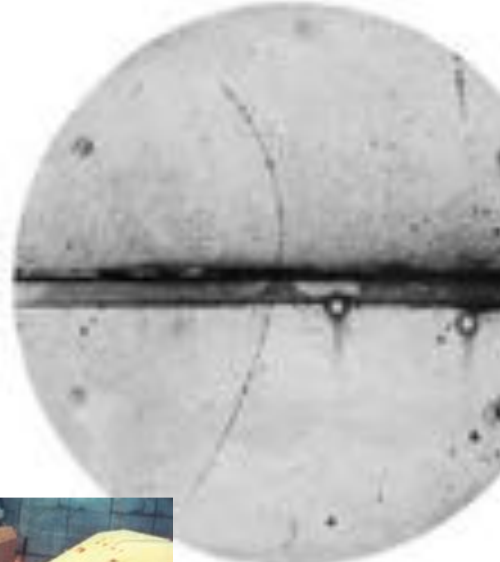
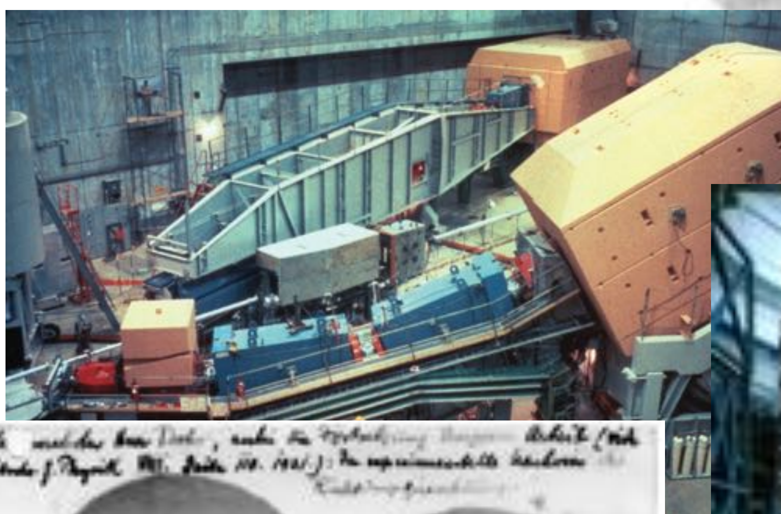
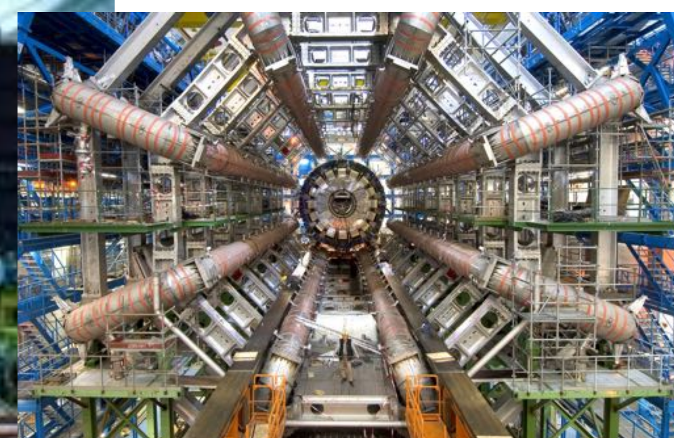
Will there ever be a Standard Model of the Brain ?

Karlheinz Meier, Ruprecht-Karls-Universität Heidelberg
Colloquium DESY & Zeuthen May 2017

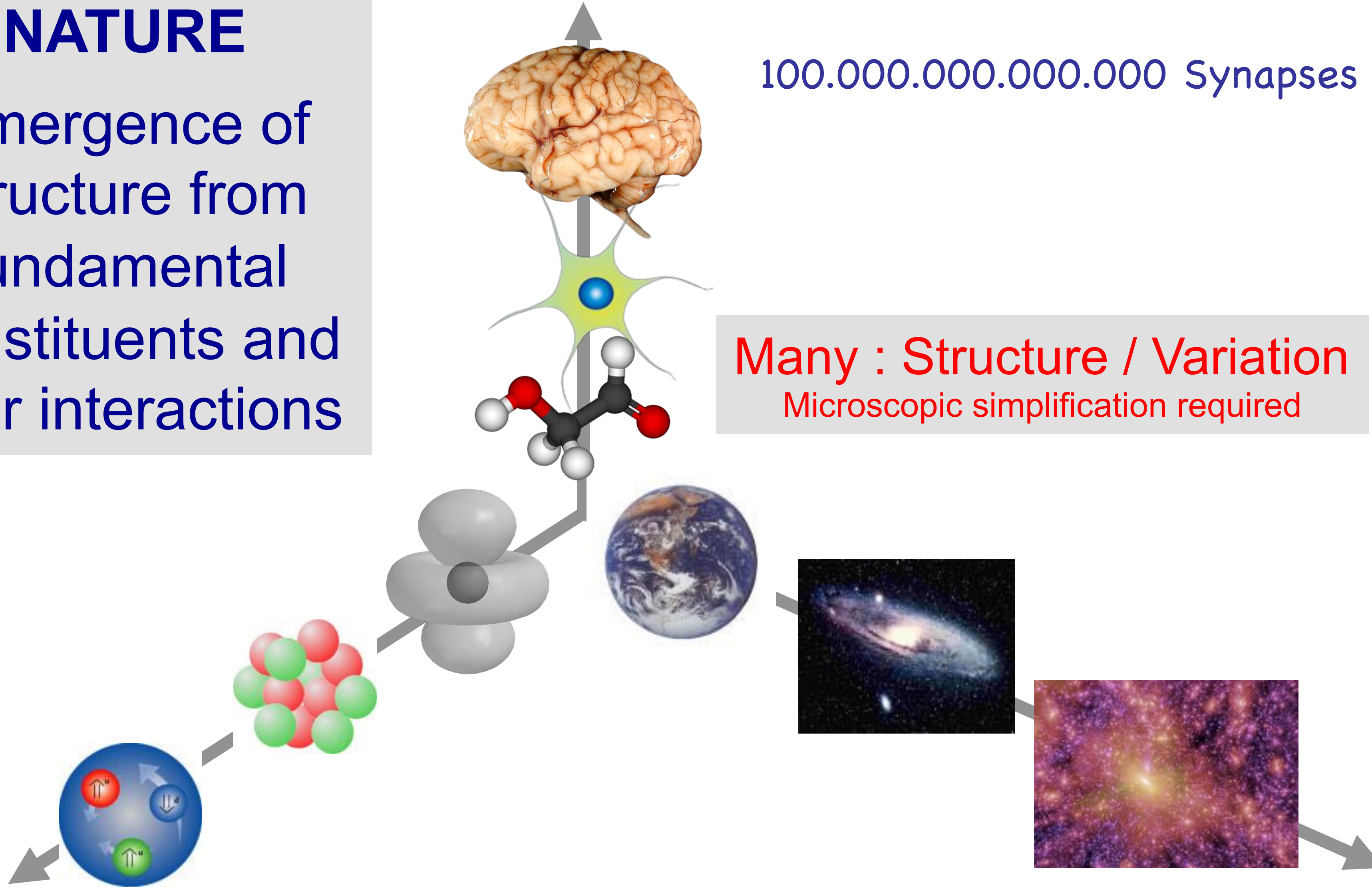








NATURE

Emergence of structure from fundamental constituents and their interactions



100.000.000.000.000 Synapses

Many : Structure / Variation

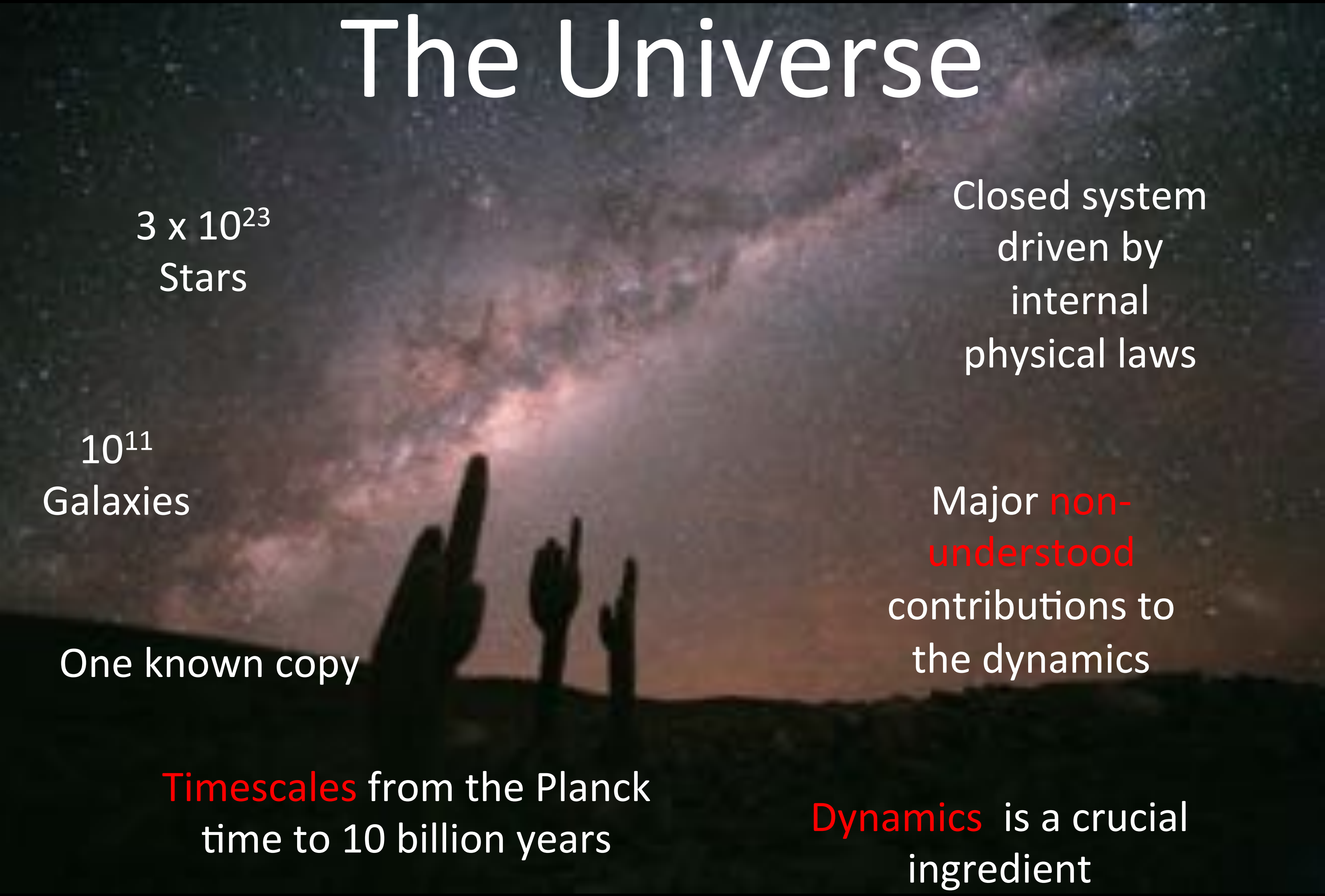
Microscopic simplification required

0.000000000000000001 Meters

100.000.000.000.000.000.000.000 Stars

Few :Precision / Uniqueness

The Universe



3×10^{23}
Stars

10^{11}
Galaxies

One known copy

Closed system driven by internal physical laws

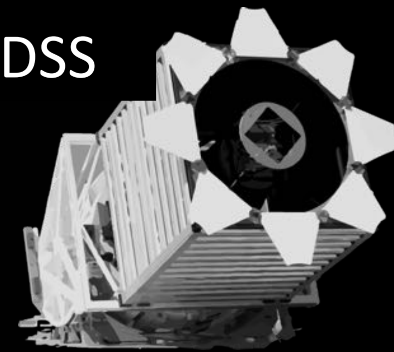
Major non-understood contributions to the dynamics

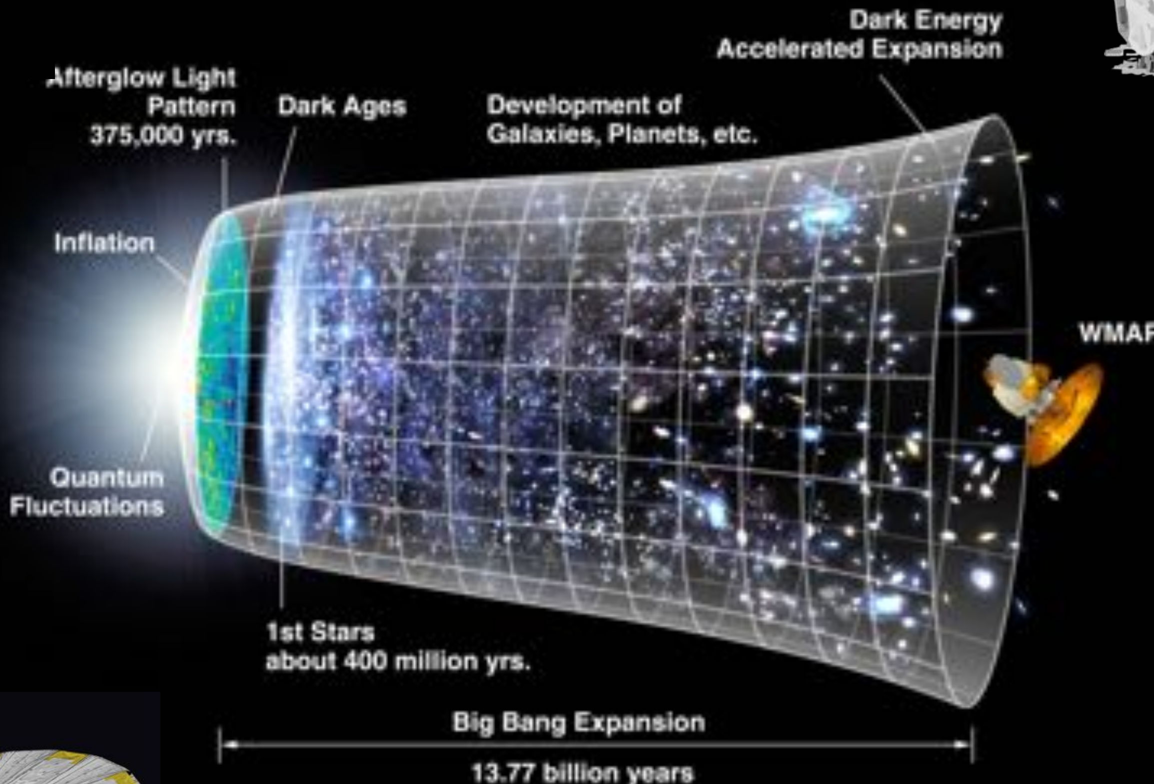
Timescales from the Planck time to 10 billion years

Dynamics is a crucial ingredient

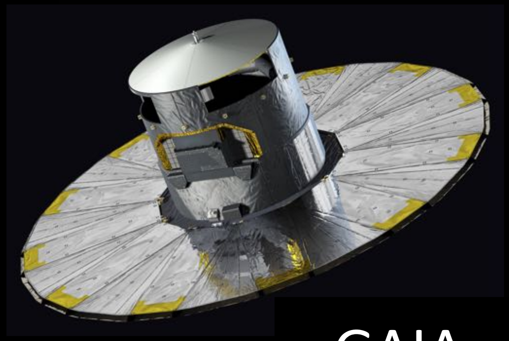
HUBBLE

SDSS





WMAP,
PLANCK



GAIA

Modern Astrophysics
Access to multiple scales
in Space and Time


z = 48.4
T = 0.05 Gyr
500 kpc

The Brain

10^{15} connections (synapses)

10^{11} nodes (neurons)

