Sparse Coding as a Generative Model

$$S = \sum_{i=1}^{M} a_i \phi_i + \epsilon$$

image vector

neural activity (sparse)

feature vector

other stuff

Find activations by descending E

$$E = \overbrace{\frac{1}{2}||S - \hat{S}||_2^2}^{\text{Preserve Information}} + \overbrace{\lambda \sum_{i=1}^{M} C(a_i)}^{\text{Limit Activations}}$$

Coefficients via gradient descent

$$-\frac{\partial E}{\partial a_k} = \sum_{i=1}^{N}\left[S_i\Phi_{i,k} - \sum_{j\neq k}^{M}\Phi_{i,k}\Phi_{i,j}a_j\right] - a_k + \lambda\frac{\partial C(a_k)}{\partial a_k}$$

$$b_k = \phi_k^T S = \sum_{i=1}^{N} S_i\Phi_{i,k}$$

Driving input (excitation)

$$G = \Phi^T\Phi \qquad G_{i,j} = \sum_{n=1}^{N}\Phi_{n,i}\Phi_{n,j}$$

Lateral inhibition

$$f_\lambda(a_k) = a_k + \lambda\frac{\partial C(a_k)}{\partial a_k}$$

Self Inhibition

$$\boxed{-\frac{\partial E}{\partial a_k} = b_k - \sum_{j\neq k}^{M} G_{k,j}a_j - f_\lambda(a_k)}$$

Network dynamics for descending E

Internal state (membrane potential): u_k
Broadcasted activity: a_k

The membrane potential follows the energy gradient:

$$-\frac{\partial E}{\partial a_k} = b_k - \sum_{j \neq k}^{M} G_{k,j} a_j - f_\lambda(a_k)$$

$$\dot{u}_k \propto -\frac{\partial E}{\partial a_k}$$

membrane leak term

$$\dot{u}_k = \frac{1}{\tau}\left[b_k - \sum_{m \neq k}^{M} G_{k,m} a_m - f_\lambda(a_k)\right]$$

Leaky integrator model

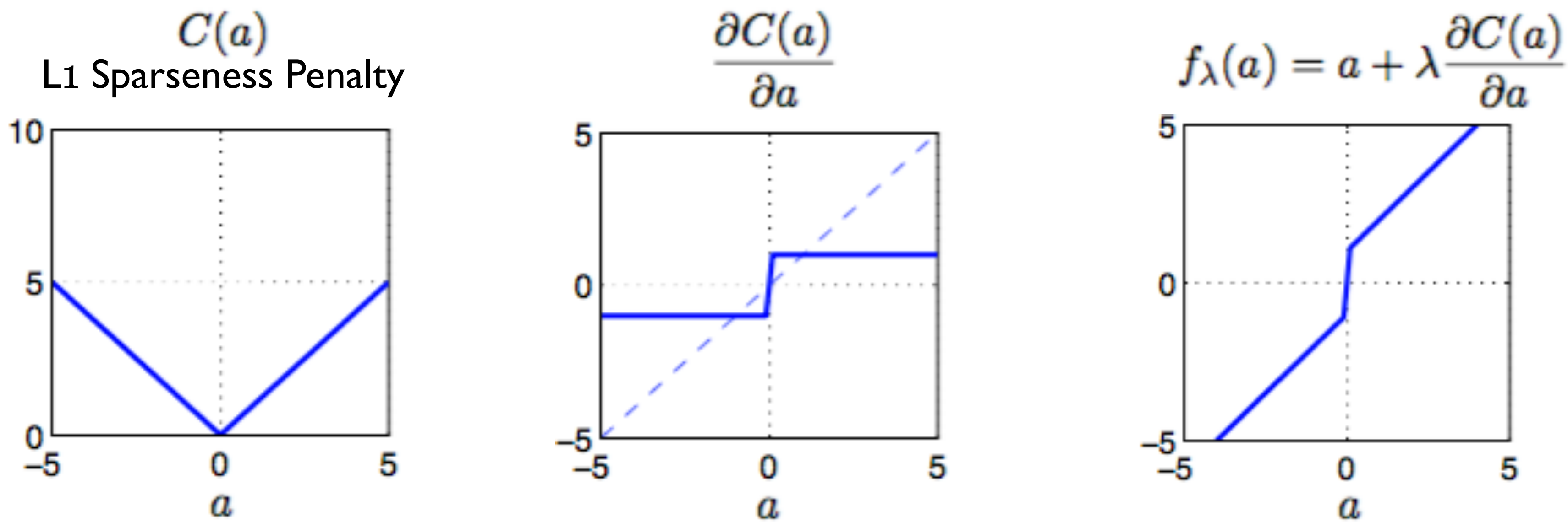
$$-\frac{\partial E}{\partial a_k} = b_k - \sum_{j \neq k}^{M} G_{k,j} a_j - f_\lambda(a_k)$$

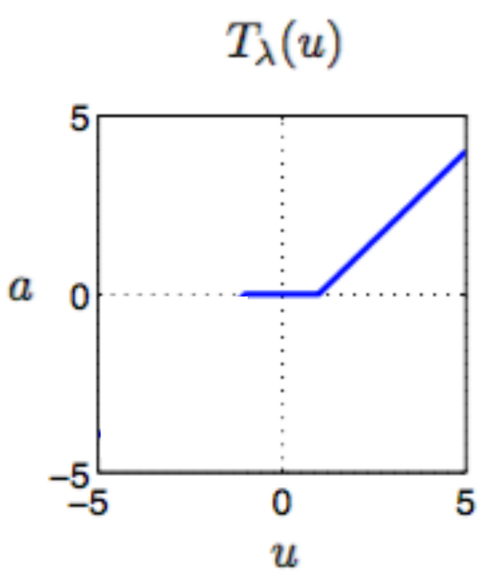
$$f_\lambda(a_k) = a_k + \lambda \frac{\partial C(a_k)}{\partial a_k}$$

$$u_k = f_\lambda(a_k) \qquad a_k = f_\lambda^{-1}(u_k) := T_\lambda(u_k)$$

$$\boxed{\tau \dot{u}_k = -\frac{\partial E}{\partial a_k} = \overbrace{b_k}^{\text{Driving excitation}} - \overbrace{\sum_{m \neq k}^{M} G_{k,m} a_m}^{\text{Lateral Inhibition}} - \overbrace{u_k}^{\text{Leak}}}$$

Determining the Thresholding Function





Other transfer functions can also be used!

As long as the thresholding function is monotonically increasing.

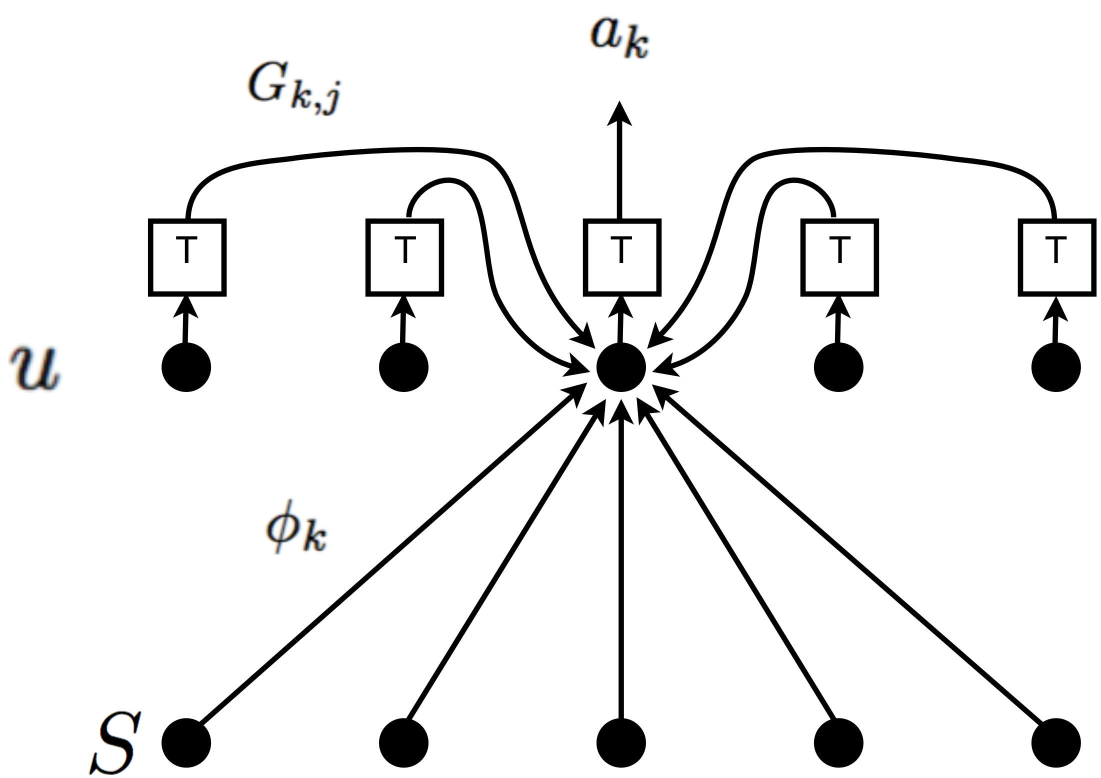
Neuron output is thresholded membrane potential

