Reproducible Models and Replicable Implementations

Current Trends in Computational Neuroscience

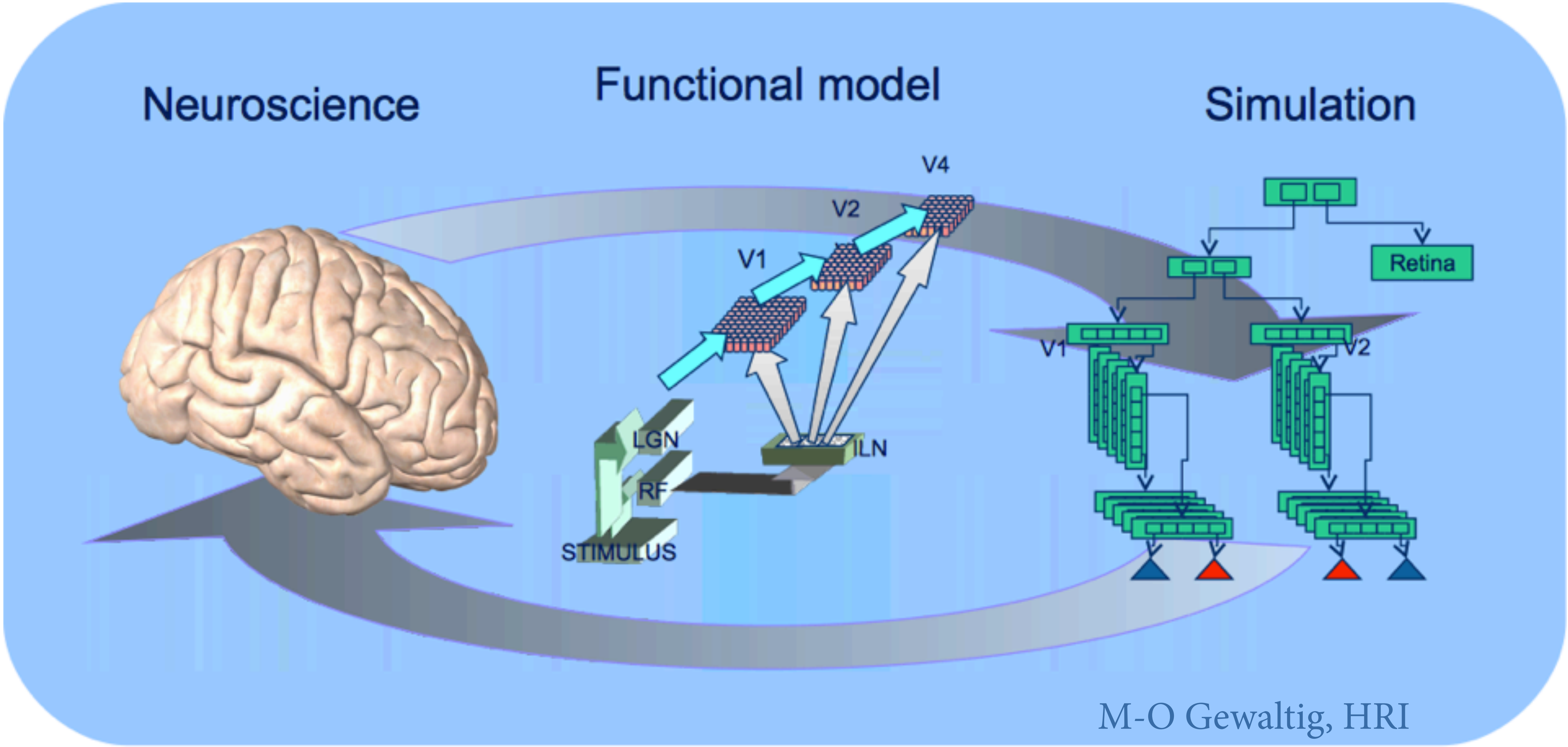
Hans Ekkehard Plesser, UMB
Sharon Crook, ASU
Andrew Davison, CNRS

Presented at SIAM Computational Science & Engineering
Reno, NV, 4 March 2011

UNIVERSITETET FOR MILJØ- OG BIOVITENSKAP
MDCCCLIX

Computational Neuroscience

What is Computational Neuroscience?



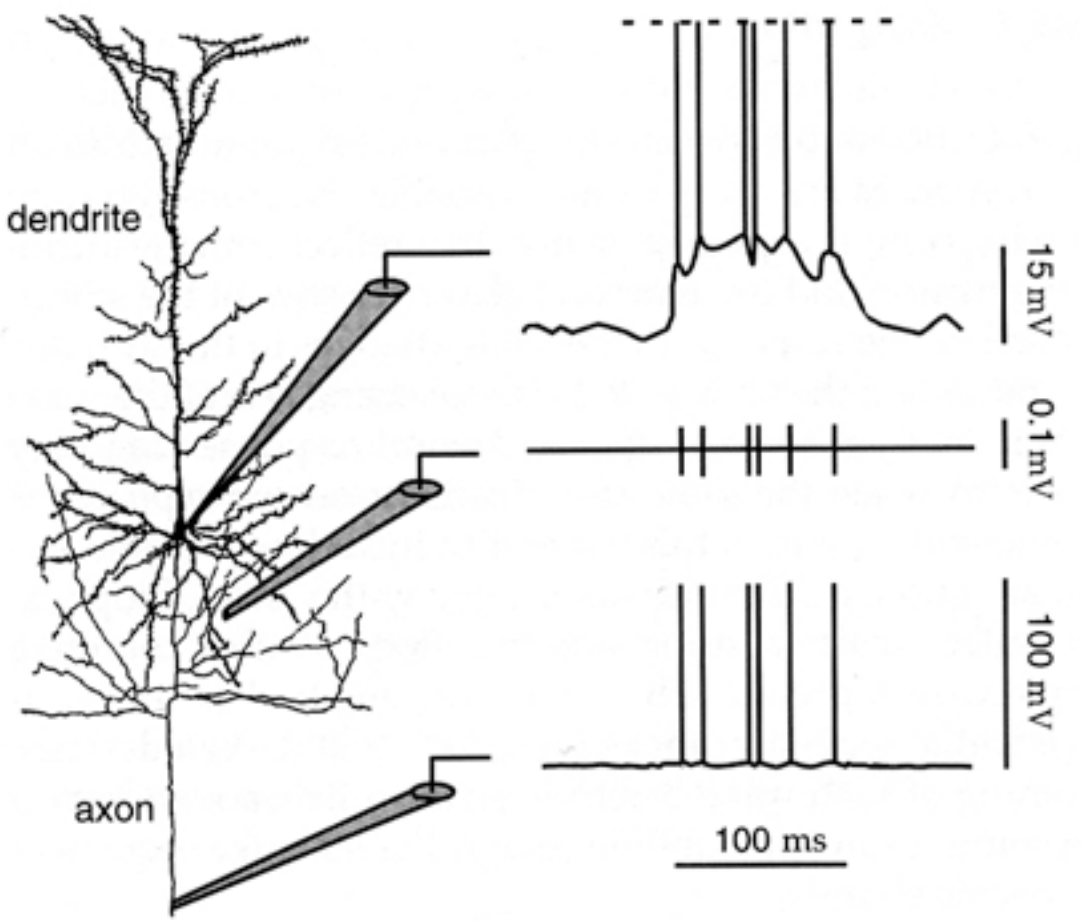
M-O Gewaltig, HRI

The goal of neural modeling is to relate, in nervous systems, function to structure on the basis of operation.

