12:00 - 1:00	intro: the why of normative approaches part 1: sensory coding
1:20 - 2:20	part 2: inference part 3: action selection
2:40 - 3:40	hands-on problems
3:40 - 4:00	outlook

wifi | Hilton Honors Meeting
csnventi20

TAs

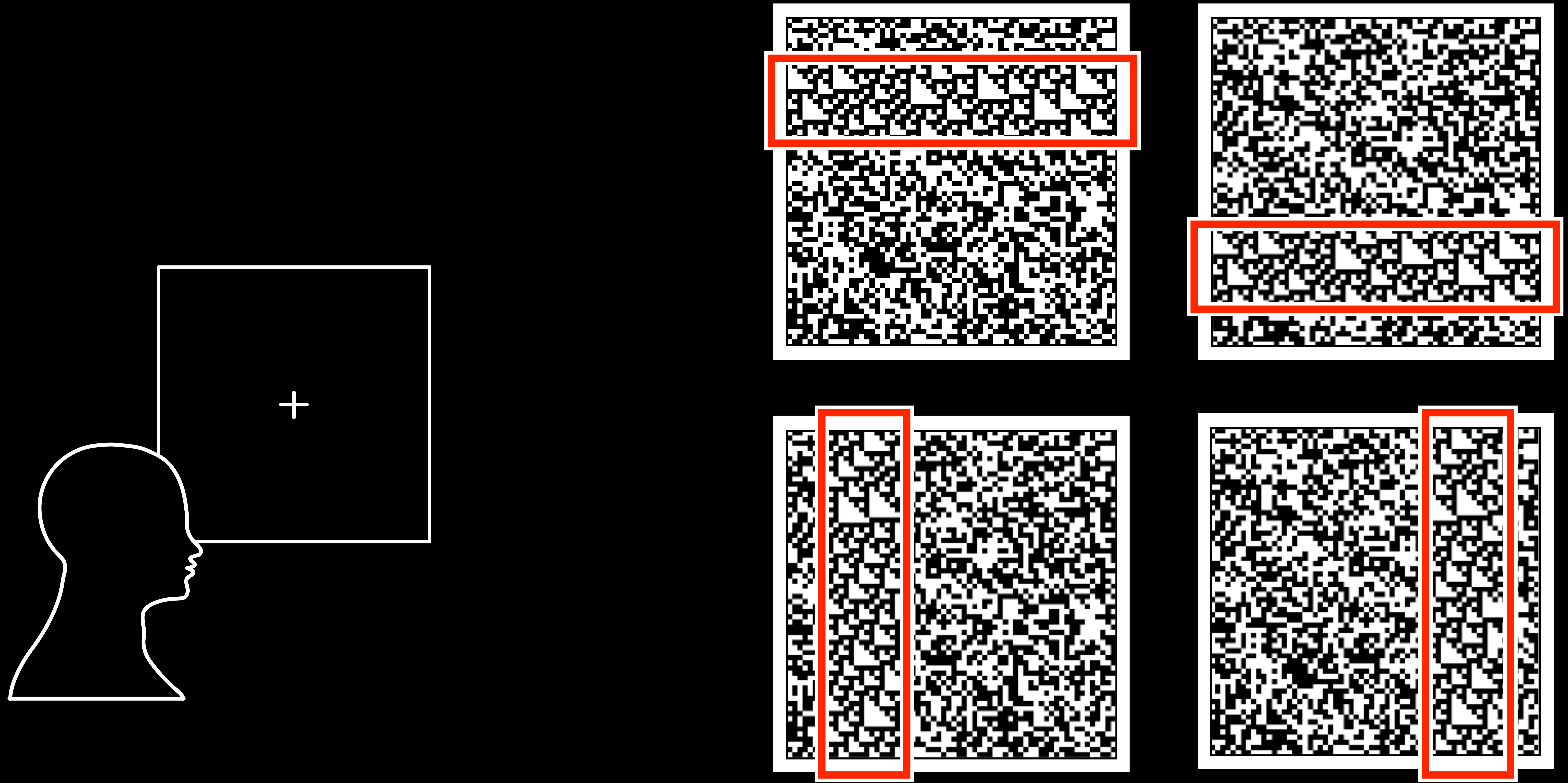
Diego Arribas
Zoe Ashwood
Pierre-Etienne Fiquet
Caroline Haimerl
Anna Kutschireiter
Tzuhsuan Ma (Maz)
Jorge Menendez
Josue Nassar
Marcella Noorman
Sashank Pisupati
Satpreet Singh
Charline Tessereau

A wing would be a most mystifying structure if one did not know that birds flew. One might observe that it could be extended a considerable distance, that it had a smooth covering of feathers with conspicuous markings, that it was operated by powerful muscles, and that strength and lightness were prominent features of its construction. These are important facts, but by themselves they do not tell us that birds fly. Yet without knowing this, and without understanding something of the principles of flight, a more detailed examination of the wing itself would probably be unrewarding.

Horace Barlow, 1961

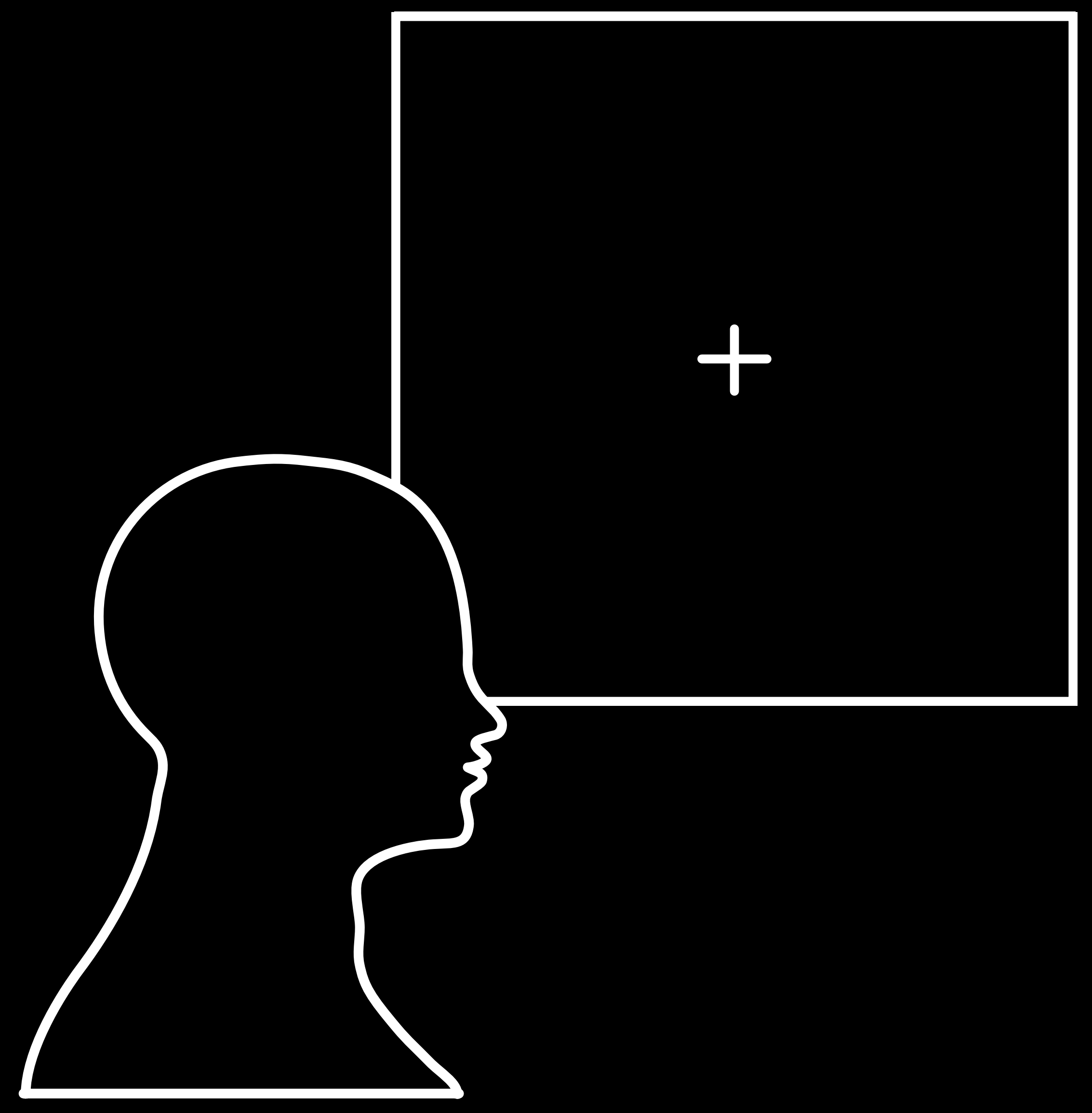
psychophysics & stimulus generation: Jonathan Victor, Mary Conte

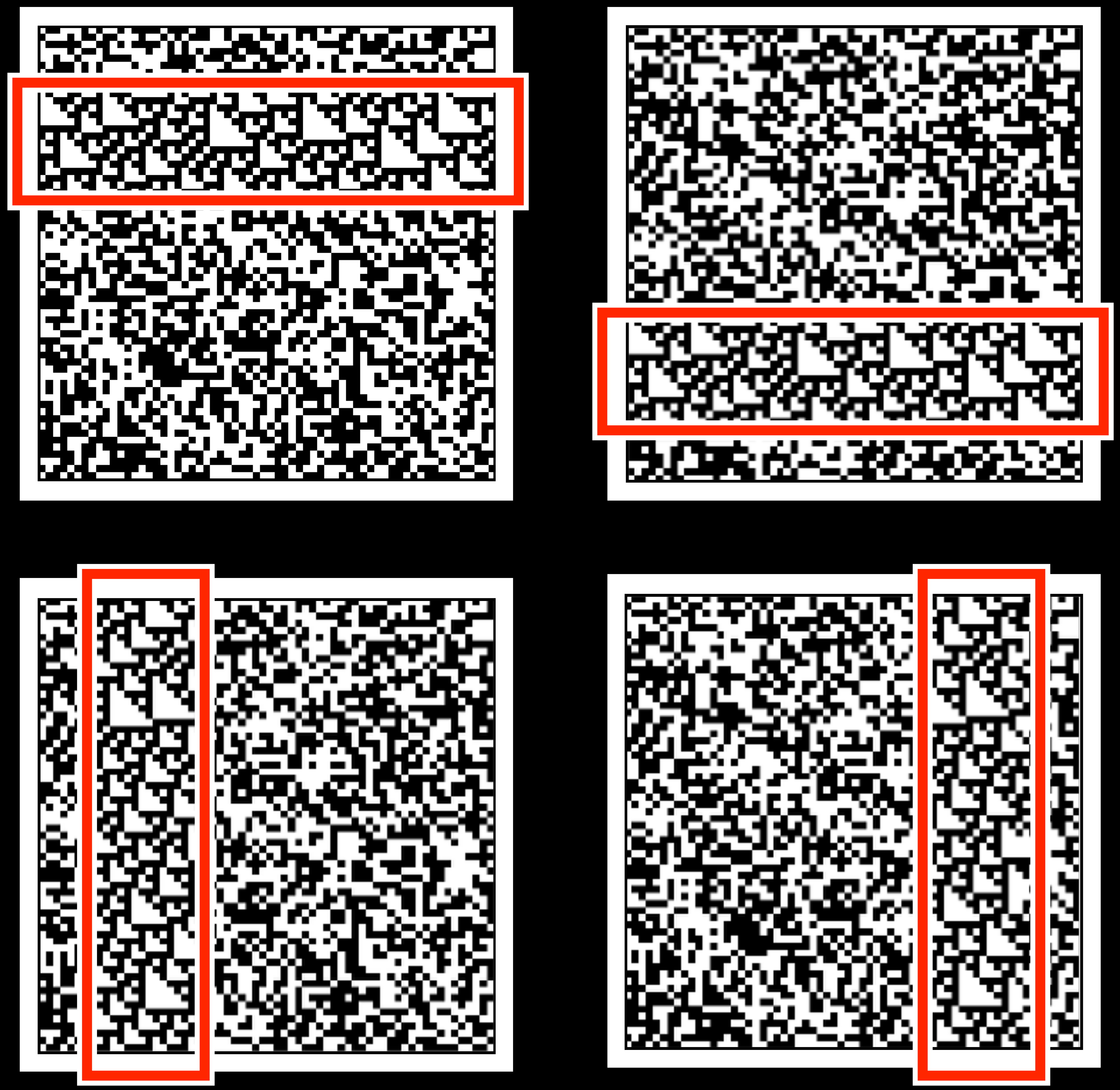
Victor & Conte (2012), JOSA A; Victor, Thengone, & Conte (2013), J Vision



psychophysics & stimulus generation: Jonathan Victor, Mary Conte

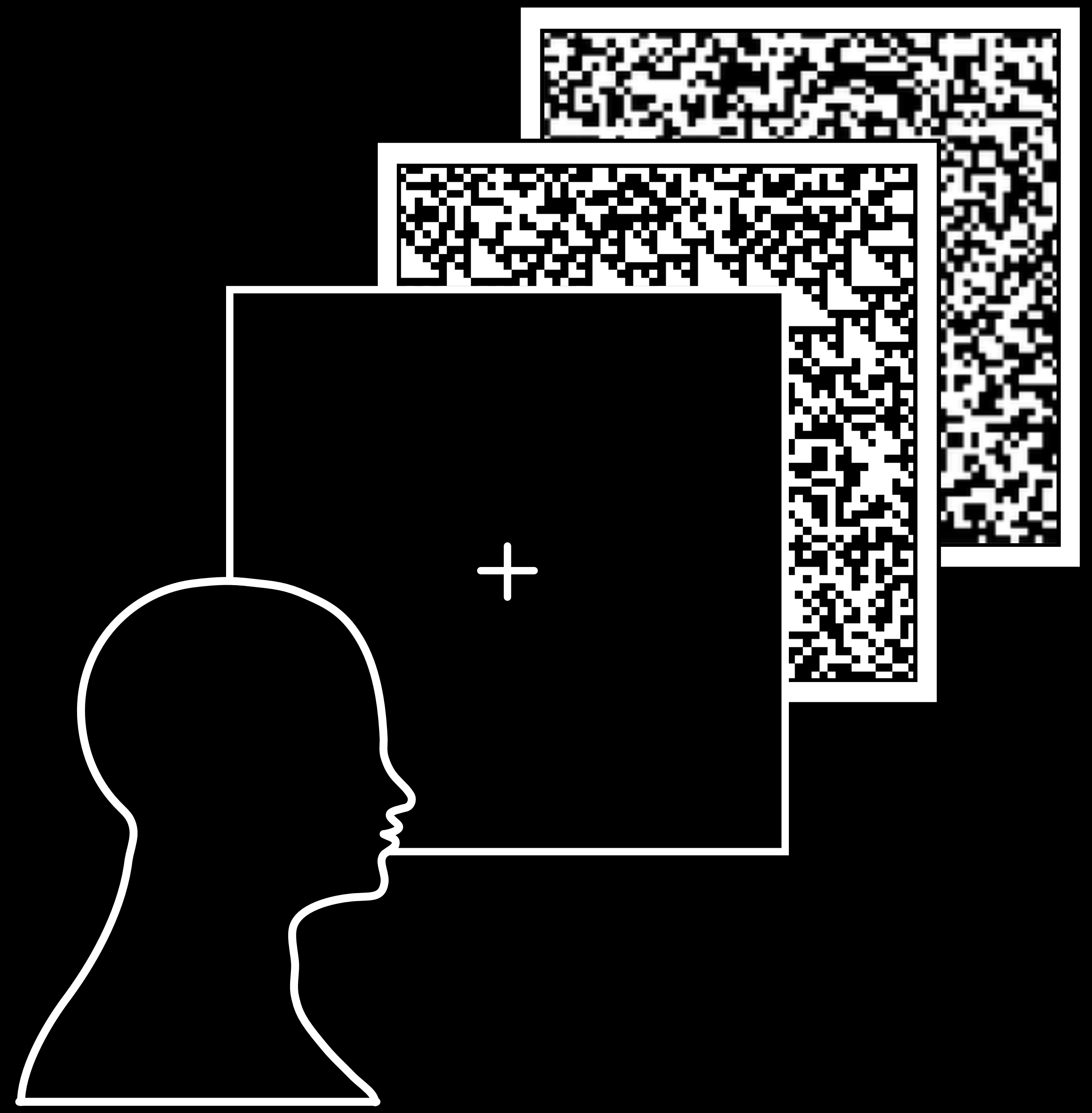
Victor & Conte (2012), JOSA A; Victor, Thengone, & Conte (2013), J Vision

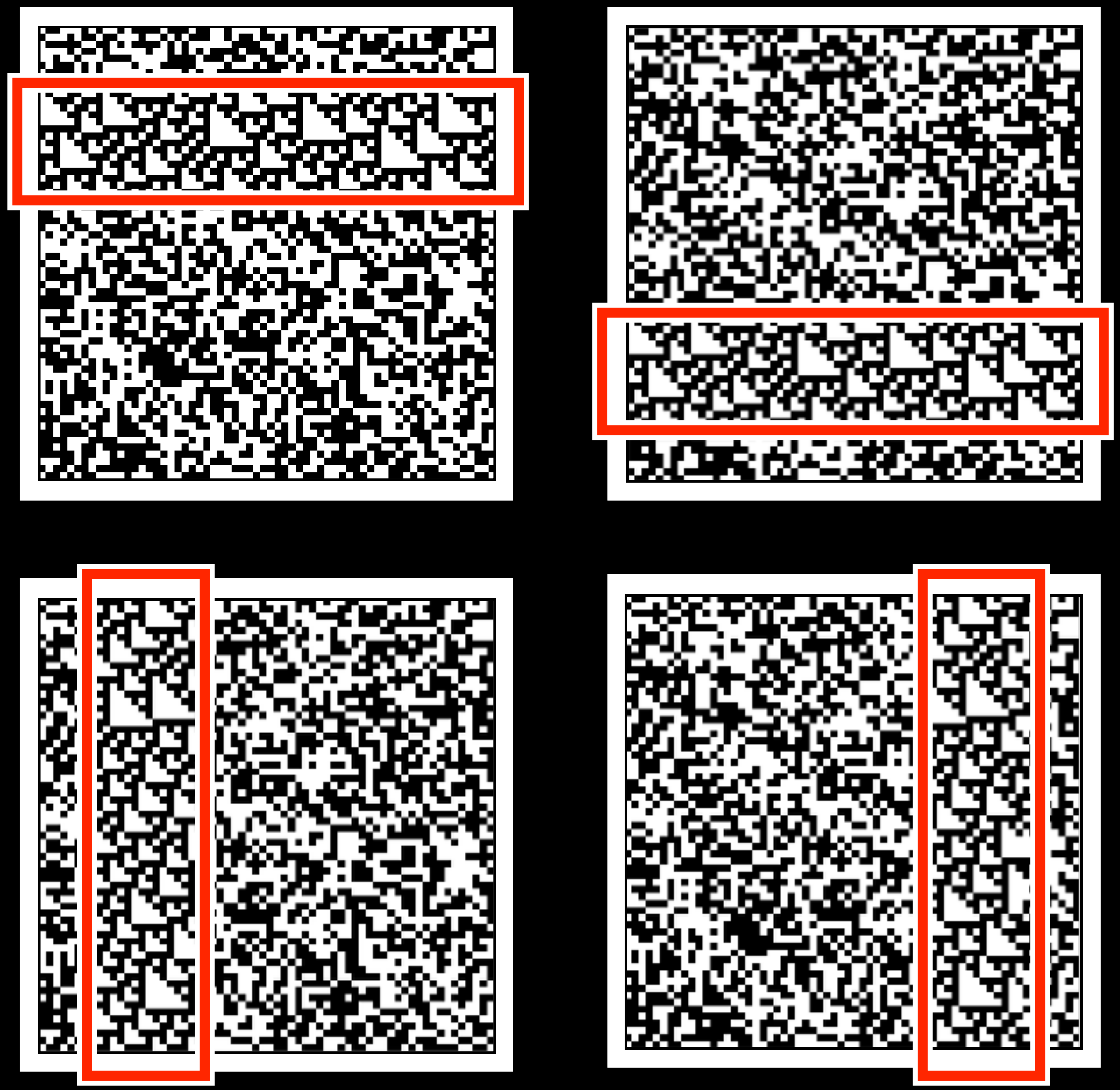




psychophysics & stimulus generation: Jonathan Victor, Mary Conte

Victor & Conte (2012), JOSA A; Victor, Thengone, & Conte (2013), J Vision



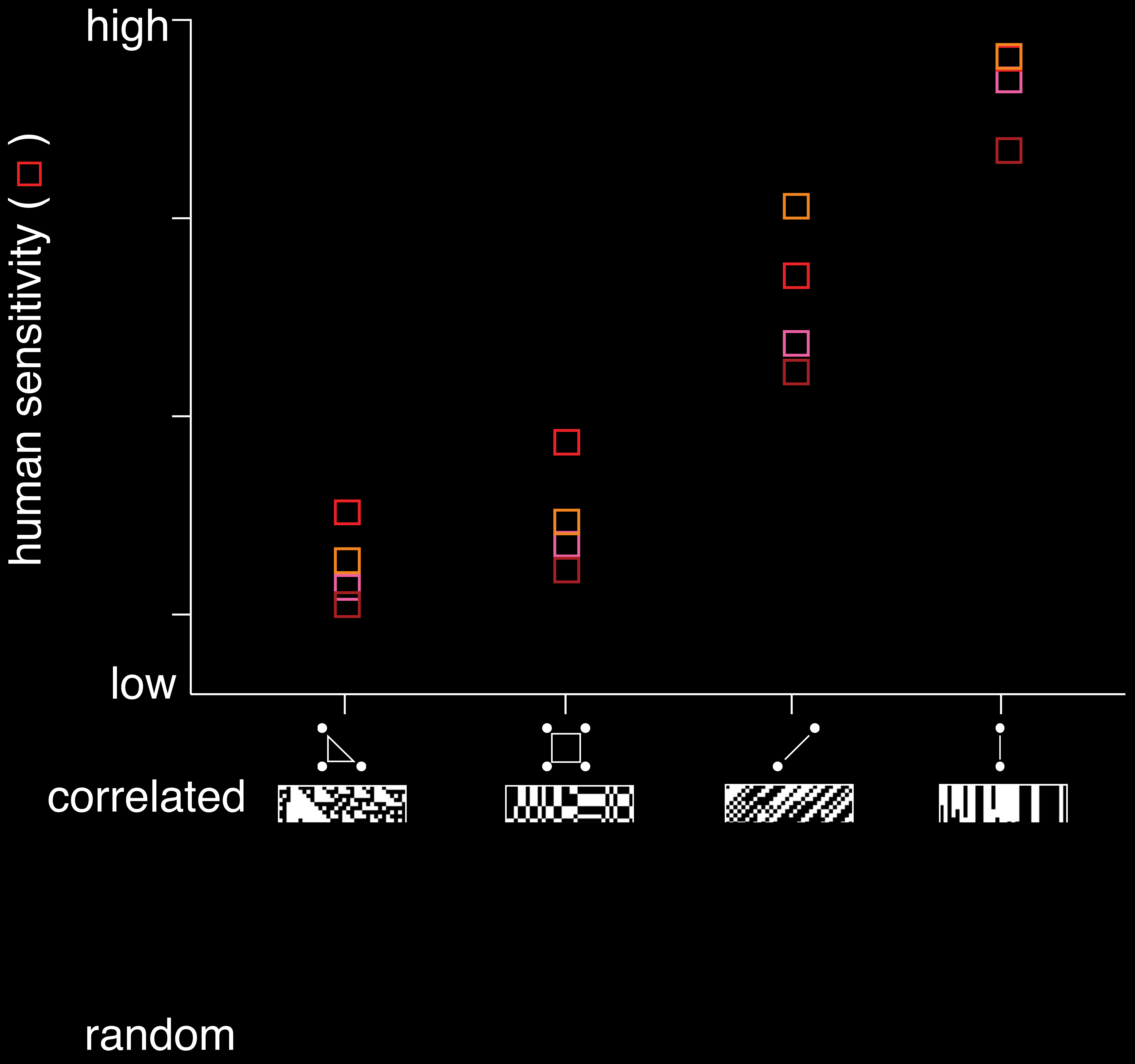


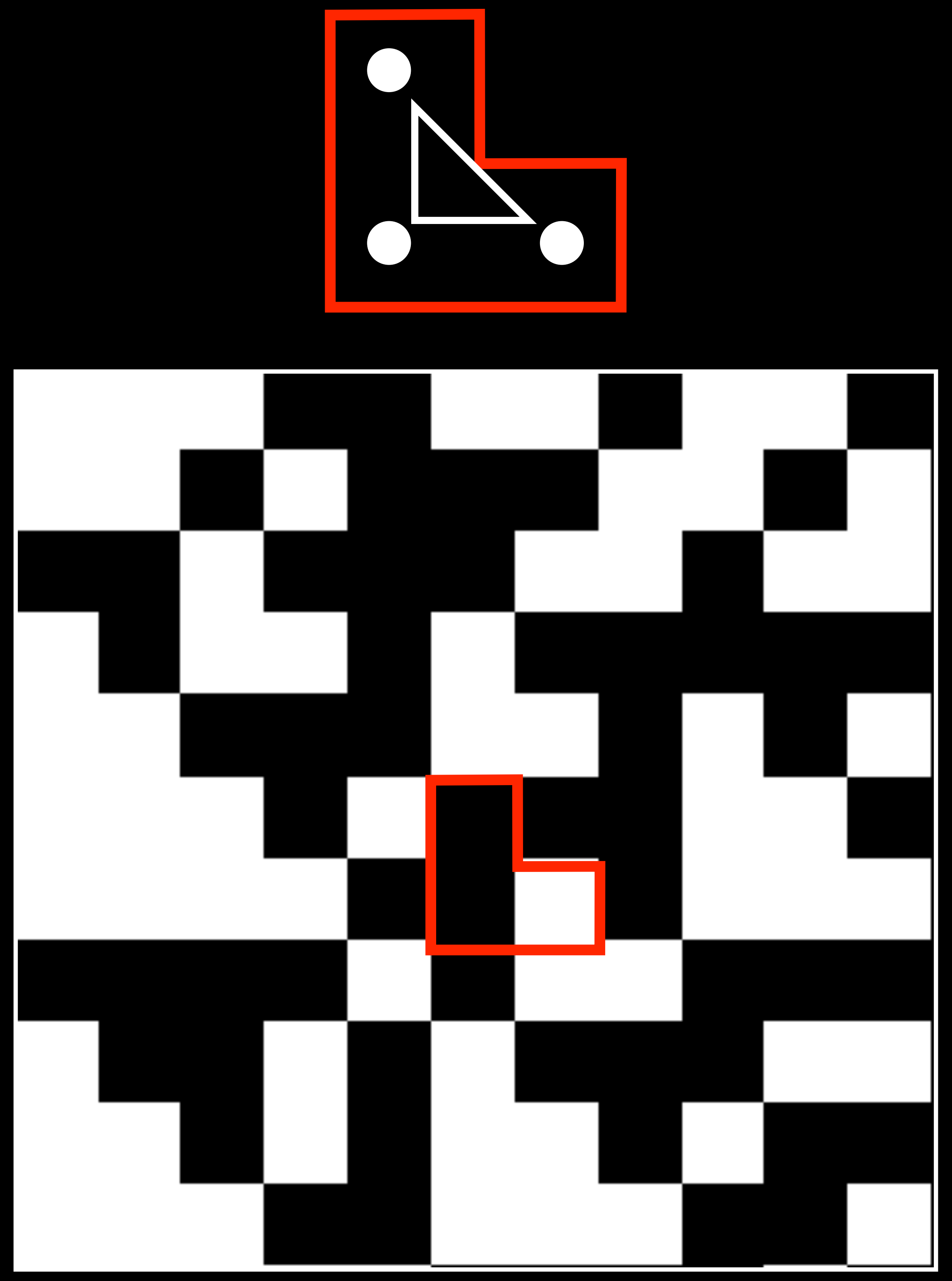
psychophysics & stimulus generation: Jonathan Victor, Mary Conte

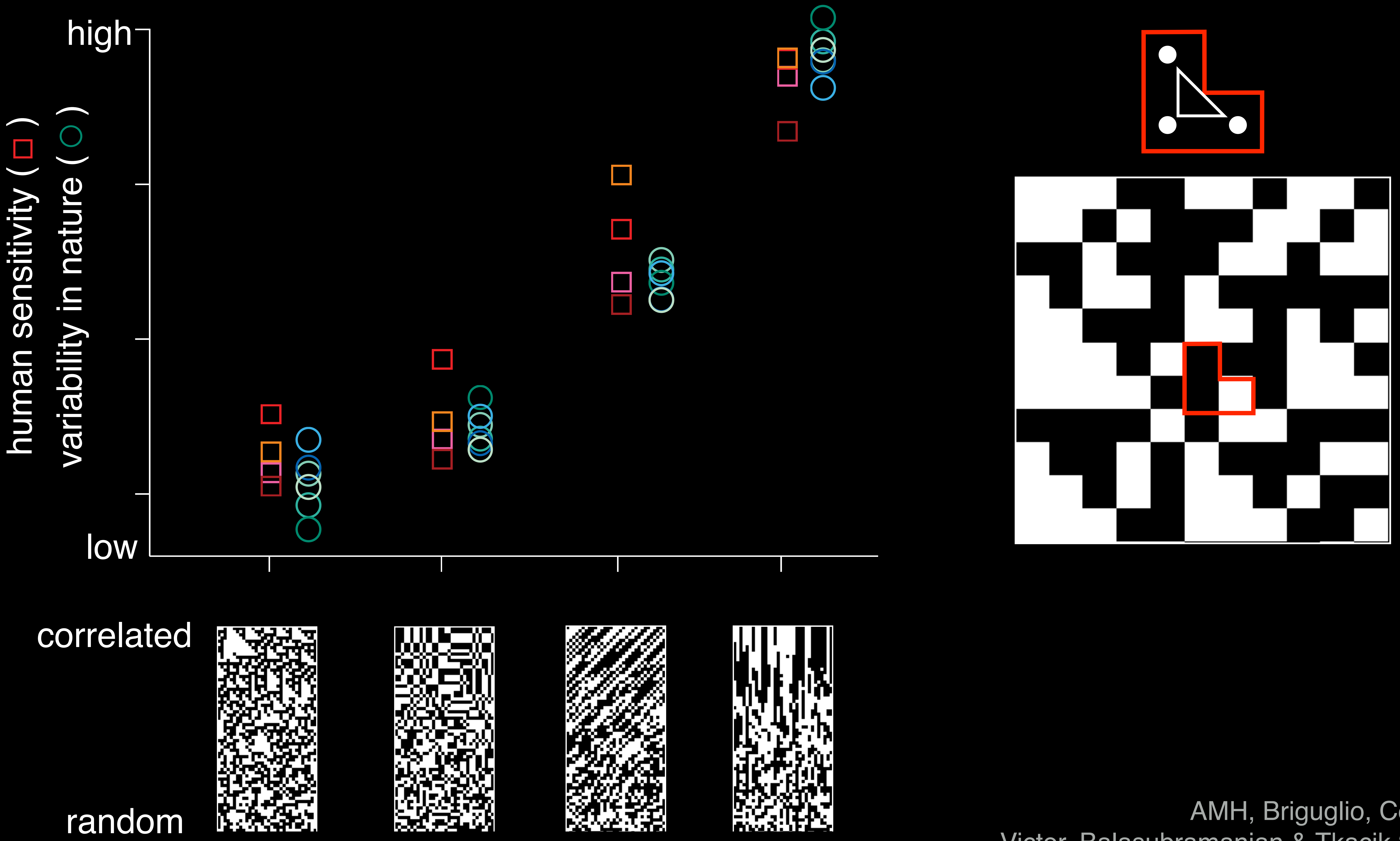
Victor & Conte (2012), JOSA A; Victor, Thengone, & Conte (2013), J Vision

+

+


high
low
human sensitivity ( )
correlated
random





AMH, Briguglio, Conte,
Victor, Balasubramanian & Tkacik 2014

claim:

you should be **more sensitive** to visual features
that are **more variable**

…because they are **more informative**

claim:

you **should be** more sensitive to [variable] visual features …

performance advantage

…**because** they are more informative

lawfulness of the world

the world is lawful

animals & brains can exploit the *lawfulness of the world*
to achieve a *performance advantage*

how might a system do this?

reduce redundancy / build compact representations

combat noise / correct errors

resolve ambiguities / reduce uncertainty

make predictions / improve future performance

the normative approach

lawfulness of the world
performance of a system

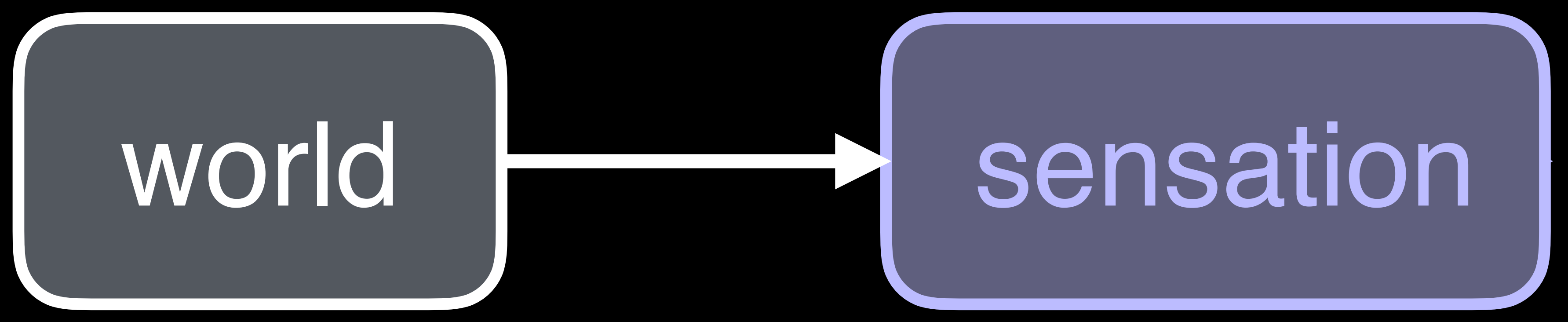
specify:

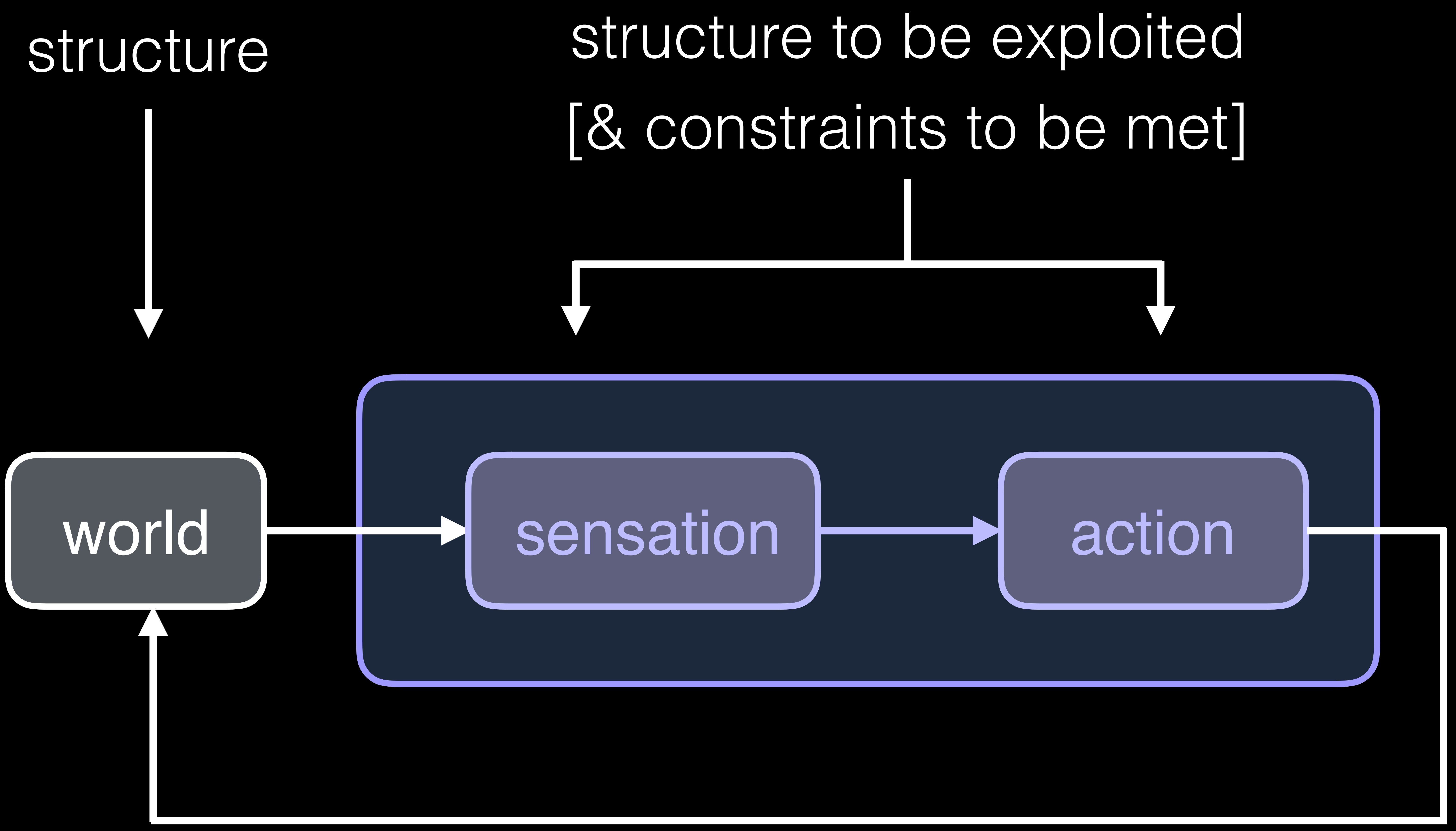
function to be performed [maximizing information about patterns]

