Bayesian Computing in Biology

Jacques Droulez
Institut des Systèmes Intelligents et de Robotique
CNRS - UPMC, Paris

Deep Blue beats Garry Kasparov (1997)

LOGIC WORLD ≠ REAL WORLD

Computers outperform human in all logical & arithmetic operations.

Living organisms outperform computers and robots in all tasks involving uncertainty, *e.g.* action & perception in the real world.

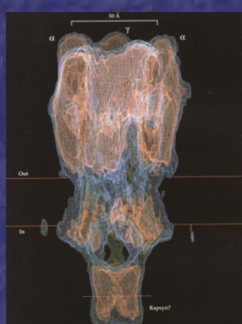
A difference exploited in the « captcha » tests.

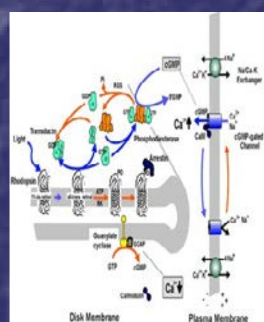
Bayesian Computing in Biology

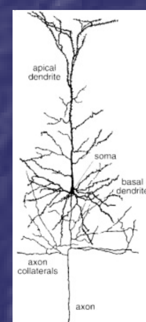
To be understood at very different space (and time) scales

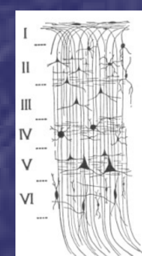
Space Scale

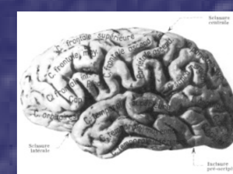
Single molecule | Systems biology | Neuron | Column | Brain | Behavior











OVERVIEW

Space Scale

Single molecule

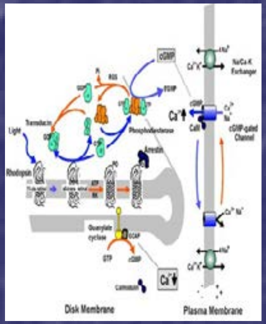
Systems biology

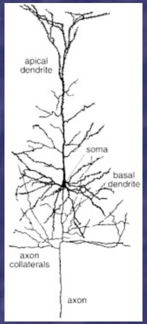
Neuron

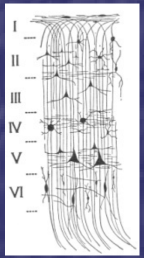
Column

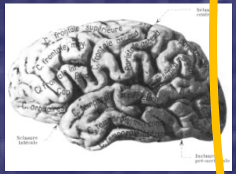
Brain

Behavior







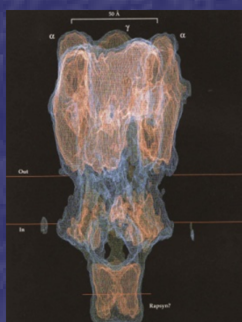


Ψ

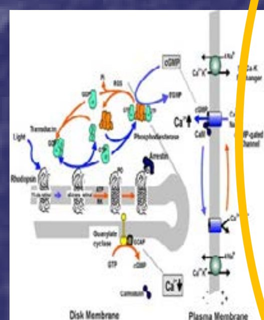
OVERVIEW

Space Scale

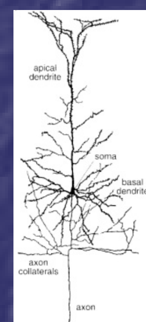
Single molecule



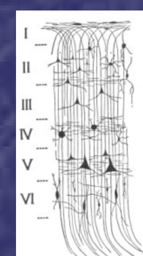
Systems biology



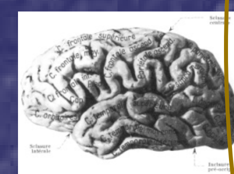
Neuron



Column



Brain

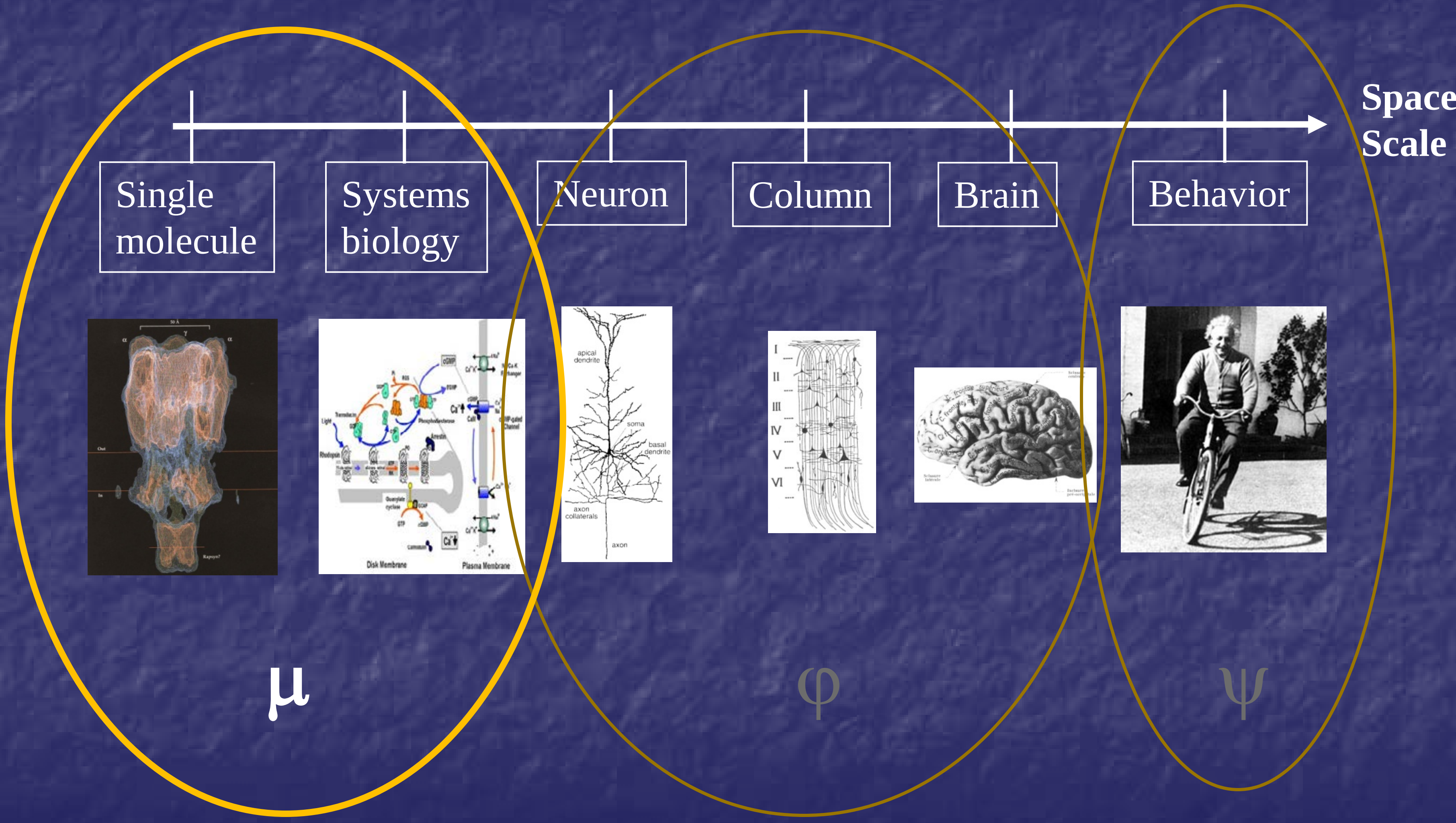


Behavior

φ

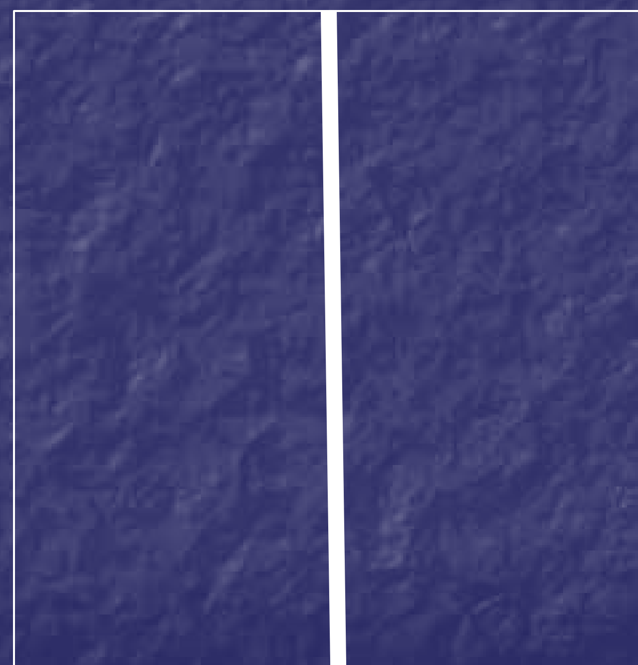
ψ

OVERVIEW

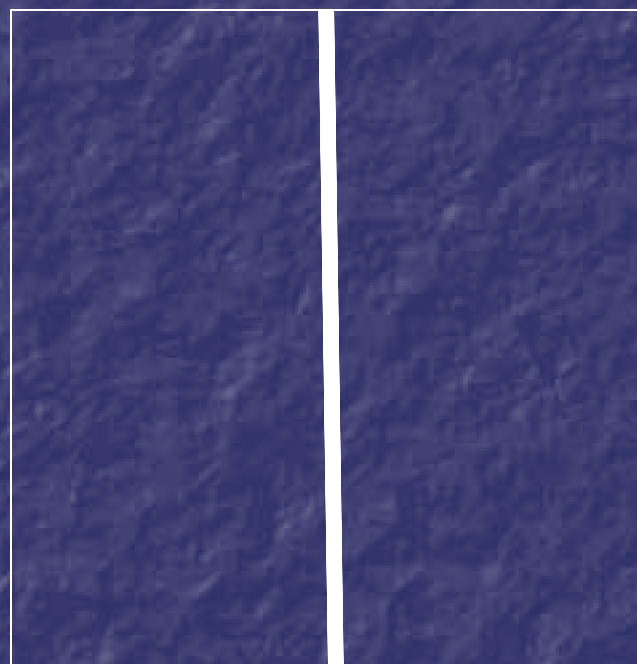

Space Scale
Single molecule
Systems biology
Neuron
Column
Brain
Behavior
apical dendrite
soma
basal dendrite
axon collaterals
axon
I
II
III
IV
V
VI
Disk Membrane
Plasma Membrane
μ
φ
ψ

Part 1: Perception as Bayesian inference: an old idea …

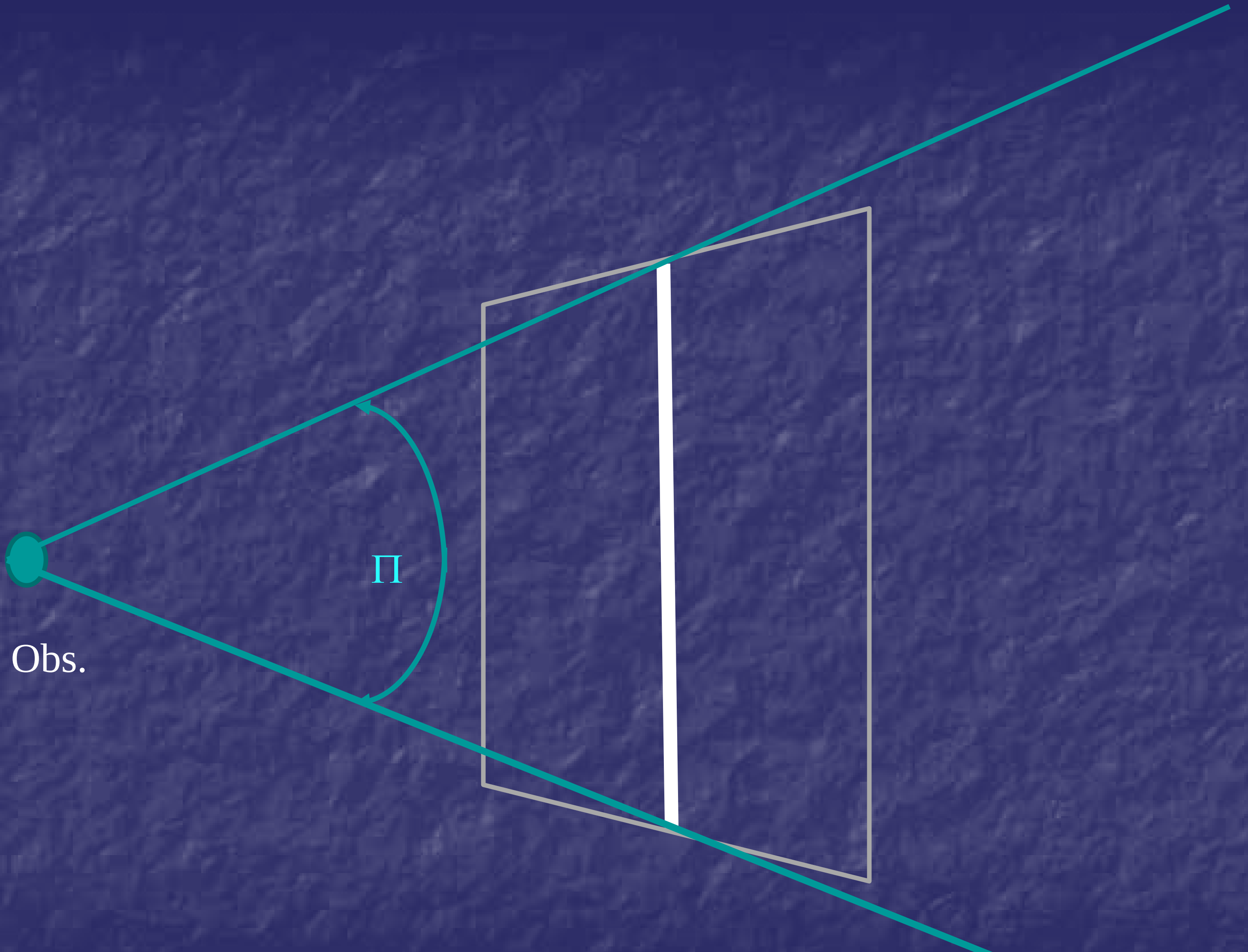
H. Helmholtz (1867), E. Mach (1897), …
Knill & Richards (1996), Kersten, Mamassian & Yuille (2004), …



