

Neuroscience & big-data:

Collective behavior in neuronal ensembles

Ulisse Ferrari



Research career:



SAPIENZA
UNIVERSITÀ DI ROMA

PhD in Theor. Physics of
disordered systems (G. Parisi)



ENS
ÉCOLE NORMALE
SUPÉRIEURE

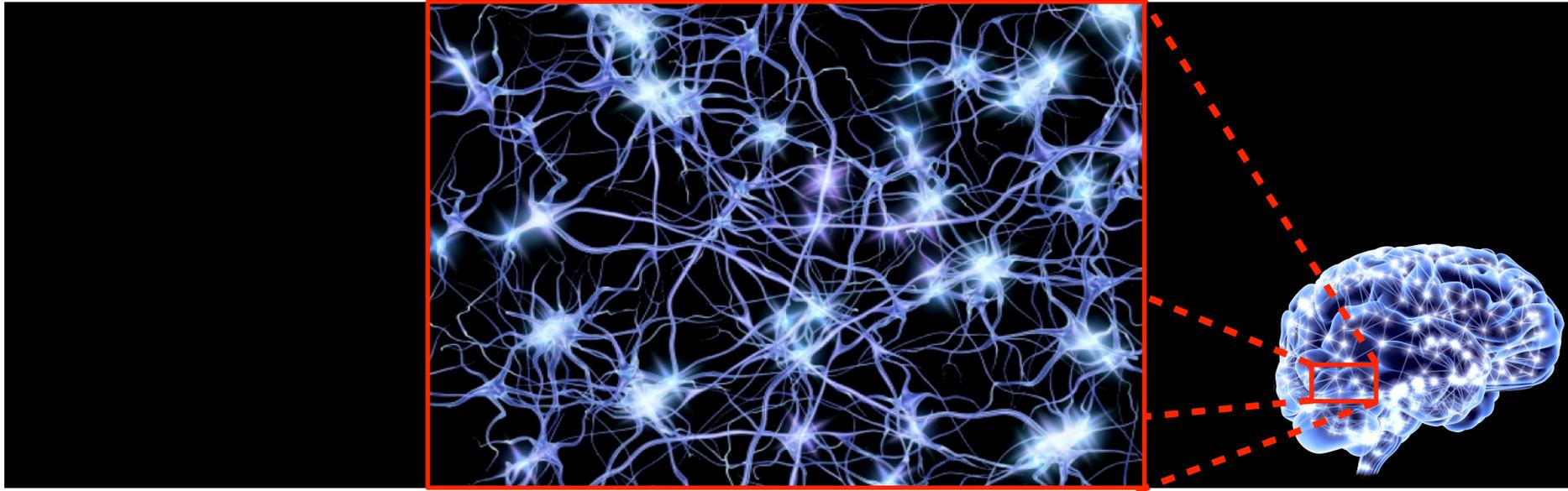
Inverse problems:
neuronal collective behavior



INSTITUT DE
LA VISION
★ PARIS

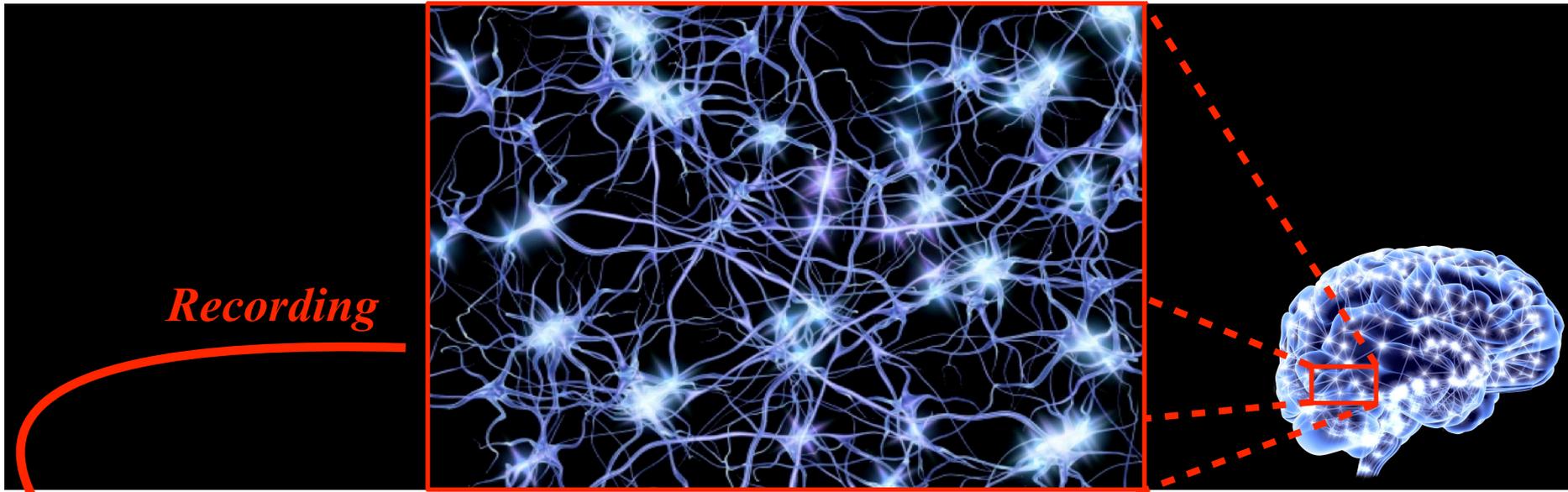
Contact with experiments and
retinal collective encoding

Collective behavior in neuronal ensembles

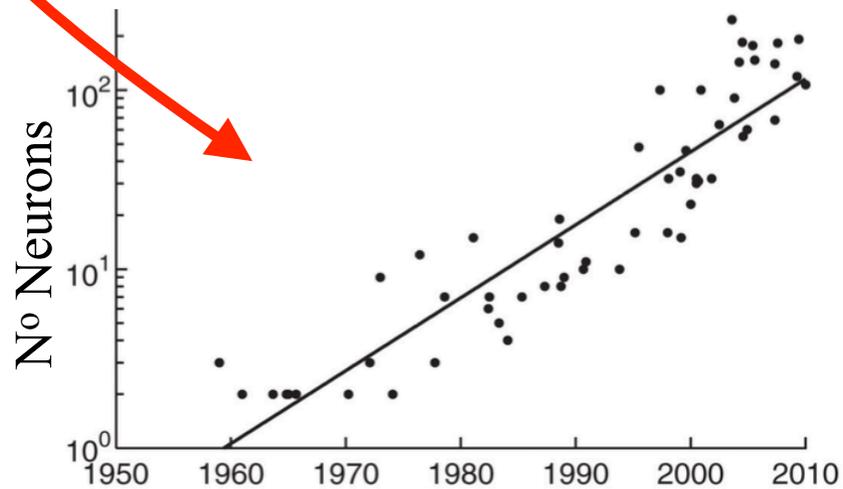


Human: $\sim 10^3$ synapses/neuron

Collective behavior in neuronal ensembles

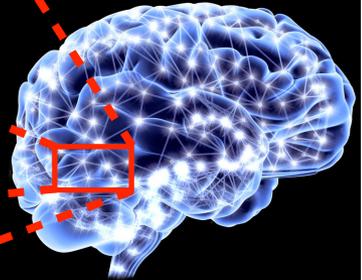
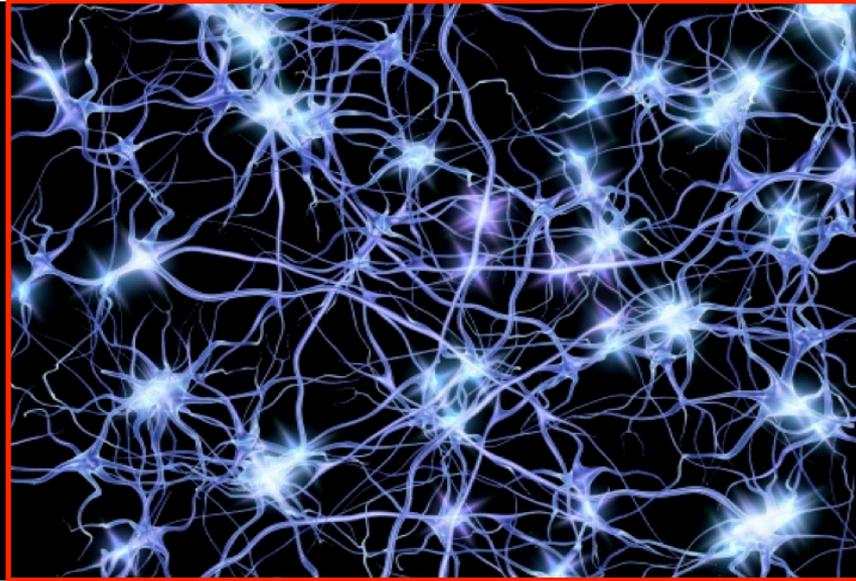


Human: $\sim 10^3$ synapses/neuron

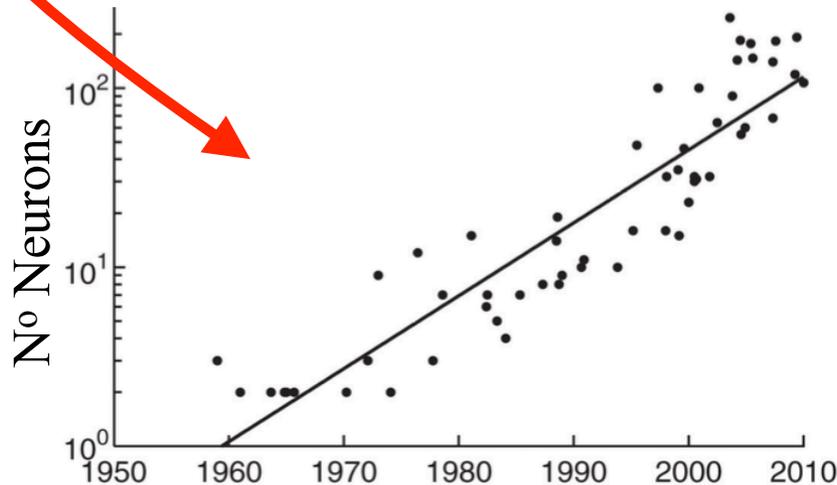


Collective behavior in neuronal ensembles

Recording

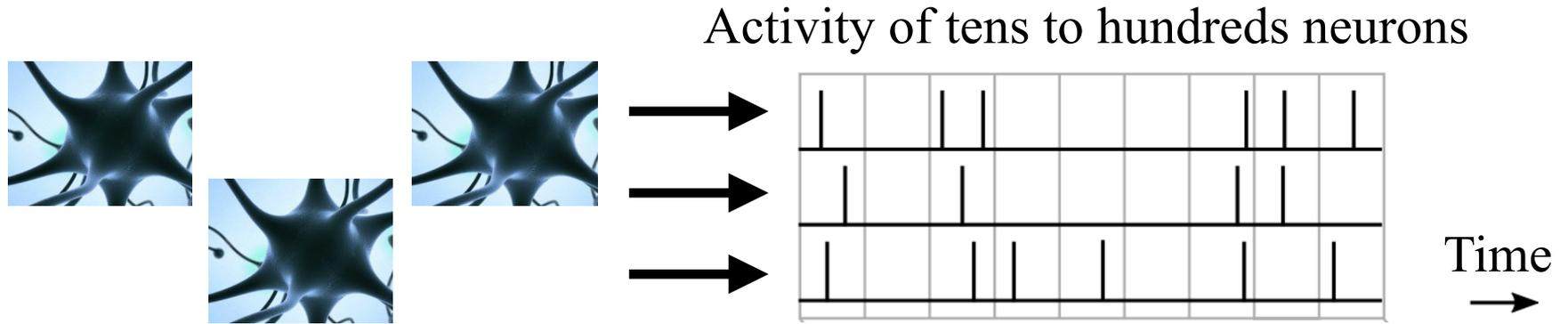


Human: $\sim 10^3$ synapses/neuron

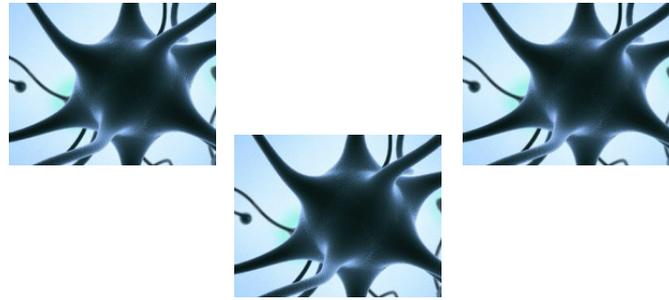


Challenge:
*develop new
concepts and tools
to make sense
of network activity*

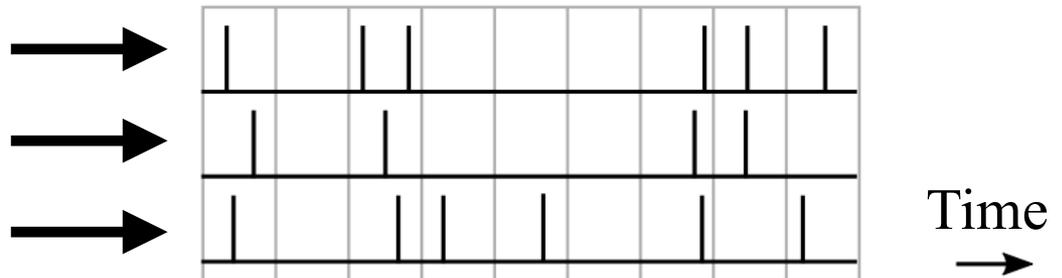
How to characterise neuronal collective activities?



How to characterise neuronal collective activities?



Activity of tens to hundreds neurons



ΔT

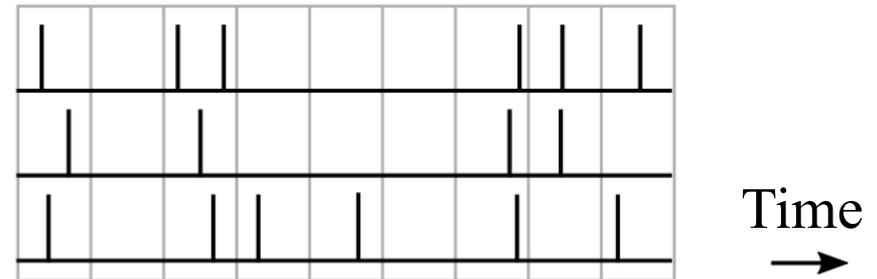
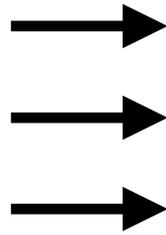
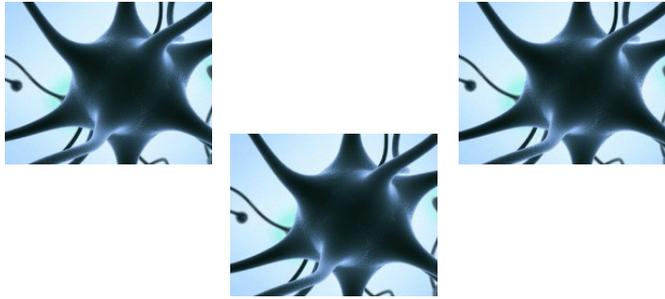
σ_1	:	1	0	1	0	0	0	1	1	1
σ_2	:	1	0	1	0	0	0	1	1	0
σ_3	:	0	0	1	1	1	0	1	0	1

$\vec{\sigma} \sim 10^5$ samples

$$P(\vec{\sigma}) \propto ?$$

How to characterise neuronal collective activities?

Activity of tens to hundreds neurons



Maximum Entropy Principle

$$P(\vec{\sigma}) \propto \exp \left(\sum_{i < j} J_{ij} \sigma_i \sigma_j + \sum_i h_i \sigma_i \right)$$

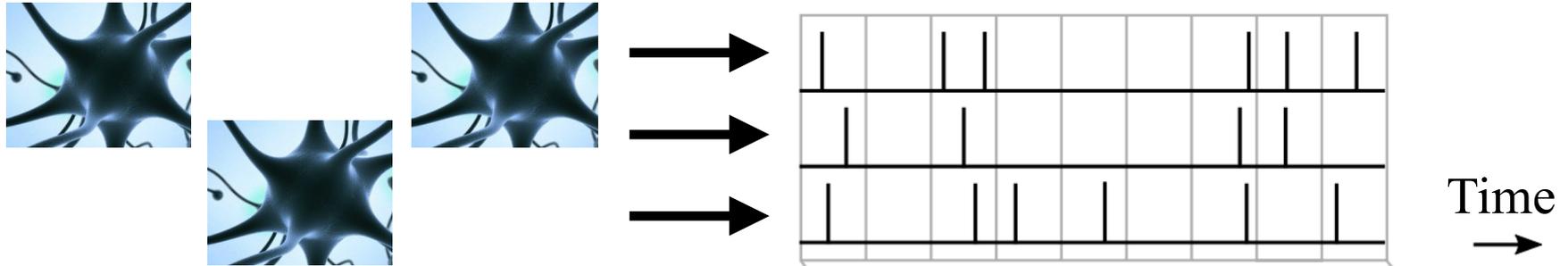
σ_1 :	1	0	1	0	0	0	1	1	1
σ_2 :	1	0	1	0	0	0	1	1	0
σ_3 :	0	0	1	1	1	0	1	0	1

$\vec{\sigma} \sim 10^5$ samples

Exponential model on interaction graph

How to characterise neuronal collective activities?

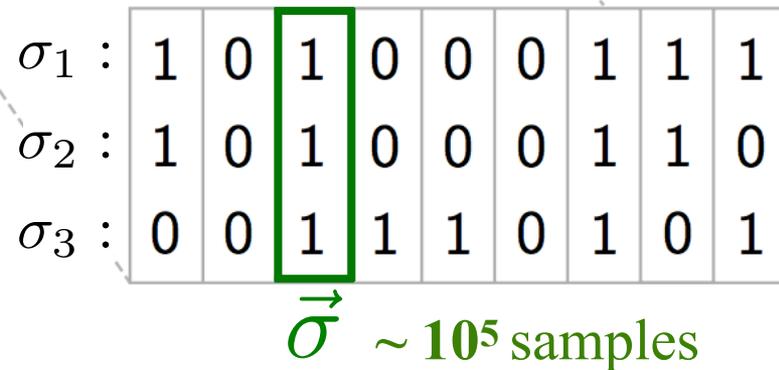
Activity of tens to hundreds neurons



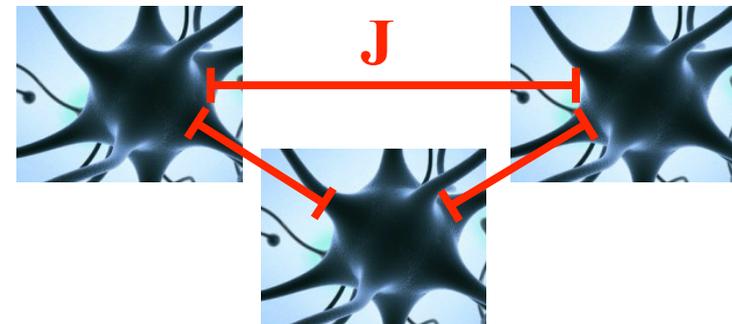
Maximum Entropy Principle

$$P(\vec{\sigma}) \propto \exp \left(\sum_{i < j} J_{ij} \sigma_i \sigma_j + \sum_i h_i \sigma_i \right)$$

Exponential model on interaction graph



Interaction model



Maximum entropy models

Choose sufficient statistics: $T_a(\vec{\sigma})$

Find the largest entropy $P(\vec{\sigma})$ that reproduces their averages:

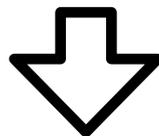
$$\min_{\mathbf{J}, \lambda} \max_{P(\vec{\sigma})} \left[- \sum_{\vec{\sigma}} P(\vec{\sigma}) \log P(\vec{\sigma}) + \lambda \left(\sum_{\vec{\sigma}} P(\vec{\sigma}) - 1 \right) + \sum_a J_a (Q_a - \overline{P}_a) \right]$$
$$\overline{P}_a \equiv \langle T_a(\vec{\sigma}) \rangle_{\text{data}} \quad Q_a \equiv \langle T_a(\vec{\sigma}) \rangle_P$$

Maximum entropy models

Choose sufficient statistics: $T_a(\vec{\sigma})$

Find the largest entropy $P(\vec{\sigma})$ that reproduces their averages:

$$\min_{\mathbf{J}, \lambda} \max_{P(\vec{\sigma})} \left[- \sum_{\vec{\sigma}} P(\vec{\sigma}) \log P(\vec{\sigma}) + \lambda \left(\sum_{\vec{\sigma}} P(\vec{\sigma}) - 1 \right) + \sum_a \mathbf{J}_a (Q_a - \overline{P}_a) \right]$$
$$\overline{P}_a \equiv \langle T_a(\vec{\sigma}) \rangle_{\text{data}} \quad Q_a \equiv \langle T_a(\vec{\sigma}) \rangle_P$$



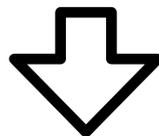
$$P(\vec{\sigma} | \mathbf{J}) = g(\mathbf{J}) \exp(\mathbf{J} \cdot T(\vec{\sigma}))$$

Maximum entropy models

Choose sufficient statistics: $T_a(\vec{\sigma})$

Find the largest entropy $P(\vec{\sigma})$ that reproduces their averages:

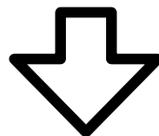
$$\min_{\mathbf{J}, \lambda} \max_{P(\vec{\sigma})} \left[- \sum_{\vec{\sigma}} P(\vec{\sigma}) \log P(\vec{\sigma}) + \lambda \left(\sum_{\vec{\sigma}} P(\vec{\sigma}) - 1 \right) + \sum_a \mathbf{J}_a (Q_a - \overline{P}_a) \right]$$
$$\overline{P}_a \equiv \langle T_a(\vec{\sigma}) \rangle_{\text{data}} \quad Q_a \equiv \langle T_a(\vec{\sigma}) \rangle_P$$



$$P(\vec{\sigma} | \mathbf{J}) = g(\mathbf{J}) \exp(\mathbf{J} \cdot T(\vec{\sigma}))$$

----- *example* -----

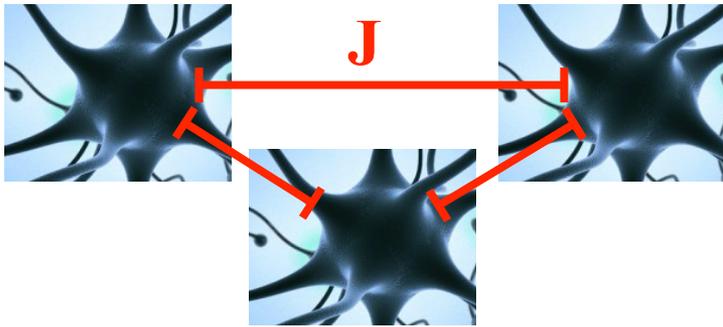
$$\langle \sigma_i \rangle_{\text{Md}} = \langle \sigma_i \rangle_{\text{Dt}} \quad \& \quad \langle \sigma_i \sigma_j \rangle_{\text{Md}} = \langle \sigma_i \sigma_j \rangle_{\text{Dt}}$$



$$P(\vec{\sigma}) \propto \exp \left(\sum_{i < j} \mathbf{J}_{ij} \sigma_i \sigma_j + \sum_i h_i \sigma_i \right)$$

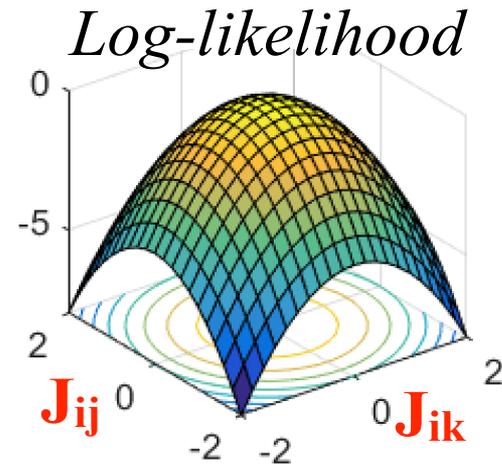
How to learn these models on neuronal data?

$$P(\vec{\sigma}) \propto \exp \left(\sum_{i < j} J_{ij} \sigma_i \sigma_j + \sum_i h_i \sigma_i \right)$$



Exponentially hard problem

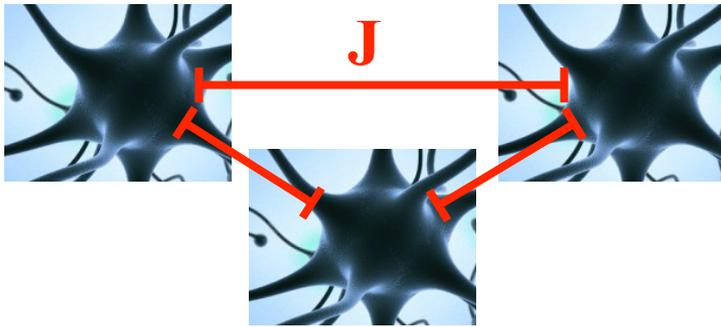
- Exponential family
- Large datasets



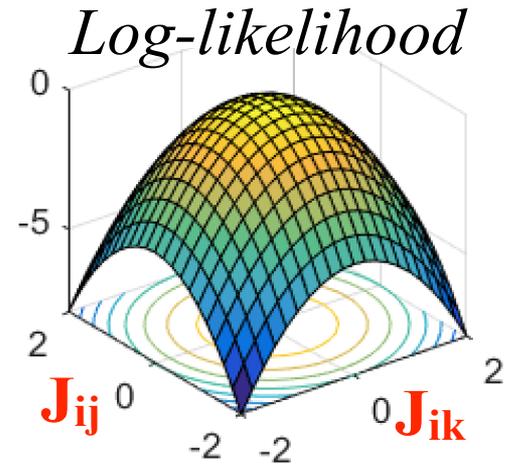
How to learn these models on neuronal data?

$$P(\vec{\sigma}) \propto \exp \left(\sum_{i < j} J_{ij} \sigma_i \sigma_j + \sum_i h_i \sigma_i \right)$$