Many billion copies worldwide

Timescales from milliseconds to years

Stochastic on the microscopic level

Open system driven by external I/O for Information processing

Major non-understood contributions to the dynamics

Dynamic long-range and short-range interactions

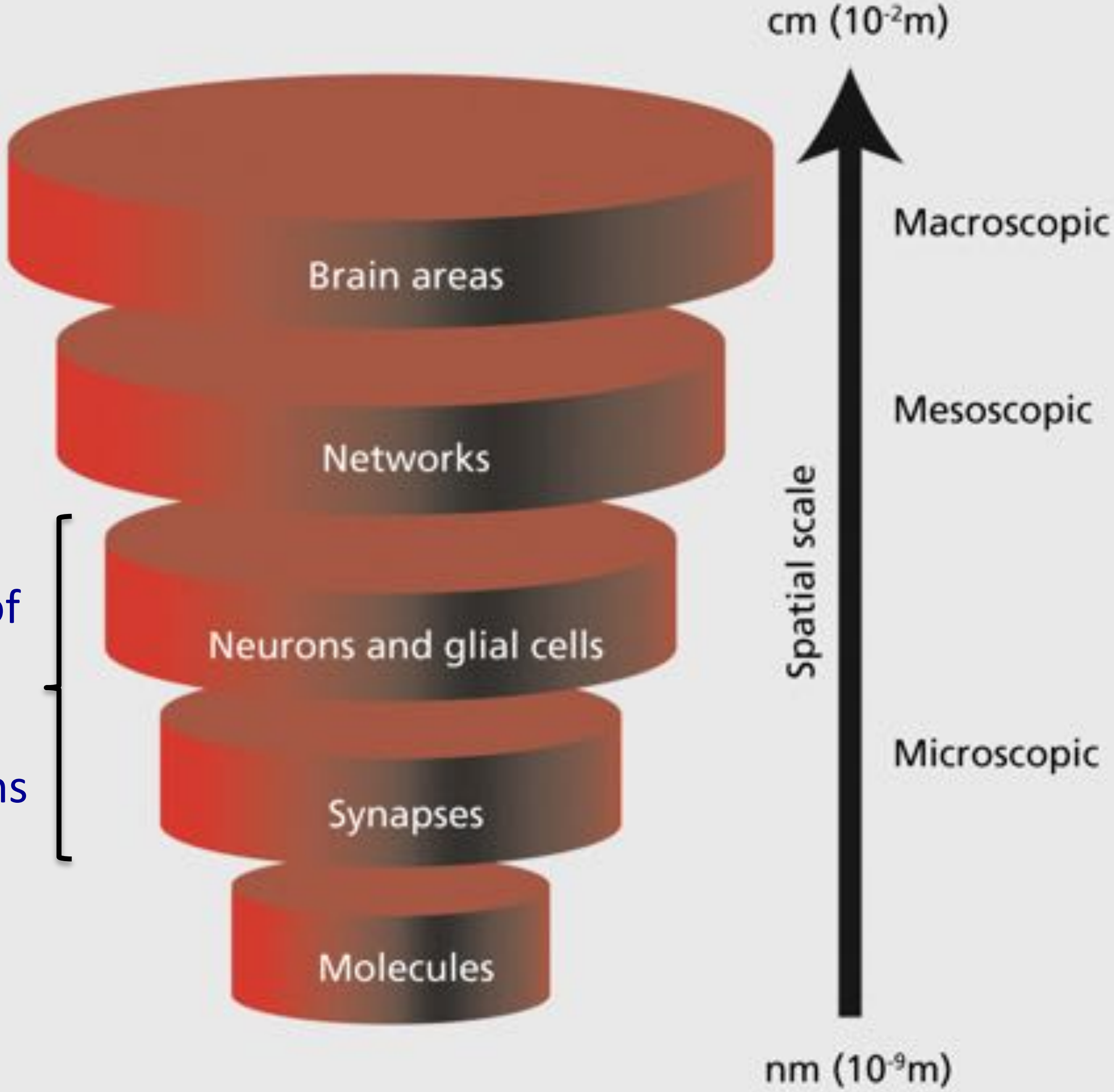
What can we hope for ?

- Identify relevant simplified constituents
- Describe structure at all relevant spatial scales
- Understand dynamics at all relevant time scales

- Understand memory
- Understand spatio-temporal pattern detection
- Understand prediction making
- Understand who you are

THE BRAIN IS AN ACTIVE INFORMATION PROCESSING SYSTEM !

„Atoms“ of the brain and their interactions ?



7 orders of magnitude in spatial scale

From constituent structure to „uniformity“

Flavio Fröhlich, Dialogues Clin Neurosci. 2014;16(1):93-102.

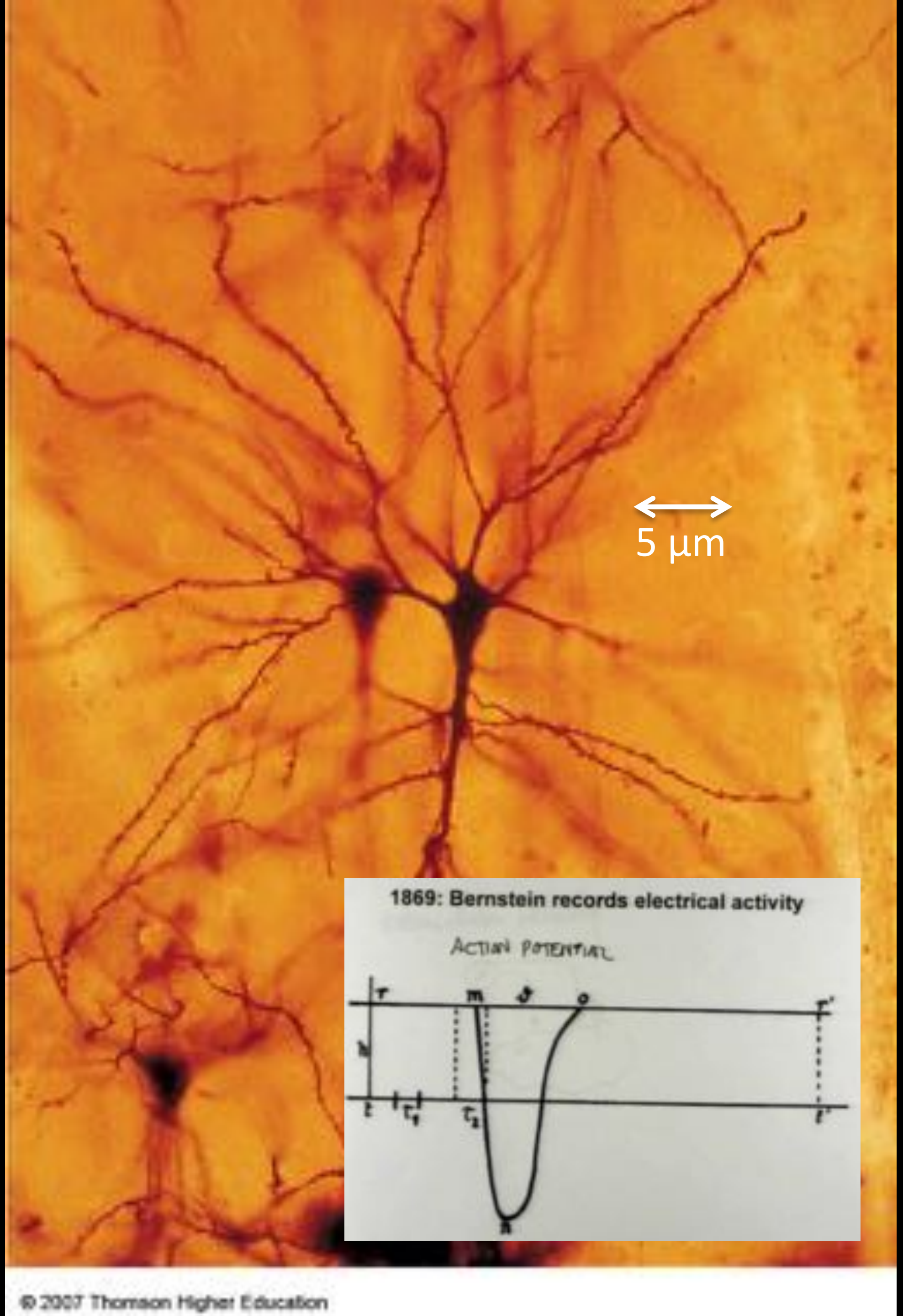
Herrmann v. Helmholtz (1821-1894)
Julius Bernstein (1839-1917)
Santiago Ramón y Cajal (1852-1934)

Individual cells in the brain are spatially separated constituents

“interaction over a distance”

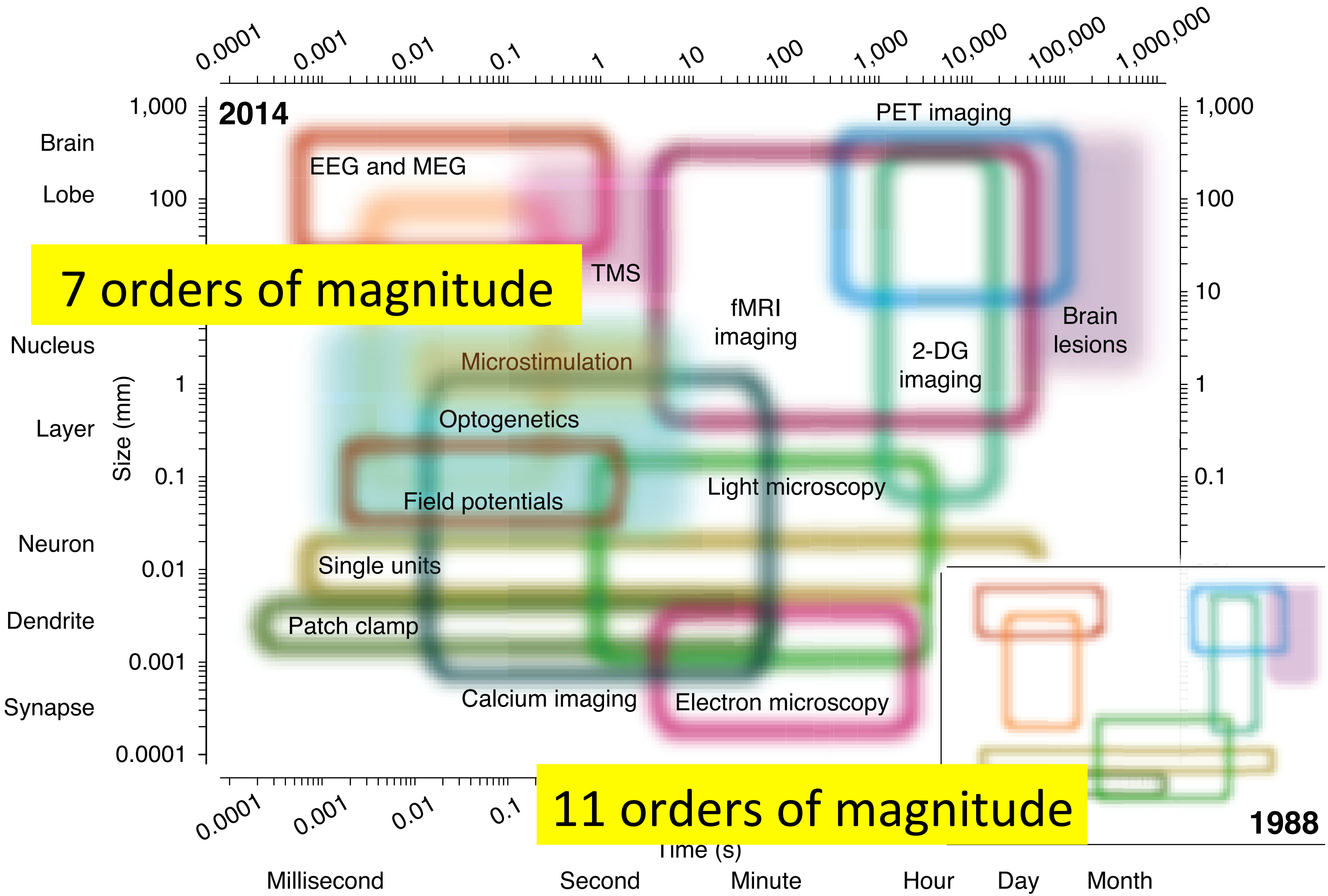
and

“spatial and temporal integration”



K. Amunts et al., Science (2013), FZ Jülich

Modern Neuroscience : Access to multiple Scales in Space and Time

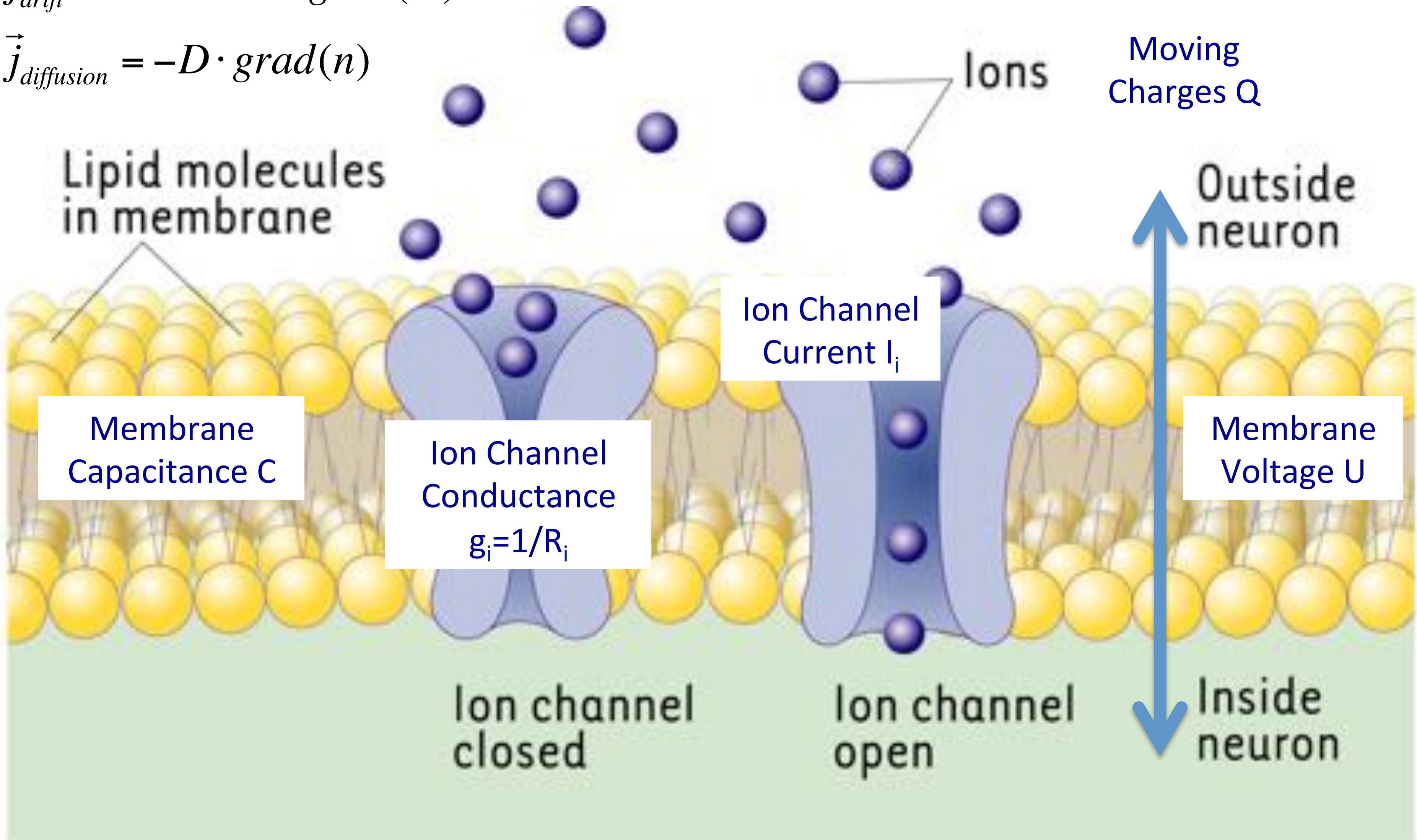


Sejnowski et al, Nature Neuroscience, 2014

Some Electrical Quantities of a real Neuron Membrane

$$\vec{j}_{drift} = \sigma \cdot \vec{E} = -\sigma \cdot grad(\Phi)$$

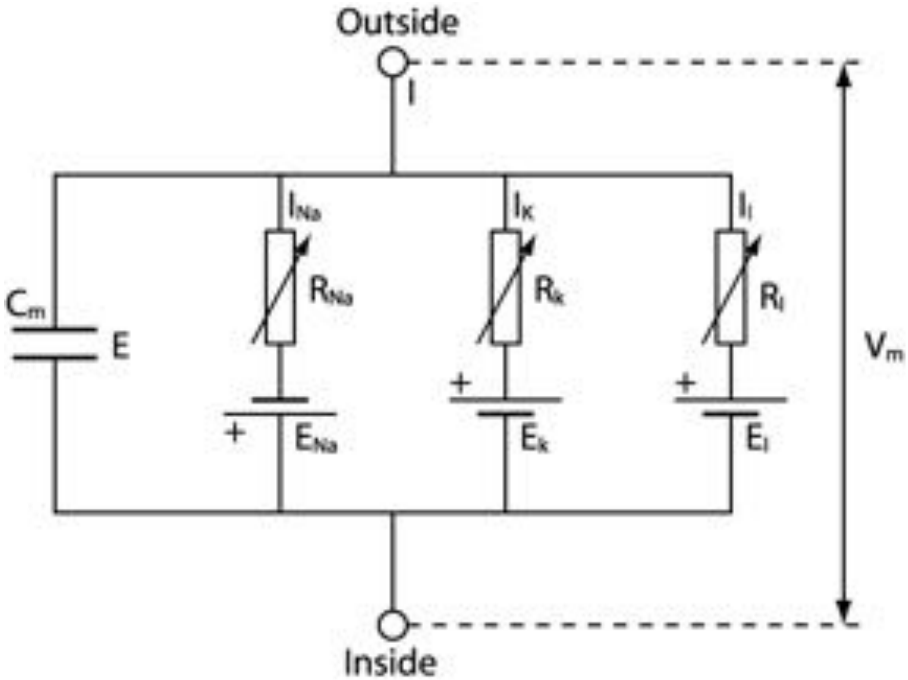
$$\vec{j}_{diffusion} = -D \cdot grad(n)$$

