NSF CRCNS Annual Meeting, October 24-26, 2016, Paris, France

US-German Collaboration: Strategy Change
in Cognitive Biological and Technical Systems (NSF-1311165)

Functional Reorganization via Phase Transitions in Early Sensory Areas During Category Learning

Robert Kozma*, Yury Sokolov
University of Memphis, Memphis, TN, & *Umass, Amherst, MA, USA

Walter J. Freeman, University of California at Berkeley
Berkeley, CA, USA

Miklos Ruszinko
Renyi Institute of Mathematics, Hungarian Academy, Budapest, Hungary

Sanqing Hu
Hangzhou Dianzi University, Hangzhou, China

Frank Ohl, Andreas Schultz, Tim Wanger
Leibniz Institute for Neuroscience, Magdeburg, Germany

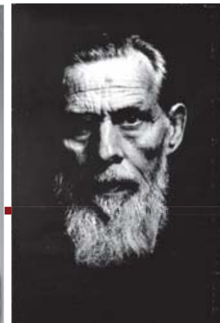
His Legacy - Freeman Neurodynamics

Walter Freeman

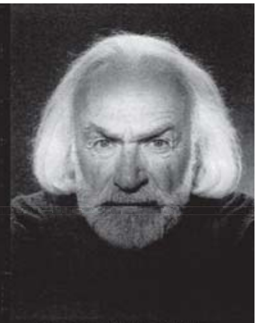
(January 30, 1927, DC – April 24, 2016, Berkeley, CA)

Charles Sherrington 1857-1951

John Fulton 1899-1960



Embodiments of Mind Warren S. McCulloch



Languages of the Brain Karl Pribram

Ross Ashby, Warren McCulloch, Grey Walter, Norbert Wiener. 1956

Ilya Prigogine 1917-2003

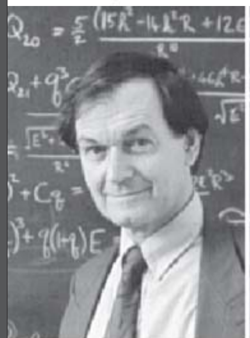
Hermann Haken 1929-

Alfred Renyí, 1920 -1970

Paul Erdös, 1913 -1996

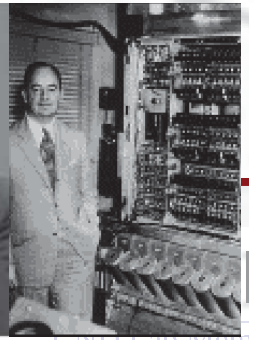
Martin Heidegger 1889-1976

M. Merleau-Ponty 1908-1961



R. Penrose: Q M

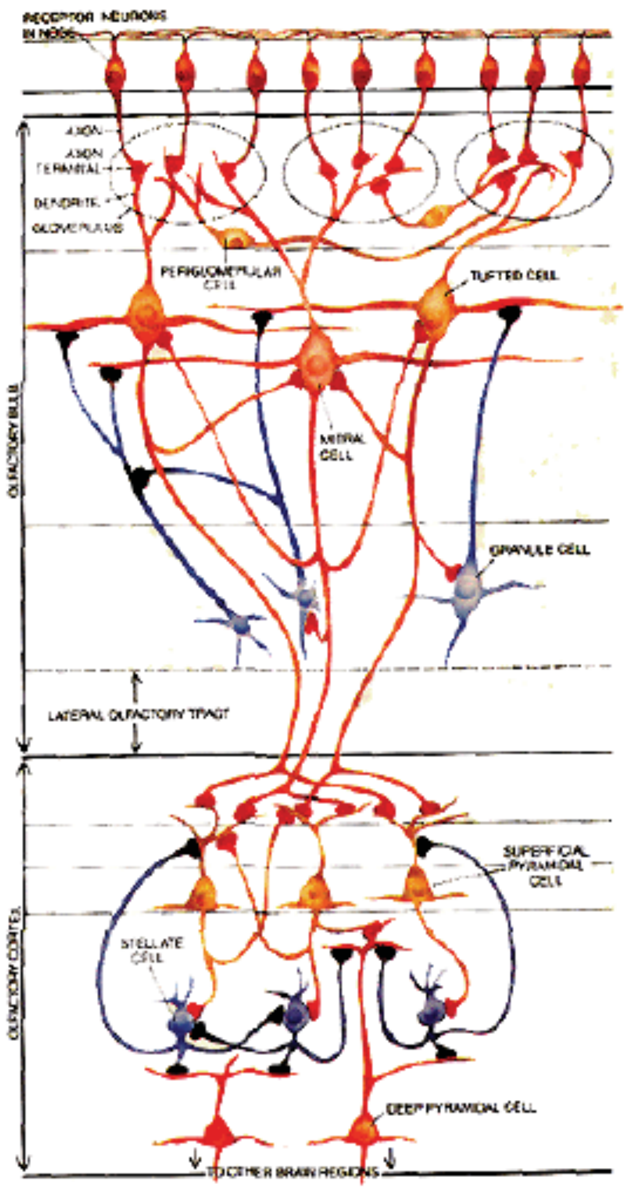
H. Umezawa: Q F T

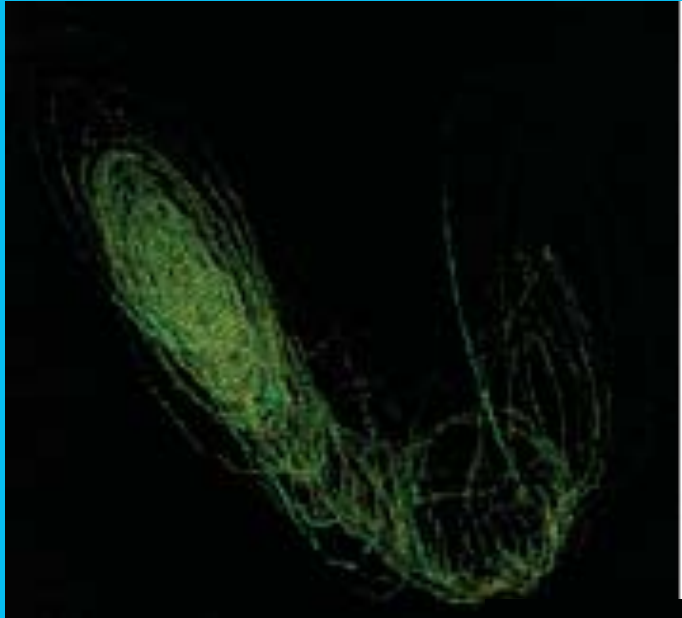


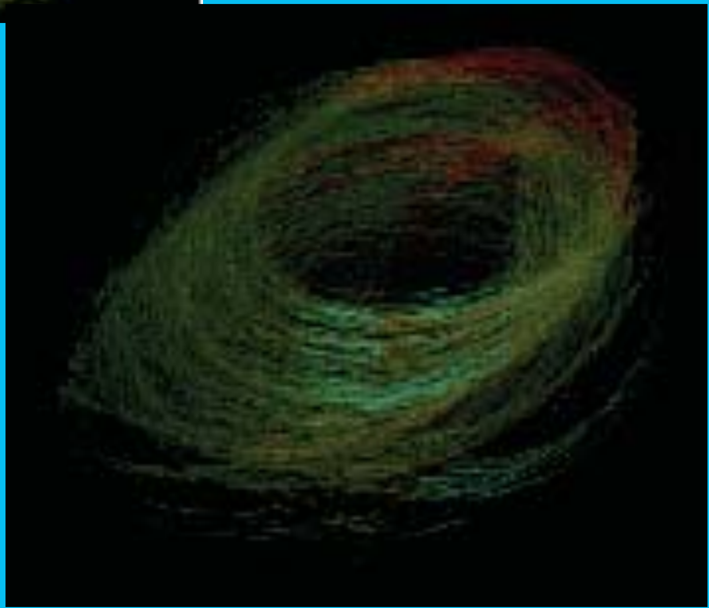
"Brains lack the arithmetic and logical depth that characterize our computations… ."

"We require exquisite numerical precision over many logical steps to achieve what brains accomplish in very few short steps." J.Neumann, The Computer and the Brain, 1958.