- Exponential family
- Large datasets



Exponentially hard problem

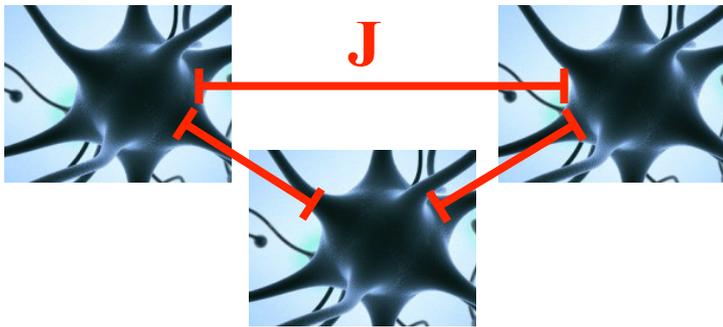


Exp. Family $\longrightarrow \mathcal{H}[J] = \text{Cov}_J[\text{sufficient stat.}]$

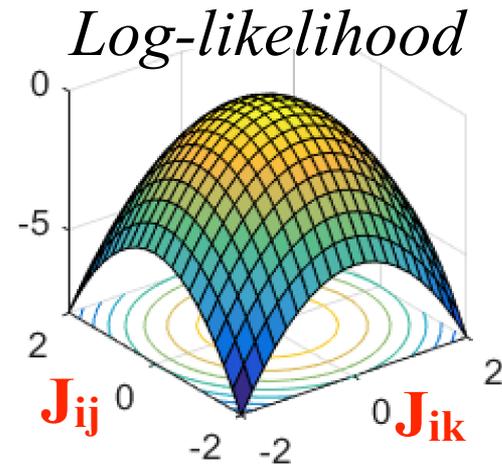
How to learn these models on neuronal data?

$$P(\vec{\sigma}) \propto \exp \left(\sum_{i < j} J_{ij} \sigma_i \sigma_j + \sum_i h_i \sigma_i \right)$$

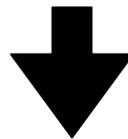
- Exponential family
- Large datasets



Exponentially hard problem



$$\text{Exp. Family} \longrightarrow \mathcal{H}[J] = \text{Cov}_J[\text{sufficient stat.}]$$



$$\mathcal{H}[J^{\text{sol.}}] \approx \text{Cov}^{\text{data}}[\text{sufficient stat.}]$$

Newton method without additional cost

Collective behavior in neuronal ensembles

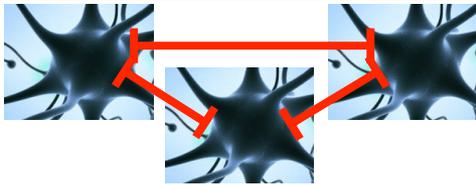
Max-Entropy

$$H[P] = - \sum P \log P$$

Neuronal network



Interaction Model



Collective behavior in neuronal ensembles

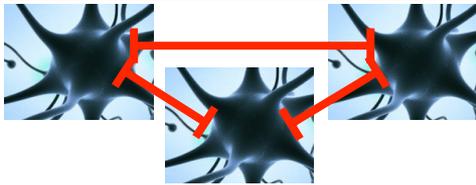
Max-Entropy

$$H[P] = - \sum P \log P$$

Neuronal network

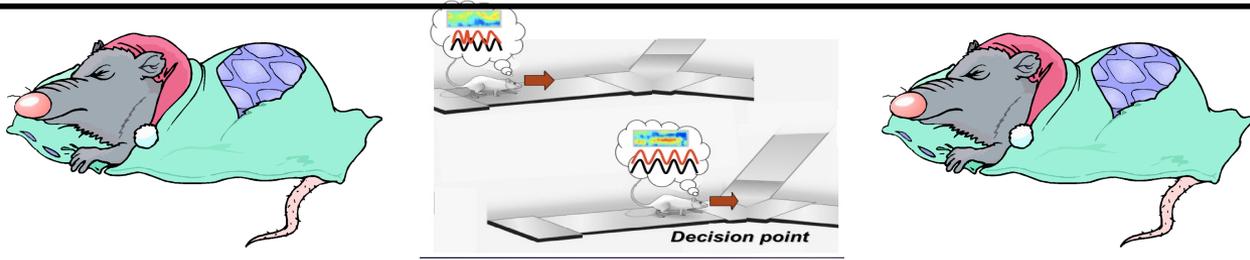


Interaction Model

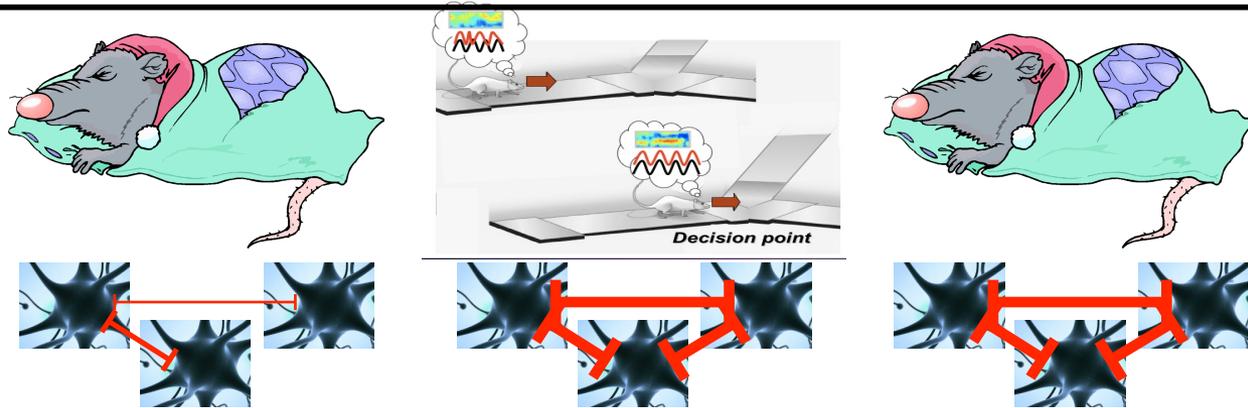


Are these models useful for biology?

What's the neuronal basis (engram) of memory?

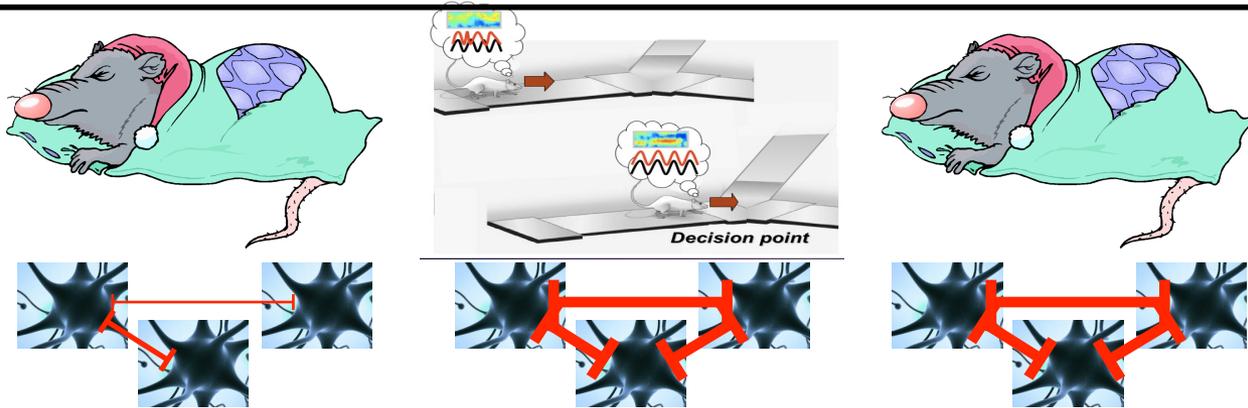


What's the neuronal basis (engram) of memory?

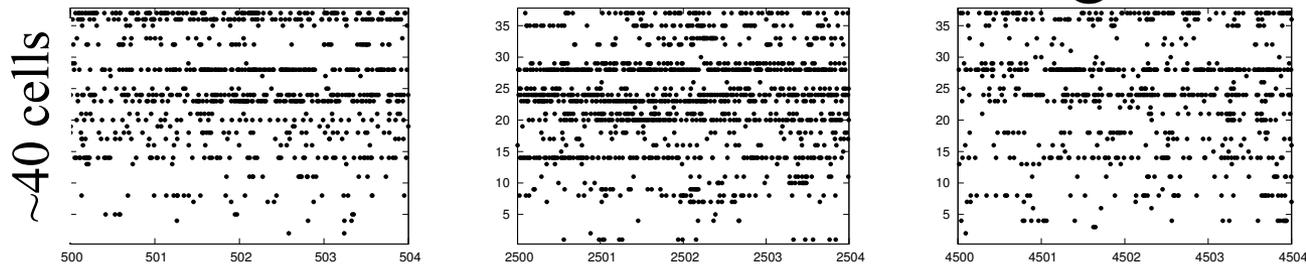


Hebb's Hypothesis:
reinforcement of interactions during learning

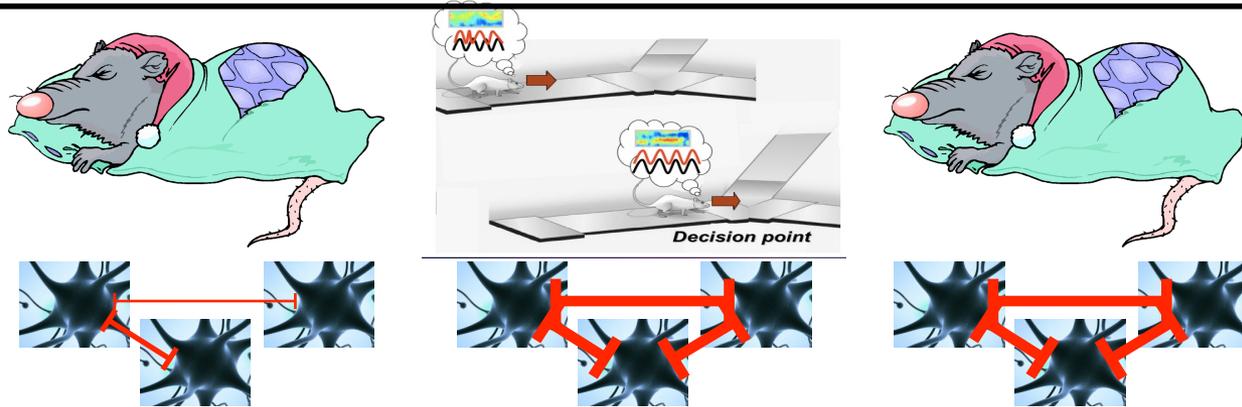
What's the neuronal basis (engram) of memory?



Hebb's Hypothesis:
reinforcement of interactions during learning

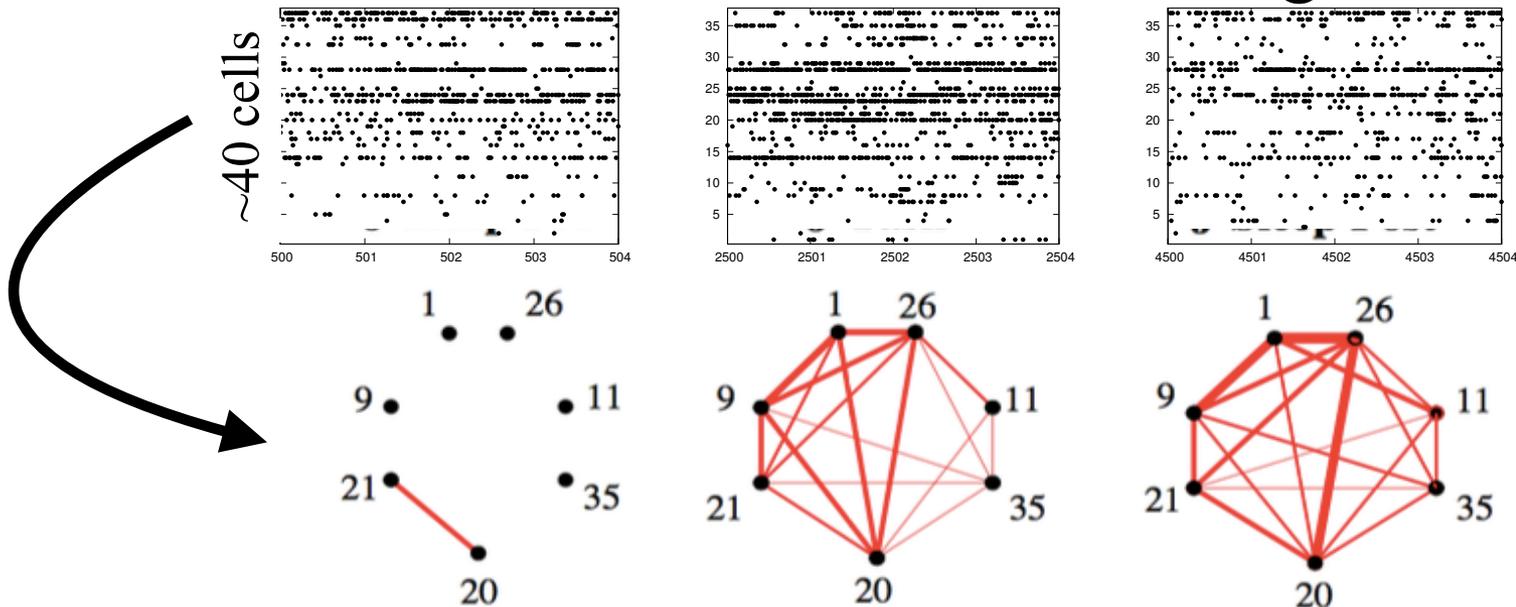


What's the neuronal basis (engram) of memory?



Hebb's Hypothesis:

reinforcement of interactions during learning



Collective behavior in neuronal ensembles

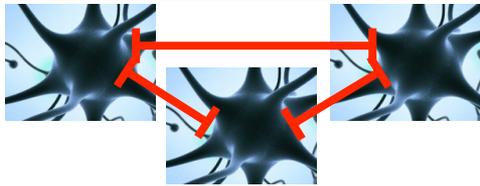
Max-Entropy

$$H[P] = - \sum P \log P$$

Neuronal network



Interaction Model



Effective approach:

- Capable of unveiling **functional** coupling structure

Collective behavior in neuronal ensembles

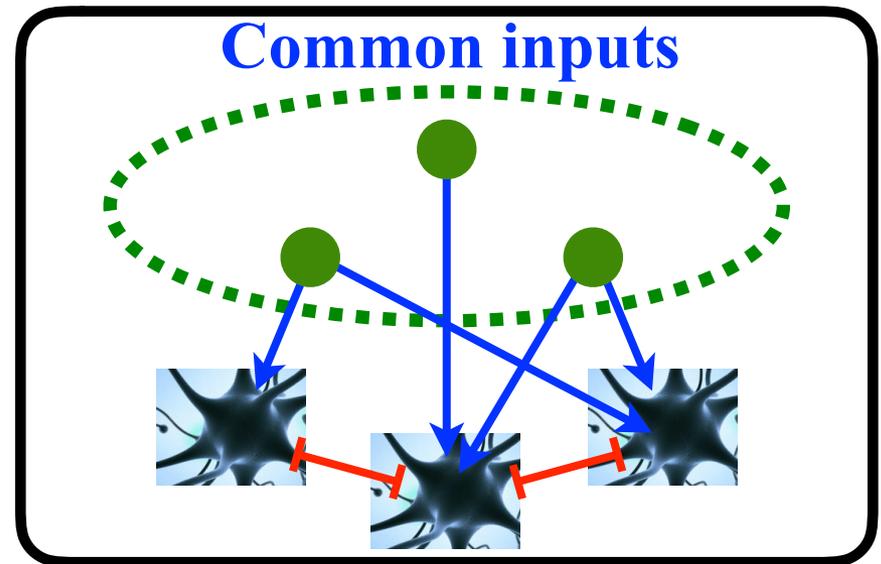
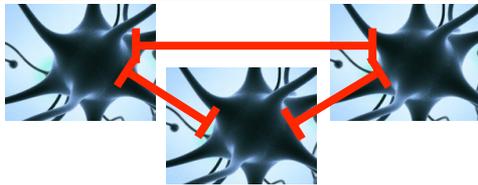
Max-Entropy

$$H[P] = - \sum P \log P$$

Neuronal network



Interaction Model



Effective approach:

- Capable of unveiling **functional** coupling structure
- Couplings reflects **common inputs** & **real connections**

Can I disentangle the two contributions?

Collective behavior in neuronal ensembles

Max-Entropy

$$H[P] = - \sum P \log P$$

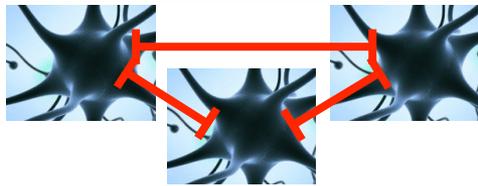
Neuronal network



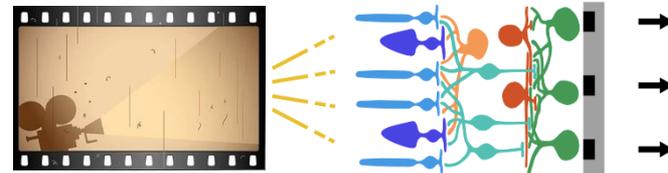
Visual Information



Interaction Model



Stimulus processing

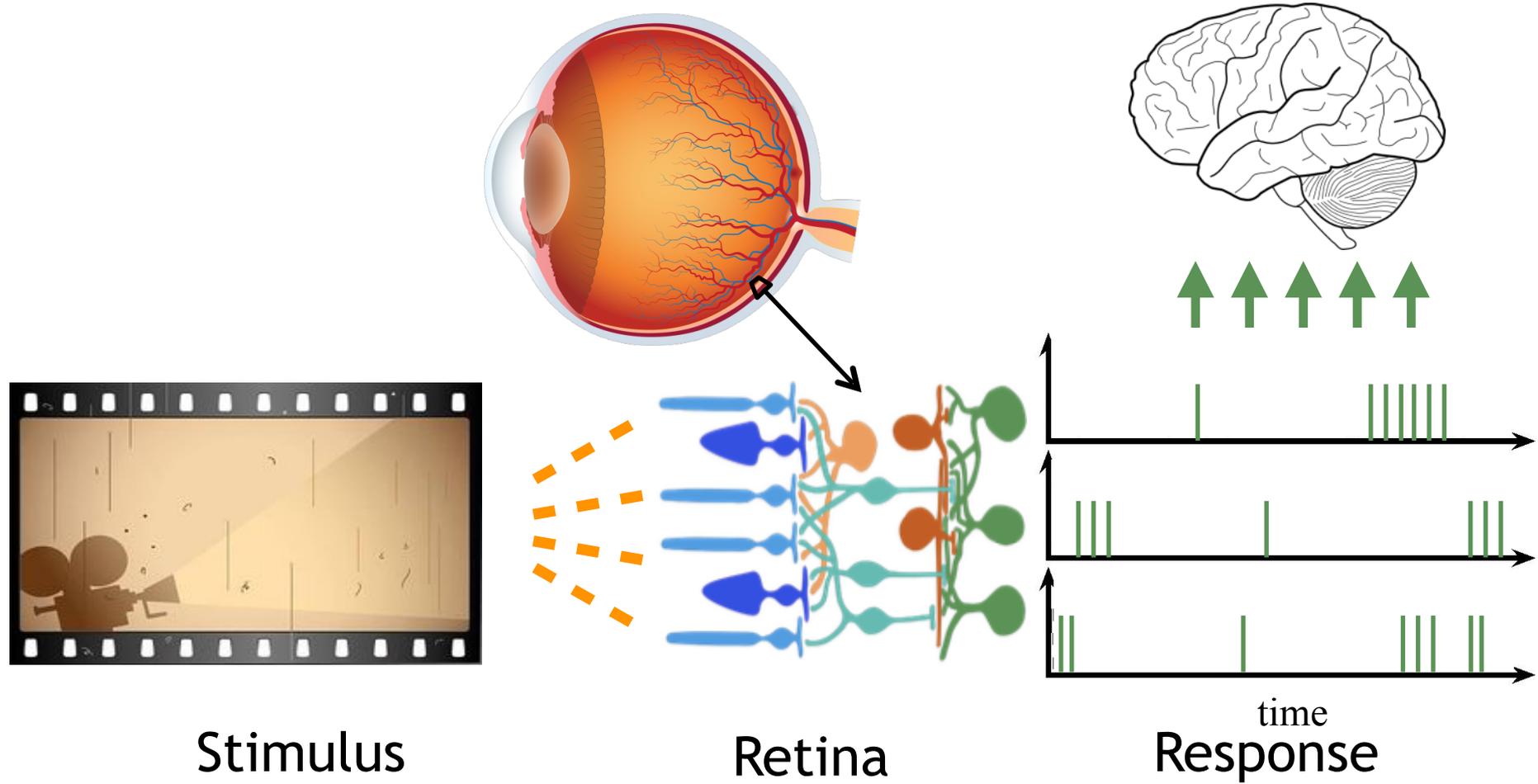


Effective approach:

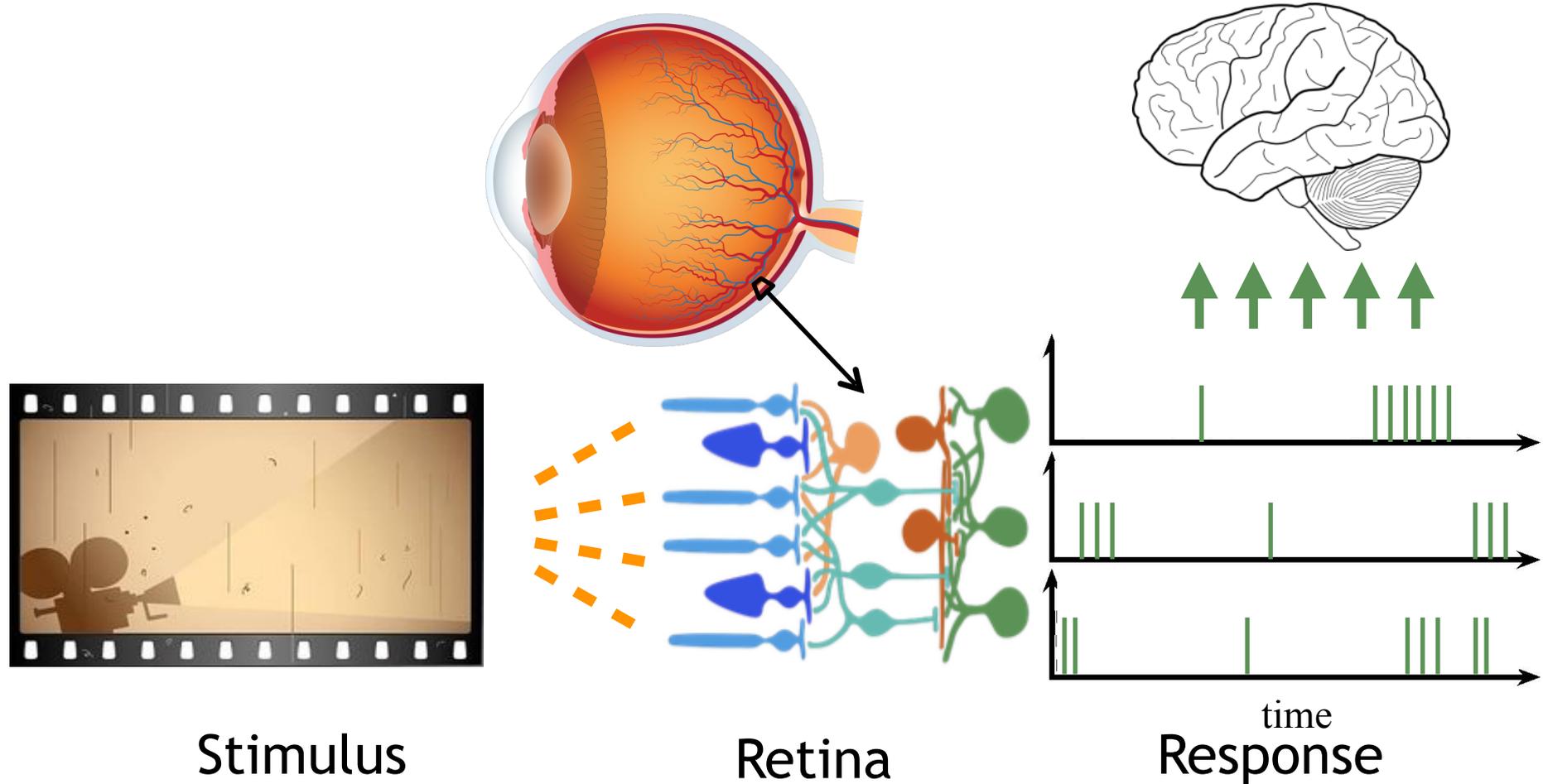
- Capable of unveiling **functional** coupling structure
- Couplings reflects **common inputs** & **real connections**

Can I disentangle the two contributions?

