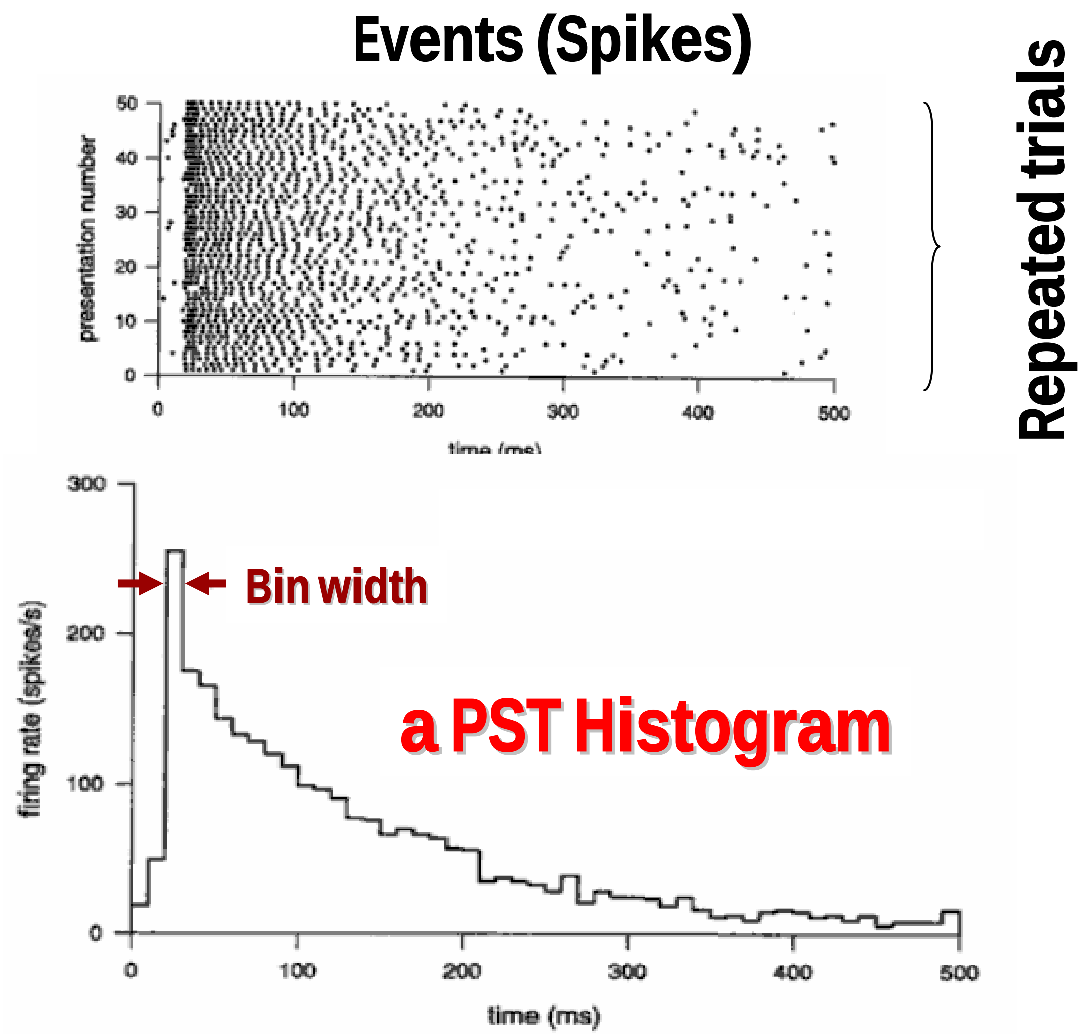
A Recipe for Constructing a Peri-stimulus Time Histogram

Hideaki Shimazaki
Dept. of Physics Kyoto University, Japan

March 1, 2007 Johns Hopkins University
March 2, 2007 Columbia University

Peri-stimulus Time Histogram

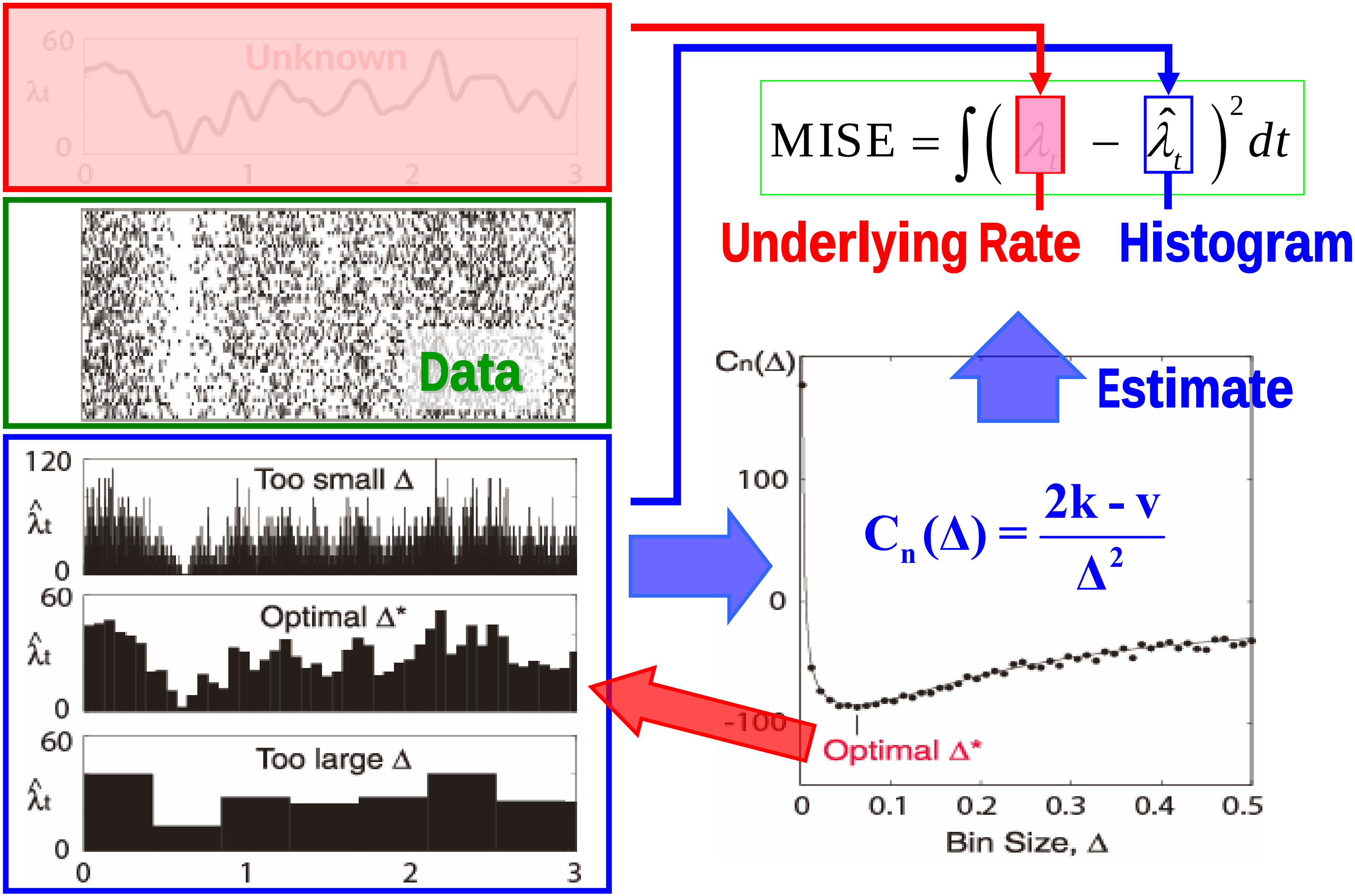


From *Spikes: Exploring the Neural Code*, Rieke et al. 1997

Adrian, E. (1928). *The basis of sensation: The action of the sense organs*.

George L. Gerstein and Nelson Y.-S. Kiang (1960)
An Approach to the Quantitative Analysis of Electrophysiological Data from Single Neurons, Biophys J. 1(1): 15–28.

1. Bin Width Selection



Method for Selecting the Bin Size

- Divide the data range into *N* bins of width *Δ*. Count the number of events *ki* in the *i* th bin.

- Compute the cost function

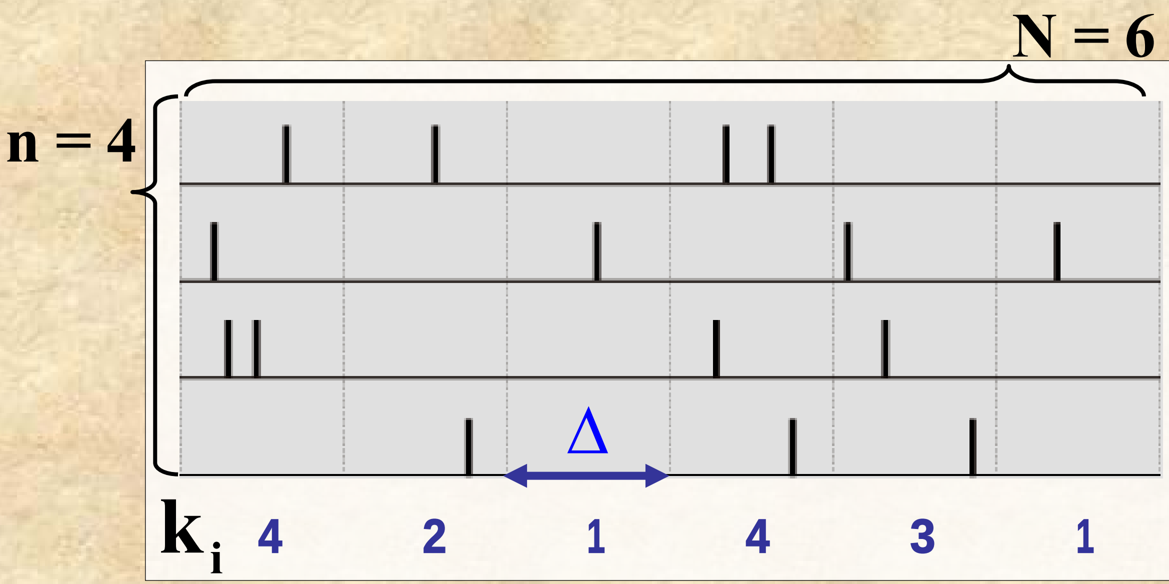
$$C_n(\Delta) = \frac{2k - v}{(n\Delta)^2},$$

