Lecture 2: Connectionist approach

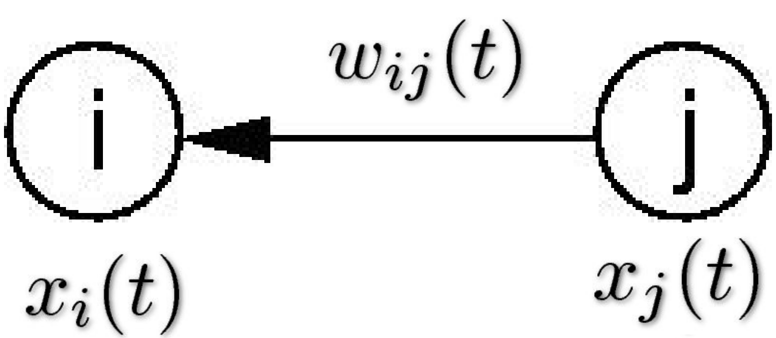
# Elements and feedforward networks

Dr. Thomas Nowotny
University of Sussex

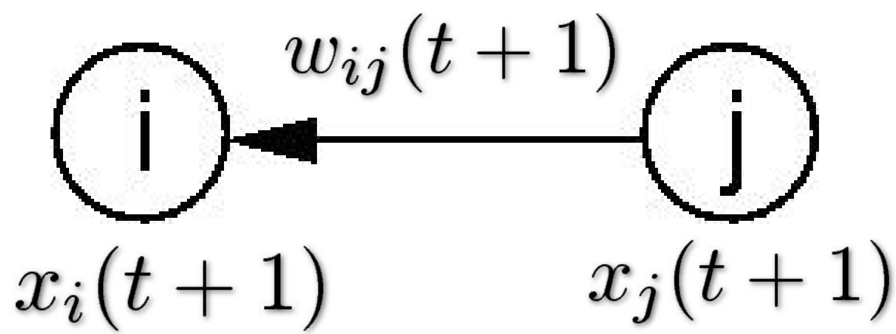
# Connectionist approach

- The connectionist approach is an approach of minimal assuptions:
  - Neurons have two states “on” (1) or “off (0)
  - Time can be discretized in discrete steps
  - Neurons are either connected (1) or not (0)

At time $t$




At time $t+1$



# McCulloch-Pitts neurons

In a connectionist approach, neurons are described by the McCulloch-Pitts neuron model:

$$x_i(t+1) = \Theta\left(\sum_{j=1}^{N} w_{ij}(t)x_j(t) - \theta\right)$$

$x_i(t) \in \{0,1\}$ - state of neuron i at time t

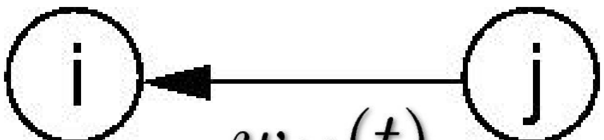
$w_{ij}(t) \in \{0,1\}$ - state of connection (synapse) from neuron j to neuron i at time t

$$\Theta(x) = \begin{cases} 1 & x > 0 \\ 0 & \text{otherwise} \end{cases}$$ - Heaviside function

$\theta \in \mathbb{N}$ - firing threshold

# McCulloch-Pitts neurons

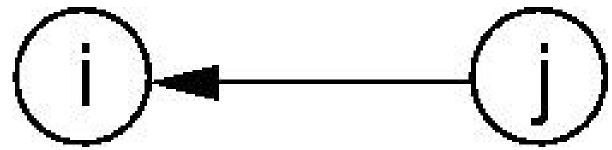
At time $t$




$w_{ij}(t)$

$x_i(t)$ $x_j(t)$

At time $t+1$




$$x_i(t+1) = \Theta(w_{ij}(t)x_j(t) - \theta)$$

Note:

- Need to take values from previous time step
- The notation of $w_{ij}(t)$ may seem awkward at first:
  $w_{ij}(t) = w_{i \leftarrow j}(t)$


Dr. Thomas Nowotny,
Centre for Computational Neuroscience and Robotics
University of Sussex

# Connectionist approach

- Advantages
  - Can be used even if details are not known
  - Remains valid if knowledge about details changes
  - Can often be applied to many systems, even if details differ

# Connectionist approach

- Disadvantages
  - Some things are awkward to implement (e.g. mutual inhibition in layers)
  - Some things are almost impossible to include (e.g. sub-threshold oscillations)
  - Does not include intrinsically active neurons well
  - Intrinsic neuron dynamics not described (e.g. refractory period)
  - ...

# McCulloch-Pitts neurons are hyperplanes

The equation $\sum_{j=1}^{N} w_{ij} x_j = \theta$ defines a plane in N dimensional space.

"McCulloch-Pitts neurons fire to the right of a hyper-plane and are silent on the left."

For example N=2:

$$w_{i1} x_1 + w_{i2} x_2 = \theta$$

$$\Leftrightarrow x_2 = -\frac{w_{i1}}{w_{i2}} x_1 + \frac{\theta}{w_{i2}}$$

$$\Leftrightarrow y = ax + b \qquad \left( a = -\frac{w_{i1}}{w_{i2}},\ b = \frac{\theta}{w_{i2}} \right)$$

# Random connections

- In connectionist approaches connections are often chosen to be *random* (“generic”):

$$w_{ij} = \begin{cases} 1 & \text{with probability } p_c \\ 0 & \text{with probability } q_c = 1 - p_c \end{cases}$$

- Similarly input neurons are often assumed to fire with a fixed probability

$$x_j = \begin{cases} 1 & \text{with probability } p_{\text{in}} \\ 0 & \text{with probability } q_{\text{in}} = 1 - p_{\text{in}} \end{cases}$$

# Propagation of probabilities

- With random connections and random input firing, other neurons i have a probability to fire

$$P\left(x_i = 1\right) = P\left(\sum_{j=1}^{N} w_{ij} x_j \geq \theta\right)$$

$$= \sum_{k=\theta}^{N} P\left(\sum_{j=1}^{N} w_{ij} x_j = k\right)$$

# Statistical independence

Definition of **statistical independence**:

Event A is independent from event B if and only if

$$P(A \cap B) = P(A) \cdot P(B)$$

**Conditional probabilities**:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Read: “P of A given B”

# Bayes theorem

Why these definitions make sense: **If A and B independent**,

$$P(A\,|\,B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A)$$

**Bayes theorem**:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

$$= \frac{P(A \cap B)}{P(A)} \frac{P(A)}{P(B)} = P(B|A) \frac{P(A)}{P(B)}$$

# Binomial distribution

If $x_i \in \{0, 1\},\ i = 1, \ldots, N$ are independent random variables with distribution $\{p, 1-p\}$ then the probability distribution for the sum is

$$P\Big(\sum_{i=1}^{N} x_i = k\Big) = \binom{n}{k} p^k (1-p)^{N-k}$$

Quick proof: The sum is k if we put k “1” into the N $x_i$

| $x_i$ | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 1 |

$$P = \mathrm{p} \cdot \mathrm{q} \cdot \mathrm{q} \cdot \mathrm{p} \cdot \mathrm{q} \cdot \mathrm{p} \cdot \mathrm{q} \cdot \mathrm{p} \cdot \mathrm{p} \cdot \mathrm{q} \cdot \mathrm{q} \cdot \mathrm{p} = p^k (1-p)^{N-k}$$

Dr. Thomas Nowotny,
Centre for Computational Neuroscience and Robotics
US
University of Sussex

