Three-dimensional reverse engineering of neuronal microcircuits: *The barrel cortex in silico*

Dissertation

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

der Mathematisch-Naturwissenschaftlichen Fakultät

und

der Medizinischen Fakultät

der Eberhard-Karls-Universität Tübingen

vorgelegt von

Daniel Udvary

aus Stuttgart-Bad Cannstatt, Deutschland

Juli 2020

Tag der mündlichen Prüfung: 22. Februar 2021

Dekan der Math.-Nath. Fakultät: Prof. Dr. W. Rosenstiel

Dekan der Medizinischen Fakultät: Prof. Dr. I. B. Autenried

1. Berichterstatter: Dr. M. Oberlaender

2. Berichterstatter: Prof. Dr. C. Schwarz

Prüfungskommission: Prof. Dr. A. Burgalossi

Prof. Dr. C. Schwarz

Prof. Dr. J. H. Macke

Dr. M. Oberlaender

Ich erkläre, dass ich die zur Promotion eingereichte Arbeit mit dem Titel: "Three-dimensional reverse engineering of neuronal microcircuits: The barrel cortex in silico" selbstständig verfasst, nur die angegebenen Quellen und Hilfsmittel benutzt und wörtlich oder inhaltlich übernommene Stellen als solche gekennzeichnet habe. Ich versichere an Eides statt, dass diese Angaben wahr sind und dass ich nichts verschwiegen habe. Mir ist bekannt, dass die falsche Abgabe einer Versicherung an Eides statt mit Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft wird.

Tübingen, den ....

Datum Unterschrift

Abstract

The mammalian neocortex is one of the most complex biological tissues and the center of higher brain functions. Currently, the specific distribution of neurons and neurites, as well as their intricate wiring within an entire neocortical area that emerge during development and are then refined throughout life, are not accessible. Here, I present a reverse engineered model of one neocortical area, the rat barrel cortex. First, I created a model of its structural composition constraint by measurements of cortex geometry, neuron distributions, and morphological reconstructions. This model provided anatomically realistic and robust estimates of the area's neuron and neurite distributions and captured the structural principles preserved across individuals. Second, I used the model's distribution of neurites to constrain synapse formation. Specifically, I introduced a stochastic synapse formation strategy that predicts the area's wiring diagrams if they were solely shaped by the area's structural composition in the absence of any learning or plasticity rules. I find that the predicted wiring diagrams are sparse, heterogeneous, correlated, and structured unlike random networks — all of which are either observed or speculated properties of neocortical wiring. A systematic comparison between predicted and empirical wiring properties on the subcellular, cellular, and network level revealed a high degree of consistency. This demonstrates that the structural organization of the neuropil provides strong constraints for synapse formation. For the consistently predicted wiring properties, such as connection probabilities, it cannot be ruled out that they were shaped by the area's structural composition, i.e., implicitly by the developmental mechanisms that positioned neurons and neurites within the neuropil. A more sophisticated synapse formation strategy is not necessarily required. In contrast, such a sophisticated strategy might underlie the inconsistently predicted wiring properties, e.g., the frequency of certain circuit motifs. The herein presented approach can hence act as a starting point to identify wiring correlates of sensory experience or learning and provide a foundation to explore the relationship between synapse formation, an area's structural composition, and network architecture.

Acknowledgments

I would first like to thank my supervisor, Dr. Marcel Oberlaender, for his guidance, support, trust, and great expertise, and for giving me the opportunity to pursue my doctoral studies in his group. Also, thank you for always keeping your door open.

I would like to thank my advisory committee, Prof. Dr. Matthias Bethge and Prof. Dr. Cornelius Schwarz, for helpful comments and discussions.

I would like to thank all collaborators for their outstanding teamwork (in alphabetical order): Arco Bast, Dr. Vincent J. Dercksen, Dr. Robert Egger, Ela Elsholz, Prof. Dr. Dirk Feldmeyer, Jason M. Guest, Philipp Harth, Prof. Dr. Hans-Christian Hege, Prof. Dr. Christiaan P. J. de Kock, Prof. Dr. Jakob Macke, Dr. Hanno S. Meyer, Dr. Rajeevan T. Narayanan, Prof. Dr. Bert Sakmann, Dr. David A. Slabik, and Felipe Yanez. Further thanks go to Dr. Marlene Arzt, Dr. Moritz Helmstaedter, Dr. Mike Hemberger, and Dr. Christian Koelbl for providing reconstructions of inhibitory neurons.

Special thanks to former and current members of Dr. Marcel Oberlaender's group in both Tübingen and Bonn for the many helpful discussions and friendships. It was a pleasure working with you all over the past six years.

Last but not least, I would like to express my very profound gratitude to my family and friends from far and near, and in particular to my girlfriend Tanja. Thank you all for providing me with unfailing support, continuous encouragement, and keeping me sane over the years. This accomplishment would not have been possible without you.

Finally, I would like to leave the remaining space in memory of
Katrin Kutscheidt, a wonderful friend and fellow student,
Alison Smith, a friend and team member of the lab,
and my grandmother, who I will be ever grateful for her assistance, and I am
sorry that she has not lived to see me graduate.

Contents

List of Figures

List of Tables

List of abbreviations

CV	Coefficient of variations
DSC	Dense structural composition
EX	Excitatory neuron
GranAsc	Granular-Ascending
GranSym	Granular-Symmetric
IN	Inhibitory interneuron
InfraAsc	Infragranular-Ascending
L1/2/3/4/5/6	Cortical layer 1/2/3/4/5/6
L2PY/L3PY/L4PY	L2 pyramidal neuron/L3 pyramidal neuron/L4 pyramidal neuron
L4sp/L4ss	L4 star pyramidal neuron/L4 spiny stellate neuron
L5IT/L5PT	L5 slender-tufted intratelencephalic neuron/L5 thick-tufted pyramidal tract neuron
$L6_A$CC/$L6_B$CC/L6CT	L6A corticocortical pyramidal neuron/L6B polymorphic corticocortical pyramidal neuron/L6 corticothalamic neuron
MLE	Maximum likelihood estimate
PRE	Presynaptic structures (i.e., boutons)
POST	Postsynaptic structures (e.g., spines of excitatory neurons)
SAD	Summed absolute difference
SD	Standard deviation
SEM	Standard error of the mean

SupraDesc	Supragranular-Descending
VPM	Ventral posterior medial nucleus of the thalamus
vS1	vibrissal part of the primary somatosensory cortex (informally: barrel cortex)
WM	White matter tract

Glossary

Dense structural model	3D model that approximates the dense structural composition of a certain brain volume (i.e., distribution of all neurons and neurites including distribution of pre- and postsynaptic structures) (Egger, Dercksen, Udvary, Hege, & Oberlaender, 2014)
Dense statistical connectome	Model that provides probabilities of each pre- and postsynaptic structure being connected by a synapse (Egger et al., 2014). Allows deriving a statistical wiring diagram.
Geometrical model	3D model of the average geometry of a certain brain volume (e.g., pial and white matter surfaces) (Egger, Narayanan, Helmstaedter, de Kock, & Oberlaender, 2012).
In silico	Experiment performed on computer
In vitro	"Within the glass", i.e., experiments performed in a controlled environment outside of the living organism
In vivo	"Within the living", i.e., experiments performed in the living organism

1 | Introduction

The functionality of the neocortex relies on the formation of complex networks, which are the result of millions of years of evolution (Rakic, 2009; Kaas, 2011). Once development is completed, the overall structure, such as the distribution of neurons and neurites and their network architecture, remains largely stable. Still, individual wiring patterns are refined by learning and memory throughout life (Bennett, Kirby, & Finnerty, 2018). Over the past decades, enormous progress has been made in understanding the biological and cellular mechanisms underlying the formation of certain wiring patterns (Sanes & Yamagata, 2009; Shen & Scheiffele, 2010; Yogev & Shen, 2014; De Wit & Ghosh, 2016; Südhof, 2018). Yet deciphering the logic of neocortical networks remains a fundamental challenge in neuroscience (Hassan & Hiesinger, 2015; Schröter, Paulsen, & Bullmore, 2017; Zador, 2019; Takahashi, 2019; Barabási & Barabási, 2020). This is mainly because the neocortex is a vastly complex and heterogeneous biological tissue (DeFelipe, 2015): Within a cubic millimeter of neocortex, several kilometers of dendrites and axons (Braitenberg & Schüz, 1998) that originate from a functionally and morphologically diverse spectrum of hundreds of thousands of neurons (Meyer et al., 2013; Narayanan et al., 2015; Harris & Shepherd, 2015), form on the order of a billion connections (Santuy, Rodriguez, DeFelipe, & Merchan-Perez, 2018) of varying lifetimes (Berry & Nedivi, 2017).

1.1 Development gives rise to a densely and diversely structured neocortex

The stereotypical development of the neocortex and its division into highly specialized areas is the result of genetically induced, evolutionarily conserved processes (Rash & Grove, 2006; Rakic, 2009). Guided by radial glial cells, neurons migrate from their birthplace to their final position in the neocortex (Kriegstein & Noctor, 2004). Excitatory neurons that comprise about 80% of all neurons in the neocortex are born in the ventricular zone of the embryonic neocortex and migrate radially in an inside-first, outside-last manner: Excitatory neurons of the deep layers are born first, while later-born excitatory neurons migrate past them towards more superficially located layers (see Figure 1.1). In contrast, inhibitory neurons that comprise the remaining 20% of neocortical neurons are born in different subcortical proliferative zones. Inhibitory neurons migrate over long distances, first tangentially, and then radially to their final positions (Kepecs & Fishell, 2014).

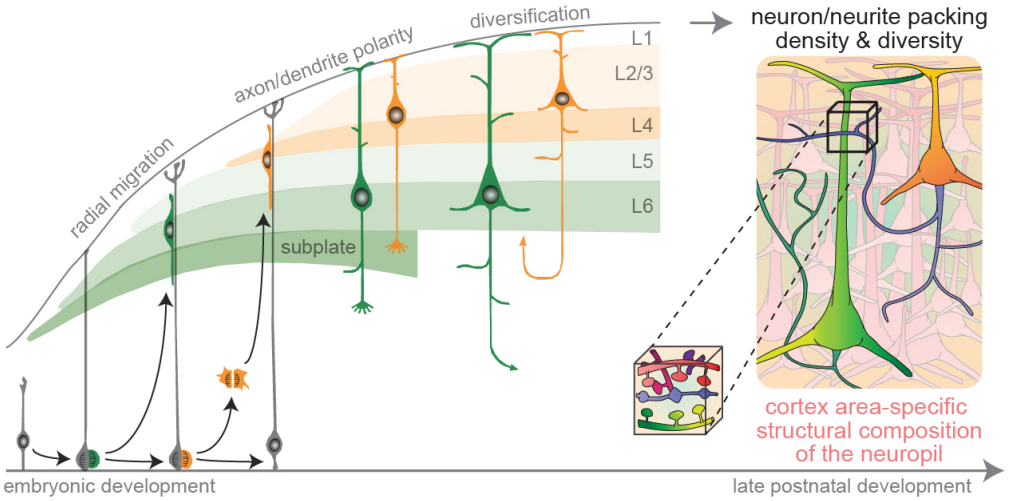

Figure 1.1. Schematic of neocortical development. Genetically induced programs shape the structural composition of the neuropil resulting in an area-specific densely and diversely packed neuropil. The positioning of neurons and neurites during development provides structural constraints for which neurons could in principle form synapses with each other. Schematic illustration on the left modified from Kwan et al. (2012). Figure adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020).

During migration, the neuron's shape starts polarizing, forming a leading process and a trailing process. Depending on the cell type, each process becomes either the dendrite or the axon (A. P. Barnes & Polleux, 2009; Cheng & Poo, 2012). While dendrites receive and integrate information from other neurons via structures called synapses, axons transmit electrical signals in form of action potentials generated at the soma to other target neurons by forming synapses with the targets' dendrites and/or somata. Once axon and dendrite are differentiated, the neurites mature in their final shapes in a highly dynamic process. An array of cell-intrinsic and cell-extrinsic processes such as extracellular molecular gradients and activity-dependent processes guide the neurons' axons and dendrites through the neuropil into particular subvolumes (Lefebvre, Sanes, & Kay, 2015; Wamsley & Fishell, 2017; Jabaudon, 2017; Stoeckli, 2018).

During certain temporal windows, referred to as "critical periods", lack of sensory inputs can disrupt the developmental of neuron morphologies resulting in functional deficits of the respective cortical area (Feldman & Brecht, 2005; Fox & Wong, 2005; Erzurumlu & Gaspar, 2012). Axons are primarily guided by receptors on their elongating tip, called "growth cone", that respond to repulsive and attractive guidance cues in their environment (Stoeckli, 2018). While axons in the primary visual cortex develop horizontal projection patterns interconnecting columns of similar orientation tuning (Hubel & Wiesel, 1959; Bosking, Zhang, Schofield, & Fitzpatrick, 1997), axons in the vibrissal part of the primary somatosensory cortex (vS1) develop asymmetric horizontal projection patterns forming two orthogonal strata (Narayanan et al., 2015). The general shape of excitatory dendrites is predominately specified by cell-intrinsic processes. In contrast, cell-extrinsic processes such as area-specific environmental cues determine the detailed branching of excitatory dendrites, starting with an overproduction of dendrite branches that then either stabilize or retract (Lefebvre et al., 2015). The majority of neurons in layer 4 of vS1 first develop symmetrically shaped dendrites that then orient their dendritic arbor towards where sensory input arrives by branch addition and elimination (Van der Loos & Woolsey, 1973; Espinosa, Wheeler, Tsien, & Luo, 2009; Simi & Studer, 2018). Area-specific cues are especially crucial in determining the shape and thus the function of inhibitory neurons as shown by transplanting migrating inhibitory neurons in different cortical environments (e.g., Quattrocolo,

Fishell, & Petros, 2017; Ishino et al., 2017) (for review, see Fishell & Kepecs, 2019).

Within each neocortical area, a densely and diversely packed neuropil develops comprising highly specialized neuron morphologies that can be grouped into distinct morphological types (e.g., Staiger, Möck, Proenneke, & Witte, 2015; Narayanan et al., 2015; Harris & Shepherd, 2015; Gouwens et al., 2019) (see Figure 1.2). The specialization of neuron morphologies for each neocortical area is accompanied by the emergence of complex synaptic networks within and across areas. Excitatory synapses are typically formed between protrusions along dendrites, called spines, and swellings along axons or bulbs at the end of an axon, called boutons. While spines are highly dynamic during development (Alvarez & Sabatini, 2007), bouton are less dynamic (De Paola et al., 2006). First, synapses between neurons are already formed prenatally, and synapse numbers surge during early development, followed by a phase of synapse elimination and refinement (Bourgeois & Rakic, 1993; Alvarez & Sabatini, 2007; Erzurumlu & Gaspar, 2012). Neurons in layer 4 of vS1 are highly interconnected before connections are eliminated in later periods of development (Tarusawa et al., 2016). Synapse formation itself might be involved in refining neuron morphologies, known as the "synaptotrophic hypothesis" (Vaughn, Barber, & Sims, 1988; Niell, Meyer, & Smith, 2004; Cline & Haas, 2008). Once developed, a complex network of trillions of synaptic connections has formed with individual neurons receiving thousands of synaptic inputs (Schoonover et al., 2014).

1.2 The neocortex is largely stable but refined throughout life

Development aims to ensure that each neocortical area can elicit and maintain its functionality. Therefore, the neocortex is highly plastic during development and able to respond to extrinsic changes by modifying its structural composition (Fox, 2002; Erzurumlu & Gaspar, 2012; Feldman & Brecht, 2005). However, after development is completed and critical periods have ended, plasticity is substantially

Figure 1.2. Neocortex comprises diverse neuron morphologies whose neurites innervate different subvolumes. Here, the area-specific morphologies of excitatory neurons found within the vibrissal part of the primary somatosensory cortex (vS1) are depicted. Axons are color-coded by cell type (Narayanan et al., 2015), dendrites and exemplary thalamic axon in black. For more details on the different cell types of vS1, see Section 2.1.1. Figure adapted from Egger et al. (2020).

reduced[1]. Then, the structural components, such as the shape and position of neurons and their wiring, represent the result of cortical development that is sufficient for proper function. These structural components, specifically the distribution of neurons and neurites, are mostly maintained throughout the remaining life, hence providing lifelong constraints for synapse formation.

The vast majority of axon branches and boutons, as well as dendrite branches and spines, are stable. In vS1 of adolescent mice, 90% of excitatory boutons persist over at least two weeks (Qiao et al., 2016), and 55% of spines persist over at least one month (Trachtenberg et al., 2002; Holtmaat et al., 2005). Similarly, but wider varying rates of persistence were observed for inhibitory synapses (Marik, Yamahachi, McManus, Szabo, & Gilbert, 2010; Eavri et al., 2018; Chen et al., 2012). Together with findings that connection probabilities involving inhibitory neurons do not differ between young and mature animals (Fino & Yuste, 2011; Packer, McConnell, Fino, & Yuste, 2013) (but see Jiang et al., 2015; Scala et al., 2019), these results indicate that neurites of inhibitory neurons, as well as their synapses, are

[1]Note that under pathological conditions large-scale remodeling of the neocortex occurs (e.g., Emoto, 2011).

fairly stable during adulthood. Throughout life, however, wiring between neurons and their pre- and postsynaptic structures undergo fine-scale refinements (e.g., De Paola et al., 2006; Oberlaender, Ramirez, & Bruno, 2012; Qiao et al., 2016), potentially reflecting structural correlates of experiences, learning, and memory processes (e.g., van Kerkoerle, Marik, Meyer Zum Alten Borgloh, & Gilbert, 2018; Albieri et al., 2015) (for review, see Bailey & Kandel, 1993; Chklovskii, Mel, & Svoboda, 2004; Fox & Wong, 2005; S. J. Barnes & Finnerty, 2010; van der Zee, 2015; Bennett et al., 2018).

1.3 What strategies of synapse formation shape neocortical networks?

The wiring of the developed neocortex is highly complex and specialized for each neocortical area. The formed intracortical connections vary between neurons based on their cell types (e.g., Bruno & Sakmann, 2006; Constantinople & Bruno, 2013), long-range projection targets (e.g., Brown & Hestrin, 2009), neuronal lineages and birth times (e.g., Yu, Bultje, Wang, & Shi, 2009; Gao, Sultan, Zhang, & Shi, 2013), or by laminar, absolute, or relative position of their somata (e.g., C. C. Petersen & Sakmann, 2000; Lefort, Tomm, Floyd Sarria, & Petersen, 2009; Perin, Berger, & Markram, 2011; Avermann, Tomm, Mateo, Gerstner, & Petersen, 2012). On the subcellular level, synaptic inputs vary along dendrites (e.g., Schoonover et al., 2014; Qi & Feldmeyer, 2016) and depend on the subcellular domain (e.g., Motta et al., 2019; Karimi, Odenthal, Drawitsch, Boergens, & Helmstaedter, 2020). Subcellular target-specificity is such a prominent feature among inhibitory neurons, that they are traditionally characterized based on the subcellular domain of excitatory neurons they target (DeFelipe, 1997; Somogyi, Tamás, Lujan, & Buhl, 1998; Kubota, 2014). Most notably are Chandelier cells that target the axon initial segment of excitatory neurons (Somogyi, 1977; Inan & Anderson, 2014). On the network level, the observed topologies of networks are unlike those expected in a randomly connected network (e.g., Song, Sjöström, Reigl, Nelson, & Chklovskii, 2005; Rieubland, Roth, & Häusser, 2014). For example, recurrent network motifs are more frequently observed in the neocortex than expected in a random network

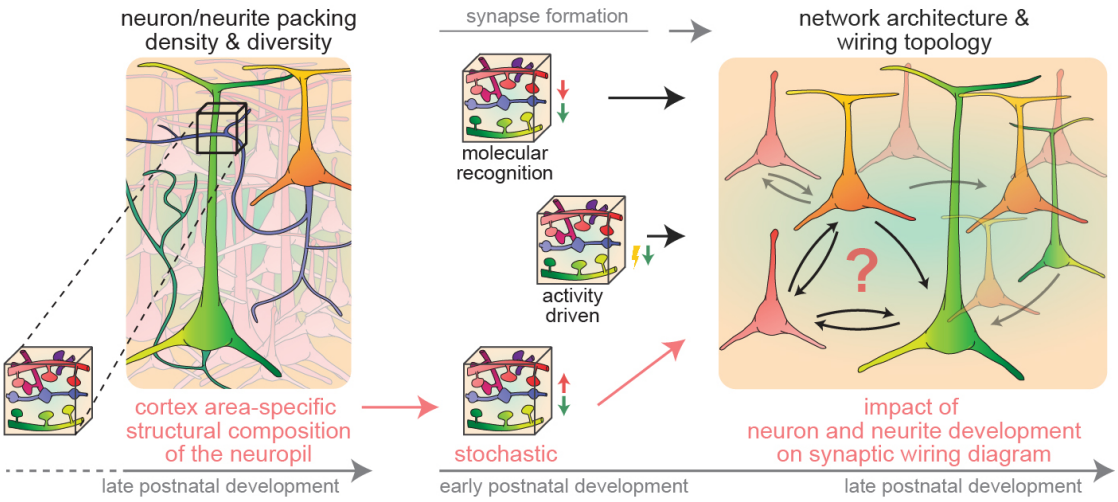

Figure 1.3. Approach for testing the impact of the area's structural composition (i.e., distribution of neurons and neurites as established by development) on neocortical wiring. Given the characteristic structural composition of one neocortical area, I investigated how wiring diagrams would look like in the absence of recognition-dependent synapse formation strategies that, for example, rely on activity or molecular recognition between specific pre- and postsynaptic compartments. See main text for more details. Figure adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020).

(e.g., Song et al., 2005; Kampa, Letzkus, & Stuart, 2006; Perin et al., 2011).

The discoveries of complex and convoluted mechanisms at the cellular and molecular level that are directly involved in synapse formation and elimination[2] have contributed enormously to our understanding of individual connectivity patterns (Sanes & Yamagata, 2009; Shen & Scheiffele, 2010; Yogev & Shen, 2014; De Wit & Ghosh, 2016; Südhof, 2018). However, the logic underlying the complex wiring observed in the neocortex is still largely a mystery. The complex wiring observed is hence commonly attributed to strategies of synapse formation that instruct attraction or repulsion between specific synaptic partners (e.g., Song et al., 2005; Sanes & Yamagata, 2009; Kasthuri et al., 2015) (but see Hassan & Hiesinger, 2015) (see Figure 1.3).

One envisioned scenario of how such a recognition-dependent synapse formation strategies could be implemented is molecular recognition, also known as

[2]in particular in the peripheral nervous system (Sanes & Lichtman, 1999; Lanuza et al., 2018)

the "chemoaffinity hypothesis" (Langley, 1895; Sperry, 1963). According to the chemoaffinity hypothesis, synapses are formed between structures that recognize each other by matching pairs of receptors or adhesion molecules. Evidence for this mostly genetically encoded mechanism has been described, for example, in the olfactory system of the *Drosophila* (Hong, Mosca, & Luo, 2012; Ward, Hong, Favaloro, & Luo, 2015) and in the mammalian retina (Krishnaswamy, Yamagata, Duan, Hong, & Sanes, 2015) (for review, see Zipursky & Sanes, 2010).

Another scenario reflecting a recognition-dependent synapse formation strategy is activity-driven synapse formation and elimination. Spontaneous or sensory-evoked neuronal activity is especially implicated in synapse elimination and stabilization (Katz & Shatz, 1996; Flores & Méndez, 2014; Fauth & Tetzlaff, 2016). During development, neural activity is crucial for the proper wiring of the neuromuscular junction (Sanes & Lichtman, 1999; Favero, Busetto, & Cangiano, 2012) but also in other systems such as sensory systems (Katz & Shatz, 1996; Hooks & Chen, 2006; Fox & Wong, 2005). In adult vS1, neural activity evoked by increased sensory experience leads to the formation of both excitatory and inhibitory synapses (Knott, Quairiaux, Genoud, & Welker, 2002). Theoretical work has demonstrated that plasticity and learning mechanisms, induced by activity, can give rise to many of the observed connectivity patterns (e.g., Miner & Triesch, 2016; D. Zhang, Zhang, & Stepanyants, 2019). Beyond inducing structural changes to the wiring diagram by forming and eliminating synapses, activity is involved in inducing functional synaptic plasticity, i.e., by modifying synaptic strength (Hebb, 1949; Feldman, 2009; Magee & Grienberger, 2020). It is important to note that: First, molecular recognition and activity are not mutually exclusive mechanisms (Cline & Haas, 2008) as activity can induce genes (Nedivi, 1999) and guide adhesion molecules (Cantallops & Cline, 2008). Second, any single synapse in the neocortex is most likely not the result of one but of many convoluted mechanisms (Sanes & Yamagata, 2009).

In contrast to synapse formation strategies that connect specific synaptic partners based on activity and/or molecular identity, there is the hypothesis that structural features such as spatial overlap or proximity between individual axon and dendrite pairs imply synapse formation, often summarized under the term "Peters' rule" (Peters & Feldman, 1976; Braitenberg & Schüz, 1991; Rees, Moradi,

& Ascoli, 2017). In the peripheral nervous system, structural features between neurons or muscle fibers are implicated in synapse formation (in neuromuscular junction: Willshaw, 1981; Sanes & Lichtman, 1999; Lanuza et al., 2018; in spinal sensory-motor circuits: Li et al., 2007; Balaskas, Abbott, Jessell, & Ng, 2019). In the neocortex, structural features between neurons are implicated in pair-wise connectivity patterns (e.g., Fino & Yuste, 2011; Packer et al., 2013) and even in specific network topologies (Gal, Perin, Markram, London, & Segev, 2019). However, the proximity between individual axons and dendrites is only a necessary but not a sufficient condition for synapse formation as demonstrated by numerous studies at the subcellular and cellular level in the neocortex (e.g., Brown & Hestrin, 2009; Petreanu, Mao, Sternson, & Svoboda, 2009; Mishchenko et al., 2010; Kasthuri et al., 2015; Motta et al., 2019). Therefore, Peters' rule, irrespective of its many definitions and interpretations (Rees et al., 2017), can generally not capture the intricate wiring properties observed across different levels of detail in the neocortex. While all these studies have probed the relationship between structural properties of individual neurons and their connectivity in isolation and often in a binary fashion, a systematic investigation to which degree the characteristic structural composition of an entire neocortical area — once established by development — provides potentially lifelong constraints for synapse formation that are preserved across animals is missing.

Here, I aim to perform such a systematic investigation and explore the impact of a synapse formation strategy according to which synaptic connectivity is implicitly shaped by the stereotyped outcomes of developmental programs that positioned neurites of desired synaptic partners in close proximity to each other (see Figure 1.3). Once developed, axons and dendrites might form synapses in an *unspecific* manner with their locally available target structures. Such a strategy can be phenomenologically described as *stochastic* and formulated independently of the executing mechanisms at the cellular and molecular level (Hassan & Hiesinger, 2015).

1.4 The impact of an area's characteristic structural composition on its network properties

Exploring the impact of the characteristic structural composition of the neuropil of an entire neocortical area on its network architecture comes with several challenges. In particular, the complexity and high packing density of the neocortex make it difficult to access experimentally and to tract and probe its structural composition and networks over long periods of time. The main challenge hereby is that the distribution of neurons and neurites and their wiring diagrams of an entire neocortical area across animals and developmental stages is currently inaccessible. In theory, this data could be provided by multiple dense reconstructions of the neuropil using an electron microscopic approach and identifying the structural properties preserved across animals. However, state-of-the-art dense electron microscopic reconstructions are currently limited to small volumes that only comprise truncated neurons and are more than four orders of magnitude smaller than the entire rat vS1 (e.g., Kasthuri et al., 2015; Motta et al., 2019). Furthermore, the development of the structural composition of the neuropil and its networks are not two distinct processes; instead, they are convoluted (e.g., codependent), occur partly simultaneously, and can depend on the same mechanisms (e.g., activity). Also, wiring patterns of the neocortex remain to a lesser degree plastic throughout life (Bennett et al., 2018; see Section 1.2). Thus, it is infeasible to separate them or abolish, for example, activity-dependent synapse formation strategies without affecting the development of neuron and neurite distributions. Here, I overcome these challenges by applying a computational modeling approach that has been developed in the group of Dr. Marcel Oberlaender (Egger et al., 2014).

The goals of this thesis are:

1. creating an anatomically detailed model of the characteristic dense structural composition of the neuropil of an entire neocortical area post critical periods of neuron and neurite development. The model should:

   (a) capture the structural principles characteristic for the respective neocortical area with respect to its subcellular, cellular, morphological, and

cytoarchitectonic organization.

(b) provide robust and realistic estimates of the structural composition of the neuropil, such as neuron and neurite distributions, that are preserved across individuals.

2. exploring how a model of the structural composition of the neuropil can advance our understanding of neocortical connectivity and potentially function. Specifically, the aim is to answer the question: *How does the structural composition of the neuropil constrain neocortical networks?* To answer this question, I aim to:

   (a) predict how wiring diagrams of an entire neocortical area would look like if they were solely shaped by the structural composition of the neuropil.

   (b) quantitatively assess which properties of such predicted wiring diagrams represent observed properties of neocortical wiring.

In the following, I provide an overview of my strategy to achieve the above-stated goals, followed by a brief review of the model system used, the rat vibrissal system.

1.4.1 Strategy to explore the impact of the structural composition of the neuropil on neocortical networks

To achieve the previously stated goals and explore how the structural composition of a neocortical area constrains its networks, I used the following strategy that can be subdivided into four components (see Figure 1.4): (1) dense structural model, (2) synapse formation strategy, (3) dense statistical connectome model, and (4) probability distribution of observables and comparison to empirical wiring properties. I provide a short introductory overview of the four components and my contribution to each of them on the following pages. The components are explained in detail in the respective methods section (see Section 2).

(1) Dense structural model: The dense structural model provides access to the characteristic structural composition of a brain volume for a particular devel-

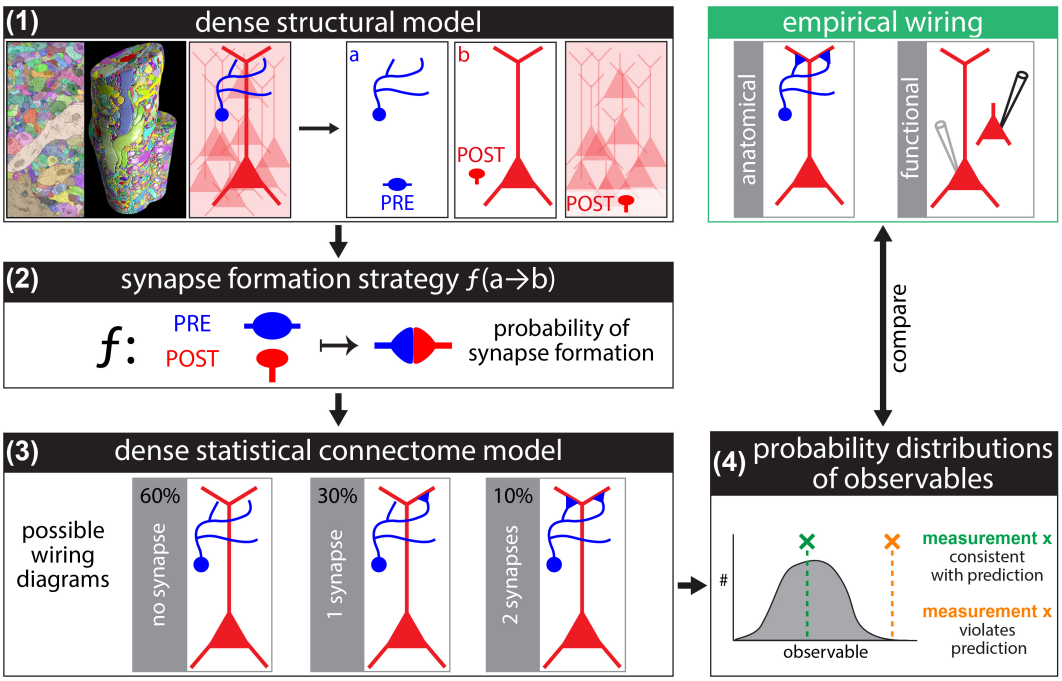

Figure 1.4. Strategy to explore the impact of the structural composition of the neuropil on neocortical networks. (1) Dense structural model: Anatomically detailed model of the characteristic structural composition of the neuropil that provides access to the 3D distribution of all pre- and postsynaptic structures (PRE and POST), including the entire morphology of the neuron they originate from in a brain volume. Exemplified for a presynaptic neuron a (blue), postsynaptic neuron b (red), and all other postsynaptic neurons in the same volume (shaded red). Two left most images adapted from Kasthuri et al. (2015). (2) Synapse formation strategy: Mathematical formulation that transforms the amount of pre- and postsynaptic structures into a probability of a synapse being formed from neuron a on neuron b. (3) Dense statistical connectome model: Prediction of all possible wiring diagram consistent with distribution of pre- and postsynaptic structures (1) and synapse formation rule (2), here illustrated for neuron pair (a, b). (4) Probability distribution of observables: Predicted probability distribution of an observable wiring property (e.g., connection probability) can be compared to its empirically determined value. For more details, see main text.

opmental stage in an animal's life. The characteristic structural composition refers to the distribution of neuron somata, its cell types, and neurites, including pre- and postsynaptic structures. Here, I used an empirically constraint dense structural model of rat vS1 that was developed in the group of Dr. Marcel Oberlaender (Egger et al., 2014). The model is constraint by measurements of cortex geometry and cytoarchitecture (Egger et al., 2012), neuron distribution (Meyer et al., 2013), neuron morphologies, and densities of pre- and postsynaptic structures (Narayanan et al., 2015; Oberlaender, de Kock, et al., 2012) across multiple animals of the same strain and developmental stage past neurite and neuron development. By integrating measurements across multiple animals, the model provides an estimate of the neuron and neurite packing densities and diversities that are preserved across vS1 of all probed animals. I refer to this model as the "dense structural model". Given its estimates, constraints for synapse formation can be derived, i.e., 3D locations and amount of pre- and postsynaptic structures (denoted by PREs and POSTs).

Contribution: In collaboration with my colleague Dr. Robert Egger (Max Planck Institute for Biological Cybernetics, Tübingen), I contributed to generating the dense structural model by incorporating inhibitory neurons (INs) (this is an ongoing project; see Section 2.1.2 and Egger et al., 2014). I assessed the interanimal variability of the empirical constraints and the subsequent robustness and limits of the model. I analyzed the structural estimates of the model (see Section 2.1.3 and Narayanan, Udvary, & Oberlaender, 2017), and assessed how consistent the structural estimates are compared to measurements that were not used to constrain the model, e.g., synapse densities and branchlets ratios obtained by electron microscopy (see Section 2.1.3).

(2) Synapse formation strategy: The synapse formation strategy describes the transformation of structural estimates provided by the dense structural model (e.g., PREs and POSTs) into a probability of a synapse being formed between them. The strategy is formulated as a mathematical formula. As introduced in Section 1.3, the synapse formation strategy used in this thesis is as follows: Synapses between PREs and POSTs are formed in an unspecific or stochastic manner based on their local availability (see Section 2.2). The mathematical foundation for this synapse formation strategy was formulated previously in the group

of Dr. Marcel Oberlaender (Egger et al., 2014). Note that the strategy is merely phenomenological and thus, independent of the underlying cellular and molecular mechanisms involved in synapse formation and elimination.

Contribution: In Egger et al. (2014) the mathematical formula underlying the herein described synapse formation strategy was part of an approach to test the validity of Peters' rule for a small set of empirically determined wiring properties. Here, I used the same mathematical formula but in a broader context. I reinterpreted the formula as part of a stochastic synapse formation strategy that allows testing the impact of the structural composition of the neuropil (see Section 2.2). In collaboration with my colleague Dr. Robert Egger (Max Planck Institute for Biological Cybernetics, Tübingen), I extended the formula to be applicable to INs, and in collaboration with the group of Prof. Dr. Hans-Christian Hege (Zuse Institute Berlin), I developed an approach that allows modification of the synapse formation strategy in a web application (see Section 2.6).

(3) Dense statistical connectome model: The dense statistical connectome model predicts all possible wiring diagrams including their probabilities of occurring (Egger et al., 2014), that are consistent with the structural estimates (i.e., PREs and POSTs) as provided by the dense structural model (1) and with the synapse formation strategy (2). I refer to the dense statistical connectome model as the "dense statistical connectome".

Contribution: I analyzed predictions of the dense statistical connectome at the subcellular, cellular, and network level (see Section 2.3). In collaboration with the group of Prof. Dr. Hans-Christian Hege (Zuse Institute Berlin), the dense statistical connectome and my analysis tools were incorporated into two interactive web applications (see Section 2.5 and 2.6).

(4) Probability distributions of observables: The dense statistical connectome allows deriving the probability distributions of "observables" given the structural estimates provided by the dense structural model (1) and the synapse formation strategy (2). An observable refers to a wiring property that, in an ideal case, is experimentally accessible (e.g., the connection probability between two neuron populations is observable by performing multiple *in vivo* or *in vitro* paired

recordings). The predicted probability distributions of wiring properties can then be compared to empirical wiring properties, obtained, for example, by anatomical reconstructions or inferred from functional recordings.

Contribution: To validate the predictions, I screened the scientific literature for empirical wiring properties and replicated the experimental conditions of each measurement in the dense statistical connectome. I systematically compared empirical and predicted wiring properties at the subcellular, cellular, and network level and assessed the consistency between the model's predictions and the observations (see Section 2.4).

Viewed from a network perspective, the dense structural model provides the distribution of the nodes (i.e., pre- and postsynaptic structures). The synapse formation strategy provides the probability of a pair of nodes being connected by an edge (i.e., synaptically connected). The dense statistical connectome describes all possible network configurations, including their probabilities, when applying the synapse formation strategy to all the nodes of the dense structural model.

The thesis is similarly organized as the presented strategy. The thesis starts with the dense structural model, followed by the dense statistical connectome, and a systematic comparison between empirical and predicted wiring properties. The thesis ends with a description of two web applications that allow the scientific community and public to use both models, and in the future to explore the impact of not only of the herein formulated synapse formation strategy on wiring but of any strategy that can be formulated in mathematical terms.

1.4.2 The rat vibrissal system as a model system

Due to the wealth of structural and connectivity data available, the vibrissal part of the primary somatosensory cortex (vS1 or barrel cortex) is an ideal starting point to (i) create an anatomically detailed model thereof and (ii) to predict wiring diagrams solely shaped by the structural composition of the neuropil and compare those to empirical data. vS1 is part of the vibrissal system of the rodent. Its experimental accessibility and anatomically well-defined somatotopy made it one of the most widely studied and well described mammalian model systems in neuroscience

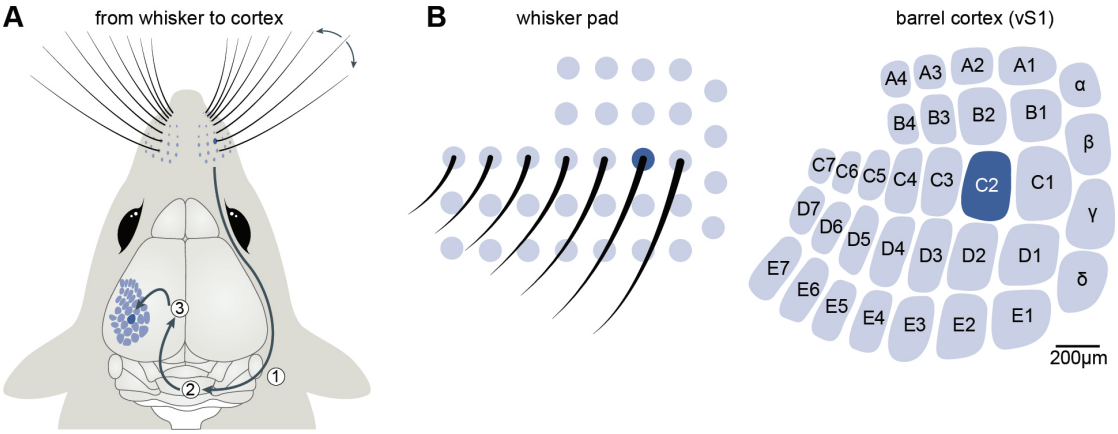

Figure 1.5. Rat vibrissal system. (A) Signal evoked by the deflection of a facial whisker is relayed via the brain stem (1) to the thalamus (2) and then to the vibrissal part of the primary somatosensory cortex (vS1) (3). (B) The spatial arrangement of the whisker follicles (only C-row whiskers are shown) at the rat's snout is conserved as anatomical structures termed "barrels" in layer 4 of vS1 (i.e., somatotopic arrangement). Barrels and whiskers are arranged into rows A to E and numbered arcs. C2 whisker follicle and barrel highlighted in dark blue. Figure adapted from C. C. H. Petersen (2019).

(for review, see C. C. Petersen, 2007; Feldmeyer et al., 2013). Many rodents, such as rats and mice, are nocturnal animals that explore their environment using the whiskers at their snout. Sensory information, such as the deflection of a whisker, is relayed from the whisker follicle in a somatotopic fashion through the brain stem and the thalamus before reaching the cortex (see Figure 1.5).

At each stage, neurons are predominantly somatotopically arranged into neuron-dense structures, called "barrelettes" in the brain stem, "barreloids" in the ventral posterior medial nucleus of the thalamus (VPM), and "barrels" in layer 4 of vS1. Barrels are organized in rows and arcs, resembling the spatial arrangement of the facial whiskers at the animal's snout (Van der Loos & Woolsey, 1973). Barrels are separated by neuron-sparse areas called septa. Neurons located within the barrel as well as neurons above and below respond most strongly to deflections of the somatotopically aligned whisker. This kind of discrete functional unit is common across sensory modalities and species and referred to as "cortical column" (Mountcastle, 1957; Hubel & Wiesel, 1959). In vS1, the anatomical correlate of a cortical column is called a "barrel column" (i.e., a vertical extrapolation of the

barrel to the pial and white matter surface) and spans all six neocortical layers.

As previously described in detail in Section 1.1, the development of vS1 like that of other neocortical areas is the result of various genetically induced developmental programs (for review, see Erzurumlu & Gaspar, 2012; Feldmeyer et al., 2013; Yang et al., 2018). By the end of the second postnatal week, active whisking has started and the animal opens its eyes. By then, neuron morphologies are established but are still subject to fine-scale remodeling over the next two weeks. Synapses between neurons are added, eliminated, and refined until development is completed by the fourth postnatal week. Past the first month, structural changes to morphologies and synapses are limited to fine-scale readjustments (for more details, see Section 1.2).

The developed rat vS1 contains about half a million neurons of which 87% are excitatory and the remaining 13% are inhibitory (Meyer et al., 2013). Excitatory neurons are well characterized and can be divided into ten functionally and morphologically distinct cell types (Oberlaender, de Kock, et al., 2012; Narayanan et al., 2015) (see Figure 1.2). Inhibitory neurons in vS1, like in other areas, are highly diverse with respect to morphological, functional, molecular, and genetic properties (Cadwell et al., 2016; Feldmeyer, Qi, Emmenegger, & Staiger, 2018), hampering classification approaches (Ascoli et al., 2008; DeFelipe et al., 2013; Staiger et al., 2015; Gouwens et al., 2019). Nonetheless, morphological and functional properties of inhibitory neurons grouped by their laminar position in rat vS1 have been characterized in various studies (e.g., Helmstaedter, Sakmann, & Feldmeyer, 2009a, 2009b, 2009c; Koelbl, Helmstaedter, Lübke, & Feldmeyer, 2013; Egger et al., 2015; Arzt, Sakmann, & Meyer, 2018; Emmenegger, Qi, Wang, & Feldmeyer, 2018).

2 | Methods

I used custom-written routines in `C++` using the Insight Toolkit[1] and Visualization Toolkit[2] libraries, Python, and Matlab software (Mathworks, Natick, MA, USA) for analysis. The TreesToolbox[3] (Cuntz, Forstner, Borst, & Häusser, 2010), an open-source software package, was used to process the neuron morphologies in Matlab. Amira software (FEI) was used for visualization (Stalling, Westerhoff, & Hege, 2005). The *NeuroNet* software package (Egger et al., 2014; Lang, Dercksen, Sakmann, & Oberlaender, 2011) for Amira was used to generate the dense structural model and the dense statistical connectome. The color maps for visualizing the distribution of somata, dendrites, and axons in some Figures are based on `www.ColorBrewer2.org`, by Cynthia A. Brewer, Penn State. Box plots were generated with the Matlab built-in `boxplot` where the bottom and top of the box represent the 25th and 75th percentiles, and the line within the box the median. The lines extend to the adjacent values. Outliers are all values more than 1.5 times the interquartile range away from the top or bottom of the box.

2.1 Model of dense structural composition of rat barrel cortex

I generated an anatomically detailed digital model of the dense structural composition (referred to as dense structural model) of the vibrissal part of the primary somatosensory cortex (vS1 or barrel cortex) of the rat post critical periods of

[1] `https://itk.org/`
[2] `https://vtk.org/`
[3] `https://www.treestoolbox.org/`

neuron and neurite development. The aim of the dense structural model was to capture the structural principles characteristic for vS1 and to provide realistic and robust structural estimates to constrain synapse formation between branches of axons and dendrites and neuron somata. In this Section, I describe how the dense structural model was generated and constraint (see Section 2.1.1), how I integrated the IN morphologies into the dense structural model (see Section 2.1.2), and how I extracted and analyzed structural estimates predicted by the dense structural model and assessed their robustness (see Section 2.1.3). To test how realistic the structural estimates were, I compared observable (i.e., experimentally accessible) structural estimates to available empirical data, that was not used to constrain the model.

2.1.1 Generation of dense structural model

Note: The following Section is in part adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020).

The dense structural model was generated by using *NeuroNet*, a custom-designed extension package for Amira software, and was described in detail previously (Egger et al., 2014). Note, the dense structural model does not provide estimates for synaptic connectivity, and the model is ignorant of glial cells. In the following, I summarize how the dense structural model is generated and constraint by empirical data (see Figures 2.1 and 2.2). The model incorporates measurements of the following four components: (1) geometrical, (2) cellular, (3) morphological, and (4) pre- and postsynaptic. All measurements were from rats of the same strain and developmental stage and housed in standard laboratory environments. The integration of IN morphologies into the dense structural model is described in detail in Section 2.1.2. The analysis of the robustness of the empirical constraints is described in Section 2.1.3.