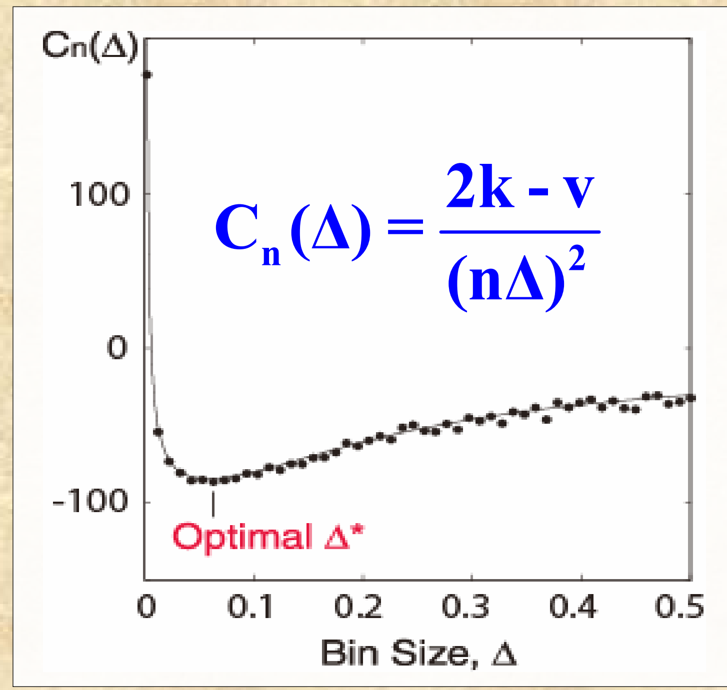
while changing the bin size *Δ*.

- Find *Δ** that minimize the cost function.

Mean $\quad k = \frac{1}{N}\sum_{i=1}^{N} k_i,$

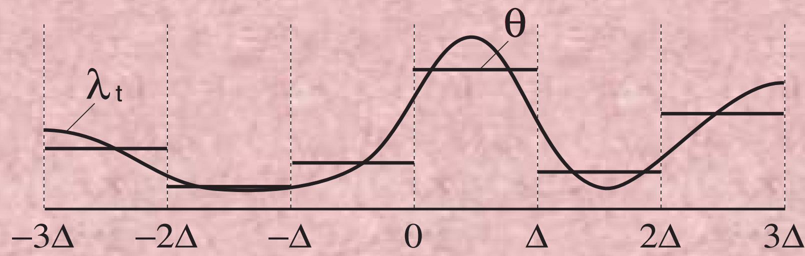
Variance $\quad v = \frac{1}{N}\sum_{i=1}^{N} (k_i - k)^2$



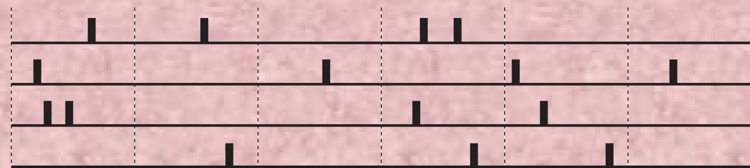


Theory for the Method

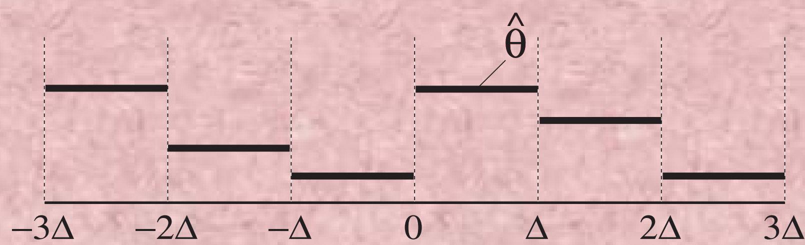
Time-Varying Rate



Spike Sequences



Time Histogram



The mean underlying rate in an interval [0, Δ]:

$$\theta = \frac{1}{\Delta}\int_0^{\Delta} \lambda_t \, dt.$$

The spike count in the bin obeys the Poisson distribution*:

$$p(k \mid n\Delta\theta) = \frac{\left(n\Delta\theta\right)^k}{k!} e^{-n\Delta\theta}.$$

A histogram bar-height is an estimator of θ:

$$\hat{\theta}_n = \frac{k}{n\Delta}$$

*When the spikes are obtained by repeating an independent trial, the accumulated data obeys the Poisson point process due to a general limit theorem.

Theory: Selection of the Bin Size

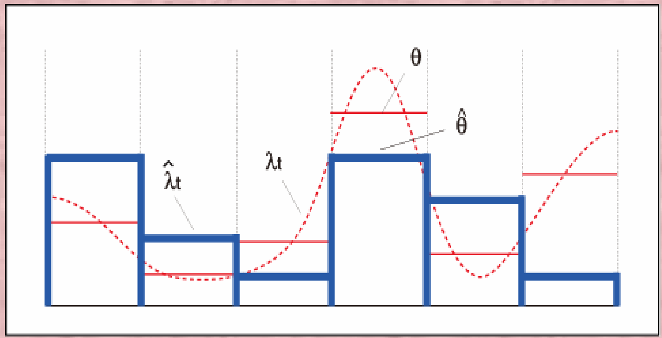
$$\mathrm{MISE} \equiv \frac{1}{T}\int_0^T E(\hat{\lambda}_t - \lambda_t)^2 dt = \left\langle \frac{1}{\Delta}\int_0^\Delta E(\underbrace{\hat{\theta}_n - \lambda_t}_{\theta - \theta})^2\, dt \right\rangle.$$

Expectation by the Poisson statistics, given the rate.

Average over segmented bins.

$$\mathrm{MISE} = \left\langle E(\hat{\theta}_n - \theta)^2 \right\rangle + \frac{1}{\Delta}\int_0^\Delta \left\langle (\lambda_t - \theta)^2 \right\rangle dt.$$

Sampling Error **Systematic Error**



Decomposition of the Systematic Error

$$\frac{1}{\Delta}\int_0^\Delta \left\langle (\overbrace{\lambda_t - \theta}^{\langle\theta\rangle - \langle\theta\rangle})^2 \right\rangle dt = \frac{1}{\Delta}\int_0^\Delta \left\langle (\lambda_t - \langle\theta\rangle)^2 \right\rangle dt - \left\langle (\theta - \langle\theta\rangle)^2 \right\rangle$$

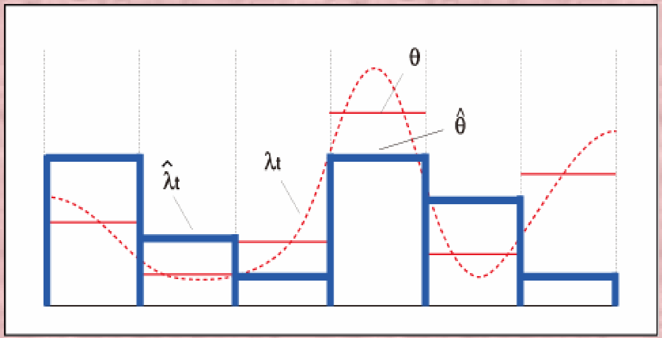
Systematic Error

Variance of the rate

Independent of Δ

Variance of an ideal histogram

Introduction of the cost function:



$$C_n(\Delta) \equiv \mathrm{MISE} - \frac{1}{T}\int_0^T \left(\lambda_t - \langle\theta\rangle\right)^2 dt$$

$$= \left\langle E(\hat{\theta}_n - \theta)^2 \right\rangle - \left\langle \left(\theta - \langle\theta\rangle\right)^2 \right\rangle.$$

Sampling error Unknown: Variance of an ideal histogram

The variance decomposition: $\left\langle E\left(\hat{\theta}_n - \left\langle E\hat{\theta}_n \right\rangle\right)^2 \right\rangle = \left\langle E(\hat{\theta}_n - \theta)^2 \right\rangle + \left\langle \left(\theta - \langle\theta\rangle\right)^2 \right\rangle.$

