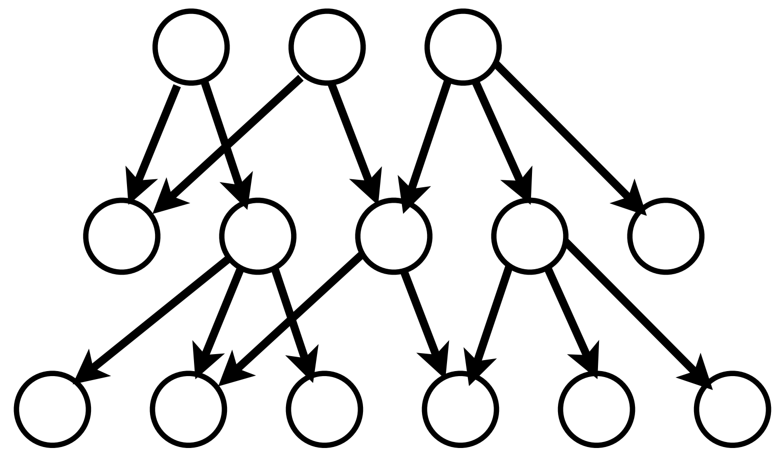
Statisztikus tanulás az idegrendszerben

ORBÁN GERGŐ

http://golab.wigner.mta.hu

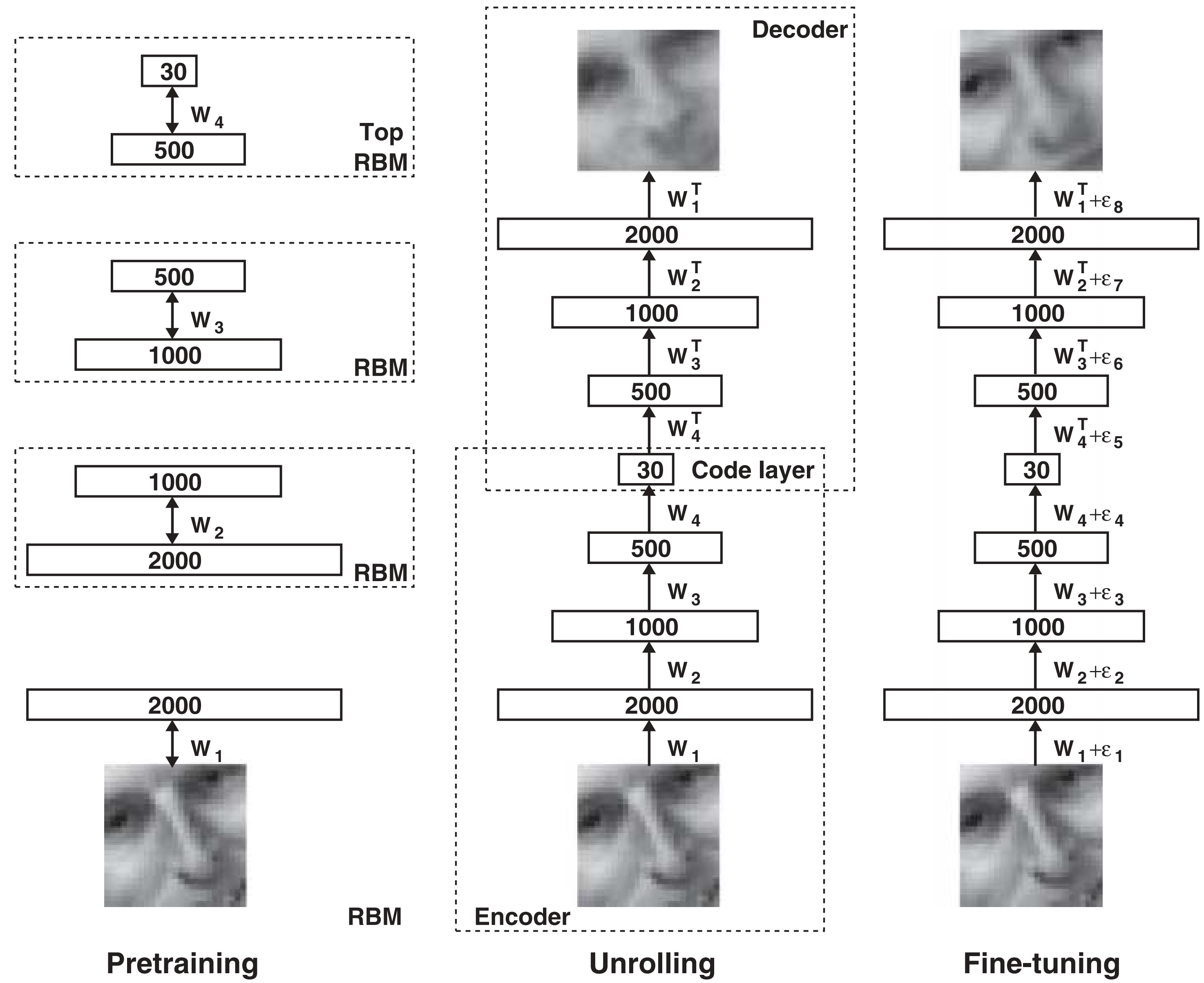
WIGNER RESEARCH CENTRE FOR PHYSICS

Hierarchikus grafikus modellek



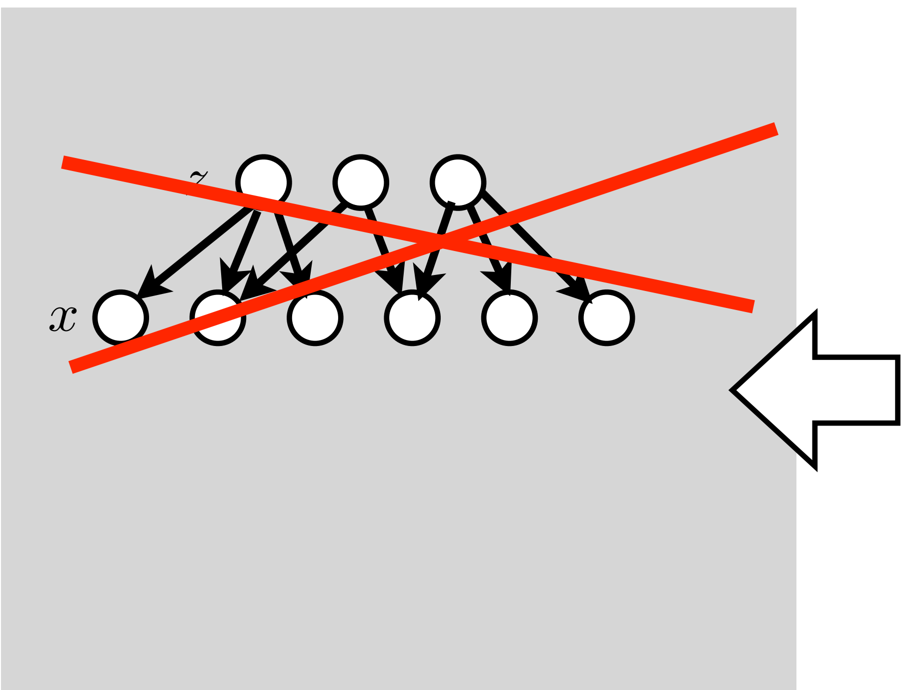
- Nehéz a nemlineáris optimalizálás hierarchikus rendszerekben:
- Amennyiben erős függéseket tételezek fel, akkor lokális minimumokban ragad meg a tanulás
- Amennyiben gyenge függéseket tételezek fel, akkor a grádiensek a rétegek között egészen elenyészővé válnak és a hálózat számára nincsen “szignál” ami alapján tanulni tudna

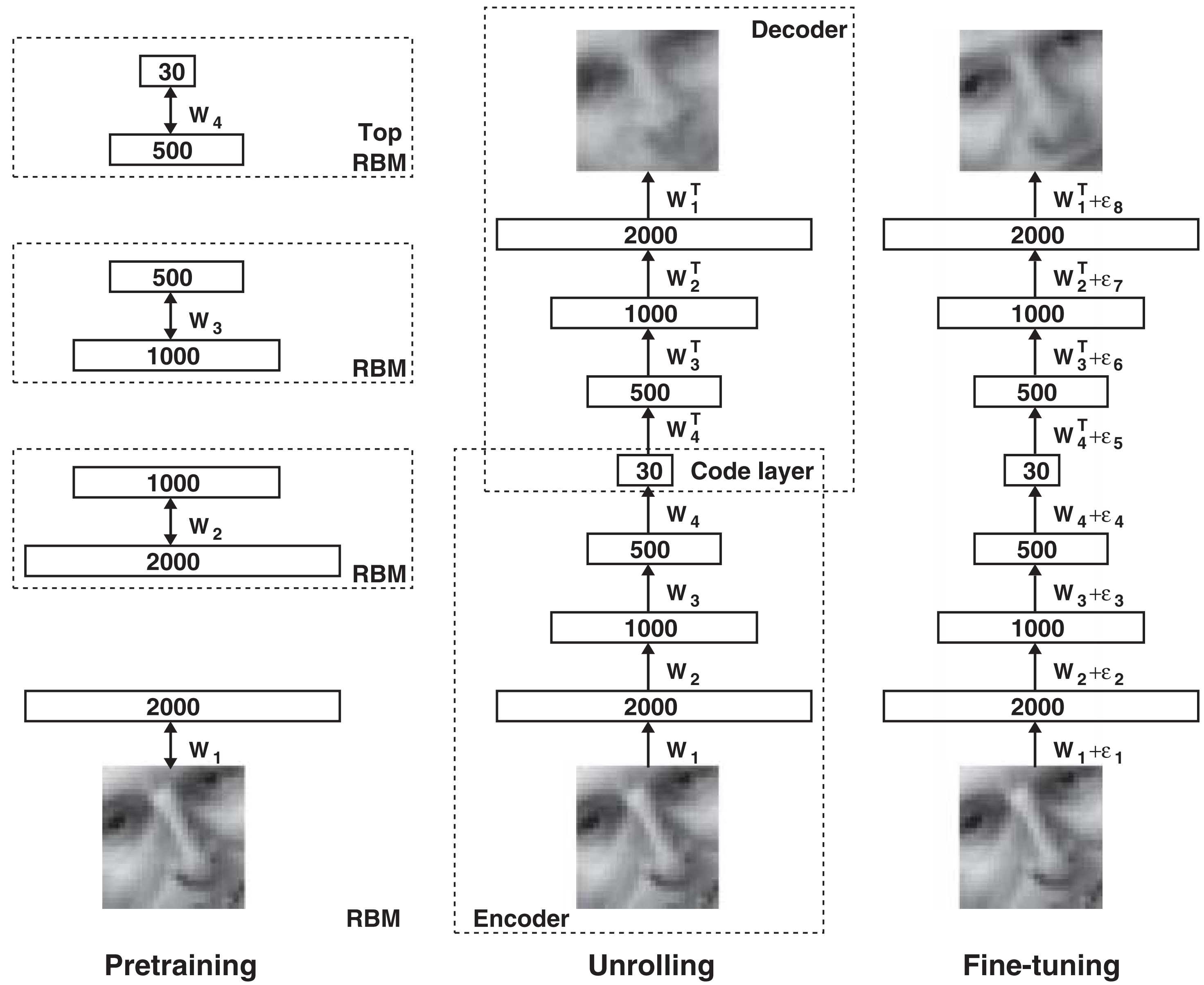
Deep Belief Networks



Hinton & Salakhutdinov, 2006

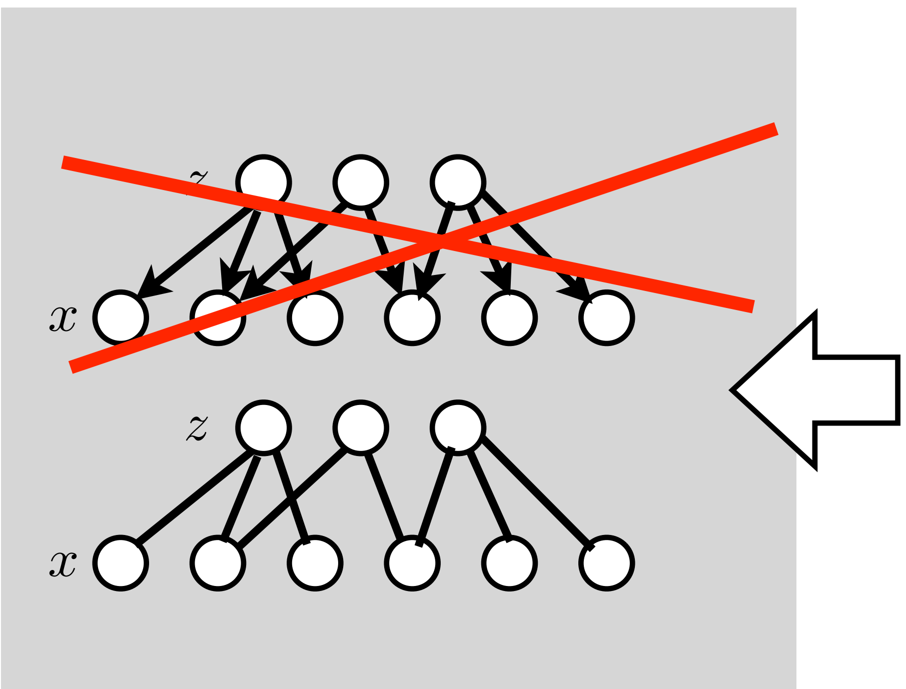
Deep Belief Networks

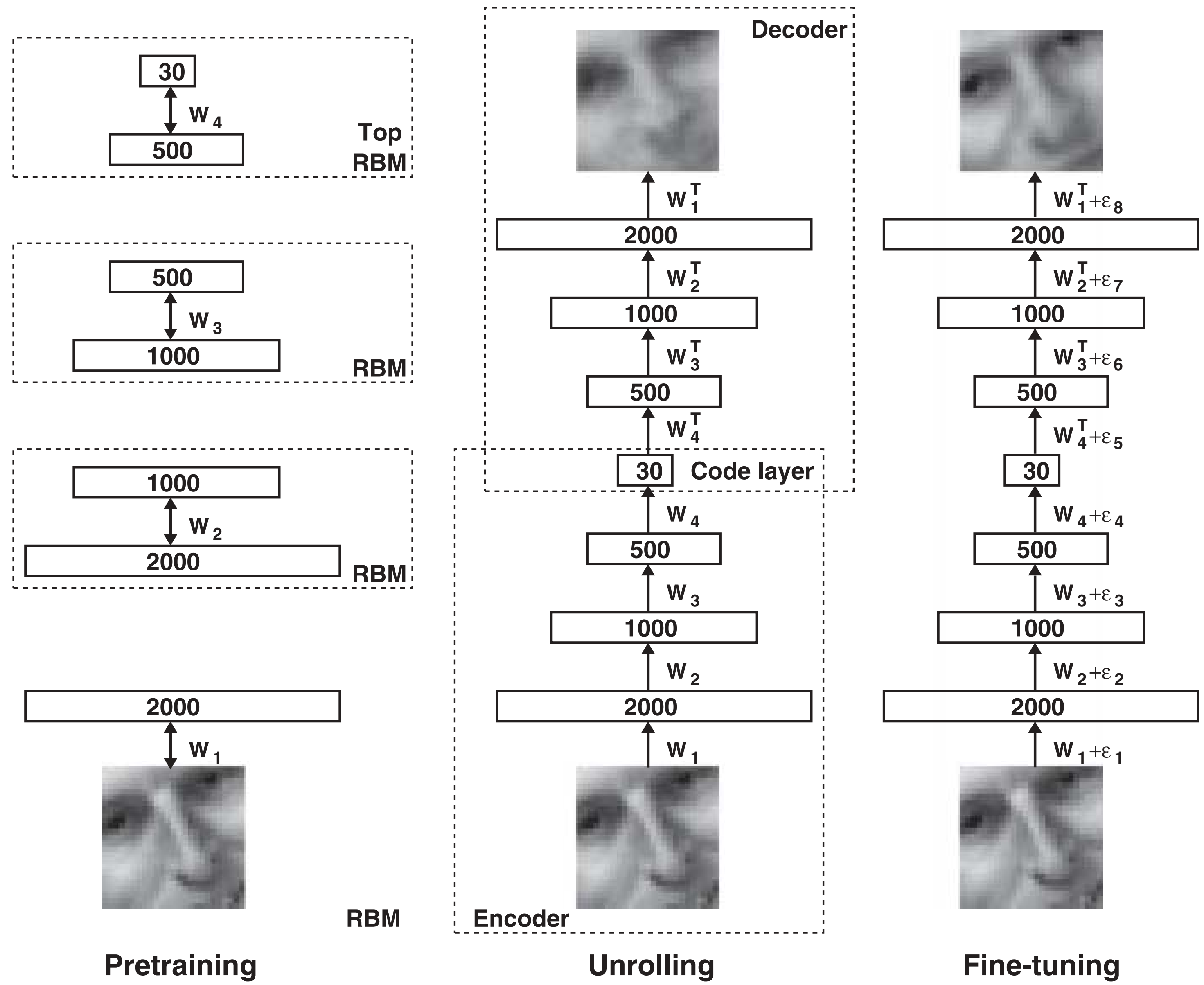




Hinton & Salakhutdinov, 2006

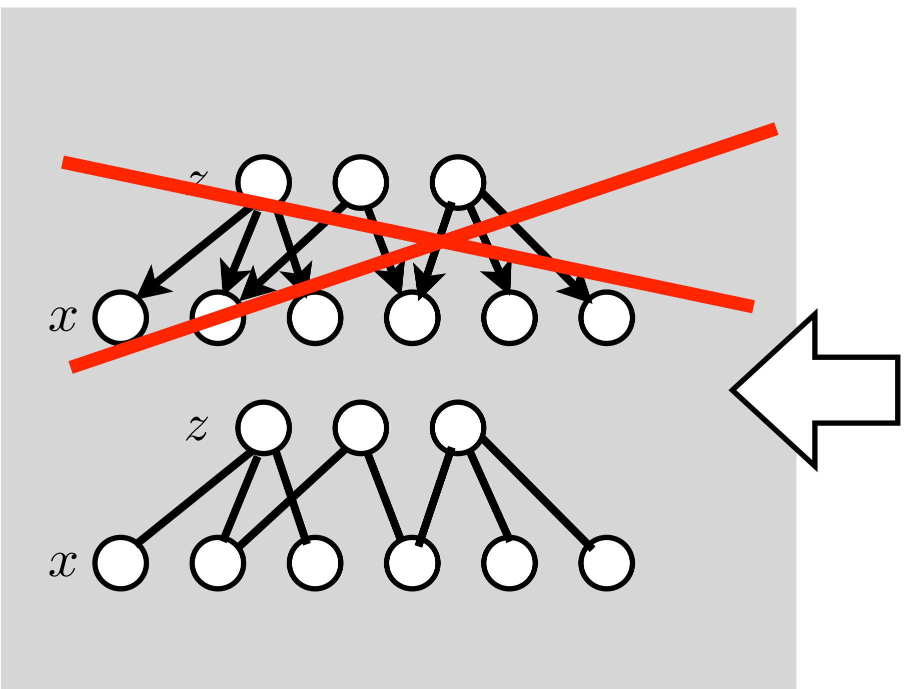
Deep Belief Networks

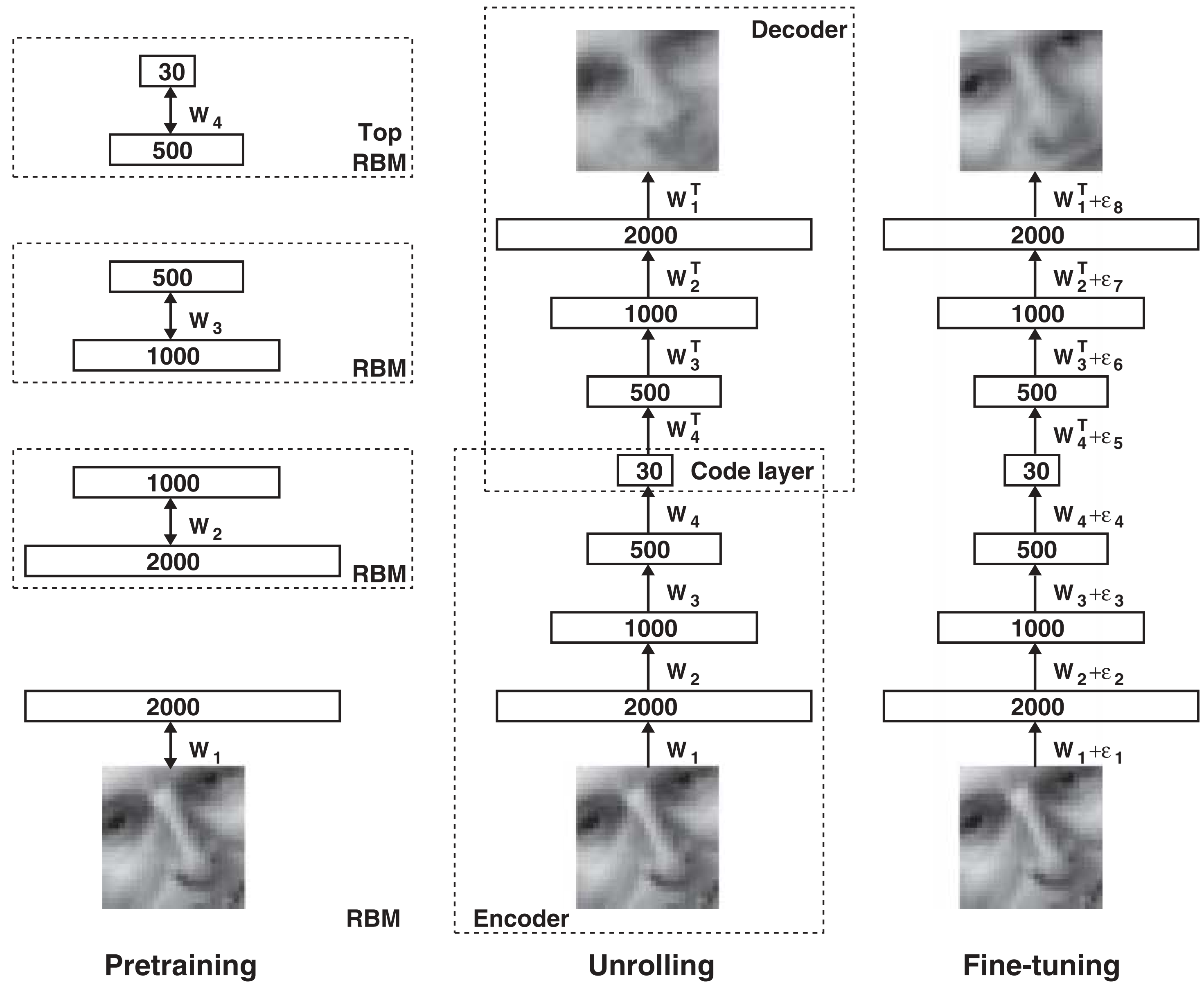




Hinton & Salakhutdinov, 2006

Deep Belief Networks





Hinton & Salakhutdinov, 2006

Energia függvény:

$$E(\mathbf{v}, \mathbf{h}) = -\sum_{i \in \text{pixels}} b_i v_i - \sum_{j \in \text{features}} b_j h_j - \sum_{i,j} v_i h_j w_{ij}$$

Aktivációs függvény:

$$P(h_j \mid \mathbf{v}) = \text{sigmoid}(b_j + \sum_i v_i w_{ij})$$

Deep Belief Networks - pretraining

Energia függvény:

$$E(\mathbf{v}, \mathbf{h}) = -\sum_{i \in \text{pixels}} b_i v_i - \sum_{j \in \text{features}} b_j h_j - \sum_{i,j} v_i h_j w_{ij}$$

Aktivációs függvény:

$$P(h_j \mid \mathbf{v}) = \text{sigmoid}(b_j + \sum_i v_i w_{ij})$$

Deep Belief Networks - pretraining

Energia függvény:

$$E(\mathbf{v}, \mathbf{h}) = -\sum_{i \in \text{pixels}} b_i v_i - \sum_{j \in \text{features}} b_j h_j - \sum_{i,j} v_i h_j w_{ij}$$

Aktivációs függvény:

$$P(h_j \mid \mathbf{v}) = \text{sigmoid}(b_j + \sum_i v_i w_{ij})$$

Kontrasztív divergencia:

$$\Delta w_{ij} = \varepsilon\left(\langle v_i h_j \rangle_{\text{data}} - \langle v_i h_j \rangle_{\text{recon}}\right)$$

Deep Belief Networks - pretraining

Energia függvény:

$$E(\mathbf{v},\mathbf{h}) = -\sum_{i\in \text{pixels}} b_i v_i - \sum_{j\in \text{features}} b_j h_j - \sum_{i,j} v_i h_j w_{ij}$$

Aktivációs függvény:

$$P(h_j \mid \mathbf{v}) = \mathrm{sigmoid}(b_j + \sum_i v_i w_{ij})$$

Kontrasztív divergencia:

$$\Delta w_{ij} = \varepsilon\left(\langle v_i h_j\rangle_{\text{data}} - \langle v_i h_j\rangle_{\text{recon}}\right)$$

