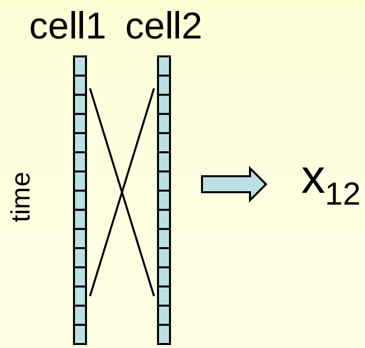
- Select cells that do not belong to the same tetrode: N

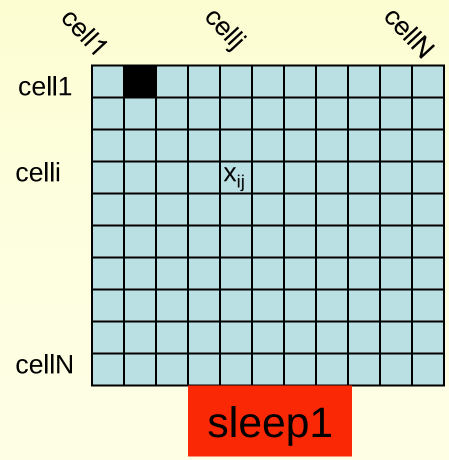
- For epochs task, sleep1 and sleep2 compute the correlation matrix:
    - For each cell, compute the firing histogram (*bins*) in the epoch.

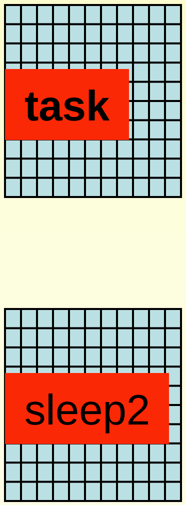


    - Compute the cross correlation matrix (NxN) of the firing histograms.

      (Returns significance level P)



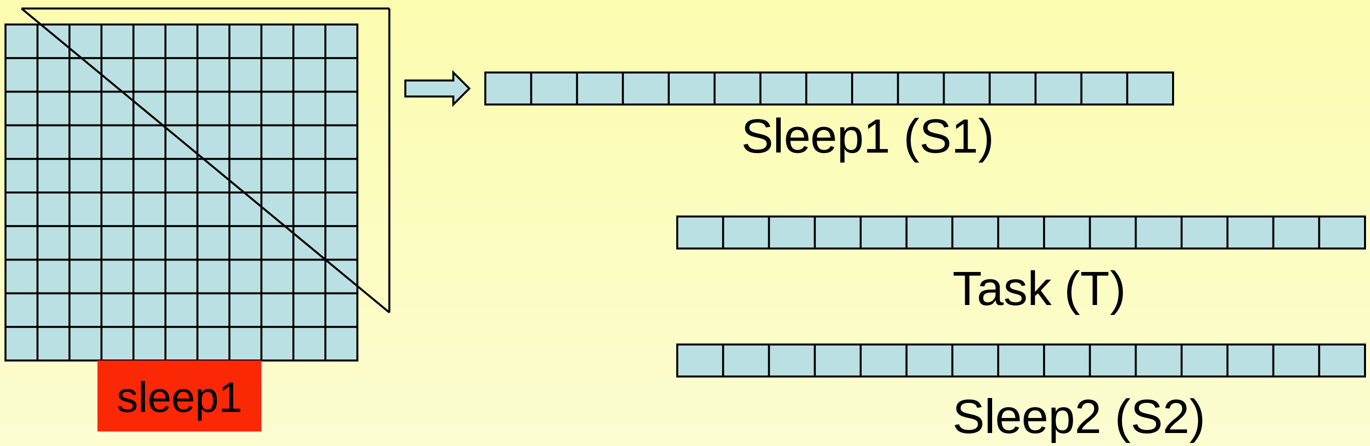




- (Find the correlation coefficients that are significant (p<0.1) across all three epochs)

‘Classic’ EV method: High Level Description

- Select the upper triangle values (excluding diagonal) of all three matrices and vectorize (S1, T, S2)



- Compute EV/REV (‘Explained Variance’):

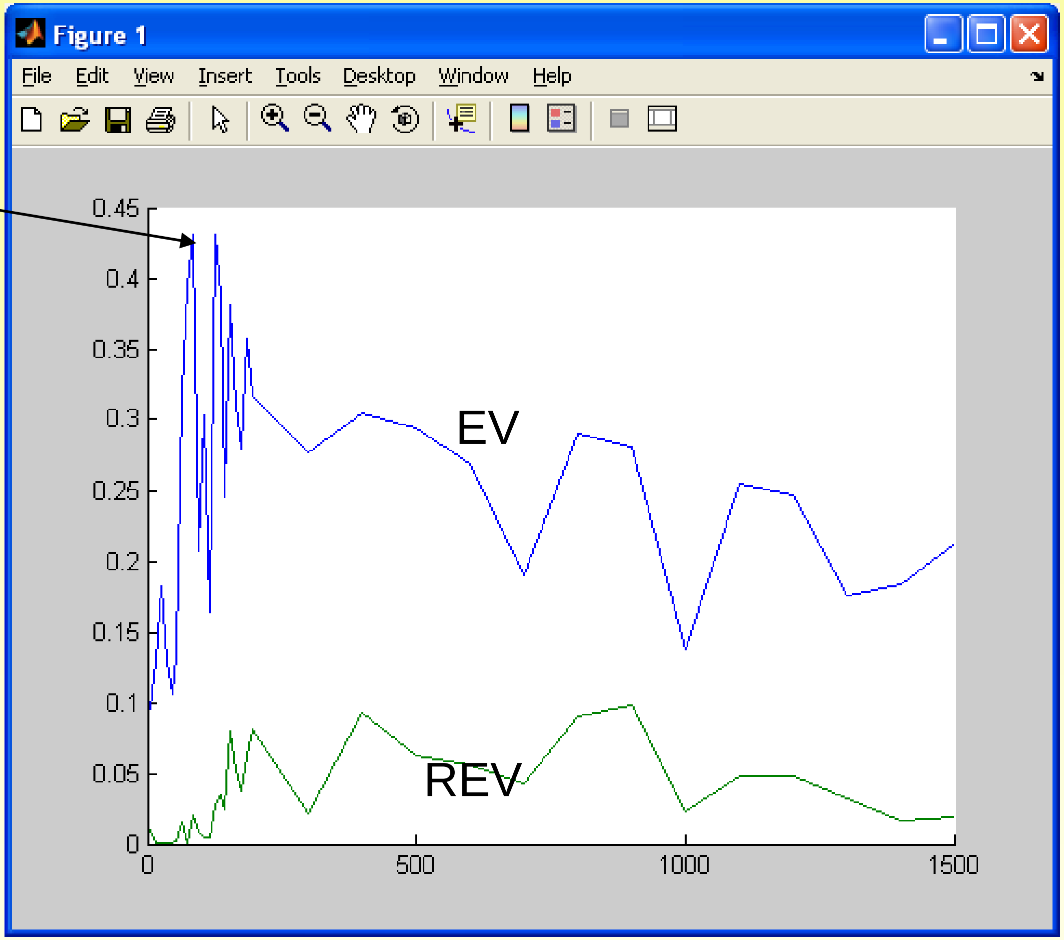
    - Compute the crosscorrelations: c_{1T}=xcorr(S1,T), c_{2T}=xcorr(S2,T) and c_{12}=xcorr(S1,S2)

$$EV = \left[\frac{c_{2T} - c_{1T}c_{12}}{\sqrt{\left(1-c_{1T}^2\right)\left(1-c_{12}^2\right)}} \right]^2 \qquad REV = \left[\frac{c_{1T} - c_{2T}c_{12}}{\sqrt{\left(1-c_{2T}^2\right)\left(1-c_{12}^2\right)}} \right]^2$$

=> Only one parameter: Bin size (typically 100ms)

‘Classic’ EV method: The bin size problem

EV at 100 ms: **0.44**



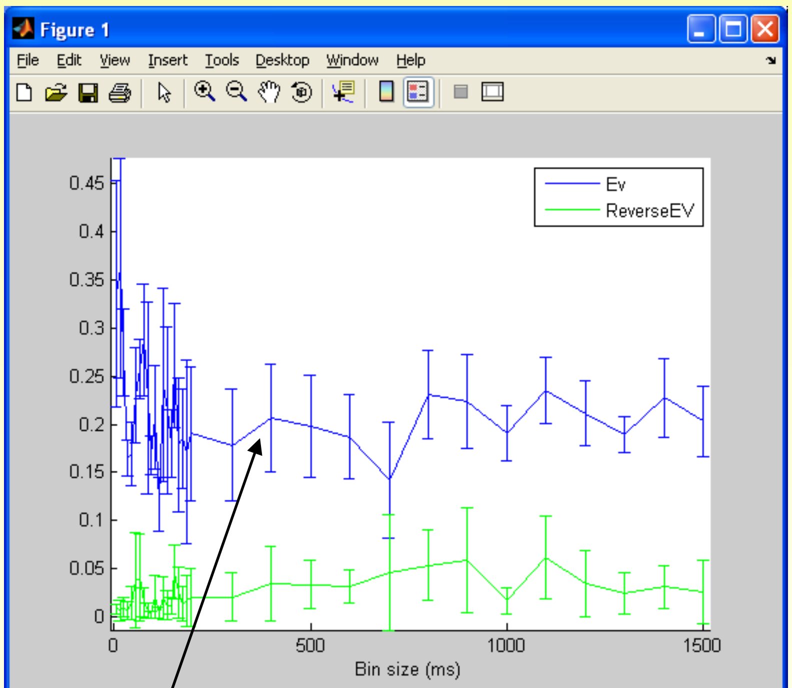
Bin size (ms)

Correcting for ‘noise’ – Classic ‘EV’

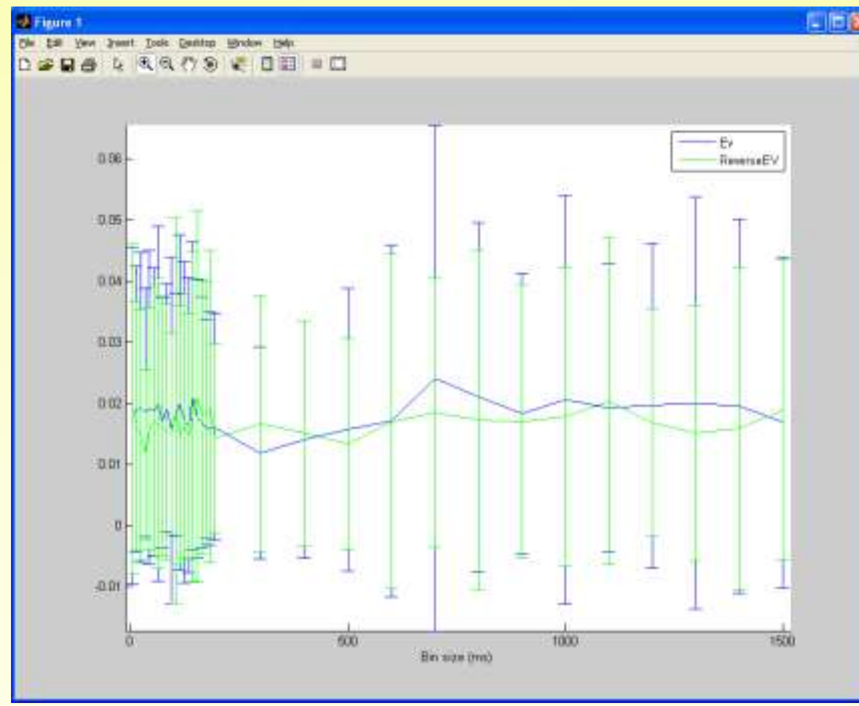
How to obtain a more reliable estimate of reactivation?