Variance of a histogram Sampling error

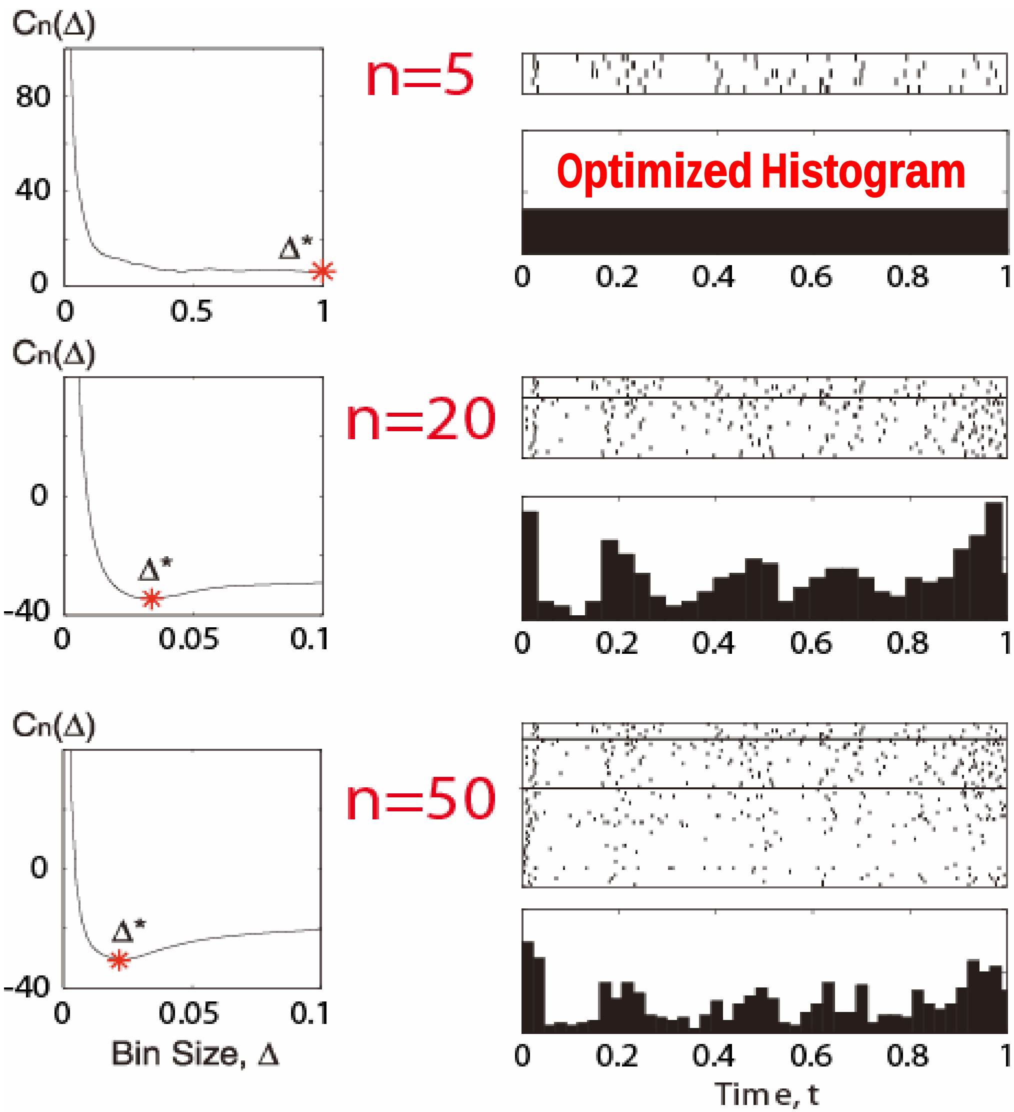
$$C_n(\Delta) = 2\left\langle E\left(\hat{\theta}_n - \theta\right)^2 \right\rangle - \left\langle E\left(\hat{\theta}_n - \left\langle E\hat{\theta}_n \right\rangle\right)^2 \right\rangle.$$

The Poisson statistics obeys: $E(\hat{\theta}_n - \theta)^2 = \frac{1}{n\Delta} E\hat{\theta}_n.$

$$C_n(\Delta) = \frac{2}{n\Delta}\left\langle E\hat{\theta}_n \right\rangle - \left\langle E\left(\hat{\theta}_n - \left\langle E\hat{\theta}_n \right\rangle\right)^2 \right\rangle.$$

Mean of a Histogram Variance of a Histogram

Application to an MT neuron data



2 Too few to make a Histogram !

1 Optimal bin size decreases

Data : Britten et al. (2004) neural signal archive

Theories on
the Optimal Bin Size

If you know the underlying rate, you can compute

See also Koyama and Shinomoto *J. Phys. A*, 37(29):7255–7265. 2004

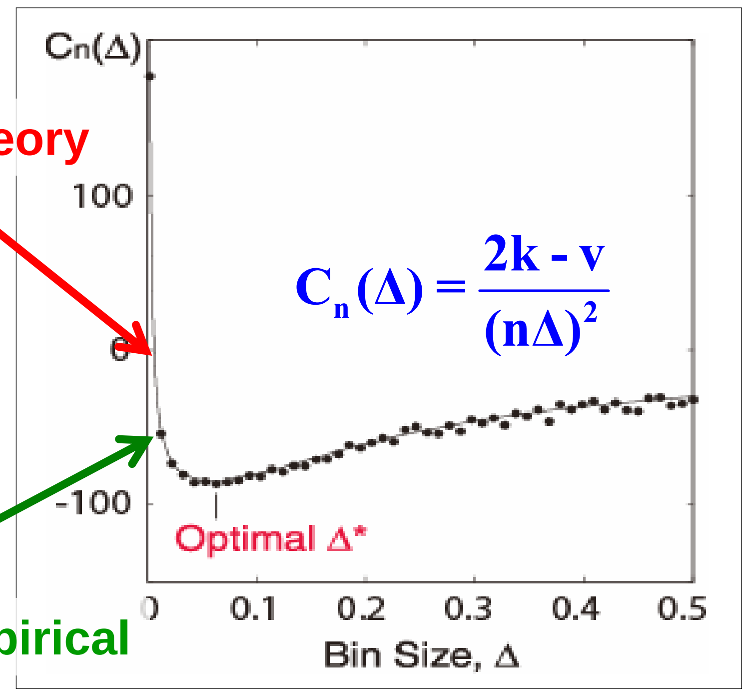
Theoretical cost function:

$$C_n(\Delta) = \frac{\langle \theta \rangle}{n\Delta} - \left\langle \left(\theta - \langle \theta \rangle\right)^2 \right\rangle$$
$$= \frac{\mu}{n\Delta} - \frac{1}{\Delta^2}\int_0^{\Delta}\int_0^{\Delta} \phi(t_1 - t_2)\,dt_1 dt_2 .$$

Mean μ

Correlation function $\phi(t_1 - t_2)$

Theory



Empirical

(i) Expansion of the cost function by Δ:

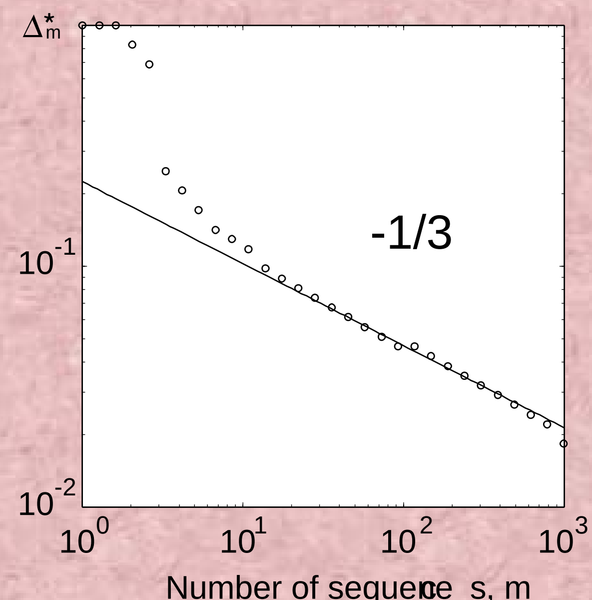
Theoretical cost function: $C_n(\Delta) = \frac{\mu}{n\Delta} - \frac{1}{\Delta^2}\int_0^\Delta \int_0^\Delta \phi(t_1 - t_2)\, dt_1 dt_2.$

When the number of sequences is large, the optimal bin size is very small

The expansion of the cost function by Δ:

$$C_n(\Delta) = \frac{\mu}{n\Delta} - \phi(0) - \frac{1}{3}\phi'(0_+)\Delta - \frac{1}{12}\phi''(0)\Delta^2 + O(\Delta^3).$$

Scaling of the optimal bin size



Scaling of the optimal bin size:

$$\Delta^* \sim \left(-\frac{6\mu}{\phi''(0)\,n}\right)^{1/3}.$$

Ref. Scott (1979)

(ii) Divergence of the optimal bin size

When the number of sequences is small, the optimal bin size is very large.

The expansion of the cost function by 1/Δ:

$$C_n(\Delta) \sim \frac{\mu}{n\Delta} - \frac{1}{\Delta}\int_{-\infty}^{\infty} \phi(t)\,dt + \frac{1}{\Delta^2}\int_{-\infty}^{\infty} |t|\,\phi(t)\,dt$$

$$= \mu\left(\frac{1}{n} - \frac{1}{n_c}\right)\frac{1}{\Delta} + u\,\frac{1}{\Delta^2}$$

The second order phase transition.

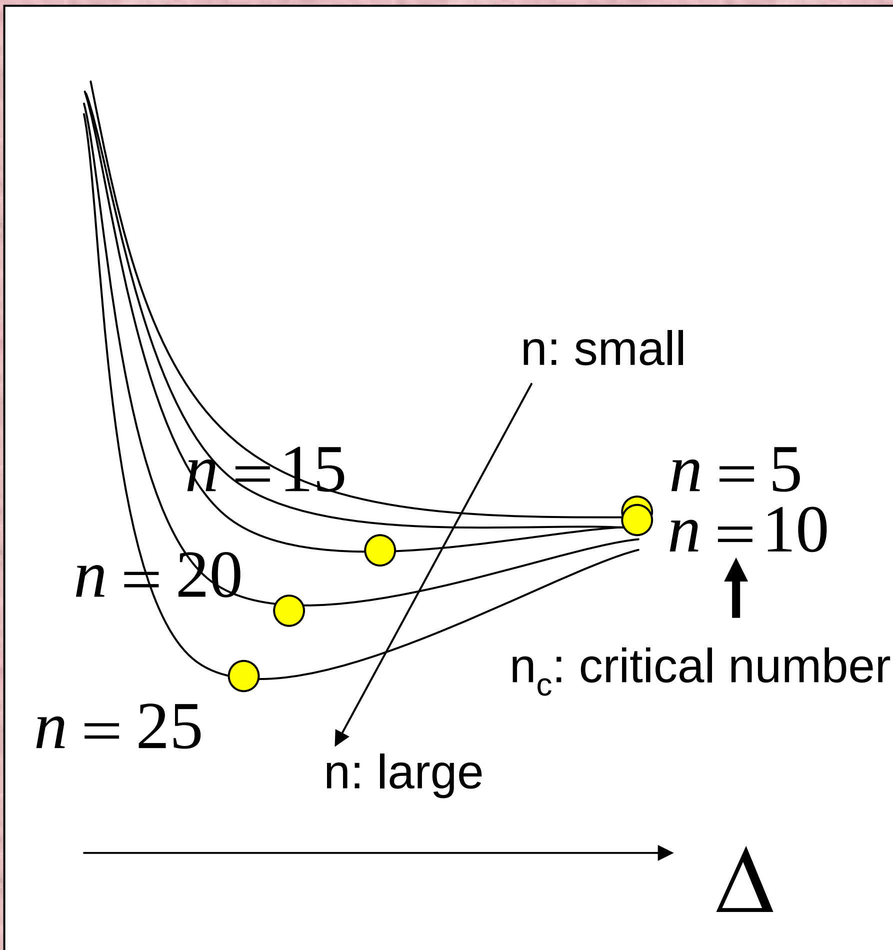
Critical number of trials: $n_c = \mu \Big/ \int_{-\infty}^{\infty} \phi(t)\,dt$

$n < n_c$ Optimal bin size diverges.