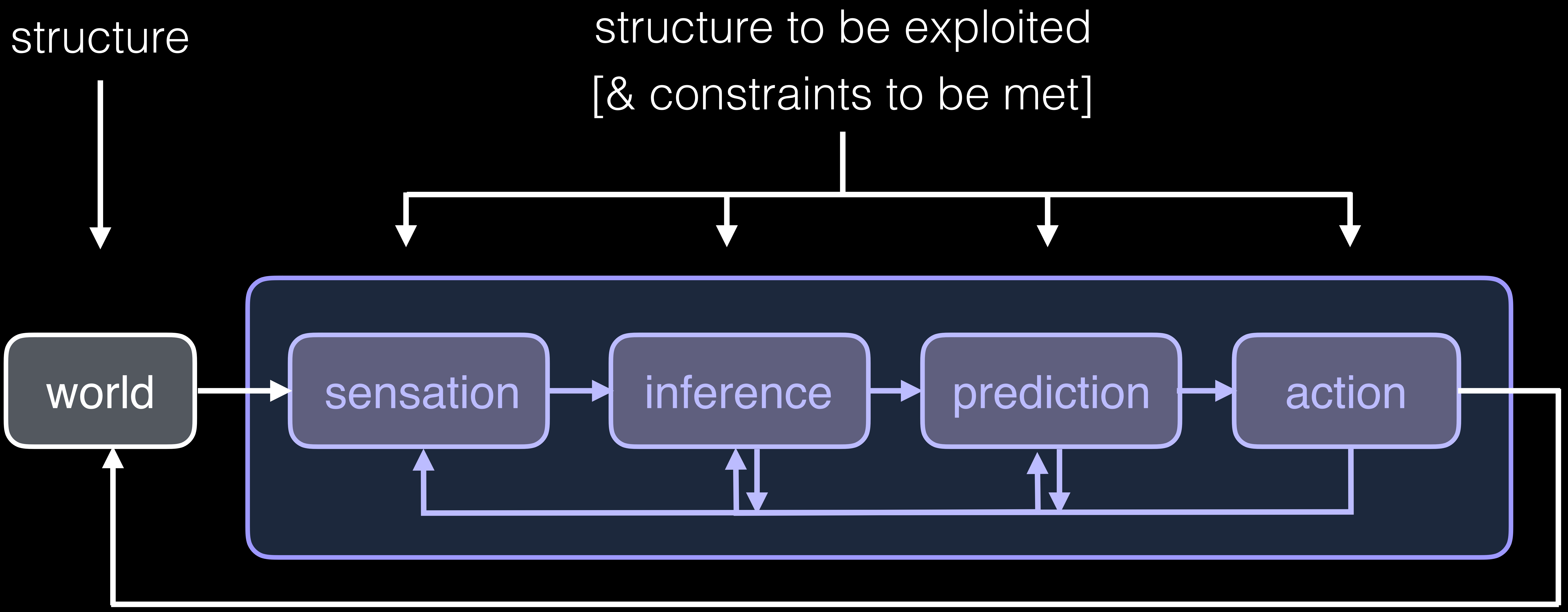
context in which function will be performed [natural visual world]

constraints on the system that performs this function [precision, accuracy, speed, energy, …] [bandwidth & noise constraints]

determine **best solution** for achieving particular function in particular context & subject to particular constraints [tune sensitivity to variability]


world
sensation


structure
structure to be exploited
[& constraints to be met]
world
sensation
action


structure
structure to be exploited
[& constraints to be met]
world
sensation
inference
prediction
action

PART 1

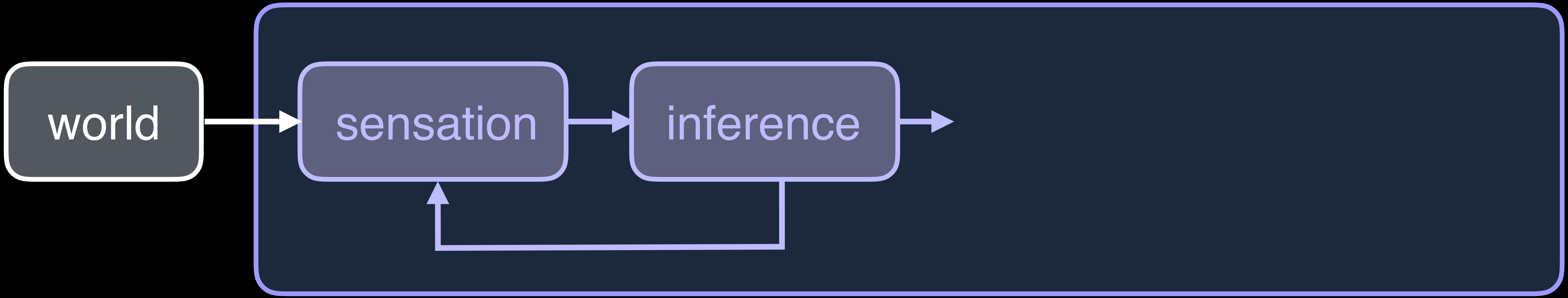
remove redundancy
combat noise

PART 1

remove redundancy
combat noise

PART 2

resolve ambiguity



PART 1

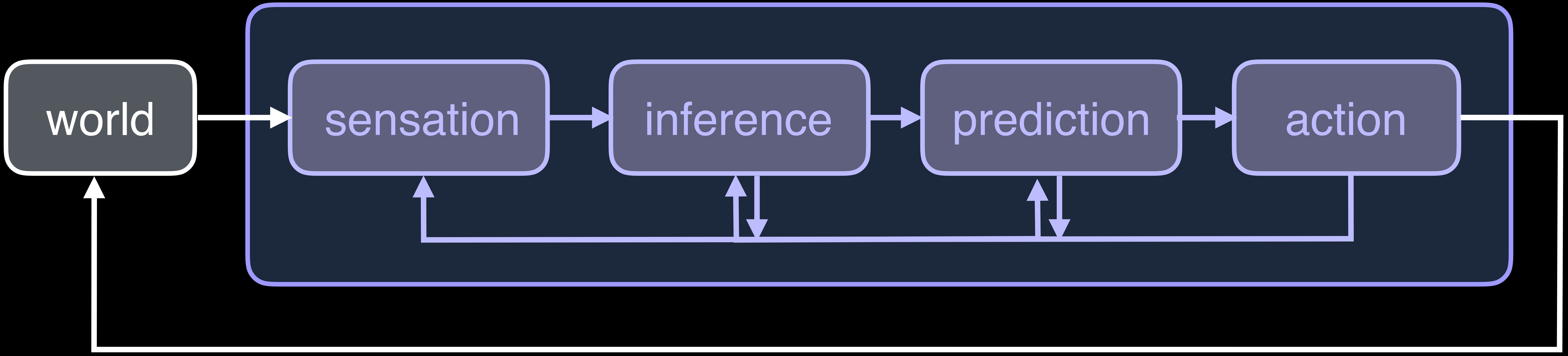
remove redundancy
combat noise

PART 2

resolve ambiguity

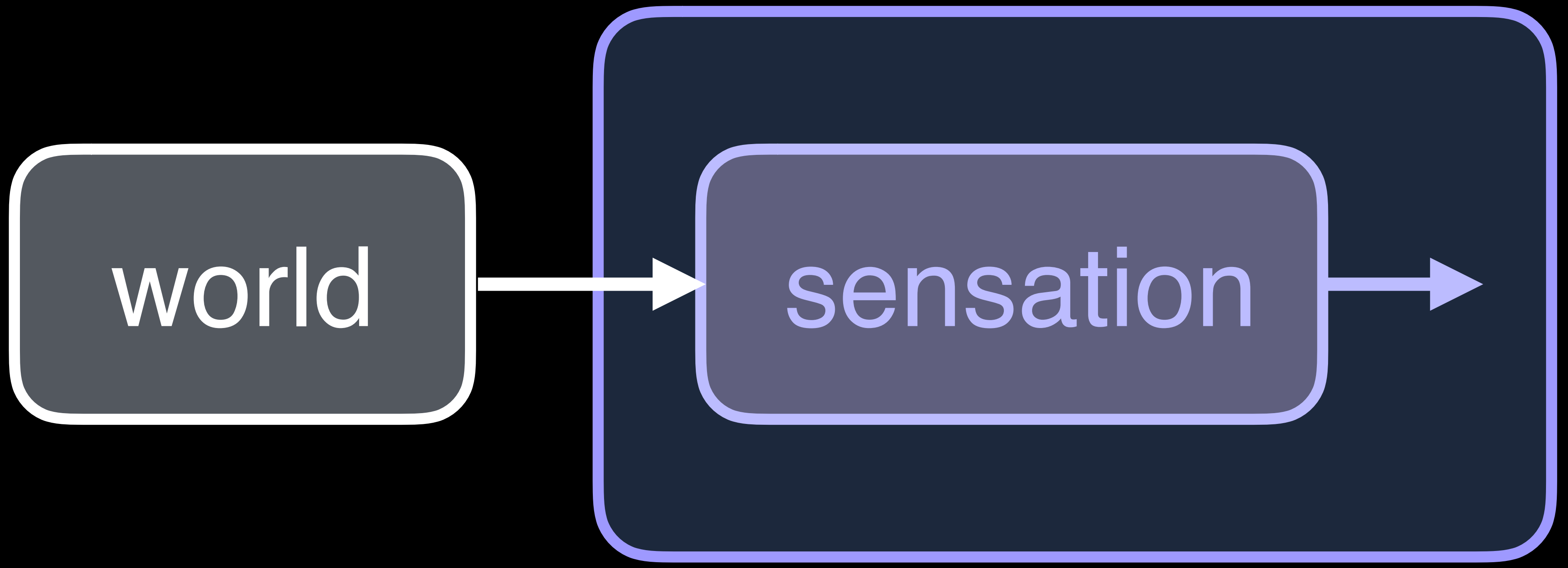
PART 3

make & use
predictions



PART 1

remove redundancy
combat noise



When we begin to consider perception as an information-handling process, it quickly becomes clear that much of the information received by any higher organism is *redundant*.

[this means that] if we know at a given moment the states of a limited number of receptors (i.e., whether they are firing or not firing), we can make better-than-chance inferences with respect to the prior and subsequent states of these receptors, and also with respect to the present, prior, and subsequent states of other receptors.

[this is] precisely equivalent to an assertion that the world as we know it is lawful.

It appears likely that a major function of the perceptual machinery is to strip away some of the redundancy of stimulation, to describe or encode incoming information in a form more economical than that in which it impinges on the receptors.

Fred Attneave 1954

Barlow's redundancy reduction hypothesis

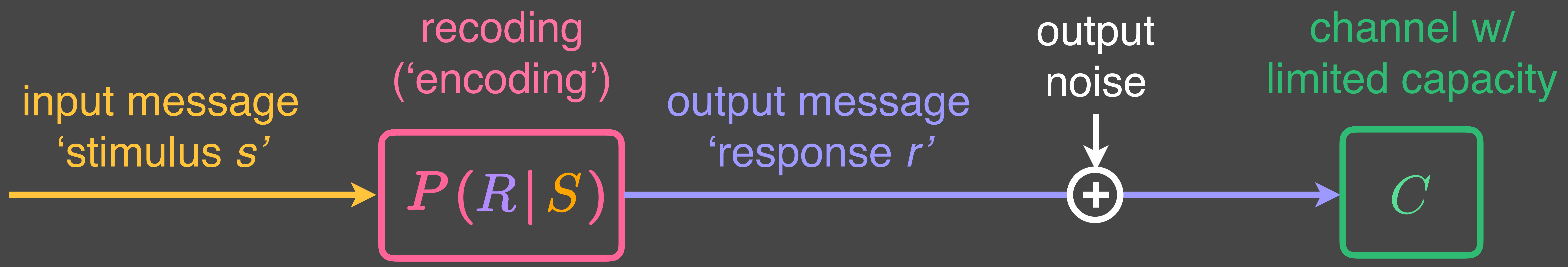
goal: maximize information $I(R;S) = H(R) - H(R|S)$

$H(R|S) \rightarrow 0$ low input noise

maximize response entropy $= H(R)$

OR

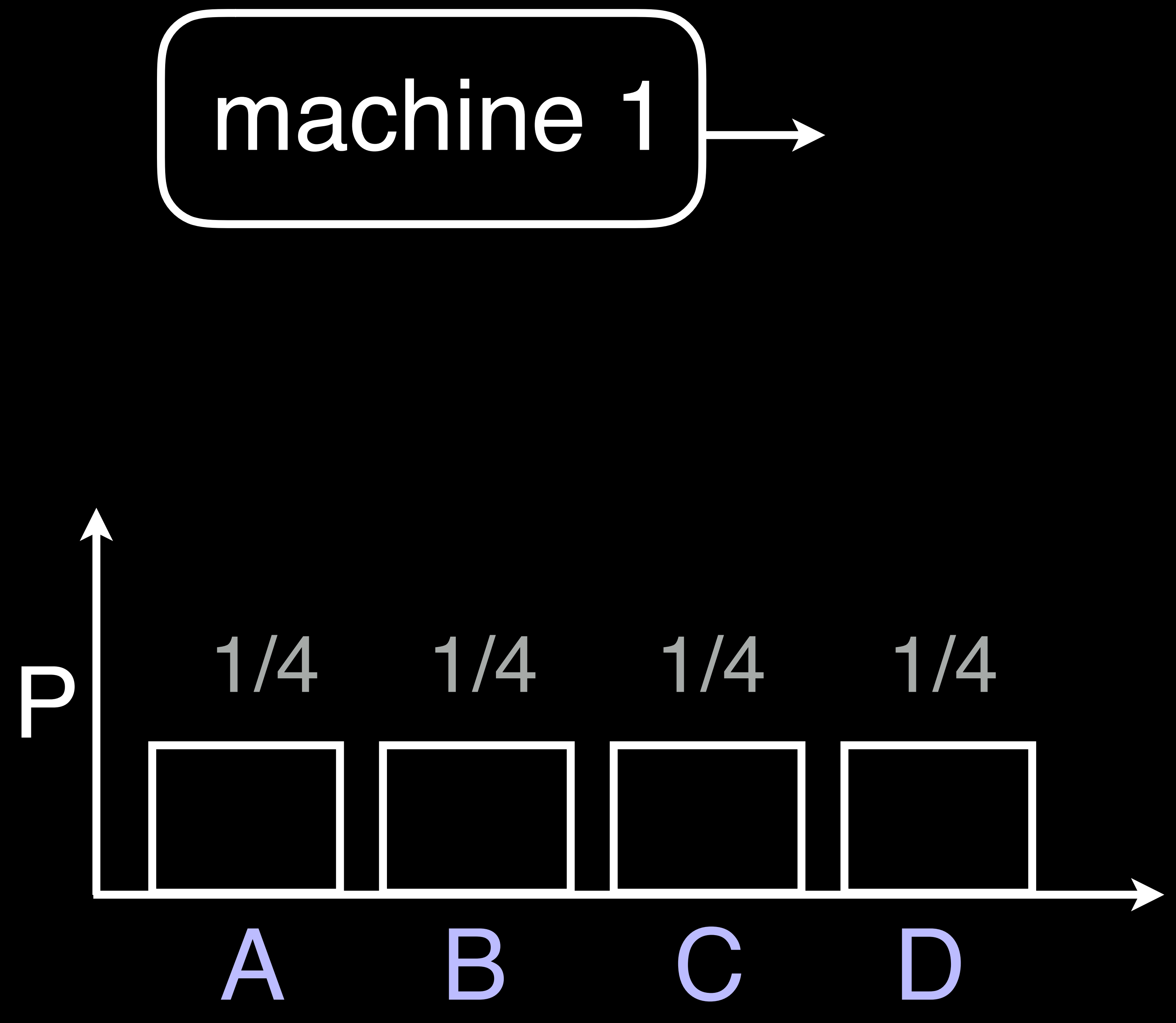
minimize redundancy $\mathcal{R} = 1 - H(R)/C$

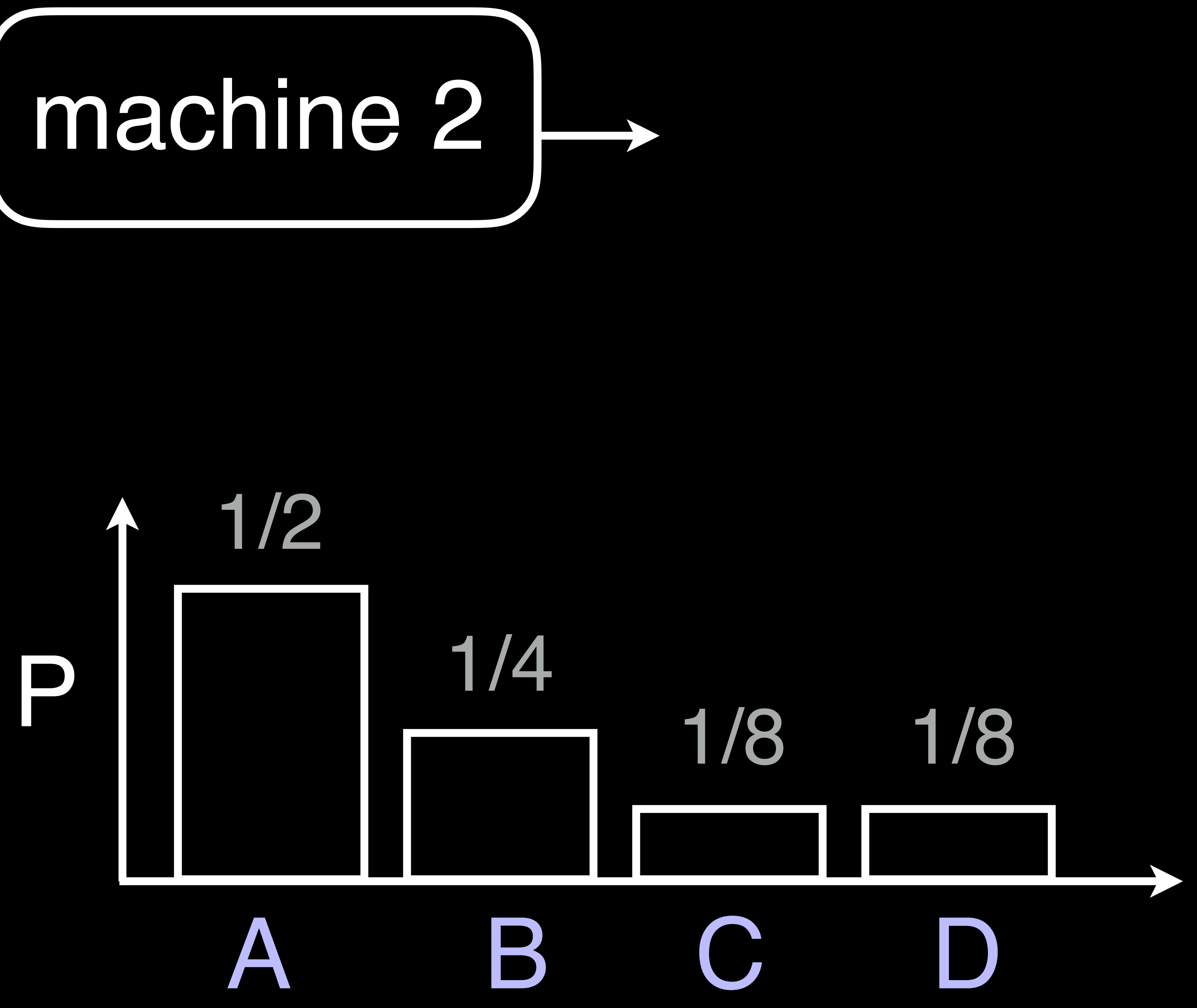


Horace Barlow 1961

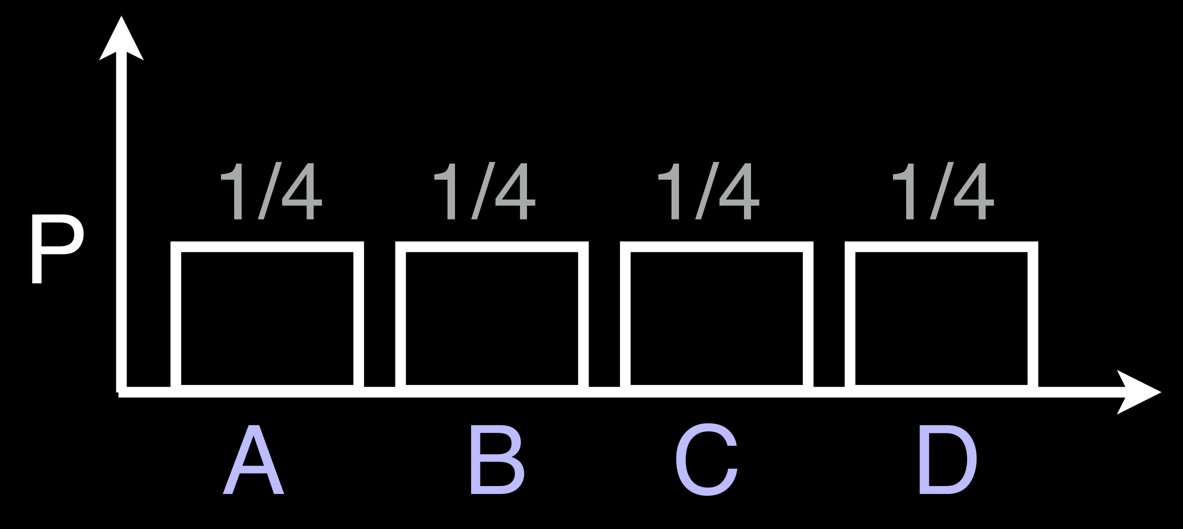
entropy $H(R)$

average # yes/no questions needed
to determine output with certainty

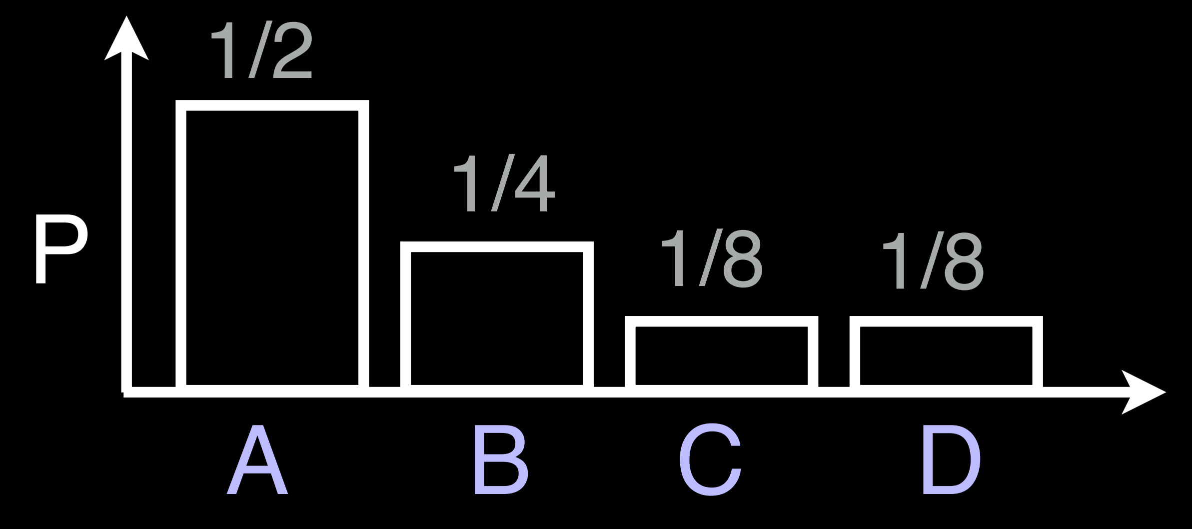




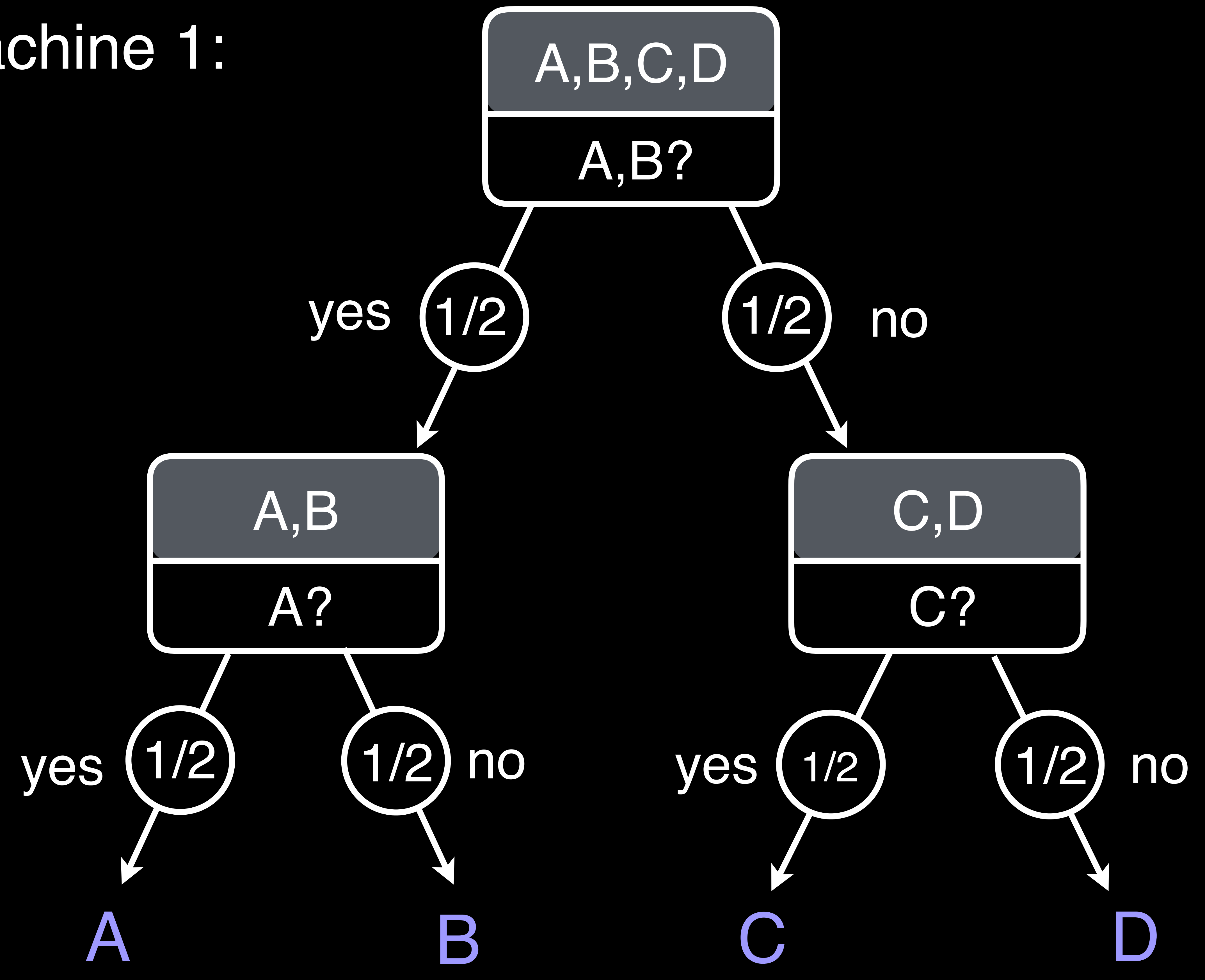
machine 1



machine 2



machine 1:



entropy

average # yes/no questions needed
to determine output with certainty

$$= \sum \text{\# questions per option} * \text{probability of option}$$

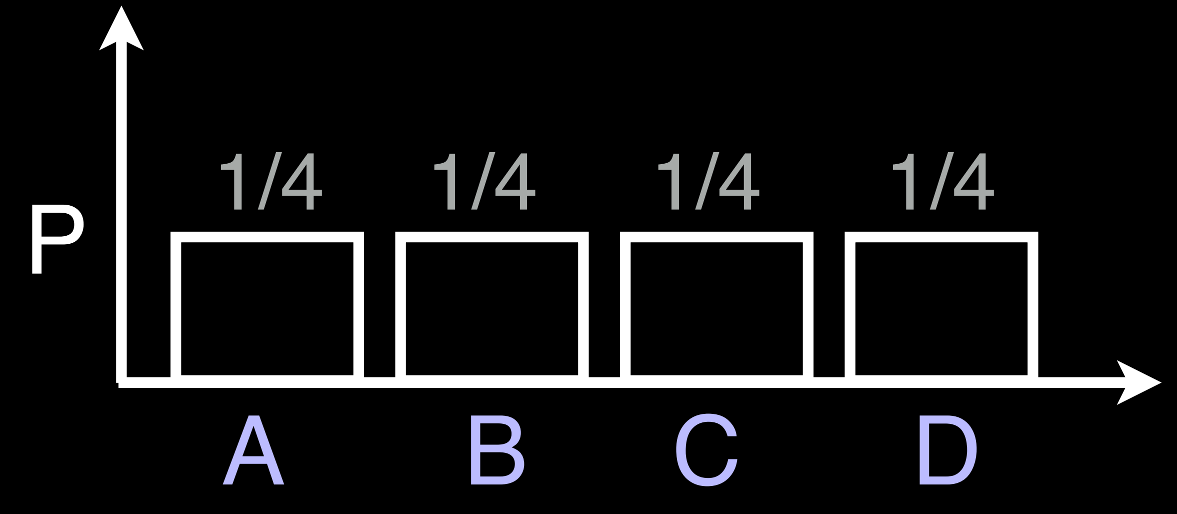
$$= \#Q_A * P_A + \#Q_B * P_B + \#Q_C * P_C + \#Q_D * P_D$$

$$\mathbf{H_1} = 2*(1/4) + 2*(1/4) + 2*(1/4) + 2*(1/4)$$

$$= 2 \text{ [bits]}$$

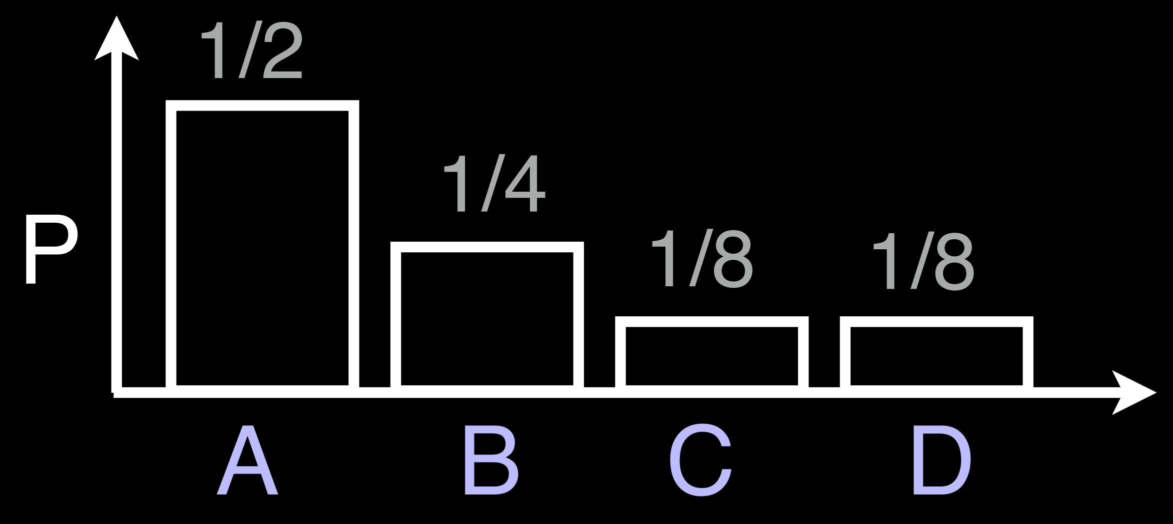
$$\mathbf{H_2} = ?$$

machine 1



H = 2 bits

machine 2



H = 1.75 bits

entropy

average # yes/no questions needed to determine output with certainty

$$= \sum \text{\# questions per option} * \text{probability of option}$$

$$= \#Q_A * P_A + \#Q_B * P_B + \#Q_C * P_C + \#Q_D * P_D$$

$$\mathbf{H_1} = 2*(1/4) + 2*(1/4) + 2*(1/4) + 2*(1/4)$$

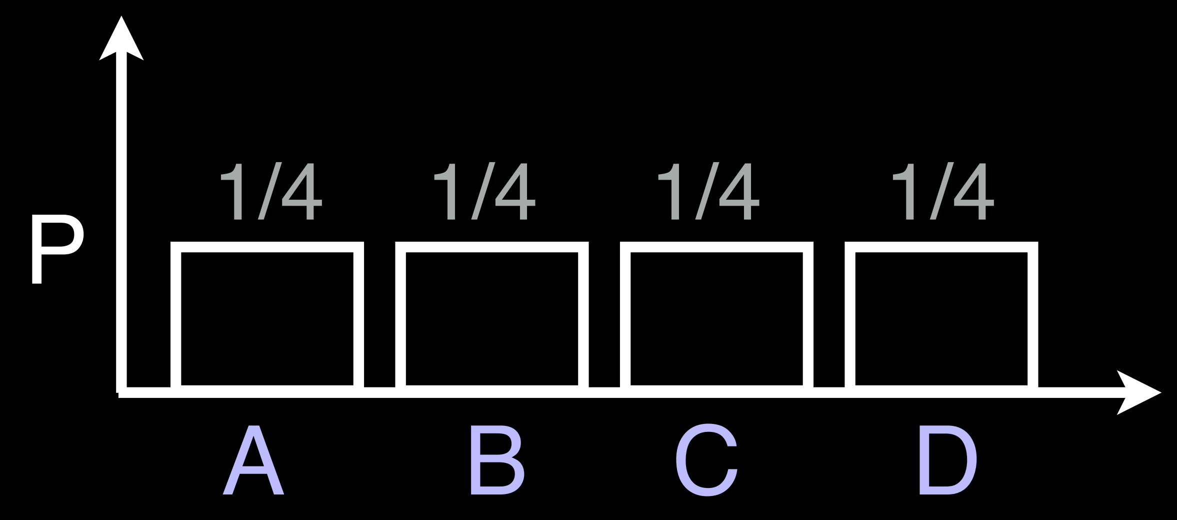
$$= 2 \text{ [bits]}$$

$$\mathbf{H_2} = 1.75 \text{ [bits]}$$

$$\#\,\mathrm{Q} = -\log_2 P$$

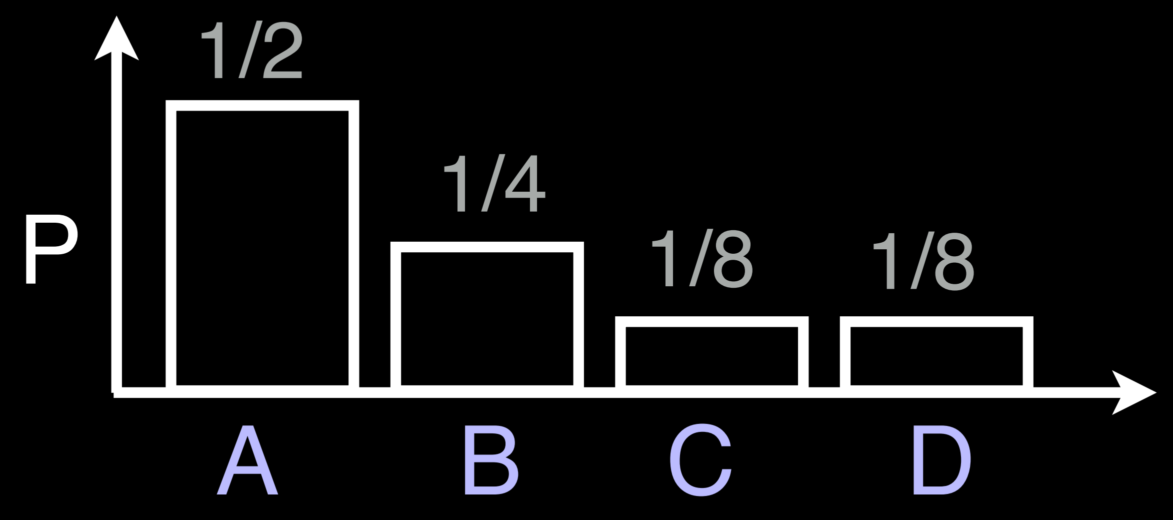
$$H = -\sum P \log_2 P$$

machine 1



H = 2 bits

machine 2



H = 1.75 bits

entropy

average # yes/no questions needed
to determine output with certainty

higher entropy

↓

more uncertainty

↓

greater reduction in uncertainty
by making a measurement

“classic” efficient coding hypothesis (low input noise)

