Modelling stochastic neural learning

Computational Neuroscience
András Telcs

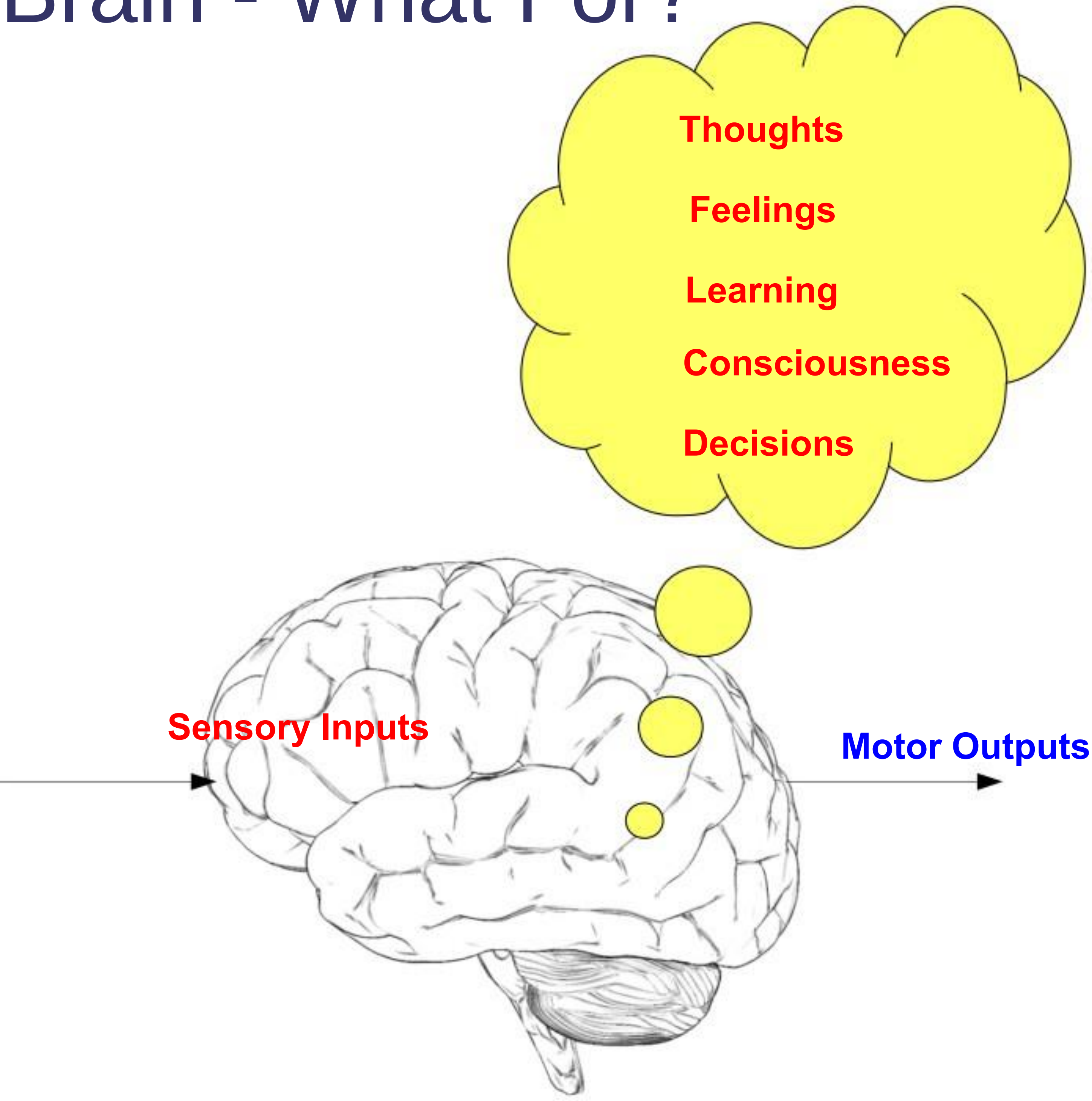
telcs.andras@wigner.mta.hu
www.cs.bme.hu/~telcs
http://pattern.wigner.mta.hu/participants/andras-telcs

Compiled from lectures of

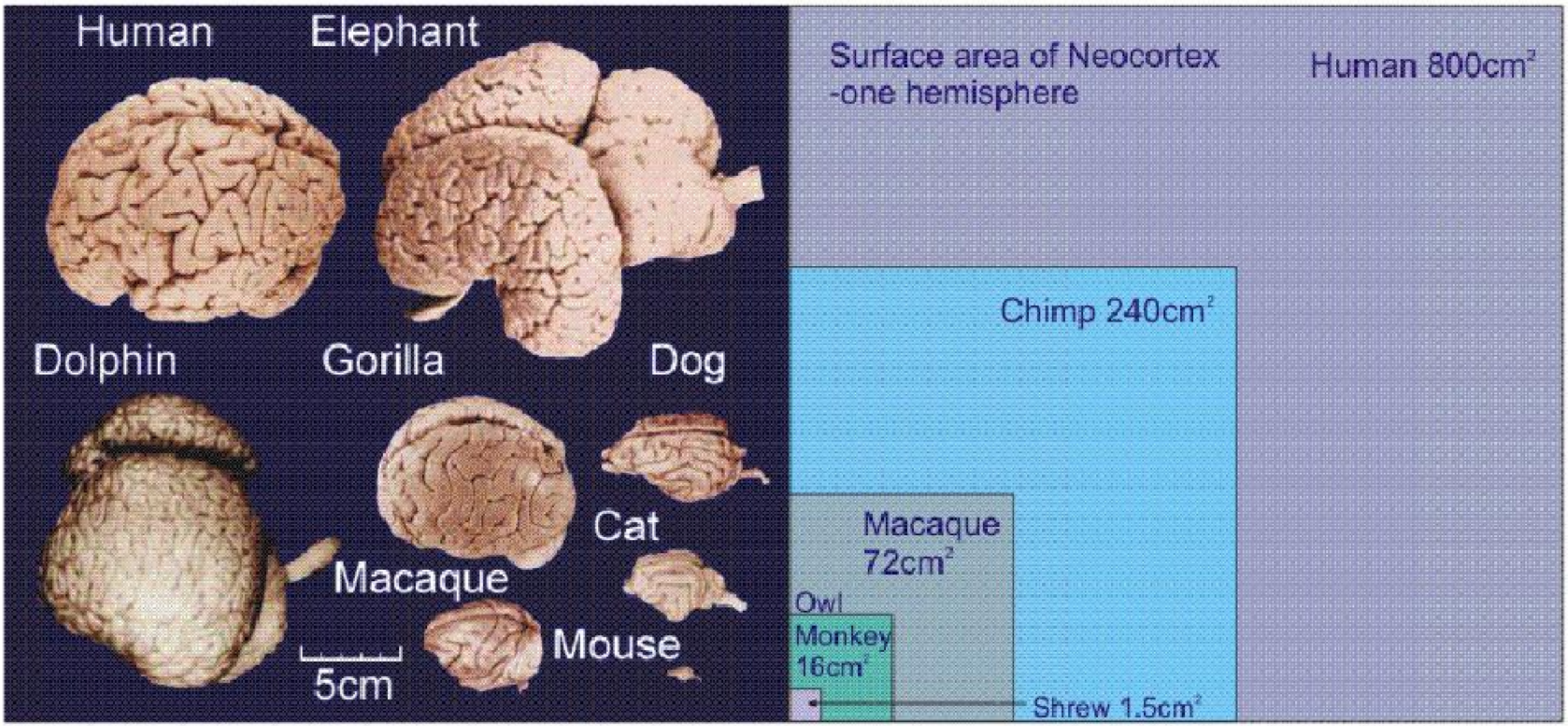
Prashant Joshi
Jotahna Pillow

Review of
“Statistically optimal perception and learning: from behavior to neural representations.” Fiser, Berkes, Orban & Lengyel Trends in Cognitive Sciences (2010)
“Spontaneous Cortical Activity Reveals Hallmarks of an Optimal Internal Model of the Environment.” Berkes, Orban, Lengyel, Fiser. *Science* (2011)
“Neural Dynamics as Sampling: A Model for Stochastic Computation in Recurrent Networks of Spiking Neurons” Lars Buesing , Johannes Bill, Bernhard Nessler, Wolfgang Maass Published: November 3, 2011 DOI: 10.1371/journal.pcbi.1002211

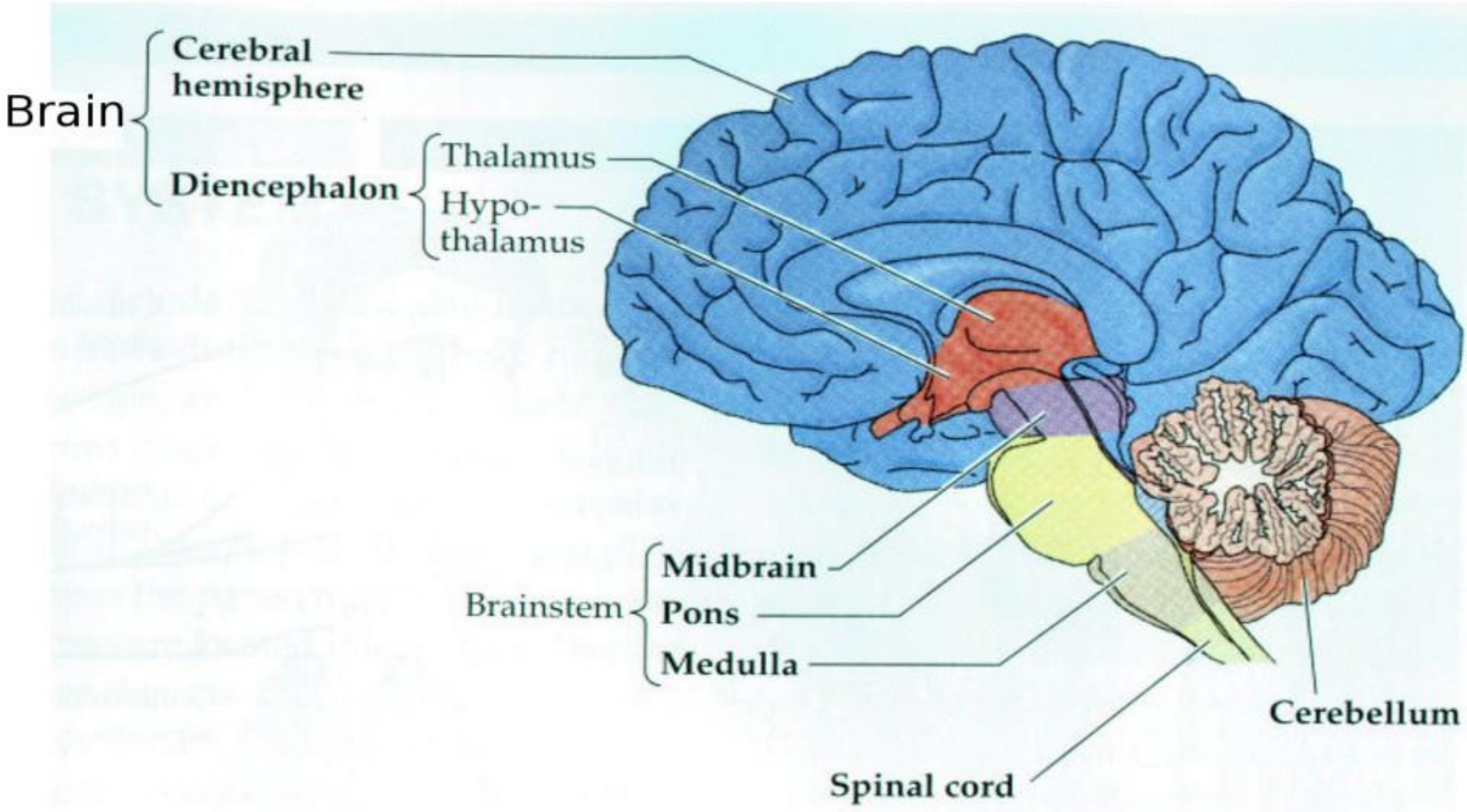
Brain - What For?



Mammalian Brains

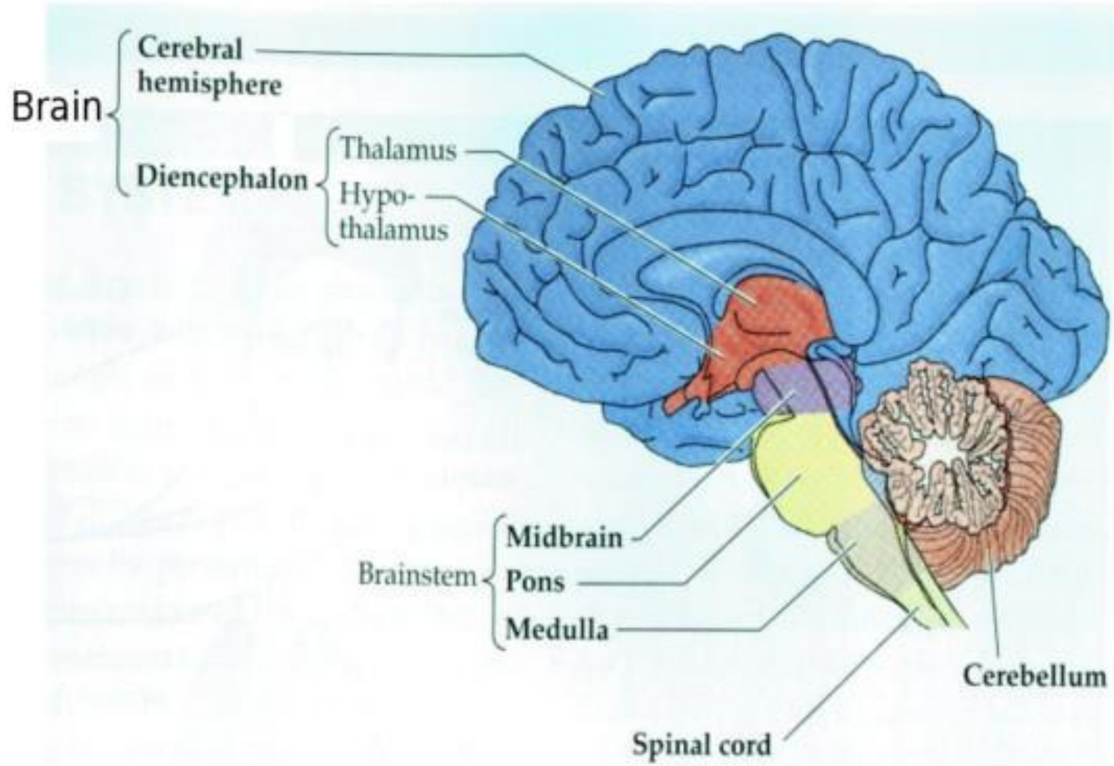


Broad Classification of



Each Region Has a Specific

population activity research unit



Spinal Cord - Sensory & motor information and control of reflexes

Brainstem - Information gateway between body and brain, integrative functions (cardiovascular, respiratory, and pain-sensitivity control and also consciousness)

