Statistical Physics Approaches to High-Dimensional Inference

applications to biological data

Rémi Monasson

Laboratory of Theoretical Physics,
CNRS & Ecole Normale Supérieure, Paris

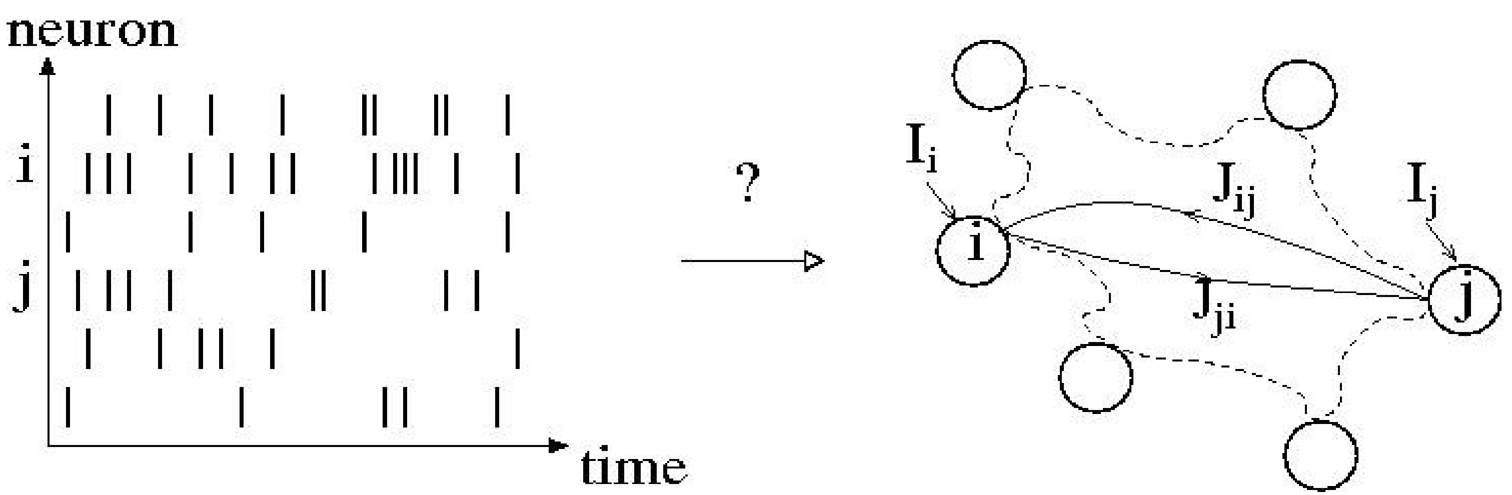
Winter School on Quantitative Biology, ICTP, December 2012

- Monday: Theoretical framework for model inference
  (special case: many interacting & stationary variables)
  Mean-field inference
  Applications covariation in protein families (I)

- Yesterday: Issues & advanced statistical physics methods
  Inverse Hopfield model & Random Matrix Theory
  Applications to covariation in proteins (II)
  to neural data (I)

- Today: Case of interacting & non-stationary variables
  Applications to neural data (II)

Getting functional interactions from the activity



Data (spiking times)

Functional Interactions

Approaches

- Principal Component Analysis or Independent Component Analysis

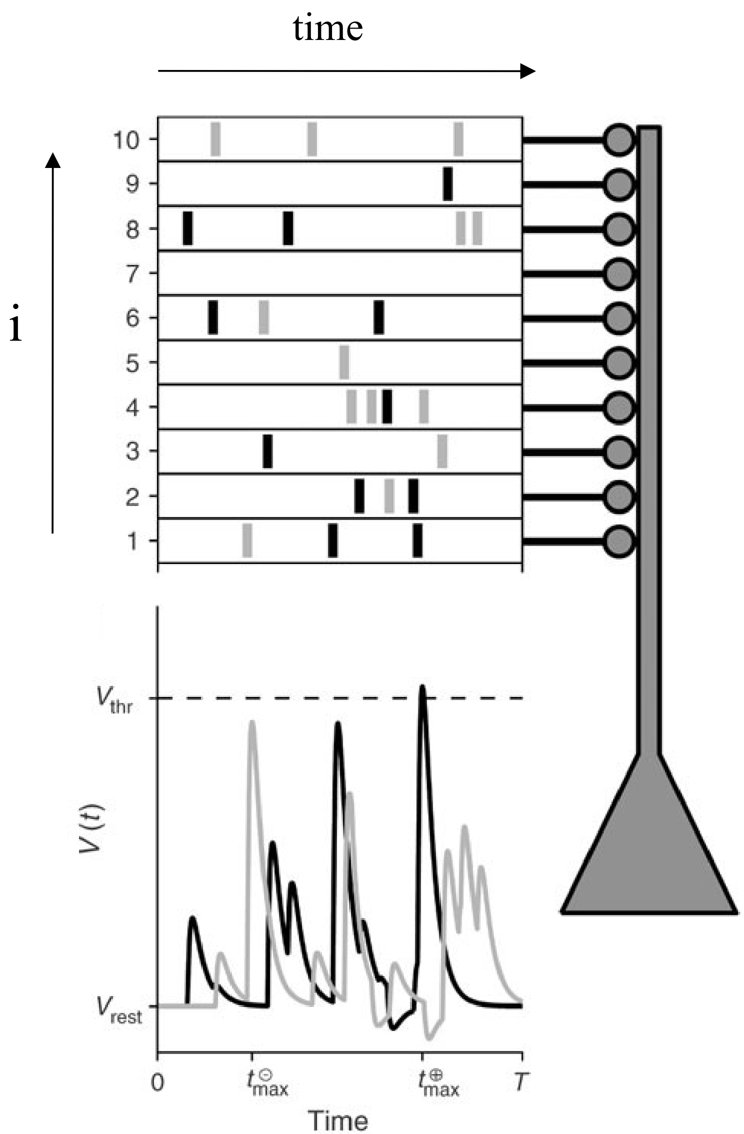
- Granger Causality (regressive models)