# Proof of binomial distribution

There are $\binom{N}{k} = \frac{N!}{k!(N-k)!}$ ways to put the “1” and “0”,

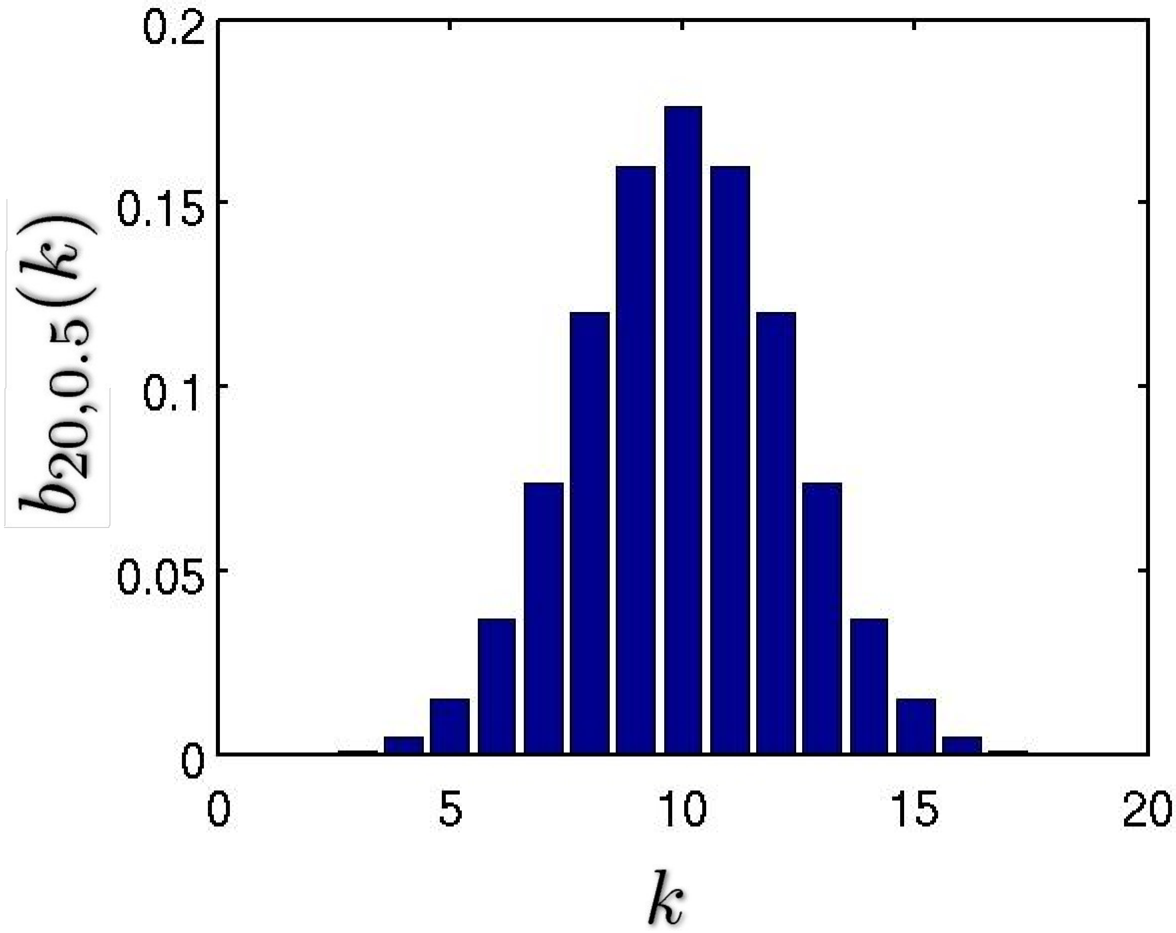
and all are mutually exclusive, therefore

$$P\Big(\sum_{i=1}^{N} x_i = k\Big) = \binom{N}{k} p^k (1-p)^{N-k}$$

The binomial distribution is often denoted as

$$P\Big(\sum_{i=1}^{N} x_i = k\Big) = b_{N,p}(k)$$

# Binomial distribution properties




- Expectation value

$$\mathbb{E}\sum_{i=1}^{N} x_i = N \cdot p$$

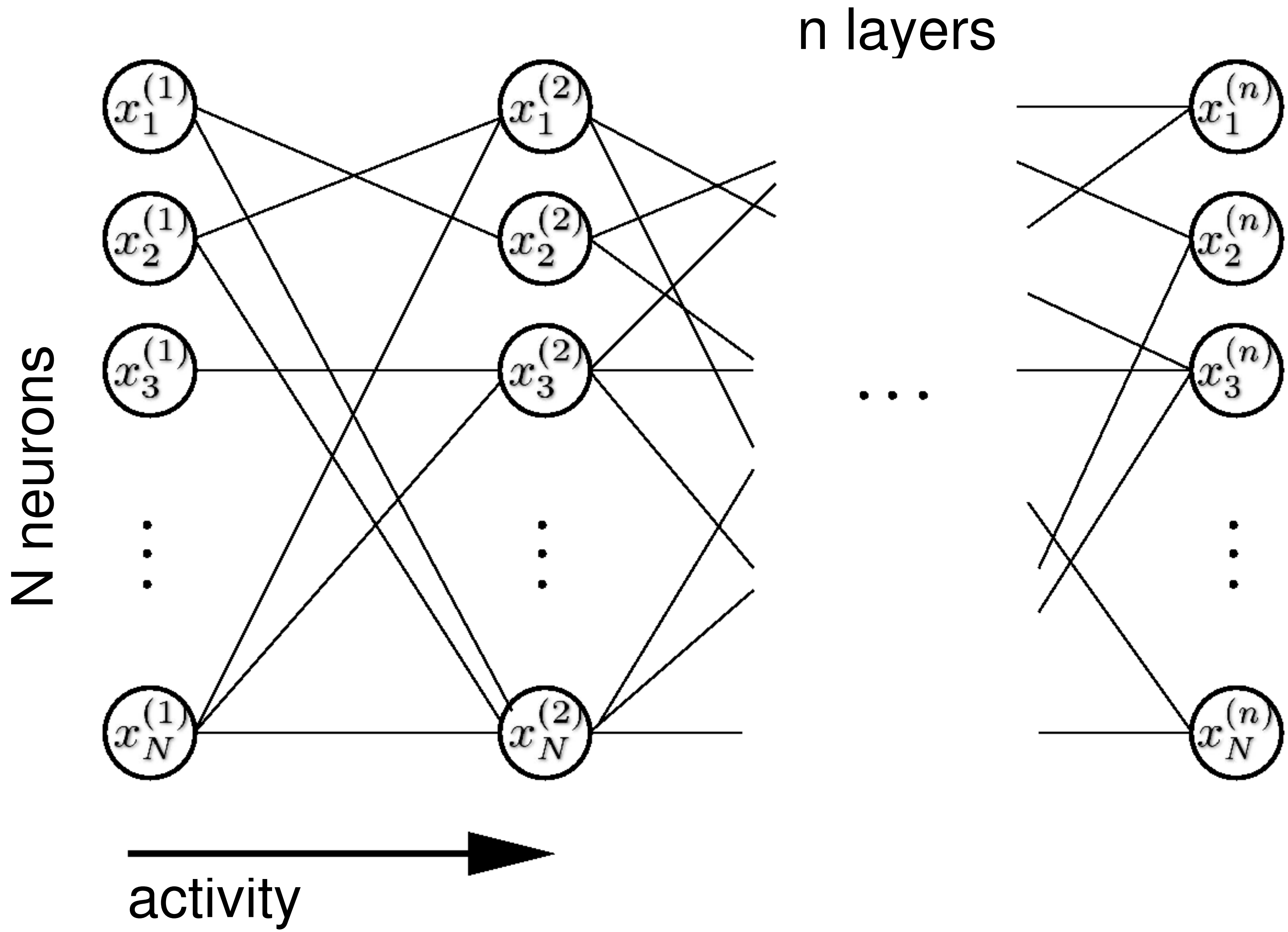
- Standard deviation

$$\sigma_{\sum_{i=1}^{N} x_i} = \sqrt{Np(1-p)}$$

Expectation value and maximum are *not* the same

More on this: Exercises

# Feedforward networks



# Feedforward networks

- Denote $X^{(j)} := \sum_{i=0}^{N} x_i^{(j)}$
- Assume that $a_j$ input neurons fire in Layer j
- Then the probability of a neuron in layer j+1 to fire is

$$p^{(j+1)}(a_j) := P(x_i^{(j+1)} = 1 \,|\, X^{(j)} = a_j) = \sum_{k=\theta}^{a_j} b_{a_j, p_c}(k)$$

# Feedforward networks

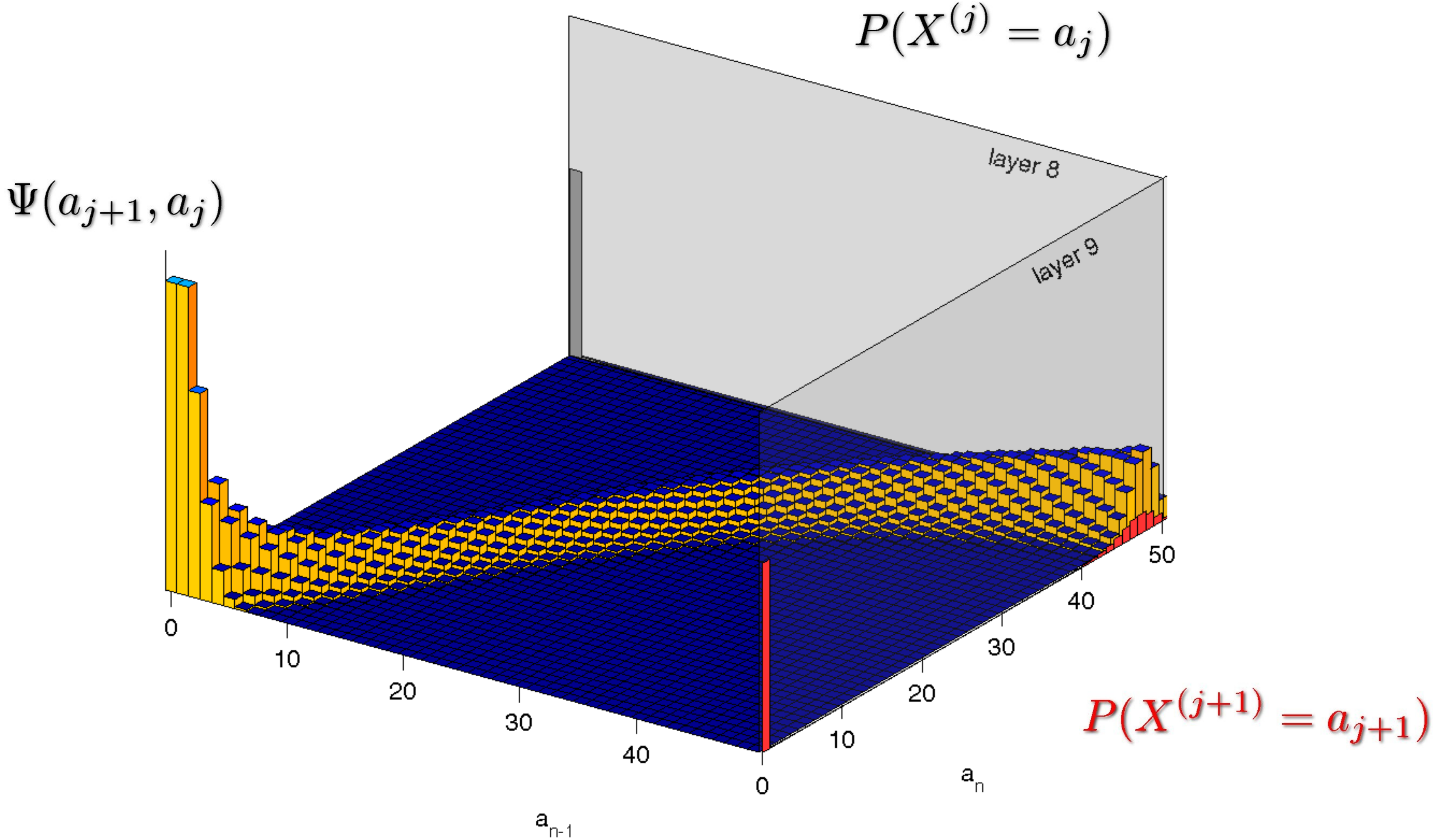
- Because the connections are chosen independently,