…Rest1-Rest1-Rest1-Rest1 – task – **Rest2**

Multiple sleep1 epochs



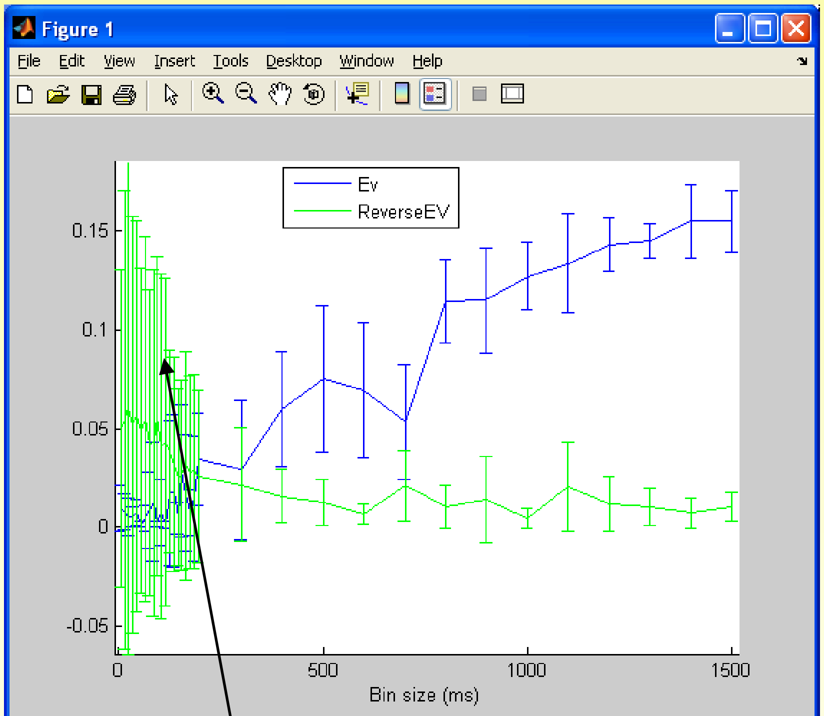
ISI Shuffled – 50 trials

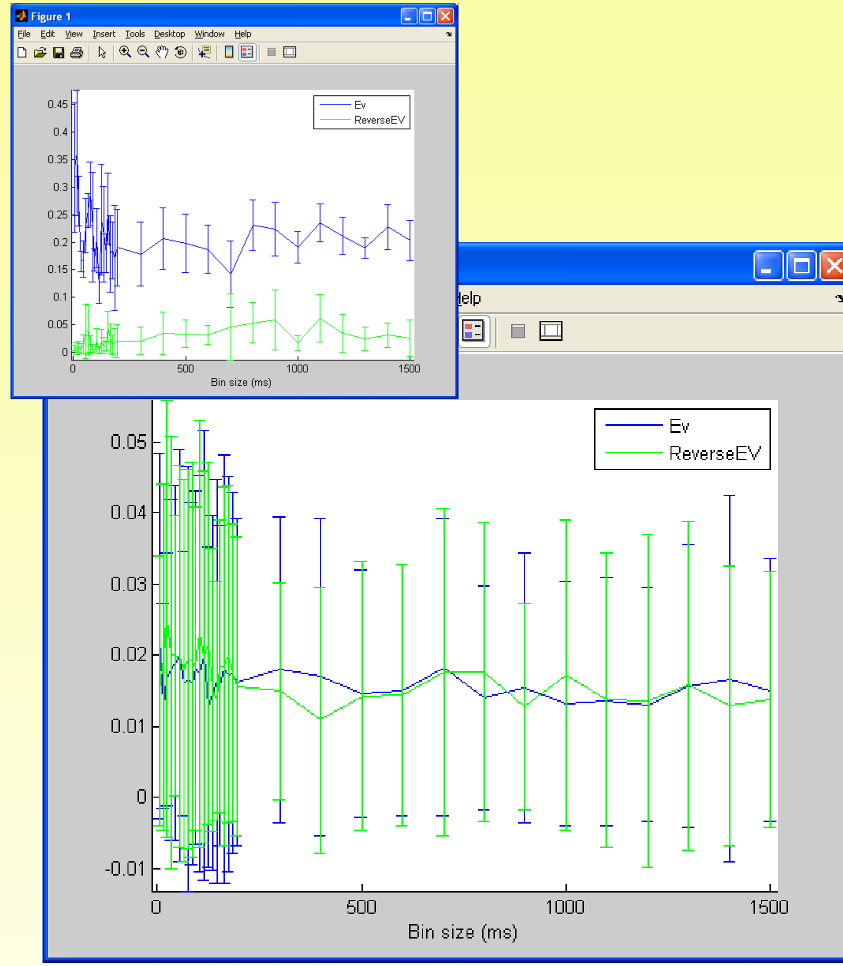


Apparent EV: ~**0.2**

Establishing the conditions in which reactivation occurs

'Classic' EV – during non REM sleep



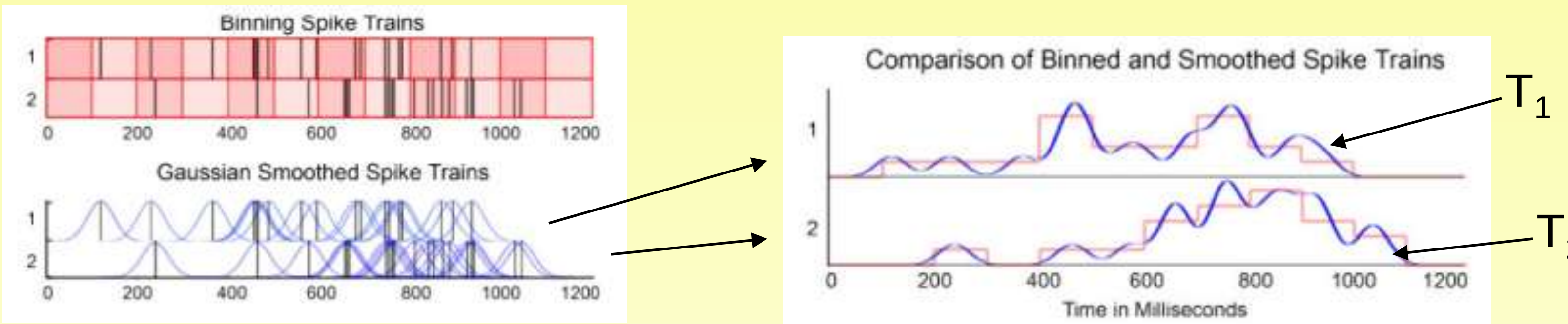


At 100 ms bin → No reactivation outside of REM sleep

Can we do these analyses with no binning?

‘Smooth’ EV method: High Level Description

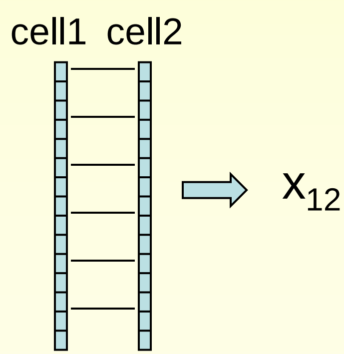
- Select cells that do not belong to the same tetrode: N
- For task, sleep1 and sleep2 compute the *similarity* matrix:
    - For each cell, convolve the spike train with gaussian (width sigma)

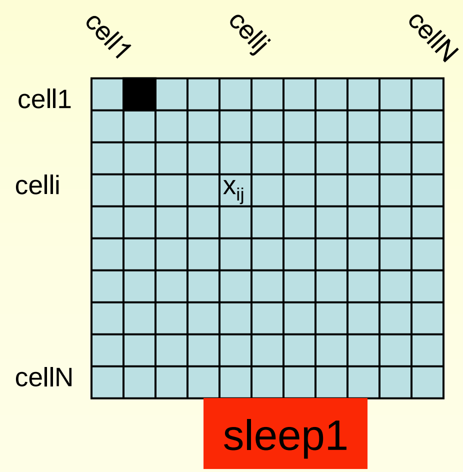


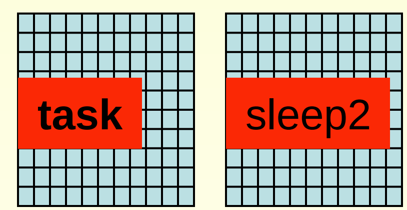
(Fellous et al 2004, Kruskal et. al. 2007)

- Compute the pairwise dot product between convolved spike trains

$$S_{ij} = \frac{\vec{T}_i.\vec{T}_j}{\left\|\vec{T}_i\right\|\left\|\vec{T}_j\right\|} = \cos(\vec{T}_i, \vec{T}_j)$$

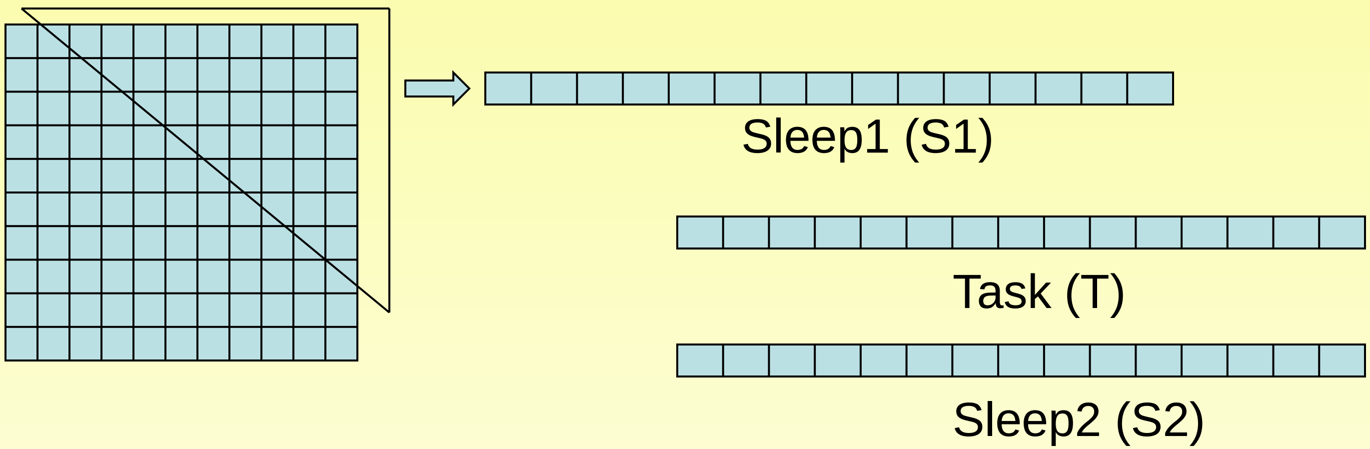






‘Smooth’ EV method: High Level Description

- Select the upper triangle values (excluding diagonal) of all three matrices and vectorize (S1, T, S2)