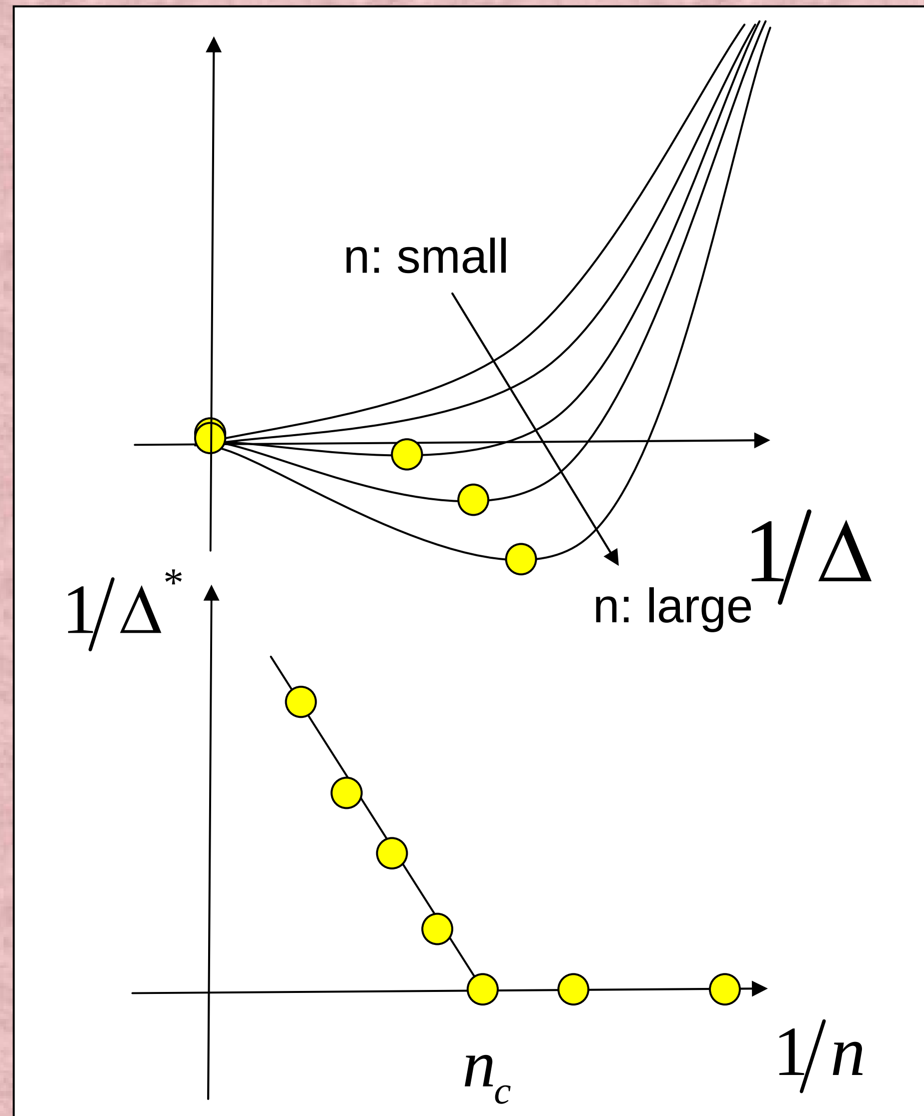
$n > n_c$ Finite optimal bin size.

Not all the process undergoes the first order phase transition.
Others undergo the second order (discontinuous) phase transition.

Phase transitions of optimal bin size

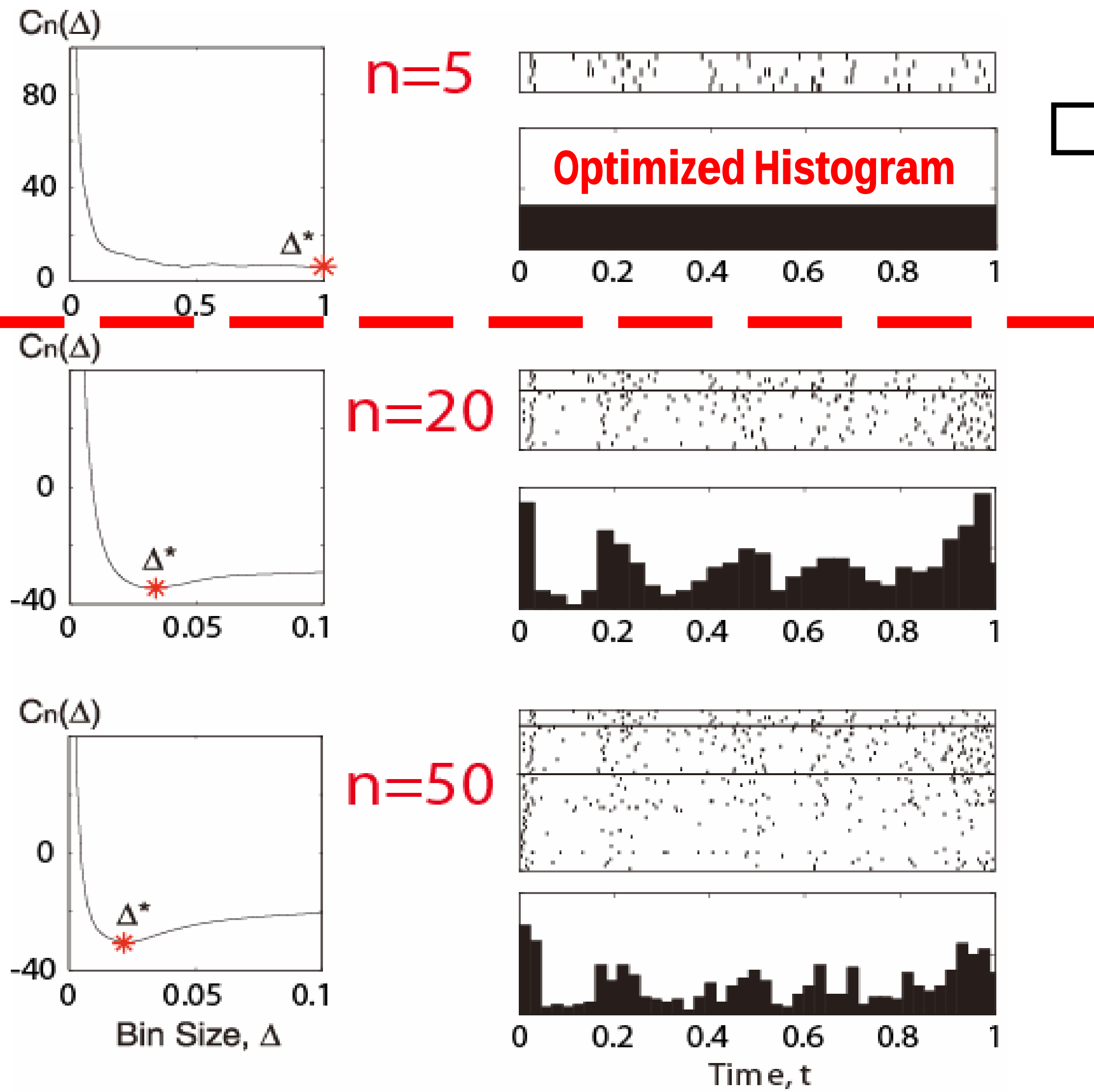


n_c: critical number of trials



Back to Practice!

2. The extrapolation method



Too few to make a Histogram !

Extrapolation

$$C_m(\Delta \mid n) = \left(\frac{1}{m} - \frac{1}{n} \right) \frac{k}{n\Delta^2} + C_n(\Delta)$$

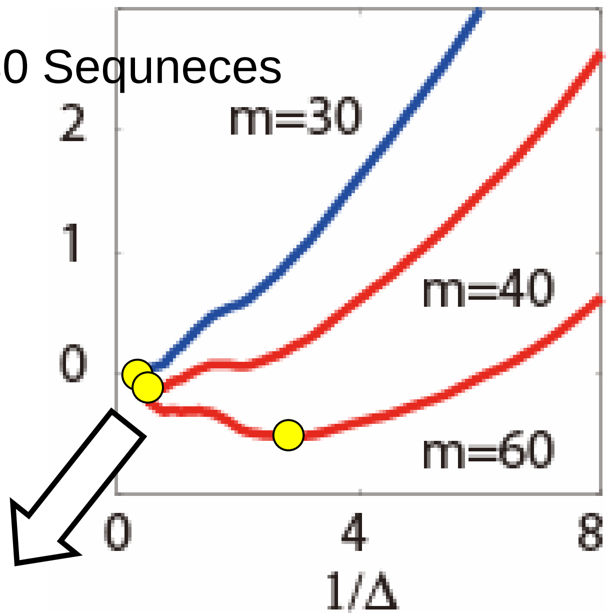
Estimation:

At least 12 trials are required.

