Statisztikus tanulás az idegrendszerben

ORBÁN GERGŐ

golab.wigner.mta.hu

WIGNER CENTRE RESEARCH FOR PHYSICS

Introduction

Knowledge representation

Probabilistic models

Bayesian behaviour

Approximate inference I (computer lab)

Vision I

Approximate inference II: Sampling

Measuring priors

Neural representation of probabilities

Structure learning

Vision II

Decision making and reinforcement learning

elméleti

Introduction

Knowledge representation

Probabilistic models

Bayesian behaviour

Approximate inference I (computer lab)

Vision I

Approximate inference II: Sampling

Measuring priors

Neural representation of probabilities

Structure learning

Vision II

Decision making and reinforcement learning

elméleti

kognitív

Introduction

Knowledge representation

Probabilistic models

Bayesian behaviour

Approximate inference I (computer lab)

Vision I

Approximate inference II: Sampling

Measuring priors

Neural representation of probabilities

Structure learning

Vision II

Decision making and reinforcement learning

Introduction

Knowledge representation

Probabilistic models

Bayesian behaviour

Approximate inference I (computer lab)

Vision I

Approximate inference II: Sampling

Measuring priors

Neural representation of probabilities

Structure learning

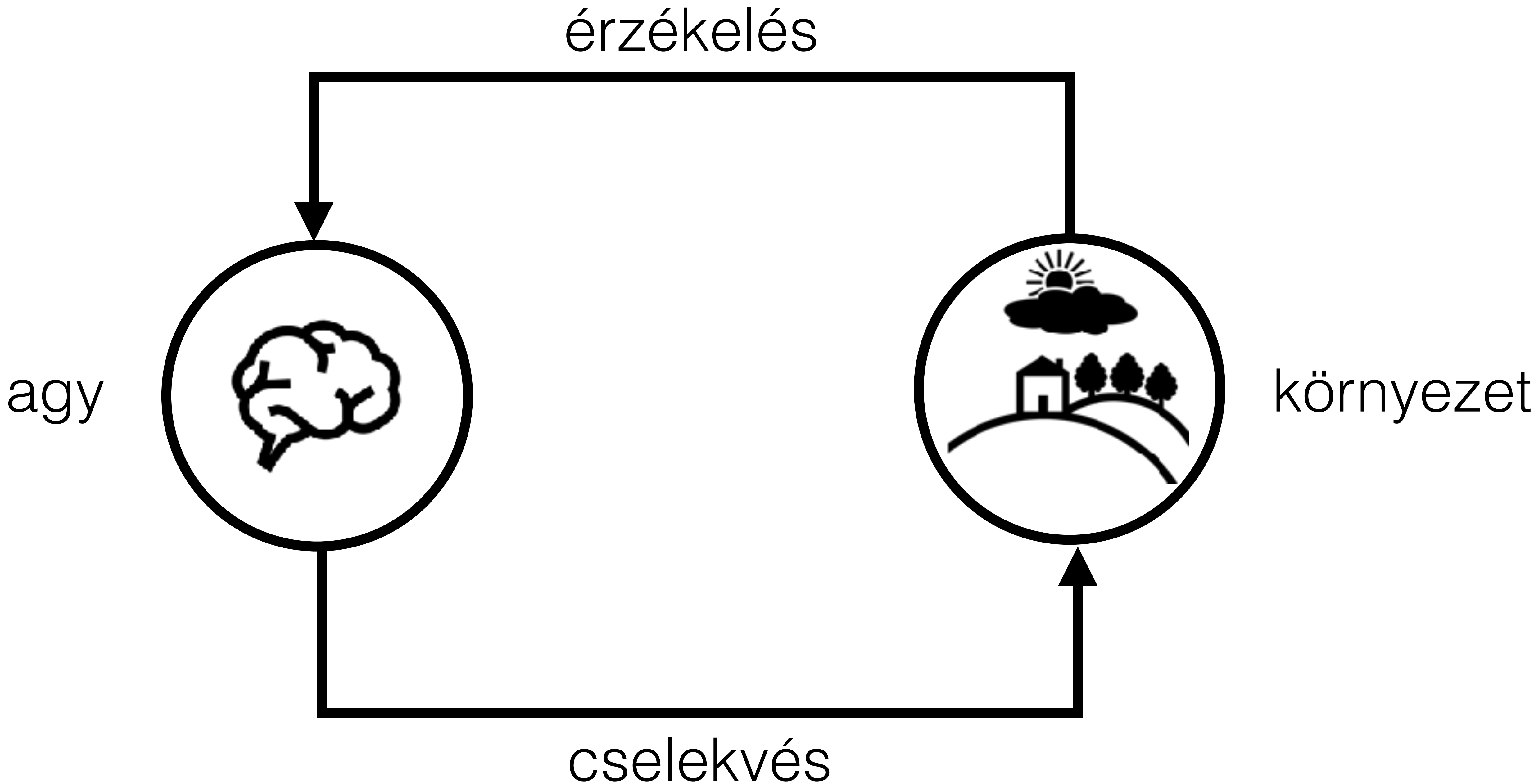
Vision II

Decision making and reinforcement learning

elméleti

kognitív

neurális

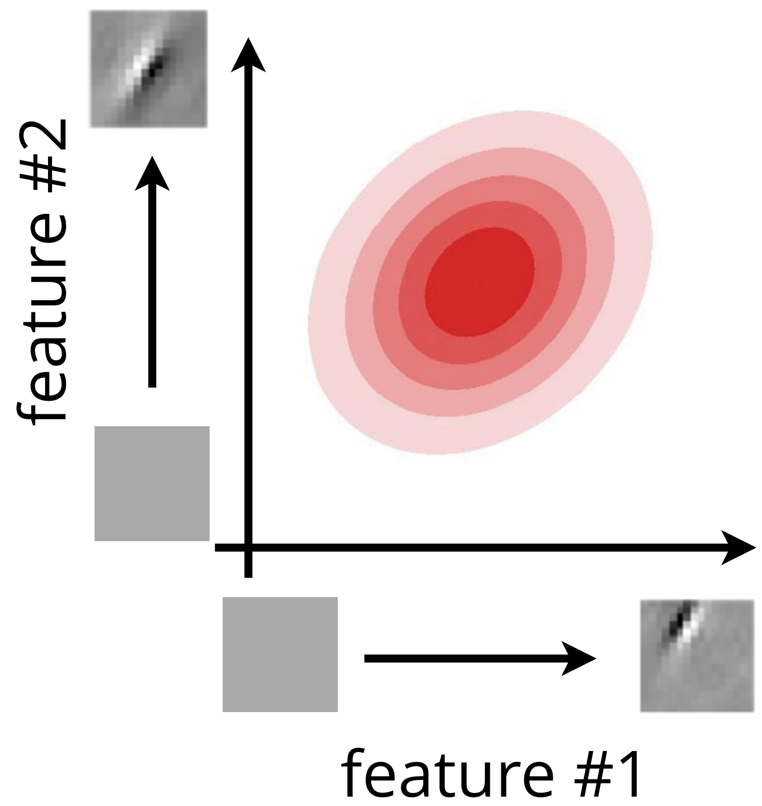

érzékelés
agy
környezet
cselekvés

mean responses

$$\mathrm{P}(a_1, a_2 \,|\, \mathrm{image}, c)$$

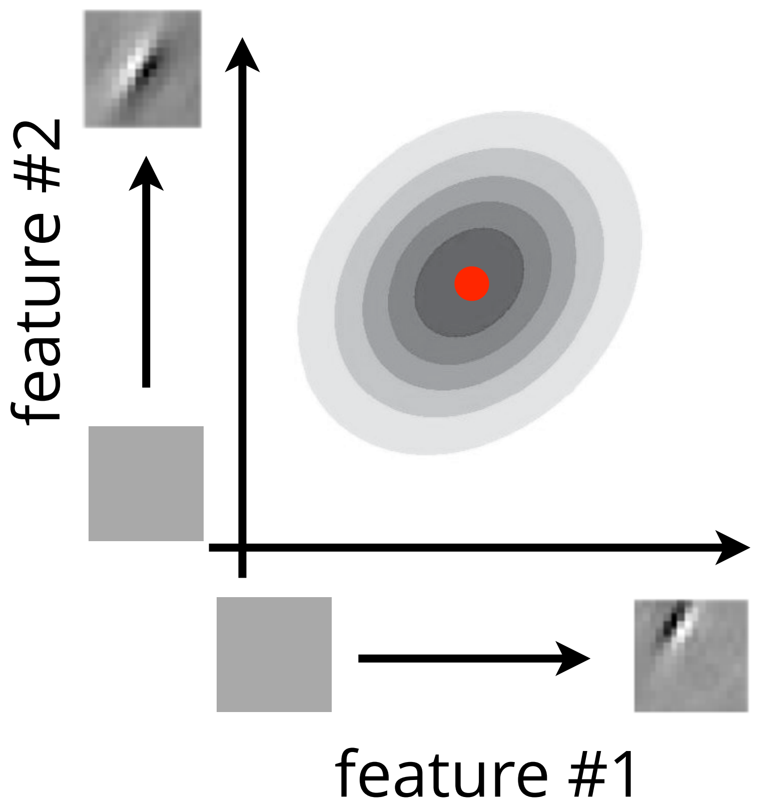
mean responses

$\mathrm{P}(a_1, a_2 \,|\, \mathrm{image}, c)$

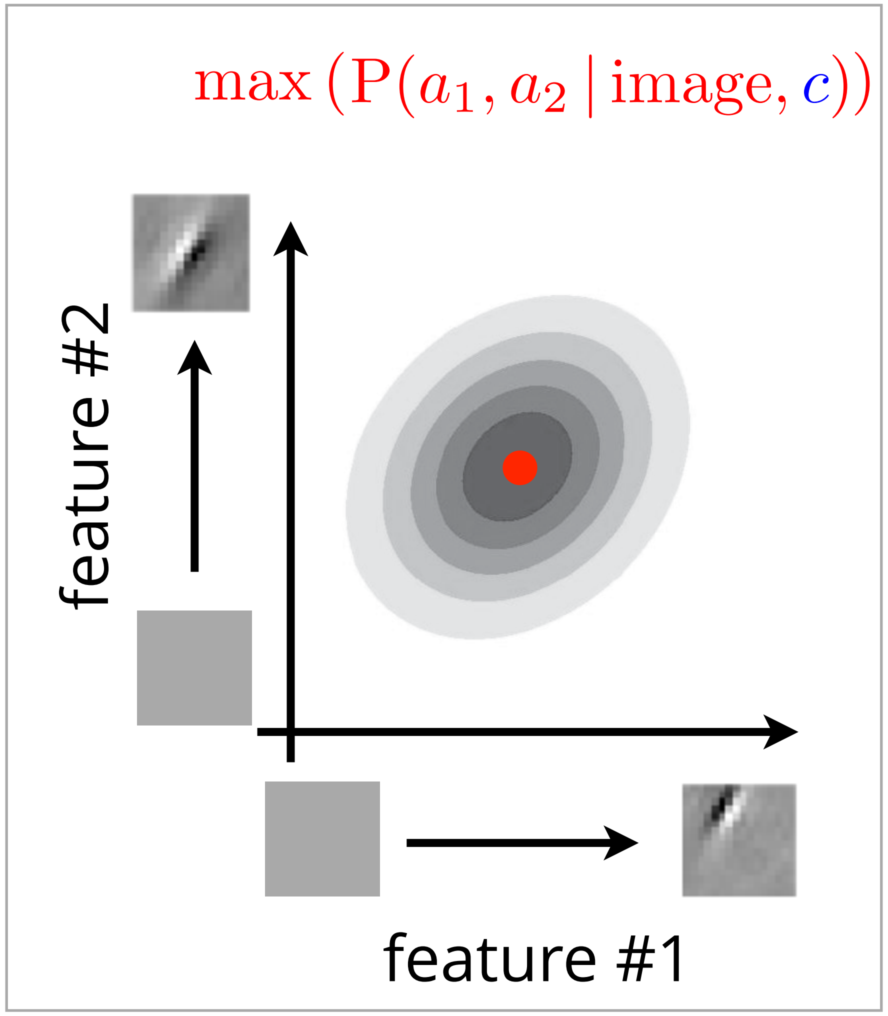


mean responses

$$\max\left(\mathrm{P}(a_1, a_2 \,|\, \text{image}, c)\right)$$



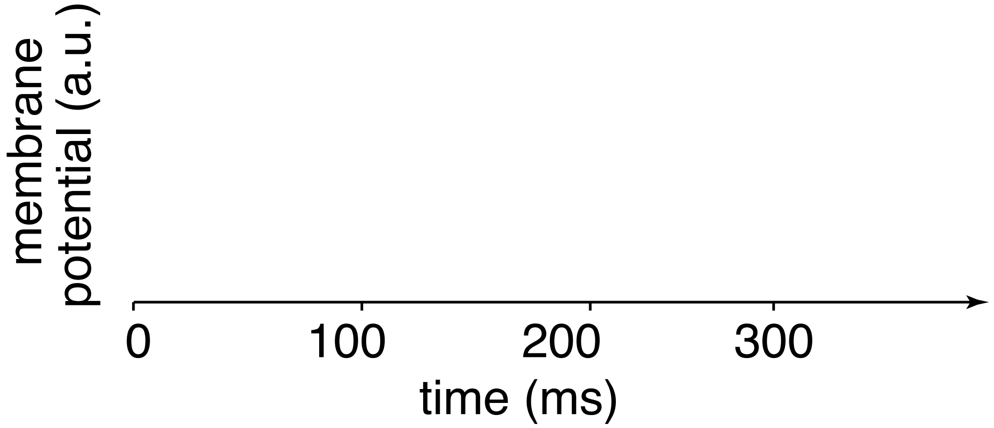
mean responses



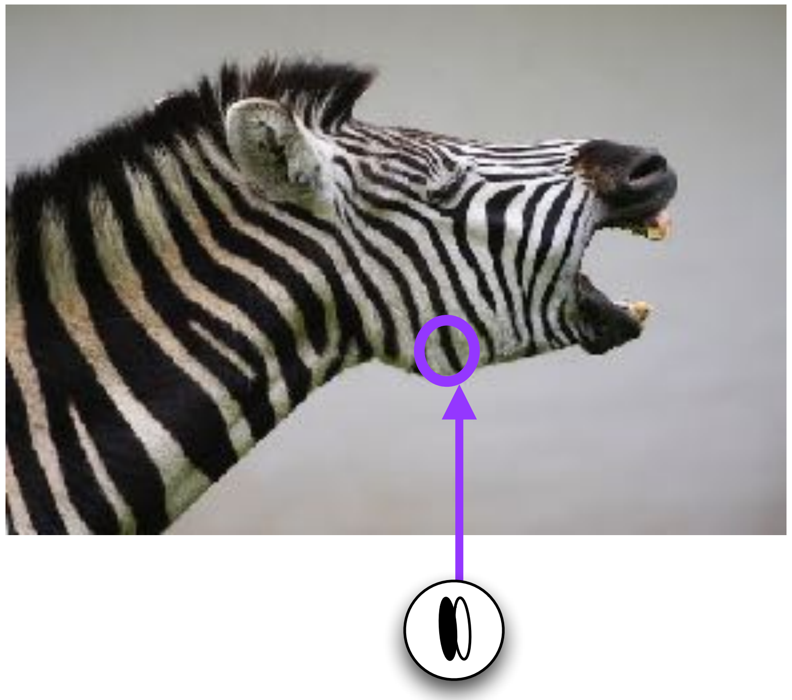
traditional theories
e.g. Olshausen & Field, Nature 1996, Schwartz & Simoncelli, Nat Neurosci 2001

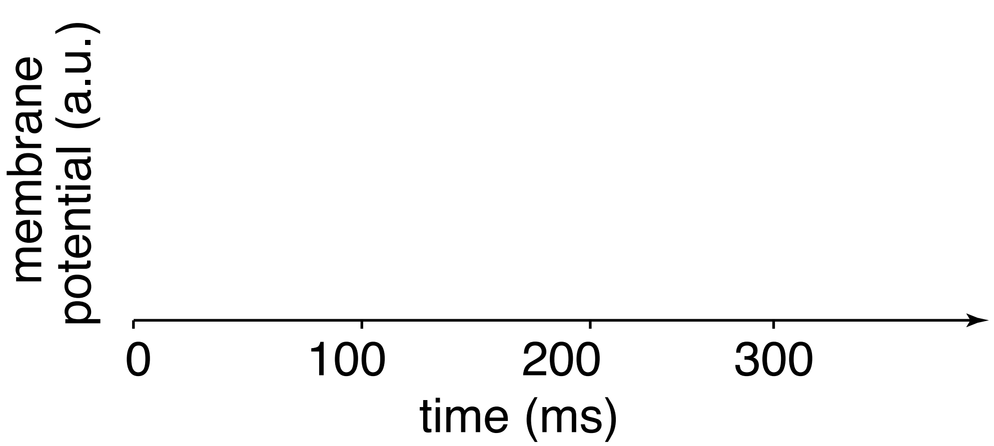
mean response $\rightsquigarrow$ maximum a posteriori inference

stochastic sampling

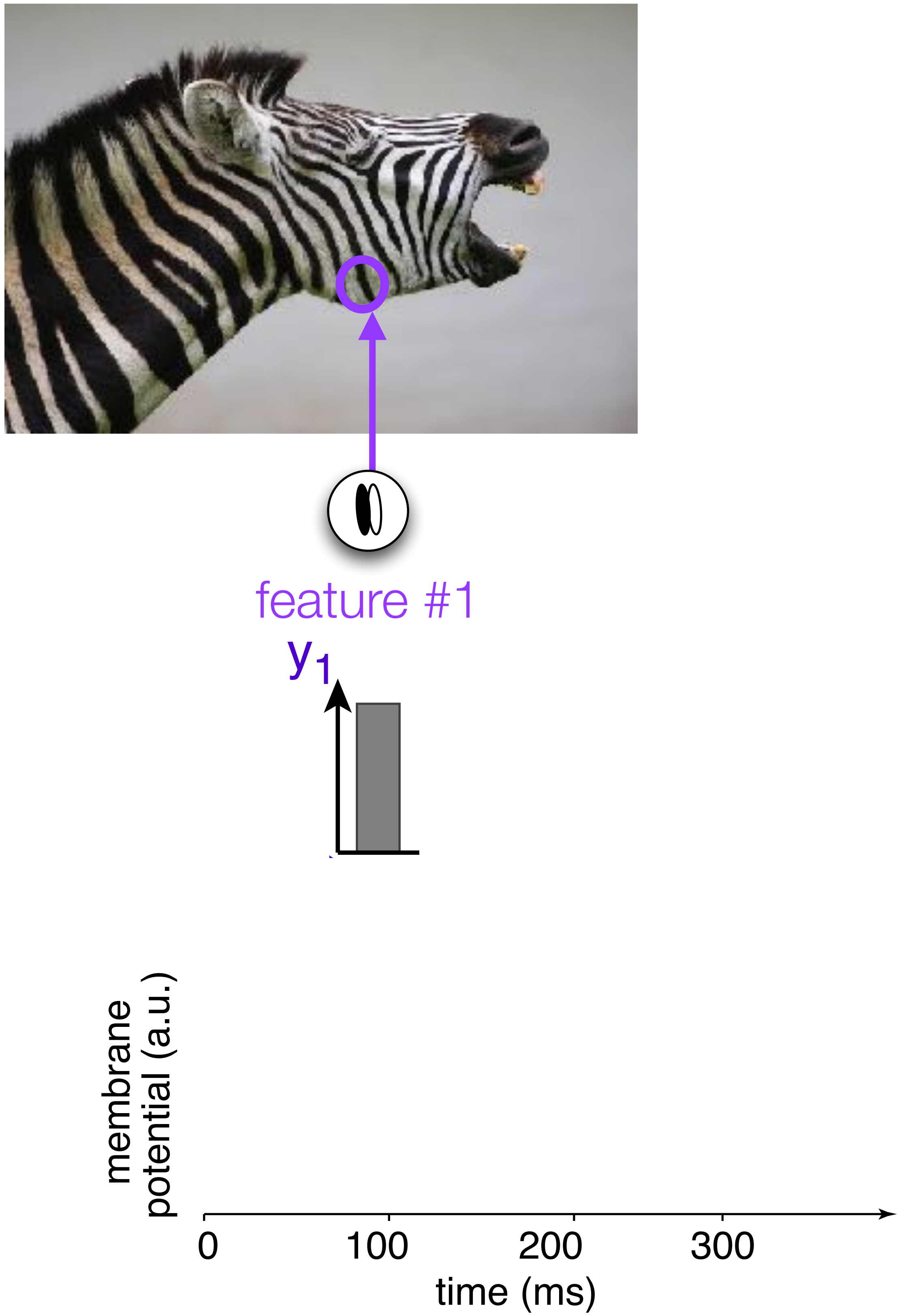


stochastic sampling

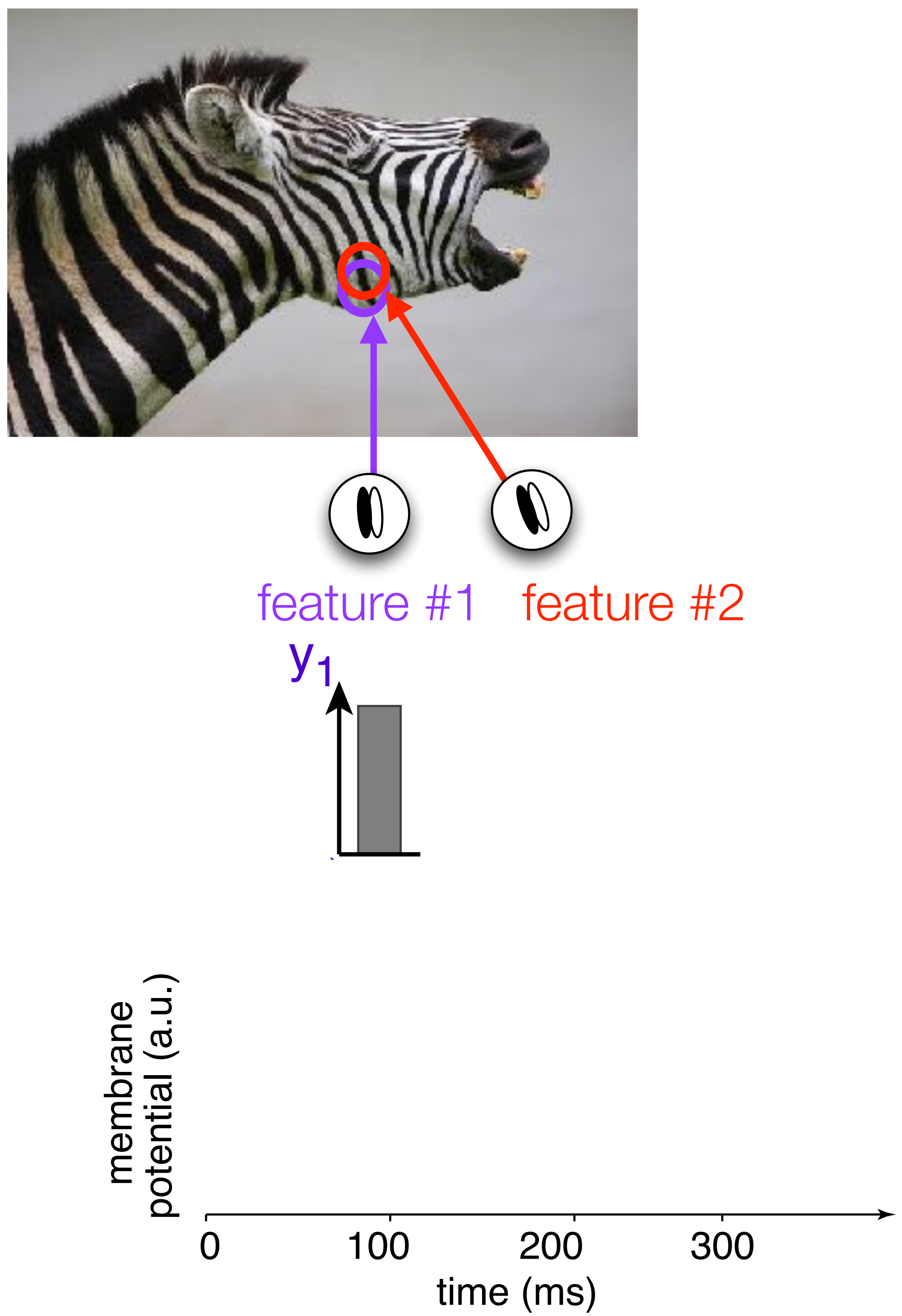




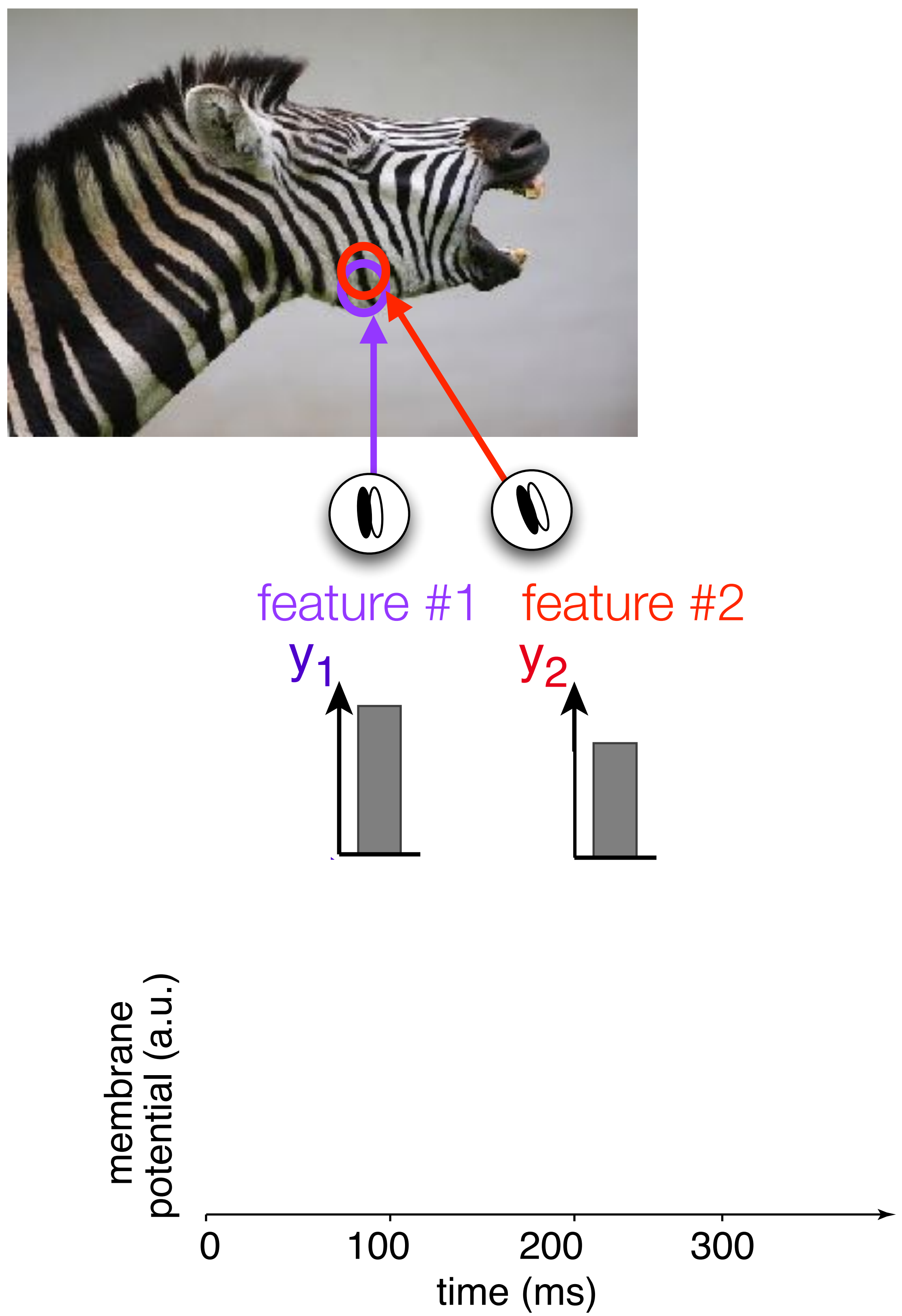
stochastic sampling



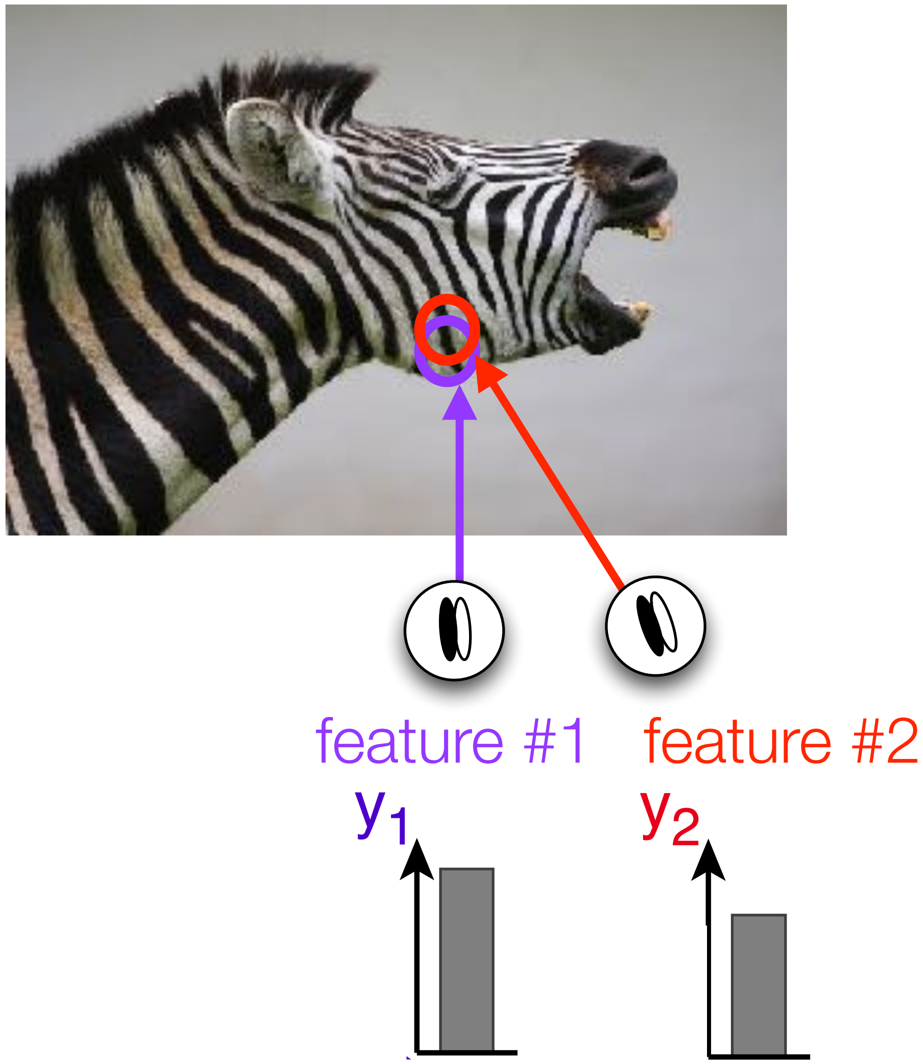
stochastic sampling

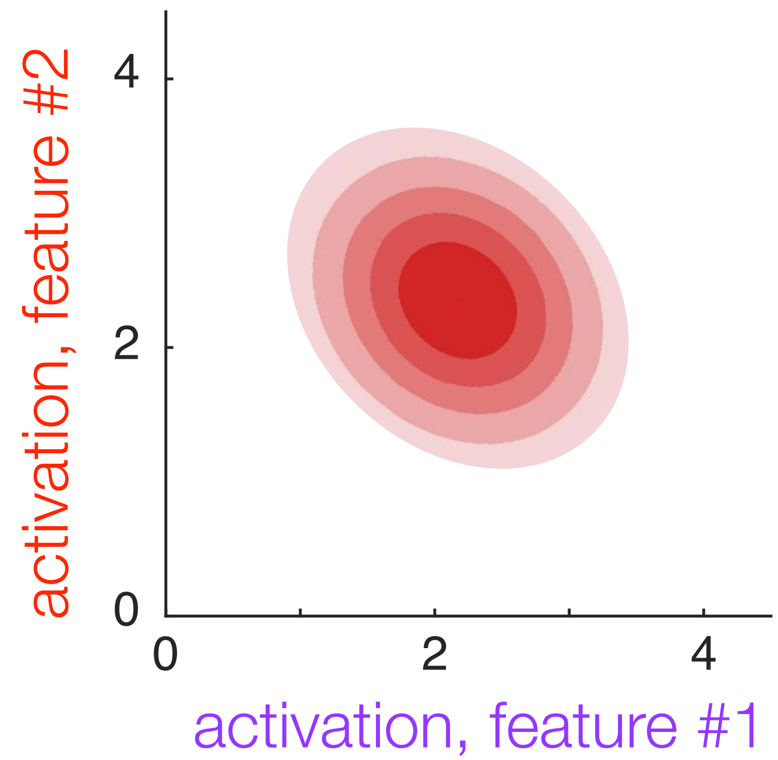


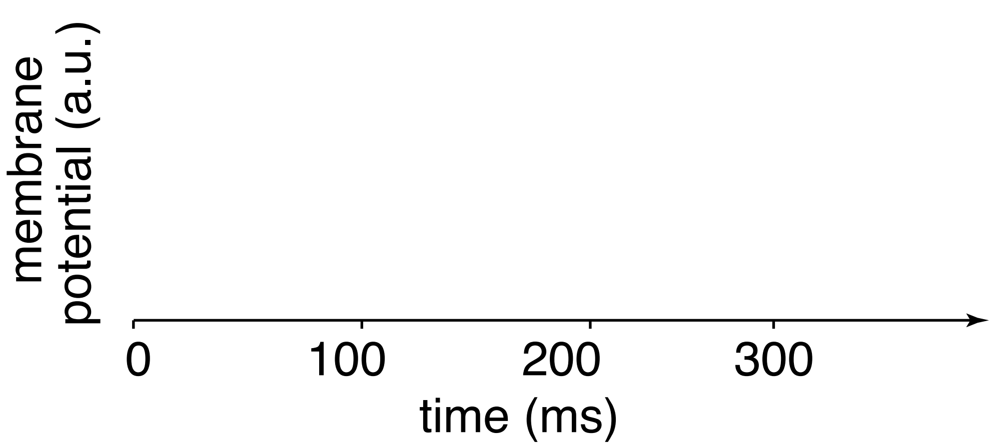
stochastic sampling



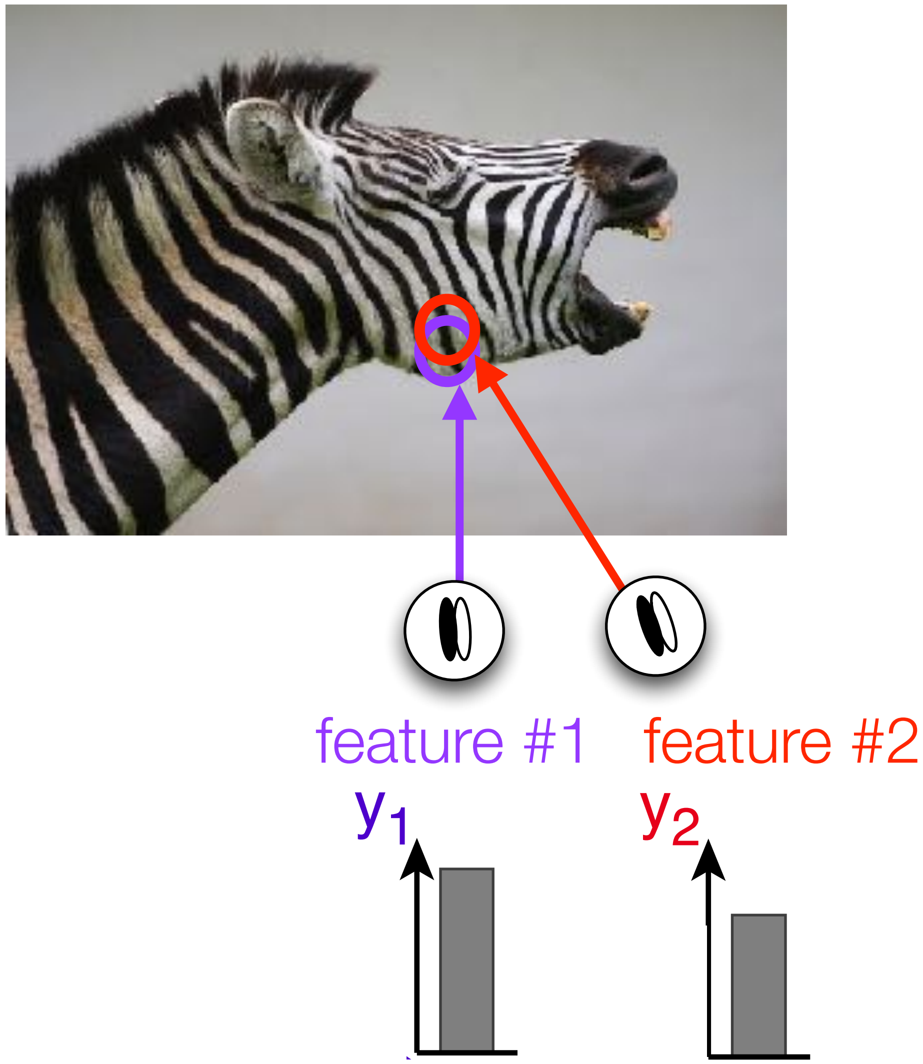
stochastic sampling

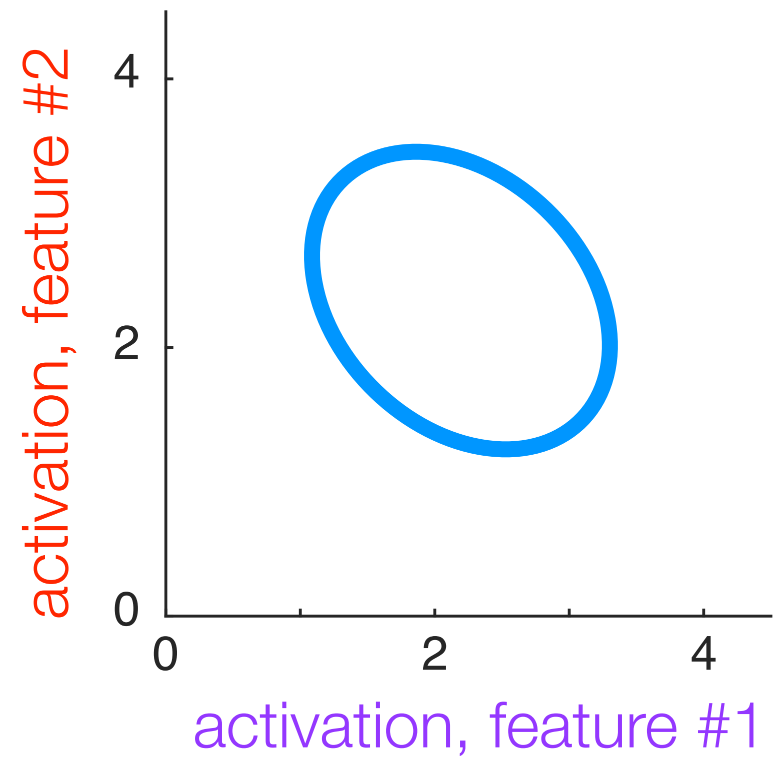


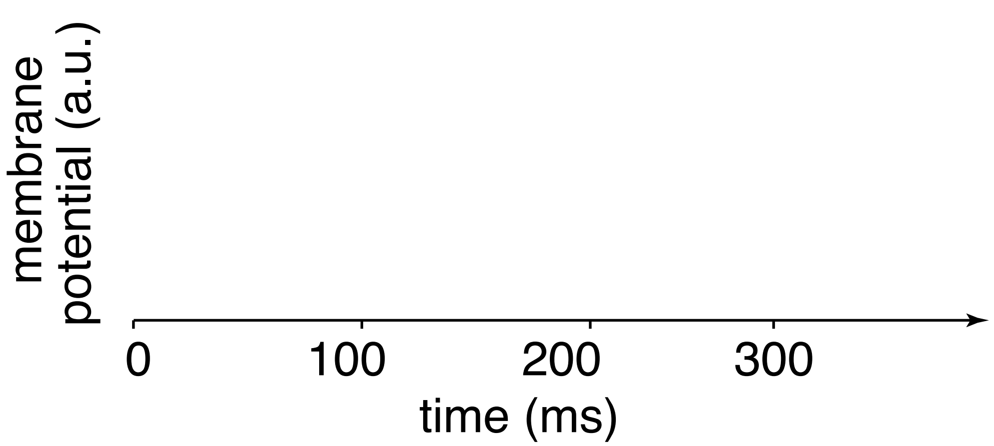




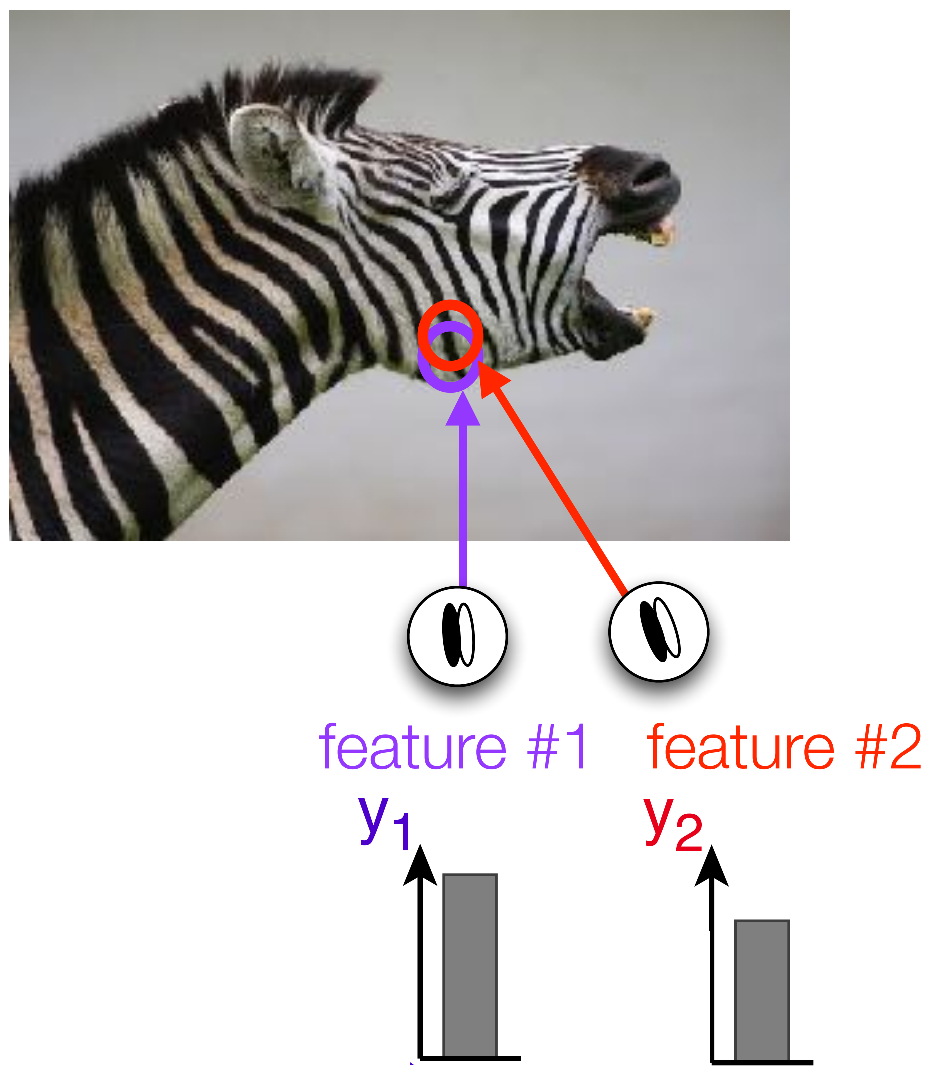
stochastic sampling

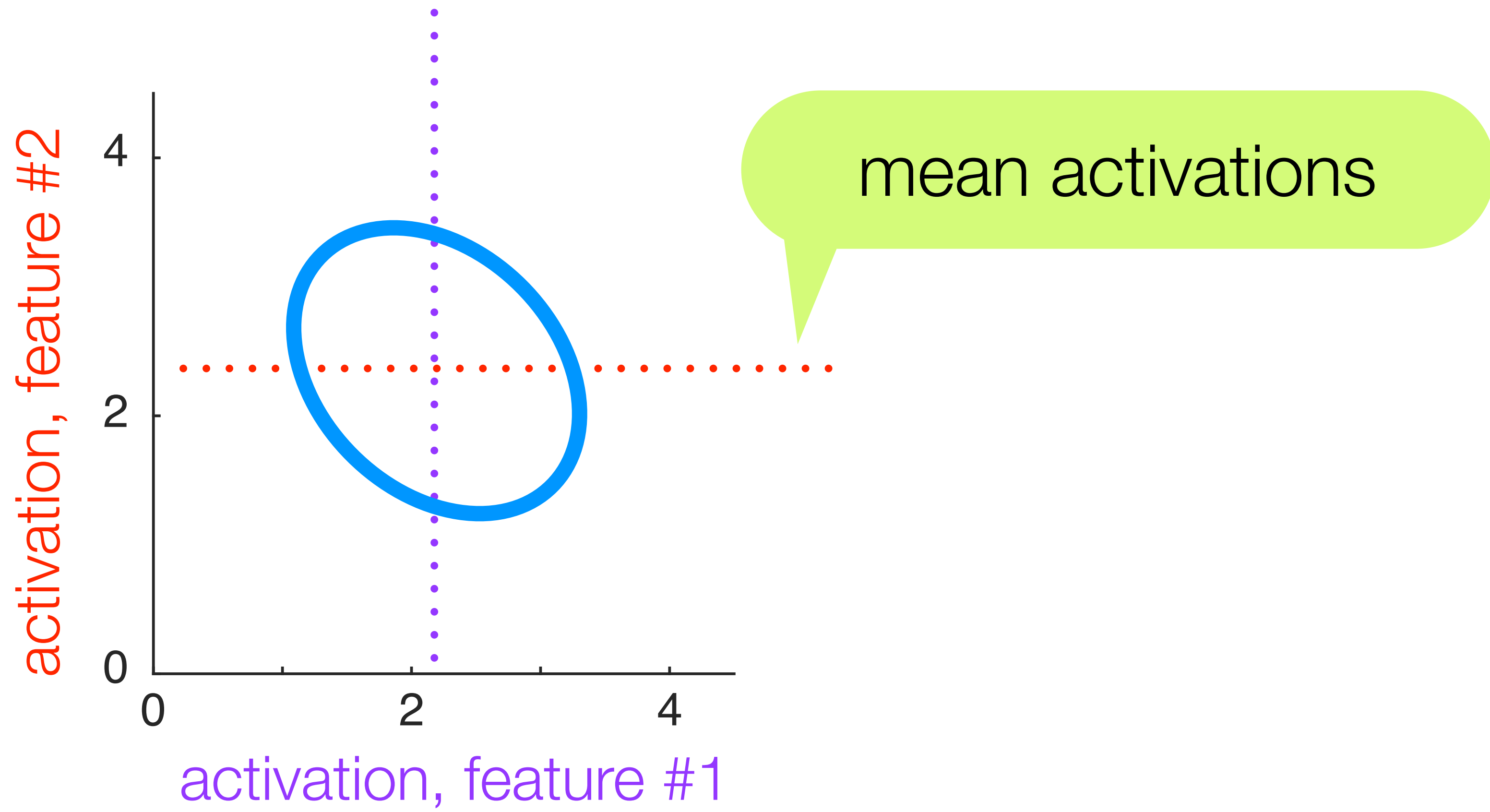


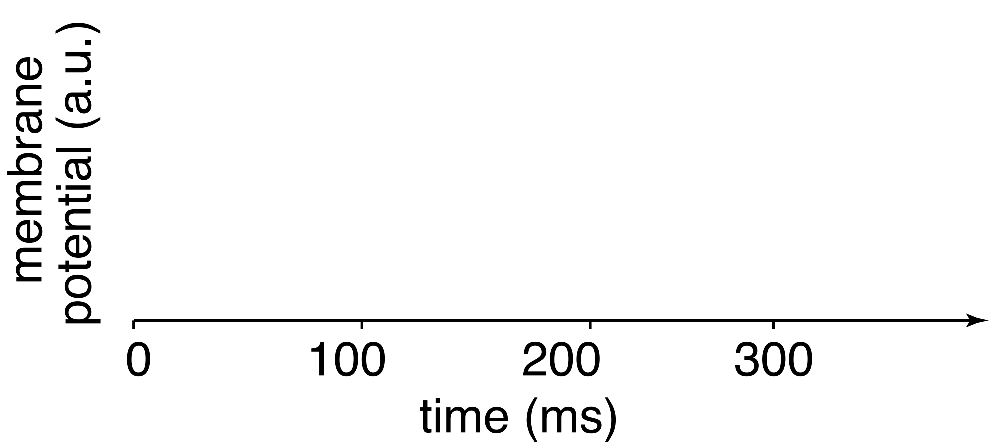




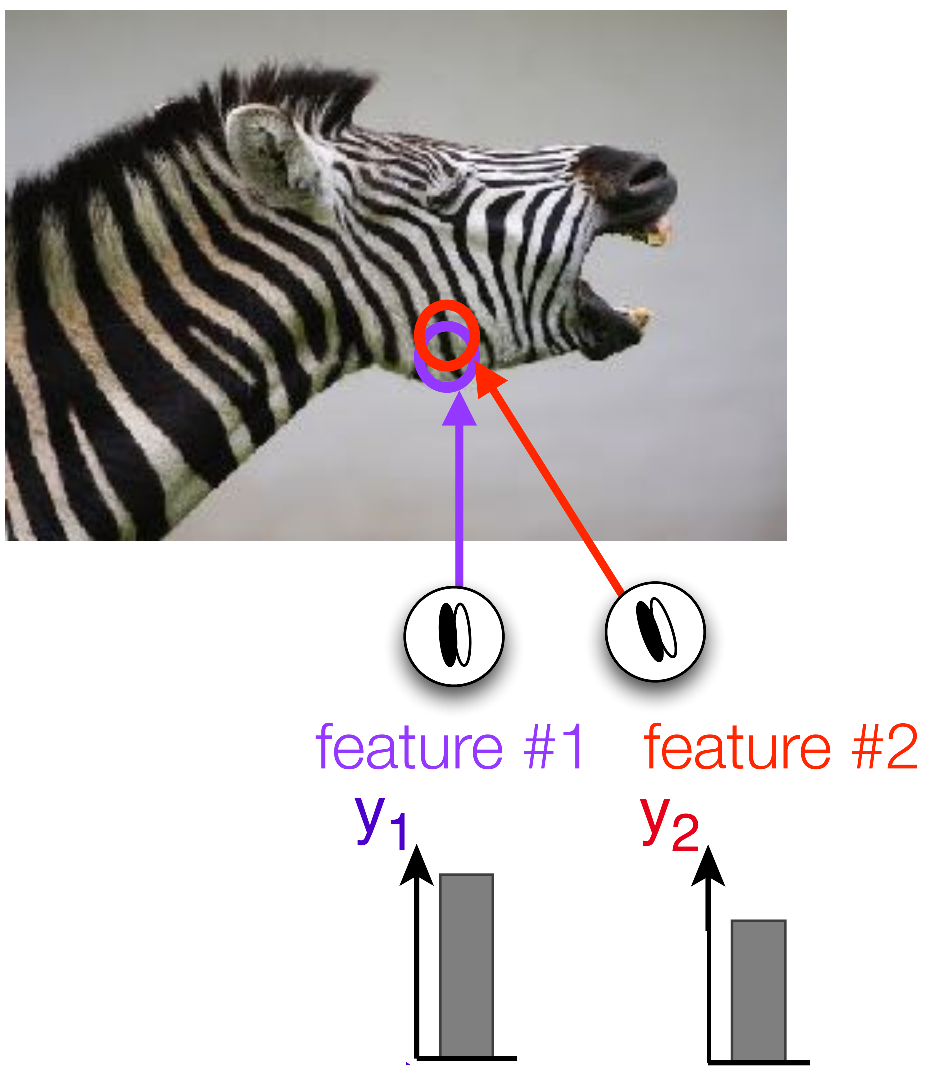
stochastic sampling

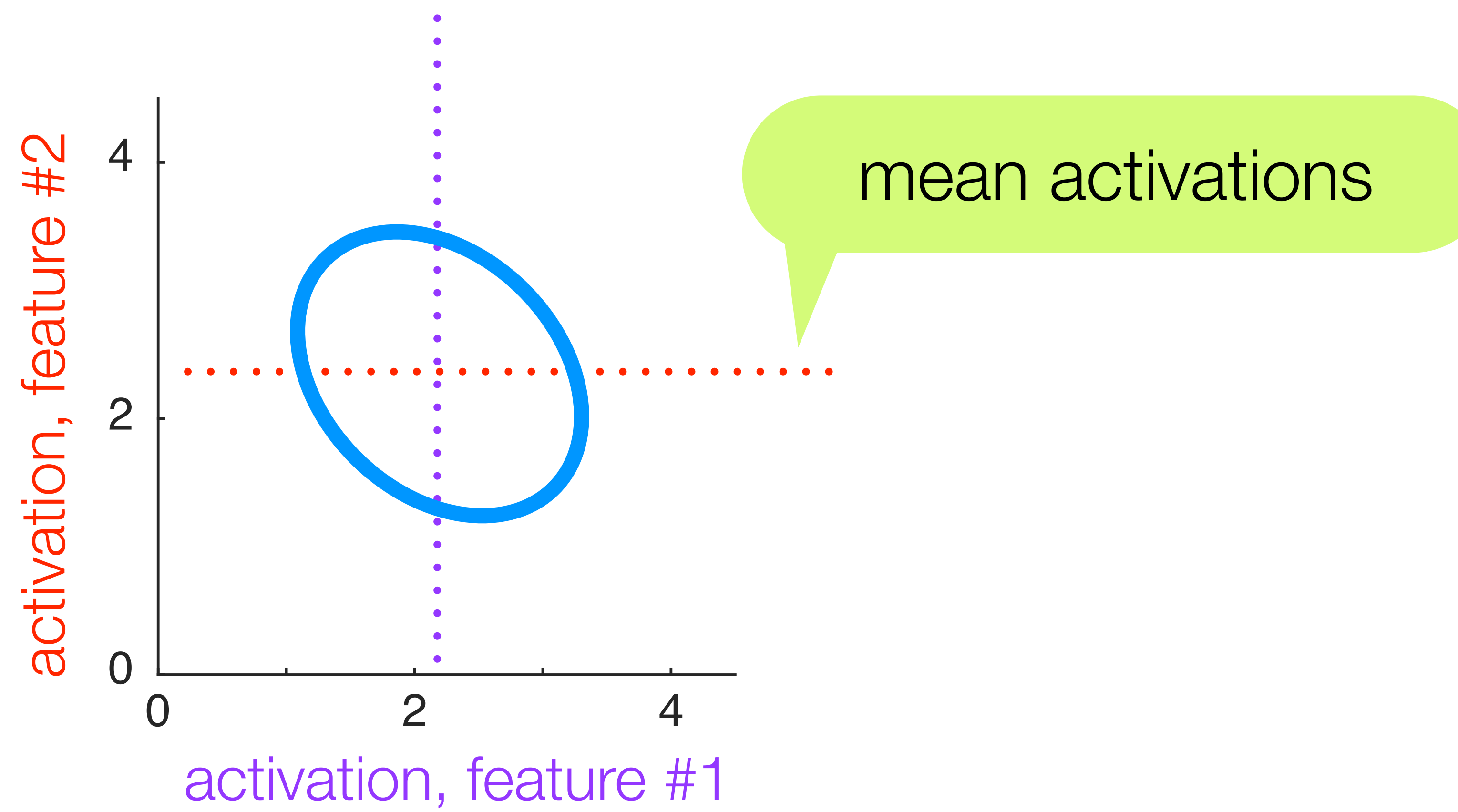


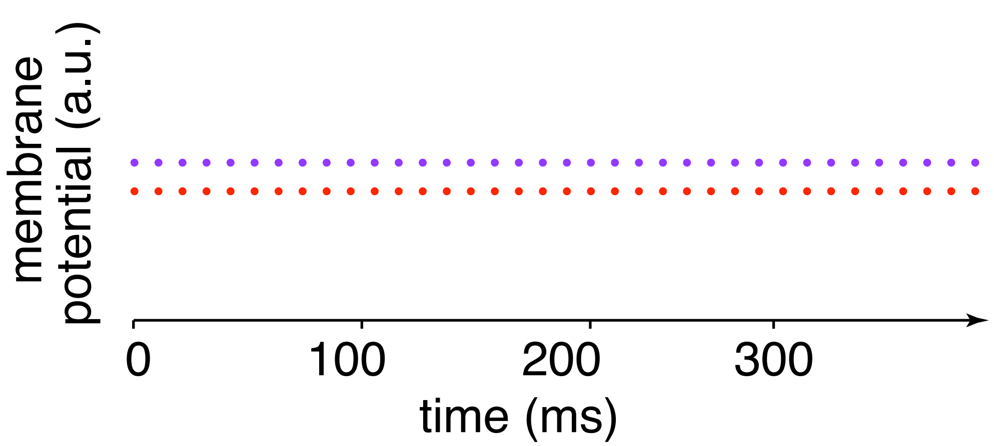




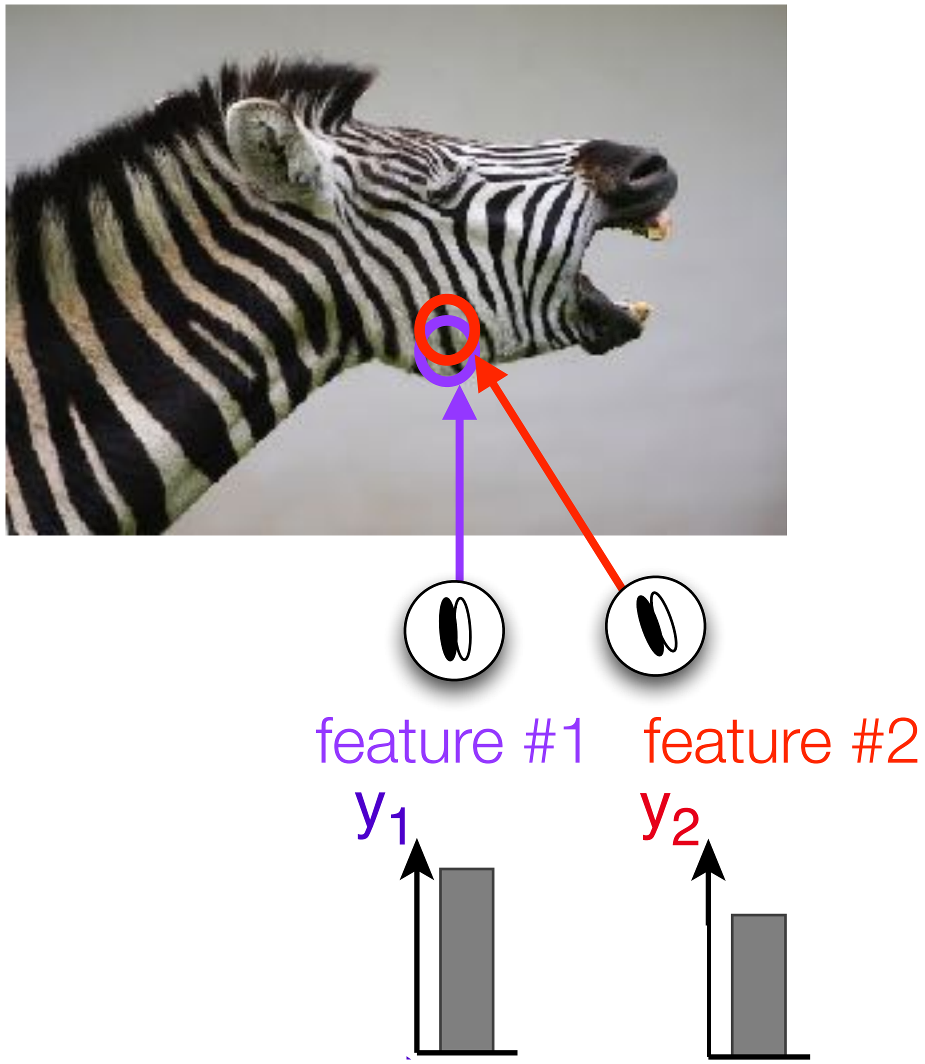
stochastic sampling

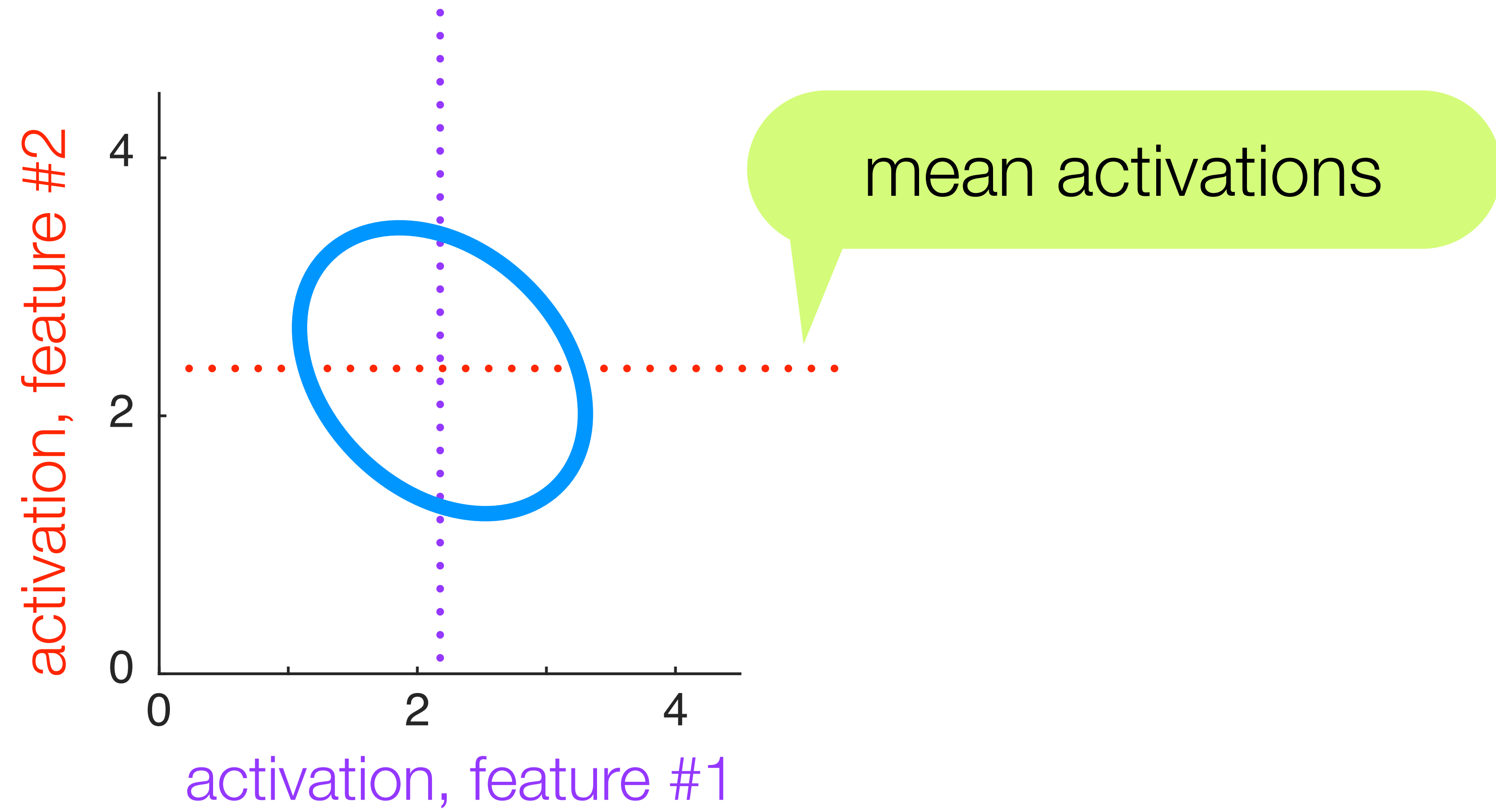


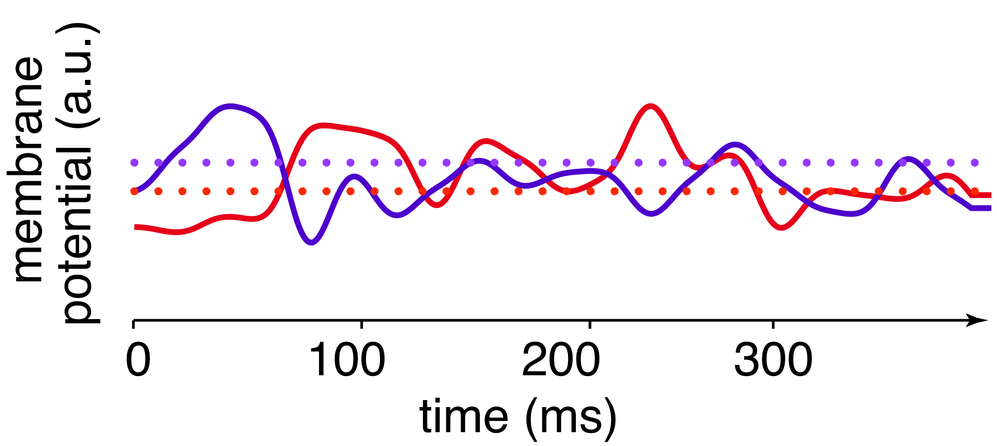




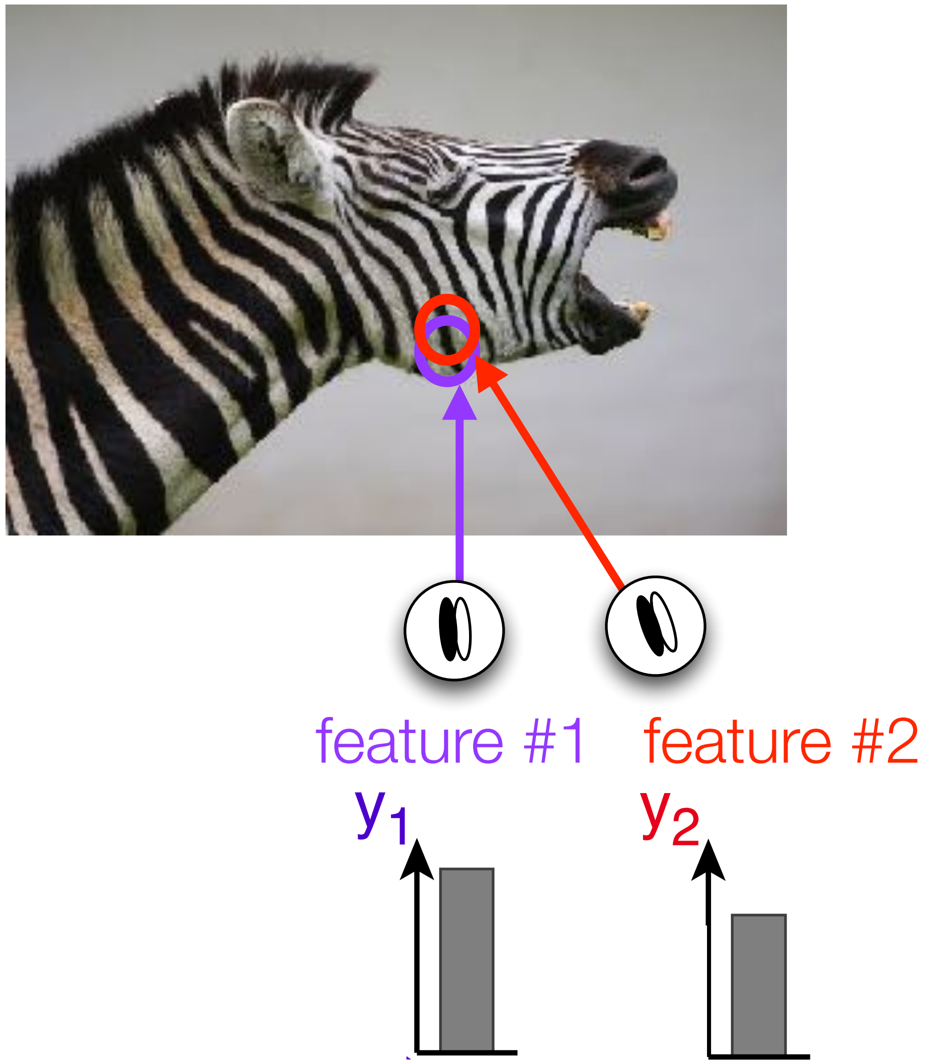
stochastic sampling

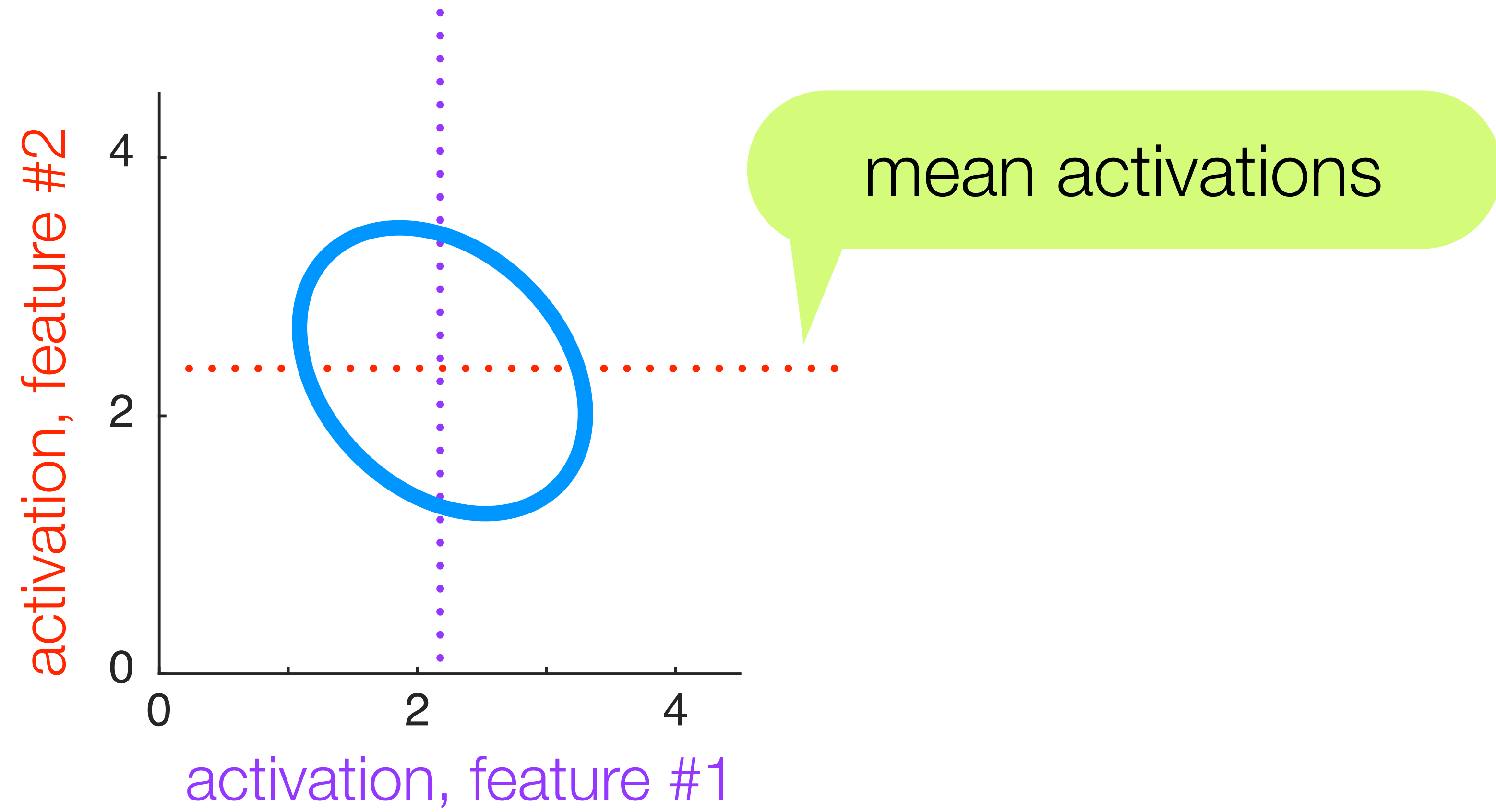


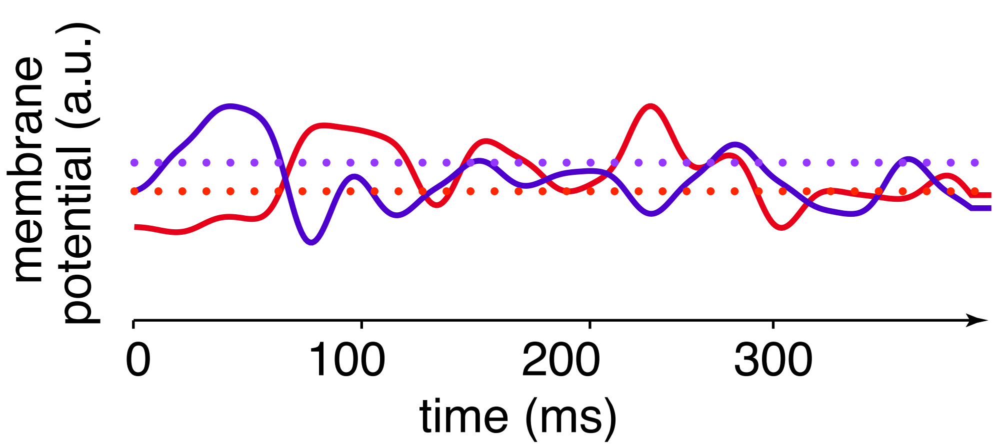


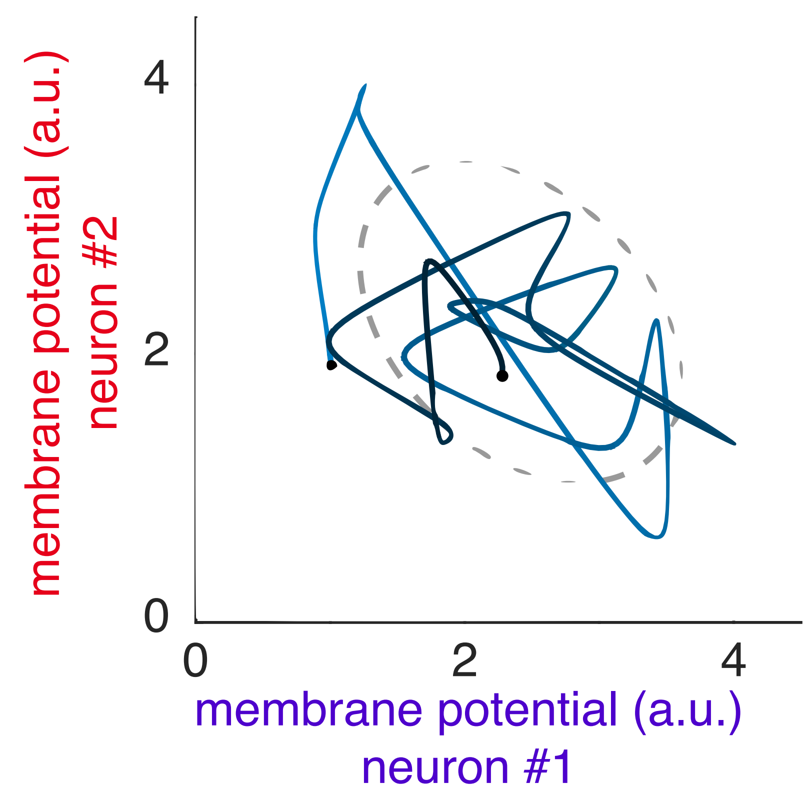


stochastic sampling

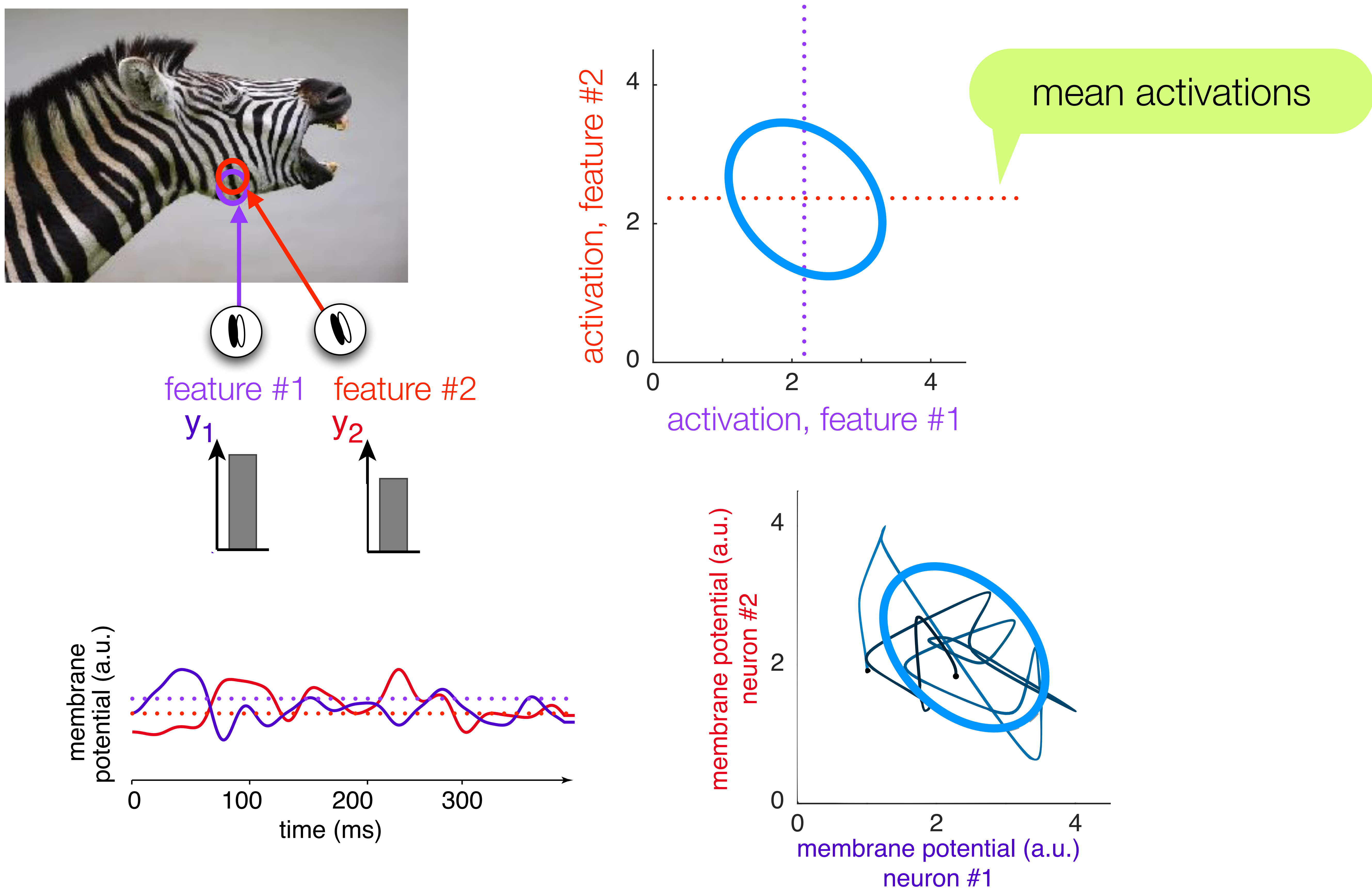




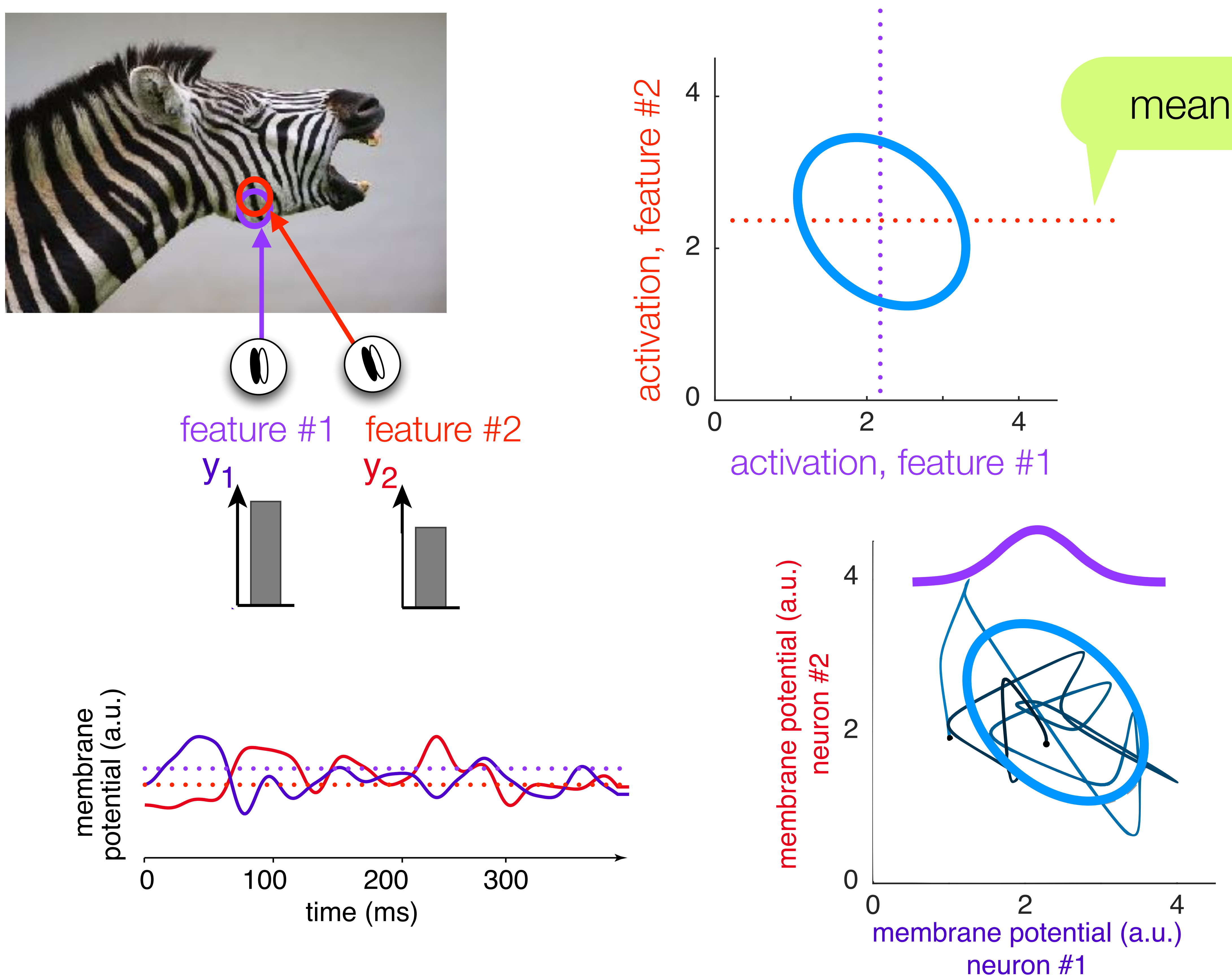


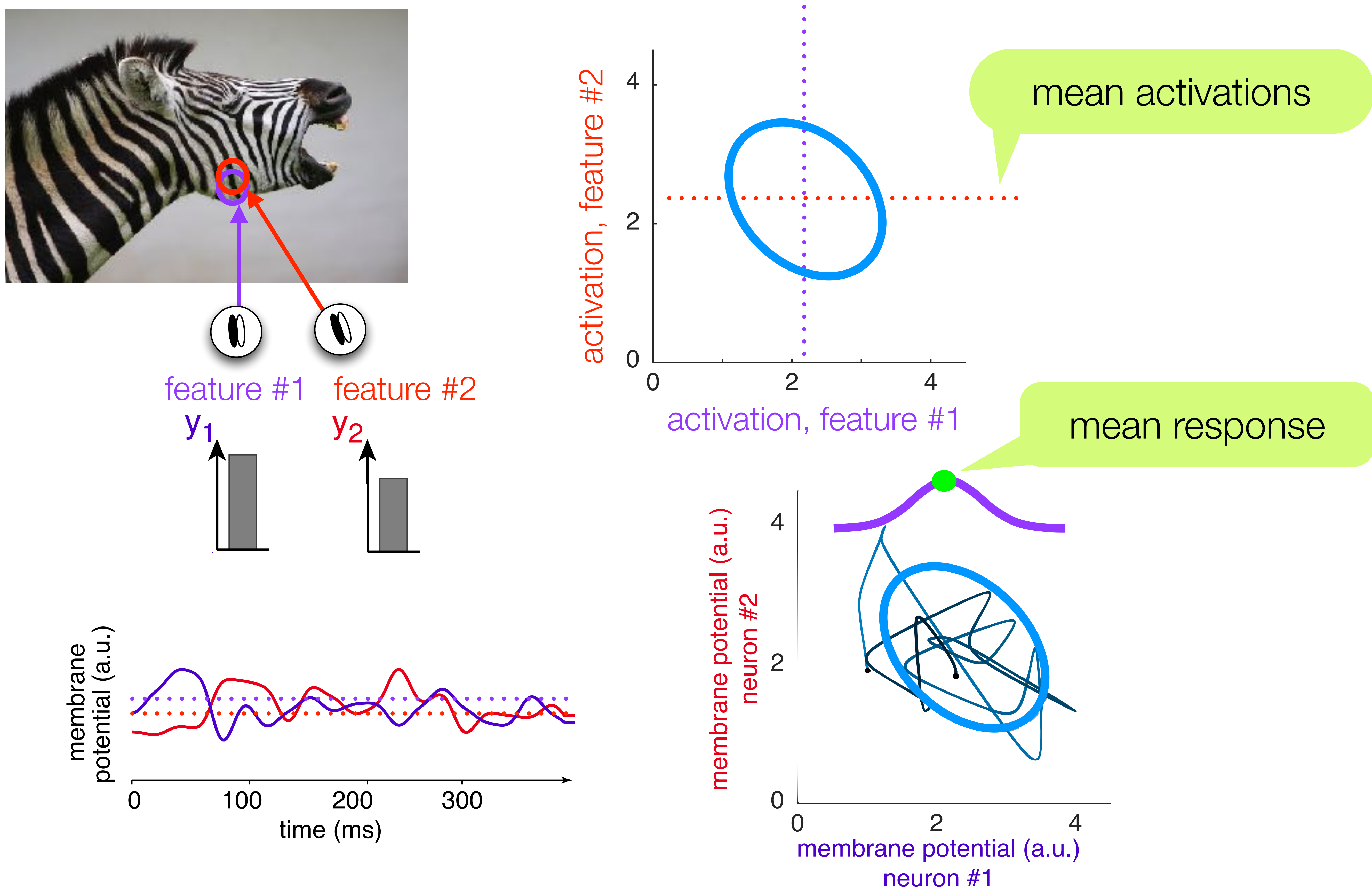


stochastic sampling



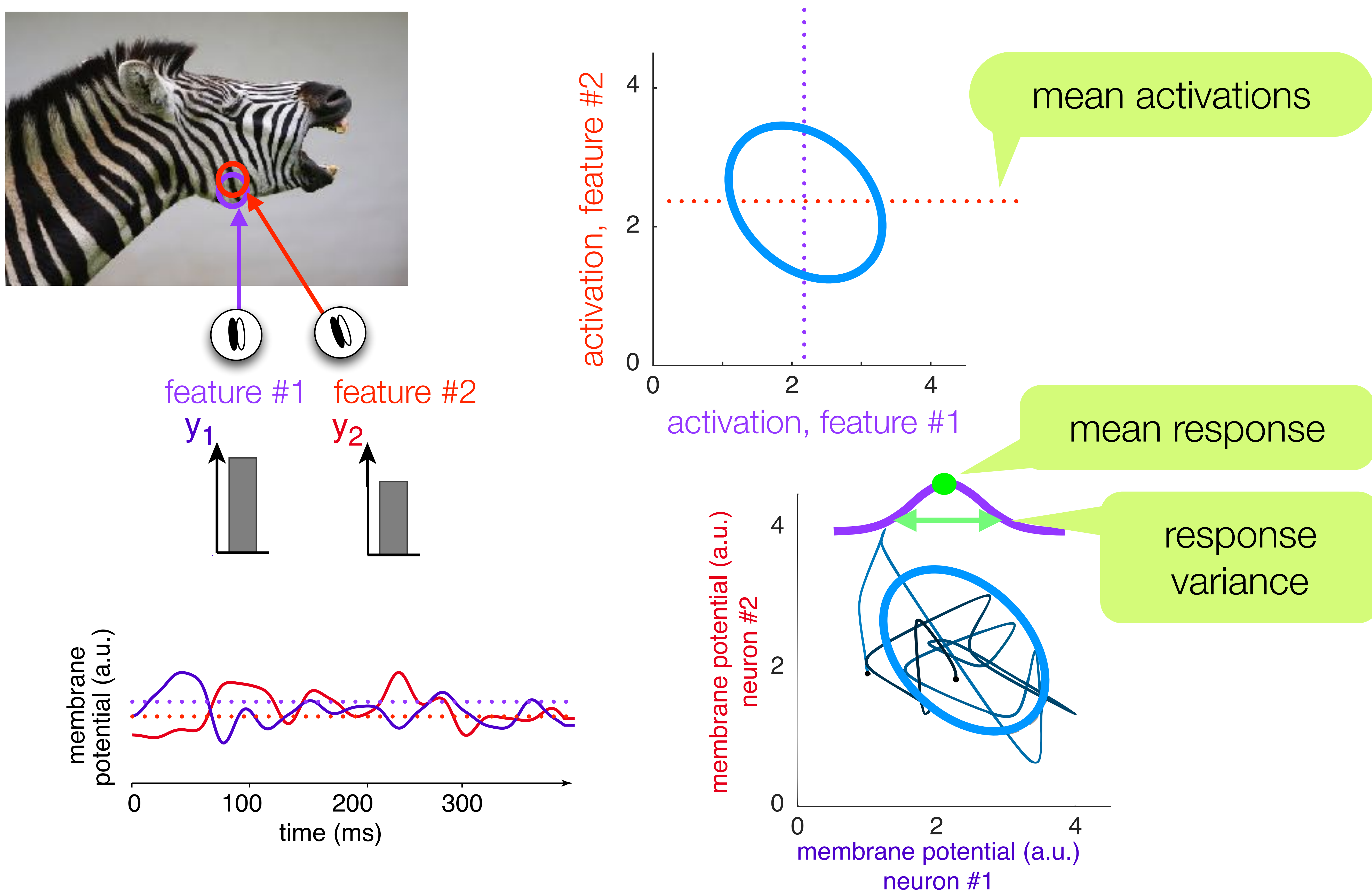
stochastic sampling

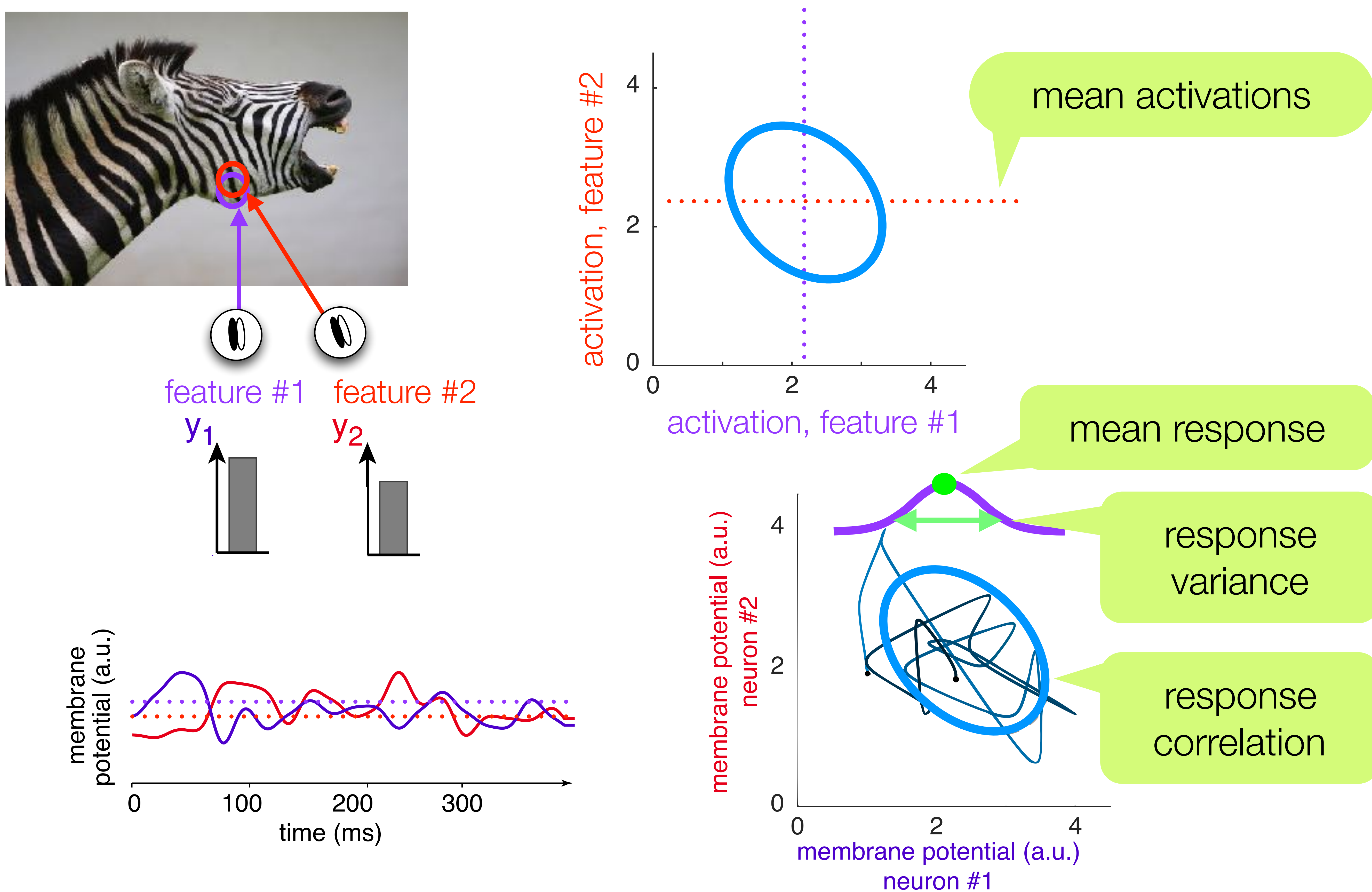




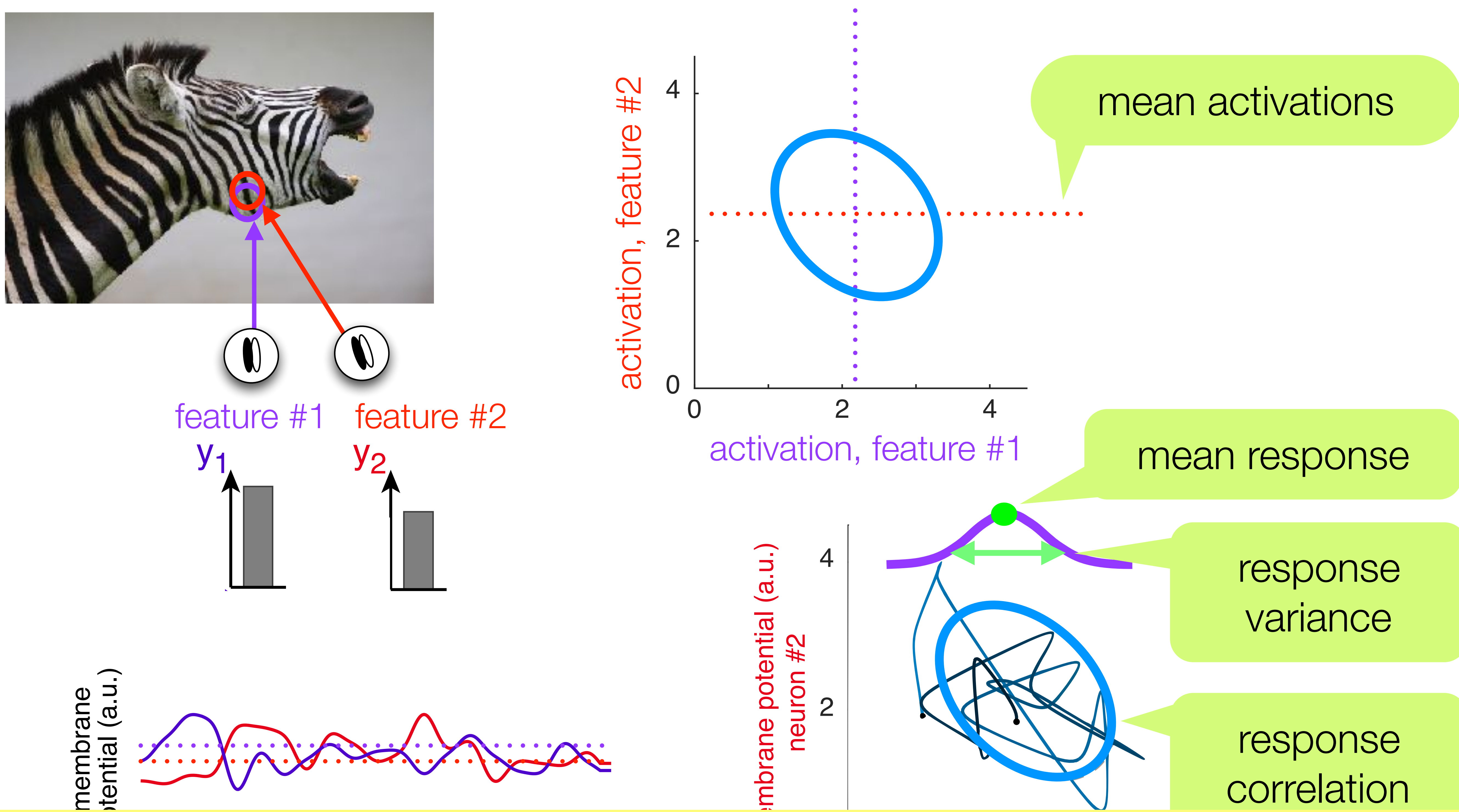
feature #1
feature #2
y_1
y_2
membrane potential (a.u.)
0
100
200
300
time (ms)
activation, feature #2
4
2
0
0
2
4
activation, feature #1
mean activations
mean response
membrane potential (a.u.) neuron #2
membrane potential (a.u.) neuron #1