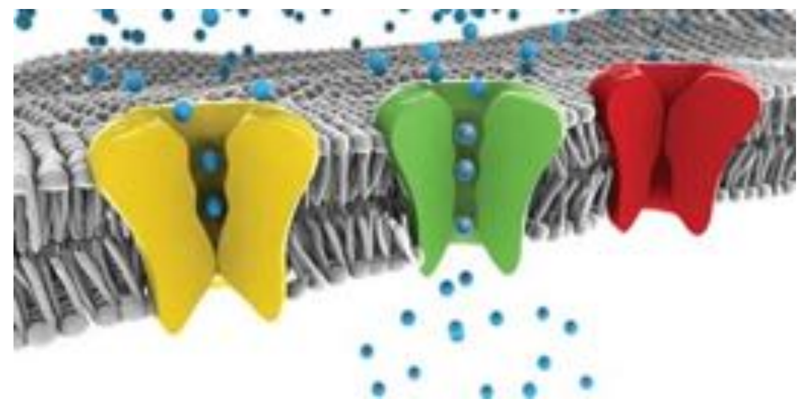


U, I and g are functions of time in an operating network !

Current theories and modelling are treating these quantities only (few exceptions)

Hodgkin-Huxley 1952
Describing the non-linearity



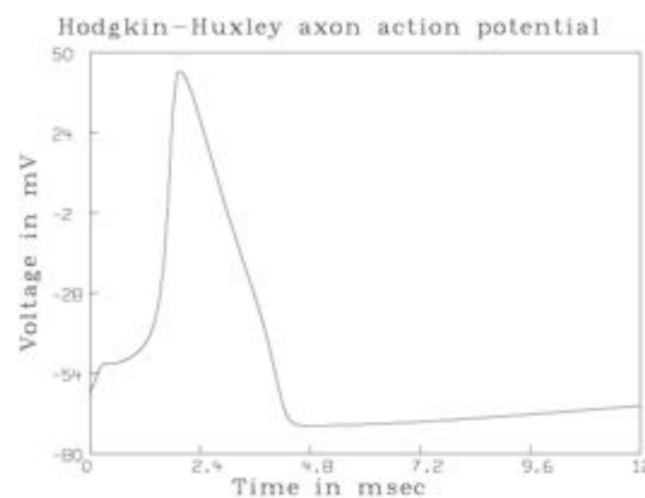


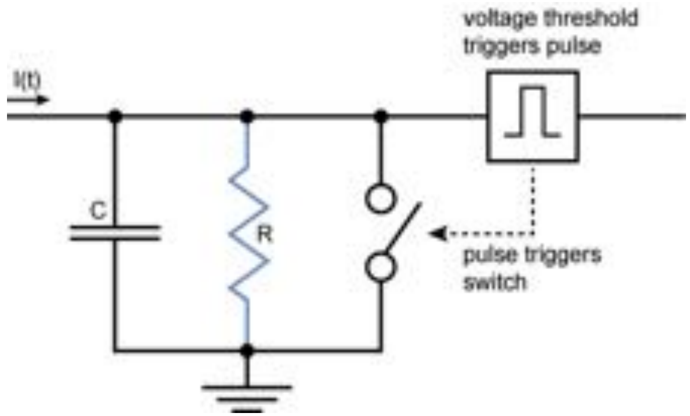
$$I = C_m \frac{\mathrm{d}V_m}{\mathrm{d}t} + \bar{g}_{\mathrm{K}} n^4 (V_m - V_K) + \bar{g}_{\mathrm{Na}} m^3 h (V_m - V_{Na}) + \bar{g}_l (V_m - V_l),$$

$$\frac{dn}{dt} = \alpha_n(V_m)(1-n) - \beta_n(V_m)n$$

$$\frac{dm}{dt} = \alpha_m(V_m)(1-m) - \beta_m(V_m)m$$

$$\frac{dh}{dt} = \alpha_h(V_m)(1-h) - \beta_h(V_m)h$$

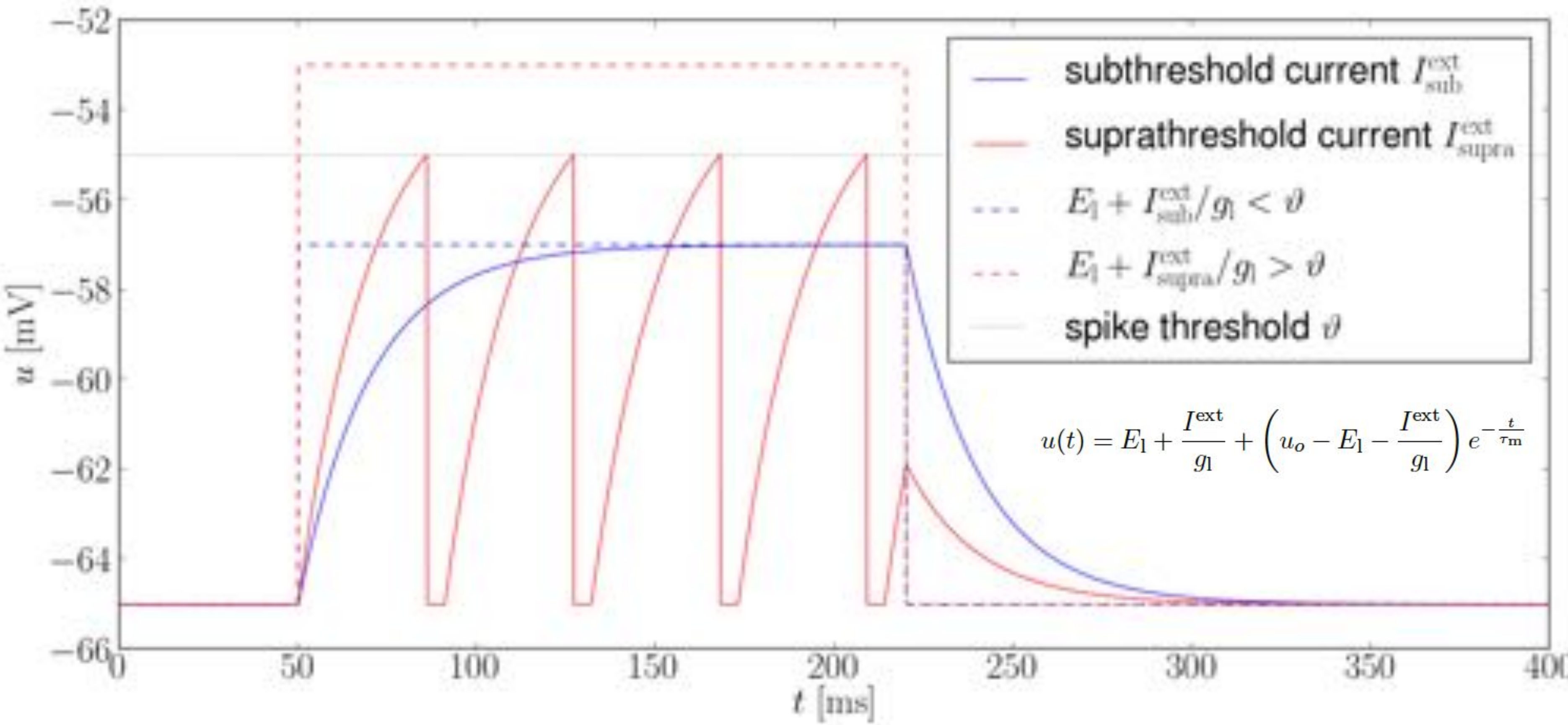




Leaky-integrate-and-fire (LIF)

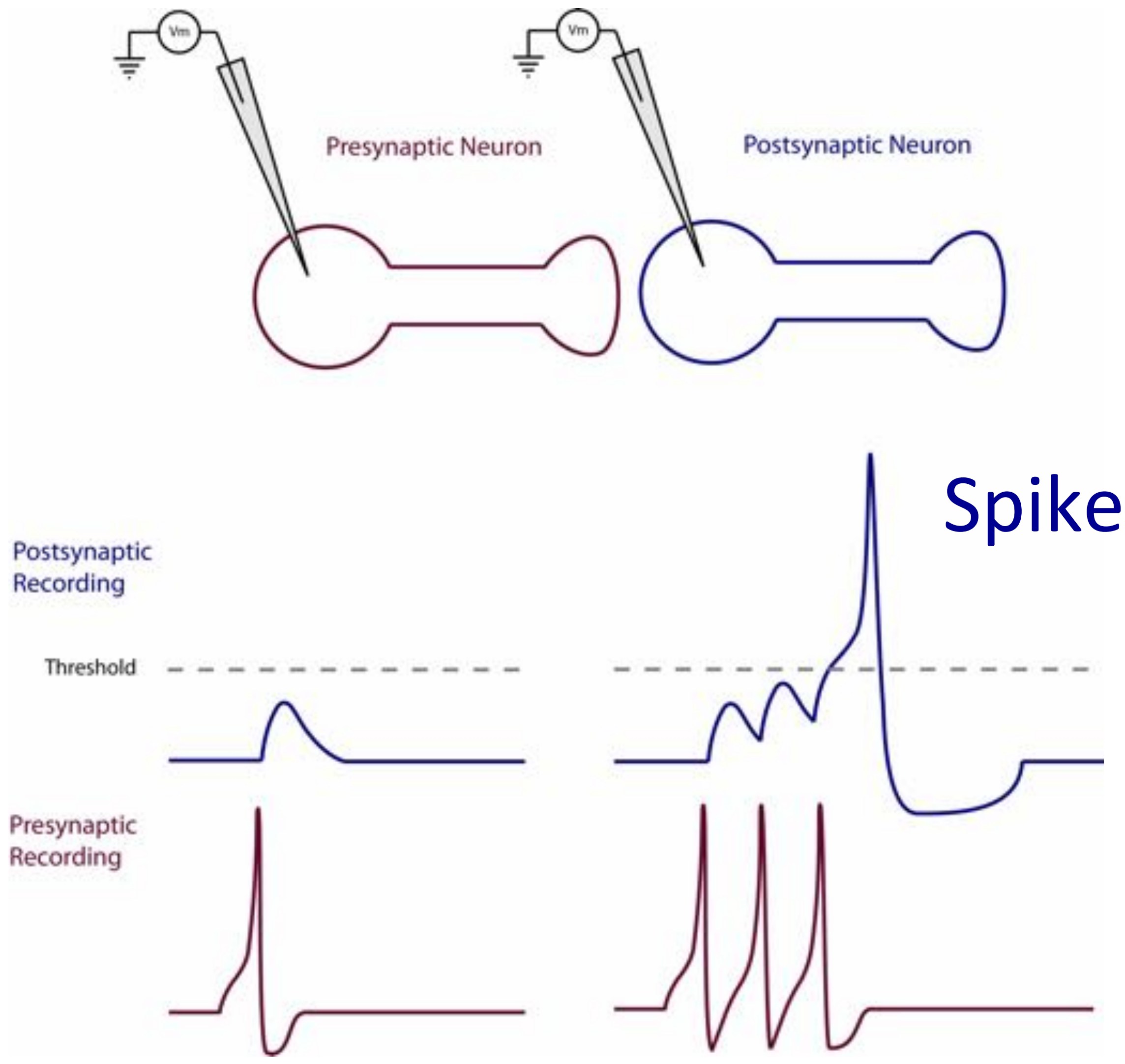
$$C_\mathrm{m}\frac{du}{dt} = g_\mathrm{l}(E_\mathrm{l} - u) + I^\mathrm{syn} + I^\mathrm{ext}$$