$$P(X^{(j+1)} = a_{j+1} \,|\, X^{(j)} = a_j) = \underbrace{b_{N,p^{(j+1)}(a_j)}(a_{j+1})}_{\Psi(a_{j+1},a_j)}$$

- Then, by definition of conditional probabilities

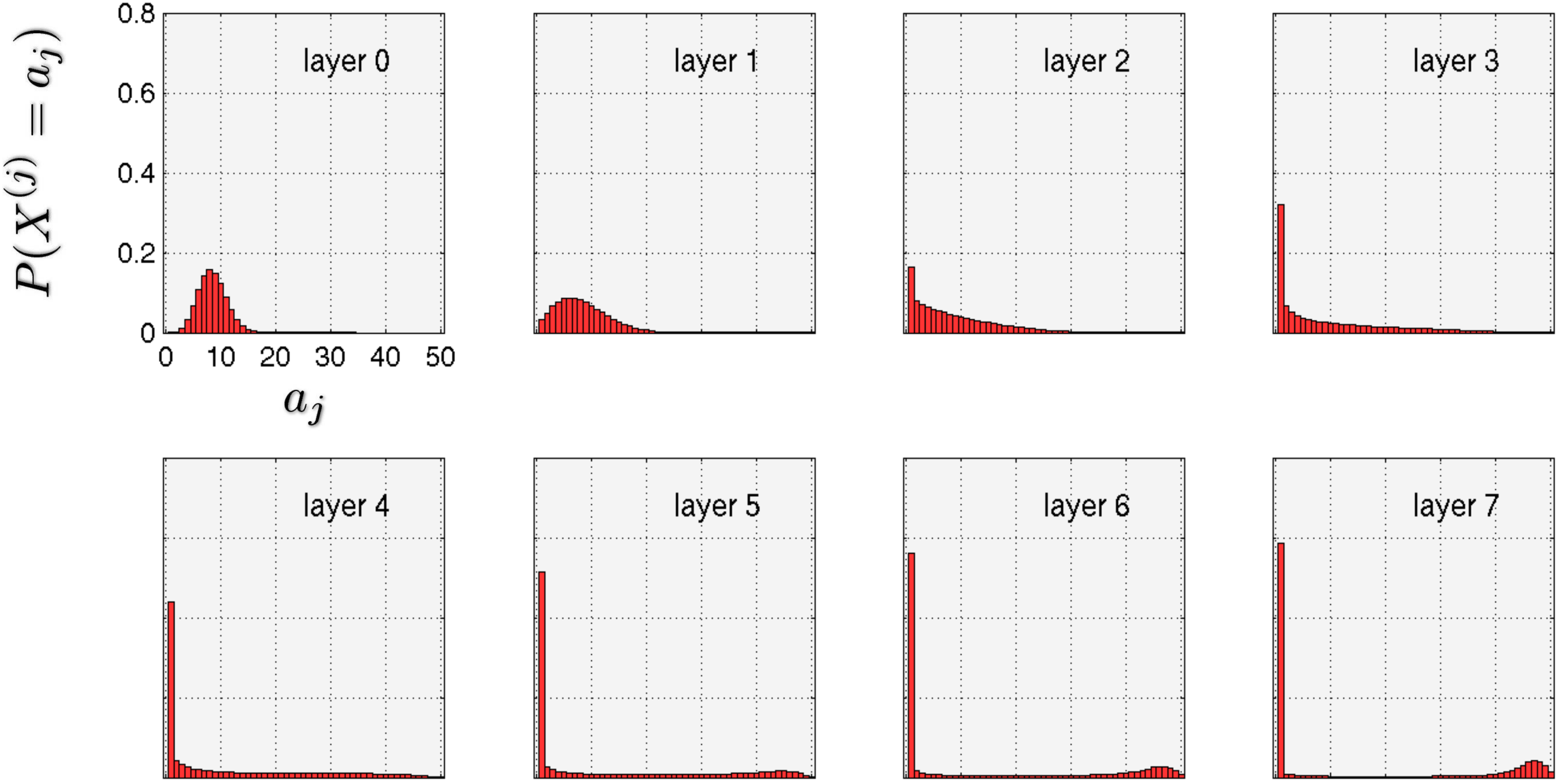
$$P(X^{(j+1)} = a_{j+1}) = \sum_{a_j=0}^{N} \Psi(a_{j+1},a_j)\, P(X^{(j)} = a_j)$$

- We get an iteration equation!



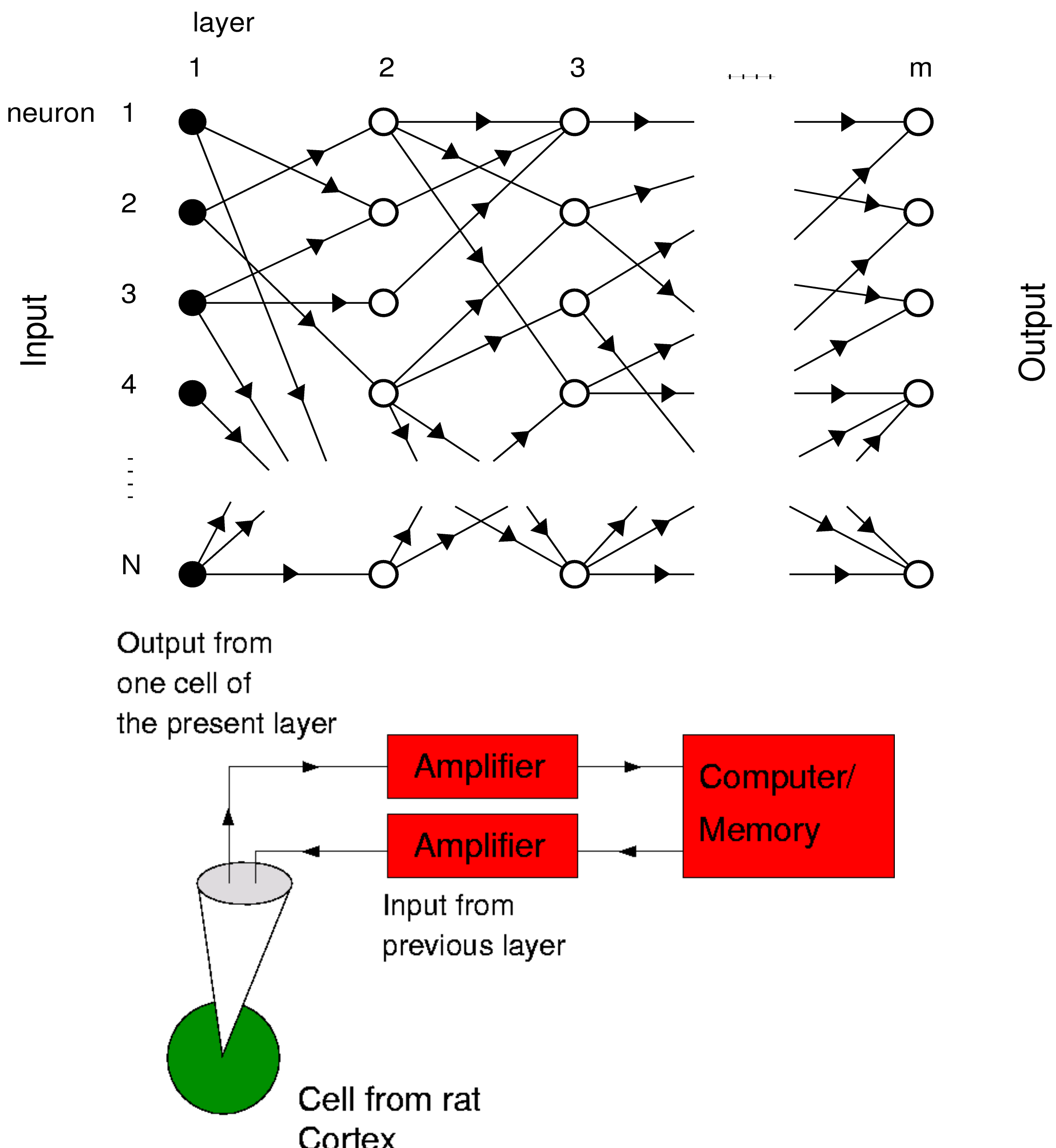
$P(X^{(j)} = a_j)$
layer 8
layer 9
$\Psi(a_{j+1}, a_j)$
0
10
20
30
40
$a_{n-1}$
0
10
20
30
40
50
$a_n$
$P(X^{(j+1)} = a_{j+1})$

# Iterated probability distribution




Neurons either all fire, or all are silent in deeper layers.

