State-space Model of Dynamic Correlations in Parallel Spike Sequences

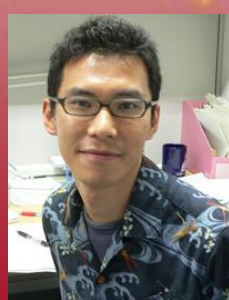
Speaker: Hideaki Shimazaki

Nonlinear Dynamics Seminar, Jun. 1, Kyoto

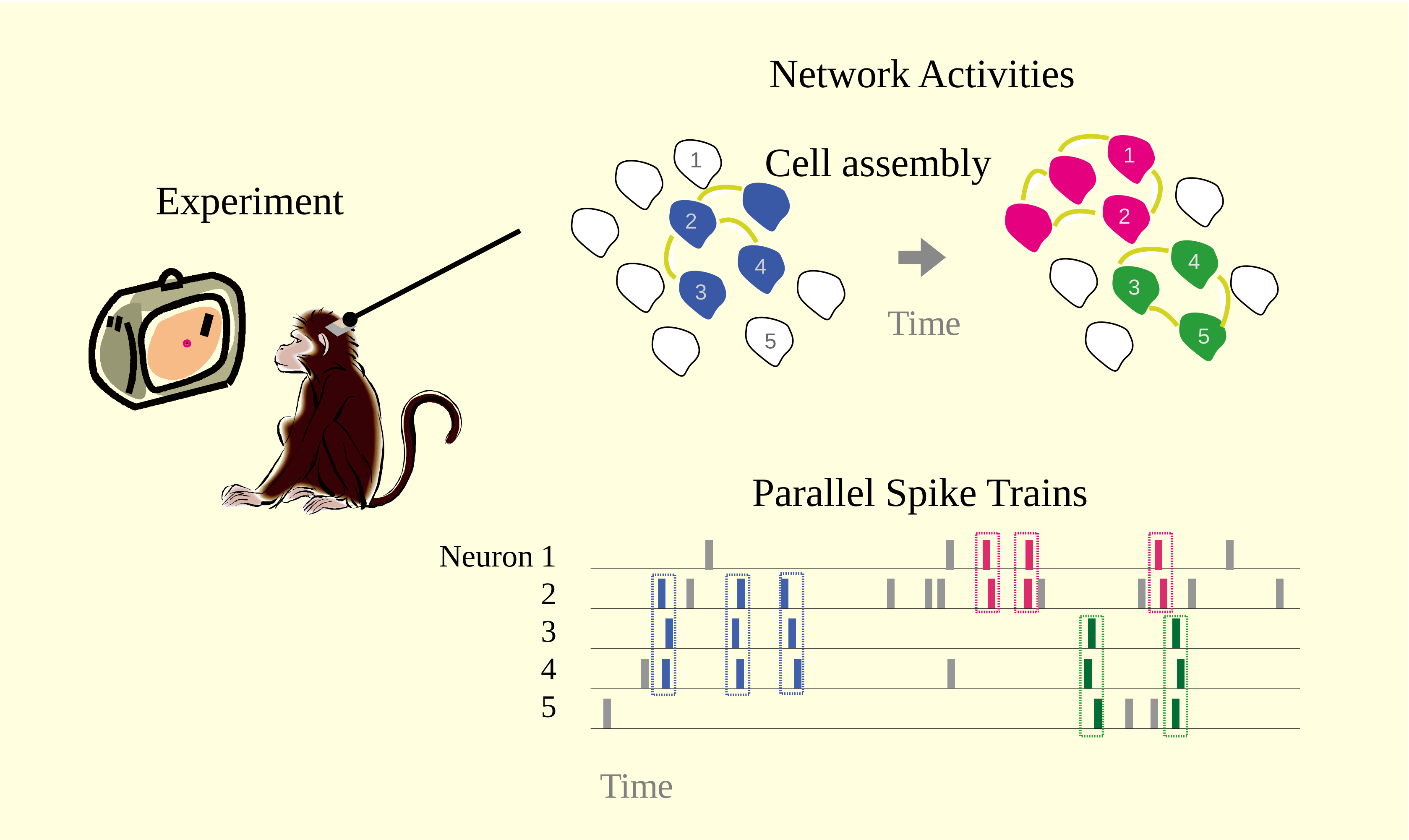
Hideaki Shimazaki[1], Shun-ichi Amari[1], Emery N. Brown[2], Sonja Grün[1]

1 RIKEN Brain Science Institute, Wako-shi, Saitama, Japan

2 Anesthesia and Critical Care, Massachusetts General Hospital, Boston, MA, USA
Harvard-MIT Division of Health Sciences and Technology, Cambridge, MA, USA



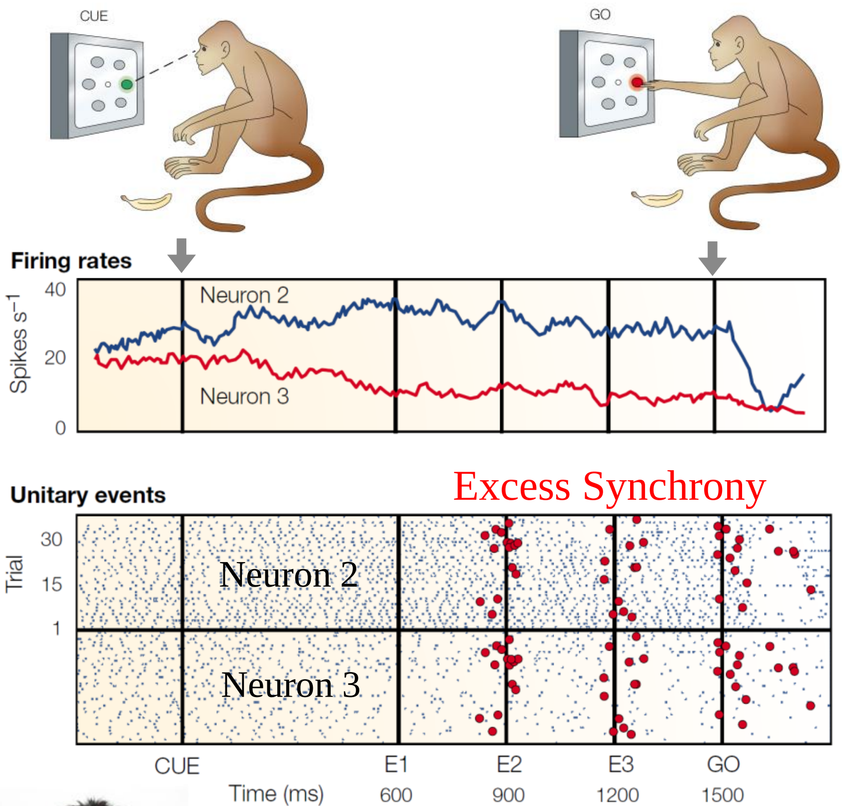

Cooperative Activities of Neurons



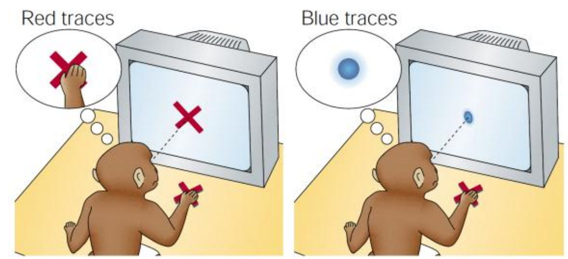
From the observation of synchronous activities,
we wish to trace dynamically correlated neuronal groups.

Evidence for Dynamic Correlations

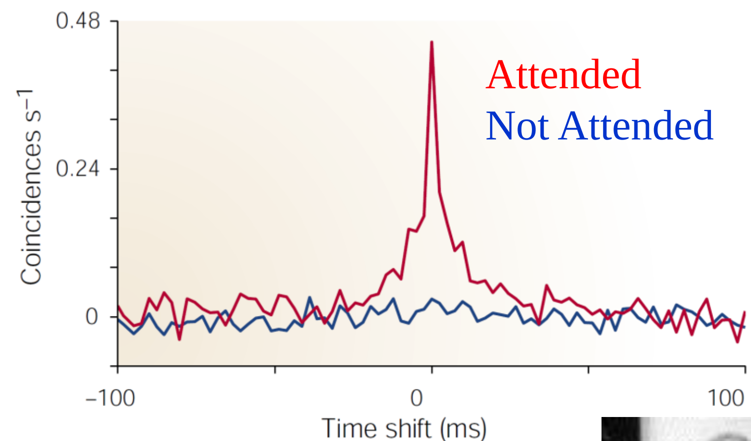
Expectation



Attention

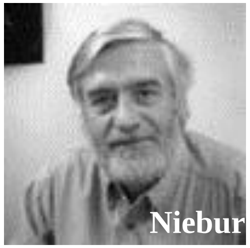


Spike correlogram



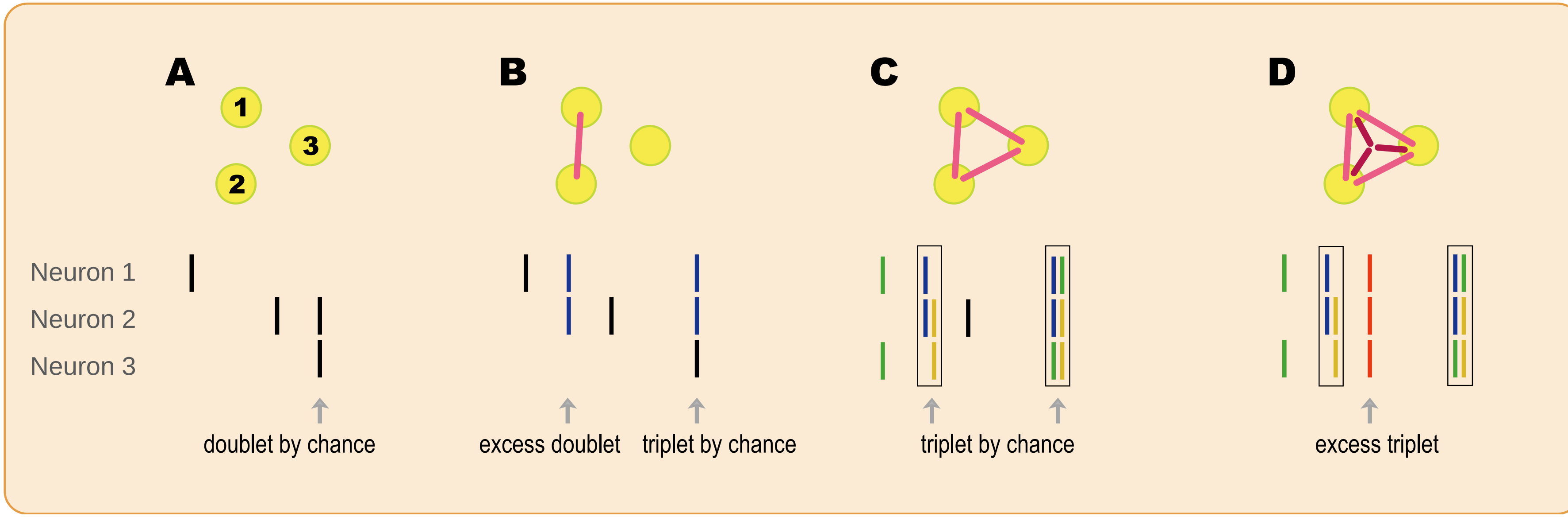
Riehle, Grün, Diesmann, Aertsen
Science 278: 1950-1953, 1997.

Steinmetz, Roy, Fitzgerald, Hsiao, Johnson, Niebur
Nature, *404*, 187-190, 2000.



Spike correlation in ms precision occurs at behaviorally relevant instances.

What is the Higher-order Correlation?



A Neurons are independent.
Synchrony may appear by chance.

B Neuron 1 and 2 are correlated.
Triplets may appear by chance.

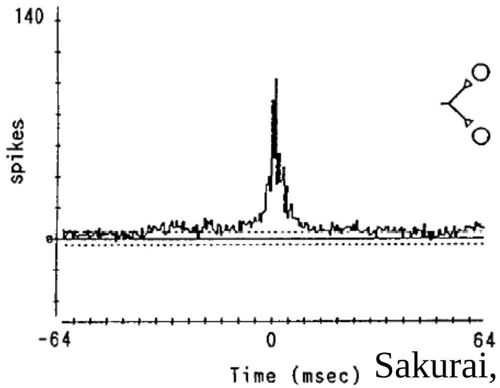
C Neurons are pairwisely correlated.
Triplets may appear by chance.