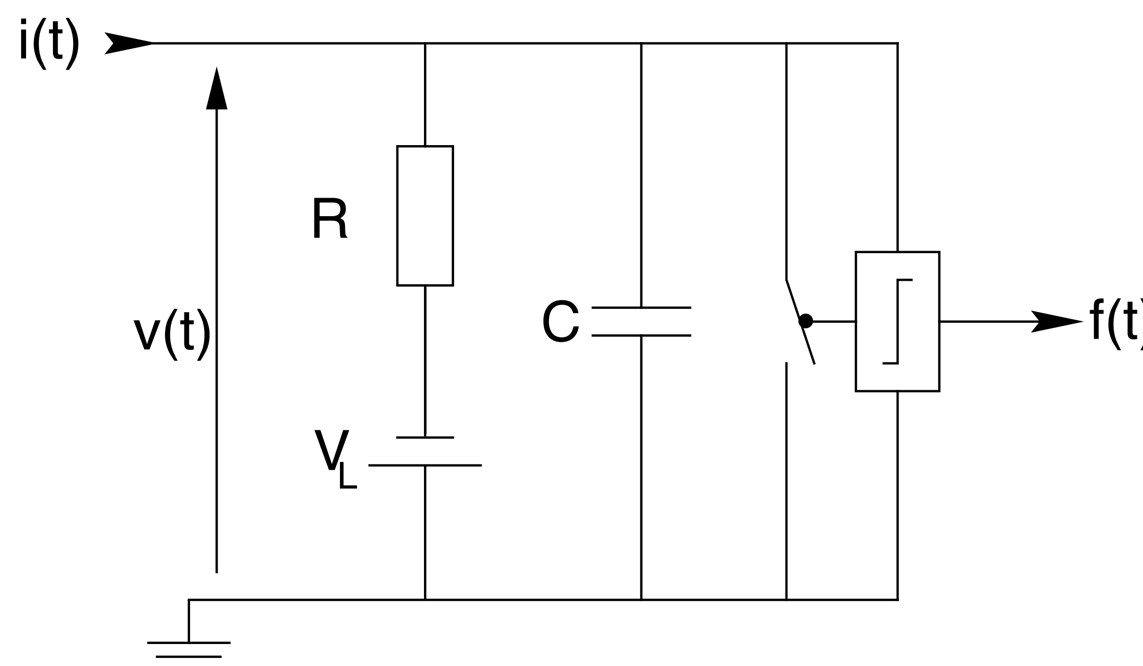
MacGregor & Lewis, 1977

Keeping it simple: point neurons

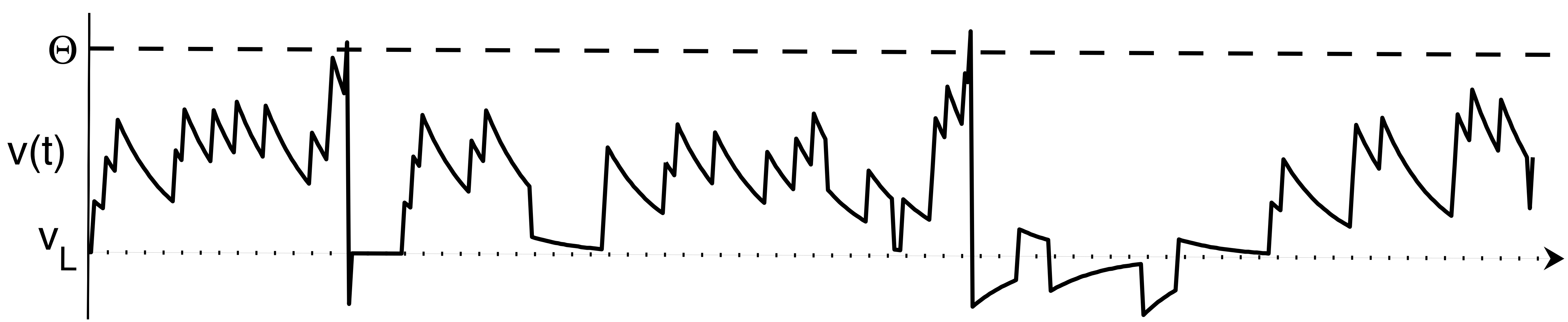
Complexity of the neuron



… abstracted as simple RC-circuit

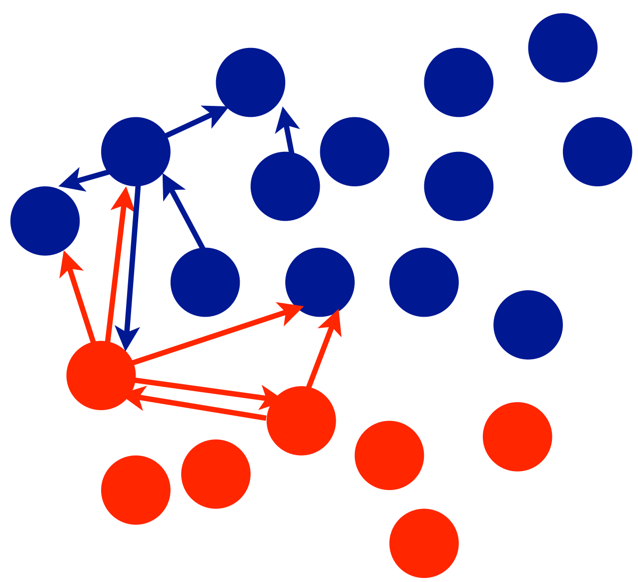


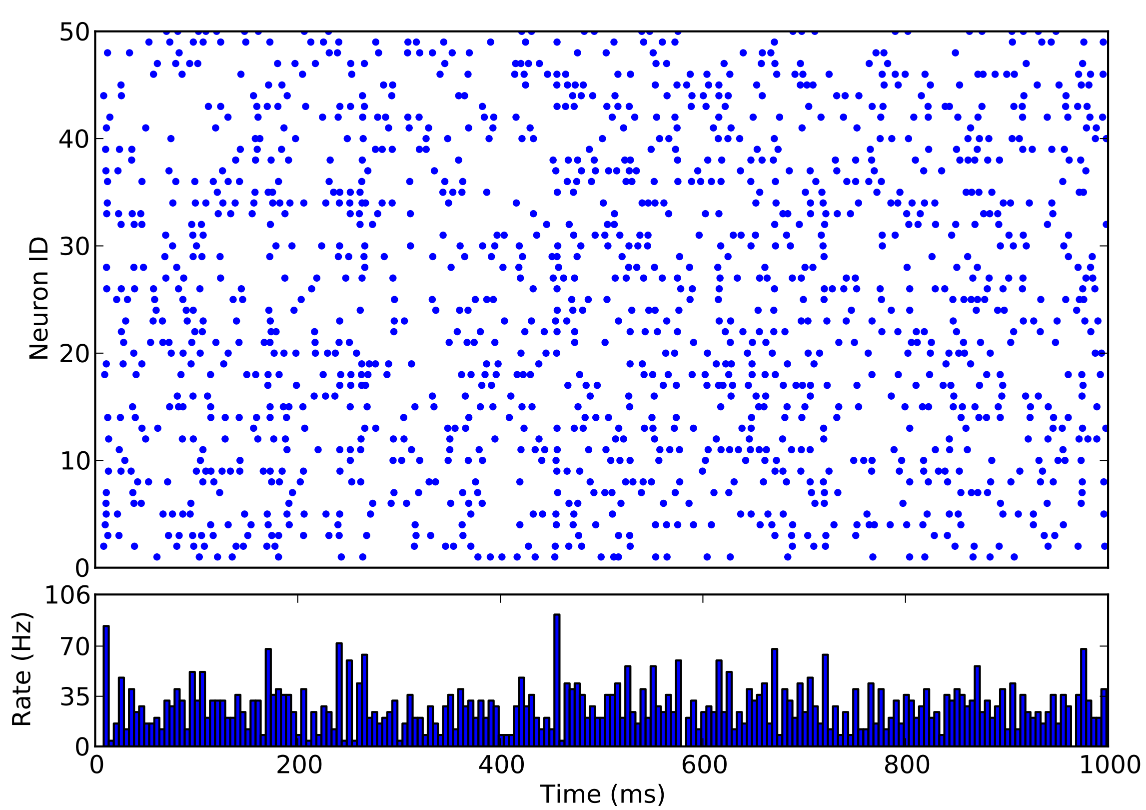
