Nonlinear approaches to neural system identification

September 5, 2014

Lucas Theis

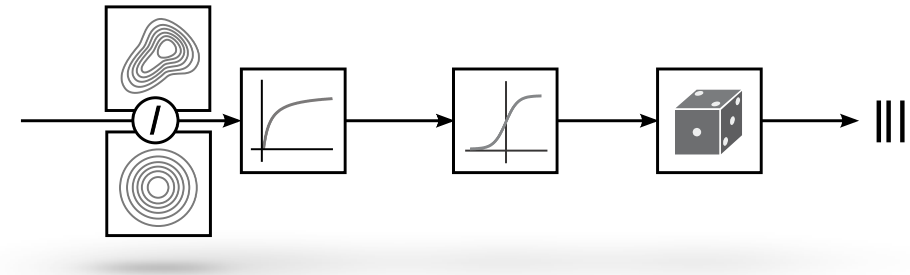
CIN Werner Reichardt
Centre for Integrative Neuroscience

bccn tübingen

Overview

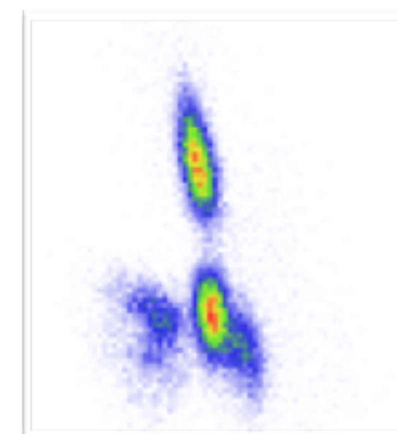
1 **Beyond generalized linear models**

The spike-triggered mixture model



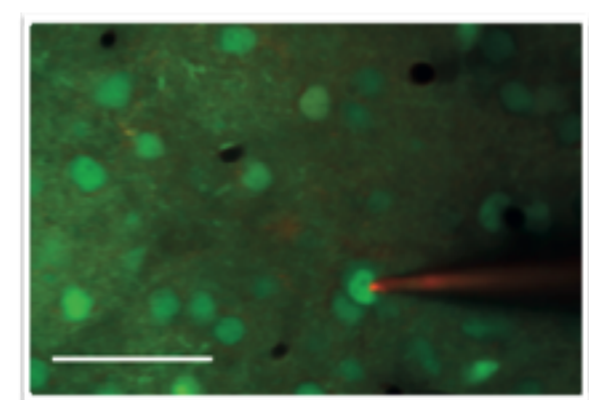
2 **From stimulus to spikes**

Modeling the stimulus-response relationship of whisker receptors

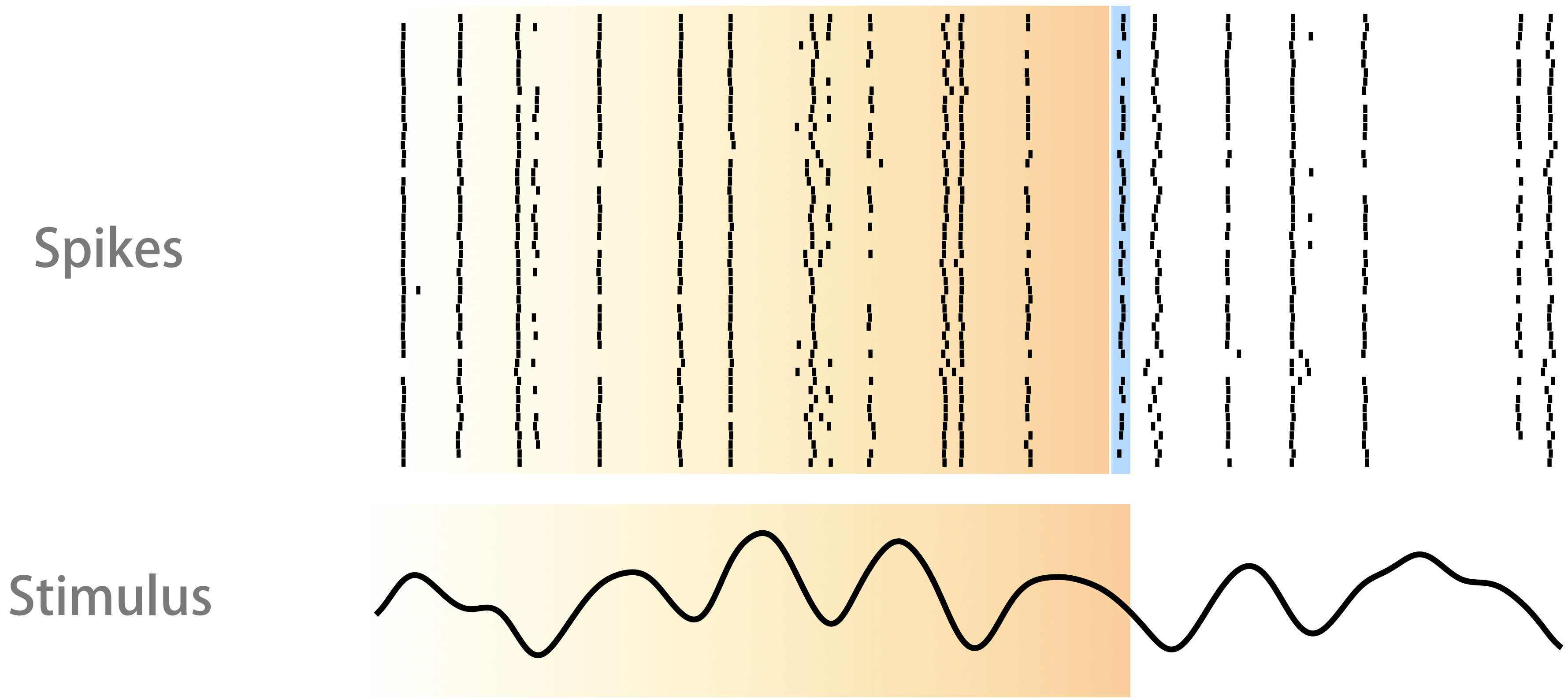


3 **From calcium to spikes**

Improved spike estimation for two-photon imaging

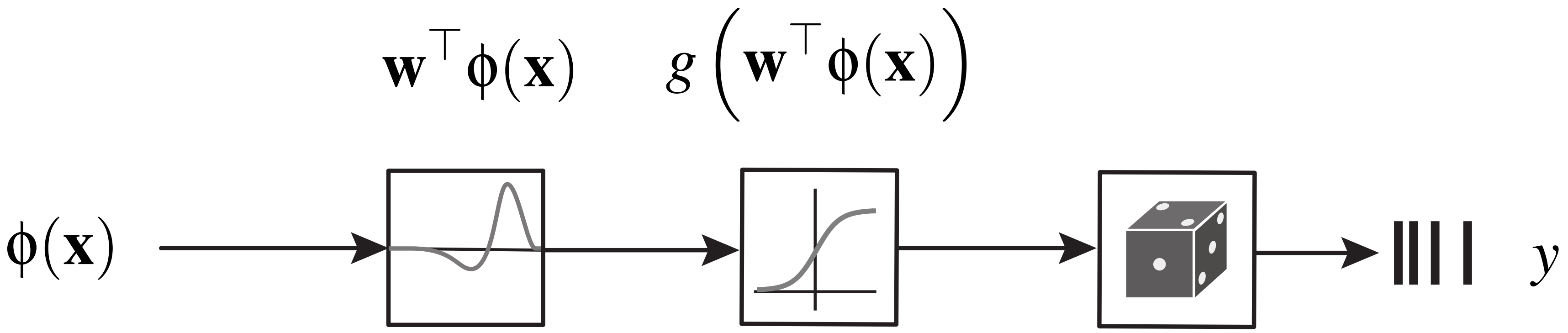


Neural system identification

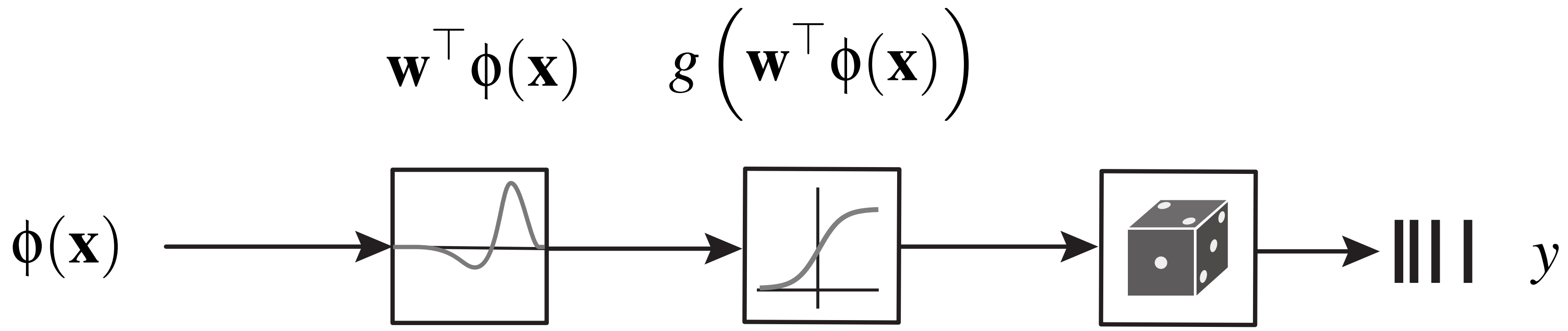


$$p(\text{spike} \mid \mathbf{x})$$

Generalized linear models



Generalized linear models

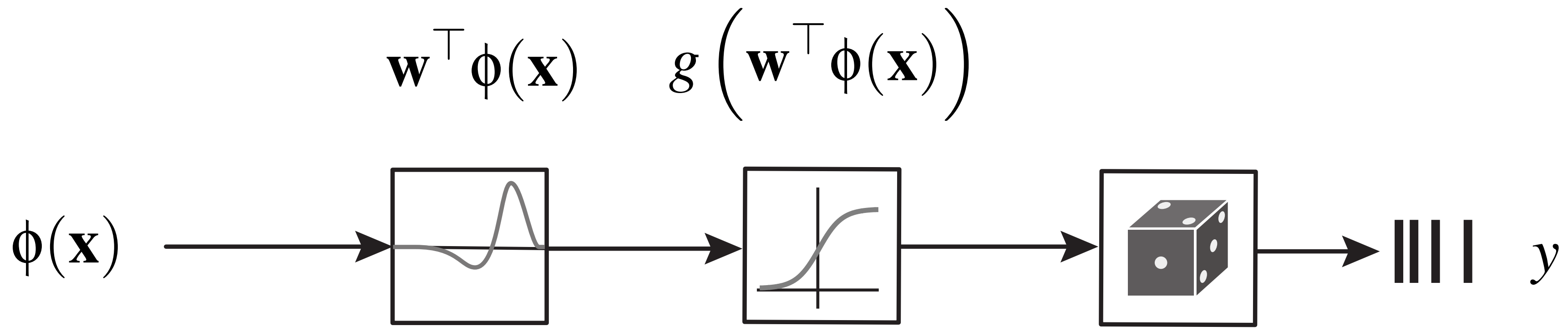


Linear-nonlinear-Poisson

$$p(y \mid \mathbf{x}) = \frac{\lambda^y}{y!} e^{-\lambda}$$

$$\lambda = \exp(\mathbf{w}^\top \phi(\mathbf{x}))$$

Generalized linear models



Linear-nonlinear-Poisson

$$p(y \mid \mathbf{x}) = \frac{\lambda^y}{y!}e^{-\lambda}$$

$$\lambda = \exp(\mathbf{w}^\top\phi(\mathbf{x}))$$

Linear-nonlinear-Bernoulli

$$p(y \mid \mathbf{x}) = r^y(1-r)^{1-y}$$

$$r = \left(1+\exp\left(-\mathbf{w}^\top\phi(\mathbf{x})\right)\right)^{-1}$$

Generalized linear models

$$p(\mathbf{x} \mid y = 1) = h(\mathbf{x}) \exp\left(\boldsymbol{\eta}_1^\top \phi(\mathbf{x}) - A(\boldsymbol{\eta}_1)\right)$$

$$p(\mathbf{x} \mid y = 0) = h(\mathbf{x}) \exp\left(\boldsymbol{\eta}_0^\top \phi(\mathbf{x}) - A(\boldsymbol{\eta}_0)\right)$$

$$p(y = 1 \mid \mathbf{x}) = \sigma\left(\mathbf{w}^\top \phi(\mathbf{x}) + b\right)$$

$$\mathbf{w} = \eta_1 - \eta_0$$

$$b = A(\eta_0) - A(\eta_1) + \log \frac{p(y=1)}{p(y=0)}$$

Generalized linear models

$$p(\mathbf{x} \mid y=1) = h(\mathbf{x}) \exp\left(\boldsymbol{\eta}_1^\top \phi(\mathbf{x}) - A(\boldsymbol{\eta}_1)\right)$$

$$p(\mathbf{x} \mid y=0) = h(\mathbf{x}) \exp\left(\boldsymbol{\eta}_0^\top \phi(\mathbf{x}) - A(\boldsymbol{\eta}_0)\right)$$

$$p(y=1 \mid \mathbf{x}) = \sigma\left(\mathbf{w}^\top \phi(\mathbf{x}) + b\right)$$

$$\mathbf{w} = \eta_1 - \eta_0$$

$$b = A(\eta_0) - A(\eta_1) + \log \frac{p(y=1)}{p(y=0)}$$

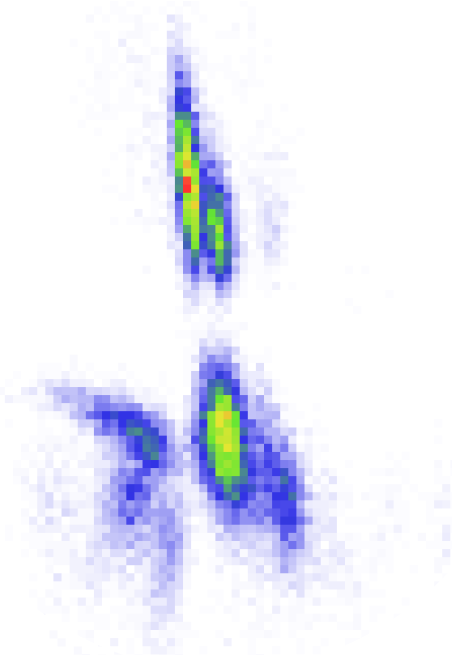
Generalized linear models

$$p(\mathbf{x} \mid y = 1) = h(\mathbf{x}) \exp\left(\boldsymbol{\eta}_1^\top \phi(\mathbf{x}) - A(\boldsymbol{\eta}_1)\right)$$
$$p(\mathbf{x} \mid y = 0) = h(\mathbf{x}) \exp\left(\boldsymbol{\eta}_0^\top \phi(\mathbf{x}) - A(\boldsymbol{\eta}_0)\right)$$

$$p(y = 1 \mid \mathbf{x}) = \sigma\left(\mathbf{w}^\top \phi(\mathbf{x}) + b\right)$$

$$\mathbf{w} = \boldsymbol{\eta}_1 - \boldsymbol{\eta}_0$$
$$b = A(\boldsymbol{\eta}_0) - A(\boldsymbol{\eta}_1) + \log \frac{p(y = 1)}{p(y = 0)}$$

Generalized linear models



$$p(\mathbf{x} \mid y=1) = h(\mathbf{x}) \exp\left(\boldsymbol{\eta}_1^\top \phi(\mathbf{x}) - A(\boldsymbol{\eta}_1)\right)$$

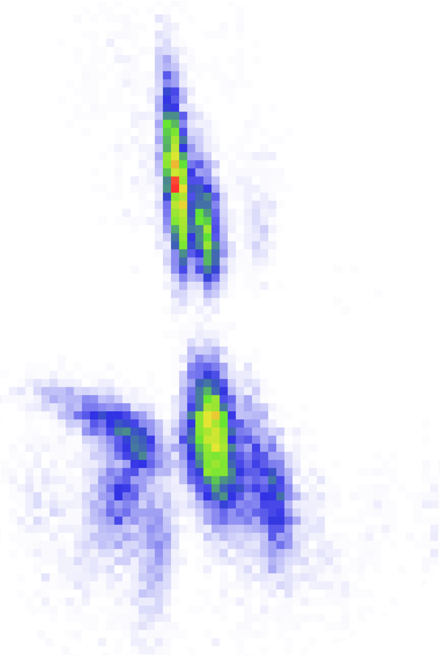
$$p(\mathbf{x} \mid y=0) = h(\mathbf{x}) \exp\left(\boldsymbol{\eta}_0^\top \phi(\mathbf{x}) - A(\boldsymbol{\eta}_0)\right)$$

$$p(y=1 \mid \mathbf{x}) = \sigma\left(\mathbf{w}^\top \phi(\mathbf{x}) + b\right)$$

$$\mathbf{w} = \boldsymbol{\eta}_1 - \boldsymbol{\eta}_0$$

$$b = A(\boldsymbol{\eta}_0) - A(\boldsymbol{\eta}_1) + \log \frac{p(y=1)}{p(y=0)}$$

Spike-triggered mixture model



$$p(\mathbf{x} \mid y = 1) = \sum_k \pi_k \mathcal{N}(\mathbf{x}; \boldsymbol{\mu}_{1k}, \boldsymbol{\Sigma}_{1k})$$

$$p(\mathbf{x} \mid y = 0) = \mathcal{N}(\mathbf{x}; \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$$

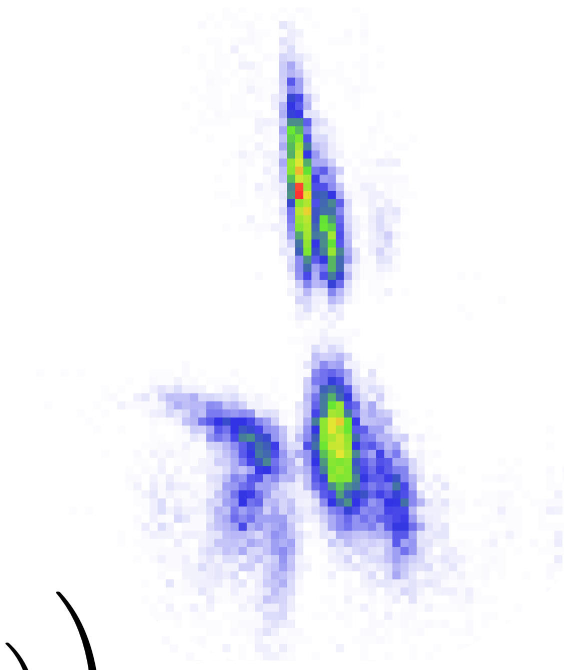
$$p(y = 1 \mid \mathbf{x}) = \sigma\left(\log \sum_k \exp\left(\mathbf{x}^\top \mathbf{Q}_k \mathbf{x} + \mathbf{w}_k^\top \mathbf{x} + b_k\right)\right)$$

Theis et al. (2013)

Spike-triggered mixture model

$$p(\mathbf{x} \mid y = 1) = \sum_k \pi_k \mathcal{N}(\mathbf{x}; \boldsymbol{\mu}_{1k}, \boldsymbol{\Sigma}_{1k})$$

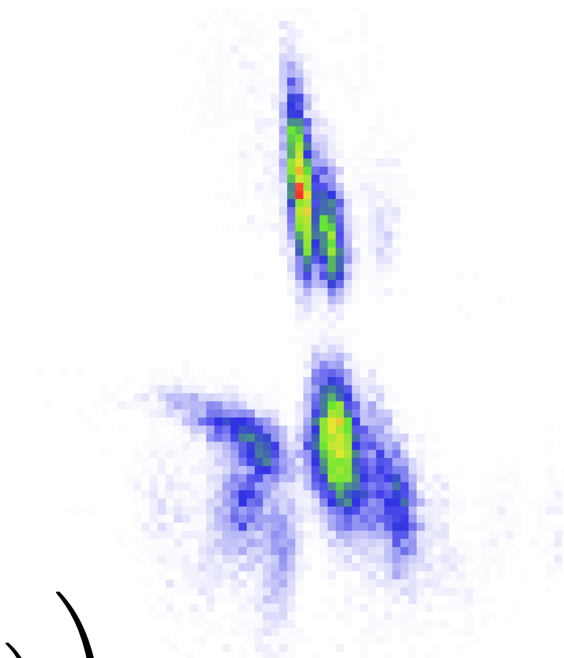
$$p(\mathbf{x} \mid y = 0) = \mathcal{N}(\mathbf{x}; \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$$



$$p(y = 1 \mid \mathbf{x}) = \sigma\left(\log \sum_k \exp\left(\mathbf{x}^\top \mathbf{Q}_k \mathbf{x} + \mathbf{w}_k^\top \mathbf{x} + b_k\right)\right)$$

Theis et al. (2013)

Spike-triggered mixture model



$$p(\mathbf{x} \mid y = 1) = \sum_k \pi_k \mathcal{N}(\mathbf{x}; \boldsymbol{\mu}_{1k}, \boldsymbol{\Sigma}_{1k})$$

$$p(\mathbf{x} \mid y = 0) = \mathcal{N}(\mathbf{x}; \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$$

$$p(y = 1 \mid \mathbf{x}) = \sigma\left(\log \sum_k \exp\left(\mathbf{x}^\top \mathbf{Q}_k \mathbf{x} + \mathbf{w}_k^\top \mathbf{x} + b_k\right)\right)$$