- Stationary Models:
  Maximum Entropy Models

- Dynamical Models:
  Generalized Linear Models
  Integrate-and-Fire Models

Bayesian approaches

The Leaky Integrate-&-Fire model



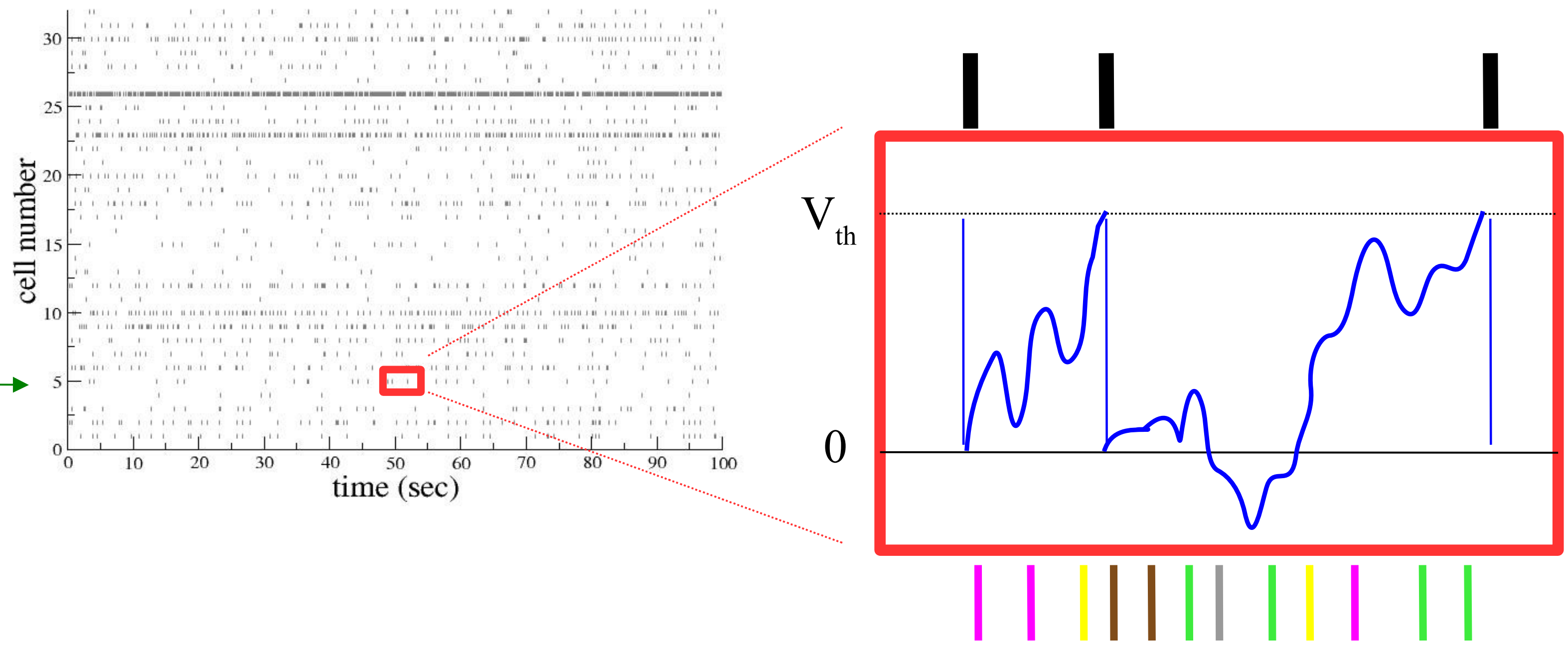
$$\boxed{C \frac{dV}{dt}(t) = - g\, V(t) + \sum_i J_i \sum_n K(t-t_{i,n}) + \eta(t)}$$

+ comparison with threshold

active or **silent**

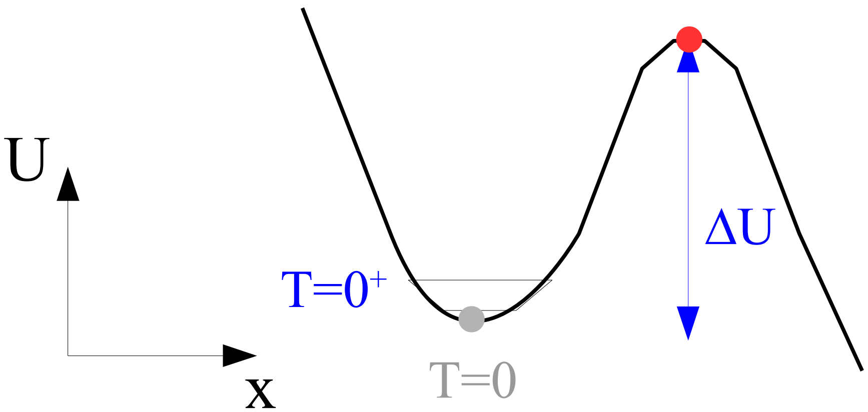
Inference of Synaptic Weights from Raster Plots

Couplings
from neurons
i to j=5?

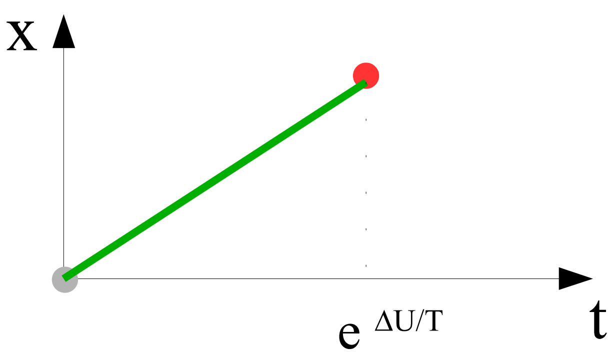


- $P[\{t_{i,n}\}|\{J_i\}]$ = product of FPT probabilities
- $\{J_i\}$ maximizing $P[\{J_i\}|\ \{t_{i,n}\}] = P[\{t_{i,n}\}|\{J_i\}]\ P_0[\{J_i\}] / P[\{t_{i,n}\}]$?
- Numerical calculation of P: limited to small data sets (Paninski et al., 2004)

Optimal Paths for the Potential and the Noise (1)



Small T:

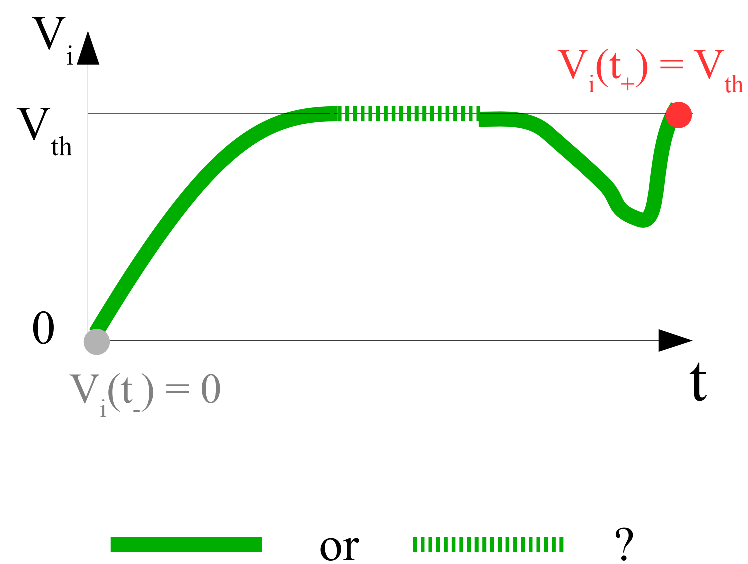


U defines the optimal paths for x(t) and f(t)

(Langer 1967)