(1) Geometrical component: *NeuroNet* generated a 3D model of the average geometry of rat vS1 (Egger et al., 2012), referred to as the geometrical model. The geometrical model was based on precise reconstructions of the location, orienta-

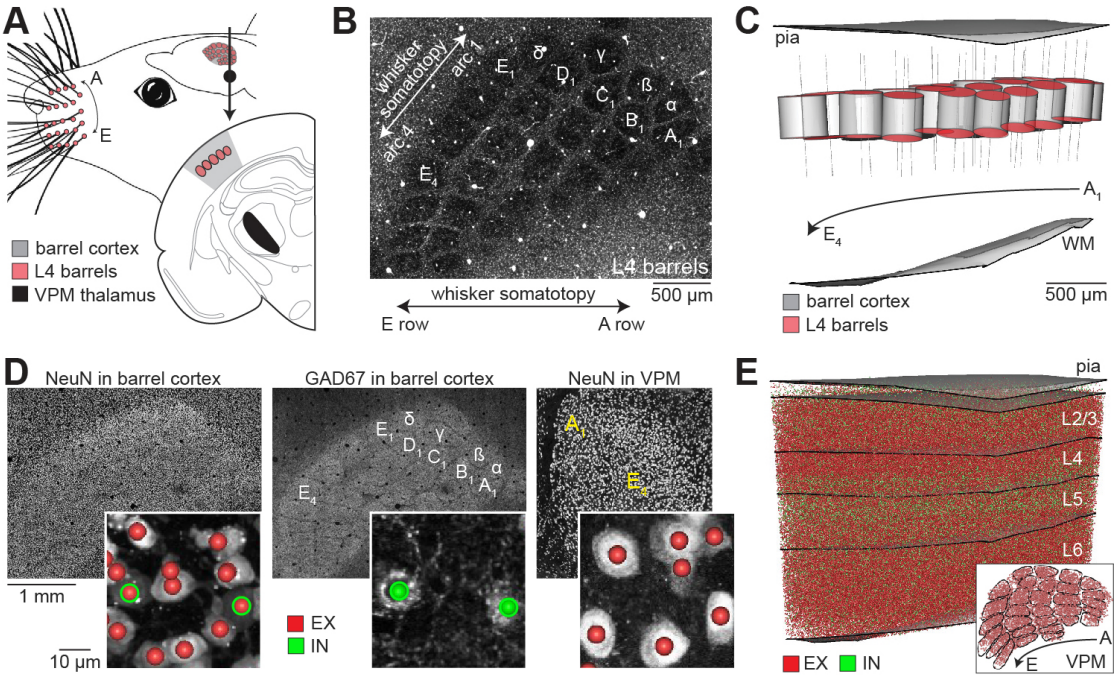

Figure 2.1. Laminar, columnar, and cellular organization of rat vS1. (A) Schematic of rat whisker system (cf. Figure 1.5 for details on rat whisker system). (B) Somatotopic whisker map of barrels in layer 4 (L4), visualized by cytochrome oxidase. Panel adapted from Egger et al. (2012). (C) Reconstructions of the location, orientation, and diameter of all 24 cortical barrel columns and pial and white matter (WM) surfaces across 12 rats were used to generate a digital 3D model of the average barrel map, pial, and white matter of rat vS1. (D) Excitatory (EX) and inhibitory (IN) somata in vS1 and in the ventral posterior medial nucleus (VPM) of the thalamus (Panel D adapted from Meyer et al., 2013) were visualized by NeuN and GAD67 labeling. Zoom-ins show examples of automated neuron detection. (E) 3D neuron counts in vS1 across four rats and in VPM thalamus across three rats were used to generate a digital 3D model of the average cellular and laminar organization of rat vS1 and VPM thalamus. Figure adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020).

tion, and diameter of cortical barrel columns that represent the 24 major facial whiskers (A1-E4 and α-δ), and of the pial and white matter surfaces above and below vS1, respectively (see Figure 2.1BC). These anatomical reference structures were determined in twelve rats (Wistar, male/female, postnatal day 28), which revealed that cortex geometry and somatotopy vary by $\pm$50 µm across animals. This variability of vS1 geometry defines the spatial resolution limit of the dense structural model (i.e., cubes of 50 µm edge length) (see also Section 3.1.1).

(2) Cellular component: *NeuroNet* combined the geometrical model with empirically determined 3D distributions of excitatory and IN somata in rat vS1, and in the ventral posterior medial nucleus (VPM) of the thalamus (Meyer et al., 2013) (see Figure 2.1D and E). Soma distributions were measured with respect to the anatomical reference structures in VPM for three rats and vS1 for four rats, respectively (Wistar, male, postnatal day 28), converted into density distributions at the resolution of the geometrical model (i.e., 50 µm cubes), and averaged. The average number of excitatory and IN somata in the volume of vS1 that represents the 24 major facial whiskers (6.77 mm^3) was 460,981 $\pm$ 51,583 and 68,734 $\pm$ 14,160 (n = 4 rats). The respective average number of excitatory somata in the VPM was 6,225 $\pm$ 173. Given the observed average 3D densities of excitatory and inhibitory somata, neuron somata were randomly placed within each 50 µm cube in the dense structural model. The dense structural model used here comprised 477,537 excitatory and 69,810 inhibitory neurons in vS1, and 6,225 in the VPM.

(3) Morphological component: *NeuroNet* replaced each excitatory soma in the vS1 model by a soma/dendrite morphology that was reconstructed from an *in vivo* labeled neuron (Wistar, male/female, postnatal day 25-45; n = 153 neurons; Narayanan et al., 2015), whose soma depth location was within 50 µm of its model location (see Figure 2.2). The sample of morphologies represents $\sim$1% of the excitatory population of one barrel column. Based on its soma depth location and 20 dendrite features, each excitatory neuron was objectively assigned to one of the ten major morphological cell types of the neocortex (Narayanan et al., 2015): pyramids in layers 2, 3, and 4 (L2PY, L3PY, and L4PY), spiny-stellates (L4ss) and star-pyramids (L4sp) in layer 4, slender-tufted intratelencephalic (L5IT) and thick-

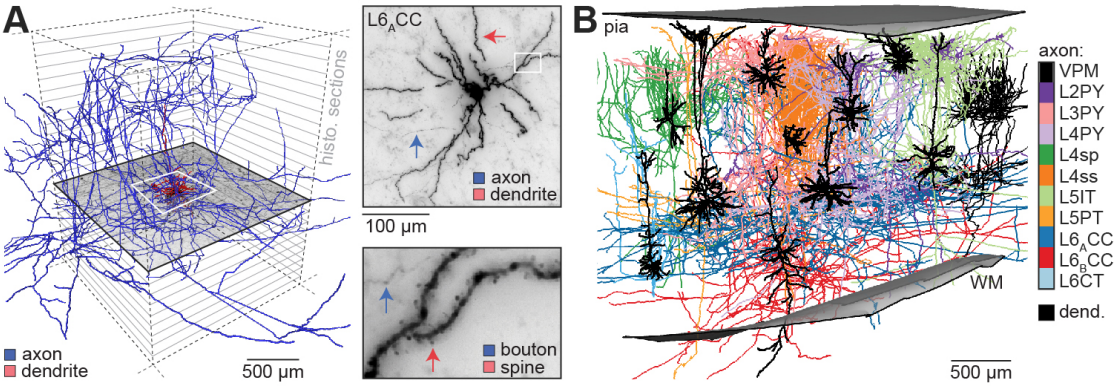

Figure 2.2. Excitatory dendrite and axon morphologies in rat vS1. (A) Reconstruction of an exemplary neuron that was labeled *in vivo* with biocytin via cell-attached recording. Dendrites (red) and axons (blue) were reconstructed across consecutive histological sections ranging from the pia to the WM. Image shows projection of optical sections from the histological section that contains the soma. Zoom-in shows that the labeling quality is sufficient to identify axonal boutons and dendritic spines. (B) 153 dendrites and 88 axons were labeled *in vivo*, reconstructed (Narayanan et al., 2015), and used to populate the model (Egger et al., 2014). Exemplary dendrite (black) and axon morphologies of the major axodendritic cell types of rat vS1 plus VPM axon. Figure adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020).

tufted pyramidal tract neurons (L5PT) in layer 5, corticocortical neurons at the border between layers 5 and 6 ($L6_ACC$), and polymorphic corticocortical ($L6_BCC$) and corticothalamic neurons (L6CT) in layer 6. Matching the resultant numbers of excitatory neurons per barrel column and cell type within the vS1 model, axon morphologies ($n = 74$) from *in vivo* labeled neurons were placed, such that their respective soma positions and axon orientations remained unchanged (Egger et al., 2014). Similarly, the intracortical part of *in vivo* labeled VPM axon morphologies (Wistar, male/female, postnatal day 25-170, $n = 14$ neurons; Oberlaender, de Kock, et al., 2012; Oberlaender, Ramirez, & Bruno, 2012) were placed to match the number of neurons per VPM barreloid (Meyer et al., 2013). The sample of VPM axon morphologies represents ∼5% of the neurons located within a barreloid in VPM thalamus. For the cell type-specific numbers of somata in the model please refer to Table 3.1. This procedure of duplicating and placing single neuron morphologies in the dense structural model to match the total number of distributed

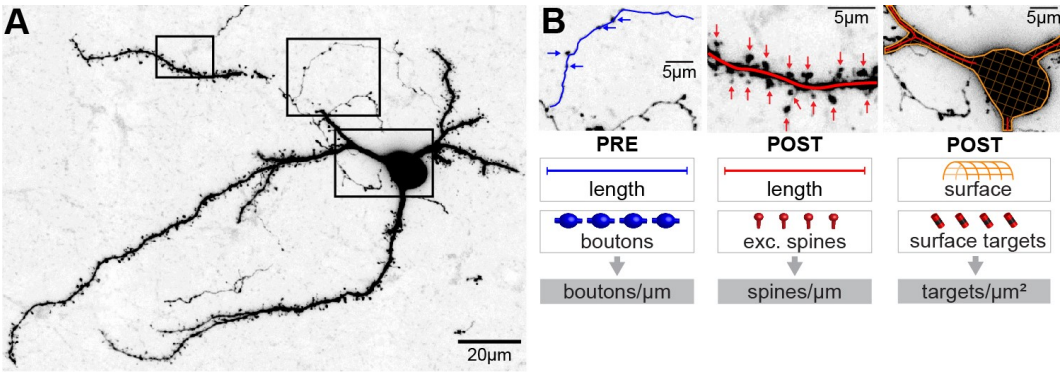

Figure 2.3. Pre- and postsynaptic structures along axons and dendrites. (A) z-projection of a 50 µm thick section containing the soma, dendrites and axon branches of a L4ss neuron. (B) From left to right: Close-up view of axon branch including boutons (indicated by arrows), dendrite branch including spines (indicated by arrows), and neuron somata with digital reconstruction. Left: Bouton densities were empirically determined and used to map boutons onto axons in the model by assuming proportionality between axon length and the number of boutons (Narayanan et al., 2015). Center: Assuming proportionality between dendrite length and number of spines, the spine densities were derived and mapped onto the excitatory dendrites in the model. Right: Postsynaptic structures that are not spines (i.e., along INs or along excitatory neurons if targeted by INs) were derived by assuming proportionality between postsynaptic structures and the surface area of dendrites and somata. For more details, see main text. Figure adapted from Egger et al. (2014)

.

neuron somata is denoted by the term "upscaling". The upscaling of single neuron morphologies to the total numbers of neurons provided estimates for the amount of dendrites and axon that each neuron contributes to every 50 µm cube of rat vS1 (i.e., for animals around postnatal day 28), depending on its soma location, somatotopic position (i.e., closest barrel column or barreloid), cell type, and particular morphology. To avoid underestimation of dendrite densities at the borders of the vS1 model, the model volume was extended by a sufficiently large surrounding volume, which was excluded from the analyses. The procedure of incorporating IN morphologies of L2-6 into the dense structural model is described in detail in the following Section 2.1.2.

Table 2.1. Cell type-specific meta connectivity. The number of postsynaptic structures is assumed to be proportional to either the dendrite length (if the presynaptic neuron is excitatory) or dendrite and soma surface areas (in all other cases).

Presynaptic type	Postsynaptic type	Number of postsynaptic structures proportional to
excitatory	excitatory	dendrite length
excitatory	inhibitory	dendrite and soma surface area
inhibitory	excitatory	dendrite and soma surface area
inhibitory	inhibitory	dendrite and soma surface area

(4) Pre- and postsynaptic component: In the dense structural model, axons and dendrites are not represented by their respective trajectories. Instead, morphologies are transformed into densities per 50 µm cube (i.e., reflecting the spatial resolution limit for comparing structural data across animals Egger et al., 2012) of pre- and postsynaptic structures, respectively. The densities of presynaptic structures (i.e., axonal boutons) were derived by multiplying the axon length that each neuron contributes to a particular 50 µm cube with the number of boutons per length, as measured for the respective cell type and layer (Narayanan et al., 2015) (see Table 2.2 and Figures 2.2A and 2.3; for IN bouton densities, see Section 2.1.2). The densities of the postsynaptic structures depend on the type of the presynaptic neuron (i.e., excitatory vs. inhibitory; see Table 2.1) (Egger, 2015). Presynaptic structures of excitatory neurons are assumed to form synapses exclusively with postsynaptic structures along dendrites of excitatory neurons. Excitatory synapses do not target excitatory soma in line with empirical data (Schoonover et al., 2014). The number of postsynaptic structures along excitatory dendrites (i.e., spines) was derived by assuming that spine densities are proportional to dendrite length. In all other cases (i.e., pre- and/or postsynaptic structures of IN) presynaptic structures are assumed to form synapses with postsynaptic structures along both dendrites and somata as observed empirically (Staiger, Zilles, & Freund, 1996; Markram et al., 2004). The respective number of postsynaptic structures along dendrites and somata was derived by assuming proportionality to their respective surface areas (for more details, see Section 2.1.2). The total number of postsynaptic structures in the vS1 model was scaled to match the predicted number of axonal boutons along

Table 2.2. Cell type-specific bouton densities.

Cell type	Bouton density [1/µm]		
	infragranular	granular	supragranular
L2PY	0.31	0.31	0.36
L3PY	0.23	0.25	0.25
L4PY	0.18	0.25	0.22
L4sp	0.21	0.28	0.24
L4ss	0.24	0.27	0.25
L5IT	0.19	0.24	0.28
L5PT	0.20	0.25	0.19
$L6_ACC$	0.20	0.27	0.26
$L6_BCC$	0.29	0.26	0.23
L6CT	0.27	0.27	0.27
VPM	0.28	0.31	0.34
L1IN	0.39	0.39*	0.39*
L2-L6IN	0.20	0.20	0.20

(*) assumed density, no horizontal projection axon branches could be found (Egger, 2015; Egger et al., 2015)

the cortical depth. The derived density of postsynaptic structures for excitatory neurons ranged from 1.04 to 1.68 spines per µm dendrite length, depending on the subcellular compartment and cell type (see Table 2.3). These derived densities are consistent with empirical spine density measurements (Larkman, 1991; Kawaguchi, Karube, & Kubota, 2006). The respective numbers for IN are described in detail in Section 2.1.2.

2.1.2 Integration of inhibitory neuron morphologies

In collaboration with my colleague Dr. Robert Egger (Max Planck Institute for Biological Cybernetics, Tübingen), I incorporated 204 *in vitro* labeled and reconstructed IN morphologies of L2-6 into the dense structural model of vS1. Morphologies of *in vivo* labeled L1 INs were incorporated by Dr. Robert Egger previously (Egger et al., 2015; Egger, 2015). In this Section, I first describe the integration of IN morphologies into the model, and then how the densities of pre- and postsynaptic structures required to constrain synapses from and/or onto INs

Table 2.3. Cell type-specific spine densities.

Cell type	Spine density [1/µm]	
	apical	basal
L2PY	1.68	1.68
L3PY	1.68	1.68
L4PY	1.68	1.17
L4sp	1.17	1.17
L4ss	1.17	1.17
L5IT	1.68	1.04
L5PT	1.68	1.04
$L6_ACC$	1.04	1.04
$L6_BCC$	1.04	1.04
L6CT	1.04	1.04

were estimated.

Morphological component of inhibitory neurons

The reconstructed IN morphologies are from five different data sets of rat vS1 (Wistar, male/female, postnatal day 18-35) spanning the entire cortical sheet. The sample represents ∼8% of the inhibitory population of one barrel column:

- 48 axon/dendrite plus 3 axon reconstructions of *in vitro* labeled INs in L2/3 (Helmstaedter et al., 2009a, 2009b, 2009c).
- 40 axon/dendrite reconstructions of *in vitro* labeled fast-spiking INs in L4 (Koelbl et al., 2013).
- 68 axon/dendrite reconstructions of *in vitro* labeled INs in L6 (Arzt et al., 2018).
- 45 axon/dendrite reconstructions of *in vitro* labeled INs in L5 (Dr. Marlene Arzt, Dr. Mike Hemberger, Dr. Hanno S. Meyer, and Prof. Dr. Bert Sakmann at Max Planck Florida Institute for Neuroscience, Florida, USA, and Max Planck Institute for Medical Research, Heidelberg; unpublished).

- 10 axon/dendrite reconstructions of *in vivo* labeled INs in L1 (Egger et al., 2015).

I incorporated "curated" IN morphologies of the *in vitro* labeled morphologies (i.e., excluding L1 INs) into the dense structural model of rat vS1: Due to tissue shrinkage caused by fixation and mounting on microscope slides, all IN morphologies were scaled along the cutting dimension to the original tissue thickness, i.e., 300 µm or 350 µm, as defined by the reported vibratome settings (Koelbl et al., 2013; Helmstaedter et al., 2009a; Arzt et al., 2018). The scaled IN reconstructions were registered into the geometrical model (Egger et al., 2012). Registration was performed by matching the corresponding available anatomical landmarks (e.g., pial surface, WM surface, and L4 barrels) of an IN reconstruction and the D2 column in the geometrical model and by calculating an optimal transformation for the IN morphology that corrected for varying slicing orientations. Each IN morphology was placed at its location with a precision of 50 µm (Egger et al., 2012). INs in L2/3 lacked pial and WM surface contours and were therefore placed at the center of L2/3.

I made use of the classification routine I had developed in my master thesis (Udvary, 2014) and refined during my time as a doctoral student (see also Udvary et al., 2016). The classification is based on the neuron's soma depth location and eight morphological and topological features of its axon projection pattern. INs of L2-6 were objectively classified into five different projection types (see Figure 2.4): local projecting INs ($n = 103$) and four nonlocal projecting IN types, "Supragranular-descending" (SupraDesc; $n = 16$), "Granular-Symmetric" (GranSym; $n = 11$), "Granular-Ascending" (GranAsc; $n = 9$), "Infragranular-ascending" (InfraAsc; $n = 65$). Their naming reflects the laminar location of the neurons' somata and direction in which their axon projects to. INs of the L1 data set (Egger et al., 2015) were manually assigned as a separate cell type.

Before incorporating the INs into the dense structural model, the *in vitro* labeled INs were curated to compensate for the truncation introduced by the slicing procedure. I made use of a tool that I had developed during a laboratory rotation in Dr. Marcel Oberlaender's group in 2013 at the Max Planck Institute for Biological Cybernetics in Tübingen (Udvary, 2013). The tool extrapolated the truncated

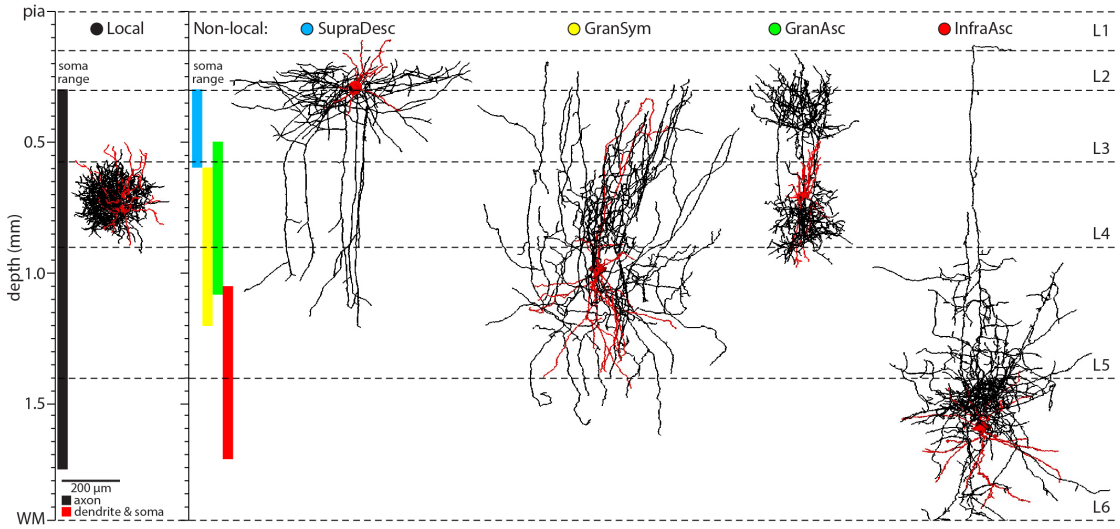

Figure 2.4. Vertical axon projection types of INs in rat vS1. Exemplary morphologies of five IN axon projection types. The colored vertical bars denote the respective depth range of somata assigned to each projection type. From left to right: local projecting INs (black) and four nonlocal projecting IN types, "Supragranular-descending" (SupraDesc; blue), "Granular-Symmetric" (GranSym; yellow), "Granular-Ascending" (GranAsc; green), "Infragranular-ascending" (InfraAsc; red). Dendrites and axons are shown in red and black, respectively. Dashed lines represent layer borders.

IN morphologies by assuming that the neuron morphologies were radial symmetric in the horizontal (i.e., tangential) plane.

The local projecting type was subdivided into six subtypes, spread equally along the cortical axis to preserve depth-dependent morphological features such as an increase in the horizontal and vertical axon bounding box with increasing cortical depth (Udvary, 2014). IN morphologies were incorporated into the dense structural model the same way as the excitatory neurons: each inhibitory soma in the vS1 model is replaced by either a registered *in vivo* soma/dendrite morphology (INs of L1; Egger et al. (2015)) or by a registered and curated soma/dendrite morphology (i.e., INs of layers 2 to 6), whose soma depth location is within 50 µm of its model location. IN axon morphologies are placed by matching the number of placed soma/dendrites for each IN projection type (Egger et al., 2014). For each inhibitory soma, an IN axon morphology is chosen at random from all IN morphologies that are of the same projection type as the projection type assigned

to the soma, and that are registered to the same barrel column. Unlike dendrite morphologies, axon morphologies are not transformed to the new soma locations to prevent that rotation and translation leads to a loss of location-specific projection patterns (e.g., that translations would result in inappropriate innervation of septal areas). A refined incorporation of INs into the dense structural model is currently an ongoing effort in the group of Dr. Marcel Oberlaender (see Discussion Section 4.8 for details).

Pre- and postsynaptic components of inhibitory neurons

To constrain synapse formation involving INs, I estimated their densities of pre- and postsynaptic structures (see Tables 2.1 and 2.2).

Presynaptic densities of inhibitory neurons: I determined the bouton density from samples of branches of five INs (one in L2/3 and four in L6) whose boutons were reconstructed. The bouton density ranged from 0.16 to 0.24 (0.20 ± 0.03) per µm, in line with previously reported values (Karube, Kubota, & Kawaguchi, 2004; Markram et al., 2004; Wang, Gupta, Toledo-Rodriguez, Wu, & Markram, 2002). The bouton density in the dense structural was set to 0.2 per µm for INs in L2-6 (see Table 2.2) assuming that bouton densities of INs in L2-6 are independent of the IN type and laminar location (but see Markram et al., 2004). Bouton densities of L1 INs were determined empirically by Egger et al. (2015).

The number of postsynaptic structures along INs was assumed to be proportional to their dendrite and soma surface areas. However, not all excitatory and inhibitory neuron morphologies had reconstructed radii along their dendrites. Therefore, Egger (2015) developed a database of cell type- and soma-distance-dependent dendrite branch radii for the dendrites with reconstructed radii. The dendrites without reconstructed radii were assigned radii values by an exponential function with parameter values from the cell type- and soma-distance-dependent database (for more details, see Egger, 2015).

Postsynaptic densities of inhibitory neurons if presynaptic neuron is excitatory: Inhibitory neurons are usually aspiny and empirical data shows

that excitatory synapses are found on both inhibitory dendrites and somata (e.g., White, Benshalom, & Hersch, 1984; Keller & White, 1987; Ahmed, Anderson, Martin, & Nelson, 1997; Staiger et al., 1996). Therefore, the number of postsynaptic structures along inhibitory dendrites and somata that could be contacted by excitatory boutons was derived by assuming proportionality to their respective surface areas (described in detail in Egger, 2015). The total number of postsynaptic structures of excitatory and INs in the dense structural model was scaled to match the predicted number of excitatory boutons along the cortical depth. The derived density of postsynaptic structures onto inhibitory dendrite and soma surface areas was assumed not to vary between cell types and subcellular structures. The derived density of postsynaptic structures for INs was 0.74 per μm^2 of dendrite or soma surface area (Egger, 2015). The derived densities were consistent with empirical synapse density measurements on inhibitory somata (Keller & White, 1987; Ahmed et al., 1997).

Postsynaptic densities of inhibitory and excitatory neurons if presynaptic neuron is inhibitory: Inhibitory synapses are found on dendrites (mostly on dendritic shafts) and somata of both excitatory and inhibitory neurons (e.g., White & Rock, 1981; White et al., 1984; Keller & White, 1987; Markram et al., 2004). Therefore, the number of postsynaptic structures along inhibitory and excitatory dendrites and somata that could be contacted by inhibitory boutons was derived by assuming proportionality to their respective surface areas (described in detail in Egger, 2015). The total number of postsynaptic structures of both excitatory and inhibitory neurons in the dense structural model was scaled to match the predicted number of inhibitory boutons along the cortical depth. The derived density of postsynaptic structures for both excitatory and INs was 0.06 per μm^2 of dendrite or soma surface, assuming that the density does not vary between cell types and subcellular structures. The derived densities were consistent with several empirical synapse density measurements on inhibitory and excitatory somata (Keller & White, 1987; White et al., 1984; Ahmed et al., 1997). However, there is empirical evidence that the derived density might be underestimated and vary between cell types and subcellular compartments (Ahmed et al., 1997; Shigematsu et al., 2016; Hioki et al., 2013; Kameda et al., 2012).

2.1.3 Analysis of structural estimates

In this Section, I describe first how I tested the robustness of the empirical constraints of the dense structural model as well as of the resultant structural estimates. I then describe how I analyzed the structural estimates and compared observable estimates (i.e., experimentally accessible ones) to available empirical data. While the first part of this Section focuses exclusively on excitatory neurons, in the last part a limited analysis of INs in the dense structural model is presented.

Robustness of dense structural model

To test the robustness of the dense structural model, I first investigated the interanimal variability of the empirical constraints of the dense structural model. Second, I investigated the robustness of the sparse sample of axon and dendrite morphologies. Third, I investigated the robustness of the dense structural model when upscaling the sparse sample of neuron morphologies to all neurons in the model. *Note:* The following paragraphs are in part adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020).

Variability of empirical constraints of model: I calculated the standard deviations (SDs) and coefficient of variations (CVs) of the empirical constraints per barrel column in vS1 and per barreloid in VPM across animals (Egger et al., 2012; Meyer et al., 2013). The precision of the rat vS1 barrel map was determined as the SD of the 2D positions of the barrel top and bottom along the arc and row across 12 rats when registering to the geometrical model of vS1 (Egger et al., 2012). The layer precision was determined as the SD of cortical depth of the layer borders between L1 and L2/3, L2/3 and L4, L4 and L5, and L5 and L6 across four rats (Meyer et al., 2013).

Robustness of sample of axon and dendrite morphologies: First, I assessed how the cell type-specific volume innervated by dendrites changed with the number of morphologies per cell type (see Table 2.4 for number of morphologies per cell type). The dendrite morphologies of each excitatory cell type were regis-

Table 2.4. Cell type-specific number of morphologies used for upscaling in the vS1 model. Morphological sample published in Narayanan et al. (2015).

Cell type	Number of	
	axon morphologies	dendrite morphologies
VPM	14	n/a
L2PY	9	16
L3PY	14	30
L4PY	4	7
L4sp	5	15
L4ss	9	22
L5IT	5	18
L5PT	7	16
$L6_ACC$	6	11
$L6_BCC$	4	5
L6CT	11	13

tered to the D2 barrel column (Egger et al., 2012), aligned by their lateral soma position, and their dendrite innervation volume per 50 µm cube was calculated. For each cell type, I determined all possible combinations of subsamples of dendrite morphologies. If there were more than 500 possible combinations for a given subsample size of morphologies, a random sample of 500 combinations was used. For each combination of morphologies, the change of their total innervation volume to the volume innervated by all morphologies (referred to as maximal dendrite innervation) was calculated, i.e., one minus the ratio between the number of cubes innervated by the respective combination of morphologies and those innervated by all morphologies.

Second, I assessed how the cell type-specific length densities per 50 µm cube changed with the number of morphologies per cell type. For each combination of morphologies, their respective dendrite length contribution per cube was determined. For each subsample size of morphologies, the mean, SD, and CV of the dendrite length per cube across all possible combinations of morphologies was calculated. I repeated the same analysis for the axon morphologies of each cell type but without aligning the axon morphologies by their somata.

Robustness of model predictions due to upscaling of morphological sample: I assessed how the neurite density and diversity changed with the number of morphologies per cell type used for upscaling. I created multiple dense structural models of vS1 where subsamples of excitatory morphologies were used for upscaling, i.e., vS1 models based on only one morphology per cell type, two morphologies per cell type, and so on. Specifically, I determined all possible combinations of subsamples of morphologies per cell type. If there were more than 500 possible combinations for a given number of morphologies per cell type, a random sample of 500 combinations was used. Upscaled versions of morphologies that were not in the selected combination were removed from the vS1 model. To compensate for the removed morphologies, the remaining morphologies in the vS1 model were duplicated to match the overall number of neurons in the model in a cell type and column-specific manner. If the number of morphologies per cell type equaled or exceeded the number of morphologies of a particular cell type, no morphologies of the respective cell type were removed (e.g., in case of 6 morphologies per cell type all axon morphologies of L4sp, L4PY, L5IT, $L6_ACC$, and $L6_BCC$ are used, while for all other cell types subsamples are used). For a subvolume of 50 µm cubes across L2-6 of the C2 barrel column, I determined for each cube the following summary statistics across all combinations: mean, SD, and CV of the density and diversity of axons and dendrites per number of morphologies per cell type used for upscaling. Diversity is defined as the number of cell types that contribute structures to a cube.

Comparison with empirical data

I investigated how realistic the estimates of the dense structural model are. To do so, I compared the model's estimates to available empirical data that was not used to constrain the model. *Note:* The following paragraphs are in part adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020).

Axon and dendrite branchlets: The ratio between axon and dendrite branchlets in a small subvolume of mouse somatosensory cortex was reported in one electron microscopy study (Kasthuri et al., 2015). I compared this observation to the model's estimates: The number of excitatory axon and dendrite branchlets within

each 50 µm cube across L2-6 of the C2 barrel column in the dense structural model were converted into histograms with bin widths of 1000 and 100 branchlets, respectively. Axon/dendrite branchlet ratios were calculated per 50 µm cube. The number of axon and dendrite branchlets within a particular 50 µm cube of the model reflect estimates from upscaling *in vivo* labeled neuron morphologies to the total numbers of neurons in the dense structural model (see Section 2.1.1).

Synapse densities across layers: I compared empirical layer-specific excitatory synapse densities in small subvolumes of rat and mouse somatosensory cortex (Santuy et al., 2018; Motta et al., 2019; Kasthuri et al., 2015) to the model's estimates. To do so, the synapse densities within the dense structural model were determined by calculating the number of excitatory boutons per 50 µm cube for the D2 barrel column, and grouping the 50 µm cubes by their respective laminar positions within L1-6. Layer borders were defined by the excitatory somata density profile along the cortical depth, as reported previously (Meyer et al., 2013). The border between L2 and L3 was defined by the inhibitory somata density profile along the cortical depth, as reported previously (Meyer et al., 2011; Narayanan et al., 2017). To test whether the empirical and predicted synapse densities were different, I performed a two-sample Kolmogorov-Smirnov test with $\alpha = 0.01$ using the Matlab built-in function `kstest2` for each layer. The analysis was repeated for inhibitory boutons and synapses, respectively.

Definition of cortical layers in rat vS1

Note: In collaboration with my colleague Dr. Rajeevan T. Narayanan (Max Planck Institute for Biological Cybernetics, Tübingen, and research center caesar, Bonn), I analyzed the structural estimates of excitatory neurons by the model with respect to the layer organization of rat vS1. The following Sections are adapted from Narayanan et al. (2017).

Empirical properties of the structural composition of the neocortex or the synaptic connectivity between neurons are often reported with respect to layers. I used the column-specific layer border definitions as determined by the measured number

Table 2.5. Layer borders as distances from the pial surface for the D2 and C2 barrel columns as reported by Meyer et al. (2013).

Layer border	D2 barrel column [µm]	C2 barrel column [µm]
L1-L2/3	157 ± 16	157 ± 14
L2/3-L4	575 ± 57	576 ± 52
L4-L5	900 ± 50	855 ± 25
L5-L6	1411 ± 28	1349 ± 32
L6-WM	1973 ± 44	1892 ± 64

and 3D distribution of excitatory and inhibitory somata in rat vS1 (see Table 2.5; Meyer et al., 2010, 2013).

Two caveats about these layer border values should be noted. First, it was not possible to define the border between L2 and L3 based on the overall soma density distribution. In the results reported by Meyer et al. (2013), the vertical distribution of inhibitory somata was, in general, different compared to the layer-defining distribution of excitatory neuron somata. Inhibitory somata were densest in upper L2/3 and at the border between L4 and L5 (Meyer et al., 2011). The distribution of inhibitory somata thus provided a quantitative basis for an anatomical separation between L2 and L3. The average border between L2 and L3 was hence calculated as 296 ± 30 µm. Note that for the analysis of predicted connectivity (see Section 2.3), layer borders between L2 and L3 and sublayer border in L5 and L6 (i.e., between L5A and L5B, and L6A and L6B) were defined as the center of layer borders reported by Meyer et al. (2013) (see Table 2.5).

The second caveat is that layer borders deviate between barrel columns. The differences do not reflect a linear scaling with cortical thickness. Instead, Meyer et al. (2013) found that the diameter of cortical barrel columns compensates for differences in cortical thickness across vS1, resulting in largely the same volume for barrel columns that represent whiskers of the same row along the animal's snout (Egger et al., 2012). The analysis was mostly restricted to either the D2 or C2 barrel column which represents an average across all barrel columns with respect to their geometrical properties.

Layer-specific organization of excitatory cell types

Note: The following Section is adapted from Narayanan et al. (2017).

The sample of excitatory neuron morphologies ($n = 153$) described in Narayanan et al. (2015) was used to investigate the layer-specific organization of excitatory cell types. I grouped all neurons registered to the D2 barrel column of each cell type together and computed the isosurfaces for the respective dendrite volumes at the 99th percentile. I computed the 1D length density profile of axons and dendrites for each neuron in 50 µm steps along the cortical axis and the relative contribution of dendrite and axon length of each cell type to each of the six cortical layers.

To quantify the cell type-specific structural composition of each layer, I analyzed the estimated structural composition of the dense structural model with respect to the cortical layers. I extracted the number of somata, as well as the dendrite and axon path lengths, that each of the ten cell types contributed to each of the six layers to the D2 barrel column. To provide a comprehensive overview of the degrees to which dendrite and axon distributions of the different cell types are organized with respect to layers, I calculated the relative amount of dendrite and axon path length within each layer for each cell type. These calculations of the dendrite and axon length represented all neurons in the dense structural model of vS1 whose respective dendrites and/or axons extended into the volume of the D2 barrel column.

Path length and conduction velocity of VPM axons

In collaboration with my colleagues Dr. Robert Egger and Dr. Rajeevan T. Narayanan (Max Planck Institute for Biological Cybernetics, Tübingen, and research center caesar, Bonn), I analyzed the path length distribution of VPM axons ($n = 14$) (Egger et al., 2020). We wanted to test whether differences in path lengths of VPM axons could explain temporal delays in sensory-evoked responses between neurons in the deep layers (L6) and neurons located more superficial within vS1. Therefore, the aim was to estimate the temporal delay between sensory-evoked signals from VPM reaching the L5/6 border and the L4/5 border. Specifically,

I calculated the path length of VPM axons from the L5/6 border to the L4/5 border in the C2 barrel column. Next, I estimated the conduction time along VPM axons from the WM to the L5/6 border and to the L4/5 border. I divided the path lengths by the measured intracortical conduction velocity of VPM axons ($0.33\,\mathrm{m\,s^{-1}}$; Salami, Itami, Tsumoto, & Kimura, 2003).

Structural organization of inhibitory neurons

Similar to the analysis of excitatory neurons, I analyzed the structural estimates of INs predicted by the dense structural model (for more details, see Sections 2.1.1 and 2.1.2). For the following analysis another realization of the dense structural model of vS1 was used. This model was restricted to nine barrel columns and the septa between them (rows: B-D, arcs: 1-3) and comprised 196,553 neurons. The model estimated the number of somata that each of the five inhibitory projection types plus INs in L1 contribute to each of the six layers of the C2 barrel column. I grouped all neurons of each projection type together and computed the isosurfaces for the respective dendrite and axon volumes at the 99th and 95th percentile, respectively. I computed the 1D length density profile of axons and dendrites for each neuron in 50 µm steps along the cortical axis.

2.2 The stochastic synapse formation strategy

Note: This Section is in part adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020). The synapse formation strategy was first reported in Egger et al. (2014).

Using the structural constraints provided by the dense structural model, I applied a strategy of stochastic synapse formation that provides probabilities for synapses being formed. I made use of a synapse formation strategy derived mathematically in previous work (Egger et al., 2014). This strategy assumes that axons (specifically their presynaptic structures PREs) randomly connect to any of the locally available dendrites (specifically their postsynaptic structures POSTs). Here, the subvolumes represent the resolution limit of the dense structural model (i.e., cubes

with 50 µm edge length). Within each 50 µm cube, I calculated the "dense structural composition" (DSC) between PREs and POSTs. DSC represents the pre- and postsynaptic structures that two neurons a and b contribute to a cube $\hat{x}$ with respect to the total amount of postsynaptic structures contributed by all neurons N.

$$\mathrm{DSC}(a, b, \hat{x}) = \mathrm{PRE}(a, \hat{x}) \cdot \frac{\mathrm{POST}(b, \hat{x})}{\sum_{i \in N} \mathrm{POST}(i, \hat{x})} \tag{2.1}$$

Here, $\sum_{i \in N} \mathrm{POST}(i, \hat{x})$ represents the number of postsynaptic structures that all excitatory and INs in the dense model of vS1 contribute to a 50 µm cube $\hat{x}$, whereas $\mathrm{PRE}(a, \hat{x})$ and $\mathrm{POST}(b, \hat{x})$ represent pre- and postsynaptic structures only from neurons a and b, respectively.

Based on the DSC, I formulated two constraints for synapse formation:
Synapse constraint 1: Synapses form stochastically. All presynaptic structures within a cube $\hat{x}$ have equal probability to form n connections with any of the available postsynaptic ones.

$$p(a, b, \hat{x}, n) = \frac{(\mathrm{DSC}(a, b, \hat{x}))^n}{n!} \cdot \exp(\mathrm{DSC}(a, b, \hat{x})) \tag{2.2}$$

Here, $p(a, b, \hat{x}, n)$ is the probability of neuron a forming n synapses with neuron b in cube $\hat{x}$.

Synapse constraint 2: Synapses form independently. The formation of synapses within a cube $\hat{x}$ does not affect synapse formation elsewhere.

$$P(a, b) = 1 - \exp(-\sum_{\hat{x}} \mathrm{DSC}(a, b, \hat{x})) = 1 - \prod_{\hat{x}} \exp(-\mathrm{DSC}(a, b, \hat{x})) \tag{2.3}$$

Here, $P(a, b)$ is the probability that neuron a is connected to neuron b by at least one synapse across all cubes. Intuitively, $\mathrm{DSC}(a, b, \hat{x})$ represents the expected average number of synapses connecting neuron a to neuron b in cube $\hat{x}$ and $\mathrm{DSC}(a, b) = \sum_{\hat{x}} \mathrm{DSC}(a, b, \hat{x})$ the expected average number of synapses connecting

neuron a to neuron b. Note that in the following lowercase letters a and b denote individual neurons, while uppercase letters A and B denote neuron populations, unless stated otherwise.

This strategy represents stochastic processes of synapse formation, elimination, and replacement and purposefully neglects recognition-dependent strategies, for example, based on the identity between pre- and postsynaptic structures (e.g., implemented by molecular recognition or activity) (Sanes & Yamagata, 2009). The strategy is also consistent with the "synaptotrophic hypothesis", which states that synapse formation itself refine neuron morphologies (Vaughn et al., 1988; Niell et al., 2004; Cline & Haas, 2008).

Application of this synapse formation strategy yields a distribution of all possible wiring diagrams, as well as their respective probabilities, that are consistent with the synapse constraints, and with the dense distributions of pre- and postsynaptic structures. I refer to the probability distribution of wiring diagrams as "dense statistical connectome" (Egger et al., 2014). This particular dense statistical connectome predicts how the wiring diagram of the dense structural model would look like if it were shaped solely based on the characteristic structural composition of the neuropil.

2.3 Dense statistical connectome of rat barrel cortex

The dense statistical connectome is the result of applying the synapse formation strategy to the estimates provided by dense structural model of vS1 (described in the two previous Sections 2.1 and 2.2). For each pair of neurons (i.e., 553,572 presynaptic neurons in vS1 and VPM and 547,347 postsynaptic neurons in vS1) in the vS1 model, the respective dense structural composition (DSC) was computed per 50 µm cube (see Equation 2.1). The dense structural model comprised 51,200 such cubes. The connection probability p for each neuron pair in each cube, as well as P across all cubes were calculated (see Equations 2.2 and 2.3).

In this Section, I describe the methods used to analyze and assess the range of possible wiring diagrams and their respective probabilities predicted by the dense

statistical connectome. I first describe the analysis of the pair-wise connectivity (see Section 2.3.1), in-degree correlations (see Section 2.3.2), and network topologies (see Section 2.3.3). Second, I describe a mathematical model of how pair-wise and correlated connectivity is related to the network topology (see Section 2.3.4). *Note:* Most of the descriptions in this Section are adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020).

2.3.1 Pair-wise connectivity

I analyzed the pair-wise connectivity predicted by the dense statistical connectome. To do so, I grouped all excitatory neurons based on their somatotopic position (i.e., barrel columns or barreloids) and cell type identity. I computed the mean, SD, CV, and *skew* (denoting Pearson mode skewness[4]) of the connection probabilities between these populations. The color map for the matrix representation was mapped on the respective percentiles of the connection probability distribution; the color map for the zoom-in of the matrix representation was limited to 95% of the connection probability values.

Approximating the connection probability distributions with common distributions

I investigated whether the predicted distributions of the connection probabilities follow any common distribution. To do so, I computed the histogram of the connection probabilities (bin width: $p = 1\%$) of all possible excitatory cell type combinations within the C2 barrel column. I normalized the area of the histogram to 1, resulting in the probability mass function of the connection probabilities p_{pmf}. I analyzed the shape of the predicted connection probability distributions by determining the best fit with seven different functions: four continuous probability density functions (Gaussian, Half-Normal, Exponential, and Gamma function), and three discrete probability mass functions (zero-inflated Poisson, zero-inflated Negative Binomial, and Binomial function)[5]. For discrete mass functions, each bin

[4]mode was defined as the most frequent connection probability value when rounded four digits to the right of the decimal point

[5]A zero-inflated Poisson or Negative Binomial model allow for frequent zero-valued observations (i.e., nonoverlapping neuron pairs)

(from 0 % to 100 %) of p_{pmf} was mapped on discrete values from 0 to 100. In the case of the Gamma function, all zero connection probabilities were replaced with $\epsilon = 2^{-52}$. The best fit was determined by maximum likelihood estimation (MLE) either using the Matlab built-in function `fitdist` or in case of the zero-inflated models using Python 3.7.4 (statsmodels 0.10.1[6] and NumPy 1.16.5[7]). In case of the continuous probability density functions, the resultant distributions were binned with the same bin width as p_{pmf}. I calculated the summed area difference SAD between each of the seven functions with their best fit and p_{pmf} or the mapped p_{pmf}. SAD is bounded between 0 and 2, with 0 indicating a perfect fit and 2 indicating no overlap between the fit and p_{pmf}. The best-fitting function was the one resulting in the lowest SAD.

Relationship between connection probability and axodendritic overlap

Brown and Hestrin (2009) showed that differences in the axodendritic overlap between two neuron populations cannot fully account for the differences in their measured connection probabilities. Given this finding, I assessed the relationship between the (sparse) axodendritic overlap and the predicted connection probabilities in a similar manner as reported by Brown and Hestrin (2009). To do so, I first aligned all excitatory neuron morphologies registered to the D2 barrel column by their somata in the horizontal plane (i.e., preserving their cortical depth). I transformed each axon and dendrite morphology in a 3D axon or dendrite length density with a resolution of 50 µm cubes and calculated their axodendritic overlap by multiplying each axon length density with each dendrite length density. Hence, the resultant overlap(a,b,$\hat{x}$) denotes the product of the axon length of neuron a and the dendrite length of neuron b in each cube $\hat{x}$. For each neuron pair, I summed the resulting axodendritic overlap across all cubes: $\text{overlap}(a, b) = \sum_{\hat{x}} \text{overlap}(a, b, \hat{x})$. I calculated the mean and SD across all cell-to-cell overlaps for each cell type combination (A, B), denoted by $\text{overlap}(A, B)$.

To compare the overlaps and the respective connection probabilities predicted by the model, the following analysis was performed: First, I calculated the correlation coefficient R between the mean connection probabilities $\text{P}(A, B)$ and the mean

[6] https://www.statsmodels.org/stable/index.html
[7] https://numpy.org/

axodendritic overlap(A, B). Second, I calculated the ratios between all overlaps across all cell type combinations and the ratios between all connection probabilities across all cell type combinations. Next, I assessed whether these ratios are linearly related with a slope of one, i.e., whether an increase of the axodendritic overlap between two cell type combinations is reflected in the same increase of the respective connection probabilities. Specifically, I performed 100 trials of 1,000 randomly sampled ratios, fitted a linear model with an intercept, and tested whether the fitted slope was significantly different from a slope of one using the Matlab built-in `coefTest`. I repeated this analysis and compared the mean overlap to mean connection probabilities derived from neurons grouped by different intersomatic distances in the horizontal plane. I also repeated the analysis for the overlap calculated in 2D, i.e., one lateral dimension was collapsed, and the overlap was the product of the 2D axon and dendrite length densities.

Emulation of slice experiments

Many anatomical and functional studies are performed in tissue slices *in vitro*. Assessing the impact of slicing on observed properties is difficult. Therefore, I implemented an approach to emulate slicing within the model that allows a systematic assessment of the impact of slicing (i.e., neuron truncation). I created ten slices of 300 µm thickness through the dense structural model of vS1. Each slice contained either the entire C2 barrel column or parts of the C2 barrel column (i.e., approximately the center of vS1). These *in silico* brain slices were shifted by 20 µm with respect to each other along the rostro-caudal axis. Axons and dendrites of all neurons whose somata were located within a slice were truncated, i.e., branchlets were cut at their intersection with the slice surface, and branchlets that became disconnected from the soma were removed (see Figure 2.5). I computed the connection probabilities between each neuron pair in the slice models as described previously but PRE and POST of the neuron pairs were determined based on the truncated morphologies, whereas the total number of POST in each 50 µm cube was identical with the one from the complete vS1 model. I investigated the impact of slicing on the structural estimates (i.e., neurite length) and predicted connectivity (i.e., connection probability between neurons).