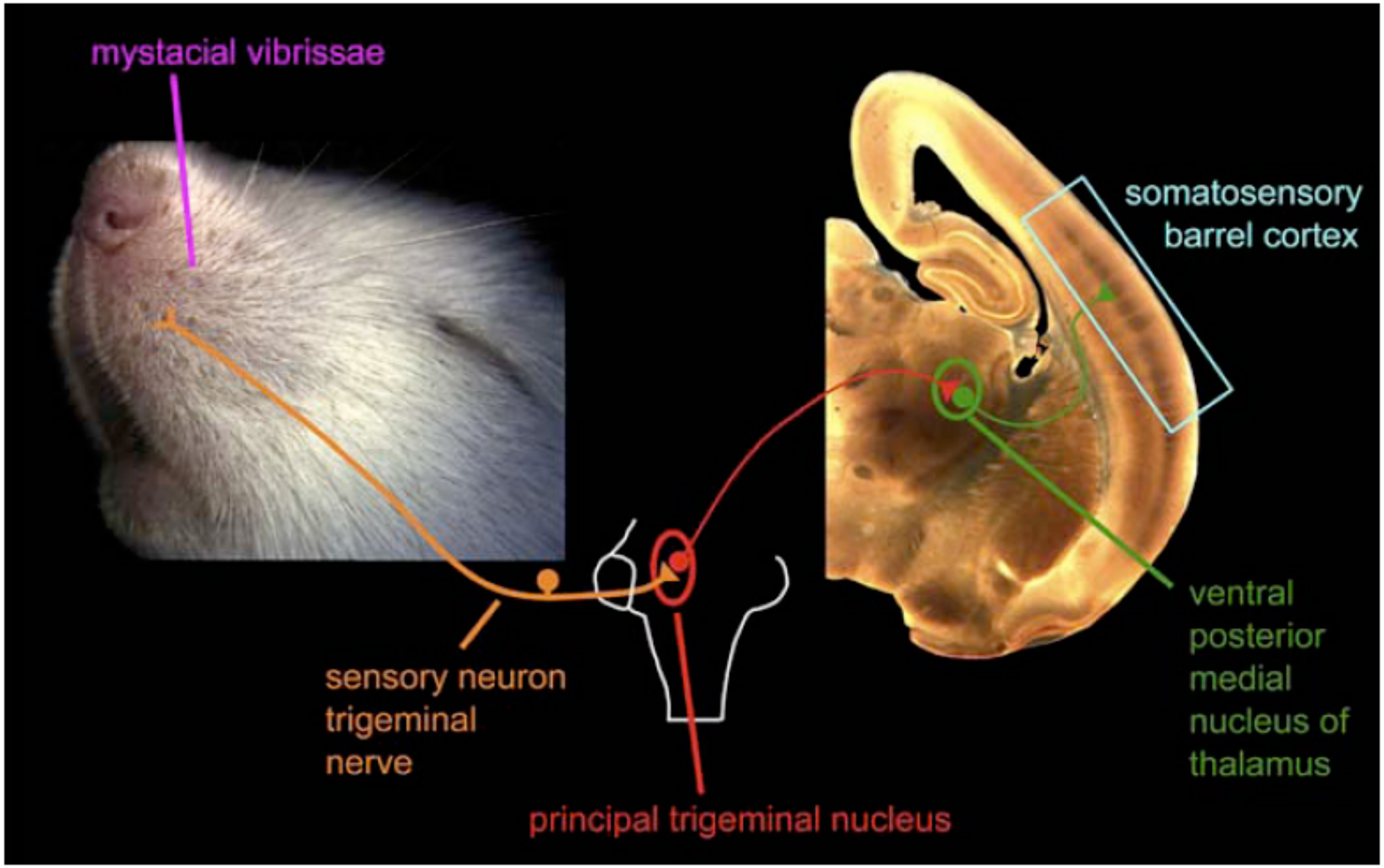
$$\mathbf{Q}_k = \sum_m \alpha_{km} \mathbf{u}_m \mathbf{u}_m^\top$$

$$p(y = 1 \mid \mathbf{x}) = \sigma\left(\log \sum_k \exp\left(\alpha_{km}\left(\mathbf{u}_m^\top \mathbf{x}\right)^2 + \mathbf{w}_k^\top \mathbf{x} + b_k\right)\right)$$

Theis et al. (2013)

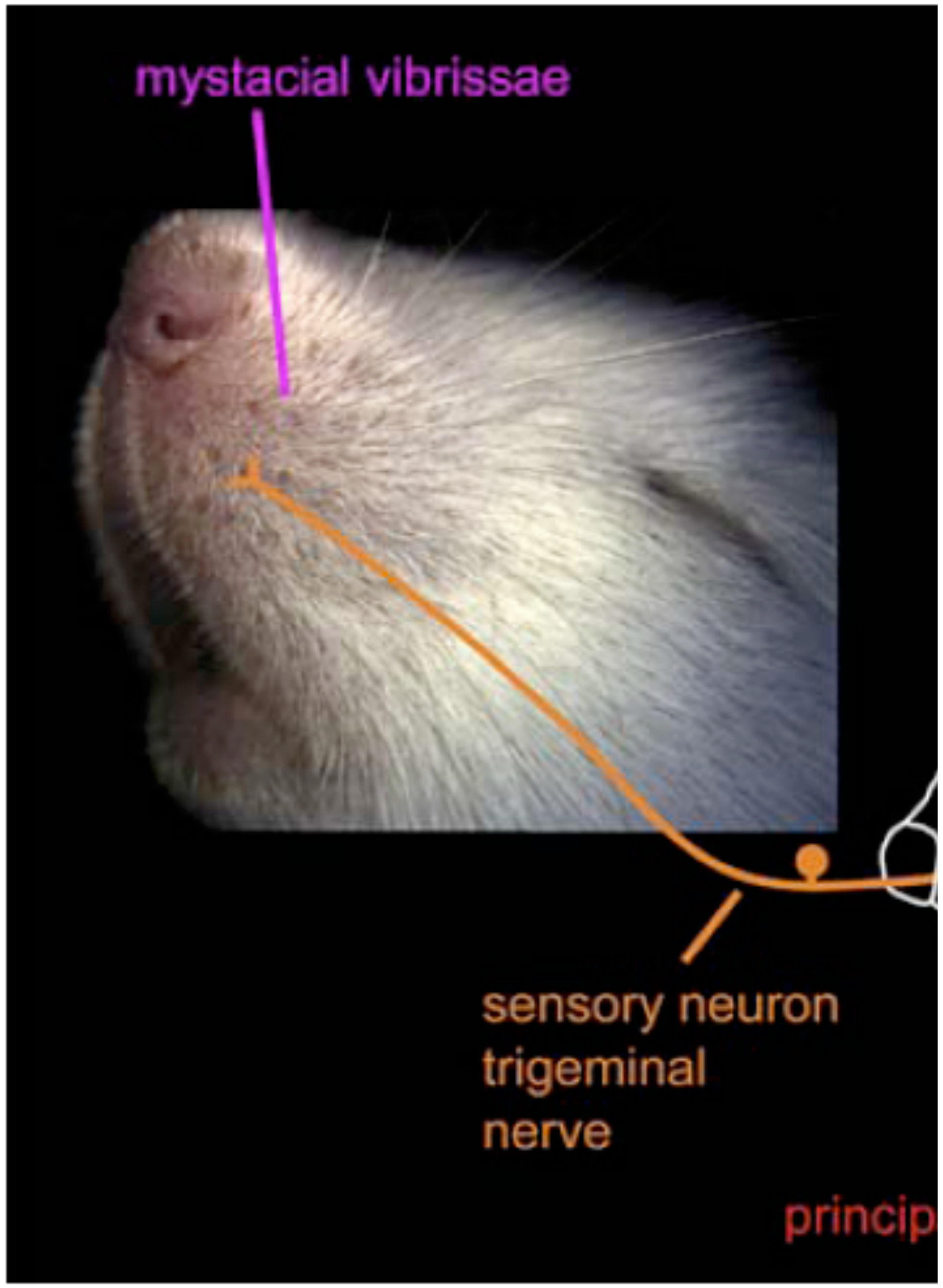
From stimulus to spikes

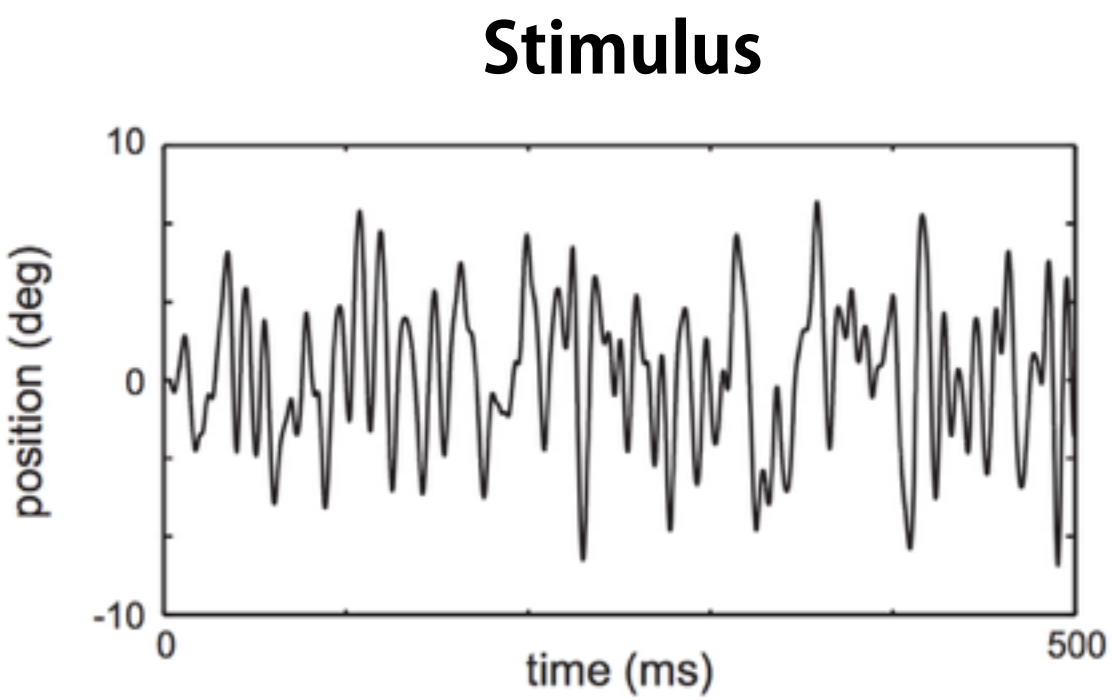
Trigeminal ganglion cells



Chagas et al. (2013)

Trigeminal ganglion cells

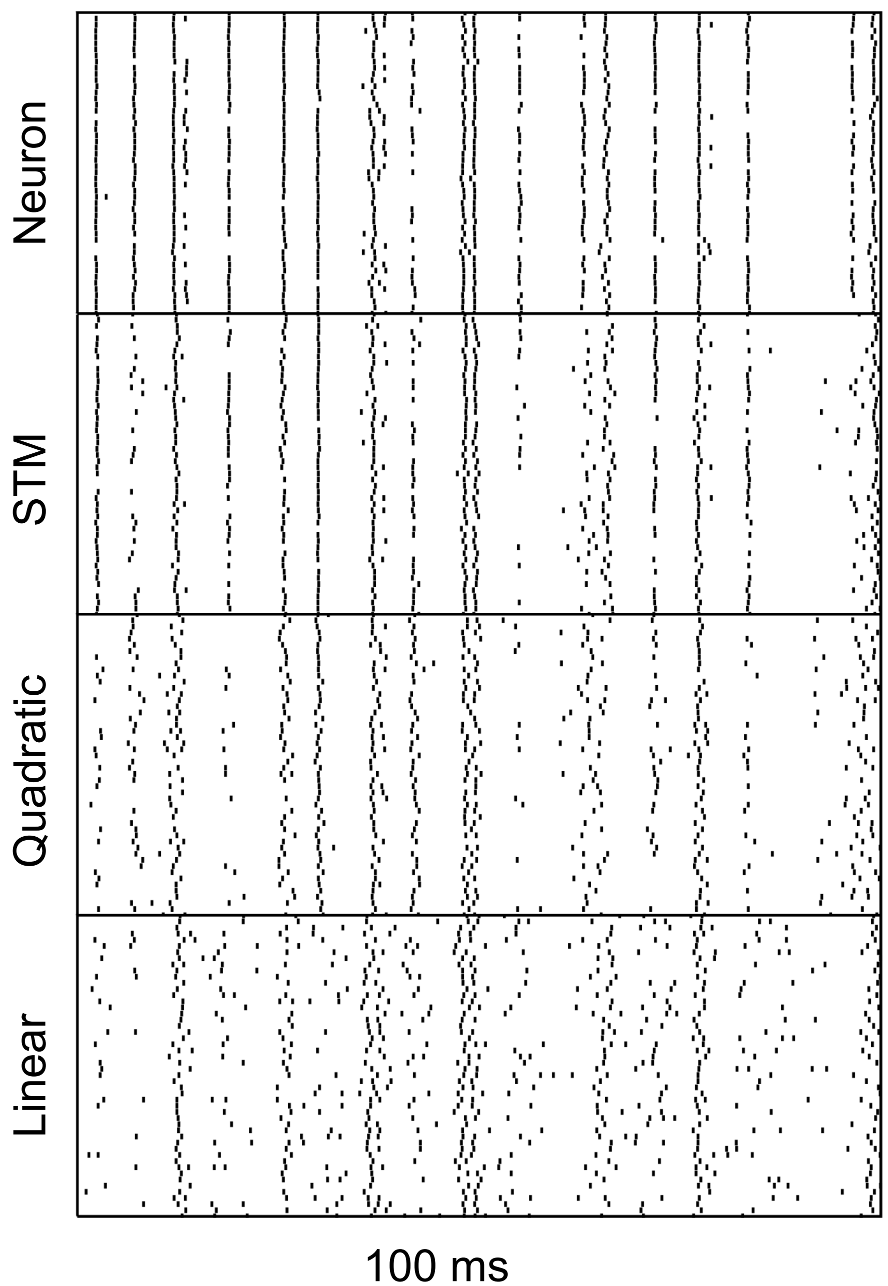




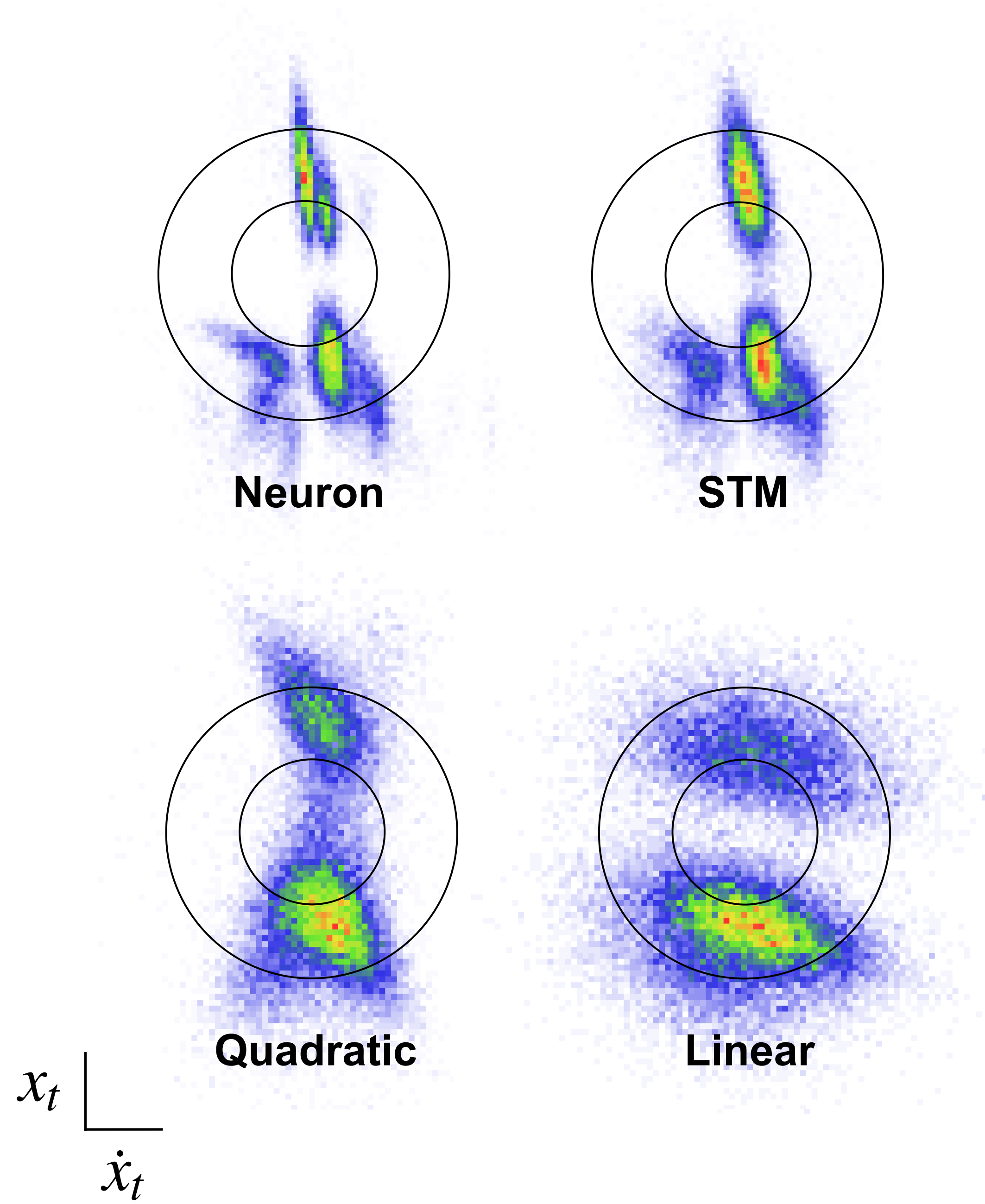
Chagas et al. (2013)

Qualitative evaluation

Spike trains



Spike-triggered distribution

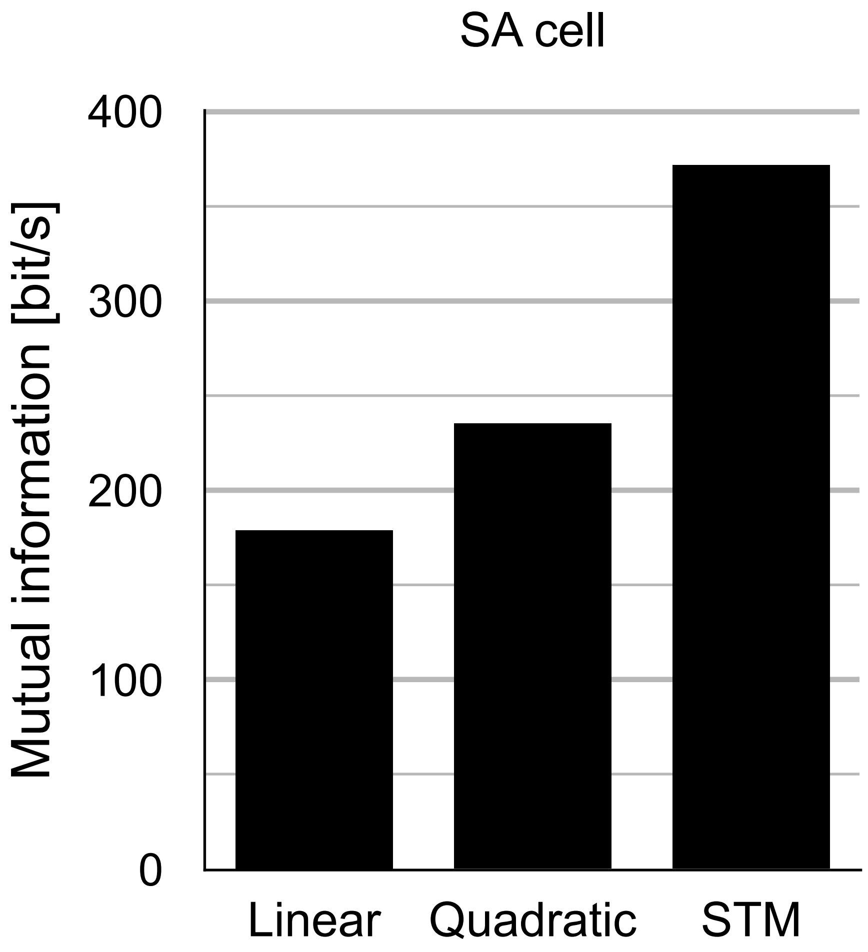


Theis et al. (2013), Chagas et al. (2013)

Quantitative comparison

$$I[y, \mathbf{x}] = H[y] - H[y \mid \mathbf{x}]$$
$$\geq H[y] + E[\log p(y \mid \mathbf{x})]$$

Expected log-likelihood

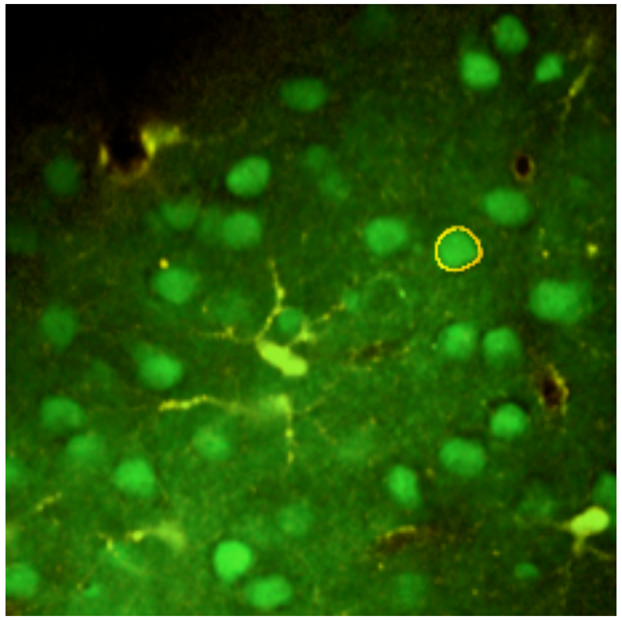


Theis et al. (2013), Chagas et al. (2013)

