Lecture 15: *Introduction to the information bottleneck approach*

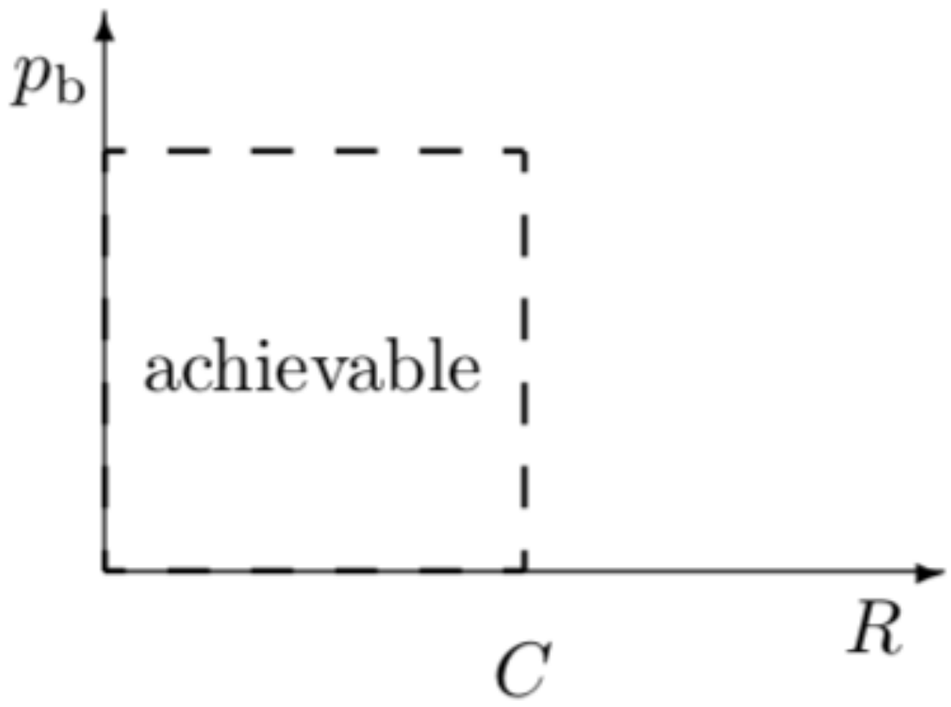
The main results of information theory and coding are more substantial:

source coding Th'm: gives bounds on compressibility of signals given their entropy

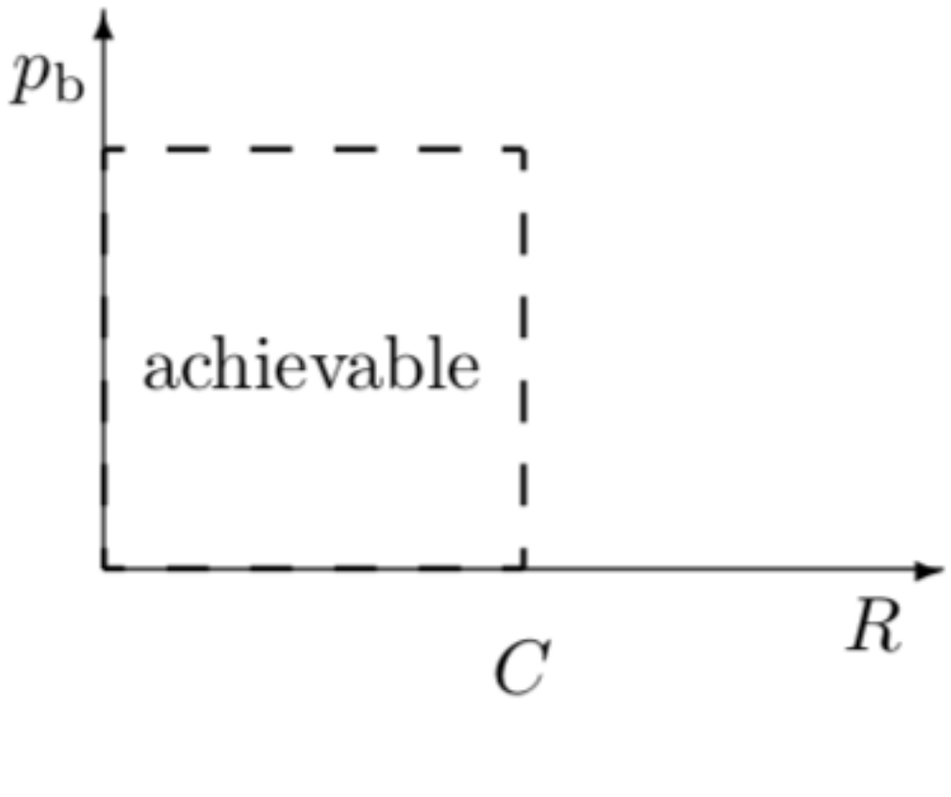
channel coding Th'm: describes the limit of the information rate of a noisy channel, and the maximal efficiency of error-correcting codes

Recall our introduction to channel coding:

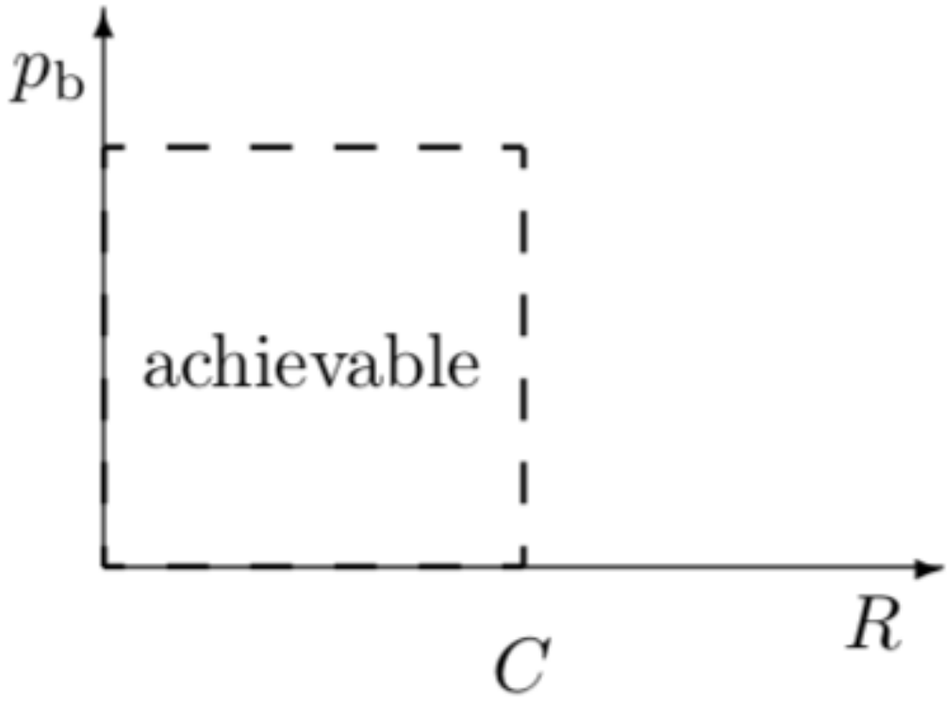
Recall our introduction to channel coding:

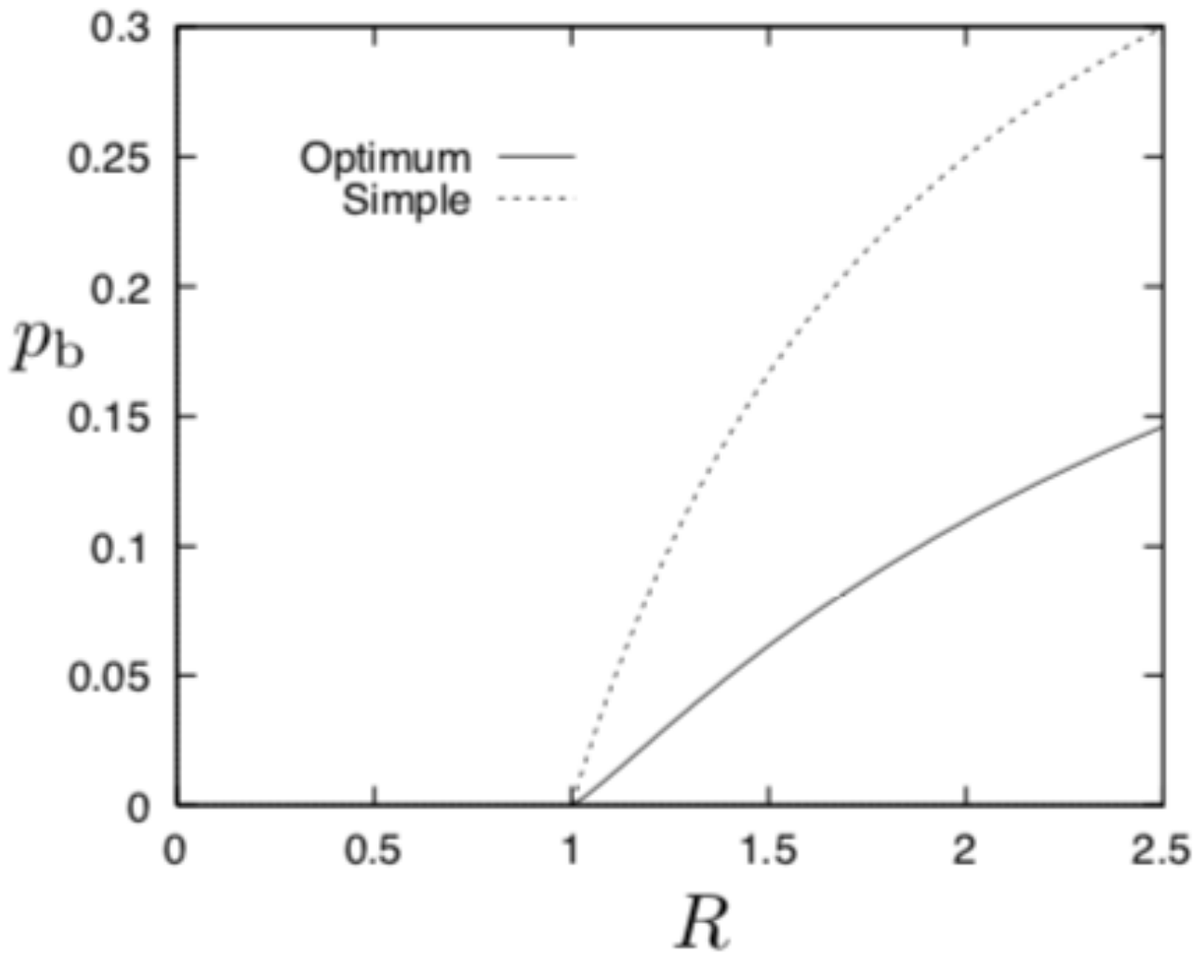


Recall our introduction to channel coding:



Recall our introduction to channel coding:

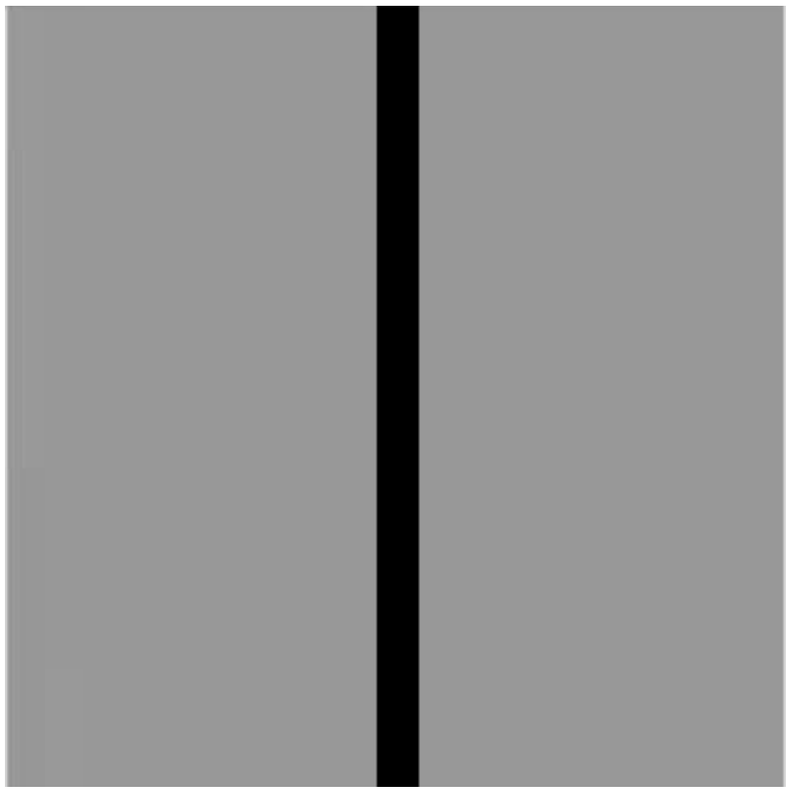




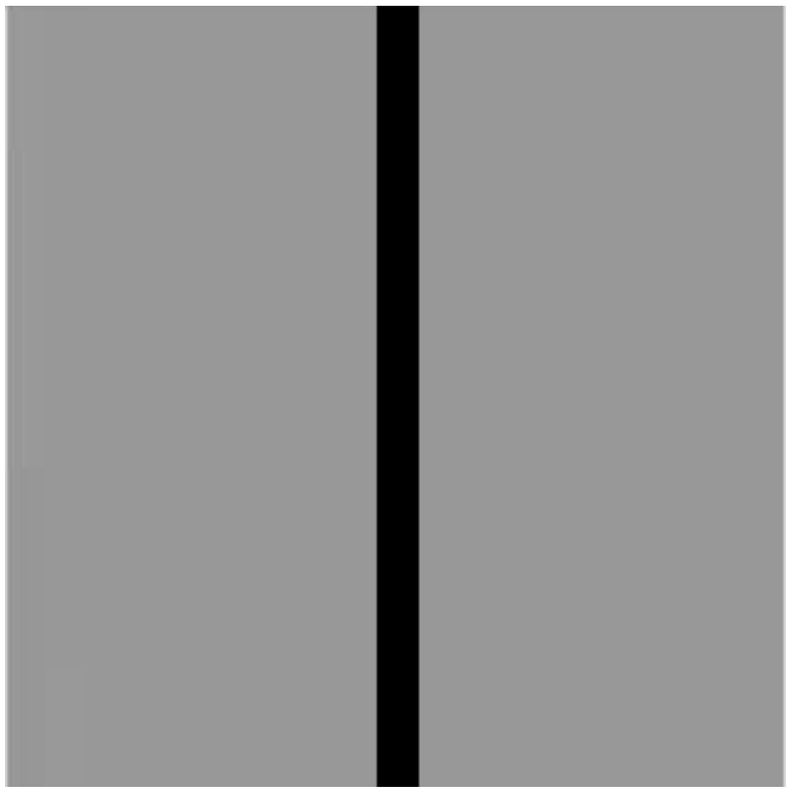
Processing delays mean the brain has to make predictions:

Processing delays mean the brain has to make predictions:

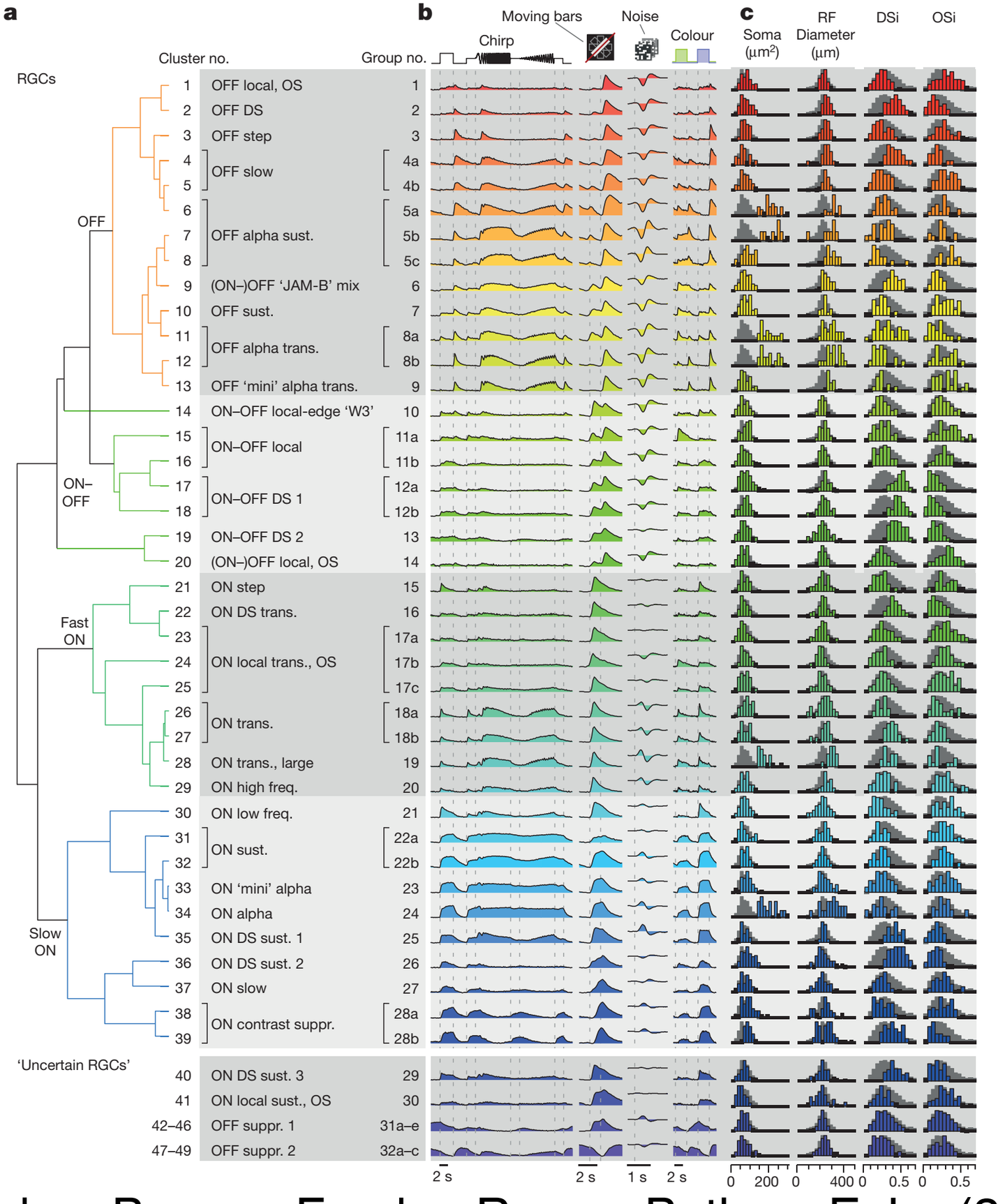
Towards more natural motion stimuli:



Towards more natural motion stimuli:

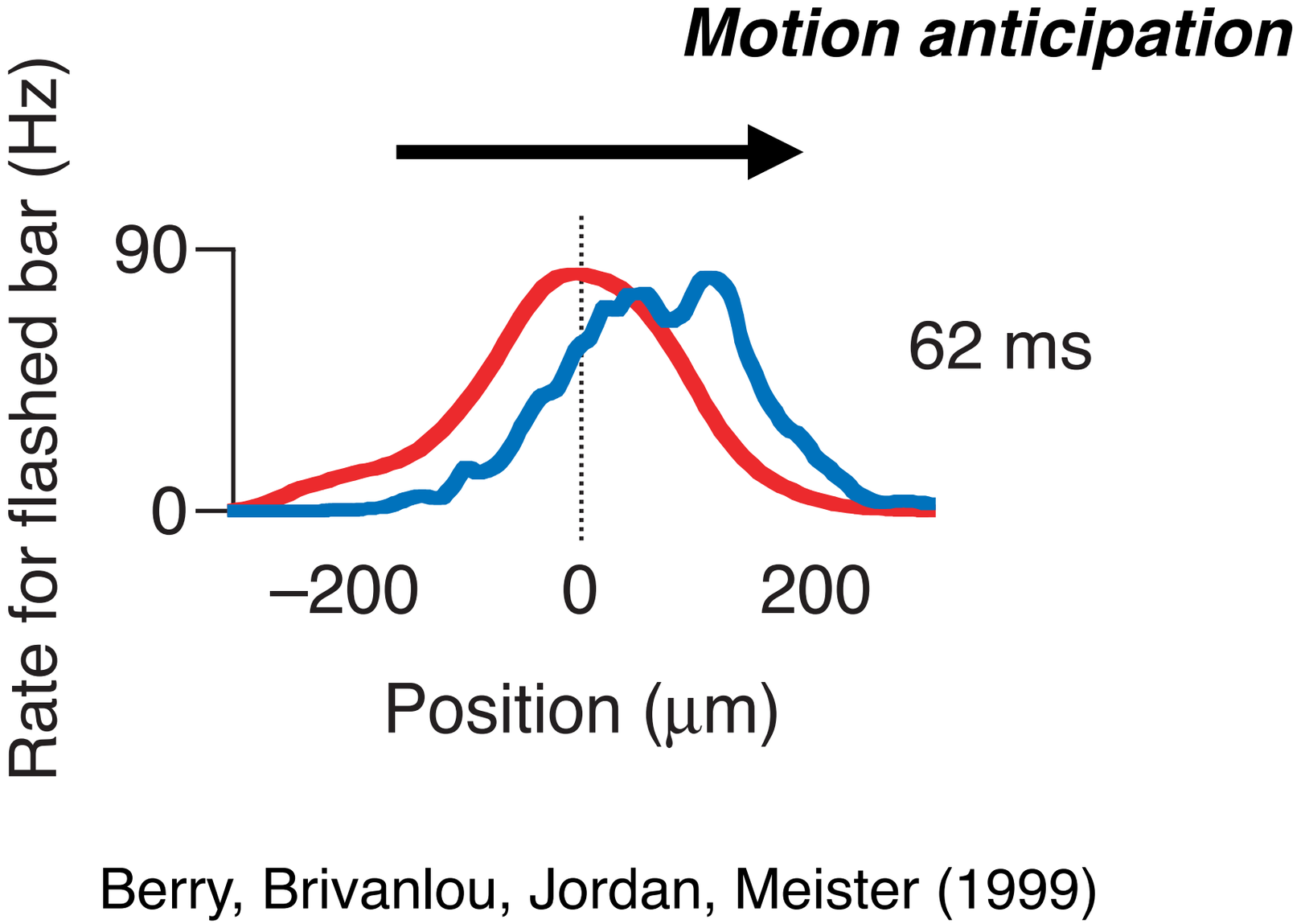


The retina is a complex piece of neural tissue:



Baden, Berens, Franke, Roson, Bethge, Euler (2016)

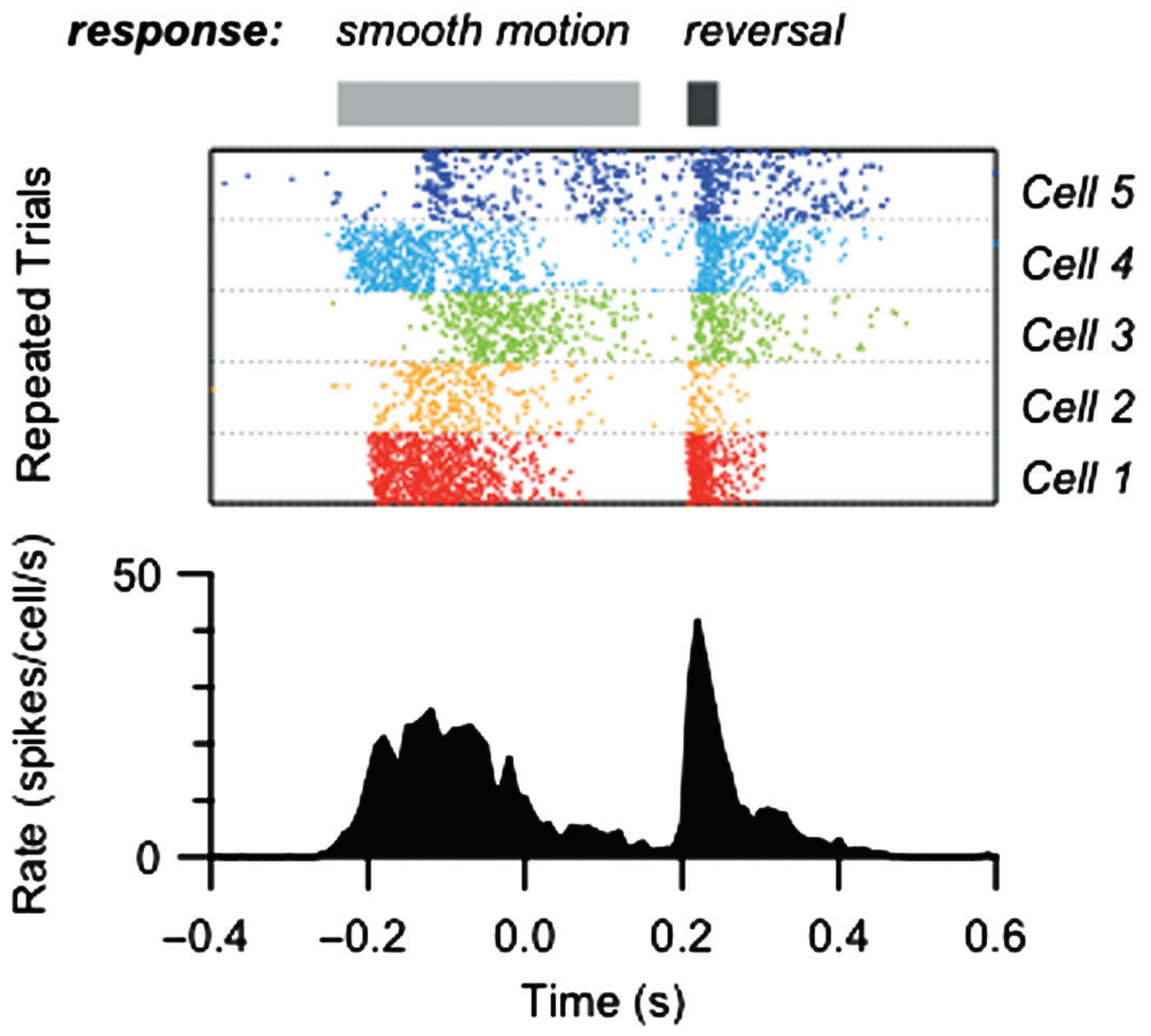
The retina does a lot of interesting things:



Berry, Brivanlou, Jordan, Meister (1999)

The retina does a lot of interesting things:

