STRUCTURE AND PLASTICITY POTENTIAL OF NEURAL NETWORKS IN THE CEREBRAL CORTEX

A dissertation presented

by

Tarec Edmond Fares

to
The Department of Physics

In partial fulfillment of the requirements for the degree of
Doctor of Philosophy

in the field of
Physics

Northeastern University
Boston, Massachusetts
June 2010
STRUCTURE AND PLASTICITY POTENTIAL
OF NEURAL NETWORKS IN THE CEREBRAL CORTEX

by

Tarec Edmond Fares

ABSTRACT OF DISSERTATION

Submitted in partial fulfillment of the requirement
for the degree of Doctor of Philosophy in Physics
in the Graduate School of Arts and Sciences of
Northeastern University, June, 2010
Abstract

Learning and memory formation in the brain depend on the plasticity of neural circuits. In the adult and developing cerebral cortex, this plasticity can result from the formation and elimination of dendritic spines. New synaptic contacts appear in the neuropil where the gaps between axonal and dendritic branches can be bridged by dendritic spines. Such sites are termed potential synapses. In this thesis, we first described a theoretical framework for the analysis of spine remodeling plasticity. We provided a quantitative description of two models of spine remodeling in which the presence of a bouton is either required or not for the formation of a new synapse. We derived expressions for the density of potential synapses in the neuropil, the connectivity fraction, which is the ratio of actual to potential synapses, and the number of structurally different circuits attainable with spine remodeling. We calculated these parameters in mouse occipital cortex, rat CA1, monkey V1, and human temporal cortex. We found that on average a dendritic spine can choose among 4-7 potential targets in rodents and 10-20 potential targets in primates. The neuropil’s potential for structural circuit remodeling is highest in rat CA1 (7.1-8.6 bits/µm$^3$) and lowest in monkey V1 (1.3-1.5 bits/µm$^3$). We also evaluated the lower bound of neuron selectivity in the choice of synaptic partners. Post-synaptic excitatory neurons in rodents make synaptic contacts with more than 21-30% of pre-synaptic axons encountered with new spine growth. Primate neurons appear to be more selective, making synaptic connections with more than 7-15% of encountered axons.

We next studied the role neuron morphology plays in defining synaptic connectivity. As previously stated it is clear that only pairs of neurons with closely positioned axonal and dendritic branches can be synaptically coupled. For excitatory neurons in the cerebral cortex,
such axo-dendritic oppositions, or potential synapses, must be bridged by dendritic spines to form synaptic connections. To explore the rules by which synaptic connections are formed within the constraints imposed by neuron morphology, we compared the distributions of the numbers of actual and potential synapses between pre- and post-synaptic neurons forming different laminar projections in rat barrel cortex. Quantitative comparison explicitly ruled out the hypothesis that individual synapses between neurons are formed independently of each other. Instead, the data are consistent with a cooperative scheme of synapse formation, where multiple-synaptic connections between neurons are stabilized, while neurons that do not establish a critical number of synapses are not likely to remain synaptically coupled.

In the above two projects, analysis of potential synapse numbers played an important role in shaping our understanding of connectivity and structural plasticity. In the third part of this thesis, we shift our attention to the study of the distribution of potential synapse numbers. This distribution is dependent on the details of neuron morphology and it defines synaptic connectivity patterns attainable with spine remodeling. To better understand how the distribution of potential synapse numbers is influenced by the overlap and the shapes of axonal and dendritic arbors, we first analyzed uniform disconnected arbors generated \textit{in silico}. The resulting distributions are well described by binomial functions. We used a dataset of neurons reconstructed in 3D and generated the potential synapse distributions for neurons of different classes. Quantitative analysis showed that the binomial distribution is a good fit to this data as well. All distributions considered clustered into two categories, inhibitory to inhibitory and excitatory to excitatory projections. We showed that the distributions of potential synapse numbers are universally described by a family of single parameter ($p$) binomial functions, where $p = 0.08$, and for the inhibitory and $p = 0.19$ for the excitatory projections.
In the last part of this thesis an attempt is made to incorporate some of the biological constraints we considered thus far, into an artificial neural network model. It became clear that several features of synaptic connectivity are ubiquitous among different cortical networks: (1) neural networks are predominately excitatory, containing roughly 80% of excitatory neurons and synapses, (2) neural networks are only sparsely interconnected, where the probabilities of finding connected neurons are always less than 50% even for neighboring cells, (3) the distribution of connection strengths has been shown to have a slow non-exponential decay. In the attempt to understand the advantage of such network architecture for learning and memory, we analyzed the associative memory capacity of a biologically constrained perceptron-like neural network model. The artificial neural network we consider consists of robust excitatory and inhibitory McCulloch and Pitts neurons with a constant firing threshold. Our theoretical results show that the capacity for associative memory storage in such networks increases with an addition of a small fraction of inhibitory neurons, while the connection probability remains below 50%. Though, the theoretical distribution of connection strengths is shown to be made of truncated Gaussian functions, in practice, results of finite-time numerical simulations based on the perceptron learning rule exhibit slow non-exponential decay.
Acknowledgments

First and foremost I want to thank my advisor Professor Armen Stepanyants for his guidance, support, and care throughout my PhD experience. I appreciate all his time, and patience which made my road to the PhD all the less harsh. He introduced me to the world of computational neuroscience and with his enthusiasm, clarity of ideas, and skill to explain things clearly, made it even more fun. Thank you for a great three research filled years!

I would like to thank all the teachers and professors that made my stay in Northeastern enjoyable, intellectually stimulating, and fun filled experience of learning and growing.

I am grateful to all my colleagues and wonderful friends that I have made, with whom I share the long winding road to the PhD! Their company, great discussions, and camaraderie have made the road all the more enjoyable.

I wish to thank all the wonderful staff and secretaries of the department of Physics, those who have left and who are still here, for all their help and assistance they have provided.

I owe my deepest gratitude to my parents who, since my childhood, have raised me, taught me, aided me, and facilitated my interest and curiosity in science, without which I would not have succeeded.

Lastly, and most importantly, I am eternally indebted to my wife, Hiba, for her love and support on our long journey together. Her presence by my side, away from home, has made it all possible! To you this thesis is dedicated!
5 Effect of biological constraints on associative memory storage capacity of artificial neural networks 58
  5.1 Introduction ................................................................. 58
  5.2 The model ................................................................. 61
  5.3 The solution ............................................................... 64

6 Conclusions 71

Appendix A 75
Appendix B 88
Appendix C 97
Appendix D 100
Bibliography 117
List of Figures

2.1 Potential synapse ................................................................. 17
2.2 Probability of potential connection for line segments .................. 19
2.3 Characterization of structural spine remodeling ......................... 26
2.4 Structural entropy per spine as a function of connectivity parameter, \(f^*\) .... 30
2.5 Potential for structural plasticity .............................................. 31
3.1 Potential and actual synapses ................................................... 36
3.2 Distributions of dendritic path lengths from actual and potential synapses to the soma are not statistically different ...................... 38
3.3 Individual synapses between pre- and post-synaptic neurons are not formed independently of each other ........................................... 42
3.4 Synapses between pre- and post-synaptic neurons are formed in a cooperative Manner .......................................................... 46
4.1 Examples of neurons reconstructed in 3D from cat and rat cortices .... 50
4.2 Potential connectivity for uniformly distributed arbors .................. 51
4.3 Potential synapse numbers for uniformly distributed segments are well described by a Poisson distribution .................................... 52
4.4 Distributions of potential synapse numbers for uniformly distributed segments ...... 53
4.5 Potential connectivity between neuronal arbors reconstructed in 3D ................ 54
4.6 Potential synapse numbers for neurons reconstructed in 3D are well described with binomial distributions.

4.7 The distribution of potential synapse numbers for reconstructed neurons is well with a binomial function.

5.1 Perceptron models and their critical capacity.

5.2 Critical capacity of the presented model.

5.3 Dependence of critical capacity, $\alpha_c$, on the parameters of the presented model: $N_i/N$, $p$, and $\kappa$.

5.4 Dependence of connection probability, $P_{con}$, on the parameters of the presented model: $N_i/N$, $p$, and $\kappa$.

5.5 Dependence of the average absolute connection strength, $<|J|>$, on the parameters of the presented model: $N_i/N$, $p$, and $\kappa$.

5.6 Distribution of connection strengths.
List of Tables

2.1 Anatomical parameters of synaptic circuits from mouse, rat, monkey, and human cortices ................................................. 22

2.2 Parameters of structural circuit remodeling in mouse, rat, monkey, and human cortices ..................................................... 25

3.1 Results of the cooperative synapse formation model ................................................. 47

5.1 Examples of variability in the distributions of connection strengths ($J$) and firing thresholds ($h$) ............................................. 60
Chapter 1

Introduction

Structural Plasticity of Cortical Circuits

Many important brain functions such as learning and memory depend on the plasticity of neural circuits. Neurogeometric analysis of cortical neuropil shows that there are typically many axons within spine reach of a given dendrite (1, 2). Hence, cortical neuropil holds a potential for structural circuit reorganization through the retraction of some of the preexisting spines and the formation of new spines and synapses. This spine remodeling plasticity has a large capacity for modifying neural circuits (1) and is thought to play an important role in learning and long-term memory formation (3). There have been numerous \textit{in vivo} experimental observations of spine remodeling plasticity in different cortical areas of developing and adult animals (4-7). Yet, it remains unclear how the new dendritic spines find their presynaptic targets and establish synaptic connections.
Cooperative Synapse Formation in the Neocortex

Our understanding of the rules governing synaptic connectivity in the brain is hindered by the complexity of neuron morphology. To circumvent this problem, it is often necessary to explicitly account for the shapes of axonal and dendritic arbors in the analysis of synaptic connectivity [see e.g. (8-11)]. In the cerebral cortex, the majority of synaptic connections between excitatory neurons are made on dendritic spines (12). Therefore, individual synapses between neurons occur in places where axonal branches of the pre-synaptic neuron are located within spine reach from the dendritic branches of the post-synaptic cell. Such axo-dendritic oppositions are termed potential synapses (1, 2). Because in the adult cerebral cortex, axonal and dendritic arbors of excitatory neurons form an extremely stable scaffold (5, 13-15) the resulting matrix of potential synapses is stable as well [see however (15-17), where small, layer specific, changes in terminal axonal and dendritic branches had been observed]. What is more, due to the stereotypic morphologies of dendritic and local axonal arbors of cortical neurons (same species, brain region, layer, etc.) (18), the matrix of potential synapses is expected to be similar among different brains (2, 19). This matrix constraints possible connectivity patterns in the adult brain and provides the main avenue for the formation of new synaptic connections. A potential synapse can be converted into an actual synapse if the gap between the pre- and the post-synaptic branches is bridged by a dendritic spine. In Chapter 3 we propose a cooperative scheme for potential to actual synapse formation.
Universality of the Distribution of Potential Synapse Numbers

Neuron morphology plays an important role in defining synaptic connectivity patterns in the brain. Traditionally, Sholl analysis (20) has been used as a quantitative method of studying the shapes of individual arbors. However, this method has no direct bearing on synaptic connectivity, which is dependent on the correlation between the shapes of the axonal and dendritic arbors. A better descriptor of synaptic connectivity between neurons is the number of potential synapses they form. In Chapter 3 we will see how different connectivity patterns between neurons can be characterized with the distribution of potential synapse numbers. The question that we ask next is what is the functional form of this distribution? In Chapter 4 we will utilize a large dataset of neurons reconstructed in 3D from different species and brain areas to show that the distribution of potential synapse numbers has a characteristic universal shape.

Effect of biological constraints on associative memory storage capacity of artificial neural networks

Over the years, there has been a great deal of interest in McCulloch and Pitts (21) model neural networks and their capacity for associative memory storage (see e.g. (22-31)). Most of the theoretical and computational models that have been used thus far ignore the above mentioned biological constraints on network architecture. In Chapter 5 of this thesis we therefore investigate a neural network model with more biologically plausible properties. We show that, unlike in traditional perceptron-like models, where network capacity is independent of the network architecture, the capacity in the considered model exhibits a more interesting behavior.
Chapter 2

Structural Plasticity of Cortical Circuits

2.1 Introduction

In this chapter, we describe a theoretical framework for the analysis of the neuropil’s potential to undergo structural synaptic remodeling. We considered two models of synapse formation by structural spine remodeling. In the first model, dendritic spines could establish new synapses in the neuropil wherever the pre-synaptic axonal branch is located sufficiently close to the post-synaptic dendritic segment. In the second model, the presence of a bouton on the pre-synaptic axonal branch is required for the formation of new synapses. We introduced two measures of the neuropil’s potential for structural synaptic reorganization: the connectivity fraction and the structural synaptic entropy. These structural parameters of neural circuit plasticity depend on the shape of dendritic spine length distribution function and other neuropil characteristics that are routinely measured with light or electron microscopy. The connectivity fraction and structural
synaptic entropy can be used to assess the learning and memory capacity of neural tissue in its normal or diseased state and make inter-areal or inter-species comparisons.

Based on the published anatomical data we evaluated the connectivity fraction in mouse occipital cortex, rat CA1, monkey V1, and human temporal cortex. The connectivity fraction is defined as a spine length dependent-ratio of the numbers of actual to potential synapses. As a result, the connectivity fraction is equivalent to the probability of finding an actual synaptic connection at a potential synaptic site. We found that the average connectivity fraction in rodent cortex, 0.14-0.24, is significantly higher than that in primate cortex, 0.051-0.10. Hence, an average dendritic spine can choose among 4-7 potential targets in rodents and 10-20 potential targets in primates. As our comparisons involved different species and cortical areas it is not clear whether these differences arise from differences between species, cortical areas, or are the result of the combination. New experiments are needed to resolve these questions.

2.2 Models of structural synaptic plasticity

A potential synapse is a location in the neuropil where the distance $s$ between an axon and a dendrite is small enough to be bridged by a dendritic spine (see Figure 2.1 A,B). We made this definition of a potential synapse more precise by considering the following two models. In model A (see Figure 2.1 C), we assumed that a spine can bridge the gap between the axon and the dendrite regardless of the presence of a bouton on the axon (1, 2). As a result, all dendrites that lay within the reach of a dendritic spine from the given axon are potential to that axon.
Figure 2.1. Potential synapse. A. 3D reconstructions of a layer 4 spiny stellate cell axon (blue) and a layer 3 pyramidal cell dendrite (red) from the cat primary visual cortex. Potential synapses between the arbors are shown with small black circles. Scale bar is 100μm. Modified from (19).

B. Schematic illustration of a 5μm x 5μm x 5μm volume of cortical neuropil (based on the data from mouse occipital cortex, Table 2.1). The axonal and dendritic segments shown are potentially connected if they can be bridged by dendritic spines. Densities of axons and dendrites were reduced 6-fold to avoid clutter. C and D illustrate two models of potential connectivity. In C a potential synapse is defined as a site in the neuropil where a dendritic branch is located a distance s away from an axon (model A). Presence of a synaptic bouton on the axon is not required. In D (model B) a potential synapse is defined as a location in the neuropil where a dendritic branch is present a distance s away from an existing synaptic bouton (blue sphere on the axon).

In this model it was assumed that spine outgrowth and spine–axon contact precede
bouton and synapse formation. Of course, some spines can by chance make contacts with pre-existing boutons, leading to the formation of multiple synapse boutons. There is however some direct and indirect experimental evidence that the presence of a bouton on an axon is required for establishing a new synaptic connection (9, 32). Hence, in model B (see Figure 2.2D), we assumed that the preexistence of a bouton on the axon is required for establishing a new synaptic connection (9, 32). In this model, we defined a potential synapse as proximity between a bouton and a dendritic branch. In other words, all of the dendrites that lay within the reach of a dendritic spine from the given bouton are potential to that bouton.

2.3 The density of potential synapses in the neuropil

In the following we calculate the density of potential synapses in the neuropil. We began by calculating the probability $P^4(d)$ that two long straight segments of given orientations (making angle $\theta$ with each other) and lengths $l_a$ and $l_d$ ($l_a, d \gg d$), randomly positioned inside a large volume $V$ will be within distance $d$ of each other (Figure 2.2A). Clearly, the two segments are located within distance $d$ of each other if one segment penetrates the imaginary cylinder of radius $d$ surrounding the other segment. The probability of this event is equal to the probability for the origin of the first segment to fall inside the right prism (2d by $l_a$ by $l_d$ and angle $\theta$ at the base) as shown in Figure 2.2A. Hence, $P^4(d)$ is equal to the ratio of the prism’s volume and volume $V$,

$$P^4(d) = \frac{2dl_a l_d \sin \theta}{V}.$$ \hspace{1cm} (2.1)
Figure 2.2. Probability of potential connection for line segments. A. The probability that a random line segment of a given orientation and length $l_a$ (red) is located within distance $d$ from another segment (blue) of length $l_a$ ($l_{a,d} >> d$) is equal to the probability that the origin of the first segment falls inside the prism shown in the figure. B. The probability that a random line segment (red) of length $l_a$ is located within distance $d$ from a given point (blue dot) is equal to the probability that the origin of this segment falls inside the cylinder of length $l_d$ and radius $d$ as shown.

Similarly, one can calculate the probability $P^\beta(d)$ that a long straight segment of a given orientation and length $l_d$, randomly positioned inside a large volume $V$ will be located within distance $d$ of a given point (Figure 2.2B). The segment is located within distance $d$ of the point if it intersects an imaginary sphere of radius $d$ centered at the point. The probability of this event is equal to the probability for the origin of the segment to be located inside the cylinder of radius $d$ and length $l_d$ as shown in Figure 2.2B. As a result, $P^\beta(d)$ is equal to the ratio of the cylinder’s volume and volume $V$,

$$P^\beta(d) = \frac{\pi d^2 l_d}{V}. \quad (2.2)$$

Now we considered a straight axonal segment of length $l^a$ and radius $r_a$ located inside a large neuropil volume $V$ (model A). A randomly chosen dendritic segment of length $l^d$, radius
and a given orientation in the neuropil (making an angle $\theta_j$ with the axon) makes a potential synapse that can be bridged by a spine in the $[0, s]$ length range, if the axis of this segment is located within distance $d$ from the axis of the axon, such that $r_a + r_d \leq d \leq s + r_a + r_d$. This means that the surfaces of the axonal and dendritic cylinders must be closer than a distance $s$ without touching each other. The probability of this happening is equal to the difference of probabilities $P_{ij}^{A}(s + r_a + r_d) - P_{ij}^{A}(r_a + r_d)$, where subscripts are added in reference to segments $i$ and $j$. In model B, a dendritic segment and a bouton of radius $r_b$ can be bridged by a spine in the $[0, s]$ length range if the axis of the dendritic segment is located within distance $d$ from the center of the bouton, such that $r_b + r_d \leq d \leq s + r_b + r_d$. As a result, the probability of potential connection is $P_{ij}^{B}(s + r_b + r_d) - P_{ij}^{B}(r_b + r_d)$.

Using Equations 2.1 and 2.2 for models A and B, the probabilities of potential connection corresponding to spine lengths in the $[0, s]$ range reduce to,

$P_{ij}^{A}(s + r_a + r_d) - P_{ij}^{A}(r_a + r_d) = \frac{2s l_d^i l_d^j \sin \theta_j}{V}$

$P_{ij}^{B}(s + r_b + r_d) - P_{ij}^{B}(r_b + r_d) = \frac{\pi \left[ s^2 + 2s \delta \right] l_d^j}{V}$

(2.3)

In the second equation $\delta = r_d + r_b$ is the sum of dendritic and bouton radii (see Table 2.1). The expected cumulative number of potential synapses in the neuropil volume $V$ that can be bridged by spines shorter than $s$ is the sum of the probabilities in Equations 2.3 over all the dendritic segments ($i$) and potentially post-synaptic elements ($j$, axonal segments in model A or boutons in model B):
Due to the assumption of no correlation in the layout of axonal and dendritic branches, the sum in the first equation breaks down into a product of axonal and dendritic components. As a result, the cumulative number of potential synapses reduces to:

\[
N^A_{pot}(s) = \frac{2s}{V} \sum_{ij} l^i_j l^j_i \sin \theta_{ij} \\
N^B_{pot}(s) = \frac{\pi \left(s^2 + 2s\delta\right)}{V} \sum_{ij} l^i_j .
\]  

(2.4)

In these expressions \( L_a, L_d \) are the combined axonal and dendritic lengths, \( N_b \) is the total number of boutons in the neuropil volume \( V \), and \( \overline{\sin \theta} \) is the average sine of angles between axonal and dendritic branches.

As \( L_a, L_d, N_b \), and the cumulative number of potential synapses, \( N^A_{pot}(s) \), are extensive quantities of volume, Equations 2.5 can be conveniently rewritten in terms of densities,

\[
n^A_{pot}(s) = 2\overline{\sin \theta} s L_a L_d \\
n^B_{pot}(s) = \pi \left(s^2 + 2s\delta\right) n_b L_d .
\]  

(2.5)

(2.6)

where \( \rho_{a,d} \) are the axonal and dendritic length densities (combined length of axons or dendrites in a unit volume of neuropil) and \( n_b \) denotes the volume density of boutons. If relative orientations of axonal and dendritic branches are isotropically distributed (which is already true if either axonal or dendritic branches are isotropic), then \( \overline{\sin \theta} = \pi / 4 \) (1). Note that models A and B result in different dependence of the cumulative potential synapse density on spine length \( s \).

In model A, a potential synapse was defined symmetrically with respect to axonal and dendritic
branches (Figure 2.1C). As a result, the cumulative density of potential synapses in the neuropil, $n_{pot}^A(s)$, depends on the product of axonal and dendritic length densities, $\rho_a$ and $\rho_d$, of excitatory neurons. Length density is the combined length of neurites (axons or dendrites) in a unit volume of neuropil. In model B (Figure 2.1D), a potential synapse was defined as an opposition between a bouton and a dendrite, and thus $n_{pot}^B(s)$ depends on the product of the volume density of boutons, $n_b$, and the dendritic length density.

<table>
<thead>
<tr>
<th>Species, brain area, age</th>
<th>Distribution of spine lengths, $p(s)$</th>
<th>Average spine length, $\bar{s}$ [µm]</th>
<th>Sum of the average dendritic and bouton radii, $\bar{\delta}$ [µm]</th>
<th>Inter-bouton interval along an axon, $1/b_a$ [µm]</th>
<th>Spine density on a dendrite, $1/b_d$ [µm$^{-1}$]</th>
<th>Density of asymmetric synapses, $n_s$ [µm$^{-3}$]</th>
<th>Dendritic length density, $\rho_d \approx n_b b_d$ [µm$^{-2}$]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mouse occipital, layer 3, adult</td>
<td>(33)</td>
<td>0.99±0.01 (0.01)</td>
<td>0.70 (34)</td>
<td>4.5±1.4 (34-36)</td>
<td>1.94±0.42 (34, 37)</td>
<td>0.91±0.25 (34, 36)</td>
<td>0.48±0.10 (0.10)</td>
</tr>
<tr>
<td>Rat CA1, stratum radiatum, adult</td>
<td>(38)</td>
<td>1.08±0.03 (0.03)</td>
<td>0.50 (39-42)</td>
<td>3.7±0.6 (43)</td>
<td>3.41±1.05 (38, 41, 44)</td>
<td>2.0±0.30 (45)</td>
<td>0.59±0.08 (0.08)</td>
</tr>
<tr>
<td>Monkey V1, layer 3, adult</td>
<td>(1)</td>
<td>1.86±0.05 (0.05)</td>
<td>0.95</td>
<td>5.6±2.4 (46)</td>
<td>0.55±0.07 (47-50)</td>
<td>0.26±0.04 (51, 52)</td>
<td>0.47±0.05 (0.05)</td>
</tr>
<tr>
<td>Human temporal, layer 3, Adult</td>
<td>(33)</td>
<td>1.42±0.01 (0.01)</td>
<td>1.35</td>
<td>-</td>
<td>2.62±0.34 (53)</td>
<td>1.07±0.31 (54)</td>
<td>0.42±0.09 (0.09)</td>
</tr>
</tbody>
</table>

Table 2.1: Anatomical parameters of synaptic circuits from mouse, rat, monkey, and human cortices. The data is shown in mean±std(sem) format. Details are provided in Appendix A.

We note that in model A, the cumulative density of potential synapses depends linearly on the spine length, $s$, while in model B the dependence is quadratic. As the majority of excitatory synapses are made on dendritic spines (12) and the majority of spines bear a single excitatory synapse, the dendritic length density can be estimated as the product of the asymmetric synapse density and the average inter-spine interval along excitatory dendrites, $\rho_d \approx n_s b_d$. The estimated
values of $\rho_a$ for the four considered systems are shown in Table 2.1. Similarly, the axonal length density can be estimated as the product of the density of asymmetric synapses and the average inter-bouton interval on excitatory neuron axons, $\rho_a \approx n_h a$. Because of the presence of multiple synapse boutons this expression may slightly overestimate the axonal length density. (Chapter 2.4)

2.4 The connectivity fraction

In order to evaluate the potential of cortical neuropil for structural circuit reorganization by spine remodeling, we introduced the connectivity fraction, $f_{A,B}(s)$. The connectivity fraction is a measure of plasticity potential associated with spine remodeling. It is defined as the fraction of potential synapses (for a given spine length, $s$) which had been converted into actual ones. As the cumulative number of potential synapses depends on distance $s$, the connectivity fraction, in general, will be dependent on $s$ as well, $f(s)$.

To calculate $f_{A,B}(s)$ we note that in a unit volume of neuropil, there are $\Delta n_{act}(s) = n_{spine} p(s) \Delta s$ actual synapses on spines in the $[s - \Delta s / 2, s + \Delta s / 2]$ length range. Here, $p(s)$ is the spine length distribution function and $n_{spine}$ is the volume density of spines. These $\Delta n_{act}(s)$ spines are distributed among $\Delta n_{pot}^{A,B}(s)$ potential synaptic sites (derived from Equations 2.6). The connectivity fraction $f_{A,B}(s)$ can be obtained as the ratio of the numbers of spines and potential synapses in the $[s - \Delta s / 2, s + \Delta s / 2]$ interval:
The above expressions are not convenient for making quantitative estimates of the connectivity fraction. This is because axonal and dendritic length densities are not typically measured in experiments. To get around this problem we note that the ratio \( n_{\text{spine}} / \rho_d \) is equal to the average linear density of spines on the excitatory neuron dendrites, or the inverse of the inter-spine interval on a dendrite, \( 1/b_d \). Similarly, \( \rho_a \) can be expressed as the product of the average inter-bouton interval on the excitatory neuron axons, \( b_a \), and volume density of these boutons, \( n_b \) (\( \rho_a = b_a n_b \)). Finally, by denoting the ratio between the number of asymmetric synapses, \( n_s \), and the number of boutons, \( n_b \), on excitatory neuron axons as \( m (n_s = n_b m) \), we arrive at the final expressions for the connectivity fractions in models A and B,

\[
\begin{align*}
 f_A(s) &= \frac{n_{\text{spine}} p(s)}{2 \sin \theta \rho_a \rho_d} = f_A^* \overline{s} p(s); \\
 f_B(s) &= \frac{n_{\text{spine}} p(s)}{2 \pi (s + \delta) n_b \rho_d} = f_B^* \overline{s}^2 p(s).
\end{align*}
\]

(2.9)

All the components in these expressions are routinely measured with electron or light microscopy. For convenience we broke down the connectivity fractions, in Equations 2.8, into a product of two parts: dimensionless parameters \( f_{A,B}^* \), which depend on the anatomical details of neuropil organization, and a dimensionless part which is primarily dependent on the shape of the spine length distribution function. In our calculations we use Equations 2.8 with the values of \( m = 1 \) and \( \sin \theta = \pi / 4 \), following the assumptions and approximations stated in Appendix A.

A parameter similar to \( f_A^* \) was initially introduced in (1) under the name of filling parameters.
fraction. One difference being that in the earlier work the spine length was measured between the
tip of the spine head and the dendritic axis as opposed to the base of the spine. The value of this
dimensionless parameter was estimated to be in the 0.1-0.3 range for many species and cortical
areas consistent with the results of this study.

The connectivity fraction depends on the shape of the spine length distribution function,
\( p(s) \), as well as on average anatomical parameters of cortical micro-architecture (Equations
2.9). The latter dependence is mainly captured by a single parameter \( f_{AB}^{\ast} \) which is referred to as
the connectivity parameter (Table 2.1).

We evaluated the expressions for the connectivity fraction based on the experimentally
measured spine length distribution functions and the values of anatomical parameters (Table 2.1)
for the four considered cortical areas. The results for models A and B are shown in Figures 2.3B
and Table 2.2.

<table>
<thead>
<tr>
<th>Species, brain area, age</th>
<th>Connectivity parameter, ( f^* )</th>
<th>Average connectivity fraction, ( \bar{f} )</th>
<th>Maximum connectivity fraction, ( f_{\text{max}} )</th>
<th>Average entropy per spine, ( i/n_i ) [bits]</th>
<th>Average entropy per volume, ( i ) [bits/( \mu m^3 )]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mouse occipital, layer 3, adult</td>
<td>A 0.32 ± 0.08</td>
<td>B 0.36 ± 0.08</td>
<td>A 0.19 ± 0.05</td>
<td>B 0.14 ± 0.03</td>
<td>A 0.25 ± 0.06</td>
</tr>
<tr>
<td>Rat CA1, stratum radiatum, adult</td>
<td>A 0.27 ± 0.04</td>
<td>B 0.23 ± 0.03</td>
<td>A 0.24 ± 0.04</td>
<td>B 0.16 ± 0.02</td>
<td>A 0.30 ± 0.05</td>
</tr>
<tr>
<td>Monkey V1, layer 3, adult</td>
<td>A 0.14 ± 0.01</td>
<td>B 0.10 ± 0.03</td>
<td>A 0.10 ± 0.01</td>
<td>B 0.051 ± 0.006</td>
<td>A 0.15 ± 0.05</td>
</tr>
<tr>
<td>Human temporal, layer 3, adult</td>
<td>A -</td>
<td>B 0.20 ± 0.05</td>
<td>A -</td>
<td>B 0.072 ± 0.016</td>
<td>A -</td>
</tr>
</tbody>
</table>

Table 2.2: Parameters of structural circuit remodeling in mouse, rat, monkey, and human
cortices. The data is shown in mean±sem format. Calculation details are provided in Appendix A.
Figure 2.3. Characterization of structural spine remodeling. A. Spine length distributions: (A1) mouse occipital cortex, (A2) area CA1 of rat hippocampus, (A3) monkey visual area V1, (A4) human temporal cortex. B1-B4. Connectivity fraction as a function of spine length. Connectivity fraction is the ratio of the numbers of actual to potential synapses. Results based on model A are shown in red and those based on model B are shown in green. C1-C4. Structural entropy density per spine as a function of spine length. The overall structural entropy per spine can be calculated as the area under the curves.

The most salient feature of the results in Figures 2.3B is that the average (or peak, $f_{\text{max}}$) connectivity fractions, $\bar{f}$, in rodents are significantly higher than those in the primates ($p<0.05$ for all pairwise comparisons, Student’s t-test for samples with unequal variance). There are no significant differences in $\bar{f}$ within rodent and primate groups. For example, in mouse occipital cortex the average connectivity fraction is 0.19 for model A (see Table 2.2) indicating that on average a spine can choose among about 5 potentially pre-synaptic axons. For model B $\bar{f} = 0.14$
which means that a single bouton can be contacted by about 7 potentially post-synaptic dendrites. In contrast, the average connectivity fraction in human temporal cortex is 0.072 in model B, which is significantly smaller (p<0.03). Here a single bouton can be contacted by an astounding 14 different post-synaptic dendritic spines.

2.5 The number of cortical circuits that can be built with spine remodeling

Next, we calculated the entropy associated with building different synaptic circuits out of the scaffold of overlapping axonal and dendritic branches in the neuropil. This entropy reflects the number of structurally different circuits that can be achieved by spine remodeling. It is expected that this number of structurally different circuits exceeds the number of circuits that have different functional properties. This is because some structural connectivity patterns that differ in the placement of individual synapses will result in the same functional circuit on the level of individual neurons or neural networks. As a result, the structural synaptic entropy provides only the upper bound to the amount of useful (functional) entropy in the neural network that can be accessed with spine remodeling. Yet, structural synaptic entropy is expected to be correlated with the amount of useful entropy and, as such, could be viewed as a measure of circuit plasticity and memory storage capacity. We found that the neuropil’s potential for structural circuit remodeling is highest in rat CA1 (7.1-8.6 bits/µm³) and lowest in monkey V1 (1.3-1.5 bits/µm³). We do not know which of the considered models of synaptogenesis (A or B) is more appropriate. It is likely that both mechanisms of new synapse formation are utilized in the brain. In this case, our results only describe the limiting scenarios. Yet, the average difference in structural synaptic entropy for the two models is only about 15%, suggesting that our estimates are informative irrespective of
The neuropil volume $V$ contains $\Delta n_{\text{pot}}^{A,B}V$ potential and $\Delta n_{\text{act}}V$ actual synapses in the $[s - \Delta s / 2, s + \Delta s / 2]$ range of spine lengths. In this range, the number of ways to choose actual synapses out of the pool of potential ones is given by the binomial coefficient,

$$
\Delta \Omega_{A,B}(s) = \binom{\Delta n_{\text{pot}}^{A,B}V}{\Delta n_{\text{act}}V}.
$$

(2.11)

The structural synaptic entropy, $\Delta I_{A,B}(s)$, is defined as the log₂ of this number. In the limit of large numbers of actual and potential synapses (large $V$) the binomial coefficient can be approximated using Stirling’s formula (55), and

$$
\Delta I_{A,B}(s) = \log_2 \left( \Delta \Omega_{A,B}(s) \right) = V i_{A,B}(s) \Delta s
$$

$$
i_{A,B}(s) = -n_{\text{spine}} p(s) \left[ \log_2 \left( f_{A,B}(s) \right) + \frac{1 - f_{A,B}(s)}{f_{A,B}(s)} \log_2 \left( 1 - f_{A,B}(s) \right) \right].
$$

(2.12)

In this expression $i_{A,B}(s)$ is referred to as entropy density distribution.

Because the structural synaptic entropy, $\Delta I_{A,B}(s)$, is proportional to the volume of neuropil, it is natural to introduce the volume density of this quantity. Adding the entropic contributions arising from different spine length intervals, $\Delta s$, we obtain the overall structural synaptic entropy per volume of neuropil,

$$
i_{A,B} = \int_0^\infty i_{A,B}(s) ds = -n_{\text{spine}} \int_0^\infty \left[ \log_2 \left( f_{A,B}(s) \right) + \frac{1 - f_{A,B}(s)}{f_{A,B}(s)} \log_2 \left( 1 - f_{A,B}(s) \right) \right] p(s) ds.
$$

(2.13)

Model dependence in this expression is contained in the connectivity fraction $f(s)$, Equations 2.8. The structural synaptic entropy per volume scales with the volume density of spines, $n_{\text{spine}}$. It is a functional of the spine length distribution function $p(s)$ and depends on
anatomical characteristics of neuropil micro-architecture through \( n_{\text{spine}}, n_s, b_a, b_a, \) and \( m \). In our calculations we use a simplified version of Equation 2.11, where volume density of spines, \( n_{\text{spine}} \), is replaced with the volume density of asymmetric synapses, \( n_s \) (see Appendix A).

The number of structurally different circuits attainable by remodeling of dendritic spines of length \( s \) is described by the structural entropy density per spine, \( i_{A,B}(s)/n_s \),

\[
\frac{i_{A,B}(s)}{n_s} = -p(s) \left[ \log_2 \left( \frac{f_{A,B}(s)}{f_{A,B}(s)} \right) + \frac{1-f_{A,B}(s)}{f_{A,B}(s)} \log_2 \left( 1-f_{A,B}(s) \right) \right].
\]  (2.14)

The results of calculations based on this expression are shown in Figures 2.3C. The shapes of the entropy density curves roughly follow the shapes of the spine length distribution functions. For large values of spine length, \( s \), the entropy density per spine in model B is larger than that in model A. This trend, which is exactly opposite to the trend in the connectivity fractions shown in Figures 2.3B, is due to the fact that large connectivity fractions correspond to low entropy densities per spine and vice versa (see Equation 2.14 and Figure 2.4). In mouse occipital cortex, the structural entropy density per spine peaks at about 3bits/\( \mu \m \) for 1\( \mu \m \) spines. This means that, for example, spines in the range of lengths from 0.95\( \mu \m \) to 1.05\( \mu \m \) contribute 3bits/\( \mu \m \times 0.1\mu \m = 0.3 \)bits of entropy per spine to structural connectivity. The peak entropy density per spine is highest in rat CA1; 3.0bits/\( \mu \m \) in model A and 3.7bits/\( \mu \m \) in model B.
Figure 2.4. Structural entropy per spine as a function of connectivity parameter, $f^*$. Panels A and B show the results based on models A and B for mouse occipital cortex (red), rat CA1 (blue), monkey V1 (green), and human temporal cortex (black). The dots in these figures indicate the connectivity parameters, $f^*$, and the corresponding values of structural entropy per spine, $i/n_s$, calculated based on data from Tables 2.1 and 2.2.

Integrating Equation 2.14 over spine lengths (areas under the curves in Figures 2.3C) we arrived at the expression for the overall structural entropy per dendritic spine:

$$\frac{i_{A,B}}{n_s} = \sum_0^\infty \log_2 \left( \frac{f_{A,B}(s)}{1 - f_{A,B}(s)} \right) p(s)ds \ . \ (2.15)$$

The results of this calculation are shown in Figure 2.5A and Table 2.2. In rodent cortex a single spine can contribute 3.6-4.5 bits of entropy to the patterns of synaptic connectivity. In monkey V1 and human temporal cortex the amount of entropy available to a single spine is significantly higher, 4.9-5.9 bits (p<0.04 for all pairwise comparisons). In addition, we did not detect any significant differences in structural entropy per spine within rodent and primate groups. The fact that a single spine in primate cortex has high entropy for structural plasticity
seems satisfactory. Yet, in comparison, low 3.6-4.3 bits of entropy per spine in the rat CA1, an essential learning and memory area, could seem surprising. However, this lack of structural entropy per spine in rat CA1 is compensated by a large density of synapses.

**Figure 2.5.** Potential for structural plasticity. **A.** Structural entropy per spine in mouse occipital cortex, rat CA1, monkey V1, and human temporal cortex. White bars show the results based on model A and gray bars correspond to model B. **B.** Structural entropy per volume of cortical neuropil.

Multiplying the entropy per spine, $i_{A,B}/n_s$, with the density of asymmetric synapses, $n_s$, we obtain the structural synaptic entropy per volume of neuropil, $i_{A,B}$ (see Figure 2.5B and Table 2.2). This quantity reflects the number of structurally different circuits that can be achieved with spine remodeling in the unit volume of cortical neuropil. With the exception of comparison between mouse and human in model B, all within-model differences in Figure 2.5B are significant ($p<0.05$). Structural entropy per volume is highest in the rat CA1, 7.1-8.6 bits/$\mu$m$^3$ and lowest in the monkey V1, 1.3-1.5 bits/$\mu$m$^3$, primarily due to the high and low densities of asymmetric synapses in these cortical areas.
The average structural entropy per spine, $i_{A,B}/n_s$, according to Equations 2.13 and 2.14, depends only on the shape of the spine length distribution function $p(s)$ and the connectivity parameter $f_{A,B}^*$. Figure 2.5 shows the dependence of $i_{A,B}/n_s$ on $f_{A,B}^*$ in models A and B. The dots indicate the results of the calculations based on the data from Tables 2.1 and 2.2. For a given spine length distribution function, the structural entropy per spine is a decreasing function of the connectivity parameter.

2.6 Selectivity of post-synaptic neurons in their choice of pre-synaptic partners

We considered a model of synaptogenesis by structural spine remodeling where creation of new synapses between axonal and dendritic branches requires the following three steps. First, the pre-synaptic axon has to be within the spine reach of the post-synaptic dendrite, i.e. the two branches must be in potential contact with each other. Second, a dendritic spine or filopodium has to find and establish an initial contact with the axon. Finally, based on the functional properties of the neurons this connection could be stabilized and transformed into a synapse or be eliminated.

Consider a potentially connected pair of axonal and dendritic branches. This potential connection may or may not contain an actual synaptic contact. The probability that potentially connected branches are synaptically coupled is given by the connectivity fraction, $f(s)$. According to the above model of synaptogenesis, this probability is equal to the product of the probability for a dendritic spine of length $s$ to find the axon, $p(find|s)$, and the probability for this initial contact to stabilize and transform into an actual synaptic connection,
\[ p(\text{stabilize} \mid \text{found}) \]. Hence,

\[ f(s) = p(\text{find} \mid s) p(\text{stabilize} \mid \text{found}). \quad (2.16) \]

Probabilities \( f(s) \) and \( p(\text{find} \mid s) \) depend on the distance \( s \) between the branches. The conditional probability \( p(\text{stabilize} \mid \text{found}) \), on the other hand, is independent of the geometrical details of circuit organization. This probability is only dependent on functional properties of the neurons. It reflects the selectivity of cortical neurons in their choice of synaptic partners.

The fact that the probability \( p(\text{find} \mid s) \) has to be less or equal to one for all values of the parameter \( s \) allows us to obtain the lower bound for the probability \( p(\text{stabilize} \mid \text{found}) \) in the following way,

\[
\begin{align*}
  p(\text{stabilize} \mid \text{found}) & \geq f(s), \quad \forall s \\
  p(\text{stabilize} \mid \text{found}) & \geq f_{\text{max}}
\end{align*}
\quad (2.17)
\]

In the last expression \( f_{\text{max}} \) is the maximum value of the connectivity fraction, \( f(s) \), Figure 2.3B.

Excitatory neurons in regions of the cerebral cortex differ in their functional properties. In the primary visual cortex, for example, neurons differ in their ocular dominance and orientation preference (56). These neurons connect to each other based on the similarity of their functional properties (57-59). It is not entirely clear how many different neuron types are there in the visual cortex. In other words, what is the probability that two nearby randomly chosen excitatory neurons would establish a synaptic contact given the opportunity? One may think that this probability is quite small, less than 10% based on the dual intracellular recordings from nearby neuron pairs (60). This view is also supported by the traditional Hubel and Wiesel model of the
hypercolumn (61), which contains neurons with different ocular dominance and orientation properties. The question of selectivity of excitatory neurons becomes even more obscure in the non-primary cortical areas where functional classification of neurons is often unknown. We estimated the lower bound of selectivity of neurons in their choice of synaptic partners based on the shape of the spine length distribution function and average anatomical parameters of neuropil micro-architecture. We estimate the lower bound of $p(\text{stabilize} \mid \text{found})$ based on the values of $f_{\text{max}}$ from Table 2.2.

Our results show that rodent post-synaptic excitatory neurons are not very selective in their choice of pre-synaptic targets. These neurons make synaptic contacts with more than 21-30% of pre-synaptic axons encountered with new spine growth. Primate neurons appear to be more selective making synaptic connections with more than 7-15% of encountered axons. This estimate allows us to define the number of functionally different neuron classes in a given region of the cerebral cortex as the inverse of neuron selectivity. According to this measure, in the considered areas of rodent cortex, there are at most 3-5 functionally different classes of neurons. In the primate cortical areas, the number of functionally different classes of neurons could be much higher. It is only limited from above by our estimate of 7-14 neuron classes.
Chapter 3

Cooperative Synapse Formation in the Neocortex

3.1 Introduction

In this chapter, we examine the rules of excitatory connectivity within the constraints imposed by the morphologies of neurons, i.e. within the matrix of potential synapses. Different connectivity patterns can be built within this matrix by converting different sets of potential synapses into actual. However, not all such connectivity patterns are biologically plausible. This is because, even within the matrix of potential synapses, synaptic connectivity is built according to specific rules (9). We ask and attempt to answer two questions. Are the individual synaptic connections between potentially connected pre- and post-synaptic neurons established independently of each other? If not, what are the possible rules governing synaptic connectivity within the potential connectivity matrix?
Figure 3.1. Potential and actual synapses. A. 3D reconstructions of a layer 4 spiny stellate cell axon (red) and a layer 2/3 pyramidal cell dendrite (blue) from rat barrel cortex. Potential synapses between the arbors are shown with small black circles. Adopted from (8) (their Figure 3a). B. Distribution of the numbers of potential synapses (blue) and distribution of the numbers of actual synapses for synaptically coupled neurons (red) forming local L4→L2/3 projection. C. Same for pairs of nearby neurons in L5. D. Same for pairs of nearby neurons in L4.

Pairs of excitatory neurons in the cerebral cortex typically share several potential synapses (see Appendix B for more on definition of potential synapse) whenever there is a significant overlap of their axonal and dendritic arbors (Figure 3.1A). This is true for most
nearby excitatory neuron pairs and for some local inter-laminar projections in the neocortex [see e.g. (2, 9, 19, 62-66)]. In particular, in rat barrel cortex, neuron pairs forming local (separated by less than 50 µm laterally) layer 4 to layer 2/3 (L4→L2/3), L5→L5, and L4→L4 projections are known to have numerous potential synapses (8, 10). Hence, it is not surprising that the numbers of actual synapses between such synaptically coupled neurons are numerous as well (67-69) (Figure 3.1B-D).

What comes to us as a surprise, is that the distributions of actual synapse numbers, between synaptically coupled neurons forming these projections, are highly tuned, i.e. the ratios between variance and mean (the Fano factor) of these distributions are significantly less than one: 0.060 for L4→L2/3, 0.21 for L5→L5, and 0.31 for L4→L4 projections. In particular, these distributions are manyfold less variable than the corresponding distributions of potential synapse numbers (Figure 3.1B-D), for which the Fano factors are greater than one: 2.0, 3.2, and 2.2 respectively. This suggests that individual actual synapses between pre- and post-synaptic neurons may have been chosen from the potential synapses not randomly, e.g. not independently of each other. Is it possible that individual cells somehow regulate the numbers of synapses made with every synaptic partner? We examined this idea quantitatively by analyzing experimentally observed numbers of actual synapses between neurons (67-69) together with the dataset of neuron morphologies reconstructed in 3D (8) [obtained from http://NeuroMorpho.Org (70)].

3.2 Potential synapses are converted into actual with no specificity to synapse location

We assumed that actual synapses, for a given projection, are created at individual potential synaptic locations with a constant probability, \( p \), independent of synapse location on axonal and
dendritic arbors, as well as its cortical position. Though, there are numerous examples of synaptic specificity within the potential connectivity matrix [see e.g. (8, 71-73)], we think that the above assumption holds for the considered projections. As a possible test, we compared the distribution of dendritic path lengths from actual synapses to the soma to the similar distribution for potential synapses. We identified all the potential synapses for every pair of reconstructed neurons, and calculated the dendritic path length from each potential synapse to the soma. Next, we compared these path lengths to similar experimental measurements performed for actual synapses, Figure 3.2.

![Figure 3.2](image)

**Figure 3.2.** Distributions of dendritic path lengths from actual and potential synapses to the soma are not statistically different. **A.** L4→L2/3 projection. The path length distribution for potential synapses (blue line) was rescaled to match the total number of events in the path length distribution for actual synapses (red bars). **B.** Same for pairs of nearby neurons in L5. **C.** Same for pairs of nearby neurons in L4.

The distribution of actual synapse path lengths for L4→L2/3 projection was obtained from (68) (their Figure 14A). In that study path length measurements were performed on camera lucida drawings of 2D projected neuron images. Likewise, when generating the distribution of path lengths for potential synapses, the dendritic arbors were projected on the 2D plane of the slice. The comparison did not show any significantly differences between the distributions (p =
0.16, Kolmogorov-Smirnov test). In addition, the mean path lengths to the soma were not significantly different (p = 0.38, two-tailed Student’s t-test for samples with unequal variances) for actual, 67 ± 34 µm (mean ± SD), and potential, 73 ± 55 µm, synapses.

For L5 neurons measurements from (74) (their Figure 4B) were used for the comparison of actual and potential synaptic path lengths. In both cases the path lengths were measured in 3D. The results are shown in Figure 3.2B. The distributions of path lengths are not significantly different from each other (p = 0.13, Kolmogorov-Smirnov test). Also, the comparison of the mean path lengths for actual, 107 ± 63 µm (mean ± SD), and potential, 100 ± 97 µm, synapses yielded no significant difference (p = 0.65, two-tailed Student’s t-test for samples with unequal variances). Because the Frick et al. study was performed in L5A of rat barrel cortex, matching the dataset of reconstructed neurons, it was used for the comparison of path lengths. Unfortunately, the small number (n = 6) of connections analyzed in that study precluded its use in our calculations. Instead, data from Markram et al. (69), which contains a higher number (n = 19) of connections but comes from L5B of rat somatosensory cortex, was used. Because L5B neurons, in comparison to L5A, have longer apical dendrites, with more extensive apical tufts, and larger numbers of oblique branches, the path lengths for actual synapses (based on L5B neurons) turned out to be significantly longer than those for potential synapses (based on L5A neurons). For the same reason, it is expected that if L5B neurons were available for the analysis of potential connectivity, the numbers of potential synapses would be up to 30% higher (63) than what is shown in Figure 3.1C (based on L5A neurons). However, the distributions of potential synapse numbers for different projections and different values of the parameter s appear to have a universal shape (see Appendix B). As a result, this distribution for L5B neurons is expected to be similar to that for L5A neurons but calculated for up to 30% larger value of s. Since the
parameters of the cooperative synapse formation model are not sensitive to particular values of \( s \) (see Appendix B), the morphological differences between L5A and L5B neurons are not expected to affect the main conclusions of our work.

Measurements from (67) (their Figure 15A) were used for L4→L4 comparison. Since these measurements were done on 2D projections, the path lengths for potential synapses were 2D projected as well. Figure 3.2C shows that the distributions of path lengths for actual and potential synapses are not significantly different (\( p = 0.27 \), Kolmogorov-Smirnov test). Similarly, the mean path lengths are not significantly different (\( p = 0.57 \), two-tailed Student’s t-test for samples with unequal variances) for actual, 69 ± 37 µm (mean ± SD), and potential, 66 ± 52 µm, synapses.

3.3 Synapses between neurons are not formed independently of each other

Synaptic connectivity between nearby neurons or neurons forming local inter-laminar projections is very sparse, i.e. the probabilities of finding such connected neurons are low [0.03-0.3 (75)]. In spite of this sparseness, synaptically coupled neurons are typically interconnected with several synapses (67-69). Clearly, if actual synapses were formed completely randomly at the potential synaptic sites, low probabilities of connection would entail small (one or sometimes two) numbers of actual synapses between synaptically coupled neurons (9), which is inconsistent with the experimental evidence. To reconcile the numbers one could consider the fact that neurons, even in the same small cortical region, may differ in their functional properties. Hence, a given neuron may only connect to some of its (synaptically compatible) potential partners, establishing individual synaptic connections with them probabilistically and independently of one another. As
a result, it may be possible to have a low probability of connection on the one hand and a high connectivity between synaptically coupled neurons on the other.

In order to examine this idea quantitatively, we simulated the model of independent synapse formation in which individual synapses between synaptically compatible neurons are formed at the potential synaptic locations probabilistically and independently of each other. We used a Monte Carlo procedure and generated the distributions of actual synapse numbers for pairs of synaptically compatible neurons. To this end, we first randomly selected reconstructed neuron pairs, and for every pair identified the potential synaptic sites with a computer search algorithm. Next, we randomly converted individual potential synapses into actual with a fixed probability $p$, which was assumed to be independent of synaptic locations on axonal and dendritic arbors, as well as their cortical positions (see Chapter 3.2).

In result, for different values of the parameter $p$, we generated the distributions of actual synapse numbers, which take into account the morphological characteristics of neurons belonging to different cortical layers, Figure 3.3. For every value of the parameter $p$ and for all the projections, the generated and the experimental distributions of synapse numbers are statistically different ($p$-value $\leq 10^{-5}$ for $L4 \rightarrow L2/3$ and $L5 \rightarrow L5$, and $\leq 0.002$ for $L4 \rightarrow L4$, see Appendix B). We conclude that the idea of independent synapse formation is irreconcilable with the experimental data.

The above Monte Carlo simulation can be formulated theoretically. In this model, cortical cells forming a particular projection may differ in their functional properties which determine their preferences in establishing synaptic connections with each other. Some pairs of cells would couple synaptically when given the opportunity. Such pairs are referred to as synaptically compatible. The probability that a randomly chosen pair of neurons is synaptically compatible is
denoted with $\kappa$.

**Figure 3.3.** Individual synapses between pre- and post-synaptic neurons are not formed independently of each other. **A.** Experimental distribution of the numbers of actual synapses (red bars) between synaptically coupled neurons forming local L4→L2/3 projection. Individual lines show the distributions of actual synapse numbers obtained from the model of independent synapse formation for different values of the parameter $p$. **B.** Same for pairs of nearby neurons in L5. **C.** Same for pairs of nearby neurons in L4.

A synaptically compatible pair of neurons has a chance of establishing a synaptic connection only if it is potentially connected. These neuron pairs convert their individual potential synapses into actual synapses according to a fixed probability, $p$ (see SI text for a test of this assumption). Hence, the probability that a pair of these cells will establish precisely $N_s$ actual synapses, given $N_p$ potential ones, is equal to the binomial probability of choosing $N_s$ out of $N_p$,

$$ B(N_s \mid N_p) = \frac{N_p!}{(N_p-N_s)!N_s!} p^{N_s} (1-p)^{N_p-N_s}. \quad (3.1) $$

Synaptically compatible neurons forming a given projection may share different numbers of potential synapses. This variability is captured by the distribution of potential synapse numbers, $P(N_p)$ (Figure 3.1). The probability that a randomly chosen pair of synaptically compatible cells shares exactly $N_s$ actual synapses can be obtained by calculating the average of
$B(N_s \mid N_p)$ weighted with $P(N_p)$.

The above steps can be summarized into the following concise expression for the probability that a randomly selected pair of cells shares exactly $N_s$ actual synapses:

$$A^{IM}(N_s) = (1 - \kappa)\delta_{0,N_s} + \kappa \sum_{N_p=N_s}^{\infty} P(N_p)B(N_s \mid N_p).$$  \hspace{1cm} (3.2)

In this expression, $\delta_{0,N_s}$ is the Kronecker symbol, which equals 0 for all values of $N_s \geq 1$ and is 1 for $N_s = 0$. One may show that, because the distribution of potential synapse numbers, $P(N_p)$, is normalized to one, the distribution of actual synapse numbers, $A^{IM}(N_s)$, is normalized to one as well.

Two experimental measures are typically reported regarding the connectivity between neurons. The first measure is the probability of connection, $P_{con}$. This is the probability that a randomly chosen pair of neurons is synaptically coupled. Within the model of independent synapse formation, $P_{con}$ can be obtained by adding the probabilities for a randomly selected neuron pair to have one or more actual synapses,

$$P_{con}^{IM} = \sum_{N_s=1}^{\infty} A^{IM}(N_s) = \kappa \sum_{N_s=1}^{\infty} \sum_{N_p=N_s}^{\infty} P(N_p)B(N_s \mid N_p).$$  \hspace{1cm} (3.3)

The second experimental measurement is the distribution of actual synapse numbers between randomly selected, but synaptically connected, neurons, $A^{IM}(N_s \mid con)$. This conditional probability can be deduced from Bayes’ rule,

$$A^{IM}(N_s \mid con) = \frac{A^{IM}(N_s,con)}{P_{con}^{IM}}. \hspace{1cm} (3.4)$$

Since we are only interested in the synaptically coupled neurons ($N_s \geq 1$), the joint probability $A^{IM}(N_s, con)$ can be replaced with $A^{IM}(N_s)$, and the distribution of synapse numbers for synaptically connected neurons reduces to:

43
The shape of $A^{IM}(N_s \mid \text{con})$ depends only on a single parameter, $p$. Equation 3.5 was used to generate $A^{IM}(N_s \mid \text{con})$ for different values of $p$, Figure 3.3.

### 3.4 Model of cooperative synapse formation

To reconcile the experimental observations of highly tuned connectivity between synaptically coupled neurons with broadly distributed potential synapse numbers, we next explored the possibility of cooperative (not independent) synapse formation. We hypothesized that, in order to be synaptically coupled, pre- and post-synaptic cells must be connected with greater than some critical number of synapses. To model such cooperative synapse formation we introduced a monotonically increasing function, $f(N_s)$, which is the probability that a cell pair, initially connected with $N_s$ synapses, will remain synaptically coupled. Specifically, we used a two-parameter sigmoidal function,

$$f(N_s) = \frac{1}{1 + \exp \left[ -\frac{4}{\Delta} (N_s - N_s^c) \right]}$$

where $N_s^c > 0$ and $\Delta > 0$ specify the position of the inflection point and the inverse slope (cotangent) at that position. Because $N_s^c$ marks the position of the cooperative transition, we refer to this parameter as the critical number of synapses. The parameter $\Delta$ controls the sharpness of the cooperative transition. For example, in the transition window of width $\Delta$ centered at $N_s^c$, the value of $f(N_s)$ change from 0.12 to 0.88. For this reason $\Delta$ is referred to as the width of the
cooperative transition. Similar to the model of independent synapse formation, individual potential synapses between neurons are transformed into actual synapses with a fixed probability $p$. However, in this model, the sigmoidal function makes it less likely that neurons coupled with low numbers of synapses will retain their connections.

Thus, in this model, the probability for synaptically coupled neurons to make $N_s$ synapses (Equation 3.5) is modulated with a monotonically increasing sigmoidal function $f(N_s)$ (Equation 3.6).

$$A^{CM}(N_s | con) = \frac{1}{P_{con}^{CM}} f(N_s) \sum_{N_p=N_s}^{\infty} P(N_p)B(N_s | N_p), \quad N_s \geq 1$$

$$P_{con}^{CM} = \sum_{N_s=1}^{\infty} f(N_s) \sum_{N_p=N_s}^{\infty} P(N_p)B(N_s | N_p)$$

$P_{con}$ in these expressions is the probability of connection.

Next, we determined the values of the model parameters $p$, $N_s^c$, and $\Delta$, which resulted in distribution functions $A^{CM}(N_s | con)$ that are statistically similar (at 5% significance level) to the experimentally observed distributions of synapse numbers.

Exploring the space of the model parameters $p$, $N_s^c$, and $\Delta$, we identified parameter values for which the model distributions of synapse numbers were not significantly different from the experimental distributions (see Appendix B). Figure 3.4A, C, and E show surfaces that enclose the 95% confidence regions for the values of $p$, $N_s^c$, and $\Delta$. The confidence regions for the three projections are similar in shape but differ in size. Because there are parameter values that are common to all three 95% confidence regions, it is possible that all three projections are governed by the same set of parameters. In the limit of small values of $N_s^c$ and $\Delta$ ($p$-axis in Figure 3.4A, C, and E), the cooperativity function, $f(N_s)$, approaches one for all non-zero numbers of synapses, and the model transforms into the model of independent synapse formation. Figure 3.4A, C,
Figure 3.4. Synapses between pre- and post-synaptic neurons are formed in a cooperative manner. 

A. 95% confidence region for the model parameters $p$, $N^c_s$, and $\Delta$ for local L4→L2/3 projection. Left-right arrow on the $p$-axis indicates the parameter region of the independent synapse formation model. 

B. Bar plot shows the experimental distributions of the numbers of actual synapses for local L4→L2/3 projection. The distributions of actual synapse numbers obtained from the model for the average values of the parameters $p$, $N^c_s$, and $\Delta$ is shown with a green line. Function $f(N_s)$ for this parameter values is shown in black. Gray strip of width $\Delta$, centered at $N^c_s$, marks the cooperative transition window. 

C, D. Same for pairs of nearby neurons in L5. 

E, F. Same for pairs of nearby neurons in L4.
and E clearly show that this parameter regime (left-right arrows) falls outside the 95% confidence regions, again illustrating that the model of independent synapse formation is inconsistent with the experimental data.

Table 3.1 contains the average values of the model parameters $p$, $N_s^c$, and $\Delta$, as well as their 95% confidence intervals. The most salient feature of the results is that the width of the cooperative transition, $\Delta$, is on the order of a single synapse for all projections. Hence, the cooperative transitions are extremely sharp. The high degree of cooperativity becomes evident from the inspection of functions $f(N_s)$ in Figure 3.4B, D, and F (black lines), where these functions were plotted for the average values of $N_s^c$ and $\Delta$. In these figures, gray windows of widths $\Delta$, centered at the critical numbers of synapses, $N_s^c$, demarcate the cooperative transition regions. At $N_s = N_s^c$, the probability for a neuron pair to maintain its connection is 0.5. A single synapse above or below the critical number of synapses transforms this probability into nearly one or zero. In other words, a single synapse in excess of the critical number of synapses is sufficient to stabilize the connection, and, conversely, a single synapse below the critical value will result in the connection loss. Hence, cells appear to have a mechanism for effectively thresholding the numbers of synapses formed with their individual synaptic partners.

<table>
<thead>
<tr>
<th>Projection</th>
<th>$p$</th>
<th>$N_s^c$</th>
<th>$\Delta$</th>
<th>$P_{con}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>L4→L2/3</td>
<td>0.18 ± 0.090 (0.025-0.40)</td>
<td>3.8 ± 0.38 (3.1-5.3)</td>
<td>0.47 ± 0.28 (0.10-1.2)</td>
<td>0.020 ± 0.023 (0-0.11)</td>
</tr>
<tr>
<td>L5→L5</td>
<td>0.19 ± 0.10 (0.025-0.40)</td>
<td>4.9 ± 0.73 (3.1-7.6)</td>
<td>0.98 ± 0.58 (0.10-2.3)</td>
<td>0.040 ± 0.051 (0-0.25)</td>
</tr>
<tr>
<td>L4→L4</td>
<td>0.27 ± 0.13 (0.025-0.60)</td>
<td>2.9 ± 0.90 (1.1-6.4)</td>
<td>1.3 ± 0.79 (0.10-3.5)</td>
<td>0.16 ± 0.15 (0-0.62)</td>
</tr>
</tbody>
</table>

Table 3.1. Results of the cooperative synapse formation model. Numerical values are shown in the mean ± SD (95% confidence interval) format.
Knowing the values of the model parameters $p$, $N_s^c$, and $\Delta$ makes it possible to calculate the probability, $P_{con}$, that a pair of neurons, randomly selected from one of the considered projections, is synaptically coupled. For all the projections, the values of $P_{con}$ calculated based on the cooperative model (see Table 3.1) are in good agreement with the reported experimental probabilities of connection, which are 0.03 for L4→L2/3, 0.10 for L5→L5, and 0.20-0.30 for L4→L4 projections (68, 69, 76). These results provide independent evidence in support of the cooperative synapse formation model.
Chapter 4

Universality of the Distribution of Potential Synapse Numbers

4.1 Introduction

What is the functional form of the distribution of potential synapse numbers, and what is the effect of neuron morphology on this distribution? To answer these questions we first analyzed a model potential synapse distribution generated by two simulated arbors of straight disconnected segments. We showed that the binomial function well describes the resulting distribution. Next, we calculated potential synapse distributions for real neuronal arbors. Here, we made use of a dataset of neurons reconstructed in 3D from different species and brain areas (501 neurons, Figure 4.1 and Appendix D) [http://NeuroMorpho.Org (70)]. Again, the binomial function provided a good fit to the distributions of potential synapse numbers.
4.2 Potential synapse numbers for uniformly distributed arbors

The complexity of neuronal shapes (see Figure 4.1) makes it difficult to understand the relationship between neuron morphology and the distribution of potential synapse numbers. To simplify the problem we generated in silico arbors of straight disconnected axonal and dendritic segments (Figure 4.2), as an alternative to using real reconstructed neurons. Such generated arbors were made of large numbers, \( n_a \) and \( n_d \), of axonal and dendritic segments, randomly distributed and oriented in a box of volume \( V \). Segment lengths, \( l_a \) and \( l_d \), were much shorter than the box dimensions but much longer than the spine length \( s \). The total lengths of axonal and
dendritic arbors inside the box are $L_a = n_a l_a$ and $L_d = n_d l_d$.

Figure 4.2. Potential connectivity for uniformly distributed arbors. Unconnected 5 µm long axonal (black) and dendritic (red) segments are randomly distributed and oriented in a 100x100x100 µm$^3$ box. This simplified model is used for the comparison of the theoretical and numerical results on potential connectivity.

We confirmed (not shown) that the average numbers of potential synapses for such arbors are consistent with the theoretical expression (1),

$$\bar{N}_p = \frac{\pi}{2} \frac{L_a L_d}{V} s .$$

(4.2)

Theoretically, it is also expected that the distribution of potential synapse numbers for such uniform arbors will be binomial in form:
\[ B(N_p | p, n_a n_d) = \binom{n_a n_d}{N_p} p^{N_p} (1 - p)^{n_a n_d - N_p} . \]  

\[ p = \frac{\pi l_a l_d s}{V} . \]  

In this expression, \( p \) is the probability that a single pair of randomly chosen axonal and dendritic segments in the box is potentially connected.

If \( p \) is small and numbers of axonal and dendritic segments are large, which is the case in our numerical simulations (Figure 4.2), the above distribution reduces to the Poisson:

\[ P(N_p | \bar{N}_p) = \frac{e^{-\bar{N}_p} \bar{N}_p^{N_p}}{N_p!} , \]  

where the average number of potential synapses, \( \bar{N}_p \), is provided by Equation 4.2.

**Figure 4.3.** Potential synapse numbers for uniformly distributed segments are well described by a Poisson distribution. **A.** Variance versus average of potential synapse distributions. Blue points correspond to individual distributions and the red line is the best fit. **B.** Fano factor (variance over mean) versus average for potential synapse distributions. The data is consistent with a Poisson distribution where variance equals average and Fano factor is 1.

We confirmed this result in a numerical simulation. Here, distributions of potential synapses were calculated for a large number of realizations of dendritic and axonal arbors. We varied arbors’ length densities, as well as segment lengths, while keeping the spine length constant at \( s = 2\mu m \) (77). Figure 4.3A shows variance vs. average plot for the resulting
distributions. The distribution are consistent with the Poisson, for which the variance equals the average and the Fano factor is 1 (Figure 4.3B).

We next verified that the Poisson distribution indeed provides the best fit to the data. The results are shown in Figure 4.4.

**Figure 4.4.** Distributions of potential synapse numbers for uniformly distributed segments. **A.** Distributions of potential synapse numbers for 1000 realizations of axonal and dendritic arbors. White line is the average of all 1000 distributions. Blue asterisk is the Poisson distribution of the same average number of potential synapses. **B.** Distributions of potential synapse numbers for axonal and dendritic arbors of different densities (blue curves) have different averages. Red lines are the corresponding Poisson distributions.

### 4.3 Potential synapse numbers for pairs of real neurons

Having established that uniform arbors give rise to binomial distributions of potential synapse numbers, we next generate the distributions for real neuronal arbors reconstructed in 3D. Some reconstructed neurons were obtained from the NeuroMorpho.org online database (n = 398 cells). Here, neurons were chosen according to a criterion of having both axonal and dendritic arbors fully reconstructed. Additional cells (n = 103) were taken from the cat V1 and rat barrel cortex datasets (see Appendix C), previously used in Chapters 2 and 3 of this thesis. Since cells in these datasets were shrinkage-corrected and morphed to a standard cat V1 and rat barrel cortex
templates (8, 19), we were able to analyze potential connectivity for all cell pairs within each dataset. Neurons from the cat dataset were also translated laterally in the plane of the slice and rotated around the axis normal to the pia to create more distributions of potential synapse numbers. Overall potential connectivity was analyzed for about 160,000 arbor pairs.

**Figure 4.5.** Potential connectivity between neuronal arbors reconstructed in 3D. A. Axonal (red) and dendritic (blue) arbors of L4 and L2/3 neurons from rat barrel cortex form multiple potential synapses (black circles) (8). B. We generate the distribution of potential synapse numbers by randomly shifting the dendritic arbor 1000 times within a 10 µm box and counting the numbers of potential synaptic clusters for every shift.

For each pair of arbors, the distribution of potential synapse numbers was generated by shifting the dendritic arbor 1000 times, and for each shift, counting the number of potentially connected branch segments. Next, a clustering algorithm was applied to differentiate between segments that belong to the same potential synapse cluster and those that do not. The algorithm calculates the distance (along the axon) between potentially connected segments, and pairs of segments that are more than 10µm apart are grouped into different clusters. The 10µm clustering
threshold roughly corresponds to the average inter-bouton interval along the axon (77). Our results are not sensitive to small variations in this parameter (not shown).

We count the number of clusters for each arbor shift to generate the distribution of potential synapse numbers (Figure 4.5). The variances vs. averages for all the generated distributions are shown in Figure 4.6. The straight line fits have slopes of less than 1, consistent with binomial distributions, for which the slope is $1 - p$ (see Equation 4.3). The scatter in the data is comparable to that for the case of uniform arbors (see Figure 4.3A). We calculated the binomial probabilities of success, $p$, from the slopes of the fits. This resulted in $p = 0.19$ for the excitatory to excitatory and $p = 0.08$ for the inhibitory to inhibitory projections.

![Figure 4.6](image)

**Figure 4.6.** Variance versus average of potential synapse distributions for neurons reconstructed in 3D. **A.** Pairs of excitatory neurons. Blue points correspond to individual distributions and the red line is the best fit, $p = 0.19$. **B.** Same plot for pairs of inhibitory neurons. Inhibitory best fit $p = 0.08$.

Because the distance along the axon was used for clustering, the axonal arbor is effectively broke down into a set of 10 μm segments. Thus, the low value of $p$ for the inhibitory
projection is a consequence of a relatively sparse dendritic arbor of inhibitory neurons (see e.g. Figure 4.1).

![Image: Figure 4.7](image)

**Figure 4.7.** The distributions of potential synapse numbers for reconstructed neurons are well fit (red lines) with a binomial function. **A.** Representative distribution for excitatory neurons, $\bar{N}_p = 4.0, p = 0.20$. **B.** Representative distribution for inhibitory neurons, $\bar{N}_p = 4.9, p = 0.08$. **C.** All potential synapse distributions for excitatory neurons were fit with binomial functions. Goodness of these fits are captured by the $R^2$-coefficients. **D.** Same for all pairs of inhibitory neurons.

Finally, individual distributions of potential synapse numbers were fit with binomial functions by varying a single parameter, $p$. Figures 4.7 A, B show representative fits for inhibitory and excitatory projections. The resulting $R^2$-coefficients, which indicate the goodness
of the fitting procedures, are reported in Figure 4.7 C, D.
Chapter 5

Effect of biological constraints on associative memory storage capacity of artificial neural networks

5.1 Introduction

Many basic features of the synaptic connectivity in cortical networks have become evident over the years from electrophysiological recordings and the analysis of neuron morphology. First, in spite of a large variety of identified neuron classes (34), cortical neurons can be classified into two major groups – excitatory and inhibitory neurons. In the adult cortex, synapses made by inhibitory neurons are believed to be all inhibitory while those made by excitatory neurons – all excitatory (78). The resulting connectivity is predominantly excitatory with only about 15%-20% of inhibitory neurons and inhibitory synapses (34). Second, cortical networks, in addition to
being partitioned into a large numbers of cortical areas, are believed to made of smaller units, from a few tens to tens of thousands of neurons, such as mini-columns (79, 80), structural (19, 81), and functional (56, 61) columns. Third, neurogeometric analysis of cortical neurons (2, 19) shows that nearby neurons within these units have the potential of being interconnected by structural synaptic plasticity (1, 5, 82). In spite of this potential, connectivity within the units is extremely sparse. Even nearby neurons have less than 50% probability of being synaptically coupled (60, 67-69, 83, 84), and this probability decays rapidly with the increase in distance between the neurons beyond 100 µm range (83).

Fourth, electrophysiological recordings from synaptically connected neuron pairs reveal that the distribution of connection strengths between neurons has a long-tail, non-exponential decay (28, 69, 74, 83-88). Finally, it is commonly assumed that similar to many biological networks (see e.g. (89, 90)), cortical networks need to be robust in order to process sensory information reliably, store and retrieve memories. Robustness in the cortical network function is necessary to withstand synaptic failure, fluctuations in post-synaptic potentials, failure in generation or propagation of action potentials, and spontaneous activity. Our analysis of published experimental data (see Table 5.1) revealed that the distributions of connection strengths between neurons in different systems are much more variable than the distributions of neurons’ firing thresholds. For example, coefficients of variation in the connection strength distributions, 0.71-1.1 mV [0.88 ± 0.03 (mean ± standard error), n = 11 systems] (28, 69, 74, 83-88), are significantly higher than those for the distributions of firing thresholds, 0.11-0.33 mV (0.22 ± 0.02, n = 12 systems) (91-95). This suggests that learning in a neural network is primarily mediated by synaptic plasticity, while the threshold of firing remains relatively constant.
<table>
<thead>
<tr>
<th>Brain region</th>
<th>Connection/neuron type</th>
<th>#</th>
<th>Mean [mV]</th>
<th>SD [mV]</th>
<th>CV</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>rat barrel cortex</td>
<td>PC barrel L2/3 → PC barrel L2/3</td>
<td>35</td>
<td>1.0</td>
<td>0.73</td>
<td>0.71</td>
<td>(85)</td>
</tr>
<tr>
<td>rat somatosensory cortex</td>
<td>PC L5 → PC L5</td>
<td>138</td>
<td>1.3</td>
<td>1.1</td>
<td>0.83</td>
<td>(69)</td>
</tr>
<tr>
<td>rat barrel cortex</td>
<td>PC barrel L4 → PC barrel L2/3</td>
<td>64</td>
<td>0.7</td>
<td>0.6</td>
<td>0.86</td>
<td>(68)</td>
</tr>
<tr>
<td>rat barrel cortex</td>
<td>PC barrel L5A → PC barrel L5A</td>
<td>27</td>
<td>1.3</td>
<td>1.3</td>
<td>1.0</td>
<td>(74)</td>
</tr>
<tr>
<td>rat visual cortex</td>
<td>PC L2/3 → PC L2/3</td>
<td>47</td>
<td>0.55</td>
<td>0.49</td>
<td>0.88</td>
<td>(87)</td>
</tr>
<tr>
<td>rat visual and somatosensory cortex</td>
<td>PC L2/3 → PC L2/3</td>
<td>83</td>
<td>0.65</td>
<td>0.64</td>
<td>0.98</td>
<td>(83)</td>
</tr>
<tr>
<td>rat visual and somatosensory cortex</td>
<td>PC L2/3 → FS L2/3</td>
<td>79</td>
<td>3.5</td>
<td>2.5</td>
<td>0.72</td>
<td>(83)</td>
</tr>
<tr>
<td>rat visual and somatosensory cortex</td>
<td>FS L2/3 → PC L2/3</td>
<td>109</td>
<td>3.0</td>
<td>2.5</td>
<td>0.85</td>
<td>(83)</td>
</tr>
<tr>
<td>guinea pig hippocampus</td>
<td>PC CA3 → PC CA1</td>
<td>73</td>
<td>0.13</td>
<td>0.11</td>
<td>0.83</td>
<td>(88)</td>
</tr>
<tr>
<td>rat cerebellar cortex</td>
<td>granule cell → Purkinje cell</td>
<td>104</td>
<td>0.072</td>
<td>0.064</td>
<td>0.89</td>
<td>(28, 96)</td>
</tr>
<tr>
<td>rat visual cortex</td>
<td>PC L5 → PC L5</td>
<td>931</td>
<td>0.77</td>
<td>0.84</td>
<td>1.1</td>
<td>(84)</td>
</tr>
<tr>
<td>rat barrel cortex</td>
<td>PC barrel L3</td>
<td>13</td>
<td>29</td>
<td>4</td>
<td>0.14</td>
<td>(92)</td>
</tr>
<tr>
<td>rat barrel cortex</td>
<td>PC barrel L2</td>
<td>3</td>
<td>31</td>
<td>10</td>
<td>0.32</td>
<td>(92)</td>
</tr>
<tr>
<td>rat barrel cortex</td>
<td>PC septal L3</td>
<td>2</td>
<td>25</td>
<td>4</td>
<td>0.16</td>
<td>(92)</td>
</tr>
<tr>
<td>rat barrel cortex</td>
<td>PC septal L2</td>
<td>6</td>
<td>31</td>
<td>7</td>
<td>0.23</td>
<td>(92)</td>
</tr>
<tr>
<td>rat barrel cortex</td>
<td>PC barrel L5A</td>
<td>16</td>
<td>21</td>
<td>5.6</td>
<td>0.27</td>
<td>(94)</td>
</tr>
<tr>
<td>rat barrel cortex</td>
<td>PC barrel L5B</td>
<td>11</td>
<td>21</td>
<td>3.3</td>
<td>0.16</td>
<td>(94)</td>
</tr>
<tr>
<td>rat barrel cortex</td>
<td>PC L2/3</td>
<td>16</td>
<td>38</td>
<td>4</td>
<td>0.11</td>
<td>(95)</td>
</tr>
<tr>
<td>rat hippocampus</td>
<td>CA1 PC</td>
<td>8</td>
<td>24</td>
<td>8</td>
<td>0.33</td>
<td>(93)</td>
</tr>
<tr>
<td>rat hippocampus</td>
<td>CA1 interneurons</td>
<td>17</td>
<td>25</td>
<td>8</td>
<td>0.32</td>
<td>(93)</td>
</tr>
<tr>
<td>cat spinal cord</td>
<td>low-rheobase α-motoneuron</td>
<td>46</td>
<td>14</td>
<td>3.7</td>
<td>0.26</td>
<td>(91)</td>
</tr>
<tr>
<td>cat spinal cord</td>
<td>medium-rheobase α-motoneuron</td>
<td>38</td>
<td>19</td>
<td>3.5</td>
<td>0.19</td>
<td>(91)</td>
</tr>
<tr>
<td>cat spinal cord</td>
<td>high-rheobase α-motoneuron</td>
<td>44</td>
<td>20</td>
<td>2.4</td>
<td>0.12</td>
<td>(91)</td>
</tr>
</tbody>
</table>

Table 5.1: Examples of variability in the distributions of connection strengths (J) and firing thresholds (h). Abbreviations used in the table: SD – standard deviation, CV – coefficient of variation, PC – pyramidal cell, FS – fast-spiking non-accomodating interneuron, L – layer, barrel/septal indicate barrel/septum related columns.
5.2 The model

We consider a recurrent McCulloch and Pitts neural network model of $N_e$ excitatory and $N_i = N - N_e$ inhibitory neurons. The strengths of the synaptic connections in the network are governed by a connectivity matrix $J$, where a matrix element $J_{i,j}$ describes the strength of connection between the pre-synaptic neuron $i$ and the post-synaptic neuron $j$. We note that a synaptic connection between neurons can be mediated by multiple synaptic contacts, and $J_{i,j}$ describes the overall connection strength. Connection strength is zero for neurons that are not connected, or for neuron pairs connected only with silent synapses.

As all synapses made by inhibitory neurons have to be inhibitory and those made by excitatory neurons must be excitatory, the signs of the elements of the connectivity matrix in every row, $i$, are determined by the pre-synaptic neuron type, $g_i$ (-1 for inhibitory and +1 for excitatory neurons):

$$J_{i,j}g_i \geq 0; \quad i, j = 1, ..., N.$$  

Such networks with constrained connectivity matrices are referred to as “sign-constrained” networks (27). Alternatively, neural networks with no connectivity constraints are termed “unconstrained”.

The state of a neuron $j$ at a time step $t$ is described by a binary number $x_j(t)$, where $x_j(t) = 0$ corresponds to the resting state and $x_j(t) = 1$ to the firing state of the neuron. Neural activity is described by an average firing probability, $p$, which is assumed to be the same for all neurons in the network.

The states of all the neurons are updated simultaneously at each time step by adding the inputs to each neuron and passing the results through a threshold,
\[ x_j(t+1) = \theta \left( \sum_{i=1}^{N} J_{i,j} x_i(t) - h \right); \quad j = 1, \ldots, N. \quad (5.2) \]

Here \( \theta \) is the Heaviside function and \( h \) is the firing threshold, which is assumed to be the same for every neuron. Taking advantage of the binary nature of Equation 5.2 one can conveniently rewrite it as a set of inequalities:

\[ \left( 2x_j(t+1) - 1 \right) \left( \sum_{i=1}^{N} J_{i,j} x_i(t) - h \right) > 0; \quad j = 1, \ldots, N. \quad (5.3) \]

To impose the requirement of robustness with respect to small fluctuations in synaptic strength \( (J_{i,j}) \) and action potential failures or spontaneous neural activity (fluctuations in \( x_j \)) we followed a standard notation \((24, 28)\) and made the above inequalities more stringent by introducing a stability parameter, \( \kappa > 0 \) (see Supplementary Information for details):

\[ \left( 2x_j(t+1) - 1 \right) \left( \sum_{i=1}^{N} J_{i,j} x_i(t) - h \right) > \frac{\kappa h}{\sqrt{N}}; \quad j = 1, \ldots, N. \quad (5.4) \]

For non-zero \( \kappa \), slight perturbations that can affect the system do not necessarily change its dynamics.

Neural networks in Equation 5.4 can learn to transition from a given state at time \( t \), described by the activities of all neurons, \( \zeta_i \), to another predetermined state, \( \omega_i \), at time step \( t+1 \). Such associations of consecutive network states may be important for the high fidelity in the propagation of neural activity observed \textit{in vivo} \((97)\) and for the recall of memories. In fact, neural networks described by Equation 5.4 can learn to implement a predetermined set of \( m \) associations:

\[ \left( 2\alpha_j - 1 \right) \left( \frac{1}{N} \sum_{i=1}^{N} J_{i,j} \zeta_i^{\mu} - h \right) > \frac{\kappa}{\sqrt{N}} \quad \text{for} \quad \mu = 1, \ldots, m \quad (5.5) \]

Index \( \mu \) in this set of inequalities enumerates the pairs of associations \( \{ \zeta^{\mu}, \omega^{\mu} \} \).
For independent associations, the set of inequalities (5.5) can be decoupled along the presynaptic index \( j \), and the problem breaks down into \( N \) independent perceptron-like systems:

\[
(2\alpha^\mu - 1) \left( \frac{1}{N} \sum_{j=1}^{N} J_j \xi_j^\mu - 1 \right) > \frac{\kappa}{\sqrt{N}}; \quad \mu = 1, \ldots, m
\]

\[
J_i g_i > 0, \quad i = 1, \ldots, N
\]

\[
\xi_i^\mu, \omega^\mu \in P = \begin{cases} 
0, & 1 - p \\
1, & p 
\end{cases}
\]

(5.6)

**Figure 5.1.** Perceptron models and their critical capacity. A. Solid lines show how the probability of finding a solution in different perceptron-like models depends on the number of presented associations, \( m \) (\( \alpha = m/N \)). Critical capacity, \( \alpha_c \), is defined as the number of associations per input which can be implemented with 50% probability (dashed arrows). Critical capacity is 2 for an unconstrained and 1 for a sign-constrained perceptron model. Critical capacities in these models do not depend on the fraction of inhibitory inputs, \( N_i/N \), or the value of the robustness, \( \kappa \). In the presented model, critical capacity depends strongly on \( N_i/N \), \( \kappa \), and \( p \). Solid lines show theoretical results (22) and dotted lines are the results of numerical simulations for \( N = 50 \) inputs. B. Geometric interpretation of the presented model for a neuron receiving only 2 inputs. The two sets of inequalities in Equation 5.6 parse the space of possible connection strengths, \( J_1 \) and \( J_2 \), with hyper-planes: padded (dashed) lines for the first set and the figure axes for the second. The red points mark all possible solution regions (if solution exists). Any point from the region of solutions satisfies the constraints set by Equation 5.6.

The ratio between the number of associations a network can learn and the number of neurons it contains is referred to as the network’s capacity, \( \alpha = m/N \). Figure 5.1A shows how the
probability of such learning depends on the number of associations, \( m \).

Because with increasing \( m \) it becomes progressively more difficult to find connection strengths which will fulfill Equation 5.6, the probability of successful learning is a decaying function of \( m \). Network capacity, which can be learned with 50% probability is known as the critical capacity, \( \alpha_c \) (24). Figure 5.1A illustrates that \( \alpha_c = 2 \) for the standard unconstrained perceptron (22, 24), and \( \alpha_c = 1 \) in the sign-constrained case (27, 29), independent of the fraction of inhibitory neurons, \( N_i/N \), and robustness, \( \kappa \). This is not the case for a biologically constrained network considered here.

5.3 The solution

The solution to the problem of Equation 5.6 (if it exists) is typically not unique. This is illustrated geometrically in Figure 5.1B. The axes in this figure parse the space of all connection strengths into positive (excitatory) and negative (inhibitory) regions. The solution has to be confined to one of these regions due to the constraints on the signs of all connections. The dashed lines in Figure 5.1B represent hyper-planes which additionally divide the space as specified by the first set of inequalities in Equation 5.6. The set of solutions (any of the regions in Figure 5.1B marked with a red dot) is necessarily convex. This is because for any two networks \( J_a \) and \( J_b \) from the set of possible solutions of Equation 5.6, their linear combination in the form \( \theta J_a + (1 - \theta) J_b \), where \( 0 \leq \theta \leq 1 \), is a solution as well. In order to select a unique solution we take into account the fact that stronger synapses are larger in volume (98) and, as a result, are more costly to the organism (99). To keep this cost to the minimum, we choose the solution which minimizes the absolute connection strength \( \sum_{i=1}^{N} |J_i| \). Needless to say, that this
choice has no effect on the critical capacity of the network.

Figure 5.2. Critical capacity of the presented model. A. Critical capacity as a function of the fraction of inhibitory neurons for $p = 0.5$ and different values of robustness, $\kappa$. B. Critical capacity as a function of the fraction of inhibitory neurons for $\kappa = 1$ and different values of firing probability, $p$. Solid lines show theoretical results according to Equations D.27 of Appendix D, and dotted lines are the results of the corresponding numerical simulations for $N = 50$ excitatory and inhibitory neurons.

Figures 5.2 and 5.3 show how the critical capacity of the biologically constrained model, Equation 5.6, depends on the model parameters (see Appendix D for details): the fraction of inhibitory neurons, $N_i/N$, the robustness, $\kappa$, and the firing probability, $p$. For networks of mostly excitatory neurons ($N_i/N < 0.5$), the critical capacity is a decreasing function of robustness. This trend is due to the fact that some of the solution regions (see Figure 5.1B) close as $\kappa$ increases, decreasing the probability of finding a solution for a random set of associations. For the fully robust networks, $\kappa \to \infty$, all inner regions of solutions close and $\alpha_c$ converges to its minimum value (black line in Figure 5.2A). For mostly inhibitory networks ($N_i/N > 0.5$), the regions of solutions are open, Figure 5.1B, and $\alpha_c$ is independent of $\kappa$ (see Figure 5.2A). As expected, the
critical capacity, by enlarge, increases with increasing correlations in the set of associations (deviations of $p$ from 0.5), Figure 5.2B.

![Figure 5.3](image)

**Figure 5.3.** Dependence of critical capacity, $\alpha_c$, on the parameters of the presented model: $N_i/N$, $p$, and $\kappa$. **A.** $\kappa = 0$, **B.** $\kappa = 0.1$, **C.** $\kappa = 1$, and **D.** $\kappa = \infty$. Black pixels in the colormap correspond to infinite critical capacity. Gray pixels mark parameters for which critical capacity cannot be defined.

A characteristic feature of sign-constrained neural networks is the finite fraction of zero-strength connections (27, 28). In the brain, the mechanisms of excitation and inhibition differ in the type of neurotransmitter used and, regardless of the level of fluctuations in the synaptic
transmission, inhibitory synapses can never become excitatory and vice versa. This clear
delineation of synapses into excitatory and inhibitory (the axes in Figure 5.1B are not padded),
results in solutions with zero-strength connections. Large fractions of zero-strength connections
are consistent with the experimentally observed low connection probabilities in many cortical
networks. Figure 5.4 illustrates how the connection probability, $P_{\text{con}}$, in the considered model
depends on $N_i/N$, $\kappa$, and $p$.

![Figure 5.4](image)

**Figure 5.4.** Dependence of connection probability, $P_{\text{con}}$, on the parameters of the presented
model: $N_i/N$, $p$, and $\kappa$. **A.** $\kappa = 0$, **B.** $\kappa = 0.1$, **C.** $\kappa = 1$, and **D.** $\kappa = \infty$. Gray pixels mark parameters
for which connection probability cannot be defined.
For all values of these parameters, $P_{\text{con}}$ remains below 0.5. In networks of mostly excitatory neurons, $P_{\text{con}}$ decreases with increasing $\kappa$, as the interior solution regions in Figure 5.1B, which do not contain zero-strength connections, close.

Figure 5.5. Dependence of the average absolute connection strength, $\langle|J|\rangle$, on the parameters of the presented model: $N_i/N$, $p$, and $\kappa$. A. $\kappa = 0$, B. $\kappa = 0.1$, C. $\kappa = 1$, and D. $\kappa = \infty$, $N_i = 0$. In the latter case, $\langle|J|\rangle$ diverges for any non-zero fractions of inhibitory neurons. Black pixels in the colormap correspond to infinite $\langle|J|\rangle$. Gray pixels mark parameters for which $\langle|J|\rangle$ cannot be defined.
The distribution of connection strengths in the presented model becomes wider with the increase in robustness and with the increase in the fraction of inhibitory neurons, Figure 5.5. This is understandable, because by increasing $\kappa$ the solutions get pushed away from the zero connection strength point, Figure 5.1B, and an increase in $N_i/N$ must be accompanied with an increase in excitation to insure that neurons would reach the threshold of firing. For some combinations of $N_i/N$, $\kappa$, and $p$, the solution becomes unstable and the absolute average synaptic strength diverges (see the interface of stability in Figure 5.5). This instability results from the fact that the solution regions are open for some values of the model parameters, and $<|J|>$ diverges due to the assumption of the uniform prior distribution of connection strengths (see Appendix D for details).

Figure 5.6. A. Distribution of connection strengths. A. Red line shows the theoretical distribution of connection strengths in the presented model with 10 inhibitory and 40 excitatory neurons, $p = 0.5$, $\kappa = 0.1$. Red bar at zero indicates the theoretical number of zero-strength connections. Blue bars are the corresponding result of numerical simulation. B. A log-log plot of the connection strength distribution for excitatory connections. Blue dots correspond to the results of the numerical simulation. The cyan line shows the same data binned into 10 equal-sized bins on the logarithmic scale. Gaussian fit (red line) fails to depict the long-tail of the connection strength distribution obtained numerically. Error-bars indicate standard errors assuming Poisson distribution of counts in each bin.
It has been theorized that the distribution of connection strengths in all-excitatory, sign-constrained networks is a truncated Gaussian with a finite fraction of zero-strength connections (27, 28). In contrast, electrophysiological recordings in different cortical systems reveal that the distribution of connection strengths has a long non-Gaussian tail (28, 69, 74, 83-88). Similar to (27, 28), our theoretical analysis of networks of excitatory and inhibitory neurons (see Appendix D for details) resulted in connection strength distributions that are composed of two truncated Gaussians (one for excitatory and one for inhibitory connections) and a finite fraction of zero-strength connections. An example of such a distribution is shown in Figure 5.6A (red line). Figure 5.6A (blue bars) also shows the distribution of connection strengths obtained numerically for the corresponding network of 10 inhibitory and 40 excitatory neurons. The central region of the distribution can be fit well with two truncated Gaussian functions; however, this fit fails in the tail regions. This becomes evident from the examination of the distributions on a log-log plot (Figure 5.6B), where the numerical distribution displays the scale-free feature observed experimentally (84).
Chapter 6

Conclusions

In the first part of the thesis, we developed a theoretical framework which describes the capacity of cortical circuits to change with structural spine remodeling. We considered and compared two possible scenarios of structural plasticity: one where the initial synaptic contact can be made between a dendritic spine and a point on an axon, preceding bouton formation (1), and the other where the initial synaptic contacts are preferentially established with pre-existing synaptic boutons (32). Locations in the neuropil where synaptic contacts between axonal and dendritic branches can be made with new spine growth are termed potential synapses. Such axo-dendritic oppositions occur at different relative distances between neurites and, as a result, can be bridged by spines of different lengths. Hence, the shape of spine length distribution function is expected to reflect the neuropil’s potential for spine remodeling.

We calculated the density of potential synapses in the neuropil for the two quantitative models of spine remodeling plasticity. We derived an expression for the probability of an actual
synapse being present at a potential synaptic site. This probability, or the connectivity fraction [formerly referred to as the filling fraction (1)], was calculated as a function of distance between axonal and dendritic branches. We estimated the number of structurally different circuits that can be achieved by spine remodeling. This number, or structural synaptic entropy, is a measure of the neuropil’s potential for circuit reorganization. Furthermore, we evaluated the selectivity of cortical neurons in their choice of synaptic partners. This selectivity, or probability for one neuron to establish a synaptic connection with another neuron encountered with a new spine growth, gives a sense of the number of functionally different classes of excitatory neurons in a small cortical region. Based on published anatomical data, we estimated and compared the above parameters of neuropil micro-architecture in mouse occipital cortex, rat CA1, monkey V1, and human temporal cortex. Our results showed marked differences in the potential for structural synaptic reorganization in these systems.

In Chapter 3, the results are two-fold. First, we showed that the idea of independent synapse formation between pre- and post-synaptic neurons is inconsistent with the experimental datasets of neuron morphology and synaptic connectivity. Second, we proposed a cooperative model of synapse formation, which led to a good agreement with the experimentally observed distributions of synapse numbers and the probabilities of finding synaptically coupled pairs of neurons. Our numerical results suggest that cells effectively threshold the numbers of synapses formed with their individual synaptic partners. Strong multiple-synaptic connections, which exceed the critical number of synapses, are stabilized, whereas weak connections are degraded. The critical values of synapse numbers may differ among various projections.

In Chapter 4, we examined the effect of arbor morphology on the shape of the distribution of potential synapse numbers. We first analyzed randomly distributed axonal and dendritic arbors
generated *in silico*, and we showed that the distributions of potential synapse numbers are well described with the binomial function. Next, we applied the same methodology to a large dataset of axonal and dendritic arbors reconstructed in 3D from different species and brain regions. We showed that the binomial distribution provides a good fit in the case of real neurons as well. The distributions clustered into two main groups: inhibitory to inhibitory projections, with a binomial probability of success $p = 0.08$, and excitatory to excitatory projections with $p = 0.19$. Our results show that the apparent complexities of axonal and dendrite arbor shapes of cortical neurons do not give rise to a complicated description of structural connectivity. The distributions of potential synapse numbers are universally depicted with the binomial function of two parameters. One of these parameters, the average number of potential synapses, is defined by the arbor densities and the volume of the overlap region. The second parameter, $p$, is determined by the type of the structural projection. This result should make it possible to perform structurally realistic large-scale neural network simulations without the need of the time-consuming neuronal reconstructions.

In the last chapter of the thesis, we presented a neural network model which qualitatively explains many basic features of cortical connectivity. Analogous to real cortical networks, the presented robust model of excitatory and inhibitory neurons has a high fraction of zero-strength synapses with the probability of connection in the 0 to 50% range. The critical capacity of our model improves from increasing the fraction of inhibitory neurons up to 50%. However, this improvement comes at a price; it is associated with an increase in the average absolute connection strength. Our results suggest that sparse connectivity and small fraction of inhibitory neurons and synapses in cortical networks may result from the trade-off among such factors as network capacity, robustness, and metabolic cost. What is more, the distributions of connection
strengths which are obtained in numerical simulations exhibit the long-tail non-exponential decay observed in experiments.
Appendix A

Assumptions and approximations

Analysis of Chapter 2 relies on several realistic assumptions and approximations. In the following we provide a list of assumptions specifying with A or B the model of potential connectivity for which these assumptions are needed.

- We treated axonal (A) and dendritic (A, B) branches as straight segments on the scale of the average spine length (1-2μm). This assumption holds for most excitatory neurons, including those in neocortex and hippocampus. It is justified by low tortuosity of dendritic and axonal branches of excitatory neurons on a micrometer scale (62).

- In calculating the numbers of potential synapses we used an approximation where the average axonal (A) and dendritic (A, B) branch lengths were assumed to be much longer than the average spine length. Branch here was defined as a neuron process extending from soma or a bifurcation point to a successive bifurcation or an end point. This approximation ignores the
corrections to the potential synapse count due to branch tips and bifurcation points. This approximation is certainly valid for the excitatory neurons in the cerebral cortex.

- We made a simplifying assumption that axonal (A) and dendritic (A, B) branches of excitatory neurons are randomly distributed in the neuropil. Inhomogeneous distributions of branches would lead to spatial variations in the values of the calculated parameters. Yet, this should not affect the average values calculated in this Chapter 2. This is justified in part by the absence of correlations between the positions of axonal and dendritic branches (axons on average are not “attracted to” or “repelled from” dendrites) (62).

- We assumed that synaptic contacts can occur anywhere along axonal (A) and dendritic (A, B) branches. In other words, there are no hot-spots for spine and bouton formation. The presence of hot-spots on dendritic branches would not change the results of our calculations because the average inter-spine interval is already smaller than the average spine length.

- Even though below in this section the results for model A are derived for an arbitrary distribution of angles between axonal and dendritic branches in the neuropil, in our calculations we made a simplifying assumption that this distribution is isotropic. This assumption had been verified (unpublished) on the dataset of cortical excitatory neurons reconstructed in 3D (19). It is already valid if the orientations of axonal and/or dendritic branches are isotropic. Anisotropic distribution of angles could lead to a slightly different numerical coefficient.

- In our calculations we made a simplifying approximation in which the number of asymmetric synapses is approximately equal to the number of boutons on excitatory neuron axons (A, B). This is because the majority of boutons on excitatory axons contain only a single asymmetric synapse. This approximation is supported by low fractions of multiple synapse boutons: 0.026 in cat primary visual cortex (12, 100), 0.16 in rat motor cortex (101), 0.18-0.24 in
rat hippocampus (102, 103), and 0.14 in mouse barrel cortex (32). The average asymmetric synapse to bouton ratio is 1.08 in mouse neocortex (34). For asymmetric synapse to bouton ratios much larger than one, it would be necessary to use more precise expressions Equations 2.9.

- Another simplifying approximation made in our calculations is that the volume density of spines of excitatory neuron dendrites is equal to the volume density of asymmetric synapses (A, B). This approximation is supported by generally low fractions of asymmetric synapses that are not made on dendritic spines of excitatory neurons. For example, in mouse neocortex this fraction is estimated at 0.13 (34). For large fractions of such synapses it would be necessary to use Equation 2.13.

- We did not take into consideration volume exclusion effects in the neuropil (A, B), assuming that neurites are flexible and can be easily deformed to accommodate synaptic connections. Such effects could be important near the dendritic shafts where the space is partly occupied by inhibitory axons and their shaft synapses. This may effectively push the excitatory axons farther out. Hence, our results may not be accurate for small values of the parameter $s$ (less than 0.5µm).

- We assumed that dendritic spines are straight segments that extend perpendicularly to the dendritic shafts (A, B). In model A we made an additional assumption that dendritic spines bridge the gaps between potentially connected axons and dendrites along the shortest paths. Thus, in this model a spine connecting an axon and a dendrite is perpendicular to both branches. We address these limitations of our models later in Appendix A.

- Optical measurements of spine length distribution functions may not be accurate. First, due to the limited resolution, spine lengths shorter than 0.5µm cannot be measured accurately. Second, some dendritic spines can be overshadowed by the dendrite. Third, spines that do not lie
entirely in the focal plane will appear shorter, shifting the spine length distribution function to
the left. In the experimental data used in this Chapter 2 the last two effects were minimized by
only measuring spines which lie in the focal plane on both sides of the dendritic branch.

- One of the approximations made in this Chapter 2 is that the density of asymmetric
synapses is close to the density of excitatory boutons. Though the expressions for the
connectivity fractions, Equations 2.9, were derived for an arbitrary asymmetric synapse to
bouton ratio, in all the numerical results this fraction was set to one. In general, asymmetric
synapse to bouton ratio is greater than one and it would proportionally increase the connectivity
parameter, Equations 2.9, and decrease the structural synaptic entropy per spine. However, this
change is expected to be small due to the slow dependence of the structural synaptic entropy on
the connectivity parameter in the 0.1-0.4 range (see Figure 2.4).

- Another approximation made in our model was in treating dendritic spines as straight
segments that extend perpendicularly from pre-synaptic dendritic branches connecting them to
post-synaptic targets. Though by and large dendritic spines in light microscopy images appear to
be quite straight and extend predominantly perpendicularly to dendritic shafts, our results
underestimate the amount of structural synaptic entropy. This is because we did not account for
the excess entropy associated with spine conformations which could result in new connectivity
patterns. The number of potential synapses accessible to spines in the $[s-\Delta s/2, s+\Delta s/2]$ length range is proportional to $\Delta A_4 = 2b_{s}\Delta s$ in model A, the cross-sectional area per spine of a
cylindrical shell surrounding an axon where the pre-synaptic dendritic branches are located (see
Chapter 2.3). In model B this number is proportional to $\Delta A_b = 2\pi(s+\delta)\Delta s$, which is the cross-
sectional area of a spherical shell surrounding a bouton. With new spine conformations these
cross-sectional areas increase to $\tilde{A}_{A,B}$, leading to a proportional increase in the number of
potential synapses and a decrease in the connectivity fraction, \( \tilde{f}_{A,B}(s) = \frac{f_{A,B}(s) A_{A,B}}{\hat{A}_{A,B}} \). To estimate the increase in structural entropy per spine due to this effect, we note that in the limit of small connectivity fractions, the second term in Equation 2.15 can be neglected and

\[
\tilde{i}_{A,B} / n_s \approx -\int_0^\infty \log_2 \left( \tilde{f}_{A,B}(s) \right) p(s) ds = \hat{i}_{A,B} / n_s + \int_0^\infty \log_2 \left( \frac{\tilde{A}_{A,B}}{\hat{A}_{A,B}} \right) p(s) ds .
\]

Hence, due to different spine conformations, structural entropy per spine increases in the amount equal to the average logarithm of \( \frac{\tilde{A}_{A,B}}{\hat{A}_{A,B}} \), and can be large in absolute terms. However, the role of this term in comparing structural entropy per spine between different systems is much smaller. To illustrate this point consider comparison between two systems, with spine length distribution functions \( p_1(s) \) and \( p_2(s) \), in model A. Difference in the conformational entropy terms in these systems is

\[
\int_0^\infty \log_2 \left( \frac{\hat{A}_A}{\hat{A}_A} \right) [p_1(s) - p_2(s)] ds .
\]

Due to the fact that the spine length distribution functions are normalized to unity this expression reduces to

\[
\int_0^\infty \log_2 \left( \frac{\hat{A}_A}{\hat{A}_A} \right) [p_1(s) - p_2(s)] ds .
\]

It is not clear how \( \hat{A}_A \) depends on \( s \) but, because of the slow nature of the logarithm function, the resulting integral is small (fraction of one bit for the considered systems) for all reasonable dependences. Moreover, as \( \hat{A}_A \) is expected to increase with \( s \), accounting for conformational entropy of dendritic spines only enhances the difference in structural entropy per spine between rodent and primate groups.

- The values of parameters of structural plasticity calculated in this Chapter 2 are only as good as the anatomical data that were used for the calculations. There is substantial variability in some anatomical measurements performed in different laboratories. One part of this variability is biological in nature. Parameters of cortical neuropil can be highly variable among individual
animals (same species, age, and brain area). This inter-brain variability is reflected in the error bars reported in this Chapter 2. We made conservative estimates of standard errors to the mean values of all the calculated parameters. As described in Chapter 2.3 true standard errors are expected to be much smaller. Another part of variability in anatomical measurements, not captured by our error bars, is due to different biases introduced by different experimental procedures. To minimize the effect of experimental biases, when possible, we used results that are corrected for experimental artifacts (tissue shrinkage or hidden spines), are based on large numbers of animals, and results that are agreed upon by several experimental laboratories. For example, the volume density of asymmetric synapses in layer 3 of mouse occipital cortex reported in (34, 36) is $0.91\mu\text{m}^{-3}$ (n=3 mice). However, the density of asymmetric synapses averaged across all cortical layers of mouse visual cortex was estimated at $2.2\mu\text{m}^{-3}$ (n=1 mouse) in (54). We chose not to use the latter density because it is based on measurements from a single animal, and it is averaged across all cortical layers, thus the exact value for layer 3 is not known. It is likely that the discrepancy in this case could have resulted from methodological differences.

The results for the mouse occipital cortex would change significantly if the $2.2\mu\text{m}^{-3}$ density of asymmetric synapses was used. In this case the connectivity parameters would decrease to $f_A^* = 0.13\pm(0.03)$, $f_B^* = 0.14\pm(0.02)$, and the average connectivity fractions to $\bar{f}_A = 0.075\pm(0.015)$, $\bar{f}_B = 0.056\pm(0.009)$. The average entropy per spine would increase to $5.3\pm(0.3)$ in model A and to $5.9\pm(0.2)$ in model B, and the average entropy per volume to $11.7\pm(1.6)$ and $12.9\pm(1.7)$ respectively.

By analyzing anatomical parameters measured in different brains we ignored possible correlations among these parameters. These correlations could significantly affect the results of
our calculation. For example, animals with low densities of asymmetric synapses could have higher dendritic or axonal length densities. In result, the connectivity fraction in these animals would be lower than predicted. To resolve this issue, all the parameters present in Table 2.1 must be measured in the same tissue. Such efforts are already under way. One promising approach here is an automated 3D reconstruction of neuropil volume imaged with electron microscopy (104-106). This method has the potential to measure reliably spine length distribution, inter-bouton and inter-spine intervals, and the density of asymmetric synapses, all in the same small neuropil volume. It will undoubtedly lead to more precise estimates of the connectivity fraction and structural synaptic entropy.

Estimation of error-bars

In Chapter 2 we compared parameters of structural synaptic plasticity in different systems. Such comparisons are hindered by generally large variability in anatomical measurements, and significant differences (if any) can only be observed on the level of the mean values. Hence, it is necessary to provide quantitative estimates of the standard errors to the mean values (sem). As the raw experimental data for the citations in Table 2.1 were not available to us, our strategy here was to estimate the upper bounds of the sem’s from the reported standard deviations (std) and the numbers of measurements.

Consider an experiment where a statistical measurement is preformed in \(n\) brains from multiple samples per brain. The samples are then pooled together and the mean and total variance (\(\text{var} = \text{std}^2\)) are reported as results of the experiment. Table 2.1 shows a number of such experiments where the inter-bouton interval along an axon, the spine density on a dendrite, and the density of asymmetric synapses in different systems were measured. To access inter-brain
variability we would like to estimate the variance in the average measurements for different brains. For this we note that the total variance is equal to the sum of the average within-brain variance and variance in the averages for different brains. Hence, variance in the averages for different brains is always less or equal to the total variance. Then the upper bound of the standard error to the mean is equal to the square root of the total variance divided by \( n \), sem \( \leq \sqrt{\text{var/n}} = \text{std} / \sqrt{n} \). Standard deviation of all measurements pooled together divided by the square root of the number of brains provides a conservative estimate of inter-brain sem.

When calculating the product of two or more measured variables we propagated their sem’s using the Monte Carlo procedure. Here, we took advantage of the central limit theorem (55) and for each variable generated a set of 10,000 elements from a Gaussian distribution with the measured mean value and the sem estimated as described above. Next, we calculated the product between all members of the sets and obtained the mean and the sem for the product. This is how the mean and the sem values were estimated for all the calculated quantities, including the dendritic length density \( \rho_d \) from Table 2.1 and all the parameters from Table 2.2. To test for the pairwise difference in the means of the calculated parameters and to obtain the corresponding p-values we used the Student’s t-test for samples with unequal variance.

**Anatomical data**

To evaluate the connectivity fraction and the capacity of neural circuits to undergo structural remodeling, in Chapter 2 we used previously published anatomical data from some of the best-studied cortical systems. Our choice of the cortical systems was primarily restricted by the availability of the spine length distribution function and we confined our analysis to mouse
occipital cortex, rat CA1, monkey V1, and human temporal cortex. The anatomical parameters of circuit micro-architecture in these areas are provided in Table 1. Below we give a detailed account of how these parameters were selected. Unless stated otherwise, the numerical values are shown in the mean±std(sem) format.

**Mouse occipital cortex:** Anatomical data for the mouse occipital cortex was based on measurements from adult mice. The distributions of spine neck lengths and spine head areas were measured on basal dendrites of layer 3 pyramidal cells of 2 mice (33). As there is no correlation between spine neck length and spine head area (33), these distributions were used to generate the spine length (neck length + head diameter) distribution. This was done with the following Monte Carlo procedure. 10,000 spine neck lengths and head areas were sampled from their corresponding distributions. The average spine head diameters were calculated for each spine head. These diameters were randomly associated with different spine neck lengths to generate the spine length distribution function (Figure 2.3A1). The average spine length obtained from this distribution is $0.99±(0.01)\mu m$ ($n=2$ mice, 1226 spines). The correction for hidden spines (spines above and below the dendrite) was deemed unnecessary, as only lateral spines were reconstructed in (33). Parameter $\delta$ was estimated at 0.70$\mu m$ which is the sum of the average radius of dendritic branches, 0.45$\mu m$ (34), and the average radius of synaptic boutons, 0.25$\mu m$ (34). Though, this value of the parameter $\delta$ was based on several cortical regions and does not contain error bars, it does not affect strongly the results of our theory. This parameter is only present in model B, Equations 2.9, and has only an effect on our results in the region of small spine lengths. The average inter-bouton interval was estimated to be $4.5±1.4(0.47)\mu m$ ($n=9$ mice, 20 cells) based on measurements in layers 2-4 of different cortical areas (34-36). As no significant variations in $b_a$ between cortical areas is reported (35), we used this value to
represent the inter-bouton interval in the mouse occipital cortex. Based on (34, 37) the average spine density along a dendrite, $1/b_d$, is $1.94\pm0.42(0.24)\mu m^{-1}$ (n=3 mice, 10 cells). The volume density of asymmetric synapses in layer 3 of mouse occipital cortex is $0.91\pm0.25(0.15)\mu m^{-3}$ (n=3 mice, 3 blocks of tissue). This value was calculated as the product of the total synapse density, $1.05\pm0.29(0.17)\mu m^{-3}$, and the 0.87 fraction of asymmetric synapses (34, 36). We would like to mention that a significantly higher estimate of the average density of asymmetric synapses, $2.2\mu m^{-3}$ (n=1 mouse), was reported in (54). We did not use this measurement as it is based on a single animal and is averaged across all cortical layers. However, the discrepancy in the densities of asymmetric synapses may have resulted from methodological differences. This point is further addressed in the discussion section. The dendritic length density calculated as the product of the inter-spine interval and the density of asymmetric synapses is $0.48\pm(0.10)\mu m^{-2}$, implying that one cubic micrometer of neuropil in layer 3 of adult mouse occipital cortex contains on average $0.48\mu m$ of dendritic length.

**Rat CA1:** Anatomical data for the rat hippocampus was based on measurements in stratum radiatum of CA1 region in adult Long-Evans and Wistar rats. Here, the spine length distribution function (Figure 2.3A2) was derived from the work of (38) (n=2 rats, 485 spines). In this work the cumulative distribution is scaled to the median spine length of $1\mu m$. Hence, in generating the distribution in Figure 2.3A2 we differentiated the original cumulative distribution and rescaled it to the average spine length of $1.08\pm(0.03)\mu m$ (n=4 rats, 351 spines). This value of spine length was determined by averaging the results from (39), $0.95\pm0.42\mu m$ (n=3 rats, 100 spines), and (40), $1.21\pm0.43\mu m$ (n=1 rats, 251 spines). The value of the average radius of dendritic branches in stratum radiatum, $0.30\mu m$, was obtained by averaging the results from (39),

84
0.28µm (n=3 rats, 7 dendritic segments), (40), 0.36µm (n=3 rats, 3 cells), and (41), 0.25µm (n=7 rats, 26 dendritic segments). This value was added to the average radius of synaptic boutons, 0.2µm (n=2 rats, 224 varicosities) (42), resulting in δ = 0.5µm. The average inter-bouton interval in stratum radiatum of rat CA1, 3.7±0.6(0.3)µm (n=5 rats, 1909 varicosities), was based on measurements for CA3 axons projecting to CA1 (43). This value of inter-bouton interval is consistent with the value of 3.0±1.4µm reported in (42) which has to be corrected for tissue shrinkage by an estimated 10-25% (42). The density of spines on a dendrite, 1/b_d, in stratum radiatum of rat CA1, 3.41±1.05(0.40)µm⁻¹ (n=7 rats, 20 cells), was obtained from the work of (41). This value is roughly in the middle between the two other reported measurements of spine density, 3.03±0.83(0.59)µm⁻¹ (n=2 rats, 15 dendrites) (44), and 3.80±0.76(0.54)µm⁻¹ (n=2 rats, 26 dendrites) (38). Pooling all the data together would not result in a significant change in the density of spines. Using the density of asymmetric synapses of 2.0±0.30(0.13)µm⁻³ (n=5 rats, 5 blocks of tissue) (45) we estimated that the dendritic length density in stratum radiatum of rat CA1 is 0.59±(0.08)µm⁻². This result is similar to that in the mouse occipital cortex.

**Monkey V1:** This data was based on measurements from layer 3 of adult Macaque monkey primary visual cortex. Here, the distribution of spine lengths on basal dendrites was obtained from the work of (1), where the spine length was measured between a point on the dendritic axis closest to the base of the spine and the tip of the spine head. As in the present work spine length was measured from the dendritic surface, the original distribution was shifted to the left by 0.70µm, the amount corresponding to the average radius of basal dendrites (1). The resulting distribution (Figure 2.3A3) has the average spine length of 1.86±(0.05)µm (n=3 monkeys, 233 spines). Parameter δ was estimated as the sum of the average radius of basal
dendritic branches, 0.70\(\mu\)m, and the generic average radius of synaptic boutons, 0.25\(\mu\)m. As stated above, this parameter does not strongly affect the results of our calculations. We obtained the average inter-bouton interval in layer 3 of monkey V1 from the work of (46). The inter-bouton histogram in this Chapter 2 is consistent with an average inter-bouton interval of 5.6\(\pm\)2.4(1.2)\(\mu\)m (n=4 monkeys, 400 varicosities) reported as 180 boutons/mm. To estimate the average spine density on basal dendrites we pooled together data from four different studies (47-50) performed in Macaque primary visual cortex. Our estimate of \(1/b_d\), which was based on the reported measurements of spine density and the number of branches as a function of distance from soma, resulted in 0.55\(\pm\)0.07(0.03)\(\mu\)m\(^{-1}\) (n=5 monkeys). The volume density of asymmetric synapses was calculated as the product of the volume density of all synapses in layer 3 of adult Macaque monkey V1, 0.34\(\pm\)0.04(0.02)\(\mu\)m\(^3\) (3 monkeys, 9 blocks of tissue) (51), and the fraction of asymmetric synapses, 0.76\(\pm\)0.09(0.06) (n=2 monkeys, 5 blocks of tissue) (52), resulting in 0.26\(\pm\)0.04(0.03)\(\mu\)m\(^3\). This estimate lead to the average dendritic length density of 0.47\(\pm\)(0.05)\(\mu\)m\(^2\).

**Human temporal cortex:** The spine length distribution function was derived from the work of (33). Here the distributions of spine neck lengths and spine head areas were measured on basal dendrites of layer 3 pyramidal cells in temporal cortices of 2 adult male patients. The spine length distribution (Figure 2.3A4) was generated in the same way as for mouse occipital cortex, resulting in the average spine length of 1.42\(\pm\)(0.01)\(\mu\)m (n=2 humans, 2768 spines). The value of parameter \(\delta\), which in this case is equal to the sum of the average radii of second order dendritic branches and synaptic boutons, was estimated as 1.1\(\mu\)m +0.25\(\mu\)m. The former number was based on the analysis of published neuron images (33) and the latter is the generic value of the average bouton radius. We did not find a reliable estimate of the average inter-bouton interval in human
temporal cortex and thus, we provided results for model B only. The value of the average spine density on basal dendrites of layer 3 pyramidal cells in human temporal cortex, 2.62µm$^{-1}$ (n=1 human, 73 dendrites), comes from the work of (53). This type of data is very difficult to come by with and in Chapter 2 we used measurements from only a single human subject. We did not use the measurements of the spine density reported in (33) for 2 human patients, as in that comparative study no correction was made for hidden spines, i.e. spines located directly above or below the dendrite. To have an estimate of the extent of inter-subject variability we applied the variability observed in the monkey spine density to the human data. This estimate is justified in part by the fact that a similar coefficient of variation had been observed in human data from (33). This resulted in a spine density of 2.62 ±0.34(0.34)µm$^{-1}$. The volume density of asymmetric synapses in layer 3 of human temporal cortex was calculated by multiplying the density of all synapses with the reported fraction of asymmetric synapses resulting in 1.07±0.31(0.18)µm$^{-3}$ (n=3 humans, 60 blocks of tissue) (54). As a result, the dendritic length density calculated as the product of the inter-spine interval and the density of asymmetric synapses was 0.42±(0.09)µm$^{-2}$. 
Appendix B

The distribution of potential synapse numbers

In Chapter 3, a potential synapse was defined as an opposition between axonal and dendritic branches of two neurons, where the distance between the branch centerlines is less than \( s = 2 \) \( \mu \)m. Assumptions and approximations related to the concept of the potential synapse, as well as different definitions of potential connectivity were previously described (2, 77). According to the definition used in Chapter 3, at most one potential synapse can exist between a pair of axonal and dendritic branches, independent of their lengths and relative layout. A branch here is defined as a neuron process connecting a cell body or a bifurcation point to a successive bifurcation or an end point. The above constraint is necessary in order to disambiguate situations where potentially connected axonal and dendritic branches ran in parallel over long distances. Because this is a rare occurrence for cortical excitatory neurons, the numbers of potential synapses calculated according to this definition are in good agreement with other estimates [see e.g. Figure 3C in (8)]. The value of \( s = 2 \) \( \mu \)m was chosen for the definition of a potential synapse, as it corresponds
to the average distance between centerlines of synaptically connected axonal and dendritic branches (1), i.e. this value is roughly equal to the sum of the average dendritic spine length, the average dendritic branch radius, and the average axonal bouton radius. Though the average number of potential synapses between cells scales linearly with $s$ (2, 62), and, hence, the shape of the distribution of potential synapse numbers is also dependent on $s$ (Figure B1), the main results of Chapter 3 are not very sensitive to particular values of this parameter. This point is illustrated in Figure B2, where all the results were replicated for $s = 1.5 \, \mu m$ and $s = 2.5 \, \mu m$.

The numbers of potential synapses were calculated for all pairs of neurons by using a computer search algorithm (19) which finds and counts close oppositions between axonal and dendritic branches in the arbors’ overlap region. As the number of potential synapses is very sensitive to exact placement of individual branches within the arbors’ overlap region, 1,000 different realizations of potential connectivity were considered for each neuron pair. This was done by randomly shifting the arbors around their original positions in the slice template of rat barrel cortex. Random shifts of the neurons were performed in the plane of the slice and were confined to 20 μm square regions surrounding the neurons’ original positions. Because axonal arbors of the neurons were significantly truncated due to tissue slicing, the neurons were not moved in the direction perpendicular to the slice surface. Likewise, the neurons were not shifted over large distances in the plane of the slice, as the shapes of axonal and dendritic arbors depend on neuron’s positions with respect to the laminar and barrel boundaries.

To emulate the conditions of the Sakmann lab experiments (67-69), only neuron pairs with separations of no more than 50 μm in the direction along the cortical surface (in the plane of the slice) were considered. This restriction resulted in 70 neuron pairs for L4→L2/3 projection, 104 pairs for L5→L5 projection, and 116 pairs for L4→L4 projection. For each projection, the
distribution of the numbers of potential synapses was obtained by pooling the 1,000 potential synapse numbers generated for each neuron pair. As a result, the distributions of the numbers of potential synapses for $L4 \rightarrow L2/3$, $L5 \rightarrow L5$, and $L4 \rightarrow L4$ projections contained 70,000, 104,000, and 116,000 elements (Figure 3.1B, C, D). The results of Chapter 3 are based on the morphologies of neurons reconstructed from barrel and septum-related columns (8). Because dendritic and local axonal morphologies of neurons in these regions are similar (8), the distributions of potential synapse numbers for the considered projections were also similar in the barrel and septum-related columns. We separately analyzed barrel and septum related projections, and because no significant differences in the results were detected, these projections were pooled.

**P-values and confidence regions**

Inspection of the results shown in Figure B1 makes it evident that differences between the generated and the experimental distributions of actual synapse numbers can be formulated in terms of their mean values and widths. Some generated distributions of actual synapse numbers can match the mean values but are much wider than the experimental distributions. Conversely, the generated distributions which match the experimentally observed variances, significantly underestimate the average observed synapse numbers. To capture these trends in both independent and cooperative synapse formation models, we combined two significance tests: comparison of the mean values and comparison of the Fano factors (ratio between variance and mean). Alternative statistical tests, based on the coefficient of variation and the mean, the standard deviation and the mean (results not shown), and best fitting (see below), were performed and did not lead to significant differences in any of the results.
To examine whether the mean values of the generated and the experimental distributions of actual synapse numbers are statistically different, we calculated, for each set of the model parameters, p-values based on the two-tailed Student’s t-test for samples with unequal variances. Because there is no standard statistical test for comparing the Fano factors of two distributions, we designed a test based on a bootstrap procedure [see section 15.6 in (74)]. For each considered projection, we generated 10,000 new bootstrap sets of actual synapse numbers. Each such set was of the same length as the corresponding experimental set of synaptically coupled neurons. These sets were created by randomly sampling with replacement numbers of synapses from the generated distributions of synapse numbers. The p-value for this test was calculated as the fraction of the bootstrap sets which yielded Fano factors less or equal to the Fano factor for the corresponding experimental distribution.

Next, the results of the above two tests were combined by using Bonferroni multiple hypotheses testing procedure. Here, we used the sequential rejective Bonferroni procedure (107), which makes no assumptions about the statistical dependence of the tested hypotheses. According to this procedure, if the minimum of the two p-values is greater than 2.5%, then neither of the hypotheses can be rejected at 5% significance level. In this case, the comparison tests for the model and the experimental distributions of actual synapse numbers do not detect differences between the distributions at 5% significance level.

For the model of independent synapse formation, the above procedure was carried out for different values of the parameter $p$, which was varied in steps of 0.025 in the $[0,1]$ range. The maximum combined p-value was reported in the Results section for each projection. For the cooperative synapse formation model, statistical tests were repeated for all values of the model parameters $p$, $N_s c$, and $\Delta$ on the following 3D grid: $p \in [0,1]$ step size of 0.025, $N_s c \in [0,10]$ step
size of 0.1, \( \Delta \in [0,4] \) step size of 0.1. Significance tests at 5% level were performed for every point on the grid, which determined the 95% confidence region in the 3D space of the model parameters (Figure 3.4). Upper and lower 95% confidence bounds for \( p, N^c_s \), and \( \Delta \) (Table 3.1) were obtained by projecting the 3D 95% confidence region onto the corresponding axis.

**Alternative statistical analysis of the cooperative synapse formation model**

*Figure B1.* Dependence of the distribution of the numbers of potential synapses on the value of the parameter \( s \).  
**A.** Parameter \( s \) dependence of the distribution of potential synapse numbers for neurons forming local L4→L2/3 projection.  
**B.** Same for pairs of nearby neurons in L5.  
**C.** Same for pairs of nearby neurons in L4.  
**D.** Comparison of the distributions from A, B, and C in 1-5 \( \mu \)m range of parameter \( s \). Individual distributions, rescaled along the x- and y-axes to have the average number of potential synapses and the area under the curve equal 1, have a universal shape.
To further validate the idea of the cooperative synapse formation (Chapter 3) we performed an additional statistical test, alternative to the one described above. First, for each considered projection, we randomly selected 10,000 bootstrap sets of reconstructed neuron pairs. Each such set was of the same length as the experimental set of synaptically coupled neurons. The bootstrap sets were created by randomly sampling neuron pairs with replacement from all nearby reconstructed neurons. Second, we calculated the numbers of potential synapses for all the selected neuron pairs leading to 10,000 bootstrap sets of potential synapse numbers. Third, every bootstrap set of potential synapse numbers was used as an input to Equations 3.8, to generate the distributions of actual synapse numbers for different values of the model parameters, $p$, $N^c_s$, and $\Delta$. All combinations of these parameters on the three-dimensional grid were considered. The parameter values, which resulted in the best fits of the experimental distribution of actual synapse numbers, were selected. This procedure resulted in 10,000 sets of the best fit parameters, $p$, $N^c_s$, and $\Delta$, each corresponding to a different realization of the bootstrap sampling.

<table>
<thead>
<tr>
<th>Projection</th>
<th>$p$</th>
<th>$N^c_s$</th>
<th>$\Delta$</th>
<th>$P_{con}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>L4→L2/3</td>
<td>0.055 ± 0.11 (0.025-0.10)</td>
<td>4.3 ± 0.17 (4.1-4.4)</td>
<td>0.33 ± 0.14 (0.10-0.50)</td>
<td>0.0021 ± 0.022</td>
</tr>
<tr>
<td>L5→L5</td>
<td>0.32 ± 0.11 (0.15-0.58)</td>
<td>4.2 ± 0.13 (4.1-4.5)</td>
<td>0.42 ± 0.24 (0.20-0.90)</td>
<td>0.11 ± 0.060</td>
</tr>
<tr>
<td>L4→L4</td>
<td>0.40 ± 0.15 (0.18-0.75)</td>
<td>2.2 ± 0.41 (2.0-2.7)</td>
<td>0.51 ± 0.29 (0.10-1.4)</td>
<td>0.30 ± 0.13</td>
</tr>
</tbody>
</table>

**Table B1.** Results of the alternative statistical analysis of the cooperative synapse formation model. Numerical values are shown in the mean ± SD (95% confidence interval) format.

The mean, standard deviation, and confidence intervals of these parameters are provided in the Table B1. These parameter values were used to calculate the connection probabilities, $P_{con}$. 

93
Figure B2. Robustness of the results with respect to different values of the parameter $s$. A. Dependence of the probability $p$ on $s$. Average values for different projections are shown with bars of different colors. Error bars indicate the standard deviations. B. Same for the critical number of synapses, $N_s^c$. C. Same for the cooperative transition width, $\Delta$. D. Same for the probability of connection, $P_{con}$.

All the results shown in the Table B1 are not significantly different from those reported in Table 3.1, confirming our conclusions. However, a caveat of the above described statistical analysis is in the fact that by only considering the best fit parameter values we may significantly underestimate the confidence intervals. The fits to the data do not need to be the best, they have to be good enough, i.e. result in the distributions that are not significantly different from the experimental distributions. For this reason, the majority of confidence intervals for $p$, $N_s^c$, and $\Delta$
in the Table B1 are much narrower than the corresponding values from Table 3.1. Therefore, the statistical test described in the Chapter 3 provides a better estimate of the variability in the model parameter values.

**Experimental Procedures**

The results of Chapter 3 are based on the combined analysis of datasets of neuron morphology and synaptic connectivity in rat barrel cortex. The dataset of 3D neuron morphology used in this work was originally reconstructed in the context of the Shepherd et al. study (8). This dataset was obtained from http://NeuroMorpho.Org (70). Because the experimental procedures related to this dataset were previously described (8), in the following we only provide a brief summary of facts that are relevant to our work. Young adult Sprague-Dawley rats (26–36 days postnatal) were used in the experiment. 300 µm thick slices of barrel cortex were cut perpendicular to barrel rows. Excitatory cells located 50 to 100 µm (mean of 78 µm) below slice surface were labeled with biocytin and reconstructed in 3D with Neurolucida system (MicroBrightField Inc, Colchester, VT). In total, 13 pyramidal cells from L2/3, 13 spiny stellate and 13 pyramidal cells from L4, and 23 pyramidal cells from L5A were reconstructed from barrel (33 cells) and septum (29 cells) related columns. All reconstructions were corrected for tissue shrinkage and conformed to the template of a barrel cortical slice. This was an essential step, because for the analysis of potential connectivity it is typically necessary to put together neurons reconstructed from different brains.

The distributions of synapse numbers, for pairs of synaptically coupled neurons forming L4→L2/3, L5→L5, and L4→L4 projections, were measured in a series of studies performed in
the Sakmann group (67-69). In these experiments, 300-400 µm thick cortical slices were obtained from juvenile (12-23 days postnatal) Wistar rats. The L4→L2/3 (13 neuron pairs in the barrel column) (68) and L4→L4 (11 barrel related neuron pairs) (67) projections were studied in rat barrel cortex and the L5→L5 (19 thick tufted pyramidal cell pairs) (69) projection in rat somatosensory cortex. Dual whole-cell recordings were used to identify synaptically coupled neurons. Such neurons were typically separated by no more than 50 µm in the direction along the cortical surface and were located up to 120 µm deep in the slice (80 µm on average). The neurons were labeled with biocytin and their putative synaptic contacts were identified as close oppositions of a bouton and dendrite in the same focal plane. To confirm that the putative synapses, detected with light microscopy, were indeed synaptic contacts, a subset of them was analyzed with serial electron microscopy. Since more than 80% of putative synapses were confirmed at the electron microscopy level (68, 69, 108), light microscopic predictions were deemed accurate. However, very small synapses, or synapses hidden by thick dendrites, may have been omitted, which would result in an underestimation of synapse numbers. We do not expect small underestimations to affect the conclusions of Chapter 3, since larger synapse numbers would lead to even greater differences between the distributions of actual and potential synapse numbers.
Appendix C

Dataset of cat V1 neurons used in Chapter 4

**Figure C1.** Excitatory cells. Laminar distribution and morphology of 20 pyramidal and 4 spiny stellate cells from the cat V1 used for the analysis of potential connectivity. All neurons have been reconstructed in 3D and conformed to the common template of the cat V1. Axons are shown in black, dendrites are in red, positions of soma centers are marked with black dots. Axonal branches which fall beyond 1000μm (in radius) cylinders surrounding the somata have been truncated. (Adopted from Stepanyants et al. (19))
Figure C2. Inhibitory basket cells. Laminar distribution and morphology of 17 inhibitory basket cells from the cat V1 used for the analysis of potential connectivity. All neurons have been reconstructed in 3D and conformed to the common template of the cat V1. Axons are shown in black, dendrites are in red, positions of soma centers are marked with black dots. Axonal branches which fall beyond 1000μm (in radius) cylinders surrounding the somata have been truncated. (Adopted from Stepanyants et al. (19))
Dataset of neurons from rat barrel cortex dataset used in Chapters 3 and 4

Figure C3 Excitatory neurons from rat barrel cortex

A. Examples of reconstructed neurons from different layers projecting axons to L2/3. B. Dendrites of reconstructed L2/3 pyramidal neurons ($n=13$), aligned by soma position. Bar, 0.5 mm. C. Axons of L4 neurons ($n=26$). D. Axons of L5A neurons ($n=23$). Ascending branches reached peak densities in L2 (arrow). (Adopted from Shepherd et.al. (8))
Appendix D

Here we formulate and solve the perceptron like model of Chapter 5 which takes into account a number of biologically inspired constraints. Related models, including only some of the constraints considered here were previously described in a number of studies [see e.g. (22, 24, 25, 28-30, 109, 110)].

Model of a robust sign-constrained perceptron with a fixed threshold

In this model we assumed that:

1. The weights of individual perceptron synapses, \( \tilde{J}_j \) \((j = 1, \ldots, N)\), are sign-constrained, of which \( N_i \) are inhibitory and \( N_e \) are excitatory, \( N_i + N_e = N \).

2. For every association, \( \mu \) \((\mu = 1, \ldots, m)\), synaptic inputs, \( \zeta_j^\mu \), and the associated outputs, \( o^\mu \), of the perceptron are binary \( \{0,1\} \), and are randomly drawn from the distribution in which 1 appears with probability \( p \) and 0 with probability \( 1 - p \).
3. The firing threshold of the perceptron, $h$, is fixed and does not change during learning.

Perceptron model with the above constraints can be summarized mathematically as:

$$
\theta \left( \sum_{j=1}^{N} \tilde{J}_j \zeta_j^\mu - h \right) = \sigma^\mu, \quad \mu = 1, \ldots, m
$$

$$
\tilde{J}_j g_j > 0, \quad j = 1, \ldots, N
$$

$$
\zeta, \sigma \in \mathcal{P} = \begin{cases} 
0, & 1 - p \\
1, & p 
\end{cases}
$$

(D.1)

In this expression, $\theta$ is the step function, $m$ denotes the number of associations, and the parameter $g_j$ specifies if input $j$ is excitatory ($g_j = 1$) or inhibitory ($g_j = -1$).

We also assume that:

4. The associations, $\{\zeta_j^\mu, \sigma^\mu\}$, are learned robustly

$$
\sigma^\mu \left( \sum_{j=1}^{N} \tilde{J}_j \zeta_j^\mu - h \right) > \tilde{\kappa}, \quad \mu = 1, \ldots, m
$$

$$
\sigma^\mu = 2\sigma^\mu - 1
$$

$$
\tilde{J}_j g_j > 0, \quad j = 1, \ldots, N
$$

$$
\zeta \in \mathcal{P}_\zeta = \begin{cases} 
0, & 1 - p \\
1, & p 
\end{cases}
$$

$$
\sigma \in \mathcal{P}_\sigma = \begin{cases} 
-1, & 1 - p \\
1, & p 
\end{cases}
$$

(D.2)

and

5. The perceptron operates in the vicinity of the firing threshold. This means that the average synaptic input to the perceptron is of the order of $h$, and the fluctuations in the average input are of the order of $\tilde{\kappa}$:
Assuming that the firing probability, \( p \), does not scale with \( N \), we rewrite Equations D.2 in a more convenient notation, where the dependence on \( N \) is shown explicitly:

\[
p = \text{const} \\
\tilde{J} = \frac{J_h}{N} \\
\tilde{\kappa} = \frac{k_h}{\sqrt{N}}
\]  

(D.4)

The result is a model governed by the number of inputs, \( N \), and three intensive parameters, \( N_i/N, p, \) and \( \kappa \):

\[
\sigma^\mu \left( \frac{1}{N} \sum_{j=1}^{N} J_j \zeta_j^\mu - 1 \right) > \frac{\kappa}{\sqrt{N}} , \quad \mu = 1, \ldots, m \\
\sigma^\mu = 2^{\mu^2} - 1 \\
J_j g_j > 0 , \quad j = 1, \ldots, N \\
\zeta \in P_\zeta = \begin{cases} 
0, & 1 - p \\
1, & p 
\end{cases} \\
\sigma \in P_\sigma = \begin{cases} 
-1, & 1 - p \\
1, & p 
\end{cases}
\]  

(D.5)

Replica theoretical formulation of the problem

Following the previously developed methodology (24, 25, 28-30, 110), we calculated the volume of the solution space, \( \Omega(\zeta^\mu, \sigma^\mu) \), given a set of associations, \( \{\zeta^\mu, \sigma^\mu\} \),
\[ \Omega(\zeta^\mu_j, \sigma^\mu) = \int \prod_{j=1}^N dJ_j \prod_{\mu=1}^m \theta \left( \sigma^\mu \left( \frac{1}{N} \sum_{j=1}^N J_j \zeta^\mu_j - 1 \right) - \frac{\kappa}{\sqrt{N}} \right) \prod_{j=1}^N \theta(J_j g_j) . \tag{D.6} \]

Here, we assumed that weights of individual synapses are independent of each other and the magnitudes of the weights come from a uniform prior probability distribution.

The typical volume of solutions, \( \Omega_{\text{typical}} \), is defined through the averaging of the logarithm of \( \Omega(\zeta^\mu_j, \sigma^\mu) \) over the associations,

\[ \ln(\Omega_{\text{typical}}) = \langle \ln(\Omega(\zeta^\mu_j, \sigma^\mu)) \rangle_{\zeta^\mu_j, \sigma^\mu} . \tag{D.7} \]

Following the replica method, the average of the logarithm is calculated in the limit

\[ \ln(\Omega_{\text{typical}}) = \lim_{n \to 0} \frac{\langle \Omega(\zeta^\mu_j, \sigma^\mu)^n \rangle_{\zeta^\mu_j, \sigma^\mu} - 1}{n} . \tag{D.8} \]

Furthermore, \( \langle \Omega(\zeta^\mu_j, \sigma^\mu)^n \rangle_{\zeta^\mu_j, \sigma^\mu} \) is calculated by introducing \( n \) replica systems, quenched with respect to the set of associations \( \{ \zeta^\mu_j, \sigma^\mu \} \),

\[ \langle \Omega(\zeta^\mu_j, \sigma^\mu)^n \rangle_{\zeta^\mu_j, \sigma^\mu} = \left( \int \prod_{a,j=1}^{n,N} dJ_j^a \prod_{\mu,a=1}^{m,n} \theta(\sigma^\mu \left( \frac{1}{N} \sum_{j=1}^N J_j^a \zeta^\mu_j - 1 \right) - \frac{\kappa}{\sqrt{N}} \prod_{j,a=1}^{N,n} \theta(J_j^a g_j) \right)_{\zeta^\mu_j, \sigma^\mu} \tag{D.9} \]

Below, we outline the main steps of calculation of Equation D.9.

Input and output associations, \( \zeta^\mu_j \) and \( \sigma^\mu \), are decoupled through the introduction of a new variable, \( \frac{1}{N} \sum_{j=1}^N J_j^a \zeta^\mu_j = 1 + \lambda^a \mu \),

103
\[ \langle \Omega(\xi_j^\mu, \sigma_j^\mu) \rangle_{\xi_j^\mu, \sigma_j^\mu, \sigma_j^\mu} = \int \prod_{\mu, a, j=1}^{m, n} dJ_j^\mu \prod_{\mu, a, j=1}^{m, n} d\lambda_j^a, \mu \left\langle \prod_{\mu, a=1}^{m, n} \theta \left( \sigma_j^\mu \lambda_j^a, -\frac{\kappa}{\sqrt{N}} \right) \right\rangle_{\sigma_j^\mu} \times \]
\[ \left\langle \prod_{\mu, a=1}^{m, n} \delta \left( 1 + \lambda_j^a, \mu - \frac{1}{N} \sum_{j=1}^{N} J_j^a \xi_j^\mu \right) \right\rangle_{\xi_j^\mu, \sigma_j^\mu, \sigma_j^\mu} \prod_{\mu, j=1}^{N, a} \theta (J_j^\sigma g_j) \]

The first averaging in Equation D.10 is performed after replacing the step function with its integral representation:

\[ \left\langle \prod_{\mu, a=1}^{m, n} \theta \left( \sigma_j^\mu \lambda_j^a, \mu - \frac{\kappa}{\sqrt{N}} \right) \right\rangle_{\sigma_j^\mu} = \prod_{\mu=1}^{m} \left( \prod_{a=1}^{n} \int_0^\infty du, a, \mu \delta \left( \sigma_j^\mu \lambda_j^a, \mu - \frac{\kappa}{\sqrt{N}} - u, a, \mu \right) \right)_{\sigma_j^\mu} = \]
\[ \prod_{\mu=1}^{m} \left( \prod_{a=1}^{n} \int_0^\infty du, a, \mu \right) \int_0^\infty e^{\frac{i}{\sigma_j^\mu} \int_0^\infty \frac{d\lambda_j^a, \mu}{2\pi} \left( \lambda_j^a, \mu - \frac{\kappa}{\sqrt{N}} - u, a, \mu \right)} \right)_{\sigma_j^\mu} = \]
\[ \int \prod_{\mu, a=1}^{m, n} d' u, a, \mu \int \prod_{\mu=1}^{m} \exp \left( \frac{i}{\sigma_j^\mu} \int_0^\infty \frac{d\lambda_j^a, \mu}{2\pi} \right) \prod_{\mu=1}^{m} \left( pe^{\frac{1}{\sigma_j^\mu} \int_0^\infty \frac{d\lambda_j^a, \mu}{2\pi}} + (1-p)e^{-\frac{1}{\sigma_j^\mu} \int_0^\infty \frac{d\lambda_j^a, \mu}{2\pi}} \right) \]

In this expression, symbol \(d'\) denotes integration from 0 to \(\infty\) and symbol \(d\) from \(-\infty\) to \(\infty\).

The second averaging in Equation D.10 can be performed in a similar manner, leading to:

\[ \left\langle \prod_{\mu, a=1}^{m, n} \delta \left( 1 + \lambda_j^a, \mu - \frac{1}{N} \sum_{j=1}^{N} J_j^a \xi_j^\mu \right) \right\rangle_{\xi_j^\mu} = \left\langle \prod_{\mu, j=1}^{m, N} \int \frac{d\lambda_j^a, \mu}{2\pi} e^{\frac{i}{\sigma_j^\mu} \int_0^\infty \frac{d\lambda_j^a, \mu}{2\pi} \left( 1 + \lambda_j^a, \mu - \frac{1}{N} \sum_{j=1}^{N} J_j^a \xi_j^\mu \right)} \right\rangle_{\xi_j^\mu} = \]
\[ \prod_{\mu=1}^{m} \left( \prod_{a=1}^{n} \int_0^\infty e^{\frac{i}{\sigma_j^\mu} \int_0^\infty \frac{d\lambda_j^a, \mu}{2\pi}} \right) \left( e^{\frac{1}{\sigma_j^\mu} \int_0^\infty \frac{d\lambda_j^a, \mu}{2\pi}} \prod_{\mu=1}^{m} \left( 1 - p + pe^{-\frac{1}{\sigma_j^\mu} \int_0^\infty \frac{d\lambda_j^a, \mu}{2\pi}} \right) \right) \]

In calculating the above two averages we accounted for the fact that the associations \(\{ \xi_j^\mu, \sigma_j^\mu \}\) are the same for all the replicas, and, before the averaging, we moved the products over the replicas into the exponentials as sums.
Plugging Equations D.11 and D.12 into D.10 we arrive at:

\[ \left( \Omega(\xi^\mu, \sigma^{\mu}) \right)^n_{\xi^\mu, \sigma^{\mu}} = \prod_{j=1}^{n} dJ_j \prod_{\alpha, \beta=1}^{m} \frac{d\lambda^{\alpha, \beta} d\hat{\lambda}^{\alpha, \beta}}{2\pi} \prod_{\mu, \nu=1}^{n} \frac{d\alpha^{\mu} d\hat{\alpha}^{\mu}}{2\pi} \times \]

\[ i \sum_{\mu, \nu=1}^{m} \left( \chi^{\mu}(1+\lambda^{\mu}) - \hat{\chi}^{\mu} \left( \frac{\nu^{\mu} + \kappa}{\sqrt{N}} \right) \right) \prod_{\mu=1}^{m} \left( pe^{\frac{1}{2} \sum_{\nu=1}^{n} \hat{\alpha}^{\mu} \chi^{\mu}} + (1-p)e^{-\frac{1}{2} \sum_{\nu=1}^{n} \hat{\alpha}^{\mu} \chi^{\mu}} \right) \times \] (D.13)

\[ \prod_{\mu, j=1}^{m,N} \left( 1 - p + pe^{\frac{1}{N} \sum_{\nu=1}^{n} \hat{\alpha}^{\mu} J_j^{\nu}} \right) \prod_{j, a=1}^{N,N} \theta(J_j^a g_j) \]

We note that the term \( \frac{1}{N} \sum_{a=1}^{n} \hat{\lambda}^{\alpha, \mu} J_j^{a} \) is small in the \( n \to 0 \) limit, and expand the exponent containing this term up to the second order,

\[ \left( \Omega(\xi^\mu, \sigma^{\mu}) \right)^n_{\xi^\mu, \sigma^{\mu}} = \prod_{\mu, a=1}^{m} \frac{d\lambda^{\alpha, \beta} d\hat{\lambda}^{\alpha, \beta}}{2\pi} \prod_{\mu, a=1}^{m} \frac{d\alpha^{\mu} d\hat{\alpha}^{\mu}}{2\pi} \times \]

\[ i \sum_{\mu, a=1}^{m} \left( \chi^{\mu}(1+\lambda^{\mu}) - \hat{\chi}^{\mu} \left( \frac{a^{\mu} + \kappa}{\sqrt{N}} \right) \right) \prod_{\mu=1}^{m} \left( pe^{\frac{1}{2} \sum_{a=1}^{n} \hat{\alpha}^{\mu} \chi^{\mu}} + (1-p)e^{-\frac{1}{2} \sum_{a=1}^{n} \hat{\alpha}^{\mu} \chi^{\mu}} \right) \times \] (D.14)

\[ \prod_{j=1}^{n} \prod_{a=1}^{m} dJ_j^{a} \prod_{\mu=1}^{m} e^{-\frac{1}{N} \sum_{\nu=1}^{n} \hat{\alpha}^{\mu} J_j^{\nu}} \left( \frac{p^{a} - p^{\mu} \left( \frac{1}{N} \sum_{\nu=1}^{n} \hat{\alpha}^{\mu} J_j^{\nu} \right)^2}{2} \right) \prod_{a=1}^{n} \theta(J_j^a g_j) \]

At this point we introduced two order parameters,

\[ \frac{1}{N} \sum_{i=1}^{N} J_i^a = \frac{1}{p} + \frac{S^a}{\sqrt{N}}, \quad \frac{1}{N} \sum_{i=1}^{N} J_i^a J_i^b = q^{a,b} \] (D.15)

The first of these two parameters insures that the perceptron operates near the firing threshold, as described by Equations D.3. The second parameter characterizes the correlations between solutions belonging to different replicas. In particular, at the critical number of associations, \( m_c \), the volume of the solution space shrinks to zero and \( q^{a,a} = q^{a,b} \).

Introduction of the above order parameters makes it possible to decouple the integrals in Equation D.14 into of separate products over indices \( i \) and \( \mu \):
After the following renormalizations of the internal integration variables,

\[ \hat{q} \rightarrow \hat{q} ; \quad \hat{\lambda} \rightarrow \hat{\lambda} ; \quad \sqrt{N} \lambda \rightarrow \lambda ; \quad \hat{u} \rightarrow \hat{u} ; \quad \sqrt{N} u \rightarrow u , \]

(D.17)

Equation D.16 can be rewritten as:

\[
\left< \Omega \left( \xi^\mu, \sigma^\mu \right) \right>_{\xi^\mu, \sigma^\mu} = \prod_{a=1}^{n} \frac{ds^a d\hat{s}^a}{2\pi / \sqrt{N}} \prod_{a,b=1}^{n} \frac{d\theta^{a,b} d\hat{\theta}^{a,b}}{2\pi} e^{i \hat{\sum}_{a=1}^{n} \left( \frac{1}{N} \hat{\sum}_{a=1}^{n} \hat{\theta}^{a,b} q^{a,b} + a G_N \left( |x|, |q^{a,b}| \right) + G_3 \left( |x|, |\hat{q}^{a,b}| \right) \right)} \times (D.16)
\]

After these renormalizations, the integral in Equation D.18 can be calculated by using the steepest descent method combined with the replica symmetric saddle point assumption, \( s^a = s \), \( \hat{s}^a = \hat{s} \), \( q^{a,b} = q_0 \), \( q^{axb} = q \), \( \hat{q}^{a,b} = \hat{q}_0 \), and \( \hat{q}^{axb} = \hat{q} \). As a result, the integral in Equations D.18 reduces to:

\[
\begin{align*}
G_E \left( \{ s^a \}, \{ q^{a,b} \} \right) &= \ln \left[ \prod_{a=1}^{n} \frac{d\theta^{a,b}}{2\pi} e^{i \hat{\sum}_{a=1}^{n} \hat{\theta}^{a,b} q^{a,b}} \left( p e^{i \hat{\sum}_{a=1}^{n} \hat{\theta}^{a,b}} + (1 - p) e^{-i \hat{\sum}_{a=1}^{n} \hat{\theta}^{a,b}} \right) \right] \\
G_S \left( \{ s^a \}, \{ \hat{q}^{a,b} \} \right) &= \frac{N}{N} \ln \left[ \prod_{a=1}^{n} d\hat{\theta}^{a,b} e^{i \hat{\sum}_{a=1}^{n} \hat{\theta}^{a,b} q^{a,b}} \right] + \frac{N}{N} \ln \left[ \prod_{a=1}^{n} d\hat{\theta}^{a,b} e^{-i \hat{\sum}_{a=1}^{n} \hat{\theta}^{a,b} q^{a,b}} \right]
\end{align*}
\]

In this expression \( \alpha = m / N \) is referred to as the capacity of the perceptron.

Ignoring the small terms in \( N \), the integral in Equation D.18 can be calculated by using the steepest descent method combined with the replica symmetric saddle point assumption, \( s^a = s \), \( \hat{s}^a = \hat{s} \), \( q^{a,b} = q_0 \), \( q^{axb} = q \), \( \hat{q}^{a,b} = \hat{q}_0 \), and \( \hat{q}^{axb} = \hat{q} \). As a result, the integral in Equations D.18 reduces to:
The saddle point, \((s, q_0, \hat{s}, \hat{q}_0, \hat{q})\), satisfies the following system of equations:

\[
\begin{align*}
\frac{\partial G_E(s, q_0, q)}{\partial \hat{s}} = 0; & \quad \text{in} \hat{q}_0 + \alpha \frac{\partial G_E(s, q_0, q)}{\partial \hat{q}_0} = 0; \quad -\text{in} \hat{q} + \alpha \frac{\partial G_E(s, q_0, q)}{\partial \hat{q}} = 0 \\
\frac{\text{in}}{p} + \frac{\partial G_E(s, q_0, q)}{\partial \hat{s}} = 0; & \quad \text{in} q_0 + \frac{\partial G_E(s, q_0, q)}{\partial \hat{q}_0} = 0; \quad -\text{in} q + \frac{\partial G_E(s, q_0, q)}{\partial \hat{q}} = 0
\end{align*}
\] (D.20)

To simplify Equations (D.19) we employed the Hubbard-Stratonovich transformation in the form

\[
\sum_{s, i} e^{s_{ij}} = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} dx e^{-x^2 + 2s_{ij}x_i x_j}.
\] (D.21)

and take the \(n \to 0\) limit:

\[
\begin{align*}
G_E(s, q_0, q)/n &= p \int \frac{dx}{\sqrt{\pi}} e^{-x^2} \ln \left( \int d\hat{u} d\hat{v} e^{(q_0 - q)^2 (\hat{u}^2 - p \hat{v}^2)/2 + \sqrt{\hat{u}^2 - p \hat{v}^2}} \right) + \\
(1 - p) \int \frac{dx}{\sqrt{\pi}} e^{-x^2} \ln \left( \int d\hat{u} d\hat{v} e^{(q_0 - q)^2 (\hat{u}^2 - p \hat{v}^2)/2 + \sqrt{\hat{u}^2 - p \hat{v}^2}} \right)
\end{align*}
\] (D.22)

\[
G_E(s, q_0, q)/n = N \int \frac{dx}{\sqrt{\pi}} e^{-x^2} \ln \left( \int dJ e^{i(q_0 - q)J_x + (\bar{u} + 2x_0^2)^{1/2}} \right) + \\
N_e \int \frac{dx}{\sqrt{\pi}} e^{-x^2} \ln \left( \int dJ e^{i(q_0 - q)J_x + (\bar{u} - 2x_0^2)^{1/2}} \right)
\]

After integrating over \(\hat{u}, \hat{v}\), and \(J\), and eliminating complex variables with \(r = i\hat{r}\), \(t = -i\hat{t}\) substitutions we arrived at:
\[
\frac{\partial G_E(s, q_0, q)}{\partial s} = 0; \quad -t_0 + \alpha \frac{\partial G_E(s, q_0, q)}{\partial q_0} = 0; \quad t + \alpha \frac{\partial G_E(s, q_0, q)}{\partial q} = 0
\]
\[
\frac{1}{p} + \frac{\partial G_S(r, t_0, t)}{\partial r} = 0; \quad -q_0 + \frac{\partial G_S(r, t_0, t)}{\partial t_0} = 0; \quad q + \frac{\partial G_S(r, t_0, t)}{\partial t} = 0
\]
\[
G_E(s, q_0, q) / n = p \int \frac{dx}{\sqrt{\pi}} e^{-x^2} \ln \left( H \left( -\frac{x\sqrt{(p-p^2)2q + ps - \kappa}}{(q_0-q)(p-p^2)} \right) \right) +
\]
\[
(1-p) \int \frac{dx}{\sqrt{\pi}} e^{-x^2} \ln \left( H \left( -\frac{x\sqrt{(p-p^2)2q - ps - \kappa}}{(q_0-q)(p-p^2)} \right) \right)
\]
\[
G_S(r, t_0, t) / n = \frac{N_r}{N} \int \frac{dx}{\sqrt{\pi}} e^{-x^2} \ln \left( \frac{\sqrt{\pi}}{\sqrt{1-t_0}} e^{\frac{(r_2-r)^2}{2(t_0-t)}} \right) H \left( -\frac{r + 2x\sqrt{t}}{\sqrt{2(t-t_0)}} \right) +
\]
\[
\frac{N_s}{N} \int \frac{dx}{\sqrt{\pi}} e^{-x^2} \ln \left( \frac{\sqrt{\pi}}{\sqrt{1-t_0}} e^{\frac{(r_2-r)^2}{2(t_0-t)}} \right) H \left( -\frac{r - 2x\sqrt{t}}{\sqrt{2(t-t_0)}} \right)
\]

(D.23)

Function \( H \) in this expression is defined as:

\[
H(x) = \int_x^\infty \frac{dt}{\sqrt{2\pi}} e^{-t^2/2} = \frac{1}{2\sqrt{2}} (1 - \text{erf}(x))
\]

\[
\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x dt e^{-t^2}
\]

Critical capacity

As mentioned above, at the critical number of associations \((q_0 - q)\) tends to zero. In this limit \((t_0 - t) \to +\infty\) [see Equations D.26 below], and we expand \(G_E\) and \(G_S\) asymptotically, leaving only the diverging terms:
Using new notation \( \frac{p}{2q(1-p)} = s \), \( \frac{r}{2\sqrt{t}} = z \), and \( \frac{\kappa}{\sqrt{(p-p^3)2q}} = Q \), we rewrote the replica symmetric saddle point Equations D.23 in the explicit form:
\[ p \int_0^\infty \frac{dx}{\sqrt{\pi}} e^{-(x+y-Q)^2} x^2 (x+y-Q) - (1-p) \int_0^\infty \frac{dx}{\sqrt{\pi}} e^{-(x-y-Q)^2} x^2 (x-y-Q) = 0 \]

\[ \frac{q}{(q_0 - q)^2} \left( p \int_0^\infty \frac{dx}{\sqrt{\pi}} e^{-(x+y-Q)^2} x^2 + (1-p) \int_0^\infty \frac{dx}{\sqrt{\pi}} e^{-(x-y-Q)^2} x^2 \right) + \frac{1}{2} \frac{1}{q_0 - q} \left( p \int_0^\infty \frac{dx}{\sqrt{\pi}} e^{-(x+y-Q)^2} x^2 + (1-p) \int_0^\infty \frac{dx}{\sqrt{\pi}} e^{-(x-y-Q)^2} x^2 \right) = \frac{t_0}{\alpha} \]

\[ \frac{q_0}{(q_0 - q)^2} \left( p \int_0^\infty \frac{dx}{\sqrt{\pi}} e^{-(x+y-Q)^2} x^2 + (1-p) \int_0^\infty \frac{dx}{\sqrt{\pi}} e^{-(x-y-Q)^2} x^2 \right) + \frac{1}{2} \frac{1}{q_0 - q} \left( p \int_0^\infty \frac{dx}{\sqrt{\pi}} e^{-(x+y-Q)^2} x^2 + (1-p) \int_0^\infty \frac{dx}{\sqrt{\pi}} e^{-(x-y-Q)^2} x^2 \right) + \frac{1}{q_0 - q} \left( p(y-Q) \int_0^\infty \frac{dx}{\sqrt{\pi}} e^{-(x+y-Q)^2} x^2 (x+y-Q) - (1-p)(y+Q) \int_0^\infty \frac{dx}{\sqrt{\pi}} e^{-(x-y-Q)^2} x^2 (x-y-Q) \right) = \frac{t}{\alpha} \]

\[ \frac{t-t_0}{p} + z + \frac{N_i}{N_0} \int_0^\infty \frac{dx}{\sqrt{\pi}} e^{-(x+z)^2} x^2 (x+z) - \frac{N_e}{N_0} \int_0^\infty \frac{dx}{\sqrt{\pi}} e^{-(x-z)^2} x^2 (x-z) = 0 \]

\[ -z^2 + \frac{N_i}{N_0} \int_0^\infty \frac{dx}{\sqrt{\pi}} e^{-(x+z)^2} x^2 + \frac{N_e}{N_0} \int_0^\infty \frac{dx}{\sqrt{\pi}} e^{-(x-z)^2} x^2 + \frac{(t-t_0)}{2t} \left[ \frac{N_i}{N_0} \int_0^\infty \frac{dx}{\sqrt{\pi}} e^{-(x+z)^2} + \frac{N_e}{N_0} \int_0^\infty \frac{dx}{\sqrt{\pi}} e^{-(x-z)^2} - 1 \right] = \frac{1}{2} - q_0 \frac{(t-t_0)^2}{t} \]

\[ -z^2 + \frac{N_i}{N_0} \int_0^\infty \frac{dx}{\sqrt{\pi}} e^{-(x+z)^2} x^2 + \frac{N_e}{N_0} \int_0^\infty \frac{dx}{\sqrt{\pi}} e^{-(x-z)^2} x^2 + \frac{(t-t_0)}{2t} \frac{N_i}{N} \left[ -2 \int_0^\infty \frac{dx}{\sqrt{\pi}} e^{-(x+z)^2} x^2 - 2 \int_0^\infty \frac{dx}{\sqrt{\pi}} e^{-(x-z)^2} x^2 (x+z) + \frac{t_0}{t} \int_0^\infty \frac{dx}{\sqrt{\pi}} e^{-(x+z)^2} \right] + \frac{(t-t_0)}{2t} \frac{N_e}{N} \left[ -2 \int_0^\infty \frac{dx}{\sqrt{\pi}} e^{-(x-z)^2} x^2 + 2 \int_0^\infty \frac{dx}{\sqrt{\pi}} e^{-(x-z)^2} x^2 (x-z) + \frac{t_0}{t} \int_0^\infty \frac{dx}{\sqrt{\pi}} e^{-(x-z)^2} \right] = \frac{1}{2} - q \frac{(t-t_0)^2}{t} \quad (D.26) \]

Eliminating \( t \) and \( t_0 \) from these equations and taking into account the fact that \( q_0 > q > 0 \)

we arrived at the system of four equations, which gave the value of the critical capacity in terms
of the three model parameters, $\alpha_{c}\left(\frac{N_i}{N}, p, \kappa\right)$:

$$p\left(\frac{1}{\sqrt{\pi}} e^{-(y-Q)^2} - (y-Q)(1-\text{erf}(y-Q))\right) - (1-p)\left(\frac{1}{\sqrt{\pi}} e^{-(y+Q)^2} + (y+Q)(1+\text{erf}(y+Q))\right) = 0$$

$$-\frac{2z^2 + 1 + \frac{N_i - N_c}{N} \left(\frac{2z}{\sqrt{\pi}} e^{-z^2} + (2z^2 + 1)\text{erf}(z)\right)}{\left(z + \frac{N_i - N_c}{N} \left(\frac{1}{\sqrt{\pi}} e^{-z^2} + \text{erf}(z)\right)\right)^2} = p\kappa^2 \frac{2Q^2 (1-p)}{2Q^2 (1-p)}$$

$$\alpha_{c} = \frac{1 + \frac{N_i - N_c}{N} \text{erf}(z)}{1 - \text{erf}(y-Q) p + \text{erf}(y+Q)(1-p)}$$

$$\text{sign}\left(z + \frac{N_i - N_c}{N} \left(\frac{1}{\sqrt{\pi}} e^{-z^2} + \text{erf}(z)\right)\right) \leq 0; \quad Q \geq 0$$

(D.27)

These four equations, subject to the two inequality constraints, are solved numerically in Chapter 5.

We note that one of the solutions of the four equations in D.27 is relatively trivial:
Because this is the only solution of the fully robust model, $\kappa \to \infty$, we will refer to it as the fully robust solution. This fully robust solution is symmetric with respect to $N_i \leftrightarrow N_e$ and $p \leftrightarrow 1 - p$ permutations.

At a certain values of parameters, $N_i^0 / N$, $p^0$, and $\kappa^0$, the fully robust solution becomes unstable, and is no longer valid. To find this point of instability we expanded Equations D.27 to the second order near $Q = 0$ and search for the co-existence of the two solutions. This lead to the equations for the instability point:

$$z^0 + \left(2 - \frac{N_i^0}{N} \right) \left( \frac{1}{\sqrt{\pi}} e^{-(z^0)^2} + \text{erf} \left( z^0 \right) \right) = 0$$

$$y^0 + (1 - 2p^0) \left( \frac{1}{\sqrt{\pi}} e^{-(y^0)^2} + \text{erf} \left( y^0 \right) \right) = 0$$

$$\kappa^0 = -\frac{z^0 \left(1 - 2p^0\right) \left(1 + (1 - 2p^0) \text{erf} \left( y^0 \right) \right)}{2p^0y^0 \sqrt{2p^0 \left(1 - p^0\right) \left(1 + \left(2 - \frac{N_i^0}{N} \right) \text{erf} \left( z^0 \right) \right)}}$$

$\frac{N_i^0}{N} \leq \frac{1}{2}$

Distribution of synaptic weights at the critical capacity

In this subsection we calculated the distribution of synaptic weights at critical capacity, $P(J)$, by
using the same replica theory based approach. From the first principles, the distribution of synaptic weights of a perceptron input \( l \) can be defined as:

\[
P(J) = \left( \frac{1}{\Omega(\zeta_j, \sigma^\mu)} \right) \prod_{j=1}^{N} dJ_j \delta(J - J_l) \prod_{\mu=1}^{m} \theta \left( \sigma^\mu \left( \frac{1}{N} \sum_{j=1}^{N} J_j \zeta_j^\mu - 1 \right) - \frac{\kappa}{\sqrt{N}} \right) \prod_{j=1}^{N} \theta(J_j g_j) . \tag{D.30}
\]

The only dependence of this distribution on \( l \) is due to the excitatory or inhibitory nature of the input. If input \( l \) is excitatory, the resulting distribution is of excitatory weights, and vice versa.

The average in Equation D.30 is calculated in the limit,

\[
P(J) = \lim_{n \to 0} \left( \frac{1}{\Omega(\zeta_j, \sigma^\mu)} \right) \prod_{j=1}^{N} dJ_j \delta(J - J_l) \prod_{\mu=1}^{m} \theta \left( \sigma^\mu \left( \frac{1}{N} \sum_{j=1}^{N} J_j \zeta_j^\mu - 1 \right) - \frac{\kappa}{\sqrt{N}} \right) \prod_{j=1}^{N} \theta(J_j g_j) . \tag{D.31}
\]

Assigning the \( \delta \)-function in this expression to the first replica, \( a = 1 \), we obtain:

\[
P(J) = \lim_{n \to 0} \left( \frac{1}{\Omega(\zeta_j, \sigma^\mu)} \right) \prod_{a,j=1}^{n,N} dJ_j \delta(J - J_l^a) \prod_{\mu,a=1}^{m,n} \theta \left( \sigma^\mu \left( \frac{1}{N} \sum_{j=1}^{N} J_j \zeta_j^\mu - 1 \right) - \frac{\kappa}{\sqrt{N}} \right) \prod_{j,a=1}^{N,n} \theta(J_j^a g_j) . \tag{D.32}
\]

The remaining calculation proceeds along the lines similar to the calculation of

\[
\left( \frac{1}{\Omega(\zeta_j, \sigma^\mu)} \right) . \tag{D.33}
\]

The result is:
\[ P(J) = \lim_{n \to 0} \int \prod_{a=1}^{n} \frac{ds^a d\hat{s}^a}{2\pi \sqrt{N}} \prod_{a,b=1}^{n} \frac{dq^a,b dq^a,b}{2\pi} G_{s_1}(\{\hat{s}^a\}, \{q^a,b\}) + \sum_{\alpha=1}^{N} \left( \sum_{a=1}^{n} \alpha \hat{s}^a \right) + \sum_{a,b=1}^{n} \frac{1}{2} \alpha^2 q^a,b q^a,b \]

\[ G_s(\{\hat{s}^a\}, \{q^a,b\}) = \ln \left( \prod_{a=1}^{n} \frac{du^a d\hat{u}^a}{2\pi} e^{-\frac{1}{2} \sum_{a,b=1}^{n} \hat{u}^a q^a,b \hat{u}^b} \left( \sum_{\alpha=1}^{N} \alpha^2 p^a \right) + (1-p)e^{-i \sum_{a=1}^{n} q^a \left( p^a - a^2 - k \right)} \right) \]

\[ G_{s_1}(\{\hat{s}^a\}, \{\hat{q}^a,b\} | J) = \ln \left( \prod_{a=1}^{n} d'^a \delta(J \pm J^a) e^{\frac{i}{2} \sum_{a,b=1}^{n} \hat{q}^a,b J^{a,b}} \right) \]

In this expression, the top sign in \( \pm \) corresponds to neuron \( l \) being inhibitory. After saddle point calculation of the integral in \( P \), term \( G_{s_1} \), in the \( n \to 0 \) limit, is the dominant term in the exponent, and

\[ P(J) = e^{G_{s_1}(\{\hat{s}^a\}, \{\hat{q}^a,b\} | J)} \frac{\prod_{a=1}^{n} d'^a \delta(J \pm J^a)}{\prod_{a=1}^{n} d'^a} e^{\frac{i}{2} \sum_{a,b=1}^{n} \hat{q}^a,b J^{a,b}} \]

Parameters \( \hat{s}^a, \hat{q}^a,b \) are given by the replica symmetric saddle point Equations D.20, where it is assumed that \( \hat{s}^a = \hat{s}, \hat{q}^{a,a} = \hat{q}_0, \) and \( \hat{q}^{a,b} = \hat{q} \). After the Hubbard-Stratonovich transformation, Equation D.21, we get:

\[ P(J) = \frac{1}{\sqrt{\pi}} e^{-x^2} \left( \int d' e^{i(q_0 \hat{q} - \hat{q}_0 q) + i \hat{q} \hat{q}_0} \right)^{n-1} e^{i(q_0 \hat{q} - \hat{q}_0 q) + i \hat{q} \hat{q}_0} \]

Using the previous notation, \( r = i\hat{s}, t_0 = -i\hat{q}_0, t = -i\hat{q}, \) and taking the \( n \to 0 \) limit we arrived at:
This function is properly normalized for $J < 0$ and $J > 0$. Combining the distributions for the inhibitory and excitatory synaptic weights into a single expression, and, as before, replacing $\frac{r}{2\sqrt{t}}$ with $z$ we obtain:

\[
P(J) = \theta(-J) \frac{N_i}{N} \int dx \frac{e^{-x^2}}{\sqrt{\pi}} \int d'J e^{-\left((t-t_0)J^2 - 2zJ\sqrt{J}\right)} + 
\]

\[
\theta(J) \frac{N_e}{N} \int dx \frac{e^{-x^2}}{\sqrt{\pi}} \int d'J e^{-\left((t-t_0)J^2 - 2(z-x)\sqrt{J}\right)}
\]

The integrals in Equation D.37 can be simplified in the vicinity of the critical capacity by using:

\[
e^{-a(x-a)^2} \underset{a \to a}{\overset{\pi}{\to}} \sqrt{\pi} \delta(x-a)
\]

The resulting distribution of synaptic weights is a combination of two Gaussian functions truncated at zero and a $\delta$-function at zero:

\[
P(J) = \frac{1}{\sqrt{2\pi} \sigma} e^{-\left(\frac{J}{\sqrt{2\pi} \sigma} + z\right)^2} \left(\frac{N_i}{N} \theta(-J) + \frac{N_e}{N} \theta(J)\right) + \frac{1}{2} \left(1 - \frac{N_i - N_e}{N} \text{erf}(z)\right) \delta(J)
\]

\[
\sigma = \frac{-\sqrt{2}}{z + \frac{N_i - N_e}{N} \left(\frac{1}{\sqrt{\pi}} e^{-z^2} + \text{erf}(z)z\right)}
\]

Parameter $\sigma$, which is the standard deviation of the not truncated Gaussian functions, describes the width of the synaptic weight distribution. It depends on $z$, which is given by Equations D.27. For the fully robust solution, Equations D.28, $\sigma \to \infty$, and the average absolute synaptic weight diverges. This is because, in this limit, the region of solutions becomes
The fraction of zero-weight synapses, $N_0/N$, the connection probability, $P_{con}$, and the averages of $|J|$ and $J$ can be easily obtained from Equation D.39:

\[
\begin{align*}
\frac{N_0}{N} &= \frac{1}{2} \left(1 - \frac{N_i - N_e}{N} \operatorname{erf} (z) \right) \\
P_{con} &= \frac{1}{2} \left(1 + \frac{N_i - N_e}{N} \operatorname{erf} (z) \right) \\
\langle |J| \rangle &= \frac{-1}{p} \left( \frac{z}{N} \frac{N_i - N_e}{N} \frac{1}{\sqrt{\pi}} e^{-z^2} + \operatorname{erf} (z) z \right) \\
\langle J \rangle &= \frac{1}{p}
\end{align*}
\]

(D.40)

The average synaptic weight in our model is $1/p$. This is consistent with Equations D.3 and our original assumption that the perceptron operates near the firing threshold.
Bibliography


