Decoding Neural Activity

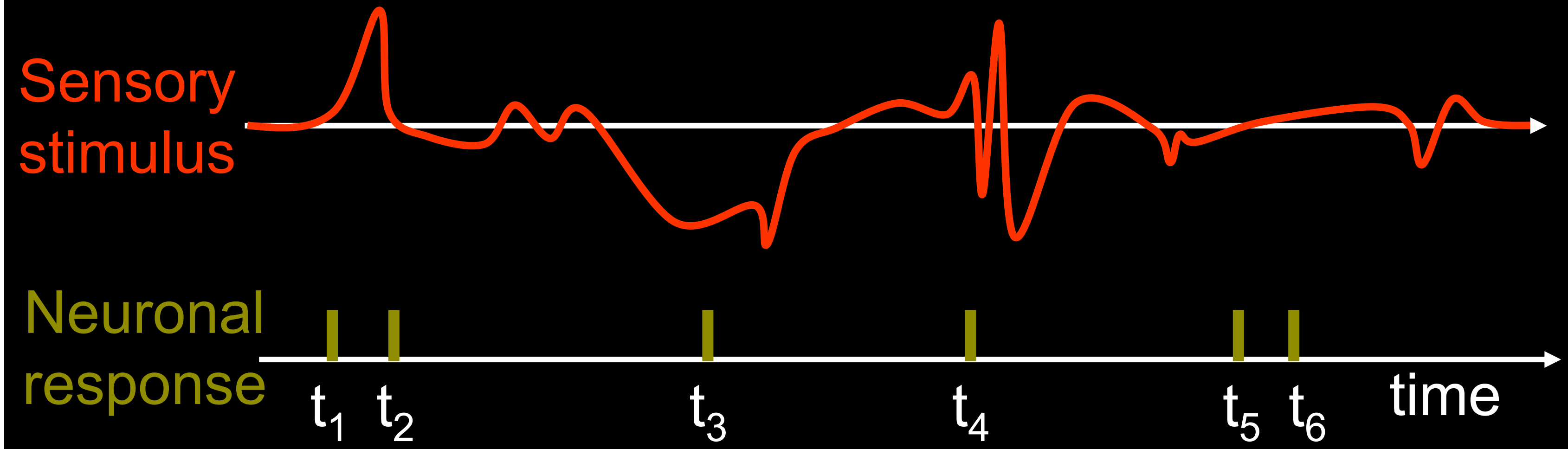
Rasmus Petersen

Mathematical Neuroscience Meeting, Edinburgh, March 2009

1. **Neuro-syntax**:
   What is basic ‘syntactic’ unit of the neuronal response?
2. **Neuro-semantics:**
   What ‘semantic’ features do these syntactic units represent?

Outline

- Neurosyntax
  - Information theory
- Neurosemantics
  - Systems identification
- Biological ‘case study’
  - Whisker system



11000001000001000011000

P[spike train | sensory stimulus]

Spike train $\mathbf{r} = \{t_1, t_2, \ldots, t_N\}$
$= (r_1, r_2, \ldots, r_L)$
Stimulus $\mathbf{s} = (s_1, s_2, \ldots, s_T)$

Possible simplifications

- Coarse-grain the response?
- Ignore higher order statistical structure?

Coarse graining

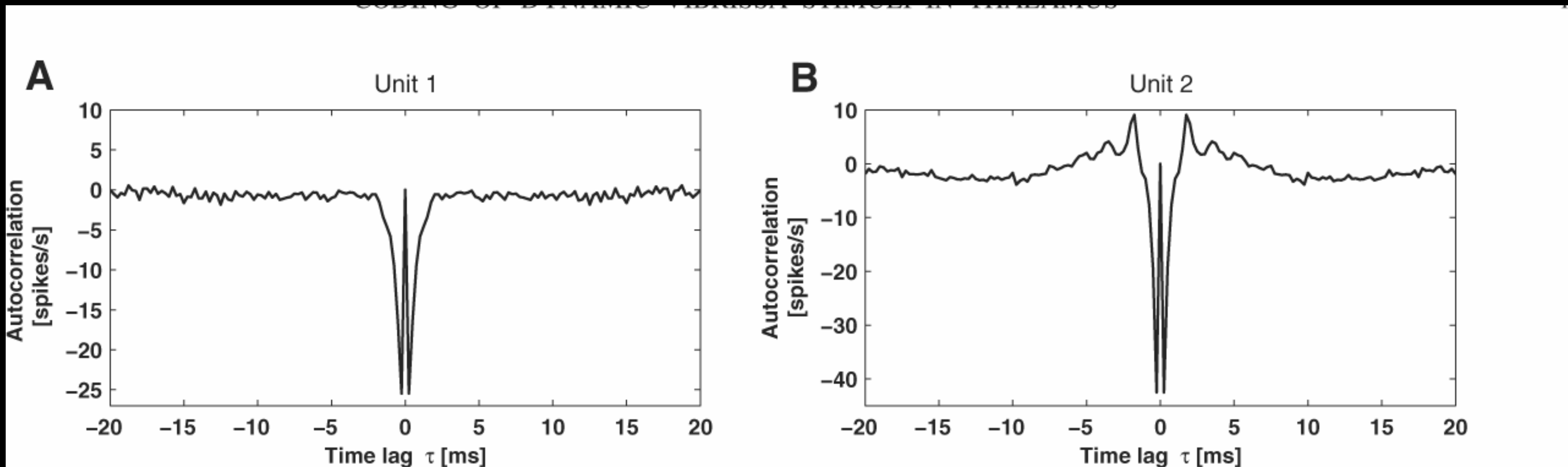
1100	0001	0000	1000	0110	000

2	1	0	1	2	0

Spike count hypothesis: The key information-bearing unit of the spike train is the number of spikes (**spike count**) fired within a time window (long enough to include several spikes). *Corollary*. Precise timing of spikes within the window is just noise.

Higher order statistical structure

- Correlation here defined as:
  - $P[(r_1, r_2, \ldots, r_N)|\mathbf{s}] \neq P[r_1|\mathbf{s}]\ P[r_2|\mathbf{s}] \ldots P[(r_N|\mathbf{s}]$
  - 'noise correlation'
- Uncorrelated 'rate' model:
  - $P[(r_1, r_2, \ldots, r_N)|\mathbf{s}] = P[r_1|\mathbf{s}]\ P[r_2|\mathbf{s}] \ldots P[r_N|\mathbf{s}]$

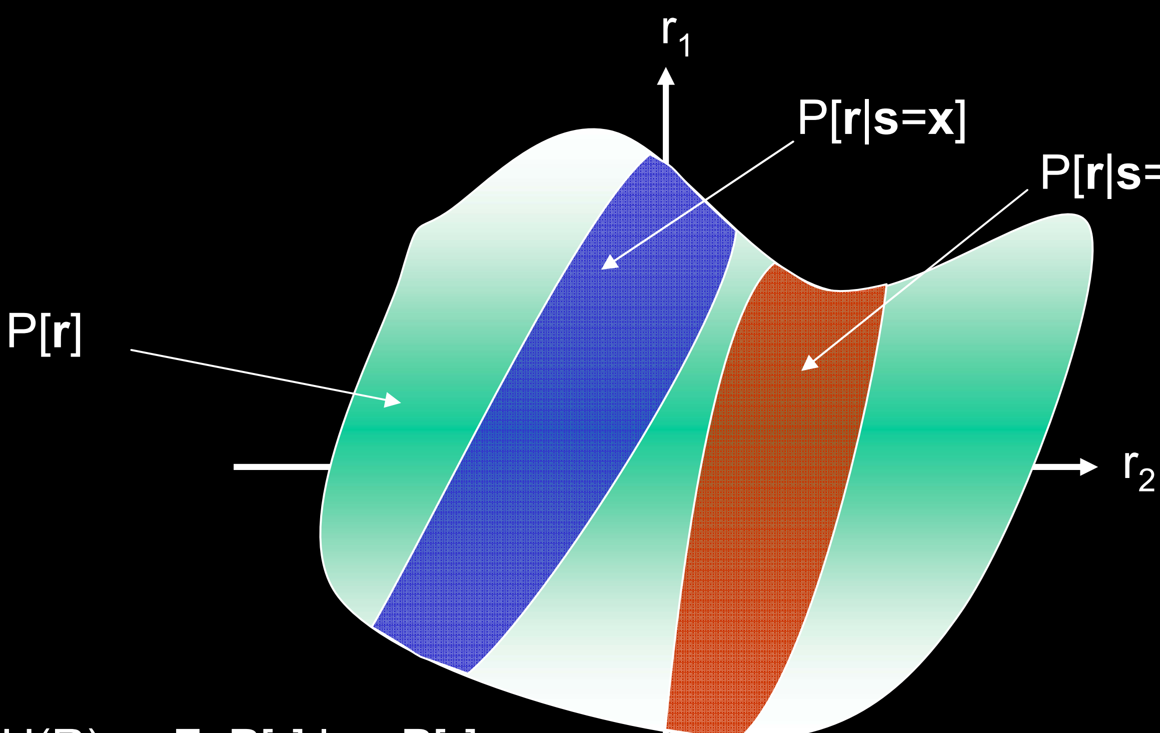


Testing coding hypotheses

- Spike count hypothesis:
  - Predicts that precise spike timing adds nothing to the “information” available from the spike count
- Rate coding hypothesis:
  - Predicts that no “information” lost by neglecting correlations

Information Theory

Stimulus **s** $S=\{\mathbf{s}\}$, $P[\mathbf{s}]$	→	Communication channel	→	Response **r** $R=\{\mathbf{r}\}$, $P[\mathbf{r}]$



$H(R|\mathbf{s}) = -\Sigma_r P[\mathbf{r}|\mathbf{s}] \log_2 P[\mathbf{r}|\mathbf{s}]$