$$u(t_\mathrm{spike} < t \leq t_\mathrm{spike} + \tau_\mathrm{ref}) = \varrho$$

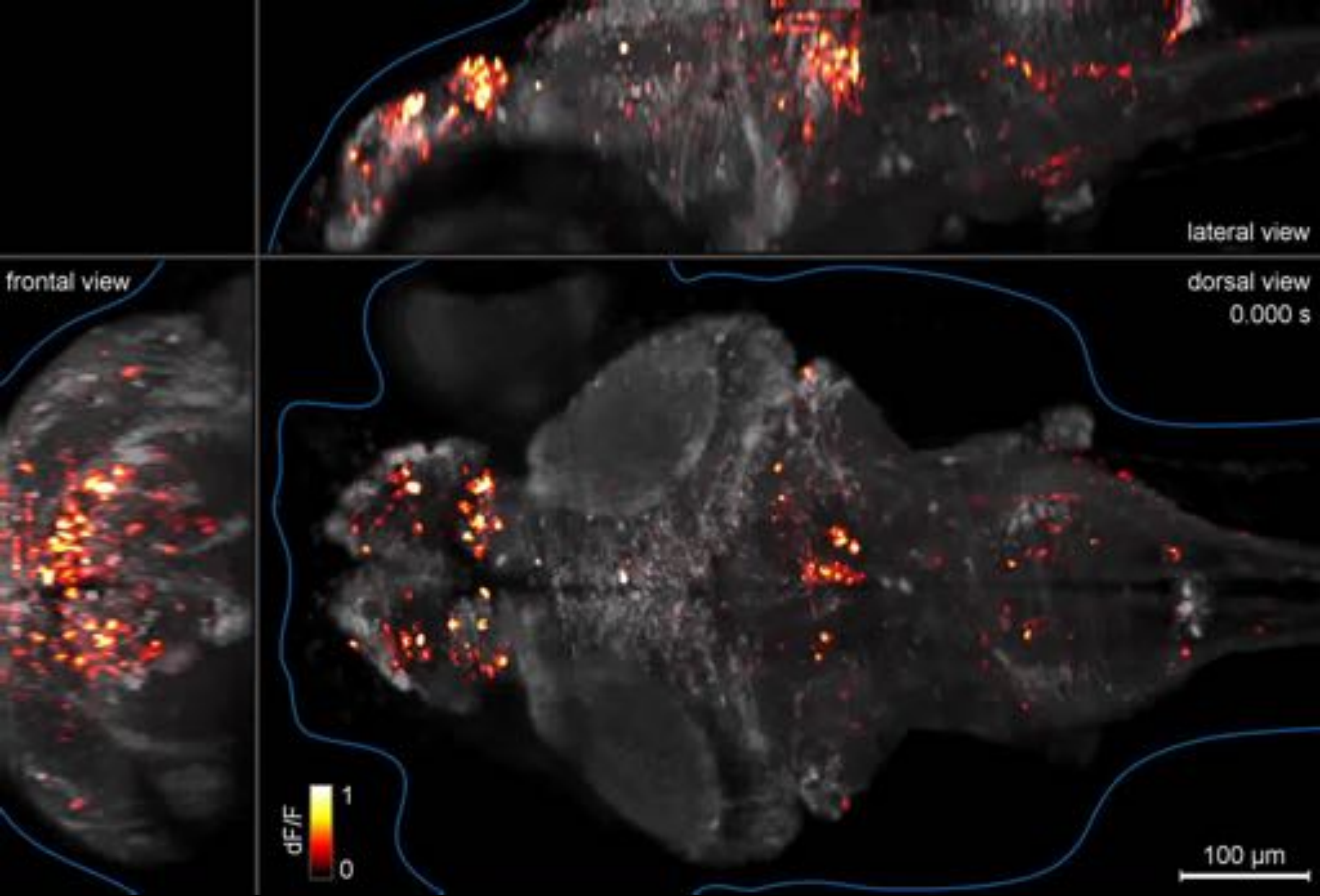


PhD, Mihai Petrovici, 2015

Time and temporal integration

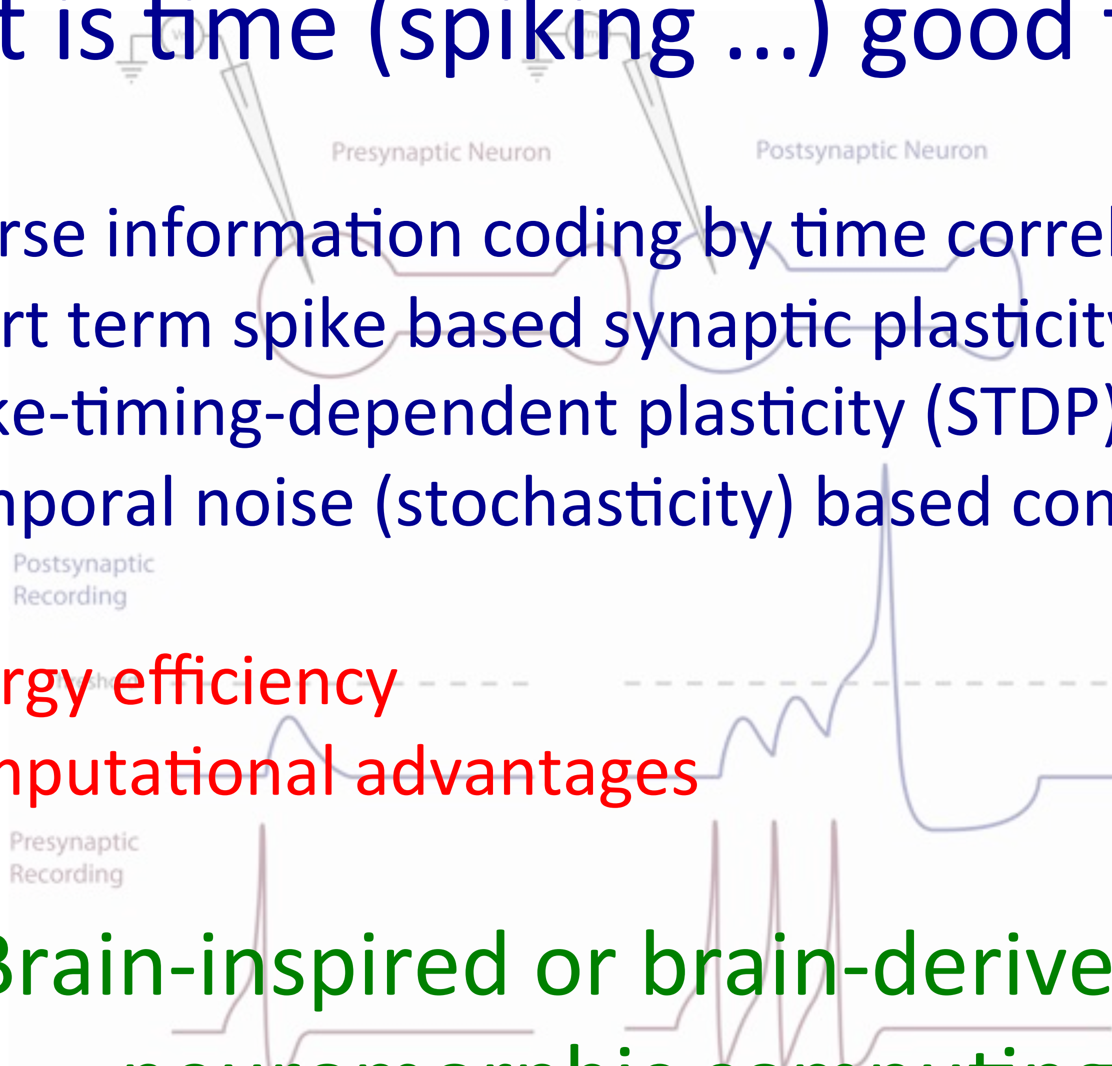


https://upload.wikimedia.org/wikipedia/commons/0/0a/Temporal_summation.JPG



Ahrens et al, *Nature Methods* **10**, 413–420 (2013)

What is time (spiking ...) good for ?



- Sparse information coding by time correlations
- Short term spike based synaptic plasticity (STP)
- Spike-timing-dependent plasticity (STDP)
- Temporal noise (stochasticity) based computing

- Energy efficiency
- Computational advantages

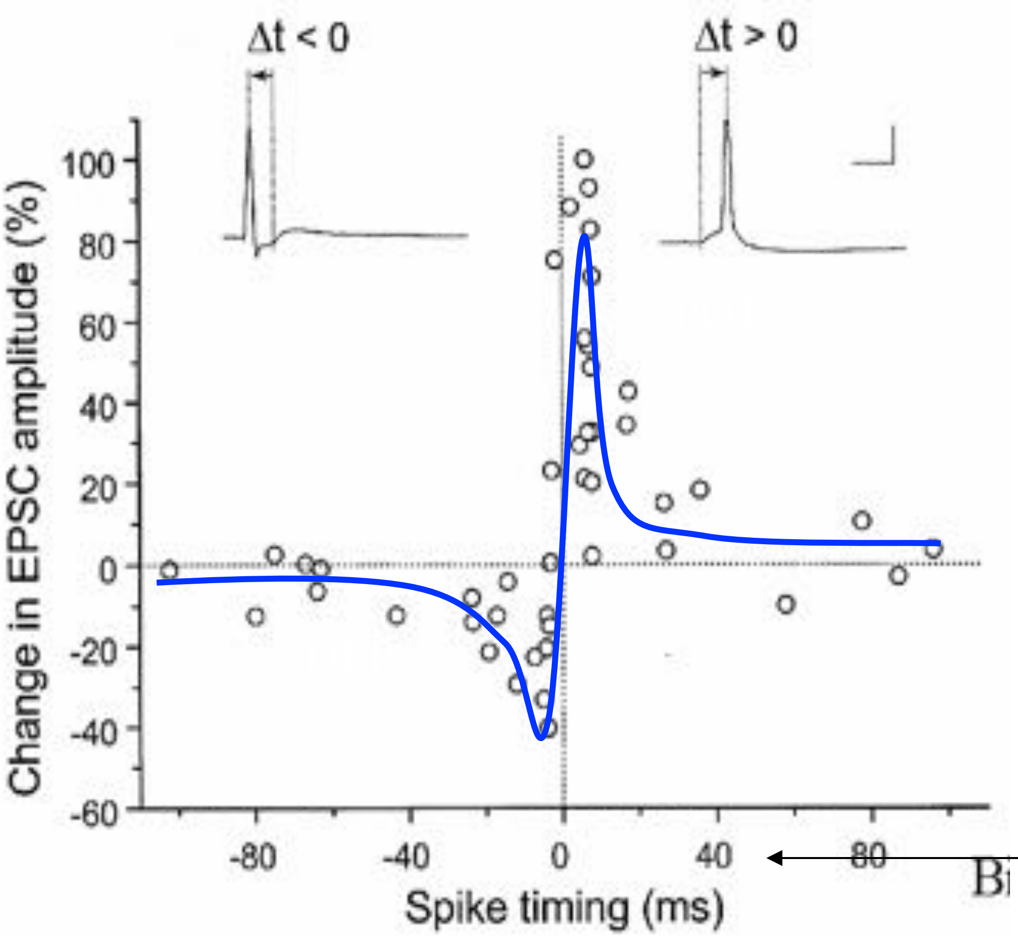
Brain-inspired or brain-derived or neuromorphic computing

https://upload.wikimedia.org/wikipedia/commons/0/0a/Temporal_summation.JPG

FAST dynamics : „Spike-Time-Dependent-Plasticity (STDP)“

In vivo intracellular recording (Adult Visual Cortex)

(Bi and Poo, Ann. Rev. Neurosci., 2001)



STDP as a Causality Detector

Extremly strong time dependence of facilitation or depression of synaptic strength

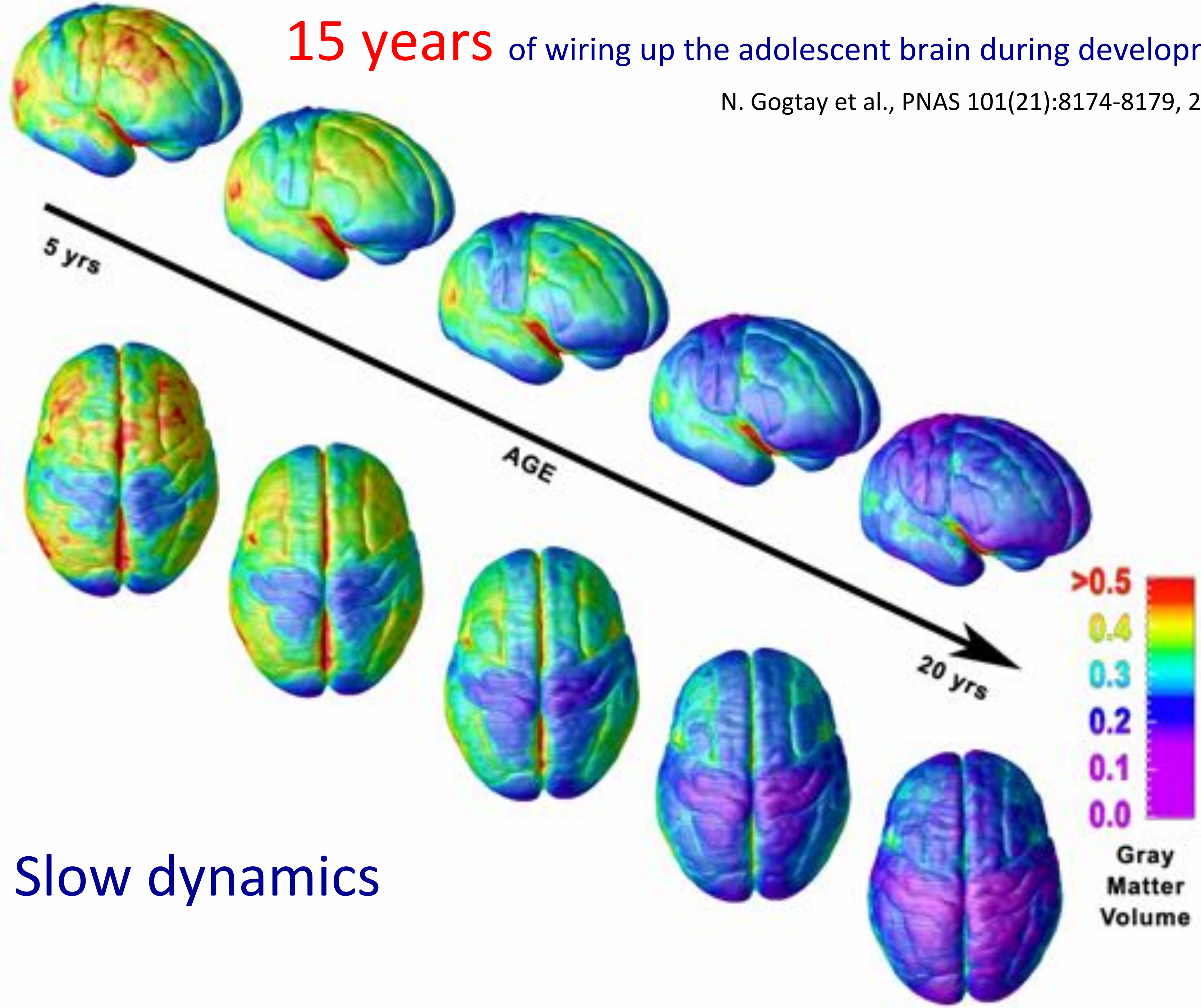
Neural circuits require asynchronous MILLISECOND timing for long term learning !

AFTER - BEFORE

„synaptic spike“

15 years of wiring up the adolescent brain during development

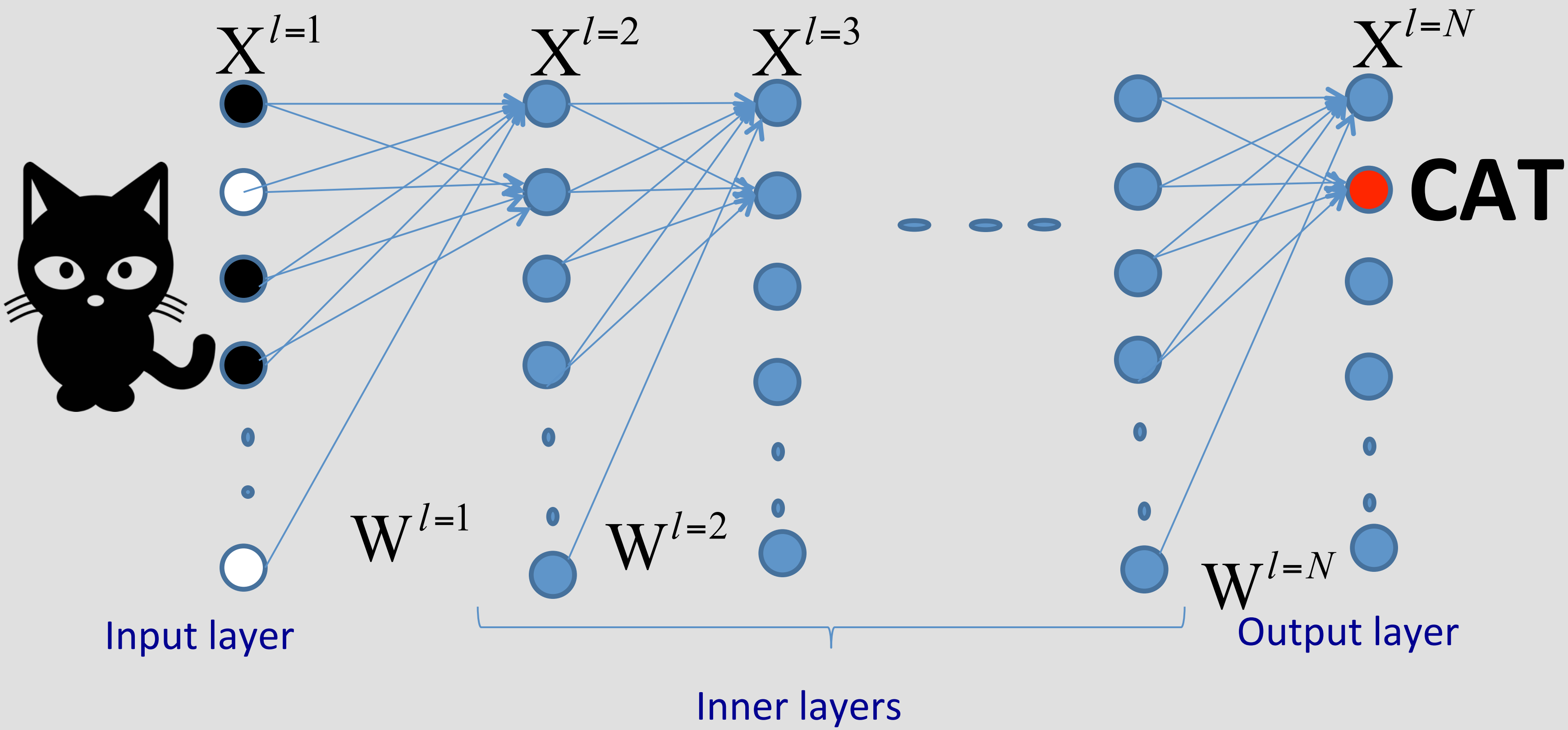
N. Gogtay et al., PNAS 101(21):8174-8179, 2004



