facets-project.org
brainscales.eu
brain-i-nets.kip.uni-heidelberg.de/
facets.kip.uni-heidelberg.de/ITN
bluebrain.epfl.ch/
apt.cs.man.ac.uk/projects/SpiNNaker/project/

Towards the FET Flagship Programme

FET FLAGSHIPS

HBP
The Human Brain Project

How to simulate what we don't understand

Karlheinz Meier
Ruprecht-Karls-Universität Heidelberg

Neuro-Inspired Computational Elements Workshop:

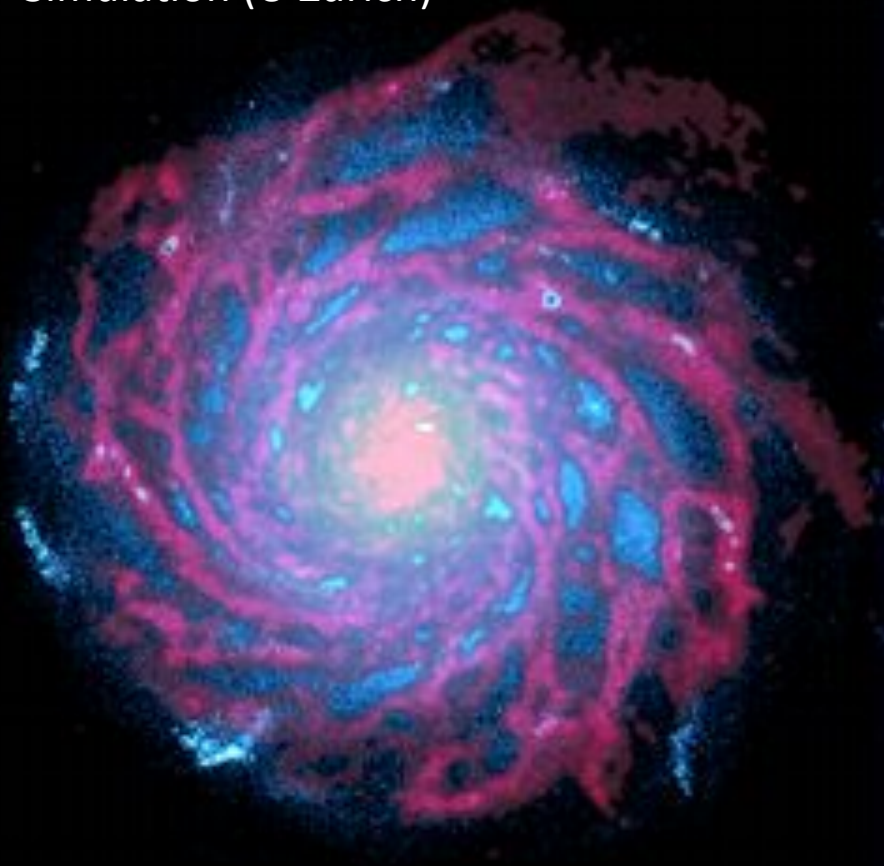
Information Processing and Computation Systems beyond von Neumann/Turing Architecture and Moore's Law Limits

February 25, 26, and 27, 2013
Albuquerque, New Mexico

Sandia National Laboratories

Simulation is a proven TOOL to UNDERSTAND the link between microscopic and macroscopic features !

Simulation (U Zürich)



M 74

©U Zürich and NASA

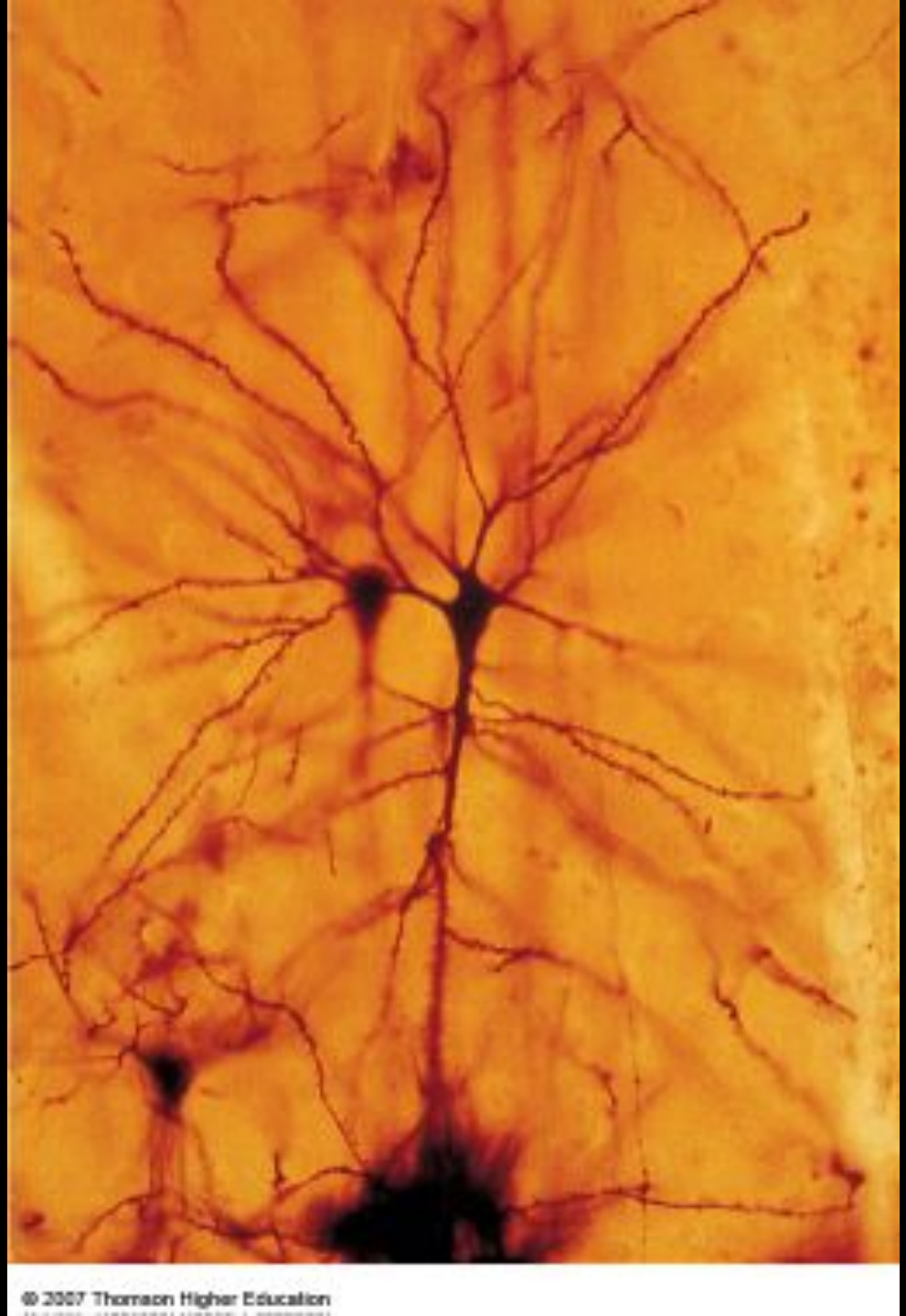
Santiago Ramon y Cajal
(1852-1934)

Individual cells in the brain are spatially separate objects

"interaction over a distance"

and

"spatial and temporal integration"



Biology vs. Numerical Simulation

A naive and not quite correct scaling exercise

Cortical Column

0.000003 Watt

100.000 Watt

Mouse Brain

0.03 Watt

1.000.000.000 Watt

Human Brain

30 Watt

1.000.000.000.000 Watt

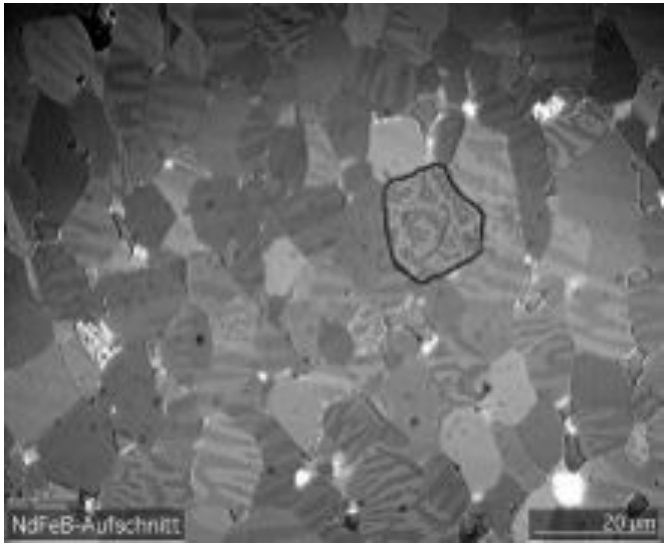
Installed in D (2010) : 170.000.000.000 Watt

AND : typically 100 times slower than biology

© BlueBrain Project (EPFL)

Condensed Matter Physics

Complex Multiparticle (10^{23}) System (e.g. Ferromagnet)

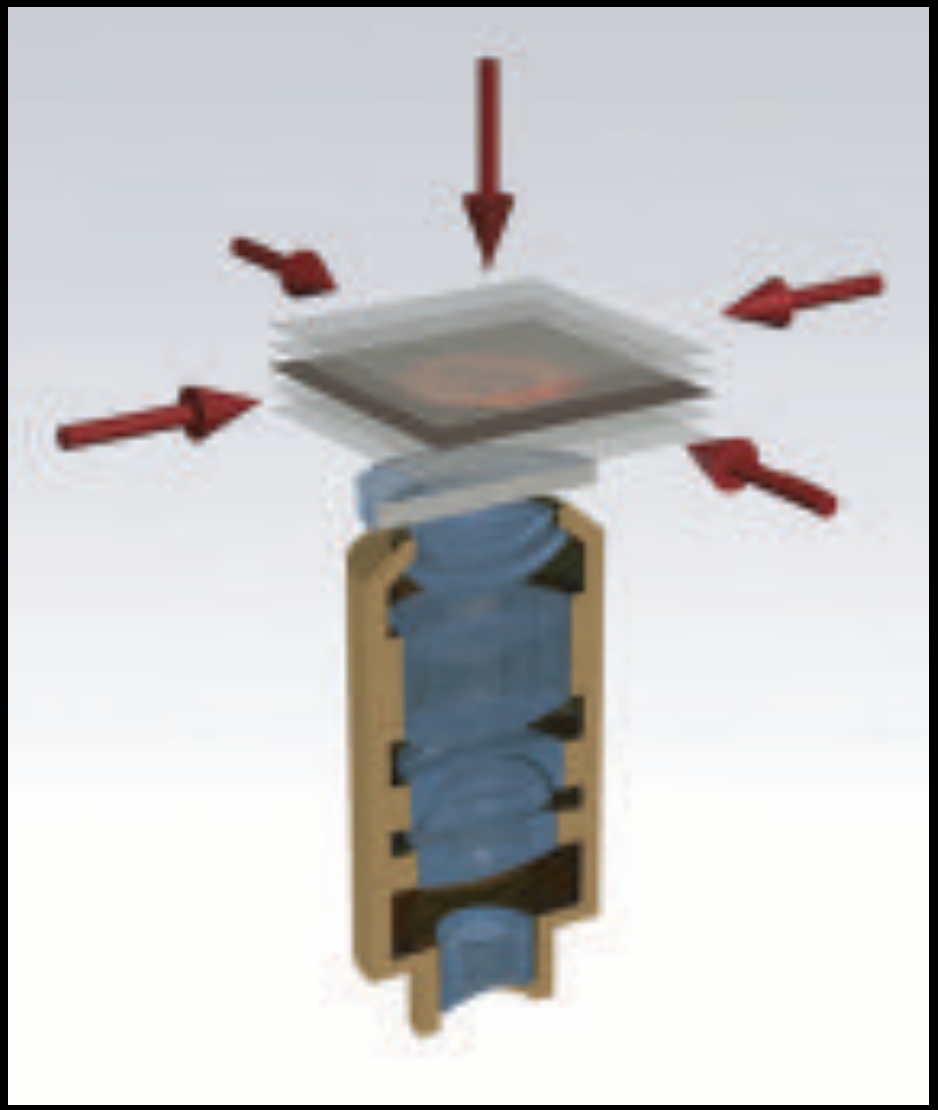


Physics approaches towards Understanding of Multiparticle Systems :

- Systematic experimental observation of collective properties (magnetisation, phase transitions)
- Probe preparation to observe (some) microscopic features
- Model building, theory and analytical treatment
- Model building, theory and numerical treatment
- Controlled, reduced size, physical model system (synthesis) ?
- e.g. real spins in a synthetic lattice ?
- QUANTUM EMULATORS ?

Quantum Emulators – Synthesize what you cannot compute

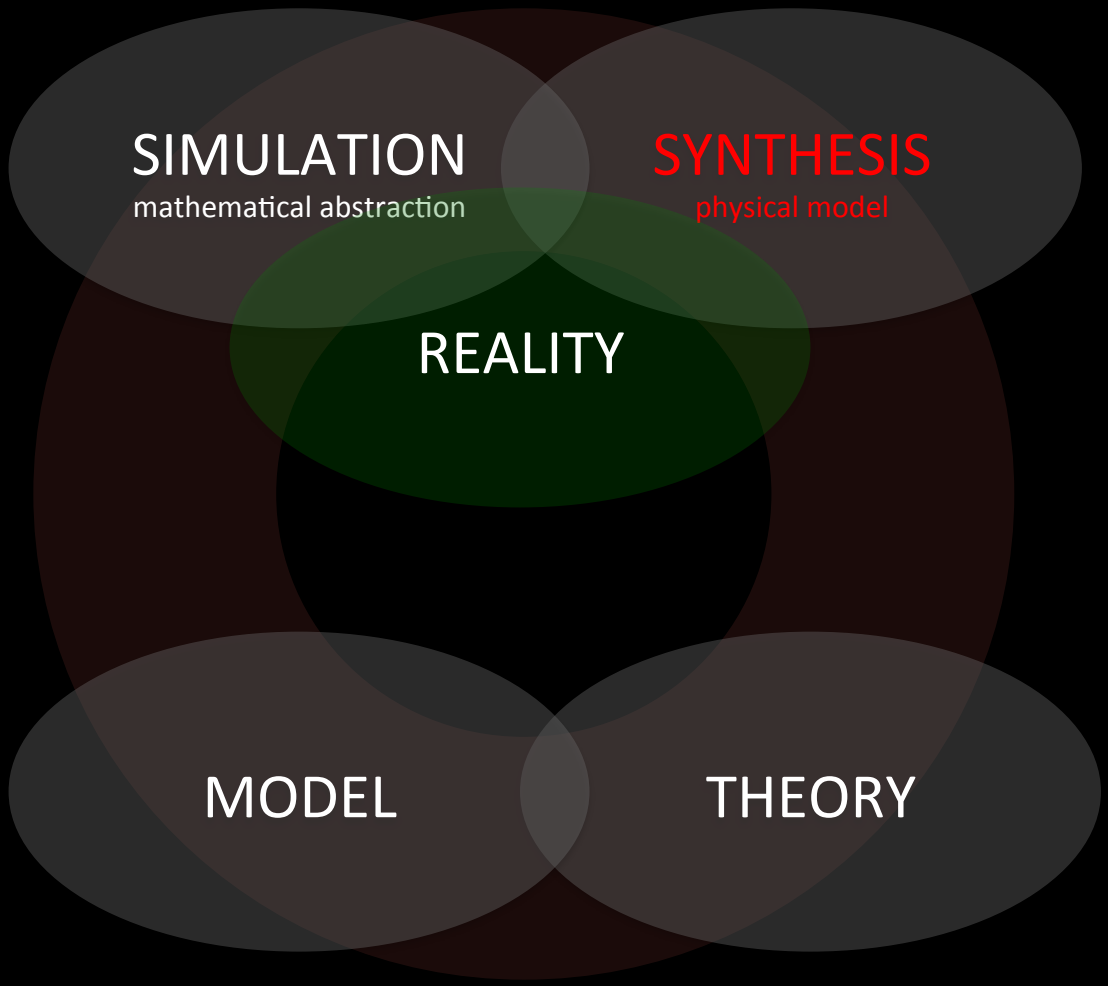
Bloch, Dalibard, Nascimbène, Nature Physics, Vol. 8, April 2012
Feynman, Int. J. Theor. Phys., Vol. 21, Nos. 6/7, 1982



„Ultracold quantum gases offer a unique setting for quantum simulation of interacting many-body systems. The high degree of controllability, the novel detection possibilities and the extreme physical parameter regimes that can be reached in these

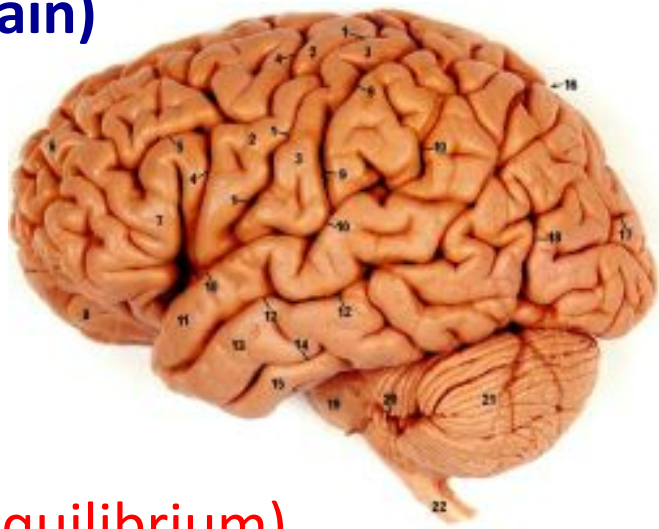
'artificial solids'

provide an exciting complementary set-up compared with natural condensed-matter systems, much in the spirit of Feynman's vision of a quantum simulator"


SIMULATION
mathematical abstraction
SYNTHESIS
physical model
REALITY
MODEL
THEORY

More difficult :

Complex Multiparticle ($10^{11}/10^{15}$) System (Brain)



Major differences !