$$C\dot{V} = -\frac{(V - V_L)}{R} + i(t)$$



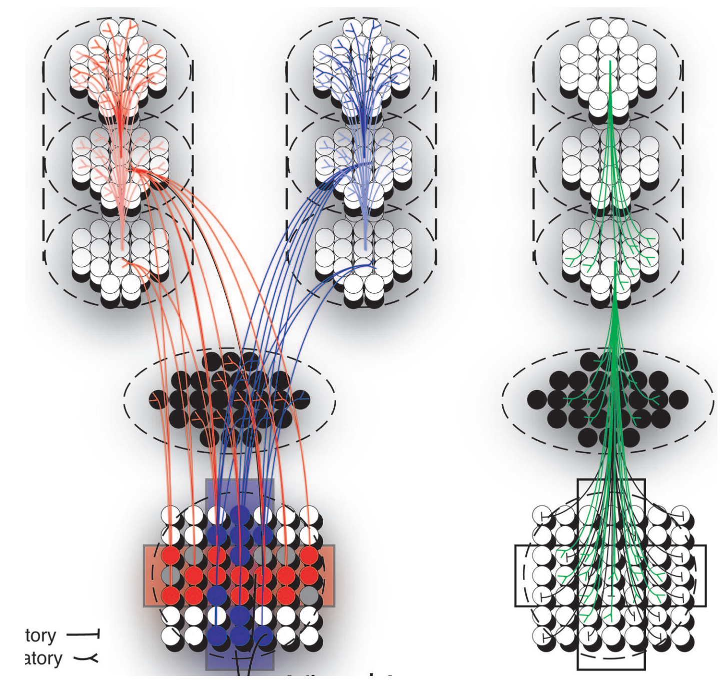
Then explore network dynamics!

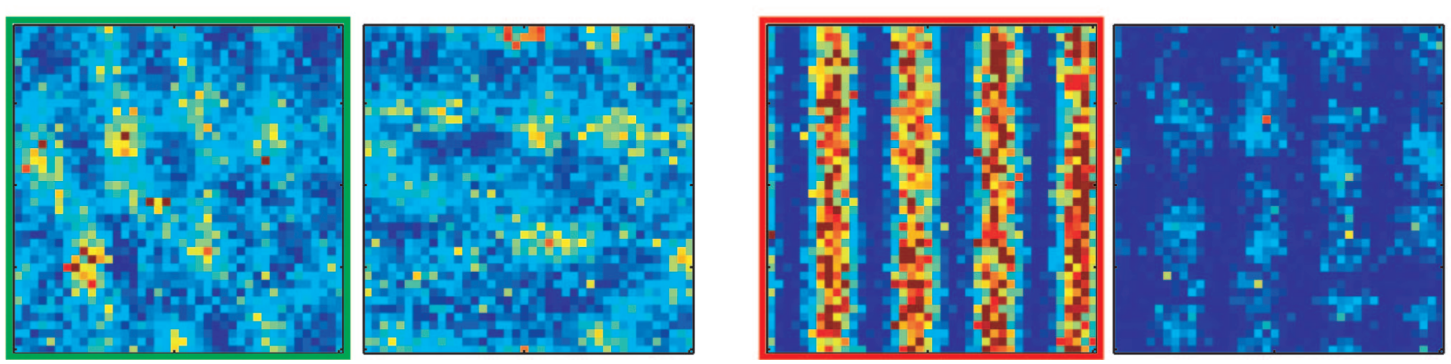
Random network





Structured network





Hill & Tononi (2005)

Simulations are exciting — but reliable?