From calcium to spikes

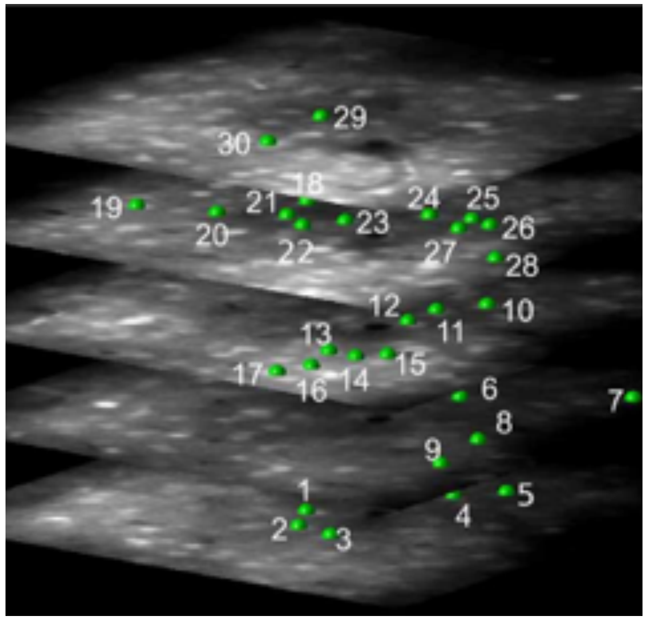
Two-photon calcium imaging

Galvanometric

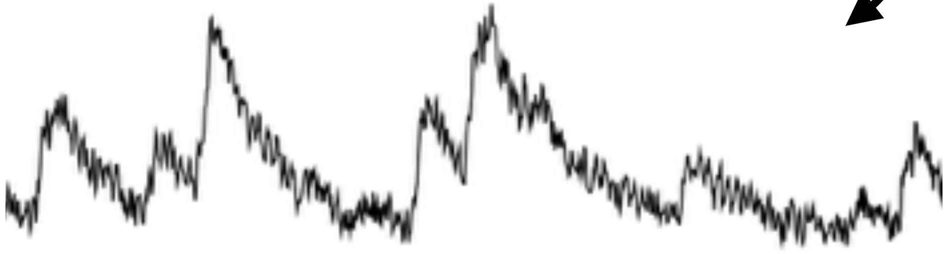


Tomek et al. (2013)

AOD

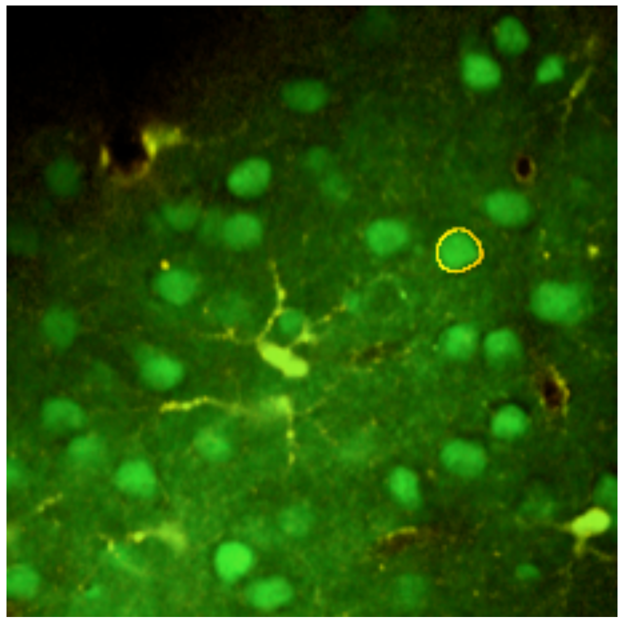


Cotton et al. (2013)



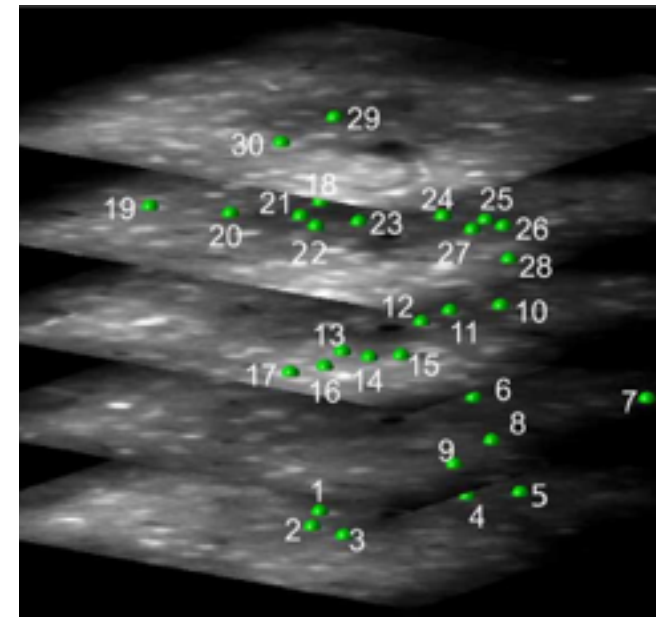
Two-photon calcium imaging

Galvanometric

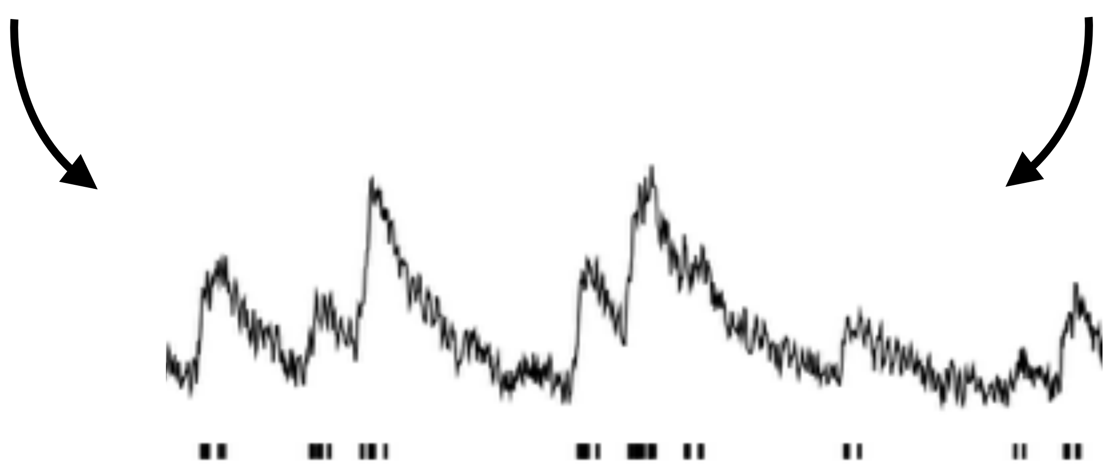


Tomek et al. (2013)

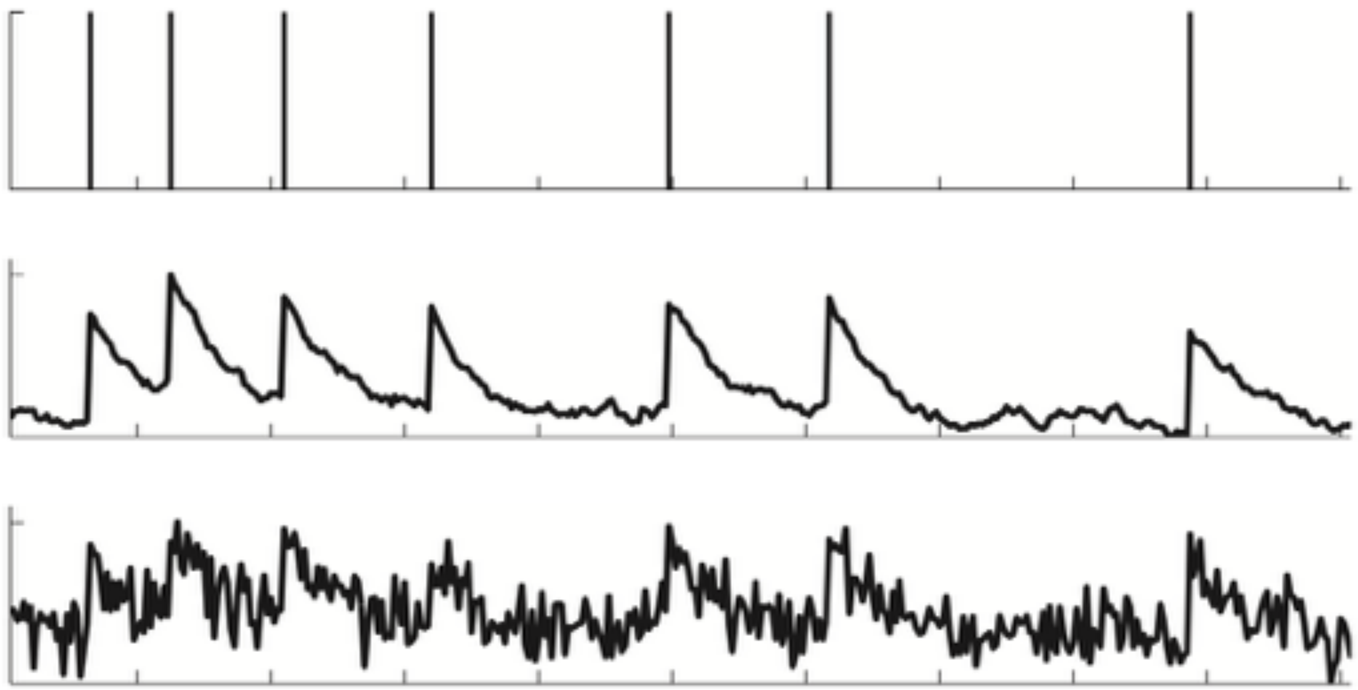
AOD



Cotton et al. (2013)



Two-photon calcium imaging



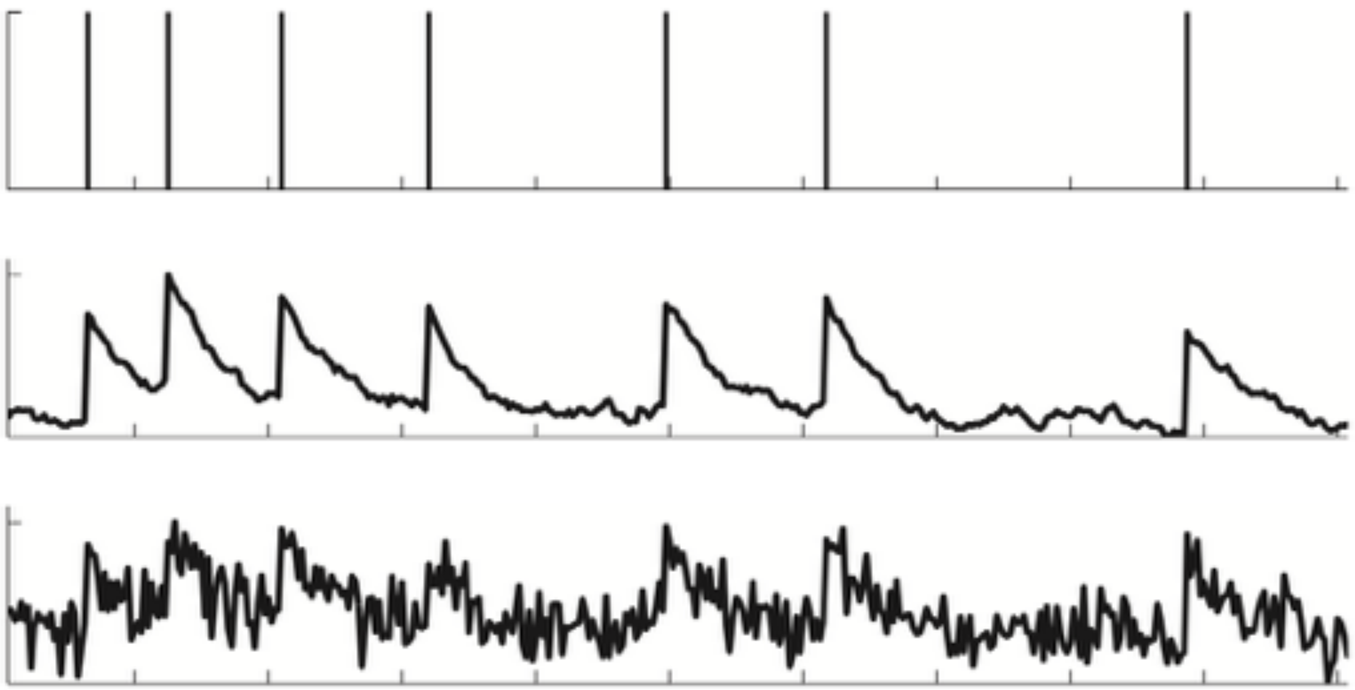
Vogelstein et al. (2010)

$$k_t \sim \text{Poisson}(k_t; \lambda)$$

$$C_t = \gamma C_{t-1} + k_t$$

$$F_t = \alpha C_t + \beta + \sigma \varepsilon$$

Two-photon calcium imaging



Vogelstein et al. (2010)

$$k_t \sim \text{Poisson}(k_t; \lambda)$$

$$C_t = \gamma C_{t-1} + k_t$$

$$F_t = \alpha C_t + \beta + \sigma\varepsilon$$

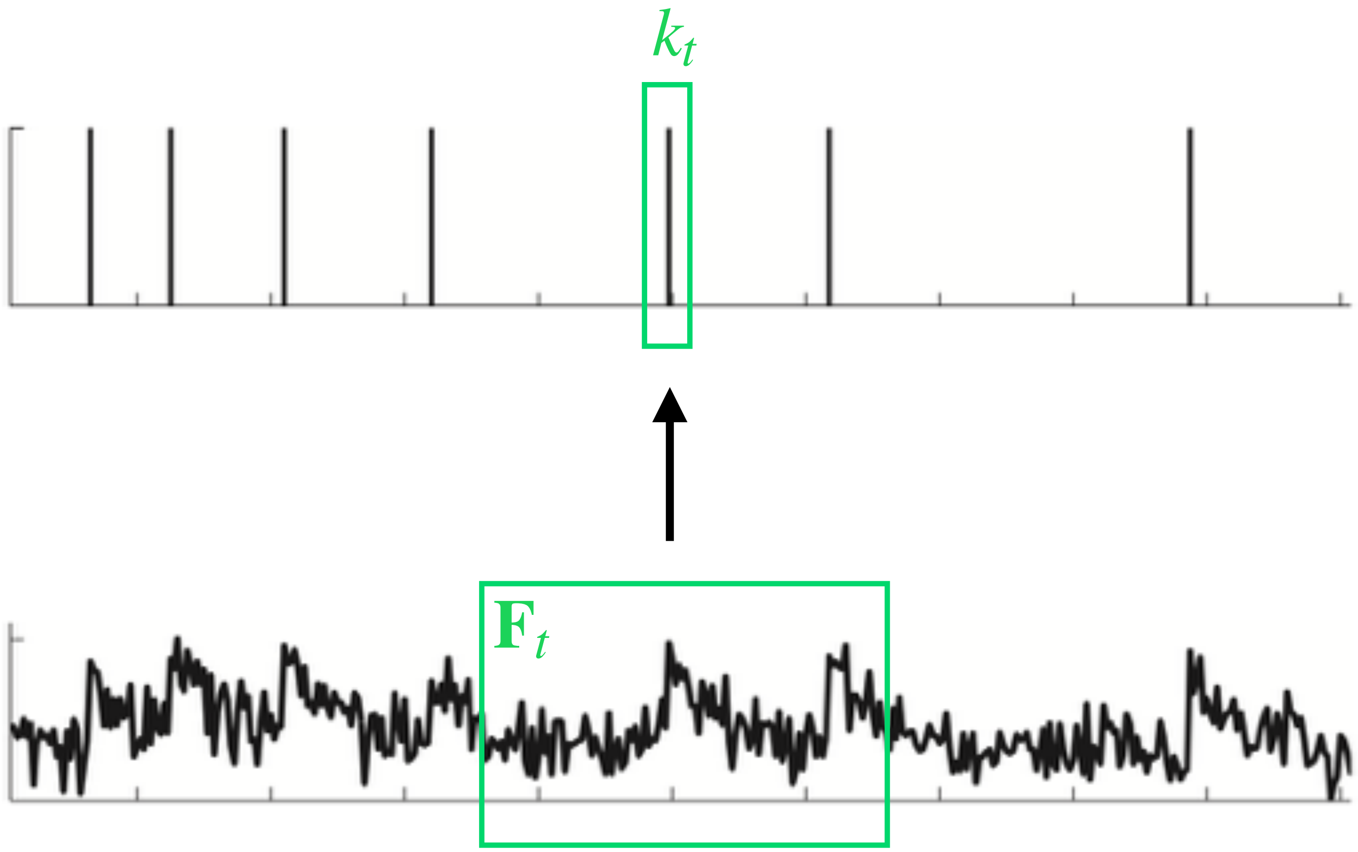
Estimating spike trains

MAP: $\hat{\mathbf{k}} = \arg\max_{\mathbf{k}} P(\mathbf{k} \mid \mathbf{F})$

SMC: $\hat{\mathbf{k}} = E[\mathbf{k} \mid \mathbf{F}]$

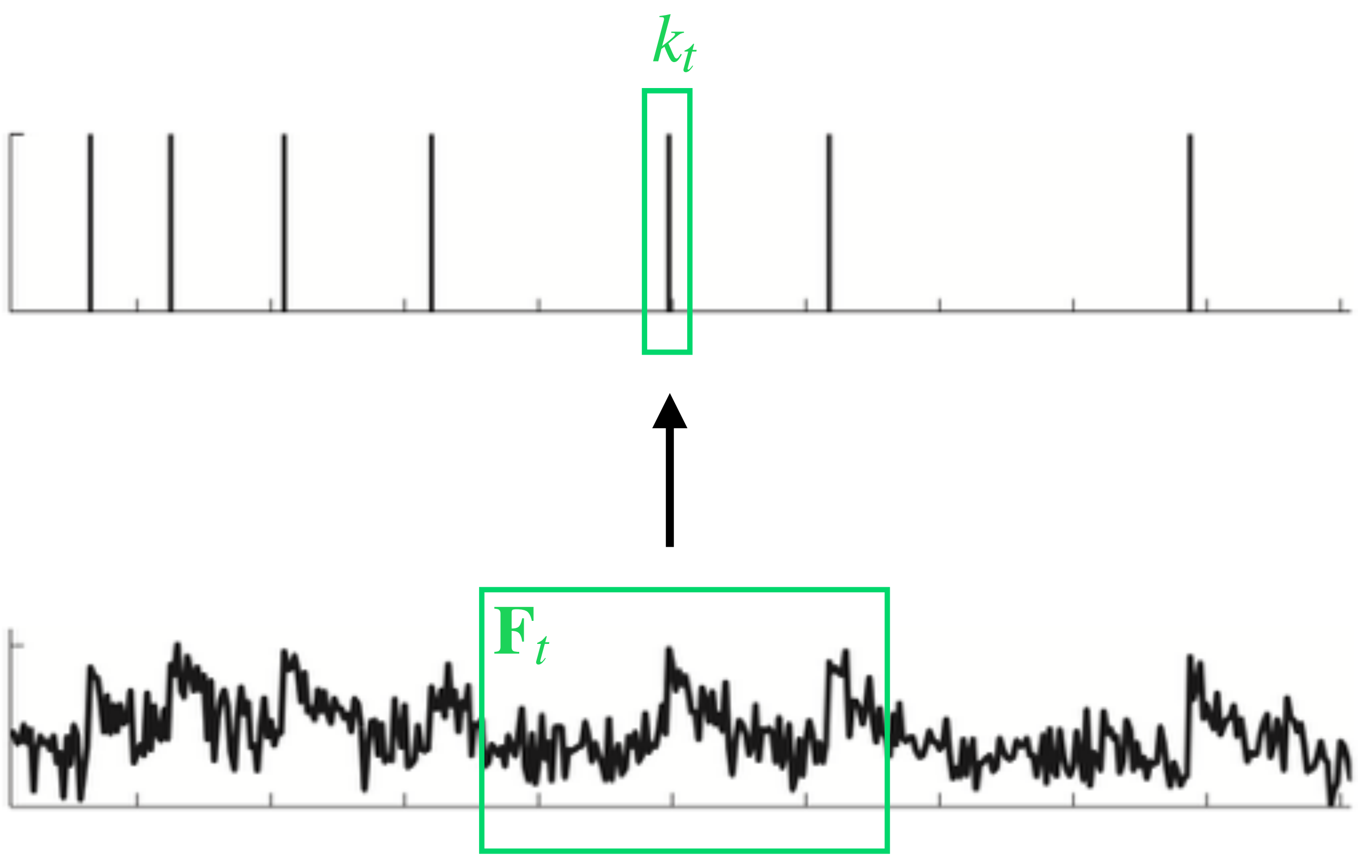
Vogelstein et al. (2010)