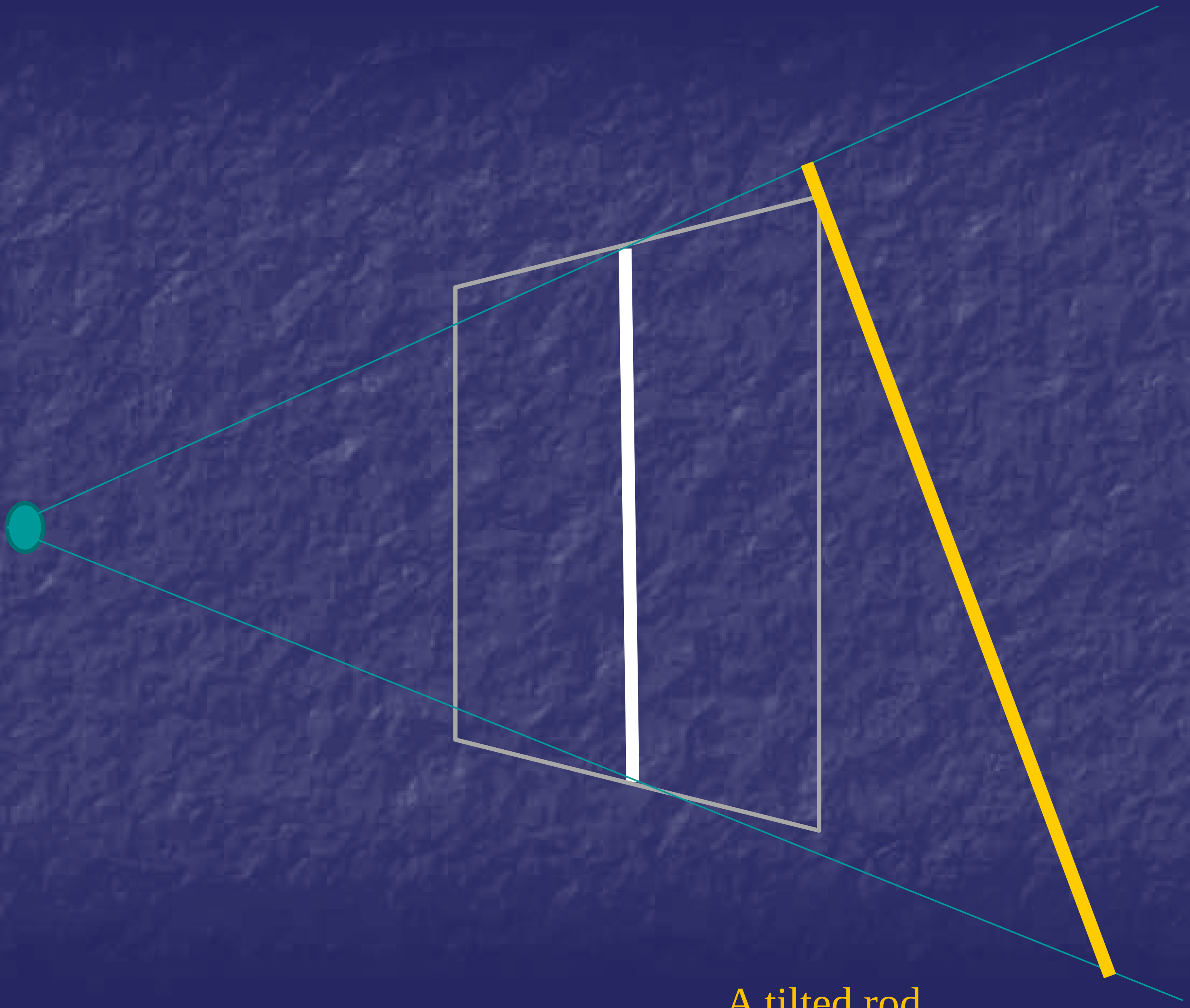
Here, an example from Ernst Mach, *The Analysis of Sensations* (1897)



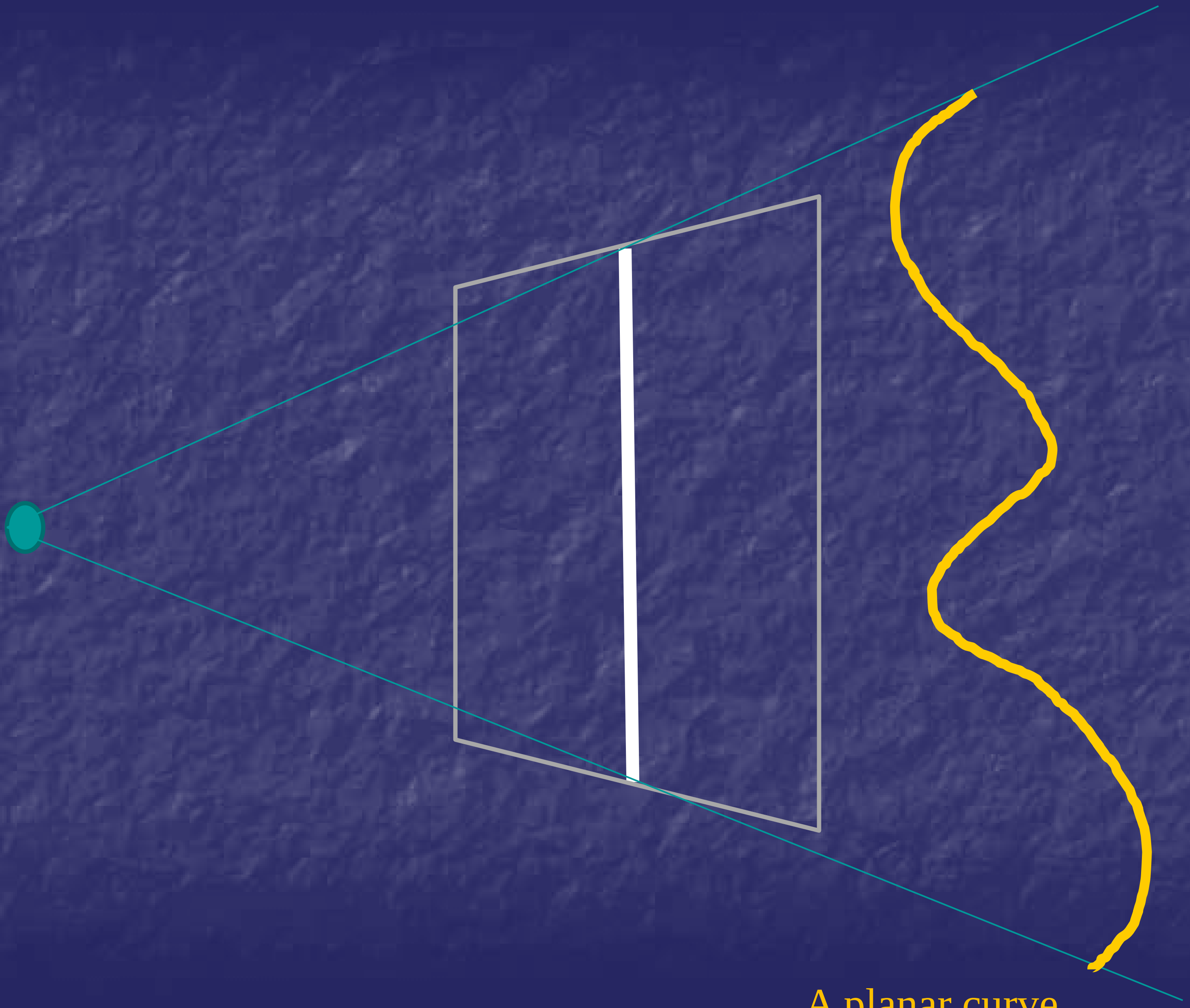
A vertical line in the image ⇒ A vertical rod in space ?



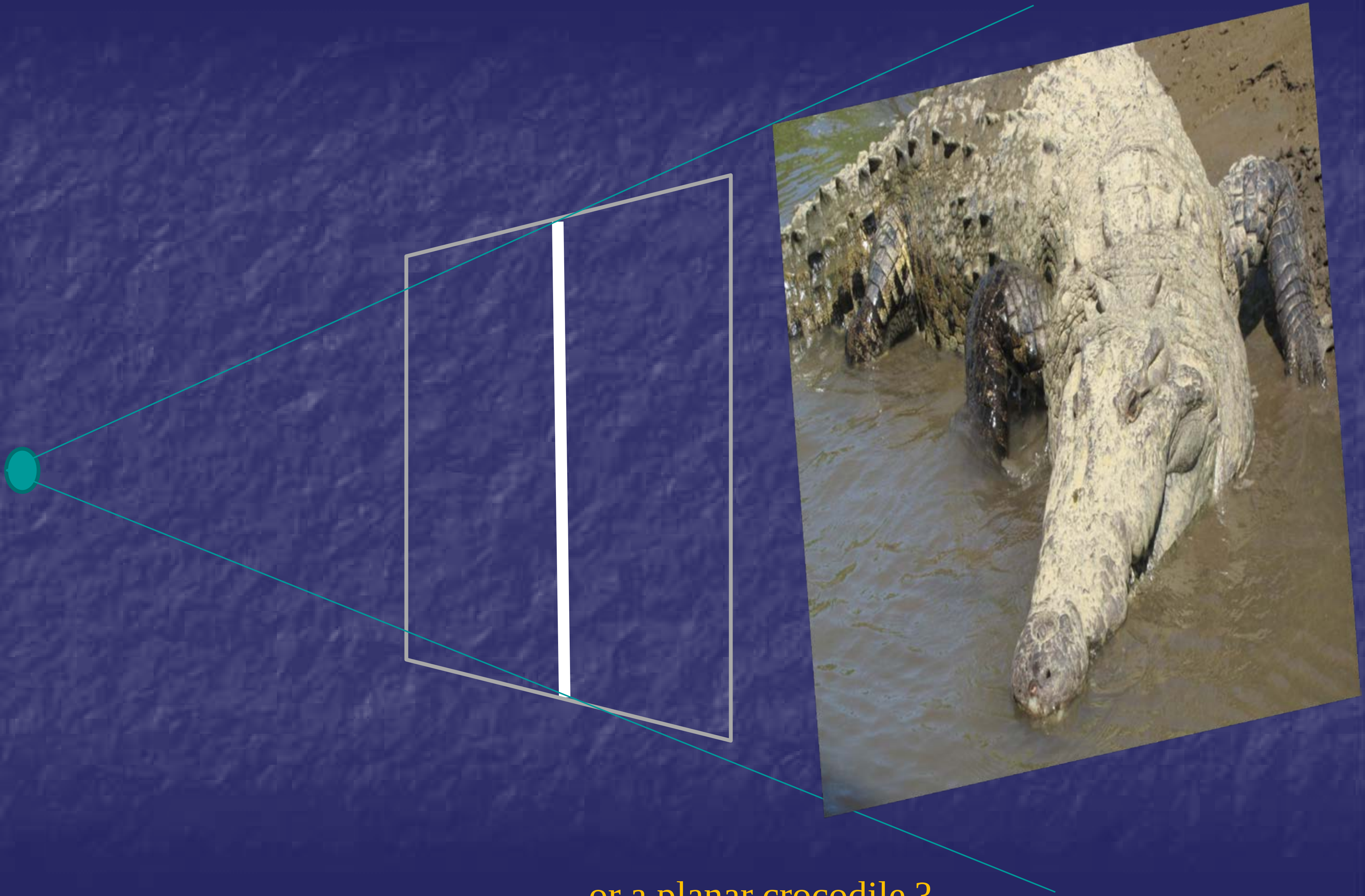
A vertical line in the image ⇒ Any object in space contained in the plane Π



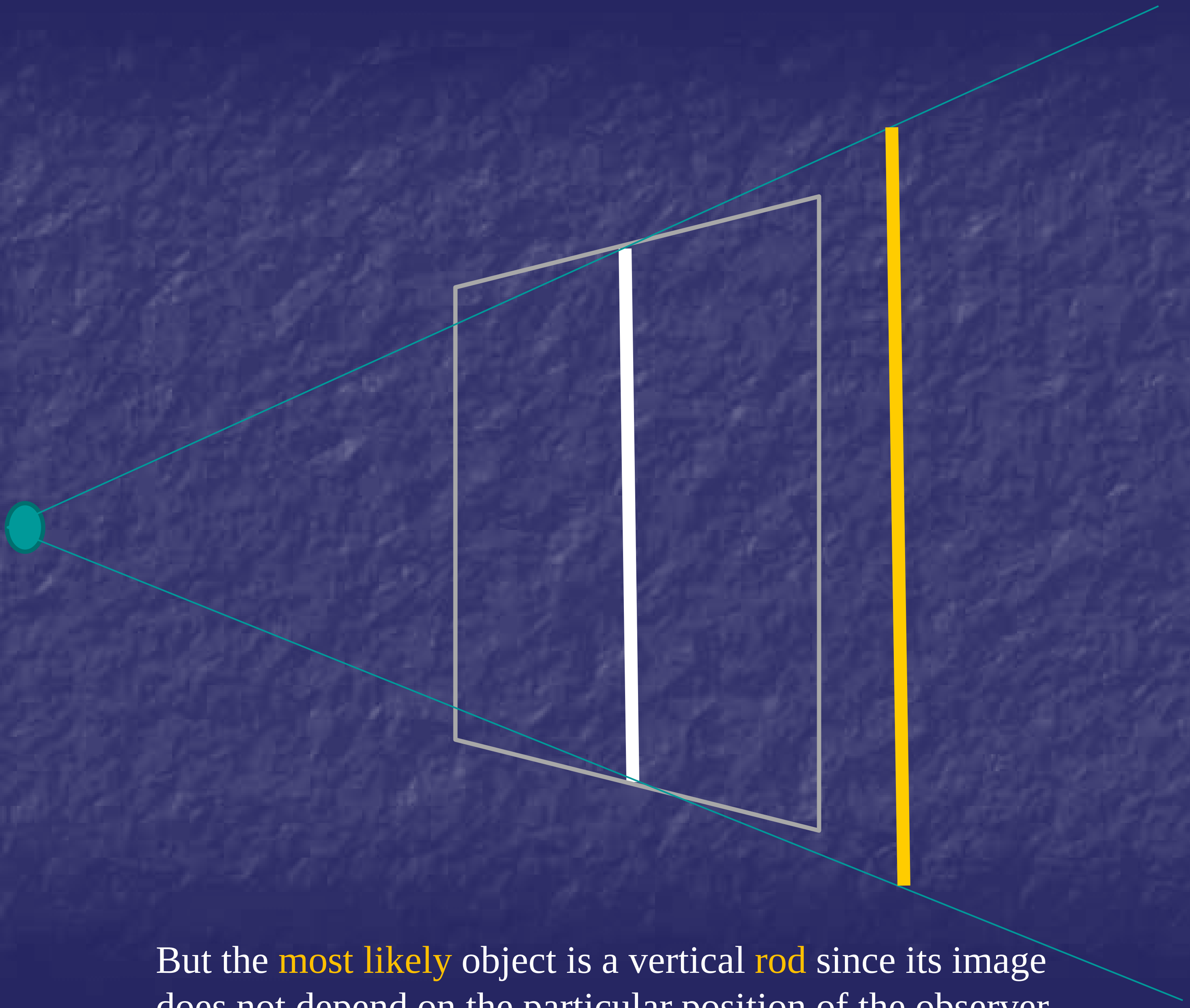
A tilted rod …



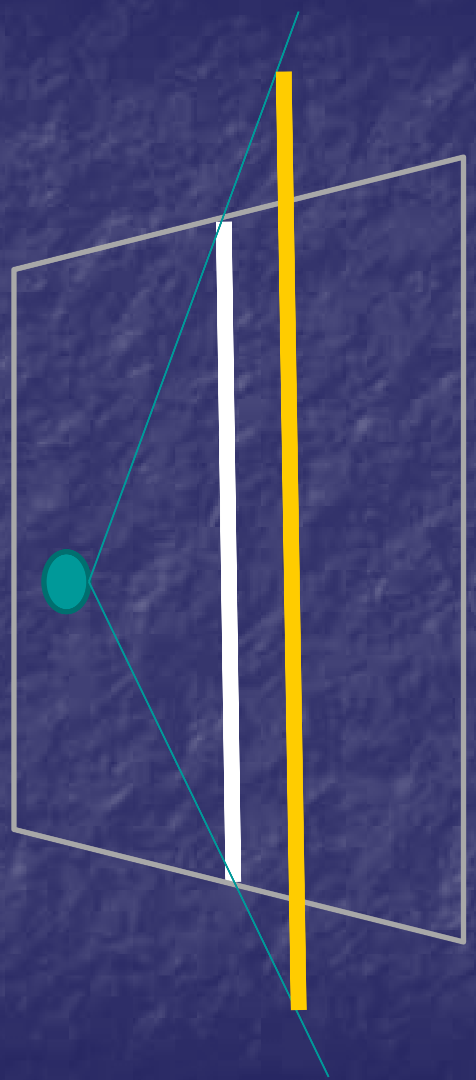
A planar curve



or a planar crocodile ?