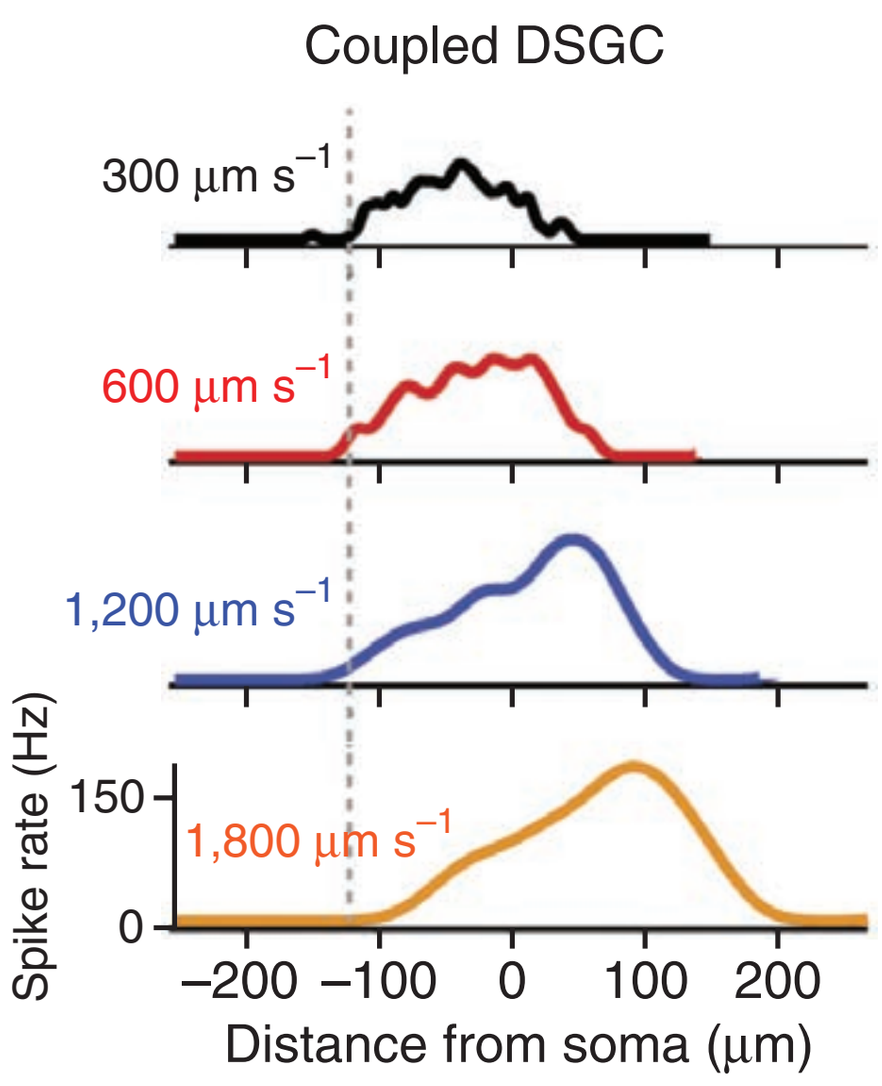
Reversal response

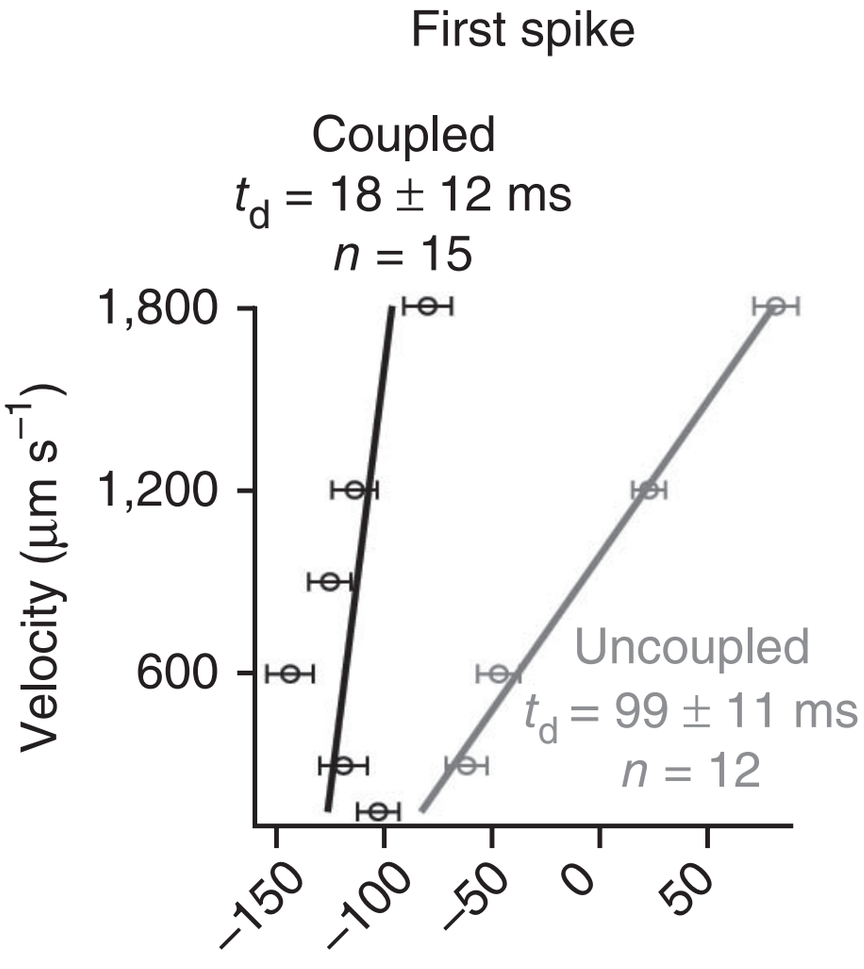


Schwartz, Taylor, Fisher, Harris, Berry (2007)

The retina does a lot of interesting things:

Lag normalization



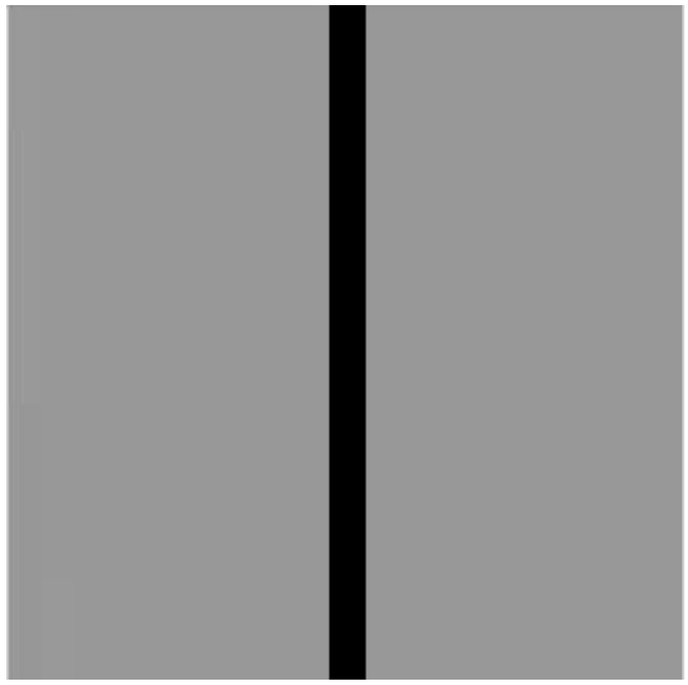


Trenholm, Schwab, Balasubramanian, Awatramani (2013)

A bar stimulus with both predictable and non-predictable motion components:

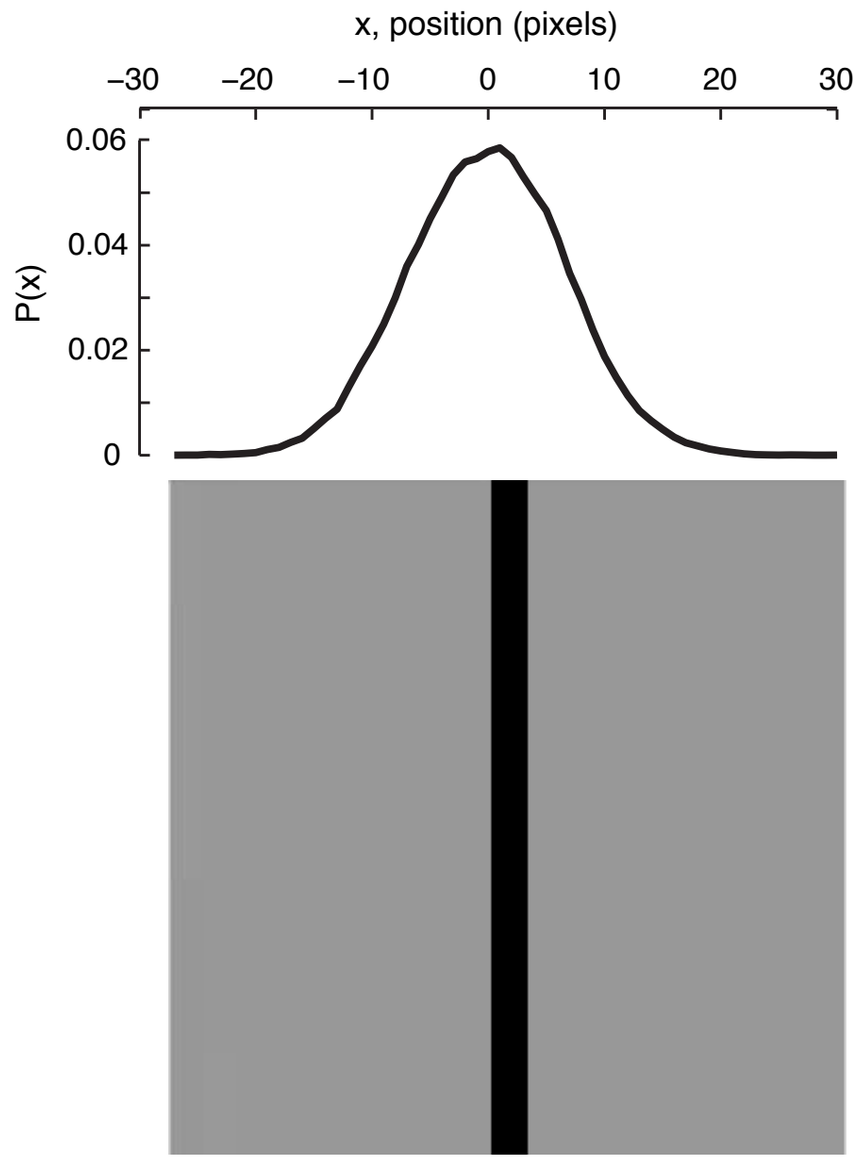
A bar stimulus with both predictable and non-predictable motion components:

A bar stimulus with both predictable and non-predictable motion components:



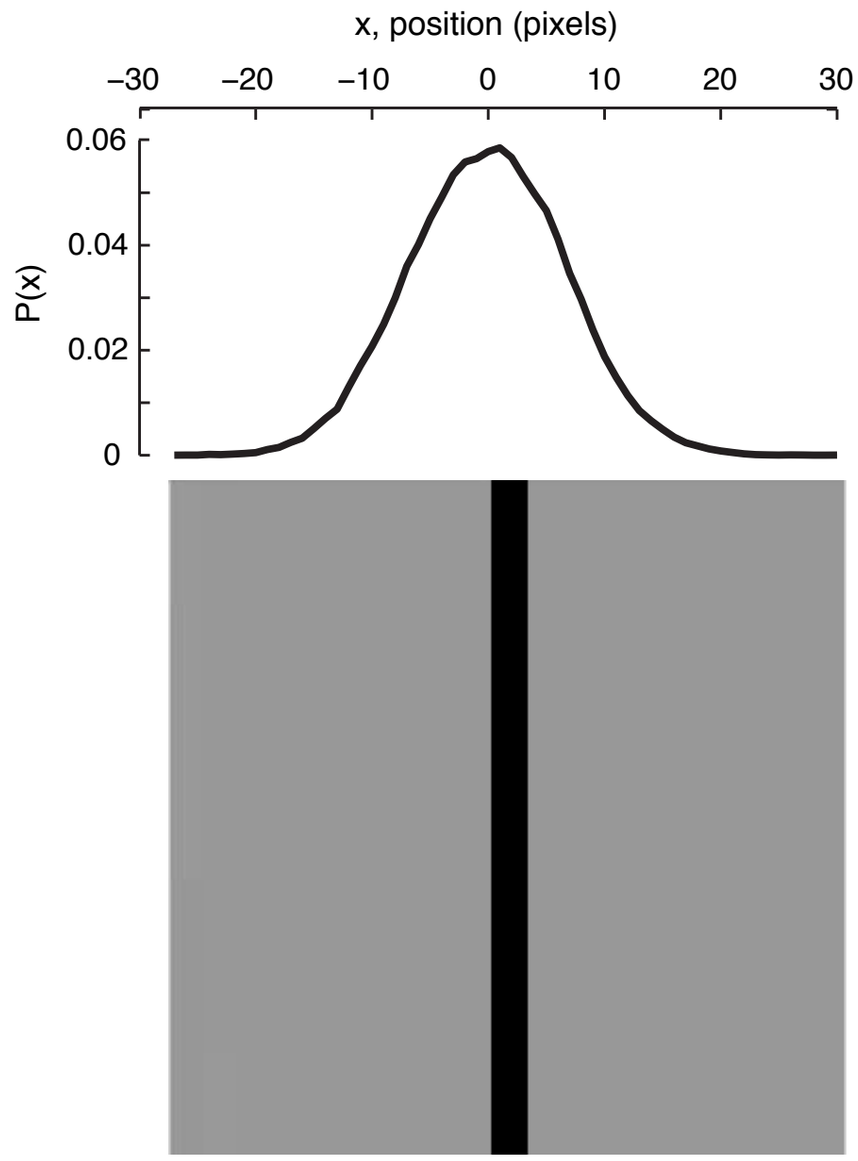
$$\frac{dv}{dt} = -\frac{v}{\tau} + D^{1/2}\Gamma(t) - \omega_0^2 x$$

A bar stimulus with both predictable and non-predictable motion components:



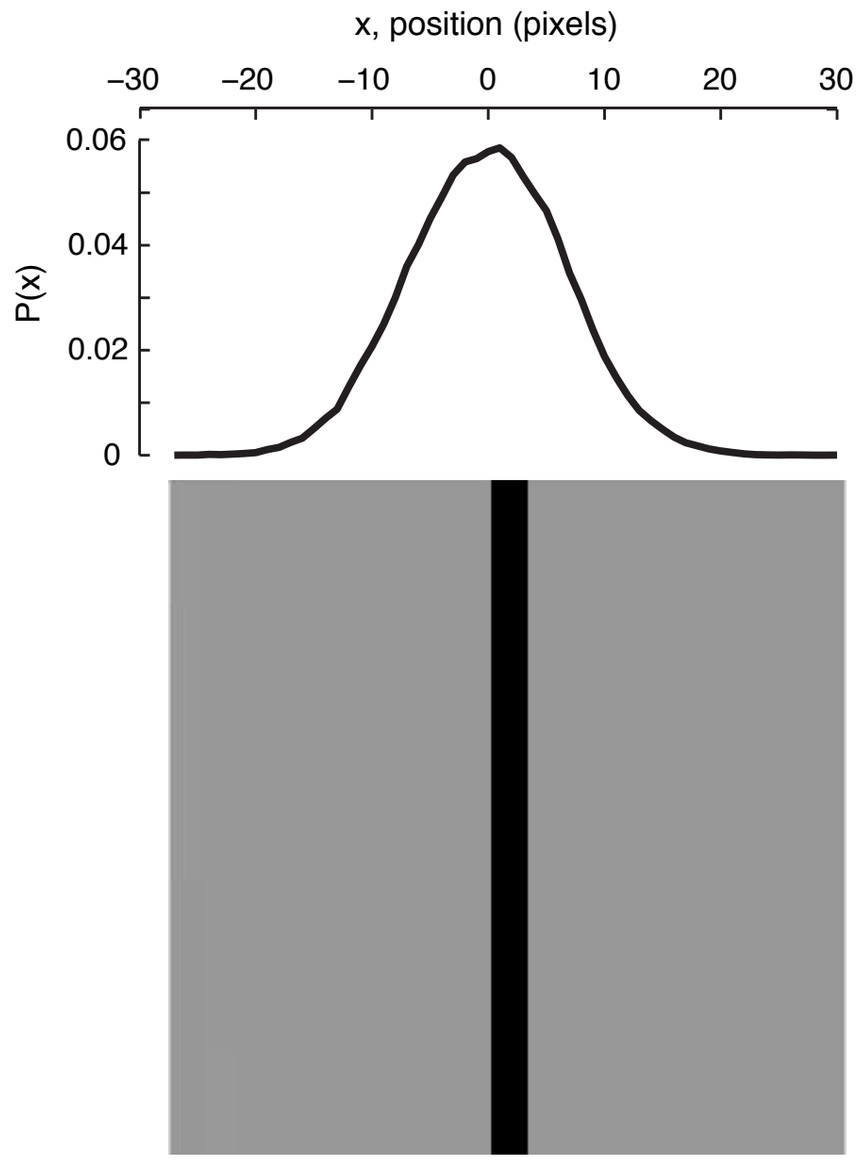
$$\frac{dv}{dt} = -\frac{v}{\tau} + D^{1/2}\Gamma(t) - \omega_0^2 x$$

A bar stimulus with both predictable and non-predictable motion components:



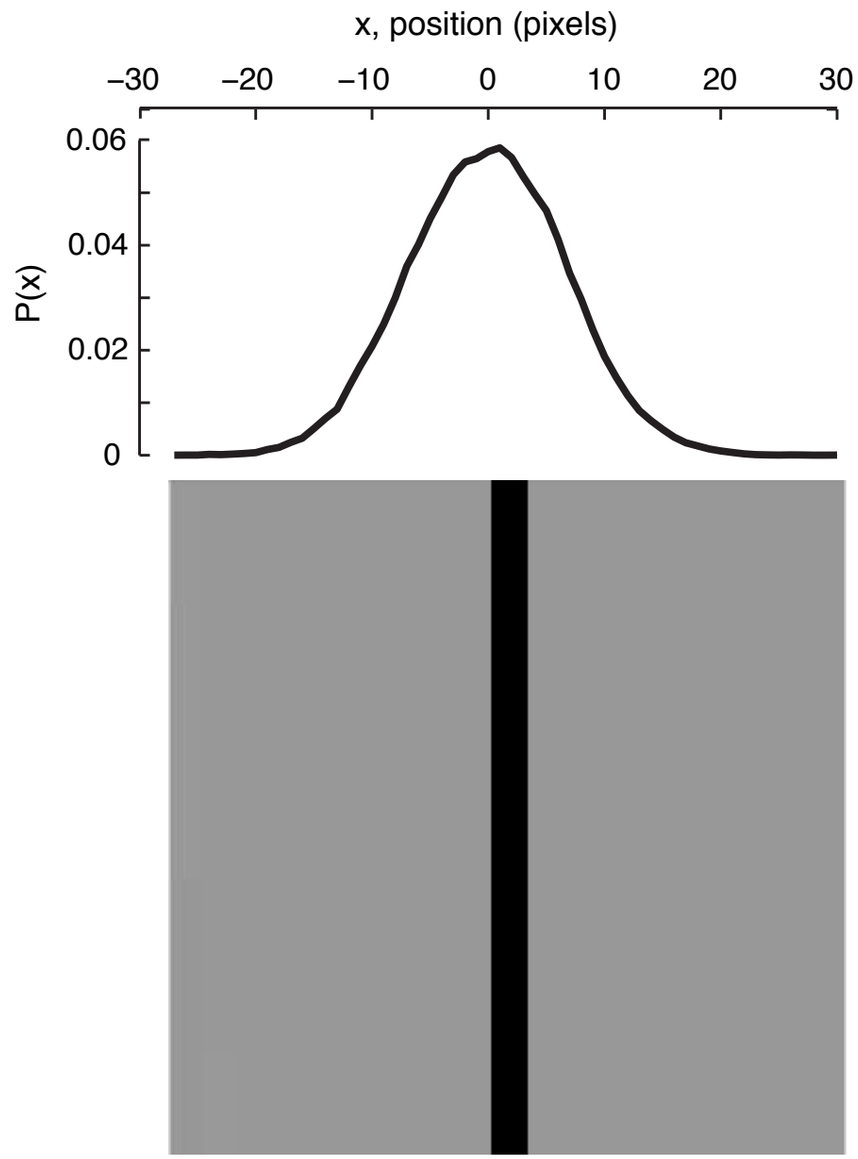
$$\frac{dv}{dt} = -\frac{v}{\tau} + D^{1/2}\Gamma(t) - \omega_0^2 x$$

A bar stimulus with both predictable and non-predictable motion components:



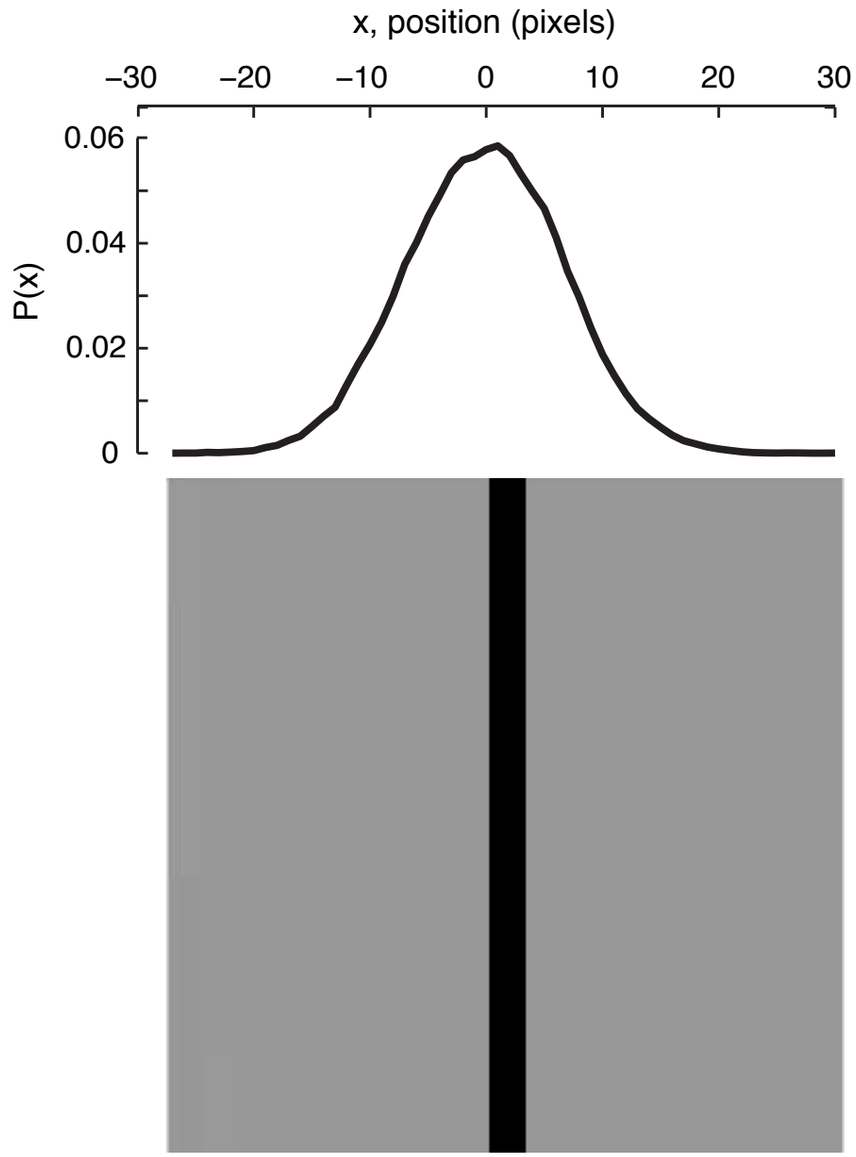
$$\frac{dv}{dt} = -\frac{v}{\tau} + D^{1/2}\Gamma(t) - \omega_0^2 x$$

A bar stimulus with both predictable and non-predictable motion components:

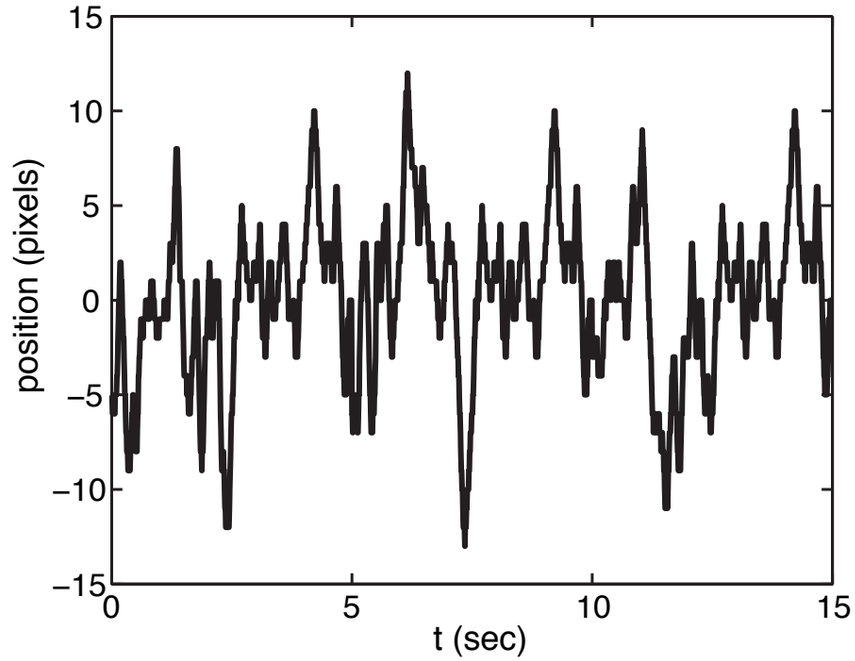


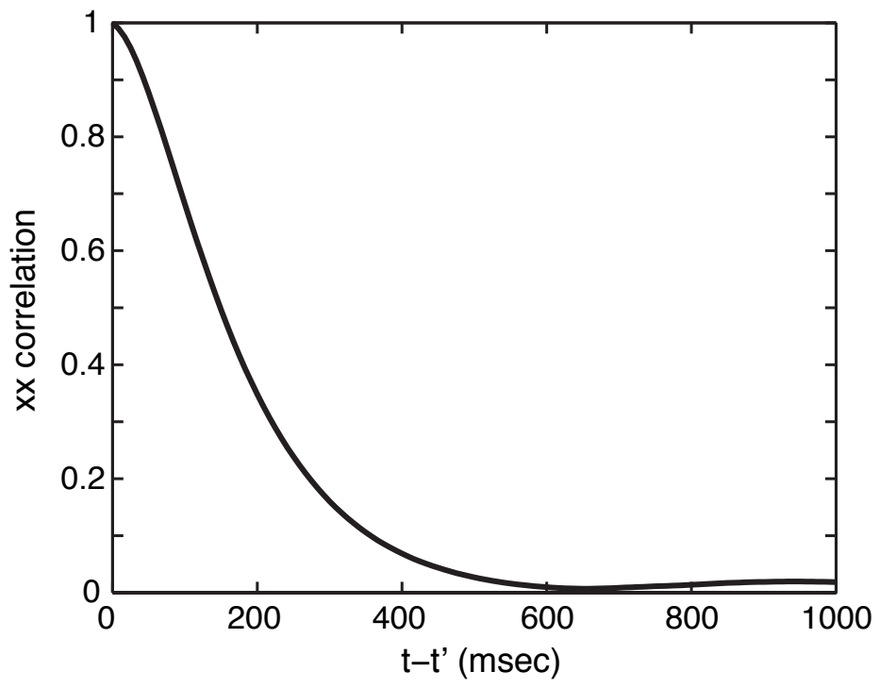
$$\frac{dv}{dt} = -\frac{v}{\tau} + D^{1/2}\Gamma(t) - \omega_0^2 x$$

A bar stimulus with both predictable and non-predictable motion components:

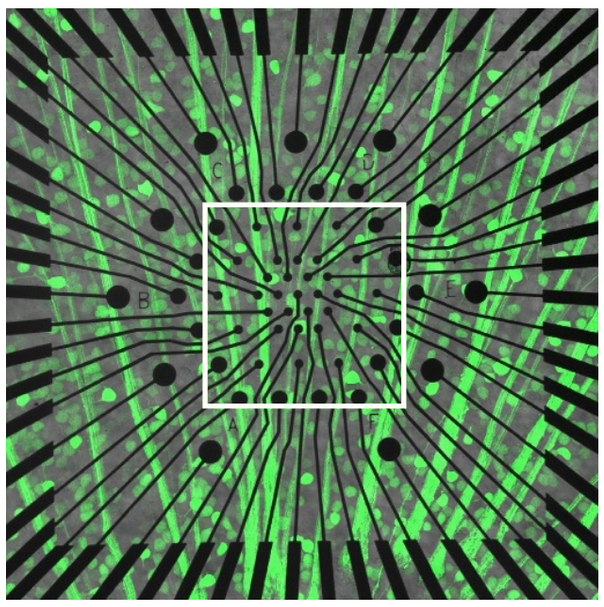


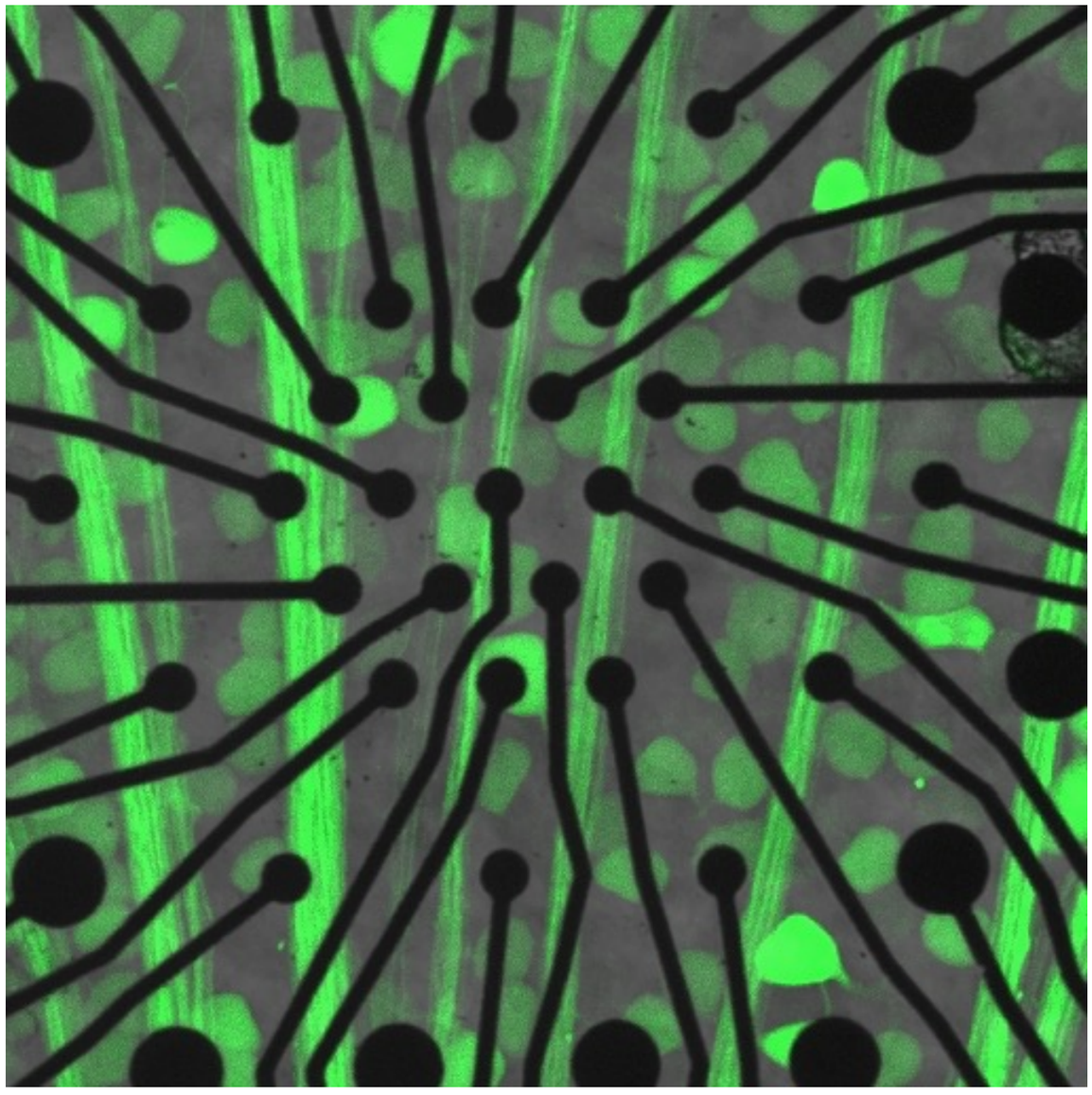
$$\frac{dv}{dt} = -\frac{v}{\tau} + D^{1/2}\Gamma(t) - \omega_0^2 x$$





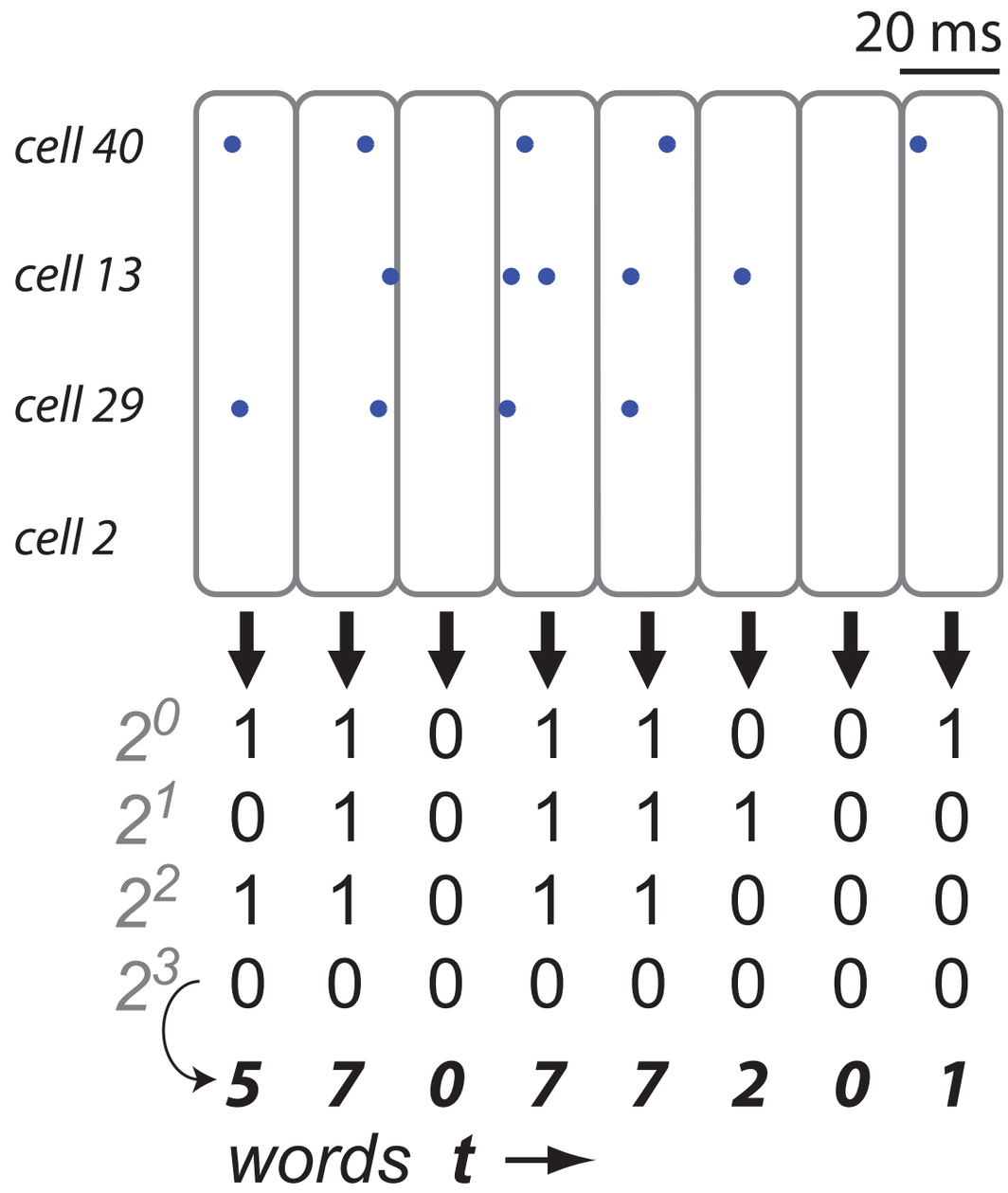
Recording from the retina:



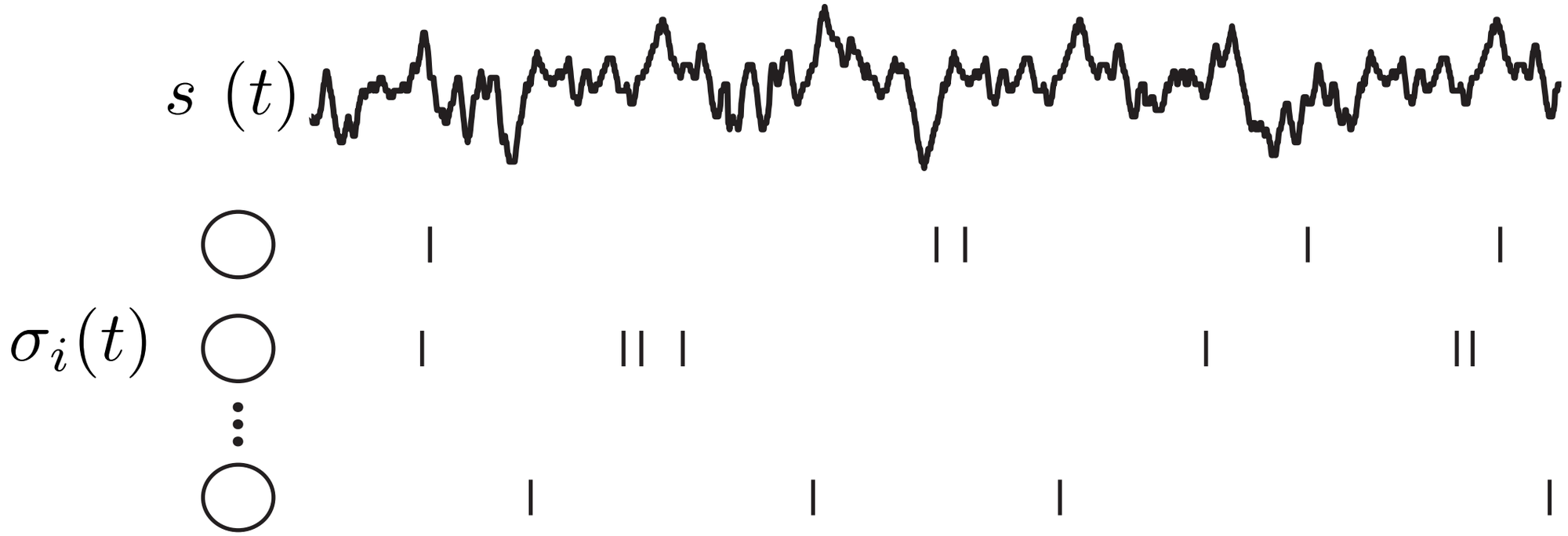


images from the Berry Lab, Ronen Segev

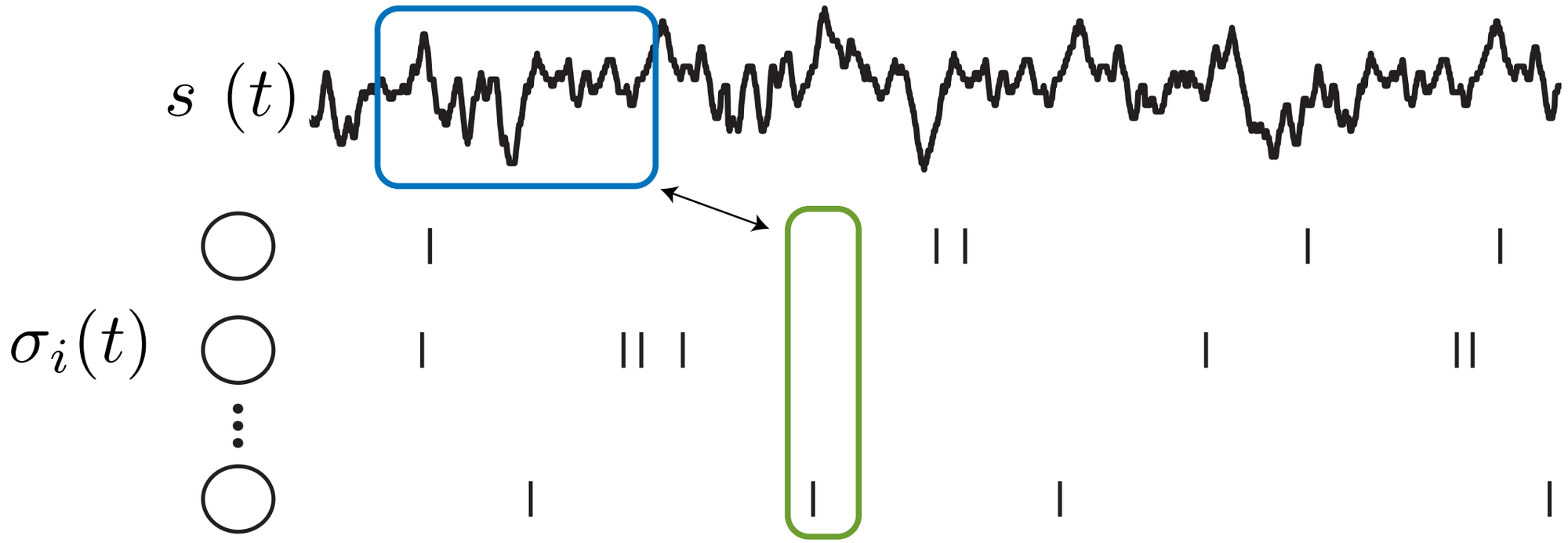
We bin time to create binary spike 'words':



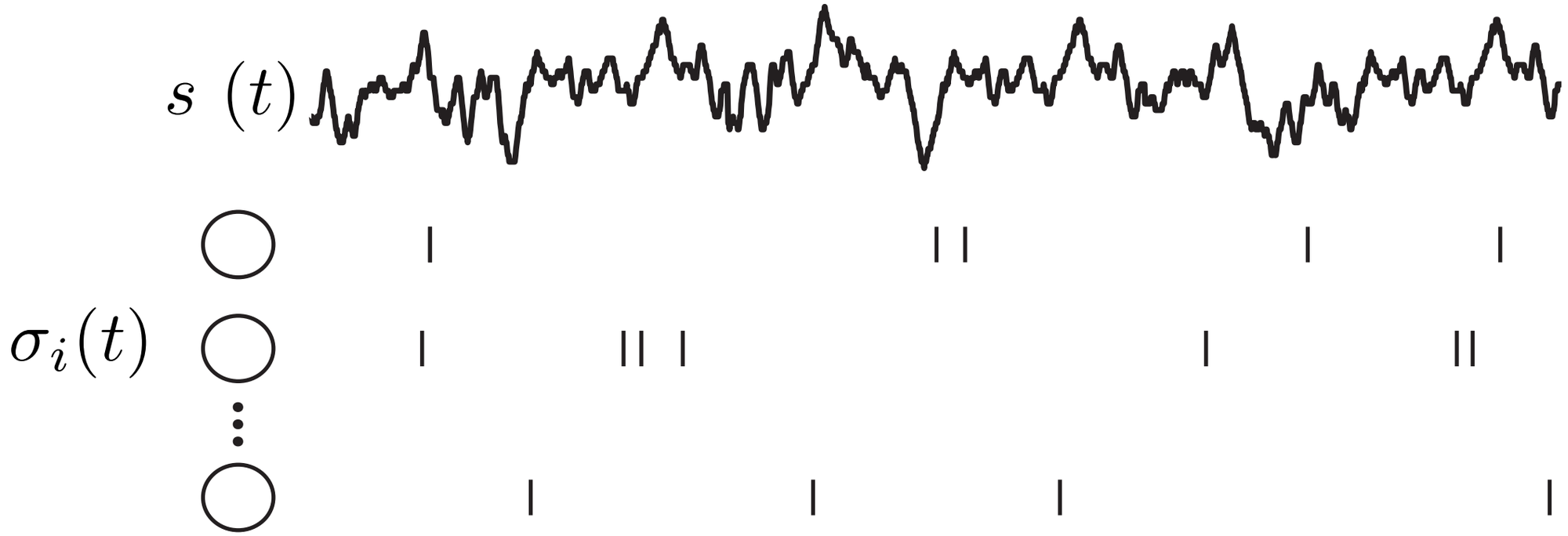
Schematic of our calculations:



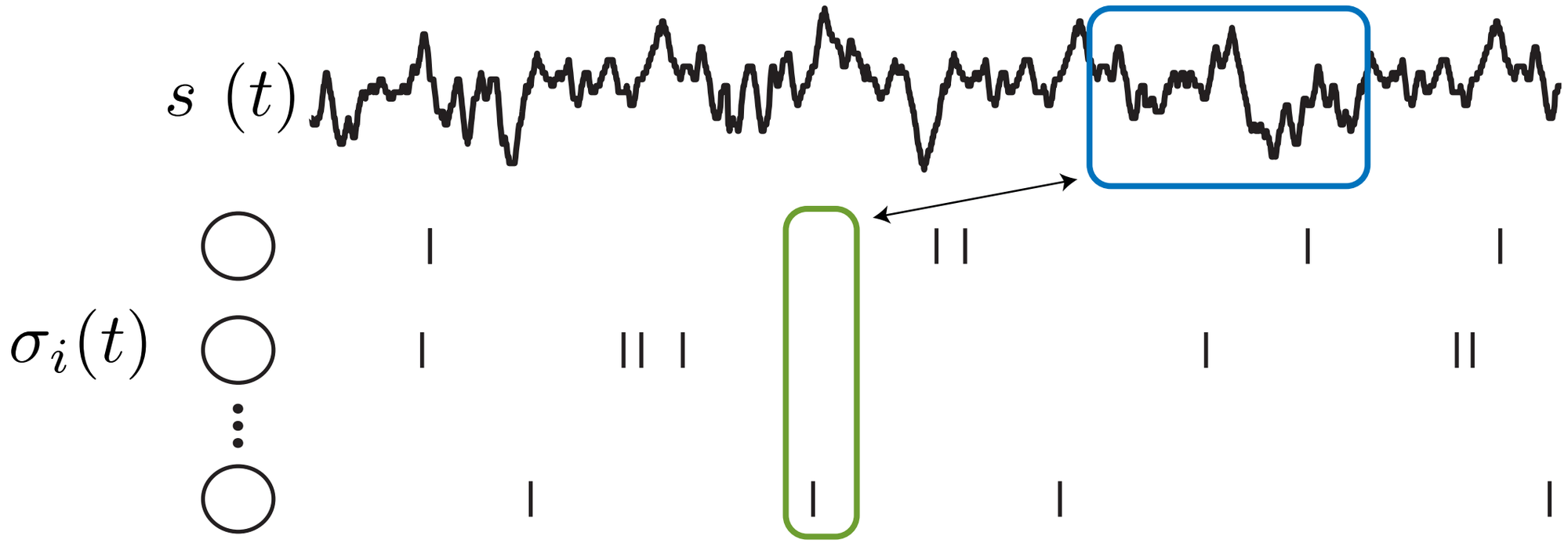
Schematic of our calculations:



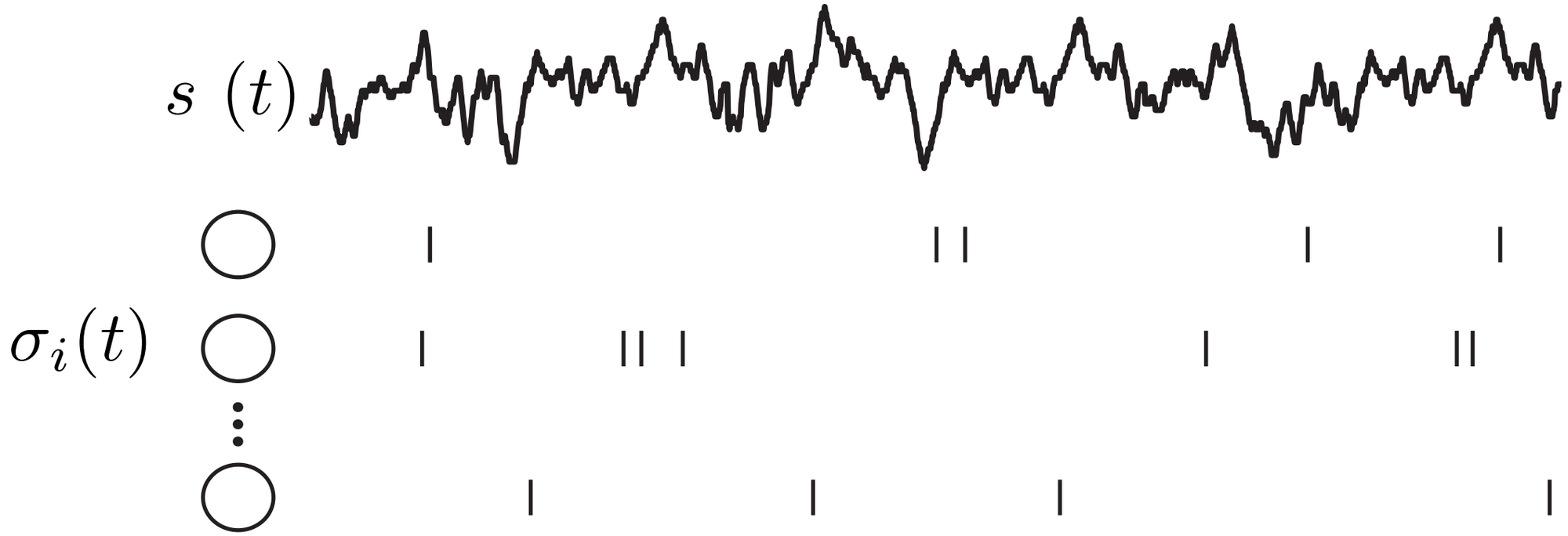
Schematic of our calculations:



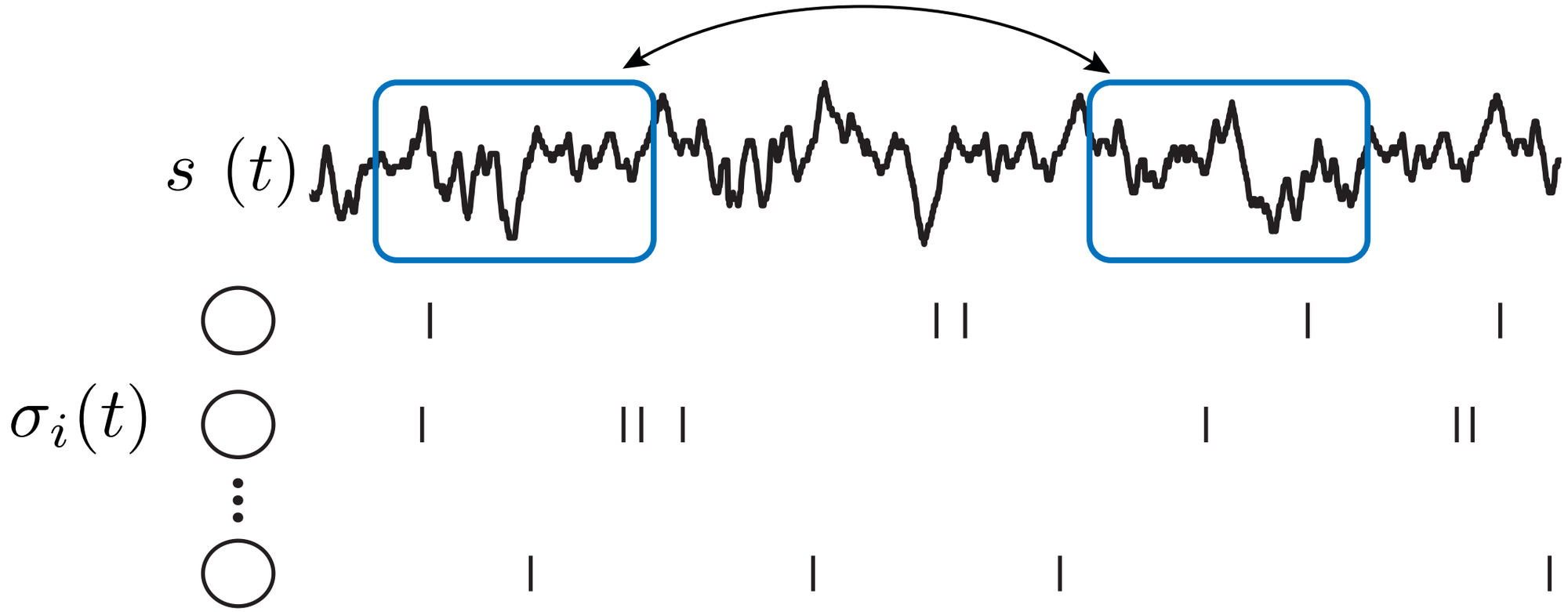
Schematic of our calculations:



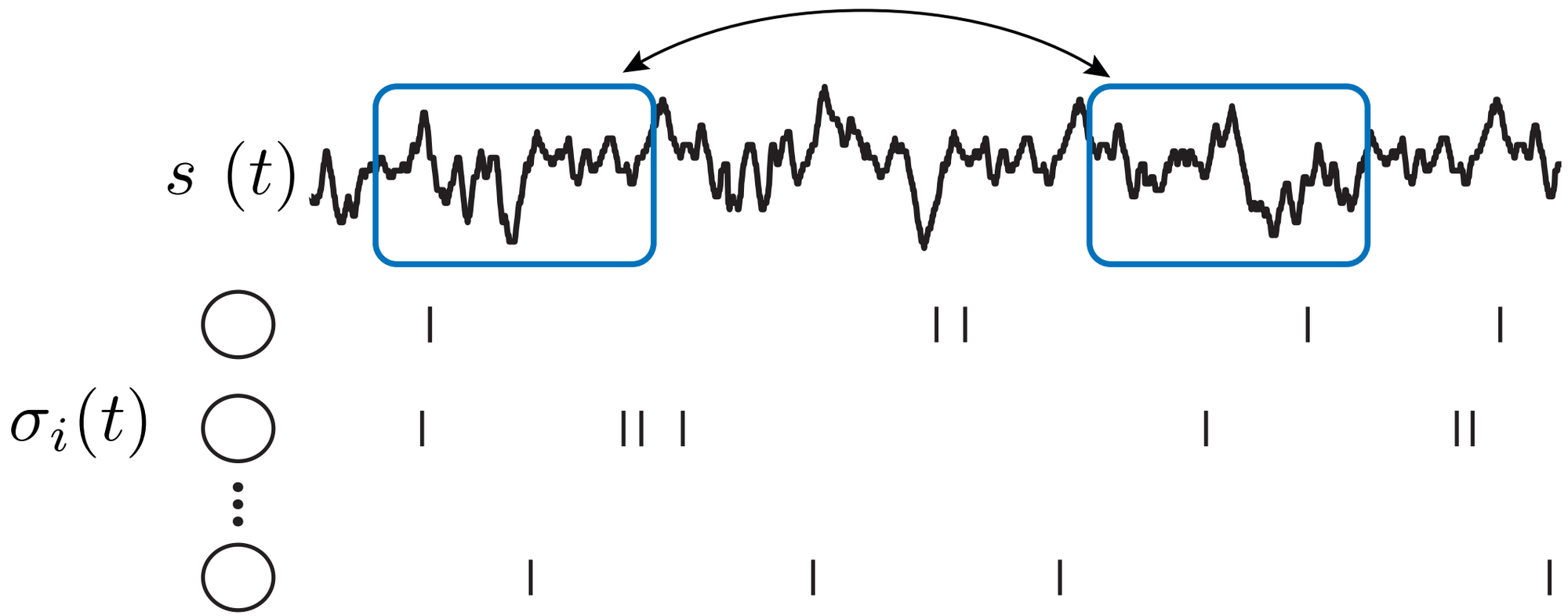
Schematic of our calculations:



Schematic of our calculations:

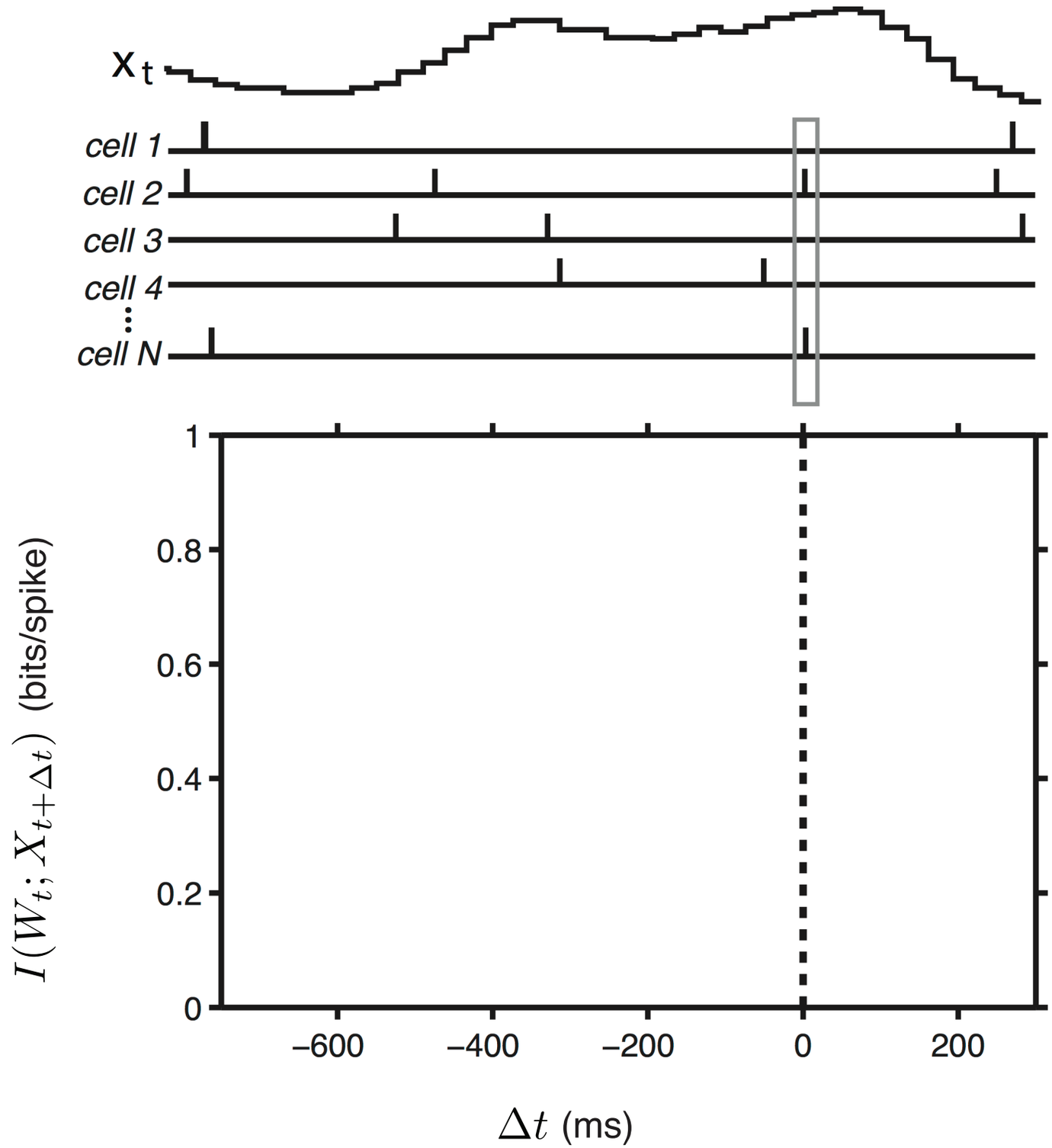


Schematic of our calculations:

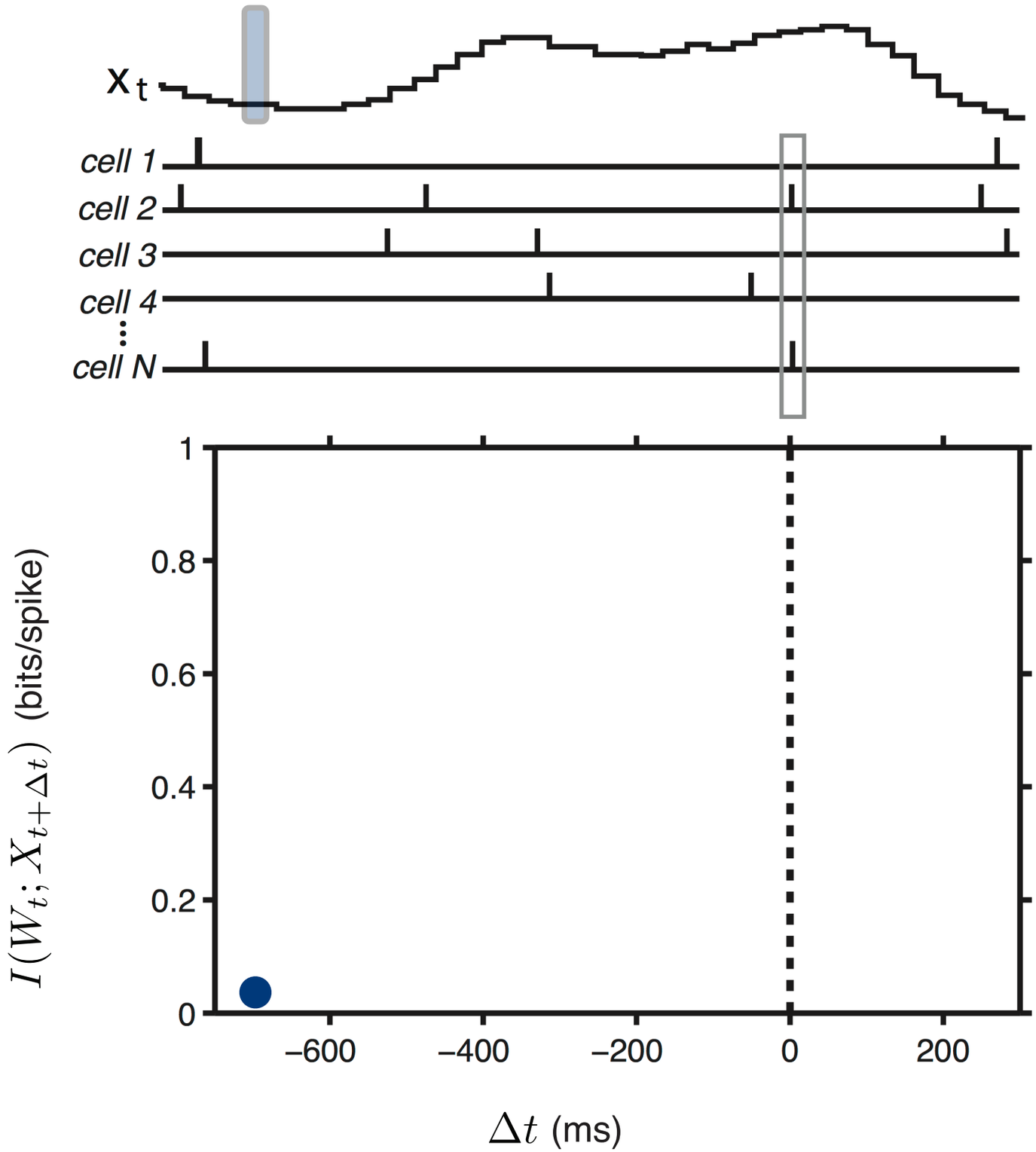


$$I(\text{past};\text{future}) = S(\text{future}) - S(\text{future}|\text{past})$$

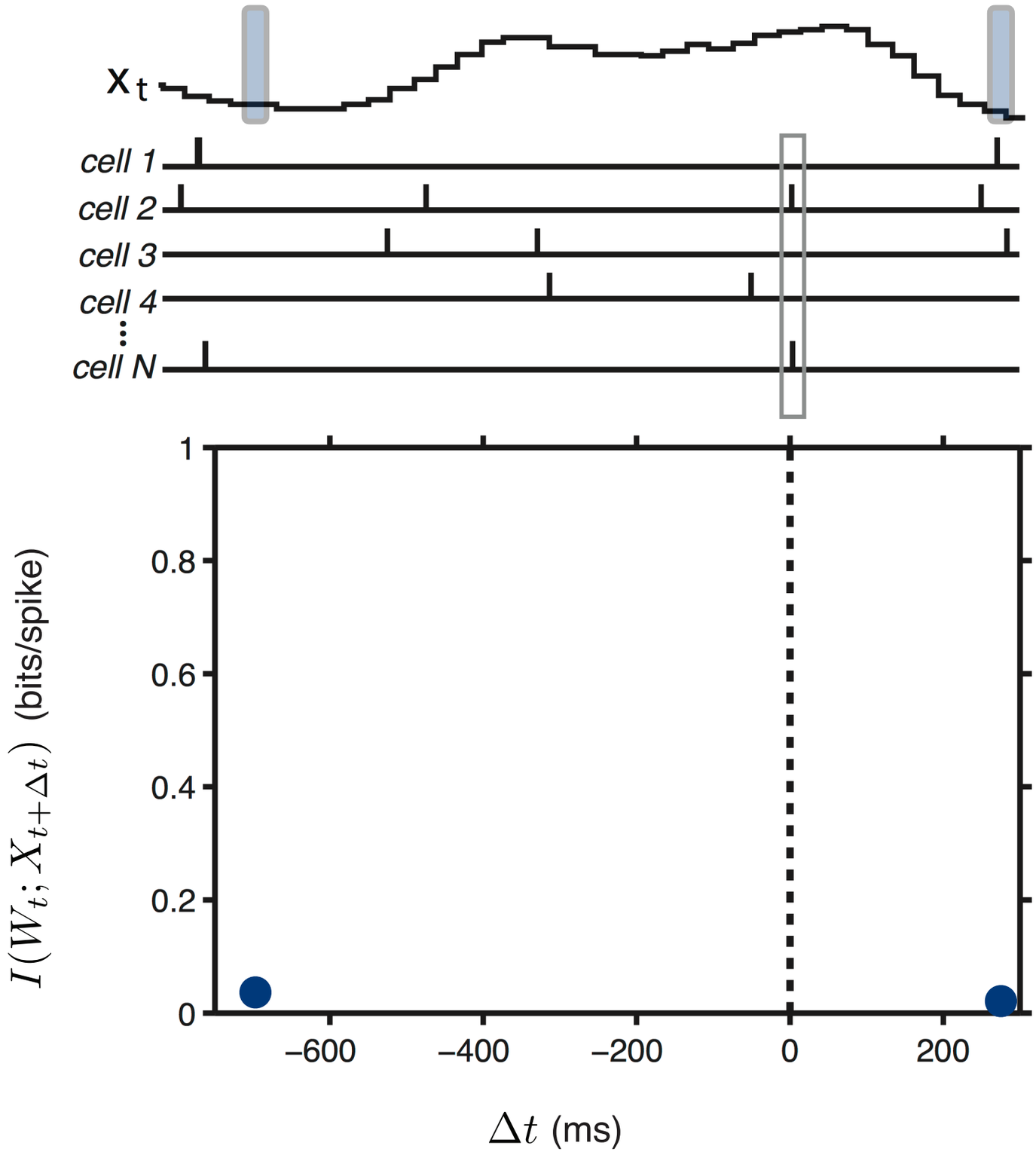
Retina populations carry info about the future:



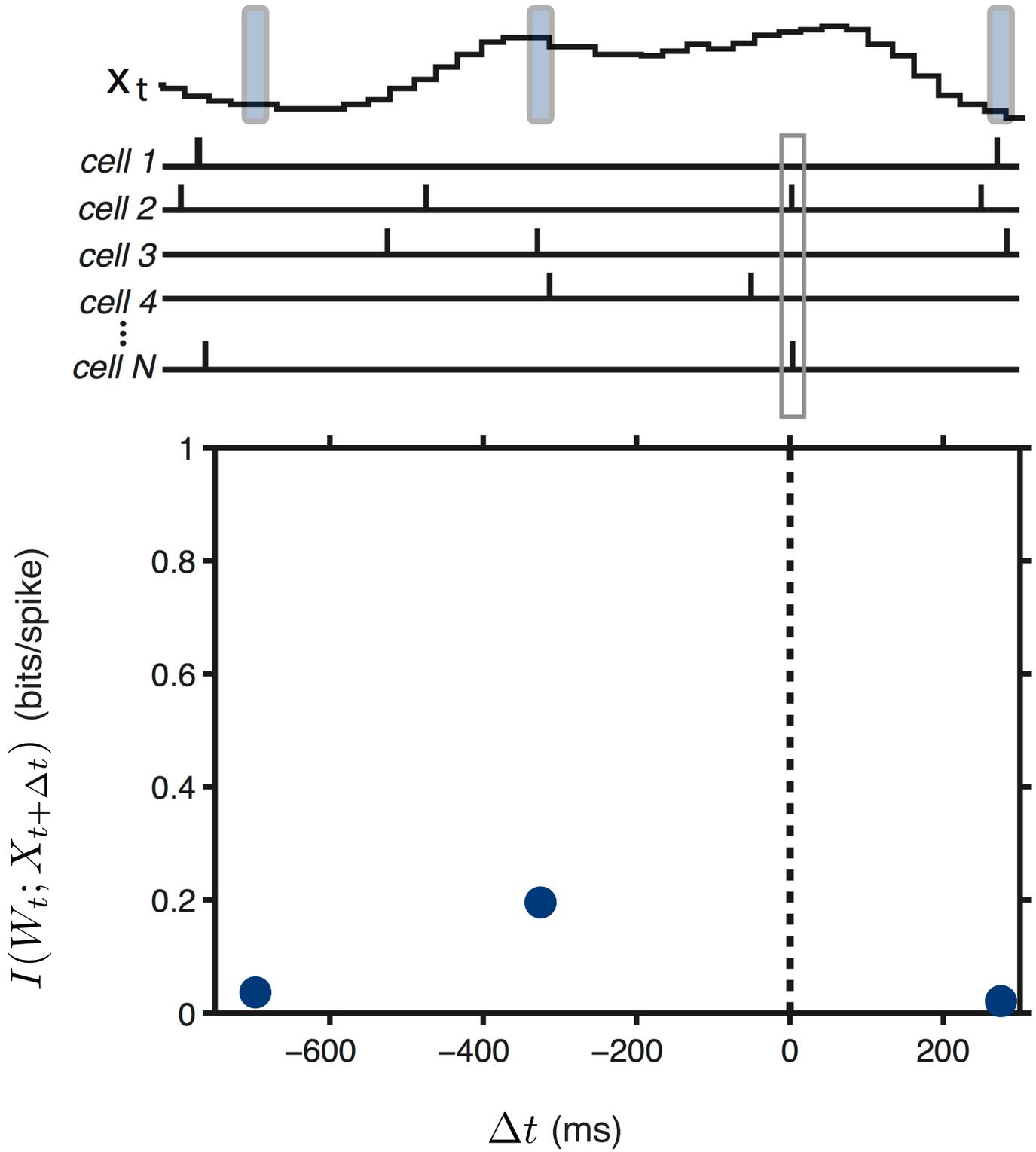
Retina populations carry info about the future:



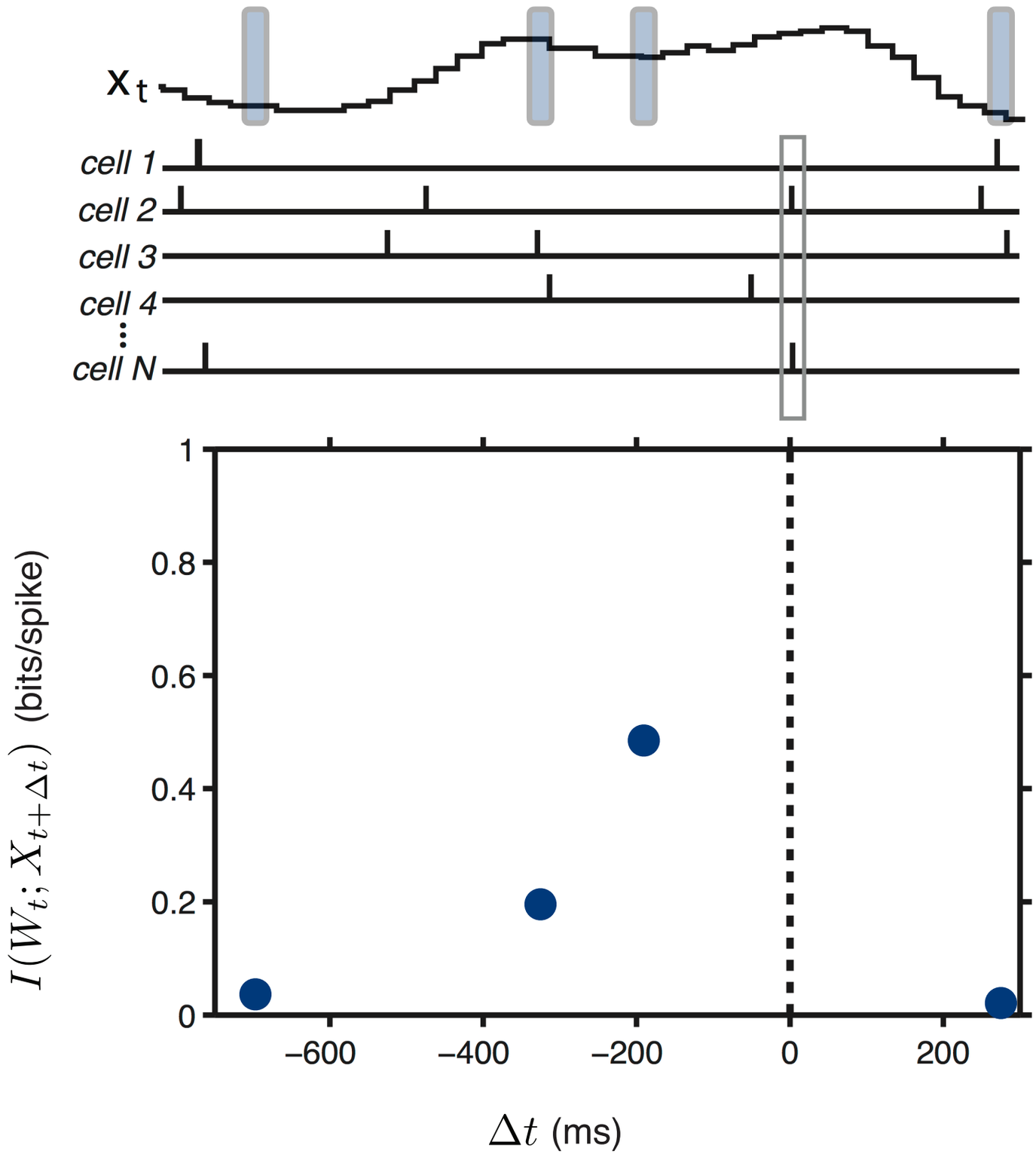
Retina populations carry info about the future:



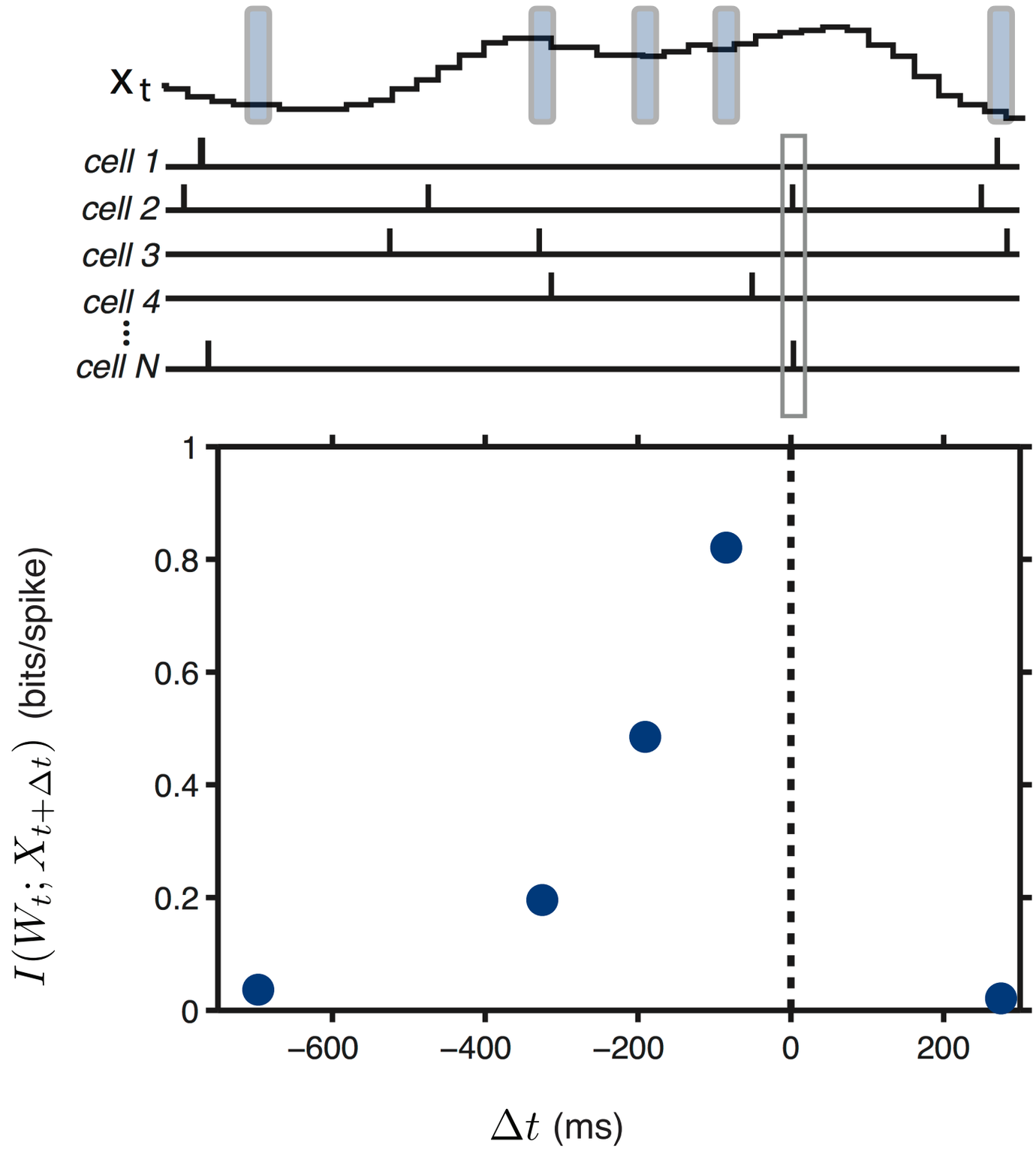
Retina populations carry info about the future:



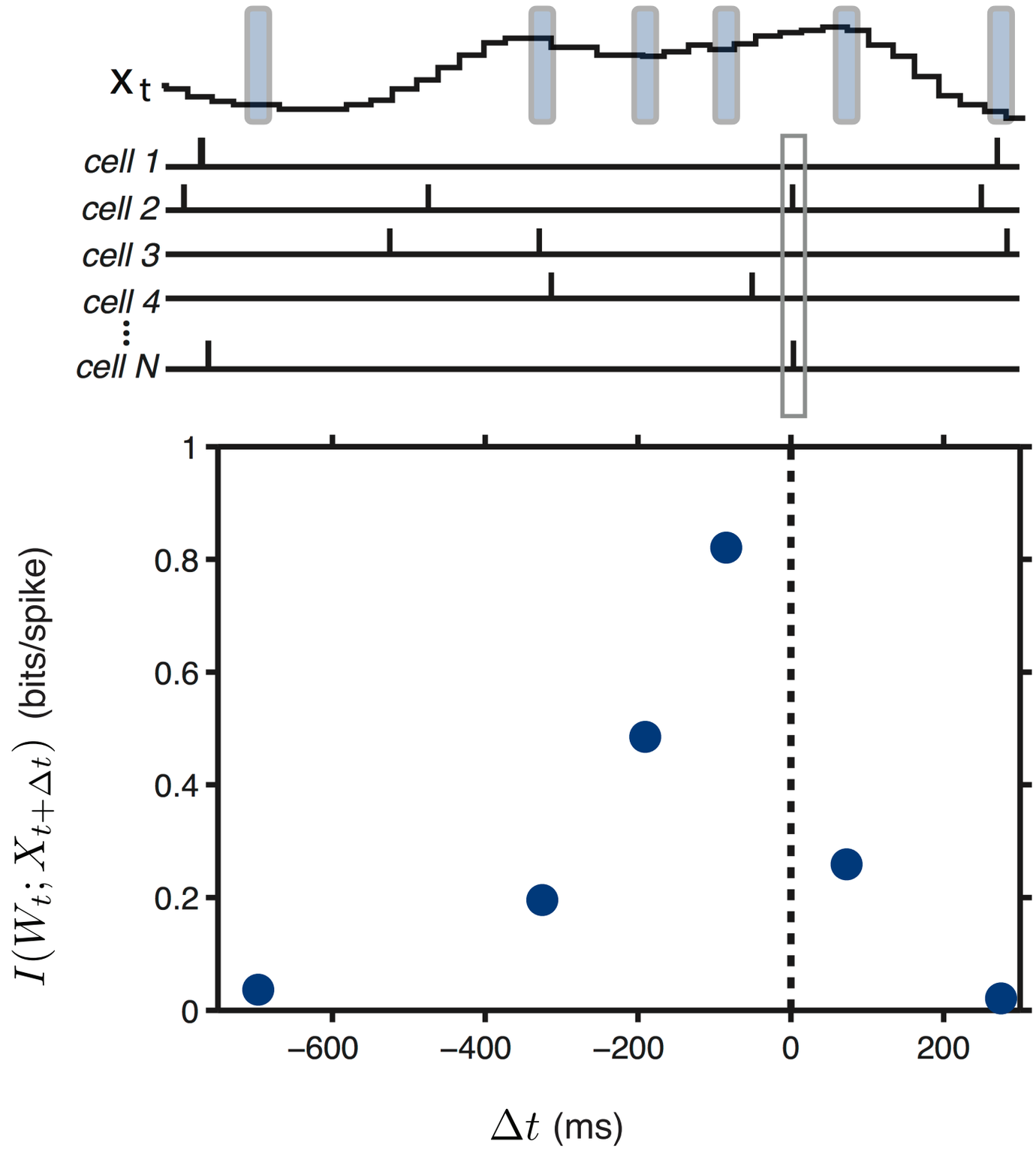
Retina populations carry info about the future:



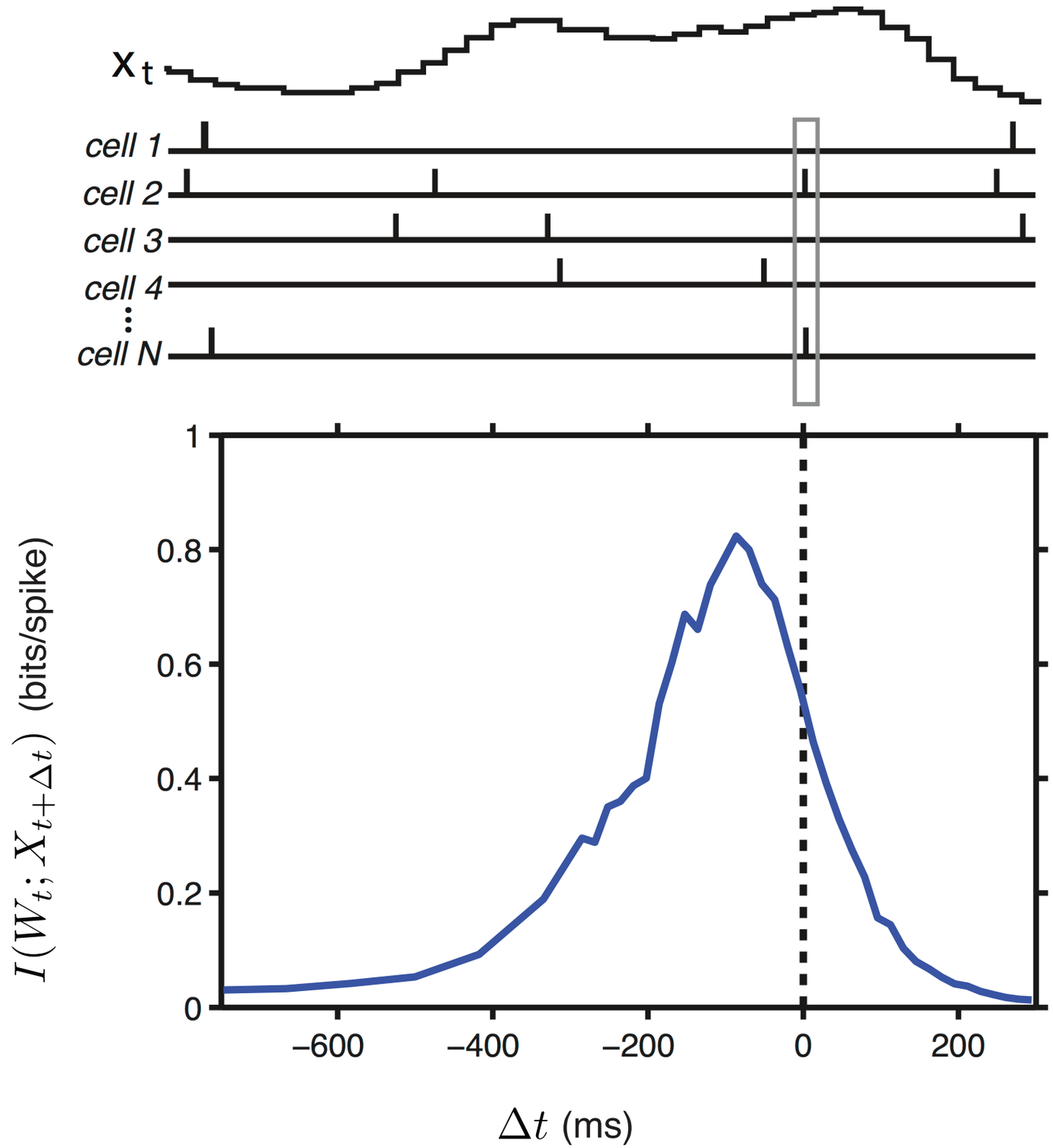
Retina populations carry info about the future:



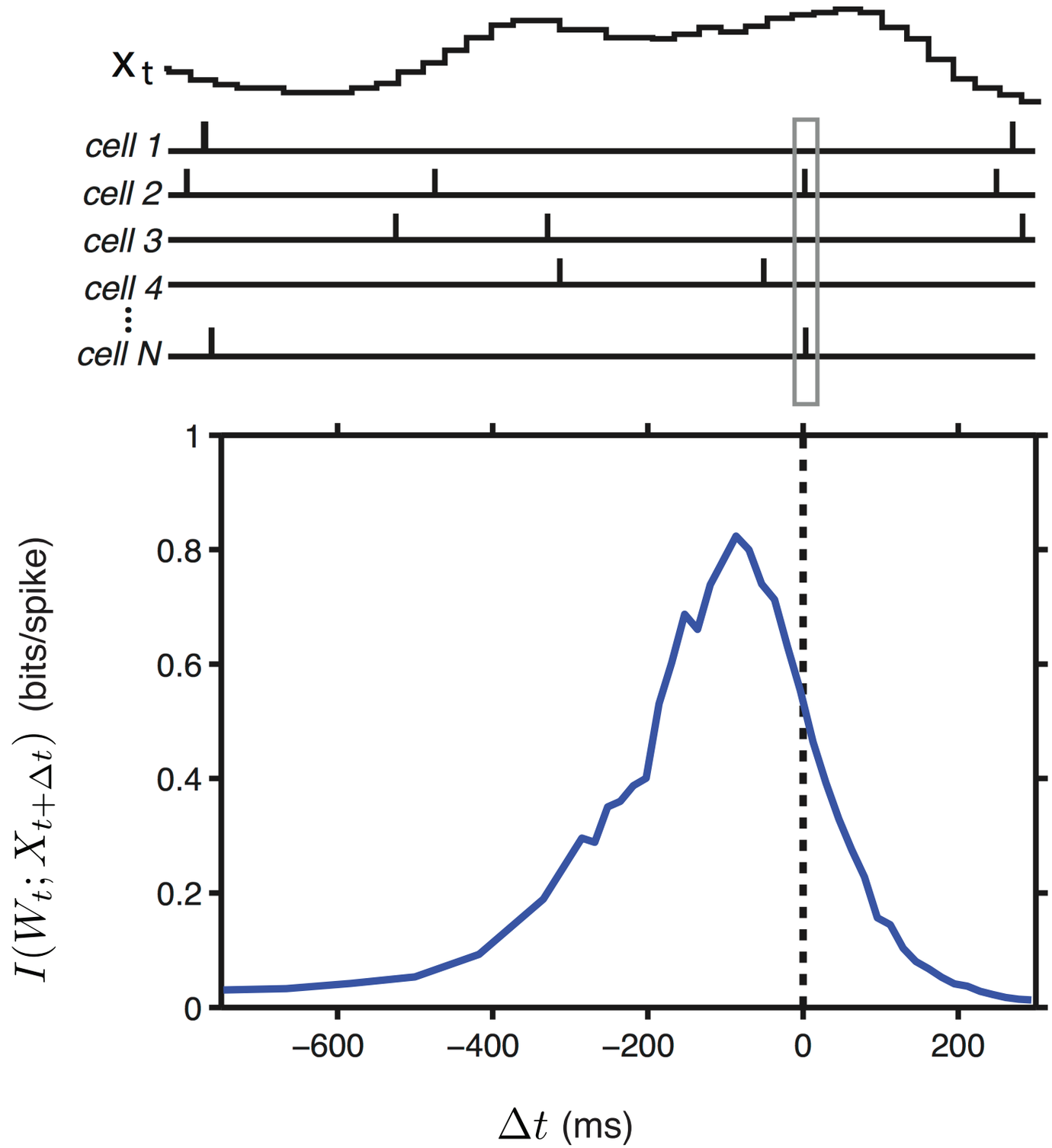
Retina populations carry info about the future:



Retina populations carry info about the future:

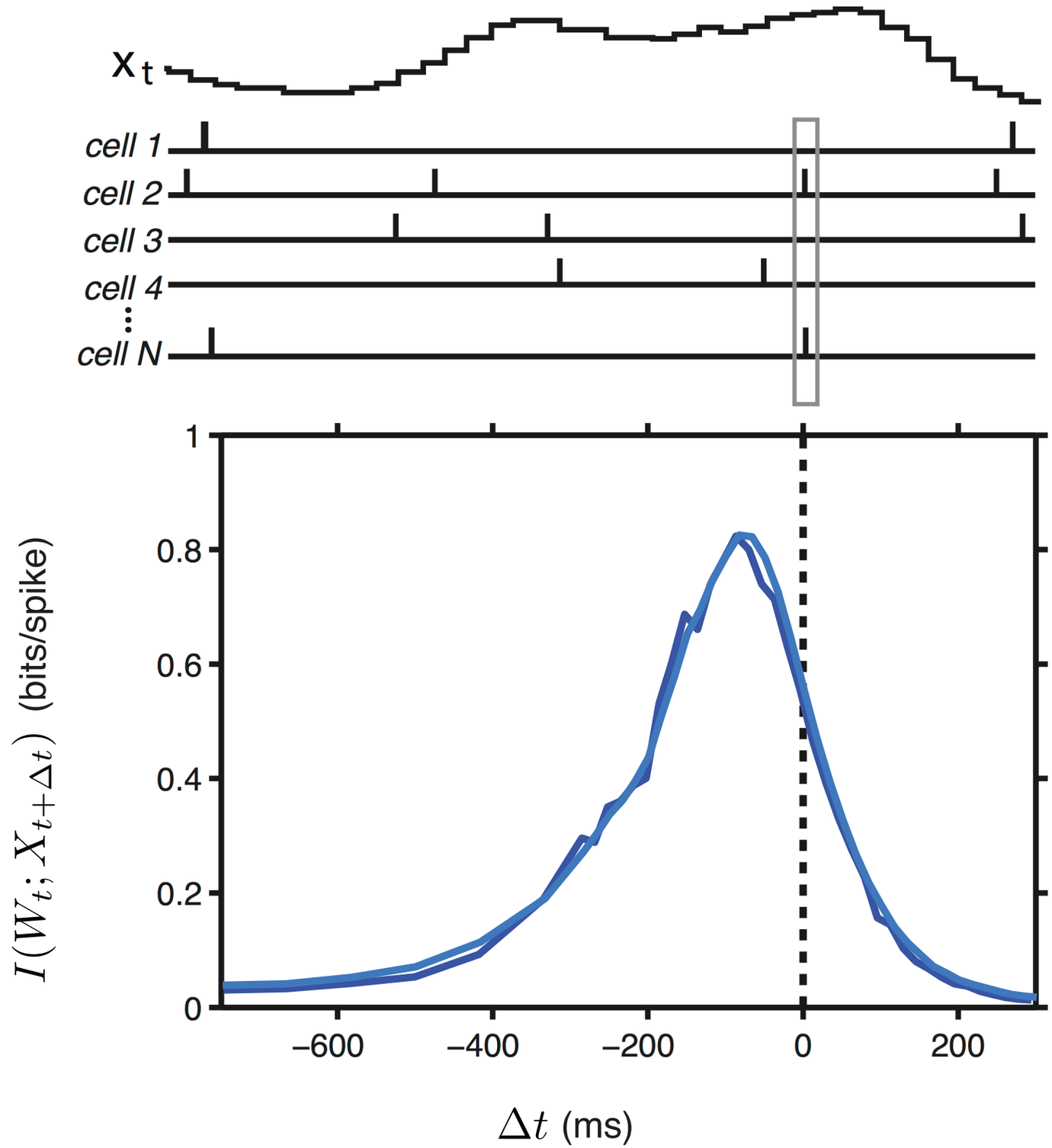


Retina populations carry info about the future:



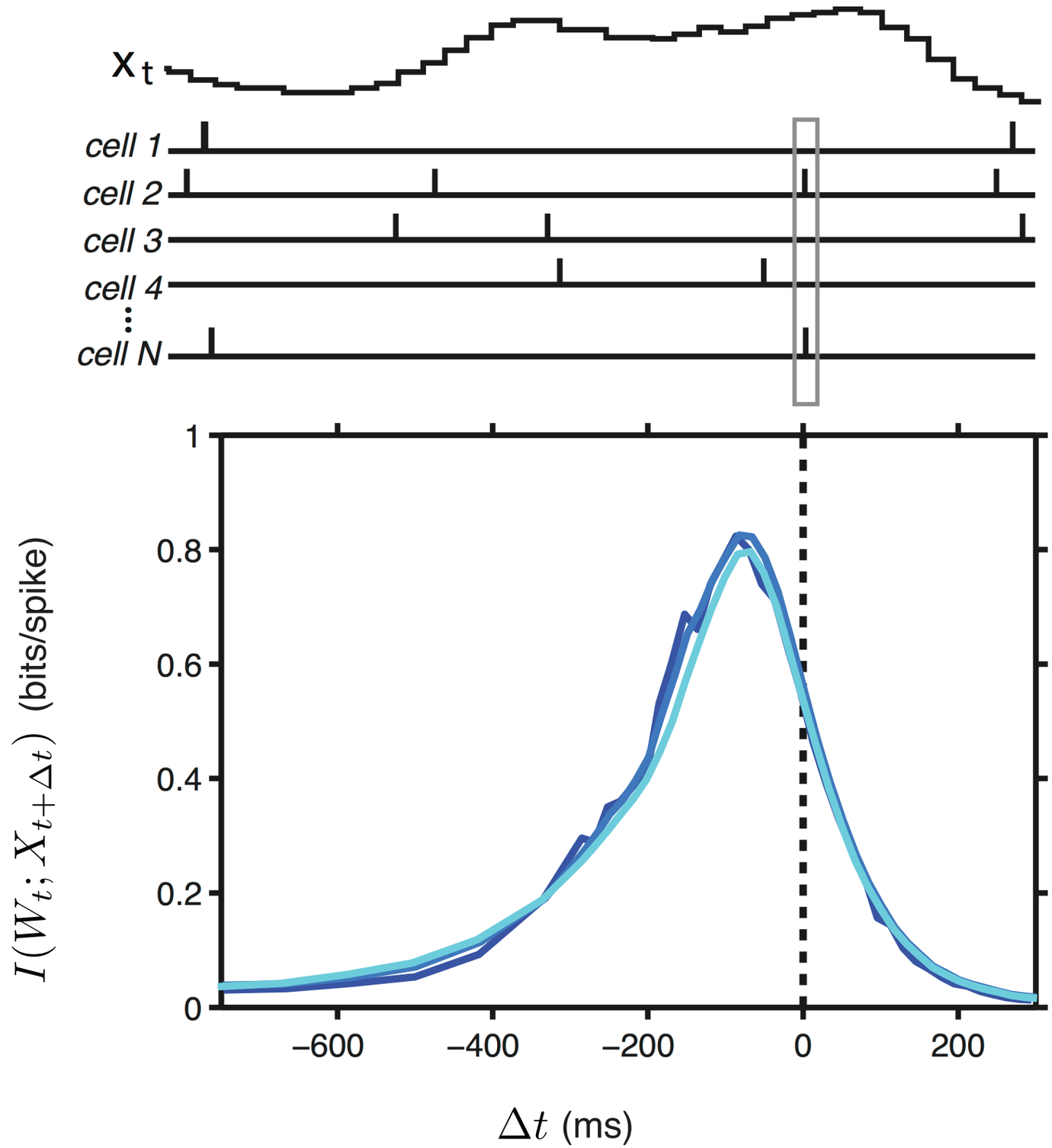
Palmer, Marre, Berry, Bialek PNAS (2015)

Retina populations carry info about the future:



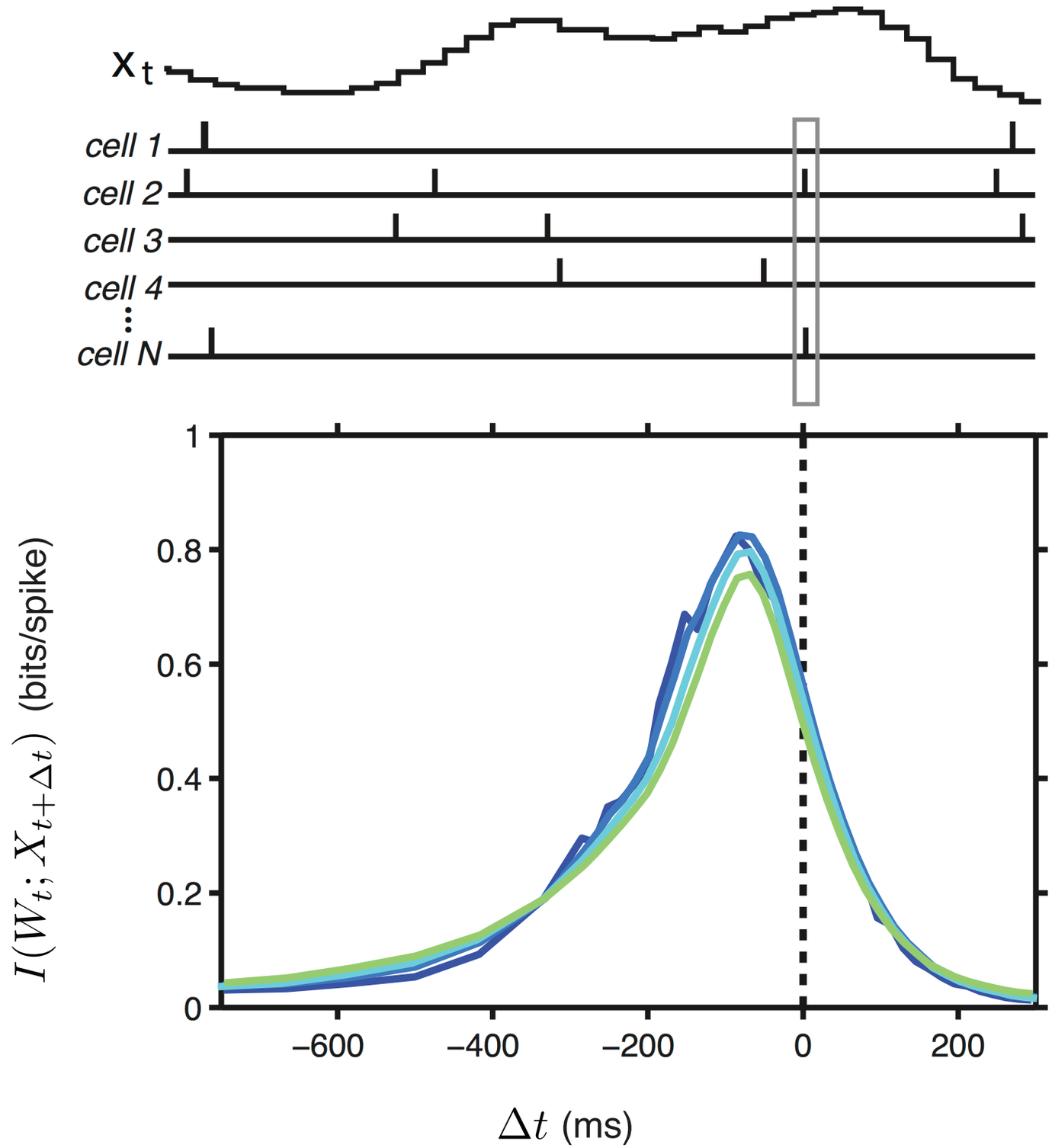
Palmer, Marre, Berry, Bialek PNAS (2015)

Retina populations carry info about the future:



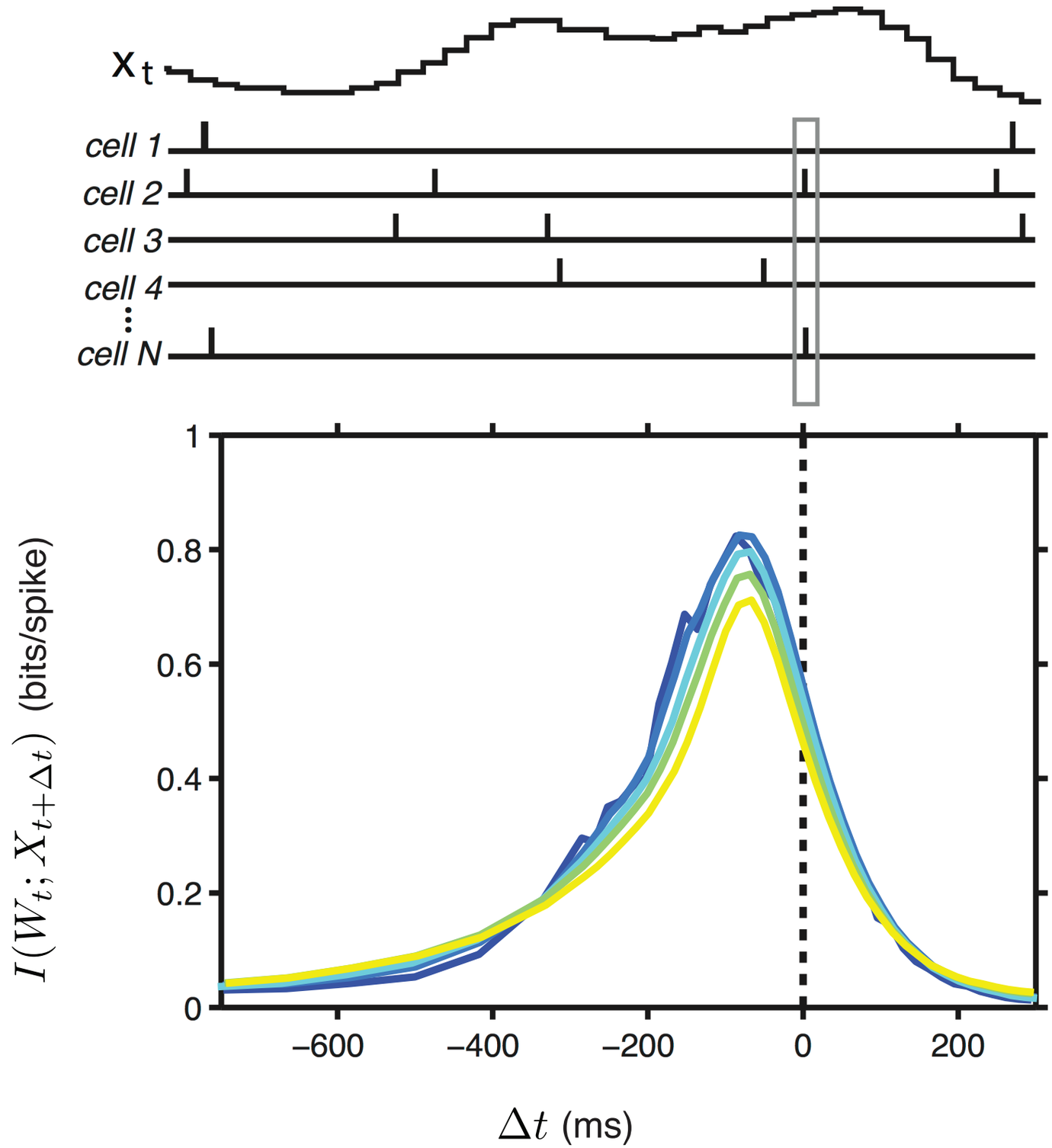
Palmer, Marre, Berry, Bialek PNAS (2015)

Retina populations carry info about the future:



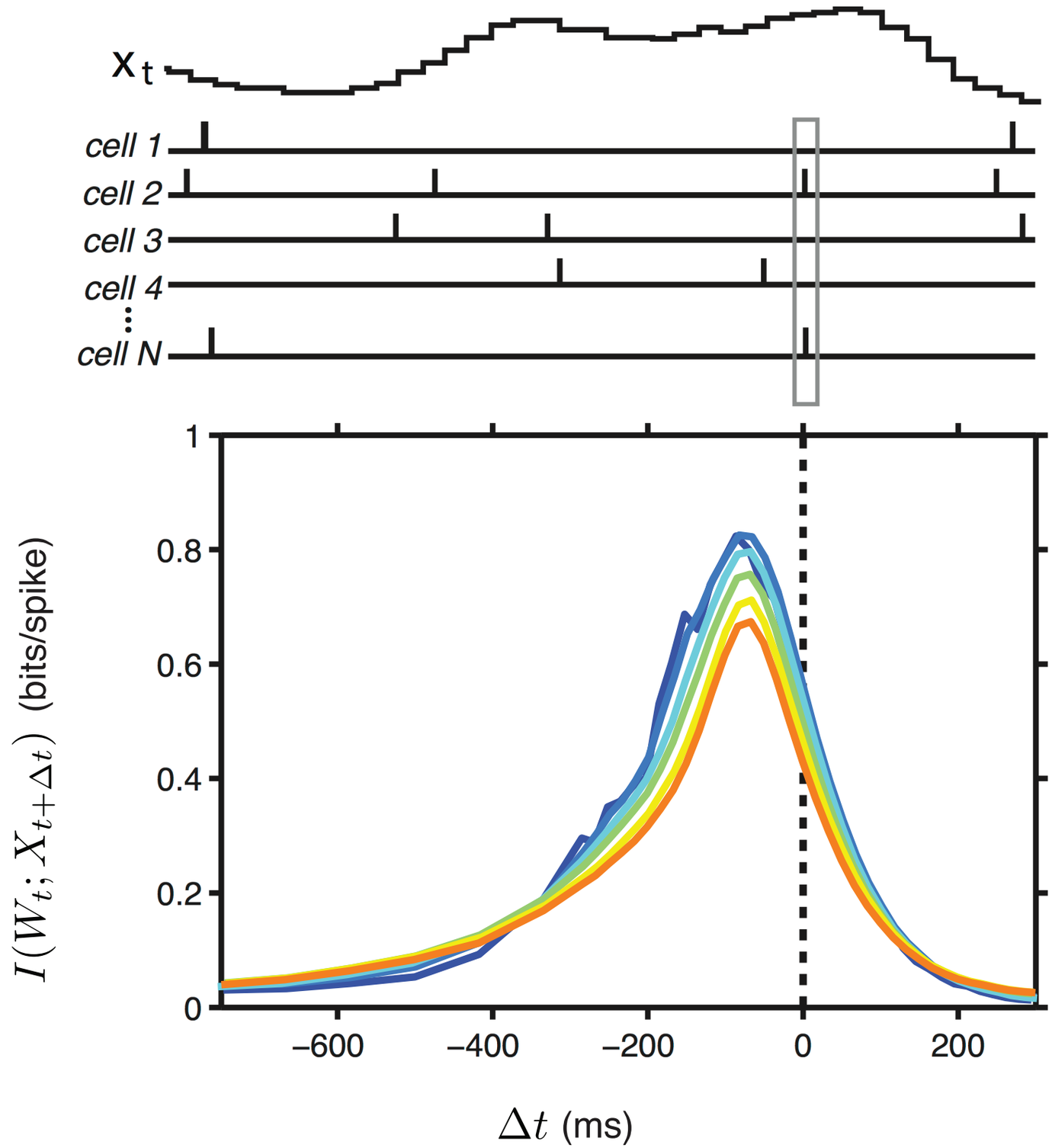
Palmer, Marre, Berry, Bialek PNAS (2015)

Retina populations carry info about the future:



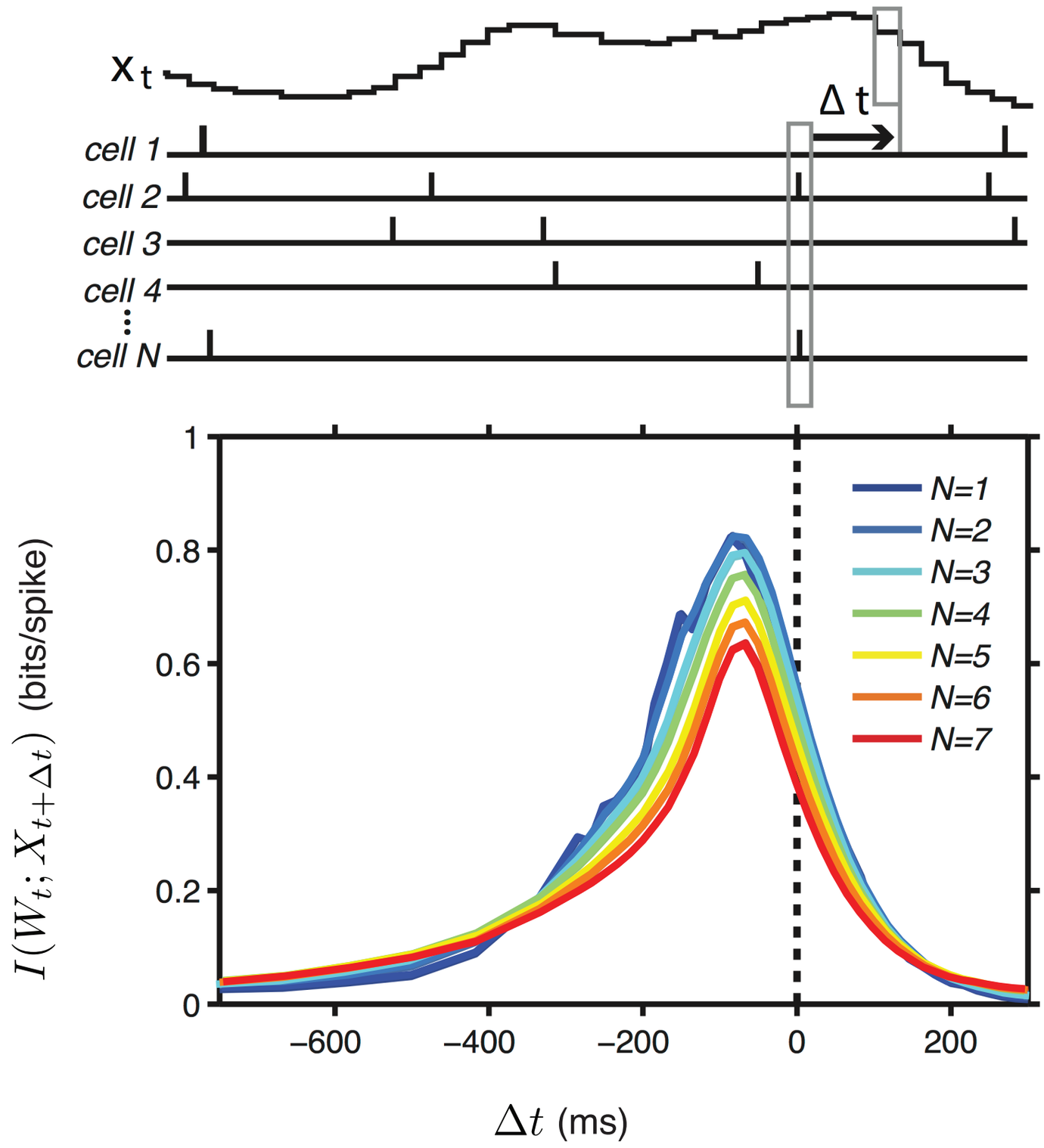
Palmer, Marre, Berry, Bialek PNAS (2015)

Retina populations carry info about the future:



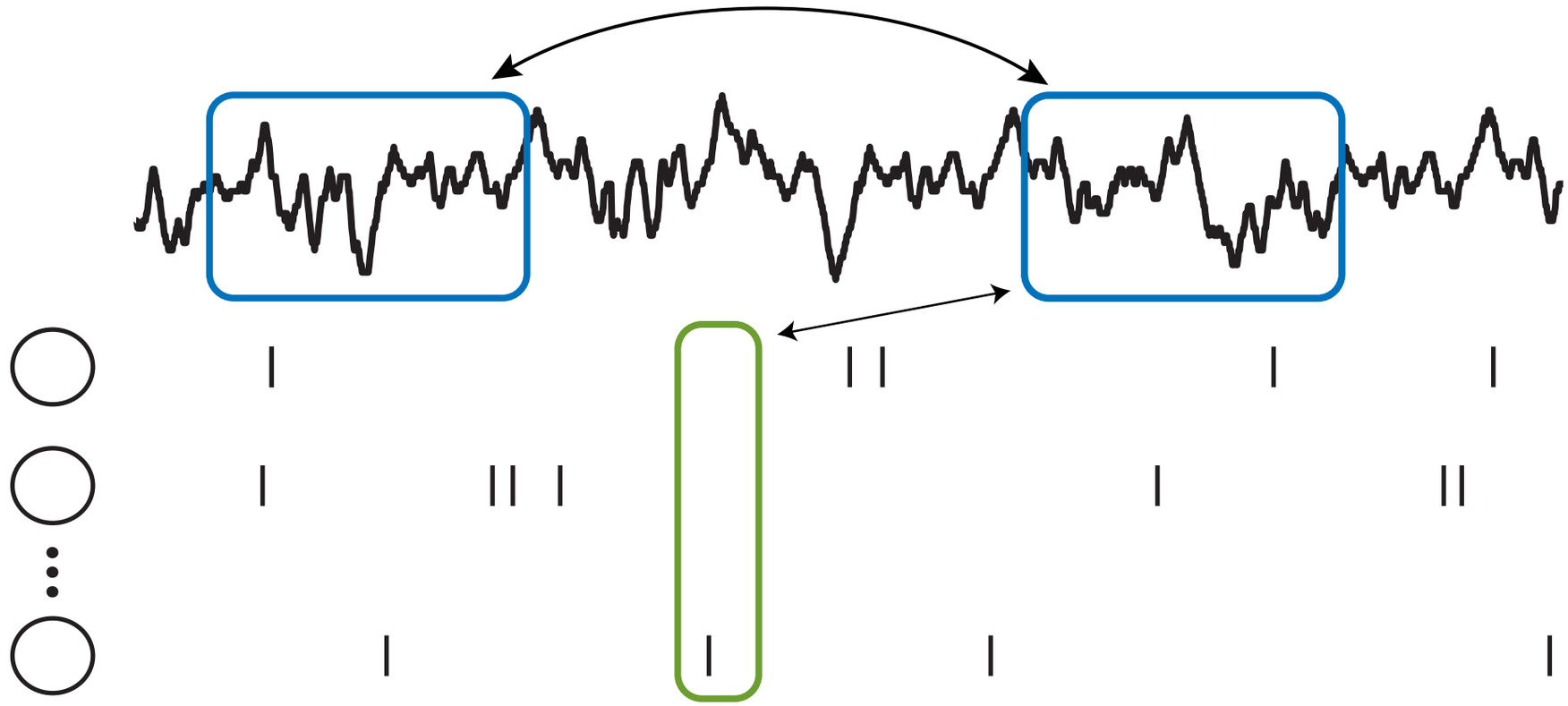
Palmer, Marre, Berry, Bialek PNAS (2015)

Retina populations carry info about the future:

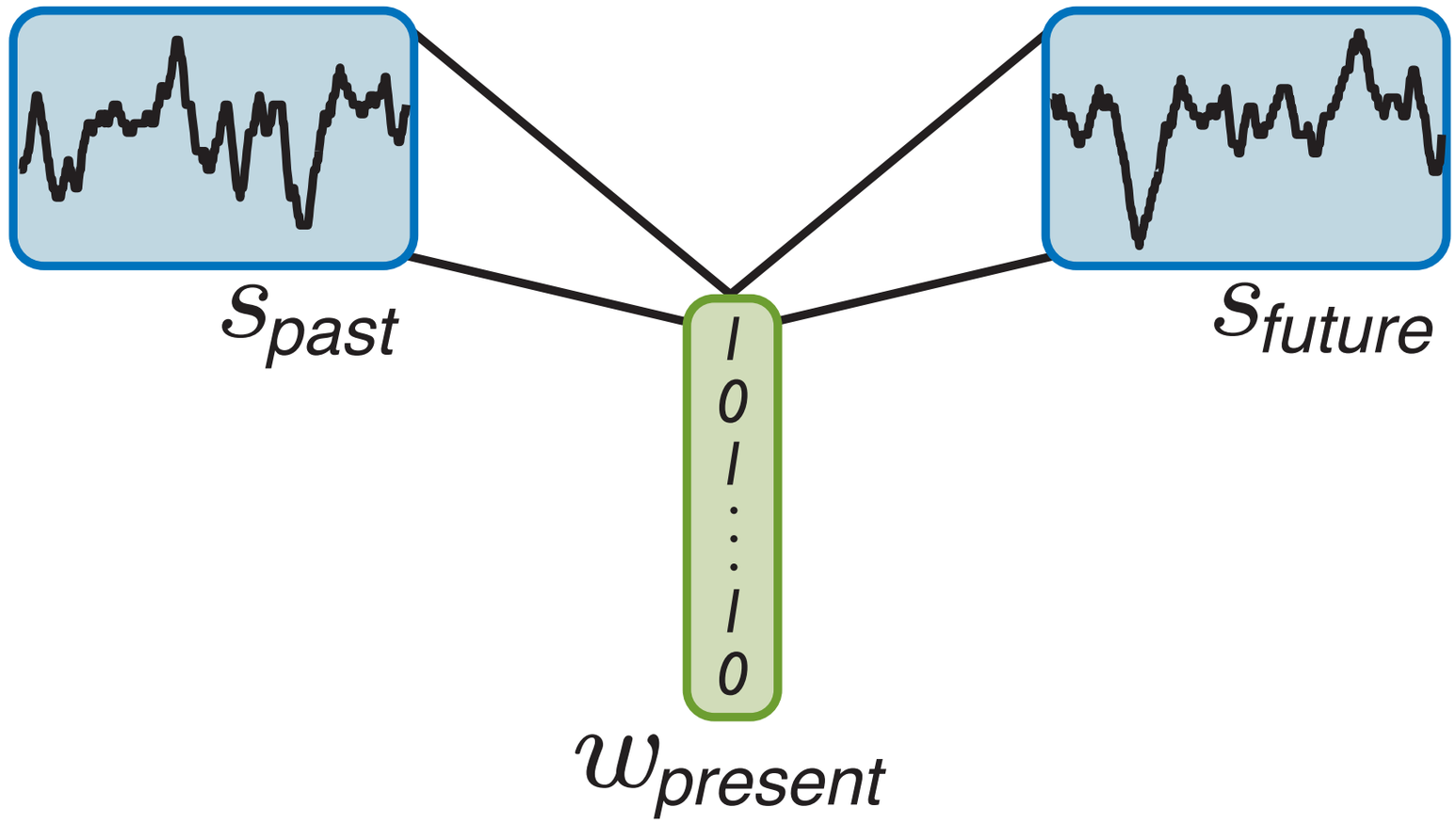


Palmer, Marre, Berry, Bialek PNAS (2015)

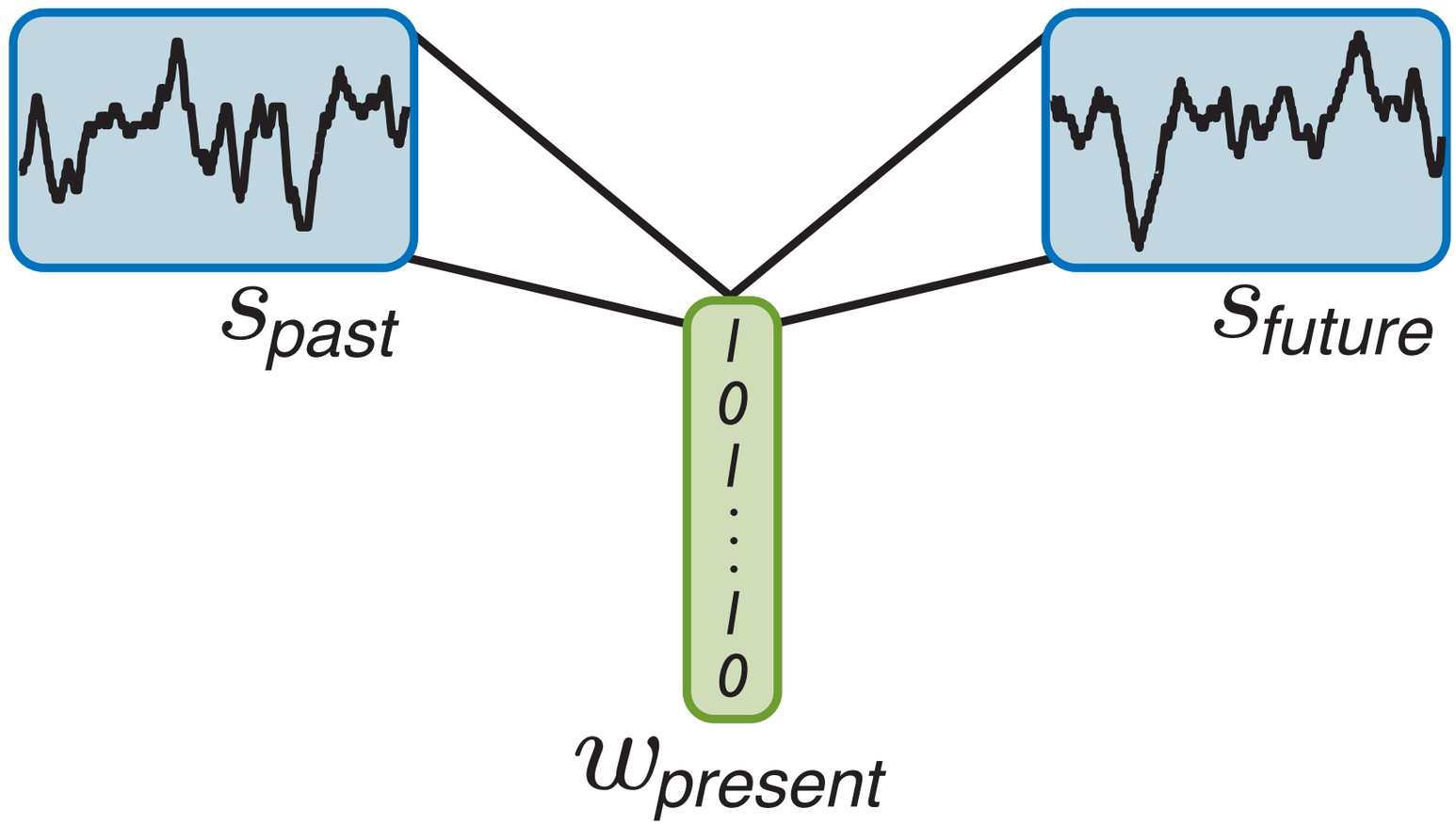
We compute the optimal code,
given the stimulus correlations:



The bottleneck problem:



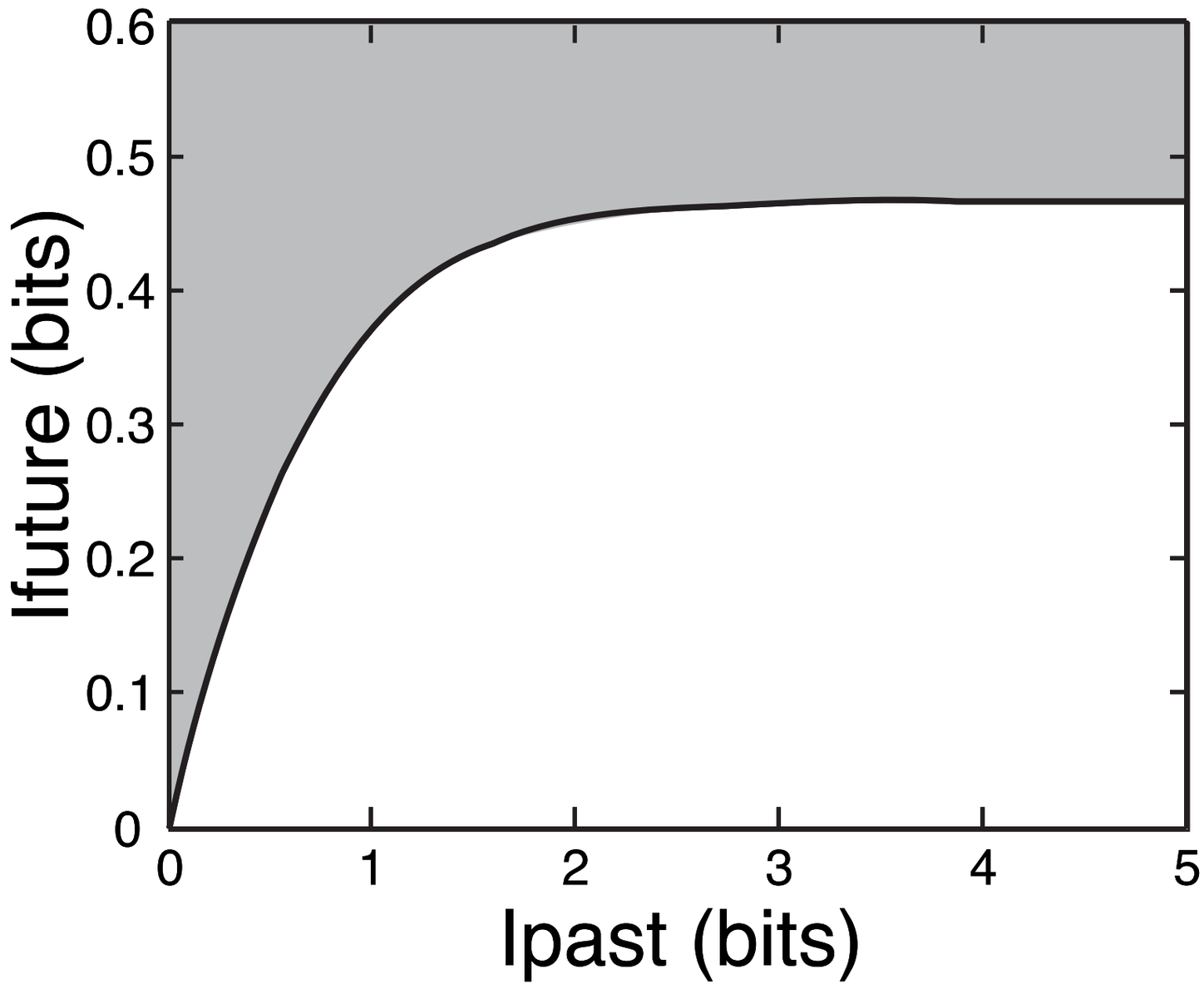
The bottleneck problem:



$$L = I_{past}(W_t; \vec{S}_{t-\Delta t}) - \beta I_{future}(W_t; \vec{S}_{t+\Delta t})$$

Optimal compression:

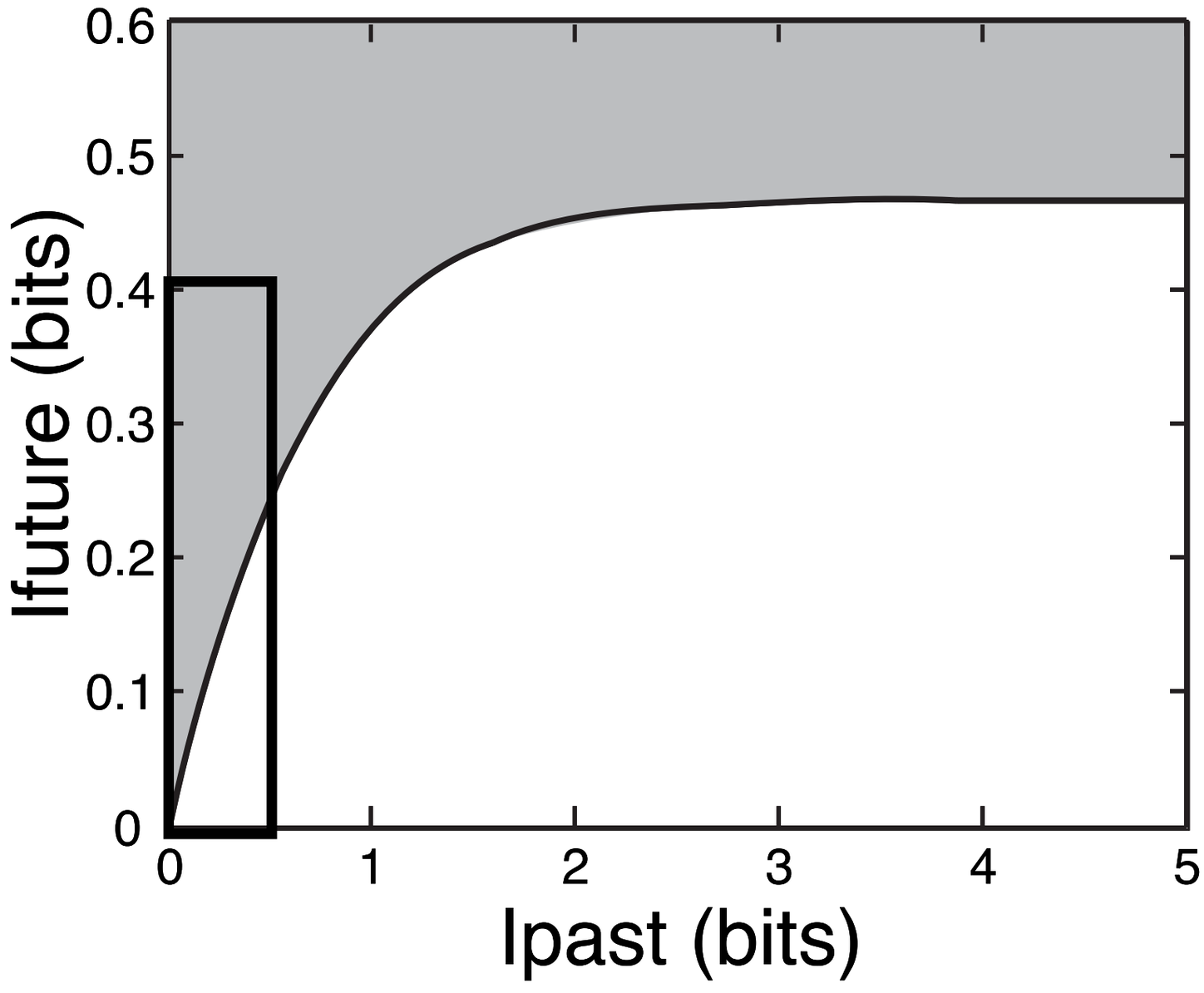
$$L = I_{past}(W_t; \vec{S}_{t-\Delta t}) - \beta I_{future}(W_t; \vec{S}_{t+\Delta t})$$



Tishby, Pereira, Bialek (1999)
Bialek, Nemenman, Tishby (2001)
Chechik, Globerson, Tishby, Weiss (2005)

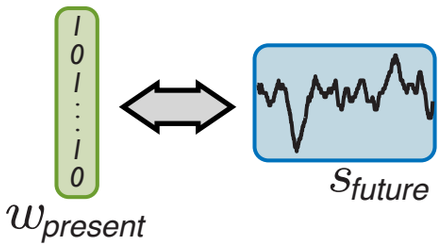
Optimal compression:

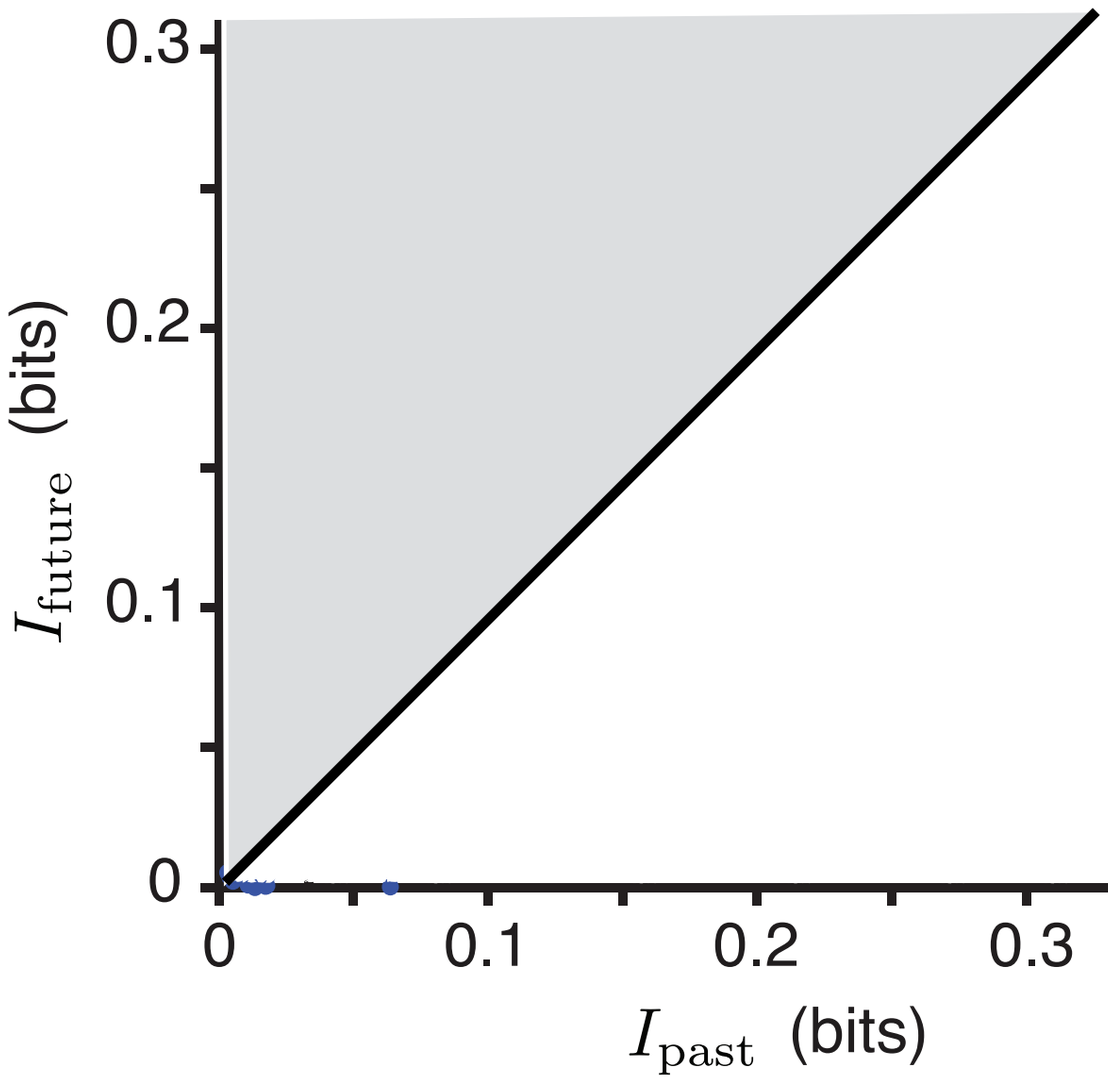
$$L = I_{past}(W_t; \vec{S}_{t-\Delta t}) - \beta I_{future}(W_t; \vec{S}_{t+\Delta t})$$

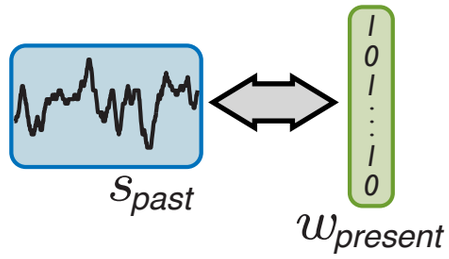


Tishby, Pereira, Bialek (1999)
Bialek, Nemenman, Tishby (2001)
Chechik, Globerson, Tishby, Weiss (2005)

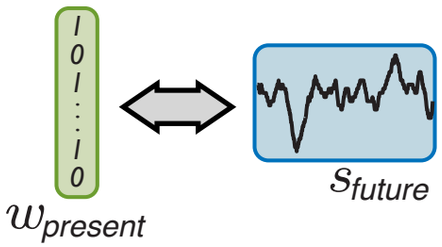
Spiking patterns sit close to the bound:

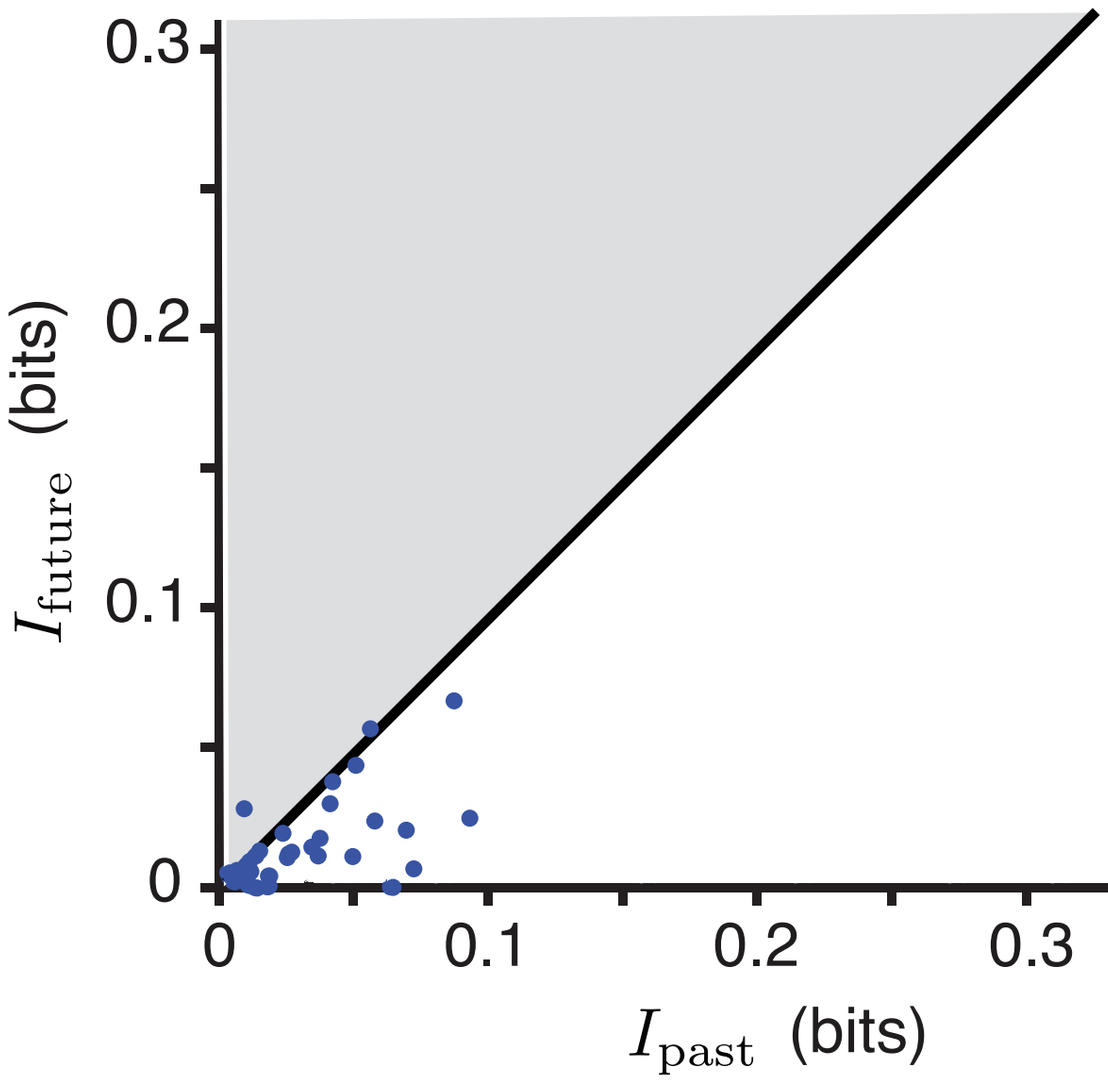


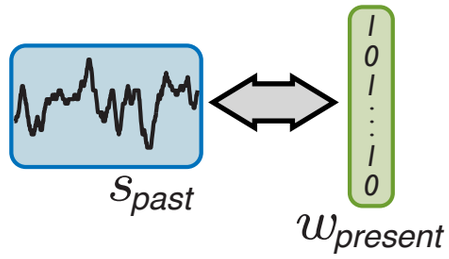




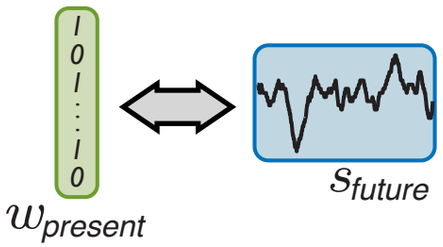
Spiking patterns sit close to the bound:

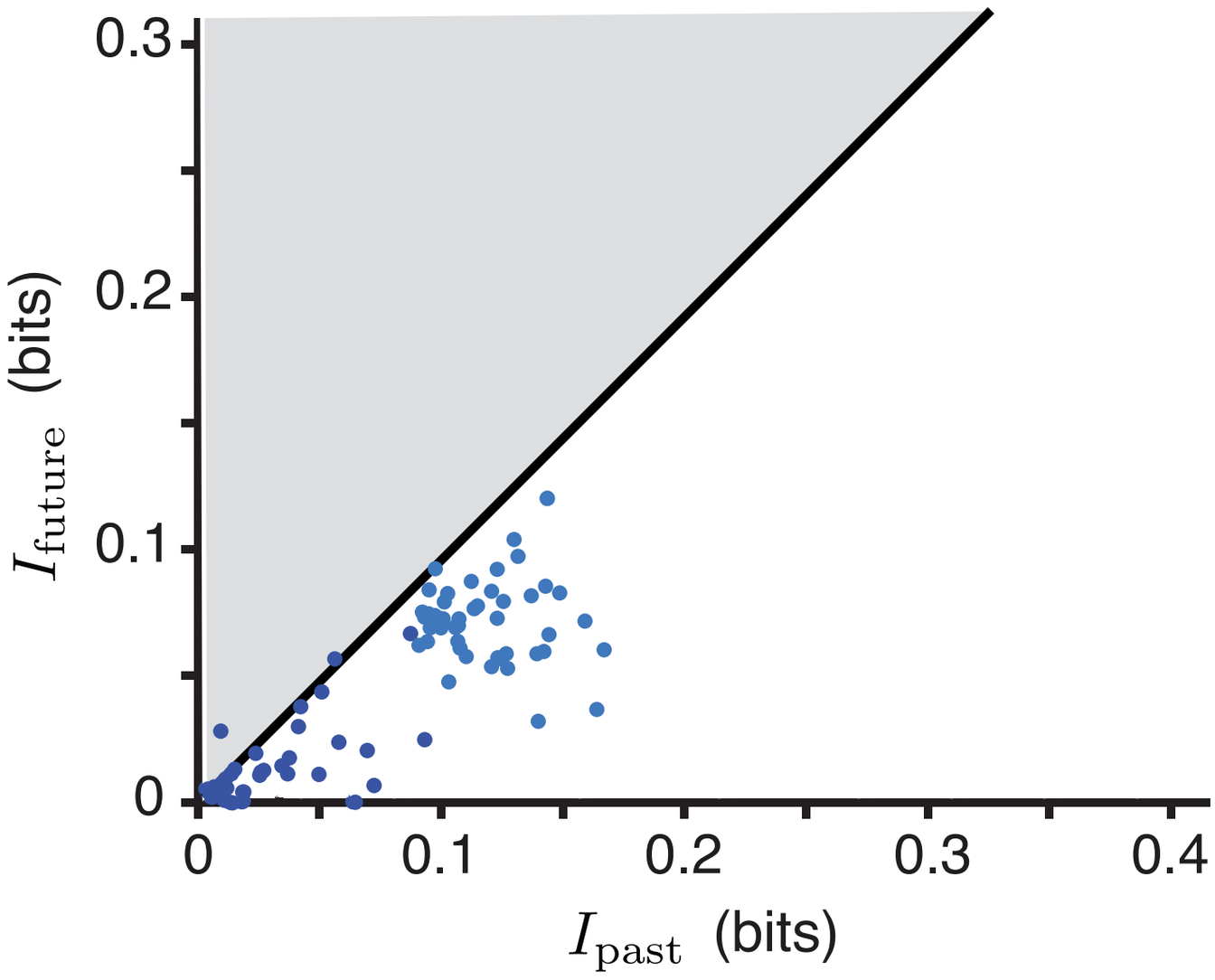


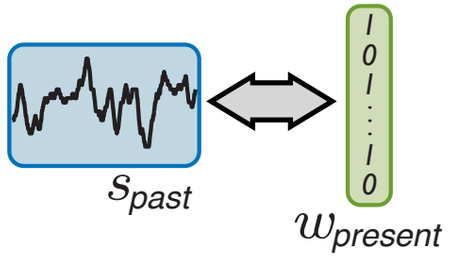




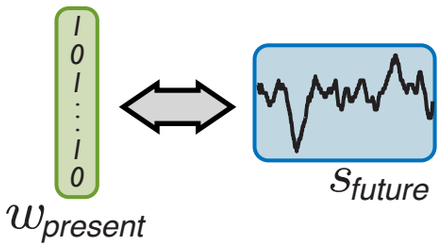
Spiking patterns sit close to the bound:

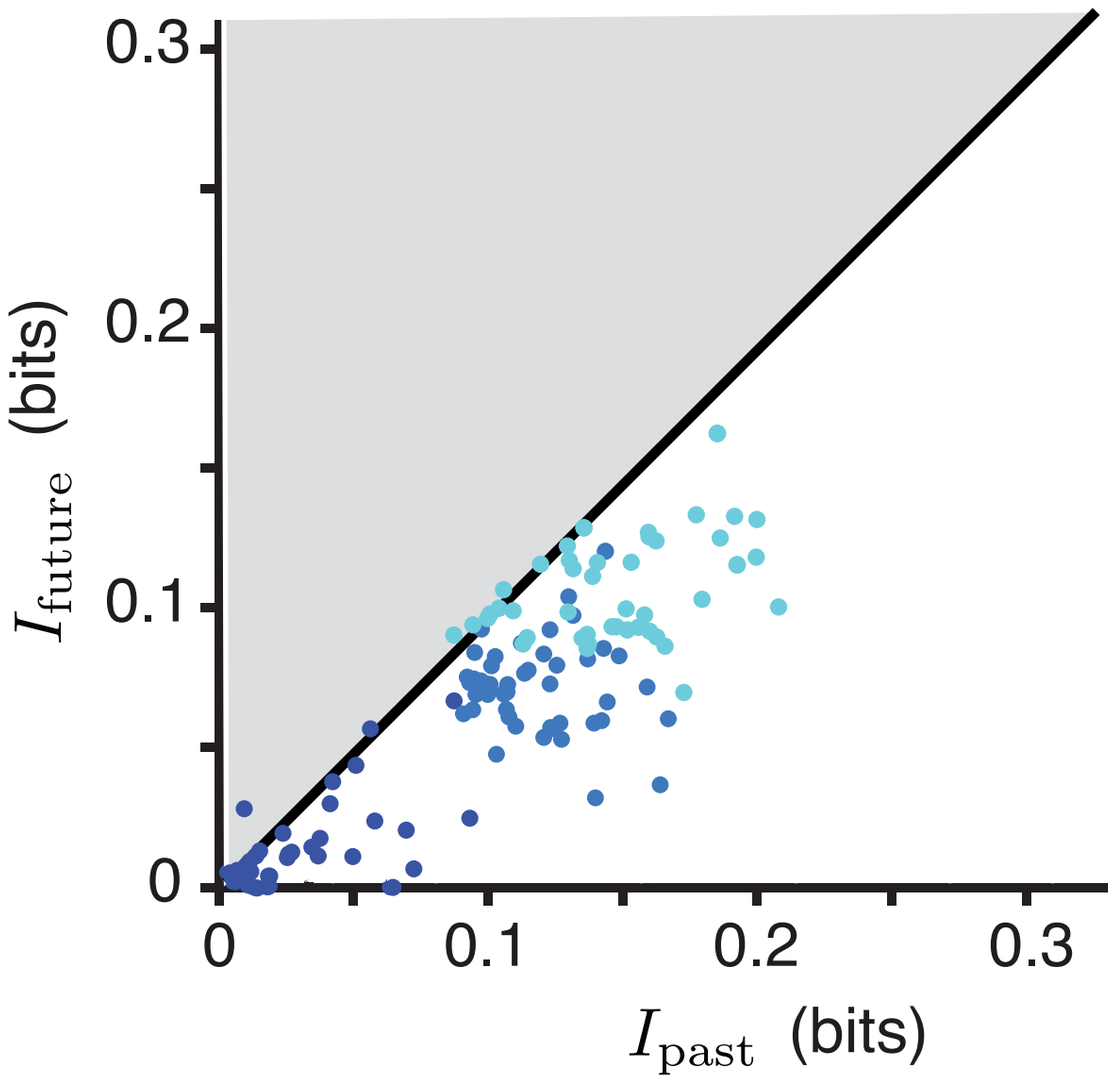


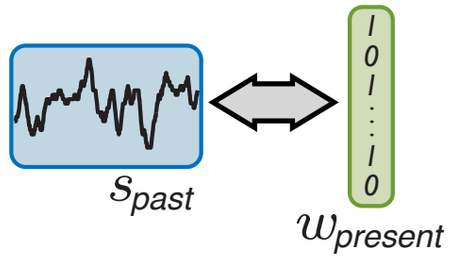




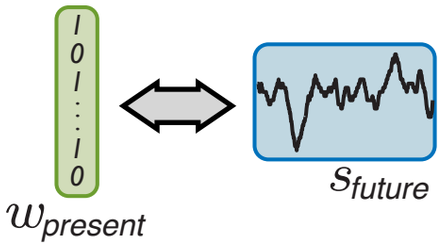
Spiking patterns sit close to the bound:

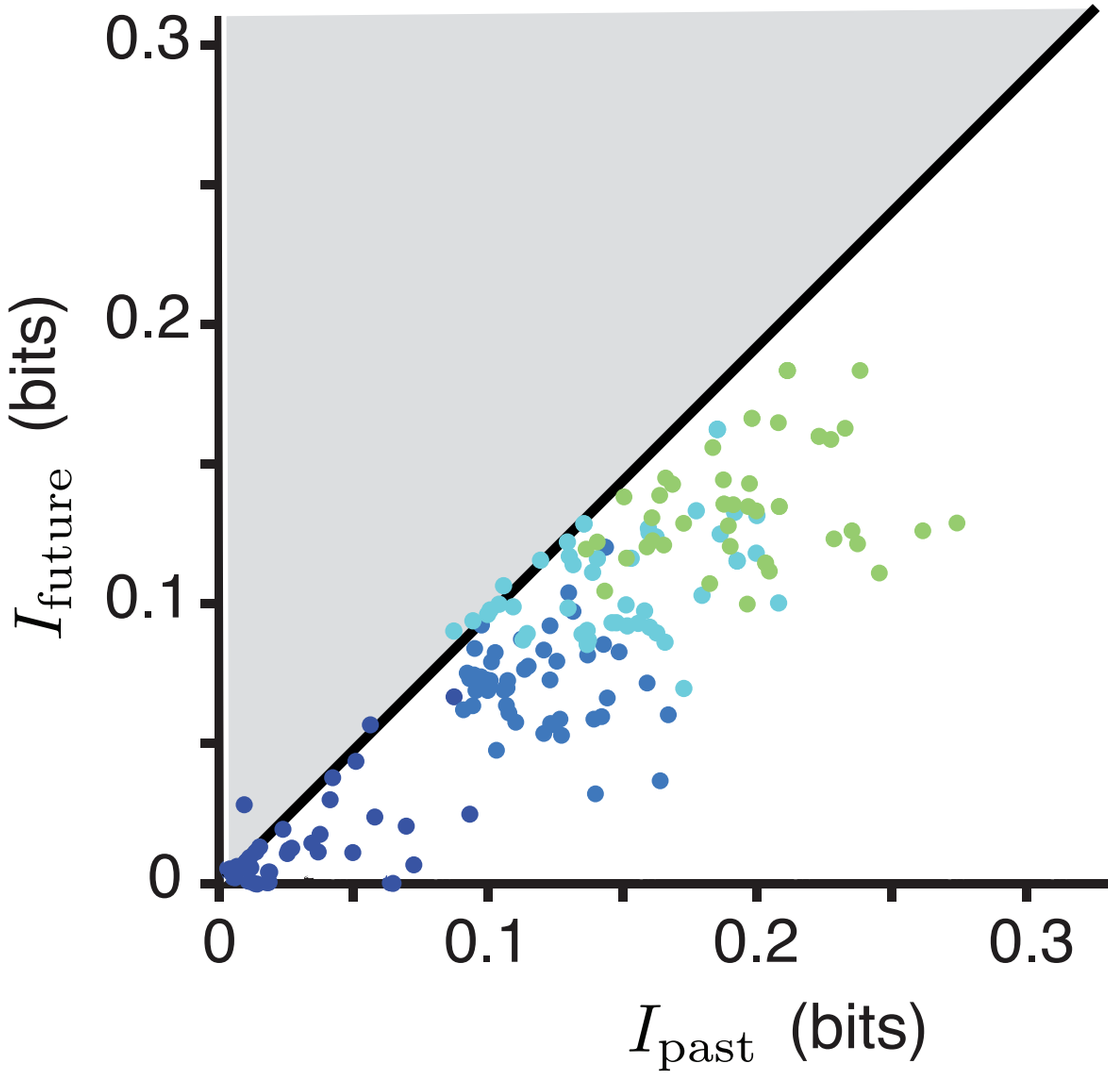


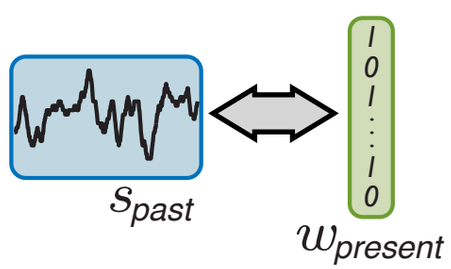




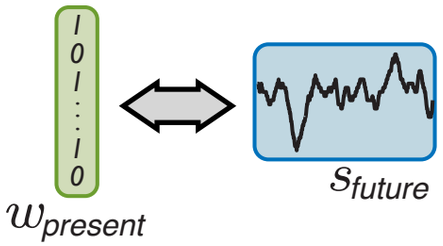
Spiking patterns sit close to the bound:

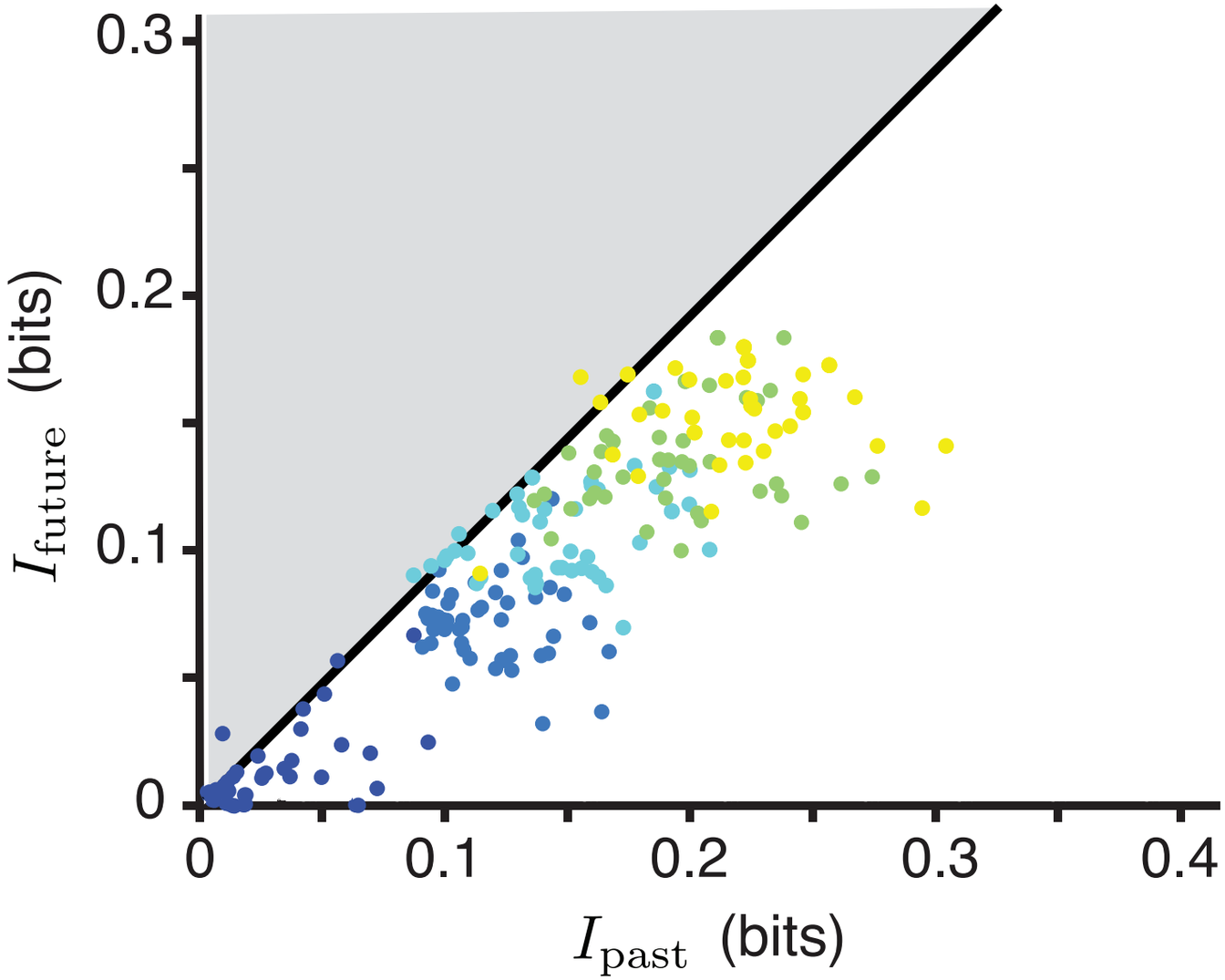


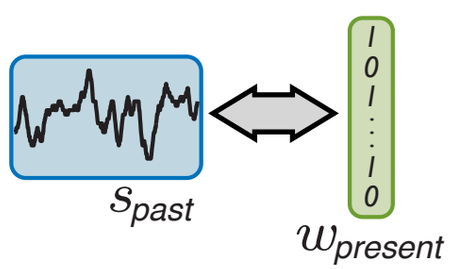




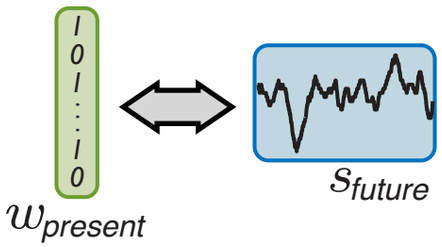
Spiking patterns sit close to the bound:

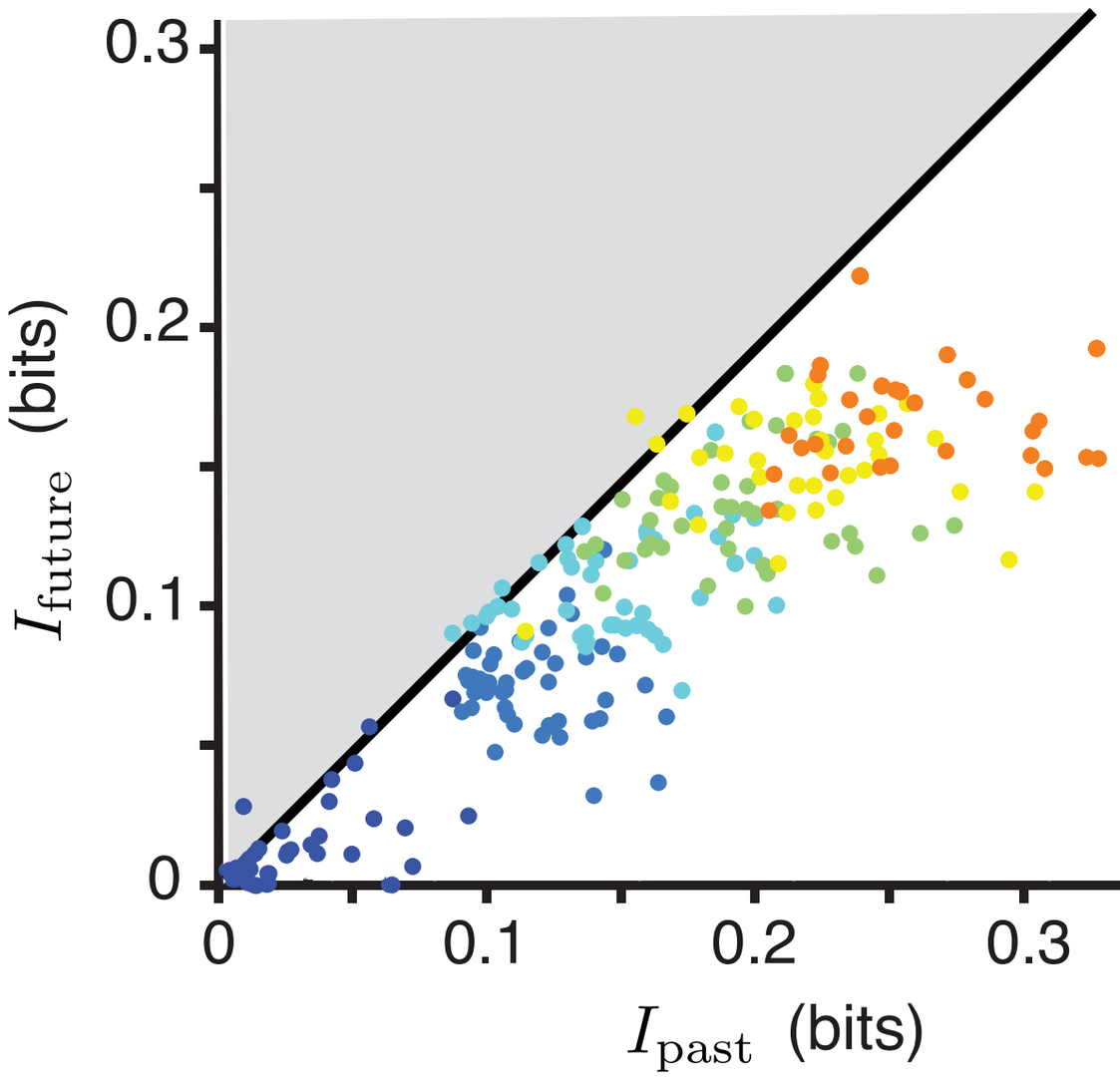


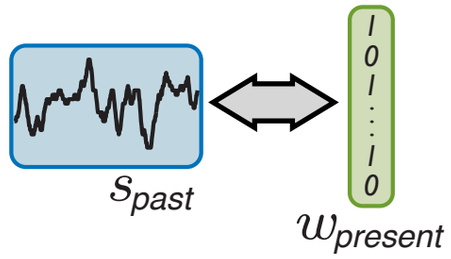




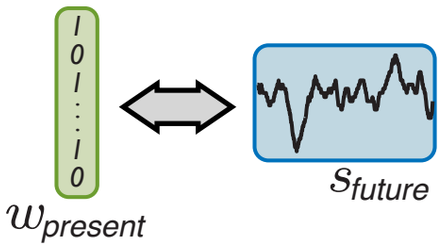
Spiking patterns sit close to the bound:

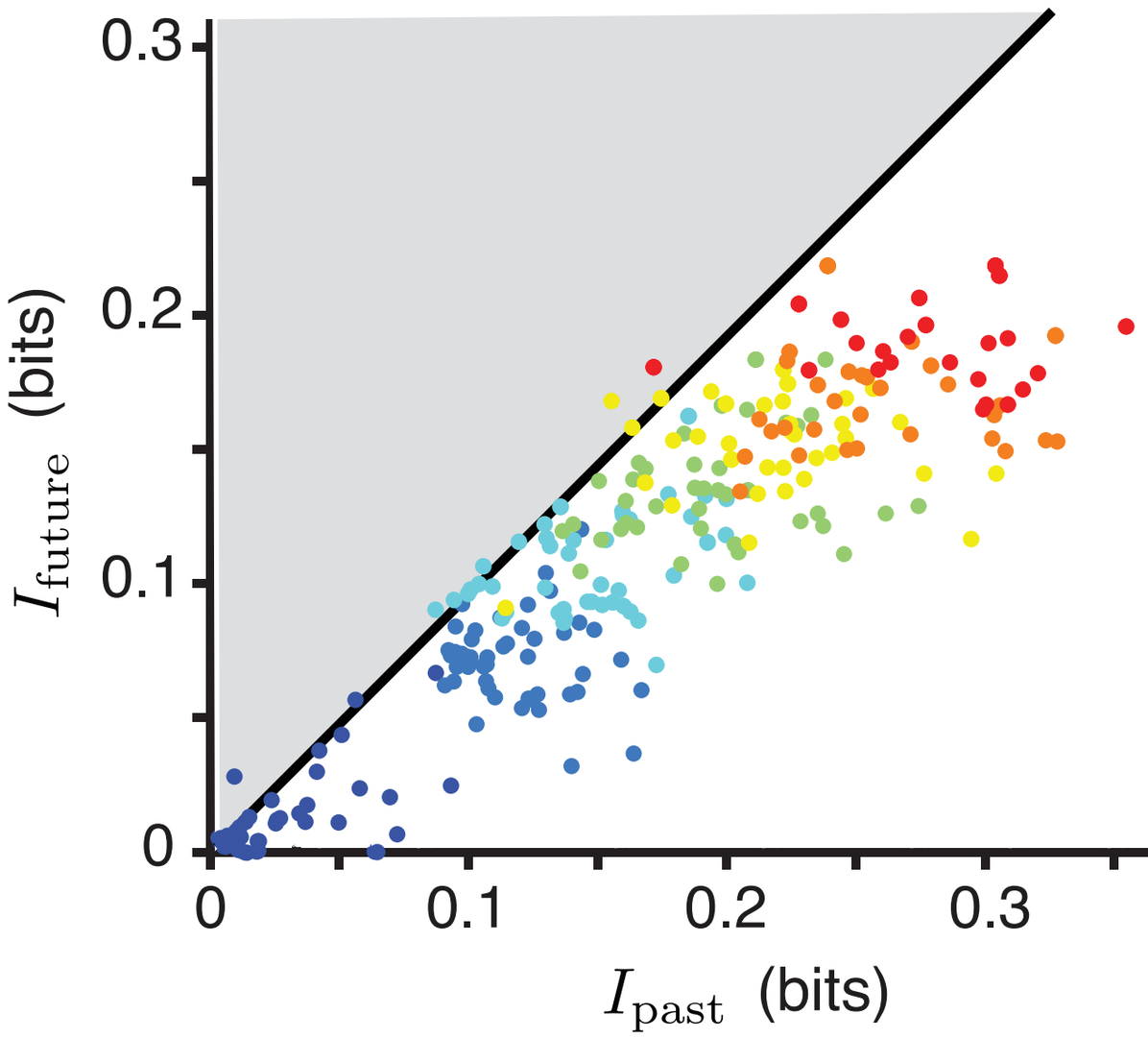


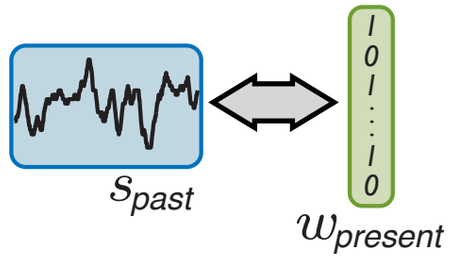




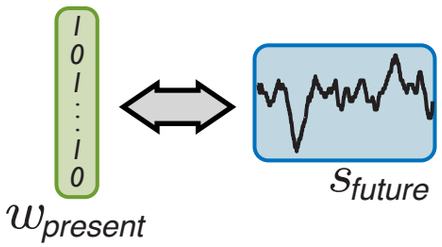
Spiking patterns sit close to the bound:

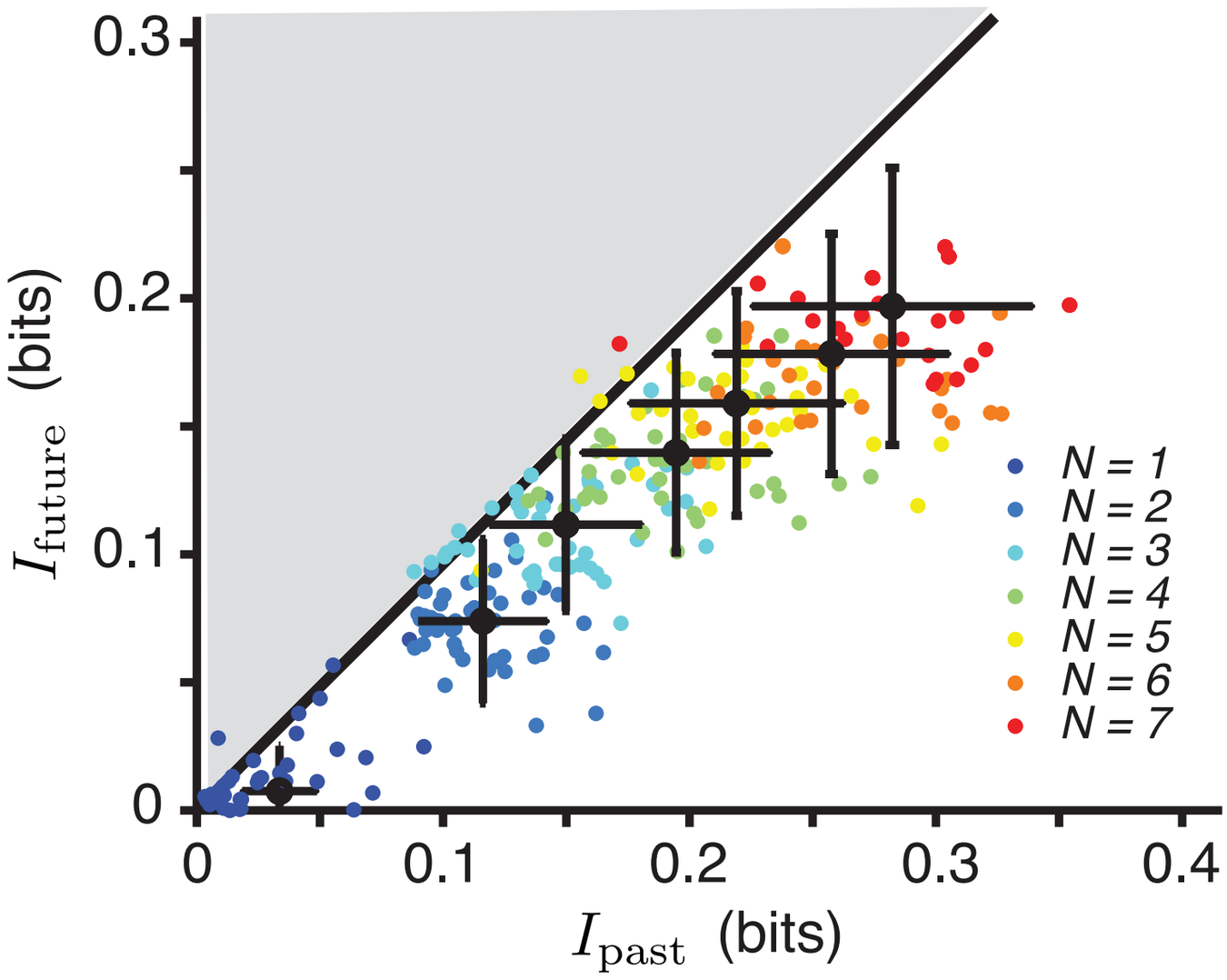


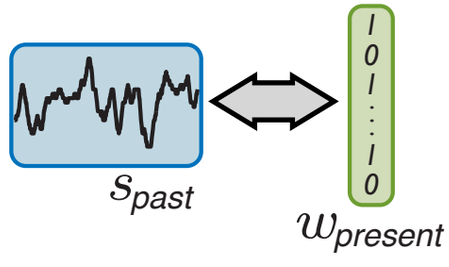




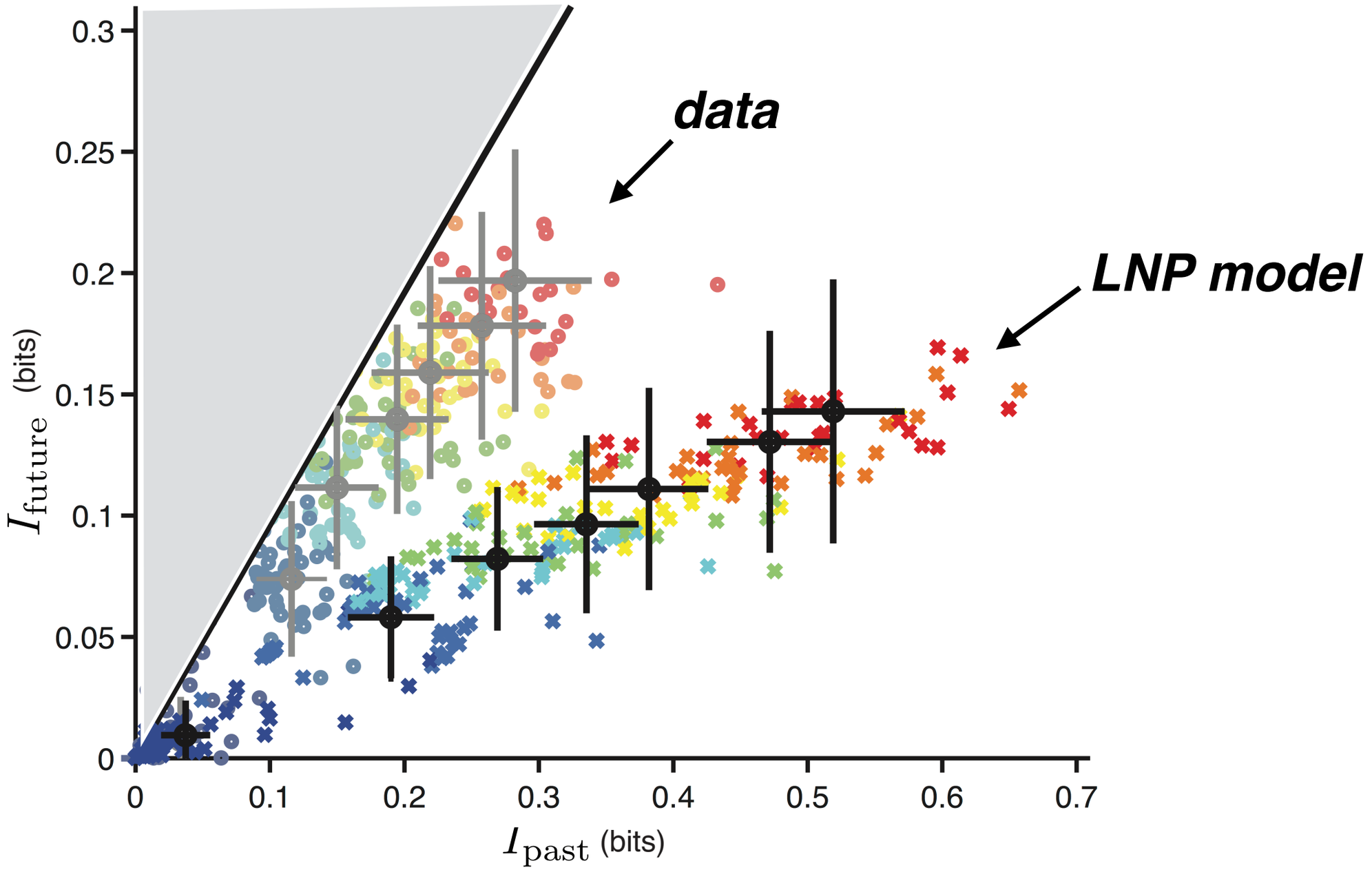
Retinal populations saturate the bound:







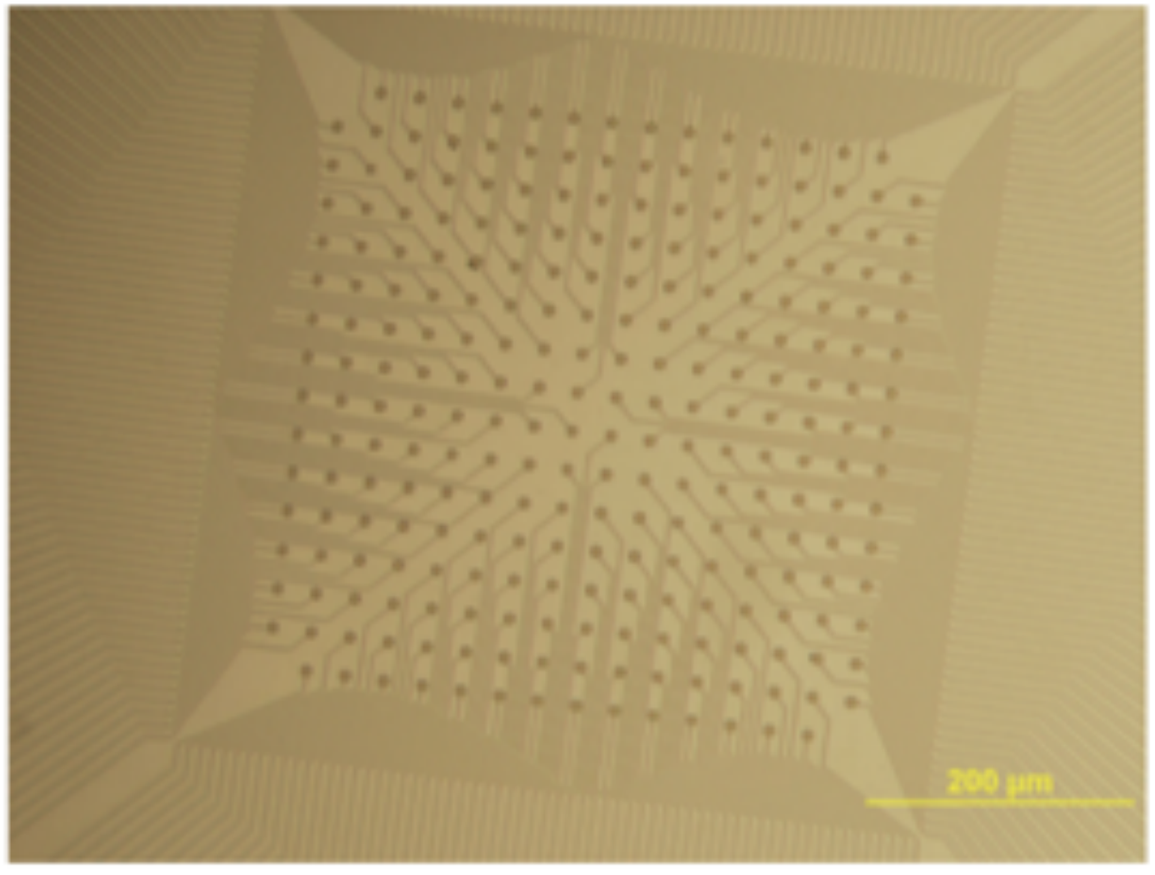
This doesn't work with just simple linear filters:



Palmer et al. PNAS (2015)

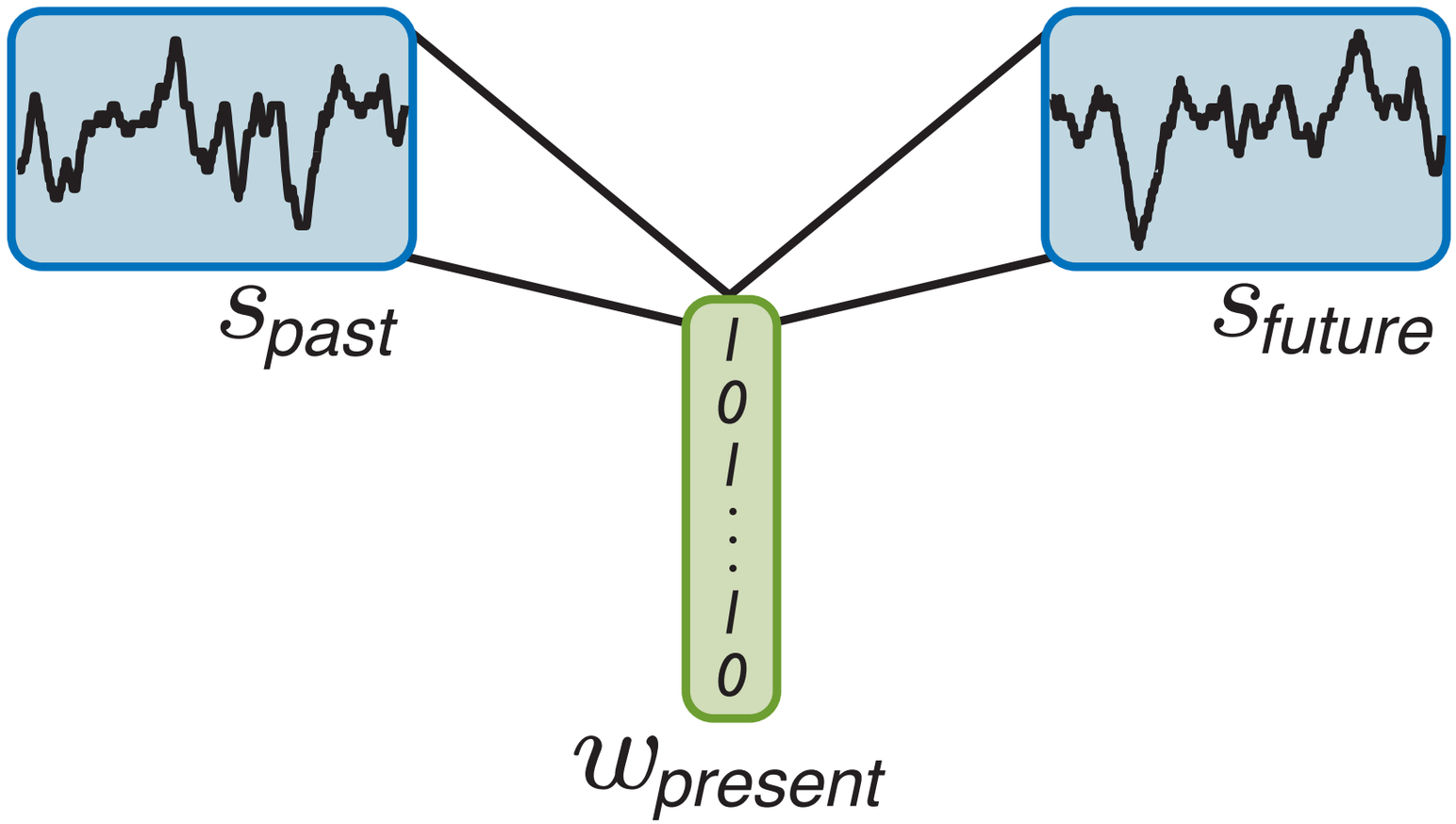
Recording from the rat retina:

wired.co.uk



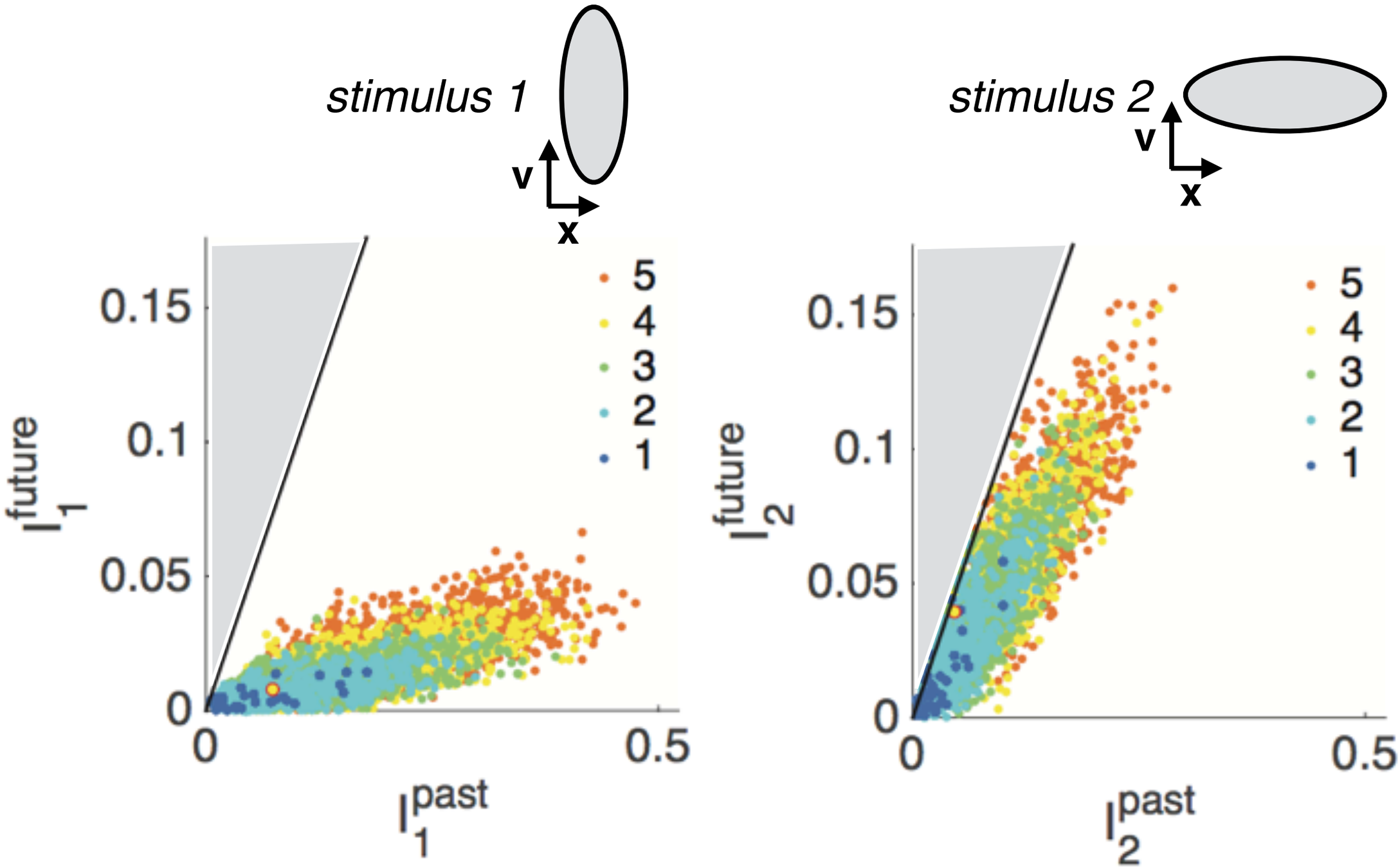
data from the Marre Lab, Institute of Vision, Paris VI University, France

Compute the bound on the predictive info:



$$L = I_{past}(W_t; \vec{S}_{t-\Delta t}) - \beta I_{future}(W_t; \vec{S}_{t+\Delta t})$$

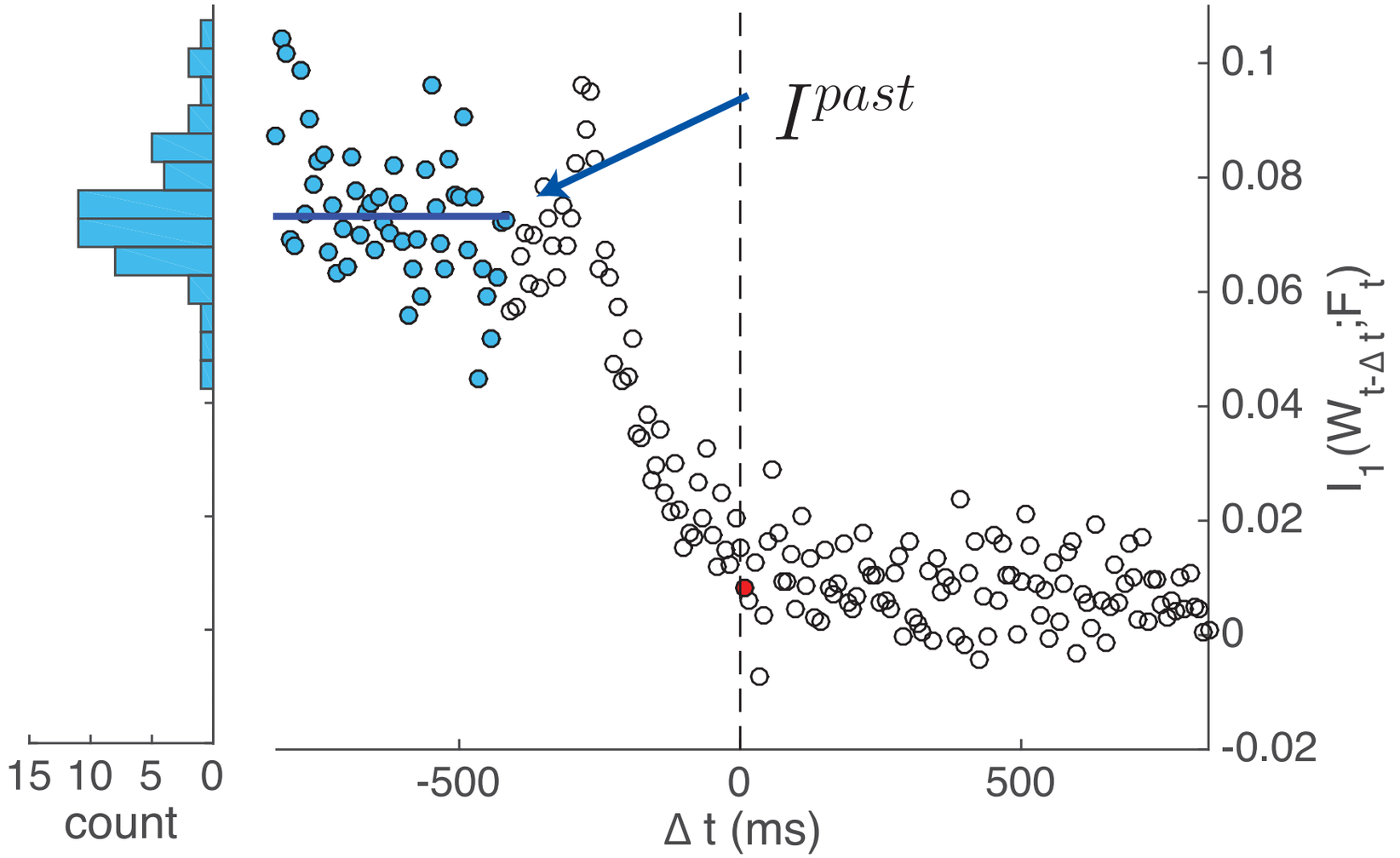
Performance relative to the bound:

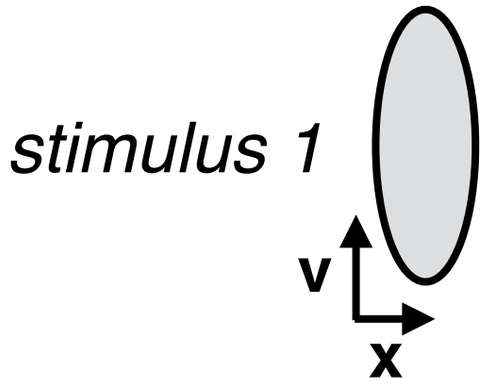


An example group of four cells:

$$I\left(W_{t-\Delta t}; F_t\right) = H\left(W_{t-\Delta t}\right) + H\left(F_t\right) - H\left(W_{t-\Delta t}, F_t\right)$$

N = 4





An example group of four cells: