System Identification

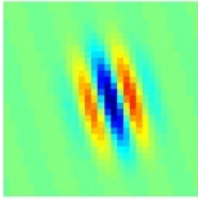
AIL087

Ján Antolík

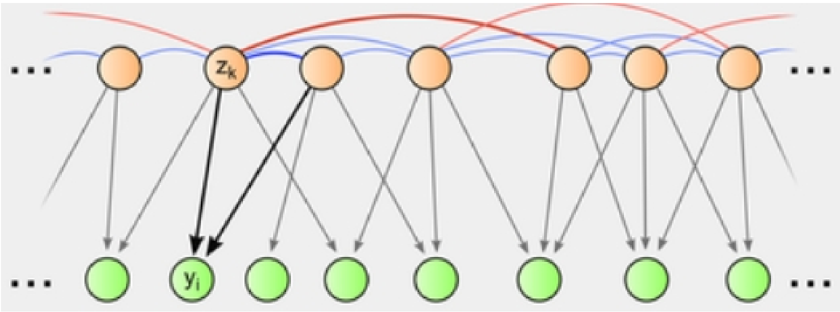
MFF UK, 2019

Computational modeling taxonomy

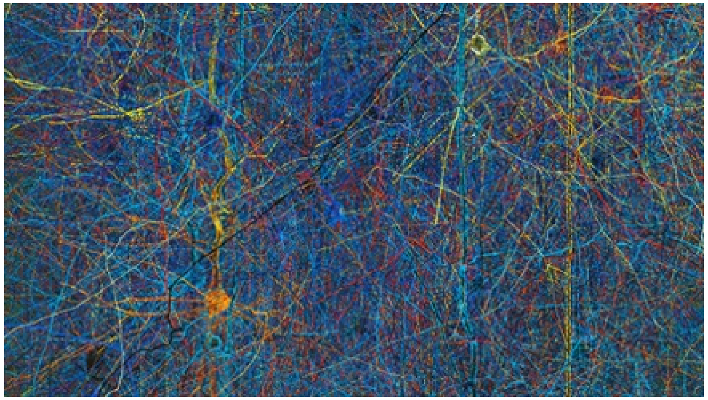
- System identification models



- Normative models
  (often Bayesian based)



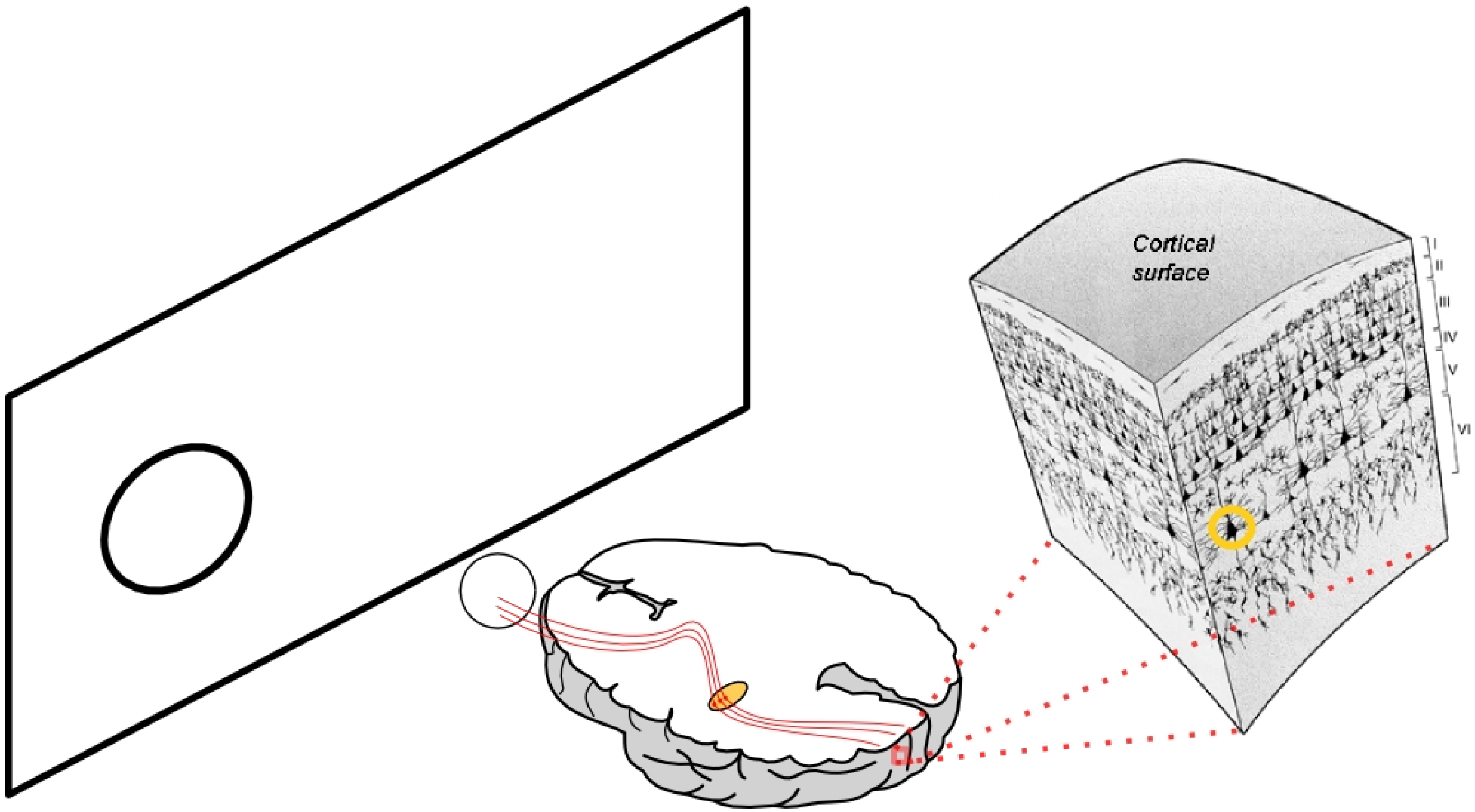
- Full dynamical models
  (aka. simulations)



System identification

- Most linked to data: we directly fit them
- Compact representation of the function: easy interpretability
- Over-simpified – we can't have too many parameters and the fitting has to be mathematically tractable
- No dynamics (mostly) – those are difficult to fit
- Overfitting

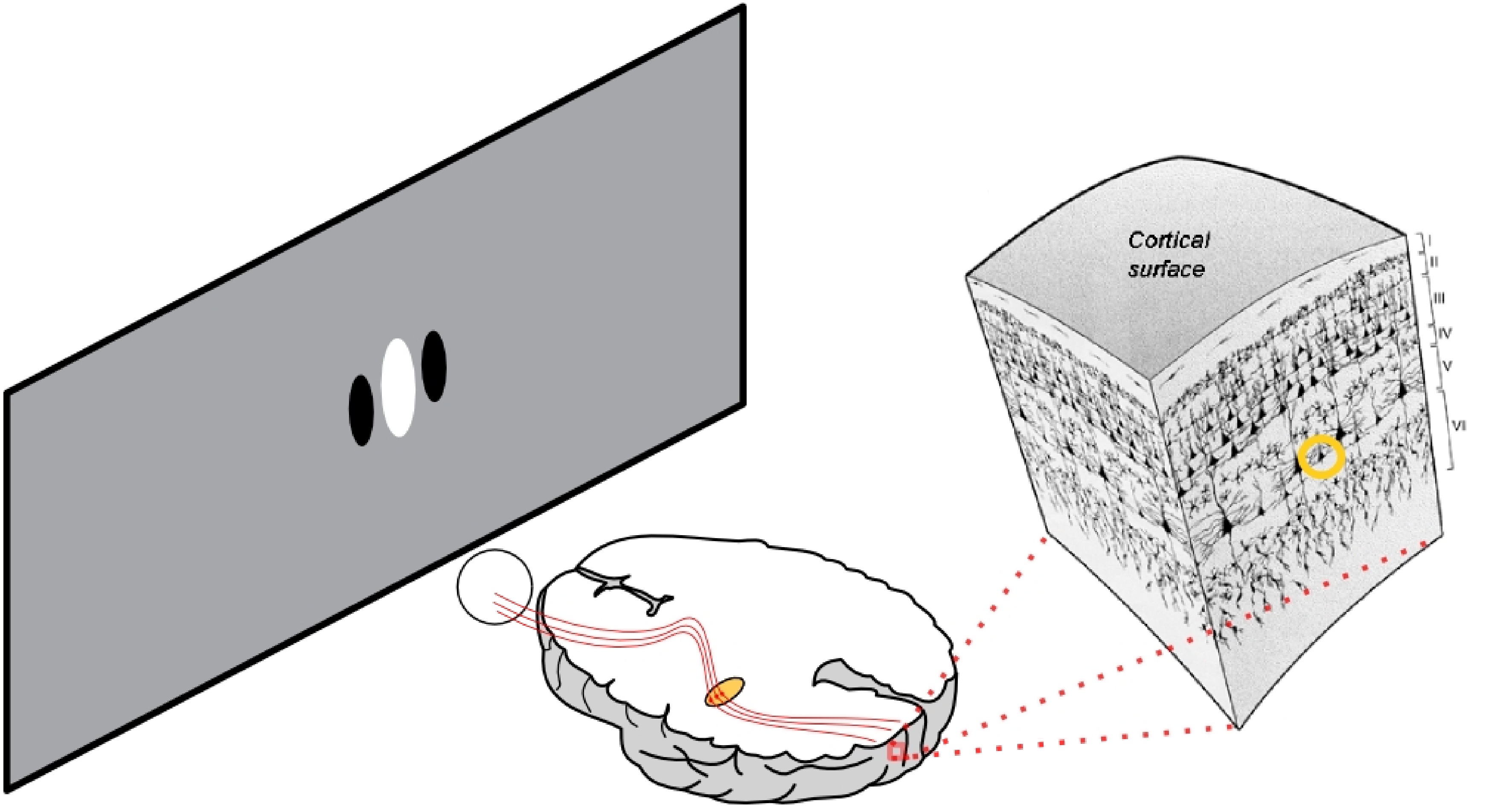
RF: position in visual field



Responses can be obtained in a given optic nerve fiber only upon illumination of a certain restricted region of the retina, termed the receptive field of the fiber.

Hartline, H K (1938)

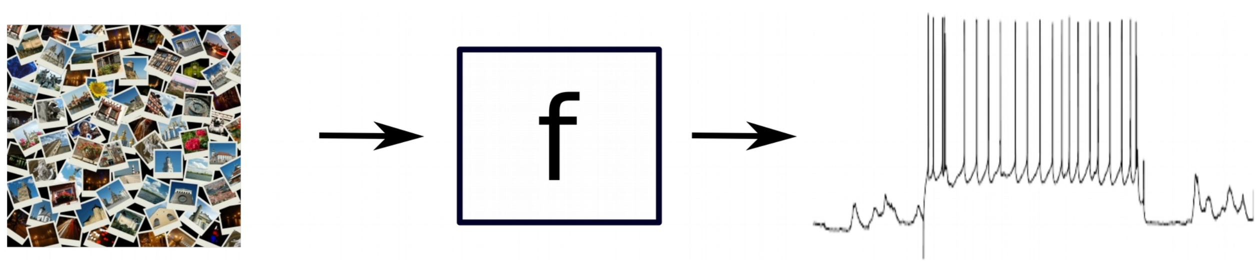
RF: map of exc/inh regions



Kuffler (1953); Hubel and Wiesel (1962)

