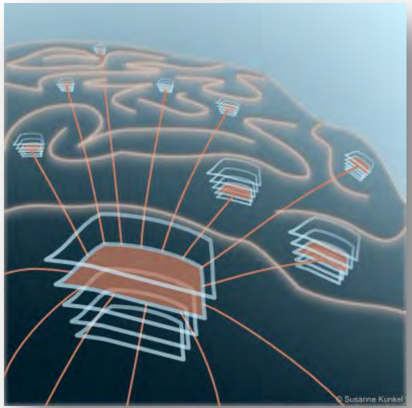


HBP The Human Brain Project

JÜLICH FORSCHUNGSZENTRUM

JARA|HPC JARA|BRAIN

NEST, simulation technology for brain-scale networks at cellular and synaptic resolution

Moritz Helias, Susanne Kunkel, Abigail Morrison, Markus Diesmann

March 18th 2014, SOS18, St. Moritz

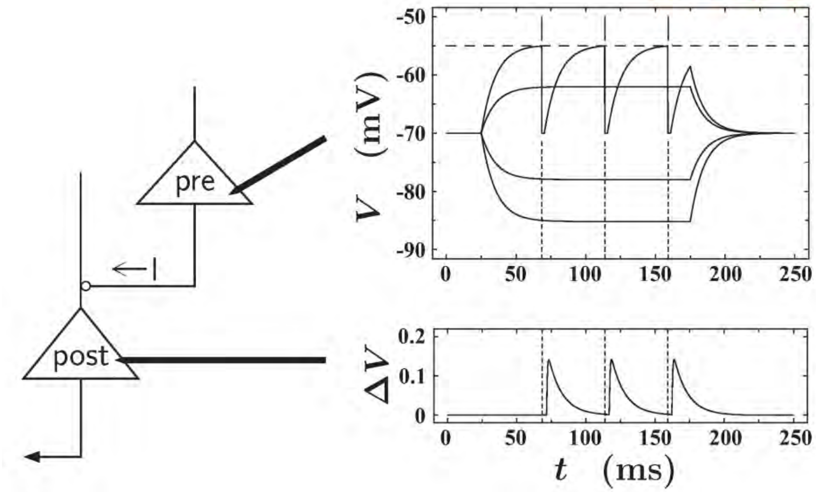
Institute of Neuroscience and Medicine (INM-6)
Institute for Advanced Simulation (IAS-6)
Jülich Research Centre

Outline

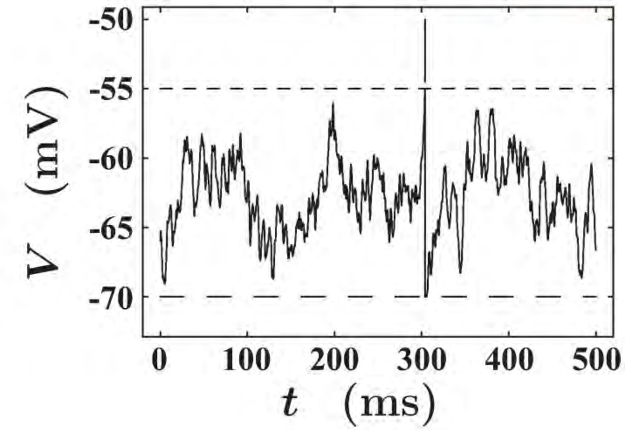
- fundamental neuronal interaction
- example of model construction
  - local cortical network
  - multi-area model
- necessity of brain-scale models
- 4^{th} generation simulation kernel of NEST

Fundamental interaction

in vitro



in vivo



- current injection into pre-synaptic neuron causes excursions of membrane potential
- supra-threshold value causes spike transmitted to post-synaptic neuron
- post-synaptic neuron responds with small excursion of potential after delay
- inhibitory neurons (20%) cause negative excursion

- each neuron receives input from 10,000 other neurons
- causing large fluctuations of membrane potential
- emission rate of 1 to 10 spikes per second

Single neuron dynamics (example)

leaky integrator equation is a *linear* differential equation

$$\frac{dV}{dt} = -\frac{1}{\tau_{\mathrm{m}}} V + \frac{1}{C} I$$

individual synaptic events t_{syn} superimpose *linearly*

$$I(t) = \sum_i w_i \, \iota(t - t^i_{\mathrm{syn}}) \qquad \iota(t) = \frac{t}{\tau_\alpha} e^{1-\frac{t}{\tau_\alpha}}, \quad t \geq 0$$

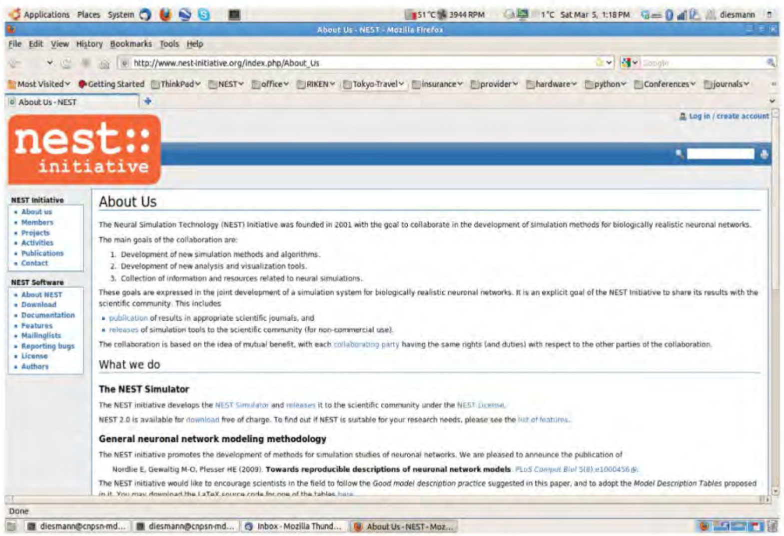
idea: rewrite *total* current *I* as differential equation

$$\frac{dI}{dt} = -\frac{1}{\tau_\alpha} I + y_1$$

$$\frac{dy_1}{dt} = -\frac{1}{\tau_\alpha} y_1 + \widehat{y}_1 \sum_i w_i \, \delta(t - t^i_{\mathrm{syn}})$$

synaptic input causes jump by $\widehat{y}_1 w_i$ in state variable y_1 at t^i_{syn}

Collaboration: the NEST Initiative



Major goals:

systematically publish new simulation technology

produce public releases under GPL

- initiated by Markus Diesmann and Marc-Oliver Gewaltig in (1994)
- collaboration of several labs (since 2001)
- registered society (since 2012)
- teaching in international advanced courses
  - Okinawa Computational Neuroscience Course OCNC
  - Advanced Course in Computational Neuroscience ACCN, Europe
- core simulation technology in EU Human Brain Project (HBP)