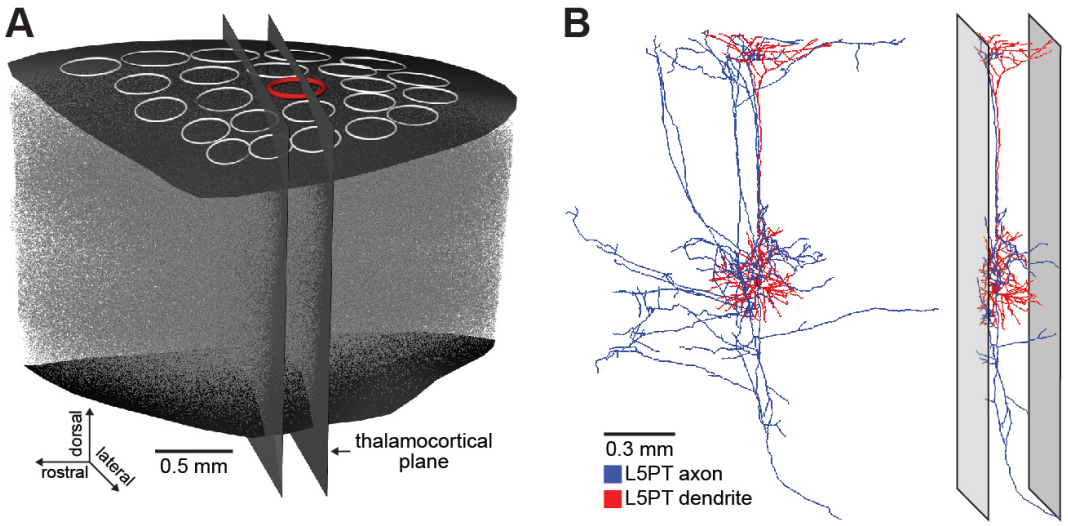

Figure 2.5. Emulation of slice experiments in model. (A) Model of vS1 with all excitatory somata in gray, C2 barrel column highlighted in red, thalamocortical planes illustrate one of the ten 300 µm thalamocortical slices through vS1. (B) Exemplary *in vivo* morphology of an L5PT (left) and its respective truncated morphology in one of the ten 300 µm thalamocortical slices (right). Only branchlets that are connected to the soma are preserved.

First, I selected all neurons whose somata were closest to the C2 barrel column and had a tissue depth between 31 µm and 130 µm[8] (i.e., shortest distance to the slicing surface). Next, I calculated the *in vivo* path length of each dendrite and axon. I repeated the calculation of the path lengths for the truncated dendrites and axons. I grouped all neurons by their cell type and calculated how much path length was preserved (i.e., dividing the truncated *in vitro* path length by their respective *in vivo* path length).

Second, I calculated the predicted connection probabilities between all neuron pairs located closest to the C2 barrel column and within the emulated slice; first without truncated dendrites/axons and then with truncated dendrites/axons. I grouped the predicted connection probabilities by their cell type identities, their intersomatic distances, and tissue depths. The mean, SD, and CV of the predicted connection probabilities for all neuron pairs whose tissue depth was between 31 µm

[8]mean tissue depth across multiple *in vitro* studies (Yoshimura, Dantzker, & Callaway, 2005; Song et al., 2005; Markram, Lübke, Frotscher, Roth, & Sakmann, 1997; Jiang et al., 2015; Feldmeyer, Lübke, & Sakmann, 2006).

and 130 µm were determined. The ratios between the *in vivo* and *in vitro* predicted connectivity statistics for each cell type combination were calculated.

Connectivity involving inhibitory neurons

Although an ongoing effort in the group of Dr. Marcel Oberlaender, I performed a limited analysis of the predicted thalamocortical input to INs of an average cortical barrel column. Using the predictions of the dense statistical connectome, I investigated how INs are innervated by axons from VPM thalamus. All INs of L2-6 in the vS1 model that were assigned to the same projection type and to the C2 barrel column were grouped to predict the mean, projection type-specific number of VPM synapses (i.e., DSC between the somatodendritic domain of each IN and VPM axons). I computed the 3D axon and bouton density of each projection type at a resolution of 50 µm cubes (see Section 2.1.2).

2.3.2 In-degree correlations

I assessed to what degree the number of excitatory synaptic inputs in the dense statistical connectome are correlated. To do so, I calculated $\mathrm{DSC}(a, b)$ for each excitatory pair of presynaptic neuron a and postsynaptic neuron b located in the C2 barrel column or barreloid:

$$\mathrm{DSC}(a, b) = \sum_{\hat{x}} \mathrm{DSC}_{(a,b,\hat{x})} \tag{2.4}$$

I grouped all presynaptic neurons by their cell type, and calculated the total number of synapses (i.e., in-degree) each postsynaptic neuron b receives from each cell type population A:

$$\mathrm{inDegree}(A, b) = \sum_{a \in A} \mathrm{DSC}_{(a,b)} \tag{2.5}$$

I computed the Pearson's linear correlation coefficient R and the corresponding p value between the in-degrees of two different cell type populations A and C onto all neurons b of one cell type population B (i.e., correlation between $\mathrm{inDegree}(A, b)$ and $\mathrm{inDegree}(C, b)$ across all neurons of population B). I repeated this computa-

tion for all possible combinations of presynaptic cell types, resulting in the distribution of 55 correlation coefficients R per postsynaptic cell type. The significance level was set to $\alpha = 0.01$. For illustration of bouton, spine, and DSC densities, the isosurfaces were calculated at the 95th, 99th, and 99th percentile, respectively.

2.3.3 Network topologies

Network topologies are commonly described with respect to topologies expected in a random network. Hence, I compared the occurrences of certain network motifs predicted by the model to those expected in a random network. I used the Matlab built-in `digraph` to illustrate a predicted network between 50 neurons located within L5 of a single cortical column as a graph. Edges between each neuron pair in the illustration were determined based on their predicted connection probabilities. In the random graph, edges were determined based on the mean connection probabilities across all 50 neurons.

The analysis of network motifs was restricted to small groups of neurons ranging from three to ten neurons. Commonly motifs between three neurons, referred to as a triplet, are studied. Triplets can be connected by 15 different connectivity motifs (see Figure 2.6). The extraction of triplet motif occurrences as predicted by the dense statistical connectome and parts of the analysis were first implemented by Dr. Robert Egger (Egger et al., 2014; Egger, 2015)[9]. I used his implementation and further refined and extended it.

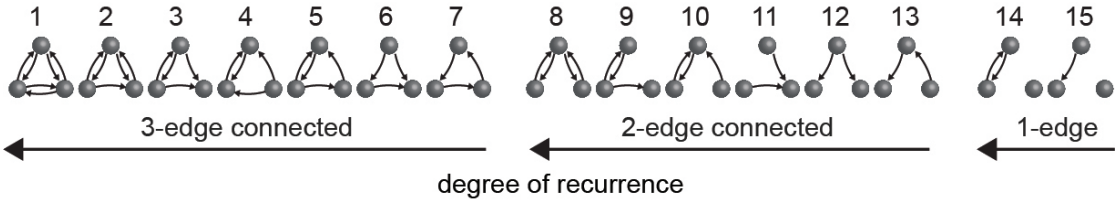

Figure 2.6. Spectra of triplet motifs sorted by their number of connected edges. Numbers above triplet motif represent motif ID.

Layer and cell type-specific triplet motif spectrum: I calculated the probability of occurrences of each triplet motif for a given set of 10,000 randomly

[9]Dr. Robert Egger limited his analysis to triplets of L4ss neurons.

selected neuron triplets. For readability's sake, I refer to the probability of occurrences of a motif as motif probability. Each neuron triplet is represented only once per set. The neuron triplets were sampled such that each neuron was either from one of the ten excitatory cell types or an excitatory or inhibitory neuron from L2/3, L4, L5, or L6, and located in the C2 barrel column. I repeated this step 10 times and calculated the mean and standard error of the mean (SEM) of each motif probability, denoted by $P_{Motif,Mean}$ and $P_{Motif,SEM}$. Motifs were sorted by their respective number of bidirectional edges (i.e., one, two, or three). I compared the mean motif probabilities $P_{Motif,Mean}$ predicted by the dense statistical connectome with those expected in a random network, where the probability of each edge is set to the respective mean probability of that edge as predicted by the dense statistical connectome:

1. I calculated the mean connection probability for each of the six edges between all three neuron populations.

2. I used these edge-wise mean connection probabilities to compute the probability of each motif, denoted by $P_{Motif,Random}$.

3. I divided $P_{Motif,Mean}$ as predicted by the dense statistical connectome by $P_{Motif,Random}$, the respective motif probability expected in a random network. The resulting ratio is referred to as motif deviation.

A motif deviation larger than one hence reflects that a motif is more frequent in the connectome than in the random network (i.e., the motif is overrepresented), while a motif deviation smaller than one reflects that a motif is less frequent in the dense statistical connectome (i.e., underrepresented). To test for significant deviations, p values were determined by computing the z-score between the probabilities in the random network, and the mean and SEM of probabilities across the 10 repetitions of sampling from the dense statistical connectome, i.e., $z = (P_{Motif,Random} - P_{Motif,Mean})/P_{Motif,SEM}$.

To illustrate the difference between the dense statistical connectome and a randomly permutated connection matrix with the same distribution shape, I randomly shuffled all values in the connection matrix between L5IT and L5PT neurons and

recalculated the in-degree correlations and motif deviations for the triplet motif spectrum.

To gain insight into the motif deviations, I analyzed the relationship between motif deviations and their underlying connection probability distributions and in-degree correlations. I computed motif deviations of all 15 triplet motifs for all 210 cell type-specific triplet combinations with at least two different excitatory cell types. For each triplet combination, I calculated the mean and CV of their connection probability distribution across all six edges and their mean correlation of in-degrees. The mean correlation of in-degrees was determined across all in-degree correlation coefficients involving the cell types of the respective triplet combination.

Most likely cell type combination for each triplet motif: I randomly selected 500,000 excitatory neuron triplets whose somata were located within the entire vS1 model and calculated the motif probabilities per triplet. Across all triplets, I determined the most likely cell type configuration for each motif as well as its probability of occurring compared to all other cell type configurations.

Network motifs between more than three neurons: To assess the occurrences of motifs between more than three neurons in the dense statistical connectome, I computed the probabilities of two exemplary motifs for up to ten neurons located within the C2 barrel column. The first motif represented "full recurrence", i.e., all neurons are connected via bidirectional edges. I randomly sampled 10,000,000 sets of neurons per motif size[10] (i.e., 3 to 10) from the dense statistical connectome, and computed the probability of the fully bidirectional motif, respectively. The respective probabilities in the random network were computed based on the mean connection probabilities across all neurons of the sample. I computed the motif deviation based on these probabilities. The second motif represented a "feedforward chain" of neurons (i.e., neurons except for two have exactly one incoming and one outgoing unidirectional edge, and the two remaining neurons have either only one incoming or one outgoing unidirectional edge). I randomly sampled 10,000 sets of neurons per motif size (i.e., 3 to 10) from the dense statisti-

[10]Number of neurons involved in probed motif, e.g., a triplet motif has a motif size of 3.

cal connectome, computed the probability of 1,000 randomly sampled feedforward chain motifs per neuron set, respectively, and computed the motif deviations from a random network.

2.3.4 Mathematical model of correlated connectivity

In collaboration with Prof. Dr. Jakob Macke (University Tübingen), we developed a simplified mathematical model of a network with correlated connectivity to investigate how properties of the connection probability distributions, such as their mean, SD, and correlations, relate to the predicted and observed motif deviations. The model assumes that the connection probabilities between any two edges are correlated. The model is closely related to a model studied in Macke, Opper, and Bethge (2011).

In the following, some notations:

- $\mathcal{N}(\mu, \lambda)$ denotes a Gaussian distribution with mean μ and variance λ.
- $\phi(t, \mu, \lambda)$ denotes the respective Gaussian probability density function evaluated at t.
- $\Phi(s, \mu, \lambda)$ denotes the respective cumulative probability density function evaluated at s.
- $L(s, \mu, \lambda) = 1 - \Phi(s, \mu, \lambda)$ defines the complementary cumulative probability density function.

The model assumes that whether there is an i-th edge between two nodes (denoted by $X_i = 1$, otherwise $X_i = 0$) is the result of a combination of a "private" source T_i, and a "shared" source S. The bigger the shared source S is relative to the private source T_i, the more correlated the resultant connection probabilities are. The mathematical model has two parameters: γ, representing the degree of connectivity, and λ, bounded between 0 and 1, and representing both the degree of correlation between shared and private source and their heterogeneity (as I demonstrate in the following). We define that the i-th edge exists (i.e., $X_i = 1$) whenever the joint input of T_i and S, denoted by Z_i, is larger than 0:

$$Z_i = \gamma + \sqrt{\lambda}S + \sqrt{\eta}T_i, \text{ where} \tag{2.6a}$$
$$\eta = 1 - \lambda \tag{2.6b}$$
$$S \sim \mathcal{N}(0, 1) \tag{2.6c}$$
$$T_i \sim \mathcal{N}(0, 1), \text{ thus} \tag{2.6d}$$
$$Z_i \sim \mathcal{N}(\gamma, \eta + \lambda) = \mathcal{N}(\gamma, 1) \tag{2.6e}$$
$$\text{cov}(Z_i, Z_j) = \lambda \tag{2.6f}$$

If $\lambda = 1$, X_i is only determined by the shared source S, while if $\lambda = 0$, X_i is only determined by the private source T_i. Given this mathematical model, the connection probability p_i for each edge X_i can be derived:

$$p_i(S) = P(X_i = 1|S) = L(0, \gamma + \sqrt{\lambda}S, \eta), \text{ then} \tag{2.7a}$$
$$\mu = \text{E}_S(p_i) = L(0, \gamma, 1) \tag{2.7b}$$
$$\sigma^2 = \text{Var}_S(p_i) \tag{2.7c}$$
$$= \int_{-\infty}^{\infty} P(X_i = 1|s)^2 \phi(s, 0, 1) ds - \mu^2 \tag{2.7d}$$
$$= \int_{-\infty}^{\infty} L(0, \gamma + \sqrt{\lambda}s, \eta)^2 \phi(s, 0, 1) ds - \mu^2 \tag{2.7e}$$

Deriving the covariance between any two connections X_i and X_j yields:

$$\text{cov}(X_i, X_j) = \text{E}(X_i, X_j) - \mu^2 = P(X_i = X_j = 1) - \mu^2 \tag{2.8a}$$
$$= \int_{-\infty}^{\infty} P(X_i = 1|s)^2 P(X_j = 1|s) \phi(s, 0, 1) ds - \mu^2 \tag{2.8b}$$
$$= \int_{-\infty}^{\infty} L(0, \gamma + \sqrt{\lambda}s, \eta)^2 \phi(s, 0, 1) ds - \mu^2 \tag{2.8c}$$
$$= \text{Var}(p_i) \tag{2.8d}$$

Thus, in this simplified, mathematical model the covariance of the connections $\text{cov}(X_i, X_j)$ is equal to the variance of connection probabilities $\text{Var}(p_i)$—the more

strongly the connection probabilities vary, the more strongly the connections themselves are correlated. Hence, the parameter λ represents a measure of both the degree of correlation and heterogeneity. To assess the impact of λ and the mean connection probability μ on motif occurrences and deviations, the probability that k out of K connections of a motif are realized is given by:

$$P(|X| = k) = \int_{-\infty}^{\infty} \binom{K}{k} L(0, \gamma + \sqrt{\lambda}s, \eta)^k \Phi(0, \gamma + \sqrt{\lambda}s, \eta)^{K-k} \phi(s, 0, 1) ds \quad (2.9)$$

I implemented this mathematical model as a numerical simulation in Matlab. I iterated over 250 γ values ranging from -2 to 2, and over 250 λ values ranging from 0 to 1. Per combination of γ and λ values, 10 trials, each with 100,000 random samples, were generated. For each trial, the mean and variance across the connection probabilities p_i, (i.e., μ and σ^2), the probability of each triplet motif $P(|X| = k)$ with $K = 6$ (i.e., maximal number of edges in a triplet), and the respective probability expected in random network based solely on μ was calculated. For details of the mathematical model for one trial, see Algorithm 1.

The deviation of each triplet motif was the ratio between the means of the motif probabilities across all trials. The deviations were mapped on a grid spanned by 20 μ values (i.e., sparsity) and 20 λ values (i.e., correlations and heterogeneity) and visualized by a log-space color map. Each of the 220 cell type-specific triplet combinations was mapped into the grid space. Specifically, for each combination, its respective λ value and motif deviation were inferred based on the variance and mean of the connection probabilities of each combination and a lookup table of μ, σ^2, and λ values as determined by the numerical simulation. The correlation coefficient R between both motif deviations was determined, i.e., between the deviation predicted by the dense statistical connectome and its respective deviation predicted by the simplified mathematical model of correlated connectivity.

Algorithm 1 A simple mathematical model of correlated connectivity

input simulator with degree of correlations and heterogeneity λ and degree of connectivity γ
randomly initialize shared source $S \sim \mathcal{N}(0, 1)$
$\eta := 1 - \lambda$
$p_i := L(0, \gamma + \sqrt{\lambda}S, \eta)$ ▷ p_i is a vector of connection probabilities
$\mu := \mathrm{E}(p_i)$ ▷ E denotes the expected value
$\sigma^2 := \mathrm{Var}(p_i)$ ▷ Var denotes the variance
$K := 6$ ▷ maximal number of edges in a triplet
for $k = 0$ to K **do**
    $P(k) := \binom{K}{k}\mathrm{E}(p_i^k(1 - p_i)^{K-k})$ ▷ probability of motif with k edges
    $P_{random}(k) := \binom{K}{k}(\mu^k(1 - \mu)^{K-k})$ ▷ probability of motif with k edges in a random network
end for
return P, P_{random}, μ, and σ^2

2.4 Comparison with empirical data

In this Section, I describe the methods used to compare empirical properties of wiring with the ones predicted by the dense statistical connectome. The section is subdivided into five parts. The first two parts describe a systematic comparisons to connectivity measurements between excitatory neurons reported in the literature at the subcellular and cellular level (see Sections 2.4.1 and 2.4.2). The third part of this Section describes a limited comparison to measured connectivity involving INs (see Section 2.4.3). The fourth part describes a comparison between observed and predicted network topologies (see Section 2.4.4). In the last part of this Section, I describe a direct comparison between putative synaptic contact patterns of individual excitatory neurons pairs determined in *in vivo* experiments in the group of Dr. Marcel Oberlaender and their respective predictions within the model (see Section 2.4.5). The descriptions in Sections 2.4.1, 2.4.2, and 2.4.4 are adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020).

When comparing empirical data to the probabilistic predictions of the dense sta-

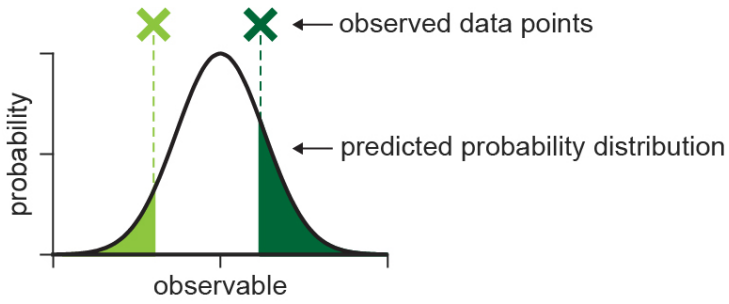

Figure 2.7. Illustration of comparing observed data vs. predicted probability distribution of observable. Two observed data points (green shaded crosses) and the predicted probability distribution. In geometrical terms, the $p_{extreme}$ values of the two observations are the respectively green shaded areas multiplied by two. In this example, the dark shaded observation is more likely to be observed in the model than the light shaded observation given the probability distribution.

tistical connectome, I deployed one particular approach multiple times (see Figure 2.7): I determined where the observable fell within the predicted probability distribution of the respective observable. Specifically, I calculated the probability of observing a value at least as extreme as the empirically observed value x given the predicted probability distribution of the respective observable. This probability is denoted by $p_{extreme}(x)$ and calculated as follows: I determined the cumulative probability of an observable x occurring in the dense statistical connectome, denoted by $F(x)$. $p_{extreme}(x)$ is then defined as:

$$p_{extreme}(x) = \begin{cases} 2 \cdot F(x), & \text{if } F(x) \leq 0.5 \\ 2 \cdot (1 - F(x)), & \text{otherwise} \end{cases} \tag{2.10}$$

Intuitively, $p_{extreme}(x)$ can be interpreted as the p value. A value of $p_{extreme}$ close to zero indicates that the observed data x has a low probability of occurring in the dense statistical connectome (i.e., the probability of observing a value at least as extreme as the observed one in the model is low), while a value close to 1 indicates the opposite, i.e., x is likely to occur.

2.4.1 Subcellular connectivity between excitatory neurons

Saturated reconstructions of small subvolumes reported occurrences of the number of synapses between the same neuron branchlets (Motta et al., 2019; Kasthuri et al., 2015). I compared these observations to the predictions of the dense statistical connectome. For each excitatory neuron pair, I calculated the predicted probabilities that they form zero, one, two, three, and (at least) four synapses per 50 µm cube located within L2-6 of the C2 barrel column, and multiplied the probabilities by the number of neuron pairs that innervate each cube, respectively. This yielded $f_{Model}(n, \hat{x})$, the occurrences of n synapses between the same neuron pairs in a cube $\hat{x}$. The hence predicted median, minimum, and maximum occurrences across cubes of $n = 1, 2, 3$, or ≥ 4 synapses per neuron pair were fitted a power law using the Matlab built-in function `fit`: $f(n) = \alpha \cdot n^{\beta}$. I repeated the fitting for $f_{Empirical}(n)$, the reported occurrences of n synapses per neuron branch from saturated reconstructions (Kasthuri et al., 2015; Motta et al., 2019) by scaling the respectively reconstructed volumes ($1.5 \times 10^{-6}\,\mathrm{mm}^3$ and $5 \times 10^{-4}\,\mathrm{mm}^3$) to a 50 µm cube (i.e., $1.25 \times 10^{-4}\,\mathrm{mm}^3$). I repeated the analysis in a layer-specific manner, i.e., for cubes in L2/3, L4, L5, and L6. In addition to the analysis of neuron pairs, the analysis was repeated for each neuron branch pair in the respective cubes.

In the next step, I calculated $p_{extreme}$, i.e., the probability of observing occurrences of synapse numbers at least as extreme as the empirical ones in the dense statistical connectome (see Equation 2.10). To do so, I determined the ratio of cubes in which the occurrences of n synapses per neuron pair was larger than its respective empirically determined occurrences, i.e., where $f_{Model}(n, \hat{x}) > f_{Empirical}(n)$. Given the ratio, I derived the probability $p_{extreme}(n)$ of observing a number of neuron pairs at least as extreme as the empirical number given the model's predicted probability distribution.

2.4.2 Cellular connectivity between excitatory neurons

In this Section, I first describe a systematic comparison of empirically determined and predicted connection probabilities, followed by a comparison between empirical and predicted synapse distributions along dendrites and number of synapses per connection.

Table 2.6. Neuron grouping for comparison. When comparing to *in vitro* data, connectivity was predicted using ten *in silico* slices. Neurons were grouped in addition to their layer position and cell type by their tissue depth. Connectivity for *in vitro* comparisons was restricted to the basal dendrite of the postsynaptic neuron population.

	in vitro	*in vivo*
model	slice model	vS1 model
neuron selection (A,B)	one barrel column layer cell type (if available) tissue depth	one barrel column layer cell type (if available)
connectivity P(A,B)	restricted to basal dendrite of B	

Connection probabilities

I compared 89 measurements of connection probabilities from 29 studies with the predicted connection probability distributions (see Table A10). Because most measurements were performed in acute brain slices *in vitro*, I emulated the respective experimental conditions in the model to better compare empirical with predicted connectivity.

Emulation of slice experiments: To emulate slicing experiments, I used the predictions provided by the *in silico* slicing model described previously in Section 2.3.1. I restricted the comparison to neurons located closest to the C2 barrel column. I then grouped the neurons as described in the respective studies (see Tables A10 and 2.6 for the grouping): First, neurons were grouped by their respective soma location within layers. Layer borders were defined as reported previously (Meyer et al., 2013)[11]. Second, if the cell type was reported in the studies, neurons were grouped additionally based on their cell type identity. Two comparisons (Bruno & Simons, 2002; Bruno & Sakmann, 2006) were restricted to postsynaptic

[11]Borders between L2 and L3, L5A and L5B, and L6A and L6B were set to the center location of L2/3, L5 and L6, respectively.

neurons located between barrel columns (i.e., within the septum). Third, I grouped neurons by their tissue depth (i.e., the shortest distance of the neuron's soma from the slicing surface). If the tissue depth was not reported, I restricted the comparison to neuron pairs within the mean reported range of tissue depths (31 µm to 130 µm). *In vitro* studies were compared to connection probability predictions on basal dendrites to mimic signal attenuation due to space clamp errors.

Comparison with empirical connection probabilities: I calculated the distribution of connection probabilities for each of the 89 measured connection probabilities using the grouping mentioned previously (see Tables A10 and 2.6). For each comparison I determined the mean and SD of the predicted connection probabilities (P_{Mean} and P_{SD}) and its respective probability coverage. The coverage was determined as follows:

$$coverage = \min(1, P_{Mean} + P_{SD}) - \max(0, P_{Mean} - P_{SD}) \tag{2.11}$$

To quantify the consistency between empirical and predicted connection probabilities, I calculated different "consistency measures". One consistency measure was Pearson's linear correlation coefficient R between the empirical and predicted mean connection probabilities (i.e., $P_{Empirical}$ and P_{Mean}, respectively). I computed the 95% confidence bounds for new observations based on a linear regression with no intercept using the Matlab built-ins `fitlm` and `predict`. I further assessed the model's consistency by determining the deviation to the empirical connection probabilities in units of SD and in units of the standard error of the mean (SEM), denoted by dev_{SD} and dev_{SEM} (see Figure 2.8). The deviation in units of SD was calculated as follows:

$$dev_{\mathrm{SD}} = \frac{P_{Empirical} - P_{Mean}}{P_{SD}} \tag{2.12}$$

The deviation in units of SEM was calculated for all comparisons where the number of neuron pairs tested empirically, denoted by $n_{empirical}$, was available[12] as follows:

[12]For two comparisons the empirical number of neuron pairs tested was not available.

$$dev_{\text{SEM}} = \frac{P_{Empirical} - P_{Mean}}{P_{SEM}} \tag{2.13}$$

where $P_{SEM} = P_{SD}/\sqrt{n_{empirical}}$. Next, I calculated the percentile $prctl$, the percentage of neuron pairs with a predicted connection probability smaller than $P_{Empirical}$:

$$prctl = \frac{\sum_{(a,b)\in\mathcal{S}_1} 1}{\sum_{(a,b)\in\mathcal{S}} 1} \tag{2.14}$$

where $\mathcal{S}$ denotes all neuron pairs that matched the grouping criteria for the respective comparison and $\mathcal{S}_1$ denotes a subset of $\mathcal{S}$ of all neuron pairs whose predicted connection probabilities are smaller or equal to its respective empirical connection probability $P_{Empirical}$. A percentile of 0% reflects that none of the predicted connection probabilities are lower than $P_{Empirical}$, while a percentile of 100% reflects that all predicted connection probabilities are smaller than $P_{Empirical}$. I derived $p_{extreme}$, the probability of observing a connection probability that is at least as extreme as the empirical connection probability $P_{Empirical}$ given the predicted connection probability distribution (cf. Equation 2.10 and Figure 2.7). However, note that $P_{Empirical}$ is a derived statistic over a sample of tested connections, while the predicted connection probability distribution reflects connection probabilities between single neuron pairs.

$$p_{extreme} = \begin{cases} 2 \cdot prctl, & \text{if } prctl \leq 0.5 \\ 2 \cdot (1 - prctl), & \text{otherwise} \end{cases} \tag{2.15}$$

Random permutation tests: I performed a random permutation test for each consistency measure (i.e., correlation coefficient R, absolute deviation values $|dev_{\text{SD}}|$ and $|dev_{\text{SEM}}|$, and $p_{extreme}$) by shuffling the 89 empirical and predicted connection probabilities and recomputing their consistency measures (see Figure 2.9). I repeated this step 100,000 times and calculated the respective histograms of the resulting consistency measures. To test for significant differences between the distributions of absolute deviation values and $p_{extreme}$ when randomly shuffling the connection probabilities, I performed a two-sample Kolmogorov-Smirnov test.

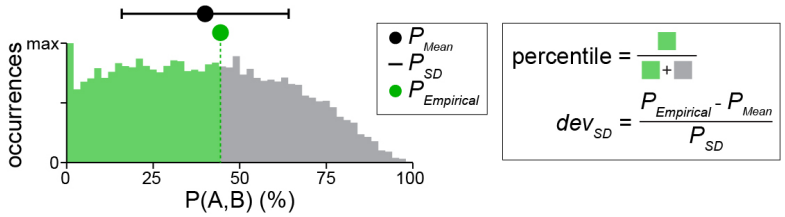

Figure 2.8. Empirical vs. predicted connection probabilities for one exemplary study (Constantinople & Bruno, 2013). Black dot reflects predicted mean connection probability, line reflects predicted SD of connection probabilities, green dot reflects empirical connection probability, and histogram represents predicted connection probabilities. Here, the deviation in units of SD and the percentile calculation are illustrated.

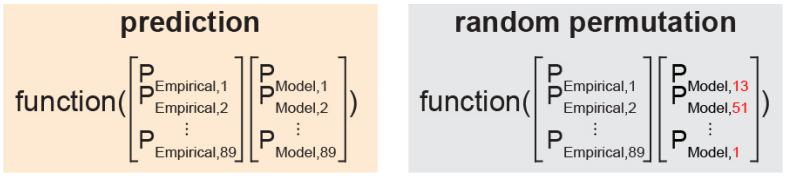

Figure 2.9. Illustration of random permutation test. Left: Each empirical connection probability is compared to its respective prediction of the model and evaluated with a consistency measure such as correlation or deviation. Right: Empirical and predicted connection probabilities are randomly paired (random permutation) and then evaluated. 100,000 random permutations were performed.

Control comparisons for different conditions: In the previously described comparison, I compared empirical connection probabilities from *in vitro* experiments to connection probabilities predicted between truncated axons and basal dendrites in emulated slices. I tested to what the degree emulating the slicing procedure and signal attenuation due to space clamp errors (i.e., restricting connectivity on the basal dendrites) affected the correlation coefficient between the empirical and predicted connection probabilities. Therefore, I repeated the comparison without emulating the slice and/or by calculating the connection probabilities between axons and the entire dendrite (i.e., basal and apical parts).

Distance-dependent connection probabilities: A few studies investigated the relationship between connectivity and intersomatic distance, revealing that in general connectivity between neuron pairs decreases with increasing intersomatic

distance (e.g., Avermann et al., 2012; Perin et al., 2011; Boucsein, Nawrot, Schnepel, & Aertsen, 2011). I compared the dense statistical connectome's predictions with three empirical studies that reported such observations *in vitro* (see Tables A5 and A6). I grouped neurons in the *in silico* slices by layer and/or cell types as described previously but additionally by their intersomatic distance along the lateral axis (i.e., the axis running parallel to the slicing surface). Boucsein et al. (2011) performed *in vitro* laser uncaging of caged glutamate in L5, thus measuring relative connection probabilities. They fitted an exponential decay model of the form $P(d) = P_0 \cdot exp(-d/\lambda_{Empirical})$ to their empirical data, where d denotes the intersomatic distance and P the measured relative connection probability. P_0 was set to 0.135 and they extracted the length constant $\lambda_{Empirical}$. I mimicked the experiment, calculated the connection probabilities between all pairs of L5PT and L5IT neurons across multiple *in silico* slices, and determined their lateral intersomatic distances. I fitted the exponential decay model to all predicted connection probabilities and their intersomatic distances and extracted the length constant λ_{Model} for comparison with $\lambda_{Empirical}$.

Synapse distributions along dendrites

Two empirical *in vitro* studies reported cell type- and layer-specific synapse distributions along dendrites (Qi & Feldmeyer, 2016; Petreanu et al., 2009). I compared the model predictions to the observed distributions. I grouped neurons in the model by layer and/or cell types but without emulating *in silico* slices. I calculated the 3D density of the DSC of each neuron pair per 50 µm cube.

Petreanu et al. (2009) reported the cortical depth profiles of synapses to L5A and L5B neurons in mouse vS1, which were normalized by the dendrite length profiles of the corresponding postsynaptic populations. I compared these measured, normalized profiles with the respectively predicted profiles of the model. Because measurements and predictions represent mouse and rat vS1, respectively, the empirically determined profiles were aligned, scaled, and interpolated so that their maximum peak and pia position matched the predicted ones.

Qi and Feldmeyer (2016) mapped the locations of putative synapses that originated from L4ss or L4sp along the dendrites of *in vitro* labeled L6A pyramids.

To compare the empirical synapse-to-soma profile with the predicted one, I transformed the distribution of the DSC to an input-weighted distribution of path length distances per neuron pair. Per presynaptic neuron a, postsynaptic dendrite branch b and 50 µm cube $\hat{x}$, I performed the following steps:

1. I determined all intersection points of branch b with cube $\hat{x}$ and all terminal points of b within cube $\hat{x}$.
2. I calculated the path length distance for each point determined in the first step.
3. I determined the range of the path length distances $L_{min}(b, \hat{x})$ and $L_{max}(b, \hat{x})$.
4. I computed the DSC between a and b.
5. I computed the input-weighted histogram of the path length distances with a bin width of 50 µm: I increased the bin value of all bins overlapping with the range of $L_{min}(b, \hat{x})$ and $L_{max}(b, \hat{x})$ by the fraction of the path length contributed to that bin and multiplied by the respective DSC.
6. I summed the input-weighted histogram over all branchlets and 50 µm cubes.

Per presynaptic axon morphology ($n_{L4ss} = 9$, $n_{L4sp} = 5$) I calculated the average input-weighted histogram (i.e., synapse-to-soma histograms) across all postsynaptic neurons. I computed the average synapse-to-soma distance for a subsample of four L4ss axons and for all five L4sp axons. The histograms of the other five L4ss axons did not differ from the average histogram of L4sp axons.

Numbers of synapses per connection

Another commonly observed quantity is the number of synapses per connection between pairs of neurons. I compared the predictions of the dense statistical connectome with seven studies that empirically determined the number of synapses per connection between two neuron populations A and B, referred to as $n(A, B)$. For comparing and predicting the synapses per connection, I grouped the neurons as in the respective studies either by layer only or by both layer and cell

type. A slice emulation was not performed. Per neuron pair, the model predicts the probability of observing any number of synapses (see Equation 2.2). Given these probabilities, I computed the 99th percentile of the number of synapses per connected neuron pair $n(a, b)$, referred to as the range of synapses per connection (Egger et al., 2014). I computed the average distribution of synapses per connection $n(A, B)$ by averaging across the individual synapse number distributions per neuron pair $n(a, b)$ as described in Egger et al. (2014):

$$n(A, B) = \langle n(a, b) \rangle_{a \in A, b \in B} = \langle Poisson(\mathrm{DSC}(a, b)) \rangle_{a \in A, b \in B} \tag{2.16}$$

where $\langle \cdots \rangle_{a \in A, b \in B}$ is the average across all neuron pairs between populations A and B. To assess the model's predictions, I determined the correlation coefficient R between the predicted average number of synapses per connection and its respective empirical number. Additionally, the distribution of the number of synapses per connection was determined for each cell type combination.

2.4.3 Connectivity involving inhibitory neurons

Although an ongoing effort in the group of Dr. Marcel Oberlaender, I performed a limited comparison between various connectivity measurements involving INs and the prediction of the dense statistical connectome, similarly as for the excitatory neurons (see Section 2.4.2).

Subcellular synapse distributions

Observations, for example, in electron microscopic studies revealed synapse numbers and densities onto inhibitory somata and dendrite branches (e.g., White & Rock, 1981; White et al., 1984; Keller & White, 1987). To compare these observations with the dense statistical connectome, I grouped all INs located closest to the C2 barrel column together. I determined the predicted number of synapses each IN receives from all the neurons in the vS1 model[13] and their location on the IN (i.e., dendrite or soma). Depending on the type and location of the presynaptic neuron, synapses were grouped by excitatory, inhibitory, or VPM. To calculate

[13] here a smaller model is used, comprised of 196,553 neurons located in nine barrel columns and the septa between them (rows: B-D, arcs: 1-3)

the respective synapse densities per µm dendrite length and per µm^2 soma surface area, I determined the total dendrite path length and soma surface area of each IN. The VPM synapse densities were determined only for INs whose somata were located in L4.

Pair-wise connectivity statistics

Like for excitatory neurons, numerous studies investigated pair-wise connectivity statistics between neuron populations involving INs. I compared measurements of connection probabilities and numbers of synapses per connected neuron with the respective values predicted by the dense statistical connectome. Specifically, I compared 73 empirical connection probabilities from 23 studies with the mean of the predicted connection probability distributions (see Table A9) and 6 empirical numbers of synapses per connected neuron pair from three studies with the predicted numbers of synapses (see Table A9). Most of these studies grouped the involved INs by different properties, such as the laminar position of their somata, electrophysiological responses (e.g., fast-spiking, non-fast-spiking, or low threshold spiking), morphological properties (e.g., Martinotti cells), chemical marker (e.g., somatostatin, parvalbumin, or vasointen- stinal peptide), or by a combination thereof.

For all comparisons except for two, I only sampled from INs of the same layer as the empirically probed layer. Slices were not emulated and all other properties were ignored. For example, while Walker et al. (2016) measured the *in vitro* connectivity between Martinotti cells in L2/3 and close-by parvalbumin-positive neurons, the predicted connectivity was calculated among all pairs of INs in L2/3. In case excitatory neurons were involved in the connectivity, these were either grouped by layer or cell type. All neurons were located within the C2 barrel column. For one study (Jiang et al., 2015), I calculated the empirical layer-specific connection probabilities and used those for comparison instead of their reported cell type-specific connection probabilities. I performed a more elaborate comparison for two connectivity measurements by Koelbl et al. (2013), as described in the following paragraph.

Emulation of an inhibitory slice experiment: For two connectivity measurements reported in Koelbl et al. (2013), I mimicked their respective experimental conditions. Koelbl et al. (2013) measured the connection probability between fast-spiking INs in L4 and L4ss in thalamocortical slices of width 300 µm (see Figure 2.10).

First, I emulated the slicing procedure: I used the reconstructed IN morphologies ($n = 40$) from Koelbl et al. (2013) and registered the morphologies to the D2 barrel column without compensating for the truncation. Neurons were integrated into the already populated dense structural model by placing them in center of the D2 barrel column in the horizontal plane and at their respective registered cortical depth. Based on the bounding box around the truncated axon of each IN morphology, I approximated the thalamocortical slice that had truncated the axon. I placed the slicing planes at their proper location with respect to each truncated axon. Based on these slice planes, I truncated the dendrites of all L4ss neurons in the vS1 model that were located within the slice.

Second, I grouped the neurons as reported by Koelbl et al. (2013): I selected all pairs of reconstructed IN morphologies and the *in silico* truncated dendrites of L4ss that met two criteria: (i) soma-to-soma distance of less than 150 µm, and (ii) a distance from the slicing planes between 20 µm to 140 µm (i.e., tissue depth). I repeated this grouping for all INs that were assigned a L4 barrel-confined inhibitory neuron (L4BIn) by Koelbl et al. (2013).

Comparison with empirical connectivity statistics: I calculated the predicted connectivity statistics (i.e., connection probability or the number of synapses per connection) for each measured connectivity using the grouping mentioned previously. I evaluated the comparison as for the pairs of excitatory neurons (for details, see Section 2.4.2). Briefly, I computed the Pearson's linear correlation coefficient R between the empirical and predicted mean connection probabilities and number of synapses, respectively. I computed the deviation dev_{SD} between each empirical and predicted mean connection probabilities (see Equation 2.12) as well as its percentile (see Equation 2.14). I derived $p_{extreme}$, i.e., the probability of observing a connection probability that is at least as extreme as the empirical one given the predicted connection probability distributions (see Equations 2.10 and

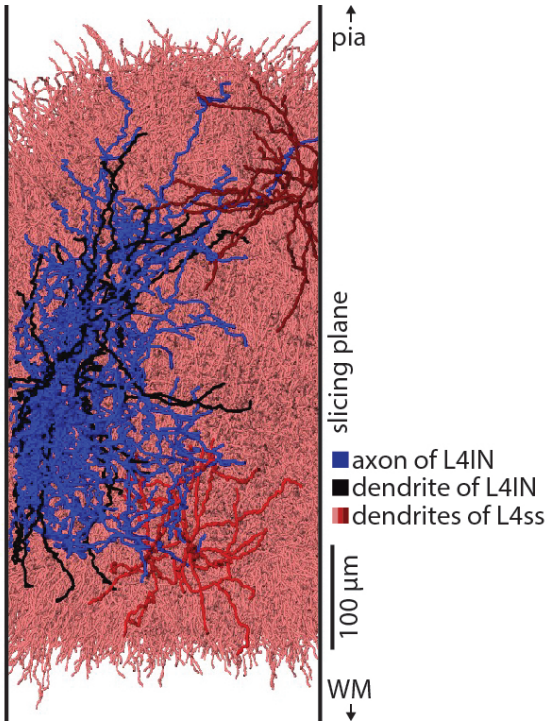

Figure 2.10. Emulation of an inhibitory slice experiment in L4. One exemplary *in vitro* L4IN morphology (axon in blue, dendrite in black) as reconstructed by Koelbl et al. (2013) located within its approximated slicing planes and integrated into the dense structural model. All surrounding L4ss dendrites in the dense structural model were truncated *in silico* given the slicing planes. Two exemplary truncated L4ss dendrites in the foreground (dark red), remaining truncated L4ss dendrites in the background (light red) as placed by the dense structural model.

2.15). I also performed a random permutation test for the correlation coefficient with 100,000 random permutations.

2.4.4 Network topologies

Various studies reported that the frequencies of observing particular motifs are unlike those expected in a random network (e.g., Milo et al., 2002; Song et al., 2005; Kampa et al., 2006; Perin et al., 2011; Rieubland et al., 2014; Jiang et al., 2015). In this Section, I describe how I repeated these experiments in the model and how I compared them to the empirical observations.

Triplet motifs: Two *in vitro* studies (Kampa et al., 2006; Song et al., 2005) empirically determined the deviations of triplet motif occurrences to those expected in a random network. I compared the empirical motif deviations to the predicted ones from the dense statistical connectome. All predictions were determined for triplets whose somata were located within the C2 barrel column. Kampa et al. (2006) empirically determined motif deviations for eight triplet motifs consisting of L2/3 pyramids and L5PTs. I mimicked their study by generating 10,000 triplets of two L5PT and one neuron located in L2/3, and 10,000 triplets of one L5PT and two neurons located in L2/3. For each triplet, I computed the probability of the empirically probed motifs. I repeated this procedure 10 times and computed the motif deviations. Song et al. (2005) empirically determined the motif deviations of each L5PT triplet motif. To mimic their study, I generated 10,000 triplets of L5PTs and computed the respective motif probabilities across all triplets. I repeated this procedure 10 times and computed the average motif deviations. I calculated the correlation coefficient between the predicted deviations and the empirically determined deviations in log scale with and without controlling for the observed frequencies of doublet motifs (for details, see Song et al., 2005). In addition, I determined the doublet motif spectrum in the model for all neurons grouped by layer and for all L5PTs of the C2 barrel column and compared those to a random network and their respective empirical values (Song et al., 2005; Lefort et al., 2009; Jouhanneau, Kremkow, Dorrn, & Poulet, 2015).

Network motifs between more than three neurons: Perin et al. (2011) reported 27 quadruplet motifs of L5PT that are more frequent than expected in a random network. I tested whether the dense statistical connectome predicts the same motifs to be overrepresented compared to a random network. To do so, I generated 10,000 quadruplets of L5PT and computed the probability of each of the 27 reported quadruplet motifs. I repeated this 10 times and computed for each of the 27 quadruplet motifs the motif deviation from a random network.

In the same study, Perin et al. (2011) probed the network topologies of motifs involving three to eight L5PTs. For each motif size, the motif deviation was determined as a function of the number of connected edges. To compare their distributions of motif deviations to the model's predictions, I randomly sampled

10,000 sets of neurons per motif size. For each set of neurons and each number of possible edges (ranging from 0 to e.g., 6 possible edges in a triplet motif or 56 possible edges in a 8-neuron motif) I computed the number of possible edge combinations (e.g., 1 possible combination of 0 or 6 edges but 20 possible combinations of 3 edges in a triplet motif). If the number of edge combinations was less than 5,000, I iterated over all possible combinations. If the number of combinations was larger than 5,000, I randomly generated 5,000 motifs that matched the number of edges. I computed the probability of each motif grouped by their number of edges in the dense statistical connectome as well as in a random network. For each motif size, I determined the distribution of motif deviations as a function of the number of edges.

2.4.5 Connectivity between *in vivo* labeled excitatory neuron pairs

In collaboration with my colleague Dr. David A. Slabik (research center caesar, Bonn), I compared distributions of empirical putative synaptic contacts between entire neuron morphologies (Slabik, 2018) with the prediction of the dense statistical connectome. By combining *in vivo* biocytin labeling and recording of multiple neurons with high-resolution confocal imaging, Dr. Slabik mapped putative synaptic contacts between completely reconstructed excitatory neuron morphologies in rat vS1. After the reconstruction of the neuron morphologies, an automated tool detected locations of close proximity between reconstructed axon and dendrites. By using a custom-designed 3D proof-editing environment, these locations were manually checked for putative synaptic contacts. A location was marked as a putative synaptic contact only if a spine and bouton were present. The resultant 3D location of putative synaptic contacts provides an upper limit of the number and location of synapses between the neuron morphologies (for more details on the methods, see Slabik, 2018). The false positive rate, i.e., the rate of putative synaptic contacts that are not synapses, is expected to be less than 15% based on a previous correlated light and electron microscopic approach (Schoonover et al., 2014). For readability's sake, I refer to the putative synaptic contacts determined by Slabik (2018) as empirical synapses.

Using the *BuildingBrains* web application (see next Section 2.5) all reconstructed excitatory neurons were registered into the dense structural model and their respective cell types were determined. To compare the empirical data to the model's prediction, I developed a tool that embeds any reconstructed and registered neuron pair (a,b) into the dense structural model of vS1 and computes its $\mathrm{DSC}(a, b, \hat{x})$. Specifically, the tool requires a presynaptic axon morphology a and its cell type, as well as the postsynaptic dendrite morphology b and its cell type. First, the tool computes the 3D bouton distribution of neuron a based on its cell type-specific bouton densities (see Table 2.2). Second, the tool computes the 3D spine density of neuron b based on its cell type-specific spine densities (see Table 2.3). Third, $\mathrm{DSC}(a, b, \hat{x})$ is calculated based on the 3D bouton distribution of a, the 3D spine distribution of b, and the total distribution of all POSTs in the dense structural model as defined by Equation 2.1. Finally, given $\mathrm{DSC}(a, b, \hat{x})$, other quantities such as $\mathrm{DSC}(a, b)$ and $p(a, b)$ are derived (see Equations 2.2 and 2.3). The predicted probability of the empirical number of synapses $n(a, b)$ is calculated, i.e., $p(a, b, n(a, b))$. The tool was applied to every registered neuron pair. The 3D locations of the empirical synapses were mapped onto the registered (i.e., transformed) neuron morphologies.