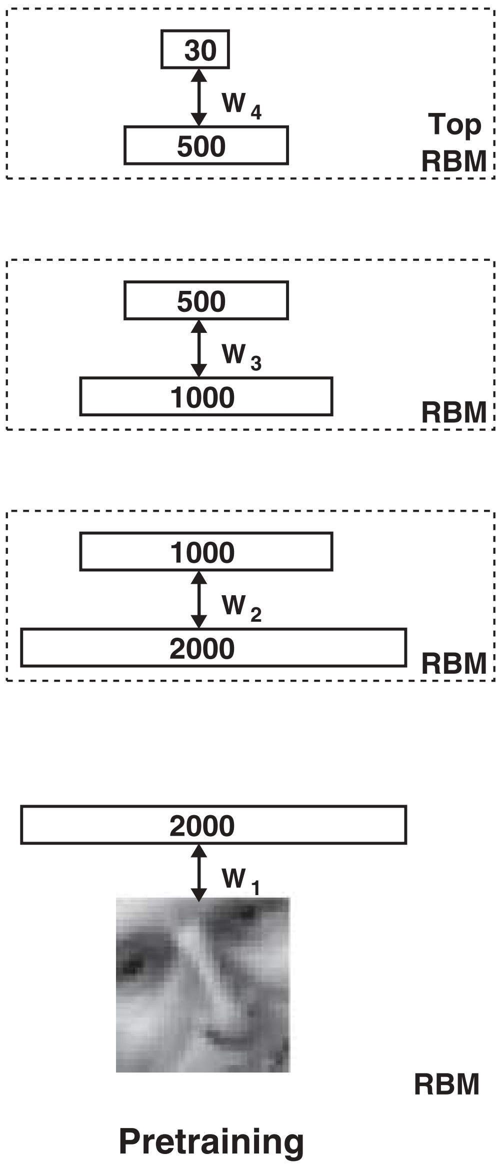
Konfabuláció:

$$P(v_i \mid \mathbf{h}) = \mathrm{sigmoid}(b_i + \sum_j h_j w_{ij})$$

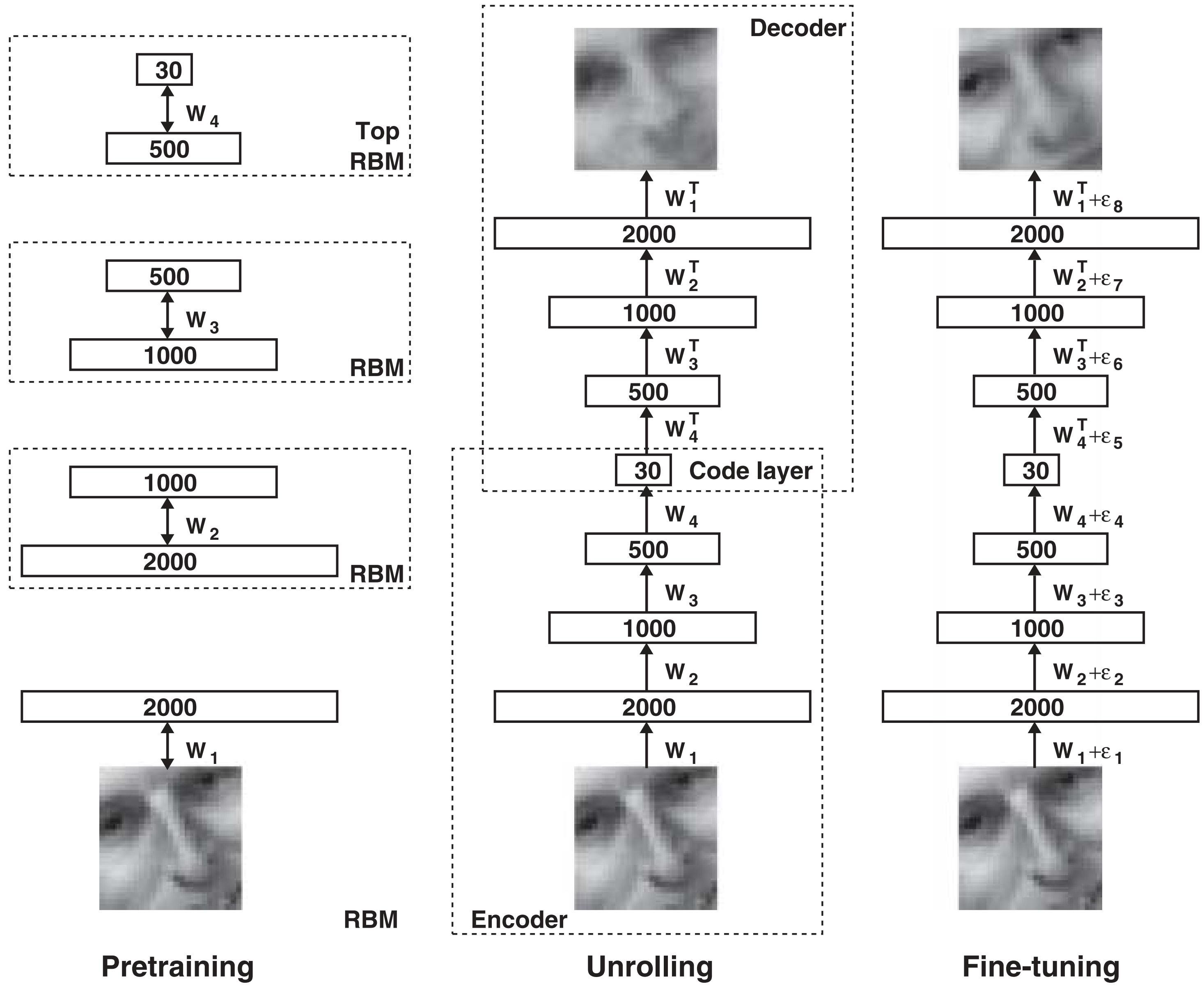
a kontrasztív divergencia második tagjában a látens súlyokat is frissíteni kell a konfabulált 'v'-knek megfelelően

Deep Belief Networks - greedy learning

- Rétegről rétegre végzem a tanulást
- Újabb réteg hozzáadásával a tréning adat likelihoodjának alsó korlátja növekszik (ha a látensek számát nem csökkentjük)
- Folytonos adat esetén a legalsó rétegben Normál eloszlású neuronokat lehet használni (egység variancia esetén a látens frissítése megegyezik a bináris esettel)

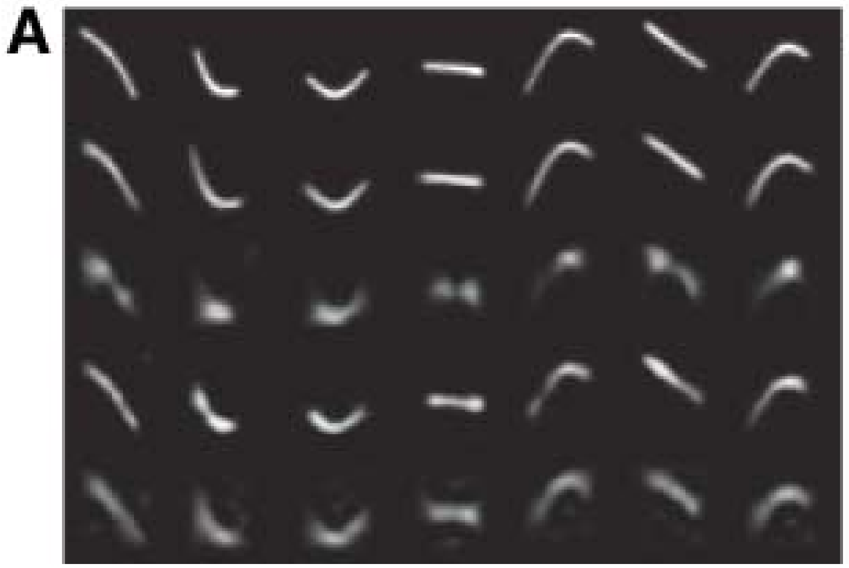


Deep Belief Networks - fine tuning



- Az RBM előtréningezés jó kezdeti feltételeket ad további tréningezéshez
- Sztenderd back propagation algoritmussal tréningezhető innen a hálózat

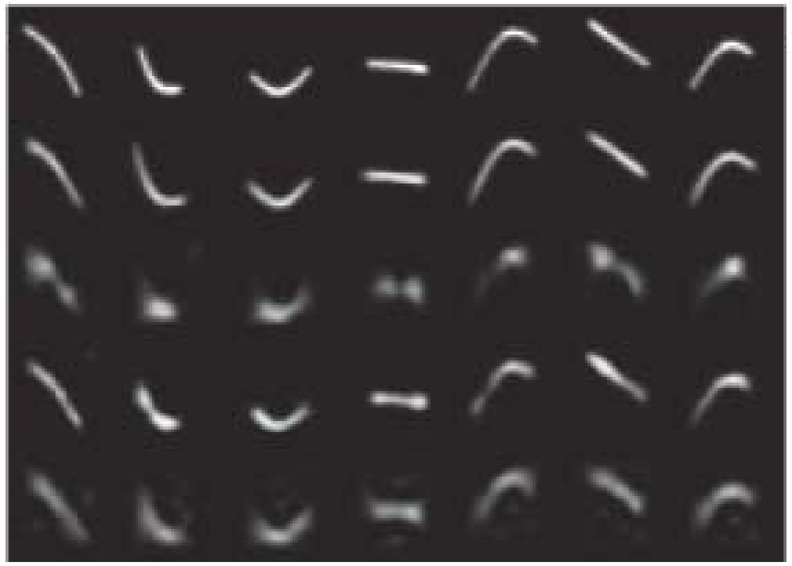
Deep Belief Networks - tanulás



- Szintetikus adat: íves görbék
- Hálózat: (28x28)-400-200-100-50-25-6
- Sorok: rekonstrukció tanulás után a 6 dimenziós DBN-nel; 6D szigmoid-PCA; 18D szigmoid-PCA, 18D sztenderd PCA

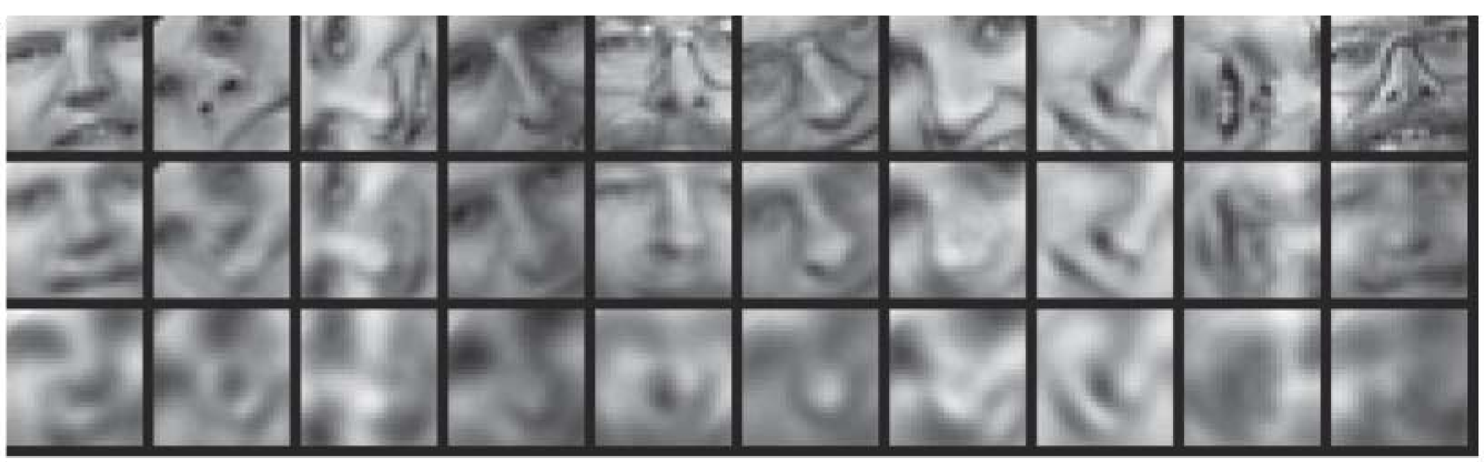
Deep Belief Networks - tanulás

A



- Szintetikus adat: íves görbék
- Hálózat: (28x28)-400-200-100-50-25-6
- Sorok: rekonstrukció tanulás után a 6 dimenziós DBN-nel; 6D szigmoid-PCA; 18D szigmoid-PCA, 18D sztenderd PCA

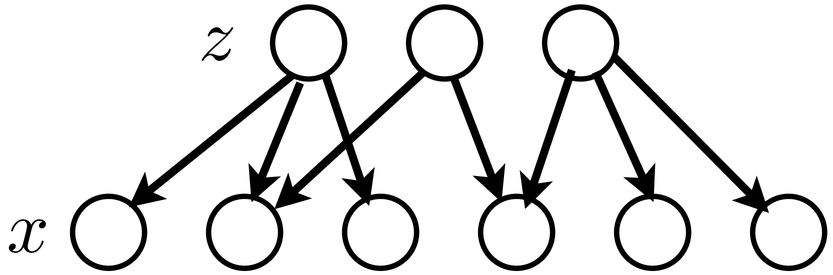
C



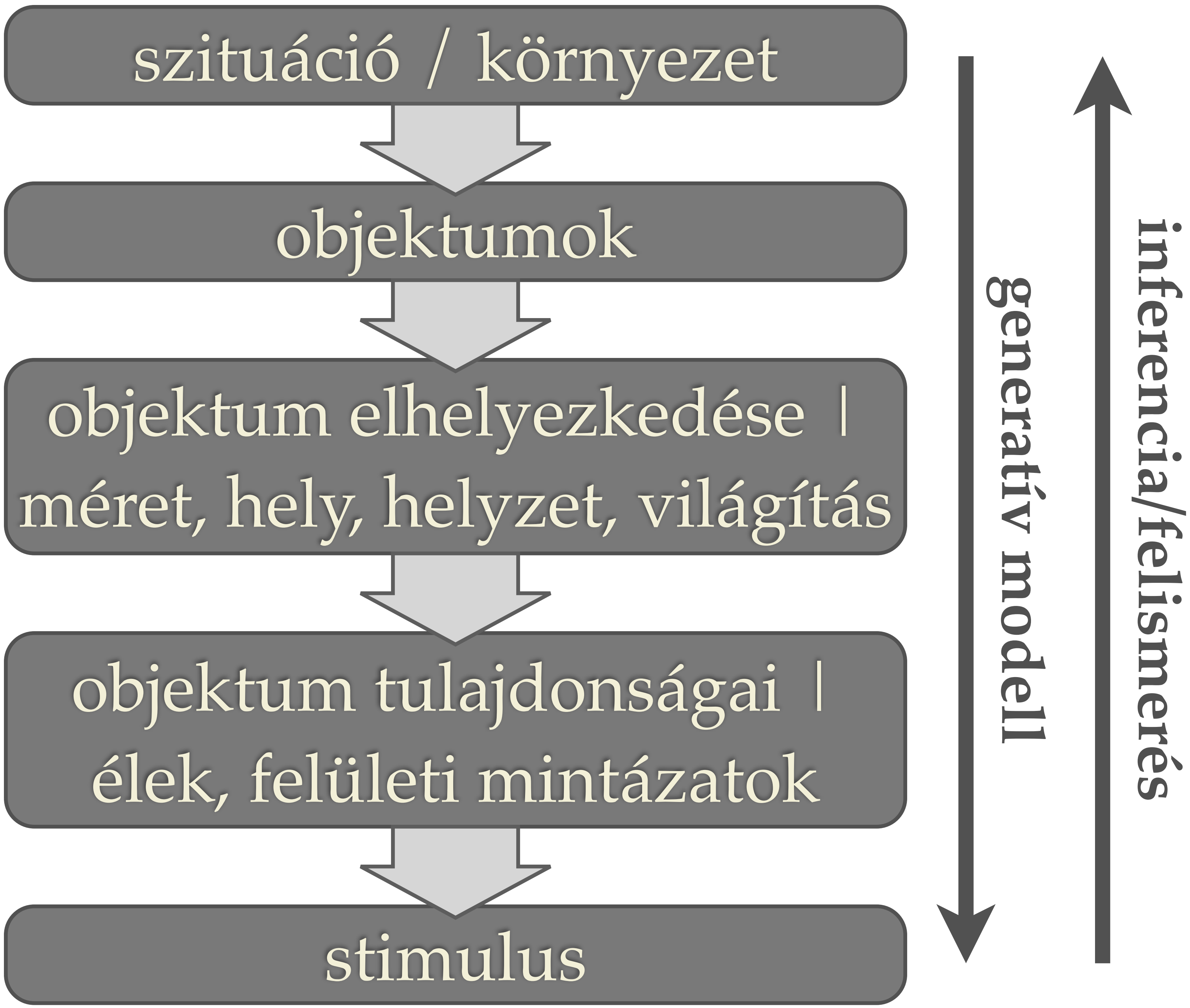
- Természetes képek: arcok
- Hálózat: (25x25)-2000-1000-500-30
- Sorok: rekonstrukció tanulás után a 30D DBN-nel; rekonstrukció 30D PCA

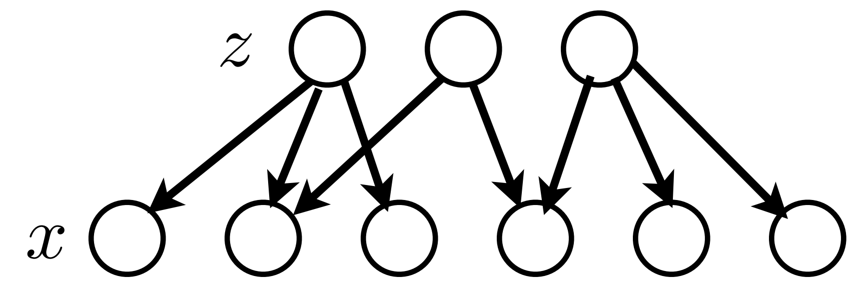
Bayes inferencia

Bayes inferencia

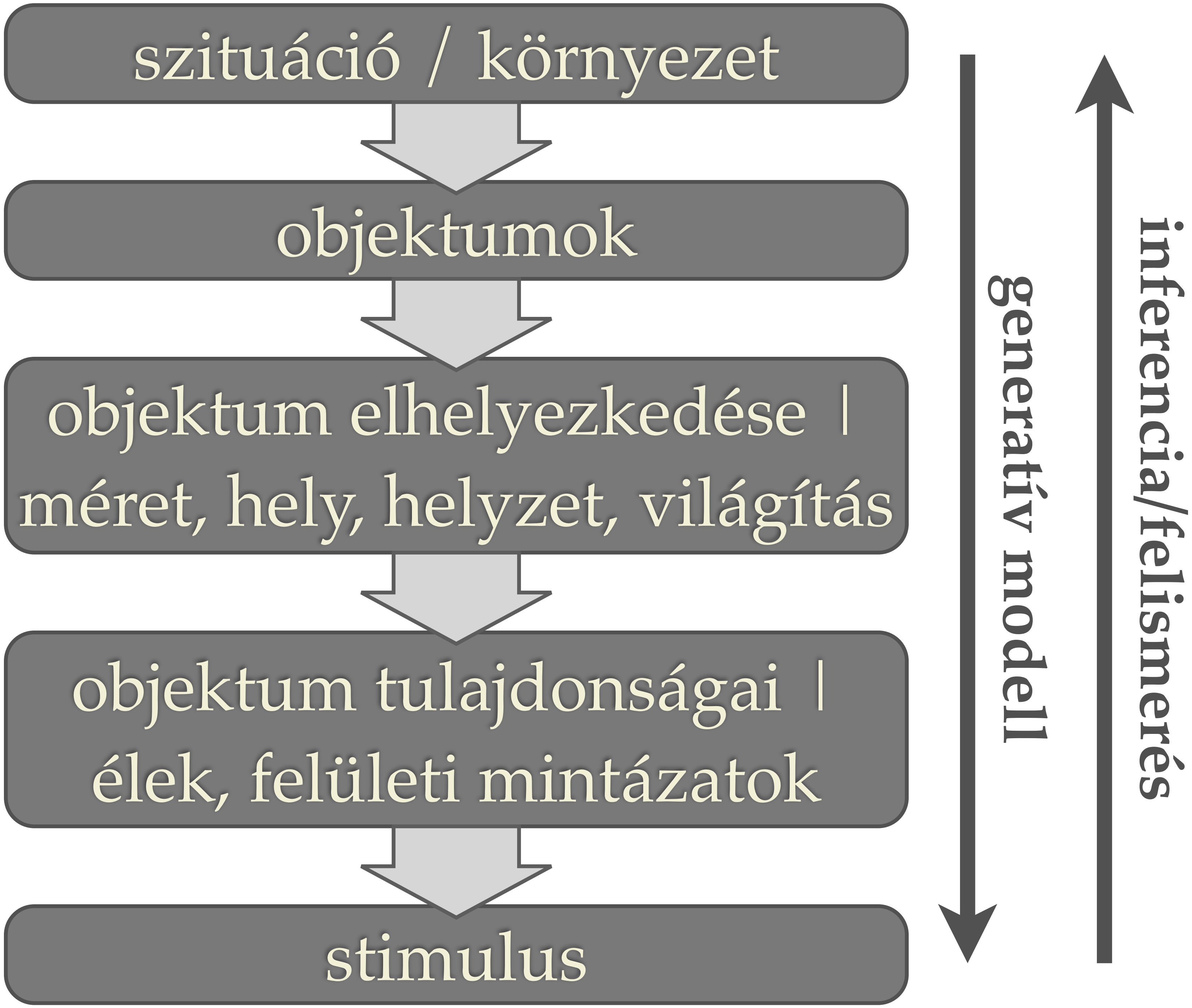


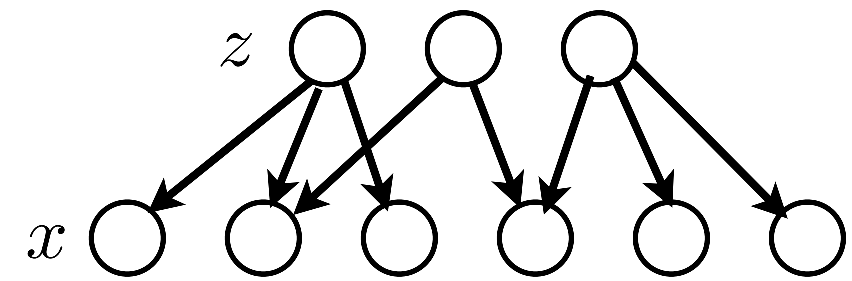
Bayes inferencia





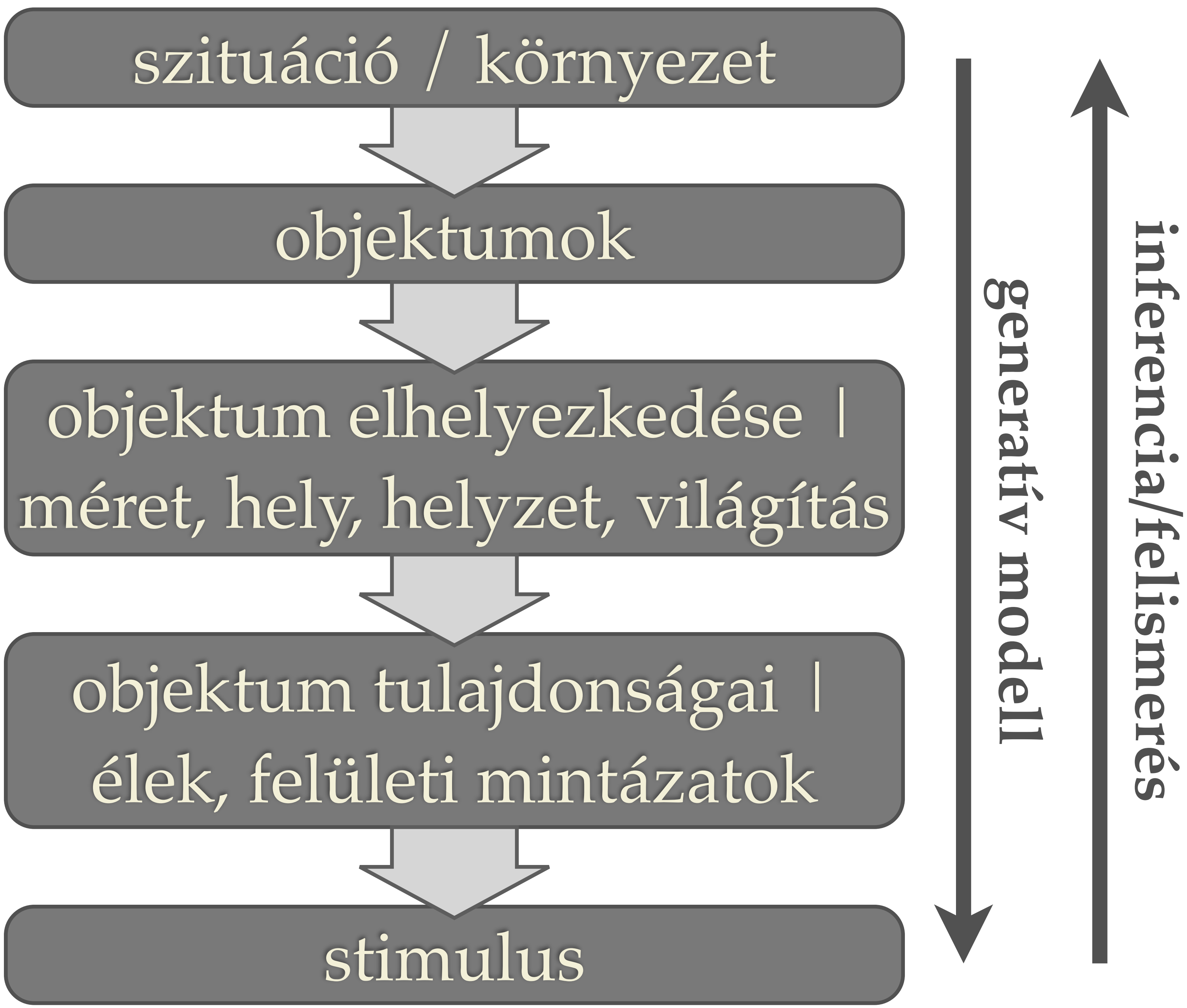
Bayes inferencia

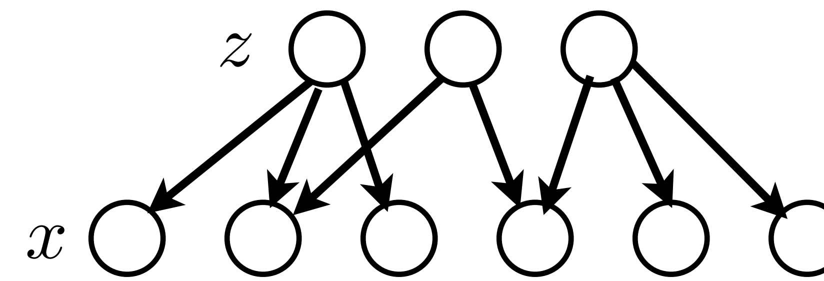




$$\mathrm{P}\left(\mathbf{z}\,|\,\mathbf{x}\right) \propto \mathrm{P}\left(\mathbf{x}\,|\,\mathbf{z}\right)\mathrm{P}\left(\mathbf{z}\right)$$