Data : Britten et al. (2004) neural signal archive

Verification of the extrapolation method

We have n=30 Sequneces



Original: $C_n(\Delta)$

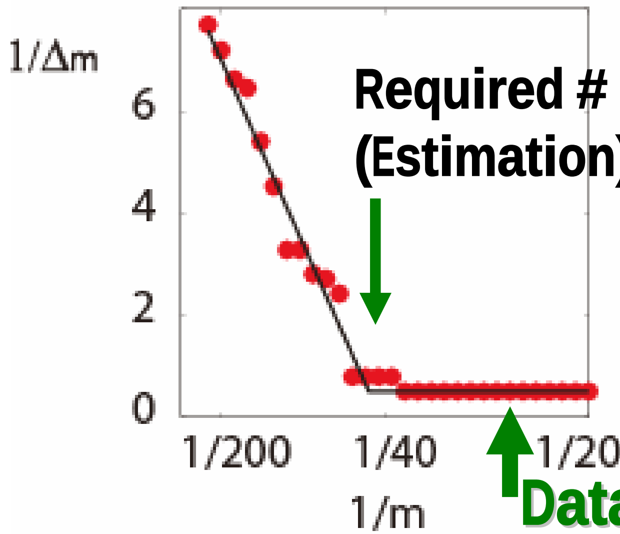
Optimal bin size diverges

Extrapolated:

$$C_m(\Delta \mid n) = \left(\frac{1}{m} - \frac{1}{n}\right)\frac{k}{n\Delta^2} + C_n(\Delta)$$

Finite optimal bin size

Optimal bin size v.s. m

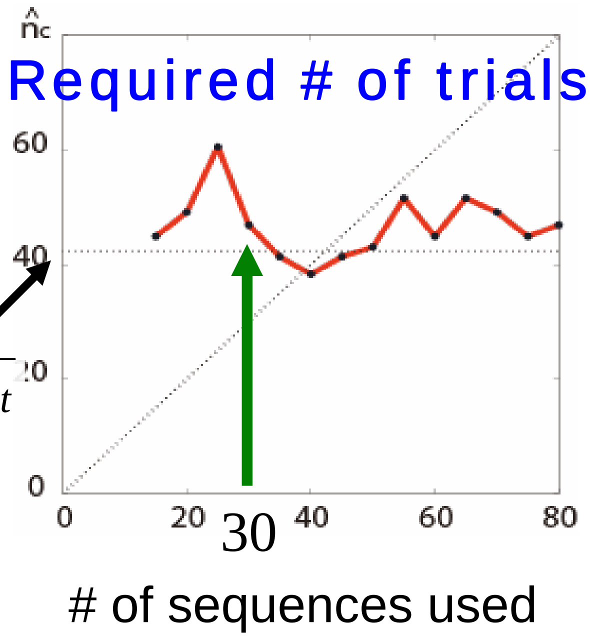


Required # of sequences (Estimation)

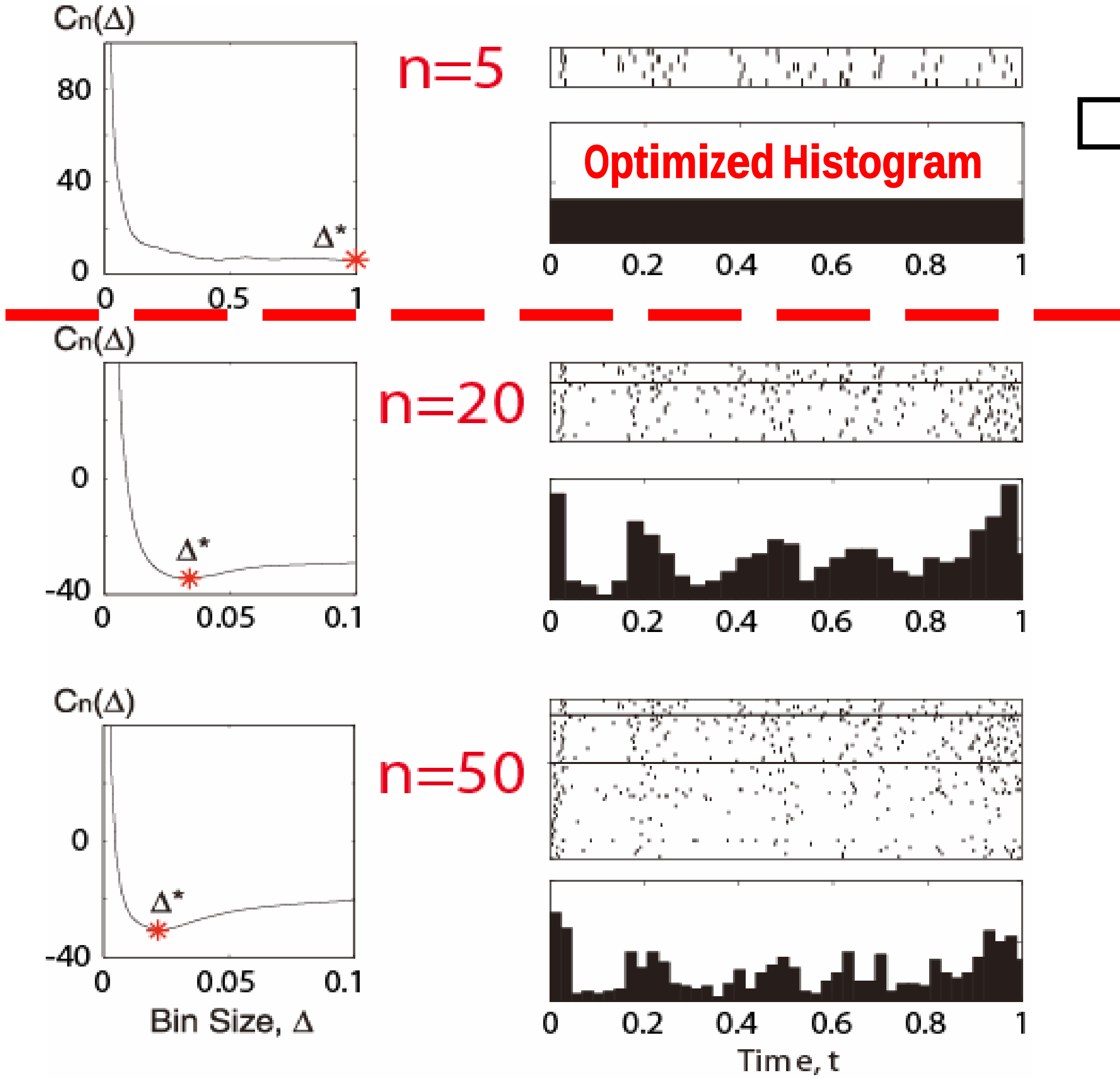
Data Size Used

Required # of trials

$$n_c = \frac{\mu}{\int_{-\infty}^{\infty} \phi(t)\,dt}$$



2. The extrapolation method



Too few to make a Histogram !

Extrapolation

$$C_m(\Delta \mid n) = \left(\frac{1}{m} - \frac{1}{n} \right) \frac{k}{n\Delta^2} + C_n(\Delta)$$

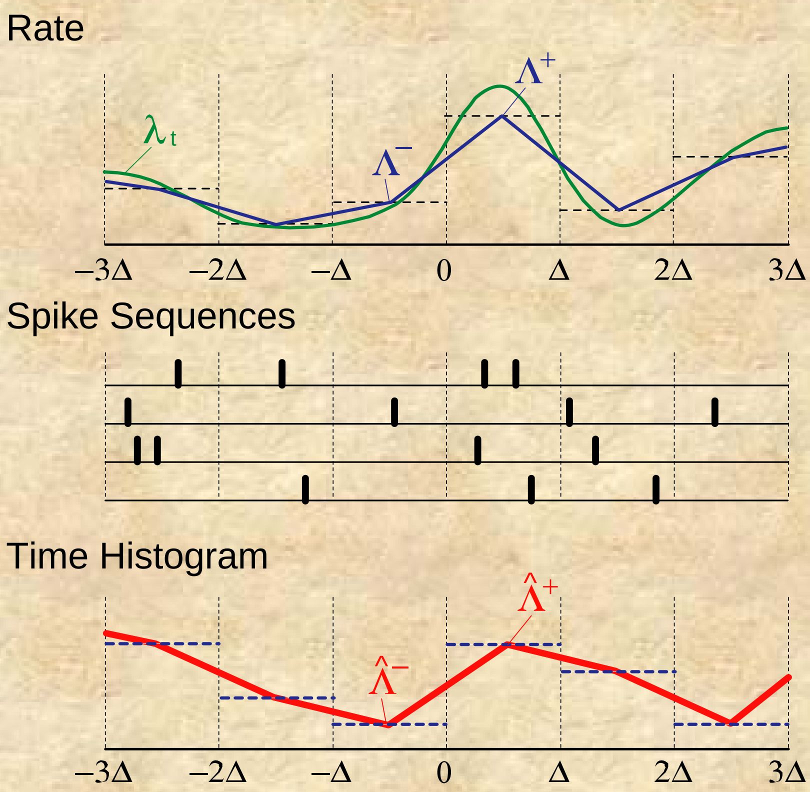
Estimation:

At least 12 trials are required.