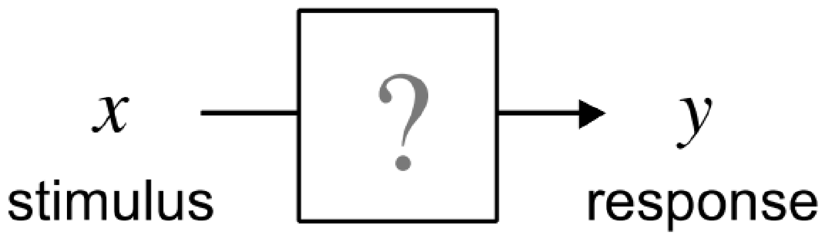
RF as a function

$$f: I^n \rightarrow (0,1)$$

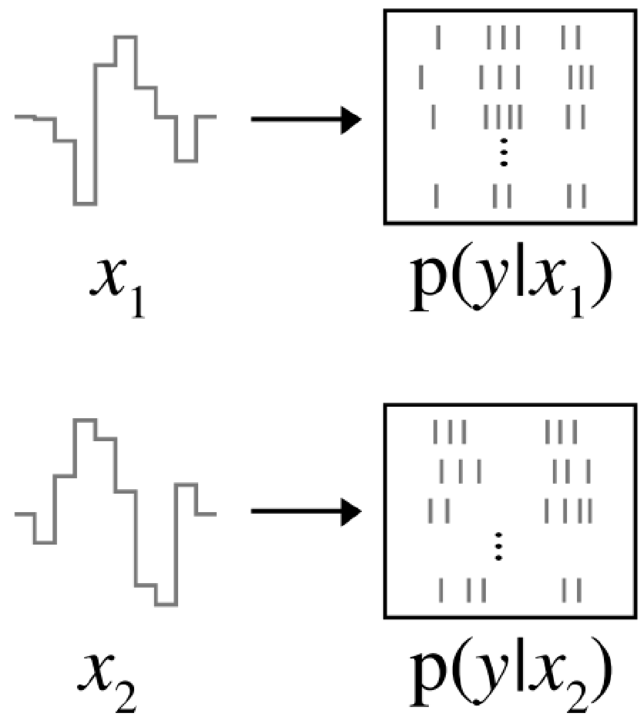


Probabilistic formulation

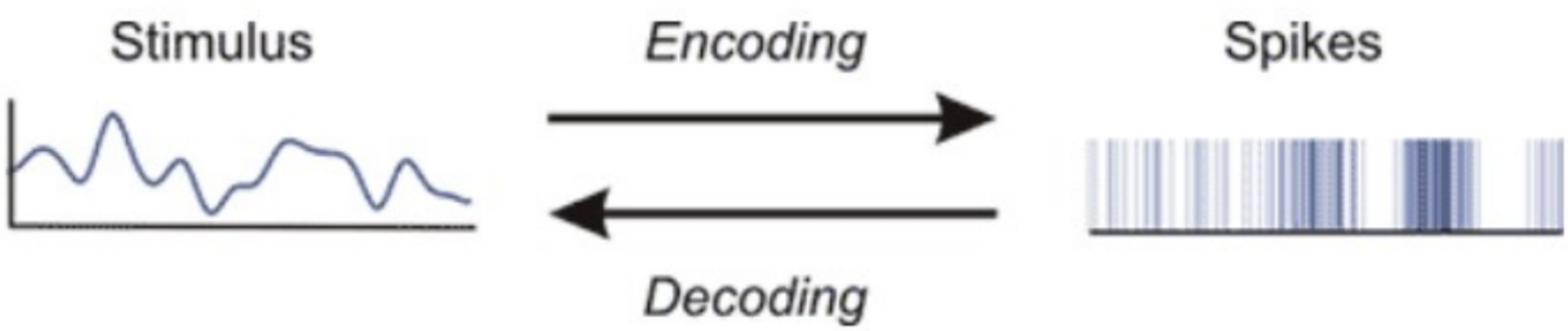
Neural Coding Problem



goal: $p(y|x)$



Encoding vs. Decoding



External world

Brain

- ENCODING: $P(y|x)$ How is information about stimulus transformed into spikes
- DECODING: $P(x|y)$ How to infer stimulus (or output variable) given spikes

Encoding vs. Decoding

S – Stimulus
R – Response

Encoding		*Decoding*
$P(R \mid S)$	vs	$P(S \mid R)$

Encoding vs. Decoding

S – Stimulus
R – Response

Encoding $P(R|S)$ vs *Decoding* $P(S|R)$

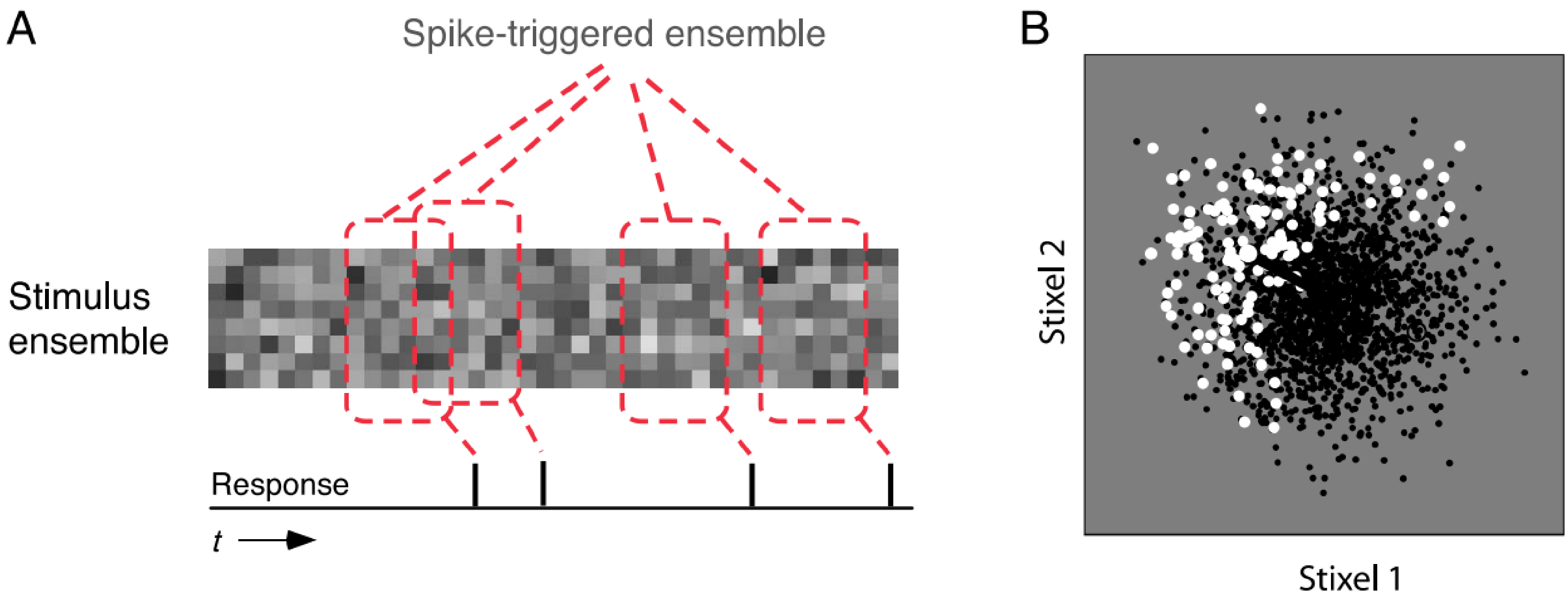
Encoding and decoding are related by **Bayes theorem**:

$$P(R,S) = P(R|S)P(S) = P(S|R)P(R)$$

$$P(S|R) = \frac{P(R|S)P(S)}{P(R)}$$

Spike Triggered Formalism

Spike–triggered ensamble



Statistical model based approach

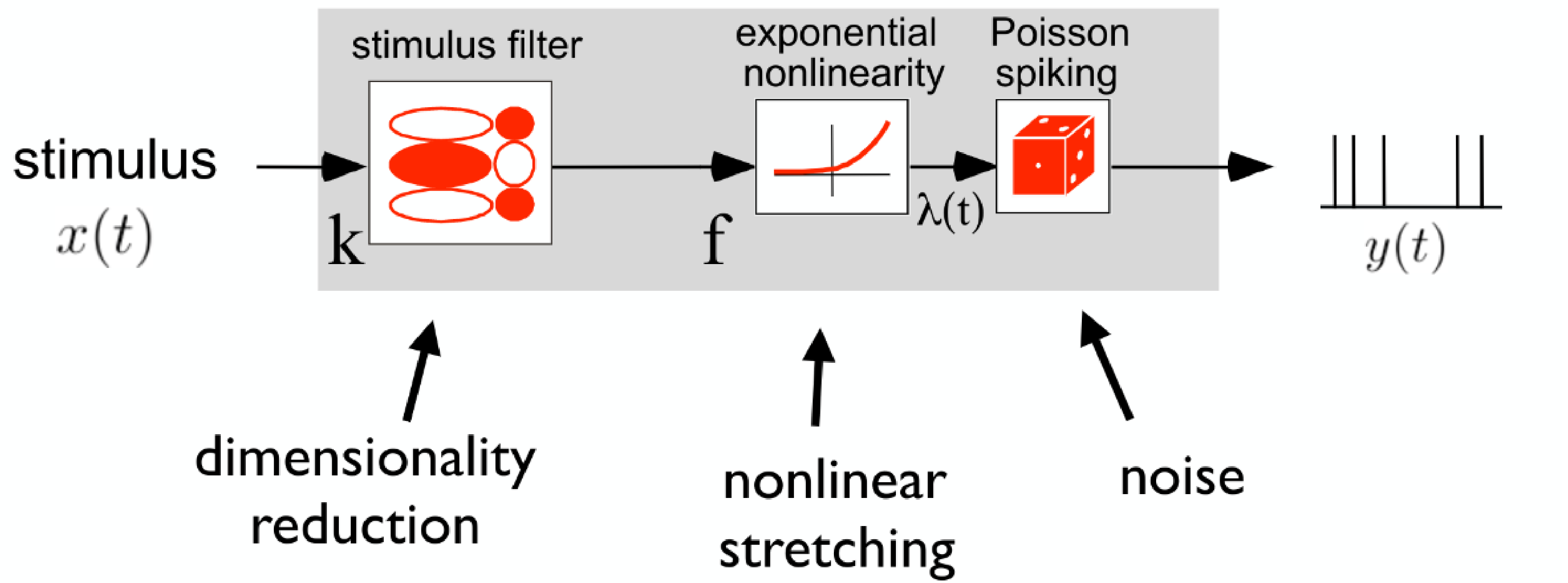
$$p(y|x) \approx \underbrace{}_{}p(y|x,\theta)$$

Response Stimulus Model

Maximum likelyhood:

$$\hat{\theta} = \arg\max_{\theta} p(\mathbf{y}|\mathbf{x},\theta).$$

Linear non–linear Poisson (LNP) model



Poisson firing

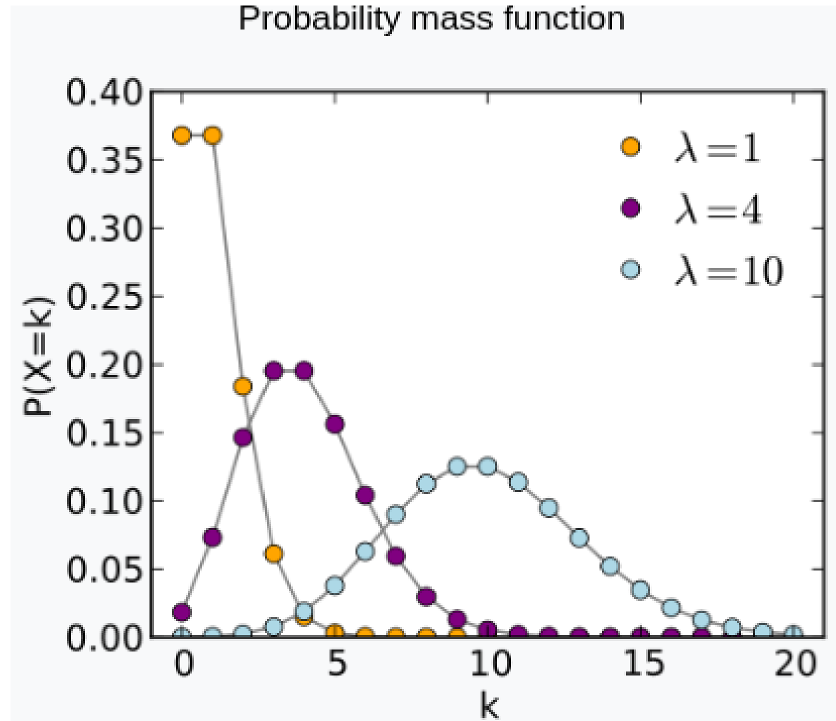
y $\sim$ Poiss(λ)