goal: maximize information $I(R;S) = H(R) - H(R|S)$

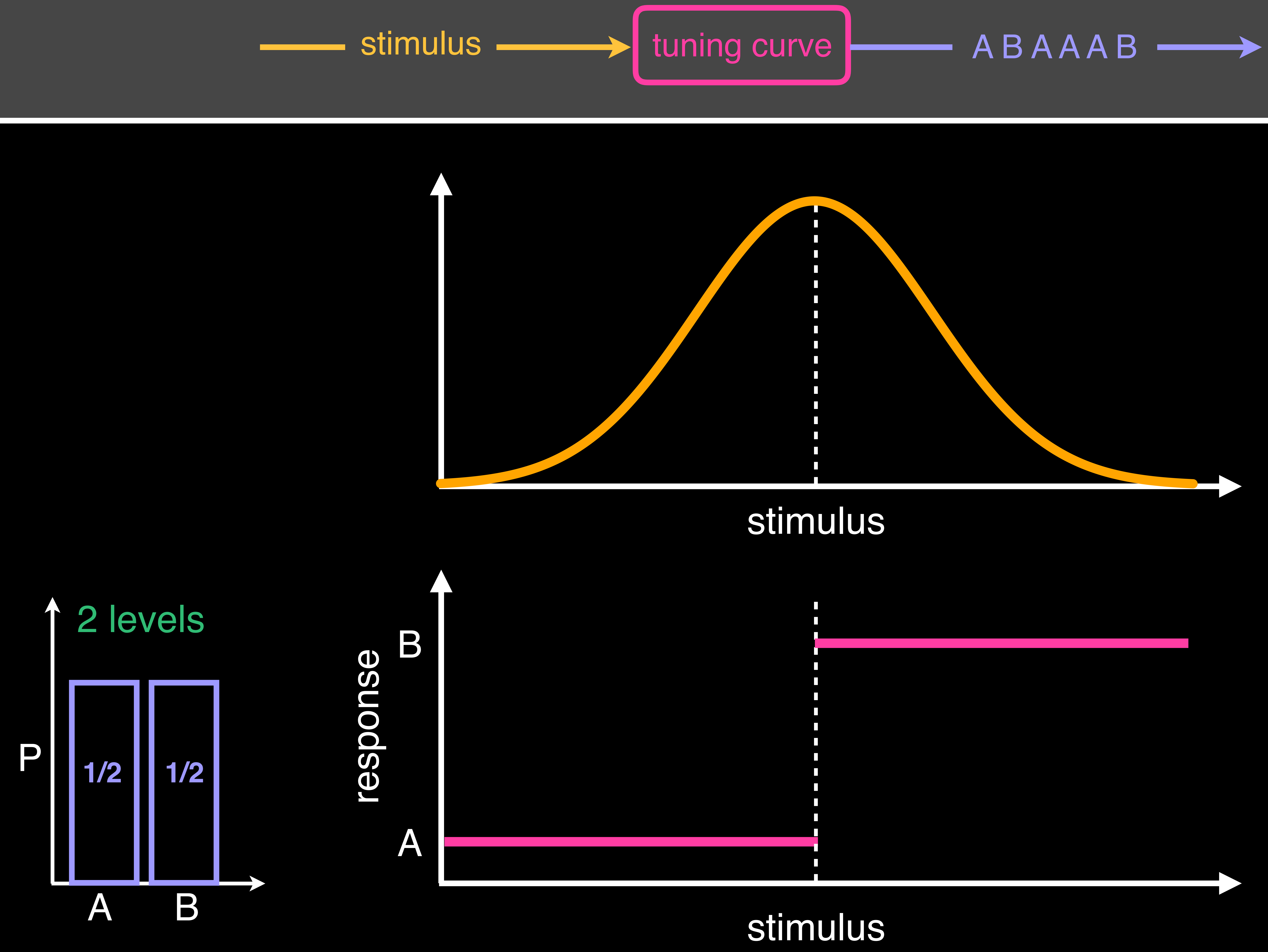
$H(R|S) \to 0$ low input noise

maximize response entropy $= H(R)$

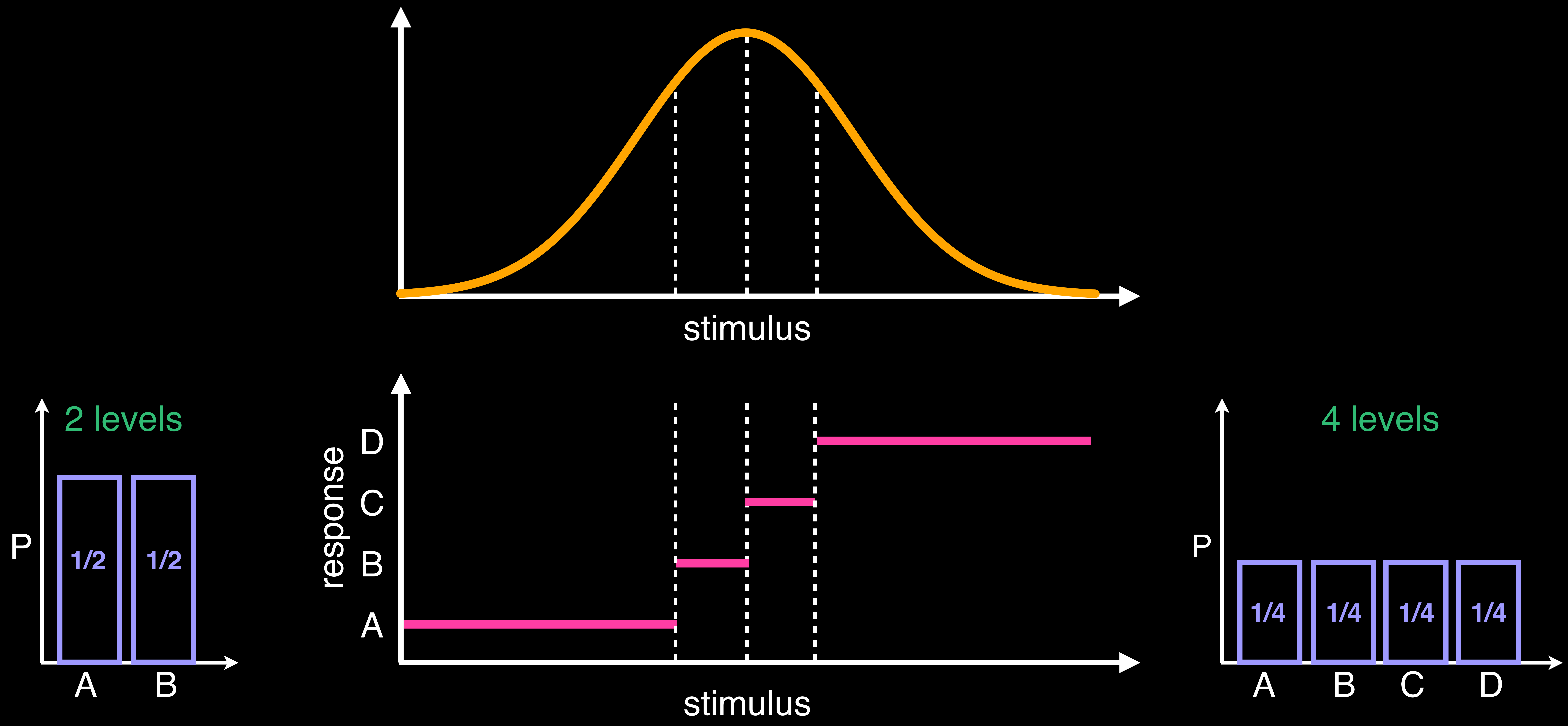
stimulus → tuning curve / machine → A B A A A B →

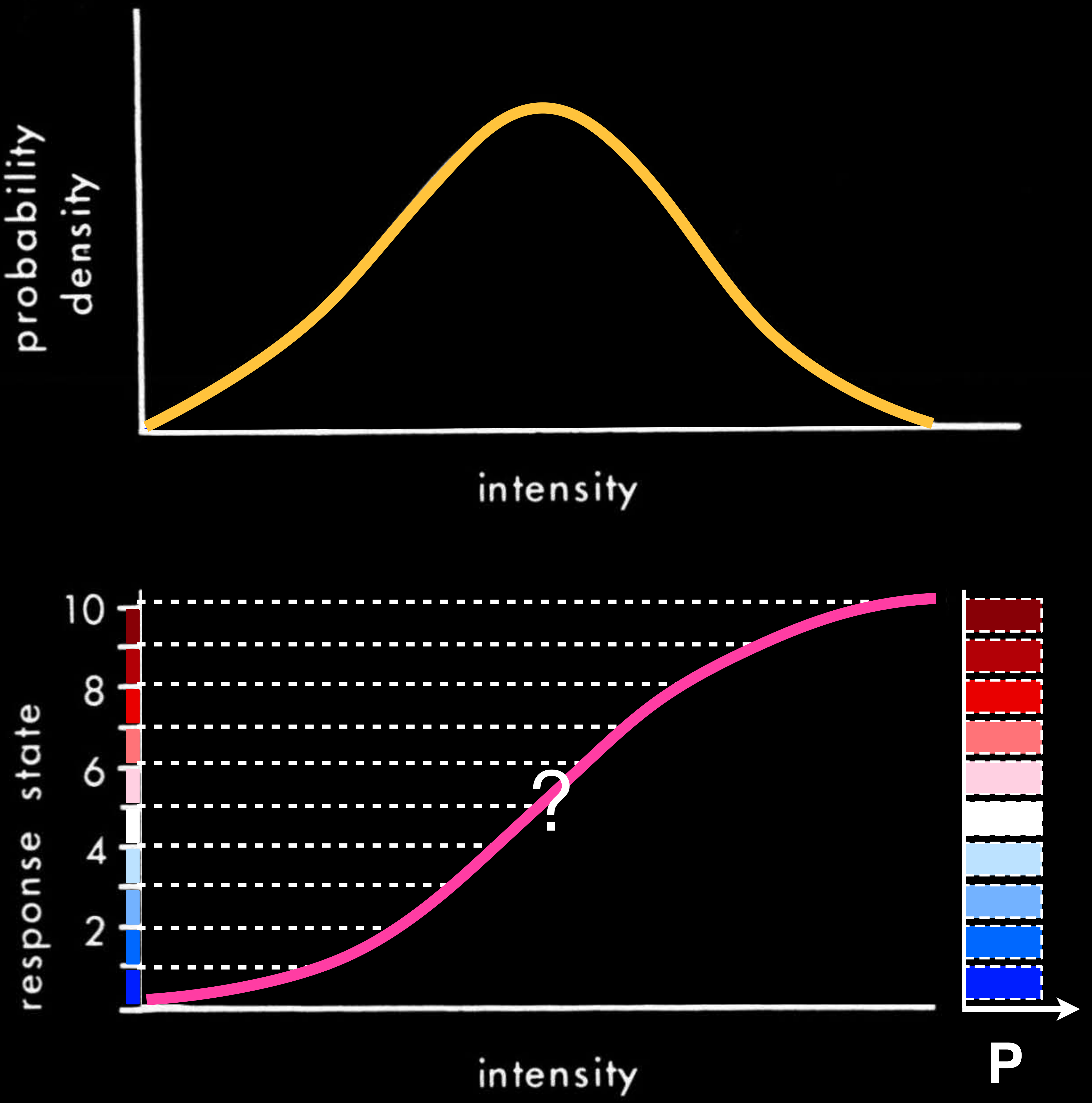
Horace Barlow 1961


stimulus
tuning curve
A B A A A B
stimulus
2 levels
P
1/2
1/2
A
B
response
B
A
?
stimulus

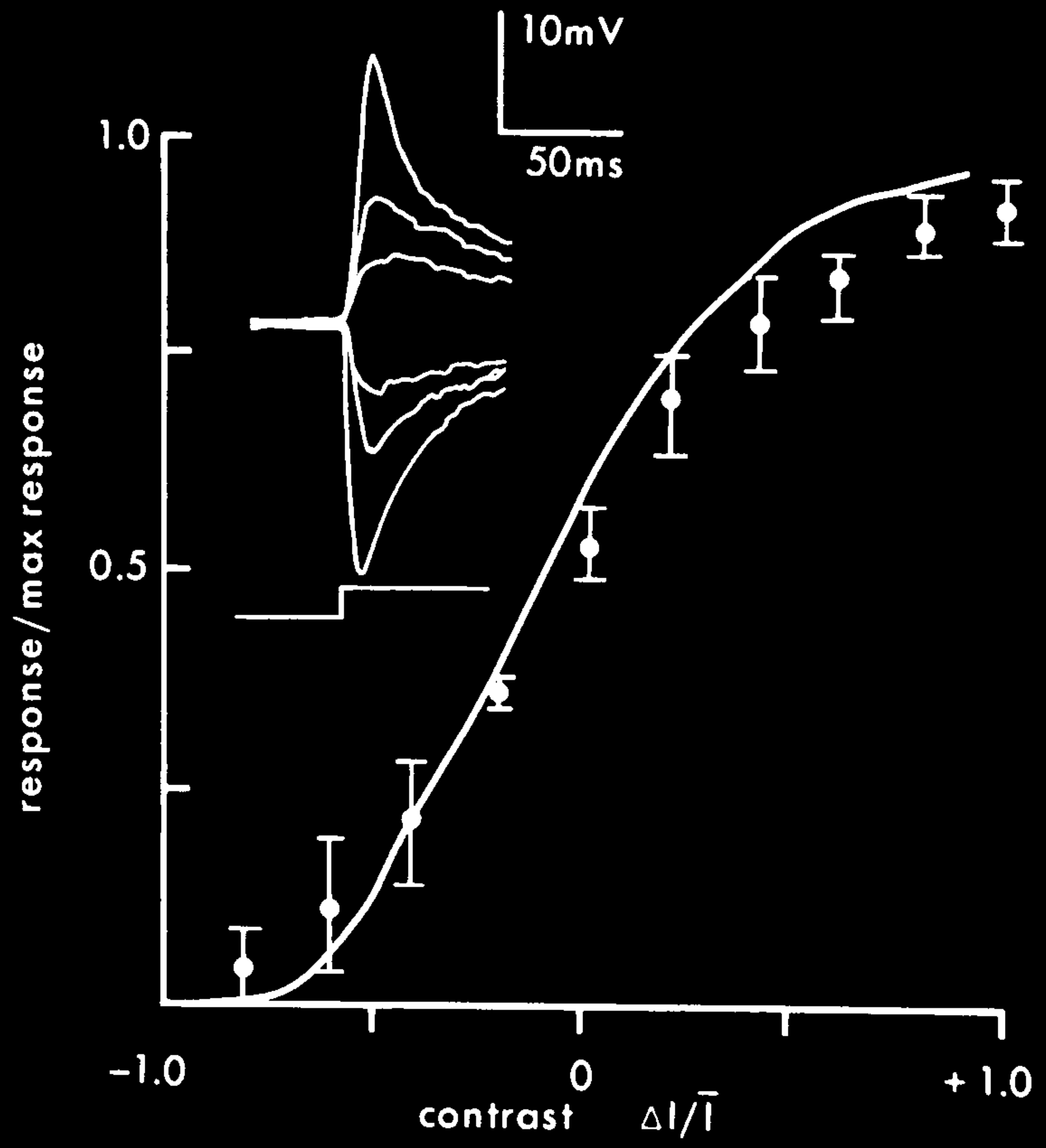

stimulus
tuning curve
A B A A A B
stimulus
2 levels
P
1/2
1/2
A
B
response
B
A
stimulus


stimulus
tuning curve
A B A A A B

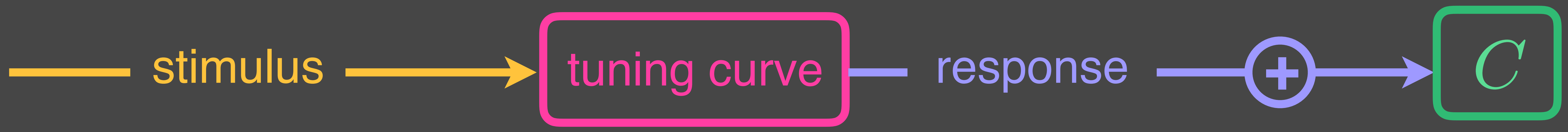

stimulus
2 levels
P
1/2
1/2
A
B
response
D
C
B
A
stimulus
4 levels
P
1/4
1/4
1/4
1/4
A
B
C
D



"histogram equalization"

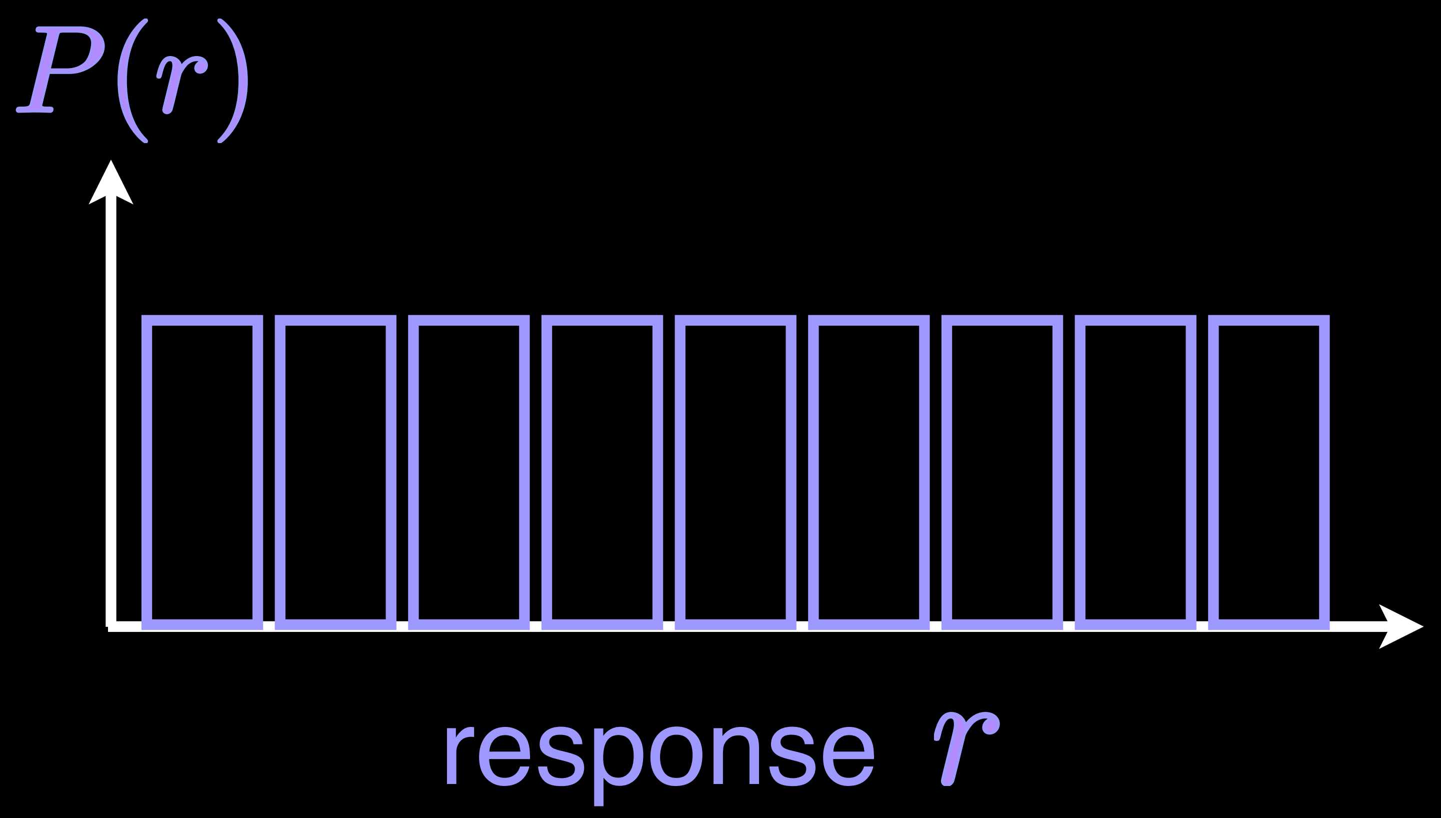


Simon Laughlin 1981



constraint on

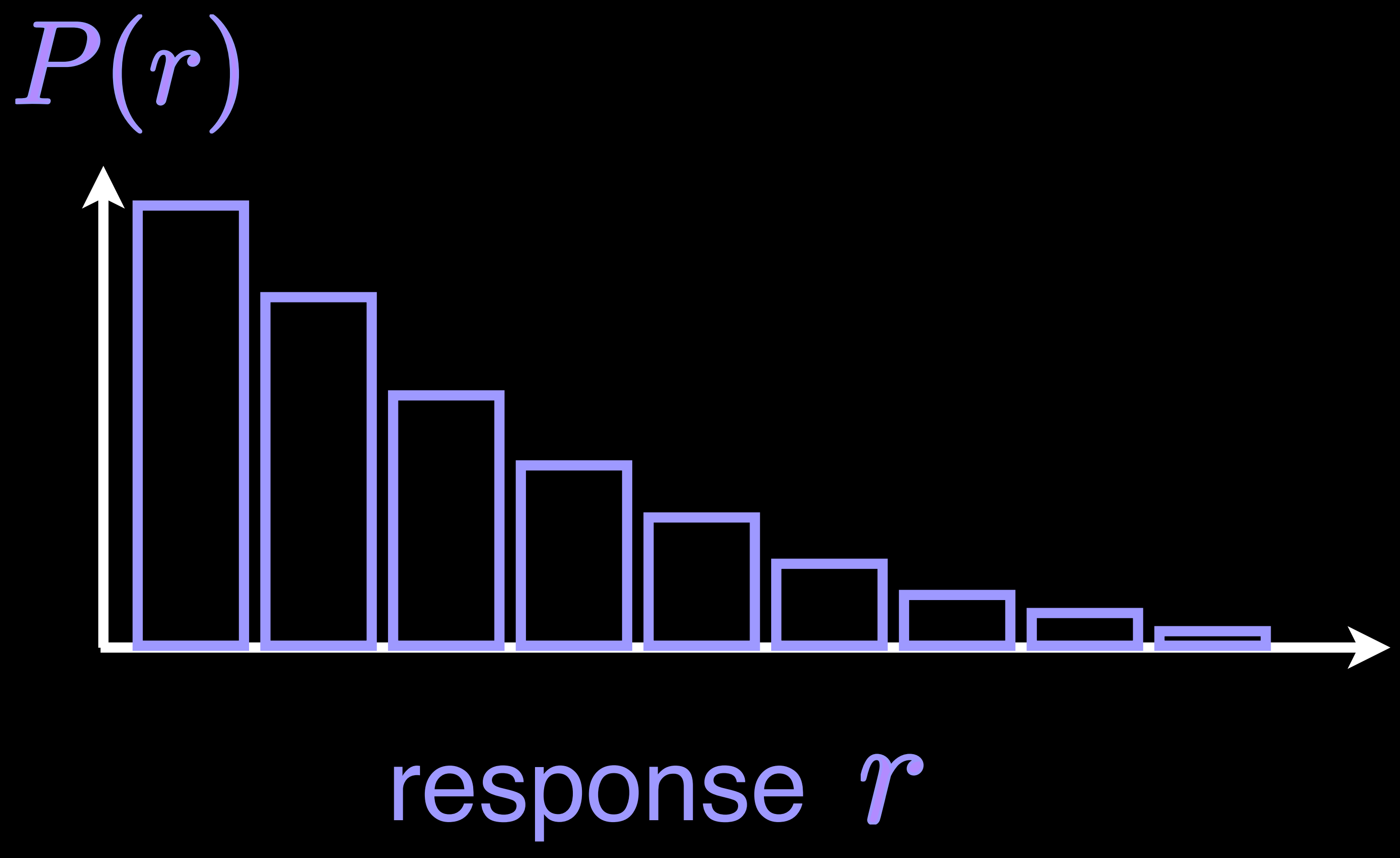
number of responses N



$$P(r) = 1/N$$

constraint on

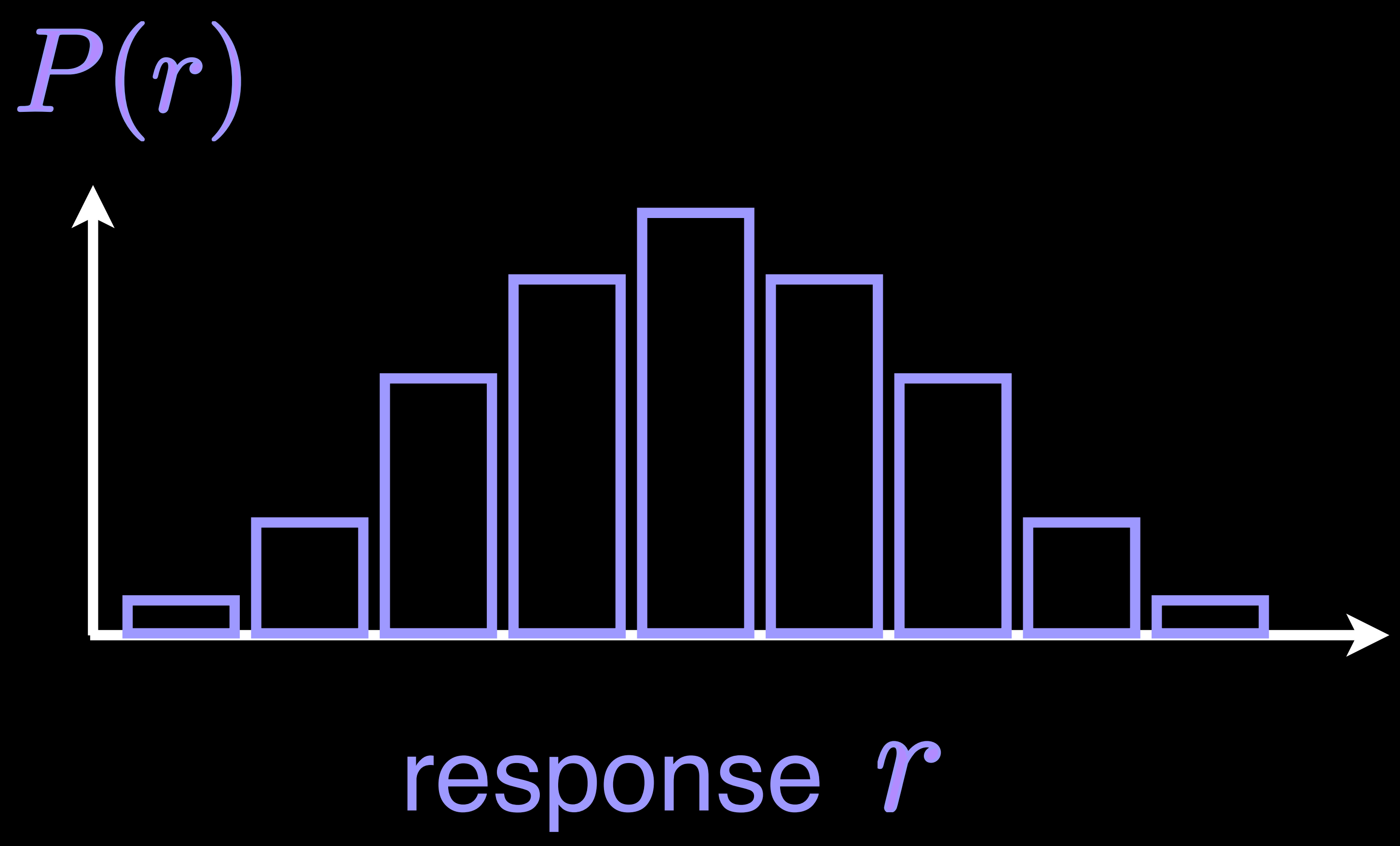
number of responses N
mean firing rate μ



$$= \frac{1}{\mu}\exp\left(-\frac{r}{\mu}\right)$$

constraint on

number of responses N
mean firing rate μ
variance in firing rate σ^2

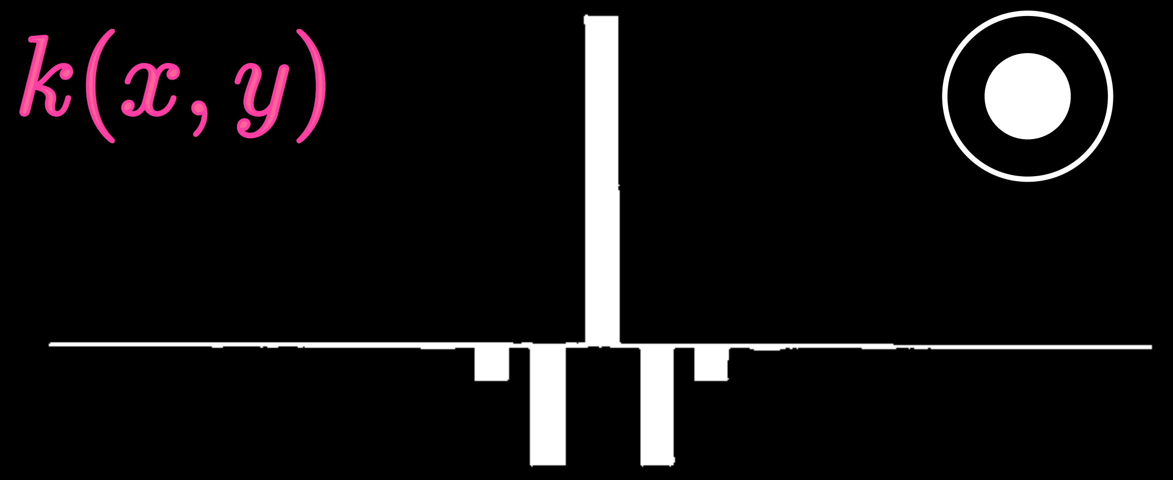


$$= \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(r-\mu)^2}{2\sigma^2}\right)$$

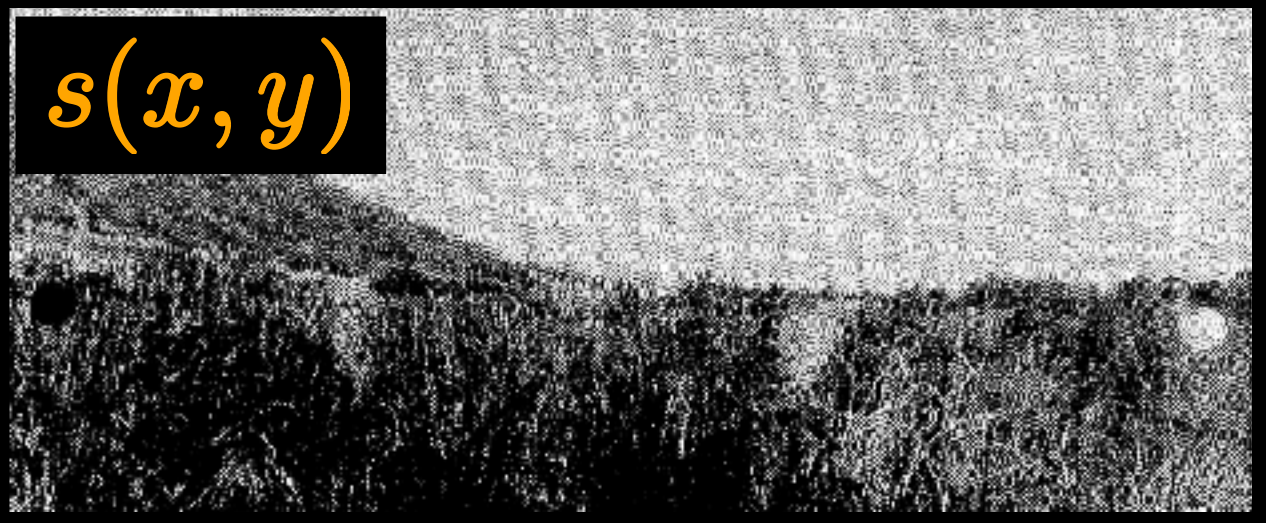
$$r(x, y) \propto k(x, y) \circledast s(x, y)$$

response | linear filter | stimulus

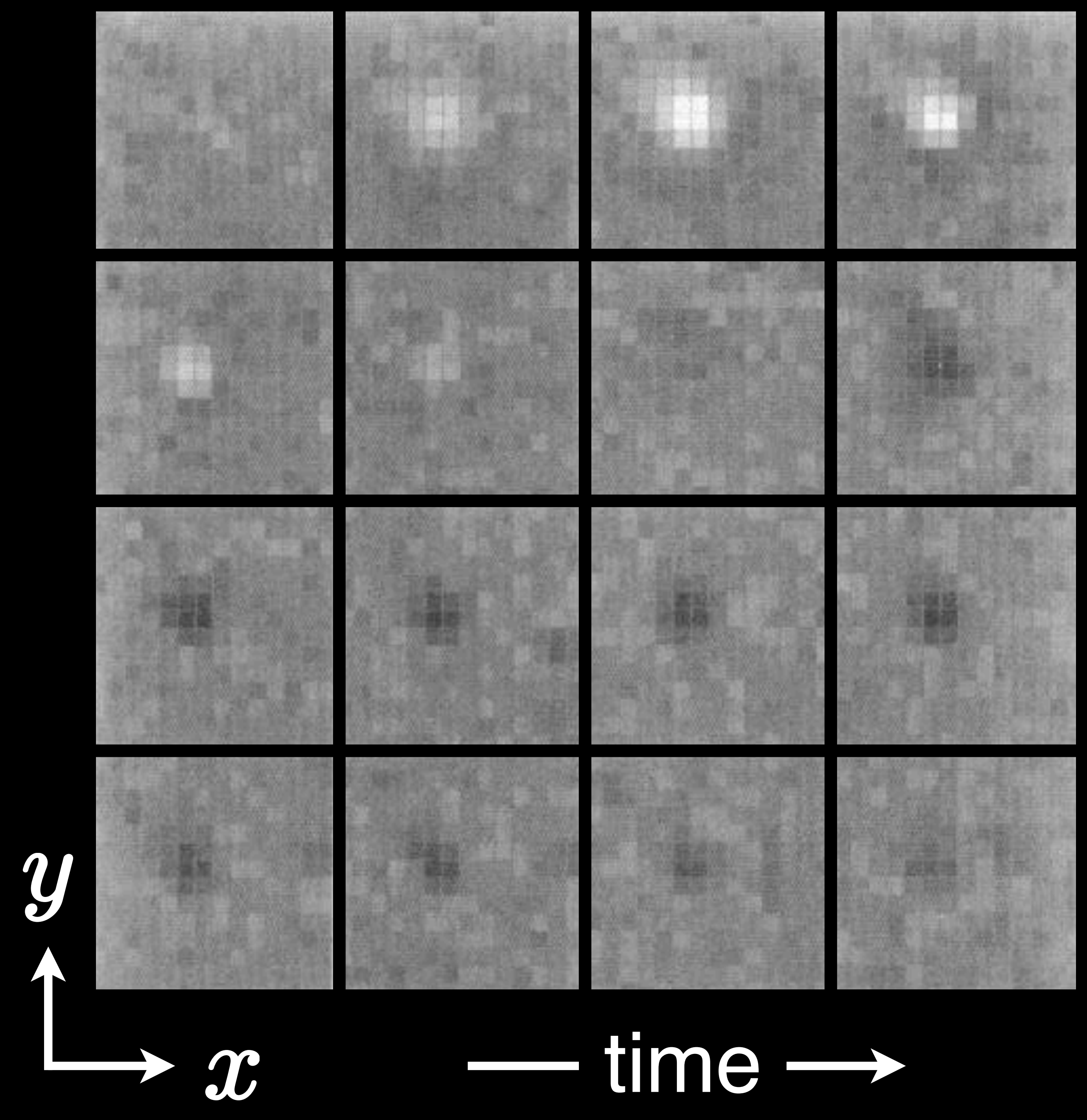
$$r(x, y) \propto k(x, y) \circledast s(x, y)$$