Slow dynamics

Artificial Neuronal Networks

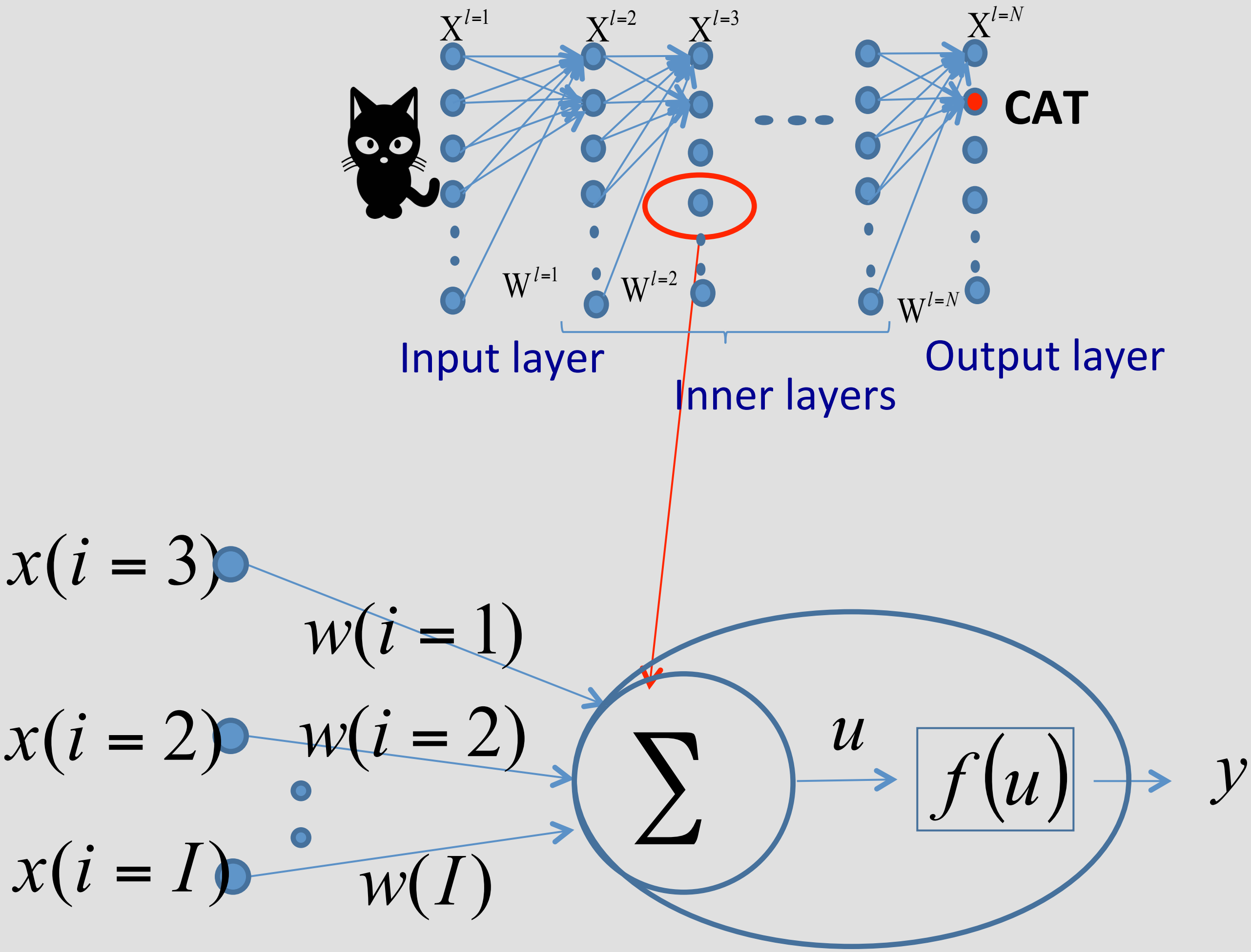
ignore time evolution ....



Here : local, no recurrency *feed-forward*

Pairs of neurons connected by weights

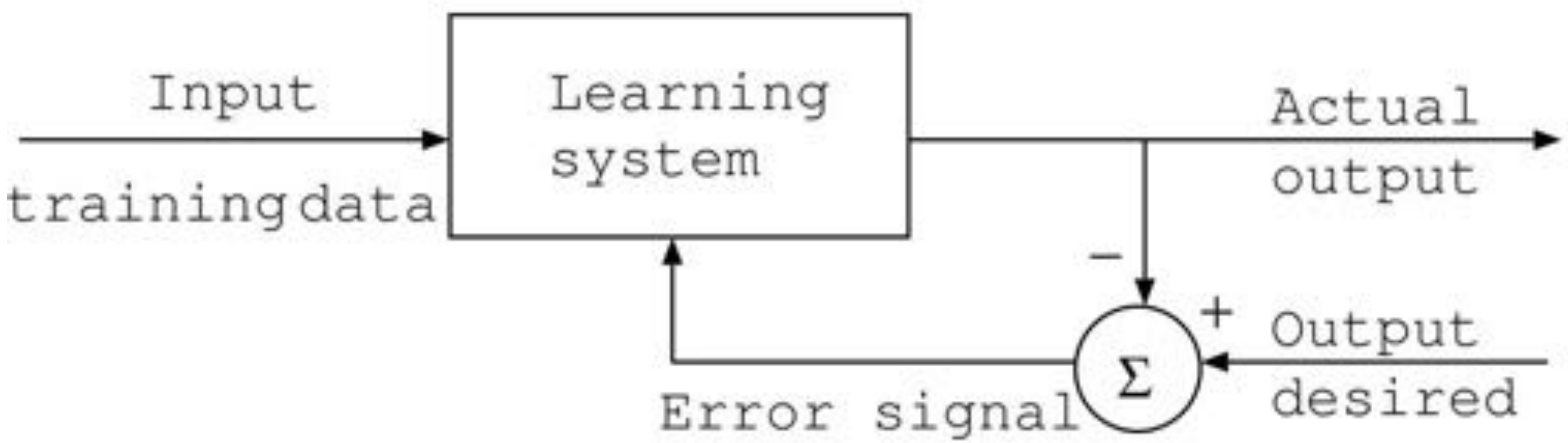
Neuron performs integration (summing)



Learning Example : Supervised

Labelled input data

Actual output



Deviation

Desired output

Jumpstart

Strategic Network

Supervised Learning

Predict human moves

database of existing matches

160.000 matches, 30 Million positions

Policy Network

Reinforcement Learning

Network self-matches

128.000.000 matches

Value Network

Combination of first 2 steps

30 Million self-matches

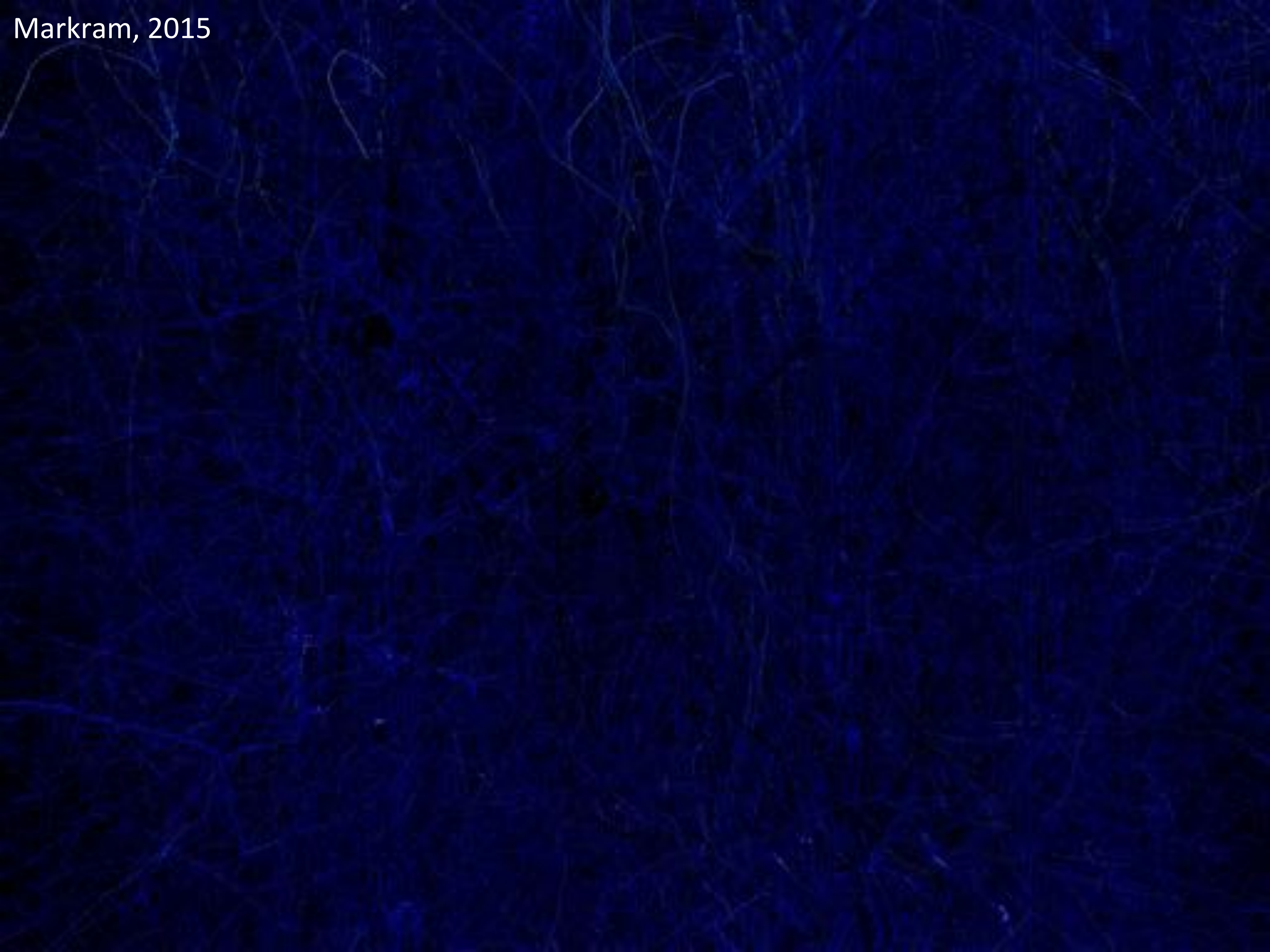
One year learning time, 0.5 MW

Energy : 183 MWh

Excessive training samples

Learning is slow and expensive

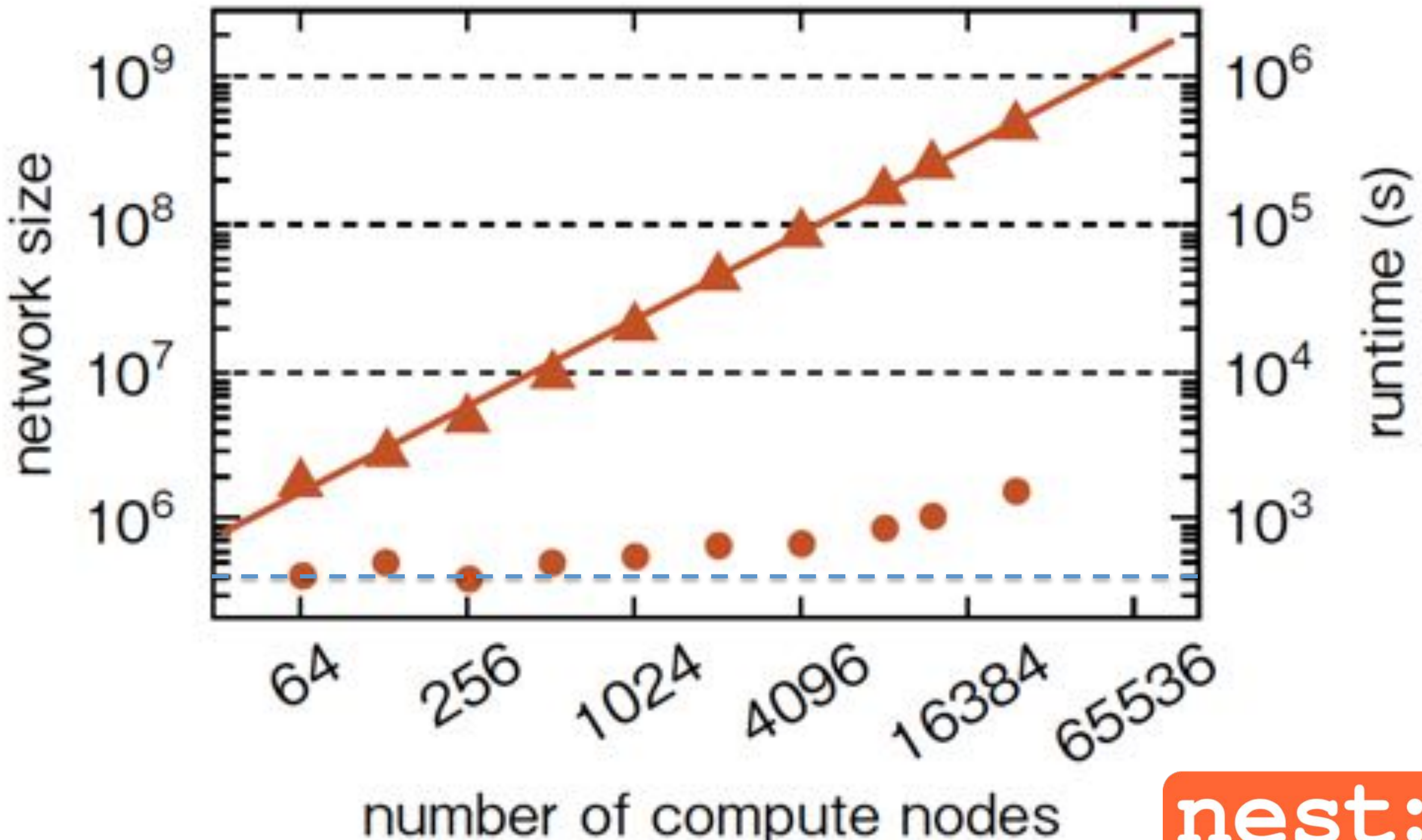
Application is fast


Markram, 2015

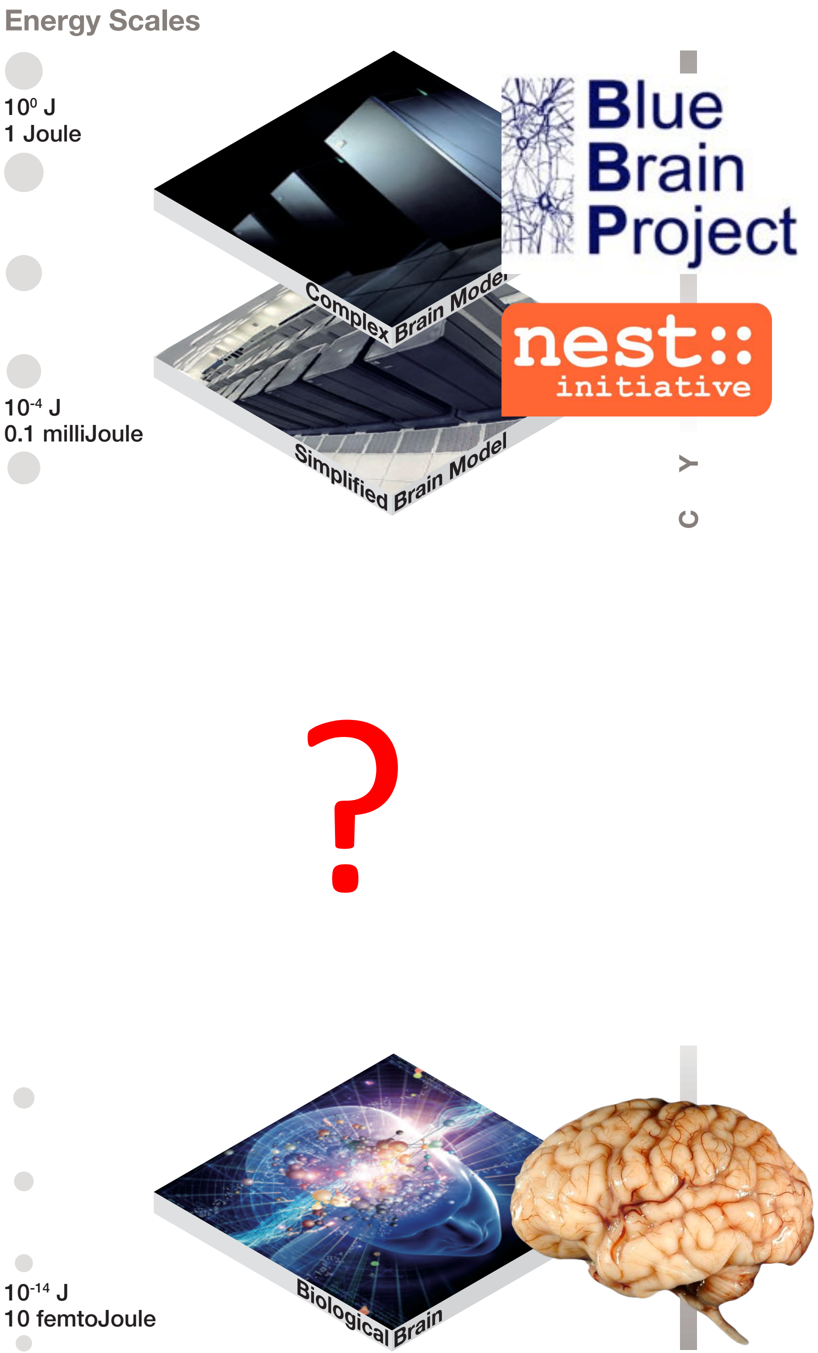
K-Computer, RIKEN Lab, 12.6 MW

Processor-to-Neural Cell Ratio 1 : 20.000

Simulation speed 1.520 : 1 compared to biological real-time



Diesmann, Proceedings of the 4th Biosupercomputing Symposium, Tokyo, 2012



Energy*Scales*

Computational Primitive :