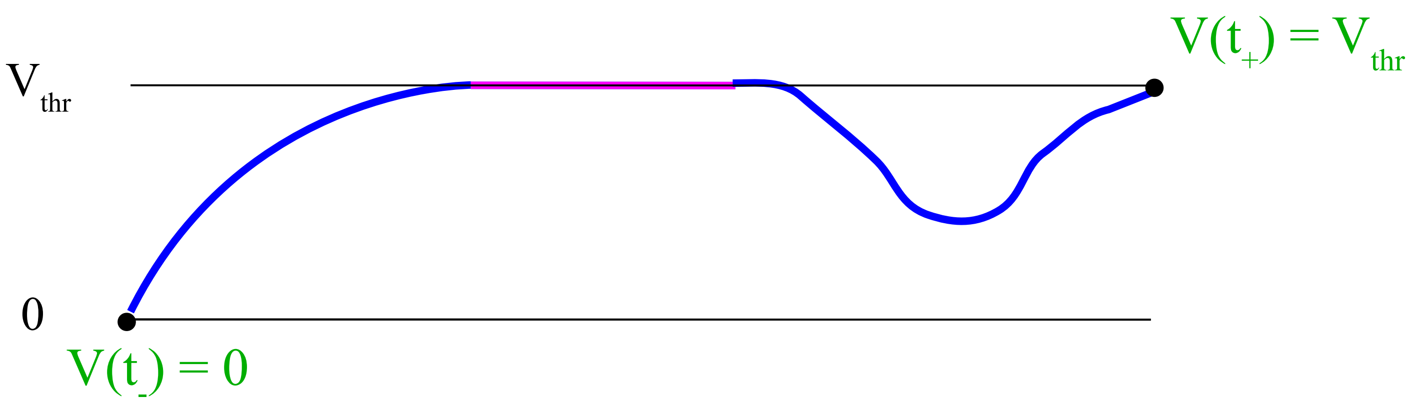
Small σ :

The likelihood P, given J,I,$\{t_{j,k}\}$, defines the optimal paths for the potential V_i and the noise (current) η_ι



Optimal Paths for the Potential and the Noise (2)

Small σ : The likelihood P, given J,I,$\{t_{j,k}\}$, defines the optimal paths for the potential V_i and the noise (current) η_i



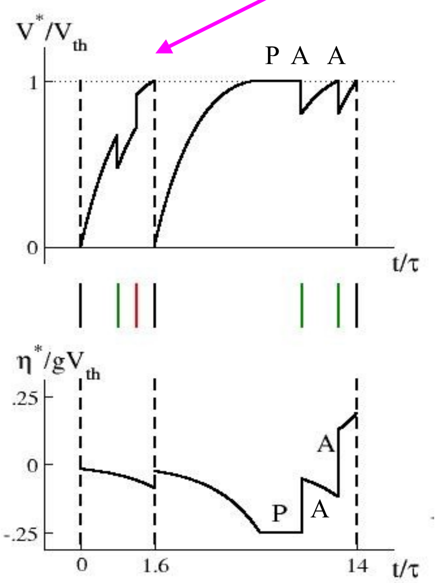
$$\left.\begin{aligned} \frac{dV}{dt}(t) &= -\, g\, V(t) + \sum_i J_i \sum_n K(t-t_{i,n}) + \eta(t) \\ \frac{d\eta}{dt}(t) &= +\, g\, \eta(t) \end{aligned}\right\} \quad \textit{or} \quad \left\{\begin{aligned} &V(t)=V_{thr} \\ &\eta(t) = g\, V_{thr} - \sum_i J_i \sum_n K(t-t_{i,n}) \end{aligned}\right.$$

Particular case: K = δ, i.e. instantaneous synaptic integration

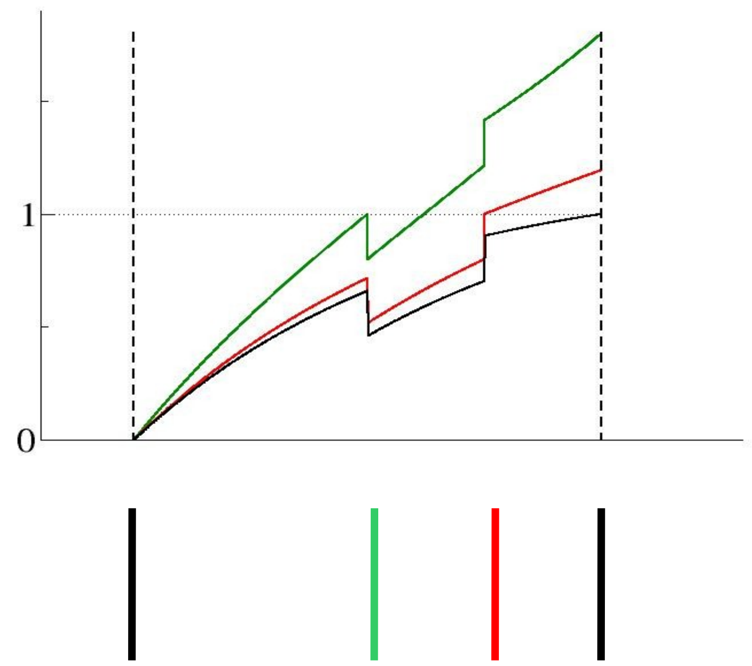
S. Cocco, S. Leibler, R.M. (2009); R.M., S. Cocco (2011)

Looking for Contacts (1)

Simplest case : no intermediate contact with the threshold potential



?