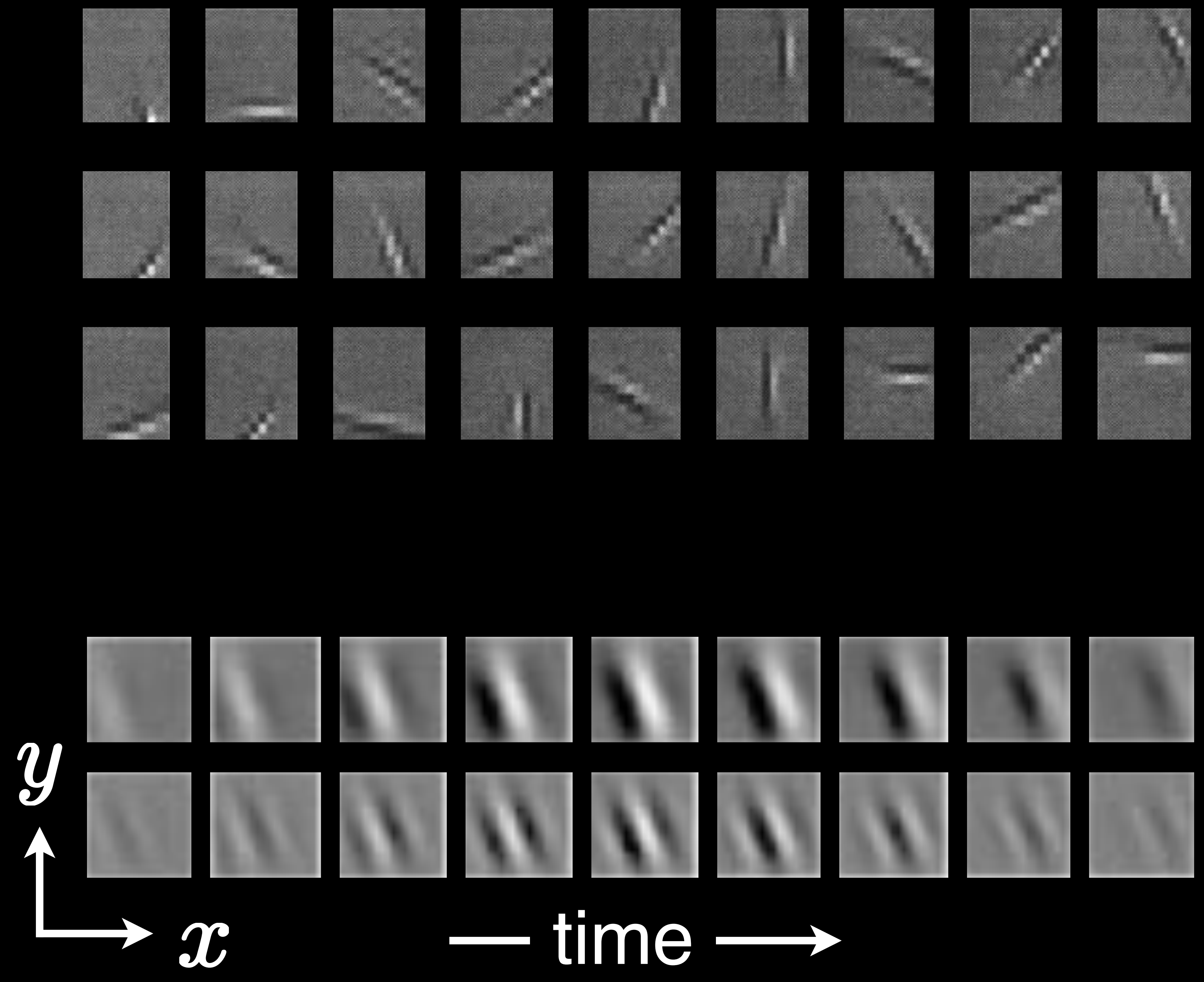


spatial location



spatial location





Srinivasan, Laughlin, & Dubs 1982
Atick & Redlich 1990

Dan, Atick, Reid 1996

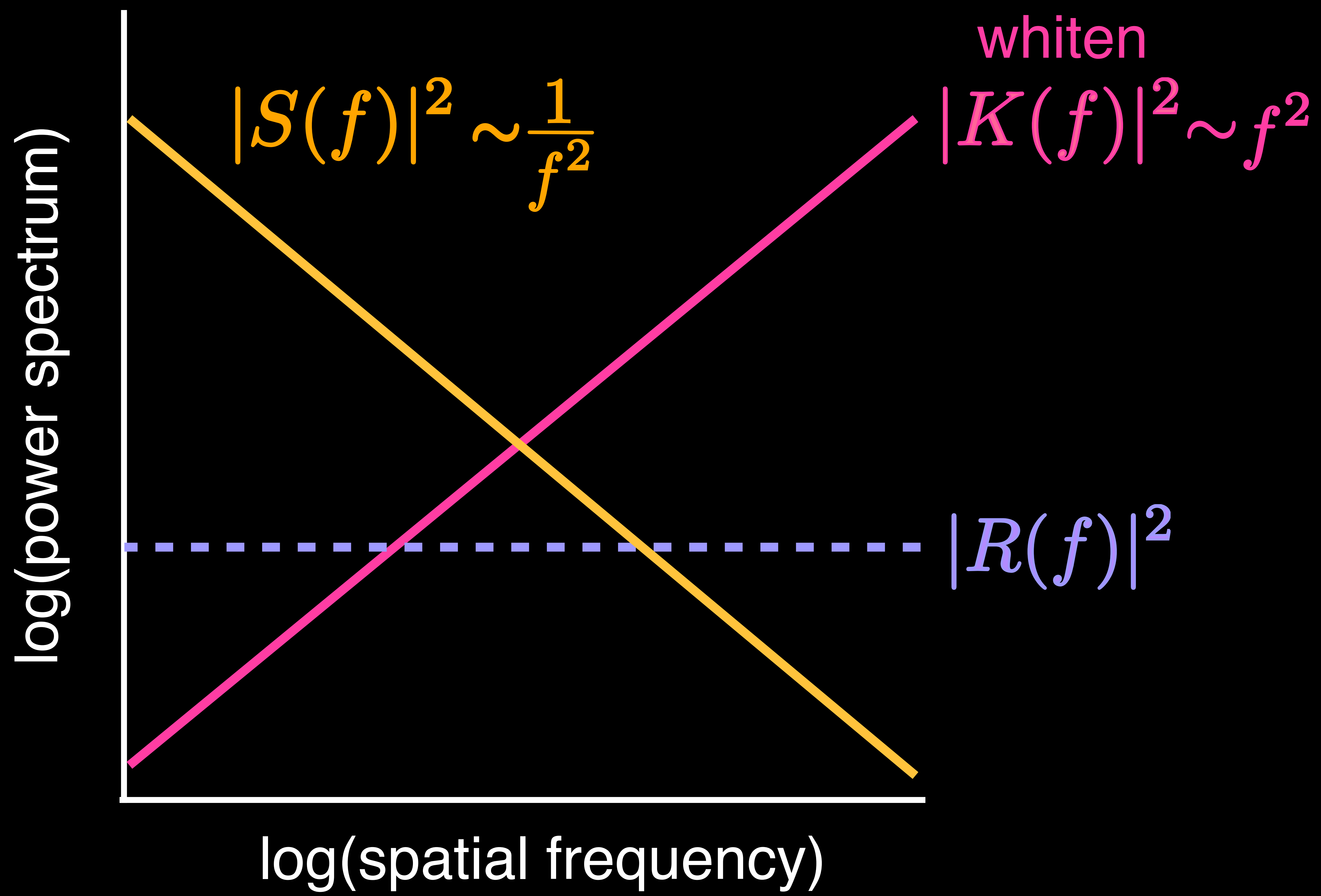
Bell & Sejnowski, 1995, 1997
Olshausen & Field, 1996
van Hateren & Ruderman, 1998

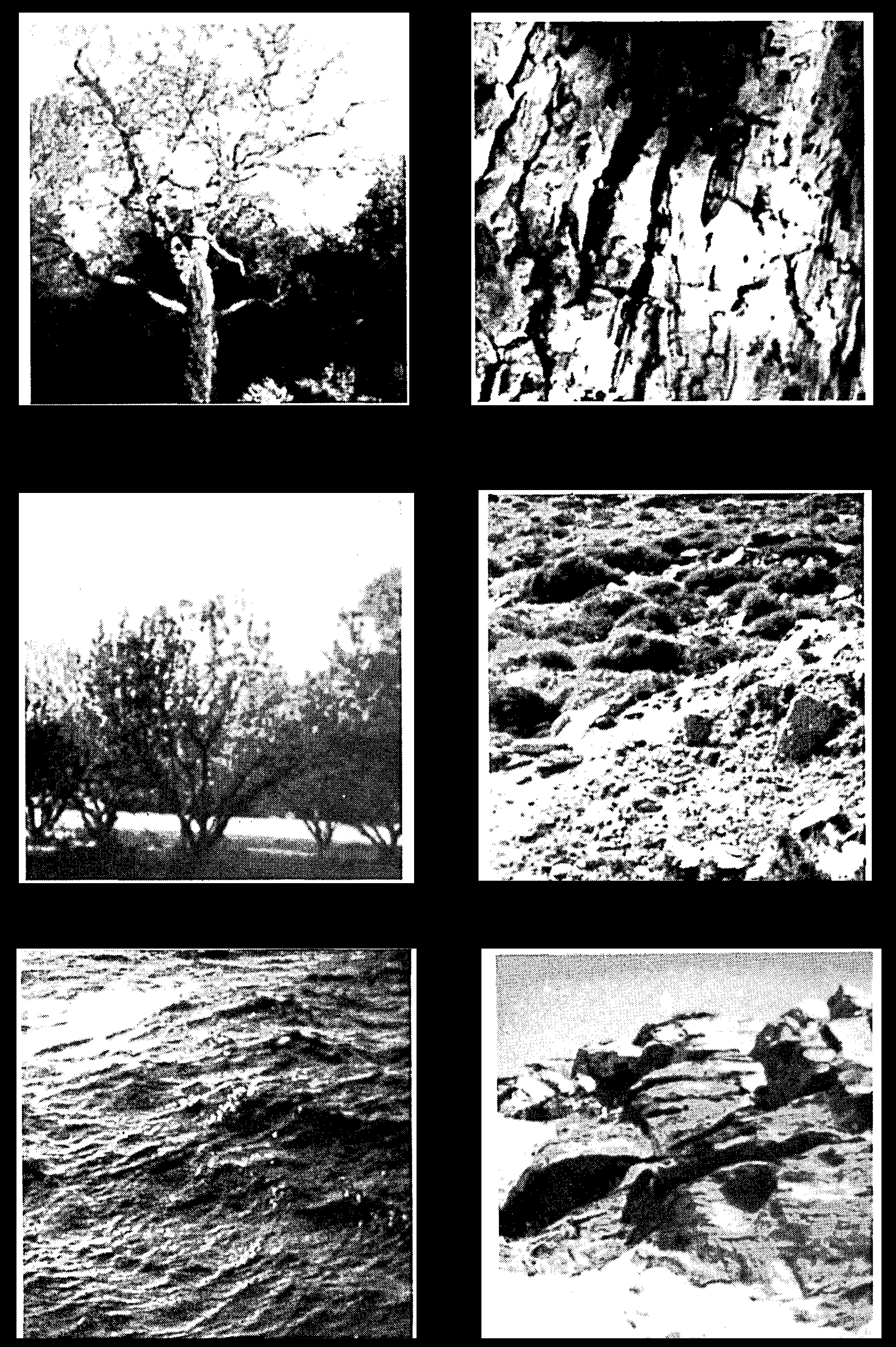
$$r(x, y) \propto k(x, y) \circledast s(x, y)$$

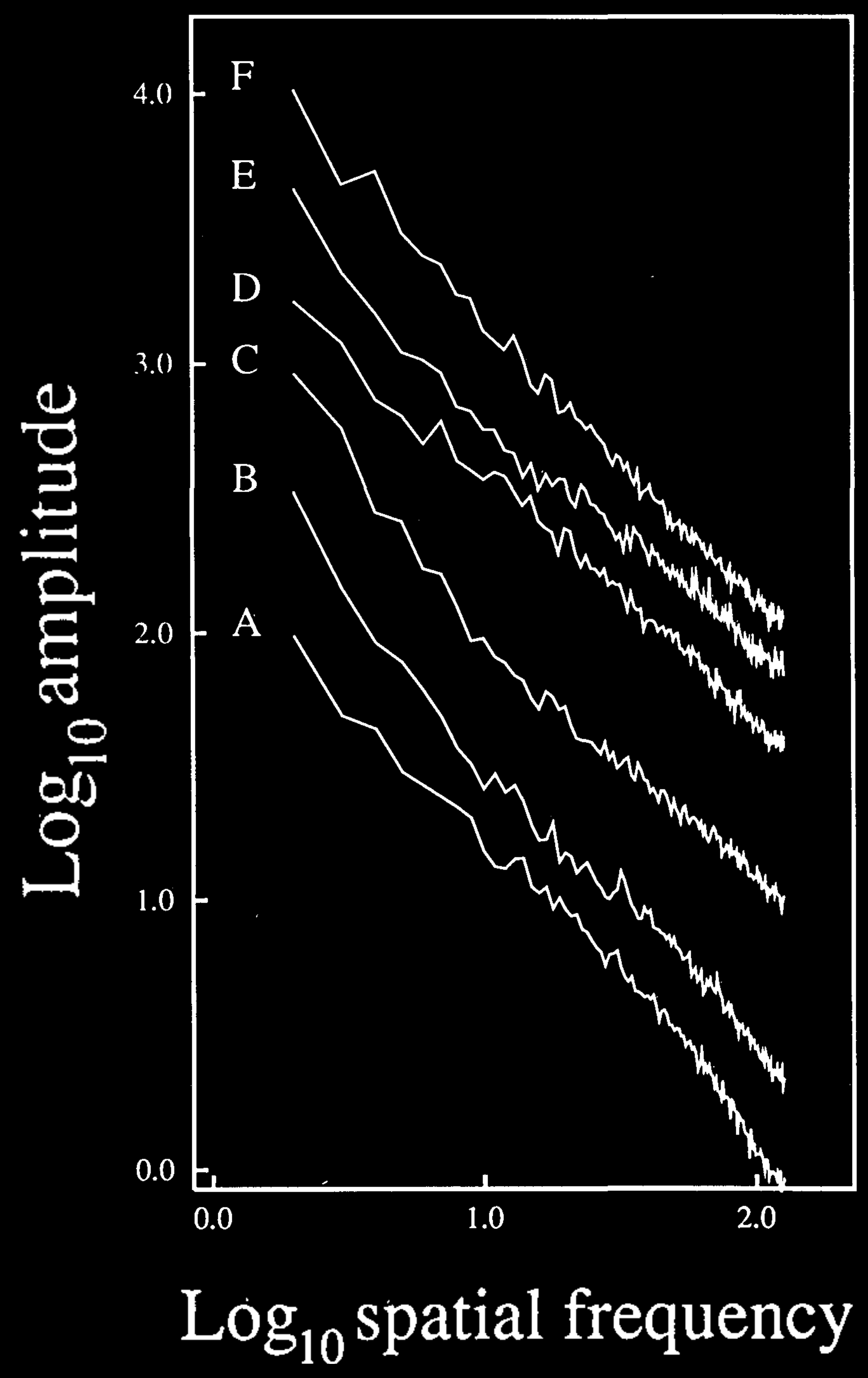
Fourier transform

$$|R(f)|^2 \propto |K(f)|^2 \cdot |S(f)|^2$$

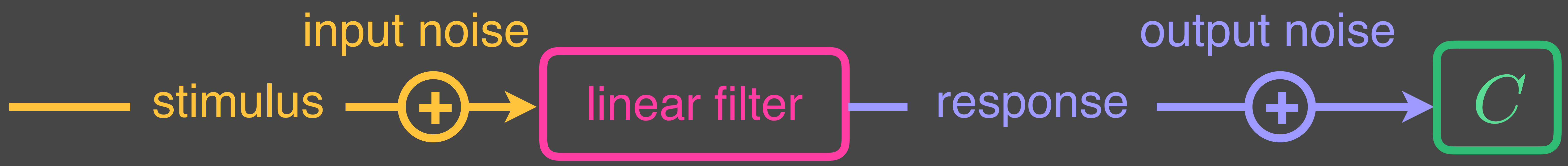
$$|R(f)|^2 \propto |K(f)|^2 \cdot |S(f)|^2$$



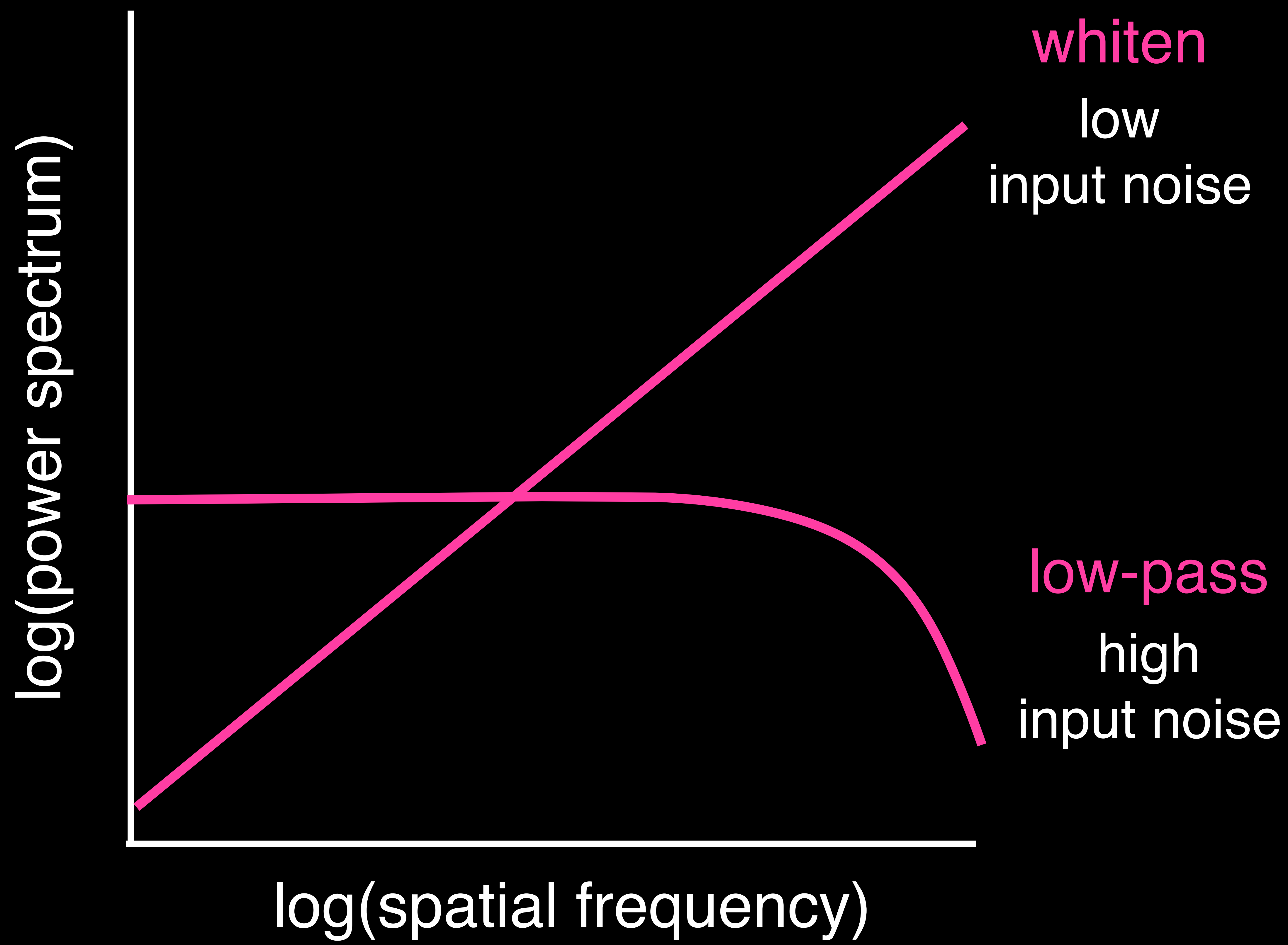


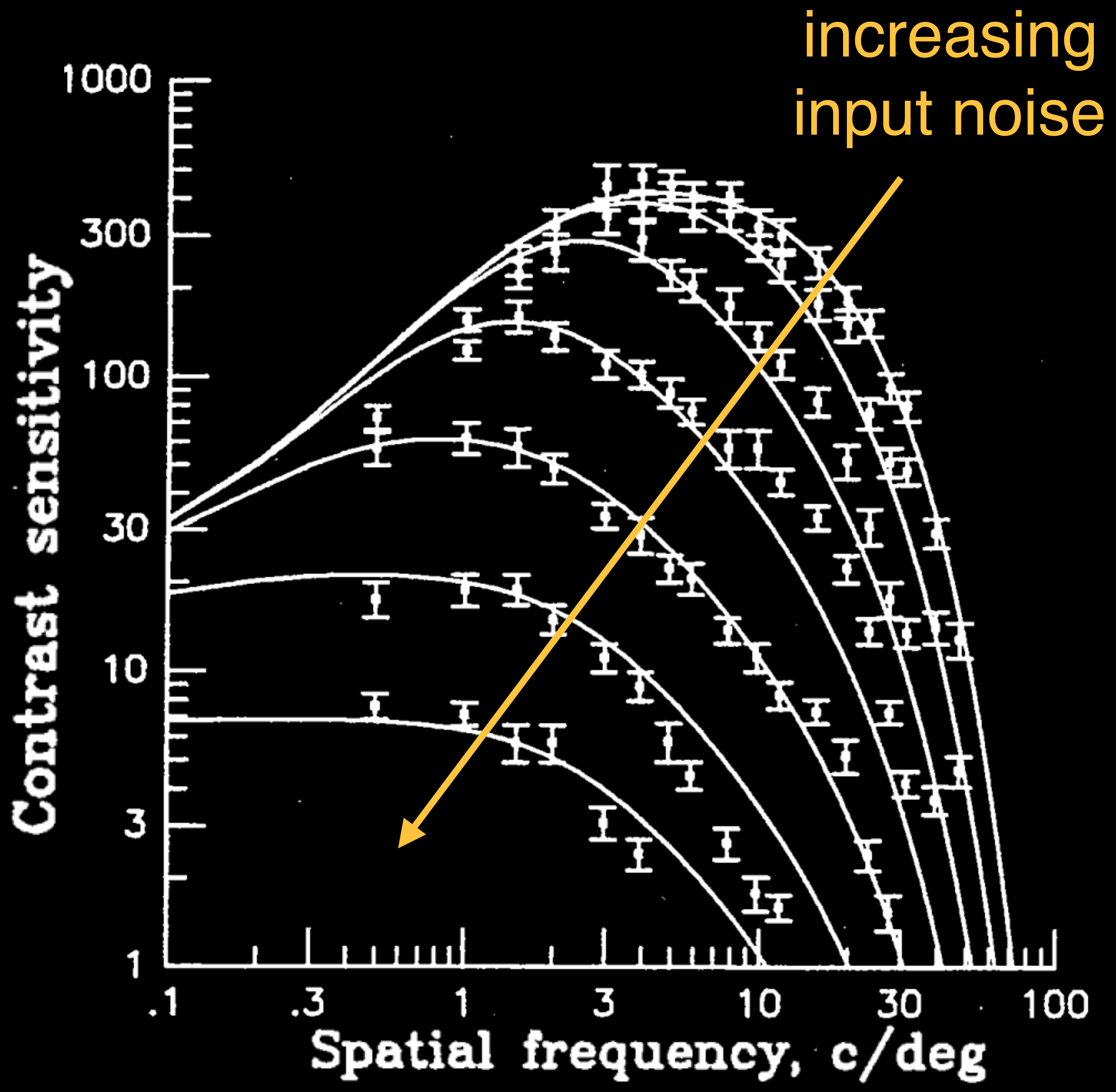


Field 1987

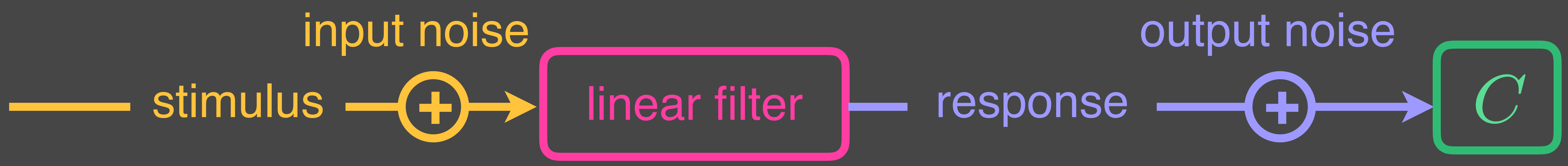


$$|R(f)|^2 \propto |K(f)|^2 \cdot |S(f)|^2$$

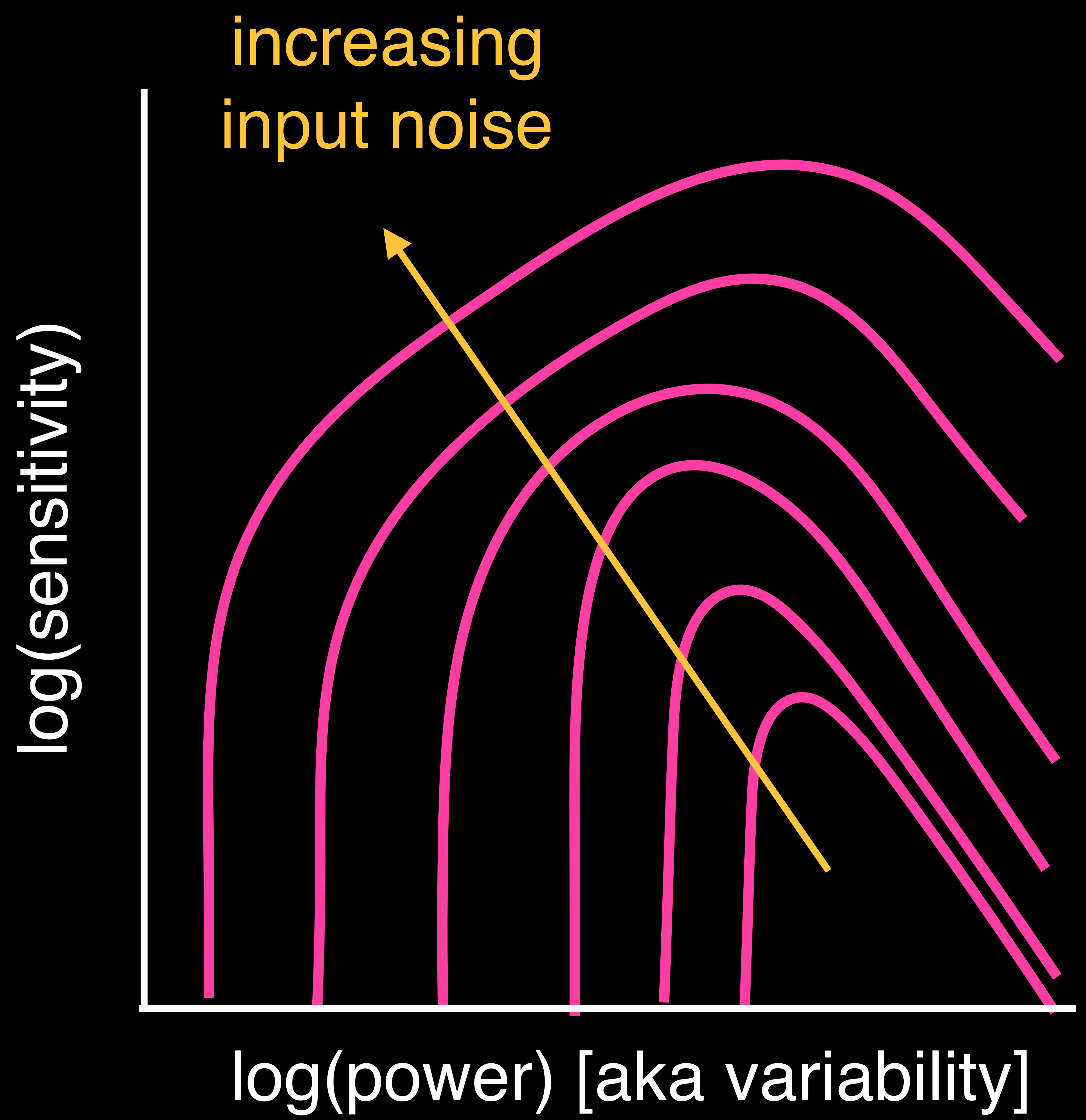


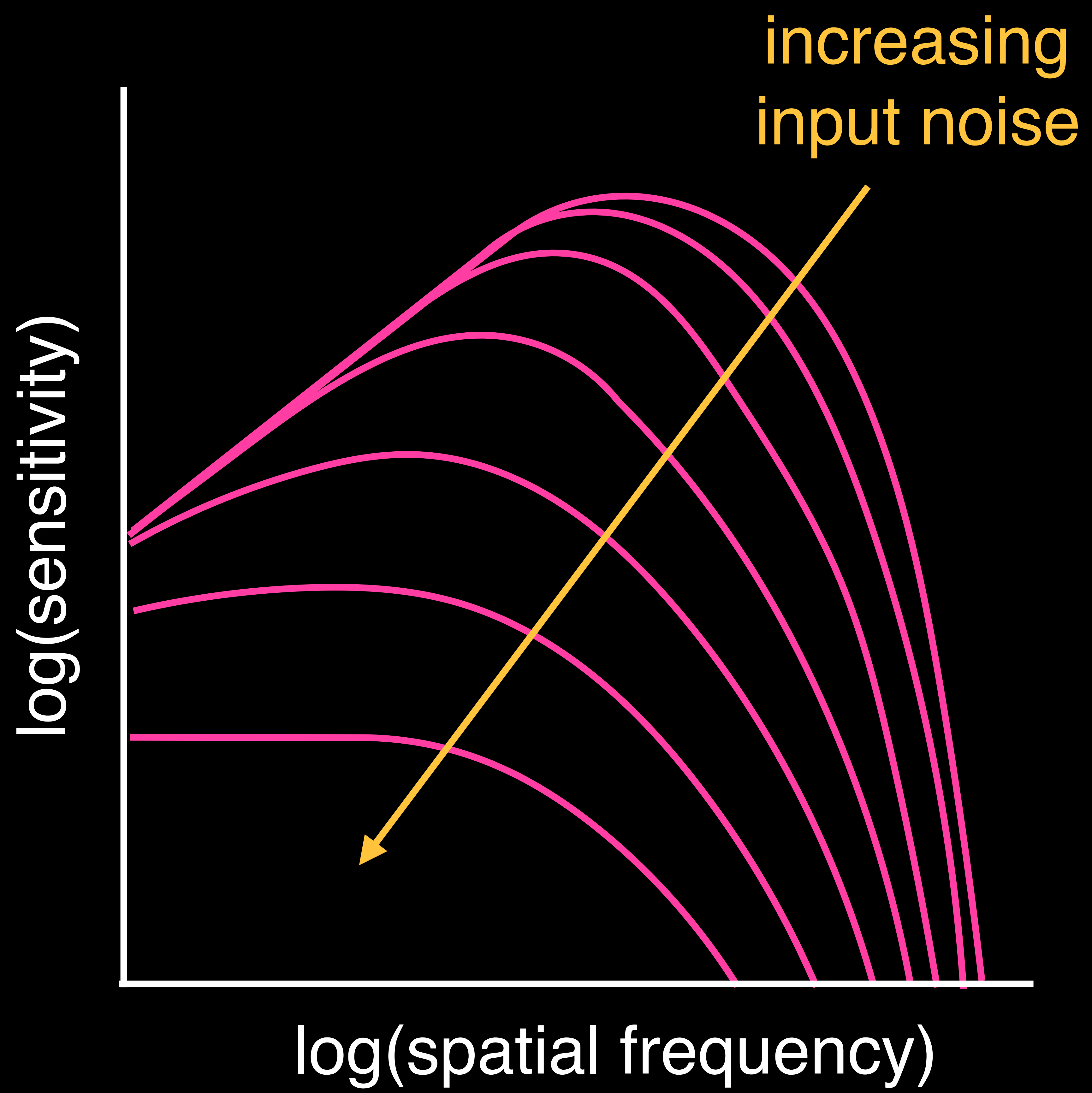


Atick & Redlich 1992



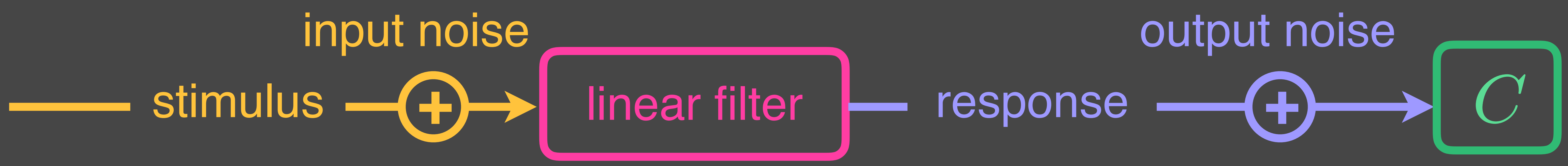
$$|R(f)|^2 \propto |K(f)|^2 \cdot |S(f)|^2$$