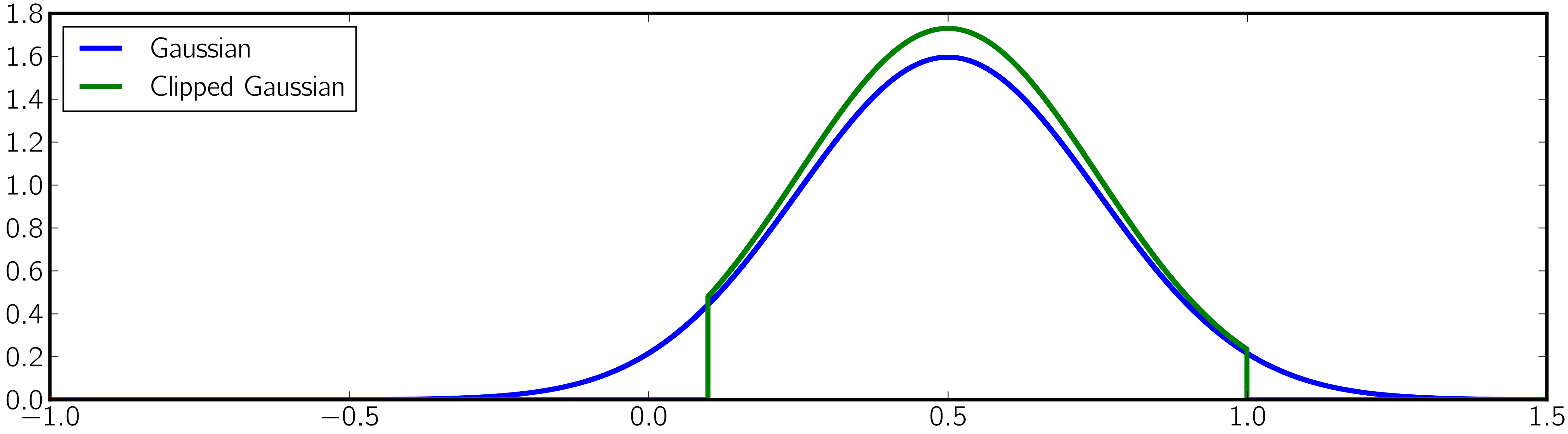
- Computational neuroscience
  - no conservation laws
  - no clear-cut separation of scales
  - no general agreement on which aspects of network activity are essential (spike rate vs spike time debate)
  - highly abstract models difficult to compare to experimental data quantitatively

➡Highly dependent on *reliable* simulations

➡Let's look at a real-life case

Case 2: The clipped Gaussian

- Well-known paper on plasticity
- Parameters chosen from Gaussian distribution, according to paper
- Results could not be reproduced independently
- Analysis of original C-code provided by authors:
  - ➡Parameters were chosen from *clipped* Gaussian



The (sad) state of the art

- Few published results can be reproduced independently
- Authors often struggle to replicate their own results
- Systematic comparison and evaluation of models are rare
- Authors rarely discuss why and how models ended up as they are
- Models are seldom re-used

Reproduction *vs* Replication

Chris Drummond

Replicability is not Reproducibility: Nor is it Good Science

ICML Montreal 2009

Replication: necessary, difficult, & insufficient

- Replication
  - Re-create identical results
  - Essentially book-keeping
  - Requires tools & discipline
  - No new insights: tests implementation, not ideas
- *Internal replication*
  Joe recreates results on original machine
- *External replication*
  Jane recreates Joe's results on her machine using Joe's code
- *Cross replication*
  Alice recreates Joe's results using a different simulator, based on a simulator-independent model description

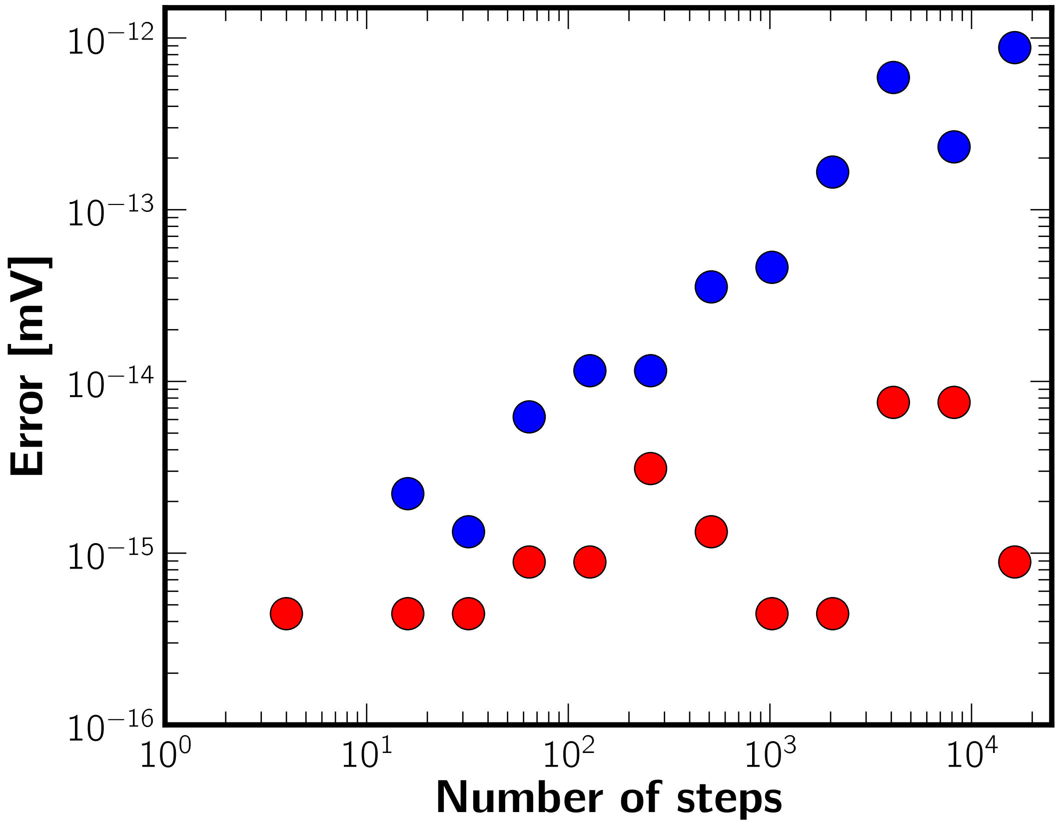
How much detail does replication require?

