Unsupervised Learning in recurrent networks

Jochen Triesch
Frankfurt Institute for Advanced Studies

FIAS Frankfurt Institute for Advanced Studies

http://fias.uni-frankfurt.de

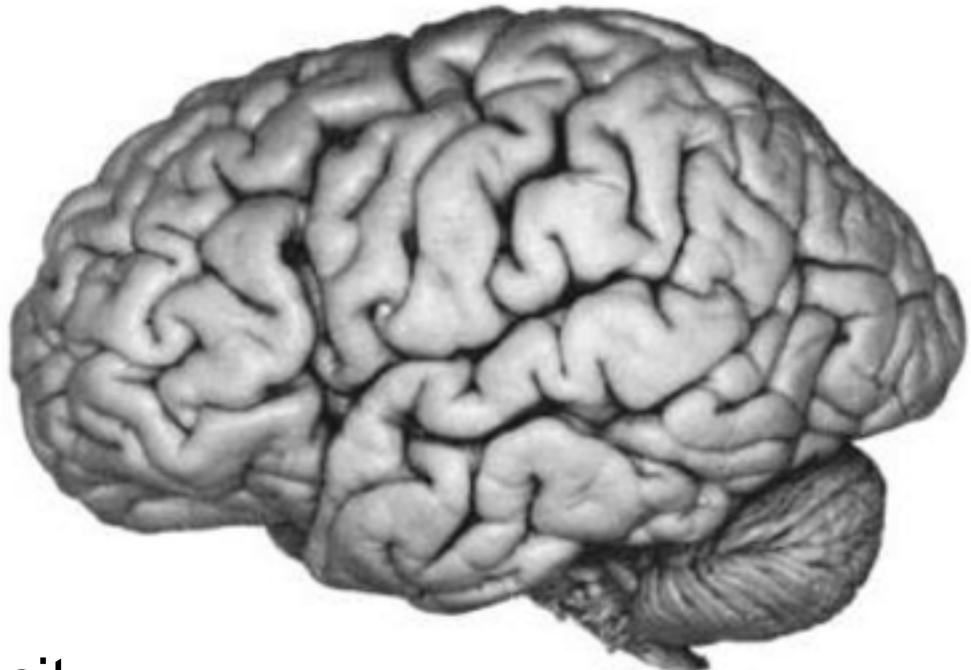
- founded in 2004, own building since 2007
- ~140 people
- interdisciplinary theoretical research in and across physics, chemistry, biology, neuroscience, and computer science
- Frankfurt International Graduate School for Science

spike-timing dependent plasticity

long-term depression
long-term potentiation

short-term facilitation
short-term depression

synaptic scaling



neuronal
adaptation

intrinsic plasticity

neuromodulation

structural plasticity

neurogenesis

spike-timing dependent plasticity

long-term depression
long-term potentiation

short-term facilitation
short-term depression

synaptic scaling

neuronal
adaptation

intrinsic plasticity

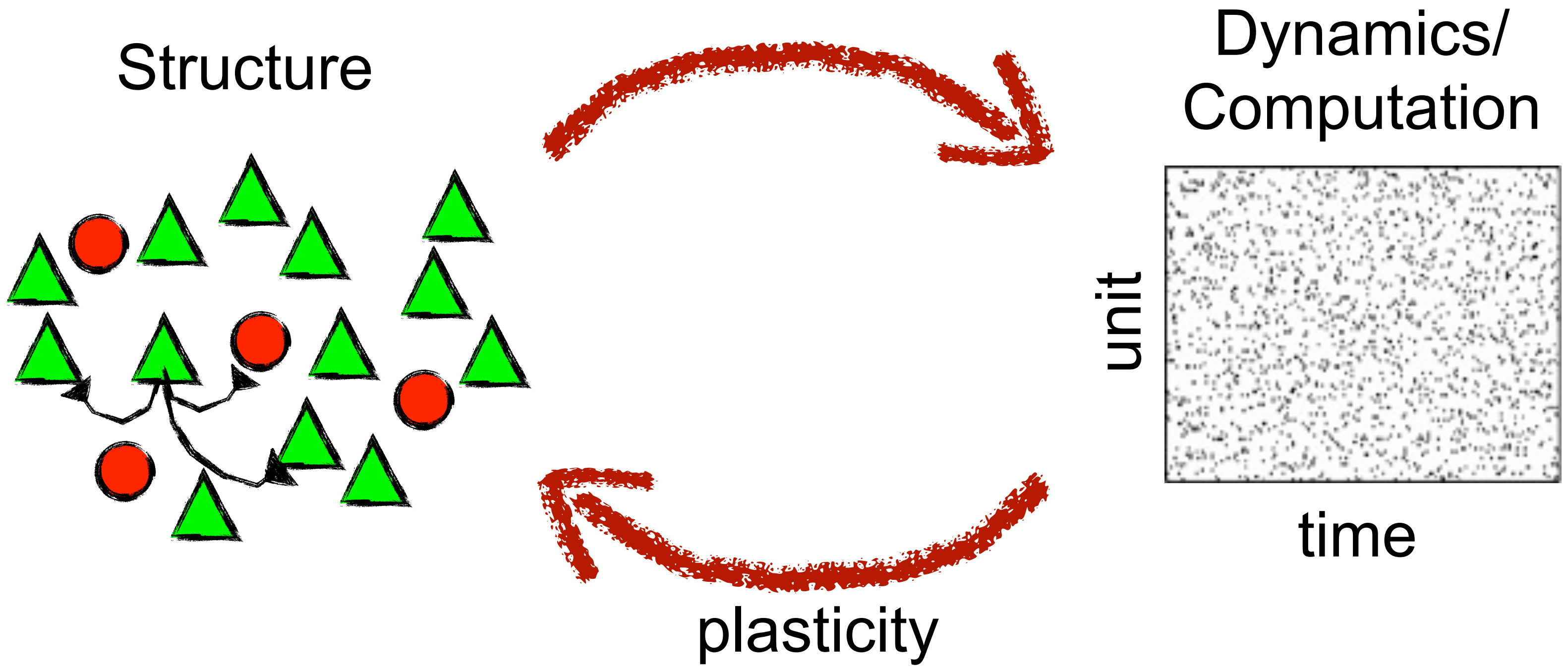
neuromodulation

structural plasticity

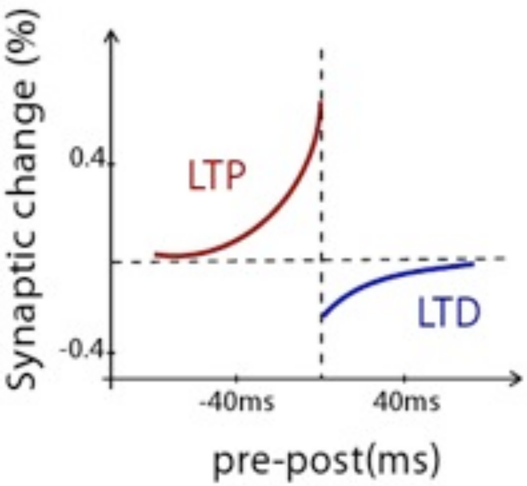
neurogenesis

how do they interact?
how do they shape neural circuits?
how do they shape neural codes?

Network Self-Organiztion

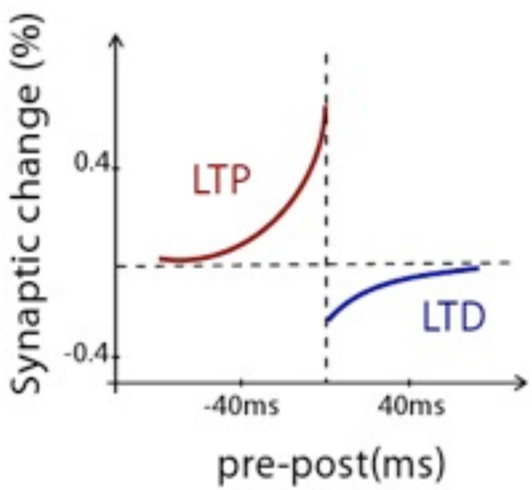


spike-timing dependent plasticity



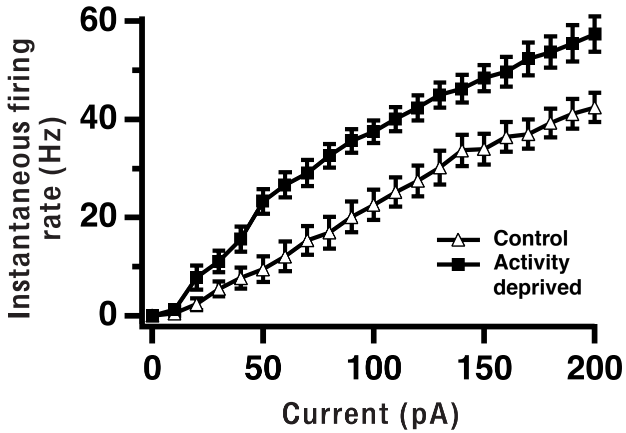
Markram et al. 1997
Bi&Poo 1998
...

spike-timing dependent plasticity



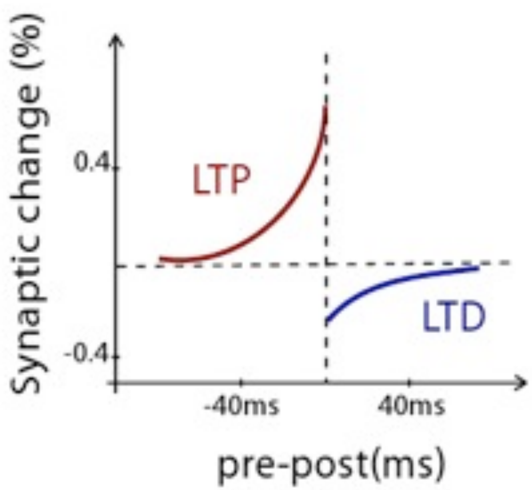
Markram et al. 1997
Bi&Poo 1998
...

intrinsic plasticity



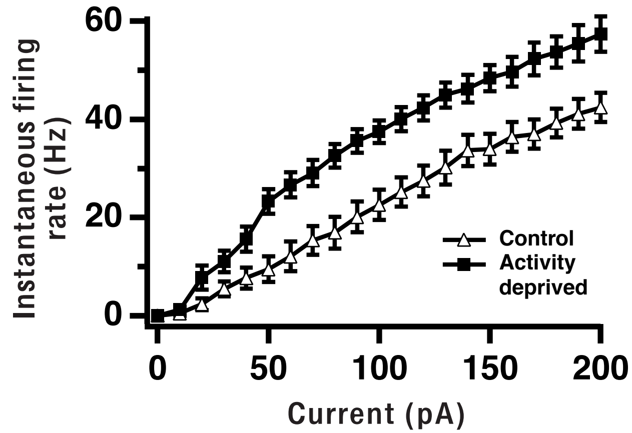
Desai et al. 1999
Zhang&Linden 2003
...

spike-timing dependent plasticity



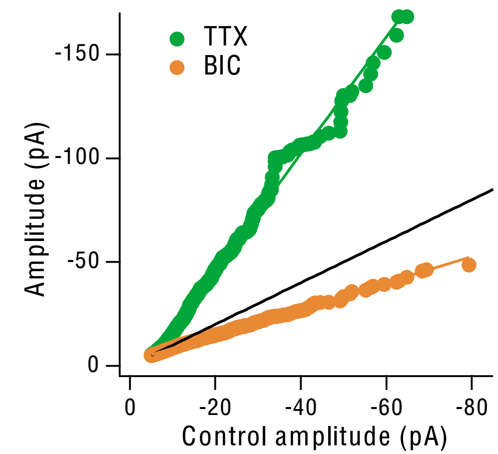
Markram et al. 1997
Bi&Poo 1998
...

intrinsic plasticity



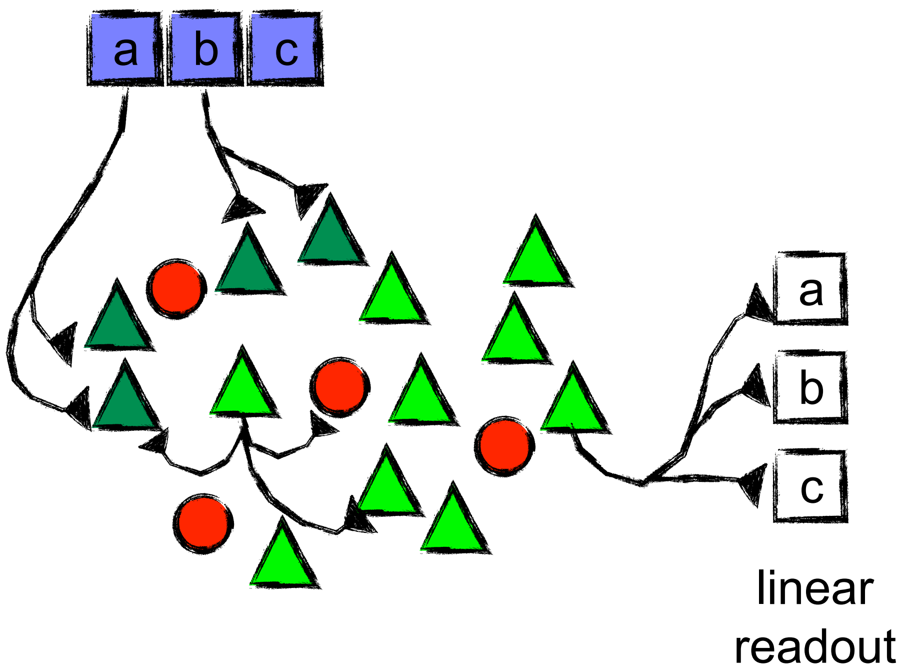
Desai et al. 1999
Zhang&Linden 2003
...

synaptic scaling

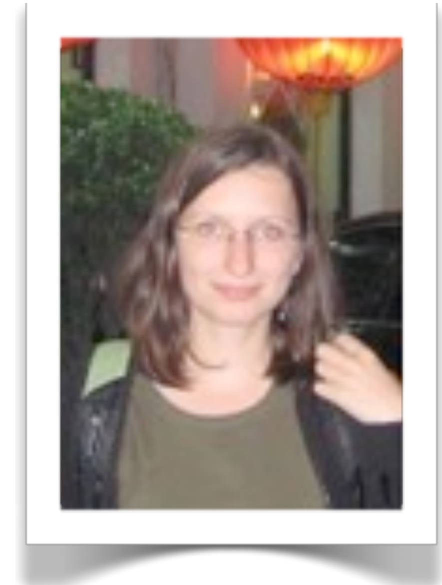


Turrigiano et al. 1998
Abbott&Nelson 2000
...