stochastic sampling



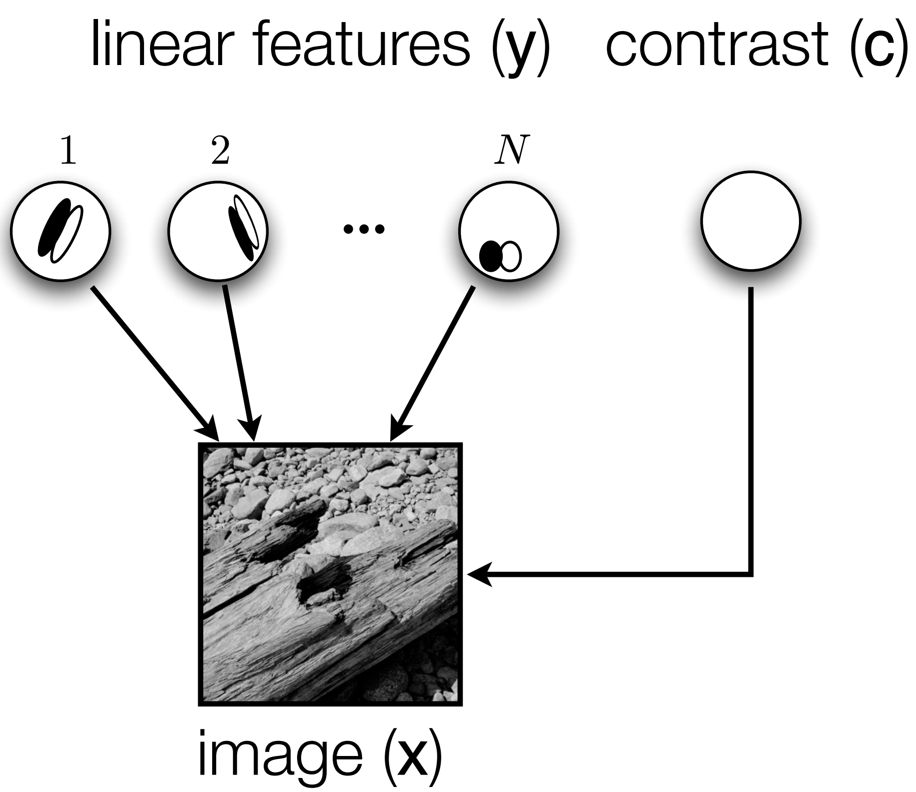

feature #1
feature #2
y_1
y_2
mean activations
activation, feature #2
activation, feature #1
0
2
4
mean response
response variance
response correlation
membrane potential (a.u.) neuron #2
membrane potential (a.u.) neuron #1
membrane potential (a.u.)
time (ms)
100
200
300

stochastic sampling

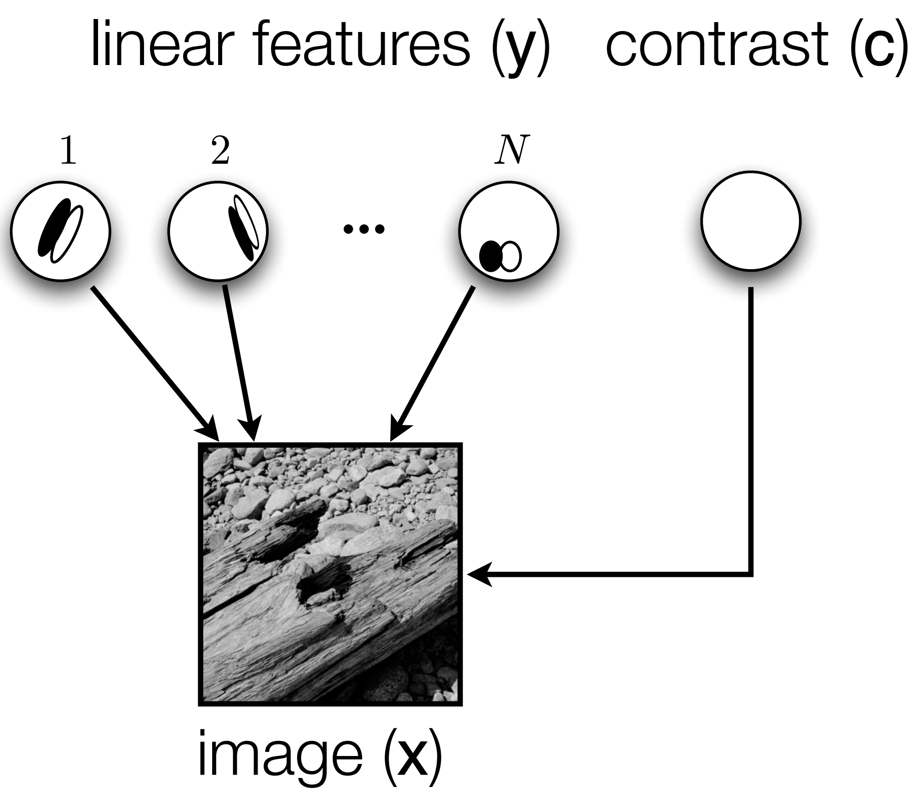


changes in inferences need to be reflected in the response statistics

Response statistics in a simple model of natural images

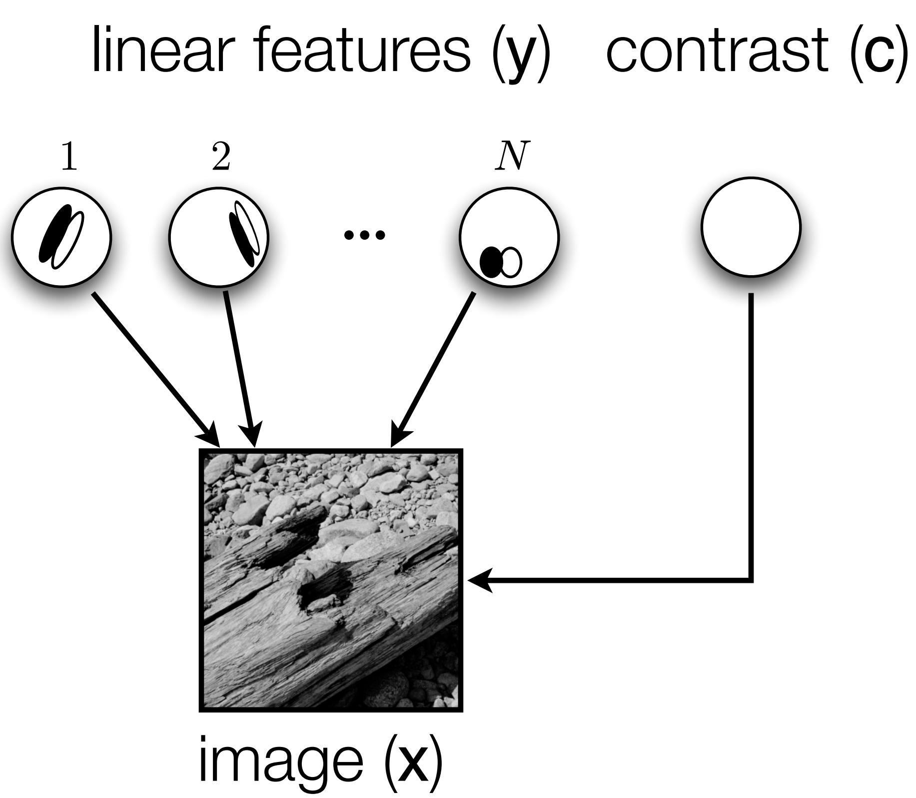


Response statistics in a simple model of natural images

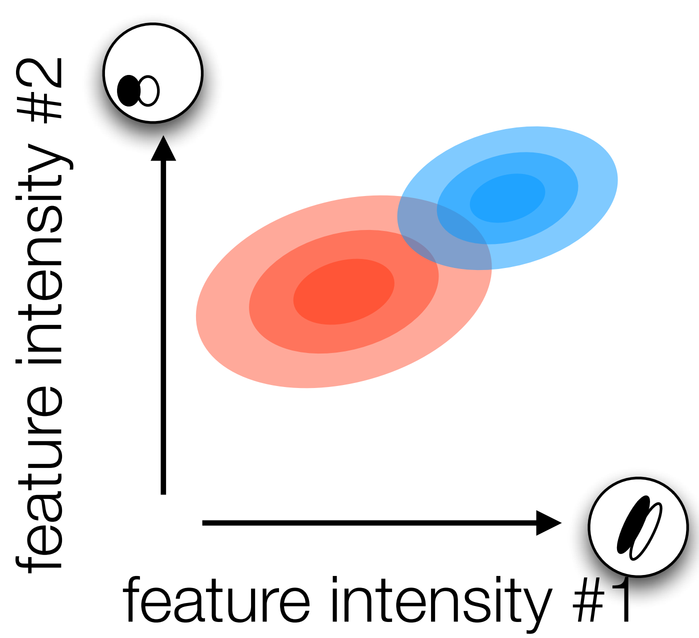


$$P(y \mid x) = \int P(y \mid x, c) P(c \mid x) \, dc$$

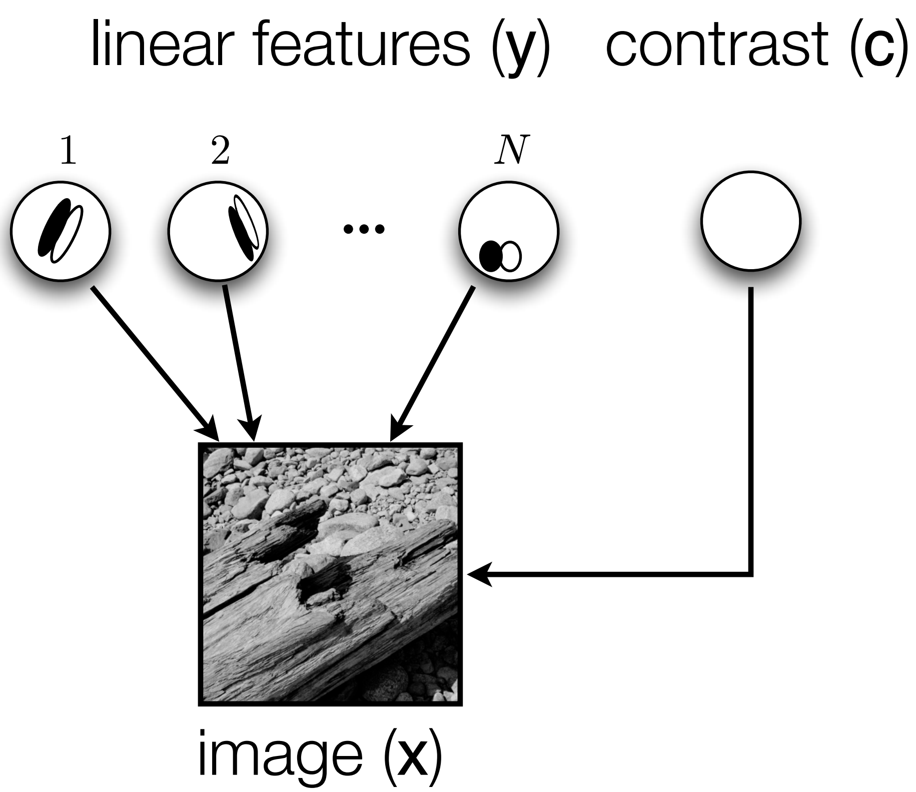
Response statistics in a simple model of natural images



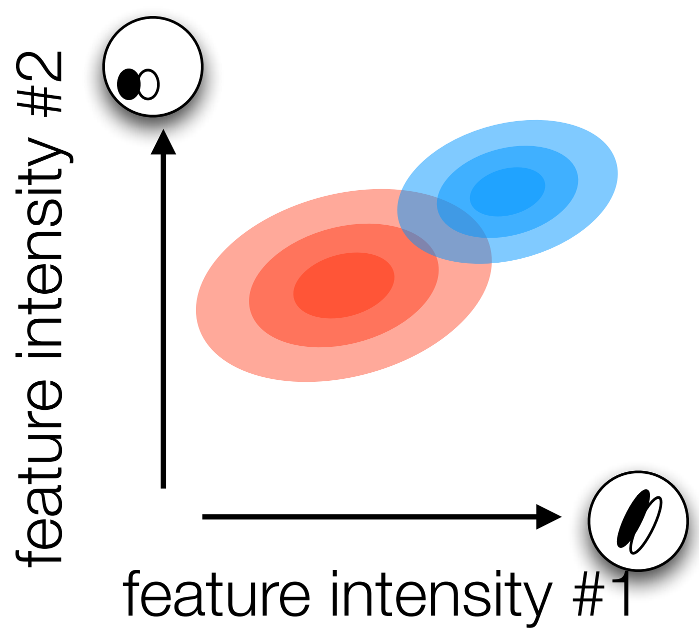
$$P(y \mid x) = \int P(y \mid x, c) P(c \mid x)\, dc$$

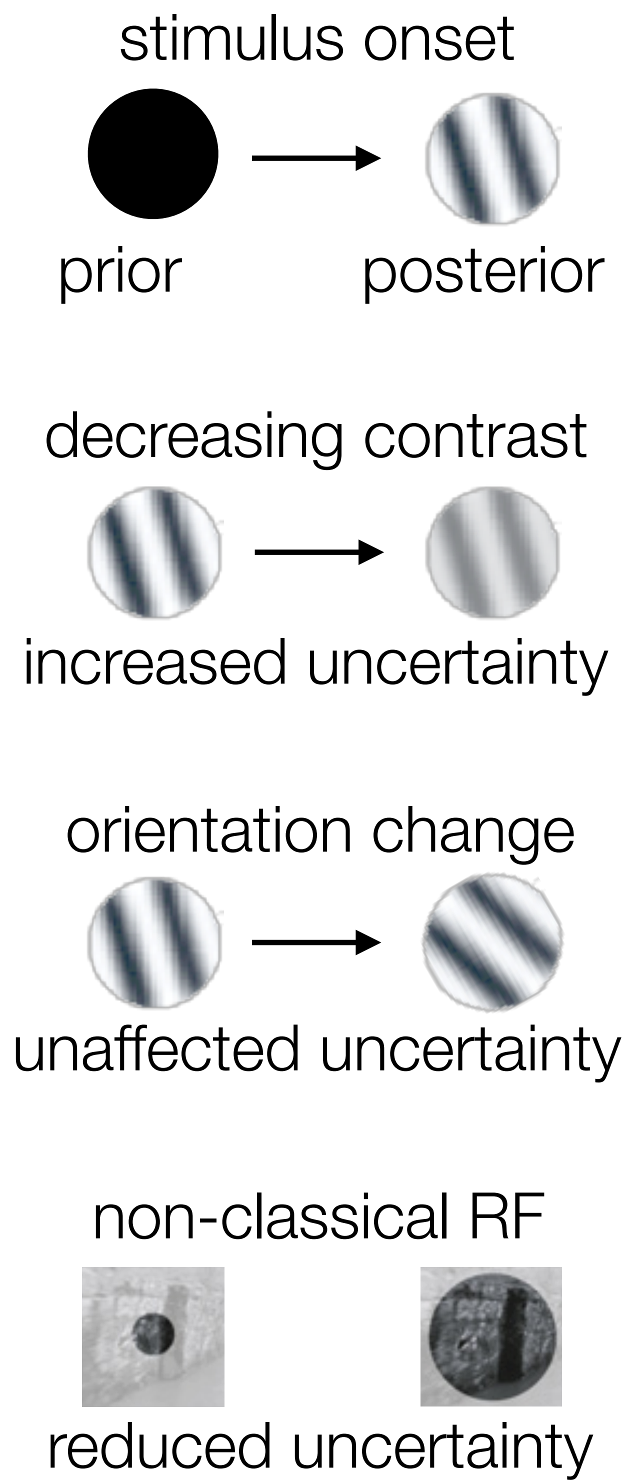


Response statistics in a simple model of natural images

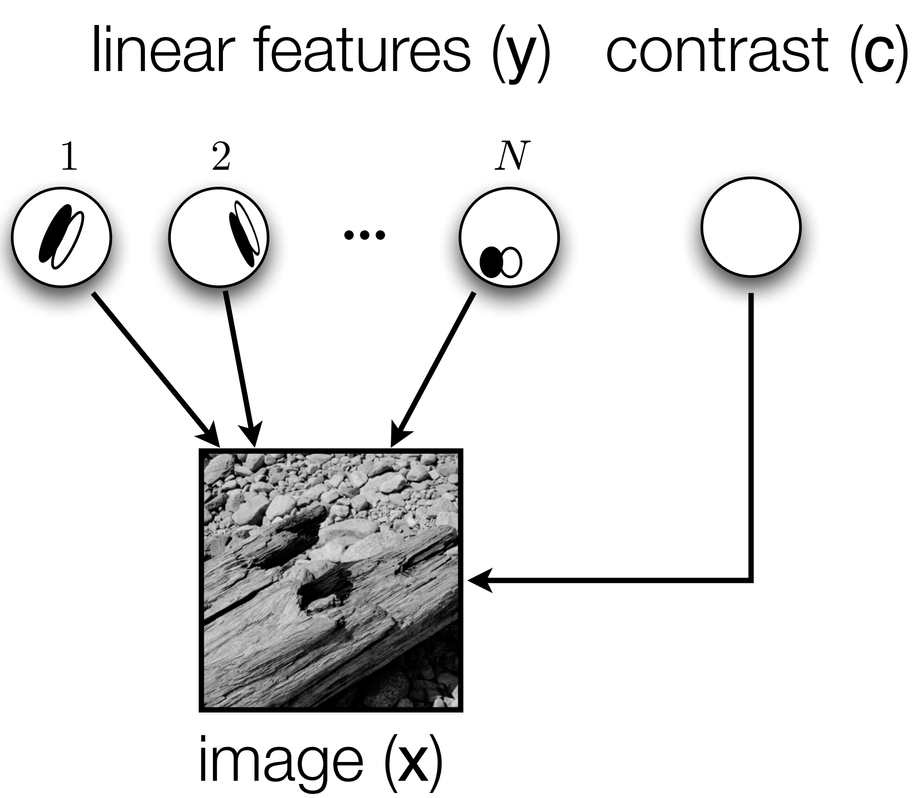


$$P(y \mid x) = \int P(y \mid x, c) P(c \mid x) \, dc$$

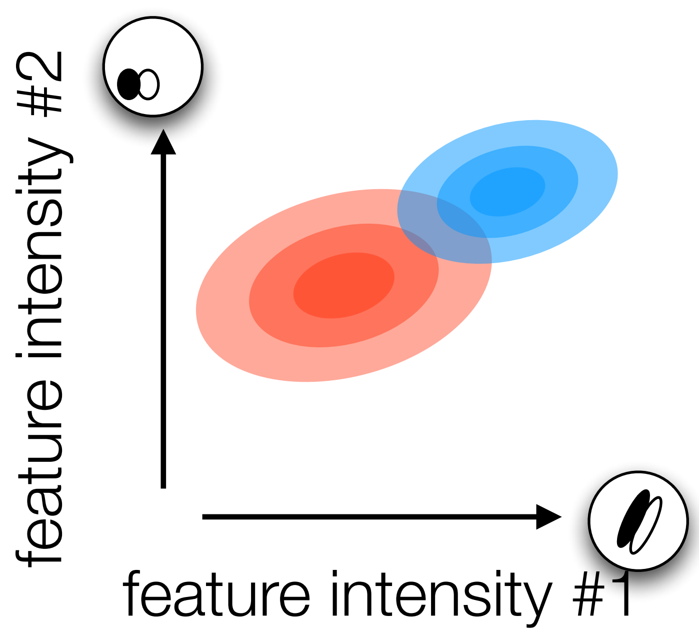


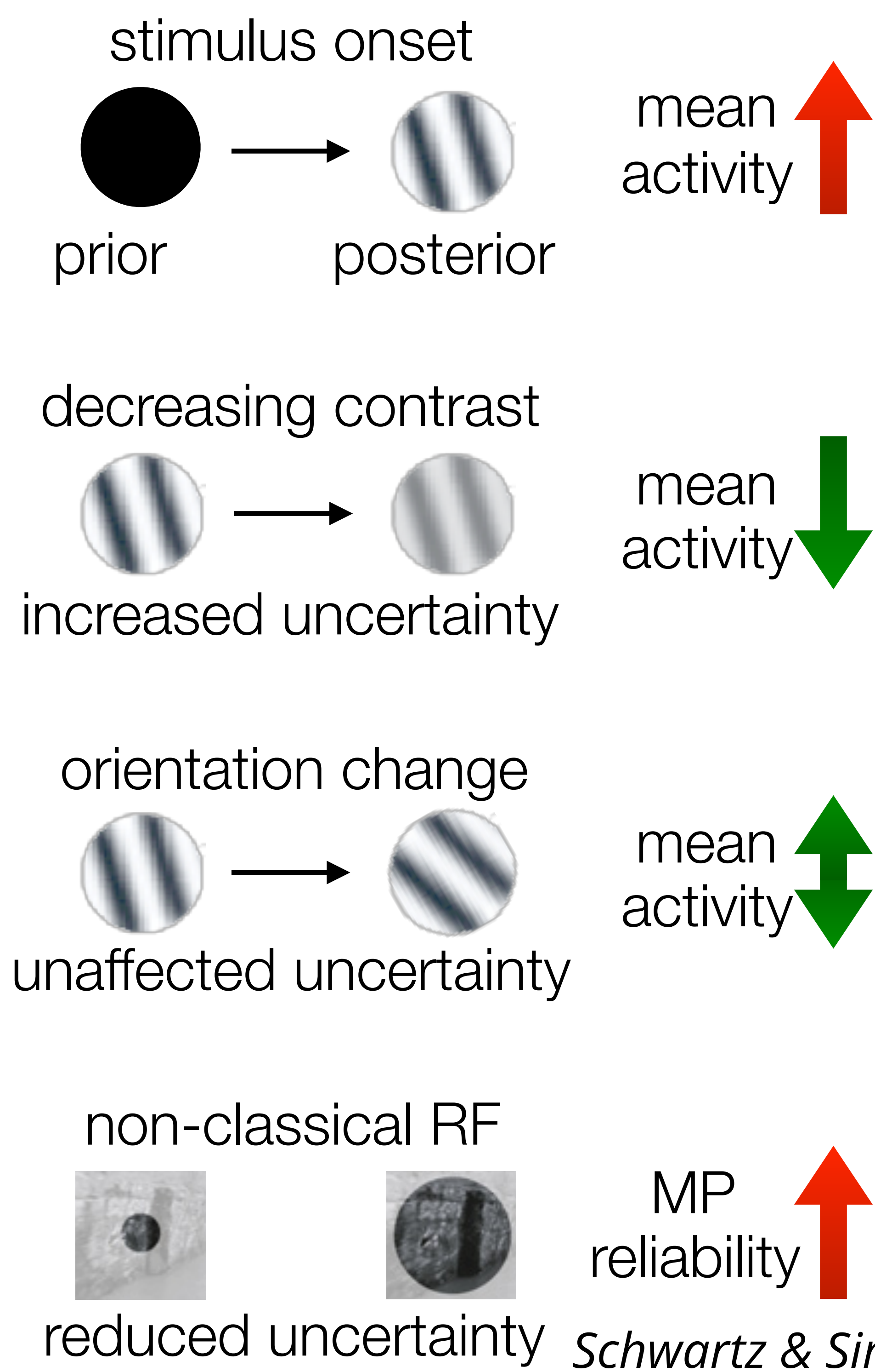


Response statistics in a simple model of natural images



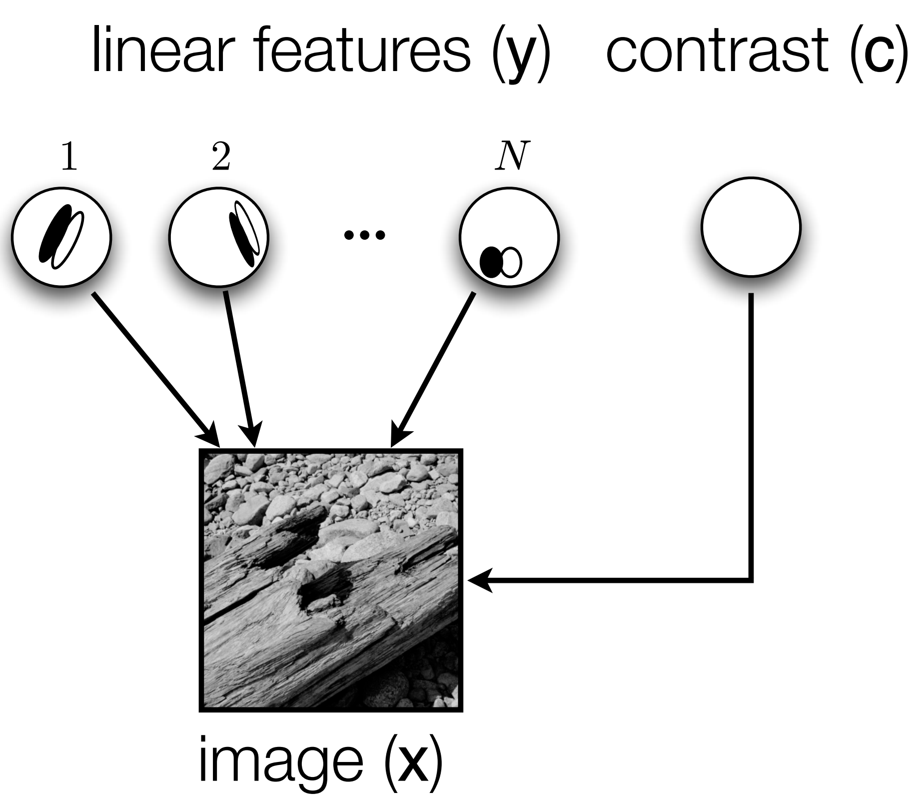
$$P(y \mid x) = \int P(y \mid x, c) P(c \mid x)\, dc$$



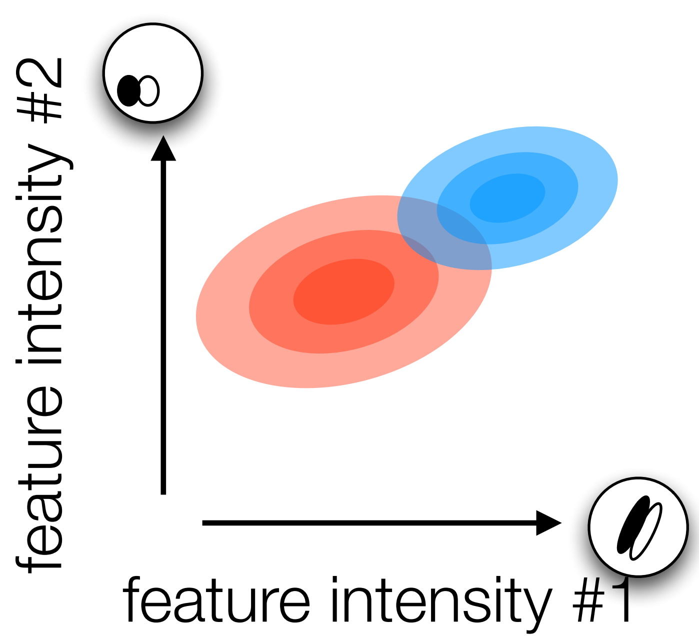


Schwartz & Simoncelli (2001) Nat Neuro

Response statistics in a simple model of natural images



$$P(y \mid x) = \int P(y \mid x, c) P(c \mid x)\, dc$$



stimulus onset

prior → posterior

mean activity

MP variance

Fano factor

decreasing contrast

increased uncertainty

mean activity

MP variance

Fano factor

orientation change

unaffected uncertainty

mean activity

MP variance

Fano factor

non-classical RF

reduced uncertainty

MP reliability

SC sparseness

response decorrelation

Schwartz & Simoncelli (2001) Nat Neuro

Orbán et al (2016) Neuron

Churchland et al (2010) Nat Neurosci; Ecker et al (2010) Science; Finn et al (2007) Neuron; Vinje & Gallant (2001) Science; Haider et al (2007) Neuron

Response statistics in a simple model of natural images

linear features (**y**) contrast (**c**)

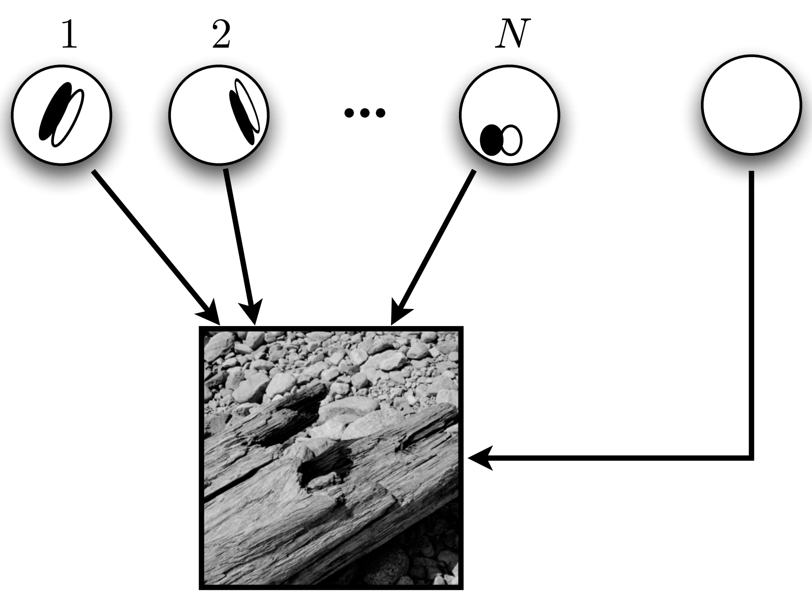
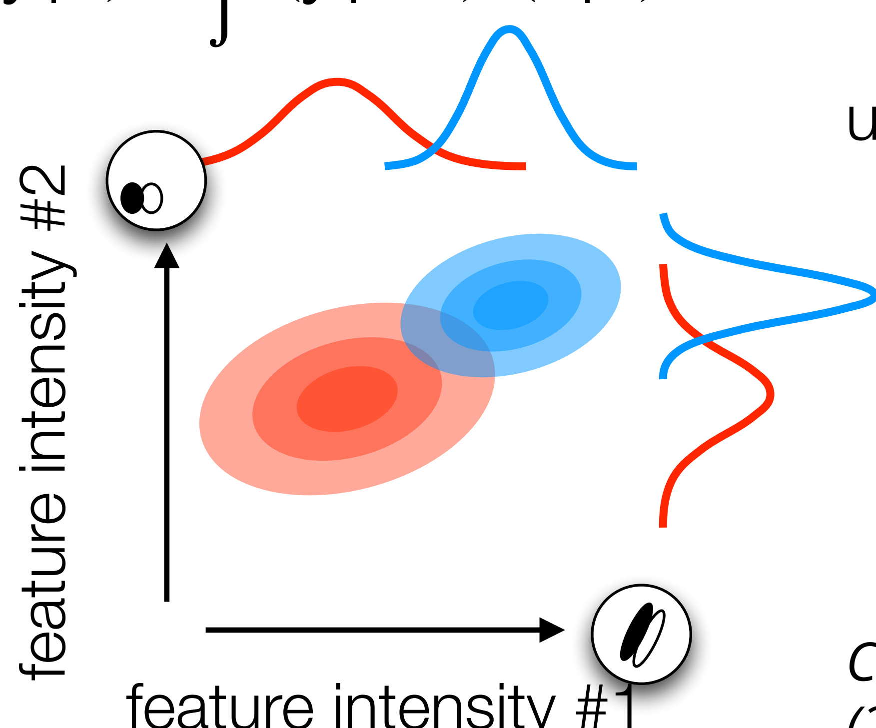
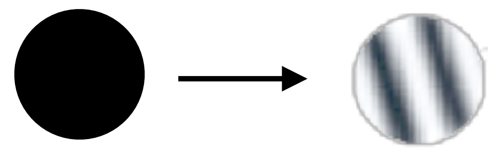


image (**x**)

$$P(y \mid x) = \int P(y \mid x, c) P(c \mid x)\, dc$$



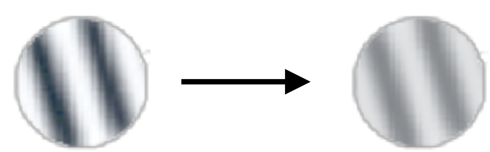
stimulus onset



prior posterior

mean activity | MP variance | Fano factor

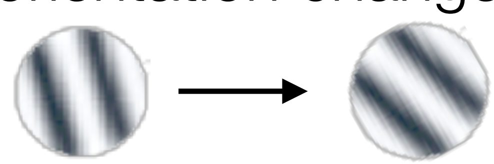
decreasing contrast



increased uncertainty

mean activity | MP variance | Fano factor

orientation change



unaffected uncertainty

mean activity | MP variance | Fano factor

non-classical RF

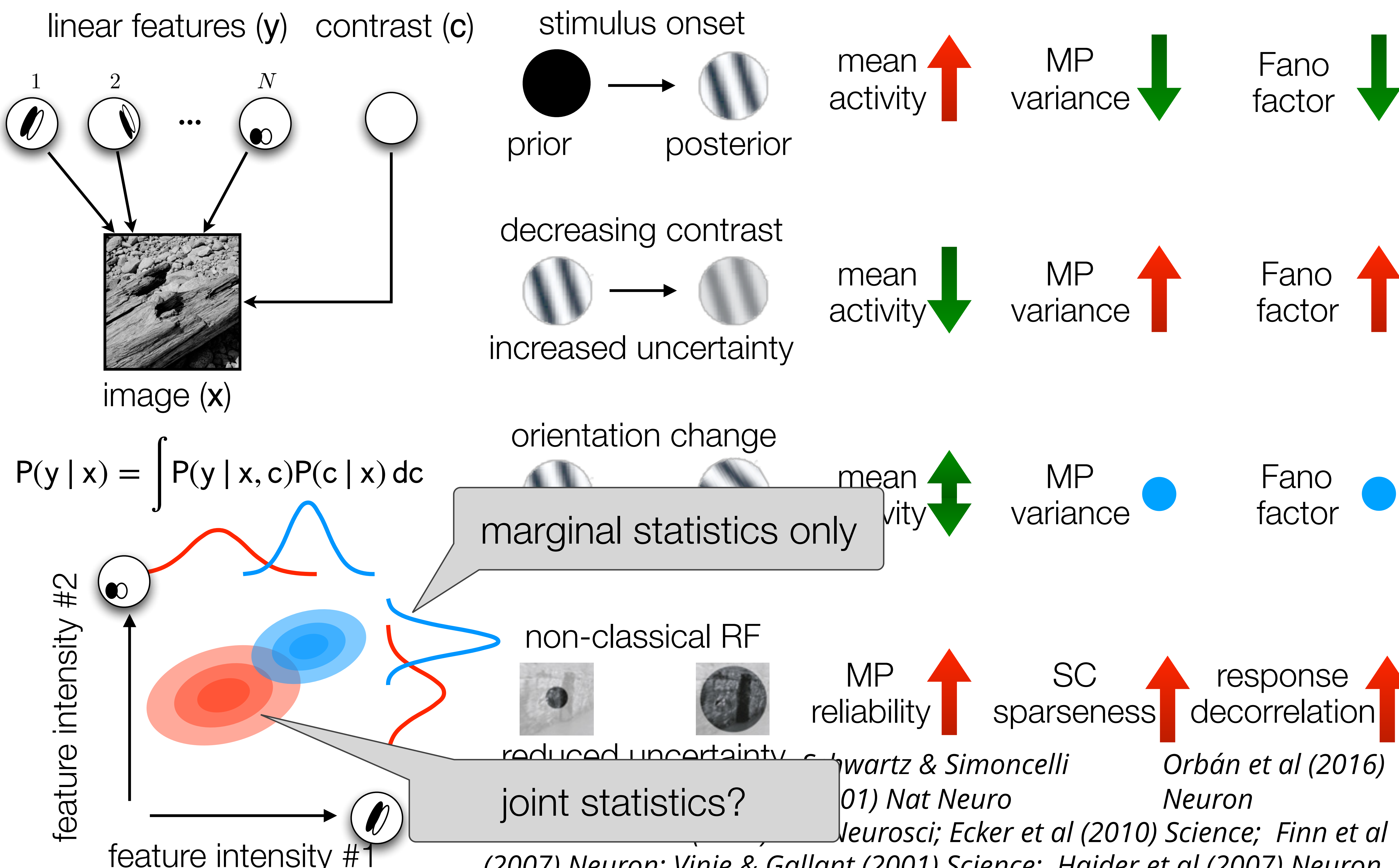
reduced uncertainty

MP reliability | SC sparseness | response decorrelation

Schwartz & Simoncelli (2001) Nat Neuro *Orbán et al (2016) Neuron*

Churchland et al (2010) Nat Neurosci; Ecker et al (2010) Science; Finn et al (2007) Neuron; Vinje & Gallant (2001) Science; Haider et al (2007) Neuron

Response statistics in a simple model of natural images



...wartz & Simoncelli ...01) Nat Neuro

Orbán et al (2016) Neuron

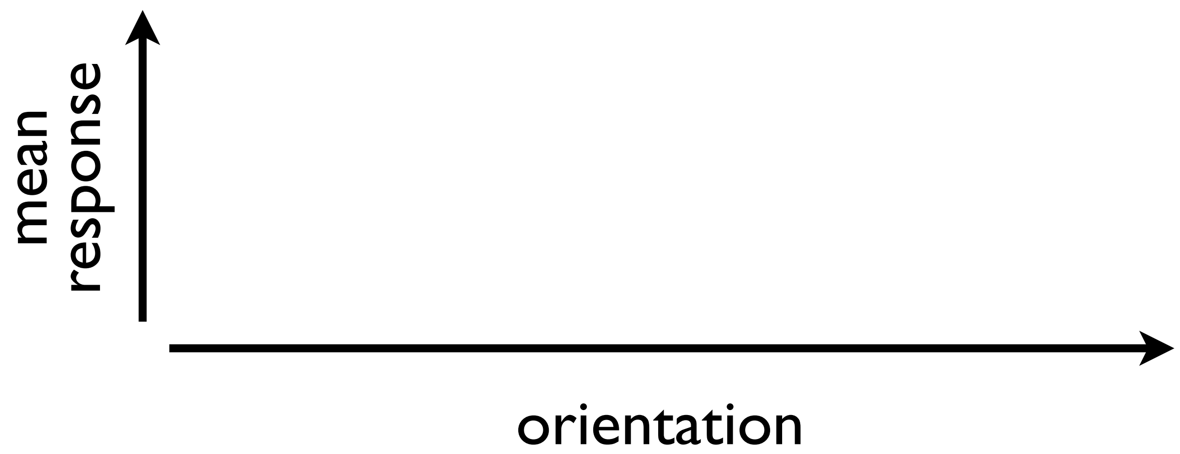
...Neurosci; Ecker et al (2010) Science; Finn et al (2007) Neuron; Vinje & Gallant (2001) Science; Haider et al (2007) Neuron

Bayes inferencia neuronhálózatokkal: PPC

V1 orientáció-szelektív neuronok

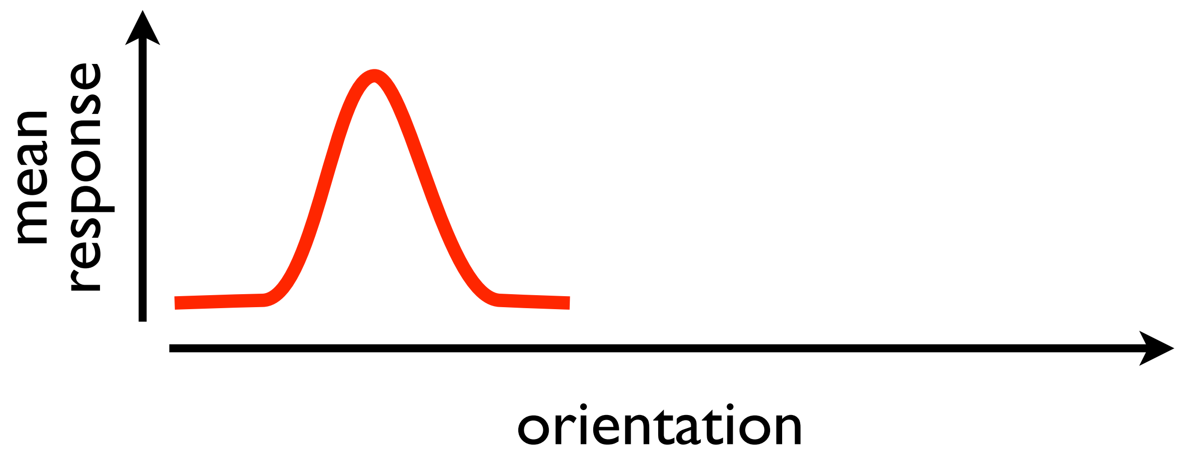
Bayes inferencia neuronhálózatokkal: PPC

V1 orientáció-szelektív neuronok



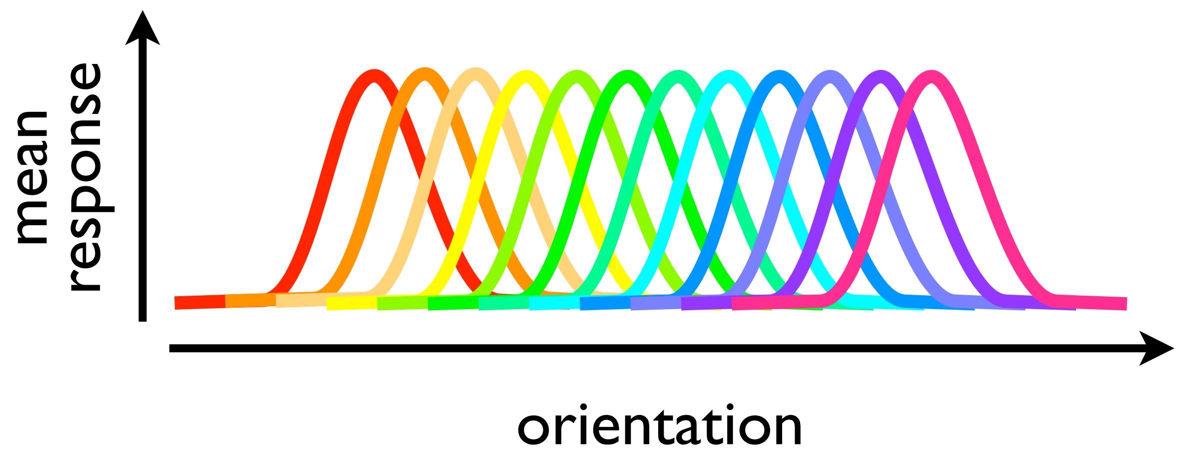
Bayes inferencia neuronhálózatokkal: PPC

V1 orientáció-szelektív neuronok



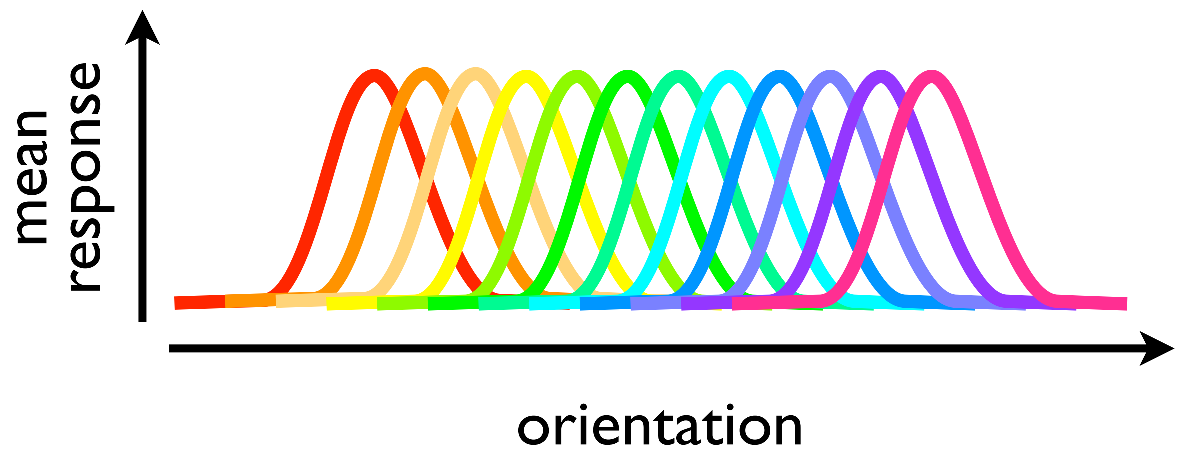
Bayes inferencia neuronhálózatokkal: PPC

V1 orientáció-szelektív neuronok



Bayes inferencia neuronhálózatokkal: PPC

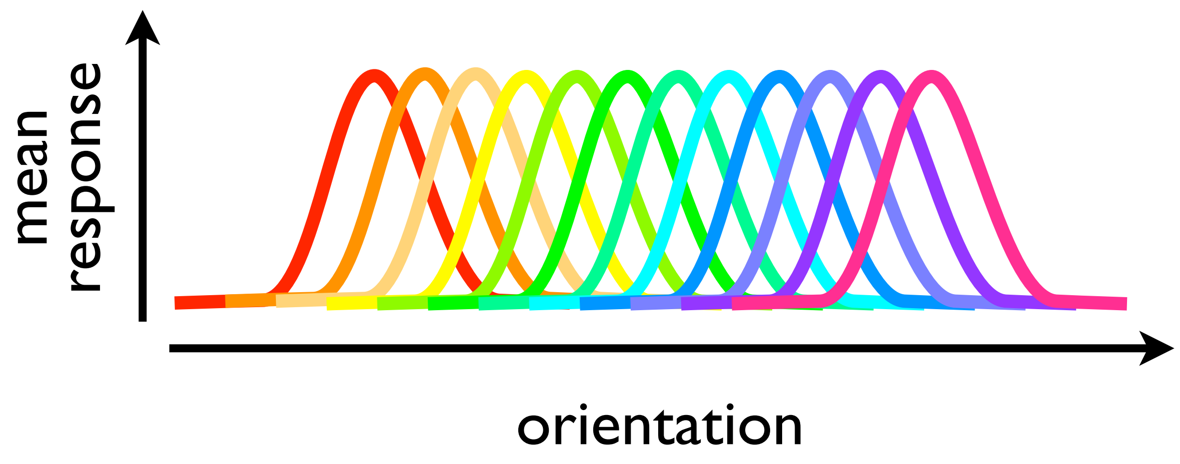
V1 orientáció-szelektív neuronok



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

Bayes inferencia neuronhálózatokkal: PPC

V1 orientáció-szelektív neuronok

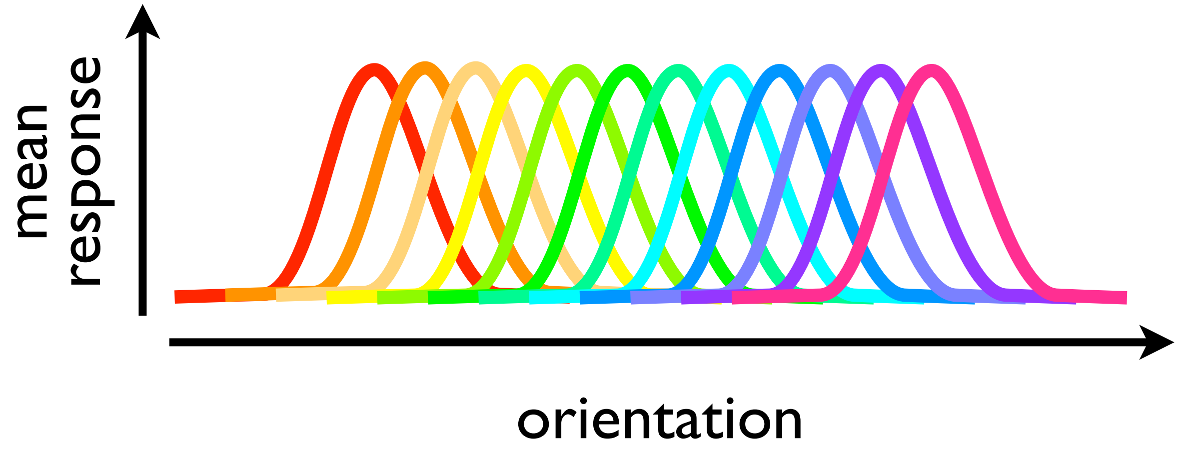


a neuronok azonban zajosak: az átlag körül az átlaggal arányos variabilitás van jelen

cél: orientáció becslése

Bayes inferencia neuronhálózatokkal: PPC

V1 orientáció-szelektív neuronok



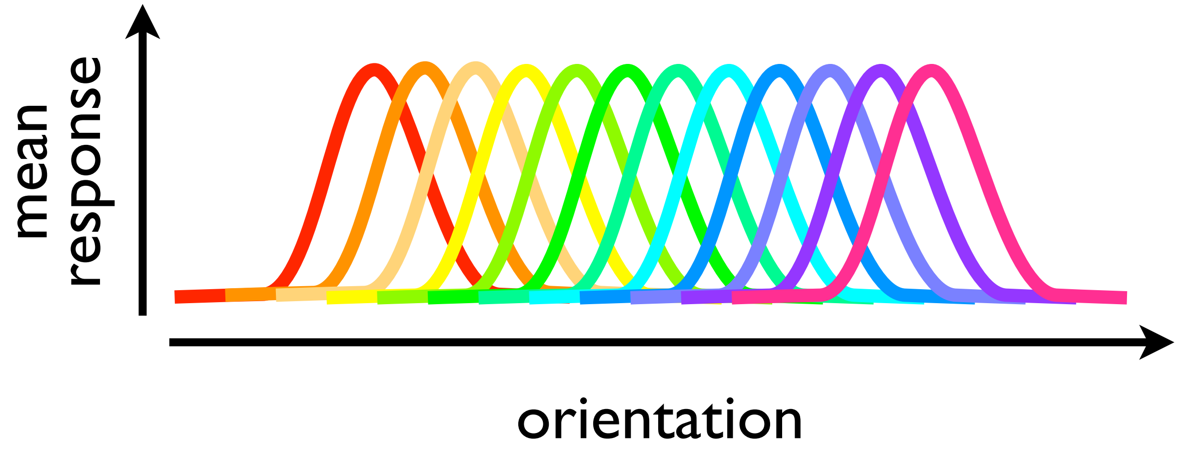
a neuronok azonban zajosak: az átlag körül az átlaggal arányos variabilitás van jelen

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

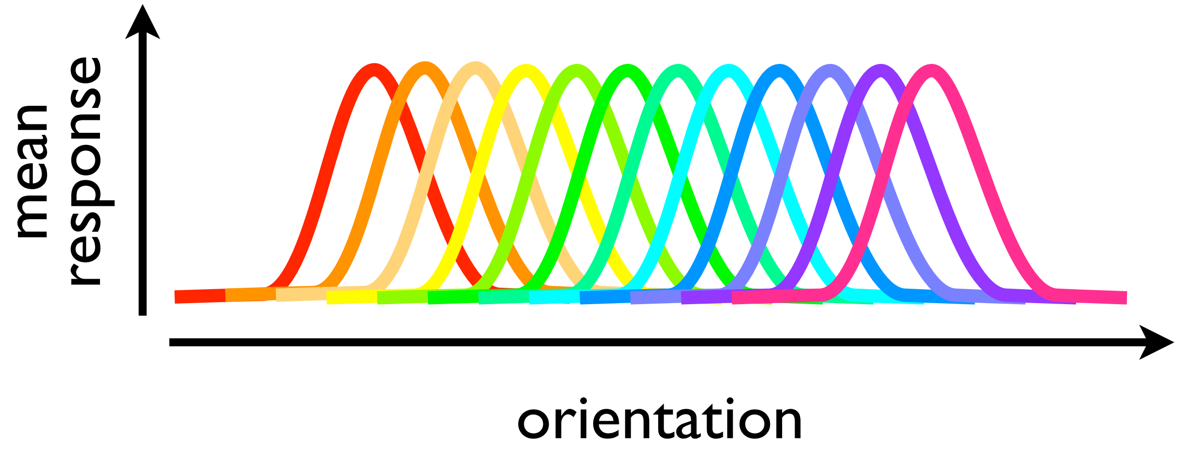
cél: orientáció becslése

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

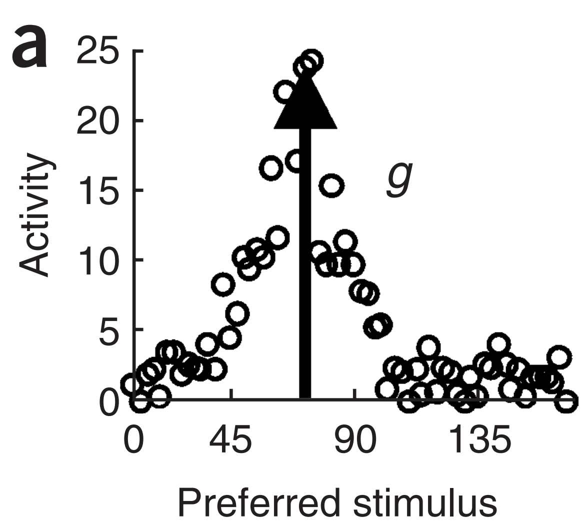
- Neurális zaj varianciája arányos az átlagos aktivitással: Poisson zaj
- Likelihood alakja:

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

Probabilistic Population Codes

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

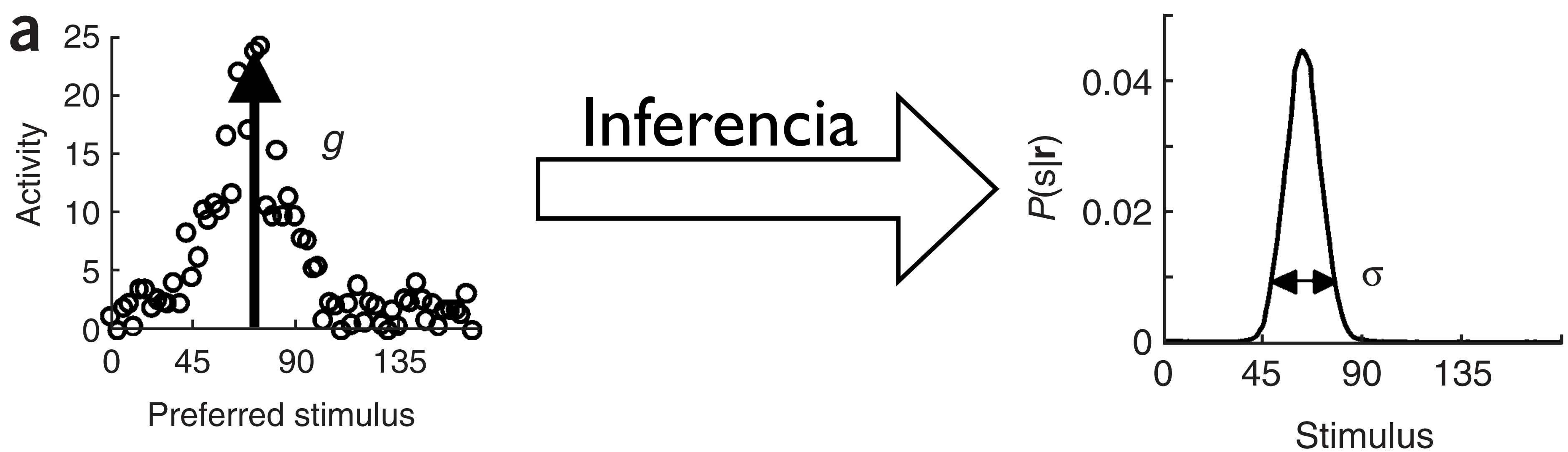
$$P(\mathbf{r} \mid s) = \prod_i \frac{e^{-f_i(s)} f_i(s)^{r_i}}{r_i!}$$



Probabilistic Population Codes

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

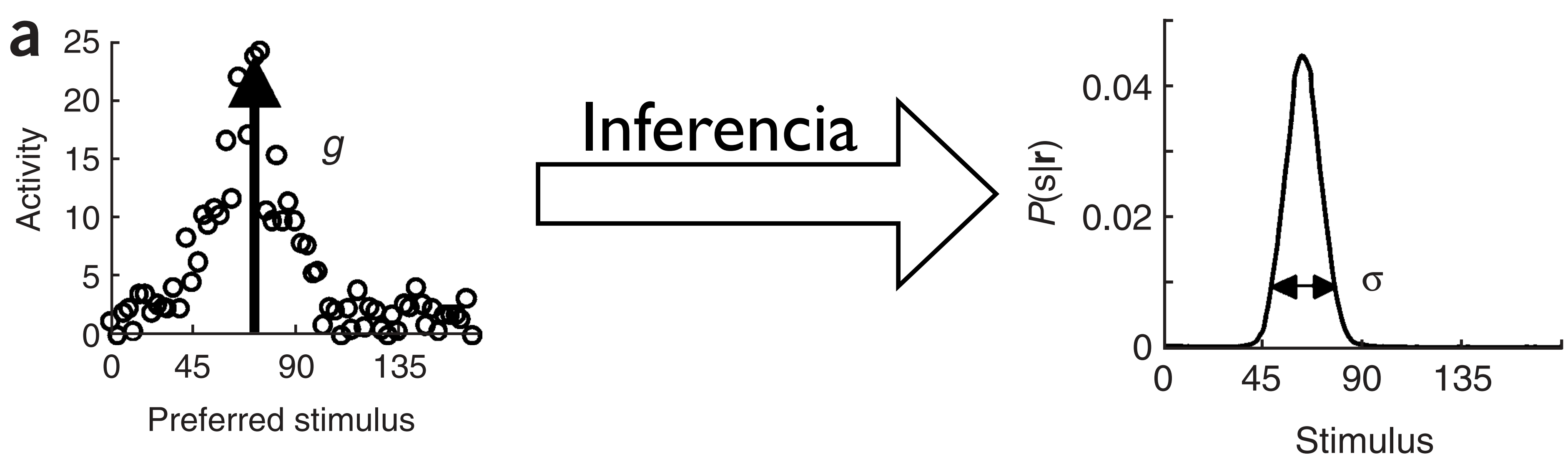
$$P(\mathbf{r} \mid s) = \prod_i \frac{e^{-f_i(s)} f_i(s)^{r_i}}{r_i!}$$



Probabilistic Population Codes

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

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

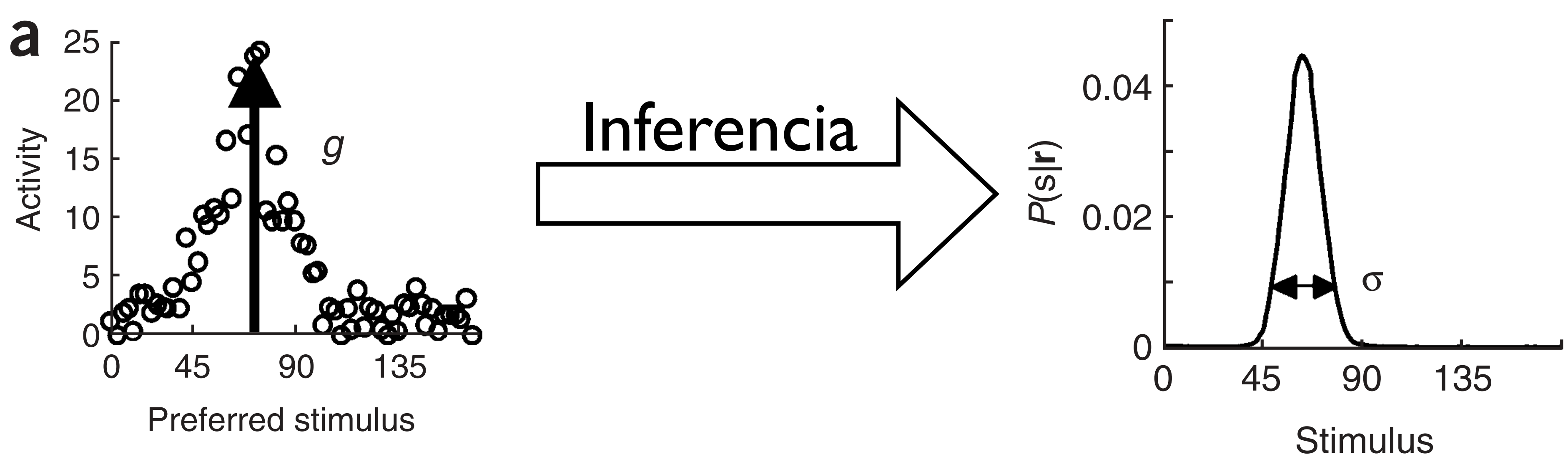


az aktivitás-intenzítás arányos a precizióval

Probabilistic Population Codes

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

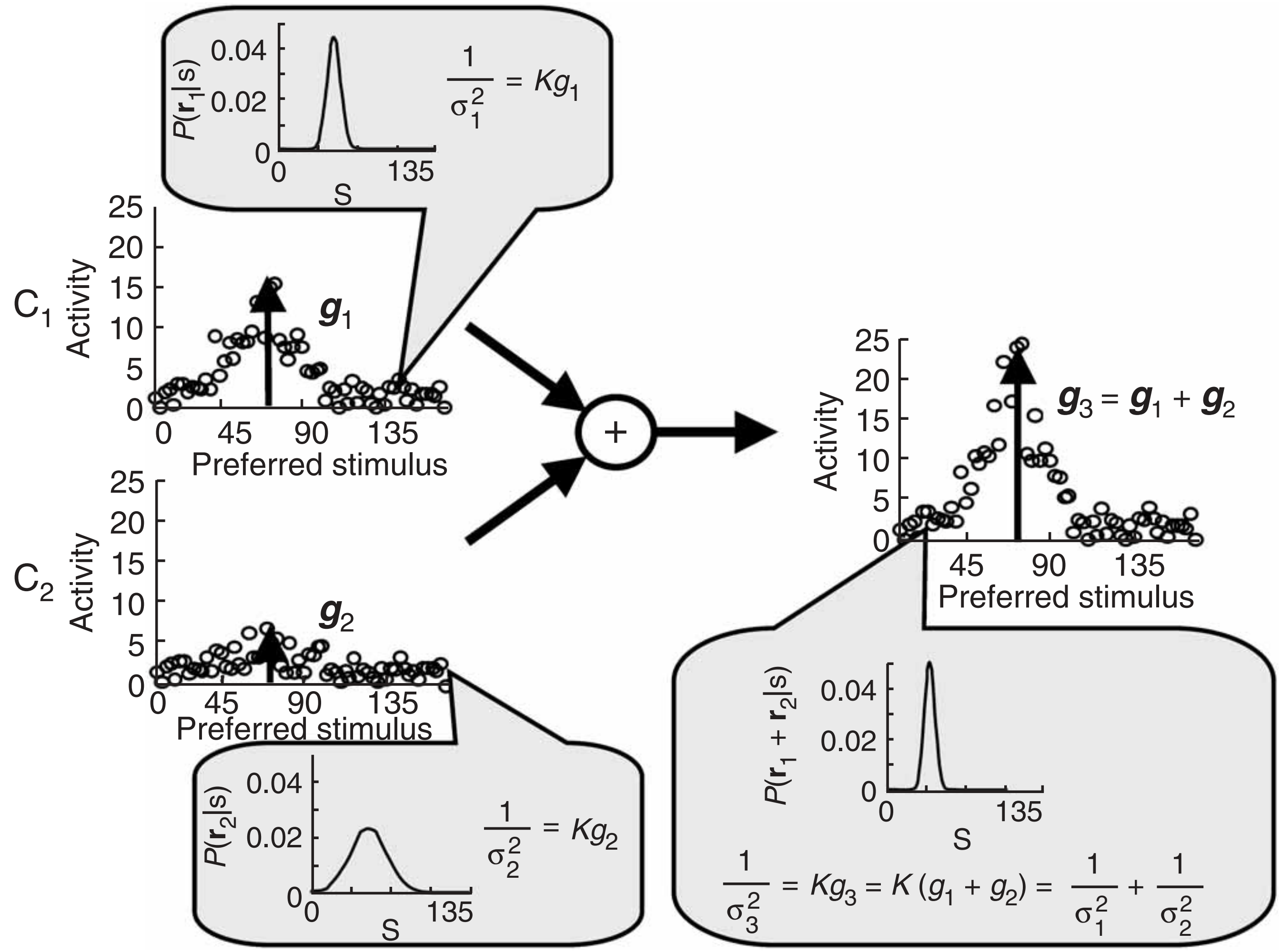
$$P(\mathbf{r} \mid s) = \prod_i \frac{e^{-f_i(s)} f_i(s)^{r_i}}{r_i!}$$



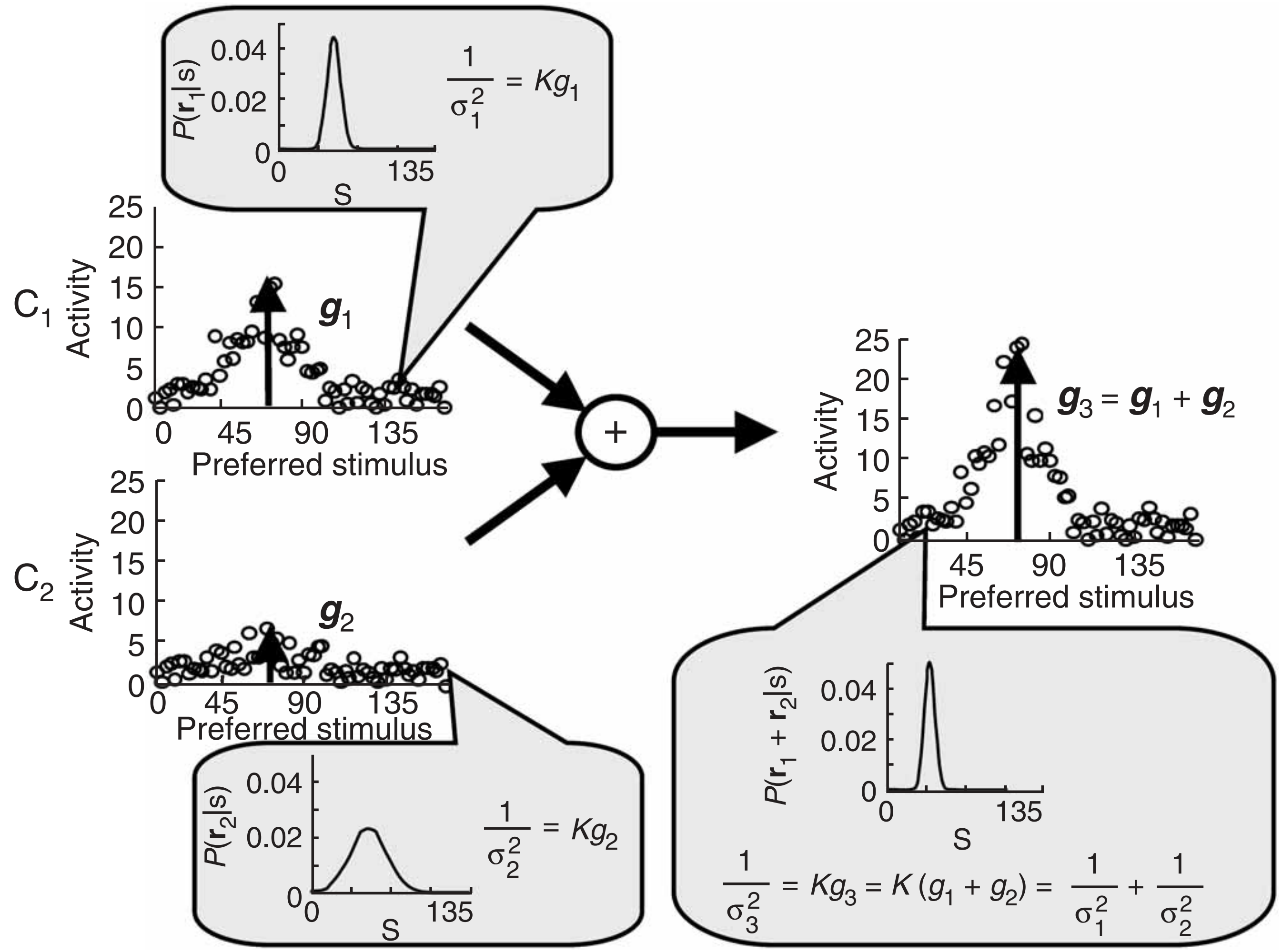
az aktivitás-intenzítás arányos a precizióval

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

PPC: Multiszenzoros integráció

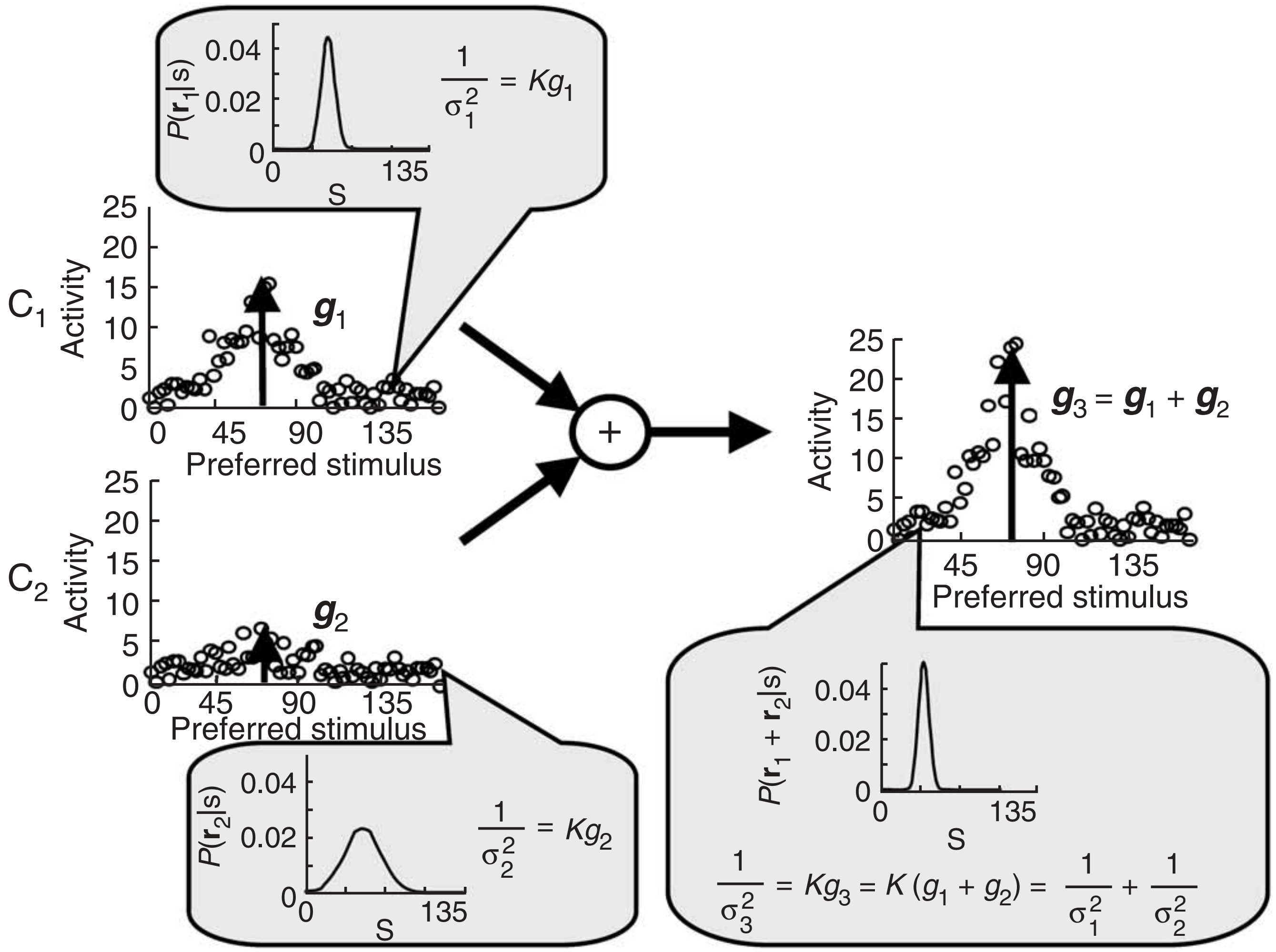


PPC: Multiszenzoros integráció



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

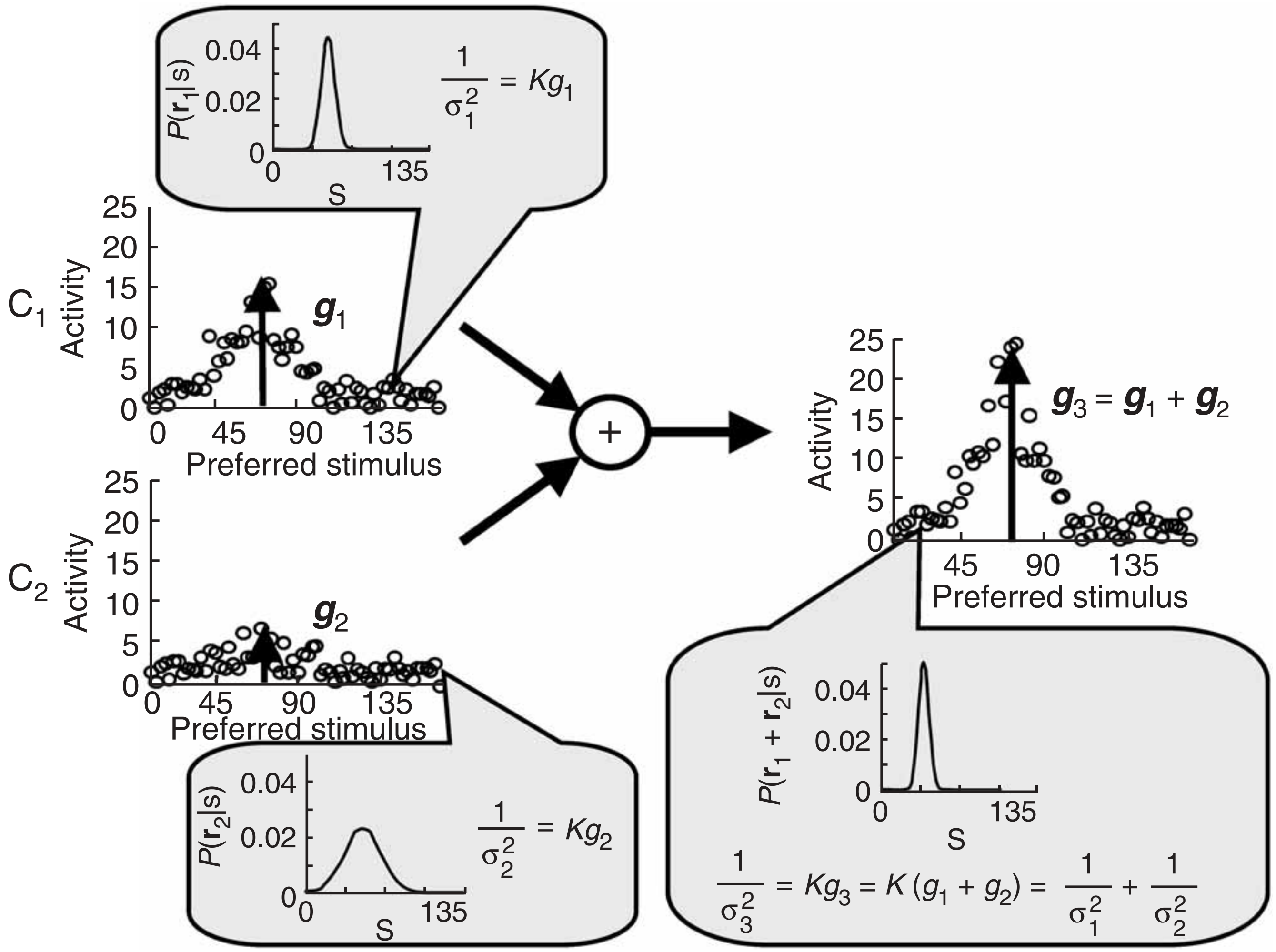
PPC: Multiszenzoros integráció



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

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

PPC: Multiszenzoros integráció



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

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

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

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

Probabilistic population codes vs sampling

Probabilistic population codes vs sampling

	PPC	sampling

Probabilistic population codes vs sampling

	PPC	sampling
neurons correspond to	parameters	variables

Probabilistic population codes vs sampling

	PPC	sampling
neurons correspond to	parameters	variables
required network dynamics	deterministic	stochastic

Probabilistic population codes vs sampling

	PPC	sampling
neurons correspond to	parameters	variables
required network dynamics	deterministic	stochastic
representable distributions	specific parametric form	nonparametric — srbitrary

Probabilistic population codes vs sampling

	PPC	sampling
neurons correspond to	parameters	variables
required network dynamics	deterministic	stochastic
representable distributions	specific parametric form	nonparametric — srbitrary
critical factor for encoding accuracy	number of neurons	time

Probabilistic population codes vs sampling

	PPC	sampling
neurons correspond to	parameters	variables
required network dynamics	deterministic	stochastic
representable distributions	specific parametric form	nonparametric — srbitrary
critical factor for encoding accuracy	number of neurons	time
instantaneous representation of uncertainty	yes, the whole distribution is represented at any given time	no, a sequence of samples is required

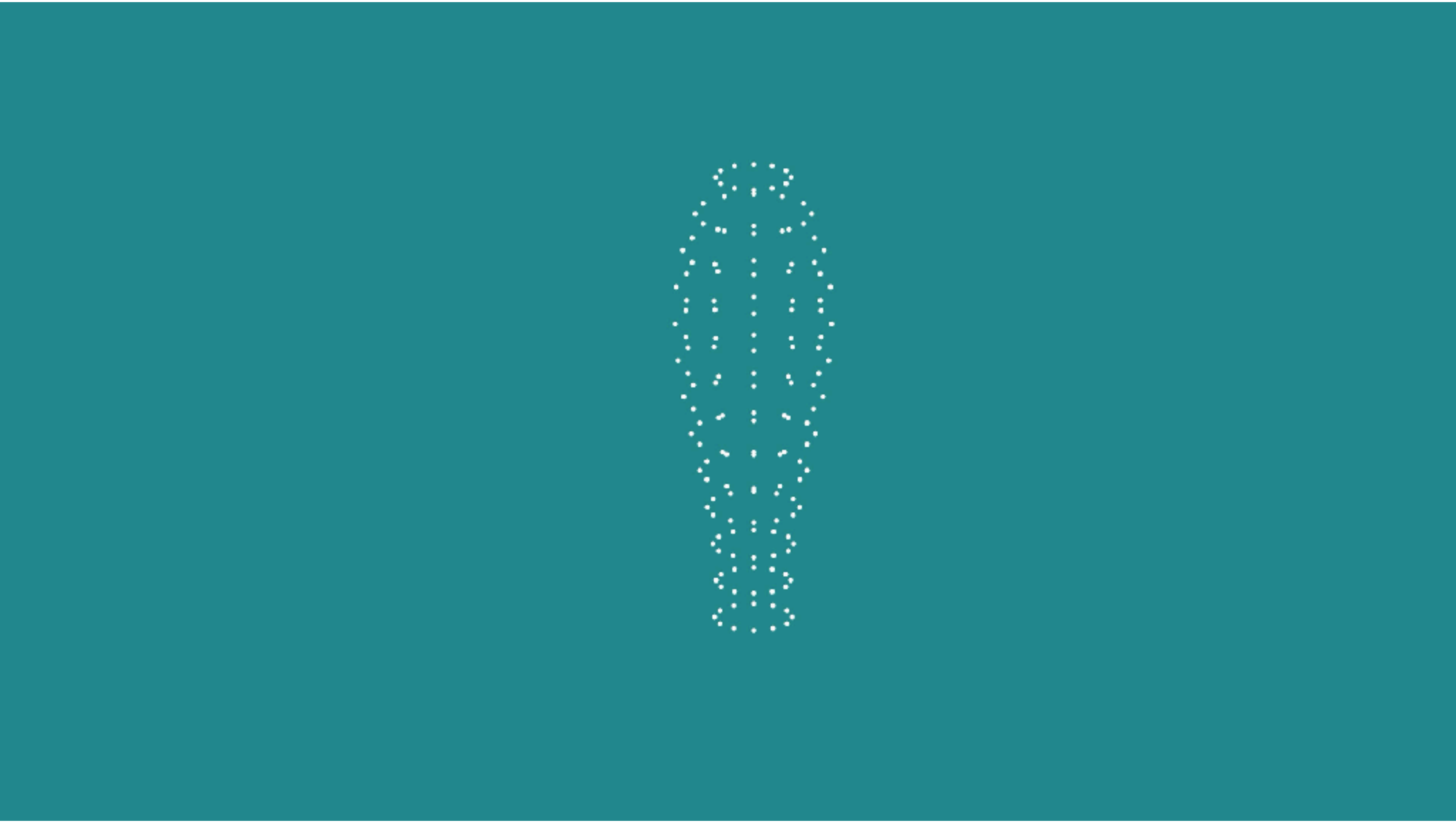
Probabilistic population codes vs sampling

	PPC	sampling
neurons correspond to	parameters	variables
required network dynamics	deterministic	stochastic
representable distributions	specific parametric form	nonparametric — srbitrary
critical factor for encoding accuracy	number of neurons	time
instantaneous representation of uncertainty	yes, the whole distribution is represented at any given time	no, a sequence of samples is required
neurons required for multivariate distributions	scales exponentially with the number of dimensions	scales linearly with the number of dimensions