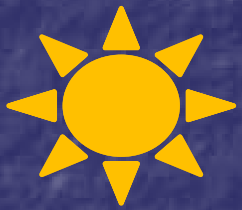
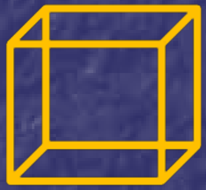


But the most likely object is a vertical rod since its image does not depend on the particular position of the observer.



High P(O | S): We do not believe in coincidences !

The Bayesian approach: priors, likelihood and free variables

Perceived Object or State S

Light source F

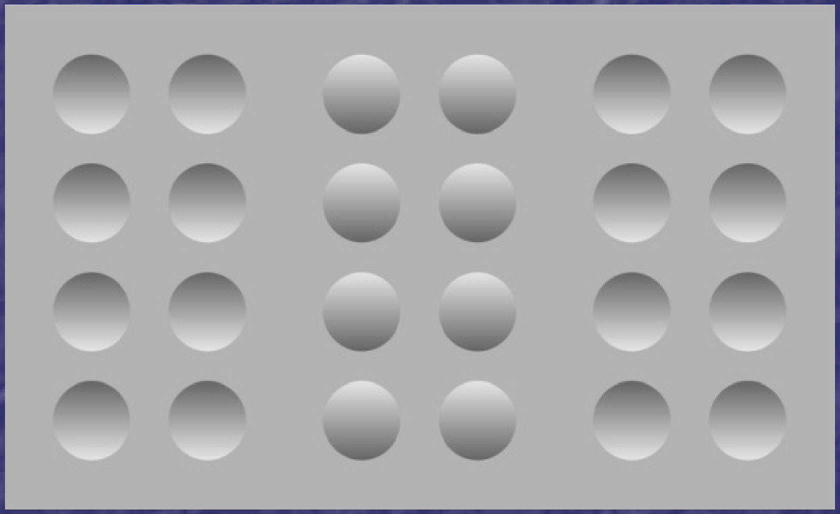
Sensory Observation O

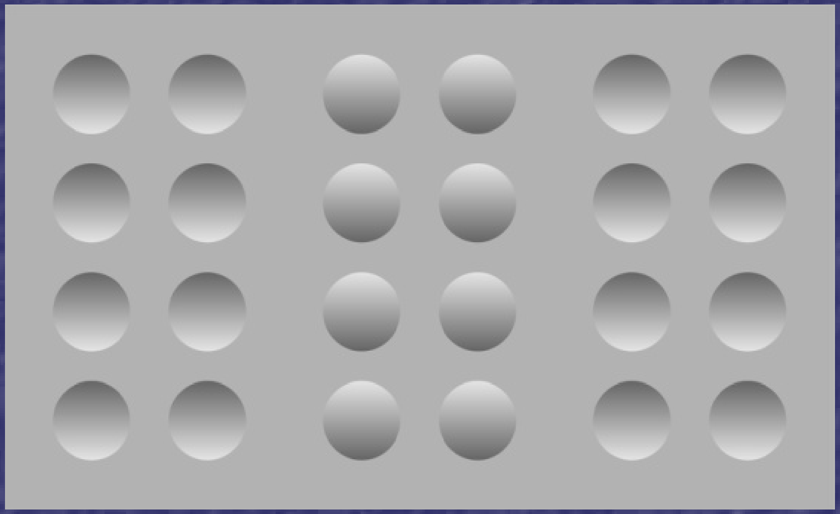
Priors P(S), P(F)
(symmetry, regularity, …)

Likelihood P(O | S, F)
(The sensor function)

Posterior $P(S \mid O) \propto P(S)\ \Sigma_F\ P(F).P(O \mid S, F)$
(Bayes' rule and marginalization rule)

3D Shape from shadow

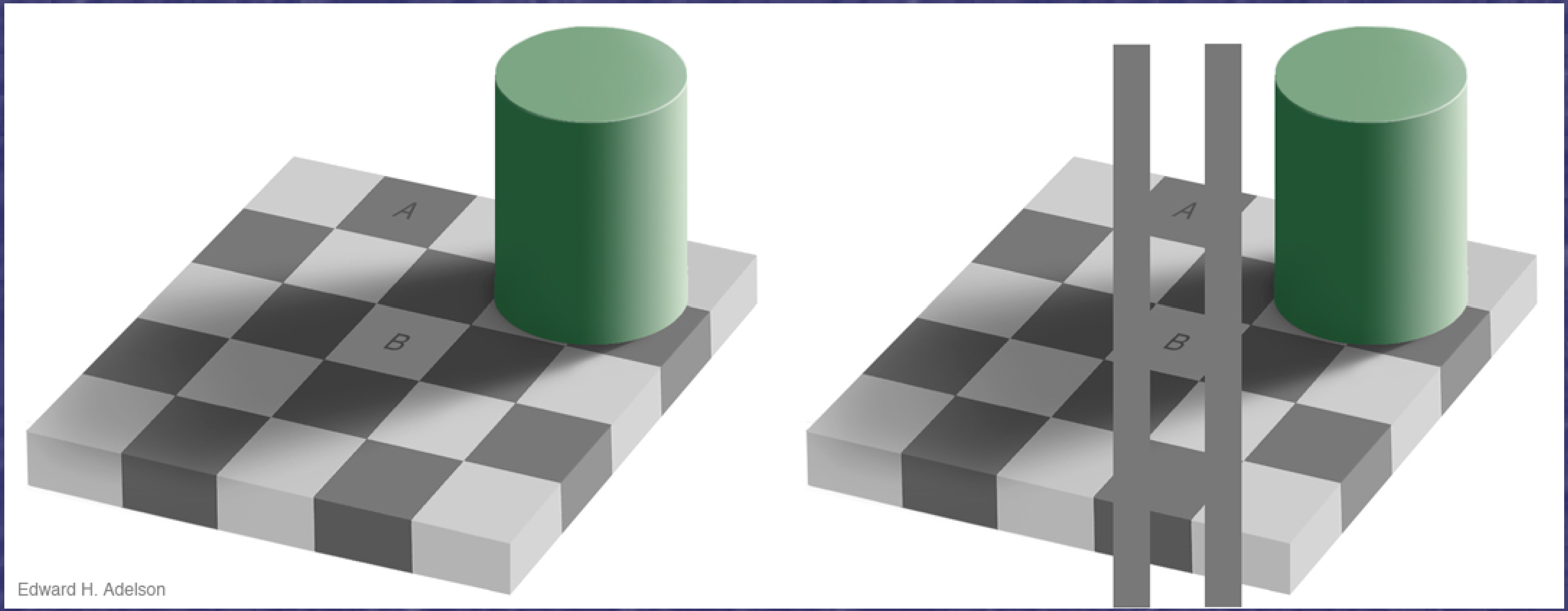




A priori, the light comes from above (The sun !): the shading is interpreted as « hollows » (if the dark part is above) or « bumps » (if the dark part is below).

Mamassian & Goutcher (2001) Prior knowledge on the illumination position. Cognition 81: B1-9

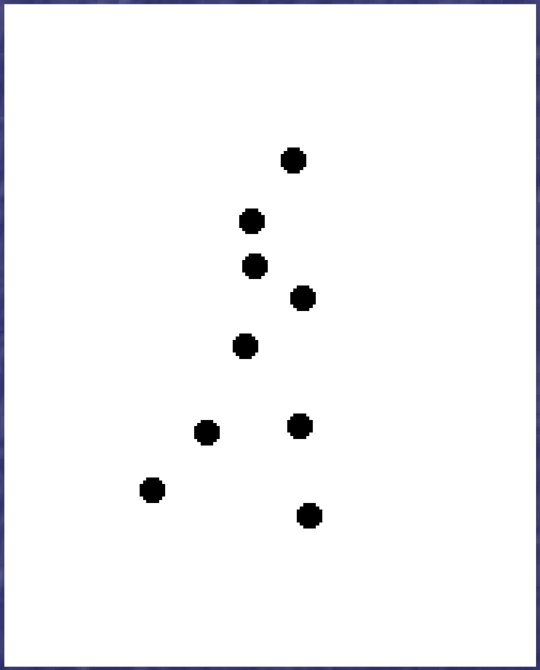
Whiteness from 3D structure



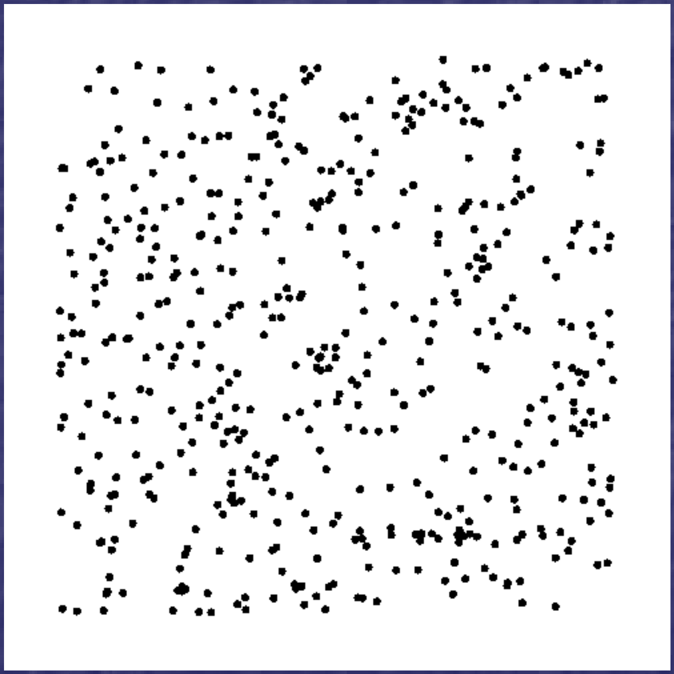
Zone B (shadowed by the green cylinder) seems whiter than zone A (unshadowed). However, both zones have the same objective luminous intensity (see right panel).

Adelson & Pentland (1996) The perception of shading and reflectance. In: Perception as Bayesian Inference (Knill & Richards, eds.) Cambridge University Press.

Priors on object shape (e.g. human) or object property (e.g. rigidity) allows to complete the otherwise undetermined visual information …

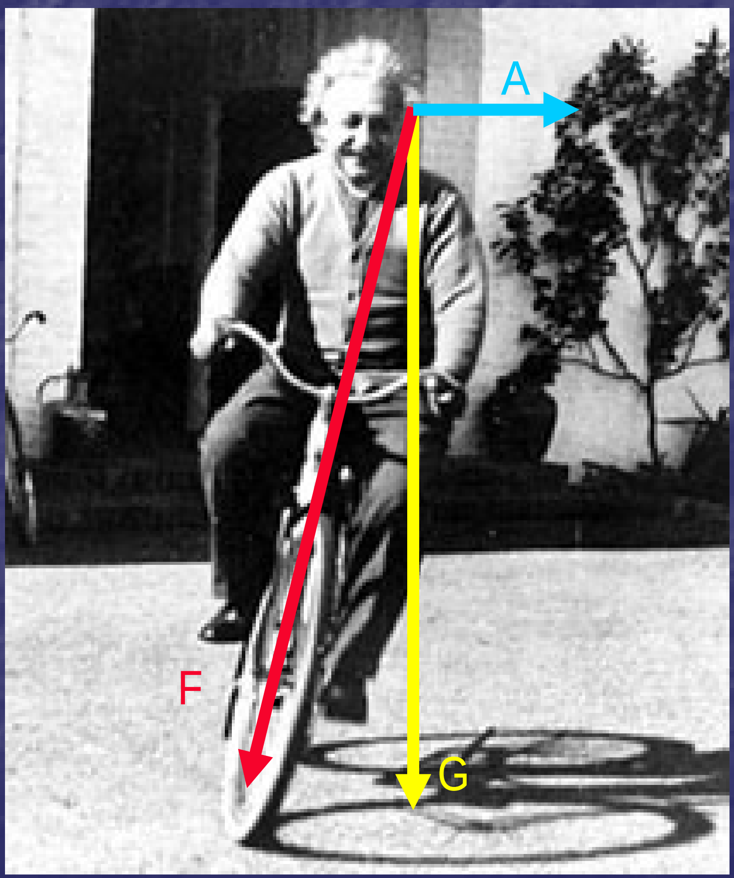


Johansson G (1973) *Perception and Psychophysics* 14:201-211

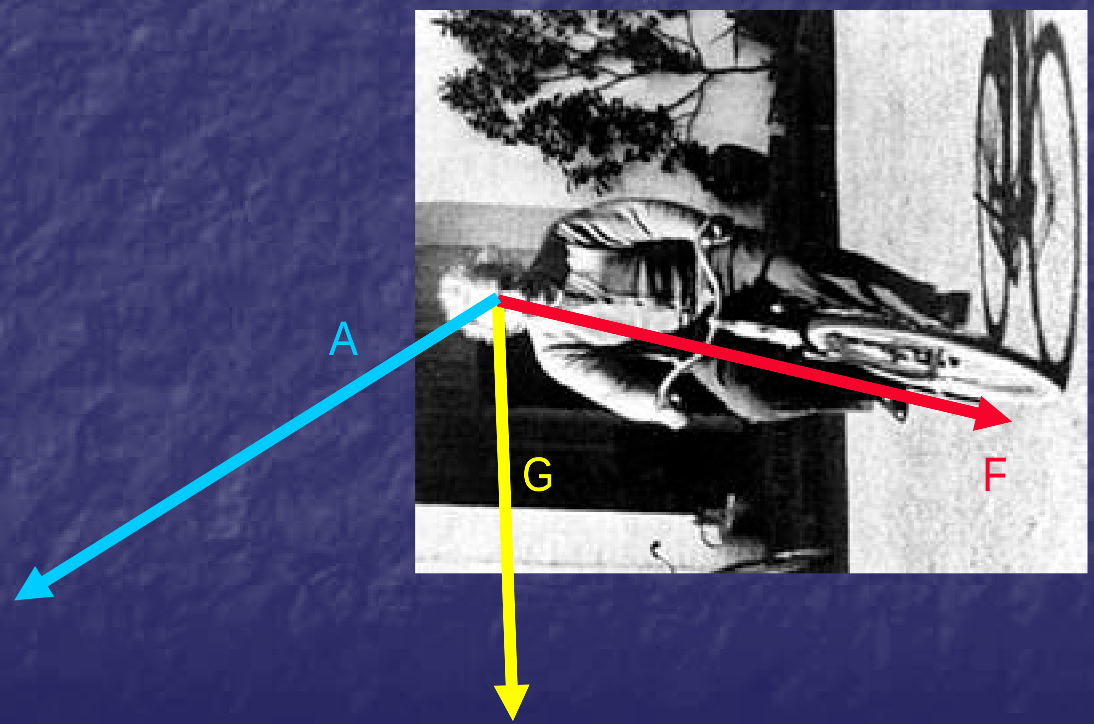


Wallach H & O'Connell DN (1953)
J. of Experimental Psychology 45(5):205-217

To solve the gravito-inertia ambiguity (F given by the vestibular sensors could result from an infinite number of combinations of gravity G and linear acceleration A), the brain uses prior favoring minimal linear acceleration.