Stimulus processing in the retina

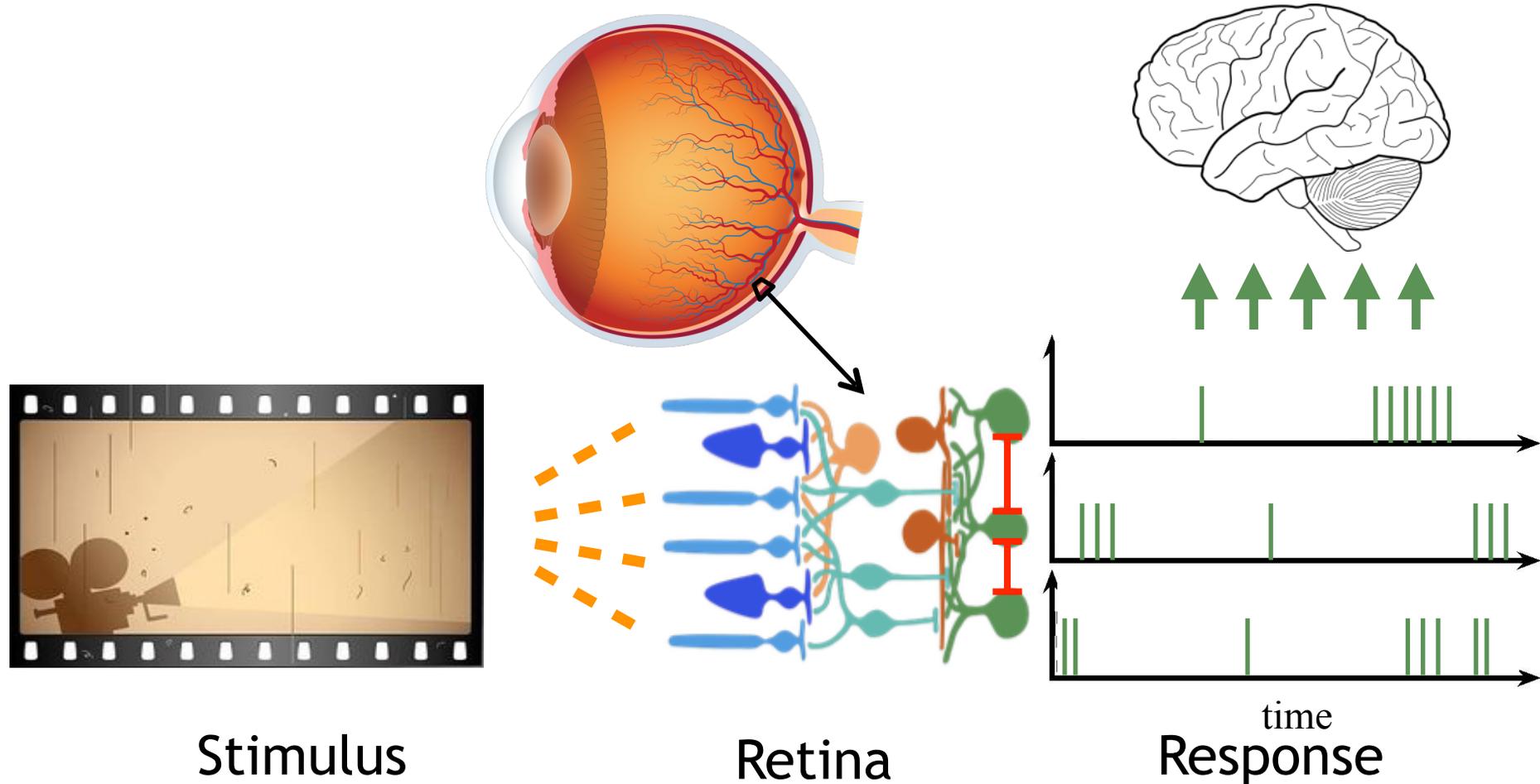


Stimulus processing in the retina



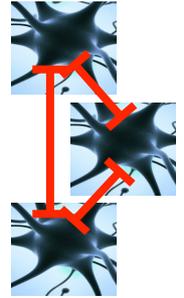
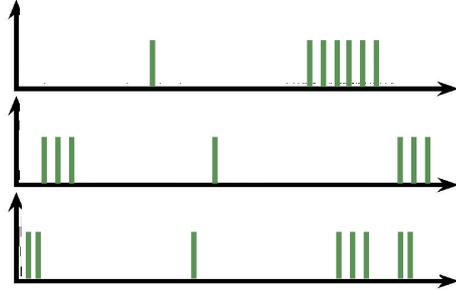
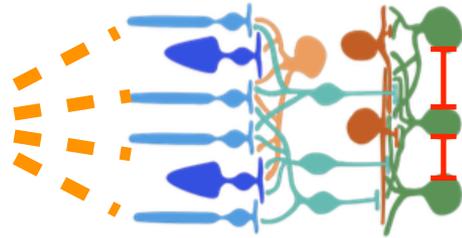
Non-linear & stochastic **feedforward system**,

Stimulus processing in the retina

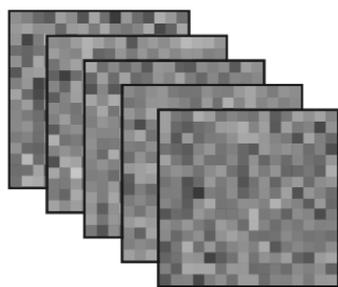
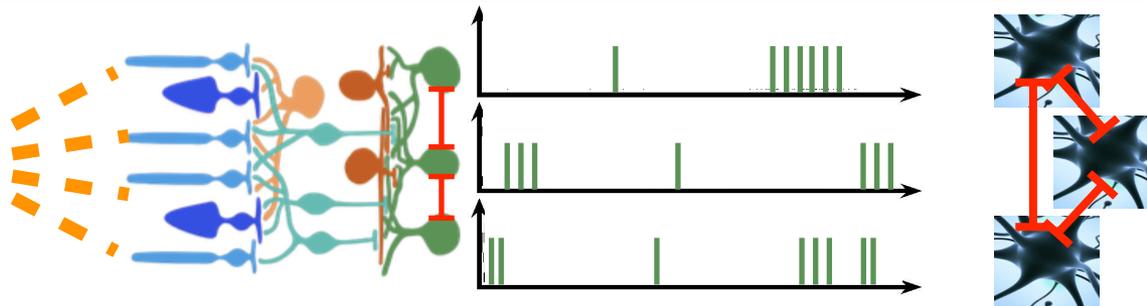


Non-linear & stochastic **feedforward system**,
with horizontal **gap-junctions**

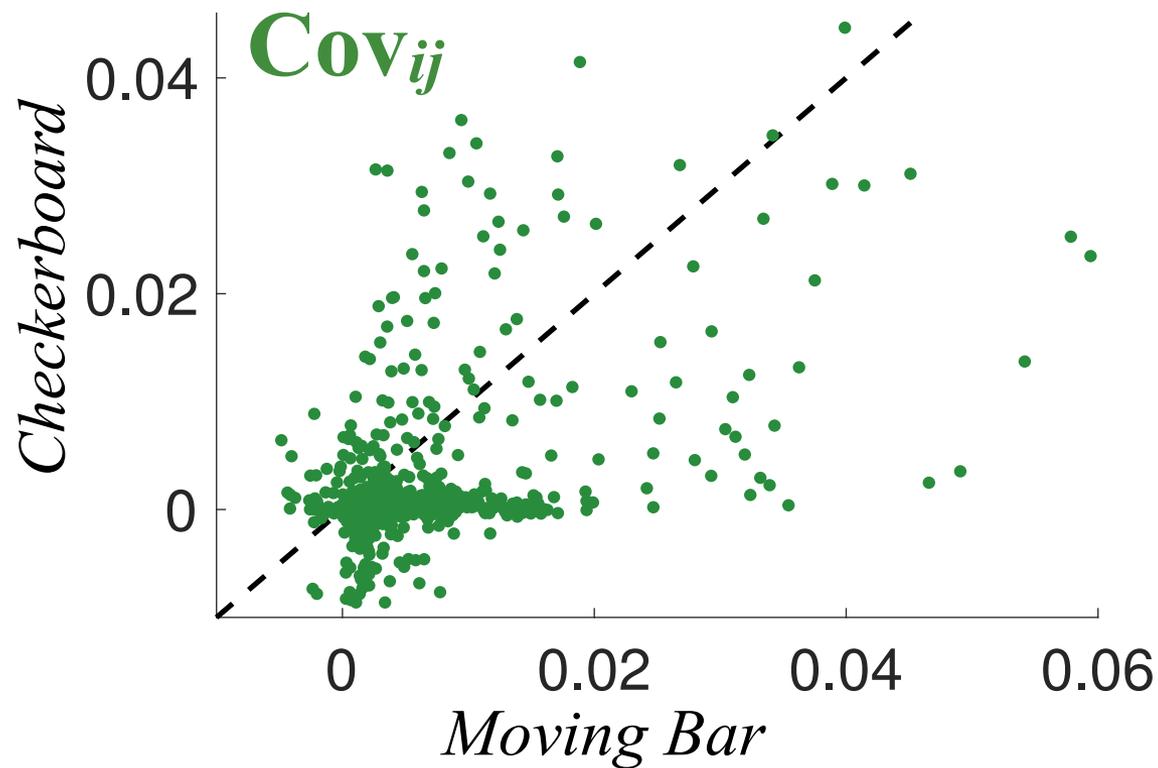
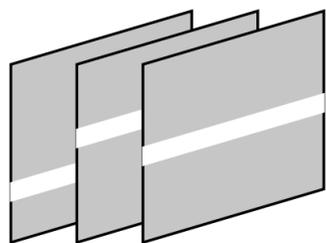
Stimulus as common input



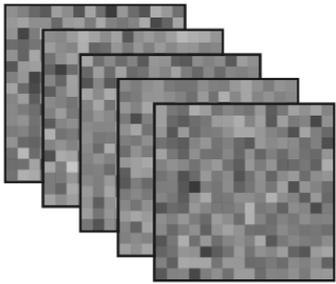
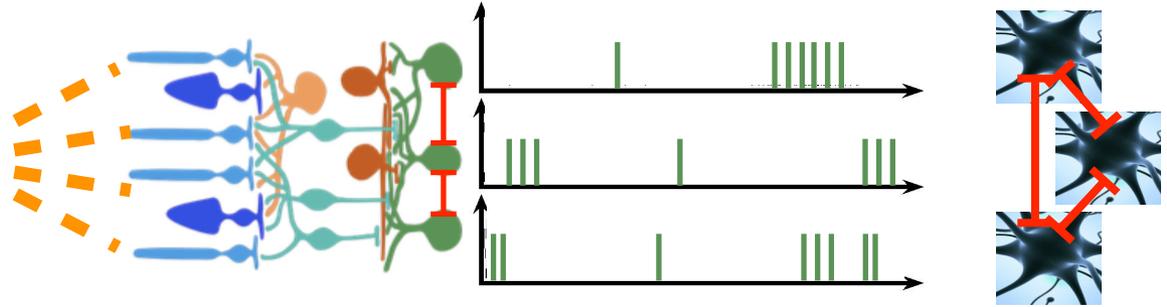
Stimulus as common input



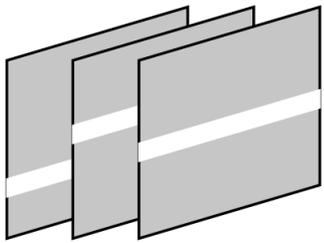
Vs



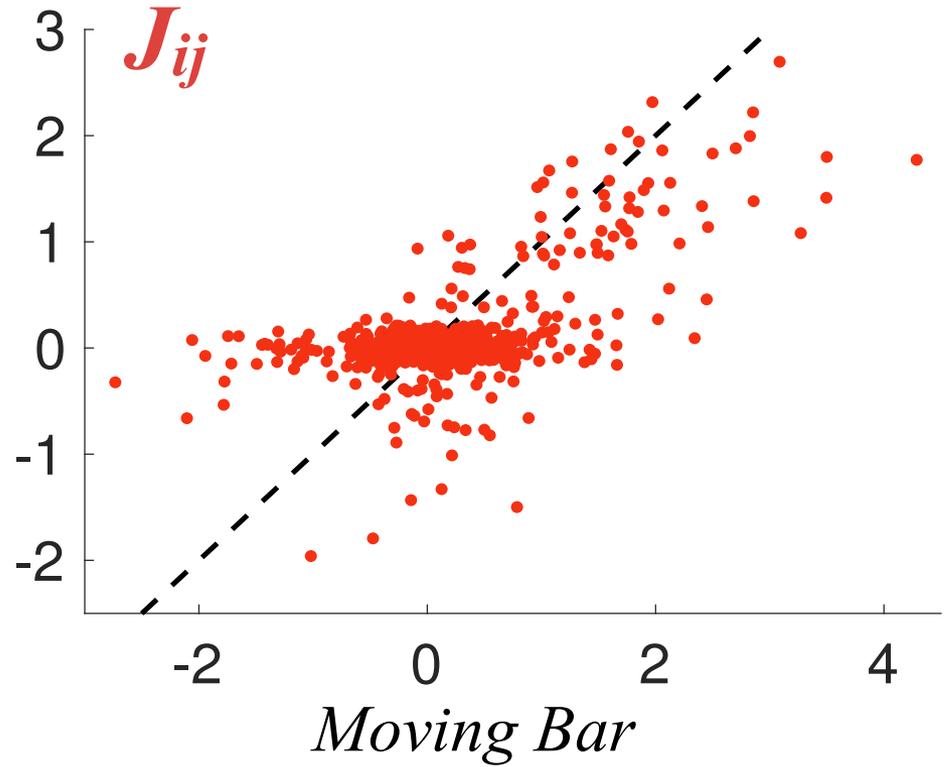
Stimulus as common input



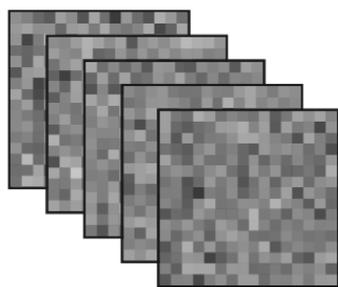
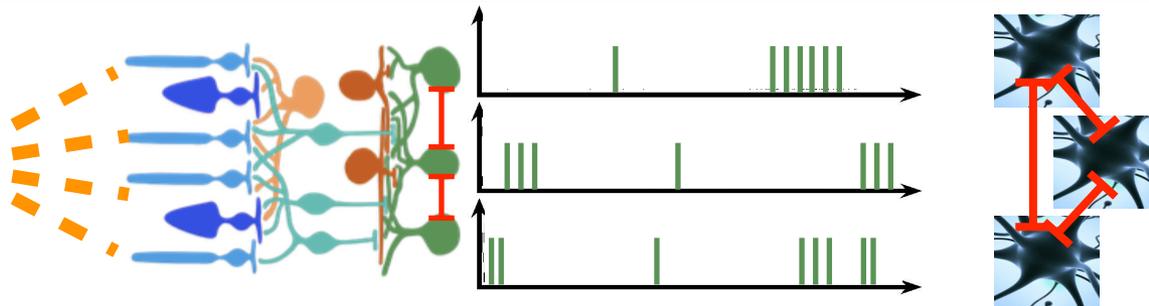
Vs



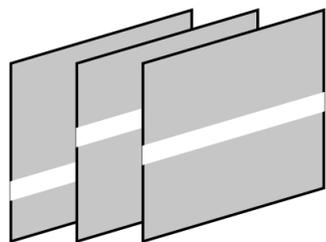
Checkerboard



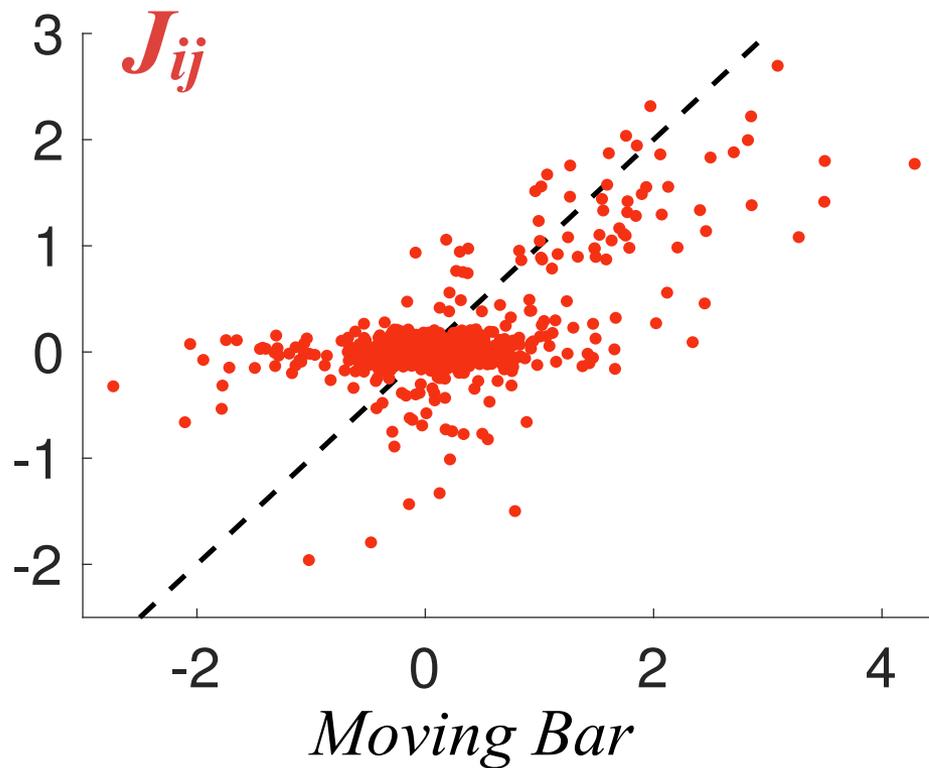
Stimulus as common input



Vs

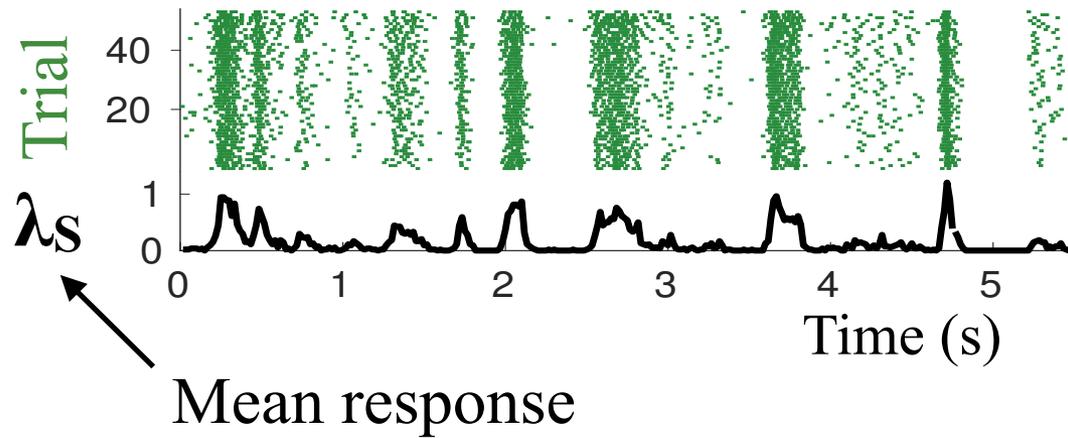


Checkerboard

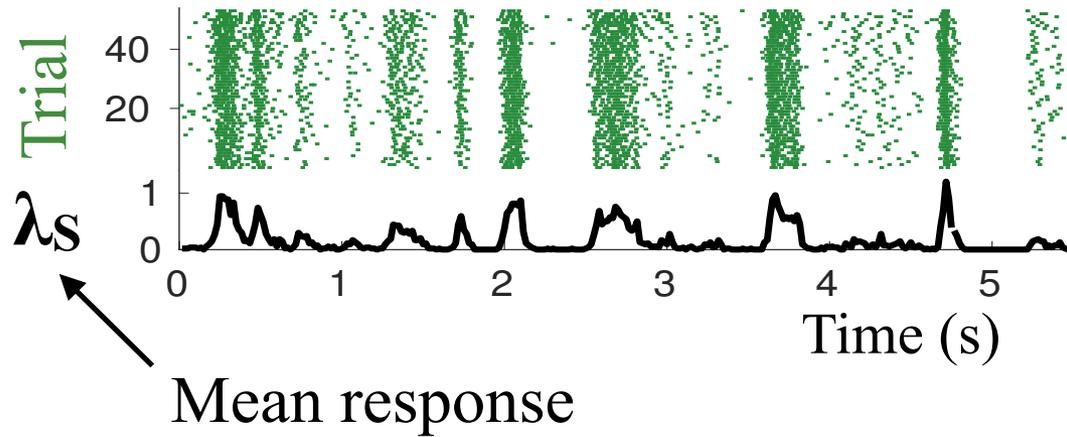


Can I disentangle the two contributions?

Non-linear & stochastic response to stimulus



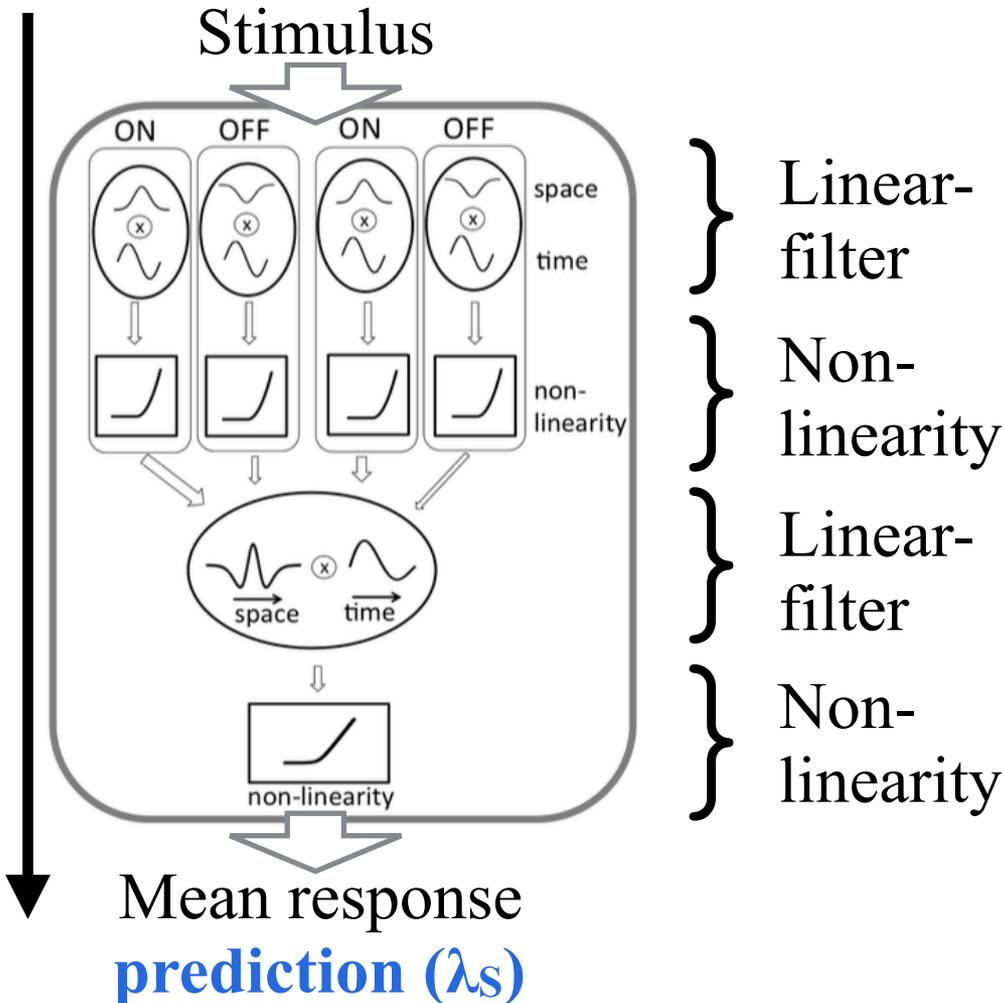
Non-linear & stochastic response to stimulus



Predict λ_S as
function
of the **stimulus**

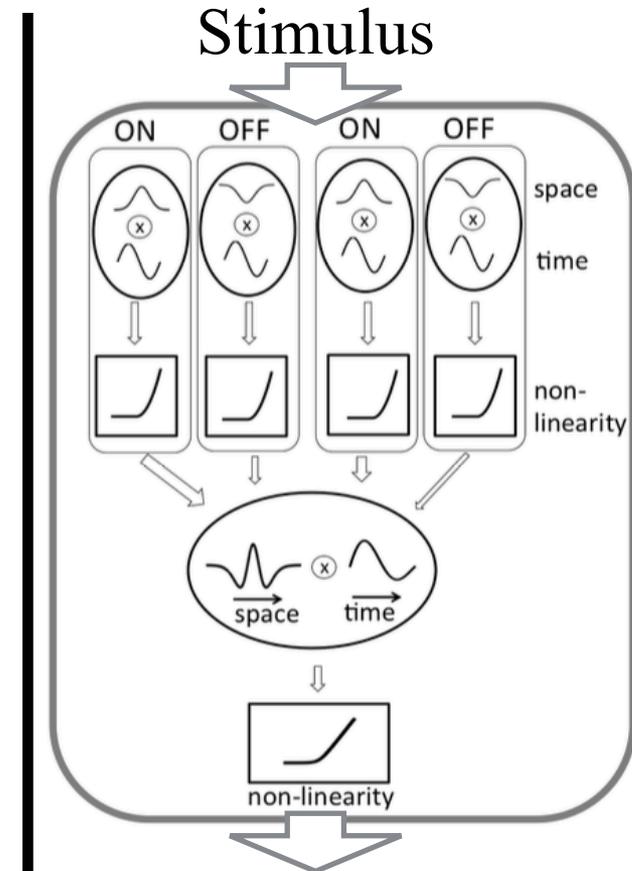
Mean output prediction from stimulus

A non-linear model...