$$\dot{u}_k = \frac{1}{\tau}\left[\overbrace{b_k}^{\text{Driving excitation}} - \overbrace{\sum_{m\neq k}^{M} G_{k,m}a_m}^{\text{Lateral Inhibition}} - \overbrace{u_k}^{\text{Leak}}\right]$$

$$a_k = T_\lambda(u_k)$$

A network of neurons following these dynamics will always lower E, or else leave it unchanged, as long as u is a monotonically increasing function of a .

Network implementation of LCA dynamics



$$\dot{u}_k = \frac{1}{\tau}\left[\overbrace{b_k}^{\text{Driving excitation}} - \overbrace{\sum_{m\neq k}^{M} G_{k,m}a_m}^{\text{Lateral Inhibition}} - \overbrace{u_k}^{\text{Leak}}\right] \qquad a_k = T_\lambda(u_k)$$

CJ Rozell et al. "Sparse Coding via Thresholding and Local Competition in Neural Circuits"
Neural Computation. 2008

Adding Inhibitory Interneurons

$$\dot{u}_k = \frac{1}{\tau}\left[\overbrace{b_k}^{\text{Driving excitation}} - \overbrace{\sum_{m\neq k}^{M} G_{k,m} a_m}^{\text{Lateral Inhibition}} - \overbrace{u_k}^{\text{Leak}}\right]$$

"G" matrix captures all of the recurrent influence

$$G = U\Sigma V^T$$

G can be decomposed by matrix factorization

Inhibitory > Excitatory Connections

Interneuron Gains (diagonal)

Excitatory > Inhibitory Connections

Simple solution does not provide biologically realistic results…

M Zhu, CJ Rozell. "Modeling Inhibitory Interneurons in Efficient Sensory Coding Models." *PLoS Comp Bio. 2015*

Alternate Solution

Decompose G matrix into low rank + sparse matrices

$$L, S = \arg\min_{L,S} \|L\|_* + \|\Lambda S\|_1, \text{ subject to } G = L + S,$$

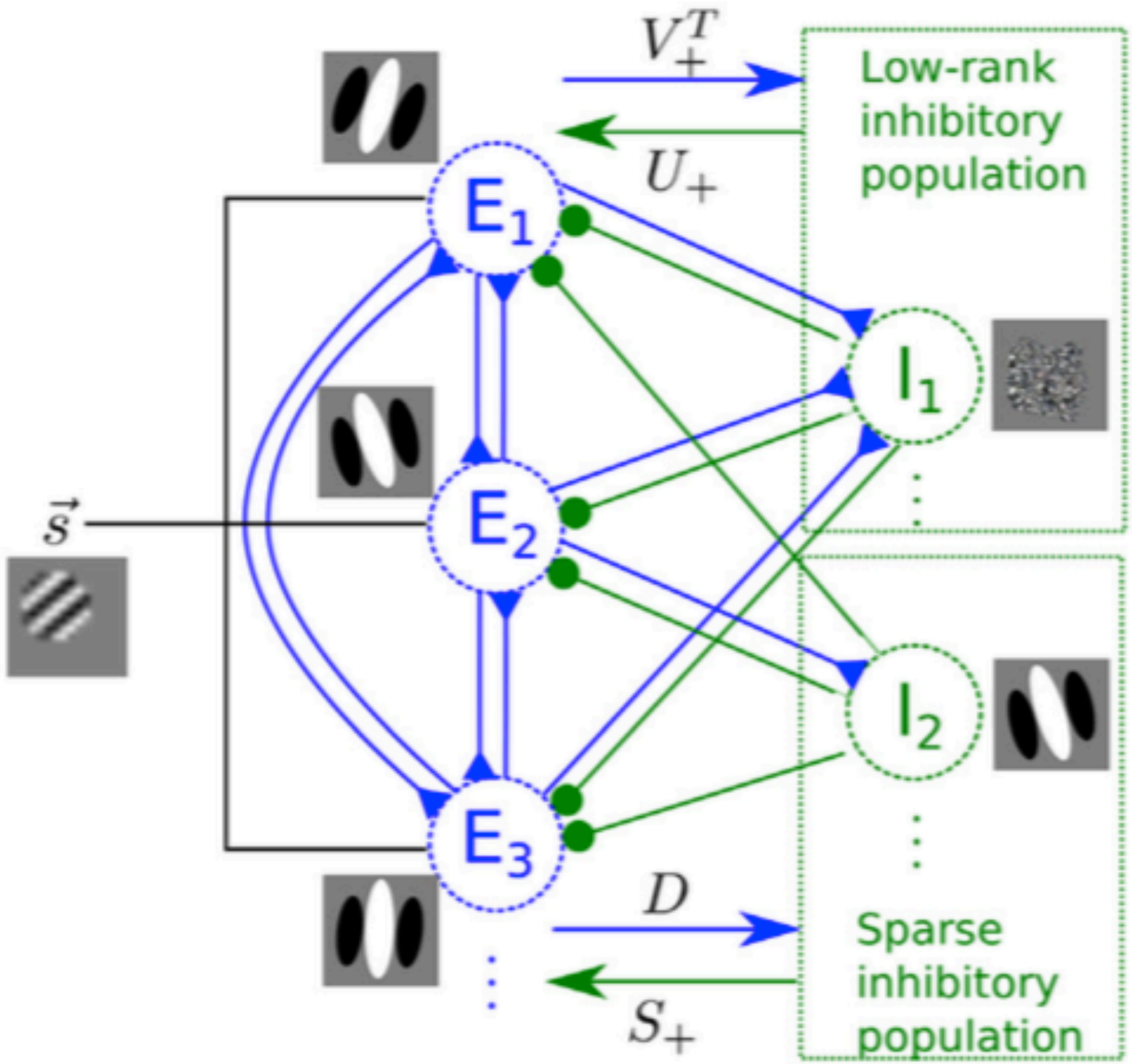
Decompose low rank matrix to separate excitatory and inhibitory connections