- Short, medium and long range interactions
- High dimensional (>> 3) network topology
- Active, adaptive nodes
- Time dependent (far away from „thermal“ equilibrium)
- **Stores and processes information !**

Approaches towards understanding :

- Observation of collective properties (memory, synchrony, ...)
- Probe preparation to observe (some) microscopic features
- Model building, theory and analytical treatment
- Model building, theory and numerical treatment
- Controlled, reduced size, physical model system ?
- NEUROMORPHIC SYSTEMS !

The Key Arguments for Neuromorphic Computing

- Low Power (energy per fundamental operation)
- Fault Tolerance
- Plasticity / Learning / Development („no algorithm“)
- Speed
- Scalability

The Key Challenges for Neuromorphic Computing

- Neuroscience Knowledge, *Flexibility*
- Configurability, *Technologies for distributed memory*
- Integration Density, *nano-components, 3D Integration*
- Circuit re-Use, *CAE Tools*
- User Access, *Unified software toolset*

But there is more to this ..

A radical departure from what is probably the most successful "science-to-technology roadmap" of humankind :

Boole (Theory) – Shannon (Circuits) – Turing (Programmability) – von Neumann (Architecture) – Kilby (Integration)

The brain does NOT :

- use Boolean logic (Boole)
- use electronic realisations of logic gates (Shannon)
- use programmed code (Turing)
- use separate memory and compute units (von Neumann)
- use quasi two-dimensional wiring (Kilby)

Where to go from here ?

Knowledge about the brain is INCOMPLETE

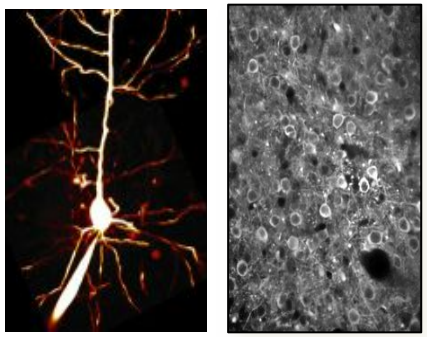
Cell complexity ?
Parameter ranges ?
Network architectures ?
Missing Pieces !

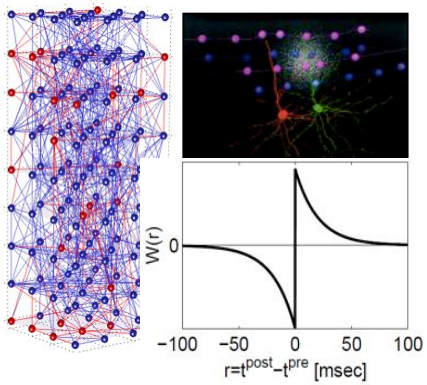
Wait *(forever)* until things are solved .....

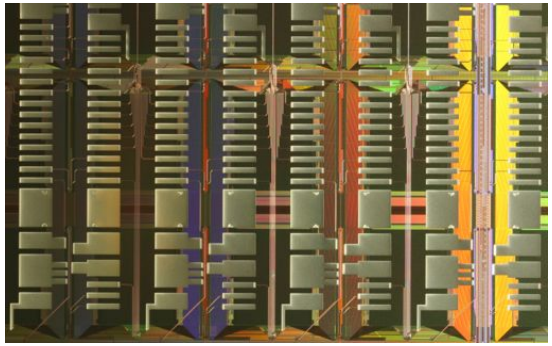
Definitely not, instead

- Build in KNOWN features NOW
- Learn HOW to build neuromorphic systems
- Make systems highly CONFIGURABLE „Neural FPGA Concept“
- Make neuromorphic systems available for NON-EXPERTS

The BrainScaleS project aims at understanding function and interaction of multiple spatial and temporal scales in brain information processing based on in-vivo experimentation and computational analysis. Generic theoretical principles will be extracted to perform artificial synthesis of cortical-like cognitive skills in numerical simulations and on a novel hybrid multiscale neuromorphic computer







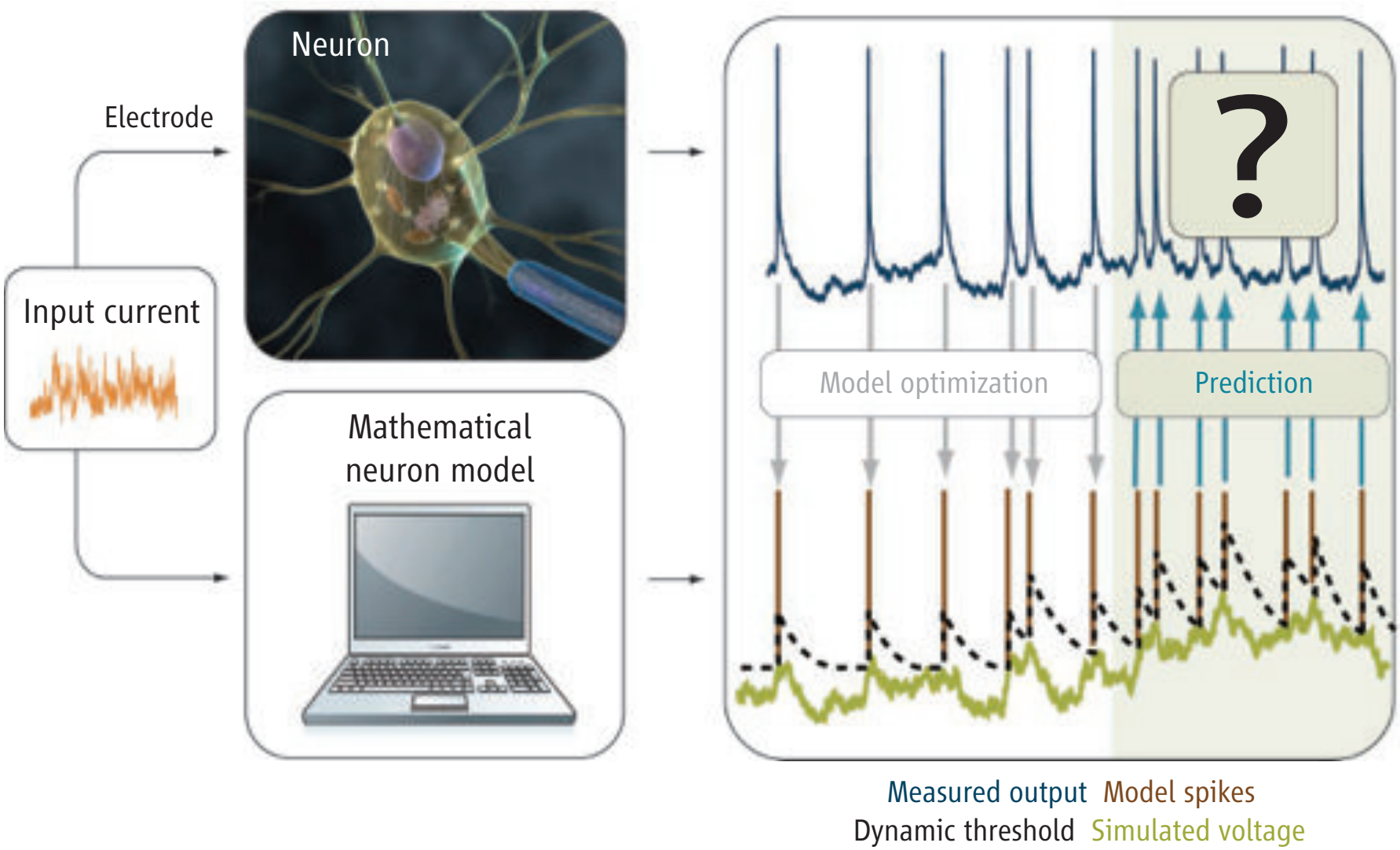
Experiments

Theoretical models
Mathematical analysis
Computer simulation

Hybrid Multiscale Facility (HMF)

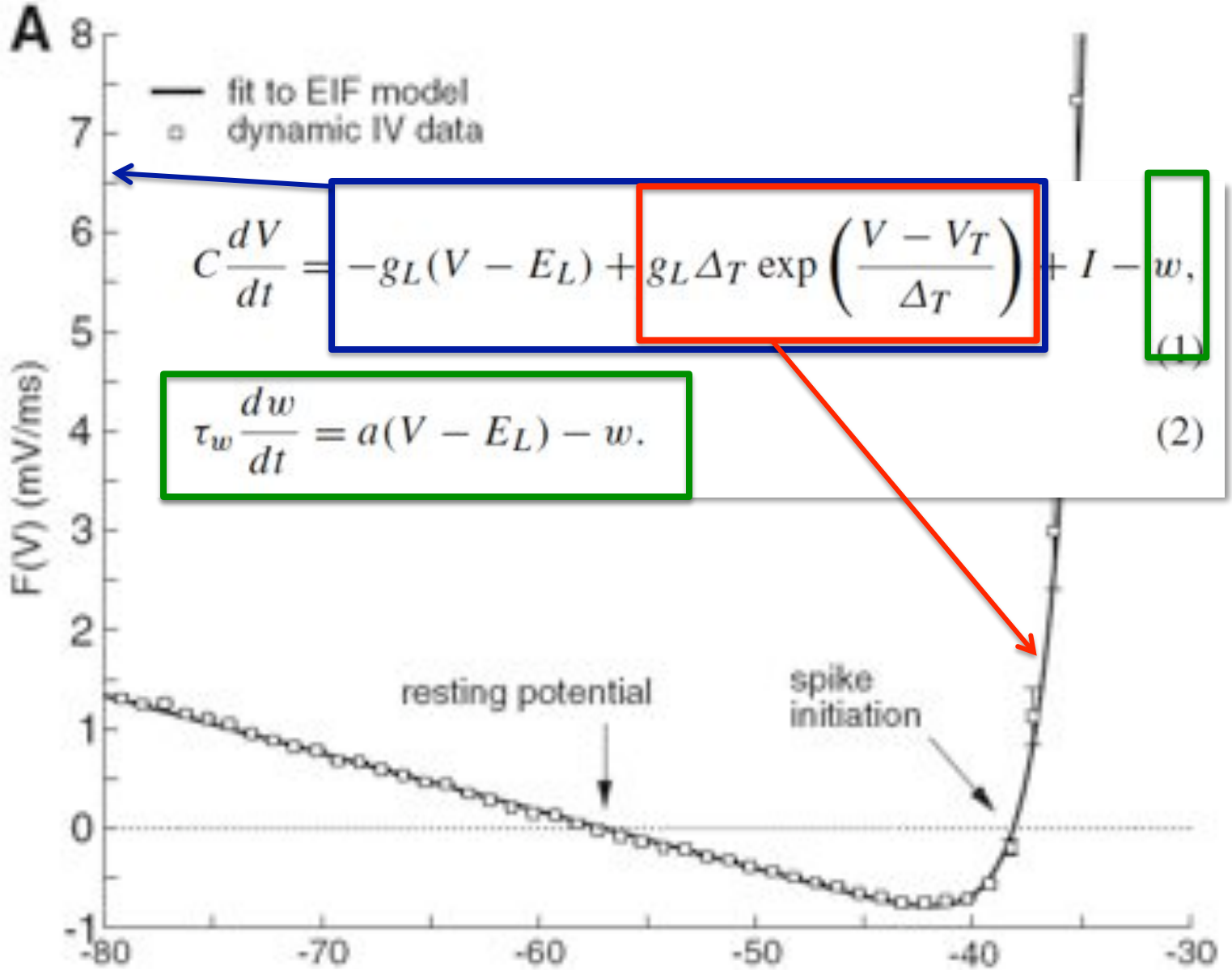
BrainScaleS

Developing and Tuning Neuron Models – From Biology to Mathematics to VLSI



Gerstner, Naud, Science 326 (2009) 379 + references therein

Adaptive-Exponential Integrate-and-Fire Neuron Model (I)



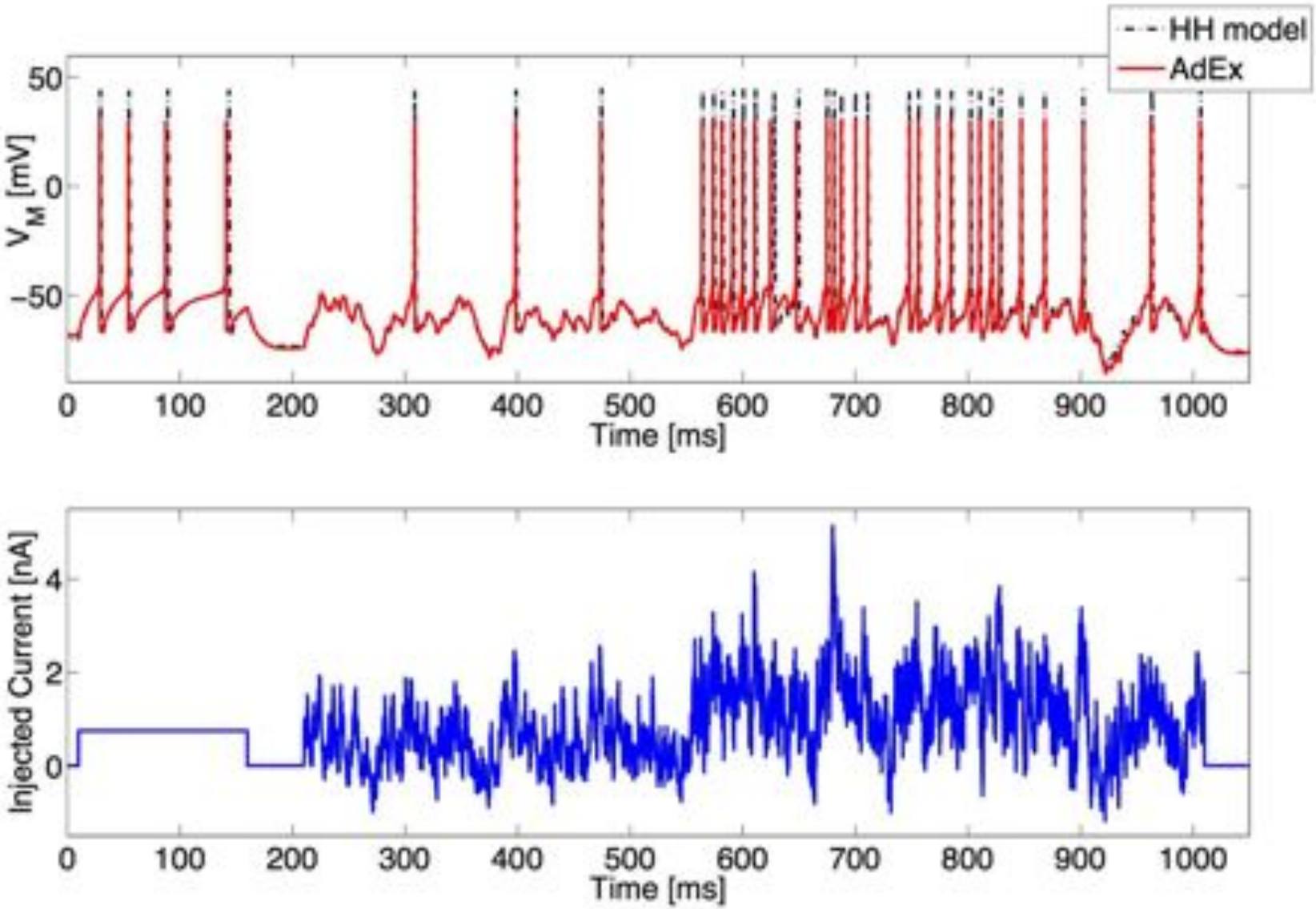
$$C\frac{dV}{dt} = -g_L(V - E_L) + g_L \Delta_T \exp\left(\frac{V - V_T}{\Delta_T}\right) + I - w, \quad (1)$$

$$\tau_w \frac{dw}{dt} = a(V - E_L) - w. \quad (2)$$

R. Naud et al.
Biol Cybern (2008) 99:335–347

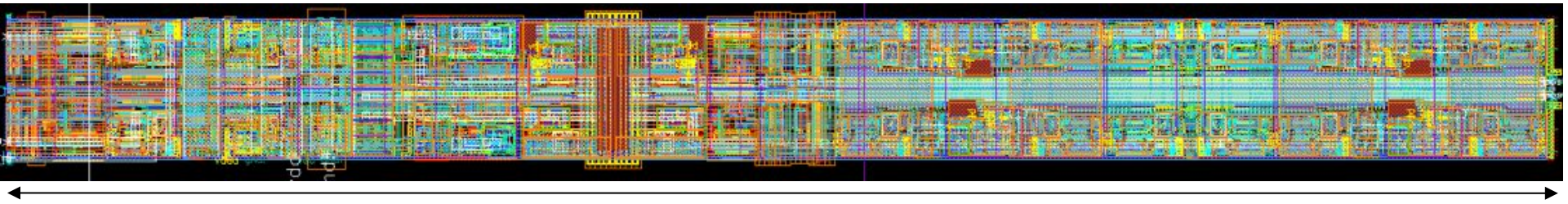
www.facets-project.org

The Adaptive-Exponential IF Neuron Model (II)