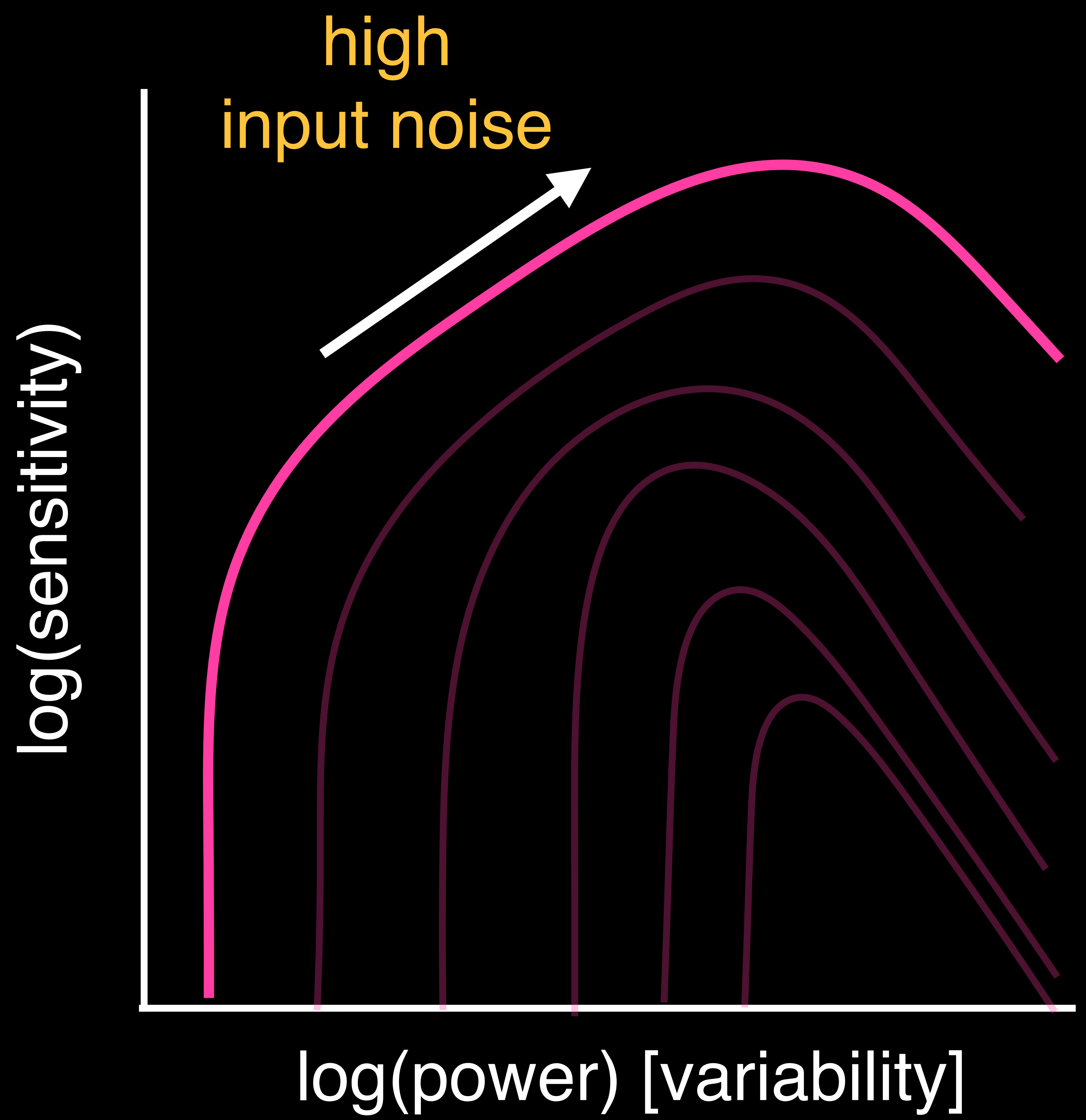


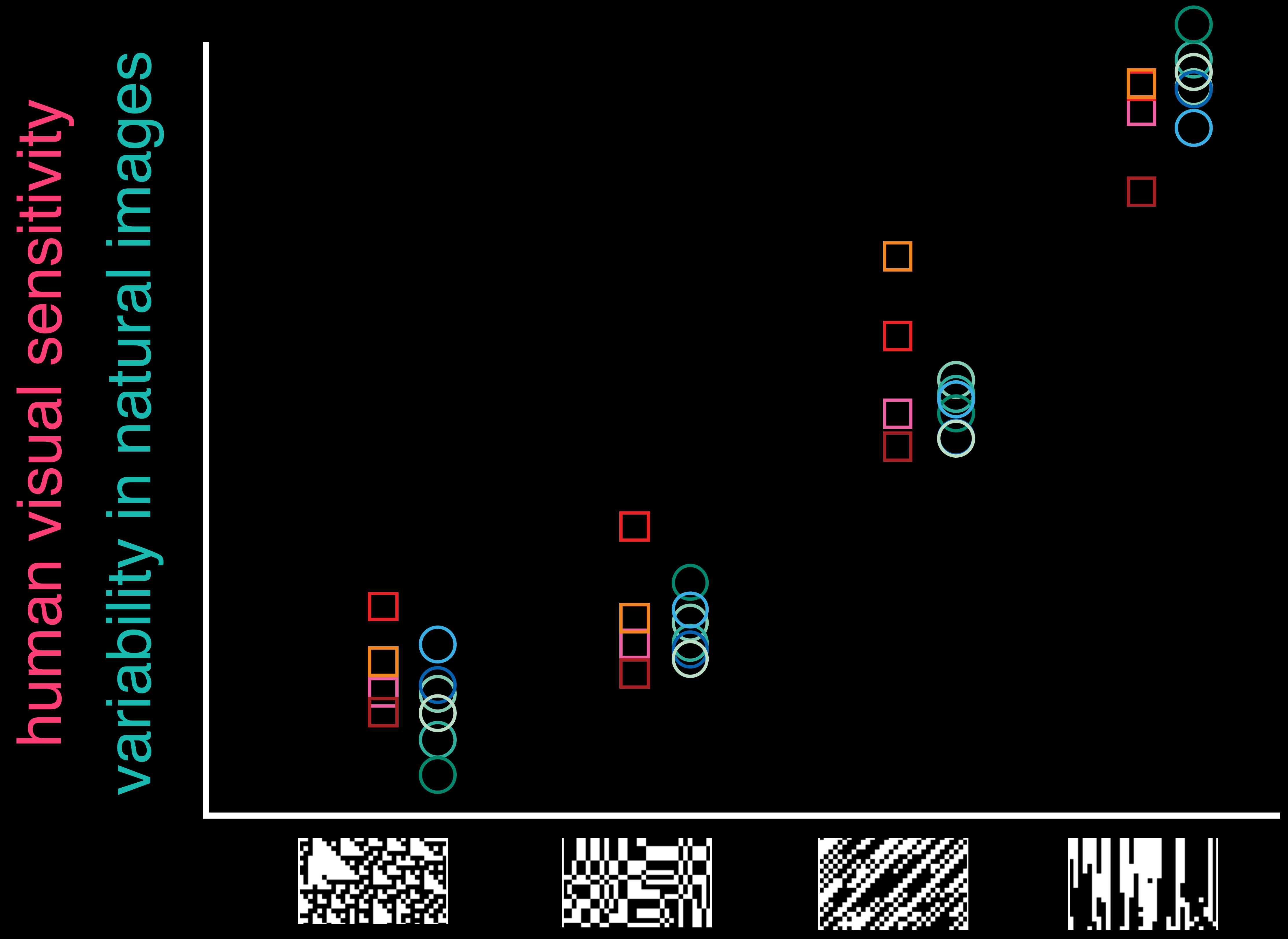
van Hateren 1992

Atick & Redlich 1992



$$|R(f)|^2 \propto |K(f)|^2 \cdot |S(f)|^2$$



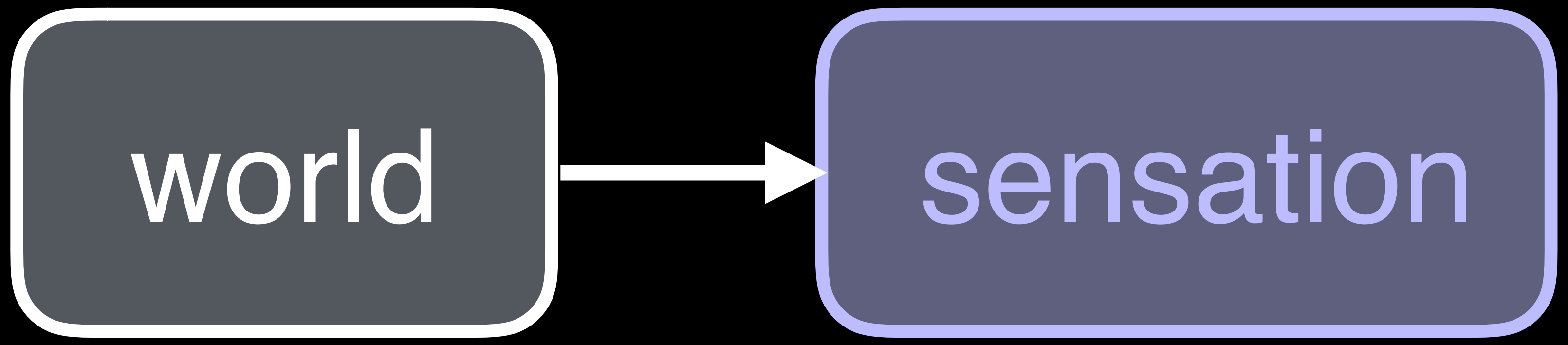


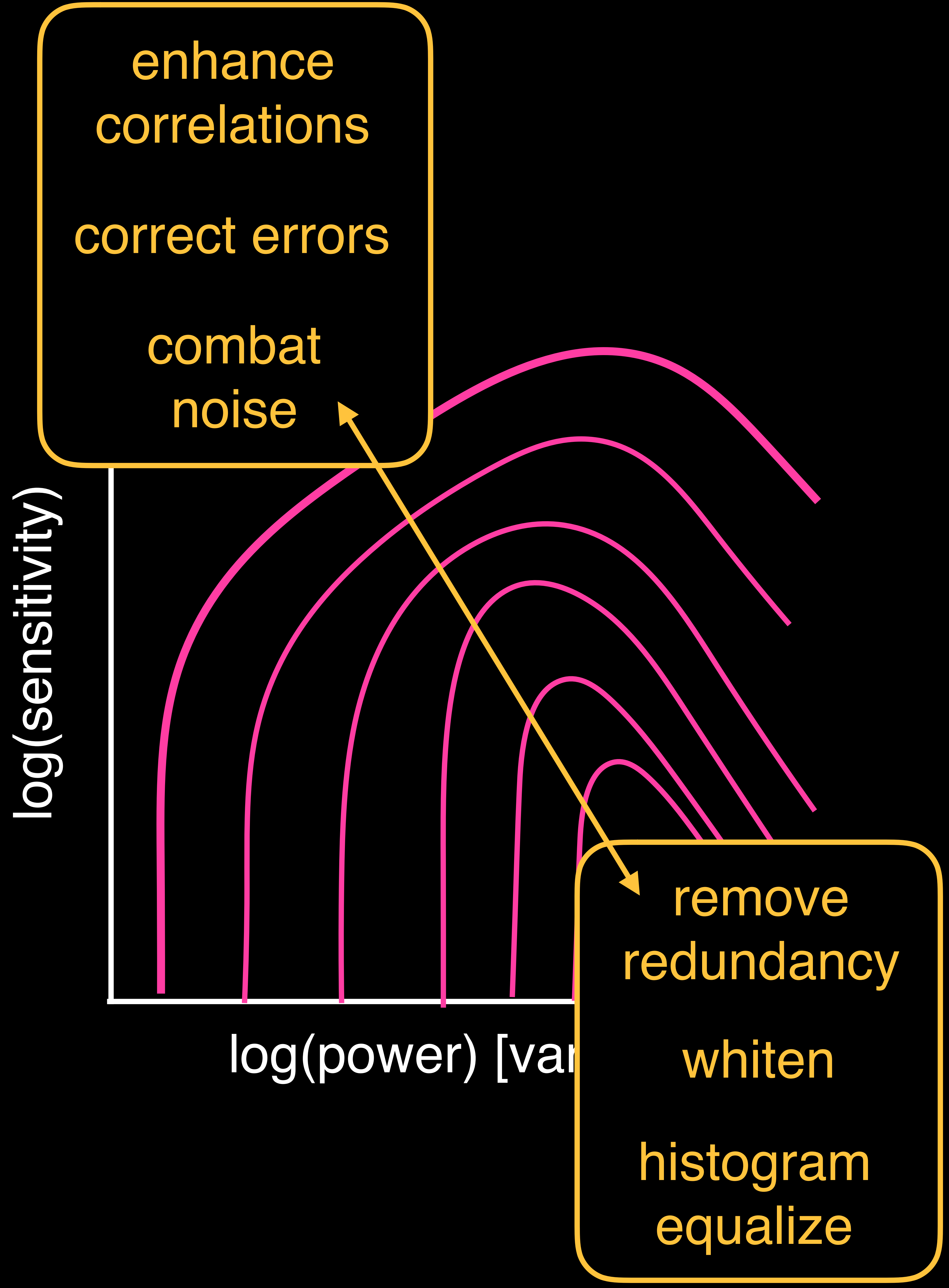
van Hateren 1992

AMH, Briguglio, Conte, Victor, Balasubramanian, Tkacik 2012

PART 1

remove redundancy
combat noise



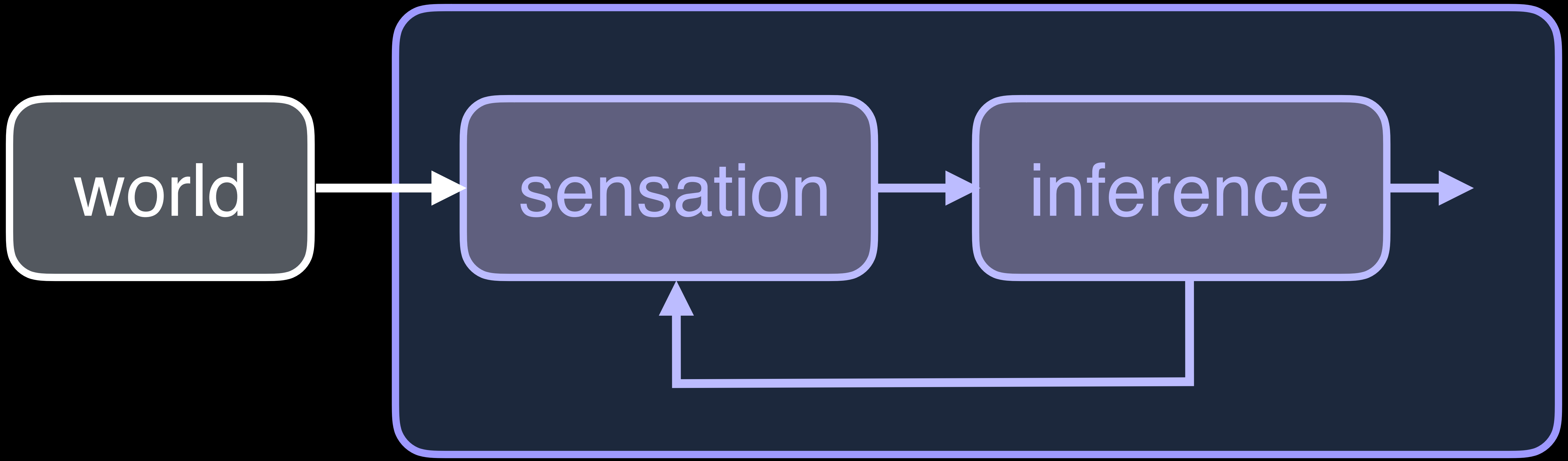


PART 1

remove redundancy
combat noise

PART 2

resolve ambiguity



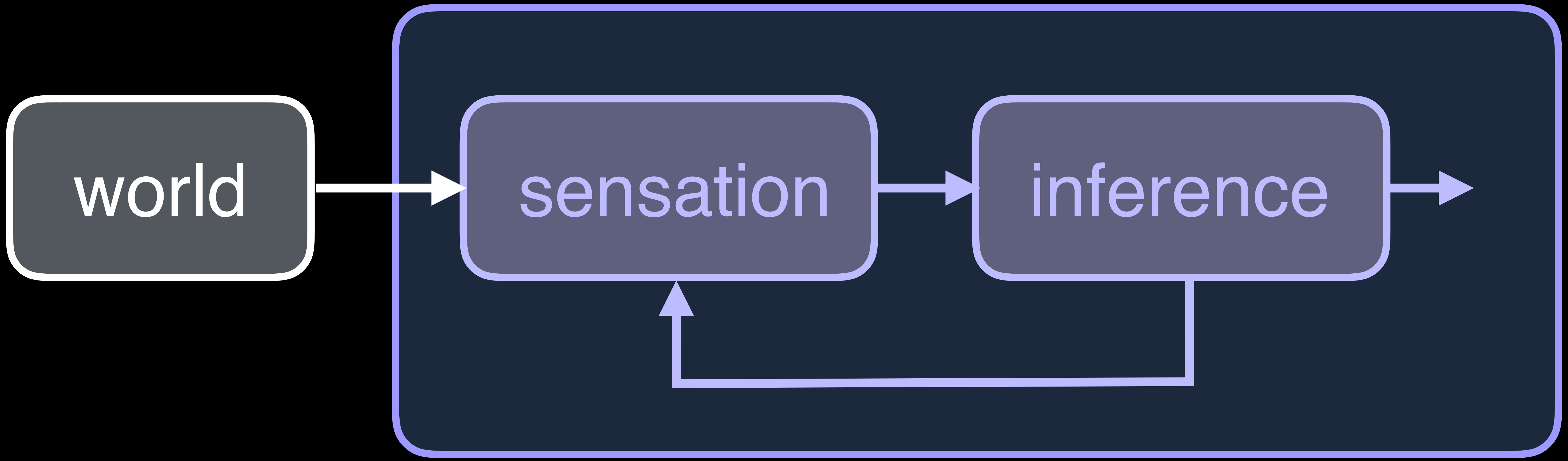
back at 1:20

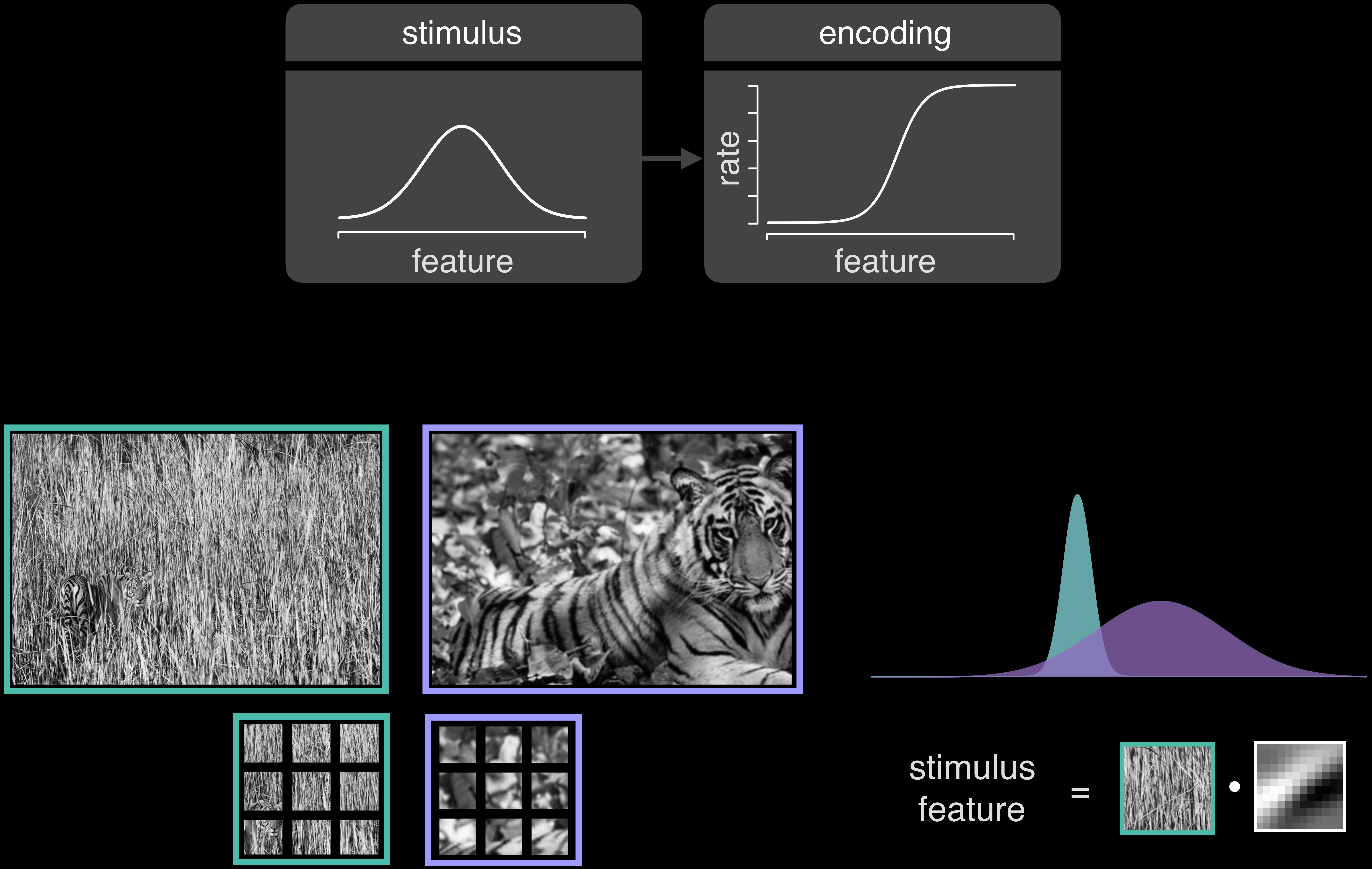
PART 1

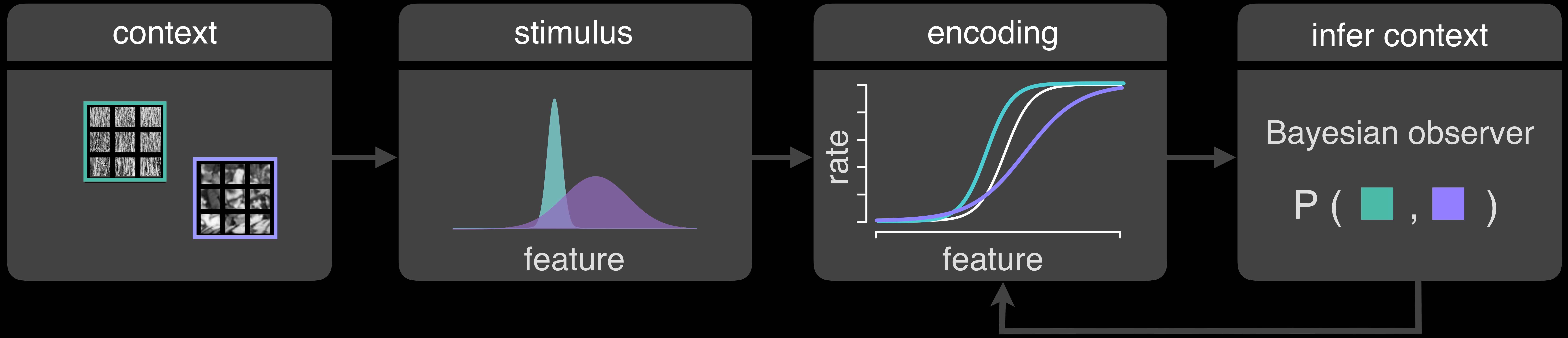
remove redundancy
combat noise

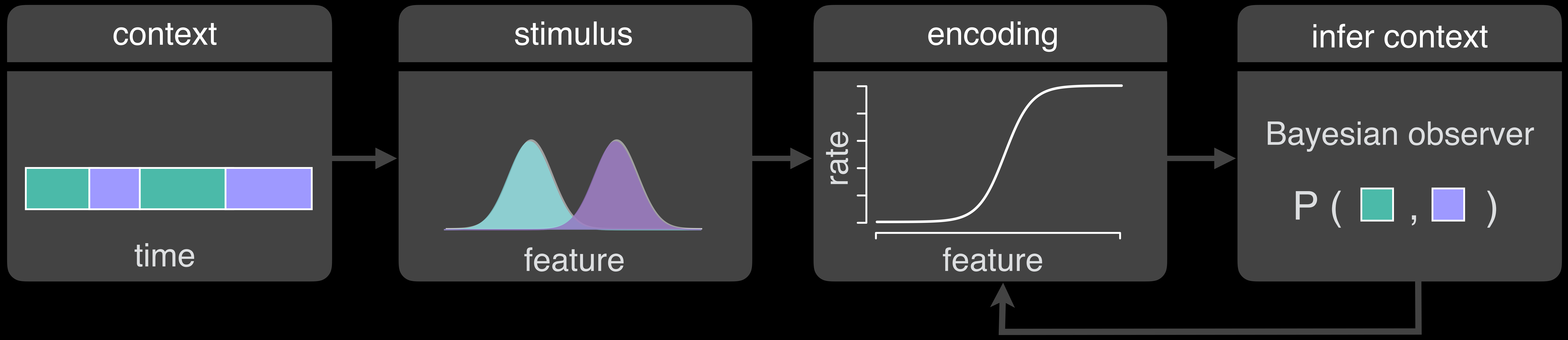
PART 2

resolve ambiguity

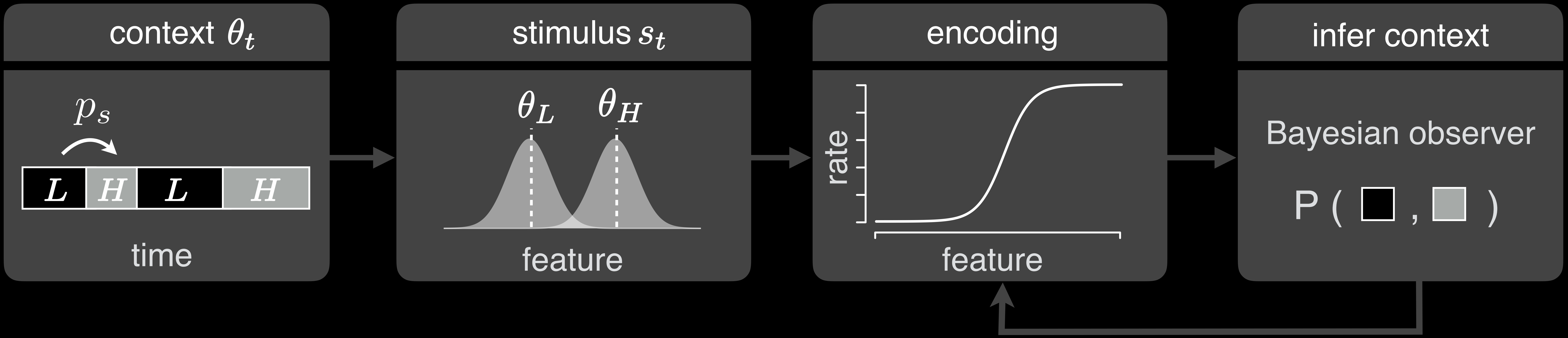



stimulus
encoding
rate
feature
feature
stimulus
feature
=


context
stimulus
feature
encoding
rate
feature
infer context
Bayesian observer
P (,)



context
time
stimulus
feature
encoding
rate
feature
infer context
Bayesian observer
P (,)



context dynamics

$$P(\theta_t|\theta_{t-1}) =$$

$$\begin{array}{c} \\ \theta_L \\ \theta_H \end{array} \begin{array}{c} \begin{bmatrix} (1-p_s) & p_s \\ p_s & (1-p_s) \end{bmatrix} \\ \theta_L \qquad \theta_H \end{array}$$

stimulus distribution

$$P(s_t|\theta_t) =$$

$$\mathcal{N}(s_t\,; \theta_t\,, \sigma^2)$$

Bayesian observer

knows $P(\theta_t|\theta_{t-1})$

$P(s_t|\theta_t)$

wants to estimate $P(\theta_t|s_t, s_{\tau<t})$

DeWeese & Zador, 1998

$$P(\theta_t | s_t, s_{\tau<t})$$

wants to estimate

$$P(s_t | \theta_t), P(\theta_t | \theta_{t-1})$$

knows

$$P(A, B) = P(B, A)$$

$$P(A|B)P(B) = P(B|A)P(A)$$

Bayes Rule

$$P(A|B) = \frac{P(B|A)\, P(A)}{P(B)}$$

$P(\theta_t | s_t, s_{\tau<t})$

wants to estimate

$P(s_t | \theta_t), P(\theta_t | \theta_{t-1})$

knows

Bayes Rule

$$P(A|B, C) = \frac{P(B|A, C)\, P(A|C)}{P(B|C)}$$

$$A = \theta_t$$
$$B = s_t$$
$$C = s_{\tau<t}$$

$$P(\theta_t | s_t, s_{\tau<t}) = \frac{P(s_t | \theta_t, s_{\tau<t})\, P(\theta_t | s_{\tau<t})}{P(s_t | s_{\tau<t})}$$

$$P(\theta_t|s_t, s_{\tau<t}) = \frac{P(s_t|\theta_t, \cancel{s_{\tau<t}})\; P(\theta_t|s_{\tau<t})}{P(s_t|s_{\tau<t})}$$

$$P(\theta_t|s_t, s_{\tau<t}) = \frac{P(s_t|\theta_t)\ P(\theta_t|s_{\tau<t})}{P(s_t|s_{\tau<t})}$$

$$\sum_{\theta_t} P(\theta_t|s_t, s_{\tau<t}) = 1 = \sum_{\theta_t} \frac{P(s_t|\theta_t)\ P(\theta_t|s_{\tau<t})}{P(s_t|s_{\tau<t})}$$

$$P(s_t|s_{\tau<t}) = \sum_{\theta_t} P(s_t|\theta_t)\ P(\theta_t|s_{\tau<t})$$

$$= \Omega$$

$$P(\theta_t | s_t, s_{\tau<t}) = \frac{1}{\Omega} P(s_t | \theta_t) \underset{*}{P(\theta_t | s_{\tau<t})}$$

$$* \; P(\theta_t \,|\, s_{\tau<t}) = P(\theta_t \,|\, s_{\tau<t})$$

$$P(\theta_t | s_t, s_{\tau<t}) = \frac{1}{\Omega} P(s_t|\theta_t) \underset{*}{P(\theta_t|s_{\tau<t})}$$

$$* \; P(\theta_t | s_{\tau<t}) = \sum_{\theta_{t-1}} P(\theta_t | \theta_{t-1}, \cancel{s_{\tau<t}}) \, P(\theta_{t-1} | s_{\tau<t})$$

$$P(\theta_t | s_t, s_{\tau<t}) = \frac{1}{\Omega} P(s_t|\theta_t)\, P(\underset{*}{\theta_t|s_{\tau<t}})$$

$$* \; P(\theta_t \,|\, s_{\tau<t}) = \sum_{\theta_{t-1}} P(\theta_t \,|\theta_{t-1})\, \underline{P(\theta_{t-1} \,|\, s_{\tau<t})}$$

$$P(\theta_{t-1} | s_{t-1}, s_{\tau<t-1})$$

$$P(\theta_t|s_t, s_{\tau<t}) = \frac{1}{\Omega} P(s_t|\theta_t) \sum_{\theta_{t-1}} P(\theta_t|\theta_{t-1})\, P(\theta_{t-1}|s_{\tau<t})$$

① ② ③ ④

① $P_t^L \equiv P(\theta_t=\theta_L|s_t, s_{\tau<t})$

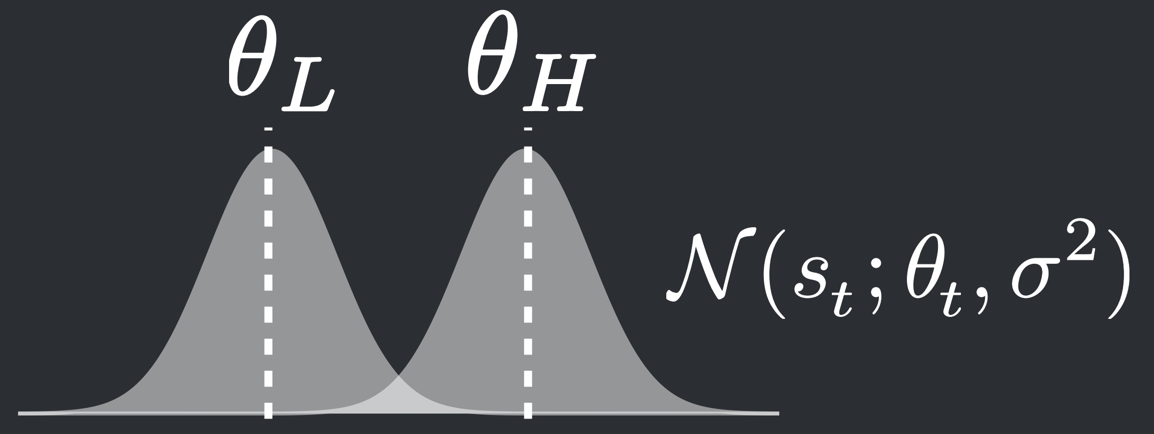
$P_t^H = (1-P_t^L)$

③•④

$$\begin{array}{c|cc|}
 & \theta_L & \theta_H \\
\theta_t=\theta_L & (1-p_s) & p_s \\
\theta_H & p_s & (1-p_s)
\end{array}
\quad
\begin{array}{c}
\theta_L \\ \theta_H
\end{array}
\begin{bmatrix} P_{t-1}^L \\ (1-P_{t-1}^L) \end{bmatrix}$$

② $\mathcal{N}(s_t;\theta_t=\theta_L,\sigma^2)$

changing context θ_t



stimulus feature s_t

$$P(\theta_t | s_t, s_{\tau<t}) = \frac{1}{\Omega} P(s_t | \theta_t) \sum_{\theta_{t-1}} P(\theta_t | \theta_{t-1}) \, P(\theta_{t-1} | s_{\tau<t})$$

① ② ③ ④

① $P_t^L \equiv P(\theta_t = \theta_L | s_t, s_{\tau<t})$

$P_t^H = (1 - P_t^L)$

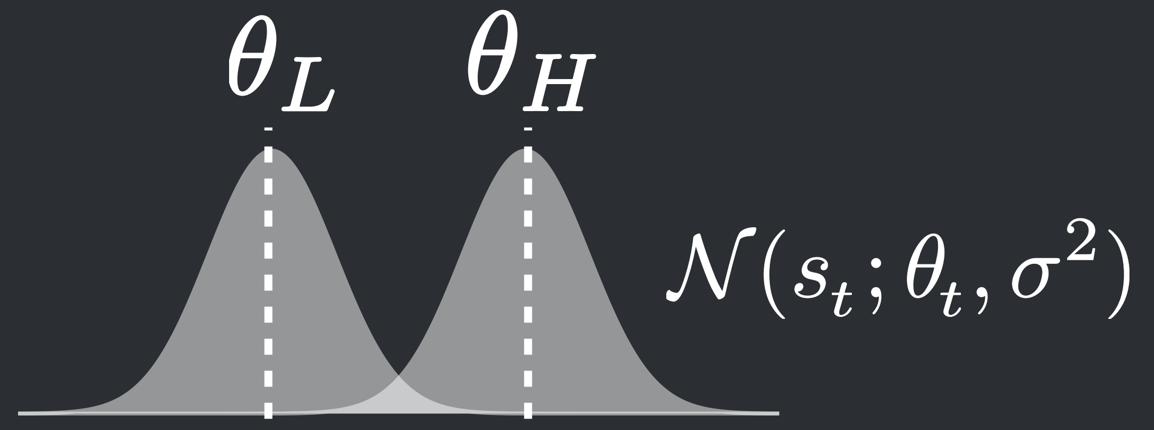
③•④

$$\begin{array}{c|cc|}
 & \theta_L & \theta_H \\
\hline
\theta_t = \theta_L & (1-p_s) & p_s \\
\theta_H & p_s & (1-p_s)
\end{array}
\quad
\begin{array}{c|c|}
\theta_L & P_{t-1}^L \\
\theta_H & (1 - P_{t-1}^L)
\end{array}$$

$$= \left[(1-p_s) \, P_{t-1}^L + p_s \, (1 - P_{t-1}^L) \right]$$

② $\mathcal{N}(s_t; \theta_t = \theta_L, \sigma^2)$

changing context θ_t



stimulus feature s_t

$$P(\theta_t|s_t, s_{\tau<t}) = \frac{1}{\Omega} P(s_t|\theta_t) \sum_{\theta_{t-1}} P(\theta_t|\theta_{t-1}) P(\theta_{t-1}|s_{\tau<t})$$

$$P_t^L = \frac{1}{\Omega} \mathcal{N}(s_t; \theta_t = \theta_L, \sigma^2) \left[(1-p_s) P_{t-1}^L + p_s (1 - P_{t-1}^L)\right]$$

$$P_t^L = \frac{1}{\Omega}\, \mathcal{N}(s_t ; \theta_t = \theta_L , \sigma^2) \left[(1 - p_s)\, P_{t-1}^L + p_s\, (1 - P_{t-1}^L)\right]$$

posterior probability of LOW context (P_t^L)

likelihood that observed stimulus was generated in LOW context ($\mathcal{N}(s_t ; \theta_t = \theta_L , \sigma^2)$)

probability that context stayed LOW ($1 - p_s$)

prior probability of LOW context (P_{t-1}^L)

probability that context changed to HIGH (p_s)

prior probability of HIGH context ($1 - P_{t-1}^L$)

$$P_t^L = \frac{1}{\Omega}\,\mathcal{N}(s_t\,;\theta_t=\theta_L\,,\sigma^2)\,\left[(1-p_s)\,P_{t-1}^L + p_s\,(1-P_{t-1}^L)\right]$$

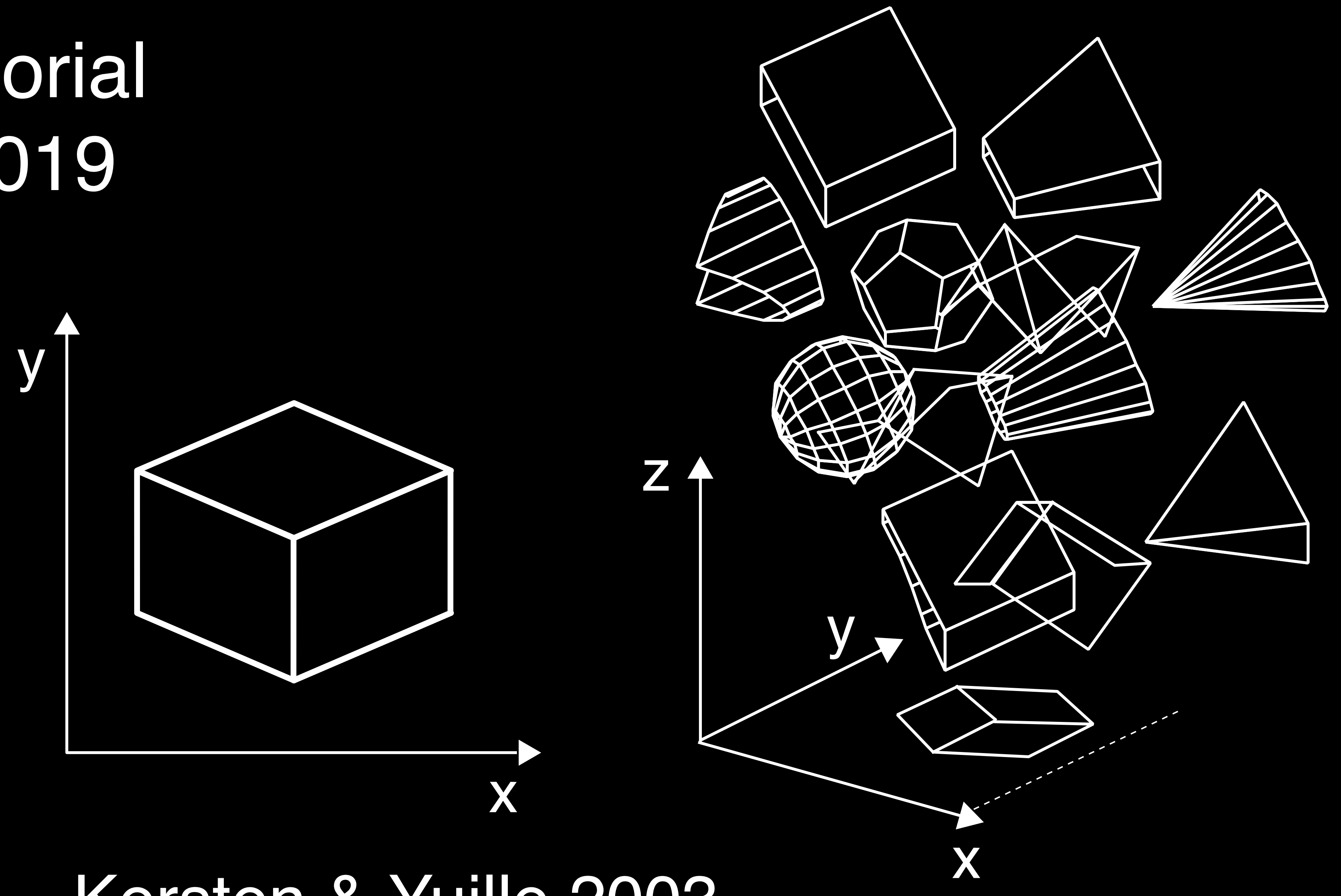
posterior ← $\propto$ — **likelihood** ← **prior**

how probable are your hypotheses about state of the world?