Data : Britten et al. (2004) neural signal archive

Advanced Topics

Line-Graph Time Histogram



Line-Graph Model

A line-graph is constructed by connecting top-centers of adjacent bar-graphs.

$$L_t = \frac{\theta^+ + \theta^-}{2} + \frac{\theta^+ - \theta^-}{\Delta} t. \qquad \Lambda^+ \equiv \frac{1}{\Delta}\int_0^\Delta \lambda_t \, dt. \qquad \theta^- \equiv \frac{1}{\Delta}\int_{-\Delta}^0 \lambda_t \, dt.$$

The spike count obeys the Poisson distribution

$$p(k \mid n\Delta\Lambda) = \frac{(n\Delta\theta)^k}{k!} e^{-n\Delta\theta}.$$

An estimator of a line-graph

$$\hat{L}_t = \frac{\hat{\theta}^+ + \hat{\theta}^-}{2} + \frac{\hat{\theta}^+ - \hat{\theta}^-}{\Delta} t.$$

A Recipe for an optimal line-graph TH

(i) Define the four spike counts,

$$k_i^{(+)}(j) \quad k_i^{(-)}(j) \quad k_i^{(0)}(j) \quad k_i^{(*)}(j) \qquad p=\{-,+,0,*\}$$

(ii) Summation of the spike count $\quad k_i^{(p)} \equiv \sum_{j=1}^{n} k_i^{(p)}(j)$

Covariations w.r.t. bins

$$s^{(p,q)} \equiv \frac{1}{N}\sum_{i=1}^{N}\left(k_i^{(p)}-\overline{k}^{(p)}\right)\left(k_i^{(q)}-\overline{k}^{(q)}\right) \qquad \overline{k}^{(p)} \equiv \frac{1}{N}\sum_{i=1}^{N} k_i^{(p)}$$

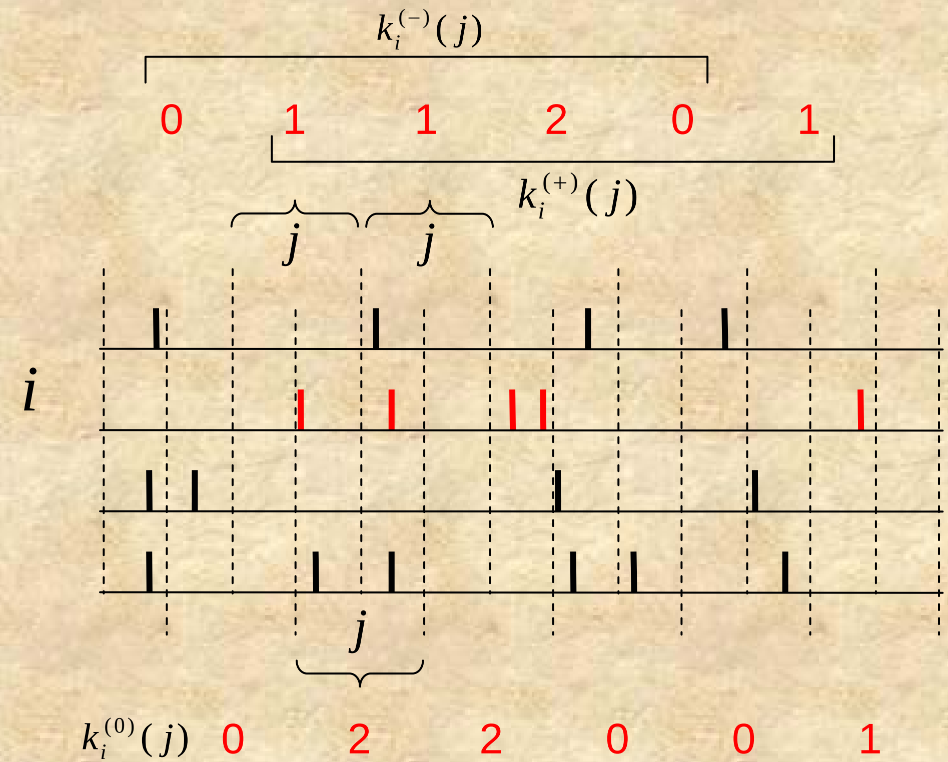
Binned-average

Bin-average of the covariation of spike count w.r.t. sequences,

$$\overline{s}^{(p,q)} \equiv \frac{1}{N}\sum_{i=1}^{N}\frac{1}{n}\sum_{j=1}^{n}\left(k_i^{(p)}(j)-\frac{k_i^{(p)}}{n}\right)\left(k_i^{(q)}(j)-\frac{k_i^{(q)}}{n}\right)$$

(iii) The covariances of an ideal line-graph model is

$$\sigma^{(p,q)} \equiv \frac{s^{(p,q)}}{(n\Delta)^2}-\frac{\overline{s}^{(p,q)}}{n\Delta^2}$$



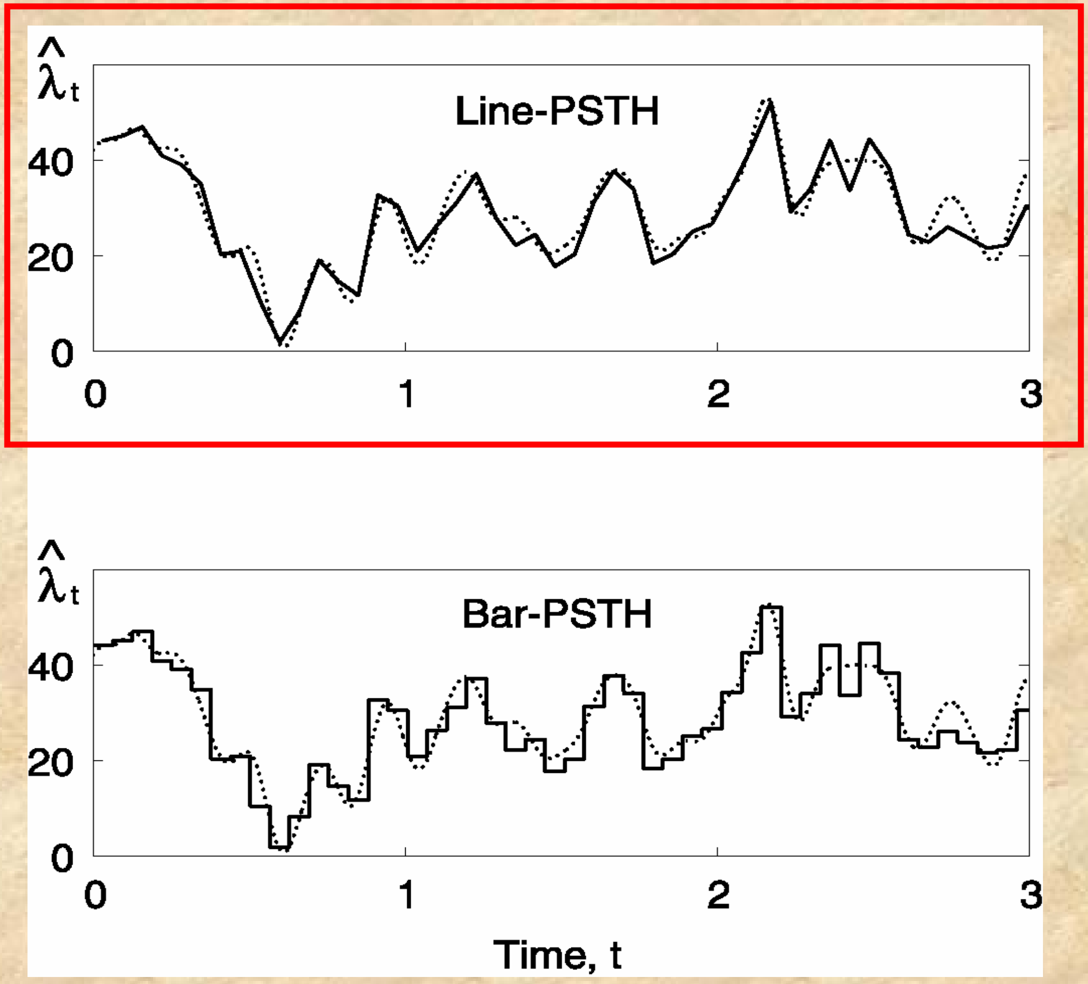
(iv) Cost function:

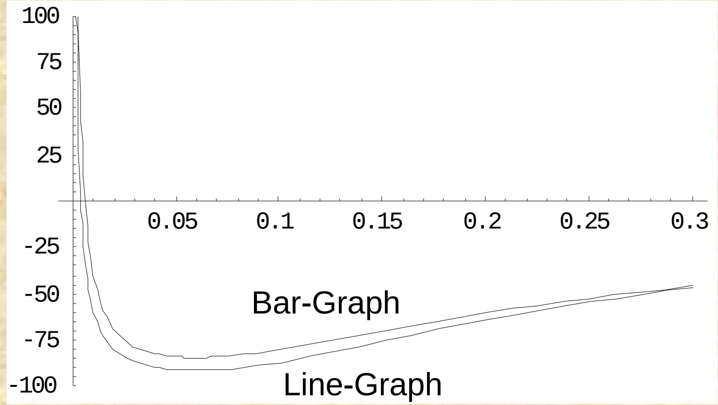
$$C_n(\Delta)=\frac{2}{3}\frac{\overline{k}^{(+)}}{(n\Delta)^2}+\frac{2}{3}\sigma^{(+,+)}+\frac{1}{3}\sigma^{(+,-)}-2\sigma^{(+,0)}-2\sigma^{(+,*)}.$$

$$k_i^{(*)}(j) \equiv 2\sum_{\ell} t_i^{\ell}(j)/\Delta$$

(v) Repatn i through iv by changing Δ. Find the optimal Δ that minimizes the cost function.

The optimal Line-Graph Histogram





The Line-graph histogram performs better if the rate is smooth.