WJ Freeman, The Physiology of Perception
Scientific American, 1991







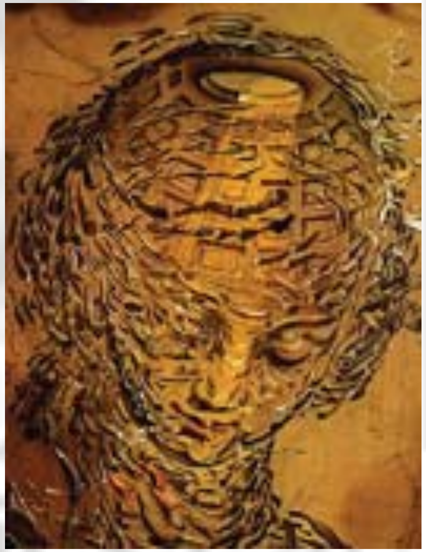
Key Dilemma
- Formal AI or/and Neuroscience -

"Controversy" concerning the top down (AI) versus bottom up (big data/deep learning) divide.

- → Instead view them as opposites
- → Integrate the available approaches

Clue from nature

- → Freeman intentional action-perception cycle
- → Metastability in brains Scott Kelso (in Haken/ Prigogine footsteps)
- → The complimentary nature



How brain Integrates symbolic and subsymbolic processing?

Metastability leads to coexisting brain states which sequentially follow each other (a few times/ second)

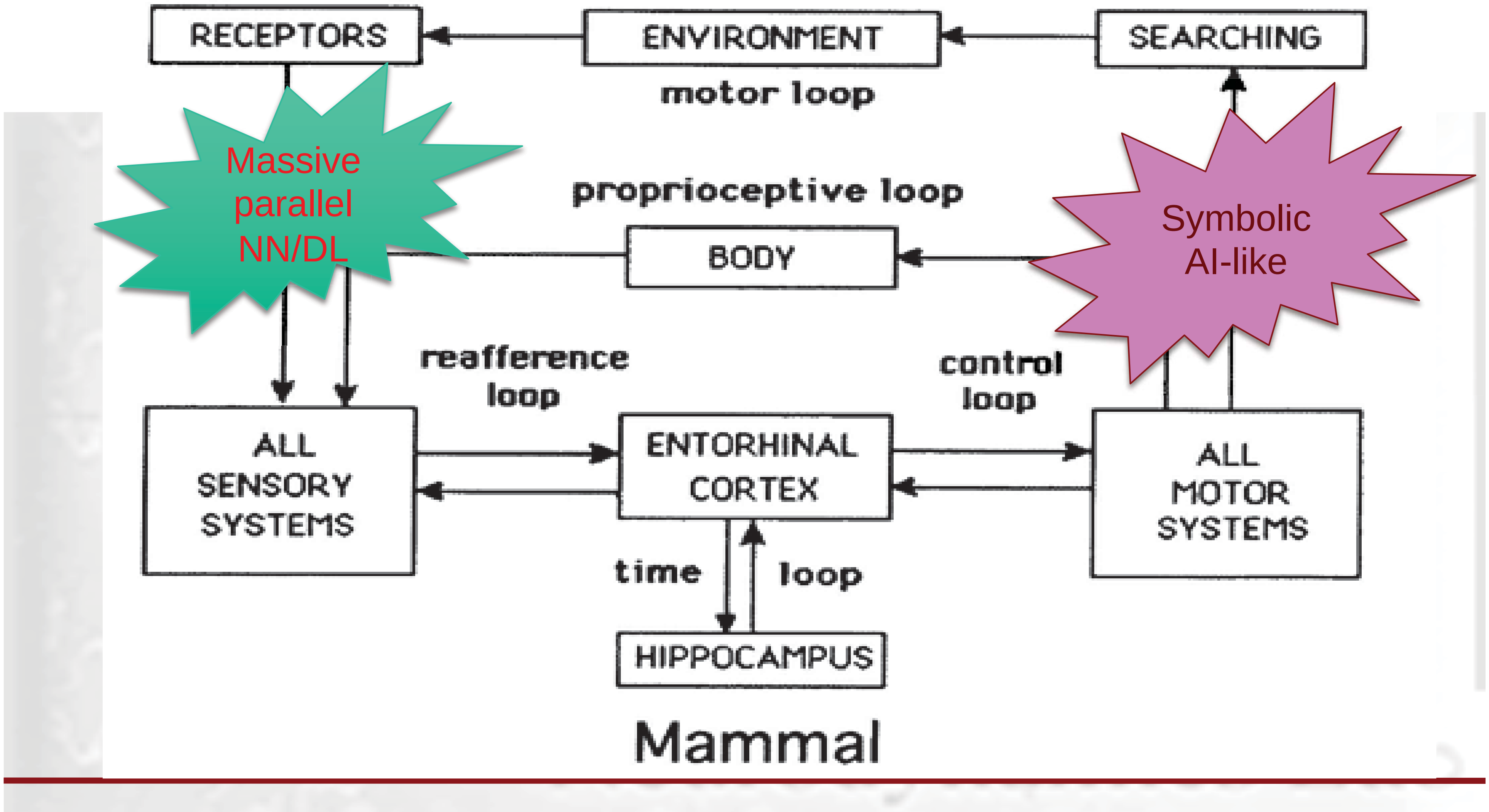
- **High dimensional:** bottom up emergence (subsymbolic)
- **Low dimensional:** top down: condensation (symbolic)
- → SINGULARITY/PHASE TRANSITION in space-time

We can combine those (brain like dynamic patterns) sequentially as in an intentional cycle

- Data-intensive bottom up → Symbol generation/rule base
- → Decision → Action → Data intensive bottom up → ….

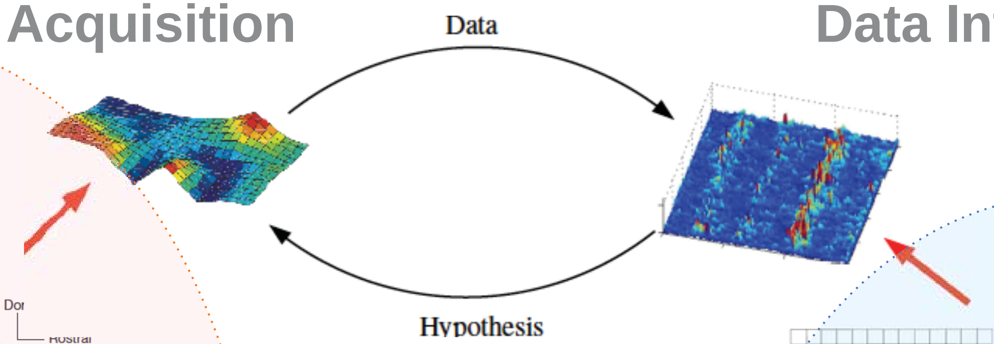
Suggested AI-Neuroscience Cycle

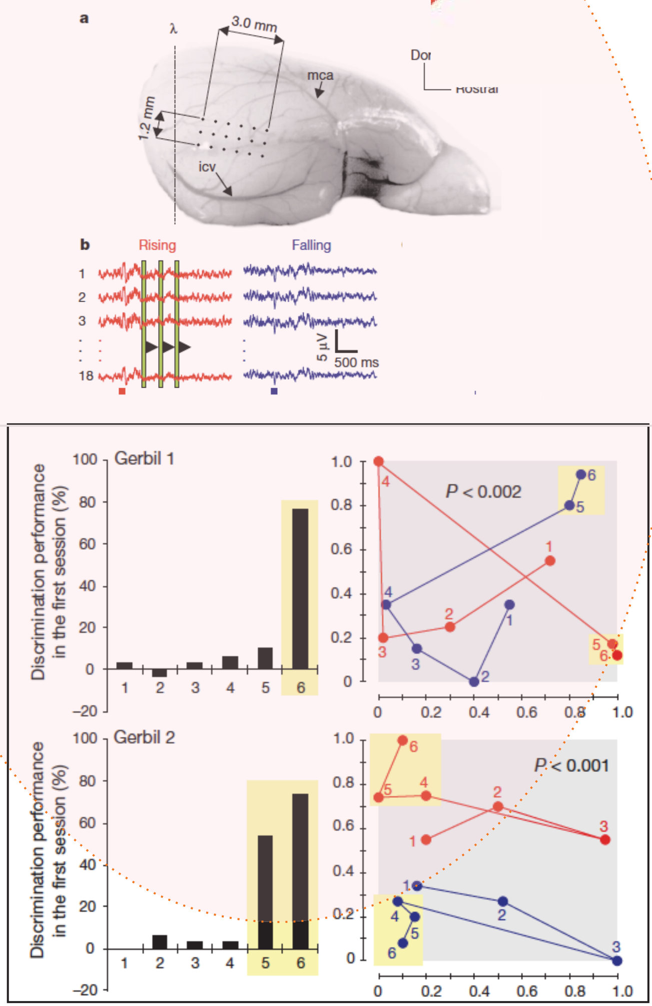
- following Freeman action-perception cycle -