D A triplewise correlation is added.
Excess triplets are generated.

Higher-order correlation generates excess synchrony which can not be explained from the lower-order correlations.

Correlation Analysis Methods

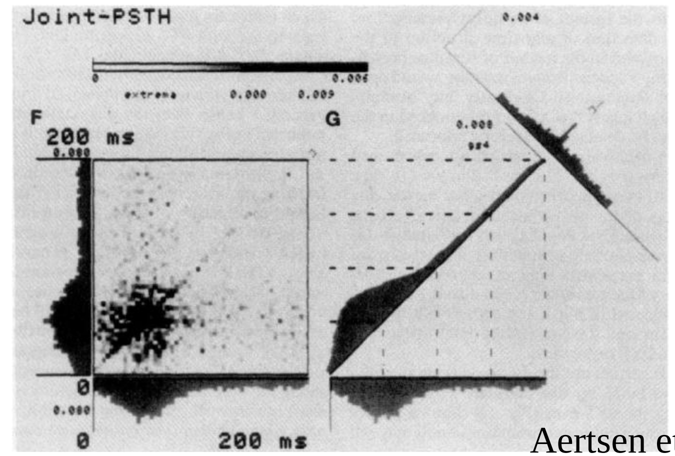
Cross-correlogram



Sakurai, 1993

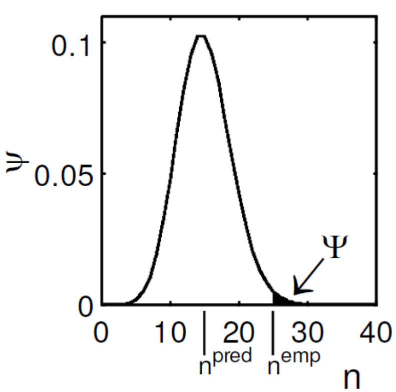
Perkel, 1967; Ahissar et al., 1992;
Sakurai, 1993; Steinmetz, 2000;
Sakurai and Takahashi, 2006; Fujisawa, 2008

Joint-PSTH



Aertsen et al. 1989

Gerstein and Perkel, 1969
Aertsen et al. 1989



Unitary Event Analysis

Riehle et al. 1997; Gruen, 2002

Test on statistical dependence of multiple neurons against the null-hypothesis of full independence

See Brown et al. 2004 for a review

None of them can trace the dynamics of correlations, incl. higher-order dependency.

Objective

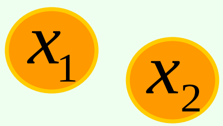
- **Estimation of the Time-dependent Higher-order Correlations in Parallel Spike Trains**

Our approach:

State-space Analysis with the Log-linear Model

The Log-linear Model

Log-linear model for 2 neurons



x_1	x_2	
0	0	p_{00}
1	0	p_{10}
0	1	p_{01}
1	1	p_{11}

$$\log p_{00} = -\psi$$
$$\log p_{10} = -\psi + \theta_1$$
$$\log p_{01} = -\psi + \theta_2$$
$$\log p_{11} = -\psi + \theta_1 + \theta_2 + \theta_{12}$$

2 neurons $$\log p(x_1, x_2) = -\psi + \theta_1 x_1 + \theta_2 x_2 + \theta_{12} x_1 x_2$$

2nd order correlation

3 neurons A *full model* $$\log p(\mathbf{x} \mid \boldsymbol{\theta}) = -\psi + \sum_i \theta_i x_i + \sum_{i<j} \theta_{i,j} x_i x_j + \theta_{123} x_1 x_2 x_3$$

3rd order correlation

A *pairwise model* $$\log p(\mathbf{x} \mid \boldsymbol{\theta}) = -\psi + \sum_i \theta_i x_i + \sum_{i<j} \theta_{i,j} x_i x_j$$

pairwise correlation

The higher-order parameters of the log-linear model indicate the higher-order correlations.

Amari, IEEE Trans. Inf. Theory, 2001

Orthogonal coordinates

Log-linear model $p(\mathbf{x}|\boldsymbol{\theta}) = \exp(-\psi + \theta_1 x_1 + \theta_2 x_2 + \theta_{12} x_1 x_2)$

x_1	x_2	P-coordinate
0	0	p_{00}
1	0	p_{10}
0	1	p_{01}
1	1	p_{11}

η-coordinate

$$\eta_1 = E[x_1] = p_{10} + p_{11}$$
$$\eta_2 = E[x_2] = p_{01} + p_{11}$$
$$\eta_{12} = E[x_1 x_2] = p_{11}$$

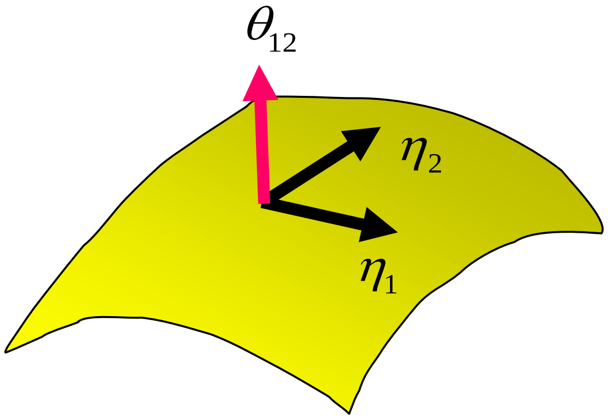
θ-coordinate

$$\log p_{00} = -\psi$$
$$\log p_{10} = -\psi + \theta_1$$
$$\log p_{01} = -\psi + \theta_2$$
$$\log p_{11} = -\psi + \theta_1 + \theta_2 + \theta_{12}$$

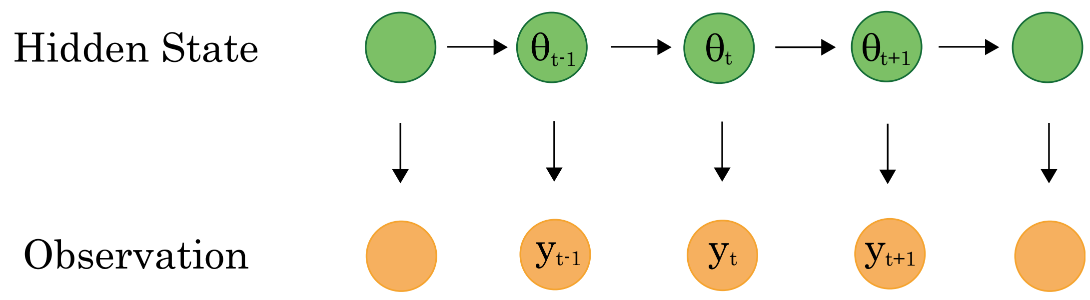
$$J(\eta_i, \theta_{12}) = E\left[\frac{\partial}{\partial \eta_i} \log p(x|\boldsymbol{\theta}) \frac{\partial}{\partial \theta_{12}} \log p(x|\boldsymbol{\theta})\right] = 0$$

$(\eta_1, \eta_2, \theta_{12})$ is an orthogonal coordinate

Amari 2001, Nakahara & Amari 2002.



State-space Analysis



State model (Gaussian) $\boldsymbol{\theta}_t = \mathbf{F}\boldsymbol{\theta}_{t-1} + \boldsymbol{\xi}_t \qquad \boldsymbol{\xi}_t \sim N(\mathbf{0}, \mathbf{Q}) \quad \boldsymbol{\theta}_1 \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

Observation model (Log-linear) $p\left(\mathbf{y}_{1:T} \mid \boldsymbol{\theta}_{1:T}\right) = \prod_{t=1}^{T} \exp\left[n\left(\mathbf{y}_t^T \boldsymbol{\theta}_t - \psi\right)\right]$

$\boldsymbol{\theta}_t = \left[\theta_1^t, \ldots, \theta_N^t, \theta_{12}^t, \theta_{13}^t, \ldots, \theta_{123}^t, \ldots, \theta_{1\cdots N}^t\right]^T$ A vector of log-linear parameters.

$\mathbf{w} = [\boldsymbol{\mu}, \boldsymbol{\Sigma}, \mathbf{Q}, \mathbf{F}]$ Hyper-parameters.

$\mathbf{y}_t = \left[y_1^t, \ldots, y_N^t, y_{12}^t, y_{13}^t, \ldots, y_{123}^t, \ldots, y_{1\cdots N}^t\right]^T$ A vector of observation for synchronous spikes rate.

$y_{123}^t = \{\text{\# of joint spikes of Neuron 1,2,3 occured at time t}\}/n$ $\quad n$ The number of total trials.

Posterior density

Posterior

Likelihood (Log-linear model)

Prior (State equation)

$$p\left(\boldsymbol{\theta}_{1:T} \mid \mathbf{y}_{1:T}, \mathbf{w}\right) = \frac{p\left(\mathbf{y}_{1:T} \mid \boldsymbol{\theta}_{1:T}\right) p\left(\boldsymbol{\theta}_{1:T} \mid \mathbf{w}\right)}{p\left(\mathbf{y}_{1:T} \mid \mathbf{w}\right)}$$

Evidence

$$\mathbf{w} = \left[\boldsymbol{\mu}, \boldsymbol{\Sigma}, \mathbf{Q}, \mathbf{F}\right]$$

Hyper-parameter

Goal: MAP estimate from $p\left(\boldsymbol{\theta}_{1:T} \mid \mathbf{y}_{1:T}, \mathbf{w}\right)$

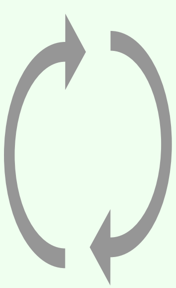
Firing rates and correlations ($\boldsymbol{\theta}_{1:T}$)

Observed spiking activities ($\mathbf{y}_{1:T}$)

Here, $\mathbf{w}$ is optimized so that it maximizes the (marginal) likelihood:

$$p\left(\mathbf{y}_{1:T} \mid \mathbf{w}\right) = \int p\left(\mathbf{y}_{1:T}, \boldsymbol{\theta}_{1:T} \mid \mathbf{w}\right) d\boldsymbol{\theta}_{1:T}.$$

Solution: EM-algorithm



E-step: Obtain $p\left(\boldsymbol{\theta}_{1:T} \mid \mathbf{y}_{1:T}, \mathbf{w}\right)$, given $\mathbf{w}$

via Recursive filtering/smoothing algorithm.

M-step: Optimize $\mathbf{w}$, given $p\left(\boldsymbol{\theta}_{1:T} \mid \mathbf{y}_{1:T}, \mathbf{w}\right)$

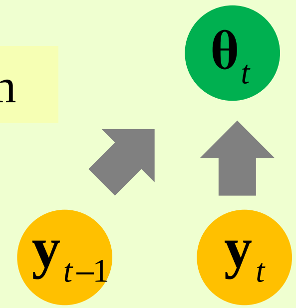
by maximizing the lower bound of evidence (Q-function).