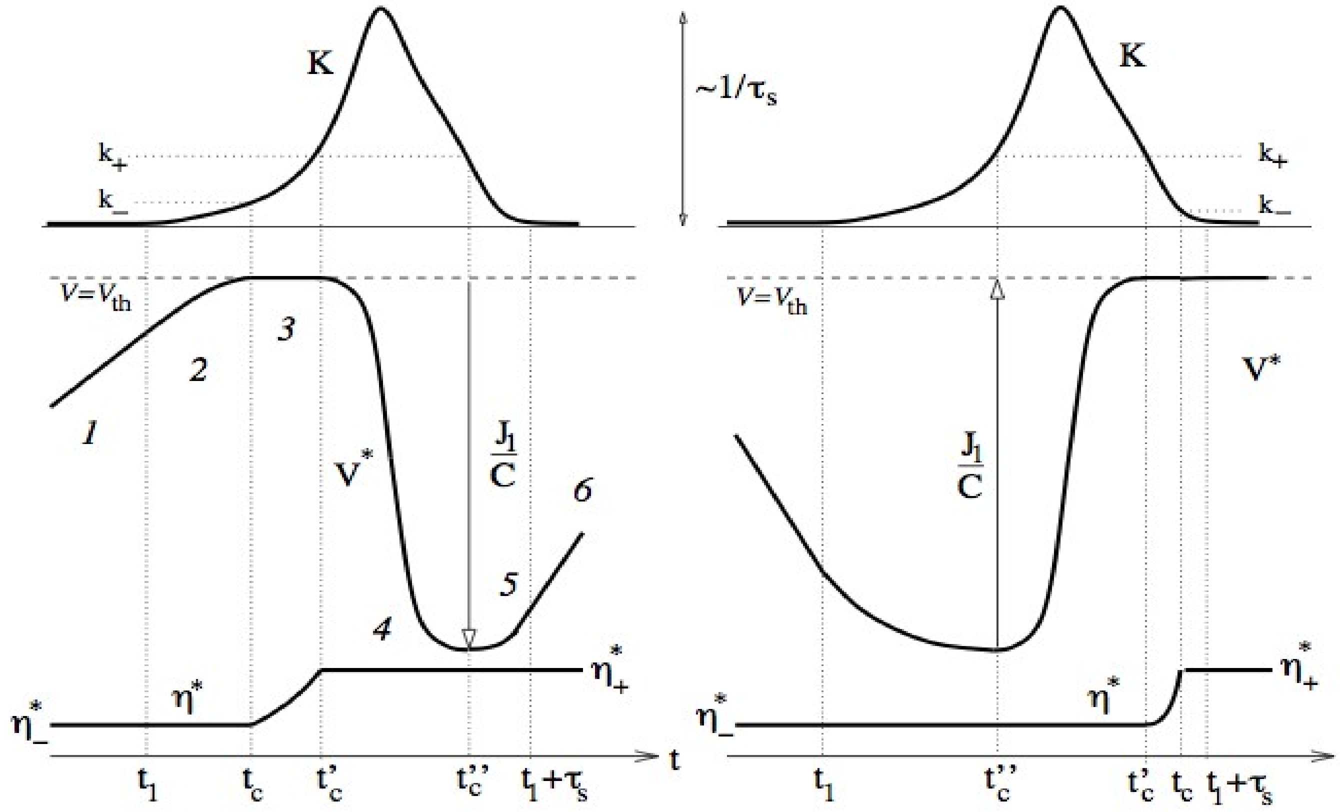


$I = 1.2\ gV_{th}$

$J_{red} = -\ J_{green} = .2\ \ CV_{th}$

Looking for Contacts (2)

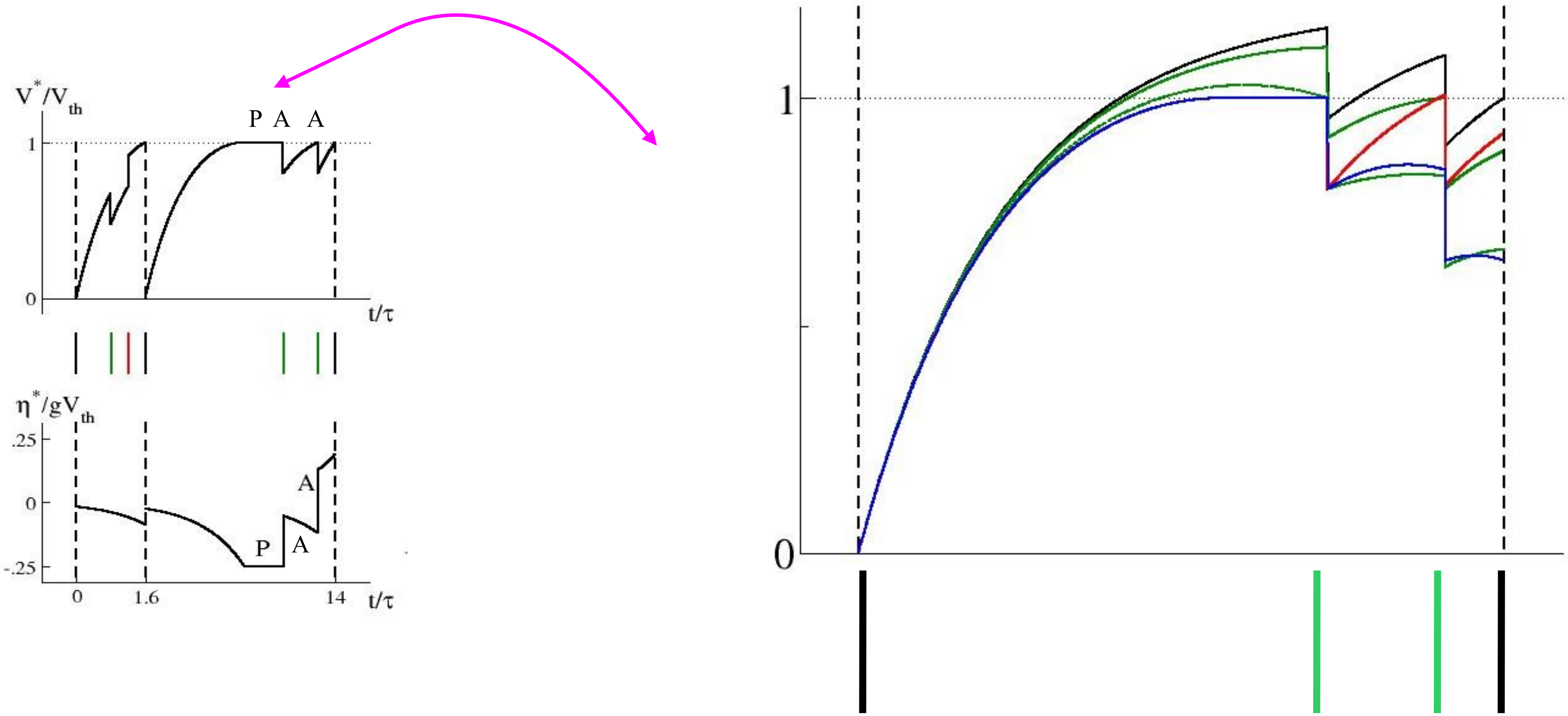
Analysis of the intermediate contact region :



<u>**Contact Rules:**</u> Contacts take place :

- Simultaneously to an input spike, with a discontinuous increase of the noise (active contacts)
- On a finite duration interval, during which noise is constant (passive contacts)

Looking for Contacts (3)

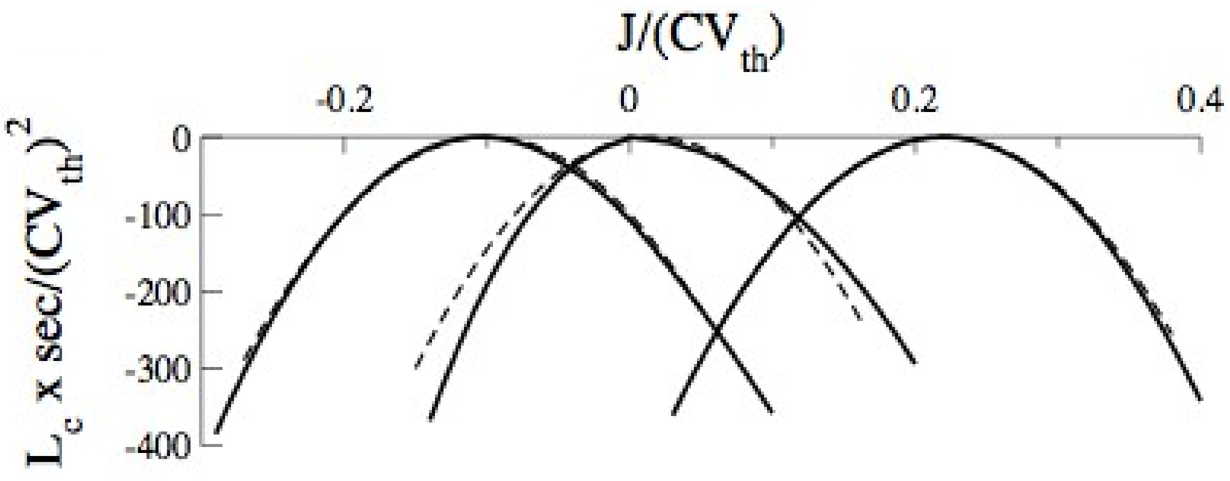


Practice:

- if the current and the interactions are known, the optimal paths for the potentials, for the noises, and the likelihood can be found in time = cst $N^2 \times S$
- optimal current and couplings are found by maximizing the likelihood (which is convex)

Properties of the inference procedure (1)

• Gives most likely interactions J and error bars ΔJ



• Tested on synthetic data generated from networks with known couplings; accurate as long as the noise variance is moderate: $\sigma^2 < f(g,C,V_{th},ISI)$

Properties of the inference procedure (2)

- Fast : recordings of 40 neurons, 170,000 spikes in 30 sec on commercial PC
  (up to N=160, 5 millions spikes with synthetic data)

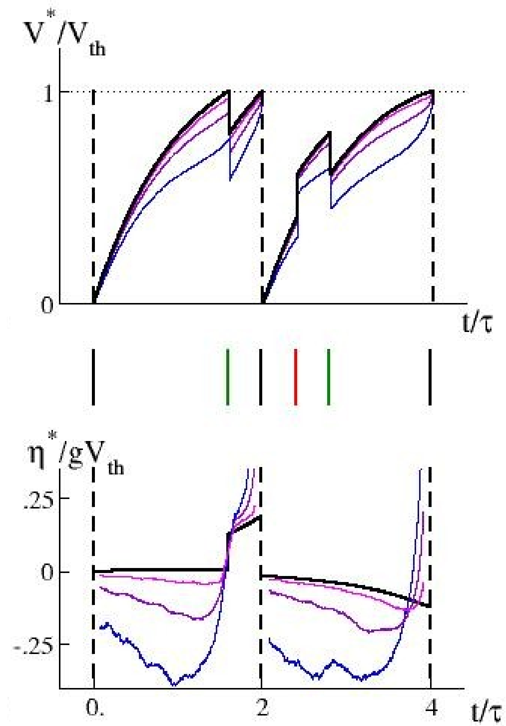
  related works: N=1 neuron, ~ 100 spikes (for any σ) (Pillow et al., 2005)
  N=5 neurons, ~ 500 spikes (σ=0) (Maretsos et al., 2004)
  time discretization $\rightarrow$ constrained optimization
  (Koyama, Paninski, 2009)

- Limitations: synaptic integration kernel is instantaneous
  no external stimulus so far
  noise must be small (empirical extension to moderate noise)

How to deal with moderate noise? (1)

- optimal paths are exact in the $\sigma \to 0$ limit

- σ = .07, .18, .36 $V_{th} (g\, C)^{1/2}$
  $I = 1.2\, gV_{th}$, $J_{red} = -\, J_{green} = .2\ CV_{th}$



B

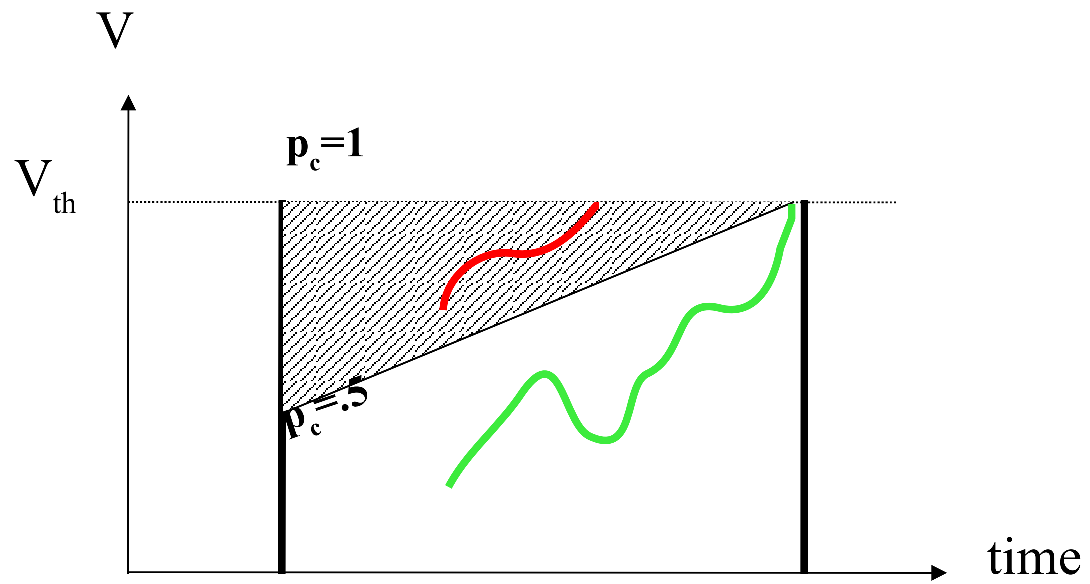
- Gaussian fluctuations around the σ=0 optimal path

$$\frac{\langle (V_i - V_i^*)^2 \rangle}{V_{th}^2} = \bar{\sigma}^2 \ \tanh\left(\frac{t_{i,k+1} - t_{i,k}}{2\tau} \right) \qquad \text{with} \qquad \bar{\sigma} = \frac{\sigma}{V_{th}\sqrt{gC}}$$

- Pessimistic point of view: Gaussian corrections to log-likelihood are independent of interaction parameters!

How to deal with moderate noise? (2)