Rodent Learning (LIN/Germany)
ECoG Data Acquisition

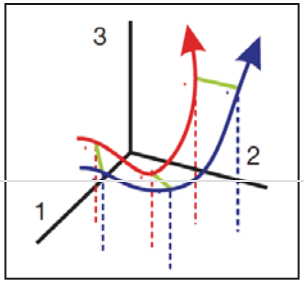
Graph Modeling (CLION/US)
Data Interpretation



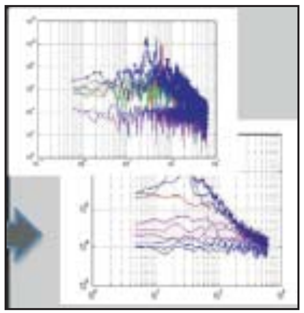


Ohl et al., Nature, 2001

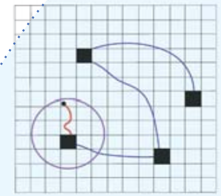
Dynamic Trajectories



Self-similarity Criticality



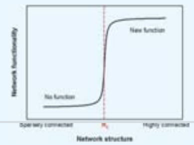
Small-world, Scale-free

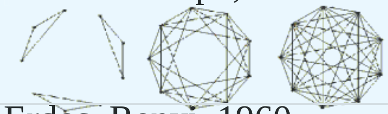


Watts, Strogatz, Nature, 1998

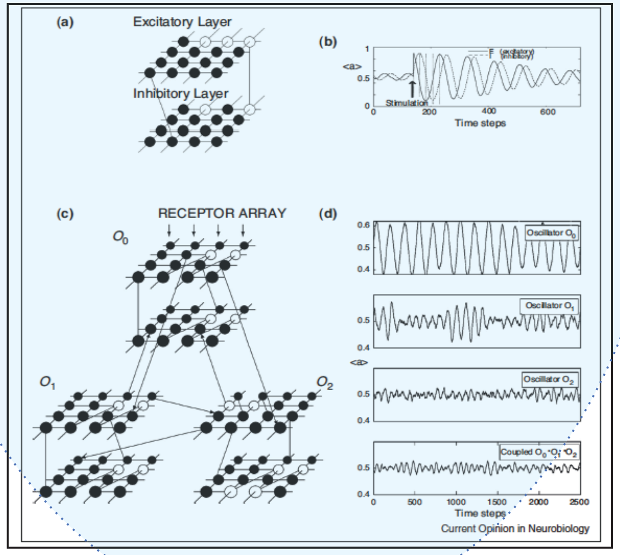
Albert, Barabasi Nature, 2000,

Random Graph, Phase trans.





Erdos, Renyi, 1960.

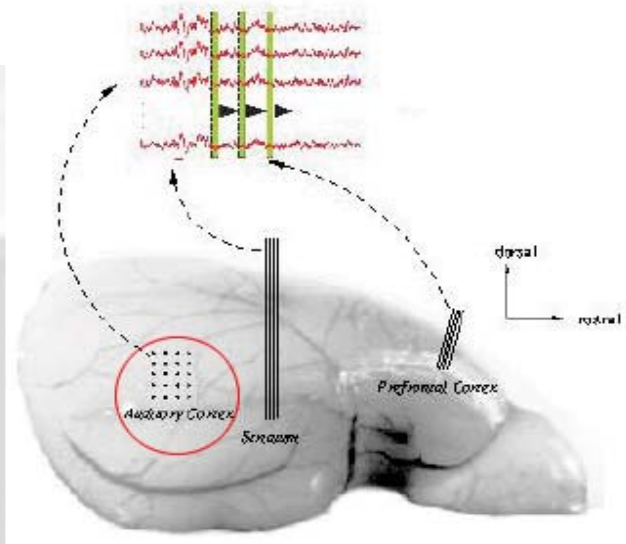


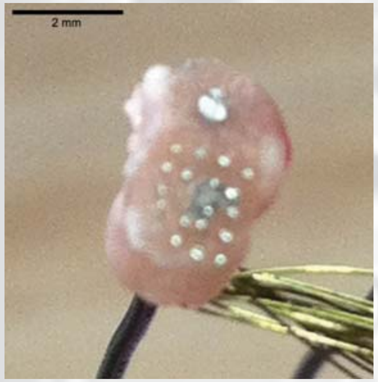
Balister, Bollobas. Kozma, RSA. 2006

Kozma Puljic, Curr.Op.Nbiol.2015

Rodent Model
Mongolian gerbil *(Meriones unguiculatus)*

Ohl, Scheich, Freeman, Nature (2001)

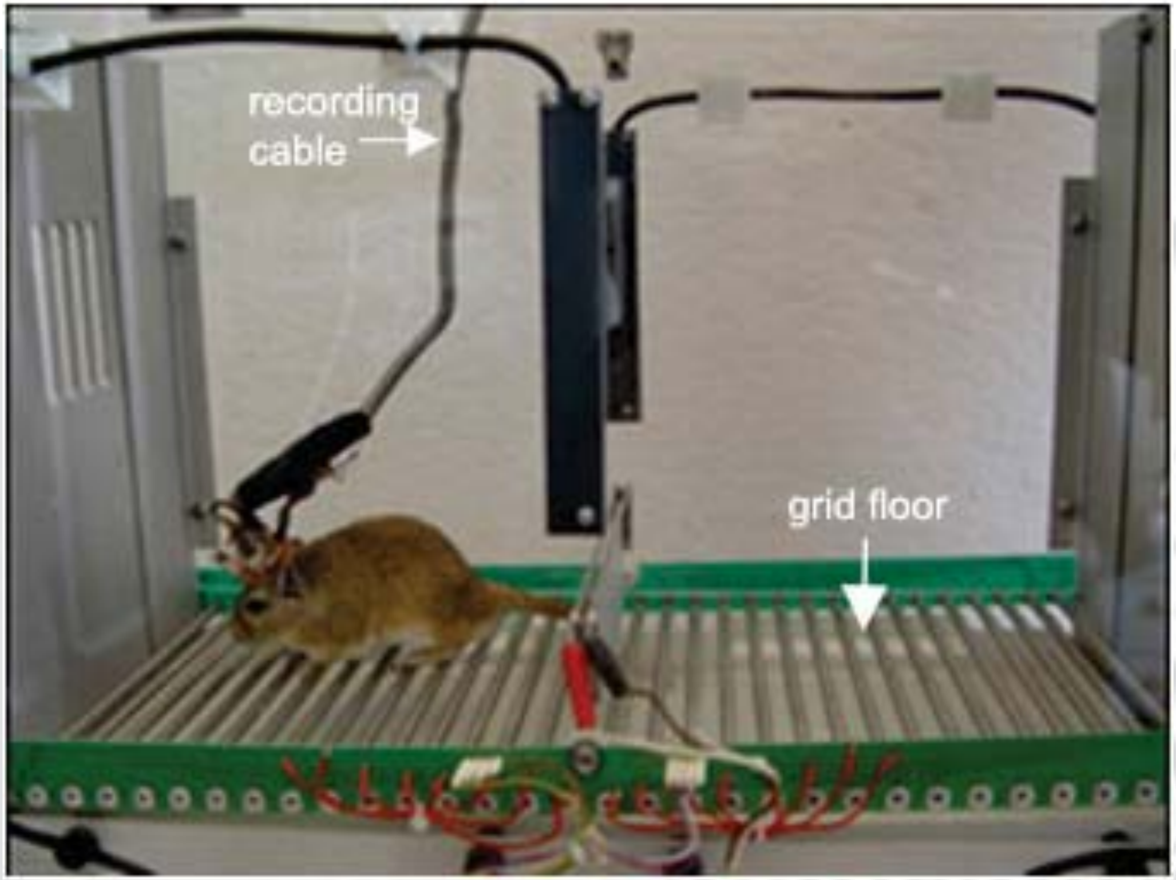




- Rectangular electrode array w/ 20 channels (5 x 4 matrix)
- Stainless steel wire, Ø 76.2 µm, impedance: 50 – 500 kΩ
- Chronic implantation of 5 x 4 electrode array on top of primary auditory cortex (A1)
- Recording of epidural potentials during behaviour