Shimazaki H., Amari S., Brown E. N., and Gruen S. Proc. IEEE ICASSP2009

Nonlinear Recursive Filtering Algorithm

One-step prediction

$p(\boldsymbol{\theta}_t \mid \mathbf{y}_{1:t-1})$



Filter

$p(\boldsymbol{\theta}_t \mid \mathbf{y}_{1:t})$

Filter Distribution

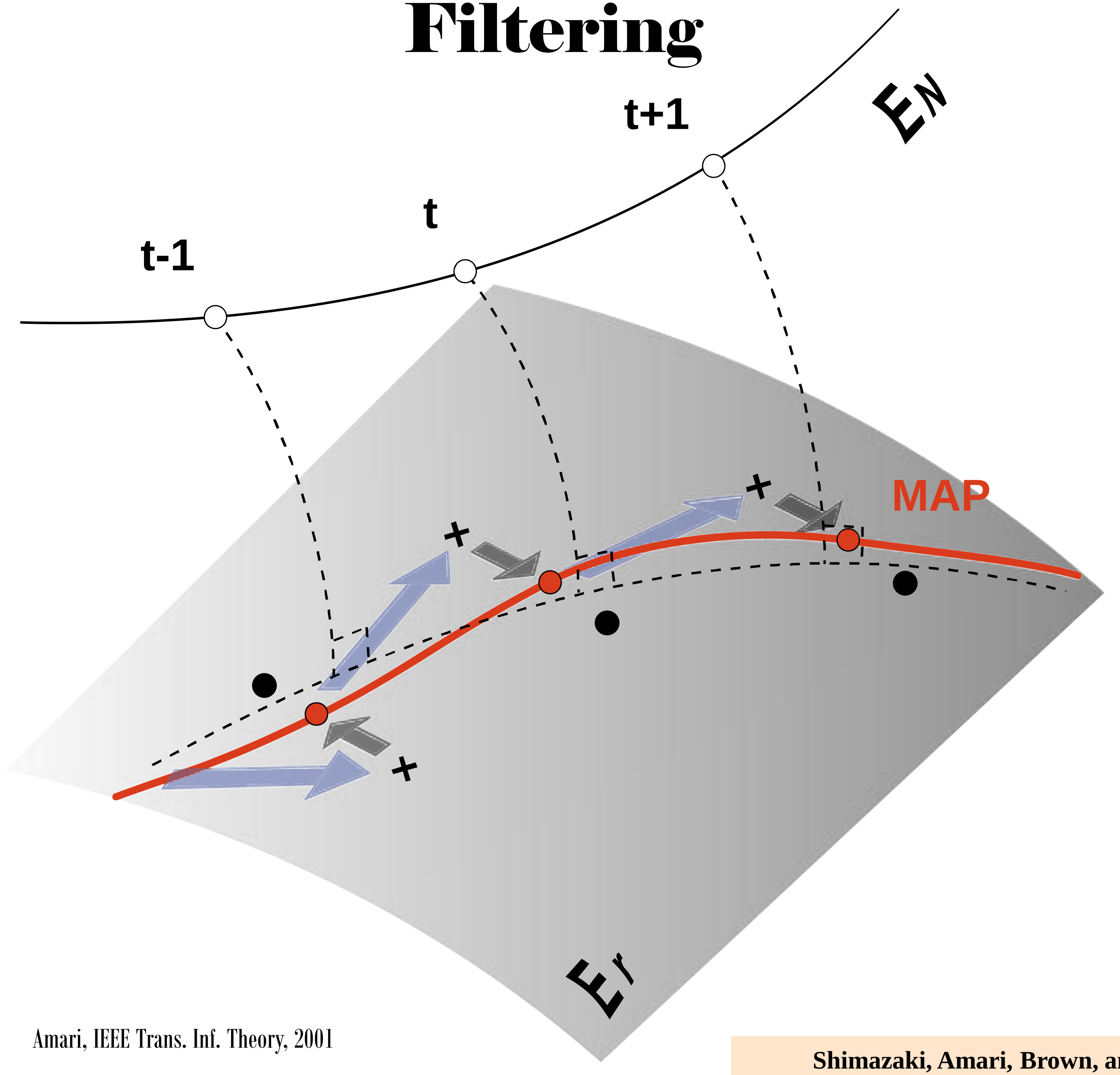
$$\underbrace{p(\boldsymbol{\theta}_t \mid \mathbf{y}_{1:t})}_{\text{Filter}} = \frac{\overbrace{p(\mathbf{y}_t \mid \boldsymbol{\theta}_t, \mathbf{y}_{1:t-1})}^{\text{Likelihood}}\ \overbrace{p(\boldsymbol{\theta}_t \mid \mathbf{y}_{1:t-1})}^{\text{Prior}}}{p(\mathbf{y}_t \mid \mathbf{y}_{1:t-1})} \qquad \text{(Bayes' rule)}$$

$$\propto \exp\left[\underbrace{n\left(\mathbf{y}_t^T \boldsymbol{\theta}_t - \psi_t\right)}_{\substack{\text{Observation at time } t \\ \text{(Log-linear distribution)}}} \underbrace{- \frac{1}{2}\left(\boldsymbol{\theta}_t - \boldsymbol{\theta}_{t|t-1}\right)^T W_{t|t-1}^{-1} \left(\boldsymbol{\theta}_t - \boldsymbol{\theta}_{t|t-1}\right)}_{\substack{\text{One-step prediction} \\ \text{(Normal distribution)}}} \right]$$

Approximate the posterior by the Gauss distribution (Laplace's method)

(ref. Smith & Brown, 2003)

Filtering



Amari, IEEE Trans. Inf. Theory, 2001

Shimazaki, Amari, Brown, and Gruen.

Laplace Approximation

Filter density $p\left(\boldsymbol{\theta}_t \mid \mathbf{y}_{1:t}\right) \propto \exp[n\left(\mathbf{y}_t^T \boldsymbol{\theta}_t - \psi\left(\boldsymbol{\theta}_t\right)\right) - \frac{1}{2}\left(\boldsymbol{\theta}_t - \boldsymbol{\theta}_{t|t-1}\right)^T W_{t|t-1}^{-1}\left(\boldsymbol{\theta}_t - \boldsymbol{\theta}_{t|t-1}\right)]$

gradient $\nabla f \equiv \nabla \log p\left(\boldsymbol{\theta}_t \mid \mathbf{y}_{1:t}\right) = -W_{t|t-1}^{-1}\left(\boldsymbol{\theta}_t - \boldsymbol{\theta}_{t|t-1}\right) + n\mathbf{y}_t - n\nabla\psi$

Hessian $H \equiv \nabla\nabla \log p\left(\boldsymbol{\theta}_t \mid \mathbf{y}_{1:t}\right) = -W_{t|t-1}^{-1} - n\nabla\nabla\psi$

Derivative of the log partition function yields cumulants:

Moment: $\partial_I \psi = E y_I = \eta_I$ $\qquad \nabla\psi = \boldsymbol{\eta}$

2nd order cumulant (Fisher metric): $\partial_I \partial_J \psi = E\left(y_I - E y_I\right) E\left(y_J - E y_J\right) = \eta_{I \cup J} - \eta_I \eta_J$ $\qquad \nabla\nabla\psi = G$

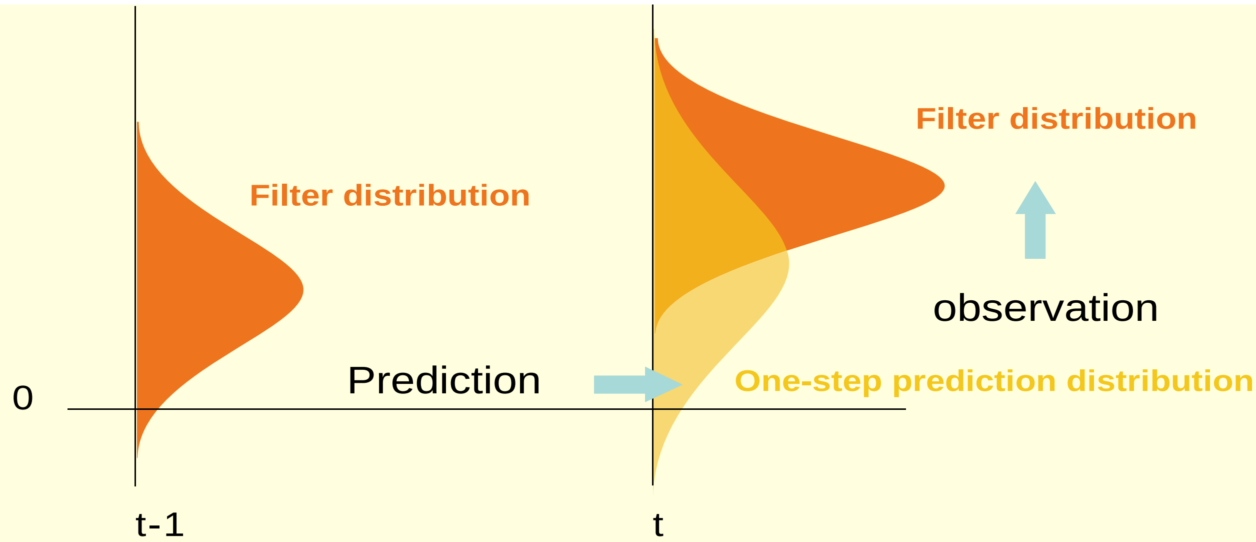
Filter Mode: Newton-Raphson method (Natural gradient method)

$$\boldsymbol{\theta}^{new} = \boldsymbol{\theta}^{old} + \varepsilon H^{-1} \nabla f$$

Filter Covariance: $-H^{-1} = -\left[-W_{t|t-1}^{-1} - nG\right]^{-1}$

Shimazaki H., Amari S., Brown E. N., and Gruen S. Proc. IEEE ICASSP2009

Summary of filtering algorithm



Recursive Nonlinear Filtering

One-step prediction

Dual coordinate to θ

Mean
$$\boldsymbol{\theta}_{t|t} = \boldsymbol{\theta}_{t|t-1} + nW_{t|t-1}\underbrace{\left[\mathbf{y}_t - \boldsymbol{\eta}_{t|t}\right]}_{\text{Observation error}}$$

Solved by
Newton-Raphson Method

Covariance
$$W_{t|t}^{-1} = W_{t|t-1}^{-1} + nG$$

One-step prediction

Fisher Information

Shimazaki H., Amari S., Brown E. N., and Gruen S. Proc. IEEE ICASSP2009

Filtering/Smoothing

One-step prediction

Shimazaki H., Amari S., Brown E. N., and Gruen S. State-space Analysis on Time-varying Correlations in Parallel Spike Sequences. Proc. IEEE ICASSP2009

Mean $$\boldsymbol{\theta}_{t|t-1} = \mathbf{F}\boldsymbol{\theta}_{t-1|t-1}$$

Covariance $$W_{t|t-1} = \mathbf{F} W_{t-1|t-1} \mathbf{F}^T + \mathbf{Q}$$

Recursive Nonlinear Filtering

One-step prediction

Dual coordinate to θ

Mean $$\boldsymbol{\theta}_{t|t} = \boldsymbol{\theta}_{t|t-1} + nW_{t|t-1} \underbrace{\left[\mathbf{y}_t - \boldsymbol{\eta}_{t|t} \right]}_{\text{Observation error}}$$