Bayes inferencia

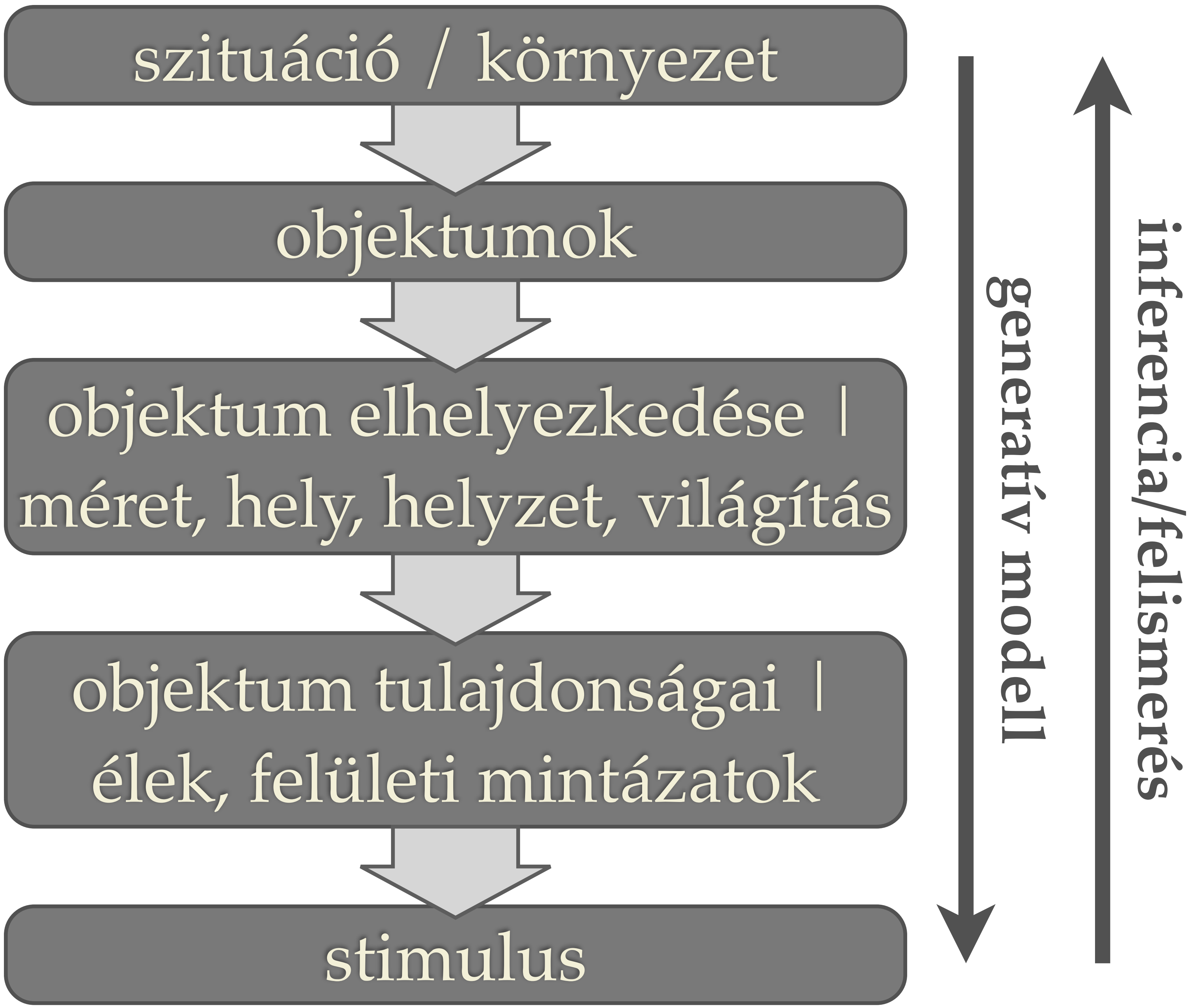


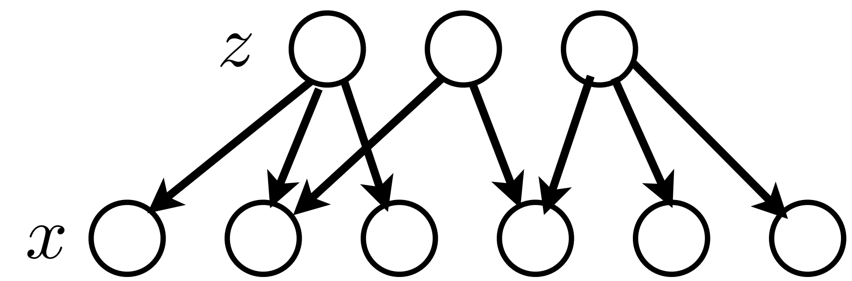


$$\mathrm{P}\left(\mathbf{z} \,|\, \mathbf{x}\right) \propto \mathrm{P}\left(\mathbf{x} \,|\, \mathbf{z}\right) \mathrm{P}\left(\mathbf{z}\right)$$

Eddig arra koncentráltunk, hogy mi a legvalószínűbb aktivitás

Bayes inferencia





$$\mathrm{P}\left(\mathbf{z}\,|\,\mathbf{x}\right) \propto \mathrm{P}\left(\mathbf{x}\,|\,\mathbf{z}\right)\mathrm{P}\left(\mathbf{z}\right)$$

Eddig arra koncentráltunk, hogy mi a legvalószínűbb aktivitás

Ez a maximum a posteriori becslés (MAP)

Bayes inferencia

Miért érdekes a poszterior eloszlás?

Bayes inferencia

Miért érdekes a poszterior eloszlás?

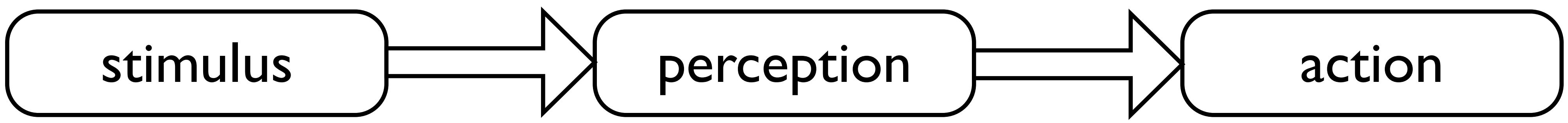
stimulus

perception

action

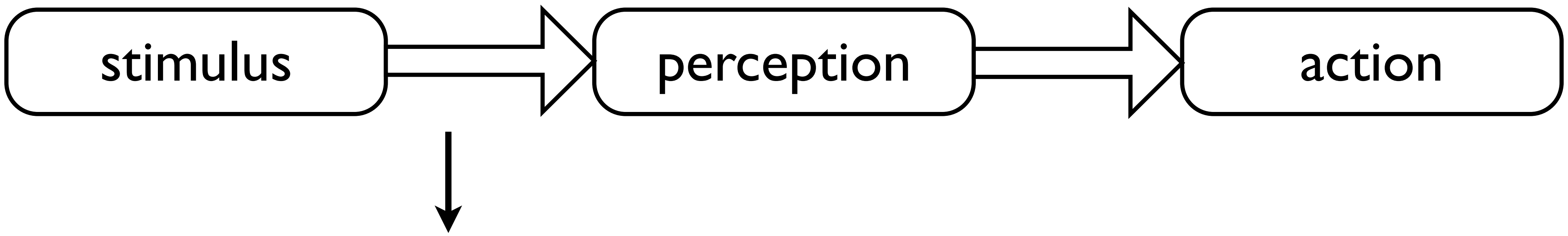
Bayes inferencia

Miért érdekes a poszterior eloszlás?



Bayes inferencia

Miért érdekes a poszterior eloszlás?



inference

$$\mathrm{P}\left(\mathbf{z}\,|\,\mathbf{x}\right) \propto \mathrm{P}\left(\mathbf{x}\,|\,\mathbf{z}\right)\mathrm{P}\left(\mathbf{z}\right)$$

Bayes inferencia

Miért érdekes a poszterior eloszlás?

stimulus $\Rightarrow$ perception $\Rightarrow$ action

inference

$$\mathrm{P}\left(\mathbf{z}\,|\,\mathbf{x}\right) \propto \mathrm{P}\left(\mathbf{x}\,|\,\mathbf{z}\right)\mathrm{P}\left(\mathbf{z}\right)$$

decision making

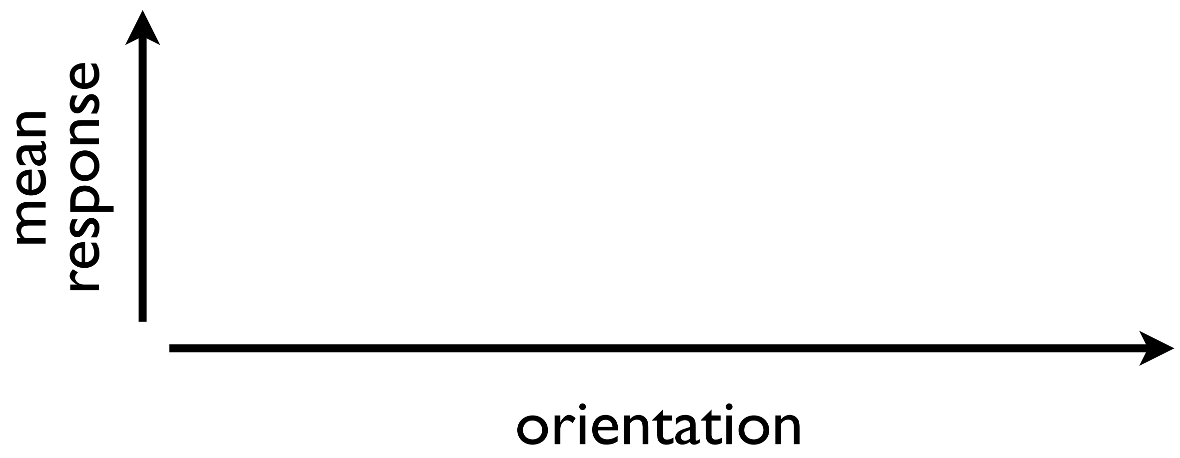
$$R(a) = \int dz\, U(z,a)\,\mathrm{P}(\mathbf{z}\,|\,\mathbf{x})$$

Bayes inferencia neuronhálózatokkal: PPC

V1 orientáció-szelektív neuronok

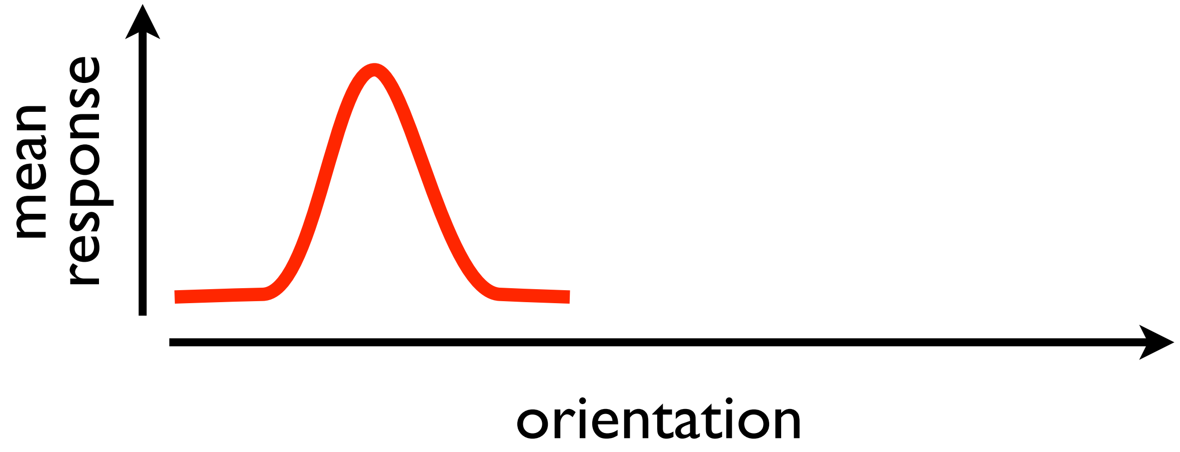
Bayes inferencia neuronhálózatokkal: PPC

V1 orientáció-szelektív neuronok



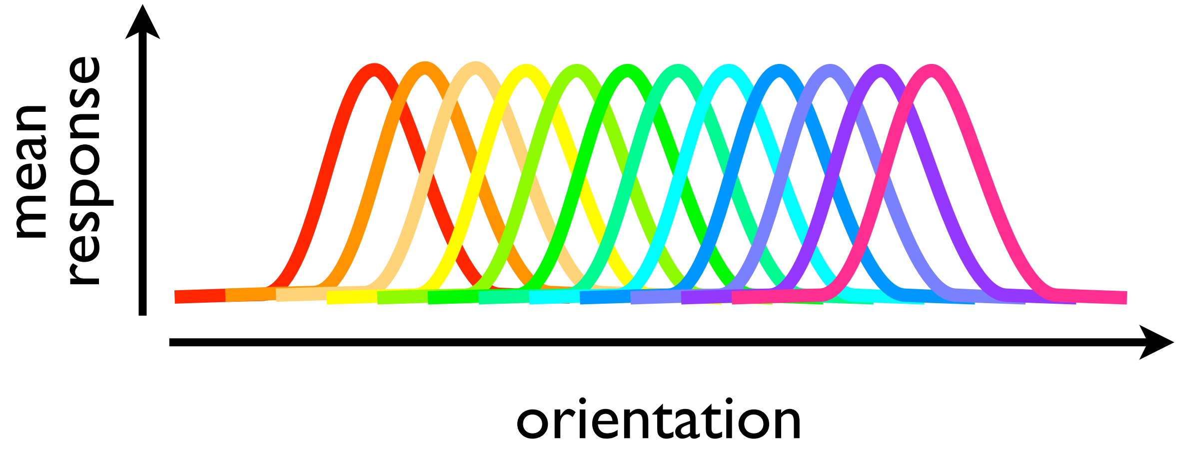
Bayes inferencia neuronhálózatokkal: PPC

V1 orientáció-szelektív neuronok



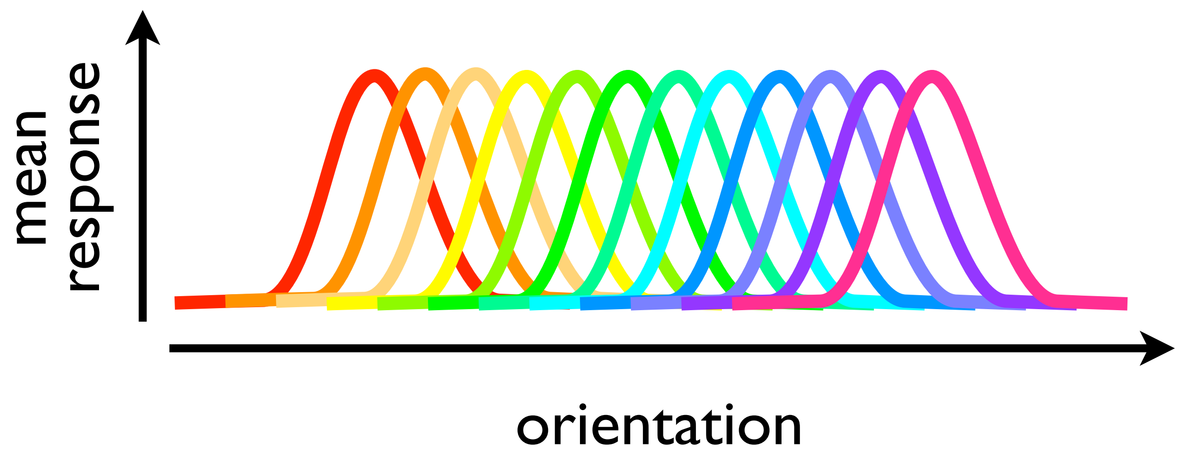
Bayes inferencia neuronhálózatokkal: PPC

V1 orientáció-szelektív neuronok



Bayes inferencia neuronhálózatokkal: PPC

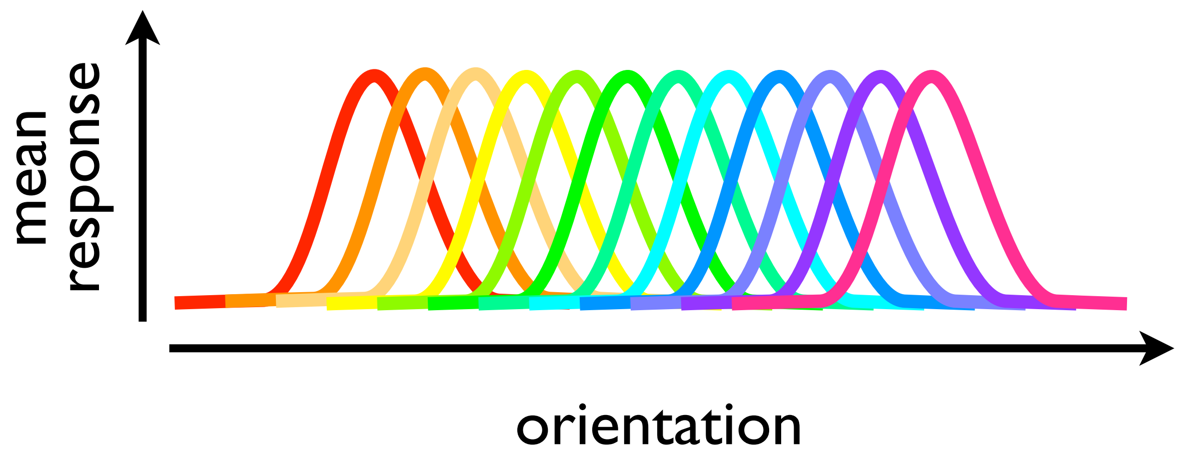
V1 orientáció-szelektív neuronok



a neuronok azonban zajosak:
az átlag körül az átlaggal
arányos variabilitás van jelen

Bayes inferencia neuronhálózatokkal: PPC

V1 orientáció-szelektív neuronok

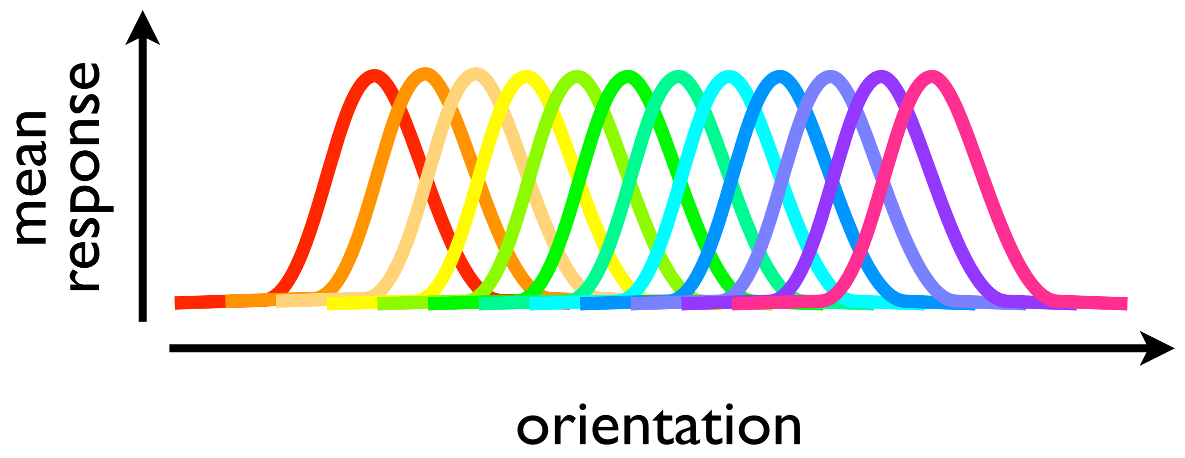


a neuronok azonban zajosak:
az átlag körül az átlaggal
arányos variabilitás van jelen

cél: orientáció becslése

Bayes inferencia neuronhálózatokkal: PPC

V1 orientáció-szelektív neuronok



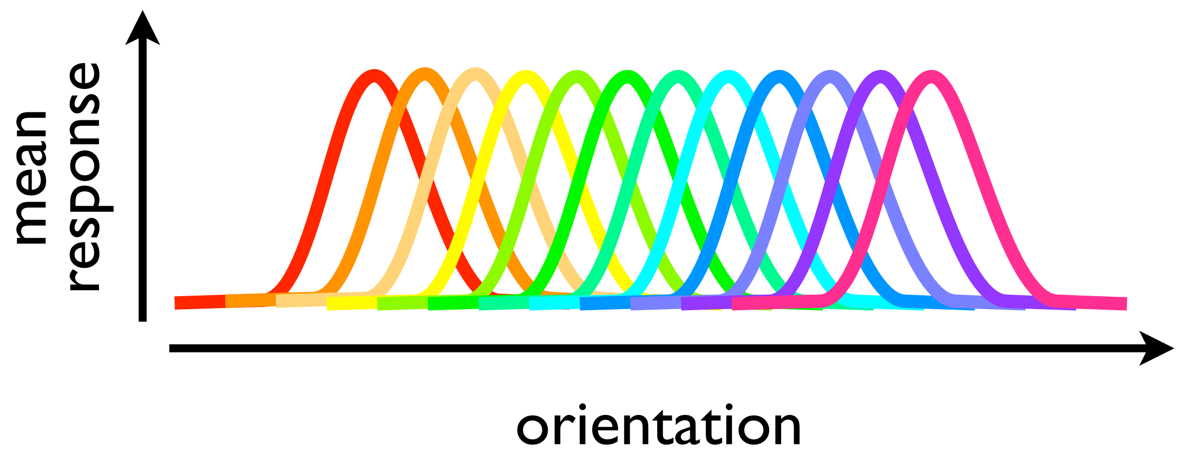
a neuronok azonban zajosak:
az átlag körül az átlaggal
arányos variabilitás van jelen

cél: orientáció becslése

megfigyelt változók: $\mathrm{r} = \{r_1, r_2, \ldots r_N\}$

Bayes inferencia neuronhálózatokkal: PPC

V1 orientáció-szelektív neuronok



a neuronok azonban zajosak: az átlag körül az átlaggal arányos variabilitás van jelen