2. Theories on the optimal bin size

Theoretical Cost Function

(Bar-Graph)

$$C_n(\Delta) = \frac{\langle\theta\rangle}{n\Delta} - \left\langle \left(\theta - \langle\theta\rangle\right)^2 \right\rangle$$

$$= \frac{\mu}{n\Delta} - \frac{1}{\Delta^2}\int_0^\Delta \int_0^\Delta \phi(t_1 - t_2)\, dt_1 dt_2 .$$

(Line-Graph)

$$C_n(\Delta) = \frac{2\mu}{3n\Delta} - \frac{2}{\Delta^2}\int_0^\Delta \int_{-\Delta/2}^{\Delta/2} \left(1 + \frac{2t_2}{\Delta}\right) \phi(t_1 - t_2)\, dt_1 dt_2$$

$$+ \frac{2}{3\Delta^2}\int_0^\Delta \int_0^\Delta \phi(t_1 - t_2)\, dt_1 dt_2 + \frac{1}{3\Delta^2}\int_0^\Delta \int_{-\Delta}^0 \phi(t_1 - t_2)\, dt_1 dt_2 .$$

Generalization of Koyama and Shinomoto *J. Phys. A*, 37(29):7255–7265. 2004

(i) Scalings of the optimal bin size

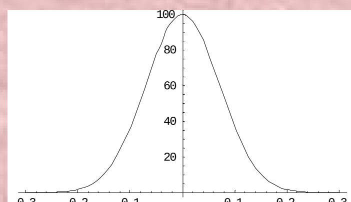
The expansion of the cost function by Δ:

(Bar-Graph) $$C_n(\Delta) = \frac{\mu}{n\Delta} - \phi(0) - \frac{1}{3}\phi'(0_+)\Delta - \frac{1}{12}\phi''(0)\Delta^2 + O(\Delta^3).$$

(Line-Graph) $$C_n(\Delta) = \frac{2\mu}{3n\Delta} - \phi(0) - \frac{37}{144}\phi'(0_+)\Delta + \frac{181}{5760}\phi'''(0_+)\Delta^3 + \frac{49}{2880}\phi''''(0)\Delta^4 + O(\Delta^5)$$

the second order term vanishes.

A smooth process: A correlation function is smooth at origin.

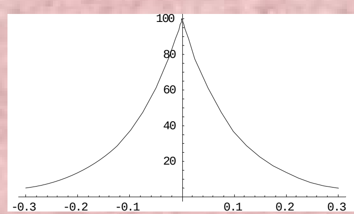


$\phi'(0_+) = 0$

(Bar-Graph) $$\Delta^* \sim \left(-\frac{6\mu}{\phi''(0)n} \right)^{1/3}.$$

(Line-Graph) $$\Delta^* \sim \left(\frac{1280\mu}{181\phi''''(0)n} \right)^{1/5}.$$

A jagged process: A correlation function has a cusp at origin.



$\phi'(0_+) \neq 0$

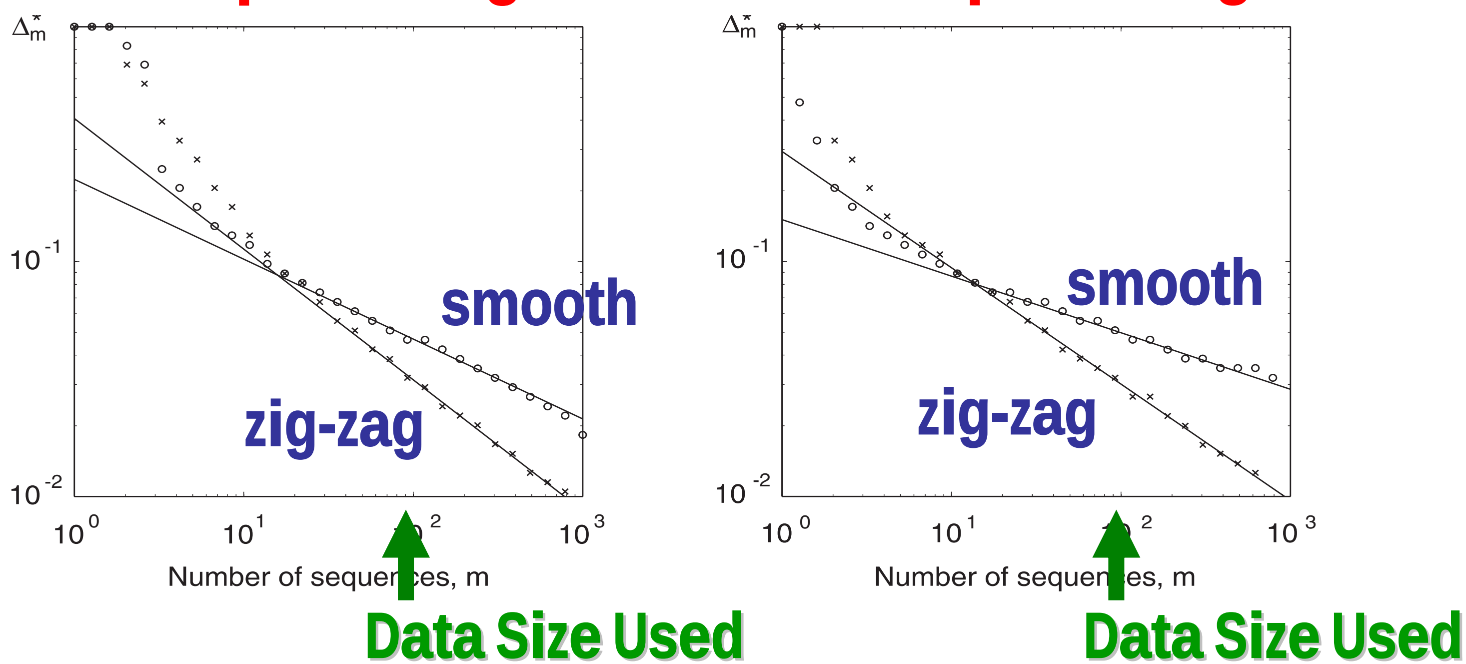
(Bar-Graph) $$\Delta^* \sim \left(-\frac{3\mu}{\phi'(0_+)n} \right)^{1/2}.$$

(Line-Graph) $$\Delta^* \sim \left(-\frac{96\mu}{37\phi'(0_+)n} \right)^{1/2}.$$

Identification of the scaling exponents

Bar-Graph Histogram

Line-Graph Histogram



	Bar-Graph Histogram	Line-Graph Histogram
smooth	$\Delta^* \sim n^{-1/3}$	$\Delta^* \sim n^{-1/5}$
zig-zag	$\Delta^* \sim n^{-1/2}$	$\Delta^* \sim n^{-1/2}$

The second and the first order phase transitions

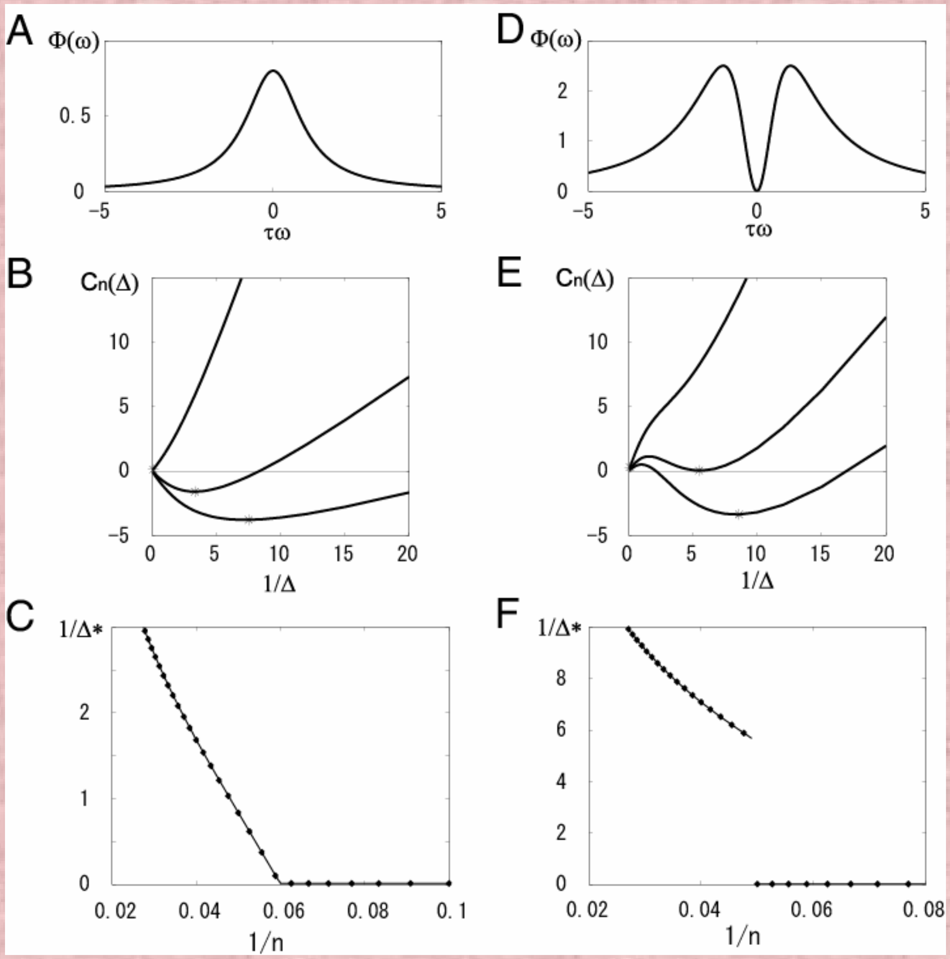
Second order (Continuous)

First order (Discontinuous)