Discriminative approach



$$P(k_t \mid \mathbf{F}_t) = \text{Poisson}\,(k_t; f(\mathbf{F}_t))$$

Discriminative approach



$$P(k_t \mid \mathbf{F}_t) = \text{Poisson}\left(k_t; f\left(\mathbf{F}_t\right)\right)$$

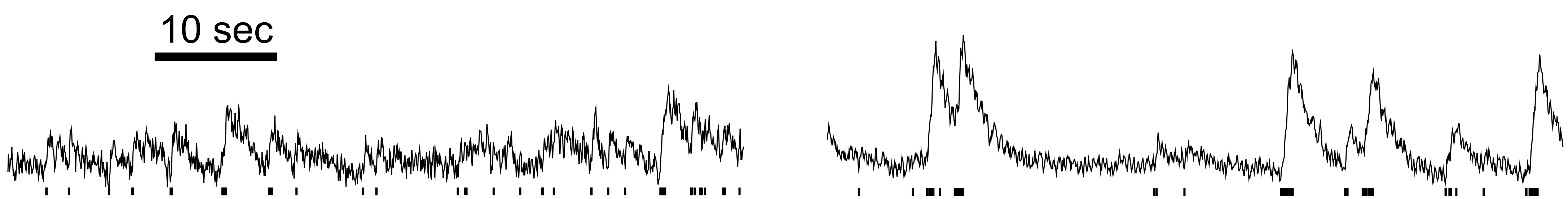
$$f_{\text{LNP}}(\mathbf{F}_t) = \exp\left(\mathbf{w}^\top \mathbf{F}_t + b\right)$$

$$f_{\text{STM}}(\mathbf{F}_t) = \sum_k \exp\left(\sum_m \beta_{km}\left(\mathbf{u}_m^\top \mathbf{F}_t\right)^2 + \mathbf{w}_k^\top + b_k\right)$$

$$f_{\text{MLP}}(\mathbf{F}_t) = \exp\left(\mathbf{w}_3^\top g\left(\mathbf{W}_2 g\left(\mathbf{W}_1 \mathbf{F}_t + \mathbf{b}_1\right) + \mathbf{b}_2\right) + b_3\right)$$

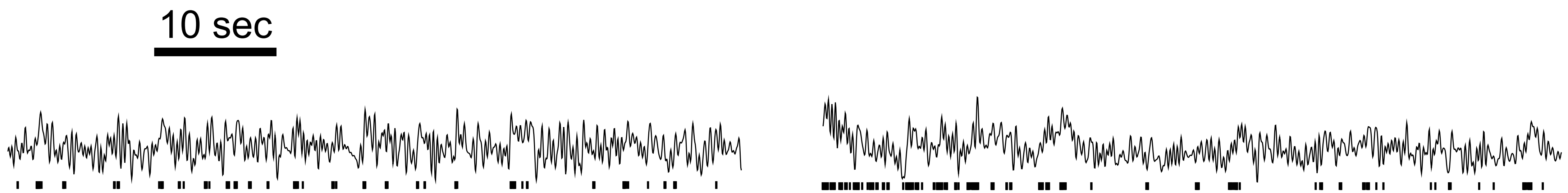
Two-photon calcium imaging

Lab	Euler	Tolias	Tolias
Region	Retina, RGC	V1	V1
Scanning	Galvanom.	AOD	Galvanom.
Rate	8 Hz	40 Hz	12 Hz
Indicator	OGB1	OGB1	OGB1
Cells	9	16	31

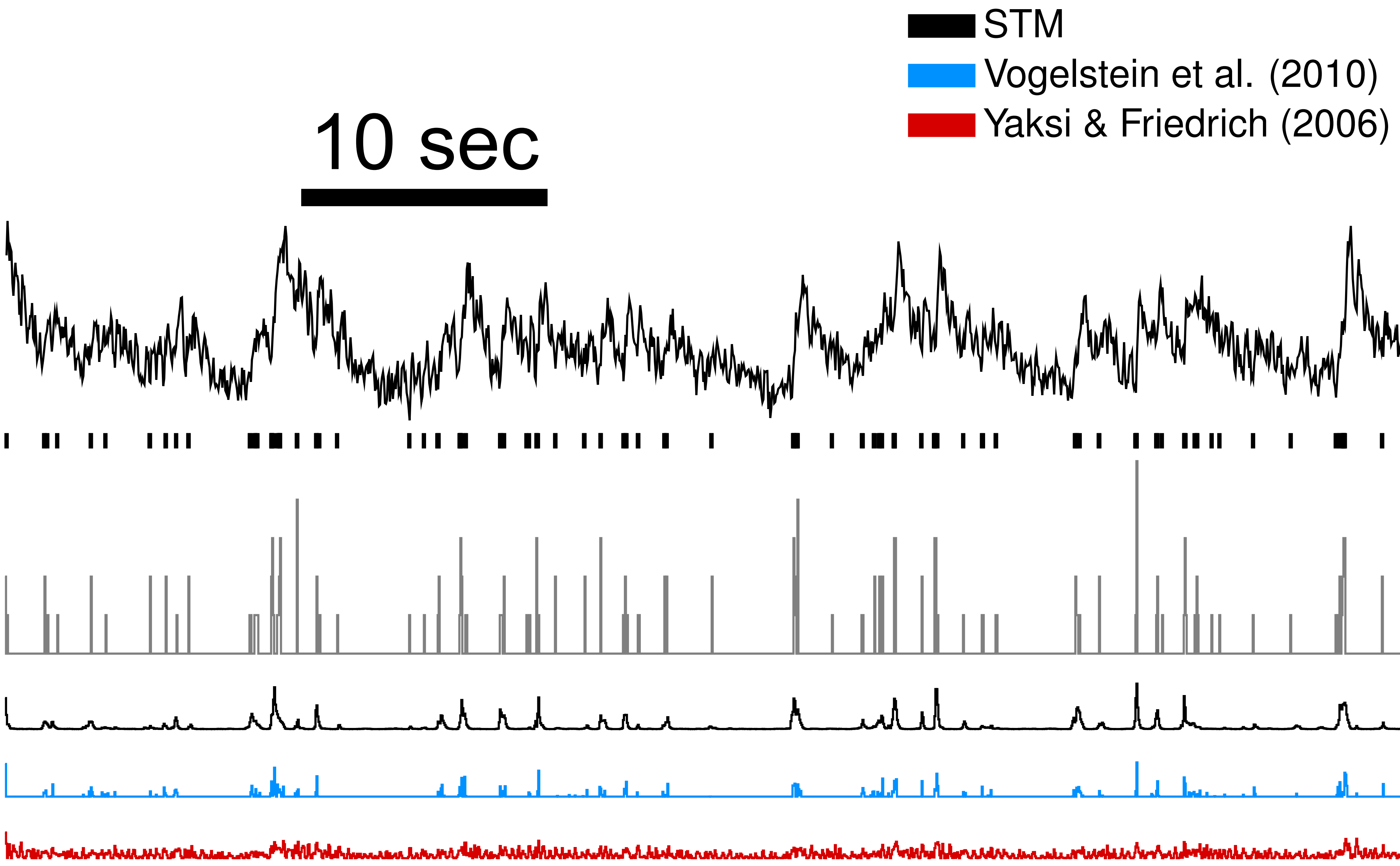


Two-photon calcium imaging

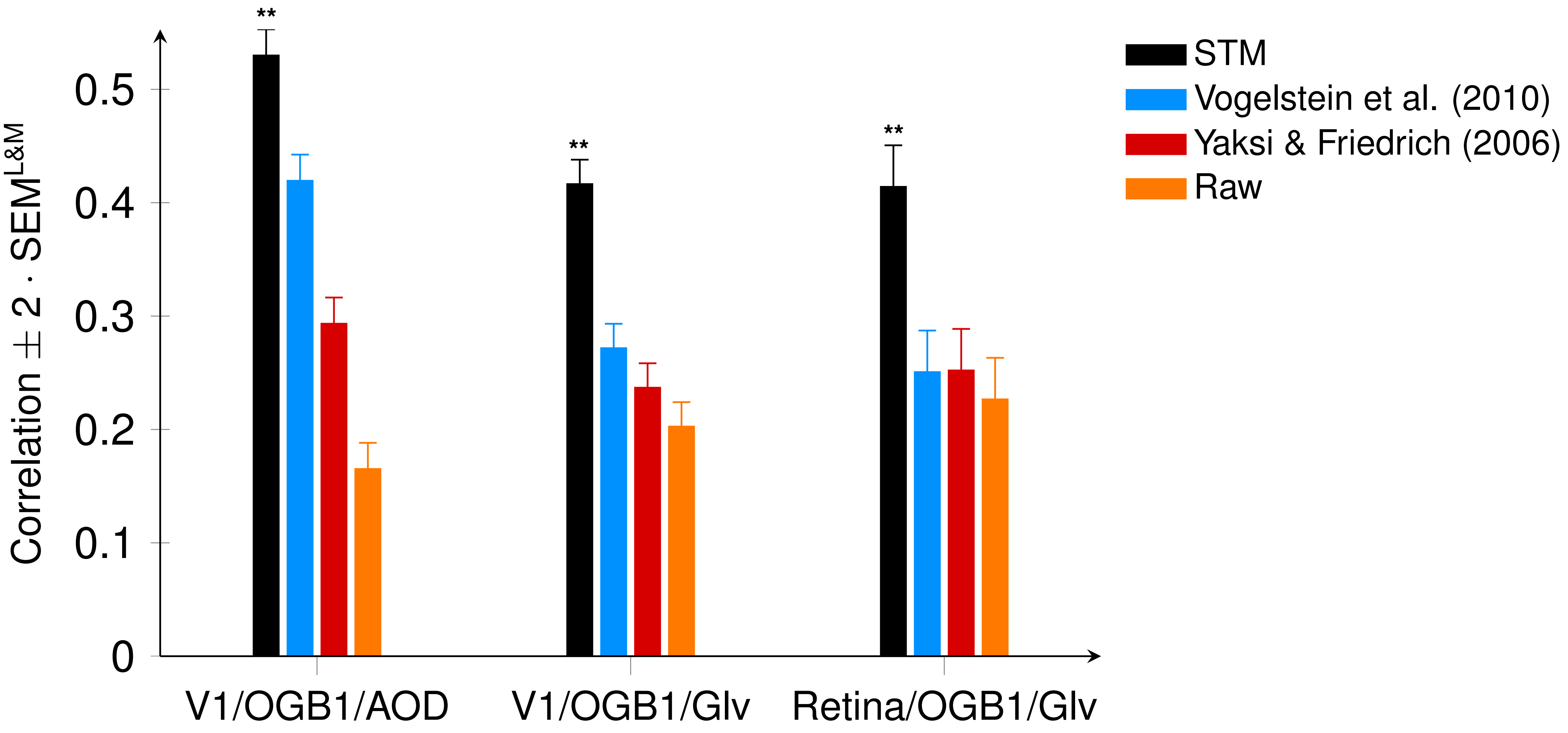
Lab	Euler	Tolias	Tolias
Region	Retina, RGC	V1	V1
Scanning	Galvanom.	AOD	Galvanom.
Rate	8 Hz	40 Hz	12 Hz
Indicator	OGB1	OGB1	OGB1
Cells	9	16	31