Energy used for a synaptic transmission

10 - 14 orders of magnitude

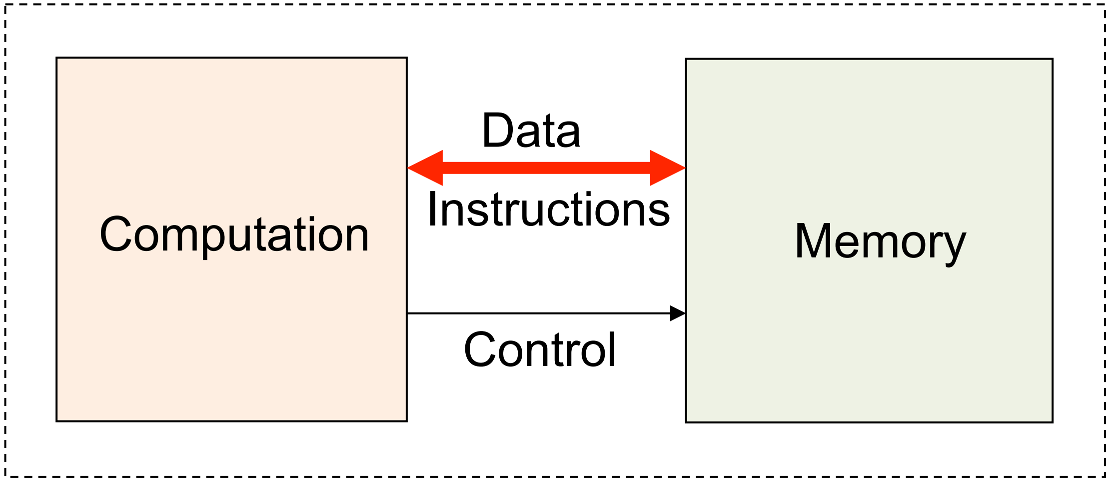
difference for *„the same thing“*

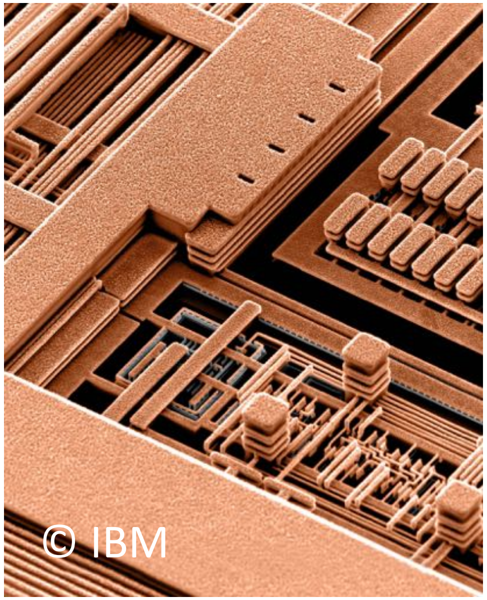
From : HBP project report

TimeScales	Nature + Real-time	Simulation
Causality Detection	10^{-4} s	0.1 s
Synaptic Plasticity	1 s	1000 s
Learning	Day	1000 Days
Development	Year	1000 Years
12 Orders of Magnitude		
Evolution	> Millenia	> 1000 Millenia
> 15 Orders of Magnitude		

von Neumann Architecture

- Data and instructions stored in memory
- Content of memory addressable by location
- Instructions executed sequentially unless order is explicitly modified
- Memory and Computation physically separated



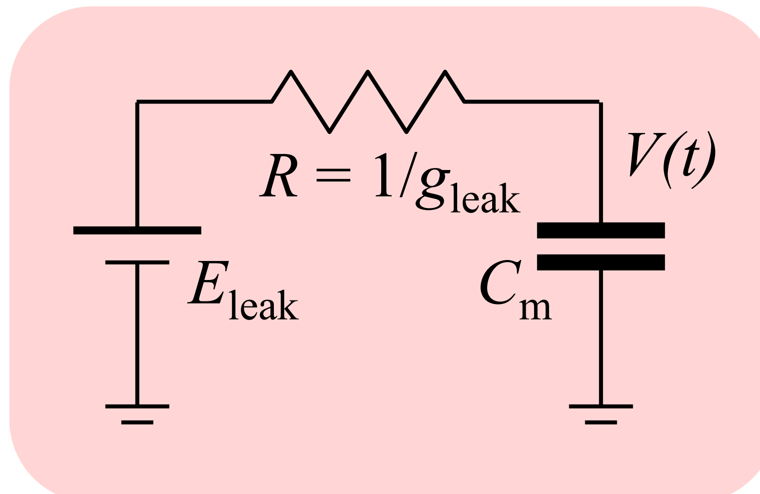


Physical Model System

Continuous Time Integrating Neural Cell Membrane

(+ non-linearity)

$$C_m \frac{dV}{dt} = -g_{leak}\left(V - E_{leak}\right)$$



	g_{leak} [S]	C_m [F]
Biology(*)	10^{-8}	10^{-10}
VLSI	10^{-6}	10^{-13}

(*) Brette/Gerstner, J. Neurophysiology, 2005

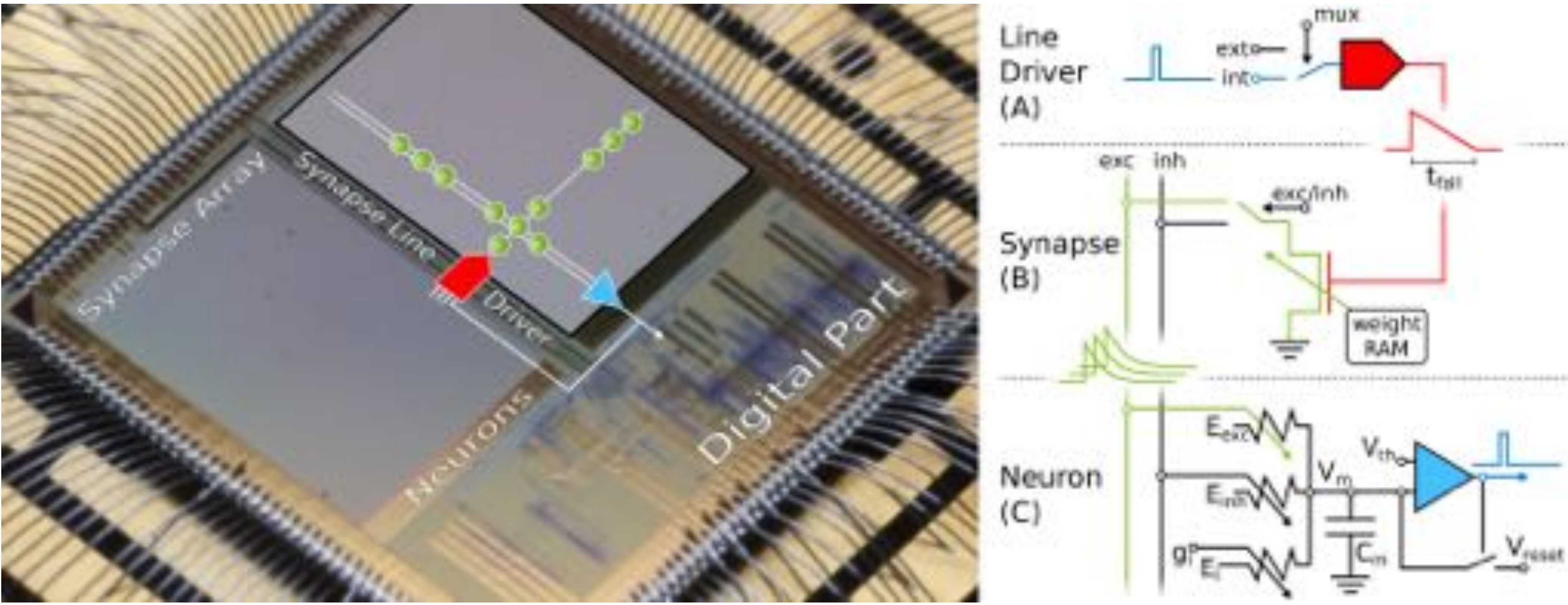
$$c_m \frac{dV}{dt} = -g_{leak}\left(V - E_l\right) + \sum_k p_k g_k \left(V - E_x\right) + \sum_l p_l g_l \left(V - E_i\right)$$

$p_{k,l}(t)$ exponential onset and decay (post-synaptic potential shape)

$g_{k,l}$ 0 to g_{max} ("weights")

effective membrane time-constant c_m / g_{total} is time-dependent

„Time" is imposed by internal physics, not by external control



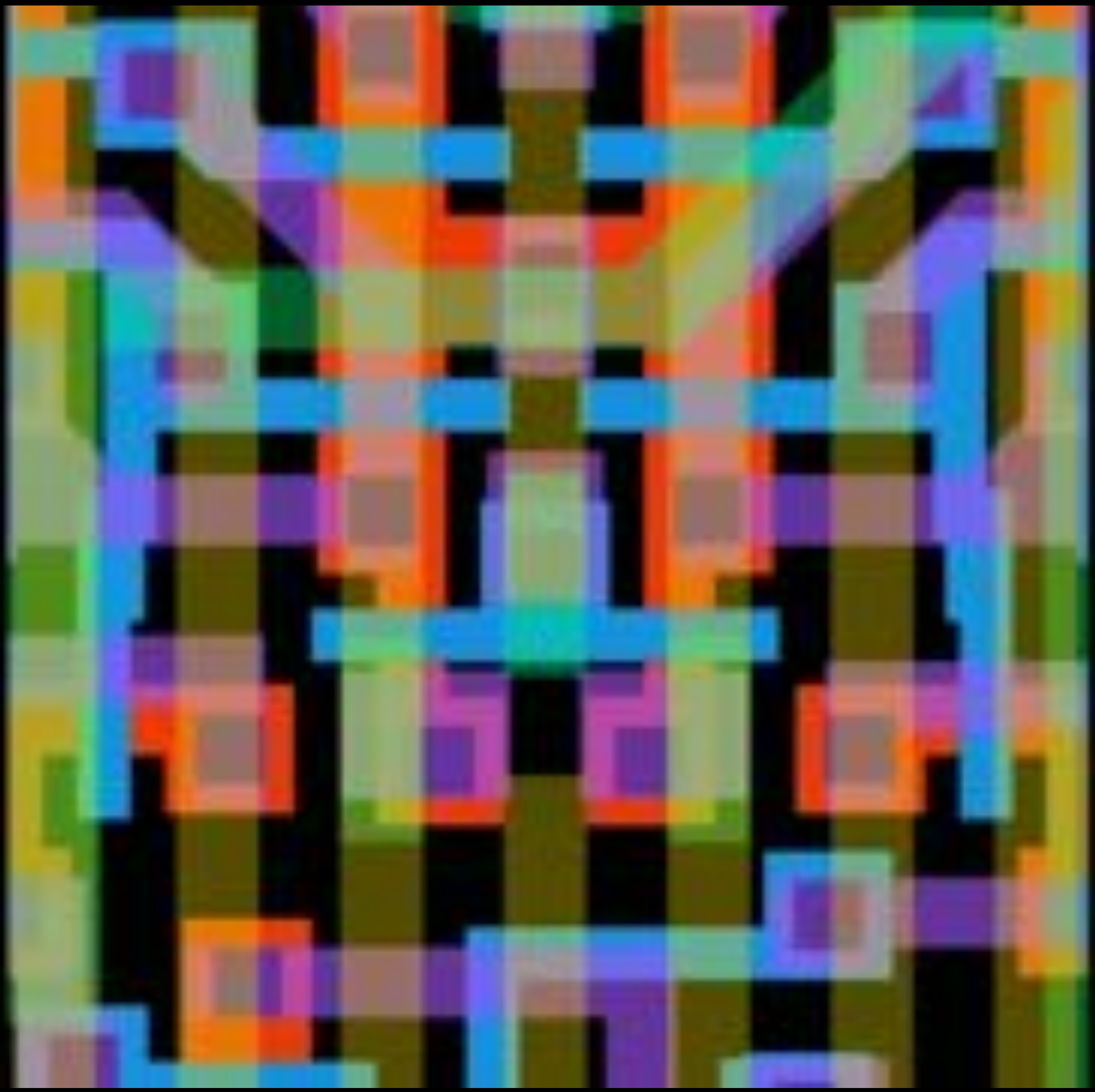
5 mm x 5 mm = 0.25 cm^2
100.000 dynamic synapses
10^6 s/cm^2 on synaptic field
4×10^5 s/cm^2 on chip

Neuron density is irrelevant when discussing VLSI neural systems

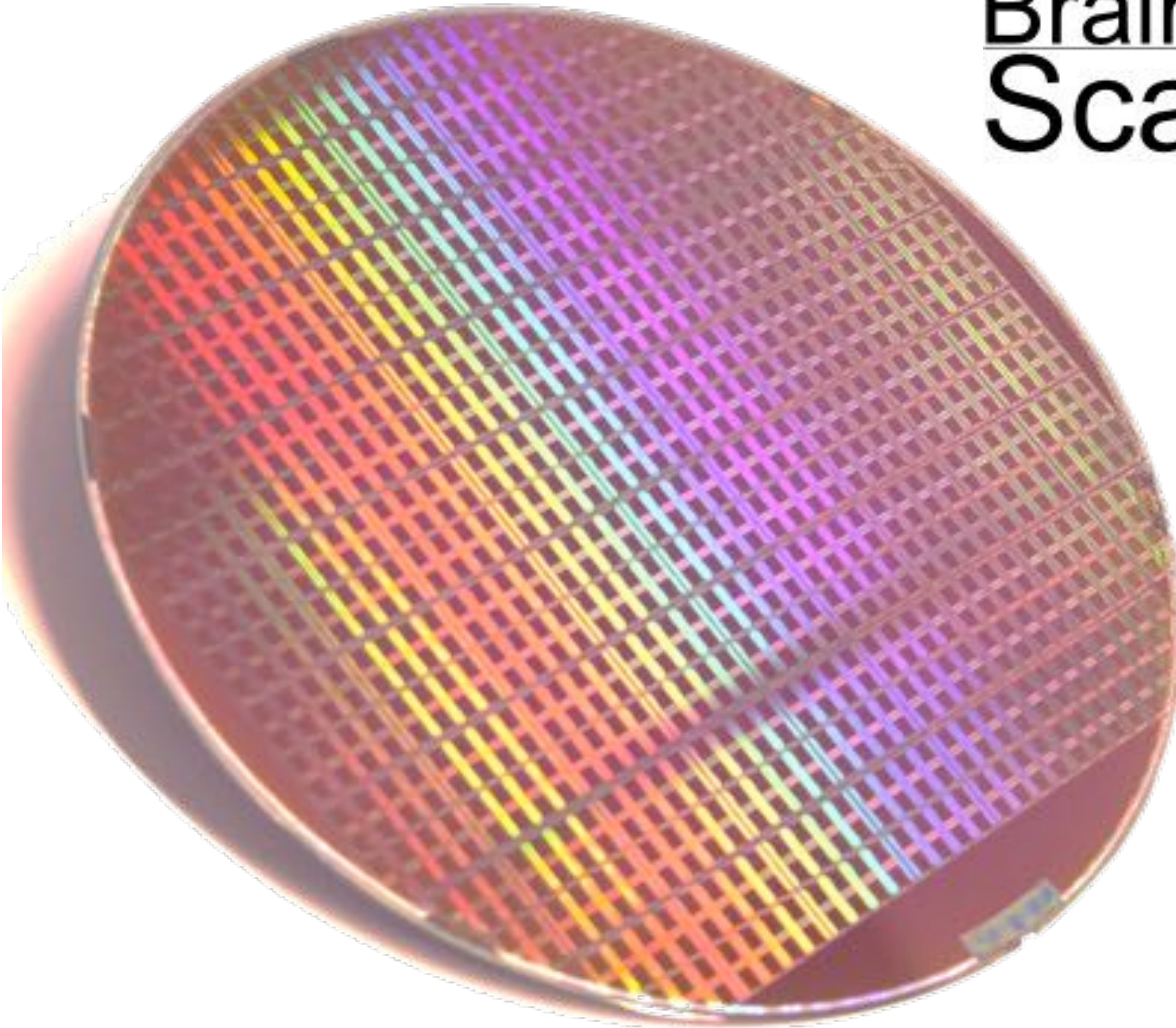
$$c_{\mathrm{m}} \frac{dV}{dt} = -g_{\mathrm{leak}}\left(V - E_{\mathrm{l}}\right) + \sum_k p_k g_k \left(V - E_{\mathrm{x}}\right) + \sum_l p_l g_l \left(V - E_{\mathrm{i}}\right)$$

Mixed-signal : local analog computation, binary, continuous time communication – „brain-like“

Thomas Pfeil, Andreas Grübl, Sebastian Jeltsch, Eric Müller, Paul Müller, Mihai Petrovici, Michael Schmuker, Daniel Brüderle, Johannes Schemmel, and Karlheinz Meier. "Six networks on a universal neuromorphic computing substrate“, Front Neurosci. 2013; 7: 11.