$$P(y) = \frac{1}{y!}\lambda^y e^{-\lambda}$$

y – spike count

λ *– spike rate*

conditional independentce

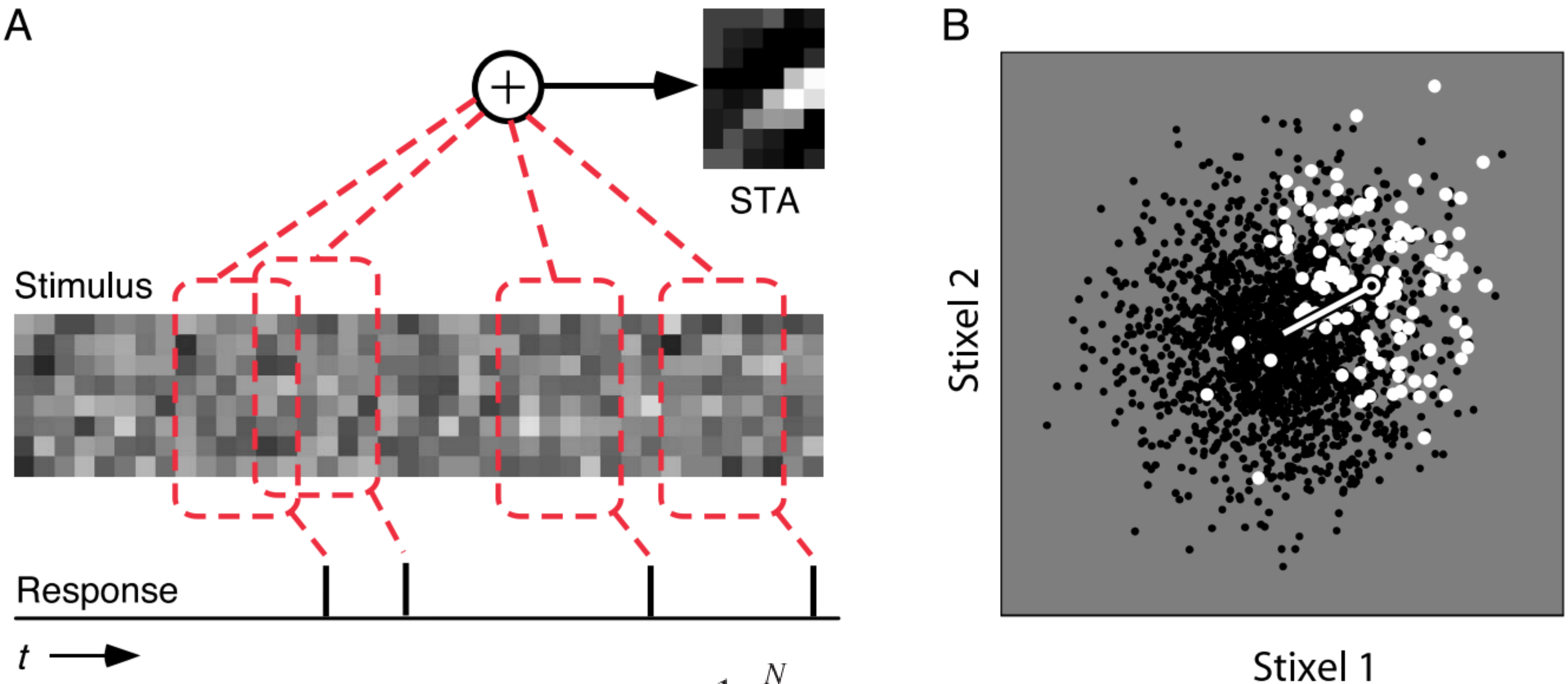


How to estimate linear filters?

- Look for deviations between raw and spike triggered stimulus.
- For LNP model, assuming
  - raw stimulus spherically symetric
  - non-linearity leads to shift of the mean of the spike-triggered ensamble relative to raw stim.
- Spike-triggered averiging:

$$\hat{A} = \frac{1}{N} \sum_{n=1}^{N} \vec{s}\,(t_n),$$

Spike−triggered averaging



$$\hat{A} = \frac{1}{N} \sum_{n=1}^{N} \vec{s}\,(t_n),$$

LNP model fitting

Maximum likelyhood:

$$\hat{\theta} = \arg\max_{\theta} p(\mathbf{y}|\mathbf{x}, \theta).$$

Assumption of conditional Independence:

$$p(\mathbf{y}|\mathbf{x}, \theta) = \prod_i p(y_i|x_i, \theta),$$

Fill in Poisson distribution:

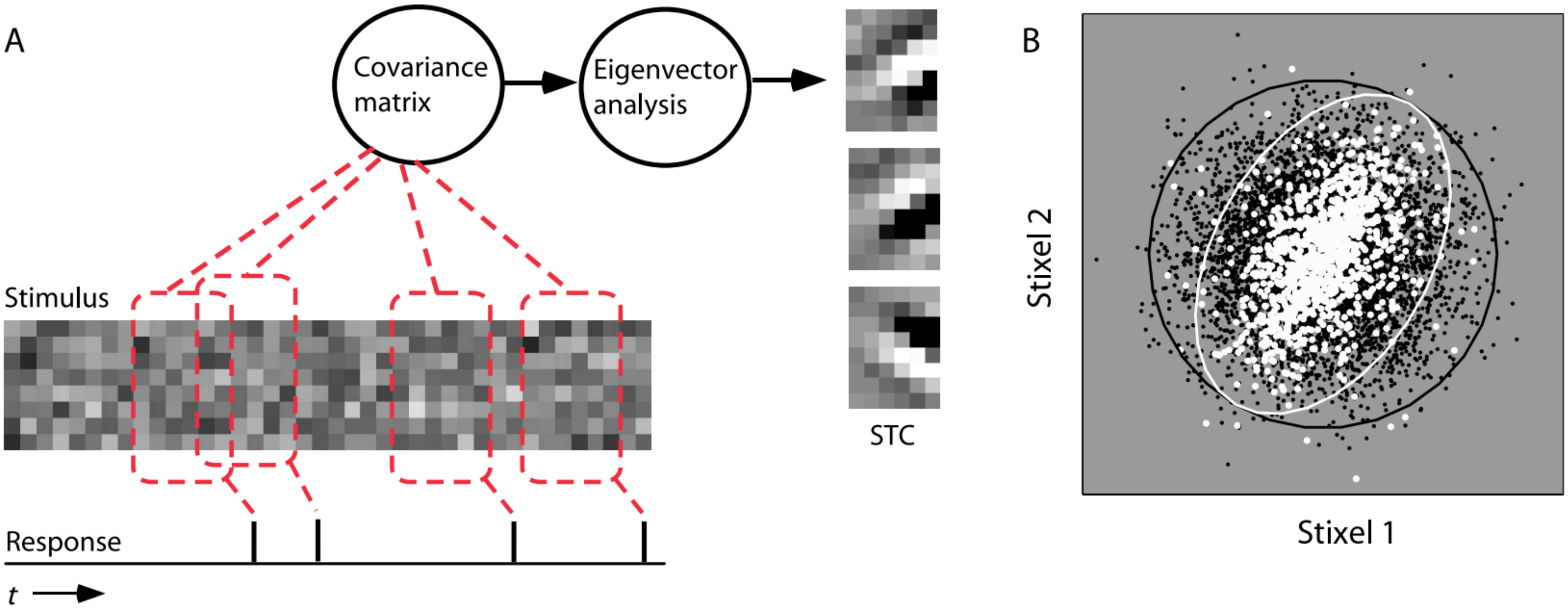
$$p(y_i|x_i, \theta) = \frac{1}{y_i!} \left[\Delta f(k \cdot x_i)\right]^{y_i} e^{-\Delta f(k \cdot x_i)}$$

$$p(\mathbf{y}|\mathbf{x}, \theta) = \Delta^n \prod_i \frac{f(k \cdot x_i)^{y_i}}{y_i!} e^{-\Delta f(k \cdot x_i)}$$

Log-likelyhood:

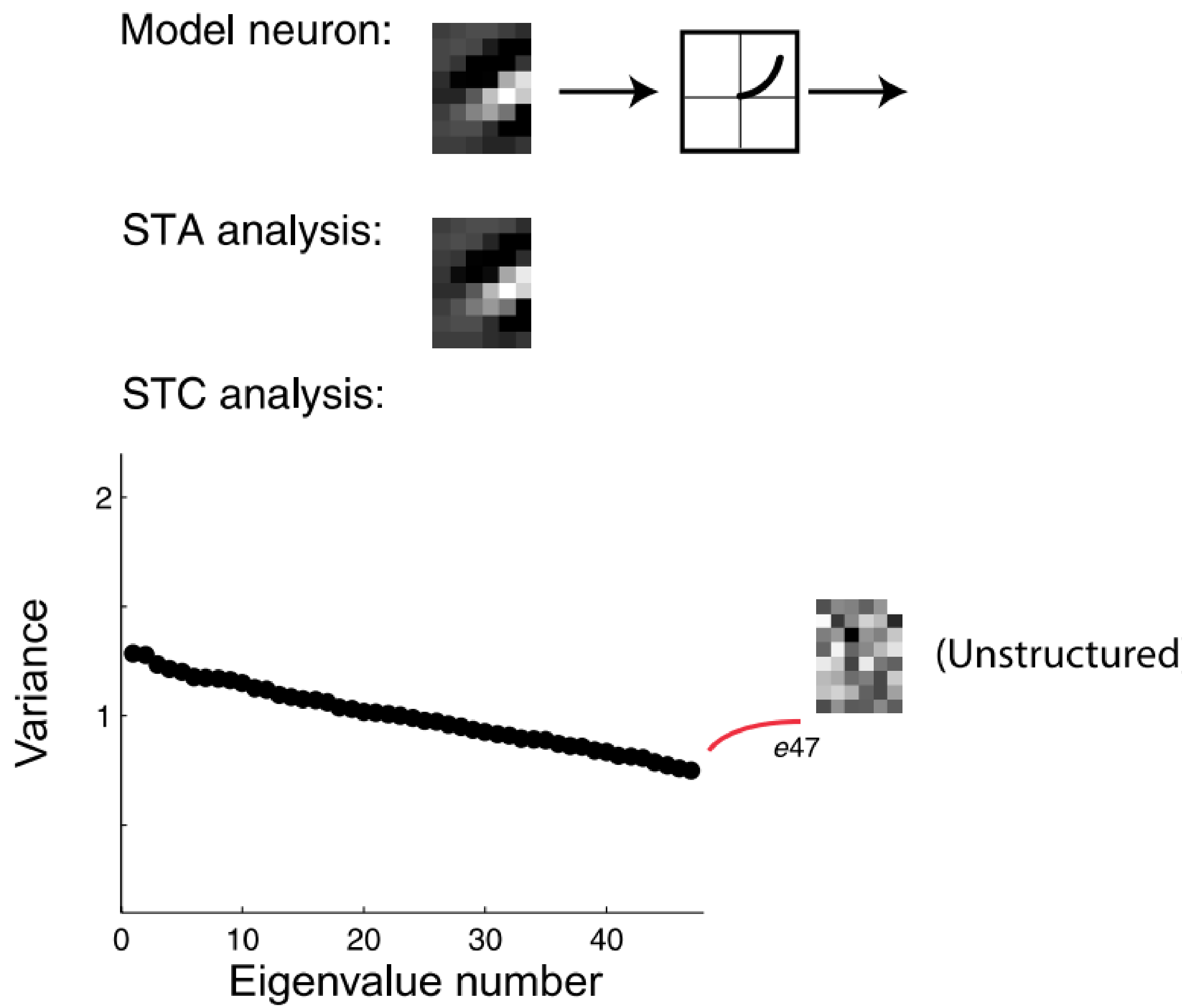
$$\log p(\mathbf{y}|\mathbf{x}, \theta) = \sum_i y_i \log f(k \cdot x_i) - \Delta \sum_i f(k \cdot x_i) + c,$$