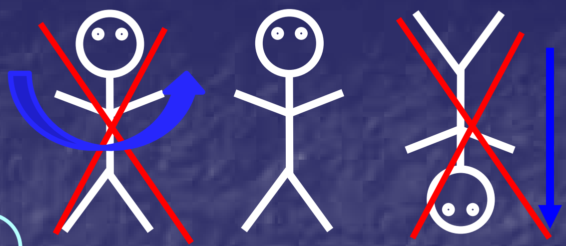


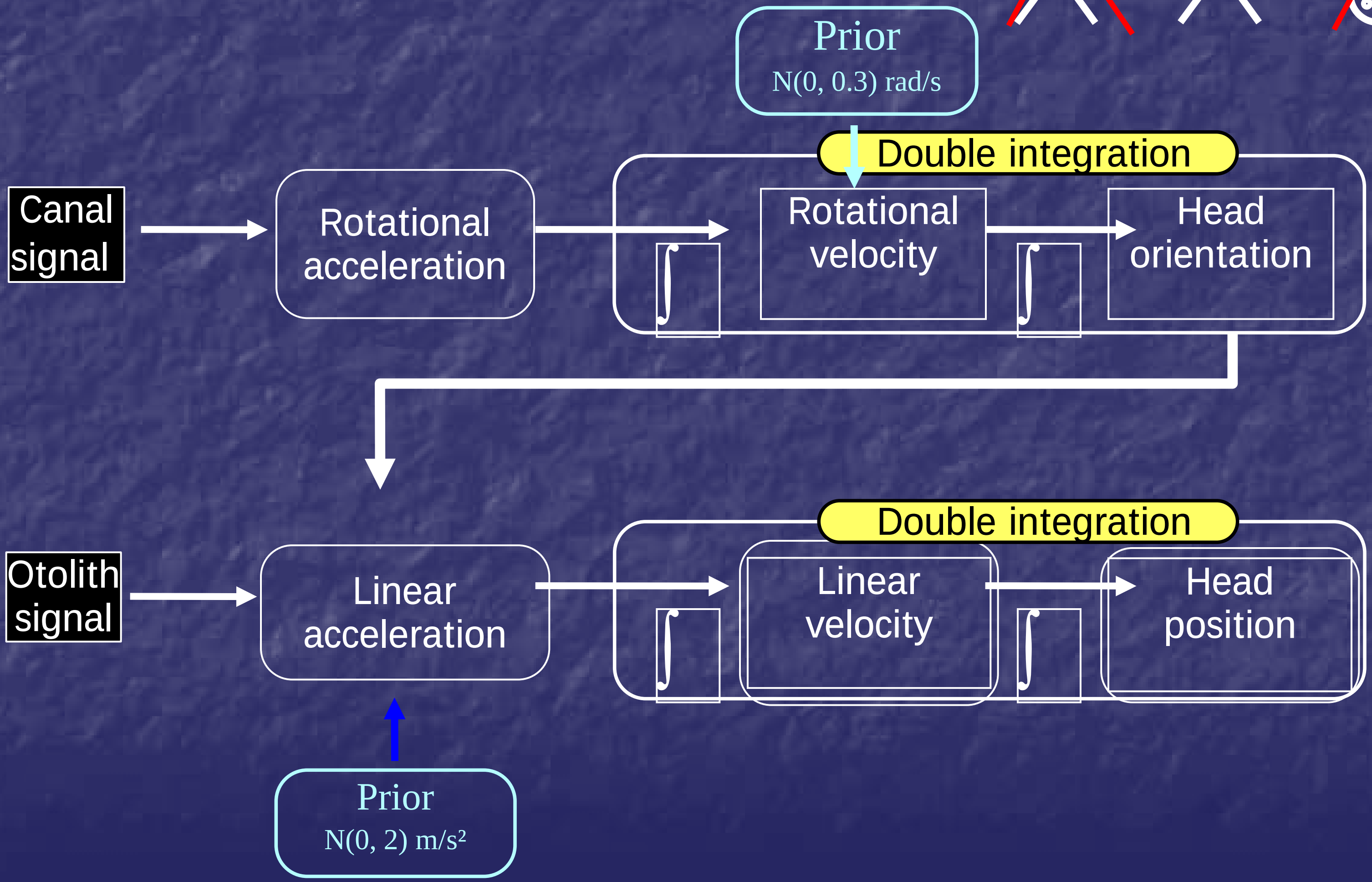
The most probable solution



Another (but less probable) solution

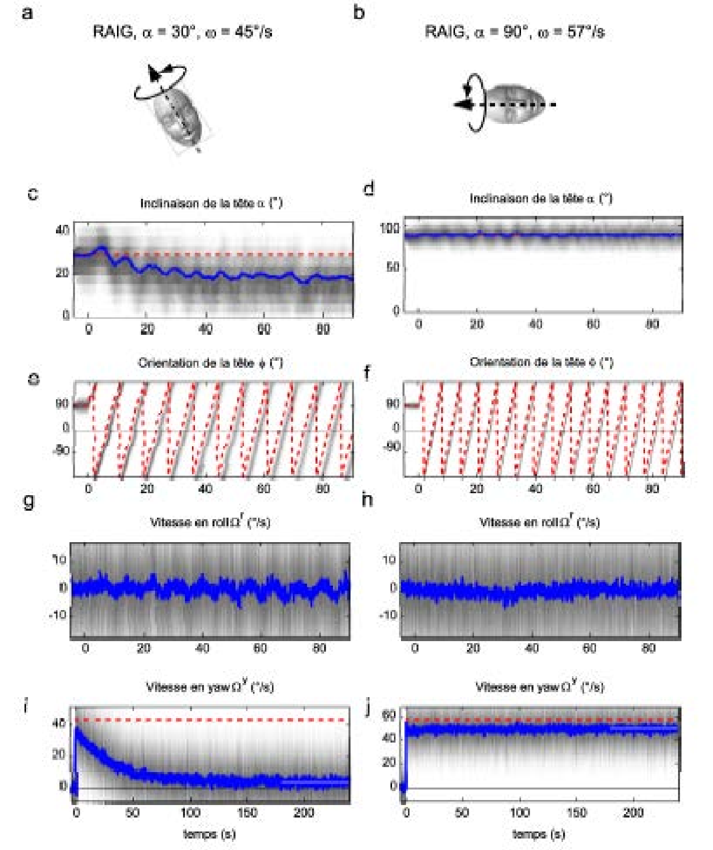
A Bayesian filter model including priors
(Low angular velocity & Low linear acceleration)



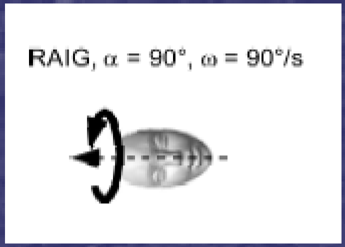


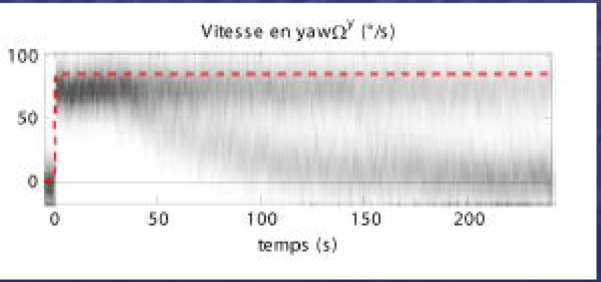
J. Laurens & J. Droulez, Biol. Cybernetics, 2007

Several effects on self-motion perception are explained, e.g.:
rotation at constant speed around an off-vertical axis



Bimodal distribution at high angular velocity

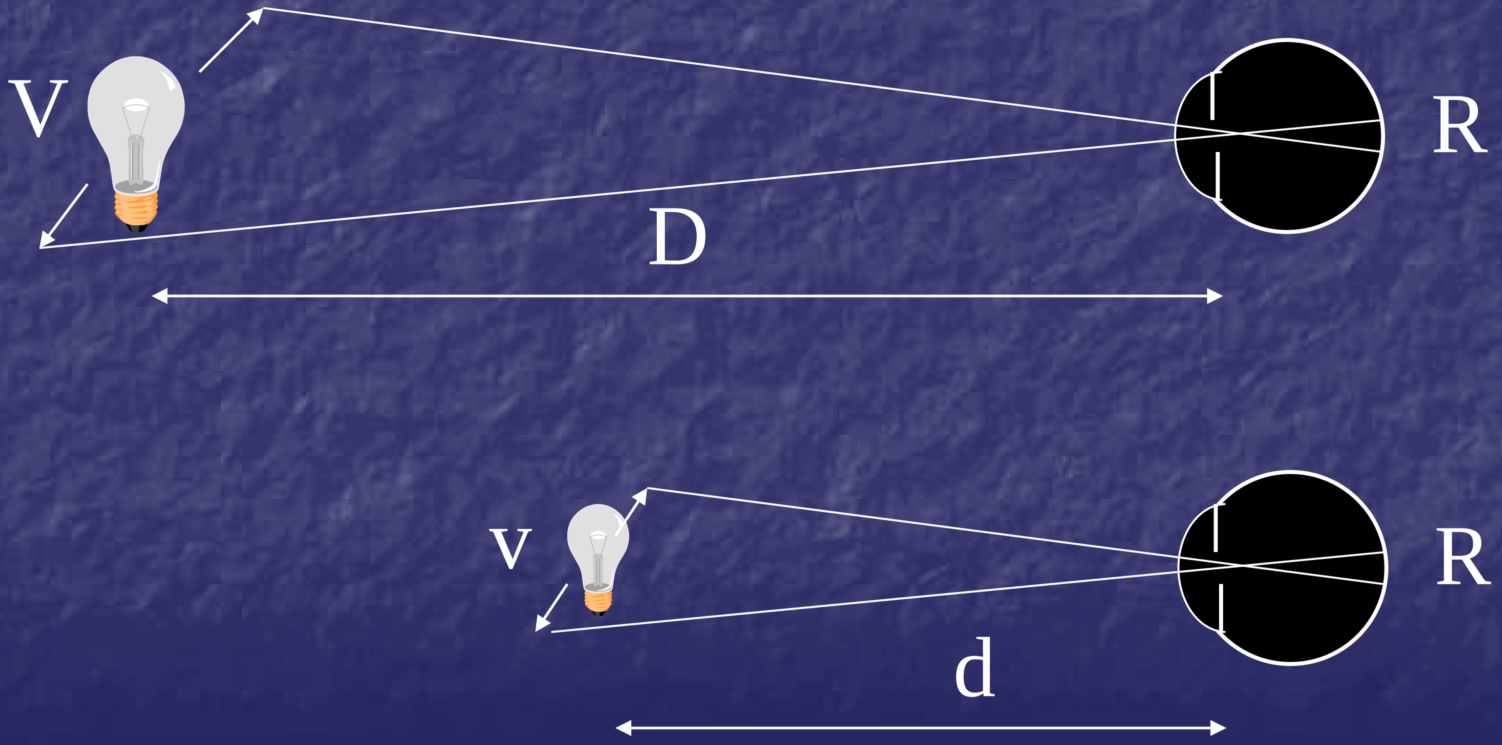




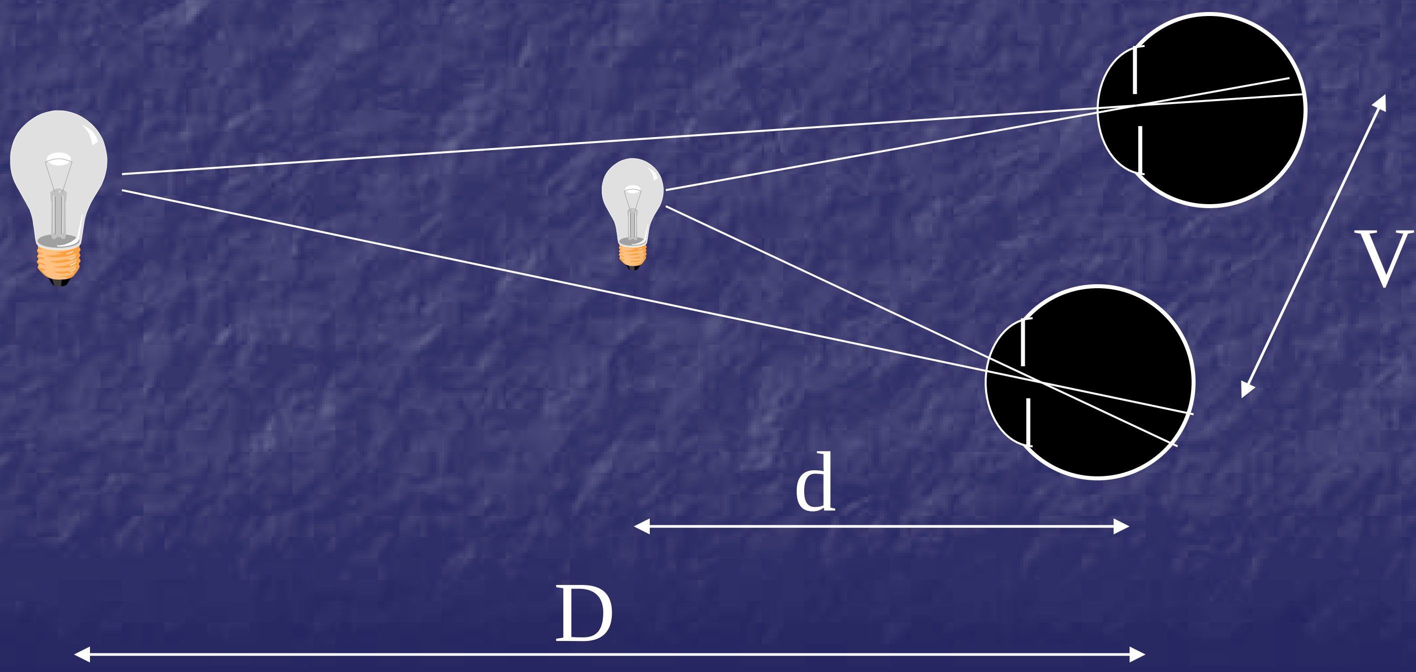
Data from Correia & Guedry (66), Lackner & Graybiel (78), Denise et al (88), …