- Probability of early crossing

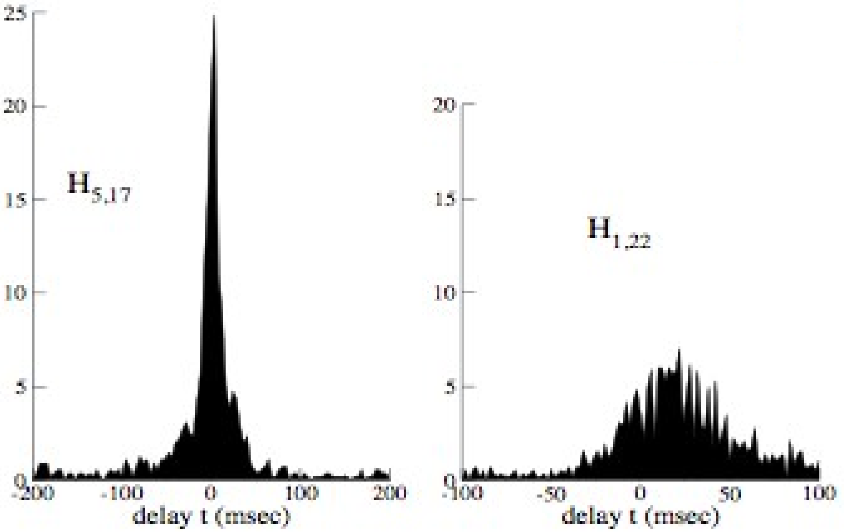


- Algorithm with Moving threshold $V_{th}(t)$ …

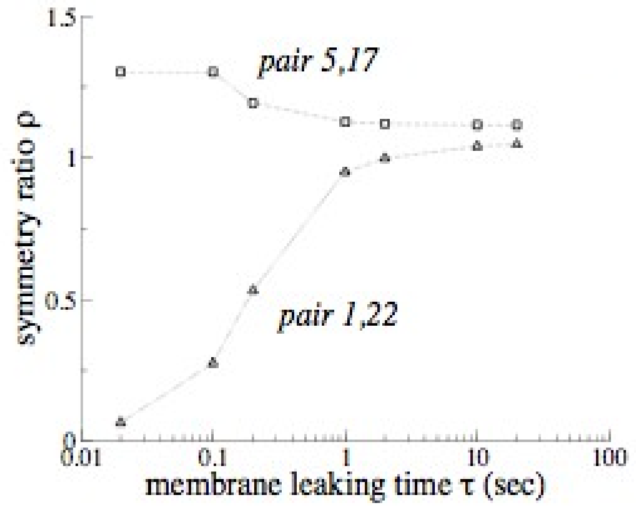
Application to Multi-Electrode Recordings (1)

Asymmetry of interactions and time-scales
(spontaneous activity)

Cross-correlograms



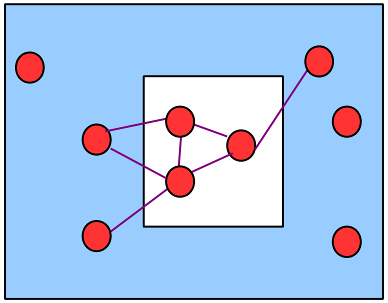
Ratio J_{ij} / J_{ji}



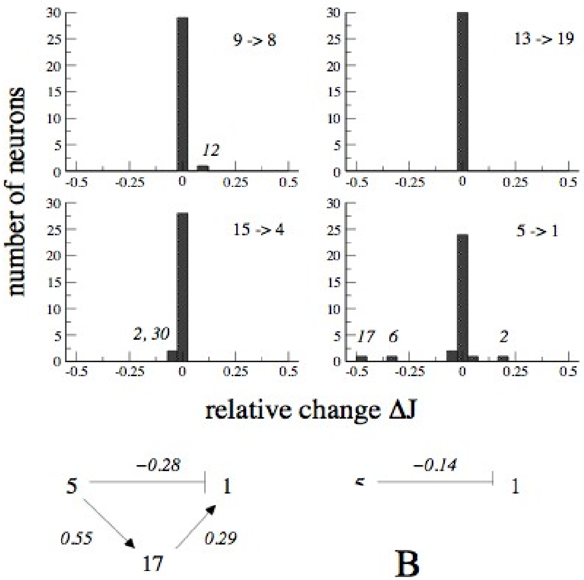
Application to Multi-Electrode Recordings (2)

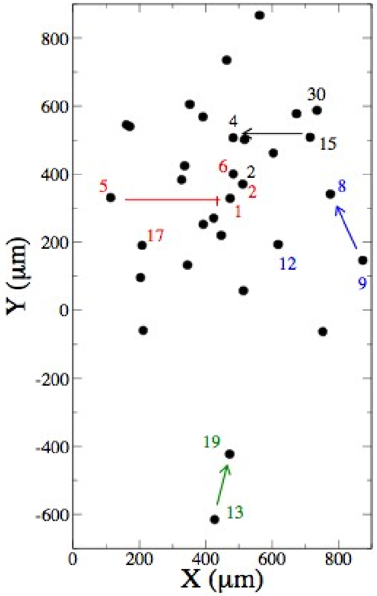
Salamander retinal ganglion cell activity: (N=32 cells, spontaneous activity and random flickers - M. Meister)

Is the inference problem well-behaved?



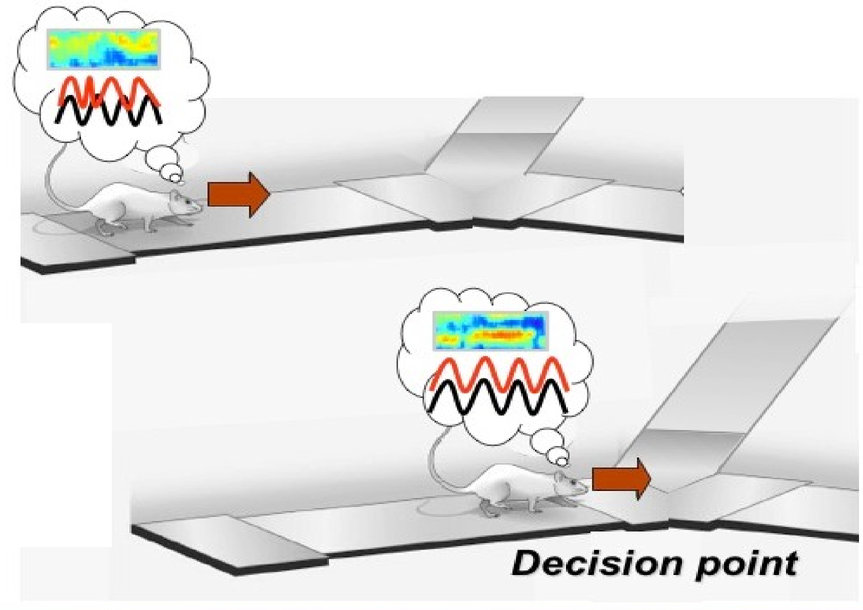
Spontaneous activity





Exploring learning and memory from in vivo cortical multi-electrode recordings

Peyrache, Battaglia et al. (2009): 37 cells
in prefrontal cortex, before, during and after a task



Principal component analysis of cortical recordings
(Peyrache, Battaglia et al. 2008)

Rat perform a rule shift task four possible rules
Activity of prefrontal cortex is recorded during:
The sleep period prior the task (PRE)- the task- and the sleep period after the task(POST)