Comparison with empirical numbers of synapses

To compare the empirical numbers of synapses with the model's predictions, the empirical number of synapses per 50 µm cube $\hat{x}$ was determined, referred to as $n(a, b, \hat{x})$. Dr. David A. Slabik and I computed the Pearson's linear correlation coefficient R between the empirical $n(a, b)$ and the predicted $\mathrm{DSC}(a, b)$ on the cellular level (i.e., between all neuron pairs) and performed a linear regression fit. I repeated this analysis on the subcellular level between the empirical $n(a, b, \hat{x})$ and $\mathrm{DSC}(a, b, \hat{x})$ (i.e., between all cubes $\hat{x}$ across all neuron pairs where $\mathrm{DSC}(a, b, \hat{x}) > 0$). To assess the resultant correlation coefficient, I calculated the range of expected correlation coefficients when the synapse numbers were drawn instead from the dense statistical connectome. Specifically, I randomly drew $n(a, b, \hat{x})$ for each cube $\hat{x}$ across all neuron pairs (a, b) given the model's predicted probabilities (see Equation 2.2). The correlation between the model-drawn

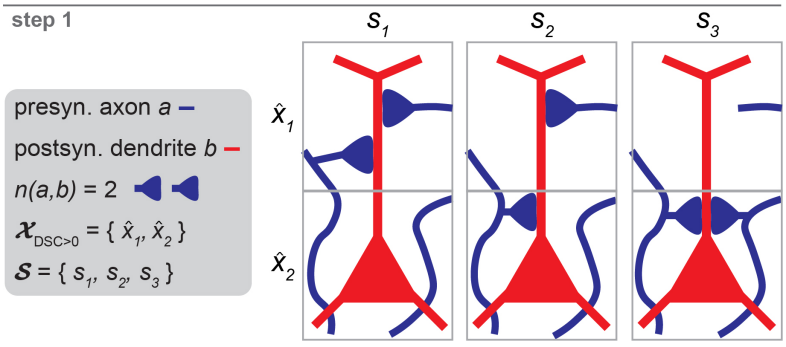

Figure 2.11. Toy example for comparing empirical vs. predicted synapse patterns between axon of presynaptic neuron a (blue) and dendrite of postsynaptic neuron b (red). Step 1: Determining the set of possible synapse patterns $\mathcal{S} = \{s_1, s_2, s_3\}$ of how two empirically observed synapses (blue triangles, $n(a, b)$) can be distributed in two cubes $\mathcal{X}_{\text{DSC}>0} = \{\hat{x}_1, \hat{x}_2\}$ (gray boxes), where axon of a overlaps with dendrite of b. Step 2: Computing the probability of each synapse pattern $p(a, b, s)$ and $p_{relative}(a, b, s)$. Step 3: Computing $prctl(s)$, i.e., summed probability of all synapse patterns that are equal or less likely than $p_{relative}(a, b, s)$.

$n(a, b, \hat{x})$ and $\text{DSC}(a, b, \hat{x})$ was determined. I repeated this step 100,000 times and compared the expected correlation coefficients with the correlation coefficient between the empirical $n(a, b, \hat{x})$ and $\text{DSC}(a, b, \hat{x})$.

Comparison with empirical synapse patterns

I compared the empirical synapse pattern with the model's predictions. For each neuron pair (a,b), I determined the set of possible synapse patterns (step 1), computed the predicted probability of each pattern in the model (step 2), and then computed the summed probability of all synapse patterns that are equal or less likely than the empirical one (step 3) (for a toy example, see Figure 2.11):

Step 1: Given the model's resolution limit of $(50\,\mu\text{m})^3$ (Egger et al., 2012), the model does not predict 3D locations of synapses but instead the probabilities of any number of synapses being located within any 50 µm cube. I determined the set

of cubes, denoted by $\mathcal{X}_{\mathrm{DSC}>0}$, where the axon of a and dendrite of b overlap (i.e., $\mathrm{DSC}(a, b, \hat{x}) > 0$). Given the set $\mathcal{X}_{\mathrm{DSC}>0}$ and the empirical number of synapses $n(a, b)$, I calculated all possible combinations of cubes where $n(a, b)$ synapses could be located in[14]. For readability's sake, I refer to the combinations of cubes as the set of possible synapse patterns $\mathcal{S}$. To compute $\mathcal{S}$, I used the Matlab built-in function `nchoosek`. The number of possible synapse patterns $|\mathcal{S}|$ (i.e., the cardinality of $\mathcal{S}$) is given by:

$$|\mathcal{S}| = \frac{(|\mathcal{X}_{\mathrm{DSC}>0}| + n(a, b) - 1)!}{n(a, b)! \cdot (|\mathcal{X}_{\mathrm{DSC}>0}| - 1)!} \tag{2.17}$$

where $|\mathcal{X}_{\mathrm{DSC}>0}|$ denotes the number of cubes where $\mathrm{DSC}(a, b, \hat{x}) > 0$. In the toy example in Figure 2.11 a neuron pair is connected by two synapses ($n(a, b) = 2$) and overlaps in two cubes ($|\mathcal{X}_{\mathrm{DSC}>0}| = 2$). Thus, there are three possible combinations of cubes ($|\mathcal{S}| = 3$) in which two synapses can be located in including the two combinations where all synapses are located in the same cube. I excluded two neuron pairs, each with seven putative contacts from further analysis due to insufficient available memory on the computer performing the analysis.

Step 2: I iterated over the set of possible synapse patterns $\mathcal{S}$. For each synapse pattern s, that is defined by $n_s(a, b, \hat{x})$, I determined its (absolute) probability $p(a, b, s)$:

$$p(a, b, s) = \prod_{\hat{x} \in \mathcal{X}_{\mathrm{DSC}>0}} p(a, b, \hat{x}, n_s(a, b, \hat{x})) \tag{2.18}$$

Next, $p_{relative}(a, b, s)$, the relative probability of each synapse pattern was calculated:

$$p_{relative}(a, b, s) = \frac{p(a, b, s)}{\sum_{i \in \mathcal{S}} p(a, b, i)} \tag{2.19}$$

yielding the probability of each synapse pattern given that the neuron pair (a,b) was connected by $n(a, b)$ synapses[15].

[14] Combinations with repetition to allow multiple synapses in one cube

[15] $p_{relative}(a, b, s)$ is a conditional probability, i.e., the probability of synapse pattern s under the condition of $n(a, b)$ synapses.

Step 3: I sorted $p_{relative}(a, b, s)$ of all synapse patterns in ascending order and determined their cumulative probabilities. The set of all synapse patterns whose $p_{relative}(a, b, s)$ was smaller or equal than $p_{relative}(a, b, s_{Empirical})$ was determined, referred to as $\mathcal{S}_1$. $s_{Empirical}$ denoted the empirical synapse pattern. I calculated $prctl(s_{Empirical})$, the predicted probability of observing either $s_{Empirical}$ or any less likely pattern:

$$prctl(s_{Empirical}) = \sum_{s \in \mathcal{S}_1} p_{relative}(a, b, s) \tag{2.20}$$

A $prctl(s_{Empirical})$ of close to 100% indicates that $s_{Empirical}$ is one of the most likely synapse patterns given the empirical number of synapses $n(a, b)$. In addition, $rank(s_{Empirical})$ was determined, i.e., the percentage of synapse patterns that have a larger probability than $s_{Empirical}$. To assess whether the empirical synapse numbers $n(a, b)$ impacts the model's predictions, the correlation coefficient R between $n(a, b)$ and $prctl(s_{Empirical})$ as well as between $n(a, b)$ and $rank(s_{Empirical})$ was determined.

2.5 *BuildingBrains* web application

In collaboration with Philipp Harth, Ela Elsholz, Dr. Vincent J. Dercksen, and Prof. Dr. Hans-Christian Hege (Zuse Institute Berlin), many tools to process and analyze reconstructed neuron morphologies that were developed in the group of Dr. Marcel Oberlaender were incorporated into a web application called *BuildingBrains*[16] (see Figure 2.12). My contribution to this collaboration was: I adapted the tools for integration into the web application, I wrote the necessary documentation and manuals of how to use the tools, and I tested the web application and the employed tools. In the following, I summarize the tools employed by *BuildingBrains*. A detailed description is available on the web page. The implementation of the web server is not described here as it was implemented by Philipp Harth, Ela Elsholz, and Dr. Vincent J. Dercksen (Zuse Institute Berlin). Note that *BuildingBrains* is an ongoing effort. At the time of this thesis, the tools employed

[16] not publically available yet

by *BuildingBrains* are limited to rat vS1. However, in the future *BuildingBrains* will be extended to other cortical areas such as the vibrissal part of the motor cortex.

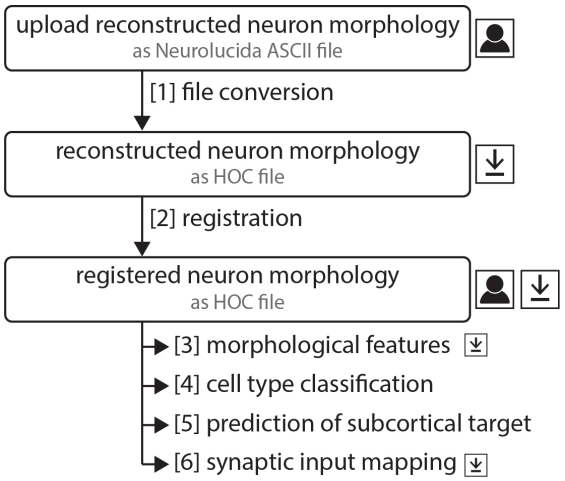

Figure 2.12. BuildingBrains pipeline. After the user uploaded a reconstructed neuron morphology, *BuildingBrains* applies a set of tools to the neuron morphology: [1] Uploaded neuron morphology is converted to the `HOC` file format using *NeuroConv* (Lang et al., 2011). [2] Neuron morphology is registered based on user-specified settings and placed into the geometrical model of vS1 (Egger et al., 2012). [3] Morphological and topological parameters are extracted (Narayanan et al., 2015). [4] Neuron morphology is automatically assigned to the most likely morphological cell type (Narayanan et al., 2015). [5] If the uploaded neuron morphology is of cell type L5PT, its most likely subcortical projection target is predicted (Rojas-Piloni et al., 2017). [6] Synapses are mapped onto the neuron morphology (Egger, 2015). User icon represents possible user interaction, download icon represents download option on web application.

BuildingBrains enables every registered user to upload a neuron morphology and specify meta-information about the uploaded neuron morphology, such as animal species, animal age, or software used for reconstruction. Reconstructed neurons have to be uploaded as `ASCII` files, as specified by the Neurolucida tracing software[17]. After the upload, the Neurolucida `ASCII` is converted to a `HOC`[18] file format using the tool *NeuroConv* developed by Lang et al. (2011). The neuron

[17] https://www.mbfbioscience.com/neurolucida
[18] compatible to the NEURON simulation environment (https://neuron.yale.edu/neuron/)

is visualized in an embedded 3D viewer. Next, *BuildingBrains* employs the registration routine developed by Egger et al. (2012) that precisely places the uploaded neuron into the geometrical model of vS1. The user can specify what registration strategy is used. The neuron morphology is registered to two barrel columns, its home column (i.e., the barrel column closest to its soma) and to the D2 barrel column. The registered neuron morphologies are visualized in an embedded 3D viewer. In the next step, *BuildingBrains* extracts morphological and topological parameters and automatically assigns the neuron to the most likely morphological cell type based on the parameters extracted for the D2 registered neuron morphology (Narayanan et al., 2015). The subcortical projection target of neurons that were classified as L5PT is predicted using a tool developed by Rojas-Piloni et al. (2017). In the last step, *BuildingBrains* embeds the neuron morphology into the dense structural model of vS1 and infers its synaptic inputs based on the dense statistical connectome (see Section 2.2 and Egger et al., 2014; Egger, 2015). Specifically, $\mathrm{DSC}(A, b, \hat{x})$ is calculated between the set of all presynaptic neurons A and the uploaded and registered neuron morphology b in all 50 µm cubes $\hat{x}$ that the dendrite of neuron b innervated. Based on the probabilities provided by Equation 2.2, the number of synapses each presynaptic neuron a forms with the uploaded neuron b was randomly drawn for each cube $\hat{x}$. The drawn number of synapses were randomly distributed on dendrite branches of b that innervated the respective cube. The predicted synapse inputs thus represent one network realization and are tailored for network embedded simulations of the uploaded neuron morphology b. The tool to map synapses onto an individual neuron morphology was developed by my colleague Dr. Robert Egger (Max Planck Institute for Biological Cybernetics, Tübingen) (Egger, 2015). The data and results of each step are available for download. In addition to mapping synapses onto neuron morphologies, *BuildingBrains* was designed to serve as a repository of reconstructed neuron morphologies and their structural information that can be shared and explored by everyone.

2.6 *Cortex in Silico* web application

In collaboration with Philipp Harth, Dr. Vincent J. Dercksen, and Prof. Dr. Hans-Christian Hege (Zuse Institute Berlin), both models, the dense structural model

and the dense statistical connectome, and their predictions will be made available to the public. The models and many of the computational routines that I developed (see Section 2.1 and 2.3) were incorporated into an interactive web application called *Cortex in Silico*[19]. Beyond access to the model and its predictions, *Cortex in Silico* provides the opportunity to formulate one's own synapse formation strategy mathematically and explore its impact on network architecture. In the following, I describe the current functionalities of *Cortex in Silico*. The implementation of the web server is not described here as it was implemented by Philipp Harth and Dr. Vincent J. Dercksen (Zuse Institute Berlin). The web application and application examples are described in Results Section 3.5.

The web application is divided into two major sections: "Digital Barrel Cortex Anatomy" and "In Silico Connectivity Experiments". In the section "Digital Barrel Cortex Anatomy" the user can explore the dense structural model, its structural estimates, and the empirical data used to generate the model. In the section "In Silico Connectivity Experiments" the user can perform various *in silico* experiments and thereby explore the connectivity predicted by the dense statistical connectome. Here, the user can also specify one's own mathematical rules for the synapse formation strategy used by the dense statistical connectome to predict connectivity (cf. Section 2.2) and compare different rule-based connectivity statistics.

Note that *Cortex in Silico* is an ongoing effort. The descriptions reflect the web application at the time of the thesis. The following sections are in parts adapted from Udvary, Dercksen, Harth, Hege, and Oberlaender (2020).

2.6.1 Digital Barrel Cortex Anatomy

In the "Digital Barrel Cortex Anatomy" section of the web application, the empirical data used to generate the dense structural model, as well as the structural estimates of the model, are presented. The section is subdivided into the four tabs, each representing the four components used to generate the dense structural model (geometrical, cellular, morphological, and pre- and postsynaptic component; see Section 2.1). Each of these components is shown in an embedded 3D viewer and

[19] not publically available yet

the user can apply specified filters ("selections") to interact with the respective structures displayed in the viewer. Data tables below the viewer are updated according to the user selection and show both empirical and model parameters. In the download area the user can download either the entire model or selected parts of it. In the following, I summarize the possible user interactions for each component or tab, respectively. A detailed description of how each component was used to generate the dense structural model is provided in Section 2.1.1.

Cortical geometry: This tab describes the generation of the geometrical model of vS1 (Egger et al., 2012) plus its laminar organization as derived from the empirical 3D soma density (Meyer et al., 2013). The user can interactively explore the geometrical model in an embedded 3D viewer. The user can apply two spatial filters selecting one layer (L1, L2, L3, L4, L5, or L6) and/or one of the 24 cortical barrels (A1-E4 or α-δ). The selected regions are highlighted in the 3D viewer (see Figure 2.13). Data tables below the viewer show (i) the empirical data used to generate the geometrical model of vS1, such as the barrel volume or column height as reported by Egger et al. (2012), and (ii) the 3D coordinates, radii, and orientations of the cortical barrels within the model.

Cellular composition: This tab describes the cellular composition of the vS1 model (Meyer et al., 2013). The user can interactively explore the 3D distribution of excitatory and inhibitory neuron somata as measured by Meyer et al. (2013) in an embedded 3D viewer. The user can apply the two spatial filters introduced in the previous paragraph (i.e., layers and barrels; see Figure 2.14). The excitatory neuron density in the respective selected regions is visualized accordingly. Data tables below the viewer show (i) the empirical neuron numbers per cortical layer and barrel column in vS1 or barreloid in VPM thalamus (Meyer et al., 2013), and (ii) the neuron counts per cortical layer and barrel column or barreloid in the model.

Morphological model: This tab describes the sample of morphologies that were used to populate the vS1 model (Narayanan et al., 2015; Oberlaender, Ramirez, & Bruno, 2012). The user can interactively explore the *in vivo* labeled and recon-

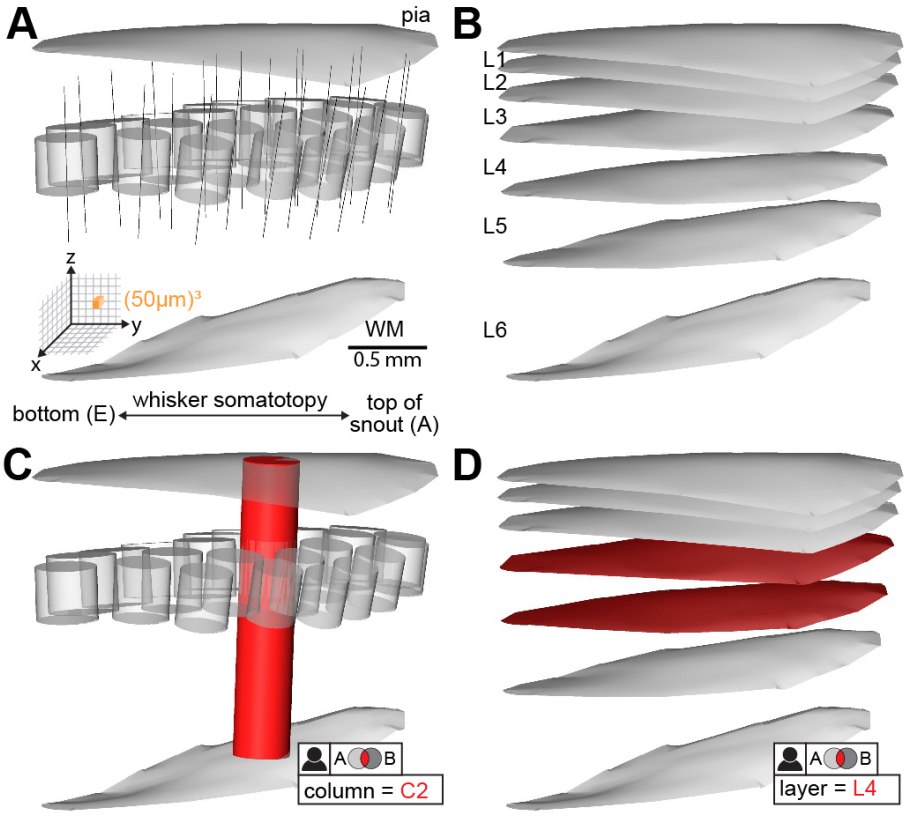

Figure 2.13. Cortex in Silico: Geometry of vS1 model. (A) Location, orientation, and diameter of all 24 cortical barrels and pial and white matter surfaces. Lines represent column axis, grid and orange 50 µm cube represent interanimal variability of anatomical landmarks (i.e., resolution limit) (B) Surfaces of layer borders. (C, D) User can apply spatial filters that highlight the respective region in red (C: C2 barrel highlighted; D: L4 borders highlighted).

structed excitatory neuron morphologies placed in the vS1 model in an embedded 3D viewer (see Figure 2.15). The user can select neurons based on their closest cortical barrel column, barreloid, and/or cell type[20]. For VPM neurons, the intracortical part of their axons was incorporated into the model (Oberlaender, de Kock, et al., 2012; Oberlaender, Ramirez, & Bruno, 2012). One representative neuron morphology per cell type is shown in the embedded 3D viewer. Data tables below the viewer show morphological parameters of the reconstructed neurons, such as mean branch lengths and cortical depths for the respective cell type.

[20] Ten morphological cell types plus VPM

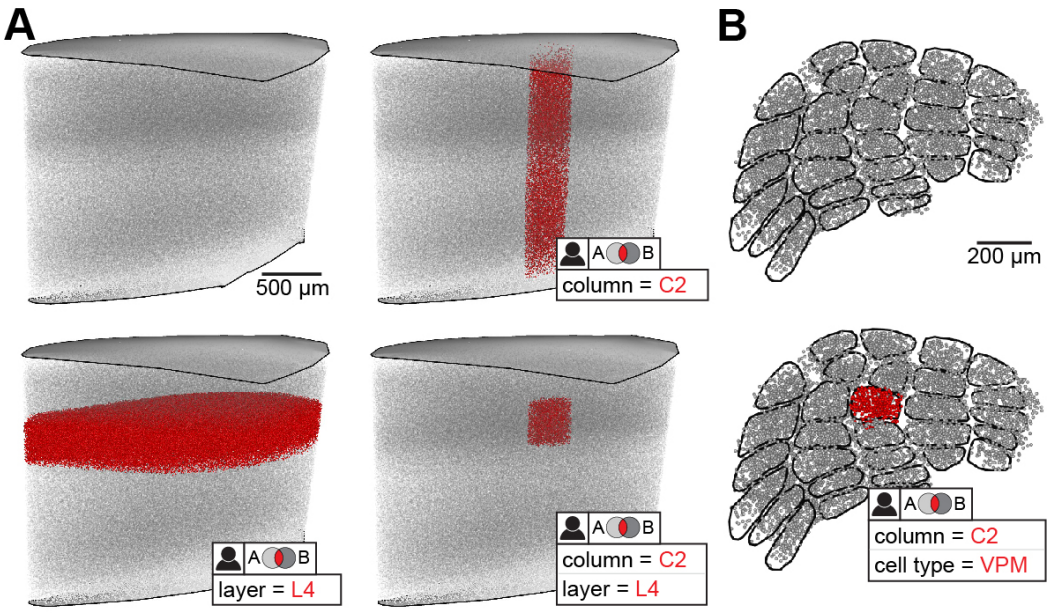

Figure 2.14. Cortex in Silico: Cellular composition of vS1 model. (A) 3D distribution of excitatory somata in vS1 model. User can apply spatial filters and combinations thereof, such as cortical barrel column (upper right), layer (lower left), and layer within barrel column (lower right). Respective selections highlighted in red. (B) Exemplary 3D soma distribution in VPM thalamus. User can filter by VPM barreloid.

Subcellular model: The last tab of "Digital Barrel Cortex Anatomy" describes how the densities of pre- and postsynaptic structures (e.g., excitatory bouton and spine densities) were determined in the vS1 model (Narayanan et al., 2015; Egger et al., 2014; Egger, 2015). In the embedded 3D viewer the user can interactively explore the 3D distribution of boutons and/or spines that match one's filters. As in the previous tab, neurons can be filtered based on their closest cortical barrel column, barreloid, and/or cell type (see Figure 2.16). Data tables below the viewer show the respective densities of boutons, spines, and postsynaptic structures along inhibitory dendrites and somata.

2.6.2 *In Silico* Connectivity Experiments

In the "*In Silico* Connectivity Experiments" section of the web application, the user can perform four different types of *in silico* experiments targeting user-defined

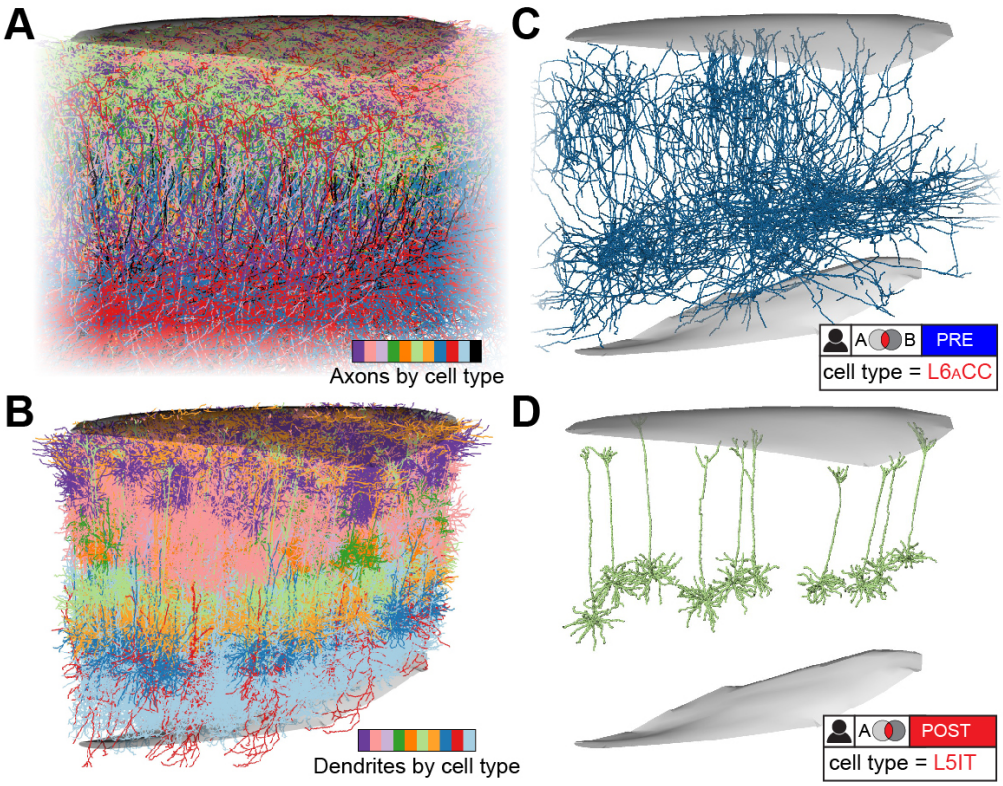

Figure 2.15. Cortex in Silico: Morphologies of vS1 model. (A,B) Representative sample of axon (A) and dendrite (B) morphologies, color-coded by cell type. (C,D) User can apply cell type-specific filter to visualize the particular cell type(s). Representative samples of axons of $L6_ACC$ (C) and dendrites of L5IT (D) shown.

regions and neurons and thereby explore the connectivity and structure predicted by the model. The experiments provide access to predictions at the subcellular, cellular, and network level and allow replication and/or modification of many of the connectivity analysis described in Sections 2.3 and 2.4. The general procedure to perform an *in silico* experiment is as follows (see Figure 2.17; details in the following sections):

1. User chooses what type of *in silico* experiment to perform.

2. User selects the neuron population(s) of interest by applying one or multiple filters.

3. User chooses the model configurations applicable to the type of experiment

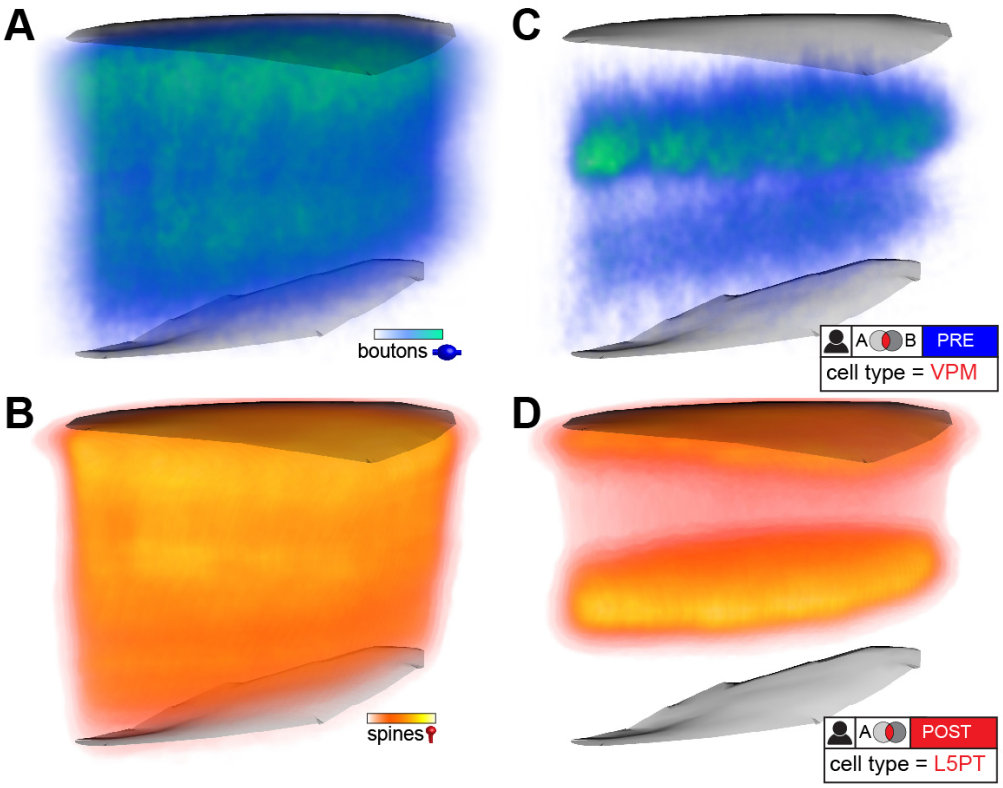

Figure 2.16. Cortex in Silico: Subcellular components of vS1 model. (A,B) 3D density distribution of excitatory boutons (i.e., PREs; A) and spines (i.e., POSTs, B). (C,D) same as in panels A and B but for boutons of VPM axons (B) and spines of L5PT (C).

(i.e., experimental conditions).

4. User may additionally specify the synapse formation strategy used for predicting connectivity.

5. User performs the *in silico* experiment.

After running an *in silico* experiment, the results are shown as a data table and as charts in the web application. The selection is visualized in an embedded 3D viewer and the results can be downloaded. The section "*In Silico* Connectivity Experiments" of the web application offers a dedicated tab to replicate many results of this thesis and of Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020) with the respective predefined selections and model configurations. In the following, I

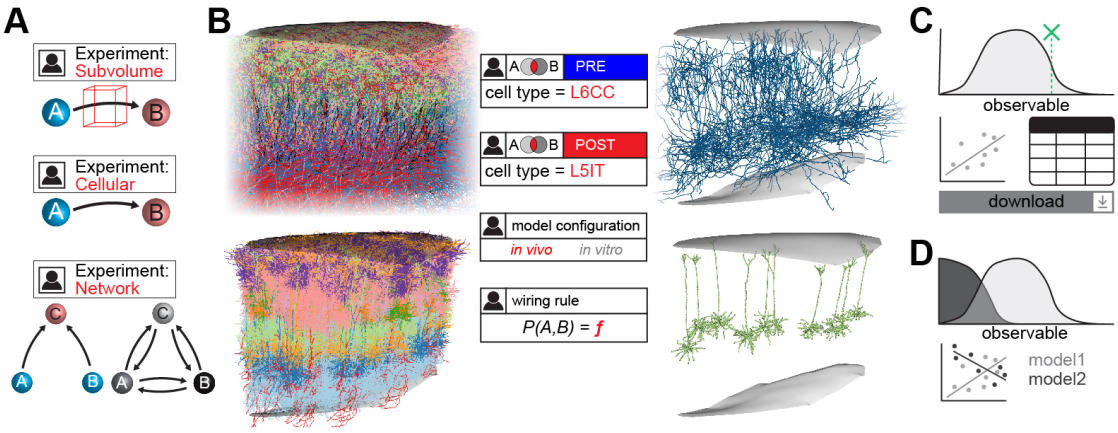

Figure 2.17. Cortex in Silico. (A) The user chooses what type of experiment to perform *in silico*. (B) The user selects the neuron population(s) of interest, chooses the applicable model configurations, and specifies the synapse formation strategy. (C) Results of the selection are shown as a data table and as charts. Predicted observables, such as connection probabilities, can be compared to empirical data. All the results can be downloaded. (D) Results can be compared between two different models (e.g., *in vitro* vs. *in vivo* predictions or two different synapse formation strategies).

explain the four different types of *in silico* experiments.

Types of *in silico* experiments

Subvolume experiment: A subvolume experiment provides access to the predicted structural composition and synaptic contacts in a user-defined subvolume of the vS1 model (see Sections 2.1.3 and 2.4.1). The user can define a subvolume X by cortical column, cortical depth, and/or by directly specifying the coordinates of its 3D bounding box. The resulting bounding box of X is visualized in an embedded 3D viewer. In addition, the user can restrict the pre- and postsynaptic populations A and B examined in subvolume X by applying one or multiple filters (listed later in this Section; see Figure 2.18). Based on the filters, all structures that match the selected neuron populations A and B and are located in subvolume X are extracted from the vS1 model. A data table displays the volume of subvolume X, number of cell bodies in X, cell type diversity[21] of cell bodies in

[21] cell type diversity denotes the number of cell types that contribute structures to a cube; see Methods Section 2.1 for details.

X, number of pre- and postsynaptic neurons that innervate X with their axons or dendrites, respectively, number of axon and dendrite branchlets in X, cell type diversity of axon and dendrites in X, length of axon and dendrite branchlets in X, and the total number of synapses in X. All predictions are with respect to the user-specified neuron populations A and B. The predictions are visualized as histograms, such as the number of cell bodies, and axon and dendrite branchlets per 50 µm cube. Synapse densities per 50 µm are shown as a box plot and the frequency of synapses per connected neuron pair as a line graph.

Cellular experiment: A cellular experiment provides access to $\mathrm{DSC}(A, B)$ and its derived connection probabilities $P(A, B)$ (see Sections 2.2 and 2.3.1). The user can define the pre- and postsynaptic neuron populations A and B by using the filters listed later in this Section (see Figure 2.18). The somata of both neuron populations A and B are visualized in two embedded 3D viewers. A data table below the viewer shows the cell type- and barrel column-specific number of neurons that match the applied filter(s). The number of neuron pairs, the number of neuron pairs that do not overlap (i.e., $\mathrm{DSC}(a, b) = 0$), and summary statistics of $\mathrm{DSC}(A, B)$ and $P(A, B)$ are shown in a separate data table and additionally visualized as histograms. $\mathrm{DSC}(A, B, \hat{x})$ is offered for download and visualized in an embedded 3D viewer.

In-degree correlation: An in-degree experiment provides access to the correlations between the in-degrees of two presynaptic neuron populations A and B onto individual neurons c of the postsynaptic population C (see Section 2.3.2). The user can define the three neuron populations A, B, and C by the filters listed later in this Section (see Figure 2.18). The somata of the three selected neuron populations are visualized in three embedded 3D viewers. A data table below the viewer shows the respective cell type- and barrel column-specific number of neurons per selected population. In a separate data table the number of postsynaptic neurons C, summary statistics of the in-degree (i.e., $\mathrm{inDegree}(A, c)$ and $\mathrm{inDegree}(B, c)$) and their respective correlation coefficient R are displayed. The results are visualized in a scatter plot where each data point reflects $\mathrm{inDegree}(A, c)$ and $\mathrm{inDegree}(B, c)$ of each postsynaptic neuron c. Optionally, the mean connection probabilities across

populations A and B per neuron c (instead of the in-degrees) can be used for the correlation calculation.

Triplet motif experiment: An triplet motif experiment provides access to the occurrence probabilities of the 15 possible connected triplet motifs (see Section 2.3.3). The user can define three neuron populations A, B, and C by the filters listed later in this Section (see Figure 2.18). The somata of the three selected neuron populations are visualized in three embedded 3D viewers. A data table below the viewer shows the respective cell type- and barrel column-specific number of neurons per population that match the filter(s). In a separate data table the probability and deviation from a random network for each triplet motif are displayed. The same results are visualized as bar charts, where each bar represents the probability or deviation of each triplet motif. By default, the calculation is performed with 3,000 random triplets drawn from the specified neuron populations; the number of samples and a random seed can be freely specified by the user.

User-specified selection of neurons

The user can select populations of neurons by applying different filters (see Figure 2.18). Possible filters in the *in silico* experiments of *Cortex in Silico* are:

- *Cortical layer*: selects neurons whose somata are located in one of six predefined layers (L1 to L6; layer borders defined by Meyer et al., 2013).
- *Cortical depth*: selects neurons whose somata are within a user-specified range of cortical depth. This filter option allows more flexibility than filtering by predefined cortical layers.
- *Home column*: selects neurons whose somata are closest to the user-specified cortical barrel column in vS1 or barreloid in VPM.
- *Inside column*: selects only neurons whose somata are either located within the barrel column or outside of the barrel column (i.e., surrounding septum).

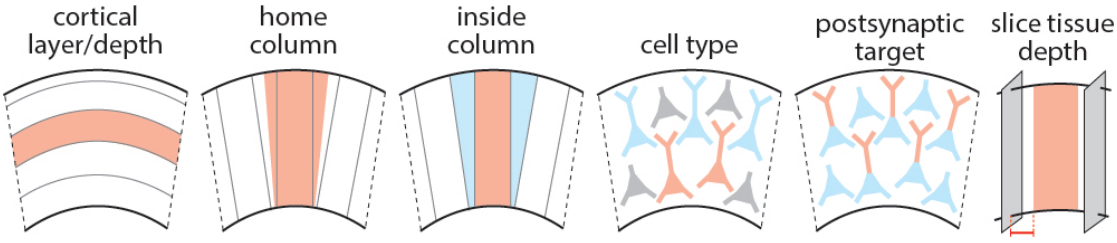

Figure 2.18. Cortex in Silico: Illustration of selectables. The user can filter neurons based on their soma position with respect to a cortical layer, a user-defined cortical depth, the closest barrel column ("home column"), and/or whether the soma is located within the barrel column (red) or in the area between the barrel columns (blue; called septa). The user can select neurons based on their cell type. Postsynaptic selections can be restricted to the postsynaptic target region, i.e., apical dendrite (red), basal dendrite (blue), or the entire dendrite (red and blue). When emulating a slice, neurons can be restricted to a user-defined slice tissue depth, i.e., the shortest distance between the somata and the slicing surface.

- *Cell type*: selects neurons of one of the ten morphological cell types (Narayanan et al., 2015), VPM thalamus, or inhibitory neurons.
- *Postsynaptic target*: selects either the entire dendrite, only the basal dendrite, or apical dendrite as postsynaptic structures. For example, this filter allows the user to explore subcellular/compartment-specific connectivity. This option is only applicable to postsynaptic neuron selections.
- *Slice tissue depth*: selects neurons whose somata are within a user-specified range of slice tissue depth, i.e., shortest distance between somata and the slicing surface. This option is only available when emulating a slice (described in the following Section).

In the subvolume experiments, the user can additionally define a bounding box of a desired subvolume[22].

Creating and comparing different model configurations

Cortex in Silico allows the user to compare two model configurations for the same *in silico* experiment. The results of both model configurations are visualized as

[22] see the previous paragraph about the subvolume experiment on page 78 for more details.

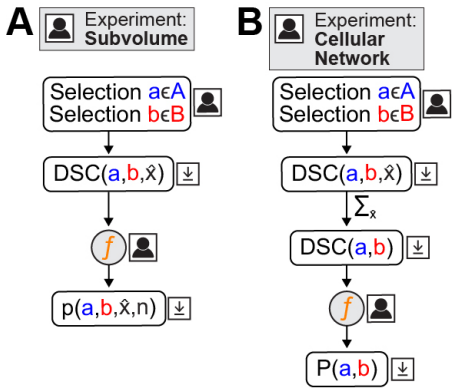

Figure 2.19. Cortex in Silico: Current options to specify the synapse formation strategy. (A) In *in silico* subvolume experiments the user can define a formula f that takes $\mathrm{DSC}(A, B, \hat{x})$ as an input and which output represents $P(A, B, \hat{x}, n)$. n reflects the number of synapses and is a predefined variable in the formula parser. (B) In *in silico* experiments at the cellular and network level the user can define a formula f that takes $\mathrm{DSC}(A, B)$ as an input and which output represents $P(A, B)$. User icon represents possible user interaction, downward arrow represents download option of data.

two-color-coded charts and in two (or more) embedded 3D viewers showing the respective selected neuron populations. The data tables show the results of both model configurations next to each other.

First, the user can choose between the full "*in vivo*" model of vS1, a downsampled "*in vivo*" model of vS1 (by a user-specifiable factor)[23], and among 10 thalamocortical slice models that emulate *in vitro* tissue slicing conditions by truncating the neuron morphologies within the selected slice. The ten slices are located at different positions within the center of vS1, each containing either the entire C2 barrel column or parts of it. They are the same 10 slice models as described in Section 2.3.1 used for the empirical comparisons (see also Figure 2.5). Emulation of a slice experiment is possible for all *in silico* experiments except for the subvolume experiment. When emulating a slice experiment, an additional filter is available allowing the user to select neurons based on their soma location within the slice (i.e., tissue depth).

Second, the user can specify one's own synapse formation strategy (see Figure

[23] *in silico* experiments are performed over a subsample of neuron somata and morphologies

2.19). By default, the synapse formation strategy described in Section 2.2 is used. Currently, the user can change the strategy by redefining the relationship between $\mathrm{DSC}(a, b, \hat{x})$ and $p(a, b, \hat{x}, n)$ (see Equation 2.2) or between $\mathrm{DSC}(a, b)$ and $P(a, b)$ (see Equation 2.3) as a mathematical formula. The user can only specify the mathematical formula for $p(a, b, \hat{x}, n)$ in the *in silico* subvolume experiment, while the mathematical formula for $P(a, b)$ can only be specified in the *in silico* experiments at the cellular and network levels. The underlying formula parser uses the Mathematical Expression Toolkit library[24] and was implemented by Philipp Harth (Zuse Institute Berlin). Depending on the user's specification, the assumptions underlying the synapse formation strategy change (see Section 2.2). Changing the synapse formation strategy affects all quantities derived from either $p(a, b, \hat{x}, n)$ or $P(a, b)$, such as synapse counts and connection probabilities. Note that at this stage, in-degrees are not affected since they are derived directly from the DSC.

[24] http://www.partow.net/about/index.html

3 | Results

3.1 Model of dense structural composition of rat barrel cortex

The generated dense structural model comprises 477,537 excitatory and 69,810 inhibitory neurons in vS1, and 6,225 neurons in the VPM (see Figure 3.1; see Table 3.1 for cell type-specific numbers of somata). In the model, the excitatory neurons give rise to 3.6 km of dendrite length and 33.0 km of axon length with a total of 5.8 billion boutons. In this Section, I describe the structural estimates provided by the dense structural model of vS1, their robustness, and how they compare to empirical data. The first four Sections 3.1.1, 3.1.2, 3.1.3, and 3.1.4 describe the resultant structural estimates for excitatory neurons (Narayanan et al., 2017; Egger et al., 2020; Udvary, Dercksen, Harth, Hege, de Kock, et al., 2020). The last Section 3.1.5 describes the results of the structural estimates for INs.

The most important results of the are:

- The interanimal variability of the empirical data used to constrain the dense structural model is low: The map of the barrel field, as well as the respective depths of layer borders within each barrel column, vary on average between animals by less than ±50 µm, providing the resolution limit of the dense structural model. Neuron densities across the barrel cortex vary by less than 5%.
- The sample of excitatory neuron morphologies is representative for each cell type, capturing ∼96% of the interanimal variability with respect to cell type-

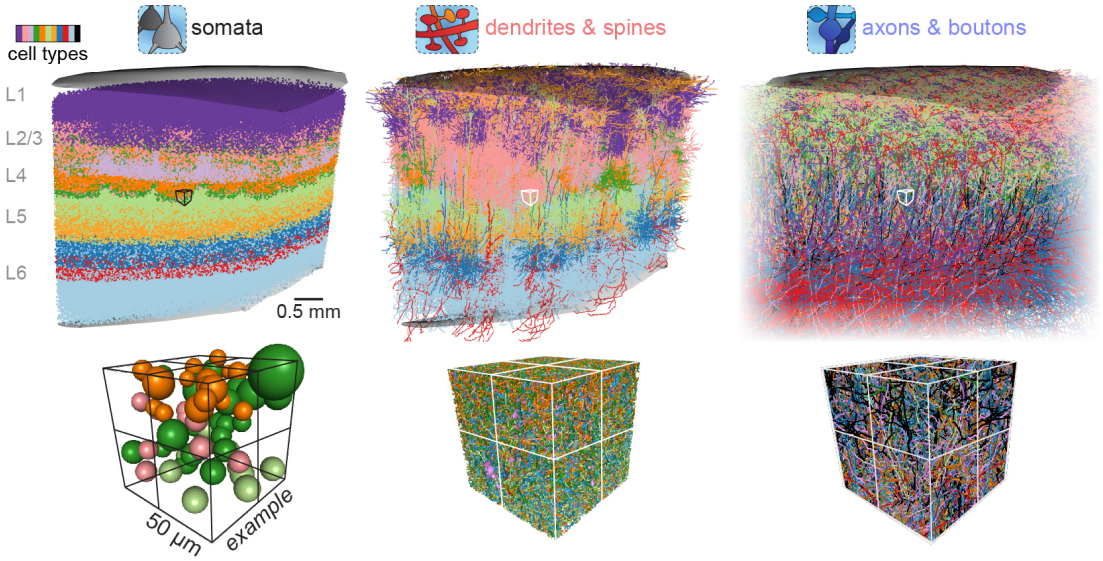

Figure 3.1. Empirically constraint model of the dense structural composition of the neuropil in rat vS1. Dense structural model provides estimates of the numbers and spatial distributions of somata (left), dendrites and spines (center), and axons and boutons (right) at a resolution of 50 µm cubes. Bottom panel show zoom-ins to eight exemplary 50 µm cubes at the border between L4 and L5 within the barrel column representing the C2 facial whisker. Colors denote the ten major excitatory cell types of vS1 and axons from the ventral posterior medial nucleus in the thalamus (VPM) in black. Figure adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020).

specific dendrite distributions, and ∼88% with respect to axon distributions.

- The upscaling of single neuron morphologies in the dense structural model provides robust estimates of the packing densities and cell type diversities of axons and dendrites within cubes of 50 µm edge length, capturing more than 90% of the dense structural composition that is preserved across animals.

- The dense structural model estimates a densely and diversely packed neuropil. Comparisons with empirical data revealed that ratios of axon/dendrite branchlets and layer-specific excitatory synapse densities are consistently estimated.

- Analysis of the dense structural model revealed that excitatory neurons are organized with respect to excitatory soma densities (i.e., in a layer-specific

manner), while inhibitory neurons are organized with respect to inhibitory soma densities.

3.1.1 Robustness of dense structural model

Note: The following paragraphs are in part adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020).

Variability of empirical constraints of model: The dense structural model was constraint with empirical data from multiple rats. First, I investigated the variability of these empirical constraints across animals as determined in previous studies (Egger et al., 2012; Meyer et al., 2013; Narayanan et al., 2015) (see Figure 3.2). The precision of the barrel field representing the 24 major facial whiskers along the animal's snout and cortical thickness per barrel column varies on average between animals by less than ±50 µm ($n = 12$) (Egger et al., 2012). The C2 and D2 barrel column — used in most of the analysis — varies in the precision of the barrel field by 66 µm and 93 µm, respectively, and in cortical thickness by 100 µm and 101 µm, respectively. The precision of the layer borders per barrel column as measured by Meyer et al. (2013) varies between animals on average by less than

Table 3.1. Cell type-specific number of excitatory somata in dense structural model.