Merging vestibular and visual information to solve the scale ambiguity:

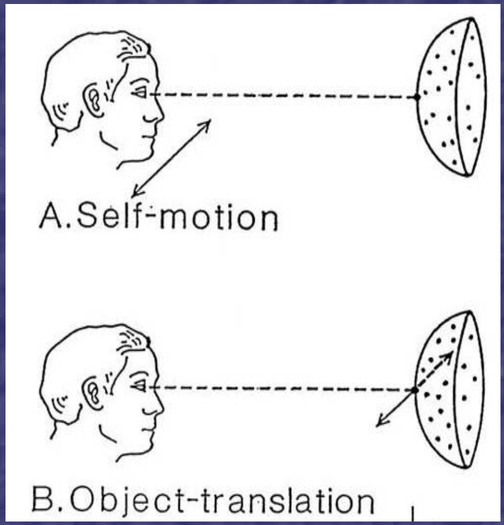
Depth, size and velocity of the object (in monocular vision) can be inferred from retinal information only to an unknown multiplicative scale factor.

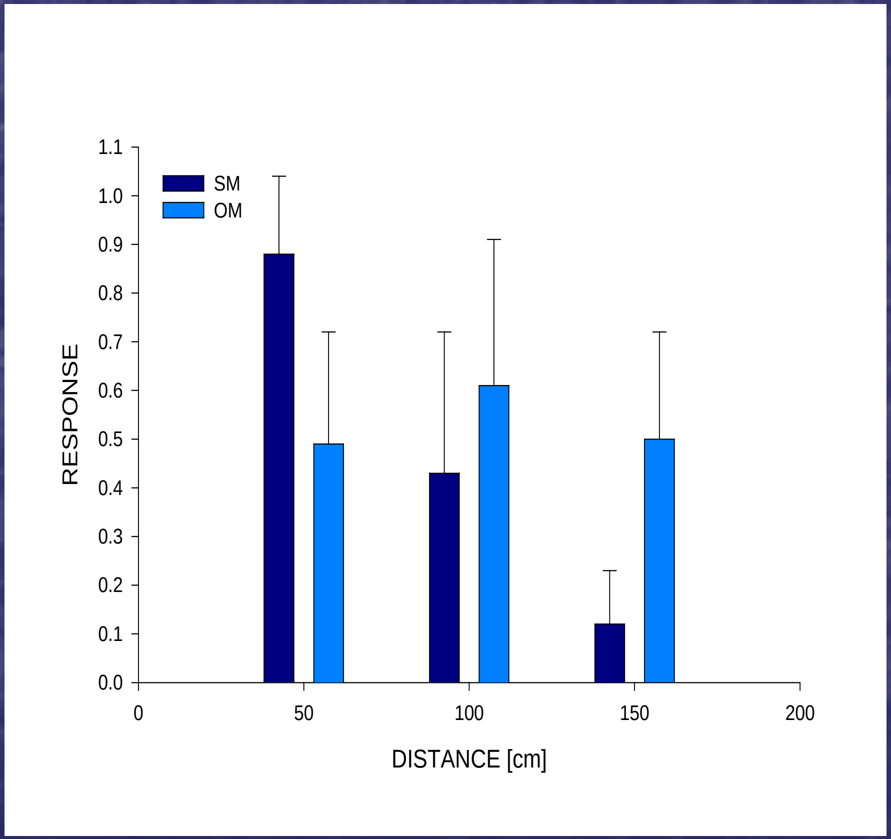


Assuming that the object is stationary, and estimating the self-motion V from vestibular signals can help to solve the scale ambiguity



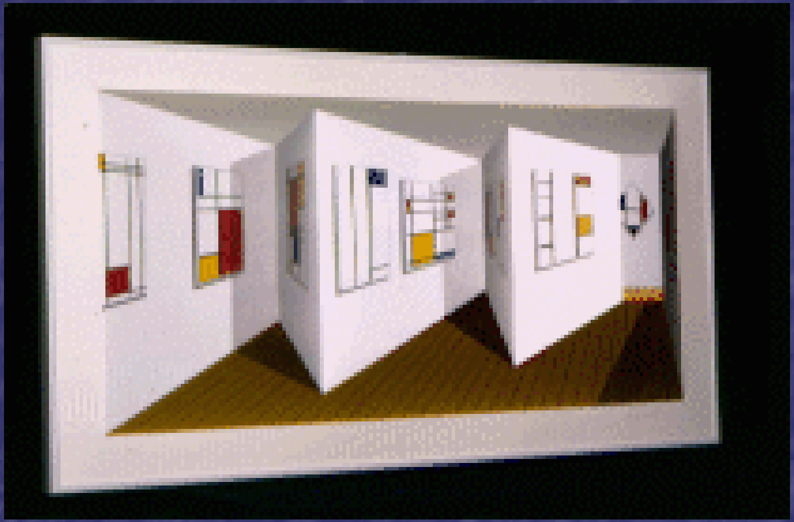
Comparison SM (subject's motion) versus OM (object's motion)
in the estimation of depth (probability of response inferior to 1 meter)



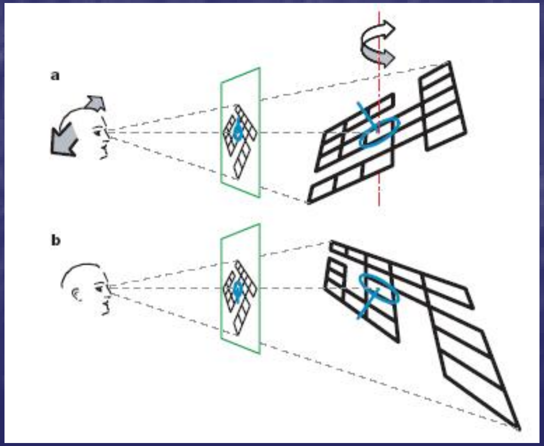


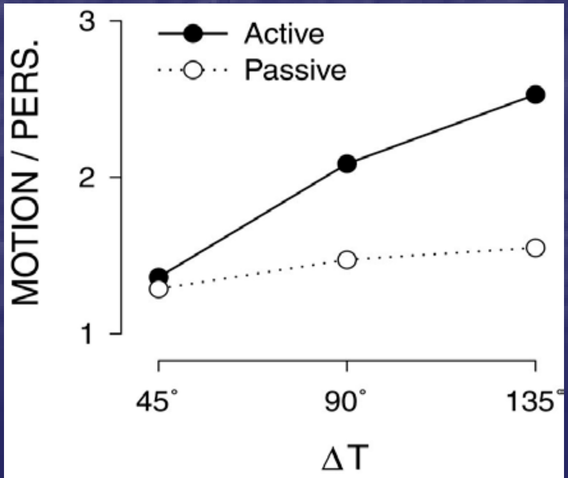
Panerai, Cornilleau-Pérès & Droulez, *Perception & Psychophysics*, 2002.

3D shape perception: the role of priors for regularity (perspective), rigidity (optic flow) and stationarity (self-motion)



Patrick Hughes
« Reverspective »

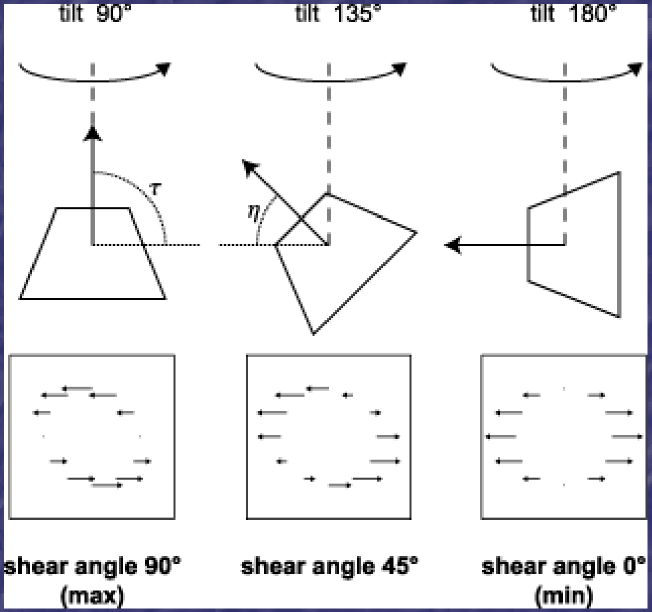


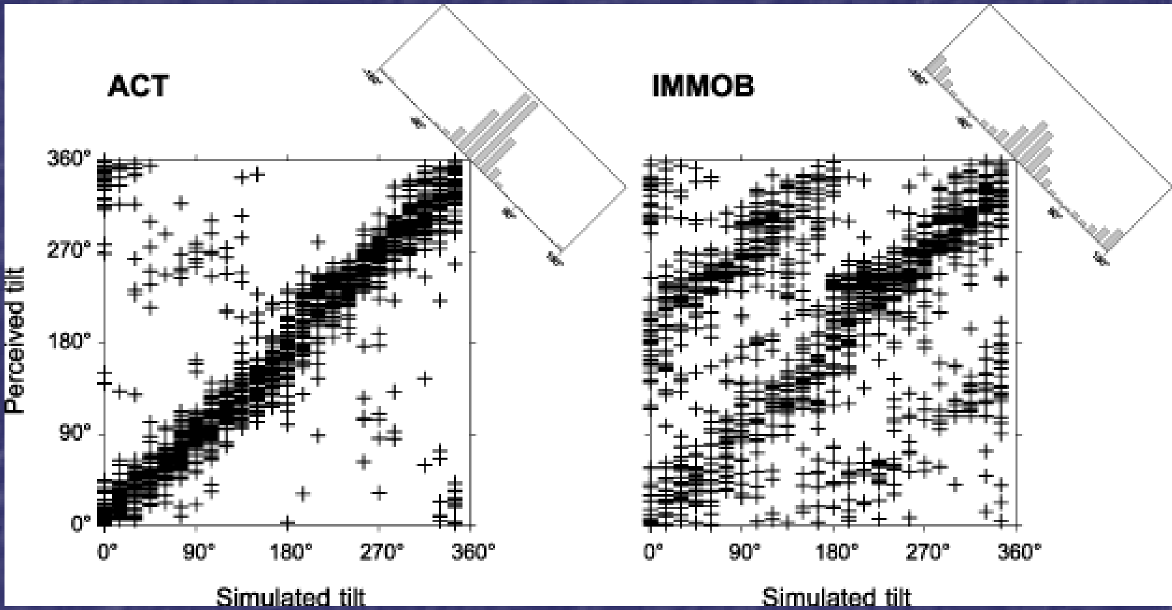


ΔT = Persp. Tilt / Motion Tilt

Wexler, Panerai, Lamouret & Droulez, *Nature*, 2001.

Perception of 3D plane orientation (« tilt ») from object and self motion





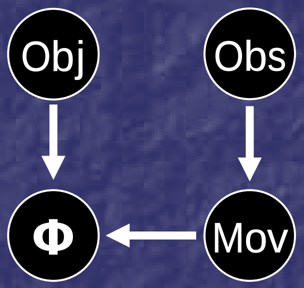
Van Boxtel, Wexler & Droulez (2003), *Journal of Vision* 3(5) : 318-332.

Variables : Object structure, Observer motion, Relative Motion, Optic Flow

Joint distribution:

P(Obj, Obs, Move, Flow) = P(Obj).P(Obs).P(Move | Obs).P(Flow | Move, Obj)

P(Obj) = regularity / perspective
P(Obs) = Self-motion information
P(Move | Obs) = Stationarity assumption
P(Flow | Move, Object) = Rigidity assumption



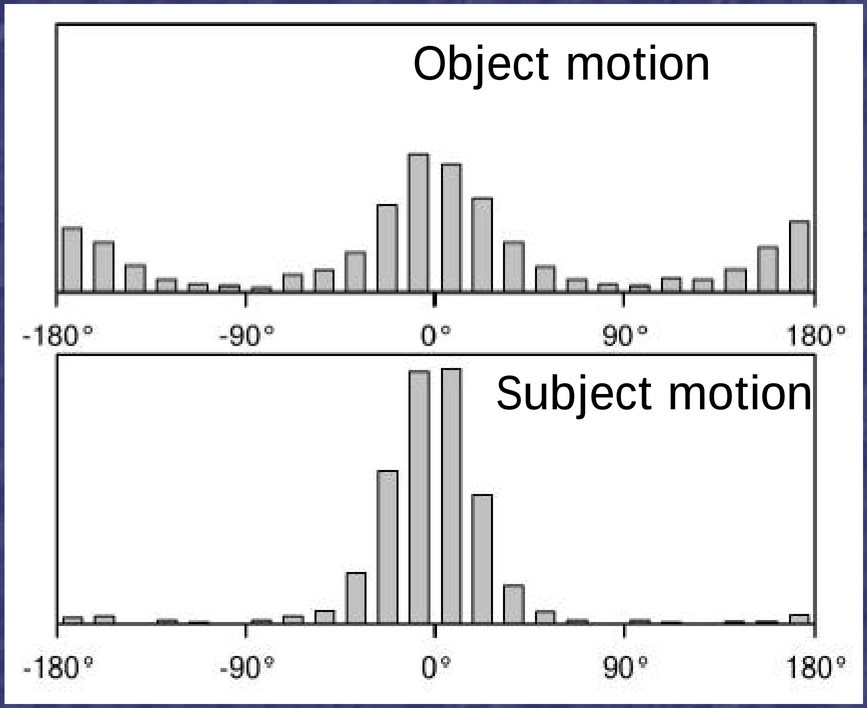
Question: P(Obj | Obs, Flow) ?