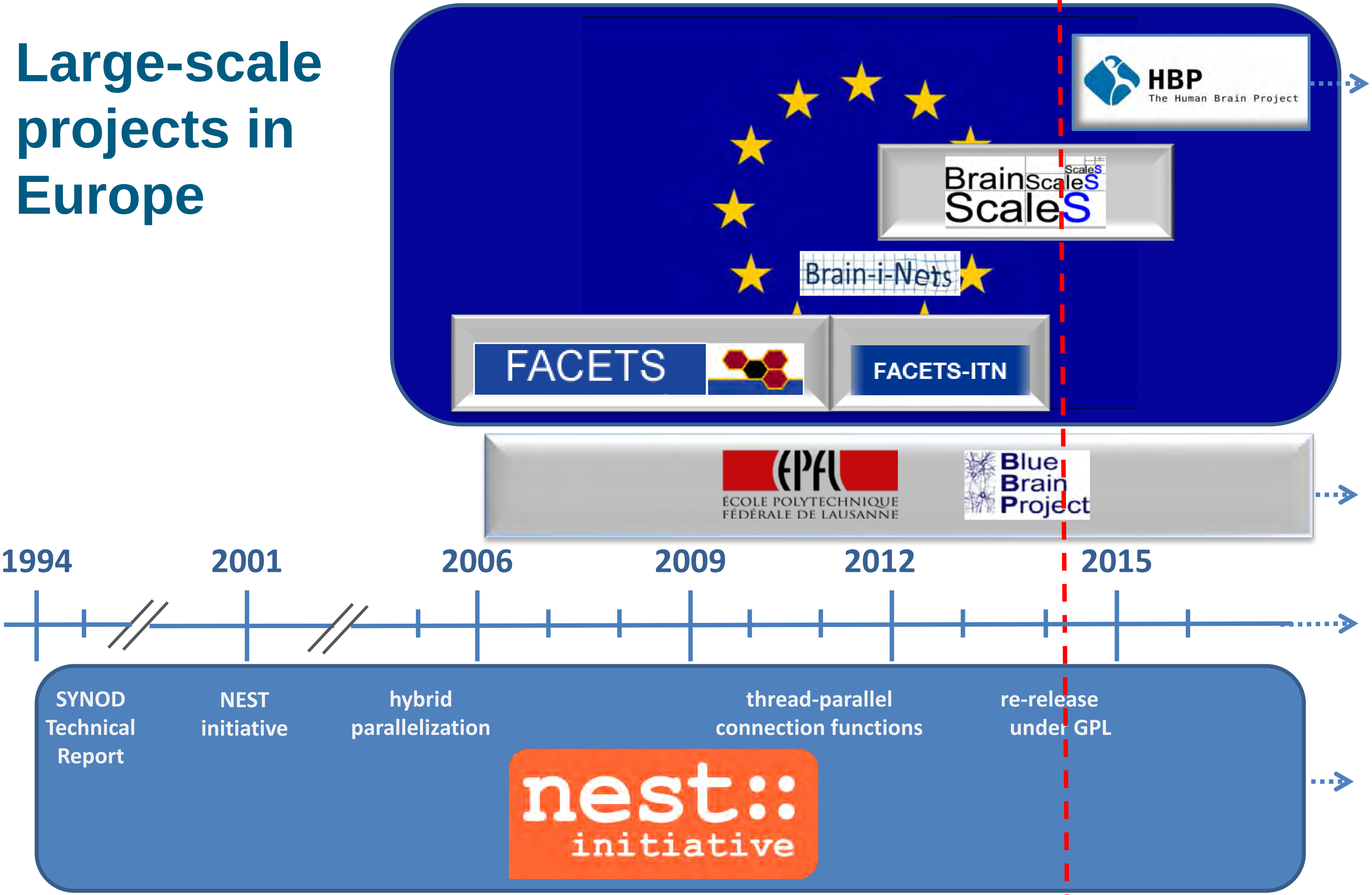
HBP The Human Brain Project

nest:: initiative

RWTH AACHEN

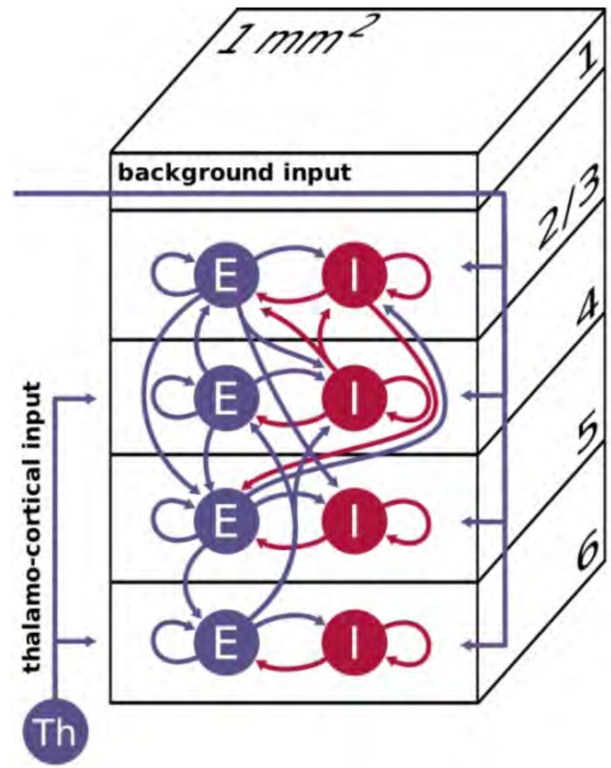
JÜLICH FORSCHUNGSZENTRUM

Large-scale projects in Europe



Minimal layered cortical network model

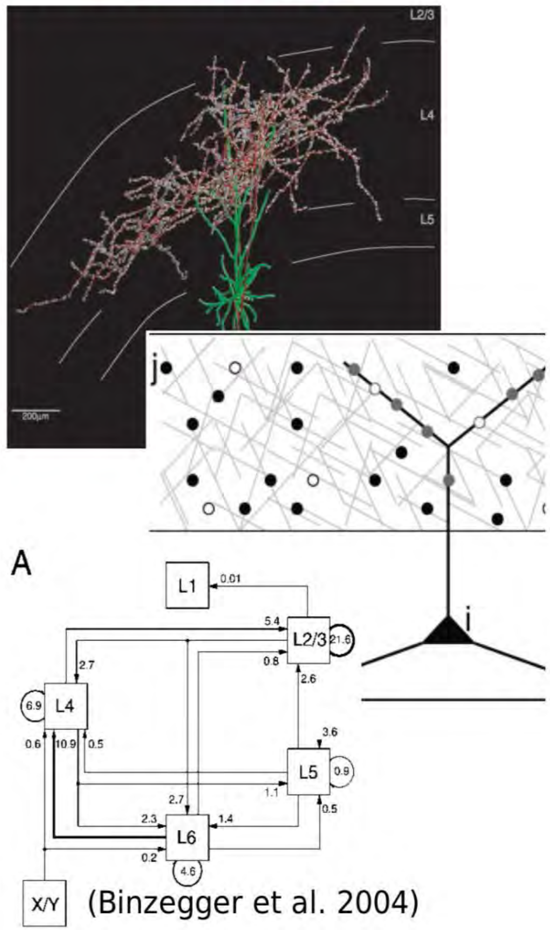
- volume: 1 mm^3
- 100,000 neurons, 1 billion synapses
- 2 populations of neurons (E,I) per layer
- E and I identical neuronal dynamics
- laterally homogeneous connectivity
- layer- and type-specific connectivity



Potjans TC & Diesmann M (2012) The cell-type specific connectivity of the local cortical network explains prominent features of neuronal activity. *Cerebral Cortex* 10.1093/cercor/bhs358

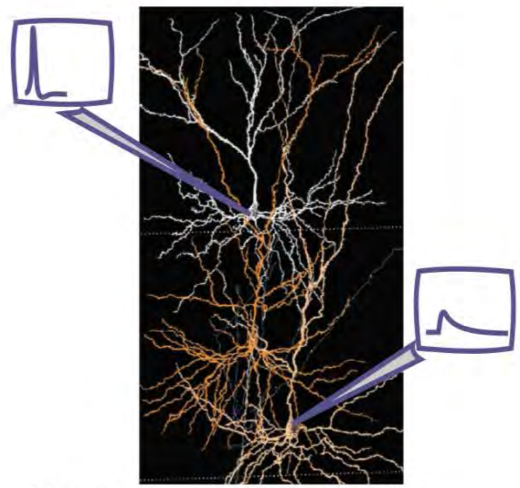
Anatomical data sets

in vivo anatomy



(Binzegger et al. 2004)

in vitro physiology



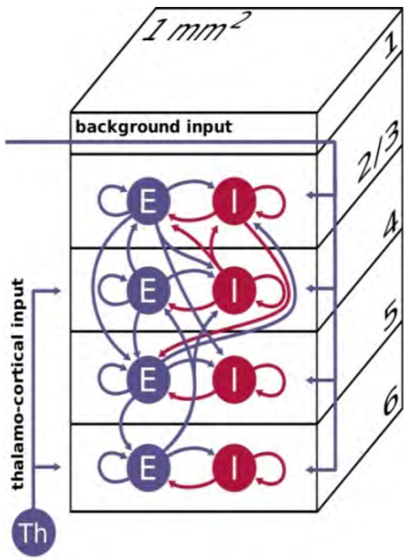
Type of connection[a]	Connectivity ratio
L5 pyramid to L5 pyramid	1:11 (15:163)
L2/3 pyramid to L2/3 pyramid	1:4 (65:247)
	1:10 (8:81)
L4 excitatory to L4 excitatory	**1:5.7 (4:23)**
L3 pyramid to L5 pyramid	1:1.8 (16:29)
[Postsynaptic apical dendrite]	
	1:1 (2:2)
L5 pyramid to L3 pyramid	1:29
L4 excitatory to L3 pyramid	1:3.6 (7:25)
(Presynaptic spiny stellates) ($n = 4$)	**1:10 (7:70)**
L5 pyramid to L5 interneuron	1:10.4 (7:73)
L5 interneuron to L5 pyramid	*1:8 (9:73)*

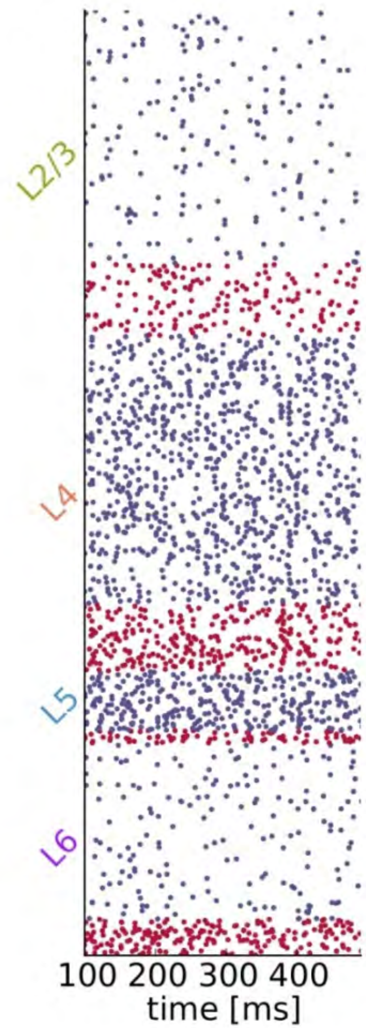
(Thomson et al. 2002)

Activity of local cortical microcircuit

taking into account layer and neuron-type specific connectivity is sufficient to reproduce experimentally observed:

- asynchronous-irregular spiking of neurons
- higher spike rate of inhibitory neurons
- correct distribution of spike rates across layers