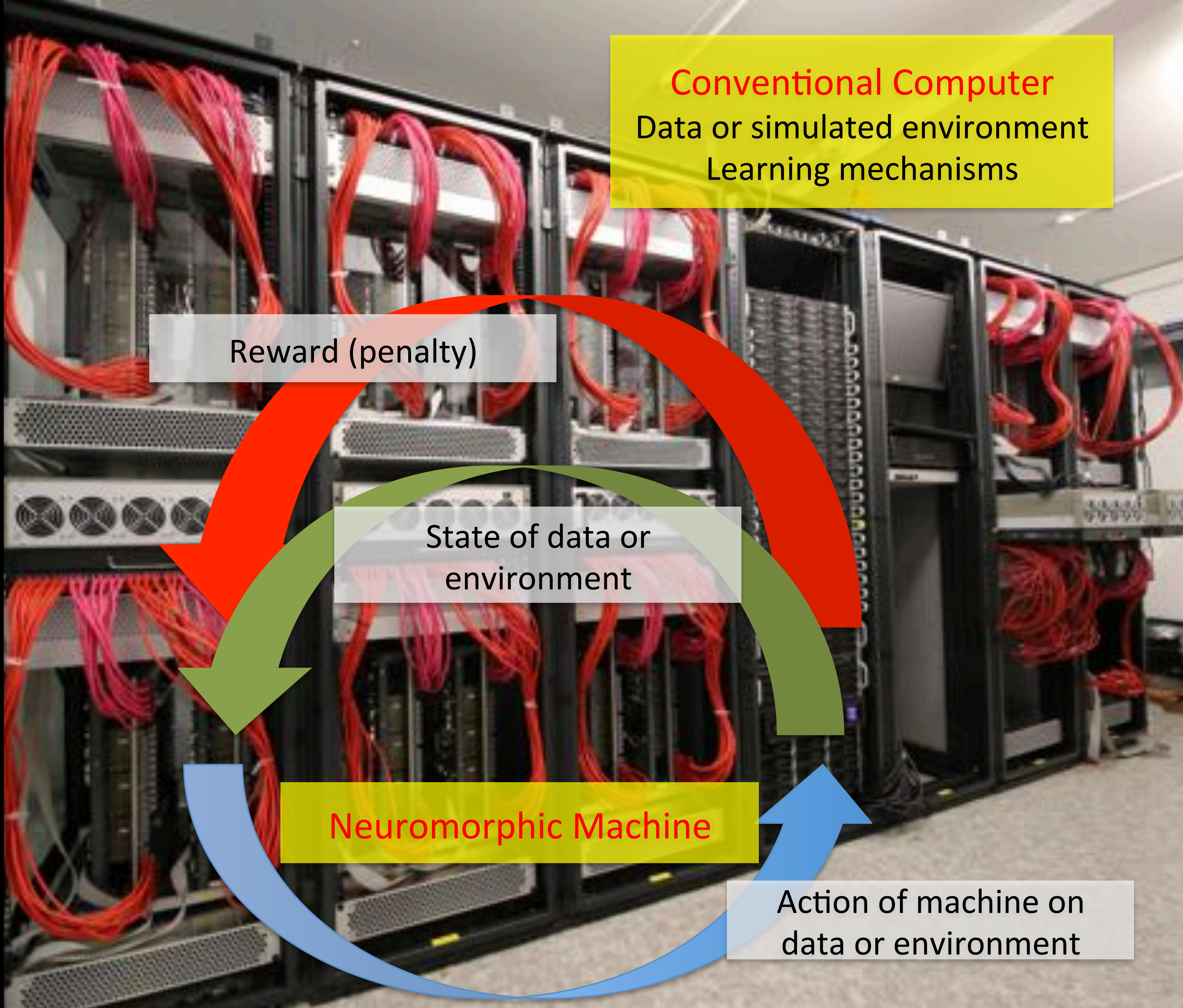
BrainScaleS



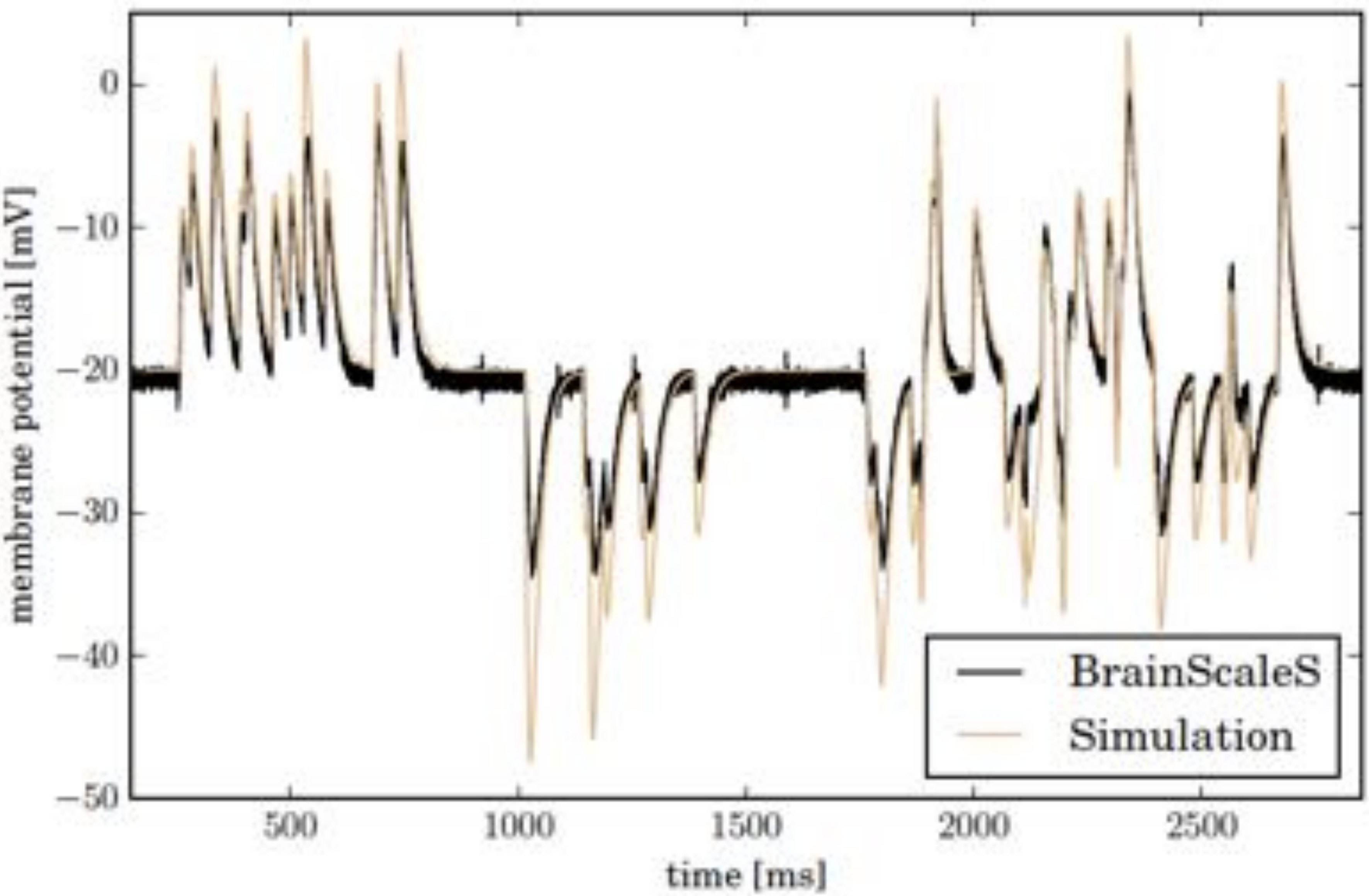
Physical Model, local analogue computing, binary continuous time communication

Wafer-Scale Integration of 200.000 neurons and 50.000.000 synapses on a single 20 cm wafer

Short term and long term plasticity, 10.000 faster than real-time

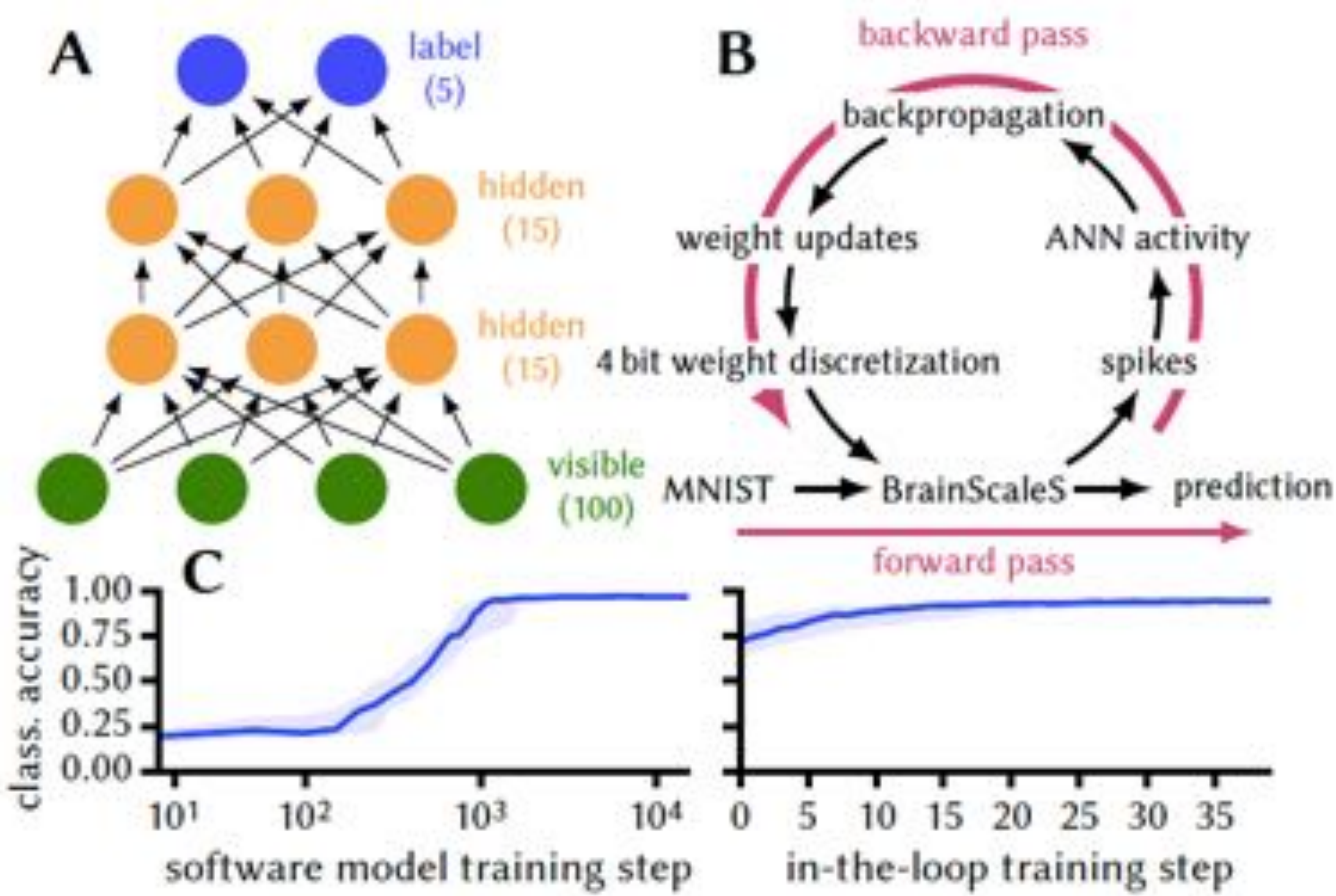

Conventional Computer
Data or simulated environment
Learning mechanisms
Reward (penalty)
State of data or environment
Neuromorphic Machine
Action of machine on data or environment

Physical model emulation



Sebastian Schmitt et al., accepted IJCNN 2017

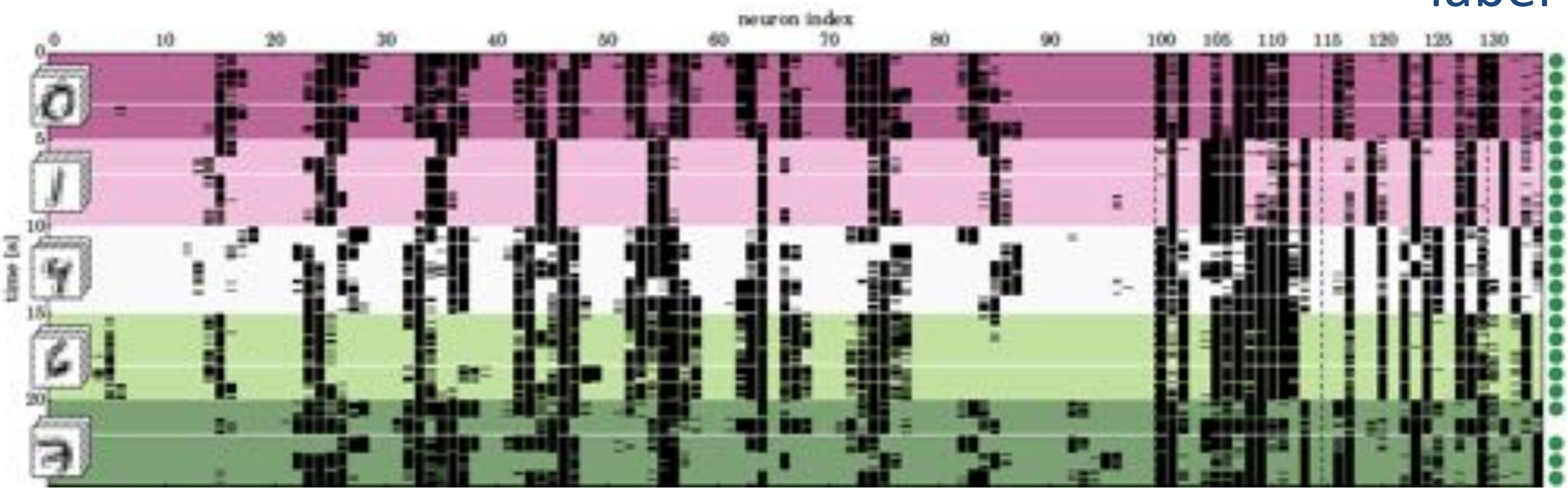
Feed-forward, rate-based. 4-layer spiking network
MNIST classification on a physical model machine
performance before and after hardware in-the-loop learning