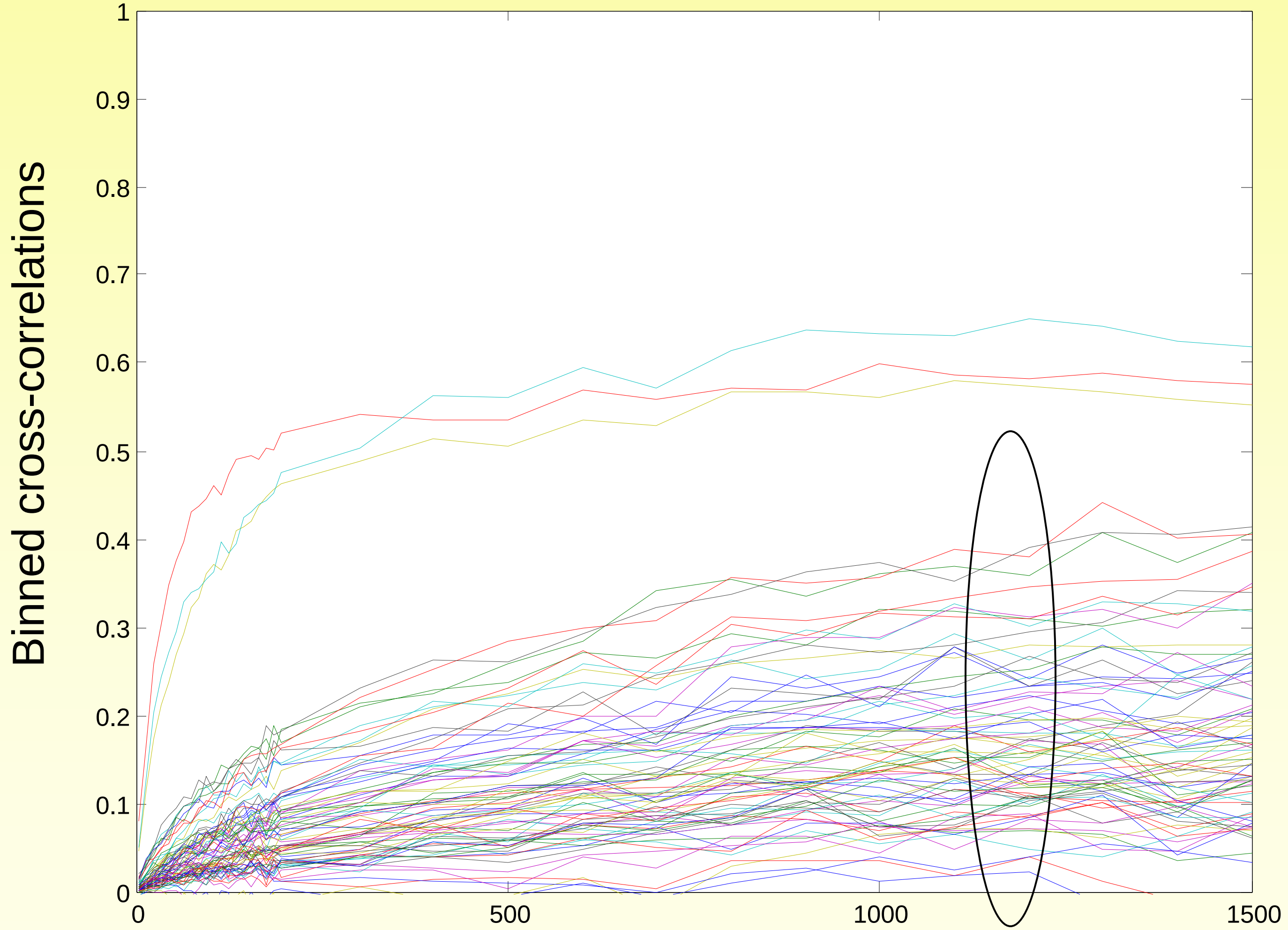
- Compute EV/REV (‘Smoothed Explained Variance’):

    - Compute the Xcorrelations: c_{1T}=xcorr(S1,T), c_{2T}=xcorr(S2,T) and c_{12}=xcorr(S1,S2)

$$EV = \left[\frac{c_{2T} - c_{1T} c_{12}}{\sqrt{\left(1 - c_{1T}^2\right)\left(1 - c_{12}^2\right)}} \right]^2 \qquad REV = \left[\frac{c_{1T} - c_{2T} c_{12}}{\sqrt{\left(1 - c_{2T}^2\right)\left(1 - c_{12}^2\right)}} \right]^2$$

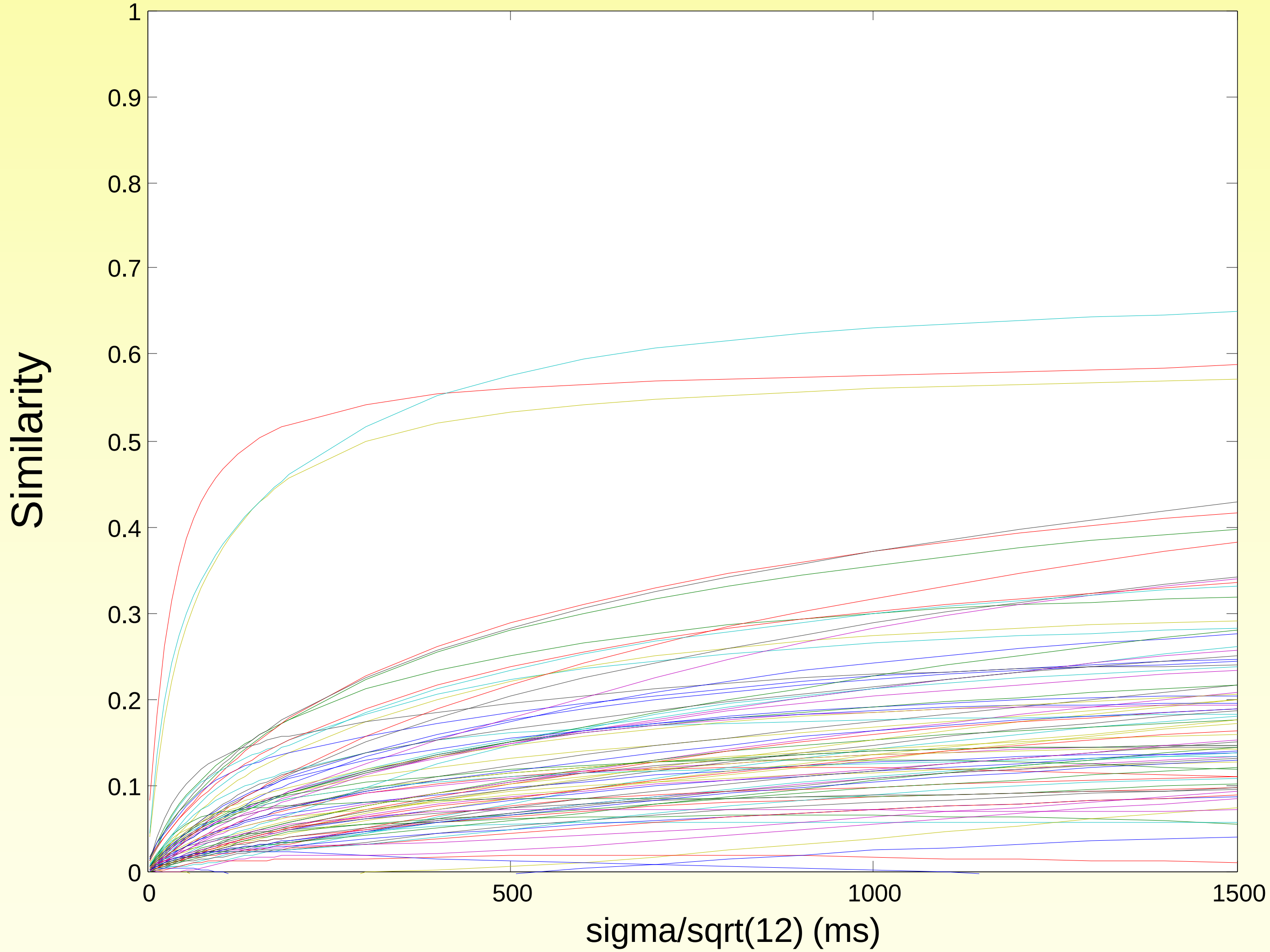
Pairwise binned histogram correlations

Cross correlations as a function of bin size:

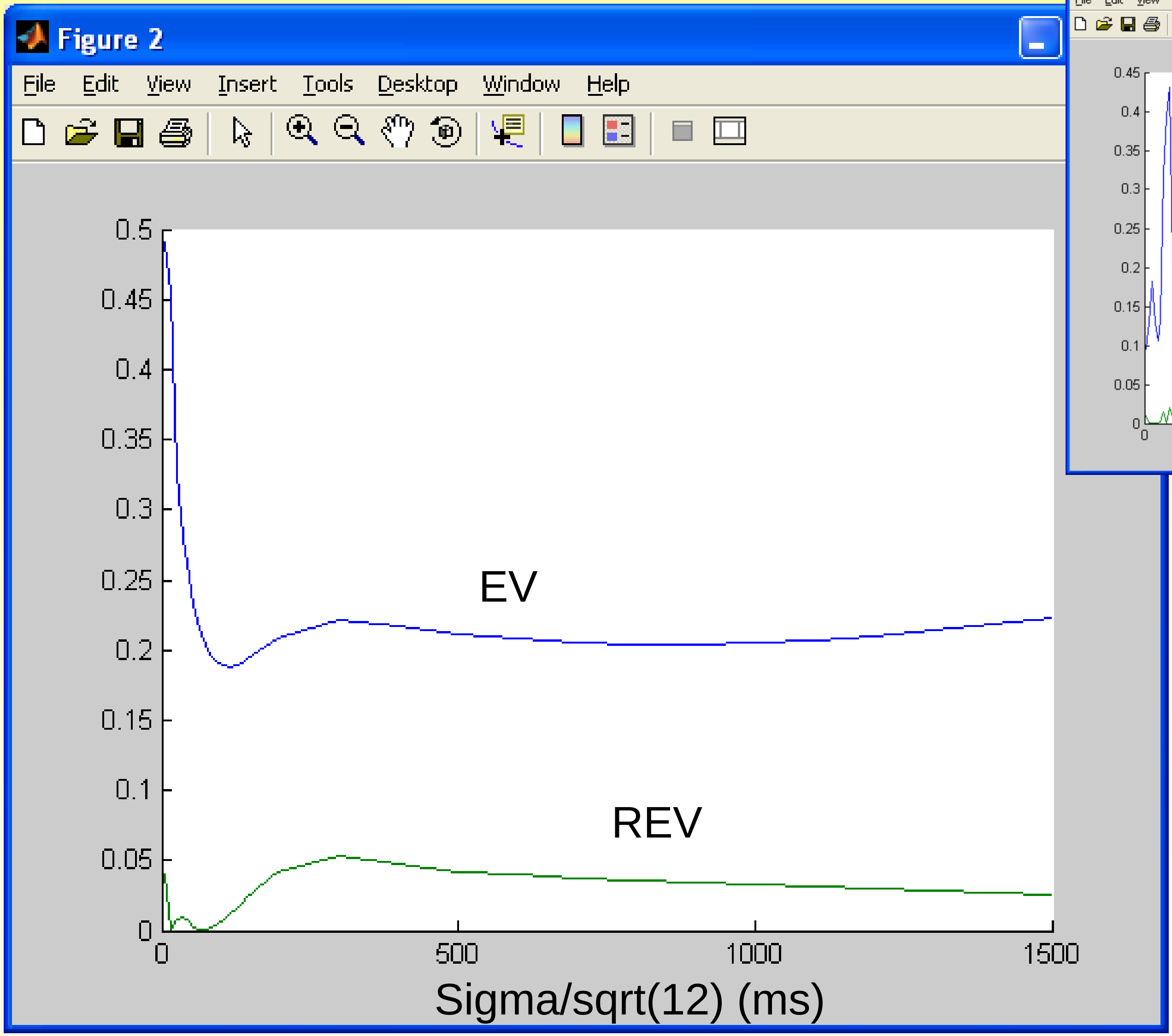


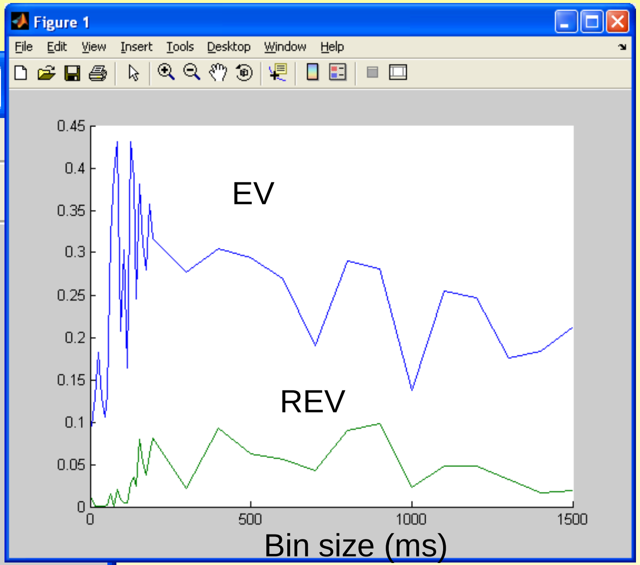
Pairwise similarity measure

Similarity as a function of ‘bin size’:

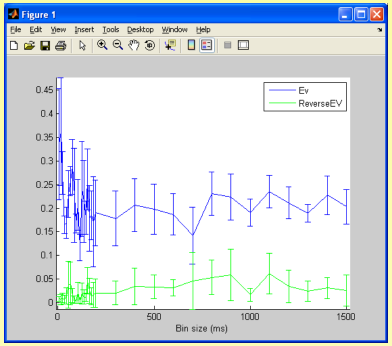


'Smooth' EV measure



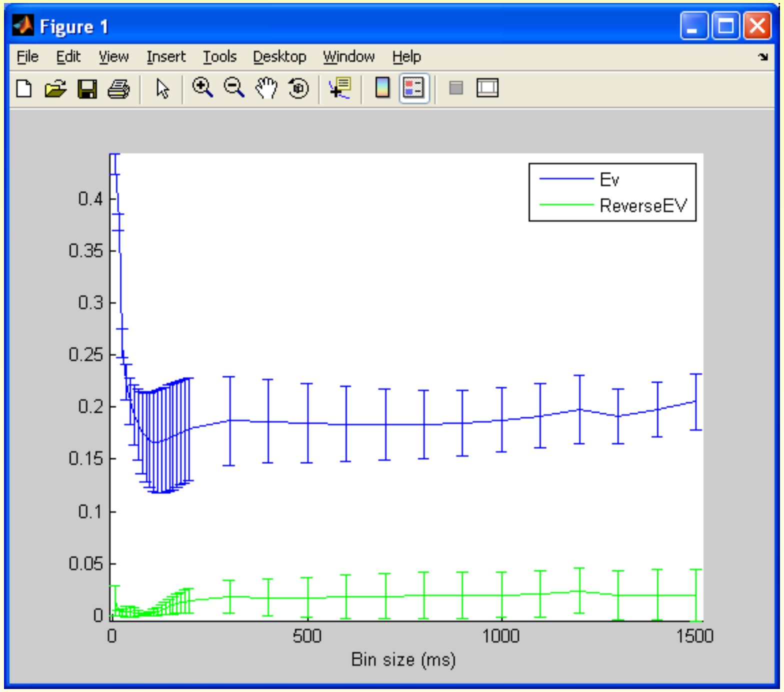


Correcting for noise: ‘Smooth’ EV

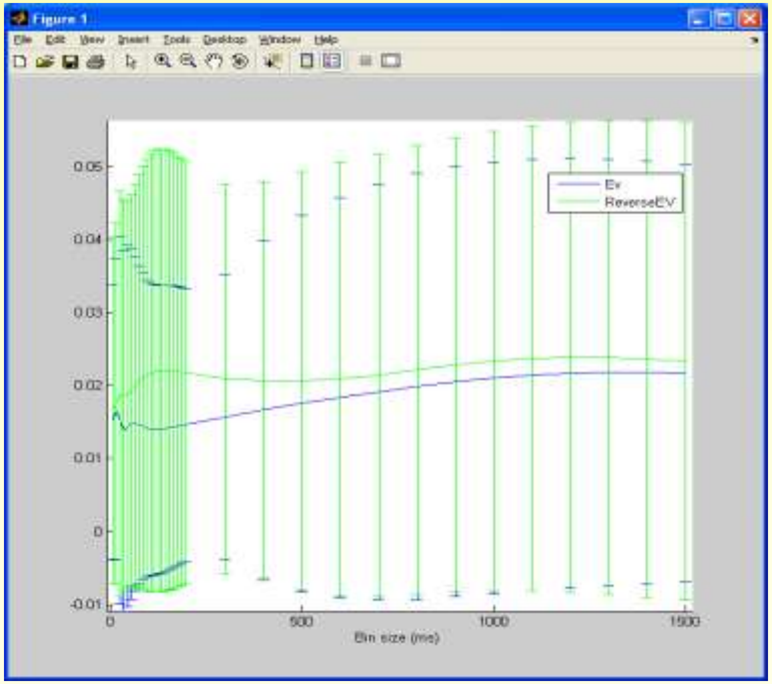


…Rest1-Rest1-Rest1-Rest1 – task – **Rest2**

Multiple sleep1 epochs



ISI Shuffled – 50 trials

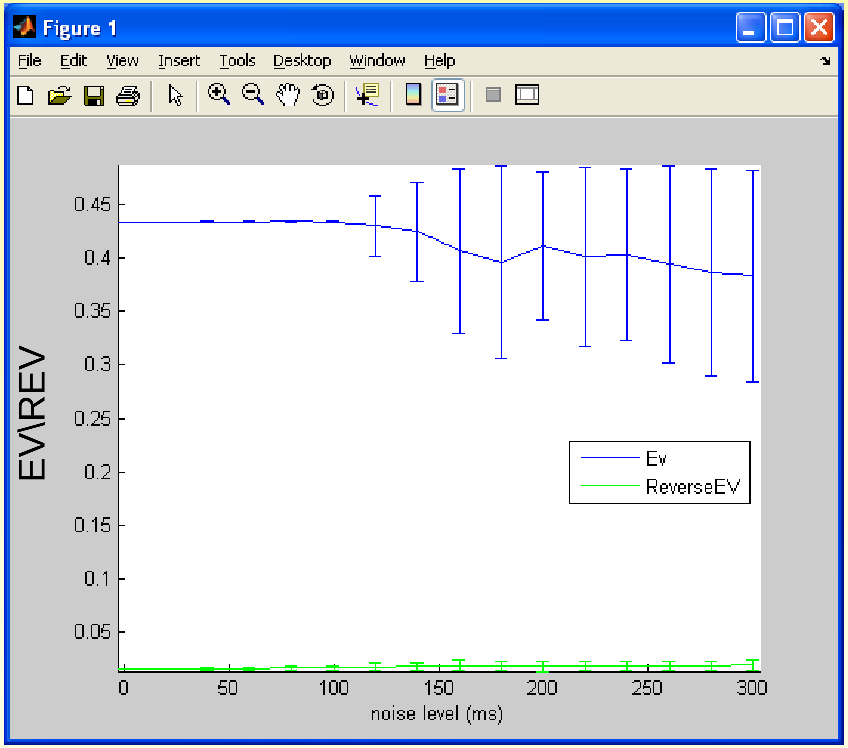


Robust to noise?