$$G \quad = L + S = U \Sigma V^T + S$$

Two subpopulations of inhibitory interneurons, one from L and one from S

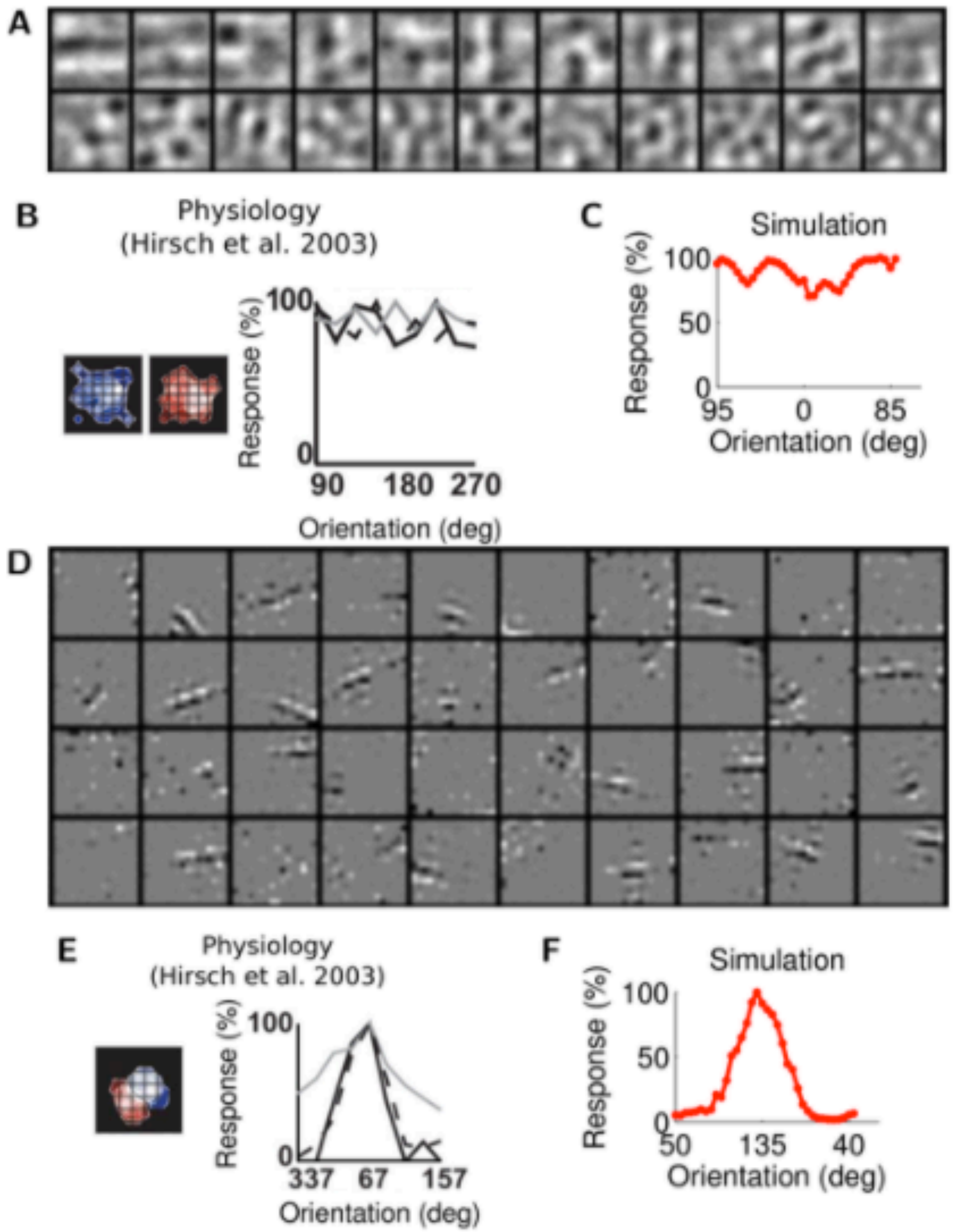
M Zhu, CJ Rozell. "Modeling Inhibitory Interneurons in Efficient Sensory Coding Models." *PLoS Comp Bio. 2015*

Inhibitory Interneurons



M Zhu, CJ Rozell. "Modeling Inhibitory Interneurons in Efficient Sensory Coding Models." *PLoS Comp Bio. 2015*

Subpopulations match experimental data



Low rank population

Sparse population

M Zhu, CJ Rozell. "Modeling Inhibitory Interneurons in Efficient Sensory Coding Models." *PLoS Comp Bio. 2015*

Inhibitory Interneurons

Solves original sparse coding problem

Respects Dale's law

Matches measured E/I cell ratios

Matches diversity of orientation tuning found in mammal study

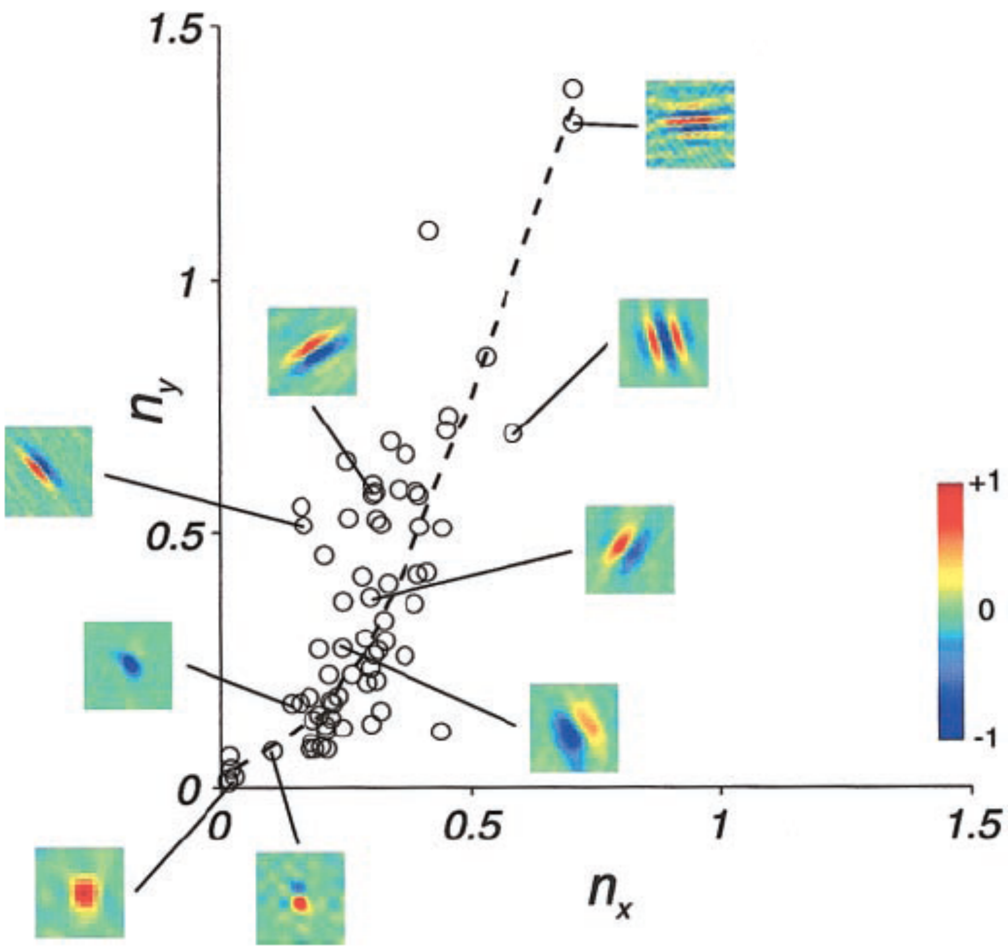
M Zhu, CJ Rozell. "Modeling Inhibitory Interneurons in Efficient Sensory Coding Models." *PLoS Comp Bio. 2015*