Learning Paradigm

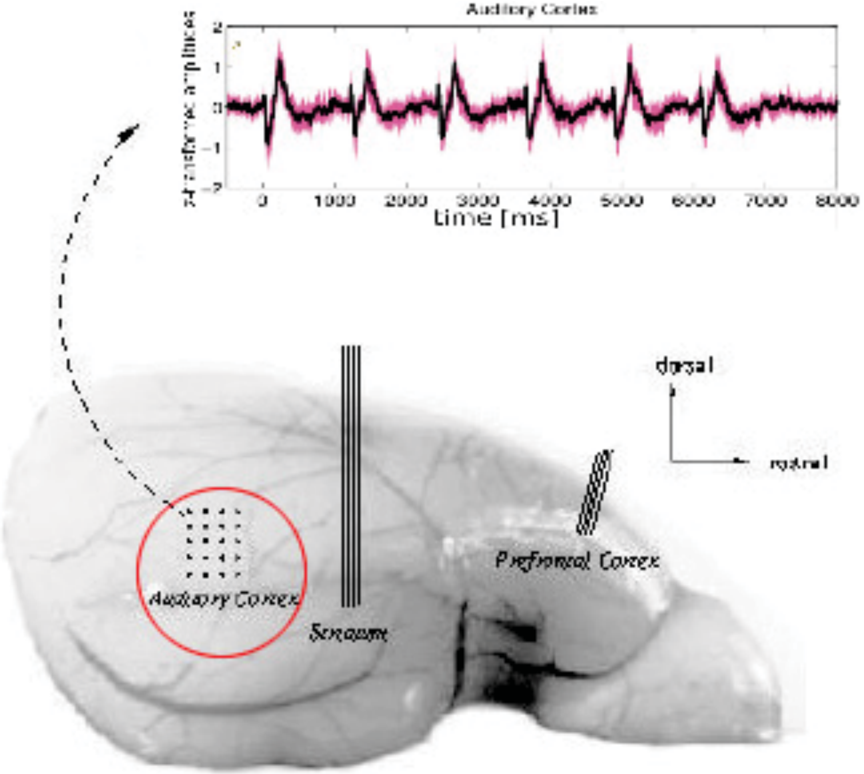
Discrimination of frequency modulated tones in the shuttle box

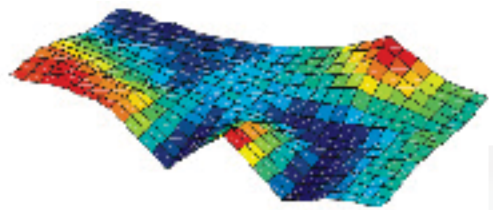


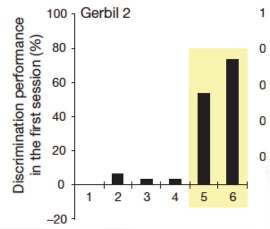
Go-tone:
sequence of ‘*rising*’ FM-tones (2–4 kHz, 200 ms duration)

NoGo-tone:
sequence of ‘*falling*’ FM-tones (4-2 kHz, 200 ms duration)

Analysis of Amplitude-Modulated (AM) Cortical Activity Patterns







Classifiable AM patterns in auditory cortex

Ohl, Scheich, Freeman, Nature (2001)

Research Goal of This US-German Project

- Understanding jump in behavior and interpret it as “Aha” moment
- Test hypothesis on “phase transition” in forming Hebbian cell assemblies
- Make predictions and design new experiments to test those predictions

Effective Interactions

Granger Causality vs New Causality (SQ Hu et al, 2011)

AR for the channels separately

$$\begin{cases} X_{1,t} = \sum_{j=1}^{m} a_{11,j} X_{1,t-j} + \epsilon_{1,t} \\ X_{2,t} = \sum_{j=1}^{m} a_{22,j} X_{1,t-j} + \epsilon_{2,t} \end{cases}$$

Joint AR model for channels X1 and X2

$$\begin{cases} X_{1,t} = \sum_{j=1}^{m} a_{11,j} X_{1,t-j} + \sum_{j=1}^{m} a_{12,j} X_{2,t-j} + \eta_{1,t} \\ X_{2,t} = \sum_{j=1}^{m} a_{21,j} X_{1,t-j} + \sum_{j=1}^{m} a_{22,j} X_{2,t-j} + \eta_{2,t} \end{cases}$$

Granger Causality (GC) is satisfied if

$$F_{X_2 \to X_1} = \ln \frac{\sigma^2_{\epsilon_1}}{\sigma^2_{\eta_1}} > 0 \quad \text{GC X2} \rightarrow \text{X1}$$

$$F_{X_1 \to X_2} = \ln \frac{\sigma^2_{\epsilon_2}}{\sigma^2_{\eta_2}} > 0 \quad \text{GC X1} \rightarrow \text{X2}$$

Weighted proportions in New causality

$$n_{X_2 \to X_1} = \frac{\sum_{t=m}^{N} \left(\sum_{j=1}^{m} a_{12,j} X_{2,t-j}\right)^2}{\sum_{h=1}^{2} \sum_{t=m}^{N} \left(\sum_{j=1}^{m} a_{1h,j} X_{h,t-j}\right)^2 + \sum_{t=m}^{N} \eta_{1,t}^2}$$

$$n_{X_1 \to X_2} = \frac{\sum_{t=m}^{N} \left(\sum_{j=1}^{m} a_{21,j} X_{1,t-j}\right)^2}{\sum_{h=1}^{2} \sum_{t=m}^{N} \left(\sum_{j=1}^{m} a_{2h,j} X_{h,t-j}\right)^2 + \sum_{t=m}^{N} \eta_{2,t}^2}$$

Causality is determined when one of these weighted proportions are larger for stimulus A and smaller for B (= no_A)

New Causality (NC) is satisfied for (X1,X2) if

$$\boxed{(n^A_{X1 \to X2} - n^A_{X2 \to X1}) \times (n^B_{X1 \to X2} - n^B_{X2 \to X1}) < 0}$$