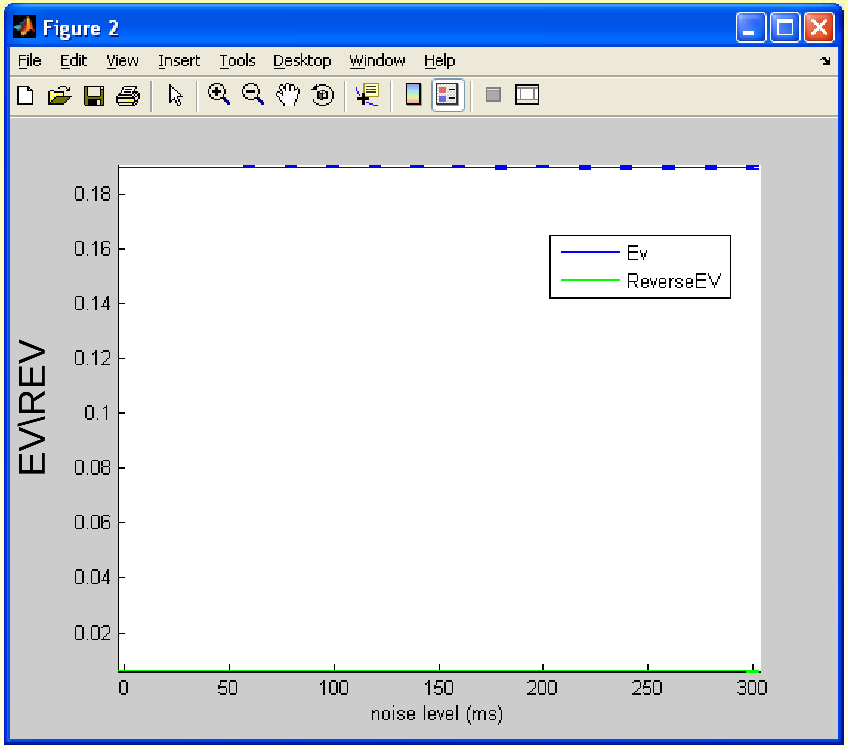
Adding noise to the data

100 ms window

'Classic' EV

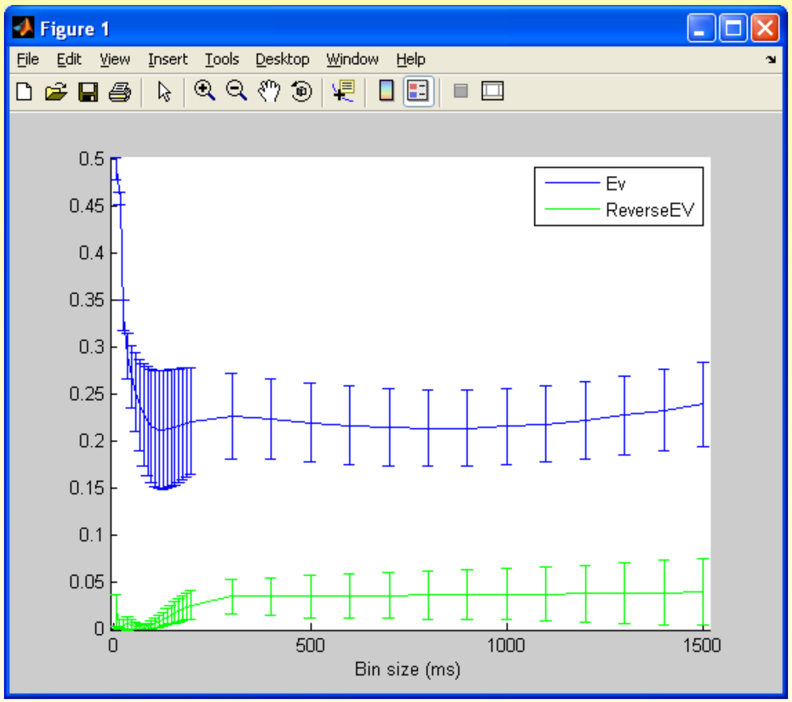


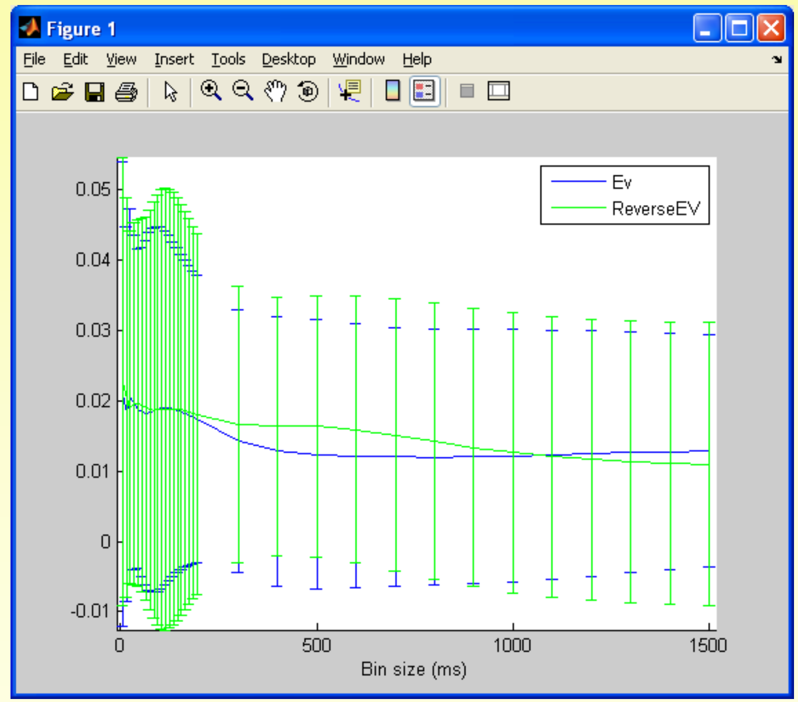
'Smooth' EV



‘Smooth’ EV measure

The issue of REM sleep





Clear reactivation outside of REM sleep

Conclusions and open problems

For a proper EV analysis:

- ‘Classic’ EV: use control experiments, use control Sleep1 epochs (i.e. record enough sleep1).

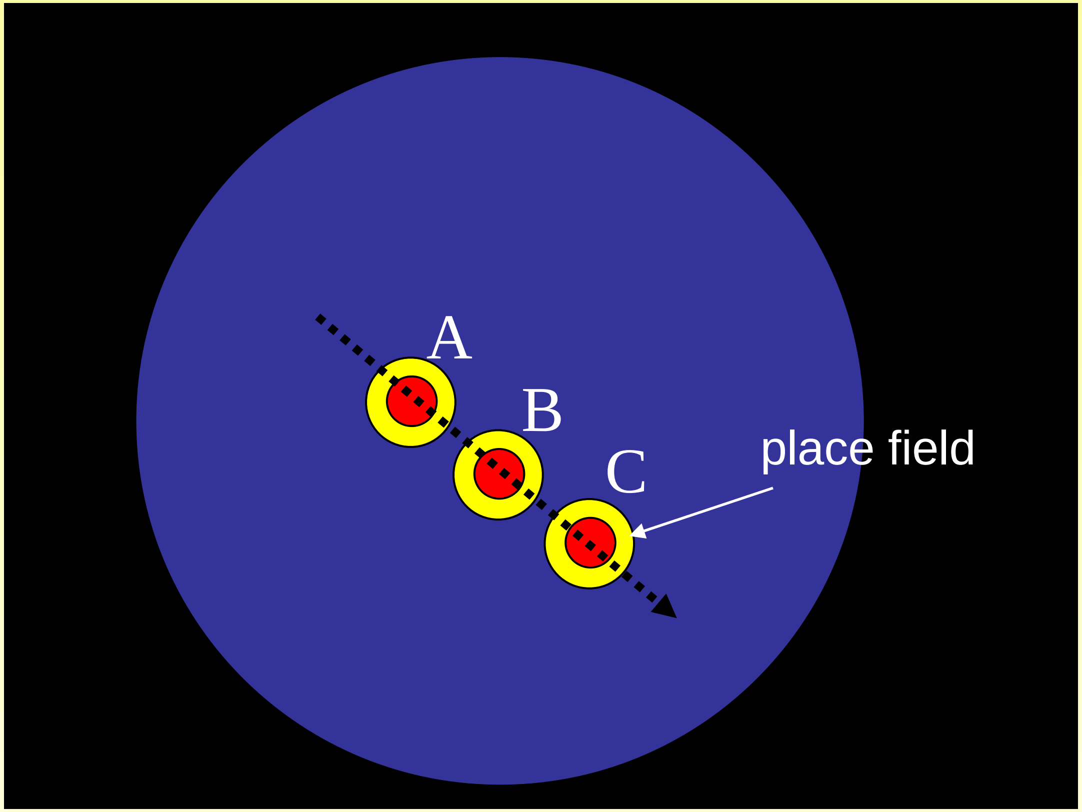
- Carefully assess the ‘time scale’ of reactivation: (bin size/sigma, length and position of sleep epochs).

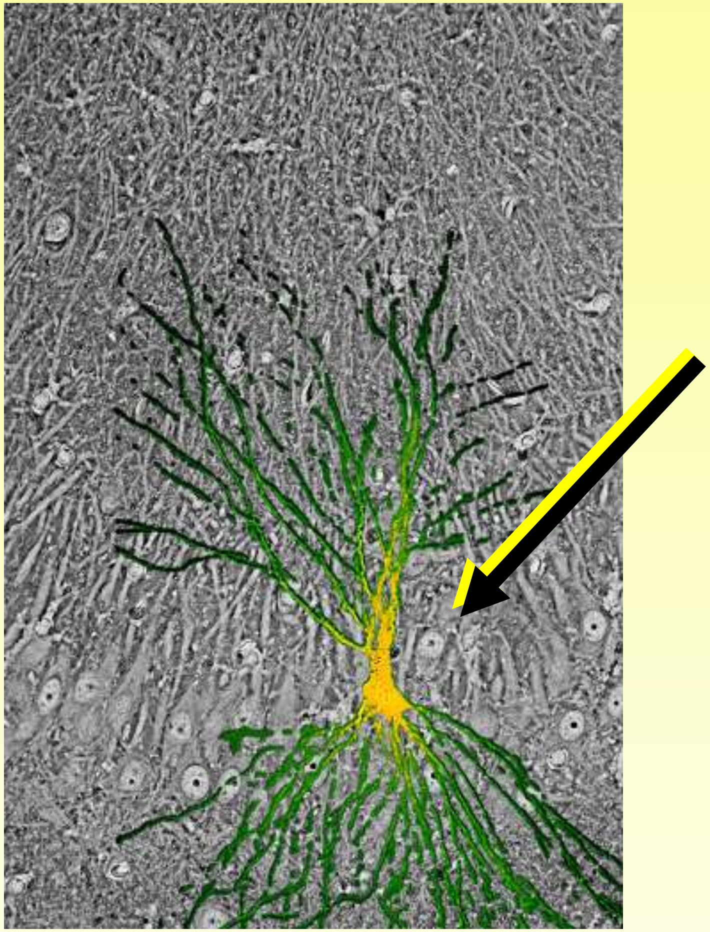
- Always compare EV and ReverseEV (significance of an EV value)

- Use enough cells (EV Vs number of cells in the population ?)

- Ev is a statistical measure: Need ways of
    - Detecting specific patterns
    - Determining when they occur
    - Determining their intrinsic time scale (compression/dilation)
    - Determining (dynamically) their cell memberships

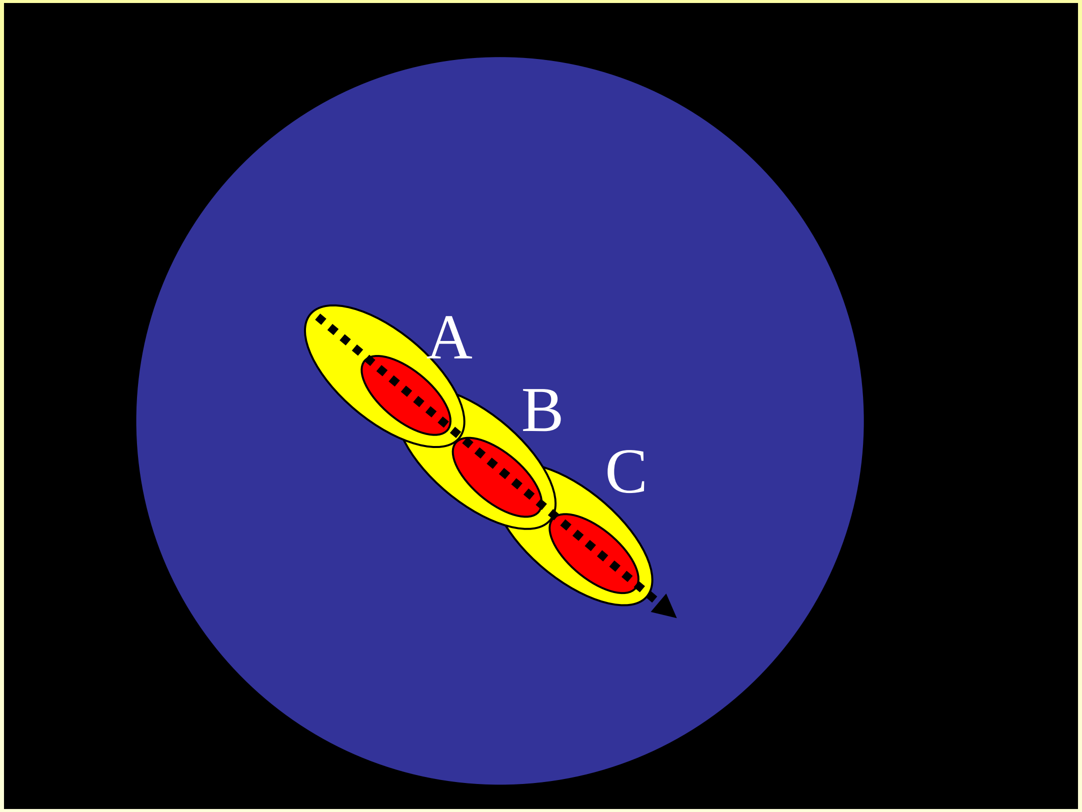
Reactivation in the hippocampus

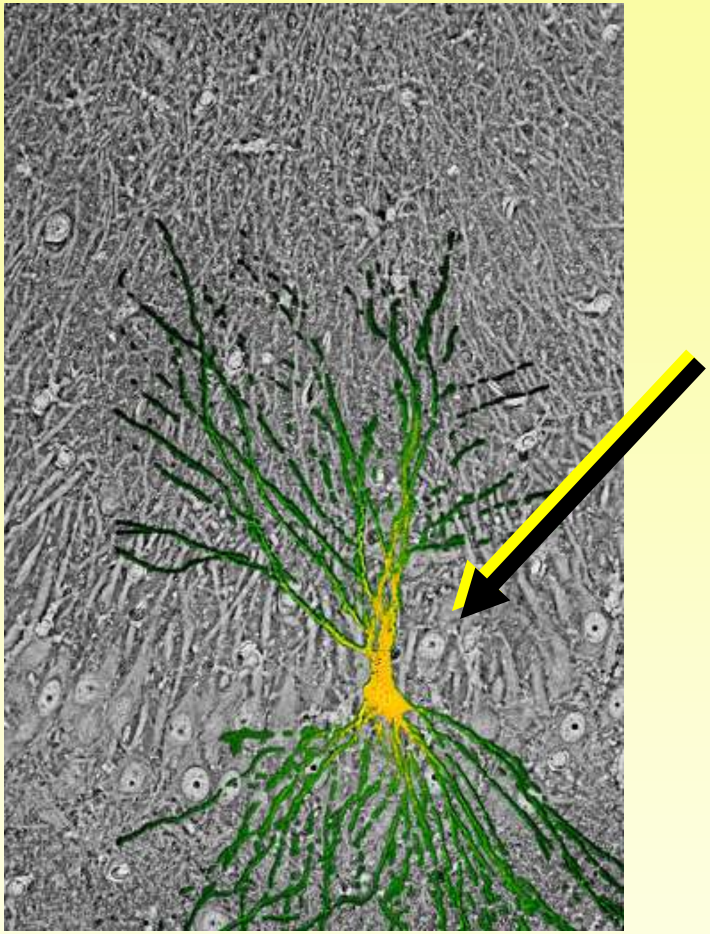




If a rat repeatedly follows the same route through the environment……

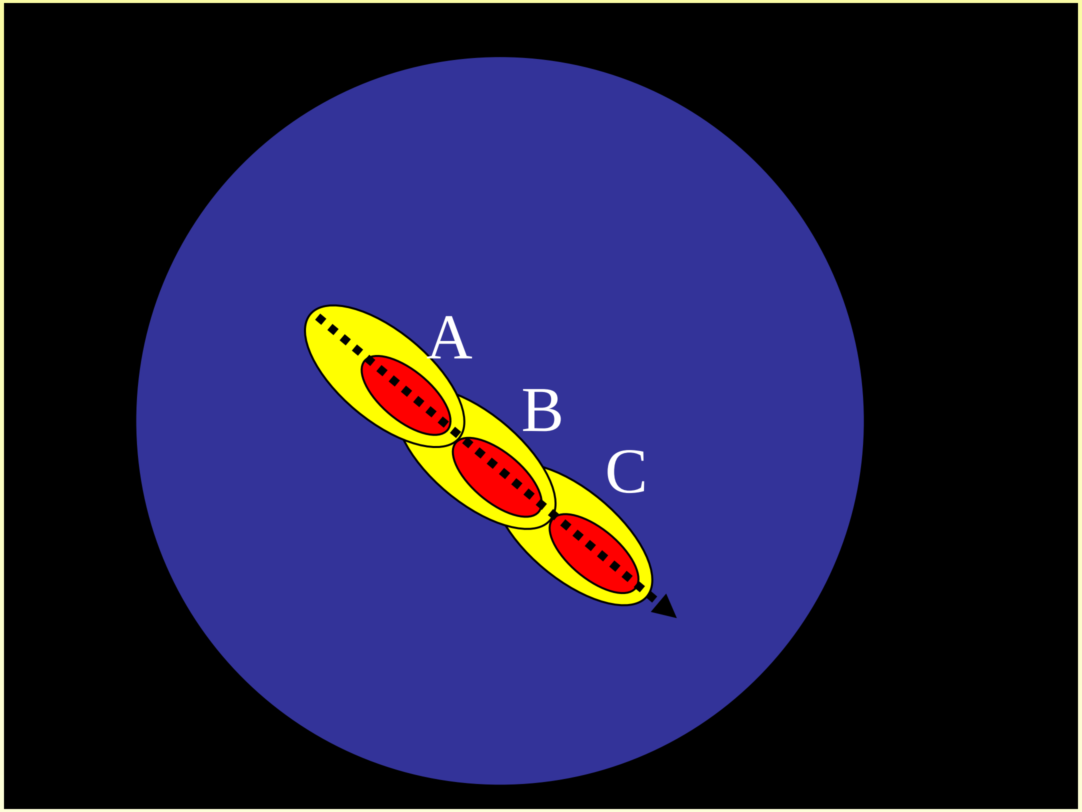
Reactivation in the hippocampus





…the place fields should expand in the direction opposite to the movement direction, creating correlations between cells.

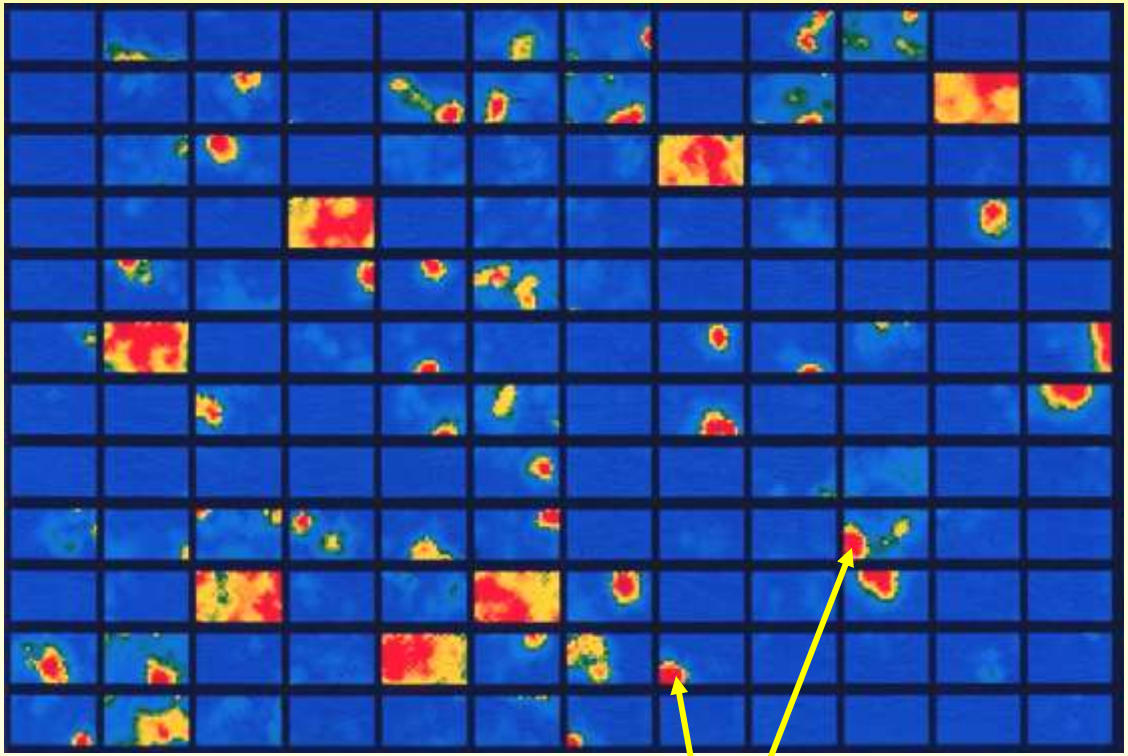
Reactivation in the hippocampus



During sleep after a period of repeated traversals, young rats spontaneously reactivate the neural activity along the route. This is part of the memory consolidation process.

Reactivation in the hippocampus

The overall pattern of activity in an environment can be characterized by the matrix of all correlations among the recorded neurons