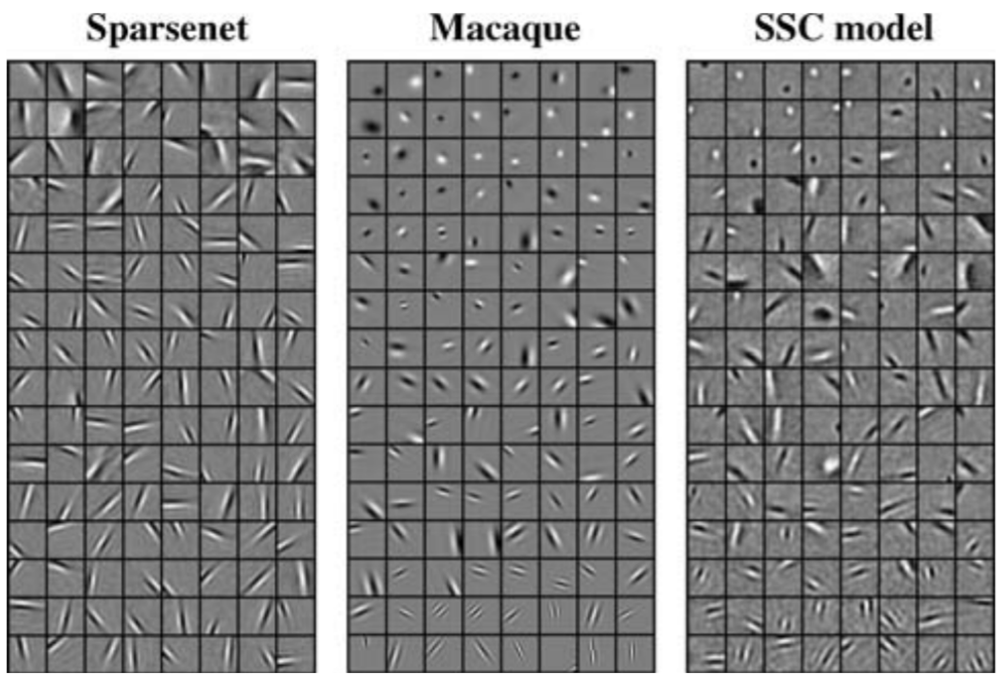
Neuroscience connections: predictions & explanations

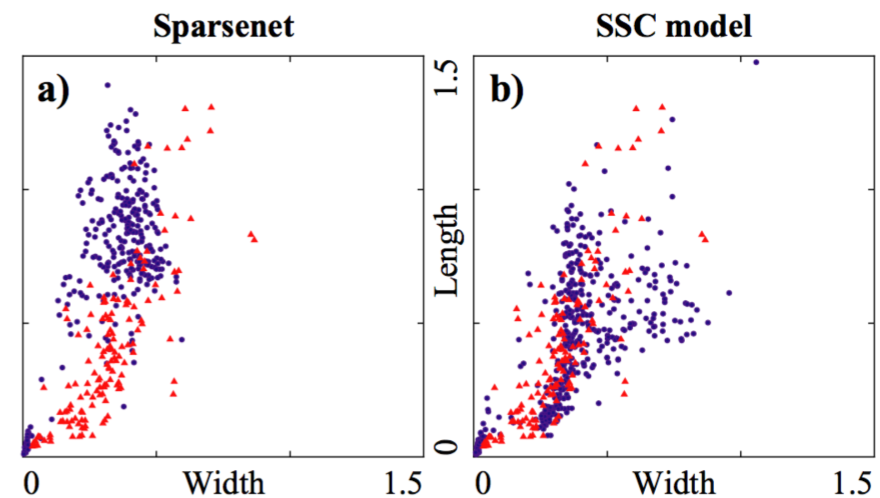
Training set images from van Hateren natural scenes database

Diversity of simple-cell RFs in macaque V1

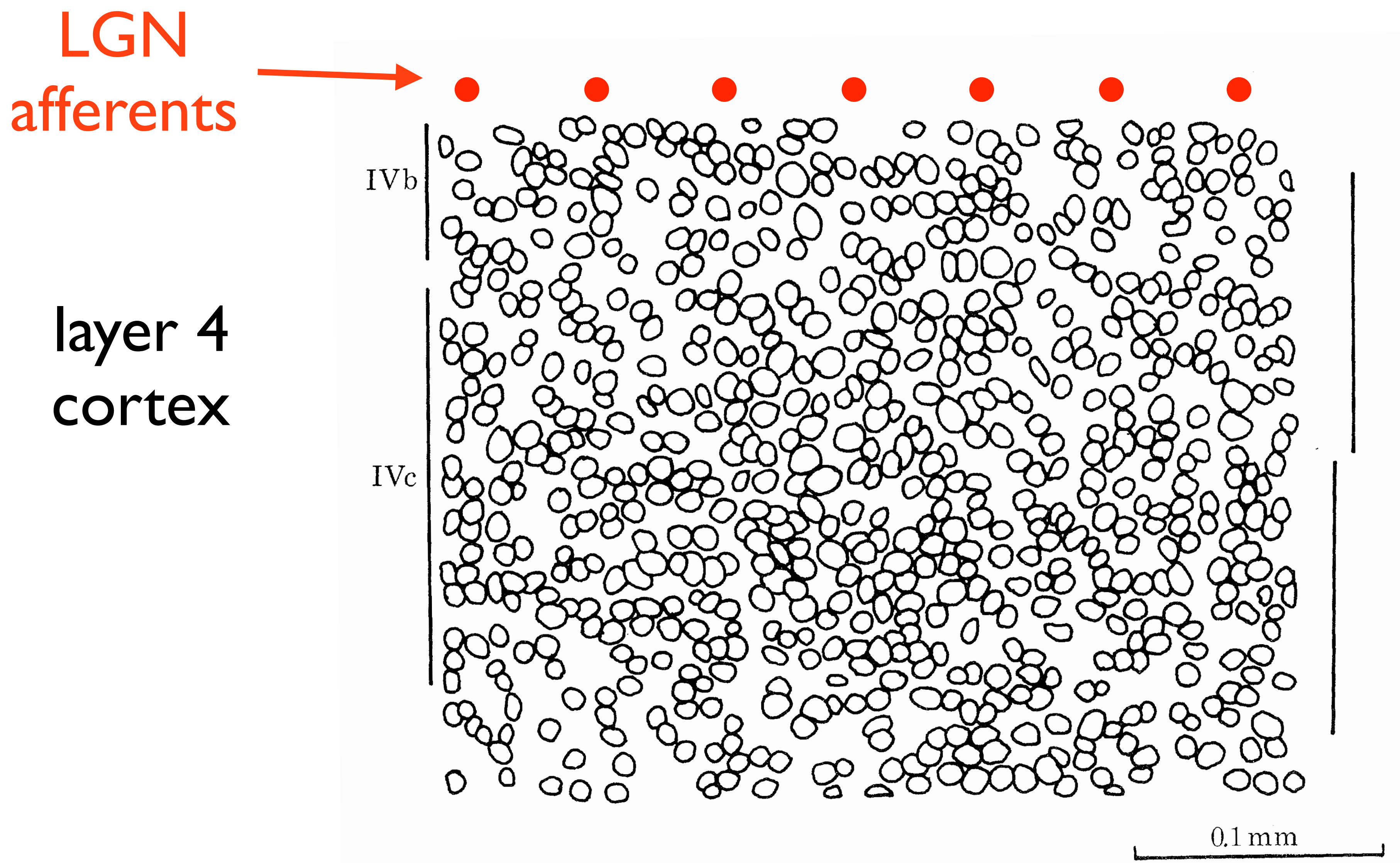
(Ringach 2002, Rehn & Sommer 2007)