Gerbil #1
Results of Learning and Collected/Core Pairs

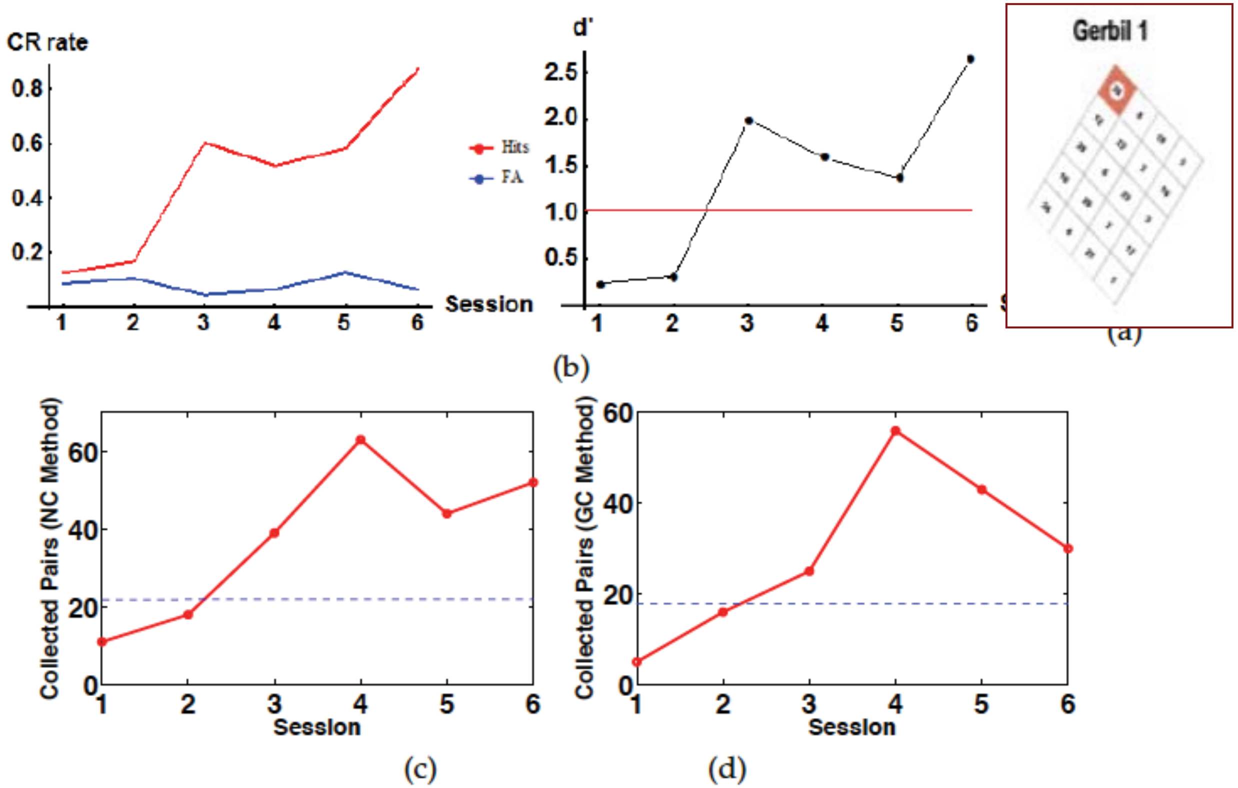


Figure 1: Analysis of data over 6 training sessions for one representative animal (Gerbil 1). (a) and (b): Conditioned response (CR) rate and detection sensitivity criterion d' per session; from Session 3 on, the gerbil can discriminate the two categories and performed above chance level ($d' > 1$) for the rest of the sessions. (c) and (d): Number of collected pairs using inverse causality determined by NC and GC methods in each session. The dashed line is drawn to guide the eye to separate Stage 1 (Sessions 1 and 2) and Stage 2 (Sessions 3 to 6).

Total Collected Pairs With Inverse Causality

Table 1: Number of collected pairs using NC and GC methods in Gerbil 1

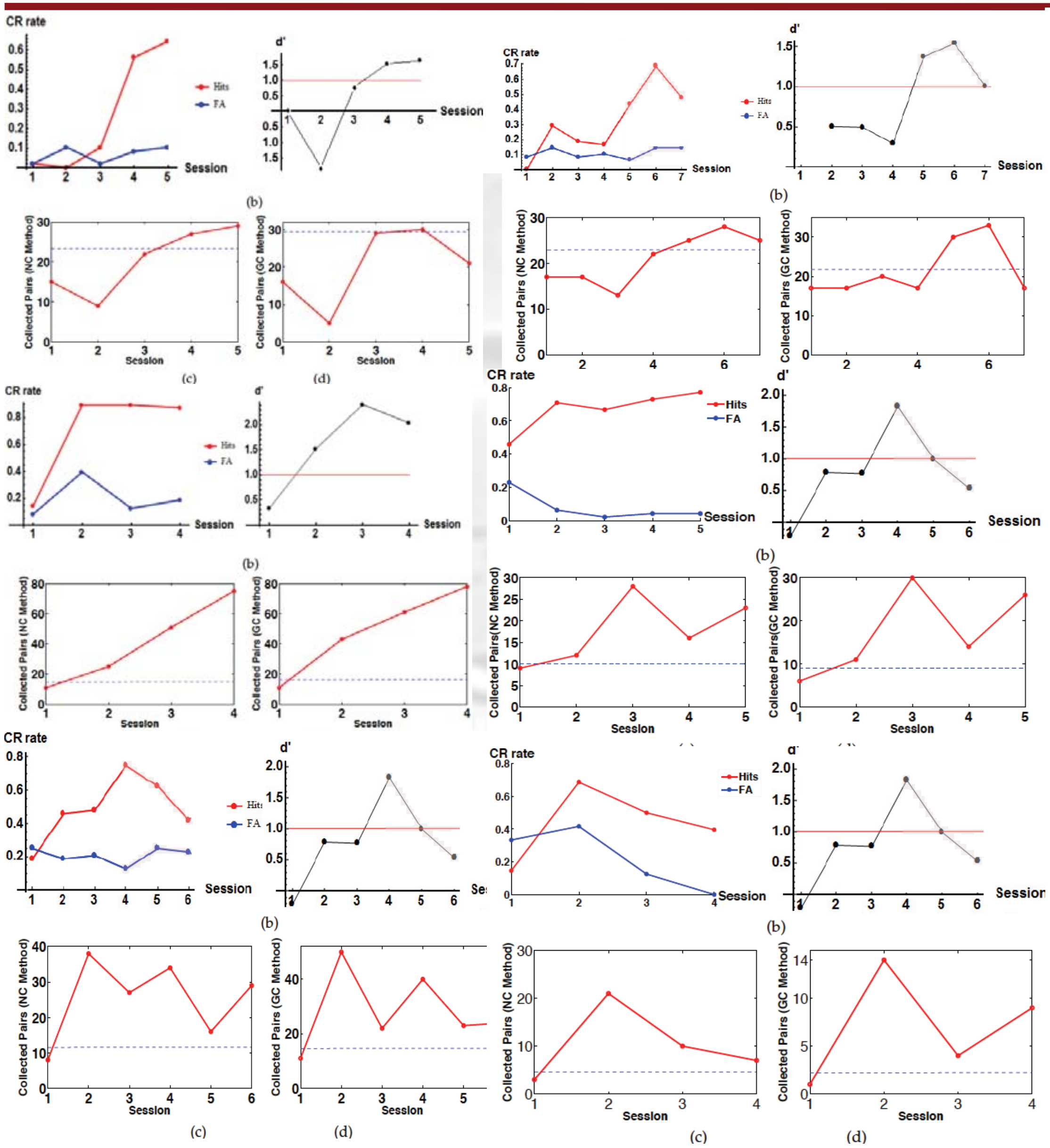
	Stage 1		Stage 2			
	S1	S2	S3	S4	S5	S6
NC	11	18	39	63	44	52
GC	5	16	25	56	43	30
Mean of stage (NC)	14.5		49.5			
Mean of stage (GC)	10.5		38.5			

Classification using Core Pairs only Appear at Stage 2

Table 4: Classification rates for the 6 core pairs obtained by NC method for Gerbil 1

	Stage 1 Mean of S1 and S2 (%)	Stage 2 Mean of S3 to S6 (%)
(C21,C23)	48.10	66.75
(C3,C18)	56.75	65.15
(C17,C18)	54.65	65.15
(C17,C22)	55.10	66.68
(C17,C23)	51.35	64.88
(C21,C22)	57.20	67.73

Group analysis of 8 gerbils



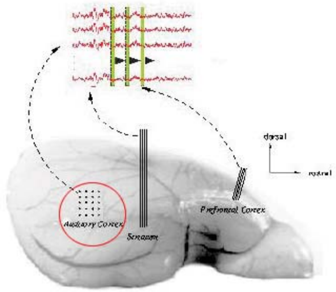
Total pairs at stages 1 & 2

Gerbil #	Method NC or GC	Session S1	S2	S3	S4	S5	S6	S7
1	NC	11	18	39	63	44	52	
	GC	5	16	25	56	43	30	
2	NC	15	9	22	27	29		
	GC	16	5	29	30	21		
3	NC	11	25	51	75			
	GC	11	43	61	78			
4	NC	17	17	13	22	25	28	25
	GC	17	17	20	17	30	33	17
5	NC	8	38	27	34	16	29	
	GC	11	50	22	40	23	24	
6	NC	9	12	28	16	23		
	GC	6	11	30	14	26		
7	NC	3	21	10	7			
	GC	1	14	4	9			
8	NC	5	14	8	9	6		
	GC	7	17	7	2	4		