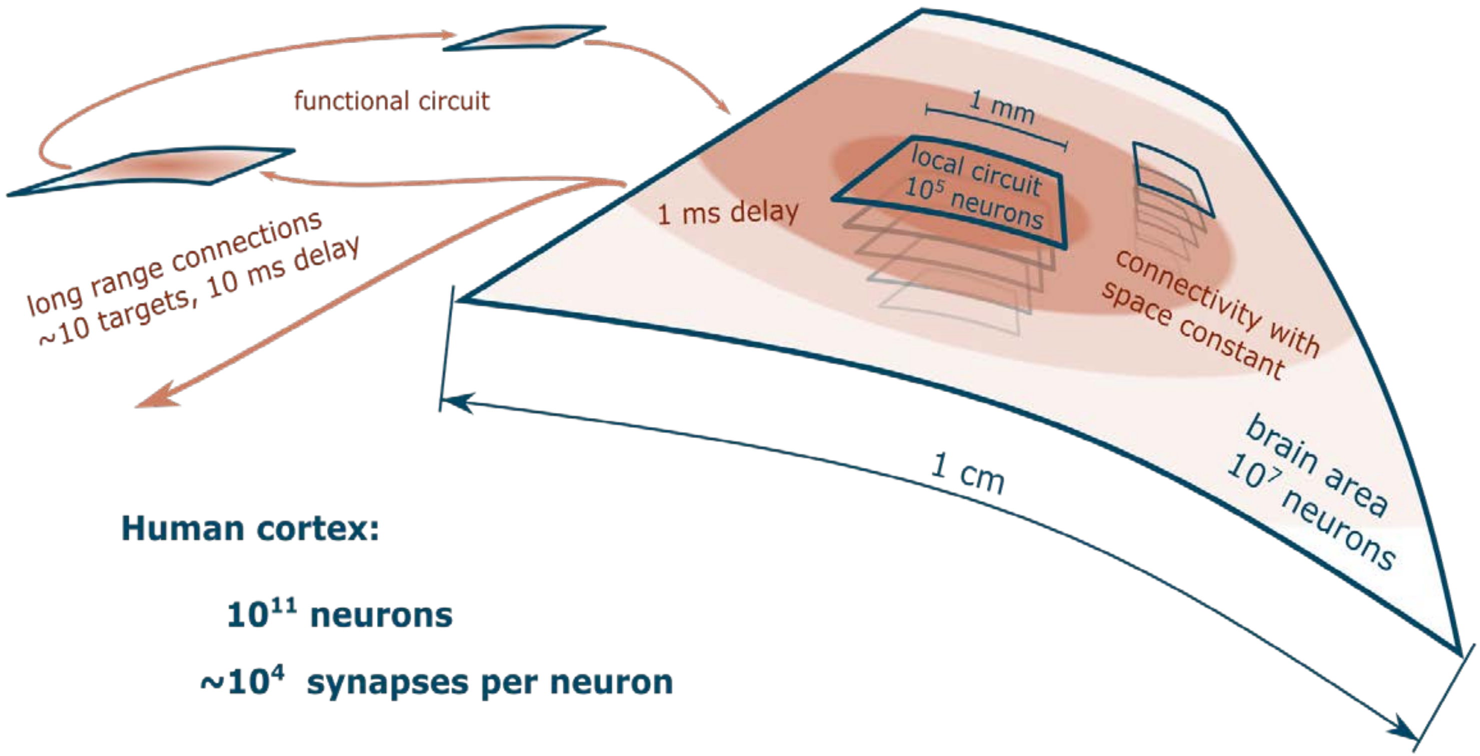




Potjans TC & Diesmann M (2012) The cell-type specific connectivity of the local cortical network explains prominent features of neuronal activity. *Cerebral Cortex* 10.1093/cercor/bhs358

Architecture of human cortex

a network of networks



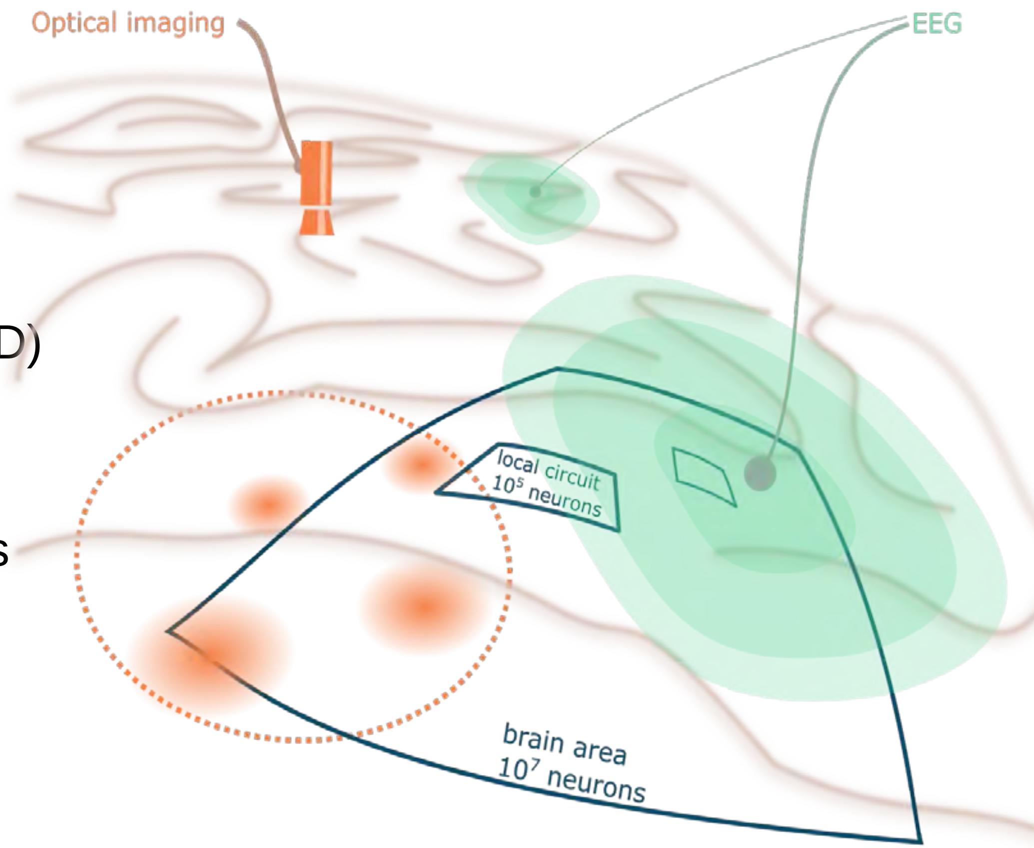
- connectivity of local microcircuit
- long-range connections between areas

Meso- and macro-scale measures

brain-scale networks
provide substrate for:

mesoscopic measures
- local field potential (LFP)
- voltage sensitive dyes (VSD)

and macroscopic measures
- EEG, MEG
- fMRI resting state networks

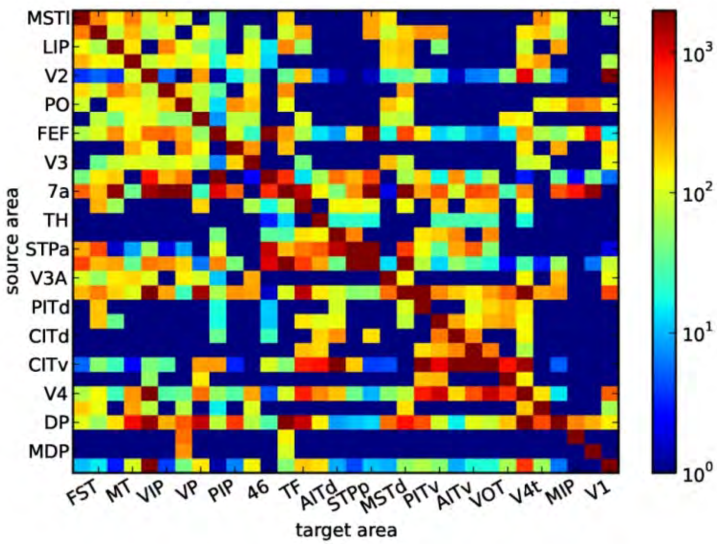


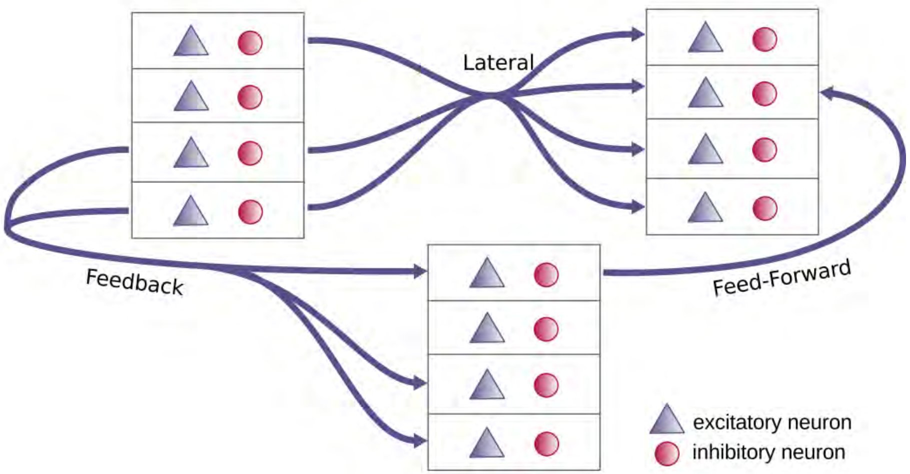
connecting microscopic models to imaging data

S Kunkel

in collaboration with Gaute Einevoll (Aas, Norway)