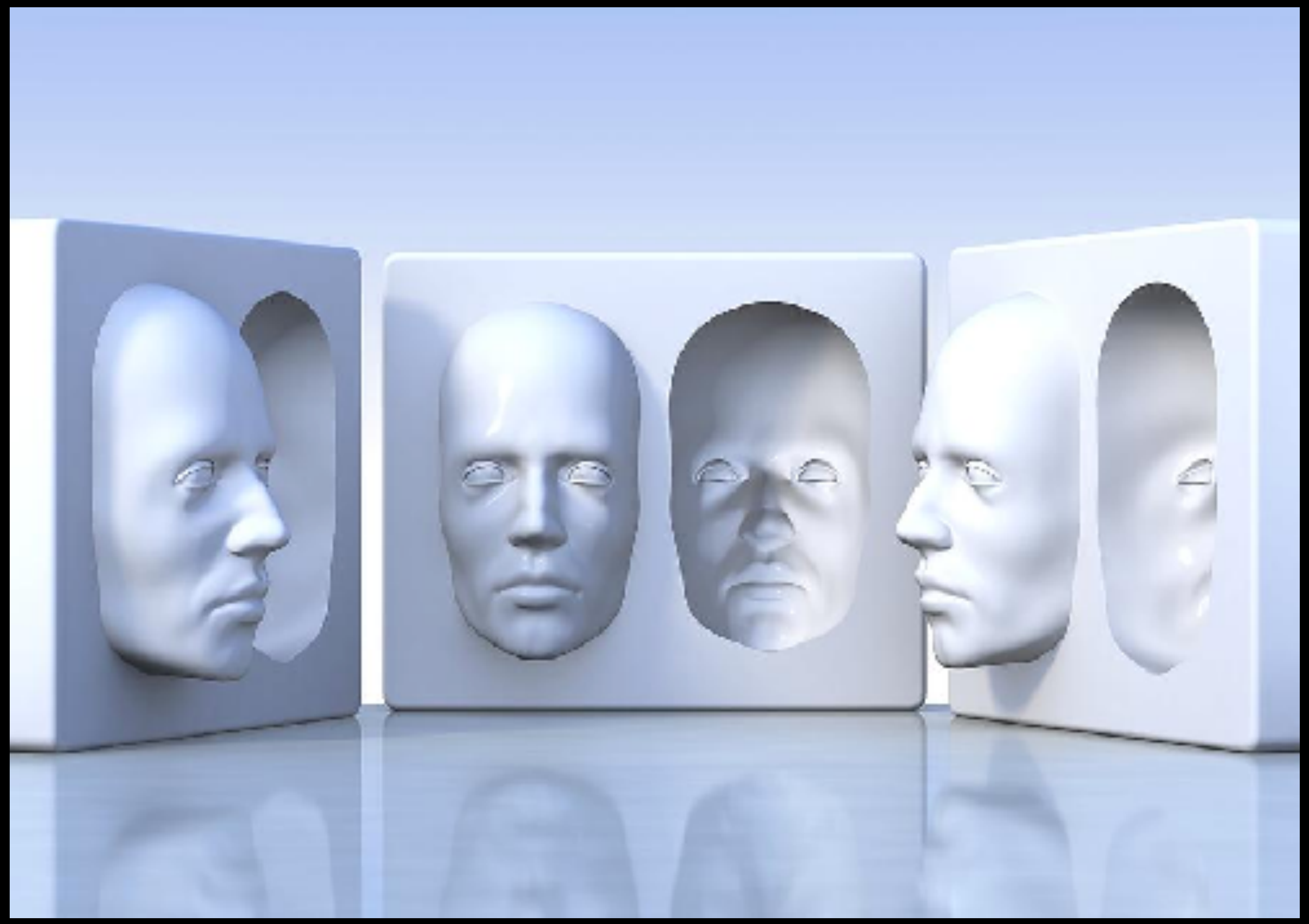
how probable are your measurements given your hypotheses?

how much do you expect your hypotheses based on your experiences?

*see Wei Ji's tutorial from Cosyne 2019



Kersten & Yuille 2003

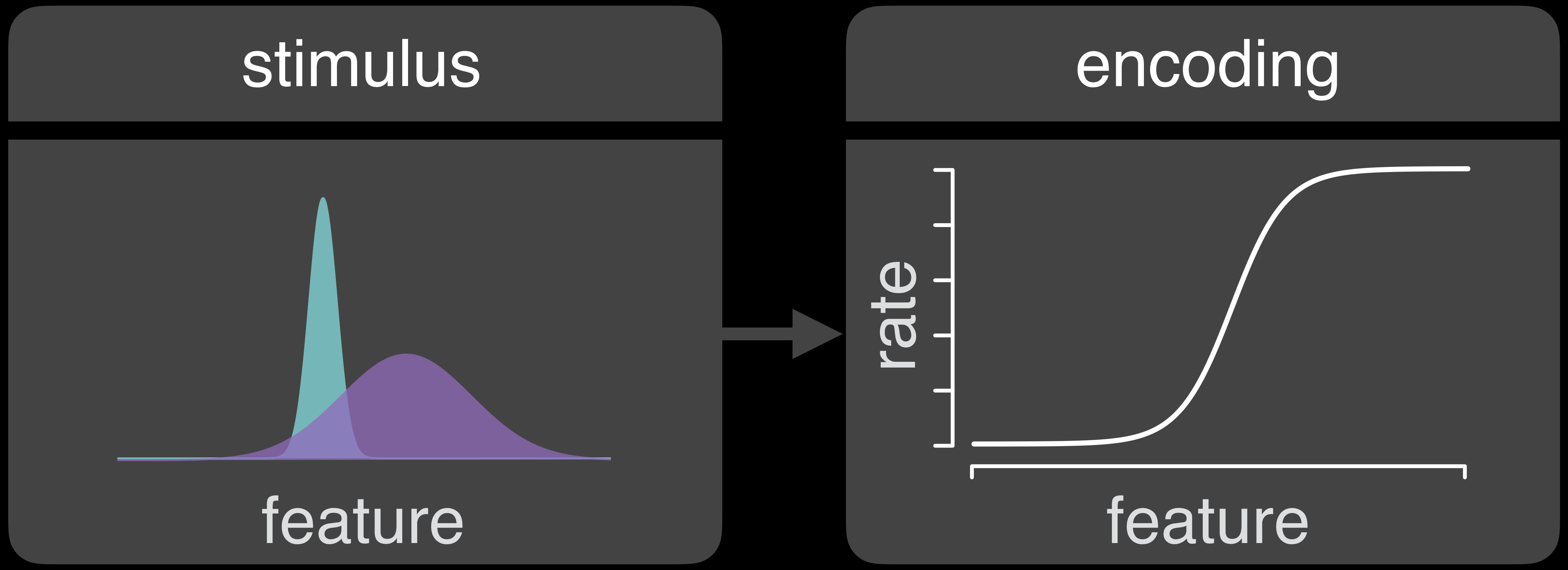


David Mack

$$P_t^L = \frac{1}{\Omega} \mathcal{N}(s_t ; \theta_t = \theta_L , \sigma^2) \Big[(1 - p_s) P_{t-1}^L + p_s (1 - P_{t-1}^L)$$

encoding

what features
should be prioritized
to maximize information?

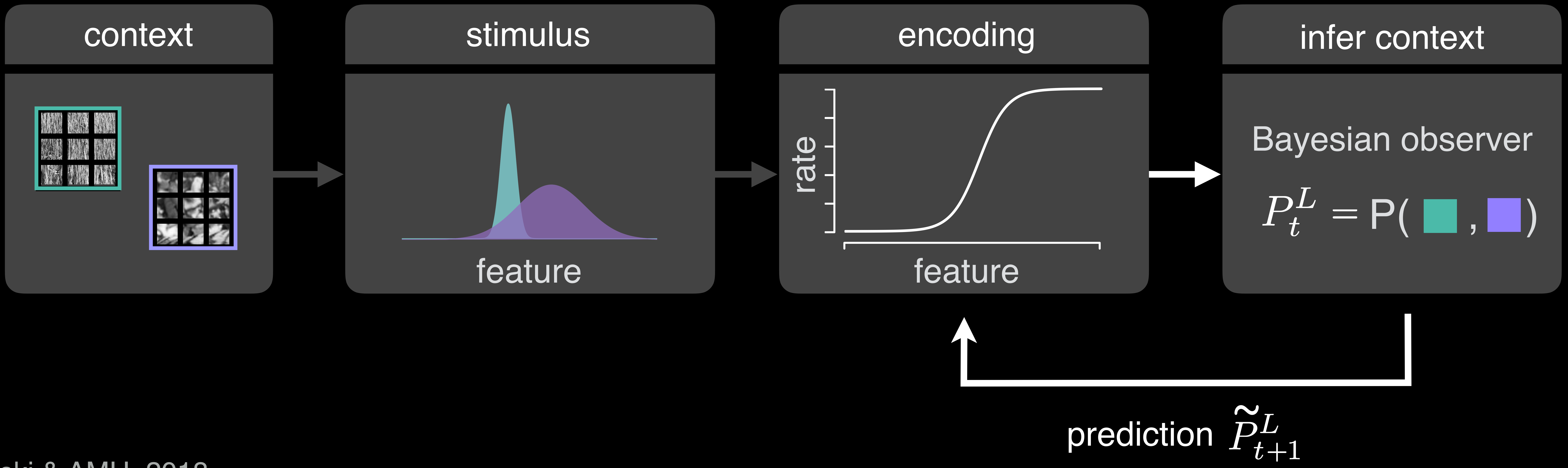


$$P_t^L = \frac{1}{\Omega} \mathcal{N}(s_t; \theta_t = \theta_L, \sigma^2) \left[(1 - p_s) P_{t-1}^L + p_s (1 - P_{t-1}^L) \right]$$

encoding → **posterior**

what features should be prioritized to maximize information?
to test hypotheses?

how probable are your hypotheses about state of the world?



Mlynarski & AMH, 2018

PART 1

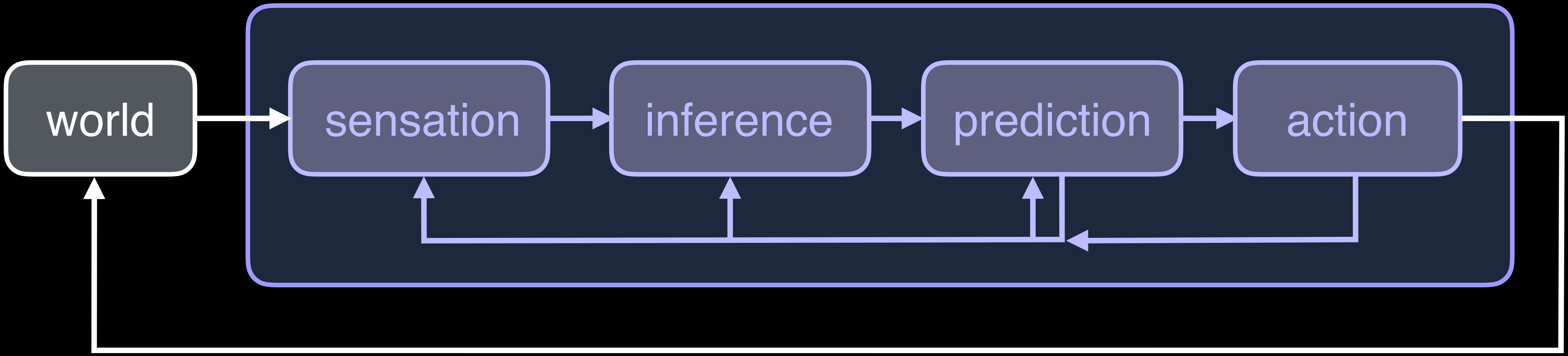
remove redundancy
combat noise

PART 2

resolve ambiguity

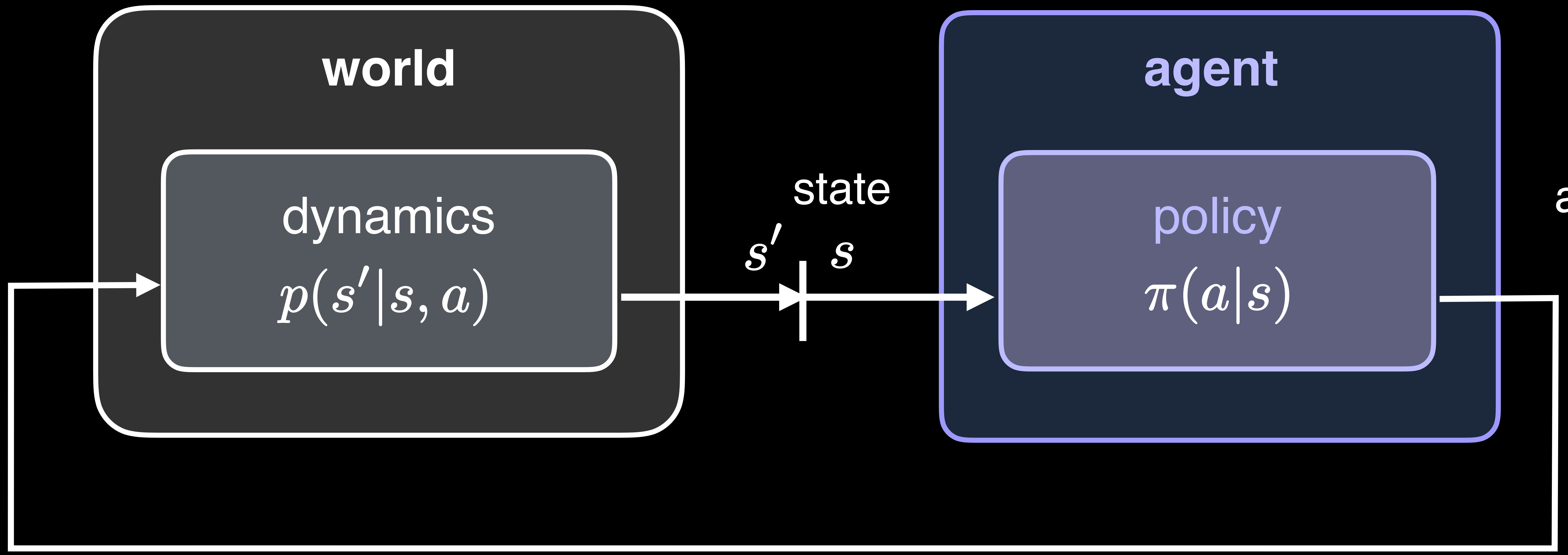
PART 3

make & use
predictions

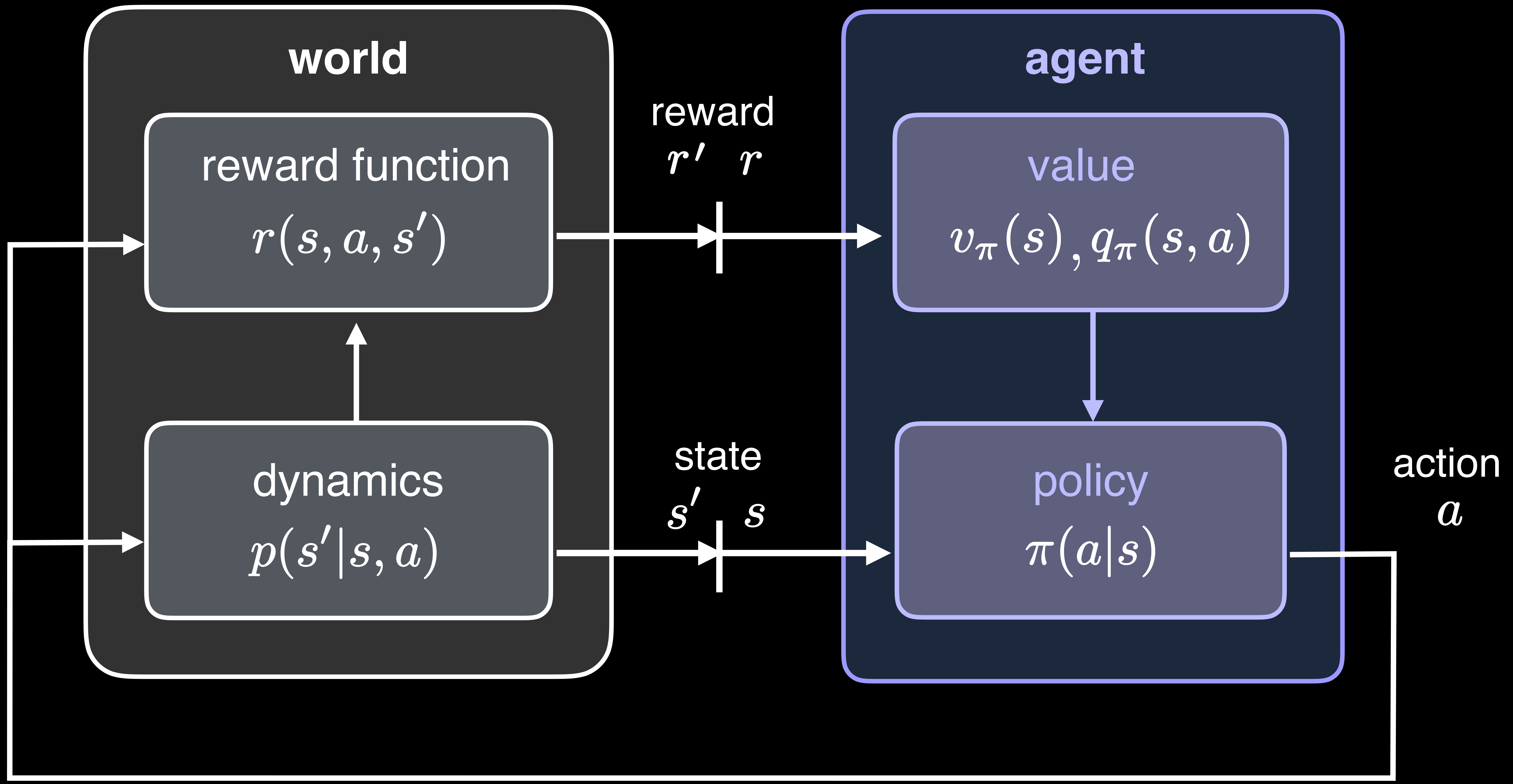


gather information (e.g. infotaxis)
maximize reward (e.g. reinforcement learning)

Vergasola, Villermaux, Shraiman, 2007
Sutton & Barto, 2018



Sutton & Barto, 2018



Sutton & Barto, 2018

day 1, trial 1

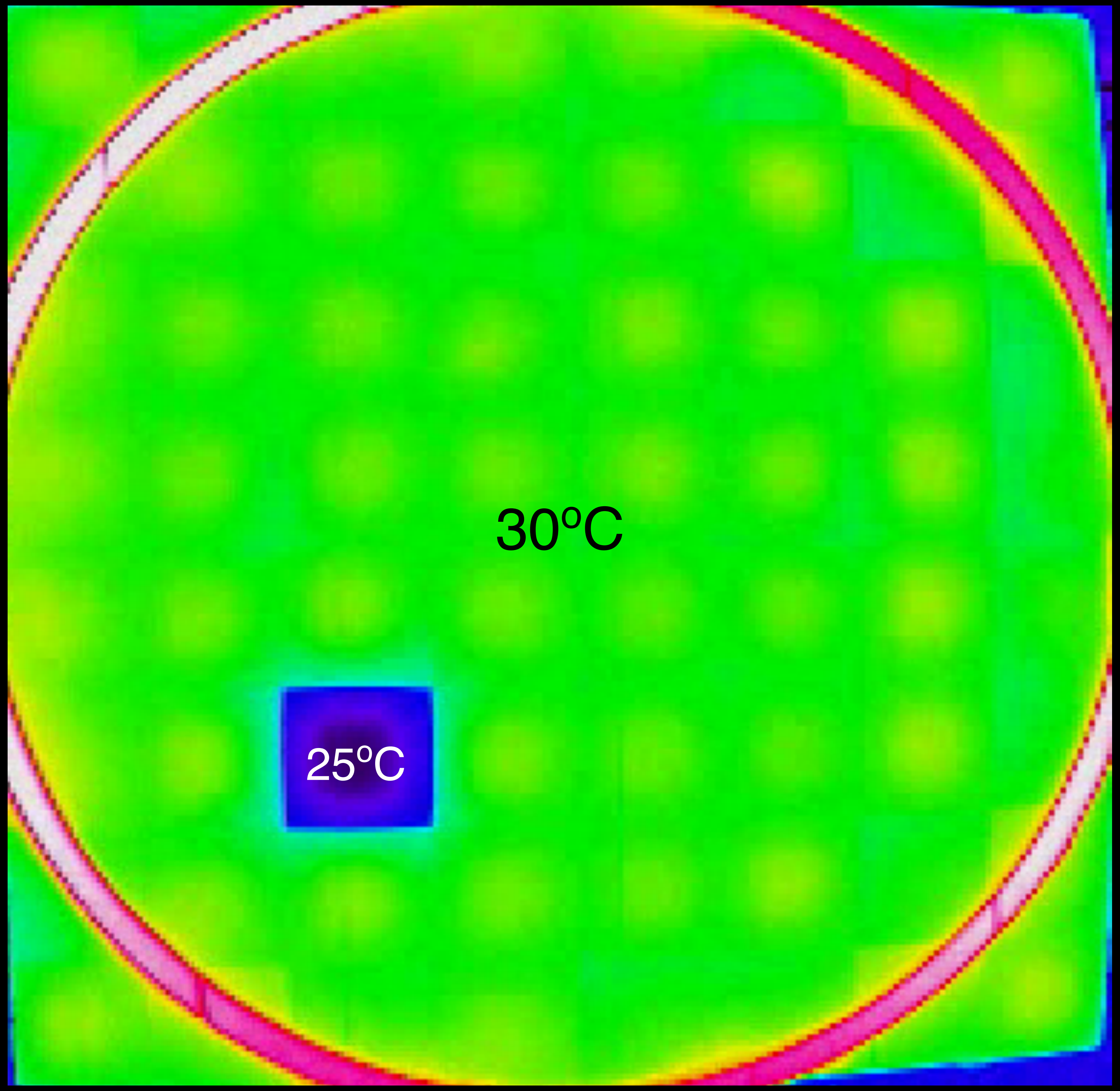
day 5, trial 10

https://tinyurl.com/vfwo8ze

day 1, trial 1

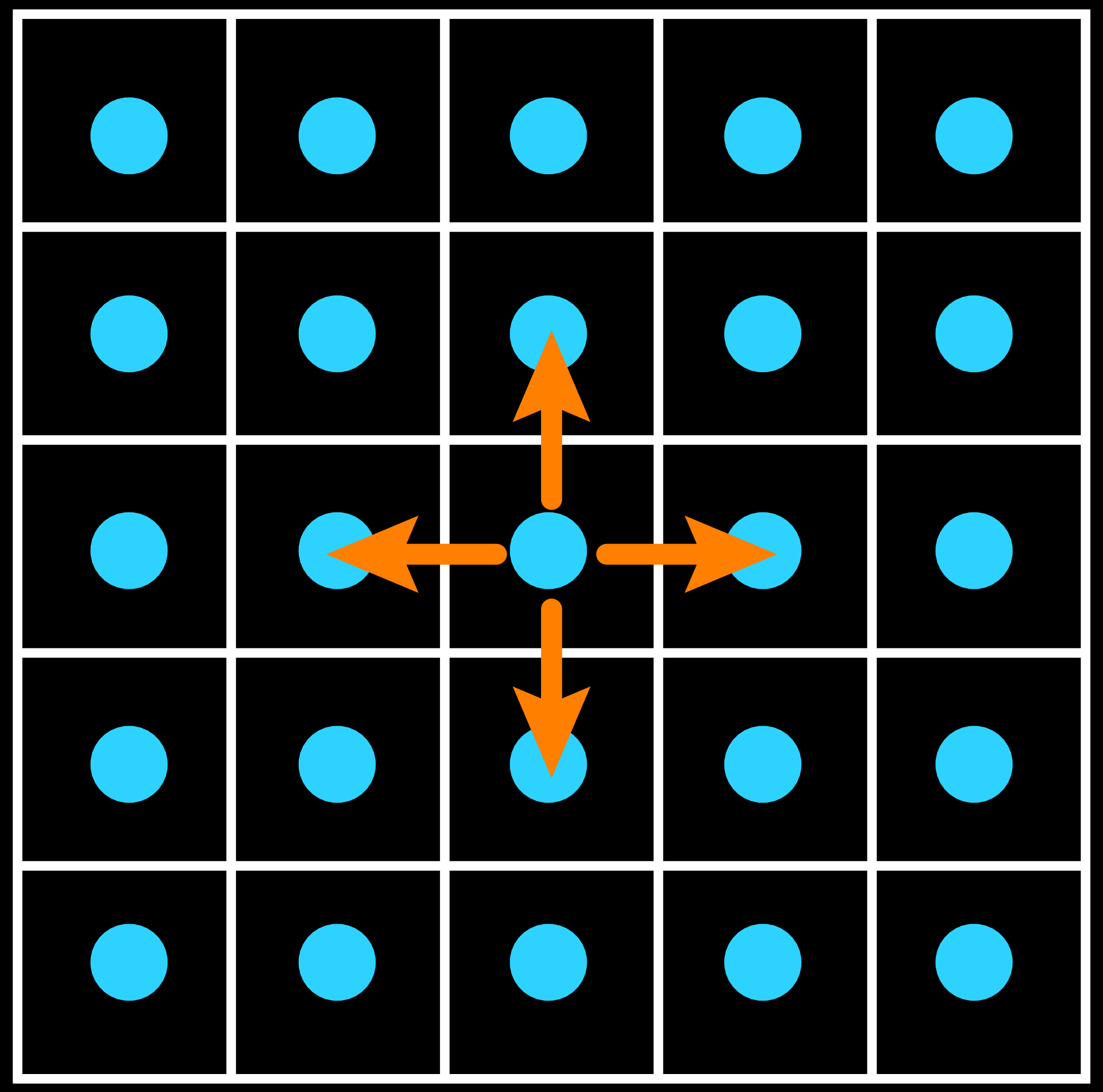
day 5, trial 10

https://tinyurl.com/vfwo8ze



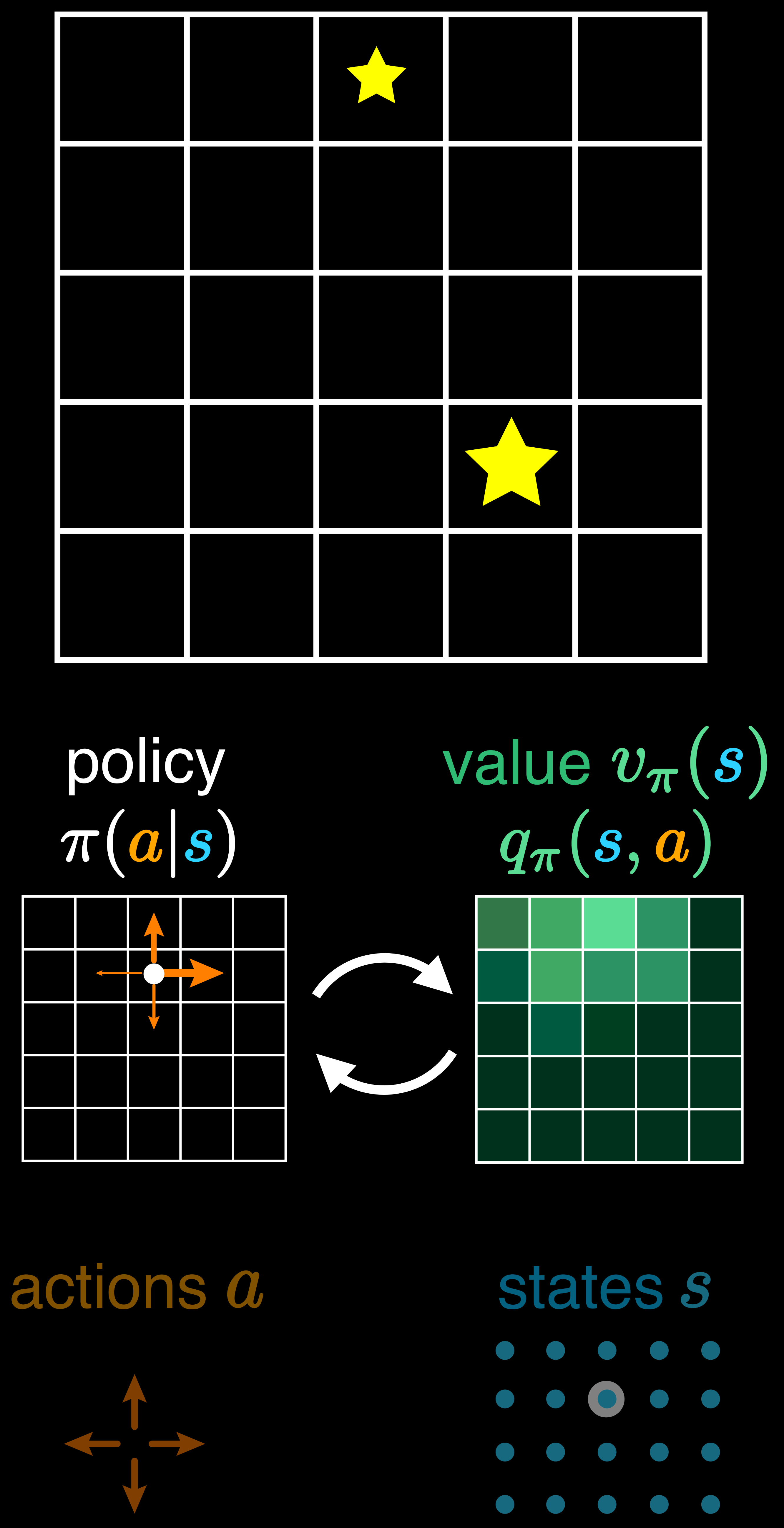
Ofstad, Zuker & Reiser, Nature (2011)

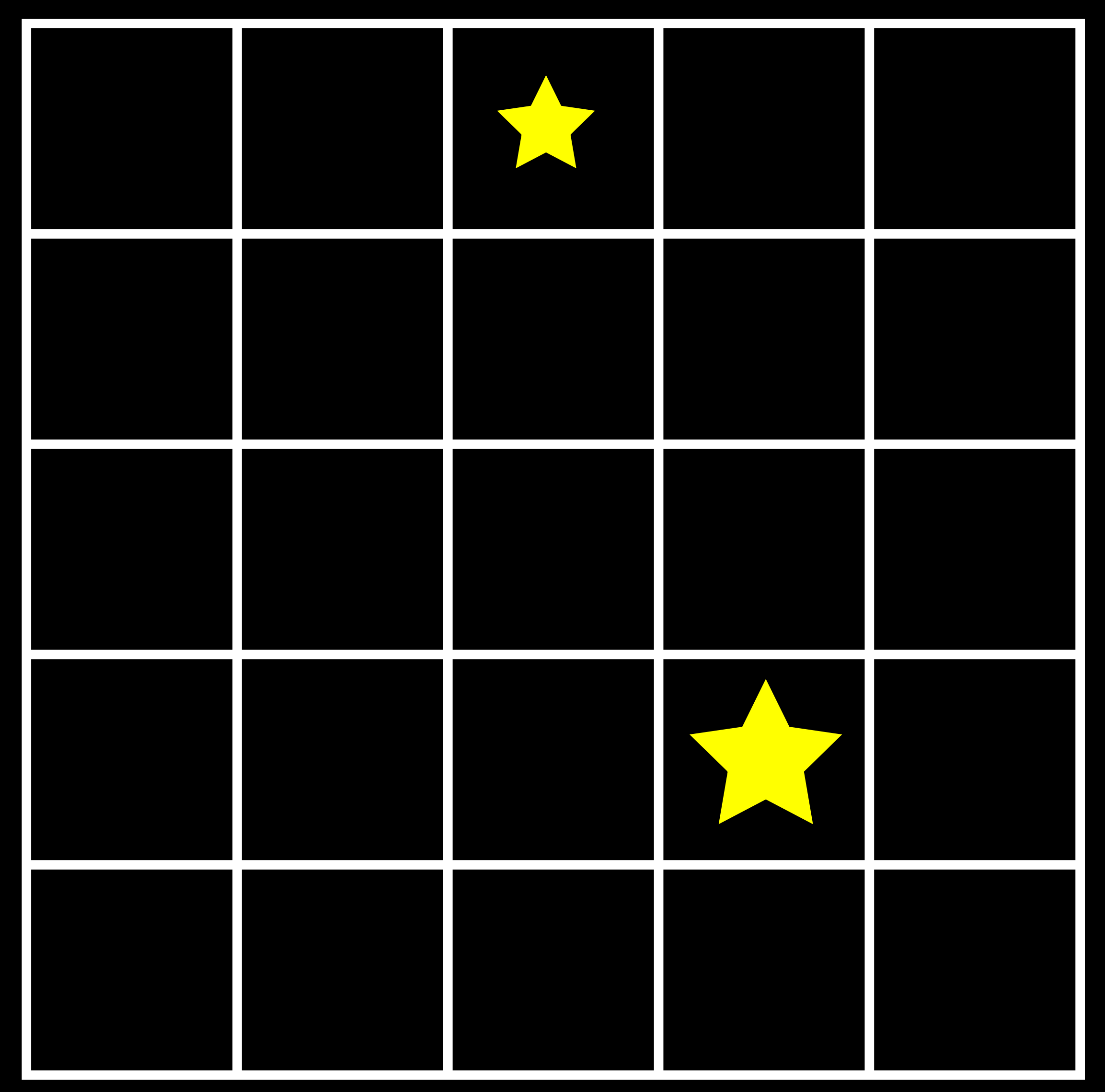
* fly tracking by Ctrax (Branson et al. 2009)

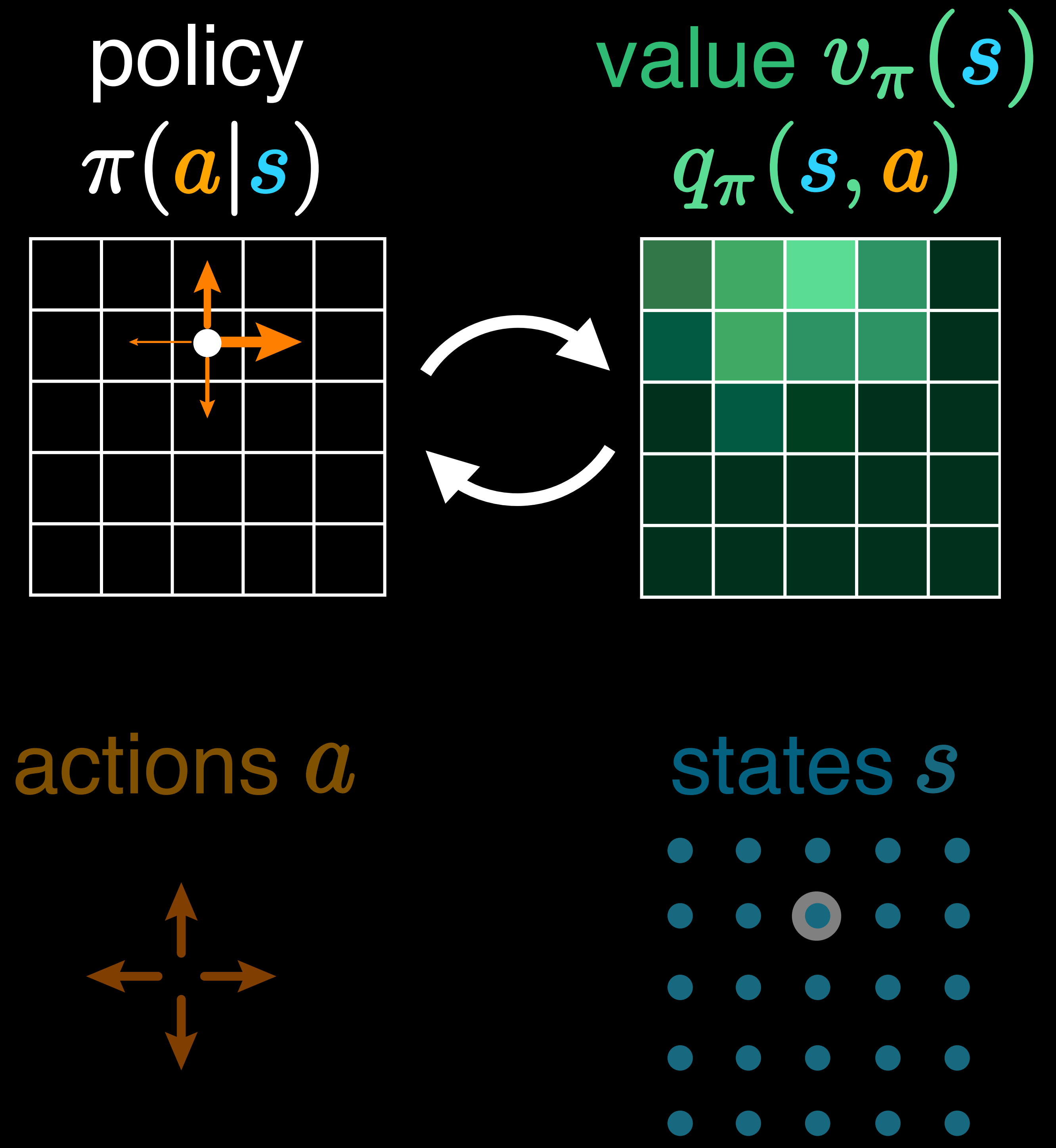


states s

actions a


policy
$\pi(a|s)$
value $v_\pi(s)$
$q_\pi(s, a)$
actions a
states s





explore / exploit tradeoff

exploit: take action that gives highest expected value

explore: take action that has lower expected value but could result in higher long-term payoff

greedy: $A = \underset{a}{\operatorname{argmax}}\ q_\pi(s,a)$ exploit

ϵ-greedy:

$(1-\epsilon)$ $A =$ greedy exploit

ϵ $A =$ random explore

softmax: $\pi(a|s) \propto \exp\left(\beta\, q_\pi(s,a)\right)$

$\beta \to \infty$ exploit

$\beta \to 0$ explore

goal: maximize expected long-term reward

return G

$G_t =$ return, starting at time t

$= R_{t+1} + R_{t+2} + R_{t+3} + \ldots$

all rewards equally important

goal: maximize expected long-term reward

return G

$G_t =$ return, starting at time t

$= R_{t+1} + R_{t+2} + R_{t+3} + \ldots$

all rewards equally important

$= R_{t+1} + R_{t+2} + R_{t+3} + \ldots$

goal: maximize expected long-term reward

return G

G_t = return, starting at time t

$= R_{t+1} + R_{t+2} + R_{t+3} + \ldots$ all rewards equally important

$= R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \ldots$ current rewards more important than distant ones

discount factor $\gamma \in [0, 1]$

$\gamma = 1$ don't discount (far sighted)