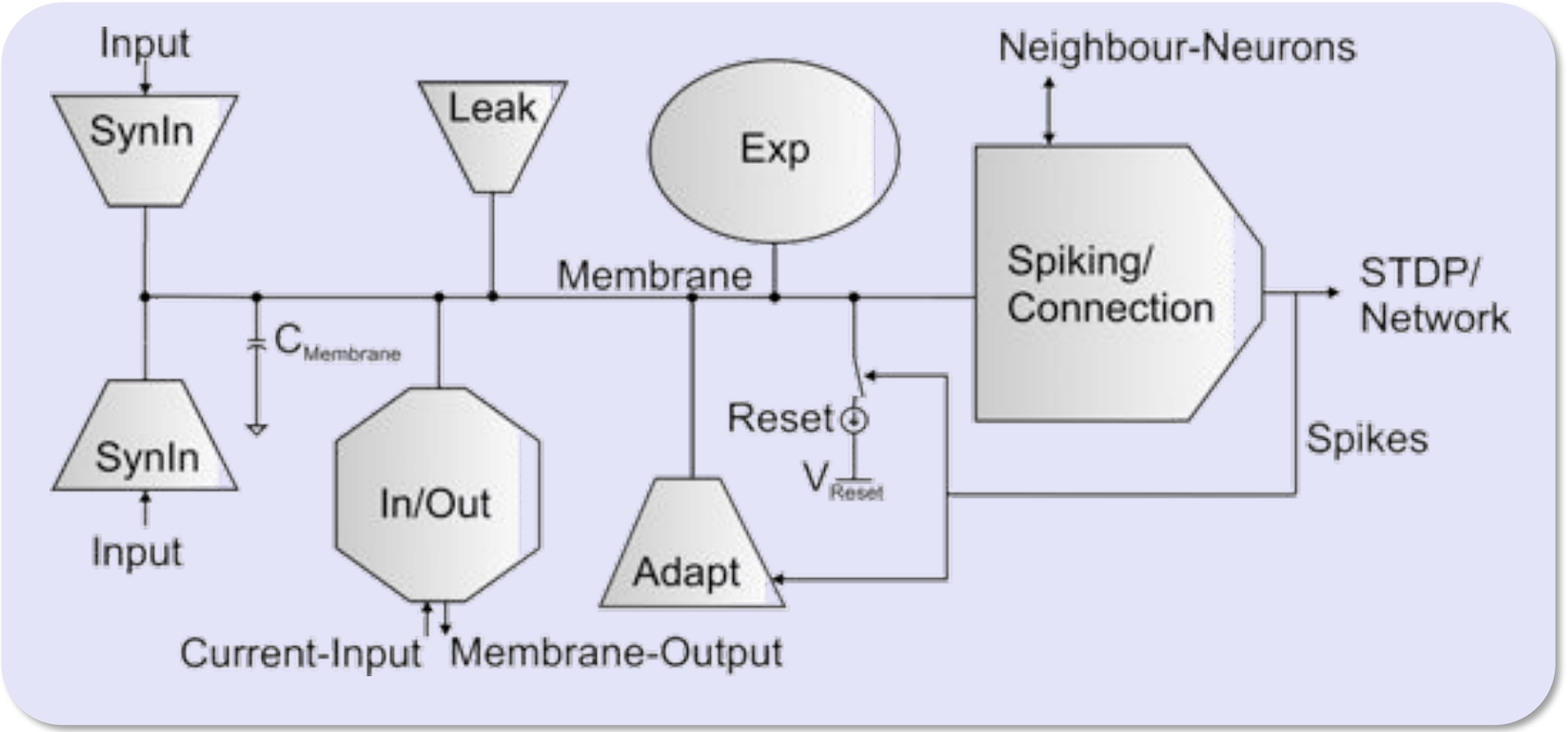


Brette, Gerstner, Adaptive Exponential Integrate-and-Fire Model as an Effective Description of Neuronal Activity, *J Neurophysiol 94: 3637-3642, 2005*

From Mathematics to Electronics

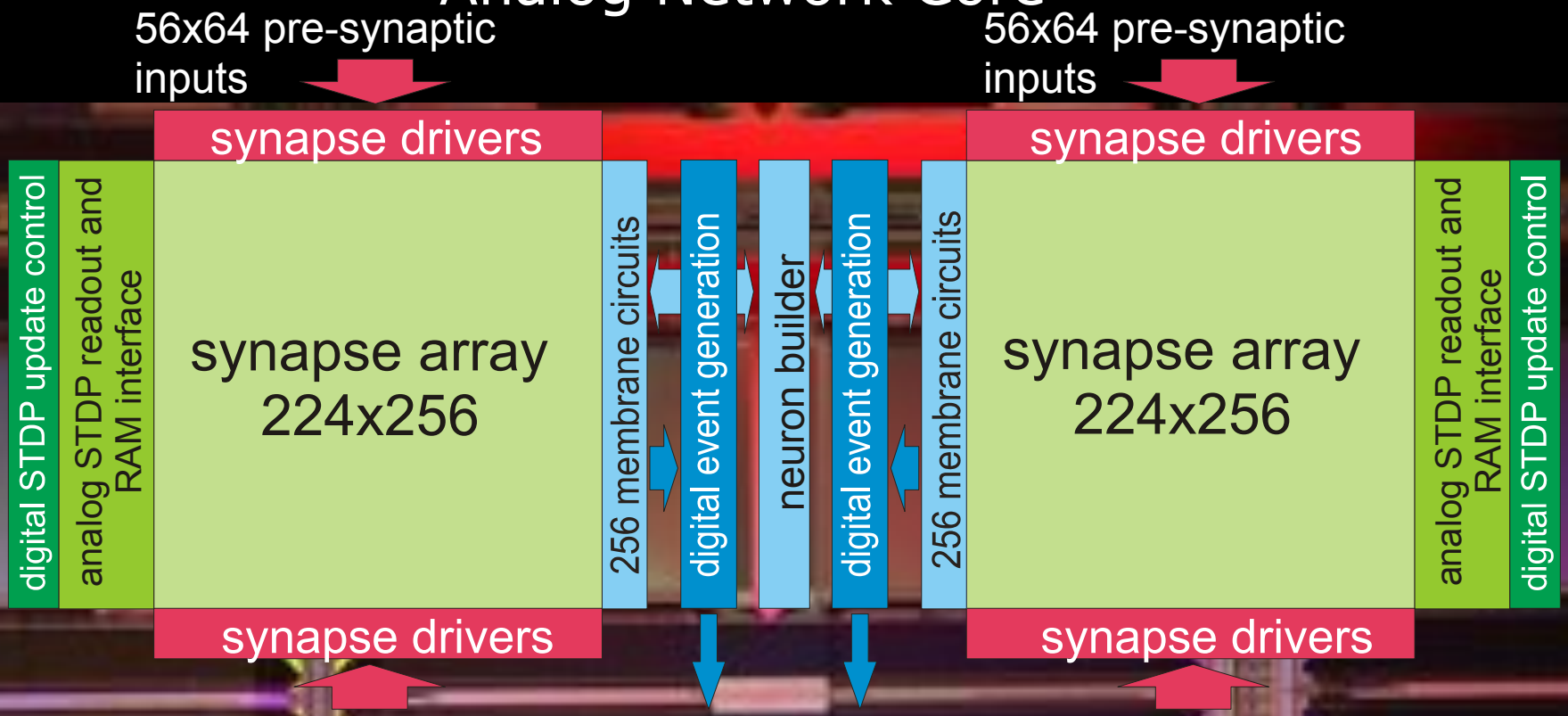


layout drawing of two neurons: 150x20 µm^2



S. Millner et al., A VLSI Implementation of the Adaptive Exponential Integrate-and-Fire Neuron Model, NIPS 2011

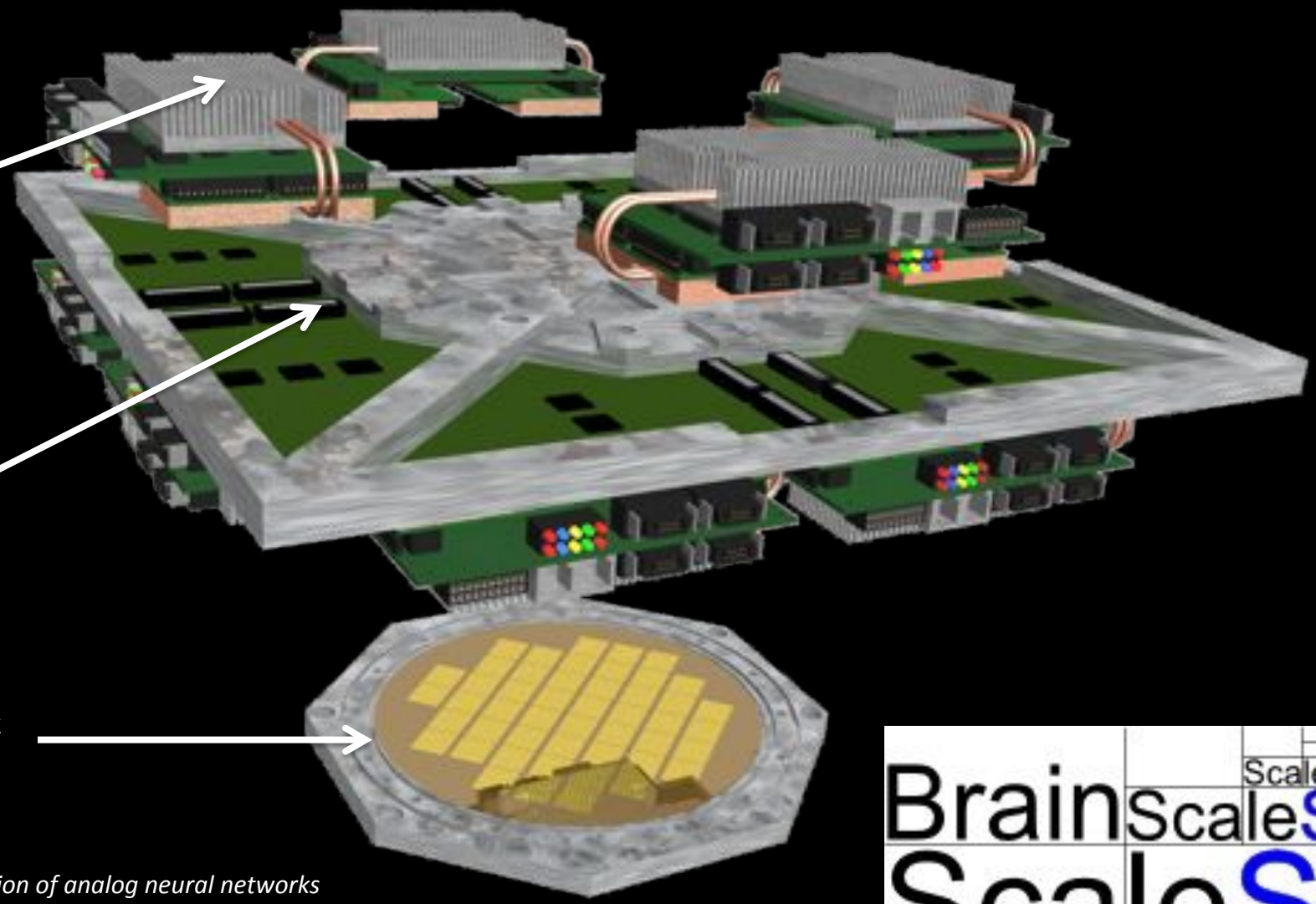
Analog Network Core



Millner, S., Grübl, A., Meier, K., Schemmel, J. and Schwartz, M.-O.
A VLSI Implementation of the Adaptive Exponential Integrate-and-Fire Neuron Model
Advances in Neural Information Processing Systems (NIPS) (2010)

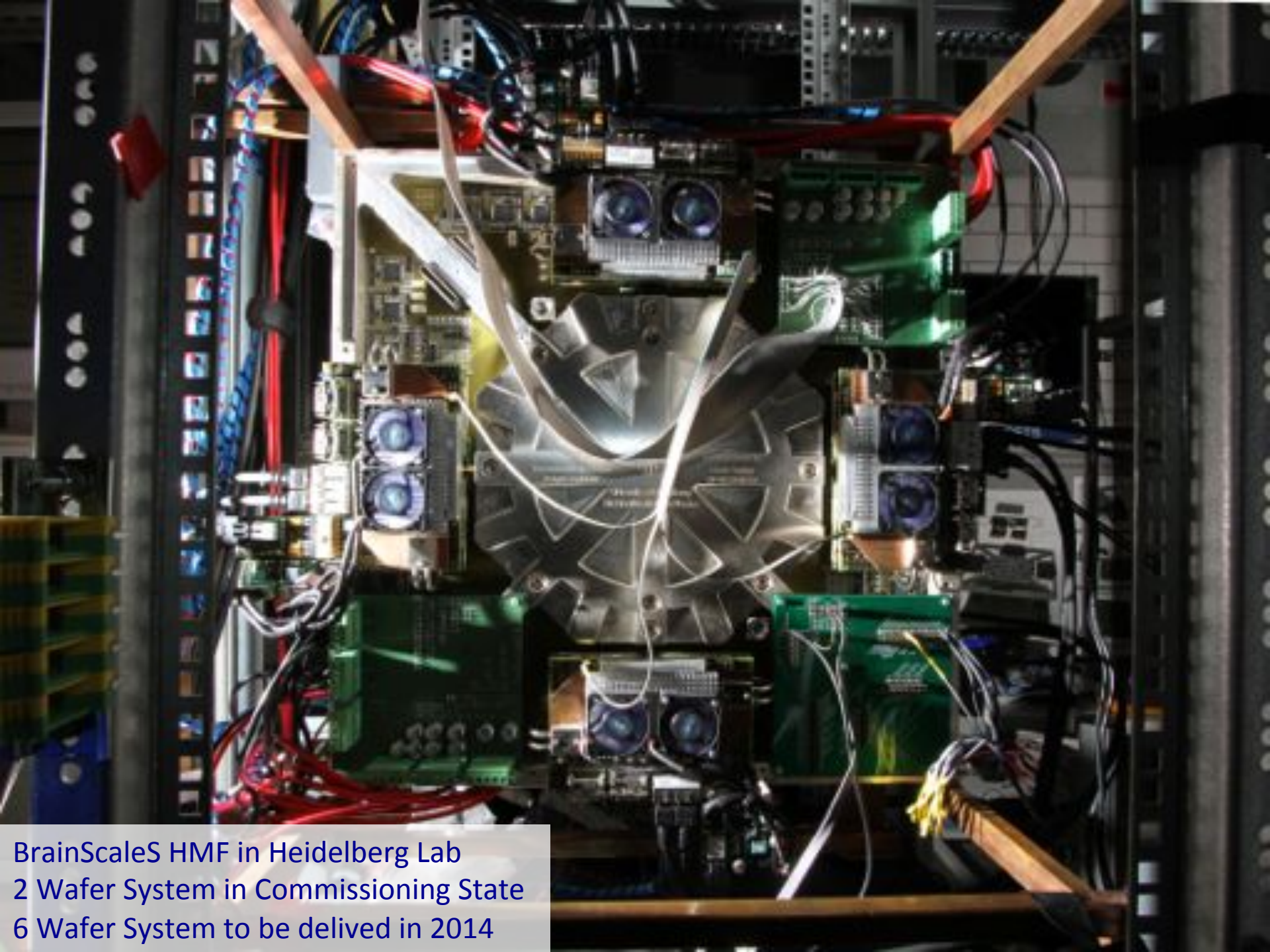
combining multiple membrane circuits with 256 synapses each allows neurons with up to 16k pre-synaptic inputs