Towards self-consistent models

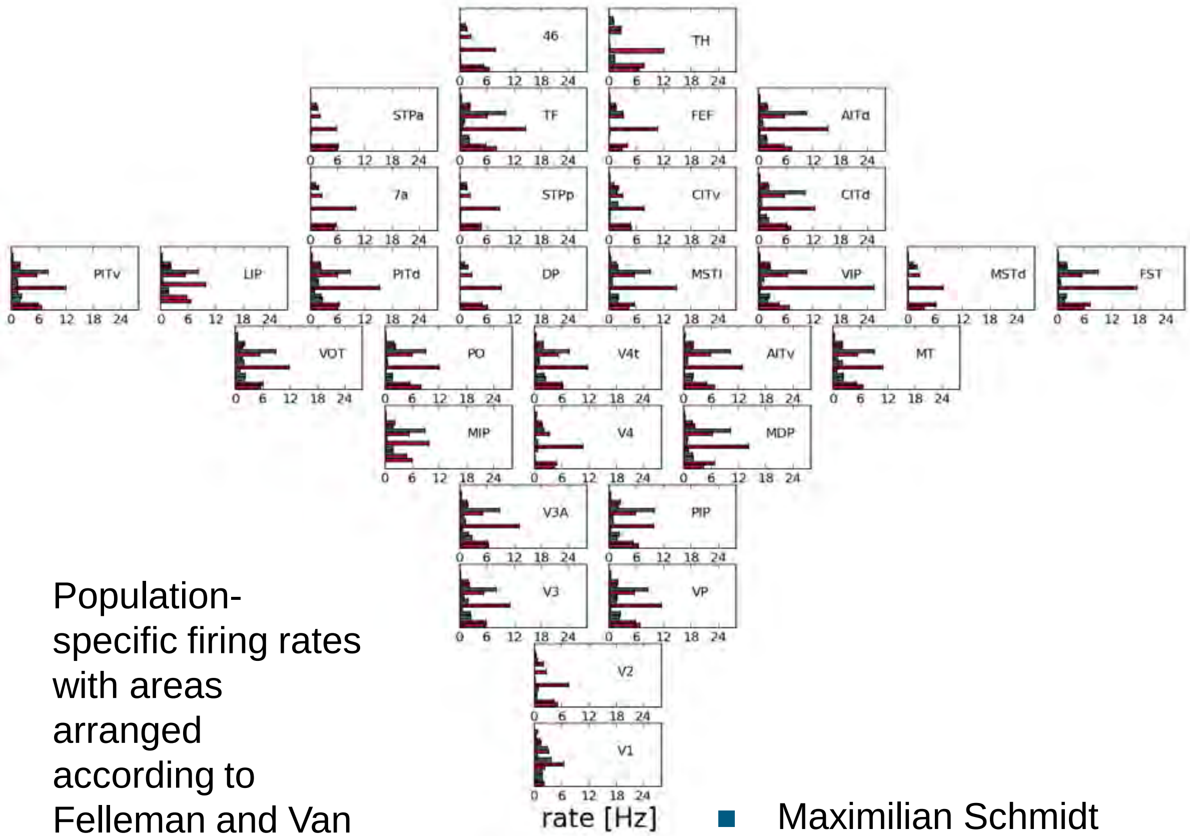


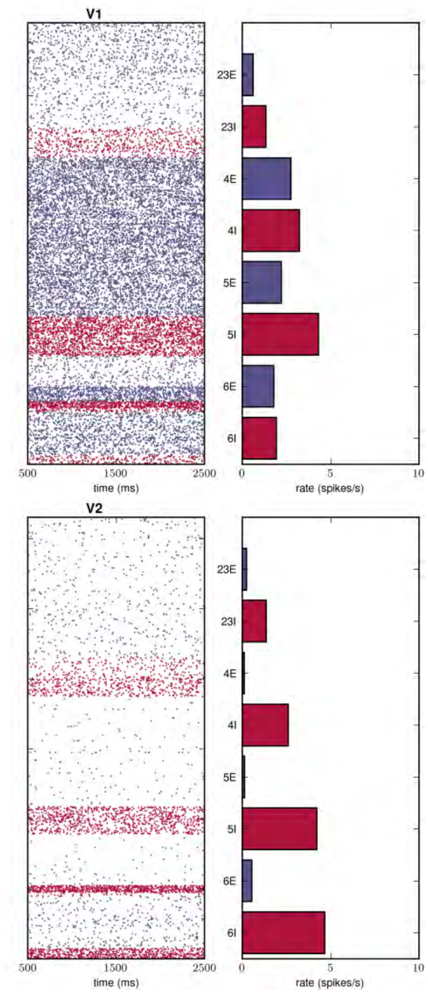


- matrix shows number of synapses between areas
- macroscopic connectivity data for macaque visual cortex most complete (CoCoMac database)
- again, model enables data integration

HBP The Human Brain Project — nest::initiative — RWTHAACHEN — JÜLICH FORSCHUNGSZENTRUM

Multi-area model of macaque visual cortex



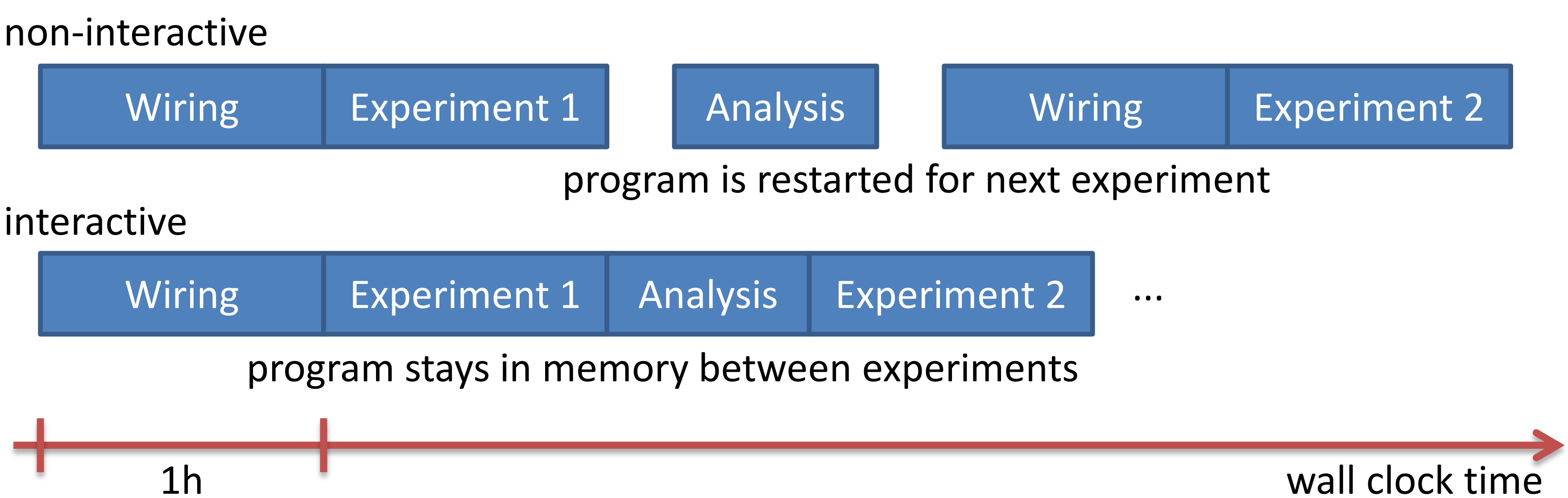


Population-specific firing rates with areas arranged according to Felleman and Van Essen hierarchy

- Maximilian Schmidt
- Sacha van Albada
- Rembrandt Bakker

Use case – Interactive experiment design

- building large networks from rules takes a lot of time (~1h)
- consecutive experiments often require the data of the last
- re-using the network structure for multiple experiments saves time
- interactive access to (preliminary) simulation results can help



2005, code for full-scale connectivity

Advancing the Boundaries of High-Connectivity Network Simulation with Distributed Computing

Abigail Morrison
abigail@biologie.uni-freiburg.de
Computational Neurophysics, Institute of Biology III and Bernstein Center for Computational Neuroscience, Albert-Ludwigs-University, 79104 Freiburg, Germany

Carsten Mehring
Carsten.Mehring@biologie.uni-freiburg.de
Department of Zoology, Institute of Biology I and Berstein Center
Neuroscience, Albert-Ludwigs-University, 79104 Freiburg, Germa

Theo Geisel
geisel@chaos.gwdg.de
Department of Nonlinear Dynamics, Max-Planck-Institute for Dy
Organization, 37018 Göttingen, Germany

Ad Aertsen
aertsen@biologie.uni-freiburg.de
Neurobiology and Biophysics, Institute of Biology III and Bernstei
Computational Neuroscience, Albert-Ludwigs-University, 79104 Freiburg, Germany

Markus Diesmann
diesmann@biologie.uni-freiburg.de
Computational Neurophysics, Institute of Biology III and Bernstein Center for Computational Neuroscience, Albert-Ludwigs-University, 79104 Freiburg, Germany

from the abstract:

software scales excellently on a wide range of tested hardware, so it can be used in an interactive and iterative fashion for the development of ideas, and results can be produced quickly even for very large networks. In contrast to earlier approaches, a wide class of neuron models and synaptic dynamics can be represented.