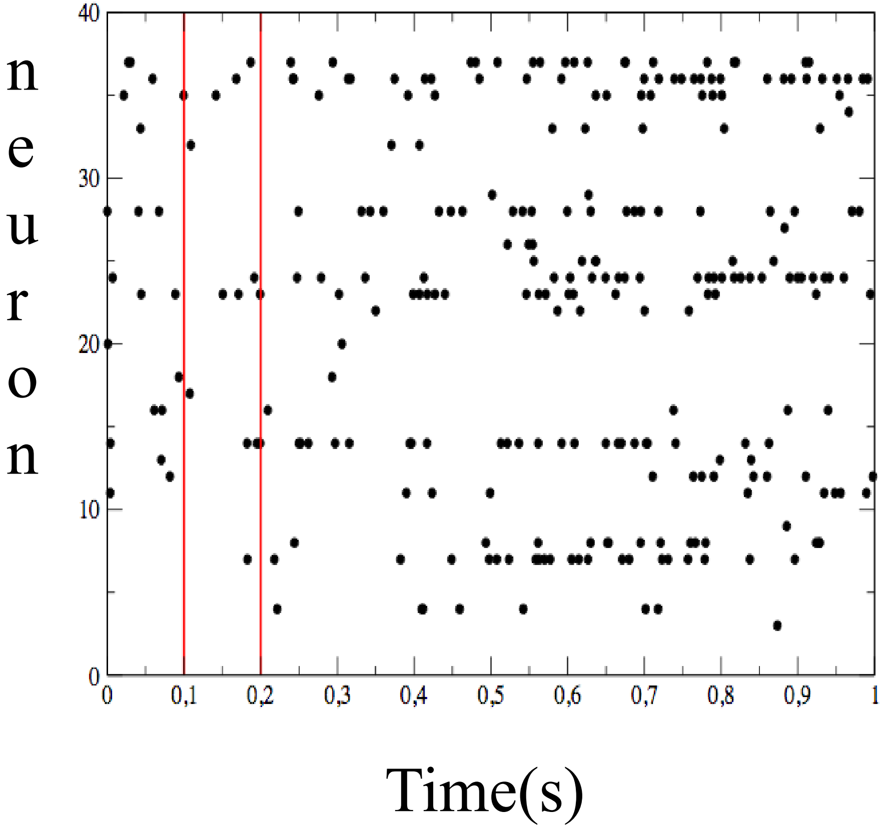
Patterns:
Starting with D. Hebb's (1949) theorists have postulated cell assemblies as the main unit of information representation which create a coherent input to downstream areas.

Replay and memory consolidation:
replay of the pattern of activity during the SWS (slow wawe sleep) in period corresponding to coordinated bursts of activity of the hippocampus (sharp waves)

Replay of rule-learning related neural patterns in the prefrontal cortex during sleep A. Peyrache.. F. Battaglia Nature Neuroscience 2008,

Principal component analysis of ensemble recordings reveals cell assemblies at high temporal resolution A.Peyrache … F. Battaglia J. Comput Neurosci 2009

1. Spike trains from the awake epoch are binned



Time is discretized in time windows of size Δt=100 ms

$$s_i^{\tau} = \begin{cases} 1 \text{ if at least one spike in time window k} \\ 0 \text{ if no spike in time windows k} \end{cases}$$

$$p_{kl} = \frac{1}{B}\sum_{\tau=1}^{B} s_k^{\tau}\, s_l^{\tau} \,,$$

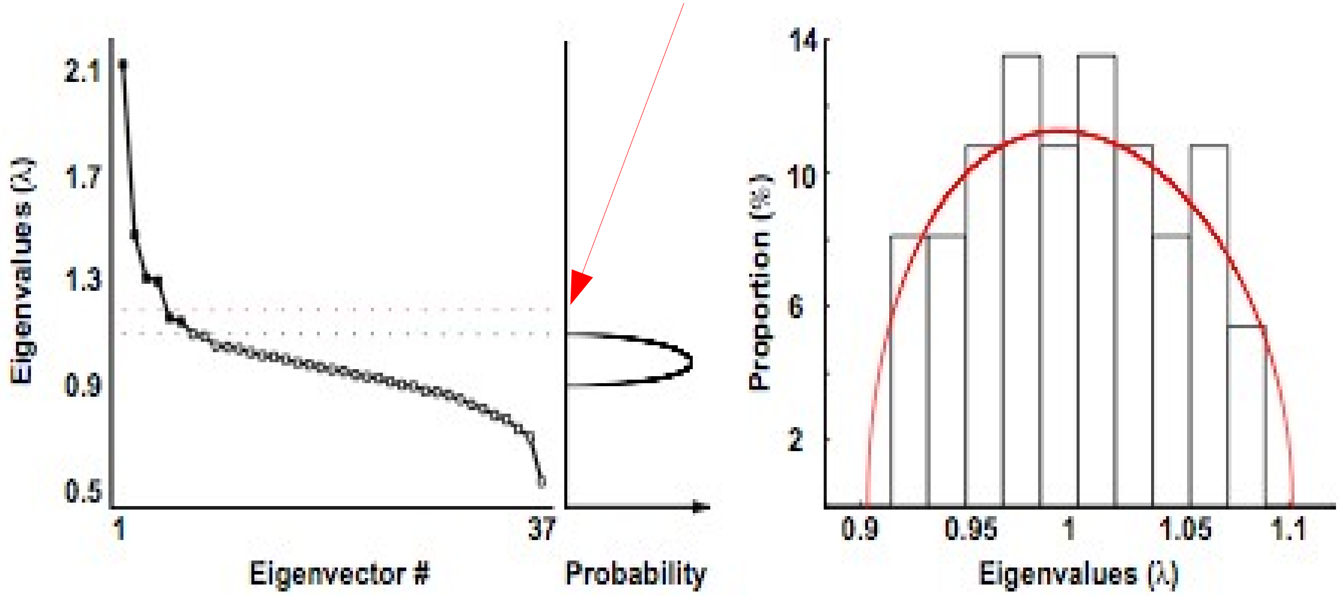
$$p_i = \frac{1}{B}\sum_{\tau=1}^{B} s_i^{\tau}$$

2. Correlation matrix computed and diagonalized

$$\chi_{ij} = \frac{p_{ij} - p_i p_j}{\sqrt{p_i p_j (1-p_i)(1-p_j)}}$$

3. Only eigenvectors associated to the largest eigenvalues are retained, threshold value from the upper bound of eigenvalues of correlation matrix of independent, normally distributed spike trains

