

# Attractor Neural Networks with Patchy Connectivity

Christopher Johansson<sup>1</sup>, Martin Rehn<sup>1</sup>, and Anders Lansner<sup>1</sup> \*

1- Royal Institute of Technology - Dept of Numerical Analysis and Computer Science  
100 44 Stockholm - Sweden

**Abstract.** We investigate the effects of patchy (clustered) connectivity in sparsely connected attractor neural networks (NNs). This study is motivated by the fact that the connectivity of pyramidal neurons in layer II/III of the mammalian visual cortex is patchy and sparse. The storage capacity of hypercolumnar attractor NNs that use the Hopfield and Willshaw learning rules with this kind of connectivity is investigated analytically as well as by simulation experiments. We find that patchy connectivity gives a higher storage capacity, given an overall sparse connectivity and regardless of learning rule.

## 1 Introduction

Attractor neural networks (NNs) have been proposed [1, 2] as a first, abstract, model of the neocortex. Topologically, these networks are similar to the neocortex in having a large number of recurrent connections, but if they approximate the dynamics of cortex is still an open question. Recently, experimental evidence of attractor states has been found in neocortical slices [3, 4].

The mammalian visual cortex is organized in columns, in addition to its layered structure. Two types of columns have been observed, small minicolumns with a diameter of  $\sim 50\mu m$  and large hypercolumns with a diameter of more than  $200\mu m$  [5, 6]. The hypercolumns in primary visual cortex were first described by Hubel and Wiesel [7]. They are also found in the somatosensory cortex [6] where they are sometimes referred to as macrocolumns, segregates, or barrels. Hubel and Wiesel showed by electrophysiological experiments in primates that the hypercolumn can function as a competitive, winner-take-all (WTA), circuitry for line orientations [7, 8]. Lateral connections in layer II/III of the visual cortex tend to terminate in clusters (patches) with a size similar to hypercolumns [9-12]. These lateral connections are highly convergent, meaning that the neurons in one hypercolumn connect to the same set of hypercolumns forming a patchy pattern of connections [11, 12]. Connectivity is sparse in the neocortex, which is a necessity to keep down the amount of wiring.

We here use a model of neocortex in which the minicolumns are the functional units [5, 6]. The minicolumns are grouped into hypercolumns that implement a WTA function over the minicolumns, controlling the activity level in the network. In the next section we describe hypercolumnar attractor NNs that use the Hopfield and Willshaw learning rules. The effect of clustered connectivity in these NNs is then investigated both analytically, with signal-to-noise analysis that is valid because of the sparse activity [13], and by simulations. Clustering is implemented such that in the

---

\* This work was financially supported by Vetenskapsrådet grant no. 612-2003-6256.

extreme case of fully clustered connectivity, two hypercolumns are either fully interconnected (at the minicolumn level) or not at all. We contrast this to random connectivity, where individual minicolumns are independently wired. We further distinguish three modes of clustered connectivity. (I) Clustering applies to incoming connections; one minicolumn tends to receive connections from either all units in a hypercolumn or from none. (II) Clustering applies to the outgoing connections; one minicolumn tends to project either to all minicolumns in a hypercolumn or to none. (III) The first two modes are combined, making connections between hypercolumns reciprocal i.e. fully clustered connectivity.

### 1.1 Attractor Neural Networks with Hypercolumns

The Hopfield learning rule leads to real-valued synaptic weights, whereas the Willshaw learning rule gives binary weights. The NN has  $N$  units grouped into  $H$  hypercolumns with  $U$  units in each (for simplicity we let all hypercolumns consist of the same number of units). Throughout the paper we will use  $H=U=40$ . In a NN with a fully ordered connectivity (100% clustering) all connections are located in a fraction of the  $H^2$  potential hypercolumn-blocks (Fig. 1, left). In the case of random connectivity the connections are randomly distributed between all units (Fig. 1, right). In the center of Fig. 1 an intermediate between these two extremes is shown. For all levels of clustering the number of inputs to a unit is constant.