- Very simple point-neuron model

$$\dot{V} = -\frac{V}{\tau} + \frac{I_E}{C}$$

- Exact updating rule ($a = I_E \tau / C$)

$$V_{k+1} = V_k e^{-h/\tau} + a(1 - e^{-h/\tau})$$



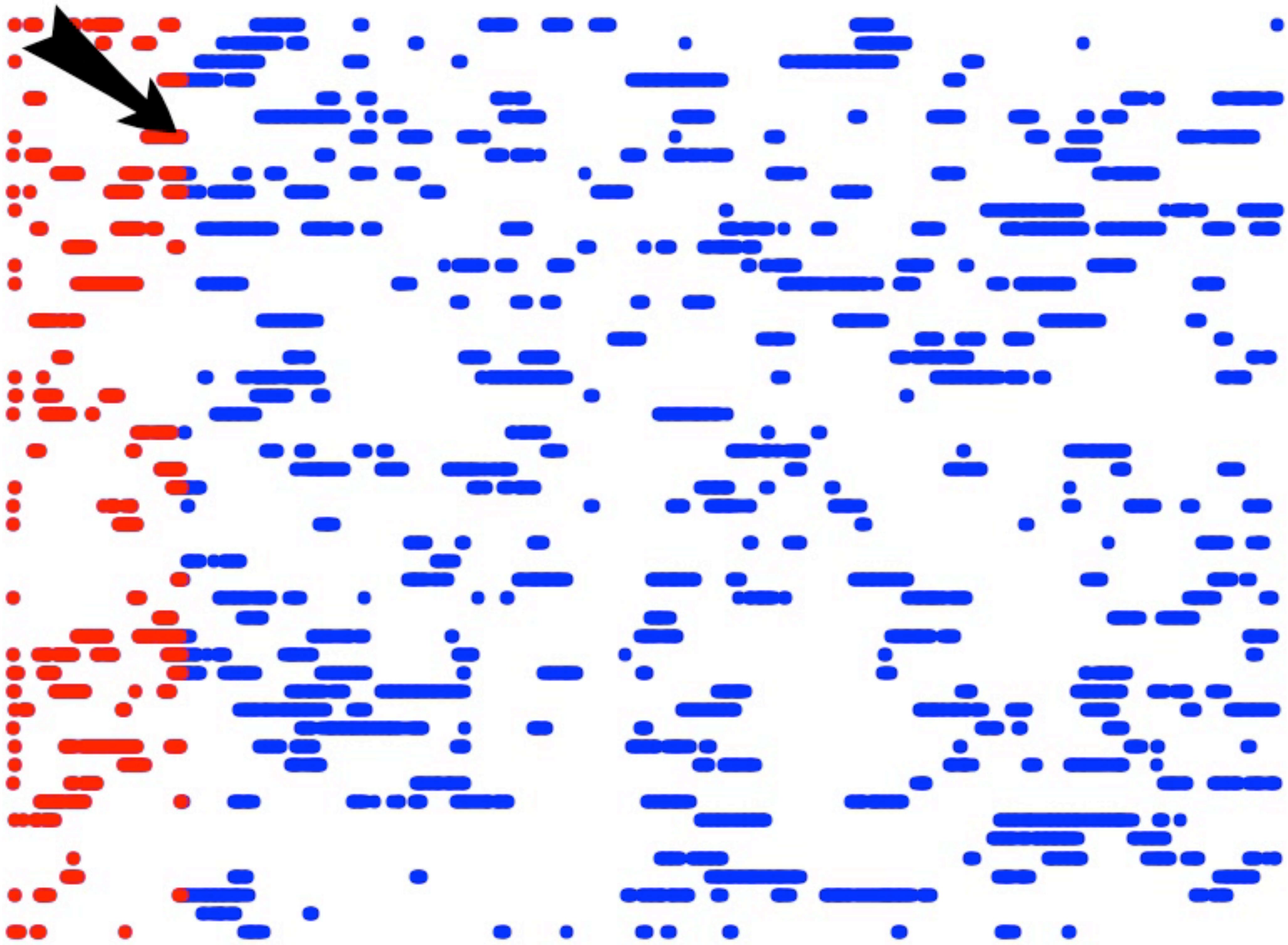
- Two different implementations

```
V[k+1] = V[k] * exp(-h/tau) - a * expm1(-h/tau)
V[k+1] = V[k] - (a - V[k]) * expm1(-h/tau)
```

- Different numerical properties

➡ Replication requires that we specify implementation!

But do such tiny differences matter?

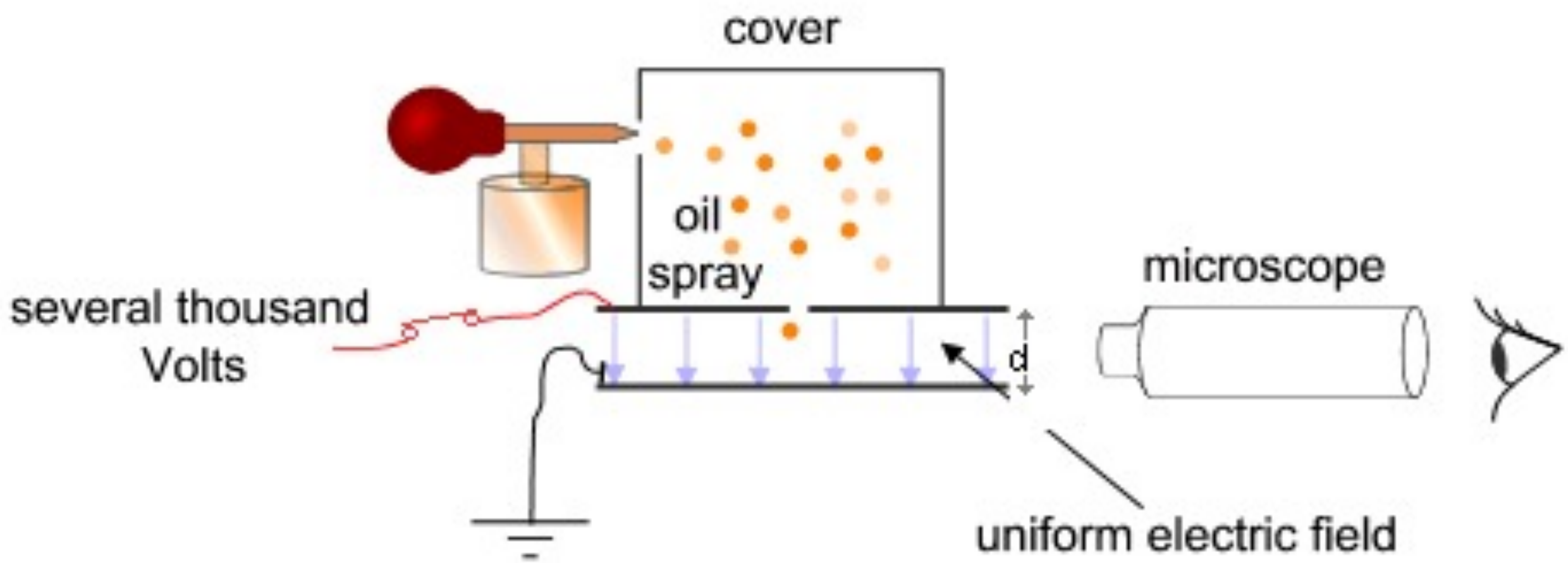


Reproduction

- Independent
- Test hypotheses and models
- Validates concepts
- Requires reflection