Neural Processing Unit,
200.000 Neurons, 50.000.000 plastic Synapses, 16.000 syn. I/P per neuron
Separation of Neural Circuits and Monitoring/Readout/Control



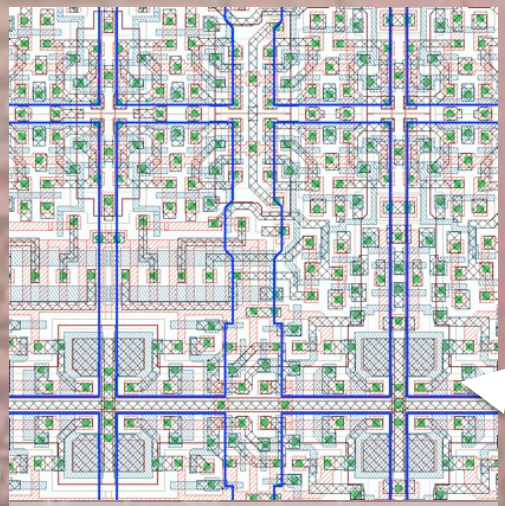
Wafer-scale integration of analog neural networks
J. Schemmel, J, Fieres and K. Meier
In : Proceedings of IJCNN (2008), IEEE Press, 431

BrainScaleS



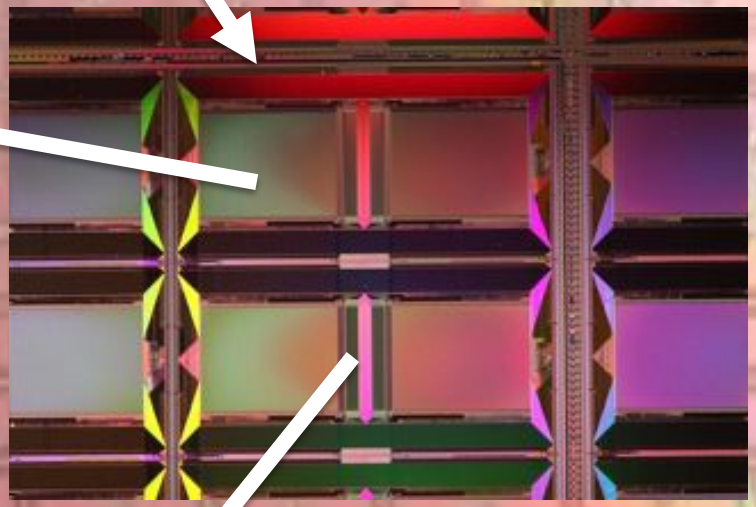
BrainScaleS HMF in Heidelberg Lab
2 Wafer System in Commissioning State
6 Wafer System to be delived in 2014

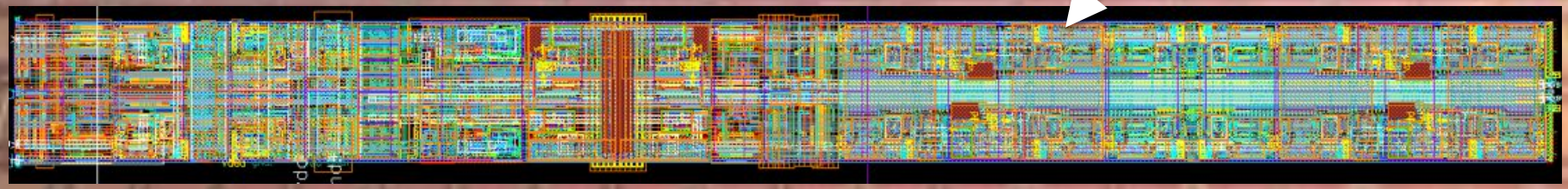
Multi-Scale Circuit Structure on 8 inch CMOS Wafer (180nm)



Plastic Synapses, 50.000.000 Million Instances on Wafer, Length Scale 10 µm 4-bit SRAM Weights, STP, STD, STDP

High Input Count Network Chips, 400 Instances on Wafer, Length Scale 1 cm network routing



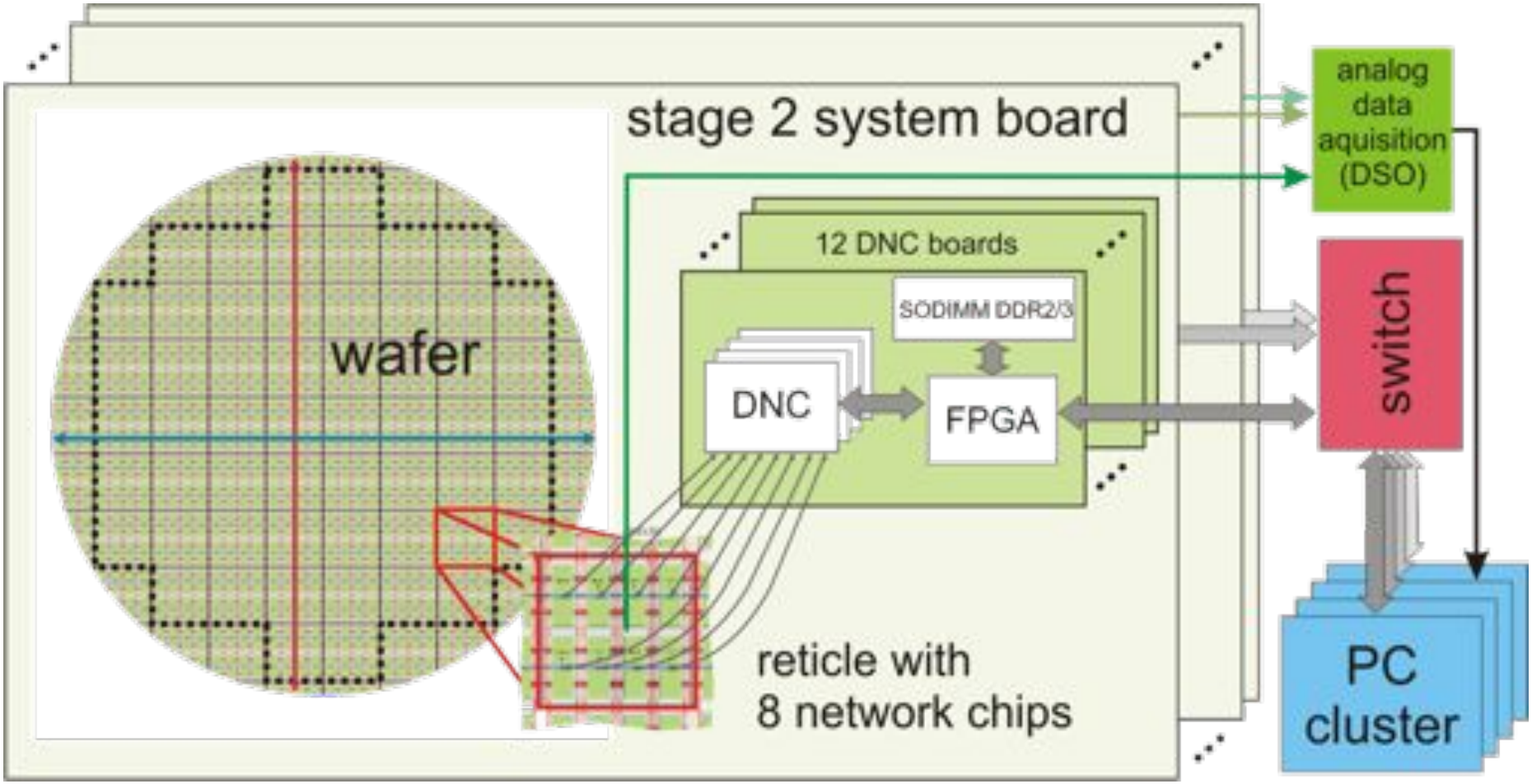


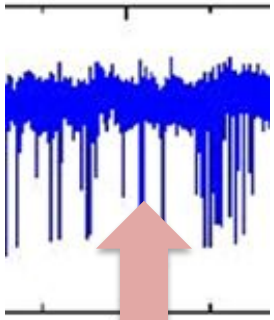
AdEx Neurons, 200.000 Instances on Wafer, Length Scale 300 µm,
Analog Floating Gate Parameter Storage
Poisson Noise Generators

Communication Architecture Overview

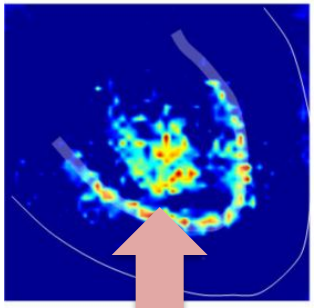
Hierarchical 2 Layer Communication Setup

- Layer 1 : **On-Wafer** continuous, asynchronous, fixed delay spike transmission
- Layer 2 : **Off-Wafer** Packet based digital communication for medium and long-range spike transmission, network set-up, read-out and control.

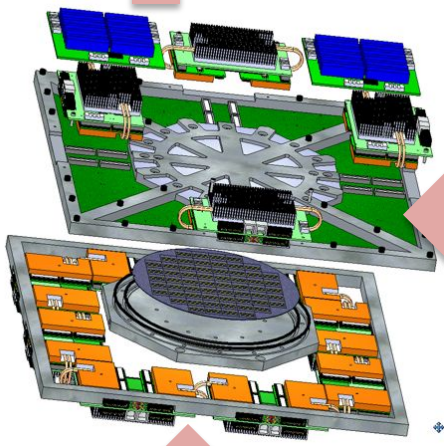




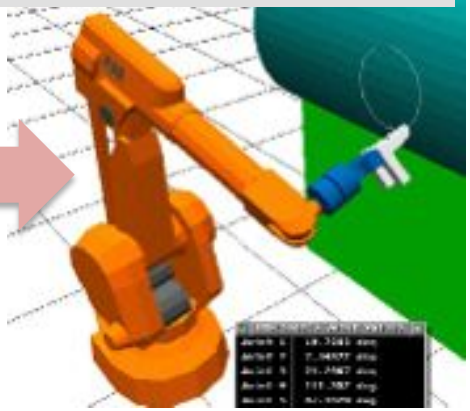
Output of
biologically
equivalent
data



Hybrid
(Neuromorphic-HPC)
Multiscale Modelling
BrainScales HMF



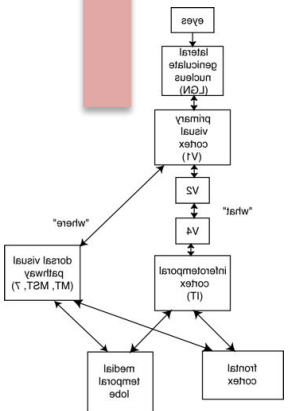
Virtual environment



$$c_m \frac{dV}{dt} = -g_{leak}(V - E_l)$$

$$\Delta t = t_{post} - t_{pre}$$

Input of
microscopic
and
macroscopic
theory

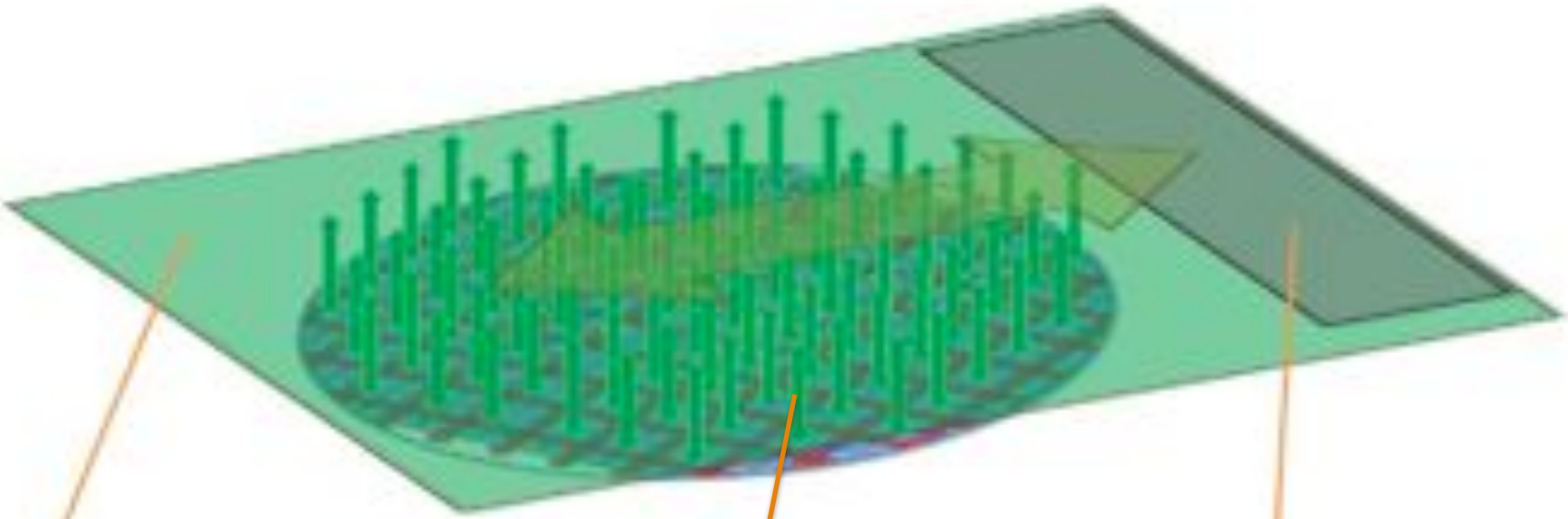


Microscopic – macroscopic

Milliseconds – years

Rapid cycling of experiments

Aggregate Bandwidth from off-Wafer Connections exceeds 1 Terabit/s
On-wafer Information Flow >> 1 Terabit/s



Printed Circuit Board

10.000 Vertical Links

Deliver approx. 500 A
to 320 cm^2 of active silicon
Send and Receive 2 Gbit/s/Link

Wafer-to-Wafer
Routing Circuits

Energy efficiency ?