Solved by
Newton-Raphson Method

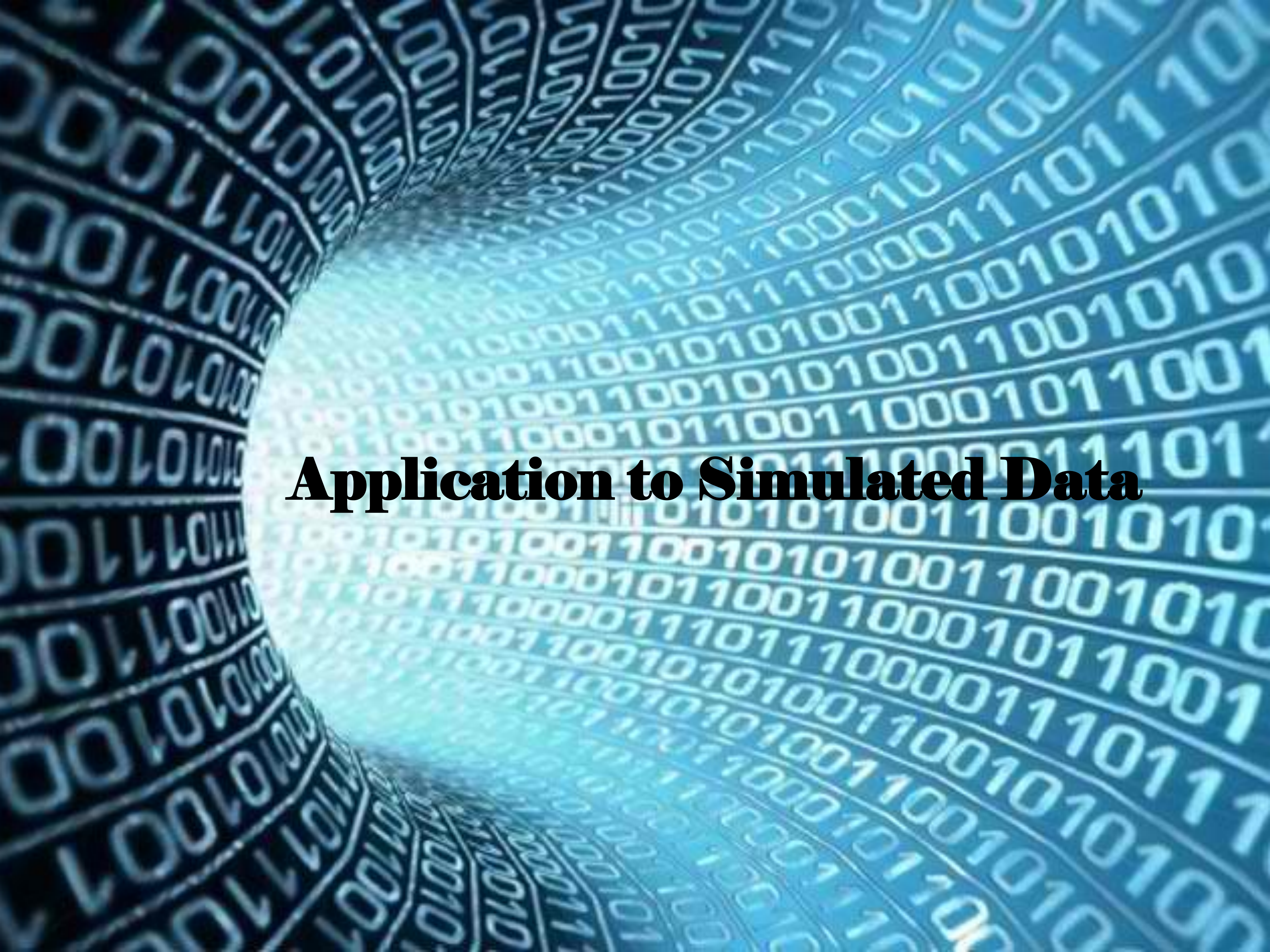
Covariance $$W_{t|t}^{-1} = W_{t|t-1}^{-1} + nG$$

One-step prediction **Fisher Information**

Fixed Interval Smoothing

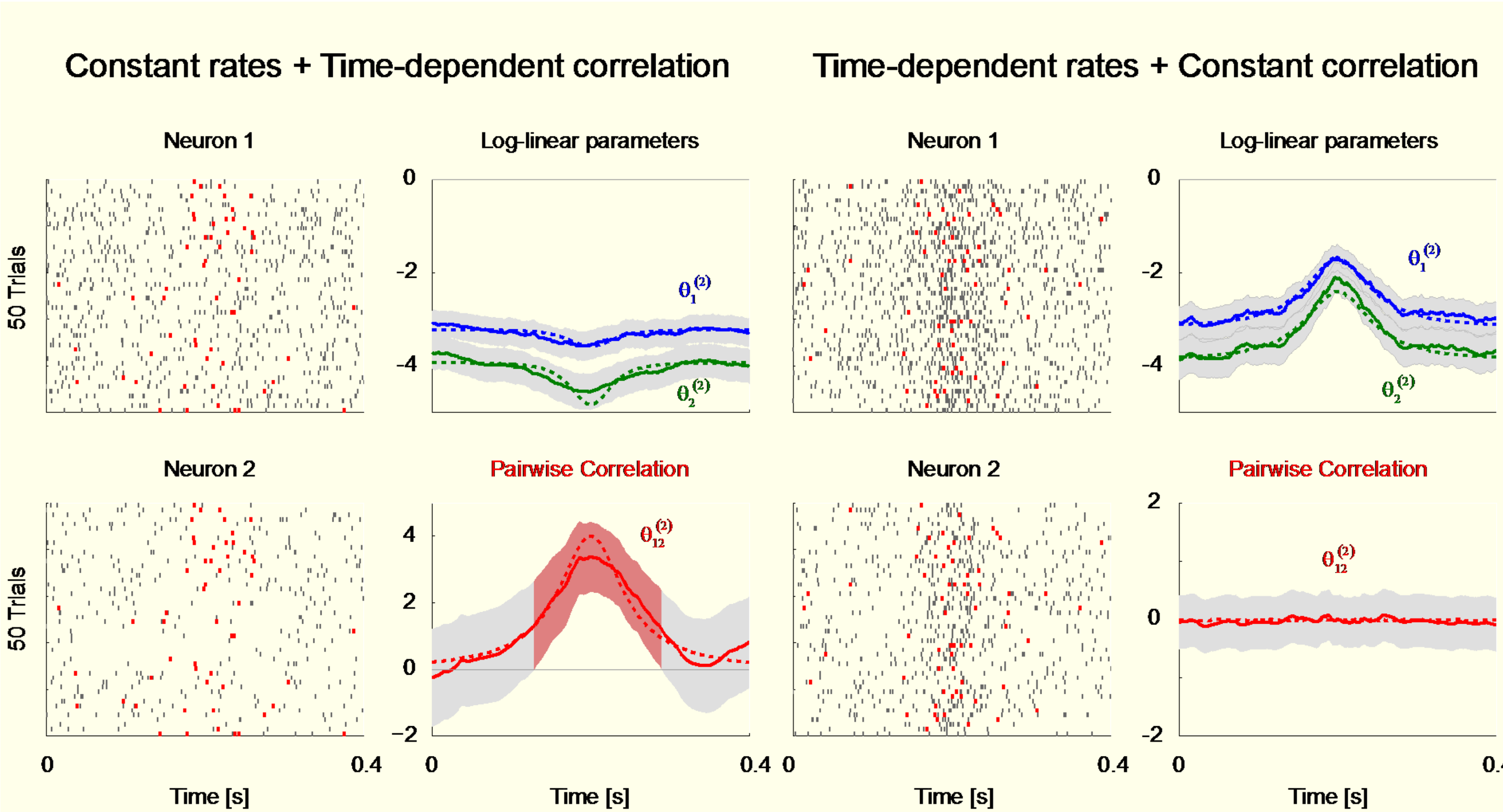
Mean $$\boldsymbol{\theta}_{t|K} = \boldsymbol{\theta}_{t|t} + A_t \left[\boldsymbol{\theta}_{t+1|K} - \boldsymbol{\theta}_{t+1|t} \right] \qquad A_t = W_{t|t} F^T W_{t+1|t}^{-1}$$

Covariance $$W_{t|K} = W_{t|t} + A_t \left[W_{t+1|K} - W_{t+1|t} \right] A_t^T$$



Application to Simulated Data

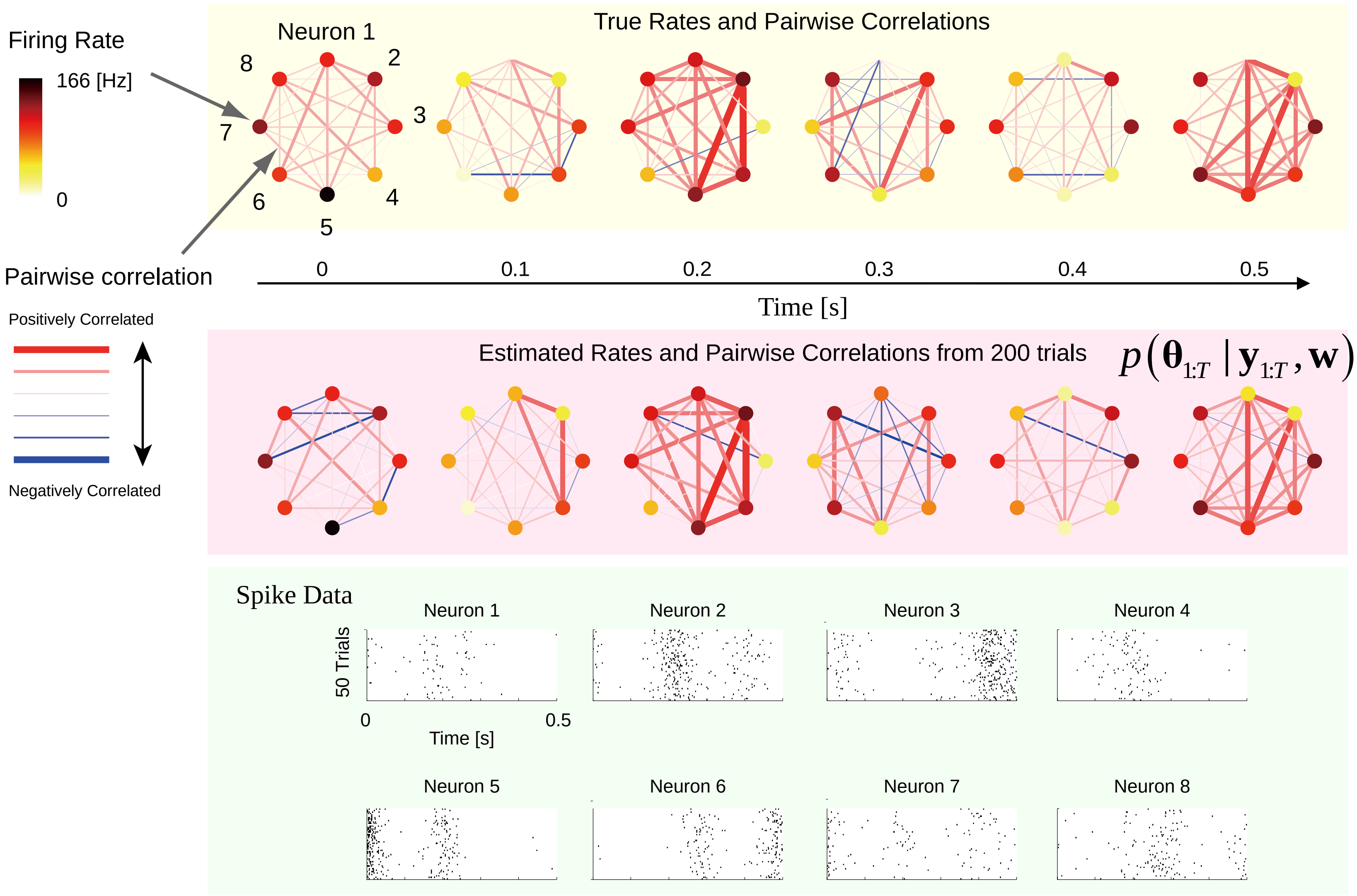
Results: 2 Neurons (Simulated)



Synchrony between two neurons is resolved into either a pairwise correlation or a chance coincidence.