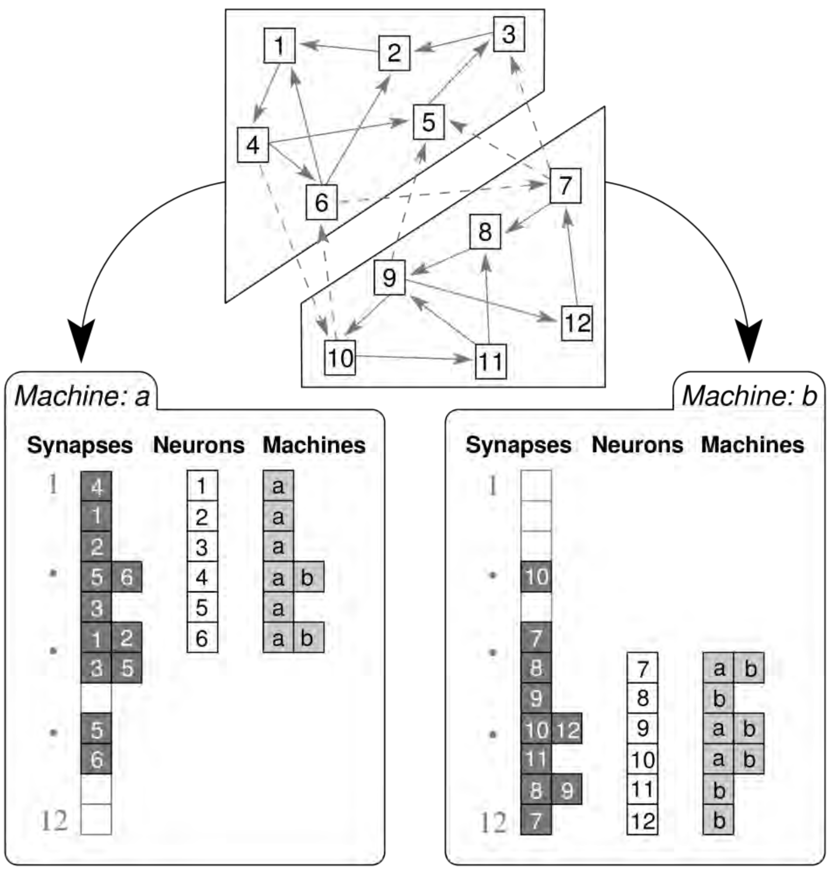
Neural Computation (2005) 17:1776-1801

Characteristics of brain simulations

- **memory not simulation time limits network size**
- intention to use full memory resources: maximum-filling scaling
- analysis based on mathematical model of memory consumption:
  - Kunkel S, Potjans TC, Eppler JM, Plesser HE, Morrison A and Diesmann M (2012) Meeting the memory challenges of brain-scale network simulation. *Front Neuroinform* 5:35
  - Serves to guide optimizations:

    at different scales different components of the software dominate memory consumption

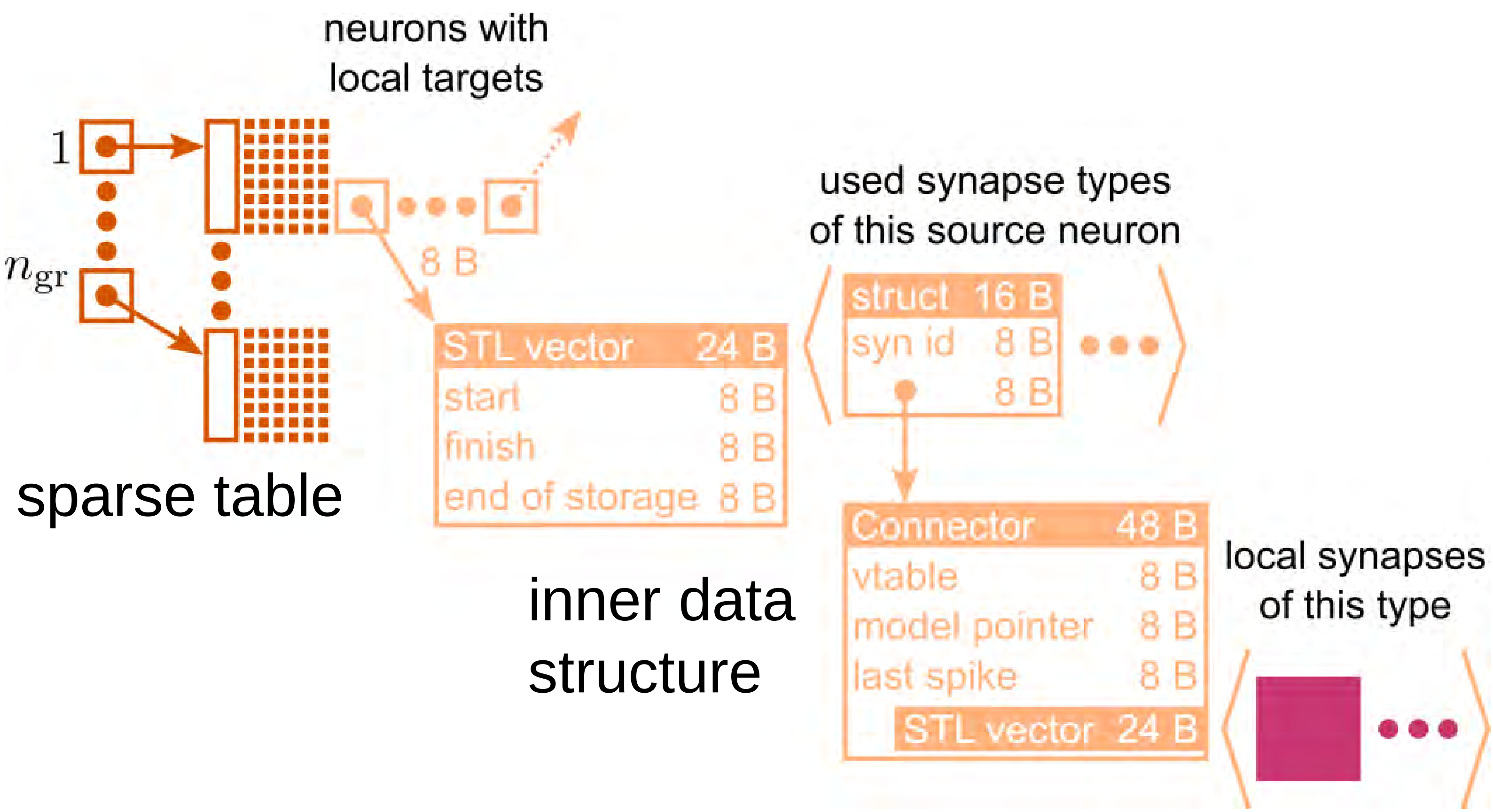
Distributed network representation



- distribution of neurons over processes
- collective communication (MPI)
- synapses represented on receiving neuron's machine
- 10^4 synapses / neuron: synapses dominate memory demands

3rd generation simulation kernel (3g)

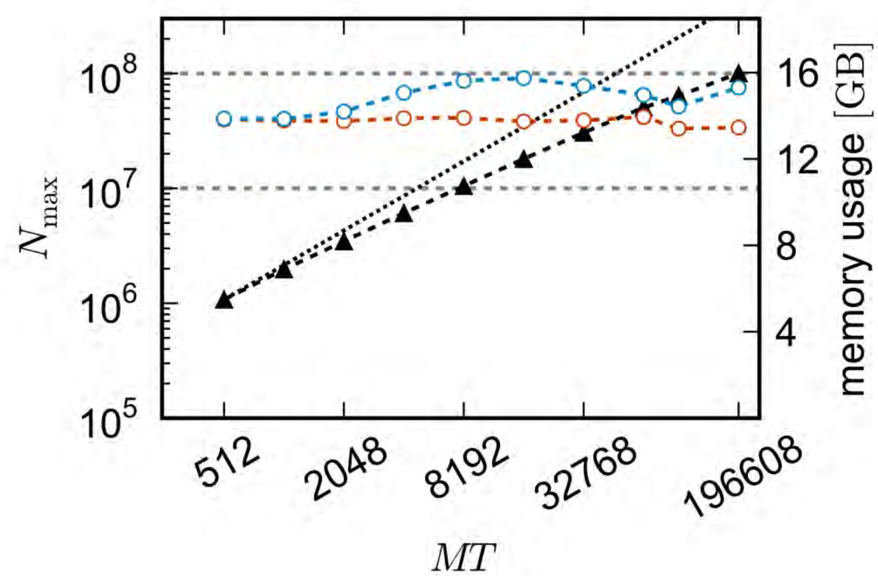
- required on each process



3g kernel – maximum filling scaling

(released with NEST 2.2 in December 2012)

- 11,250 synapses per neuron (exc-exc STDP)
- using up to M=196,608 compute nodes and T=8 threads per node
- up to 10^8 neurons on K (and JUQUEEN)
- 8 GB of memory per node