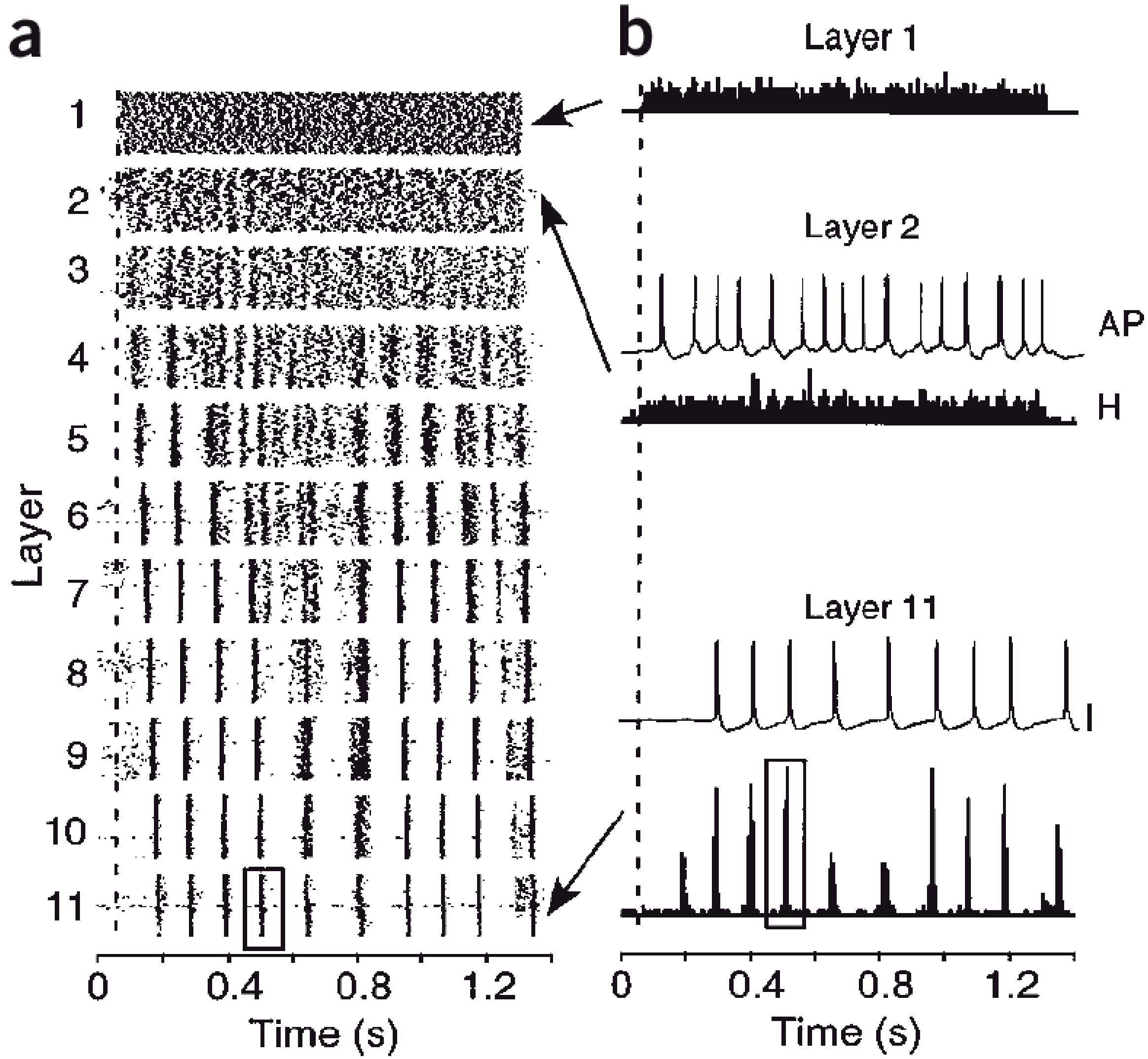
# Iteratively constructed networks (ICNs)




- One or a few cortical neurons
- Dynamic clamp synapses
- Strictly feedforward networks
- Connectivity is randomly chosen by the computer

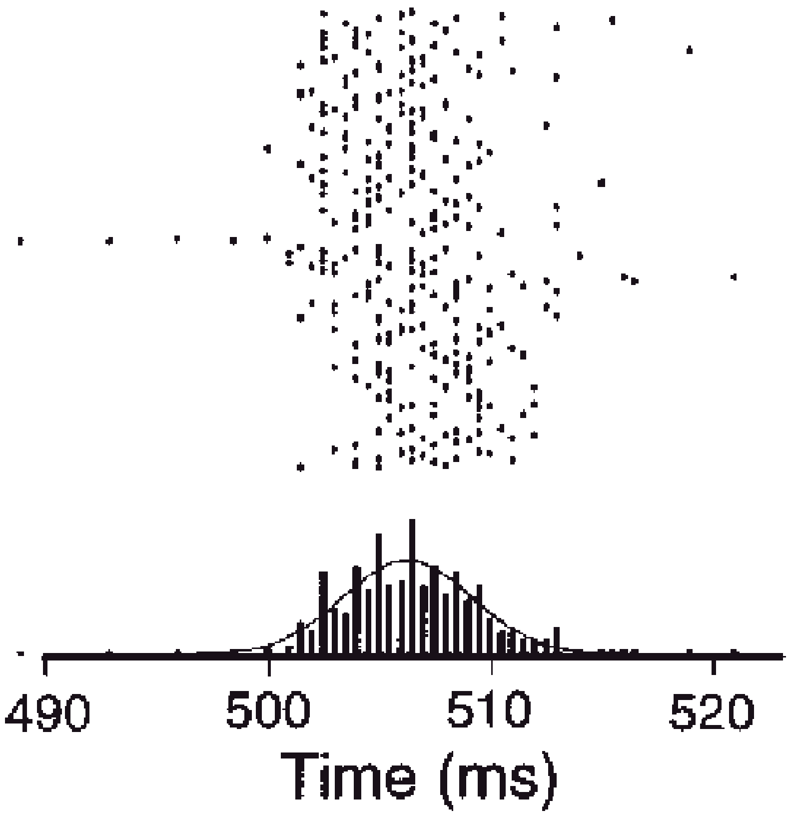
A. Reyes, Nat. Neurosci. **6**:593 (2003)