From oil drops to first-passage times

Millikan's oil drop experiment



Wikipedia

Erwin Schrödinger
Zur Theorie der Fall- und Steigversuche an Teilchen mit Brownscher Bewegung
Physikalische Zeitschrift **16**:289 (1915)

Reproduction needs replication

- Reproducible models describe scientific ideas
- Independent reproduction will generally fail to replicate original results precisely
- Requires learned judgement of discrepancies
- Requires means to understand failure

➡ Replicable implementation

Improving Scientific Practice in Computational Neuroscience

Good model description practice

- Systematic approach to model description in papers
- Standardize tables/checklists
- Standards for graphic representation of models
- Nordlie, Gewaltig, & Plesser *PLoS Comp Biol* **5**:e1000456 (2009)

A	Model Summary
Populations	Three: excitatory, inhibitory, external input
Topology	—
Connectivity	Random convergent connections
Neuron model	Leaky integrate-and-fire, fixed voltage threshold, fixed absolute refractory time (voltage clamp)
Channel models	—
Synapse model	α-shaped current inputs
Plasticity	—
Input	Independent fixed-rate Poisson spike trains to all neurons (during initial stimulation period)
Measurements	Spike activity

B	Populations	
Name	**Elements**	**Size**
E	Iaf neuron	$N_E = 4N_I$
I	Iaf neuron	N_I
E_{ext}	Poisson generator	$C_E(N_E + N_I)$

C	Connectivity		
Name	**Source**	**Target**	**Pattern**
EE	E	E	Random convergent $C_E \to 1$, weight J, delay D
IE	E	I	Random convergent $C_E \to 1$, weight J, delay D
EI	I	E	Random convergent $C_I \to 1$, weight $-gJ$, delay D
II	I	I	Random convergent $C_I \to 1$, weight $-gJ$, delay D
Ext	E_{ext}	$E \cup I$	Non-overlapping $C_E \to 1$, weight J, delay D

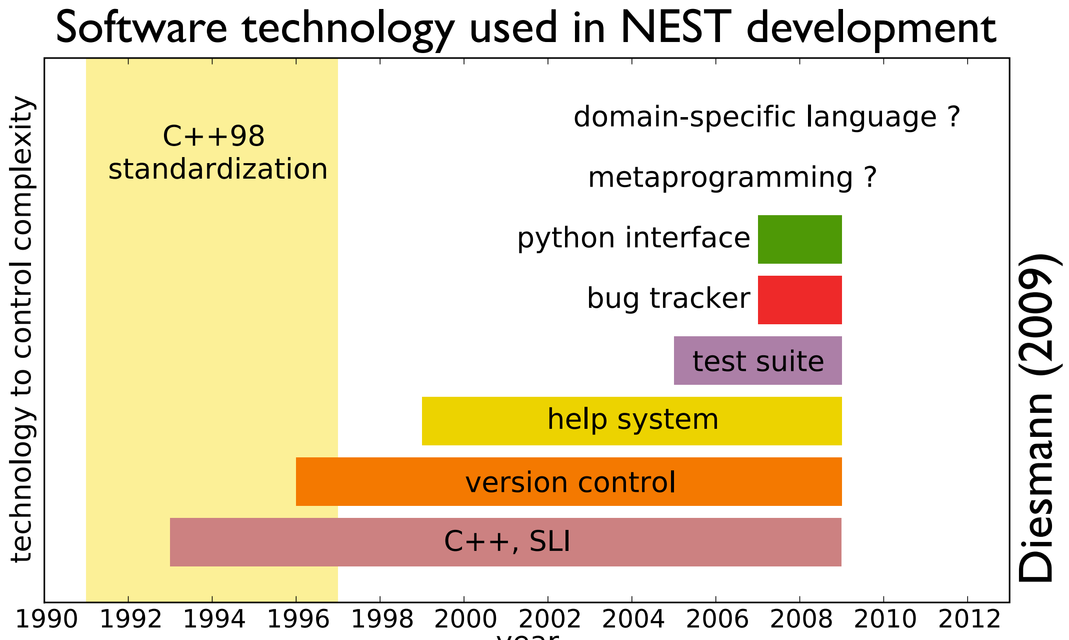
D	Neuron and Synapse Model
Name	Iaf neuron
Type	Leaky integrate-and-fire, α-current input
Subthreshold dynamics	$\tau\dot{V}(t) = -V(t) + RI(t)$ if $t > t^* + \tau_{rp}$; $V(t) = V_r$ else; $I(t) = \frac{\tau}{R}\sum_{\tilde{t}} w\alpha(t - (\tilde{t} + \Delta))\Theta(t - (\tilde{t} + \Delta))$
Spiking	If $V(t-) < \theta \wedge V(t+) \geq \theta$ 1. set $t^* = t$ 2. emit spike with time-stamp t^*

E	Input
Type	**Description**
Poisson generators	Fixed rate ν_{ext}, C_E generators per neuron, each generator projects to one neuron; active only during initial stimulation period