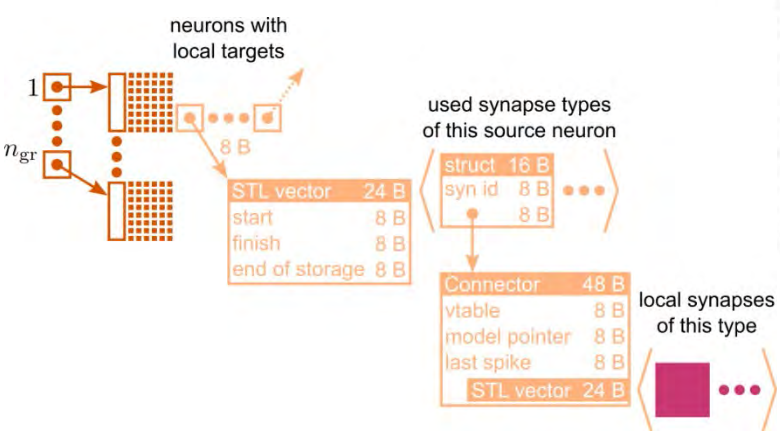


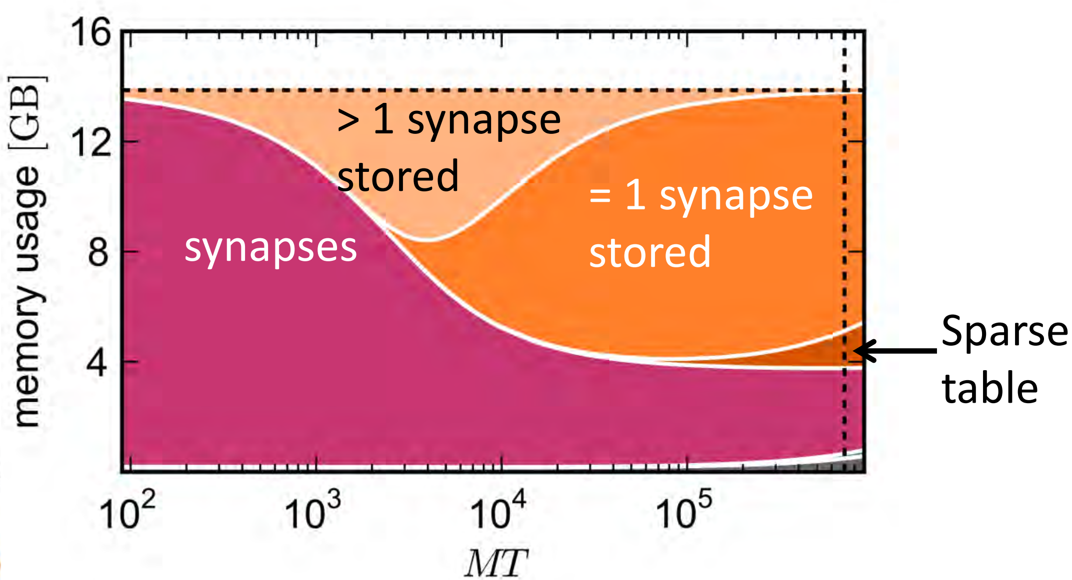
Helias et al. (2012) Supercomputers ready for use as discovery machines for neuroscience. *Front. Neuroinform.* **6**:26.

Memory footprint

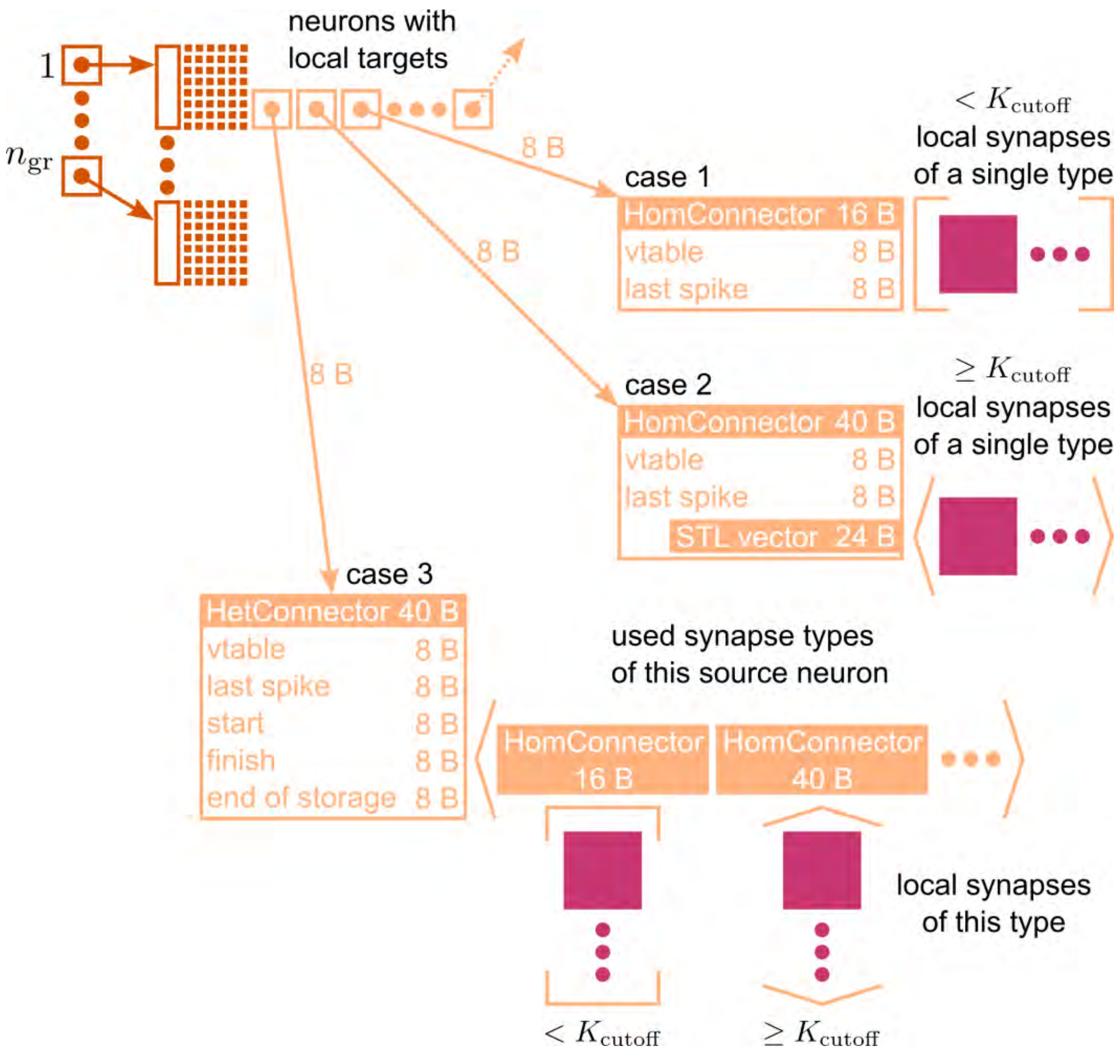
- **fundamental problem:**

  number of interaction partners << number of compute nodes
- adapt the model to account for short target lists
- potential solution: low-overhead data structure on supercomputers





New adaptive connection infrastructure (4g)

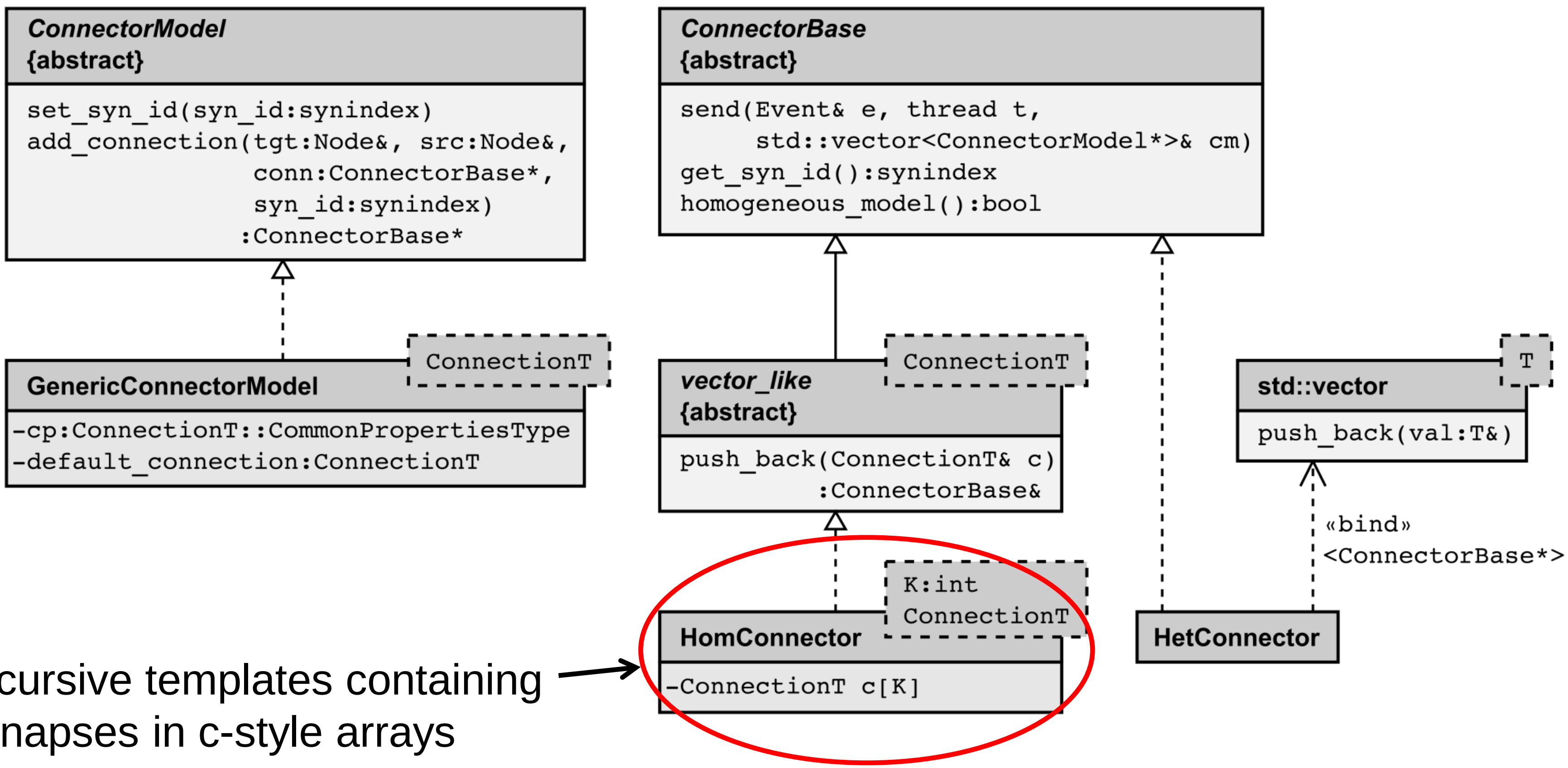


low overhead per synapse on supercomputers

use c-style arrays in limit of few synapses

full flexibility on laptops and moderately sized clusters

New adaptive connection infrastructure (4g)