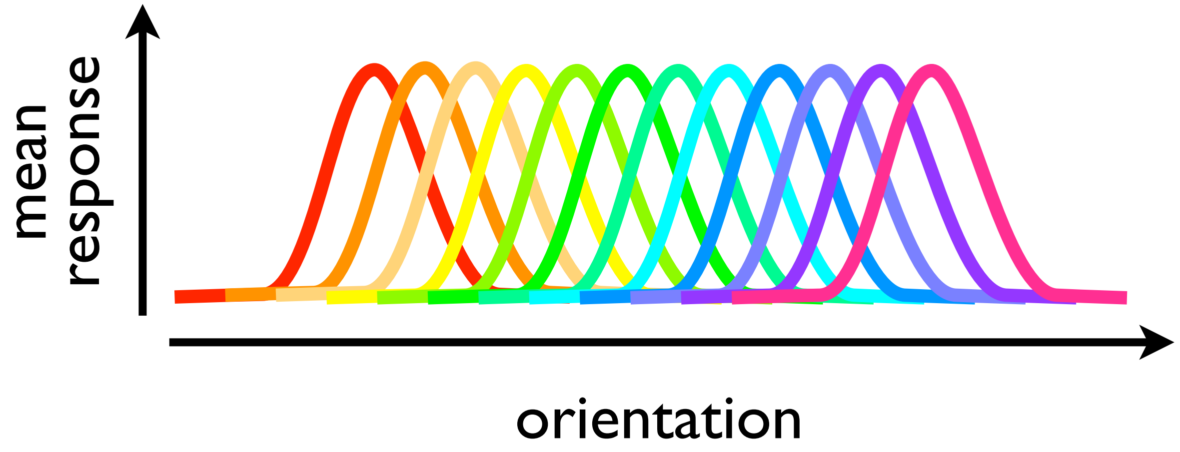
cél: orientáció becslése

megfigyelt változók: $\mathrm{r} = \{r_1, r_2, \ldots r_N\}$

nem megfigyelt változó: s

Bayes inferencia neuronhálózatokkal: PPC

V1 orientáció-szelektív neuronok



a neuronok azonban zajosak: az átlag körül az átlaggal arányos variabilitás van jelen

cél: orientáció becslése

megfigyelt változók: $\mathrm{r} = \{r_1, r_2, \ldots r_N\}$

nem megfigyelt változó: s

Bayes: $P(s \,|\, \mathbf{r}) \propto P(\mathbf{r} \,|\, s)\, P(s)$

Probabilistic Population Codes

Probabilistic Population Codes

- Neurális zaj varianciája arányos az átlagos aktivitással: Poisson zaj

Probabilistic Population Codes

- Neurális zaj varianciája arányos az átlagos aktivitással: Poisson zaj
- Likelihood alakja:

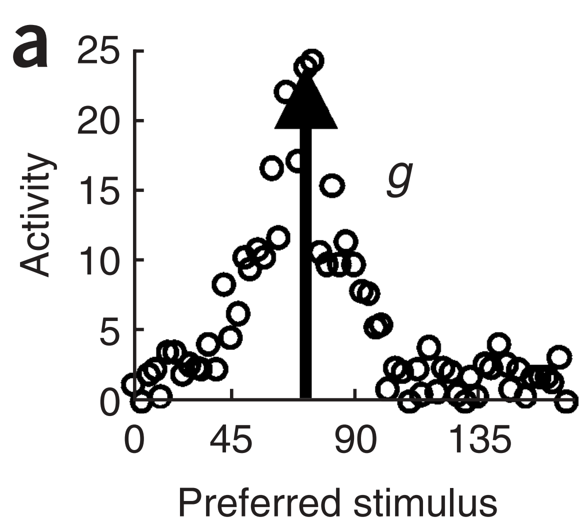
$$P(\mathbf{r} \mid s) = \prod_i \frac{e^{-f_i(s)} f_i(s)^{r_i}}{r_i!}$$

Probabilistic Population Codes

- Neurális zaj varianciája arányos az átlagos aktivitással: Poisson zaj

- Likelihood alakja:

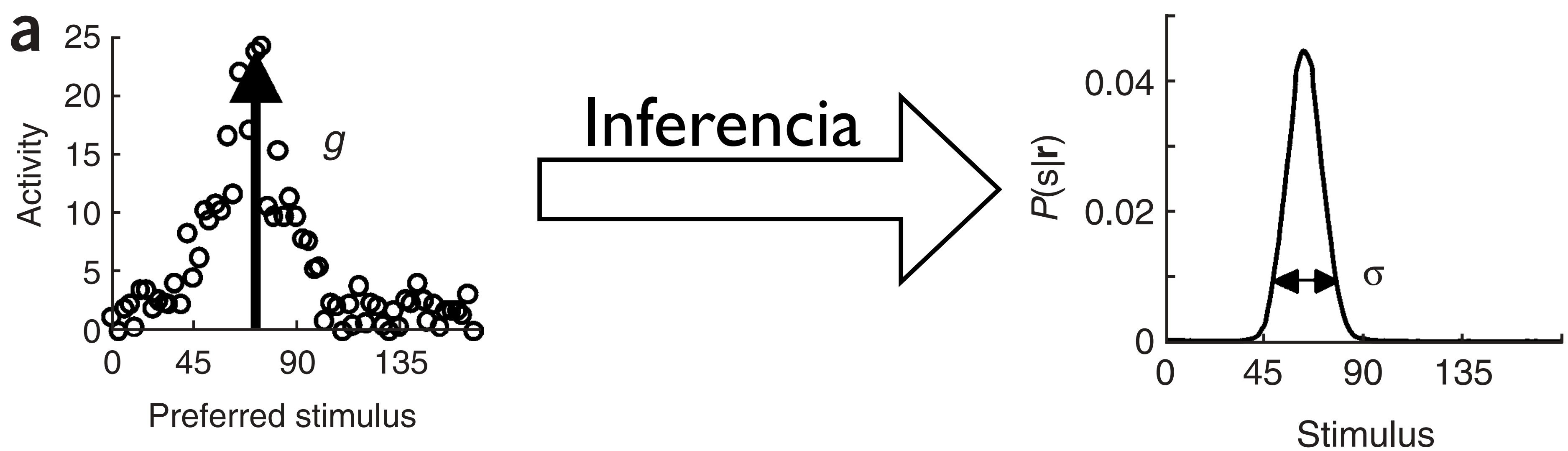
$$P(\mathbf{r} \mid s) = \prod_i \frac{e^{-f_i(s)} f_i(s)^{r_i}}{r_i!}$$



Probabilistic Population Codes

- Neurális zaj varianciája arányos az átlagos aktivitással: Poisson zaj
- Likelihood alakja:

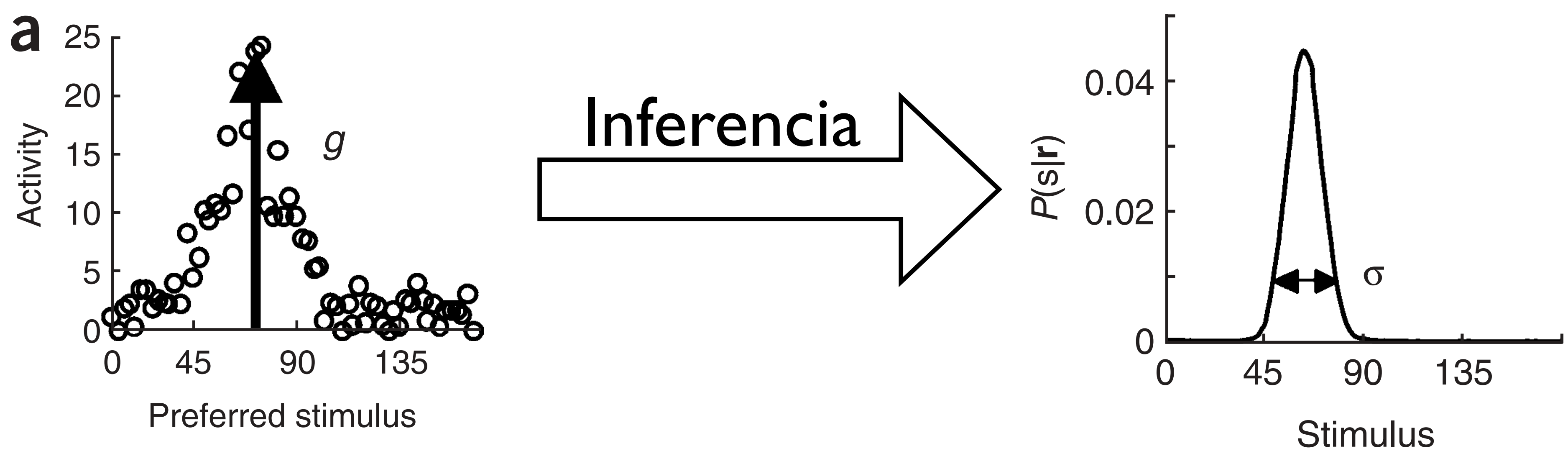
$$P(\mathbf{r}\,|\,s) = \prod_i \frac{e^{-f_i(s)} f_i(s)^{r_i}}{r_i!}$$



Probabilistic Population Codes

- Neurális zaj varianciája arányos az átlagos aktivitással: Poisson zaj
- Likelihood alakja:

$$P(\mathbf{r} \mid s) = \prod_i \frac{e^{-f_i(s)} f_i(s)^{r_i}}{r_i!}$$

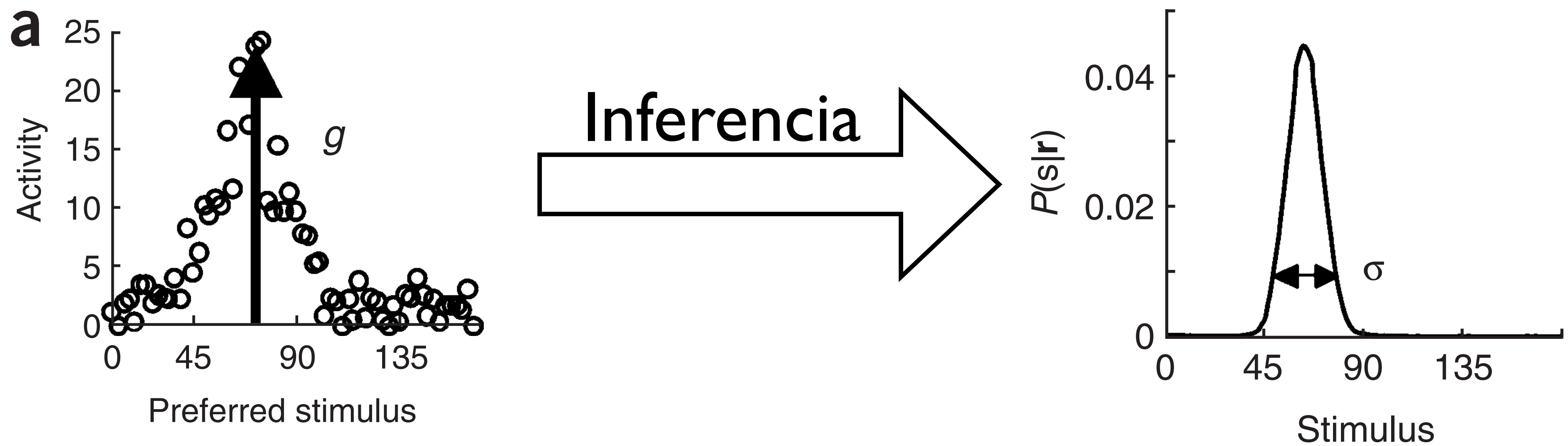


ha a stimulus-eloszlás Gauss,
akkor az aktivitás-intenzítás arányos a precizióval

Probabilistic Population Codes

- Neurális zaj varianciája arányos az átlagos aktivitással: Poisson zaj
- Likelihood alakja:

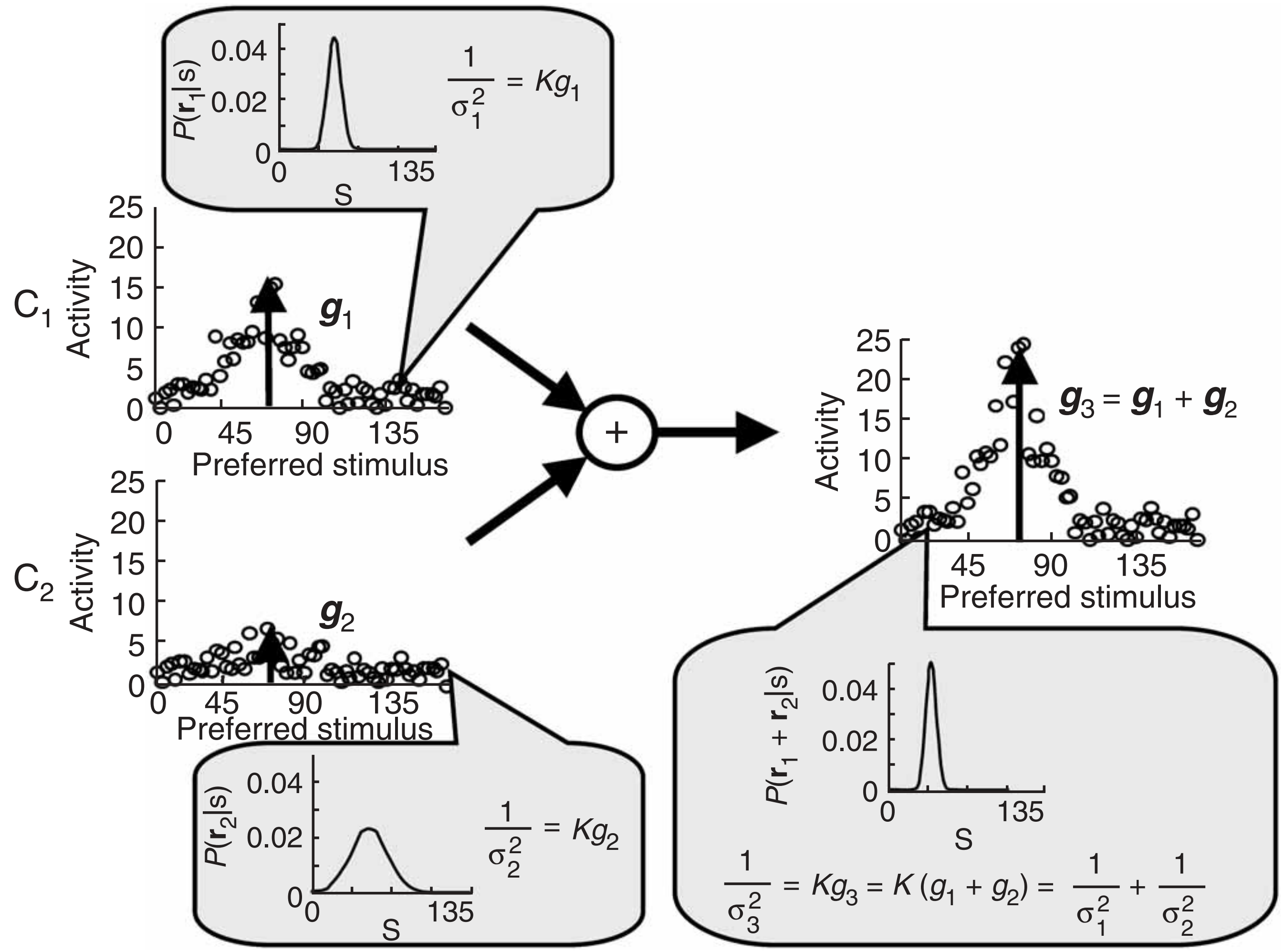
$$P(\mathbf{r} \mid s) = \prod_i \frac{e^{-f_i(s)} f_i(s)^{r_i}}{r_i!}$$



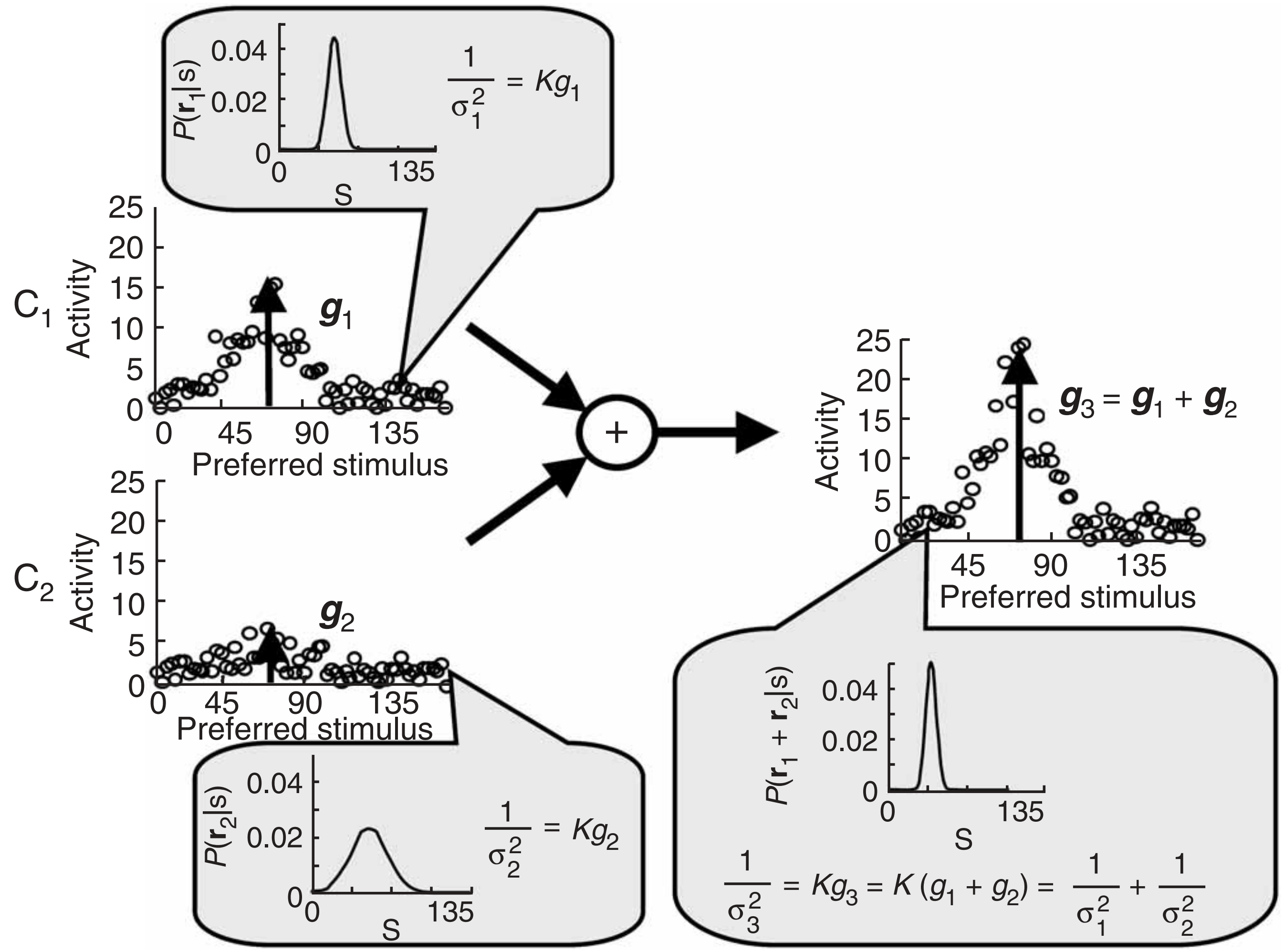
ha a stimulus-eloszlás Gauss,
akkor az aktivitás-intenzítás arányos a precizióval

$$g \propto \frac{1}{\sigma^2}$$

PPC: Multiszenzoros integráció

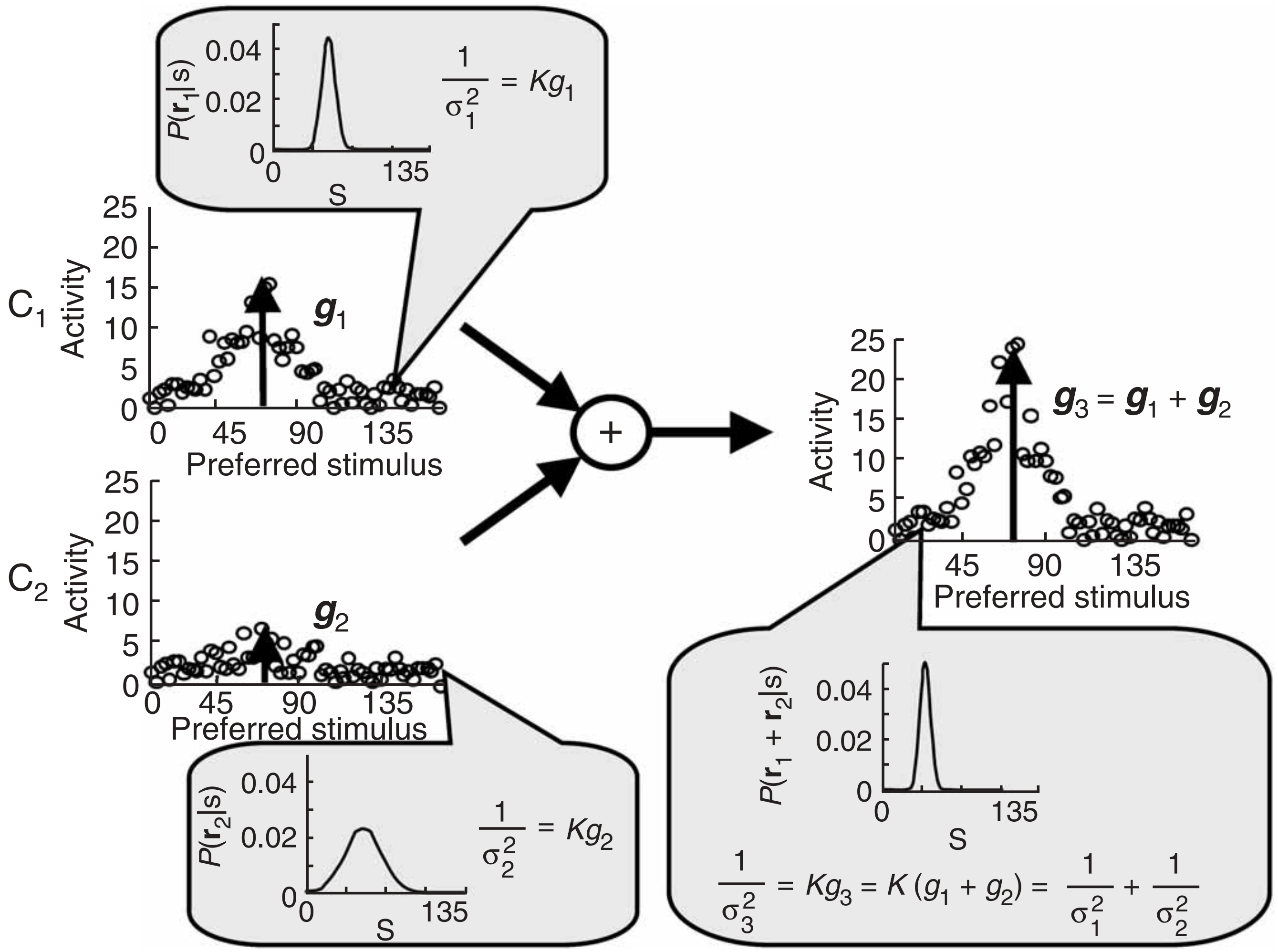


PPC: Multiszenzoros integráció



$$P(\mathbf{r} \mid s) = \prod_i \frac{e^{-f_i(s)} f_i(s)^{r_i}}{r_i!}$$

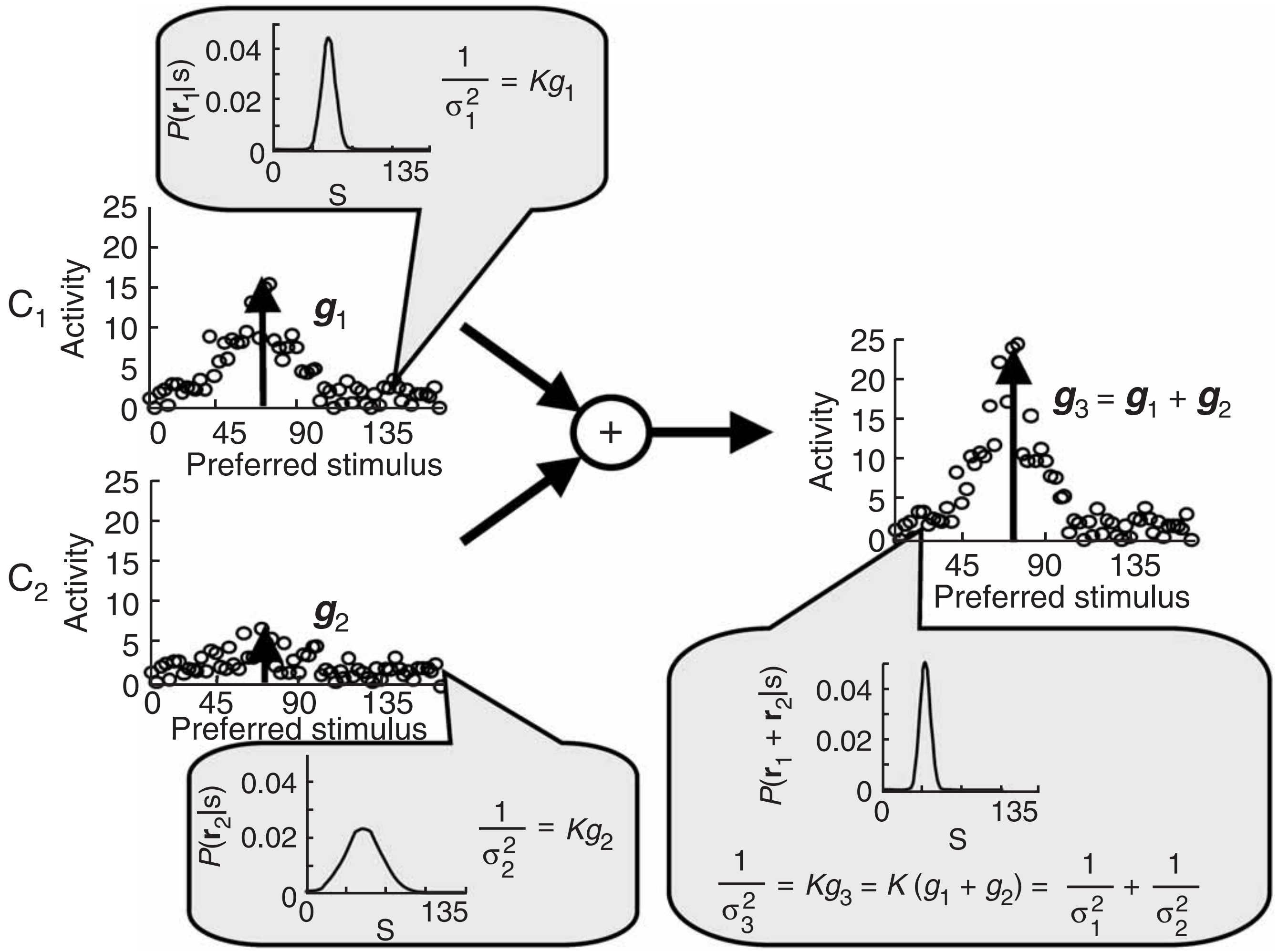
PPC: Multiszenzoros integráció



$$P(\mathbf{r} \mid s) = \prod_i \frac{e^{-f_i(s)} f_i(s)^{r_i}}{r_i!}$$

$$p(s|c_1, c_2) \propto p(c_1|s)p(c_2|s)p(s).$$

PPC: Multiszenzoros integráció



$$P(\mathbf{r} \mid s) = \prod_i \frac{e^{-f_i(s)} f_i(s)^{r_i}}{r_i!}$$

$$p(s|c_1, c_2) \propto p(c_1|s)p(c_2|s)p(s).$$

$$\frac{1}{\sigma_3^2} = \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}$$

$$\mu_3 = \frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}\mu_1 + \frac{\sigma_1^2}{\sigma_1^2+\sigma_2^2}\mu_2$$