Cell type	Number of somata
L2PY	61,369
L3PY	109,854
L4PY	19,940
L4sp	32,008
L4ss	47,214
L5IT	37,572
L5PT	28,765
$L6_ACC$	32,845
$L6_BCC$	18,204
L6CT	89,766
Total	477,537

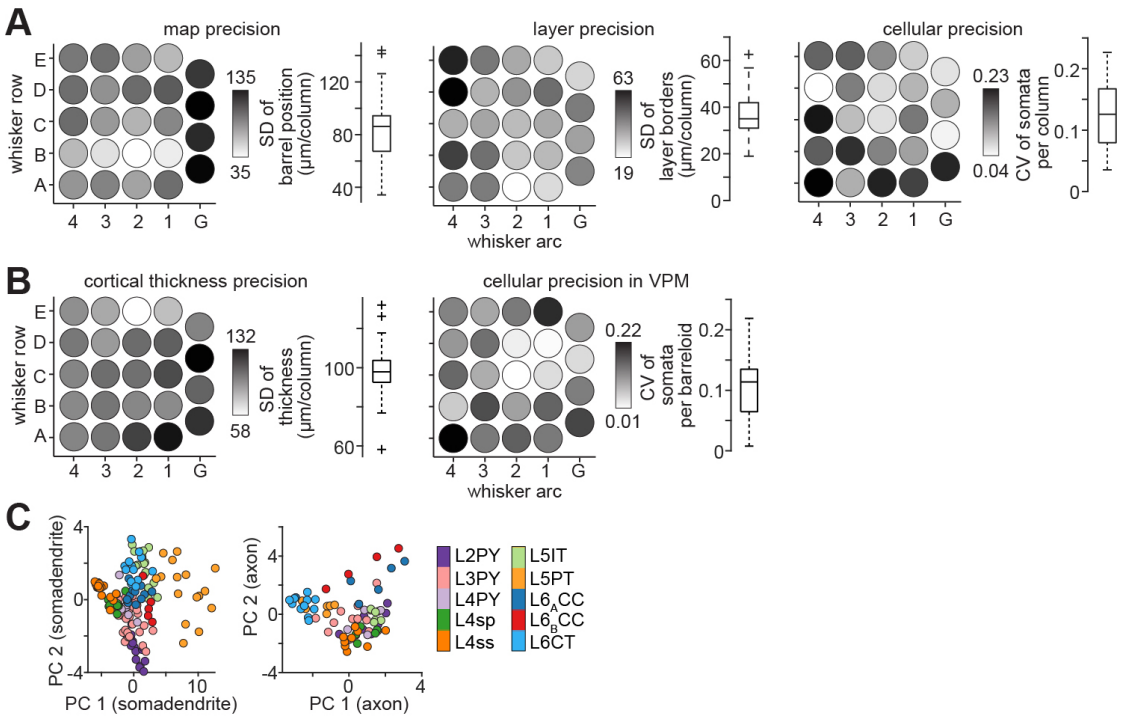

Figure 3.2. Interanimal variability of rat vS1. (A) Variability of barrel field map and layers in terms of SD of barrel and layer positions per barrel column (left and center panel). Variability of cellular organization in terms of CV of somata per barrel column (right panel). Panel A adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020). (B) Variability of cortical thickness per barrel column (left) and number of somata in VPM thalamus per barreloid (right). (C) Structural properties of dendrites and axons in the principal component space color-coded by cell types (Narayanan et al., 2015). Panels represent reanalysis of data reported in Egger et al. (2012); Meyer et al. (2013); Narayanan et al. (2015).

38 µm. In the C2 and D2 barrel column layer borders vary on average by 31 µm and 38 µm, respectively. Given the spatial precision to which the columnar and laminar layout of the barrel cortex is preserved across animals, the 6.4 mm^3 large model volume of vS1 was subdivided into cubes with 50 µm edge length. The model captures ∼82% of the interanimal variability with respect to the whisker- and whisker row-specific columnar and laminar cellular organization of vS1 (see Figure A1). The number of somata per barrel column in vS1 and per barreloid in VPM varies between animals on average by less than 13% ($n = 4$) (Meyer et al., 2013). In the C2 and D2 barrel column of vS1 soma numbers vary both by 6% and in the respective barreloids of VPM by less than 3%. The overall neuron density varies by less than 5% (Meyer et al., 2013). The structural properties of the intracortical neurons' dendrites ($n = 153$) and axons ($n = 74$) vary between cell types and to a lesser degree also within cell types, reflecting the morphological diversity across and within cell types (Narayanan et al., 2015).

Robustness of sample of axon and dendrite morphologies: The dense structural model is populated with an unbiased sample of *in vivo* labeled and reconstructed neuron morphologies that represent ∼1% of the excitatory population of one barrel column ($n = 154$) (Narayanan et al., 2015) and ∼5% of the neurons located within a barreloid in VPM thalamus ($n = 14$) (Oberlaender, Ramirez, & Bruno, 2012; Oberlaender, de Kock, et al., 2012). Given the sparse sample of neuron morphologies, I assessed how representative these morphologies are for each cell type. To do so, I calculated the respective cell type-specific dendrite and axon innervation volumes at a resolution of $(50\,\mu m)^3$ for an increasing number of reconstructed morphologies (see Figure 3.3; see Table 2.4 for an overview of the number of morphologies per cell type). In Figure 3.3A the results are exemplified for *in vivo* labeled L5PT dendrite morphologies ($n = 16$). The change of dendrite innervation decreases with the number of L5PT morphologies used. For samples of fifteen L5PT morphologies, the volume (i.e., number of 50 µm cubes; $n = 1097$) that is innervated by their dendrites changes on average by less than 3% even if additional morphologies are added. Similarly, the CV of the dendrite density per 50 µm cube decreases with the number of L5PT morphologies. For samples of fifteen L5PT morphologies, the dendrite length is 13.5 ± 1.5 µm and changes by

less than 20% (19% ± 7%) per cube.

Similar results were obtained for all cell types (see Figures 3.3BC and 3.4). For the maximum number of morphologies minus one, the dendrite innervation volume changes on average by 3.7% ± 3.2% per cube and cell type (see Figure 3.3C). For two cell types the average change is larger than 5%: L4PY and (5.9% for 7 dendrite morphologies) and $L6_BCC$ (12% for 5 dendrite morphologies) having both the lowest number of morphologies in the sample. The axon innervation volume changes by 11.6% ± 6.0% per cube and cell type. The change is larger than 20% for L4PY (21.1% for 4 axon morphologies) and $L6_BCC$ (21.3% for 4 axon morphologies). The CV of the respective length densities is on average 23.5% ± 10.0% for the dendrite innervation and 42.1% ± 13.9% for the axon innervation across cell types (see Figure 3.4). The cell types with the least number of morphologies have the largest CVs (e.g., 47.0% and 63.4% for dendrites and axons of $L6_BCC$, respectively). In summary, the sample of excitatory morphologies hence captures 96.3% ± 3.2% and 88.4% ± 6.0% of the dendrite and axon innervation volume, respectively.

Robustness of model predictions due to upscaling of morphological sample: I investigated the robustness of the structural estimates provided by the dense structural model when upscaling the sparse sample of neuron morphologies to all neurons in the model while systematically increasing the sample size of neuron morphologies (see Figure 3.5). Specifically, the length density of axons and dendrite and the cell type diversity[1] per cube were quantified. This sensitivity analysis revealed that the estimates of the structural composition converge remarkably quickly to stable values. Each cell type in the model is represented by at least five dendrite morphologies and four axon morphologies (see Table 2.4). Given these numbers, the CVs of the density and diversity per cube drops to 27.0% ± 11.5% and 22.0% ± 5.7% (dendrite and axon density), and to 6.0% ± 7.7% and 5.0% ± 5.9% (dendrite and axon diversity). The dendrite and axon densities within one barrel column (across 512 probed cubes) does not change by more 8.4% and 6.6%, respectively, and their respective cell type diversity by no more than 6.4% and 4.7% — even if the model is based on a larger sample of neuron

[1] cell type diversity denotes the number of cell types that contribute structures to a cube.

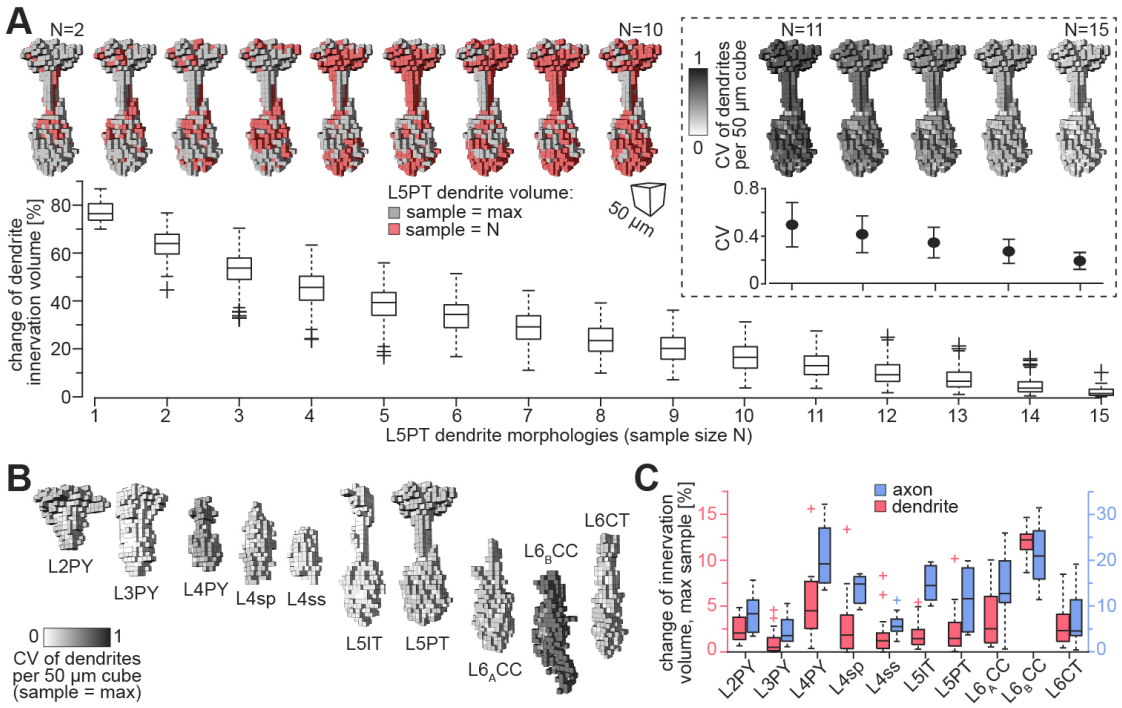

Figure 3.3. Precision of dendrite and axon morphologies in rat vS1. (A) Dendrite innervation volumes of L5PTs depending on the number of reconstructed neurons. Top: Dendrite volume innervated by all L5PT morphologies (gray) and by random subsample of L5PT morphologies (red). Bottom: Change in dendrite innervation volume across number of L5PT dendrite morphologies. Insert shows CV of dendrite densities within each 50 µm cube of the innervation volume depending on the number of reconstructed neurons. (B) Innervation volumes and CVs of dendrite densities therein for the respective maximal sample of reconstructions per cell type. (C) Quantification of the degree to which the sample of reconstructed neurons provides robust estimates of dendrite (red) and axon (blue) innervation volumes for each excitatory cell type in rat vS1. Figure adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020).

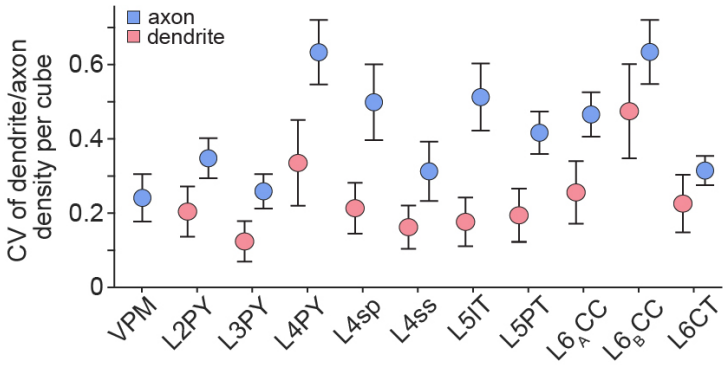

Figure 3.4. Precision of dendrite and axon length per cell type in rat vS1. Quantification of the degree to which the sample of reconstructed neurons provides robust estimates of dendrite (red) and axon (blue) length densities for each excitatory cell type in rat vS1.

morphologies. The density estimates hence capture more than 90% of the dense structural composition of vS1 that is preserved from animal to animal.

3.1.2 Structural estimates at the subcellular level

Note: The following section is adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020).

I analyzed the estimated structural composition across a subsample of 512[2] out of 51,200 50 µm cubes that comprise the dense structural model. The estimated dense structural composition varies strongly across the probed column, depending on the specific laminar and/or columnar location of each 50 µm cube. One cube comprises about 10.5 ± 4.7 neuron somata and axons and dendrites with a total path length of 0.55 ± 0.16 m and 6.89 ± 0.79 cm, respectively. For one of the four exemplary cubes (shown as zoom-ins in Figure 3.1) there are somata of 3 different cell types, dendrites of 9 different cell types, and axons of 10 different cell types present. The cell type diversity of each cube is 2.3 ± 1.0 for somata, 9.5 ± 1.5 for axons, and 6.2 ± 1.9 for dendrites (see Figure 3.6).

[2] spanning L2-6 of the C2 barrel column

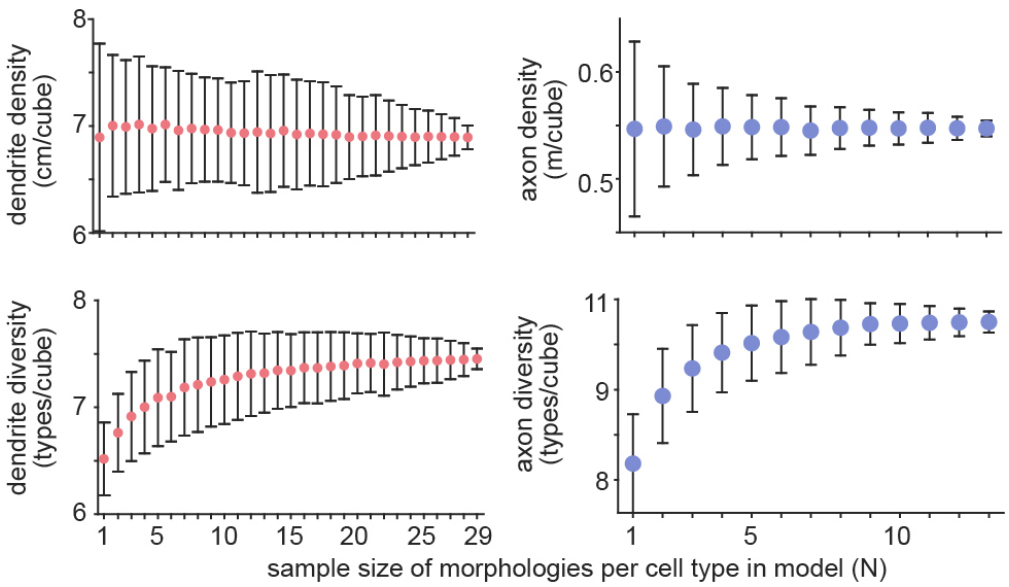

Figure 3.5. Dense structural model provides robust estimates of the dense structural composition of rat vS1. Left panel: Total dendrite length density (top) and dendrite diversity (bottom) per 50 µm cube in the model vs. number of morphologies per cell type used for upscaling in the model. Right panel: Same as in left panel but for axons. Note that the number of morphologies varies between cell types. Figure adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020).

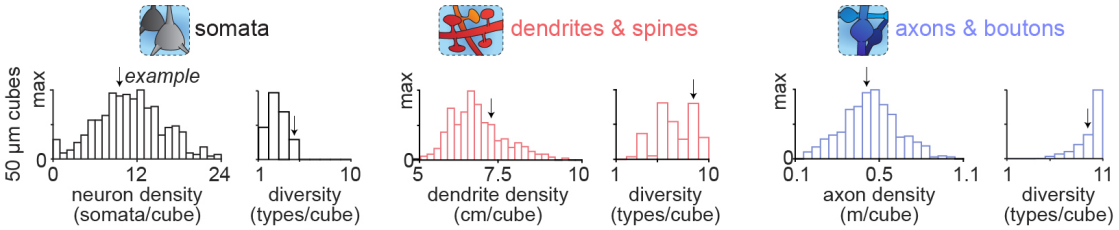

Figure 3.6. Estimates of dense structural composition in rat vS1. Dense structural model provides density distributions for the number and diversity of somata (left), dendrite length and diversity (center), and axon length and diversity (right) at a resolution of 50 µm cubes. Diversity distributions denote the number of cell types that contribute somata, dendrites, or axons to each cube. The arrows represent examples shown in Figure 3.1. Figure adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020).

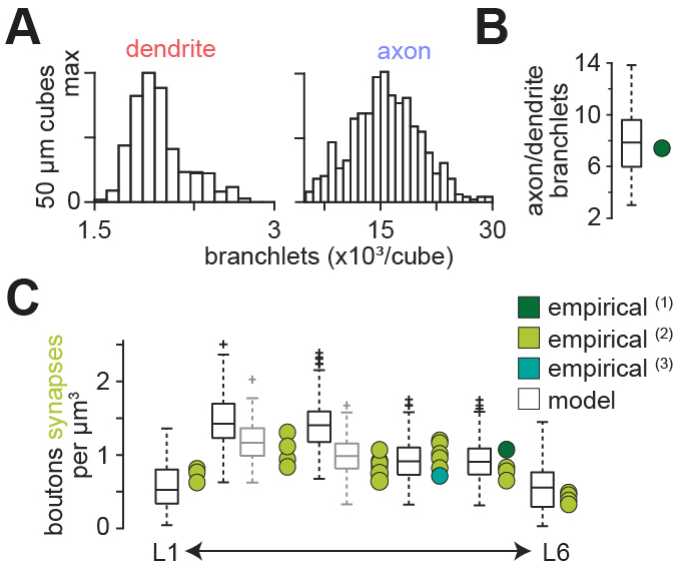

Figure 3.7. Estimates of dense structural model are consistent with empirical data. (A,B) The model estimates ratios of pre- and postsynaptic structures (dendrite/axon branchlets; panel A) per 50 µm cube that are consistent with empirical data from dense electron microscopy approaches (B). (C) The model estimated synapse densities per layer that are consistent with empirical data from electron microscopy approaches. Black boxes represent L1 to L6, gray boxes represent upper half of L2 and L4. Empirical data: (1) Kasthuri et al., 2015, (2) Santuy et al., 2018, (3) Motta et al., 2019. Figure adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020).

Comparison with empirical neurite packing and synapse densities: I tested whether the estimates of the dense structural model were consistent with measurements from saturated reconstructions of neocortical neuropil. In the model, each probed cube comprises $15,786 \pm 4,585$ axon branchlets and $2,036 \pm 230$ dendrite branchlets (see Figure 3.7A). The resultant estimates for packing densities are consistent with measured axon/dendrite branchlet ratios (Kasthuri et al., 2015) (predicted ratio of 7.8 ± 2.3 vs. empirical ratio of 7.4; see Figure 3.7B). Axonal excitatory bouton distributions are within the range of layer-specific synapse density measurements reported from electron tomography of rat barrel cortex (Santuy et al., 2018) and electron microscopy of mouse somatosensory cortex (Motta et al., 2019; Kasthuri et al., 2015) (see Figure 3.7C). A two-sample Kolmogorov-Smirnov test yielded that the estimated and empirical synapses densities in L2 and L3 are

significantly different (L2: $D = 0.77$, $p = 0.0077$; L3: $D = 0.62$, $p = 0.0005$) reflecting either a potential underestimation of the estimated synapses in these layers or differently defined layer borders. When restricting the estimated synapse densities to the upper half of L2 and L4, respectively, the estimates become consistent with the empirical densities of L2 and L3, respectively. The results with respect to the inhibitory boutons and synapses are described in Section 3.1.5 on page 107.

3.1.3 Layer-specific analysis of excitatory neurons

Note: In collaboration with my colleague Dr. Rajeevan T. Narayanan (research center caesar, Bonn), I analyzed the predictions of the dense structural model with respect to the layer organization of vS1. The following Section is adapted from Narayanan et al. (2017).

Layer-specific organization of excitatory cell types in rat vS1

Excitatory somata of the respective cell types are not restricted to the layer that is suggested by their naming convention (see Figure 3.8A). Specifically, somata of neurons with the dendrite morphology of L2PY are found throughout L2 and L3, but are more frequent within L2 (see Tables 3.2, 3.3, and 3.4 for the fraction of neurons per cell type per layer). Similarly, somata of L3PY are distributed throughout L3 and L4, but are not found in L2. In contrast, somata of L4ss and L4PY are largely restricted to L4, whereas those of L4sp are also found in lower L3. L5IT and L5PT are restricted to L5, with L5IT being more abundant in upper L5 (i.e., L5A: ∼80/20% L5IT/L5PT) compared to L5PT that are more abundant in deep L5 (i.e., L5B: ∼40/60% L5IT/L5PT) (Oberlaender et al., 2011). Somata of L6CT and $L6_BCC$ are largely restricted to L6. In contrast, $L6_ACC$ are distributed around the L5/6 border (i.e., in deep L5 and upper L6). Thus, within each layer, somata from multiple excitatory cell types intermingle. The only layer border that separated between somata of different excitatory cell types was the L4/5 border (Narayanan et al., 2015).

I calculated a convex hull around the dendrites from all cells that had been assigned to a respective cell type (see Figure 3.8A). Remarkably, these dendrite

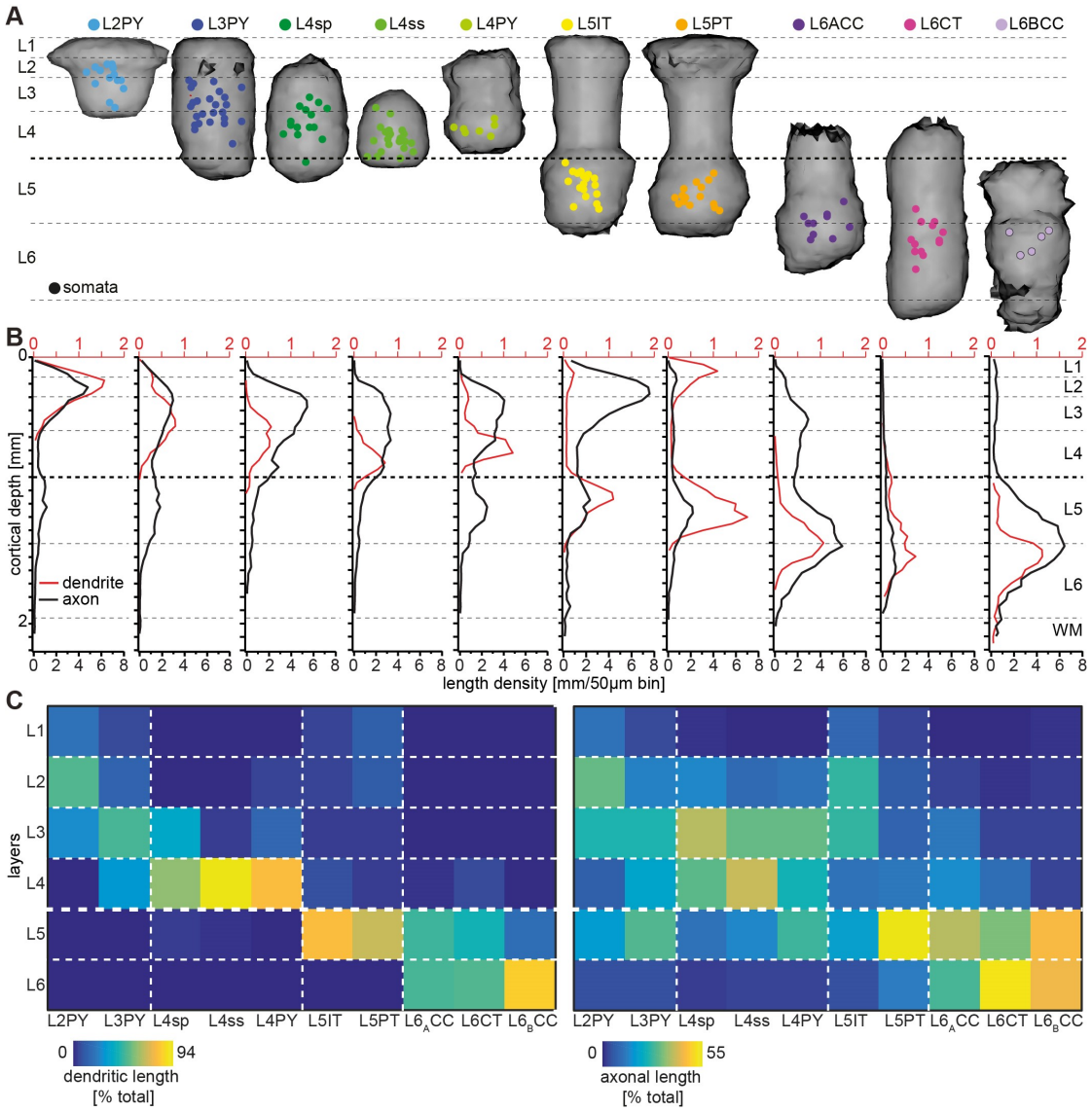

Figure 3.8. Layer-specific organization of excitatory cell types in rat vS1. (A) 3D volumes spanned by registered dendrite morphologies from the ten major excitatory cell types, respectively. Colored dots denote the soma locations of the *in vivo* labeled and reconstructed neurons. (B) 1D length density profile of dendrites (red) and axons (black) averaged across all neurons assigned to one of the 10 excitatory cell types. (C) Quantification of cell type-specific and layer-related dendrite (left panel) and axon (right panel) distributions. The colors denote the relative amount of dendrite/axon per cell type within each layer, normalized to the total dendrite/axon length of the respective type. Figure adapted from Narayanan et al. (2017).

Table 3.2. Number of somata that each of the ten major excitatory cell types contributed to each of the six layers of an average barrel column in the dense structural model. Table adapted from Narayanan et al. (2017).

Somata	L1	L2	L3	L4	L5	L6
L2PY	34	174	725			
L3PY		52	151	128	67	
L4sp			413	124	248	
L4ss			113	2336	4	
L4PY				517		
L5IT					1446	
L5PT					116	
$L6_ACC$					47	96
L6CT					133	3838
$L6_BCC$						79

"innervation volumes" showed several relationships with layers (see also Elston, Pow, & Calford, 1997; Elston, DeFelipe, Arellano, Gonzilez-Albo, & Rosa, 1999). Specifically, dendrites of L2PY are restricted to L1-3 and those of L3PY are restricted to L1-4. In contrast, dendrites of L4PY and L4sp do not innervate L1 and are restricted to L2-4. Dendrites of L4ss are largely confined to L4. Dendrites of L5IT and L5PT extend across L1 to L5 and terminate at the L5/6 border. Dendrites of $L6_ACC$ and L6CT range from L4-6, whereas those of $L6_BCC$ remain within L5 and L6. Thus, even though soma distributions of the ten cell types are only loosely related to layers and the borders between them, and dendrites of each cell type extend across multiple layers, the dendrite distributions display cell type-specific relationships with cytoarchitectonic landmarks (see Figure 3.8B).

Finally, I investigated whether intracortical axon distributions of each cell type are organized with respect to layers. In a previous study, it was shown that the intracortical axon projections of individual excitatory neurons are correlated with their respective dendrite cell type (Narayanan et al., 2015). The cell type-specific axon differences were only partly reflected by parameters such as overall path lengths or topology but originated primarily from different vertical (i.e., laminar, see Figure 3.8B) and horizontal (i.e., transcolumnar) projection patterns (Narayanan et al., 2015). Specifically, axons of L2PY densely innervate L1-3 and

Table 3.3. Length of dendrites that each of the ten major excitatory cell types contributes to each of the six layers of an average barrel column. Table adapted from Narayanan et al. (2017).

Dendrites [m]	L1	L2	L3	L4	L5	L6
L2PY	7.9	15.4	8.9	0.1		
L3PY	2.7	5.7	20.3	21.3	1.7	
L4sp		0.1	2.2	5.0	0.9	
L4ss			0.3	8.1	0.3	
L4PY		0.5	1.3	6.2		
L5IT	0.9	0.8	0.7	2.0	16.4	0.4
L5PT	3.4	2.1	0.9	1.5	21.4	0.7
$L6_ACC$				0.4	7.5	11.7
L6CT				2.3	11.3	38.8
$L6_BCC$					1.9	11.8

less densely L5, with the two innervation peaks coinciding with the centers of L3 and L5, respectively. Axons of L3PY have similar axon projection patterns compared to L2PY, but innervation of L5 is as dense as innervation of L3, where the peak coincides with the L2/3 border. Axons of the three L4 cell types deviate from most other cell types, as they do not innervate L1. Apart from this difference, the vertical axon profile of L4PY resembles the one of the L3PY. Axons of L4sp and L4ss are restricted to L2-4. However, axons of L4sp are most dense within L2/3 (the innervation peak coincides with the L2/3 border), whereas axons of L4ss are equally dense in L2/3 and L4. Axons of both, L5IT and L5PT, innervate L1-5, but in contrast to L5PT, L5IT project densely to L2 and L3 (the innervation peak coincides with the L2/3 border). Axons of $L6_ACC$ and $L6_BCC$ innervate the entire cortical depth from the pial surface to the WM tract, whereas L6CT axons are sparse and restricted to L4-6. Axons of $L6_BCC$ are sparse in L1-4, but dense in L5 and L6 (the innervation peak coincides with the L5/6 border). In contrast, axons of $L6_ACC$ are most elaborate within L5-6, but also abundant in L3-4, and sparsely innervate L1-2.

The laminar landmarks that coincide with the respective dendrite or axon distributions are different and specific for each cell type. For example, even though somata of L2PY and L3PY intermingle within L3, the vertical extents of their den-

Table 3.4. Length of axon that each of the ten major excitatory cell types contributes to each of the six layers of an average barrel column. Table adapted from Narayanan et al. (2017).

Axons [m]	L1	L2	L3	L4	L5	L6
L2PY	16.1	46.5	38.6	12.2	36.3	8.9
L3PY	15.3	45.7	84.0	73.4	108.3	29.6
L4sp	2.1	19.4	48.6	43.5	16.2	5.9
L4ss	0.3	12.3	43.6	59.5	23.5	4.9
L4PY	0.3	8.8	22.6	19.9	24.6	4.9
L5IT	14.0	45.0	47.5	22.5	40.8	11.6
L5PT	1.1	3.0	3.7	5.5	27.7	15.7
$L6_ACC$	1.2	5.7	24.7	35.0	76.4	110.1
L6CT	0.1	0.7	2.2	9.6	31.5	65.9
$L6_BCC$	0.7	1.5	2.9	3.8	36.1	77.3

drite distributions coincide with the vertical extents of L1-3 and L1-4, respectively. The vertical axon distributions of these two cell types delineate the center and extent of L5, whereas their peak axon densities in the superficial layers coincide with the center of L2 and the L2/3 border, respectively. To provide a comprehensive overview of the degrees to which dendrite and axon distributions of the different cell types are organized with respect to layers, I calculated the relative amount of dendrite and axon path length within each layer for each cell type (see Figure 3.8C).

Cell type-specific organization of cortical layers in rat vS1

I quantified the cell type-specific structural composition of each layer predicted by the dense structural model. The D2 barrel column comprises 17,816 excitatory somata (see Table 3.5 for the number of neurons per cell type). Each soma in the model is represented by a dendrite morphology of the respective cell type (see Figure 3.9B), whose registered soma depth was within 50 µm from the location in the model (Egger et al., 2014). This upscaling of the reconstructions resulted in a total dendrite path length that is found within the D2 barrel column of 246 m (i.e., originating from all excitatory neurons within entire vS1; see also Table 3.5 for dendrite path lengths contributed by each cell type). Similarly, the registered

Table 3.5. Number of somata, and dendrite and axon length that each of the ten major excitatory cell types contributes to D2 barrel column. Table adapted from Narayanan et al. (2017).

Cell type	Somata	Dendrite [m]	Axon [m]
L2PY	1,833	32.2	158.5
L3PY	2,648	51.7	356.3
L4sp	1,685	8.2	135.8
L4ss	2,453	8.7	144.0
L4PY	517	8.0	81.0
L5IT	1,446	21.1	181.4
L5PT	1,106	30.0	56.7
$L6_ACC$	1,367	19.5	253.0
L6CT	3,971	52.5	109.9
$L6_BCC$	790	13.7	122.3

axon morphologies were upscaled to the respective number of neurons per cell type (see Figure 3.9C). The resultant total (i.e., from all excitatory neurons in vS1) axon path length within the D2 barrel column is 1599 m (see Table 3.5 for axon path lengths by each cell type).

By combining the dense structural model with the quantification of the layer borders, the model provides estimates on the number of somata, as well as the dendrite and axon path lengths that each of the ten cell types contributes to each of the six layers (see Figure 3.9D). Within the D2 barrel column, the majority of dendrites in L1 originates from L2PY, followed by L5PT, L3PY, and L5IT. The majority of axons originates from L2PY, followed by L3PY and L5IT (see Table 3.4). Similar to L1, the majority of the dendrites in L2 originates from L2PY, followed by L3PY and L5PT. The majority of axons in L2 originates to almost equal amounts from L2PY, L3PY, and L5IT. In L3, L3PY, followed by L2PY and L4sp contribute the most dendrites, whereas axons in L3 originate from L3PY, closely followed by L5IT and the three cell types in L4. The cell type contributing the most dendrites to L4 are L3PY, followed by the three L4 cell types. The majority of axon in L4 originates from L3PY, L4ss, L4sp, and $L6_ACC$. L5 comprises dendrites primarily from L5PT, closely followed by L5IT and L6CT. In turn, the majority of axon in L5 originates from L3PY, followed

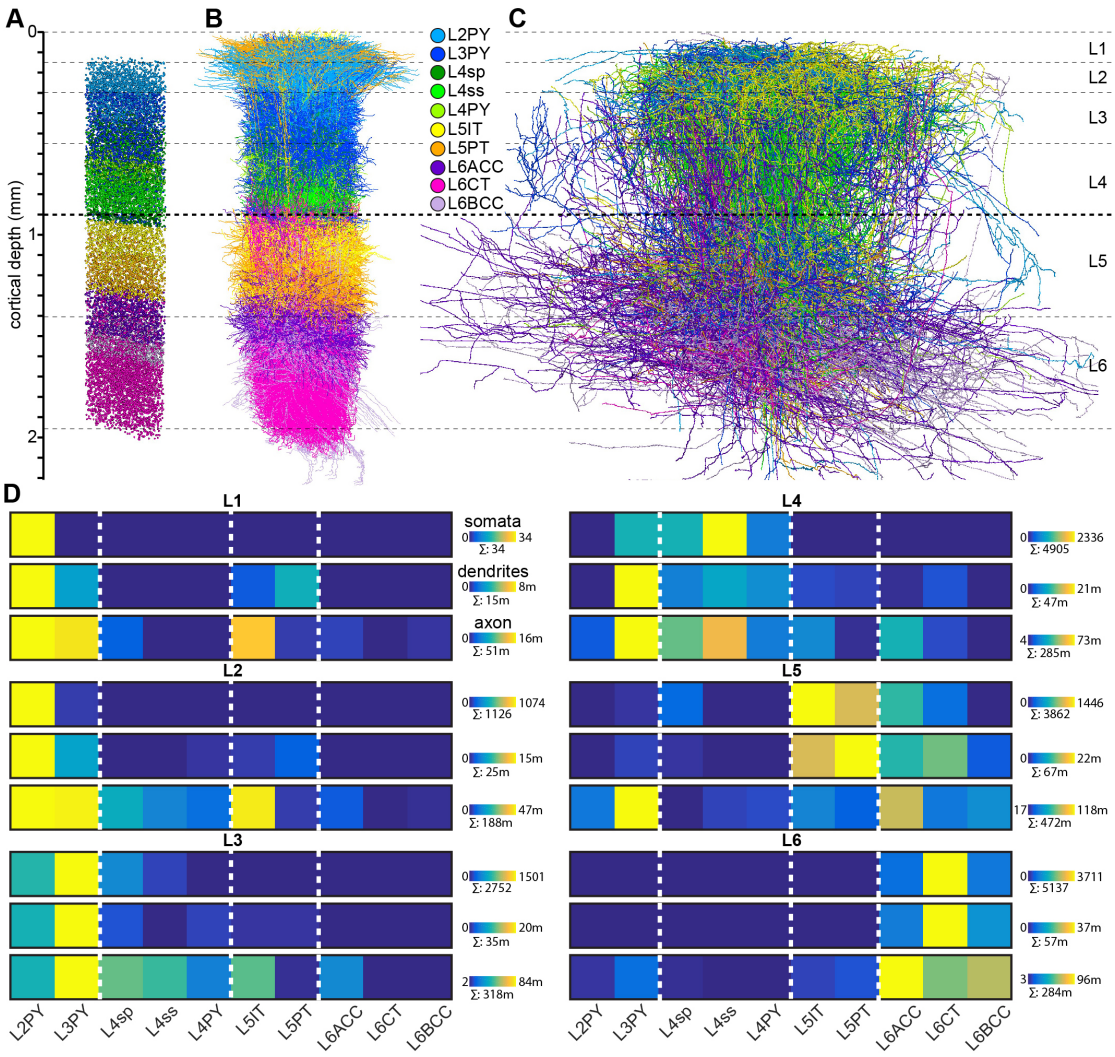

Figure 3.9. Cell type-specific organization of cortical layers in rat vS1. (A) Semi-coronal view of the cell type-specific soma distribution of the D2 barrel column in the dense structural model. (B) Corresponding to panel A: model of the cell type-specific distribution of dendrites. (C) Corresponding to A: model of the cell type-specific distribution of axons. (D) Quantification of cell type-specific structural composition of each layer with respect to somata (top rows), dendrites (center rows), and axons (bottom rows). Dendrite and axon length represent all neurons in vS1 whose respective dendrites and/or axons extend into the volume of the D2 barrel column. Figure adapted from Narayanan et al. (2017).

by $L6_A$CC, followed by similar contributions from all remaining cell types. L6 is more homogeneous than the other layers, with the majority of dendrites and axons originating from the three L6 cell types.

3.1.4 Path length and conduction velocity estimates of VPM axons

In collaboration with my colleagues Dr. Robert Egger and Dr. Rajeevan T. Narayanan (Max Planck Institute for Biological Cybernetics, Tübingen, and research center caesar, Bonn), I analyzed the path length distribution of VPM axons ($n = 14$) to estimate the temporal delay of sensory-evoked signal flow (Egger et al., 2020). Specifically, I calculated the path length of VPM axons in L6A, where somata of $L6_A$CC are located, and L4 in the C2 column (see Figure 3.10A). Combined with conduction velocity measurements (0.33 m/s; Salami et al., 2003), I predicted that sensory-evoked excitation reaches the L5/6 border 3.0 ± 1.7 ms earlier than the L4/5 border. This is in line with experimental data (see Figure 3.10BC; for more details, see Egger et al., 2020) supporting the findings that $L6_A$CC neurons respond faster to sensory-evoked excitation from VPM than neurons located more superficial.

3.1.5 Inhibitory neurons

The entire dense structural model of vS1 comprises 69,810 inhibitory neurons (see Table 3.6). The following analysis of the inhibitory composition of each layer is restricted to the C2 barrel column. The average C2 barrel column is approximated by a cylinder with a volume of $0.20\,\mathrm{mm}^3$ (Egger et al., 2012) — in the following referred to as "soma column". It comprises 16,277 neurons, including 1,986 INs in L2-6 and 25 INs in L1 (Egger et al., 2015)[3]. Based on the relative frequency of occurrence of each projection type at the respective cortical depths, the C2 barrel column comprises 1,034 local projecting INs (52%), followed by 421 InfraAsc (21%), 200 GranAsc (10%), 181 GranSym (9%), and 150 SupraDesc (8%) INs (see

[3]Note that these neuron numbers are from different vS1 model realization than presented in the other Sections; for more details, see Method Section 2.1.3.

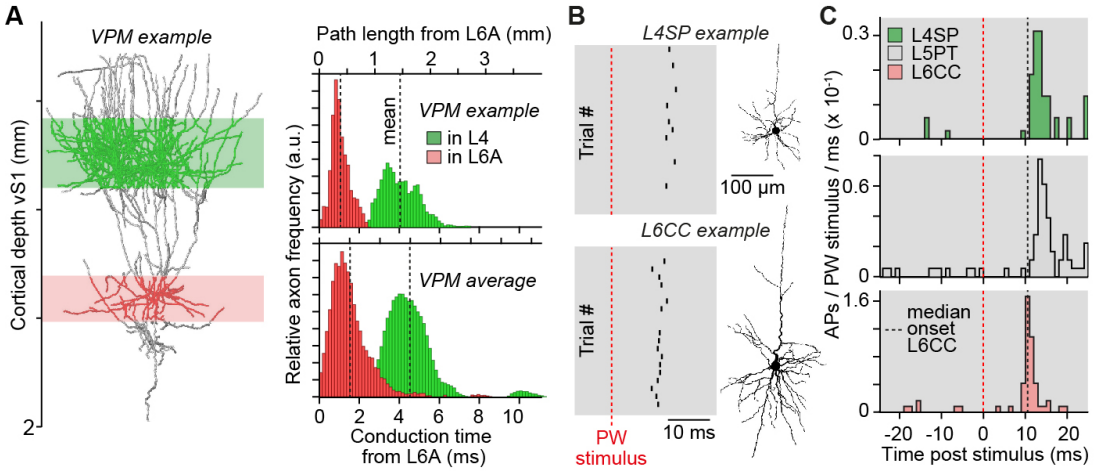

Figure 3.10. Path length distributions of VPM axons. (A) Example of *in vivo* labeled VPM axon, whose path length distribution was quantified with respect to the deepest location where somata of $L6_ACCs$ are located (here referred to as L6A). Path lengths were divided by the measured intracortical conduction velocity of VPM axons (Salami et al., 2003) (top right). Average conduction time of VPM axons ($n = 14$) to L6A and 4, respectively (bottom right), is shown. (B) Action potential responses evoked by principal whisker (PW) stimuli in exemplary L4SP and $L6_ACC$. (C) Post-stimulus time histograms of principal whisker-evoked action potentials in morphologically identified L4SPs ($n = 8$), L5PT ($n = 9$), and $L6_ACCs$ ($n = 6$). Figure adapted from Egger et al. (2020).

Table 3.7). Local INs are the dominating projection type in supragranular (241 local vs. 199 nonlocal projectors) and granular (379 local vs. 66 GranSym vs. 113 GranAsc) layers. In infragranular layers, local INs are less abundant than nonlocal projectors (414 local vs. 568 nonlocal).

I calculated the 3D distribution of the total amount of dendrites and axons originating from all INs with somata in the C2 column (see Table A1 and Figure 3.11). The resultant volume spanned by the inhibitory dendrites — in the following referred to as "IN dendrite column" — comprises a volume of 0.52 mm^3, thus exceeding the volume of the soma column by a factor of 2.5. The respective volume spanned by dendrites of excitatory neurons is similar in volume (Oberlaender, de Kock, et al., 2012). The total dendrite length within the C2 dendrite column is estimated as 7.5 m and 98.1 m for inhibitory and excitatory neurons, respectively. Compared to the density of excitatory dendrites, which is largely uniform across

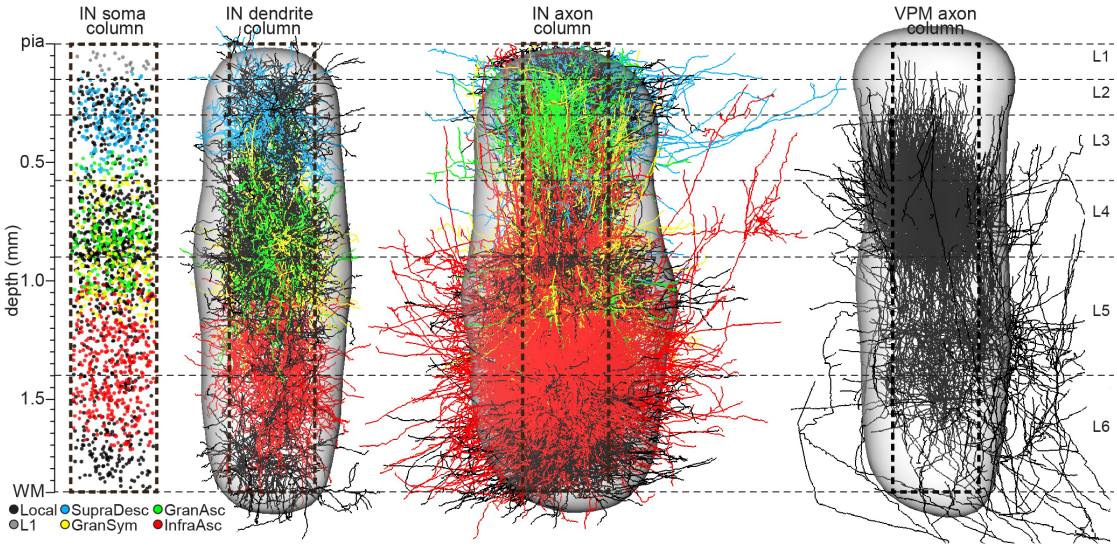

Figure 3.11. Dense structural model of inhibitory neurons in rat vS1. Left: Semi-coronal view of projection type-specific distribution of inhibitory somata in the C2 barrel column of the dense structural model (colored/black dots represent 1,986 IN somata in L2-6; gray dots represent 25 IN somata in L1). Dashed box represents dimensions of the "soma column" (i.e., the cylinder was generated by extrapolating the L4 barrel towards the pial surface and white matter (WM)). Dashed lines denote layer borders. Center-left: Projection type-specific distribution of IN dendrites representing the 1,986 INs in L2-6 of the soma column (only 10% of the dendrite morphologies are shown). The gray shaded volume represents the 3D envelop around the dendrites from all INs located within the soma column, referred to as "IN dendrite column". Center-right: Projection type-specific distribution of IN axons (axons of all 204 reconstructed INs are shown). The gray shaded volume represents the 3D envelop around the axons from all INs located within the soma column, referred to as "IN axon column". Right: 3D distribution of VPM axons originating from the barrel column-related VPM barreloid (only 10% of all VPM axons are shown). The gray shaded volume represents the 3D envelop around the dendrites of all excitatory neurons located within the soma column, referred to as "excitatory dendrite column" (adapted from Oberlaender, de Kock, et al., 2012).

Table 3.6. Projection type-specific number of inhibitory somata in dense structural model.

Projection type	Number of somata
L1	1,091
Local	37,557
SupraDesc	5,446
GranSym	5,974
GranAsc	6,261
InfraAsc	13,481
Total	69,810

Table 3.7. Number of inhibitory somata that each axon projection type contributed to each of the six layers of an average barrel column in the dense structural model. Table adapted from Udvary et al. (2016).

Layer	Local $n = 1,034$	SupraDesc $n = 150$	GranSym $n = 181$	GranAsc $n = 200$	InfraAsc $n = 421$
L1	15	6			
L2	107	58			
L3	119	86	22	33	
L4	379		66	113	
L5	235		93	54	259
L6	179				162

the cortical depth (Oberlaender, de Kock, et al., 2012), inhibitory dendrites are most densely distributed within their respective soma hot zones (i.e., resembling the inhibitory soma distribution Meyer et al., 2011). The total surface area of all inhibitory somata/dendrites within the dendrite column is estimated as 17.4 mm^2, which in the model represents an upper limit of the putative postsynaptic structures that can be innervated by axons.