https://arxiv.org/abs/1703.01909

MNIST classification on a physical model machine
Neuronal firing activity after hardware in-the-loop learning

label

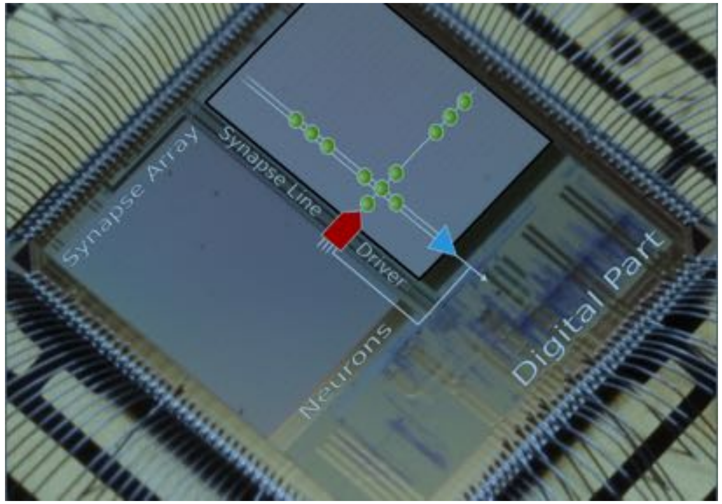


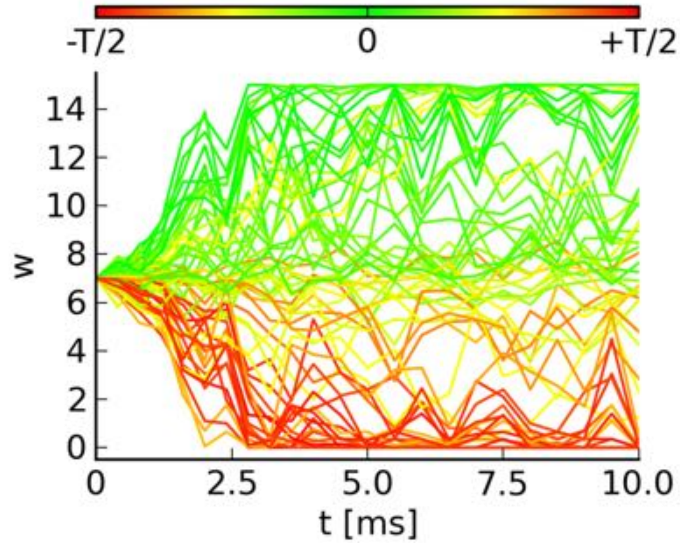
input

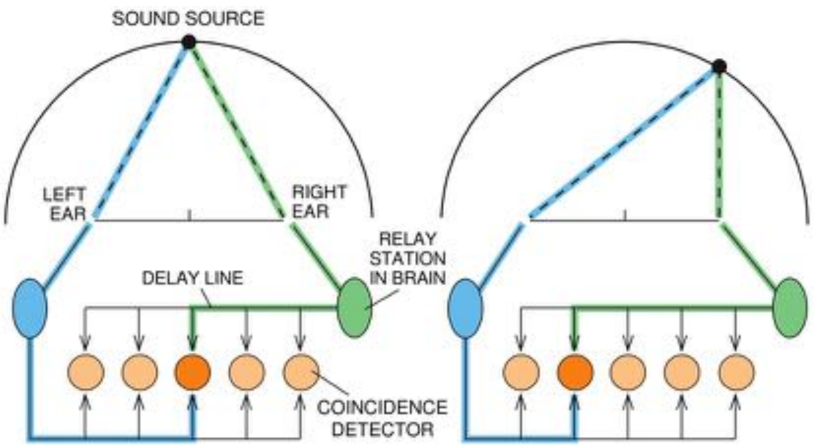
2 x hidden

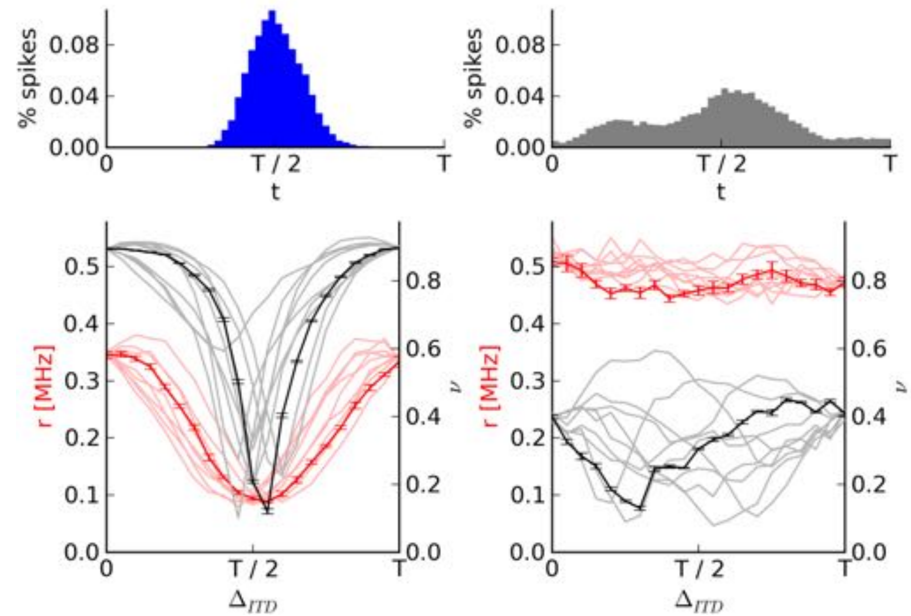
https://arxiv.org/abs/1703.01909

On-chip : Spike-Timing-Dependent-Plasticity

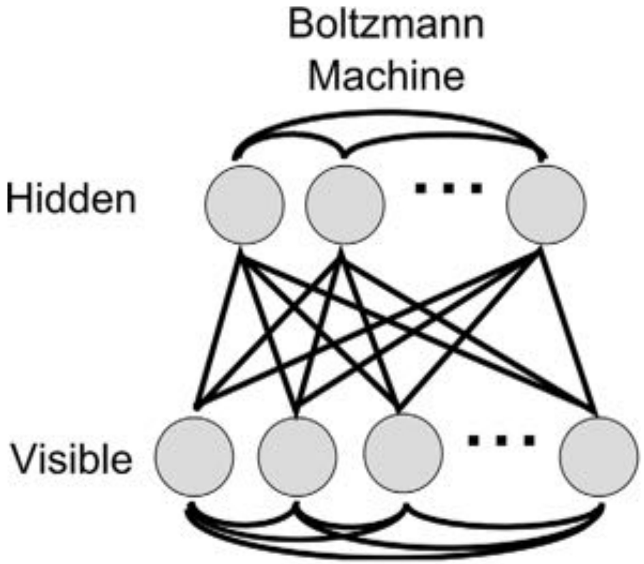


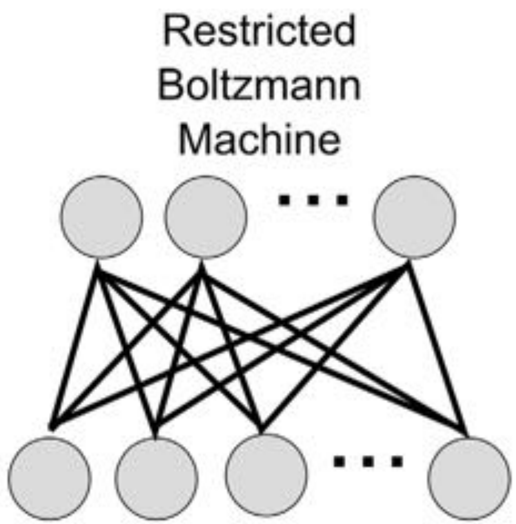






T. Pfeil, A.-C. Scherzer, J. Schemmel and K. Meier,
Neuromorphic Learning towards Nano Second Precision,
Proceedings of the 2013 International Joint Conference on Neural Networks (IJCNN).
Dallas, TX, USA: IEEE Press, 2013, pp. 869-873.





Boltzmann Machines

Networks of symmetrically connected stochastic nodes k

State of nodes described by vector of binary random variables z_k (0,1)

Probability for state-vector converges to a target Boltzmann-distribution

$$p(\vec{z}) = \frac{1}{Z} \exp\left[-E\left(\vec{z}\,\right)\right]$$

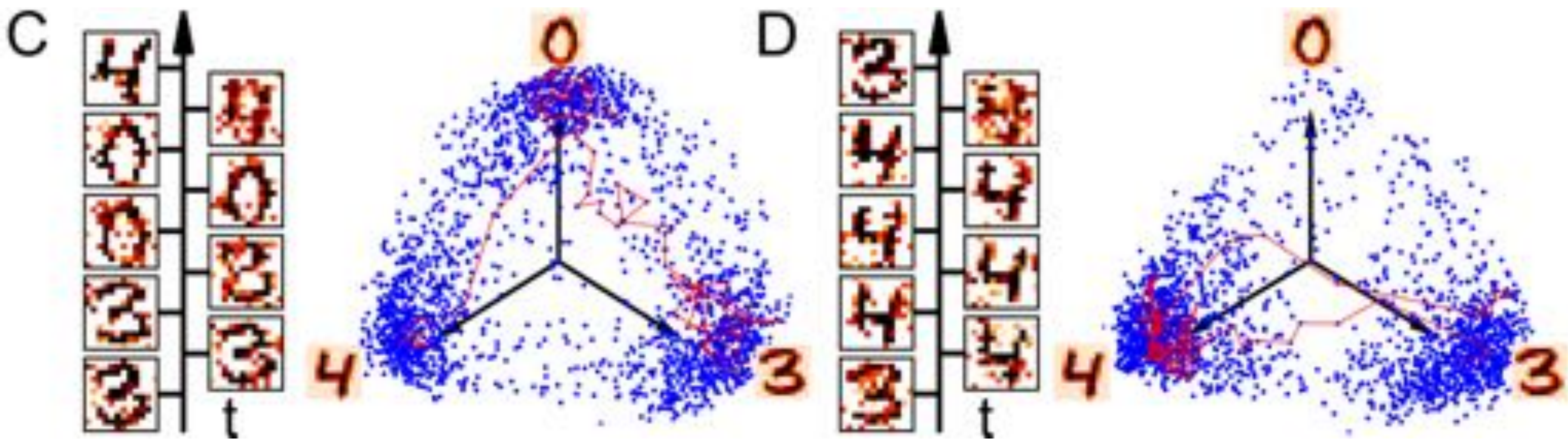
Energy function

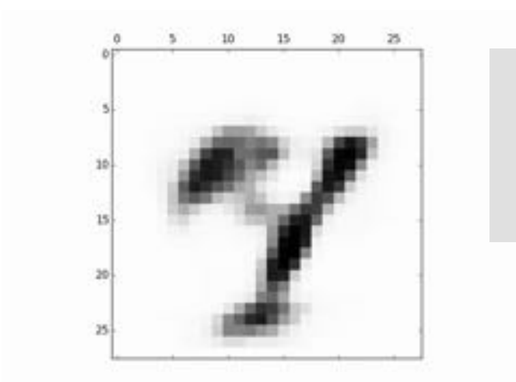
$$E(\vec{z}) = -\frac{1}{2}\sum_{i \neq j} w_{ij} z_i z_j - \sum_i b_i z_i$$

WHAT FOR ? Learn internal stochastic model of input space – Generate or discriminate

Learning specific input distributions by adjusting LOCAL interactions

- Clamp visible units to value of particular pattern – reach thermal equilibrium
- Incremement interaction between any 2 nodes that are both on
- Run network freely and sample from stored probability distribution
- Infer from clamped input

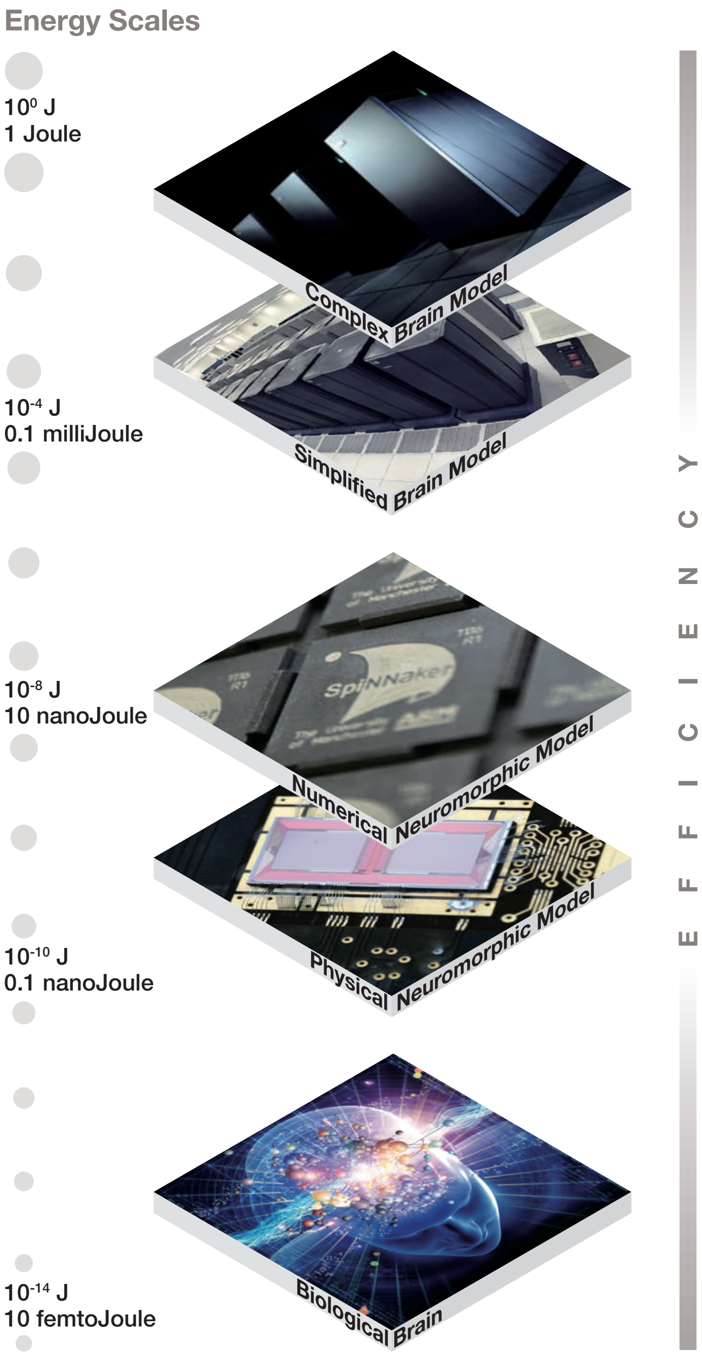




Free running
„Dreaming“
Generative

Inferring
Input incompatible with 0
Discriminative

PhD, Mihai Petrovici, BA Luziwei Leng 2015



Energy*Scales*

Energy used for a synaptic transmission

Filling the Gap

- Typically 10.000.000 times more energy efficient than state-of-the art HPC (comparable model)
- 10.000 less efficient than biology

From : HBP project report

TimeScales	Nature + Real-time	Simulation	Accelerated Model
Causality Detection	10^{-4} s	0.1 s	10^{-8} s
Synaptic Plasticity	1 s	1000 s	10^{-4} s
Learning	Day	1000 Days	10 s
Development	Year	1000 Years	3000 s
12 Orders of Magnitude			
Evolution	> Millenia	> 1000 Millenia	> Months
> 15 Orders of Magnitude			

Wide range of applications in particle physics

Offline data analysis

Spatial pattern detection and classification

- *flavour-tagging*
- *quark/gluon jet separation*
- *particle ID*
- *track / cluster finding*

Probability density calculation

- *Lifetimes*
- *Masses*

Online neural network trigger hardware systems

- *tracking*
- *background suppression*
- *topologies*

Good success, many real data publications

Substantial effort in training

How far have we come ?

- Excellent knowledge of structure
- Little knowledge of dynamics (prerequisite for a SM)
- Accelerated emulations method promising
- Very significant applications before understanding

An exciting field requiring strong experimental and theoretical skills – Physics may well play the crucial role towards a „standard model" of the brain