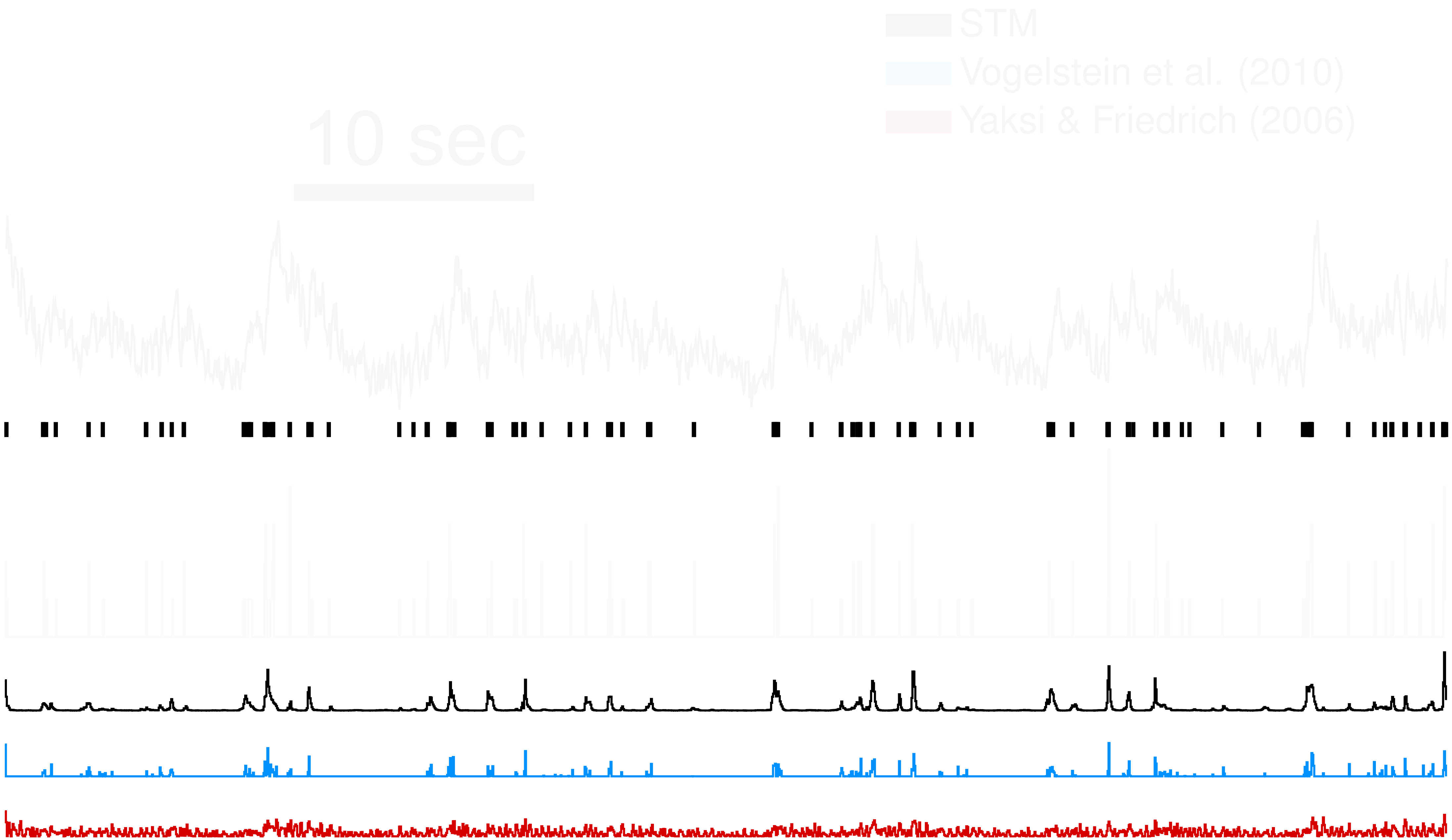
Spike predictions



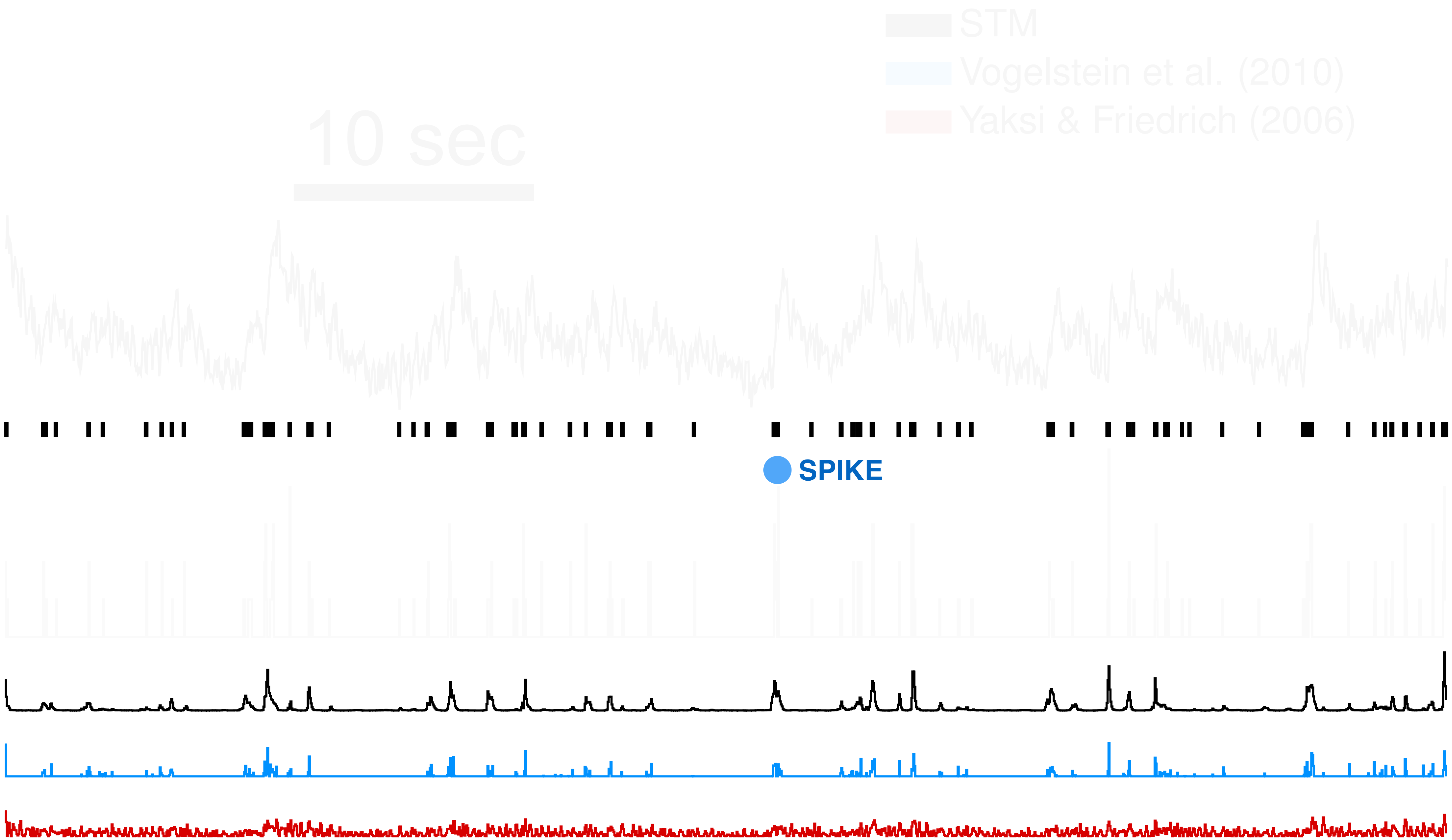
Correlation



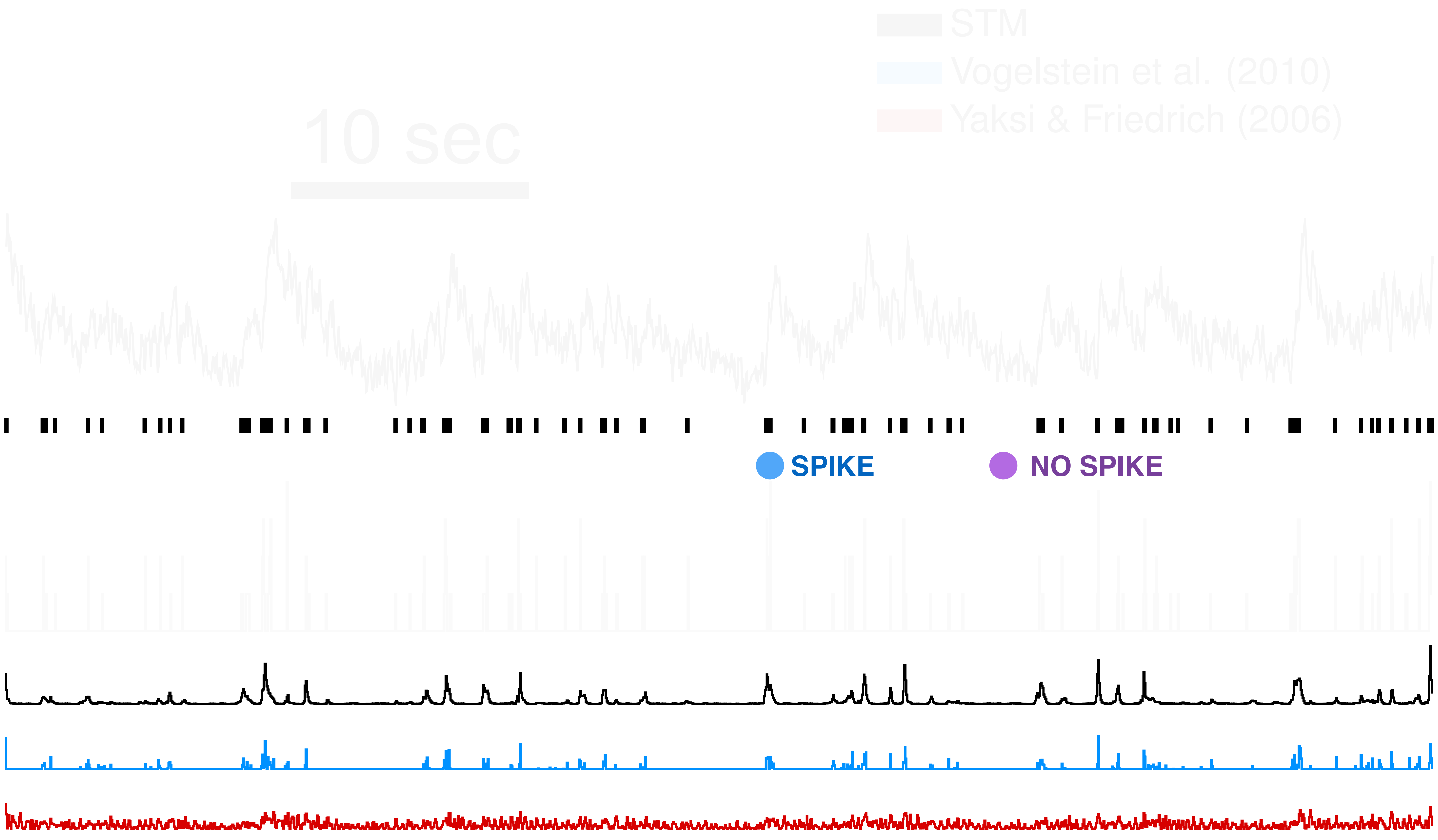
Area under curve



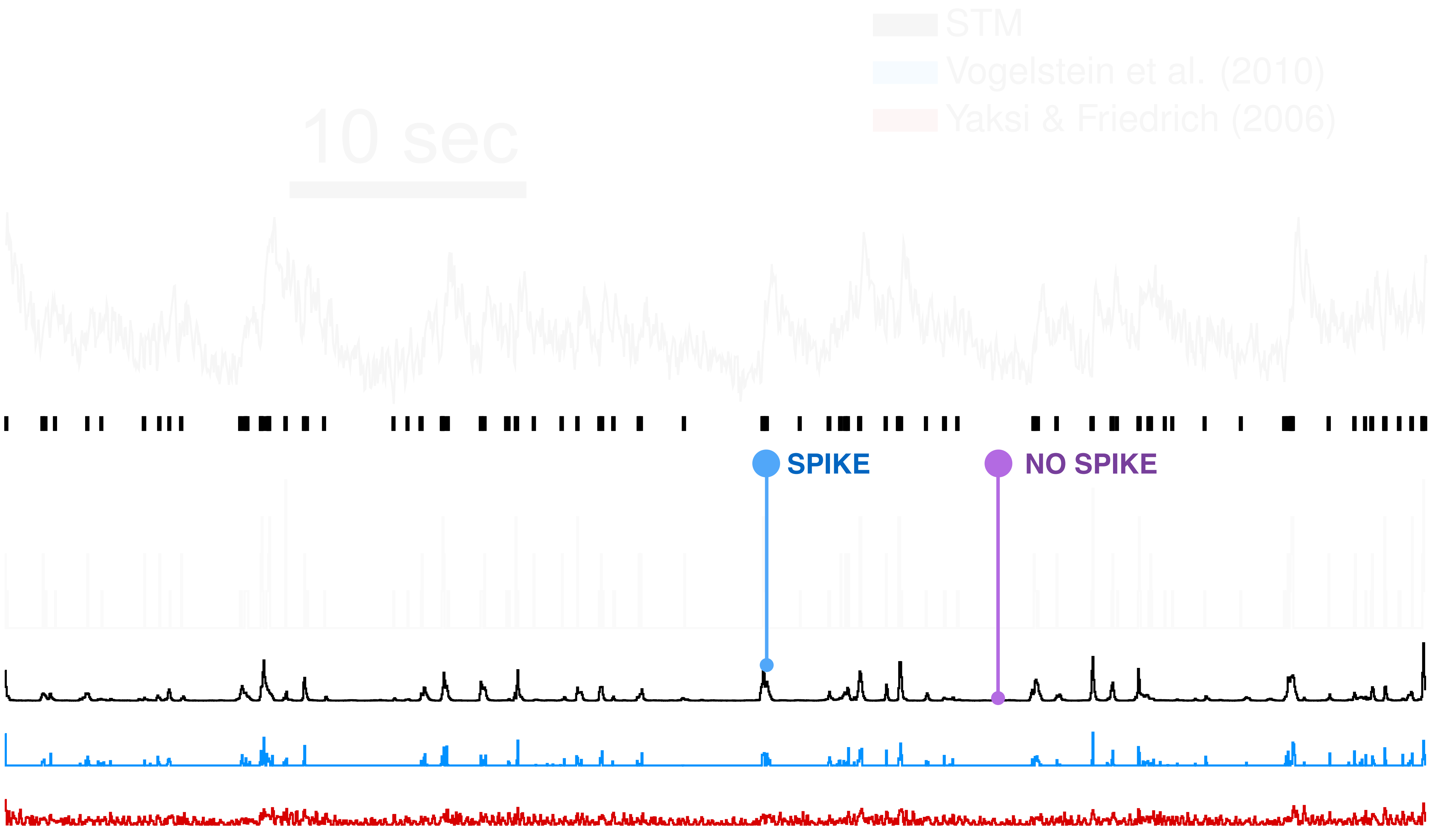
Area under curve



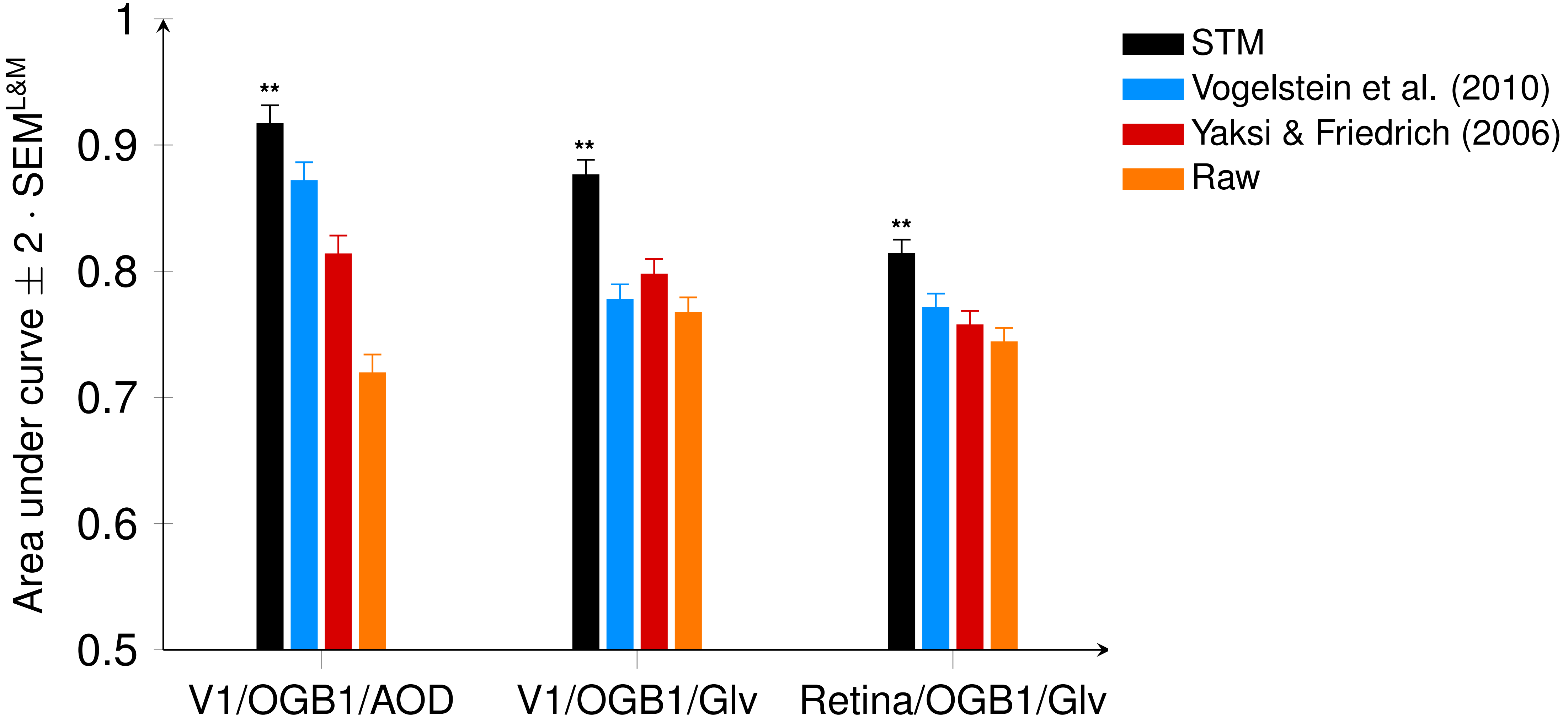
Area under curve



Area under curve

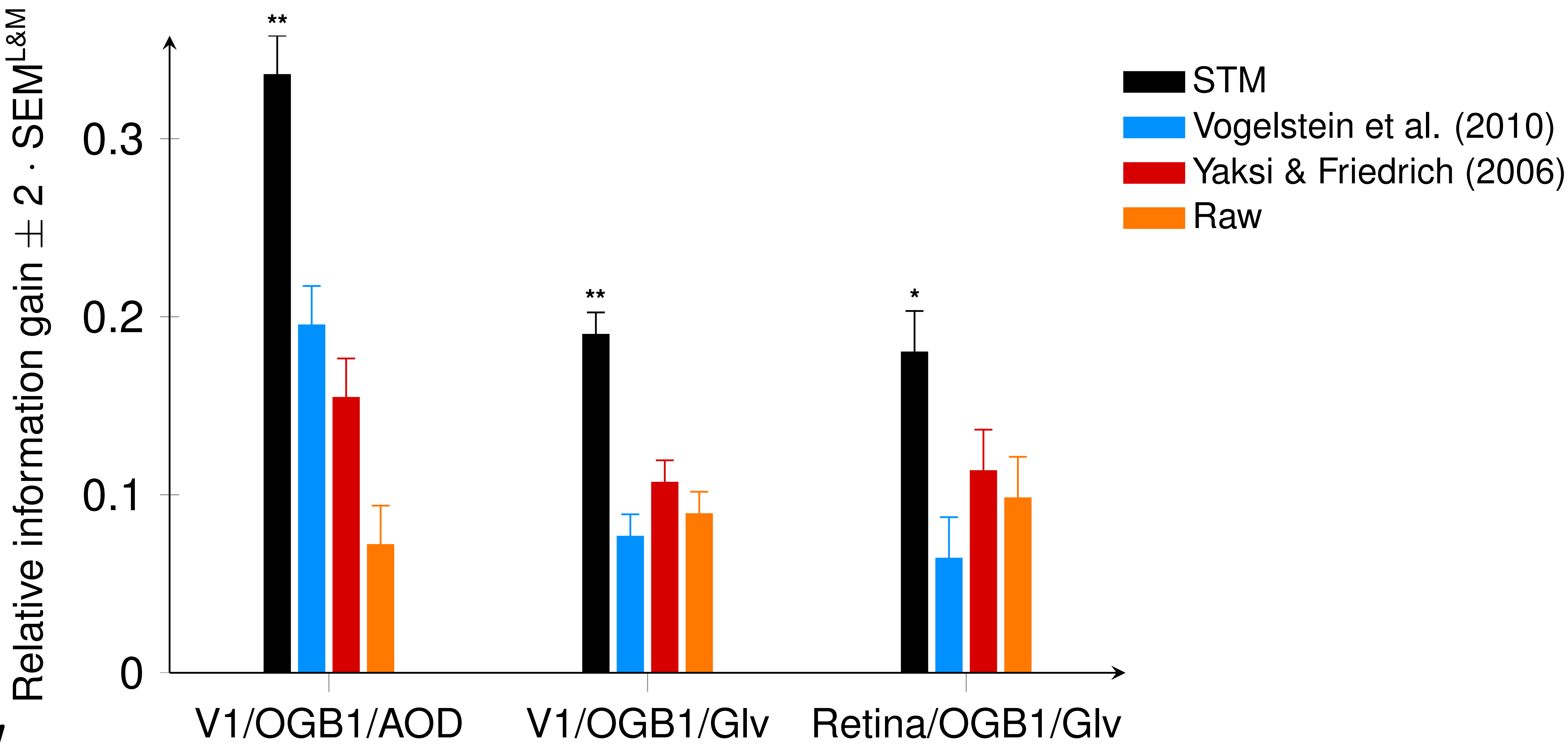


Area under curve



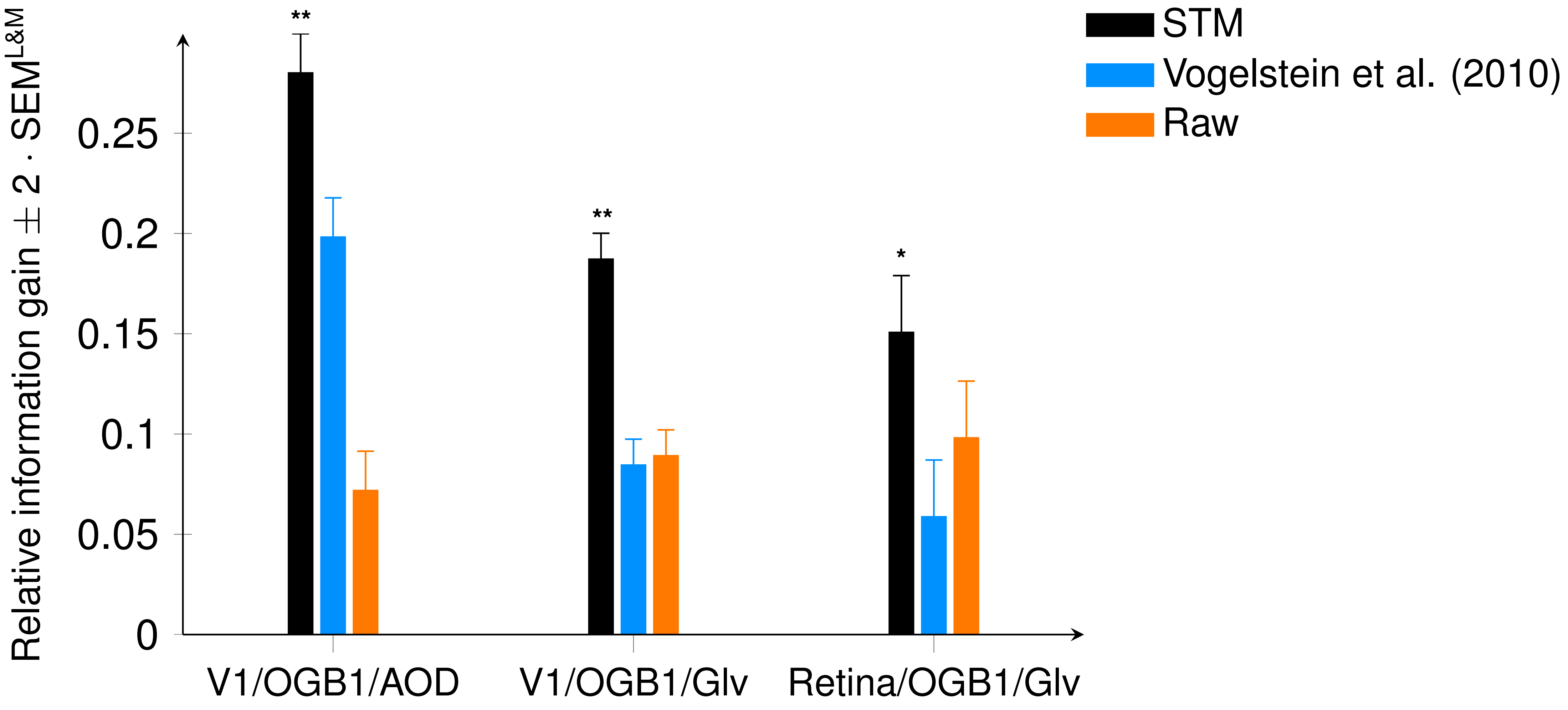
Mutual information

$$I[k,\mathbf{F}] = H[k] - H[k \mid \mathbf{F}] \geq H[k] + E\left[\log P\left(k \mid \lambda(\mathbf{F})\right)\right]$$



$$\frac{I[k,\mathbf{F}]}{H[k]}$$

Generalization



M. Bethge

Trigeminal ganglion cells

C. Schwarz

A. M. Chagas

D. Arnstein

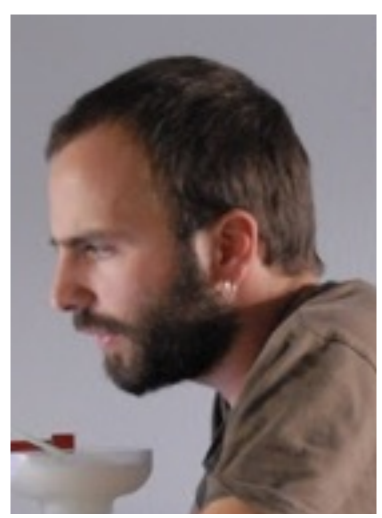
http://bethgelab.org/workshops/neurostats2014/

Thanks

Two-photon imaging

P. Berens

A. Tolias



E. Froudarakis

T. Euler

T. Baden

Generalized linear models

$$p(y=1 \mid \mathbf{x}) = \frac{1}{1+\frac{p(\mathbf{x}|y=0)p(y=0)}{p(\mathbf{x}|y=1)p(y=1)}}$$

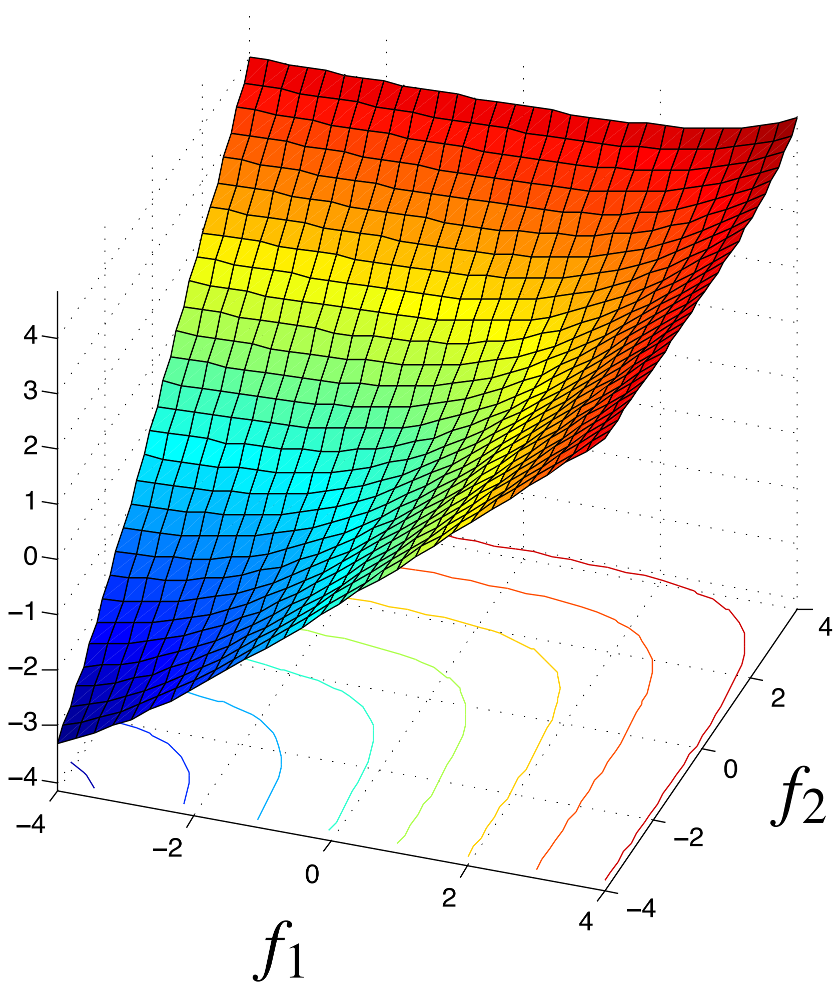
$$= \sigma\left(\log\frac{p(\mathbf{x} \mid y=1)}{p(\mathbf{x} \mid y=0)} + \log\frac{p(y=1)}{p(y=0)}\right)$$