# Reyes main results




A. Reyes, Nat. Neurosci. **6**:593 (2003)

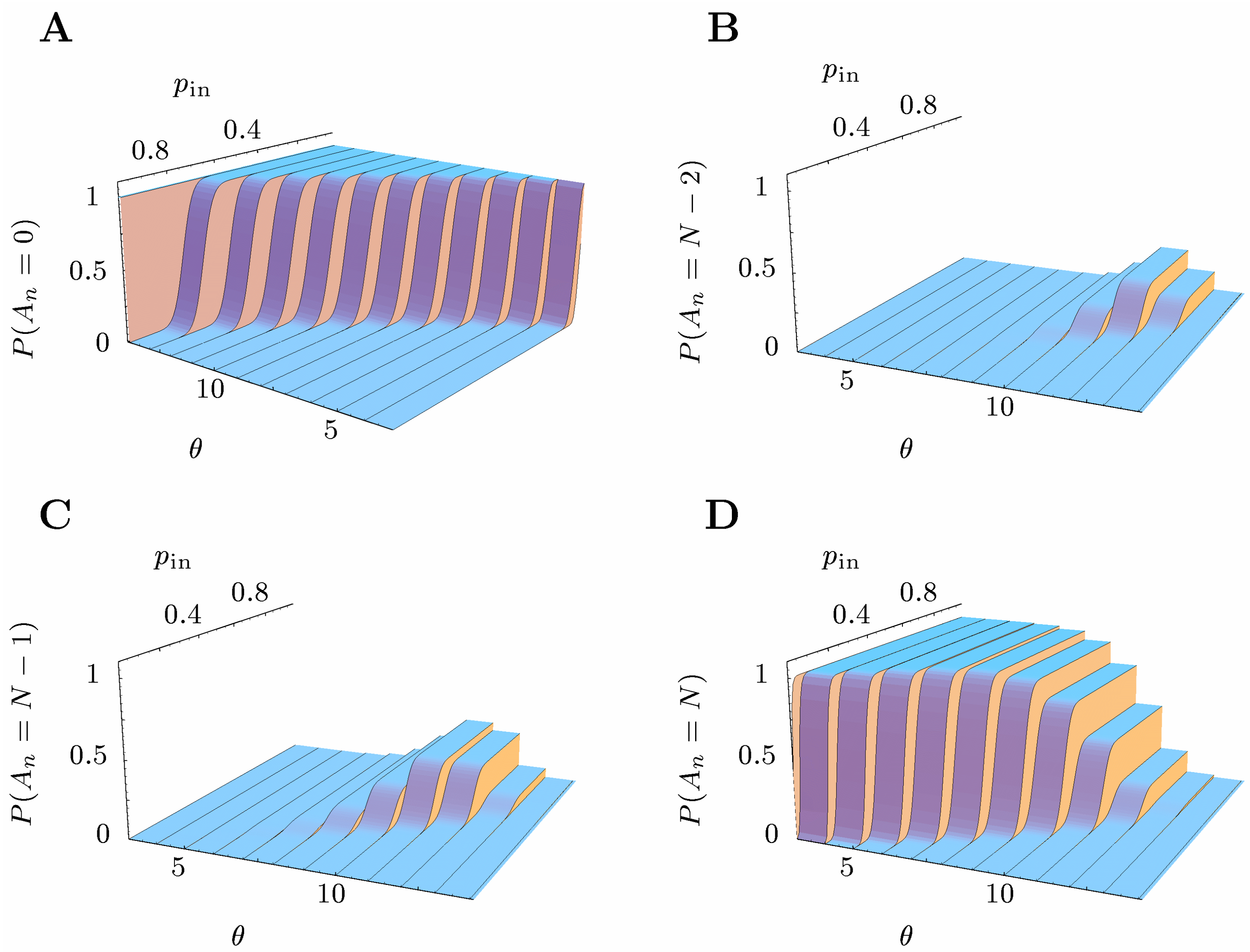
# Interpretation of Δ t




- McCulloch Pitts neurons: Δ t is the integration time of the neurons
- Our analysis: Δ t is the width of an isolated “synchronized event”

$\Delta t \approx$ 5-10 ms

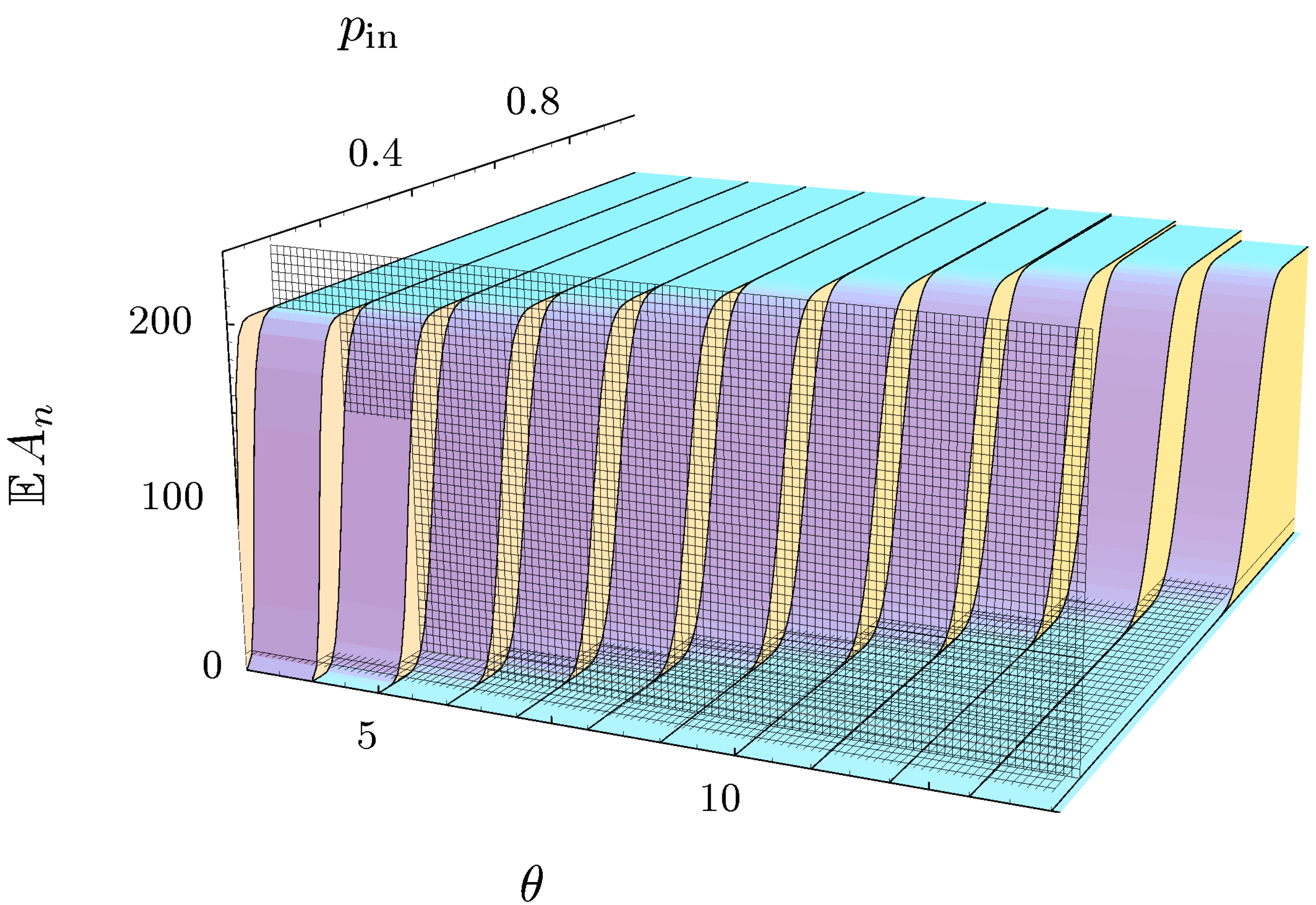
# Properties of invariant measure




n=20

N=200

# Threshold estimate for Reyes' neuron




n=20

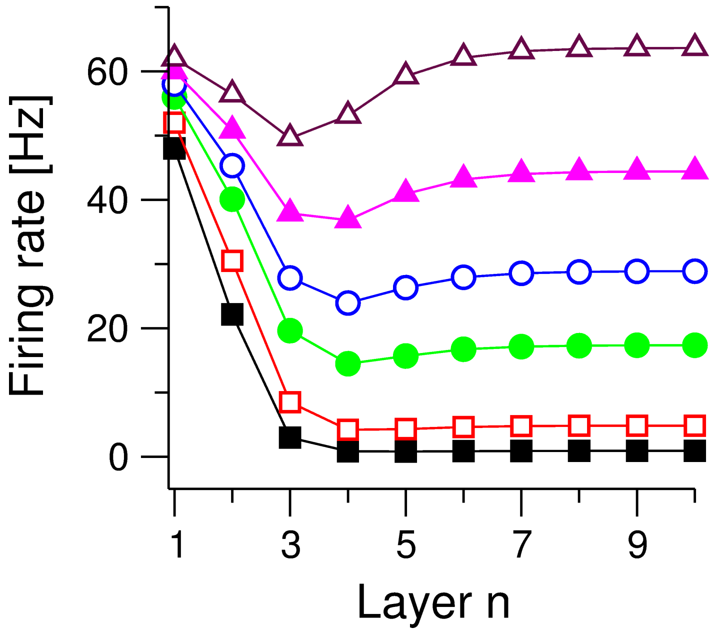
N=200

Estimated threshold is $\theta$ =5

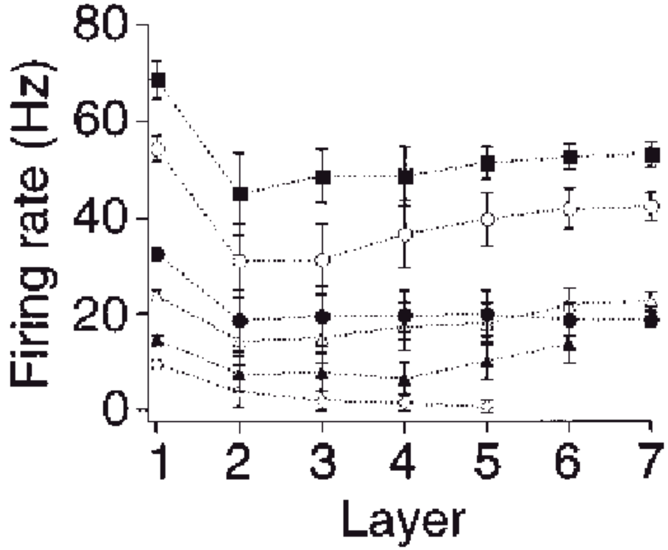
Dr. Thomas Nowotny,
Centre for Computational Neuroscience and Robotics
University of Sussex