Tracy-Widom corrections

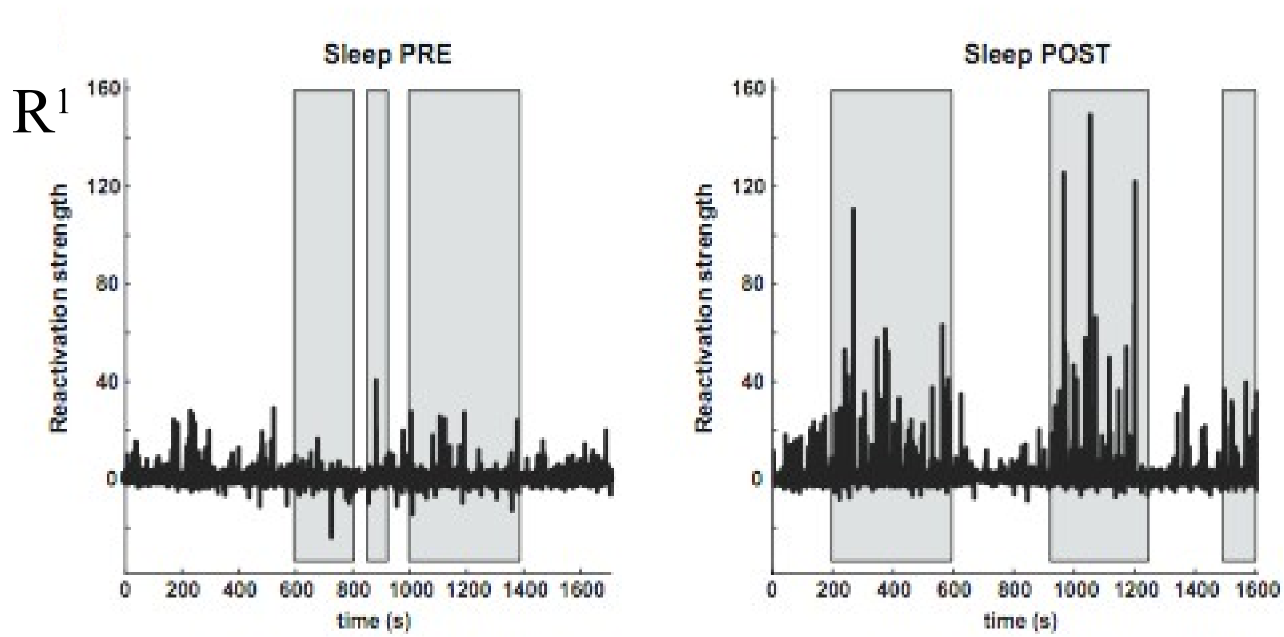


Marcenko-Pastur distribution

4. Spike trains from the sleep epoch are binned

5. The instantaneous similarity of the sleep population activity with the awake activity is computed through the

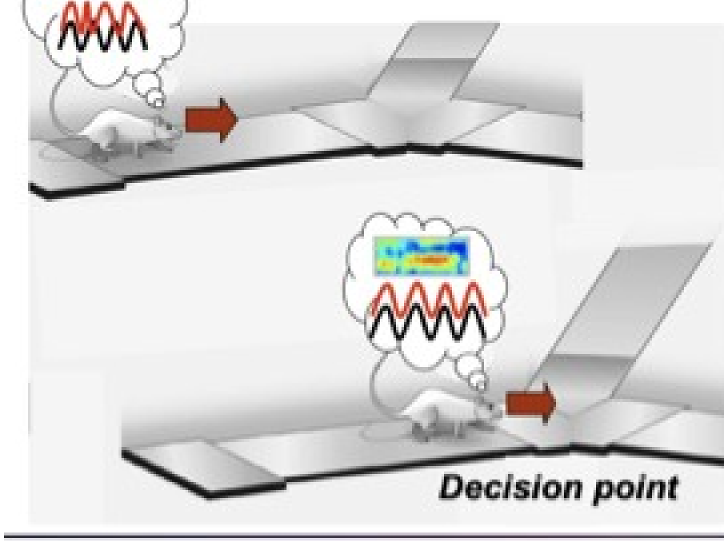
Squared overlap $R^q(\tau)=(\Sigma_i v^q_i s_i(t))^2$



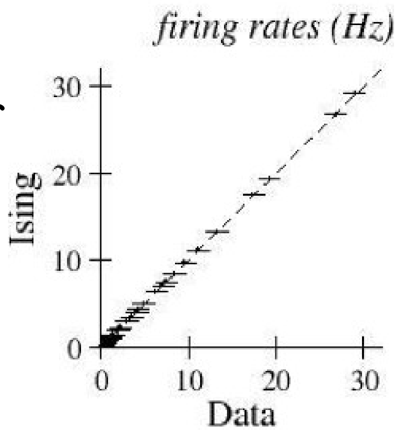
- *Instantaneous similarity is high in Slow Wave Sleep (SWS) (shaded areas) after learning of the task in coincidence with hippocampal sharp waves*

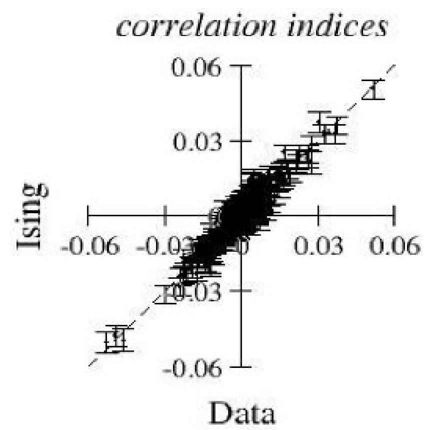
Analysis of the Cortical Activity of a Behaving Rat

Peyrache et al., 2008

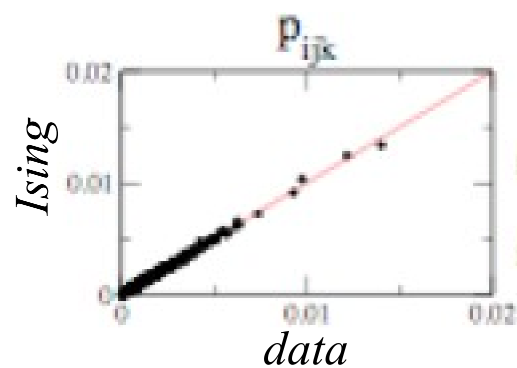


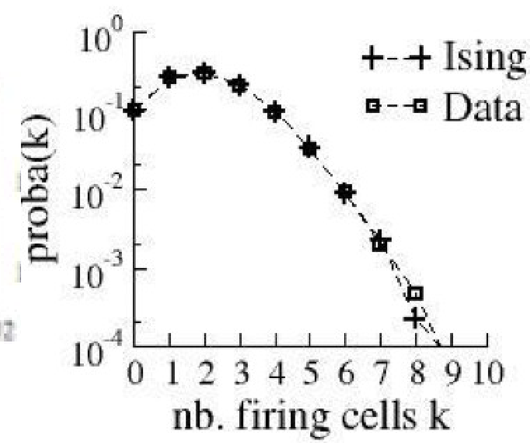
Inference of effective network J

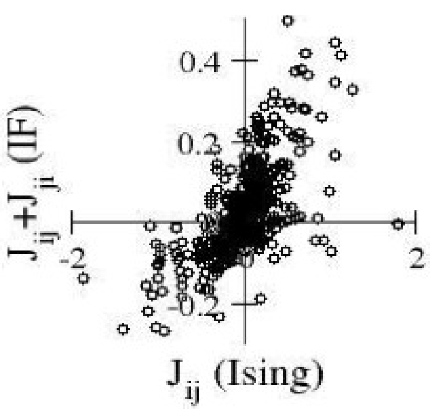




Reproduces statistics of activity







Comparison Ising - I&F

Exploring neural coding by novel optogenetic, high-density microrecordings and computational approaches

With C. Bartic (Leuven), F. Battaglia (Amsterdam), S. Cocco (Paris), M. Giugliano (Antwerpen)

Goal 1: test capability of predicting response to a specific stimulus

Goal 2: estimate changes of network due to task learning (*distributed, sparse ?)*

Goal 3: identify & manipulate cell assemblies coding for memories