$\gamma = 0$ fully discount (myopic)

goal: maximize $\underbrace{\text{expected long-term reward}}_{\text{return } G}$

$$G_t = \text{return, starting at time } t$$

$$= R_{t+1} + R_{t+2} + R_{t+3} + \ldots$$

all rewards equally important

$$= R_{t+1} + \gamma \underbrace{\left[R_{t+2} + \gamma R_{t+3} + \ldots\right]}_{G_{t+1}}$$

current rewards more important than distant ones

discount factor $\gamma \in [0, 1]$

$\gamma = 1$ don't discount (far sighted)

$\gamma = 0$ fully discount (myopic)

$$G_t = R_{t+1} + \gamma G_{t+1}$$

G_t represents actual future rewards (unknown to agent)

can instead compute **expected** future rewards, starting in state s, following policy π

consider **one timestep** in the future:

$$\mathbb{E}_\pi[R_{t+1}|S_t = s] = \sum_a \underbrace{\pi(a|s)}_{\text{policy}} \sum_{s'} \underbrace{p(s'|s,a)}_{\text{dynamics of environment}} \underbrace{r(s,a,s')}_{\text{reward function}}$$

$$G_t = R_{t+1} + \gamma G_{t+1}$$

define **state-value function**, starting in state s , following policy π

$$v_\pi(s) \equiv \mathbb{E}_\pi\left[G_t \;\middle|\; S_t = s \right]$$

prediction of rewards to come

$$G_t = R_{t+1} + \gamma G_{t+1}$$

$$\mathbb{E}_\pi[R_{t+1}|S_t = s] = \sum_a \pi(a|s) \sum_{s'} p(s'|s,a)\, r(s,a,s')$$

define **state-value function**, starting in state s , following policy π

$$v_\pi(s) \equiv \mathbb{E}_\pi\left[R_{t+1} + \gamma G_{t+1} \,\middle|\, S_t = s \right]$$

$$\mathbb{E}_\pi[R_{t+1} | S_t = s] = \sum_a \pi(a|s) \sum_{s'} p(s'|s,a)\; r(s,a,s')$$

define **state-value function**, starting in state s , following policy π

$$v_\pi(s) \equiv \mathbb{E}_\pi\left[R_{t+1} + \gamma G_{t+1} \,\middle|\, S_t = s\right]$$

$$= \sum_a \pi(a|s) \sum_{s'} p(s'|s,a) \left[r(s,a,s') + \gamma \underbrace{\mathbb{E}_\pi\left[G_{t+1} \,\middle|\, S_{t+1} = s'\right]}_{v_\pi(s')} \right]$$

define **state-value function**, starting in state s , following policy π

$$v_\pi(s) \equiv \mathbb{E}_\pi\left[R_{t+1} + \gamma G_{t+1} \;\middle|\; S_t = s \right]$$

$$= \sum_a \pi(a|s) \sum_{s'} p(s'|s,a) \left[r(s,a,s') + \gamma\, v_\pi(s') \right]$$

Bellman equation

for state-value function

define **state-value function**, starting in state s, following policy π

$$v_\pi(s) = \sum_a \pi(a|s) \underbrace{\sum_{s'} p(s'|s,a) \left[r(s,a,s') + \gamma\, v_\pi(s') \right]}_{q_\pi(s,a)}$$

state-value function

action-value function

Dynamic Programming: given $p(s'|s,a)$, $r(s,a,s')$, learn optimal v_*, q_*, π_* via **bootstrapping**

Monte Carlo: estimate v_*, q_* via sampling, learn optimal π_* from **simulated experiences**

temporal difference (TD)

learns from **experience** via **bootstrapping** define **state-value function**, starting in state s, following policy π

$$v_\pi(s) = \sum_a \pi(a|s) \sum_{s'} p(s'|s,a) \left[r(s,a,s') + \gamma\, v_\pi(s') \right]$$ state-value function

temporal difference (TD) : improve estimate of value through experience

new estimate = old estimate + step size [target − old estimate]

$$V_{t+1}(s) = V_t(s) + \alpha \left[\underline{G_t} - V_t(s) \right]$$

$$\simeq \left(R + \gamma V_t(s')\right)$$

define **state-value function**, starting in state s , following policy π

$$v_\pi(s) = \sum_a \pi(a|s) \sum_{s'} p(s'|s,a) \left[r(s,a,s') + \gamma\, v_\pi(s') \right]$$ state-value function

$$V_{t+1}(s) = V_t(s) + \alpha \left[\left(R + \gamma V_t(s')\right) - V_t(s) \right]$$

$$Q_{t+1}(s,a) = Q_t(s,a) + \alpha \left[\left(R + \gamma Q_t(s',a')\right) - Q_t(s,a) \right]$$ Q-learning, SARSA

prediction error δ_t

here, states can only be updated as they are visited

to update states that were visited in the past, we can use **eligibility traces**

$$V_{t+1}(s) = V_t(s) + \alpha\Big[\big(R + \gamma V_t(s')\big) - V_t(s)\Big] \; Z_t(s)$$

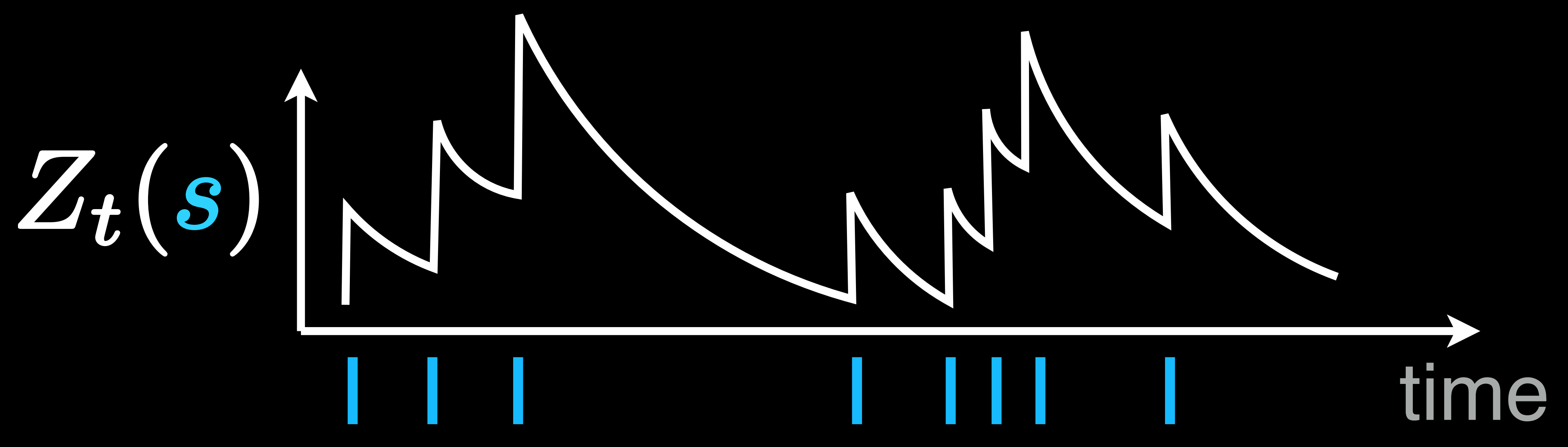
"eligibility" of state s

$$Z_t(s) = \begin{cases} \lambda\gamma Z_{t-1}(s) & s \neq S_t \\ 1 + \lambda\gamma Z_{t-1}(s) & s = S_t \end{cases}$$

trace-decay parameter $\lambda \in [0, 1]$

$\lambda = 0$ only current state can be updated

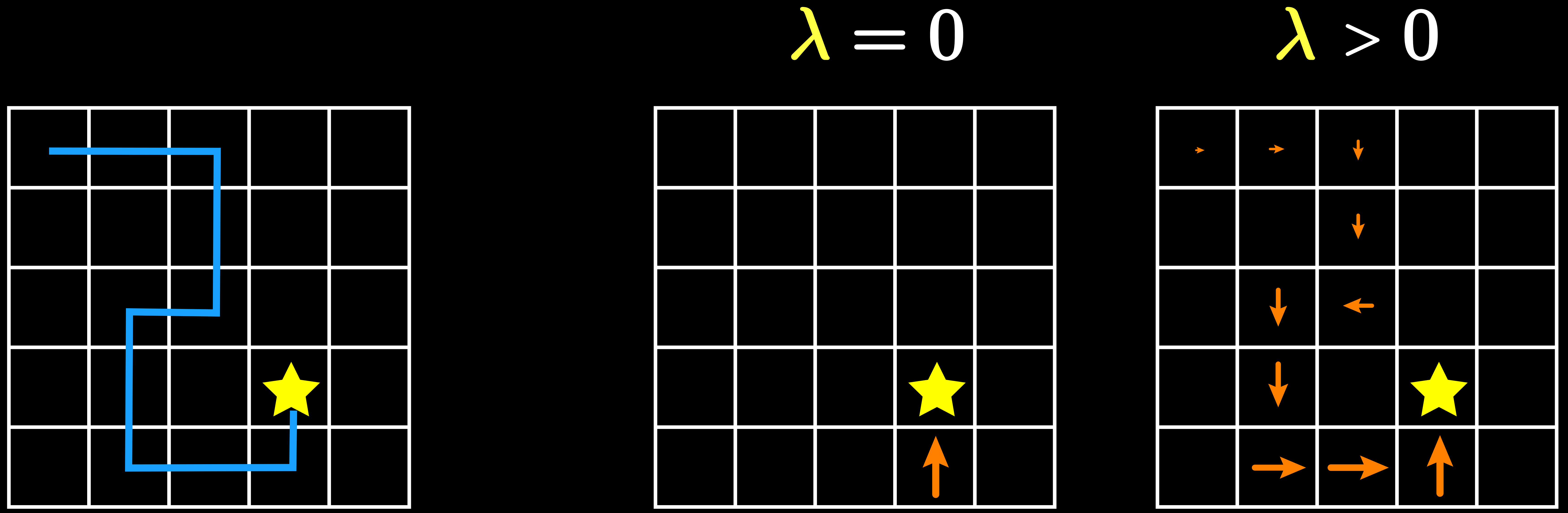
$\lambda = 1$ eligibility falls by γ each timestep

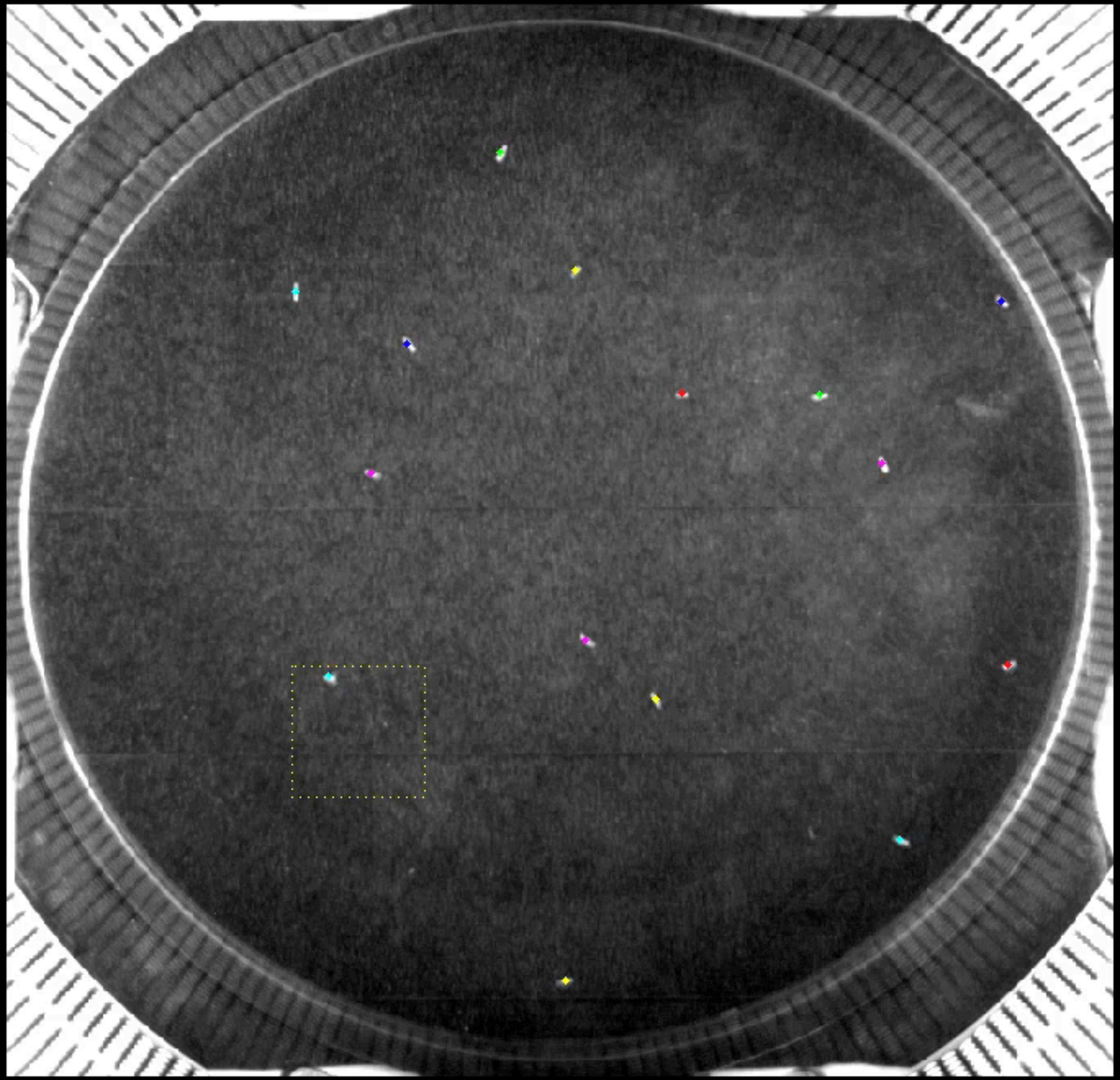


here, states can only be updated as they are visited

to update states that were visited in the past, we can use **eligibility traces**

$$V_{t+1}(s) = V_t(s) + \alpha\left[\left(R + \gamma V_t(s')\right) - V_t(s)\right] \quad Z_t(s) = \begin{cases} \lambda\gamma Z_{t-1}(s) & s \neq S_t \\ 1 + \lambda\gamma Z_{t-1}(s) & s = S_t \end{cases}$$





day 1, trial 1

day 5, trial 10

Ofstad, Zuker & Reiser, Nature (2011)

* fly tracking by Ctrax (Branson et al. 2009)

https://tinyurl.com/vfwo8ze

back at 2:40

wifi

Hilton Honors Meeting

csnventi20

problem set & code

http://bit.ly/cosyne2020-tutorial

worksheet

(handout)

outline

1 | problem setup

2 | sensory coding

3 | inference

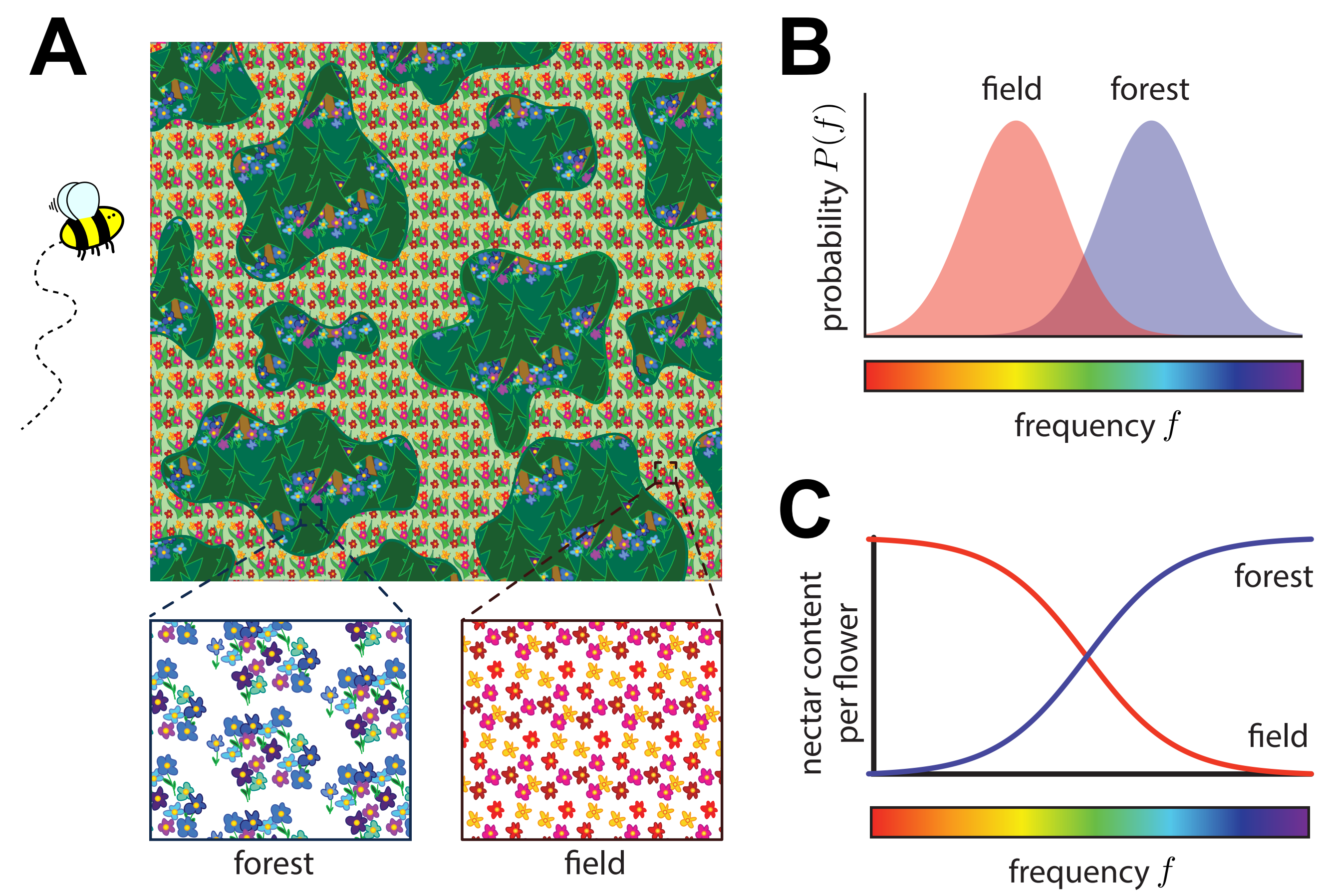
4 | action selection

do these first!

* come back to these after you've finished all 3 sections

problem set

Problem Set: Cosyne 2020



2 Sensory Coding

⋮

2.3 What would you expect . . .

2.4 Use Fig 4C to sketch . . .

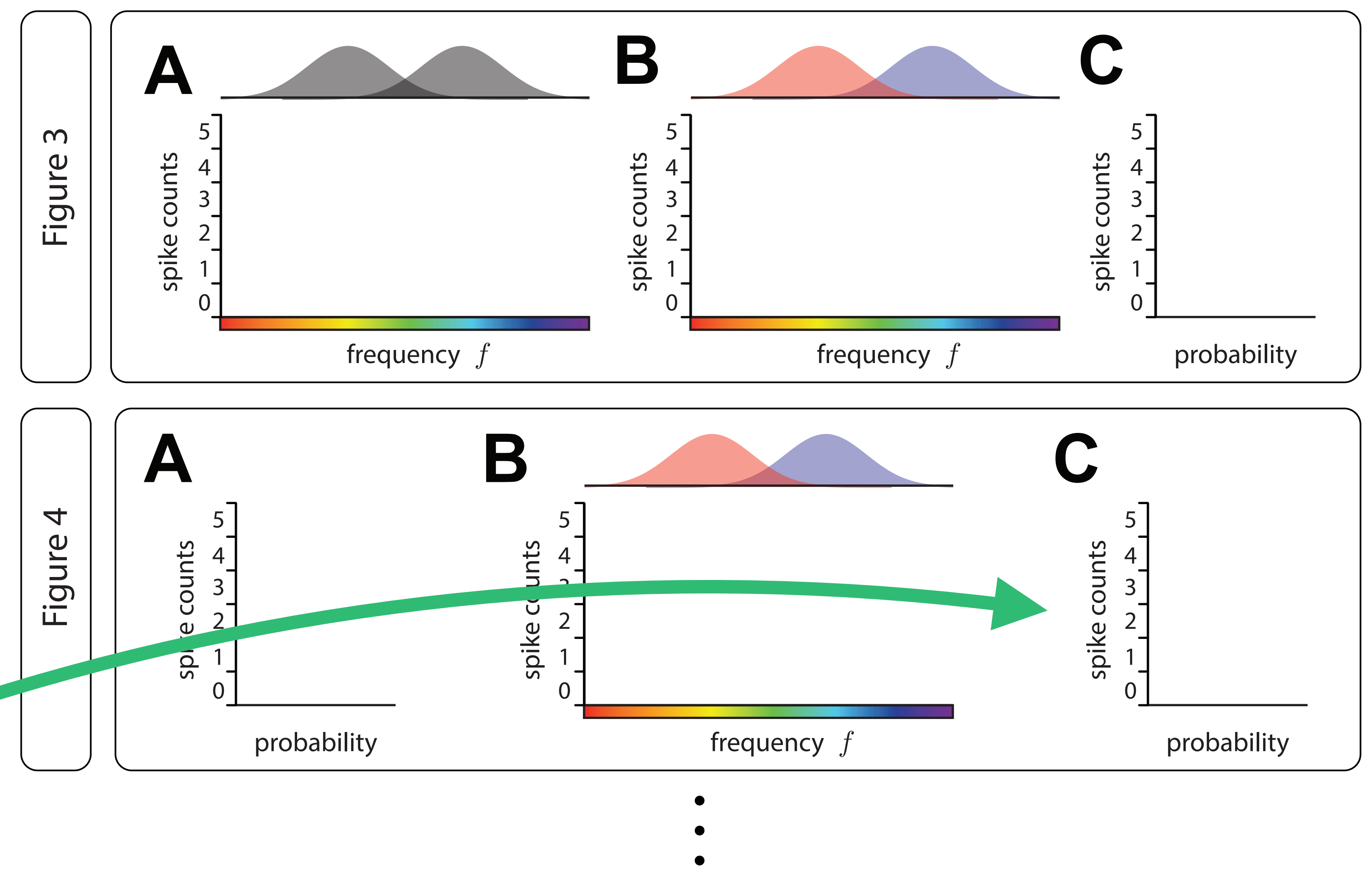
***2.5** Parameterize the nonlinearity . . .

Numerically, find the optimal . . .

worksheet

Worksheet: Cosyne 2020

Ann Hermundstad (hermundstada@janelia.hhmi.org)

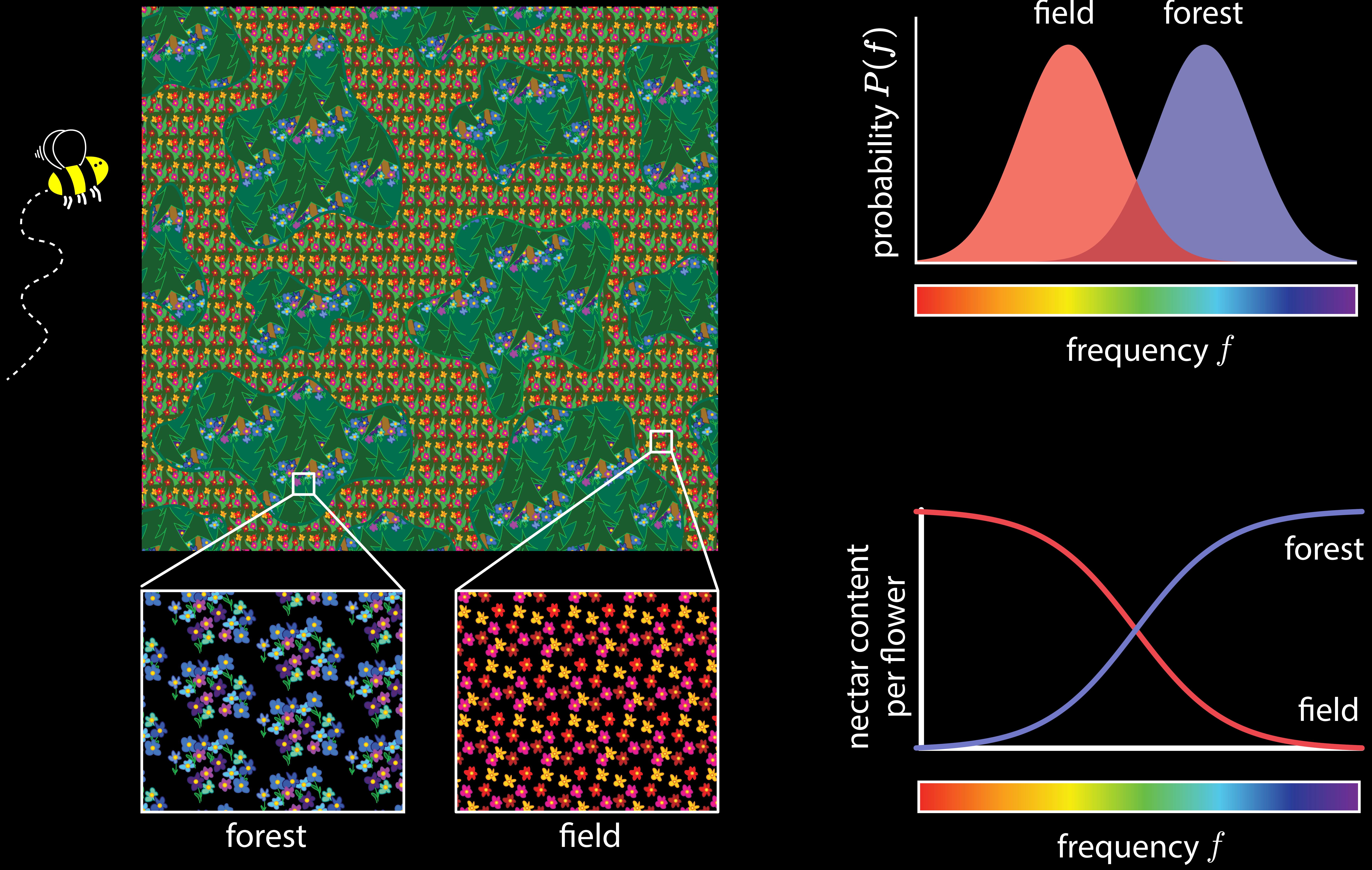


⋮

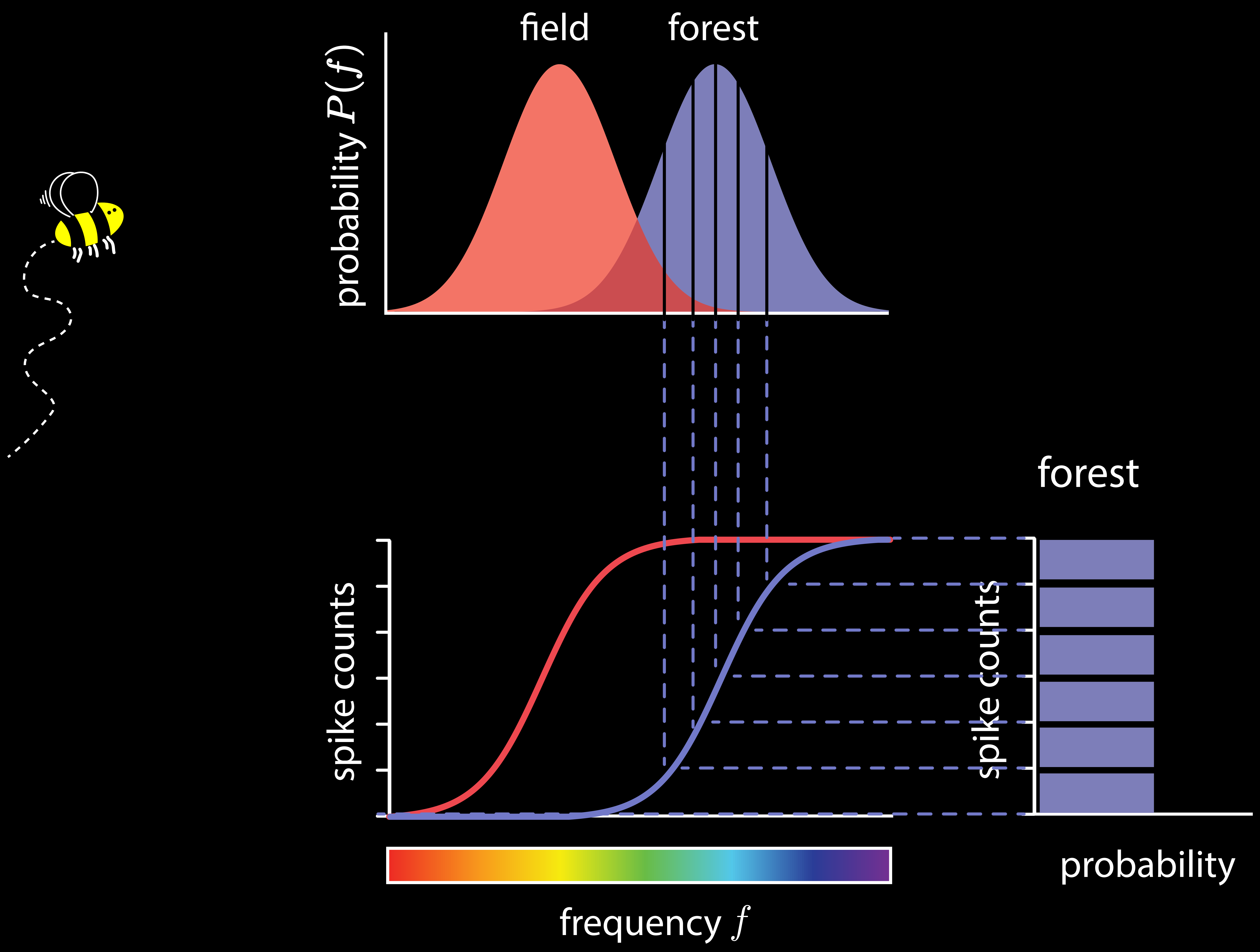
code available
(thanks to Sashank Pisupati!)

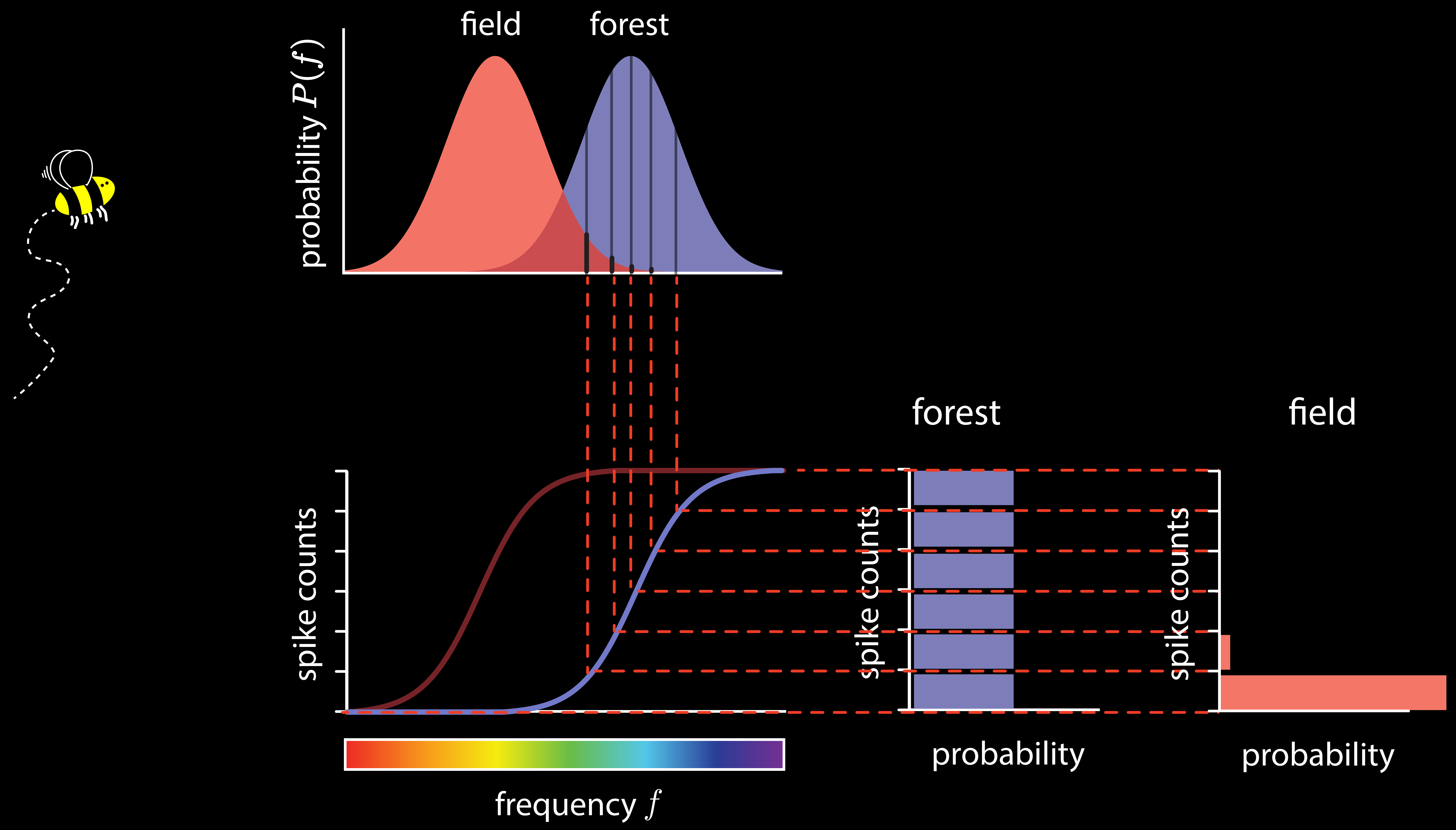
http://bit.ly/cosyne2020-tutorial

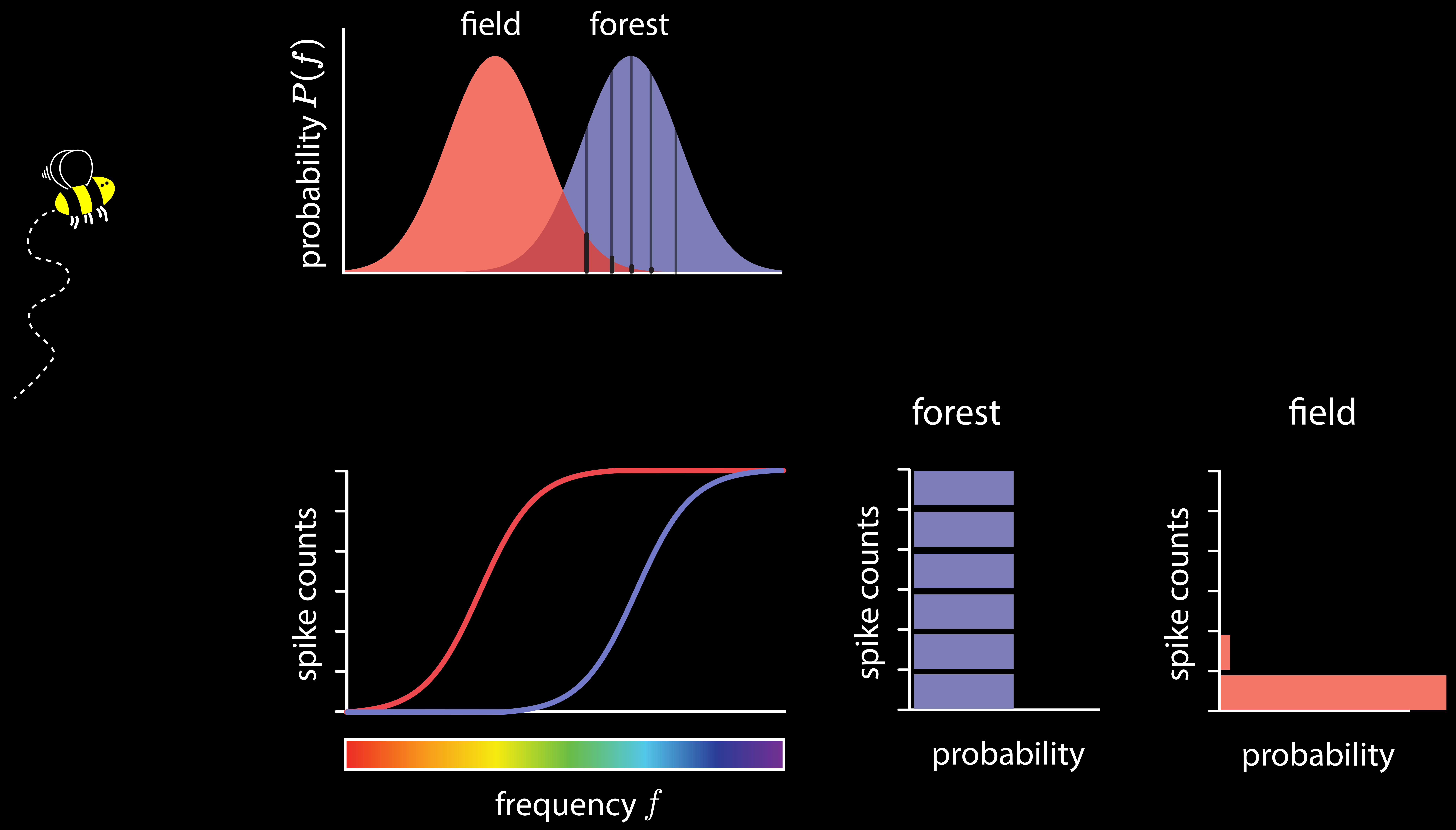
problem recap



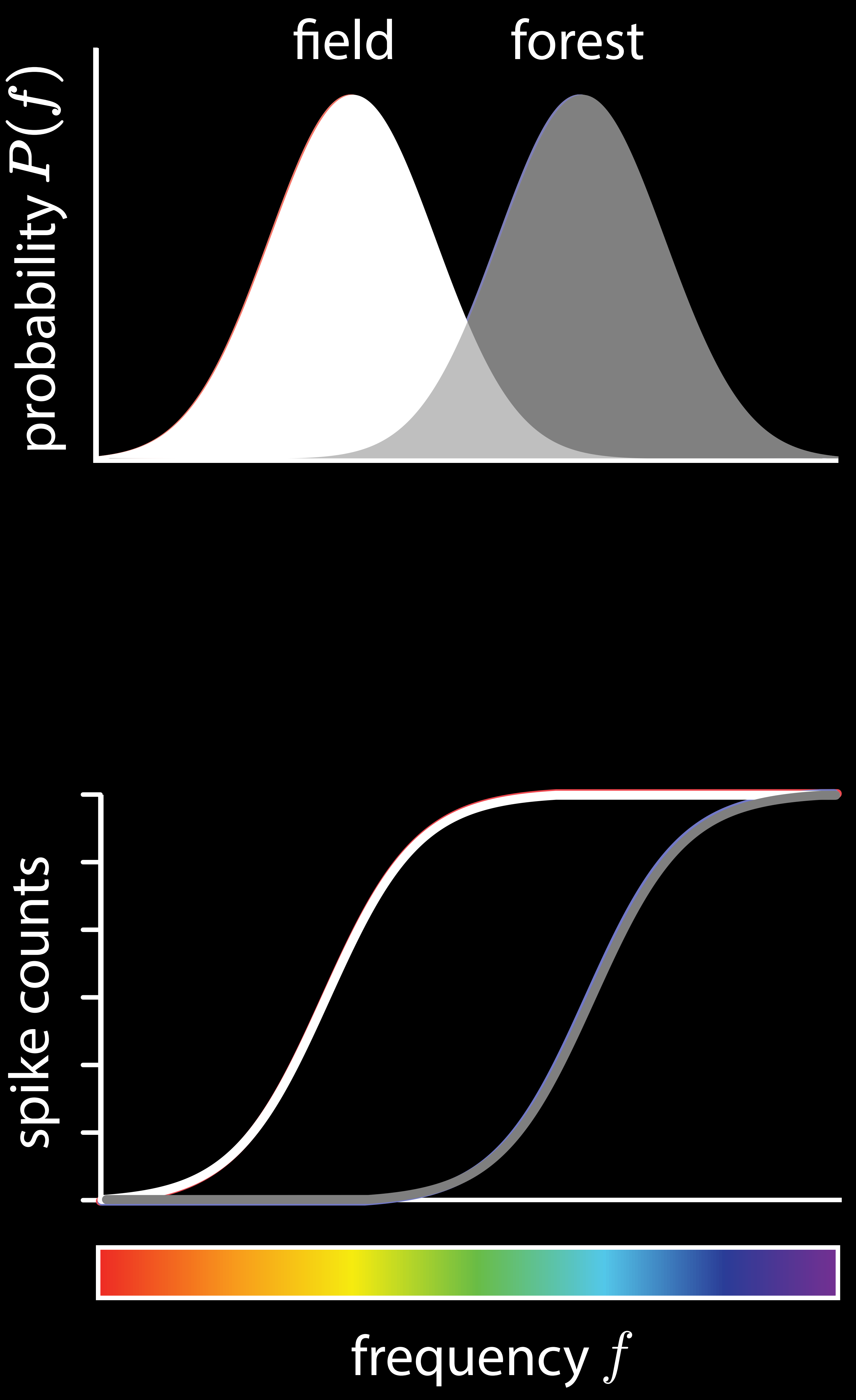
forest
field
field
forest
probability $P(f)$
frequency f
forest
field
nectar content per flower
frequency f


field
forest
probability $P(f)$
forest
spike counts
spike counts
probability
frequency f