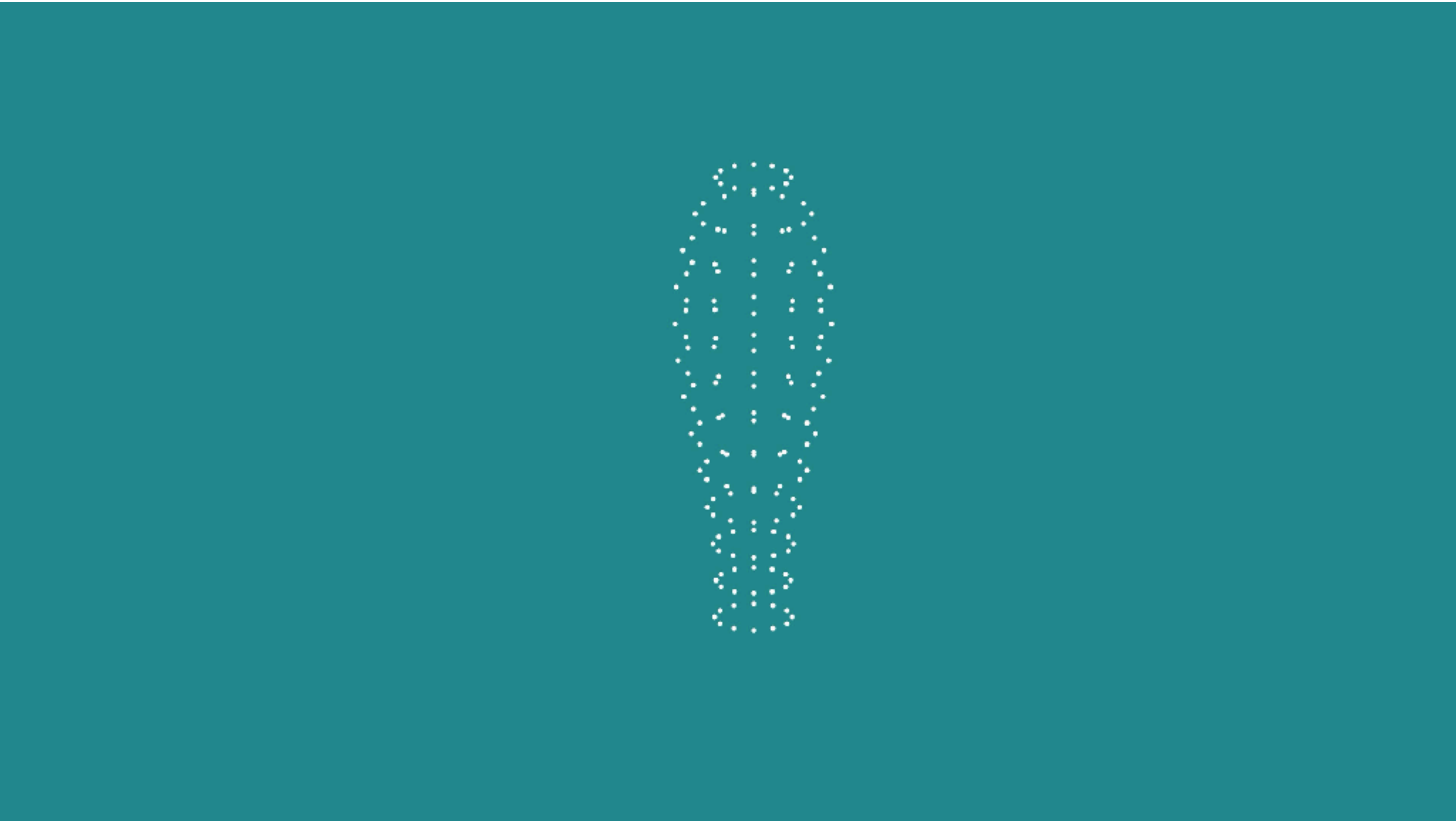
Probabilistic population codes vs sampling

	PPC	sampling
neurons correspond to	parameters	variables
required network dynamics	deterministic	stochastic
representable distributions	specific parametric form	nonparametric — srbitrary
critical factor for encoding accuracy	number of neurons	time
instantaneous representation of uncertainty	yes, the whole distribution is represented at any given time	no, a sequence of samples is required
neurons required for multivariate distributions	scales exponentially with the number of dimensions	scales linearly with the number of dimensions
implementation of learning	cumbersome	manageable

sampling in perception

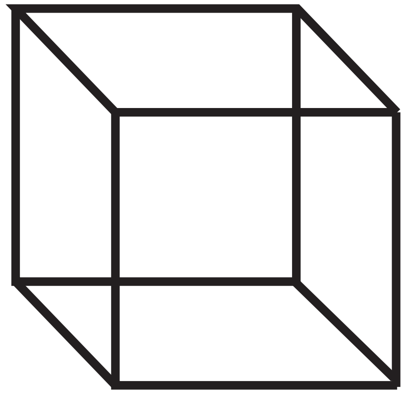


sampling in perception



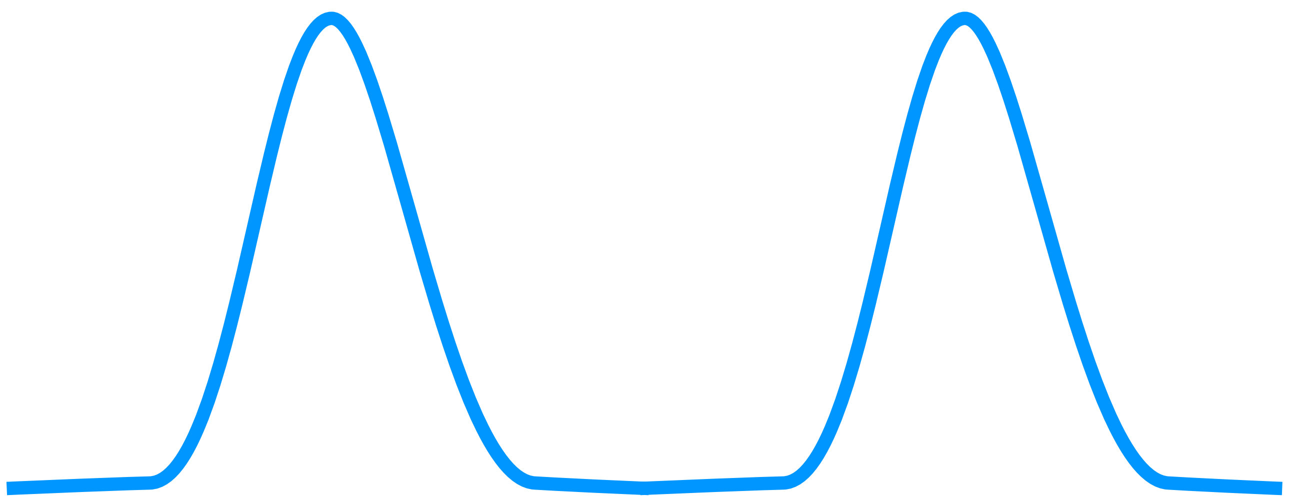
sampling in perception

Necker cube



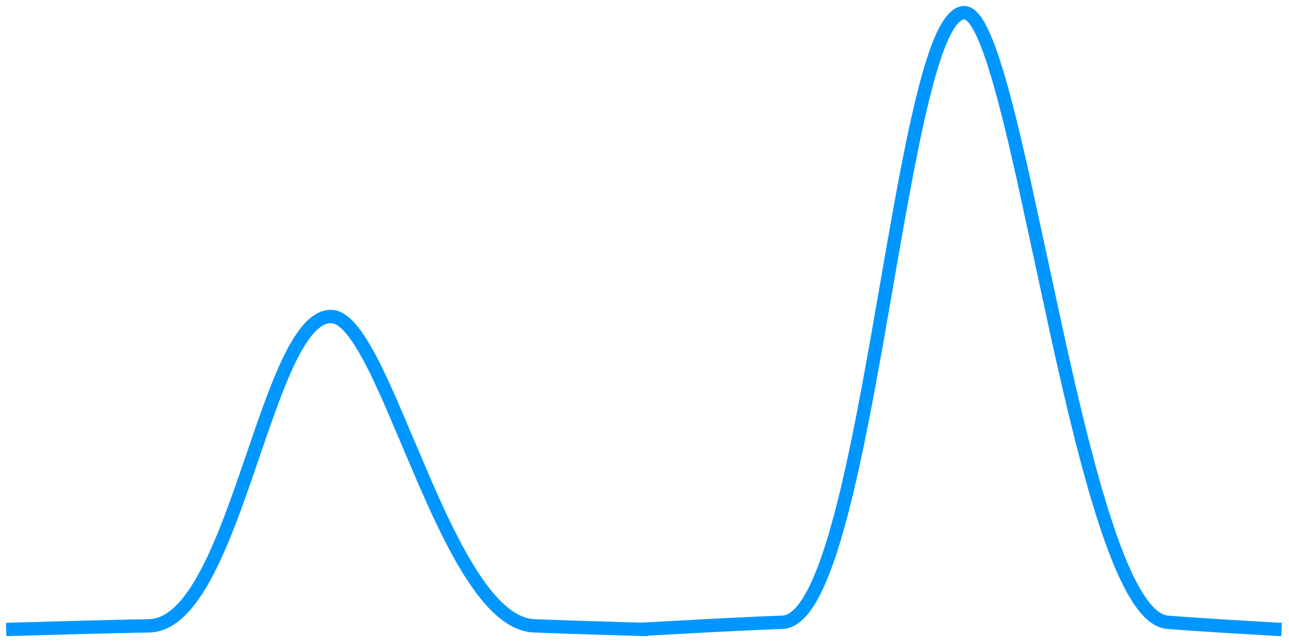
Quantitative consequences of sampling

Moreno Bote et al (2011) PNAS



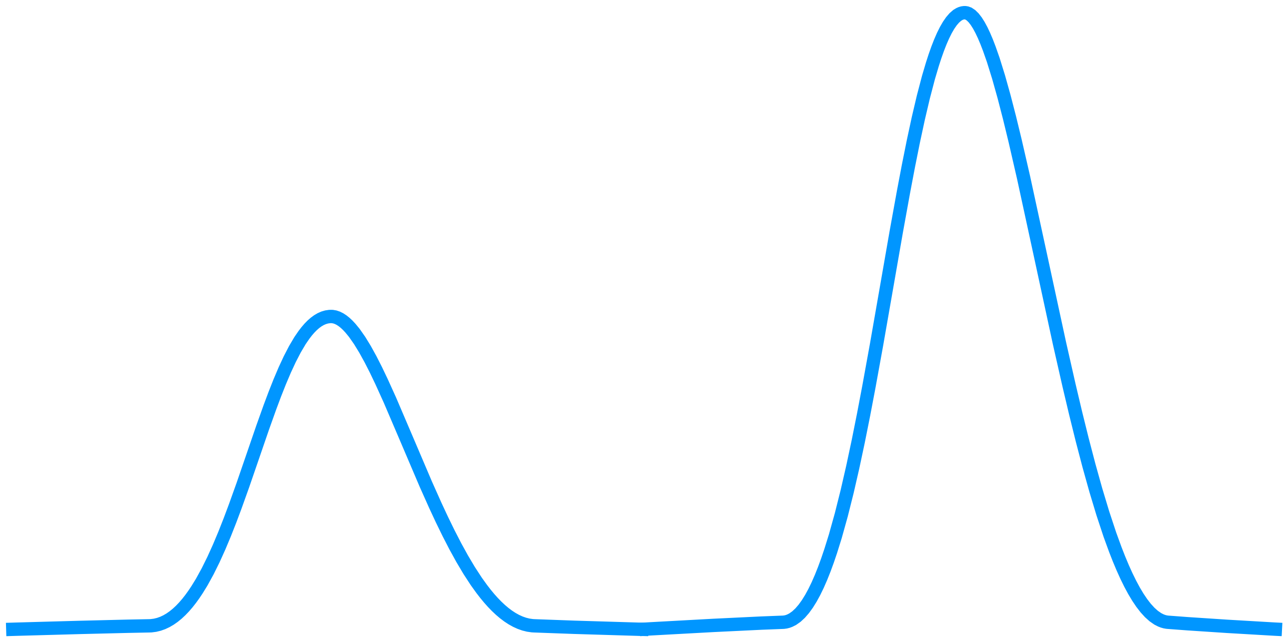
Quantitative consequences of sampling

Moreno Bote et al (2011) PNAS



Quantitative consequences of sampling

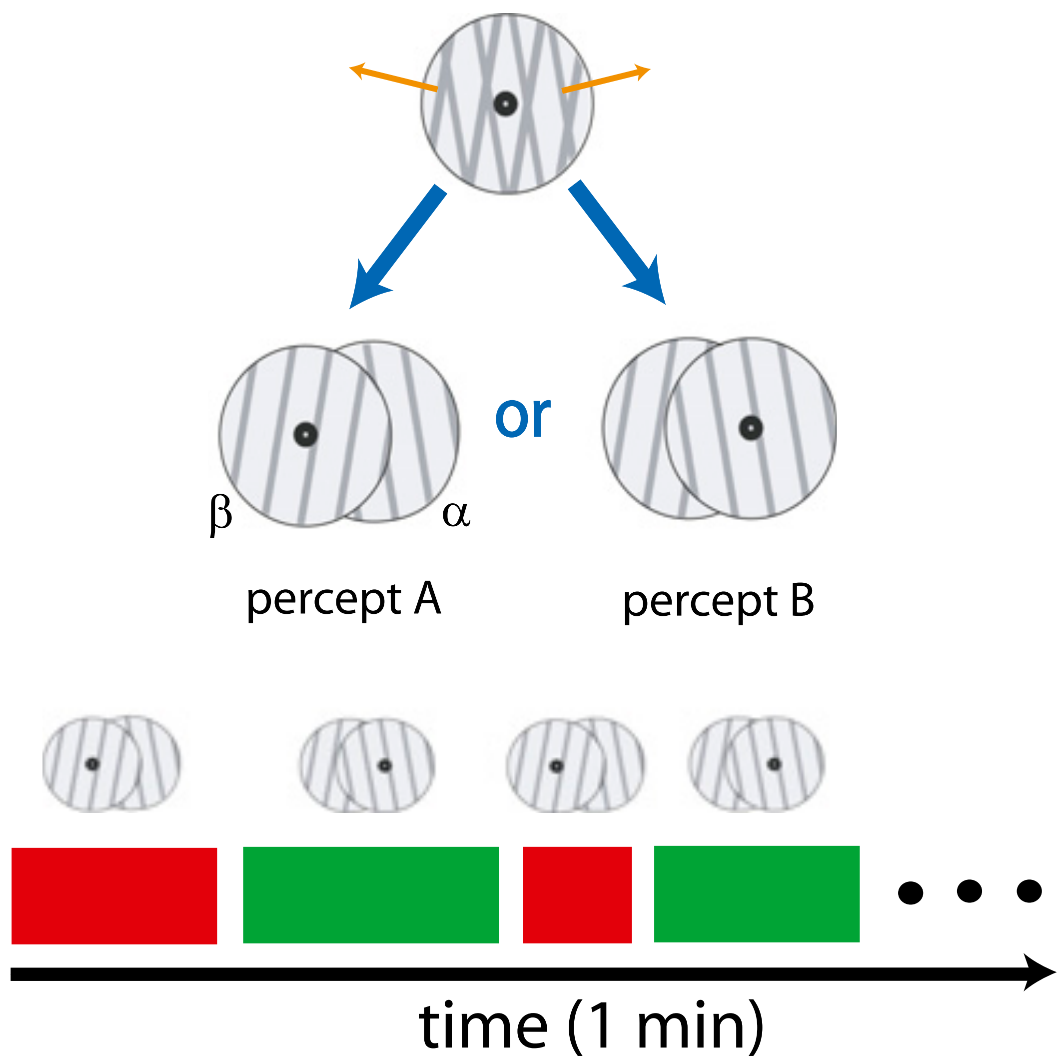
Moreno Bote et al (2011) PNAS



- Different weighings of different modes of the posterior introduce systematic variations in sampling times
- Relative dominance of percepts can be predicted

Quantitative consequences of sampling

Moreno Bote et al (2011) PNAS



- Binocularly projected moving grating images
- The proportion of one or the other perceived in the foreground is measured

Quantitative consequences of sampling

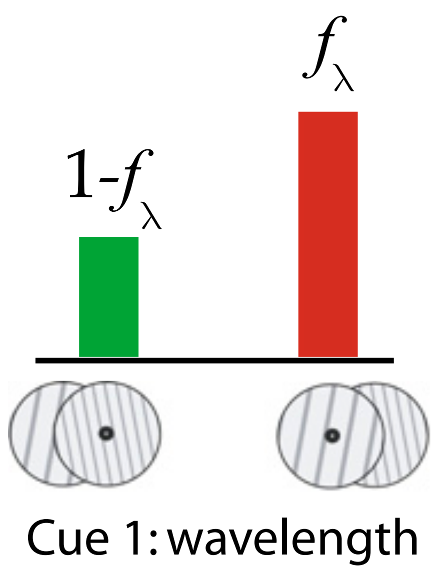
Moreno Bote et al (2011) PNAS

- Two cues can be manipulated:
  - wavelength of the grating
  - speed of the grating
- The cues are affecting independently the dominance of percepts — the weights of the modes

Quantitative consequences of sampling

Moreno Bote et al (2011) PNAS

- Two cues can be manipulated:
  - wavelength of the grating
  - speed of the grating
- The cues are affecting independently the dominance of percepts — the weights of the modes

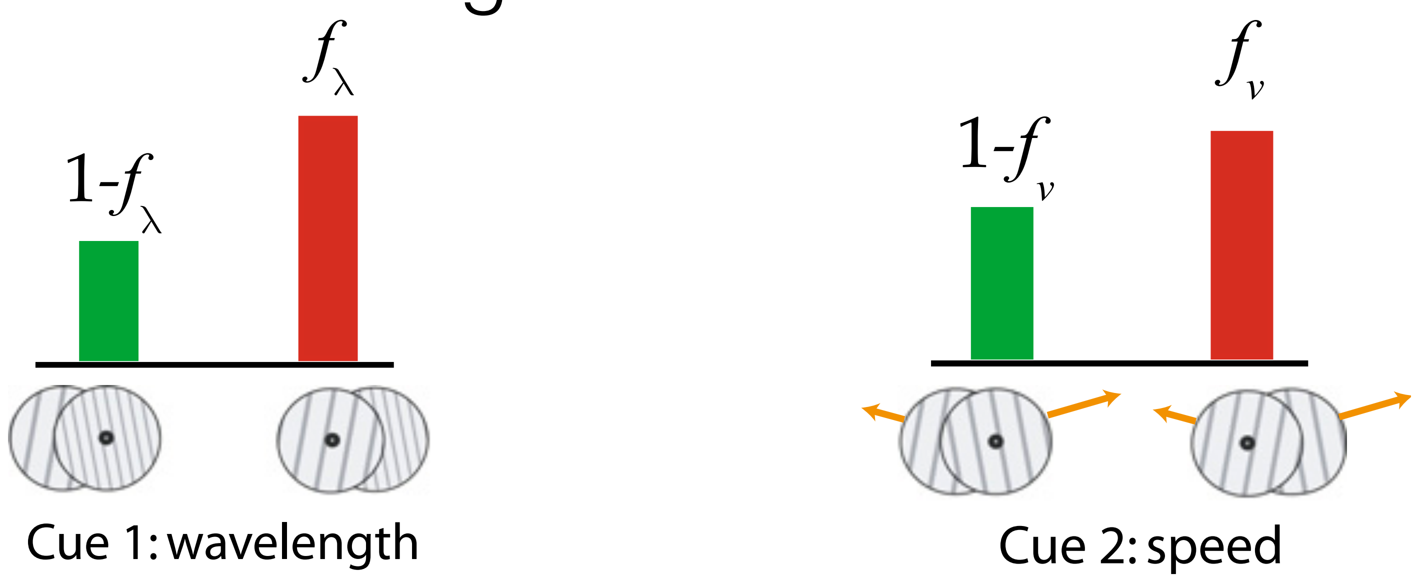


Quantitative consequences of sampling

Moreno Bote et al (2011) PNAS

- Two cues can be manipulated:
  - wavelength of the grating
  - speed of the grating

- The cues are affecting independently the dominance of percepts — the weights of the modes

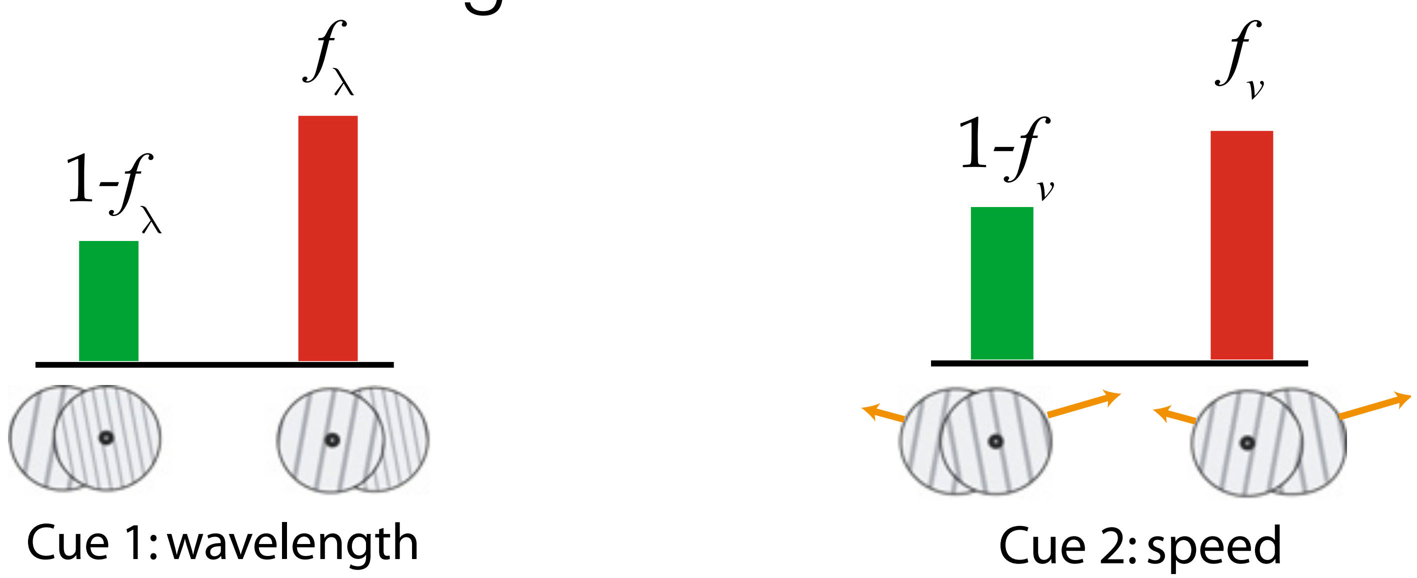


Quantitative consequences of sampling

Moreno Bote et al (2011) PNAS

- Two cues can be manipulated:
  - wavelength of the grating
  - speed of the grating

- The cues are affecting independently the dominance of percepts — the weights of the modes

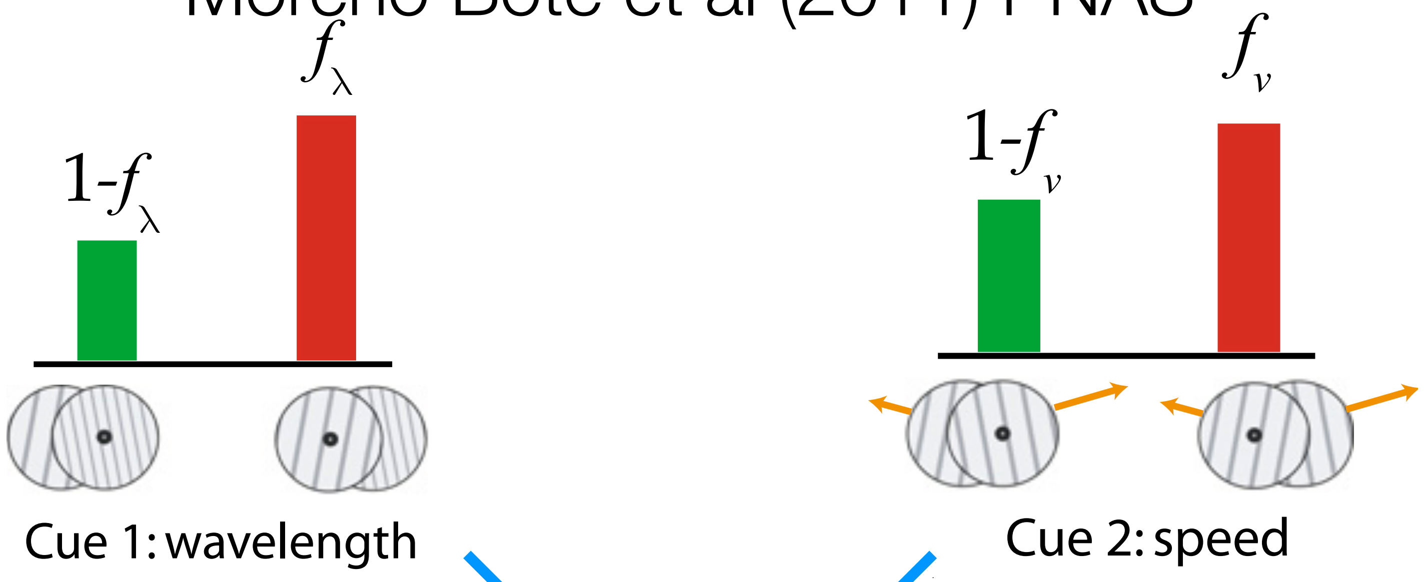


- Simultaneous presence of the two cues results in combination of the probability distributions implied by individual cues:

$$f_{\lambda v} = \frac{f_\lambda f_v}{f_\lambda f_v + (1-f_\lambda)(1-f_v)}$$

Quantitative consequences of sampling

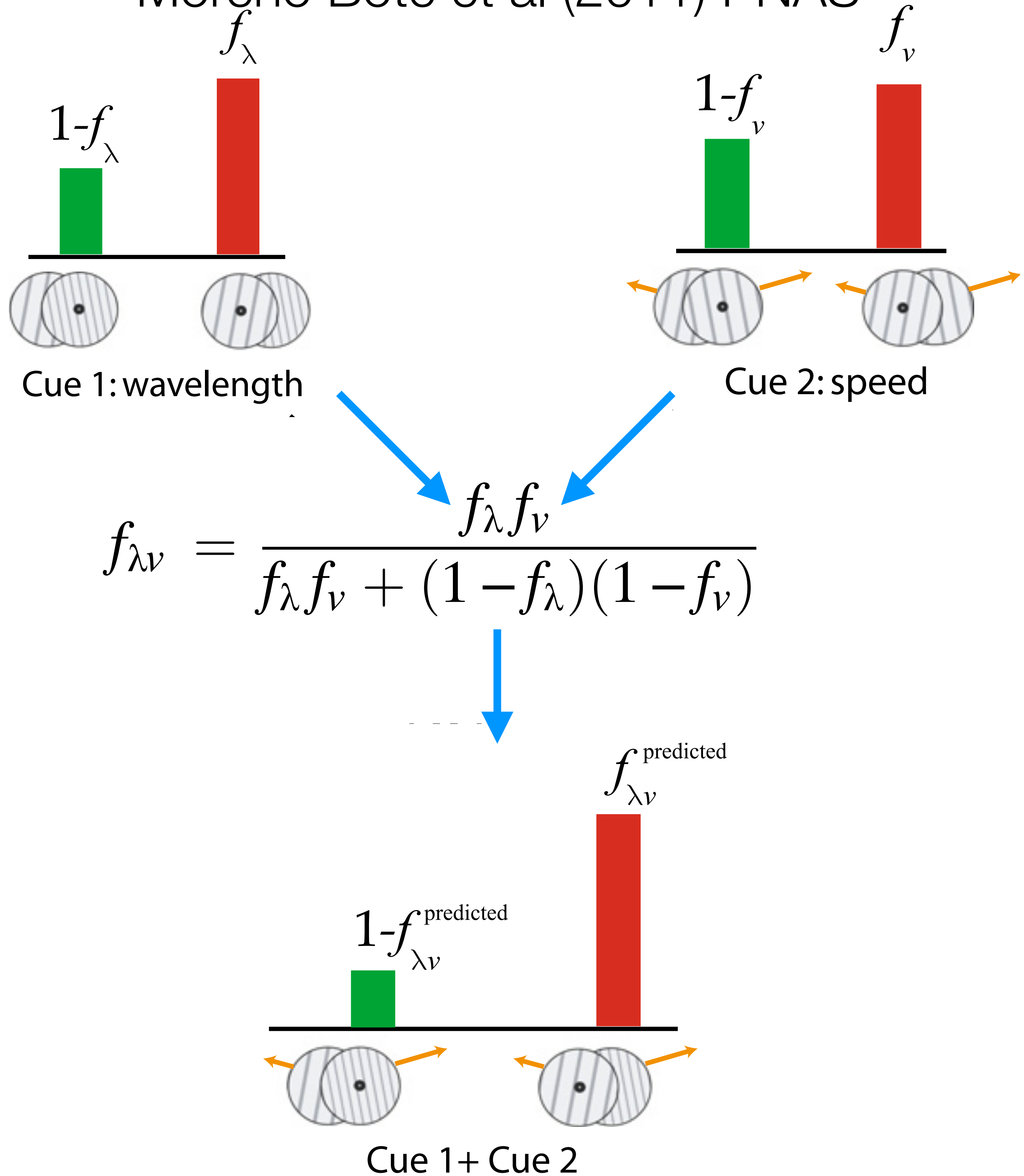
Moreno Bote et al (2011) PNAS



$$f_{\lambda v} = \frac{f_\lambda f_v}{f_\lambda f_v + (1-f_\lambda)(1-f_v)}$$

Quantitative consequences of sampling

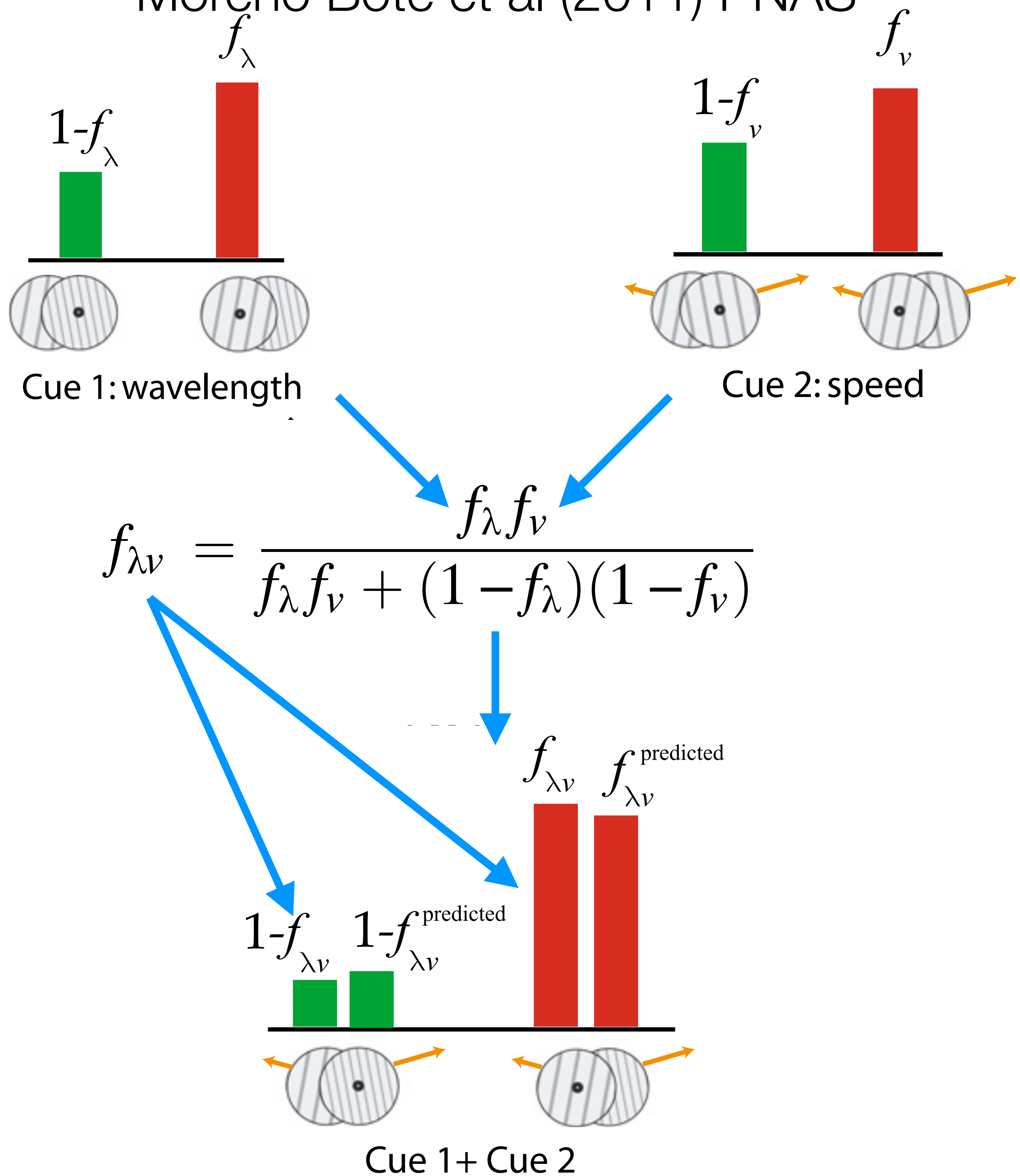
Moreno Bote et al (2011) PNAS



$$f_{\lambda v} = \frac{f_\lambda f_v}{f_\lambda f_v + (1-f_\lambda)(1-f_v)}$$

Quantitative consequences of sampling

Moreno Bote et al (2011) PNAS



$$f_{\lambda v} = \frac{f_\lambda f_v}{f_\lambda f_v + (1-f_\lambda)(1-f_v)}$$

Quantitative consequences of sampling

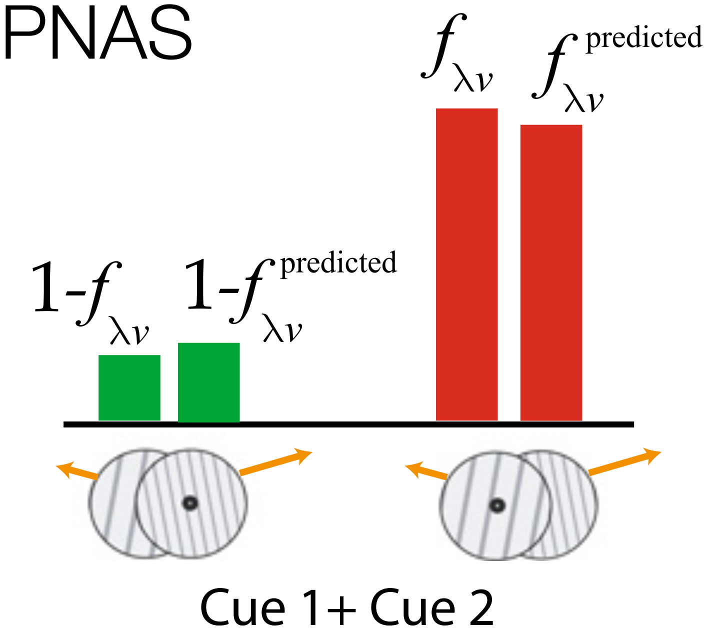
Moreno Bote et al (2011) PNAS

$$f_{\lambda\nu} = \frac{f_\lambda f_\nu}{f_\lambda f_\nu + (1-f_\lambda)(1-f_\nu)}$$

Quantitative consequences of sampling

Moreno Bote et al (2011) PNAS

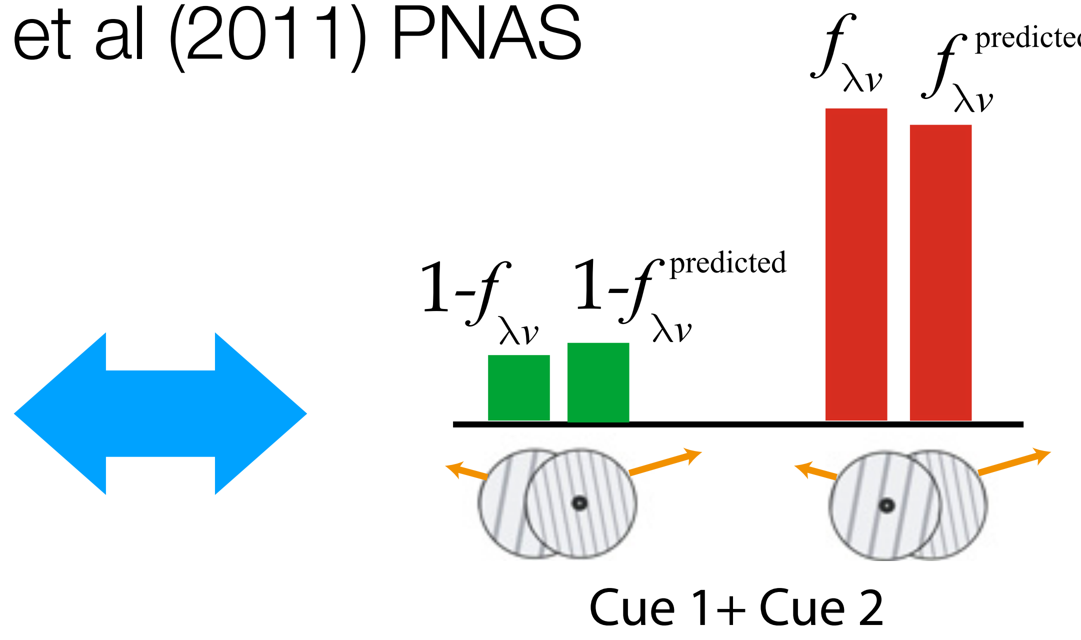
$$f_{\lambda\nu} = \frac{f_\lambda f_\nu}{f_\lambda f_\nu + (1-f_\lambda)(1-f_\nu)}$$

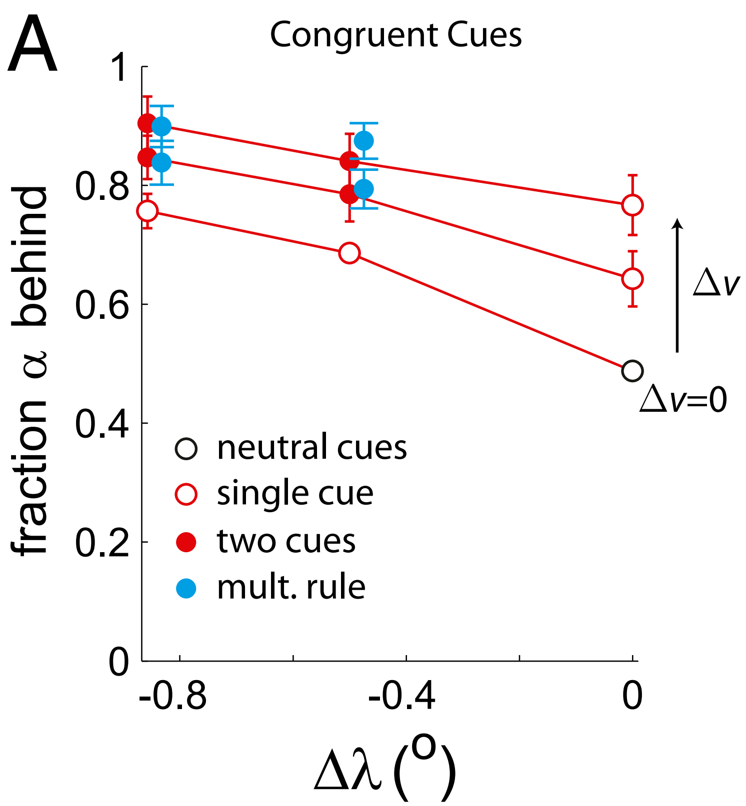


Quantitative consequences of sampling

Moreno Bote et al (2011) PNAS

$$f_{\lambda v} = \frac{f_\lambda f_v}{f_\lambda f_v + (1-f_\lambda)(1-f_v)}$$

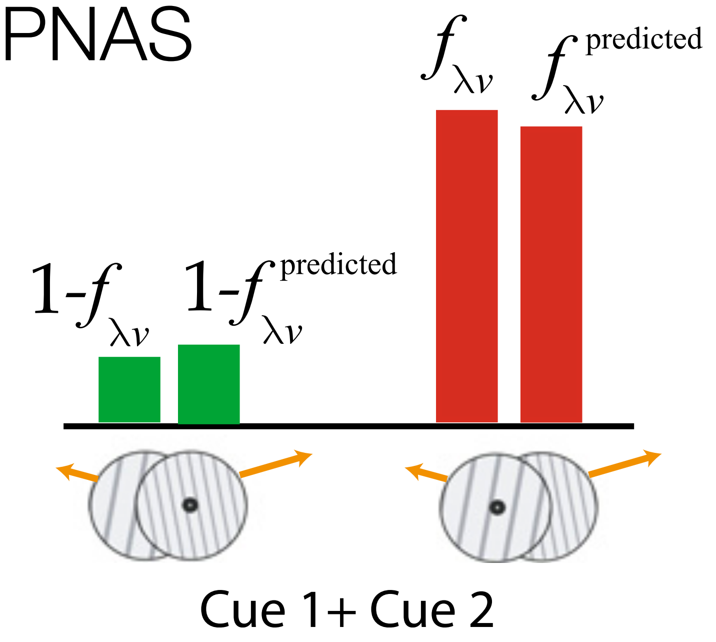


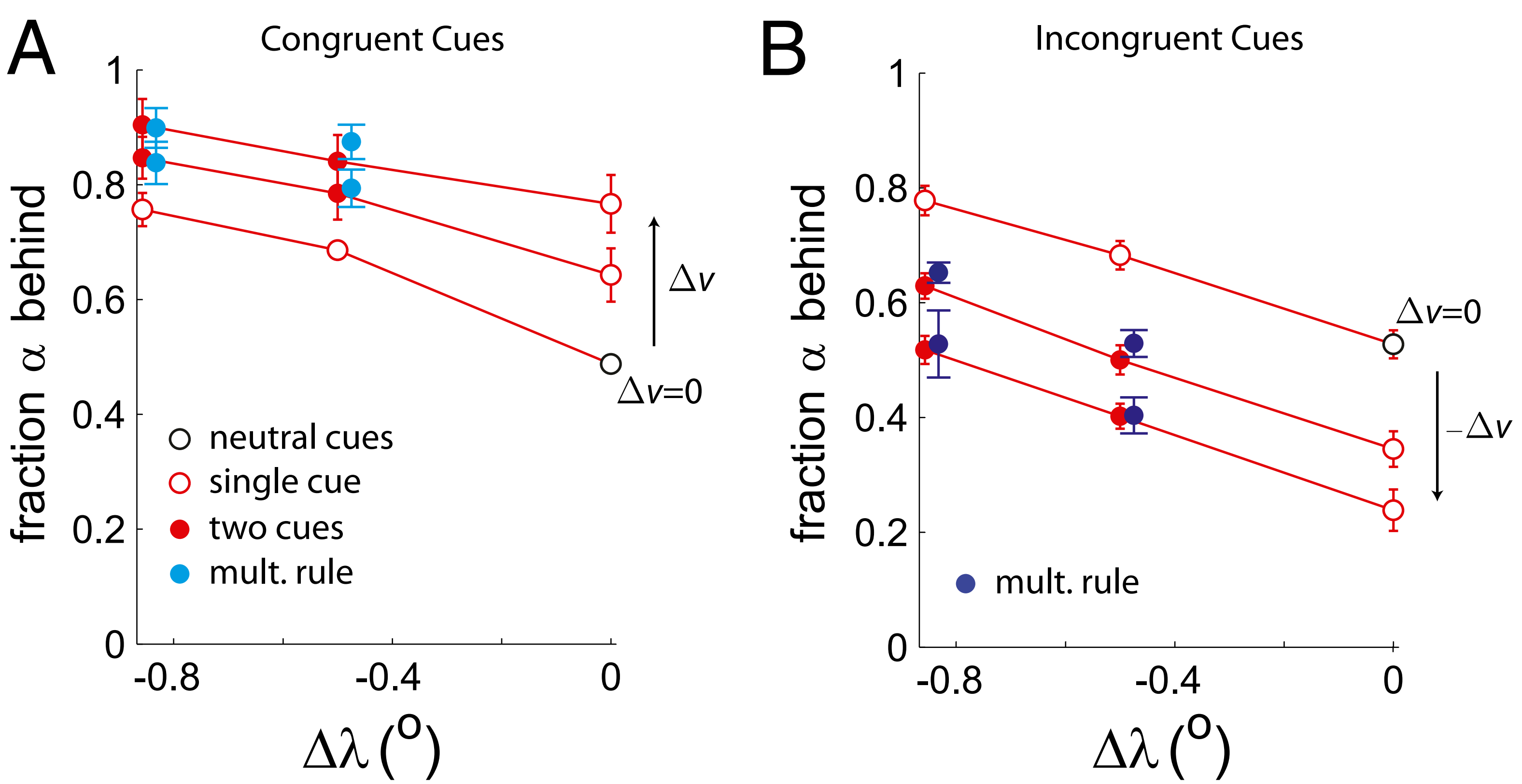


Quantitative consequences of sampling

Moreno Bote et al (2011) PNAS

$$f_{\lambda v} = \frac{f_\lambda f_v}{f_\lambda f_v + (1-f_\lambda)(1-f_v)}$$

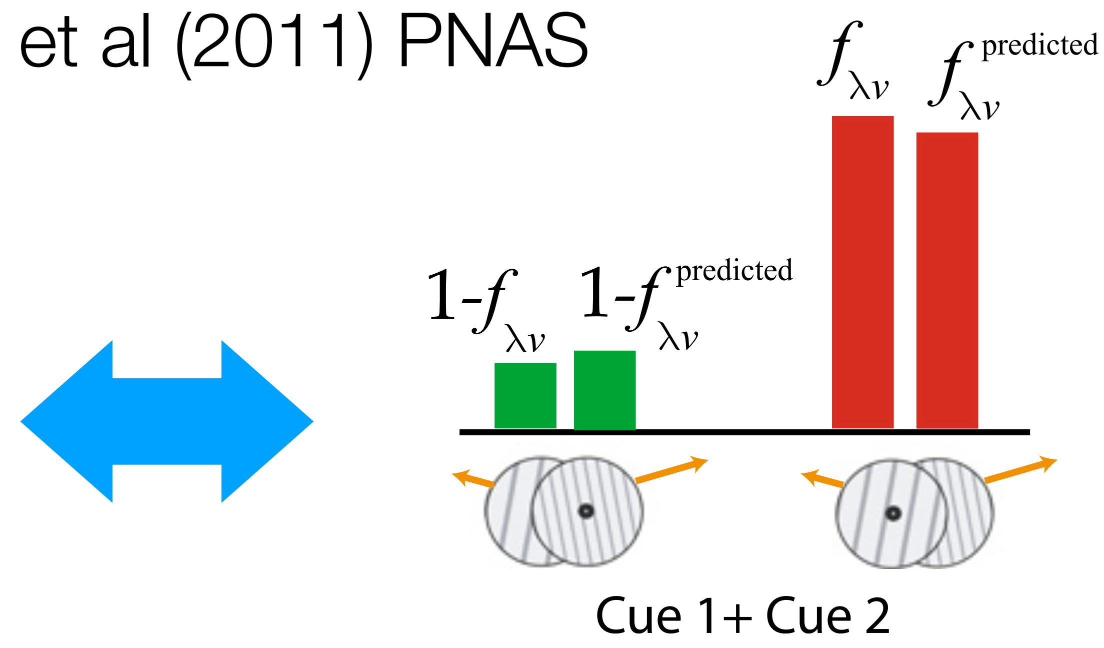


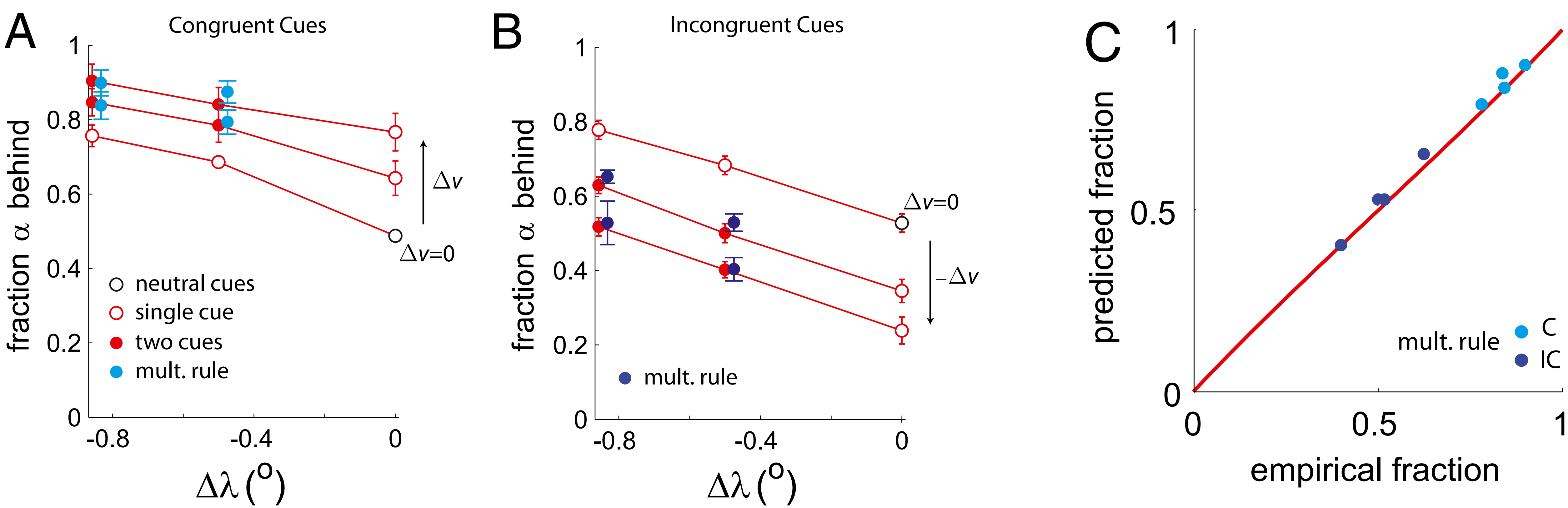


Quantitative consequences of sampling

Moreno Bote et al (2011) PNAS

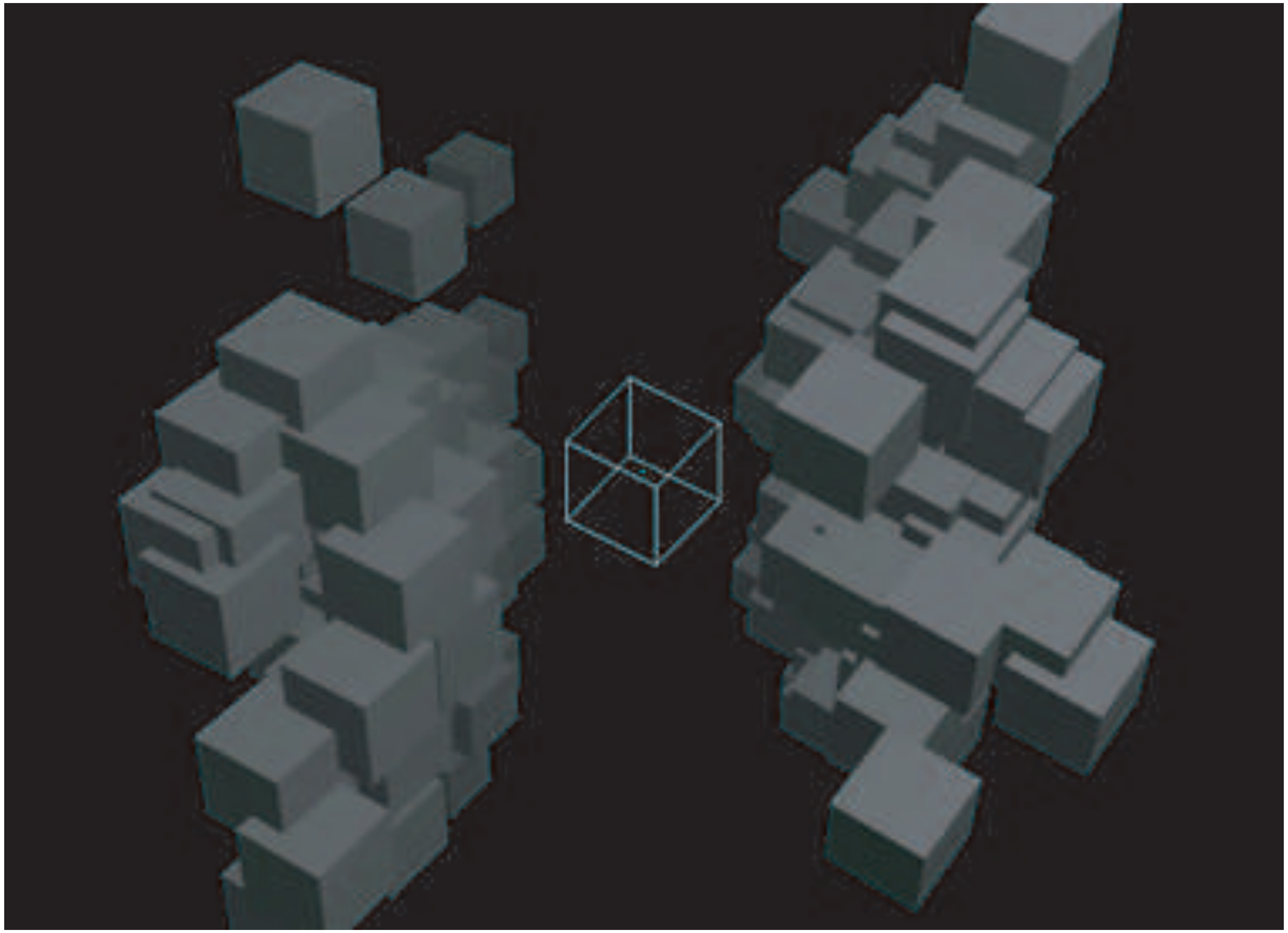
$$f_{\lambda v} = \frac{f_\lambda f_v}{f_\lambda f_v + (1-f_\lambda)(1-f_v)}$$





Contextual modulation of posterior

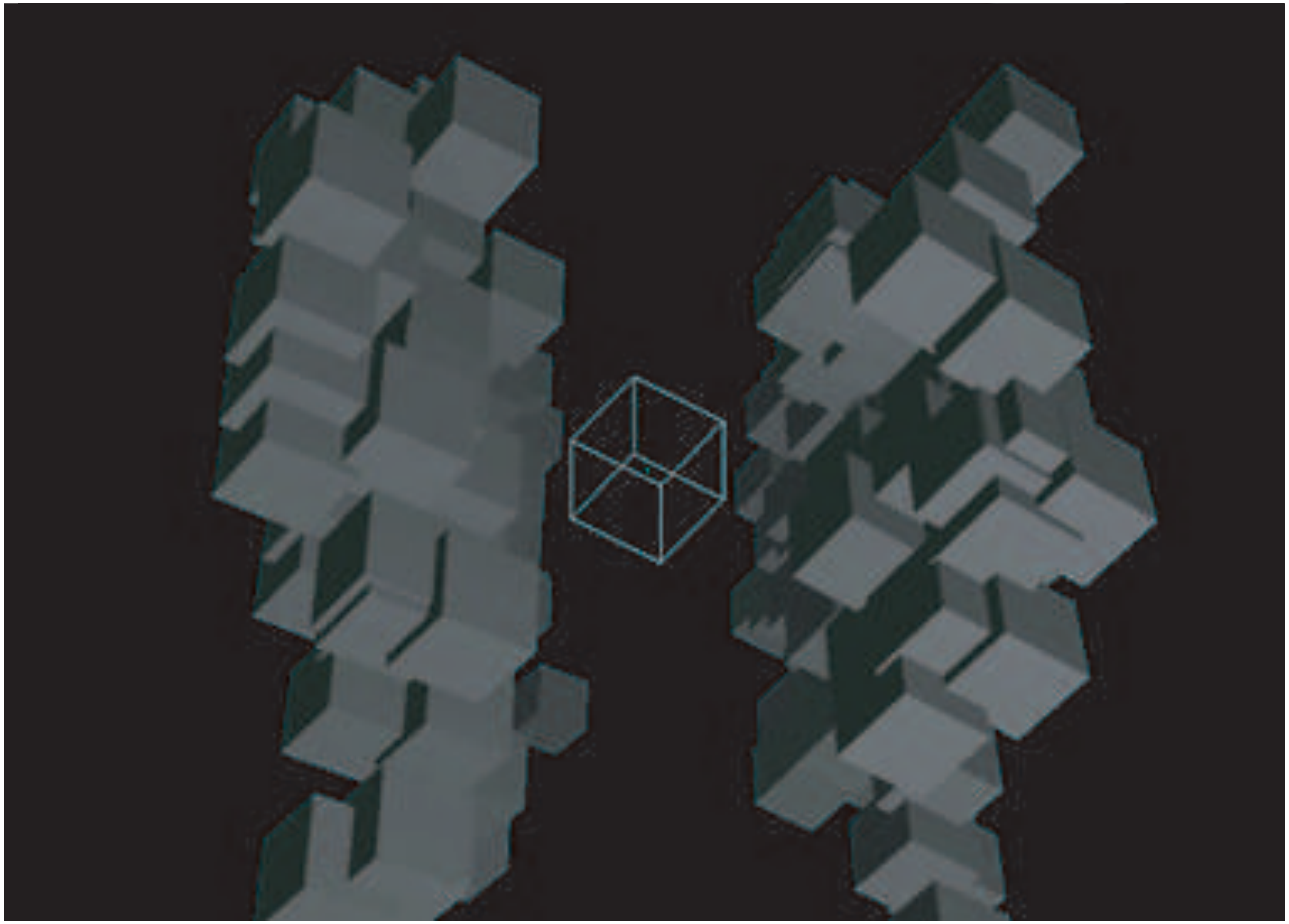
- Ambiguity can be resolved by contextual cues
- Dwelling times can be drastically modulated



Schrater & Sundereswara, NIPS, 2007

Contextual modulation of posterior

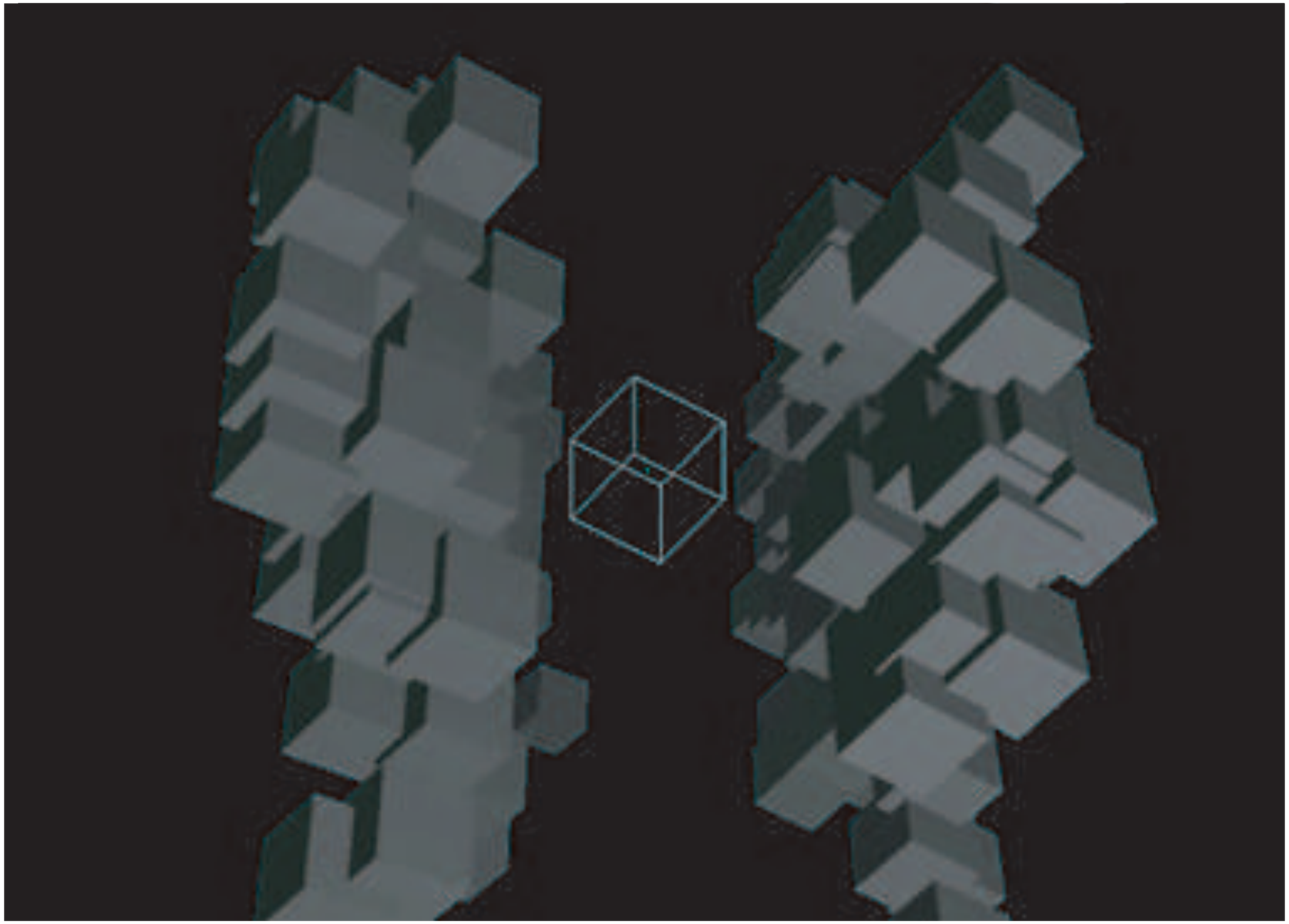
- Ambiguity can be resolved by contextual cues
- Dwelling times can be drastically modulated



Schrater & Sundereswara, NIPS, 2007

Contextual modulation of posterior

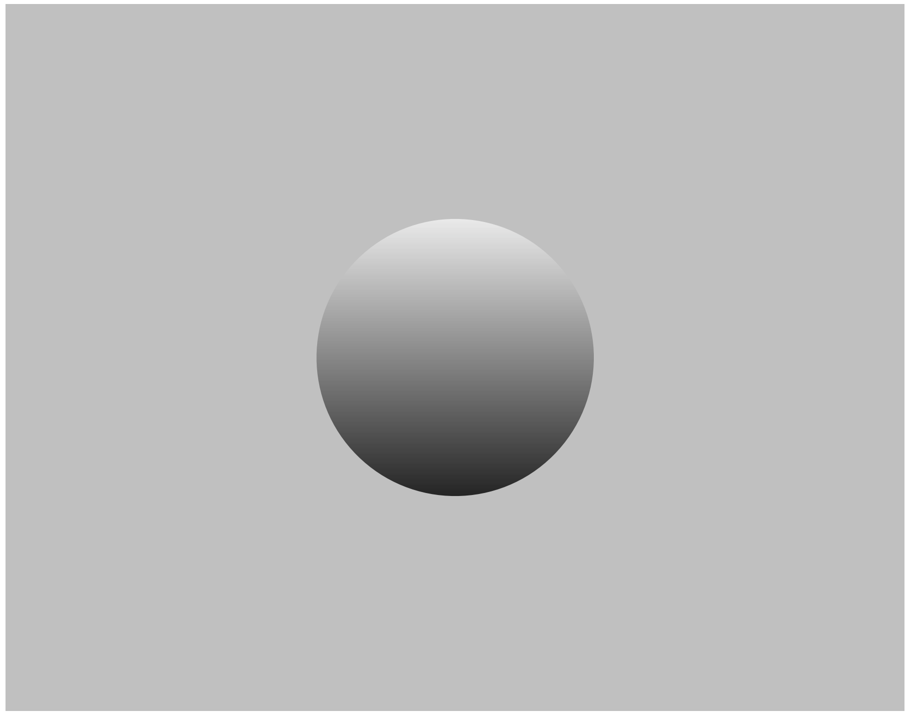
- Ambiguity can be resolved by contextual cues
- Dwelling times can be drastically modulated



Schrater & Sundereswara, NIPS, 2007

- Note: a much more delicate computation is happening here: conditioning on the context, assessment of probability of perspective

RECAP: role of priors



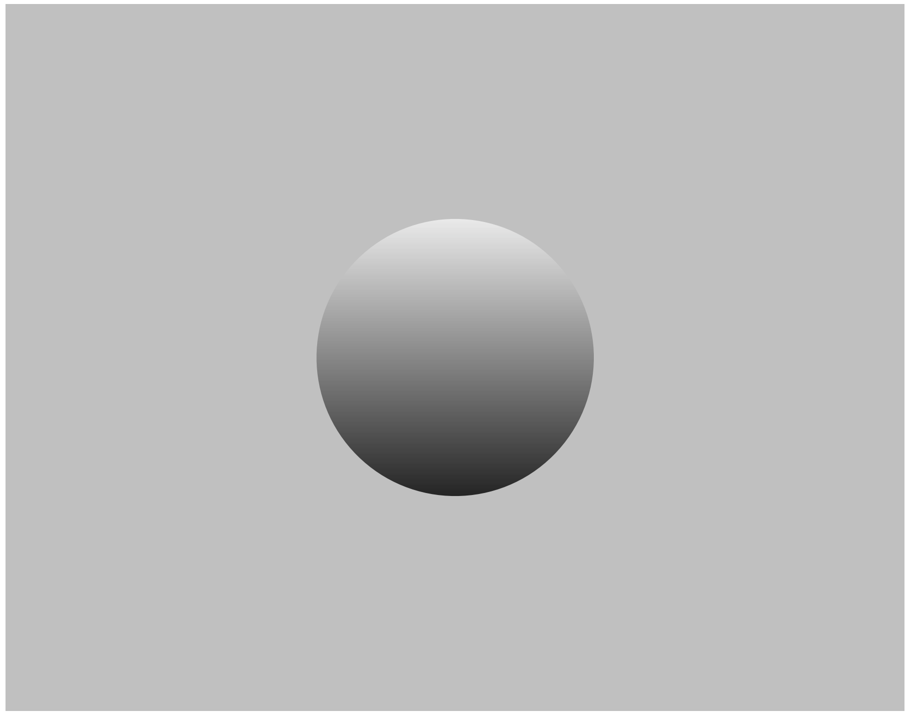
$$P(\text{feature} \,|\, \text{stimulus}) \propto P(\text{stimulus} \,|\, \text{feature}) \times P(\text{feature})$$

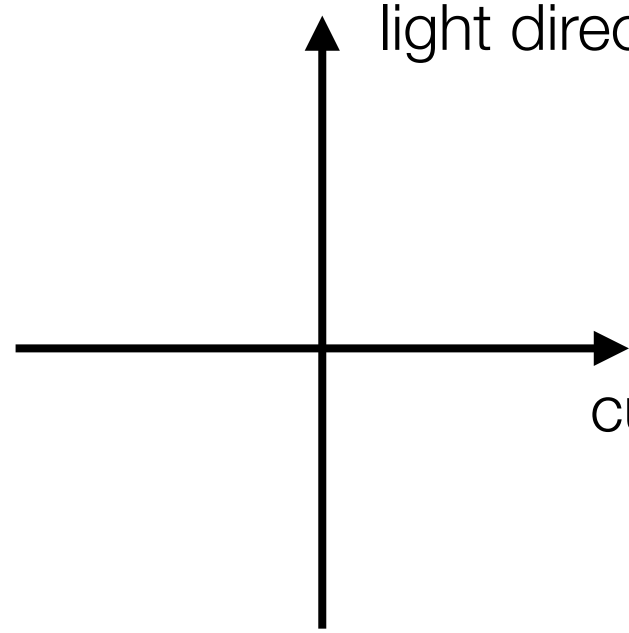
posterior: inference

likelihood: evidence

prior : expectations

RECAP: role of priors





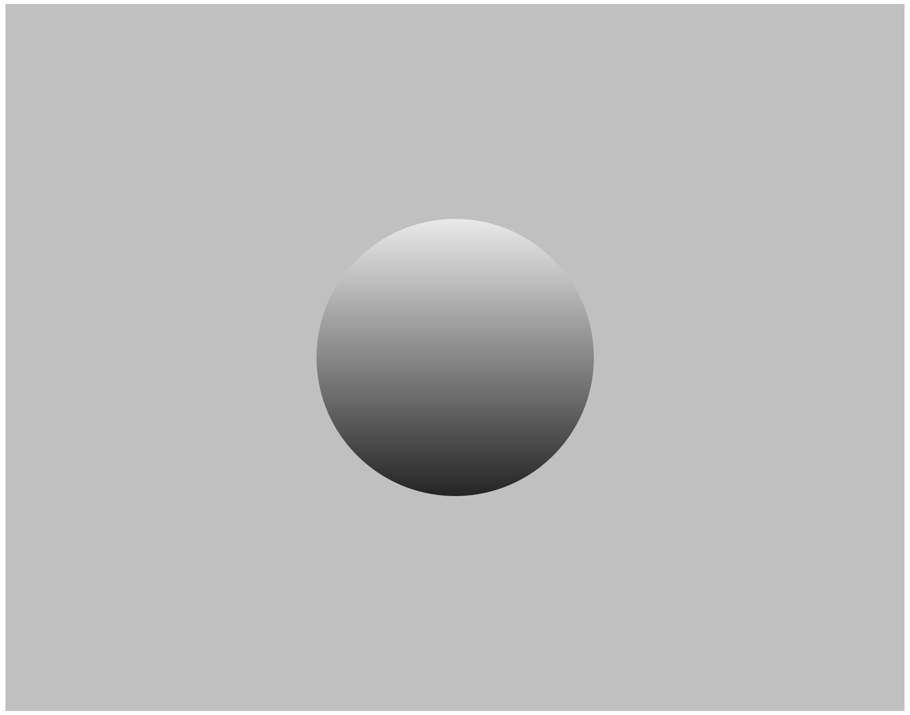
$$P(\text{feature}\,|\,\text{stimulus}) \propto P(\text{stimulus}\,|\,\text{feature}) \times P(\text{feature})$$

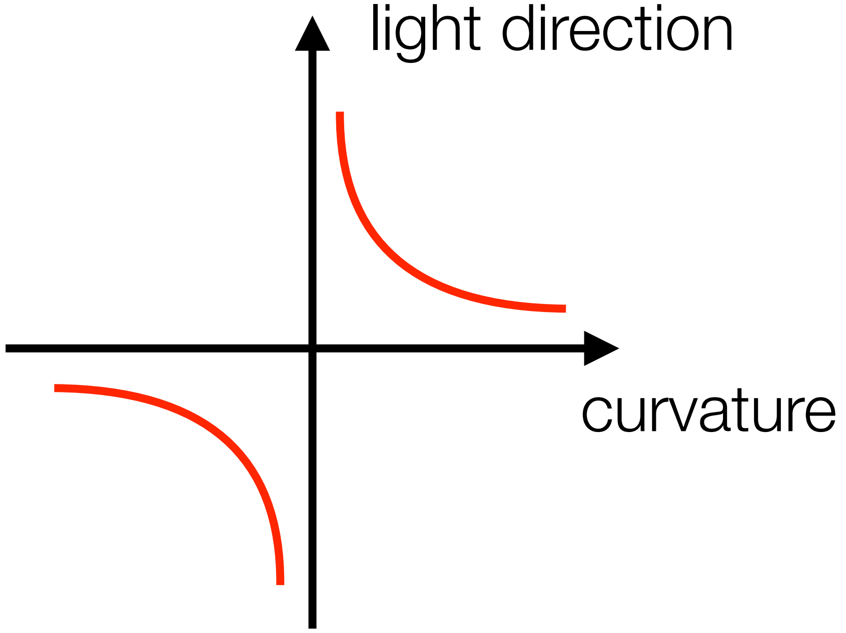
posterior: inference

likelihood: evidence

prior : expectations

RECAP: role of priors





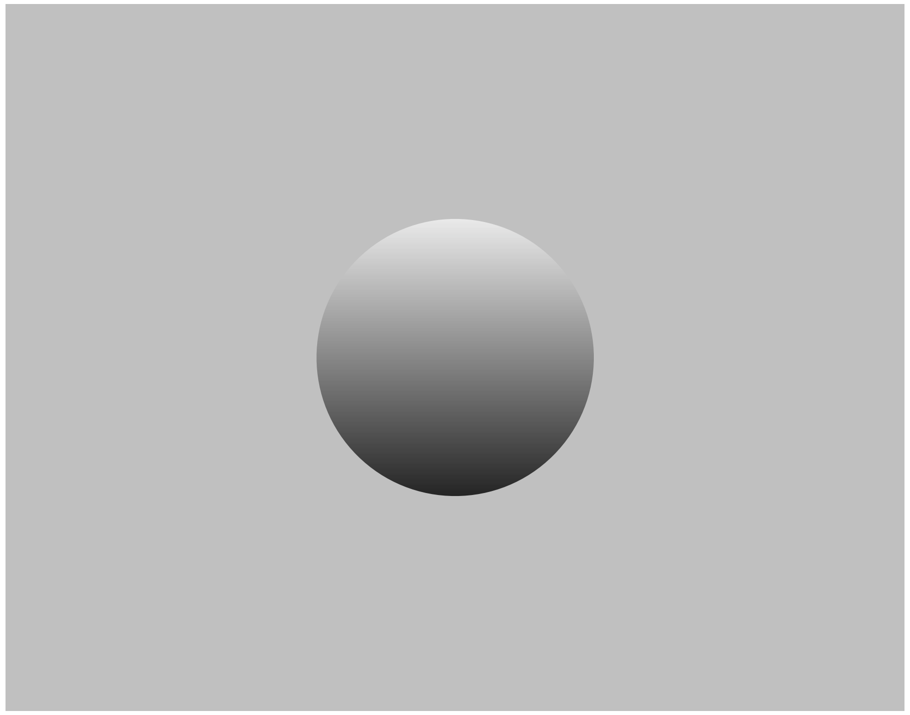
$$\color{green}{P(\text{feature} \mid \text{stimulus})} \propto \color{red}{P(\text{stimulus} \mid \text{feature})} \times \color{blue}{P(\text{feature})}$$

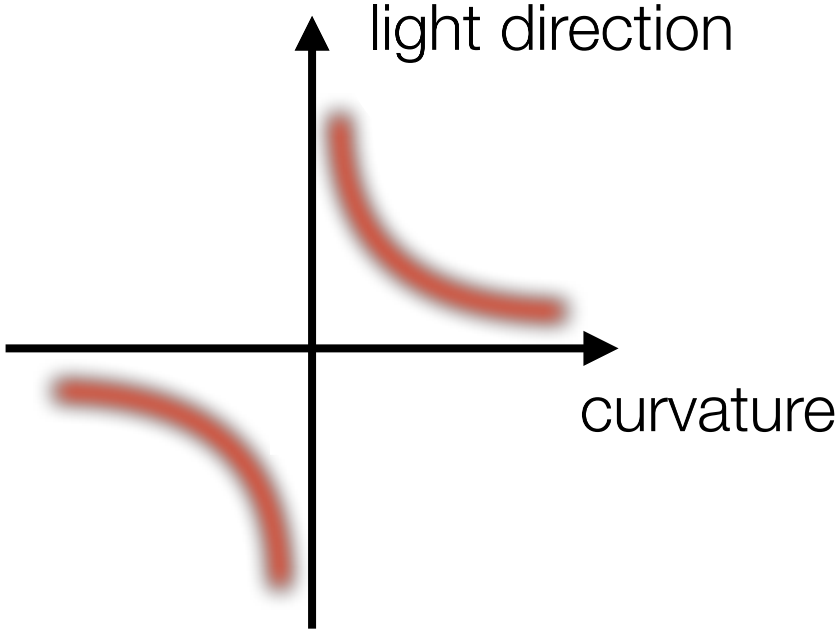
posterior: inference

likelihood: evidence

prior : expectations

RECAP: role of priors





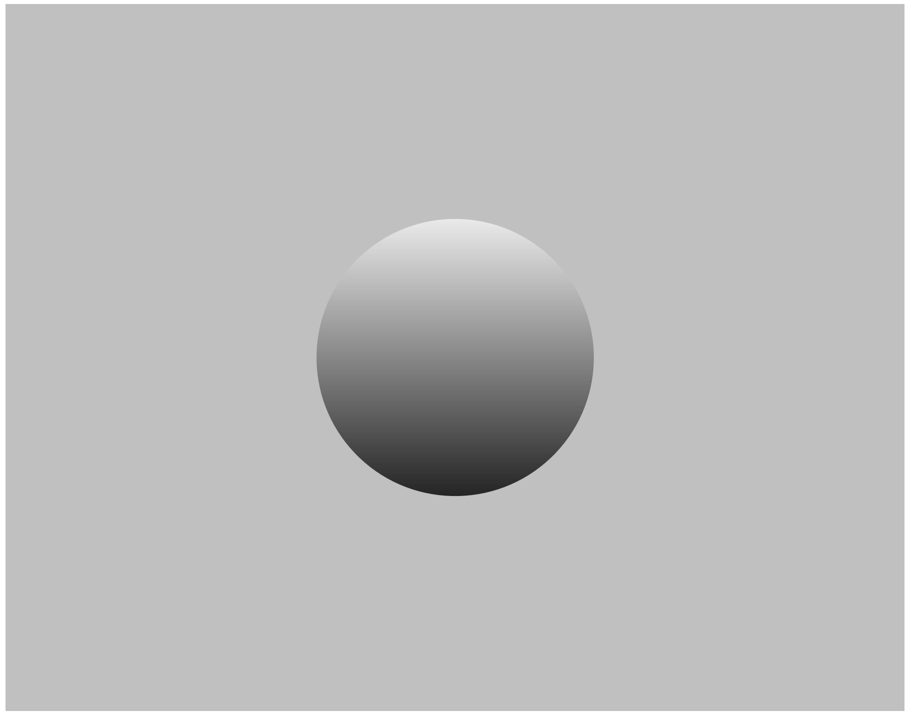
$$P(\text{feature} \mid \text{stimulus}) \propto P(\text{stimulus} \mid \text{feature}) \times P(\text{feature})$$

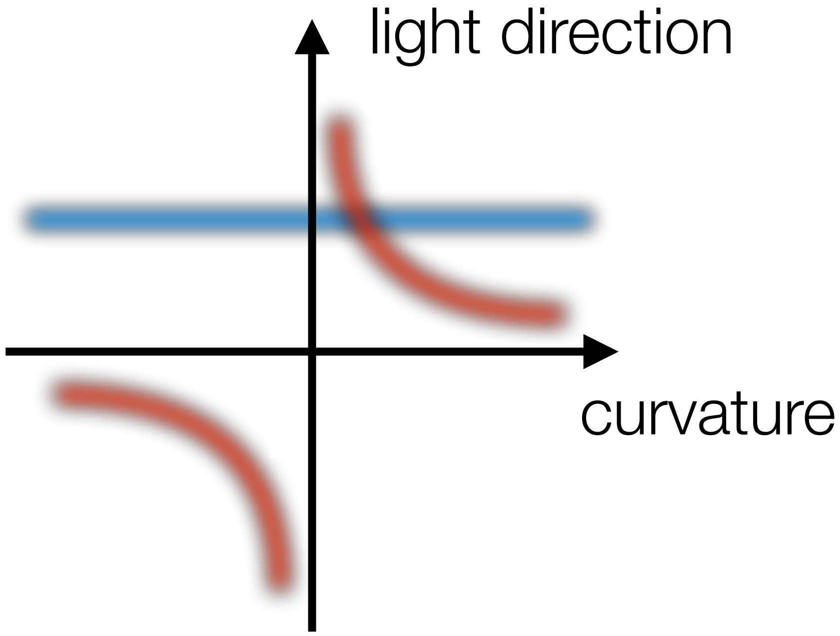
posterior: inference

likelihood: evidence

prior : expectations

RECAP: role of priors





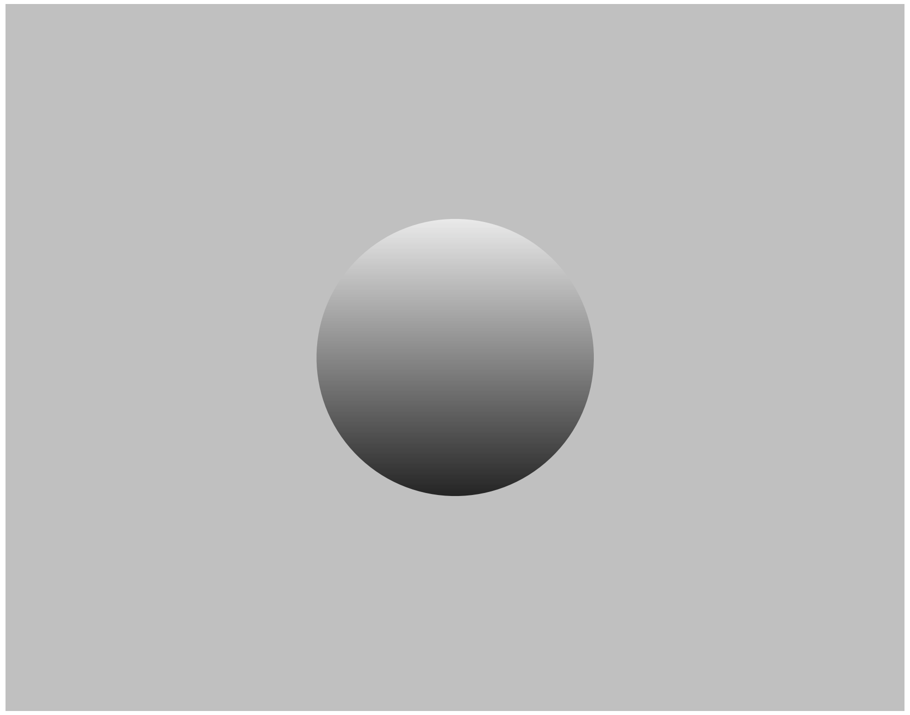
$$\color{green}{P(\text{feature}\,|\,\text{stimulus})} \propto \color{red}{P(\text{stimulus}\,|\,\text{feature})} \times \color{blue}{P(\text{feature})}$$

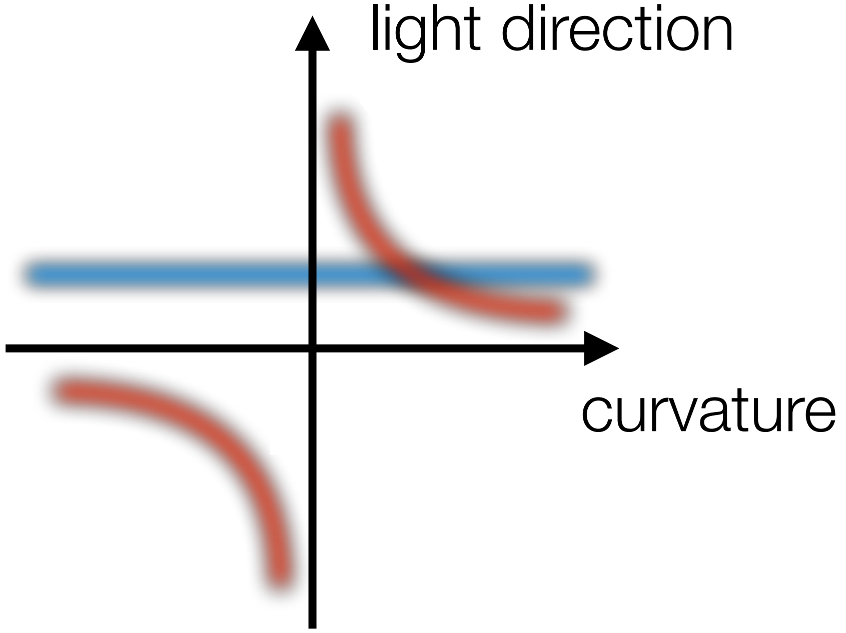
posterior: inference

likelihood: evidence

prior : expectations

RECAP: role of priors





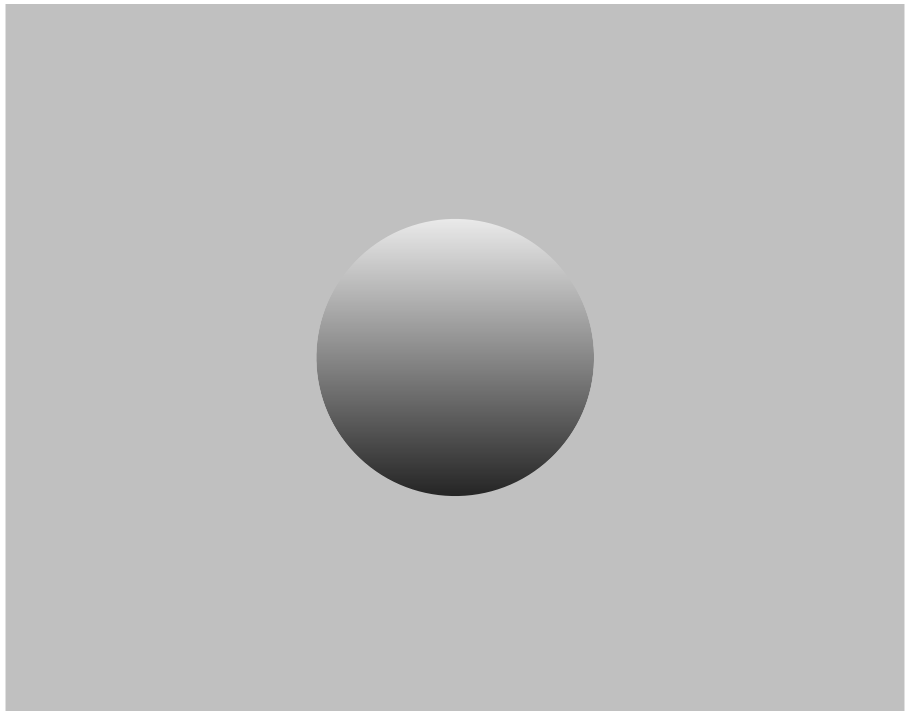
$$P(\text{feature} \mid \text{stimulus}) \propto P(\text{stimulus} \mid \text{feature}) \times P(\text{feature})$$

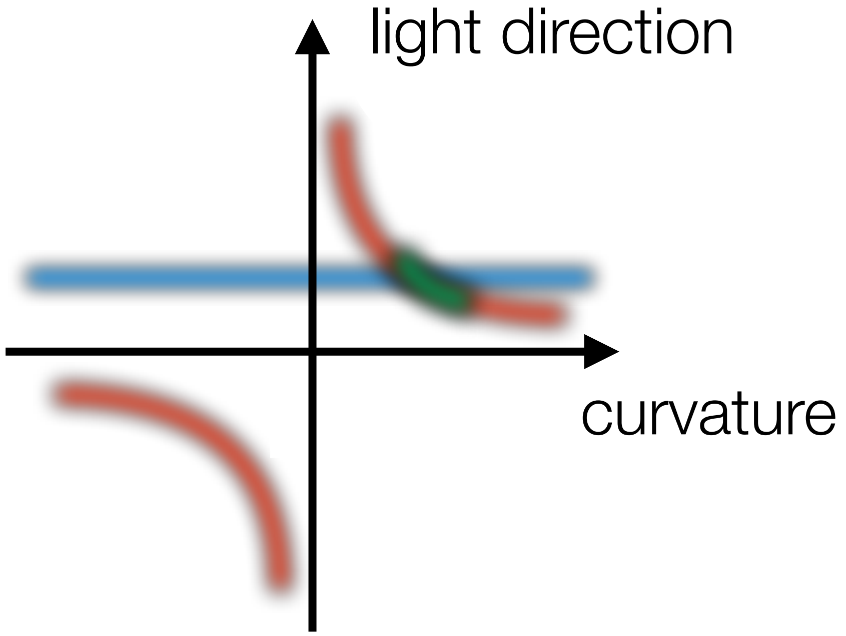
posterior: inference

likelihood: evidence

prior : expectations

RECAP: role of priors





$$P(\text{feature} \mid \text{stimulus}) \propto P(\text{stimulus} \mid \text{feature}) \times P(\text{feature})$$