SORN: self-organizing recurrent neural network

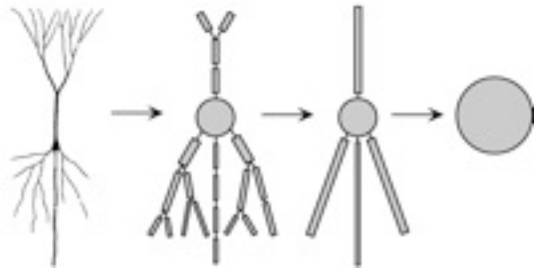


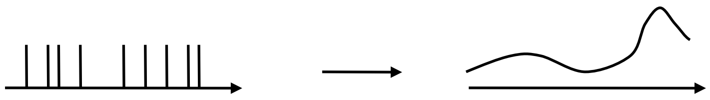
Lazar, Pipa, Triesch (2009) Frontiers in Computational Neuroscience 3:23

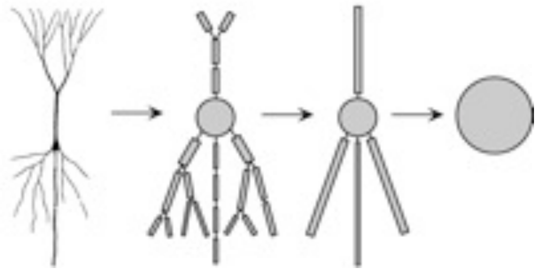


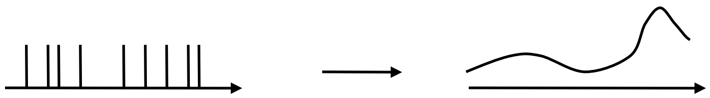
Andreea Lazar

Gordon Pipa

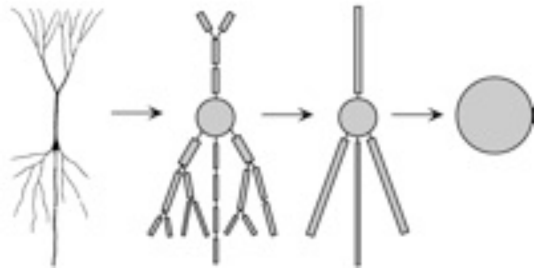


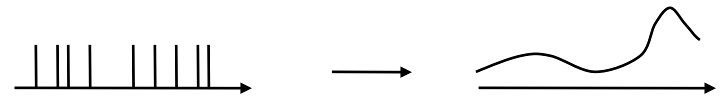






- Populations of coupled excitatory and inhibitory threshold units





- Populations of coupled excitatory and inhibitory threshold units

$$x_i(t+1) = \Theta\left(\sum_{j=1}^{N^E} W_{ij}^{EE}(t)x_j(t) - \sum_{k=1}^{N^I} W_{ik}^{EI} y_k(t) + \nu_i^U(t) - T_i^E(t)\right)$$

excitation

inhibition

a external input

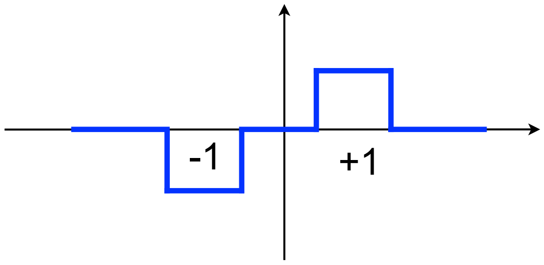
adaptive threshold

$$y_i(t+1) = \Theta\left(\sum_{j=1}^{N^E} W_{ij}^{IE} x_j(t) - T_i^I\right)$$

Plasticity Mechanisms

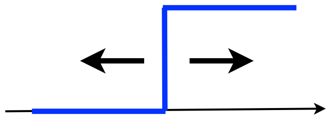
- Binarized STDP:

$$\Delta W_{ij}^{EE}(t) = \eta_{\mathrm{STDP}} \left(x_i(t) x_j(t-1) - x_i(t-1) x_j(t) \right)$$



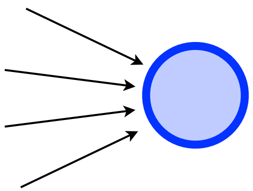
- Intrinsic plasticity:

$$T_i^E(t+1) = T_i^E(t) + \eta_{\mathrm{IP}} \left(x_i(t) - H_{\mathrm{IP}} \right)$$



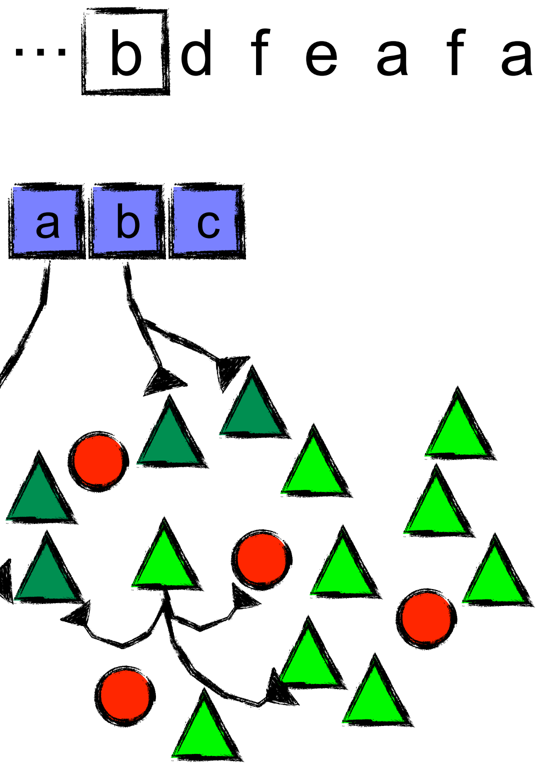
- Synaptic scaling:

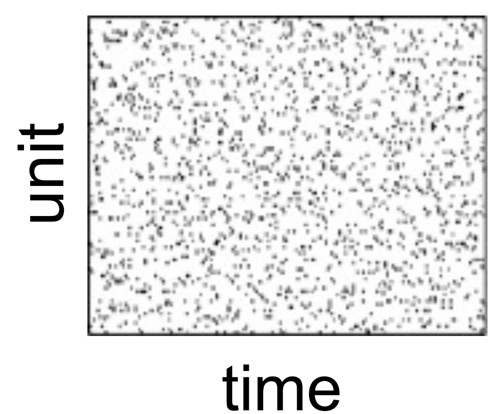
$$W_{ij}^{EE}(t) \leftarrow W_{ij}^{EE}(t) / \sum_j W_{ij}^{EE}(t)$$



A first test:

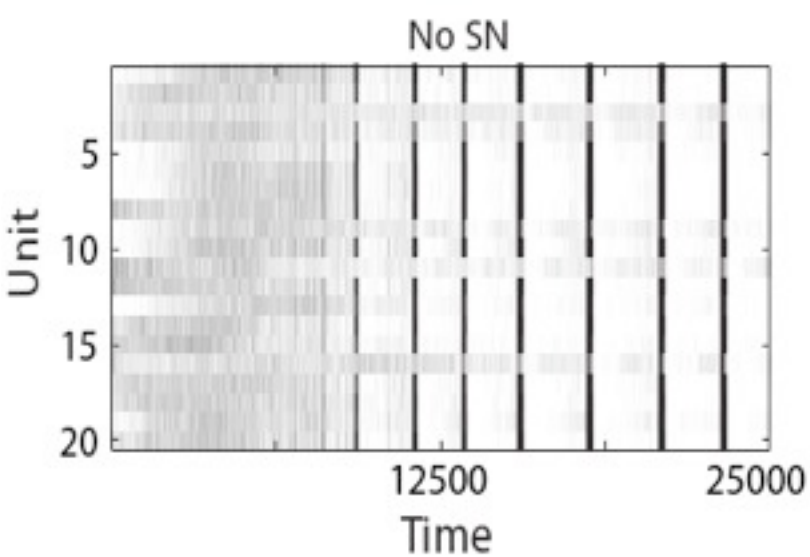
- input: six symbols (a-f) alternating randomly
- activity after self-organization is sparse and irregular

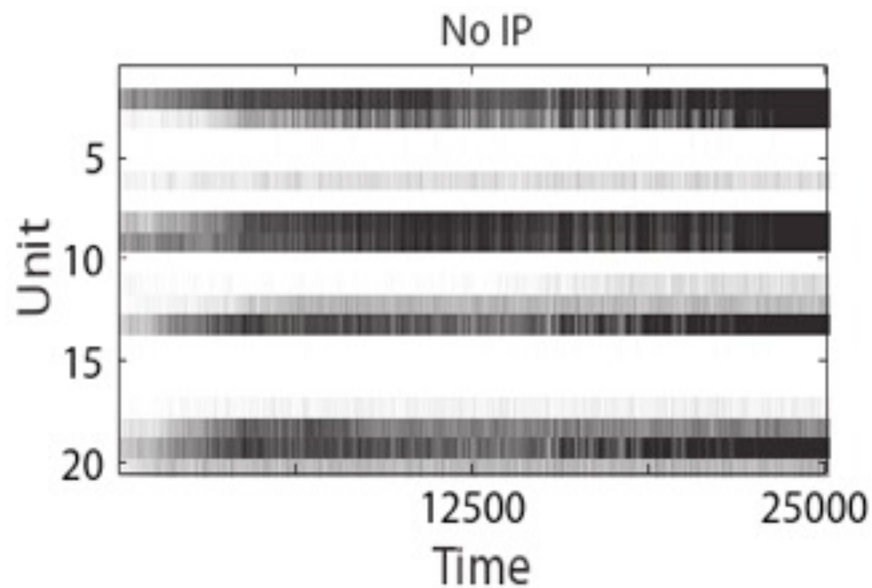




Homeostasis is important!

- abolishing intrinsic plasticity or synaptic normalization:



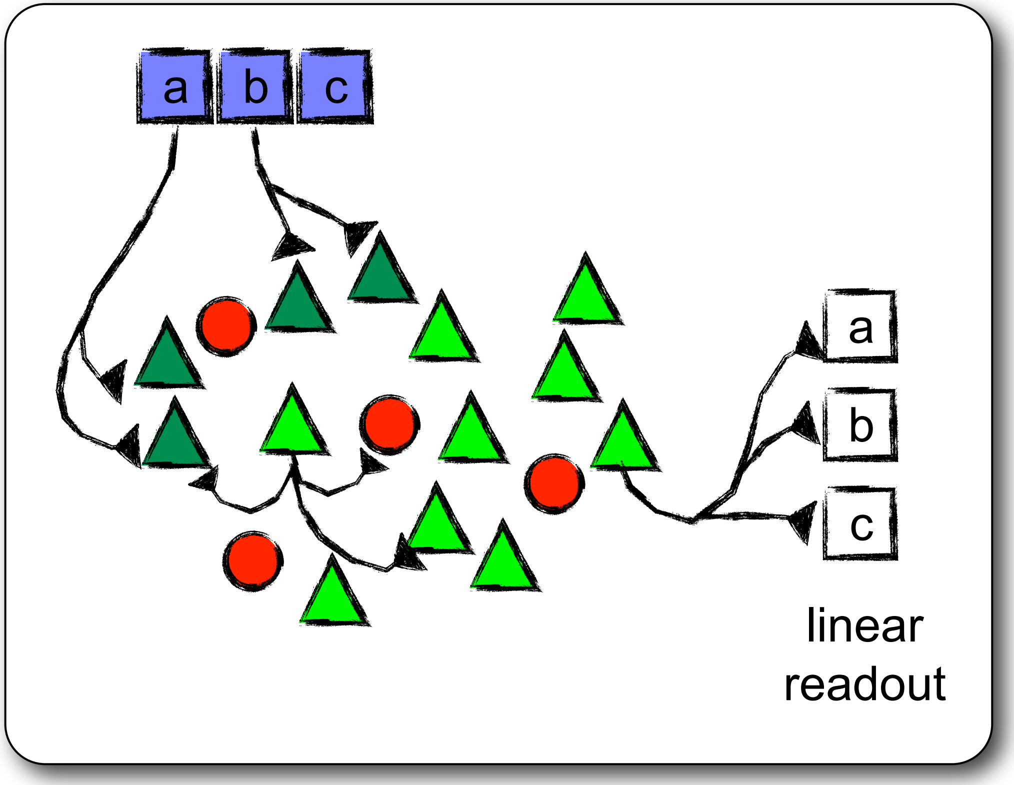


Without intrinsic plasticity or synaptic normalization pathological activity patterns develop!