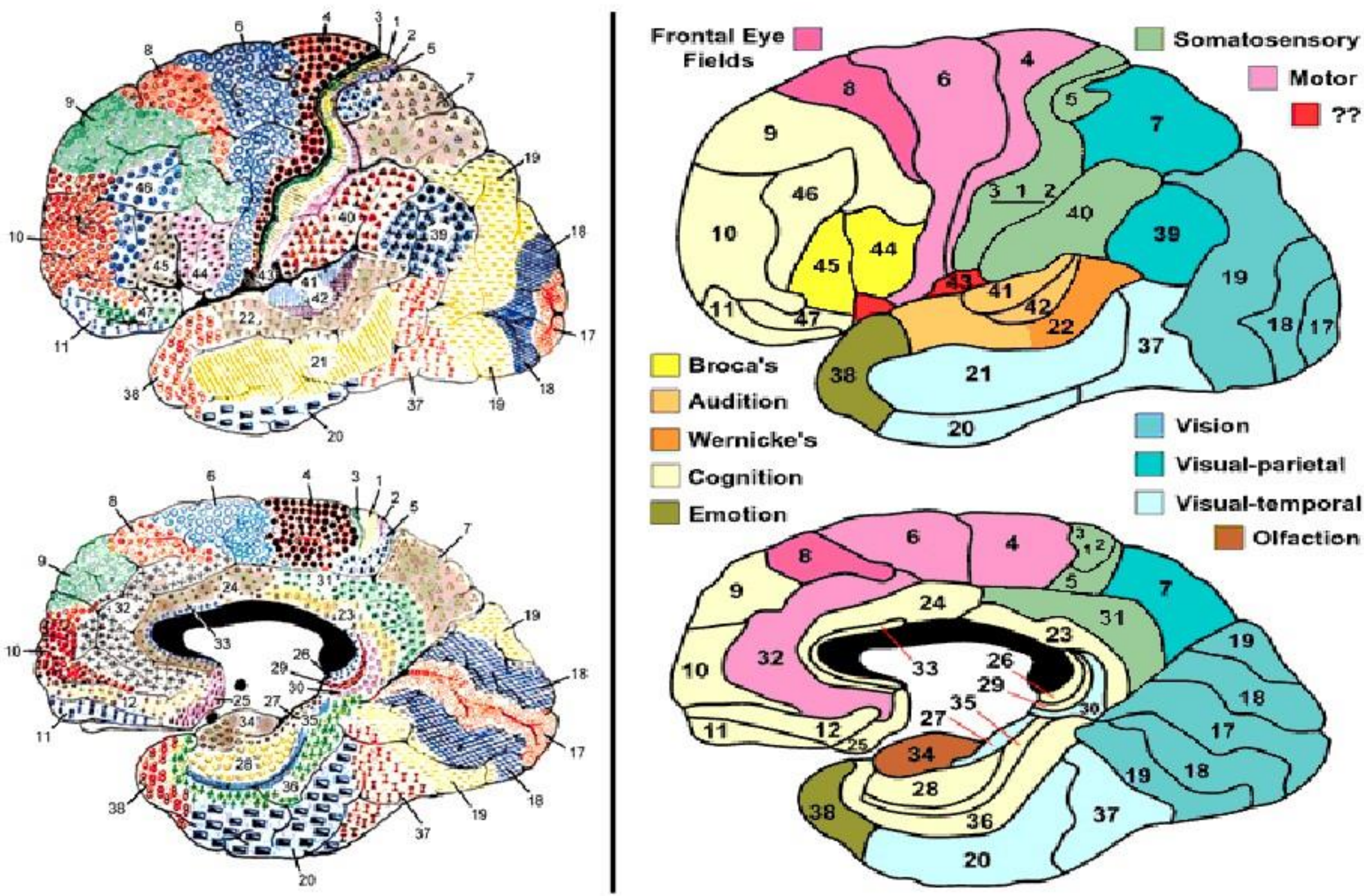
Cerebellum - Motor Control, Attention & Language etc.

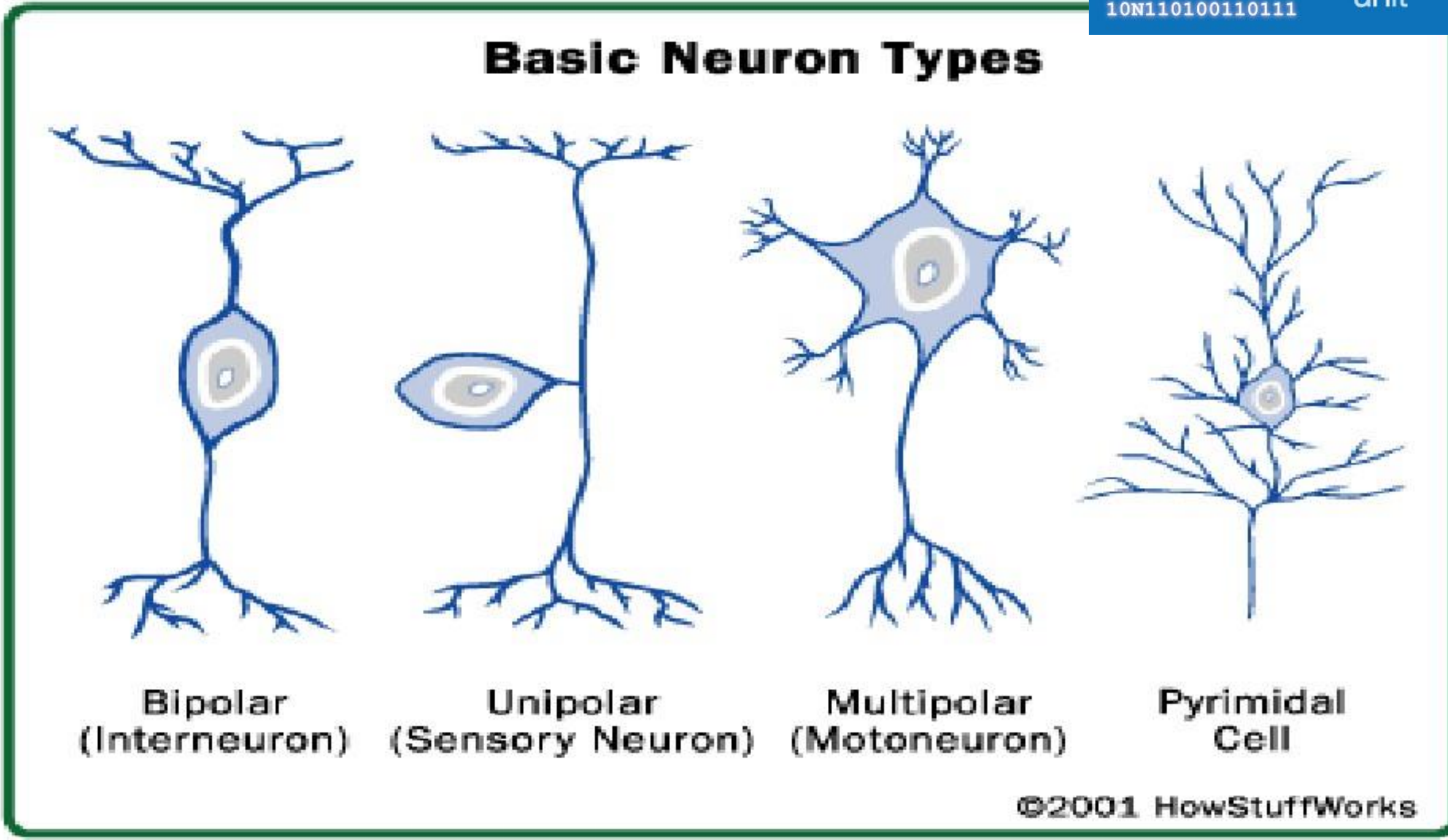
Diencephalon (Thalamus) - Inputs to cortex
Diencephalon (Hypo-Thalamus) - Control of endocrine system (hormones)

Prosencephalon (Hippocampus) - Learning, long term consolidation
Prosencephalon (Cortex) - Higher cognitive functions

Brodmann's Map

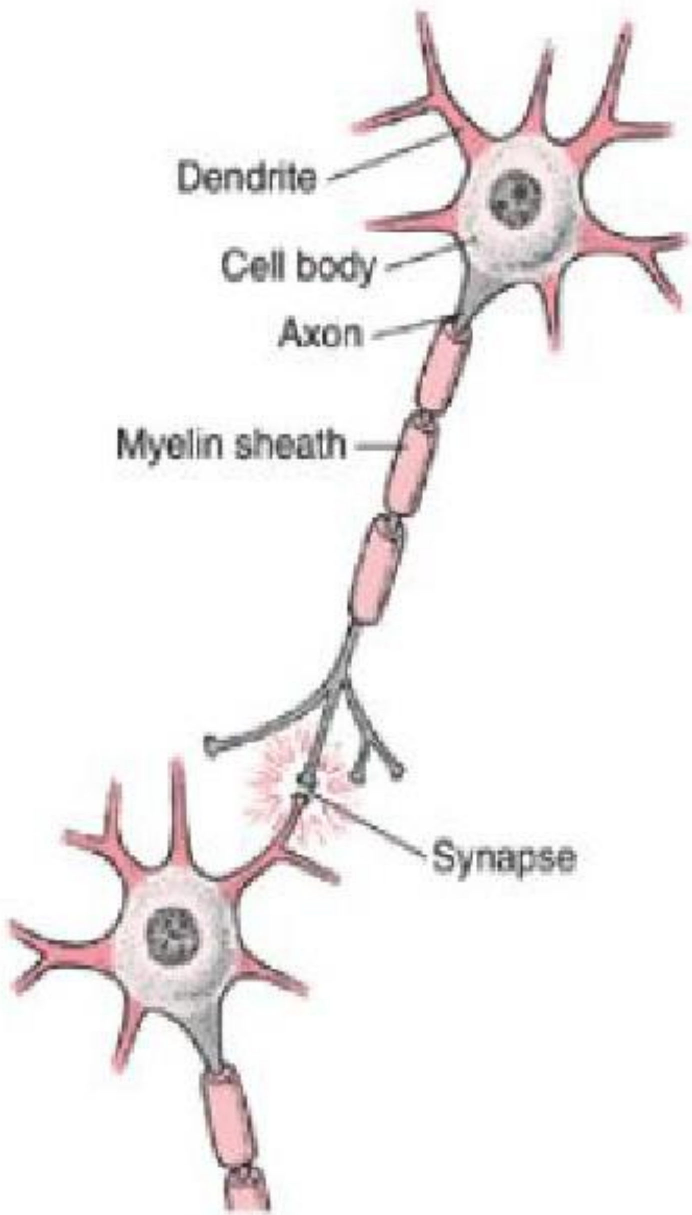
population activity research unit




Basic Neuron Types
Bipolar
(Interneuron)
Unipolar
(Sensory Neuron)
Multipolar
(Motoneuron)
Pyrimidal
Cell
©2001 HowStuffWorks

Anatomy of a Single Neuron

population activity research unit

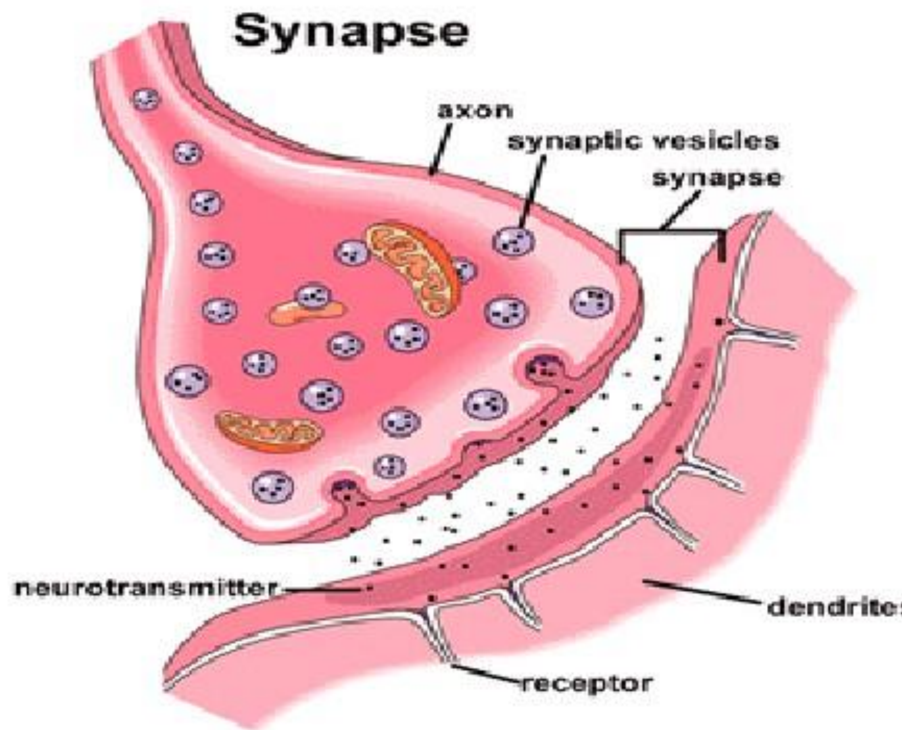


Cell Body (Soma) - Contains DNA, cell nucleus

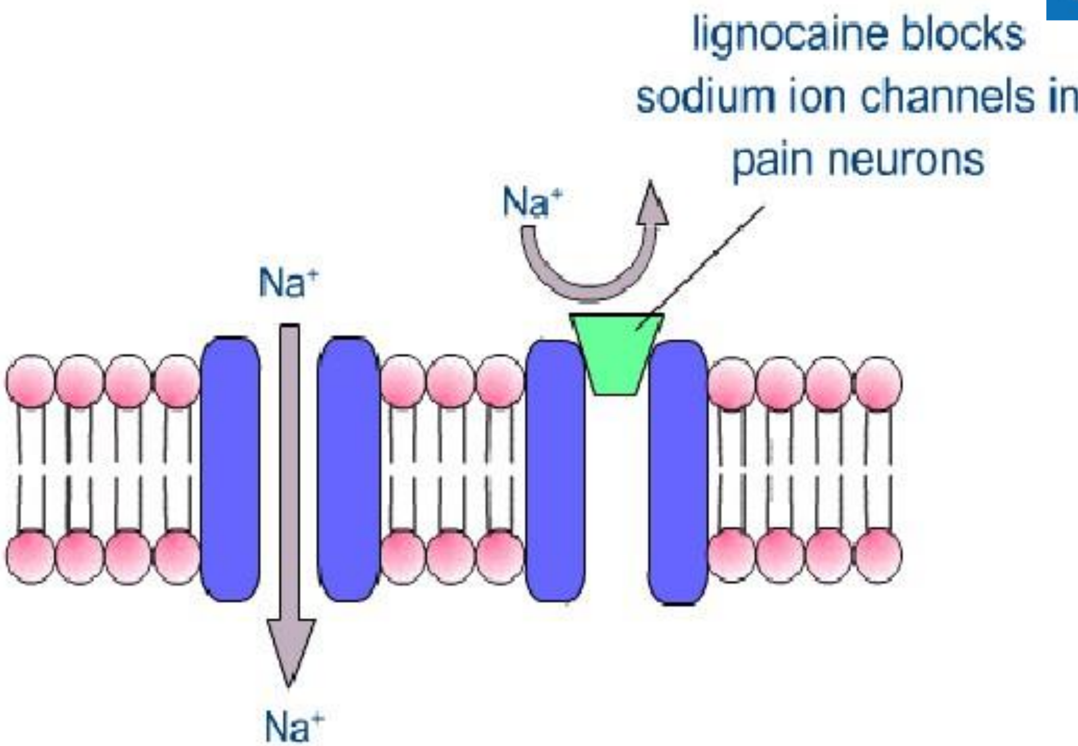
Dendrite - Brings input to the neuron

Axon - The output of a neuron

Synapse - Transfer the information from one cell to another



Ion Channels



Physiological Specialization - Membrane spanning ion channels

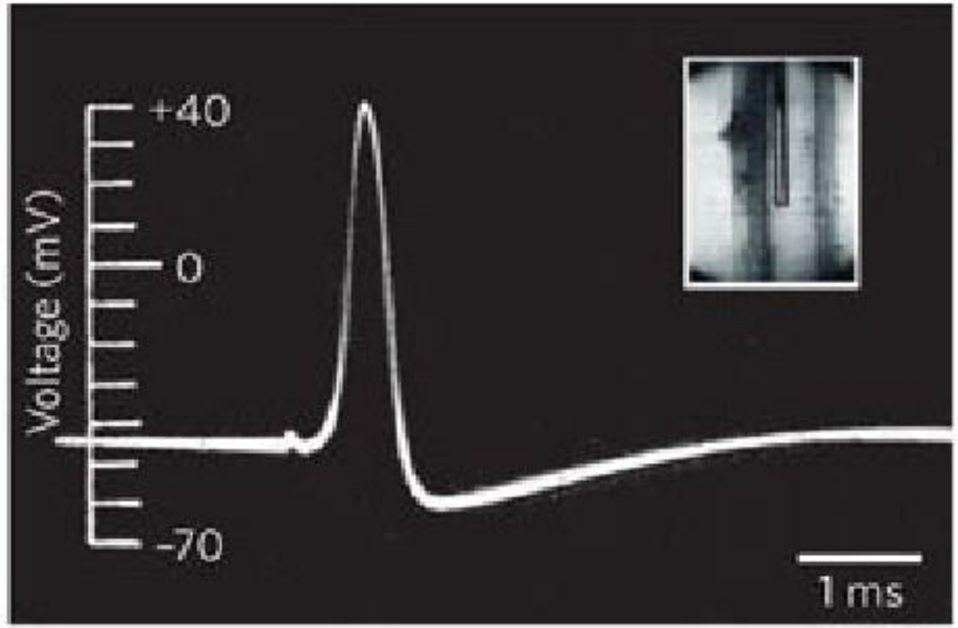
Allow ions to move in and out of cell

Work by opening and closing in response to voltage changes

Most prominent being Na+, K+, Ca2+, Cl-

Spikes - Units of Information

population activity research unit



If a neuron is **depolarized** sufficiently to raise the membrane potential above a threshold level, a **spike or action potential** is generated