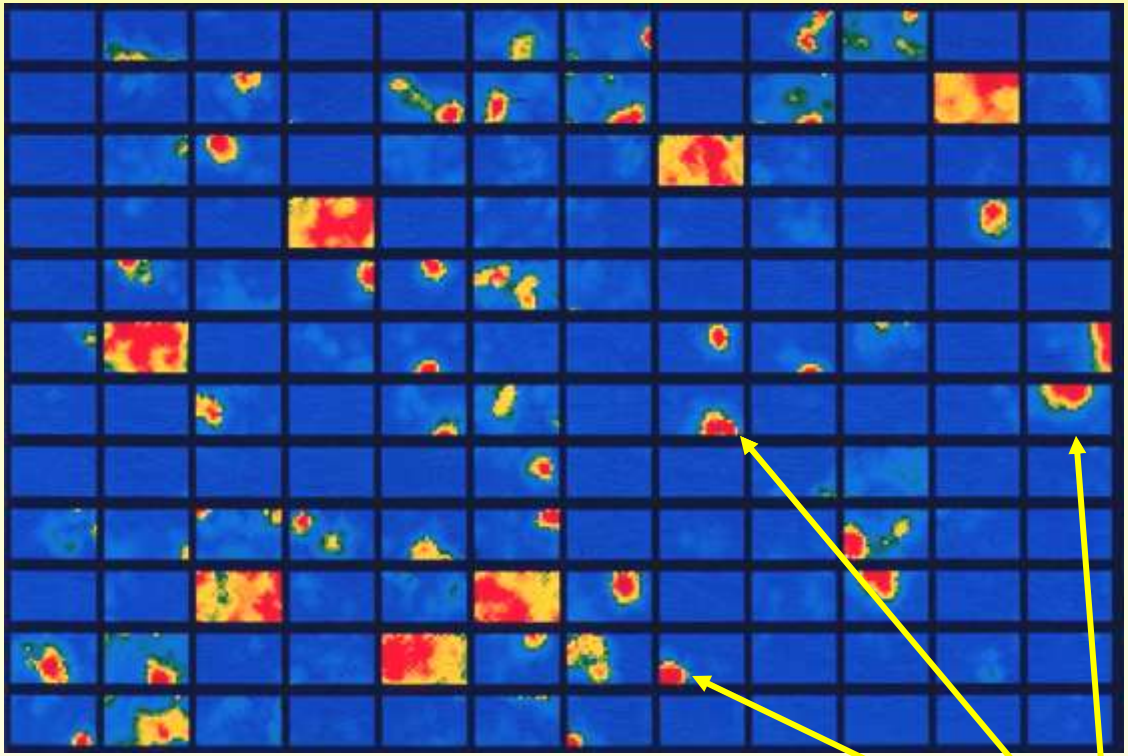
144 cells recorded simultaneously

Correlated cells

Data from Wilson & McNaughton, 1993

Reactivation in the hippocampus

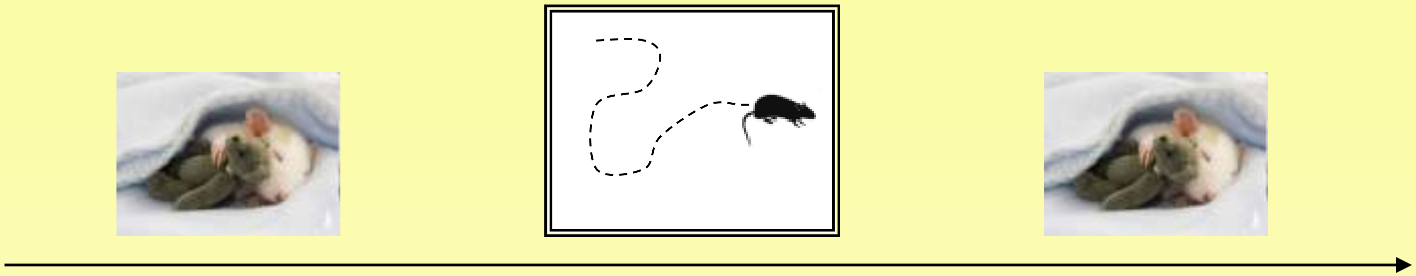
The overall pattern of activity in an environment can be characterized by the matrix of all correlations among the recorded neurons

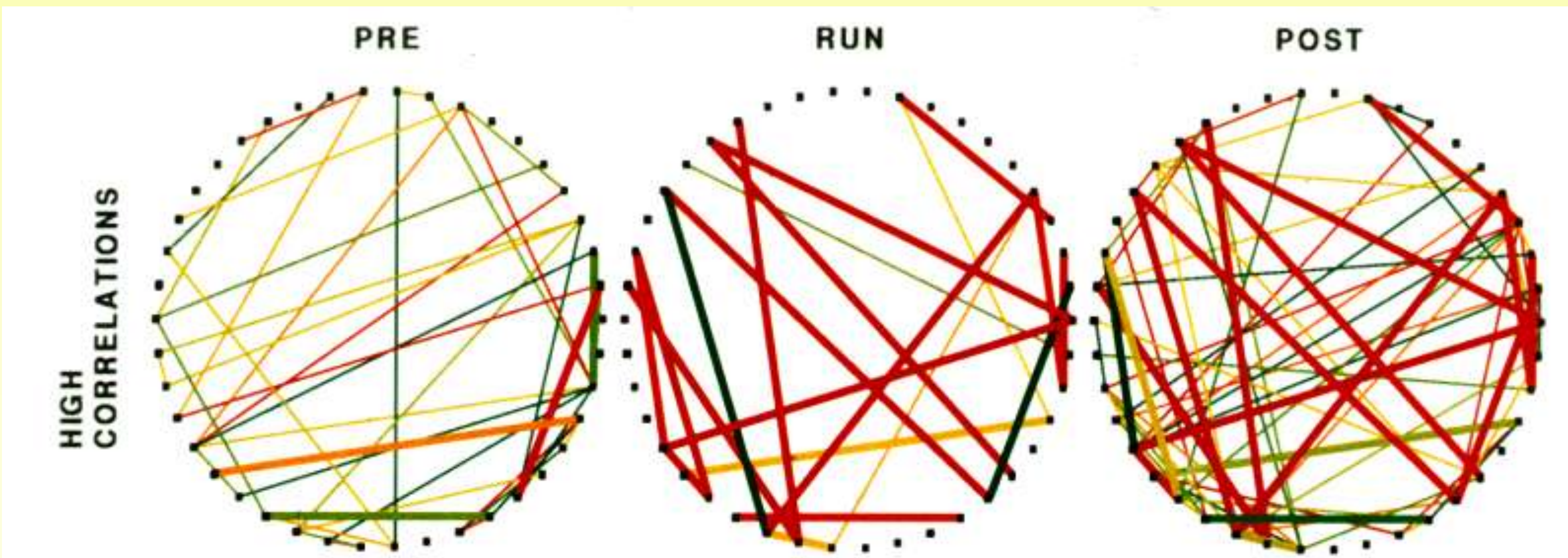


Uncorrelated cells

Data from Wilson & McNaughton, 1993

Reactivation

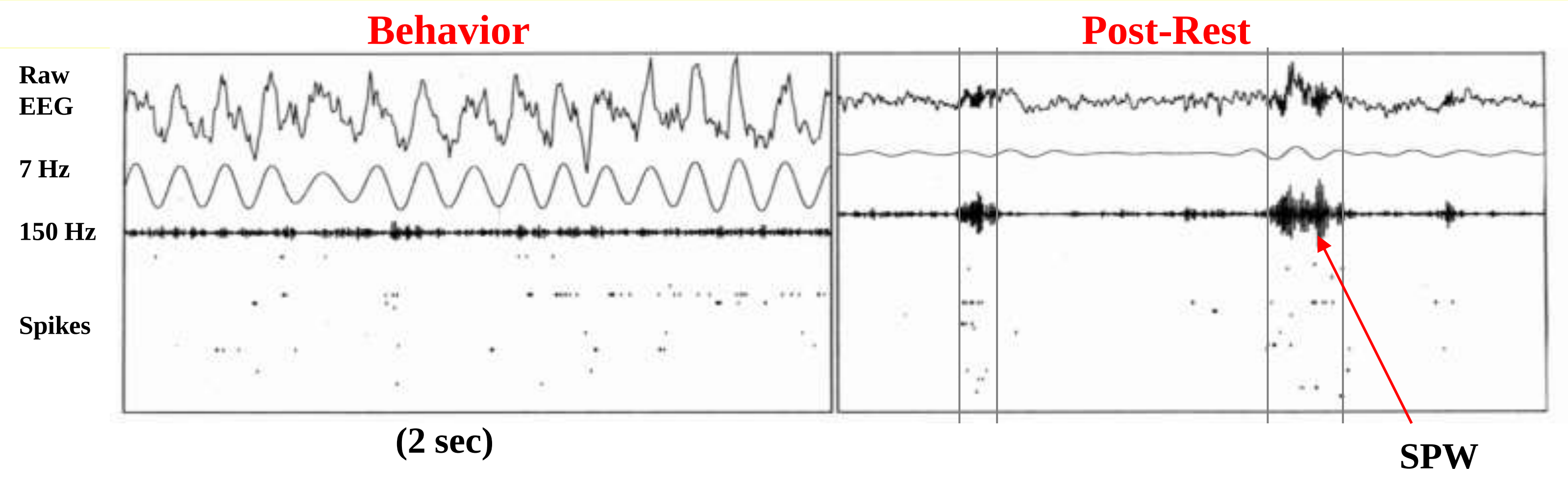


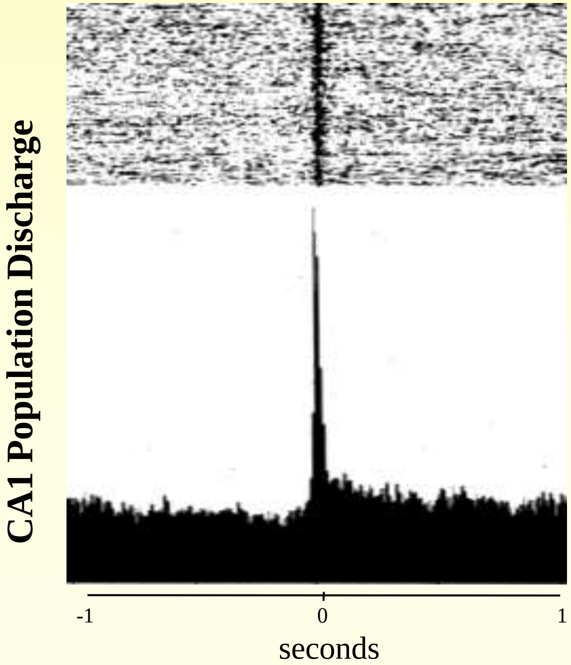


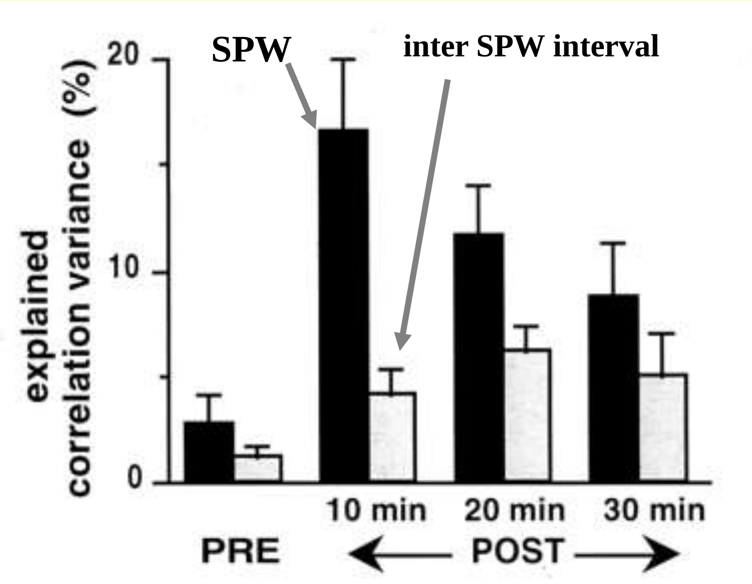
Correlation diagram for 42 simultaneously recorded CA1 pyramidal cells before, during and after a period of spatial foraging in a 70 cm^2 sq box. Lines indicate cells whose firing was correlated (red highest). The thick lines indicate correlations present during the RUN phase and also present during a rest period *either* before or after running. Note that most correlations expressed during are also expressed during rest after the running. Wilson & McNaughton, (1994), Science 265:676-679

Reactivation

Sharpwaves (SPW) are the ‘carriers’ of the ‘offline’ reactivation events in hippocampus