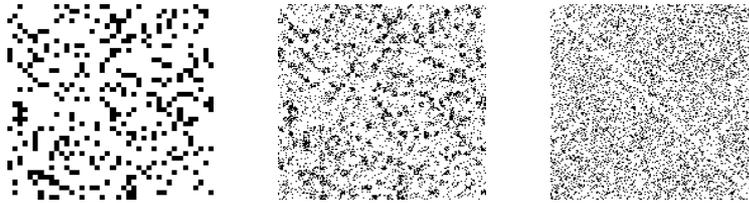


Fig. 1: Three connectivity matrices, where black represents the existence of a connection between two units; left 100% clustering, center 50% clustering, and right 0% clustering. All matrices are derived from the 100% clustered matrix. The average connectivity in all three matrices is 20%. No self-connections within a hypercolumn.

The NNs are trained with  $P$  unary coded random patterns,  $\xi^\mu$ . The activity of one unit in each hypercolumn is set to 1, the activity of the others to 0. In the following we present the Hopfield and Willshaw learning rules, where  $a=1/U$  is the mean activity:

$$\text{Hopfield: } w_{ij} = \frac{1}{N} \sum_{\mu} (\xi_i^{\mu} - a)(\xi_j^{\mu} - a), \quad \text{Willshaw: } w_{ij} = \bigvee_{\mu} \xi_i^{\mu} \wedge \xi_j^{\mu}$$

The retrieval procedure, which is based on WTA within each hypercolumn, is the same for both NNs, where  $s$  is the support of a unit,  $o$  is its activity,  $h$  is the index of a particular hypercolumn,  $Q_h$  is the set of all units belonging to hypercolumn  $h$  and  $y_{ij}$  are binary variables indicating which of the possible synaptic connections are present.

$$s_j = \sum_{i=1}^N w_{ij} y_{ij} o_i, \quad o_j \leftarrow \begin{cases} 1 & \text{if } j = \arg \max_{k \in Q_h} s_k \\ 0 & \text{otherwise} \end{cases} \quad \text{for each } h = \{1, \dots, H\}$$

## 1.2 Analysis of the Storage Capacity for the Hopfield Learning Rule

Our network uses either the Hopfield or Willshaw learning rule, masked by a matrix  $Y_{ij}$  describing which of the potential connections are present. The synaptic weight matrix will be given as  $Z_{ij}=X_{ij}Y_{ij}$  where  $X_{ij}$  is the weight matrix given by the chosen learning rule. We introduce a clustering parameter  $C$  running from 0 for the randomly connected network to 1 for the fully patchy connectivity. In the latter case we have a block structure; each unit receives input from  $K$  full hypercolumns and from those only. The masking matrix is split into a “block” part  $Y_{ij}^A$  and a “non-block” part  $Y_{ij}^B$ . For the fully clustered,  $C=1$ , case the probability for nonzero entries is one in the former and zero in the latter. In general, the masking is given by stochastic variables  $[\Pr(Y_{ij}^A=1)=p_A; \Pr(Y_{ij}^A=0)=1-p_A]$  where  $p_A=C+(1-C)d$  and  $[\Pr(Y_{ij}^B=1)=p_B; \Pr(Y_{ij}^B=0)=1-p_B]$  where  $p_B=(1-C)d$ . Here,  $d$  is the average connectivity. It follows, as it should that,  $Y_{ij}^A$  and  $Y_{ij}^B$  have the same distribution when  $C=0$ .

We now proceed to calculate the  $X_{ij}$  for the Hopfield learning rule. These weights  $X_{ij}$  can be expressed as a sum over  $P$  stochastic variables, one for each training pattern;  $X_{ij} = \frac{1}{N} \sum_{\mu}^P X_{ij}^{\mu}$ . We have three possible outcomes for the contributions from individual patterns:  $[\Pr(X_{ij}^{\mu} = (1 - \frac{1}{U})^2) = \frac{1}{U^2}; \Pr(X_{ij}^{\mu} = \frac{1}{U^2} - \frac{1}{U}) = \frac{2(U-1)}{U^2}; \Pr(X_{ij}^{\mu} = \frac{1}{U^2}) = \frac{(U-1)^2}{U^2}]$ . We will use a normal approximation for the support values for each unit and therefore we calculate the mean and variance of  $X_{ij}^{\mu}$ , yielding  $E(X_{ij}^{\mu})=0$ ,  $V(X_{ij}^{\mu})=(U-1)^2/U^4$ . The support,  $S_i$ , that each unit receives when pattern  $\xi^{\mu}$  is applied as input is

$$S_i = \sum_j^N w_{ij} y_{ij} \xi_j^{\mu} = \sum_{k=j:\xi_j^{\mu}=1}^{H-1} Y_{ik} X_{ik} = \frac{1}{N} \sum_{k=j:\xi_j^{\mu}=1}^{H-1} Y_{ik} \sum_{\mu}^P X_{ik}^{\mu} = \sum_{k=j:\xi_j^{\mu}=1}^{H-1} Z_{ik}$$