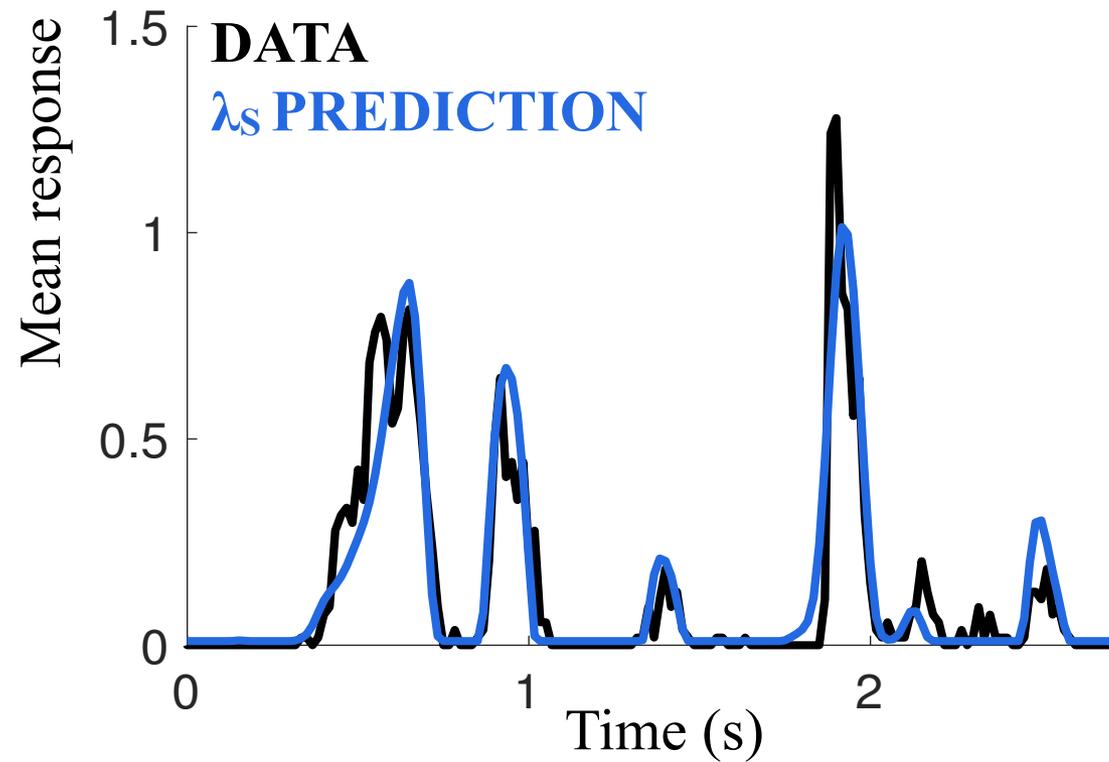


Mean output prediction from stimulus

*A **non-linear** model predicts the mean response*

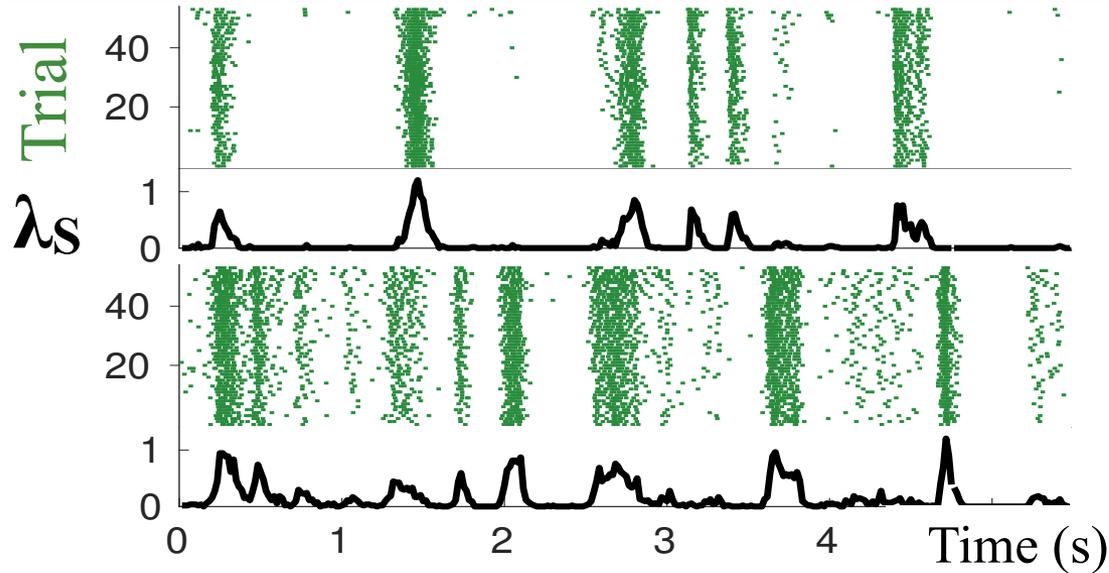


Mean response
prediction (λ_s)



BFGS learning + regularisation

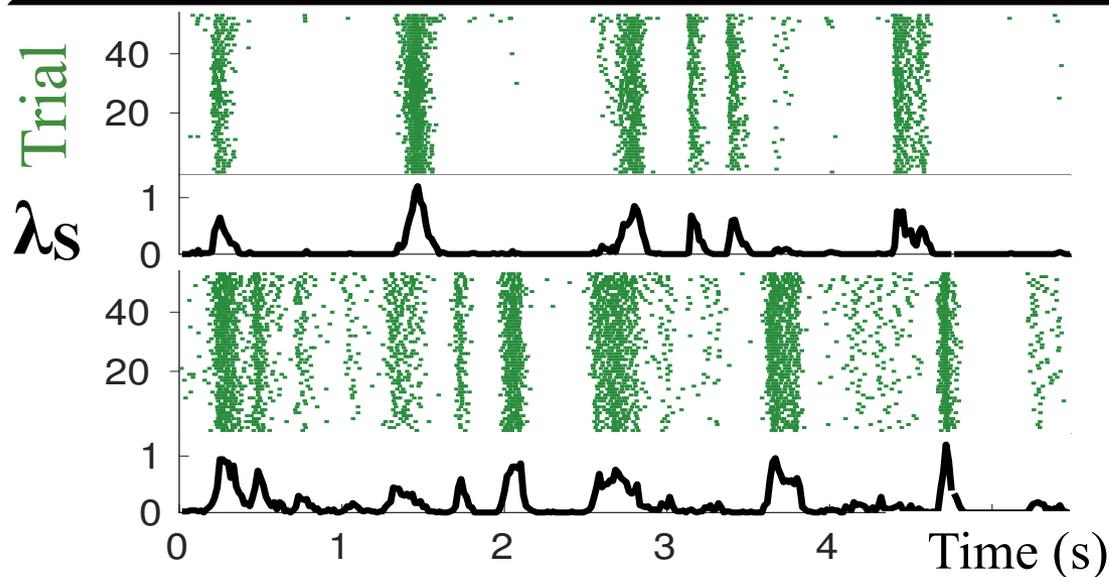
Non-linear & stochastic response to stimulus



Predict λ_S as
function
of the **stimulus**

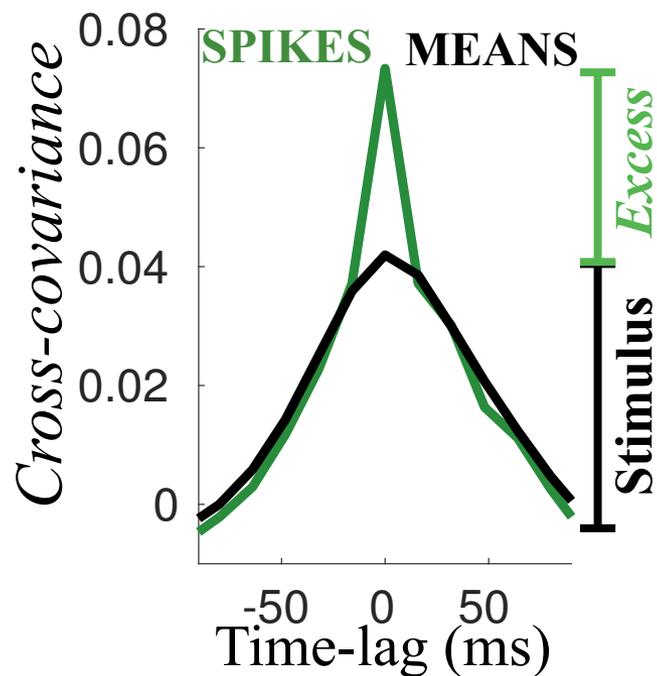
Are Neurons
more synchronous
than expected?

Non-linear & stochastic response to stimulus

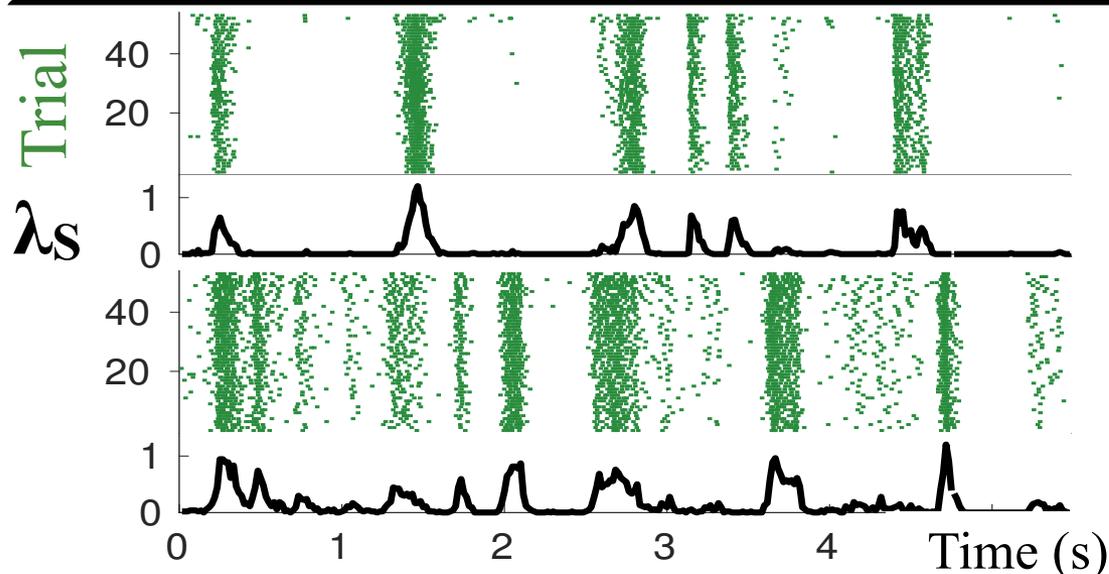


Predict λ_S as
function
of the **stimulus**

Are Neurons
more synchronous
than expected?

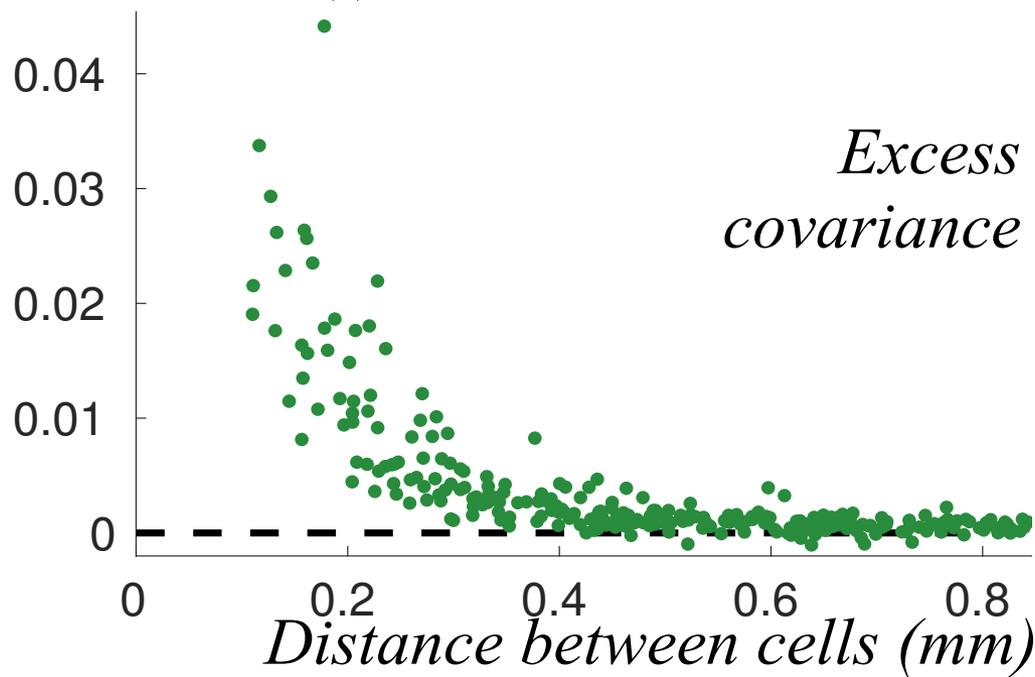
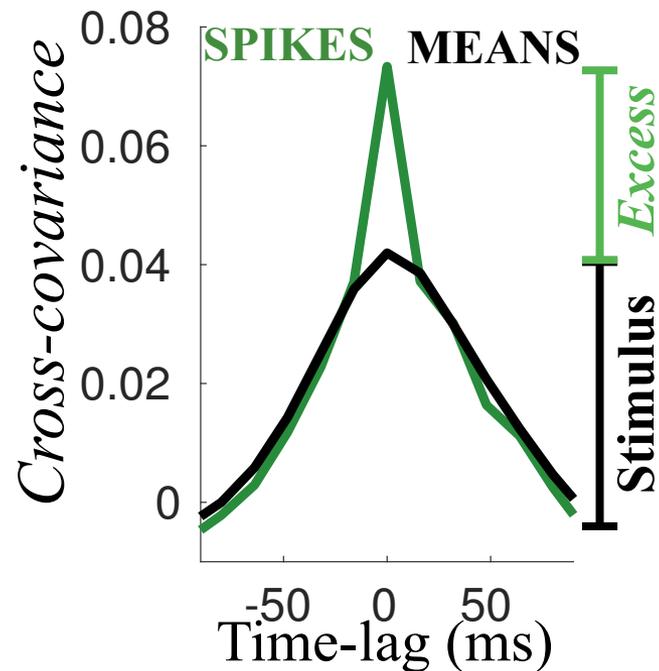


Non-linear & stochastic response to stimulus



Predict λ_S as
function
of the **stimulus**

Are Neurons
more synchronous
than expected?



Collective behavior in neuronal ensembles

Max-Entropy

$$H[P] = - \sum P \log P$$

Neuronal network



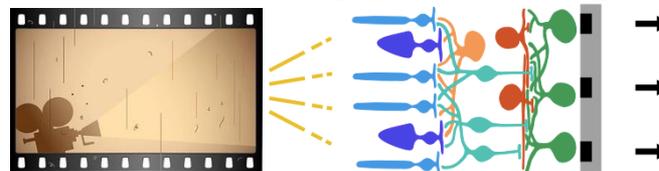
Visual information



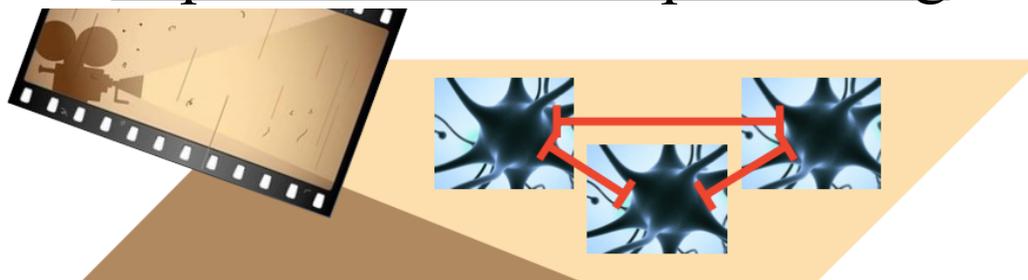
Interaction model



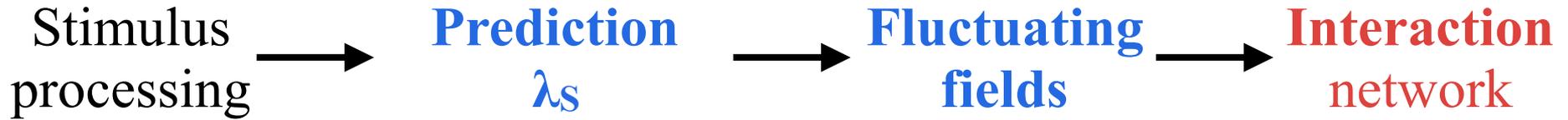
Stimulus processing



Population stimulus processing

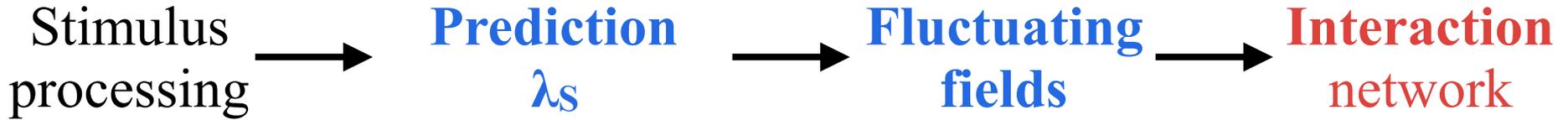


How to account for excess covariances?

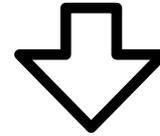


$$P_{\text{Refr}}^{\text{Pop}}(\{n_{it}\}) \sim \prod_t \left(\prod_i e^{h_i[\lambda_{S(t)}] n_{it}} \prod_{i < j} e^{n_{it} J_{ij} n_{jt}} \right)$$

How to account for excess covariances?



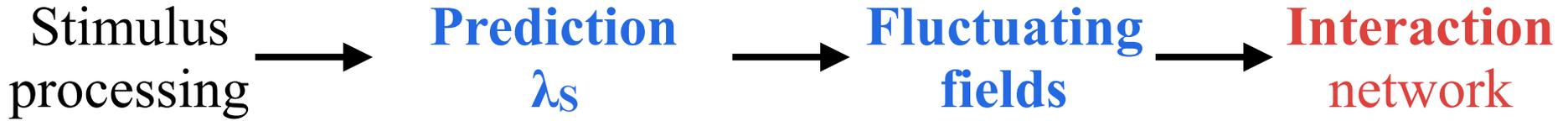
$$P_{\text{Refr}}^{\text{Pop}}(\{n_{it}\}) \sim \prod_t \left(\prod_i e^{h_i[\lambda_S(t)]n_{it}} \prod_{i < j} e^{n_{it} J_{ij} n_{jt}} \right)$$



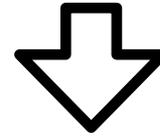
$$\prod_t \left(\prod_i e^{h_i[\text{DATA}(t)]n_{it}} \prod_{i < j} e^{n_{it} J_{ij} n_{jt}} \right)$$

& L1 regularization

How to account for excess covariances?

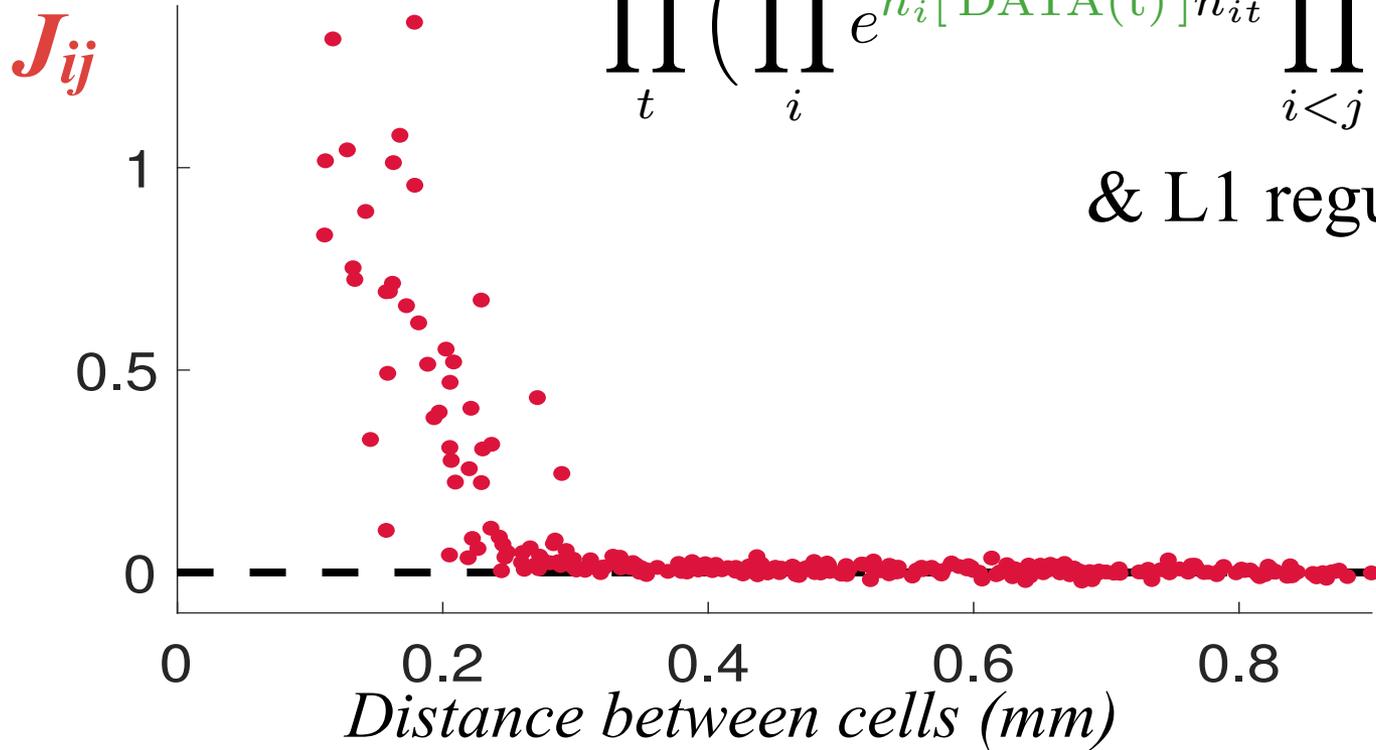


$$P_{\text{Refr}}^{\text{Pop}}(\{n_{it}\}) \sim \prod_t \left(\prod_i e^{h_i[\lambda_S(t)] n_{it}} \prod_{i < j} e^{n_{it} J_{ij} n_{jt}} \right)$$

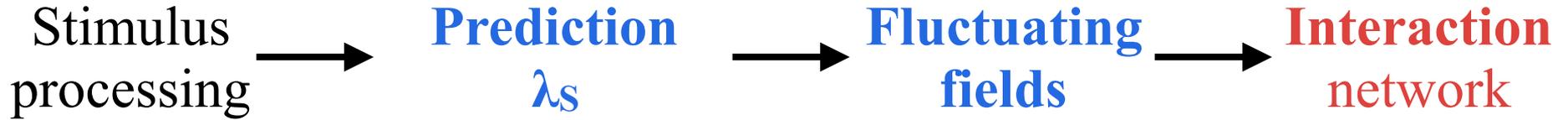


$$\prod_t \left(\prod_i e^{h_i[\text{DATA}(t)] n_{it}} \prod_{i < j} e^{n_{it} J_{ij} n_{jt}} \right)$$

& L1 regularization



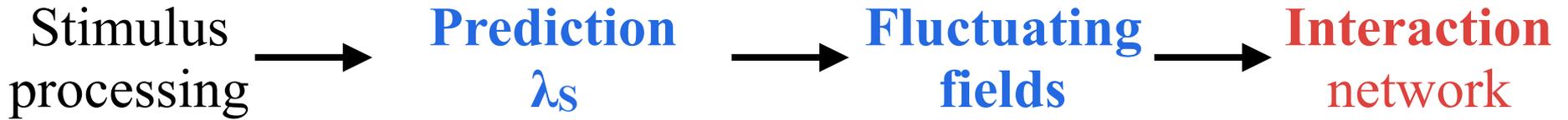
How to account for excess covariances?



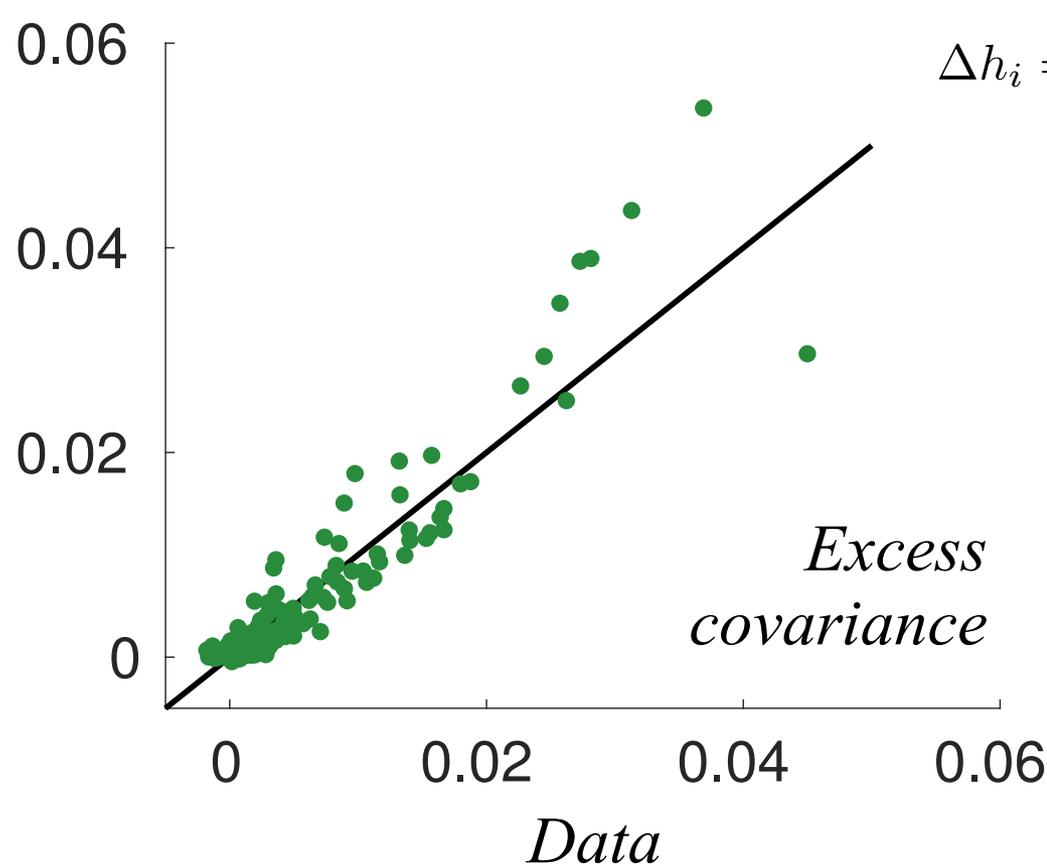
TAP/AMP correction \longrightarrow
$$\prod_t \left(\prod_i e^{h_i[\lambda_{S(t)}] n_{it}} e^{\Delta h_i[\lambda_{S(t)}] n_{it}} \prod_{i < j} e^{n_{it} J_{ij} n_{jt}} \right)$$

$$\Delta h_i = \sum_{j \neq i} J_{ij} \langle n_j \rangle + \frac{1}{2} \frac{d \langle n_i^2 \rangle_c}{d \langle n_i \rangle} \sum_{j \neq i} J_{ij}^2 \langle n_j^2 \rangle_c$$

How to account for excess covariances?

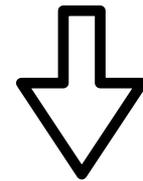


TAP/AMP correction →
$$\prod_t \left(\prod_i e^{h_i[\lambda_S(t)] n_{it}} e^{\Delta h_i[\lambda_S(t)] n_{it}} \prod_{i < j} e^{n_{it} J_{ij} n_{jt}} \right)$$



$$\Delta h_i = \sum_{j \neq i} J_{ij} \langle n_j \rangle + \frac{1}{2} \frac{d \langle n_i^2 \rangle_c}{d \langle n_i \rangle} \sum_{j \neq i} J_{ij}^2 \langle n_j^2 \rangle_c$$

Stimulus processing
&
interactions

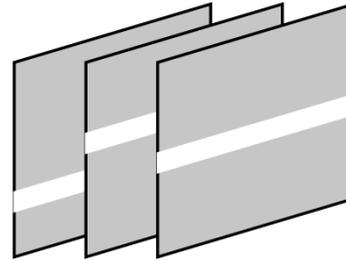


Excess covariance

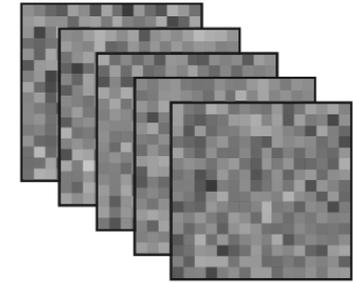
Are couplings robust to changes in common inputs?