Shimazaki, Amari, Brown, Gruen, (2012) PLoS Comp Biol, 8(3): e1002385

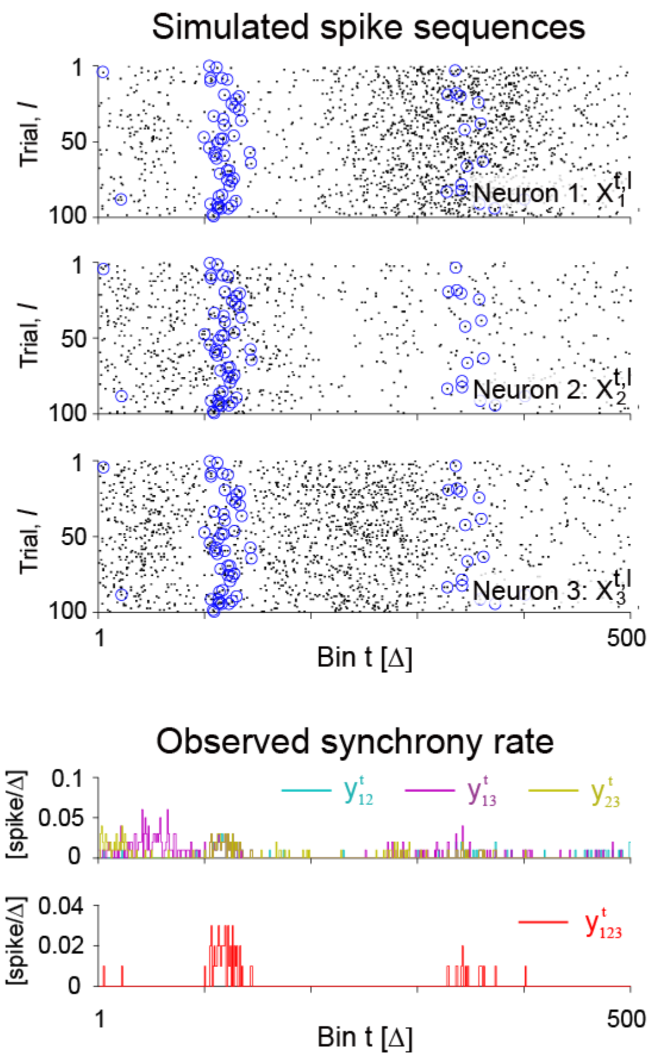
Results: Simultaneous Pairwise Analysis



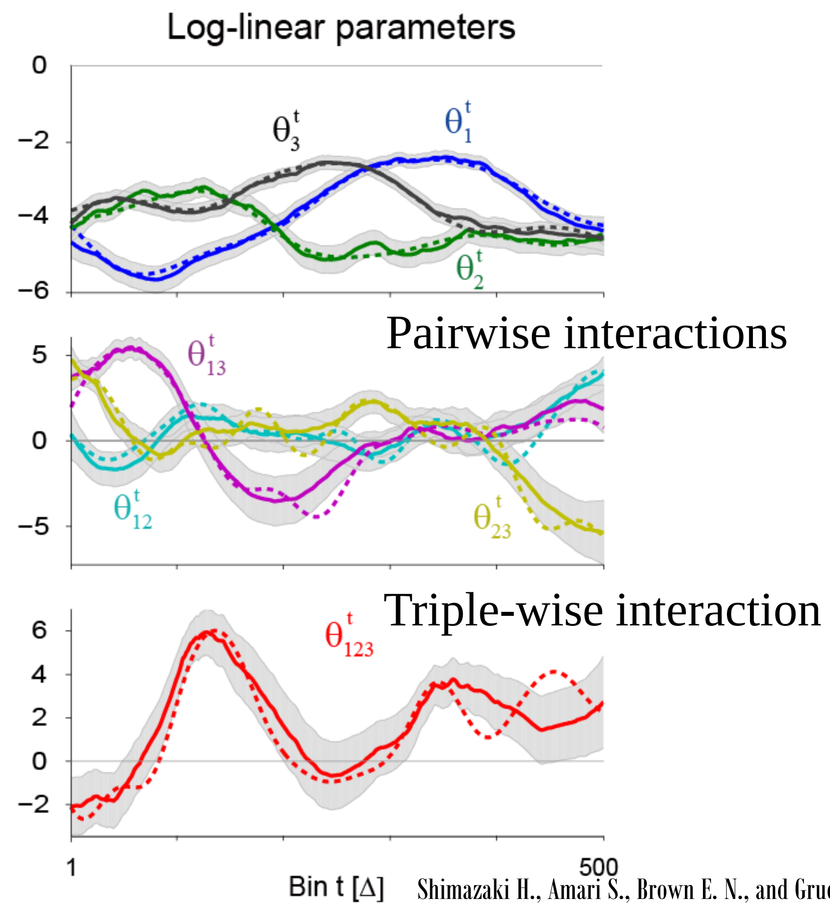
Simultaneous analysis resolves the pairwise dependency of 8 neurons.

Shimazaki, Amari, Brown, Gruen, (2012) PLoS Comp Biol, 8(3): e1002385

Results: Triple-wise interaction



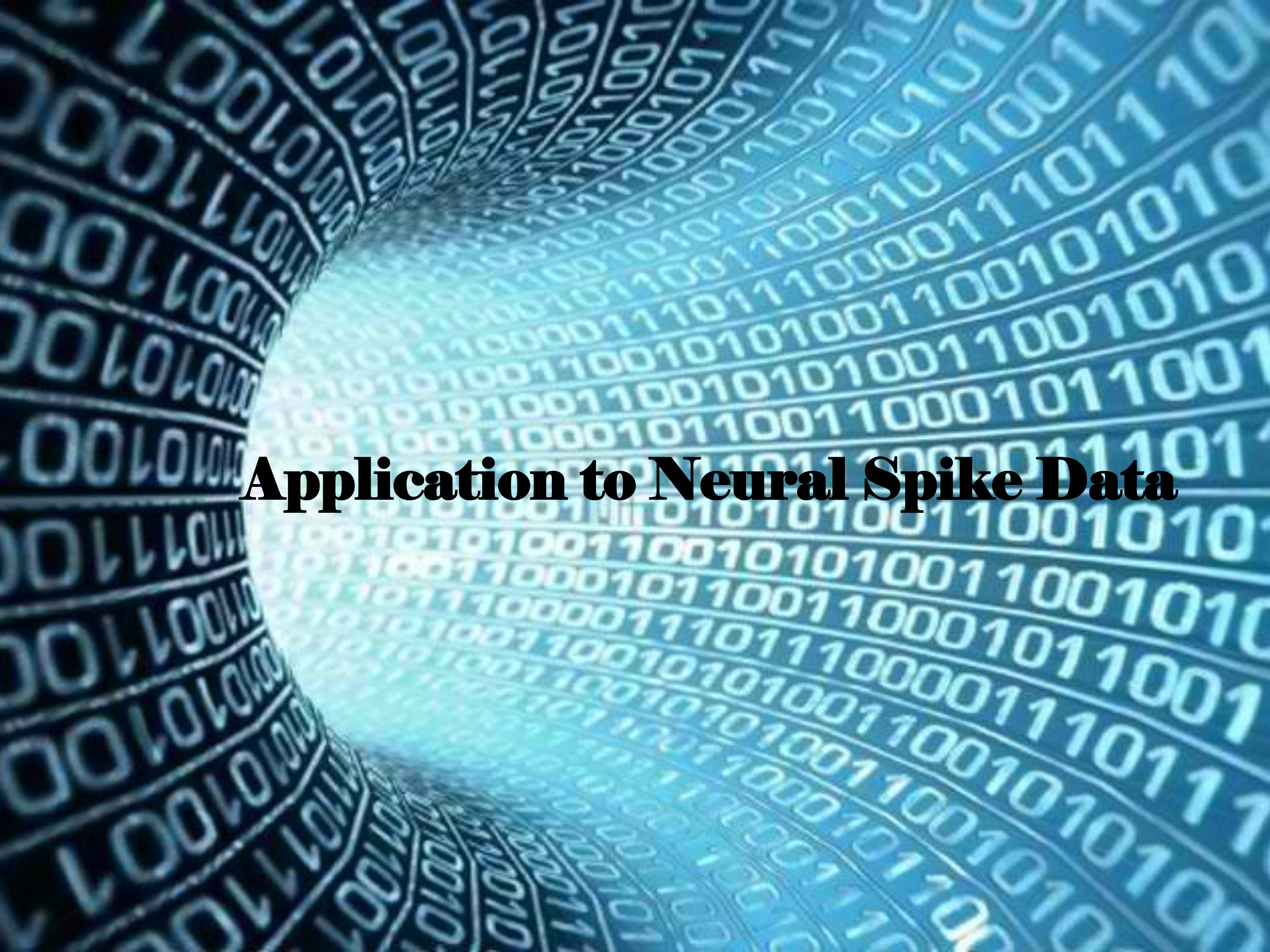
$$p\left(\boldsymbol{\theta}_{1:T} \mid \mathbf{y}_{1:T}, \mathbf{w}\right)$$



Shimazaki H., Amari S., Brown E. N., and Gruen S. State-space Analysis on Time-varying Correlations in Parallel Spike Sequences. Proc. IEEE ICASSP2009

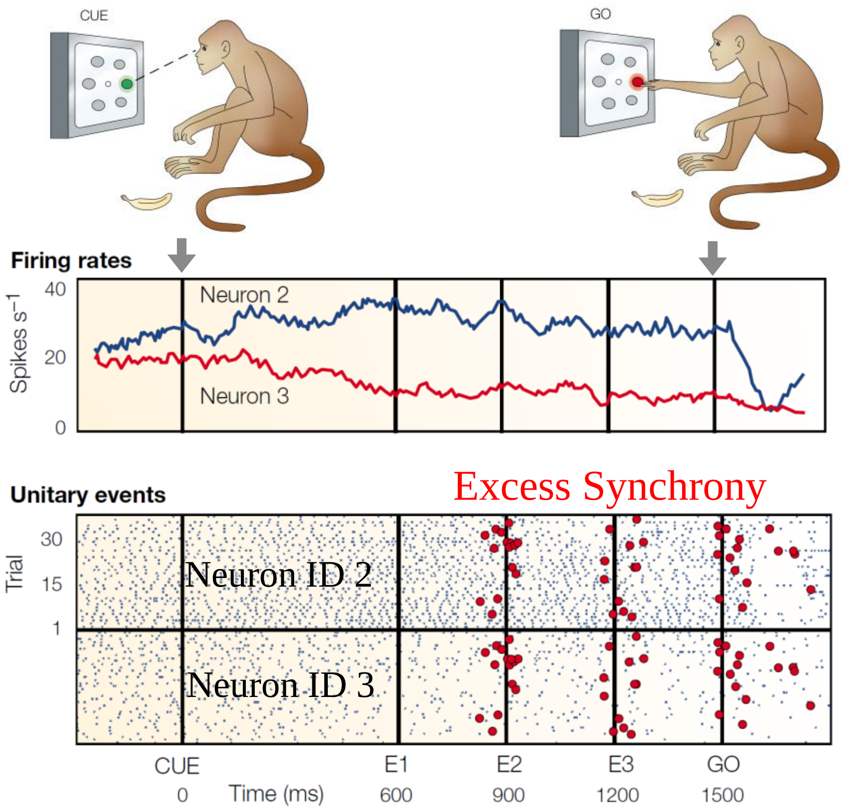
Time-dependent triple-wise interaction is estimated under time-varying firing rates and pairwise correlations.

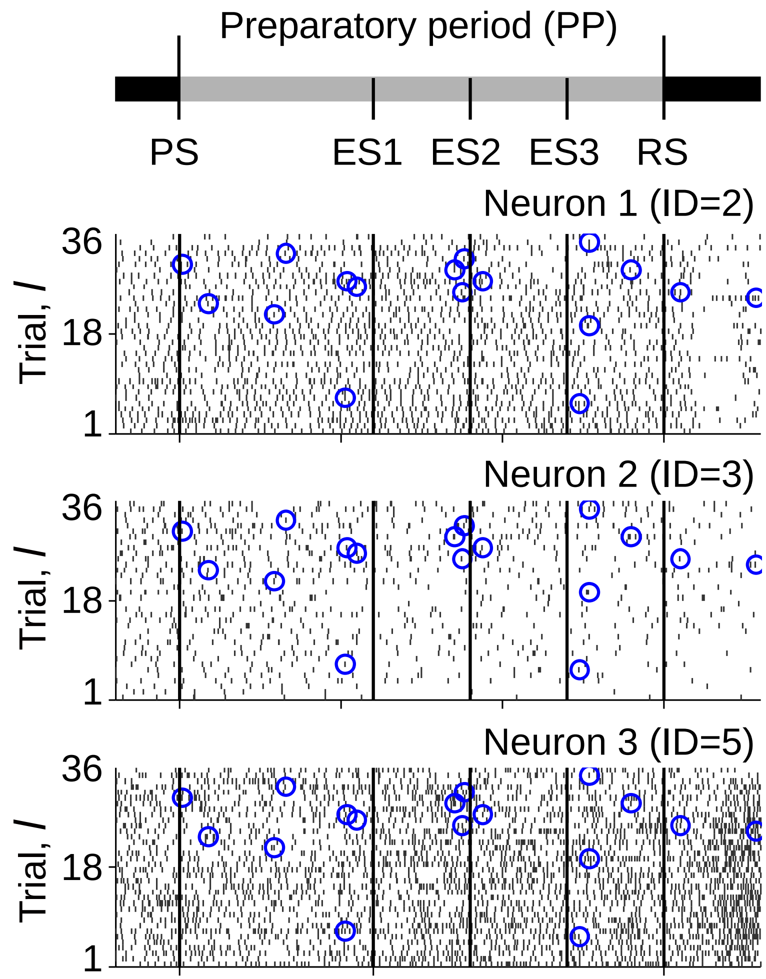
Shimazaki, Amari, Brown, Gruen, (2012) PLoS Comp Biol, 8(3): e1002385