$H[R|S] = -\Sigma_s P[\mathbf{s}] H[R|\mathbf{s}]$

Noise entropy

$H(R) = -\Sigma_r P[\mathbf{r}] \log_2 P[\mathbf{r}]$

Response entropy

$I(S;R) = H(R) - H(R|S)$

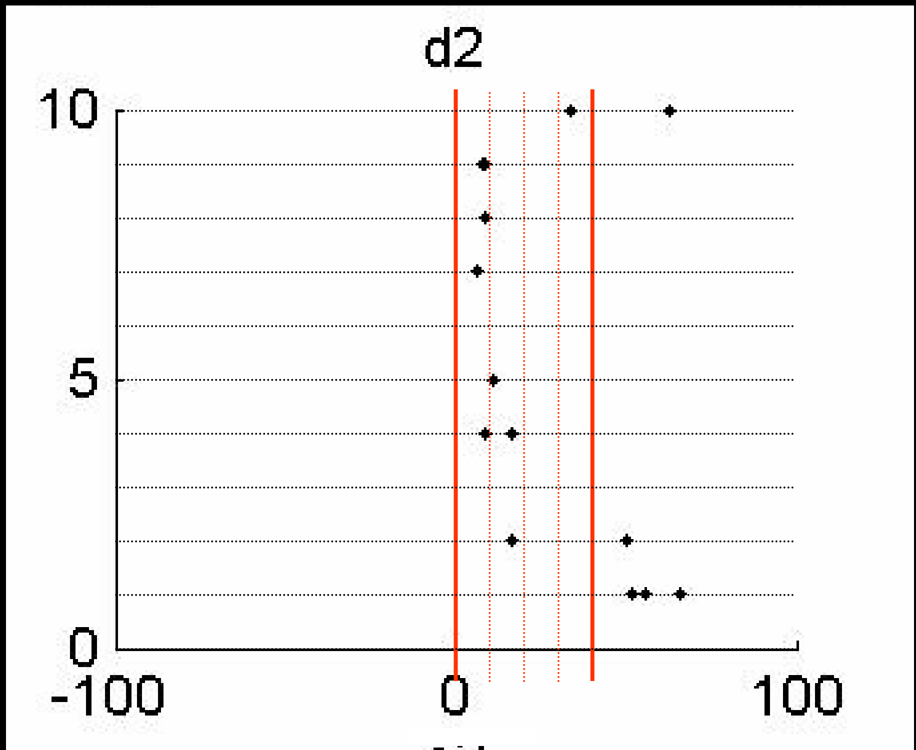
Mutual Information

Estimating mutual information

Static stimuli

Dynamic stimuli

Estimating I(S;R) – static stimuli



Stimulus trial

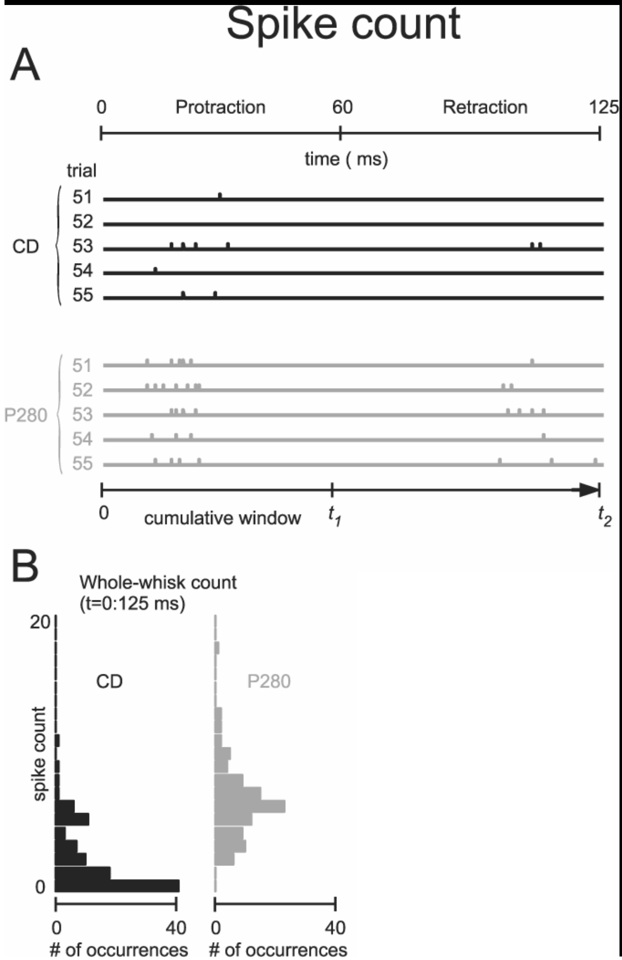
Post-stimulus time

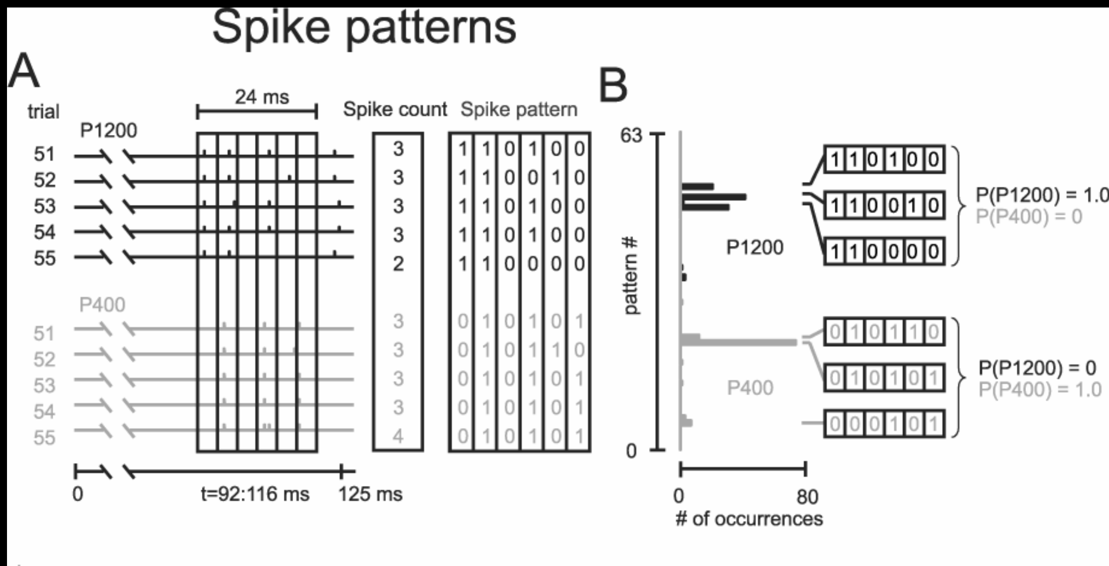
Spike count	Spike timing
1	0001
1	1000
1	1000
1	1000
0	0000
1	0100
2	1100
0	0000
1	0100
0	0000
0	0000

$$I(S|R) = H(R) - H(R|S)$$

$$H(R) = -\Sigma_{\mathbf{r}}\, P[\mathbf{r}] \log_2 P[\mathbf{r}]$$

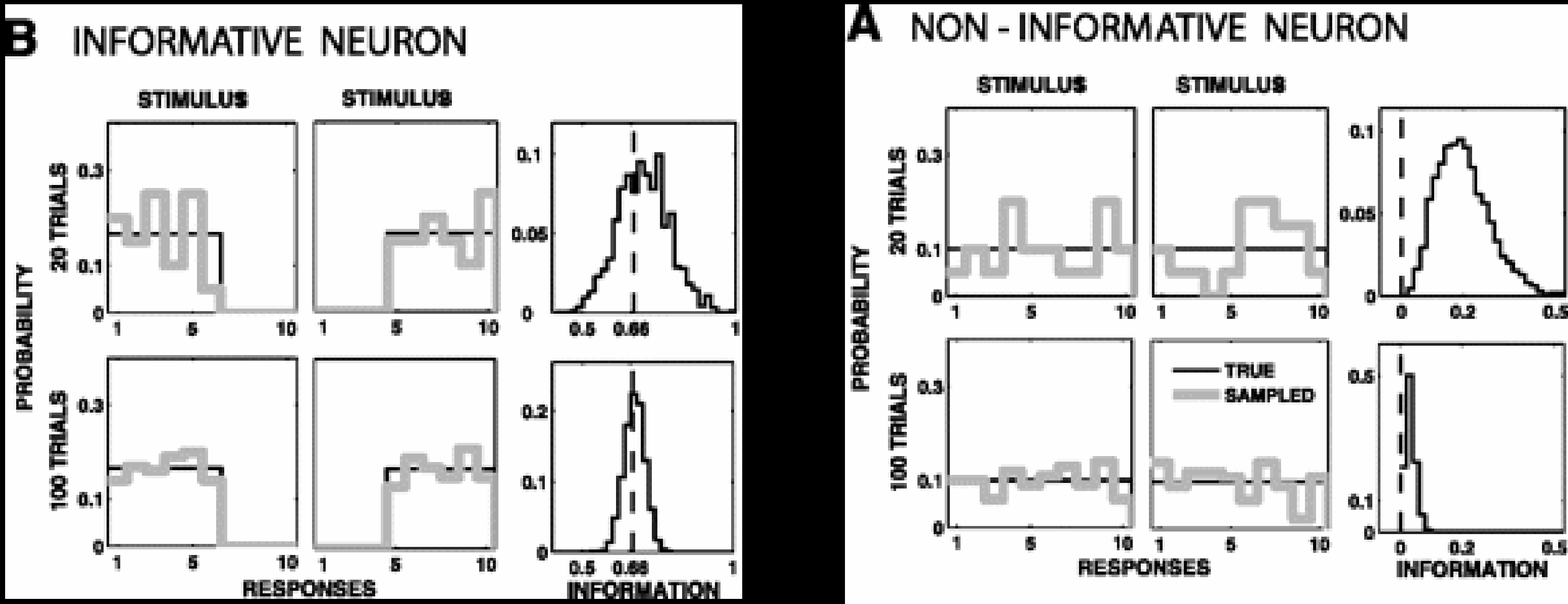
$$H(R|S) = -\Sigma_{\mathbf{s}}\, P[\mathbf{s}]\, \Sigma_{\mathbf{r}}\, P[\mathbf{r}|\mathbf{s}) \log_2 P[\mathbf{r}|\mathbf{s}]$$



