

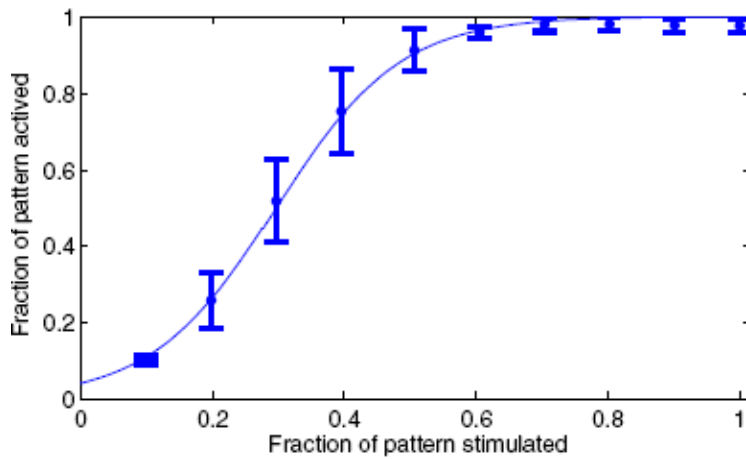
## Associative Memory—II



Potentiated synapses between neurons in a pattern enable full recall

**Fraction recalled increases faster than fraction stimulated.  
Overlapping patterns degrade recall (and memory capacity).  
Active dendrites can mitigate pattern overlap.**

## Recall performance



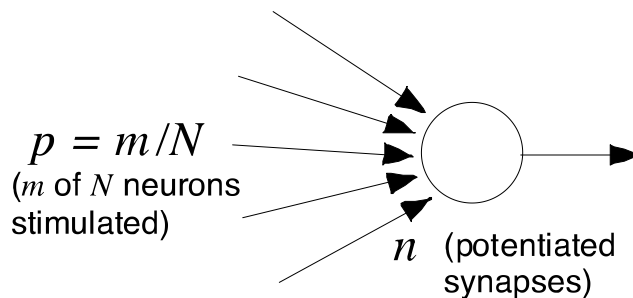
Over 90% of a pattern is recalled when half of its neurons are stimulated.

The stimulated group recruits neurons connected to it with enough potentiated synapses.

These neurons may in turn recruit more neurons—requires a high degree of connectivity.



## Simple model of recall



Fraction of potentiated synapses activated equals the fraction of pattern stimulated.

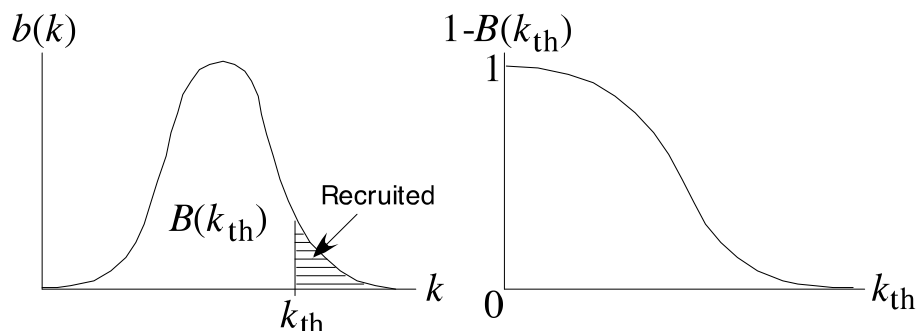
The number of a neuron's potentiated synapses that are activated,  $k$ , is modeled as a binomial distribution,  $b$ , with density:

$$b[k, n, p] = \binom{n}{k} p^k (1-p)^{n-k}$$

where  $p$  is the fraction of the pattern stimulated and  $n$  is the total number of potentiated synapses.



## Probability a neuron is recruited



A neuron fires if  $k_{th}$  or more of its potentiated synapses are activated.