We now treat separately the cases for units participating, denoted by +, and not participating, denoted by -, in the active pattern  $\xi^{\mu}$ . Next we compute the mean and variance for the four different cases of  $Z_{ij}$  as  $E(Z_{ij}^{A/B,+/-})=E(Y_{ij}^{A/B} X_{ij}^{+/-})$  and  $V(Z_{ij}^{A/B,+/-})=V(Y_{ij}^{A/B} X_{ij}^{+/-})$ ; introducing  $x:=(1-1/U)^2$ .

$$E(Z_{ij}^{A/B,-}) = p_{A/B} P E(X_{ij}^{\mu}) / N; \quad V(Z_{ij}^{A/B,-}) = p_{A/B} P V(X_{ij}^{\mu}) / N^2$$

$$E(Z_{ij}^{A/B,+}) = p_{A/B} (x + (P-1)E(X_{ij}^{\mu})) / N;$$

$$V(Z_{ij}^{A/B,+}) = p_{A/B} ((1-p_{A/B})x^2 p_{A/B} + (x - x p_{A/B})^2 + (P-1)V(X_{ij}^{\mu})) / N^2$$

In the + case these are computed from a bimodal distribution with centers at 0 and  $x$ . The mean and variance of  $S_i^-$  and  $S_i^+$  are then computed using the central limit theorem:

$$E(S_i^-) = \sum_{k=j:\xi_j^{\mu}=1}^{H-1} E(Z_{ik}^{A,-}) + E(Z_{ik}^{B,-}) = 0; \quad V(S_i^-) = \sum_{k=j:\xi_j^{\mu}=1}^{H-1} V(Z_{ik}^{A,-}) + V(Z_{ik}^{B,-}) = \frac{K P V(X_{ij}^{\mu})}{N^2}$$

$$E(S_i^+) = \sum_{k=j:\xi_j^{\mu}=1}^{H-1} E(Z_{ik}^{A,+}) + E(Z_{ik}^{B,+}) = \frac{K x}{N}$$

$$V(S_i^+) = \sum_{k=j:\xi_j^{\mu}=1}^{H-1} V(Z_{ik}^{A,+}) + V(Z_{ik}^{B,+}) =$$

$$\frac{Kx^2(-C^2H + H + KC^2 - K - 1 + C^2) + KV(X_{ij}^\mu)(H-1)}{N^2(H-1)}$$

Let  $\Phi_{S^-}$  be the normal distribution function of the stochastic variable  $S_j^-$  and  $F_N$  the cumulative distribution function of a variable  $N = \max_{i \neq j} \{S_j^-\}$  where  $i$  is the unit that participates in the active pattern, then:

$$F_N(n) = P(n \geq \max_{j \in Q_h, j \neq i} \{S_j^-\}) = \Phi_{S^-}^{U-1}(n)$$

Now we may write the probability that the unit participating in the active pattern has the largest support in a hypercolumn as;

$$p_{hyper} = P(S^+ > N) = \int_{n=-\infty}^{\infty} f_{S^+}(n) F_N(n) dn$$

where  $f$  denotes probability density functions and  $F$  cumulative probability density functions. The probability that all hypercolumns in the NN will have the same units active as those in the original, noise free, pattern is  $p_{pattern} = (p_{hyper})^H$ . The average number of correctly retrieved patterns is  $Pp_{pattern}$ , where  $P$  is the number of trained patterns. If  $N$  is approximated by a normal distribution,  $p_{hyper}$  can be computed as  $\{Z = S^+ - N, p_{hyper} = 1 - P(Z \leq 0) = \Phi(-m_Z/\sigma_Z)\}$  which requires less computations than evaluating the integral form above.