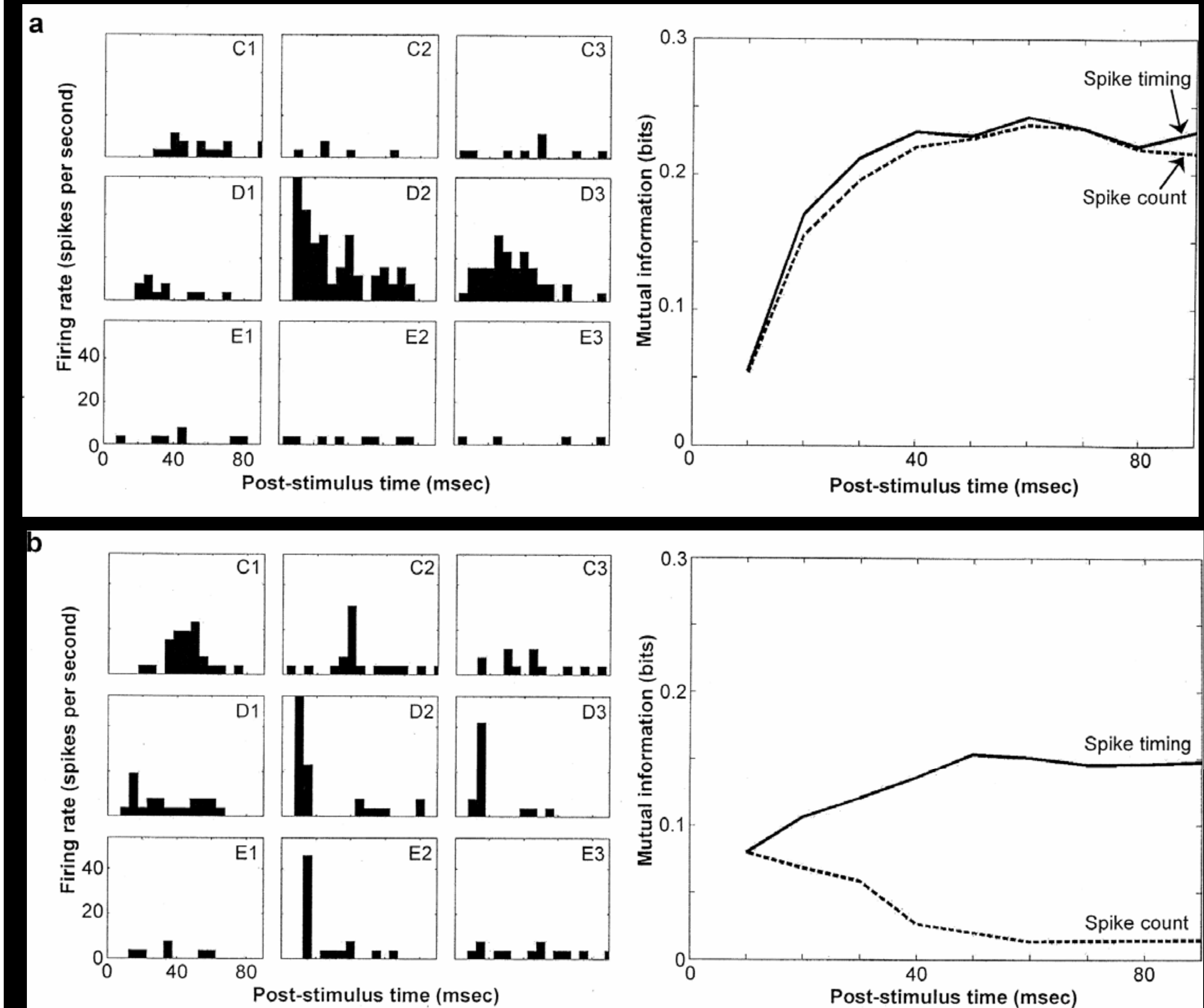


Arabzadeh et al (2006) *J Neurosci*

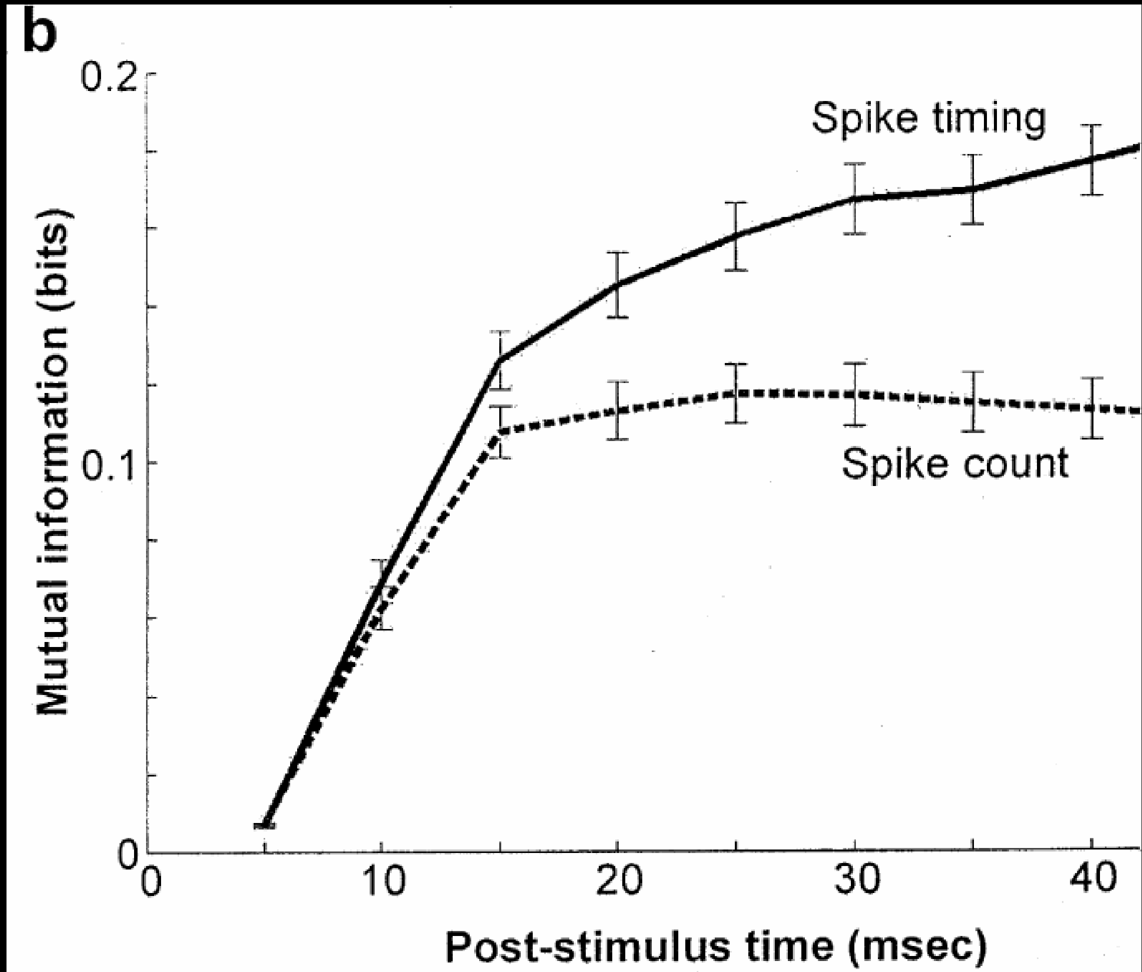
Sampling bias



Panzeri et al (2007) *J Neurophysiol*

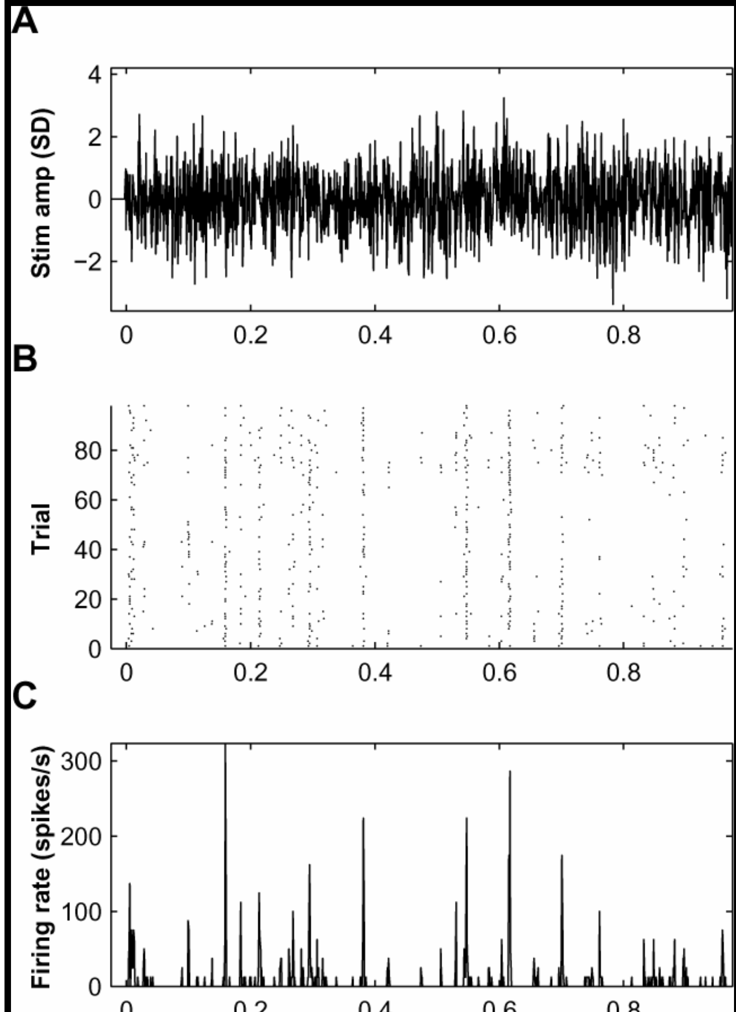


Panzeri et al (2001) *Neuron*
Petersen et al (2001) *Neuron*

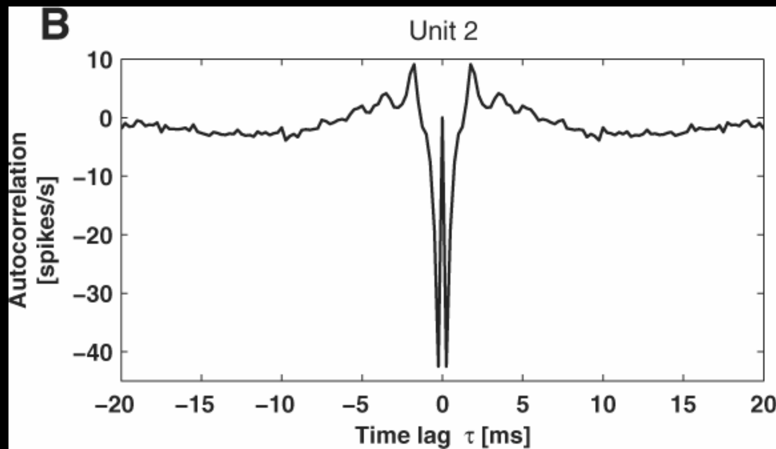


Panzeri et al (2001) *Neuron*
Petersen et al (2001) *Neuron*

Rates vs Correlations



- Correlated model:
  $P[(r_1,r_2,\ldots,r_N)|\mathbf{s}] \neq P[r_1|\mathbf{s}]\ P[r_2|\mathbf{s}]\ldots\ P[(r_N|\mathbf{s}]$
- Uncorrelated ‘rate’ model:
  $P[(r_1,r_2,\ldots,r_N)|\mathbf{s}] = P[r_1|\mathbf{s}]\ P[r_2|\mathbf{s}]\ldots\ P[r_N|\mathbf{s}]$



Assessing correlations

- Encoding framework (Schneidman et al., 2003)
- Decoding framework (Nirenberg et al., 2001)
- Information breakdown (Pola et al., 2003)

Encoding framework

- Uncorrelated model:

$$P_{ind}[(r_1,r_2,\ldots,r_N)|\mathbf{s}] = P[r_1|\mathbf{s}]\ P[r_2|\mathbf{s}]\ldots P[r_N|\mathbf{s}]$$

- Estimate I(S;R) using $P_{ind}[\mathbf{s}|\mathbf{r}]$ instead of $P[\mathbf{s}|\mathbf{r}]$
  $\rightarrow I_{ind}(S;R)$
  $I_{cor}(S;R) = I(S;R) - I_{ind}(S;R)$

- I_{cor} tells us how much the presence of noise correlations changes I(S;R) compared to a hypothetical neural system with identical PSTH but zero noise correlations.

Decoding framework

- Bayesian decoding:
  $P[\mathbf{s}|\mathbf{r}] = P[\mathbf{r}|s]\, P[\mathbf{s}] / P[\mathbf{r}]$

- Use $P_{ind}[\mathbf{r}|\mathbf{s}]$ instead of $P[\mathbf{r}|\mathbf{s}]$

- Upper bound on I(S;R) that will be lost:
  $\Delta I(S;R) = \Sigma_{\mathbf{r}}\, P[\mathbf{r}]\, \Sigma_{\mathbf{s}}\, P[\mathbf{s}|\mathbf{r}]\, \log_2 P[\mathbf{s}|\mathbf{r}] / P_{ind}[\mathbf{s}|\mathbf{r}]$

- **ΔI** tells us, If a hypothetical downstream decoding circuit attempts to decode using a model that includes PSTHs but ignores noise correlations, upper bound on how much I(S;R) will be lost.

Latham & Nirenberg (2005) *J Neurosci*

Information Breakdown

$I(S;R) = I_{ind}(S;R) + I_{cor\text{-}ind}(S;R) + I_{cor\text{-}dep}(S;R)$

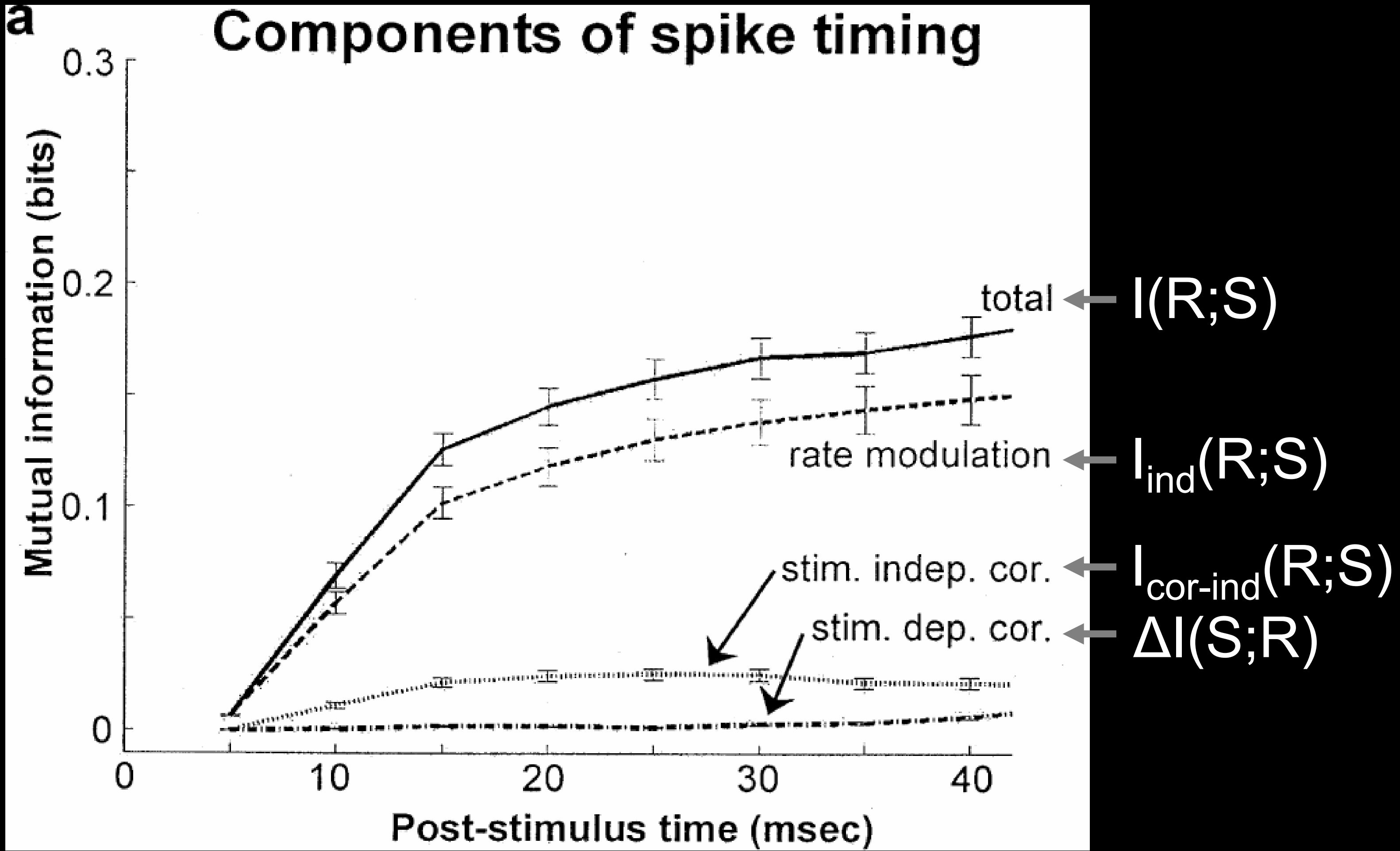
Link to encoding framework:

$I_{cor}(S;R) = I_{cor\text{-}ind}(S;R) + I_{cor\text{-}dep}(S;R)$

Link to decoding framework:

$\Delta I(S;R) = I_{cor\text{-}dep}(S;R)$

Pola et al (2003) *Neural Comp*



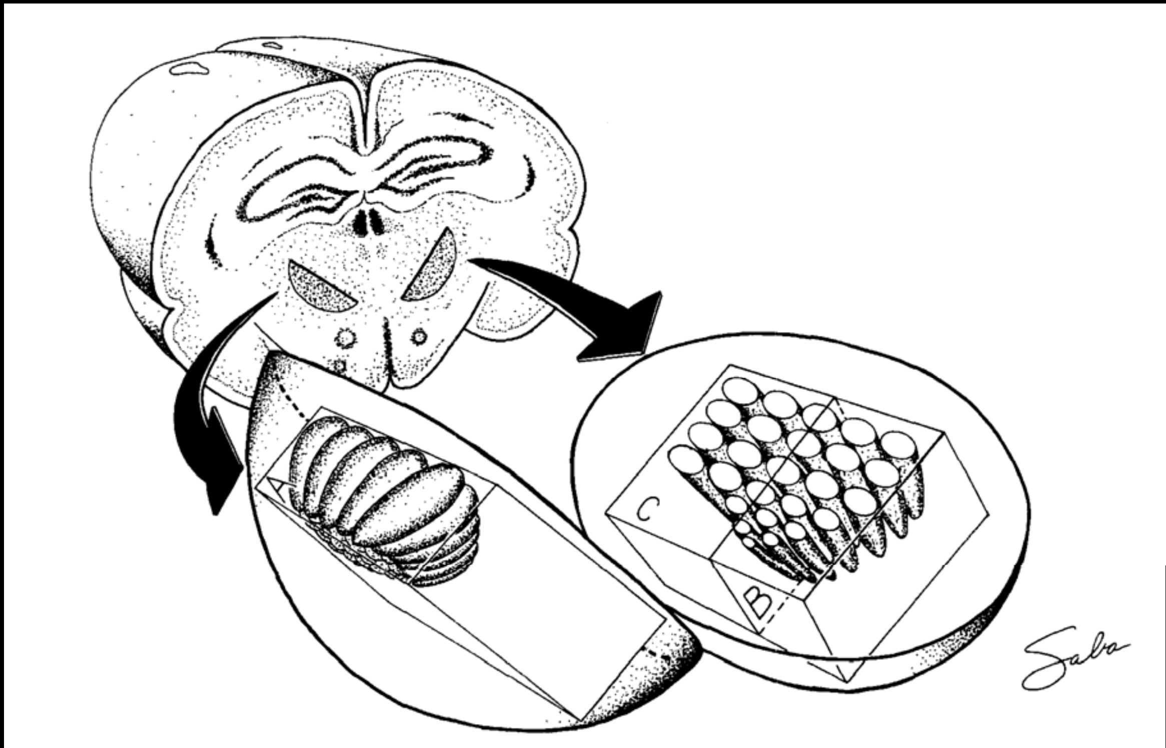
Panzeri et al (2001) *Neuron*
Petersen et al (2001) *Neuron*

Case study

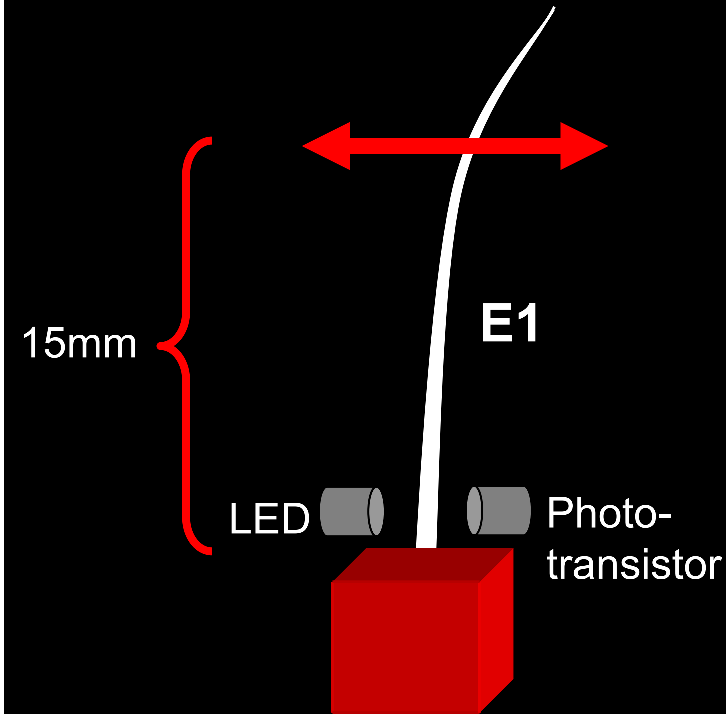
Whisker system

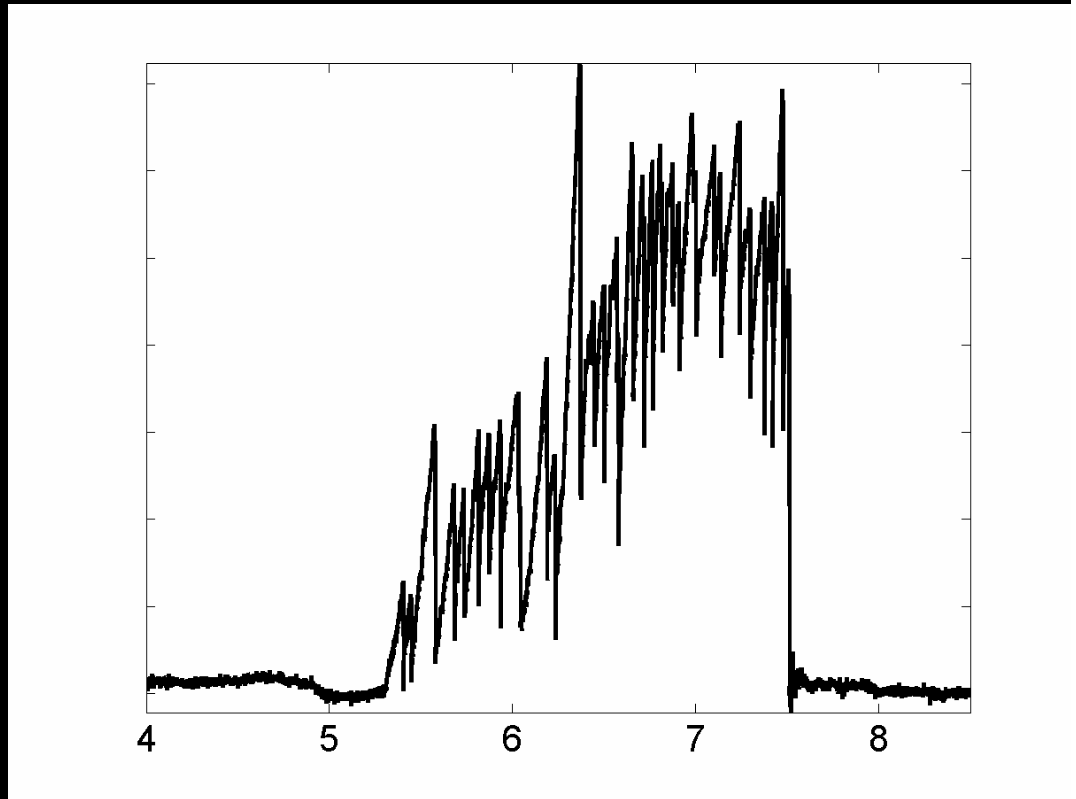
Petersen et al (2009) *Biol Cybern*

Ventro-posterior medial nucleus (VPM)



Land et al (1995)





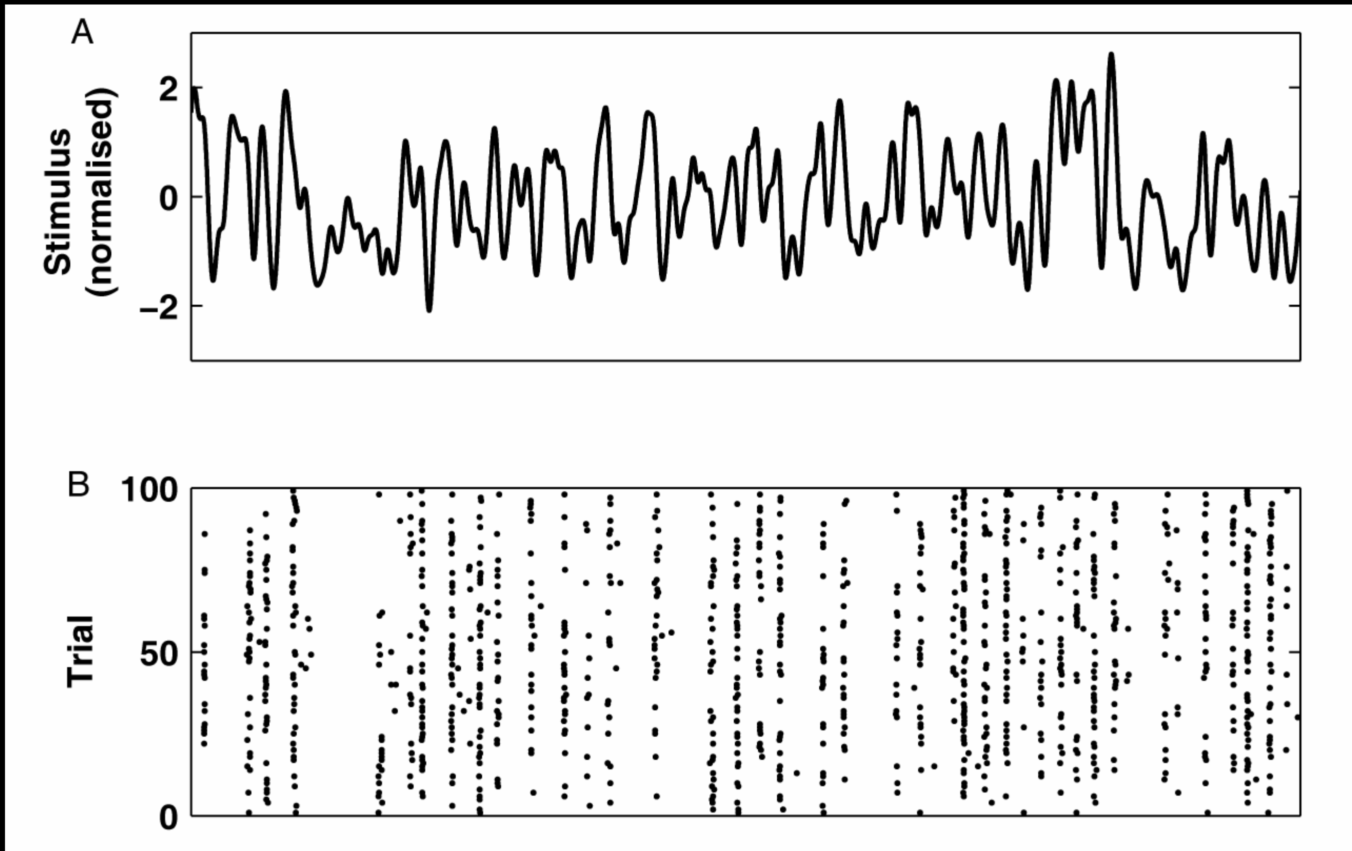
Space-to-time transformation:
Object texture to whisker vibration

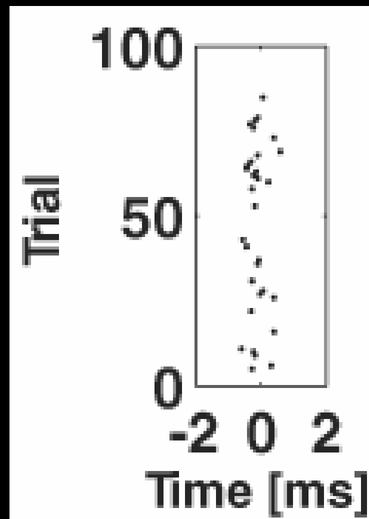
How is high band-width information about whisker vibrations squeezed through the subcortical bottleneck?