Experimental results to be explained:

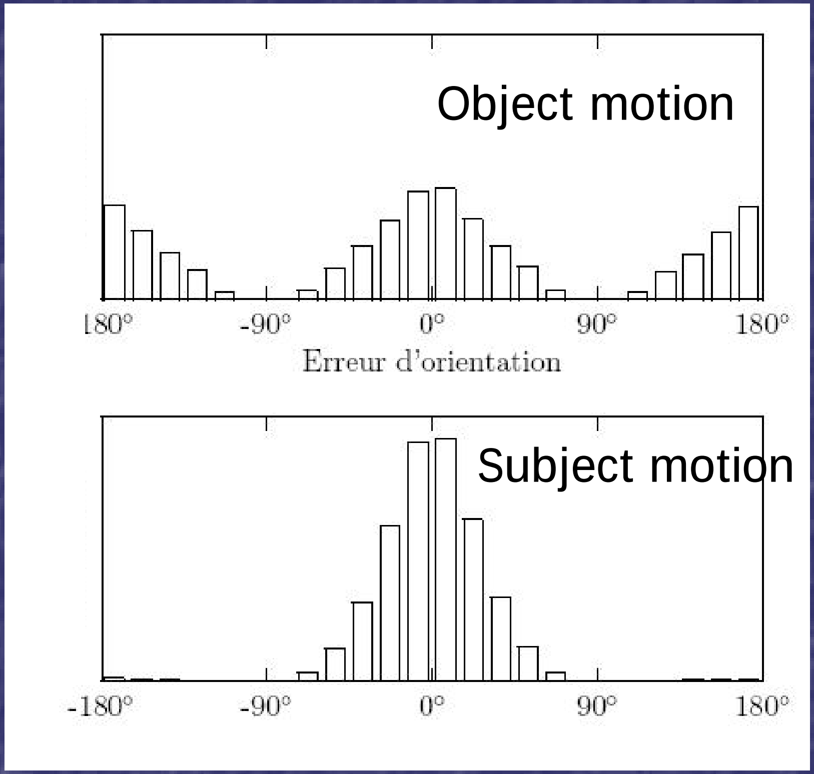
- Perceptive Inversion (suppressed in active condition)
- Perceptive variability due to shear (reduced in active condition)
- 90° Rotation of perceived orientation with added depth translation

F. Colas, J. Droulez, M. Wexler & P. Bessière, *Biol. Cybernetics* (2007)

Experiment



Model



Colas, Droulez, Wexler & Bessière (2007) *Biological Cybernetics*, 97:461-477

SUMMARY (Part 1)

- The problem in perception (e.g. 3D perception) is NOT to get rid of sensory noise, but to solve ambiguities and indeterminacies.

SUMMARY (Part 1)

- The problem in perception (e.g. 3D perception) is NOT to get rid of sensory noise, but to solve ambiguities and indeterminacies.

- In the past two decades, Bayesian models have been quite successful in explaining a large variety of perceptive effects (“illusions”).

SUMMARY (Part 1)

● The problem in perception (e.g. 3D perception) is NOT to get rid of sensory noise, but to solve ambiguities and indeterminacies.

● In the past two decades, Bayesian models have been quite successful in explaining a large variety of perceptive effects (“illusions”).

● Subjects perceptive or motor responses exhibit a large variability from trial to trial which does not result from stimulus variations, but matches quite well with the output probability distribution of Bayesian models.

SUMMARY (Part 1)

- The problem in perception (e.g. 3D perception) is NOT to get rid of sensory noise, but to solve ambiguities and indeterminacies.

- In the past two decades, Bayesian models have been quite successful in explaining a large variety of perceptive effects ("illusions").

- Subjects perceptive or motor responses exhibit a large variability from trial to trial which does not result from stimulus variations, but matches quite well with the output probability distribution of Bayesian models.

- Individual subject response could be a "sample" drawn from the internally estimated probability distribution.

Part 2: The Bayesian Brain

- How probability distributions are represented in the brain ?
- How Bayesian inferences are performed by neurons ?

1. A variety of theoretical propositions

Explicit

- Direct code : single neuronal activity ↔ probability value

$r \approx P(S = s)$ …. $r \approx Log(P(S = s))$ … $r \approx Log(P(S = 1) / P(S = 0))$

Anastasio et al (2000); Gold & Shadlen (2001); Rao (2004); Yang & Shadlen (2007); …

- Population code : ensemble of neurones ↔ linear combination of a set of basis functions

$P(S = s) \approx \Sigma_i\, r_i.h_i(s)$ or $Log(P(S = s)) \approx \Sigma_i\, r_i.h_i(s)$

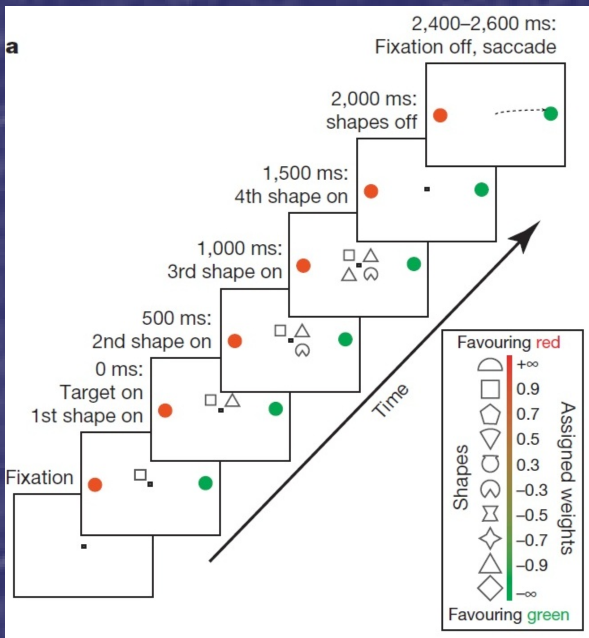
Zemel, Dayan & Pouget (1998); Ma, Beck, Latham & Pouget (2006); …

Implicit

- Sampling code: instantaneous population activity ↔ random draw from a probability distribution

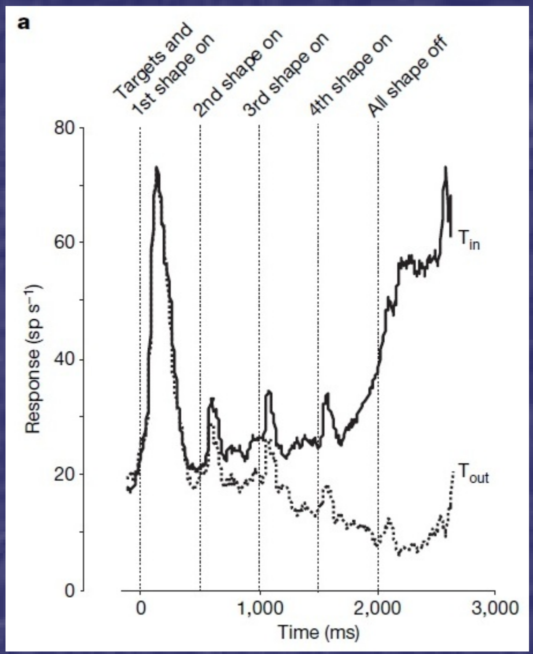
Lee & Mumford (2003); Fiser et al (2010); Maass (2014); …

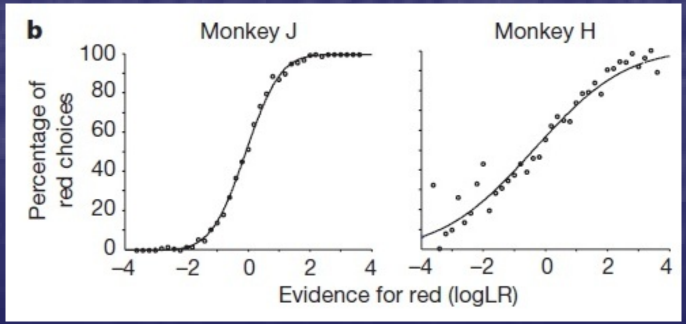
1. Evidence for a direct code (Log Likehood Ratio)

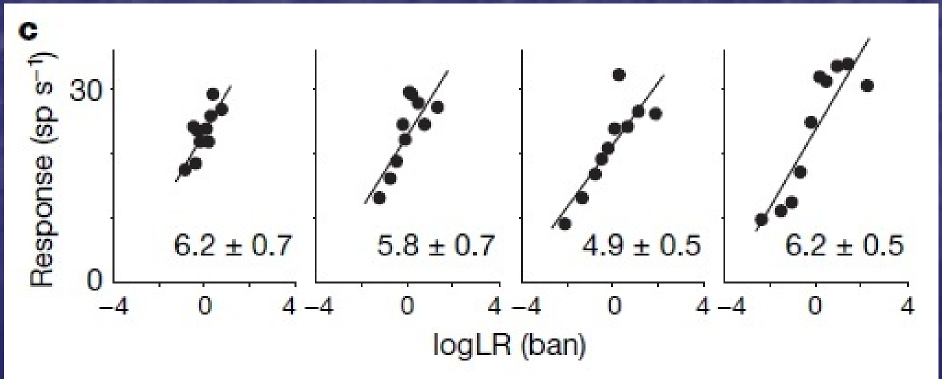


Accumulation of evidence (in LLR)

Activity in LIP (overtrained monkeys)



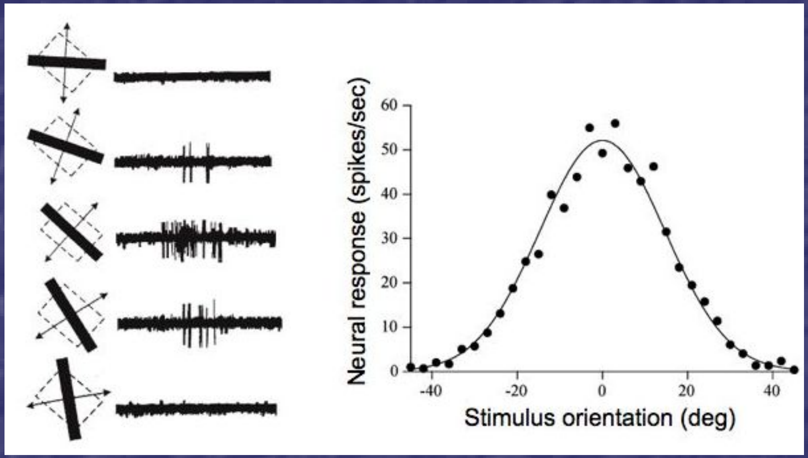




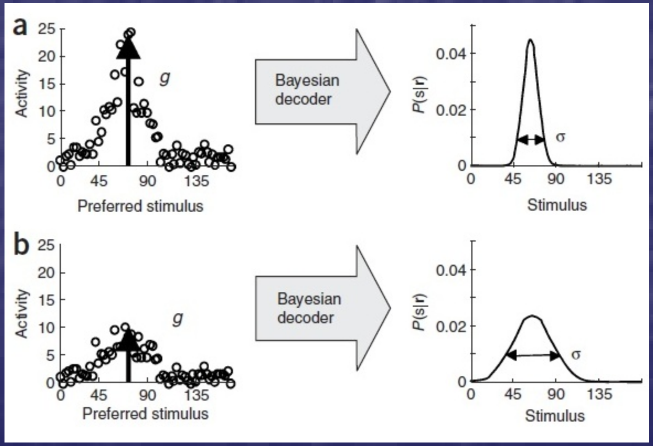
But LLR and P(Choice) are highly correlated !

Yang & Shadlen, *Nature* 447 (2007)

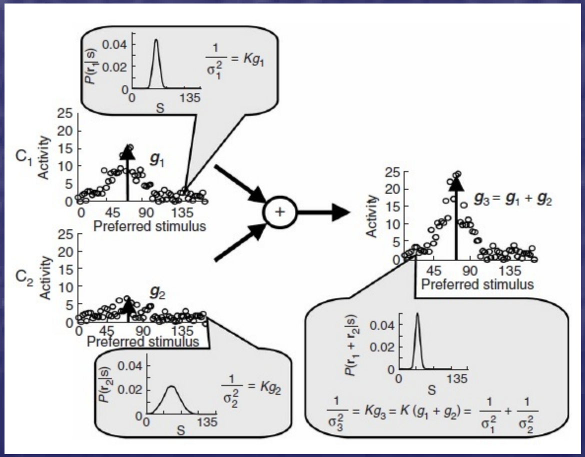
2. Evidence for a population code (Tuning curves)



In cats: Hubel & Wiesel, J. Phys. (1959). In monkeys: Hubel & Wiesel, J. Phys. (1968)



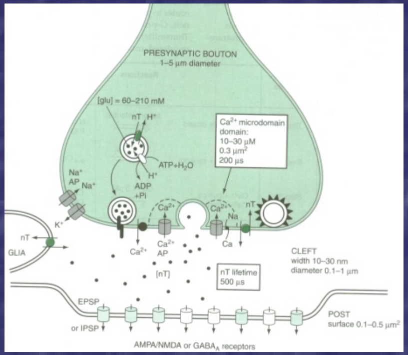
Higher gain → Lower variance

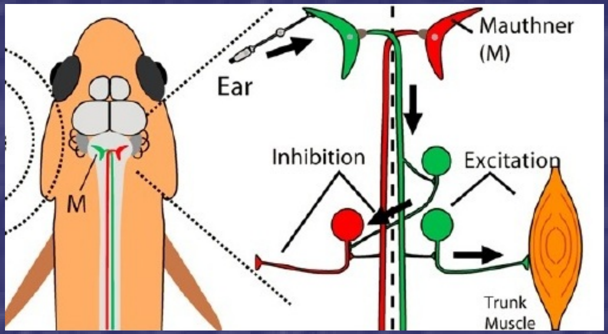


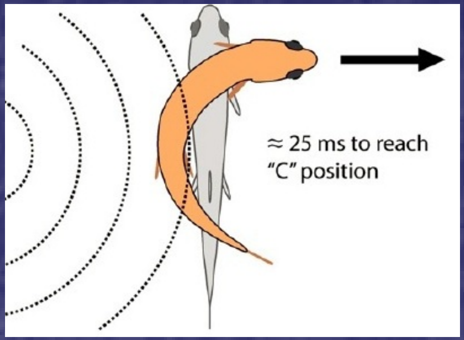
Sum of activity → Product of distribution