Spike–triggered covariance

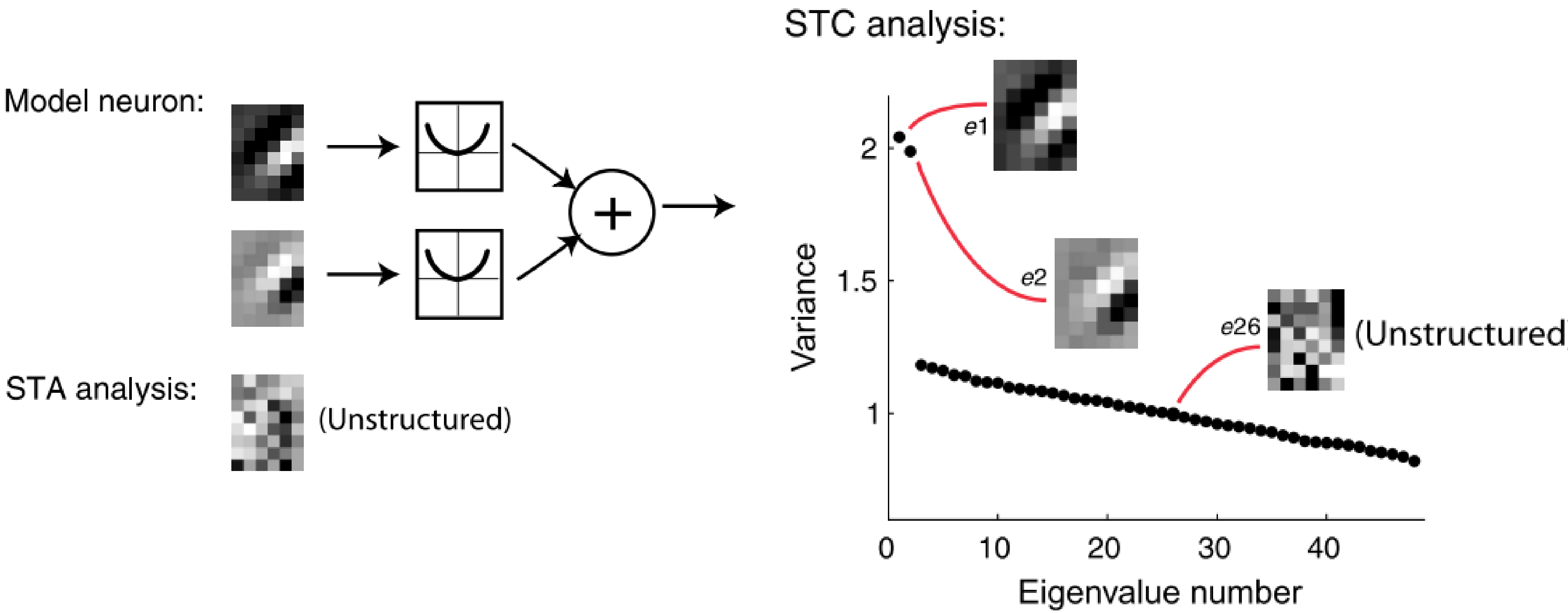


$$\hat{C} = \frac{1}{N-1} \sum_{n=1}^{N} \left[\vec{s}(t_n) - \hat{A}\right] \left[\vec{s}(t_n) - \hat{A}\right]^T,$$