Stimulus: (Smoothed) white noise

Preparation: anaesthetised, adult rat

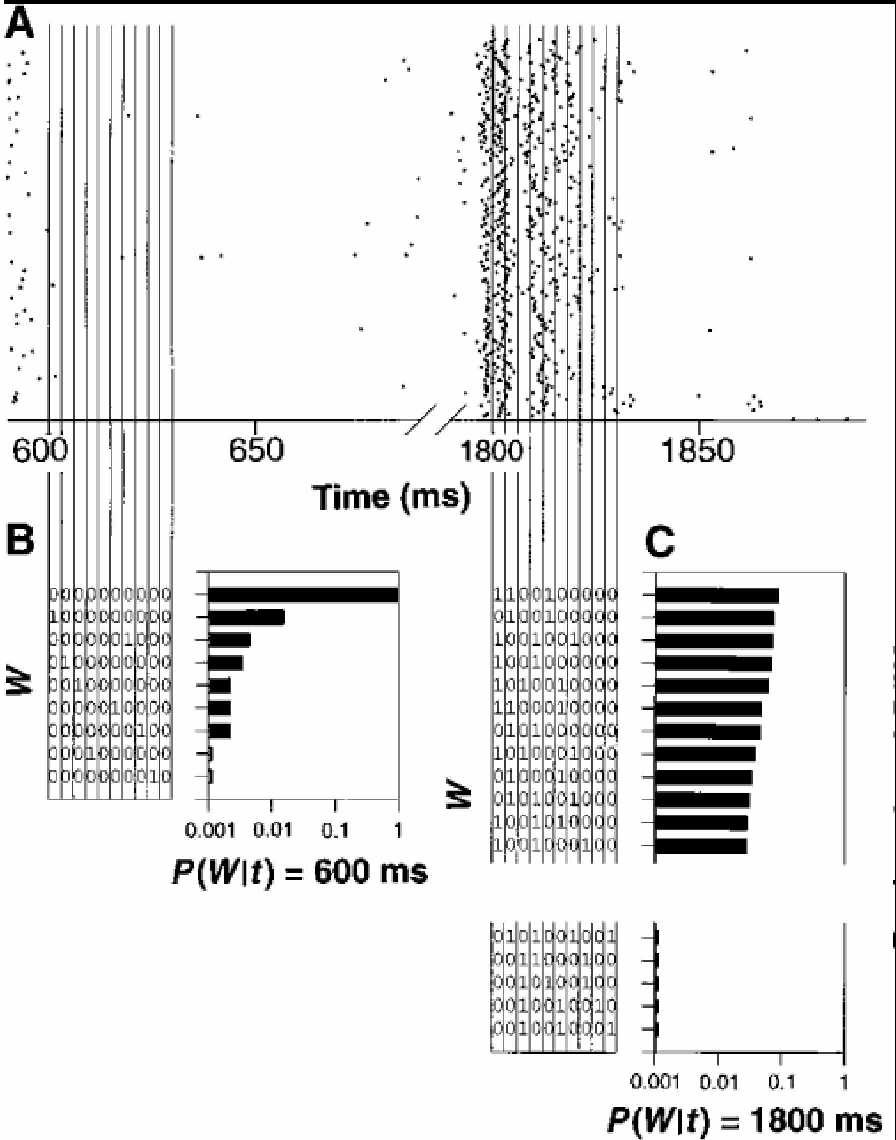
Recordings : VPM thalamus





Montemurro et al (2007) *J Neurophysiol*

Strong et al (1998)



Standard:

$I(S;R) = H(R) – H(R|S)$

$H(R|S) = \Sigma_{\mathbf{s}} P[\mathbf{s}]\ H[R|\mathbf{s}]$

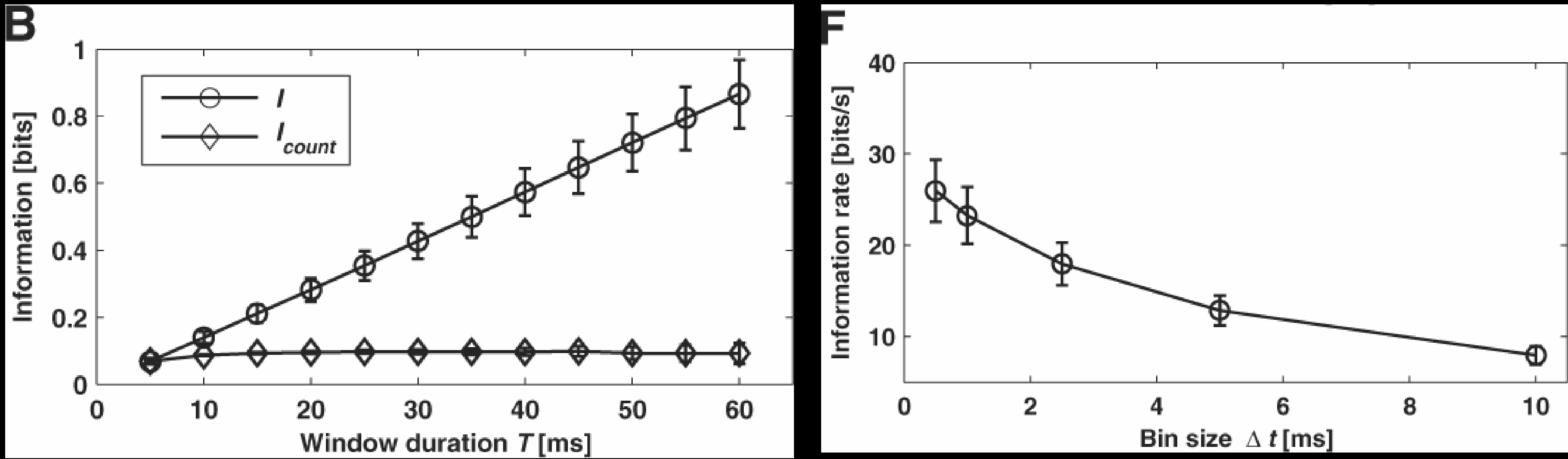
$H(R|\mathbf{s}) = -\Sigma_{\mathbf{r}}\ P[\mathbf{r}|\mathbf{s}]\ \log_2 P[\mathbf{r}|\mathbf{s}]$

Strong et al. (1998):

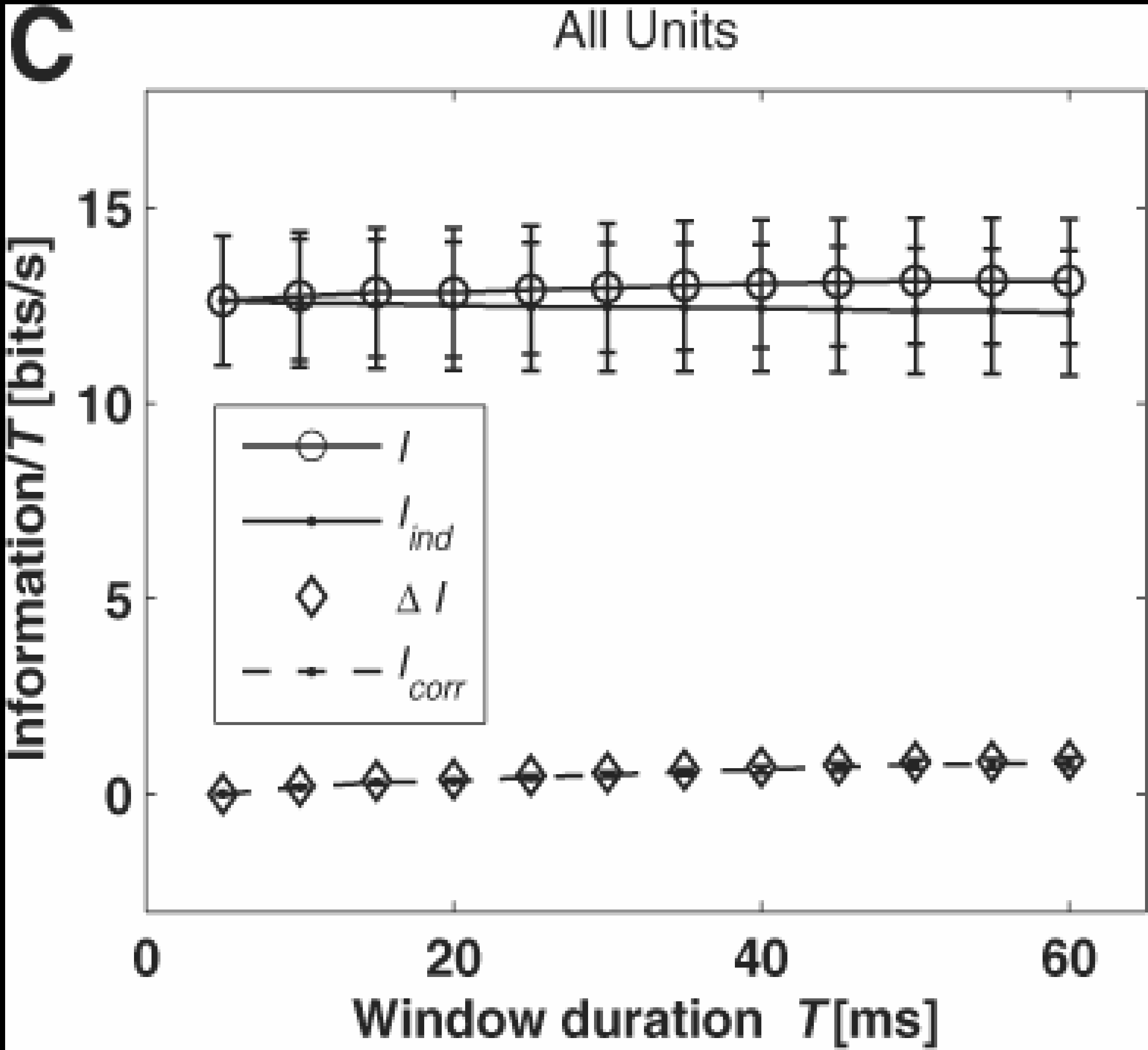
$P[\mathbf{r}|\mathbf{s}] \rightarrow P[\mathbf{r}|t]$