# Firing rate as a function of the layer

McCulloch-Pitts model

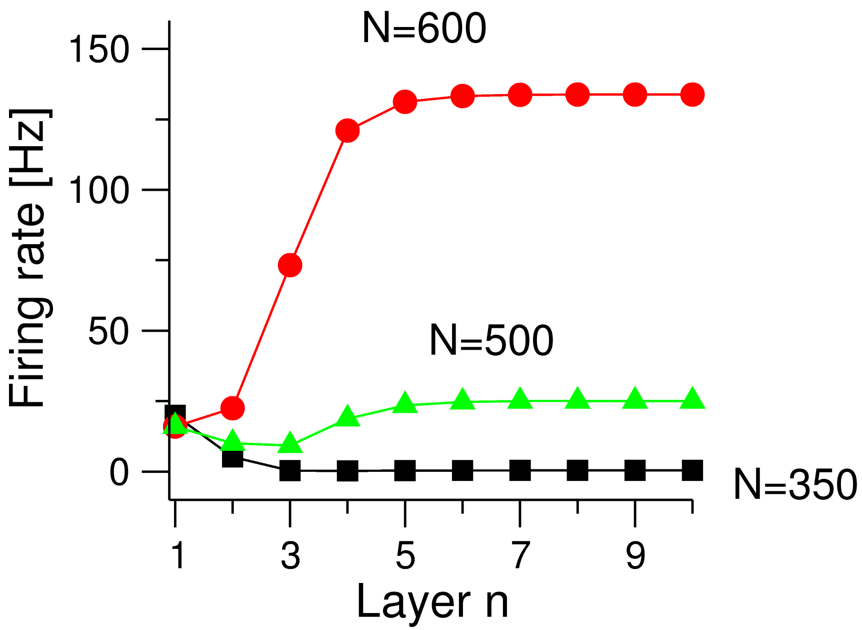



Reyes ICN

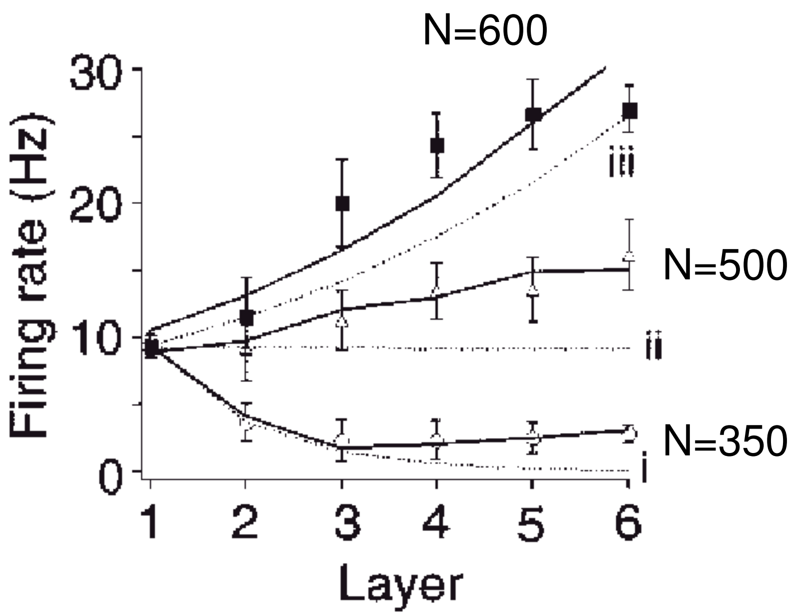


# Firing rate as function of the layer

McCulloch-Pitts model



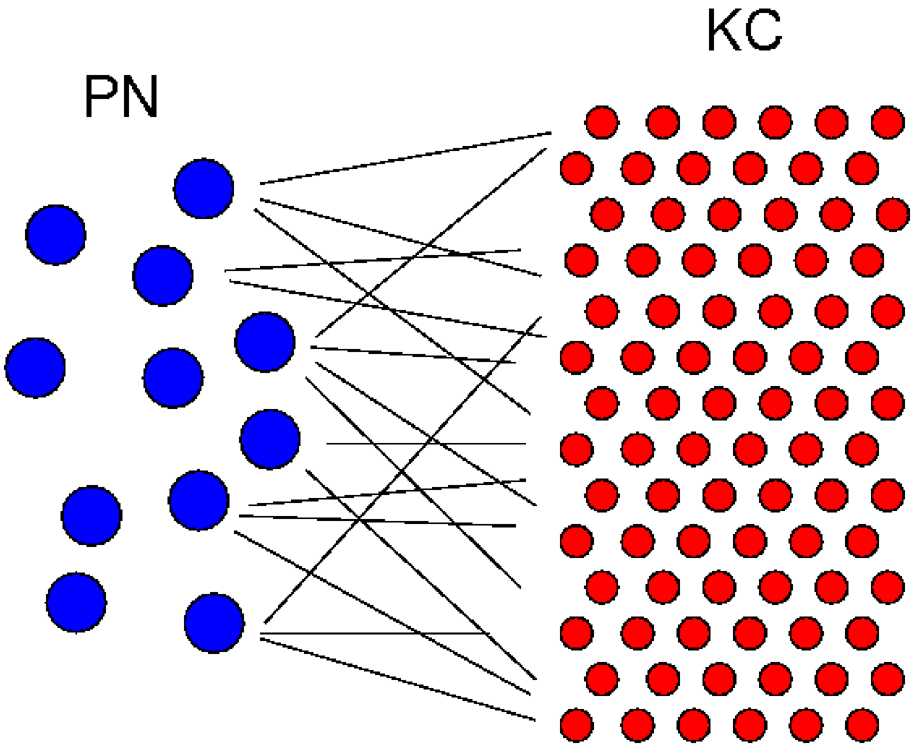

Reyes ICN




Dr. Thomas Nowotny,
Centre for Computational Neuroscience and Robotics
University of Sussex

# Application to the Insect olfactory system

# AL-MB projections are ffwd




830 PN

50000 KC

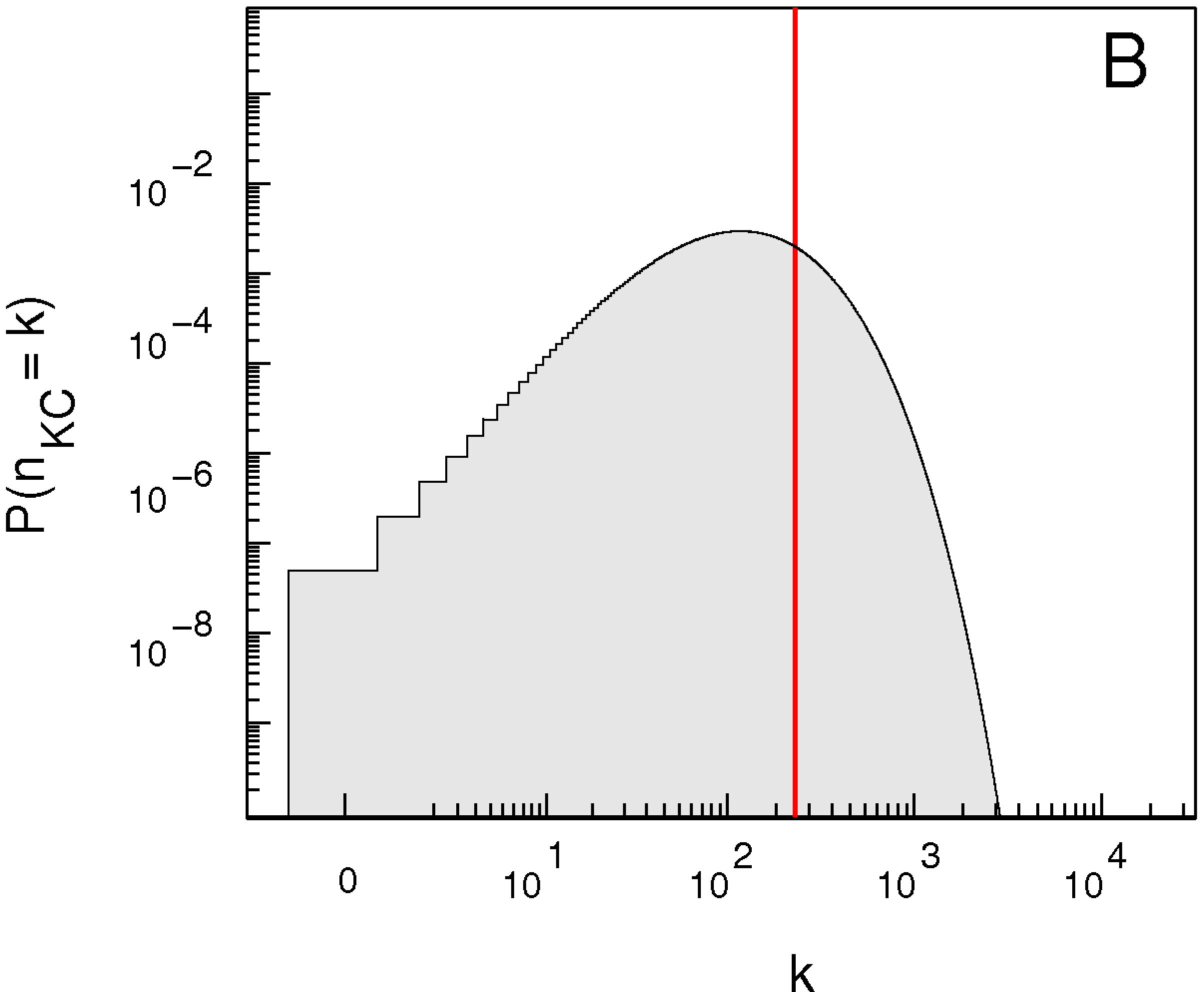
$$P(n_{\mathrm{PN}} = k) = b_{830,p_{\mathrm{PN}}}(k)$$

$$P(n_{\mathrm{KC}} = k) = ?$$

Dr. Thomas Nowotny,
Centre for Computational Neuroscience and Robotics
University of Sussex