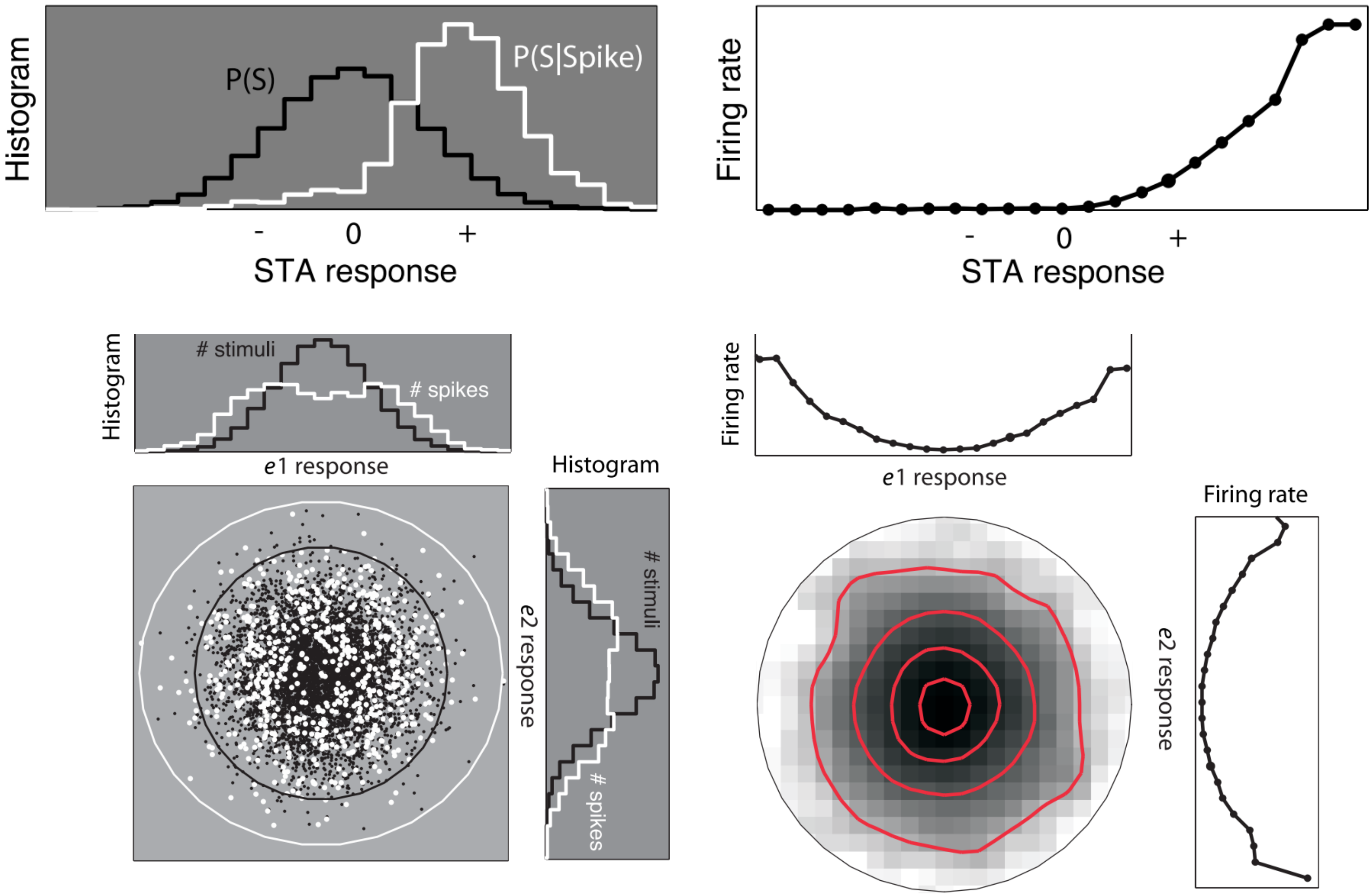
STC: fitting linear model



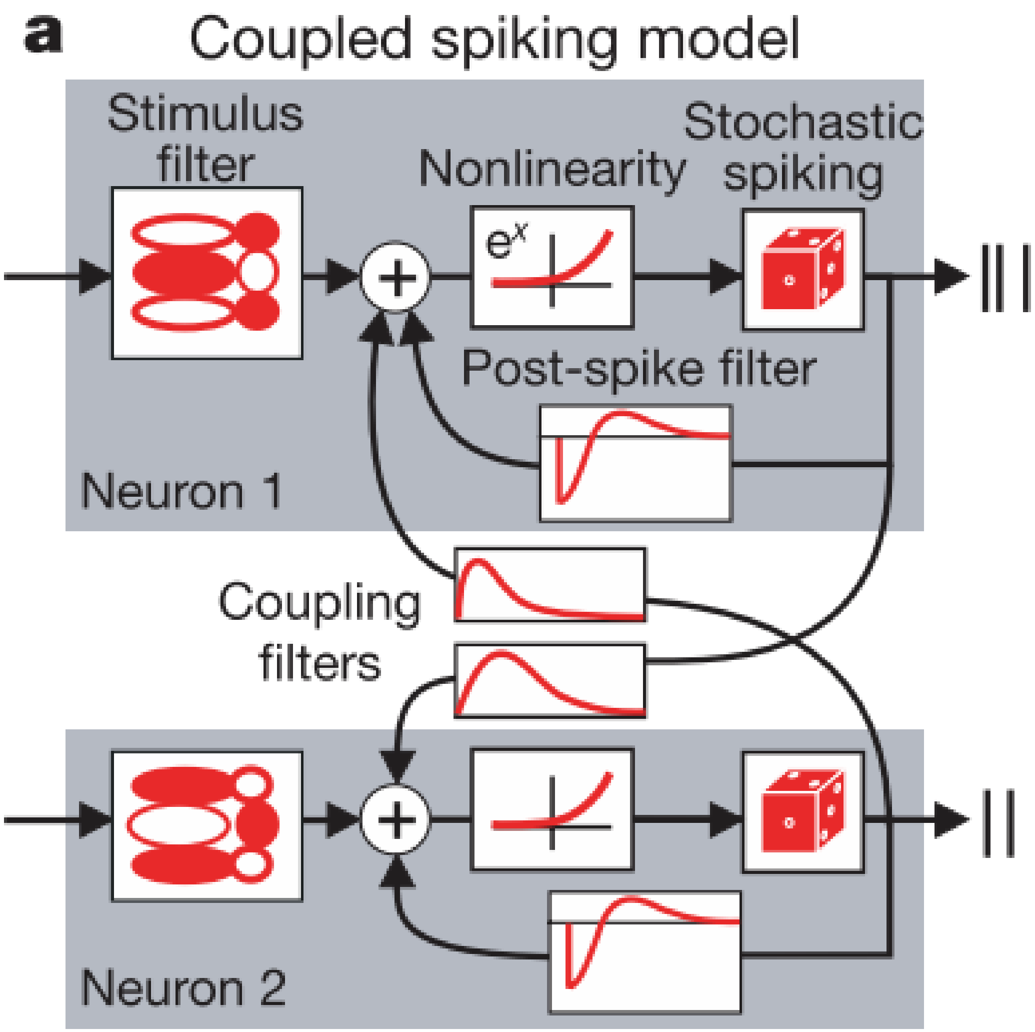
STC: fitting non−linear model



STC: estimating non-linearity



Coupling in GLMs

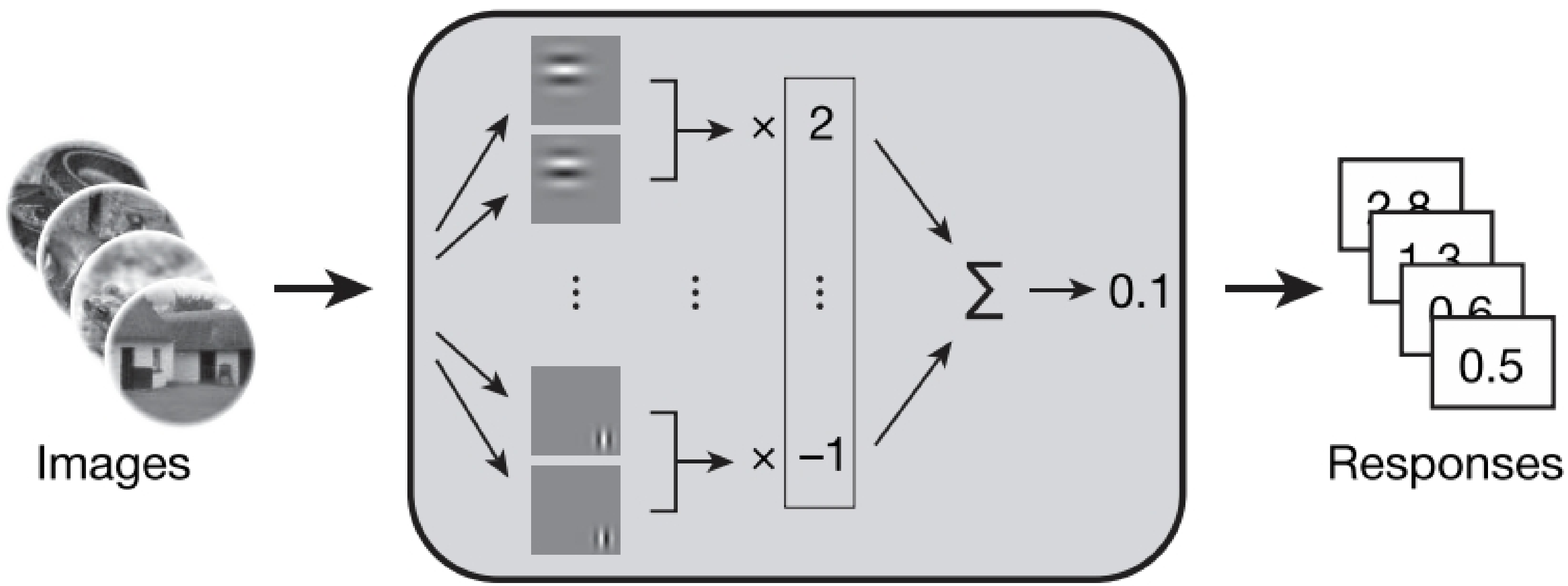


Beyond Linear Models

Barkley Wavelet Transform

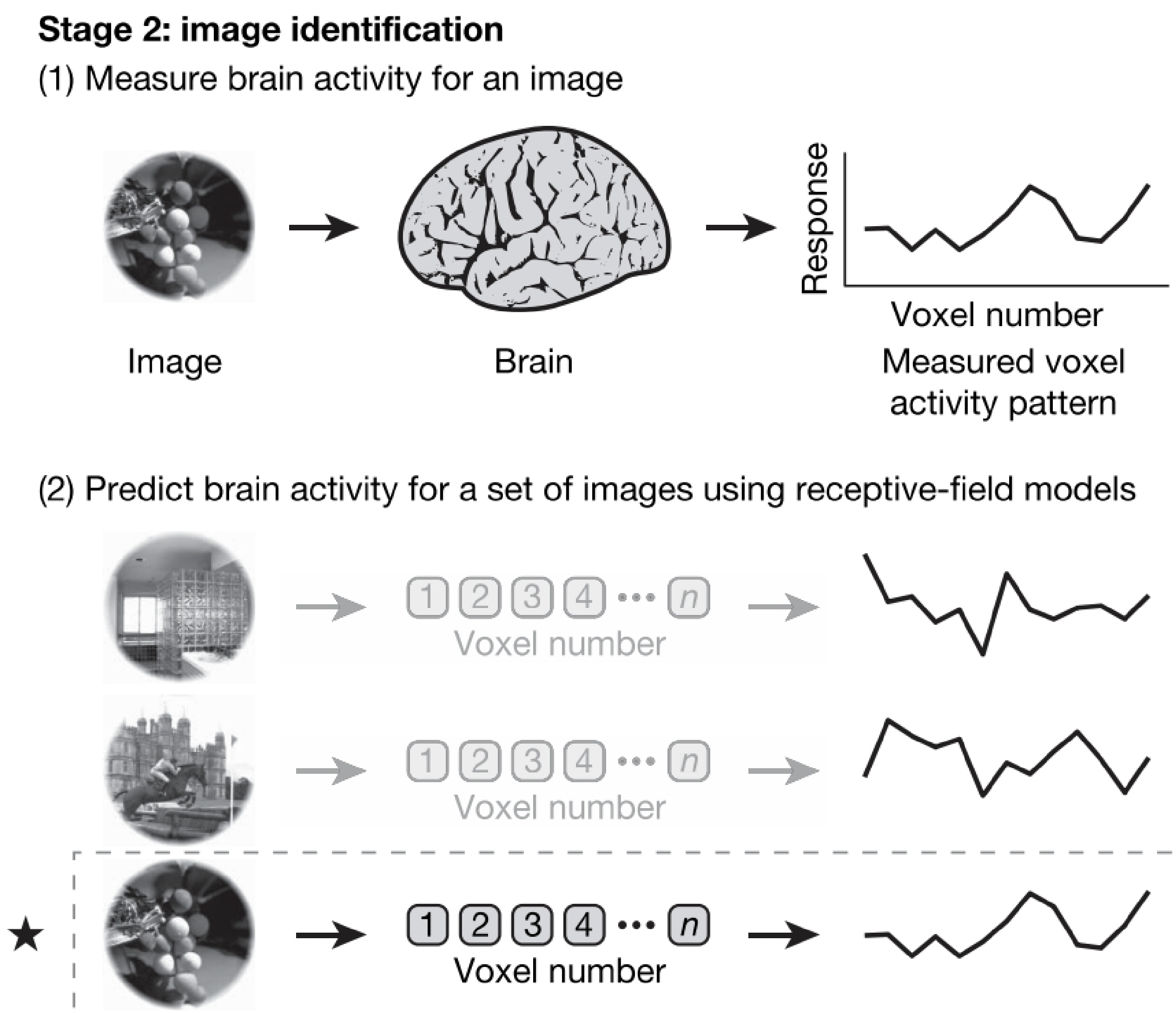
Stage 1: model estimation

Estimate a receptive-field model for each voxel



Receptive-field model for one voxel

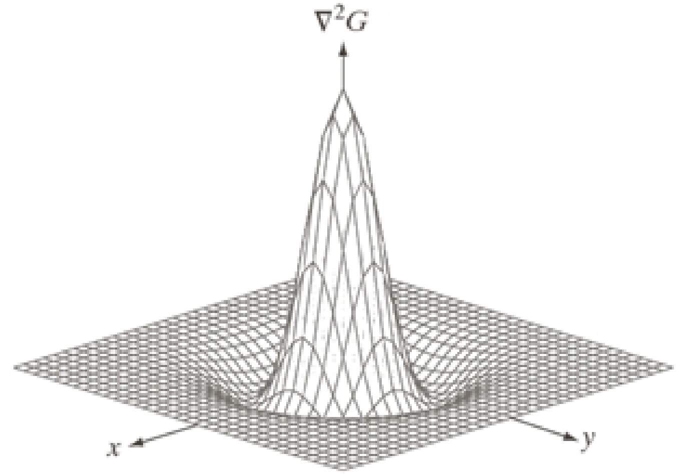
Barkley Wavelet Transform



Hierarchical structural model (HSM)

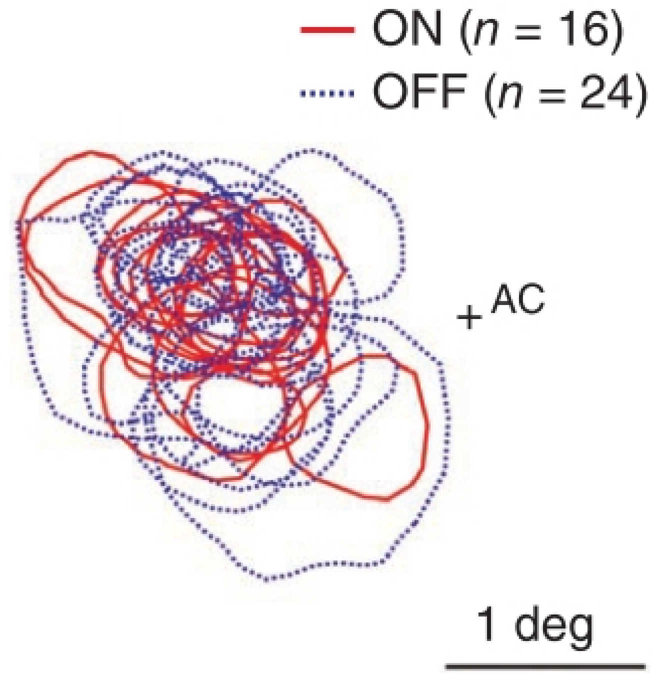
The structural priors

- Receptive fields of LGN units can be well approximated by difference-of-Gaussian function



The structural priors

- Receptive fields of LGN units can be well approximated by difference-of-Gaussian function
- Local population of V1 neurons receives common input from limited number of LGN cells



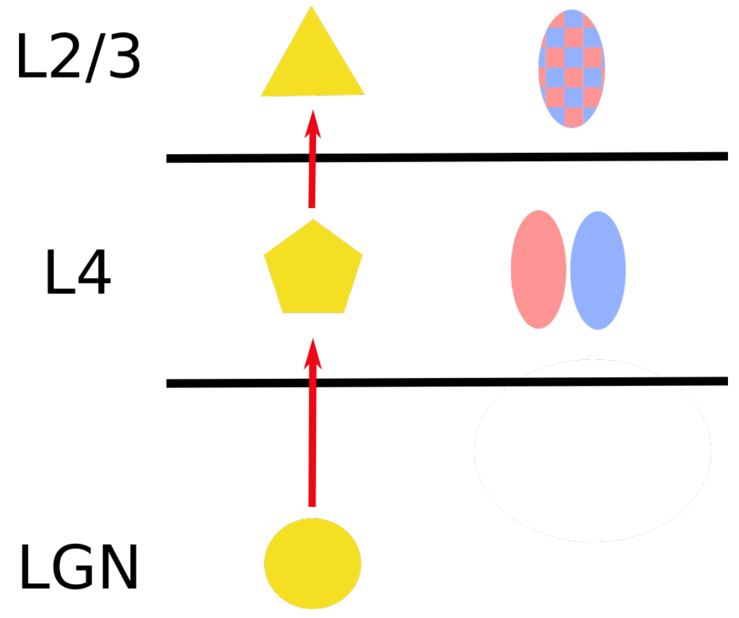
(Jin et al, 2011)

The structural priors

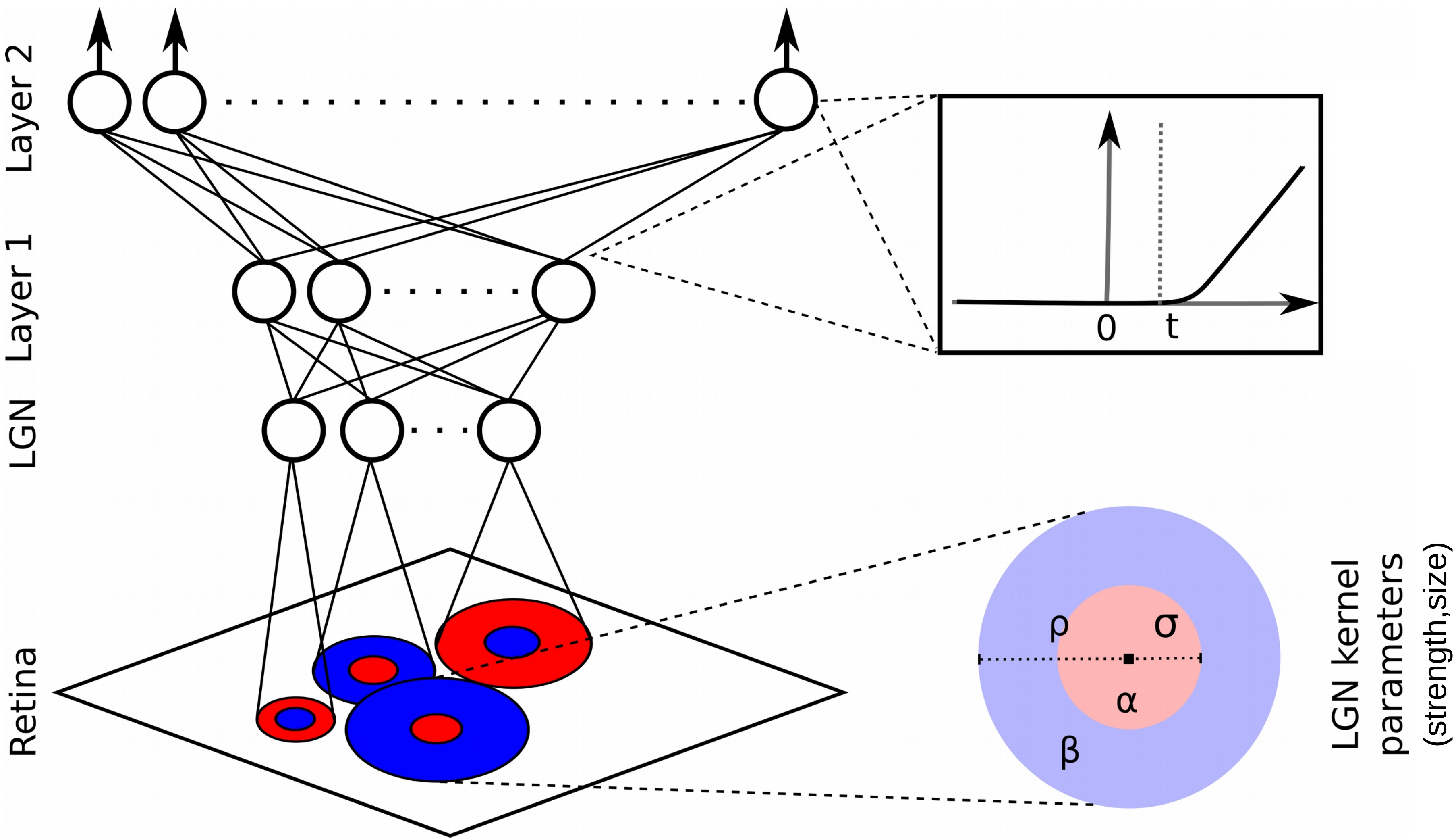
- Receptive fields of LGN units can be well approximated by difference-of-Gaussian function

- Local population of V1 neurons receives common input from limited number of LGN cells

- Hierarchical organization



The HSM structure



The model

LGN units:

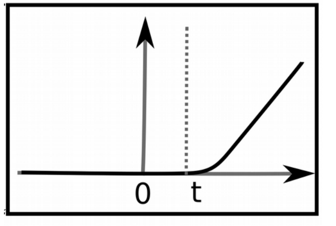
$$\psi_{i1} = \sum_{k,l} I_{kl} \left(\frac{\alpha_i}{\sigma_i^2} e^{-\frac{\left(k-\mu_i^x\right)^2 + \left(l-\mu_i^y\right)^2}{2\sigma_i^2}} - \frac{\beta_i}{\rho_i^2} e^{-\frac{\left(k-\mu_i^x\right)^2 + \left(l-\mu_i^y\right)^2}{2\rho_i^2}} \right)$$

Cortical units:

$$\psi_{il} = f\left(\sum_j w_{ij} \psi_{j(l-1)} \right)$$

Transfer function:

$$\mathrm{f}(\mathrm{x}) = \log\left(1 + \exp\left(\mathrm{x} - \mathrm{t}_\mathrm{i}\right)\right)$$



Log-likelyhood:

$$\log p\left(y|x,\phi\right) = \sum_i y_i \log M\left(\phi, x_i\right) - \sum_i M\left(\phi, x_i\right)$$

Model optimization

- Optimized with Constrained Truncated Newton Conjugate method
- Non-convex model
  - 100 restarts with different seeds of initial random parameter initialization
  - pick the best fit to **training** data
- Meta-parameters:
  - Number of LGN units (9)
  - Number of hidden units (20%)
  - Determined based on prior 1D searches based on **training** data performance

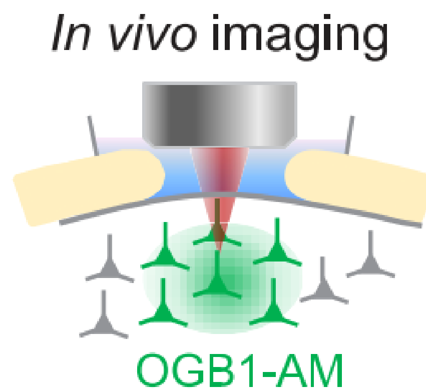
Calcium imaging of local population of neurons in mouse V1