F	Measurements
Spike activity as raster plots for subset of excitatory neurons

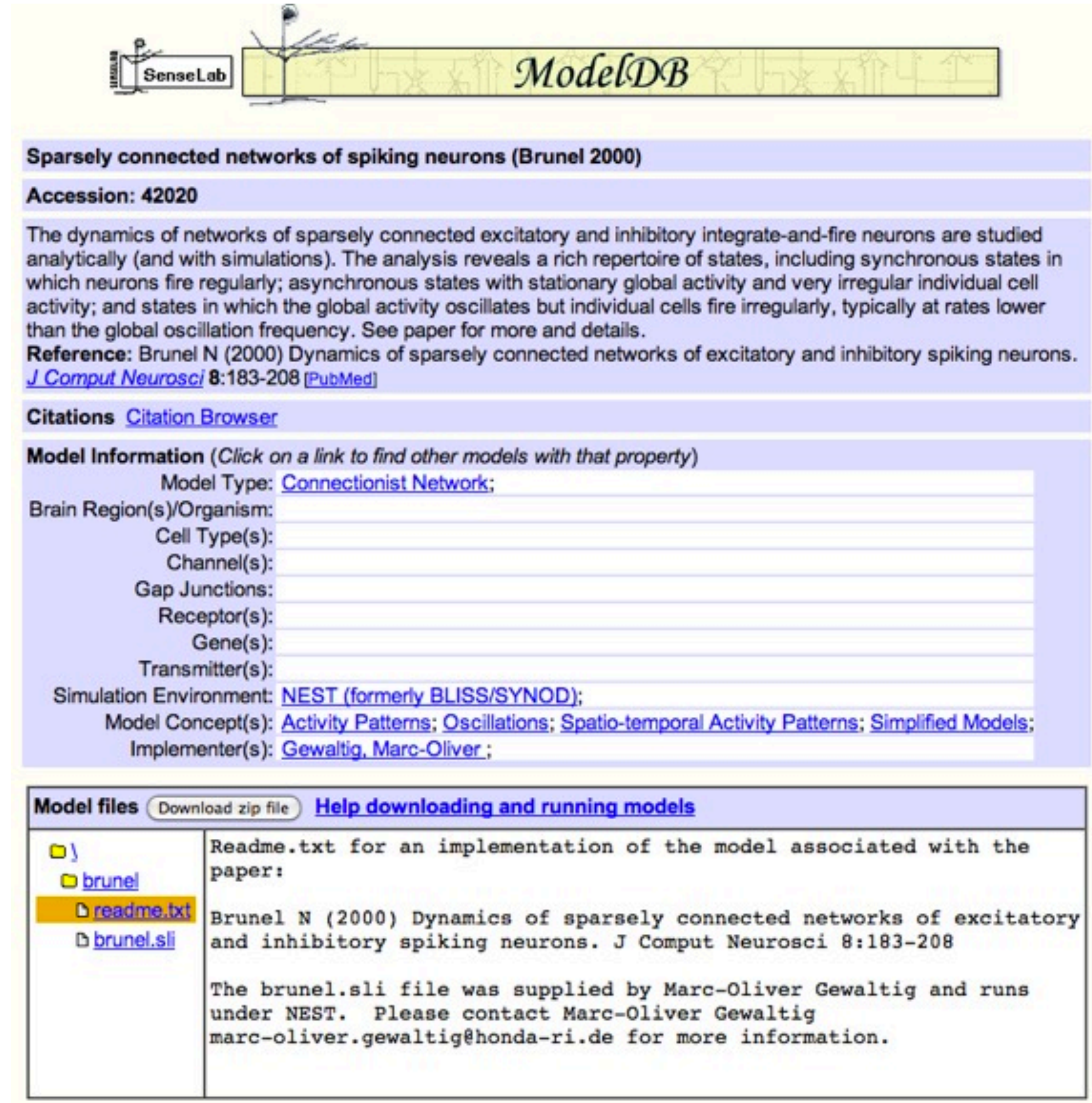
Professional, shared software

- Widely used packages replace homemade *ad hoc* code
- Currently: Neuron, NEST, Genesis, Moose, Brian, PCSim
- "Social" developments
  - Simulator review (Brette *et al*, 2007)
  - Teaching software at summer schools
  - Large-scale *scientific* projects (eg EU FACETS)
  - Neuroinformatics journals
  - Raising awareness among reviewers and editors
- "Technical" developments
  - Version control
  - Test suites



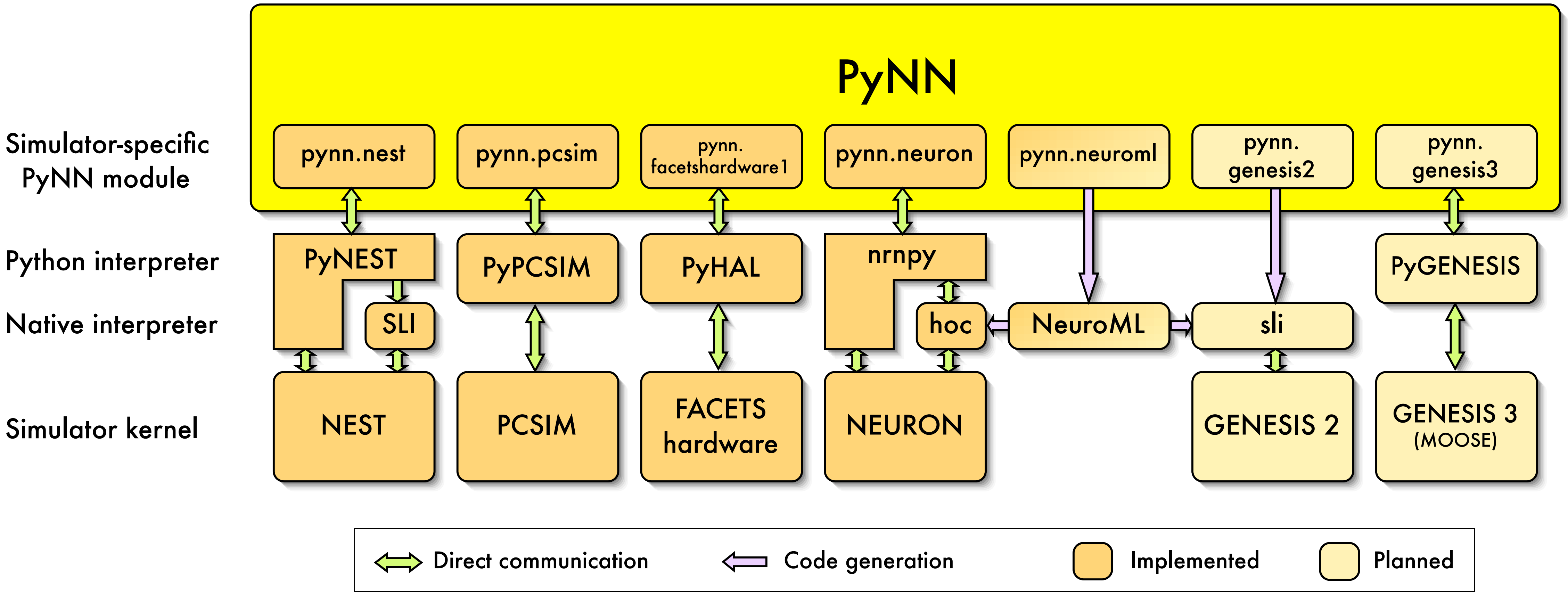
ModelDB: Sharing models

- Curated database of computational neuroscience models
- Only published models
- Open to any software
- Nearly 700 models
- http://senselab.med.yale.edu/modeldb/