Roughly 100 mV fluctuation, lasts for ~ 1ms

Of great importance, as spikes can **propagate over large distances**

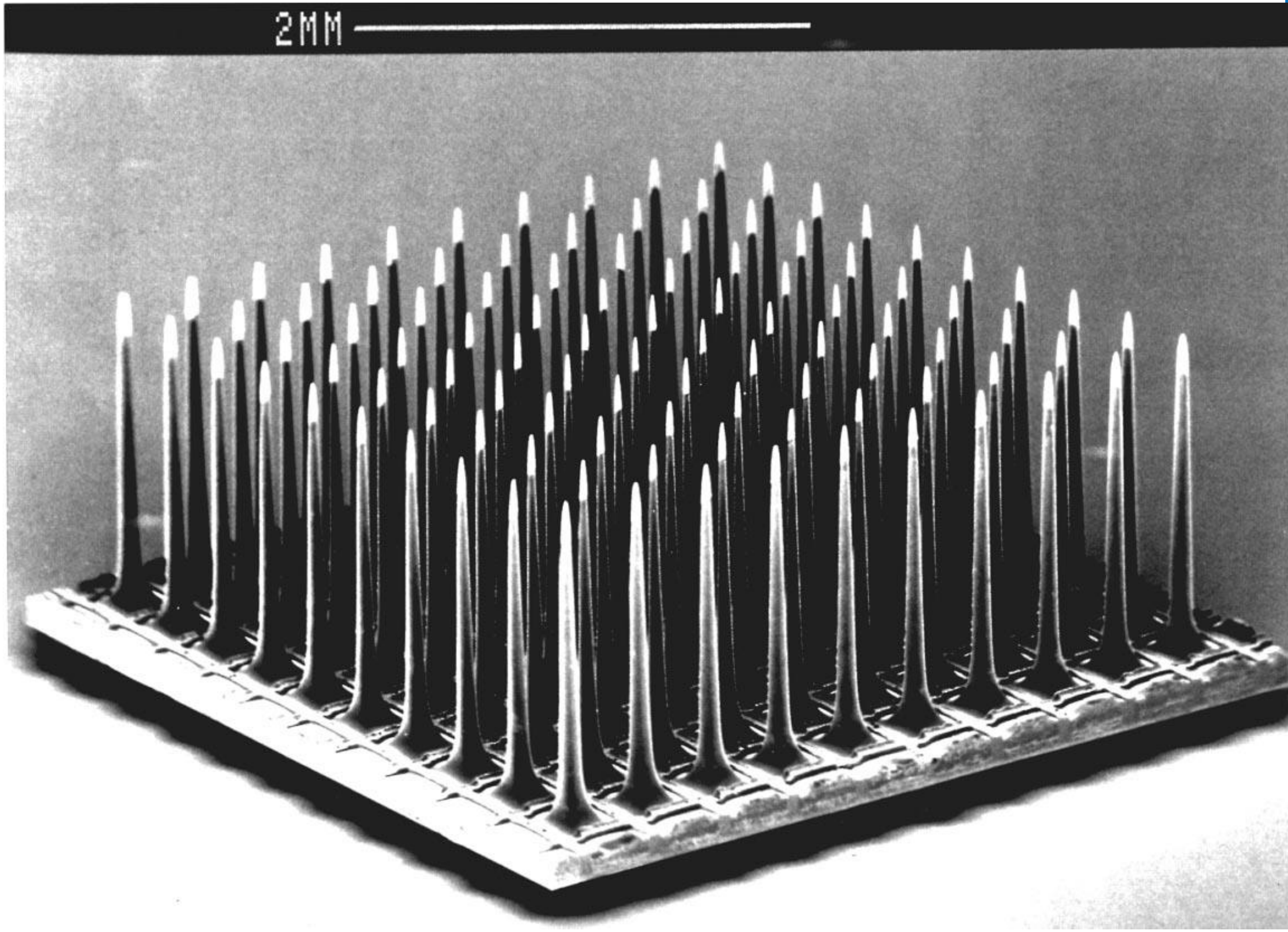
A spike causes a neuron to go into refractory period

Rough Estimates

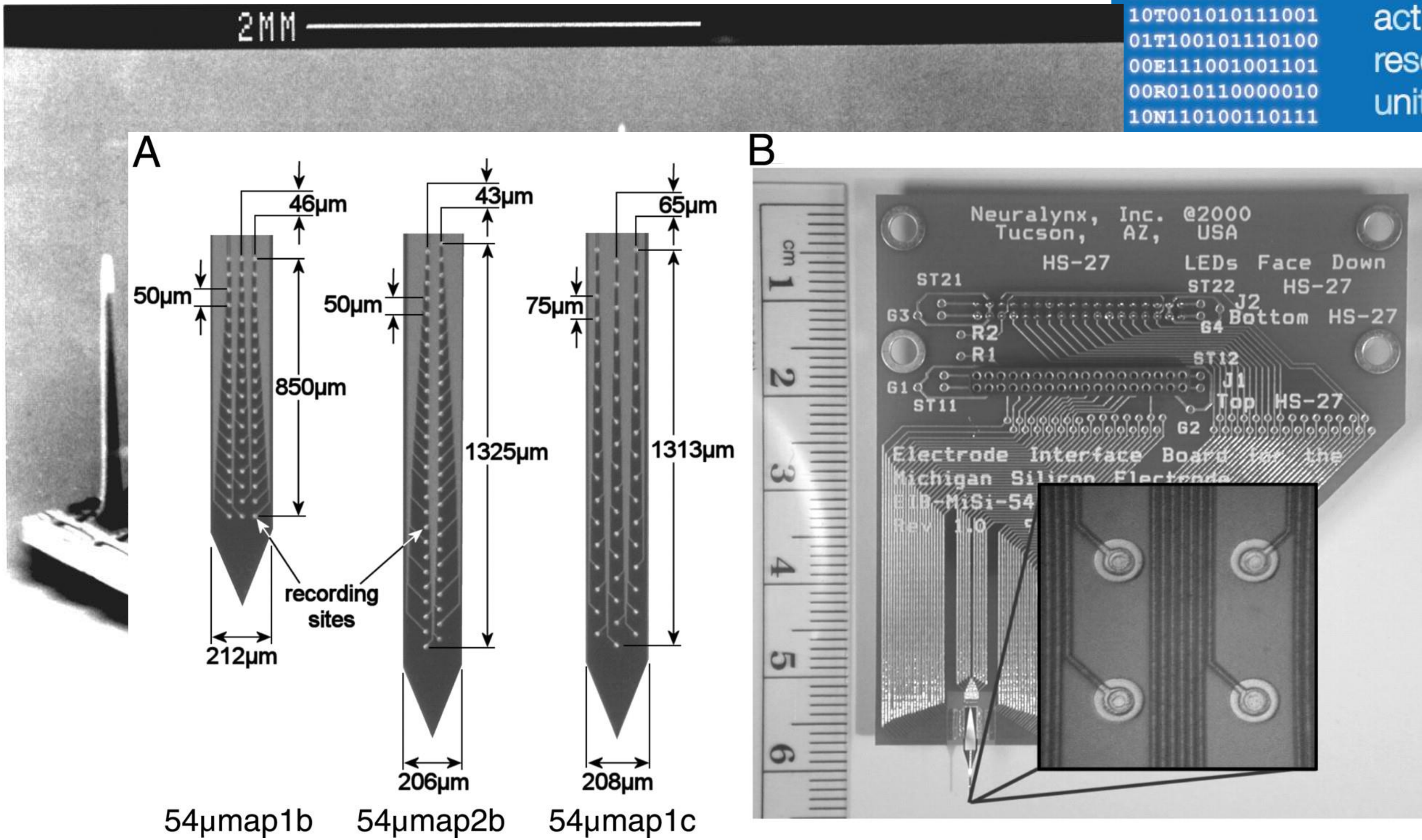
In human brain there are:

~ 10^{12} (trillion) Neurons
~ 10^{15} (quadrillion) Synapses
~ 10^{5} Neurons/mm^3
~ 10^{9} Synapse/mm^3
~ 4 Km Axon/mm^3
~ 500 million dendrites /mm^3
~ 10^{4} Input Synapses / neuron