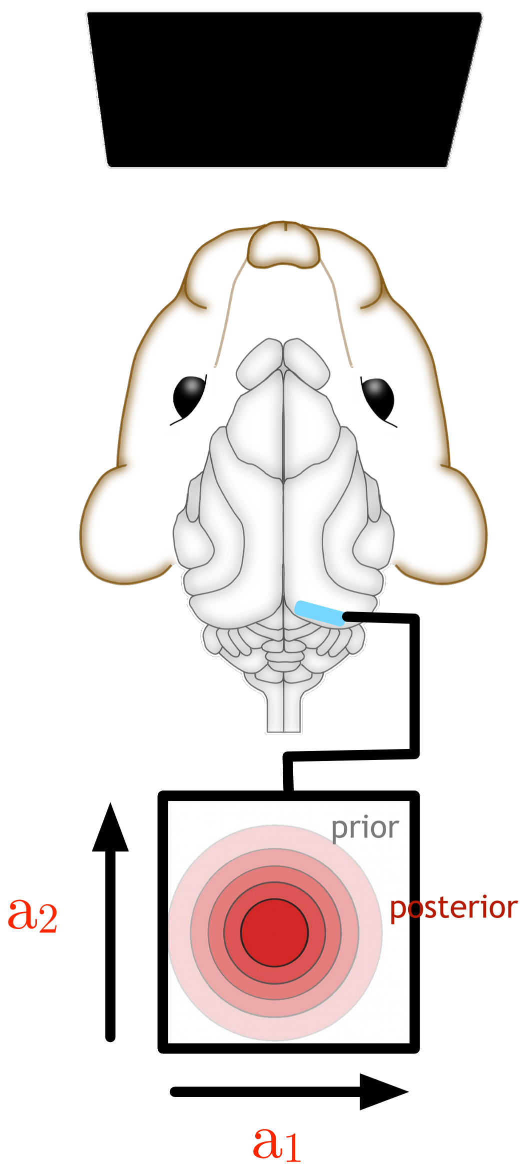
posterior: inference

likelihood: evidence

prior : expectations

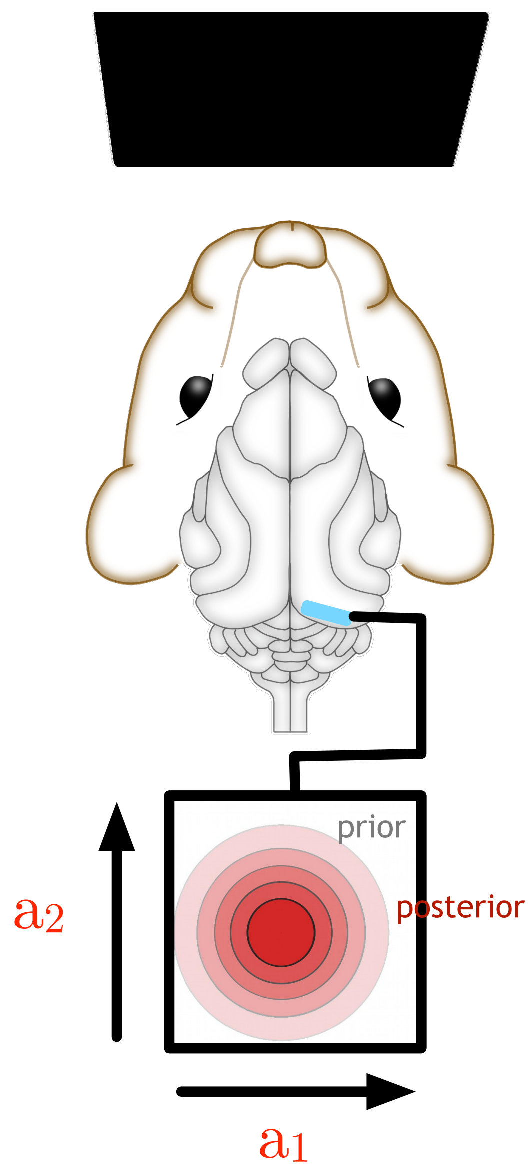
Full response statistics

prior expectations

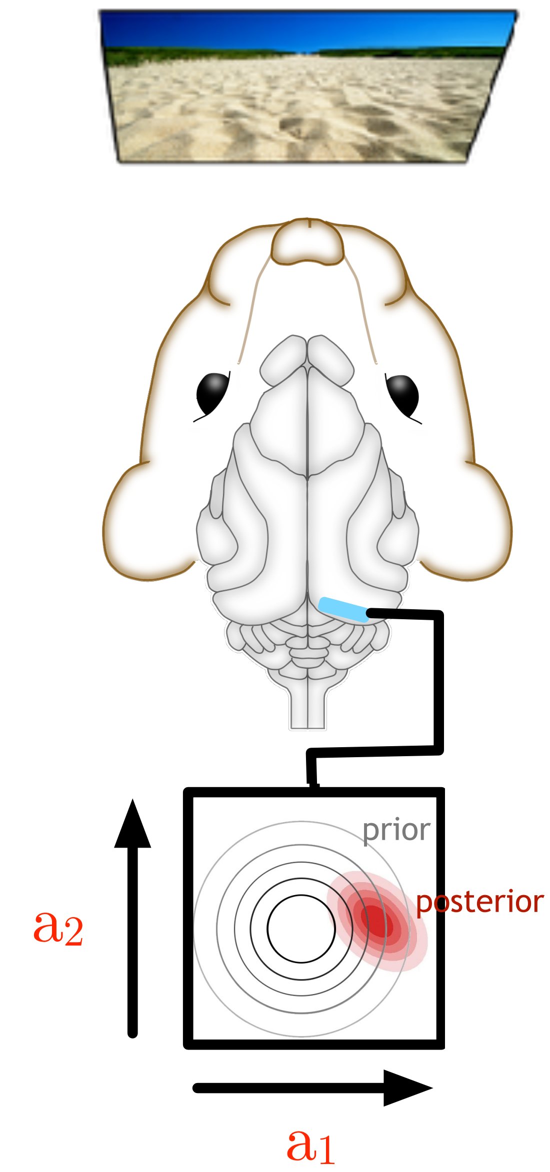


Full response statistics

prior expectations

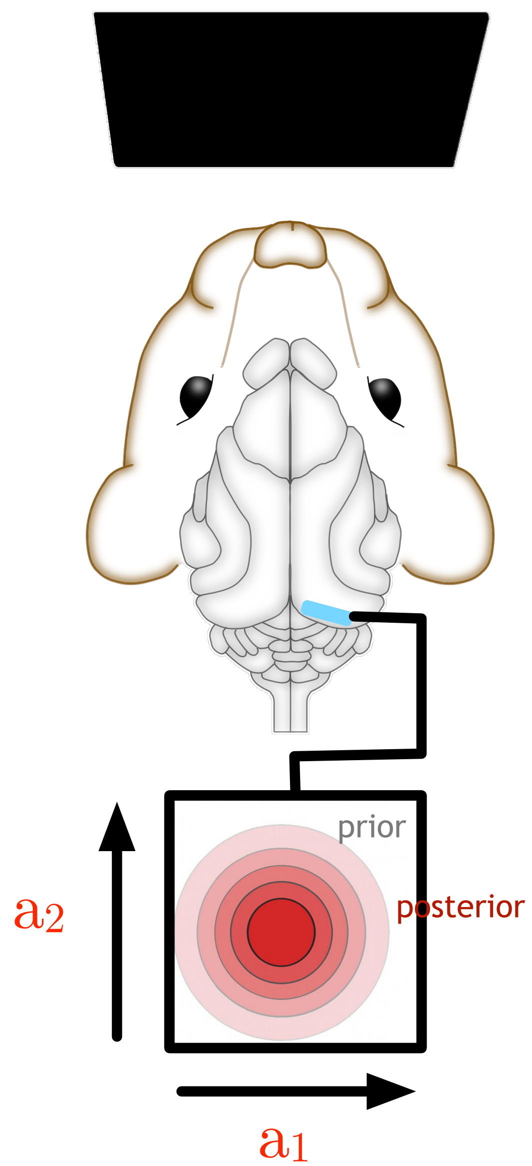


inferences

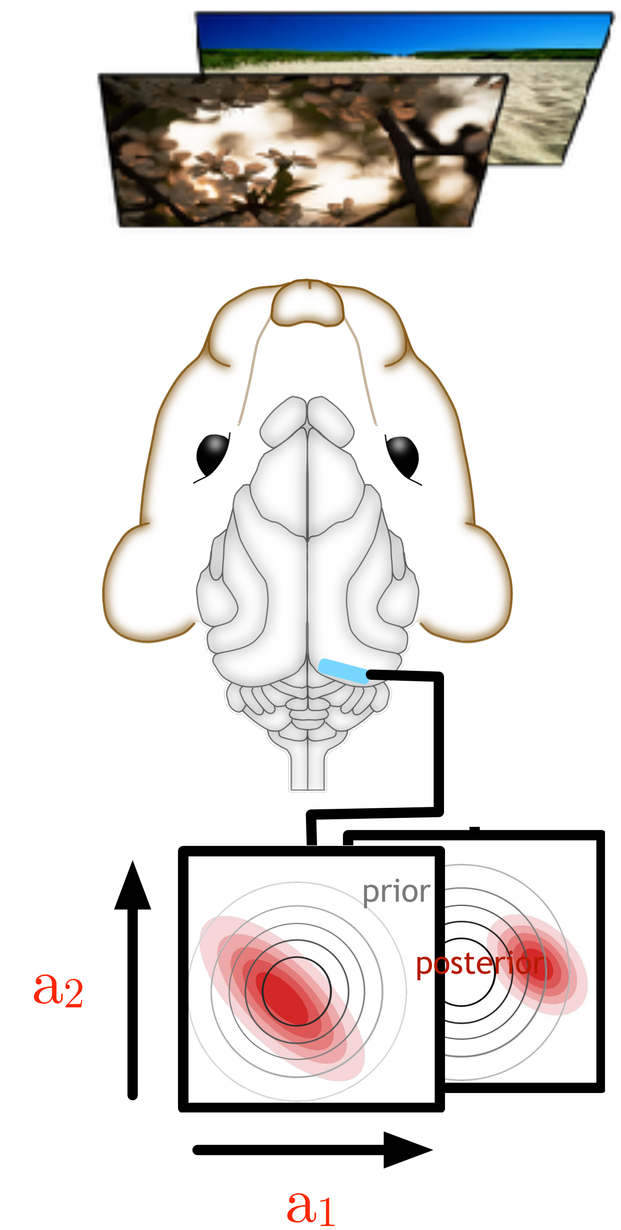


Full response statistics

prior expectations

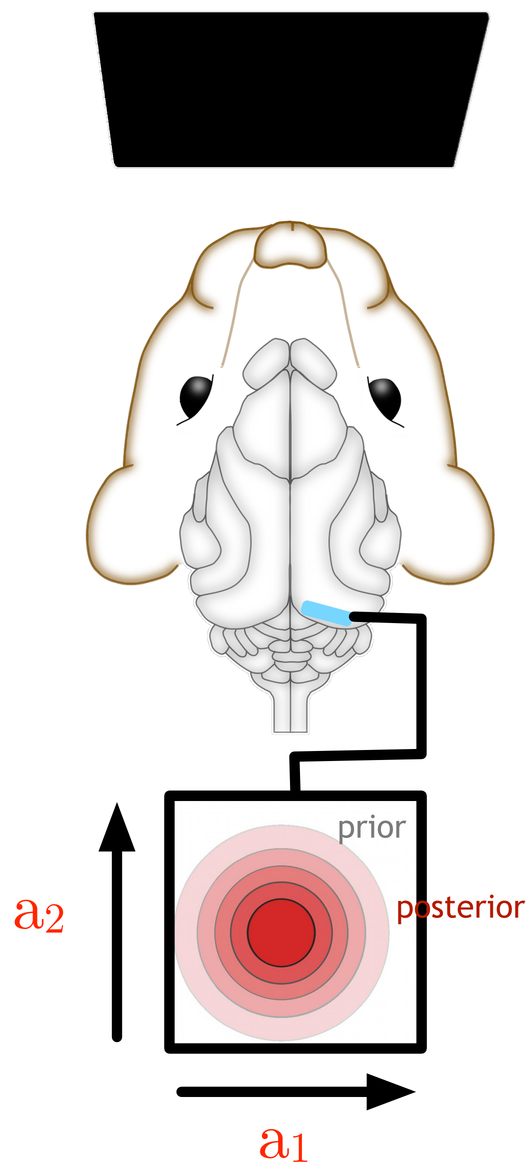


inferences

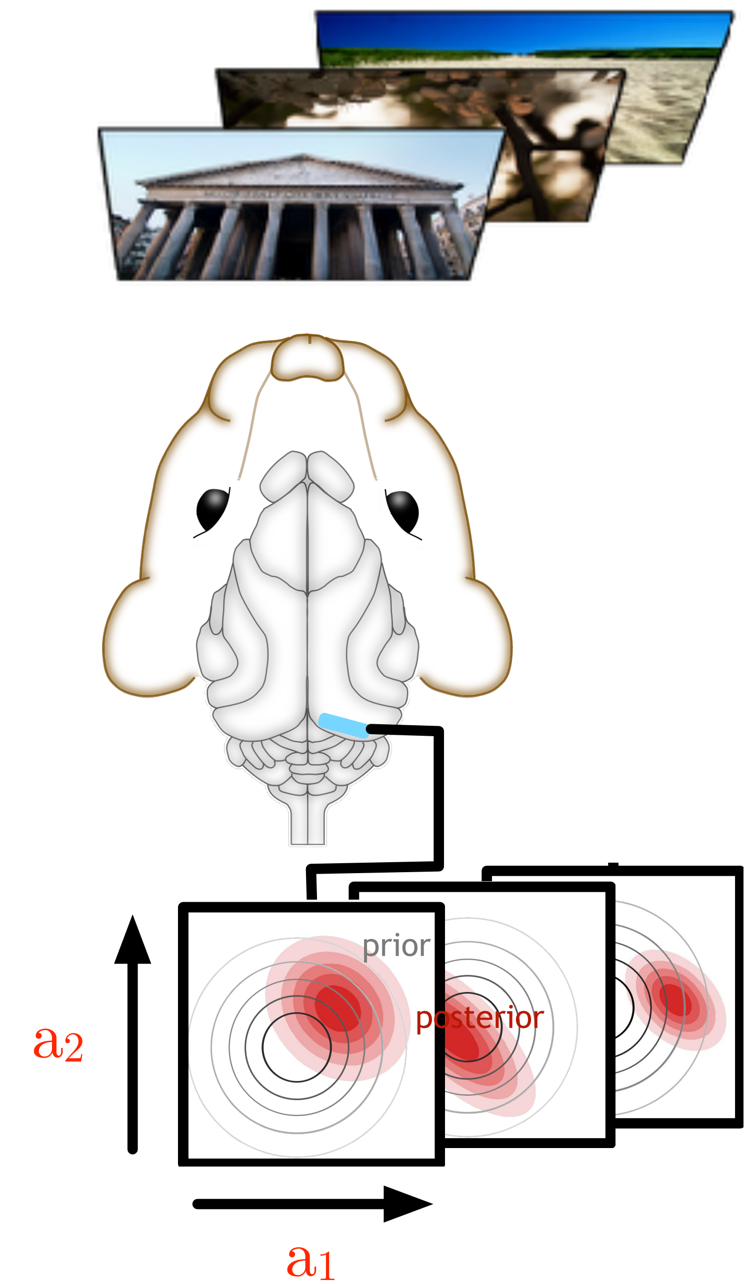


Full response statistics

prior expectations

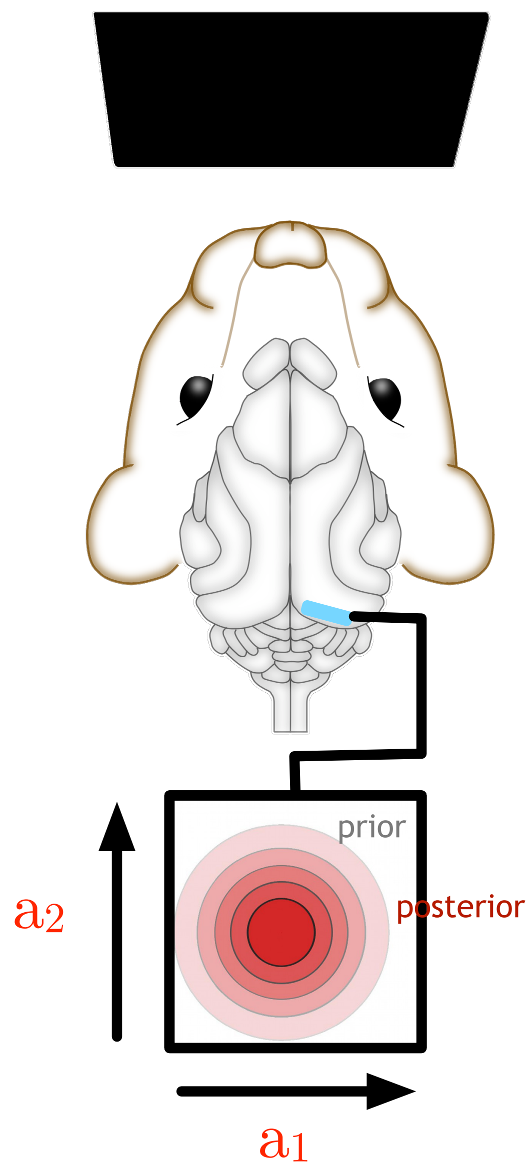


inferences

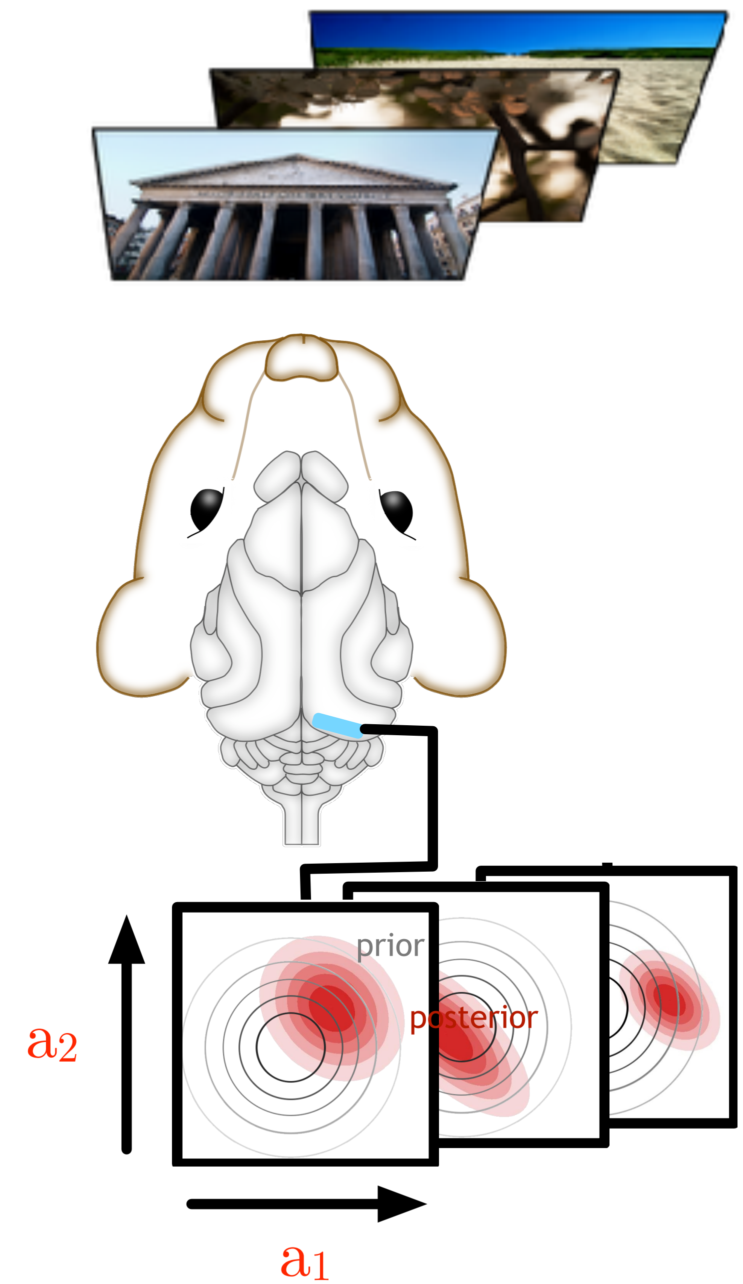


Full response statistics

prior expectations



inferences



spontaneous activity

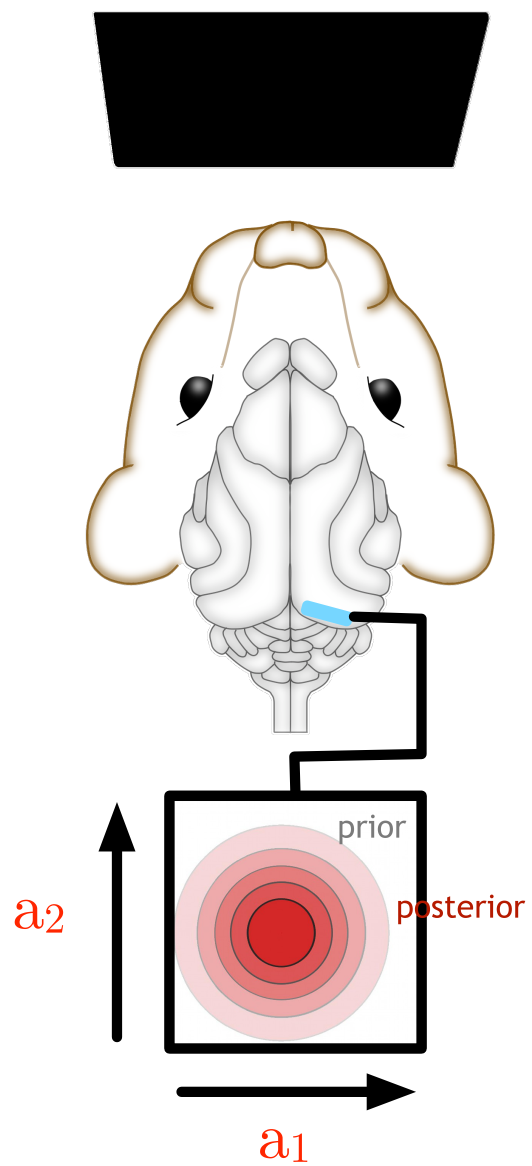
$$P(\mathbf{a})$$

evoked activity

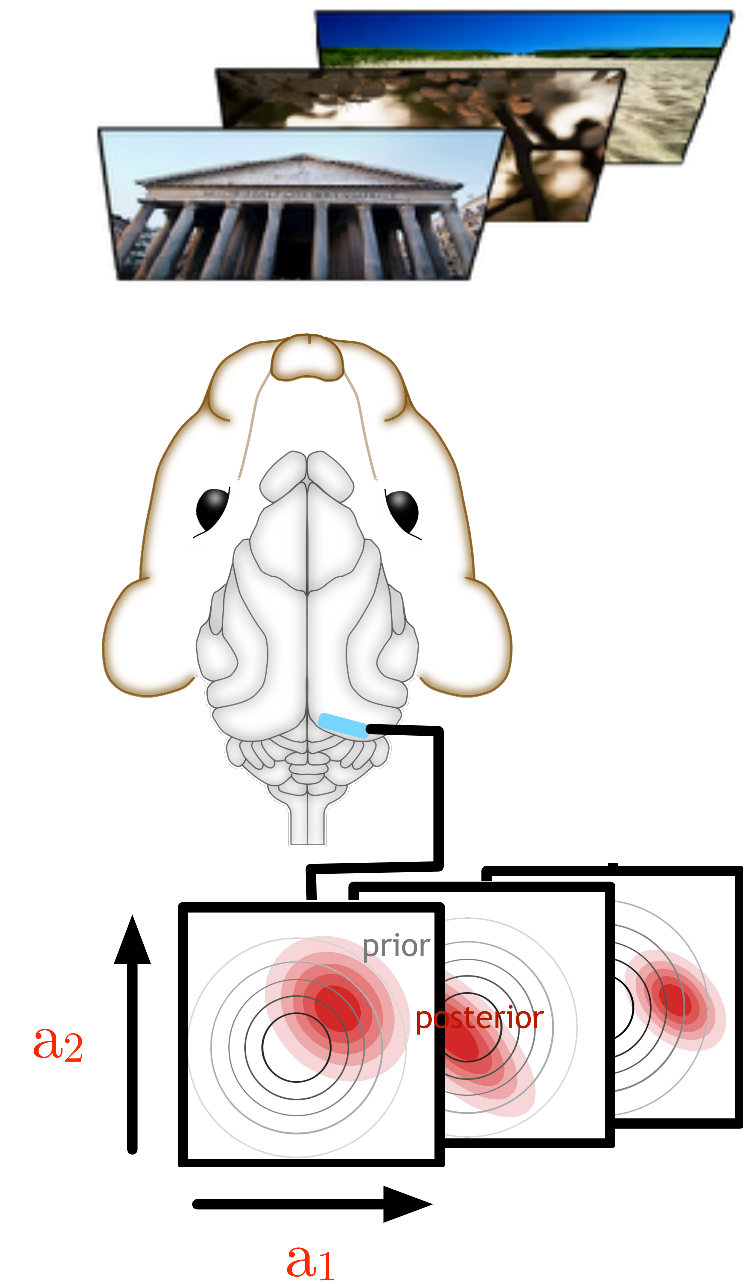
$$P(\mathbf{a} \mid \mathbf{x})$$

Full response statistics

prior expectations



inferences



$$P(\mathbf{a}) = \int dx\, P(\mathbf{a} \,|\, \mathbf{x})\, P(\mathbf{x})$$

spontaneous activity

$P(\mathbf{a})$

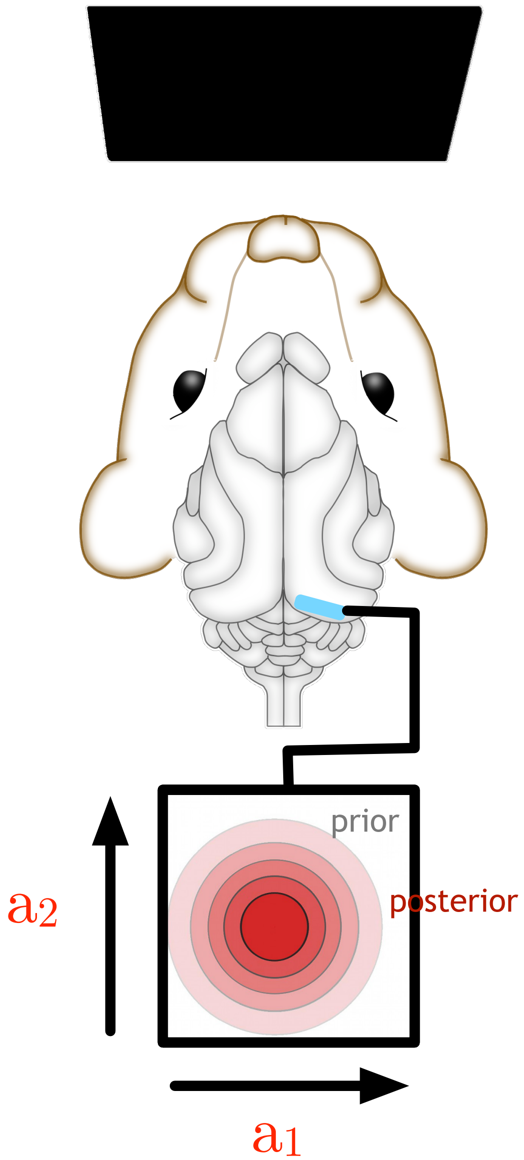
evoked activity

$P(\mathbf{a} \,|\, \mathbf{x})$

Full response statistics

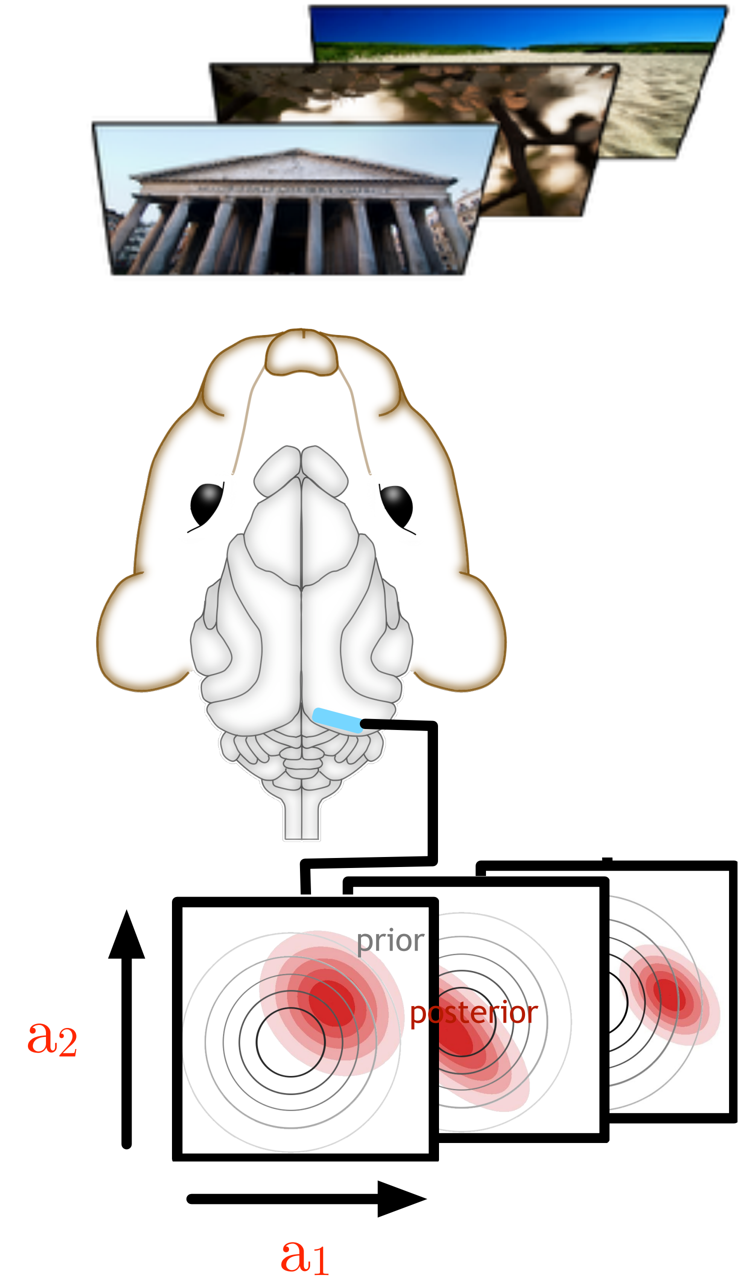
prior expectations

inferences



expectations

$$P(\mathbf{a}) = \int dx\, P(\mathbf{a}\,|\,\mathbf{x})\, P(\mathbf{x})$$



spontaneous activity

$P(\mathbf{a})$

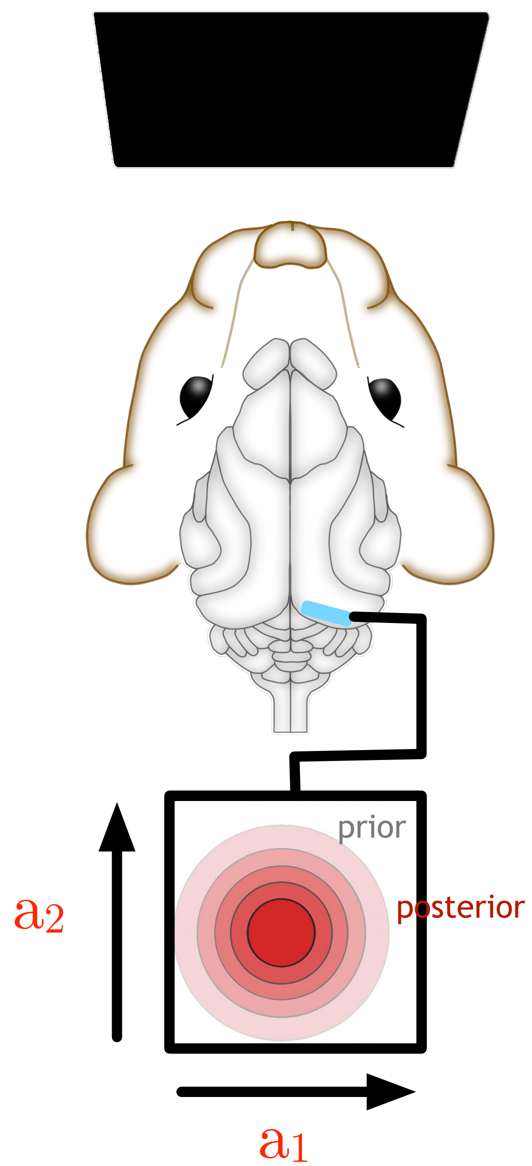
evoked activity

$P(\mathbf{a}\,|\,\mathbf{x})$

Full response statistics

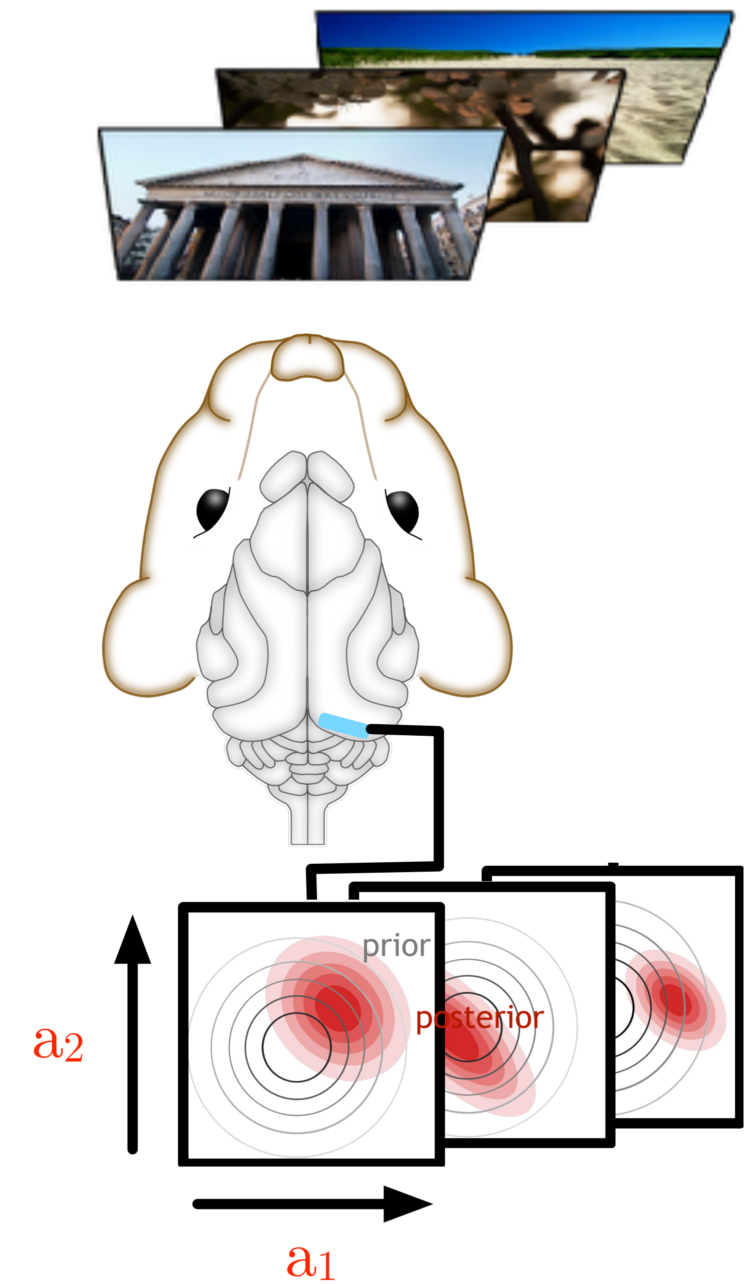
prior expectations

inferences



expectations *inference*

$$P(\mathbf{a}) = \int dx\, P(\mathbf{a}\,|\,\mathbf{x})\, P(\mathbf{x})$$



spontaneous activity

$P(\mathbf{a})$

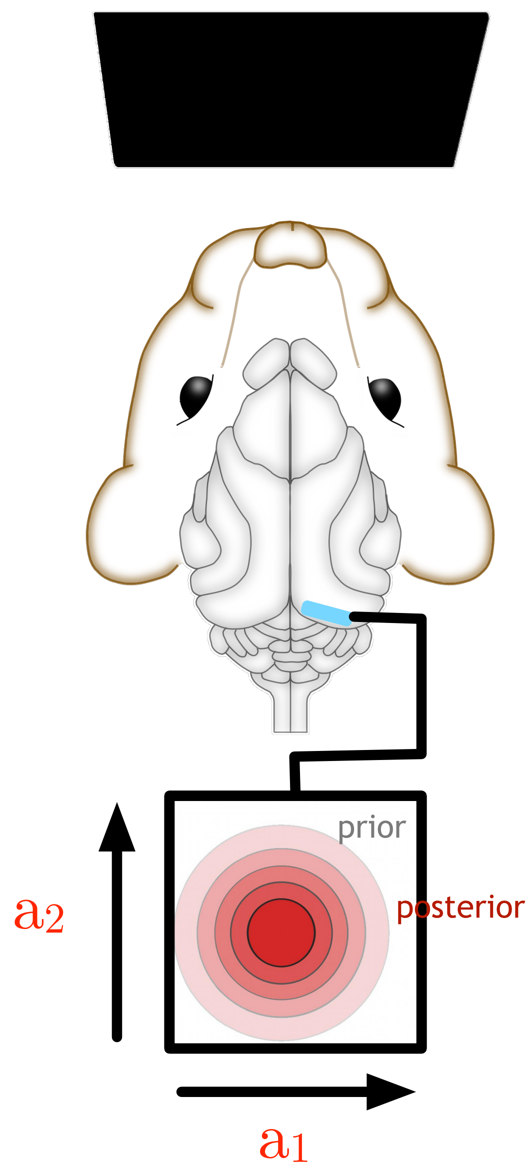
evoked activity

$P(\mathbf{a}\,|\,\mathbf{x})$

Full response statistics

prior expectations

inferences

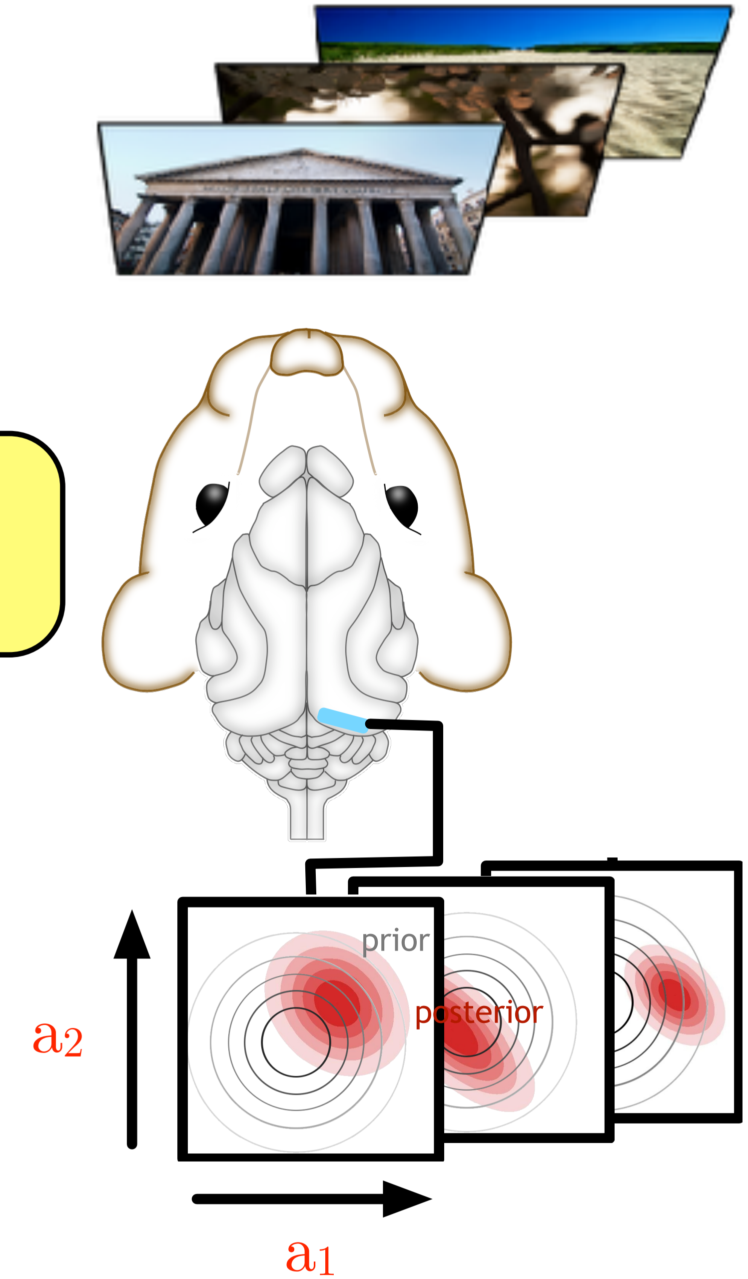


expectations

inference

stimulus statistics

$$P(\mathbf{a}) = \int dx\, P(\mathbf{a}\,|\,\mathbf{x})\, P(\mathbf{x})$$



spontaneous activity

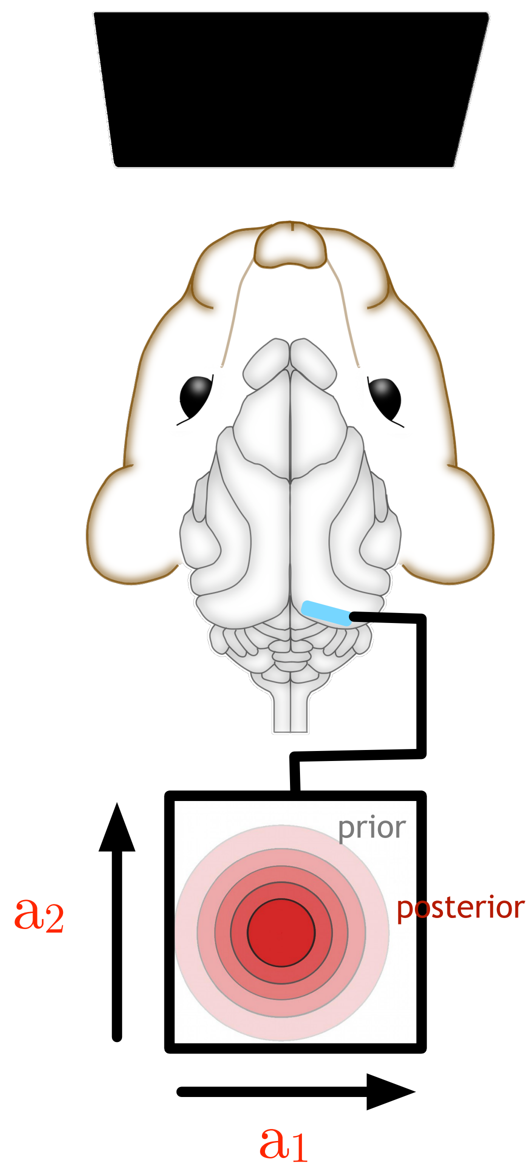
$P(\mathbf{a})$

evoked activity

$P(\mathbf{a}\,|\,\mathbf{x})$

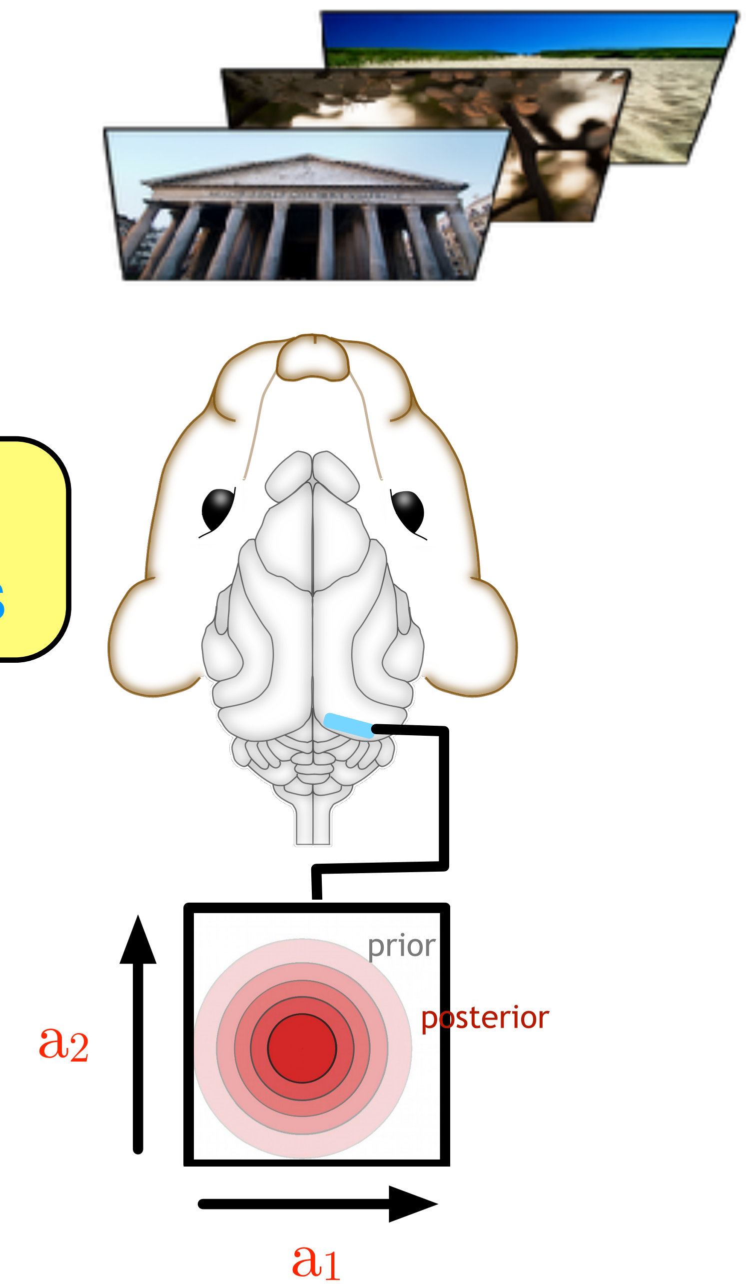
Full response statistics

prior expectations | *average inferences*



expectations | *inference* | stimulus statistics

$$P(\mathbf{a}) = \int dx\, P(\mathbf{a}\,|\,\mathbf{x})\, P(\mathbf{x})$$



spontaneous activity

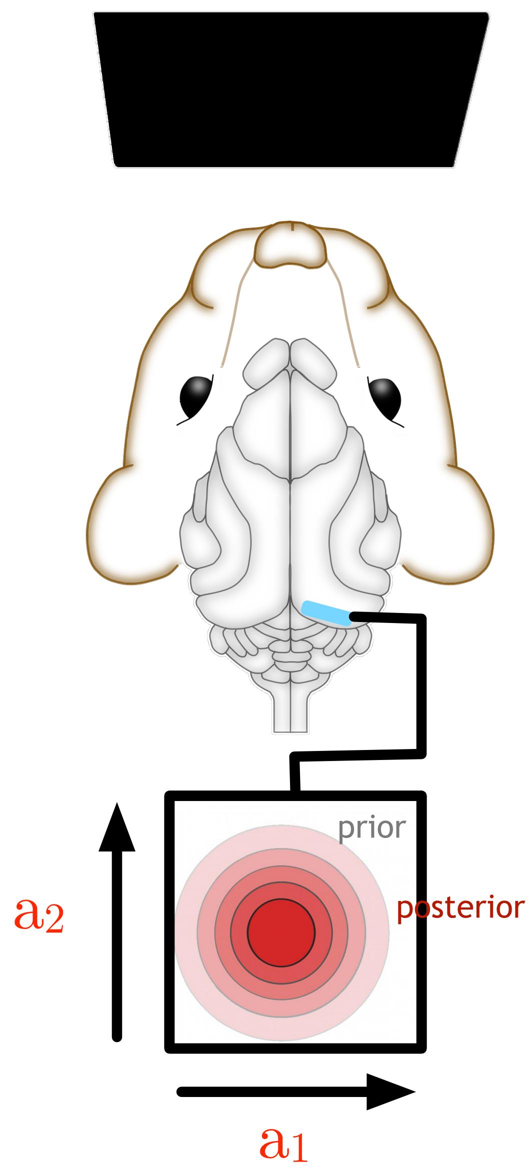
$P(\mathbf{a})$

evoked activity

$P(\mathbf{a}\,|\,\mathbf{x})$

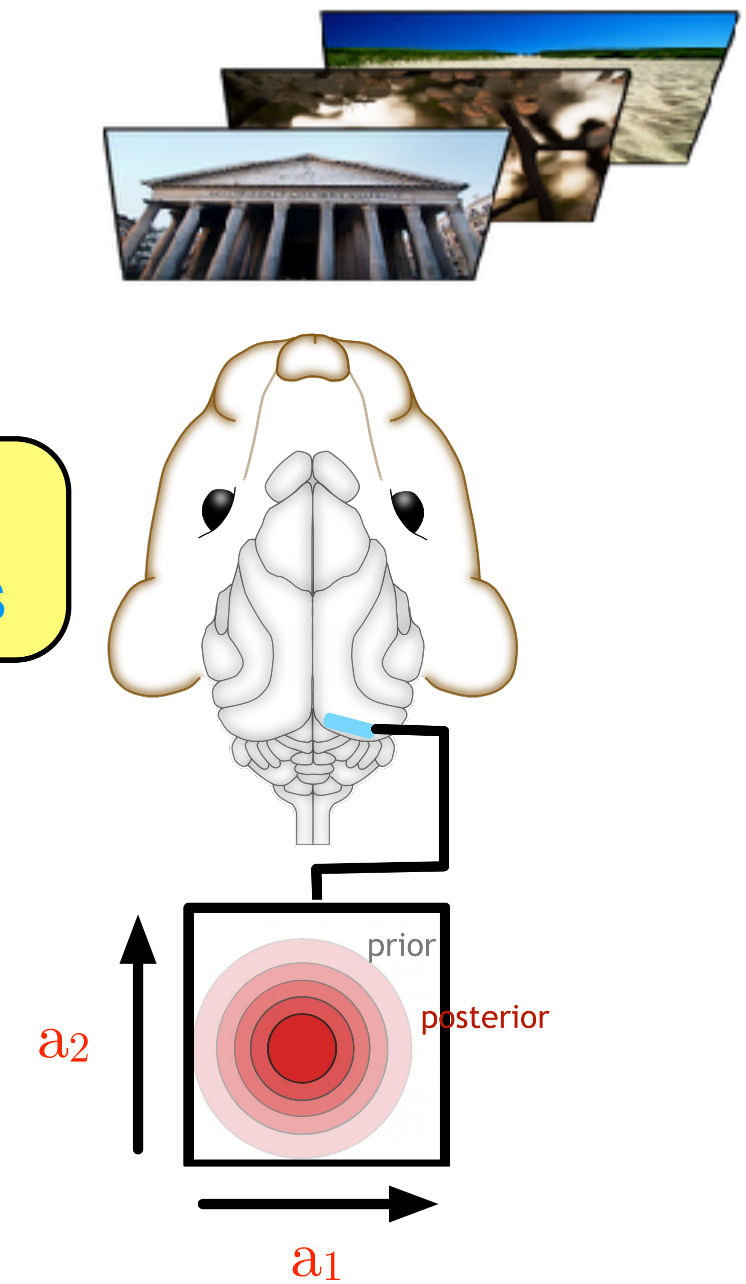
Full response statistics

prior expectations *average inferences*



expectations *inference* stimulus statistics

$$P(\mathbf{a}) = \int dx\, P(\mathbf{a}\,|\,\mathbf{x})\, P(\mathbf{x})$$

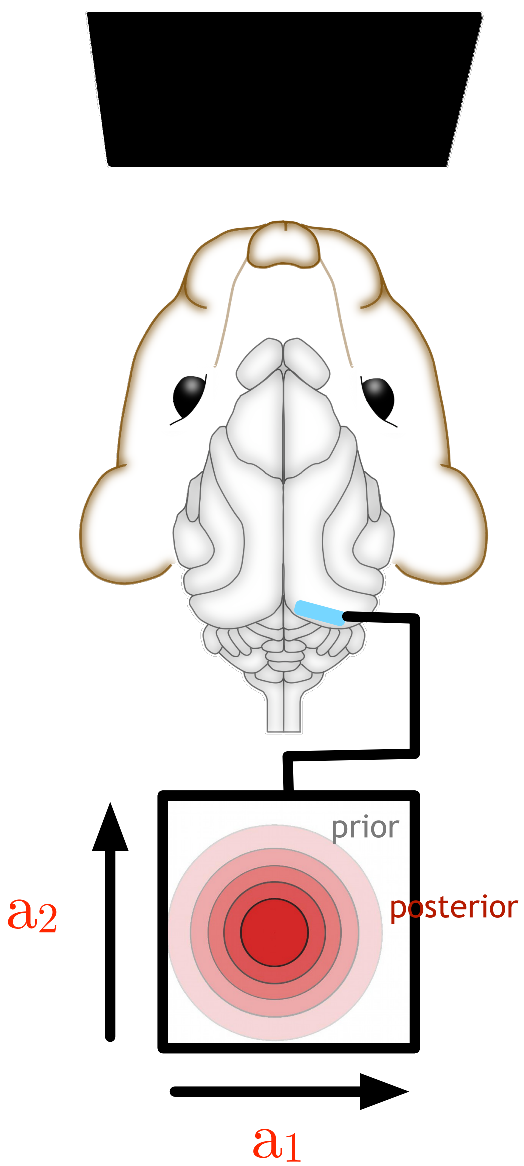


spontaneous activity $\stackrel{?}{=}$ average evoked activity

$P(\mathbf{a})$ $P(\mathbf{a}\,|\,\mathbf{x})$

Full response statistics

prior expectations *average inferences*



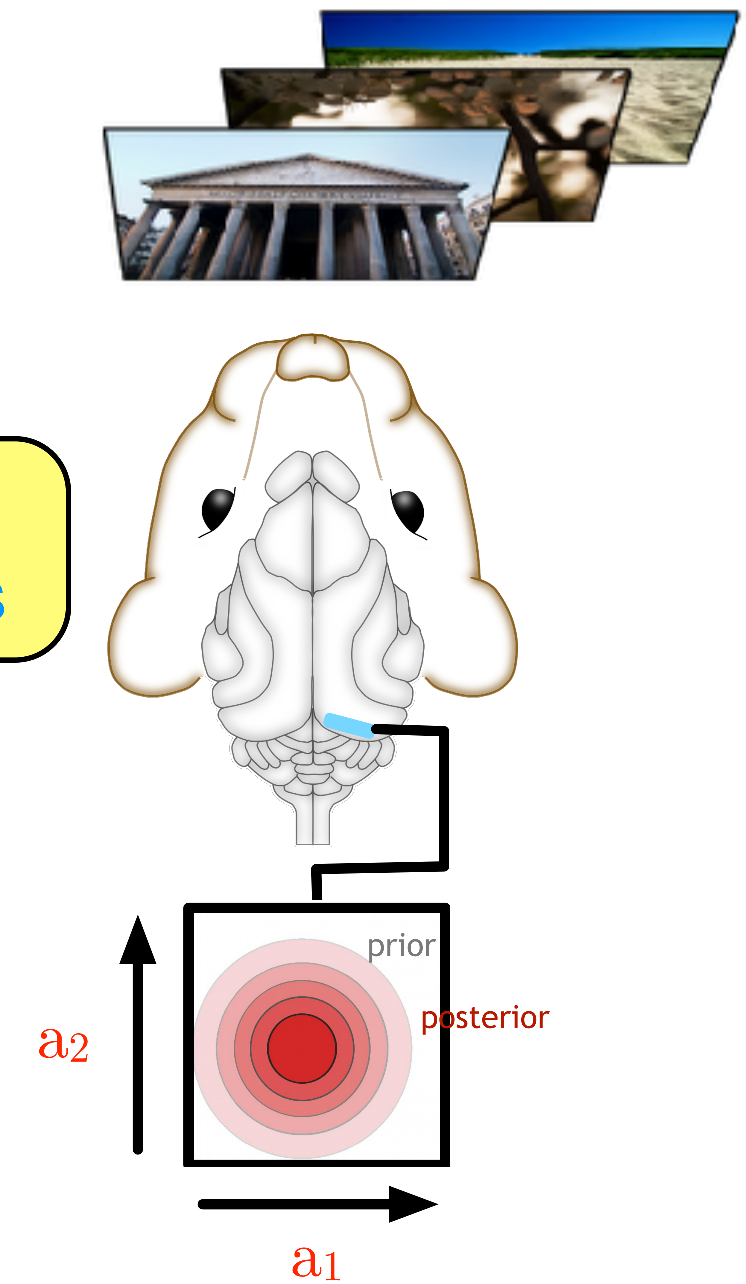
expectations

inference

stimulus statistics

$$P(\mathbf{a}) = \boxed{\int dx\, P(\mathbf{a}\,|\,\mathbf{x})\, P(\mathbf{x})}$$

$$\langle P(\mathbf{a}\,|\,\mathbf{x})\rangle_{P(\mathbf{x})}$$



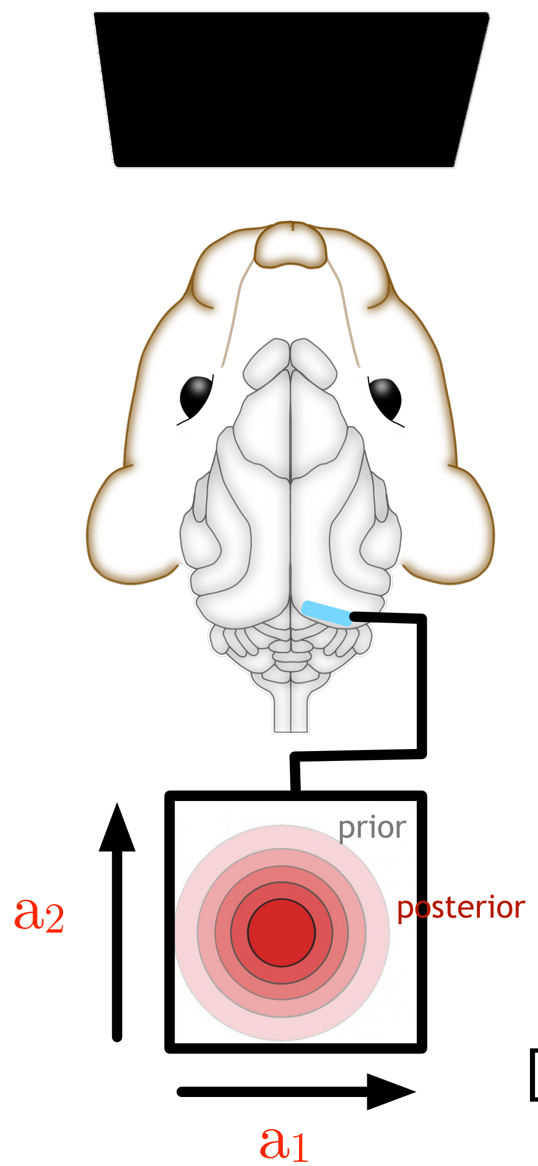
spontaneous activity $\stackrel{?}{=}$ average evoked activity

$P(\mathbf{a})$ $P(\mathbf{a}\,|\,\mathbf{x})$

Full response statistics

prior expectations

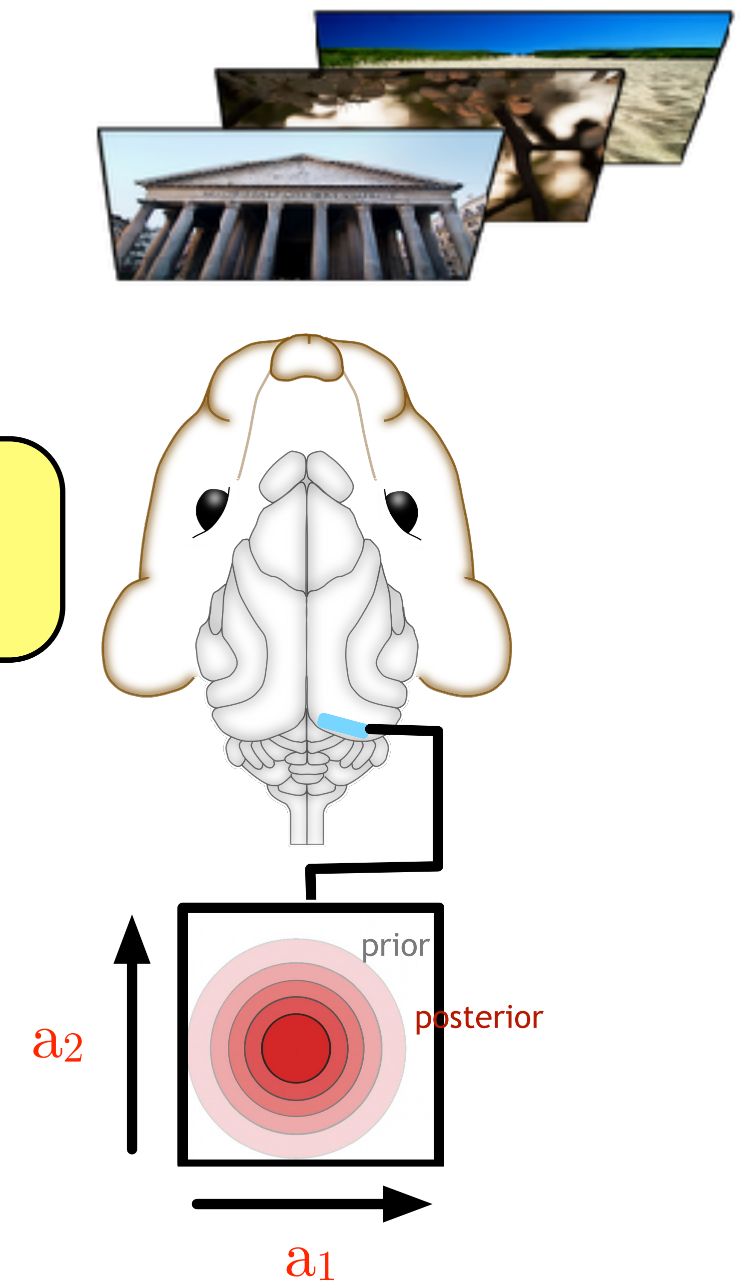
average inferences



expectations *inference* stimulus statistics

$$P(\mathbf{a}) = \boxed{\int dx\, P(\mathbf{a}\,|\,\mathbf{x})\, P(\mathbf{x})}$$

Dissimilarity: $\mathrm{KL}[P(\mathbf{a})\ ||\langle P(\mathbf{a}\,|\,\mathbf{x})\rangle_{P(\mathbf{x})}]$

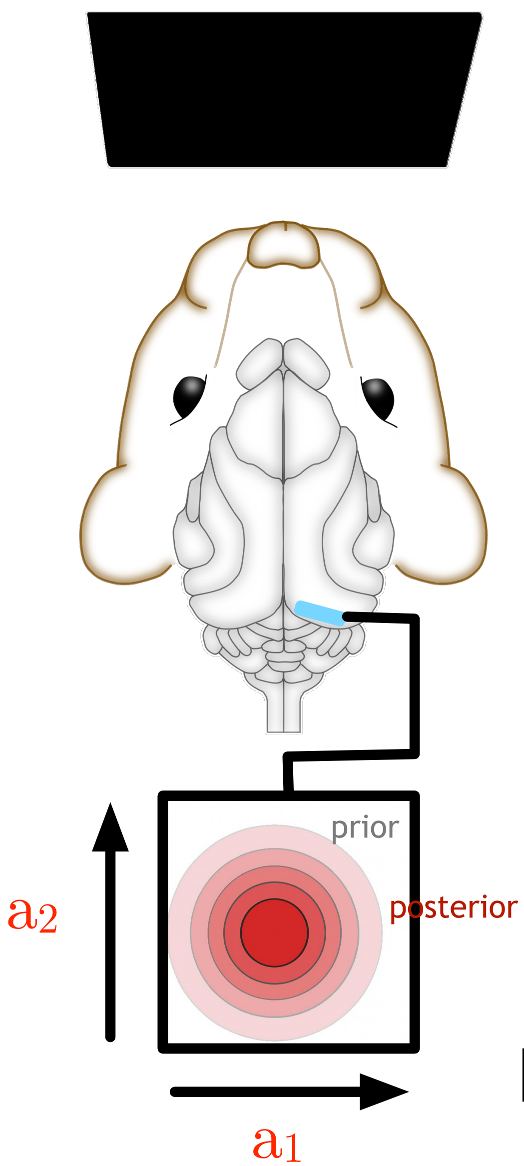


spontaneous activity $\stackrel{?}{=}$ average evoked activity

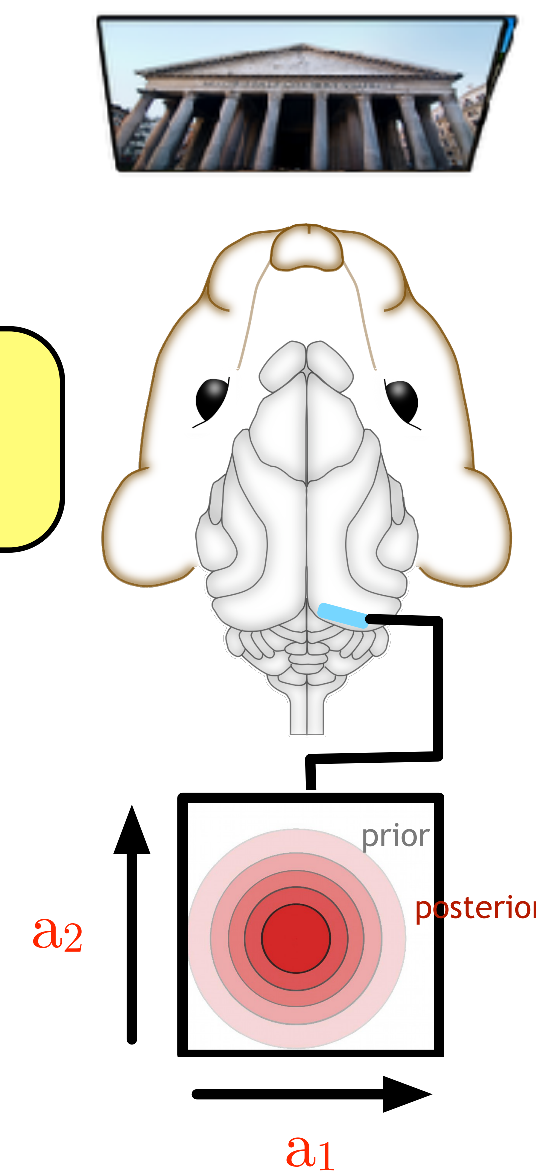
$P(\mathbf{a})$ $P(\mathbf{a}\,|\,\mathbf{x})$

Full response statistics

prior expectations



average inferences



stimulus statistics

expectations

inference

$$P(\mathbf{a}) = \boxed{\int dx\, P(\mathbf{a}\,|\,\mathbf{x})\, P(\mathbf{x})}$$

Dissimilarity: $\mathrm{KL}[P(\mathbf{a})\, ||\, \langle P(\mathbf{a}\,|\,\mathbf{x})\rangle_{P(\mathbf{x})}]$

spontaneous activity $\overset{?}{=}$ average evoked activity

$P(\mathbf{a})$ $\qquad$ $P(\mathbf{a}\,|\,\mathbf{x})$

Full response statistics

$$P(\mathbf{a}) = \int dx\, P(\mathbf{a}\,|\,\mathbf{x})\, P(\mathbf{x})$$

spontaneous activity $\stackrel{?}{=}$ average evoked activity

$P(\mathbf{a})$ $P(\mathbf{a}\,|\,\mathbf{x})$

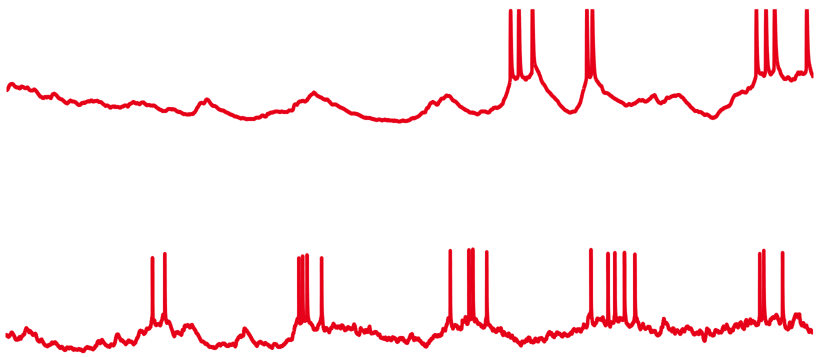
Full response statistics

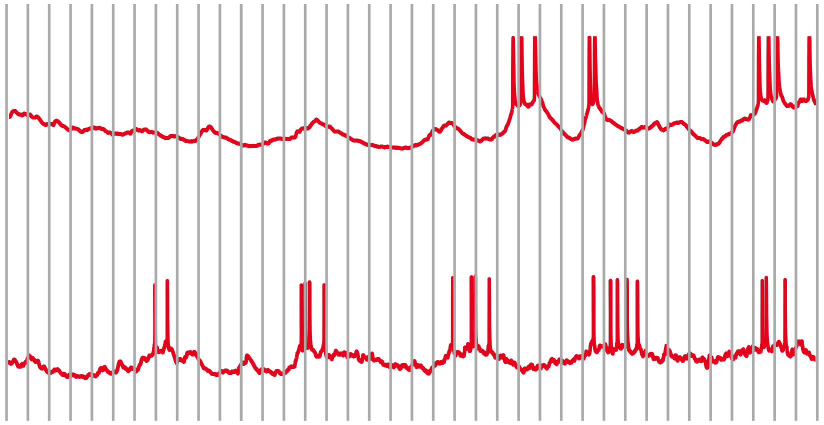
- ★ the model has been adapted to the appropriate model of the world
- ★ the stimulus statistics tested is appropriate

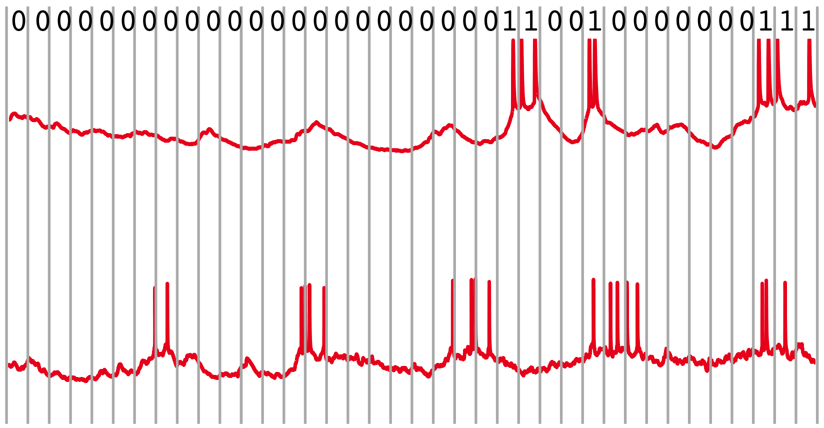
$$P(\mathbf{a}) = \int dx\, P(\mathbf{a} \,|\, \mathbf{x})\, P(\mathbf{x})$$

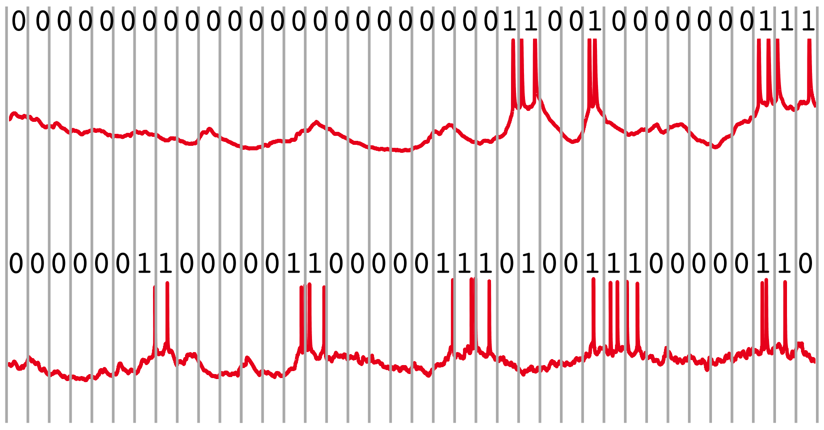
spontaneous activity $\stackrel{?}{=}$ average evoked activity

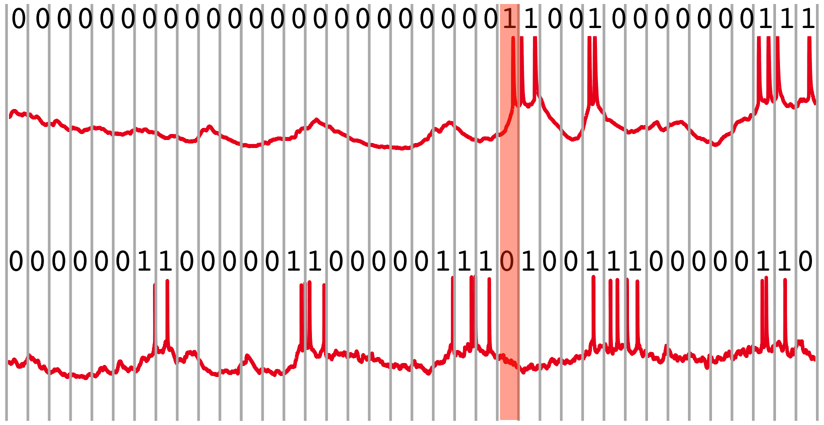
$P(\mathbf{a})$ $\quad$ $P(\mathbf{a} \,|\, \mathbf{x})$

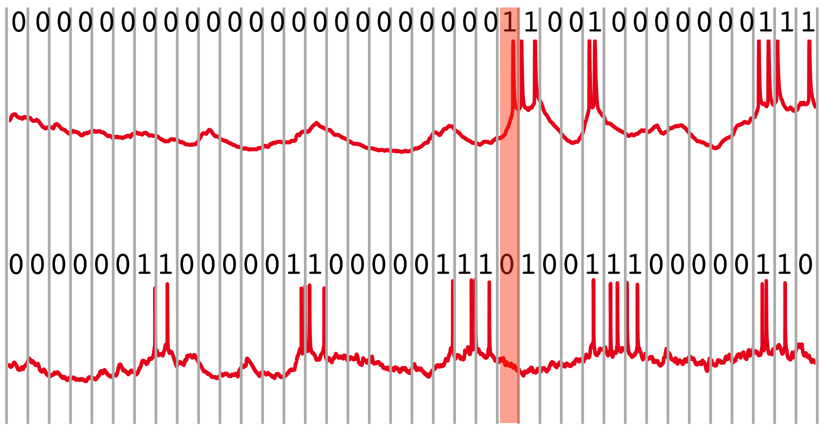


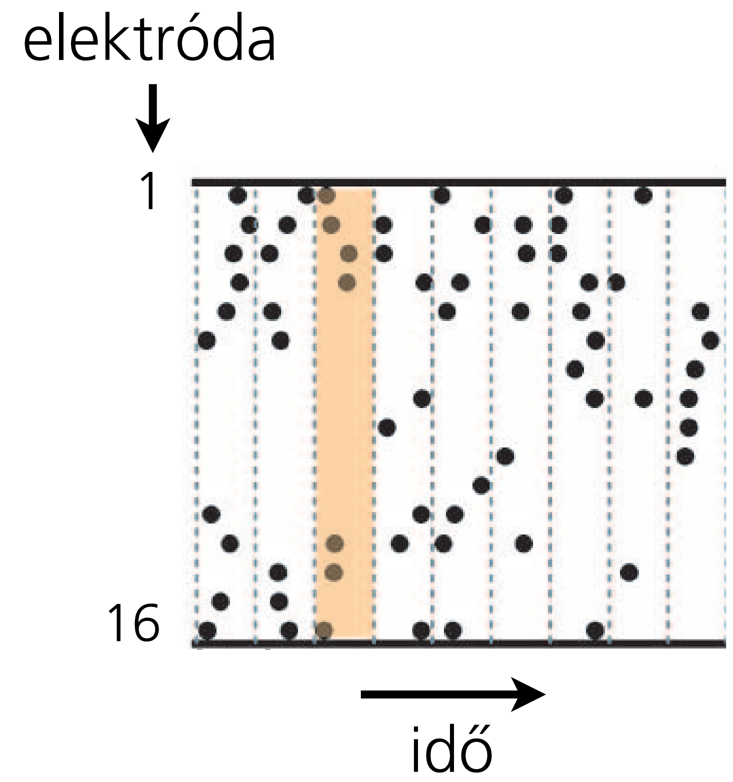


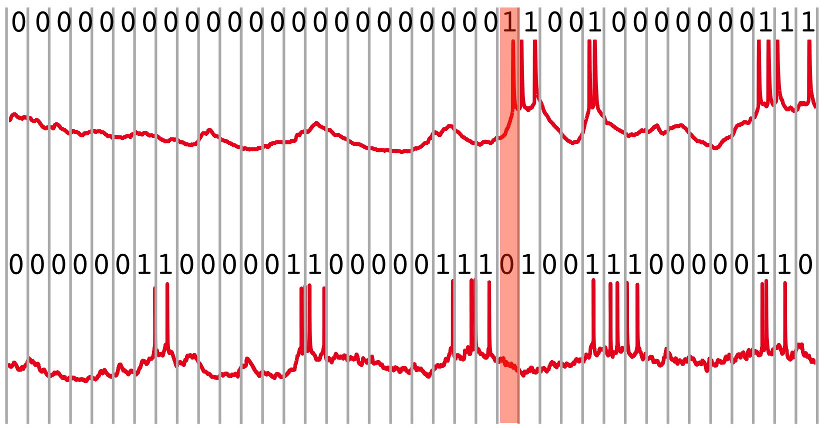

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 1 0 0 0 0 0 0 0 1 1 1

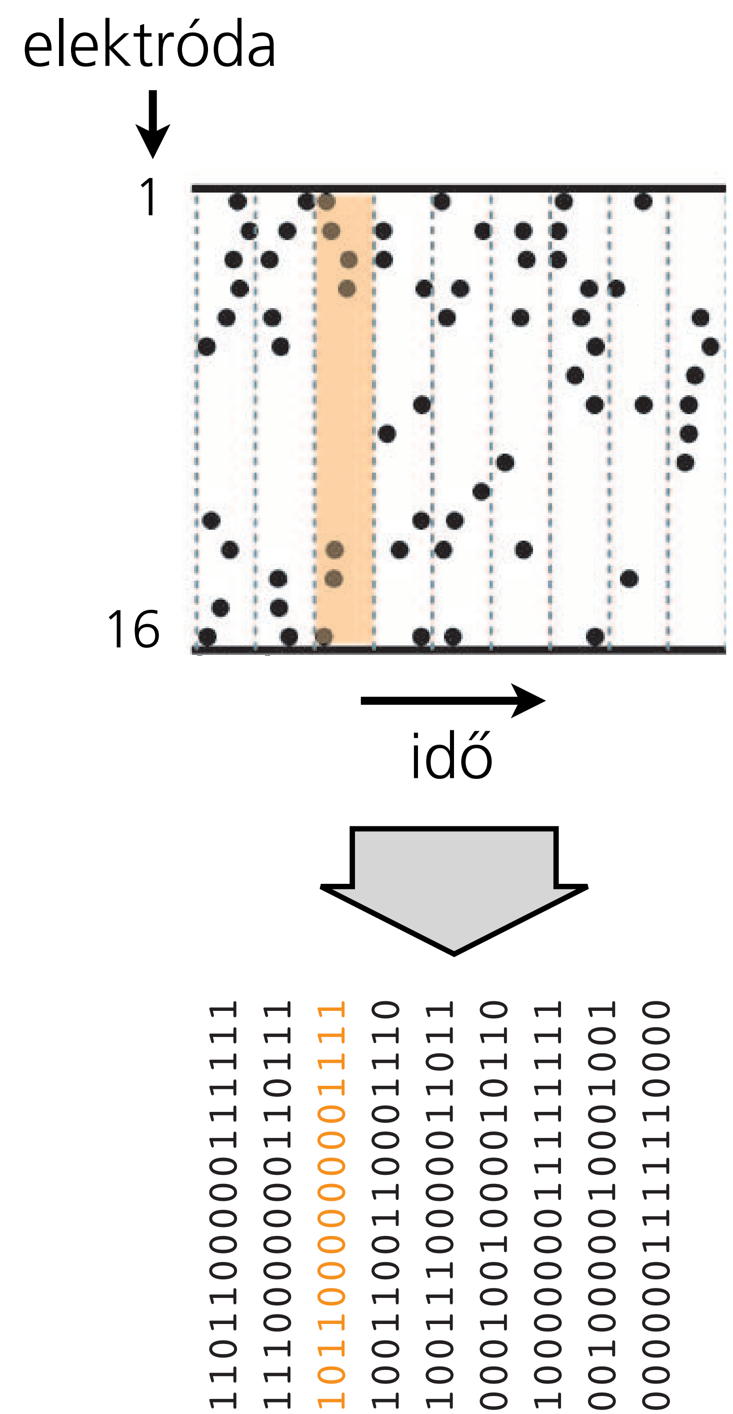

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 1 0 0 0 0 0 0 0 1 1 1
0 0 0 0 0 0 1 1 0 0 0 0 0 1 1 0 0 0 0 0 1 1 1 0 1 0 0 1 1 1 0 0 0 0 0 1 1 0

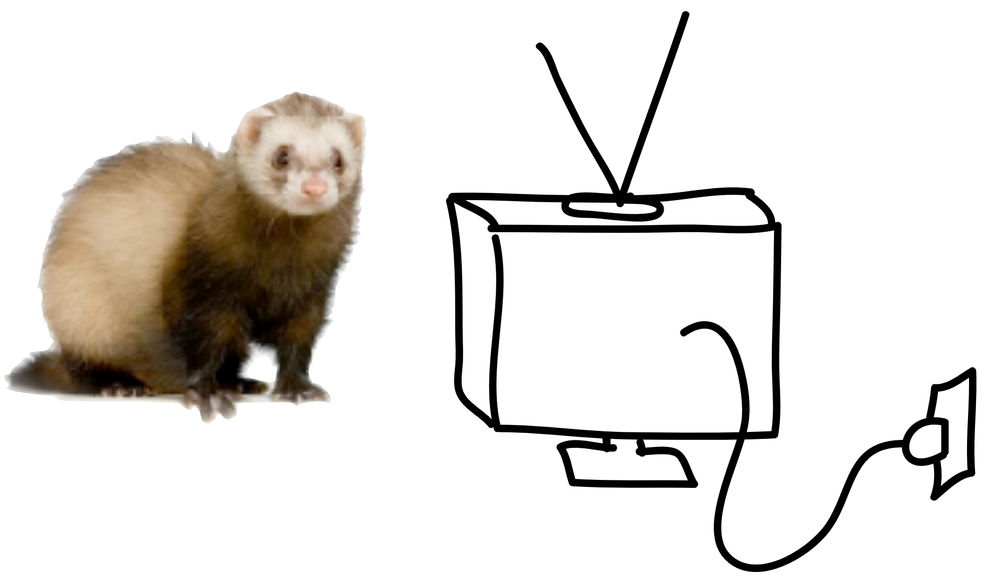

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 1 0 0 0 0 0 0 0 1 1 1
0 0 0 0 0 0 1 1 0 0 0 0 0 1 1 0 0 0 0 0 1 1 1 0 1 0 0 1 1 1 0 0 0 0 0 1 1 0