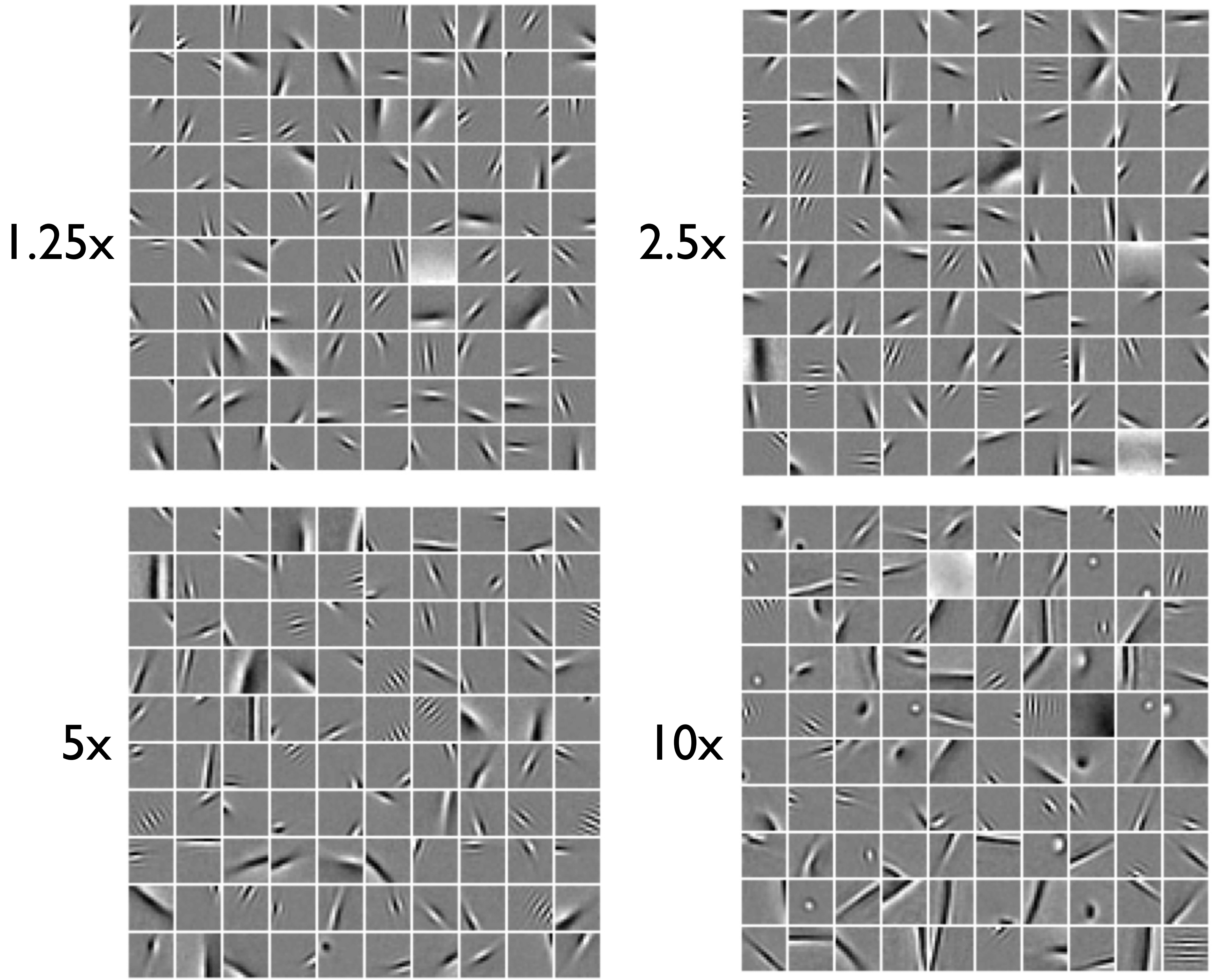




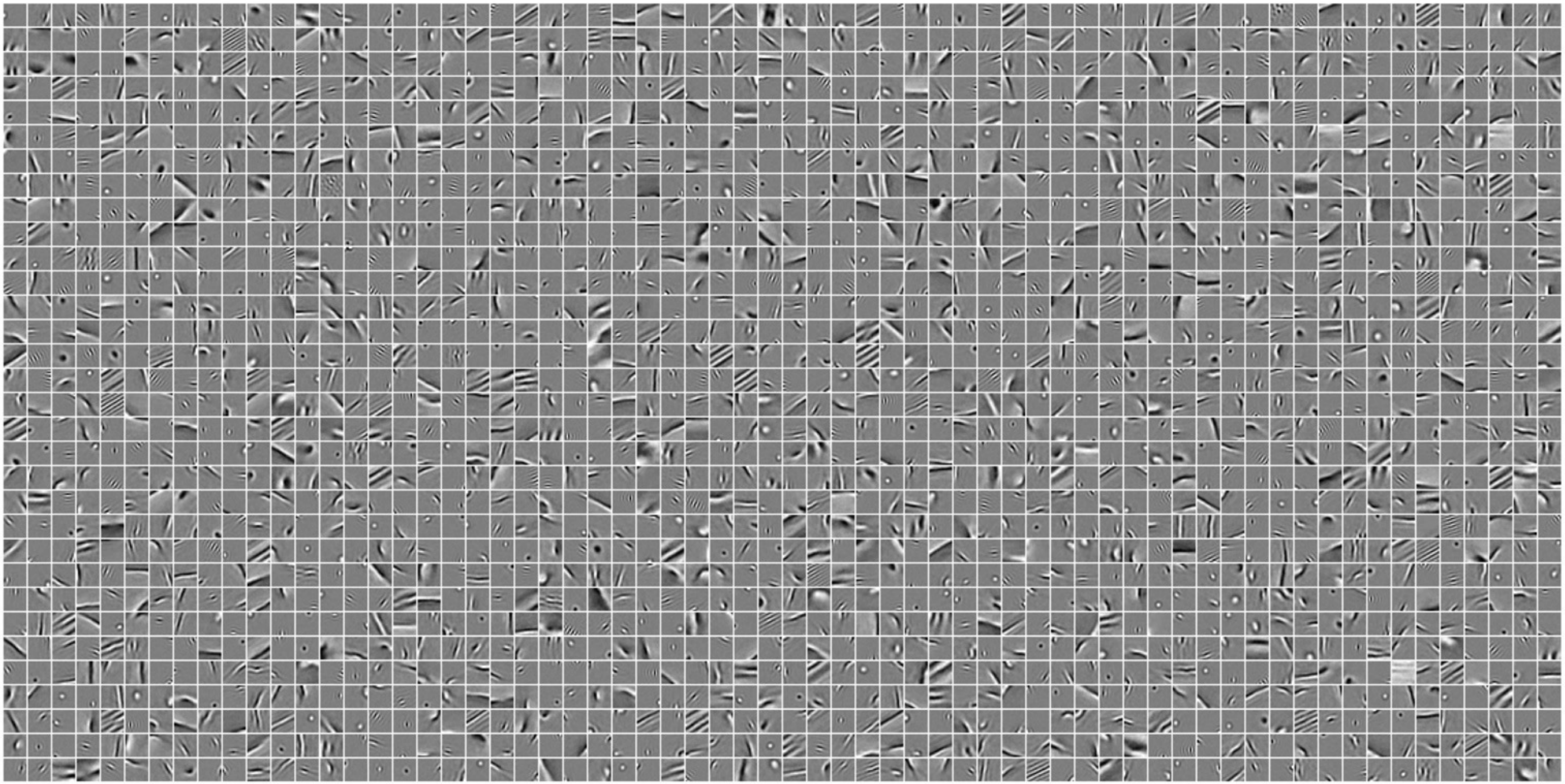
V1 is highly overcomplete

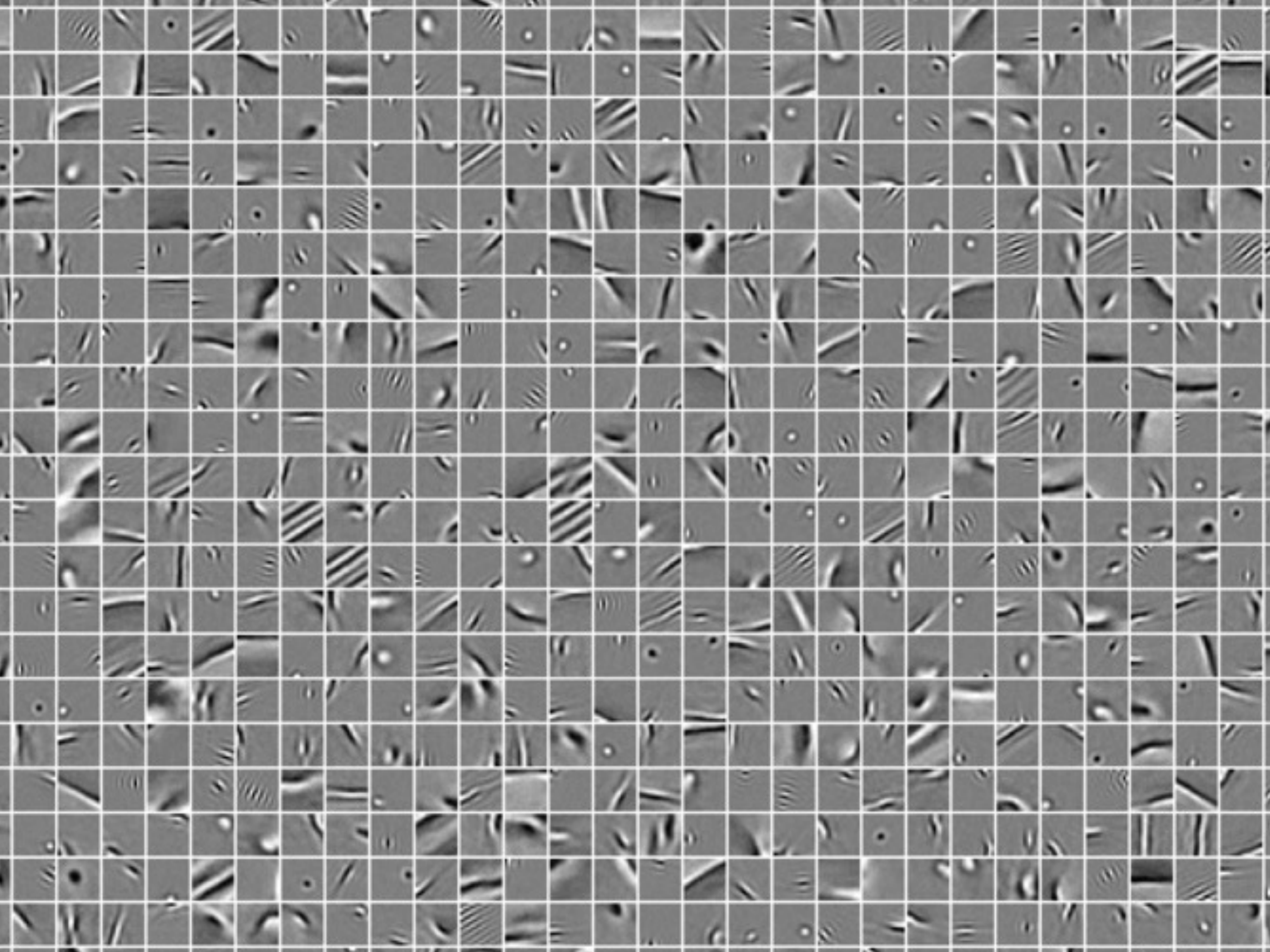