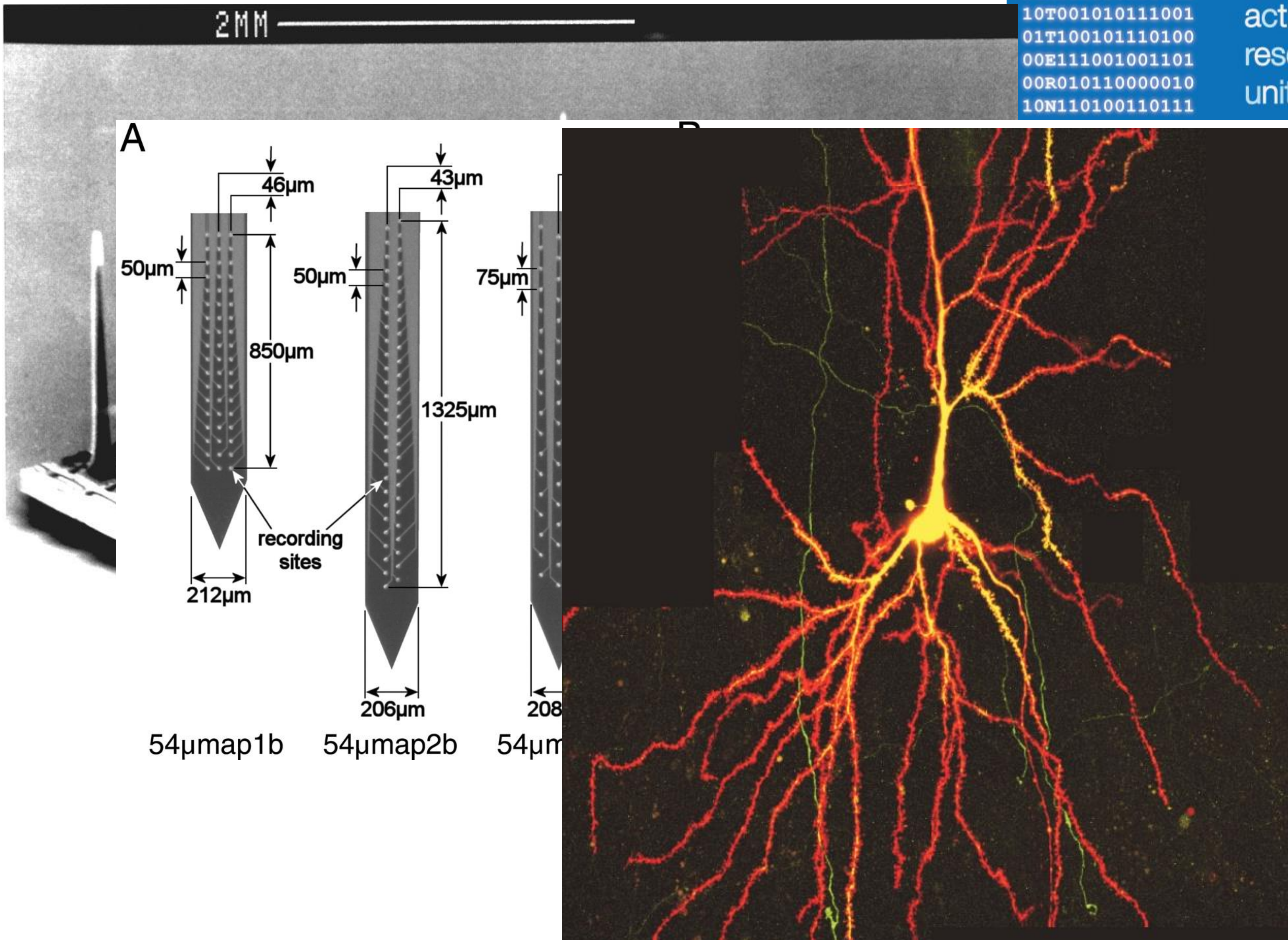
population
activity
research
unit



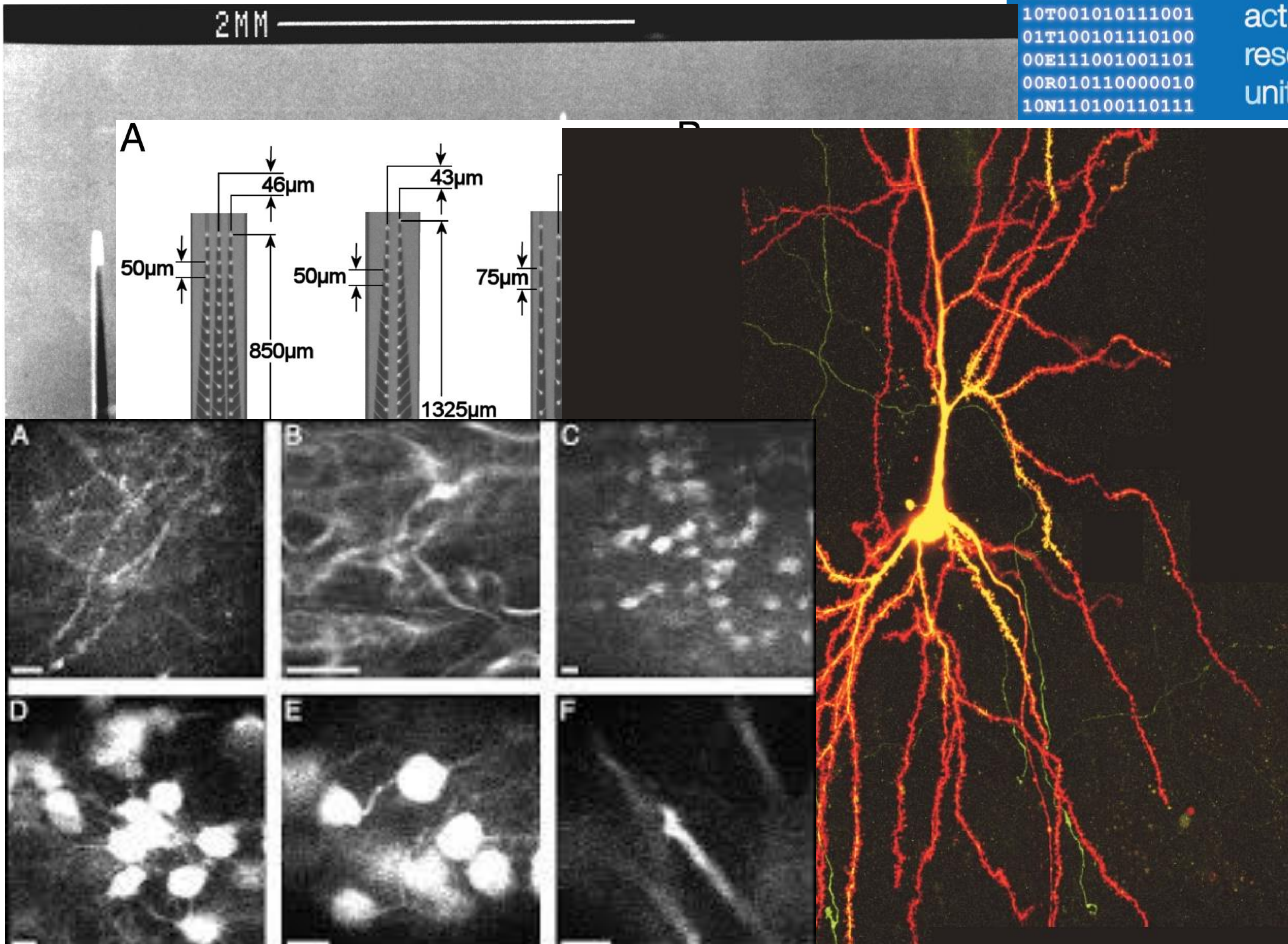
population
activity
research
unit



population activity research unit

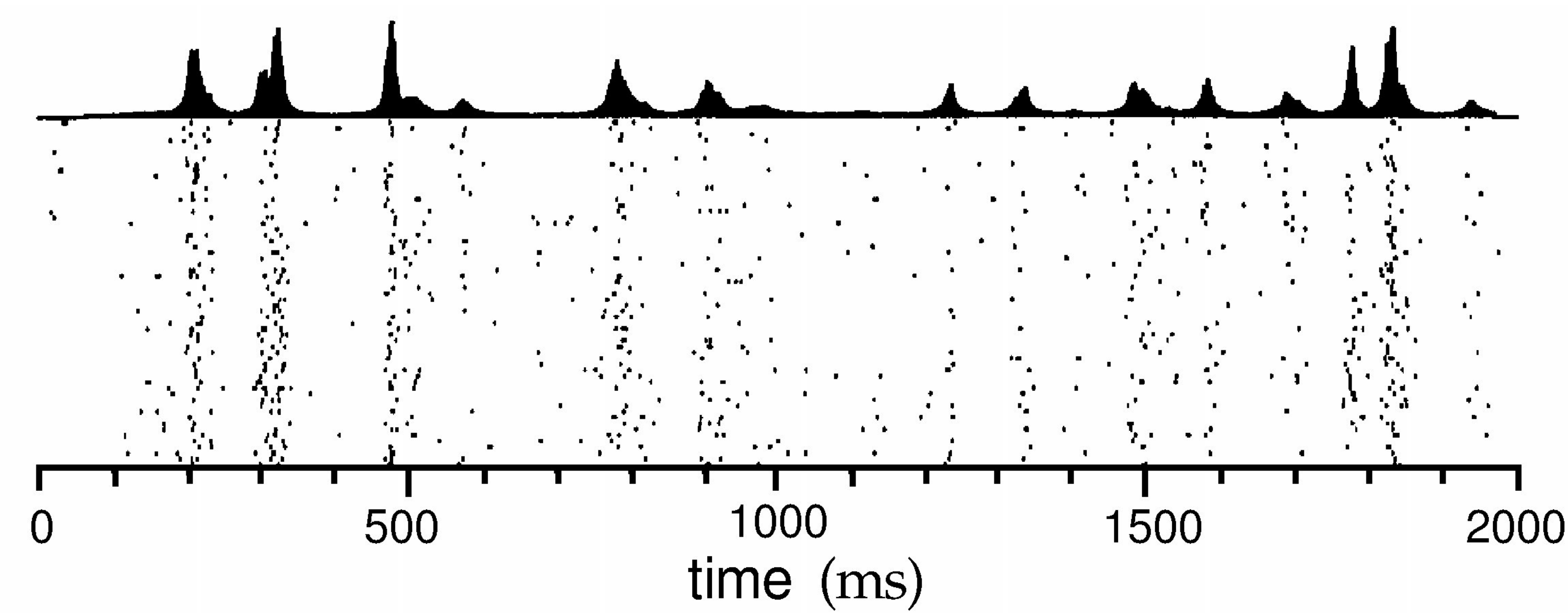


population activity research unit



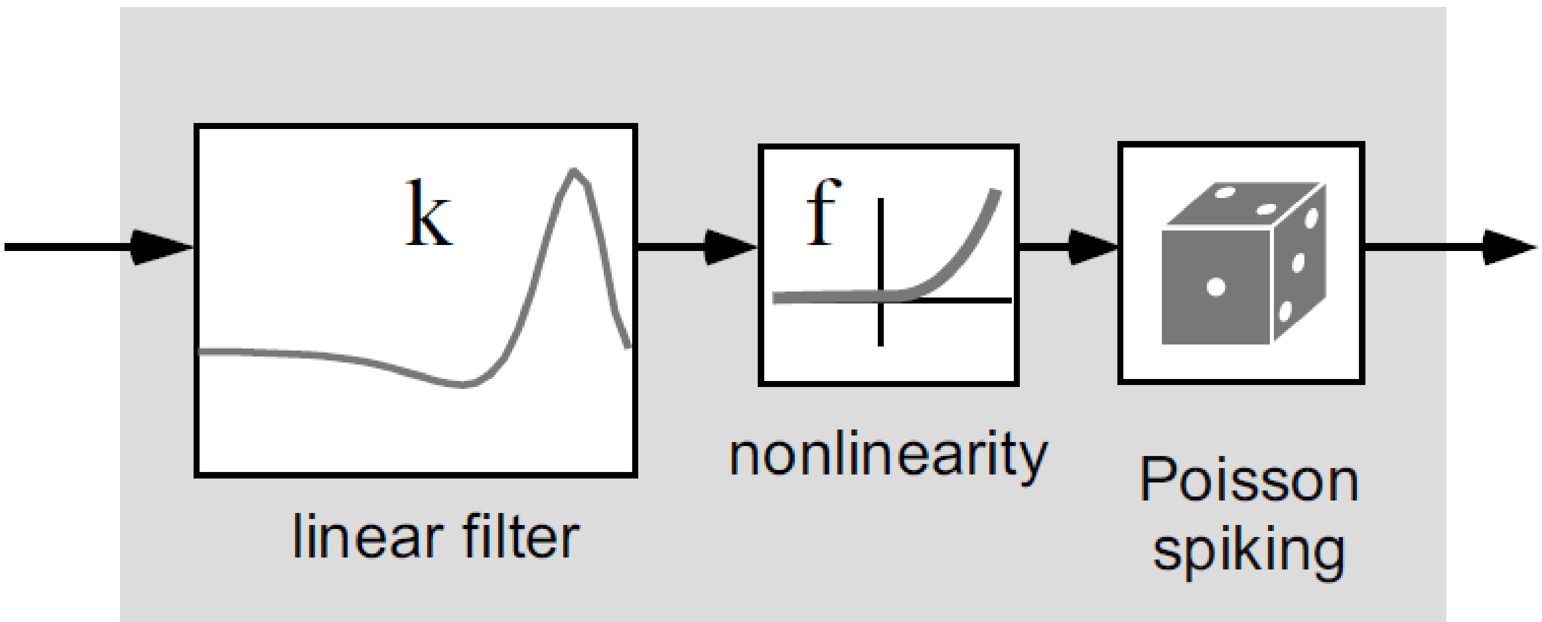
Spike train, spike raster

population activity research unit

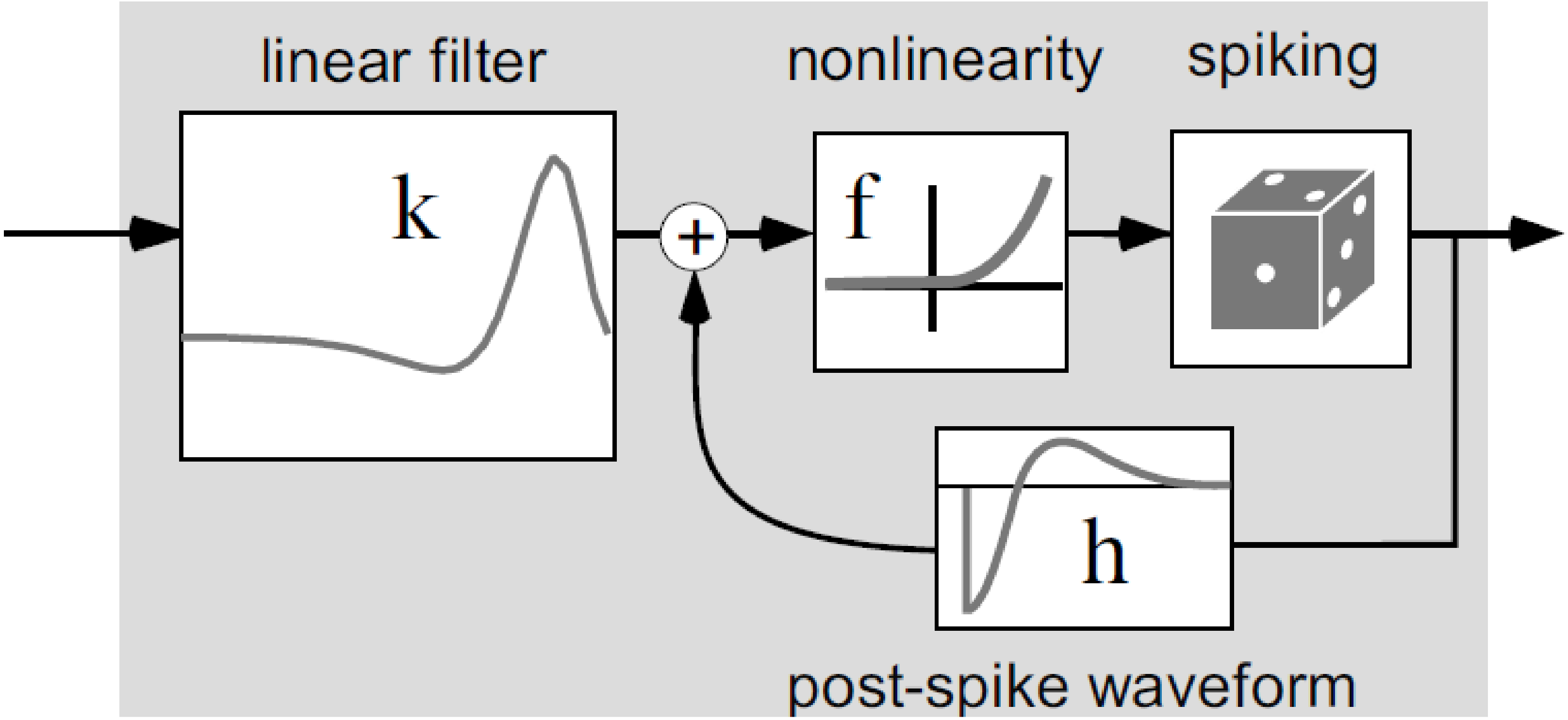


population activity research unit

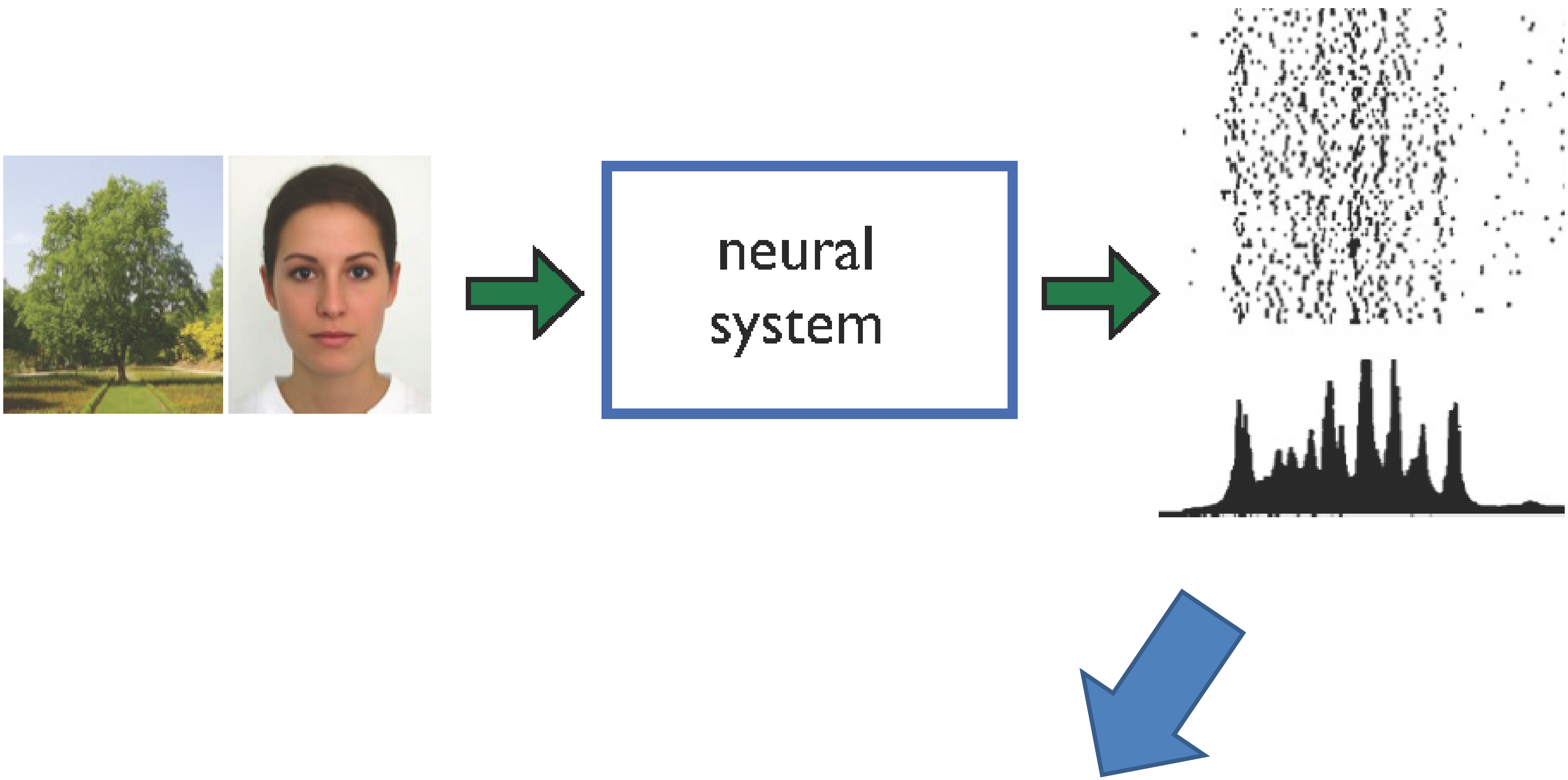
LNP model



Generalized linear model



Neural Encoding

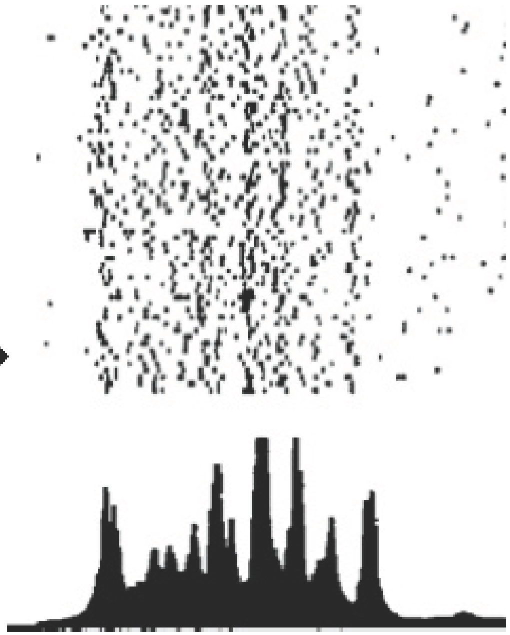


a –action/respond

Neural Encoding

population activity research unit

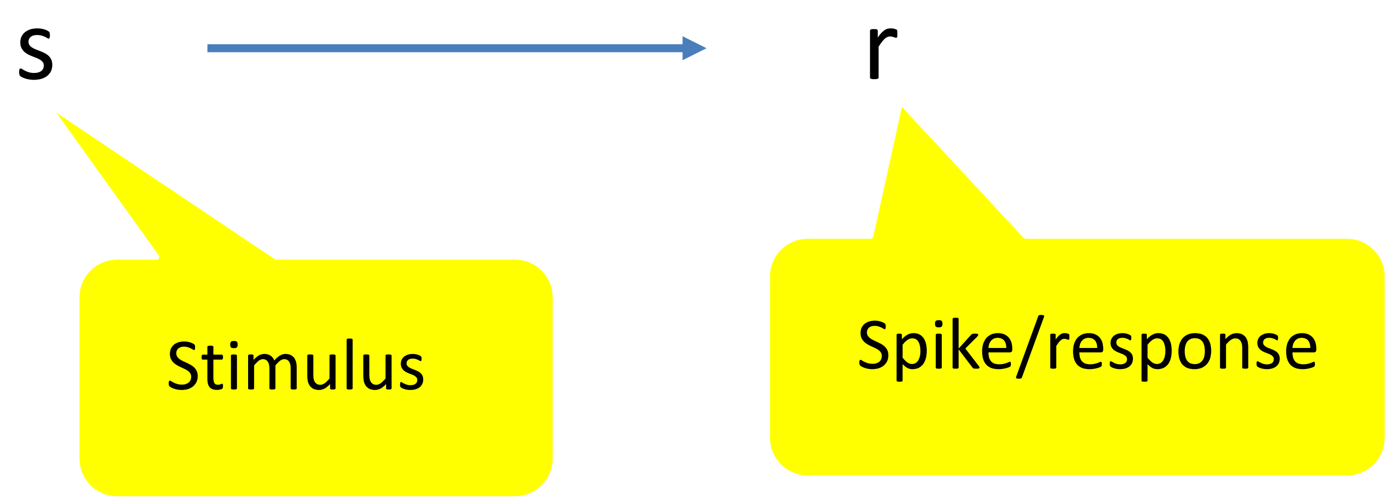
neural system



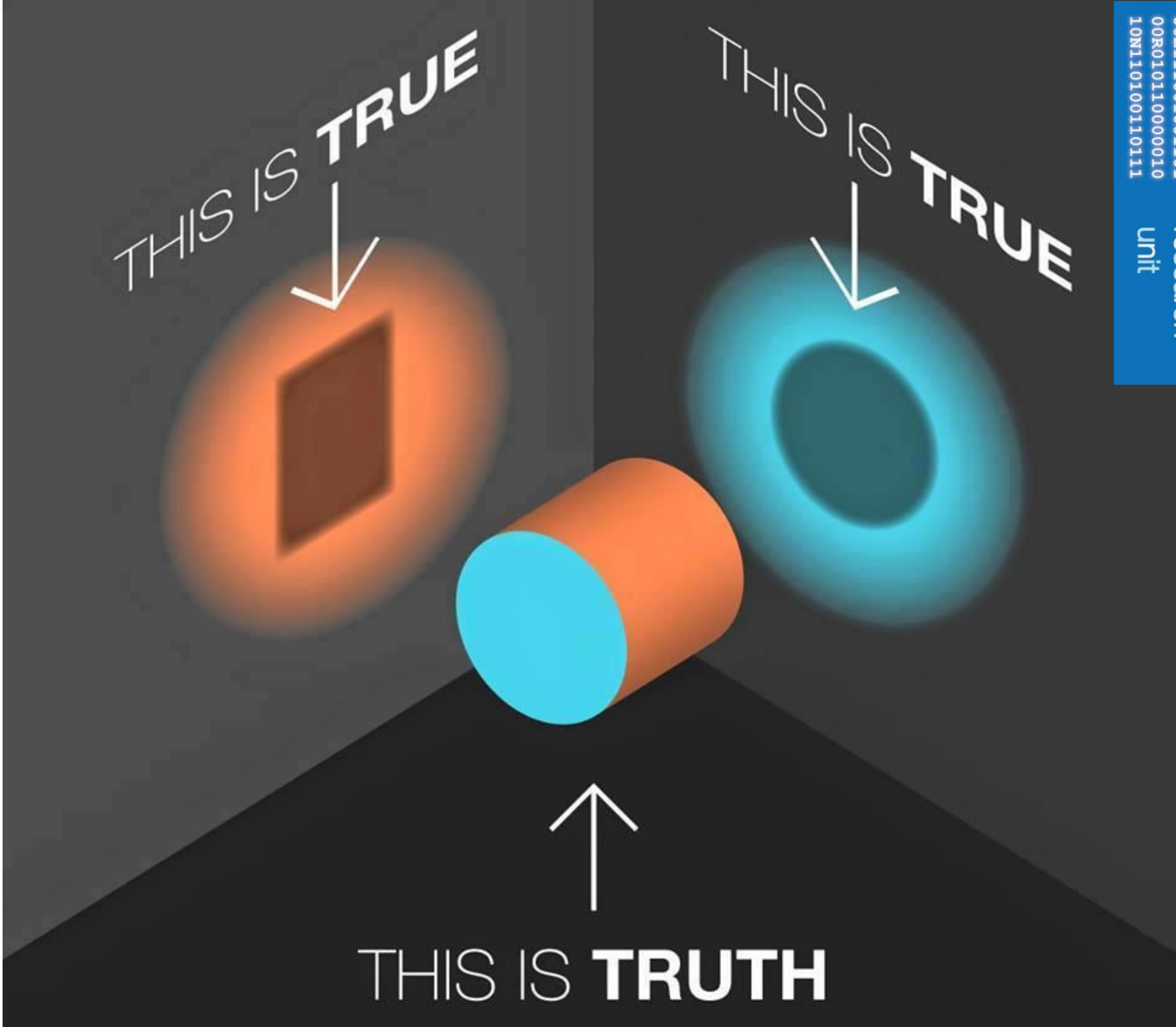
Representation learning ⟷ encoding decoding

population activity research unit

Task –
learning about the black box



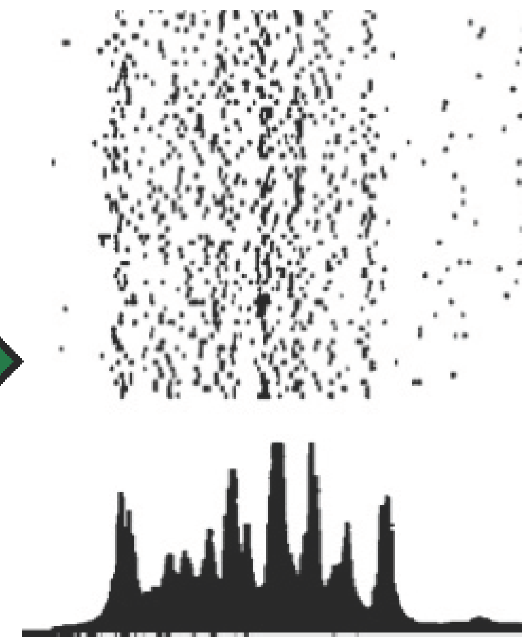