Application to Neural Spike Data

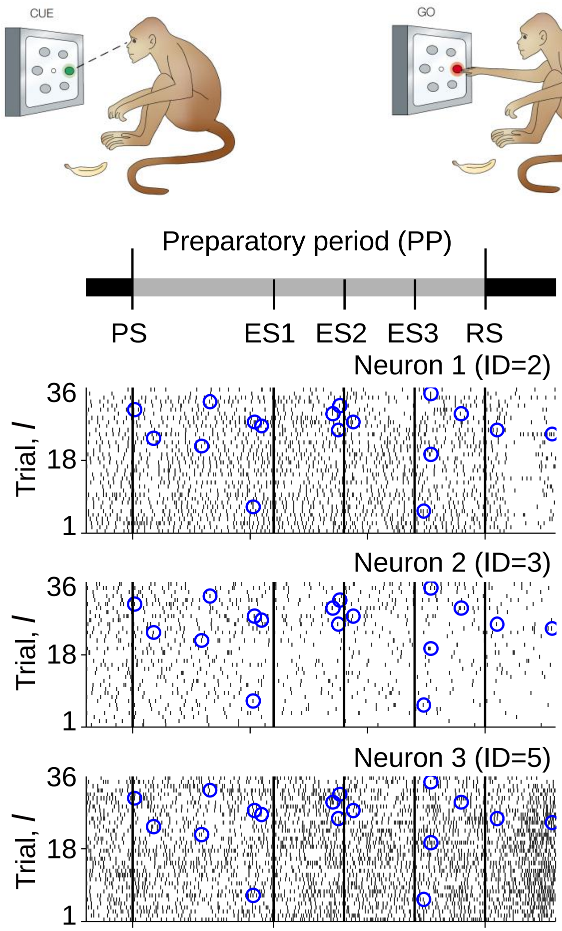
Application to Three MI neurons



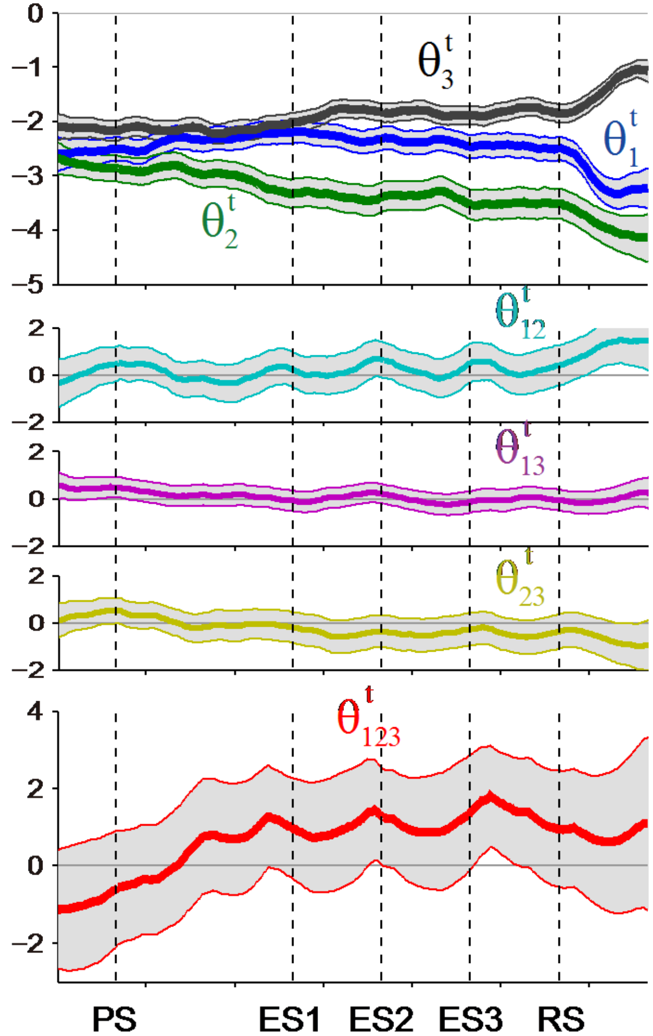


Riehle, Grün, Diesmann, Aertsen
Science 278: 1950-1953, 1997.

Application to Three MI neurons

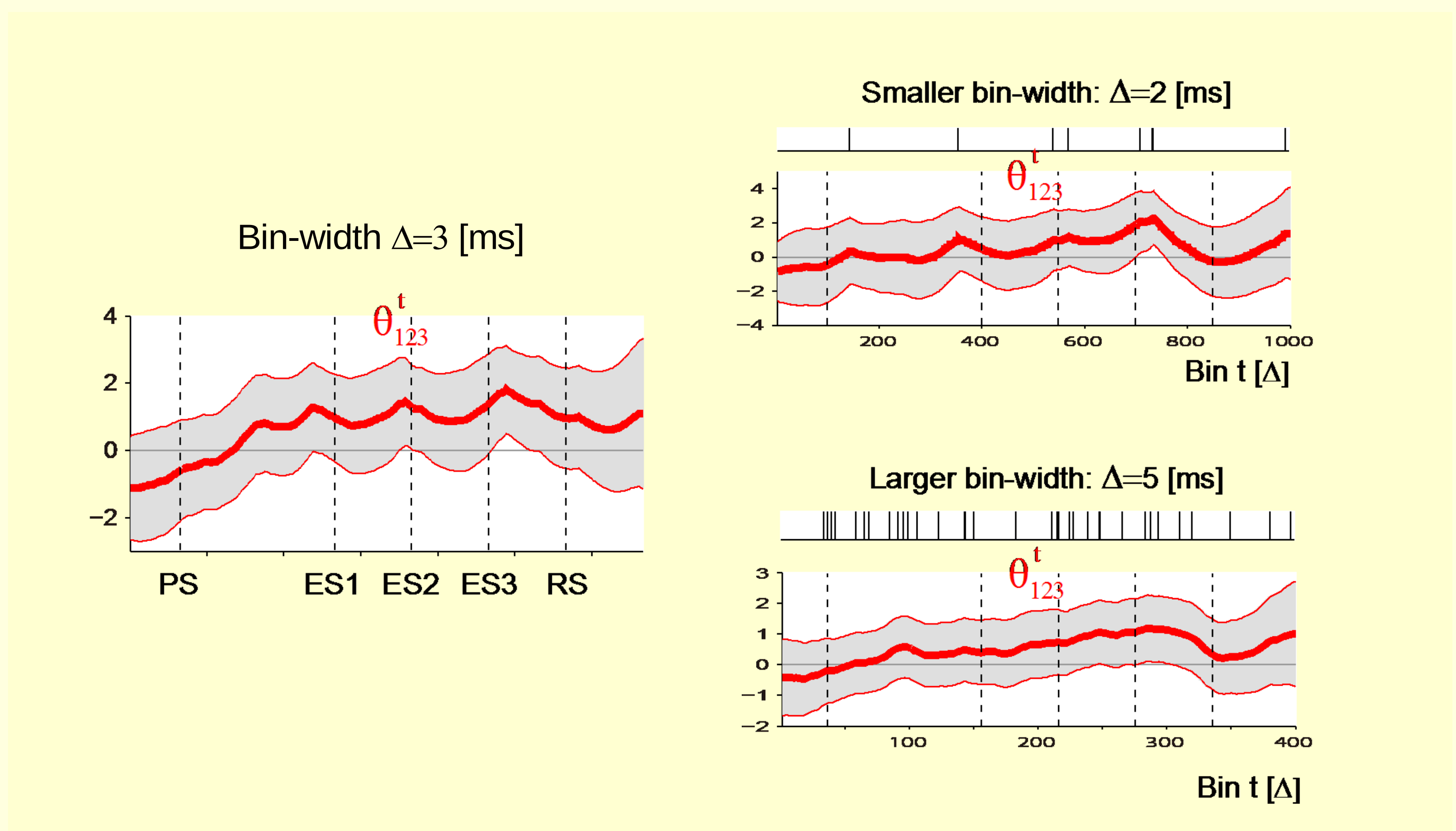


Riehle, Grün, Diesmann, Aertsen
Science 278: 1950-1953, 1997.



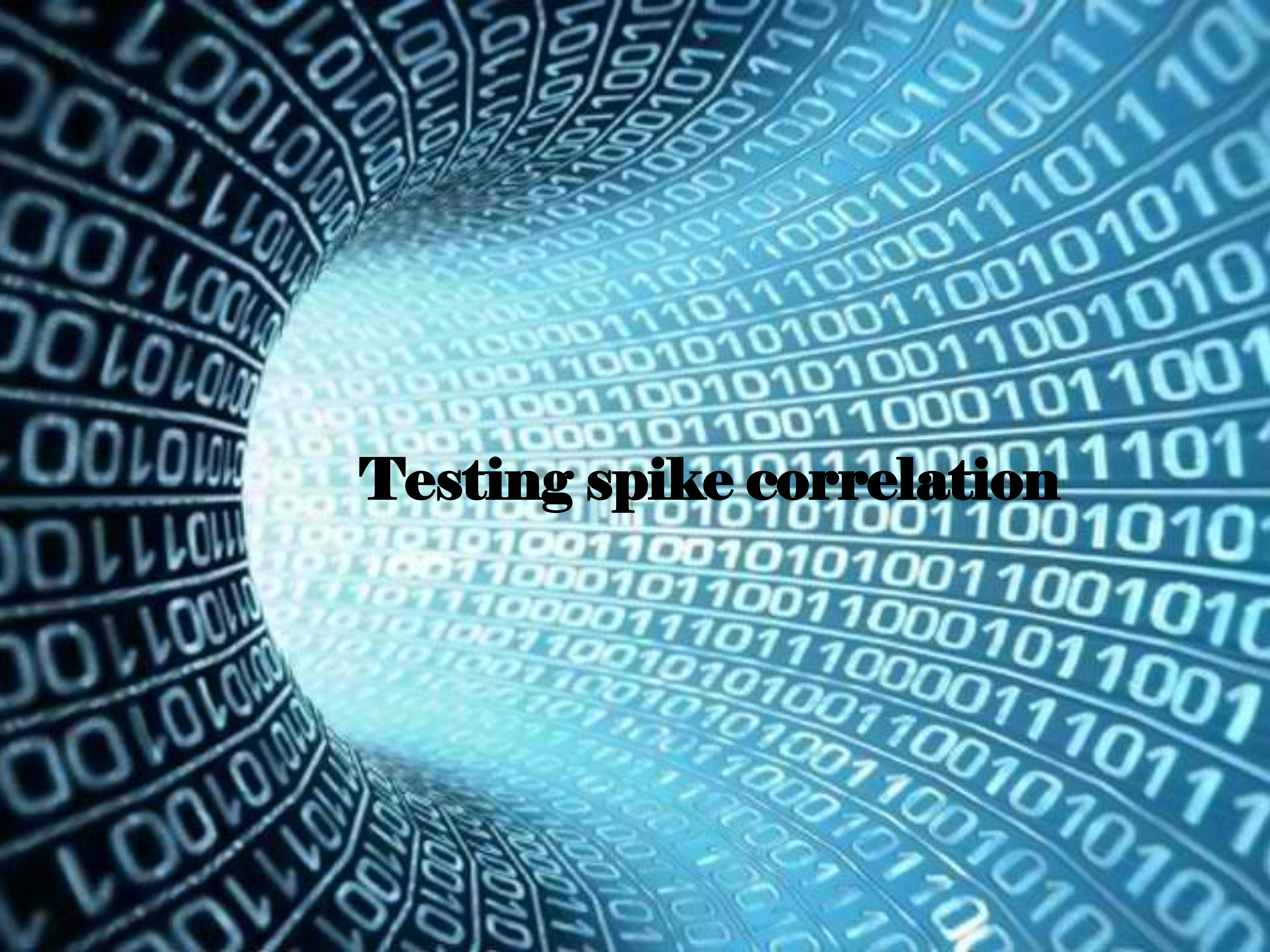
Shimazaki, Amari, Brown, Gruen, (2012) PLoS Comp Biol, 8(3): e1002385

Bin-width dependence



Gradual increase of a triple-wise interaction during the PP is observed in larger bin-width (5 ms).