Similarly, the volume containing the axons of all INs with somata in the C2 column was termed "IN axon column". The volume of the axon column (0.87 mm^3) is more than four times larger than the corresponding soma column, revealing that axon innervation of INs extends beyond the lateral borders of a barrel column. For comparison, the respective volume innervated by excitatory axons is ∼1.7-times

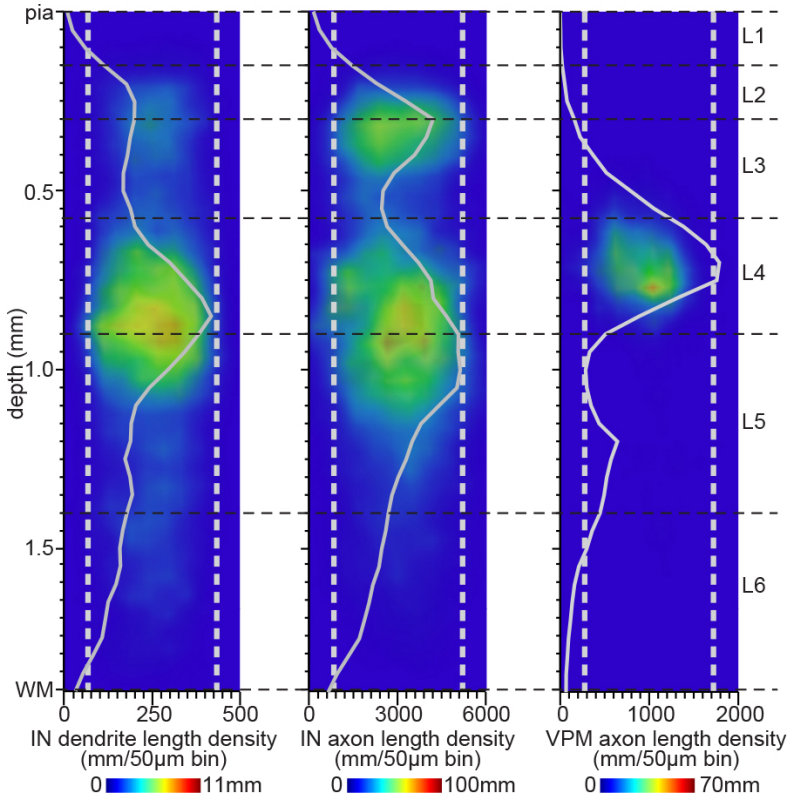

Figure 3.12. Length distributions of IN dendrites and axons in an average barrel column of vS1. Left: Projection image of the 3D density distribution (at 50 µm cube resolution) of all IN dendrites located within the soma column, superimposed with the vertical IN dendrite density profile along the cortical axis. Center: Projection image of the respective 3D IN axon density distribution, superimposed with the vertical IN axon density profile. Right: Projection image of the 3D density distribution of all VPM axons originating in the barrel column-related VPM barreloid.

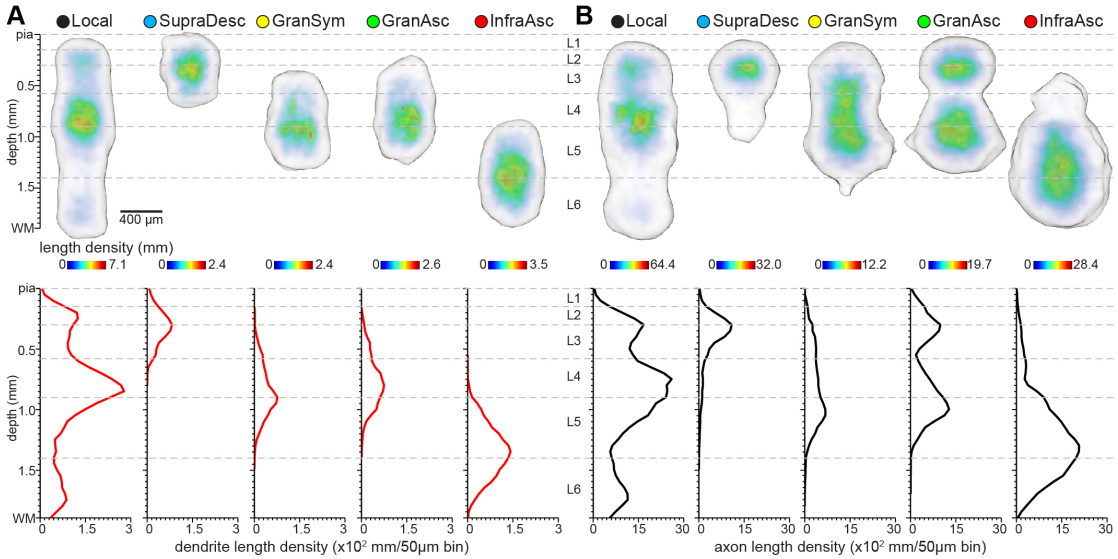

Figure 3.13. Projection type-specific 3D dendrite and axon distributions in an average barrel column of vS1. Same analysis as in Figure 3.11 but separately for each IN axon projection type. (A) Top panels show projection type-specific 3D dendrite density distributions. The gray shaded volumes represent the 3D envelops around all dendrites of a respective projection type. Bottom panels show the vertical IN dendrite density profiles for each projection type. (B) Same as in panel A but for IN axons.

larger than the inhibitory axon column (Narayanan et al., 2015). The model estimates that axons of INs with somata located in the C2 column contribute 111 m of path length to the cortical circuitry, compared to excitatory neurons that contribute 859 m. Finally, similar to the distribution of inhibitory somata and dendrites along the vertical cortex axis, inhibitory axons innervate vS1 most densely between L2 and L3 and between L4 and L5, referred to as the inhibitory hot zones (Meyer et al., 2011). The axon density peak in L2/3 is more pronounced than the peak in L4/5 compared to the respective dendrite density peaks (axon/dendrite density peak ratios between L2/3 and L4/5: 0.82/0.48, see Figure 3.12). Thus, inhibitory axons innervate cortex in a zone-specific manner (i.e., aligned to the two inhibitory hot zones; Meyer et al., 2011), whereas excitatory axons innervate cortex in a layer-specific manner (i.e. aligned to "excitatory layers"; Narayanan et al., 2015, 2017). For the projection type-specific dendrite and axons innervation, see Figure 3.13.

Comparison with empirical synapse densities: Measurements of inhibitory synapse densities revealed variations between the layers (Santuy et al., 2018; Kasthuri et al., 2015). I tested whether these variations are captured by the model. Axonal inhibitory bouton distributions in L2, L3, L4, and L6 are consistent with their respective synapse density measurements (see Figure 3.14). A two-sample Kolmogorov-Smirnov test yielded that the estimated and empirical inhibitory synapses densities in L1 and L5 are significantly different (L1: $D = 1.00$, $p = 0.0015$, L5: $D = 1.00$, $p = 0.0002$) reflecting either a potential underestimation of estimated inhibitory synapses in these layers or differently defined layer borders.

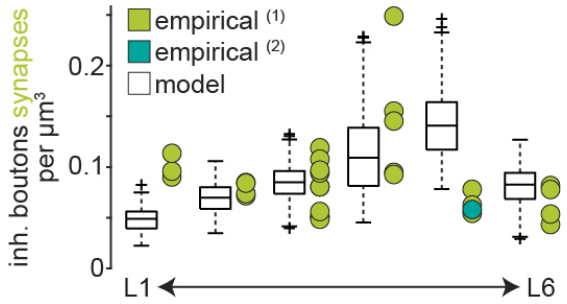

Figure 3.14. Empirical vs. estimated inhibitory synapse densities for L1-6. Comparison with empirical data from electron microscopy approaches revealed consistency with estimates for layers 2, 3, 4, and 6, and inconsistencies with layer 1 and 5. Empirical data: (1) Santuy et al., 2018, (2) Kasthuri et al., 2015.

3.2 Dense statistical connectome of rat barrel cortex

By applying the stochastic synapse formation strategy (see Method Section 2.2) to the structural estimates predicted by the dense structural model of rat vS1, a dense statistical connectome was generated (Egger et al., 2014). The dense statistical connectome provides predictions of how likely synapses are formed between any pair of pre- and postsynaptic structures within the model. From these subcellular predictions, connection probabilities between entire neurons and arbitrarily-defined neuron populations can be derived (see Figure 3.15), allowing analysis of

connectivity patterns from subcellular, cellular, and network levels. In the following Sections, I describe the analysis of the predicted pair-wise connectivity statistics between excitatory neurons (see Section 3.2.1) and those involving inhibitory neurons (see Section 3.2.2), correlations between excitatory synaptic inputs (see Section 3.2.3), network topologies (see Section 3.2.4), and the relationship between wiring properties and network topologies using a mathematical model (see Section 3.2.5).

The most important results of the are:

- The model predicts connection probabilities that are generally sparse, heterogeneous, positively skewed, and nonrandom. The majority of distributions underlying the connection probabilities cannot be approximated by a Gaussian distribution.
- Truncation of neuron morphologies through *in vitro* slicing impacts estimates of neuron length and connectivity in an unsystematic way.
- The model predicts that nearly all INs across the entire cortical sheet receive thalamocortical input from VPM. Input is predicted to be strongest in L4 and at the border between L5 and L6.
- The model predicts a correlated network with nonrandom network topologies. Recurrent motifs, for example, are more frequent in the predicted network than in a random network. The topology of the network is related to the underlying wiring properties.
- Using a simple mathematical model, I find that any sparse, heterogeneous, and correlated network results in the emergence of high degrees of recurrence.

3.2.1 Pair-wise connectivity between excitatory neurons

In this Section, I present the results of the predicted connectivity between excitatory neurons at the cellular level. The first part of this Section is adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020).

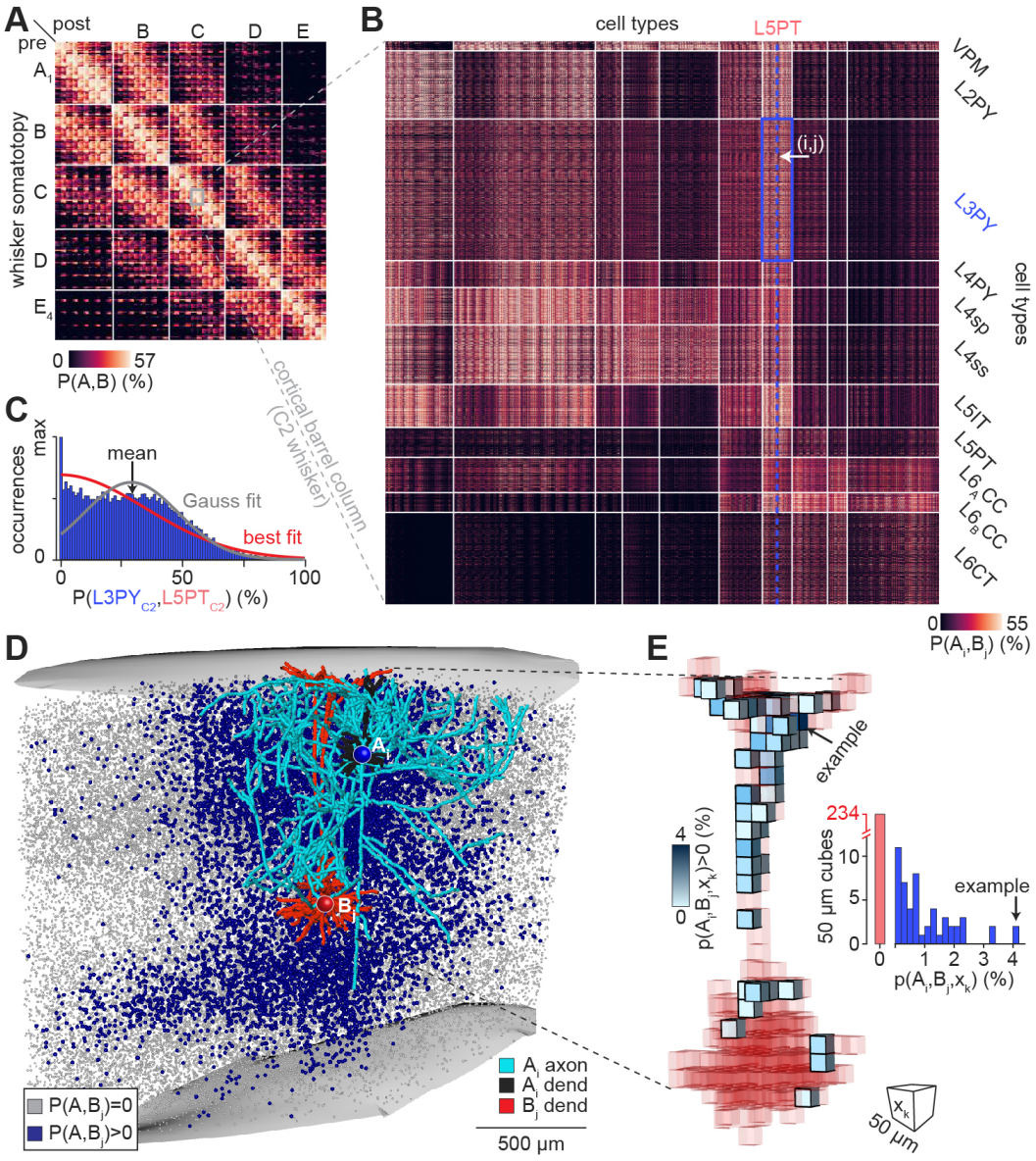

Figure 3.15. Dense statistical connectome of rat vS1. (A) Matrix representation of connection probabilities between all excitatory neurons in the model volume. The neurons in the matrix were sorted by somatotopic location (i.e., nearest barrel column) and cell type. *Caption continued on next page.*

Figure 3.15 — continued from previous page. (B) Zoom-in to the matrix represents all neurons within/around the C2 barrel column. Three examples illustrate the content of the matrix: (1) The blue rectangle comprises connection probabilities (see panel C) between presynaptic neurons of the cell type L3PY and postsynaptic neurons of the cell type L5PT. (2) The blue dashed line comprises the connection probabilities between all neurons in the model volume and one exemplary L5PT (see panel D). (3) The white arrow denotes the connection probability between one exemplary L3PY and the L5PT (see panel E). (C) Distribution of connection probabilities between all L3PY and L5PT within/around the C2 barrel column. Distributions cannot be approximated by a normal distribution (i.e., GAUSS). Here, the best fit represents a half-normal distribution. (D) Dendrite (red) and axon (light blue) morphologies representing the two example neurons (i and j) in panel B. Blue spheres represent somata of neurons whose axons overlap at the resolution of 50 µm cubes with the dendrite of the example L5PT, and which can thereby in principle be synaptically connected (i.e., connection probability > 0). Gray spheres represent neurons that cannot form synaptic connections with the exemplary L5PT. (E) Zoom-in to panel D. Dendrites and axons are represented in the model at a resolution of 50 µm cubes (right). Connection probabilities between all neurons are calculated at this resolution, resulting in distributions for each cell pair across all 50 µm cubes. The matrix of the dense statistical connectome of rat vS1 thus represents more than 1011 connection probability values and allows analyzing connectivity patterns from subcellular to network levels. Figure adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020).

Connection probabilities between neurons are sparse, heterogeneous, grouping-dependent, and nonrandom: The distribution of the predicted connection probabilities depends on how the neurons are pooled together. Figure 3.16AB illustrates these grouping-related differences for differently grouped neurons located in L5. Distribution of connection probabilities vary by pooled cell types, their absolute position within rat vS1 (e.g., inside the barrel column vs. in the septal areas), or their relative position between each other (i.e., intersomatic distances). The connection probabilities across all cell type combinations ($n = 110$) are on average low (i.e., below 50%; 17 ± 11 %), and the underlying distributions are heterogeneous (SD $= 14 \pm 6$%, CV $= 1.17 \pm 0.94$) and positively skewed (i.e., right-tailed; $skew = 0.95 \pm 0.43$) (see Figure 3.16C and A2). Only three connection probability distributions are negatively skewed (P(L4PY,L5PT): $skew = -0.08$; P(L4sp,L4PY): $skew = -0.55$; P(L4sp,L4sp): $skew = -0.09$). I

tested how well seven different distributions approximated the predicted connection probability distributions (see Figure 3.17). Of the 110 connection probability distributions, 77 (70%) are best approximated by a zero-inflated Negative Binomial (ZINB), 26 (24%) by a Half-Normal, 5 (5%) by a Gaussian, and one each by an Exponential and Gamma. The area difference SAD[4] between the predicted distributions and the approximated Gaussian distribution are 0.63 ± 0.31 reflecting that the majority of distributions are not Gaussian distributed. The predicted connectivity between neurons is hence sparse, heterogeneous, grouping-dependent, and nonrandom.

Relationship between connection probabilities and axodendritic overlap

Multiple studies have reported that the average (sparse) axodendritic overlap between two neuron populations cannot fully explain their respective measured connectivity statistics (e.g., Brown & Hestrin, 2009; Petreanu et al., 2009). Given these findings, I assessed how the predicted connection probabilities, as derived from the DSC[5], relate to the (sparse) axodendritic overlap by repeating an analysis by Brown and Hestrin (2009) for all cell type groupings ($n = 110$) (see Figure 3.18). In the following, I illustrate the analysis for L5PT and $L6_{A}CC$ neurons: The axodendritic overlap of L5PT to $L6_{A}CC$ is 1.44 times larger than of $L6_{A}CC$ to L5PT (overlap(L5PT,$L6_{A}CC$) = $163\,486 \pm 103\,126\,\mu m^2$ vs. overlap($L6_{A}CC$,L5PT) = $113\,675 \pm 79\,390\,\mu m^2$). Hence, one would expect the same relative difference between their respective connection probabilities — if the connection probabilities were solely determined by the (sparse) axodendritic overlap. However, the relationship between the predicted connection probabilities is the opposite, differing by a factor of 0.72 (P(L5PT,$L6_{A}CC$) = $15 \pm 14\%$ vs. P($L6_{A}CC$,L5PT) = $21 \pm 14\%$). The ratio of these two factors is 2.00 (i.e., 1.44 / 0.72) and not close to 1 as one would expect. The respective ratios for all 110 cell type combinations are shown as a histogram in Figure 3.18D. About 43% of the ratios are larger than $\frac{3}{2}$ or smaller than $\frac{2}{3}$. While there is a strong link between the axodendritic overlap and predicted connection probabilities in terms of their correlation and estimated slopes

[4] values bound between 0 and 2

[5] Intuitively, DSC (i.e., dense structural composition) is a measure of the dense axodendritic overlap

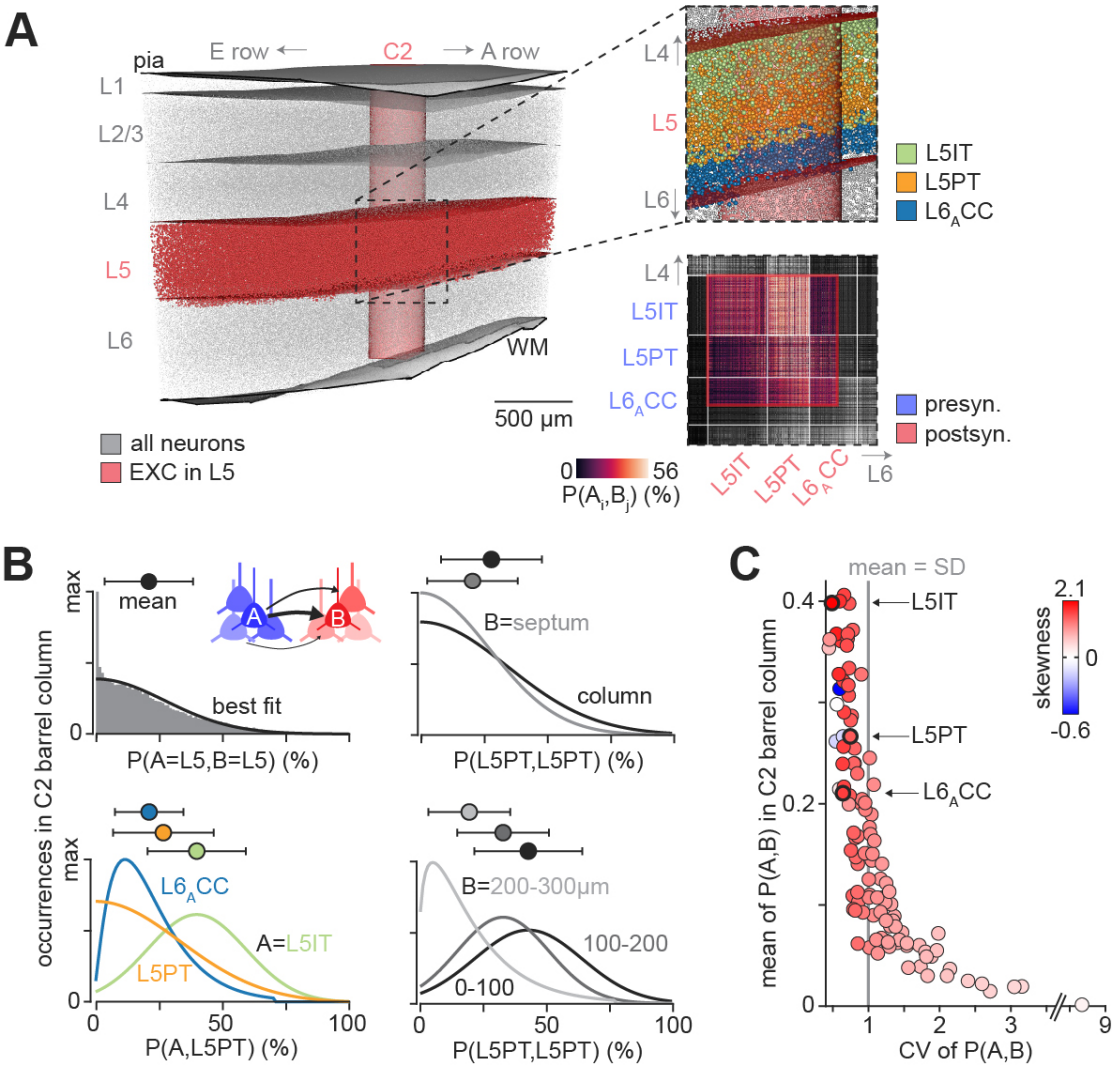

Figure 3.16. Predicted pair-wise connectivity distributions are sparse, heterogeneous, grouping-dependent, and nonrandom. (A) Model of rat vS1. C2 barrel column and somata located in L5 highlighted in red. Top right: Zoom-in to L5 of C2 barrel column with somata color coded by cell type. Bottom right: Zoom-in to matrix of connection probabilities. Neuron pairs located in L5 highlighted in red. (B) Distribution of connection probabilities between different neuron groupings visualized by their best fit. Distribution varies by cell type (bottom left), absolute position of neuron somata (top right), and their intersomatic distances (bottom right). (C) Mean, CV, and skewness of the predicted connection probability distribution for all cell type combinations within one barrel column. Connection probabilities are smaller than 50%, their distributions are broad and skewed. Figure adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020).

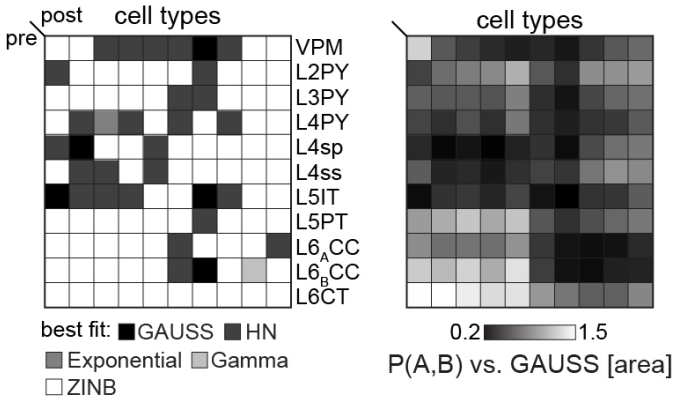

Figure 3.17. Approximation of the connection probability distributions with common distributions. Predicted connection probability distributions were fitted with different functions — Gaussian, Half-Normal (HN), Gamma, zero-inflated Negative Binomial (ZINB), Binomial, Exponential, and zero-inflated Poisson; left: Best fitting distribution; right: summed area difference (SAD) between connection probability distribution and Gaussian fit (maximal difference = 2). Figure adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020).

(see Table A2), relative differences in the average axodendritic overlap do not fully account for relative differences in the average predicted connection probabilities in line with Brown and Hestrin (2009).

Impact of slicing on neurite length and connection probabilities

Many studies infer connectivity statistics from experiments performed in tissue slices *in vitro* where the impact of slicing (i.e., neuron truncation) on the experimentally accessed parameters such as connection probabilities are unknown. Emulating *in vitro* slicing in the dense statistical connectome can provide insight into the impact of slicing. Therefore, I compared the predictions of the dense statistical connectome with its respective predictions when emulating slices *in silico*. First, the impact of slicing is cell type-dependent and affects dendrites and axons to different degrees (see Figure 3.19). The majority of the dendrite path length is preserved within a typical 300 µm wide thalamocortical slice[6] ranging on average from 85 % to 94 % across cell types. In contrast, the majority of axon path length is lost ranging on average from 17 % to 55 % across cell types.

[6] given the neuron somata is located 31 µm and 130 µm within the sliced tissue

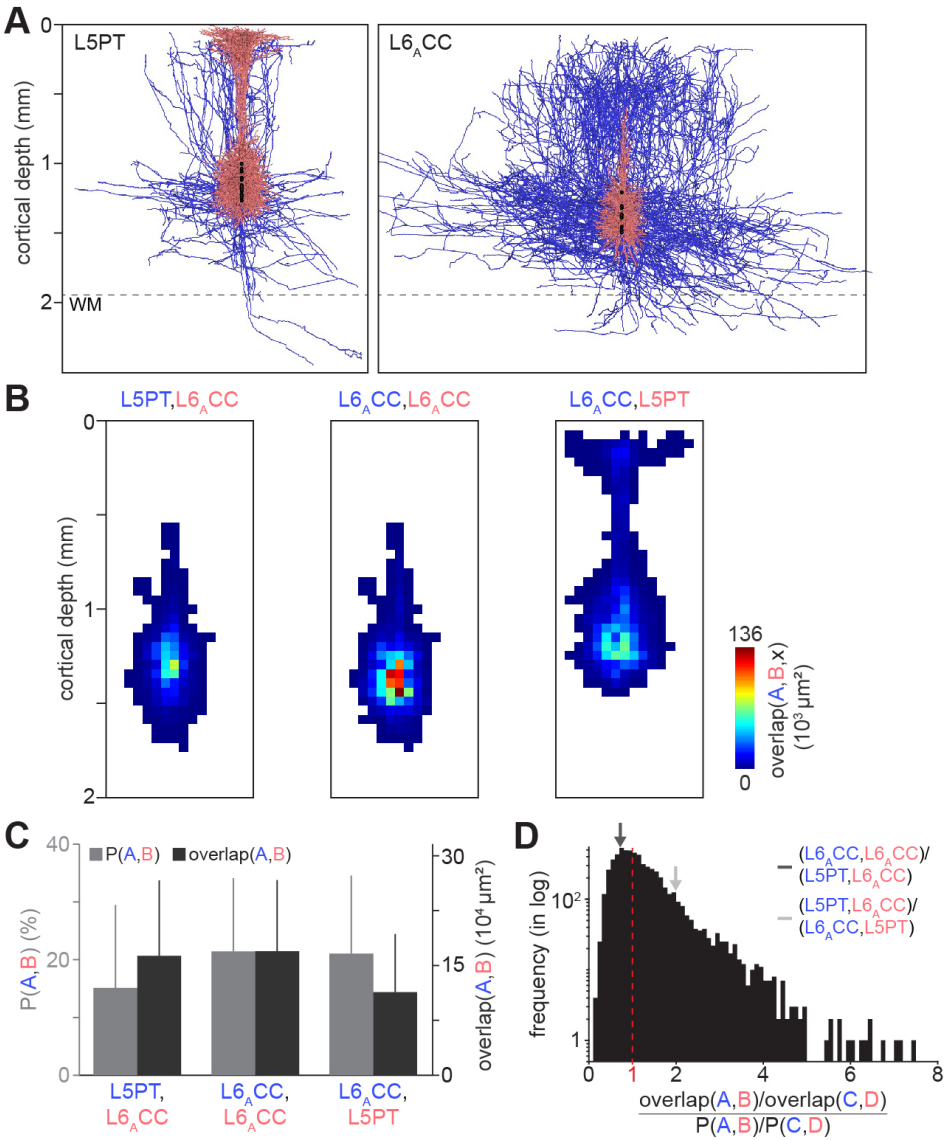

Figure 3.18. Predicted connection probabilities vs. axodendritic overlap. Analysis performed similarly to Brown and Hestrin (2009). (A) Sample of L5PT and L6$_A$CC neuron morphologies horizontally aligned by their somata. Dendrite in red, axon in blue, somata as black spheres. (B) 2D density of axodendritic overlap between three combinations of L5PT and L6$_A$CC neurons. (C) Predicted connection probabilities (gray) and axodendritic overlap (black) for the three combinations of panel B, demonstrating that relative differences in axodendritic overlap are not necessary linked to differences in predicted connection probabilities. *Caption continued on next page.*

Figure 3.18 — continued from previous page. (D) Histogram of ratios between axodendritic overlap and connection probability ratios for all cell type combinations. Ratio of one (dashed red line) indicates one-to-one relationship between differences in axodendritic overlap and connection probabilities.

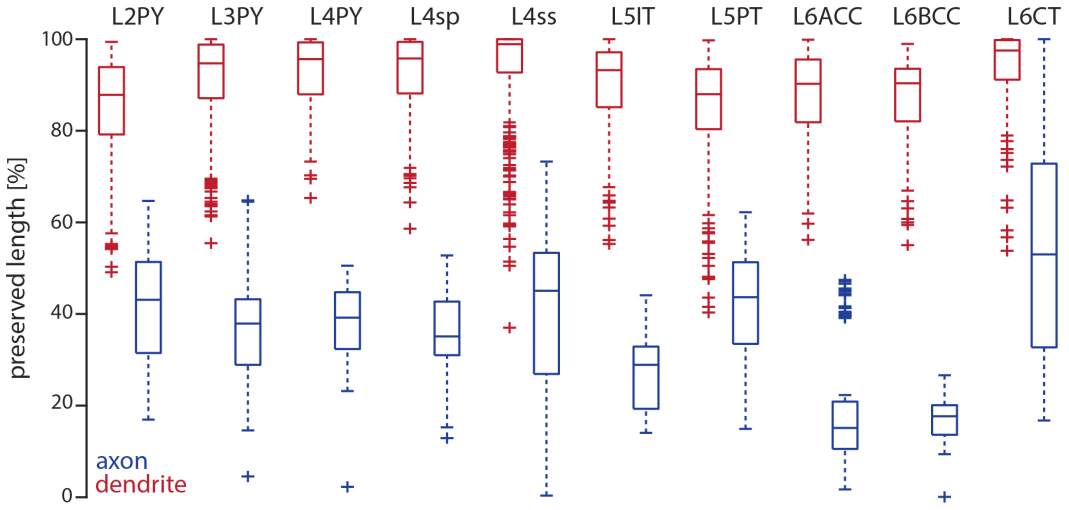

Figure 3.19. Impact of slicing on dendrite and axon length estimates. Box plots of percentages of dendrite (red) and axon (blue) length that is estimated to be preserved in emulated thalamocortical slices of width 300 µm for neurons located within 31 µm and 130 µm of the slice tissue.

Second, the predicted connection probabilities are affected unsystematically by the slicing. The decrease in connection probabilities depends — as for the dendrite and axon path lengths — on the grouping of neurons (see Figure 3.20). The connection probability depends on the soma-to-soma distance, the depth of the neuron pair within the slice, and the pre- and postsynaptic neuron morphologies (i.e., cell types). The resultant predicted histogram of the *in vitro* connection probabilities is not a down-scaled histogram of the predicted *in vivo* connection probabilities. The mean connection probabilities for each possible cell type combination drops by a factor of 2.27 ± 4.39, ranging from 1.05 to 40.03. The ratio between the SDs of connection probabilities (i.e., predicted *in vivo* connection probability divided by predicted *in vitro* connection probability) is 1.33 ± 0.99, ranging from 0.93 to 9.65. The ratio between the CVs of the connection probabilities is 0.80 ± 0.17, ranging from 0.24 to 1.01, showing on average an increase of the CV of connection

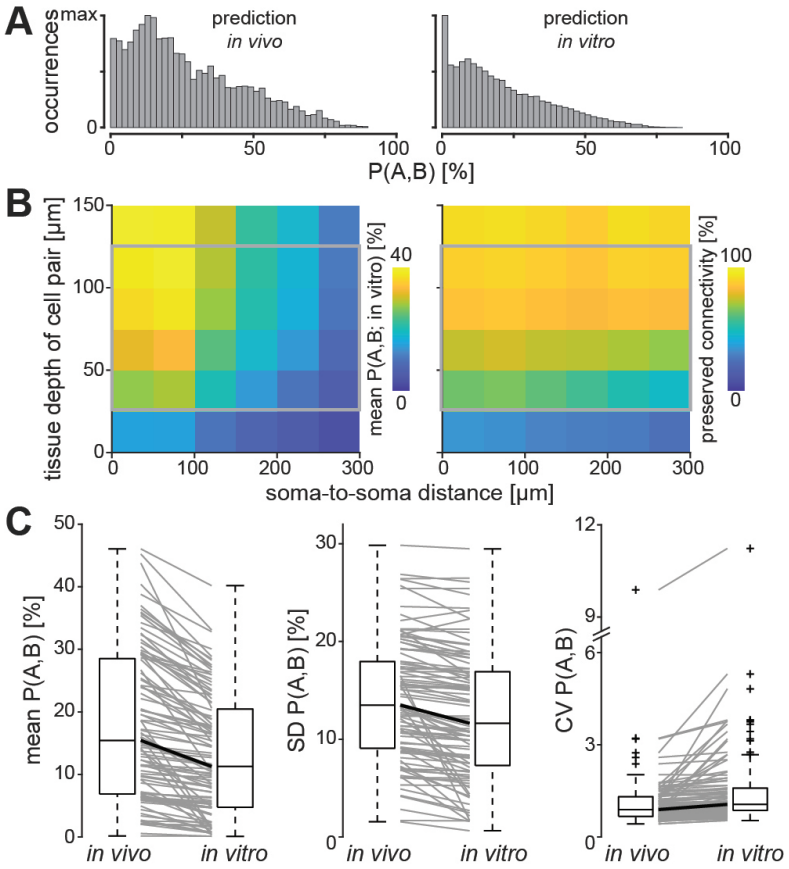

Figure 3.20. Impact of slicing on predicted connection probability statistics. (A) Example histograms of predicted *in vivo* (left) and *in vitro* (right) connection probabilities between L5 pyramids within one barrel column. (B) Left: Predicted *in vitro* connection probabilities for different soma-to-soma distances and tissue depths. Gray rectangle represents connection probabilities of right histogram in panel A. Right: Preserved connection probability in emulated slices (i.e., ratio between predicted *in vitro* and *in vivo* connection probabilities). (C) Box plots represent predicted *in vivo* and *in vitro* mean (left), SD (center), and CV (right) of the connection probabilities across all cell type combinations in one barrel column. All neuron pairs were located between 31 µm and 130 µm within the slice. Black line represents the median of connection probabilities.

probabilities in the emulated slice.

3.2.2 Pair-wise connectivity involving inhibitory neurons

In this Section, I present the results of the predicted connectivity involving INs (see Figure 3.21), specifically their predicted thalamocortical input from VPM axons. Note that the following Result section is part of an ongoing effort in the group of Dr. Marcel Oberlaender, and therefore, its conclusiveness is limited.

Thalamocortical input to inhibitory neurons

I investigated how the dense statistical connectome predicts INs to be innervated by VPM thalamus depending on their soma depth locations and intracortical axon projection patterns. To do so, the DSC between the somatodendritic domain of each IN and individual VPM axons in the dense vS1 model was calculated (see Figure 3.11). Figure 3.22A shows the soma and dendrites of an exemplary IN located in L5. The somatodendritic surface area of this exemplary IN is 7108 µm^2, comprising 5260 postsynaptic structures (POSTs). These POSTs overlap with ∼115,000 VPM boutons (PREs) and ∼6.12 million POSTs from other excitatory and inhibitory neurons. The model predicts the exemplary IN to receive 102 VPM synapses from 350 VPM neurons with a connection probability of 24%. I repeated this procedure for all 1986 INs within the C2 barrel column. In total, 10% of all predicted VPM synapses are located on INs. While the vast majority of these VPM synapses are located on INs in L4 (see Figure 3.22B), VPM synapses on INs are predicted in L2/3 and L5.

Next, I analyzed how VPM innervates each of the five inhibitory projection types (see Figure 3.12). All projection types receive VPM synapses but to different degrees (see Table 3.8; Figures 3.23 and 3.24). INs located within L4 overlap most strongly with VPM axons, resulting in 288 ± 200, 282 ± 164, and 281 ± 153 VPM synapses onto local projectors, GranSym, and GranAsc INs, respectively (see Table 3.8 and Figure 3.23). Local projectors with somata located outside the L4 barrel are predicted to receive substantially fewer VPM synapses (L2: 14 ± 13, L3: 128 ± 139, L5: 95 ± 67, L6: 33 ± 38). Similarly, SupraDesc and InfraAsc INs receive on average 52 ± 66 and 87 ± 56 VPM synapses, respectively. The number of

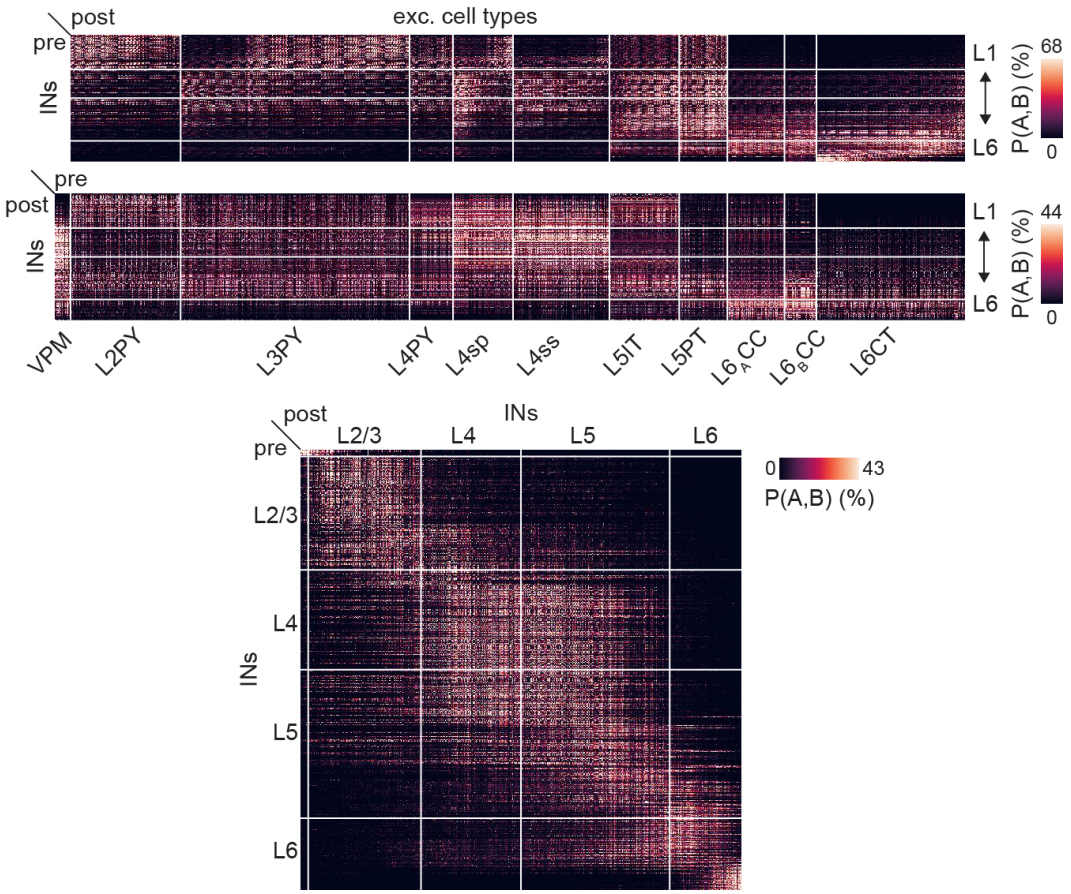

Figure 3.21. Matrix representations of connection probabilities involving INs in the C2 barrel column. From top to bottom: Connection probabilities from INs onto excitatory cell types, from excitatory cell types onto INs, and between INs. The excitatory neurons in the matrix were sorted by cell type, inhibitory neurons by cortical depth. Note the flipped pre- and postsynaptic axis in the matrix in the center; color map of each matrix representation is restricted to 95% range of the connection probabilities.

Table 3.8. Predicted number of VPM synapses onto inhibitory projection types in the C2 barrel column. "VPM syn. on L4 INs" denote the number of VPM synapses onto INs whose somata were located within the L4 barrel (i.e., 600 µm to 900 µm below the pial surface).

		Local $n = 1,034$	SupraDesc $n = 150$	GranSym $n = 181$	GranAsc $n = 200$	InfraAsc $n = 421$
VPM syn.	M±SD	149±175	52±66	188±151	214±145	87±56
	total	154,156	7,831	34,052	42,709	36,452
VPM syn. on L4 INs	M±SD	288±200	n/a	282±164	281±153	n/a
	total	109,042	0	18,640	31,806	0

putative VPM synapses onto SupraDesc and InfraAsc INs decreases with distance of the inhibitory soma to the barrel (i.e., towards L2/3 and upper L5; see Figure 3.24).

3.2.3 In-degree correlations

I analyzed to what degree the number of excitatory synaptic inputs are correlated in the dense statistical connectome, i.e., to what degree the structural composition of the neuropil can give rise to correlated networks. The analysis revealed that depending on how the neurites of the involved neuron populations are arranged and embedded within the neuropil correlations in their in-degrees emerge, ranging from positive to negative correlation (see Figure 3.25). For example, in-degrees of L5PT and $L6_ACC$ onto L5PT are positively correlated ($R = 0.64$, $n = 1086$), in-degrees of L5IT and L5PT onto L5IT are uncorrelated ($R = 0.08$, $n = 1575$), and in-degrees of L5IT and $L6_ACC$ onto $L6_ACC$ neurons are negatively correlated ($R = -0.15$, $n = 1293$). If only sampling from $L6_ACC$ neurons located within the septum, the negative correlation increases ($R = -0.49$, $n = 189$). In the model, these correlations are a direct consequence of the structural variations within grouped neurons and their spatial arrangement and embedding of neurites in the neuropil. For example, innervation volumes of L5PT and $L6_ACC$ onto L5PTs are fairly similar, giving rise to positive in-degree correlations (see Discussion Section 4.4 for details). Across all possible excitatory cell type groupings, in-degrees

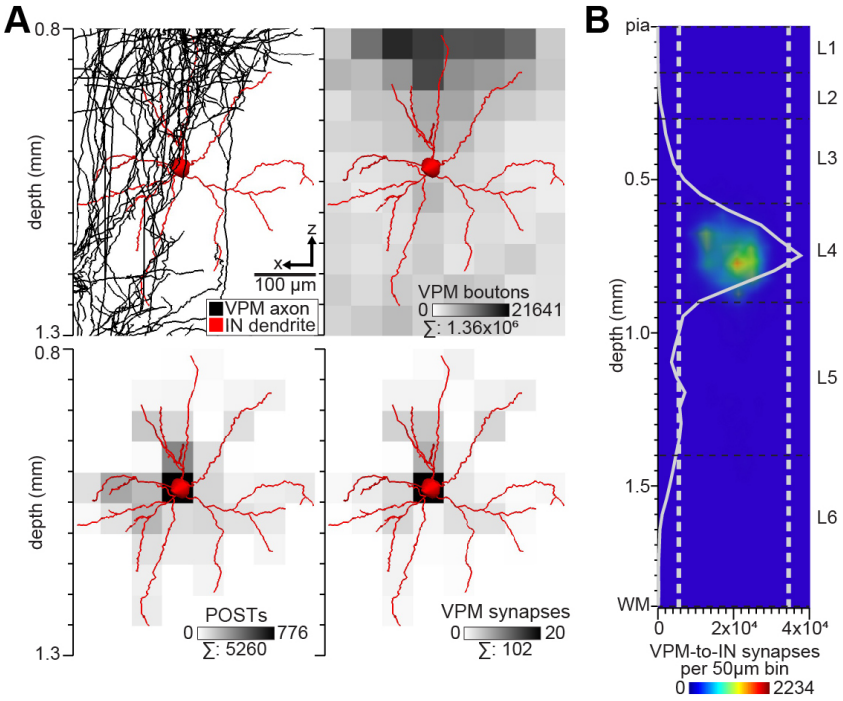

Figure 3.22. VPM input to INs across cortical depth within an average barrel column. (A) Illustration of predicting synapses between an exemplary pair of VPM axon and IN within the vS1 model. (B) Projection image of the 3D density distribution of all predicted VPM-to-IN synapses formed between VPM axons and INs.

are strongly biased towards positive correlations ($R = 0.43 \pm 0.31$, ranging from -0.67 to 0.96; see Figure A4).

3.2.4 Network topologies

In this Section, I present the results of the predicted connectivity at the network level (i.e., between three or more neurons) (see Figure 3.26). This Section is in parts adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020). Network topologies are commonly described with respect to a random network (e.g., Milo et al., 2002; Song et al., 2005; Kampa et al., 2006; Rieubland et al., 2014). Specifically, the deviation between the number of occurrences of a particular net-

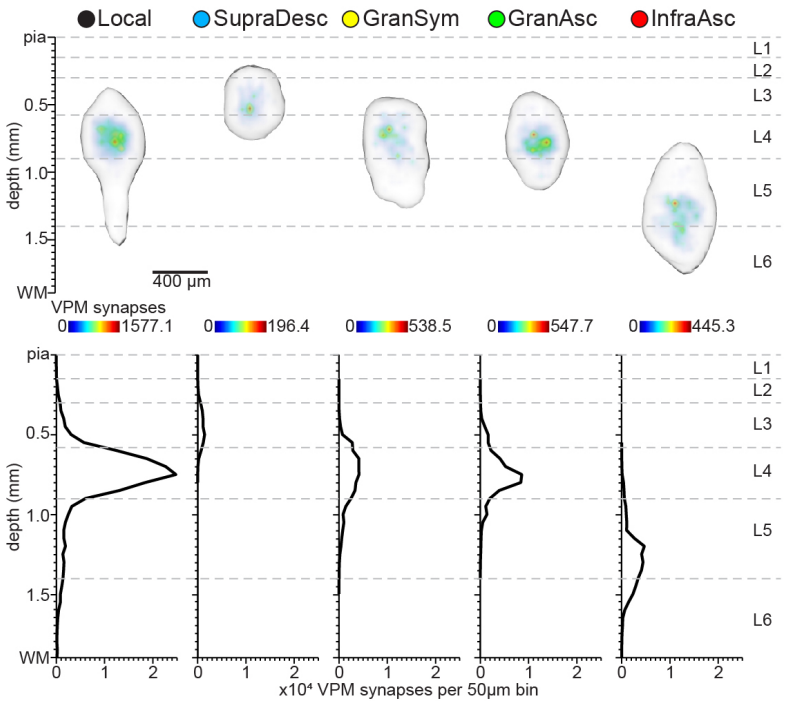

Figure 3.23. Projection type-specific VPM input. Top panel shows projection type-specific 3D distribution of predicted VPM synapses onto the respective IN projection types. The gray shaded volumes represent the 3D envelops around all VPM synapses onto the respective projection type. Bottom panel shows the vertical density profiles of VPM synapses onto each projection type.