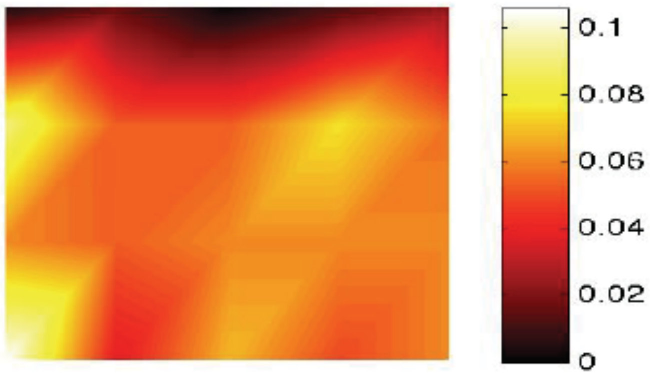
Significance test of increased # of pairs in transition from naïve to discriminating stage

Gerbil #	p_{NC}	p_{GC}
1	0.007	0.040
2	0.050	0.248
3	0.038	0.012
4	0.008	0.070
5	0.002	0.009
6	0.009	0.004
7	0.053	0.038
8	0.031	0.434

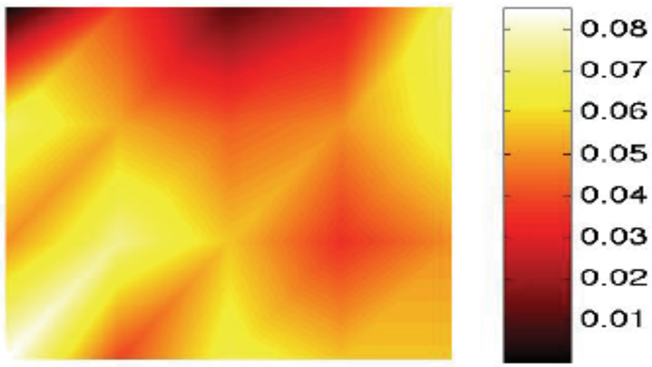
Spatial distribution of pairs
- More clustered with hot spots after learning -



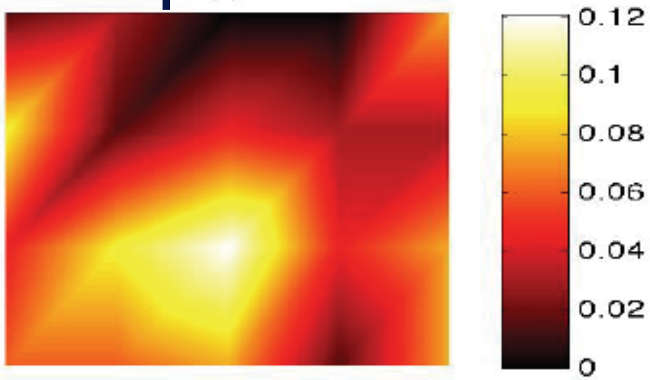
NC Stage 1
Collected pairs



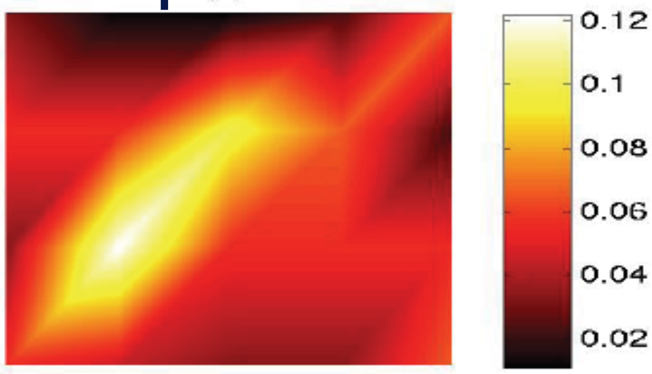
GC Stage 1
Collected pairs



NC Stage 2
Core pairs



GC Stage 2
Core pairs



Modeling the Cortical Sheet
Neuropercolation

Structure

Neuropil:

Densely connected filamentous texture of neural tissue

NxN Lattice/Torus

Basic steady-state Oscillations induced by:
- Process noise
- (Rare) non-local link
- 2 layers: excitatory-inhibitory; minicolumns
- Multilayer cortex model

Function/Dynamics

Pulse density in neuropil

- Subthreshold regime
- Possible phase transitions from connectivity/gain change

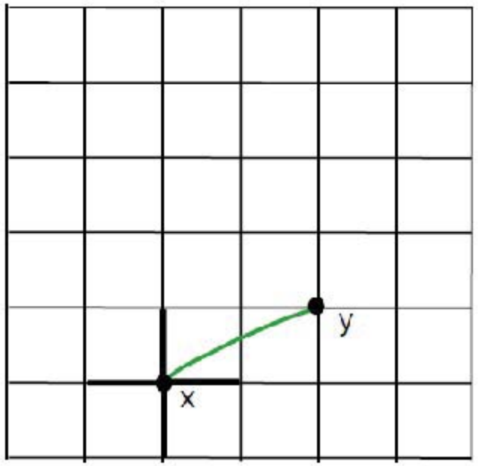
Percolation:

- Random initial → propagate
- Automata rule:
- Characterize asymptotics: zero f.p. → non-zero f.p. → periodic → chaos

Erdos, Renyi, 1960;
Bollobas, Riordan, 2002.

Random Graph Model of Lattice with few added edges (Jansen et al., 2016)

Start first with $\mathbb{Z}^2$ lattice. We take a $(N+1)\times(N+1)$ grid and for simplicity assume periodic boundary condition. Thus, we have a torus $\mathbb{T}^2 = (\mathbb{Z}/N\mathbb{Z})^2$, with the short notation $\mathbb{Z}_N^2$. The set of vertices of G consists of all the vertices of $\mathbb{Z}_N^2$ and we will not change them. All the edges from the torus $\mathbb{Z}_N^2$ are presented in the graph.



Additionally, we introduce the random edges. For any pair of vertices that are at distance d apart of each other we assign probability of edge that depends on the graph distance d:

$$p_d = \mathbb{P}\Big((x,y) \in E(G_{\mathbb{Z}_N^2,p}) \text{ and } dist(x,y) = d\Big) = \frac{c}{Nd}$$

Degree Sequence of the Introduced Random Graph

The expected number of long edges $E_\ell \subseteq E(G_{\mathbb{Z}_N^2, p_d})$ is proportional to N^2, i.e.,

$$\mathbb{E}(|E_\ell|) \sim (2c \ln 2) N^2,$$

where c is the parameter of p_d. The probability that a vertex has degree k considering only the long edges is given by

$$\mathbb{P}(W = k) = \sum_{k_2 + \ldots + k_N = k} \prod_{i=2}^{N} \binom{\Lambda_i}{k_i} \left(\frac{c}{Ni}\right)^{k_i} \left(1 - \frac{c}{Ni}\right)^{\Lambda_i - k_i}.$$

However, we can approximate it by Poisson Po(λ) distribution with $\lambda = 4c \ln 2$, that is

$$\mathbb{P}(W = k) = \frac{e^{-\lambda} \lambda^k}{k!} + O\left(\frac{1}{N}\right).$$

Diameter of the Coupled Lattice-Random Graph

Observation. Signals in the brain spread fast, hence, the diameter of any graph modeling the brain has to be small. The diameter $D(G_{\mathbb{Z}_N^2,p})$ is logarithmic in the number of vertices, i.e., it has small world property.

Theorem (JKRS'15)

There exist constants C_1, C_2, which depend on c only, such that for the diameter $D(G_{\mathbb{Z}_N^2,p})$ the following holds

$$\lim_{N\to\infty} P\left(C_1 \log N \leq D(G_{\mathbb{Z}_N^2,p}) \leq C_2 \log N\right) = 1.$$

E.g., for a network of 10^9 nodes (the size of gerbil brain) the diameter is ≈ 9.