Stimulus processing → **Prediction**
 λ_s → **Fluctuating fields** → **Interaction network**

Moving bar



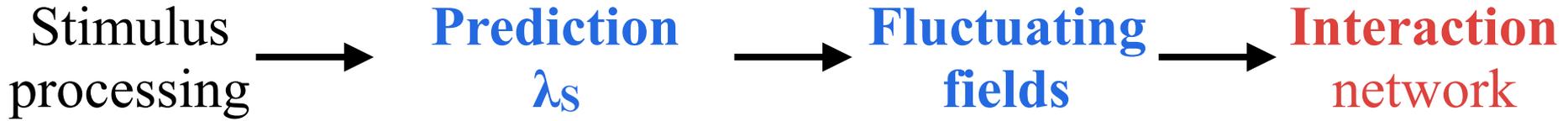
Checkerboard



Change in the
stimulus ensemble:



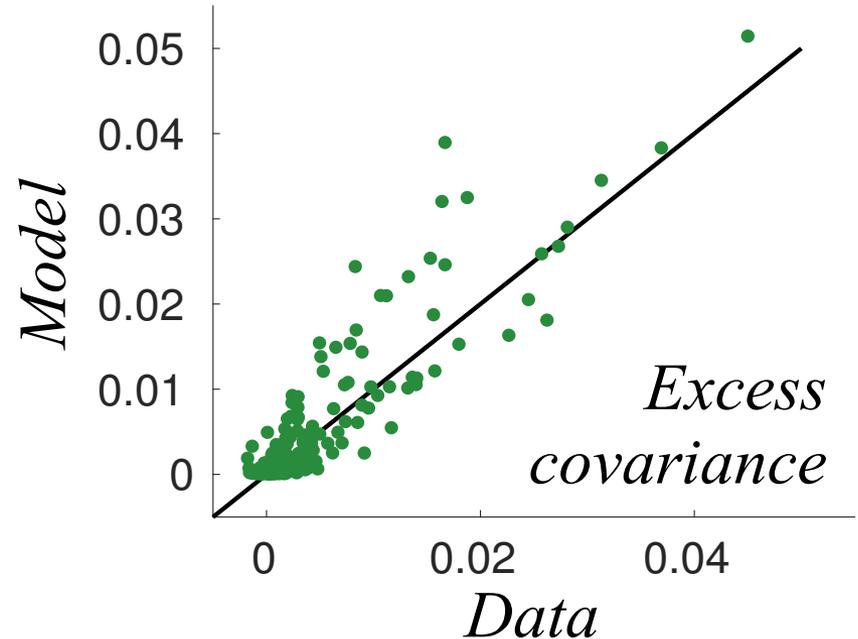
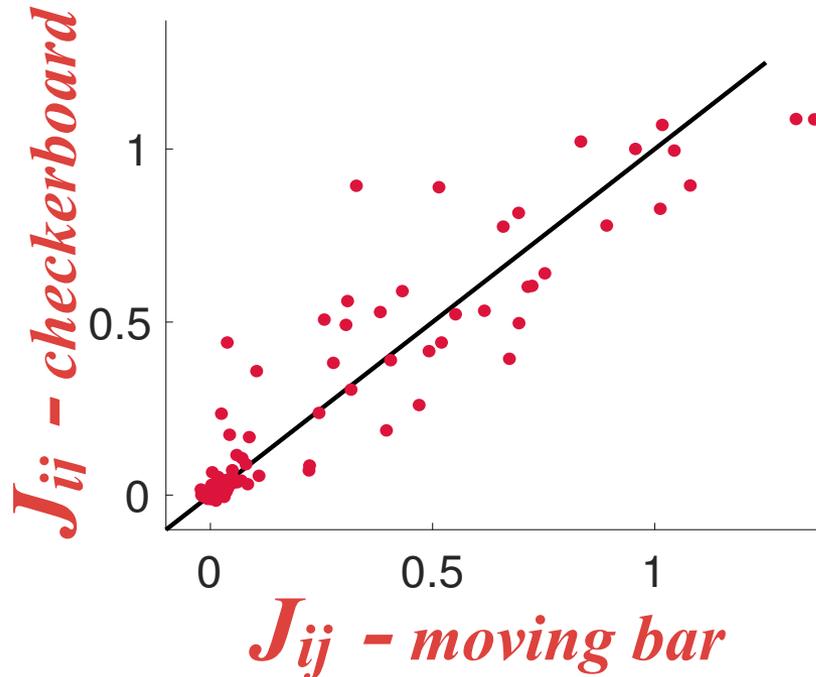
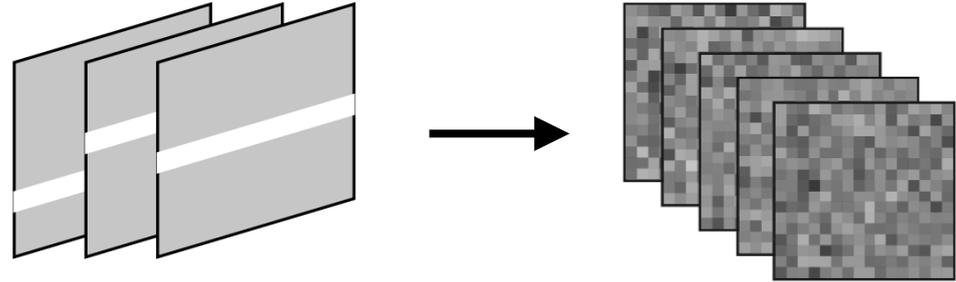
Are couplings robust to changes in common inputs?



Moving bar

Checkerboard

Change in the stimulus ensemble:



Collective behavior in neuronal ensembles

Conclusions:

- **Maximum entropy** principle allows for characterisation of network activity
- It can be used to identify **neuronal cliques**, basis for **memory formation**
- **Deep** stimulus processing model predicts mean retinal output
- **Interaction network** accounts for excess correlations in response to different stimuli

Collaborators

Theoretical side

- R. Monasson, S. Cocco, T. Mora
- G. Tkacik
- T. Obuchi



Experimental side

- O. Marre, S. Picaud
- F. Battaglia
- A. Destexhe



Thank you

Max-Entropy

Neuronal network

Visual information

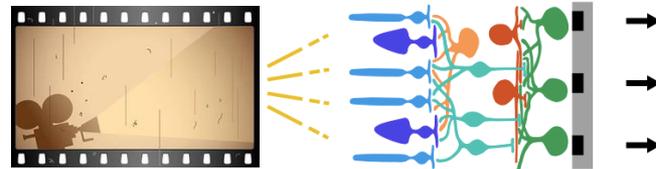
$$H[P] = - \sum P \log P$$



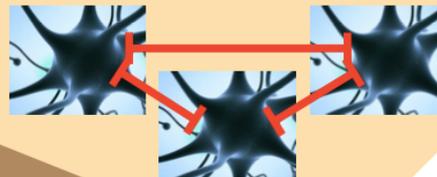
Interaction model



Stimulus processing

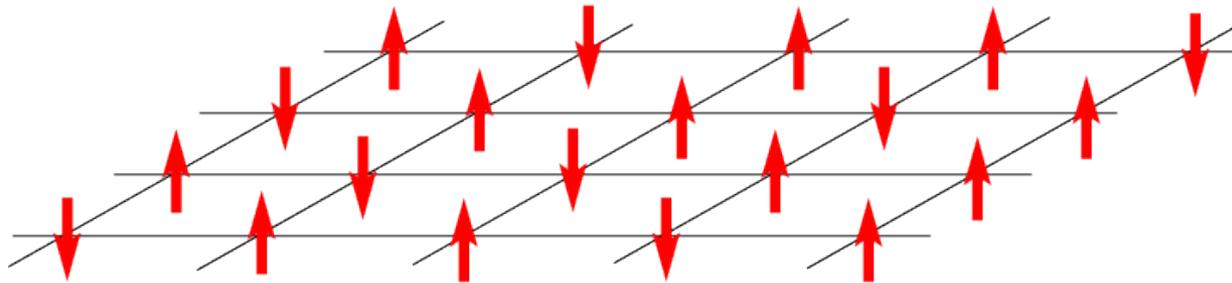


Population Stimulus processing



Spin Glasses

Statistical models with an interaction network



*Collective behavior
in disordered systems*

- 1) Phys Rev B '10;
- 2) J. Stat. Mech '11;
- 3) Phys. Rev. Lett '12;
- 4) Phys. Rev. B '12;
- 5) Phys. Rev. B '12;
- 6) Phys. Rev. B '13;
- 7) Phys. Rev. E '15

Critical Slowing Down Exponents of Mode Coupling Theory

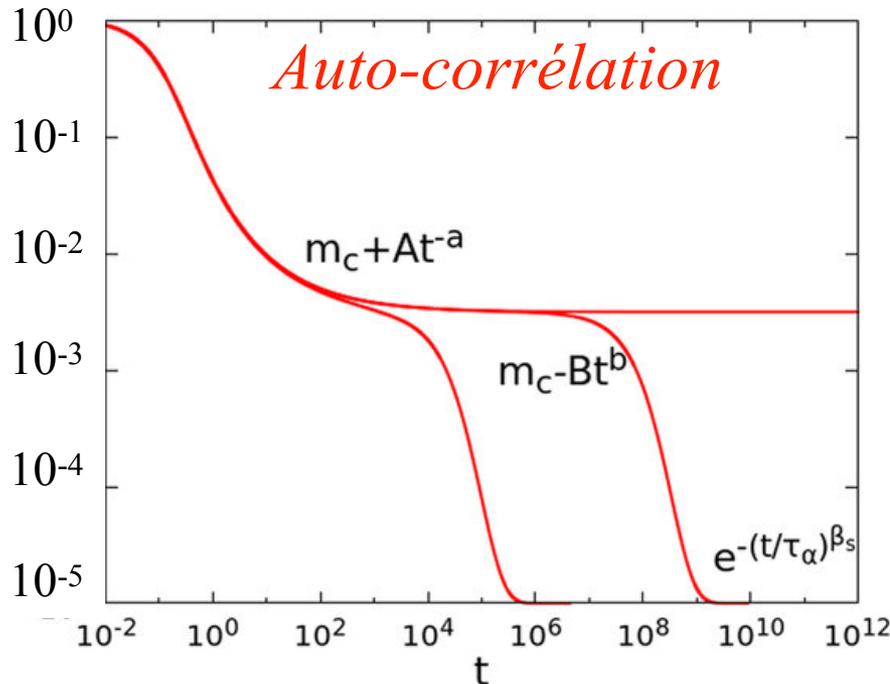
F. Caltagirone,¹ U. Ferrari,^{1,2} L. Leuzzi,^{1,2} G. Parisi,^{1,2,3} F. Ricci-Tersenghi,^{1,2,3} and T. Rizzo^{1,2}

¹*Dip. Fisica, Università La Sapienza, Piazzale A. Moro 2, I-00185, Rome, Italy*

²*IPCF-CNR, UOS Roma Kerberos, Università La Sapienza, P. le A. Moro 2, I-00185, Rome, Italy*

³*INFN, Piazzale A. Moro 2, 00185, Rome, Italy*

(Received 5 December 2011; published 24 February 2012)



Action de la theorie de champs:

$$\frac{1}{2} \sum_{(ab),(cd)} \delta Q_{ab} M_{ab,cd} \delta Q_{cd} - \frac{w_1}{6} \text{Tr} \delta Q^3 - \frac{w_2}{6} \sum_{ab} \delta Q_{ab}^3$$

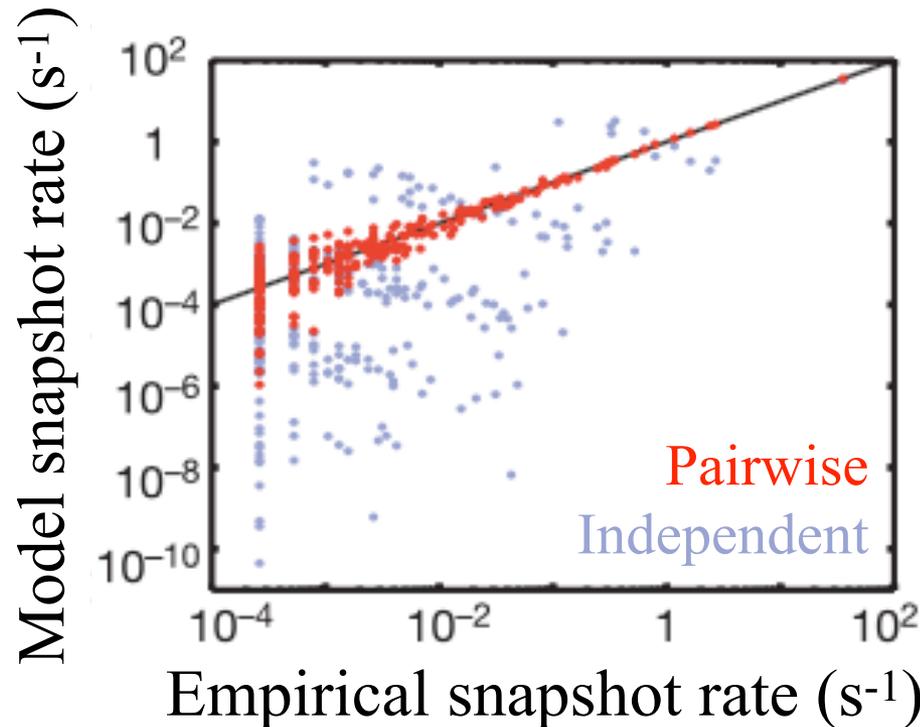
Équation pour les exposants:

$$\frac{\Gamma^2(1-a)}{\Gamma(1-2a)} = \frac{\Gamma^2(1+b)}{\Gamma(1+2b)} = \frac{w_2}{w_1}$$

Are these models good for neuronal network?

Schneidman et al. Nature 2006

Pairwise Maximum Entropy model



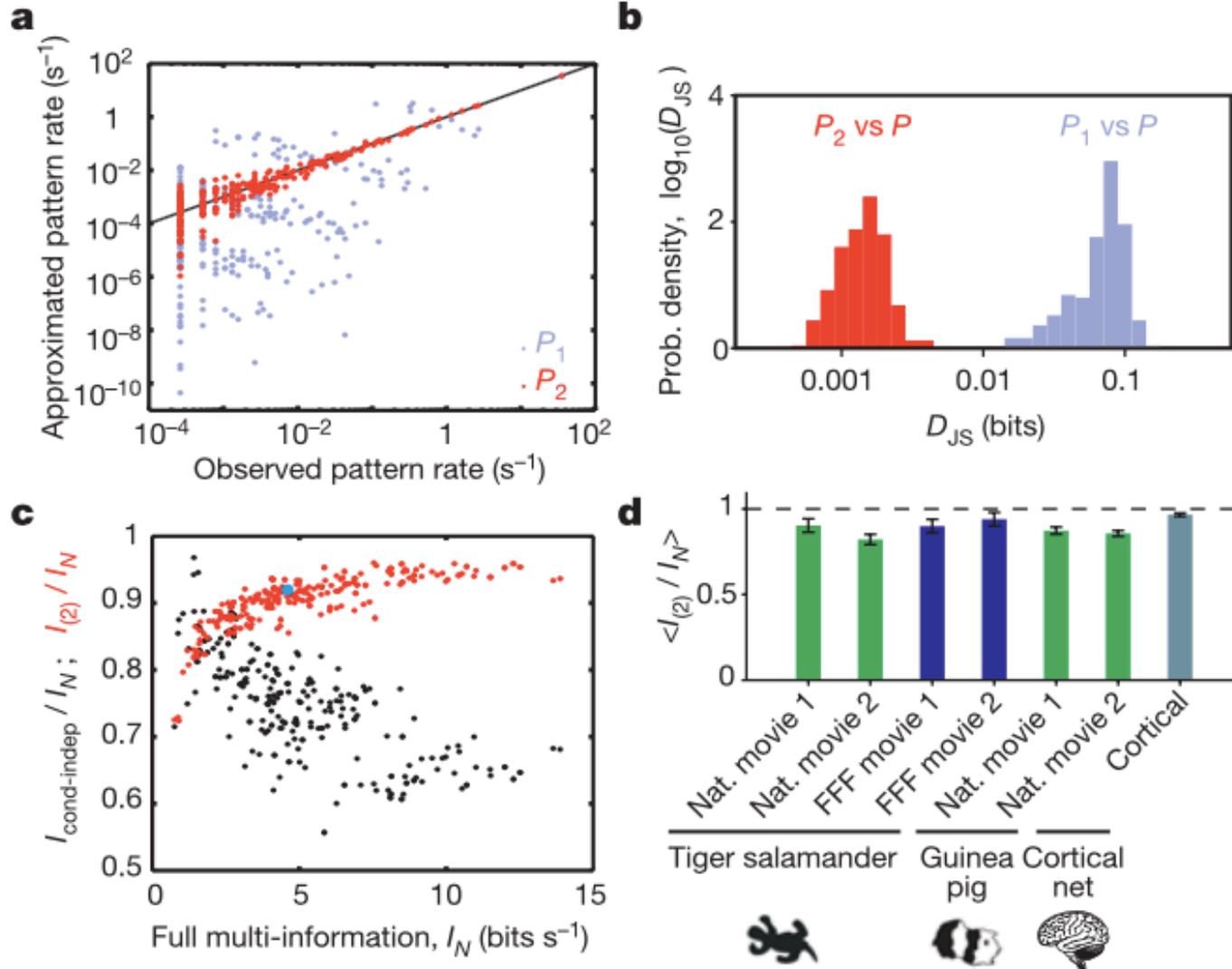
is a good model for salamander retina...

...but is any better than others?

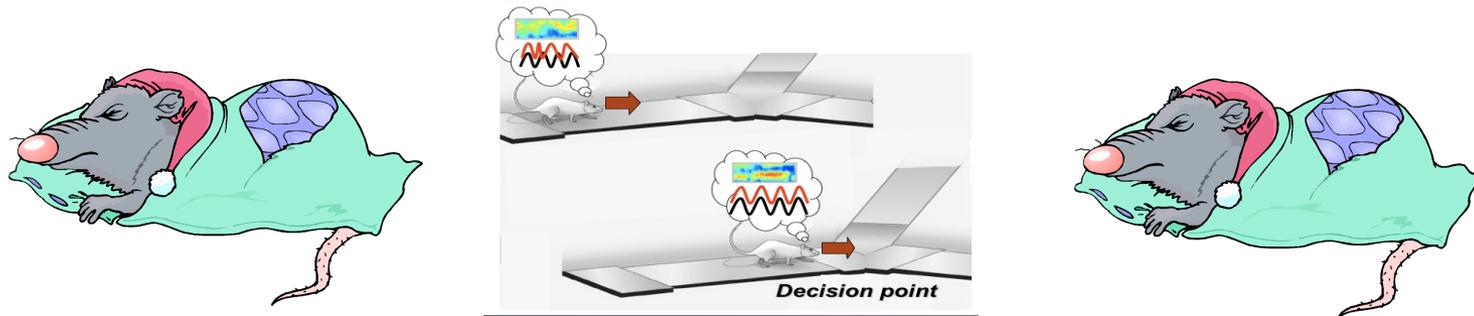
*(with the **same** data-knowledge)*

Modèles à entropie maximale

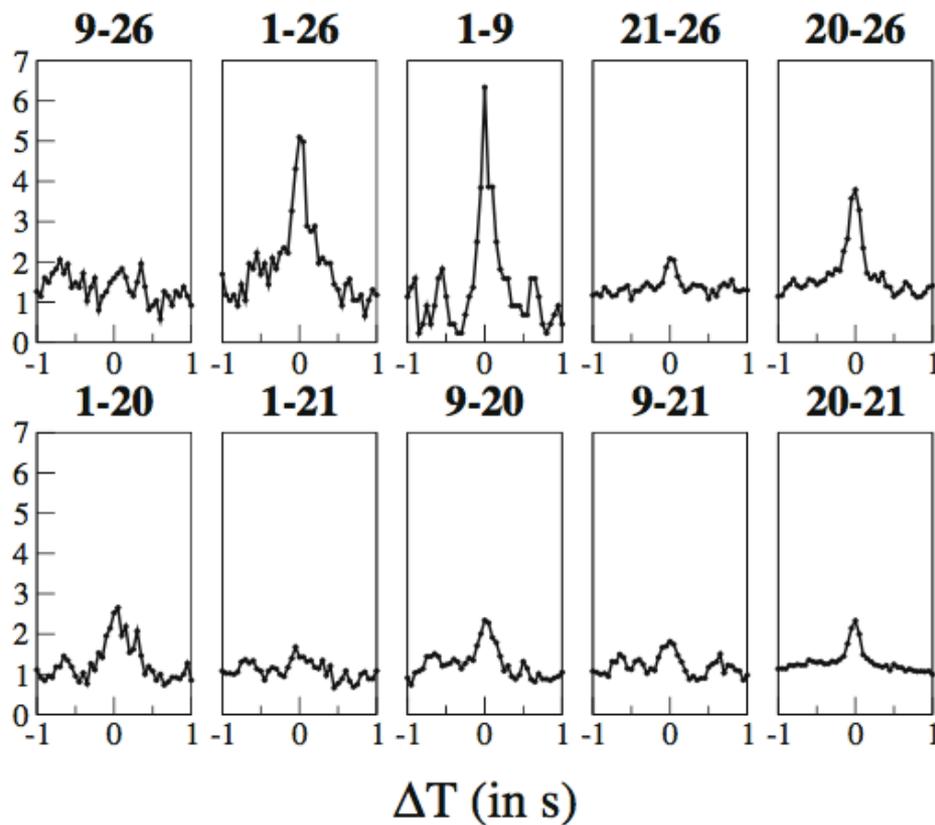
Schneidman et al. Nature 2006



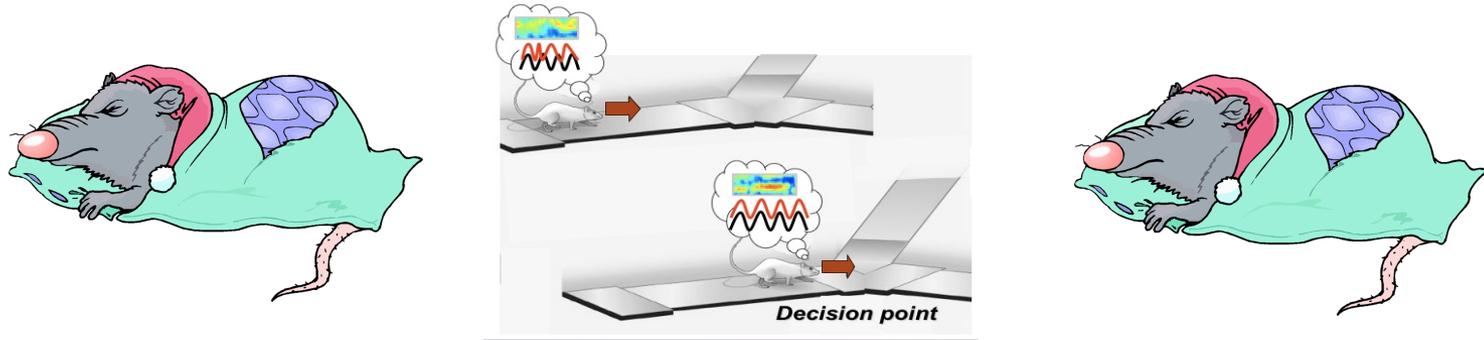
Quelle est la base neurale de la mémoire ?



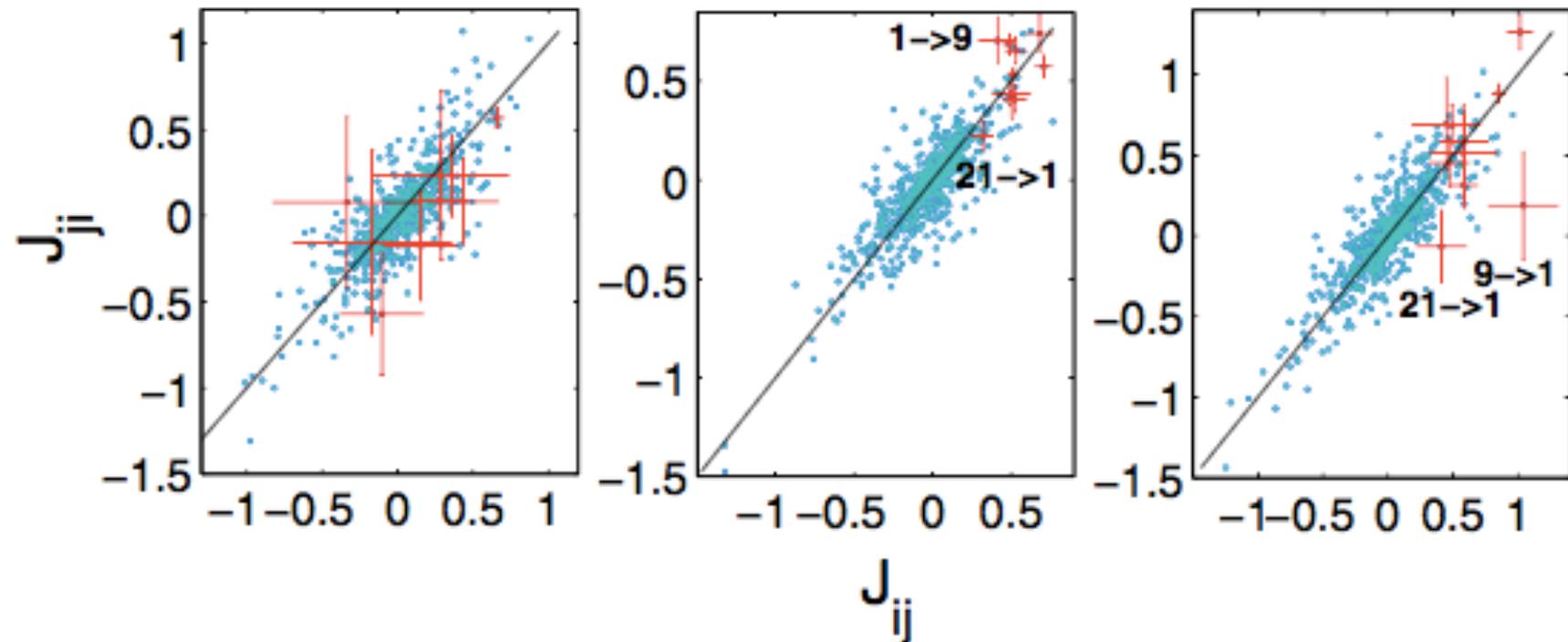
Cross correlation



Quelle est la base neurale de la mémoire ?