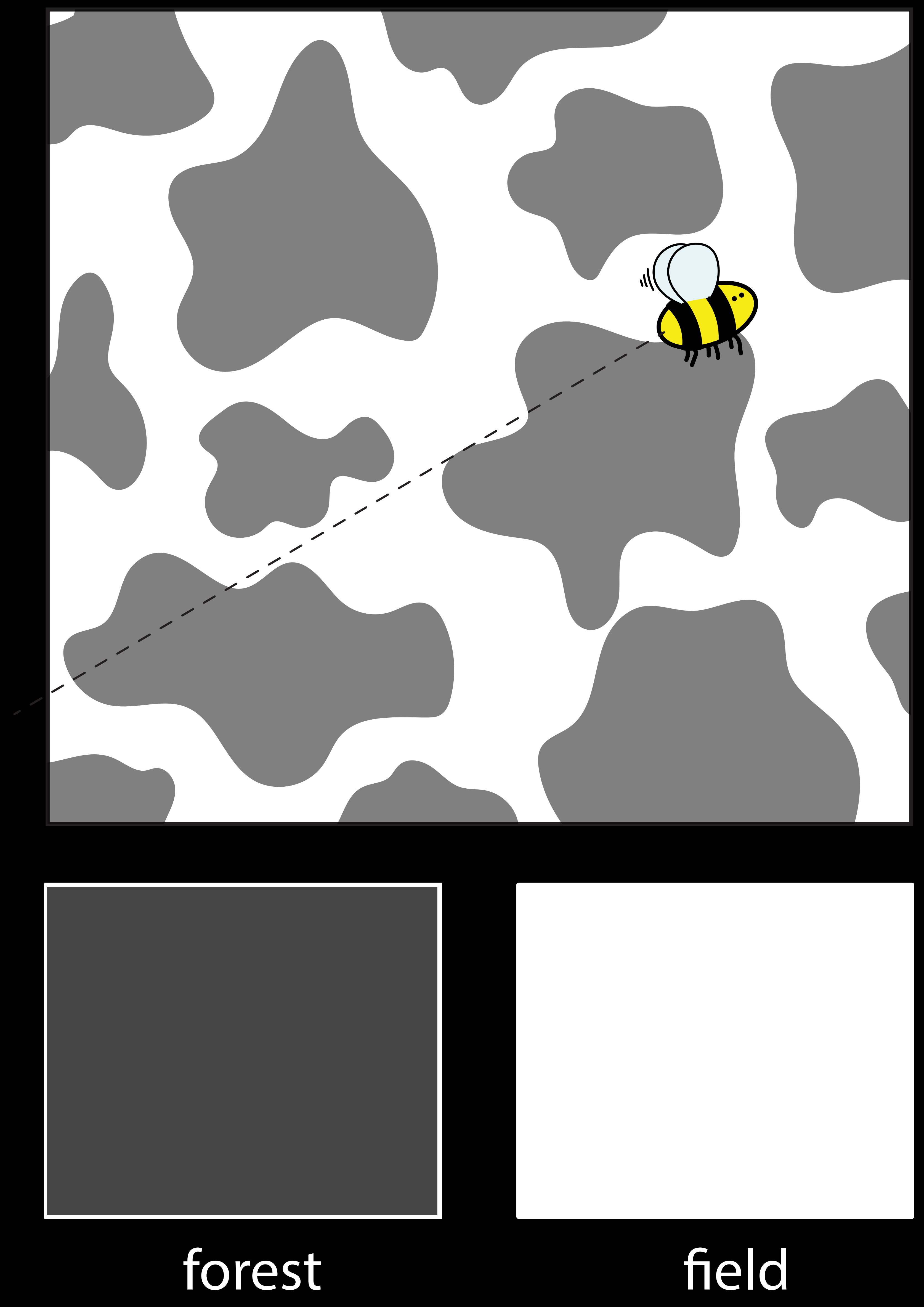

field
forest
probability $P(f)$
spike counts
frequency f
forest
field
spike counts
spike counts
probability
probability



field
forest
probability $P(f)$
spike counts
frequency f
forest
spike counts
probability
field
spike counts
probability



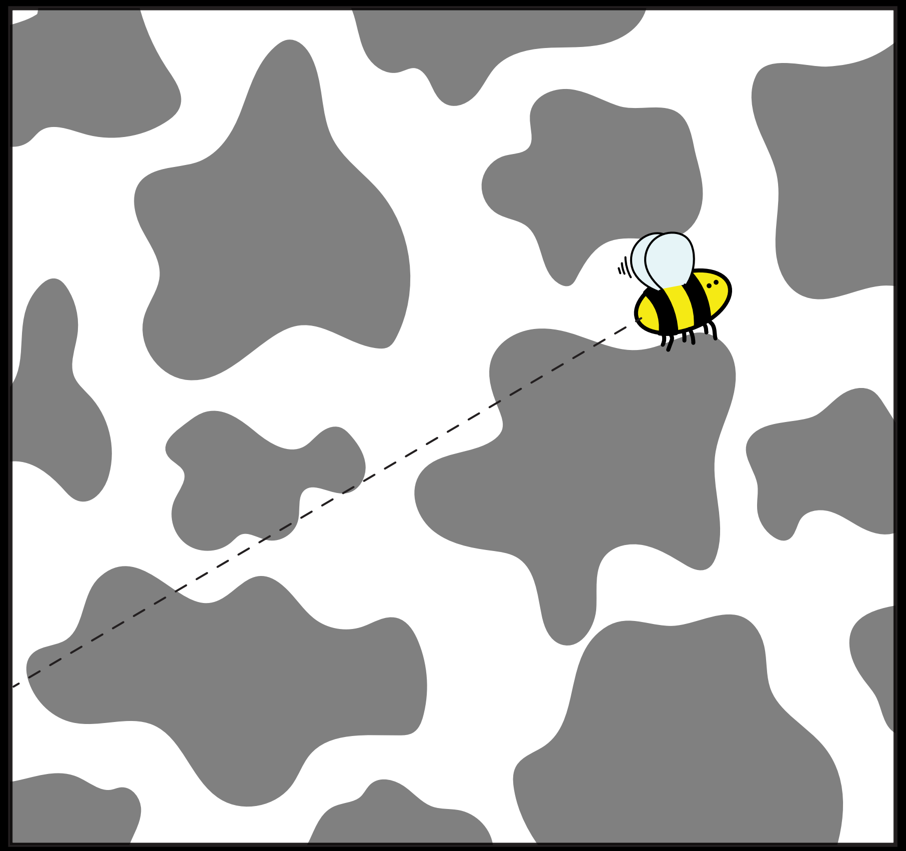
field
forest
probability $P(f)$
spike counts
frequency f



forest
field

— time →

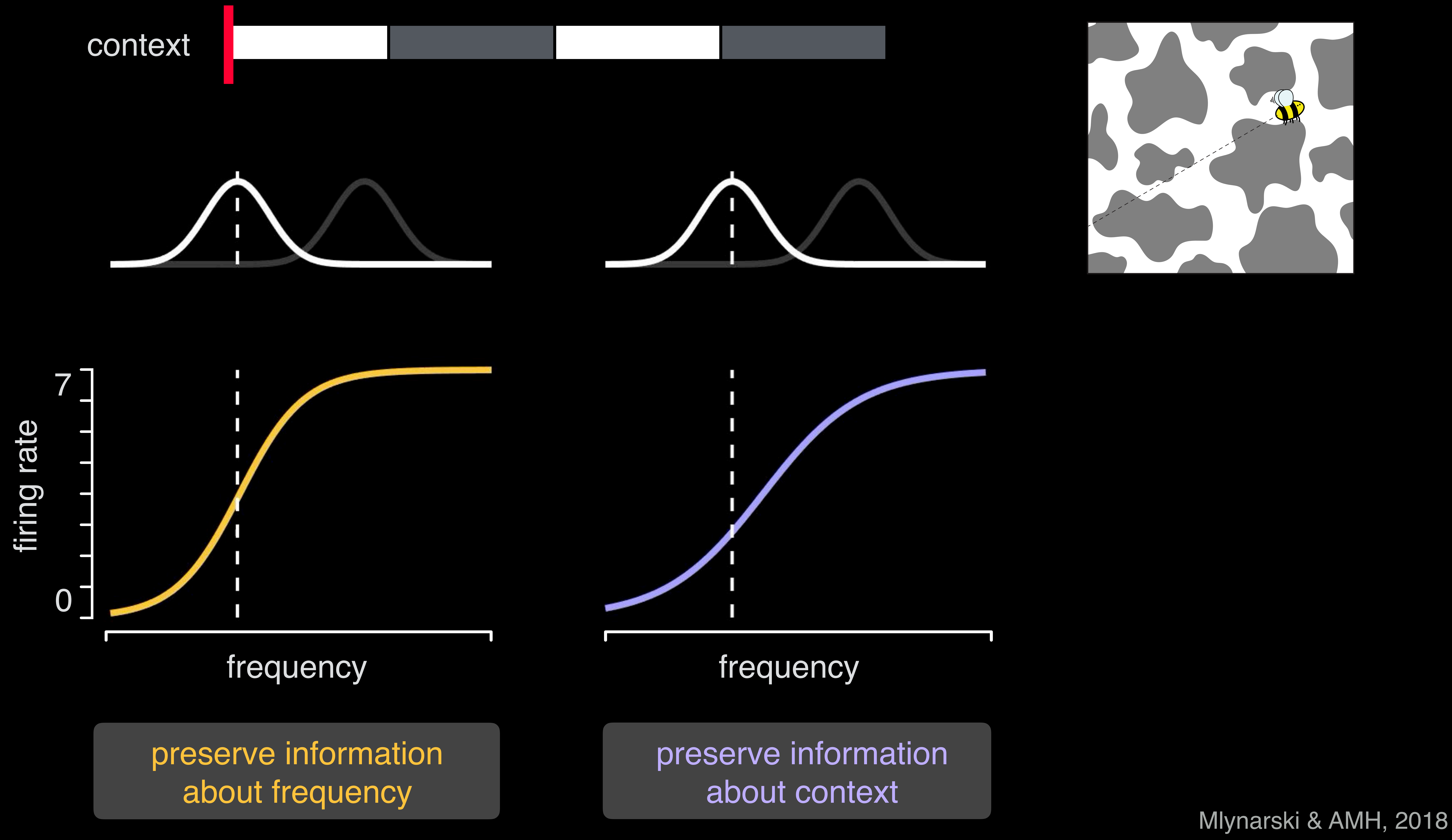
context



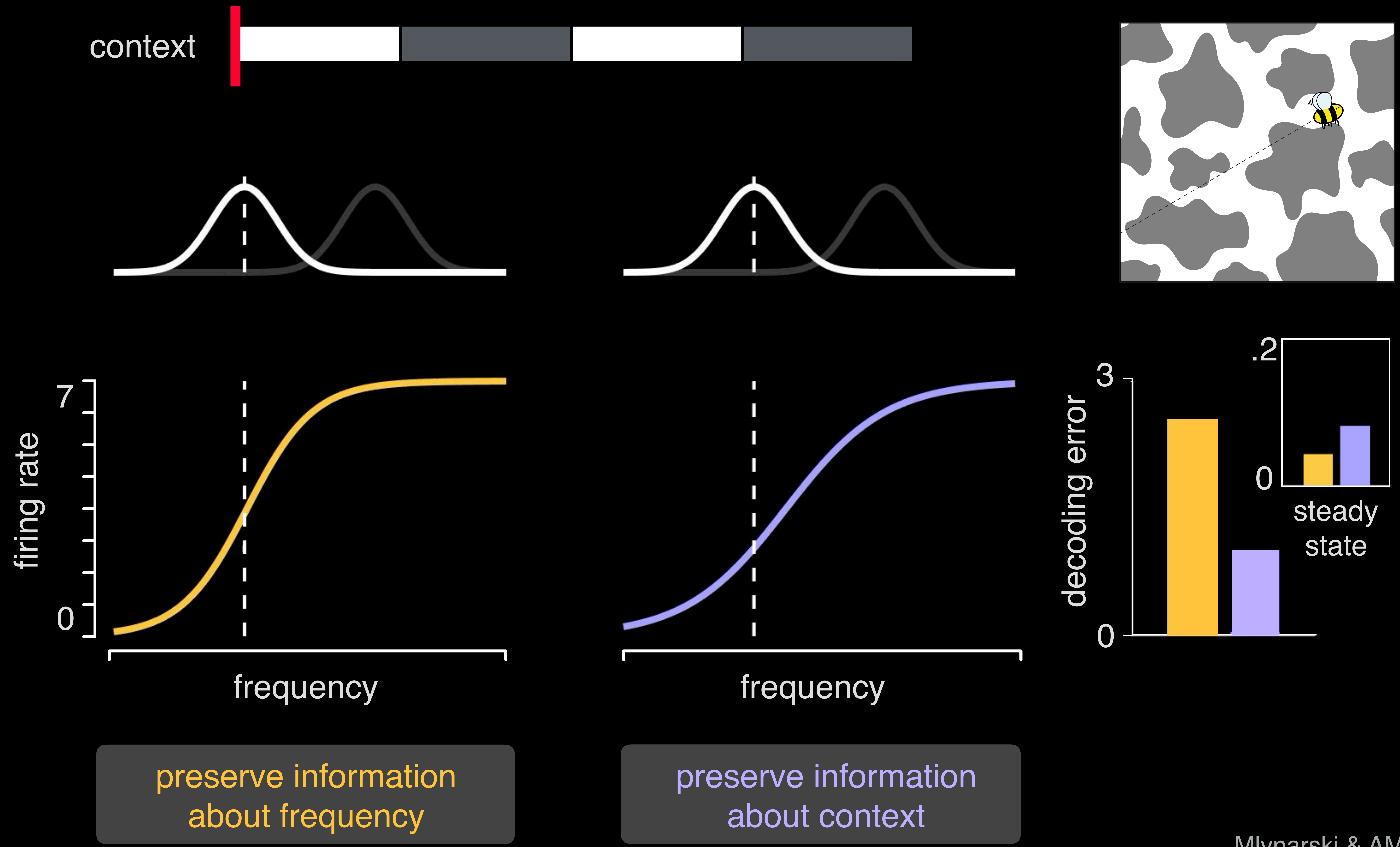
preserve information about frequency

preserve information about context

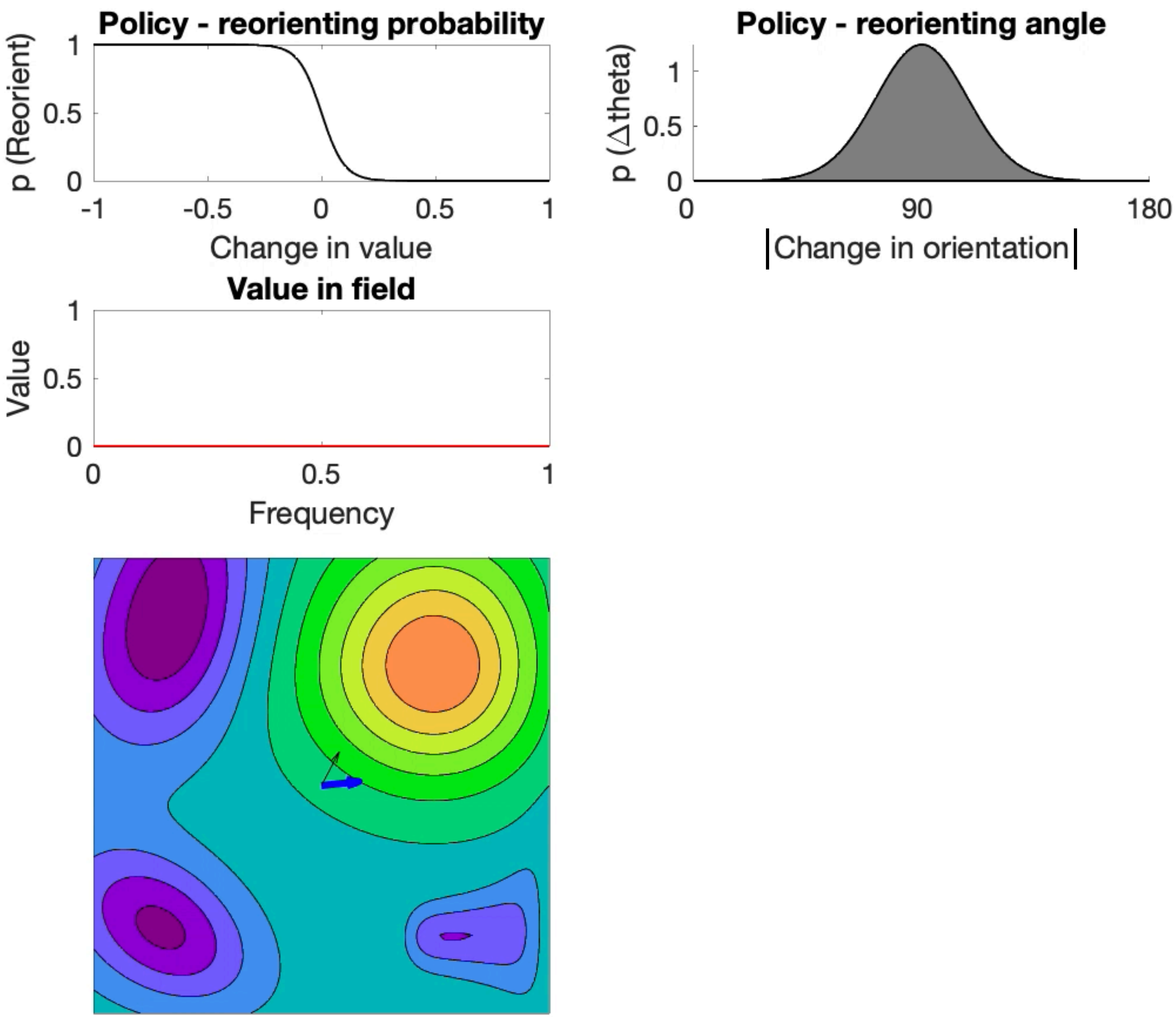
Mlynarski & AMH, 2018

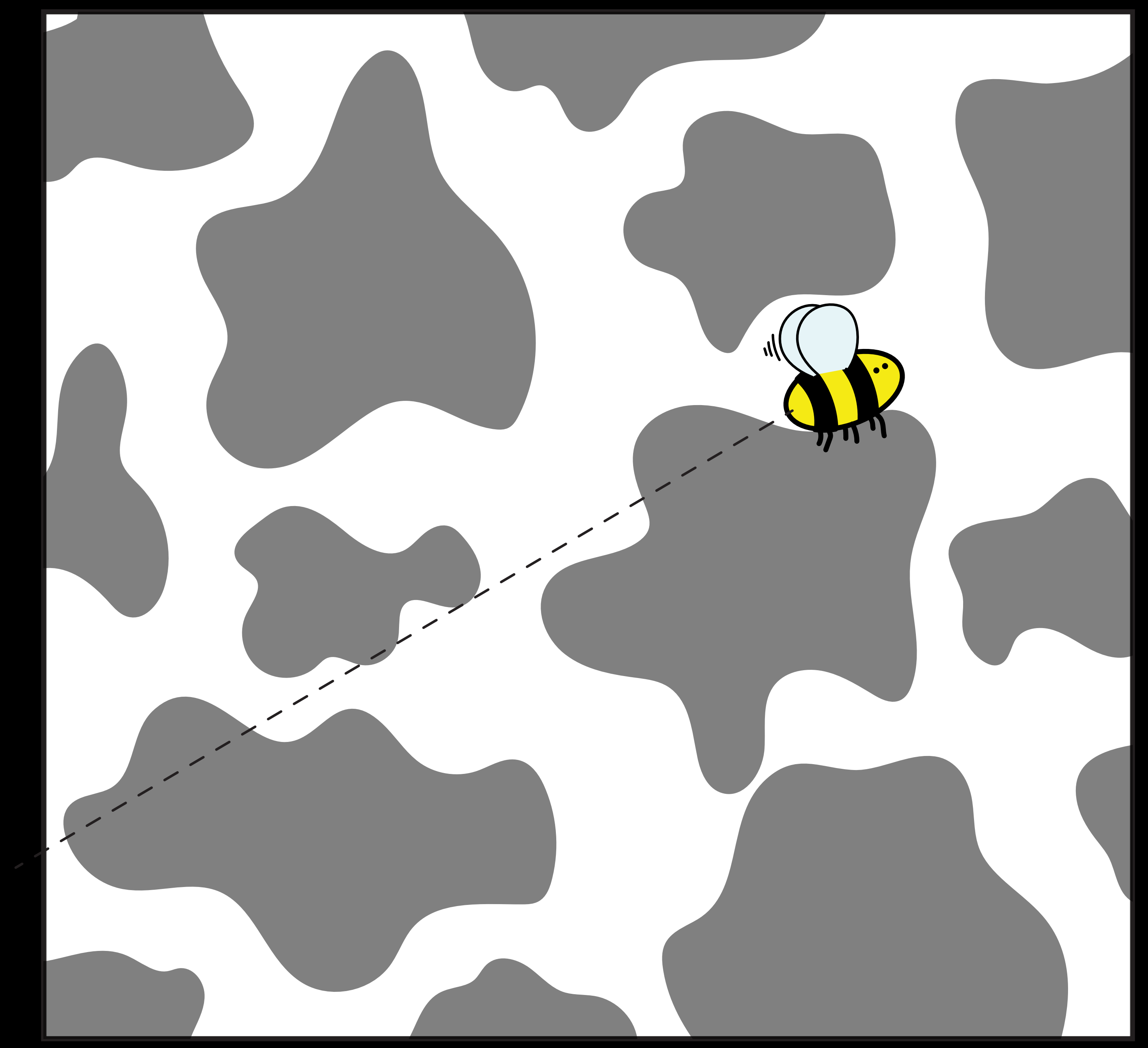


context
firing rate
7
0
frequency
frequency
preserve information about frequency
preserve information about context
Mlynarski & AMH, 2018

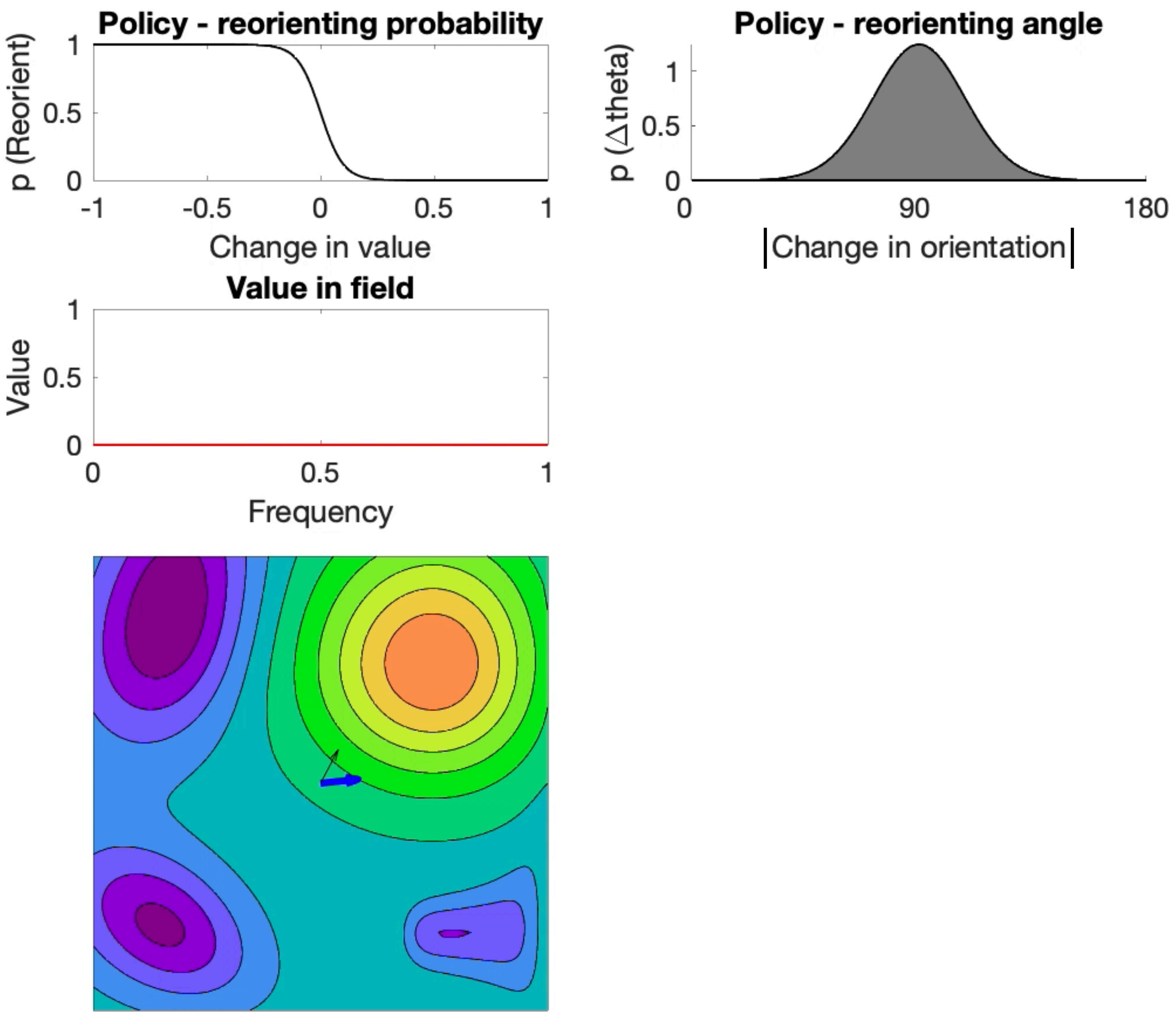


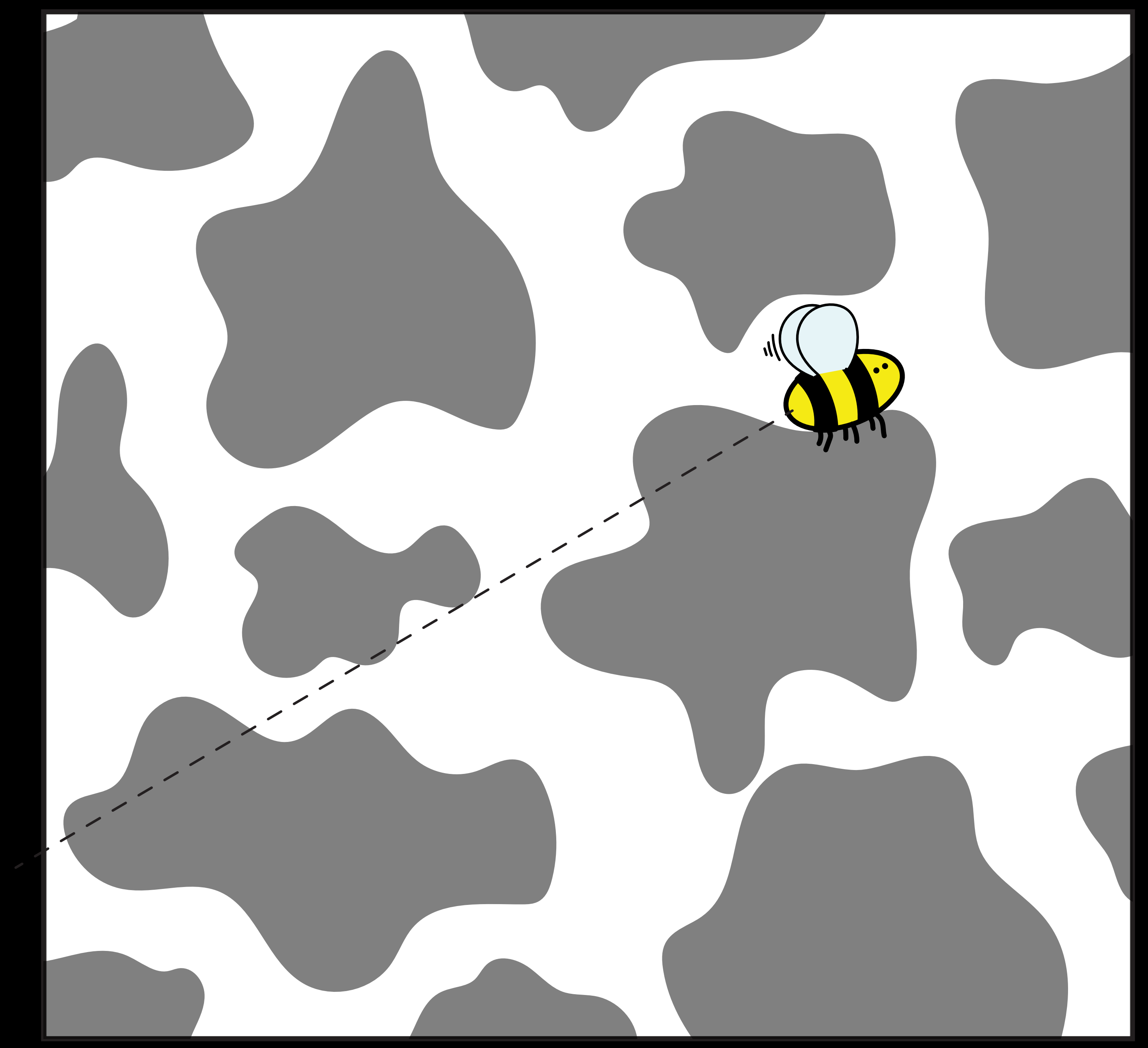
Mlynarski & AMH, 2018



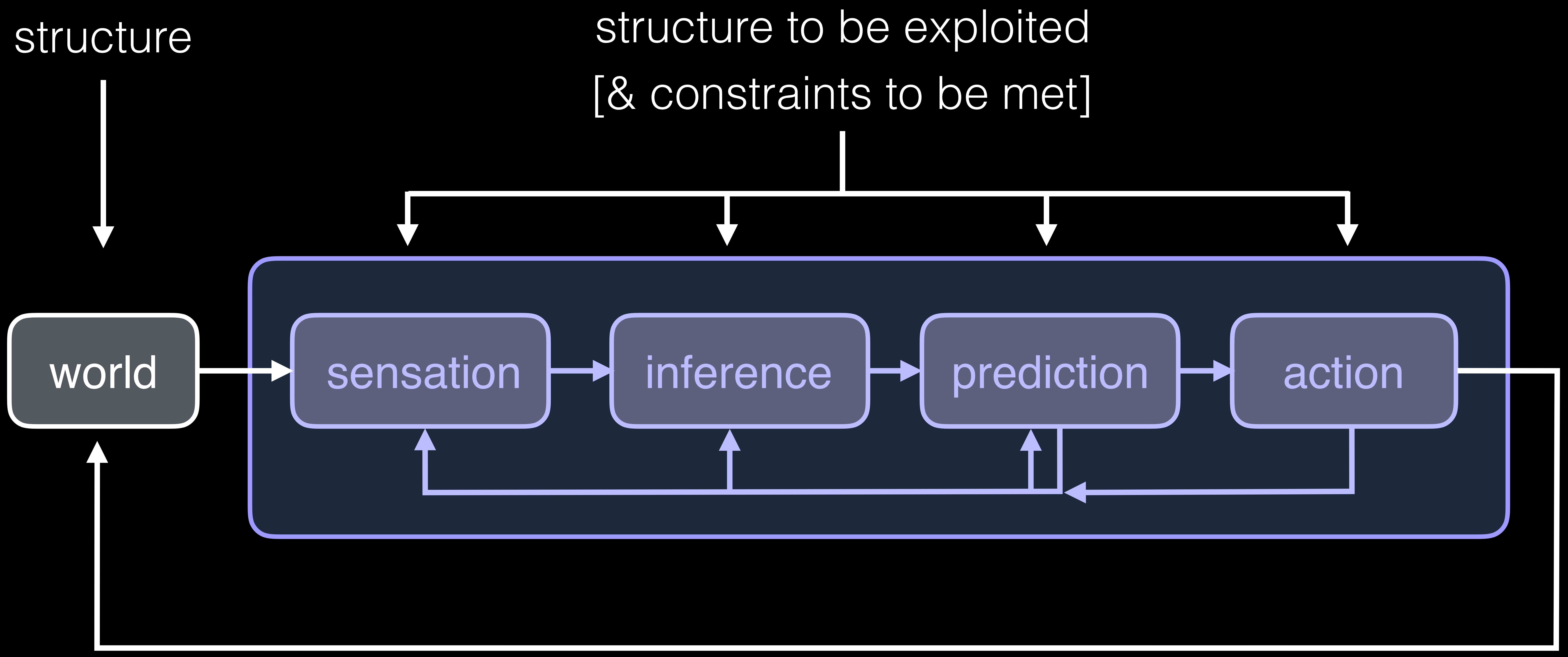


*simulations and movie
by Sashank Pisupati





*simulations and movie by Sashank Pisupati


structure
structure to be exploited
[& constraints to be met]
world
sensation
inference
prediction
action

a huge thanks to all the TAs!

Diego Arribas
Zoe Ashwood
Pierre-Etienne Fiquet
Caroline Haimerl
Anna Kutschireiter
Tzuhsuan Ma (Maz)
Jorge Menendez
Josue Nassar
Marcella Noorman
Sashank Pisupati
Satpreet Singh
Charline Tessereau