Counting Task

- Input sequences are random alternations of two words:
  - abb....bbbc
  - edd....dddf

  n repetitions of middle letter
- training in two phases:
  - reservoir
  - readout



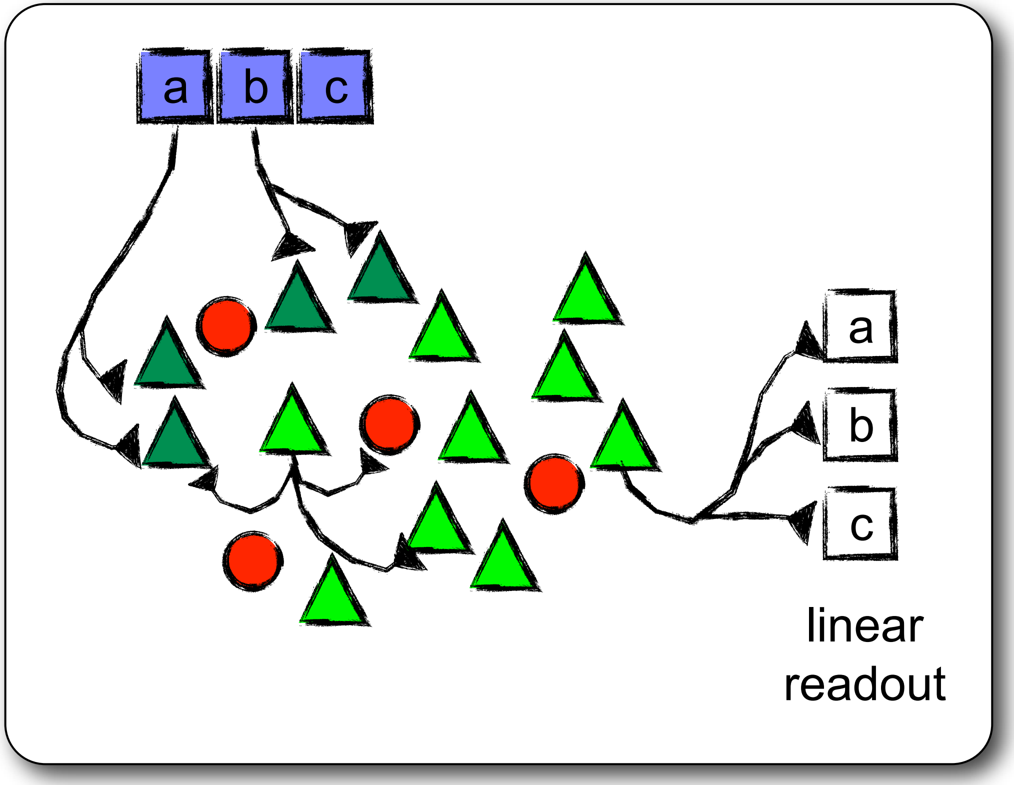
Counting Task

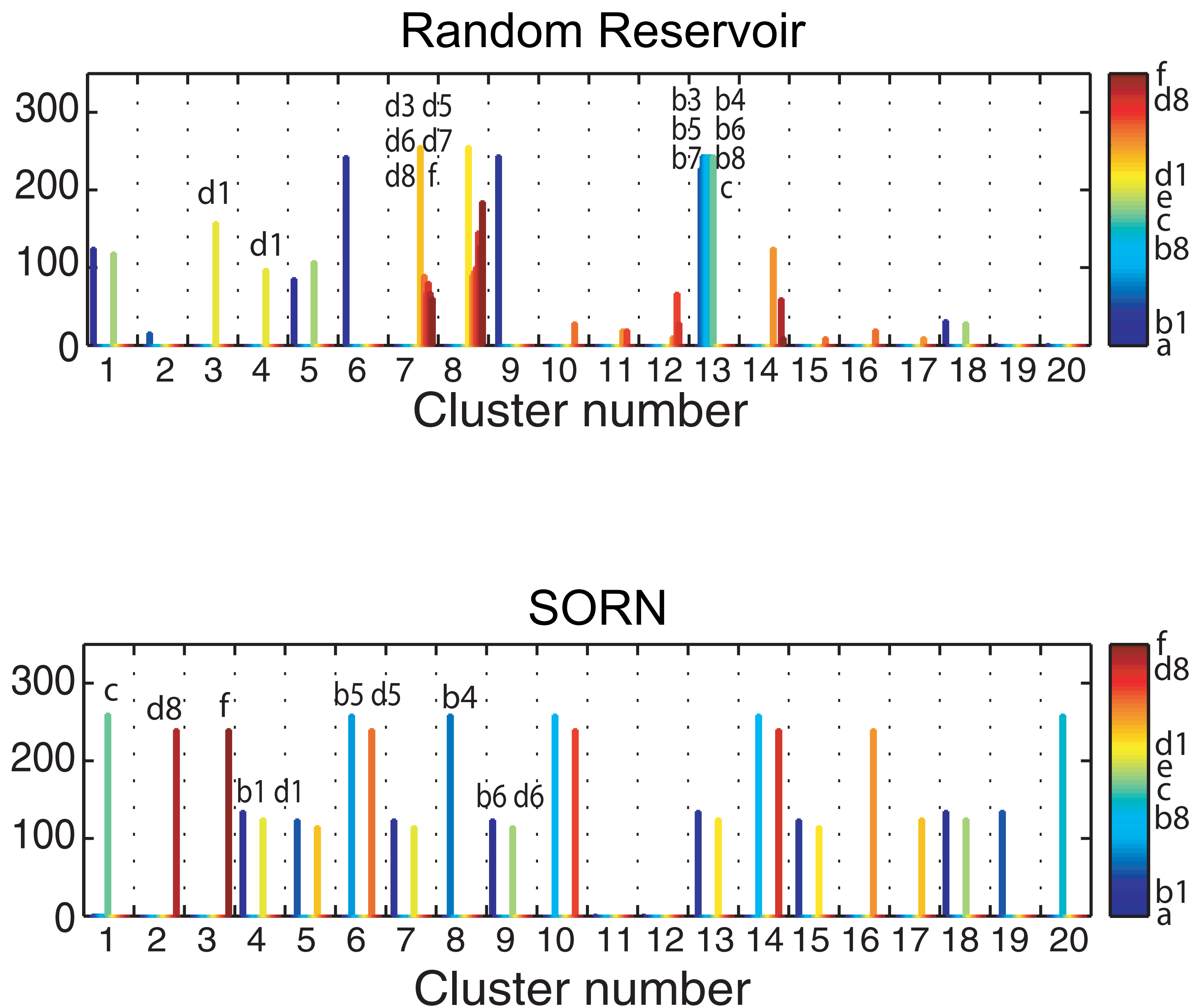
- Input sequences are random alternations of two words:
  - abb....bbbc
  - edd....dddf