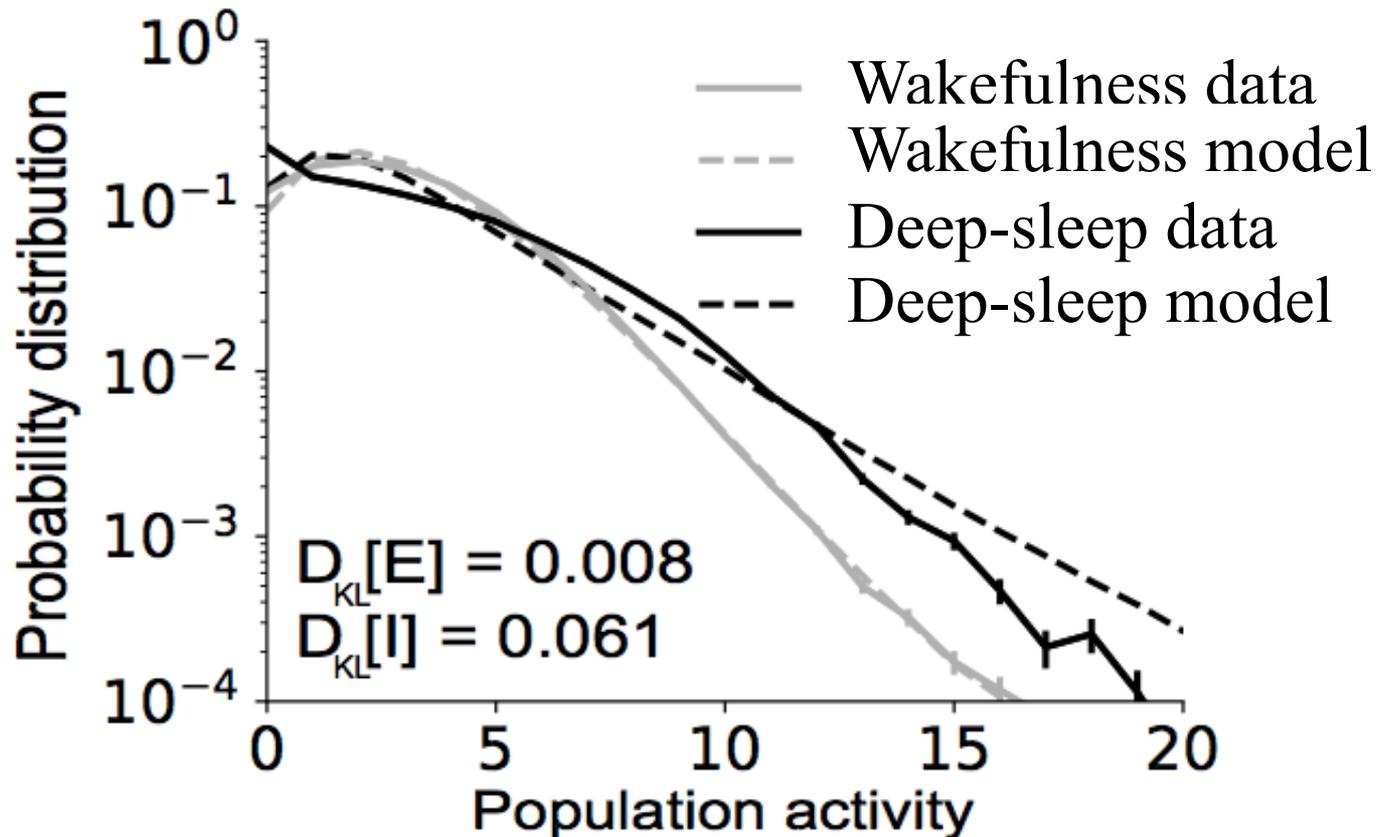
Generalised linear model



Are pairwise interactions sufficient during sleep?

Temporal cortex of human patient at $\Delta T = 50ms$

Wakefulness Vs Deep-Sleep



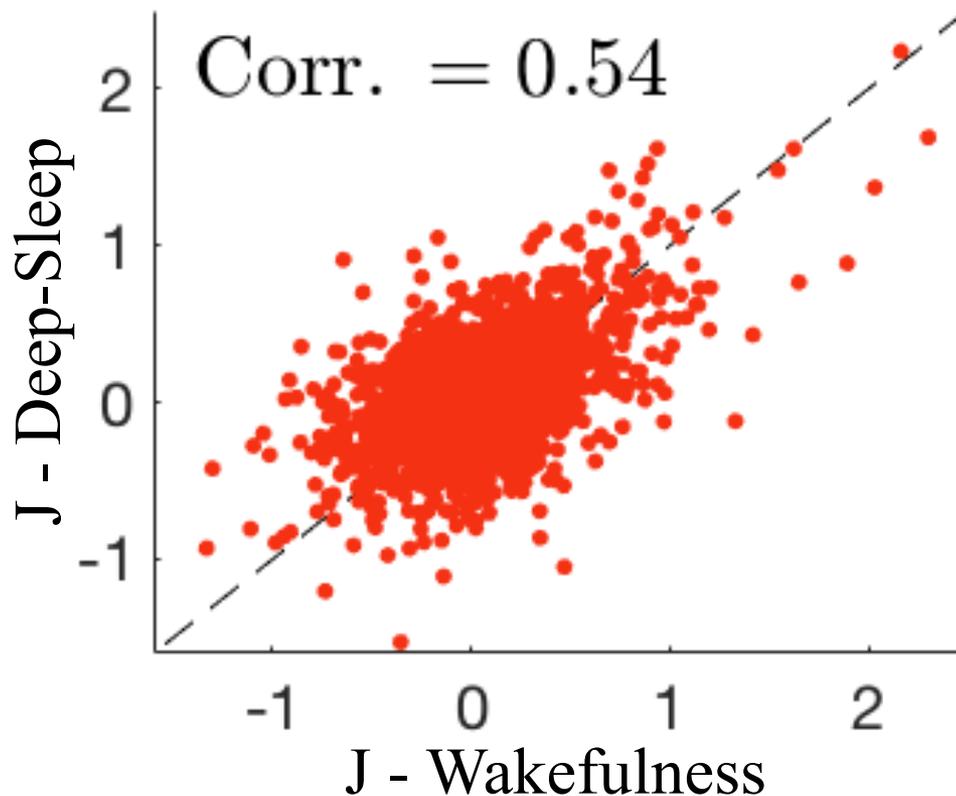
Same system, very different behavior

Nghiem, Telenzuk, Marre, Destexhe, UF, in rev. @ PRX

Are pairwise interactions sufficient during sleep?

Temporal cortex of human patient at $\Delta T = 50ms$

Wakefulness Vs Deep-Sleep

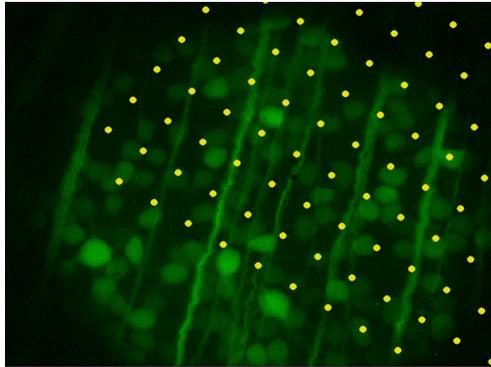


Same system, very different behavior

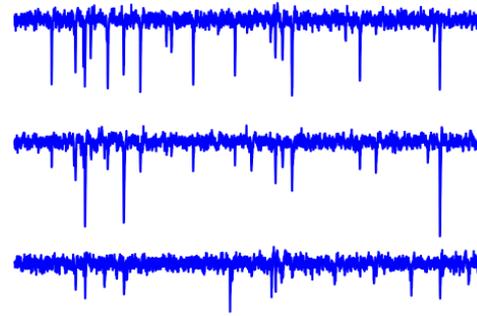
Nghiem, Telenzuk, Marre, Destexhe, UF, in rev. @ PRX

Framework: stimulus processing in the retina

Multi-electrode array

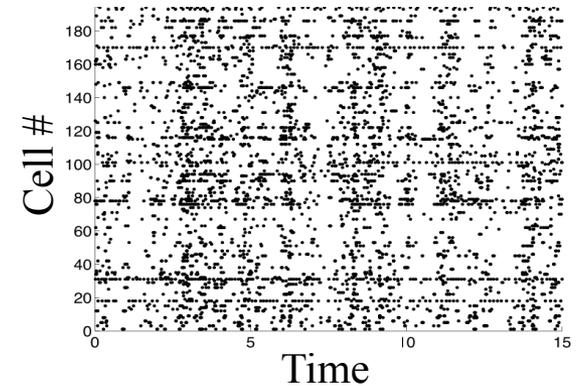


Extracellular potential

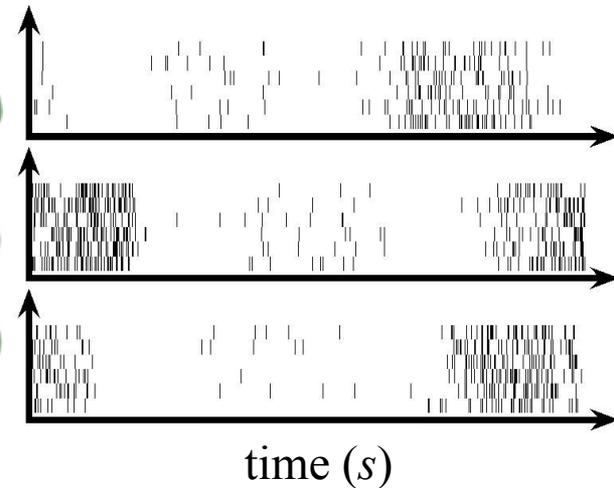
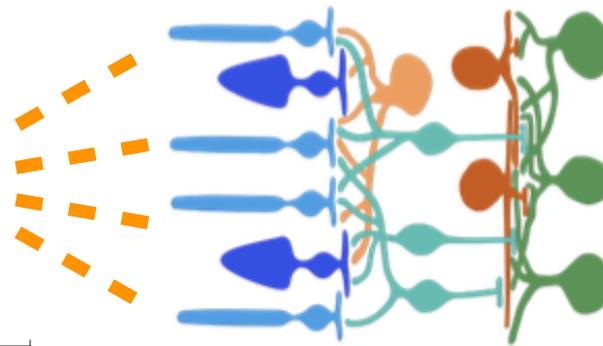
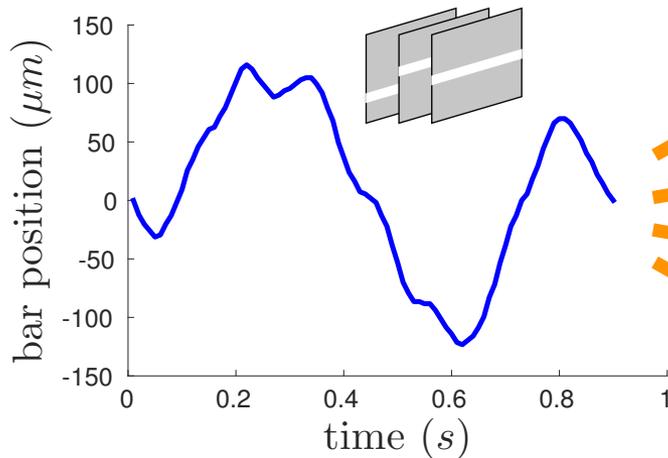


Marre '12; Yger '16

Spike sorting

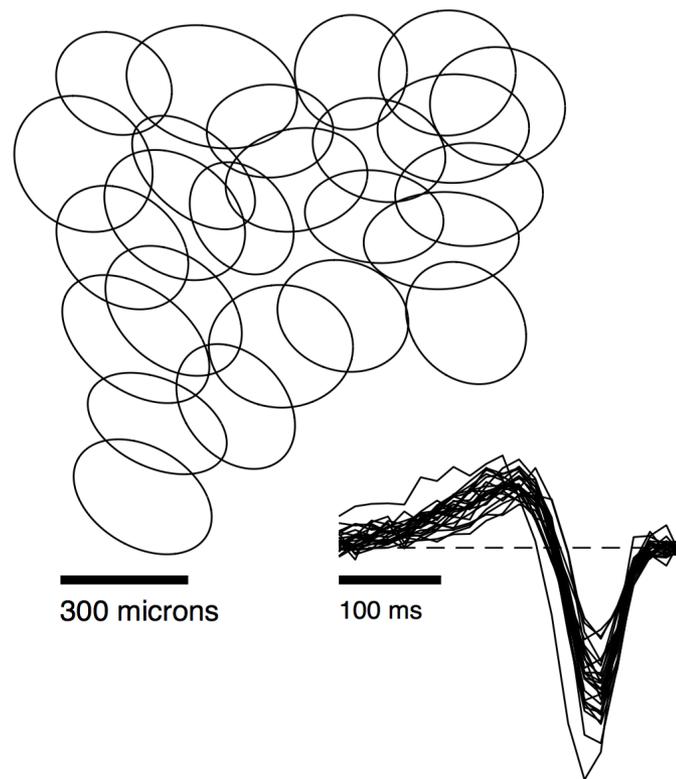
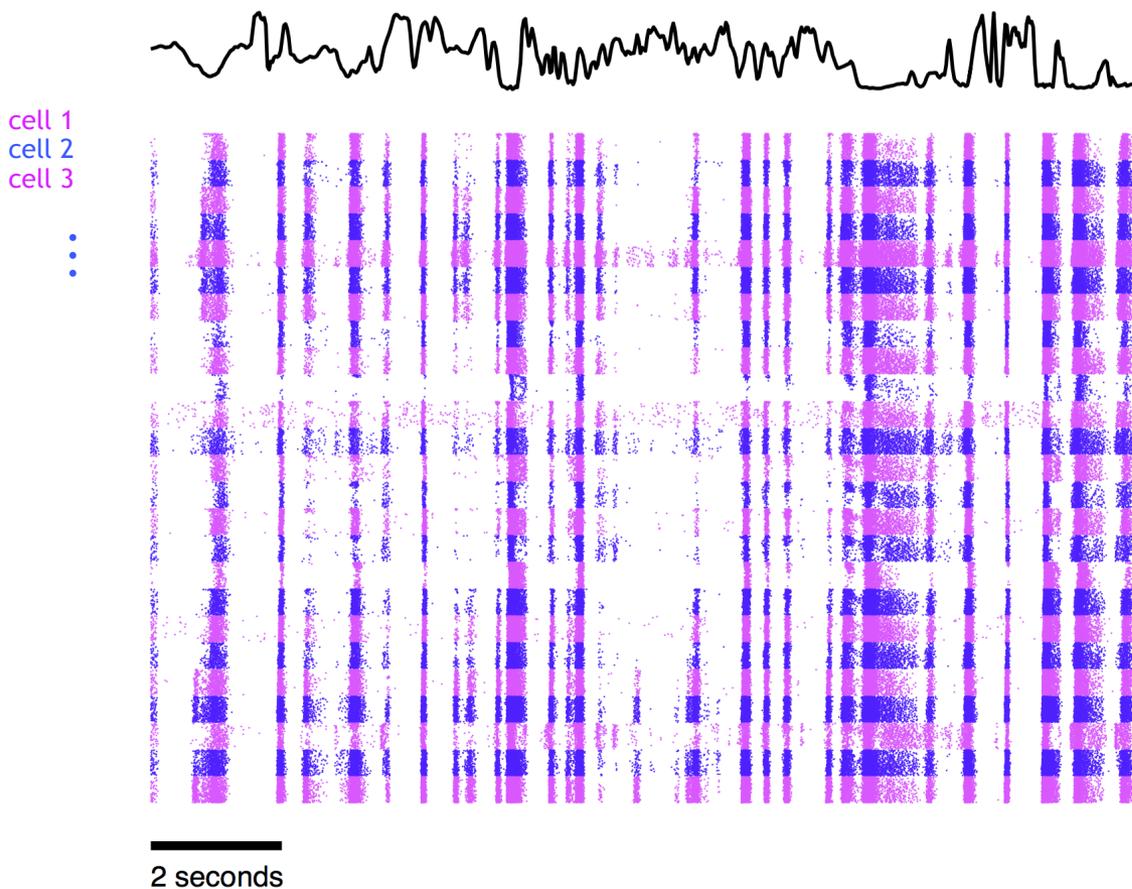


STIMULUS S

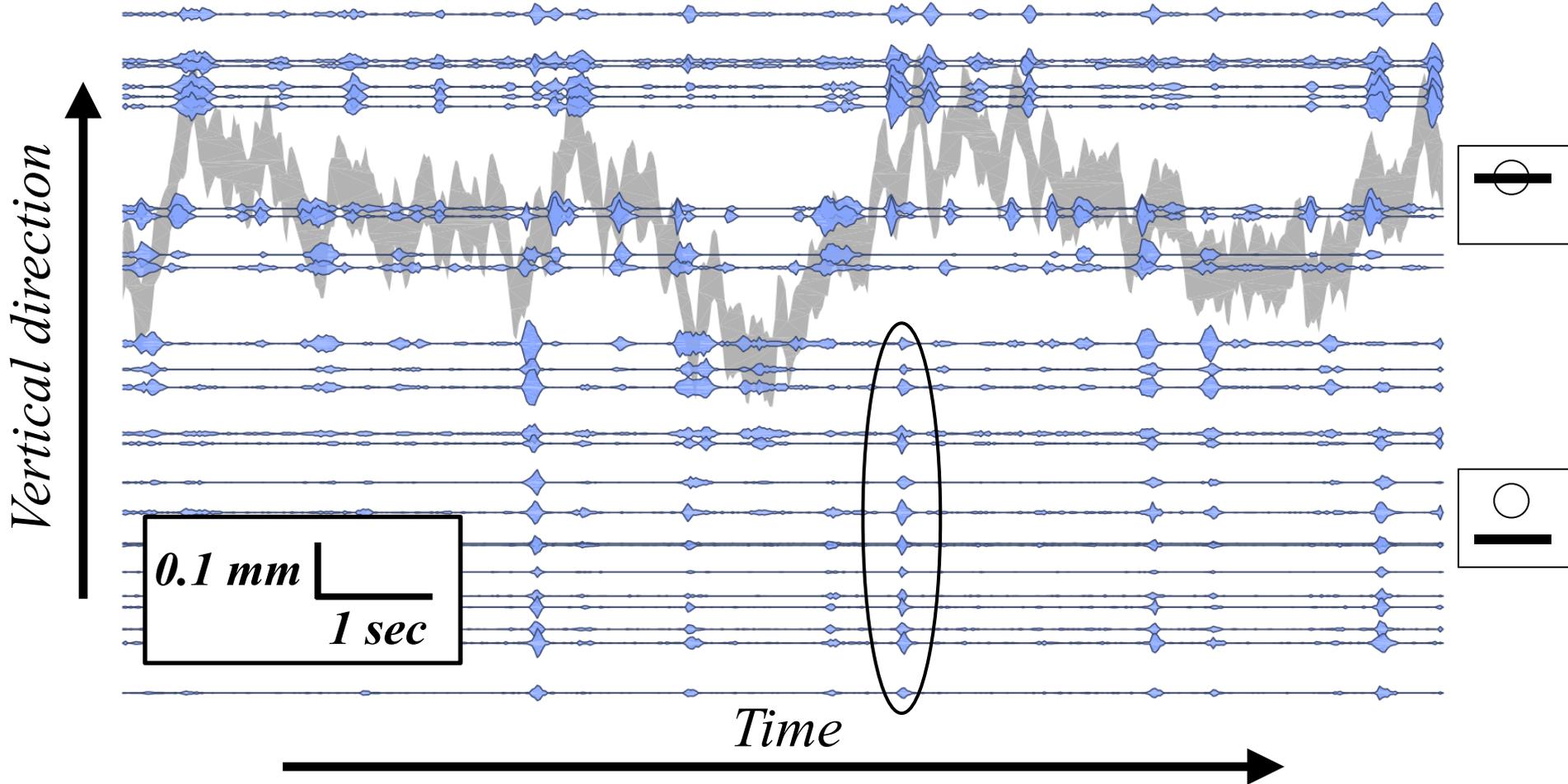


Modelling & characterising the **response to stimulus**

Strategy: input/output modelling

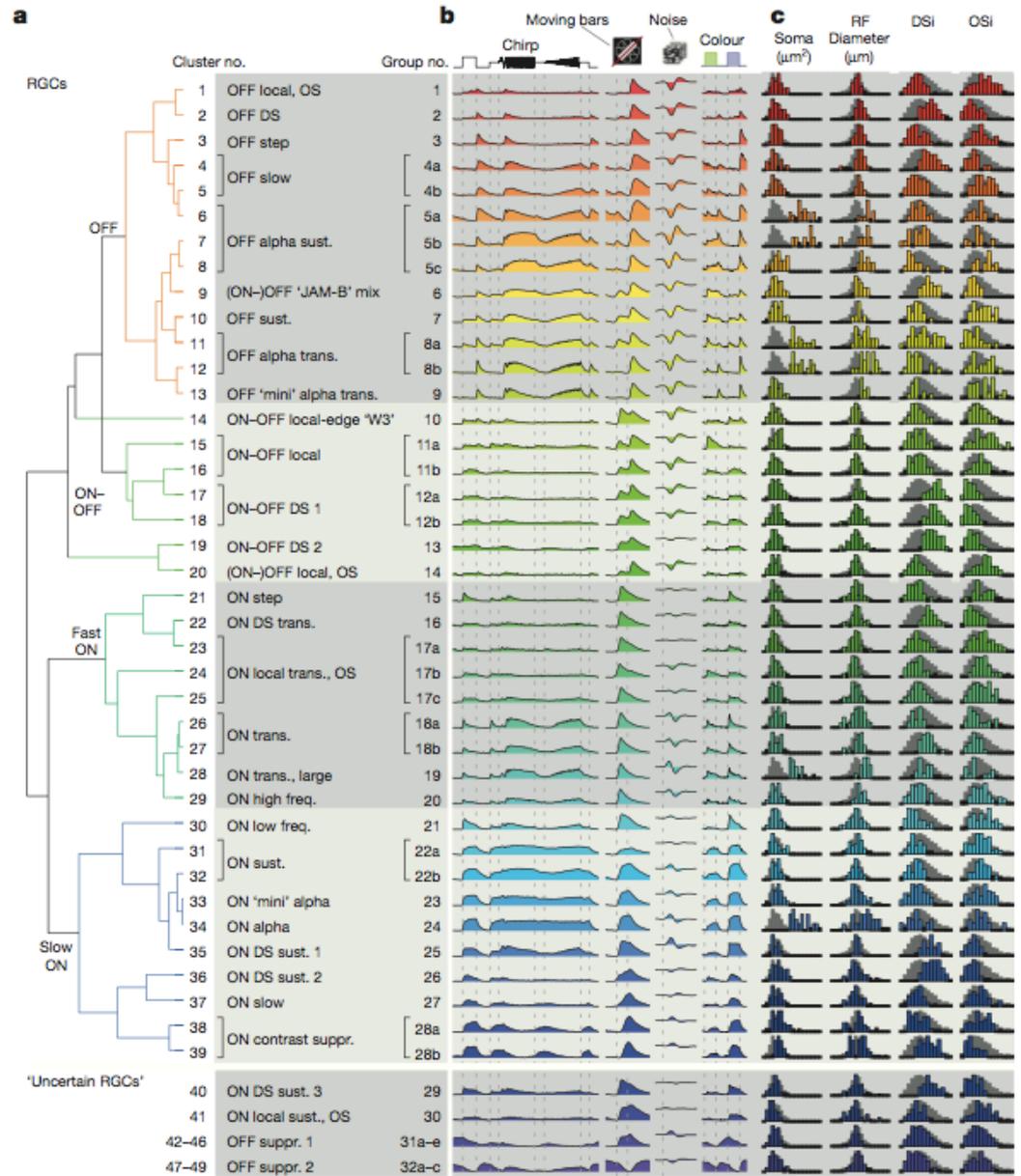
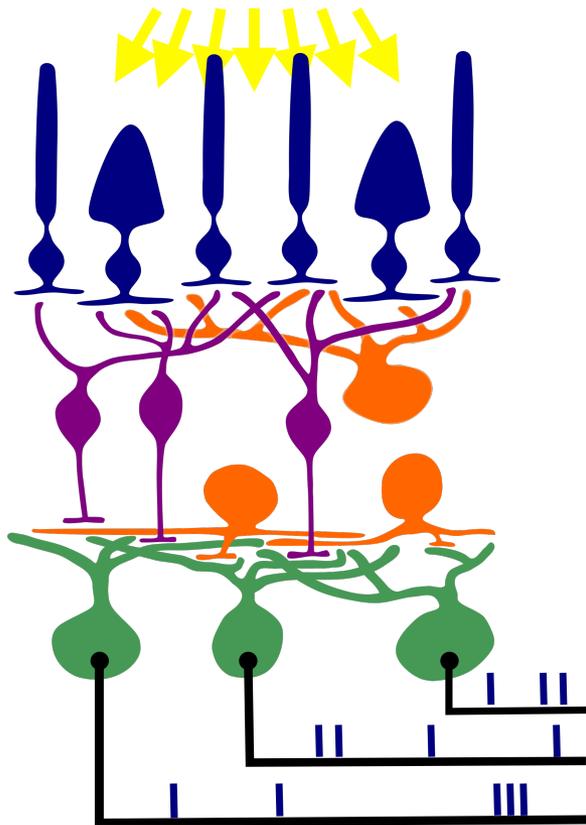