recursive templates containing synapses in c-style arrays

#synapses K is recursion parameter

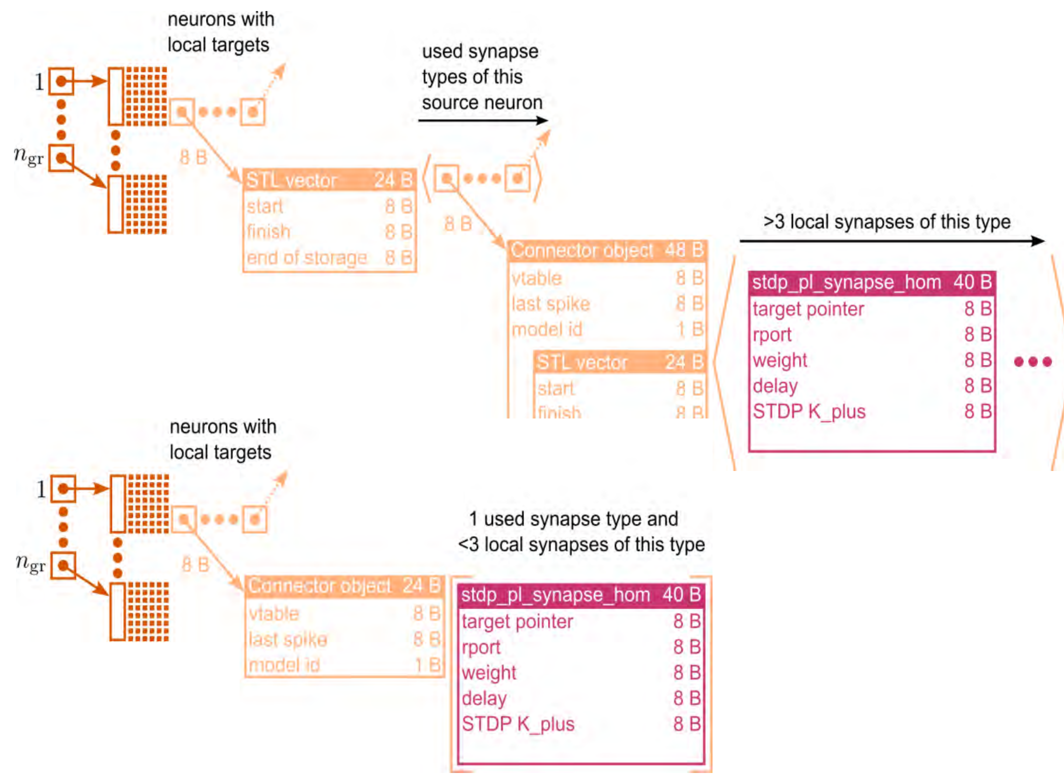
Memory layout of 3g and 4g kernel

3g memory layout

- accounts for sparseness in neuronal and connection data structures

4g memory layout

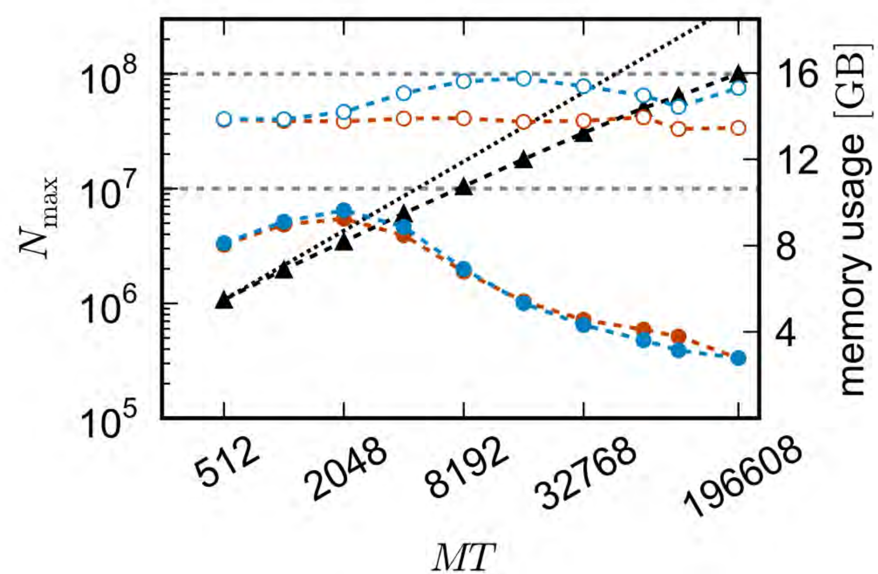
- novel adaptive data structure copes with short target lists
- not compromising on generality



HBP The Human Brain Project | nest:: initiative | RWTHAACHEN | JÜLICH FORSCHUNGSZENTRUM

Comparison of 4g to 3g kernel

- simulation of same network using 3g and 4g kernel



Reduced memory usage

- in all regimes of number of processes
- especially in the regime of 10k processes and beyond

(less than 1/3 at 100,000 cores)