Barlow (1981)


1.25x
2.5x
5x
10x

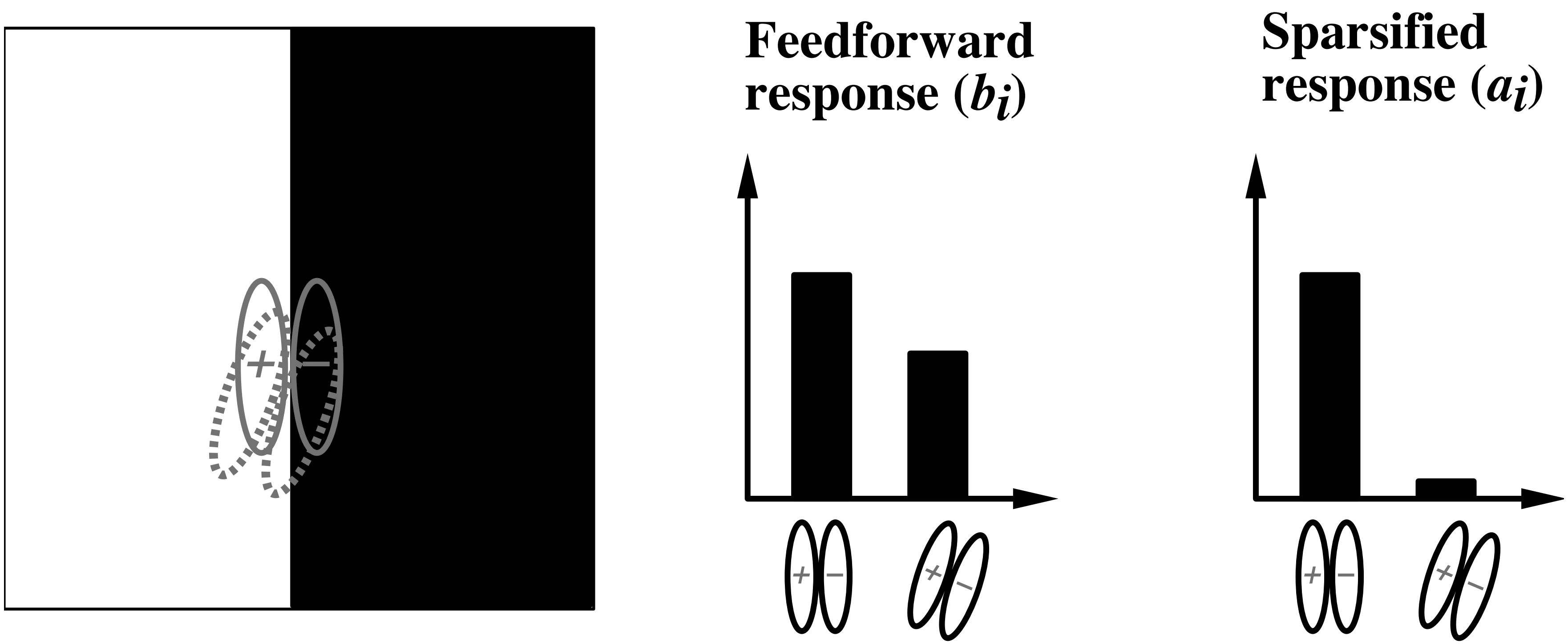
Full 10x dictionary





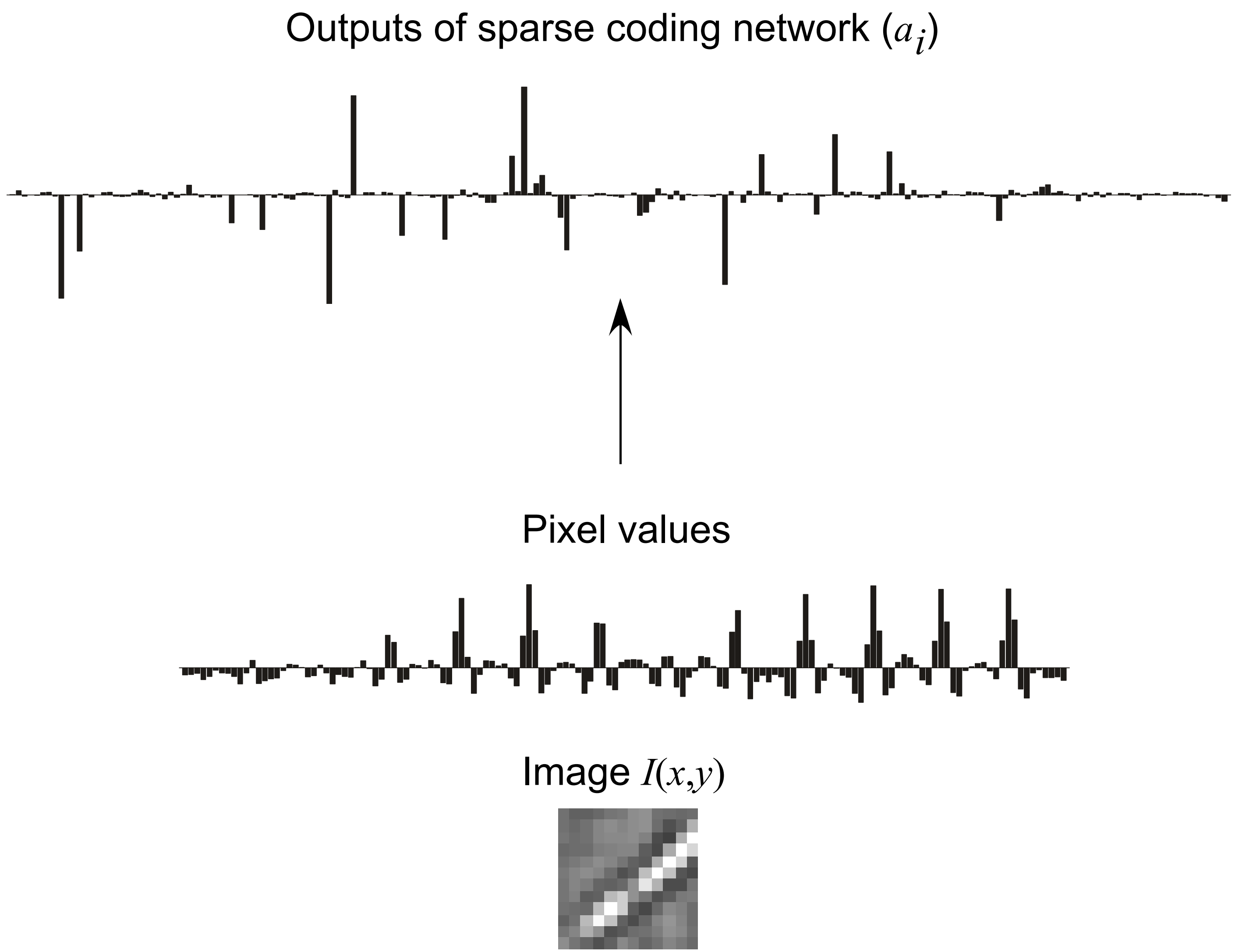
‘Explaining away’