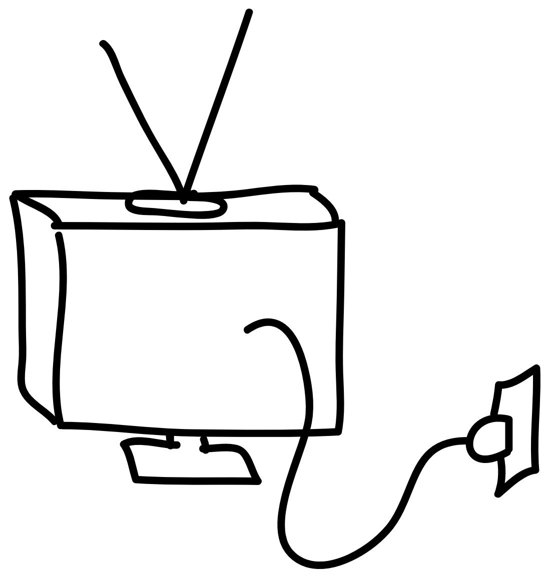

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 1 0 0 0 0 0 0 0 1 1 1
0 0 0 0 0 0 1 1 0 0 0 0 0 1 1 0 0 0 0 0 1 1 1 0 1 0 0 1 1 1 0 0 0 0 0 1 1 0


elektróda
1
16
idő




elektróda
1
16
idő





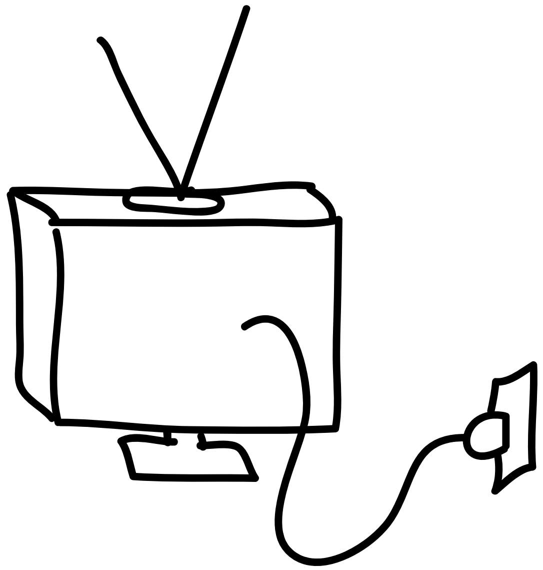
elektróda

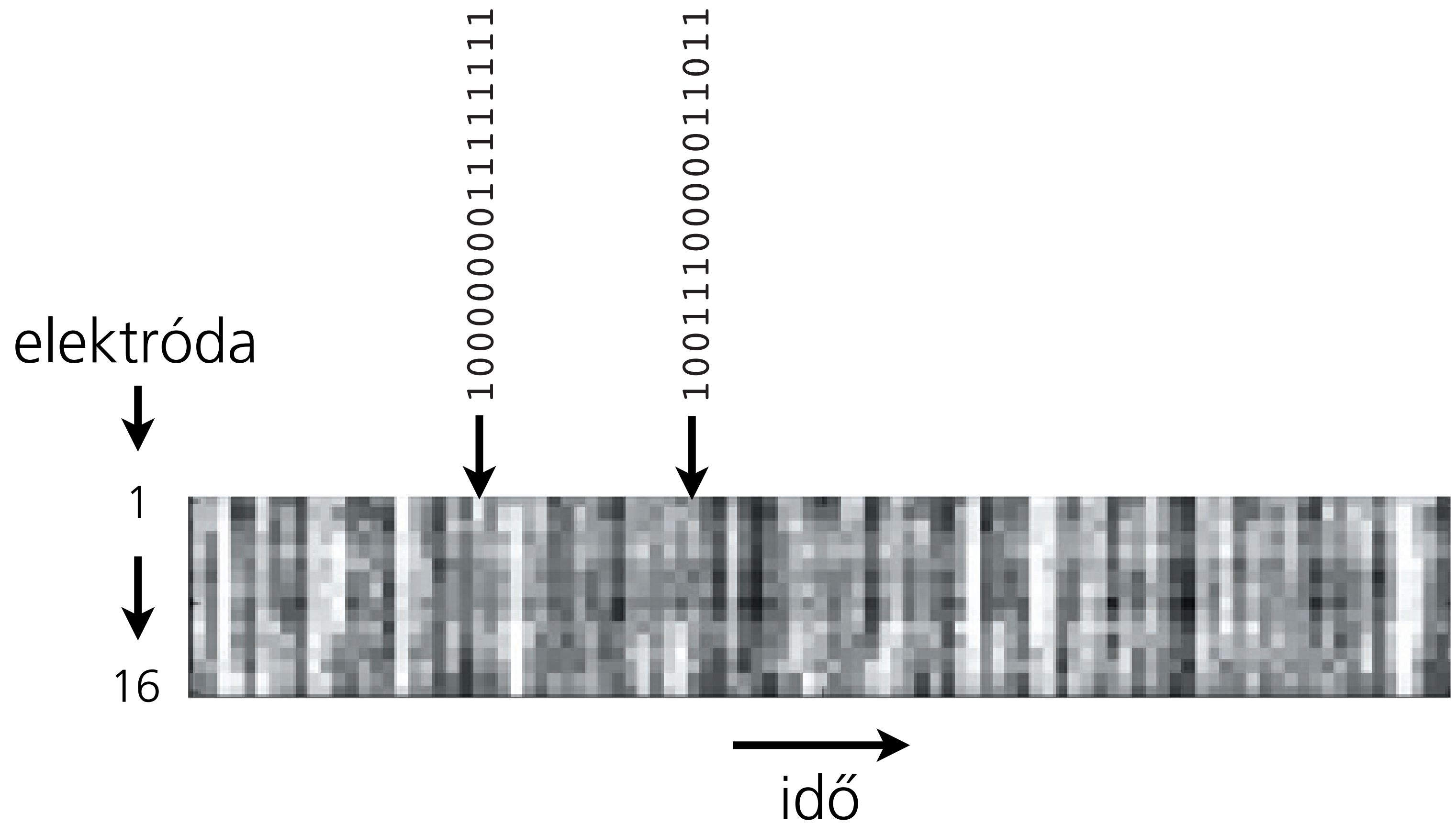
↓

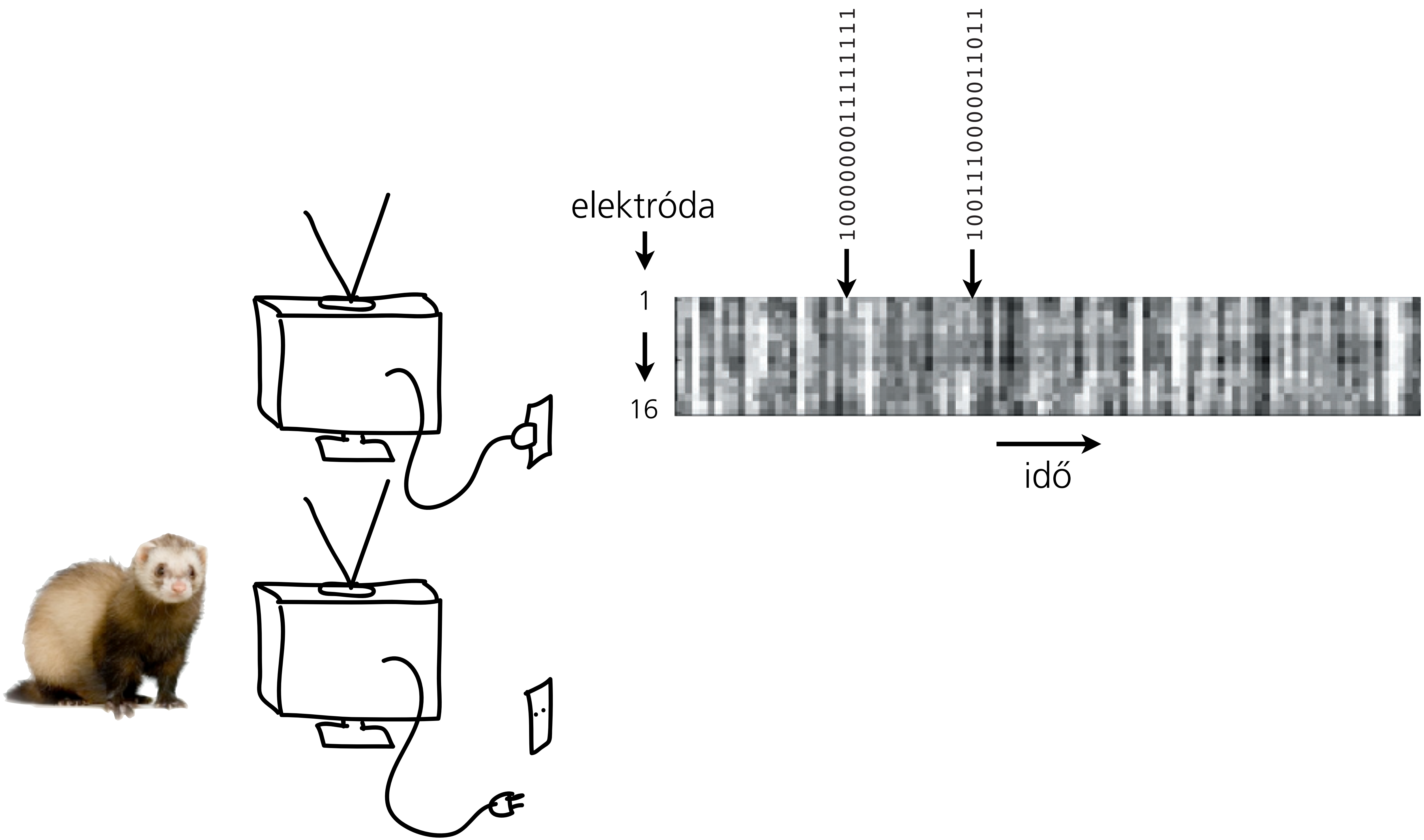
1

↓

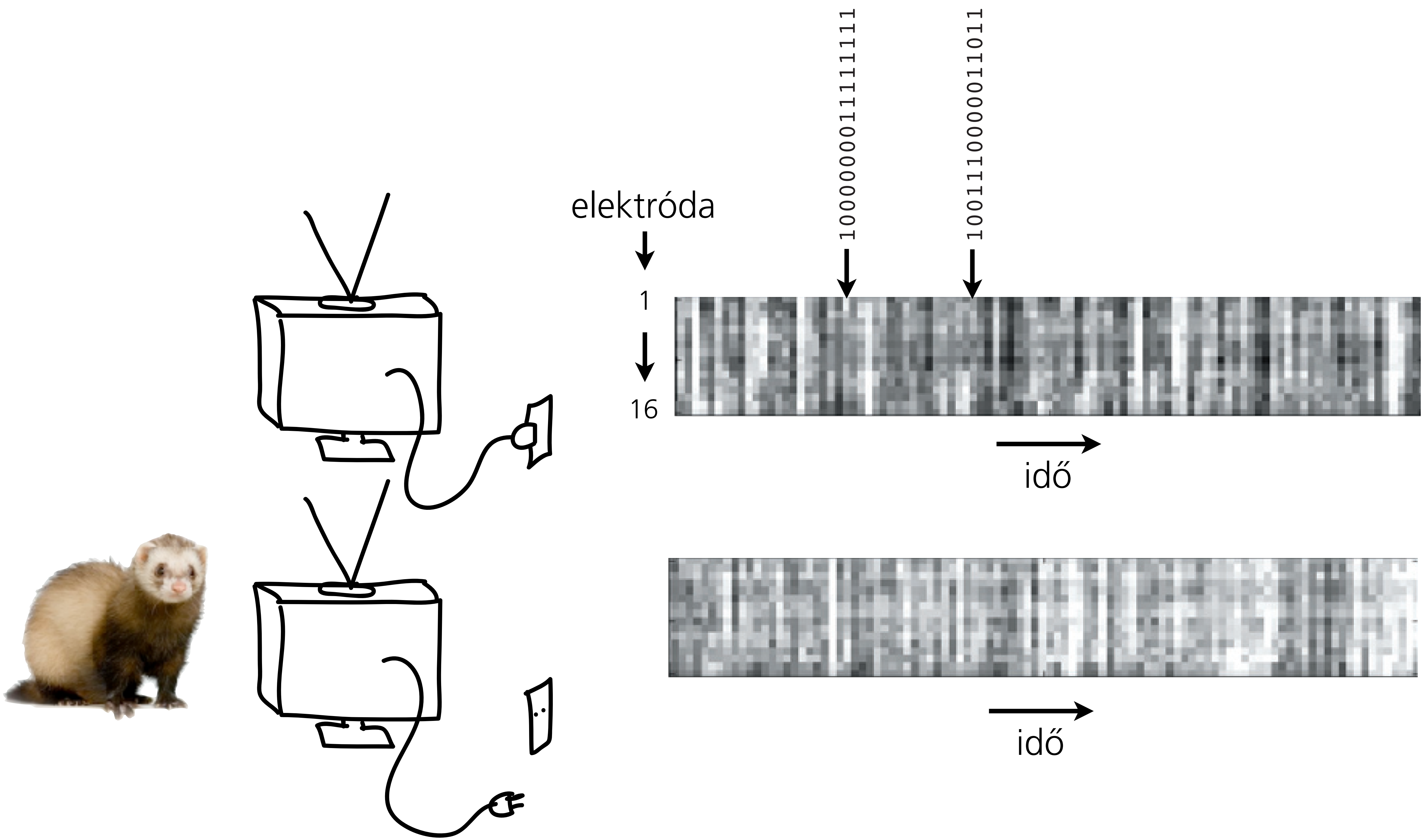
16




elektróda
1
16
1000000111111111
1001110000011011
idő

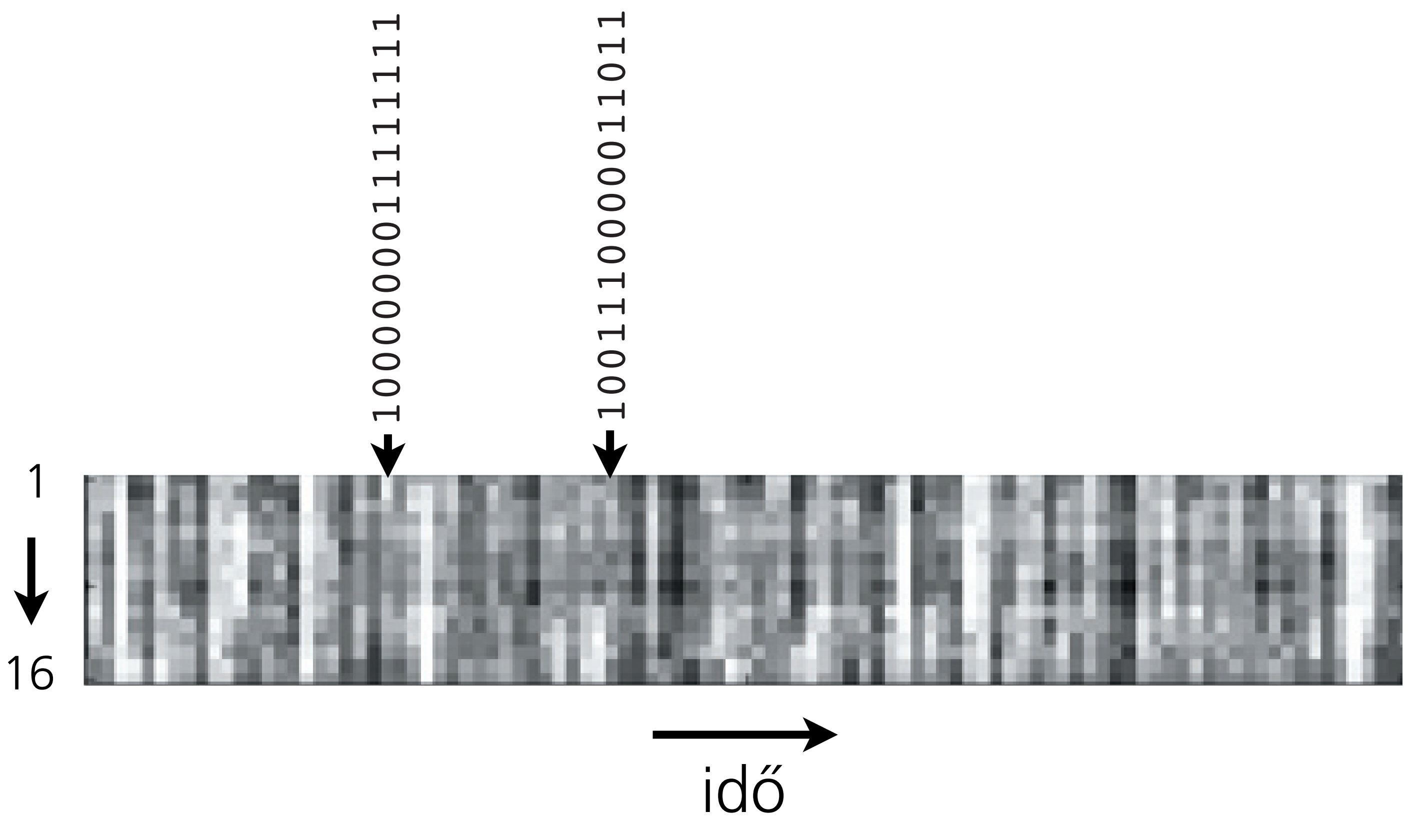


1000000011111111
1001110000011011
elektróda
1
16
idő

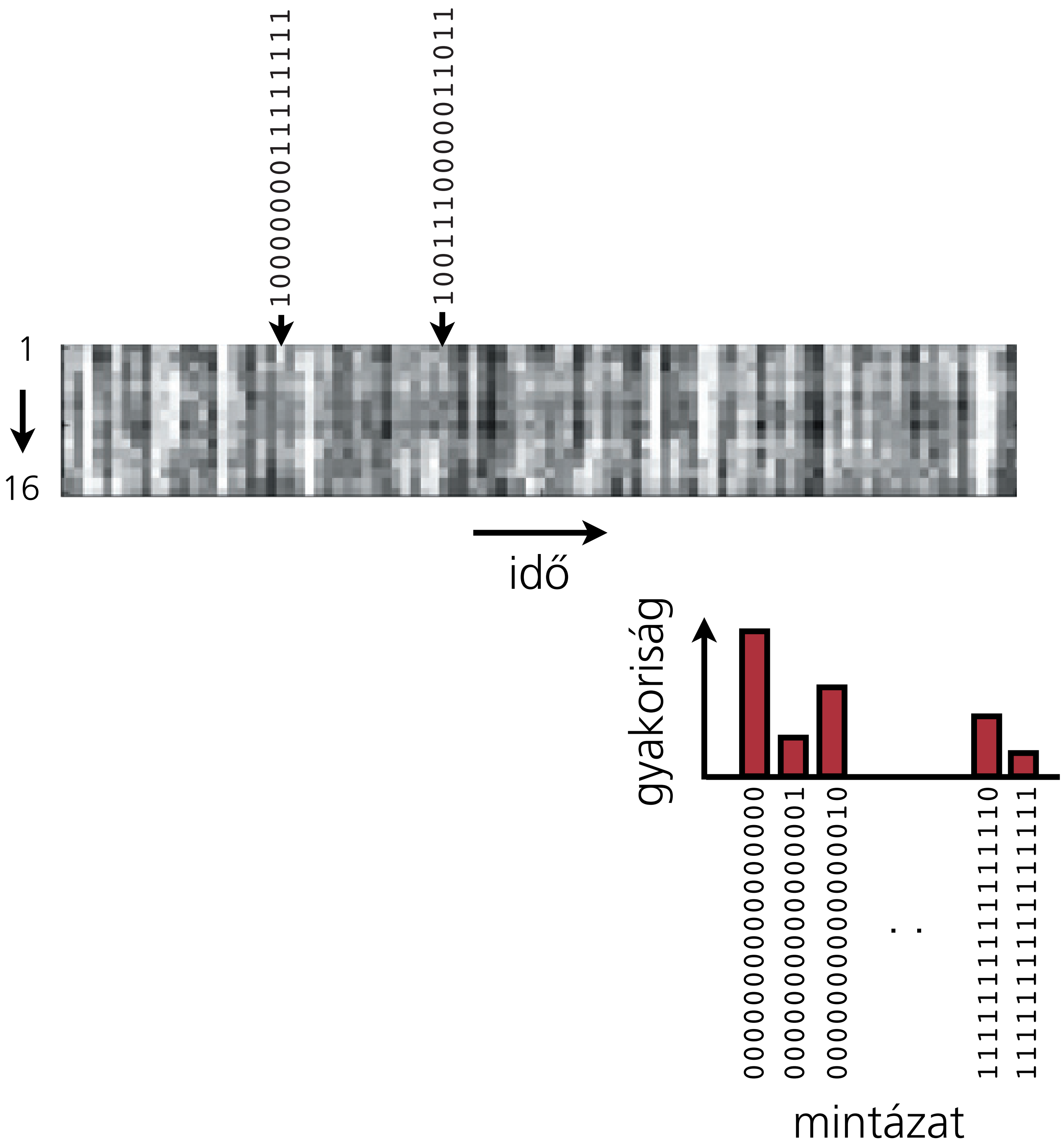


elektróda
1
16
1000000011111111
1001110000011011
idő
idő

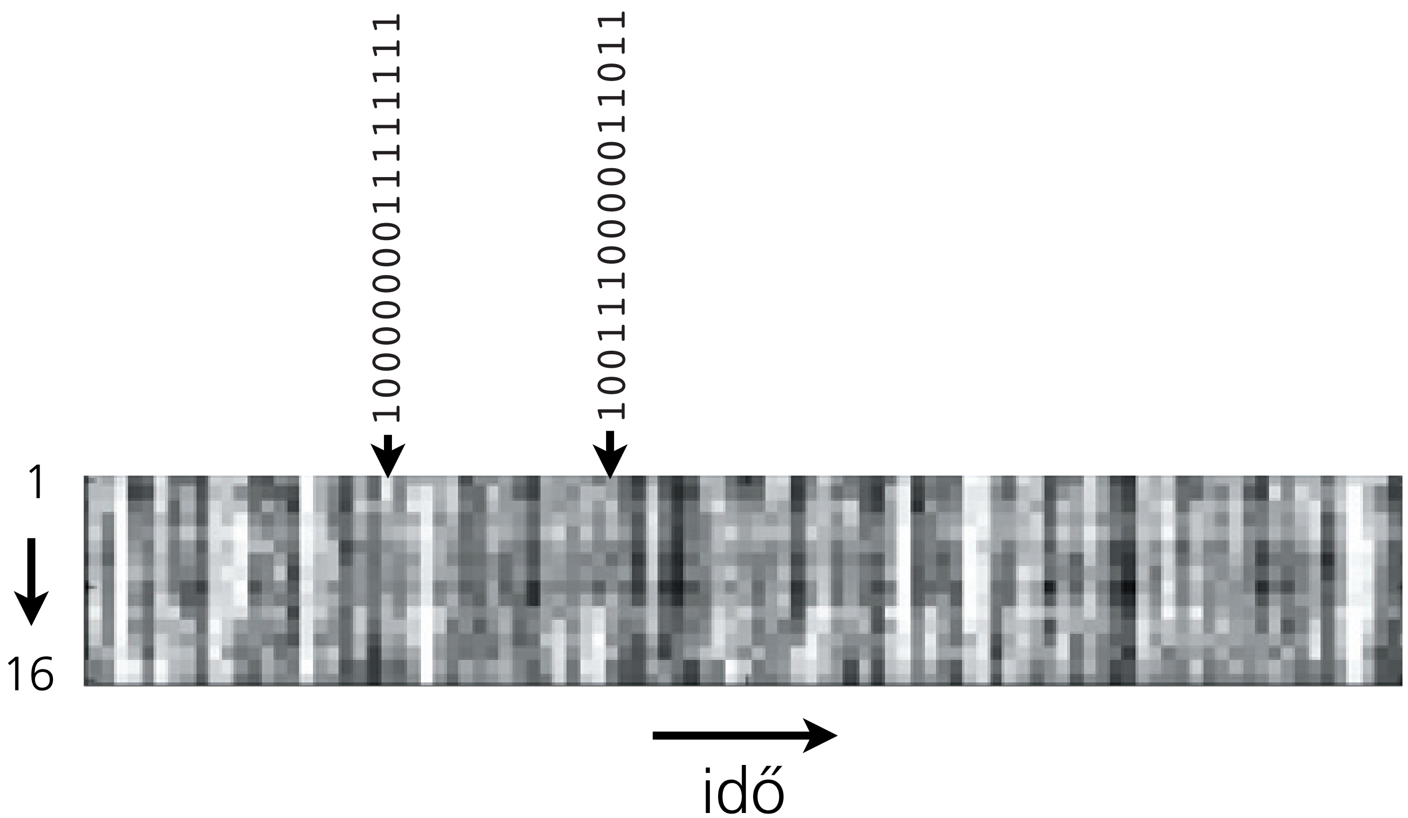
Hatékonyság?

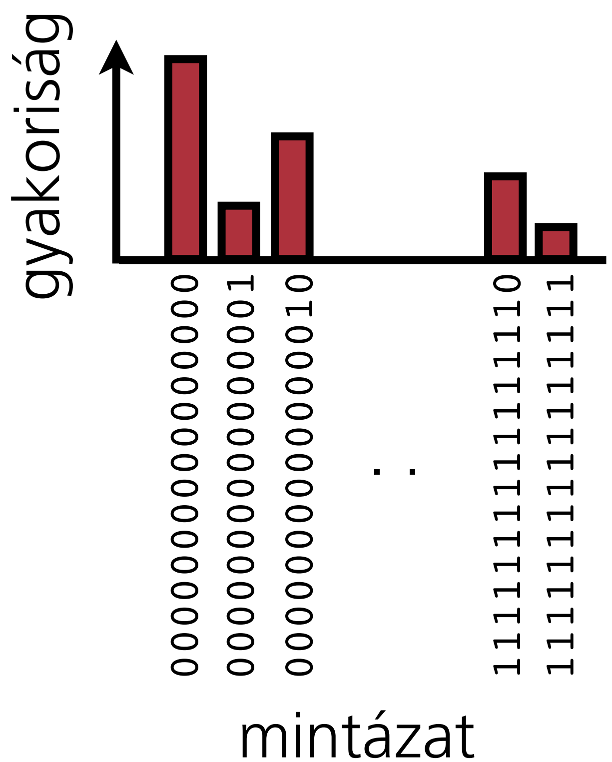


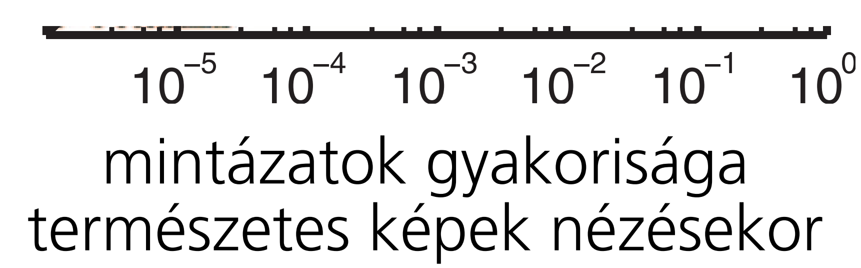
Hatékonyság?



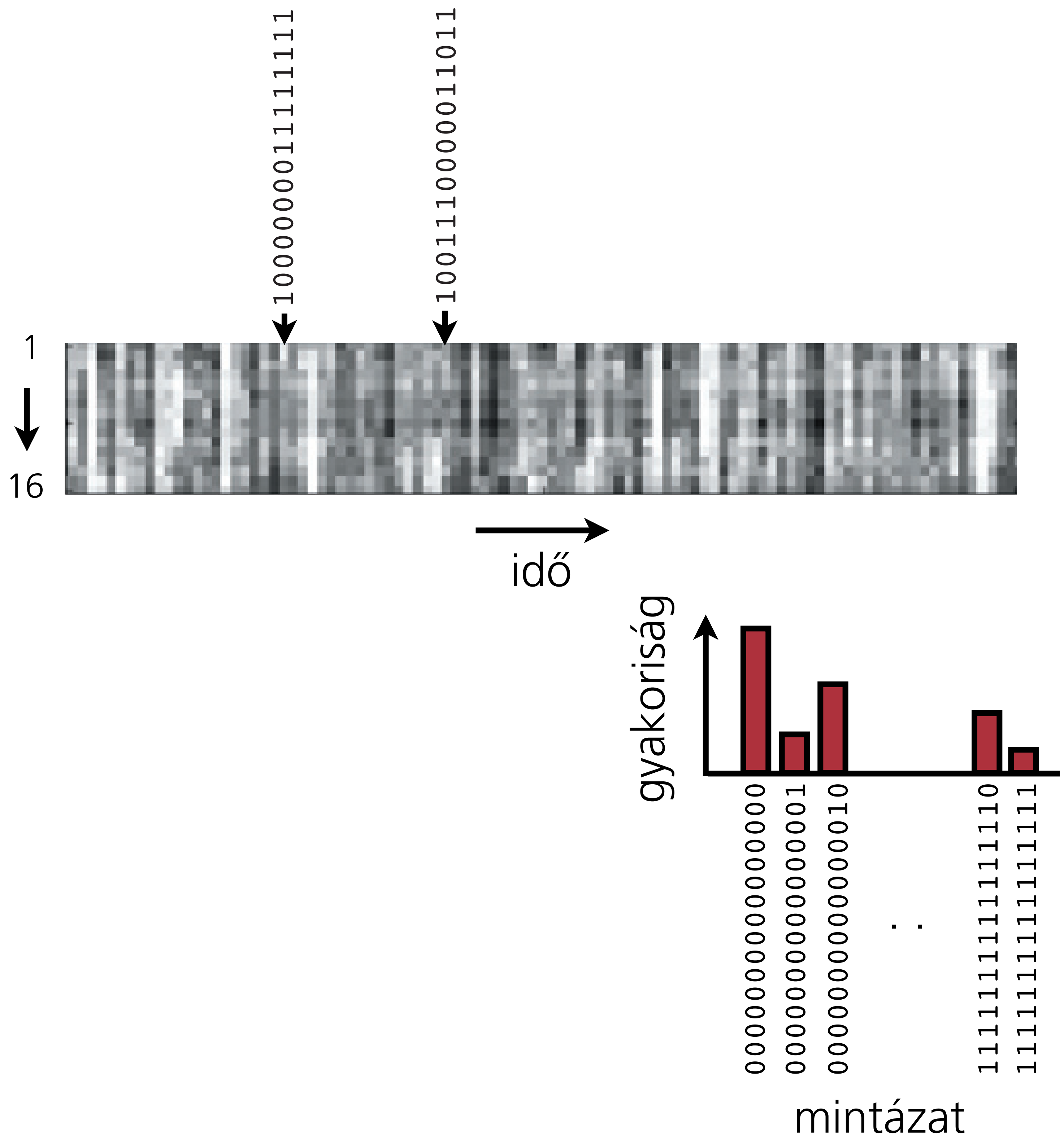
Hatékonyság?



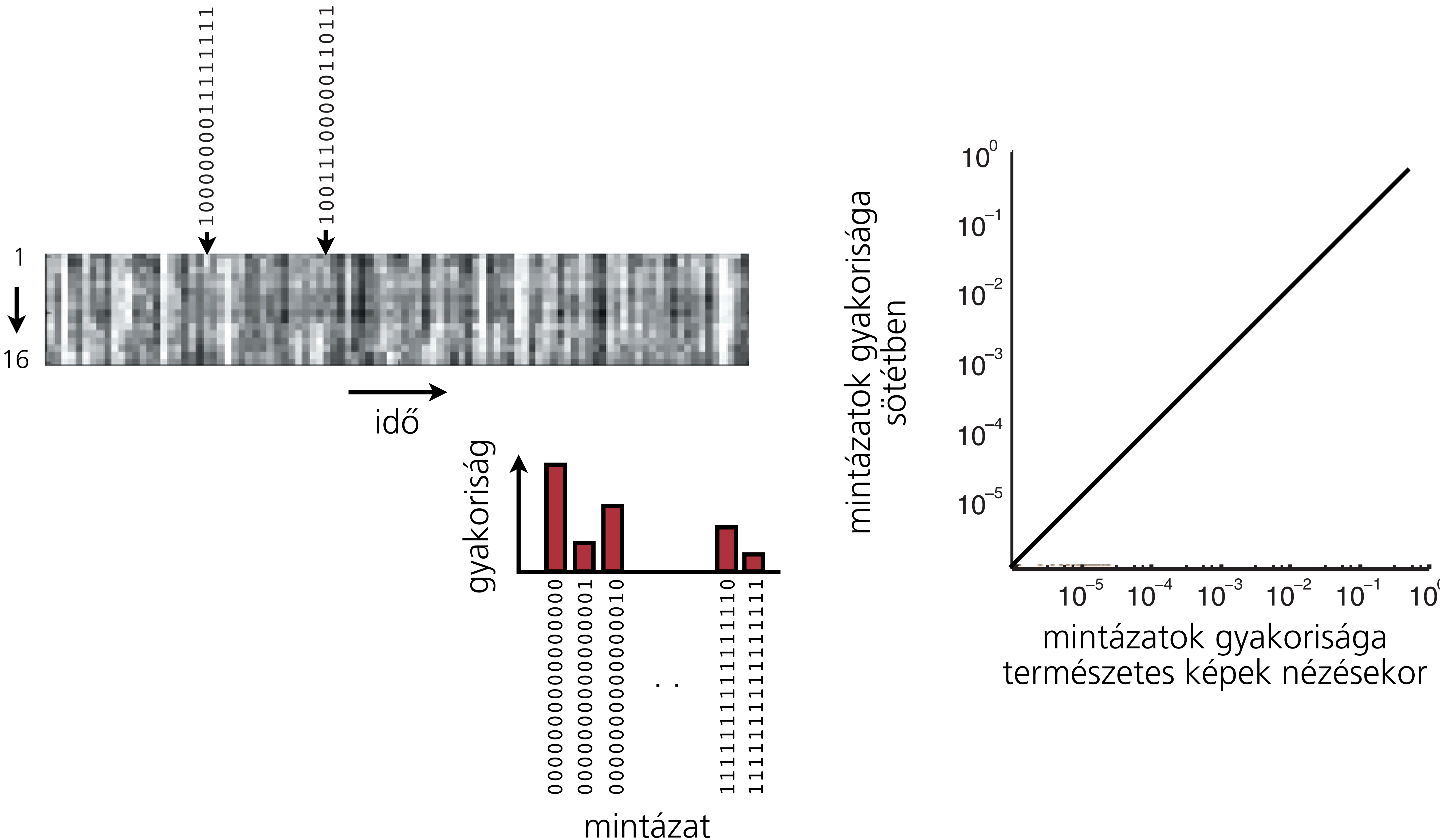




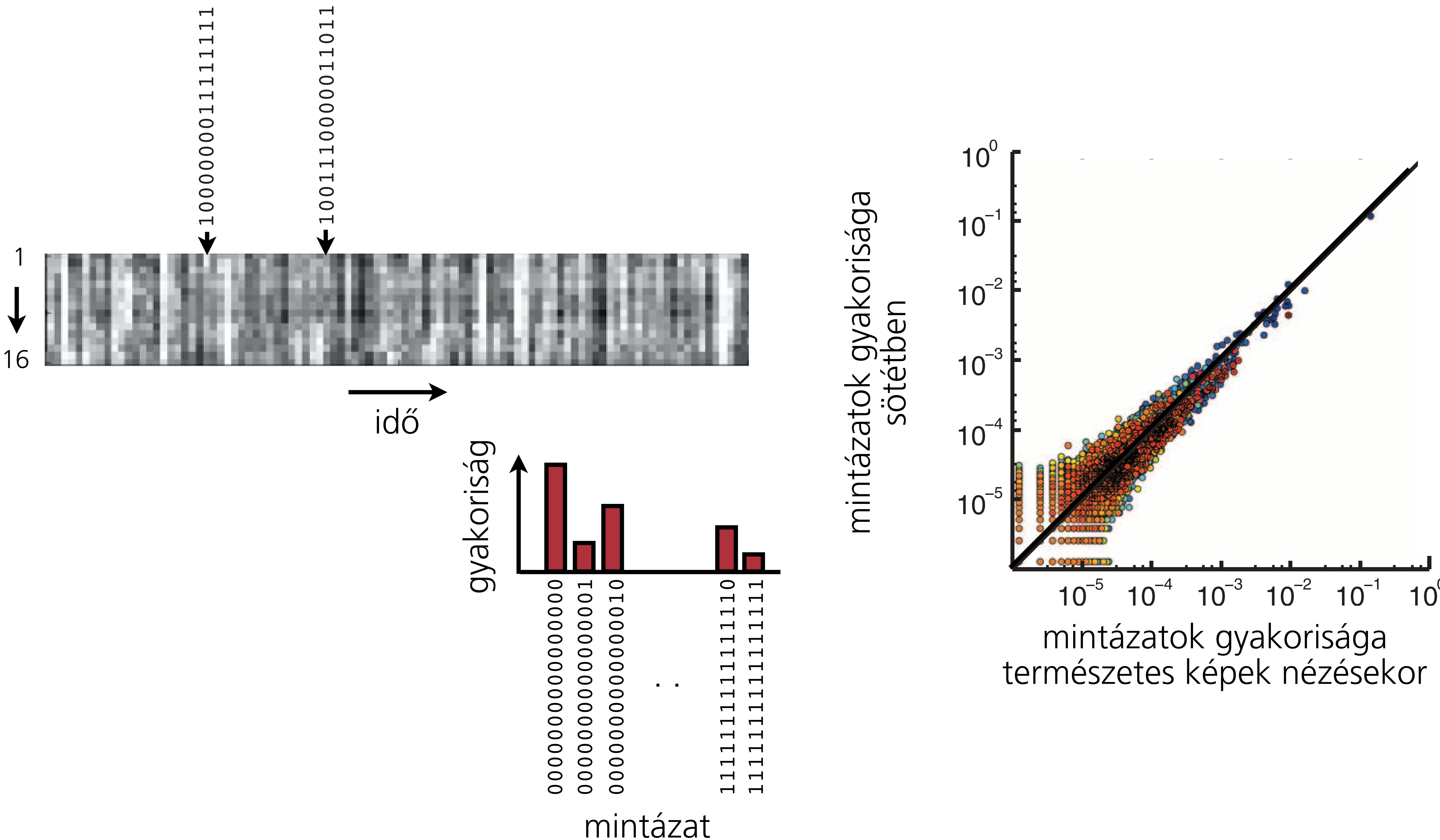
Hatékonyság?



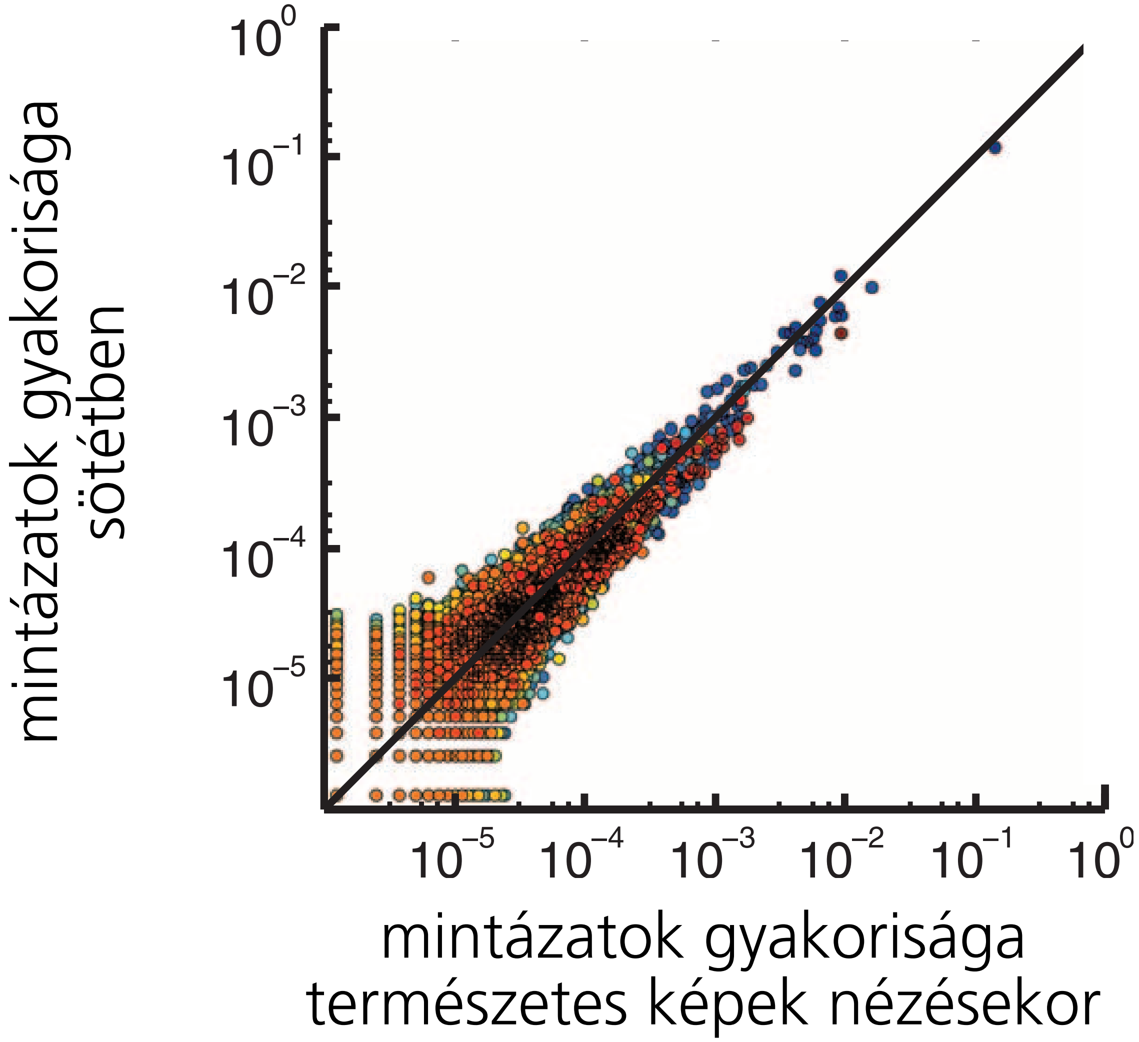
Hatékonyság?



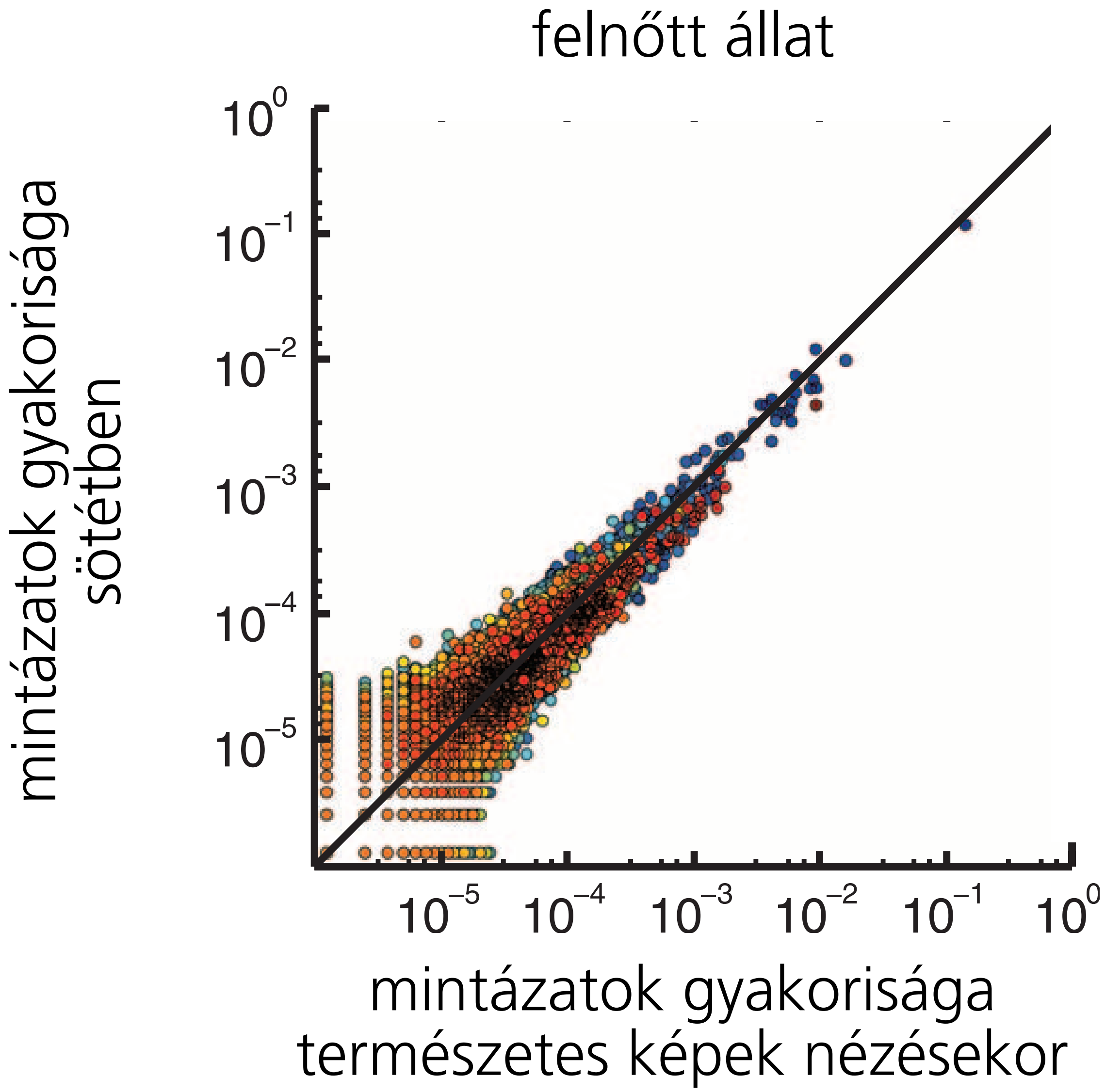
Hatékonyság?

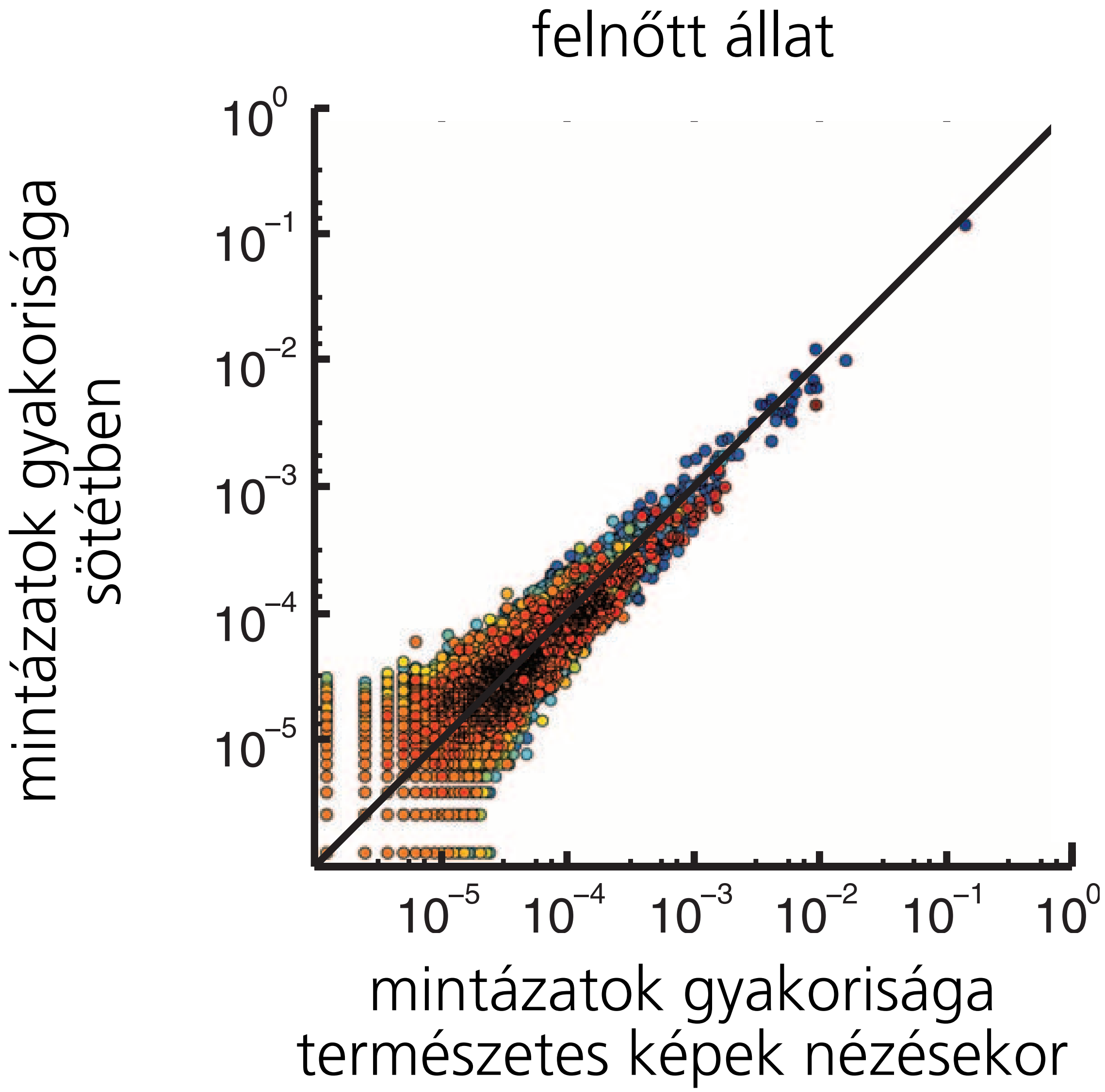


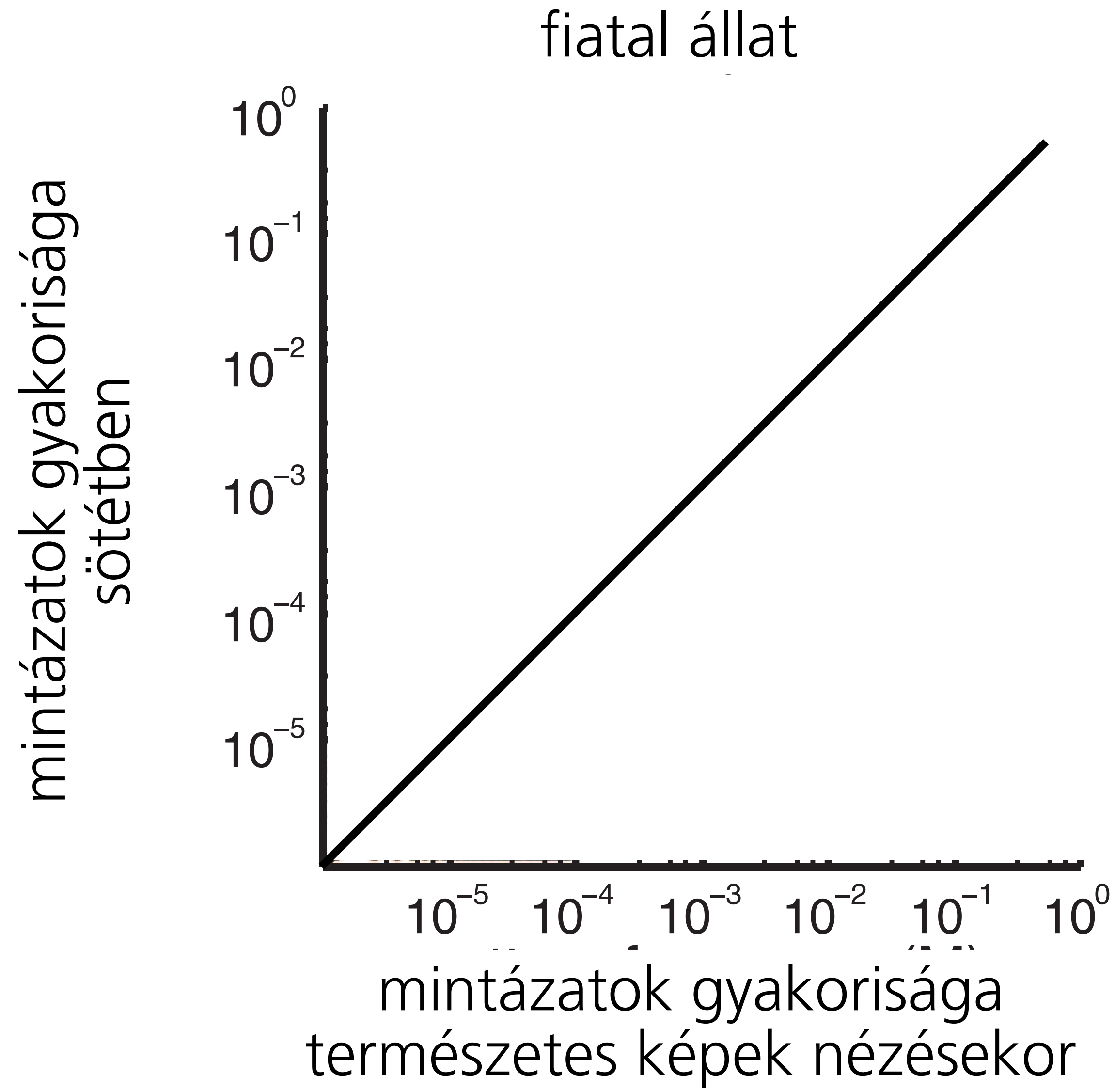
- Ha az idegrendszer ismeri a világ szerkezetét, akkor az elvárásai nem különböznek attól, amit általában érzékel

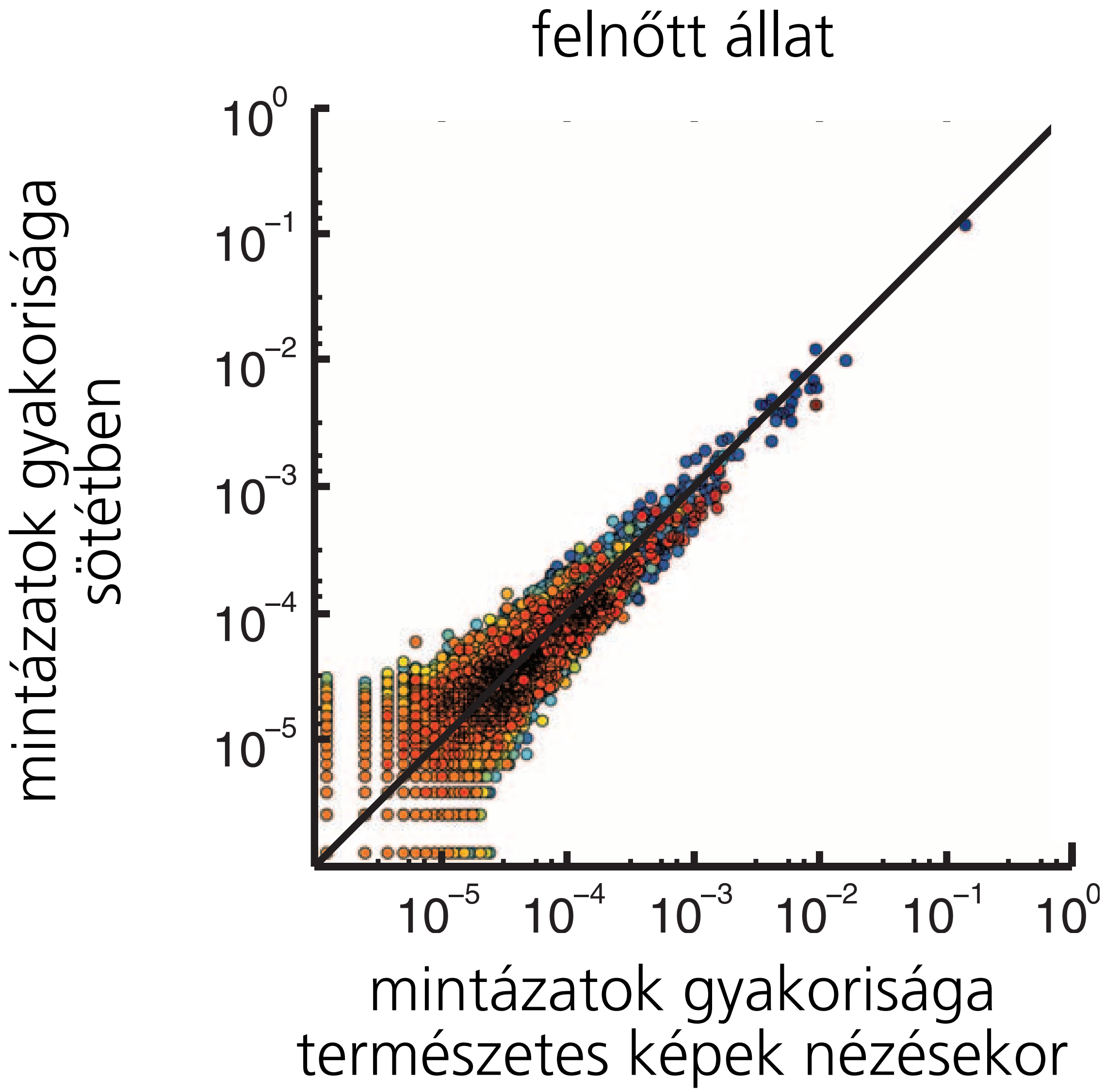


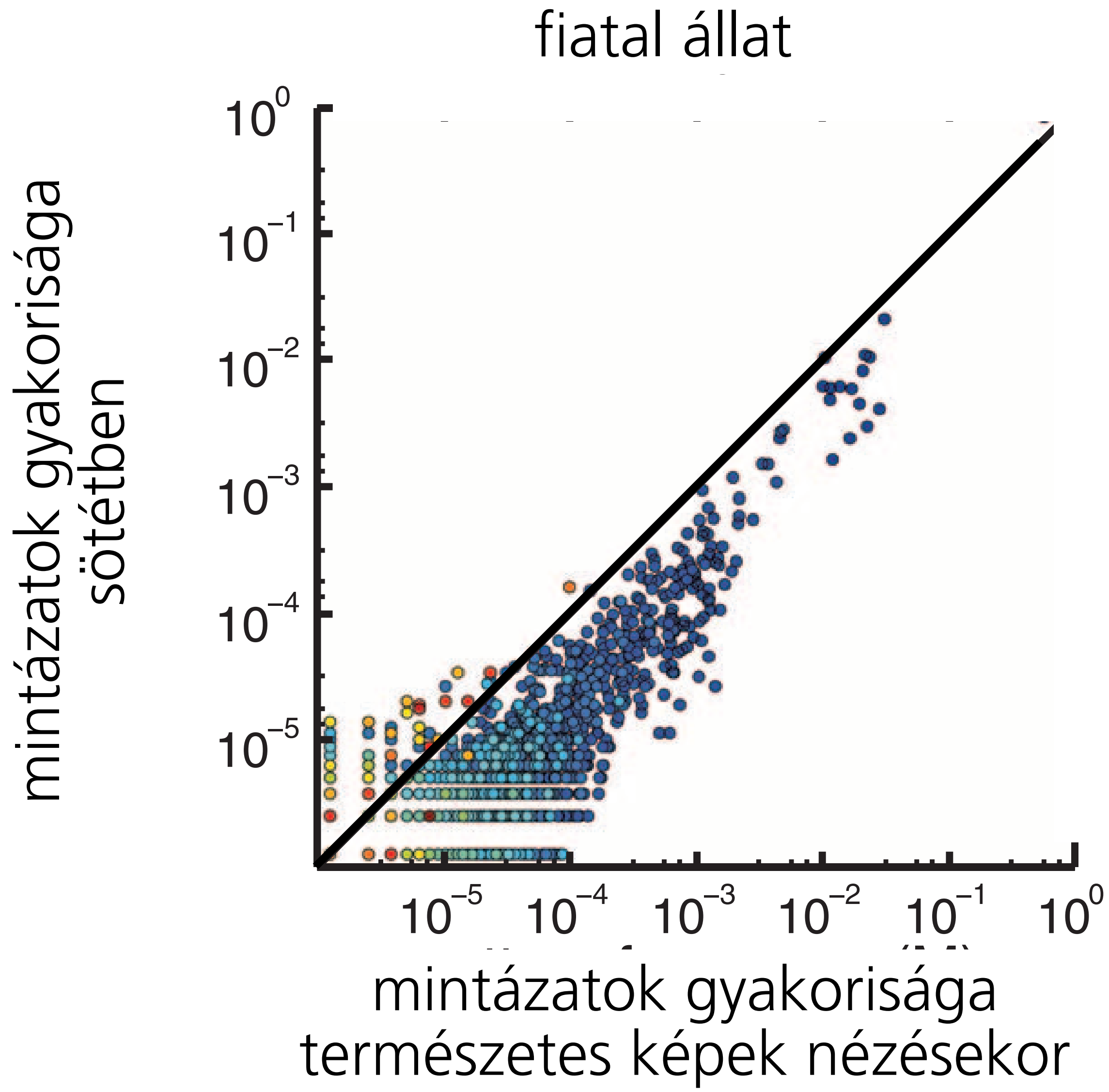
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
mintázatok gyakorisága
sötétben
10^{-5} 10^{-4} 10^{-3} 10^{-2} 10^{-1} 10^{0}
mintázatok gyakorisága
természetes képek nézésekor

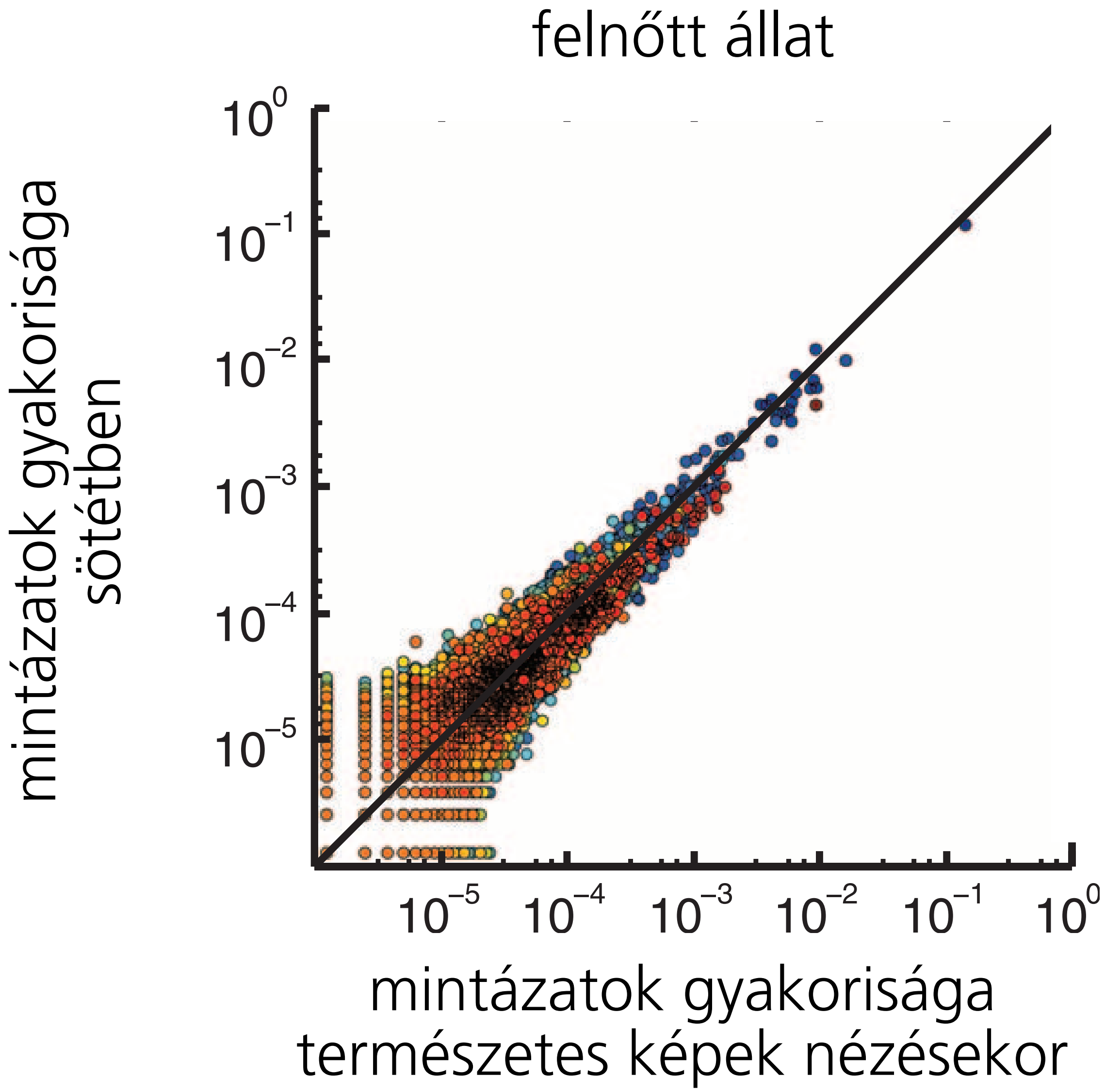

felnőtt állat
mintázatok gyakorisága sötétben
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-5} 10^{-4} 10^{-3} 10^{-2} 10^{-1} 10^{0}
mintázatok gyakorisága természetes képek nézésekor

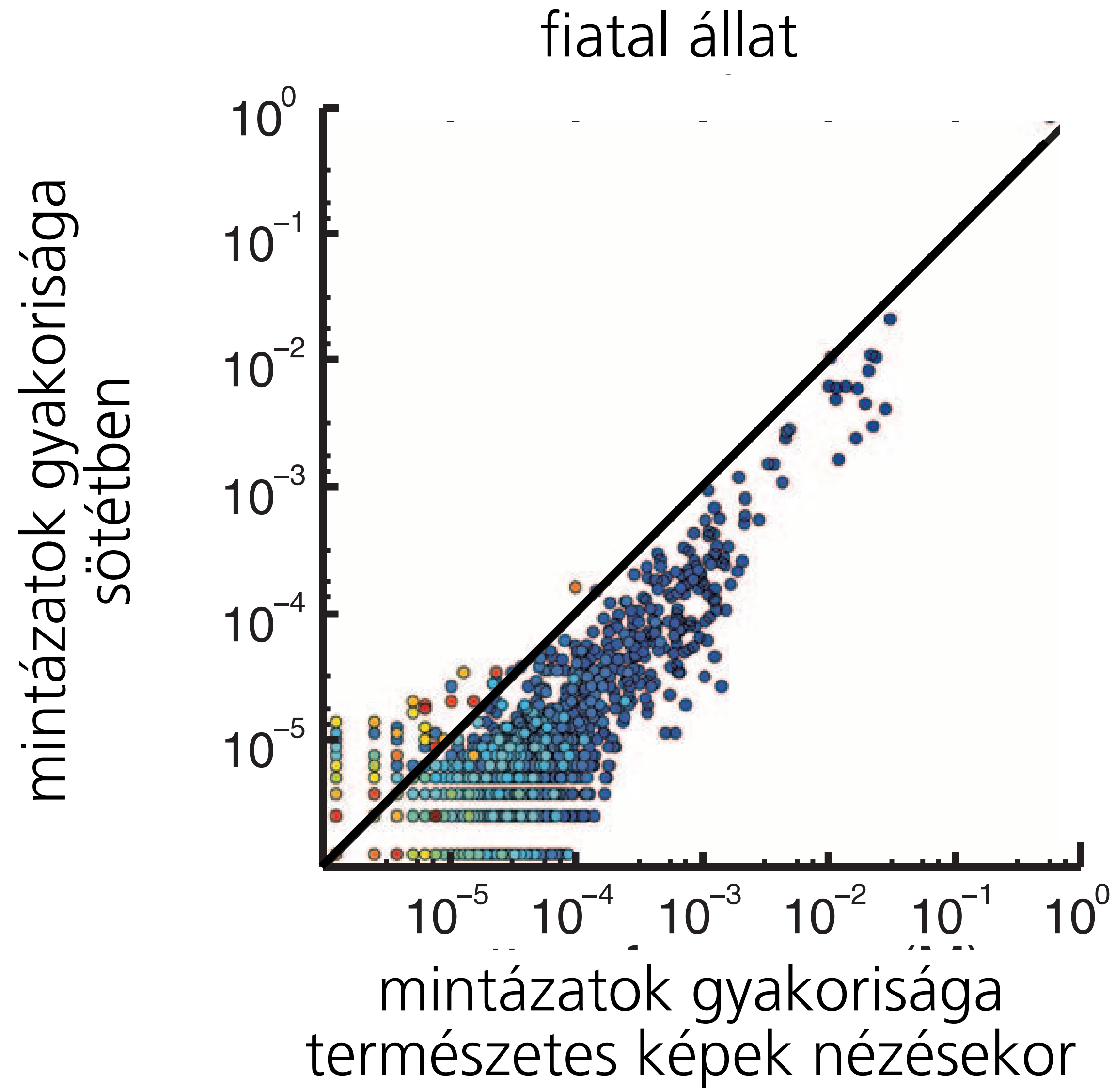

felnőtt állat
mintázatok gyakorisága sötétben
mintázatok gyakorisága természetes képek nézésekor
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}

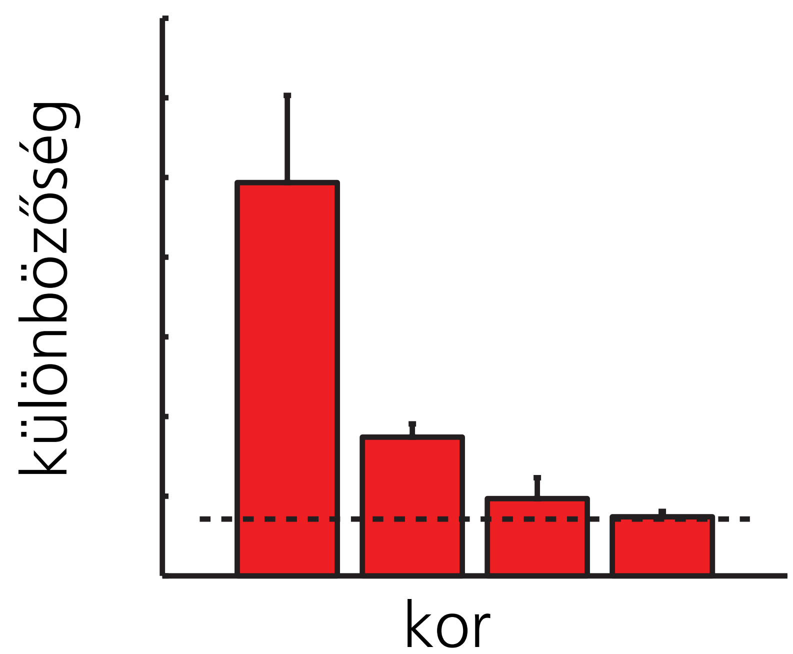

fiatal állat
mintázatok gyakorisága sötétben
mintázatok gyakorisága természetes képek nézésekor
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}

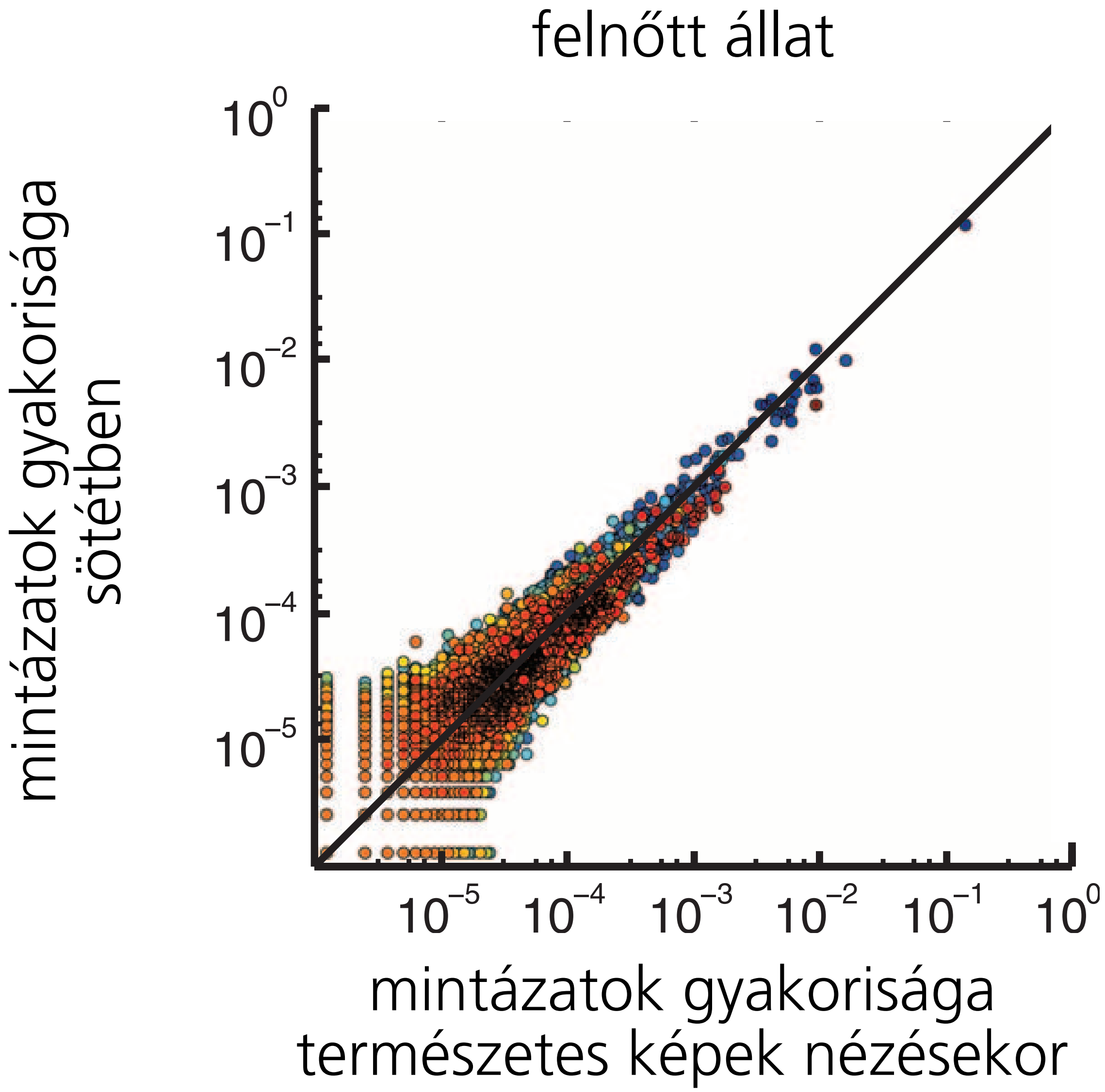

felnőtt állat
mintázatok gyakorisága sötétben
mintázatok gyakorisága természetes képek nézésekor

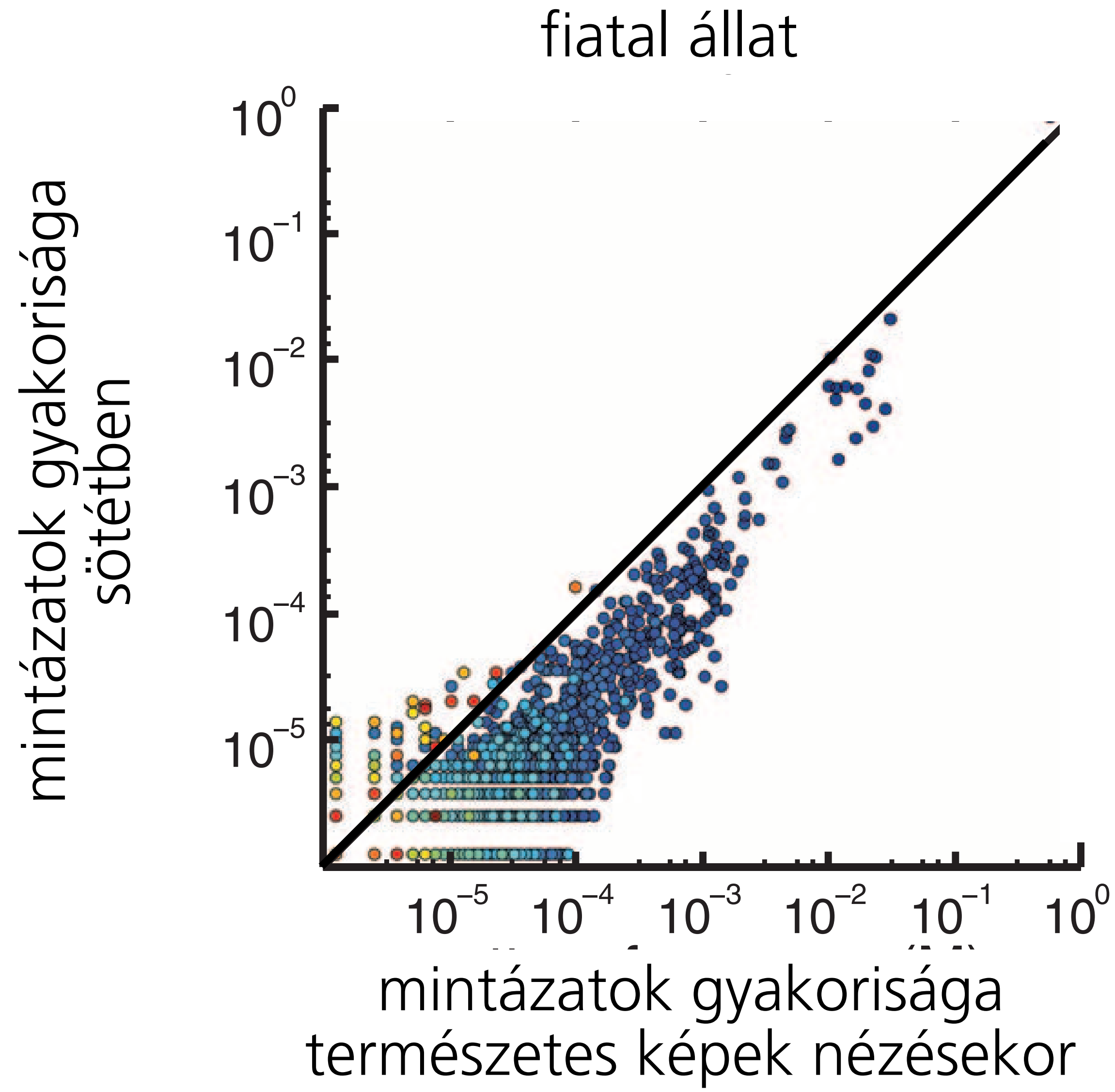

fiatal állat
mintázatok gyakorisága sötétben
mintázatok gyakorisága természetes képek nézésekor

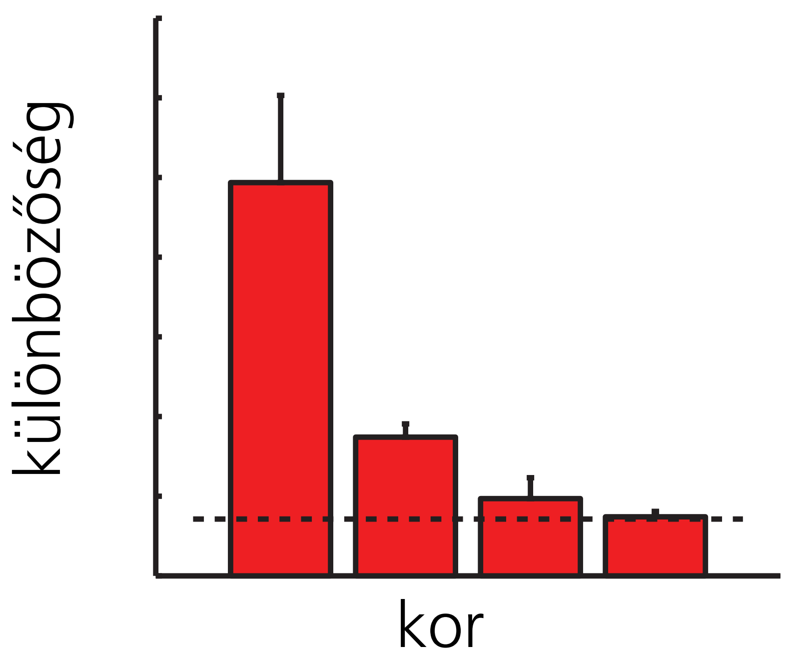

felnőtt állat
mintázatok gyakorisága sötétben
mintázatok gyakorisága természetes képek nézésekor
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}


fiatal állat
mintázatok gyakorisága sötétben
mintázatok gyakorisága természetes képek nézésekor
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}

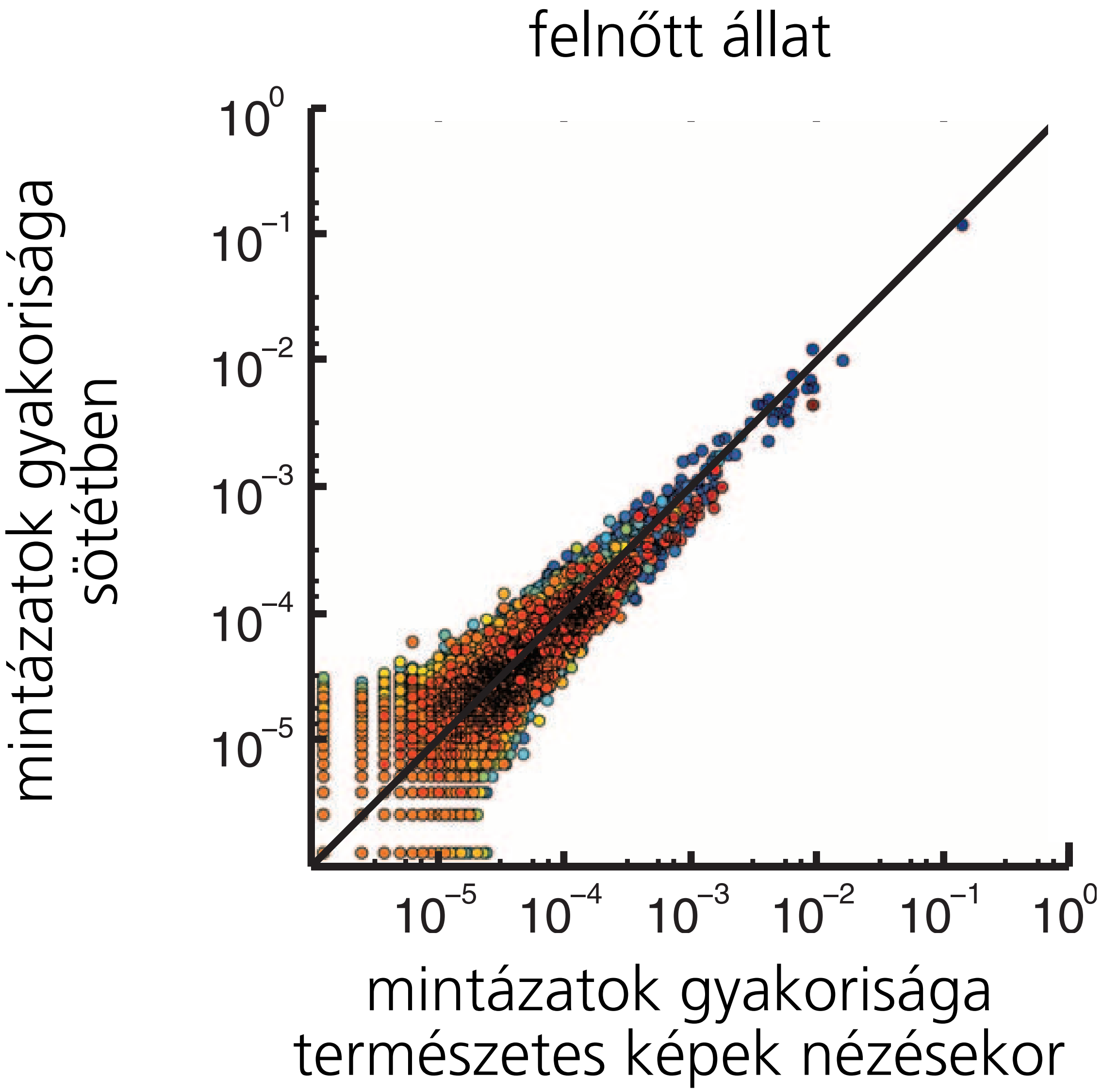

különbözőség
kor

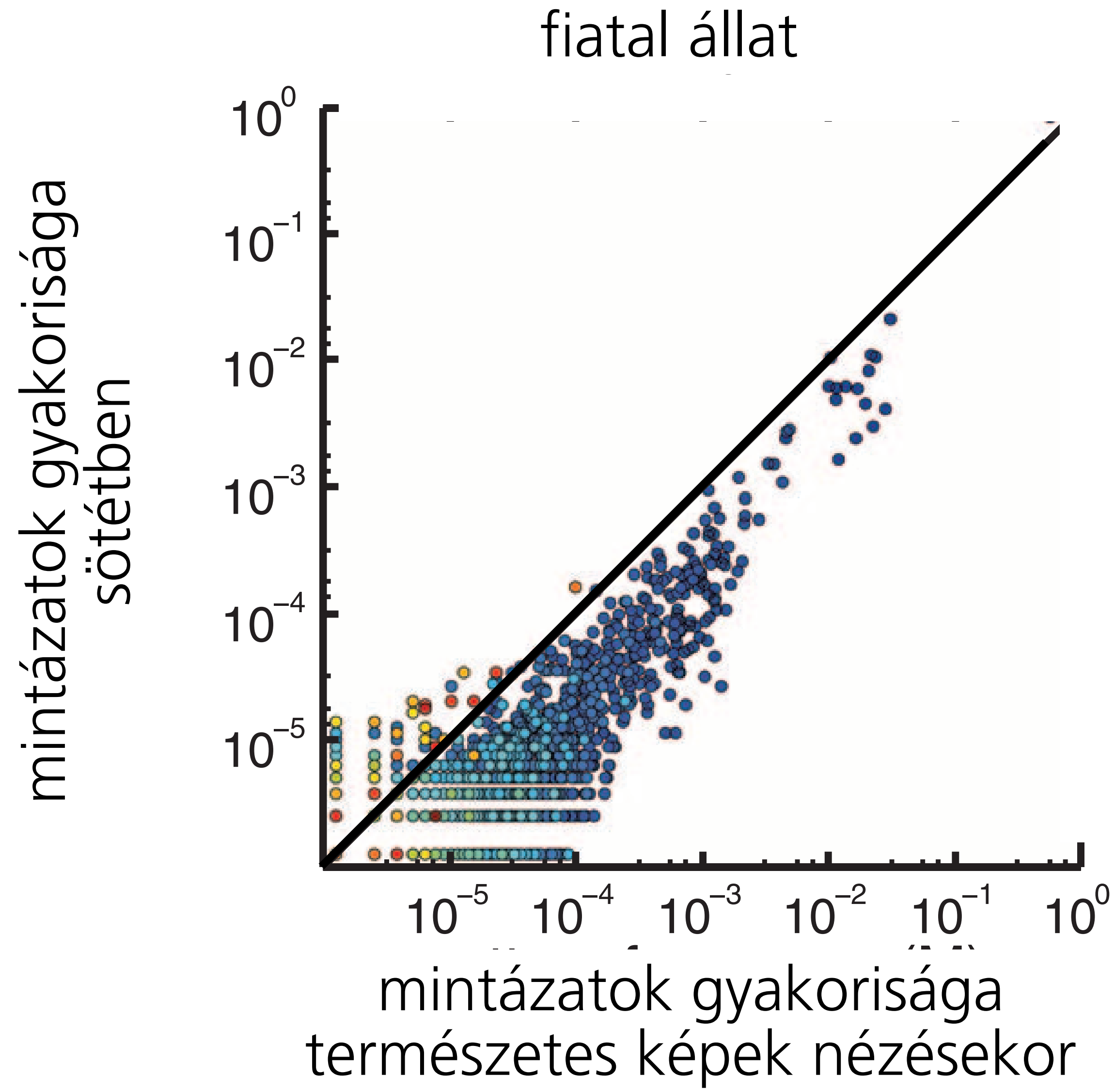


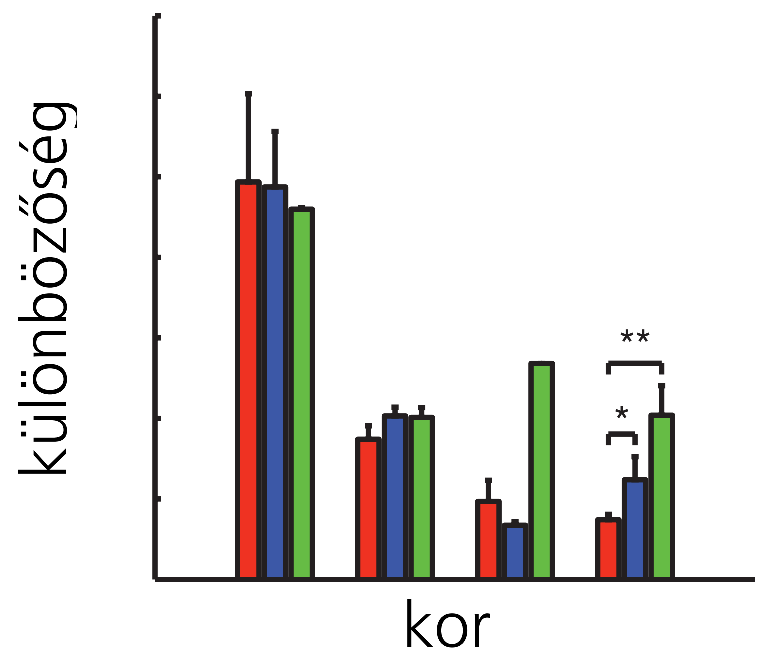




természetes képek
zaj
periodikusan sávozott képek



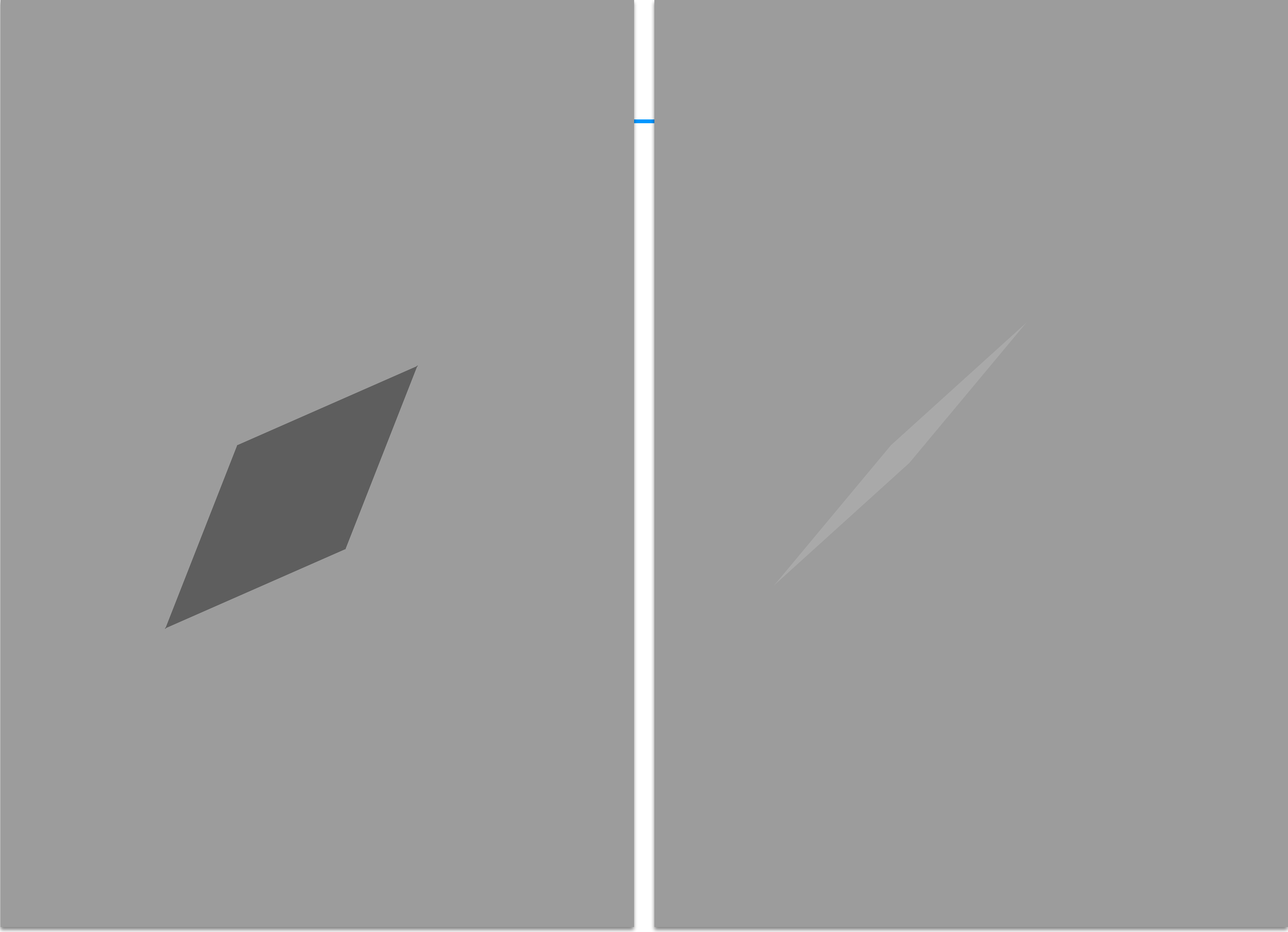


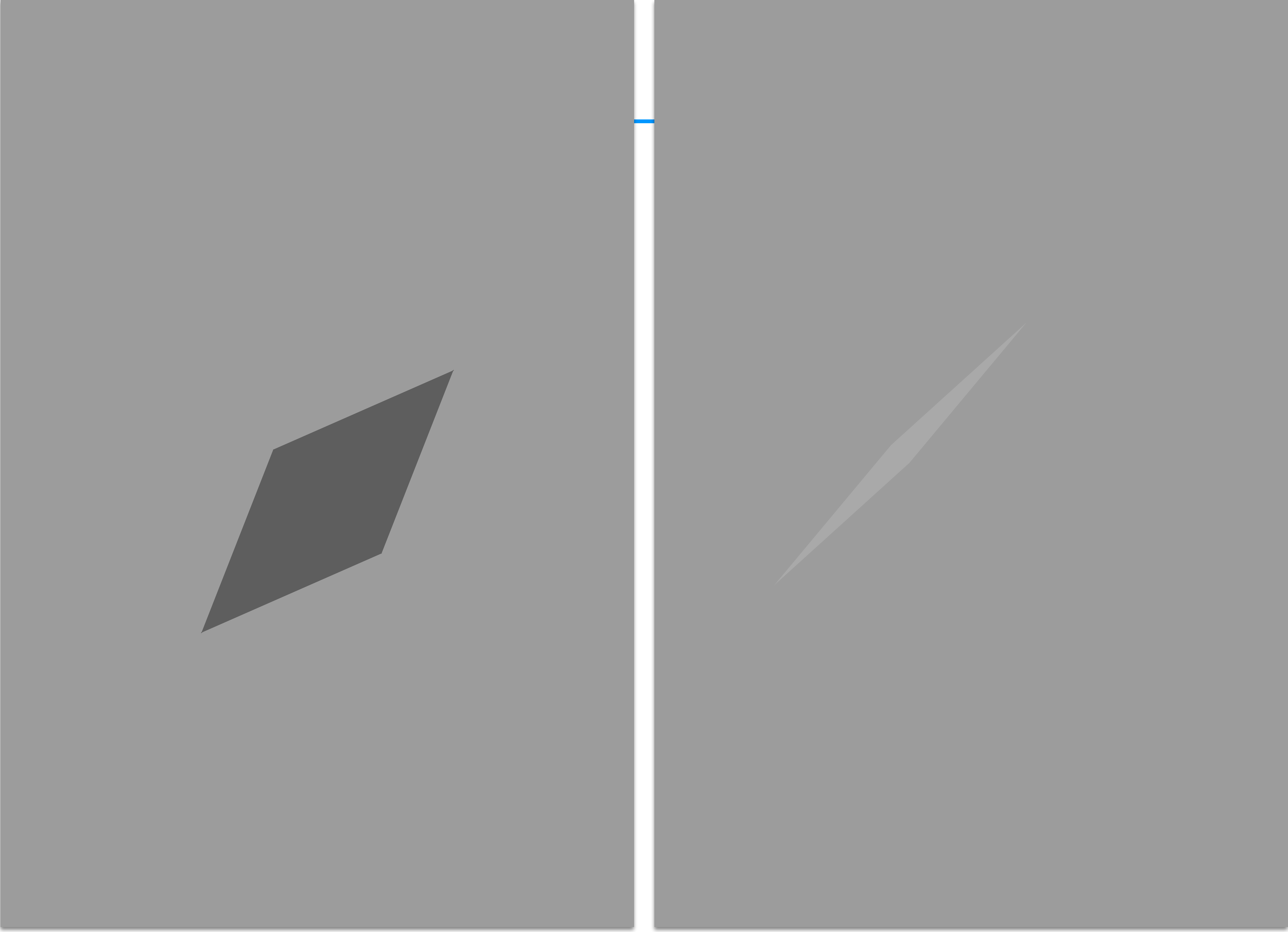


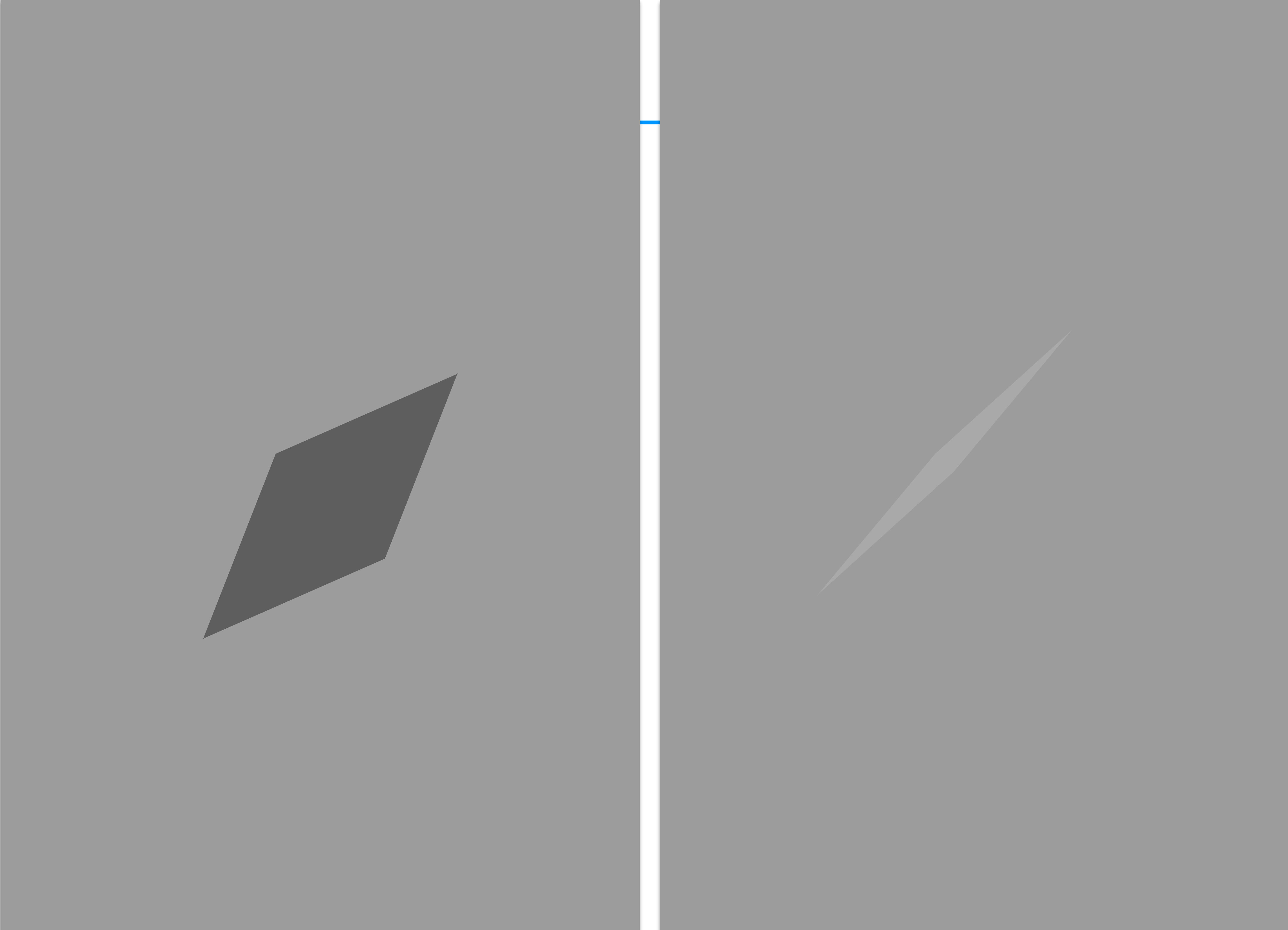
természetes képek
zaj
periodikusan sávozott képek