### 1.3 Analysis of the Storage Capacity for the Willshaw Learning Rule

The probability that a certain synapse in the  $X_{ij}$  matrix is nonzero, when storing a single pattern using the Willshaw learning rule is  $p_0 = 1/U^2$ . After storing  $P$  patterns, the density of ones in the memory matrix is  $p_1 = 1 - (1 - p_0)^P$ .

We now consider the stability of patterns. Just as in the Hopfield case, we first study one hypercolumn. We calculate the support level of the unit that is part of the active pattern,  $S^+$ , and that of the other units,  $S_i^-$ ,

$$S^+ = \sum_{h=1}^K Z_h^{A+} + \sum_{h=1}^{H-1-K} Z_h^{B+}, \quad S^- = \sum_{h=1}^K Z_h^{A-} + \sum_{h=1}^{H-1-K} Z_h^{B-}$$

where the  $Z^A$  are the contributions from the "blocks" and the  $Z^B$  are from the "non-blocks". These stochastic variables,  $Z_h$ , are in fact the same as  $Z_{ij} \cdot \zeta_j^\mu = 1$  and they take the value 1 when a connection exists and both the pre- and postsynaptic units have been activated simultaneously in one of the patterns. The latter is always true for the "+" units, in the active pattern. Assuming independence between the "blocks" and "non-blocks", the sums become binomial distributions:

$$S^+ \in \text{Bin}(K, p_A) + \text{Bin}(H-1-K, p_B), \quad S^- \in \text{Bin}(K, p_A p_1) + \text{Bin}(H-1-K, p_B p_1)$$

Similar to the case of the Hopfield learning rule we can compute the probability for stable recall in one hypercolumn

$$F_N = \left(F_{S^-}\right)^{U-1}, \quad p_{hyper} = P(S^+ > N) = \sum_{n=0}^{\infty} f_{S^+}(n) F_N(n)$$

and the probability that all hypercolumns are stable is  $p_{pattern} = (p_{hyper})^H$ , which is also the expected ratio of stable patterns for a given memory load. As in the previous section we can now determine the storage capacity as the maximum storable patterns  $Pp_{pattern}$ .

## 2 Results

In Fig. 2 (left) experimental data on the storage capacity for two different types of learning rules is shown together with analytical estimates. For both learning rules the positive effect of clustered connections is clearly seen. This effect was found to be most prominent for sparse connectivity and it was also present in a standard network using global KWTA thresholding. The storage capacity was tested with retrieval cues based on the stored patterns, in which the activity of 20% of the hypercolumns had been randomly changed. The relaxation process stopped when a stable pattern was found or when more than 20 iterations had passed. A pattern was classified as retrieved if it had a 100% match with the stored pattern. It is worth noting that although the connection matrices were asymmetric the NNs converged to a fix-point in almost all retrievals.

In Fig. 2 (right) three different types of connectivity matrices are compared. Case I, where all units in a hypercolumn connects to a single unit, gives a NN with a storage capacity very similar to that of case III when entire blocks of connections are used. Case II, where a single unit connects to an entire hypercolumn, which results in a NN with a storage capacity smaller than that of either case III and I.

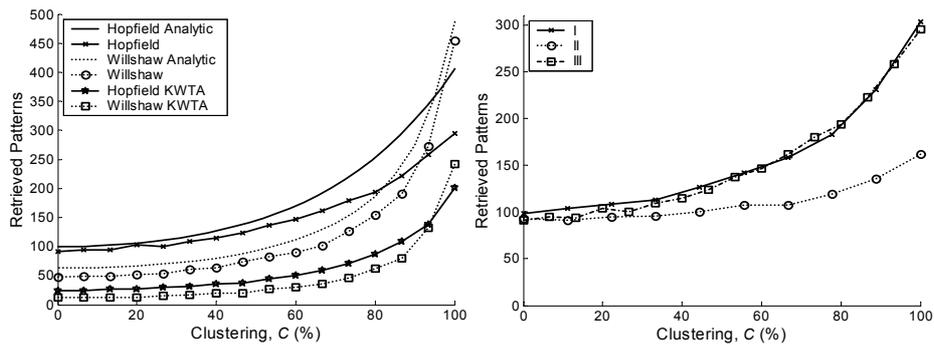


Fig. 2: (Left) The experimentally and analytically derived storage capacities plotted as a function of the clustering. (Right) The storage capacity when using the Hopfield learning rule is plotted for three different connectivity matrices. Case III is the type of connectivity matrices shown in Fig. 1, in case I a unit receives input from an entire hypercolumn at 100% clustering, and in case II a unit connects to an entire hypercolumn. The analytical estimate of the storage capacity when the Hopfield learning rule is used is based on a normal approximation which is accurate in the limit  $N, H \rightarrow \infty$ . Parameters:  $N=1600$ ,  $H=40$ ,  $U=40$ ,  $d=20\%$ .

## 3 Discussion

In this paper we have shown that an attractor NN with sparse and patchy (clustered) connectivity has a larger storage capacity than one with sparse and random connectivity. We demonstrate that this is due to a better signal to noise ratio of the support values. Based on this finding we suggest that the patchy connectivity seen in the visual cortex [9-12] have a computational advantage over a random connectivity.

This adds to previous physical arguments for patchy connectivity such that it minimizes wiring (and thus volume) while maximizing connectivity. Taken together, it is obvious that a clustered connectivity is very advantageous.

We found that the most important aspect of the clustered connectivity is that all units, in a hypercolumn, connects to the same postsynaptic units (case I). This type of connectivity resulted in a signal to noise ratio similar to that of block wise connectivity (case III). Furthermore, this type of patchy connectivity between hypercolumns has been found in the visual cortex by experiments with retrograde tracers [11, 12].

Given the physical and computational benefits of a clustered connectivity it seems very plausible that this is a generic design principle that also applies to other parts of cortex. A columnar structure is found also in e.g. the somatosensory cortex, but in associative cortical areas it is less apparent. We argue that a columnar structure and patchy connectivity is a general design principle in neocortex.

## References

- [1] E. T. Rolls and A. Treves, *Neural Networks and Brain Function*. New York: Oxford University Press, 1998.
- [2] G. Palm, "Towards a Theory of Cell Assemblies," *Biol. Cybern.*, vol. 39, pp. 181-194, 1981.
- [3] Y. Shu, A. Hasenstaub, and D. A. McCormick, "Turning on and off recurrent balanced cortical activity," *Nature*, vol. 423, pp. 288-293, 2003.
- [4] R. Cossart, D. Aronov, and R. Yuste, "Attractor dynamics of network UP states in the neocortex," *Nature*, vol. 423, pp. 283-288, 2003.
- [5] D. P. Buxhoeveden and M. F. Casanova, "The minicolumn hypothesis in neuroscience," *Brain*, vol. 125, pp. 935-951, 2002.
- [6] V. B. Mountcastle, "The columnar organization of the neocortex," *Brain*, vol. 120, pp. 701-722, 1997.
- [7] D. H. Hubel and T. N. Wiesel, "Functional architecture of macaque monkey visual cortex," *Proc. R. Soc. Lond. B.*, vol. 198, pp. 1-59, 1977.
- [8] B. Cürüklü and A. Lansner, "An Abstract Model of a Cortical Hypercolumn," presented at In Proc. of the 9th International Conference on Neural Information Processing, Singapore, 2002.
- [9] W. H. Bosking, Y. Zhang, B. Schofield, and D. Fitzpatrick, "Orientation Selectivity and the Arrangement of Horizontal Connections in Tree Shrew Striate Cortex," *J. Neurosci.*, vol. 17, pp. 2112-2127, 1997.
- [10] C. D. Gilbert and T. N. Wiesel, "Columnar specificity of intrinsic horizontal and corticocortical connections in cat visual cortex," *J. Neurosci.*, vol. 9, pp. 2432-2442, 1989.
- [11] B. Conway, J. D. Boyd, T. H. Stewart, and J. A. Matsubara, "The Projection from V1 to Extrastriate Area 21a: A Second Patchy Efferent Pathway that Colocalizes with the CO Blob Columns in Cat Visual Cortex," *Cerebral Cortex*, vol. 10, pp. 149-159, 2000.
- [12] J. Boyd and J. Matsubara, "Intrinsic connections in cat visual cortex: a combined anterograde and retrograde tracing study," *Brain Research*, vol. 560, pp. 207-215, 1991.
- [13] A. A. Frolov, D. Husek, and I. P. Muraviev, "Informational Capacity and Recall Quality in Sparsely Encoded Hopfield-like Neural Network: Analytical Approaches and Computer Simulation," *Neural Networks*, vol. 10, pp. 845-855, 1997.