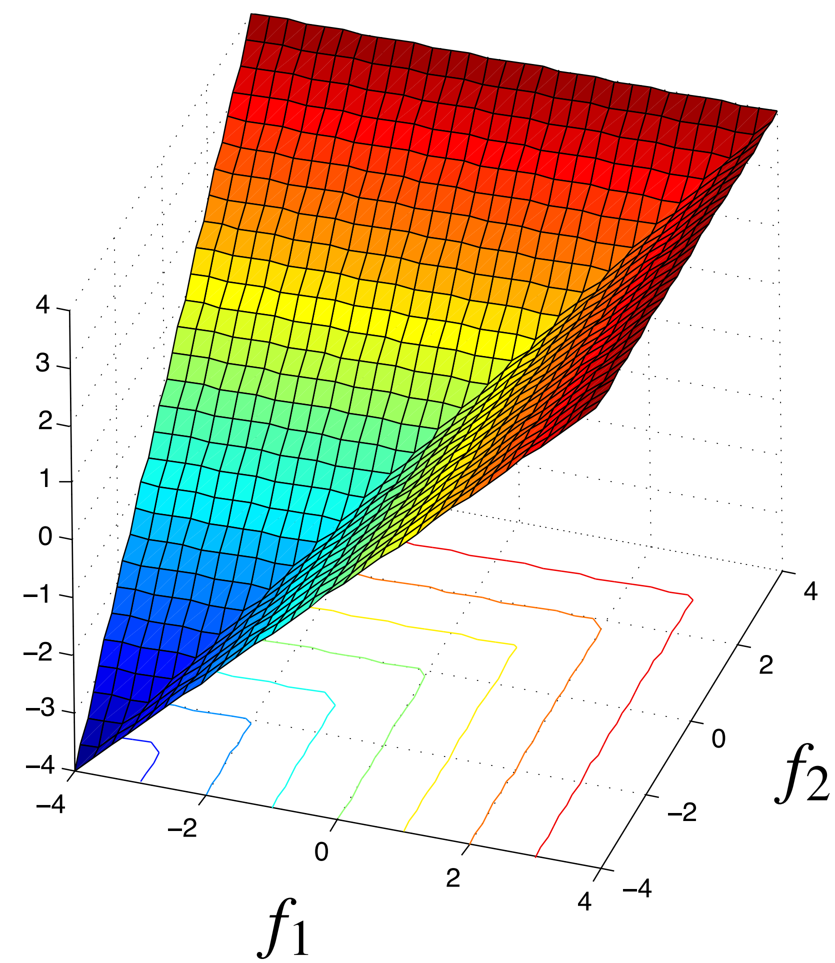
$$= \sigma\left(\mathbf{w}^\top \phi(\mathbf{x}) + b\right)$$

$$\sigma(z) = \frac{1}{1+\exp(-z)}$$

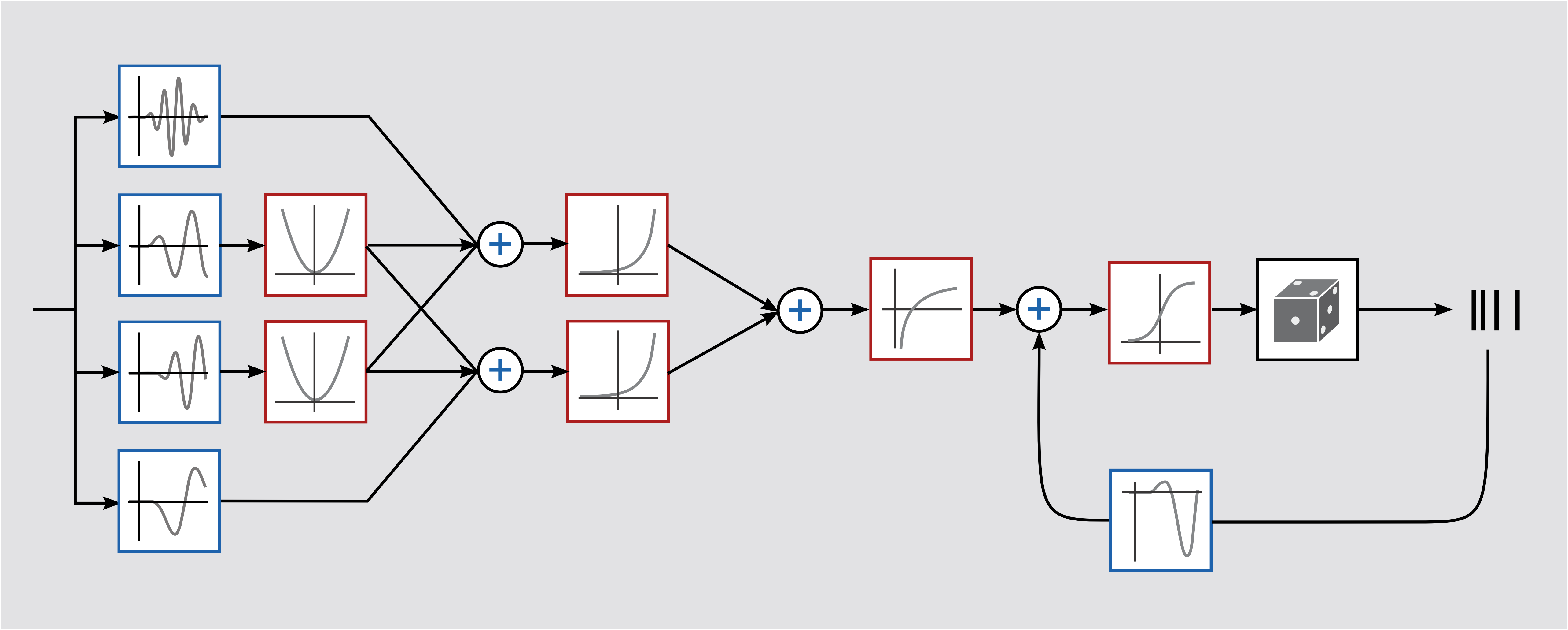
Soft-maximum

$$\log \sum_k \exp f_k \approx \max_k f_k$$

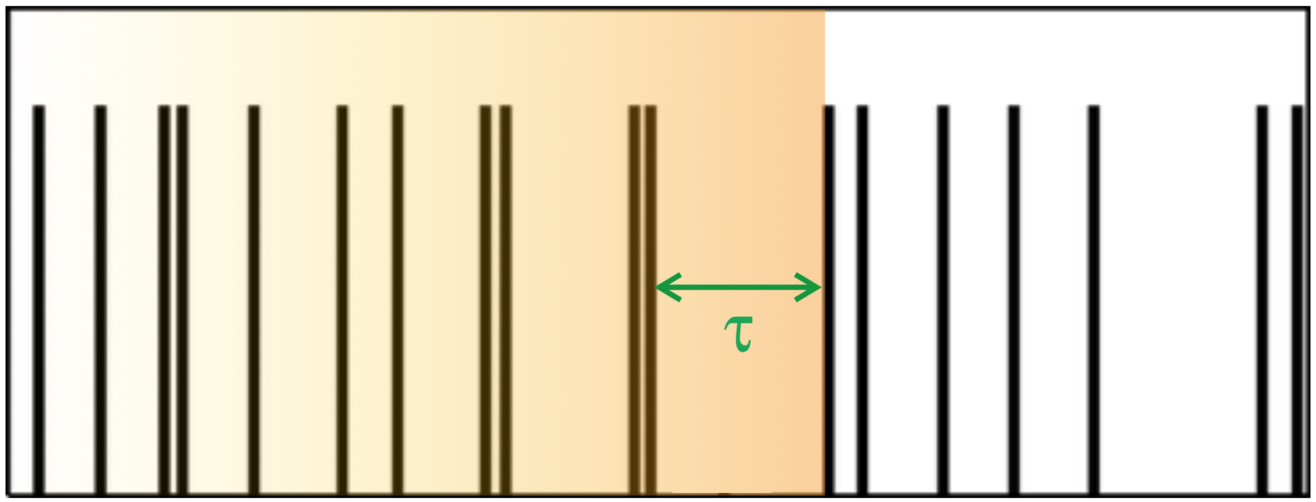




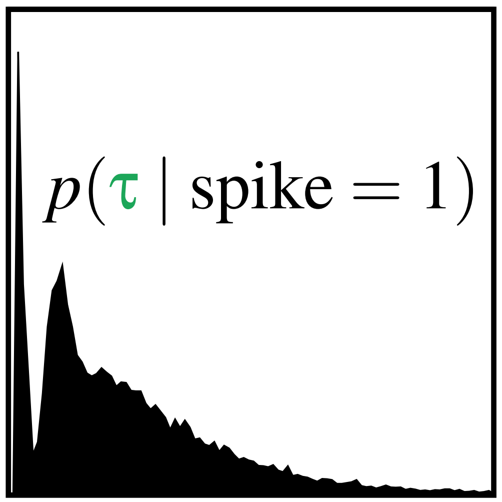
Feed-forward view



Spike history dependency



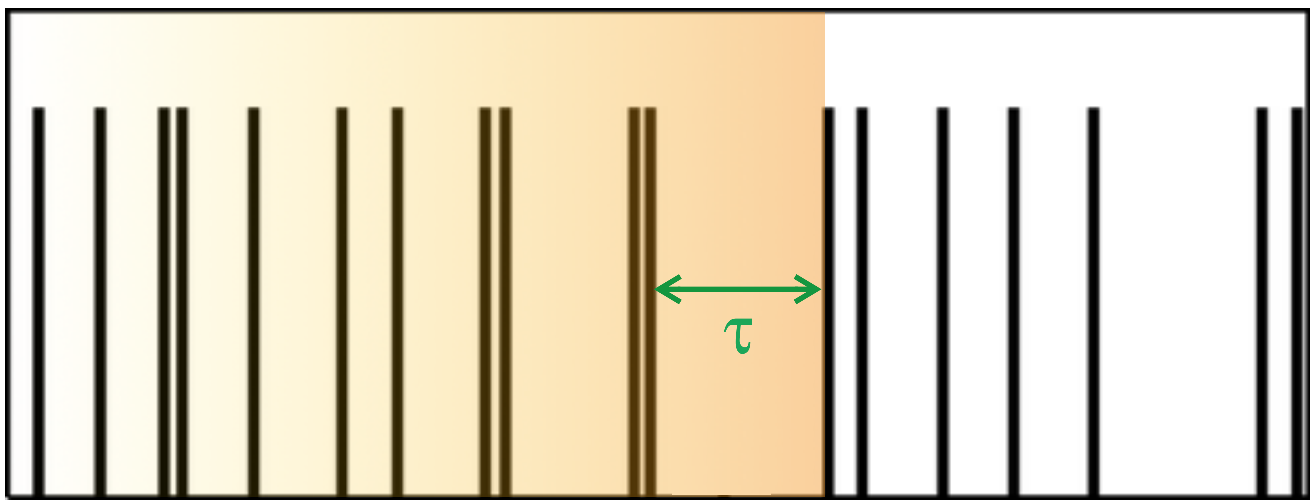
Interspike-interval distribution



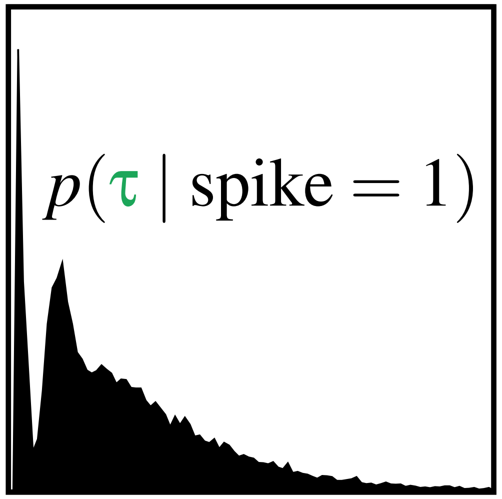
Naive Bayes assumption: $p(\mathbf{x}, \tau \mid \text{spike}) = p(\mathbf{x} \mid \text{spike})\, p(\tau \mid \text{spike})$

$$p(\text{spike} = 1 \mid \mathbf{x}, \tau) = \sigma\left(\log \frac{p(\mathbf{x} \mid \text{spike} = 1)}{p(\mathbf{x} \mid \text{spike} = 0)} + \log \frac{p(\tau \mid \text{spike} = 1)}{p(\tau \mid \text{spike} = 0)} + \log \frac{p(\text{spike} = 1)}{p(\text{spike} = 0)}\right)$$

Spike history dependency



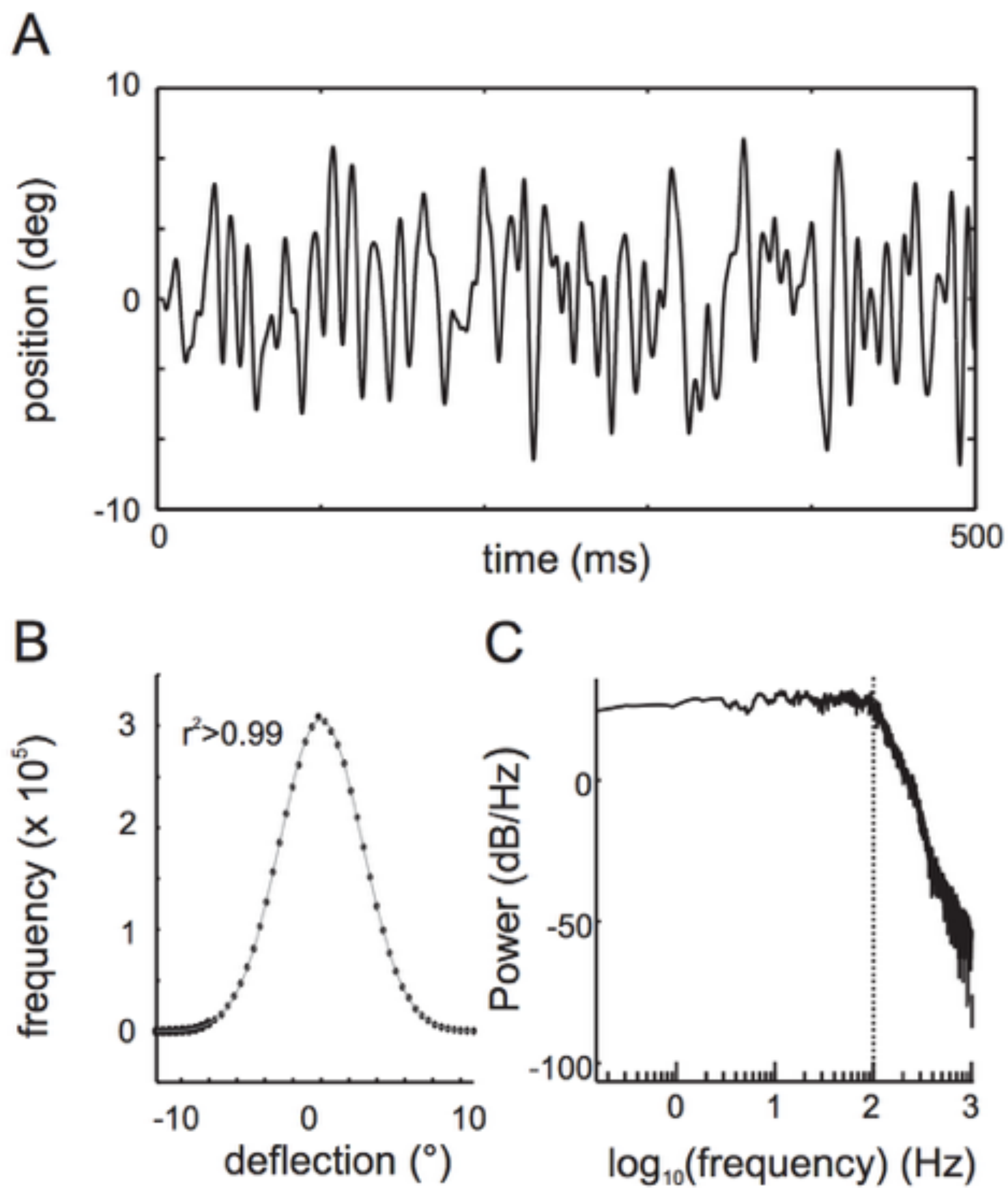
Interspike-interval distribution



Naive Bayes assumption: $p(\mathbf{x}, \tau \mid \text{spike}) = p(\mathbf{x} \mid \text{spike})\, p(\tau \mid \text{spike})$

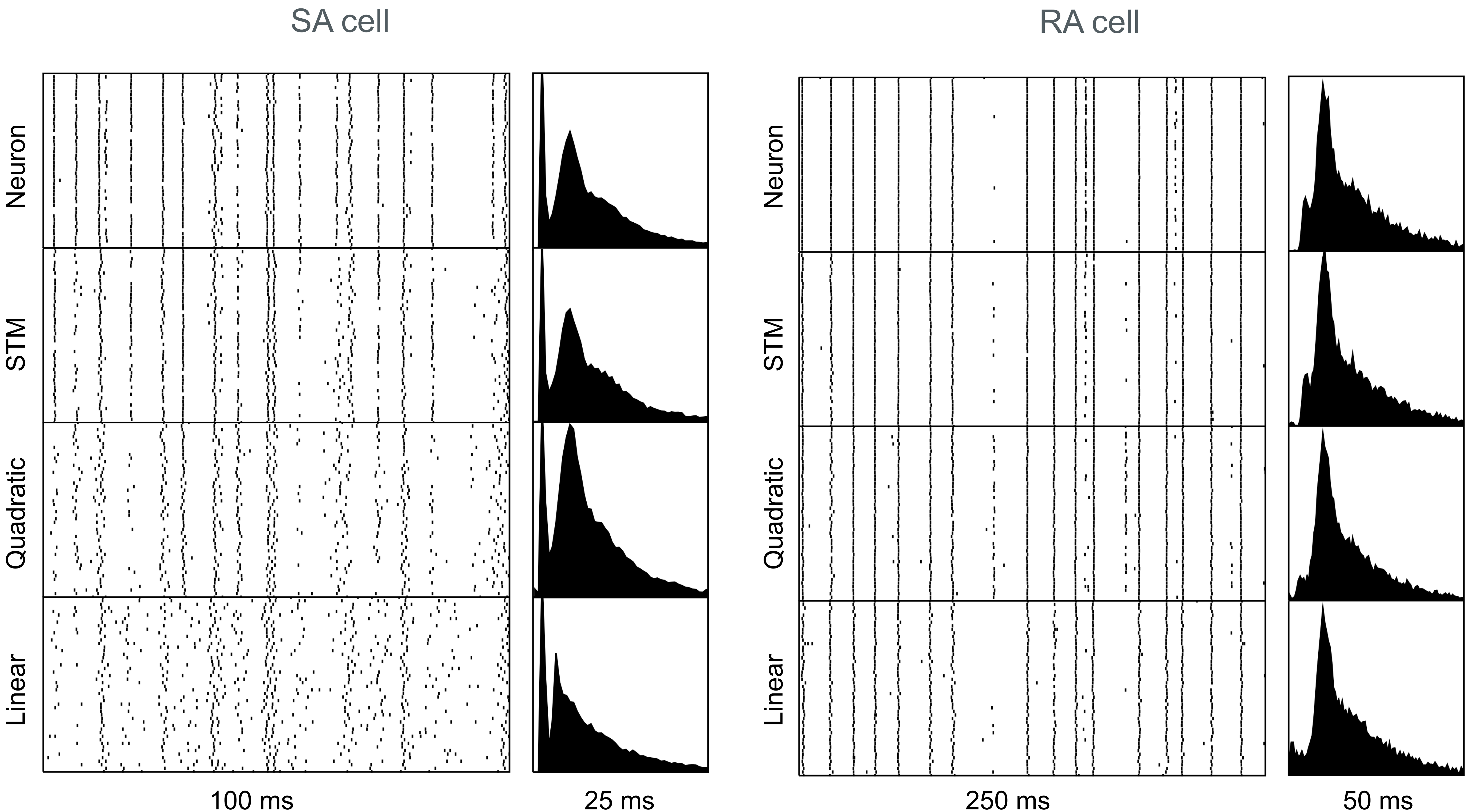
$$p(\text{spike} = 1 \mid \mathbf{x}, \tau) = \sigma\left(\log \sum_k \exp\left(\sum_m \alpha_{km} (\mathbf{u}_m^\top \mathbf{x})^2 + \mathbf{w}_k^\top \mathbf{x} + b_k \right) + \mathbf{v}^\top \phi(\tau) \right)$$

Stimulus



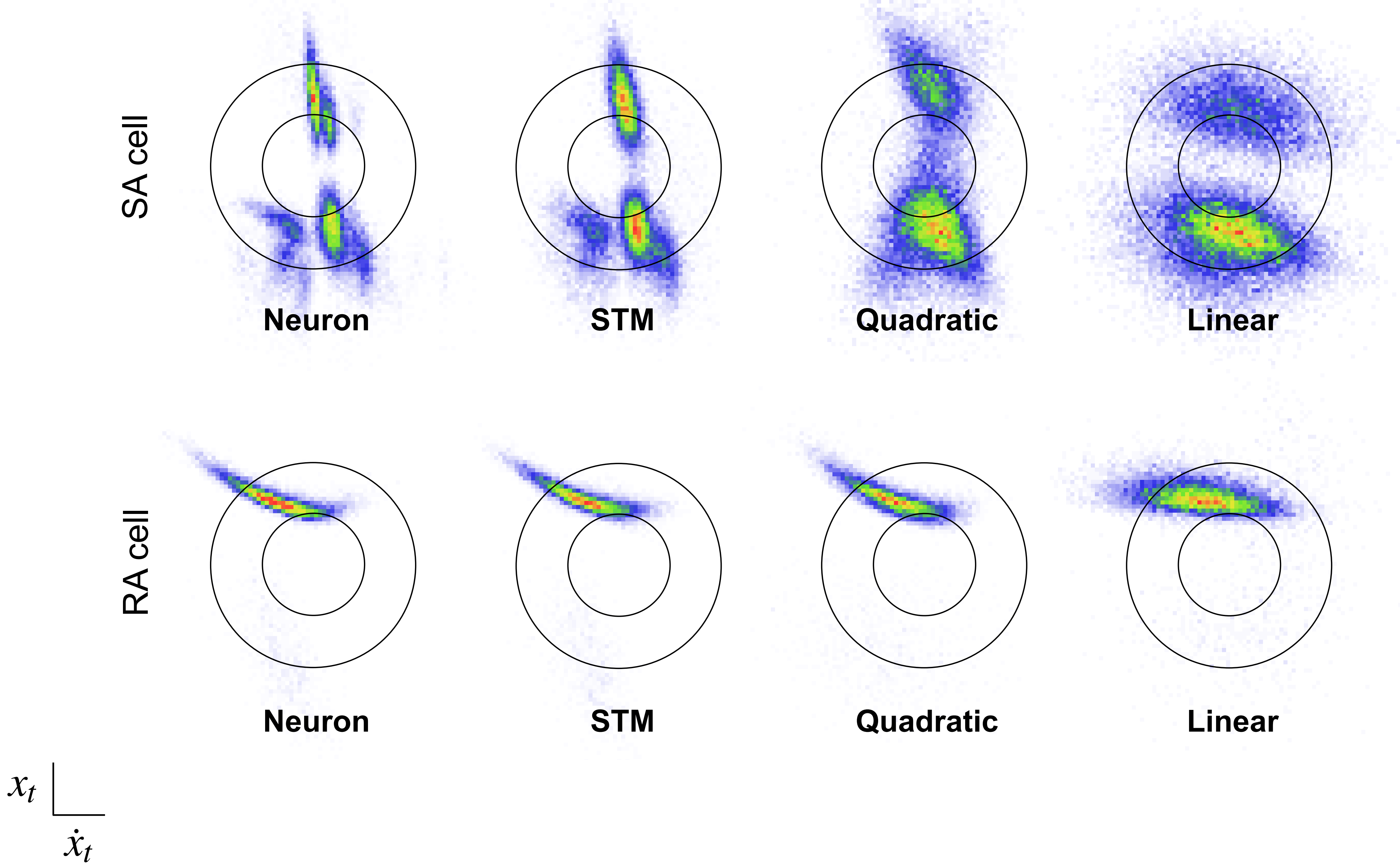
A
10
0
-10
position (deg)
0
500
time (ms)
B
r²>0.99
3
2
1
0
frequency (x 10⁵)
-10
0
10
deflection (°)
C
0
-50
-100
Power (dB/Hz)
0
1
2
3
log10(frequency) (Hz)