Shimazaki, Amari, Brown, Gruen, (2012) PLoS Comp Biol, 8(3): e1002385



Testing spike correlation

Testing spike correlations in nonstationary data

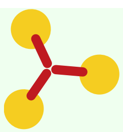
Marginal likelihood ratio (Bayes factor) (Jeffrey 61, Kass&Raftery 95)

$$B_{12}\left(\mathbf{y}_{a:b}\right)=\frac{p\left(\mathbf{y}_{a:b}\mid M_1\right)}{p\left(\mathbf{y}_{a:b}\mid M_2\right)}=\prod_{t=a}^{b}\boxed{\frac{\int_{S_1} p\left(\boldsymbol{\theta}_t\mid \mathbf{y}_{1:t}\right)d\boldsymbol{\theta}_t}{\int_{S_2} p\left(\boldsymbol{\theta}_t\mid \mathbf{y}_{1:t}\right)d\boldsymbol{\theta}_t}}\Bigg/\boxed{\frac{\int_{S_1} p\left(\boldsymbol{\theta}_t\mid \mathbf{y}_{1:t-1}\right)d\boldsymbol{\theta}_t}{\int_{S_2} p\left(\boldsymbol{\theta}_t\mid \mathbf{y}_{1:t-1}\right)d\boldsymbol{\theta}_t}}$$

Filter Odds

Prediction Odds

Models for latent signals
(spike interactions)



Test for a triple-wise interaction

Use a triple-wise model:

$$\log p\left(\mathbf{x}\mid\boldsymbol{\theta}\right)=\sum_i \theta_i x_i+\sum_{i<j}\theta_{ij}x_i x_j+\theta_{123}x_1x_2x_3-\psi$$

$$M_1:\theta_{123}>0 \quad M_2:\theta_{123}\le 0$$

$$\theta_i,\theta_{ij}\in\Re \text{ for } i,j=1,2,3$$

$B_{12}\left(\mathbf{y}_{a:b}\right)>1$ supports Model 1.

$B_{12}\left(\mathbf{y}_{a:b}\right)\le 1$ supports Model 2.

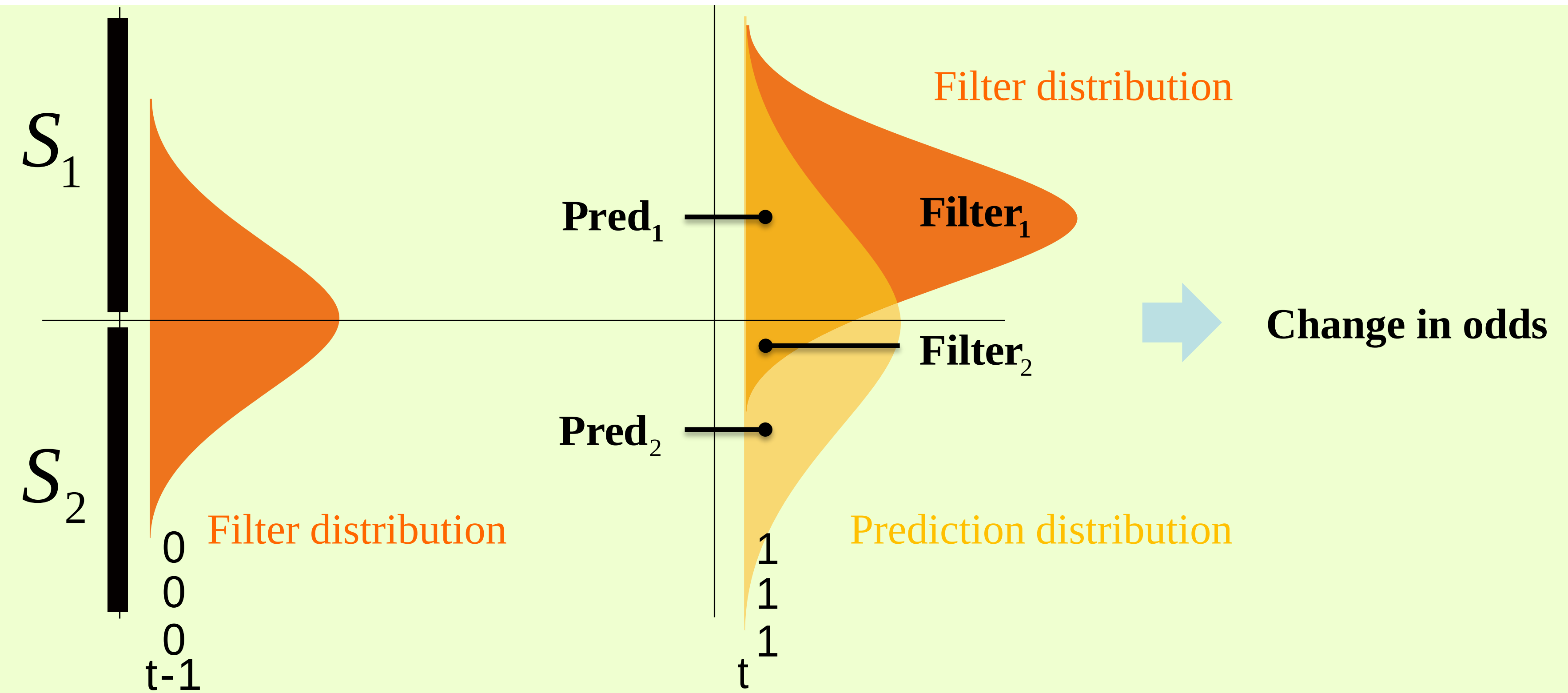
Weight of Evidence (Good, 1985)

$\log_2 B_{12}\left(\mathbf{y}_{a:b}\right)>0$ supports Model 1.

$\log_2 B_{12}\left(\mathbf{y}_{a:b}\right)\le 0$ supports Model 2.

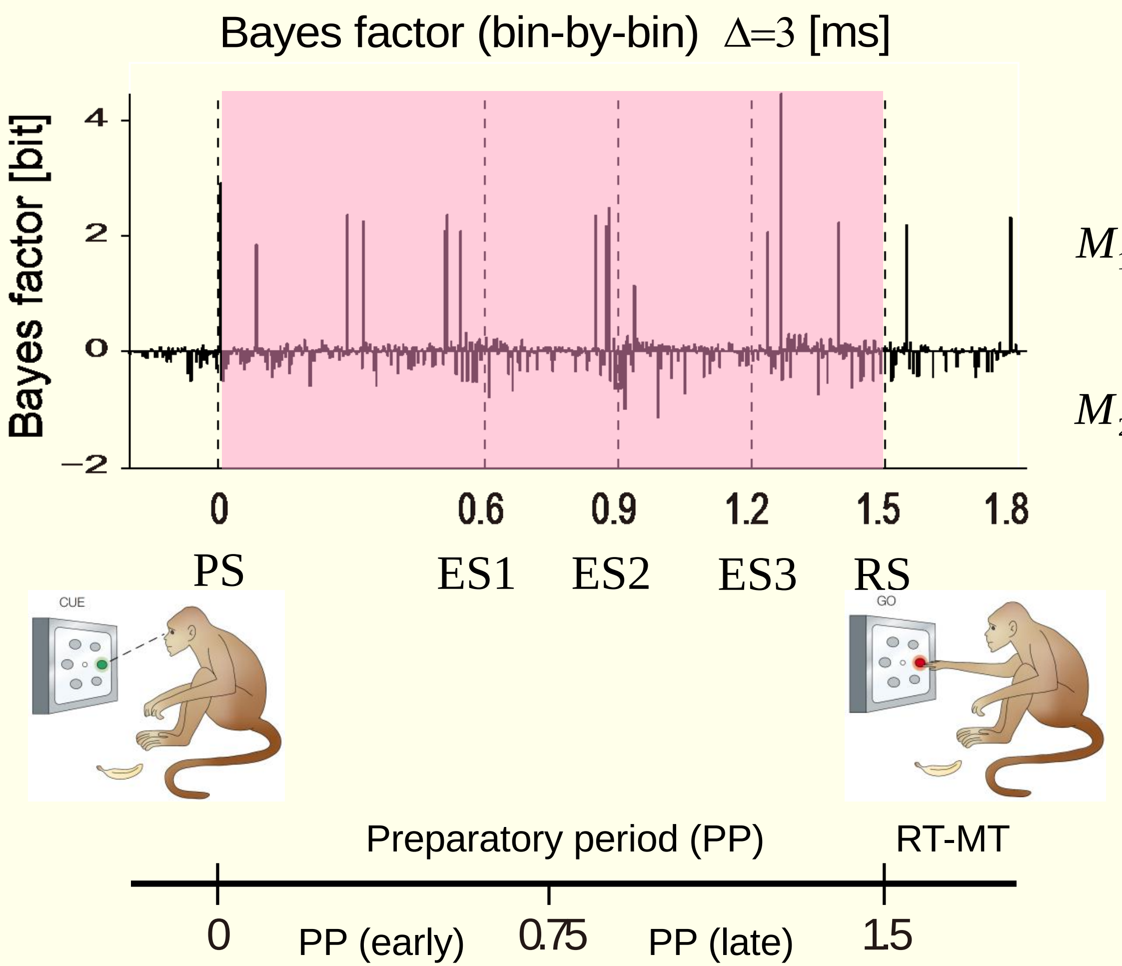
How to compute the weight of evidence

$$\log_2 B_{12}(\mathbf{y}_{a:b}) = \sum_{t=a}^{b} \log_2 \underbrace{\boxed{\frac{\int_{S_1} p(\boldsymbol{\theta}_t \mid \mathbf{y}_{1:t})\, d\boldsymbol{\theta}_t}{\int_{S_2} p(\boldsymbol{\theta}_t \mid \mathbf{y}_{1:t})\, d\boldsymbol{\theta}_t}}}_{\text{Filter Odds}} \Bigg/ \underbrace{\boxed{\frac{\int_{S_1} p(\boldsymbol{\theta}_t \mid \mathbf{y}_{1:t-1})\, d\boldsymbol{\theta}_t}{\int_{S_2} p(\boldsymbol{\theta}_t \mid \mathbf{y}_{1:t-1})\, d\boldsymbol{\theta}_t}}}_{\text{Prediction Odds}}$$



Change in odds by observing a spike pattern leads to the BF larger or smaller than 1.

Sequential computation of the BF during preparation period



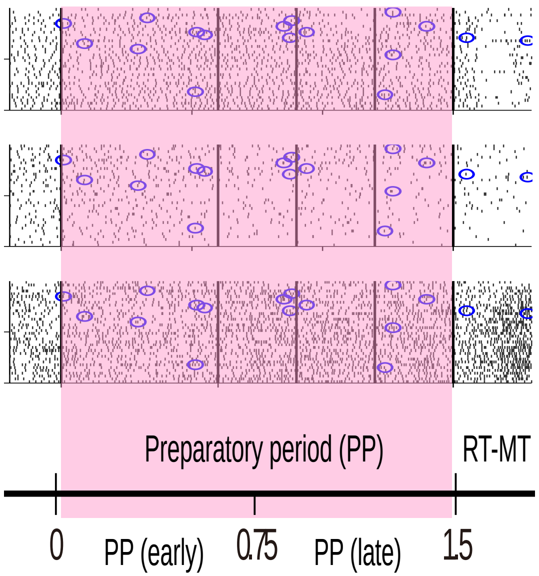
$M_1 : \theta_{123} > 0$

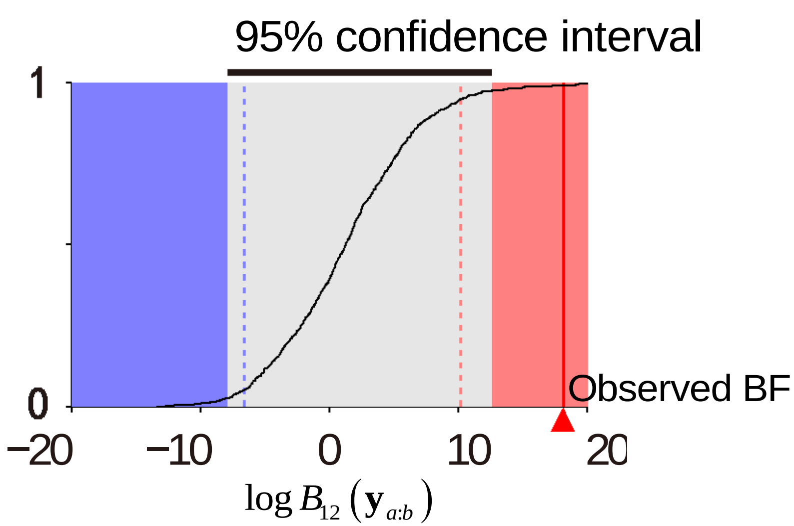
$M_2 : \theta_{123} \leq 0$

The weight of evidence during the preparatory period is 18.08 [bit].