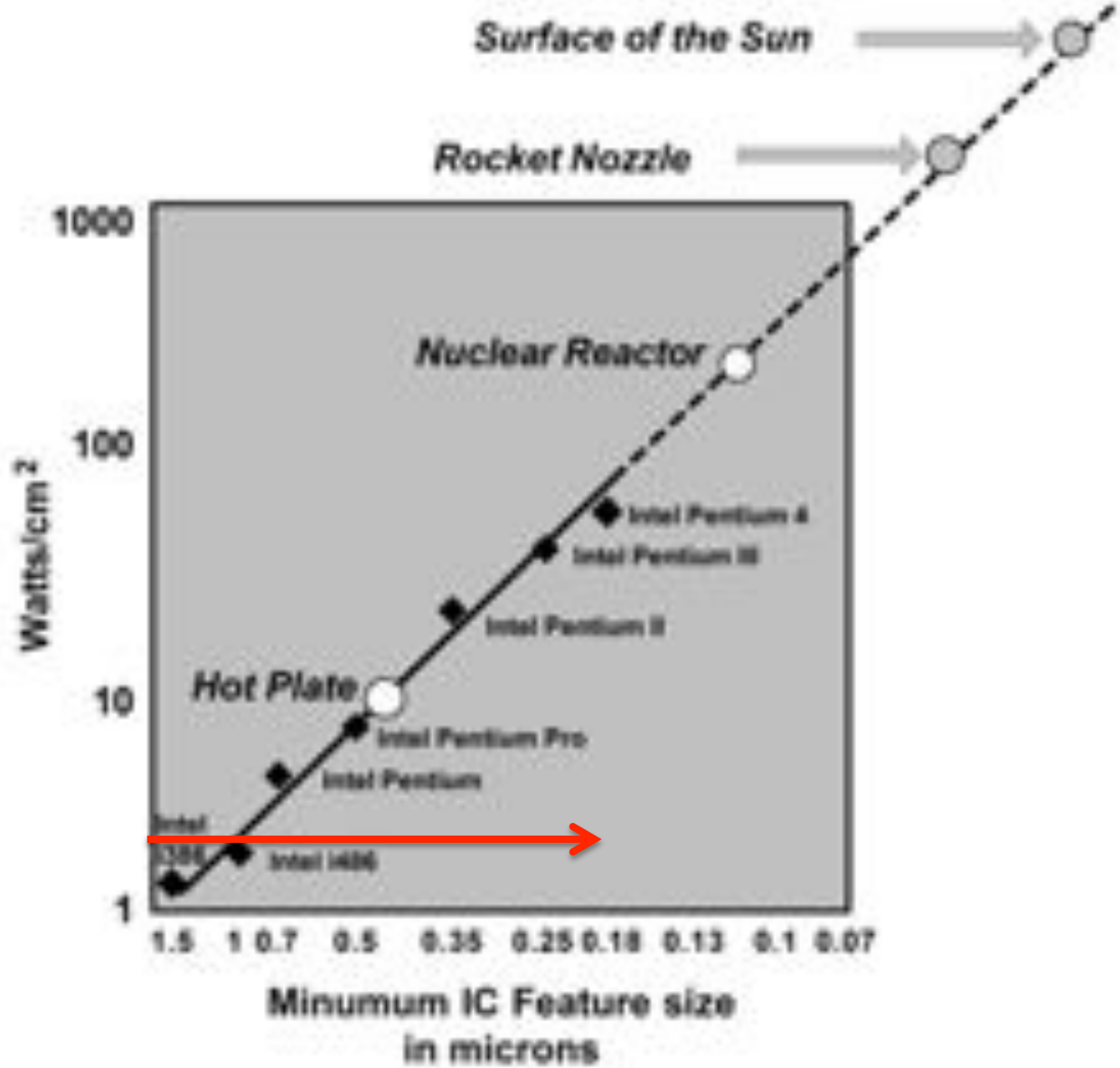
Current Status – Not optimized for low power

Accelerated BrainScaleS ANN wafer at conservative (maximum) rate : 3 Watts/cm^2

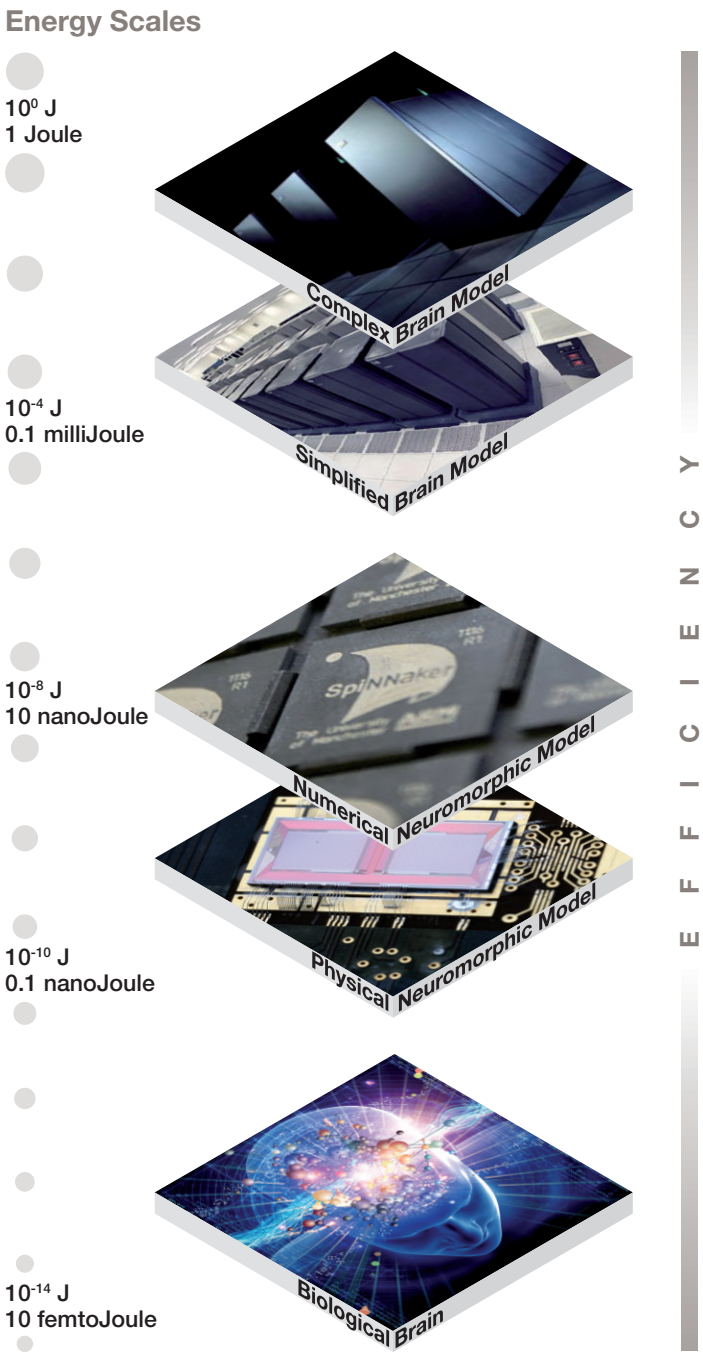
Approximately 10^{-10} J per synaptic transmission

Biology : 10^{-14} j per synaptic transmission

X 10.000 gap downwards to biology but another large upwards gap to HPC simulations



F Pollack
"New microarchitecture challenges in the coming generations of CMOS process technologies"
MICRO-32, Haifa, Israel, 1999.



Energy*Scales*

Energy used for a synaptic transmission

14 orders of magnitude difference for *„the same thing“*

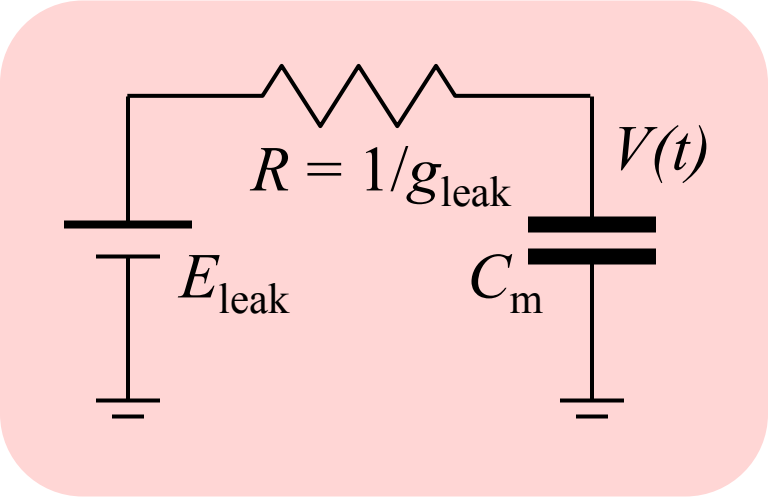
Physical models (*Neuromorphic*)

- 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

Continuous Time Integrating Neural Cell Membrane Model - *Neuromorphic*

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



	ΔV [V]	g_{leak} [S]	C_m [F]	$(g\Delta V)/C$ [V/s]
Biology(*)	10^{-2}	10^{-8}	10^{-10}	10^{0}
VLSI	10^{-1}	10^{-6}	10^{-13}	10^{6}

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

Inherent speed gap:

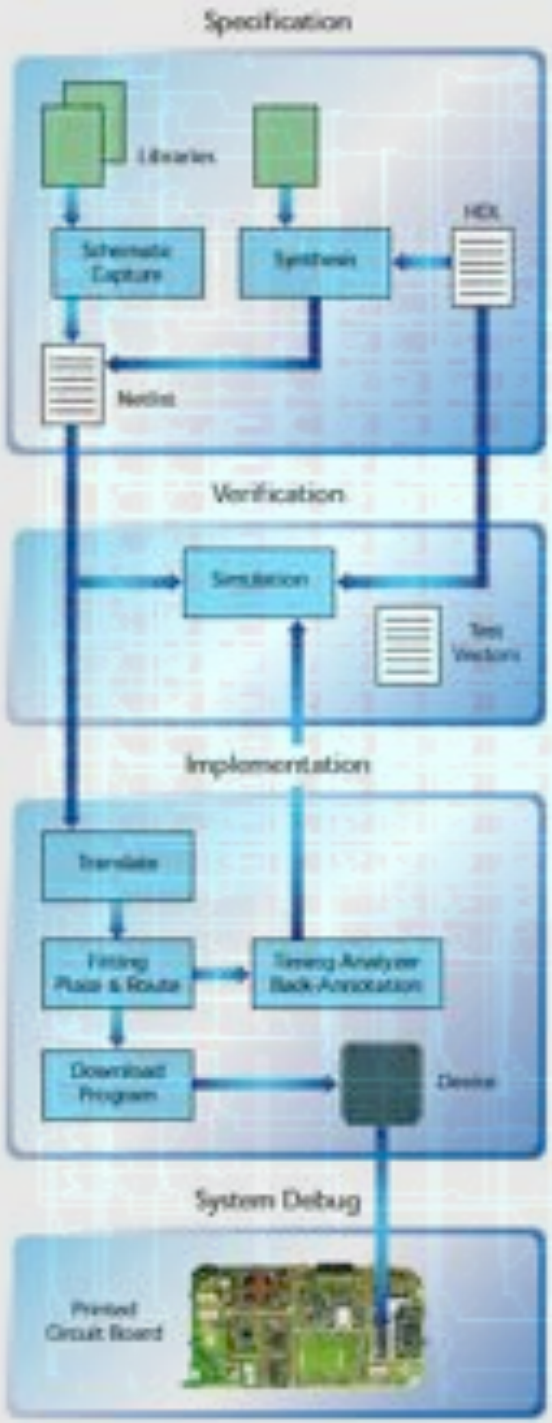
10^6 Volt/second

→ accelerated neuron model

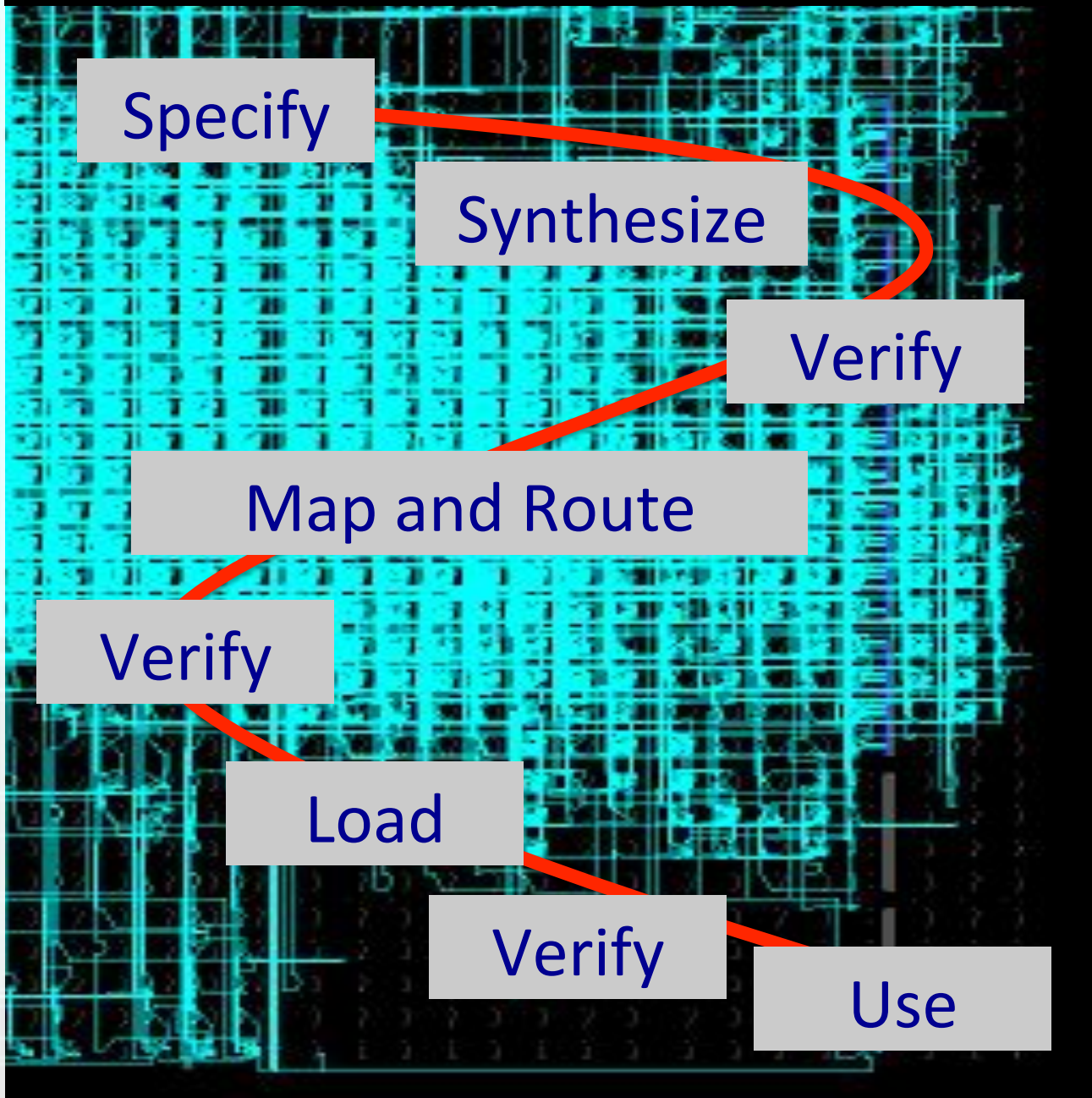
Time*Scales*

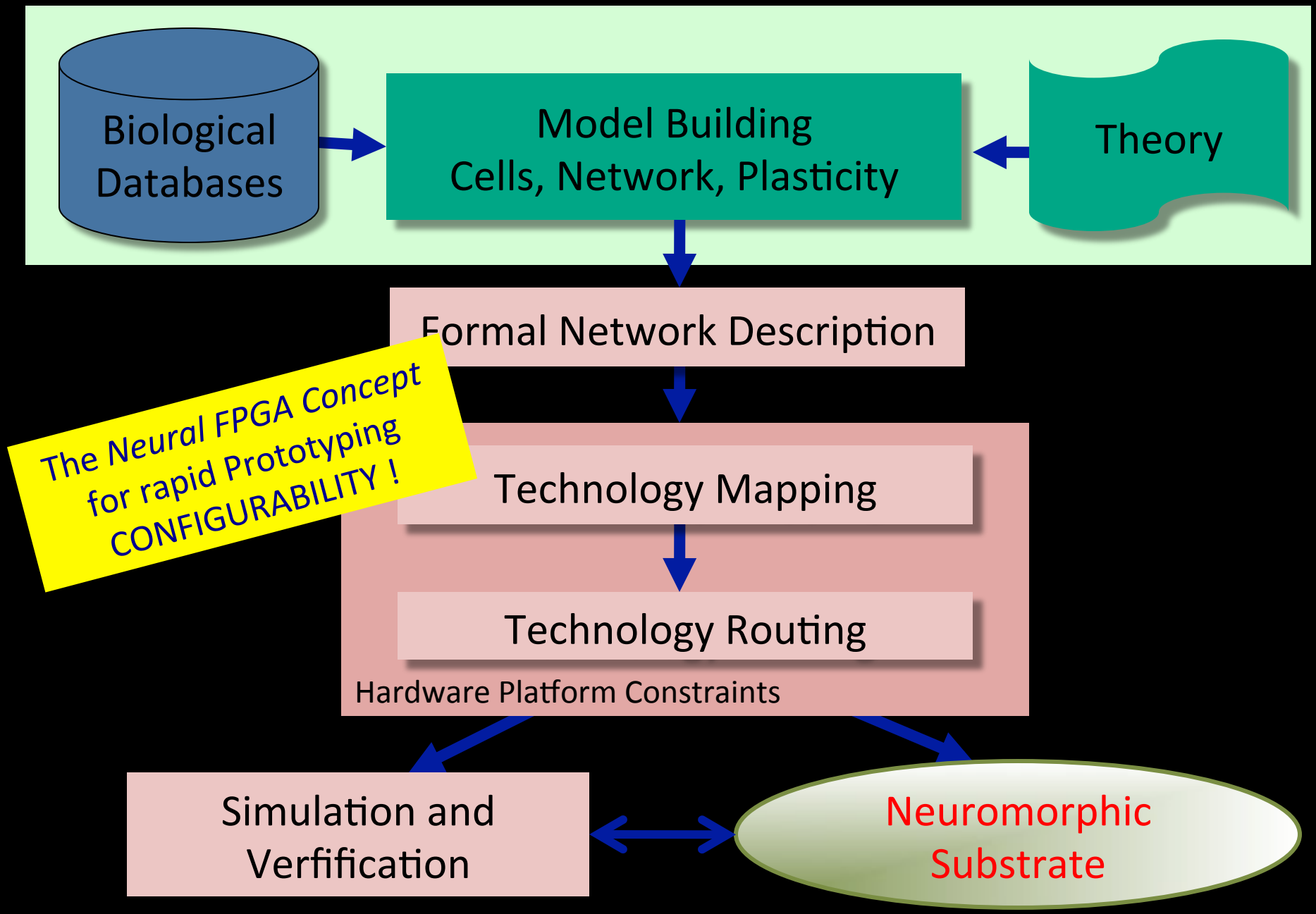
	Nature	Simulation	Accelerated physical model
Detection of causality	10^{-4} s	0.1 s	10^{-8} s
Plasticity	1 s	100 s	10^{-4} s
Learning	day	100 days	10 s
Development	year	100 years	3000 s
12 Orders of magnitude			
Evolution	> millenia	> 100 millenia	> month
> 15 Orders of magnitude			