Measuring priors

- Structurred
- Adapted to the environmental statistics
- Depends on subjective experience
- Task independence



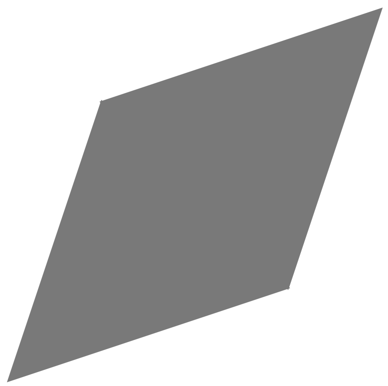




Motion illusions as optimal percepts

Weiss, Simoncelli & Adelson (2002) Nat Neurosci

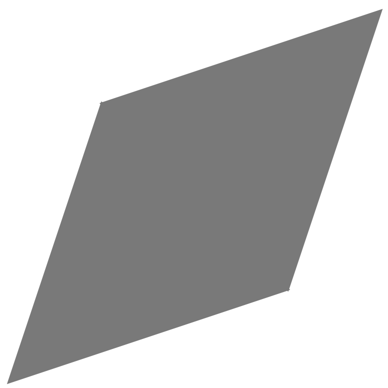
http://www.cs.huji.ac.il/~yweiss/Rhombus/rhombus.html



Motion illusions as optimal percepts

Weiss, Simoncelli & Adelson (2002) Nat Neurosci

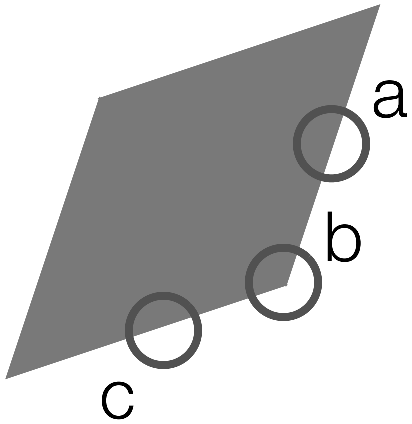
http://www.cs.huji.ac.il/~yweiss/Rhombus/rhombus.html



Motion illusions as optimal percepts

Weiss, Simoncelli & Adelson (2002) Nat Neurosci

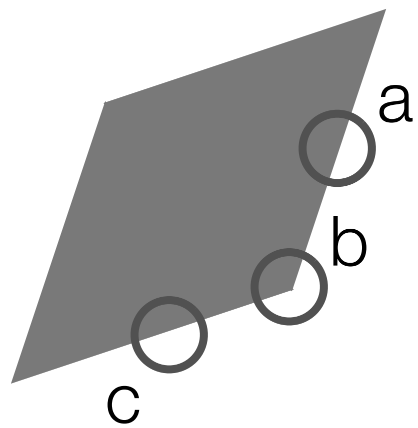
http://www.cs.huji.ac.il/~yweiss/Rhombus/rhombus.html

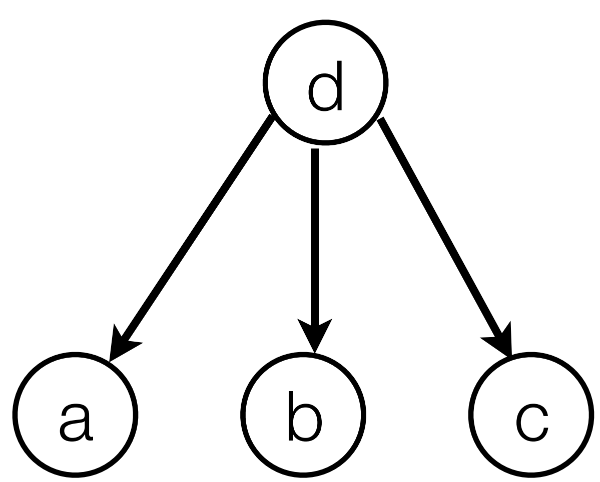


Motion illusions as optimal percepts

Weiss, Simoncelli & Adelson (2002) Nat Neurosci

http://www.cs.huji.ac.il/~yweiss/Rhombus/rhombus.html

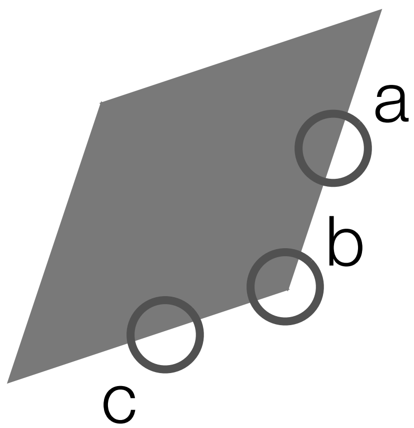


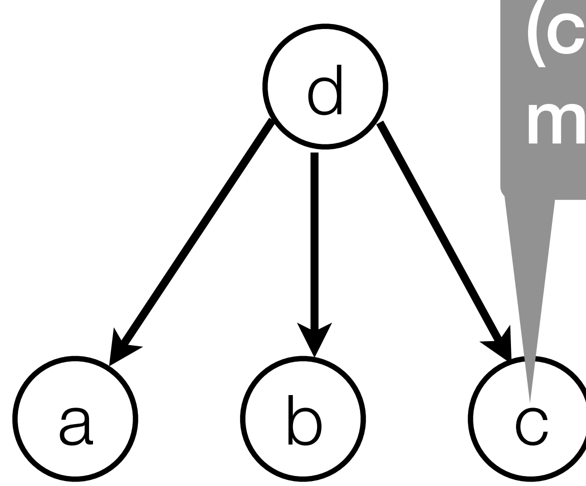


Motion illusions as optimal percepts

Weiss, Simoncelli & Adelson (2002) Nat Neurosci

http://www.cs.huji.ac.il/~yweiss/Rhombus/rhombus.html



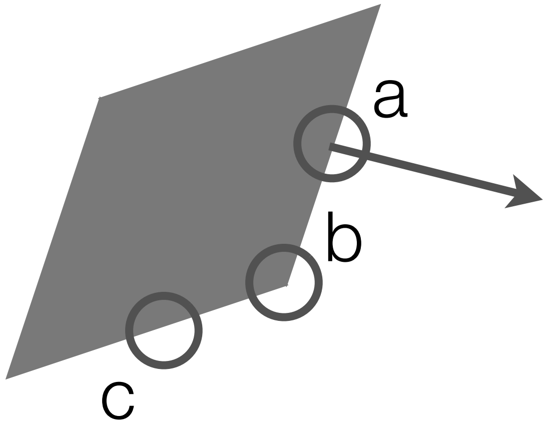


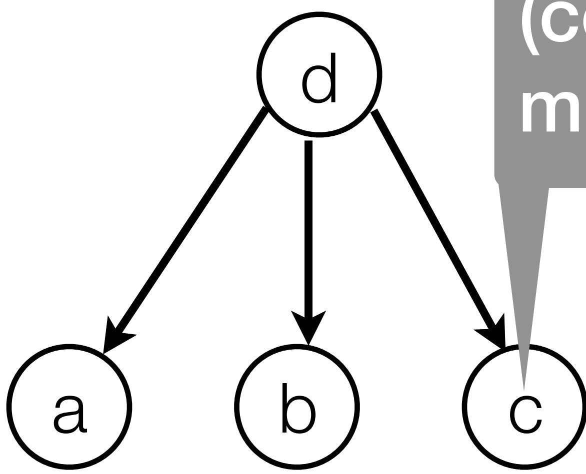
independent pieces of evidence (conditioned on the movement of the object)

Motion illusions as optimal percepts

Weiss, Simoncelli & Adelson (2002) Nat Neurosci

http://www.cs.huji.ac.il/~yweiss/Rhombus/rhombus.html



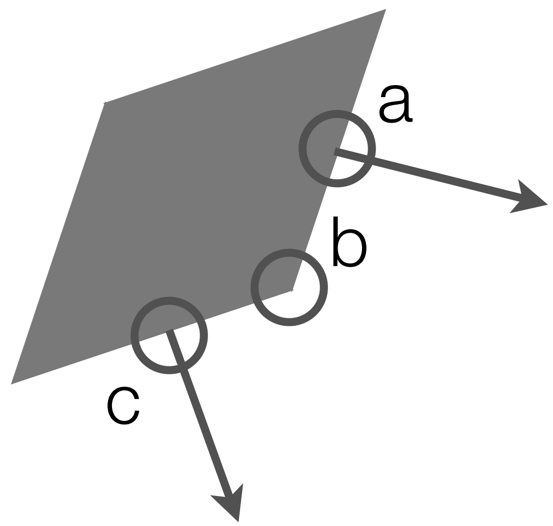


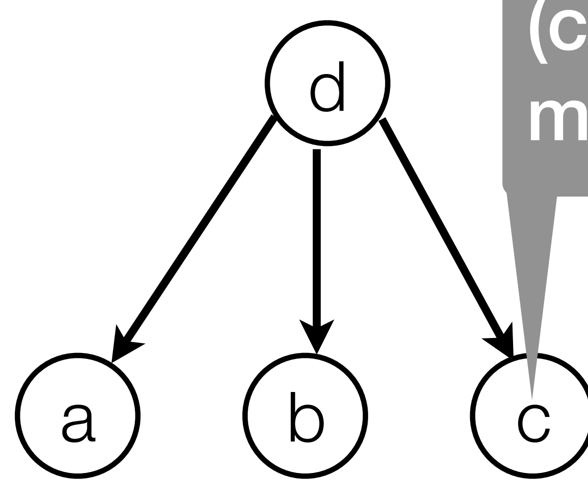
independent pieces of evidence (conditioned on the movement of the object)

Motion illusions as optimal percepts

Weiss, Simoncelli & Adelson (2002) Nat Neurosci

http://www.cs.huji.ac.il/~yweiss/Rhombus/rhombus.html



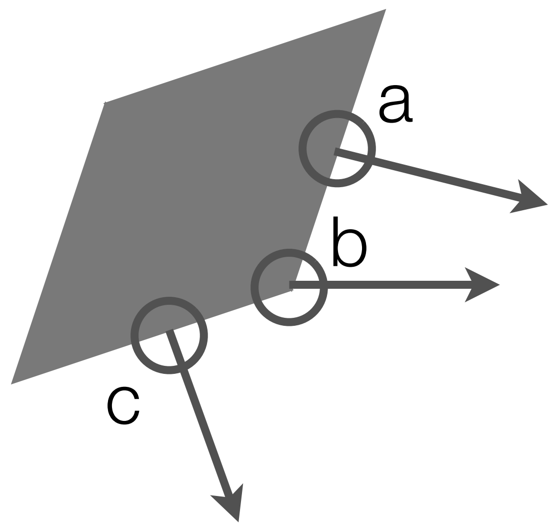


independent pieces of evidence (conditioned on the movement of the object)

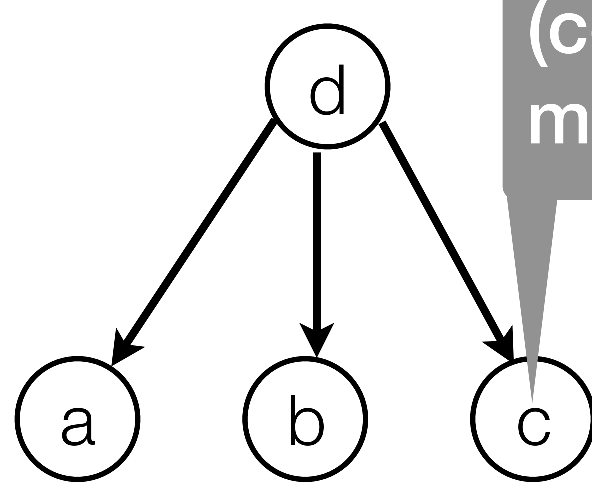
Motion illusions as optimal percepts

Weiss, Simoncelli & Adelson (2002) Nat Neurosci

http://www.cs.huji.ac.il/~yweiss/Rhombus/rhombus.html



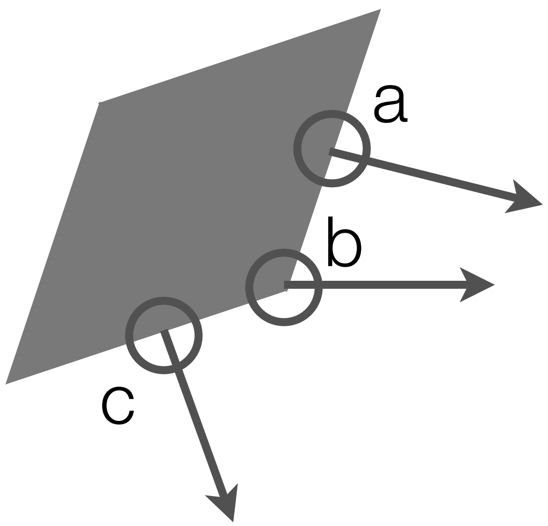
independent pieces of evidence (conditioned on the movement of the object)



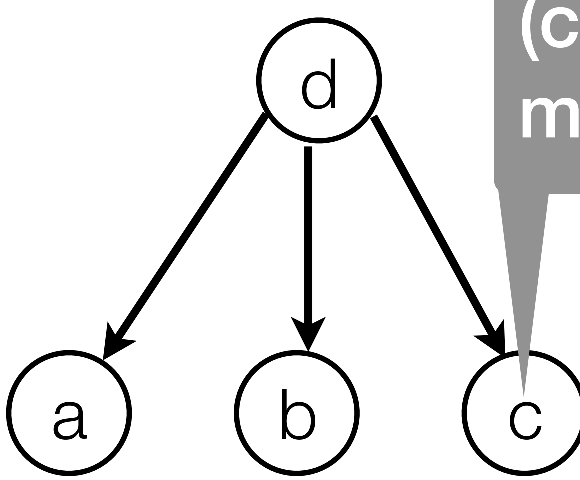
Motion illusions as optimal percepts

Weiss, Simoncelli & Adelson (2002) Nat Neurosci

http://www.cs.huji.ac.il/~yweiss/Rhombus/rhombus.html



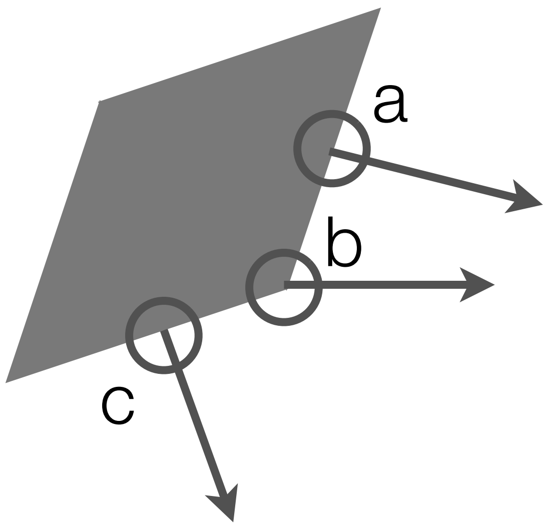
What kind of movements is the stimulus compatible with?
(which movements have high probability given the evidence?)

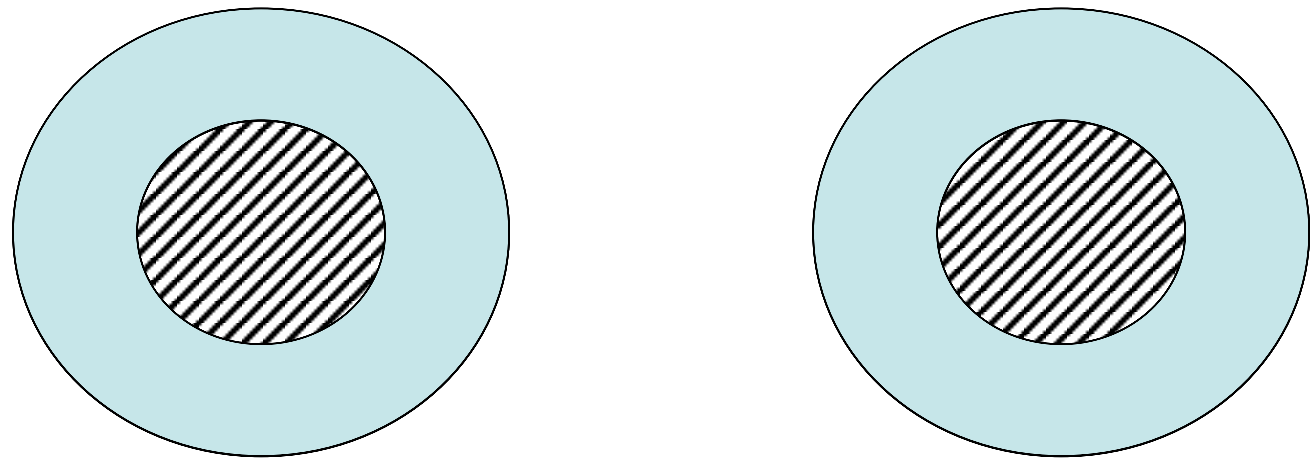


independent pieces of evidence (conditioned on the movement of the object)

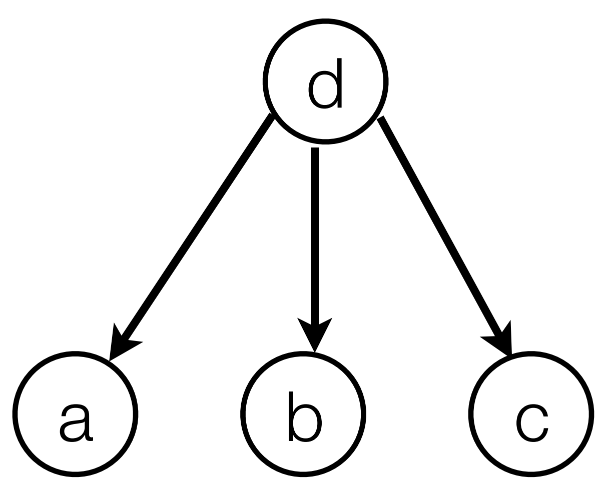
Motion illusions as optimal percepts

Weiss, Simoncelli & Adelson (2002) Nat Neurosci



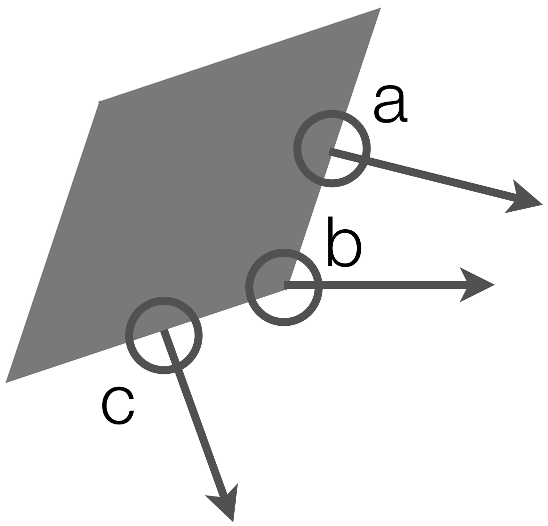


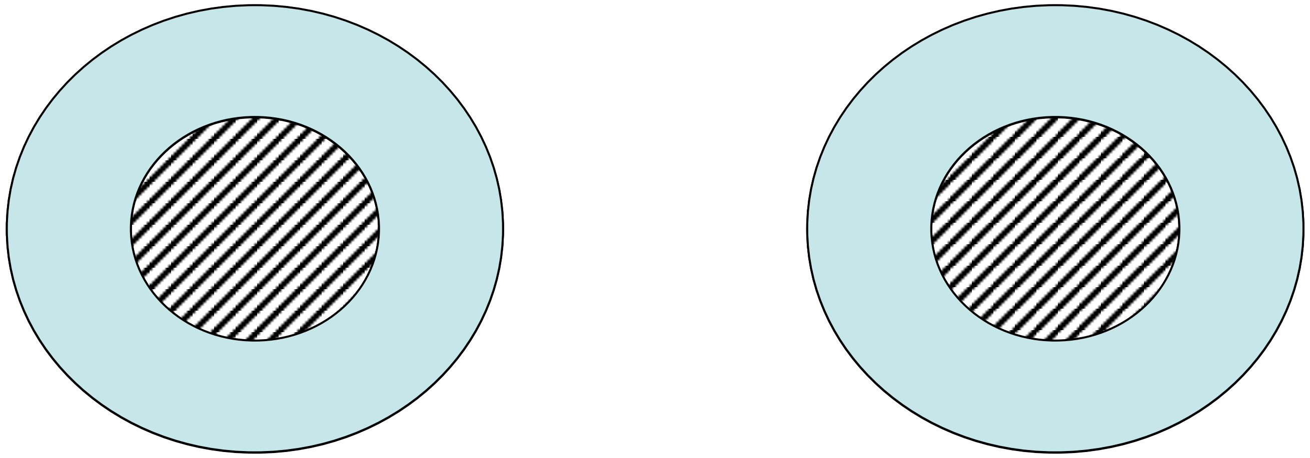
What kind of movements is the stimulus compatible with?
(which movements have high probability given the evidence?)



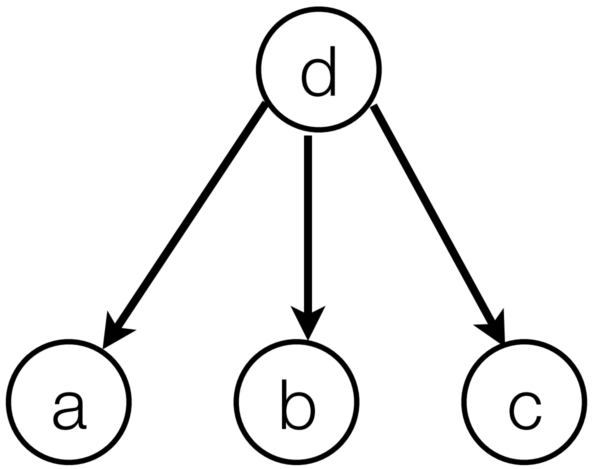
Motion illusions as optimal percepts

Weiss, Simoncelli & Adelson (2002) Nat Neurosci



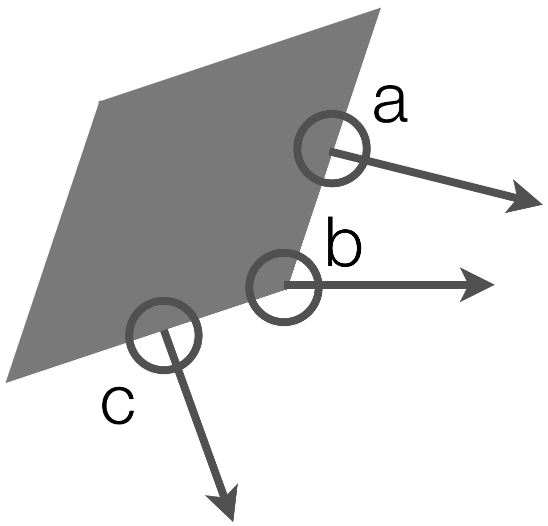


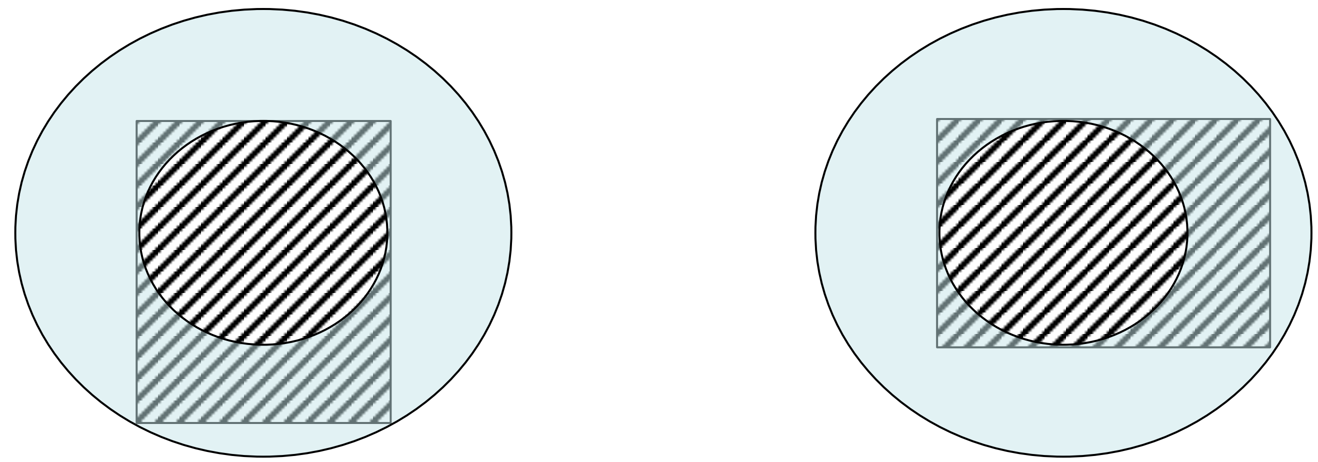
What kind of movements is the stimulus compatible with?
(which movements have high probability given the evidence?)



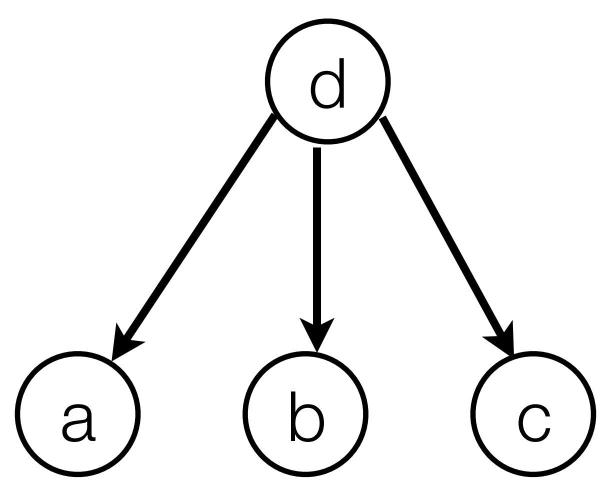
Motion illusions as optimal percepts

Weiss, Simoncelli & Adelson (2002) Nat Neurosci



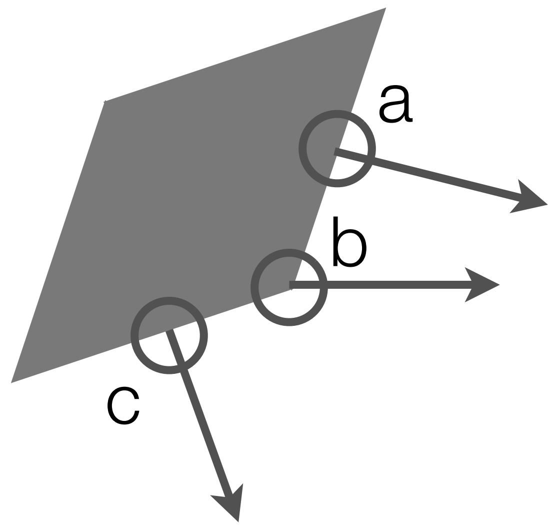


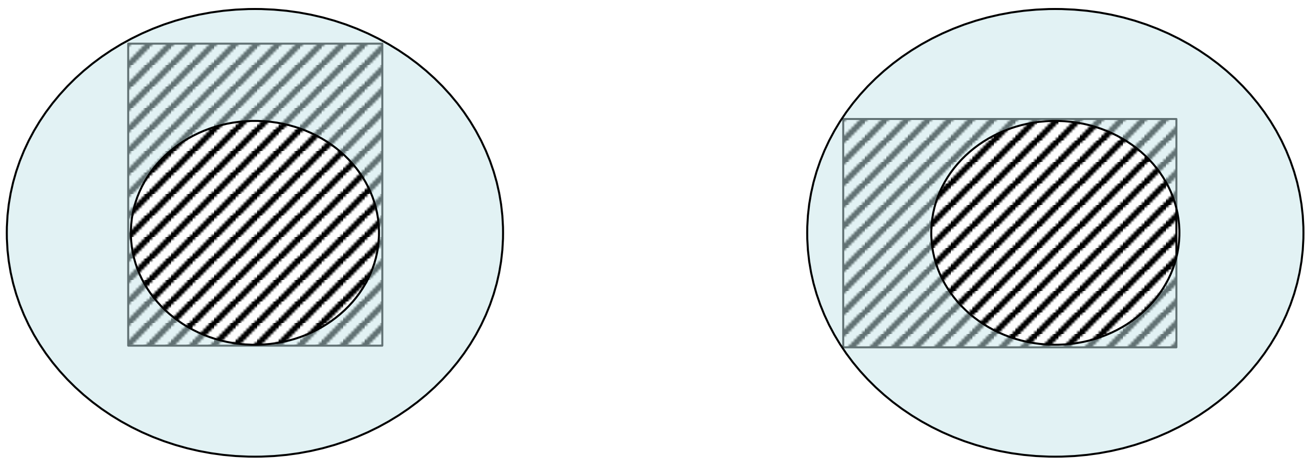
What kind of movements is the stimulus compatible with?
(which movements have high probability given the evidence?)



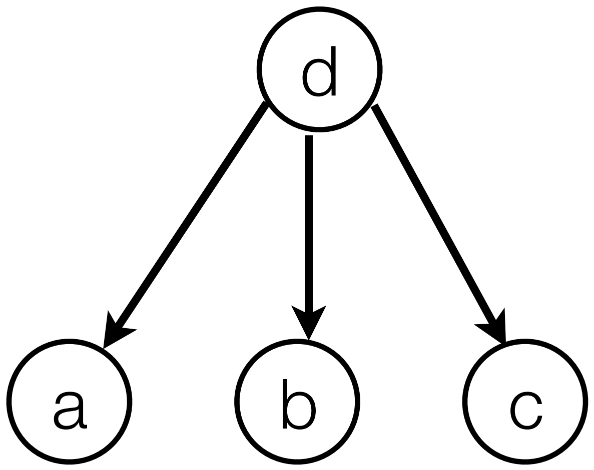
Motion illusions as optimal percepts

Weiss, Simoncelli & Adelson (2002) Nat Neurosci



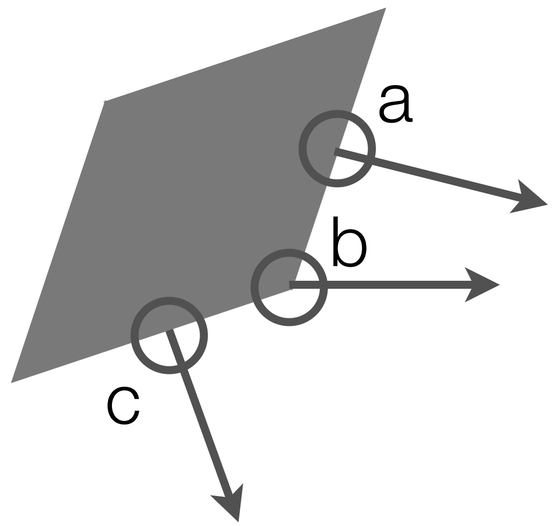


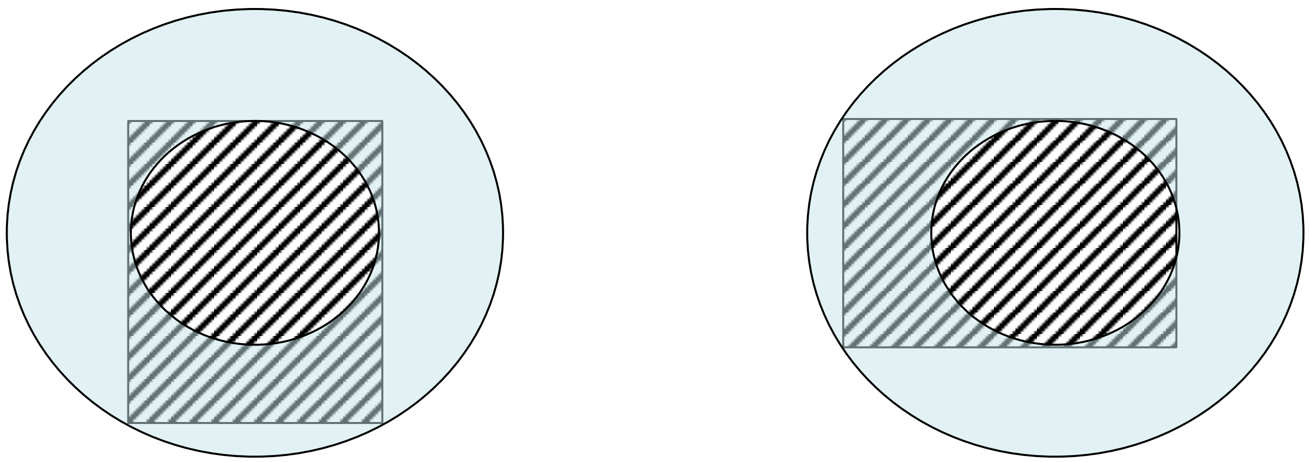
What kind of movements is the stimulus compatible with?
(which movements have high probability given the evidence?)



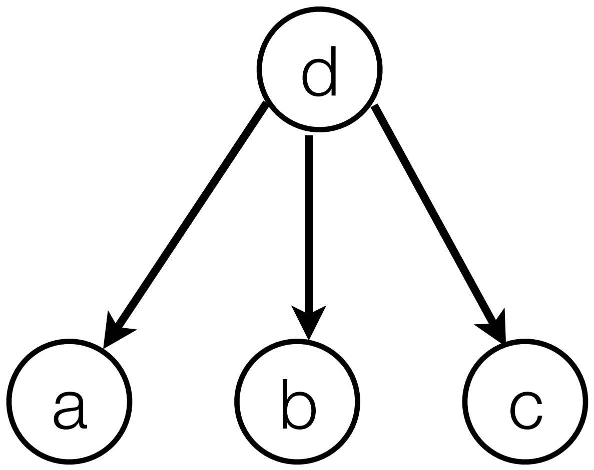
Motion illusions as optimal percepts

Weiss, Simoncelli & Adelson (2002) Nat Neurosci



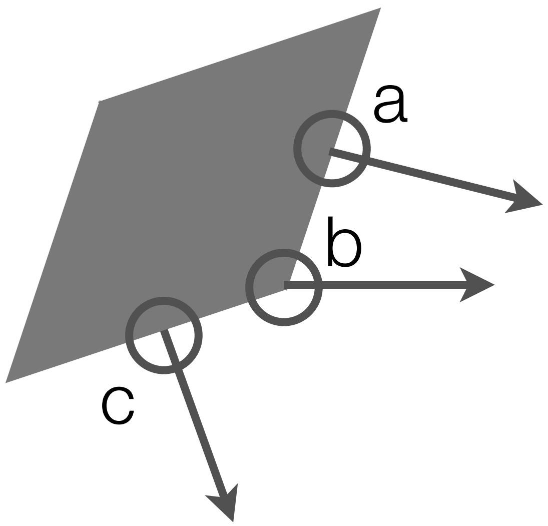


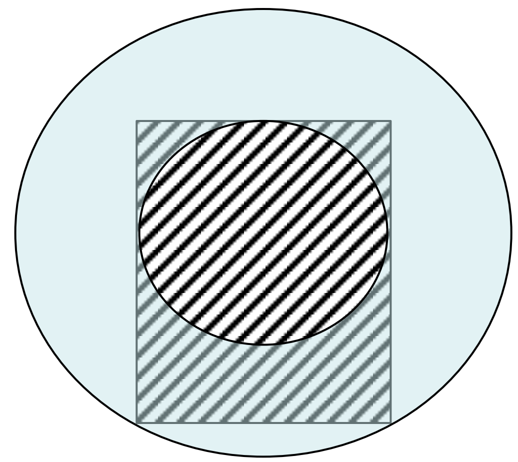
What kind of movements is the stimulus compatible with?
(which movements have high probability given the evidence?)

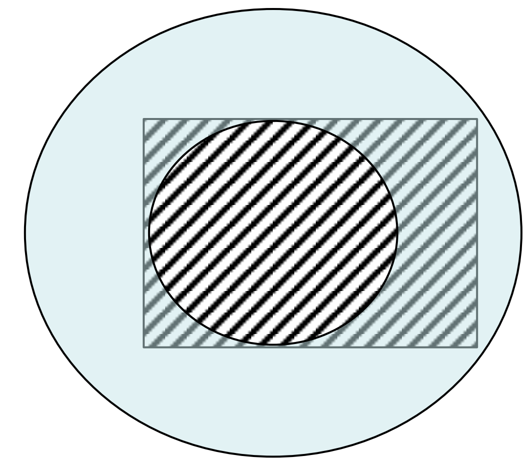


Motion illusions as optimal percepts

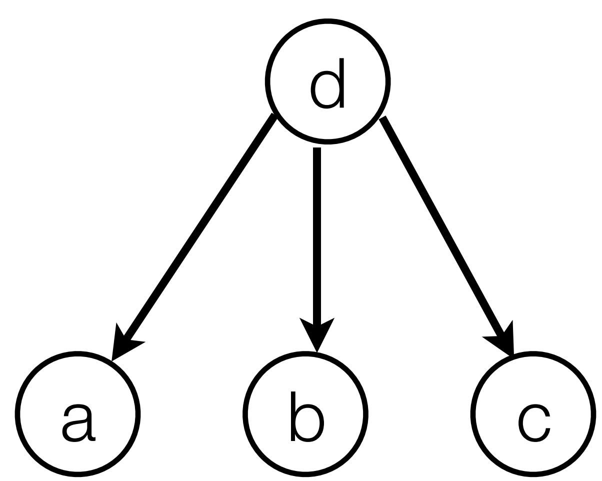
Weiss, Simoncelli & Adelson (2002) Nat Neurosci





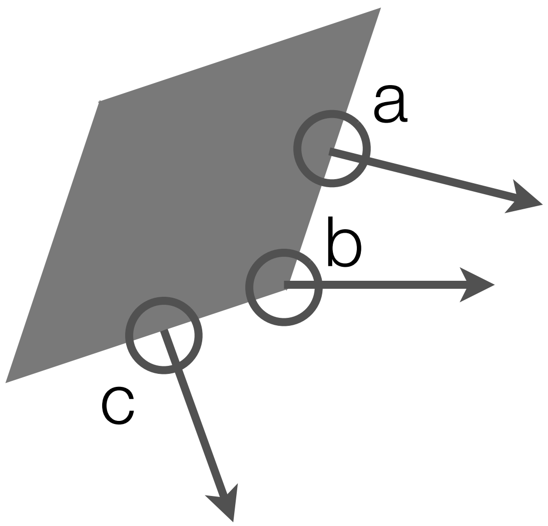


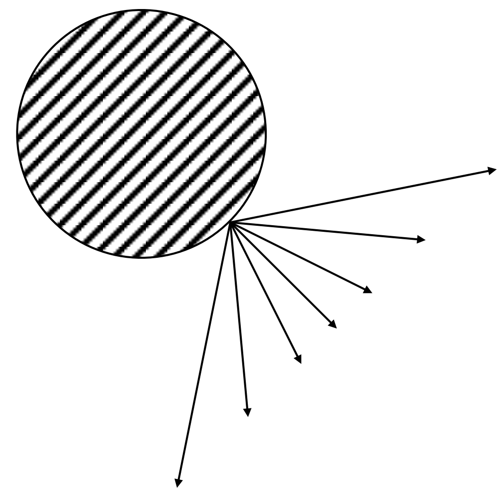
What kind of movements is the stimulus compatible with?
(which movements have high probability given the evidence?)

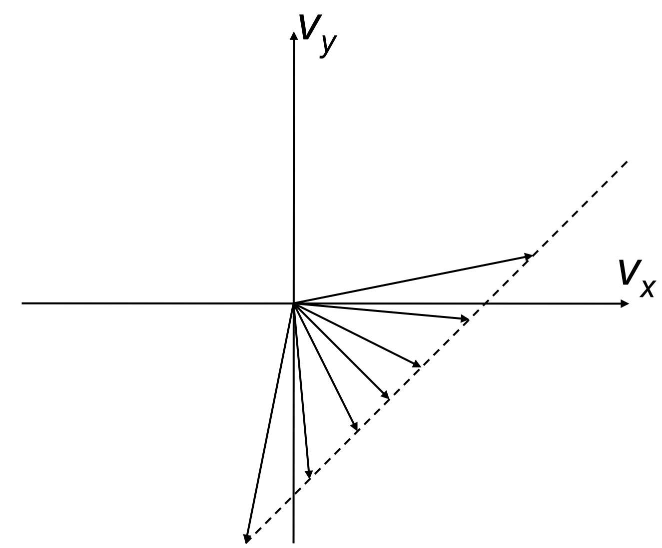


Motion illusions as optimal percepts

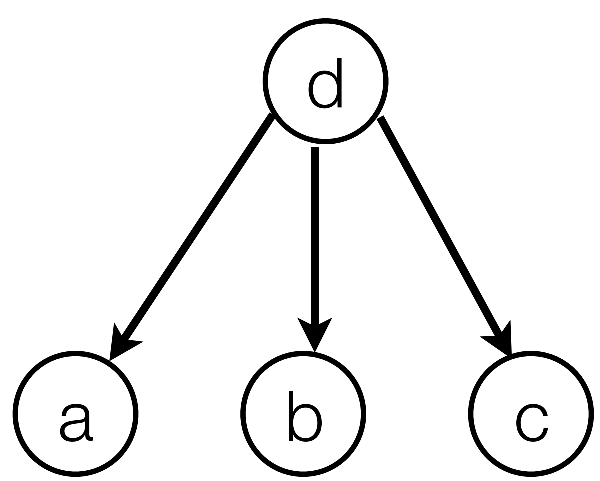
Weiss, Simoncelli & Adelson (2002) Nat Neurosci





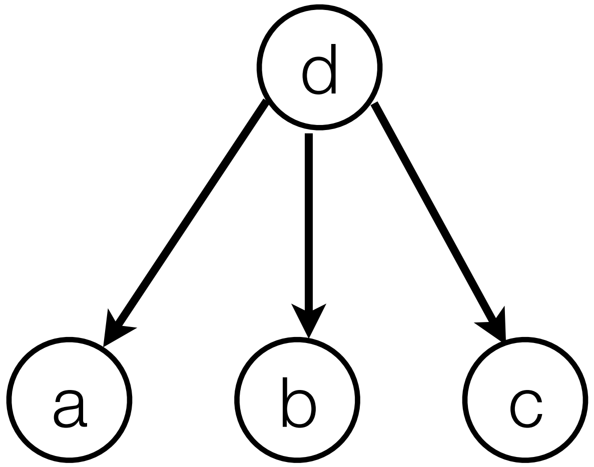


What kind of movements is the stimulus compatible with?
(which movements have high probability given the evidence?)



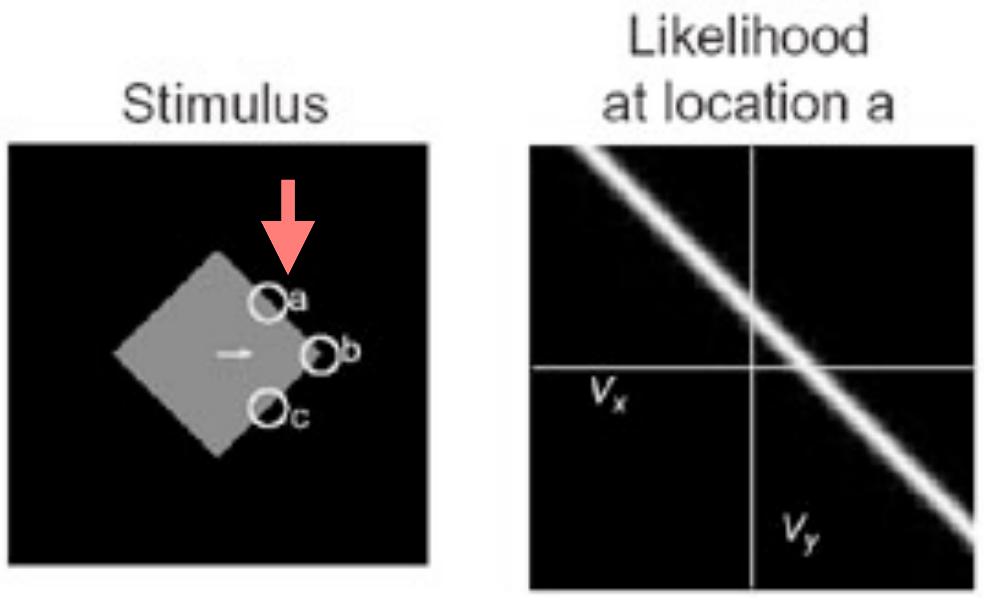
Motion illusions as optimal percepts

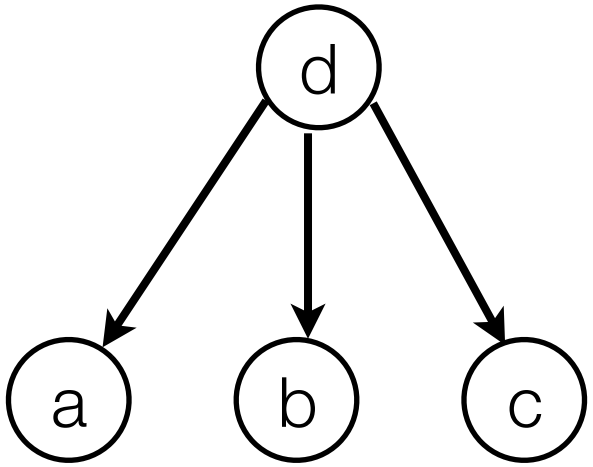
Weiss, Simoncelli & Adelson (2002) Nat Neurosci



Motion illusions as optimal percepts

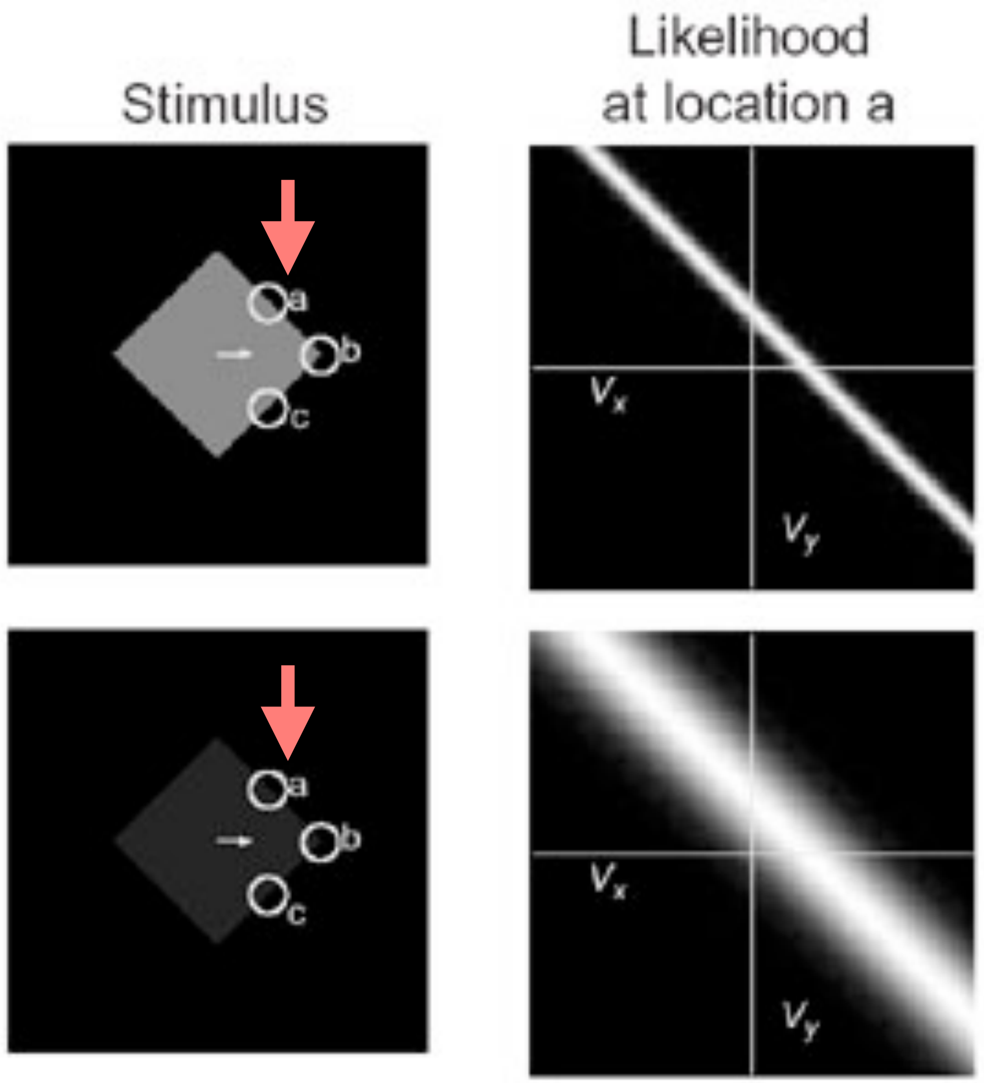
Weiss, Simoncelli & Adelson (2002) Nat Neurosci

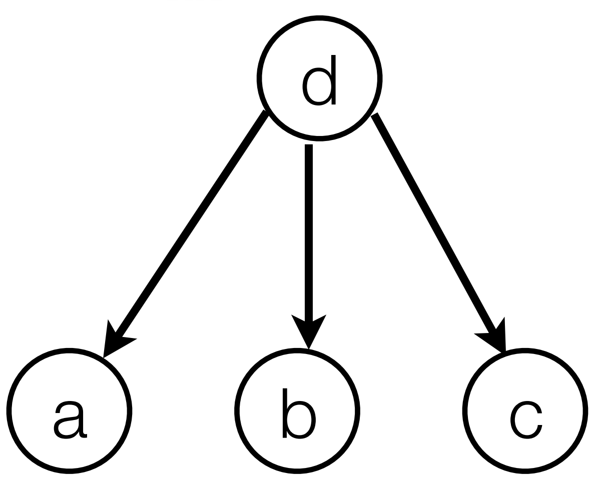




Motion illusions as optimal percepts

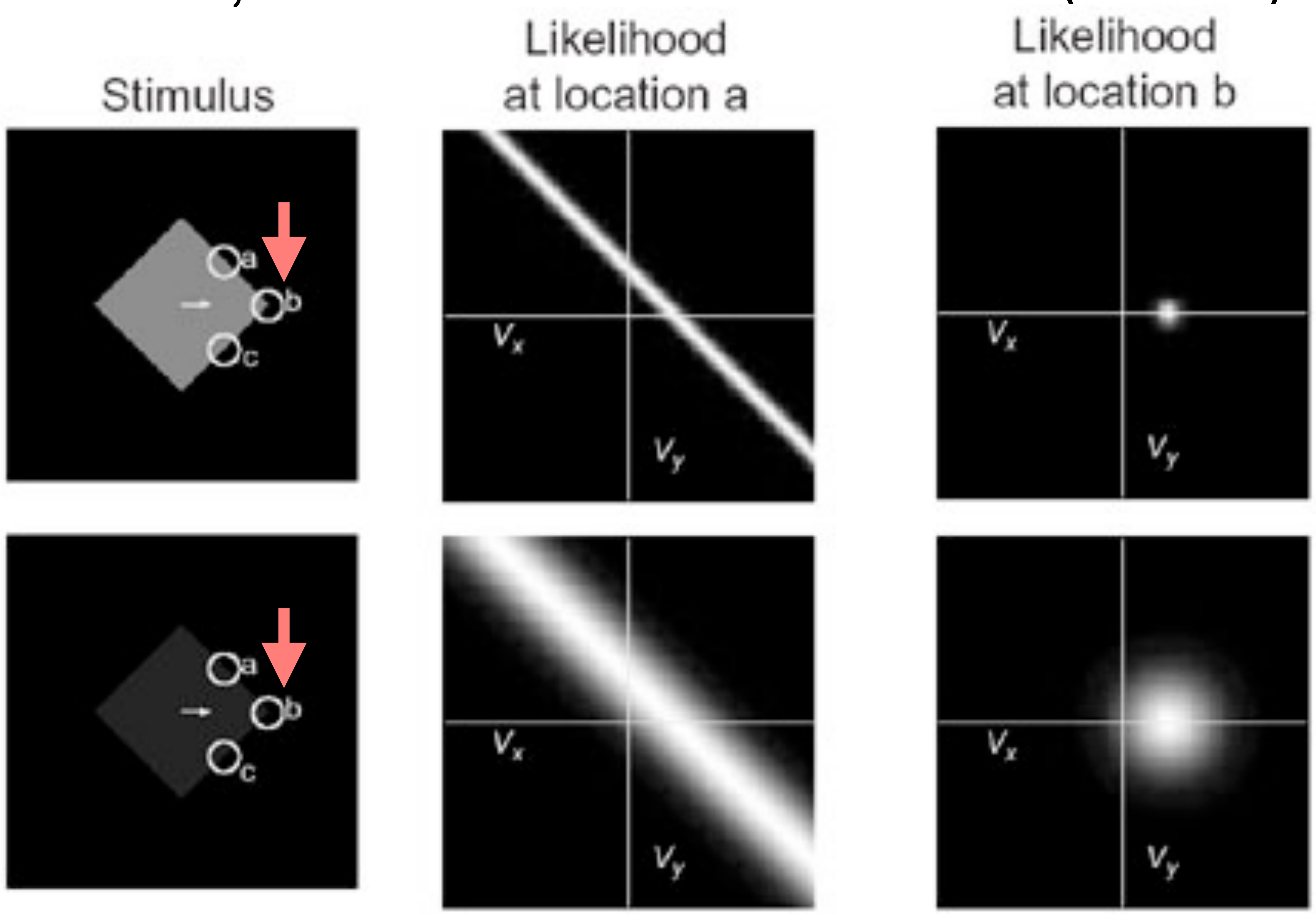
Weiss, Simoncelli & Adelson (2002) Nat Neurosci

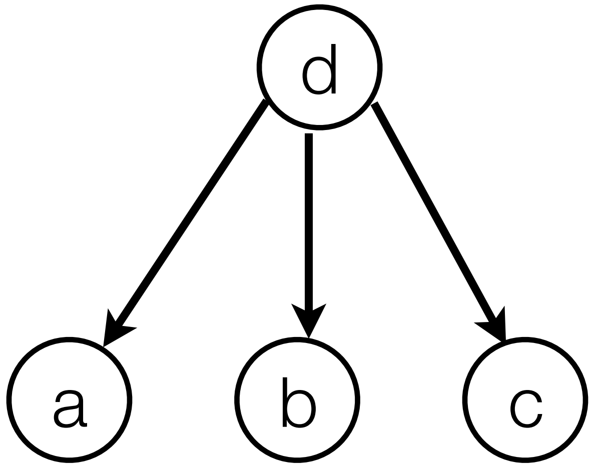




Motion illusions as optimal percepts

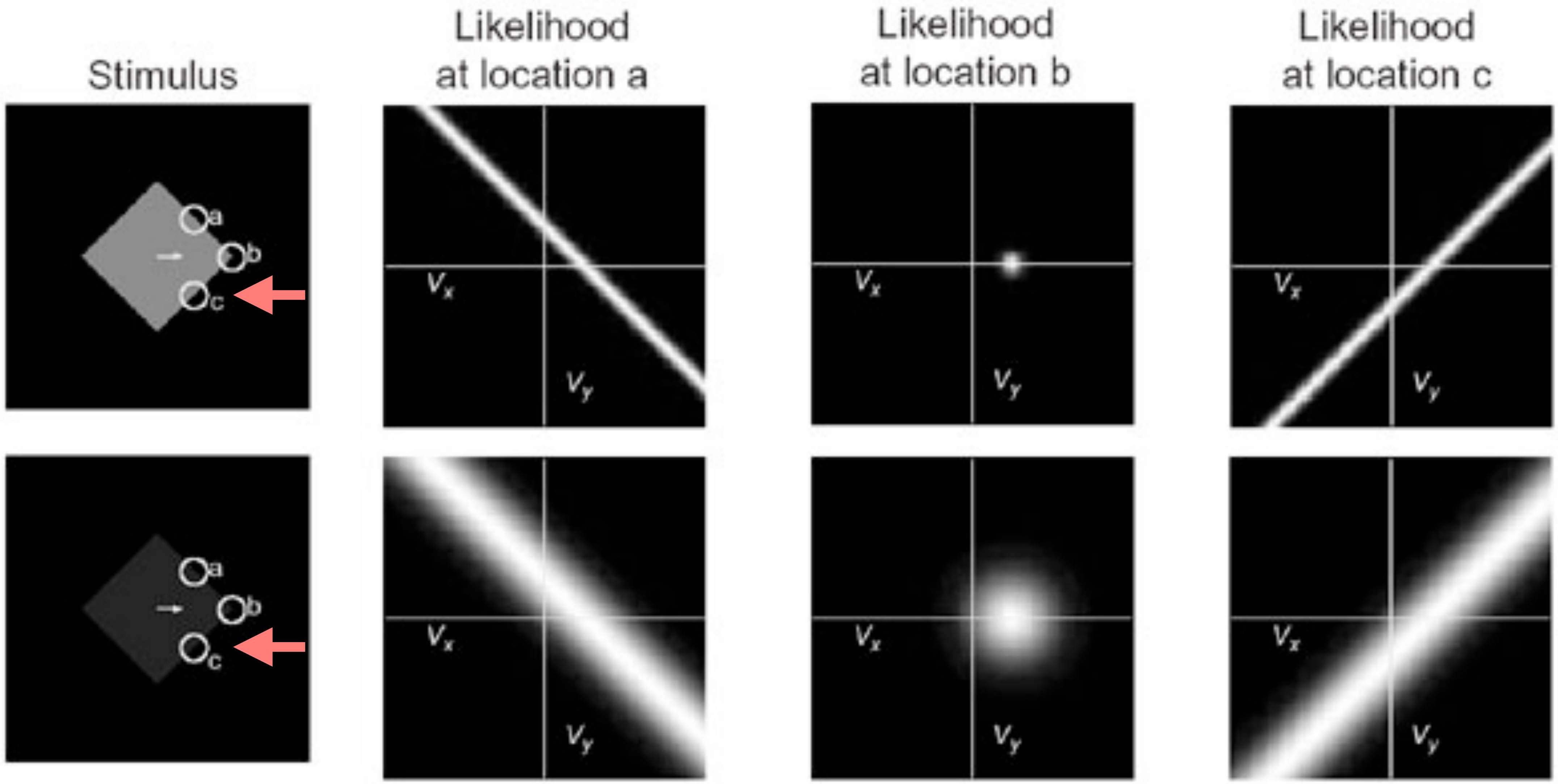
Weiss, Simoncelli & Adelson (2002) Nat Neurosci

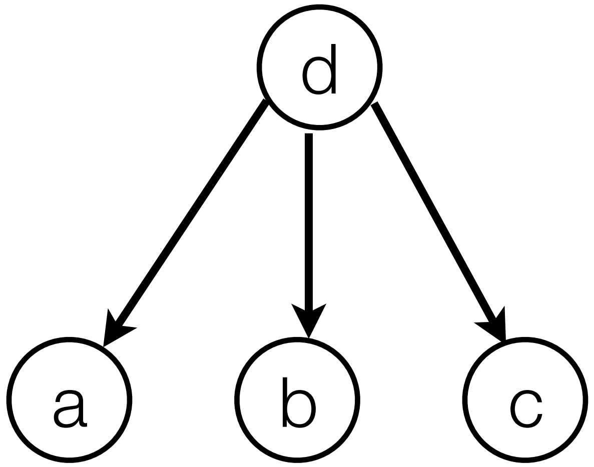




Motion illusions as optimal percepts

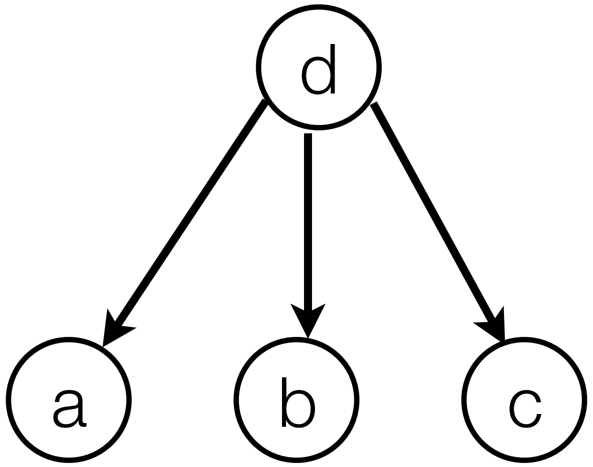
Weiss, Simoncelli & Adelson (2002) Nat Neurosci





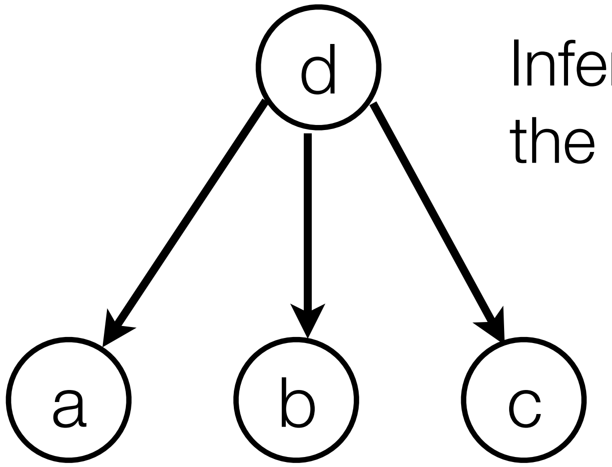
Motion illusions as optimal percepts

Weiss, Simoncelli & Adelson (2002) Nat Neurosci



Motion illusions as optimal percepts

Weiss, Simoncelli & Adelson (2002) Nat Neurosci



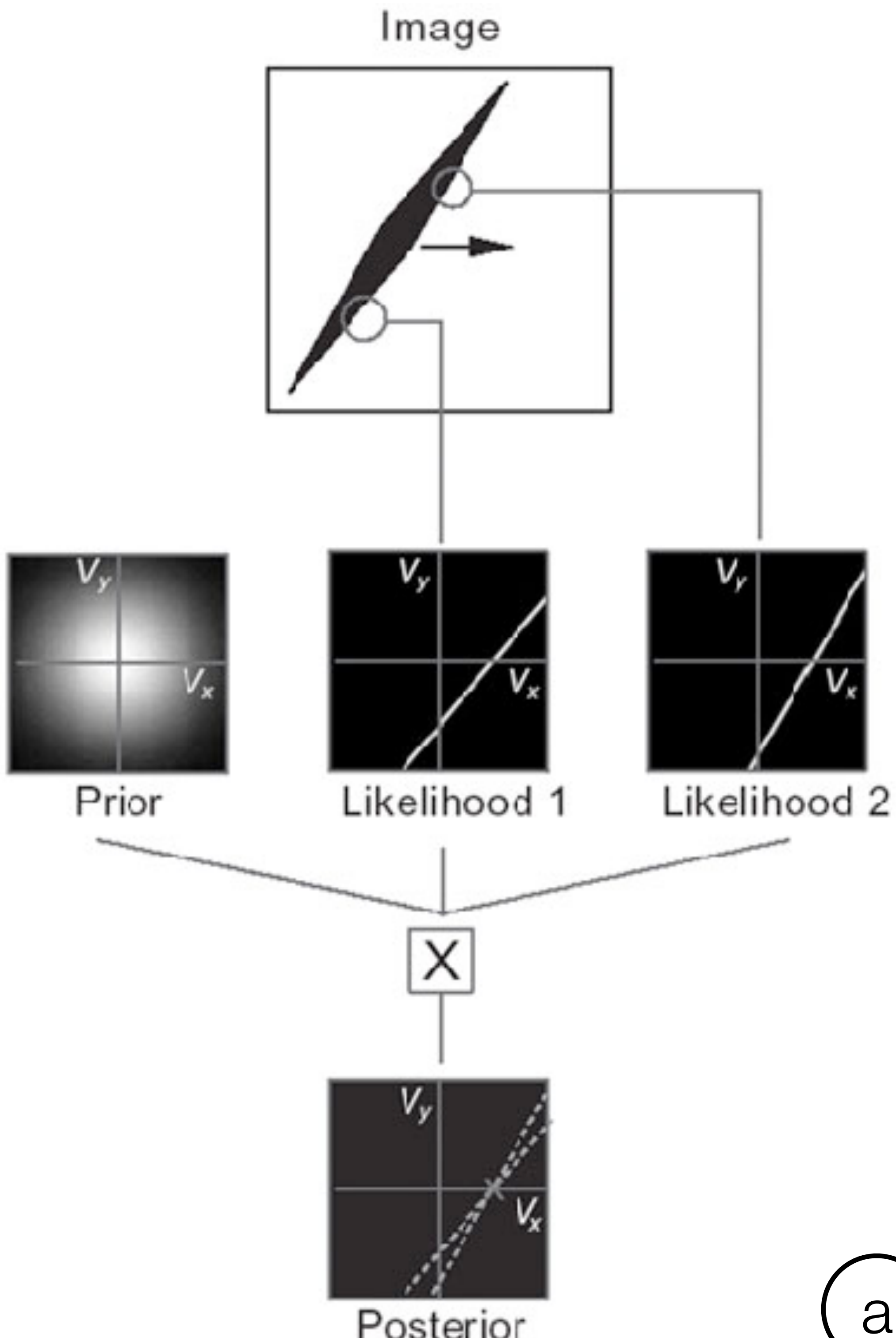
Inference of the joint underlying cause:
the movement of the object

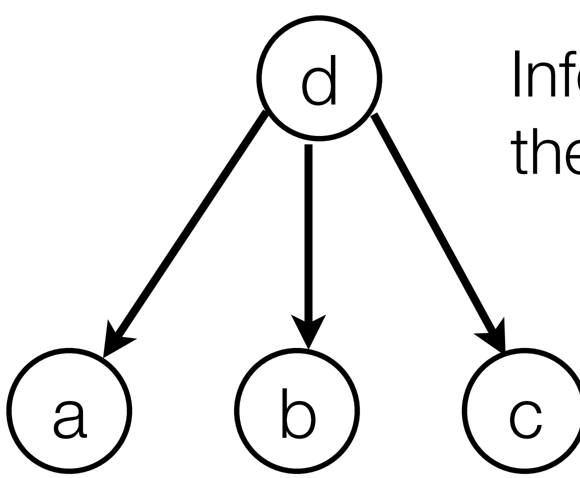
$$P(d \mid a, b, c) = P(a, b, c \mid d) P(d) =$$

$$= P(a \mid d)\, P(b \mid d)\, P(c \mid d)\, P(d)$$

Motion illusions as optimal percepts

Weiss, Simoncelli & Adelson (2002) Nat Neurosci





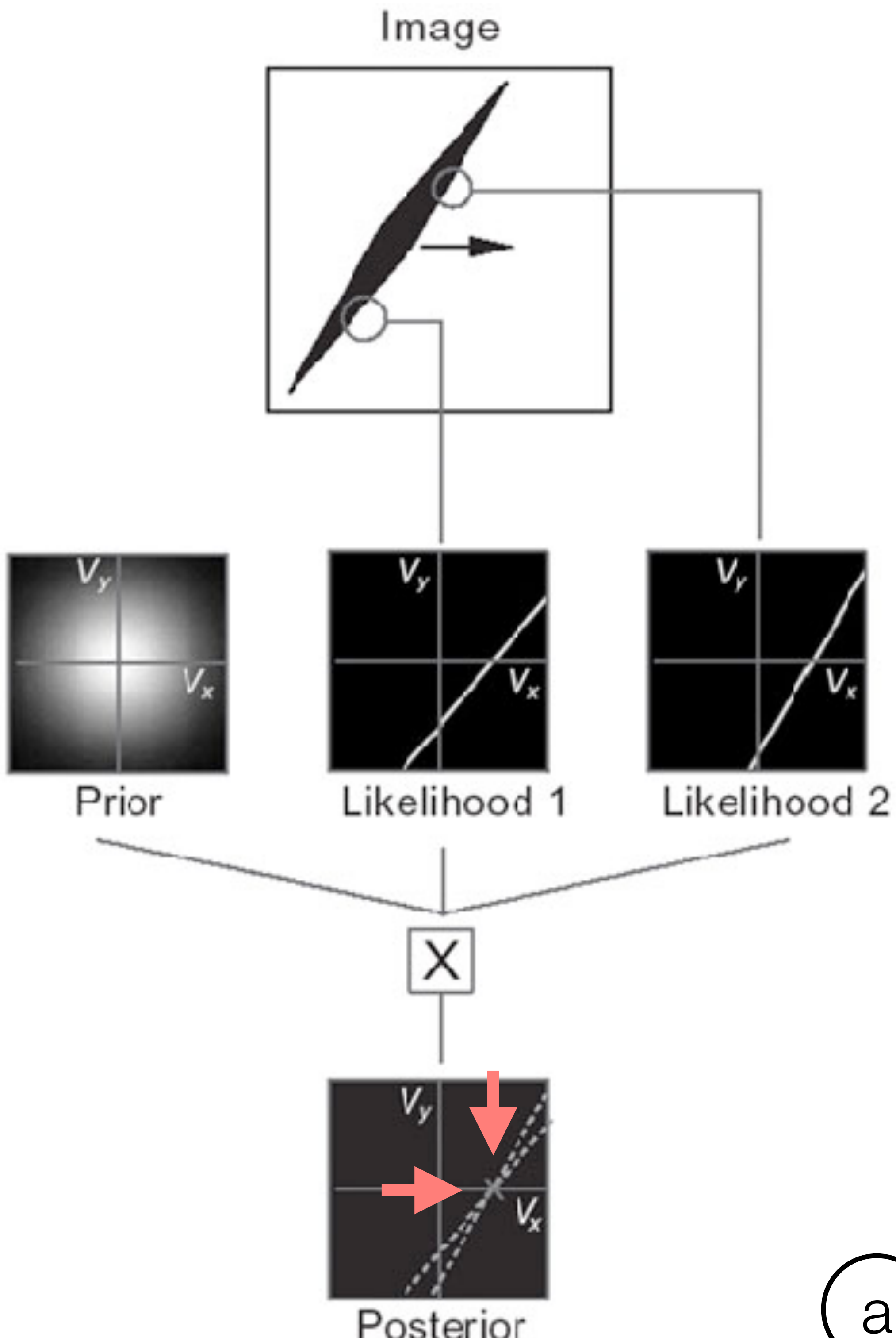
Inference of the joint underlying cause:
the movement of the object

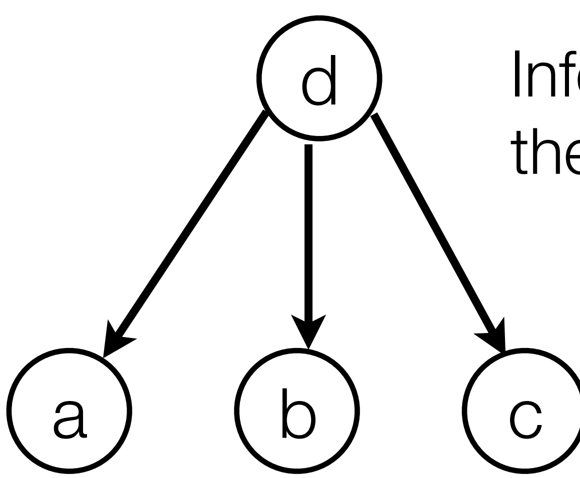
$$P(d \mid a, b, c) = P(a, b, c \mid d) P(d) =$$

$$= P(a \mid d)\, P(b \mid d)\, P(c \mid d)\, P(d)$$

Motion illusions as optimal percepts

Weiss, Simoncelli & Adelson (2002) Nat Neurosci





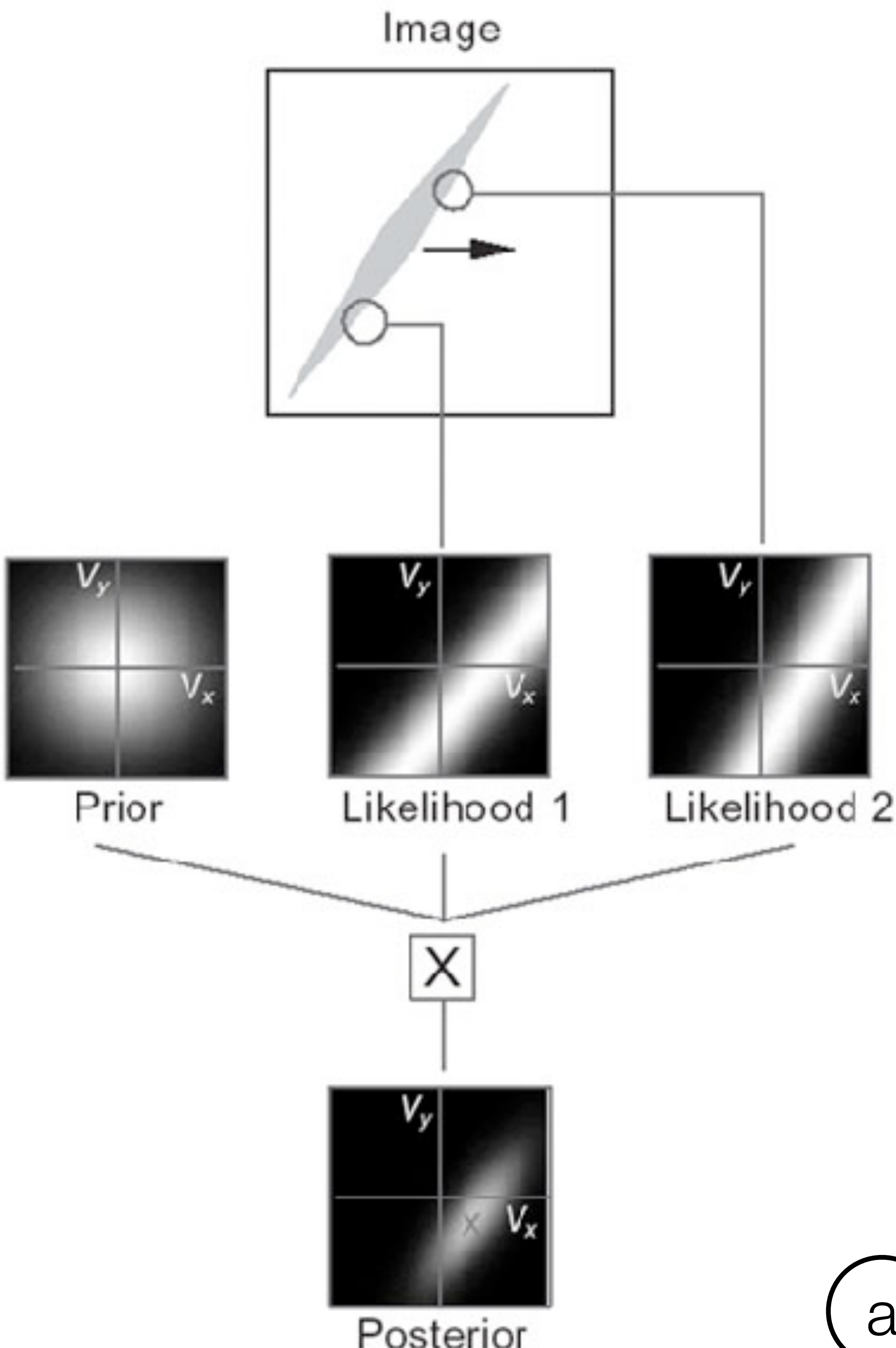
Inference of the joint underlying cause: the movement of the object

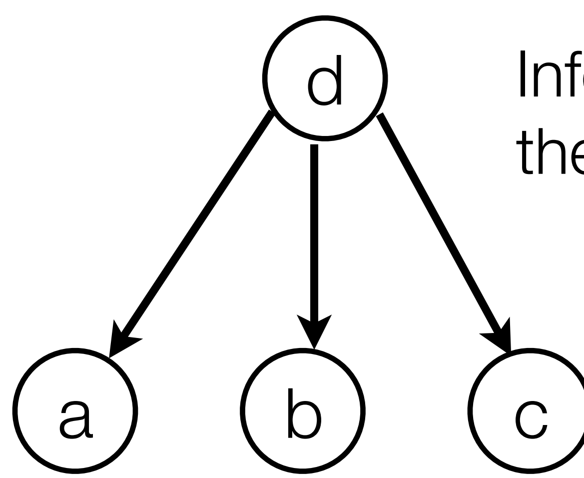
$$P(d\,|\,a,b,c) = P(a,b,c\,|\,d)P(d) =$$

$$= P(a\,|\,d)\,P(b\,|\,d)\,P(c\,|\,d)\,P(d)$$

Motion illusions as optimal percepts

Weiss, Simoncelli & Adelson (2002) Nat Neurosci





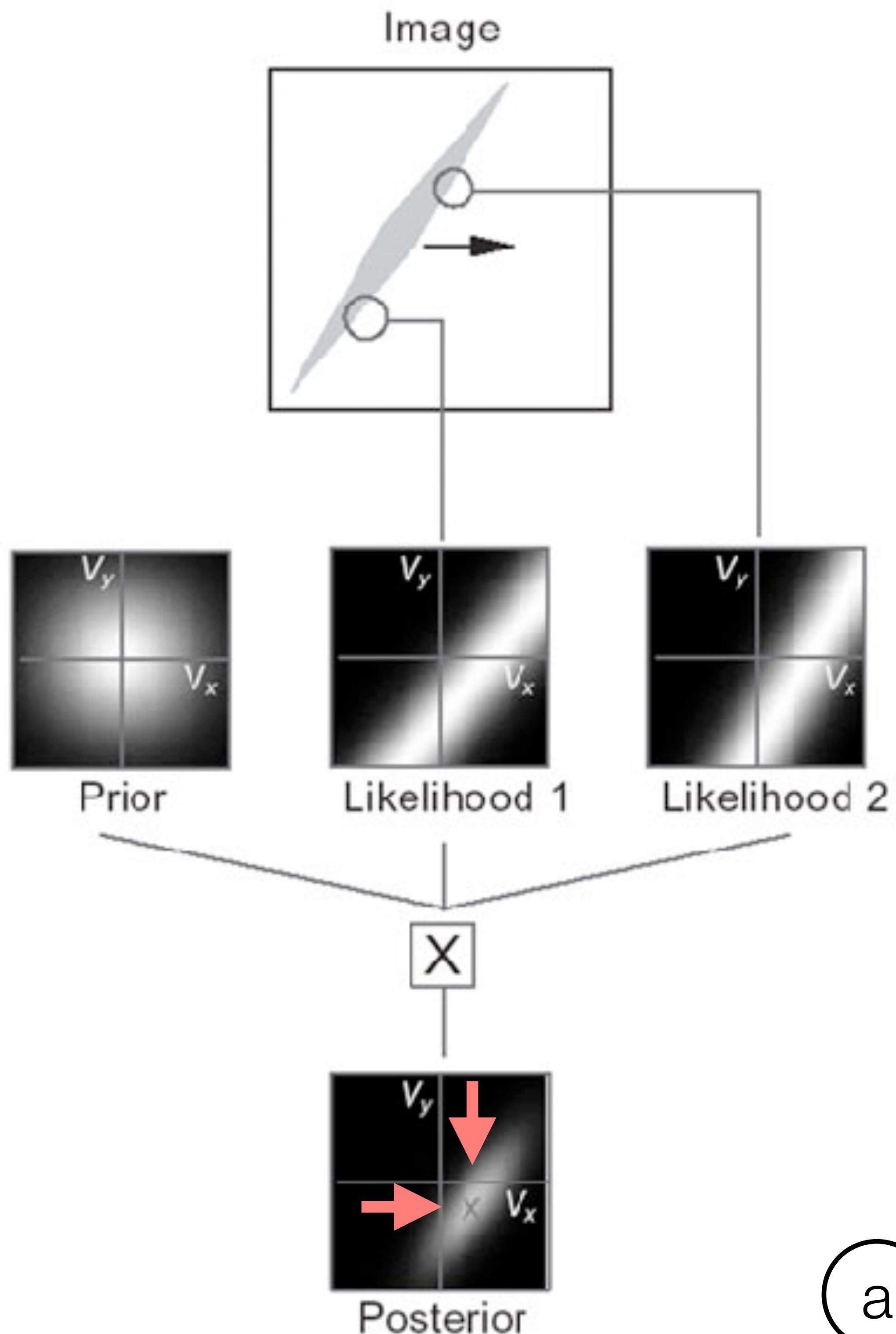
Inference of the joint underlying cause:
the movement of the object

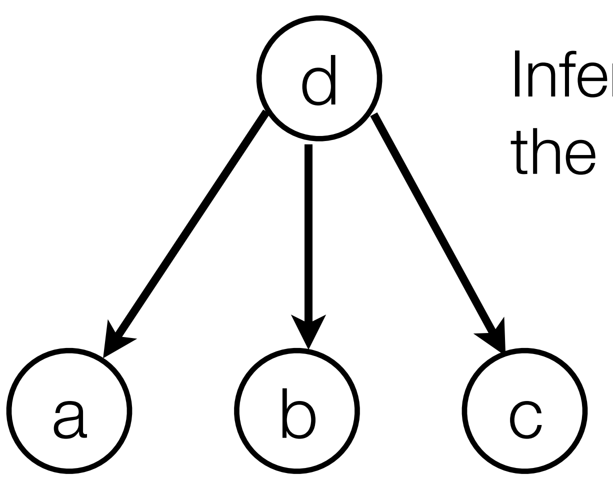
$$P(d \mid a, b, c) = P(a, b, c \mid d)P(d) =$$

$$= P(a \mid d)\, P(b \mid d)\, P(c \mid d)\, P(d)$$

Motion illusions as optimal percepts

Weiss, Simoncelli & Adelson (2002) Nat Neurosci





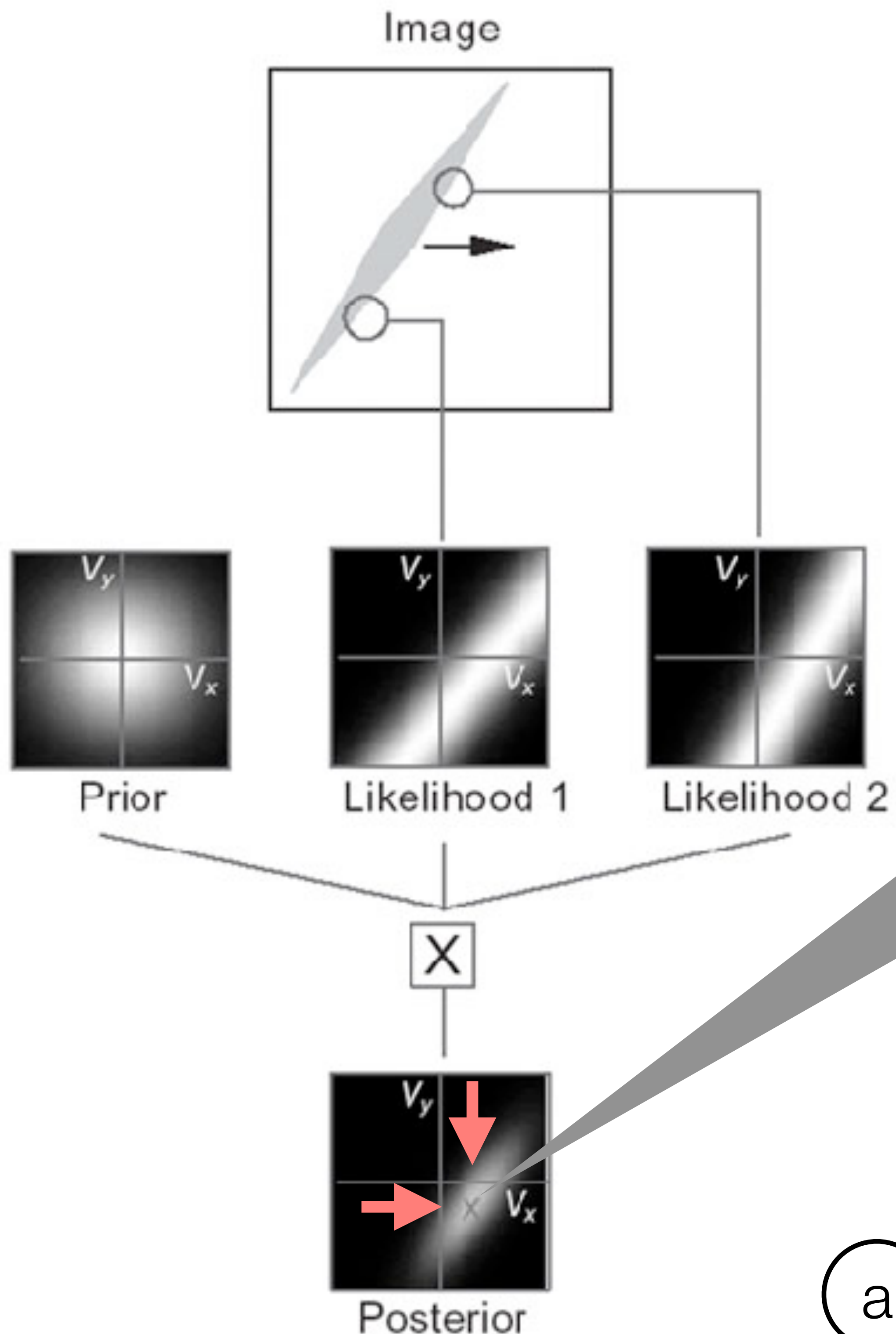
Inference of the joint underlying cause:
the movement of the object

$$P(d\,|\,a,b,c) = P(a,b,c\,|\,d)P(d) =$$

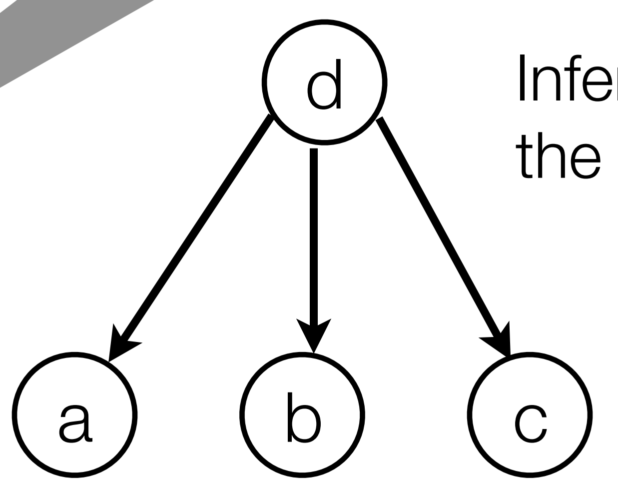
$$= P(a\,|\,d)\,P(b\,|\,d)\,P(c\,|\,d)\,P(d)$$

Motion illusions as optimal percepts

Weiss, Simoncelli & Adelson (2002) Nat Neurosci



a vertical velocity component appears



Inference of the joint underlying cause: the movement of the object

$$P(d\,|\,a,b,c) = P(a,b,c\,|\,d)P(d) =$$

$$= P(a\,|\,d)\,P(b\,|\,d)\,P(c\,|\,d)\,P(d)$$

Motion illusions as optimal percepts

Weiss, Simoncelli & Adelson (2002)

- Illusion emerges as a result of optimal computations under uncertainty
- The prior the experiment hints at reflects a simple regularity of the environment
- The prior is very generic, no subjective aspects can be revealed

Markov Chain Monte Carlo with people

Sanborn & Griffiths (2008) NIPS

IDEA:

Markov Chain Monte Carlo with people

Sanborn & Griffiths (2008) NIPS

IDEA:

- The model of a particular domain of knowledge can be directly corresponded to a prior

Markov Chain Monte Carlo with people

Sanborn & Griffiths (2008) NIPS

IDEA:

- The model of a particular domain of knowledge can be directly corresponded to a prior
- The prior distribution contains the information we now about a specific domain and we are relying on this prior to make decisions

Markov Chain Monte Carlo with people

Sanborn & Griffiths (2008) NIPS

IDEA:

- The model of a particular domain of knowledge can be directly corresponded to a prior
- The prior distribution contains the information we now about a specific domain and we are relying on this prior to make decisions
- Samples can represent a probability distribution, which obviously includes the prior distribution as well

Markov Chain Monte Carlo with people

Sanborn & Griffiths (2008) NIPS

IDEA:

- The model of a particular domain of knowledge can be directly corresponded to a prior

- The prior distribution contains the information we now about a specific domain and we are relying on this prior to make decisions

- Samples can represent a probability distribution, which obviously includes the prior distribution as well

- Design an experiment where the decisions of humans produces samples from the prior distribution

Markov Chain Monte Carlo with people

Sanborn & Griffiths (2008) NIPS

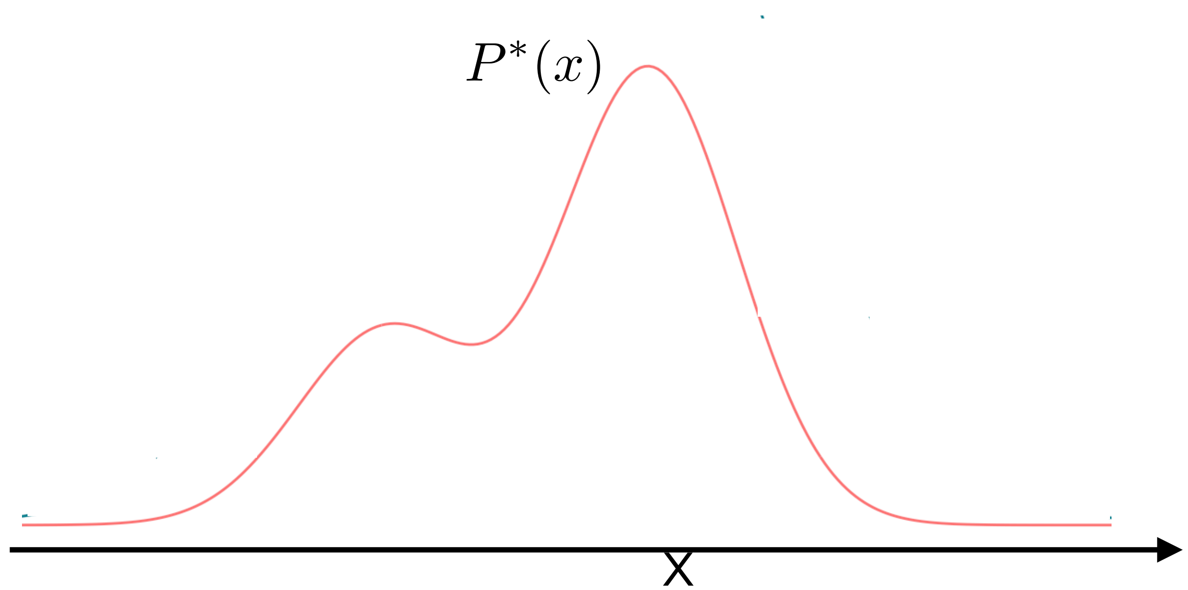
IDEA:

- The model of a particular domain of knowledge can be directly corresponded to a prior
- The prior distribution contains the information we now about a specific domain and we are relying on this prior to make decisions
- Samples can represent a probability distribution, which obviously includes the prior distribution as well
- Design an experiment where the decisions of humans produces samples from the prior distribution
- The sequence of samples will reveal the prior distribution

Markov Chain Monte Carlo with people

Sanborn & Griffiths (2008) NIPS

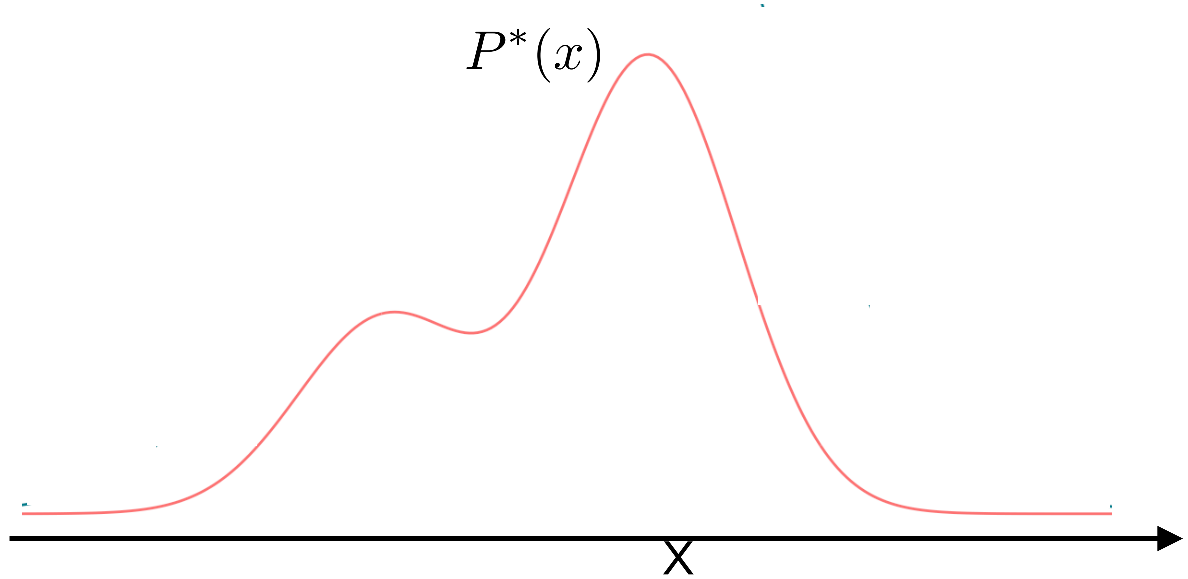
- equivalence between **choice behaviour** and a **Metropolis Hastings sampler** can be established



Markov Chain Monte Carlo with people

Sanborn & Griffiths (2008) NIPS

- equivalence between **choice behaviour** and a **Metropolis Hastings sampler** can be established



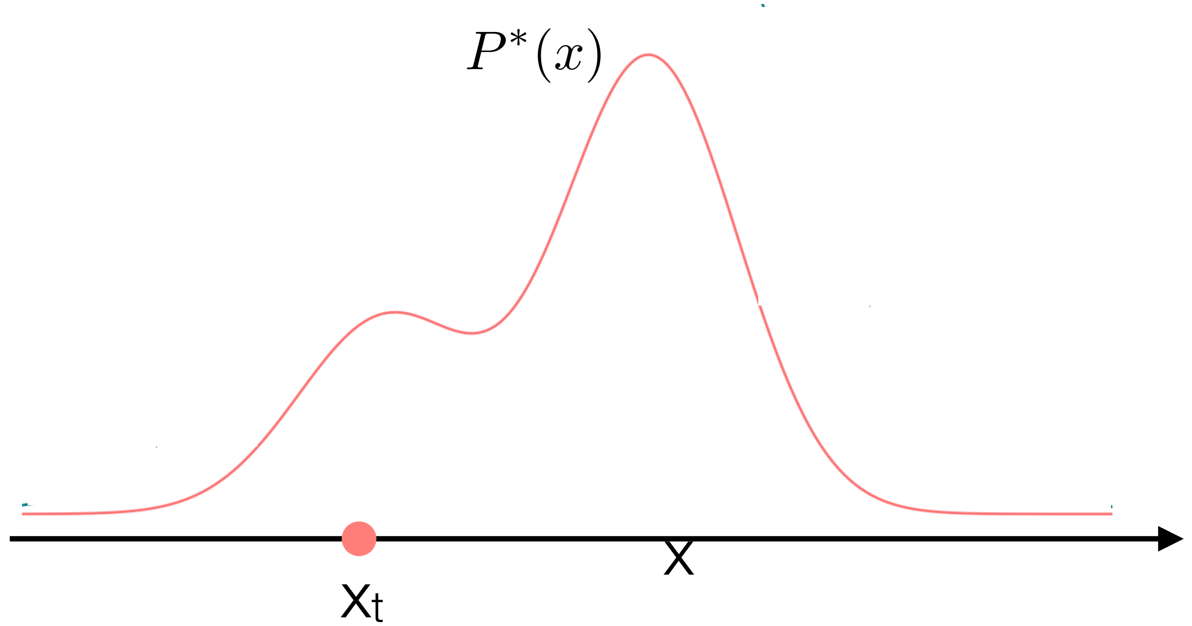
Alternative phrasing of acceptance probability: Barker dynamics

$$a(x_{t+1}, x_t) = P^*(x_{t+1}) / (P^*(x_{t+1}) + P^*(x_t))$$

Markov Chain Monte Carlo with people

Sanborn & Griffiths (2008) NIPS

- equivalence between **choice behaviour** and a **Metropolis Hastings sampler** can be established



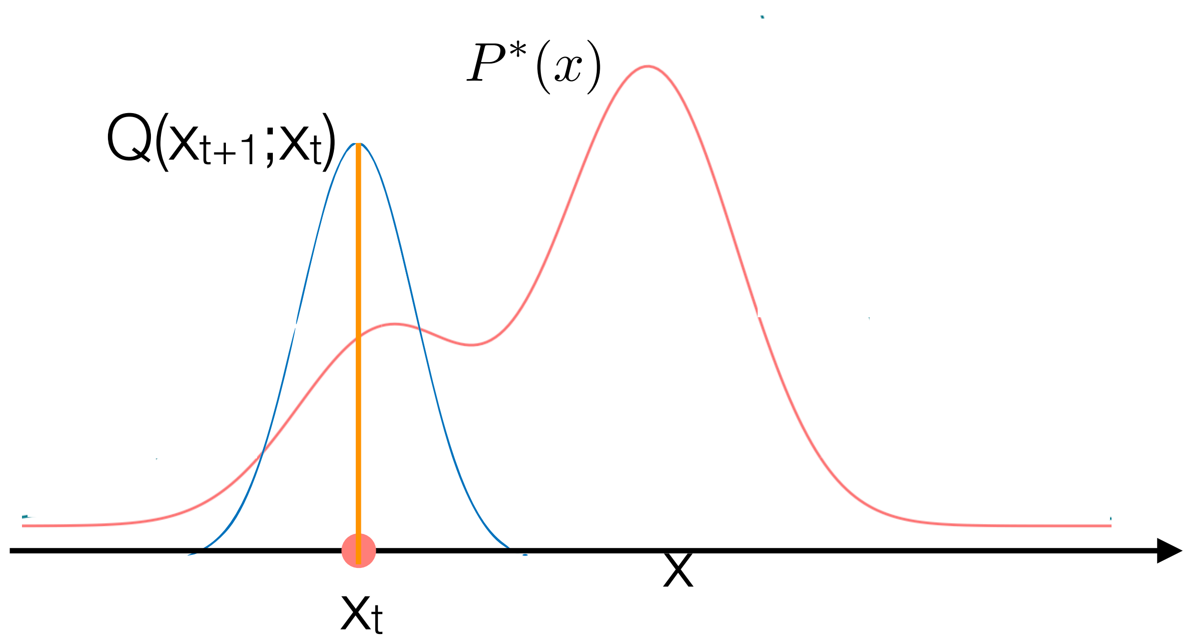
Alternative phrasing of acceptance probability: Barker dynamics

$$a(x_{t+1}, x_t) = P^*(x_{t+1}) / (P^*(x_{t+1}) + P^*(x_t))$$

Markov Chain Monte Carlo with people

Sanborn & Griffiths (2008) NIPS

- equivalence between **choice behaviour** and a **Metropolis Hastings sampler** can be established



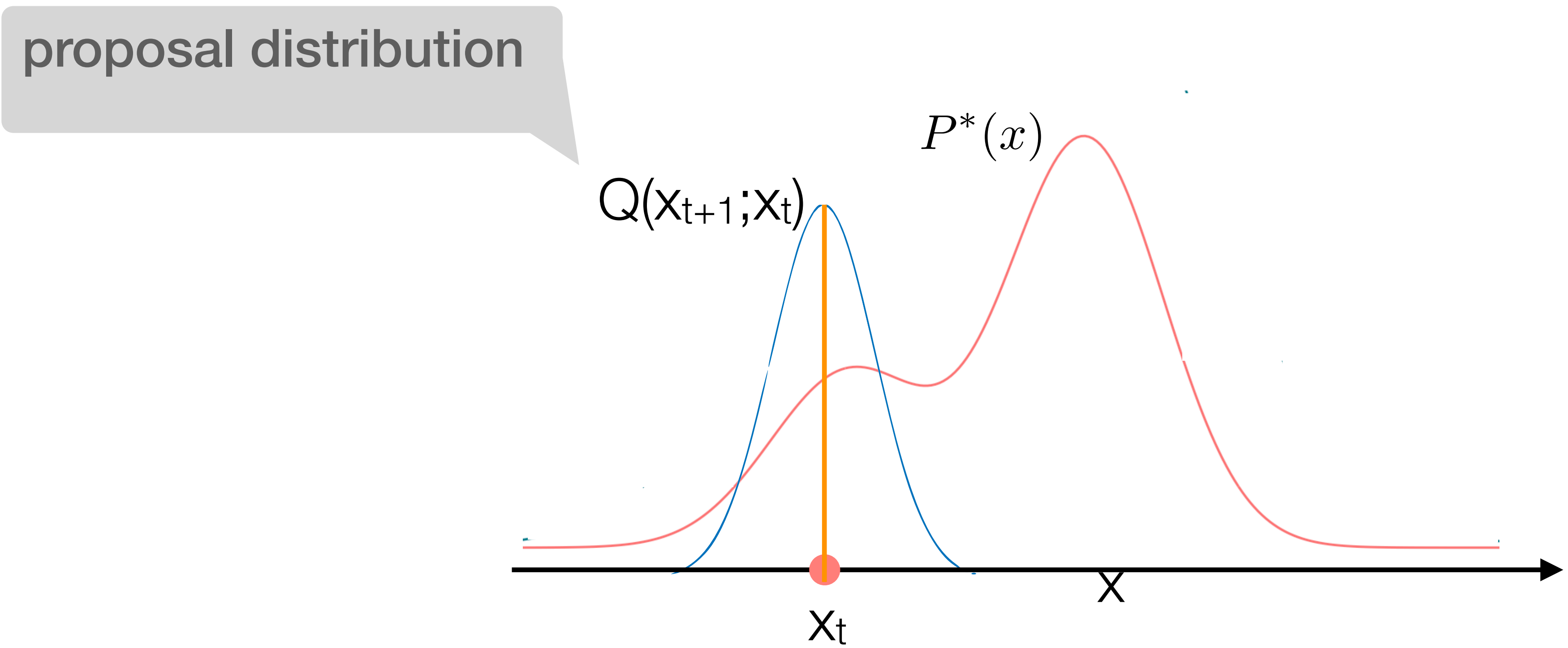
Alternative phrasing of acceptance probability: Barker dynamics

$$a(x_{t+1}, x_t) = P^*(x_{t+1}) / (P^*(x_{t+1}) + P^*(x_t))$$

Markov Chain Monte Carlo with people

Sanborn & Griffiths (2008) NIPS

- equivalence between **choice behaviour** and a **Metropolis Hastings sampler** can be established



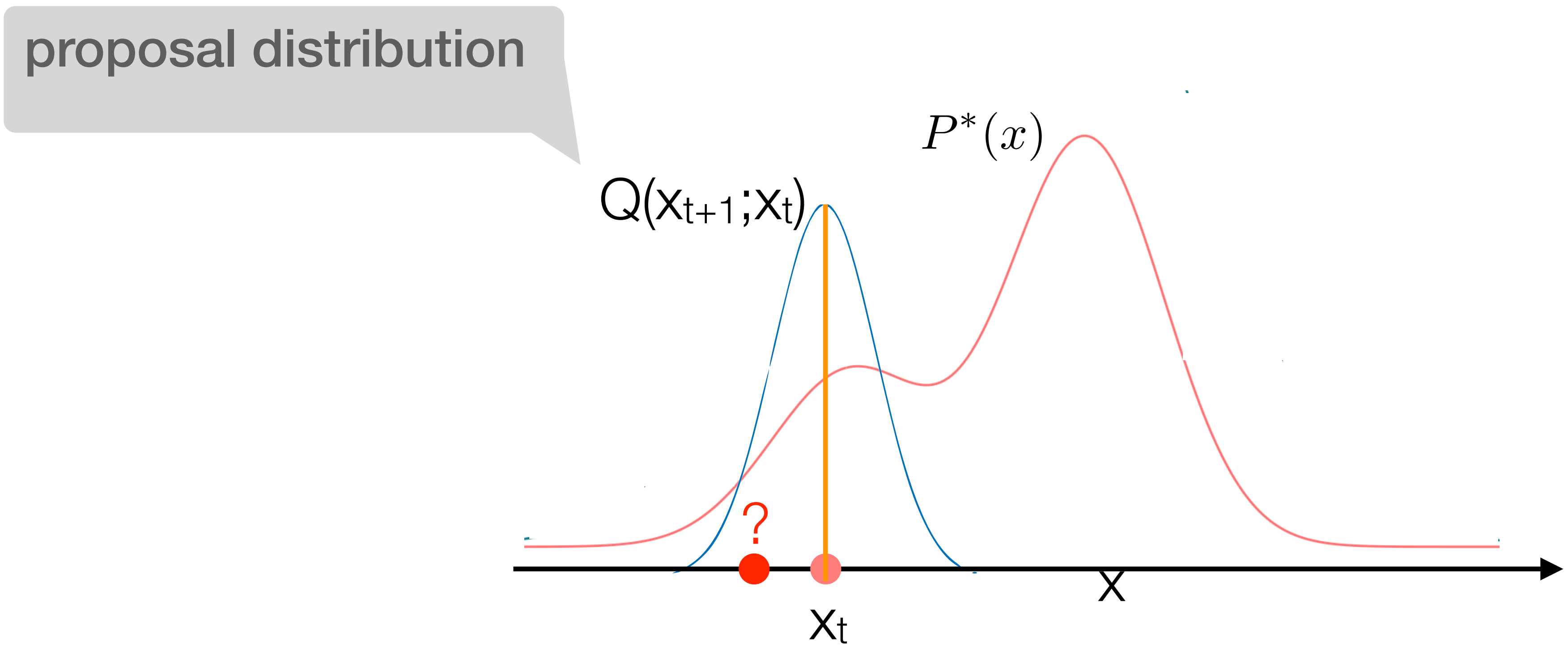
Alternative phrasing of acceptance probability: Barker dynamics

$$a(x_{t+1}, x_t) = P^*(x_{t+1}) / (P^*(x_{t+1}) + P^*(x_t))$$

Markov Chain Monte Carlo with people

Sanborn & Griffiths (2008) NIPS

- equivalence between **choice behaviour** and a **Metropolis Hastings sampler** can be established



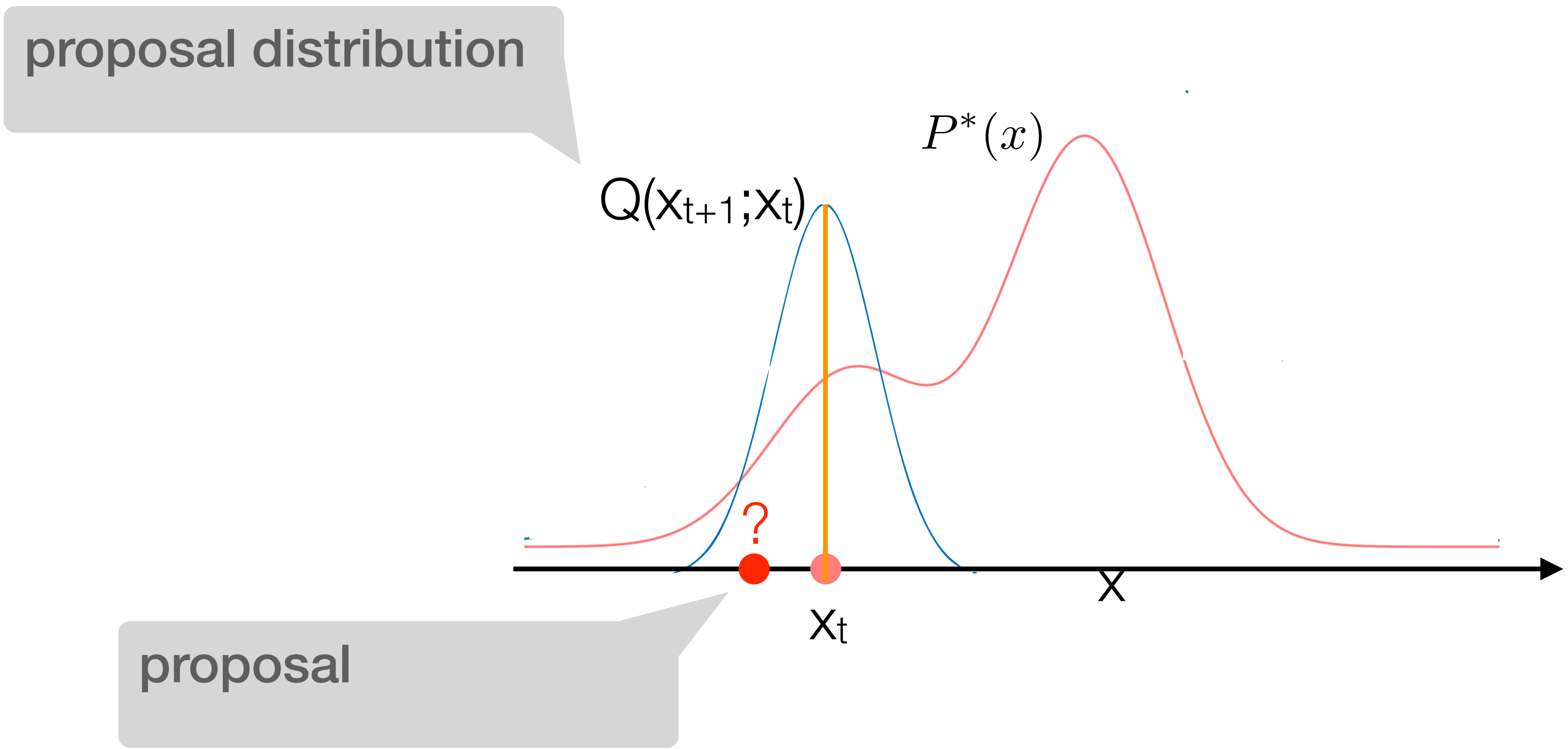
Alternative phrasing of acceptance probability: Barker dynamics

$$a(x_{t+1}, x_t) = P^*(x_{t+1}) / (P^*(x_{t+1}) + P^*(x_t))$$

Markov Chain Monte Carlo with people

Sanborn & Griffiths (2008) NIPS

- equivalence between **choice behaviour** and a **Metropolis Hastings sampler** can be established



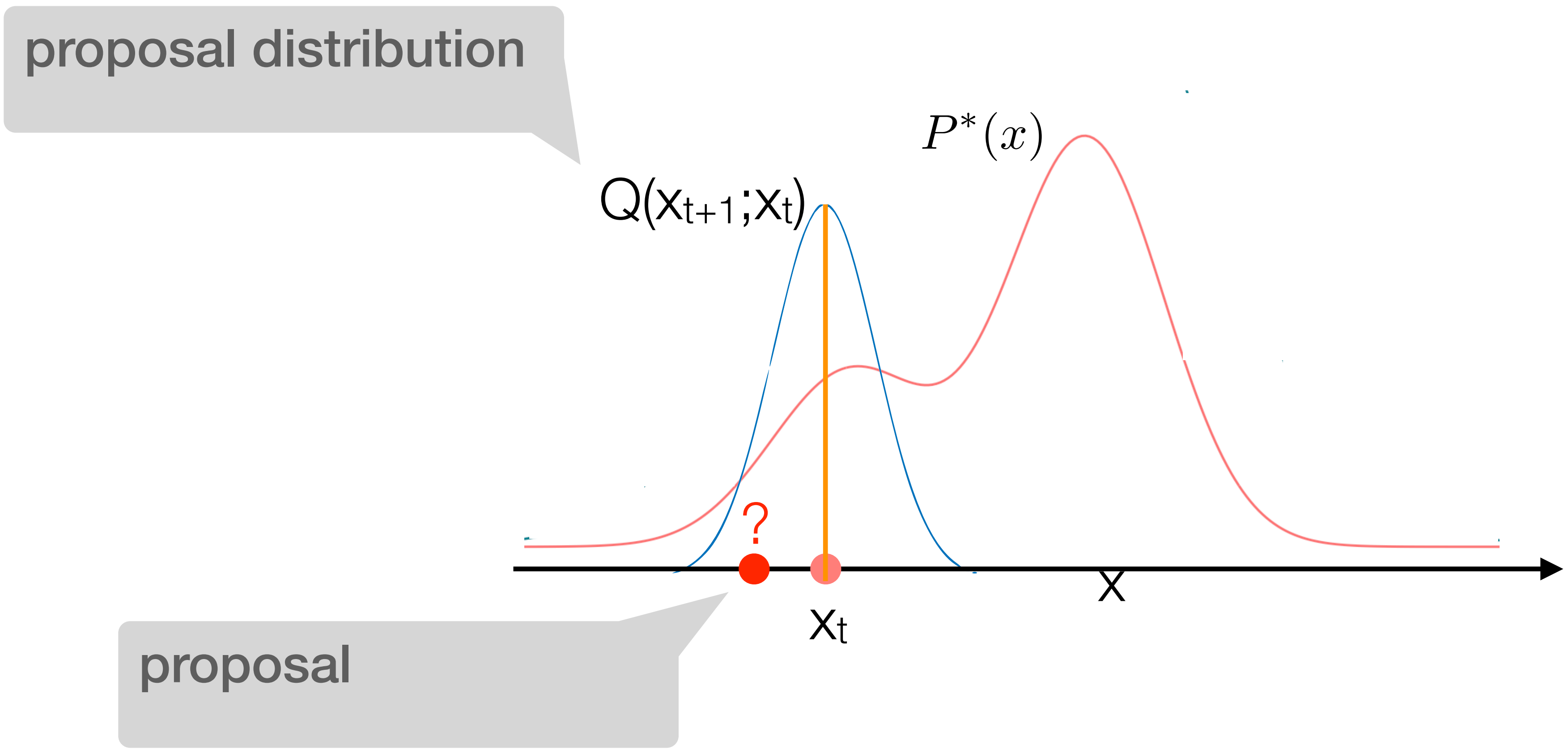
Alternative phrasing of acceptance probability: Barker dynamics

$$a(x_{t+1}, x_t) = P^*(x_{t+1}) / (P^*(x_{t+1}) + P^*(x_t))$$

Markov Chain Monte Carlo with people

Sanborn & Griffiths (2008) NIPS

- equivalence between **choice behaviour** and a **Metropolis Hastings sampler** can be established



acceptance probability is proportional to $\dfrac{P^*(x_{t+1})}{P^*(x_t)}$

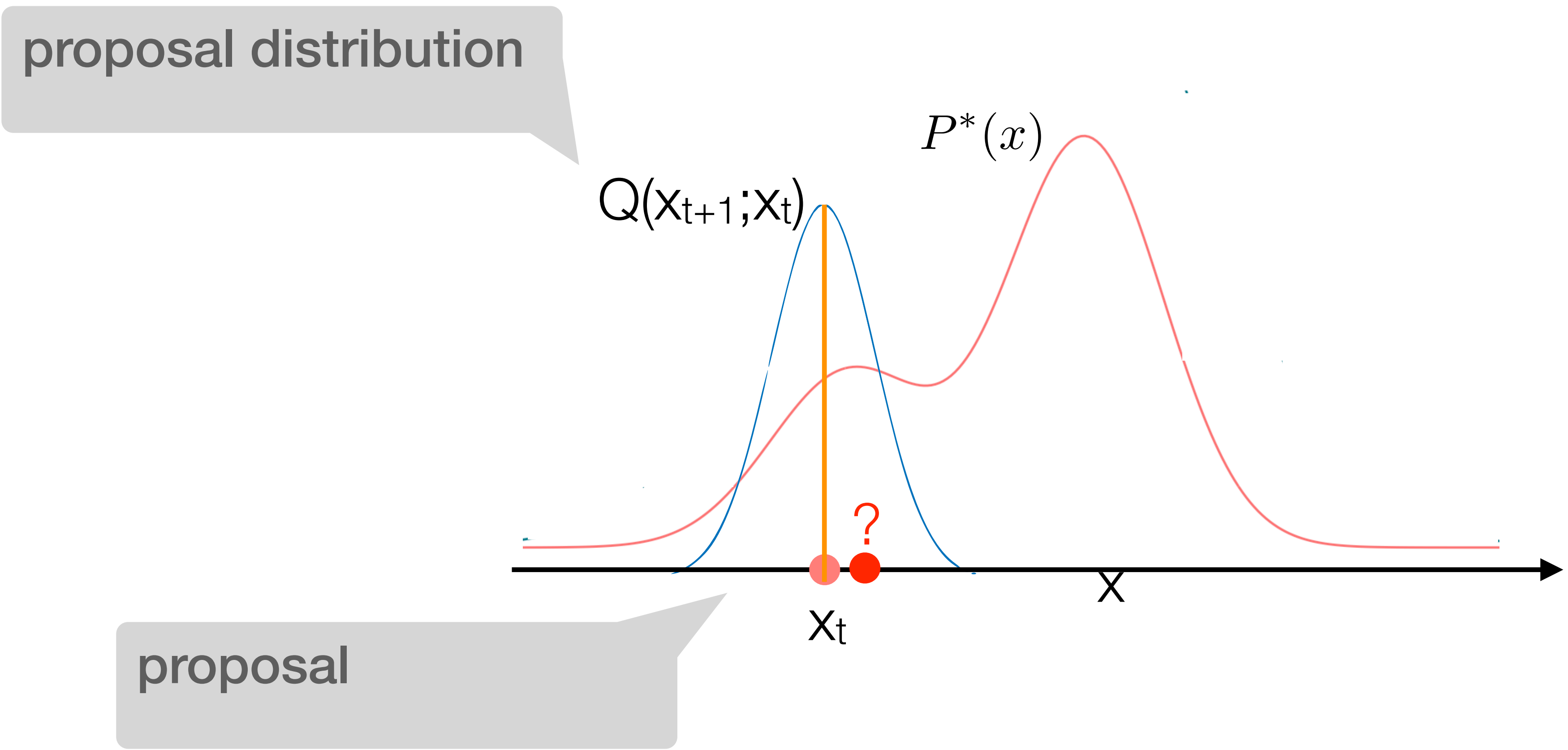
Alternative phrasing of acceptance probability: Barker dynamics

$$a(x_{t+1}, x_t) = P^*(x_{t+1}) / (P^*(x_{t+1}) + P^*(x_t))$$

Markov Chain Monte Carlo with people

Sanborn & Griffiths (2008) NIPS

- equivalence between **choice behaviour** and a **Metropolis Hastings sampler** can be established



acceptance probability is proportional to $\frac{P^*(x_{t+1})}{P^*(x_t)}$

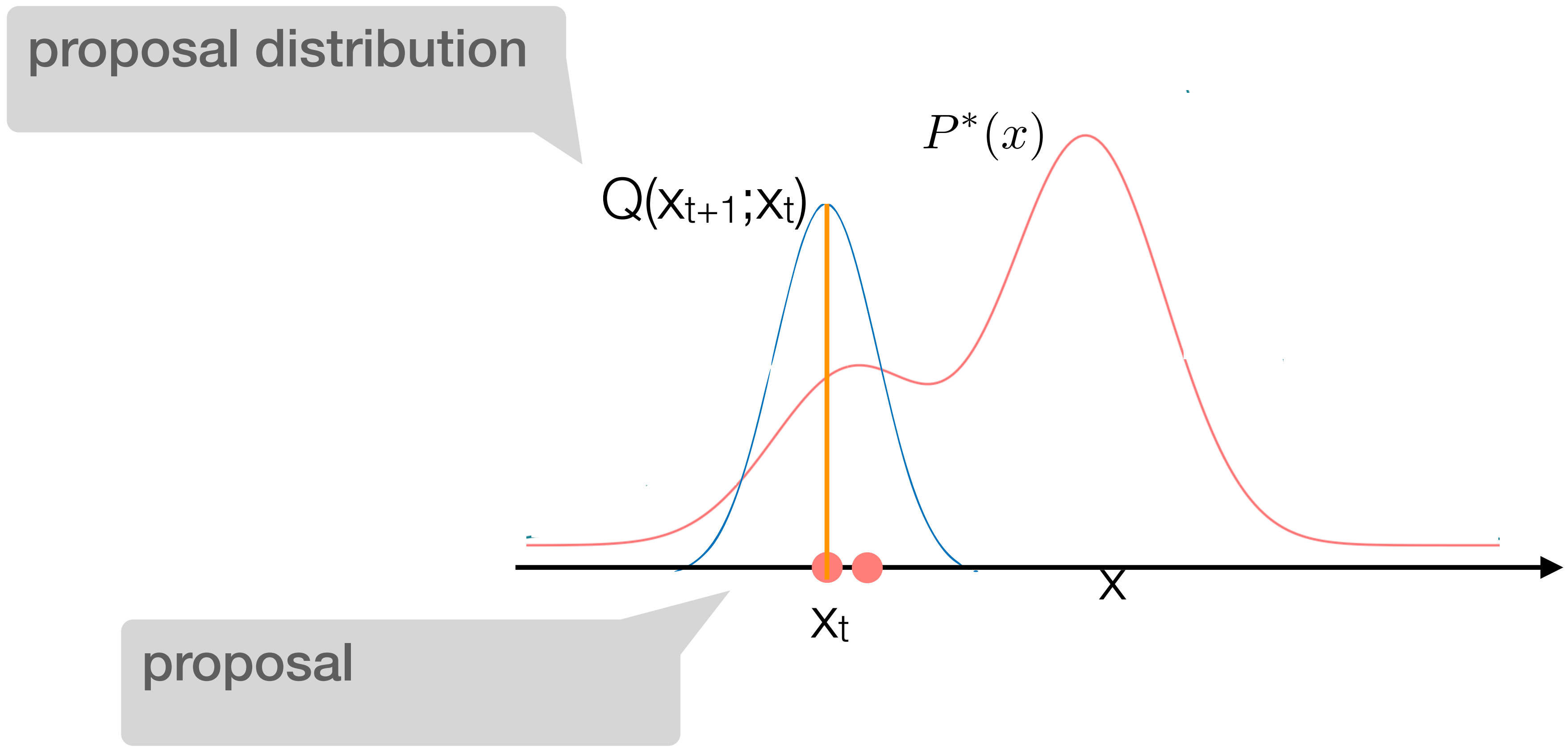
Alternative phrasing of acceptance probability: Barker dynamics

$$a(x_{t+1}, x_t) = P^*(x_{t+1}) / (P^*(x_{t+1}) + P^*(x_t))$$

Markov Chain Monte Carlo with people

Sanborn & Griffiths (2008) NIPS

- equivalence between **choice behaviour** and a **Metropolis Hastings sampler** can be established



acceptance probability is proportional to $\dfrac{P^*(x_{t+1})}{P^*(x_t)}$

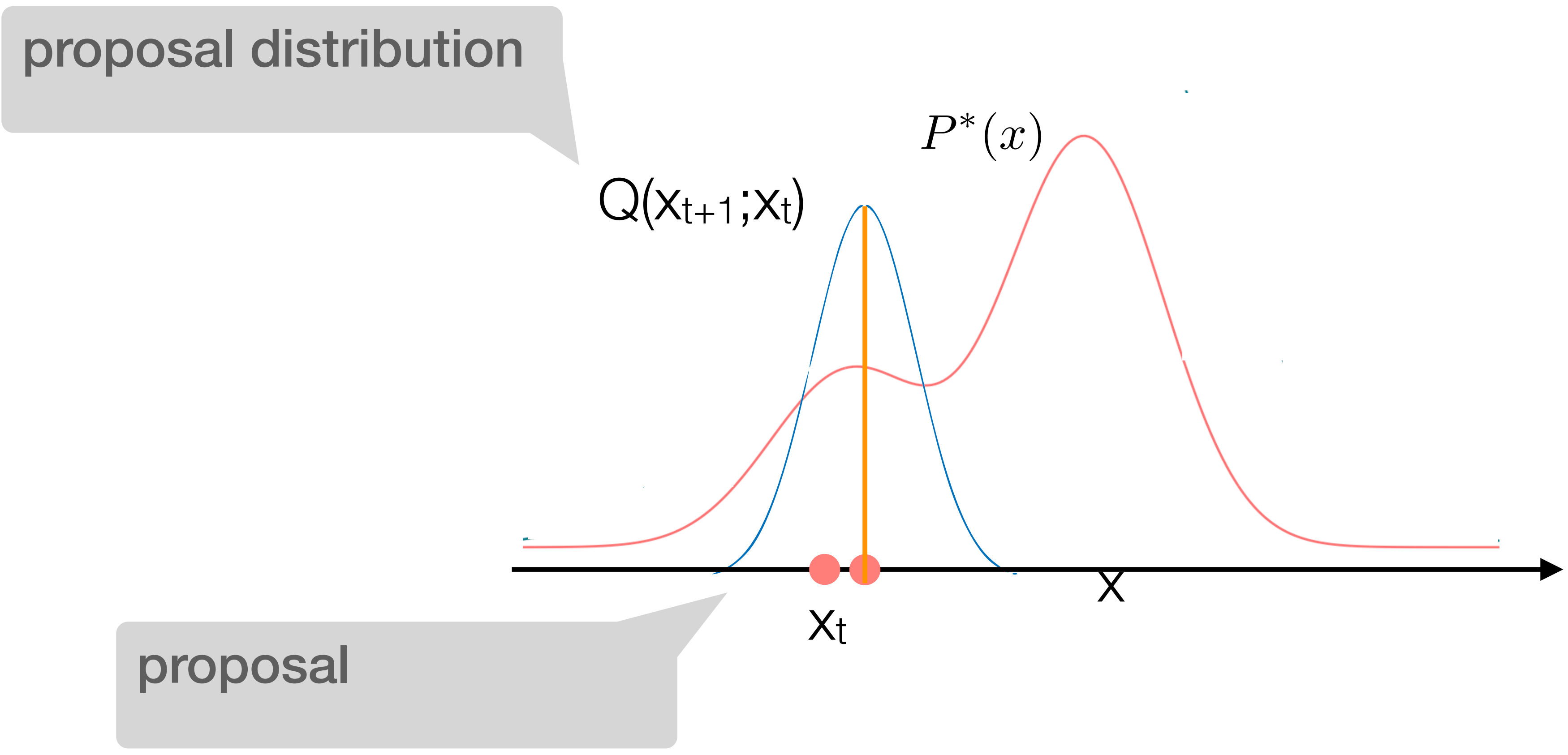
Alternative phrasing of acceptance probability: Barker dynamics

$$a(x_{t+1}, x_t) = P^*(x_{t+1}) / (P^*(x_{t+1}) + P^*(x_t))$$

Markov Chain Monte Carlo with people

Sanborn & Griffiths (2008) NIPS

- equivalence between **choice behaviour** and a **Metropolis Hastings sampler** can be established



acceptance probability is proportional to $\dfrac{P^*(x_{t+1})}{P^*(x_t)}$

Alternative phrasing of acceptance probability: Barker dynamics

$$a(x_{t+1}, x_t) = P^*(x_{t+1}) / (P^*(x_{t+1}) + P^*(x_t))$$

Markov Chain Monte Carlo with people

Sanborn & Griffiths (2008) NIPS

- two objects are shown: x1, and x2
- the subject is told that one of them is coming from a particular category c
- ‘choose the object that you think comes from category c’
- a Bayesian learner assumes two hypotheses:
  h1: x1 comes from $p(x \mid c)$, x2 comes from $g(x)$
  h2: the other way round

Markov Chain Monte Carlo with people

Sanborn & Griffiths (2008) NIPS

- two objects are shown: x1, and x2
- the subject is told that one of them is coming from a particular category c
- ‘choose the object that you think comes from category c’
- a Bayesian learner assumes two hypotheses:
  h1: x1 comes from p(x | c), x2 comes from g(x)
  h2: the other way round

posterior probability for h1:

$$p(h_1|x_1,x_2) = \frac{p(x_1,x_2|h_1)p(h_1)}{p(x_1,x_2|h_1)p(h_1)+p(x_1,x_2|h_2)p(h_2)}$$

Markov Chain Monte Carlo with people

Sanborn & Griffiths (2008) NIPS

- two objects are shown: x1, and x2
- the subject is told that one of them is coming from a particular category c
- 'choose the object that you think comes from category c'
- a Bayesian learner assumes two hypotheses:
  h1: x1 comes from p(x | c), x2 comes from g(x)
  h2: the other way round

$p(x_1,x_2|h_1)= p(x_1|c)\, g(x_2)$

$p(x_1,x_2|h_2)= g(x_1)\, p(x_2|c)$

posterior probability for h1:

$$p(h_1|x_1,x_2) \quad = \quad \frac{p(x_1,x_2|h_1)p(h_1)}{p(x_1,x_2|h_1)p(h_1)+p(x_1,x_2|h_2)p(h_2)}$$

Markov Chain Monte Carlo with people

Sanborn & Griffiths (2008) NIPS

- two objects are shown: x1, and x2
- the subject is told that one of them is coming from a particular category c
- 'choose the object that you think comes from category c'
- a Bayesian learner assumes two hypotheses:
  h1: x1 comes from p(x | c), x2 comes from g(x)
  h2: the other way round

$p(x_1,x_2|h_1)= p(x_1|c)\, g(x_2)$

$p(x_1,x_2|h_2)= g(x_1)\, p(x_2|c)$

posterior probability for h1:

$$p(h_1|x_1,x_2) = \frac{p(x_1,x_2|h_1)p(h_1)}{p(x_1,x_2|h_1)p(h_1)+p(x_1,x_2|h_2)p(h_2)}$$

g(x) is an alternative hypothesis for the origin of x

Markov Chain Monte Carlo with people

Sanborn & Griffiths (2008) NIPS

- two objects are shown: x1, and x2
- the subject is told that one of them is coming from a particular category c
- ‘choose the object that you think comes from category c’
- a Bayesian learner assumes two hypotheses:
  h1: x1 comes from p(x | c), x2 comes from g(x)
  h2: the other way round

$p(x_1,x_2|h_1)= p(x_1|c)\, g(x_2)$

$p(x_1,x_2|h_2)= g(x_1)\, p(x_2|c)$

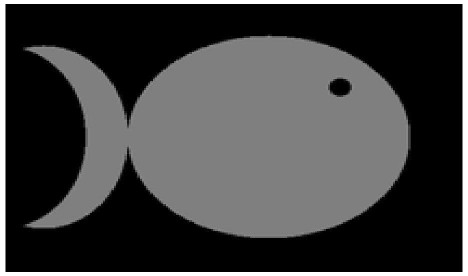
posterior probability for h1:

$$p(h_1|x_1,x_2) = \frac{p(x_1,x_2|h_1)p(h_1)}{p(x_1,x_2|h_1)p(h_1)+p(x_1,x_2|h_2)p(h_2)}$$
$$= \frac{p(x_1|c)g(x_2)p(h_1)}{p(x_1|c)g(x_2)p(h_1)+p(x_2|c)g(x_1)p(h_2)}$$

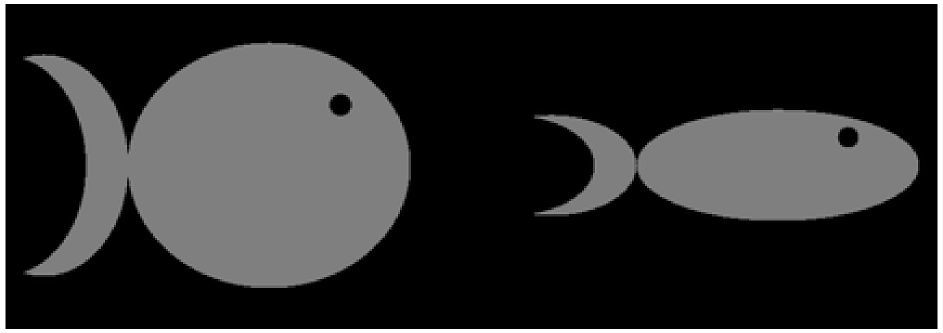
g(x) is an alternative hypothesis for the origin of x

Markov Chain Monte Carlo with people

Sanborn & Griffiths (2008) NIPS



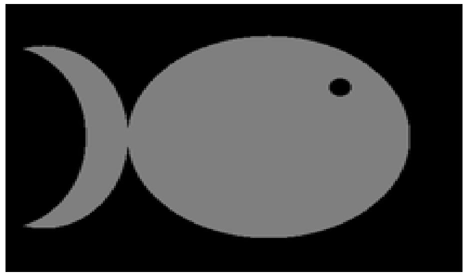
- training subjects on a novel ‘lab category’ (fish from the ocean)



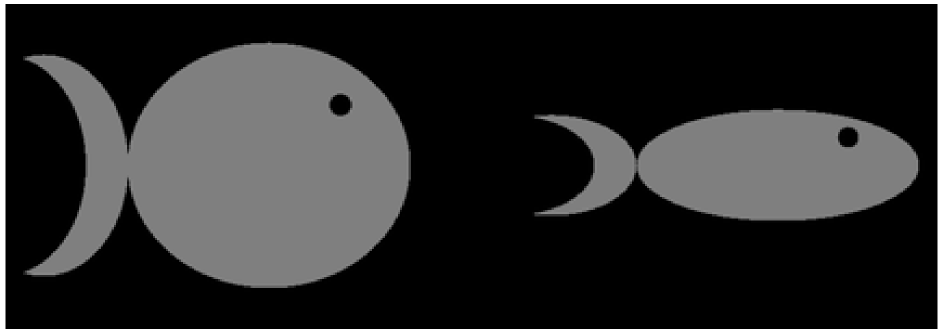
- and later test their prior with MCMC

Markov Chain Monte Carlo with people

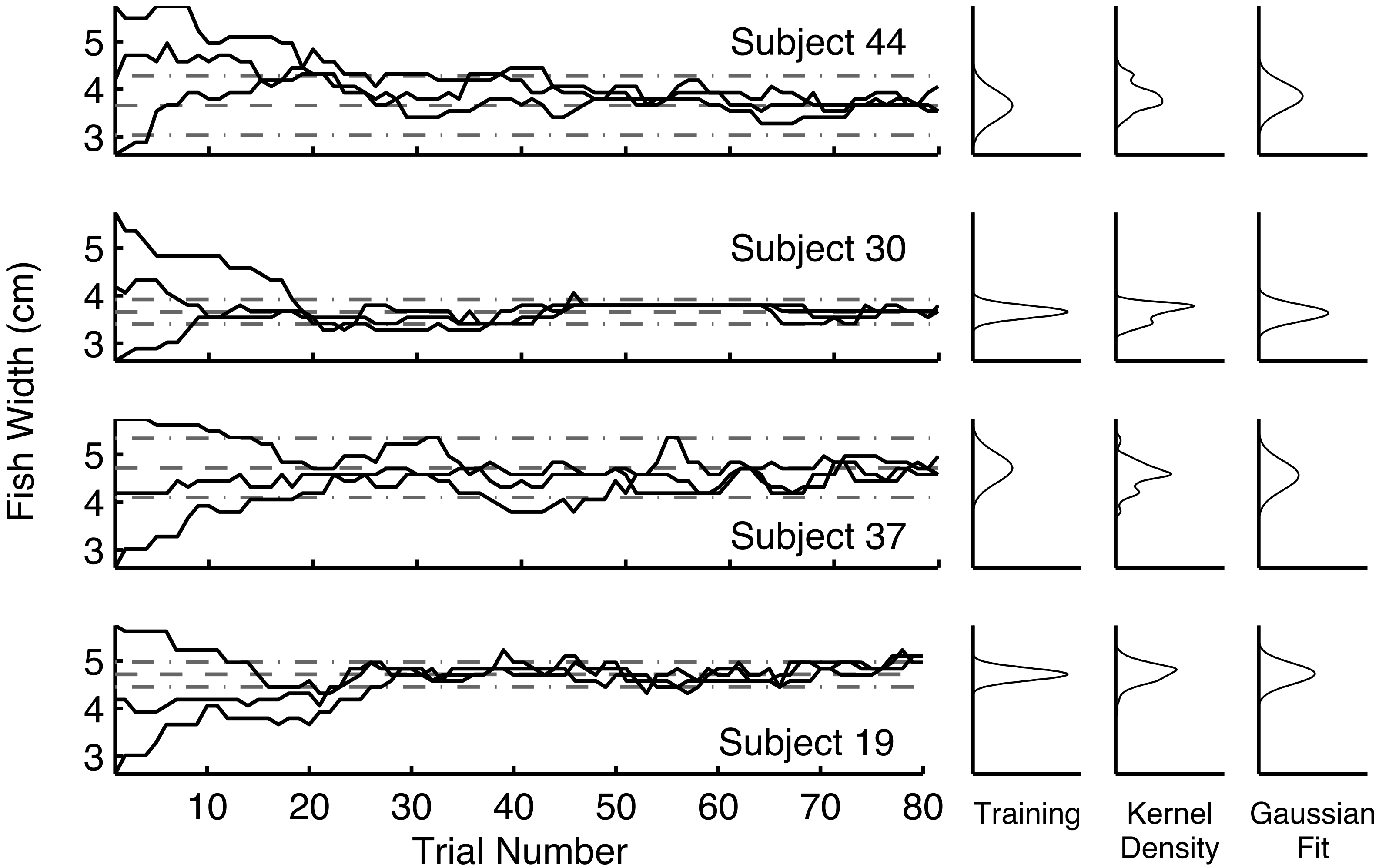
Sanborn & Griffiths (2008) NIPS



- training subjects on a novel ‘lab category’ (fish from the ocean)



- and later test their prior with MCMC



Markov Chain Monte Carlo with people

Sanborn & Griffiths (2008) NIPS

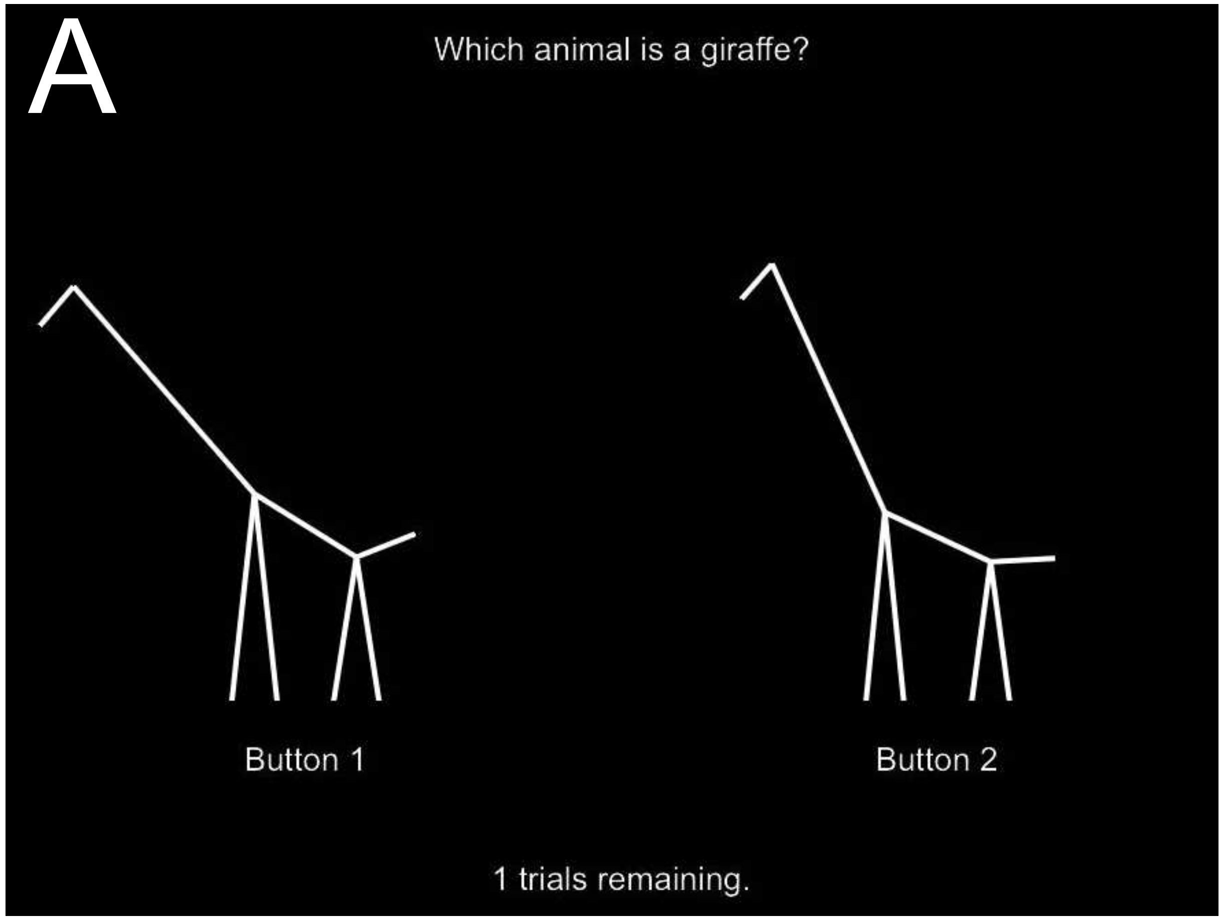
- exploring a learned category (animals)
- wire-frame animals are used which are easy to parametrise (tail length, neck length, neck angle, etc)

Markov Chain Monte Carlo with people

Sanborn & Griffiths (2008) NIPS

- exploring a learned category (animals)
- wire-frame animals are used which are easy to parametrise (tail length, neck length, neck angle, etc)

stimuli

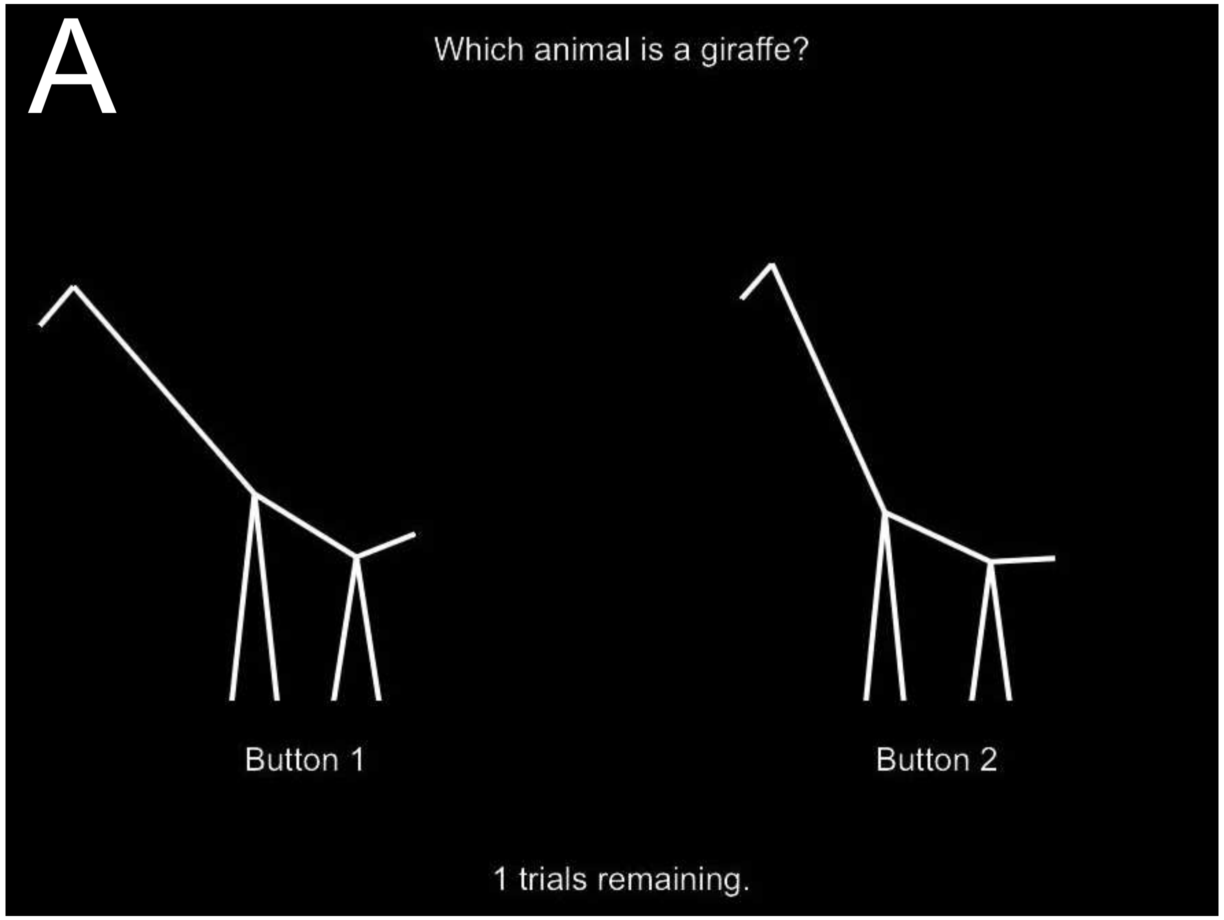


Markov Chain Monte Carlo with people

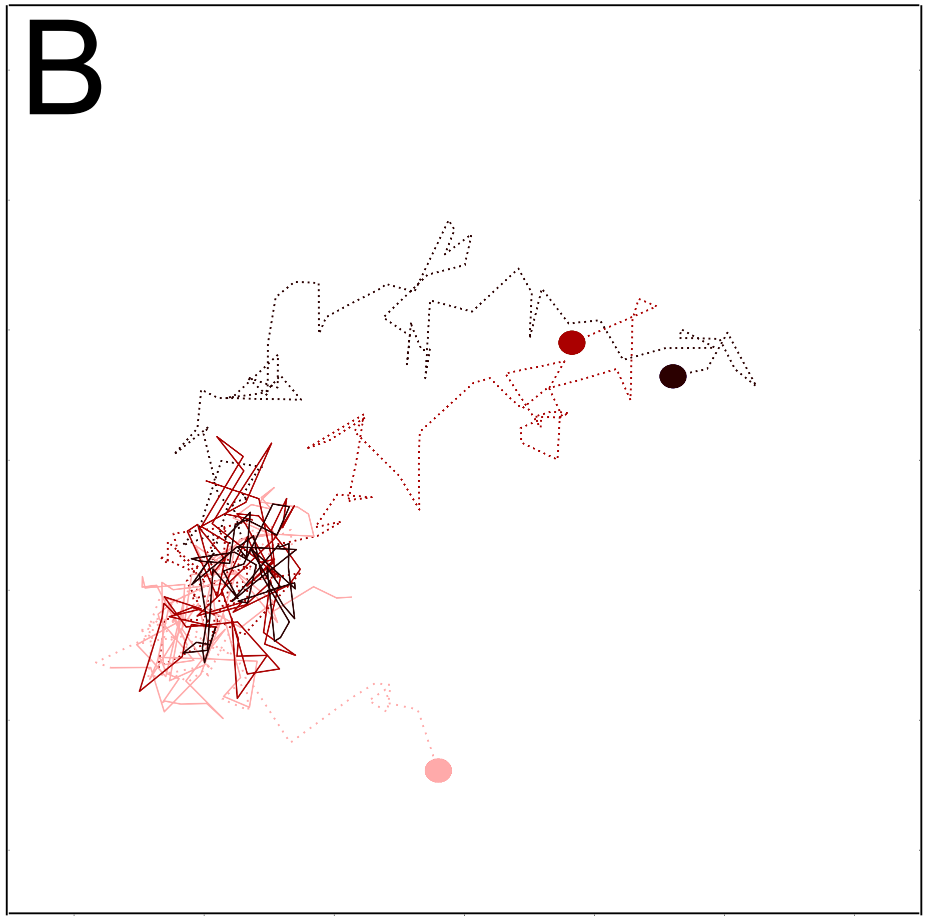
Sanborn & Griffiths (2008) NIPS

- exploring a learned category (animals)
- wire-frame animals are used which are easy to parametrise (tail length, neck length, neck angle, etc)

stimuli



sequence of stimuli
in the stimulus space



Markov Chain Monte Carlo with people

Sanborn & Griffiths (2008) NIPS

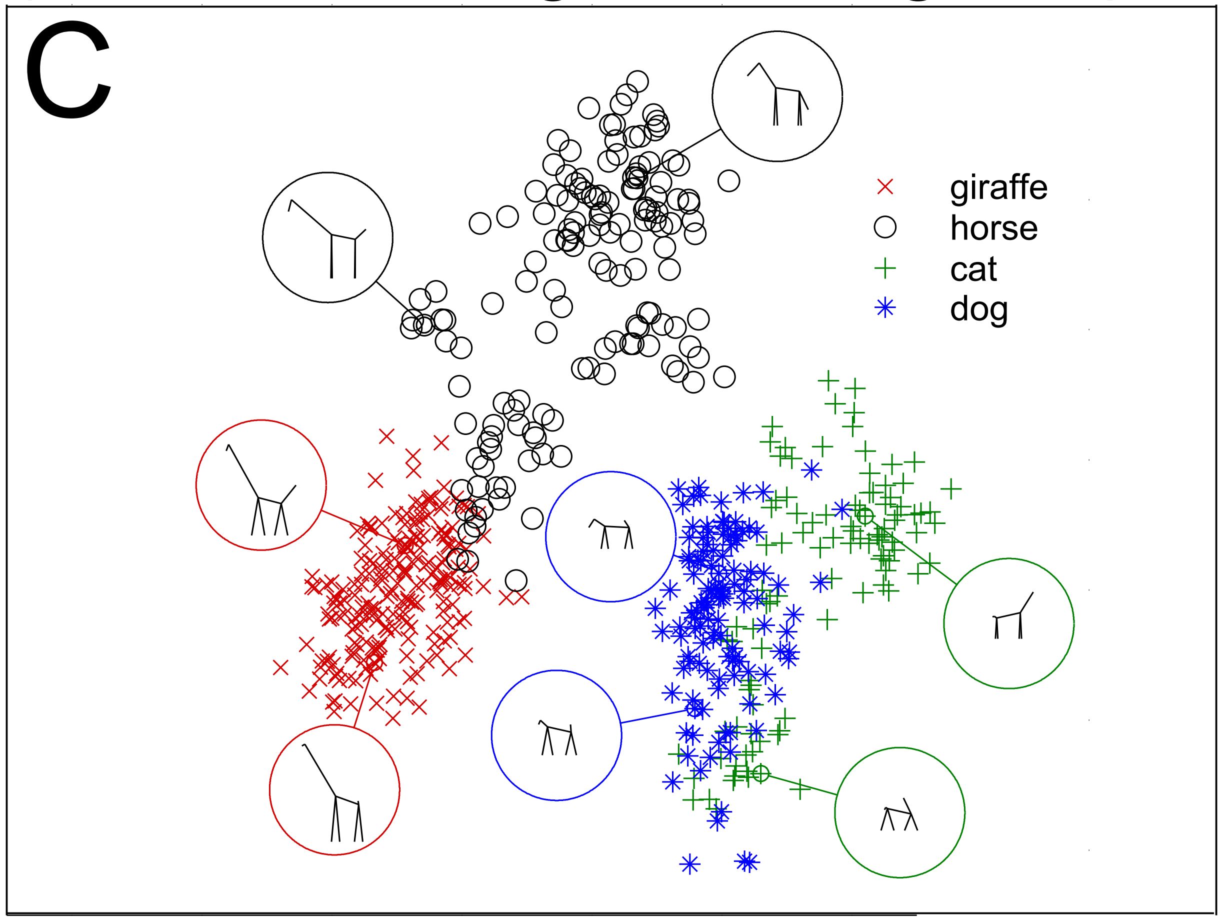
- exploring a learned category (animals)
- wire-frame animals are used which are easy to parametrise (tail length, neck length, neck angle, etc)

Markov Chain Monte Carlo with people

Sanborn & Griffiths (2008) NIPS

- exploring a learned category (animals)
- wire-frame animals are used which are easy to parametrise (tail length, neck length, neck angle, etc)

inferred priors for different categories
(2D embedding of the high-D prior)

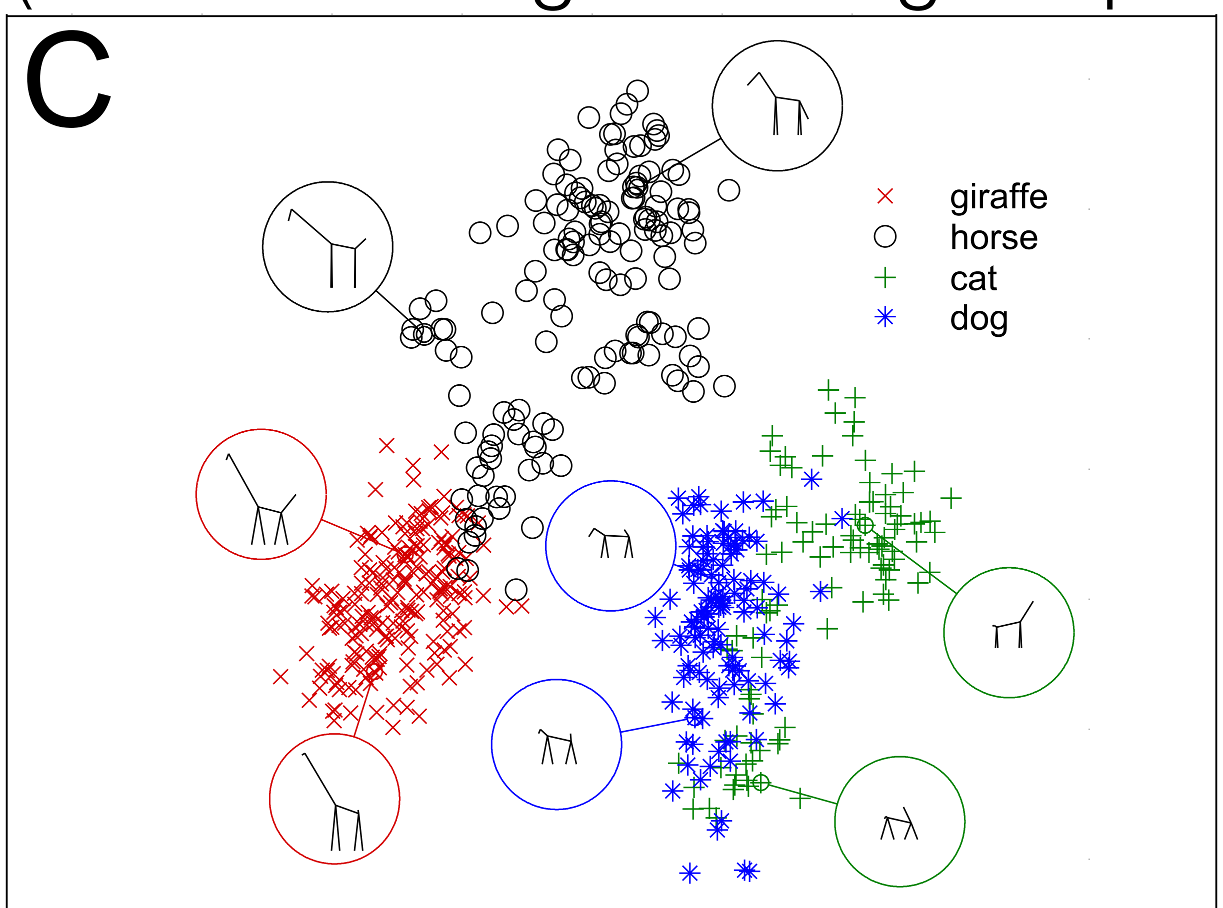


Markov Chain Monte Carlo with people

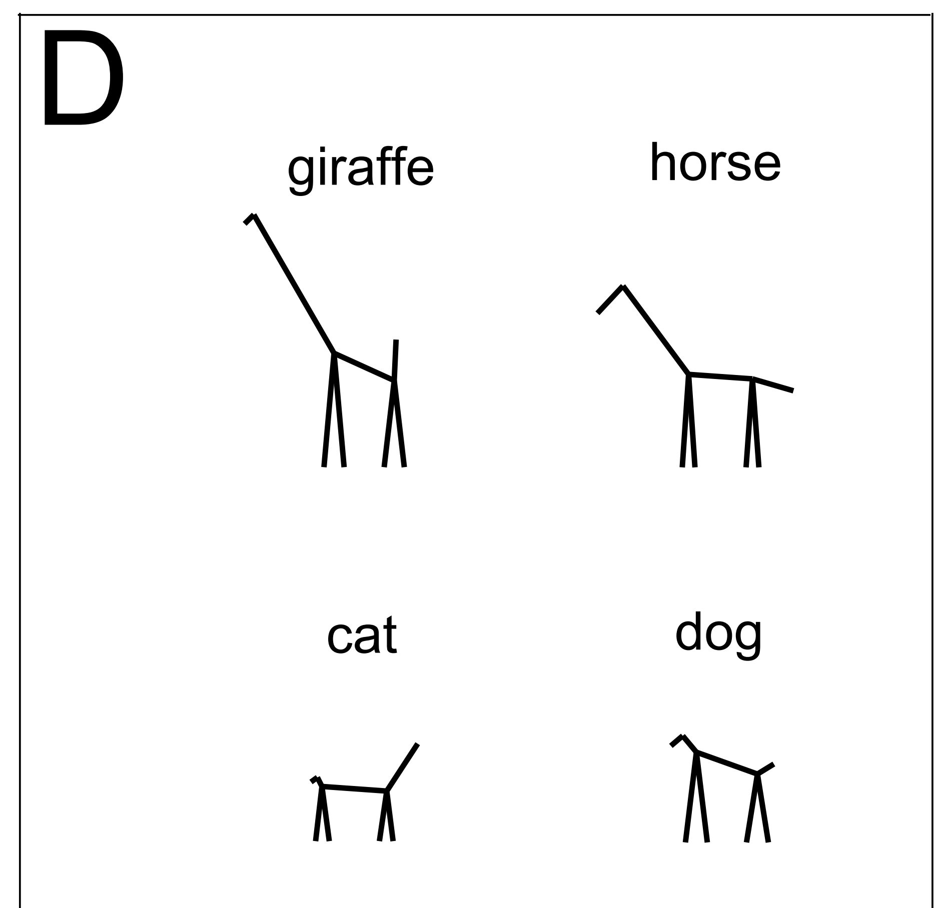
Sanborn & Griffiths (2008) NIPS

- exploring a learned category (animals)
- wire-frame animals are used which are easy to parametrise (tail length, neck length, neck angle, etc)

inferred priors for different categories (2D embedding of the high-D prior)



E[p(animals)]



Markov Chain Monte Carlo with people

Sanborn & Griffiths (2008) NIPS

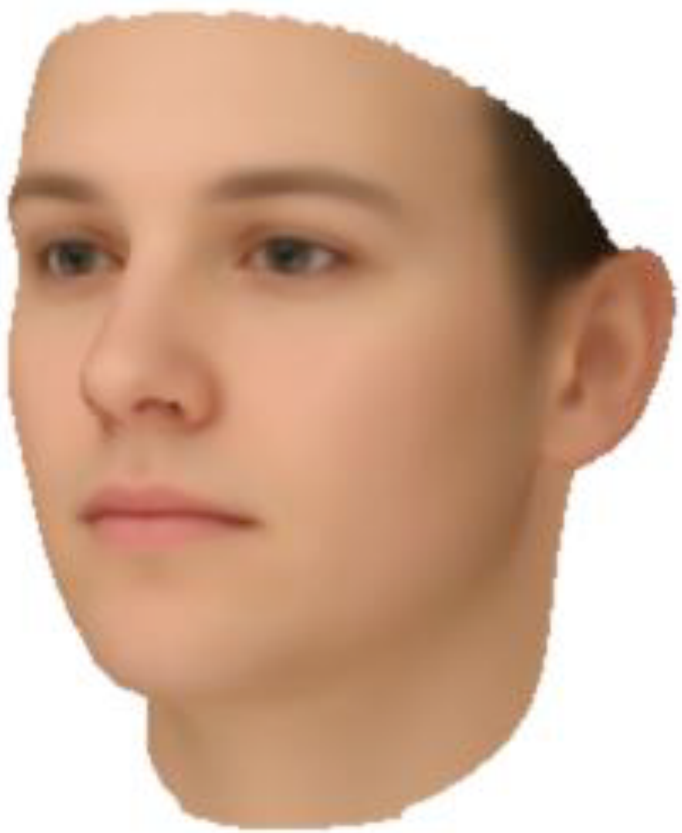
- The inferred prior can capture individual differences (subjective)
- The inferred prior is high-dimensional (fairly complex)
- The prior is task specific

Cognitive tomography

Houlsby et al (2013) Curr Bill

Cognitive tomography

Houlsby et al (2013) Curr Bill



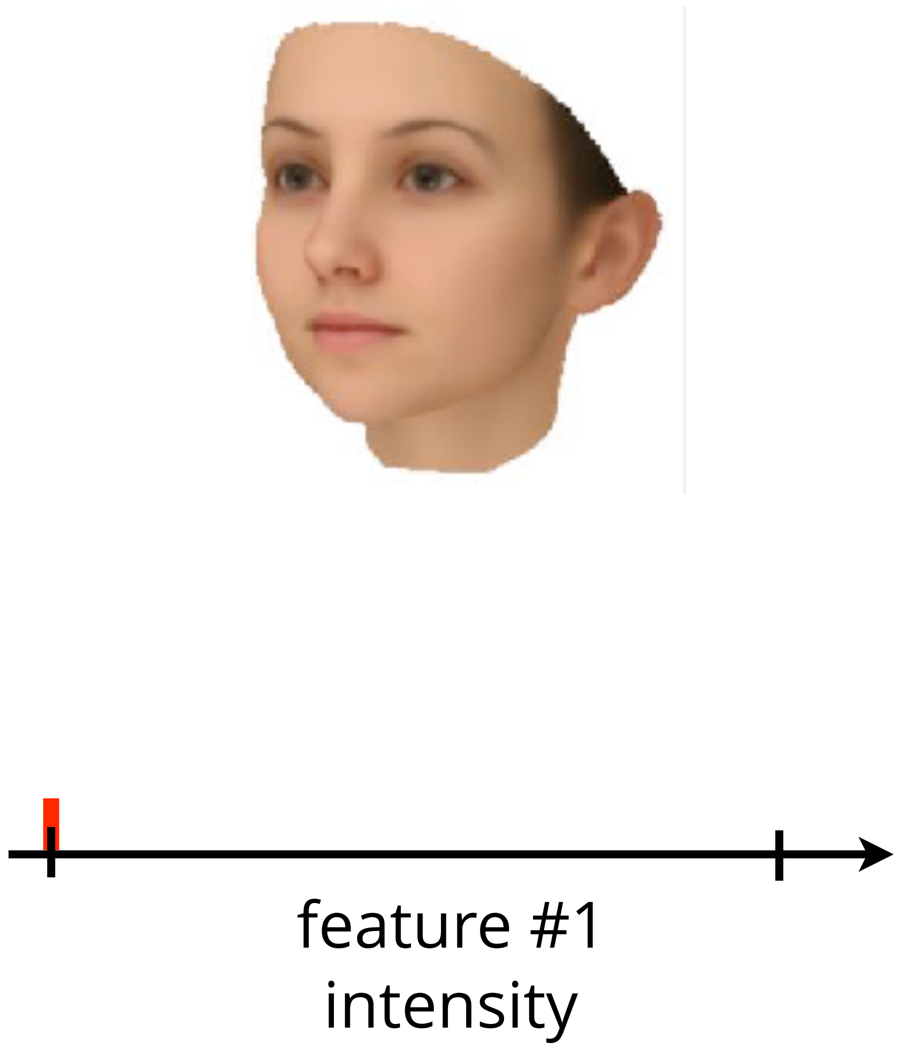
1. *complex*
2. ecologically *relevant*
3. *extensive* subjective experience
4. experience is *subjective*

Cognitive tomography

Houlsby et al (2013) Curr Biol

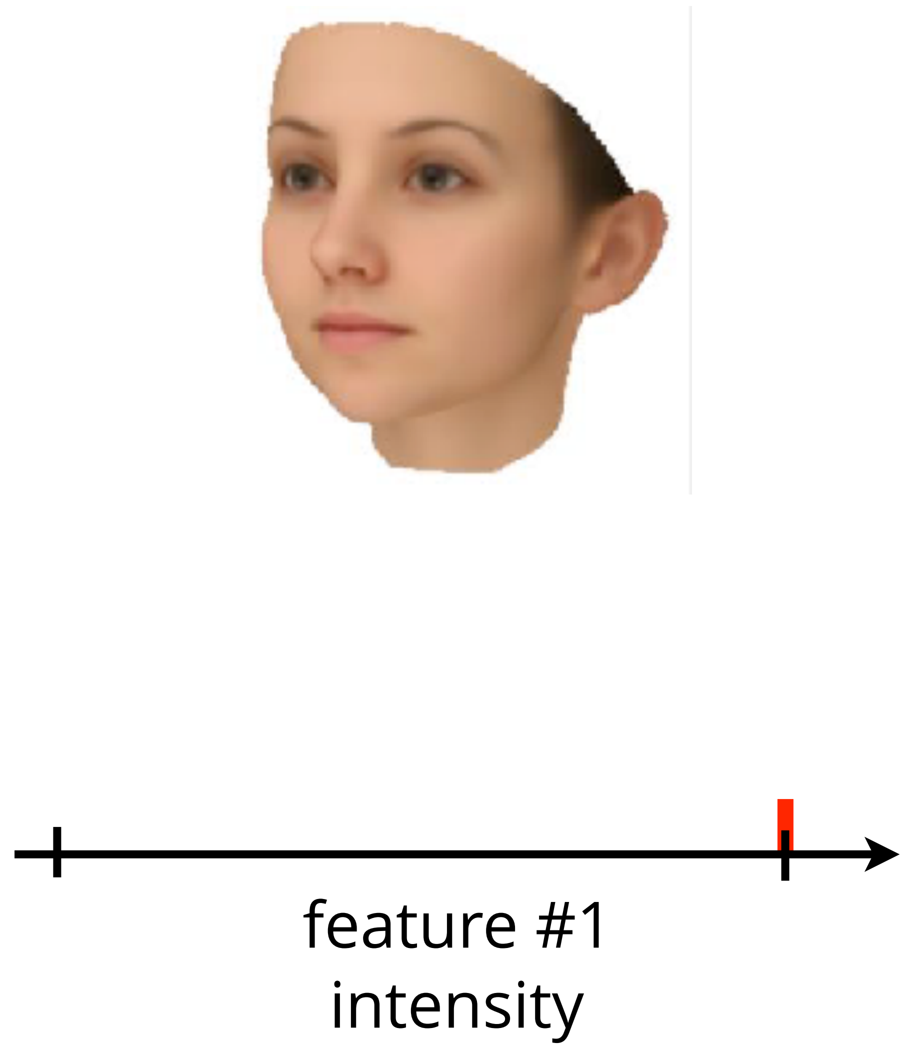
Cognitive tomography

Houlsby et al (2013) Curr Bill



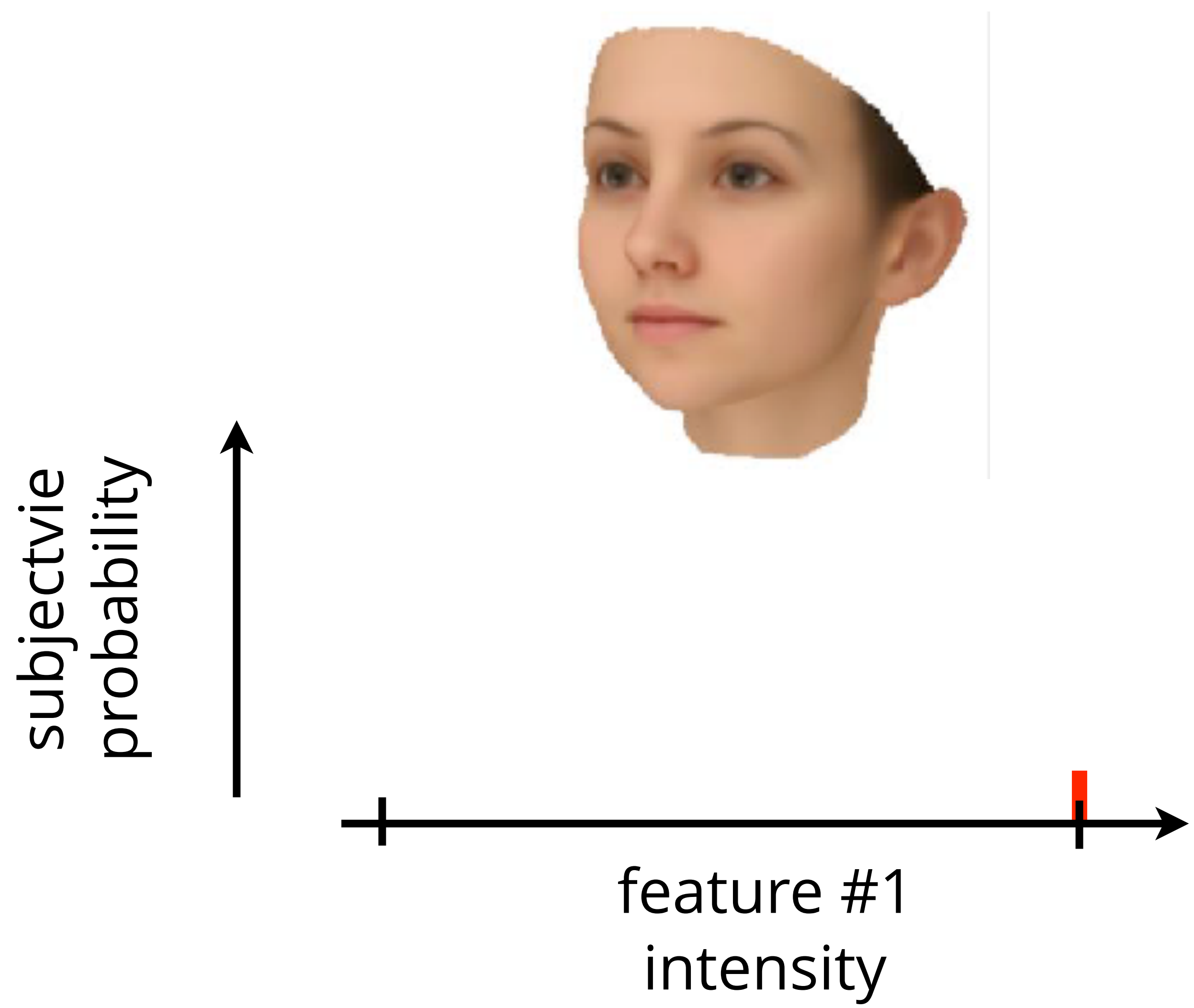
Cognitive tomography

Houlsby et al (2013) Curr Bill



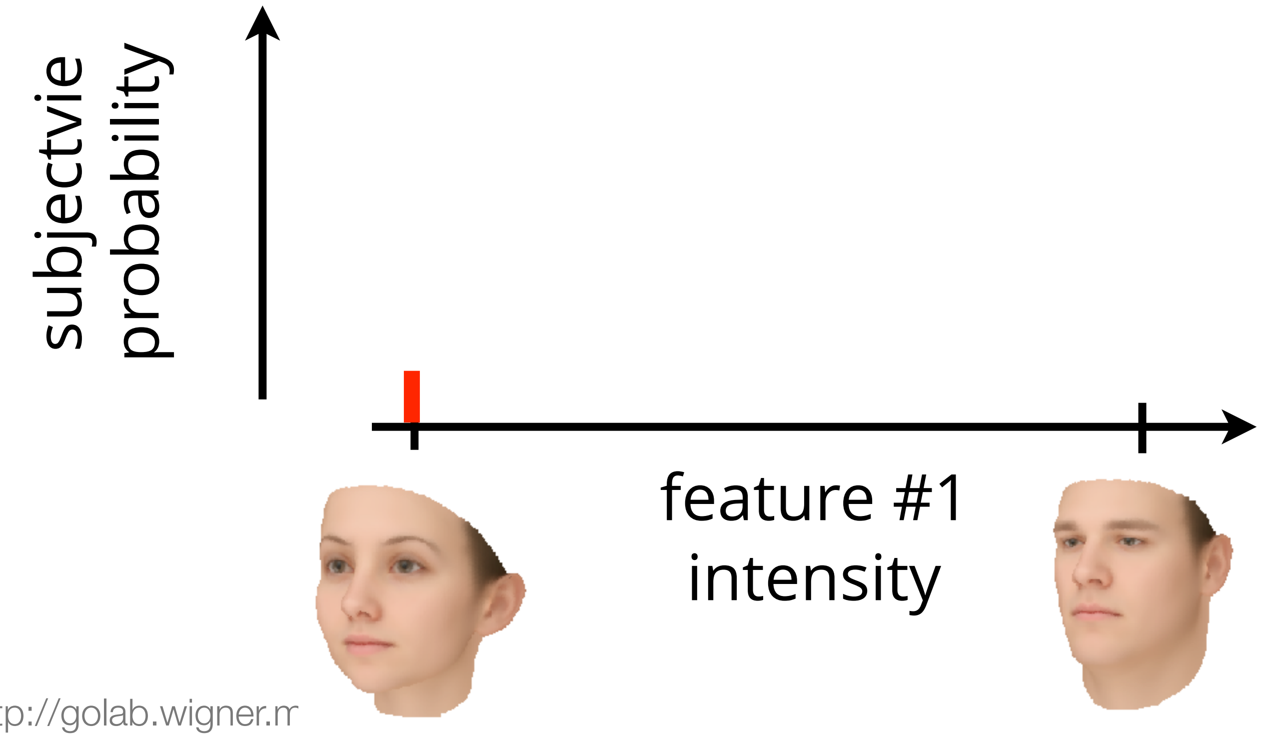
Cognitive tomography

Houlsby et al (2013) Curr Bill



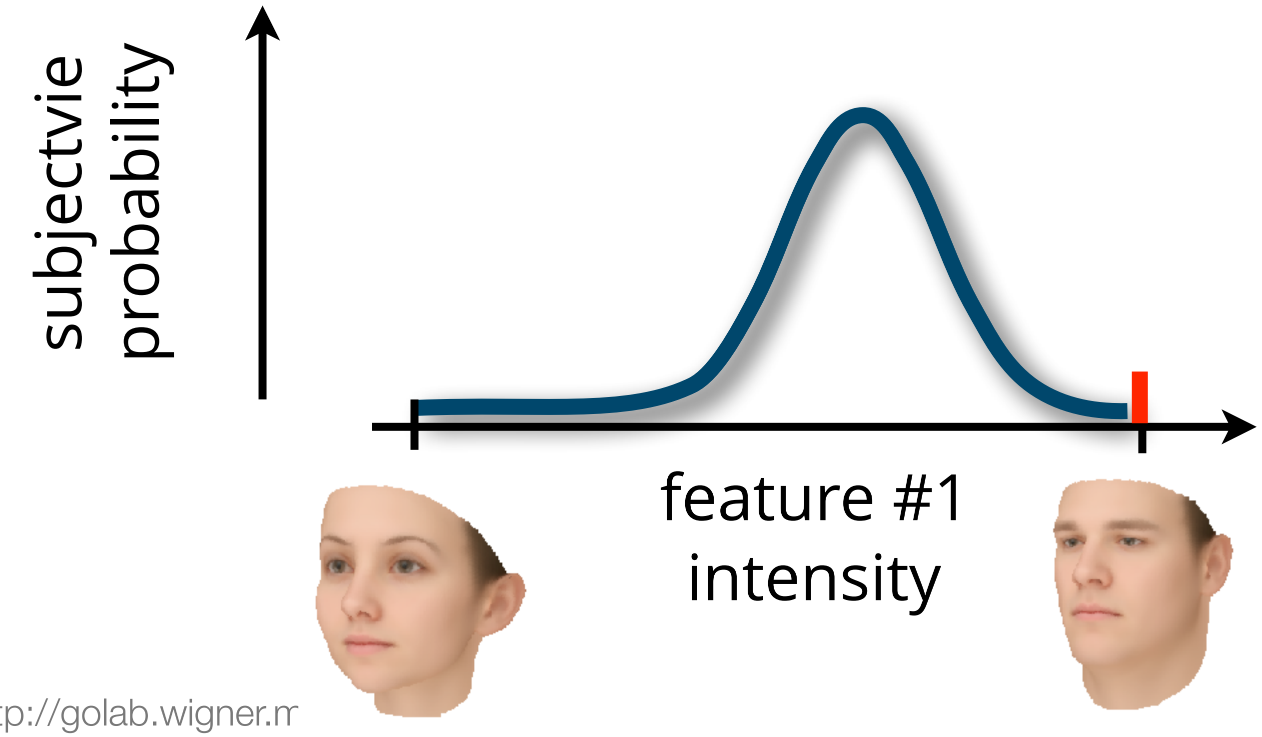
Cognitive tomography

Houlsby et al (2013) Curr Bill



Cognitive tomography

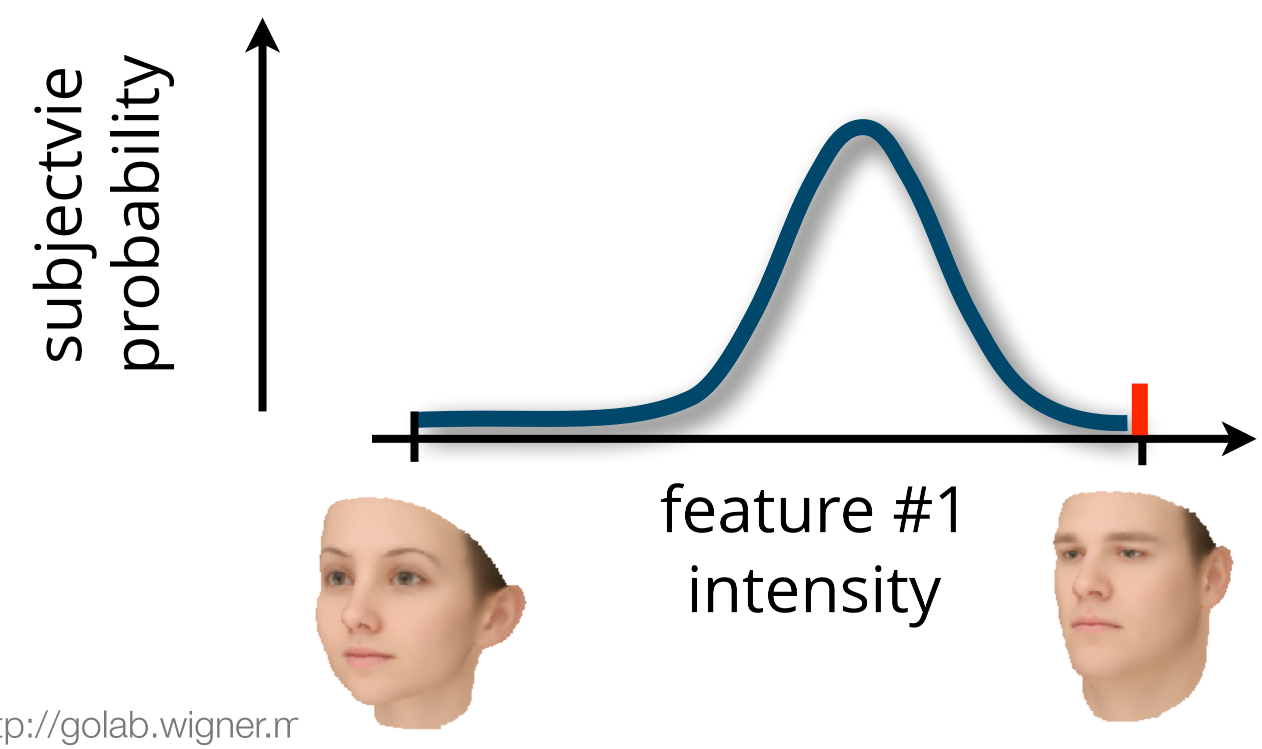
Houlsby et al (2013) Curr Bill



Cognitive tomography

Houlsby et al (2013) Curr Bill

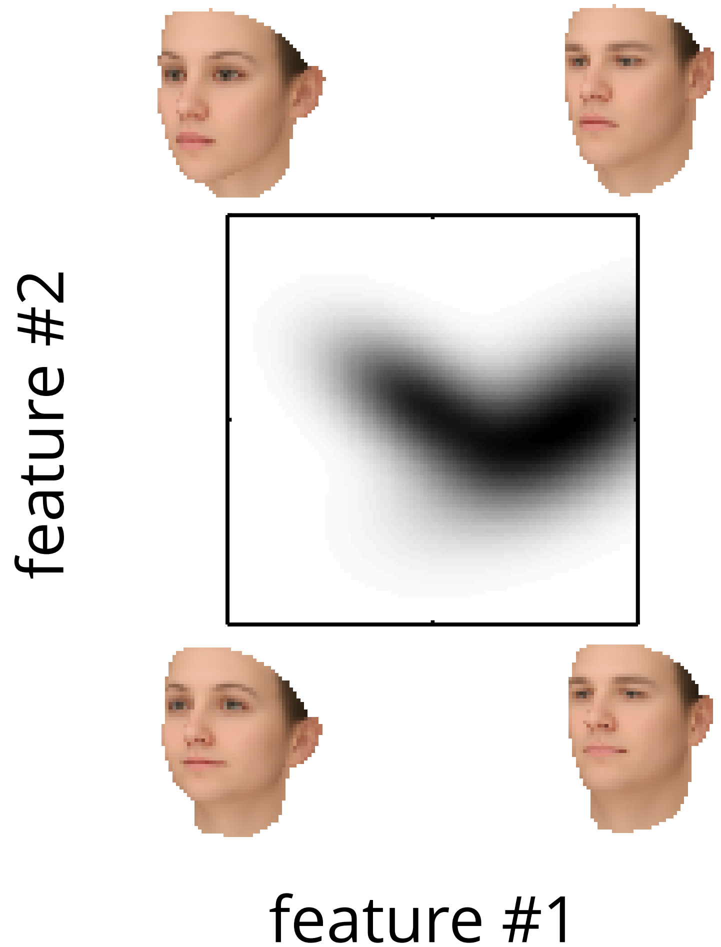
internal model — subjective distribution