STATISTICALLY OPTIMAL INTERNAL MODELS?

prior knowledge
about the visual world

↓

percept

↑

visual stimulus

STATISTICALLY OPTIMAL INTERNAL MODELS?

prior knowledge
about the visual world

$\mathbf{y}^*$
objects, features

percept

visual stimulus

STATISTICALLY OPTIMAL INTERNAL MODELS?

prior knowledge
about the visual world

$\mathrm{P}^*(\mathbf{y})$

$\mathbf{y}^*$
objects, features

percept

visual stimulus

STATISTICALLY OPTIMAL INTERNAL MODELS?

prior knowledge
about the visual world

$\mathrm{P}^*(\mathbf{y})$

$\mathbf{y}^*$
objects, features

percept

$\mathrm{P}^*(\mathbf{x}|\mathbf{y})$

$\mathbf{x}$
visual stimulus

STATISTICALLY OPTIMAL INTERNAL MODELS?

prior knowledge
about the visual world

$\mathrm{P}^*(\mathbf{y})$

$\mathbf{y}^*$
objects, features

$\mathbf{y}$
percept

$\mathrm{P}^*(\mathbf{x}|\mathbf{y})$

$\mathbf{x}$
visual stimulus

STATISTICALLY OPTIMAL INTERNAL MODELS?

prior knowledge
about the visual world

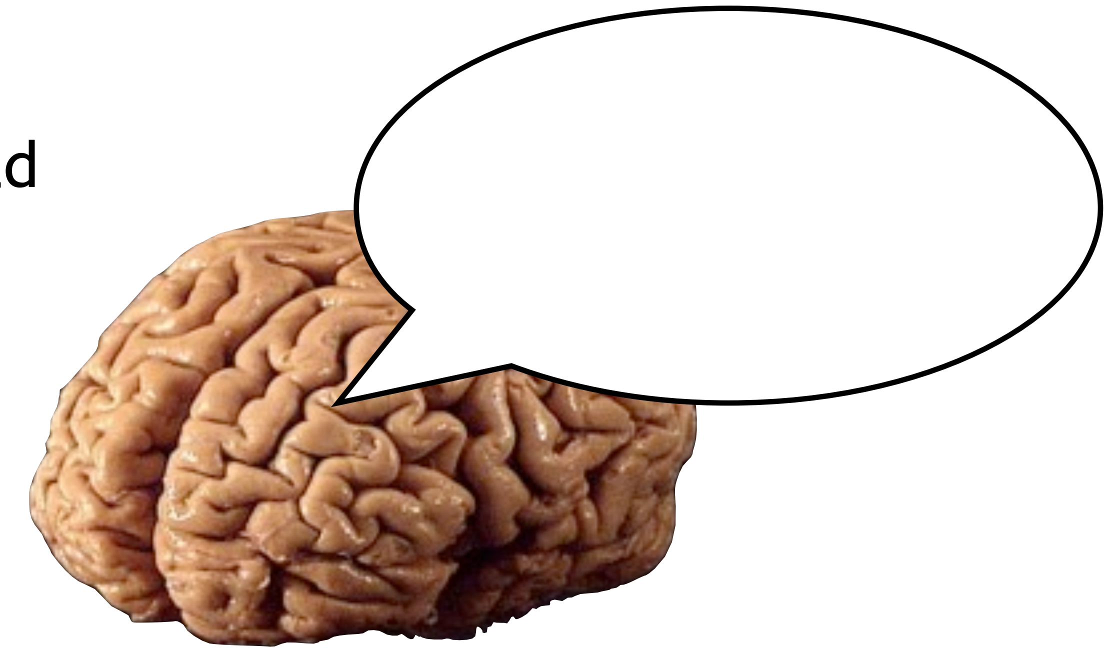
$\mathrm{P}^*(\mathbf{y})$ $\quad$ $\mathrm{P}_\theta(\mathbf{y})$

$\mathbf{y}^*$
objects, features

$\mathbf{y}$
percept

$\mathrm{P}^*(\mathbf{x}|\mathbf{y})$ $\quad$ $\mathrm{P}_\theta(\mathbf{x}|\mathbf{y})$

$\mathbf{x}$
visual stimulus



STATISTICALLY OPTIMAL INTERNAL MODELS?

prior knowledge
about the visual world

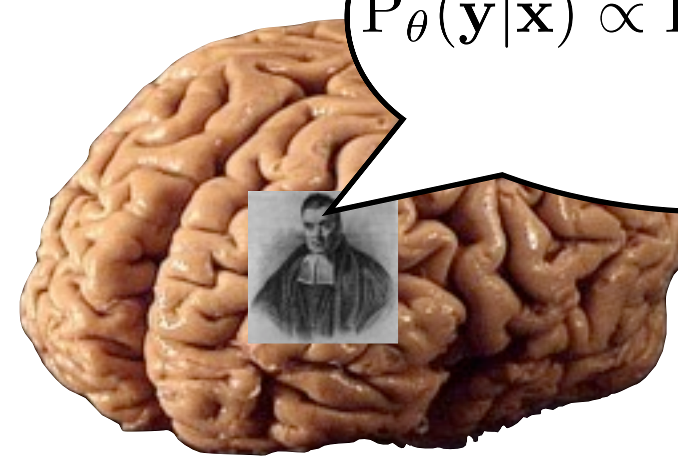
$$\mathrm{P}^*(\mathbf{y}) \qquad \mathrm{P}_\theta(\mathbf{y})$$

$\mathbf{y}^*$
objects, features

$\mathbf{y}$
percept

$$\mathrm{P}^*(\mathbf{x}|\mathbf{y}) \qquad \mathrm{P}_\theta(\mathbf{x}|\mathbf{y})$$

$\mathbf{x}$
visual stimulus



$$\mathrm{P}_\theta(\mathbf{y}|\mathbf{x}) \propto \mathrm{P}_\theta(\mathbf{x}|\mathbf{y})\,\mathrm{P}_\theta(\mathbf{y})$$

STATISTICALLY OPTIMAL INTERNAL MODELS?

prior knowledge
about the visual world

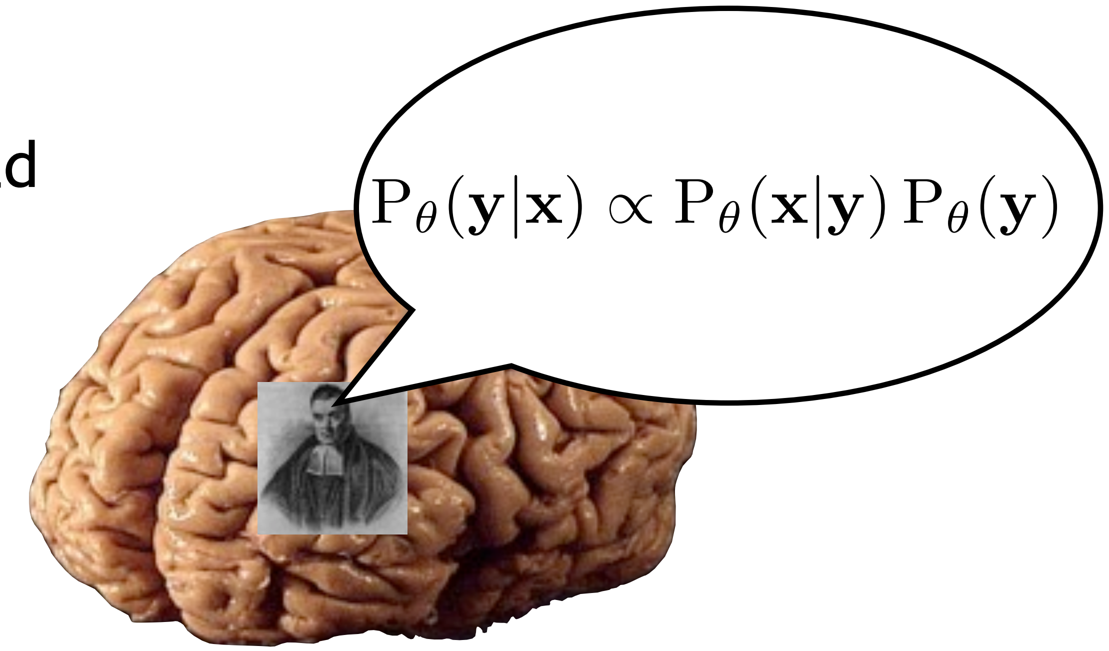
$\mathrm{P}^*(\mathbf{y})$ $\qquad$ $\mathrm{P}_\theta(\mathbf{y})$

$\mathbf{y}^*$
objects, features

$\mathbf{y}$
percept

$\mathrm{P}^*(\mathbf{x}|\mathbf{y})$ $\qquad$ $\mathrm{P}_\theta(\mathbf{x}|\mathbf{y})$

$\mathbf{x}$
visual stimulus



$$\mathrm{P}_\theta(\mathbf{y}|\mathbf{x}) \propto \mathrm{P}_\theta(\mathbf{x}|\mathbf{y})\,\mathrm{P}_\theta(\mathbf{y})$$

HOW DO WE KNOW IF AN INTERNAL MODEL IS OPTIMAL?

$$\int \mathrm{P}_\theta(\mathbf{y}|\mathbf{x})\ \mathrm{P}^*(\mathbf{x})\ d\mathbf{x} \quad = \quad \mathrm{P}_\theta(\mathbf{y})$$

STATISTICALLY OPTIMAL INTERNAL MODELS?

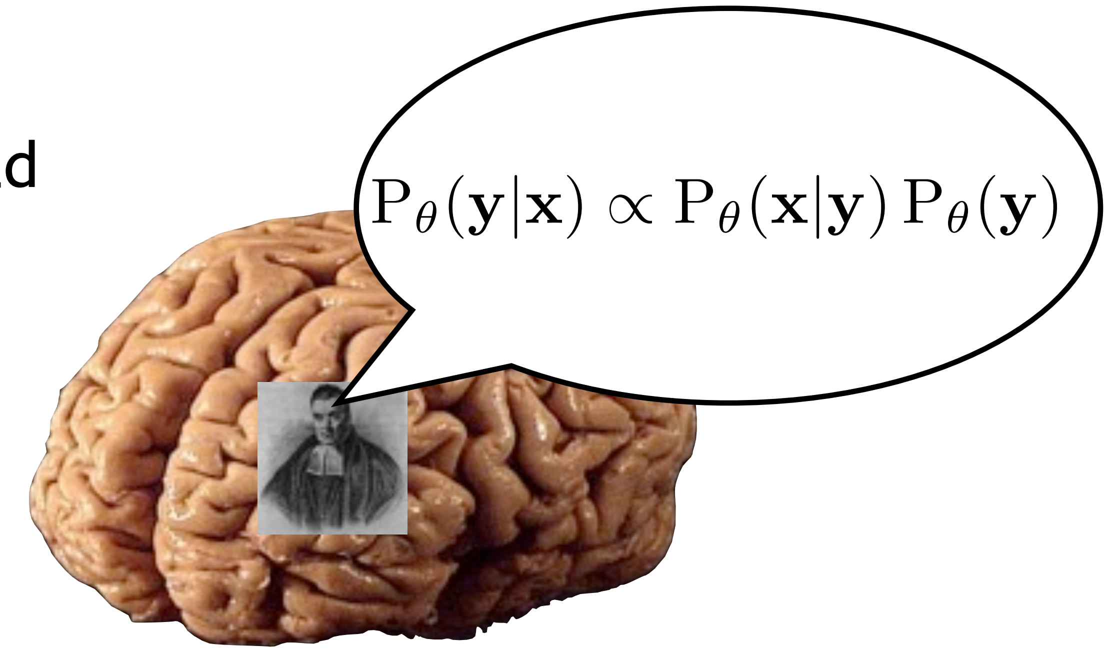
prior knowledge
about the visual world

$\mathrm{P}^*(\mathbf{y})$ ↓ $\mathbf{y}^*$ objects, features

$\mathrm{P}_\theta(\mathbf{y})$ ↓ $\mathbf{y}$ percept

$\mathrm{P}^*(\mathbf{x}|\mathbf{y})$ ↘ $\mathbf{x}$ visual stimulus

$\mathrm{P}_\theta(\mathbf{x}|\mathbf{y})$ ↑ $\mathbf{y}$ percept



HOW DO WE KNOW IF AN INTERNAL MODEL IS OPTIMAL?

"average inferences" "prior expectations"

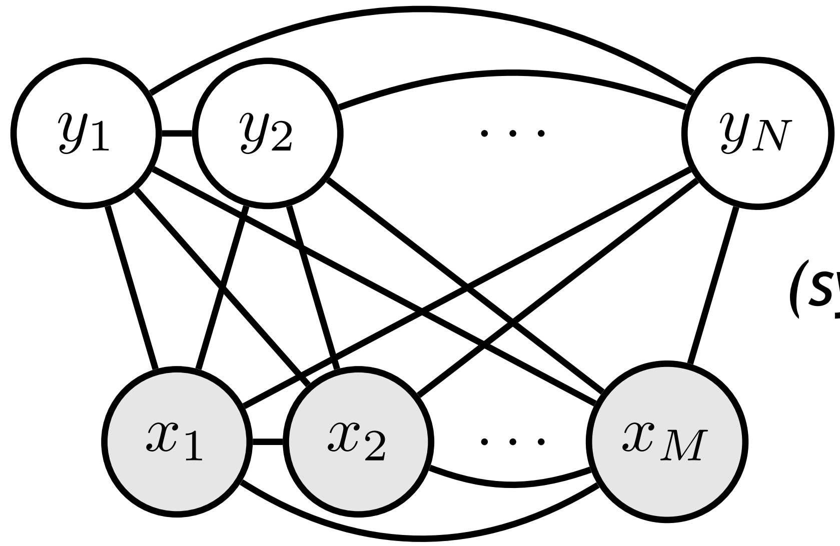
$$\int \mathrm{P}_\theta(\mathbf{y}|\mathbf{x})\ \mathrm{P}^*(\mathbf{x})\ d\mathbf{x} \quad = \quad \mathrm{P}_\theta(\mathbf{y})$$

for natural stimulus ensemble

$$\mathrm{P}^*(\mathbf{x}) = \int \mathrm{P}^*(\mathbf{y})\,\mathrm{P}^*(\mathbf{x}|\mathbf{y})\,d\mathbf{y}$$

APPLICATION TO VISUAL CORTEX

latent variables
(visual cortex)
$\mathbf{y}$



parameters
(synaptic efficacies, etc)
θ

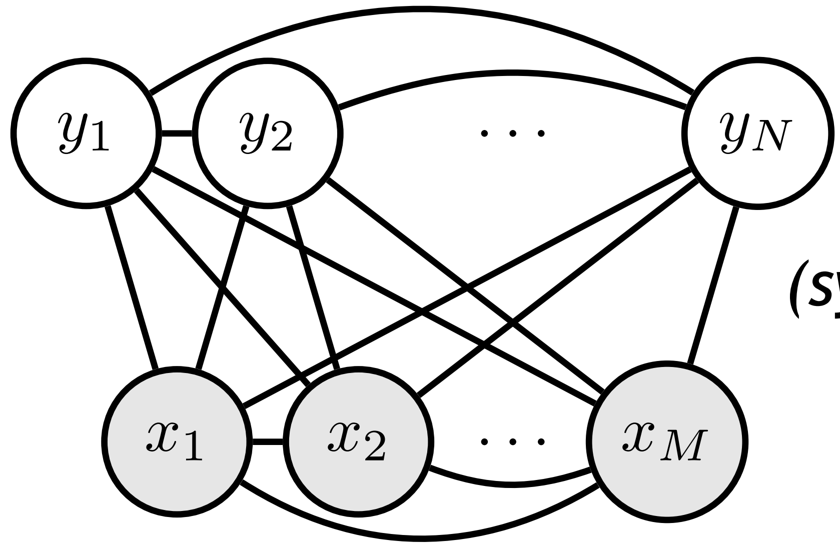
observed variables
(retina)
$\mathbf{x}$

$$\boxed{\int \mathrm{P}_\theta(\mathbf{y}|\mathbf{x})\ \mathrm{P}^*(\mathbf{x})\ d\mathbf{x} = \mathrm{P}_\theta(\mathbf{y})}$$

APPLICATION TO VISUAL CORTEX

latent variables
(visual cortex)
$\mathbf{y}$

observed variables
(retina)
$\mathbf{x}$



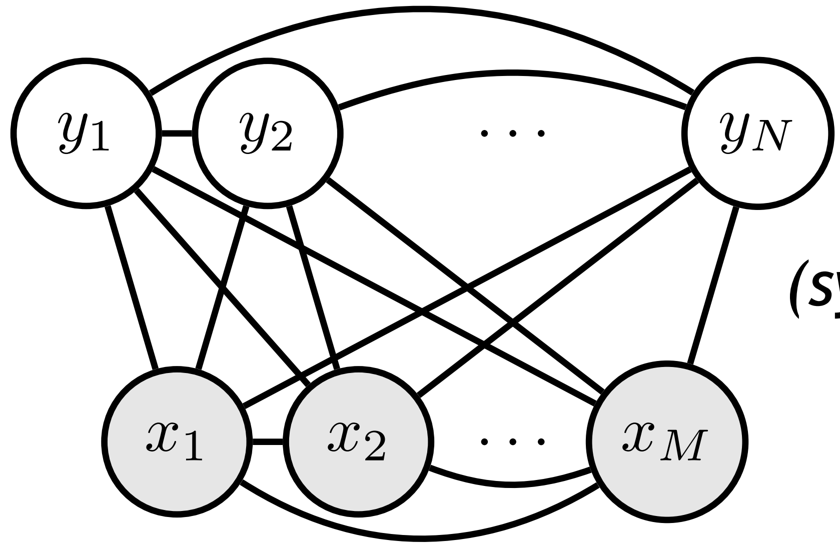
parameters
(synaptic efficacies, etc)
θ

$$\boxed{\int \mathrm{P}_\theta(\mathbf{y}|\mathbf{x})\ \mathrm{P}^*(\mathbf{x})\ d\mathbf{x} = \mathrm{P}_\theta(\mathbf{y})}$$

evoked
activity

APPLICATION TO VISUAL CORTEX

latent variables
(visual cortex)
$\mathbf{y}$



parameters
(synaptic efficacies, etc)
θ

observed variables
(retina)
$\mathbf{x}$

$$\boxed{\int \mathrm{P}_\theta(\mathbf{y}|\mathbf{x}) \ \mathrm{P}^*(\mathbf{x}) \ d\mathbf{x} = \mathrm{P}_\theta(\mathbf{y})}$$

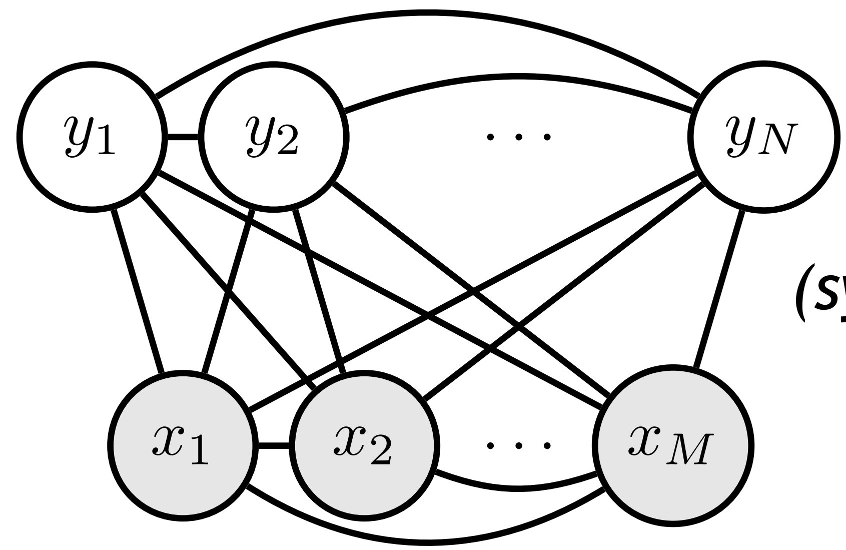
evoked
activity

spontaneous
activity

APPLICATION TO VISUAL CORTEX

latent variables
(visual cortex)
$\mathbf{y}$

observed variables
(retina)
$\mathbf{x}$



parameters
(synaptic efficacies, etc)
θ

$$\boxed{\int \mathrm{P}_\theta(\mathbf{y}|\mathbf{x})\ \mathrm{P}^*(\mathbf{x})\ d\mathbf{x} = \mathrm{P}_\theta(\mathbf{y})}$$

evoked activity

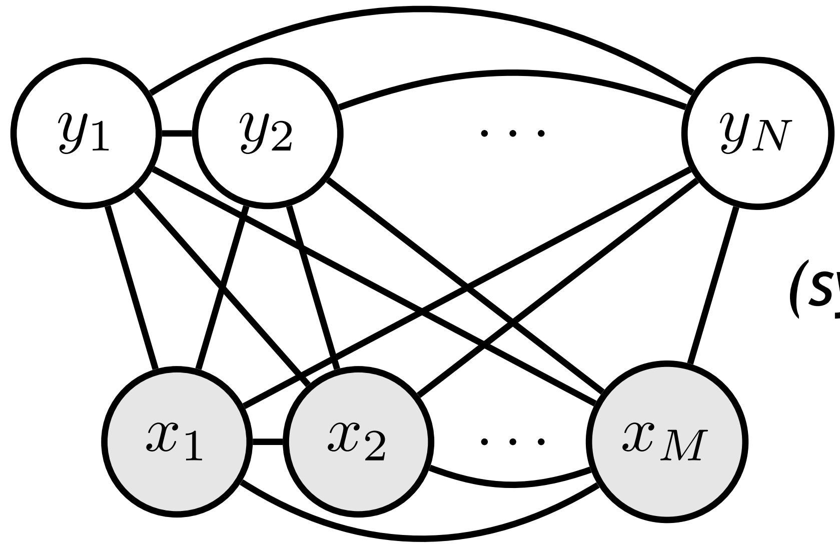
stimulus ensemble

spontaneous activity

APPLICATION TO VISUAL CORTEX

latent variables
(visual cortex)
$\mathbf{y}$

observed variables
(retina)
$\mathbf{x}$



parameters
(synaptic efficacies, etc)
θ

$$\boxed{\int \mathrm{P}_\theta(\mathbf{y}|\mathbf{x})\ \mathrm{P}^*(\mathbf{x})\ d\mathbf{x} \stackrel{?}{=} \mathrm{P}_\theta(\mathbf{y})}$$

evoked activity

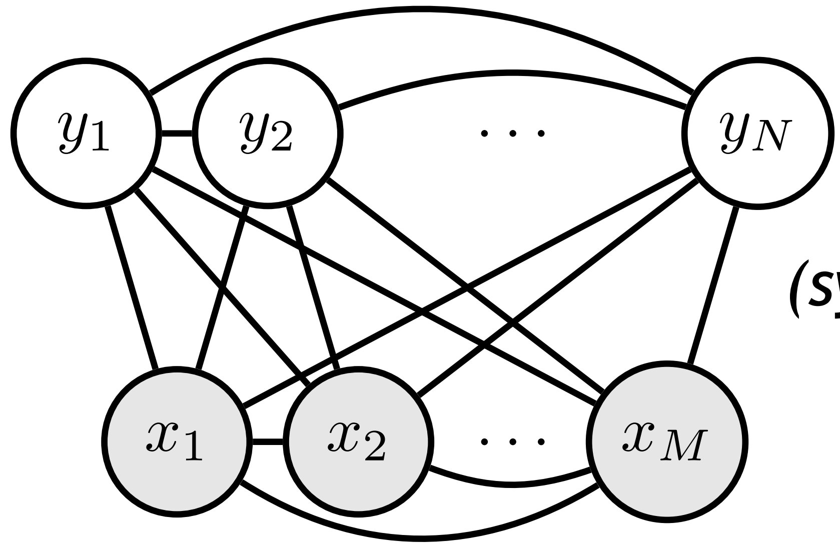
stimulus ensemble

spontaneous activity

APPLICATION TO VISUAL CORTEX

latent variables
(visual cortex)
$\mathbf{y}$

observed variables
(retina)
$\mathbf{x}$



parameters
(synaptic efficacies, etc)
θ

KL divergence

$$\boxed{\int \mathrm{P}_\theta(\mathbf{y}|\mathbf{x})\ \mathrm{P}^*(\mathbf{x})\ d\mathbf{x} \stackrel{?}{=} \mathrm{P}_\theta(\mathbf{y})}$$

evoked activity

stimulus ensemble

spontaneous activity

RECORDINGS

RECORDINGS

awake behaving ferrets

RECORDINGS

awake behaving ferrets
aged P29 (eye opening) — P151 (mature visual system)