The cumulative distribution,  $B[k, n, p]$ , yields the probability  $r$  that a neuron is recruited:

$$r = 1 - B[k_{th}, n, p]$$

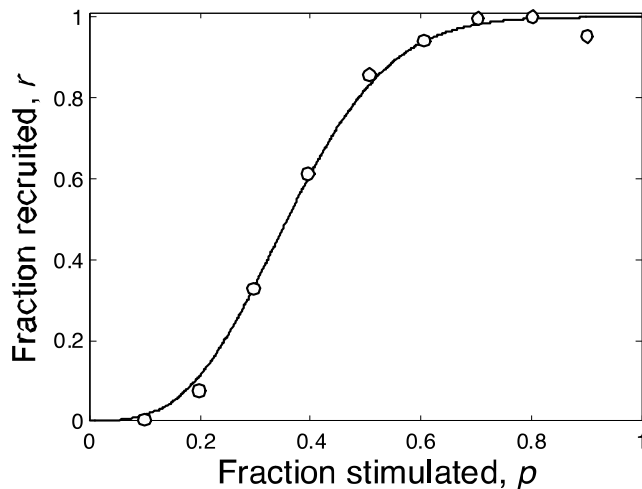
where  $k_{th}$  is the minimum number of potentiated synapses that must be activated.

Note that this result applies to feedforward as well as recurrent networks. The only difference being that a recurrent network iterates the equation—inputting a new  $p$  that includes the recruited neurons.



5 of 9

## Recruited versus stimulated fractions



Binomial distribution yields a good fit, with  $n = 16$  and  $k_{th} = 5.6$ .

Note that the stimulated ( $p$ ) and recruited ( $r$ ) fractions sum to give the fraction of the pattern recalled:

$$f = p + (1 - p) r$$

This relation was used to obtain the points plotted above ( $r$ ) from the measured data ( $f$ ):

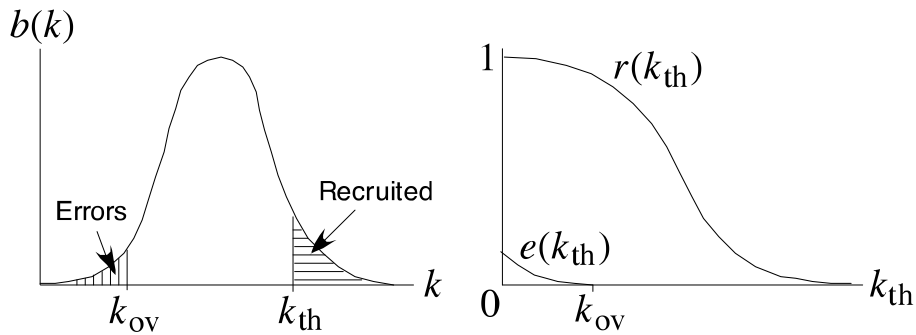
$$r = (f - p) / (1 - p)$$

The good fit suggests that iteration is insignificant—behaves like a feedforward network—most likely because connectivity is sparse (21 or less out of a 100).



6 of 9

## Overlapping patterns



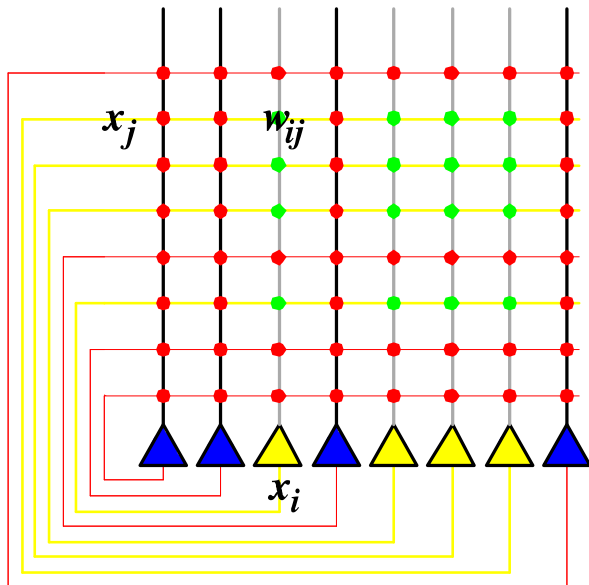
Neurons in overlapping patterns are recruited when  $k_{th} < k_{ov}$ , the pattern overlap.