Spike trains



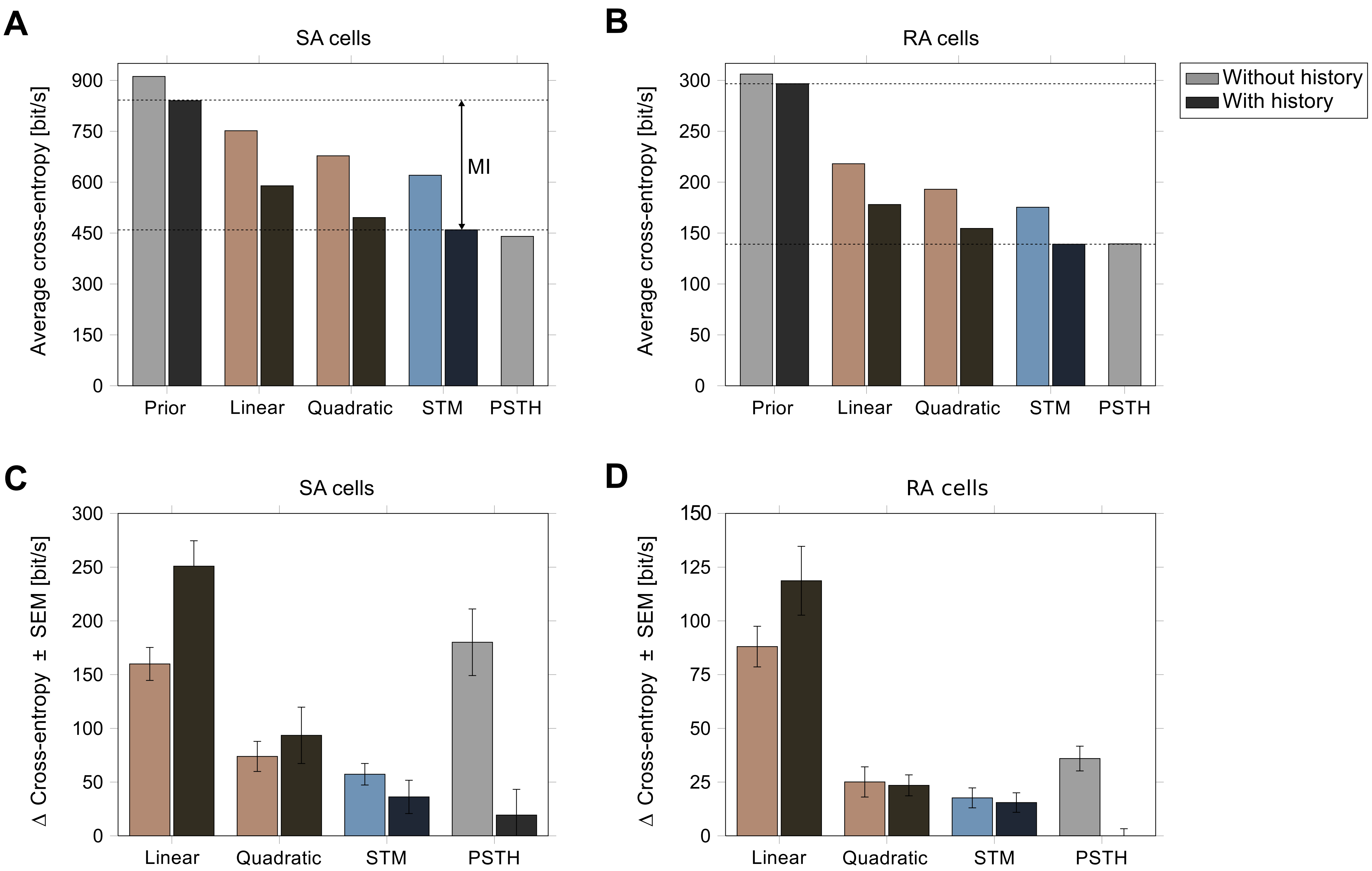
Theis et al. (2013)

Spike-triggered distributions

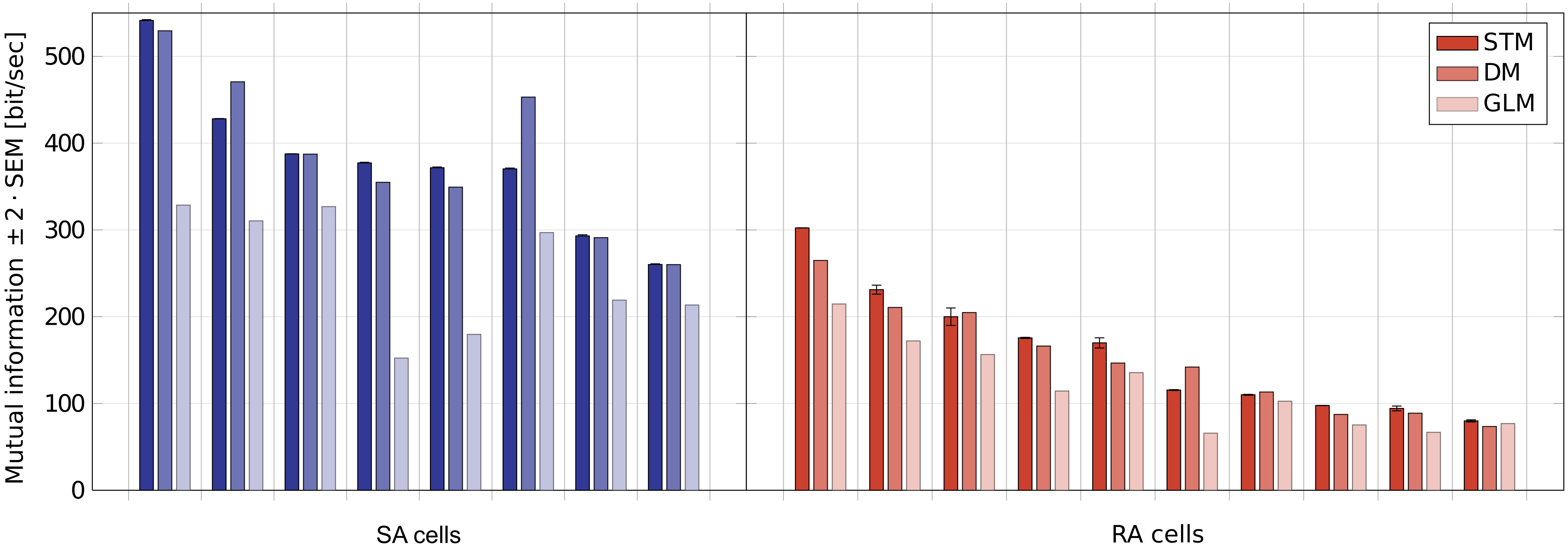


Theis et al. (2013)