$$r(t) = r(s(t-\tau, \ldots, t))$$


THIS IS TRUE
THIS IS TRUE
THIS IS TRUTH

Neural Encoding

population activity research unit

neural system



sensory stimuli

$\mathbf{s}$

$p(\mathbf{s})$ prior distribution of stimuli

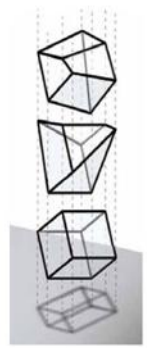
neural responses

$\mathbf{r}$

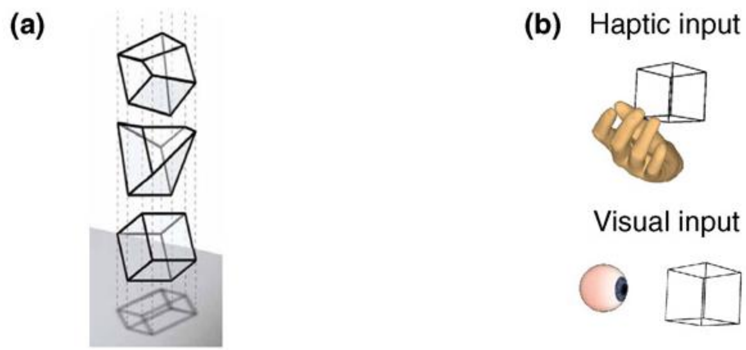
$p(\mathbf{r}|\mathbf{s})$ conditional dist. of responses

what probabilistic representations are good for

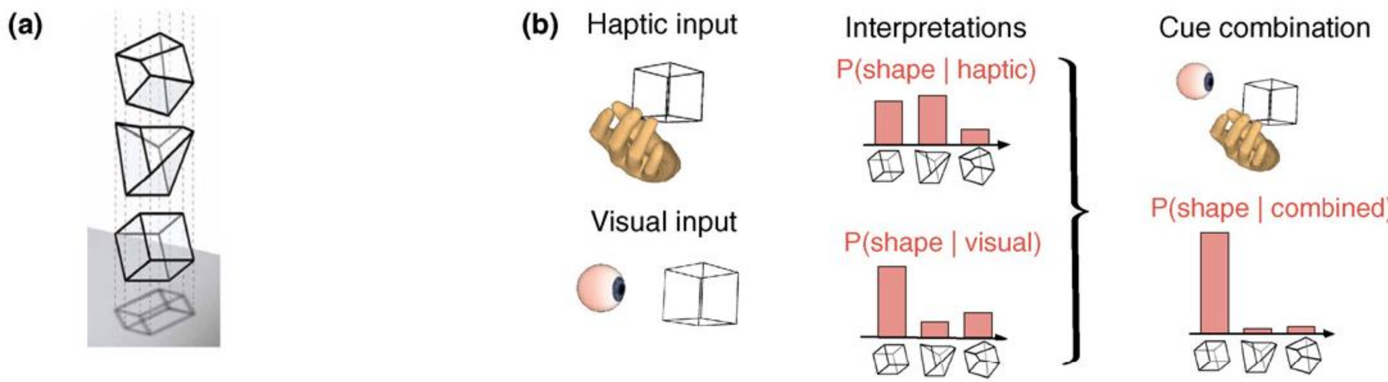
(a)



what probabilistic representations are good for

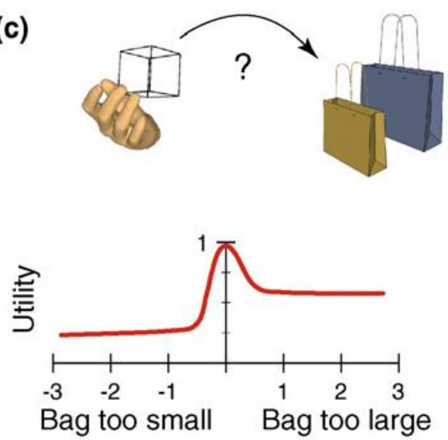


what probabilistic representations are good for



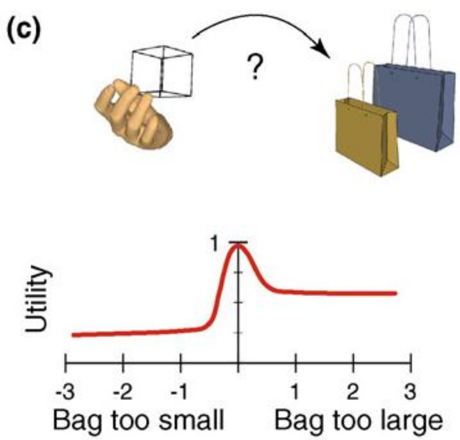
what probabilistic representations are good for

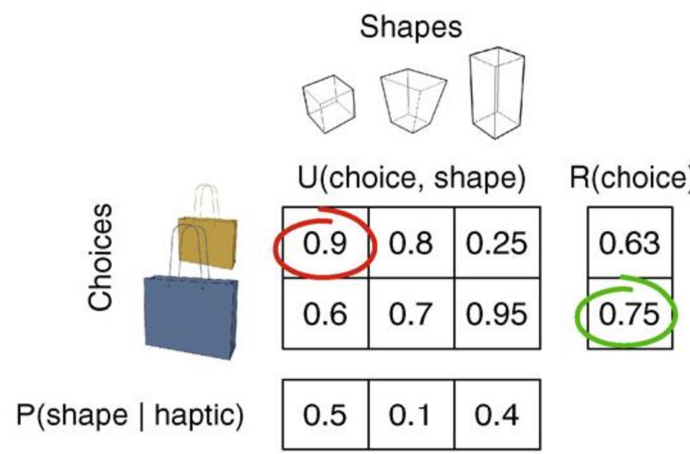
(c)



population activity research unit

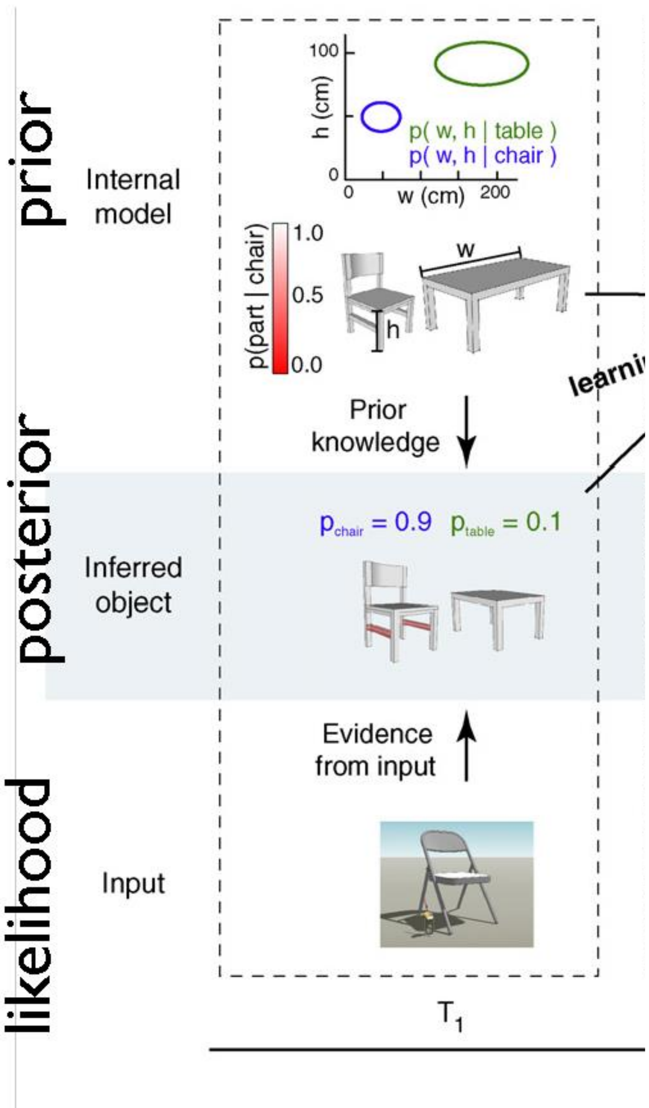
what probabilistic representations are good for