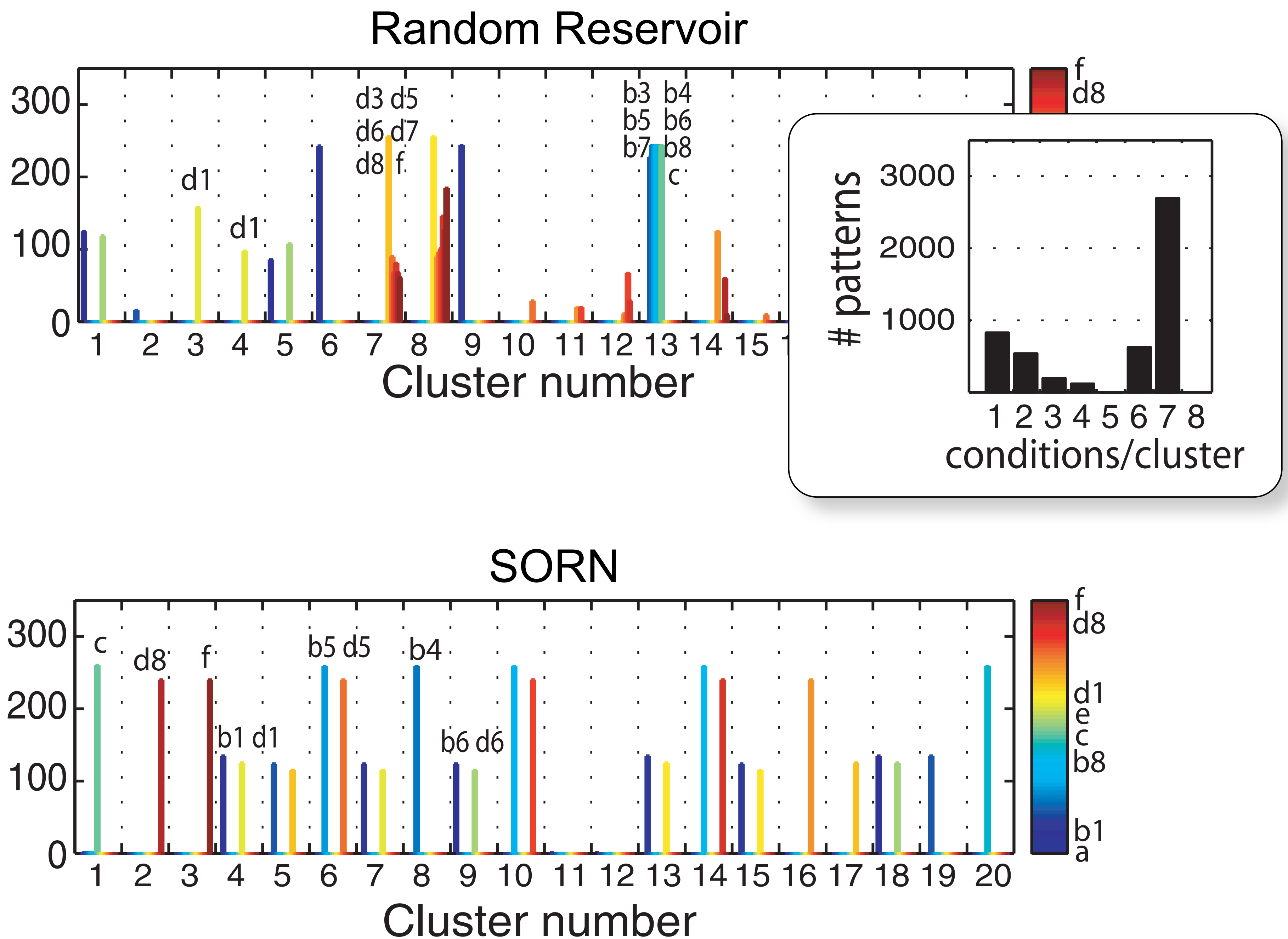
n repetitions of middle letter

- training in two phases:
  - reservoir
  - readout

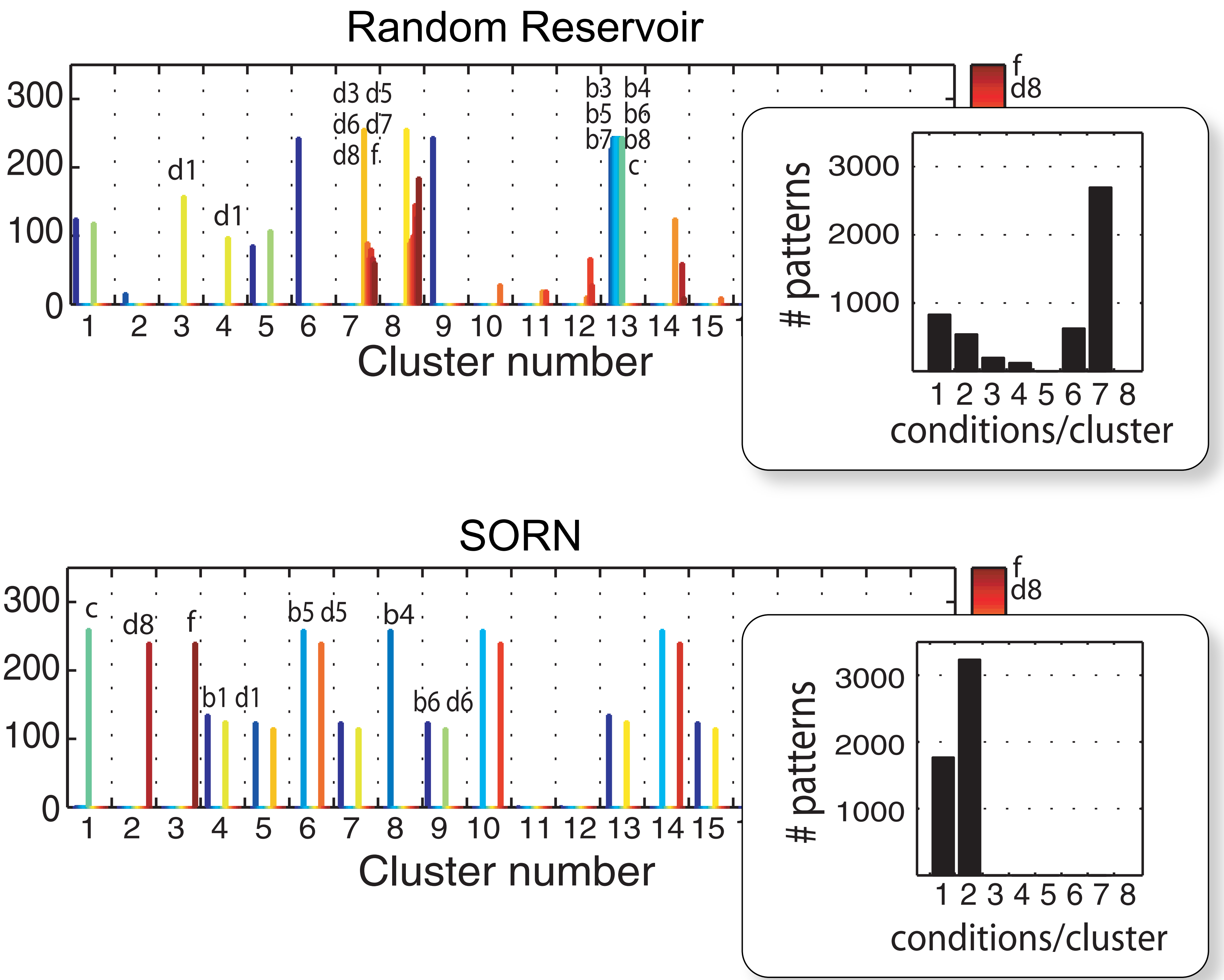
How will the network
learn to represent
these input sequences?



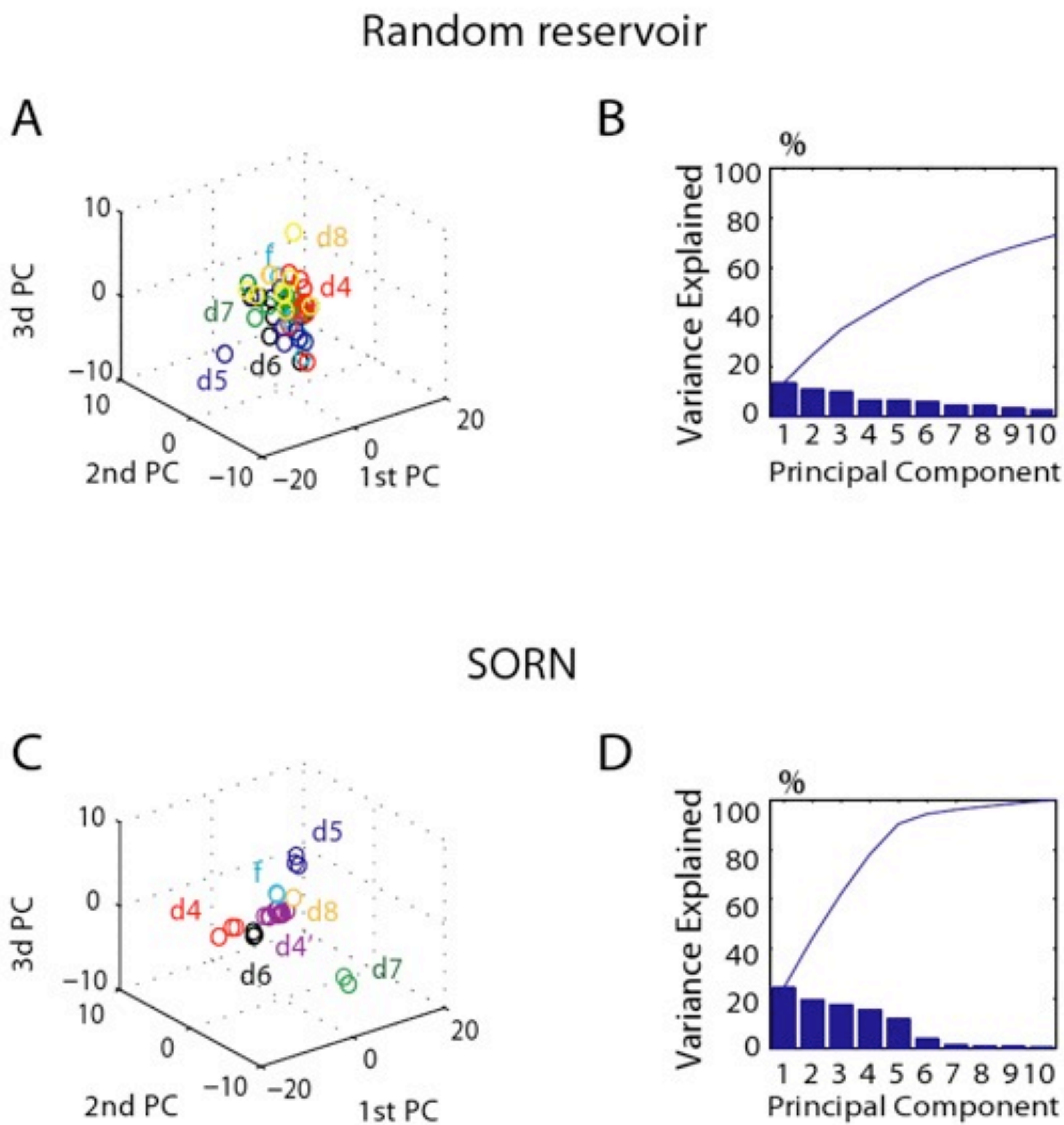

Random Reservoir
d1
d1
d3 d5
d6 d7
d8 f
b3 b4
b5 b6
b7 b8
c
Cluster number
SORN
c
d8
f
b1 d1
b5 d5
b4
b6 d6
Cluster number
f
d8
d1
e
c
b8
b1
a



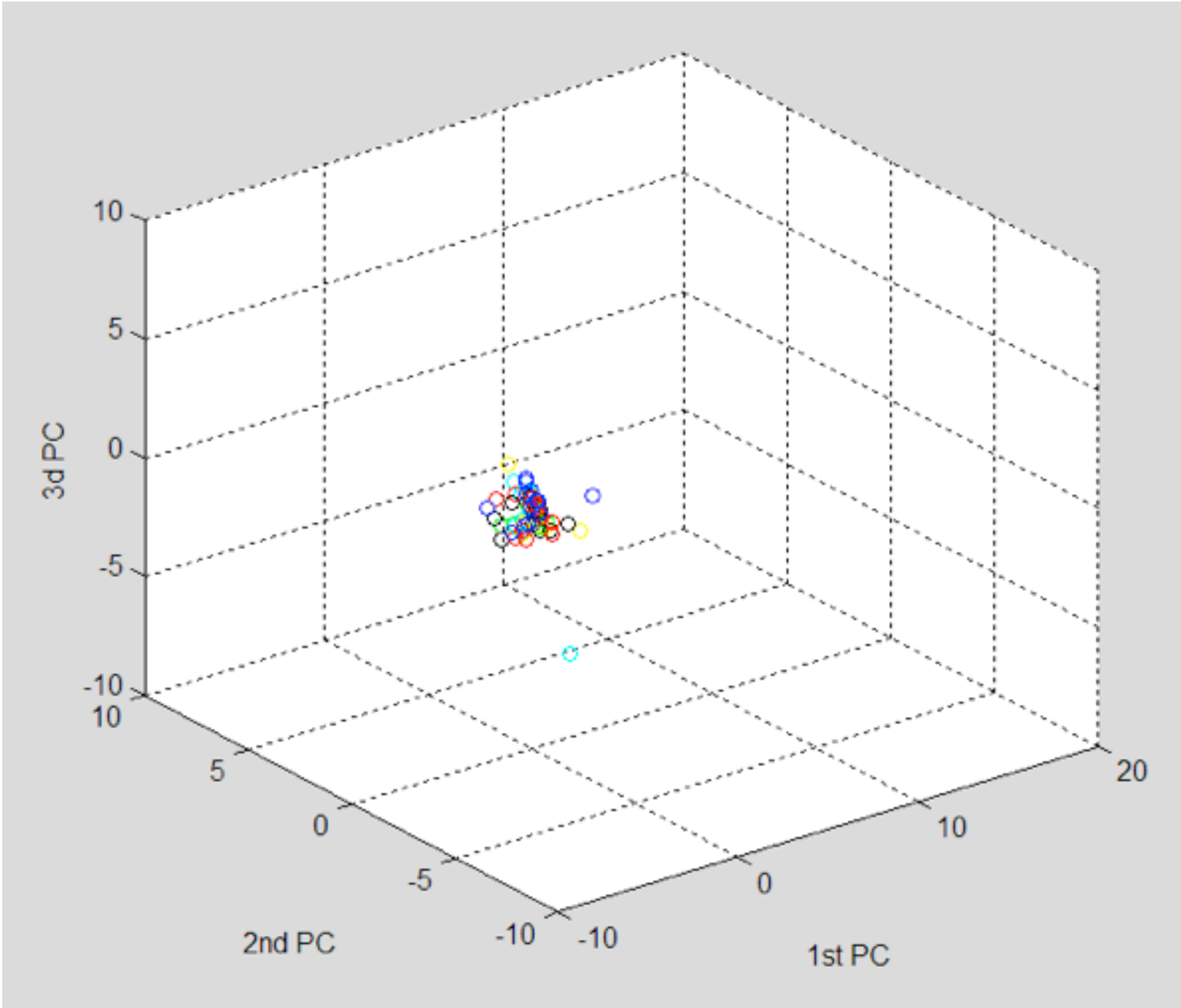
Random Reservoir
Cluster number
patterns
conditions/cluster
SORN
Cluster number
d3 d5
d6 d7
d8 f
b3 b4
b5 b6
b7 b8
c
d1
c
d8
f
b1 d1
b5 d5
b4
b6 d6
f
d8
d1
e
c
b8
b1
a