RECORDINGS

awake behaving ferrets
aged P29 (eye opening) — P151 (mature visual system)
multi-unit recordings from layers 2/3 of V1

RECORDINGS

awake behaving ferrets
aged P29 (eye opening) — P151 (mature visual system)
multi-unit recordings from layers 2/3 of V1
16 electrodes with 200 μm spacing

RECORDINGS

awake behaving ferrets
aged P29 (eye opening) — P151 (mature visual system)
multi-unit recordings from layers 2/3 of V1
16 electrodes with 200 μm spacing

conditions

- **spontaneous**

darkness
$\mathrm{S}(\mathbf{y}) = \mathrm{P}_\theta(\mathbf{y})$

RECORDINGS

awake behaving ferrets
aged P29 (eye opening) – P151 (mature visual system)
multi-unit recordings from layers 2/3 of V1
16 electrodes with 200 μm spacing

conditions

- **spontaneous**

darkness

$$\mathrm{S}(\mathbf{y}) = \mathrm{P}_\theta(\mathbf{y})$$

- **evoked**

natural image movies

$$\mathrm{M}(\mathbf{y}) = \int \mathrm{P}_\theta(\mathbf{y}|\mathbf{x})\ \mathrm{P}^*_{\mathrm{movie}}(\mathbf{x})\ d\mathbf{x}$$

RECORDINGS

awake behaving ferrets
aged P29 (eye opening) – P151 (mature visual system)
multi-unit recordings from layers 2/3 of V1
16 electrodes with 200 μm spacing

conditions

- **spontaneous**

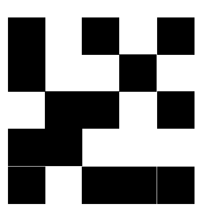
darkness

$$\mathrm{S}(\mathbf{y}) = \mathrm{P}_\theta(\mathbf{y})$$

- **evoked**

natural image movies

$$\mathrm{M}(\mathbf{y}) = \int \mathrm{P}_\theta(\mathbf{y}|\mathbf{x})\ \mathrm{P}^*_{\mathrm{movie}}(\mathbf{x})\ d\mathbf{x}$$



dynamic block noise

$$\mathrm{N}(\mathbf{y}) = \int \mathrm{P}_\theta(\mathbf{y}|\mathbf{x})\ \mathrm{P}^*_{\mathrm{noise}}(\mathbf{x})\ d\mathbf{x}$$

RECORDINGS

awake behaving ferrets
aged P29 (eye opening) – P151 (mature visual system)
multi-unit recordings from layers 2/3 of V1
16 electrodes with 200 μm spacing

conditions

- **spontaneous**

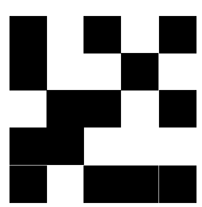
darkness

$$\mathrm{S}(\mathbf{y}) = \mathrm{P}_\theta(\mathbf{y})$$

- **evoked**

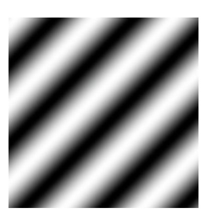
natural image movies

$$\mathrm{M}(\mathbf{y}) = \int \mathrm{P}_\theta(\mathbf{y}|\mathbf{x})\ \mathrm{P}^*_{\mathrm{movie}}(\mathbf{x})\ d\mathbf{x}$$



dynamic block noise

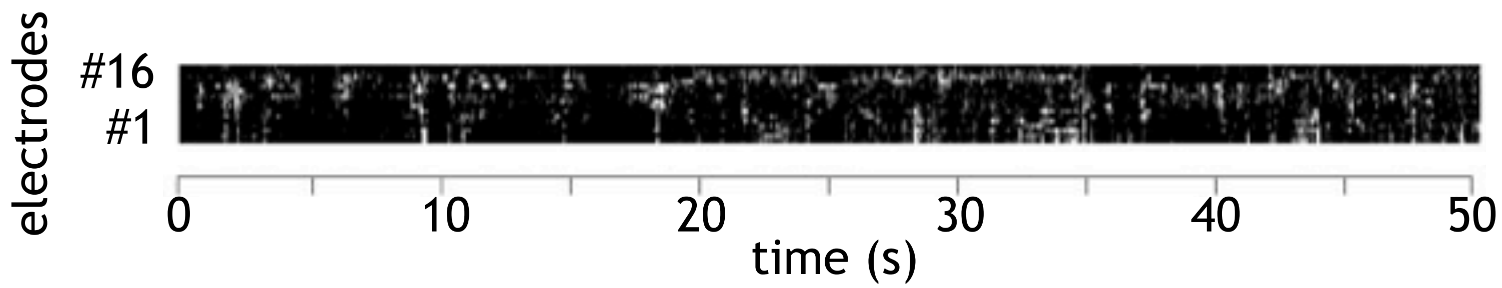
$$\mathrm{N}(\mathbf{y}) = \int \mathrm{P}_\theta(\mathbf{y}|\mathbf{x})\ \mathrm{P}^*_{\mathrm{noise}}(\mathbf{x})\ d\mathbf{x}$$



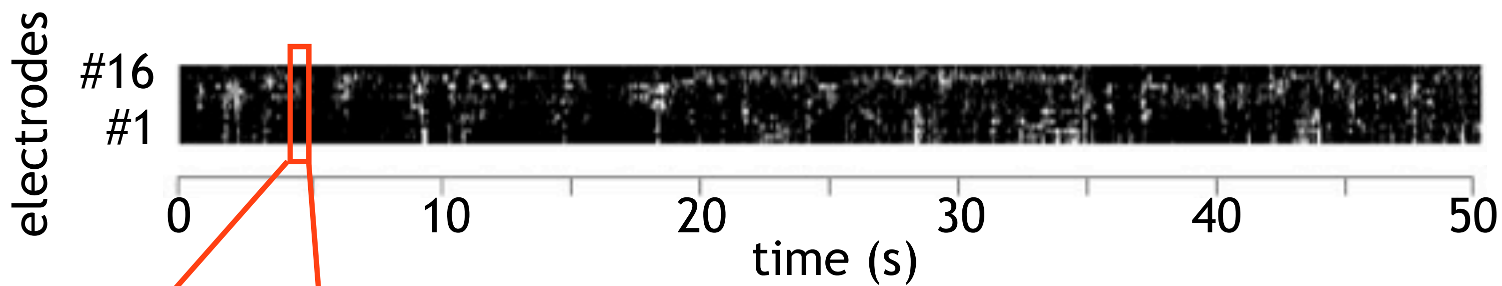
drifting full-field gratings

$$\mathrm{G}(\mathbf{y}) = \int \mathrm{P}_\theta(\mathbf{y}|\mathbf{x})\ \mathrm{P}^*_{\mathrm{grating}}(\mathbf{x})\ d\mathbf{x}$$

DATA ANALYSIS



DATA ANALYSIS



discretisation and binarisation

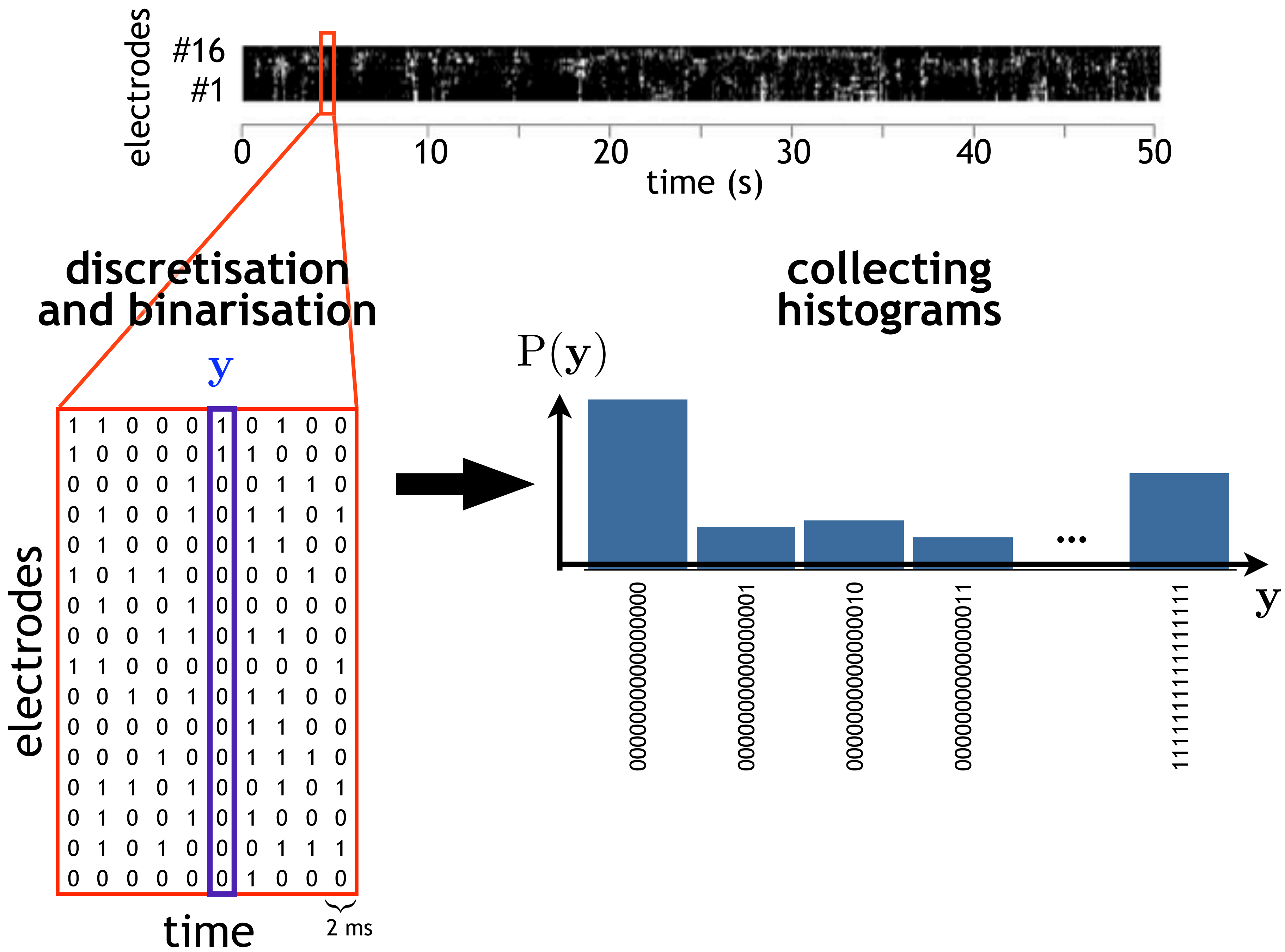
y

electrodes

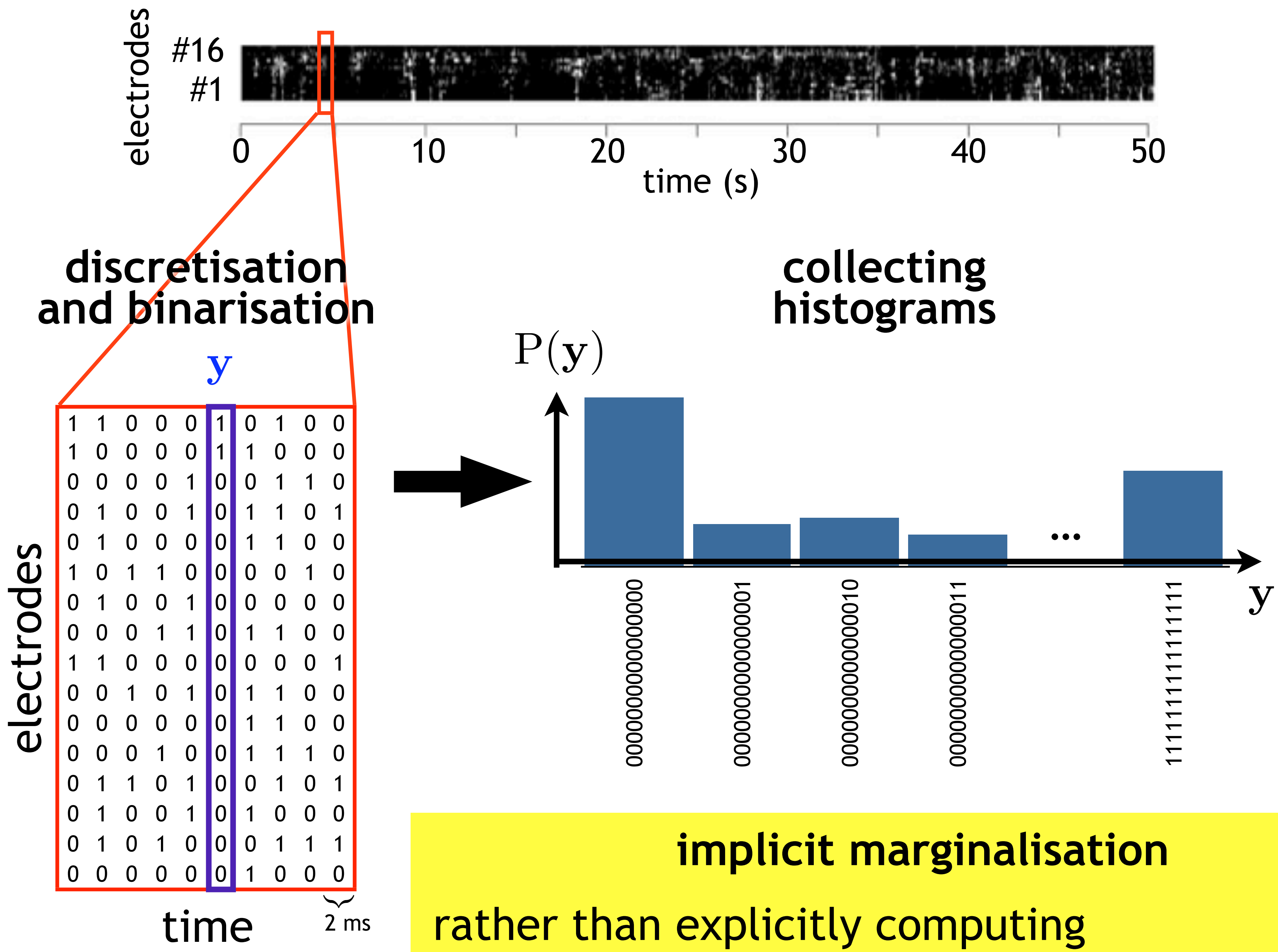
1	1	0	0	0	1	0	1	0	0
1	0	0	0	0	1	1	0	0	0
0	0	0	0	1	0	0	1	1	0
0	1	0	0	1	0	1	1	0	1
0	1	0	0	0	0	1	1	0	0
1	0	1	1	0	0	0	0	1	0
0	1	0	0	1	0	0	0	0	0
0	0	0	1	1	0	1	1	0	0
1	1	0	0	0	0	0	0	0	1
0	0	1	0	1	0	1	1	0	0
0	0	0	0	0	0	1	1	0	0
0	0	0	1	0	0	1	1	1	0
0	1	1	0	1	0	0	1	0	1
0	1	0	0	1	0	1	0	0	0
0	1	0	1	0	0	0	1	1	1
0	0	0	0	0	0	1	0	0	0

time $\underbrace{\quad}_{\text{2 ms}}$

DATA ANALYSIS



DATA ANALYSIS



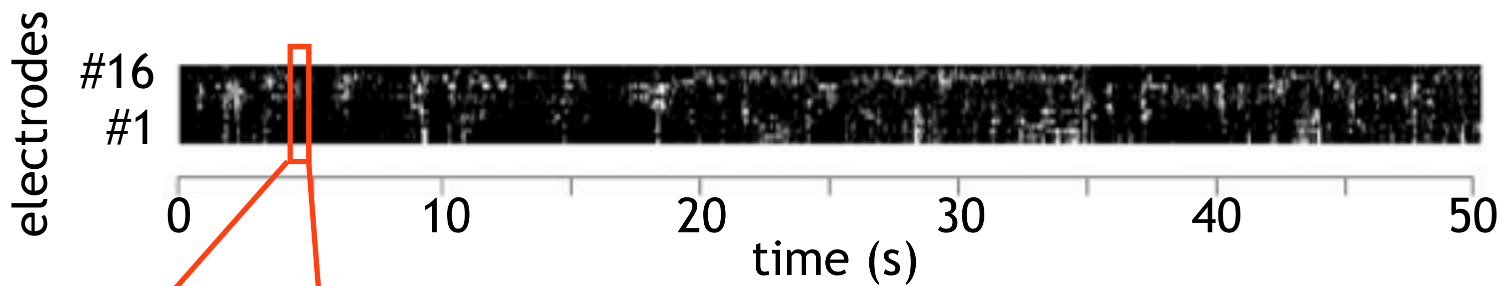
implicit marginalisation

rather than explicitly computing

$$\mathrm{P}_{\text{evoked}}(\mathbf{y}) = \int \mathrm{P}_{\theta}(\mathbf{y}|\mathbf{x})\ \mathrm{P}^{*}_{\text{stim}}(\mathbf{x})\ d\mathbf{x}$$

we collect evoked activity under each condition without regard to the actual ongoing stimulus

DATA ANALYSIS



discretisation and binarisation

$\mathbf{y}$

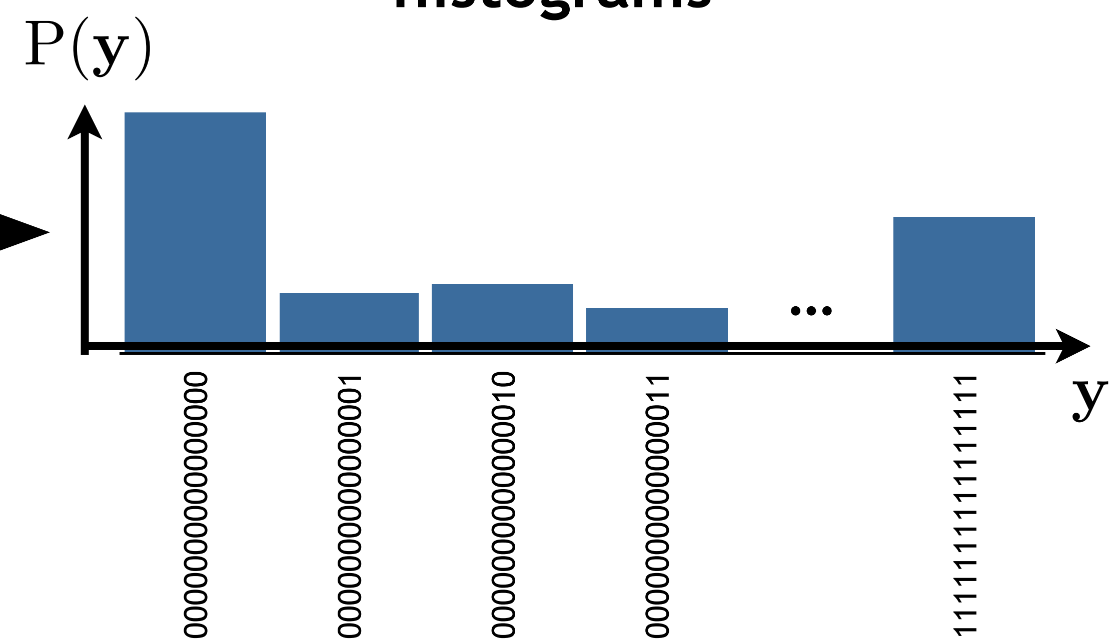
1	1	0	0	0	1	0	1	0	0
1	0	0	0	0	1	1	0	0	0
0	0	0	0	1	0	0	1	1	0
0	1	0	0	1	0	1	1	0	1
0	1	0	0	0	0	1	1	0	0
1	0	1	1	0	0	0	0	1	0
0	1	0	0	1	0	0	0	0	0
0	0	0	1	1	0	1	1	0	0
1	1	0	0	0	0	0	0	0	1
0	0	1	0	1	0	1	1	0	0
0	0	0	0	0	0	1	1	0	0
0	0	0	1	0	0	1	1	1	0
0	1	1	0	1	0	0	1	0	1
0	1	0	0	1	0	1	0	0	0
0	1	0	1	0	0	0	1	1	1
0	0	0	0	0	0	1	0	0	0

electrodes

time

2 ms

collecting histograms



comparing histograms

$$\mathrm{KL}[\mathrm{P}_1(\mathbf{y}) \parallel \mathrm{P}_2(\mathbf{y})]$$

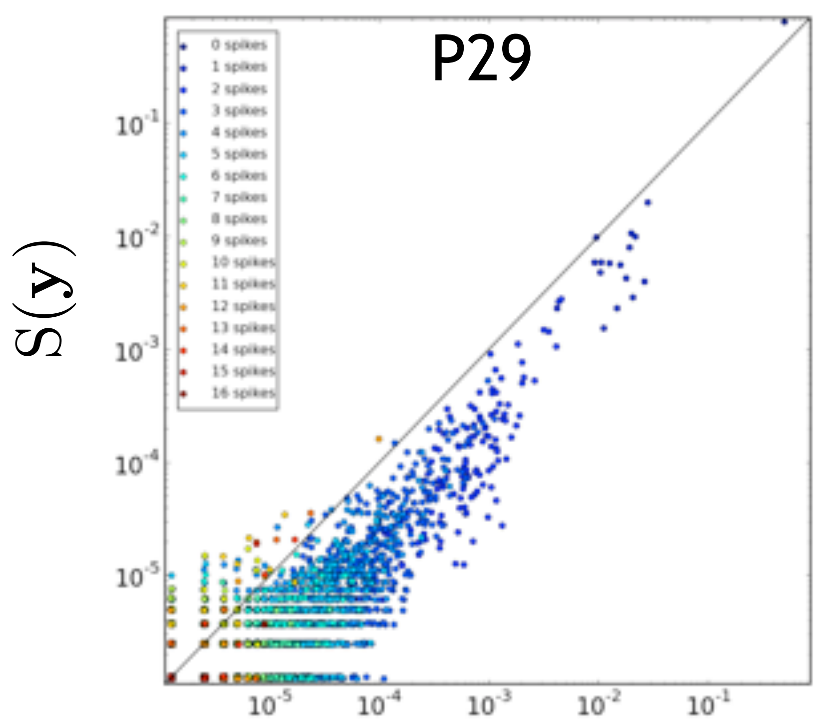
implicit marginalisation

rather than explicitly computing

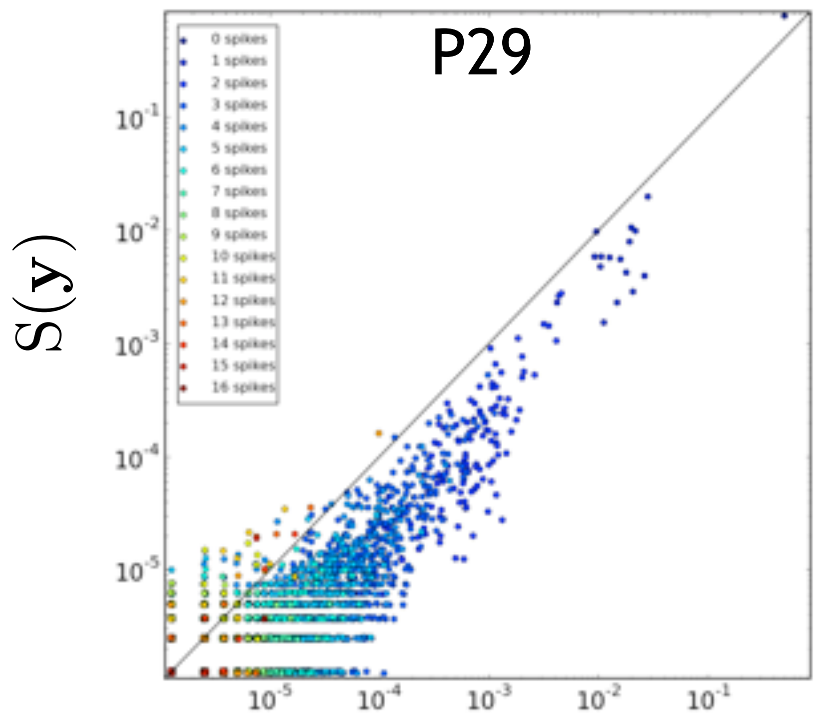
$$\mathrm{P}_{\mathrm{evoked}}(\mathbf{y}) = \int \mathrm{P}_\theta(\mathbf{y}|\mathbf{x})\ \mathrm{P}^*_{\mathrm{stim}}(\mathbf{x})\ d\mathbf{x}$$