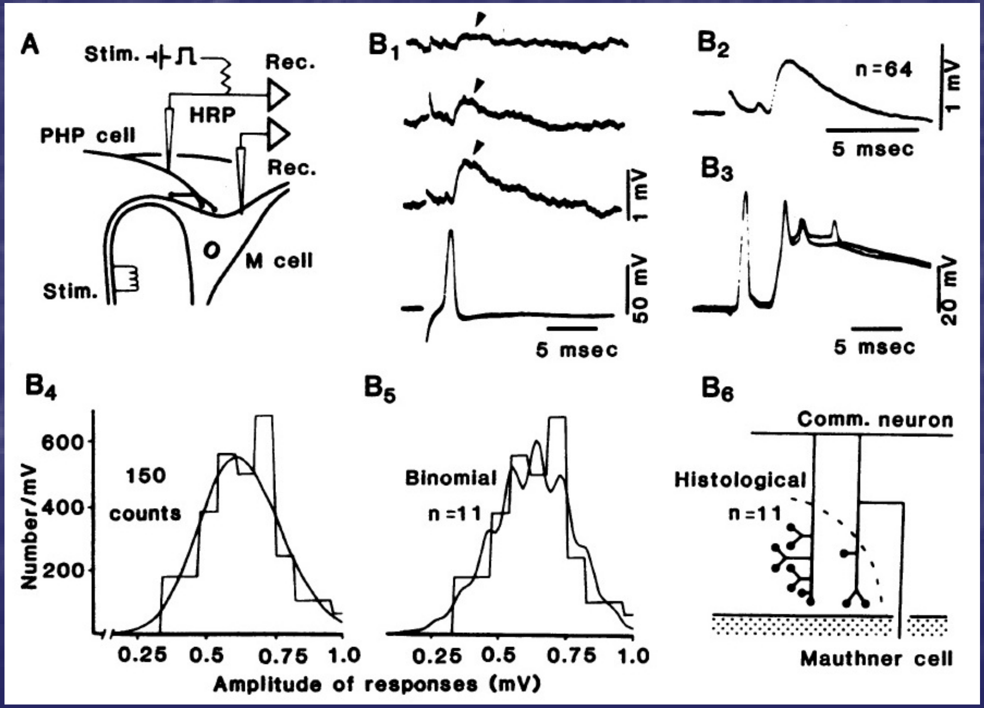
Ma, Beck, Latham & Pouget, *Nature Neurosc.* 9:1432 (2006)

3. Evidence for a sampling code (stochastic neural activity)





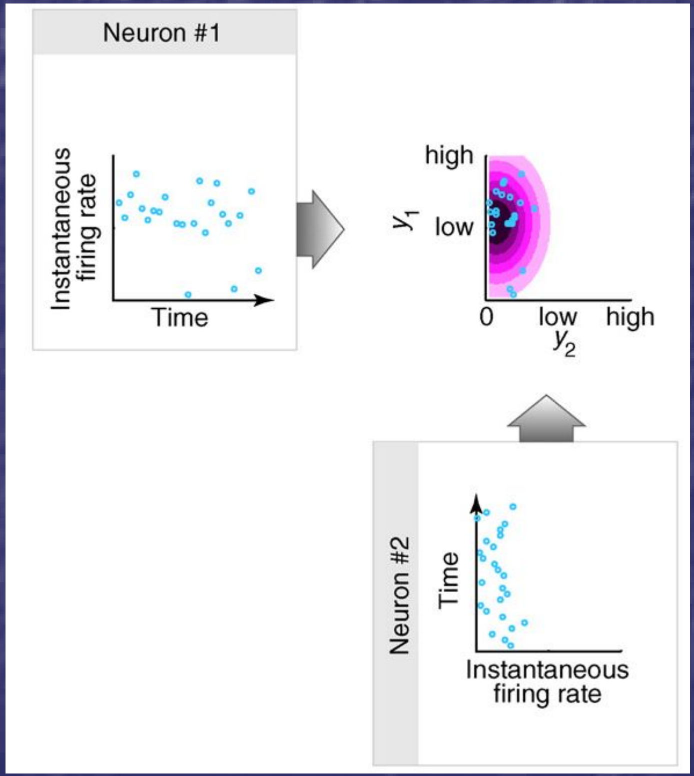




Data and model from Korn et al, Science 213 (1981)

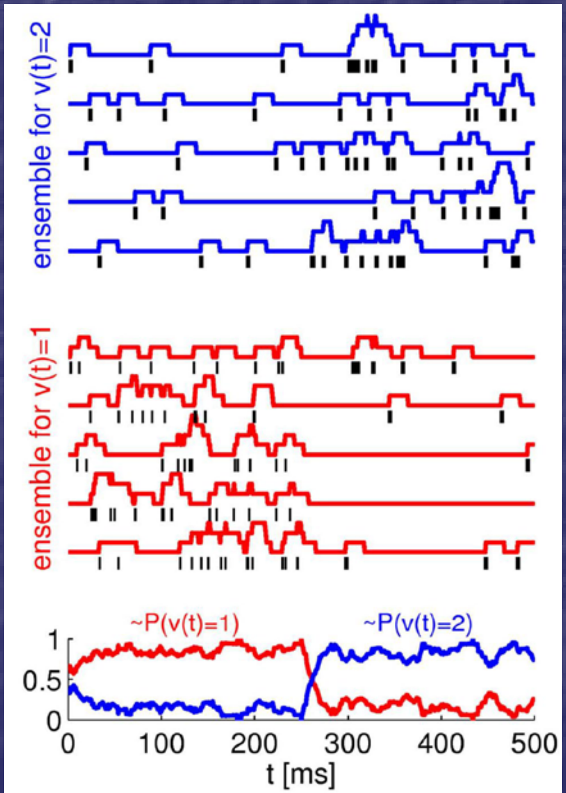
4. Examples of proposed sampling codes

One neuron per (discrete) variable



Fiser et al, Trends in Cognitive Sc. 14 (2010)

One population per (binary) variable



Legenstein & Maass, PLoS CB (2014)

SUMMARY (Part 2)

- Partial experimental evidences in favor of each of the (mutually exclusive) theoretical propositions.

SUMMARY (Part 2)

- Partial experimental evidences in favor of each of the (mutually exclusive) theoretical propositions.

- Log versions are well suited for the product rule (“Bayes’ theorem”), e.g. in naïve fusion models, but not for the sum rule (“marginalization”), e.g. in models with free variables and in temporal filters.

SUMMARY (Part 2)

- Partial experimental evidences in favor of each of the (mutually exclusive) theoretical propositions.

- Log versions are well suited for the product rule (“Bayes’ theorem”), e.g. in naïve fusion models, but not for the sum rule (“marginalization”), e.g. in models with free variables and in temporal filters.

- Direct codes and population codes aim at representing explicitly the probability distributions. Computation is based on exact inference (or close to exact inference) . Neural “noise” is conceived as a nuisance. Might be not suited for solving problems in high dimension spaces.

SUMMARY (Part 2)

● Partial experimental evidences in favor of each of the (mutually exclusive) theoretical propositions.

● Log versions are well suited for the product rule (“Bayes’ theorem”), e.g. in naïve fusion models, but not for the sum rule (“marginalization”), e.g. in models with free variables and in temporal filters.

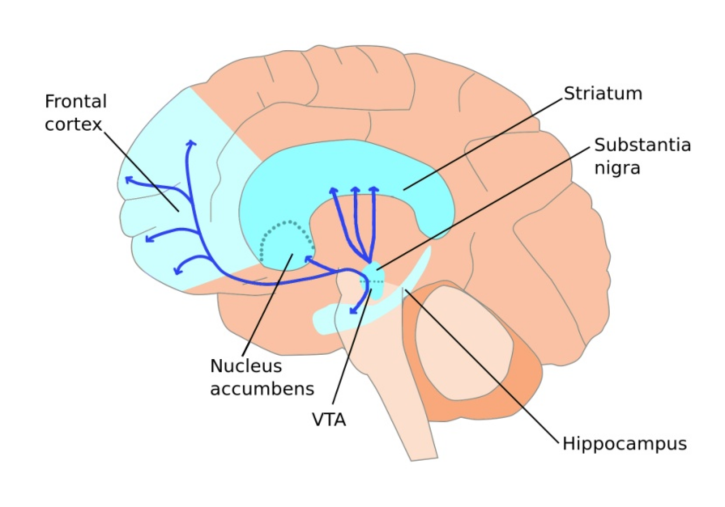
● Direct codes and population codes aim at representing explicitly the probability distributions. Computation is based on exact inference (or close to exact inference). Neural “noise” is conceived as a nuisance. Might be not suited for solving problems in high dimension spaces.

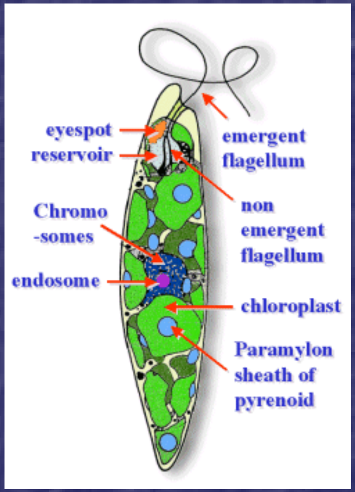
● Sampling code: accounts for biological stochasticity, well suited for hard inference problems. But the relevance of known sampling approach (*e.g.* MCMC) in neurobiology has yet to be demonstrated.

Part 3: The Bayesian Cell

Neuronal activity is also controlled by complex biochemical networks

Unicellular organisms have also developed well adapted behaviors in spite of uncertain environment



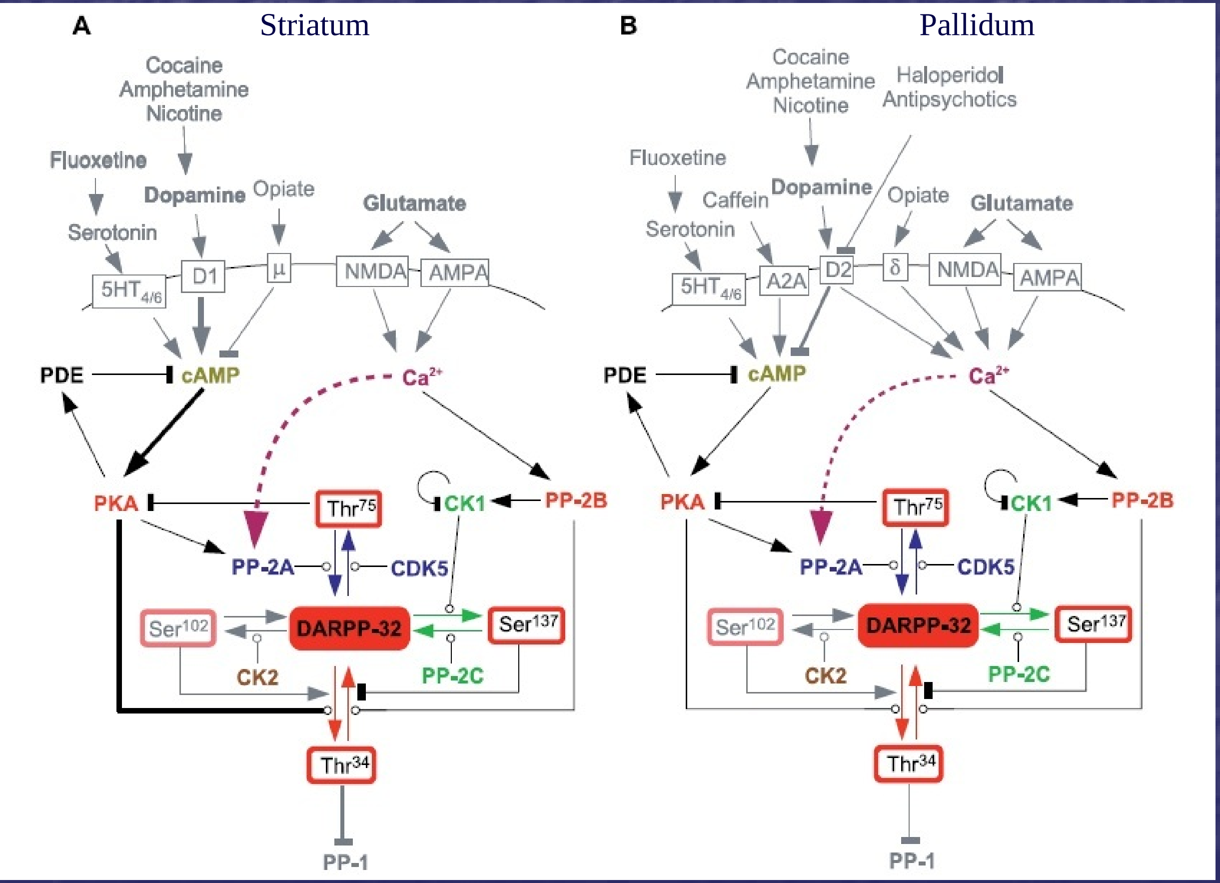


Euglena

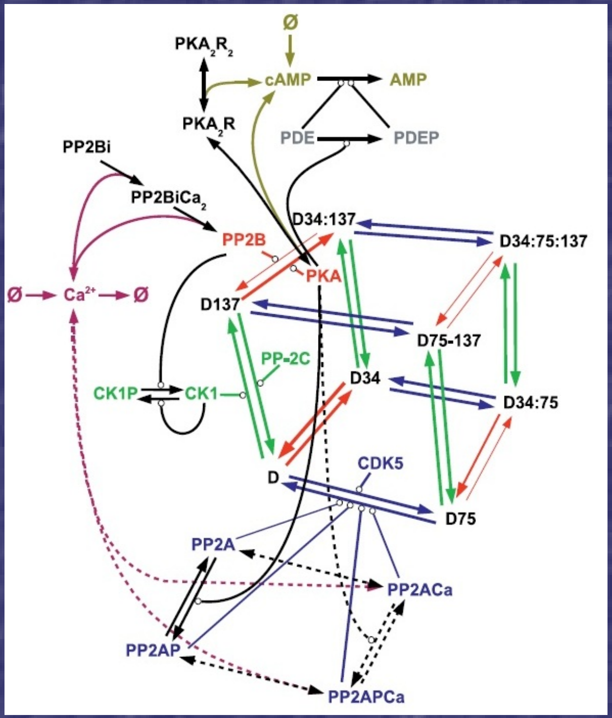
Chlamydomonas

Integration of dopamine and glutamate signals in neurons of the basal ganglia (striatum and pallidum), role in reinforcement learning. Frank et al, Nature Neurosc. (2009)