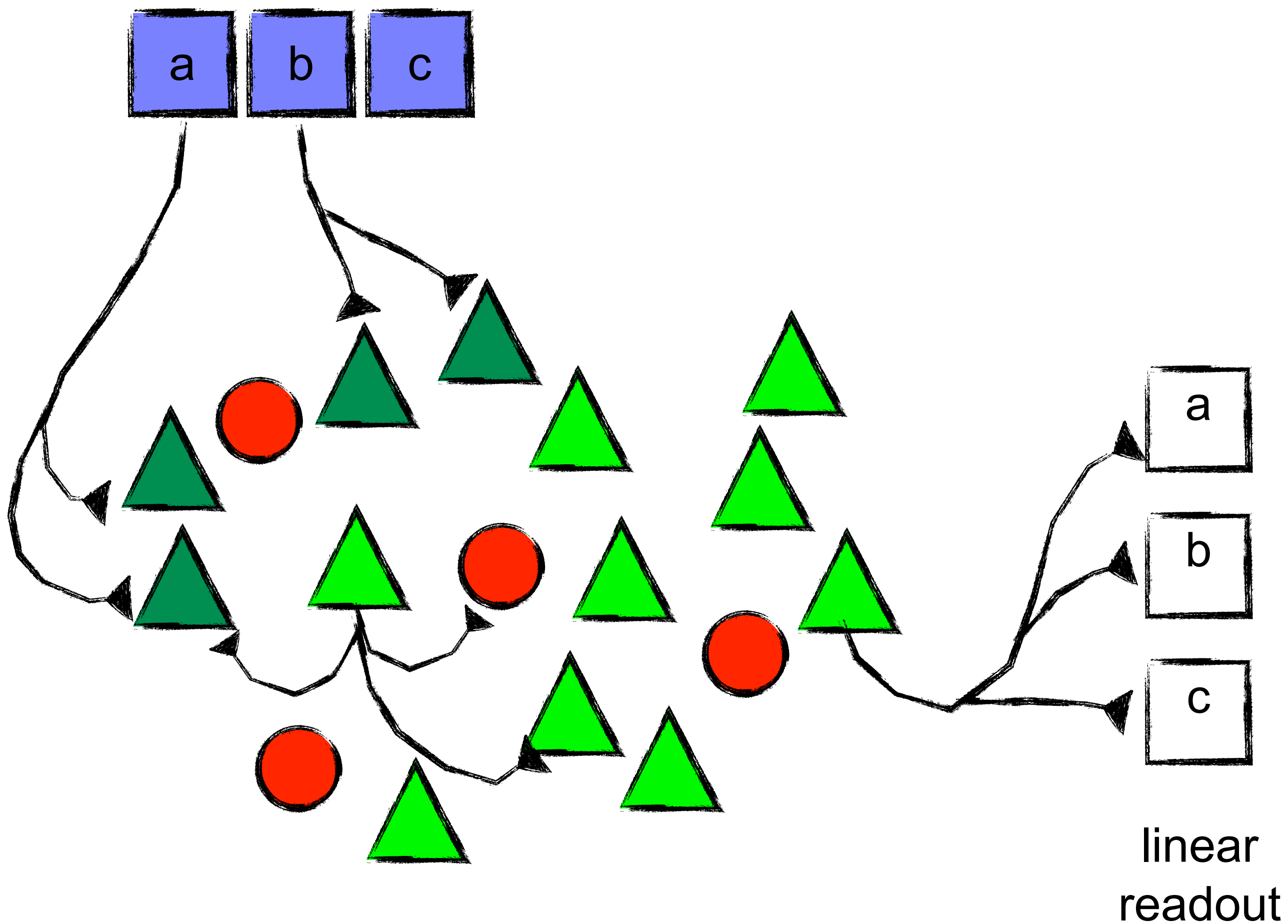


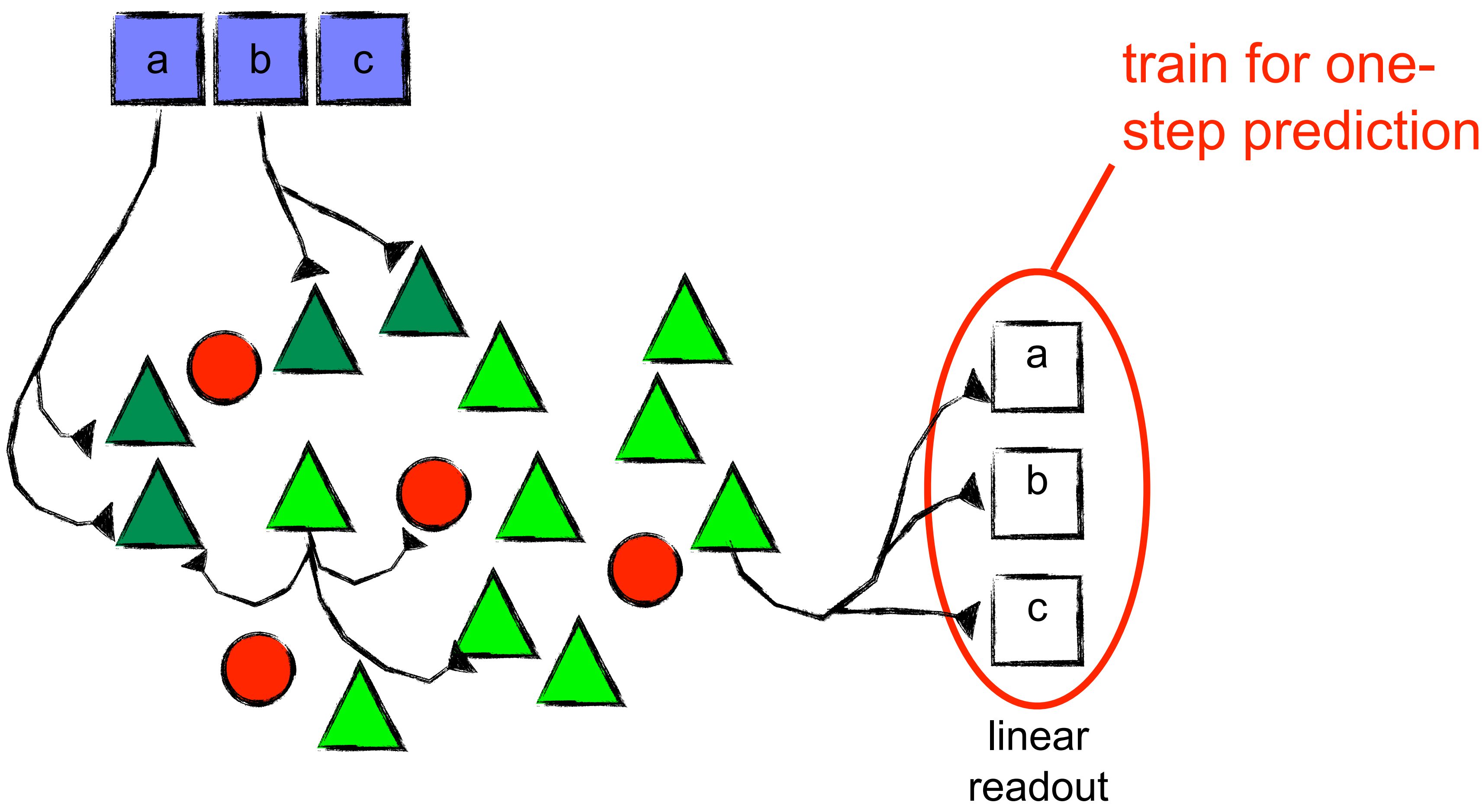
Random Reservoir
300
200
100
0
d1
d1
d3 d5
d6 d7
d8 f
b3 b4
b5 b6
b7 b8
c
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
Cluster number
f
d8
patterns
3000
2000
1000
1 2 3 4 5 6 7 8
conditions/cluster
SORN
300
200
100
0
c
d8
f
b1 d1
b5 d5
b4
b6 d6
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
Cluster number
f
d8
patterns
3000
2000
1000
1 2 3 4 5 6 7 8
conditions/cluster



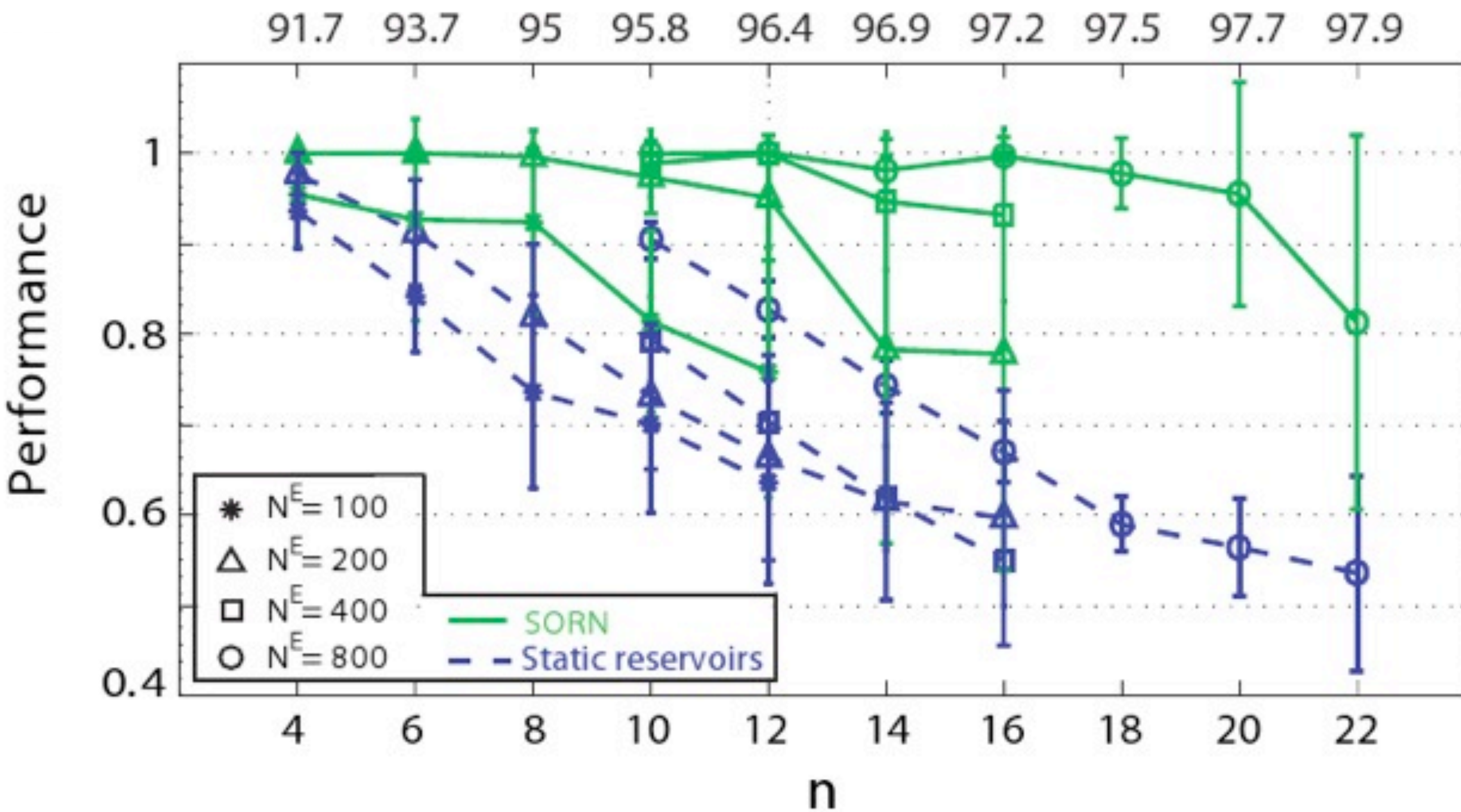
Random reservoir
A
B
%
Variance Explained
Principal Component
3d PC
2nd PC
1st PC
d8
f
d4
d7
d5
d6
SORN
C
D
d5
f
d8
d4
d4'
d6
d7