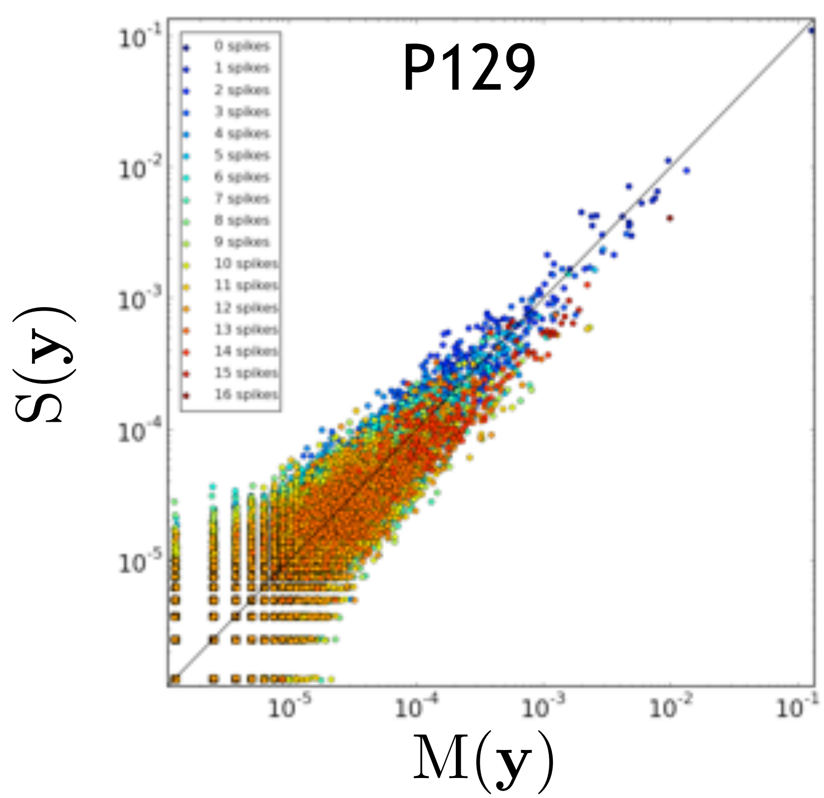
we collect evoked activity under each condition without regard to the actual ongoing stimulus

DEVELOPMENTAL CHANGES

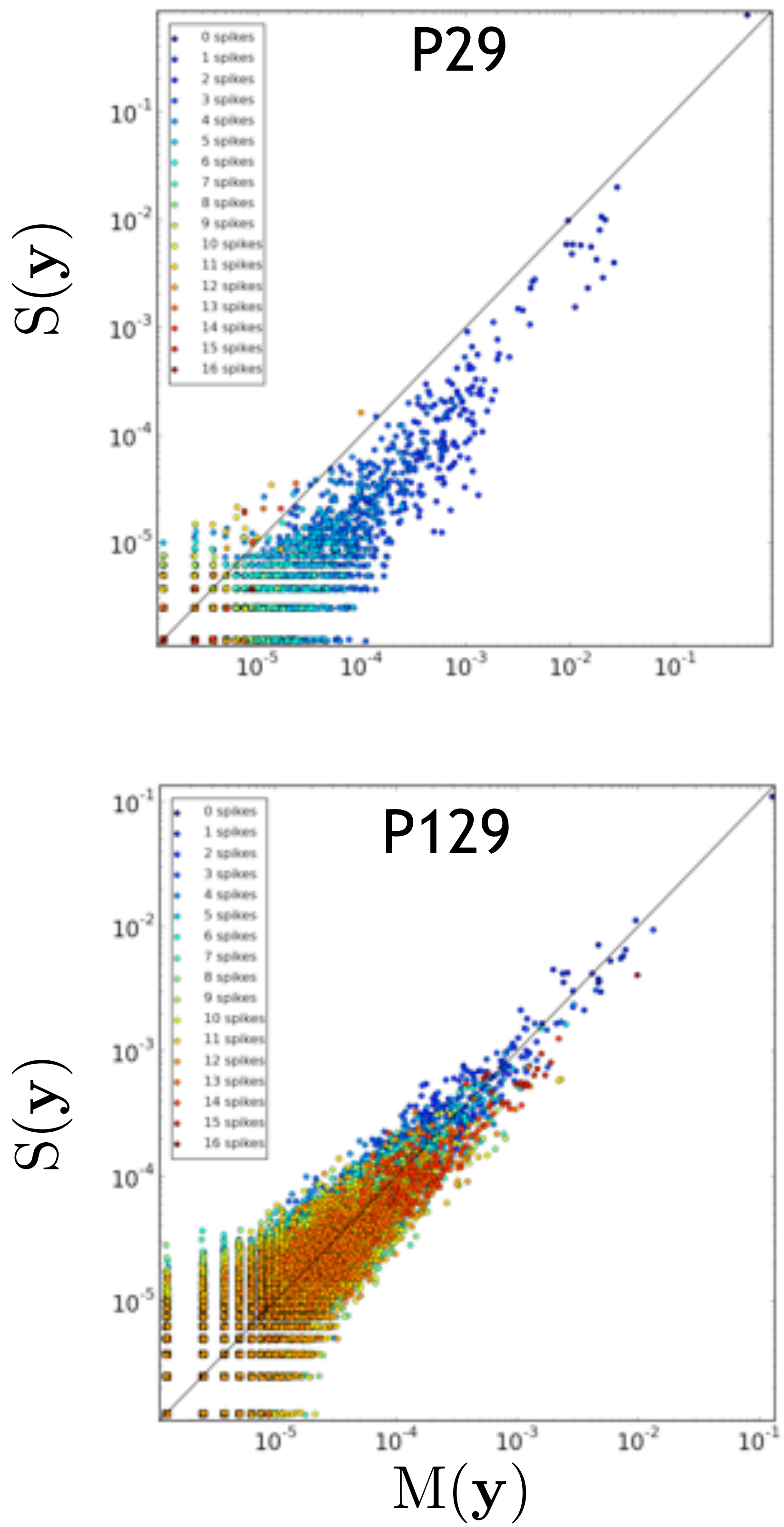


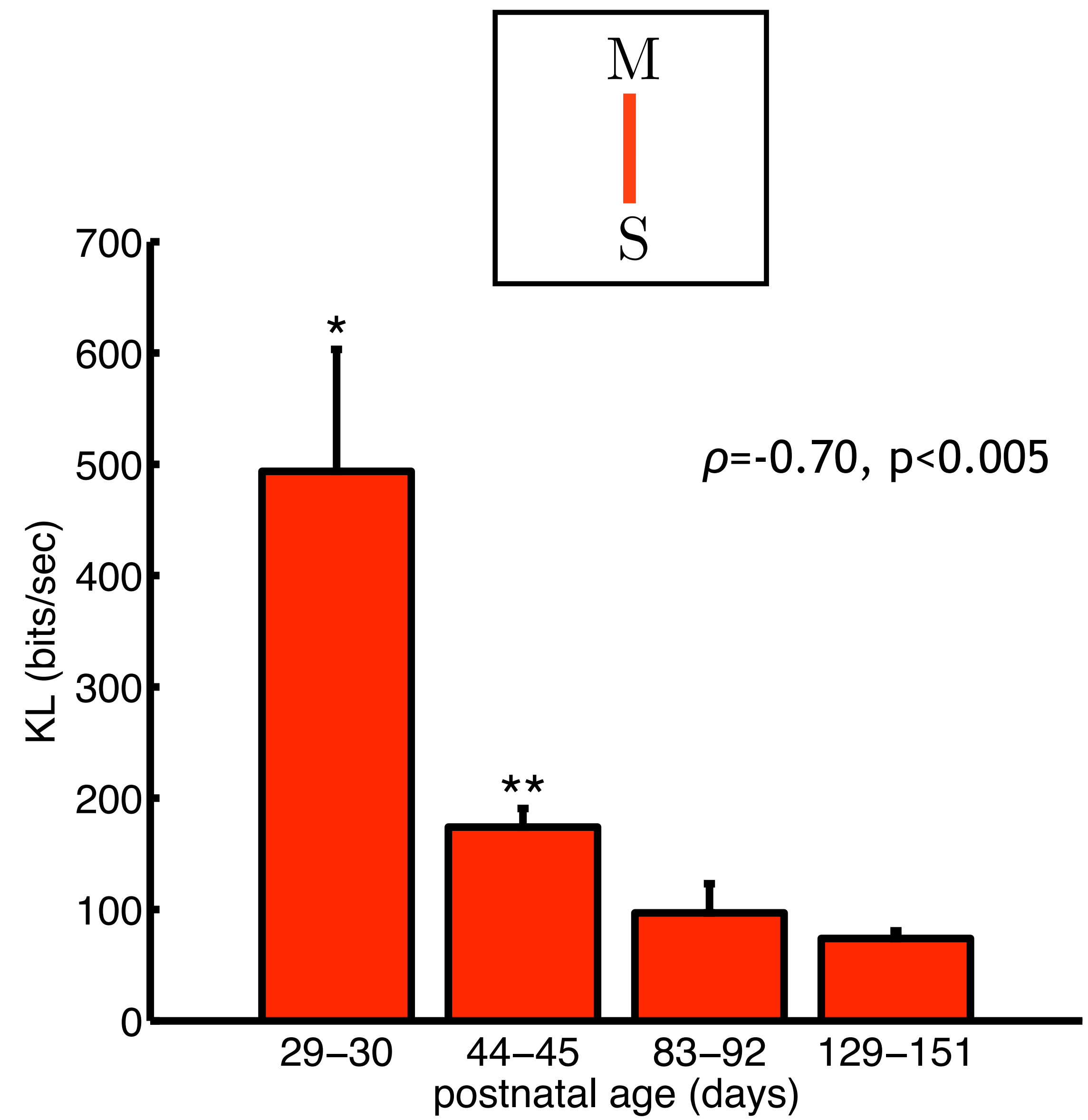
DEVELOPMENTAL CHANGES





DEVELOPMENTAL CHANGES





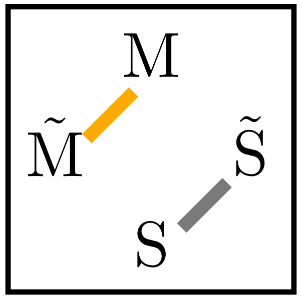
SPATIAL CORRELATIONS

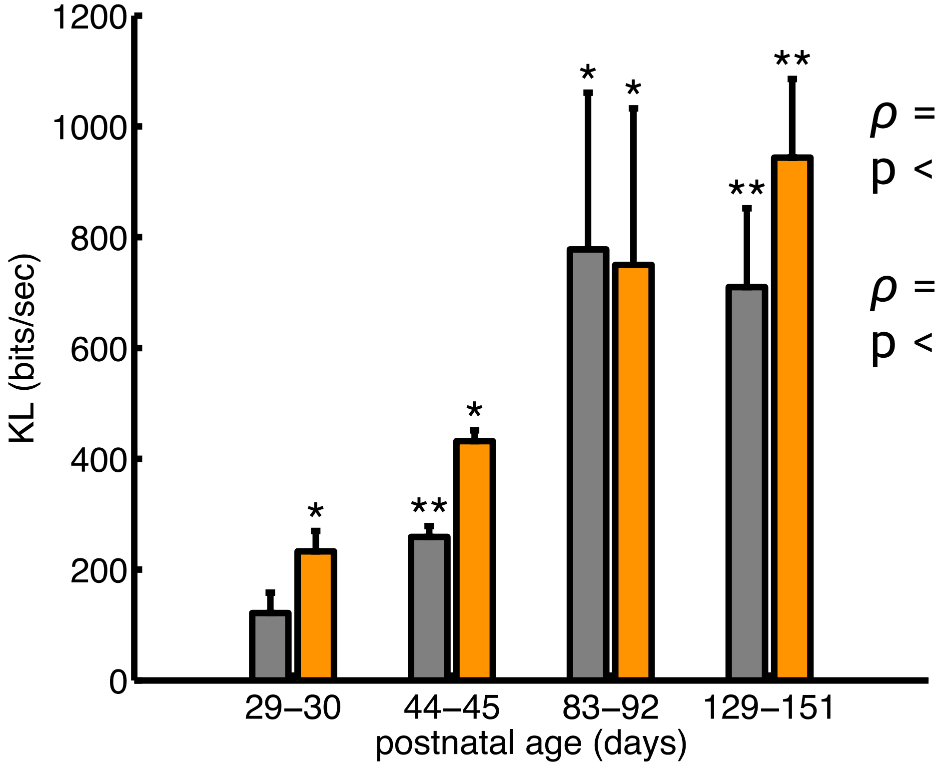
destroying spatial dependencies: $\tilde{\mathrm{P}}(\mathbf{y}) = \prod_{i=1}^{16} \mathrm{P}(y_i)$

SPATIAL CORRELATIONS

destroying spatial dependencies: $\tilde{\mathrm{P}}(\mathbf{y}) = \prod_{i=1}^{16} \mathrm{P}(y_i)$

within-condition correlations





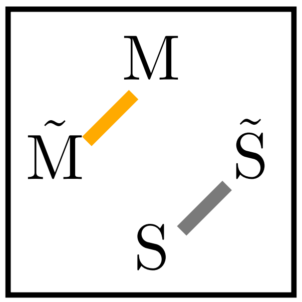
ρ = 0.73
$p < 0.005$

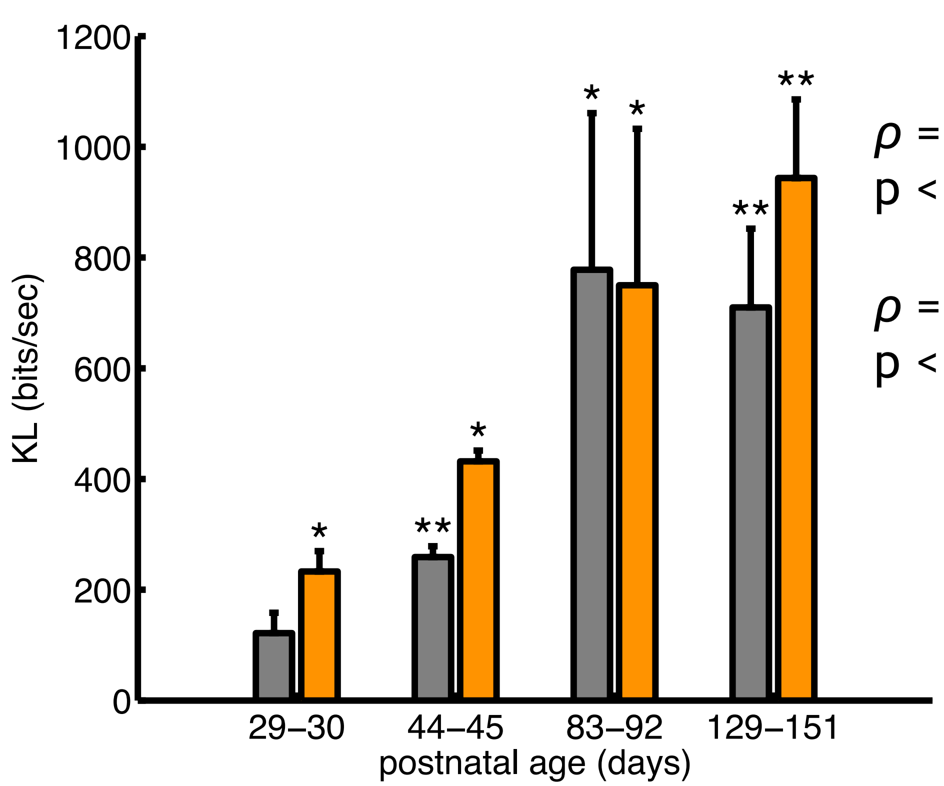
ρ = -0.76
$p < 0.005$

SPATIAL CORRELATIONS

destroying spatial dependencies: $\tilde{\mathrm{P}}(\mathbf{y}) = \prod_{i=1}^{16} \mathrm{P}(y_i)$

within-condition correlations

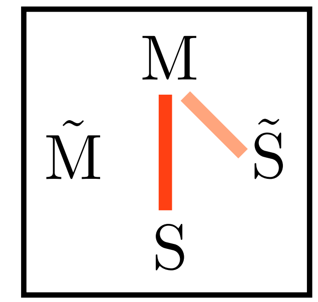


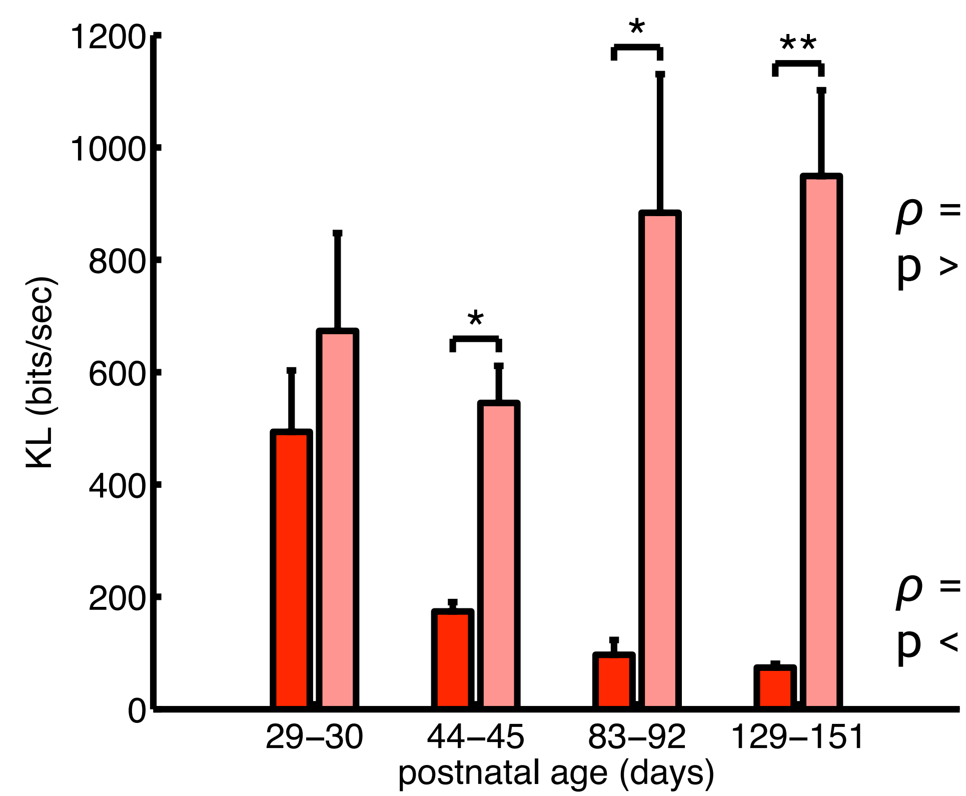


ρ = 0.73
$p < 0.005$

ρ = -0.76
$p < 0.005$

importance of correlations for match between conditions





ρ = 0.34
$p > 0.2$

ρ = -0.70
$p < 0.005$

TEMPORAL CORRELATIONS

state-transition distributions: $\mathrm{P}(\mathbf{y}_{t+\tau}|\mathbf{y}_t)$

TEMPORAL CORRELATIONS

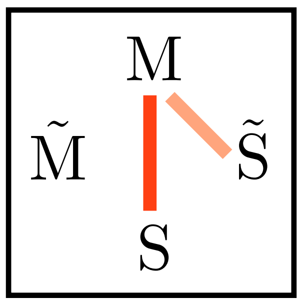
state-transition distributions: $\mathrm{P}(\mathbf{y}_{t+\tau}|\mathbf{y}_t)$

control: $\tilde{\mathrm{P}}(\mathbf{y}_{t+\tau}|\mathbf{y}_t) = \mathrm{P}(\mathbf{y})$

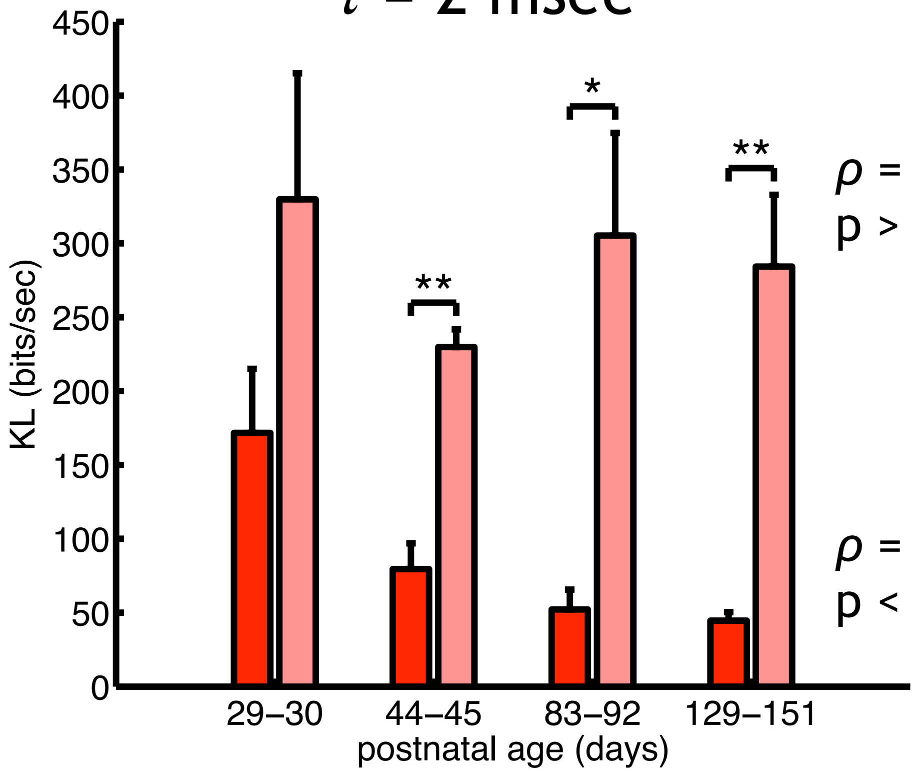
TEMPORAL CORRELATIONS

state-transition distributions: $\mathrm{P}(\mathbf{y}_{t+\tau}|\mathbf{y}_t)$

control: $\tilde{\mathrm{P}}(\mathbf{y}_{t+\tau}|\mathbf{y}_t) = \mathrm{P}(\mathbf{y})$