Temporal dynamics is key to understanding (and using) the computational paradigms of the brain

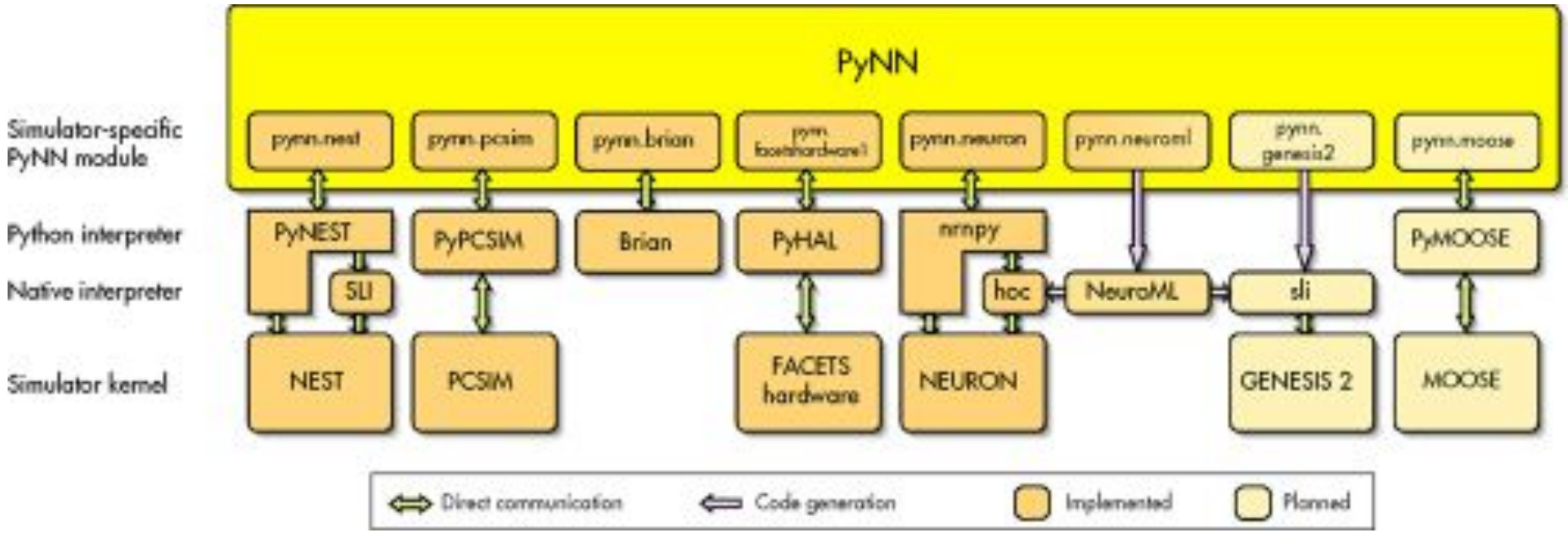


Implement Custom Digital Circuits : A Success Story for Configurable Hardware




Biological Databases
Model Building
Cells, Network, Plasticity
Theory
Formal Network Description
The Neural FPGA Concept
for rapid Prototyping
CONFIGURABILITY !
Technology Mapping
Technology Routing
Hardware Platform Constraints
Simulation and Verfification
Neuromorphic Substrate

PyNN : A Platform independent Description Language for Neural Circuits



Integrating and widely accepted toolset for a generic (i.e. simulator independent) access to neural simulation and hardware

Bringing new computer architecturs (i.e. neuromorphic systems) to the non-expert neuroscience user

Typical Configuration Space for a Neuromorphic System approx. 40 MB for a full Wafer

Scope	Name	Type	Description
Neuron circuits (A)	n/a	i_n	Two digital configuration bits activating the neuron and readout of its membrane voltage
	g_l	i_n	Bias current for neuron leakage circuit
	τ_{refrac}	i_n	Bias current controlling neuron refractory time
	E_l	s_n	Leakage reversal potential
	E_{inh}	s_n	Inhibitory reversal potential
	E_{exc}	s_n	Excitatory reversal potential
	V_{th}	s_n	Firing threshold voltage
	V_{reset}	s_n	Reset potential
Synapse line drivers (B)	n/a	i_l	Two digital configuration bits selecting input of line driver
	n/a	i_l	Two digital configuration bits setting line excitatory or inhibitory
	t_{rise}, t_{fall}	i_l	Two bias currents for rising and falling slew rate of presynaptic voltage ramp
	g_i^{max}	i_l	Bias current controlling maximum voltage of presynaptic voltage ramp
Synapses (B)	w	i_s	4-bit weight of each individual synapse
STP related (C)	n/a	i_l	Two digital configuration bits selecting short-term depression or facilitation
	U_{SE}	i_l	Two digital configuration bits tuning synaptic efficacy for STP
	n/a	s_l	Bias voltage controlling spike driver pulse length
	τ_{rec}, τ_{facil}	s_l	Voltage controlling STP time constant
	I	s_l	Short-term facilitation reference voltage
	R	s_l	Short-term capacitor high potential
STDP related (D)	n/a	i_l	Bias current controlling delay for presynaptic correlation pulse (for calibration purposes)
	$A_{+/-}$	s_l	Two voltages dimensioning charge accumulation per (anti-)causal correlation measurement
	n/a	s_l	Two threshold voltages for detection of relevant (anti-)causal correlation
	τ_{STDP}	g	Voltage controlling STDP time constants

Typical Configuration Space for a Neuromorphic System approx. 40 MB for a full Wafer

Scope	Name	Type	Description
Neuron circuits (A)			...membrane voltage
Synapse line drivers (B)			...amp
Synapses (B)			
STP related (C)			...on
STDP related (D)	n/a	i_l	Bias current controlling delay for presynaptic correlation pulse (for calibration purposes)
	$A_{+/-}$	s_l	Two voltages dimensioning charge accumulation per (anti-)causal correlation measurement
	n/a	s_l	Two threshold voltages for detection of relevant (anti-)causal correlation
	τ_{STDP}	g	Voltage controlling STDP time constants

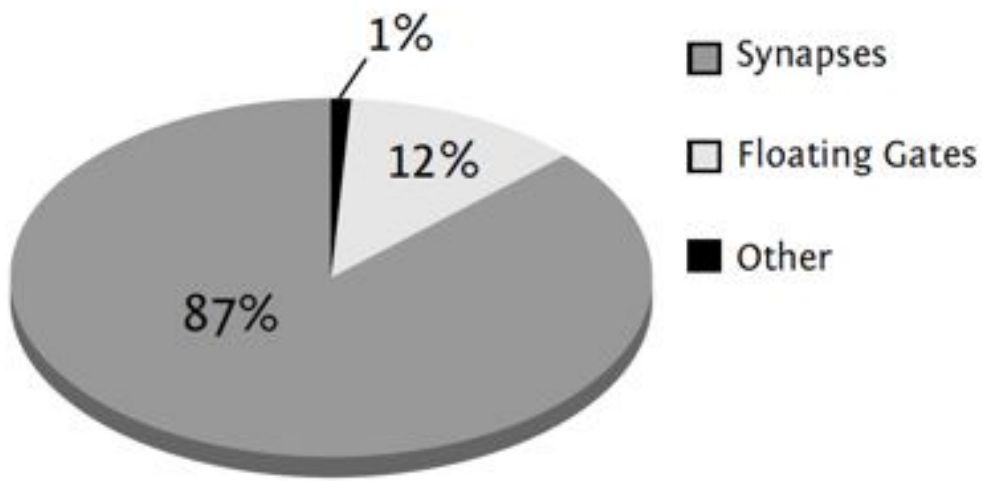
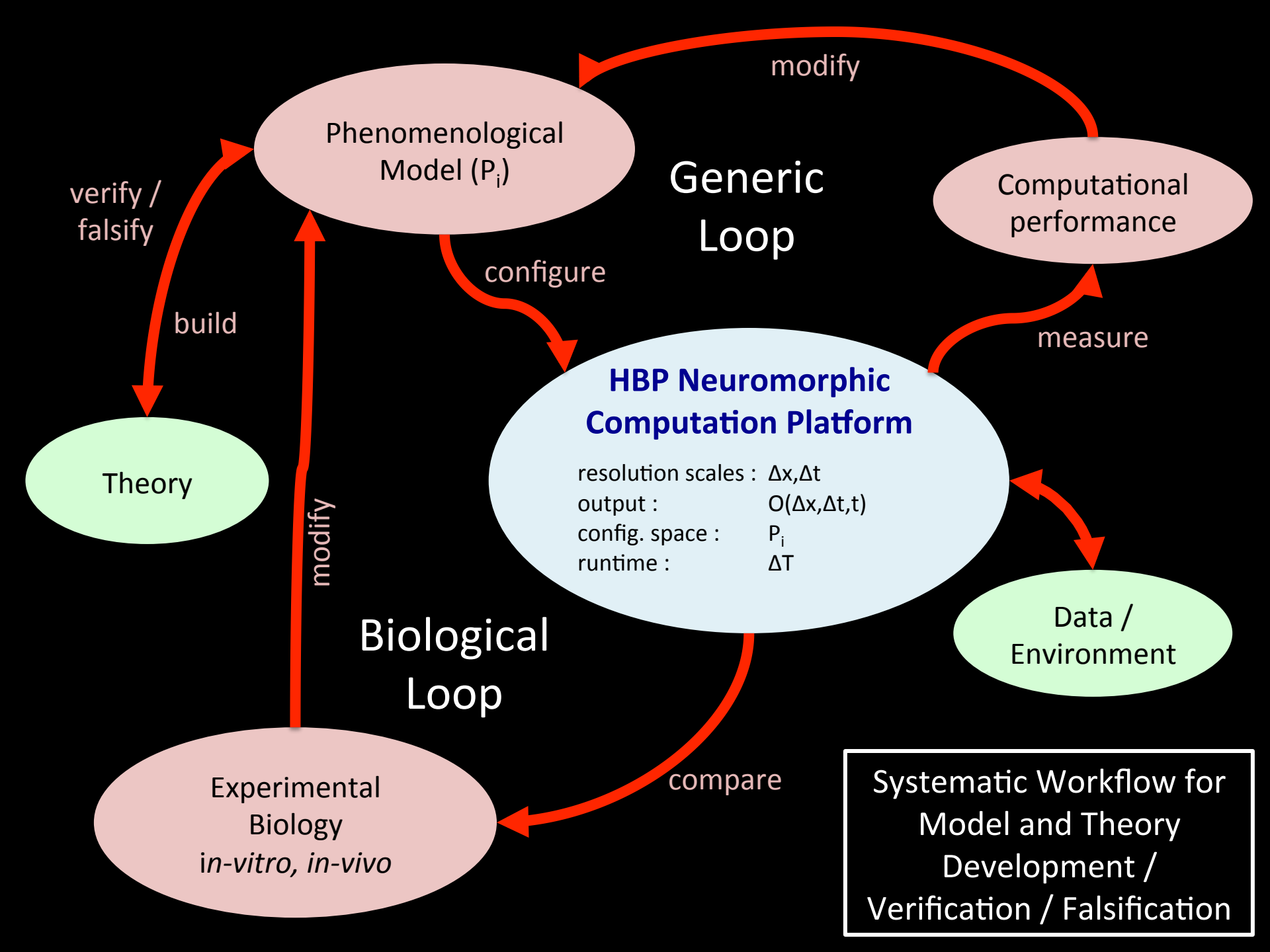


Fig. 4: Sector diagram of the parameter space to configure one HICANN chip. For a full wafer, the configuration data volume is 44 MB large.


modify
Phenomenological Model (P_i)
Generic Loop
Computational performance
verify / falsify
configure
build
measure
HBP Neuromorphic Computation Platform
resolution scales : Δx,Δt
output : O(Δx,Δt,t)
config. space : P_i
runtime : ΔT
Theory
modify
Data / Environment
Biological Loop
compare
Experimental Biology
in-vitro, in-vivo
Systematic Workflow for Model and Theory Development / Verification / Falsification

INTERNATIONAL TECHNOLOGY ROADMAP FOR SEMICONDUCTORS 2011 EDITION : EMERGING RESEARCH DEVICES, Chapter 5.3

The appeal of neuromorphic architectures lies in

i) their potential to achieve (human-like) intelligence based on unreliable devices typically found in neuronal tissue
ii) their strategies to deal with anomalies, emphasizing not only tolerance to noise and faults, but also the active exploitation of noise to increase the effectiveness of operations
iii) their potential for low-power operation.

Traditional von Neumann machines are less suitable with regard to item i), since for this type of tasks they require a machine complexity (the number of gates and computational power), that tends to increase exponentially with the complexity of the environment (the size of the input). Neuromorphic systems, on the other hand, exhibit a more gradual increase of their machine complexity with respect to the environmental complexity.

Therefore, at the level of human-like computing tasks, neuromorphic machines have the potential to be superior to von Neumann machines.

Experiments and Applications under Study In *BrainScaleS* and *HBP* - 4 Categories

I. Fundamental dynamical properties of isolated circuits
Synchronisation, coincidence detection, stability, order-chaos

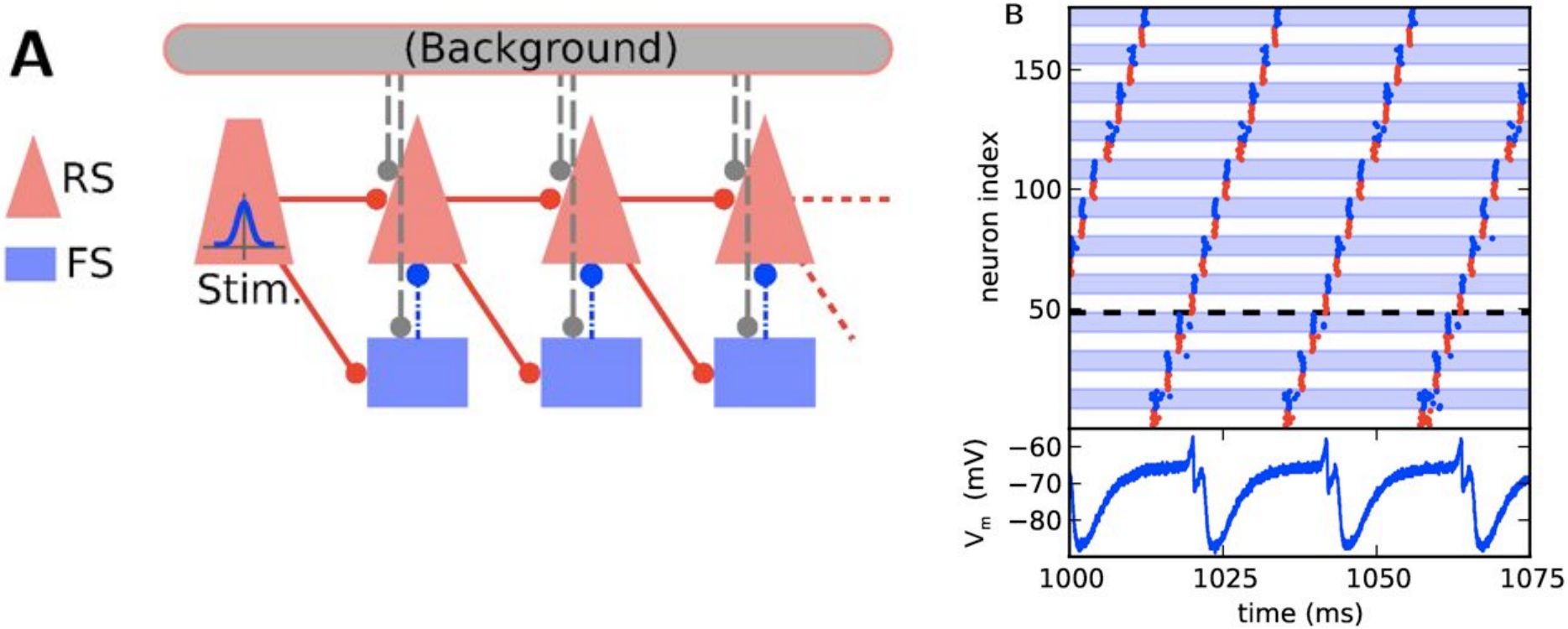
II. Implement and test fundamental, generic concepts and theories
Liquid computing, probabilistic inference, neural sampling