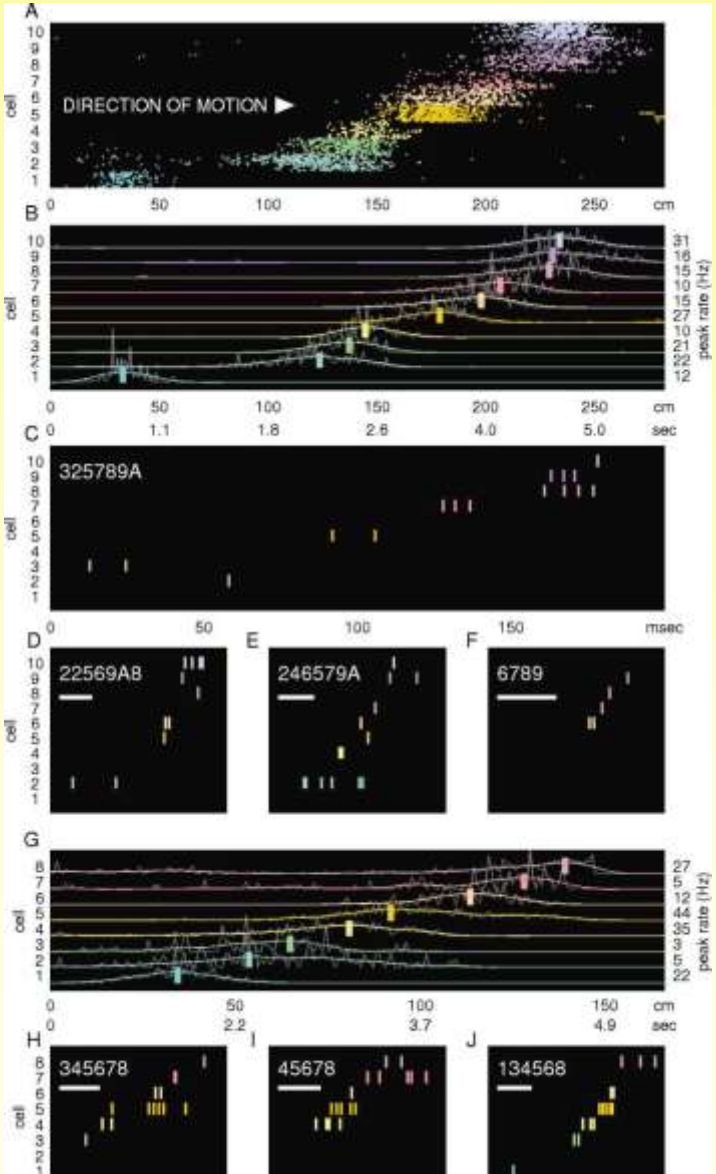






Sequence Replay

Sequence replay in hippocampus has been demonstrated using spike sequences



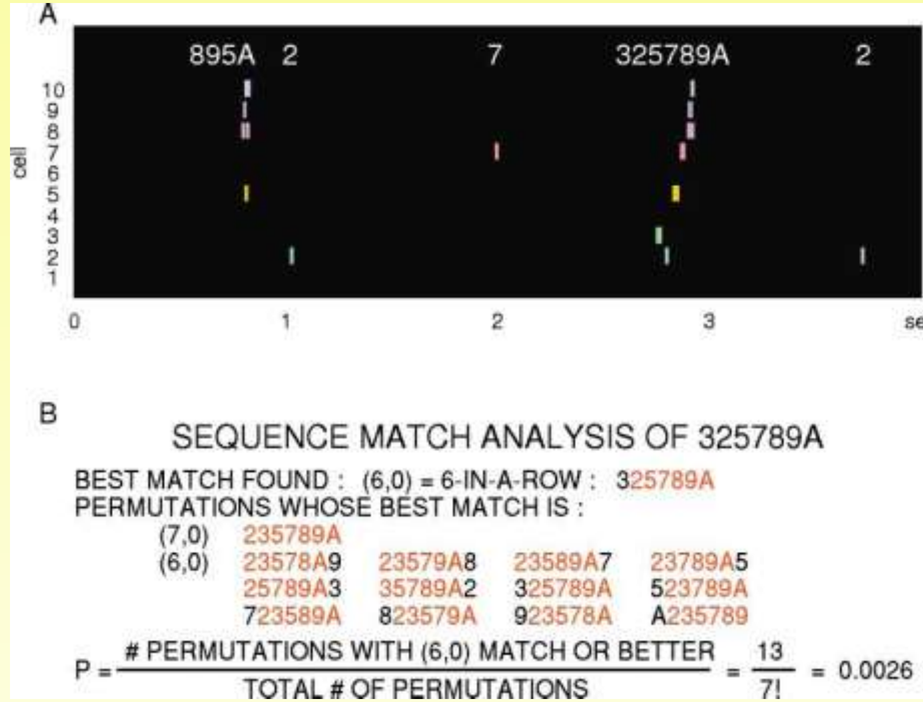


Figure 2. Example Sequences from Behavior (RUN) and Subsequent Sleep (POST)

(A–B) Determination of POS spatial sequence experienced by RAT1 in RUN. (A) Lap-by-lap rasters of all ten cells that had place fields in POS direction laps (i.e., rat running in direction of increasing position values). For each cell, laps 1–30 are stacked from bottom to top. (B) Smoothed place fields (colored lines) of these ten cells. Vertical bars mark the positions of the peaks of the smoothed fields. Smoothed firing rate (Hz) at these peaks shown to the right. Nonuniform time axis below shows time within average lap when above positions were passed. (C) A population burst from RAT1 POST SWS, showing six cells in a row firing in the same order as the POS sequence from RUN (B). Note difference in timescale. (D–F) More examples of RAT1 POST SWS population bursts that match the RUN POS sequence. (G) Same as (B), except for RAT2 POS (rat running in direction of increasing position values). (H–J) RAT2 POST SWS population bursts that match the RUN POS sequence (G). Words extracted from activity in (C)–(F) and (H)–(J) using max_isi = 50 ms and max_gap = 100 ms in upper left corner of each panel (with cell 10 represented in words by the letter A). Bar = 50 ms.

Neuron, Vol. 36, 1183–1194, December 19, 2002, Copyright ©2002 by Cell Press

Memory of Sequential Experience in the Hippocampus during Slow Wave Sleep

Albert K. Lee[1] and Matthew A. Wilson[1]
Picower Center for Learning and Memory
RIKEN-MIT Neuroscience Research Center
Department of Brain and Cognitive Sciences
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

time (Mishkin et al., 1998). A rodent hip
pyramidal cell (place cell) fires selective
mal is in a particular location (the cell'
an environment (O'Keefe and Dostrovs
different place cells fire selectively in diff
a rat's experience in moving from one loc
can be represented by the resulting seq
fields traversed. Such a sequence may

Conclusions and open problems

- ‘Classic’ EV: use control experiments, use control Sleep1 epochs (i.e. record enough sleep1).

- Carefully assess the ‘time scale’ of reactivation: (bin size/sigma, length and position of sleep epochs).

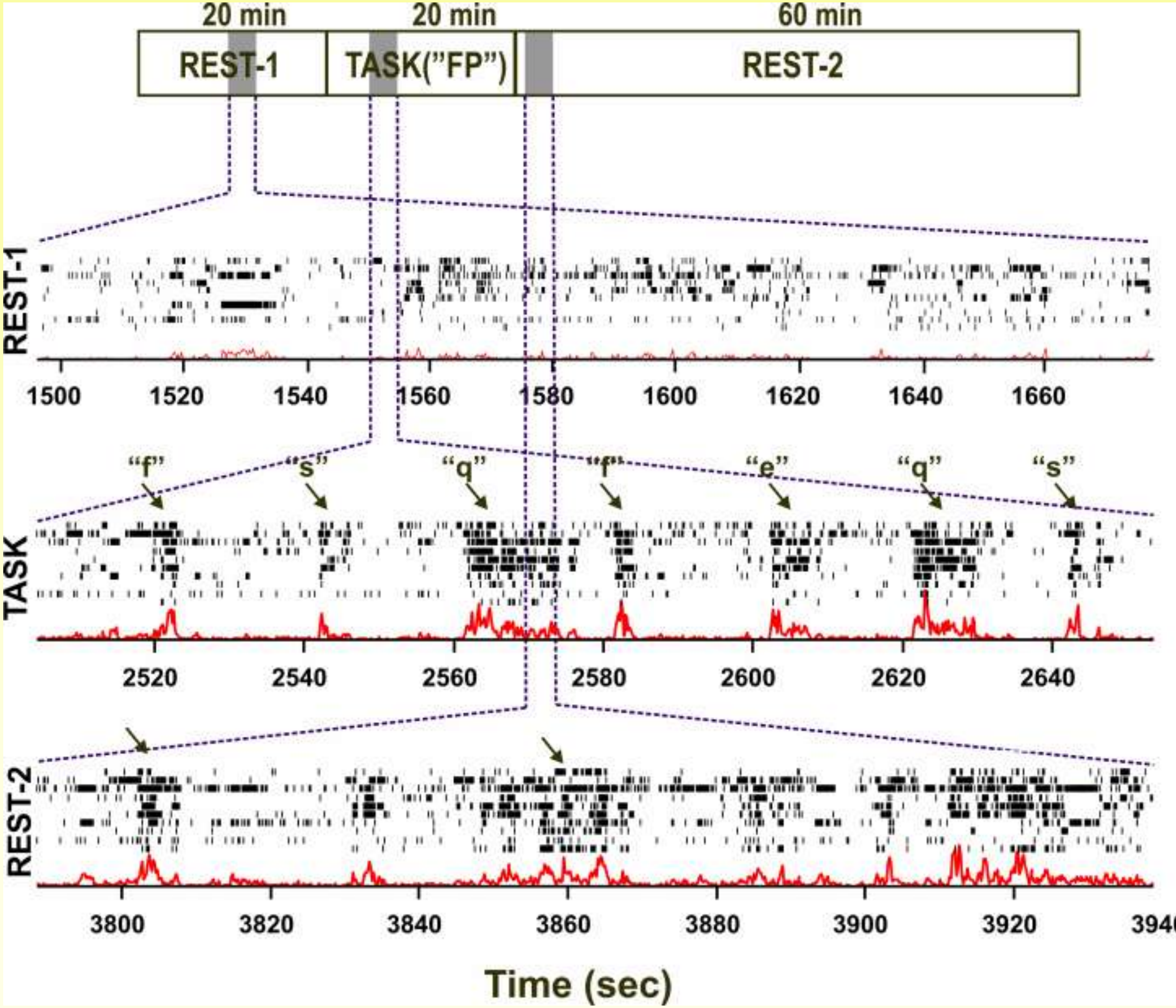
- Always compare EV and ReverseEV (significance of an EV value)

- Use enough cells (EV Vs number of cells in the population ?)

- Ev is a statistical measure: Need ways of
  - **Detecting specific patterns.**
  - **(Determining when they occur)**.
  - Determining their intrinsic time scale (compression/dilation).
  - Determining (dynamically) their cell composition.

Next!

Patterns: Multiple neurons, single trains



f = food
s = sugar
q = quinine
e = empty

'Reactivation'