τ = 2 msec



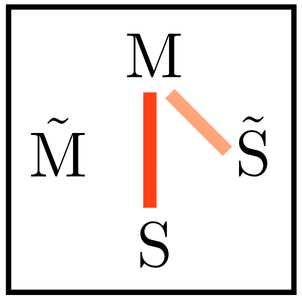
ρ = 0.03
p > 0.9

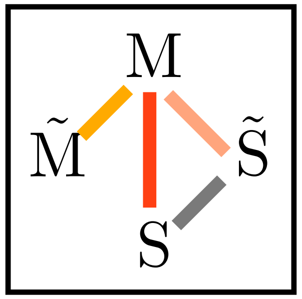
ρ = -0.72
p < 0.005

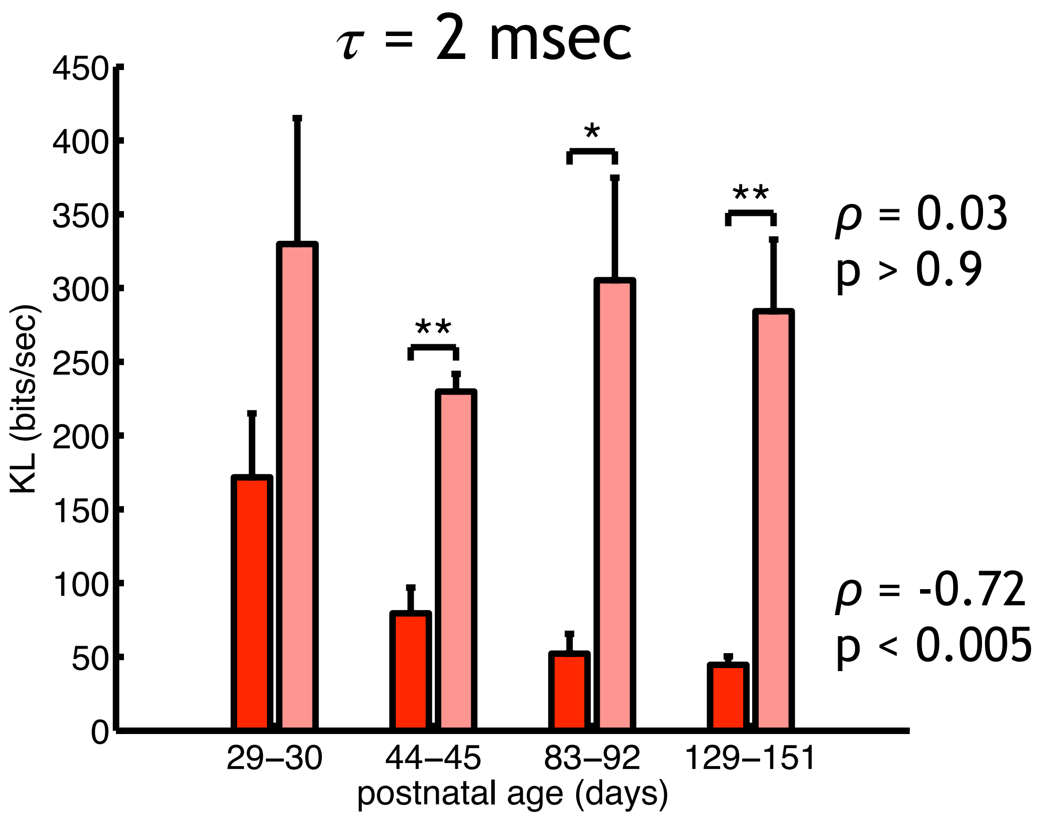
TEMPORAL CORRELATIONS

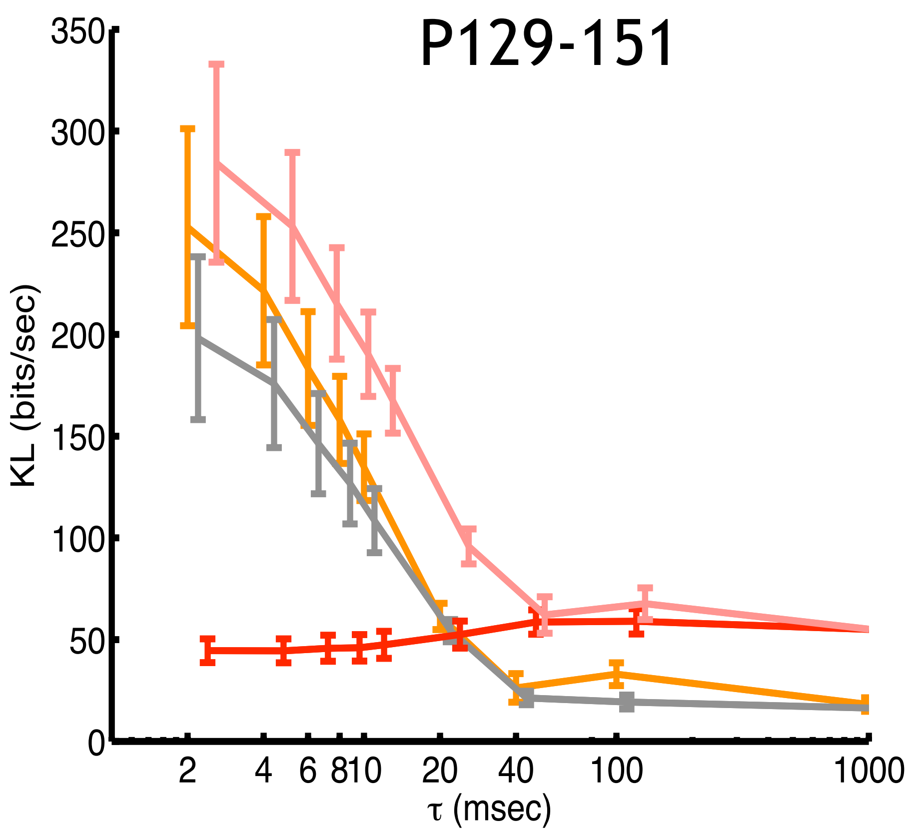
state-transition distributions: $\mathrm{P}(\mathbf{y}_{t+\tau}|\mathbf{y}_t)$

control: $\tilde{\mathrm{P}}(\mathbf{y}_{t+\tau}|\mathbf{y}_t) = \mathrm{P}(\mathbf{y})$



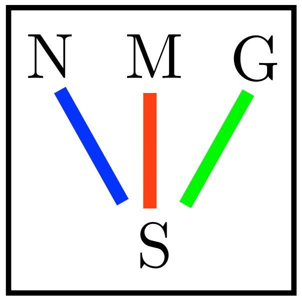


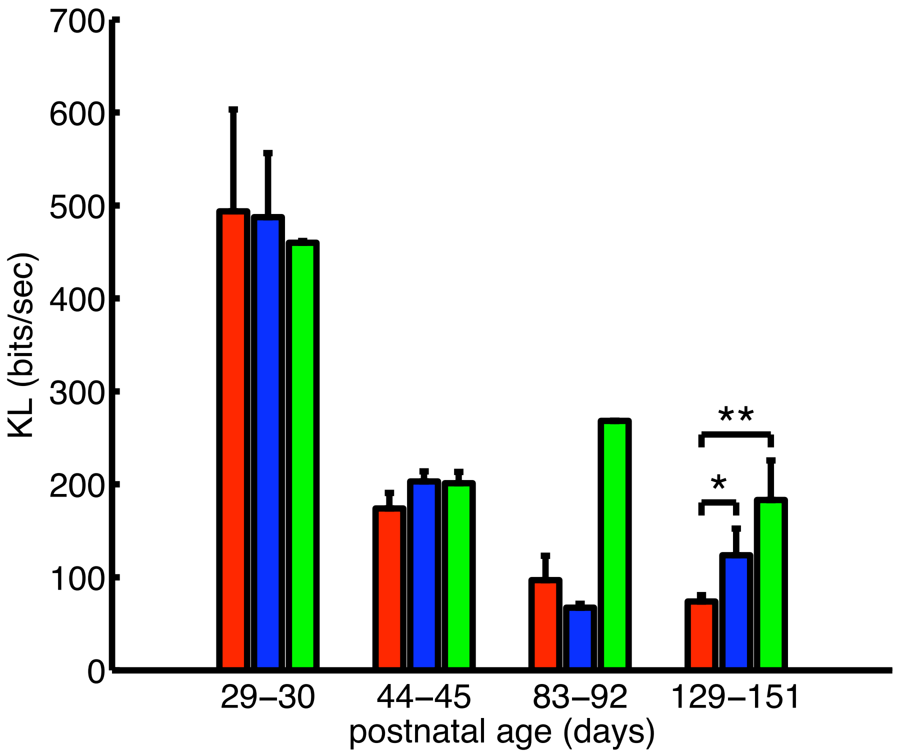




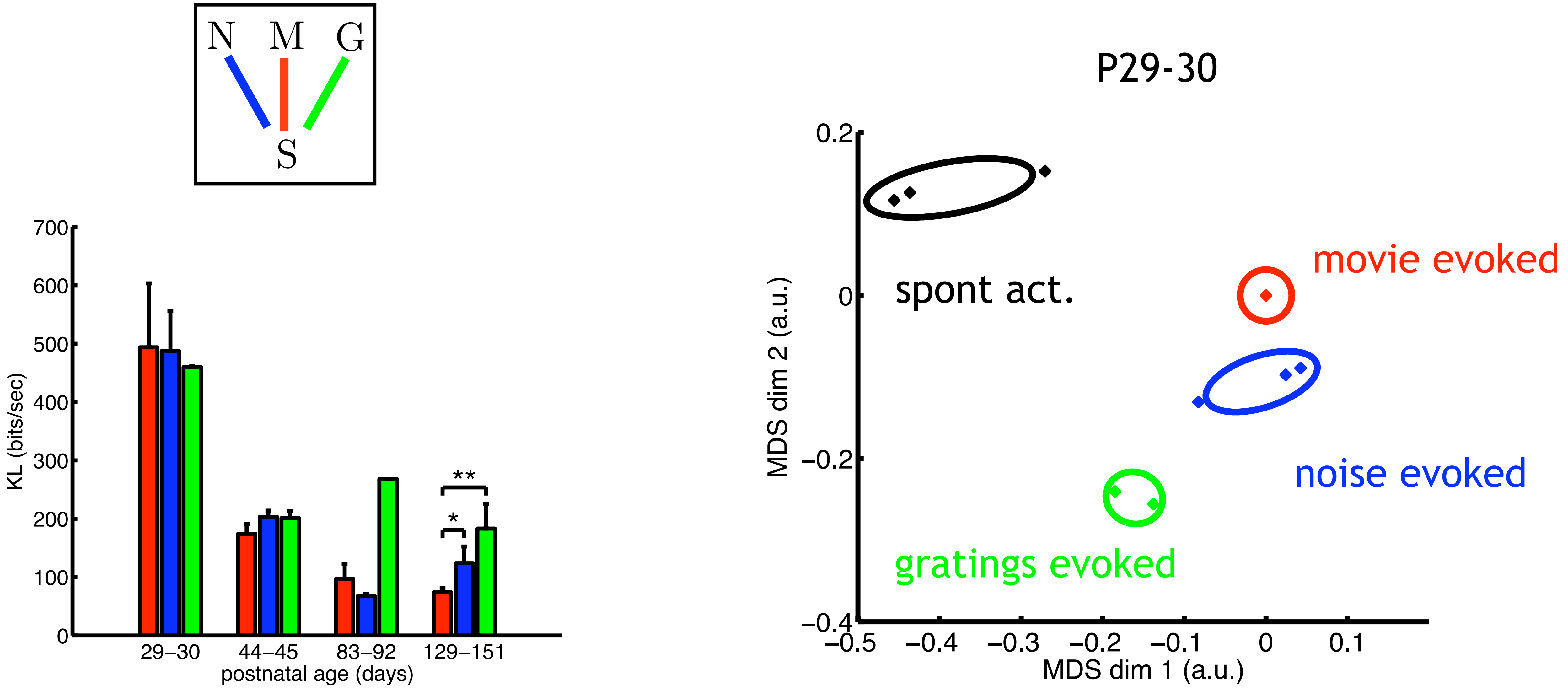
NON-NATURAL STIMULUS ENSEMBLES

NON-NATURAL STIMULUS ENSEMBLES

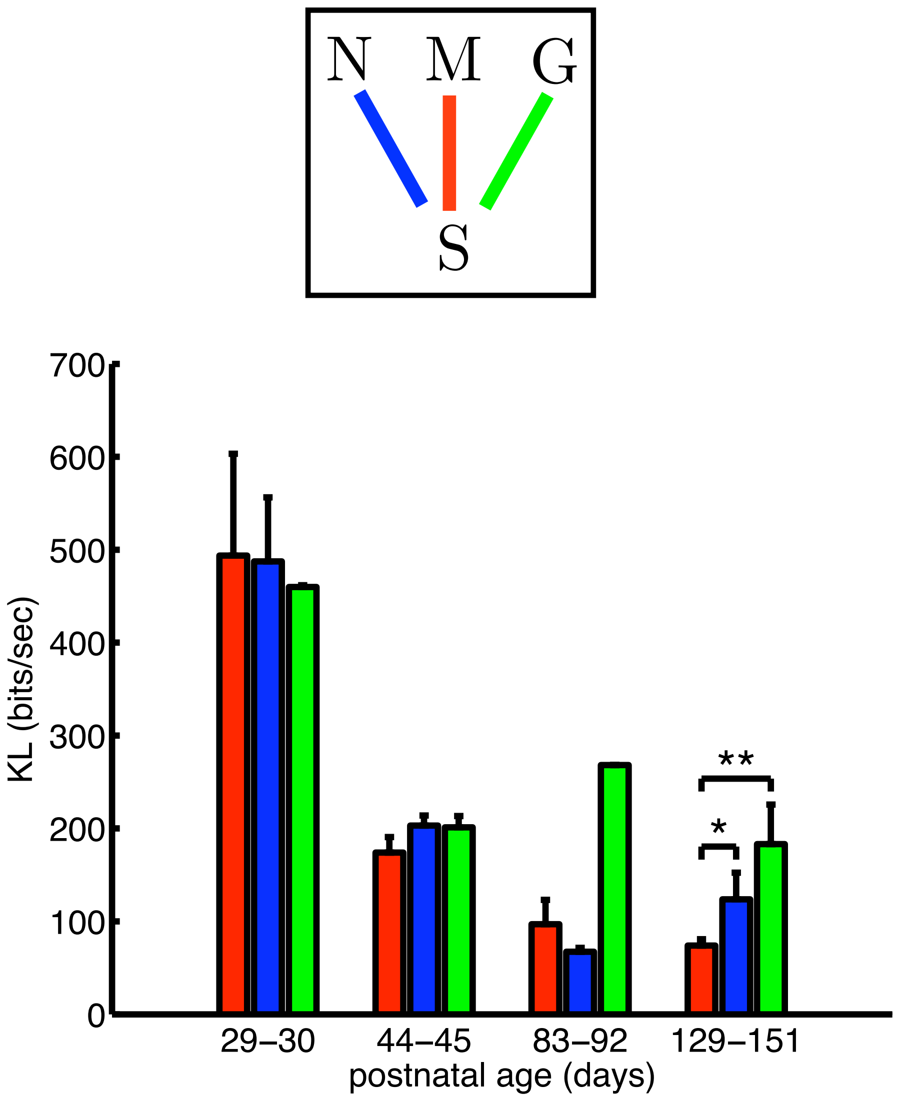


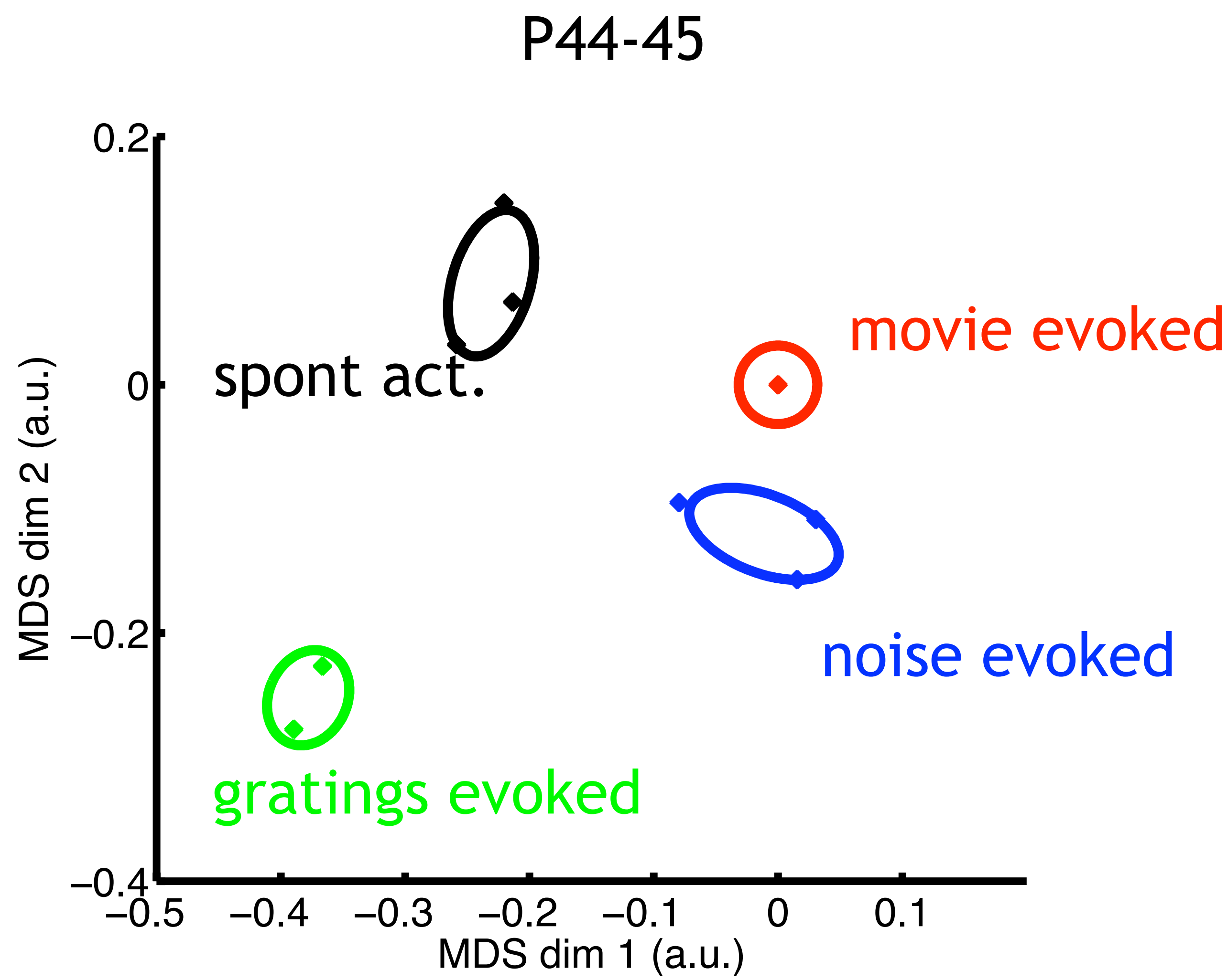


NON-NATURAL STIMULUS ENSEMBLES

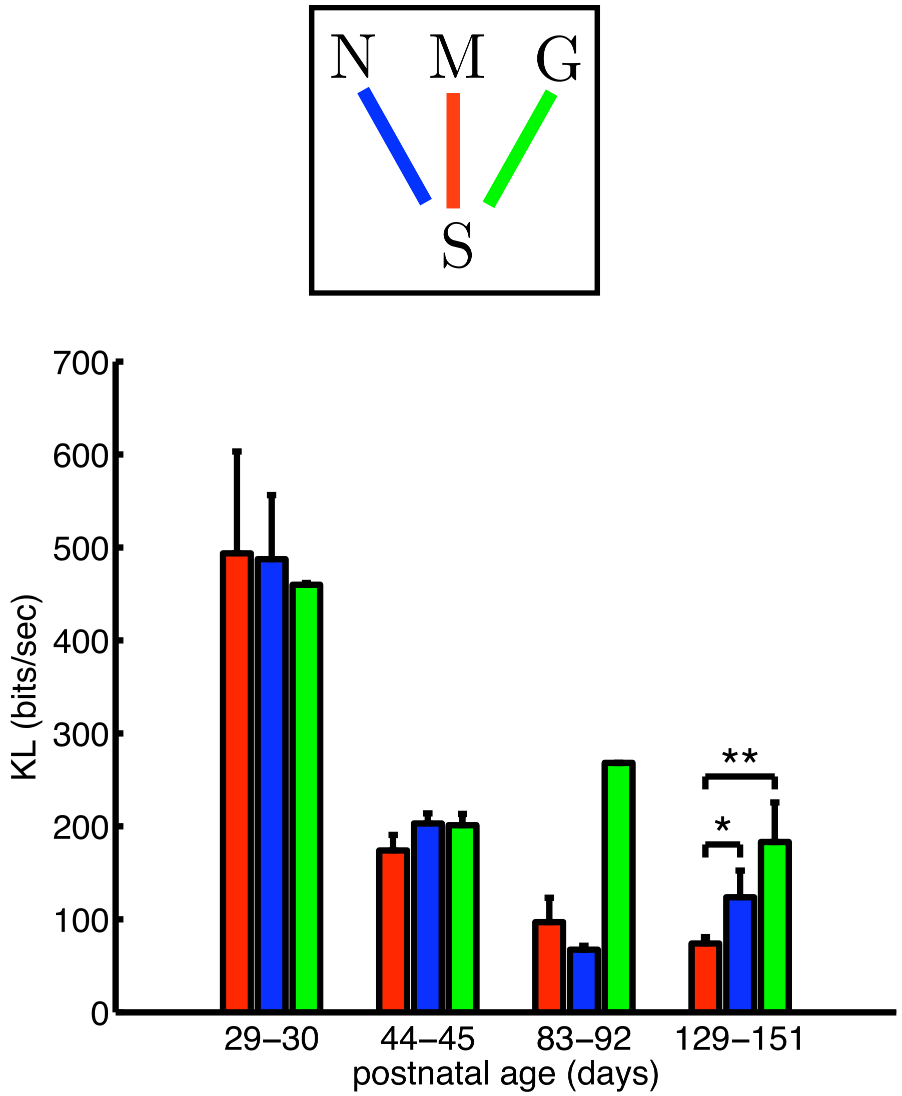


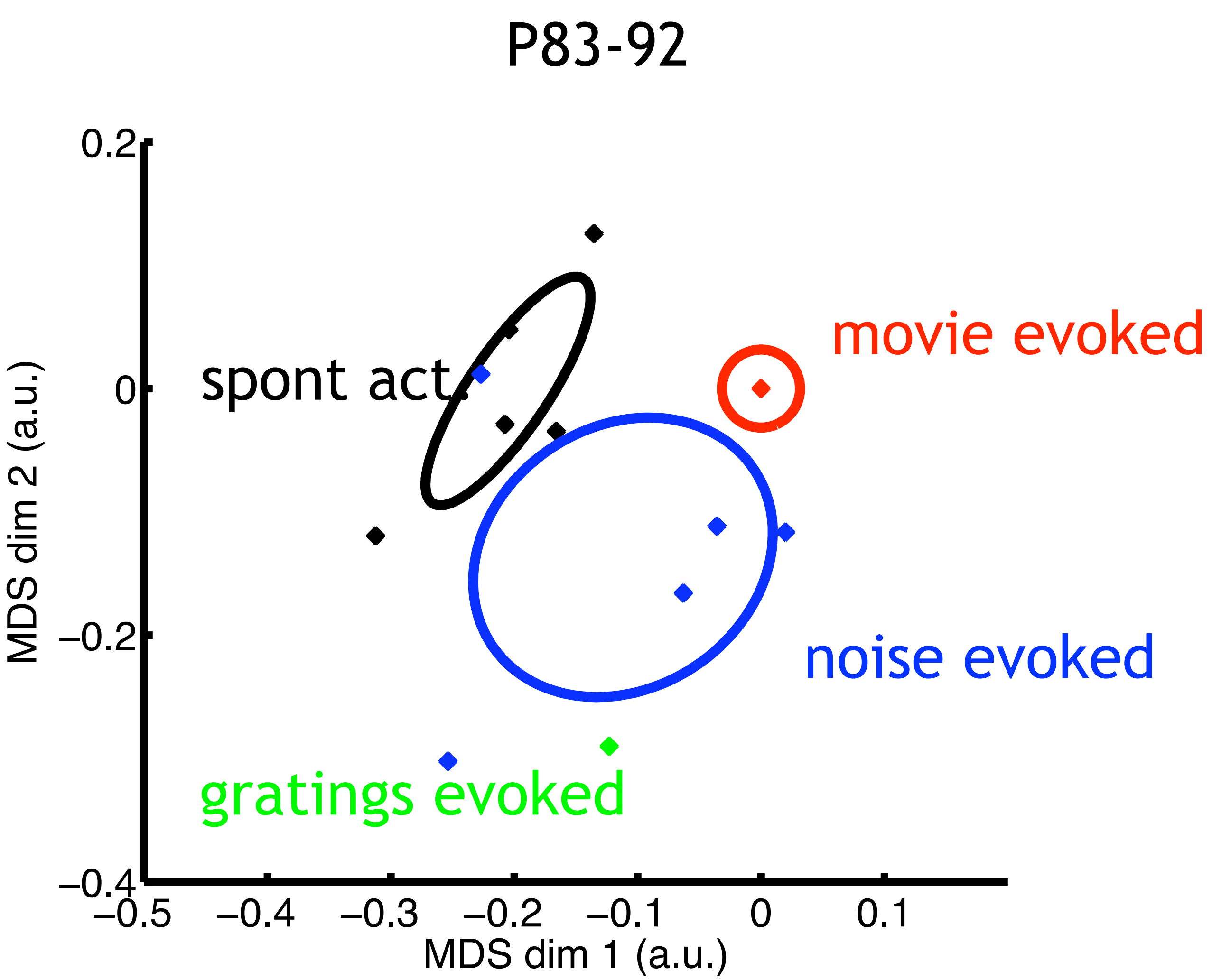
NON-NATURAL STIMULUS ENSEMBLES





NON-NATURAL STIMULUS ENSEMBLES





NON-NATURAL STIMULUS ENSEMBLES