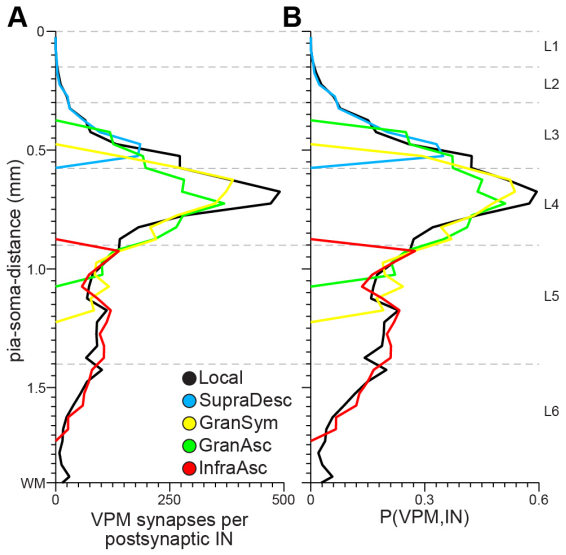

Figure 3.24. Projection type-specific depth-dependent VPM input within an average barrel column column. (A) Vertical 1D profile of the average number of VPM synapses of each IN projection type receives along their pia-soma-distance. (B) Vertical 1D profile of the average connection probability between VPM axons and INs of each projection type along their pia-soma-distance.

work motif and the number of its expected occurrences in a random network[7] is quantified (see Figure 3.27 for an illustration). A deviation of one reflects that the measured network motif occurs as often as one would expect from a random network. A deviation larger or smaller than one reflects that a network motif occurs more or less frequently than in a random network (i.e., over- or underrepresentation, respectively). Figure 3.26A illustrates a network predicted by the model and a random network with the same average connection probability (i.e., same number of edges). While in the exemplary graph network, the recurrent feedforward triplet motif (motif ID 9) is predicted to occur three times, the same motif is only expected to occur once in the random network. This yields a deviation of 3, an overrepresentation of the recurrent feedforward motif.

Triplet motifs: When determining the deviation of each triplet motif from a random network for different intracortical cell type combinations, I find that only

[7] In a random network, each edge has the same mean connection probabilities; variations and correlations are absent.

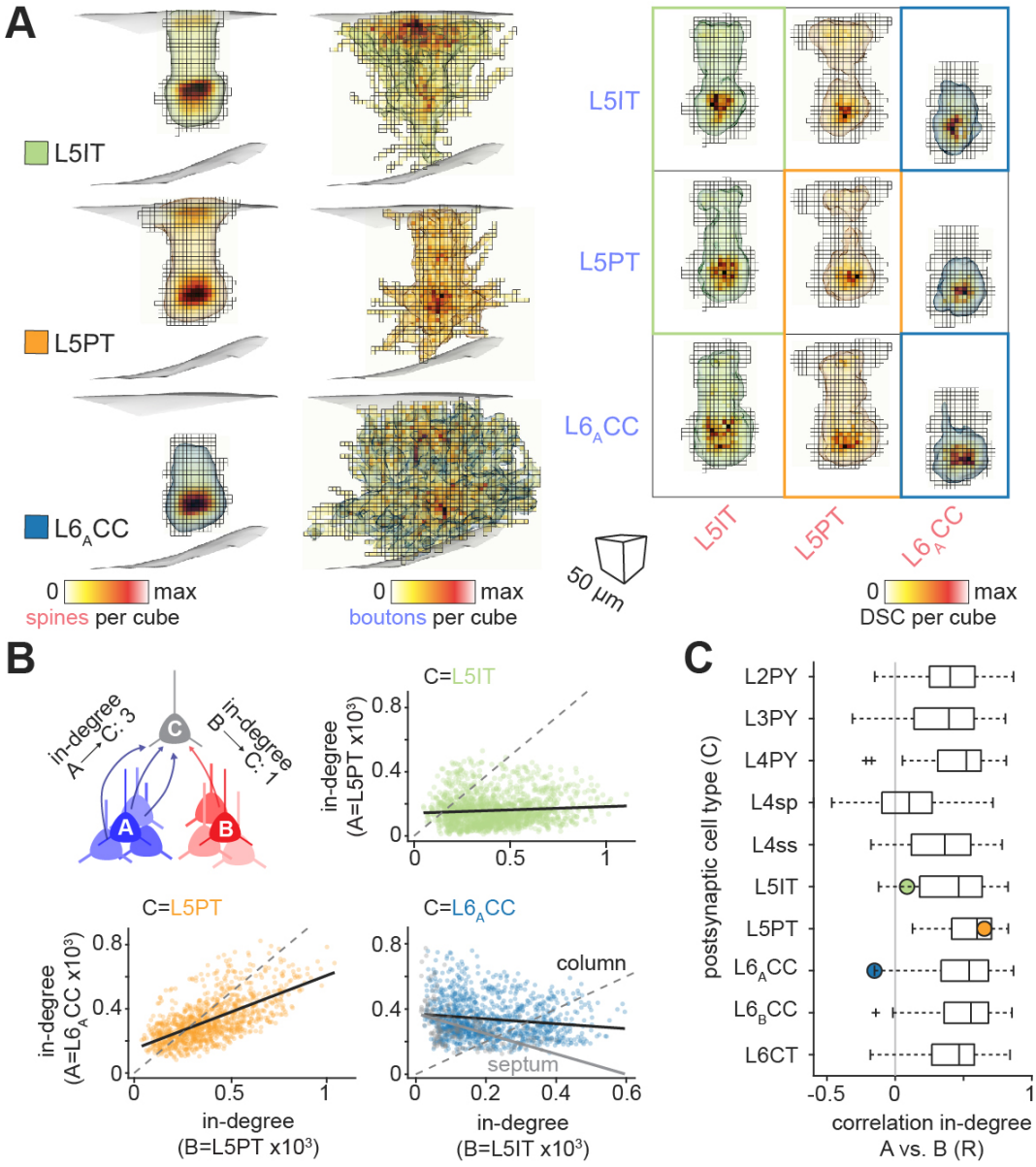

Figure 3.25. Dense statistical connectome predicts in-degree correlations. (A) Left: 3D densities of spines and boutons of L5IT, L5PT, and L6$_A$CC neurons located in C2 barrel column. Superimposed 50 µm grid shows maximal extent, isosurface represents 95th and 99th percentile of spine and bouton projection volumes, respectively. Right: Resultant DSC of all cell type combinations. (B) Correlations between the numbers of synapses of neurons from population A and B that connect to a neuron of population C (i.e., in-degree), exemplified for neuron populations of panel A. Respective DSC densities are highlighted in panel A (right). *Caption continued on next page.*

Figure 3.25 — continued from previous page. (C) In-degree correlations between all possible cell type groupings. Figure adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020).

a few motifs are as frequent as in a random network. The deviations from a random network vary between motifs and depends on the grouping of neurons (see Figures 3.28 and 3.27). For example, the deviation of triplet motifs comprising L5ITs and L5PTs neurons differs from those only comprising L5ITs. In Figure 3.28 the distribution of the motif deviation across 220 possible cell type groupings is shown. Motifs 1, 2, 3, 4, and 6 are in at least 99% of the cases overrepresented, while motifs 7, 13, 14, and 15, are in at least 98% of the cases underrepresented. A more diverse range of exemplary motif deviations is depicted in Figure A7 that includes both excitatory and inhibitory neurons grouped by different layers.

Most likely cell type combination for each triplet motif: I mimicked an experiment in which the triplet motif between three randomly targeted excitatory neurons was determined, but the cell types involved were unknown. Across 500,000 random neuron triplets, I determined the most likely cell type configuration for each triplet motif (see Figure 3.29). The dense statistical connectome predicts that the most likely cell type configuration for the fully recurrent triplet motif is one L3PY and two L2PYs. 13% of all fully recurrent triplet motifs are predicted to have that particular cell type configuration. Most triplet motifs are comprised of the most frequent cell types in vS1 (i.e., L3PY: $n = 109,854$, L6CT: $n = 89,766$, L2PY: $n = 61,369$, and L4ss: $n = 47,214$; see Table 3.1).

Network motifs between more than three neurons: By grouping all neurons in the vS1 model together, the model predicts higher-order connectivity statistics to deviate from those of random networks (see Figure 3.26C). These deviations are, however, not arbitrary. For example, the more recurrent a motif is (i.e., number of bidirectional edges), the more overrepresented it becomes compared to the random network. In contrast, chain motifs become more underrepresented the more neurons are involved.

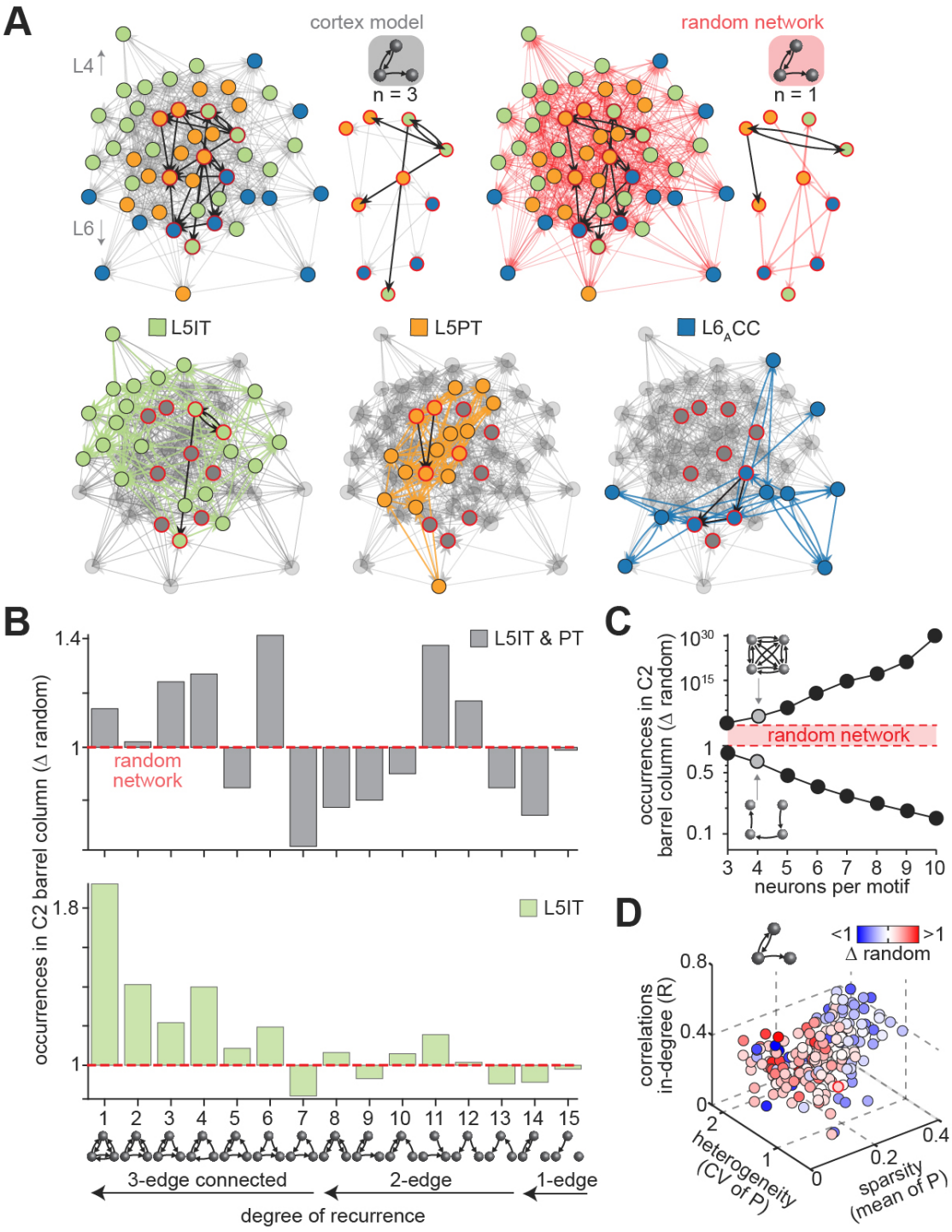

Figure 3.26. Caption on next page.

Figure 3.26. Network topologies. (A) Graph of predicted network of fifty L5 neurons (top left; gray) and its respective random network with same number of edges (top right; red). Colored nodes represent different cell types (cell type-specific subnetworks shown in bottom panel); black edges represented highlighted subnetworks. Zoom-in to subnetwork of ten nodes: Highlighted black arrows show occurrences of exemplary triplet motif as predicted in the model ($n = 3$; gray, left) and as expected in the random network ($n = 1$; red, right). (B) Occurrences of triplet motifs for two different neuron groupings compared to the occurrences expected in a random network (red dashed line). (C) Model predicts topological differences to random networks to increase with the number of neurons (i.e., nodes) per motif. (D) Deviation of recurrent feedforward motif from random network for 210 cell type-specific triplet combinations as a function of their average in-degree correlations and CV and mean of their respective connection probabilities. Deviations of remaining triplet motifs are shown in Figure A5. Figure adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020).

3.2.5 Relationship between triplet motif deviation and their underlying wiring properties

Note: The following section is in parts adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020).

I investigated the relationship between the deviations of triplet motifs from a random network and their underlying wiring properties (see Figures 3.26D and A5). For example, in case of the recurrent feedforward motif (motif ID 9) the respective motif deviation is negatively correlated with the mean of the connection probabilities (i.e., $R = -0.47$, $p = 2.61 \times 10^{-13}$), positively correlated with their CV (i.e., heterogeneity; $R = 0.40$, $p = 6.49 \times 10^{-10}$), and negatively correlated with the average in-degree correlations ($R = -0.26$, $p = 1.85 \times 10^{-4}$). In general, motif deviations increase with sparsity and heterogeneity ($R = -0.43 \pm 0.18$ and $R = 0.42 \pm 0.18$, respectively) but decrease with increasing in-degree correlations ($R = -0.26 \pm 0.26$) (see Figure 3.30). These results indicate a relationship between the sparsity, heterogeneity, and correlation of the underlying connection probability distributions and the deviation of motif occurrences from a random network.

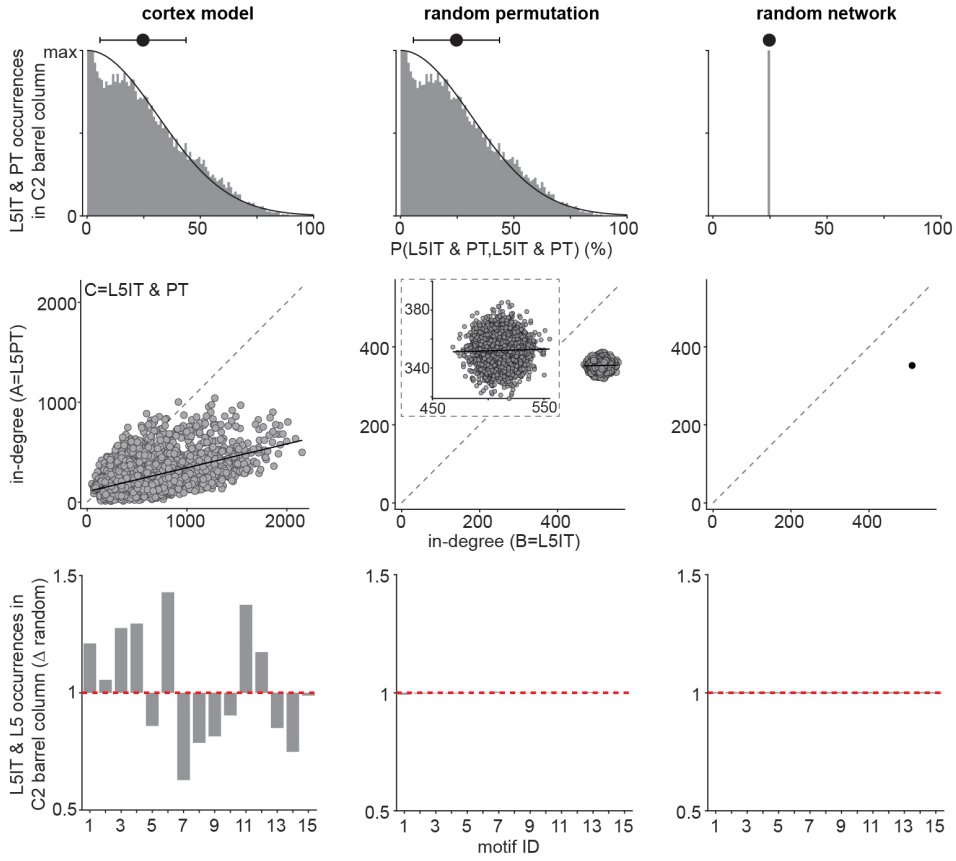

Figure 3.27. Illustration of predicted networks vs. random networks. From left to right column: dense statistical connectome, random permutation of pair-wise connectivity statistics, and random network. Top panel: Histogram of connection probabilities in the three networks. Random permutation preserves shape of the distribution, random network only preserves average connection probability. Center panel: In-degrees are only correlated in the dense statistical connectome and not in the random networks. Bottom panel: Network topologies do only deviate from a random network in the dense statistical connectome. No deviation in both random networks.

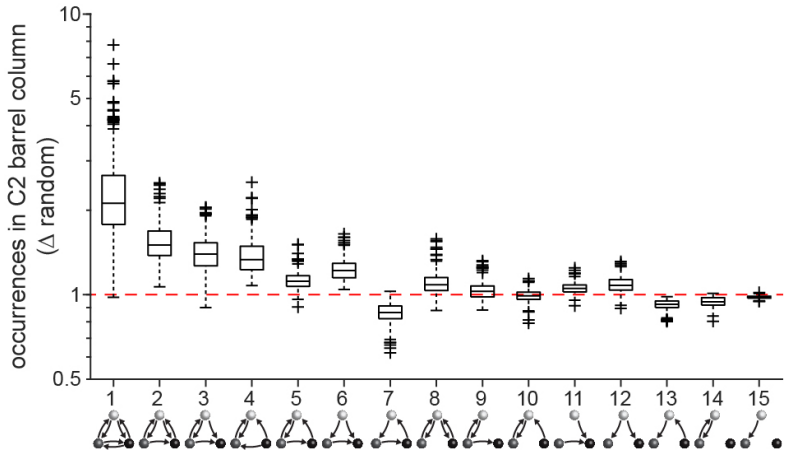

Figure 3.28. Deviations of triplet motifs from random network across intracortical cell type triplets ($n = 220$) vary.

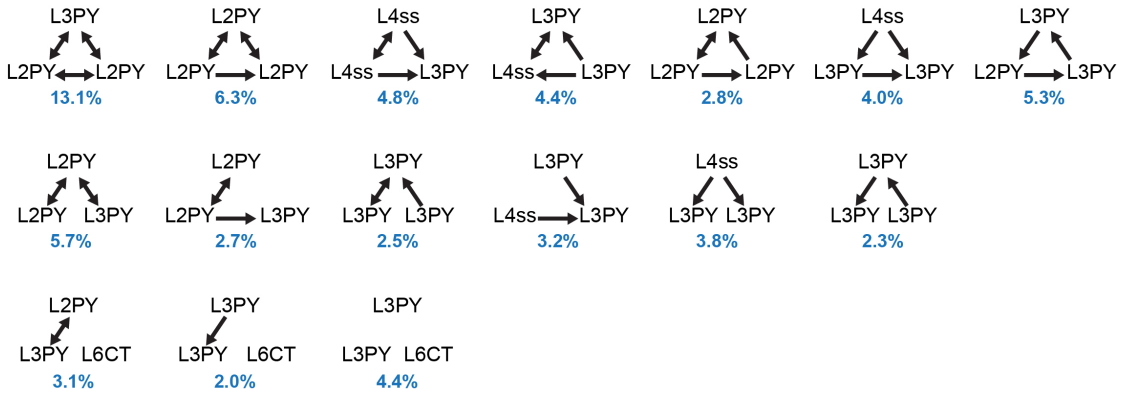

Figure 3.29. Most likely cell type configuration of each triplet motif as predicted by the model. Probability of observing the respective cell type configuration given the respective motif shown in blue.

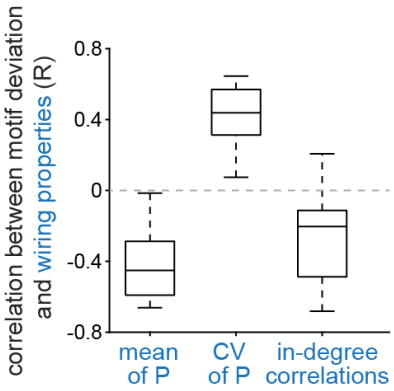

Figure 3.30. Range of correlations between triplet motif deviations and underlying wiring properties (mean connection probability, CV of connection probabilities, and in-degree correlations) across 210 cell type-specific triplet.

To better understand the impact of these wiring properties on motif deviation, a mathematical model of correlated connectivity was generated in collaboration with Prof. Dr. Jakob Macke (University Tübingen) (see Section 2.3.4 for details). This model allows relating the motif deviation with its mean connection probabilities and with a simplified measure of the degree of correlation and heterogeneity, referred to as λ (see Figures 3.31 and A6). This mathematical model revealed that the recurrent feedforward motif and the feedforward motif both have regimes of under- and overrepresentation of motif occurrences compared to a random network. The fully recurrent motif, however, is never underrepresented. I mapped the predicted properties of the connectivity distributions onto the predictions of the mathematical model of correlated connectivity. The mapped values of λ ranged from 0.18 to 0.74. The mapping revealed that the respective region of the feedforward motif are underrepresented, of the recurrent feedforward motif both under- and overrepresented, and of the fully recurrent motif overrepresented — consistent with the predictions of the dense statistical connectome. A direct comparison between the predicted motif deviation by the dense statistical connectome vs. the mathematical model yielded positive correlations ($R = 0.38 \pm 0.20$). For motif IDs 8, 11, 12, and 15 the correlations were not significant (i.e., $p > 0.01$; see Figure A6).

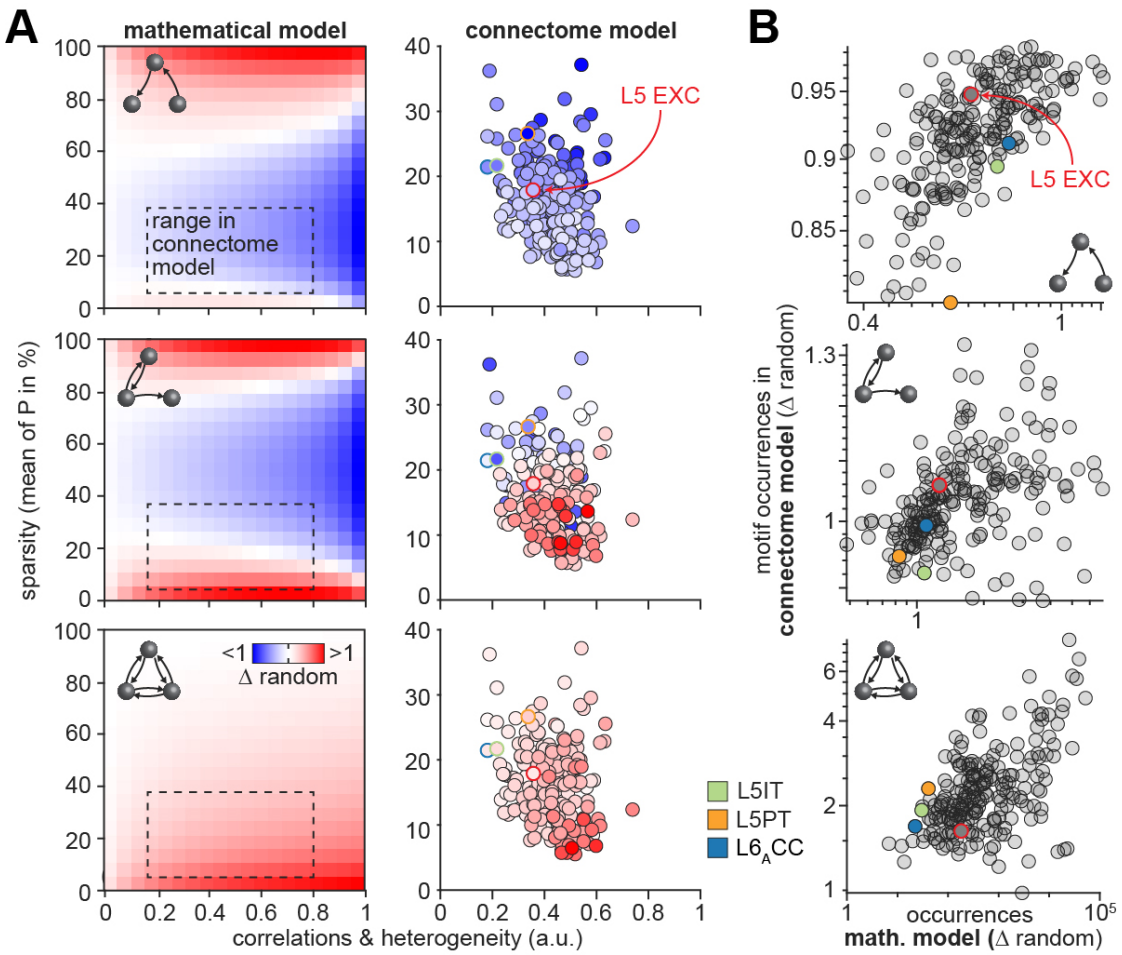

Figure 3.31. Wiring properties shape triplet motif deviations. (A) Left: Relationship predicted by mathematical model between sparsity (i.e., average connection probability), a measure of the degree of heterogeneity and correlation, and the deviation of three exemplary triplet motifs (from top to bottom: feedforward motif, motif ID 13; recurrent feedforward motif, motif ID 9; fully recurrent motif, motif ID 1). Motif deviation changes with the underlying connectivity properties. Dashed box represents predicted range from vS1 connectome model. Right: In the dense statistical connectome the occurrences of the three exemplary triplet motifs deviate from those expected in random networks for the cell type-specific groupings. (B) Deviation of each exemplary triplet motif predicted by dense statistical connectome vs. mathematical model. Remaining motifs are shown in Figure A6. Figure adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020).

3.3 Comparison with empirical data

Given the predictions of the dense statistical connectome, the question arises whether these predicted wiring diagrams represent observed neocortical wiring. To answer this question, I derived the predicted probability distributions of observable wiring properties and compared them to empirically determined wiring properties. This Section is subdivided into five parts, describing first the results of the systematic comparison between empirical and predicted connectivity of excitatory neurons at the subcellular and cellular level (see Sections 3.3.1 and 3.3.2) and of INs (see Section 3.3.3). In the last two parts, I describe the results of a comparison between observed and predicted network topologies (see Section 3.3.4) and of a direct comparison of putative synapse patterns between pairs of *in vivo* labeled excitatory neurons with the model's predictions (see Section 3.3.5). The descriptions in Sections 3.3.1, 3.3.2, and 3.3.4 are adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020).

The most important results of the are:

- Consistent with empirical data, the dense statistical connectome predicts the majority of close-by dendrite and axon branchlets to remain unconnected, despite their spatial overlap. The empirically determined occurrences of synaptic contacts between the same branchlets are consistent with the model's predictions.

- A systematic comparison of 89 empirical connection probabilities between excitatory neurons and their respective predicted connection probabilities revealed a high degree of consistency — independent of the experimental conditions and the grouping of the neurons ($R = 0.75$). Most observations of location- and subcellular-specific connectivity distributions between excitatory neurons are consistent with the model's predictions.

- A comparison between seven empirically determined ranges of the number of synapses per connection between excitatory neurons with the model's predictions is not consistent ($R = -0.44$).

- A limited comparison between 73 empirical and predicted connection probabilities involving INs revealed a high degree of consistency, that, however, is lower than for excitatory neurons ($R = 0.54$). While most observations of subcellular connectivity involving INs are within the predicted distributions, individual empirical densities of inhibitory-to-inhibitory synapses exceed those predicted by the model.

- A comparison between six empirically determined ranges of the number of synapses per connection involving INs with the model's predictions is not consistent ($R = -0.02$).

- The majority of predicted nonrandom network topologies are consistent with empirical data from various studies in terms of their over- and underrepresentation ($R \geq 0.64$). Only a few individual network motifs deviate in a different direction than predicted by the model.

- A direct comparison of putative synapse patterns between *in vivo* labeled and reconstructed neuron pairs with their respective model's predictions revealed that the consistency drops with the level of detail, i.e., from the cellular ($R = 0.75$) to the subcellular level ($R = 0.25$).

3.3.1 Subcellular connectivity between excitatory neurons

Note: The following section is in parts adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020).

A recent electron microscopic study revealed that the vast majority of close-by dendrite and axon branchlets are not synaptically connected (Kasthuri et al., 2015), refuting the idea that proximity implies connectivity. I compared this observation to the predictions of the dense statistical connectome. The vast majority of dendrites and axons are predicted to remain unconnected, despite their overlap within a particular cube. On average, the model predicts only $1.1 \pm 0.3\%$ of close-by neuron pairs to be synaptically connected — fractions that are consistent with those reported (Kasthuri et al., 2015) (see Figure 3.32A; Tables A3 and A4 for probability between individual neuron pairs and branchlets, respectively).

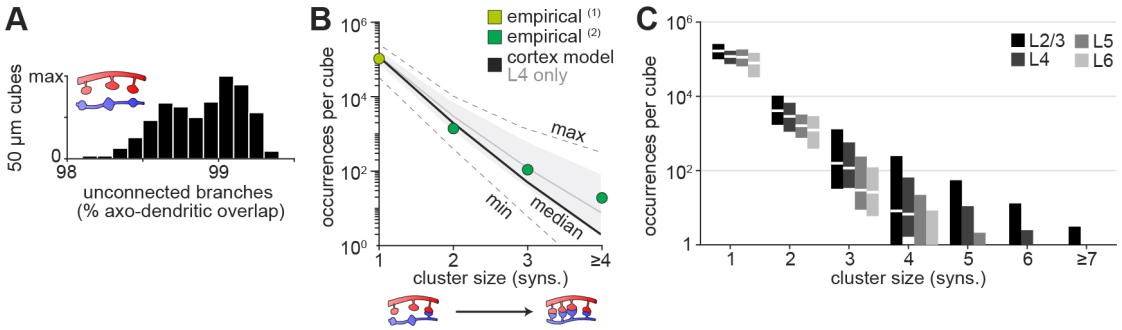

Figure 3.32. Dense statistical connectome predicts synaptic wiring properties consistent with measurements at the subcellular level. (A) The dense statistical connectome predicts that the vast majority (> 98%) of dendrites and axons that overlap within a cube are not synaptically connected. (B) The 1 % to 2 % of overlapping axons and dendrites that can be connected within a cube are predicted to form one or more synapses along the same branch. The occurrence of such clusters is predicted to decrease with the number of synapses per cluster, which is consistent with saturated reconstructions from dense electron microscopy approaches in mouse somatosensory cortex (1) (Kasthuri et al., 2015) and barrel cortex (2) (Motta et al., 2019). Black and dashed lines represent predictions from L2 to L6 (min/median/max); gray shaded area and line represent predictions for L4 only. (C) Layer-specific numbers of synapses per connection. White bars represent median predicted number of occurrences, boxes represent range of occurrences. Panels A and B adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020).

In addition, saturated reconstructions reported occurrences of two, three, and four synaptic contacts between the same branchlets, revealing a power law relationship (Kasthuri et al., 2015; Motta et al., 2019) (see Figure 3.32B). This power law relationship is consistent with the model's predictions (Kasthuri et al., 2015: $\alpha = 1.05 \times 10^5$, $\beta = -3.78$; Motta et al., 2019: $\alpha = 1.08 \times 10^5$, $\beta = -6.27$; model (minimal and maximum occurrences in brackets): $\alpha = 1.17 \times 10^5$ (3.27×10^4 to 2.55×10^5), $\beta = -5.89$ (-6.40 to -4.64). The respective occurrences varied minimally between layers (L2/3: $n_{Cubes} = 140$, $\alpha = 1.67 \times 10^5$, $\beta = -5.37$; L4: $n_{Cubes} = 84$, $\alpha = 1.19 \times 10^5$, $\beta = -5.38$; L5: $n_{Cubes} = 160$, $\alpha = 1.12 \times 10^5$, $\beta = -6.13$; L6: $n_{Cubes} = 112$, $\alpha = 7.87 \times 10^4$, $\beta = -6.01$; see Figure 3.32C). The probabilities of observing more extreme numbers of neuron pairs than the empirical ones, $p_{extreme}$, is smaller than 5% for three observations reported by Kasthuri et al. (2015): For two, three, and four synaptic contacts the probabilities are 1.96%,

1.17%, and 0.39%, respectively. $p_{extreme}$ of the observations by Motta et al. (2019) $p_{extreme}$ range between 19.9% and 62.1%. Note that the empirical data was scaled to the respective subvolumes of 50 µm cubes[8].

3.3.2 Cellular connectivity between excitatory neurons

In this Section, I present the results of the predicted connectivity between excitatory neurons at the cellular level. The first Section is adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020) and describes the results of the systematic analysis of empirical and predicted connection probabilities. In the last two sections, I present the comparison between empirical and predicted synapse distributions and number of synapses per connected neuron pair.

Connection probabilities

I compared the predictions of the dense statistical connectome with a representative set of empirical studies that measured different pair-wise connectivity statistics between excitatory neurons in the rodent sensory cortex (see Tables A5, A6, A7, and A10, and Figures 3.33, 3.37, and A8). One quantity that is commonly measured is the connection probability. Connection probabilities can be obtained with different experimental techniques and neurons are typically grouped by their cell type, long-range projection target, genetic marker, or laminar position.

One such experimental technique is *in vivo* paired recordings. For example, Bruno and Sakmann (2006); Constantinople and Bruno (2013) reported connection probabilities between VPM axons and neurons in vS1 whose laminar positions or projection types were identified (see example measurement in Figure 3.33A top panel). Their measurements revealed that connectivity from VPM thalamus depends either on the laminar location or, within L5, on the long-range projection target of the postsynaptic neuron. The shape and positioning of these two projection types are alike, i.e., they intermingle in cortical depth and their dendrites have similar innervation domains (Narayanan et al., 2015; Oberlaender, de Kock, et al., 2012). Their most distinctive morphological criteria is the thickness of their

[8]scaled by a factor of 83.3 (Kasthuri et al., 2015) and 0.25 (Motta et al., 2019); for details, see Methods Section 2.4.1.

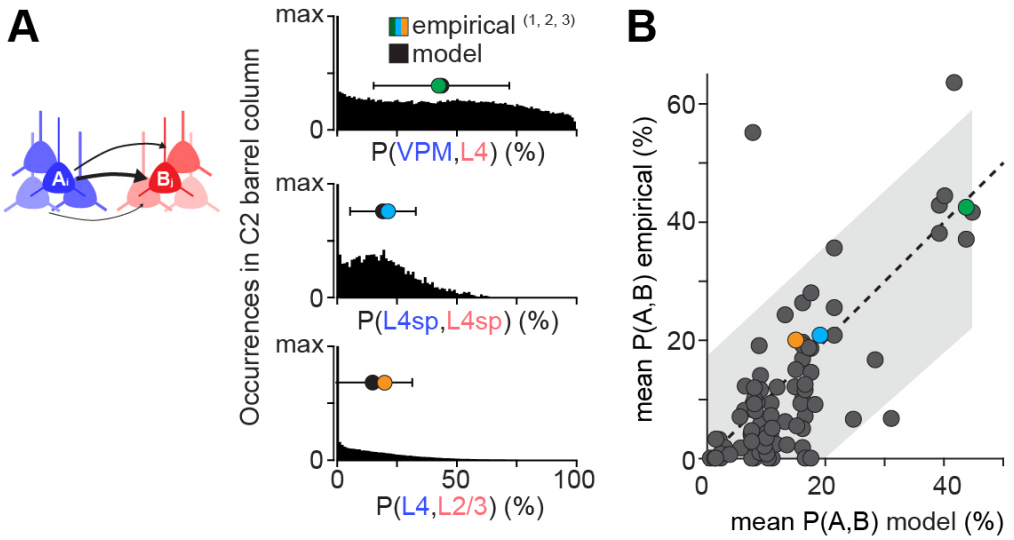

Figure 3.33. Dense statistical connectome predicts synaptic wiring properties consistent with measurements at the cellular level. The means of connection probability distributions, as predicted by the dense statistical connectome for neurons grouped by layer and/or cell type are consistent with measurements. (A) Exemplary empirical connection probabilities show connection probability distributions between different excitatory neuron populations. Empirical data: (1) Bruno & Sakmann, 2006, (2) Feldmeyer et al., 2006, (3) Yoshimura et al., 2005. (B) In total, 89 measurements (including the three examples in panel A) were compared to the respective model predictions. Figure adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020).

dendritic tuft. I mimicked their experiment by grouping the neurons accordingly (i.e., by layer and/or projection type[9]) and calculated their connection probability distributions in the model. The predicted mean connection probabilities are consistent with the measured ones, even for the two different projection types in L5 (e.g., empirical vs. model P(VPM,L4), P(VPM,L5IT), and P(VPM,L5PT) of 43% vs. 44 ± 28%, 17% vs. 28 ± 21%, and 44% vs. 40 ± 24%, respectively. For all the results and distributions, see Table A10 and Figure A8). Another, more common experimental technique to obtain connection probabilities, are *in vitro* paired recordings. There, connection probabilities are determined between neurons in slice tissue. After emulating the slicing procedure (i.e., truncating the neurons' morphologies) in the model, the predicted connectivity is consistent with the *in*

[9] L5PT and L5IT neurons

vitro measured one (e.g., Sun, Huguenard, & Prince, 2006: empirical vs. model P(L4sp,L4sp) of 21% vs. $19 \pm 14\%$; see Figure 3.33A center panel). Even connection probabilities inferred based on optical stimulations *in vitro* match with the model's prediction (e.g., Yoshimura et al., 2005: empirical vs. model P(L4,L2/3) of 20% vs. $15 \pm 17\%$; see Figure 3.33A bottom panel). Summary statistics of the predicted connection probabilities are shown in Figure 3.34. The mean and SD of the predicted connection probabilities are on the same order of magnitude (CV $= 1.34 \pm 0.72$) and the range covered by the SD of the predicted connection probabilities is on average $25 \pm 13\%$ with a maximum of 57%.

Independent of the experimental conditions and the grouping of the neurons, the majority of the 89 empirical connection probabilities reported in 29 studies that used various experimental techniques are consistent with the predicted mean connection probabilities ($R = 0.75$, $p = 2.71 \times 10^{-17}$; $p_{extreme} = 66 \pm 25\%$; $dev_{\mathrm{SD}} = -0.13 \pm 0.71$; $dev_{\mathrm{SEM}} = -2.05 \pm 8.93$[10]; see Table A10 and Figures 3.33 and A8).

The quantification of the differences between empirical and predicted connection probabilities are shown in Figures 3.33B and 3.34D. 64% of the deviation values dev_{SD} are smaller than half a SD, and 94% are smaller than one SD (4 values with $dev_{\mathrm{SD}} > 1$, one value with $dev_{\mathrm{SD}} > 4.5$). In contrast, more than 70% of dev_{SEM} are larger than two SEM. Only 14% are smaller than 1 SEM. The percentiles are $55 \pm 20\%$. The probabilities of observing a connection probability at least as extreme as the empirical one (i.e., $p_{extreme}$) are smaller than 10% for three measurements (in one case $p_{extreme} < 5\%$: empirical vs. model P(L3,L5) of 55% vs. $8 \pm 10\%$, $p_{extreme} = 0.8\%$; experiment ID 30; Thomson, West, Wang, & Bannister, 2002).

Random permutation tests: In order to test whether the high degree of consistency emerged by chance, I performed 100,000 random permutations of the 89 predicted mean connection probabilities and recalculated the different consistency measures, i.e., the correlation coefficient R, the absolute deviation values $|dev_{\mathrm{SD}}|$ and $|dev_{\mathrm{SEM}}|$, and $p_{extreme}$. Each random permutation resulted in a reduced consistency with the empirical data (see Figure 3.35). The random permutations resulted in a correlation coefficient of $R = 0.00 \pm 0.11$, ranging from -0.38 to 0.51. The

[10] $n = 87$ for dev_{SEM} due to unavailable numbers of tested neuron pairs in two measurements.

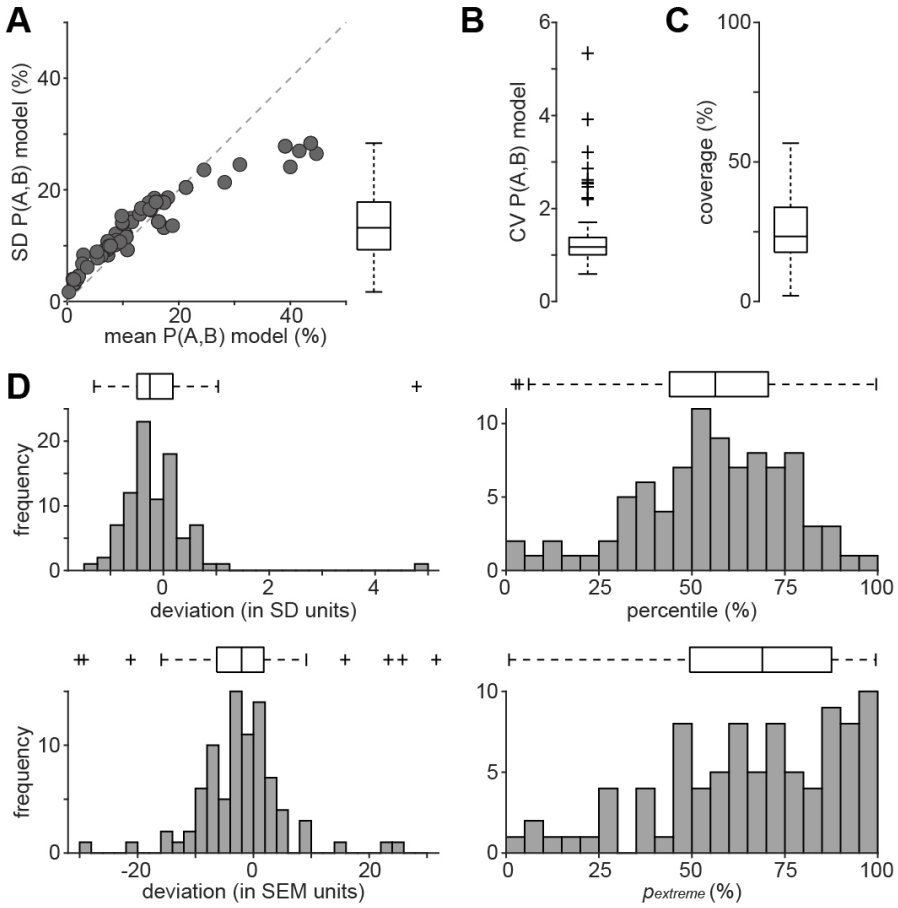

Figure 3.34. Empirical vs. predicted connection probabilities. (A) Mean and SD of predicted connection probabilities for each comparison ($n = 89$) (box plot of SD on the right). (B) CV of predicted connection probabilities for each comparison. (C) Range of probability covered by twice the predicted SD for each comparison (bounded between 0 % to 100 %). (D) Box plot (top) and histogram (bottom) of dev_{SD}, dev_{SEM}, $prctl$, and $p_{extreme}$ between empirical and predicted mean connection probabilities.

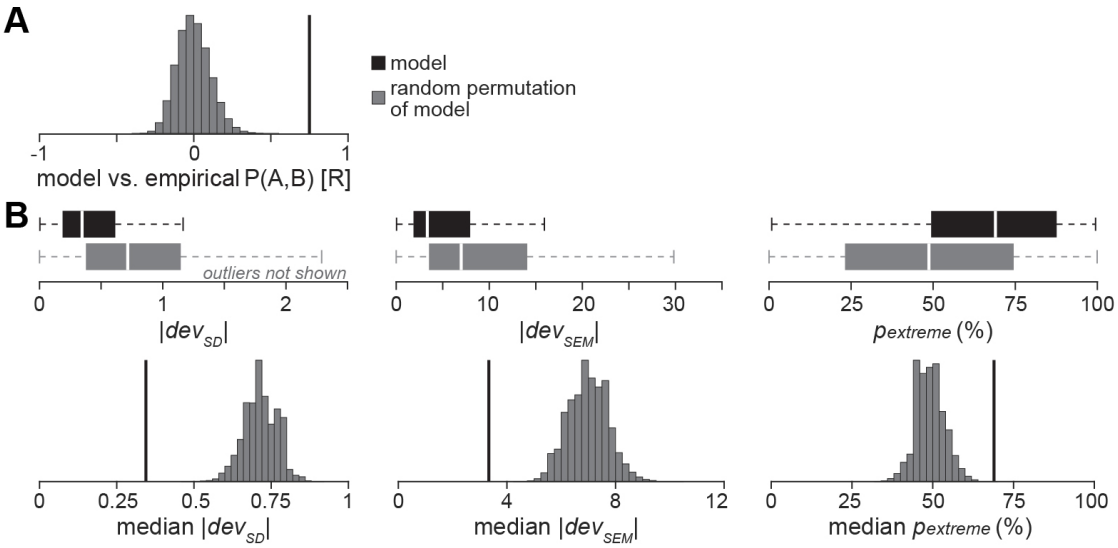

Figure 3.35. Random permutation test of consistency measures. (A) Random permutation test for correlation coefficient between empirical and predicted connection probabilities (see Figure 3.33B). Randomly shuffled predictions gray, model prediction black. Panel A adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020). (B) Top: Box plot of $|dev_{\mathrm{SD}}|$, $|dev_{\mathrm{SEM}}|$, and $p_{extreme}$ between empirical and predicted connection probabilities (black) vs. random permutation of predicted connection probabilities ($n = 8,900,000$; gray). Outliers are not shown. Bottom: Histogram of respective median values ($n = 100,000$; gray).

absolute deviation values $|dev_{\mathrm{SD}}|$ and $|dev_{\mathrm{SEM}}|$ are significantly lower compared to the randomly shuffled deviation values ($|dev_{\mathrm{SD}}|$: two-sample Kolmogorov-Smirnov test, $D = 0.36$, $p = 1.5 \times 10^{-10}$, $n = 8,900,000$; $|dev_{\mathrm{SD}}|$: two-sample Kolmogorov-Smirnov test, $D = 0.28$, $p = 2.4 \times 10^{-6}$, $n = 8,700,000$). The 89 $p_{extreme}$ values are significantly higher compared to the 8,900,000 randomly shuffled $p_{extreme}$ values (two-sample Kolmogorov-Smirnov test, $D = 0.30$, $p = 1.9 \times 10^{-7}$).