What makes it a hard task? **Flexibility!**



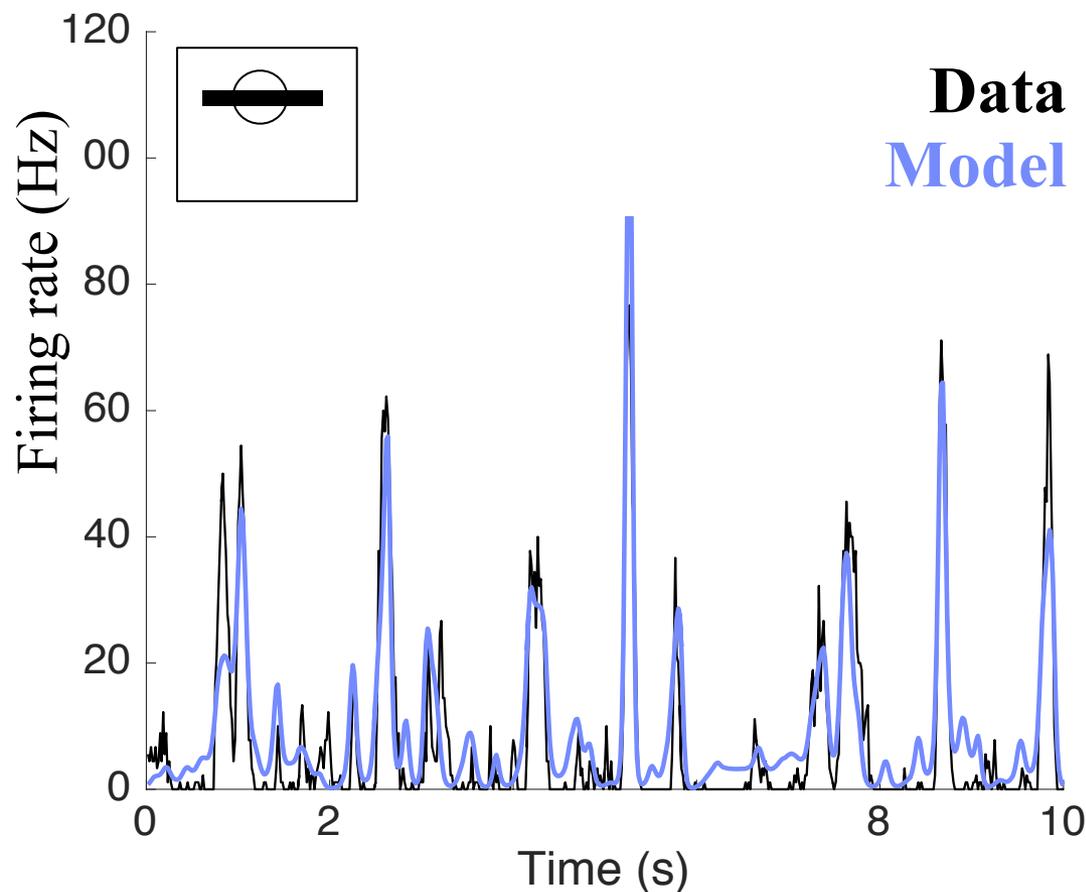
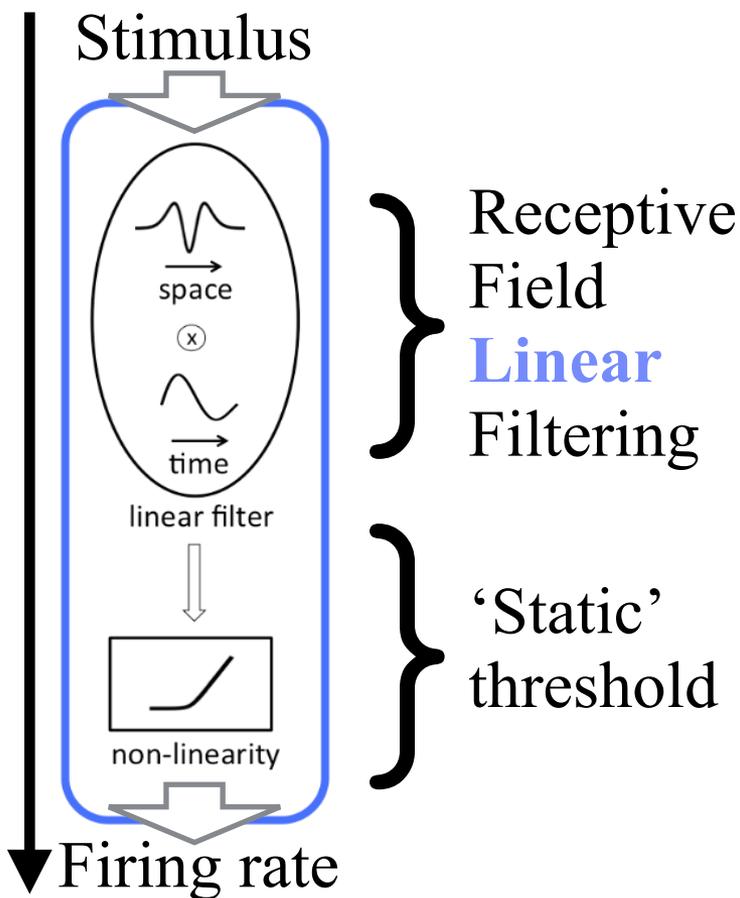
What makes the code **Flexible**?
Spoiler: **Non-linearity** of stimulus processing

La rétine



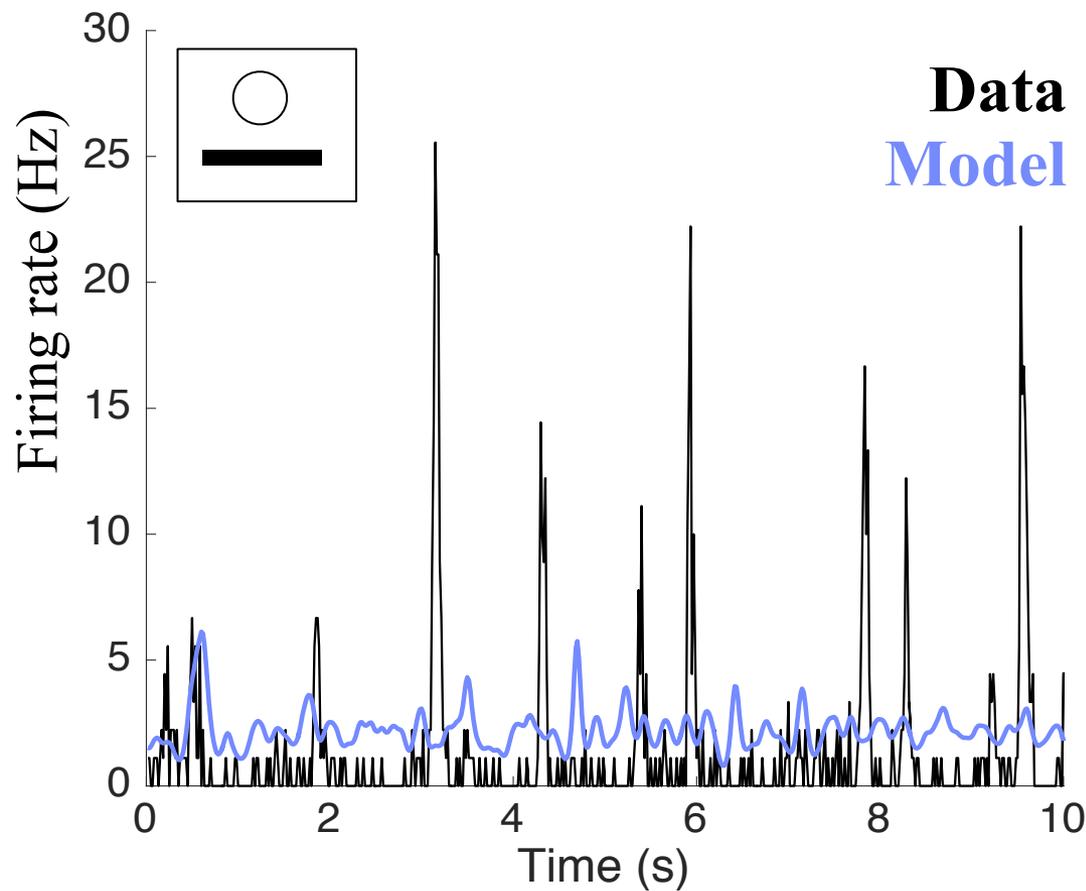
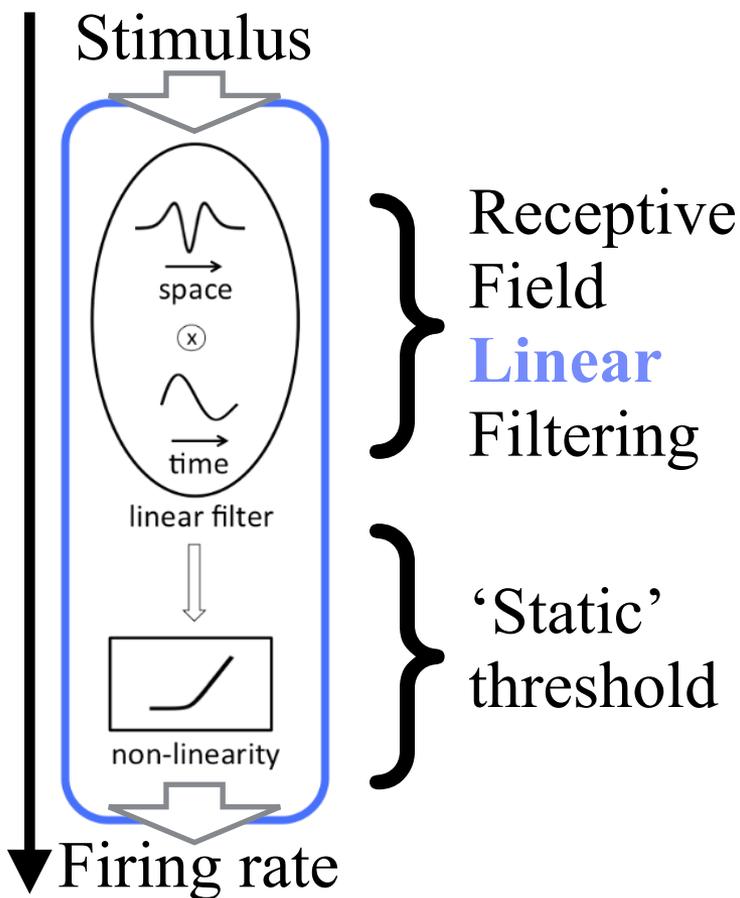
Strategy: input/output modelling

An (almost) *linear* model...



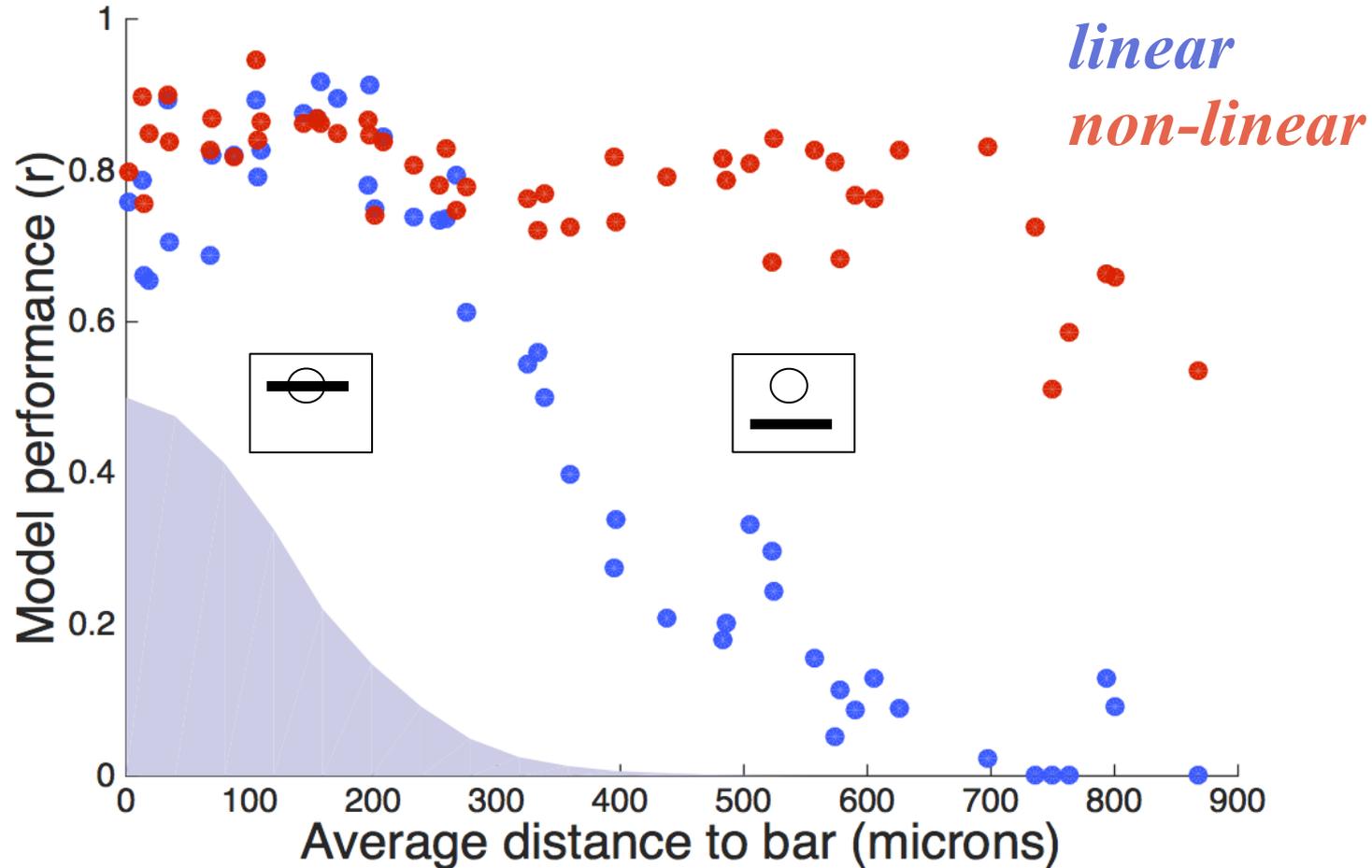
Strategy: input/output modelling

An (almost) *linear* model may not predict the response



Strategy: input/output modelling

An highly *non-linear* model predicts the response



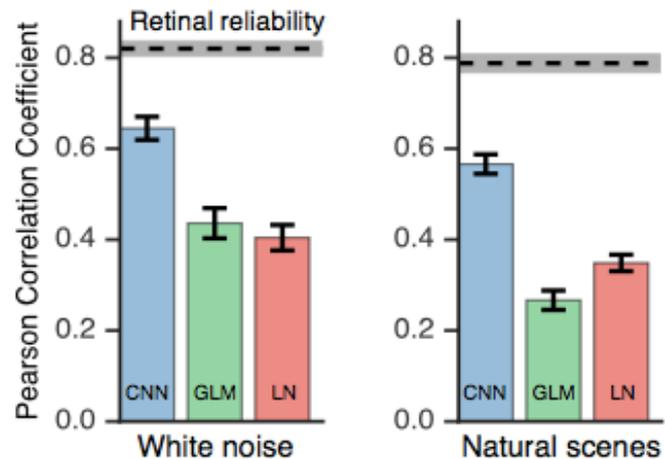
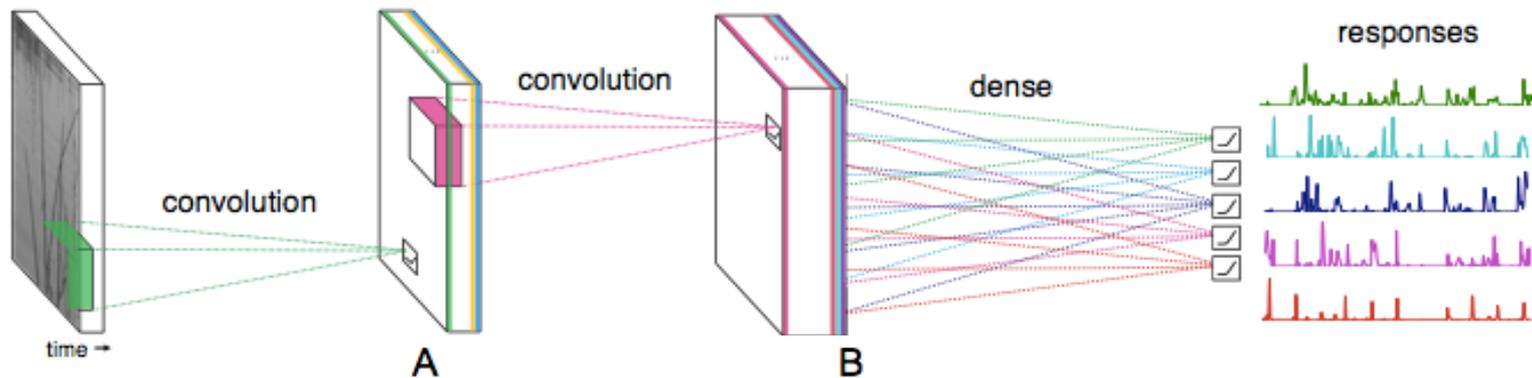
Same cell-type / two different computations

Réseau de neurones artificiels

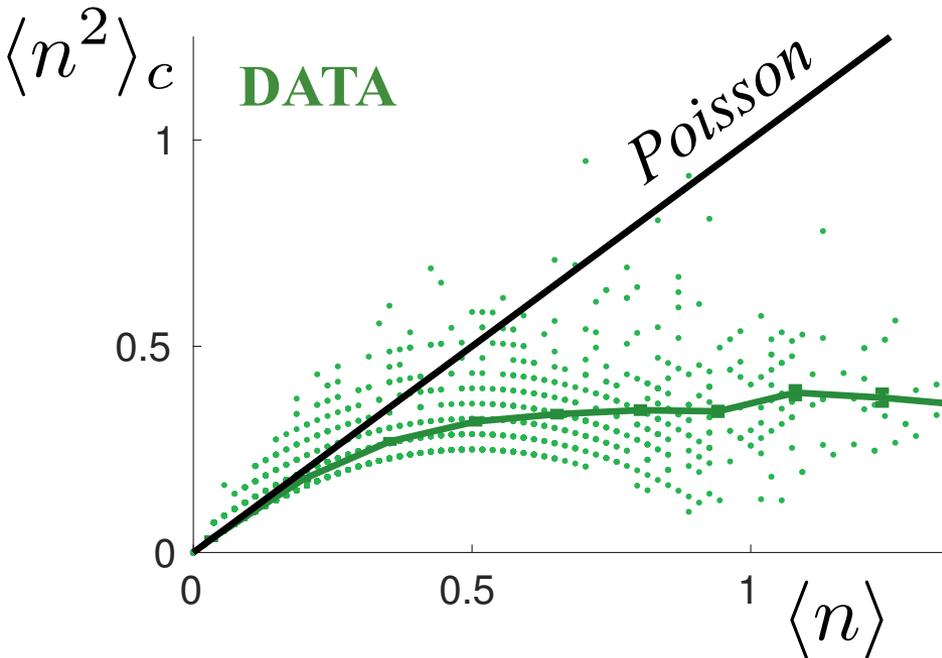
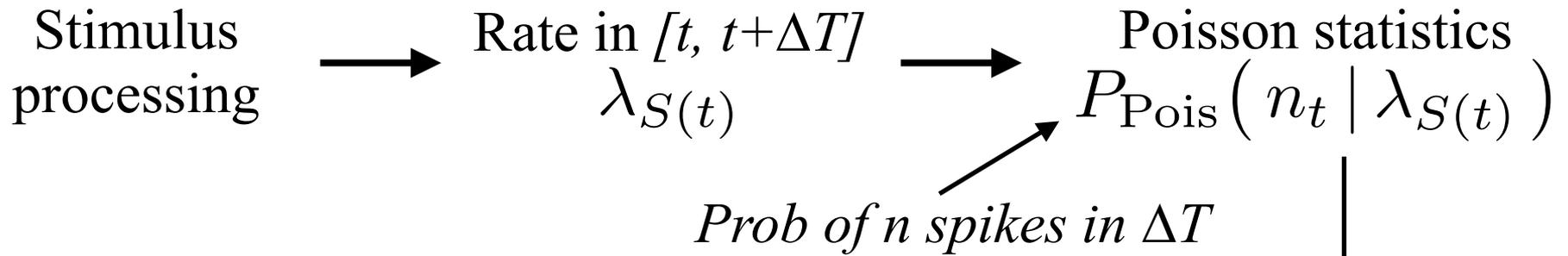
Deep Learning Models of the Retinal Response to Natural Scenes

NIPS 2016

Lane T. McIntosh^{*1}, Niru Maheswaranathan^{*1}, Aran Nayebi¹,
Surya Ganguli^{2,3}, Stephen A. Baccus³

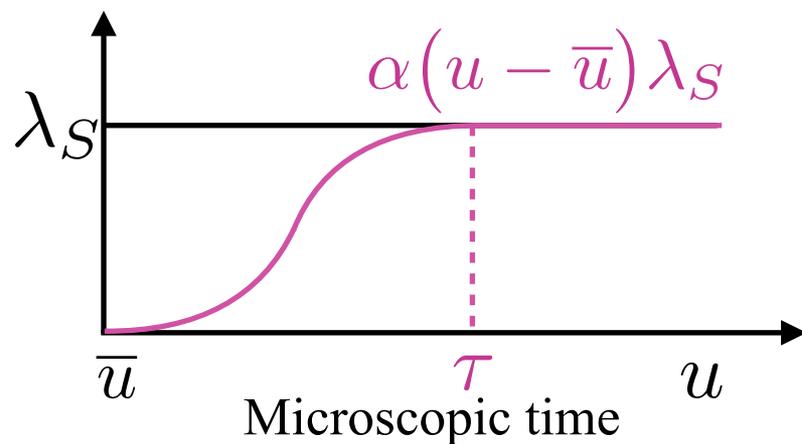
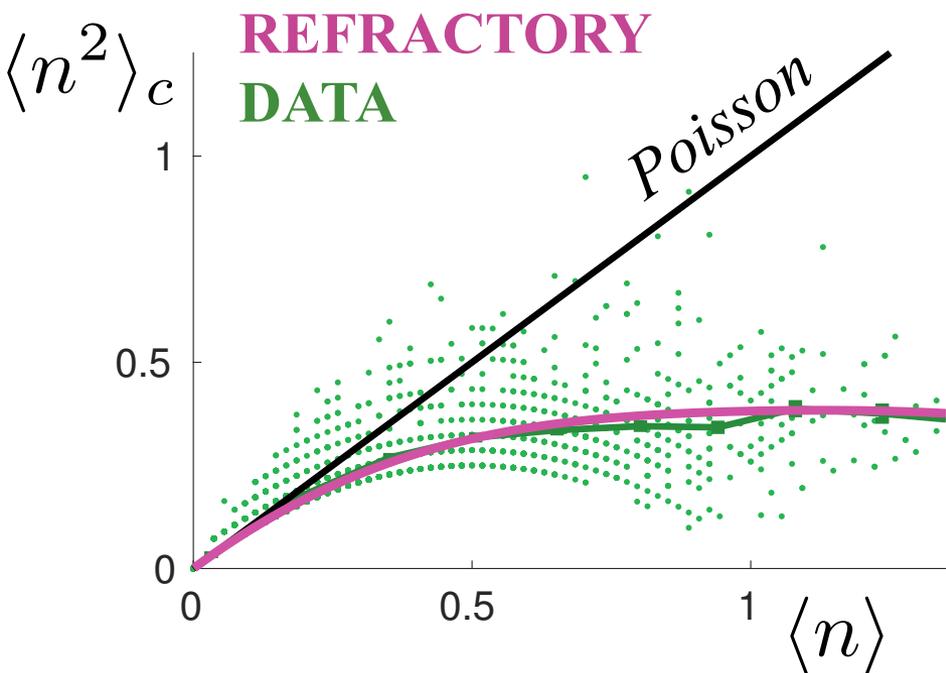
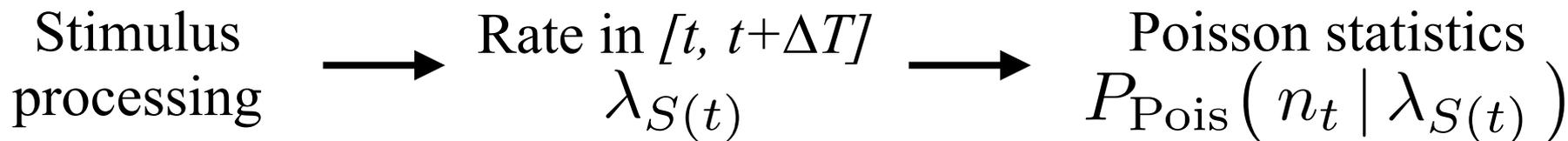


How to include **noise** in spike generation?

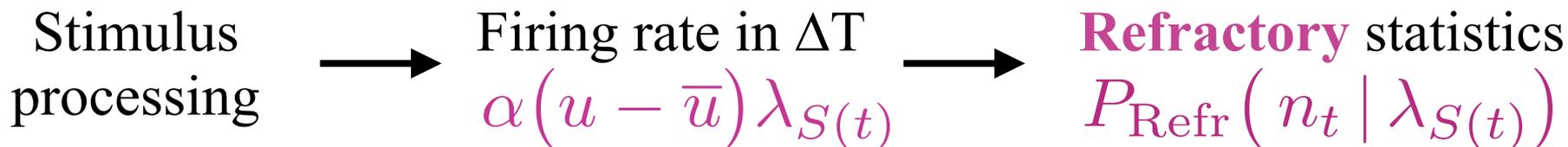


~~$\lambda = \langle n \rangle = \langle n^2 \rangle_c$~~

How to include **noise** in spike generation?



$$P_{\text{Refr}} \tau \approx 0 \approx P_{\text{Pois}} e^{-\gamma n^2 - \delta n^3}$$



How to include **noise** in spike generation?