Activation process: One type of vertices

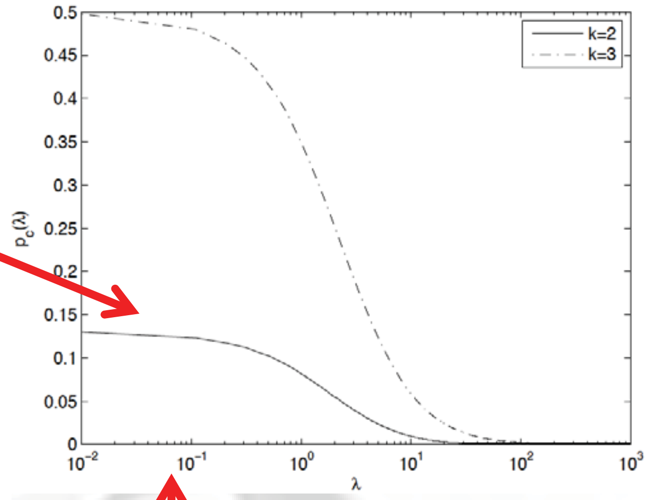
Denote by $A(t)$ the set of all active vertices at time t. We say that a vertex v is active at time t if its potential function $\chi_v(t) = 1$ and inactive if $\chi_v(t) = 0$. Therefore, $A(t) = \{v \in V(G) \mid \chi_v(t) = 1\}$. At the beginning, $A(0)$ consists of all vertices that are active with probability p_0. Each vertex may change its potential based on the states of its neighbors in the following way

$$\chi_v(t+1) = \mathbb{1}\left(\sum_{u \in N(v)} \chi_u(t) \geq k\right) \tag{6}$$

The process resembles bootstrap percolation. However, in our case active vertex can become inactive if the total activation of the vertex at some moment is less than k.

Phase Transition on the Mean Field Graph

Initial activation needed to self-sustained activity



Additional (non-local) edges

Theorem (JKRS'15)

In the mean-field approximation of the activation process $A(t)$ over random graph $G_{\mathbb{Z}_N^2, p_d}$ there exists a critical probability p_c such that for a fixed p, w.h.p., all vertices will eventually be active if $p > p_c$, while all vertices will eventually be inactive for $p < p_c$. The value of p_c is given as the function of k and λ as follows:

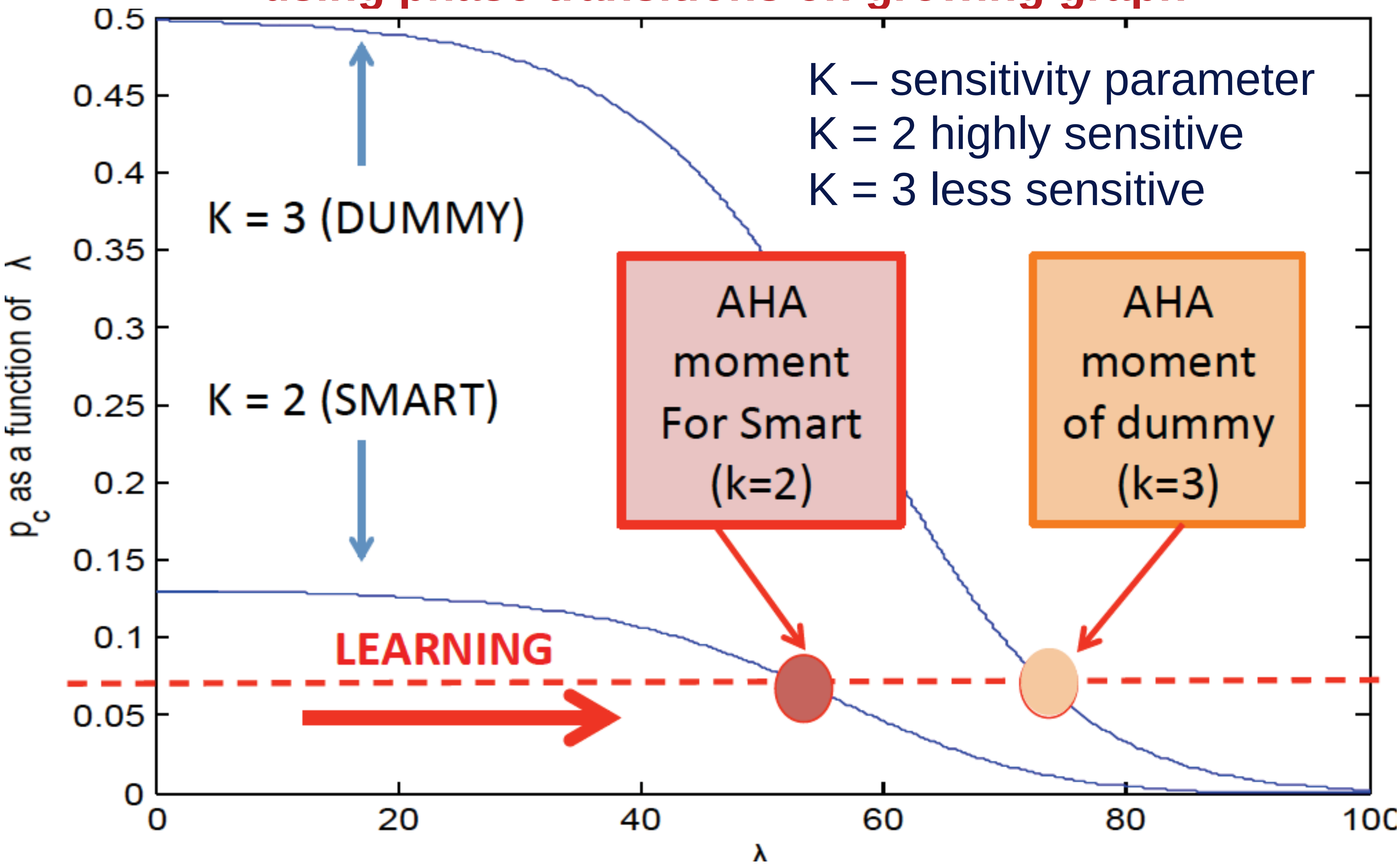
(i) *For $k = 0$ and any λ all vertices will become active in one step with high probability for any fixed $p > 0$.*

(ii) *For $k = 1$ and any λ, $p_c = 0$, i.e., for any fixed $p > 0$, all vertices will eventually become active with high probability.*

(iii) *For $k = 2$ and any λ, $p_c = x_2(\lambda)$, where $x_2(\lambda) < 0.132$ is a nontrivial solution to $x = \bar{f}_2(x)$.*

(iv) *For $k = 3$ and any λ, $p_c = x_3(\lambda)$, where $x_3(\lambda) < 0.5$ is a nontrivial solution to $x = \bar{f}_3(x)$.*

Step 1: Possible interpretation of gerbil learning using phase transitions on growing graph



Activation process with 2 types of vertices (E/I)

Update rule of Excitatory (E) node

$$\chi_v(t+1) = \mathbb{1}\left(\sum_{u \in N^E(v)} \chi_u(t) - \sum_{u \in N^I(v)} \chi_u(t) \geq k\right)$$

Update for inhibitory (I) node

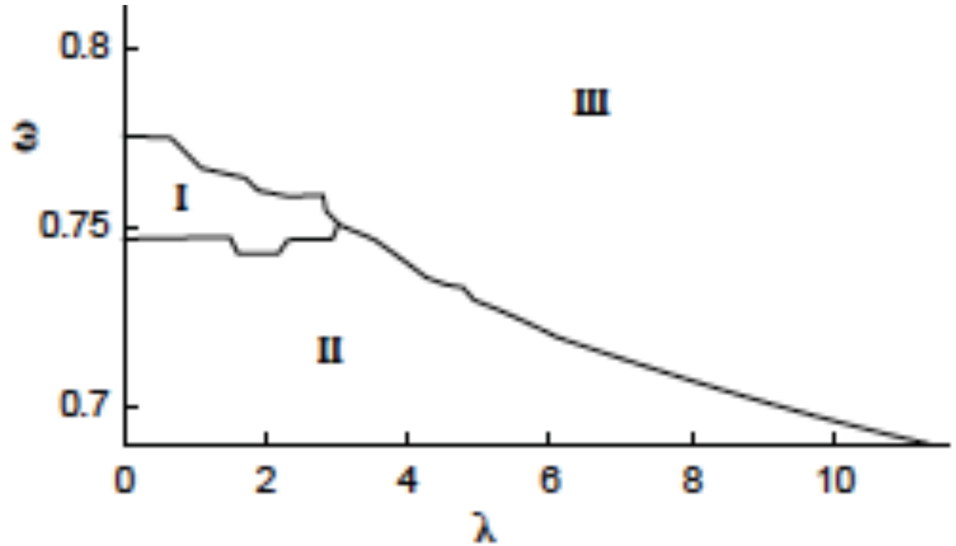
$$\chi_v(t+1) = \mathbb{1}\left(\sum_{u \in N^E(v)} \chi_u(t) + \sum_{u \in N^I(v)} \chi_u(t) \geq k\right) = \mathbb{1}\left(\sum_{u \in N(v)} \chi_u(t) \geq k\right),$$