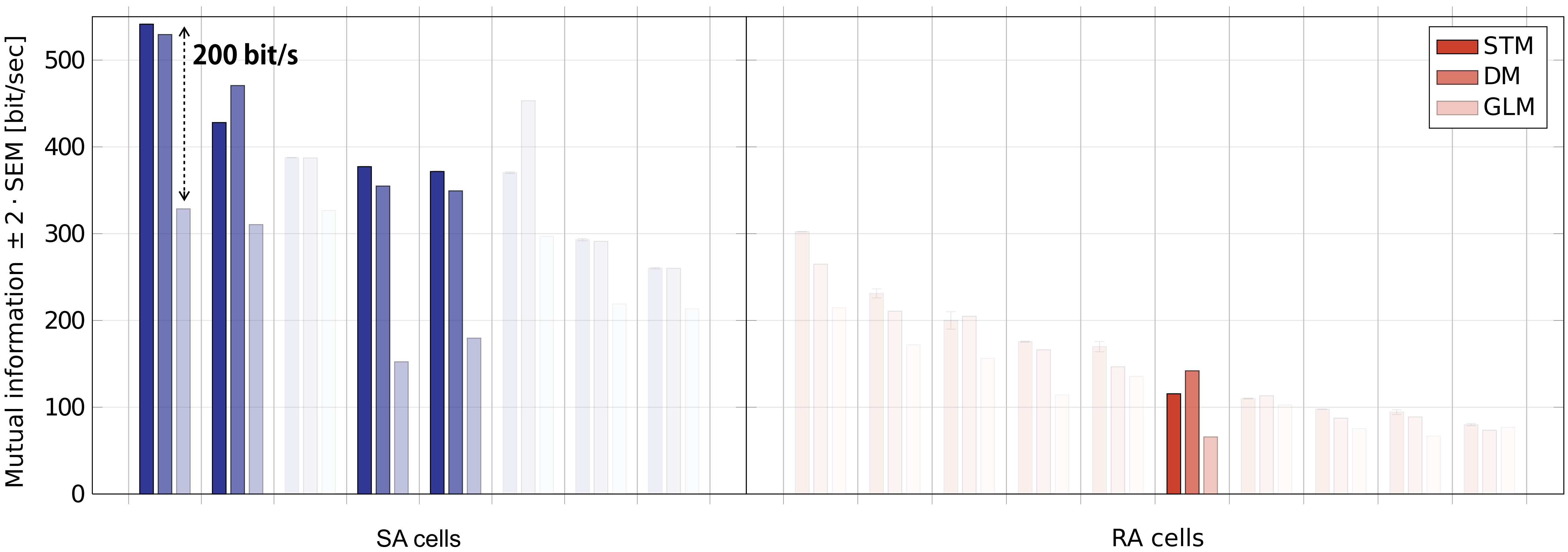
Trigeminal ganglion cells



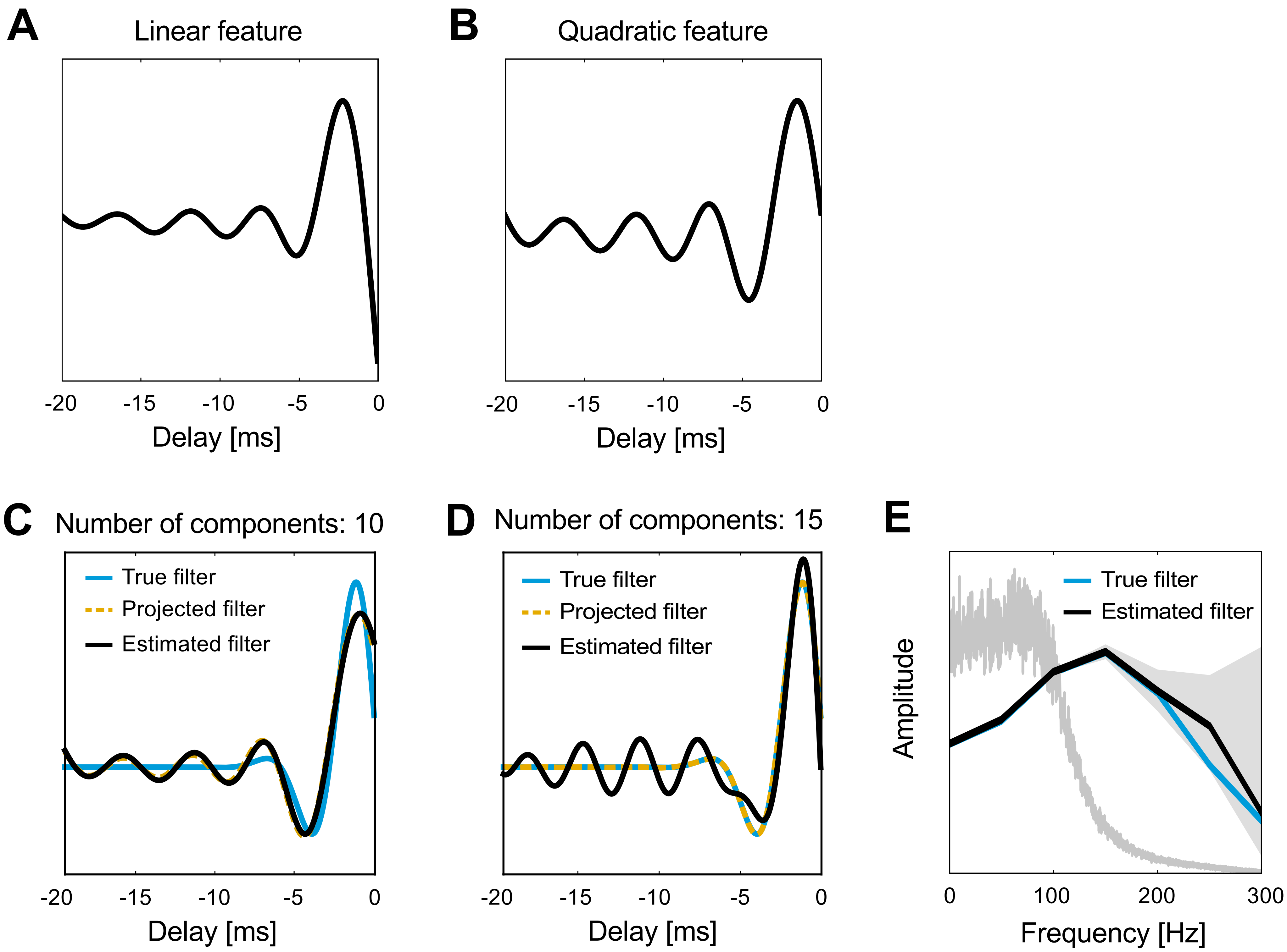
Mutual information



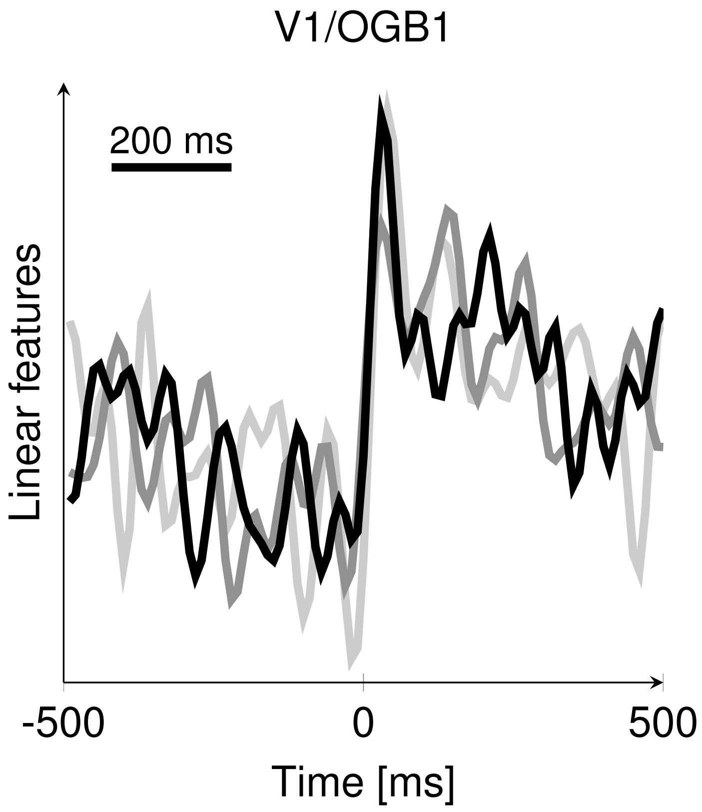
Mutual information

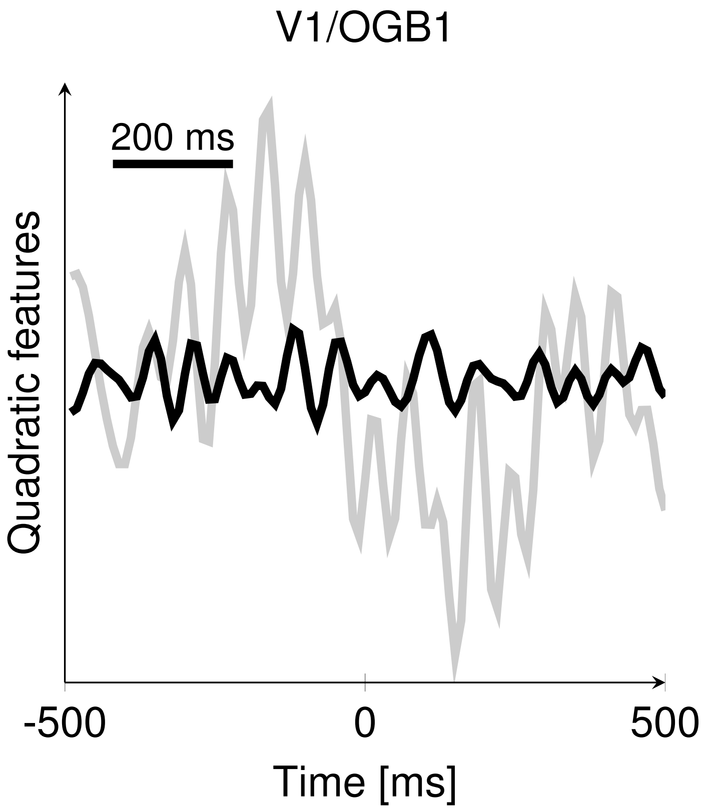


Visualizing filters

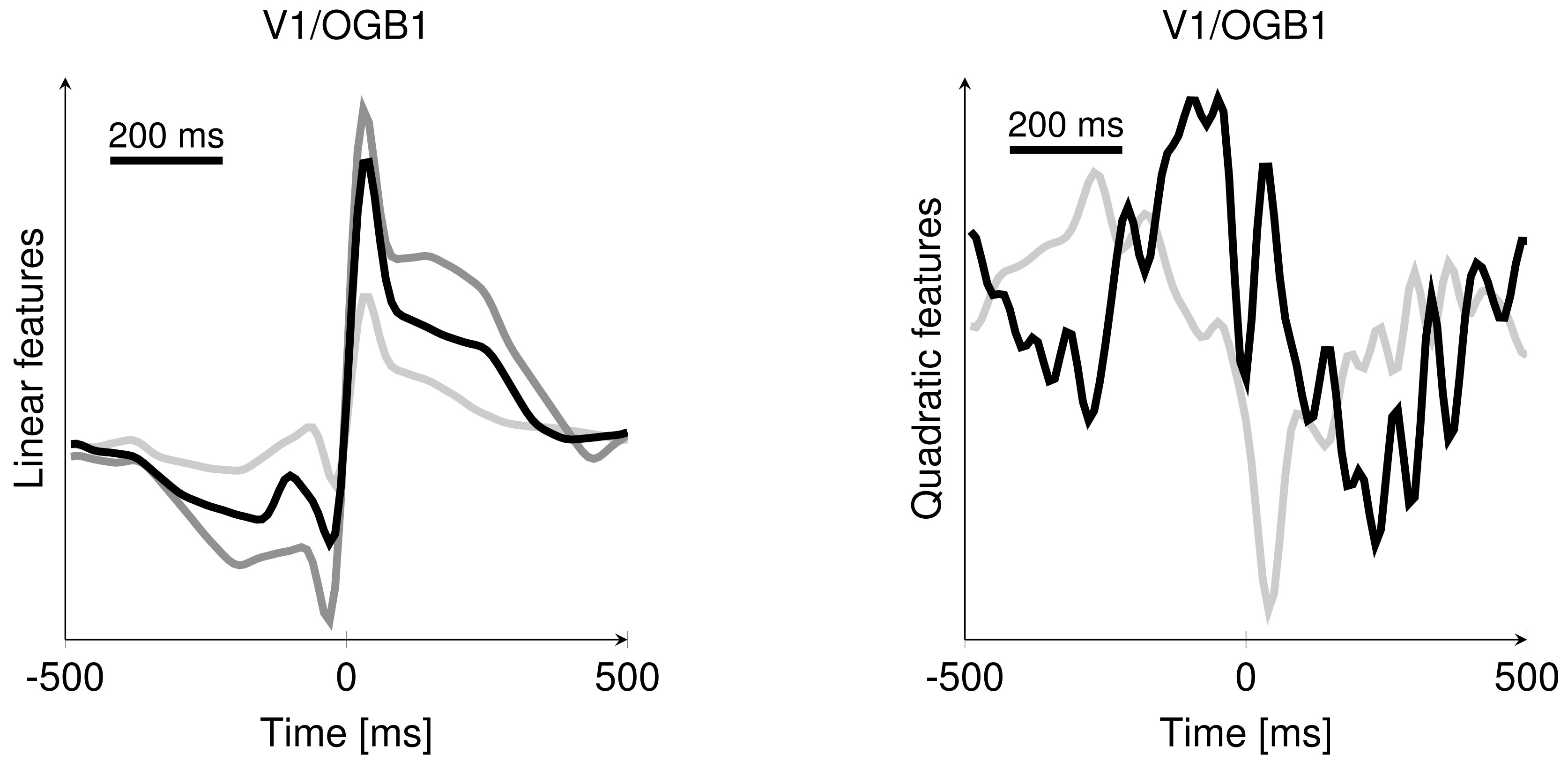


Filters for calcium imaging data





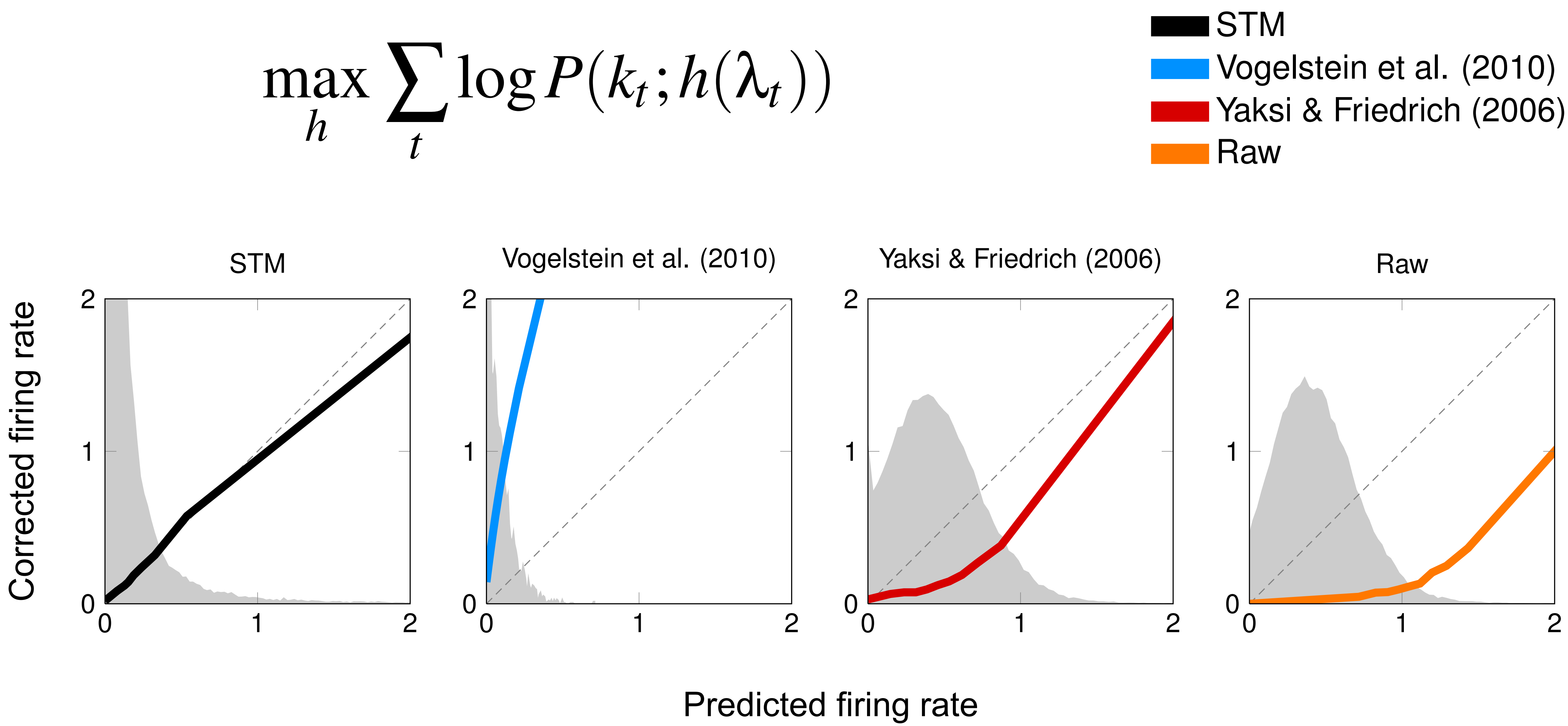
Filters for calcium imaging data



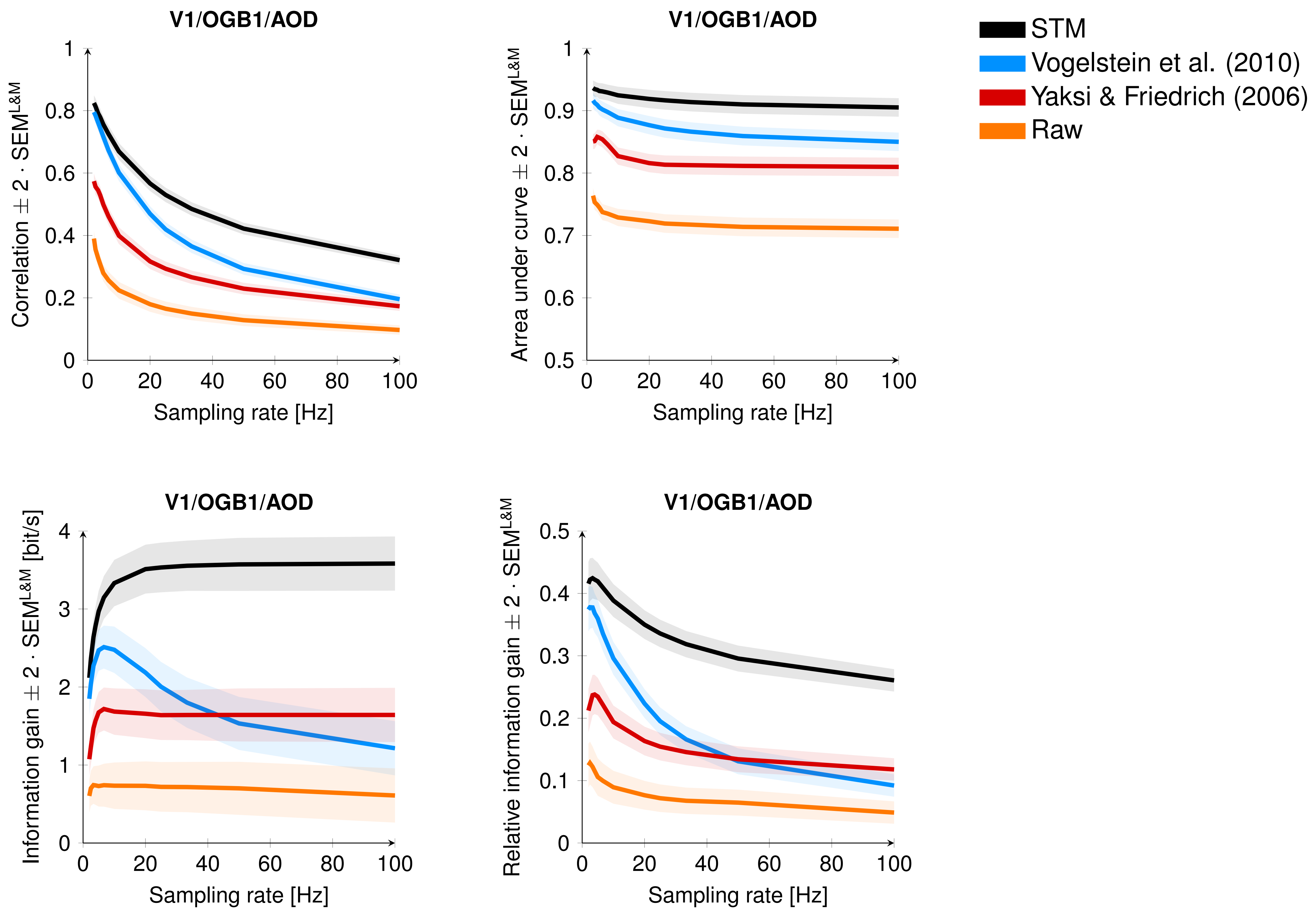
$$\frac{1}{T}\sum_t \log p(k_t \mid \mathbf{F}_t) + \lambda ||\mathbf{Aw}||_1$$

Pointwise nonlinearity

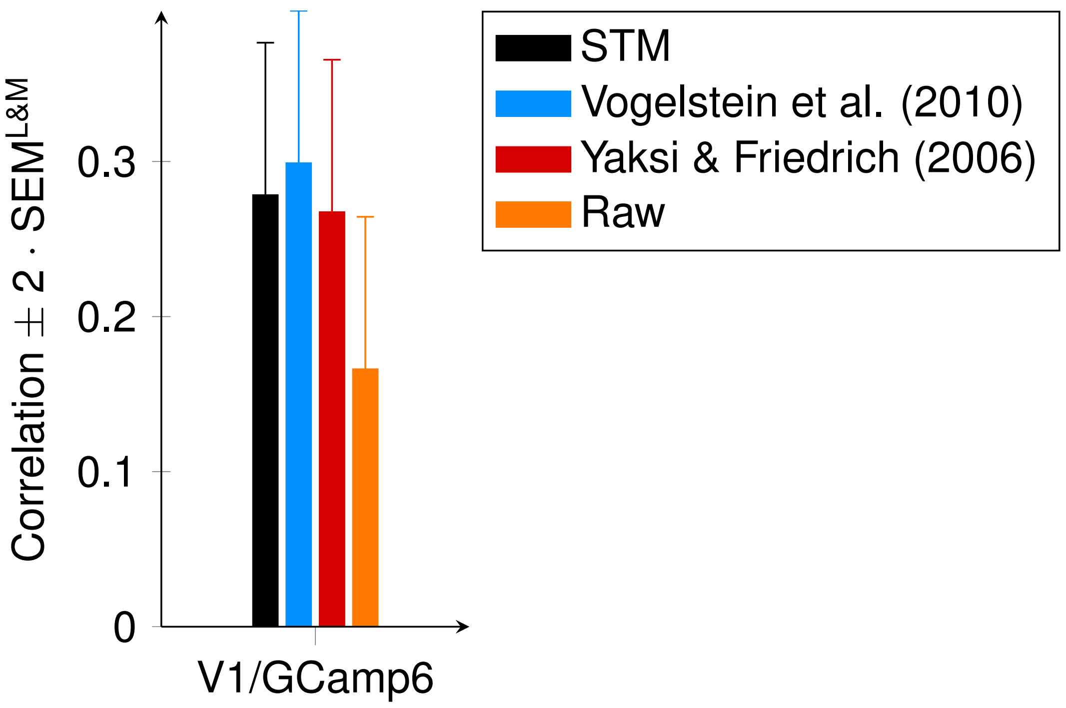
$$\max_{h} \sum_{t} \log P(k_t; h(\lambda_t))$$



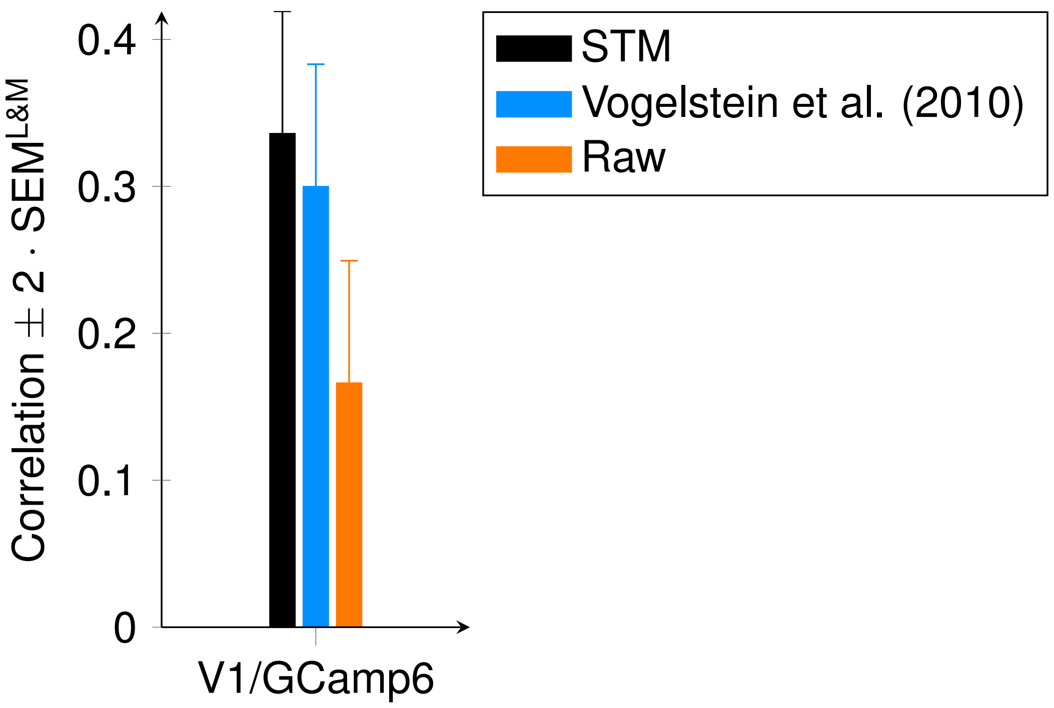
Performance as function of sampling rate



GCamp6

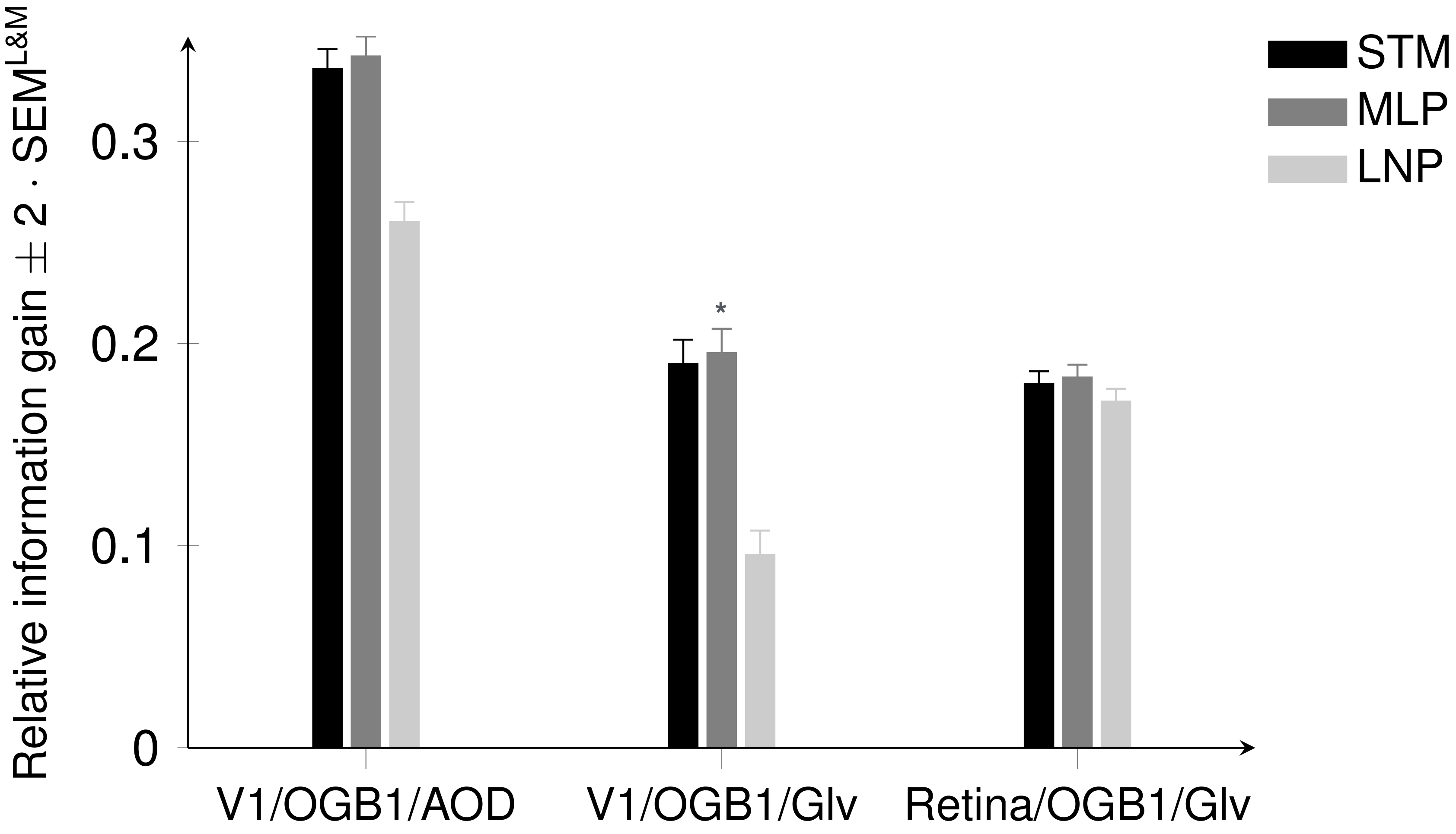


Correlation ± 2 · SEM^L&M
0.3
0.2
0.1
0
V1/GCamp6
STM
Vogelstein et al. (2010)
Yaksi & Friedrich (2006)
Raw

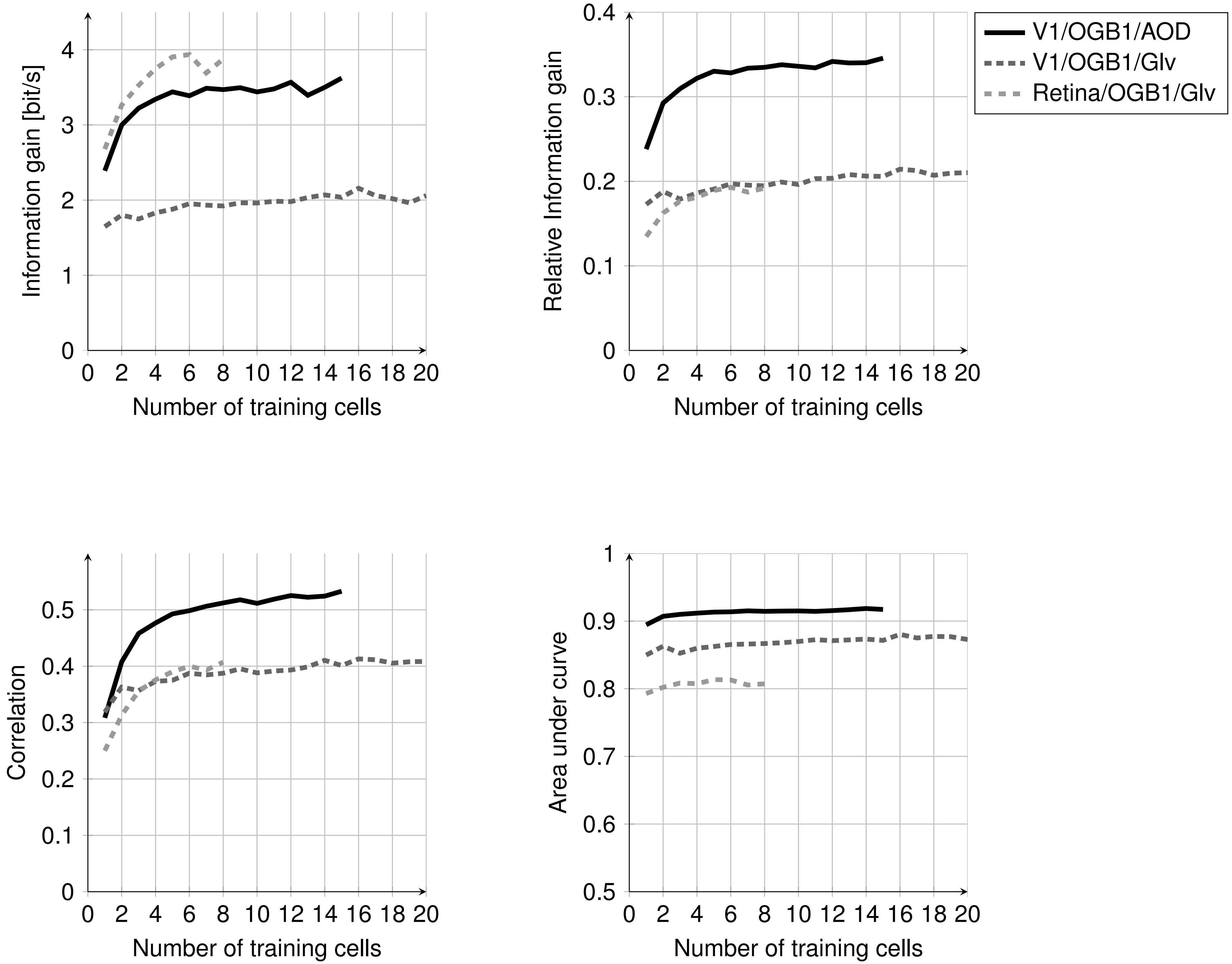


Correlation ± 2 · SEM^L&M
0.4
0.3
0.2
0.1
0
V1/GCamp6
STM
Vogelstein et al. (2010)
Raw

Comparison with other models



Training set size



Component responses