Shimazaki, Amari, Brown, Gruen, (2012) PLoS Comp Biol, 8(3): e1002385

Testing a triple-wise spike correlation





1. Select a behaviorally relevant period.

2. Apply a full LLSS. Compute the BF.

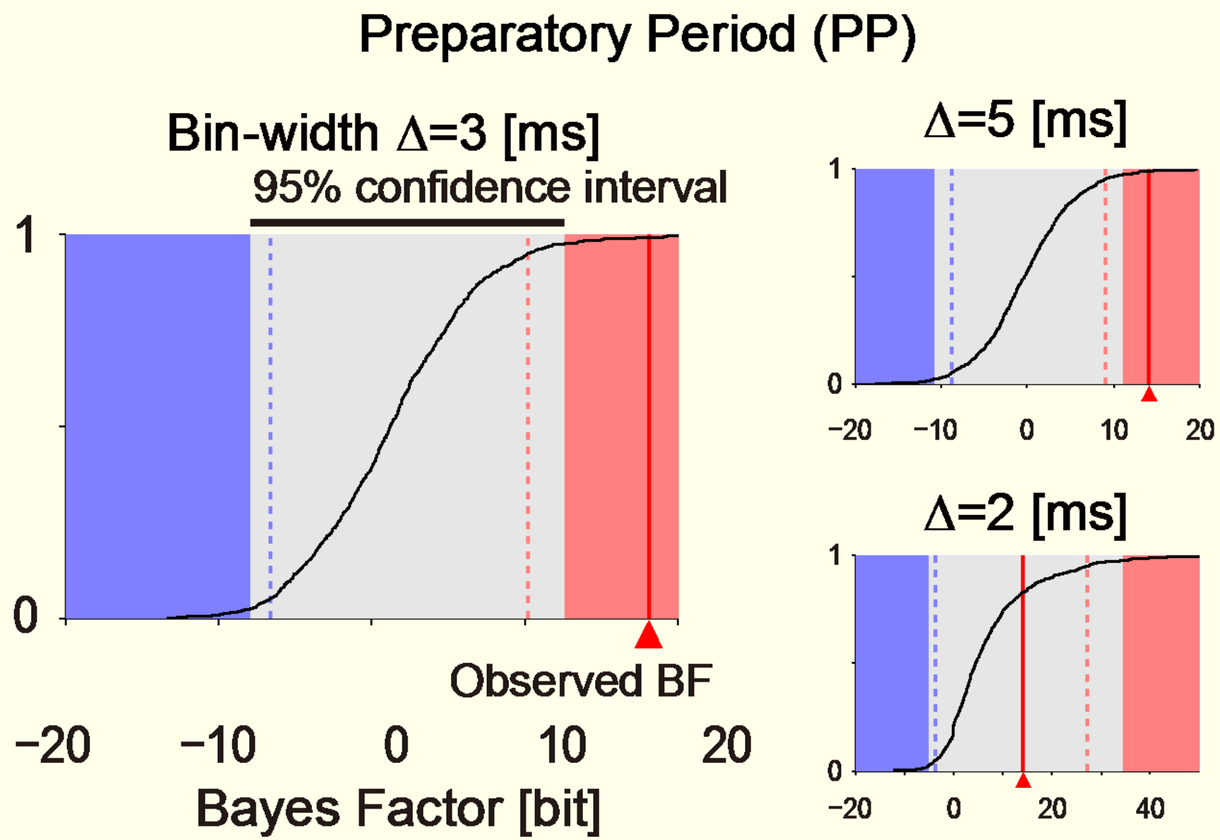
3. Compute the surrogate BF for a triple-wise correlation.

   The surrogate BF was generated under the null hypothesis of no triple-wise correlation:
   - Fit a pairwise LLSS model to the data.
   - Resample spikes from the fitted model.
   - Compute the BF of the resampled data.

4. Obtain CDF of surrogate BFs.

5. Find 95% confidence bound, and compare it with an empirical BF.

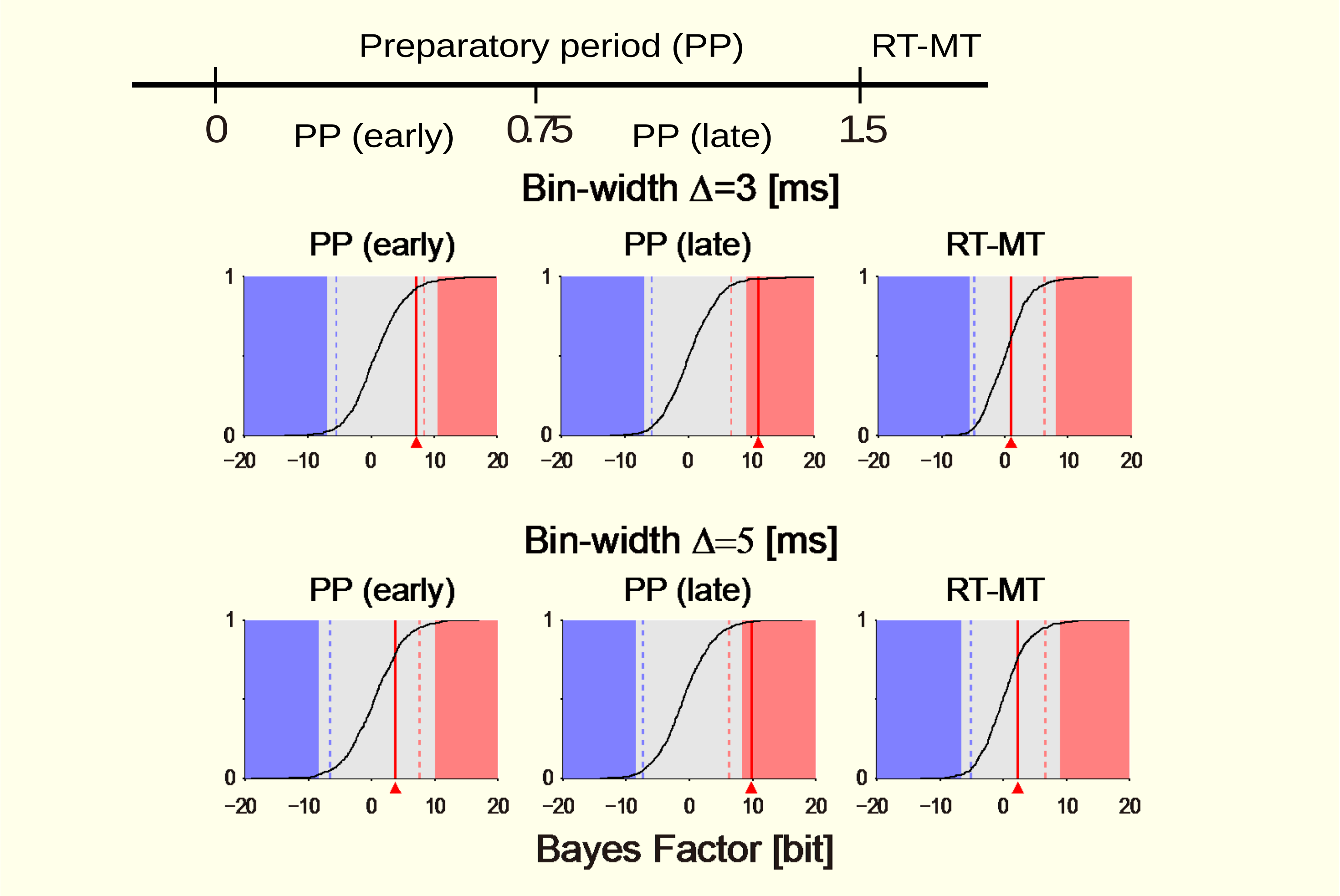
Bin-width dependence



The data supports presence of excess synchronous spiking activities of 3 and 5 ms precision during the PP that can not be explained by pairwise correlations.

Shimazaki, Amari, Brown, Gruen, (2012) PLoS Comp Biol, 8(3): e1002385

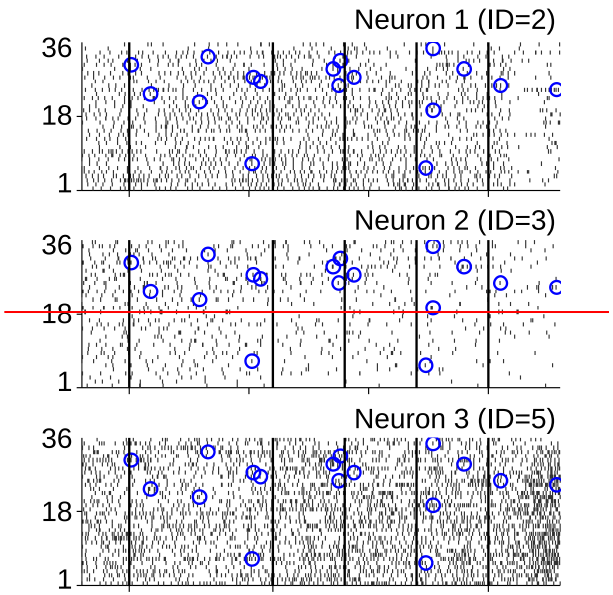
Detailed analysis

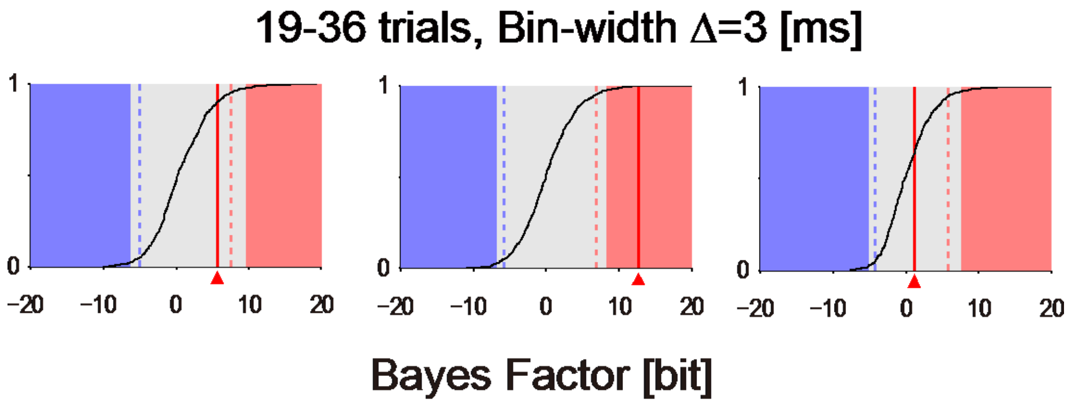


Triple-wise spike correlation was detected in late stage of the PP.

Shimazaki, Amari, Brown, Gruen, (2012) PLoS Comp Biol, 8(3): e1002385

On the across-trial nonstationarity





Our results indicate that the three neurons that expresses simultaneous spikes with a precision of 3 or 5 ms are involved in a single functional assembly that can not be explained by pair interactions, and that this assembly is dynamically organized during the PP, especially in the later sessions of an experiment.

Shimazaki, Amari, Brown, Gruen, (2012) PLoS Comp Biol, 8(3): e1002385

Summary

Toward detection of dynamic assembly activities in relation to behavior, we

- Constructed state-space model of time-varying higher-order spike correlations

- (Validated the inclusion of higher-order correlation by model selection.)

- Tested the evidence of the higher-order spike correlation.

- Applied the methods to small subset neurons from an awake monkey, and demonstrated that the higher-order correlation organizes in relation to behavior.

Acknowledgement:
We thank B. Staude, S. Koyama, Z. Chen, D. Ba, K. Sadeghi, and T. Toyoizumi for helpful discussion. We also thank Diesmann Unit for computing environment.