$H(R|\mathbf{s}) \rightarrow H(R|t)$

$H(R|S) \rightarrow \langle H(R_t)\rangle_t$



Montemurro et al (2007) *J Neurophysiol*



Montemurro et al (2007) *J Neurophysiol*

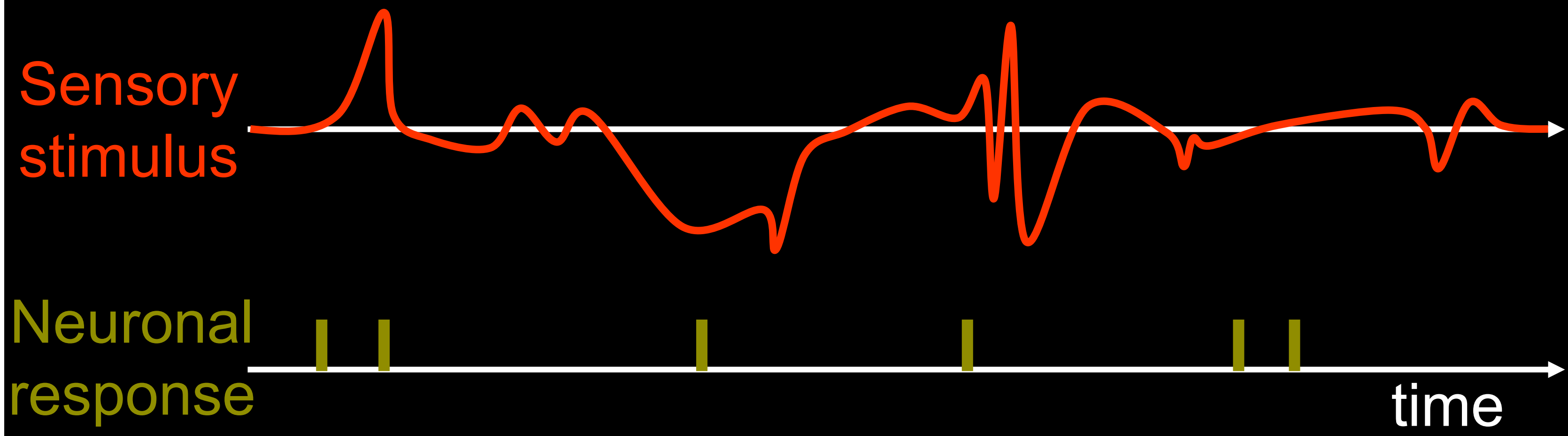
Summary: whisker-syntax

- Millisecond precise spike timing
- Major syntactic element of code is time-dependent firing rate (PSTH)
- Temporal noise correlations have small effect on coding

Role of correlations in other systems

- Retina:
  - Nirenberg et al (2001) Nature
  - Puchalla et al (2005) Neuron
- Primary visual cortex:
  - Golledge et al (2003) Neuroreport
  - Montani et al (2007) J Neurosci
- Primary somatosensory cortex:
  - Petersen et al (2001) Neuron

Neurosemantics



Assuming, $P[(r_1, r_2, \ldots, r_N)|\mathbf{s}] \approx P[r_1|\mathbf{s}]\, P[r_2|\mathbf{s}] \ldots P[r_N|\mathbf{s}]$

$P(r_t|\mathbf{s}_t) = ?$

$\mathbf{s}_t = (s_{t-L}, s_{t-L+1}, \ldots, s_t)$

Possible approaches

- LNP model

  $P[r_t|\mathbf{s}_t] = g(k_t)$

  $k_t = \mathbf{s}_t \bullet \mathbf{f}$

- Multi-feature LNP model

  $P[r_t|\mathbf{s}_t] = g(k^1{}_t, k^2{}_t, \ldots)$

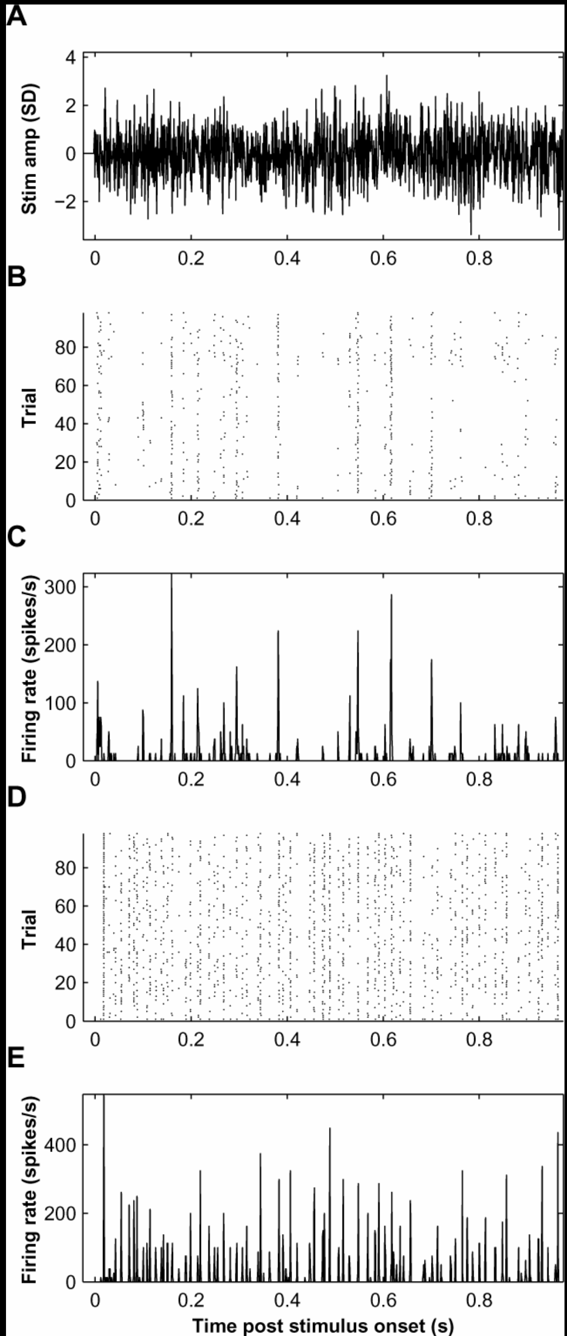
  $k^i{}_t = \mathbf{s}_t \bullet \mathbf{f}^i$

- Spike-feedback model

  $P[r_t|\mathbf{s}_t] = g(k_t, m_t)$

  $m_t = (r_{t-L}, r_{t-L+1}, \ldots, r_{t-1}) \bullet \mathbf{h}$

- Adaptive models



AIM

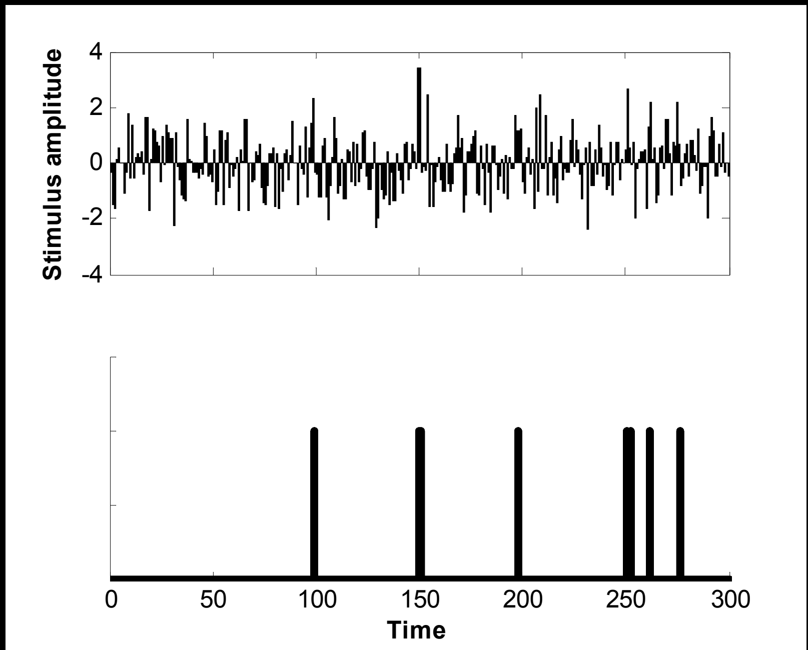
What temporal (kinetic) features of the whisker stimulus are neurons in VPM sensitive to?

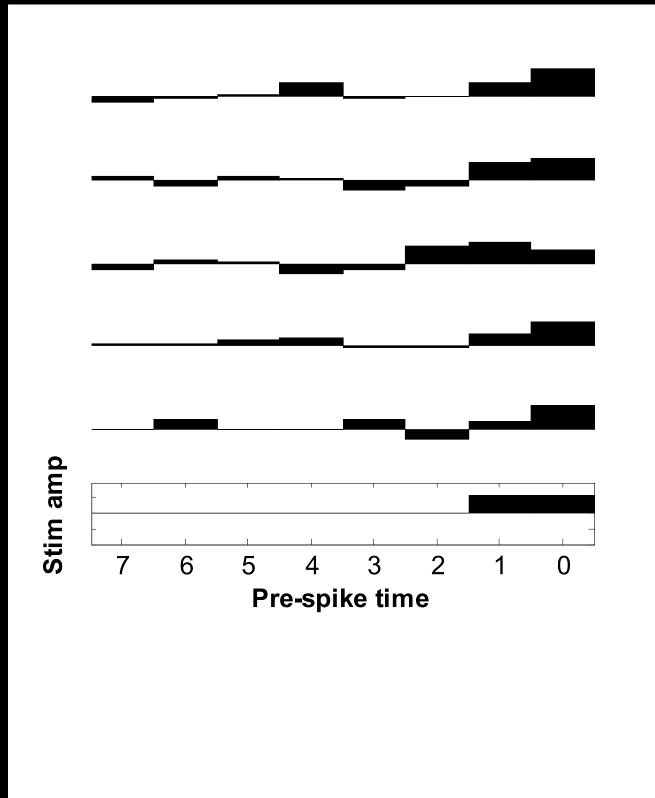
Petersen et al (2008) *Neuron*

Special cases

		kernels
Position	$r(t) = \mathrm{g}[s(t)]$	**f = [1]**
Velocity	$r(t) = \mathrm{g}[s(t) - s(t - \delta t)]$	**f = [-1,1]**
Acceler.	$r(t) = \mathrm{g}\left[\frac{1}{2}s(t) - s(t - \delta t) + \frac{1}{2}s(t - 2\delta t)\right]$	**f = [-½,1, -½]**