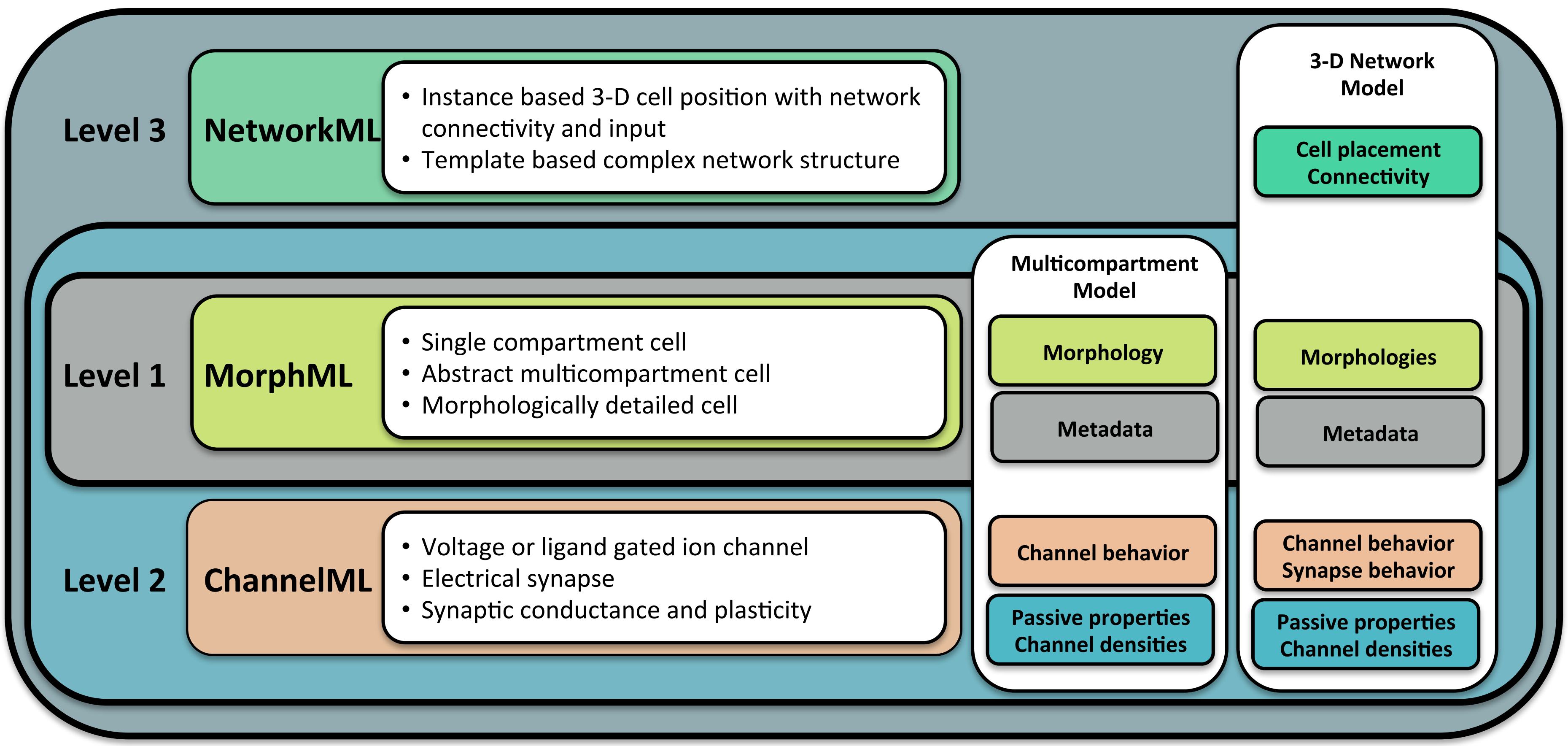
PyNN: A Meta-Simulator

- Python-based wrapper for many simulators
- Write model and simulation code once, run on all
- Facilitates model sharing and cross-validation
- Developed by Andrew Davison for FACETS project
- http://neuralensemble.org/PyNN



NeuroML: A model specification language

- XML-based language for model specification
- Multiple layers: channels, neuron morphologies, networks
- Code-generation for several simulators, including PyNN
- Facilitates model sharing and re-use
- http://www.neuroml.org



Provenance tracking: Sumatra

- Python package to enable systematic capture of the environment of numerical simulations/analyses
- Tracks simulation code, dependencies, platform information, results
- Developed by Andrew Davison as part of FACETS project
- http://neuralensemble.org/sumatra

Sumatra: TestProject: List of records

http://127.0.0.1:8002/

TestProject: List of records

Delete include data	Label	Reason	Outcome	Duration	Processes	Simulator		Script			Date	Time	Tags
						Name	Version	Repository	Main file	Version			
☐	20100709-154255		'Eureka! Nobel prize here we come.'	0.59 s		Python	2.5.2	/Users/andrew/tmp/SumatraTest	main.py	396c2020ca50	09/07/2010	15:42:55	
☐	20100709-154309			0.59 s		Python	2.5.2	/Users/andrew/tmp/SumatraTest	main.py	396c2020ca50	09/07/2010	15:43:09	
☐	haggling	'determine whether the gourd is worth 3 or 4 shekels'	'apparently, it is worth NaN shekels.'	0.59 s		Python	2.5.2	/Users/andrew/tmp/SumatraTest	main.py	396c2020ca50	09/07/2010	15:43:20	foobar

NineML: A model description standard

- Aims for community standard for declarative model descriptions
- Inspired by SBML and CellML
- Focus on networks of point neurons
- Under development by INCF Multi-scale Modeling Task Force

Perspectives

- Community increasingly aware of need for reproducibility and replicability
- Large-scale projects have led to development of valuable tools
- Summer schools educate PhD-students and post-docs in use of established modeling tools
- Neuroinformatics journals allow publication of domain-specific solutions
- International Neuroinformatics Coordinating Facility (INCF) stimulates debate and development
- NEST Initiative is devoted to furthering reliable simulations

➡ We have a long way to go, but we are (finally) moving!

Collaborators

Sharon Crook

School of Mathematical and Statistical Sciences, School of Life Sciences, Center for Adaptive Neural Systems
Arizona State University

Andrew Davison

Unité de neuroscience, information et complexité
CNRS
Gif-sur-Yvette

Eilen Nordlie

Marc-Oliver Gewaltig

Honda Research Institute Europe
Offenbach

nest:: initiative

Research Council of Norway (eVita program)