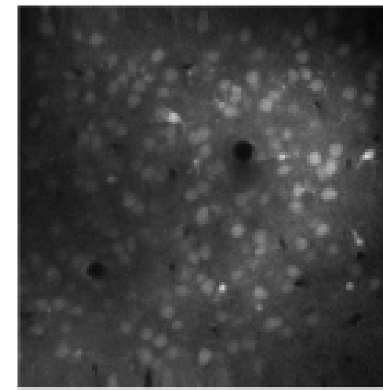
2 mice, 30–40 postnatal day

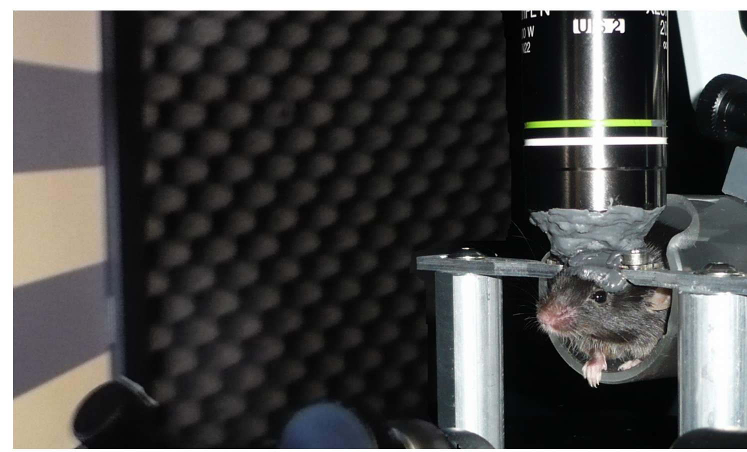
Anesthesized: isoflurane

3 imaged regions

OGB1–AM calcium indicator

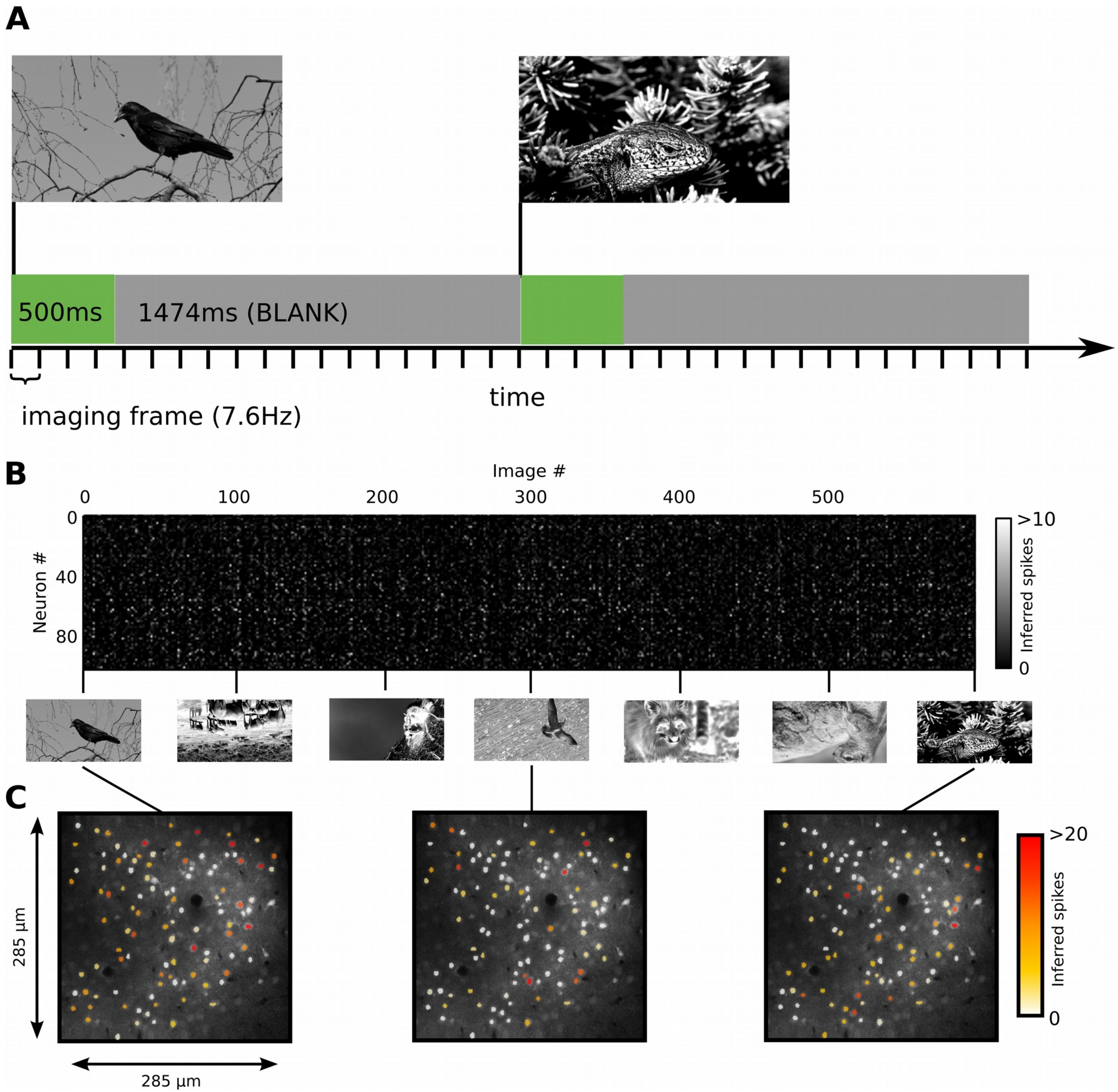


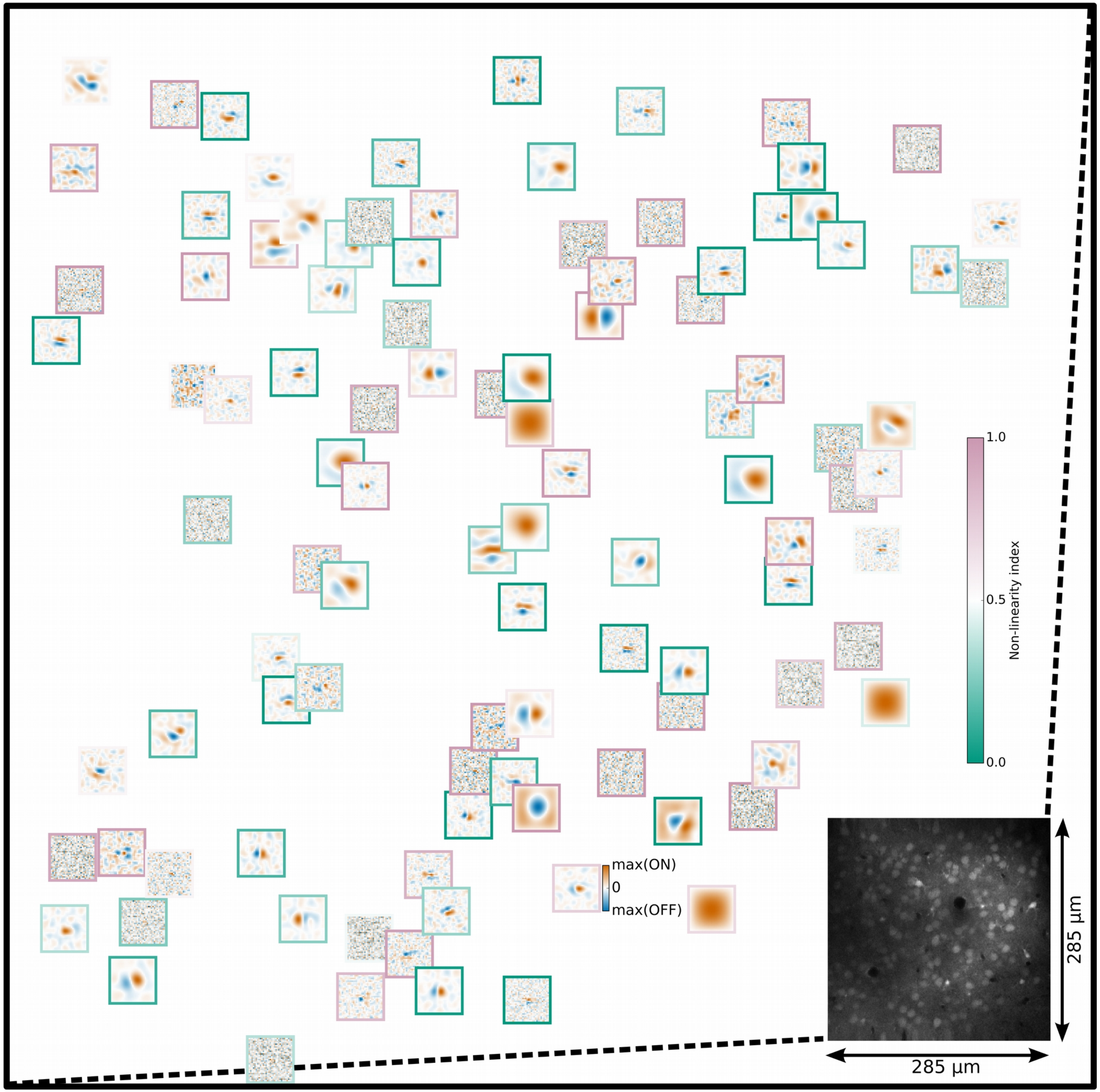




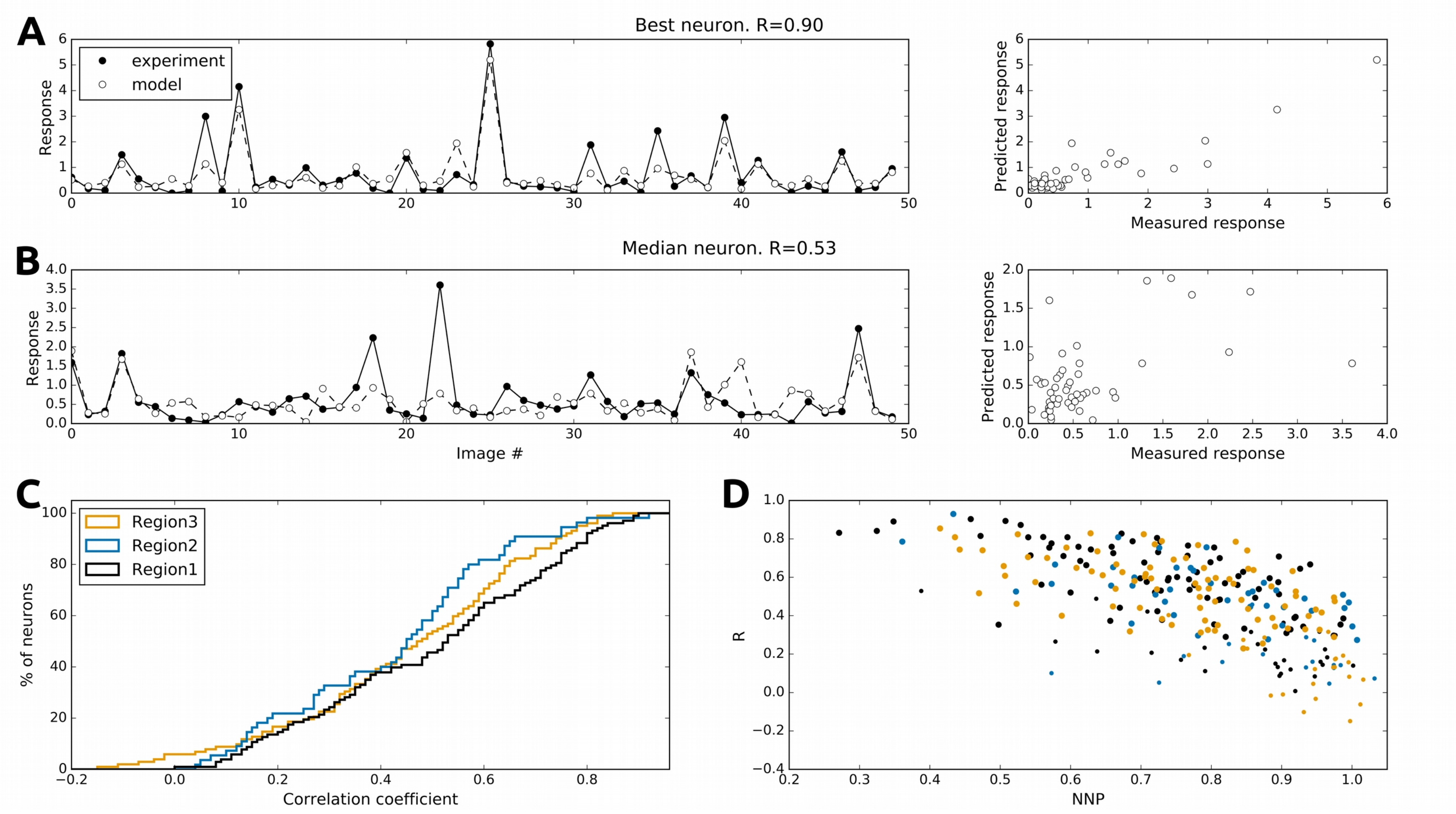
Sonja Hofer

Thomas Mrsic-Flogel


A
500ms
1474ms (BLANK)
time
imaging frame (7.6Hz)
B
Image #
0
100
200
300
400
500
Neuron #
0
40
80
>10
Inferred spikes
0
C
285 µm
285 µm
>20
Inferred spikes
0


1.0
0.5
0.0
Non-linearity index
max(ON)
0
max(OFF)
285 µm
285 µm

The model performance



Comparison: reference models

STA with laplacian regularization
(Smyth et. al, Journal of Neuroscince, 2003)

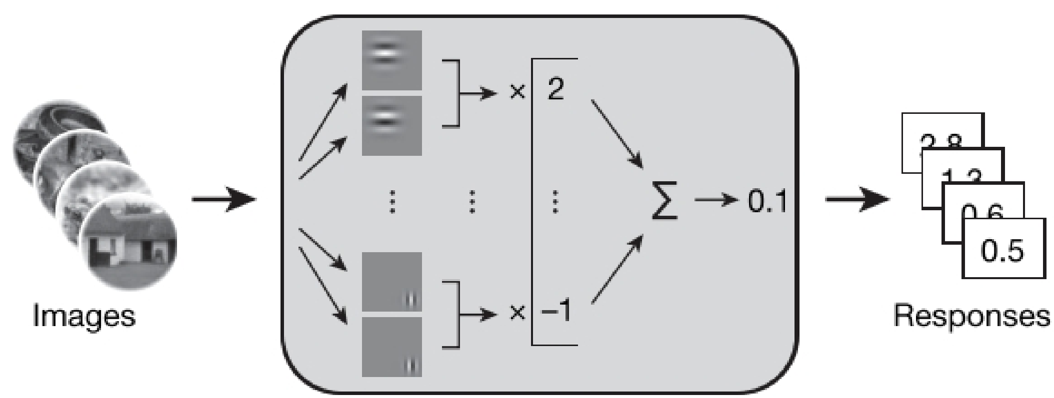
$$\mathbf{L_s} = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}. \quad \begin{bmatrix} \mathbf{S} \\ \lambda \mathbf{L} \end{bmatrix} \mathbf{f} = \begin{bmatrix} r \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

Comparison: reference models

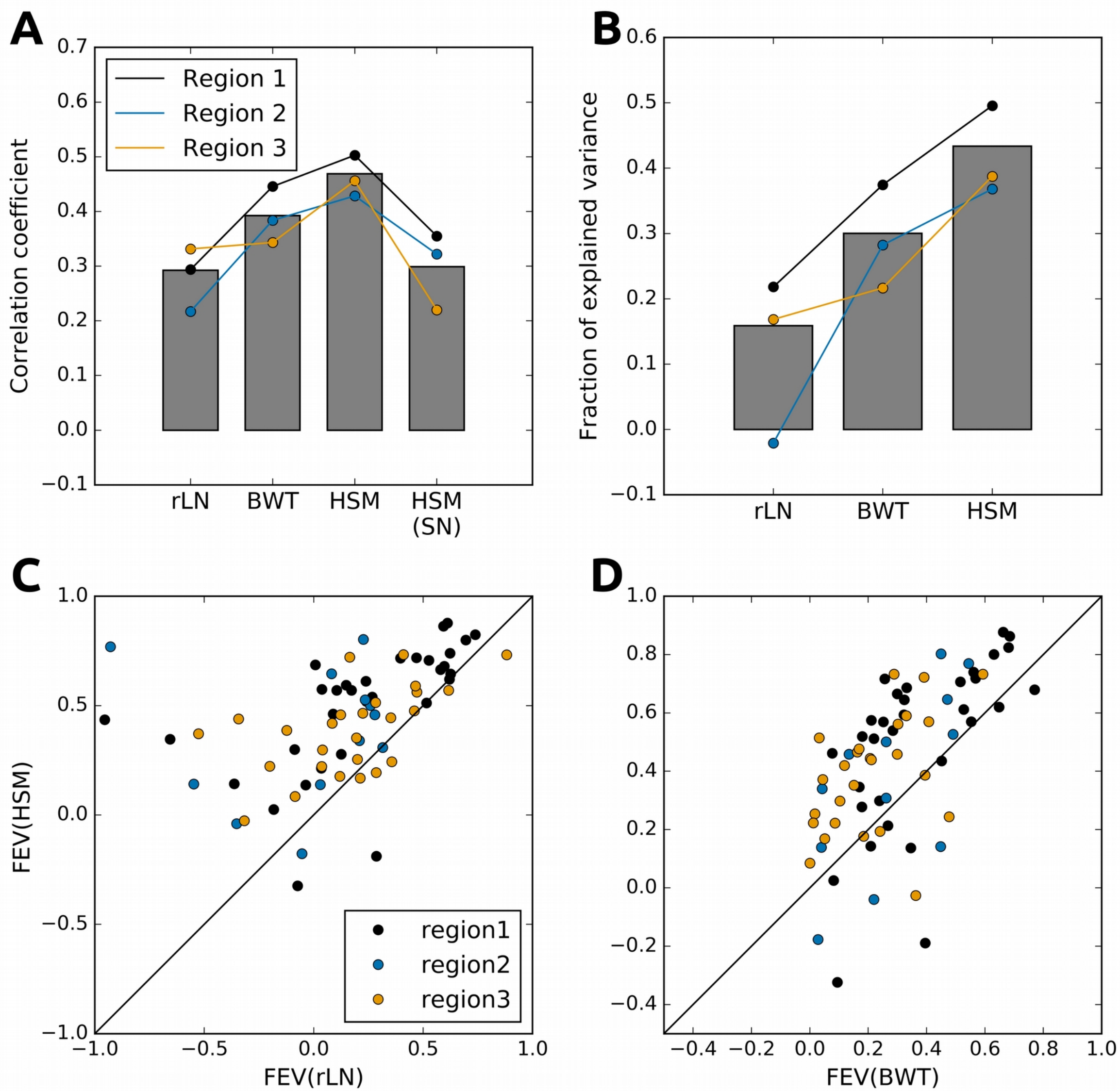
STA with laplacian regularization
(Smyth et. al, Journal of Neuroscince, 2003)

$$\mathbf{L}_s = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}. \qquad \begin{bmatrix} \mathbf{S} \\ \lambda \mathbf{L} \end{bmatrix} \mathbf{f} = \begin{bmatrix} r \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

Barkely-wavelet transform based linear model
(Kay et. al, Nature, 2008)