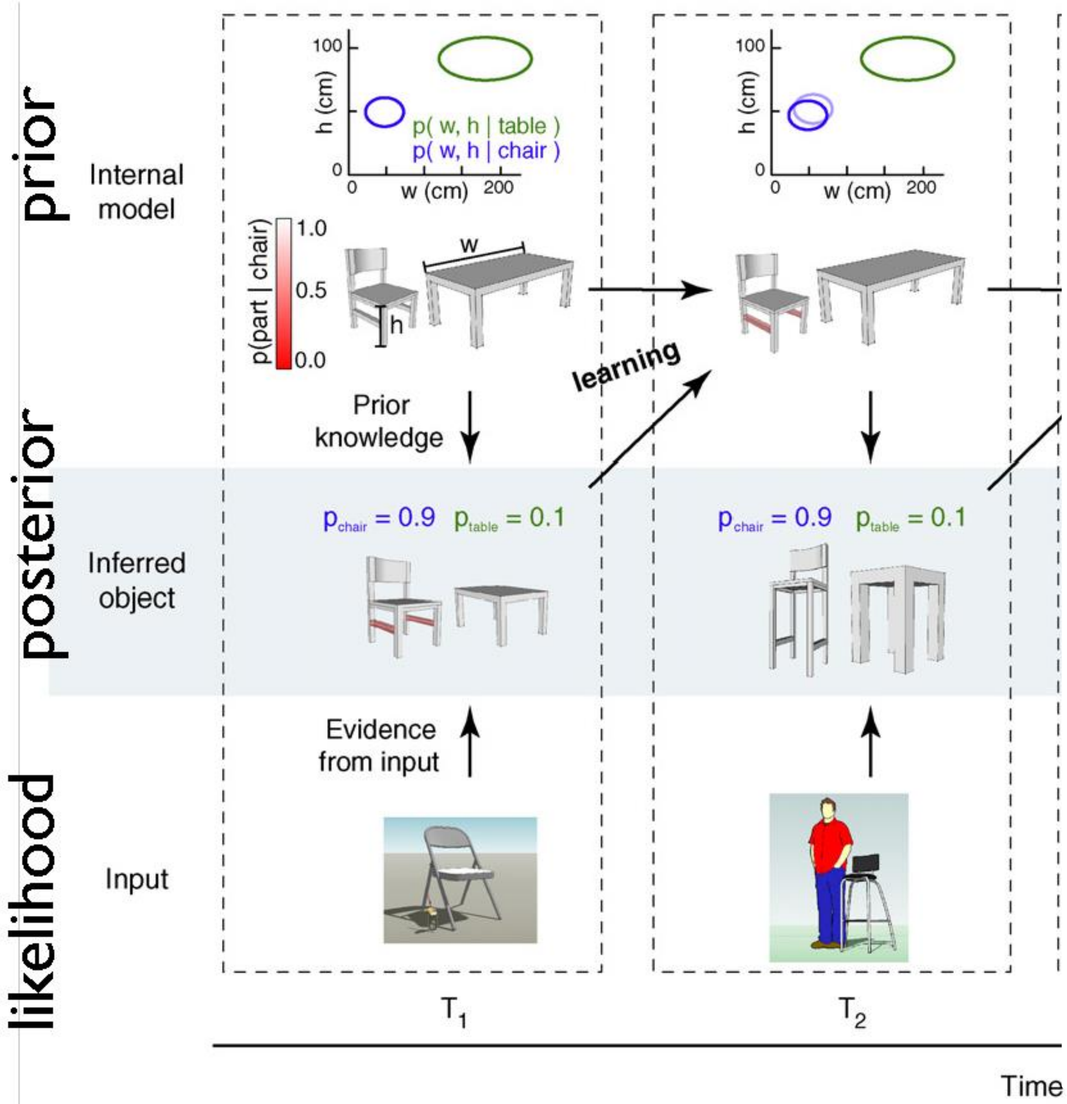
population activity research unit

learning in a probabilistic model

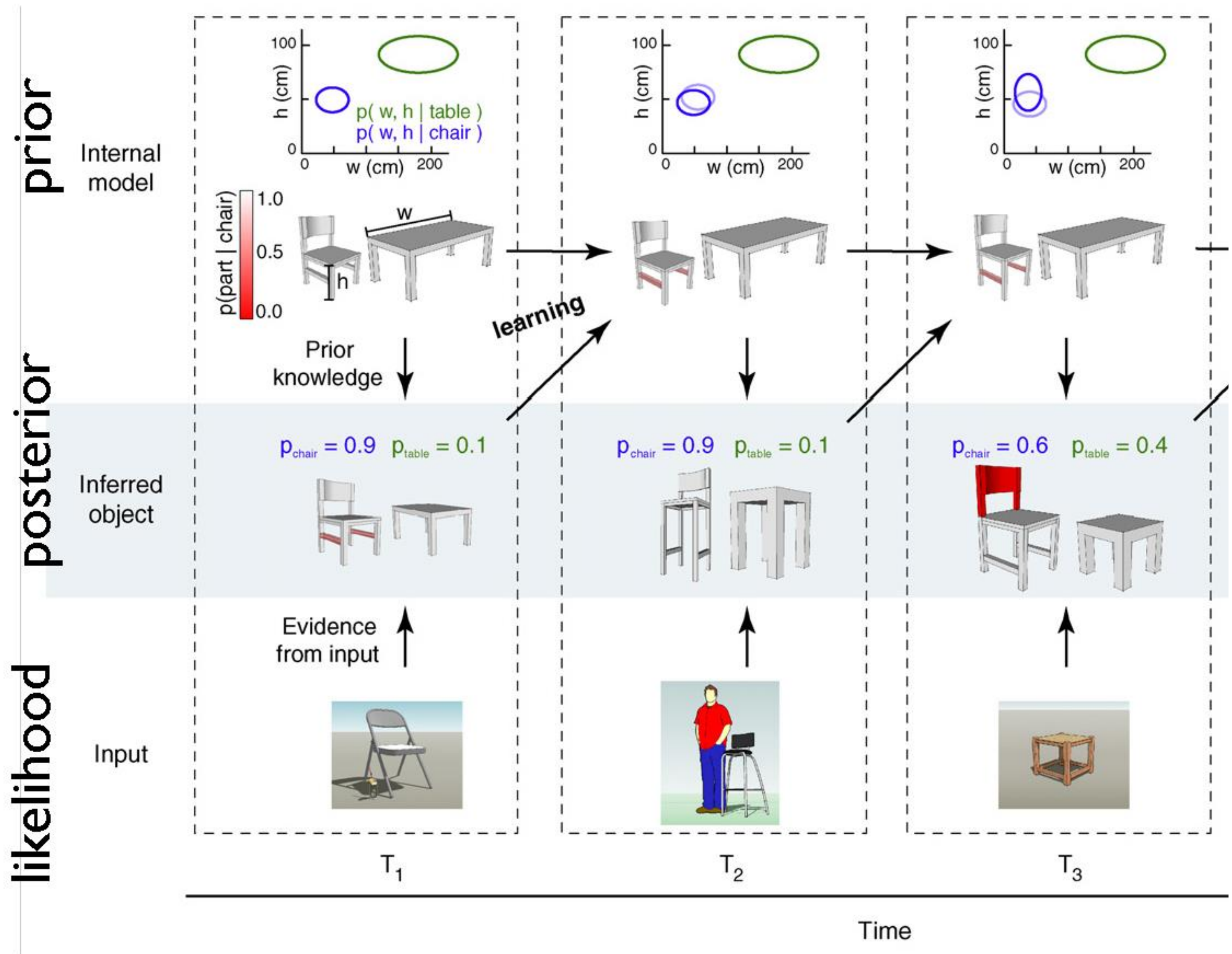


learning in a probabilistic model

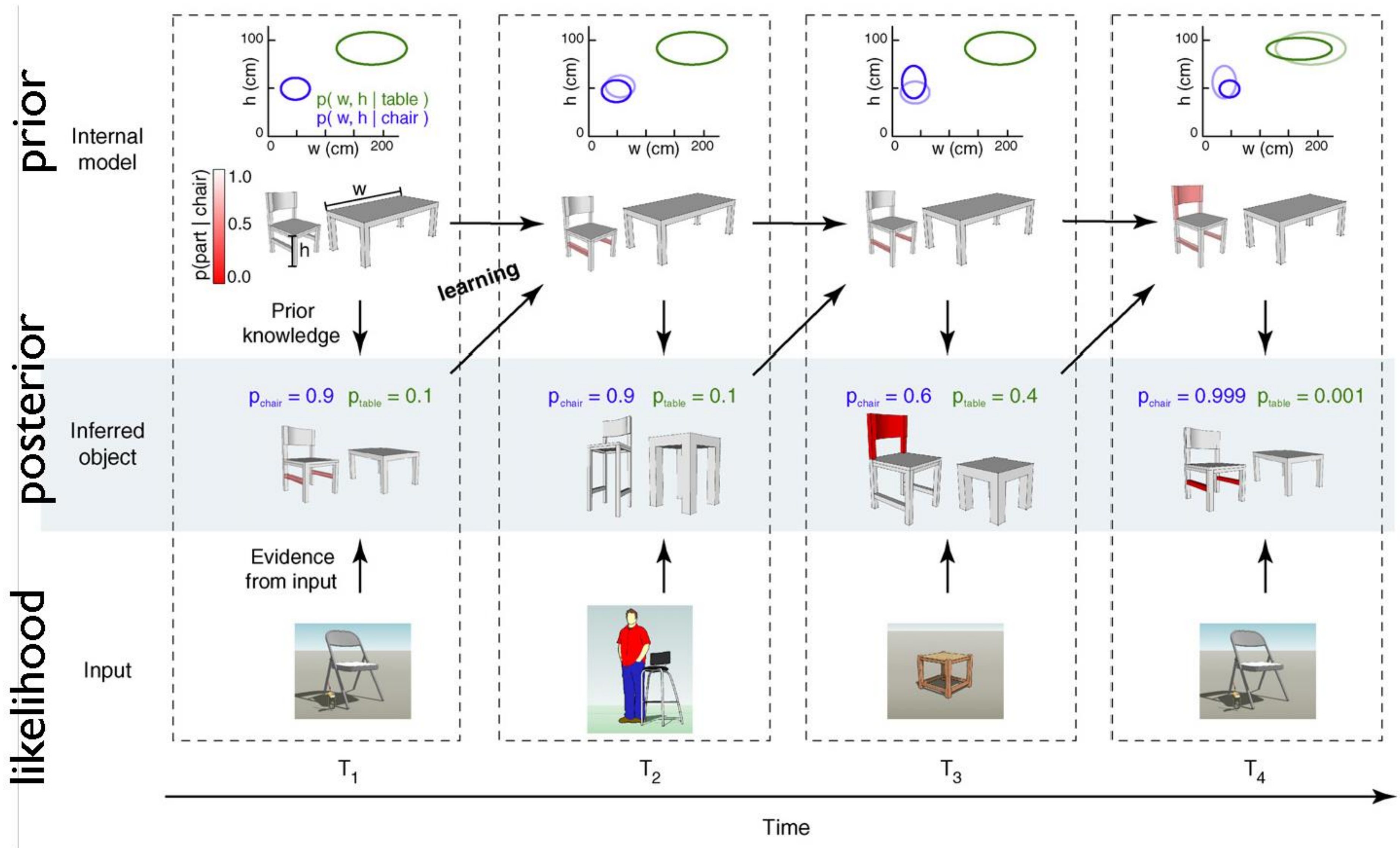
population activity research unit



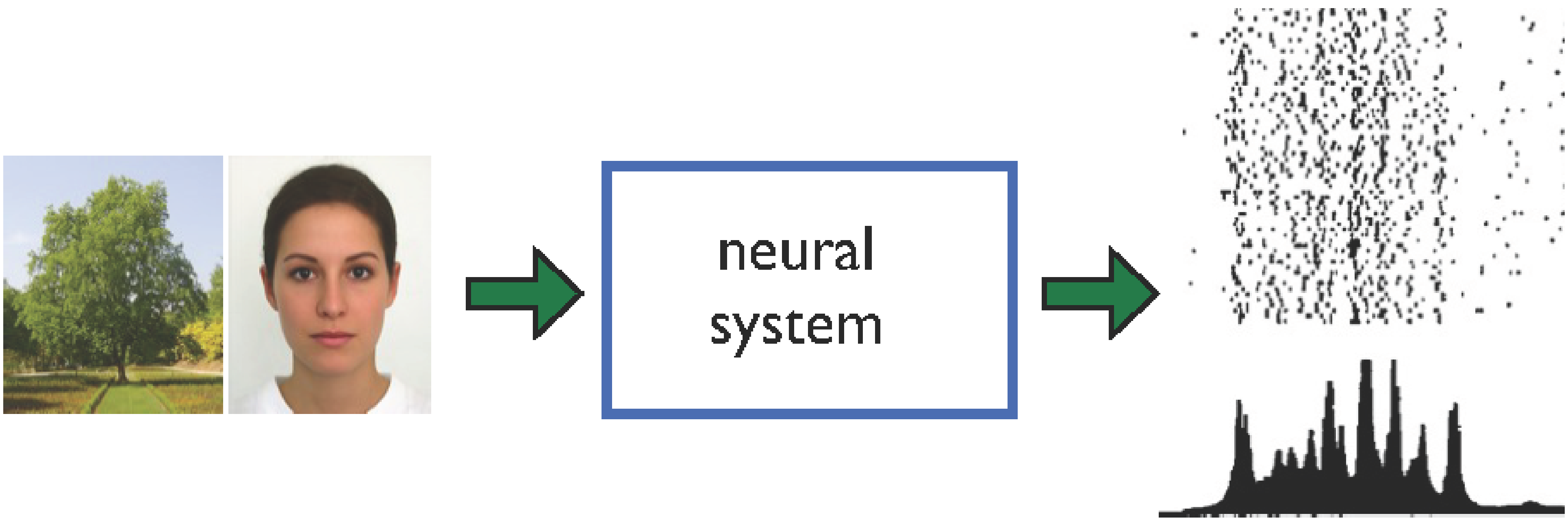
learning in a probabilistic model



learning in a probabilistic model



Neural Encoding



Neural Encoding/Decoding

joint probability

$$p(\mathbf{r}, \mathbf{s}) = p(\mathbf{r}|\mathbf{s})p(\mathbf{s})$$

sensory stimuli

$$\mathbf{s}$$

$p(\mathbf{s})$ prior distribution of stimuli

neural responses

$$\mathbf{r}$$

$p(\mathbf{r}|\mathbf{s})$ conditional dist. of responses

Neural Encoding/Decoding

joint probability

$$p(\mathbf{r}, \mathbf{s}) = p(\mathbf{r}|\mathbf{s})p(\mathbf{s})$$
$$= p(\mathbf{s}|\mathbf{r})p(\mathbf{r})$$

Neural Encoding/Decoding

joint probability

$$p(\mathbf{r}, \mathbf{s}) = p(\mathbf{r}|\mathbf{s})p(\mathbf{s})$$
$$= p(\mathbf{s}|\mathbf{r})p(\mathbf{r})$$