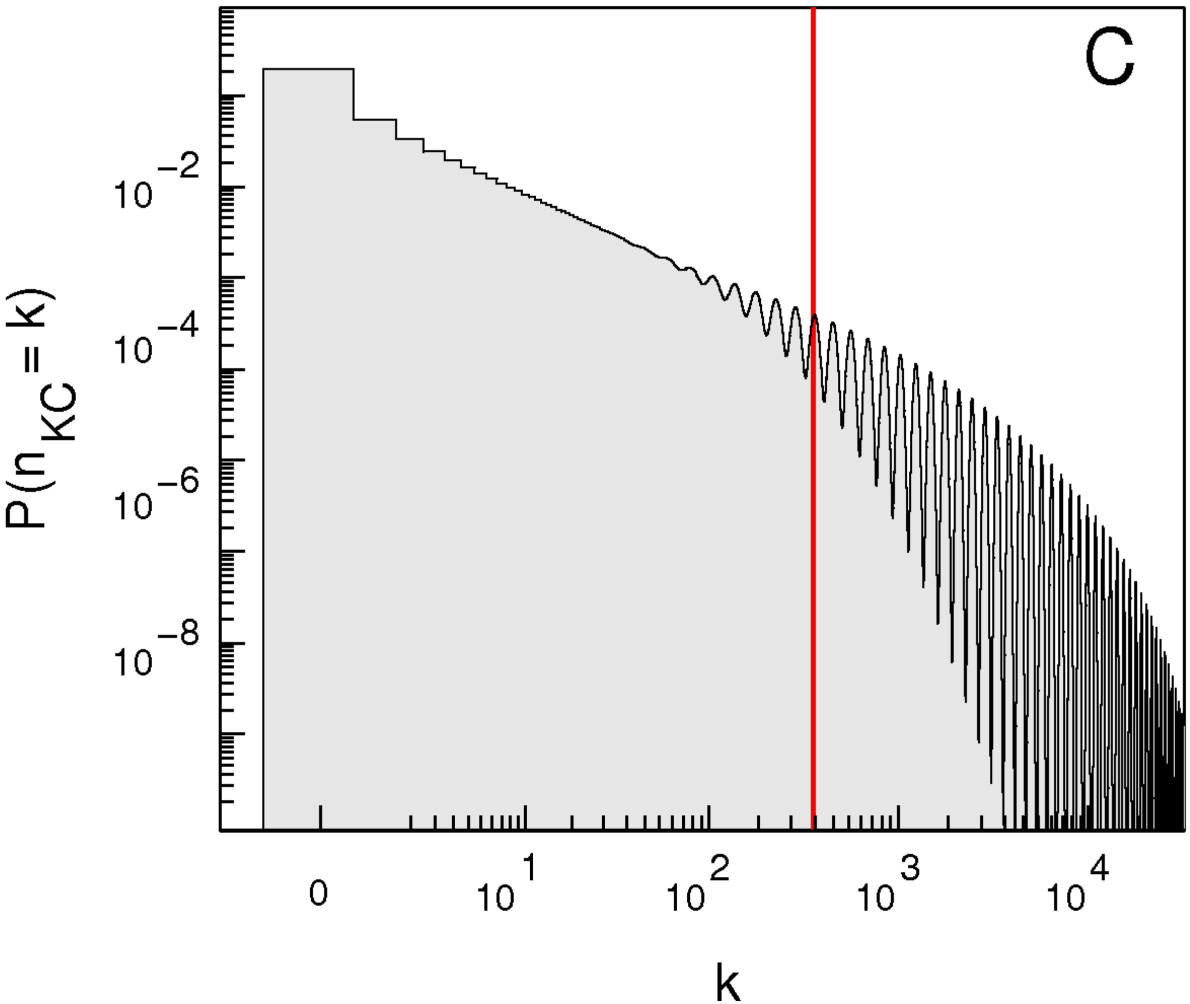
# Activity in the MB: sparse connections

$$n_{\mathrm{PN}} = 830, n_{\mathrm{KC}} = 50000, \theta_{\mathrm{KC}} = 17, p_c = 0.05, p_{\mathrm{PN}} = 0.2$$



# Activity in the MB: dense connections

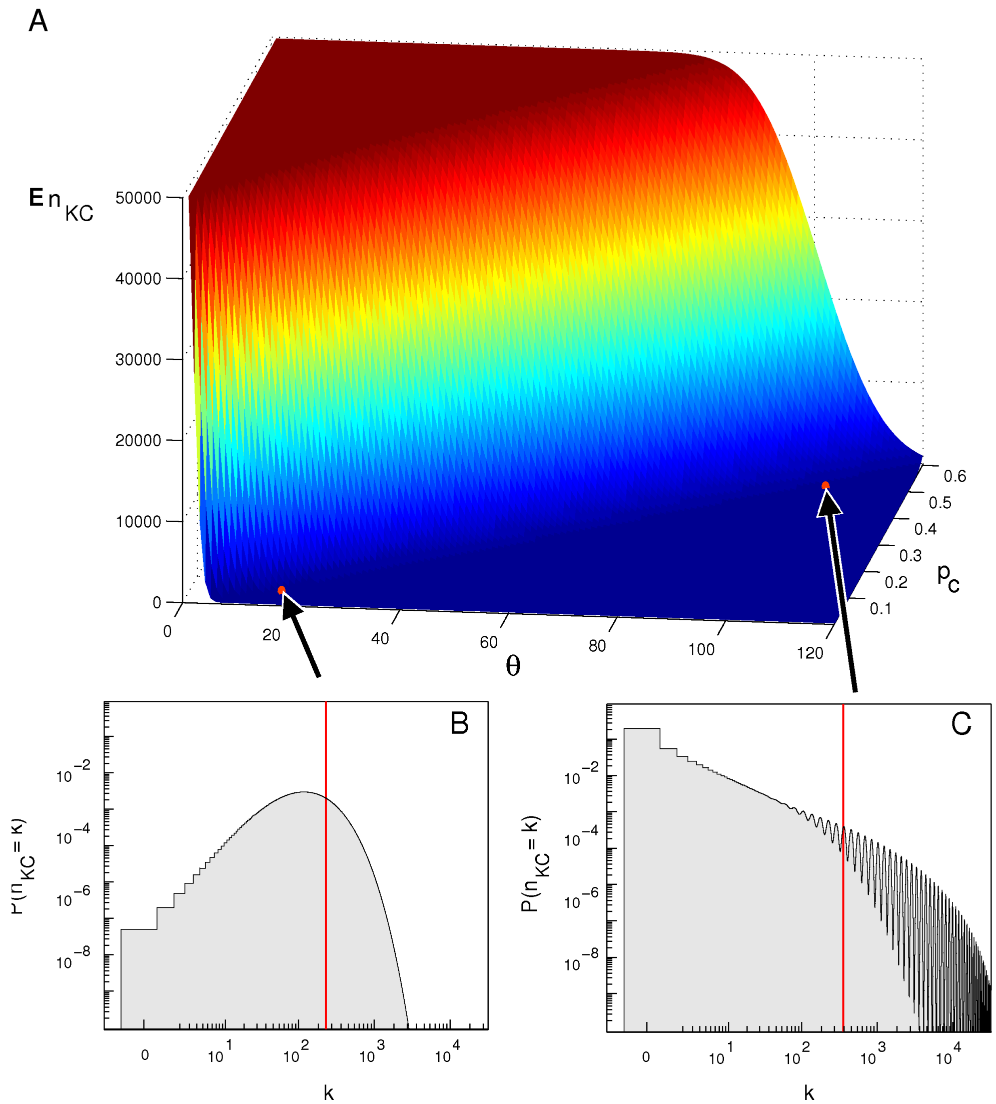
$$n_{\mathrm{PN}} = 830, n_{\mathrm{KC}} = 50000, \theta_{\mathrm{KC}} = 105, p_c = 0.5, p_{\mathrm{PN}} = 0.2$$





Dr. Thomas Nowotny,
Centre for Computational Neuroscience and Robotics
University of Sussex

# MB activity




Expectation value for the number of active KC


Dr. Thomas Nowotny,
Centre for Computational Neuroscience and Robotics
University of Sussex

# Confusion and ground state

- Similarly one can calculate the "probability of confusion"

$$P(\text{confusion}) := P(\vec{y}_1 = \vec{y}_2 \mid \vec{x}_1 \neq \vec{x}_2)$$