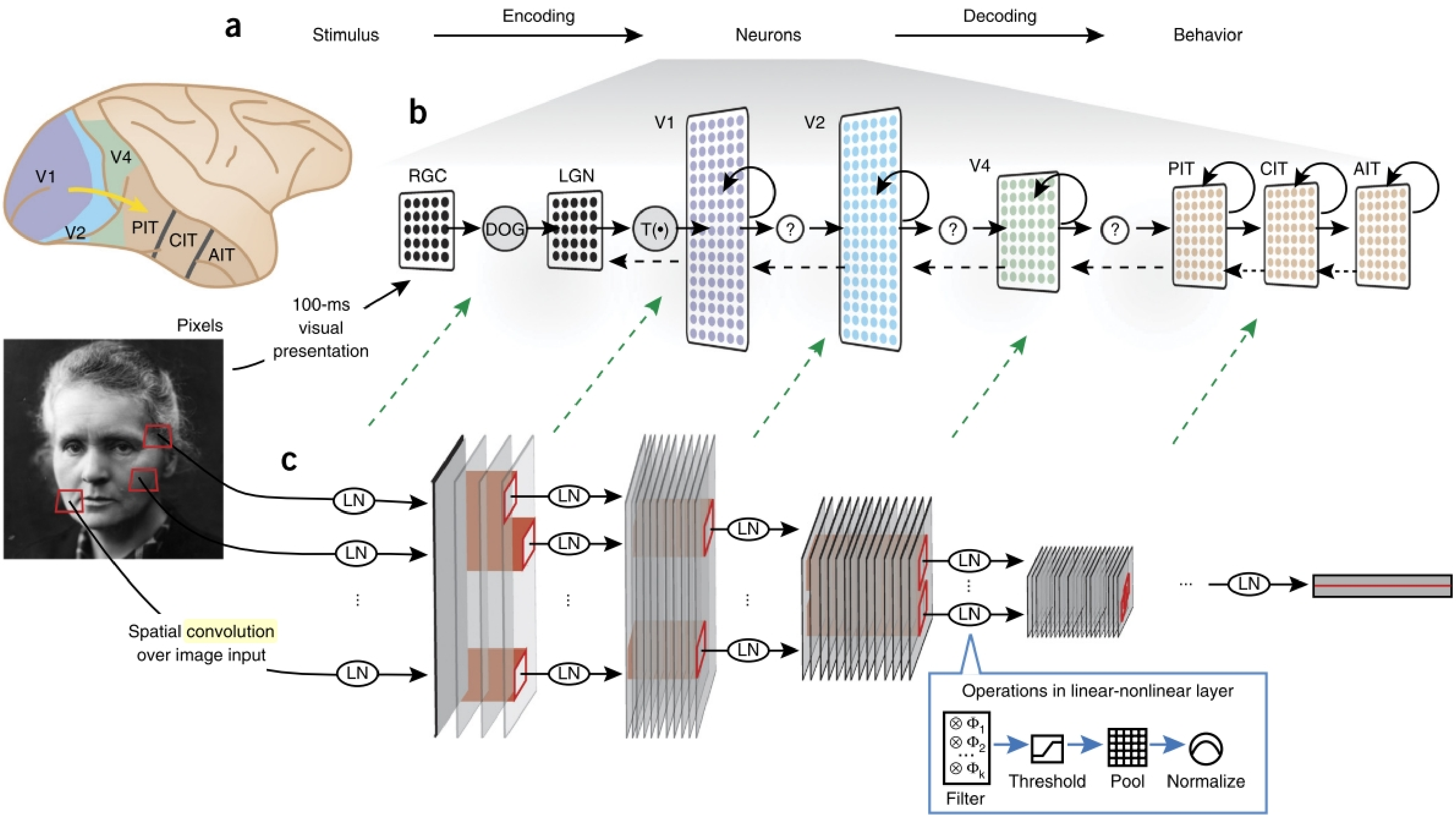


Comparison: performance

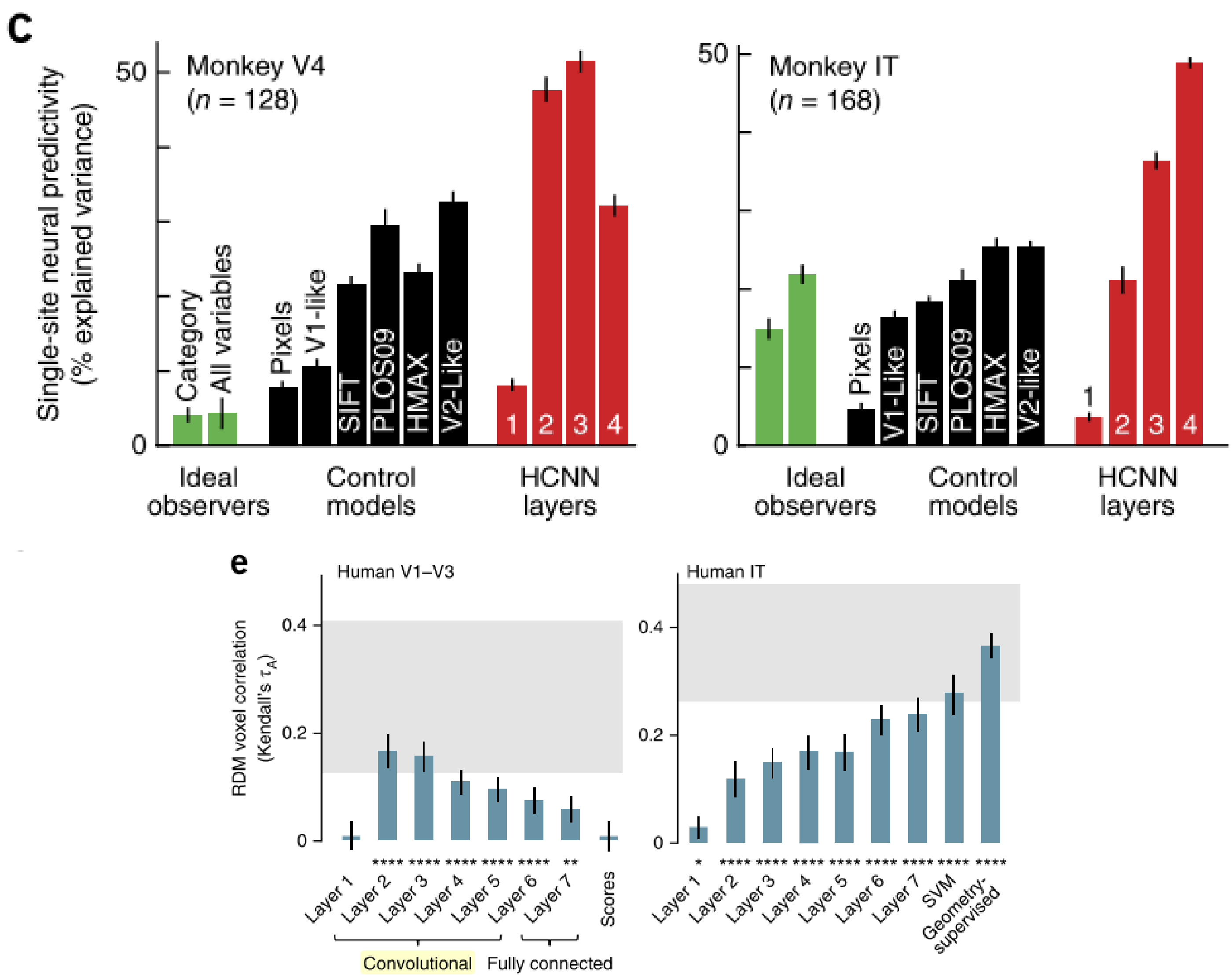


Deep Neural Networks In Neuroscience

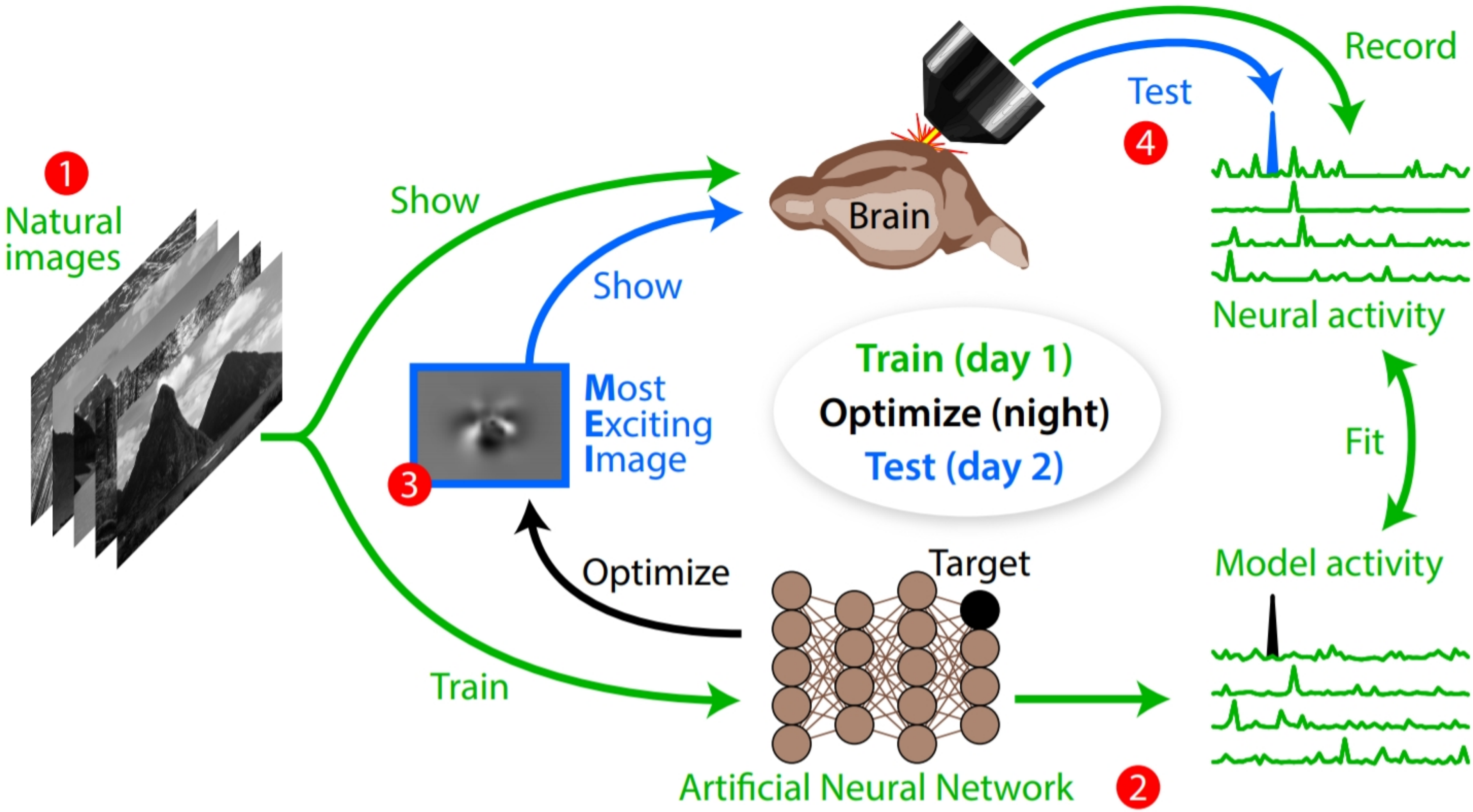
Using goal-driven deep learning models to understand sensory cortex



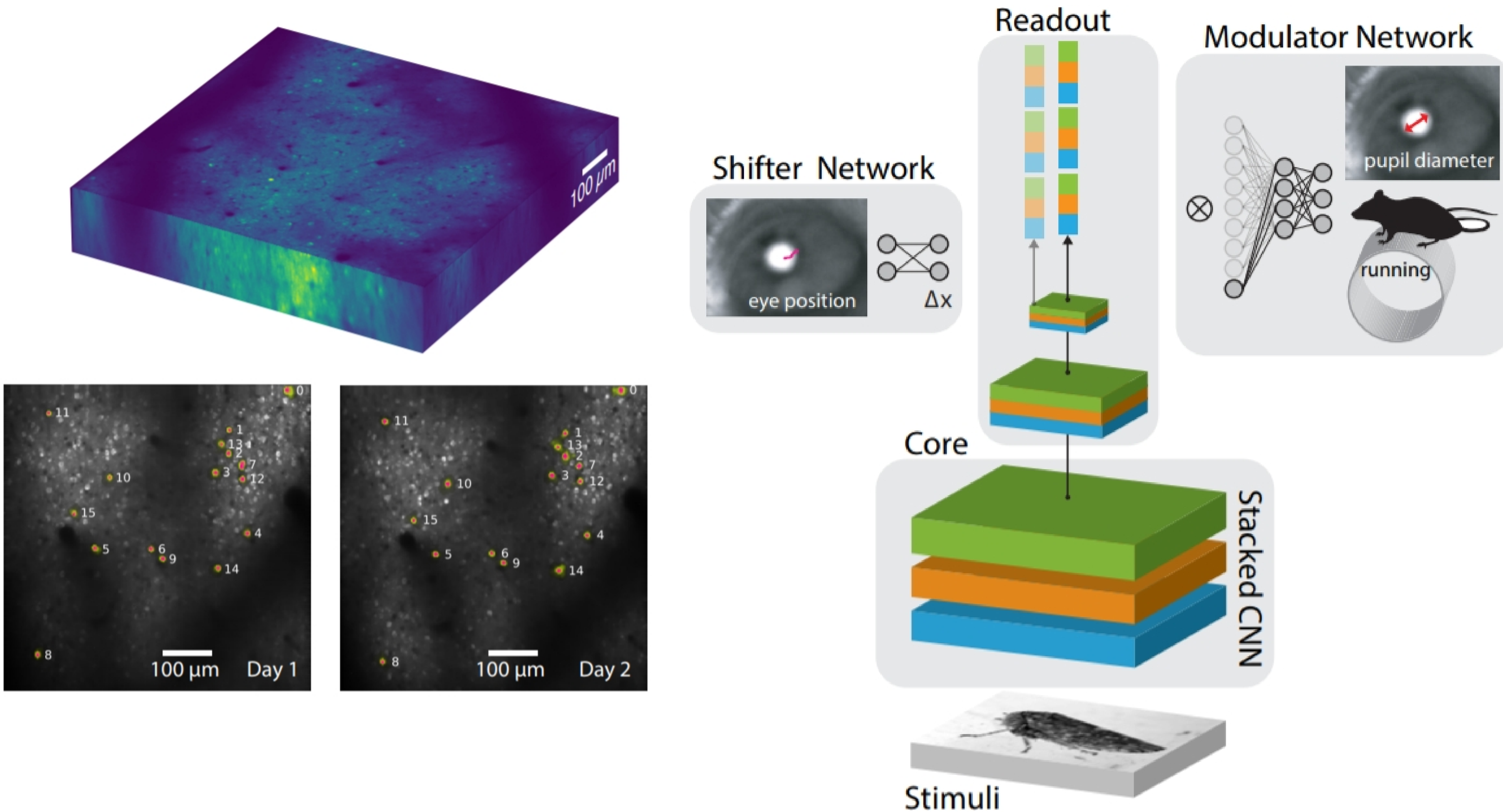
Using goal-driven deep learning models to understand sensory cortex



Inception in visual cortex: in vivo–silico loops reveal most exciting images



Inception in visual cortex: in vivo–silico loops reveal most exciting images



Inception in visual cortex: in vivo–silico loops reveal most exciting images