Bayes' theorem

$$p(\mathbf{s}|\mathbf{r}) = \frac{p(\mathbf{r}|\mathbf{s})p(\mathbf{s})}{p(\mathbf{r})}$$

prior

posterior

Neural Encoding/Decoding

joint probability

$$p(\mathbf{r}, \mathbf{s}) = p(\mathbf{r}|\mathbf{s})p(\mathbf{s})$$
$$= p(\mathbf{s}|\mathbf{r})p(\mathbf{r})$$

Bayes' theorem

$$p(\mathbf{s}|\mathbf{r}) = \frac{p(\mathbf{r}|\mathbf{s})p(\mathbf{s})}{p(\mathbf{r})}$$

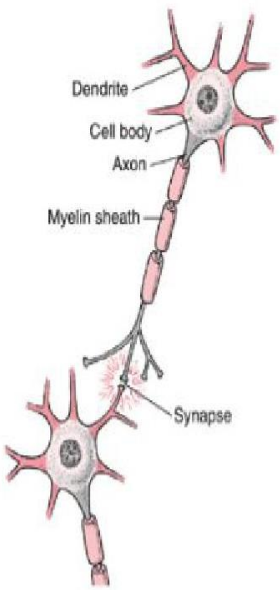
posterior

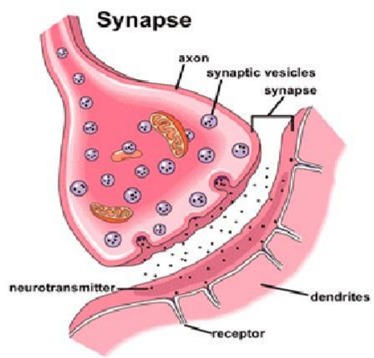
normalization factor

prior

$$p(\mathbf{r}) = \int d\mathbf{s}\, p(\mathbf{r}|\mathbf{s})p(\mathbf{s})$$

Synaptic weights

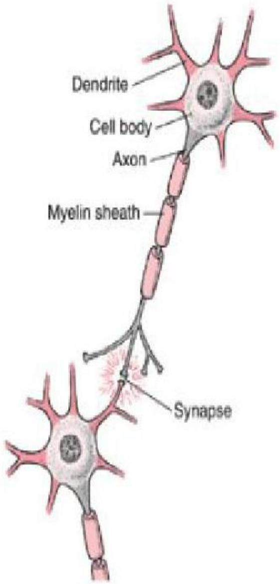


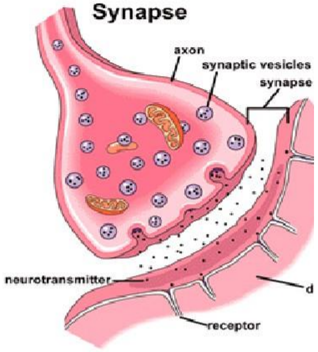


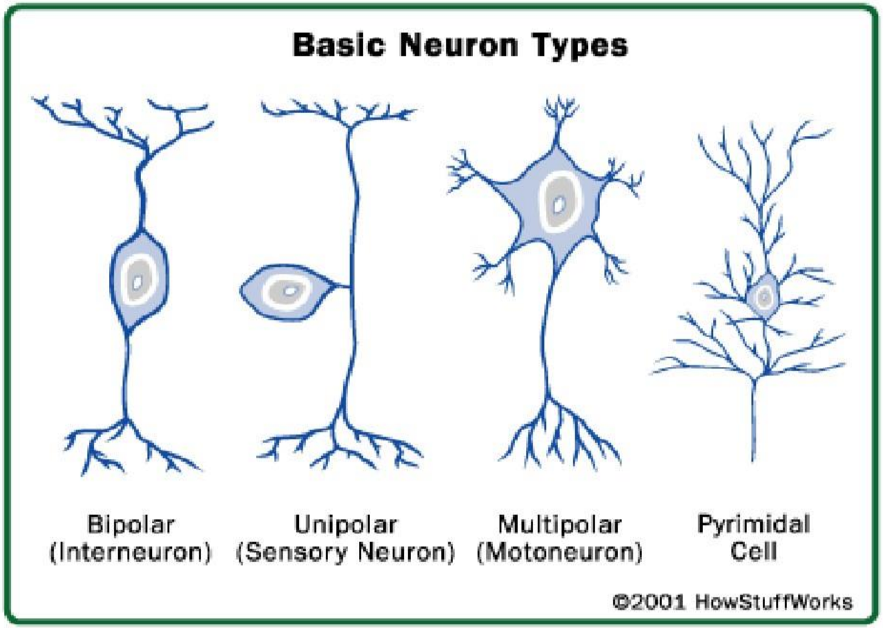
Synaptic weights

10P110100110111
00A011000001011
10T001010111001
01T100101110100
00E111001001101
00R010110000010
10N110100110111

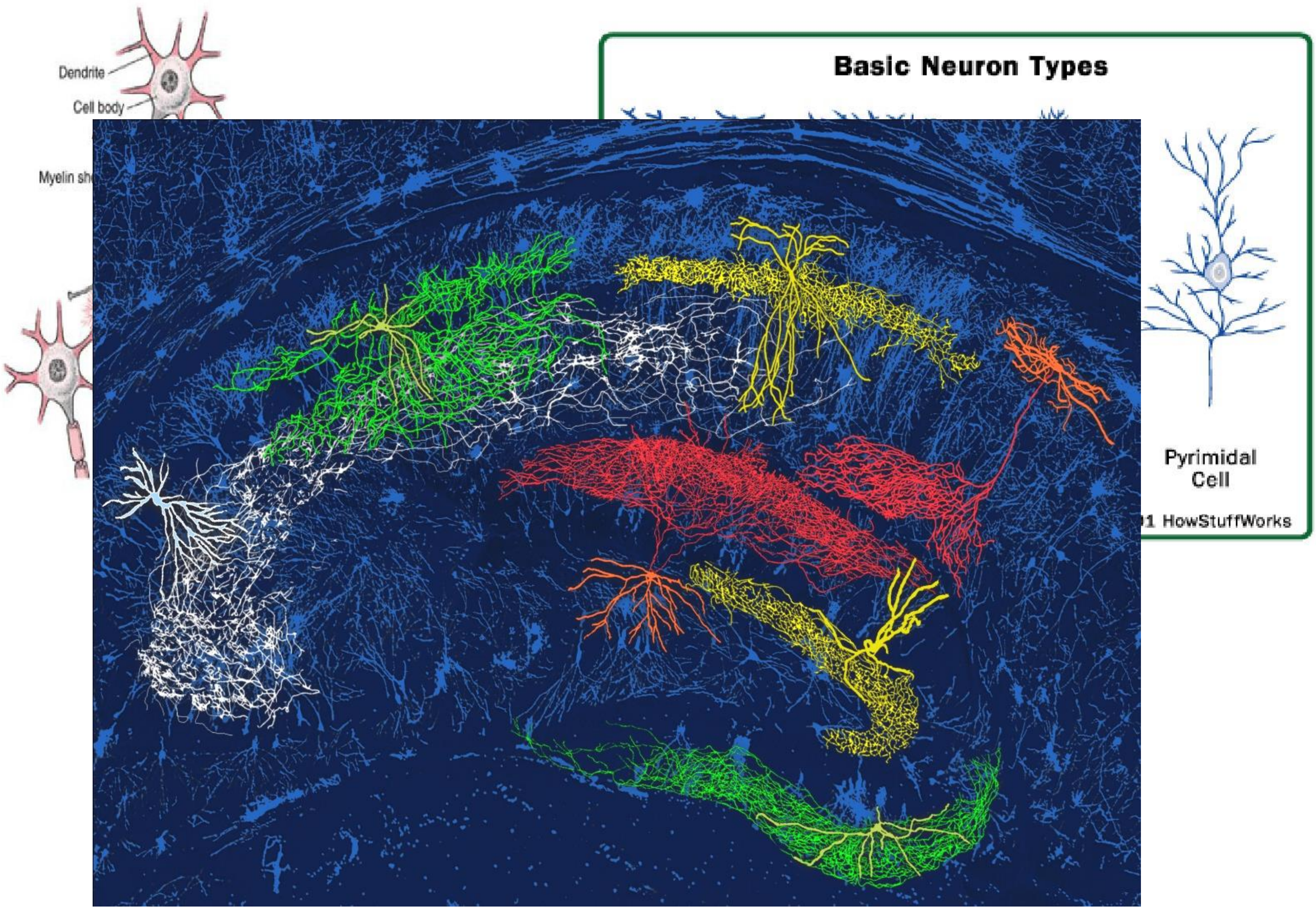
population activity research unit







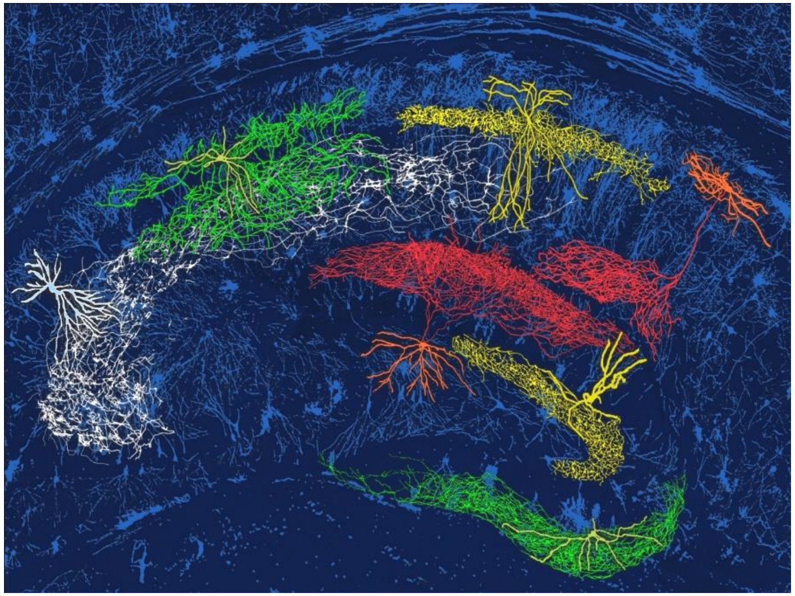
Synaptic weights



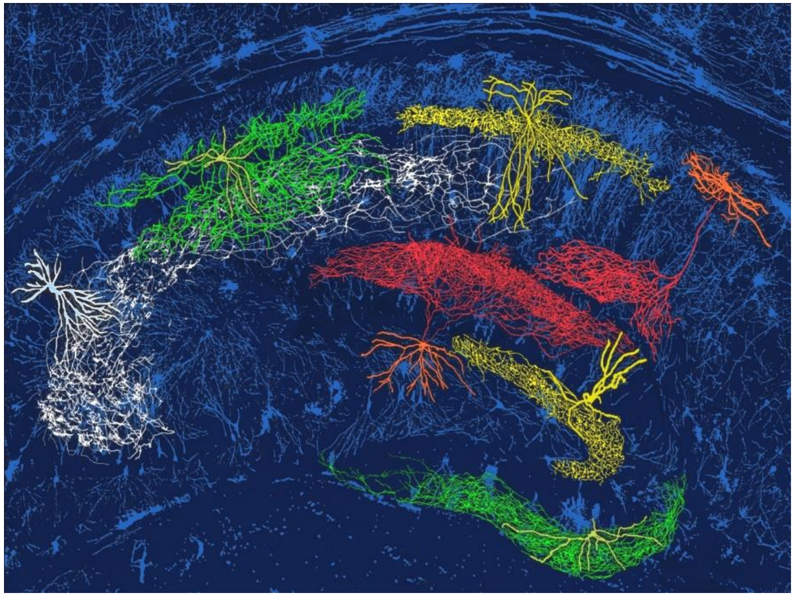
population activity research unit

Synaptic weights

If neuron A frequently contributes to fire neuron B then the synaptic connection /weight strengthen/increased.



Synaptic weights



If neuron A frequently contributes to fire neuron B then the synaptic connection /weight strengthen/increased.

- synaptic plasticity

How to store/represent and sample a distribution?

$$p\left(z_1, \ldots, z_k\right)$$

Joint distr. of k binary variables.

population activity research unit

How to store/represent and sample a distribution?

$$p\left(z_1,\ldots,z_k\right)$$

Joint distr. of k binary variables.
We construct a network $\mathscr{N}$ of k spiking neurons ν_i $i=1,\ldots,k$ which produce sample from p .

The firing of $\mathscr{N}$ will be a non-reversible MC.

Neural representation

Spike/action potential s_i with 1 ms duration

Influence on itself and other neurons
last 5-100 ms – refactory period

Neural representation

Spike/action potential s_i with 1 ms duration

Influence on itself and other neurons
last 5-100 ms – refactory period

Time is discretized. Unit is small, an integer time constant τ is defined multiple of that unit.

Neural representation

Spike/action potential s_i with 1 ms duration

Influence on itself and other neurons
last 5-100 ms – refractory period

Time is discretized. Unit is small, an integer time constant τ is defined multiple of that unit.

$$z_i(t) = 1 \Leftrightarrow s_i \quad \textit{fired in} \quad t \in (t - \tau, t]$$

i.e. The rv is set to 1 for duration of τ, for the refractory period of the neuron.

Neural representation

The Markovity needs the bookkeeping of the times:
We define

$$\left(\zeta_1,\ldots\zeta_k\right)\in\left(R^+\right)^k$$

for each *i* ζ_i denote the time when the neuron spike whitin the interval,

Neural representation

The Markovity needs the bookkeeping of the times:
We define

$$\left(\zeta_1,\ldots\zeta_k\right) \in \left(R^+\right)^k$$

for each *i* ζ_i denote the time when the neuron spike whitin the interval,

$$(t-\tau, t]$$

and setting $z_i = 1$ at time t.

Neural representation

The Markovity needs the bookkeeping of the times:
We define

$$\left(\zeta_1,\ldots\zeta_k\right)\in\left(R^+\right)^k$$

for each *i* ζ_i denote the time when the neuron spike whitin the interval,

$$(t-\tau,t]$$

and setting $z_i = 1$ at time t.

ζ_i time count down: at the spike, the var*iable* ζ_i is set to τ then decreasing linearly above 0.

Neural representation of a pd

The membrane potential of the *i*-th neuron is

$u_i(t)$

at time *t*.

Examples: Boltzmann distributions

$$p(z) = \frac{1}{Z} \exp\left(-\sum_{i,j} w_{i,j} z_i z_j - \sum_i b_i z_i \right)$$

w, b parameters

$w_{i,j} = w_{j,i}$, $w_{i,i} = 0$,

Neural representation of a pd

$$u_i(t) = b_i + \sum_j w_{i,j} z_j(t)$$

b_i individual bias (excitability)

$w_{i,j}$ synaptic weight from i to j

MC in discrete time

The time variable ζ_i *when i* spikes is set to τ (>0 integer) .

The neuron *i* can spike iff $\zeta_i <= 1$. *If* $\zeta_i > 1$ the neuron in refractory period and at each time step ζ_i is decreased by one.