Fitting the kernel: Spike-triggered averaging





Theory

For the LNP model

$P[r_t|\mathbf{s}_t] = g(k_t),\ k_t = \mathbf{s}_t \bullet \mathbf{f}$

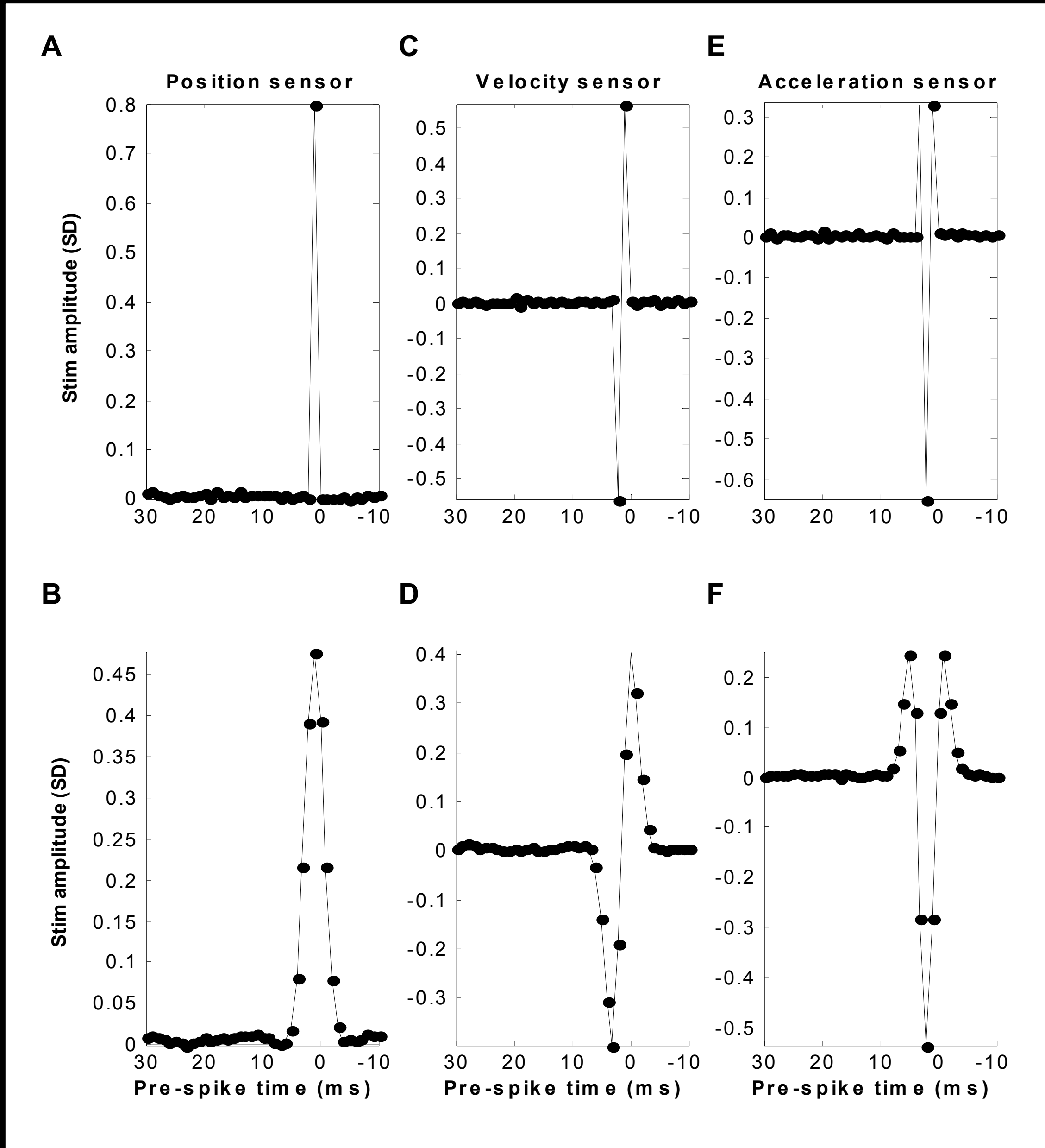
If,

- stimulus is white noise
- $\langle g[k] \rangle \neq \langle g[-k] \rangle$

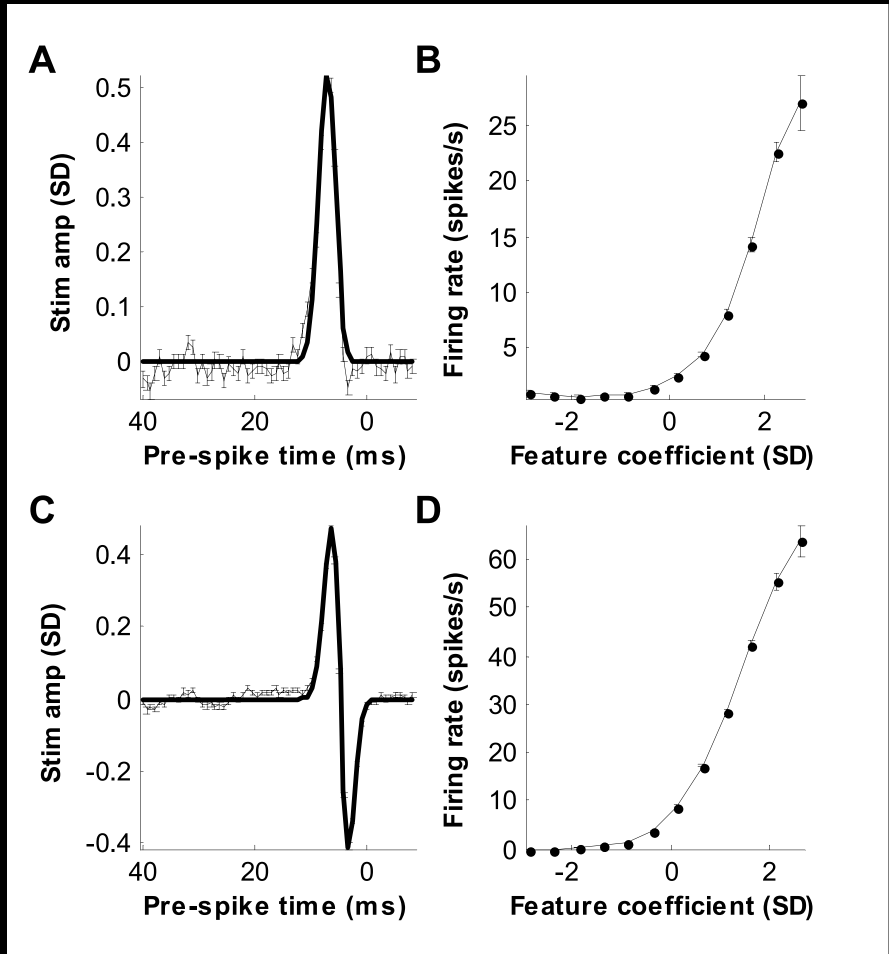
Then,

Spike-triggered average is unbiased estimate of kernel **f**

Paninski (2003) *Network*



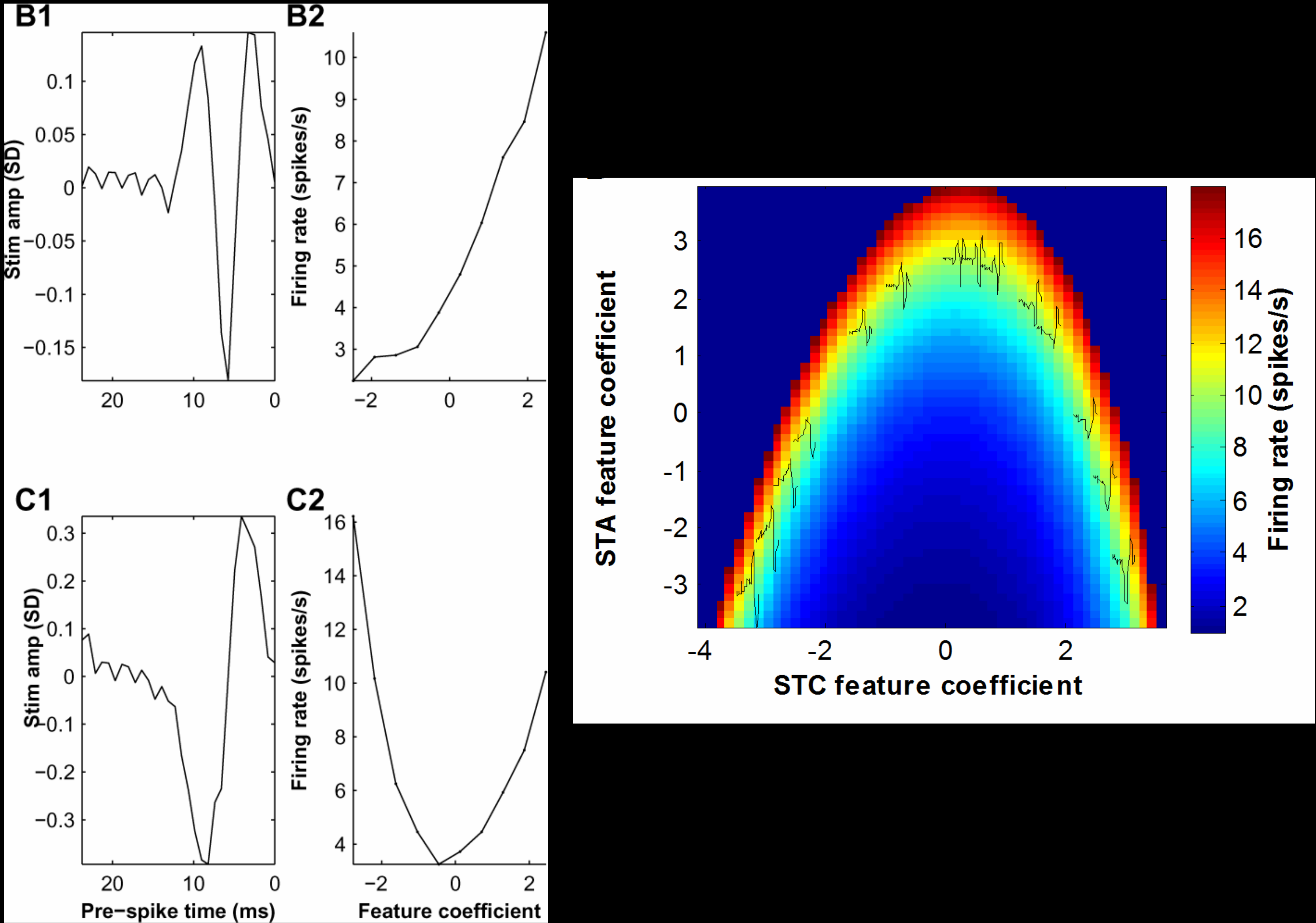
Petersen et al (2008) *Neuron*



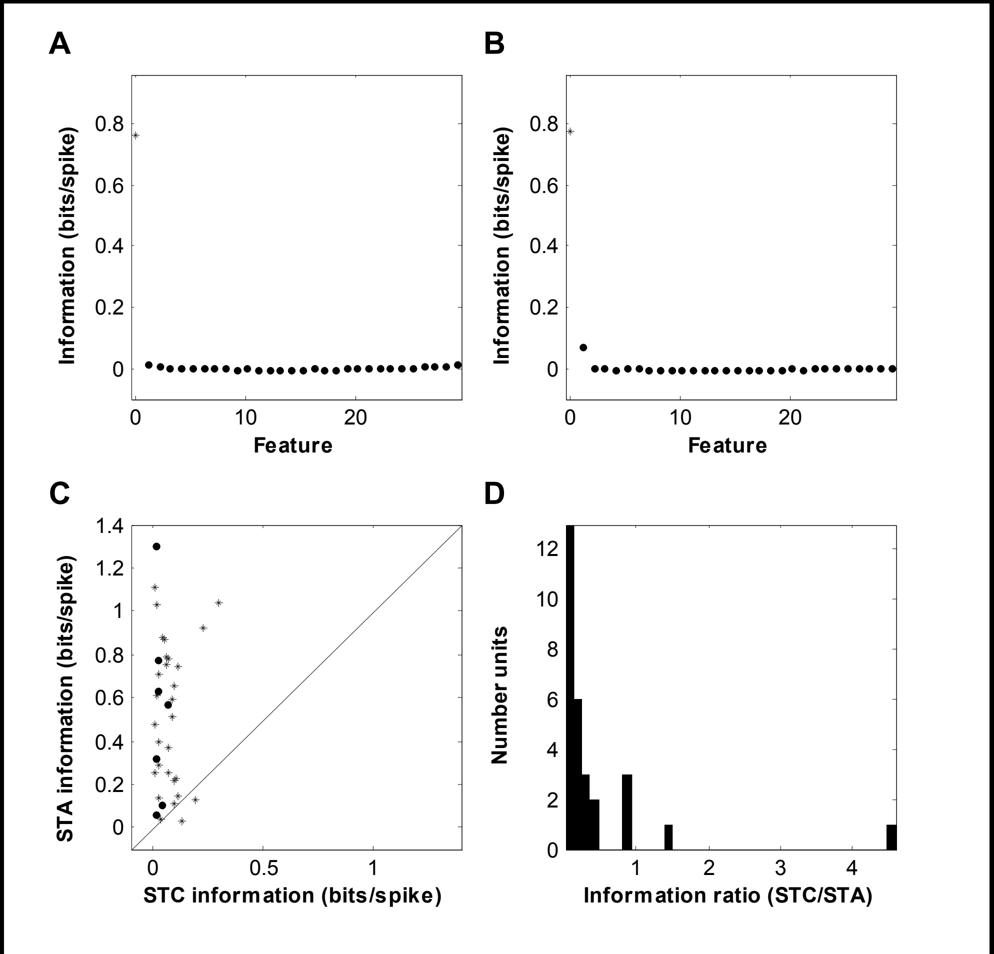
Petersen et al (2008) *Neuron*

How accurate is the LNP model?