Asymptotic regimes:

✧ I: Limit cycle

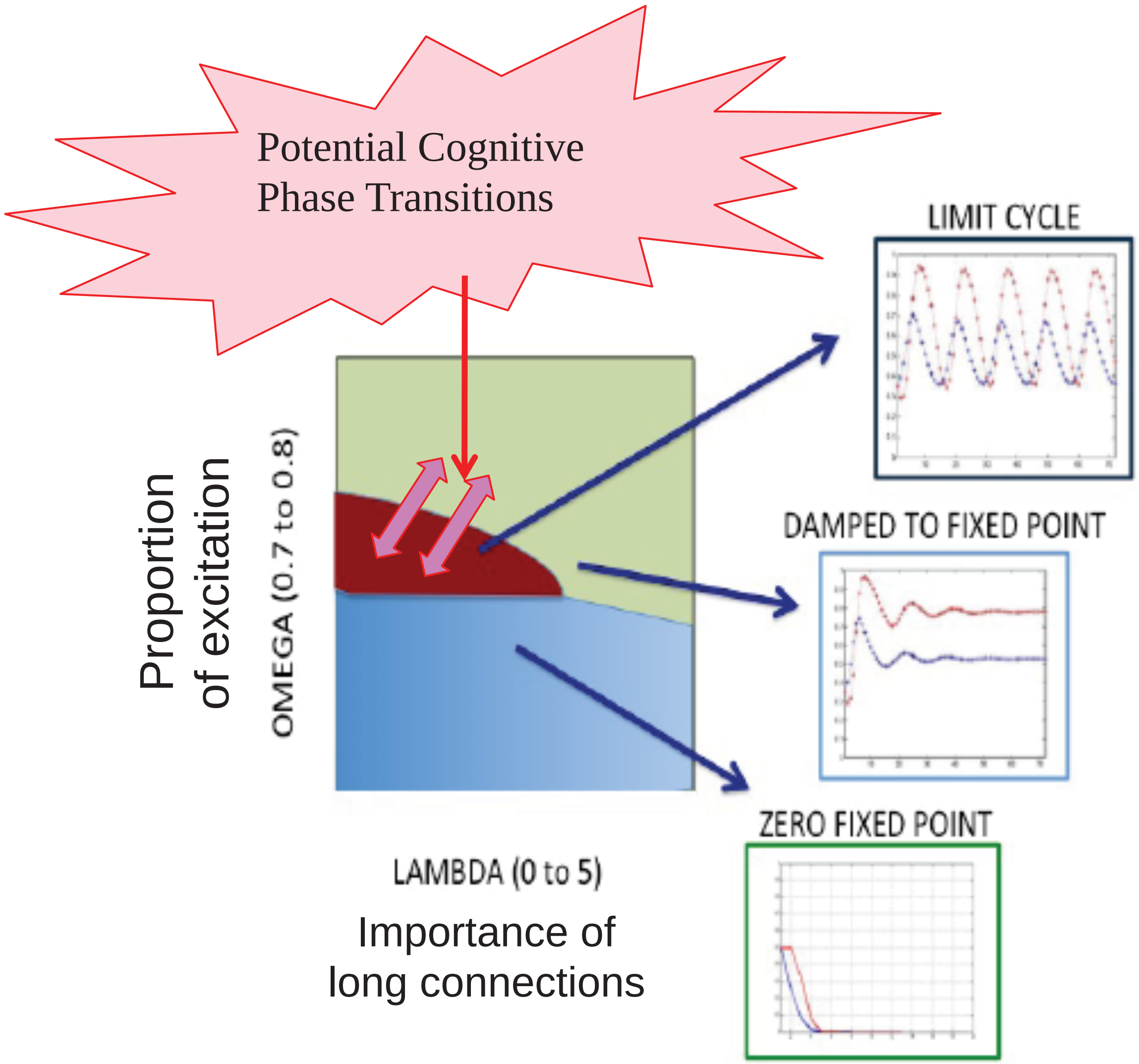
✧ II: Zero fixed point

✧ III: Nonzero fixed point



- Omega is proportion of E nodes
- Phase diagram for k1=2 & k2=3
- NB: dominant attractors are indicated

Onset of limit cycle in response to learning



- Dynamic regimes illustrated over the phase diagram in the lambda versus omega space.
- The solid dark red region correspond to conditions with limit cycle oscillations created as the result of learning.
- The blue area corresponds to conditions when collective activity dies out.
- The region corresponding to the hypothetic cognitive phase transitions is marked by pink arrows.
- Transitions from limit cycle regime toward fixed points indicate the absence of learnt stimulus.

Summary

1. Studied the evolution of the AC connectivity during learning using a generalized New Causality (NC) metric.
2. The results are indicative of emerging causal links as learning evolves.
3. Interpretation in the context of Hebbian learning and the formation of Hebbian cell assemblies (HCA).
4. Graph theory model is developed to describe phase in response to learning (narrow band oscillations)
5. Testable hypotheses for experimentations including plasticity and sensitivity control for gerbil learning experiments.

References

- F. W. Ohl, H. Scheich, and W. J. Freeman, "Change in pattern of ongoing cortical activity with auditory category learning," Nature, vol.412, pp. 733-736, 2001.
- F. W. Ohl, "Role of cortical neurodynamics for understanding the neural basis of motivated behavior - lessons from auditory category learning," Curr. Opin. Neurobiol., vol. 31: pp. 88-94, 2015.
- C. W. J. Granger, "Investigating Causal Relations by Econometric Models and Cross-spectral Methods", Econometrica, vol.37, no.4, pp. 424-438, 1969.
- M. Ding, Y. Chen, and S. L. Bressler, "Granger causality: Basic theory and applications to neuroscience," Handbook of Time Series Analysis, B. Schelter, M. Winterhalder, and J. Timmer Eds. Weinheim, Germany: Wiley-VCH, pp. 437-460, 2006.
- S. Hu, G. Dai, G. Worrell, Q. Dai, and H. Liang, "Causality analysis of neural connectivity: Critical examination of existing methods and advances of new methods," IEEE Trans Neural Networks, vol.22, no.6, pp. 829--844, 2011.
- Hu, S., Wang, H., Zhang, J., Kong, W., Cao, Y., Kozma, R. (2016) "Comparison Analysis: Granger Causality and New Causality and Their Applications to Motor Imagery," *IEEE Trans. Neur. Netw. & Learning Syst.*, 27(7), pp. 1429 – 1444.
- Jansen, S., R. Kozma, R. Miklos, Y. Sokolov (2016) "Bootstrap percolation on a random graph coupled with a lattice," (2016, submitted), Arxiv: https://arxiv.org/abs/1507.07997v2
- Kozma, R., Y. Sokolov, M. Puljic, M. Ruszinko, S.Q. Hu (2016) "Modeling Learning and Strategy Formation as Phase Transitions in Cortical Networks," IEEE Systems, Man, and Cybernetics Conf., October 9-12, 2016, Budapest, Hungary.

Credits

Collaborators:

Yury Sokolov, University of Pittsburgh

Walter J. Freeman, UC Berkeley

Miklos Ruszinko
Renyi Institute of Mathematics,
Hungarian Academy, Budapest

Sanqing Hu
Hangzhou Dianzi University,
Hangzhou, China

Frank Ohl, Andreas Schultz, Tim Wanger, Leibniz Institute for Neuroscience, Magdeburg, Germany

Bundesministerium
für Bildung
und Forschung