MC in discrete time

We define p(ζ,*z) so that*

$$p(z) = \sum_{\zeta} p(\zeta, z)$$

and the Markov transition to ensure that *p(z)* is the stationary distribution

MC in discrete time

$$p(\zeta, z) = p(\zeta \mid z)\, p(z)$$

$$p(\zeta \mid z) = \prod_{j=1}^{k} p(\zeta_j \mid z_j)$$

$$p(\zeta_j = k-1 \mid \zeta_j = k, z_j = 1) = 1 \qquad \text{if } k>1$$

$$p(\zeta_j = \tau \mid \zeta_j = 0{,}1, z_j = 0) = \sigma(u_j - \log \tau)$$

$$p(\zeta_j = 0 \mid \zeta_j = 1, z_j = 0) = 1 - \sigma(u_j - \log \tau)$$

MC in discrete time

$$p\left(z_j = 1 \mid \zeta_j > 1\right) = 1$$

MC in discrete time

$$p\left(z_j = 1 \mid \zeta_j > 1\right) = 1$$

$$p\left(z_j = 0 \mid \zeta_j = 0\right) = 1$$

MC in discrete time

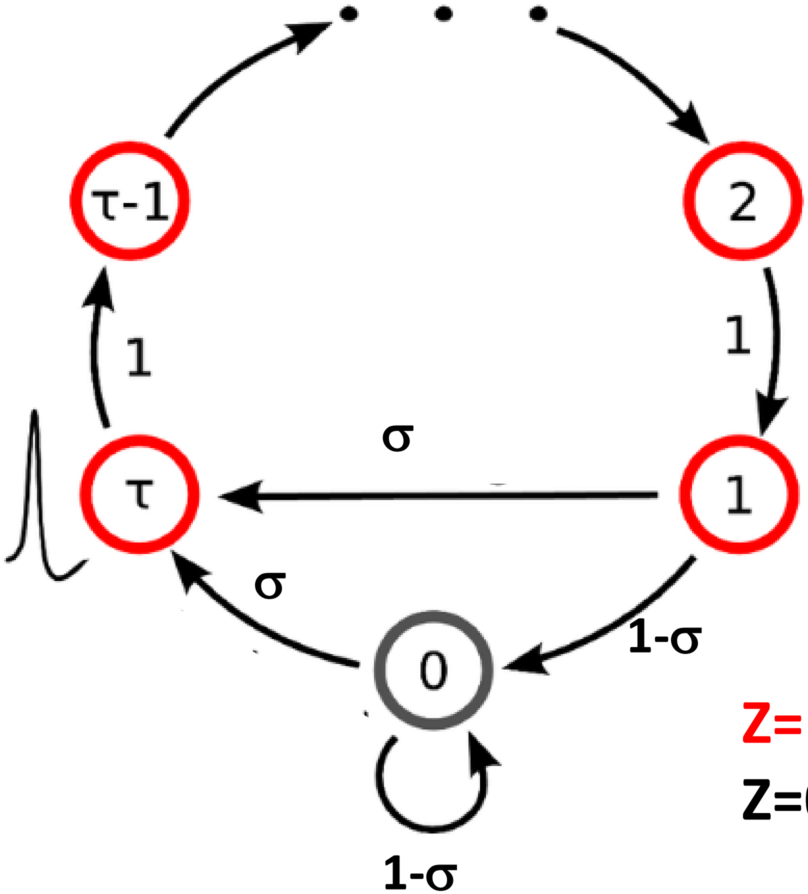
$$p\left(z_j = 1 \mid \zeta_j > 1\right) = 1$$

$$p\left(z_j = 0 \mid \zeta_j = 0\right) = 1$$

$$p\left(z_j \mid \zeta_j\right) = 0$$ In all other cases

MC in discrete time

the deterministic and random transition of ζ, z



$\sigma = \sigma(u_k - \log \tau)$

Z=1 in all the red circled states

Z=0 for the single case left

MC in discrete time

The dynamic is defined by the map T which is simply

$$T = T^{k} \circ T^{k-1} \circ \ldots \circ T^{1}$$

that is, the neuron states updated in a fixed order and at the current step, the membrane potential is based on the state of the updated or non-updated union of the set\i furthermore the new state ζ_i *depends on the previous* $\zeta'_i{}^{prec}$ *.*

MC in discrete time

The variable z_i *is determined by* ζ'_i *deterministically*

$$z_i = 1 \quad if \quad \zeta_i \geq 1$$

and

$$z_i = 0 \quad if \quad \zeta_i = 0$$

Result

1. T^i *leaves invariant* $p(\zeta',z)$,
2. *Any combination, and T as well does.*
3. *T defines an aperiodic, irreducible MC has unique invariant distribution*
4. $p(\zeta',z)$ *the unique invariant distribution.*
5. $p(z)$ *can be sampled via* $p(\zeta',z)$ *using the MC.*

Extensions

1. relative refractory state

2. continuous time

3. other distributions

Extensions

1. relative refractory state

   the readiness function *g* is introduced, the sigmoid should be replaced by a proper function *f* satisfying:

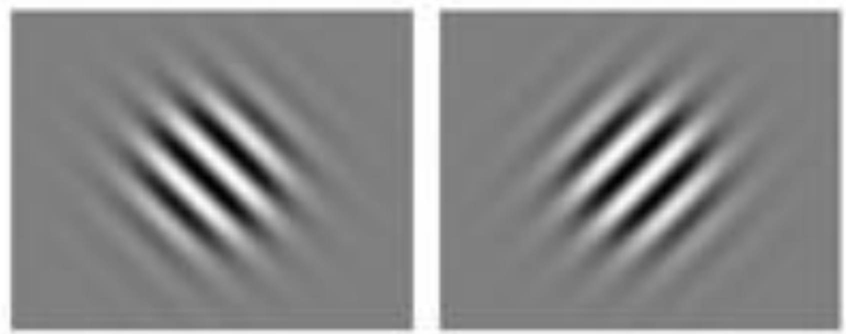
$$\exp(u) = f(u)\frac{\sum_{t=1}^{\tau}\prod_{s=t+1}^{\tau}\left(1-g(s)f(u)\right)}{\prod_{s=1}^{\tau}\left(1-g(s)f(u)\right)}$$

Remark. The pd is approximated only.

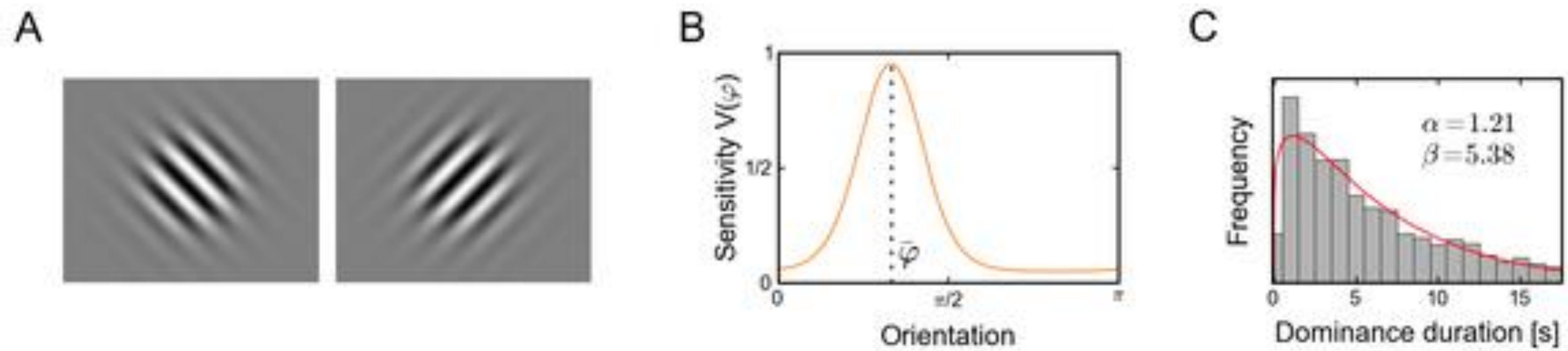
Simulation

population activity research unit

Demonstration of probabilistic inference, perceptual multistability, binocular rivalry



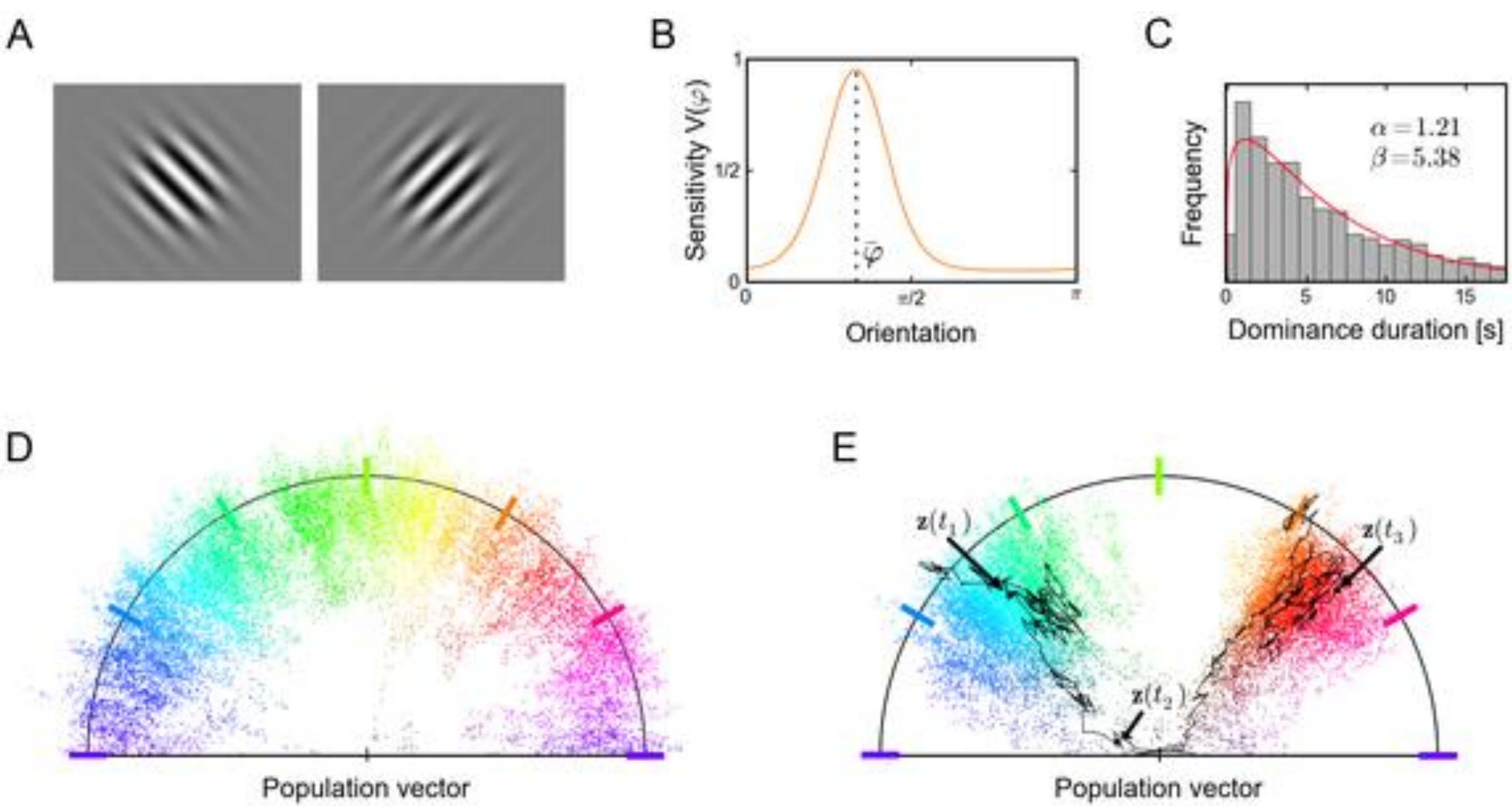
Simulation

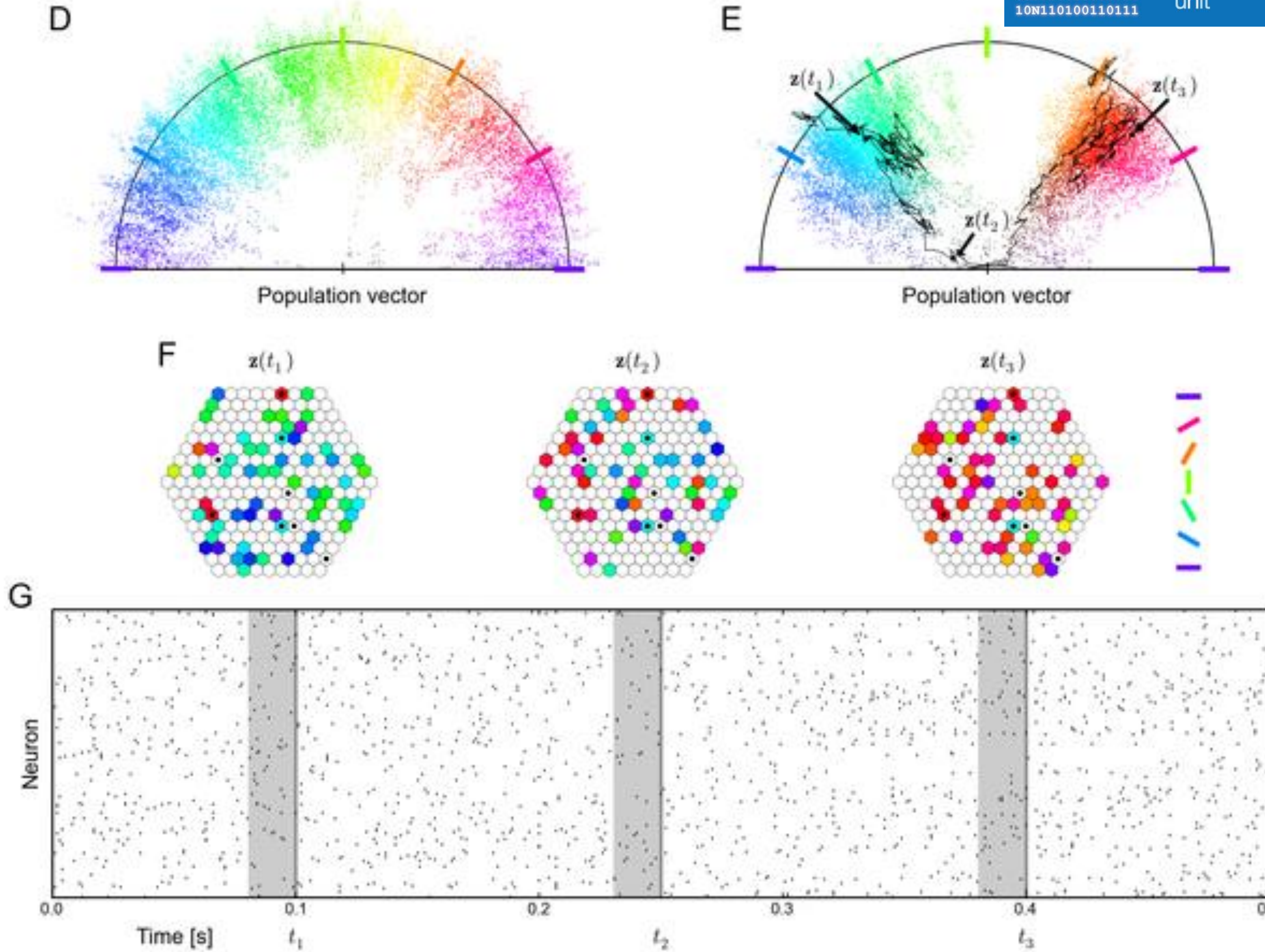

A
B
C
Sensitivity V(φ)
1
1/2
0
0
π/2
π
φ̄
Orientation
Frequency
α = 1.21
β = 5.38
0
5
10
15
Dominance duration [s]

Simulation

10P110100110111
00A011000001011
10T001010111001
01T100101110100
00E111001001101
00R010110000010
10N110100110111

population
activity
research
unit




D
Population vector
E
$\mathbf{z}(t_1)$
$\mathbf{z}(t_2)$
$\mathbf{z}(t_3)$
Population vector
F
$\mathbf{z}(t_1)$
$\mathbf{z}(t_2)$
$\mathbf{z}(t_3)$
G
Neuron
0.0
0.1
0.2
0.3
0.4
0.5
Time [s]
t_1
t_2
t_3

population activity research unit

Thanks for the attention!

Q?