- Non-linear approaches:
  - Maximum likelihood fitting of spike feedback (Paninski) model
  - Spike-triggered covariance analysis



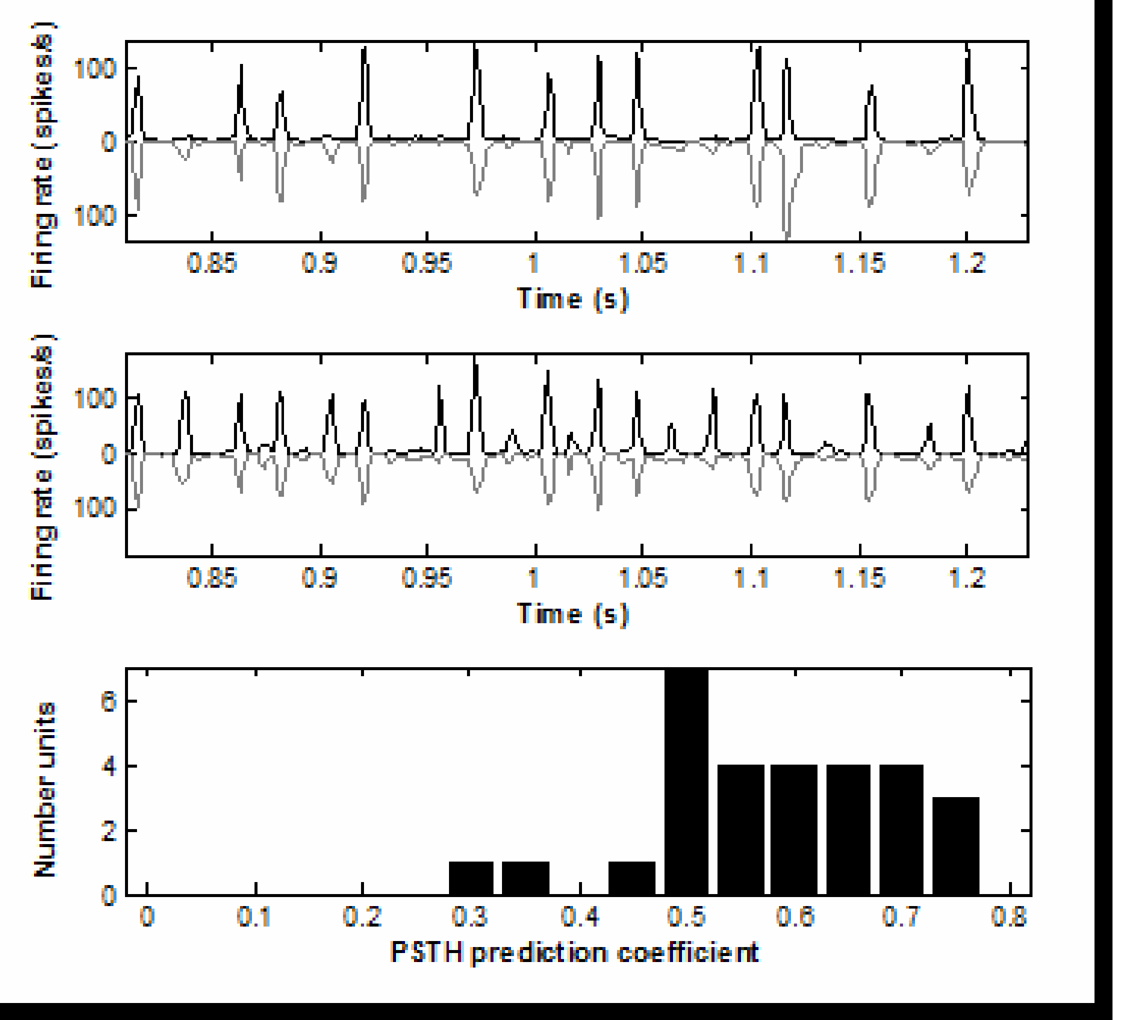
Petersen et al (2008) *Neuron*



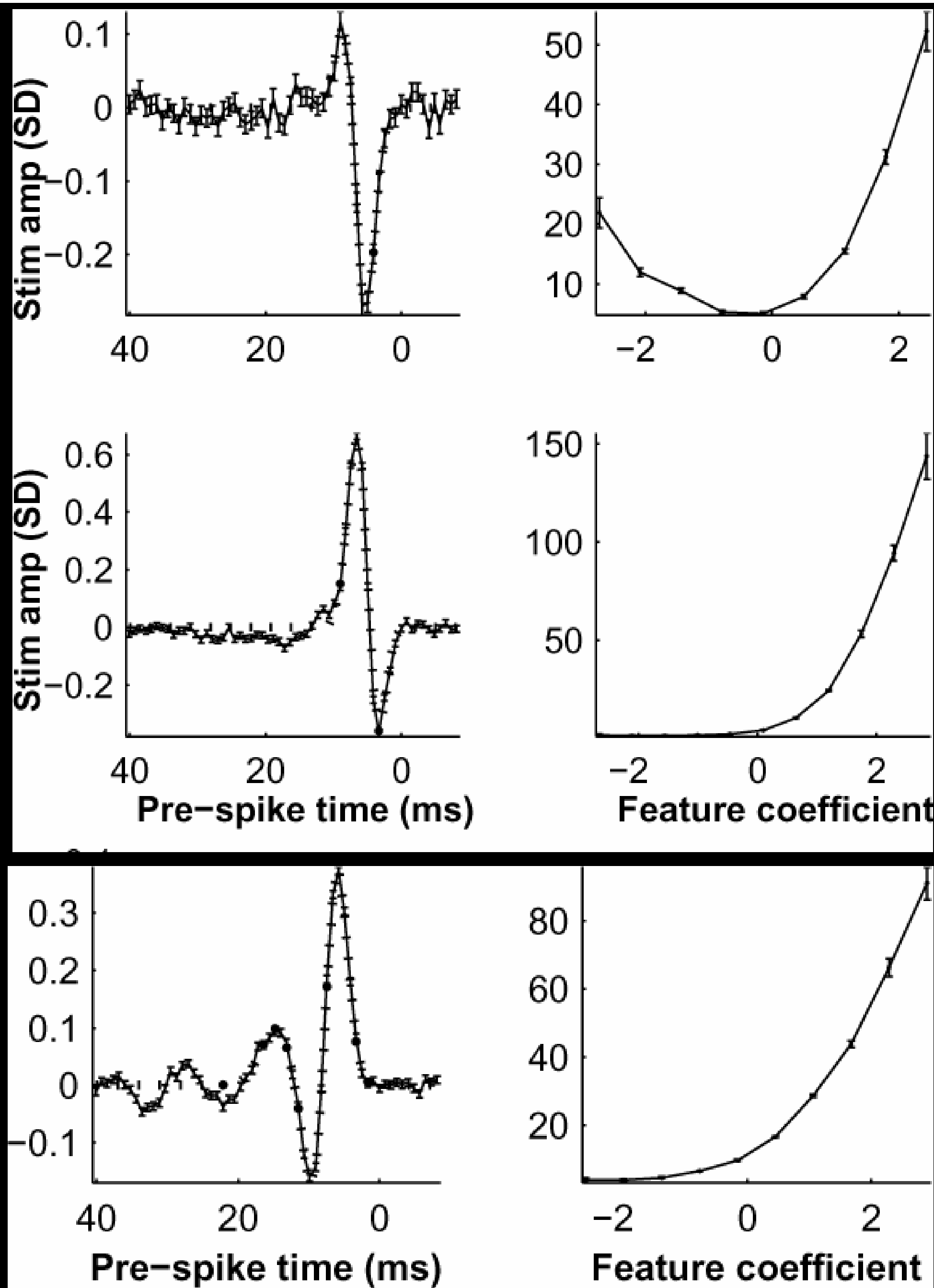
Petersen et al (2008) *Neuron*

PSTH prediction using LNP model:

$P[r_t|\mathbf{s}_t] = g(k_t),\ k_t = \mathbf{s}_t \bullet \mathbf{f}$

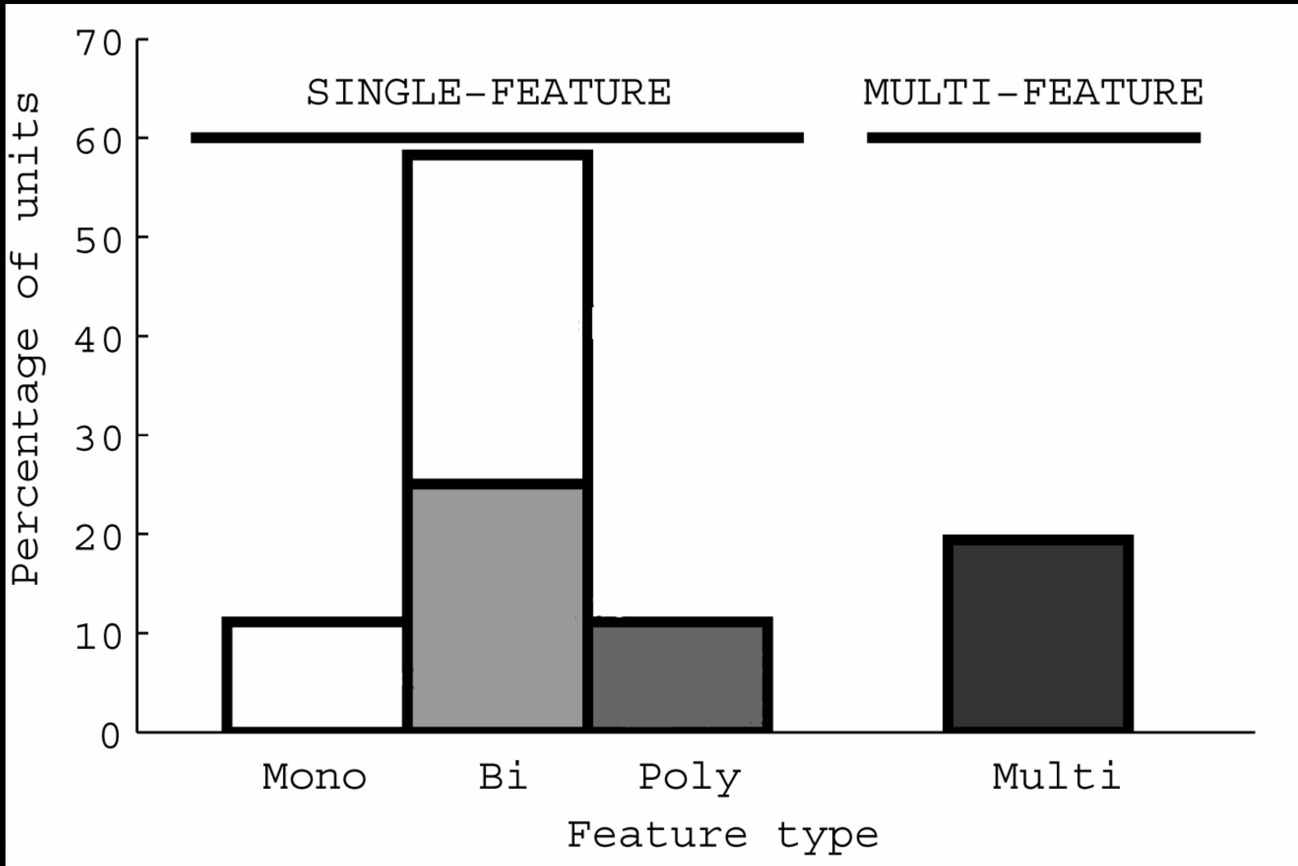


Petersen et al (2008) *Neuron*



Petersen et al (2008) *Neuron*

Distributed code for whisker motion



Petersen et al (2008) *Neuron*

Summary

- Syntax-semantics approach:
  - What is the basic ‘syntactic’ element of the response?
    - Information theory
  - What stimulus features does that syntactic element signify?
    - White noise analysis

- Whisker syntax:
  - Millisecond precise spike timing
  - Spikes code independently

- Whisker semantics:
  - Accuracy of LNP model
  - Distributed code for whisker motion