Cognitive tomography

Houlsby et al (2013) Curr Bill

internal model — subjective distribution



Cognitive tomography

Houlsby et al (2013) Curr Bill

Cognitive tomography

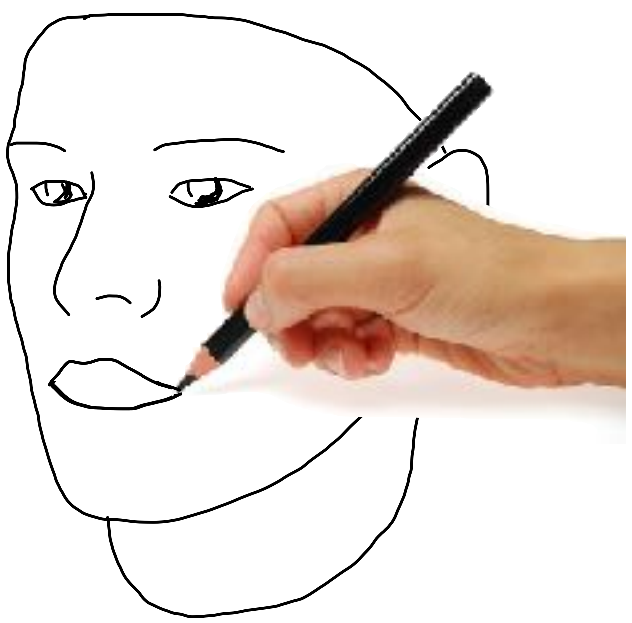
Houlsby et al (2013) Curr Bill

Cognitive tomography

Houlsby et al (2013) Curr Bill

Cognitive tomography

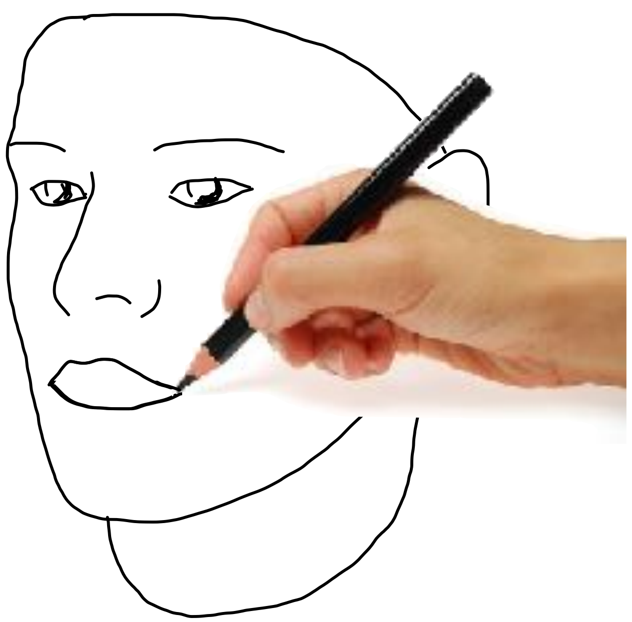
Houlsby et al (2013) Curr Bill

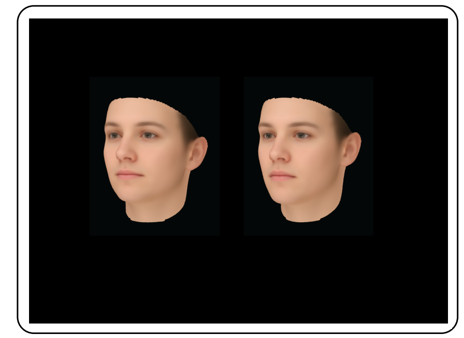


2

Cognitive tomography

Houlsby et al (2013) Curr Bill





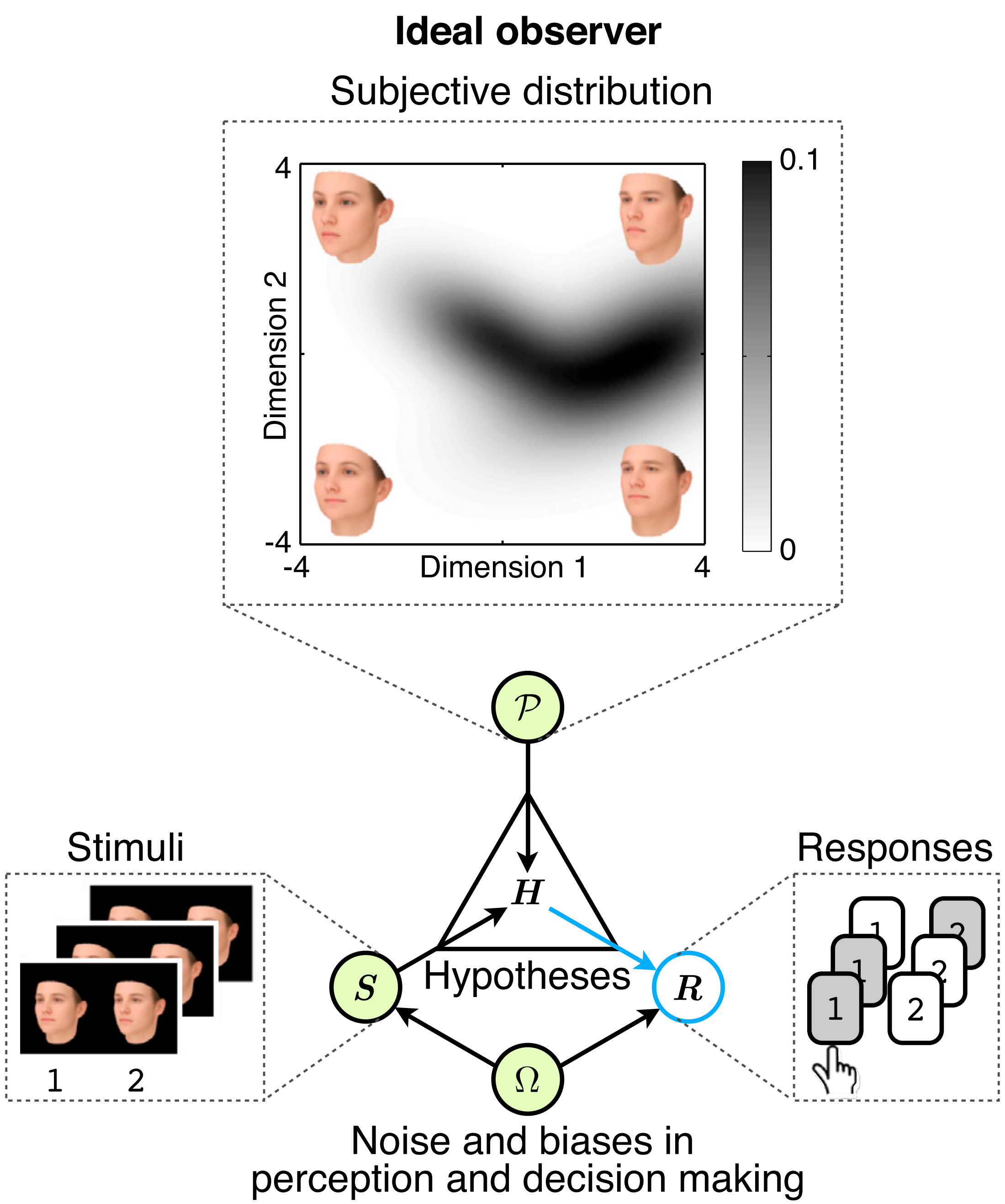
1 2

2
1
2
2
2
1
2
1
1
2
1
1
2
2
1
2
1
2

Cognitive tomography

Houlsby et al (2013) Curr Bill

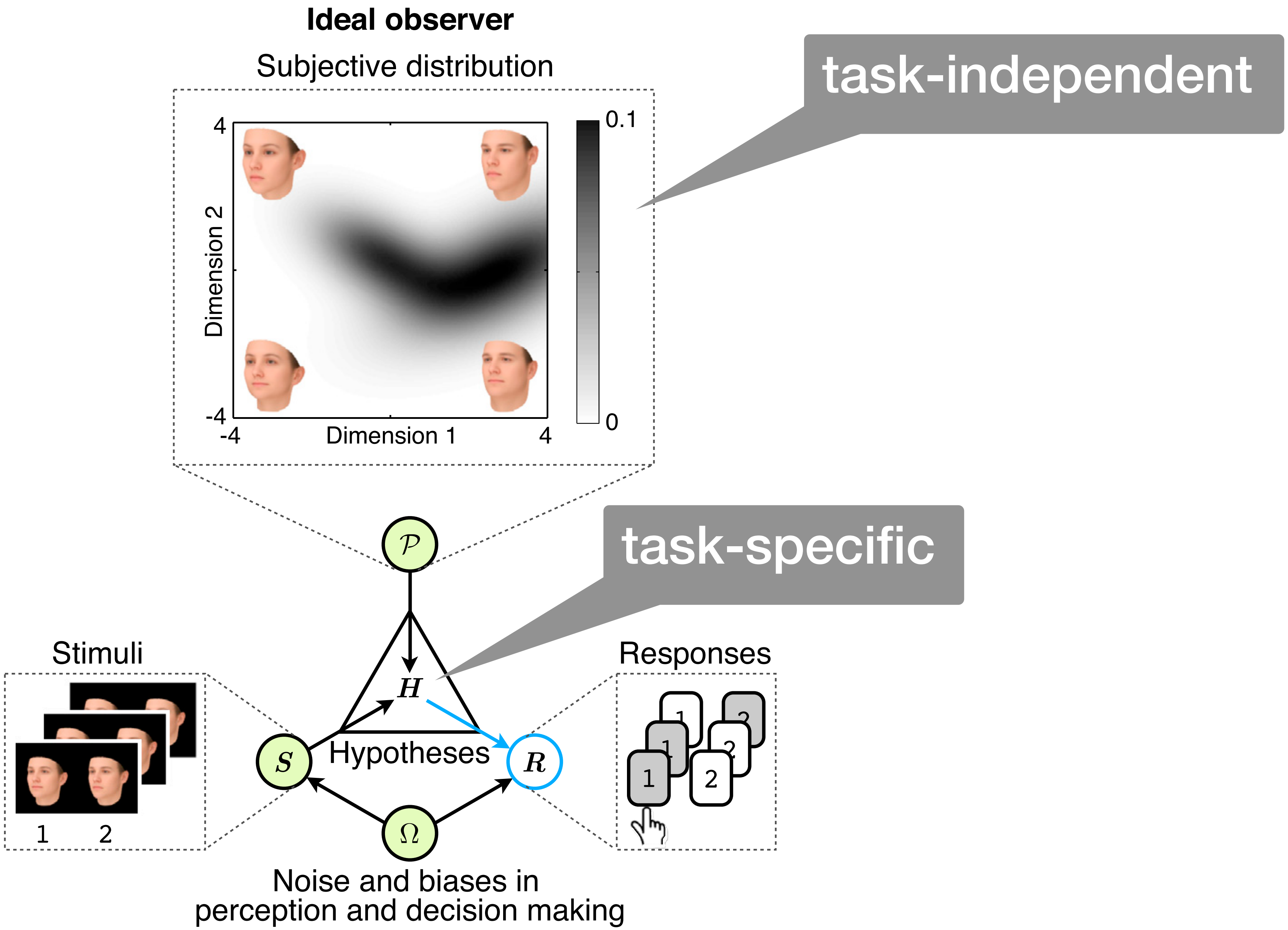
model of how responses are generated



Cognitive tomography

Houlsby et al (2013) Curr Bill

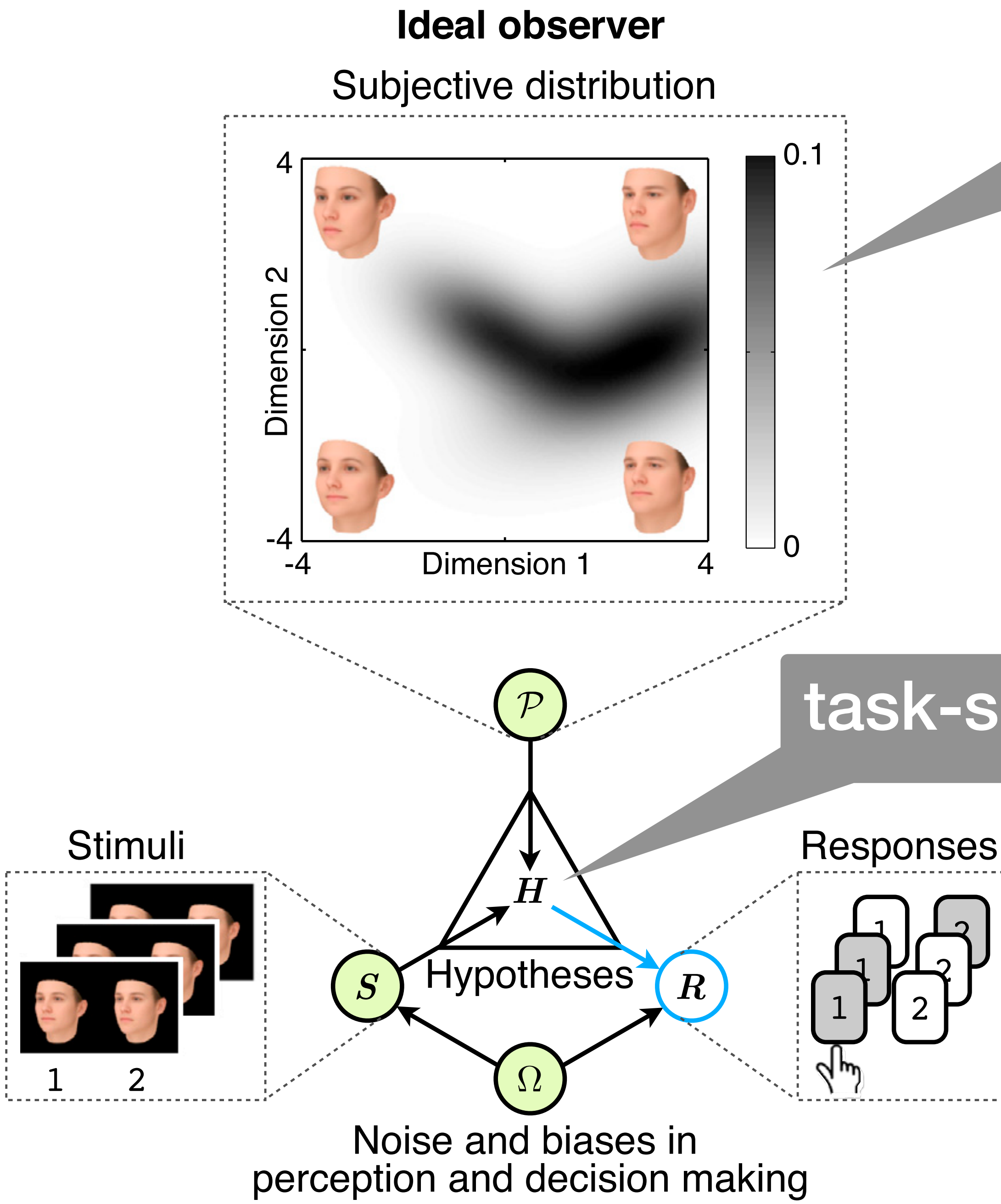
model of how responses are generated



Cognitive tomography

Houlsby et al (2013) Curr Bill

model of how responses are generated



task-independent

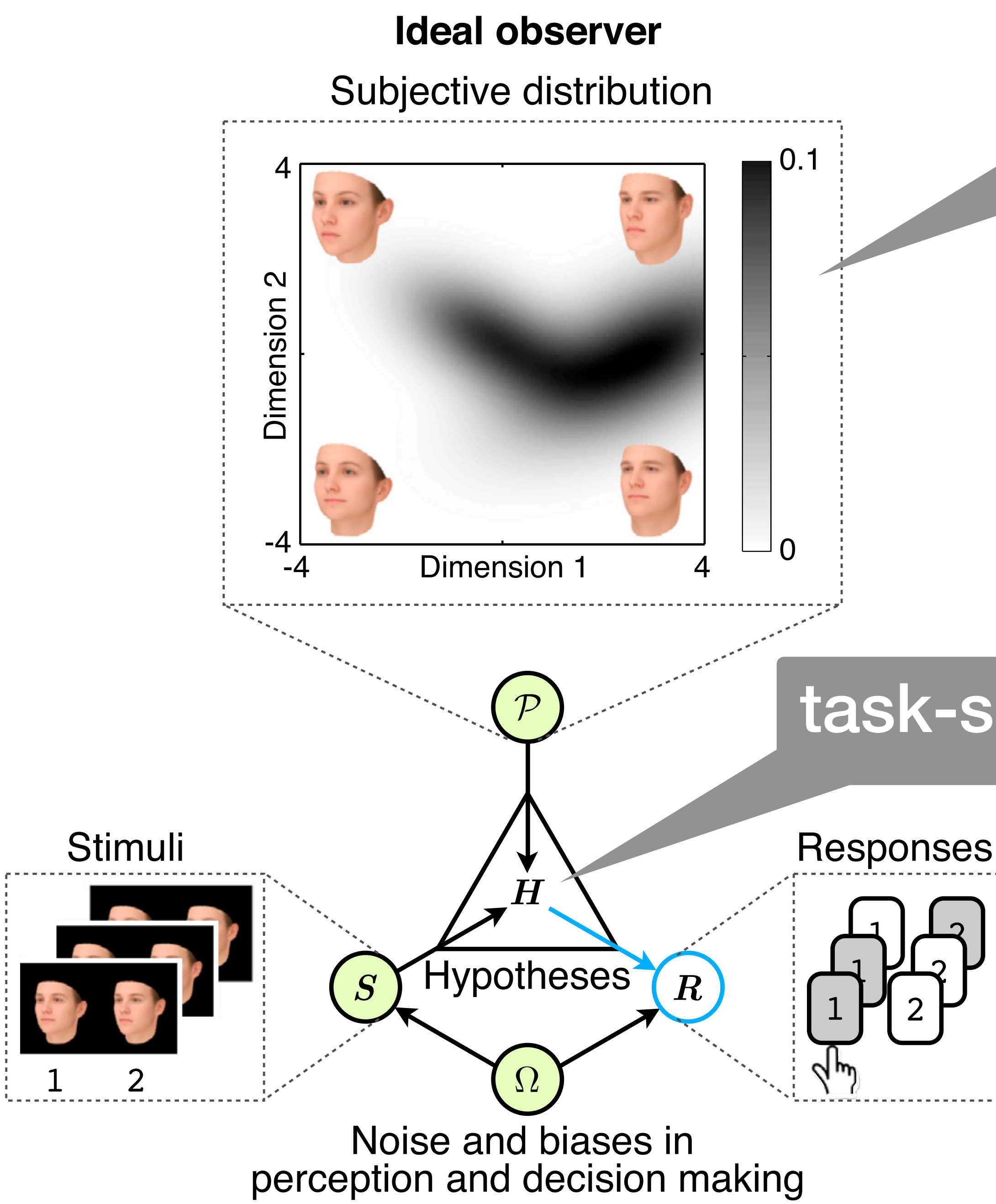
task-specific

‘cognitive tomography’:
inverting the generative model

Cognitive tomography

Houlsby et al (2013) Curr Bill

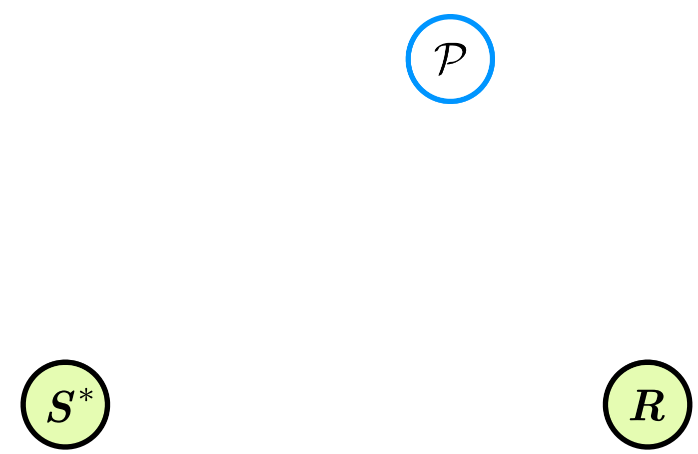
model of how responses are generated



task-independent

task-specific

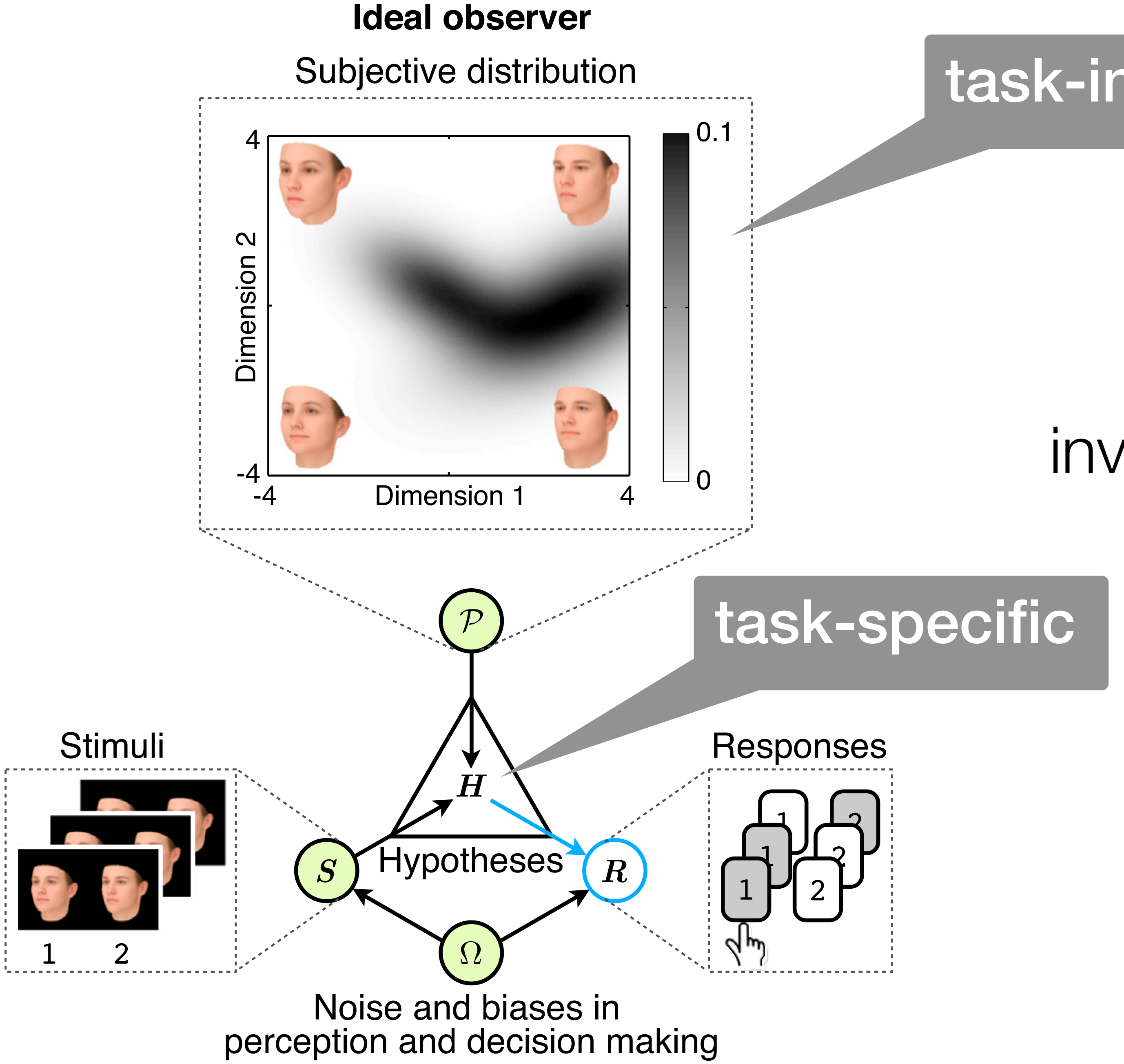
'cognitive tomography':
inverting the generative model



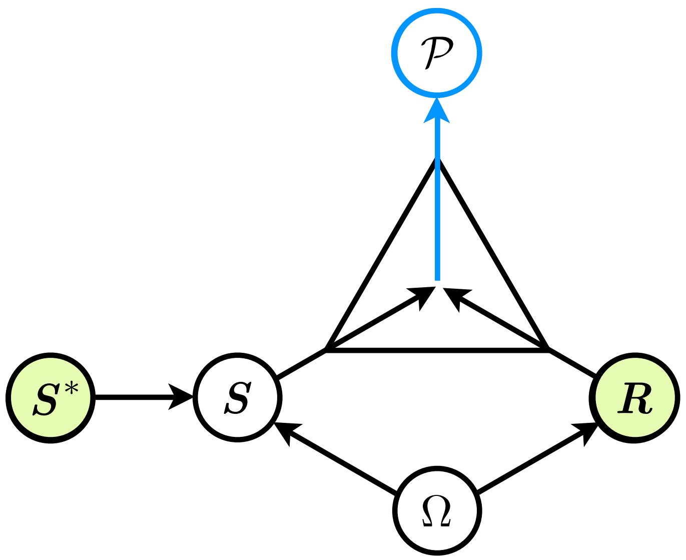
Cognitive tomography

Houlsby et al (2013) Curr Bill

model of how responses are generated

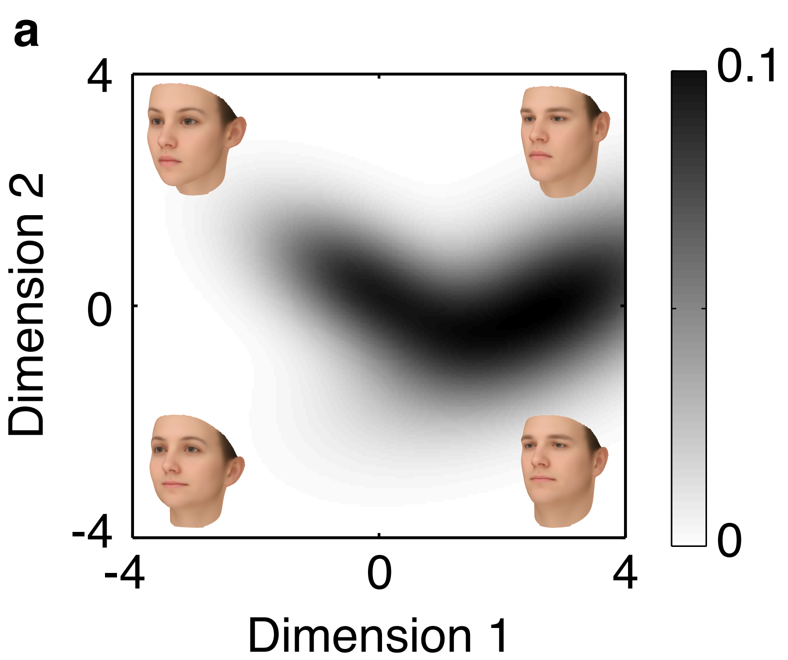


‘cognitive tomography’:
inverting the generative model



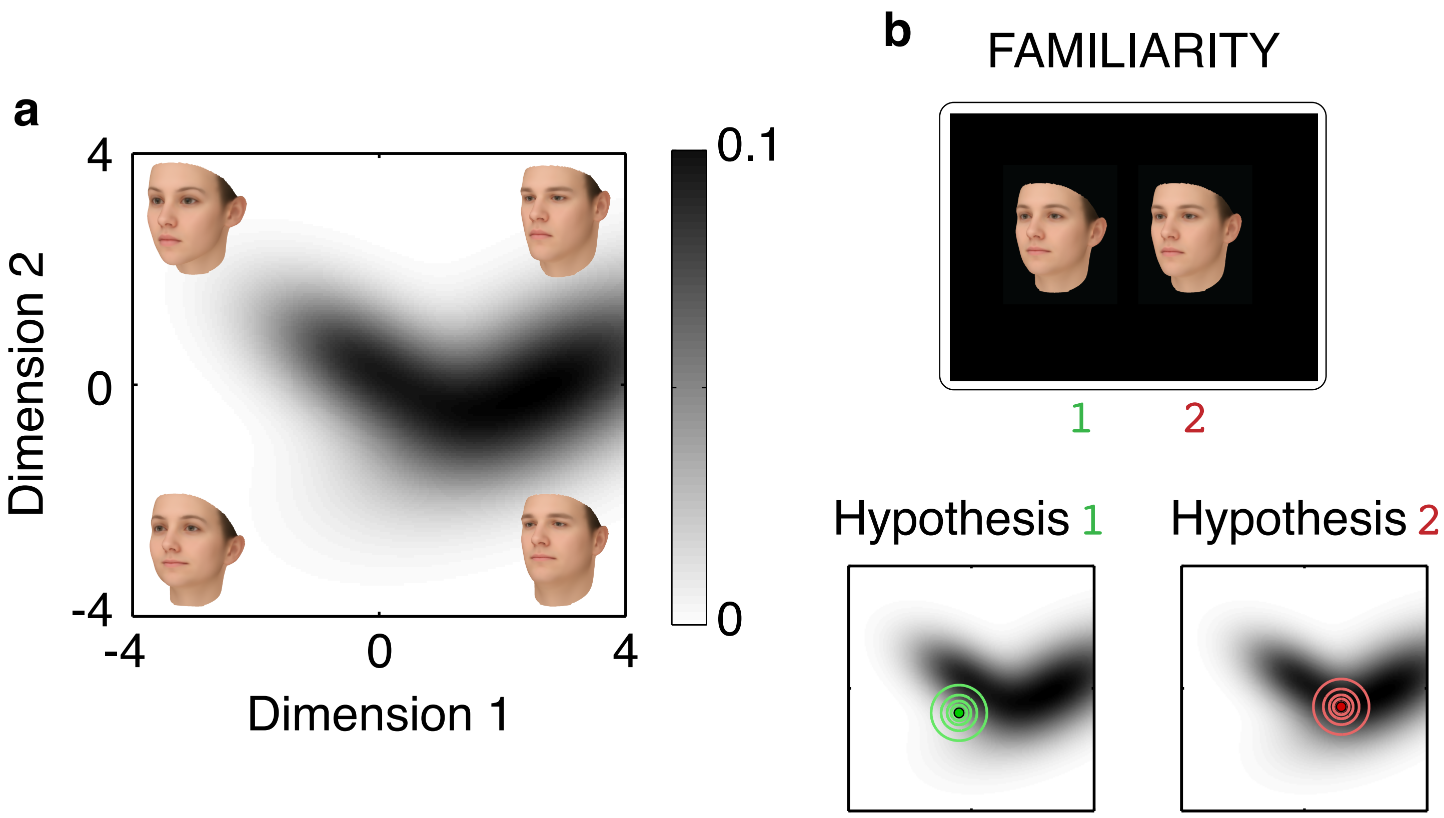
Cognitive tomography

Houlsby et al (2013) Curr Bill



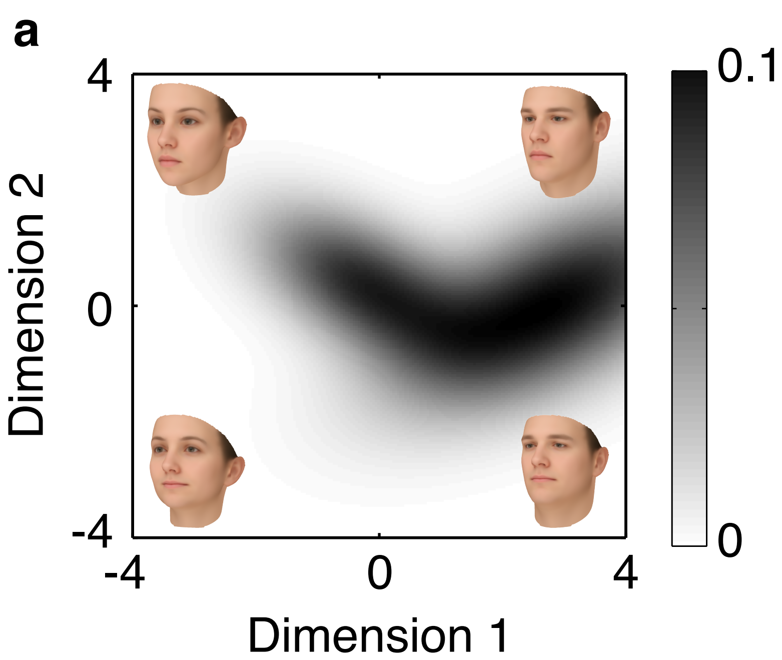
Cognitive tomography

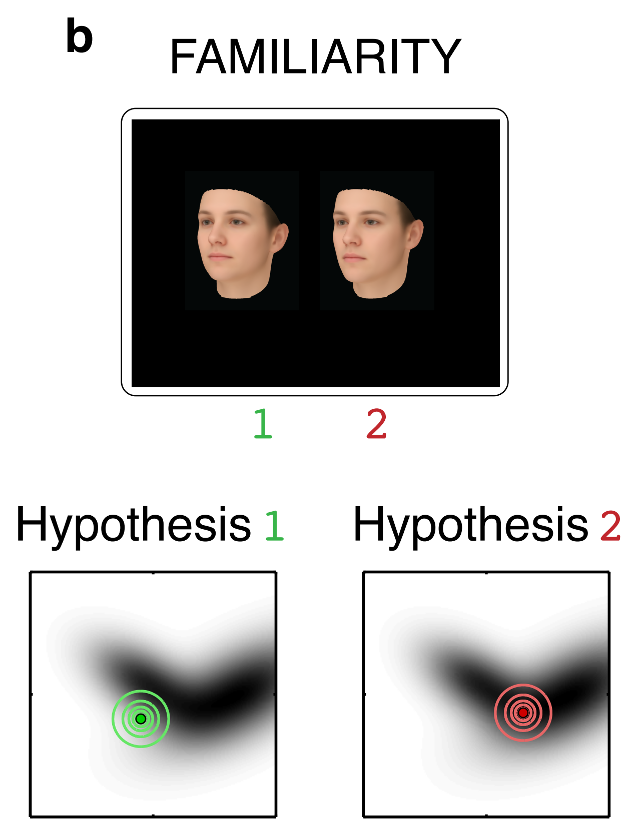
Houlsby et al (2013) Curr Bill

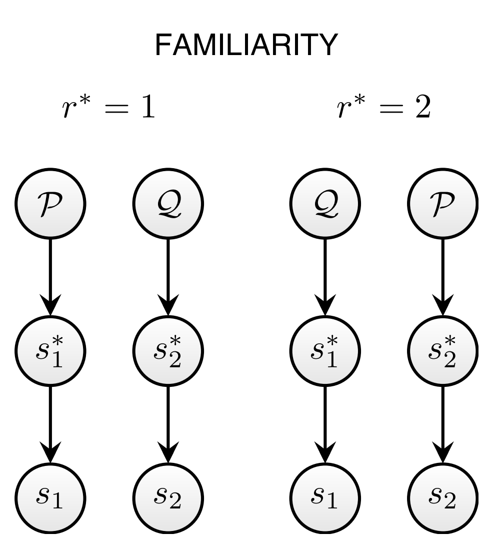


Cognitive tomography

Houlsby et al (2013) Curr Bill

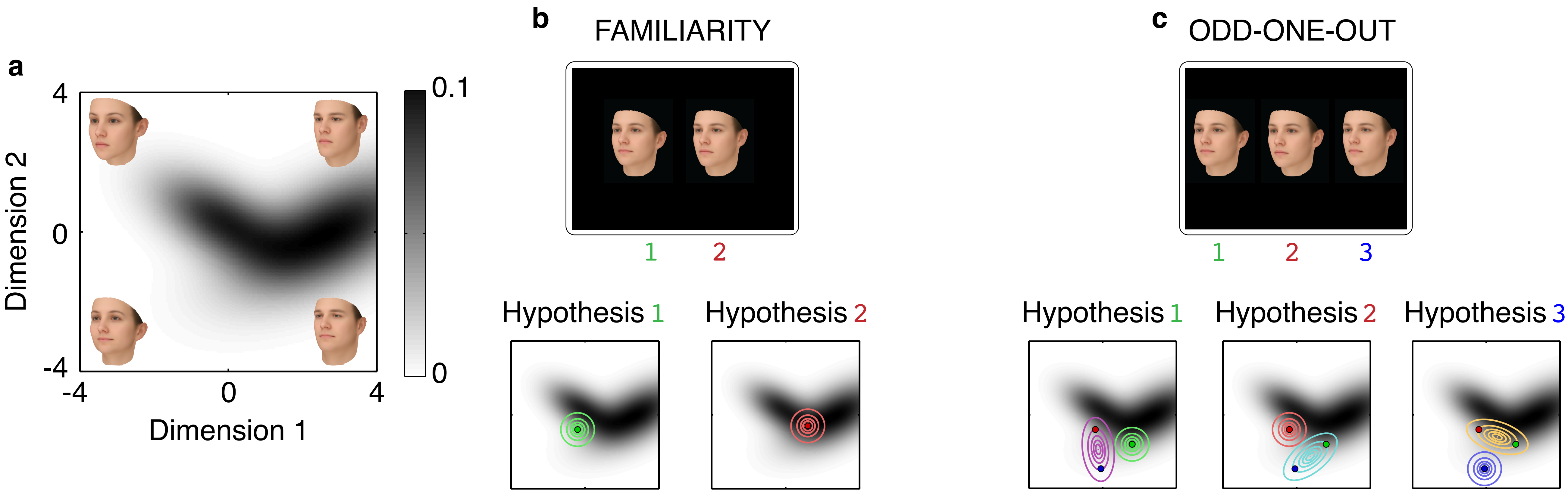


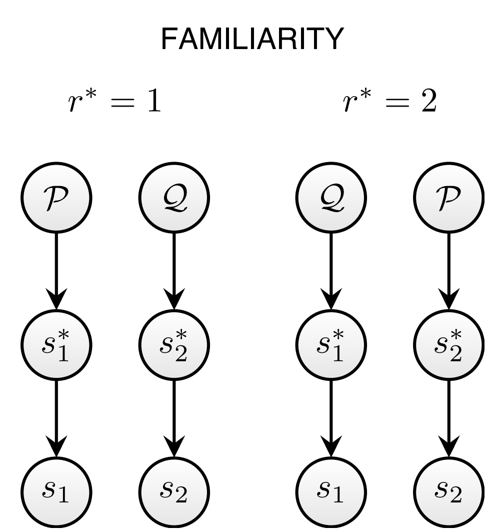




Cognitive tomography

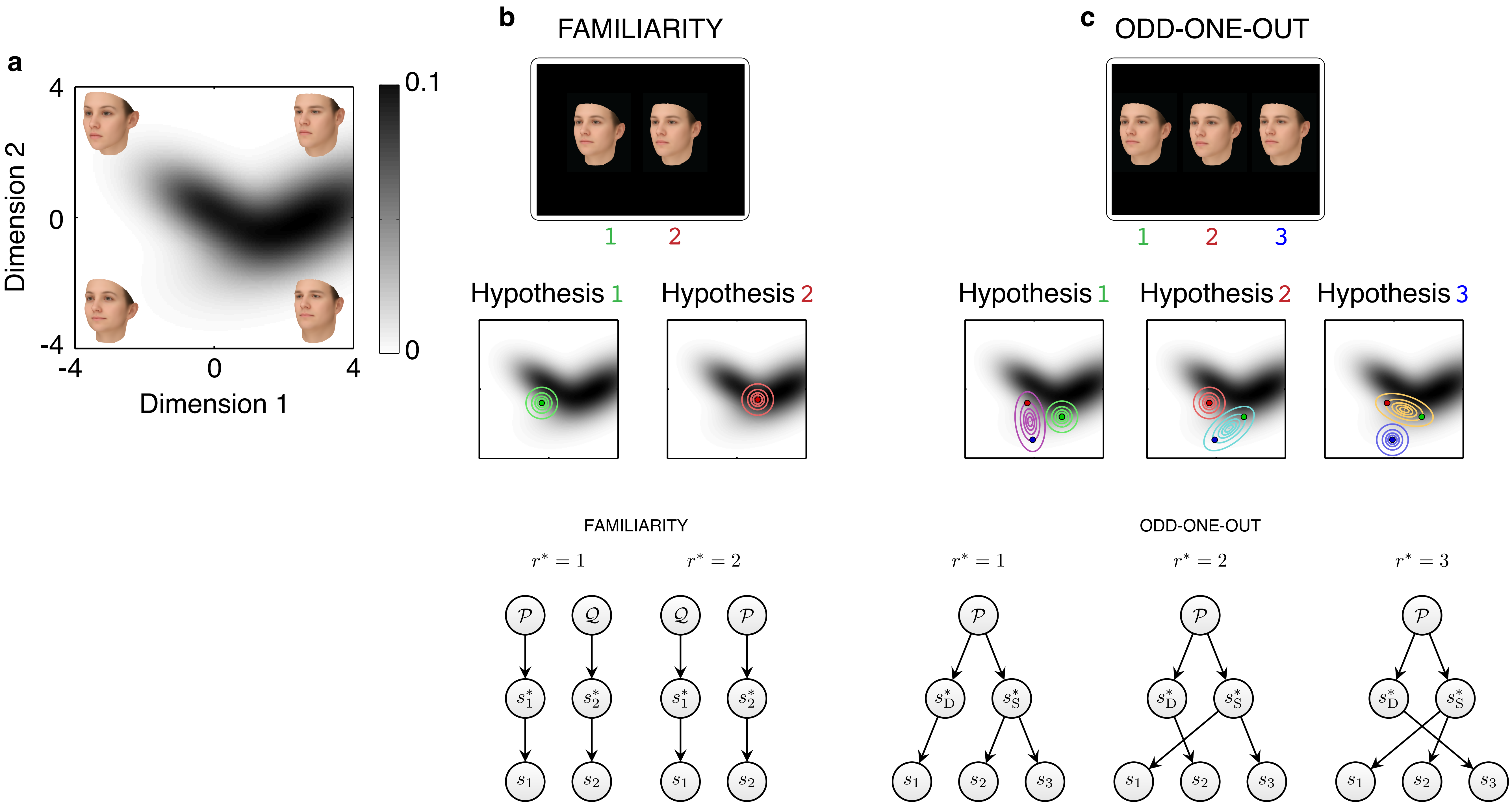
Houlsby et al (2013) Curr Bill





Cognitive tomography

Houlsby et al (2013) Curr Bill



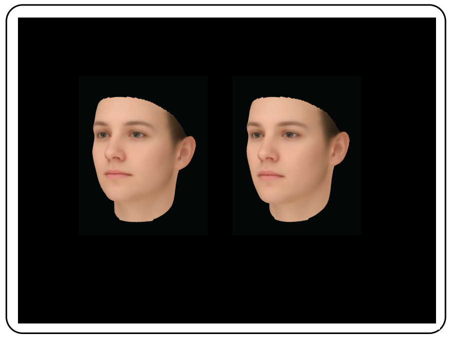
Cognitive tomography

Houlsby et al (2013) Curr Bill

Cognitive tomography

Houlsby et al (2013) Curr Bill

familiarity task

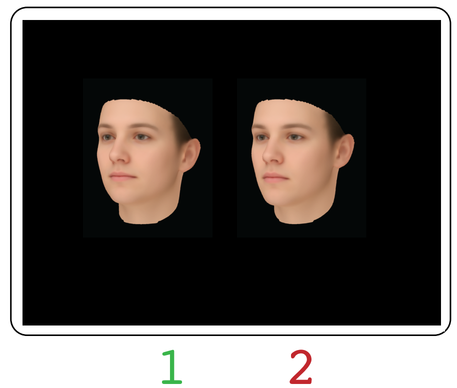


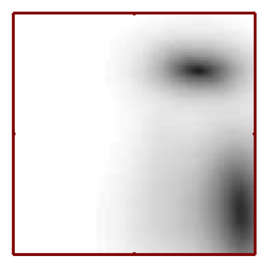
1 2

Cognitive tomography

Houlsby et al (2013) Curr Bill

familiarity task



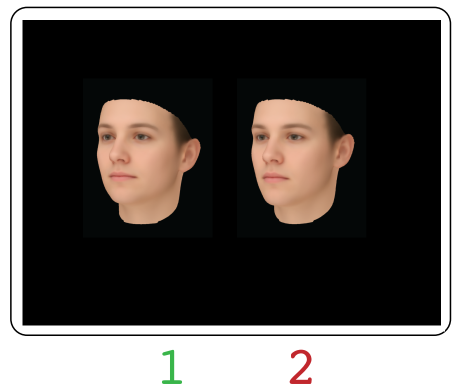


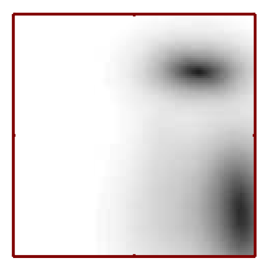
subject #1

Cognitive tomography

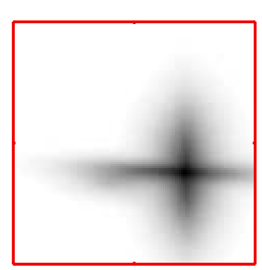
Houlsby et al (2013) Curr Bill

familiarity task





subject #1

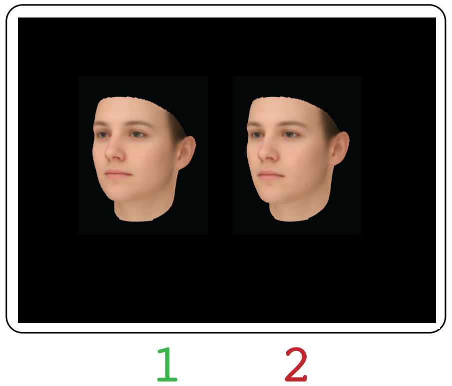


subject #2

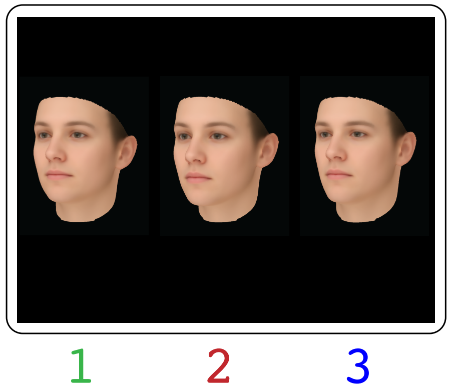
Cognitive tomography

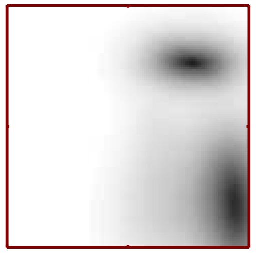
Houlsby et al (2013) Curr Bill

familiarity task

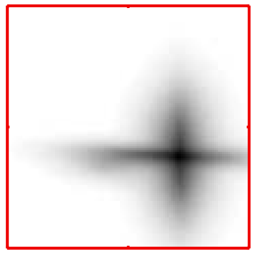


odd-one-out task





subject #1

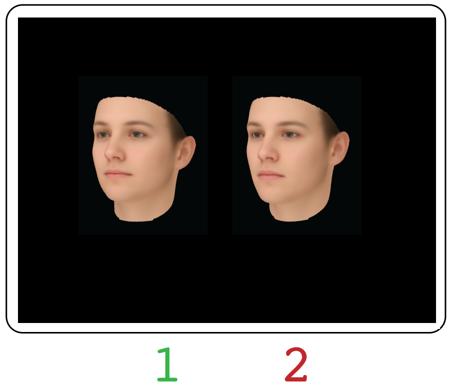


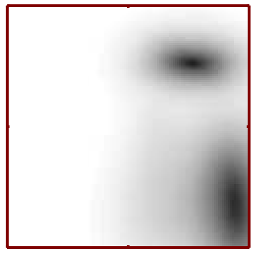
subject #2

Cognitive tomography

Houlsby et al (2013) Curr Bill

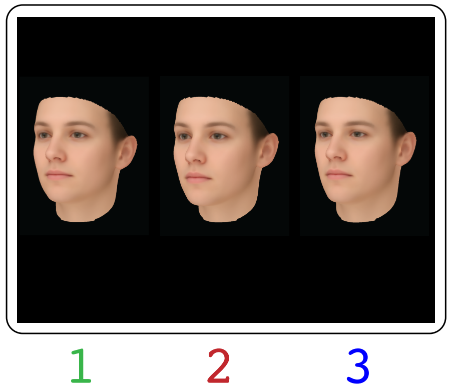
familiarity task

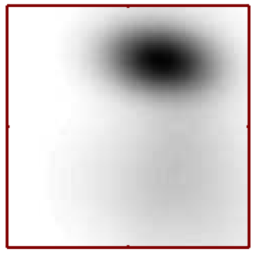




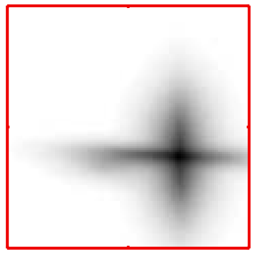
subject #1

odd-one-out task





subject #1



subject #2

Cognitive tomography

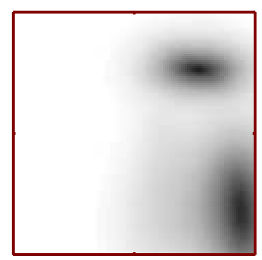
Houlsby et al (2013) Curr Bill

familiarity task

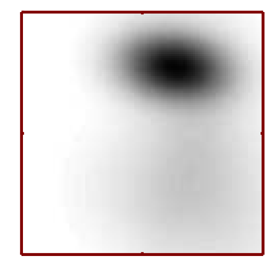
1 2

odd-one-out task

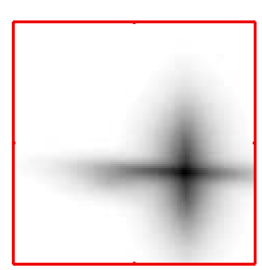
1 2 3



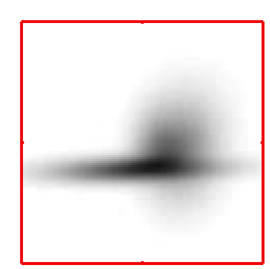
subject #1



subject #1



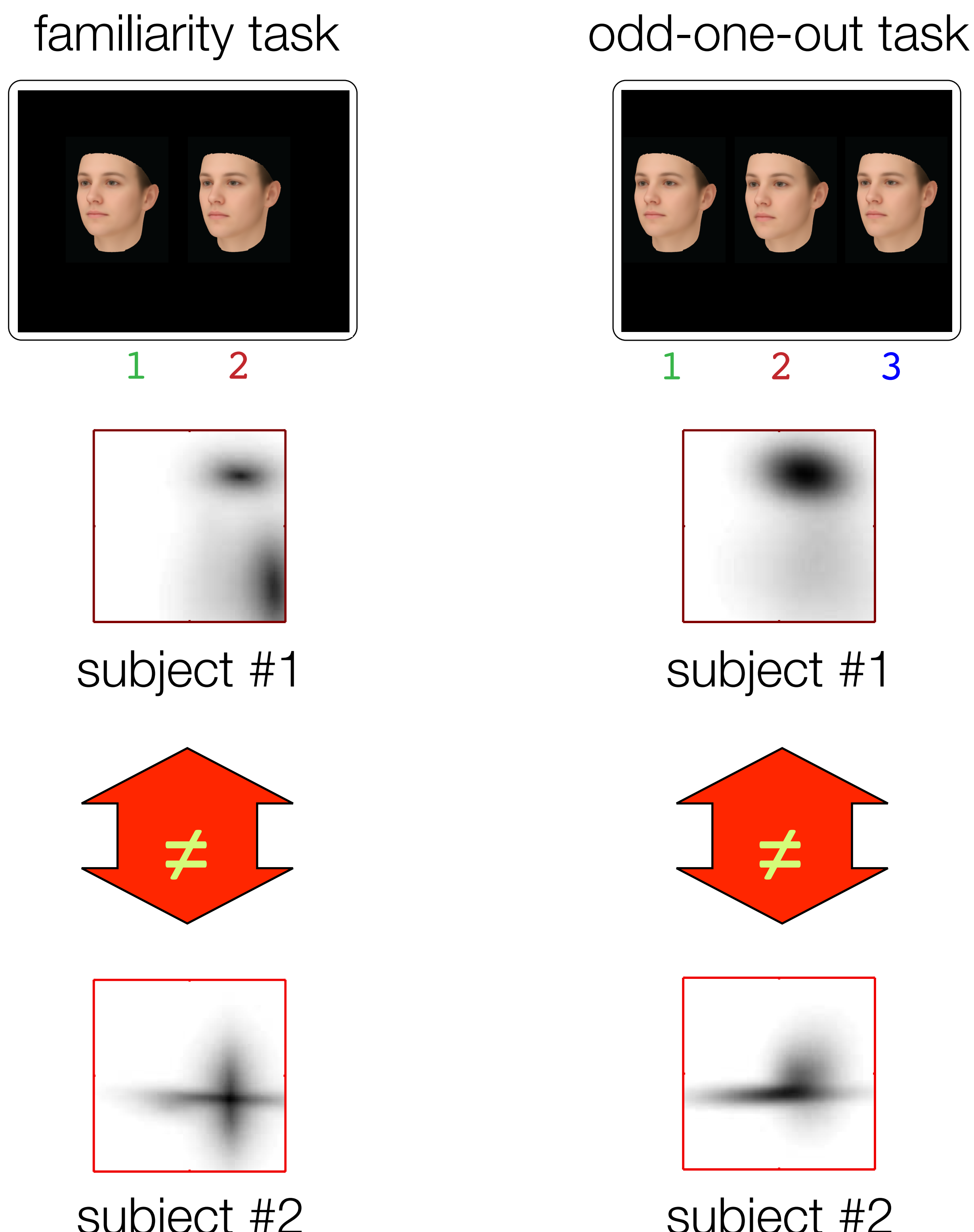
subject #2



subject #2

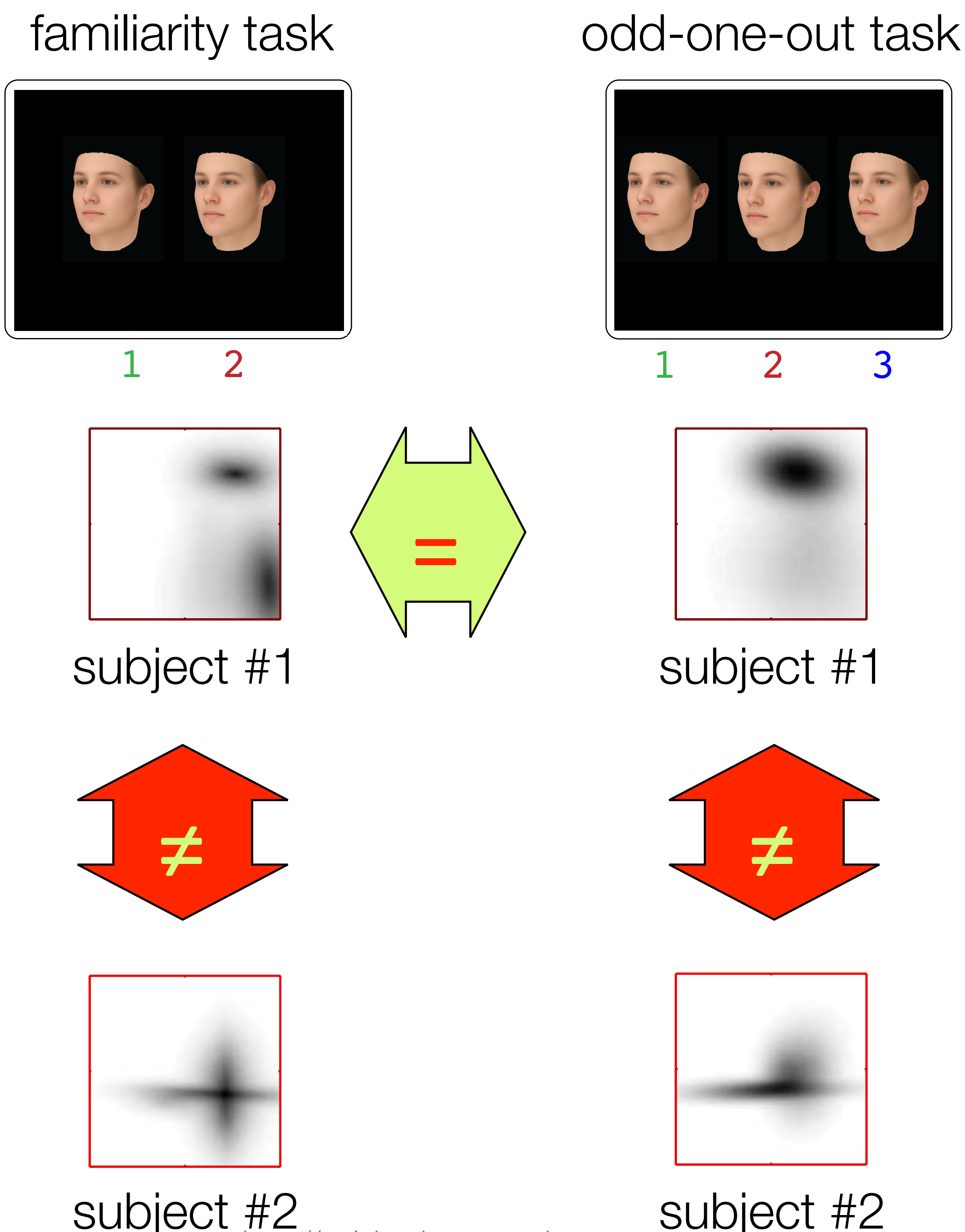
Cognitive tomography

Houlsby et al (2013) Curr Bill



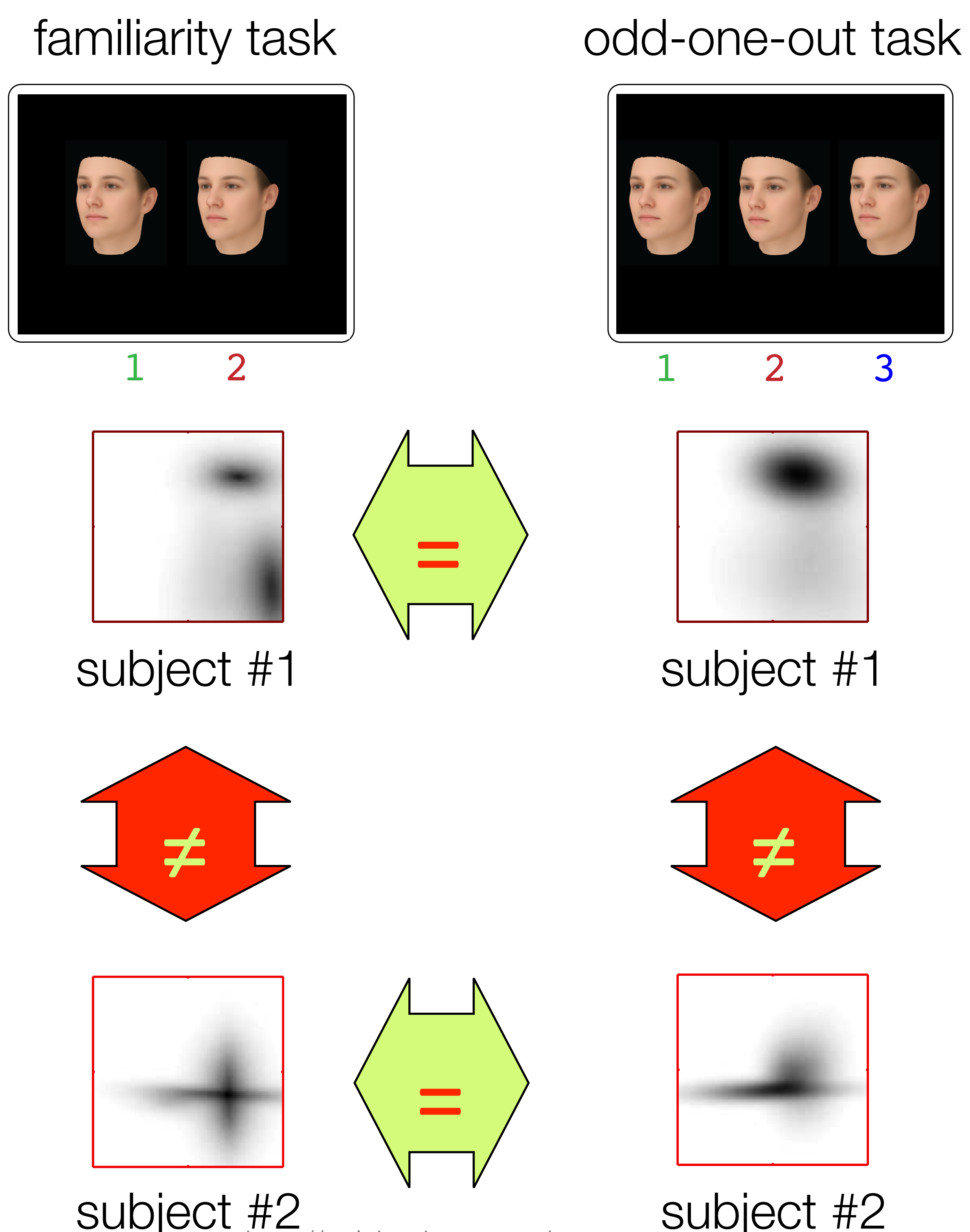
Cognitive tomography

Houlsby et al (2013) Curr Bill



Cognitive tomography

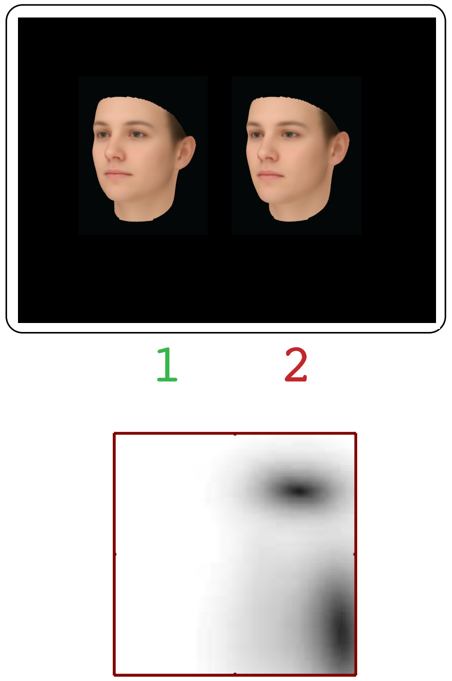
Houlsby et al (2013) Curr Bill



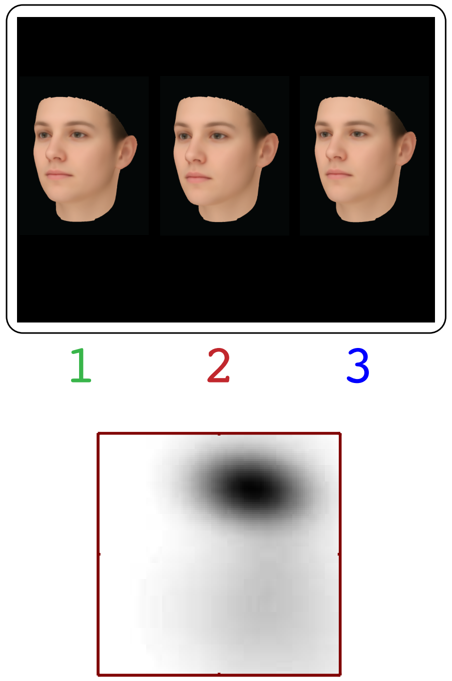
Cognitive tomography

Houlsby et al (2013) Curr Bill

familiarity task



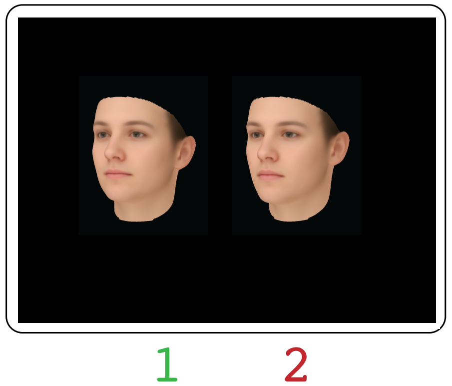
odd-one-out task

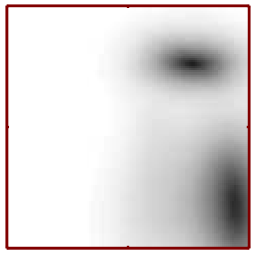


Cognitive tomography

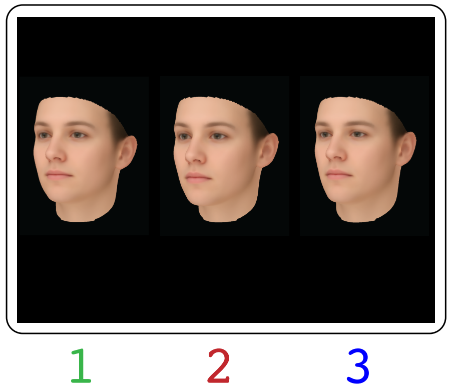
Houlsby et al (2013) Curr Bill

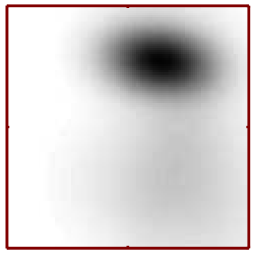
familiarity task

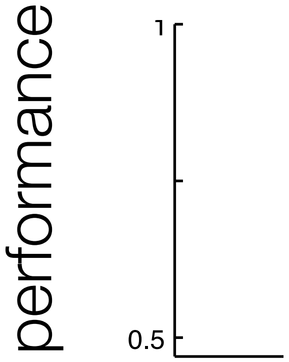




odd-one-out task





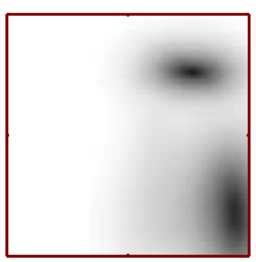


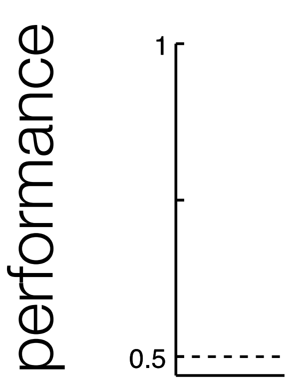
Cognitive tomography

Houlsby et al (2013) Curr Bill

familiarity task

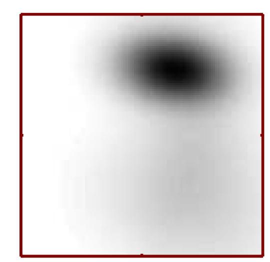
1 2





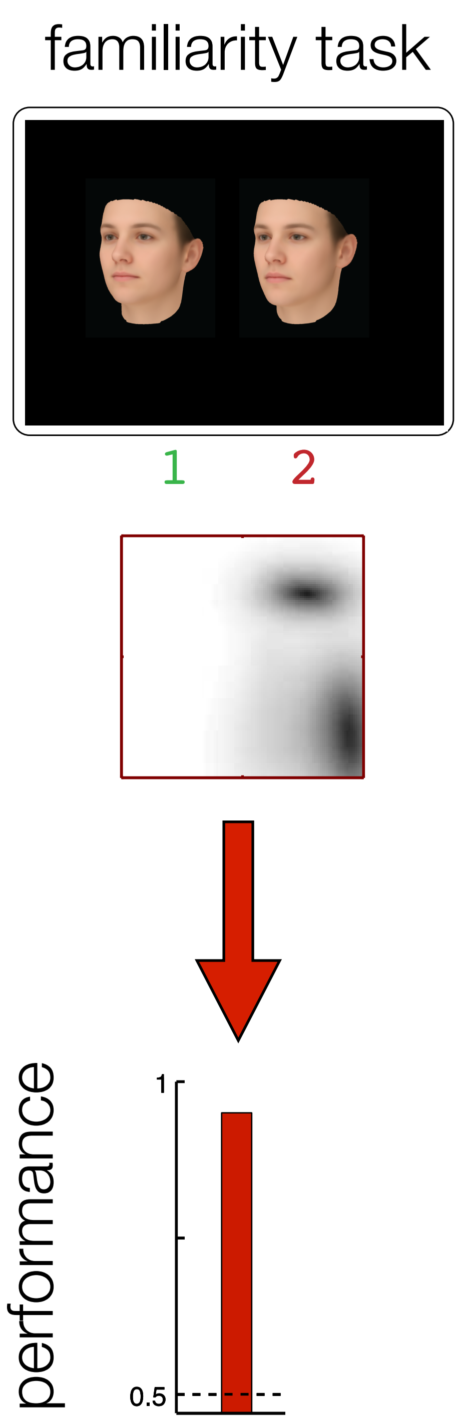
odd-one-out task

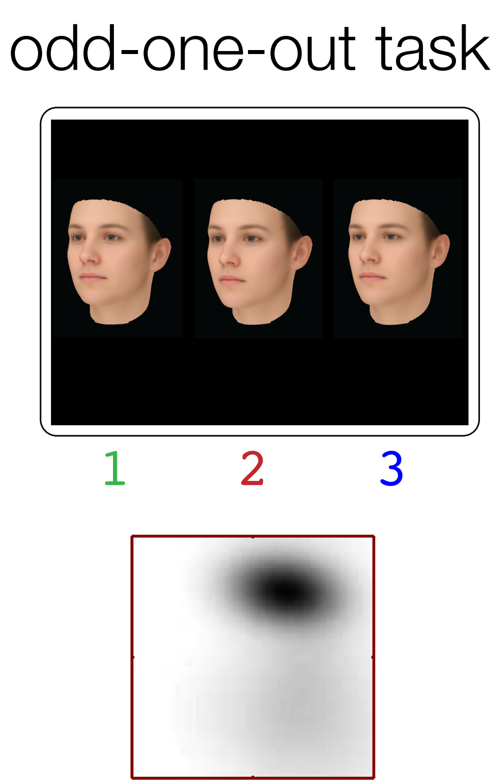
1 2 3



Cognitive tomography

Houlsby et al (2013) Curr Bill

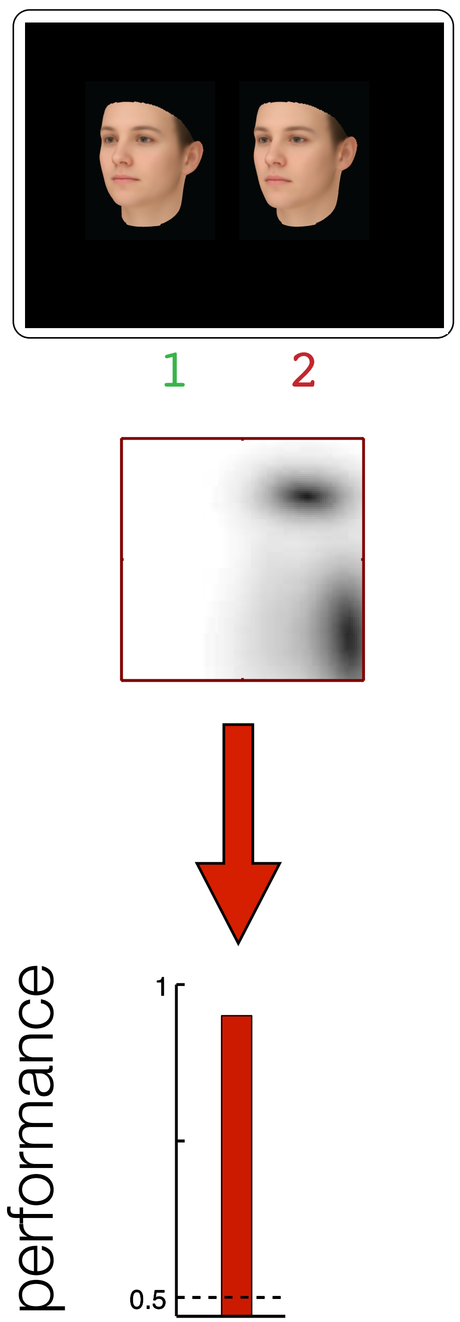




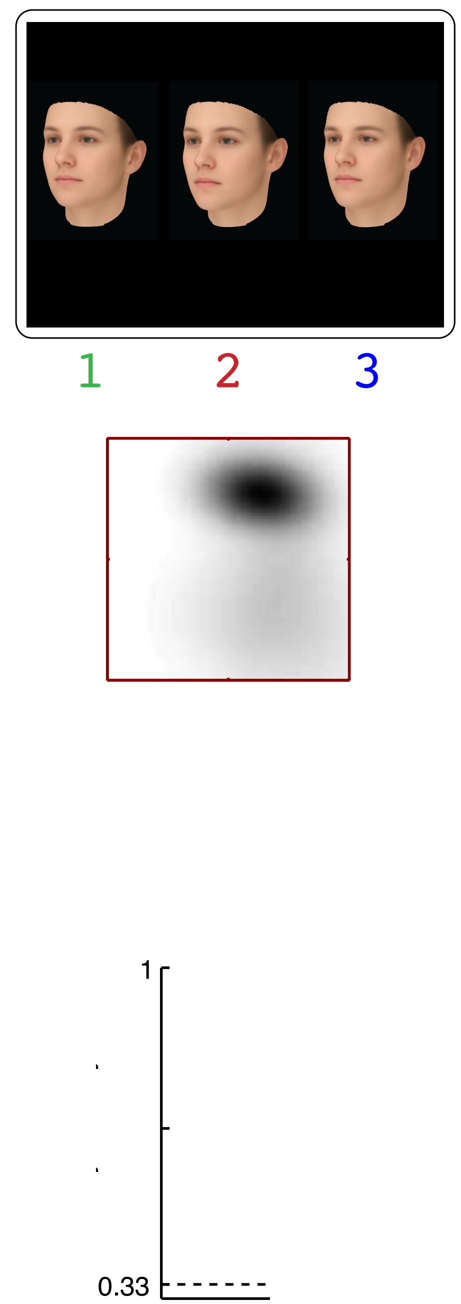
Cognitive tomography

Houlsby et al (2013) Curr Bill

familiarity task



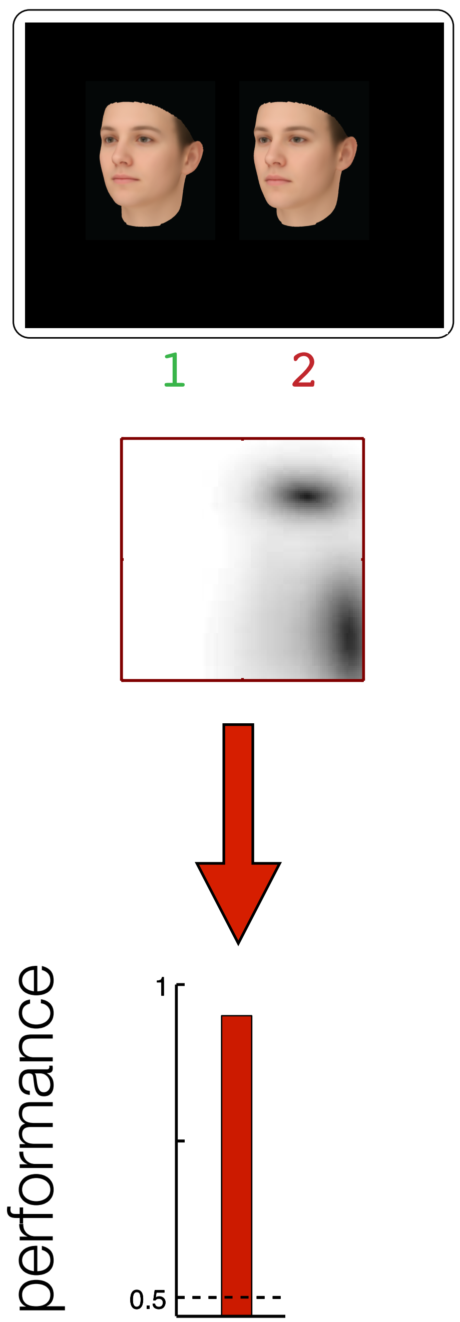
odd-one-out task



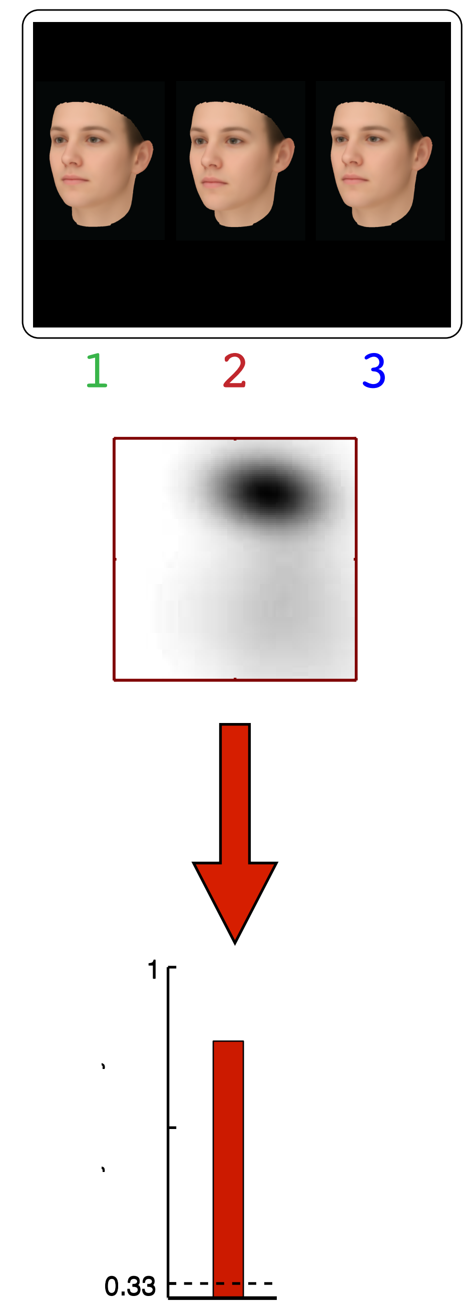
Cognitive tomography

Houlsby et al (2013) Curr Bill

familiarity task



odd-one-out task

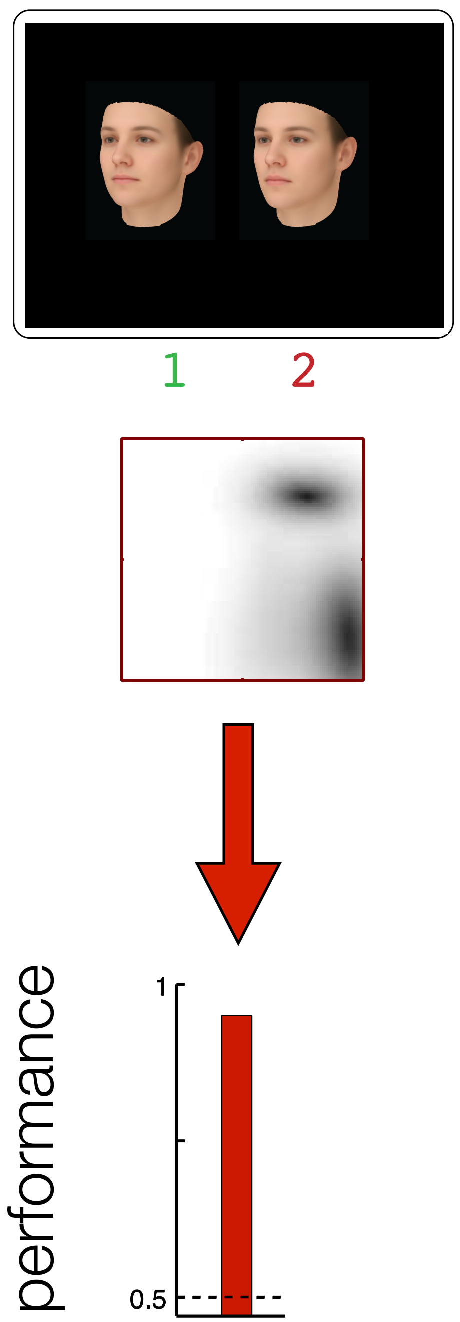


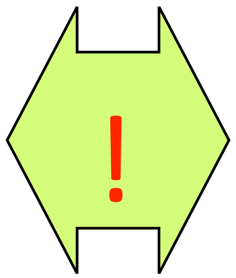
Cognitive tomography

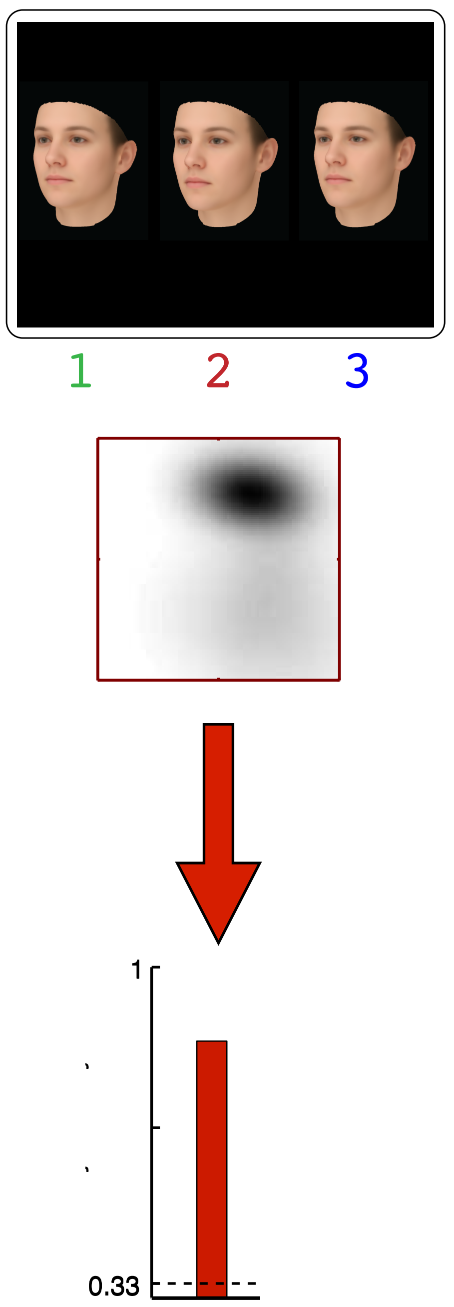
Houlsby et al (2013) Curr Bill

familiarity task

odd-one-out task

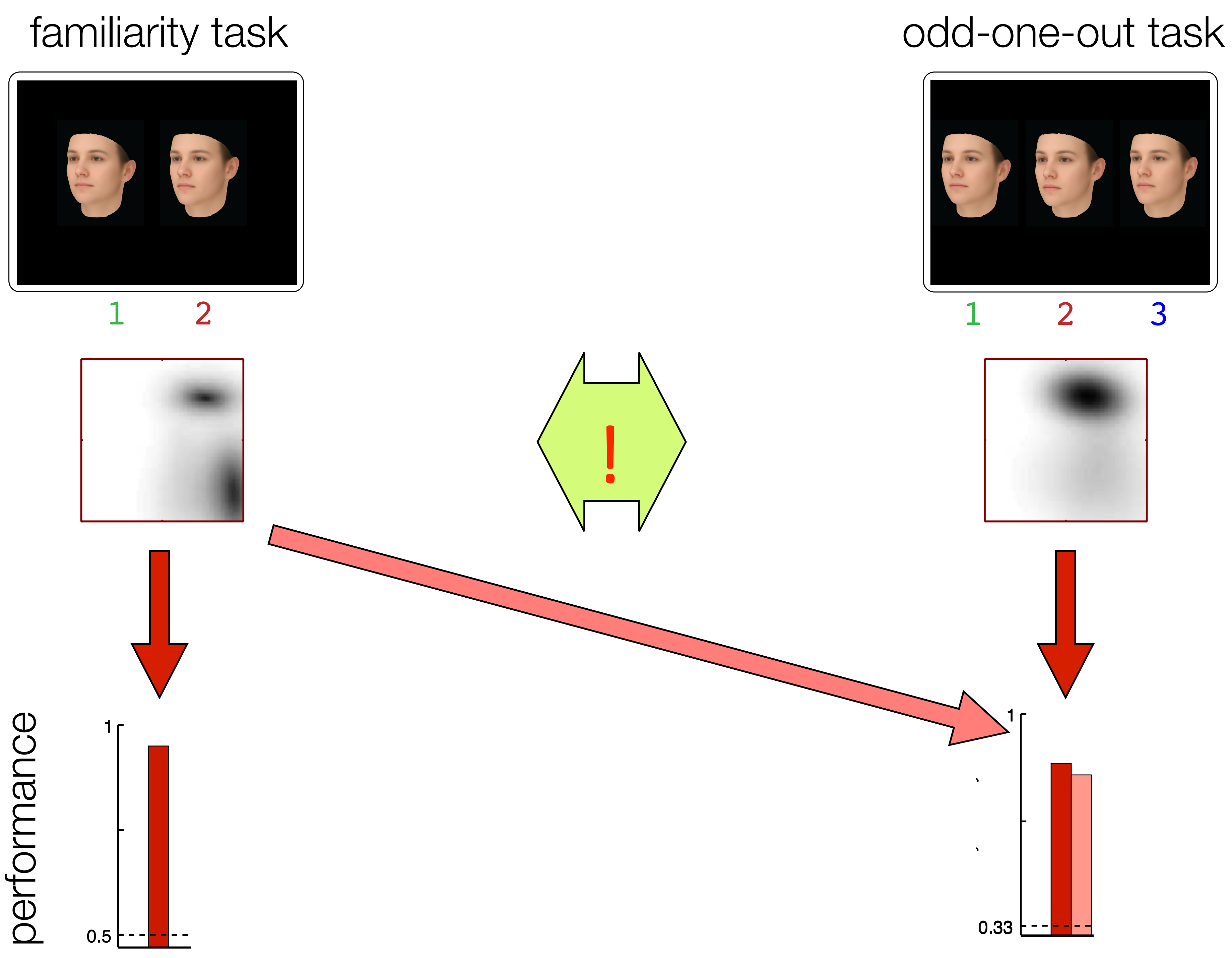






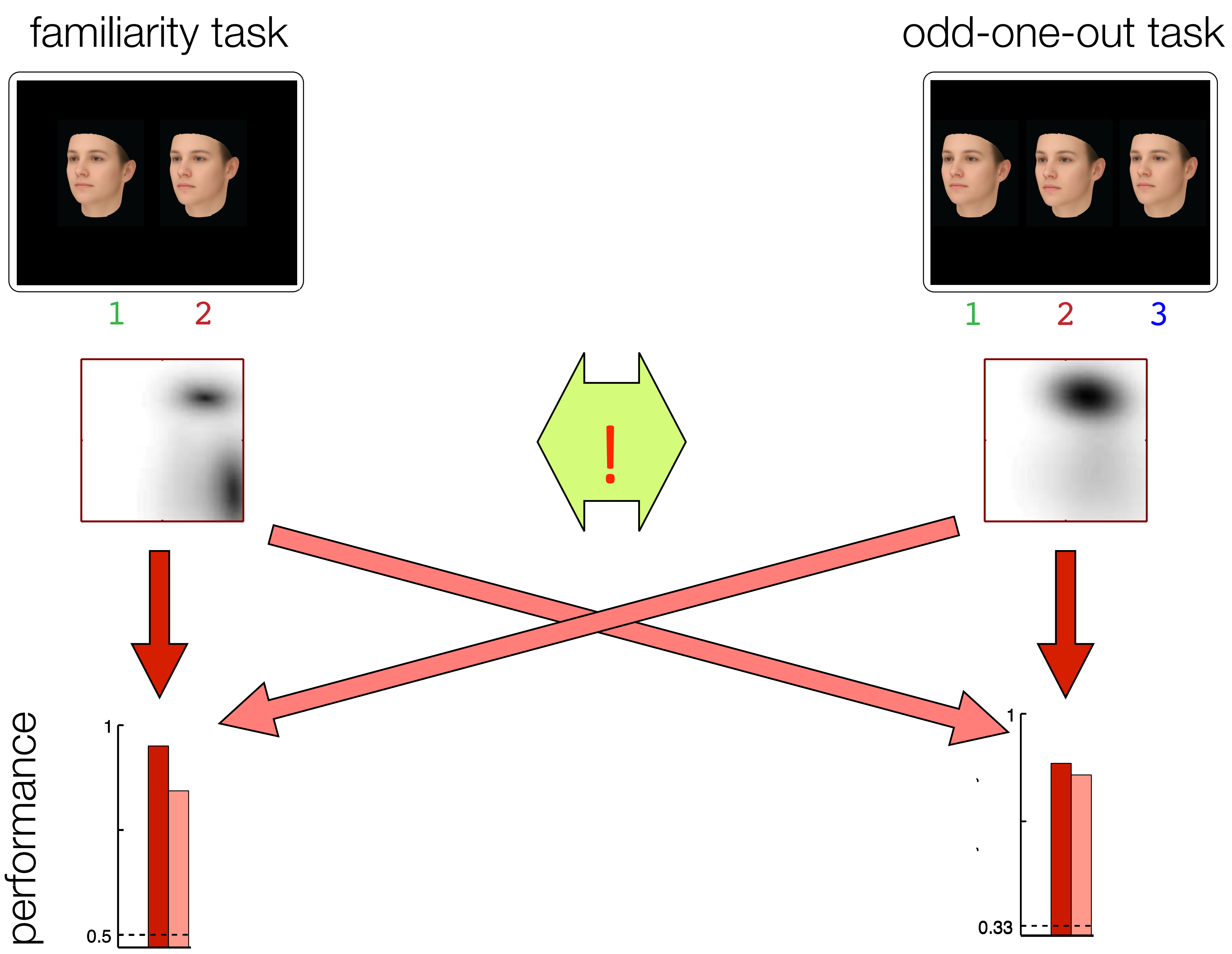
Cognitive tomography

Houlsby et al (2013) Curr Bill



Cognitive tomography

Houlsby et al (2013) Curr Bill



Cognitive tomography

Houlsby et al (2013) Curr Bill