III. Biologically realistic, reverse engineered circuits in closed loops
Cortical structures, cortical columns, functional units

IV. Generic neuromorphic computing outside neuroscience
Neuromorphic controllers, spatio-temporal pattern detection in data streams, causal relations in big data, approximate computing

All 6 experiments on the following slides run on the SAME re-configurable physical substrate (chip based system)

Synfire Chain with Feed-Forward Inhibition, Kremkov et al. 2010

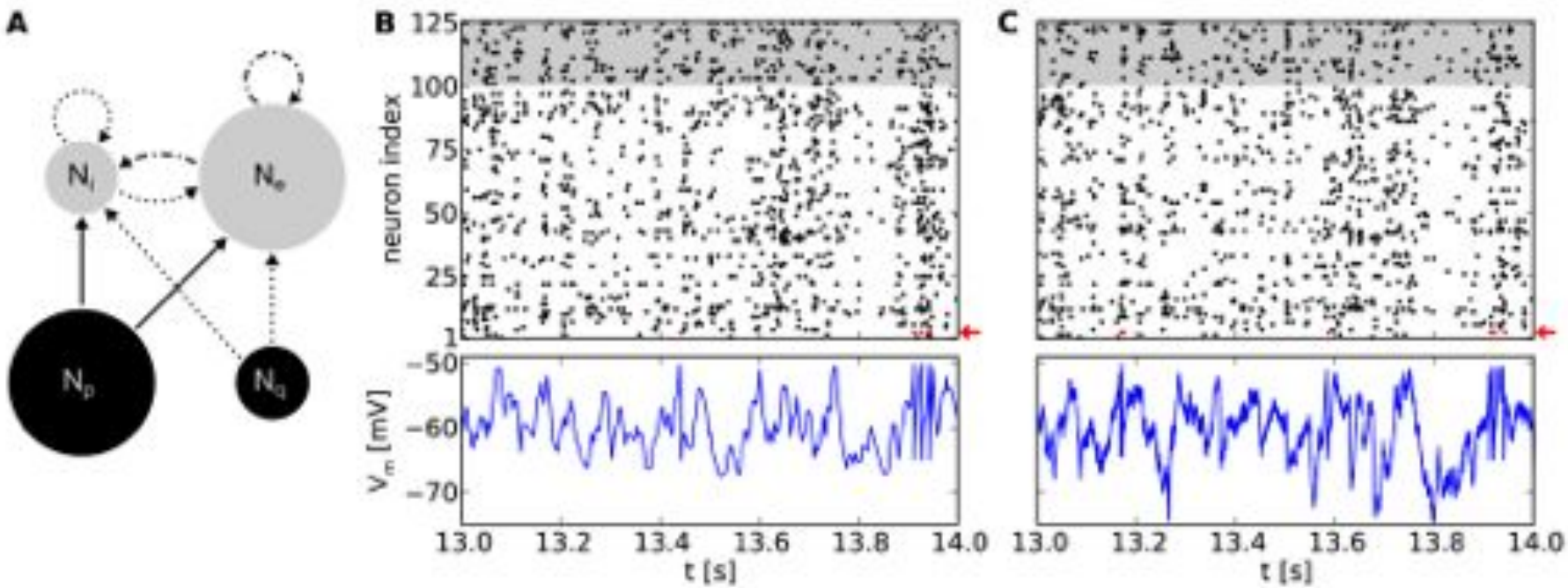


Pfeil T, Grübl A, Jeltsch S, Müller E, Müller P, Petrovici MA, Schmuker M, Brüderle D, Schemmel J, Meier K (2013). Six networks on a universal neuromorphic computing substrate. *Frontiers in Neuromorphic Engineering* 7:11

Balanced random network (Brunel 2000)
asynchronous, irregular firing
Prerequisite for information processing using stochastic inference

Hardware
10 µs

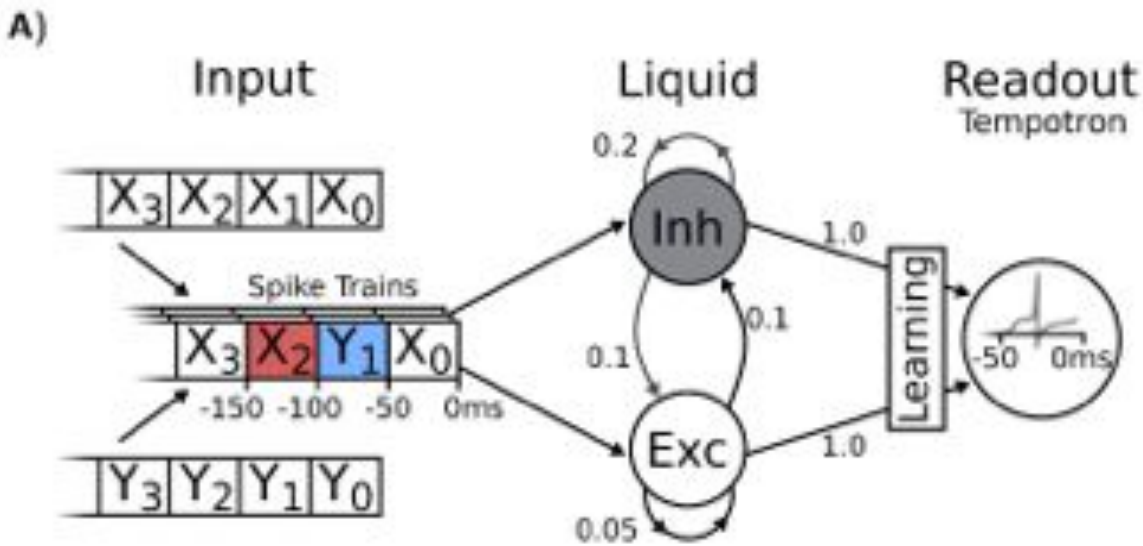
Software
1s

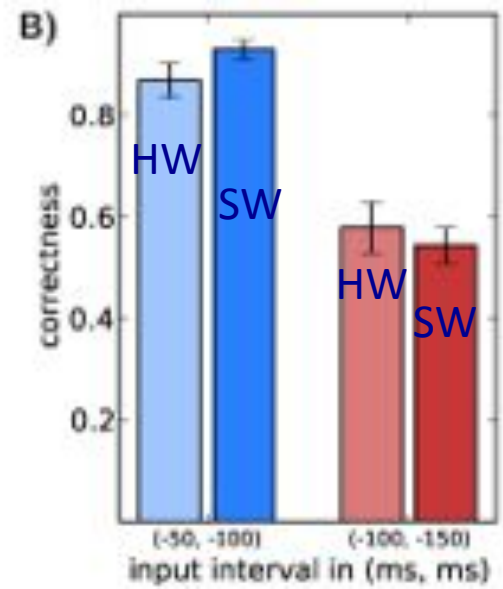


Pfeil T, Grübl A, Jeltsch S, Müller E, Müller P, Petrovici MA, Schmuker M, Brüderle D, Schemmel J, Meier K (2013). Six networks on a universal neuromorphic computing substrate. *Frontiers in Neuromorphic Engineering* 7:11

Liquid State Machine (Maass et al., 2002)
with binary tempotron classifier readout
(Gütig&Sompolinsky, 2006)

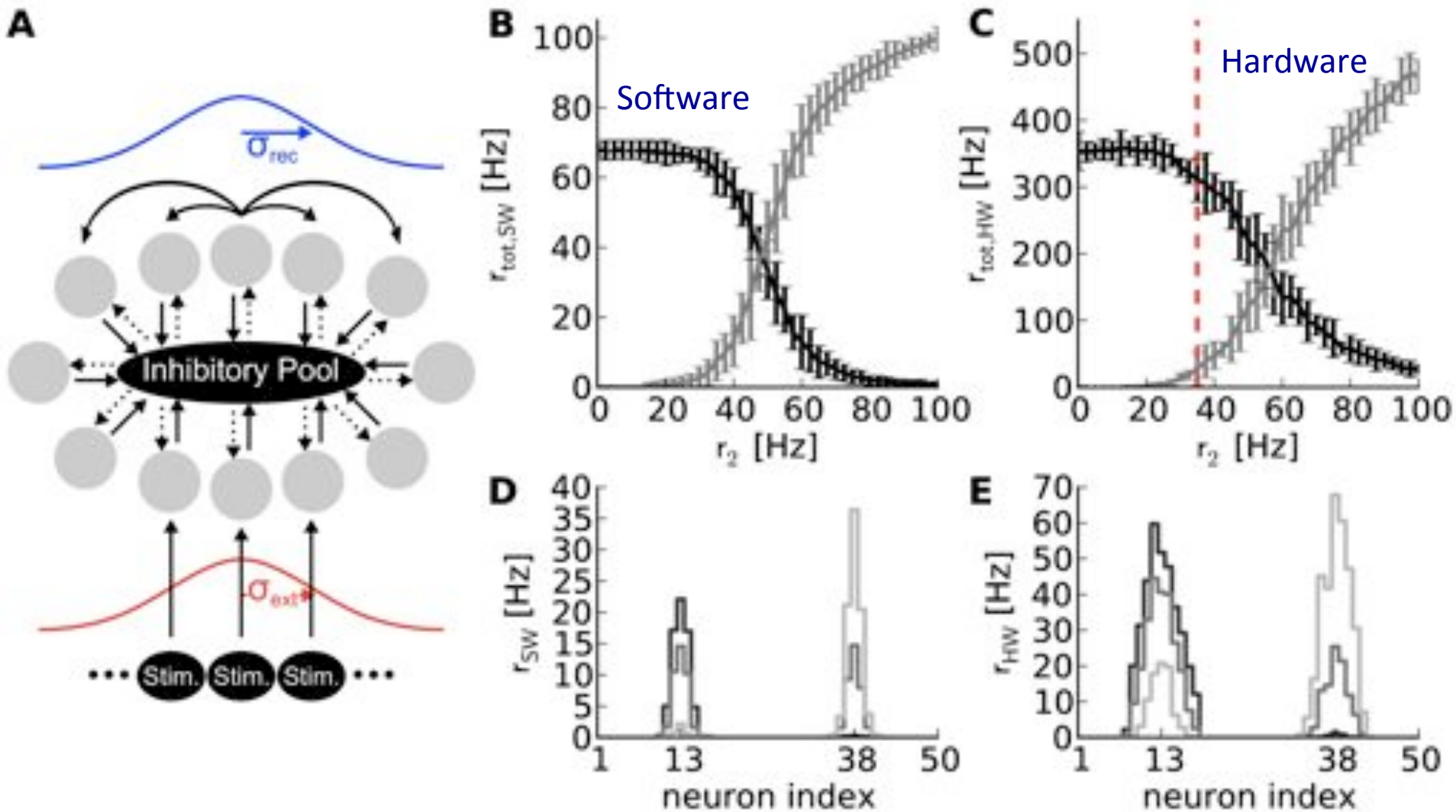
Distinguish spike train segments in a continuous data stream composed of two templates (X, Y) with identical rates





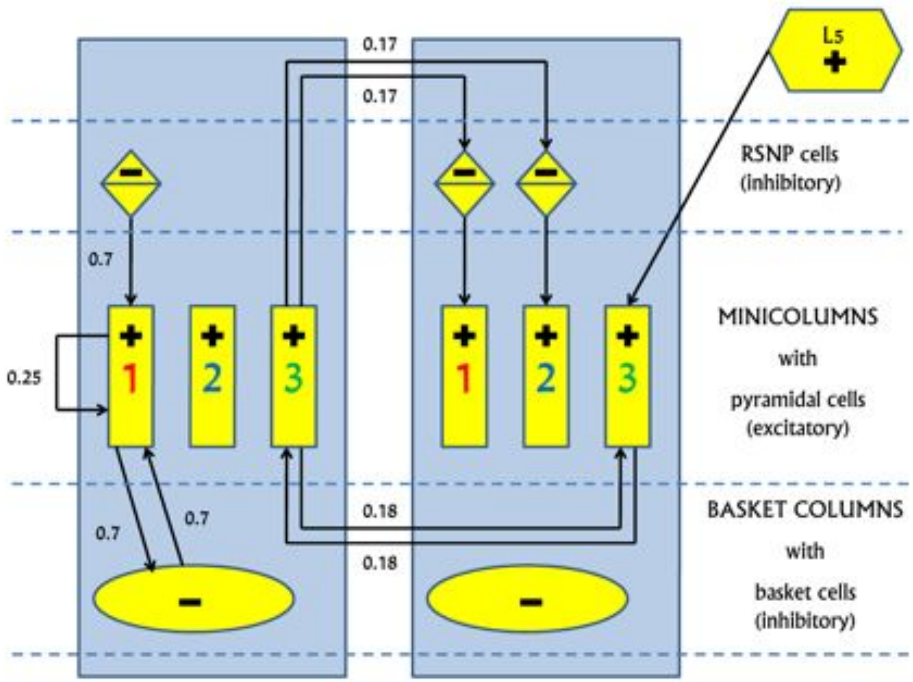
Pfeil T, Grübl A, Jeltsch S, Müller E, Müller P, Petrovici MA, Schmuker M, Brüderle D, Schemmel J, Meier K (2013). Six networks on a universal neuromorphic computing substrate. *Frontiers in Neuromorphic Engineering* 7:11

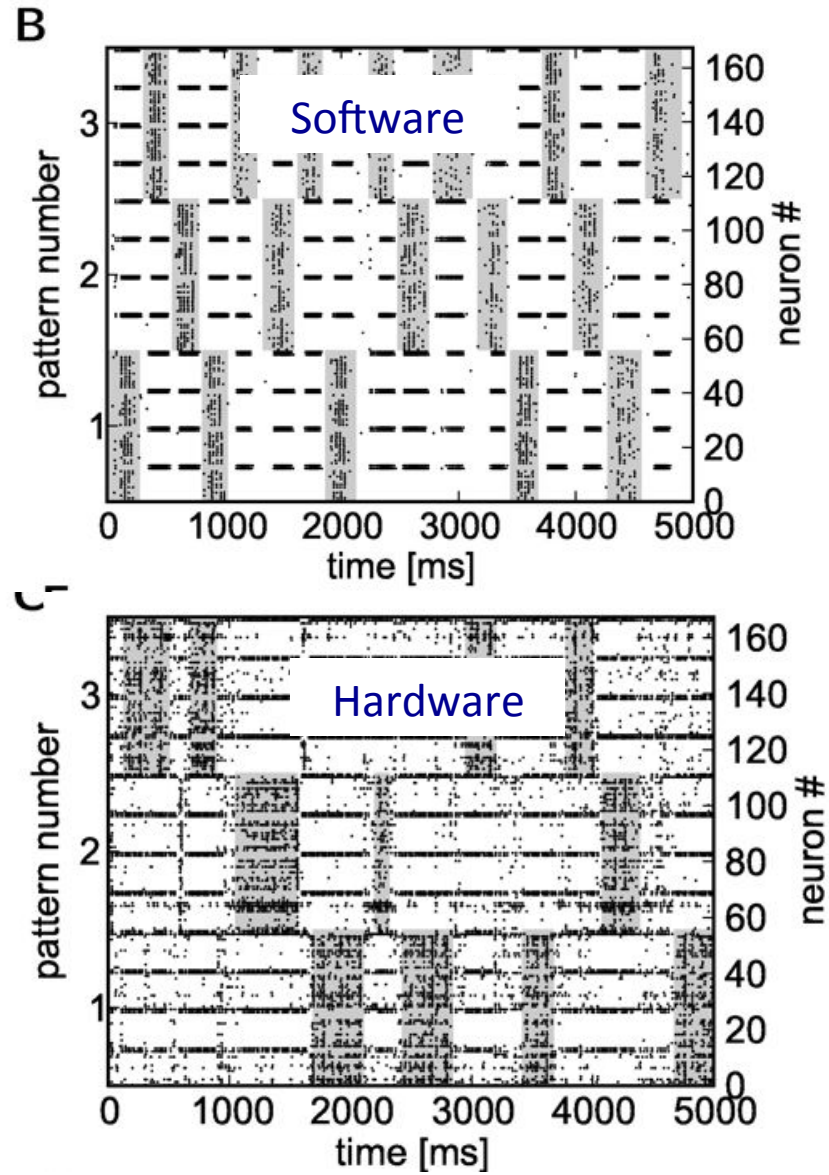
Winner-Take-All Circuit (Neftci et al., 2011)



Pfeil, Thomas, 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."
arXiv preprint arXiv:1210.7083 (2012), accepted by Frontiers in Neuroscience