Power spectrum of a rate process

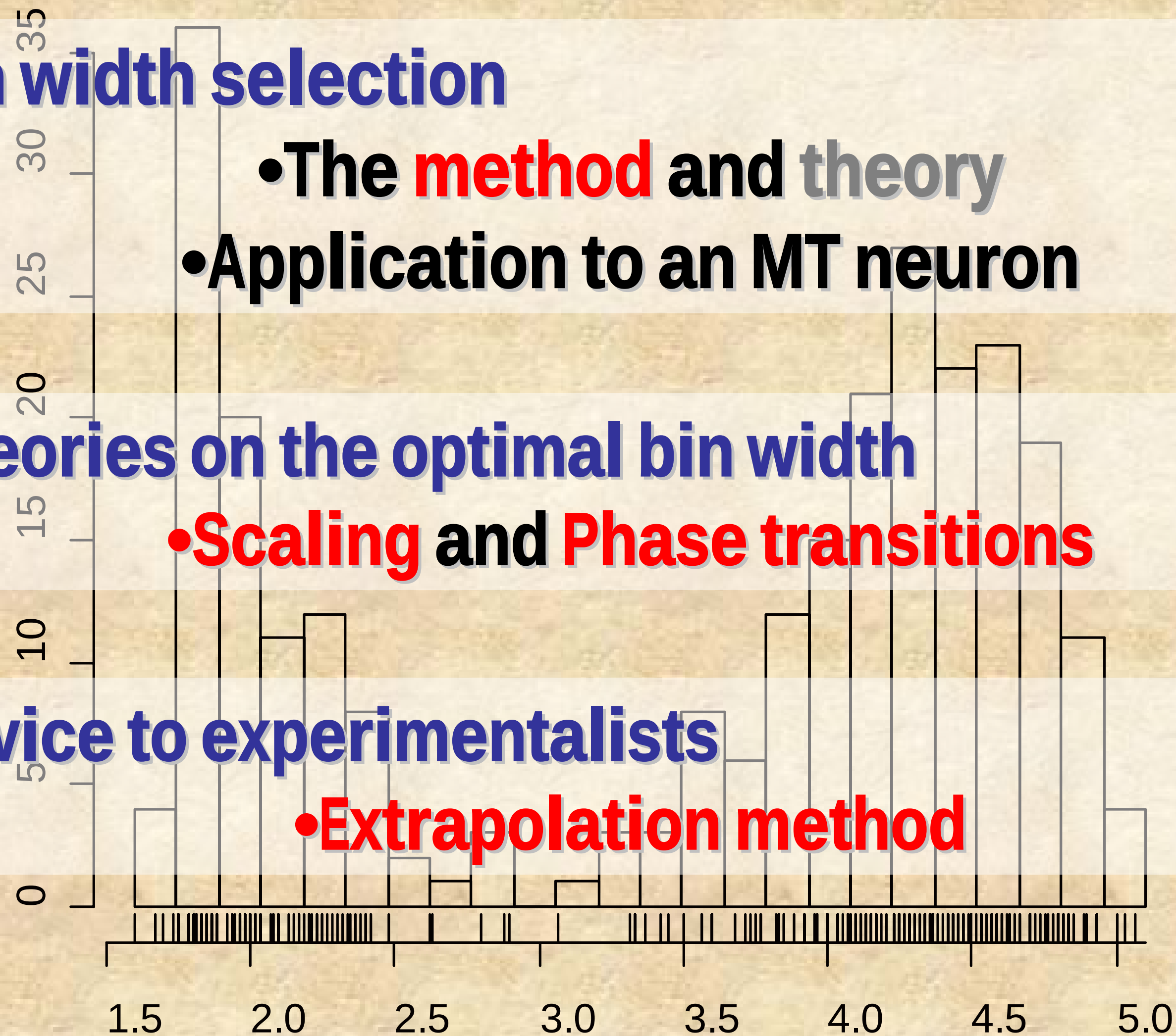
Cost functions

Optimal Bin Size



Summary of Today's Talk

1. **Bin width selection**
   - The method and theory
   - Application to an MT neuron
2. **Theories on the optimal bin width**
   - Scaling and Phase transitions
3. **Advice to experimentalists**
   - Extrapolation method

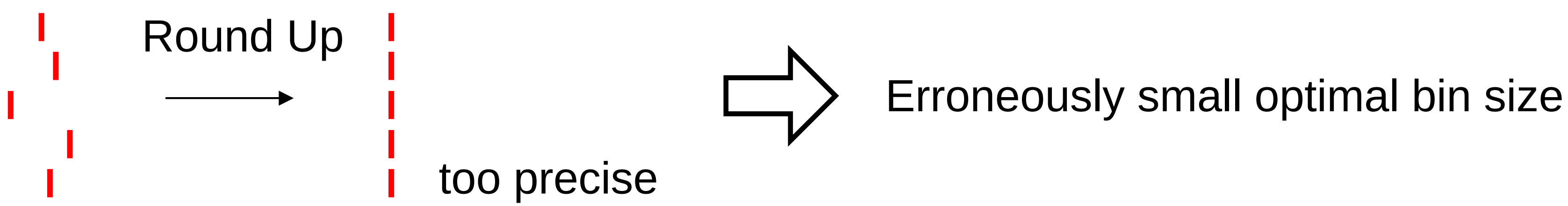


FAQ

Q. Can I apply the proposed method to a histogram for a probability distribution?
A. Yes.

Q. I want to make a 2-dimensional histogram. Can I use this method?
A. Yes.

Q. I obtained a very small bin width, which is likely to be erroneous. Why?
A. You probably searched smaller bin size than the sampling resolution.



Q. Can I use unbiased variance for computation of the cost function?
A. No.

$$C(\Delta) = \frac{2k - v}{\Delta^2}$$

$$v = \frac{1}{N}\sum_{i=1}^{N}(k_i - k)^2$$

$$\cancel{v = \frac{1}{N-1}\sum_{i=1}^{N}(k_i - k)^2}$$

Reference

A Method for Selecting the Bin Size of a Time Histogram

Hideaki Shimazaki and Shigeru Shinomoto

Neural Computation in Press

Short Summary:
Adavances in *Neural Information Processing Systems* Vol. 19, 2007

- **Web Application for the Bin Size Selection**
- **Matlab / Mathematica / R sample codes**

are available at our homepage
http://www.ton.scphys.kyoto-u.ac.jp/~shino/
/~hideaki/

See also

Koyama, S. and Shinomoto, S. Histogram bin width selection for time-dependent poisson processes. *J. Phys. A*, 37(29):7255–7265. 2004

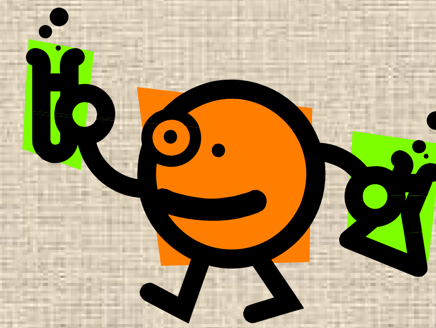
Acknowledgements

Prof. Shigeru Shinomoto
Dr. Shinsuke Koyama

JSPS Research Grant

Introducing myself…

Hideaki Shimazaki

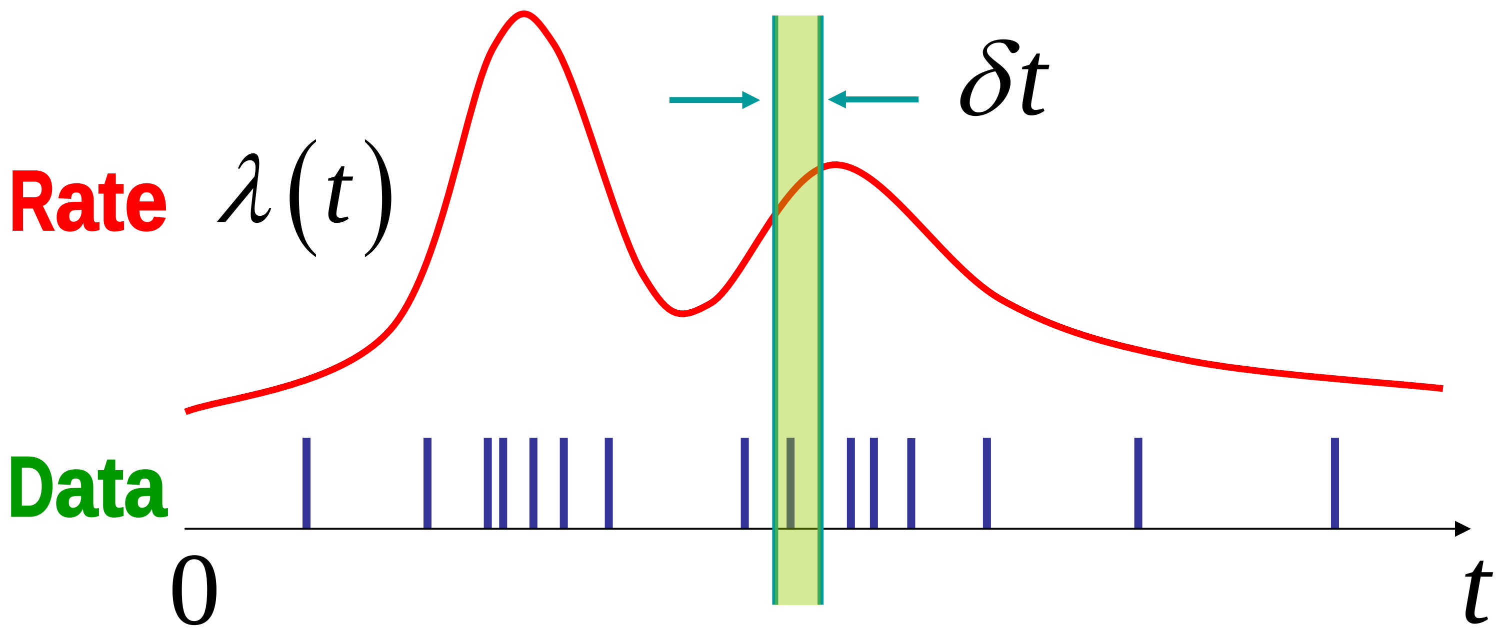


2003 MA ***Prof. Ernst Niebur***
Dept. of Neuroscience, **Johns Hopkins University**
Baltimore, Maryland

2007 Ph. D ***Prof. Shigeru Shinomoto***
Dept. of Physics **Kyoto University**, *Kyoto, Japan*

2006-2008 **JSPS Research Fellow**

The Poisson Point Process



$$\Pr[\text{One event in } \delta \text{t}] = \lambda(t)\delta t$$

$$\Pr[\text{More than one event in } \delta t] = O(\delta t)$$

⇨ Samples are independently drawn from an identical distribution.