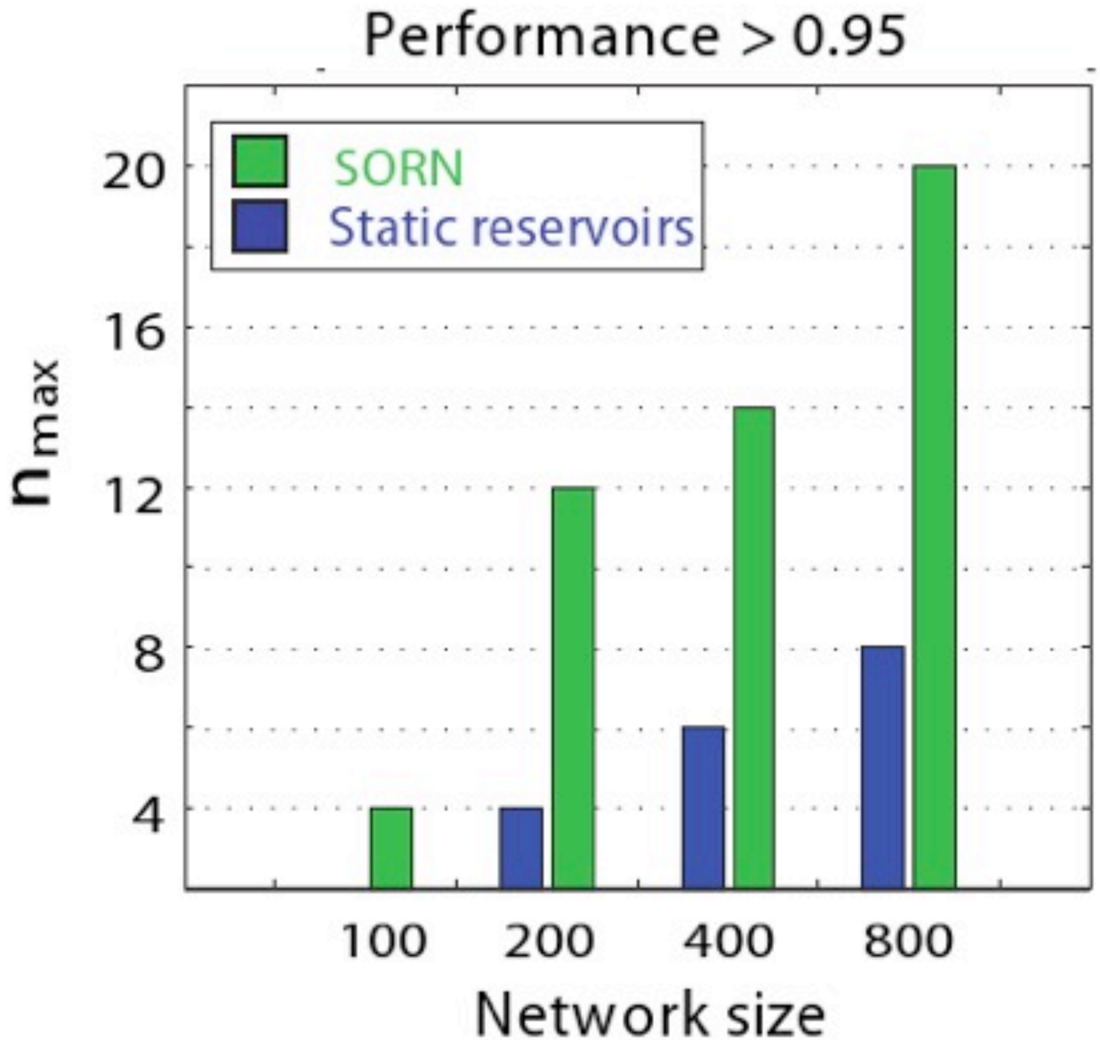

10
5
0
-5
-10
3d PC
10
5
0
-5
-10
2nd PC
-10
0
10
20
1st PC

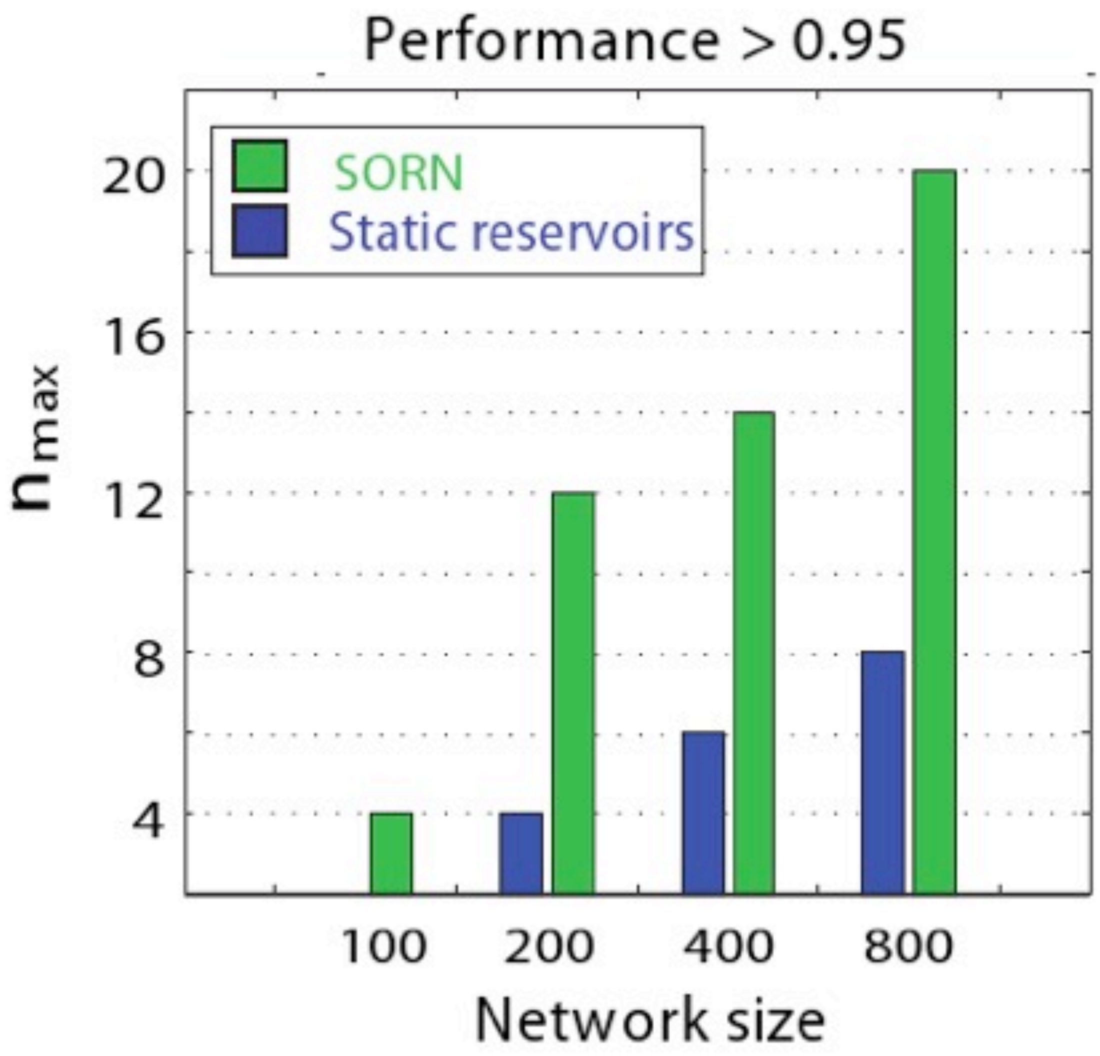

a
b
c
a
b
c
linear
readout


a
b
c
train for one-
step prediction
a
b
c
linear
readout

SORNs outperform static reservoirs




Performance > 0.95
SORN
Static reservoirs
20
16
12
8
4
n_{max}
100
200
400
800
Network size

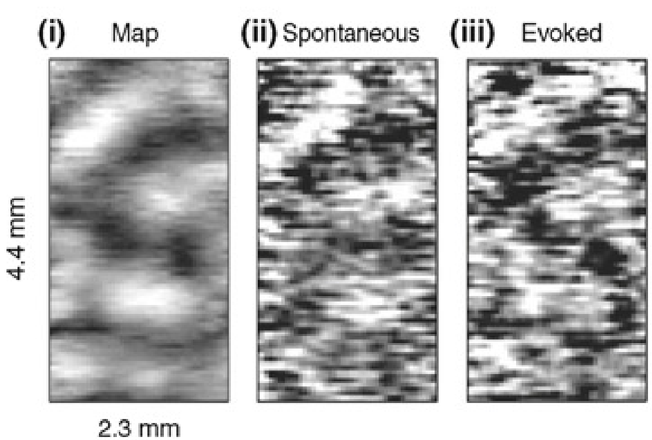


Performance > 0.95
SORN
Static reservoirs
20
16
12
8
4
n_{max}
100
200
400
800
Network size

„superfluid“

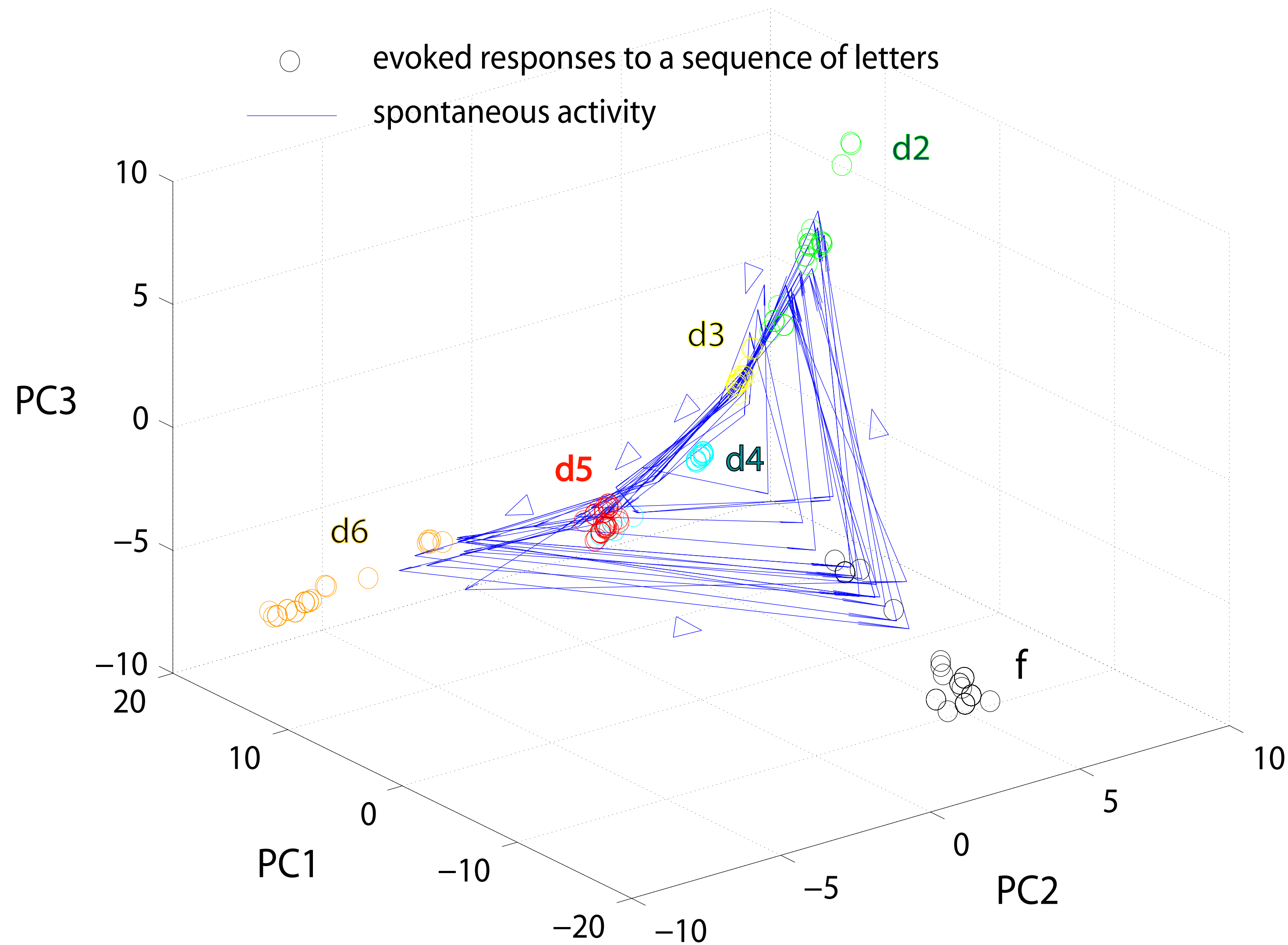
Spontaneous Activity

- Cortex exhibits patterned spontaneous activity [Tsodyks et al., 1999; Kenet et al., 2003; Fiser et al., 2004; Ringach 2009]
- Following repetetive presentation of a visual stimulus, spontaneous activity shows similarity to evoked response [Han et al., 2008]
- Spontaneous activity might represent the prior in a Bayesian inference sense [Berkes et al., 2009; Fiser et al., 2010]