Cortical Layer 2/3 attractor memory network Lundqvist et al. 2006, 2010

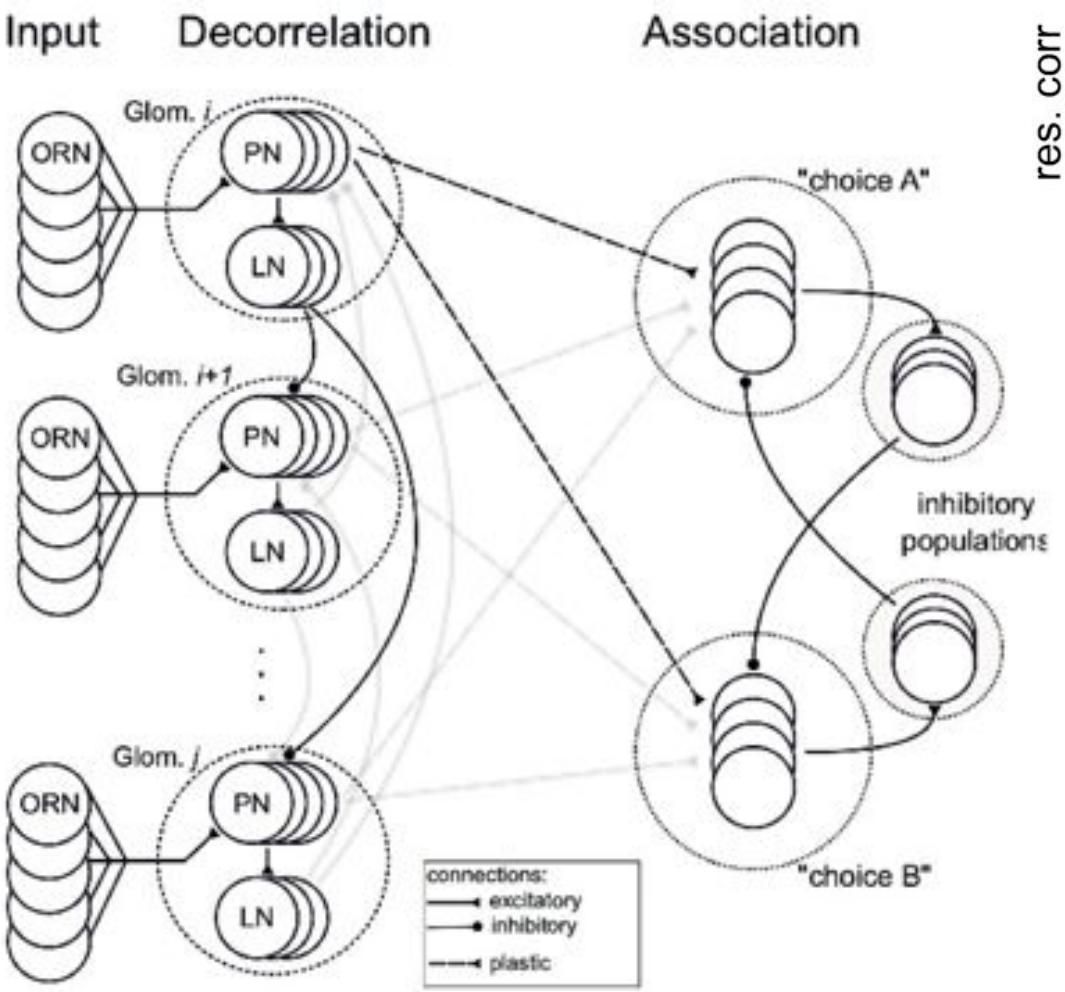


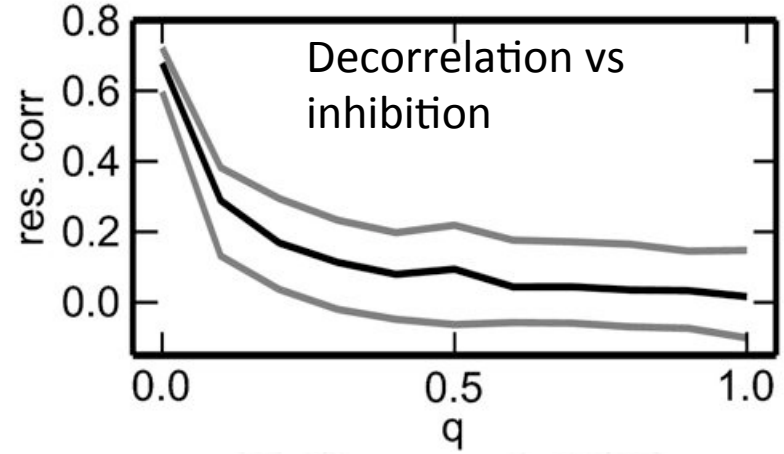


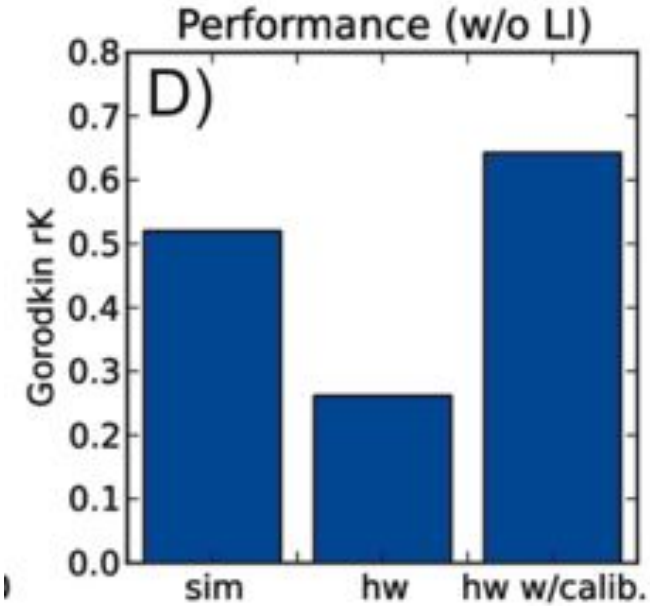
Pfeil T, Grübl A, Jeltsch S, Müller E, Müller P, Petrovici MA, Schmuker M, Brüderle D, Schemmel J, Meier K (2013). Six networks on a universal neuromorphic computing substrate. *Frontiers in Neuromorphic Engineering* 7:11

Insect antennal lobe inspired classification of multidimensional (odor) data (Schmuker&Schneider 2007)

Decorrelating inputs from sensory neurons - asscociation to choices







Pfeil T, Grübl A, Jeltsch S, Müller E, Müller P, Petrovici MA, Schmuker M, Brüderle D, Schemmel J, Meier K (2013). Six networks on a universal neuromorphic computing substrate. *Frontiers in Neuromorphic Engineering* 7:11

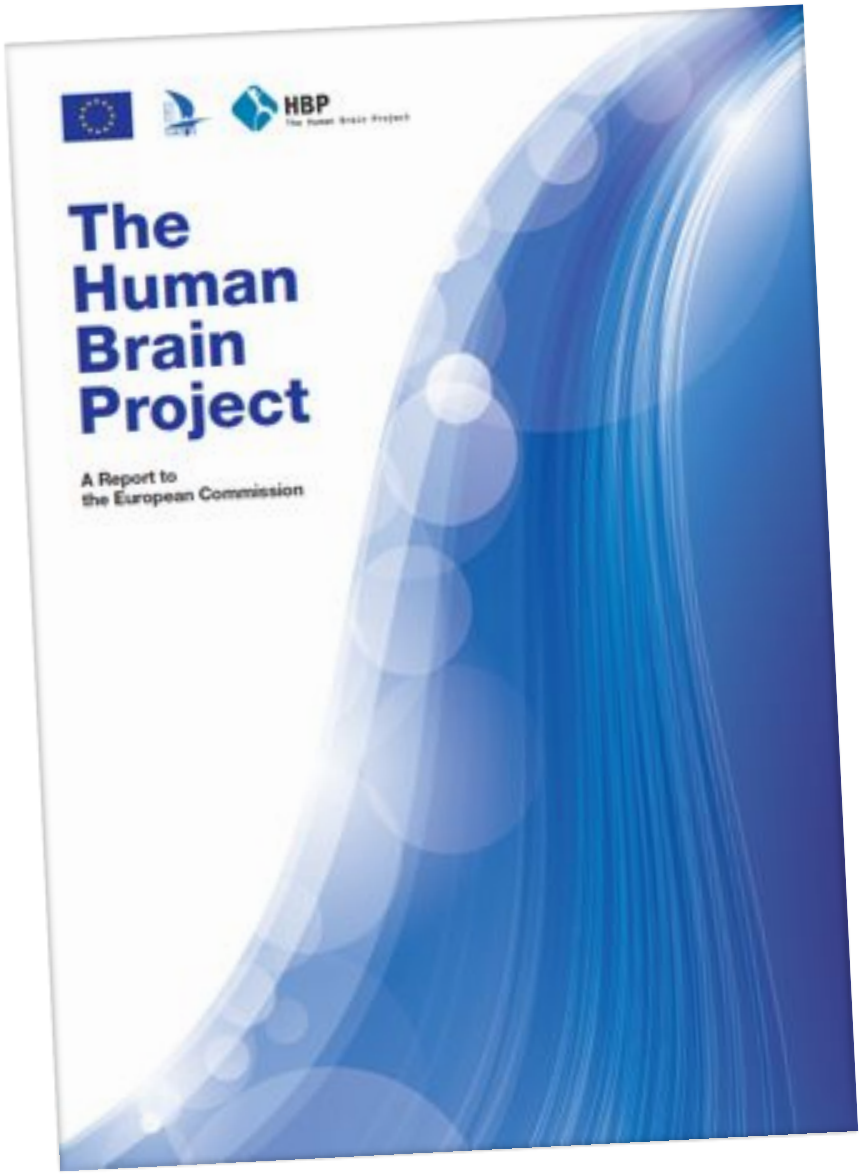
BRAIN-DERIVED COMPUTING

- Consistent concept for a novel, brain derived, non-von Neumann, non-Turing computer architecture
- Accessible to available technologies (CMOS) and attractive application for future component technologies (nanoelectronics)
- Key features : Universality, scalability, fault tolerance, power efficiency, speed, learning
- Accelerated operation : Only known approach to bridge all timescales relevant for circuit dynamics
- Important next step : Give up simulation as a reference, exploit device mismatch and noise

THANKS !

brainscales.eu

HBP
The Human Brain Project



A large-scale coordinated effort over 10 years
Budget 1.1 B€

Director : Henry Markram (EPFL)
Co-Directors : Richard Frackowiak (CHUV)
Karlheinz Meier (heidelbrg)

Public Report available :

www.humanbrainproject.org

Approval of 30 months ramp-up phase received in January 2013

Expected start : October 2013

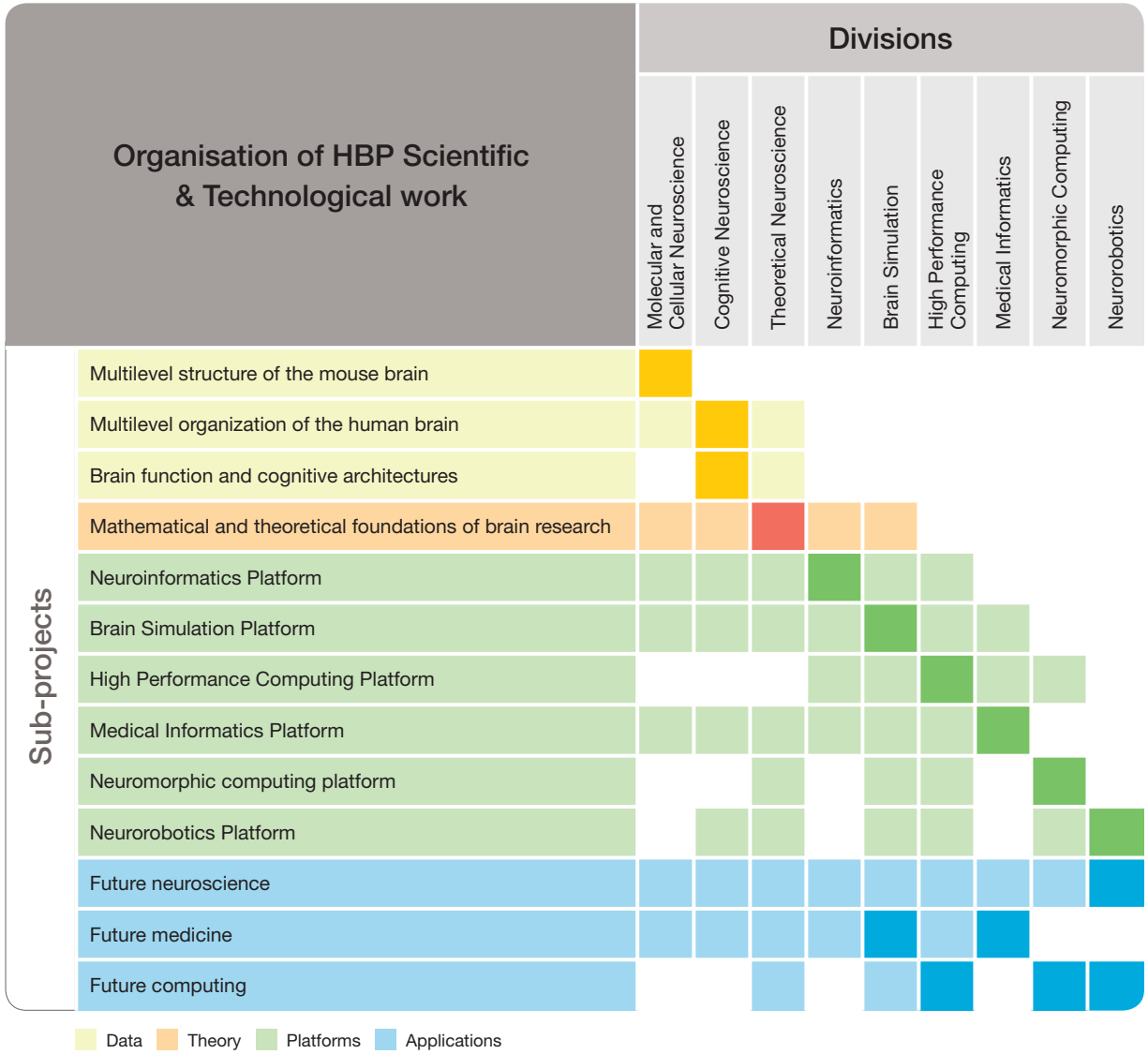


Figure 34: Organisation of work in the HBP. The nine divisions contribute to data generation, platform building and platform use. Dark shaded areas represent responsibilities for work packages; light shaded areas show contributions to the work programme.

HBP
The Human Brain Project

Future Computing Platforms

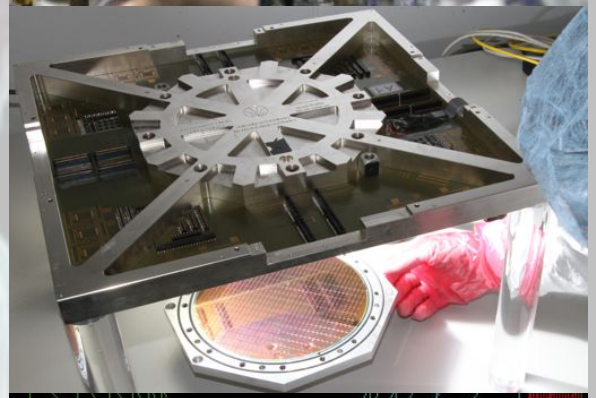
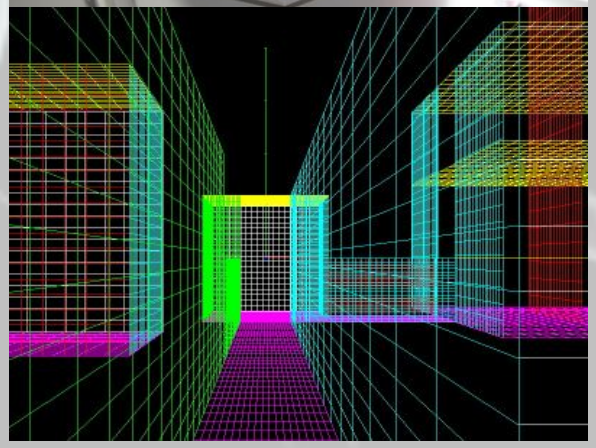
what will they provide ? – a few selected items

High Performance Computing

- Interactive, visual. Exascale supercomputing
- Massive distributed volumes of heterogeneous data
- Convergence with neuromorphic technology

Neuromorphic Computing

- First large-scale neuromorphic systems superior to HPC
- Non-von Neumann (Multicore) + non-Turing (Neuromorph)
- Technology integration (3D, non-CMOS backends)

Neurorobotics

- Virtual robots with two-way, closed loop interfaces
- Link to brain models and neuromorphic systems
- Physical prototypes and applications