- And the probability of quiescence at ground state

$$P(n_{\text{KC}} \geq N_0 \mid p_{\text{PN}} = p_{\text{baseline}})$$

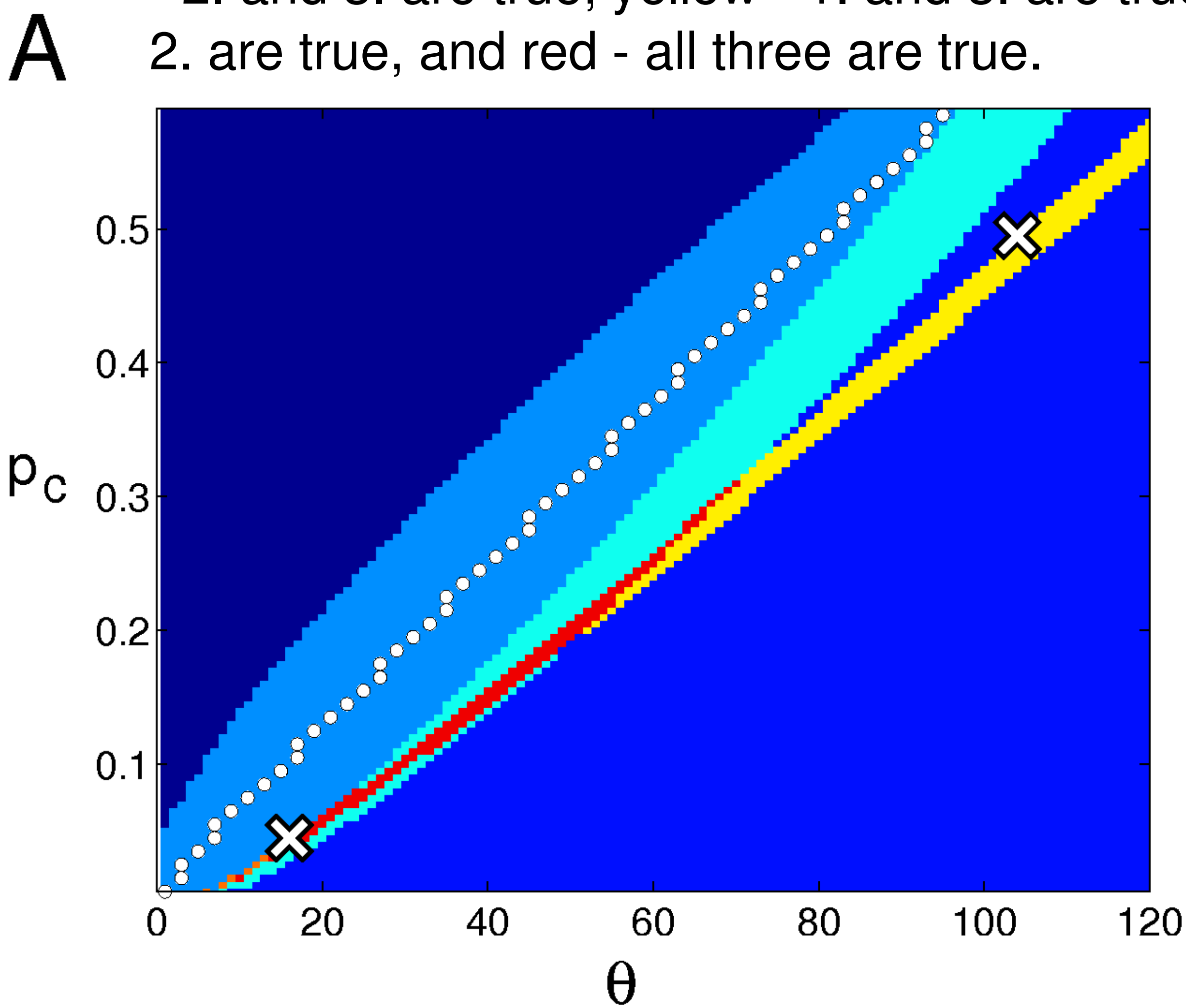
# Minimum conditions for successful operation

We can formulate minimal conditions for successful operation, e.g.:

1. $100 \leq \mathbb{E} n_{KC} \leq 500$
2. $p_{\text{confusion}} \leq 0.001$
3. $P(n_{KC} \geq 20 \mid p_{PN} = 0.13) \leq 0.01$

# Dense connections seem impossible!

Dark blue - none are fulfilled, blue - 3. is true, light blue - 2. is true, cyan - 2. and 3. are true, yellow - 1. and 3. are true, orange - 1. and 2. are true, and red - all three are true.



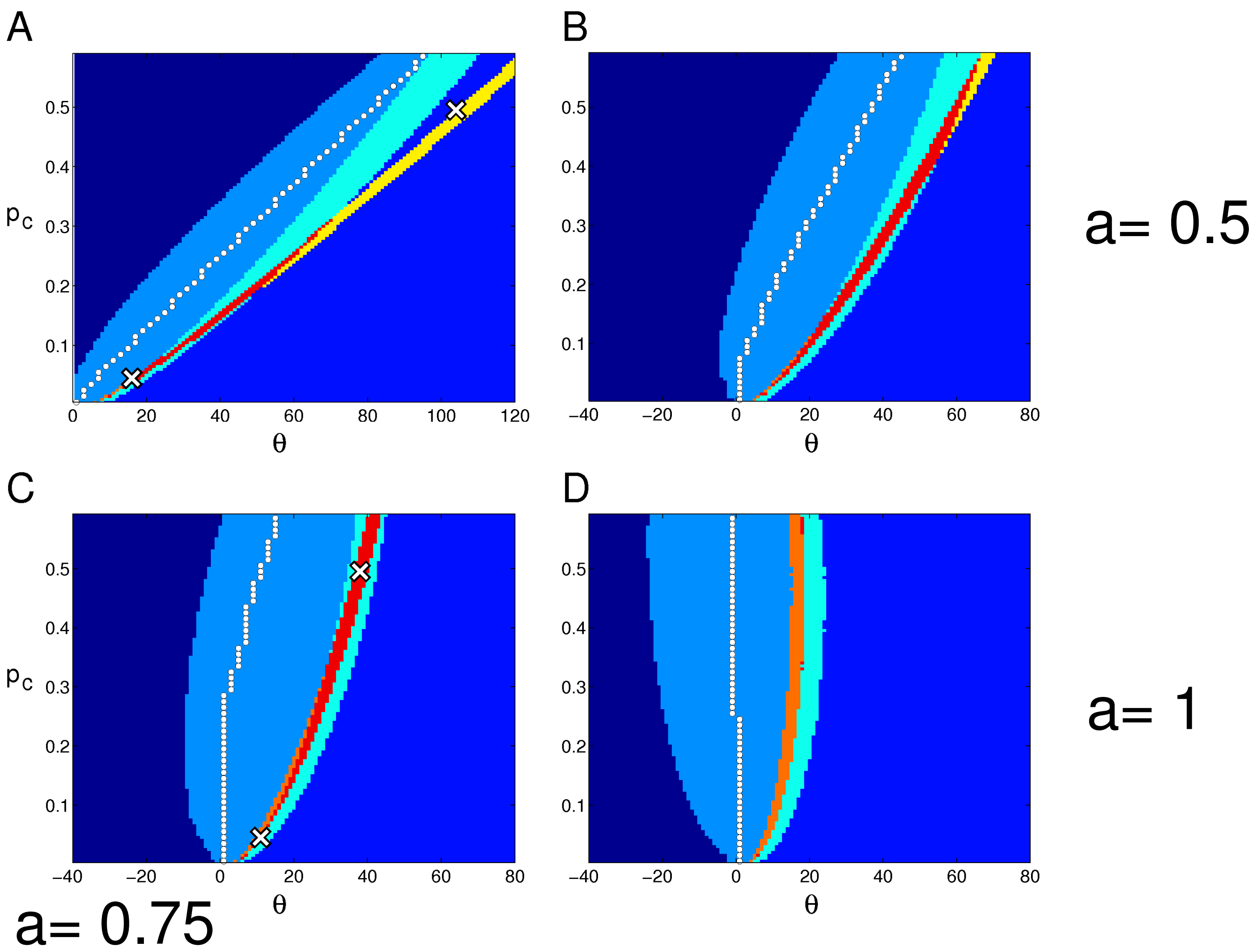

Dr. Thomas Nowotny,
Centre for Computational Neuroscience and Robotics
University of Sussex

# Fix: gain control

Substract the expected input from the input to each KC (feedforward inhibition)

$$y_i = \Theta\Big( \sum_{j=1}^{N_{\mathrm{PN}}} w_{ij} x_j - \theta - a\, p_c\, n_{\mathrm{PN}} \Big)$$

# This can be fixed by gain control




Dr. Thomas Nowotny,
Centre for Computational Neuroscience and Robotics
University of Sussex

# (Specific) Summary

- Synchrony in feedforward networks is a generic effect of connectivity
- McCulloch-Pitts approach is good enough to understand it
- For the AL-MB projections the analysis shows severe problems with dense connections
- Appropriate gain control may mediate those problems

# (General) Discussion

- McCulloch-Pitts description can be quite powerful
- It might even give interesting results in unexpected areas (here synchronization)

- On the other hand, clearly it is not for everything
- Interpretation needs to be done carefully
- Quantitative agreement is rare

- Further reading

- T. Nowotny and R. Huerta  Explaining synchrony in feedforward networks: Are McCulloch-Pitts neurons good enough? Biol Cyber **89**(4): 237-241 (2003)
- A. Reyes, Synchrony-dependent propagation of firing rate in iteratively constructed networks in vitro, Nature Neurosci. **6**:593 (2003)
- T. Nowotny et al. How are stable sparse representations achieved in fan-out systems? (in preparation)

# Next time

- Connectionist models of recognition and learning in the olfactory system of insects
- Spiking models of insect olfaction