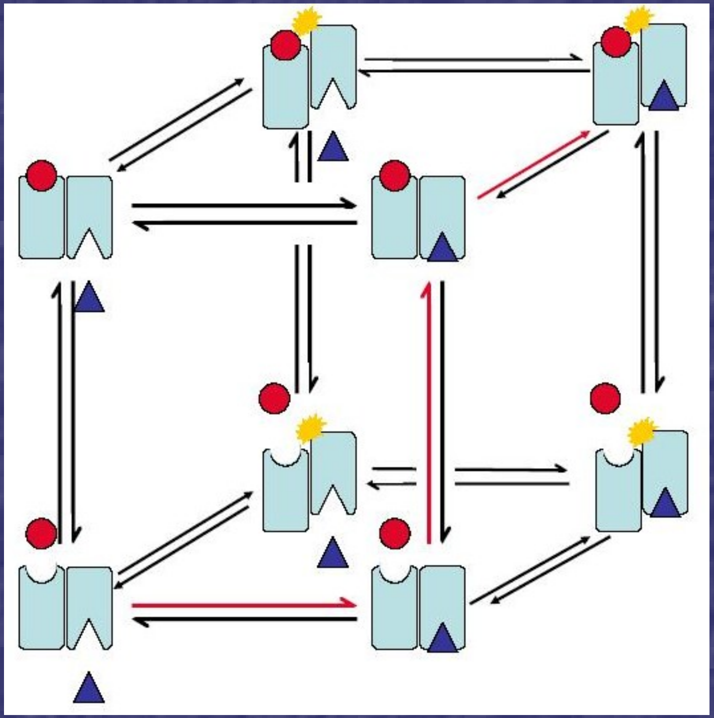
Perkins & Swain, Strategies for cellular decision-making, Mol. Syst. Biol, (2009)



Fernandez et al, DARPP32 is a robust integrator of Dopamine and Glutamate Signals. PLoS Comp. Biol. (2006)

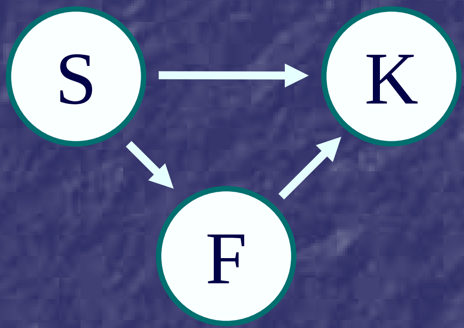


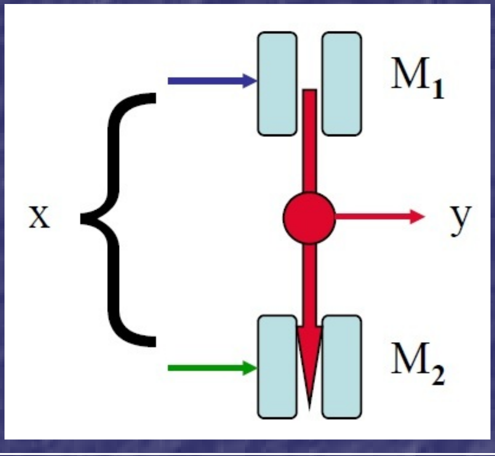
DARPP32:
3 sites of phosphorylation → 8 states
Fernandez et al (2006)



A Markov model of allosteric transitions
Droulez et al (2015)

Equivalence between Bayesian inference and cascades of biochemical systems





$$\frac{P([S=s]\,|\,k)}{P([S=0]\,|\,k)} = \frac{\sum_F P([S=s],F) \times P(k\,|\,[S=s],F)}{\sum_F P([S=0],F) \times P(k\,|\,[S=0],F)}$$

The output probability quotient is a rational function (with non negative coefficients) of likelihood quotients.

<u>Markov model of a biochemical module:</u>

N_Y = number of second messengers

$\Phi_1(x)$ = rate of release (by M_1) : a RFNC of x

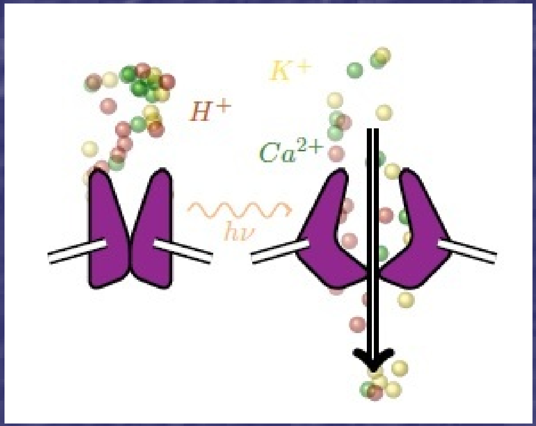
$\varphi_2(x)$ = rate of removal per messenger (by M_2)

$\Rightarrow$ At equilibrium $P(N_Y)$ is a Poisson distribution of parameter $\lambda(x) = \Phi_1(x) / \varphi_2(x)$

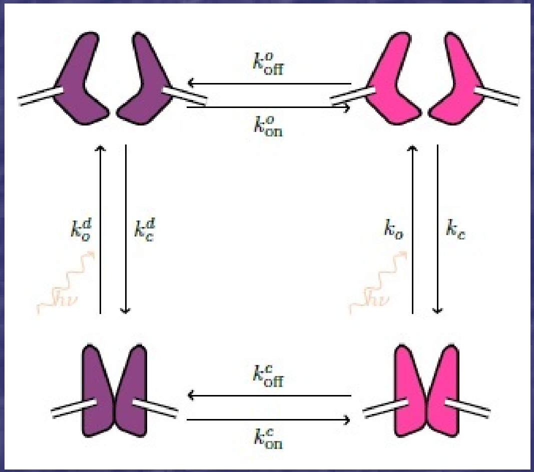
The output concentration y is a RFNC of x.

Towards a Bayesian model of sensory-motor behavior in unicellular organisms

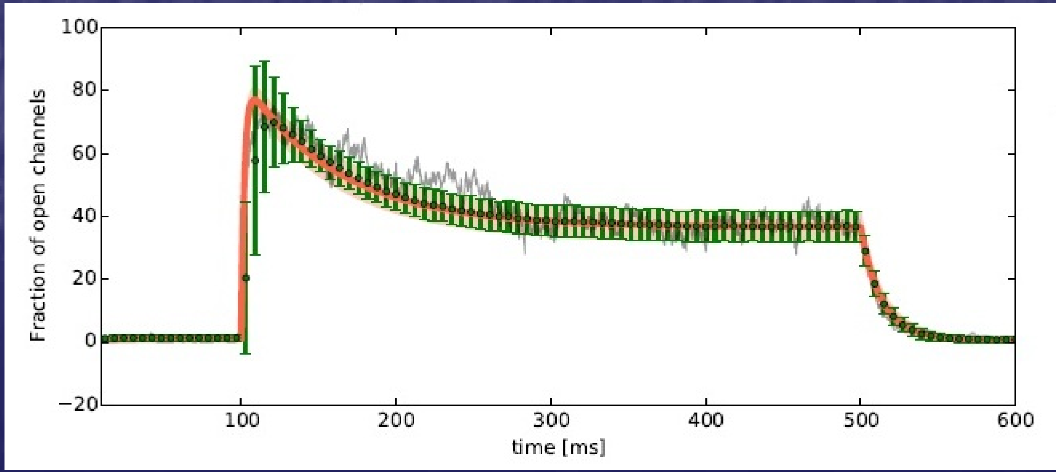
Channelrhodopsin: the molecular light sensor in the eyespot



Markov model of Channelrhodopsin (4 states)

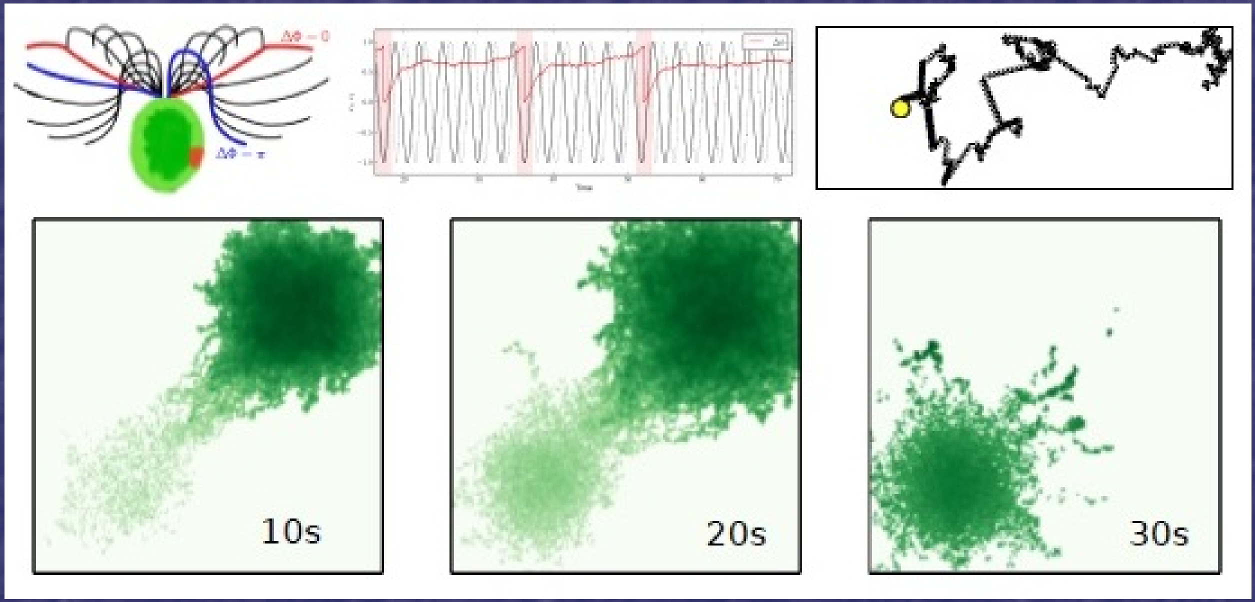


Example of simulation (Colliaux, Bessière & Droulez, SAND 2014)



Towards a Bayesian model of sensory-motor behavior in unicellular organisms

Simulation of phototaxis behavior (Colliaux et al, ECAL 2015)



SUMMARY (Part 3)

- In complement to the usual neurocomputational approach (e.g. integrate-and- fire neurons), models of the underlying biochemical signaling networks are required to understand how the brain could perform Bayesian computing.

SUMMARY (Part 3)

● In complement to the usual neurocomputational approach (e.g. integrate-and- fire neurons), models of the underlying biochemical signaling networks are required to understand how the brain could perform Bayesian computing.

● Unicellular organisms have no brain, but a number of (molecular) sensory and motor devices. They can adapt to highly changing and uncertain environments. Why such simple organisms would not use a kind of basic Bayesian computing ?

SUMMARY (Part 3)

- In complement to the usual neurocomputational approach (e.g. integrate-and- fire neurons), models of the underlying biochemical signaling networks are required to understand how the brain could perform Bayesian computing.

- Unicellular organisms have no brain, but a number of (molecular) sensory and motor devices. They can adapt to highly changing and uncertain environments. Why such simple organisms would not use a kind of basic Bayesian computing ?

- The equivalence between Bayesian inferences and the behavior of large populations of macromolecules involved in cell signaling opens new perspectives to understand how single cells and unicellular organisms could process uncertain information.

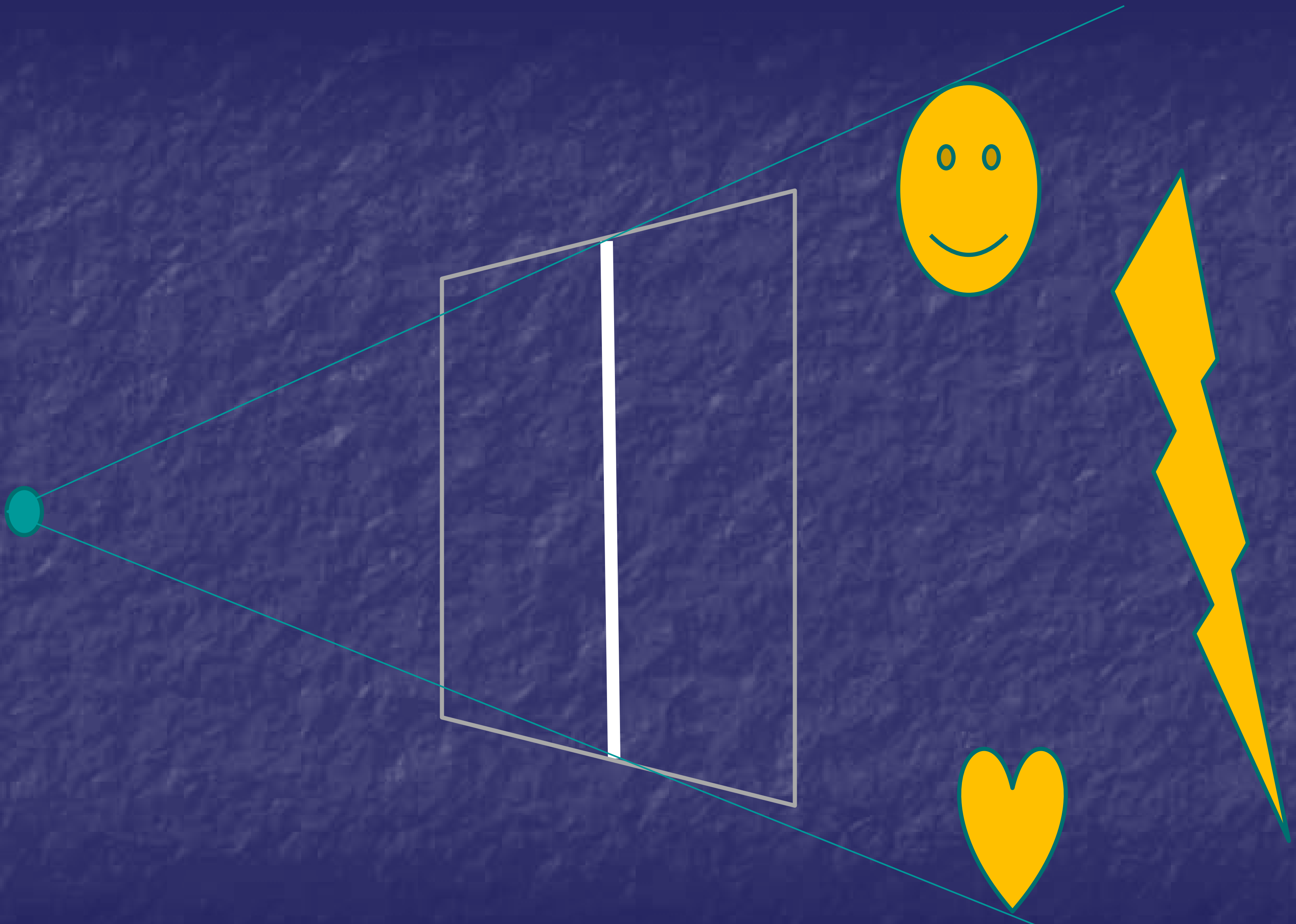
CONCLUSION

1. Bayesian theory of perception and behavior : a success story.

2. How probability is coded and processed in the brain is still a highly controversial question.

3. New perspectives might emerge from the understanding of information processing at molecular level.

Thank you for your attention !



A vertical line in the image ⇒ An infinite number of complex scenes