When multiple patterns are learned, a neuron can become active erroneously if the pattern being recalled overlaps with the one the neuron belongs to.

Such errors occur when the overlap—the number of neurons that belong to both patterns—exceeds  $k_{th}$ .

Pattern overlap imposes a lower bound on  $k_{th}$ , which translates into an upperbound on  $r$ .

## Superposition of stored patterns



Neuron  $j$  drives neuron  $i$  through a synapse with strength  $w_{ij}$ .

In the Hopfield network, neuron  $j$ 's input is given by

$$Y_j = \sum_{i=1}^N w_{ij} x_i \quad \text{where} \quad w_{ij} = \sum_{m=1}^M x_i^m x_j^m$$

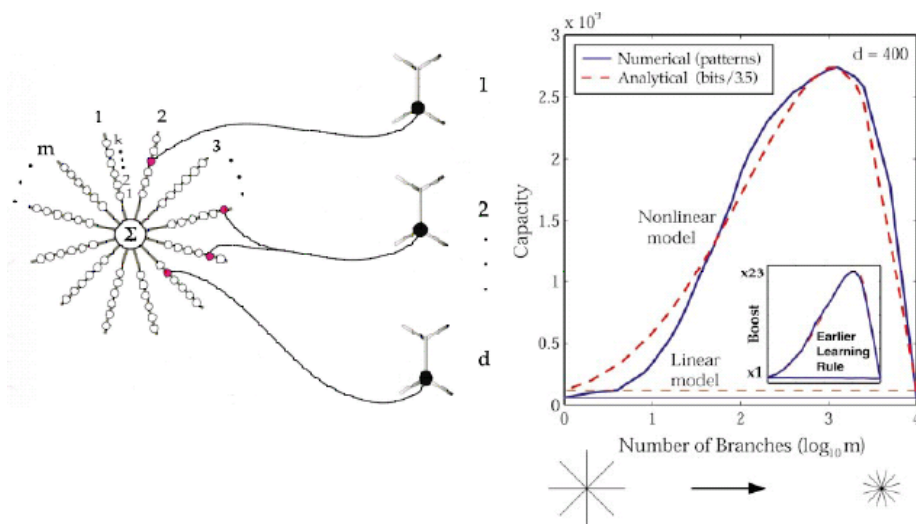
Substituting and reversing the order of summation yields:

$$Y_j = \sum_{i=1}^N \left( \sum_{m=1}^M x_i^m x_j^m \right) x_i = \sum_{m=1}^M \left( \sum_{i=1}^N x_i^m x_i \right) x_j^m$$

The inner sum is large if the neurons' activity ( $x_i$ ) is similar with that invoked by pattern  $m$  ( $x_i^m$ ). Thus, a pattern's similarity to the input determines its contribution to the output. Therefore, the network recalls a superposition of the stored patterns, each weighted by its similarity to the initial activity.

To avoid superposition, pattern overlap should be minimized, which requires that each pattern activates only a small fraction of the population (sparse coding). Then, the network will tend to recall the closest-matching pattern perfectly.

## Avoiding superposition: Active dendrites



Memory capacity increases when dendritic branches act as nonlinear subunits [Mel'01].

Another way to avoid superposition is to cluster inputs onto dendritic branches and impose a threshold that must be exceeded for a branch to become active.

The capacity of a network of neurons with such active dendrites grows logarithmically with the number of distinct ways in which the inputs can be clustered (analytical curve in figure).

For neurons with 10,000 synapses, made by  $d = 400$  distinct afferent axons, capacity was maximum when synapses were distributed across 1,250 branches with 8 synapses each.

**Next week: TBD**