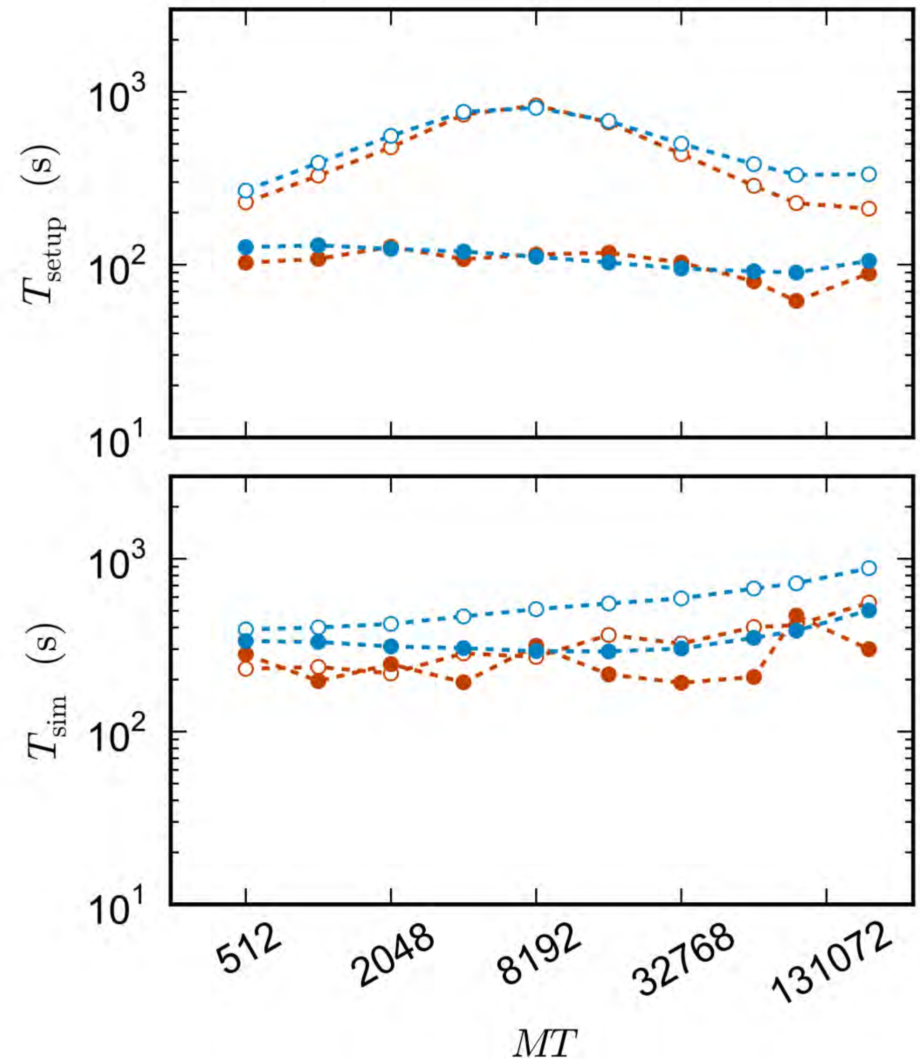
Comparison of 4g to 3g kernel

Reduced setup time

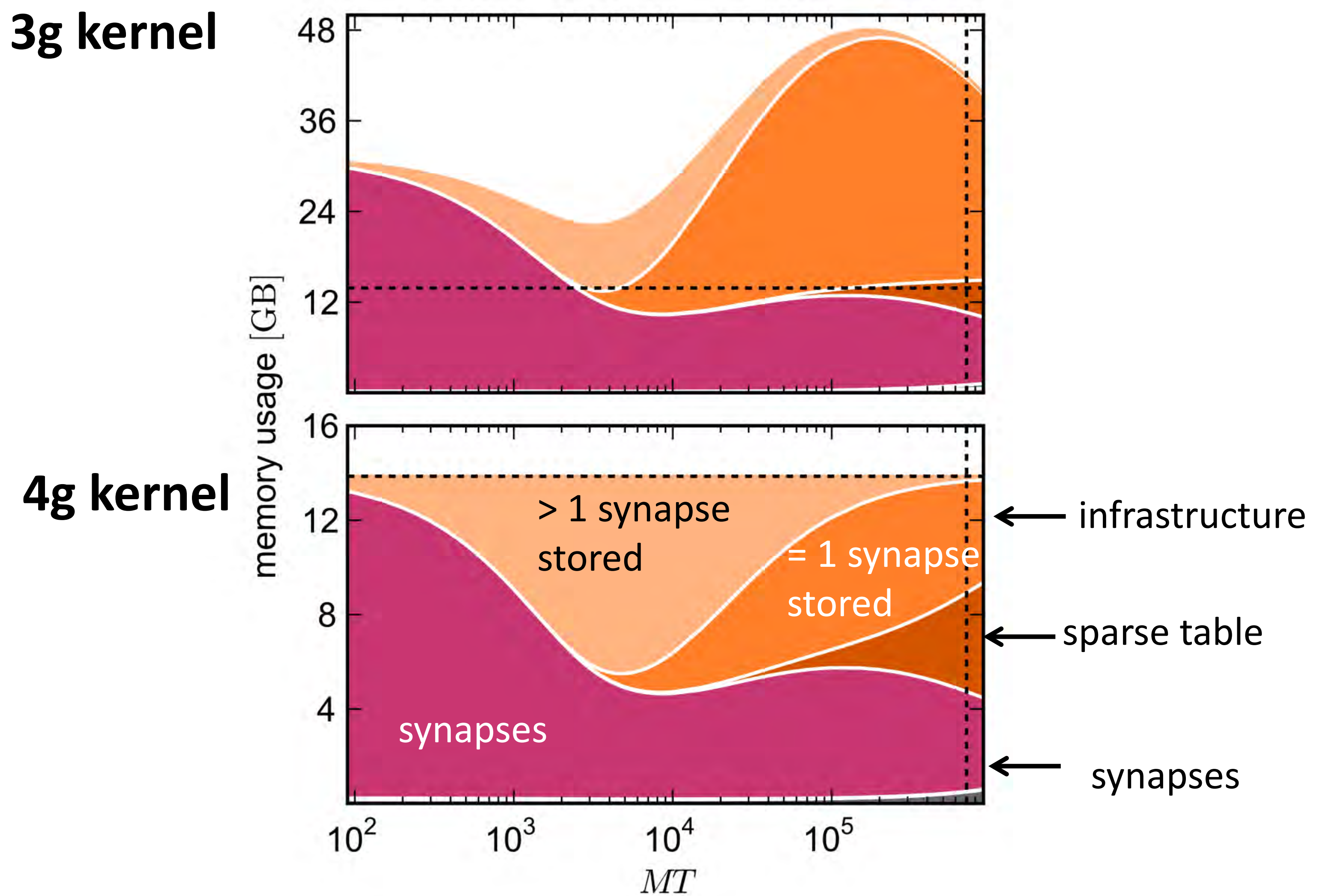
- optimization of wiring routines
- faster memory allocation using dedicated pool allocator

Reduced simulation time

- smaller objects in connection infrastructure enable more efficient use of cache

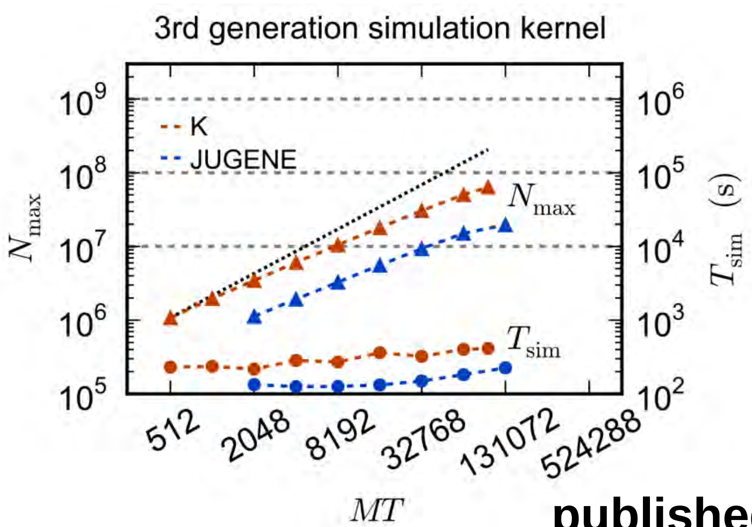


Comparison of 4g to 3g kernel

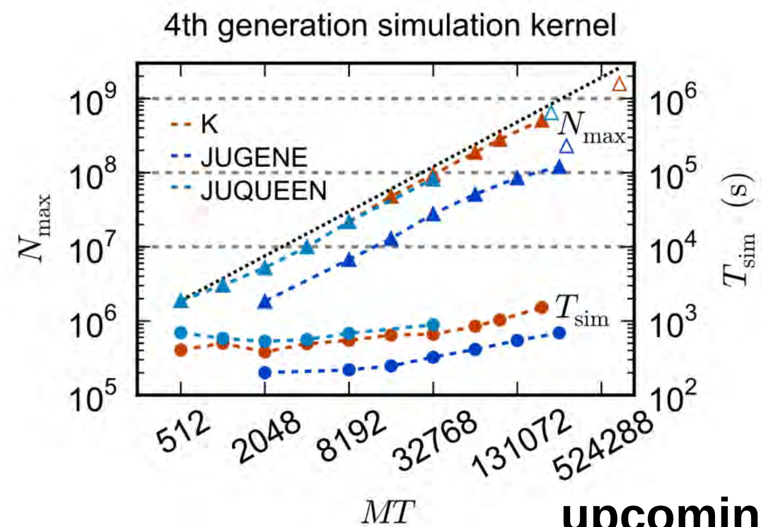


HBP The Human Brain Project · nest:: initiative · RWTHAACHEN · JÜLICH FORSCHUNGSZENTRUM

Maximum filling for 3g and 4g kernel



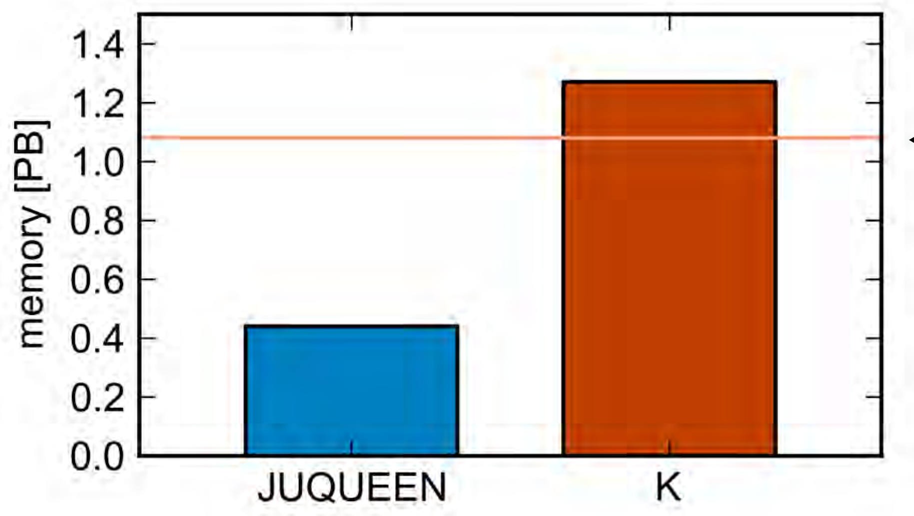
published



upcoming

- maximum-filling benchmarks (8 OpenMP threads / node on JUQUEEN,K)
- **world record August 2013:**
  - **1.85 10^9 neurons, 10^{13} synapses**
  - 1% of human brain on peta-scale computers
- almost linear scaling of network size with machine size

JUQUEEN and K



record simulation of 1.85×10^9 neurons and 6000 synapses per neuron with

- 24 B per STDP synapse
- 16 B per static synapse

- availability of JUGENE/JUQUEEN essential during K co-development
- record simulation only possible with K due to larger memory

Memory design of top-ten computers

順位	名前	国	速度(Gflops)	全メモリー量(GB)	コア数	コアあたりのメモリー量（GB）
1	Tianhe-2	China	33862700	1,024,000	3120000	0.33
2	Titan	United States	17590000	710,144	560640	1.27
3	Sequoia	United States	17173224	1,572,864	1572864	1.00
4	京	日本	10510000	1,410,048	705024	2.00
5	Mira	United States	8586612	—	786432	—
6	Stampede	United States	5168110	192,192	462462	0.42
7	JUQUEEN	Germany	5008857	458,752	458752	1.00
8	Vulcan	United States	4293306	393,216	393216	1.00
9	SuperMUC	Germany	2897000	—	147456	—
10	Tianhe-1A	China	2566000	229,376	186368	1.23

Summary

- model of local cortical network explains basic dynamical properties
- severely under-constrained; only 50% of connections are local
- mesoscopic measures require brain-scale
- production code for 10^8 neurons available (NEST 2.2, 3g)
- 10^9 prototype code for peta-scale ready (4g)
- no compromise on generality
- short run times:

  supercomputers as discovery machine for neuroscience

Challenges

- efficient use of large numbers of cores available in modern machines
- each core has only limited memory
- Exa-scale:
  - needs different communication patterns than collective MPI communication
  - new data structures to represent synapses
- equally important: work with community on reliability and reproducibility