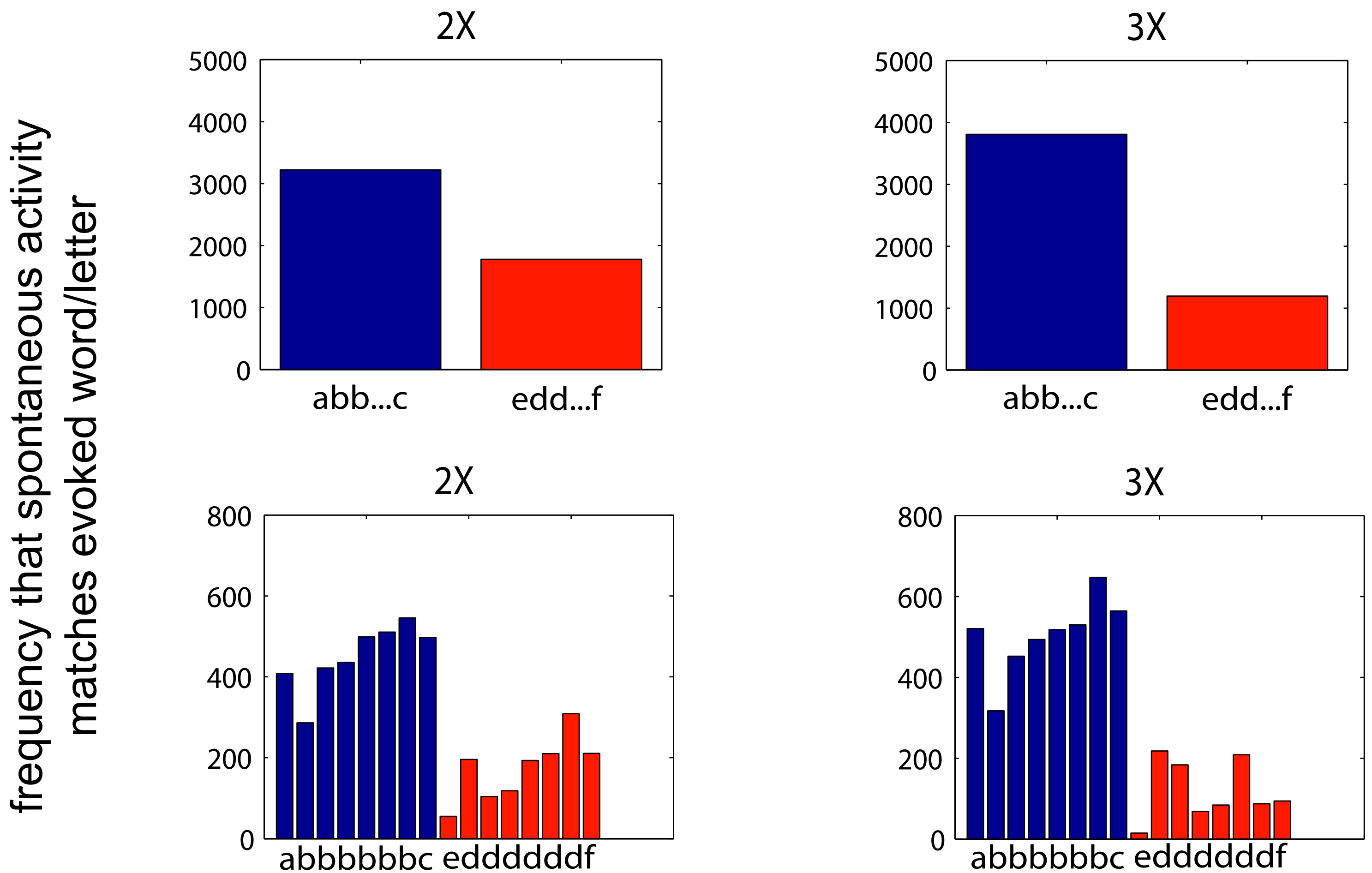


$$p(r) = \sum_s p(r \mid s)p(s)$$

Spontaneous activity patterns match evoked activity



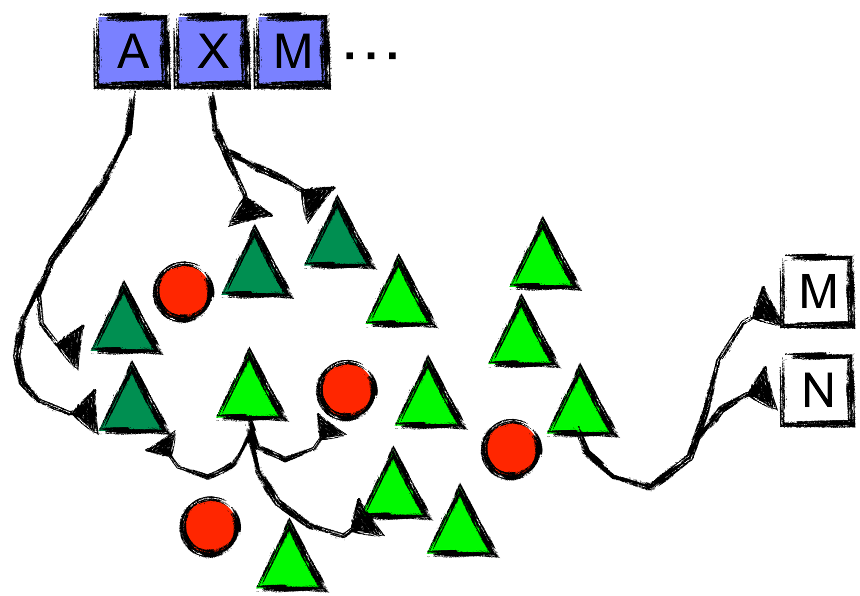
Spontaneous activity patterns reflect input statistics

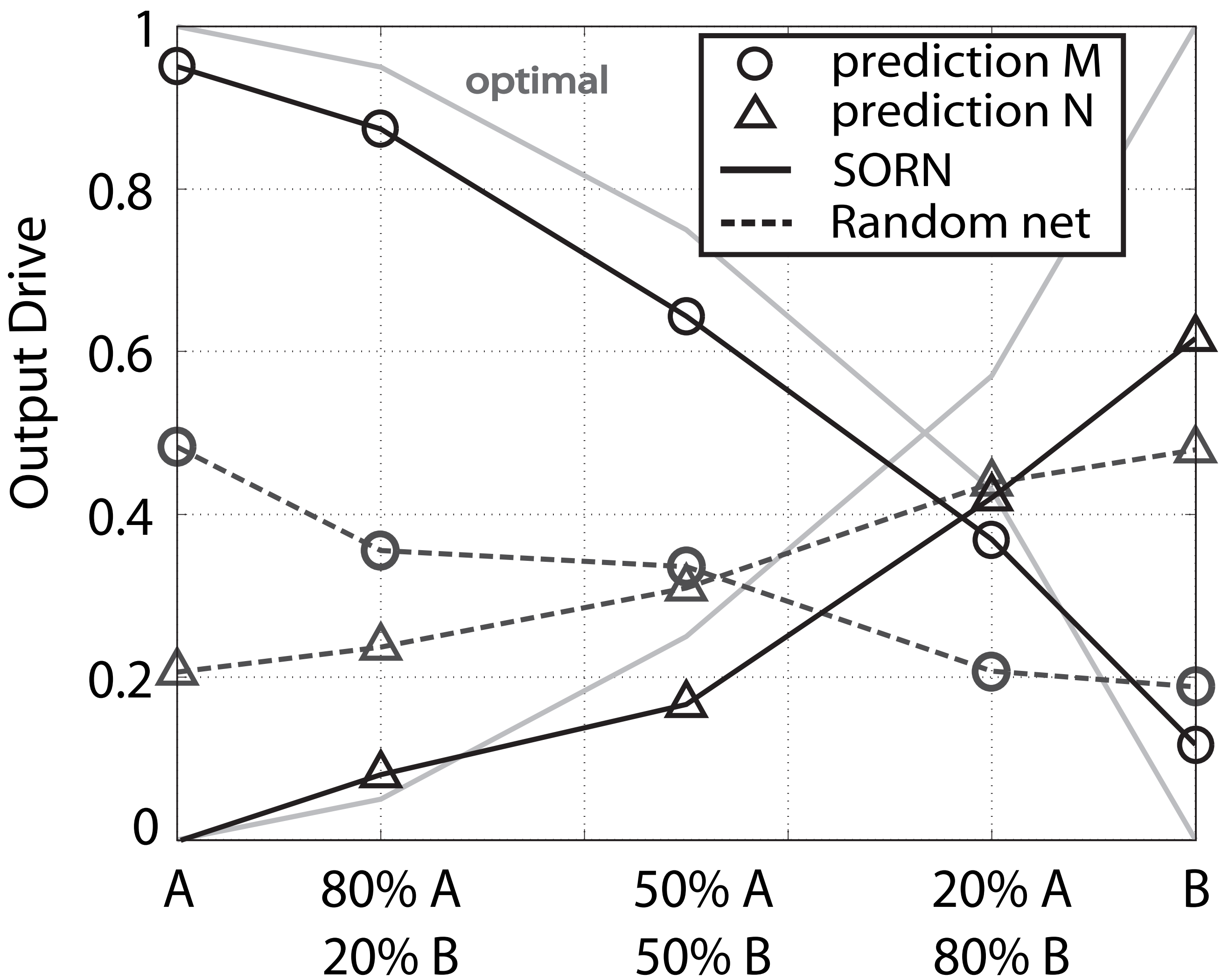


Preliminary Work: Statistical Inference?

···AXXXXXMBXXXXXNAXXXXXMAXXXXXMBXXXXXNAXXXXXMAXXXXXMAXXXXXM···

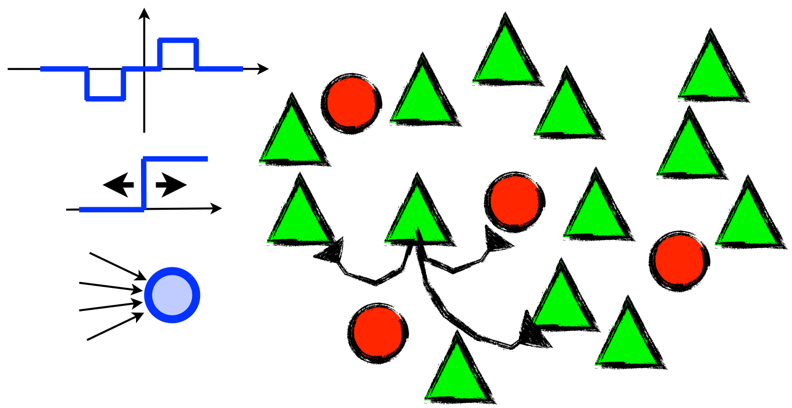
- 2 input words:
  „AXXXXXM“ (75%),
  „BXXXXXN“ (25%)
- readout trained to predict \`M´ vs. \`N´
- test: instead of \`A´ or \`B´, show mixtures of the two, e.g., 20% \`A´ and 80% \`B´

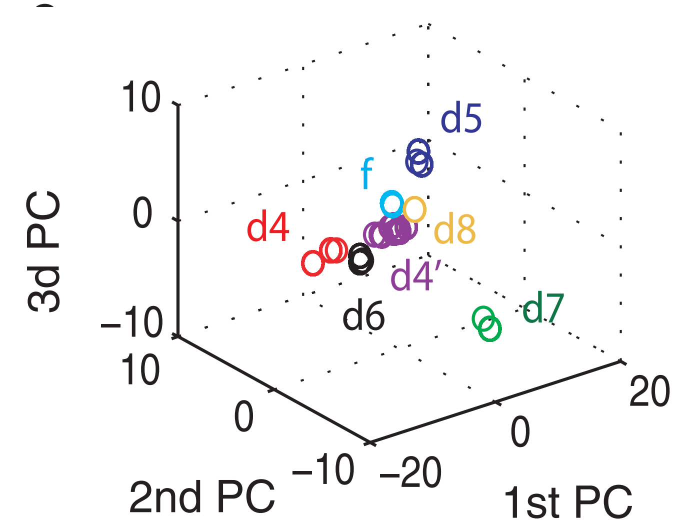




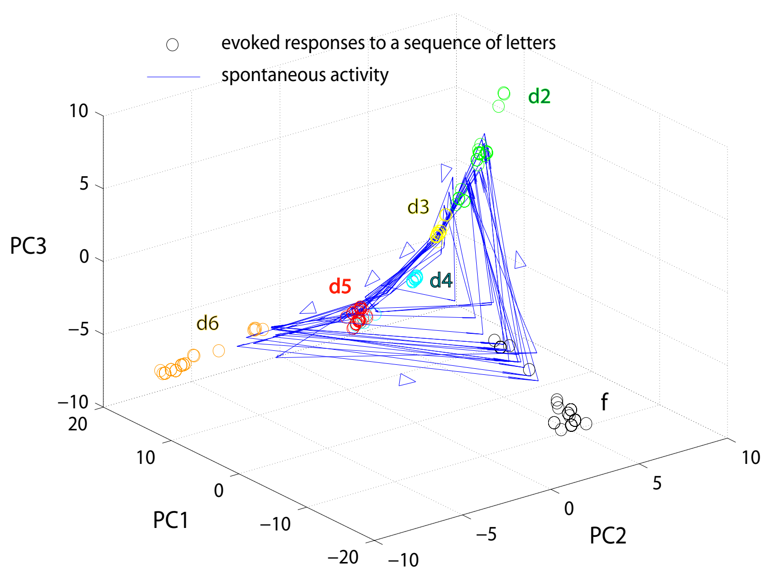
network combines ambiguous A/B input
with prior information for prediction

Discussion

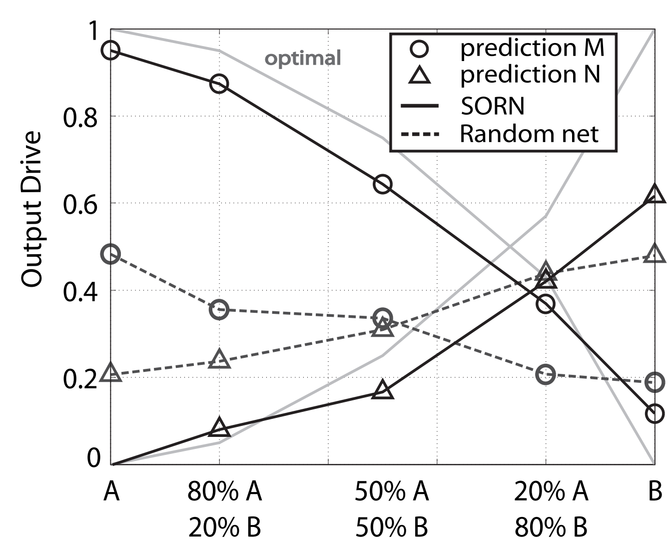




3d PC
2nd PC
1st PC
d5
f
d8
d4
d4′
d6
d7



evoked responses to a sequence of letters
spontaneous activity
PC3
PC1
PC2
d2
d3
d4
d5
d6
f



Output Drive
optimal
prediction M
prediction N
SORN
Random net
A
80% A 20% B
50% A 50% B
20% A 80% B
B

Acknowledgments

Lab:

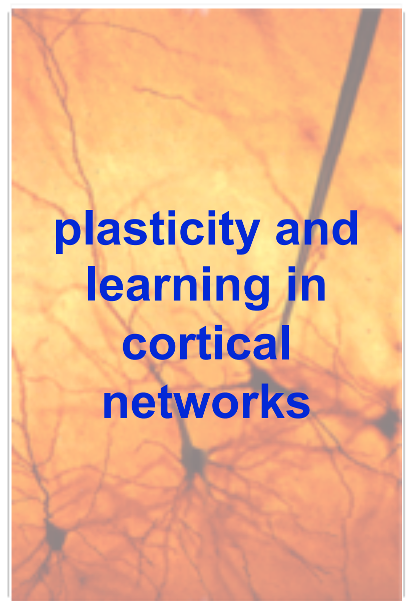
- Pramod Chandrashekhariah
- Dr. Christos Dimitrakakis
- Dr. Manu P. John
- Dr. Prashant Joshi
- Daniel Krieg
- Luca Lonini
- Daniela Pamplona
- Dr. Constantin Rothkopf
- Sohrab Saeb
- Dr. Lisa Scocchia
- Dr. Philip Sterne
- Quan Wang
- Thomas Weisswange
- Dr. Pengsheng Zheng

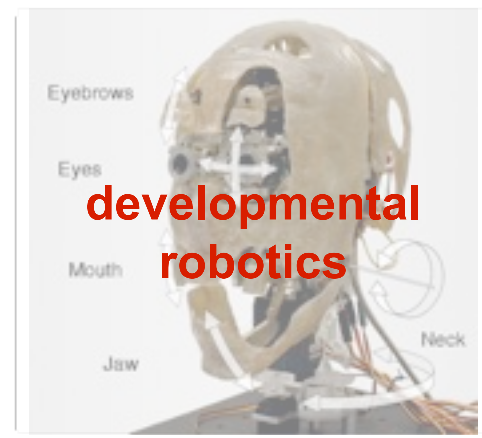
Collaborators in this project:

- Dr. Andreea Lazar
- Dr. Gordon Pipa

THANK YOU!


infant cognitive
development

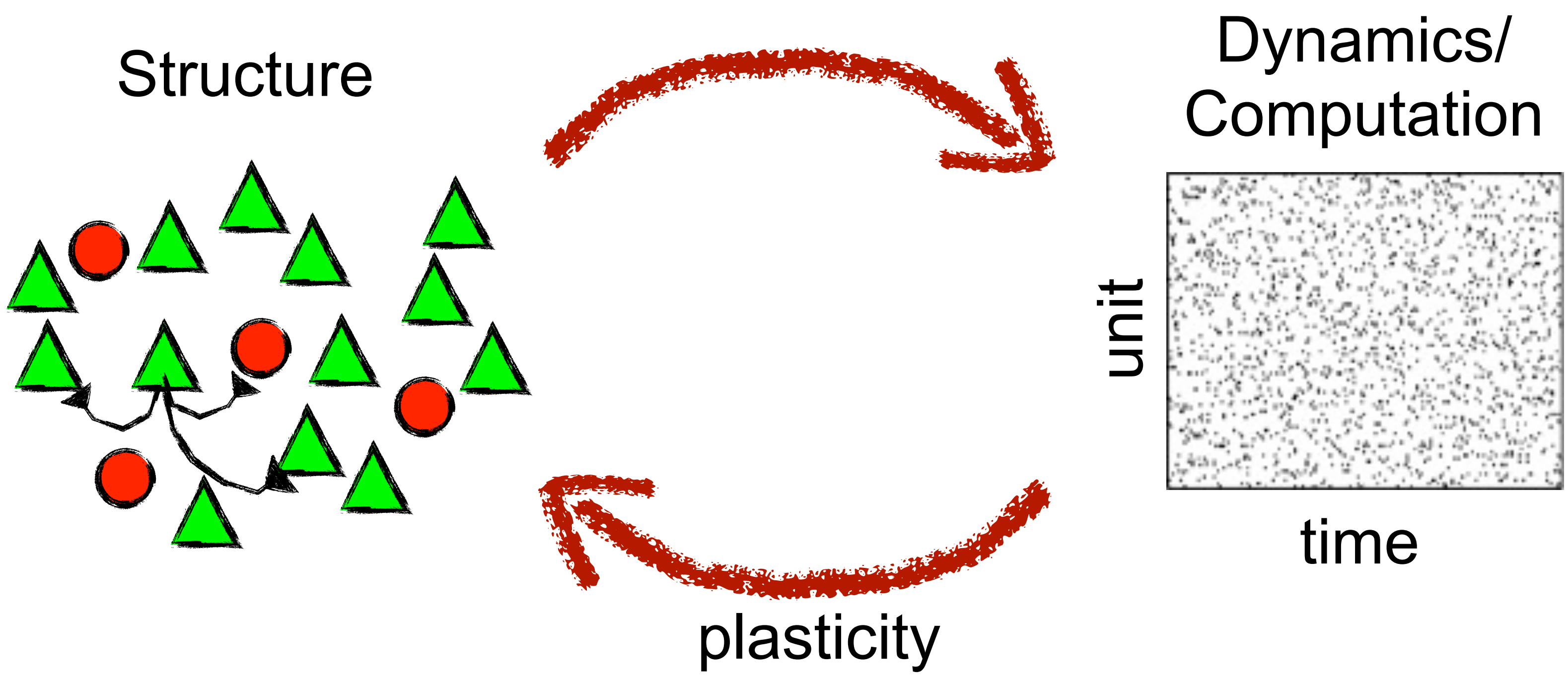

plasticity and
learning in
cortical
networks

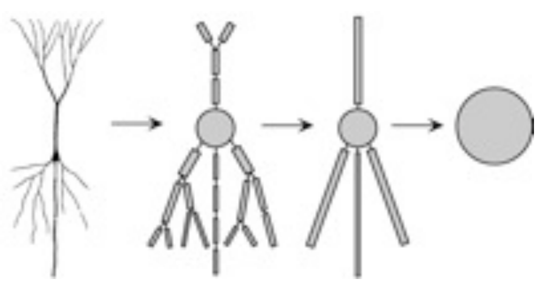

Eyebrows
Eyes
developmental
robotics
Mouth
Jaw
Neck

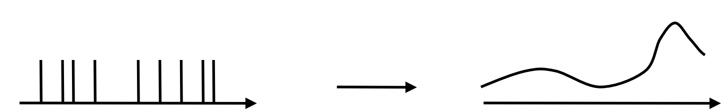

Water Bottle
biologically inspired
computer vision
Honey Beer
DVD Box

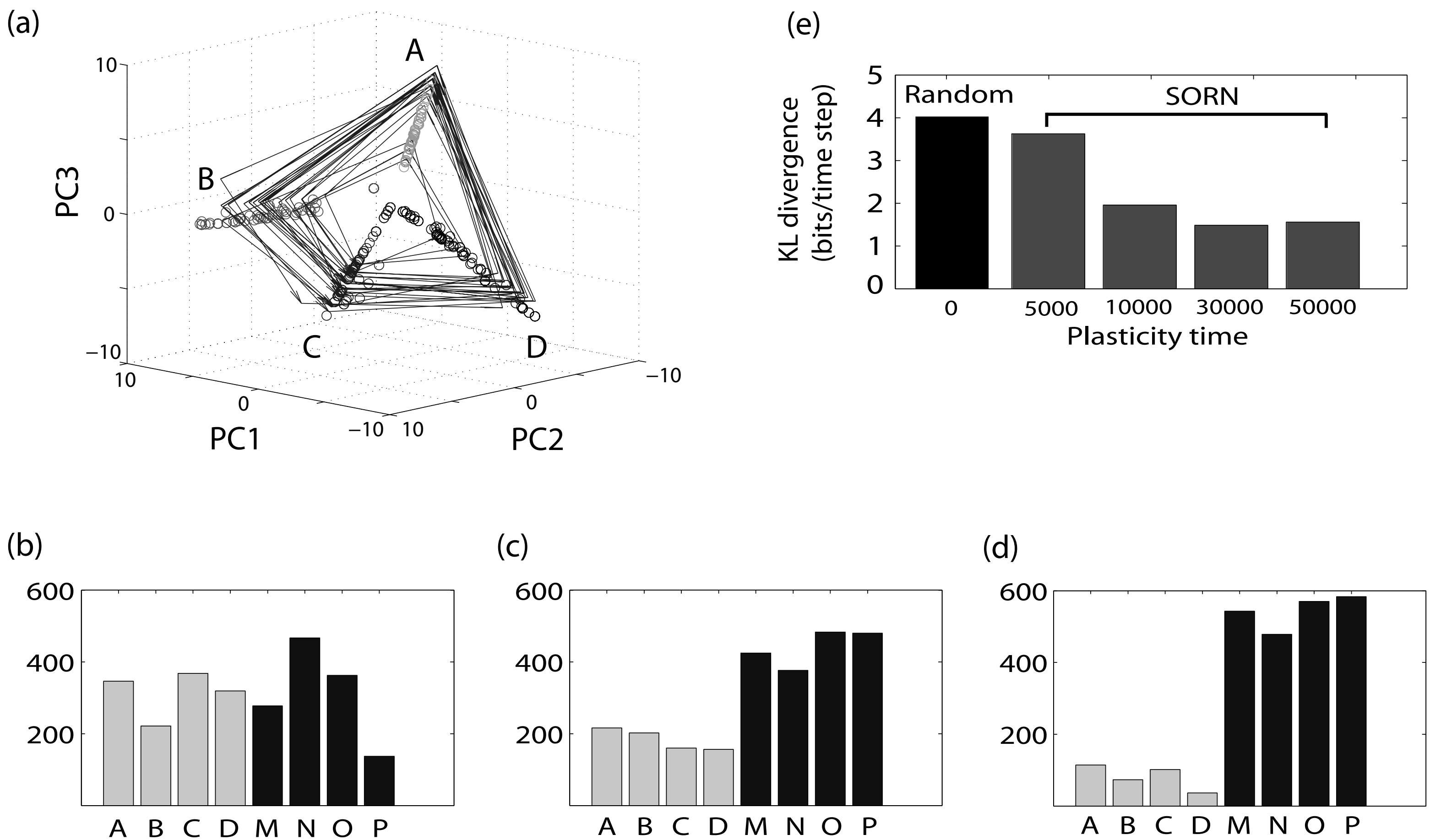

visual
psychophysics

Network Self-Organiztion





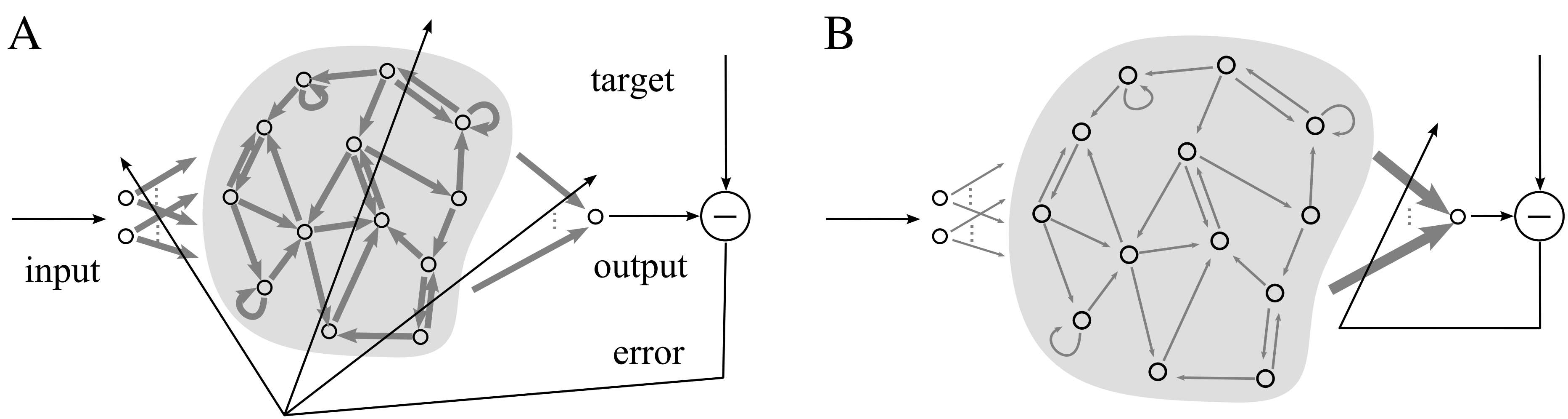




(a)
A
B
C
D
PC1
PC2
PC3
10
0
−10
(e)
KL divergence
(bits/time step)
Random
SORN
Plasticity time
0
5000
10000
30000
50000
5
4
3
2
1
0
(b)
(c)
(d)
600
400
200
A B C D M N O P

Reservoir Computing

- class of recurrent neural network architectures utilizing a „reservoir“ with fixed random connectivity (review: Lukosevicius & Jaeger, 2009)
- examples: Echo State Neworks (Jaeger, 2001); Liquid state machines (Maas et al., 2002)
- fading memory property, separation property



Reservoir Computing

- class of recurrent neural network architectures utilizing a „reservoir“ with fixed random connectivity (review: Lukosevicius & Jaeger, 2009)
- examples: Echo State Neworks (Jaeger, 2001); Liquid state machines (Maas et al., 2002)
- fading memory property, separation property