Active inference provides a more descriptive representation

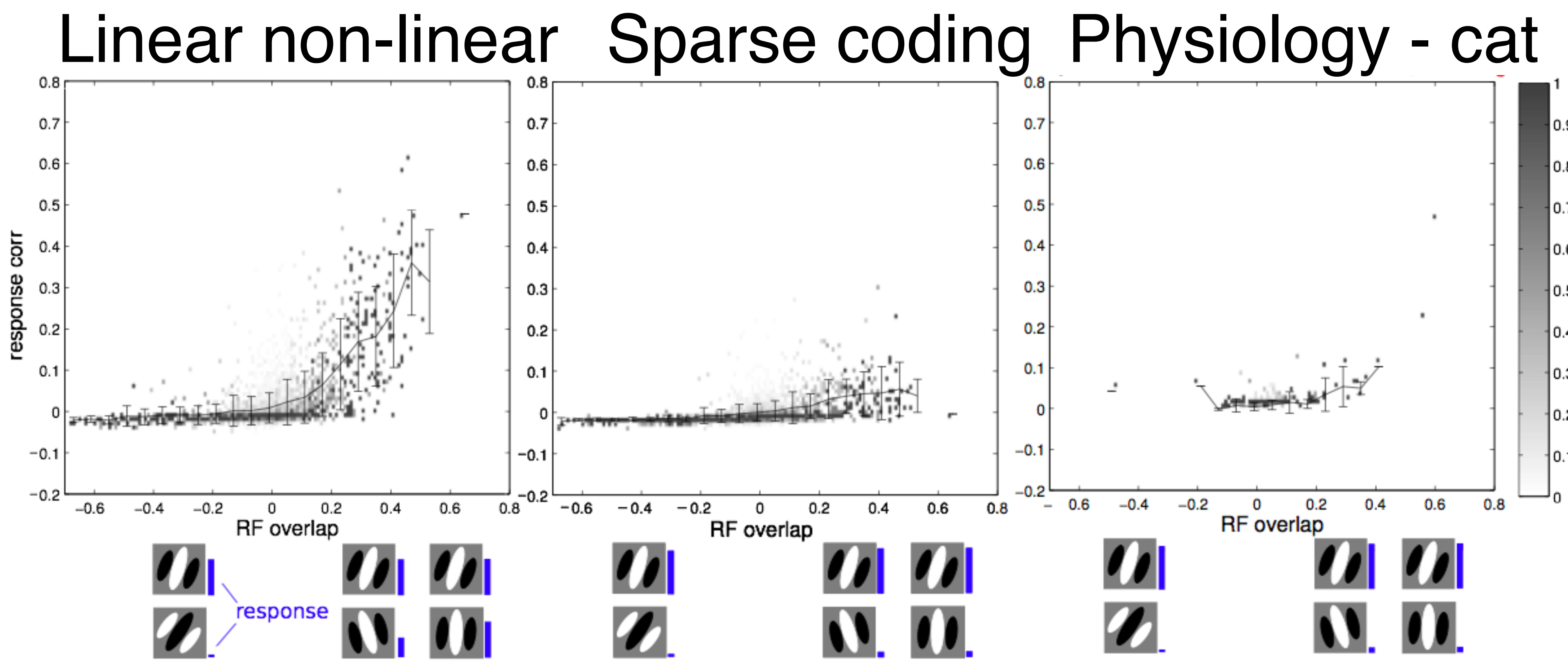


“Population” nonlinearity

Sparsification prediction



Active decorrelation

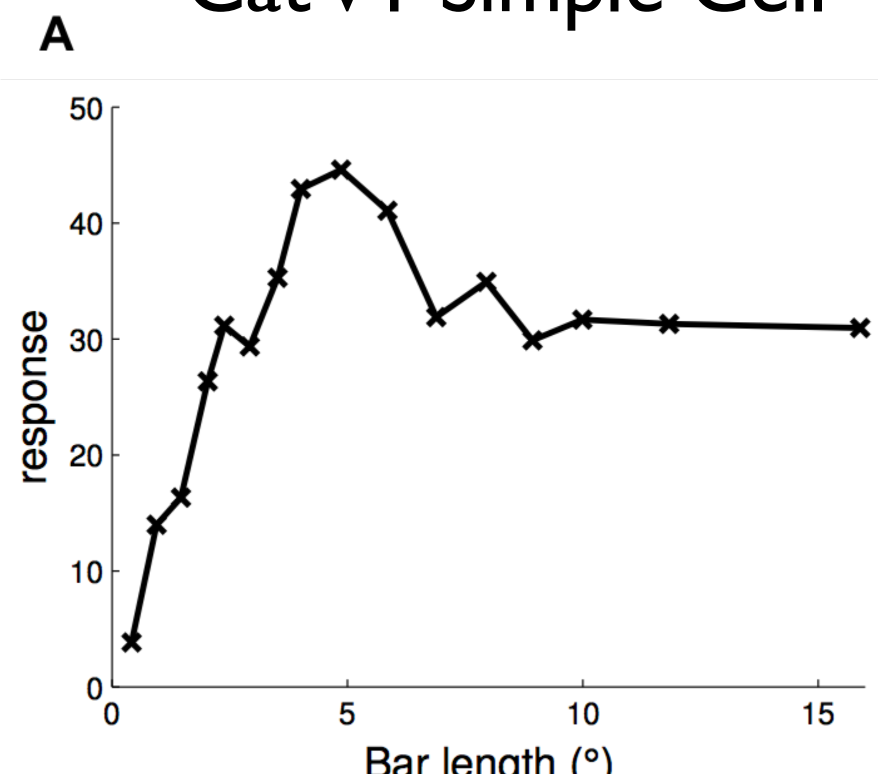


M Zhu, CJ Rozell

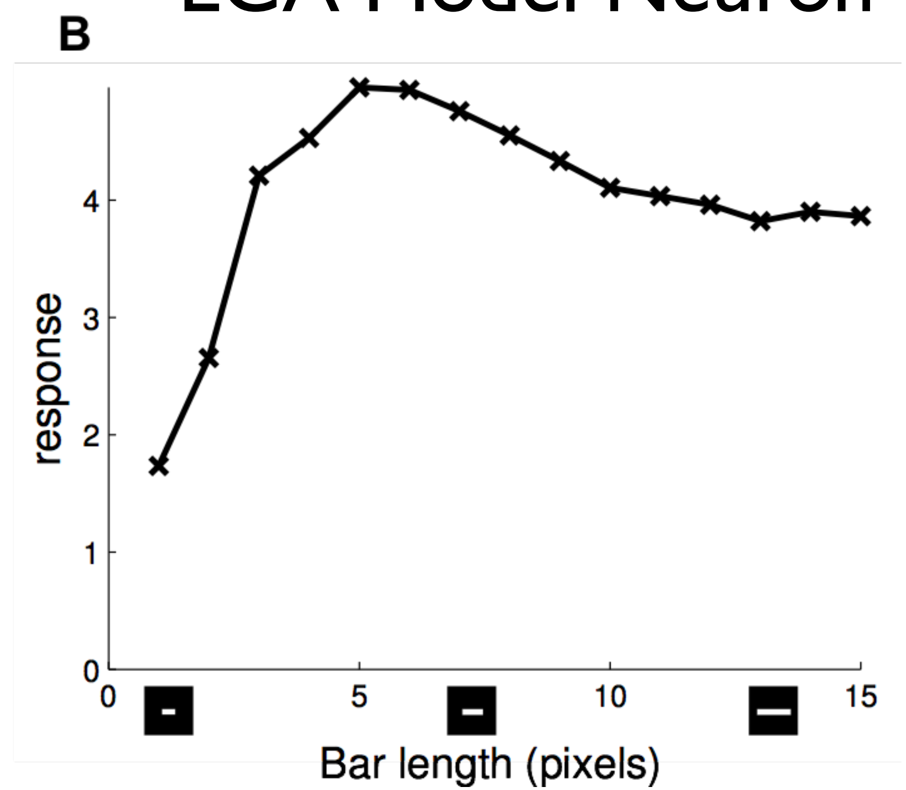
Non-Classical Receptive Field Effects

End-Stopping

Cat V1 Simple Cell



LCA Model Neuron



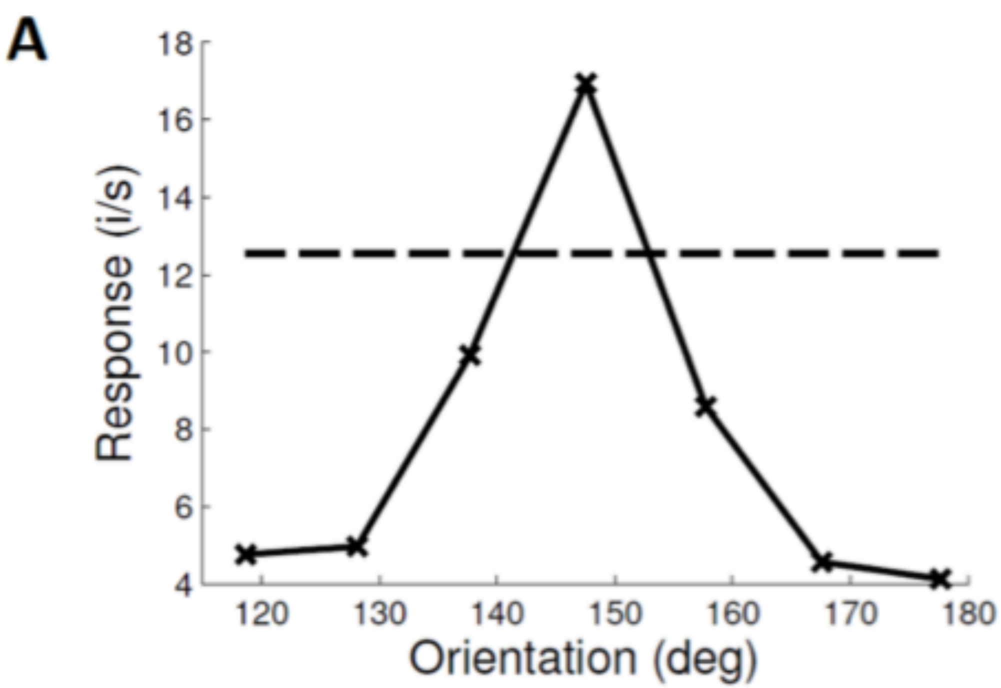
M Zhu, CJ Rozell. "Visual Nonclassical Receptive Field Effects Emerge from Sparse Coding in a Dynamical System", PLoS Comp Bio. 2013

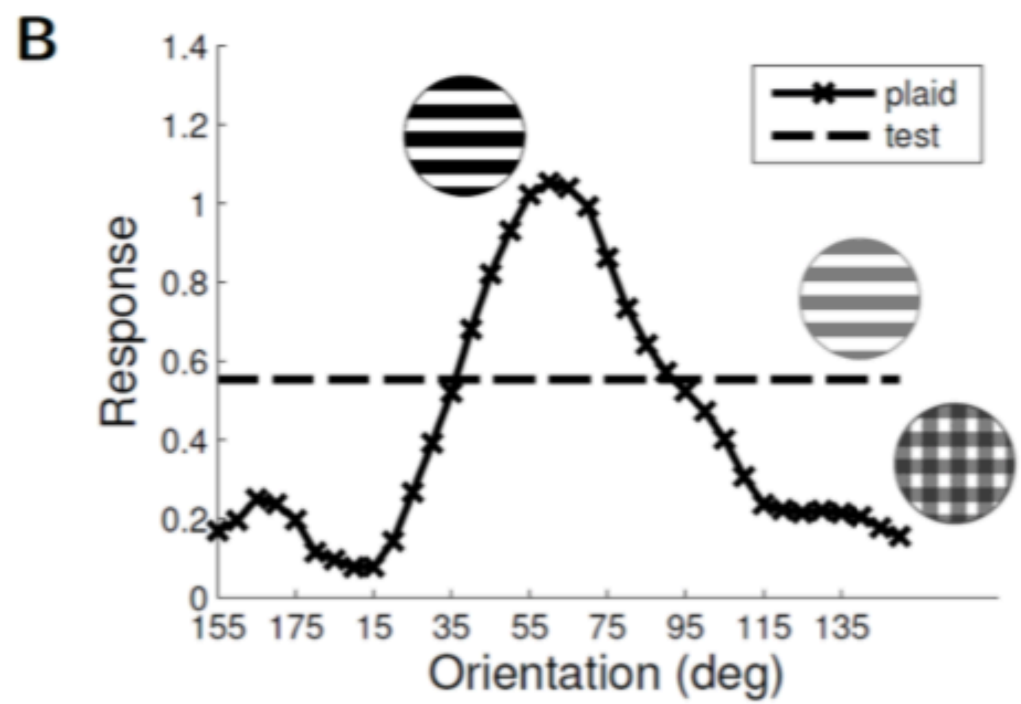
Non-Classical Receptive Field Effects

Cross Orientation Suppression

Cat V1 Simple Cell

LCA Model Neuron





M Zhu, CJ Rozell. "Visual Nonclassical Receptive Field Effects Emerge from Sparse Coding in a Dynamical System", PLoS Comp Bio. 2013

Non-Classical Receptive Field Effects

Other effects compared in this paper:

- End-stopping
- Surround suppression/facilitation
- RF expansion
- contrast invariant orientation tuning
- cross-orientation suppression

M Zhu, CJ Rozell. "Visual Nonclassical Receptive Field Effects Emerge from Sparse Coding in a Dynamical System", PLoS Comp Bio, 2013

Evidence for sparse coding

Mushroom body, locust (Laurent)

HVC, zebra finch (Fee)

Auditory cortex, mouse (DeWeese & Zador)

Hippocampus, rat/primate (Thompson & Best; Skaggs)

Motor cortex, rabbit (Swadlow)

Barrel cortex, rat (Brecht)

Visual cortex, monkey/cat (Vinje & Gallant)

Visual cortex, cat (Gray; McCormick)

Inferotemporal cortex, human (Fried & Koch)

Olshausen BA, Field DJ (2004) Sparse coding of sensory inputs. *Current Opinion in Neurobiology, 14*, 481-487.