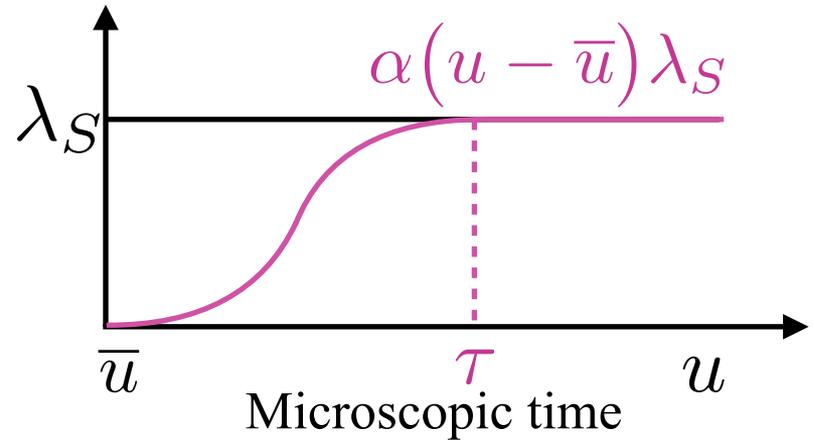
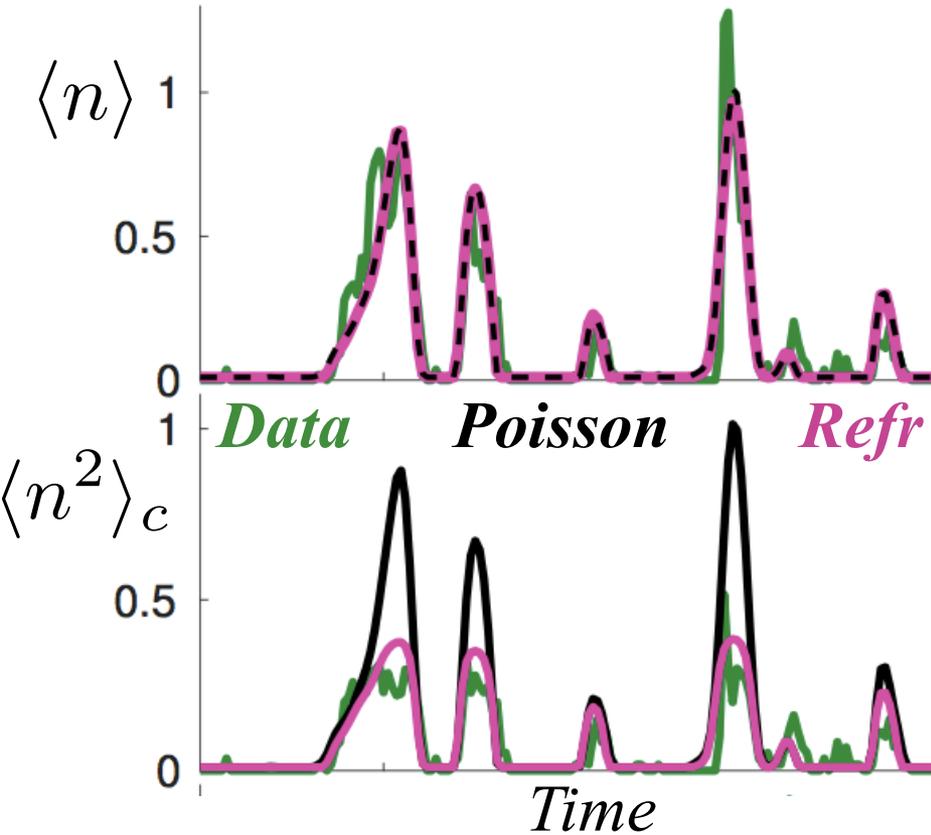
Stimulus processing



Rate in $[t, t+\Delta T]$
 $\lambda_S(t)$



Poisson statistics
 $P_{\text{Pois}}(n_t | \lambda_S(t))$



$$P_{\text{Refr}} \underset{\tau \sim 0}{\sim} P_{\text{Pois}} e^{-\gamma n^2 - \delta n^3}$$

Stimulus processing

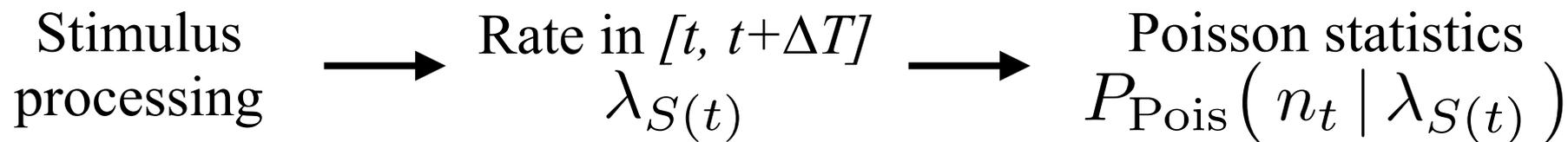


Firing rate in ΔT
 $\alpha(u - \bar{u}) \lambda_S(t)$

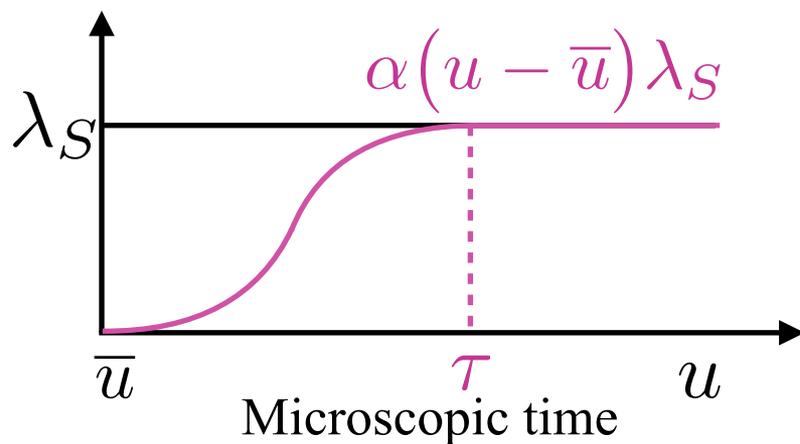
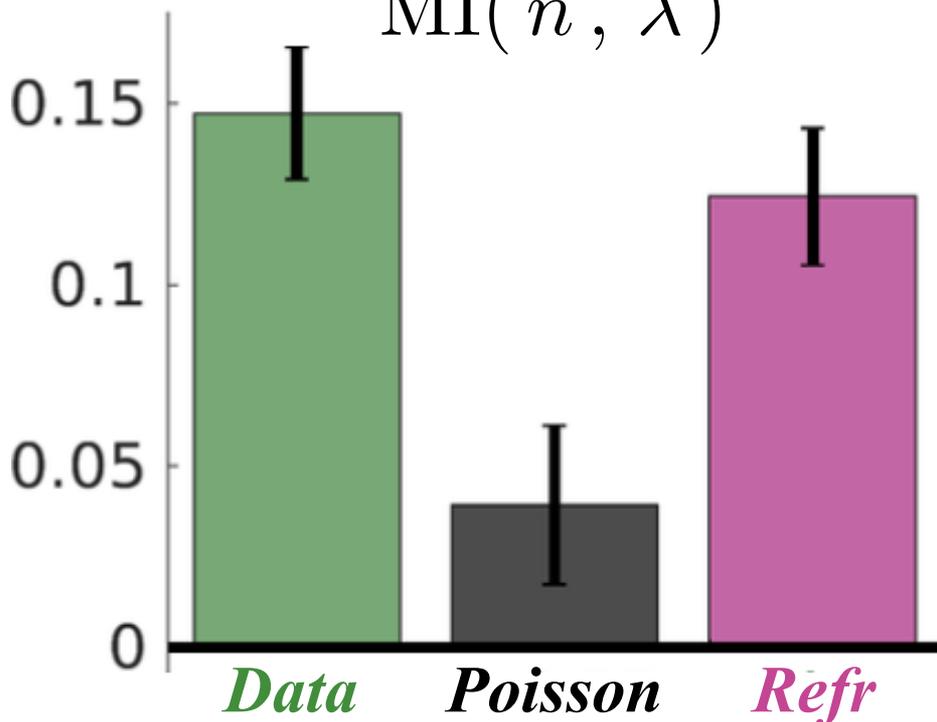


Refractory statistics
 $P_{\text{Refr}}(n_t | \lambda_S(t))$

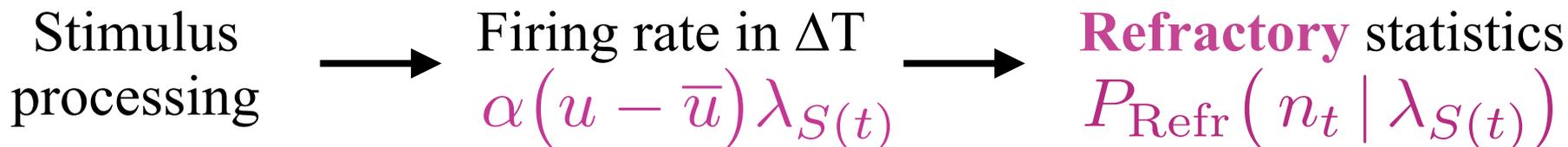
How to include **noise** in spike generation?



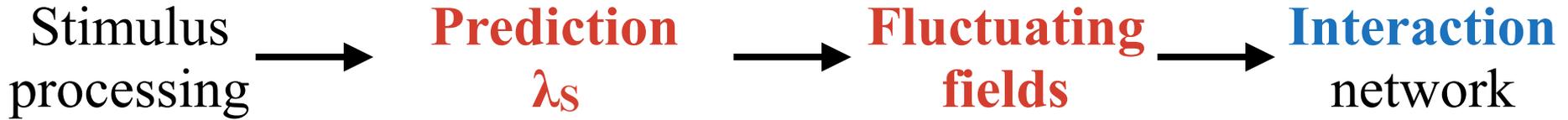
$MI(n, \lambda)$



$$P_{\text{Refr}} \tau \approx 0 \quad P_{\text{Pois}} e^{-\gamma n^2 - \delta n^3}$$



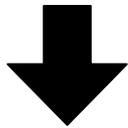
How to account for noise covariances?



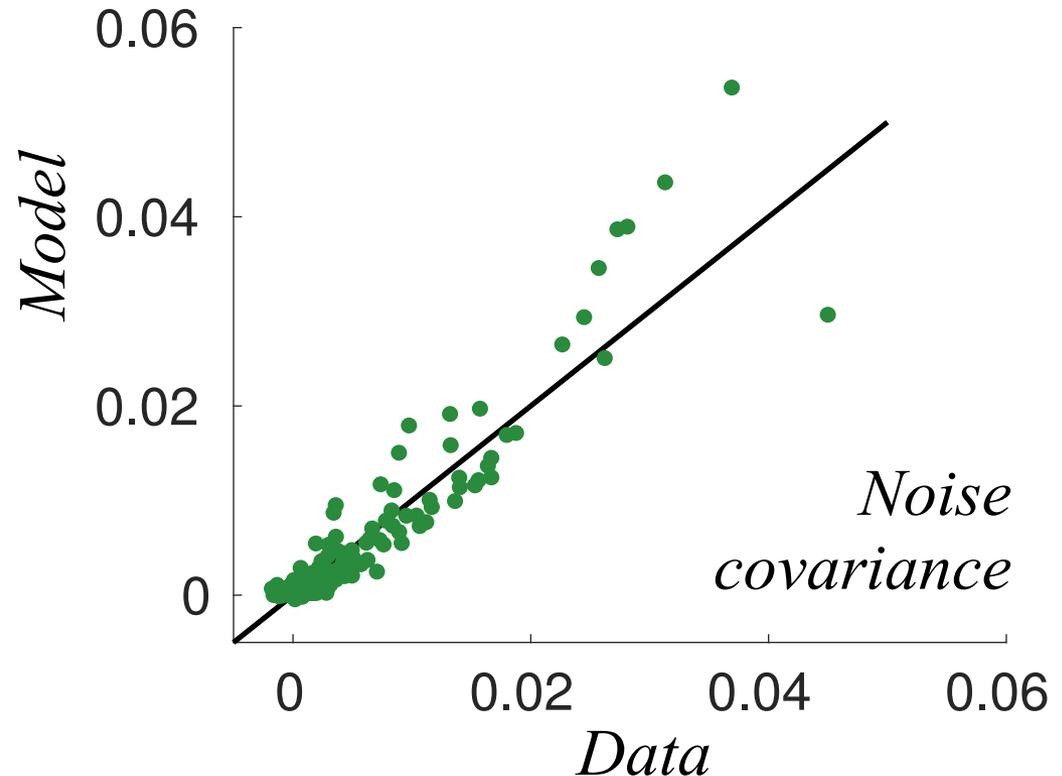
$$P_{\text{Refr}}^{\text{Pop}}(\{n_{it}\}) \sim \prod_t \left(\prod_i e^{h_i[\lambda_{S(t)}] n_{it}} \prod_{i < j} e^{n_{it} J_{ij} n_{jt}} \right)$$

TAP/AMP
correction:

$$h_i[\lambda_{S(t)}]$$



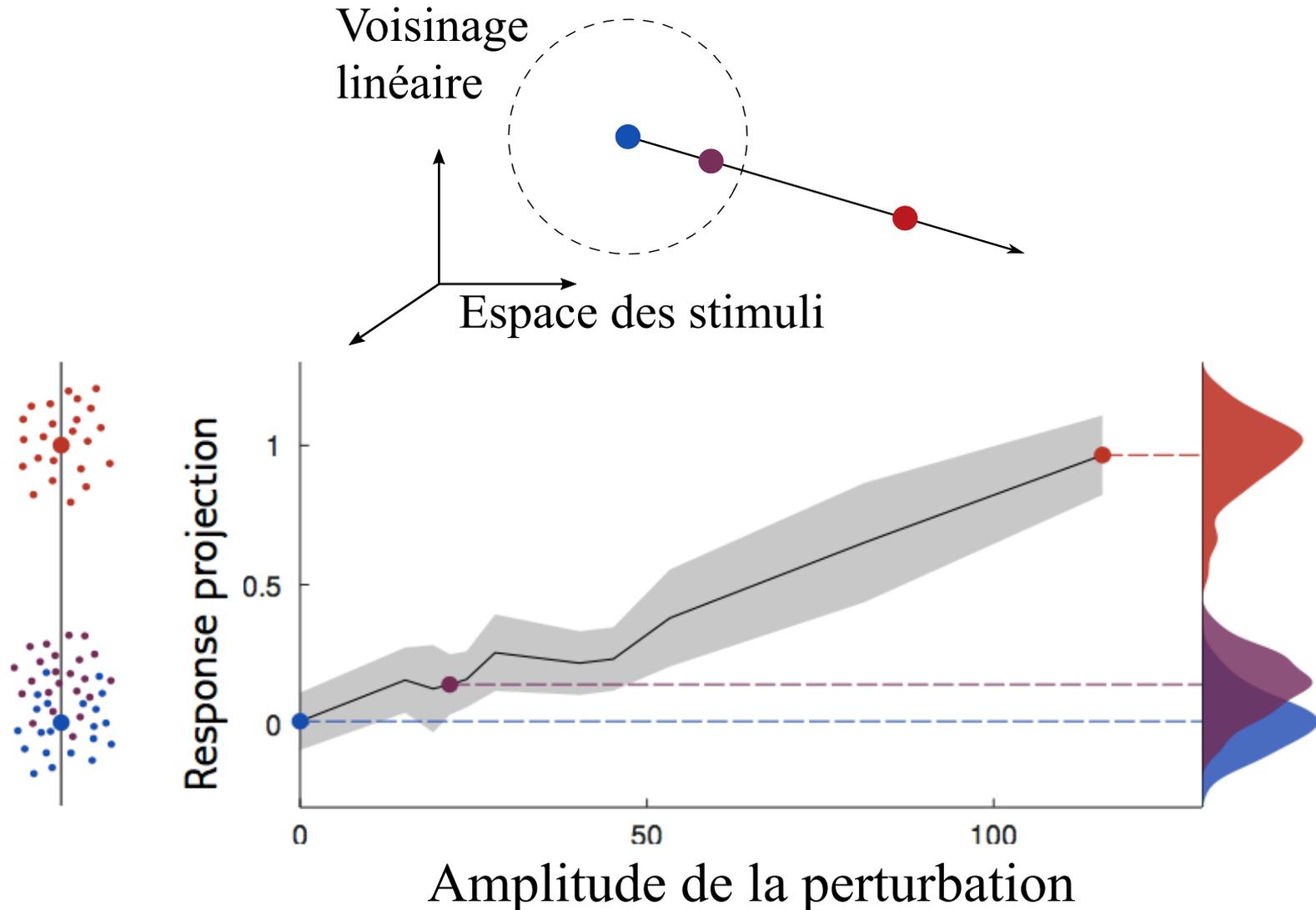
$$h_i[\lambda_{S(t)}] - \Delta h_i[\lambda_{S(t)}, J]$$



Comment décrire le traitement non-linéaire du stimulus?

Stratégie: Développement limité 'expérimental'

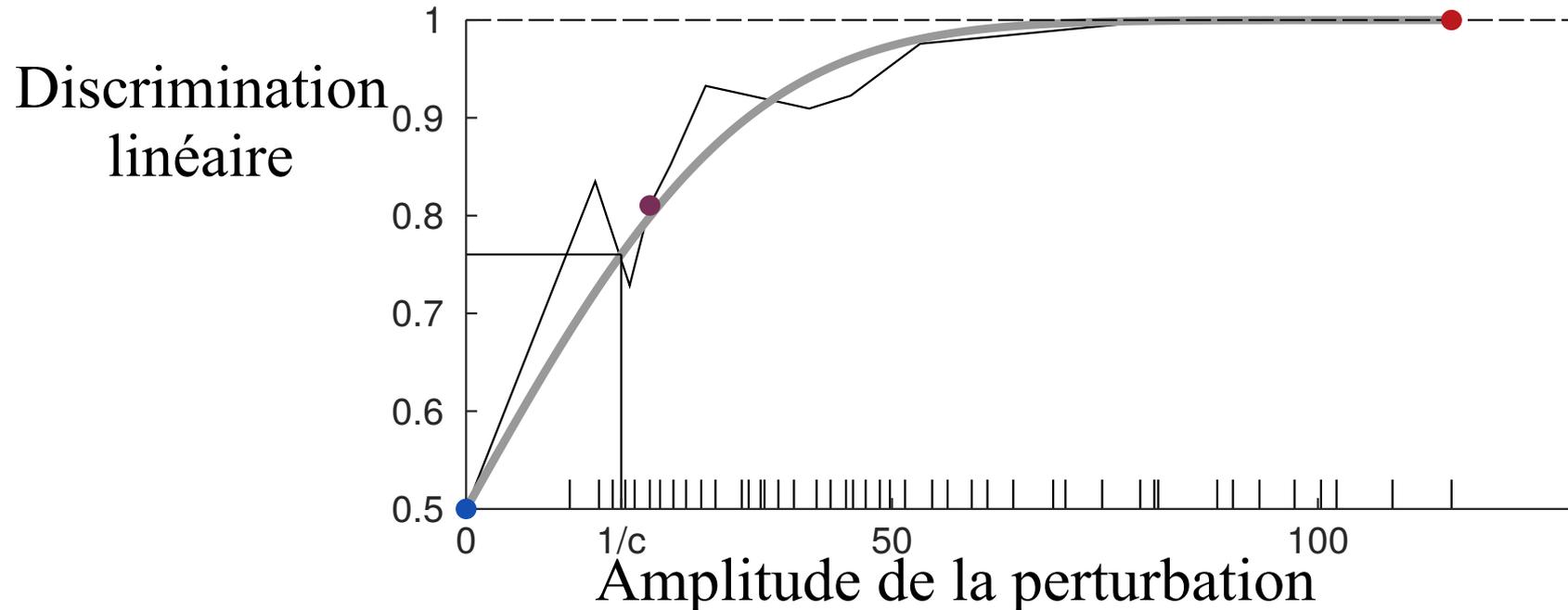
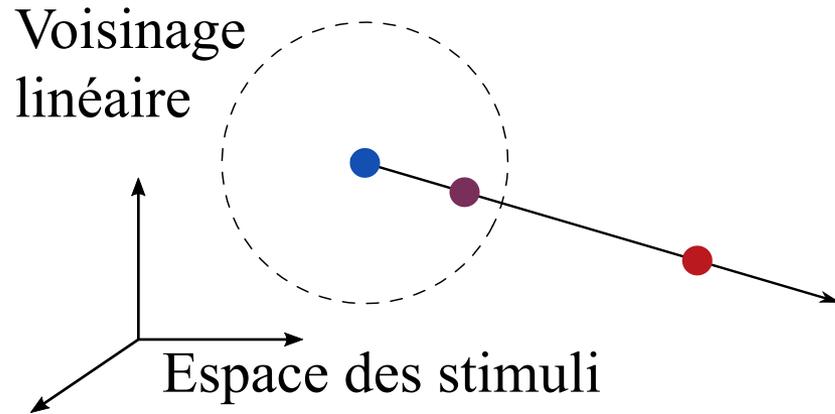
soumis



Comment décrire le traitement non-linéaire du stimulus?

Stratégie: Développement limité 'expérimental'

soumis



Comment décrire le traitement non-linéaire du stimulus?

Stratégie: Développement limité 'experimental'

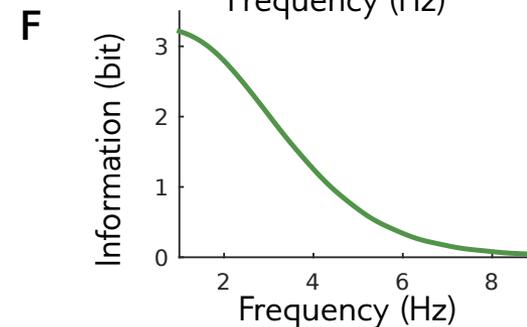
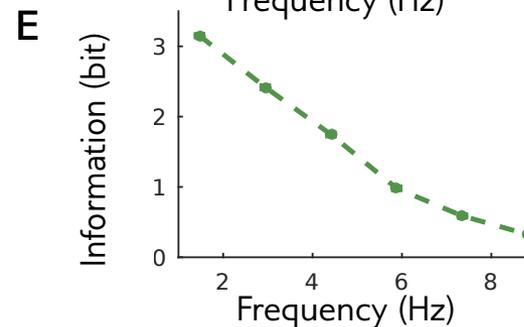
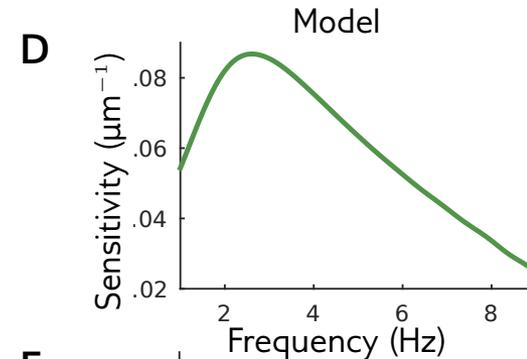
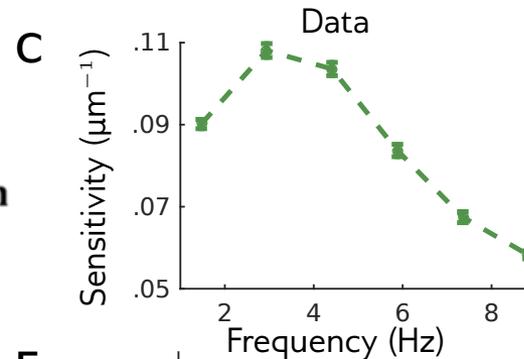
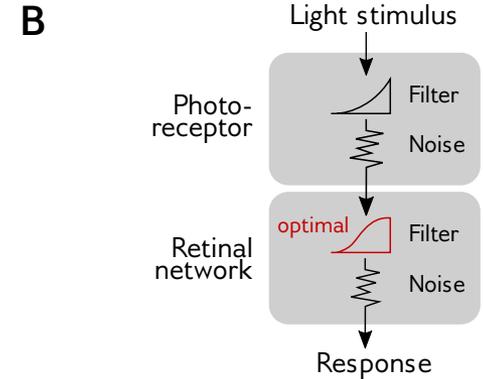
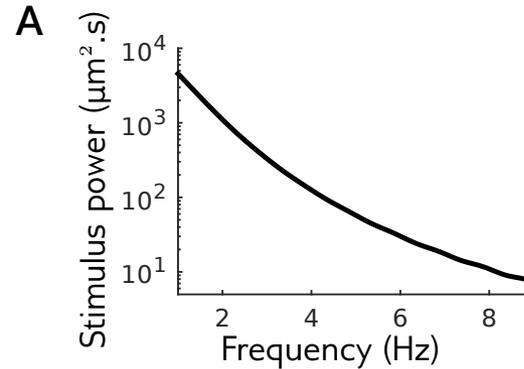
soumis

*Codage efficace:
maximisation de
l'information transmise*

A theory of maximizing sensory information

J. H. van Hateren

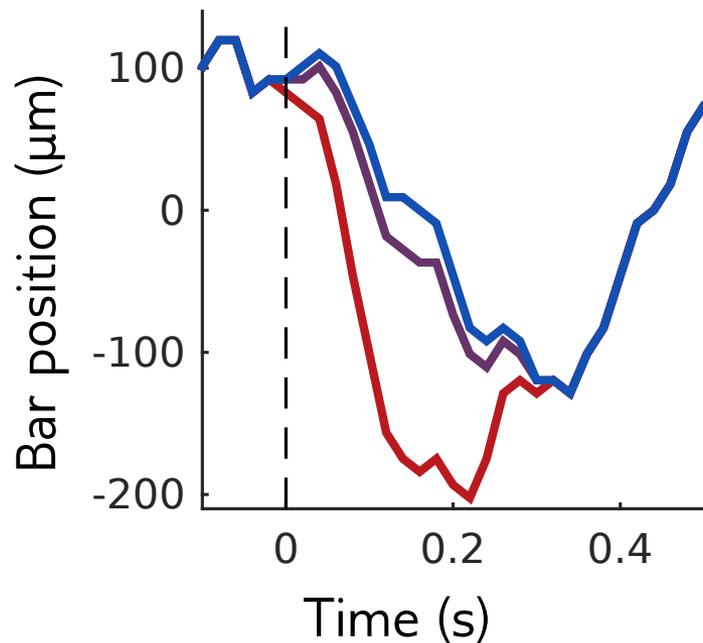
Biol. Cybern 1992



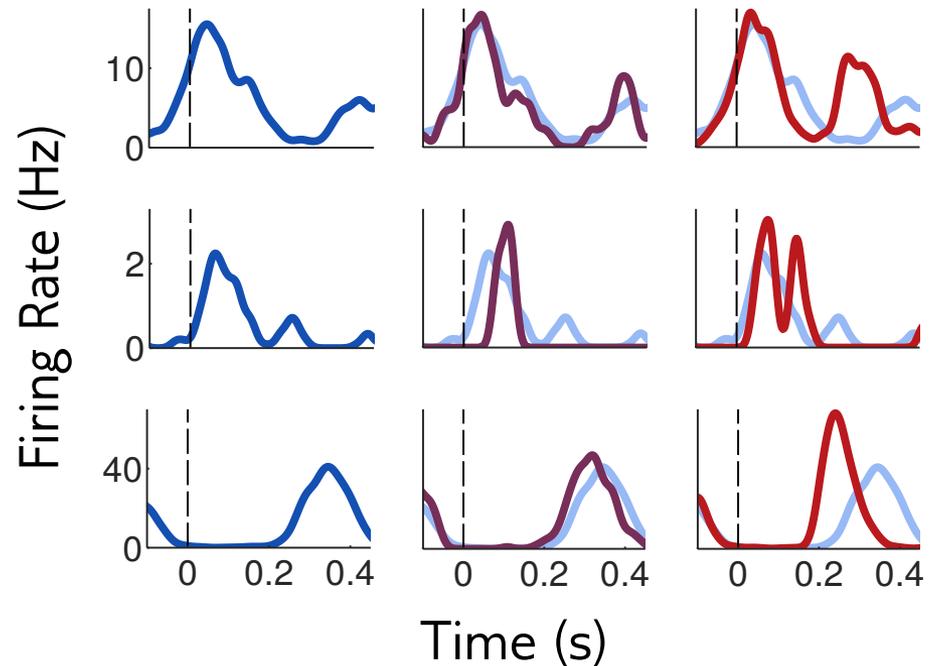
Recherche: Codage par le réseau de la rétine

Rétine: canale de communication

Stimulation visuelle



Réponse rétinienne



L'Information visuelle module l'activité du réseau

*Traitement du signal **hautement non linéaire***