Control comparisons for different experimental conditions: I assessed how the emulation of the truncation introduced by slicing and signal attenuation due to space clamping errors impacted the correlation between empirical and predicted mean connection probabilities (see Figure 3.36). Therefore, I recalculated the correlation coefficient for four different conditions: emulating the slicing procedure vs. no emulation of the slicing procedure, and using the predicted connection

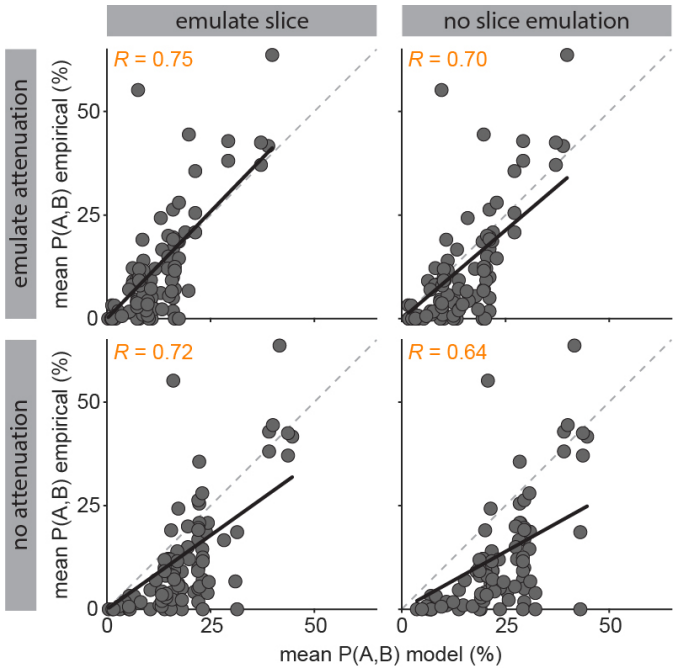

Figure 3.36. Empirical vs. predicted mean connection probabilities for slice emulation (left column) vs. no slice emulation (right column) and emulation of signal attenuation (i.e., only synapses on basal dendrite; top row) vs. no emulation of signal attenuation (i.e., synapses on entire dendrite; bottom row). Black lines represent linear fit, R the correlation coefficient.

probabilities between axons and basal dendrites only vs. axons and the entire dendrites to emulate the impact of signal attenuation. Across all conditions the correlation coefficient R between empirical and predicted mean connection probabilities is larger or equal than 0.64. The lowest correlation occurred when not taking into account the experimental conditions (i.e., no emulation of slicing and space clamp errors).

Distance-dependent connection probabilities: While most pair-wise connectivity measurements involve neurons that are in close proximity, a few studies investigated the relationship between the probability of two neurons being connected and their intersomatic distance. These studies revealed that connection probabilities change with intersomatic distance. I compared these observations to the dense statistical connectome's predictions. The measured connectivity dis-

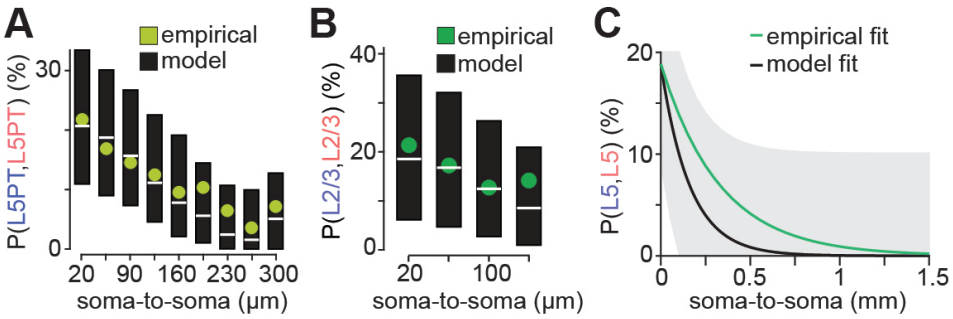

Figure 3.37. Dense statistical connectome predicts location-specific synaptic wiring properties at the cellular level. (A) Connection probability of *in vitro* paired recordings of L5PT as a function of intersomatic distance (Perin et al., 2011). (B) Connection probability of *in vitro* paired recordings in L2/3 as a function of intersomatic distance (Avermann et al., 2012). (C) *In vitro* laser uncaging of caged glutamate in L5 yields relative connection probabilities as a function of intersomatic distances (Boucsein et al., 2011). Fit of exponential decay model to empirical data (green line) and to predicted connection probabilities (black line). The gray area represents the 95% prediction interval for new observations. Figure adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020).

tributions were either obtained from paired recordings (Perin et al., 2011; Avermann et al., 2012), or from laser uncaging of caged glutamate (Boucsein et al., 2011). Consistent with the measurements, the predicted connection probabilities decrease with increasing intersomatic distances. While the decrease of connectivity measured in the laser uncaging study (Boucsein et al., 2011) is less rapid than in the model's prediction (see Figure 3.37C; $\lambda_{Empirical} = 330\,\mu m$ vs. $\lambda_{Model} = 160.8\,\mu m$, 95% confidence intervals [160.3 161.4] µm), the measured connectivity distribution of the two other studies match the predictions quantitatively (Perin et al., 2011: $p_{extreme} = 86 \pm 10\%$; $dev_{SD} = -0.16 \pm 0.12$; Avermann et al., 2012: $p_{extreme} = 90 \pm 21\%$; $dev_{SD} = -0.13 \pm 0.12$; see Tables A5 and A6 and Figure 3.37AB).

Synapse distributions along dendrites

Two empirical *in vitro* studies reported cell type- and layer-specific synapse distributions along dendrites (Qi & Feldmeyer, 2016; Petreanu et al., 2009). I compared the dense statistical connectome's prediction to those observations. Petreanu et

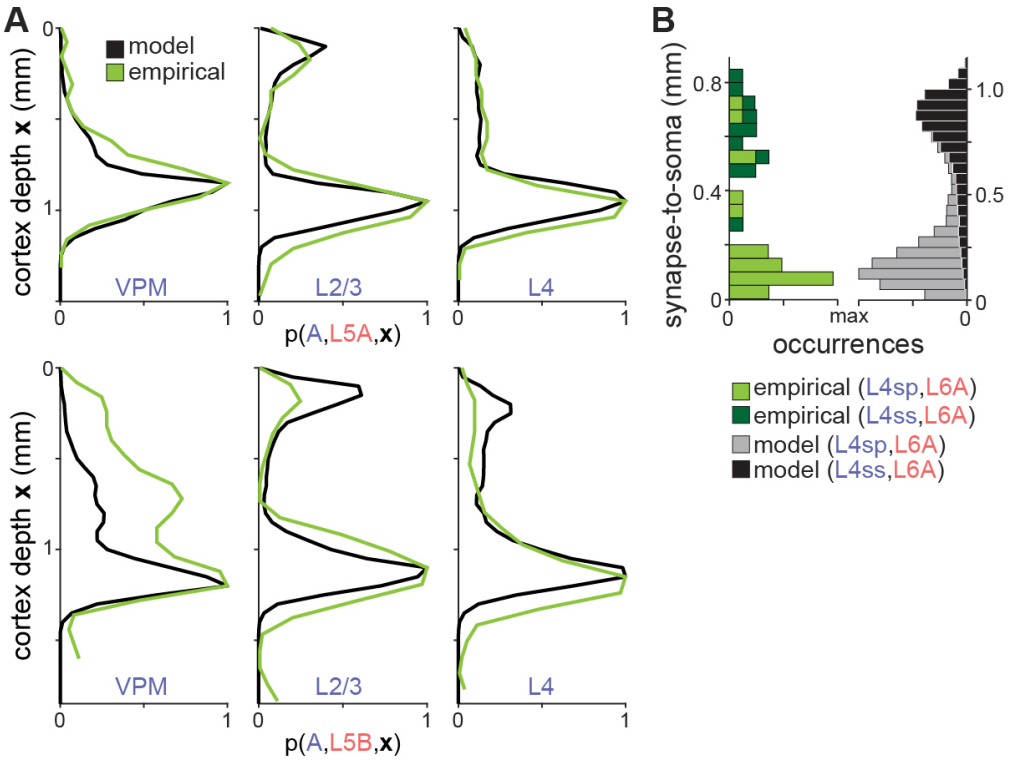

Figure 3.38. Empirical vs. predicted synapse distributions. (A) Channelrhodopsin-based *in vitro* mapping of synaptic input distributions (Petreanu et al., 2009) (green) vs. predicted distributions (black) to dendrites in L5A (top) and L5B (bottom). (B) Comparison of putative synapse distributions between *in vitro* recorded L4ss/sp and L6A neurons (Qi & Feldmeyer, 2016) (left) and predicted by the model (right).

al. (2009) performed channelrhodopsin-based synaptic input mappings in tissue slices between neurons in different cortical layers. Their technique revealed that the distributions of potential synapses onto the same neuron population depend on the laminar position of the presynaptic population. After mimicking their experiment, the model predicts distributions of synapses consistent with the measured distributions (see Table A7 and Figures 3.38A and A10).

Qi and Feldmeyer (2016) reconstructed putative synapses between *in vitro* recorded neurons. They investigated the distribution of synapses onto neurons in L6A by two cell types in L4. Their study revealed that despite having similar axon innervation domains, the two cell types in L4 preferentially target different

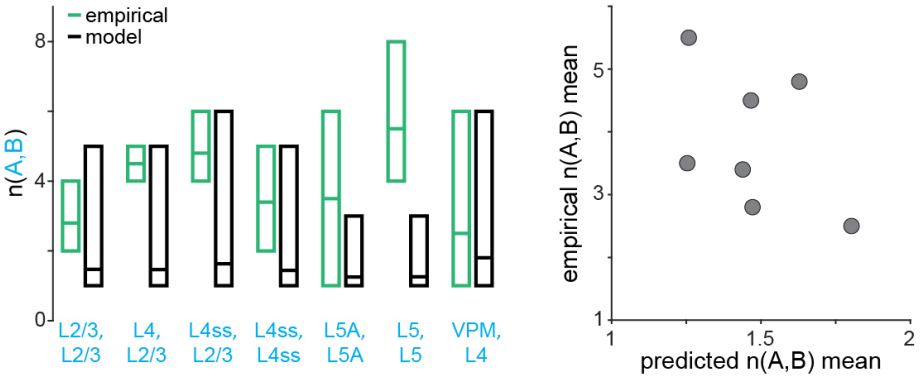

Figure 3.39. Empirical vs. predicted number of synapses per connection n(A,B). Left: Empirically determined range of the number of synapses per connection (green) vs. predicted range (black; 99% cumulative range of the average distribution). Center line represents average number of synapses per connection. The respective reference can be found in Table A8. Right: Scatter plot of empirical vs. predicted average number of synapses per connection.

dendrite compartments of L6 neurons. After mimicking their study, the model predicts synaptic distributions of L4sp and for a subset of four out of nine L4ss axon morphologies to be consistent with the measured distributions (see Figure 3.38B).

Numbers of synapses per connection

Another commonly measured wiring statistic is the number of synapses per connection. Multiple studies have reported that connected neuron pairs are typically connected not by one but by multiple synaptic contacts (e.g., Feldmeyer, Egger, Lubke, & Sakmann, 1999; Markram et al., 1997). I compared the empirically determined number of synapses per connection, referred to as $n(A, B)$, of seven studies to the model's predictions[11] (see Figure 3.39 and Table A8). Note that two comparisons (L4ss onto L4ss by Feldmeyer et al., 1999 and VPM thalamus onto L4ss by Schoonover et al., 2014) were previously reported in Egger et al. (2014).

Most of the empirically determined ranges are within the predicted ranges,

[11]Note that the prediction in this paragraph represent *in vivo* predictions (i.e., no slice emulation was performed).

except for L5 populations. However, the predicted average number of synapses per connection is not consistent with the measurements ($R = -0.44$, $p = 0.33$). In contrast to the results from paired recordings, the model predicts that the neuron pairs of the probed neuron populations are most likely only connected by a single synapse (see Figure A3 for all cell type combinations).

3.3.3 Connectivity involving inhibitory neurons

In this Section, I present a comparison between empirical and predicted connectivity statistics at the subcellular and cellular level involving INs. Note that the characterization, integration, and analysis of INs in the model are an ongoing effort in the group of Dr. Marcel Oberlaender.

Synapse distributions at the subcellular level

By examining samples of IN dendrites and somata either by electron microscopy (White & Rock, 1981; White et al., 1984; Keller & White, 1987) or by using intracellular and immunohistochemical staining combined with confocal laser scanning (Kameda et al., 2012; Hioki et al., 2013; Shigematsu et al., 2016) measurements of synapse densities were obtained. I compared these measurements with the predictions of the dense statistical connectome (see Figure 3.40).

The empirical synapse densities along dendrites is consistent with the model's predictions (empirical range vs. model synapse densities per µm: 0.09 to 3.17 vs. 1.77 ± 0.74; empirical range vs. model excitatory synapse densities per µm: 0.08 to 2.54 vs. 1.58 ± 0.68; White & Rock, 1981; White et al., 1984; Keller & White, 1987). The predicted synapse densities on neuron somata is higher than empirical values (empirical range vs. model synapse densities per µm^2: 0.04 to 0.89 vs. 0.96 ± 0.28; White et al. (1984); Keller and White (1987); empirical range vs. model excitatory synapse densities per µm^2: 0.03 to 0.8 vs. 0.85 ± 0.27; White et al., 1984; Keller & White, 1987; Shigematsu et al., 2016).

The predicted number of inhibitory synapses per dendrite length and soma surface area are within the empirical ranges reported by White et al. (1984); Keller and White (1987) (0.15 ± 0.07 vs. empirical range of 0.01 to 0.29 per µm dendrite length; 0.08 ± 0.03 vs. empirical range of 0.00 to 0.17 per µm^2 soma surface

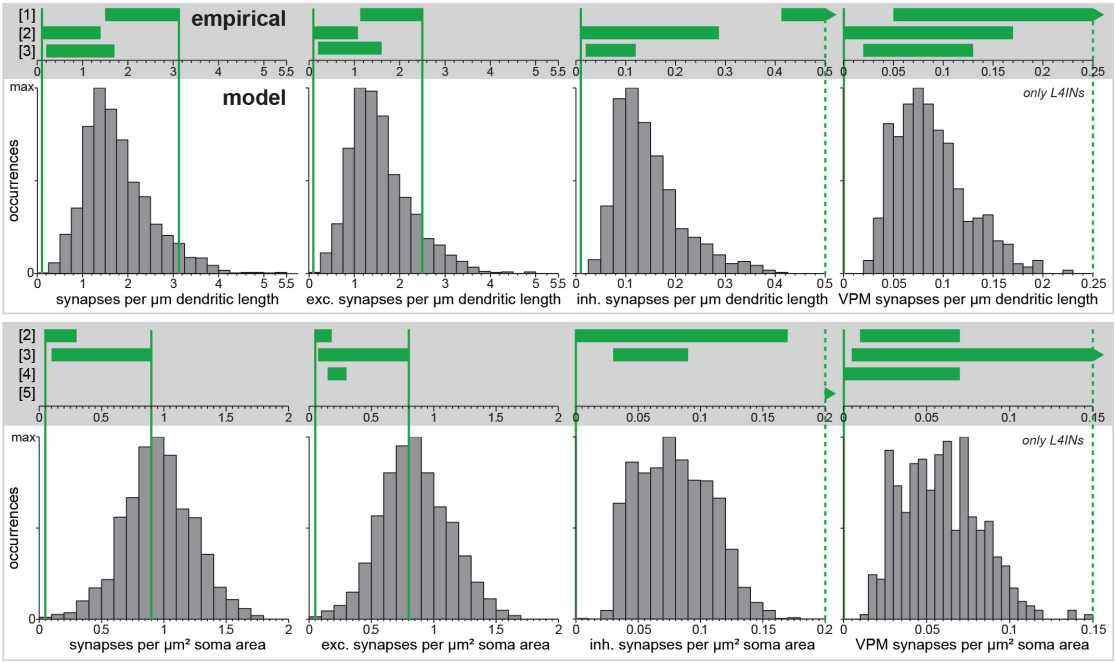

Figure 3.40. Empirical vs. predicted synapses densities on inhibitory neurons. Top: Histograms reflect predicted number of synapses (left), excitatory synapses (left-center), inhibitory synapses (right-center), and VPM synapses (right) per µm dendrite length along INs across all layers. Predicted values for VPM synapses only for INs located in L4. Green bars represent empirical ranges as reported by electron microscopic studies. [1] White and Rock (1981), [2] White et al. (1984), [3] Keller and White (1987). Dashed lines represent empirically determined range is outside of predicted histogram. Bottom: Same as top panel but for synapses per µm^2 soma surface area. Empirical ranges: [4] Shigematsu et al. (2016) and [5] Hioki et al. (2013) used intracellular and immunohistochemical staining combined with confocal laser scanning.

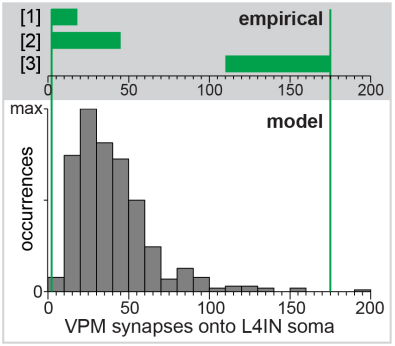

Figure 3.41. Empirical vs. predicted number of VPM synapses onto inhibitory somata in L4. Histogram reflects predicted number of VPM synapses onto each IN somata in L4. Green bars represent empirical ranges as reported by two electron microscopic studies ([1] White et al., 1984, [2] Keller & White, 1987) and one immunohistochemical study with confocal laser scanning ([3] Kameda et al., 2012).

area). However, White and Rock (1981) reported an empirical range of 0.41 to 0.63 inhibitory synapses per µm inhibitory dendrite length and Hioki et al. (2013) reported an empirical range of 0.36 to 0.47 inhibitory synapses per µm^2 inhibitory soma surface area, both exceeding the predicted numbers from the model.

The empirical numbers of VPM synapses onto L4 INs range between 0.00 and 0.55 (0.06 ± 0.05) per µm dendrite length (White & Rock, 1981; White et al., 1984; Keller & White, 1987) and between 0.00 and 0.23 (0.04 ± 0.05) per µm^2 soma surface area (White et al., 1984; Keller & White, 1987; Shigematsu et al., 2016). Consistent with these ranges, the model predicts 0.09 ± 0.04 inhibitory synapses per µm dendrite length and 0.06 ± 0.02 inhibitory synapses per µm^2 soma surface area for INs located within L4. Three studies reported absolute numbers of VPM synapses onto completely examined inhibitory somata in L4, ranging from 2 to 46 VPM synapses (White et al., 1984; Keller & White, 1987) and from 111 to 175 VPM synapses (Kameda et al., 2012). Their reported ranges are within the predicted range of 8 to 192 (42 ± 25) VPM synapses (see Figure 3.41).

Connection probabilities

Like for excitatory neurons, connection probabilities involving INs dependent on many factors, such as the location and type of the respective inhibitory neuron population. I compared 73 empirical connection probabilities involving INs reported

in 23 studies to the predictions of the dense statistical connectome. The mean and SD of the predicted connection probabilities are on the same order of magnitude (CV = 1.29±0.75) and similar to the predicted excitatory connection probabilities. However, the predicted SD and probability coverage are larger, indicating more heterogeneous connection probability distributions (inhibitory vs. excitatory SD and coverage: $22 \pm 7\%$ vs. $14 \pm 6\%$ and $45 \pm 15\%$ vs. $25 \pm 13\%$; cf. Figure 3.34). The empirically determined connection probabilities significantly correlate with the predicted mean connection probabilities ($R = 0.54$, $p = 1.01 \times 10^{-6}$; see Table A11 and Figure 3.42A-C)[12]. dev_{SD} ranges from -0.94 to 2.56 (0.40 ± 0.82). 79% of the measured connection probabilities are within 1 SD of the predicted connection probabilities. 62% ($n = 45$) of the empirical connection probabilities are higher than their respective predicted mean connection probabilities. Note that only for two measurements, the experimental conditions were mimicked; for all the other measurements, the INs were only grouped by layer (see Methods Section 2.4.3). The percentiles (i.e., percentage of connection probabilities lower than $P_{Empirical}$) are on average at $69 \pm 19\%$. The probabilities of observing connection probabilities at least as extreme as the empirical one are $56 \pm 28\%$. For no measurement $p_{extreme}$ is smaller than 5% (for four measurements $p_{extreme} < 10\%$). Random permutation of the predicted mean connection probabilities reduces consistency with the empirical data (see Figure 3.42C). The maximal correlation coefficient R for 100,000 random permutation is 0.45 ($R = 0.00 \pm 0.12$).

Numbers of synapses per connection

Similar to excitatory connections, the number of synapses per connection involving INs are typically larger than one. I compared six measurements from three studies (Koelbl et al., 2013; Helmstaedter, Staiger, Sakmann, & Feldmeyer, 2008; Silberberg & Markram, 2007) that empirically determined the number of synapses per connection of which at least one of the two connected neurons is inhibitory with the model's predictions (see Figure 3.42D and Table A9). Note that slice emulation was only performed for two predictions (L4 INs and L4 BIn onto L4ss; Koelbl et al., 2013; see Methods Section 2.4.3 for more details). For all other

[12] R increases when restricting the sample to measurements whose grouping of neurons matched the model's grouping ($R = 0.70$, $p = 7.81 \times 10^{-7}$, $n = 39$)

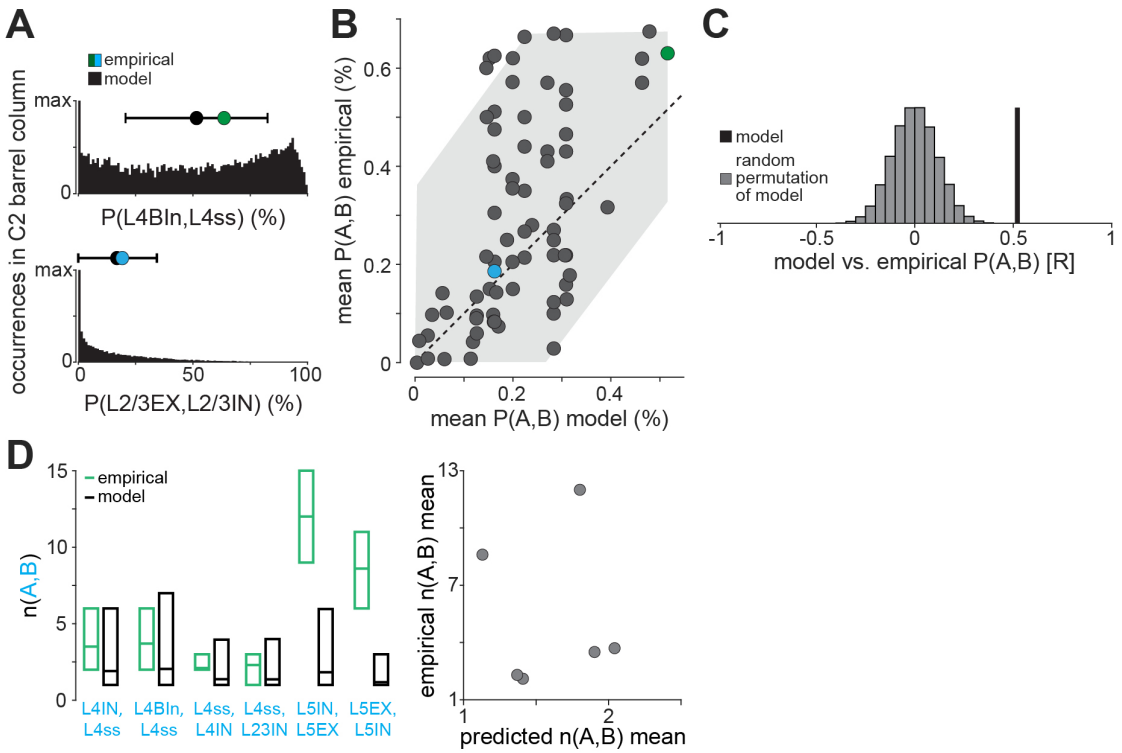

Figure 3.42. Empirical vs. predicted pair-wise connectivity involving inhibitory neurons. (A) Two exemplary connection probability measurements and its respective model's prediction (top: Koelbl et al., 2013, slice emulation; bottom: Yoshimura & Callaway, 2005, empirical subsample: L2/3 fast-spiking neurons vs. model's sample: all L2/3 INs). (B) Comparison between 73 measured connection probabilities with the respective predictions by the dense statistical connectome. The gray area represents the 95% prediction interval. Colored dots represent exemplary measurements from panel A. (C) Correlation between empirical and predicted average connection probabilities (black; panel A) vs. correlations with randomly shuffled predictions (gray). (D) Empirical vs. predicted numbers of synapses per connection $n(A, B)$ involving inhibitory neurons. Left: Empirically determined ranges of the number of synapses per connection (green) vs. predicted ranges (black; 99% cumulative range of the average distribution). Center line represents average number of synapses per connection. The respective reference can be found in Table A9. Right: Scatter plot of empirical vs. predicted average numbers of synapses per connection.

predictions slices were not emulated, although the empirical data was obtained in *in vitro* experiments. Most of the empirically determined ranges are within the predicted ranges, except for the ranges of excitatory and IN population within L5 (Silberberg & Markram, 2007). While Silberberg and Markram (2007) measured the connectivity between excitatory neurons and a subsample of INs, namely Martinotti cells, the herein presented predictions represent ranges of synapses between all INs and excitatory neurons in L5. As for the excitatory connections (see Results Section 3.3.2), the predicted average number of synapses per connection involving INs are lower than the empirical ones and uncorrelated ($R = -0.02$, $p = 0.96$). The model predicts that pairs of neurons from these neuron populations are most likely only connected by a single synapse.

3.3.4 Network topologies

Various studies investigated occurrences of circuit motifs between triplets of neurons using simultaneous multiple whole-cell recordings in slice tissue (e.g., Milo et al., 2002; Song et al., 2005; Kampa et al., 2006; Rieubland et al., 2014). By repeating multi-electrode recordings between different triplets multiple times, the number of occurrences of certain network motifs can be inferred. For each network motif, the deviation between the number of its occurrences and the number of its expected occurrences in a random network[13] is calculated. A deviation of one reflects that the measured network motif occurs as often as one would expect from a random network. A deviation larger or smaller than one reflects that a network motif occurs more or less frequently than in a random network (i.e., over- and underrepresentation, respectively). These studies revealed that motif occurrences in brain networks are unlike those expected in a random network.

Triplet motif deviations: I compared the predicted motif deviations with those determined empirically (see Figure 3.43AB). Figure 3.43A shows the empirical and predicted motif deviations of L5PT triplets (Song et al., 2005). Empirical data revealed that the unidirectional loop (motif ID 7) is underrepresented[14] compared

[13] In a random network, each edge has the same mean connection probabilities
[14] when controlling for doublet motifs. For details, see Song et al. (2005).

to a random network, while the feedforward motif with one differently directed edge (motif ID 6) is as frequent as expected in a random network. All except for one deviation predicted by the dense statistical connectome are consistent with empirical ones in terms of the direction of the deviation: In case of the unidirectional loop (motif ID 7 in Figure 3.43A), the dense statistical connectome predicts an underrepresentation, while Song et al. (2005) reported an overrepresentation. The empirical and predicted motif deviations were significantly correlated with and without controlling for doublet motifs ($R = 0.64$, $p = 0.01$ and $R = 0.68$, $p = 5.35 \times 10^{-3}$, respectively). In another study, Kampa et al. (2006) reported the motif deviations for eight triplet motifs whose neurons were located in two different layers (see Figure 3.43B). The majority of the predicted motif deviations are consistent with measured ones except for two motifs. The dense statistical connectome predicts these two network motifs to be overrepresented but Kampa et al. (2006) observed them as frequent as in a random network. Note that in contrast to the triplet motifs, doublet motifs within a cell type are randomly distributed in the model[15], hence inconsistent with empirical observations (Song et al., 2005) (see Figure 3.44A). When grouping cell types together, doublet motif deviations to random networks emerge, however, mostly inconsistent with those observed (Lefort et al., 2009; Jouhanneau et al., 2015) (see Figure 3.44B).

Network motifs between more than three neurons: A few studies investigated network motifs of more than three neurons. One such study reported 27 overrepresented quadruplet motifs for one cell type (Perin et al., 2011). I compared these results to the predictions of the dense statistical connectome (see Figure 3.43C). After computing the deviations for each of the 27 motifs, the model predicts 24 of the 27 quadruplet motifs to be overrepresented and thus, consistent with the empirical data. One motif is predicted to be random and two to be underrepresented, conflicting with the empirical data. In addition, Perin et al. (2011) determined the motif deviations of different motif sizes ranging from three to eight neurons as a function of the number of connected edges. The study reported an increase of the deviation with the number of connected edges. After reproducing

[15] due to the assumed independence between axon and dendrite morphologies within each cell type

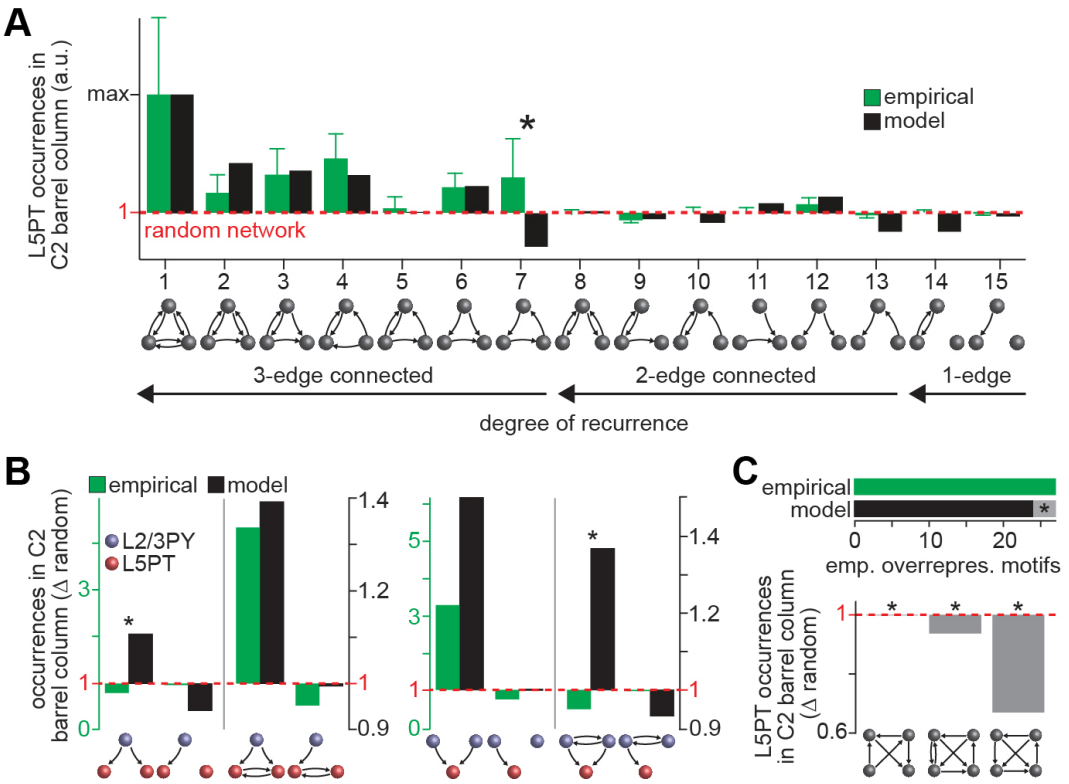

Figure 3.43. Dense statistical connectome predicts small network topologies consistent with *in vitro* measurements. (A) Comparison of empirical and predicted motif deviations of L5PT triplets from those expected in random networks (red dashed line) (Song et al., 2005). Consistent with measurements (Song et al., 2005), the model predicts that the more recurrent a motif is (i.e., number of bidirectional edges), the more frequent it is compared to a random network. Black asterisk marks prediction inconsistent with empirical deviation. Panel adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020). Note that Song et al. (2005) controlled for doublet motifs. (B) Comparison of empirical and predicted motif deviations of triplet motifs of L5PT and pyramids in L2/3 (Kampa et al., 2006). (C) Perin et al. (2011) reported 27 quadruplet motifs that are overrepresented compared to a random network. The dense statistical connectome predicts consistent with the empirical data 24 of the 27 quadruplet motifs to be overrepresented as well. Three quadruplet motifs are predicted to be not significantly overrepresented ($p < 0.0001$; black asterisk); two motifs are predicted to be significantly underrepresented compared to a random network ($p < 0.0001$).

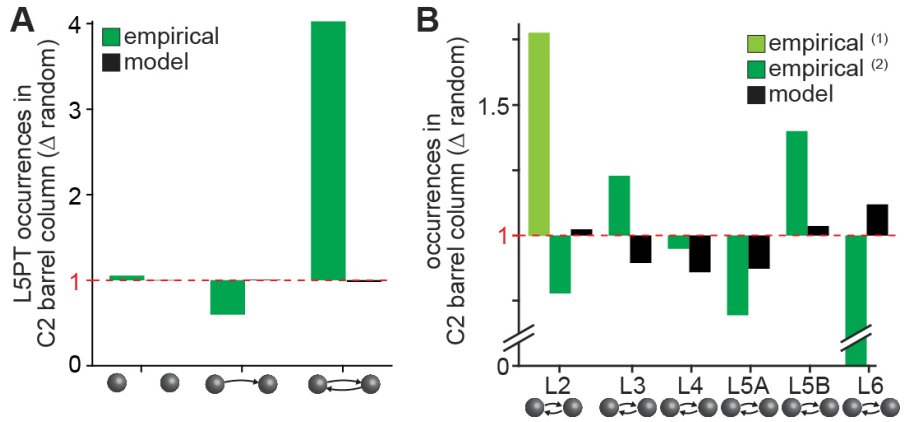

Figure 3.44. Predictions of dense statistical connectome with respect to doublet motifs. (A) Prediction of random L5PT doublet motifs inconsistent with *in vitro* measurements by Song et al. (2005). (B) Frequency of reciprocally connected neurons grouped by layer was close to their respective frequency expected in a random network. Although largely inconsistent with the measurements ([1] Jouhanneau et al., 2015, *in vivo*; [2] Lefort et al., 2009, *in vitro*), the empirical and predicted degree of deviations were relatively weak compared to those reported by Song et al. (2005). Note, Lefort et al. (2009) did not observe any reciprocally connected neurons in L6.

their experiments *in silico*, the model's predictions qualitatively resemble most of their empirical distributions (see Figure 3.45). Furthermore, the prediction that recurrent motifs are generally overrepresented, while chain motifs are generally underrepresented (see Figure 3.26C) is in line with a general theme observed in the brain networks (Milo et al., 2002; Song et al., 2005; Perin et al., 2011; Rieubland et al., 2014).

3.3.5 Connectivity between *in vivo* labeled excitatory neuron pairs

In this Section, I present a comparison of mapped putative synapses between *in vivo* labeled excitatory neuron pairs with the predictions of the dense statistical connectome using the reconstructed neuron pairs.

Dr. David A. Slabik (research center caesar, Bonn, and Max Planck Institute

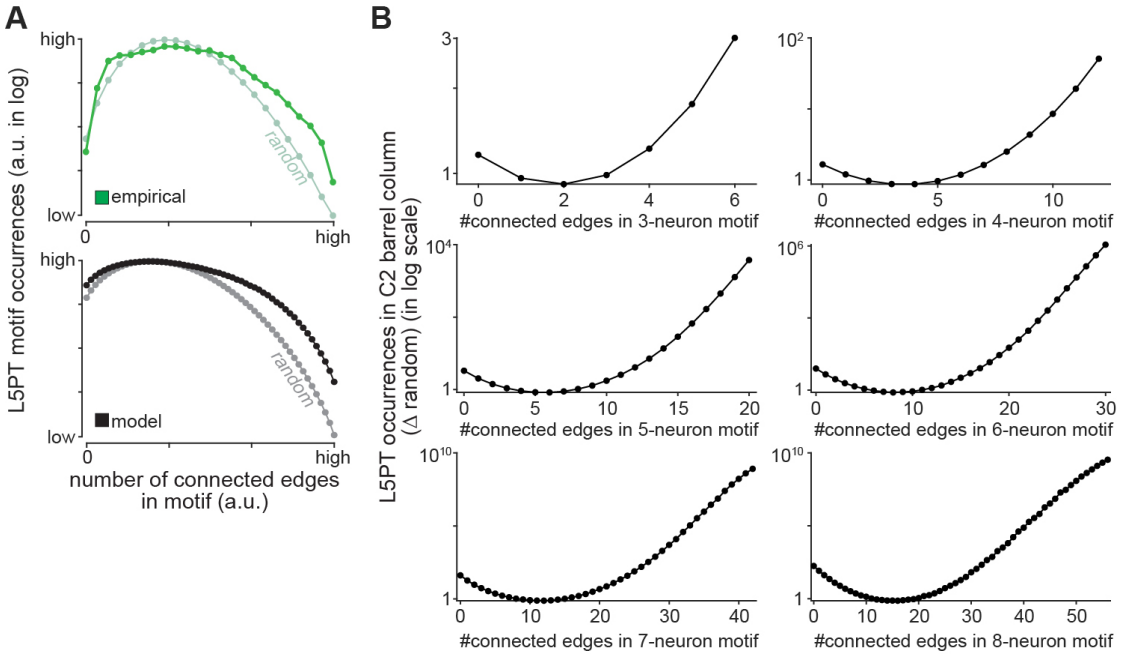

Figure 3.45. Topological differences between neocortical and random networks are predicted to increase with the number of connecting edges per motif consistent with empirical observations. (A) Top: empirical data by Perin et al. (2011) for 8-neuron motif. Bottom: respective predictions by dense statistical connectome. Bright shaded lines represent occurrences in random network, dark shaded lines empirical and predicted occurrences. Panel A adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020). (B) Predicted distribution of motif deviations as a function of the number of connected edges for motif sizes of three to eight neurons. Similar deviations were reported by Perin et al. (2011).

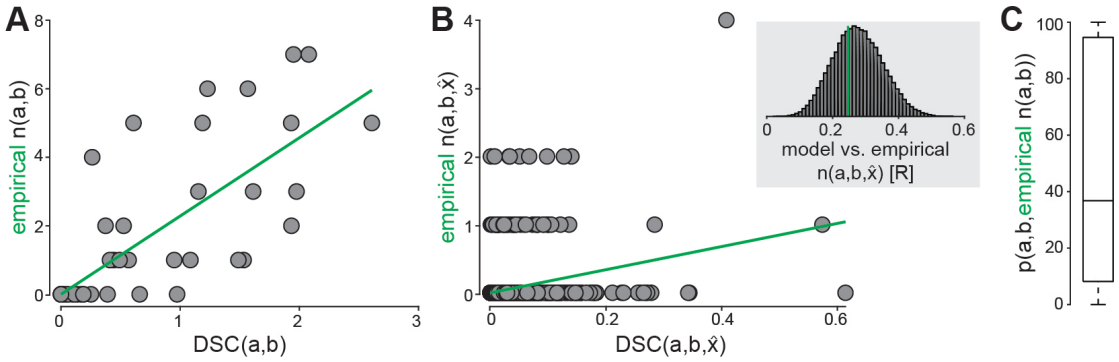

Figure 3.46. Empirical synapses between *in vivo* labeled neuron pair vs. dense structural composition. (A) Empirical $n(a, b)$ vs. $\mathrm{DSC}(a, b)$ per neuron pair. Green line represent linear regression fit. (B) Same as in panel A but per 50 µm cube $\hat{x}$. Inset: Correlation coefficients R expected from the model (i.e., between $\mathrm{DSC}(a, b, \hat{x})$ and model-drawn $n(a, b, \hat{x})$). Green arrow indicates R between $\mathrm{DSC}(a, b, \hat{x})$ and empirical $n(a, b, \hat{x})$ (from panel B). (C) Predicted probabilities of observing the empirical number of synapses.

for Biological Cybernetics, Tübingen) mapped putative synaptic contacts between completely reconstructed excitatory neuron morphologies in rat vS1 by combining *in vivo* biocytin labeling and recording of multiple neurons with high-resolution confocal imaging (Slabik, 2018). In collaboration with my colleague Dr. David A. Slabik (research center caesar, Bonn), I compared empirical putative synaptic contact distributions between pairs of entire neuron morphologies with the prediction of the dense statistical connectome. The herein analyzed data set was the result of five experiments, each with two to six *in vivo* labeled and reconstructed neuron morphologies ($n = 17$). 45 pairs of neurons were probed for putative synaptic contacts and in 23 pairs at least one putative synaptic contact was found. For readability's sake, I refer to the putative synaptic contacts determined by Slabik (2018) as empirical synapses[16].

Comparison with empirical numbers of synapses

While the empirical number of synapses $n(a, b)$ and the predicted $\mathrm{DSC}(a, b)$ on the cellular level are significantly correlated ($R = 0.75$, $p = 3.16 \times 10^{-9}$, $n = 45$ neuron

[16]Expected false positive rate of less than 15% based on a previous correlated light and electron microscopic approach (Schoonover et al., 2014)

pairs), the effect size of the correlation on the subcellular level is lower ($R = 0.25$, $p = 2.49 \times 10^{-19}$, $n = 1,292$ cubes; see Figure 3.46AB). A correlation coefficient of $R = 0.25$ is within the correlation coefficients ($R = 0.28 \pm 0.08$) expected when correlating $\mathrm{DSC}(a, b, \hat{x})$ with model-drawn $n(a, b, \hat{x})$ values (as defined in Equation 2.2). The predicted probabilities of observing the empirical number of synapses $p(a, b, n(a, b))$ range from 0.02 % to 100 % ($51 \pm 40\%$, $n = 45$ neuron pairs; see Figure 3.46C). A probability of 100% is predicted for four pairs in which the neurons do not overlap and no synapses were found. When excluding these four pairs $p(a, b, n(a, b))$ drops from an average of 51% to 46% with a maximum predicted probability of 99.8%.

Comparison with empirical synapse patterns

Given the empirical number of synapses $n(a, b)$ and the model's resolution limit of $(50\,\mu\mathrm{m})^3$, I compared the empirical synapse patterns $s_{Empirical}$ with the set of all possible combinations of cubes $\mathcal{S}$, where $n(a, b)$ synapses can be located in. For readability's sake, I refer to the combinations of cubes as synapse patterns (for more details, see Methods Section 2.4.5). Two exemplary neuron pairs and their $s_{Empirical}$, as well as the cubes predicted to most likely contain the synapses ($s_{most\,likely}$) given $n(a, b)$, are shown in Figures 3.47 and 3.48. In example I one synapse was found between a L4sp pair ($n(a, b) = 1$, $|\mathcal{S}| = 48$), consistent with the model's prediction of the most likely number of synapses ($p(a, b, n = 1) = 37\%$ compared to $p(a, b, n = 0) = 34\%$ and $p(a, b, n = 2) = 20\%$). The model predicts this one synapse to be located most likely in the same cube as the empirical one ($p(a, b, s_{Empirical}) = 4\%$, $p_{relative}(a, b, s_{Empirical}) = 11\%$, $prctl(s_{Empirical}) = 100\%$; see Figure 3.47). Synapses in any other of the 47 overlapping cubes have a lower probability than $p(a, b, s_{Empirical})$.

Example II shows another pair of L4sp, retrieved from the same experiment as Example I. This pair is connected by three synapses ($n(a, b) = 3$, $|\mathcal{S}| = 15,180$). In contrast to the empirical data, the model predicts that this pair is most likely connected by one synapse ($p(a, b, n = 1) = 32\%$), followed by two synapses ($p(a, b, n = 2) = 26\%$), no synapse ($p(a, b, n = 0) = 20\%$), and then three synapses ($p(a, b, n = 3) = 14\%$). The empirical synapses are not located inside the cubes

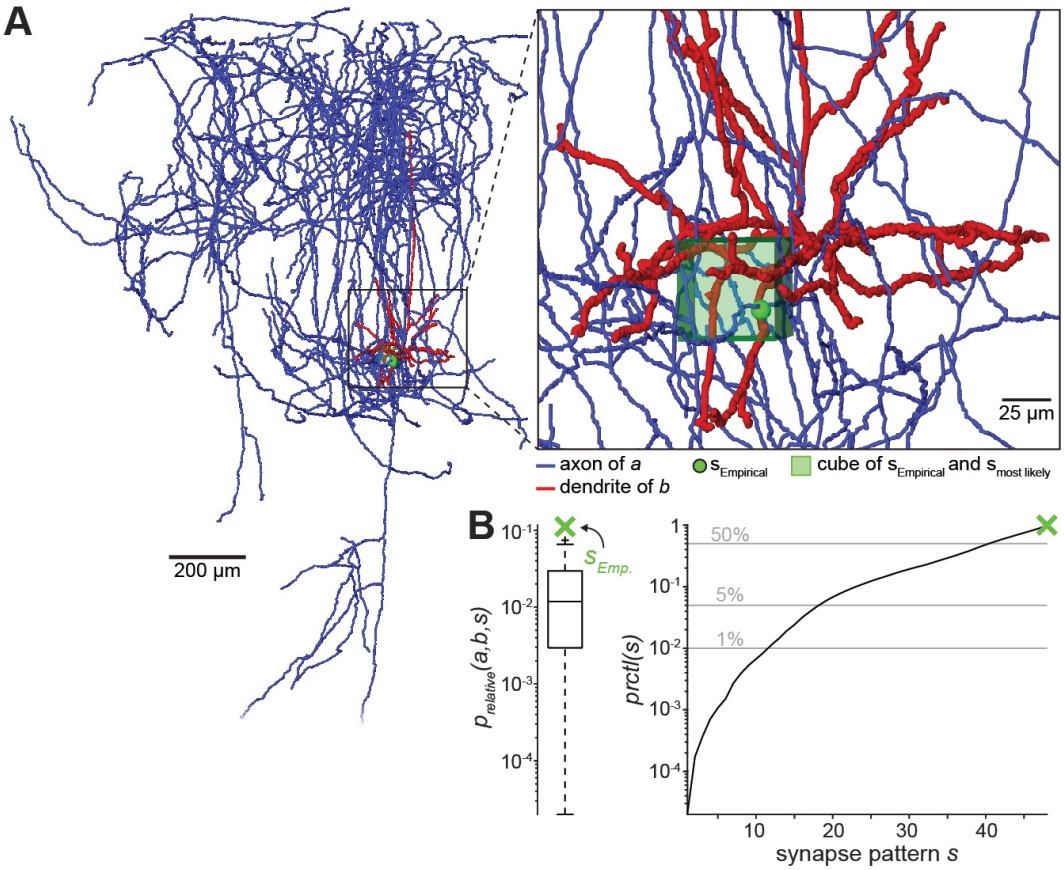

Figure 3.47. Example I: Empirical vs. predicted synapse pattern between *in vivo* labeled pair of L4sp. (A) Axon of presynaptic neuron a in blue, dendrite of postsynaptic neuron b in red. Inset: Zoom-in to the 3D location of the empirical synapse (green dot). Green shaded box represents corresponding cube $\hat{x}$, in which synapse is located in. Here, $s_{Empirical}$, in terms of cube location, matches with $s_{most\ likely}$, the cube predicted by the model to most likely contain a synapse. (B) Left: Relative probabilities of each synapse pattern s, $s_{Empirical}$ as green cross. Right: Cumulative distribution of relative probabilities of each synapse pattern s, sorted from least likely ($s = 1$) to most likely synapse pattern ($s = 48$). $s_{Empirical}$ as green cross, gray lines denote 50%, 5%, and 1% percentiles. For more details, see main text.

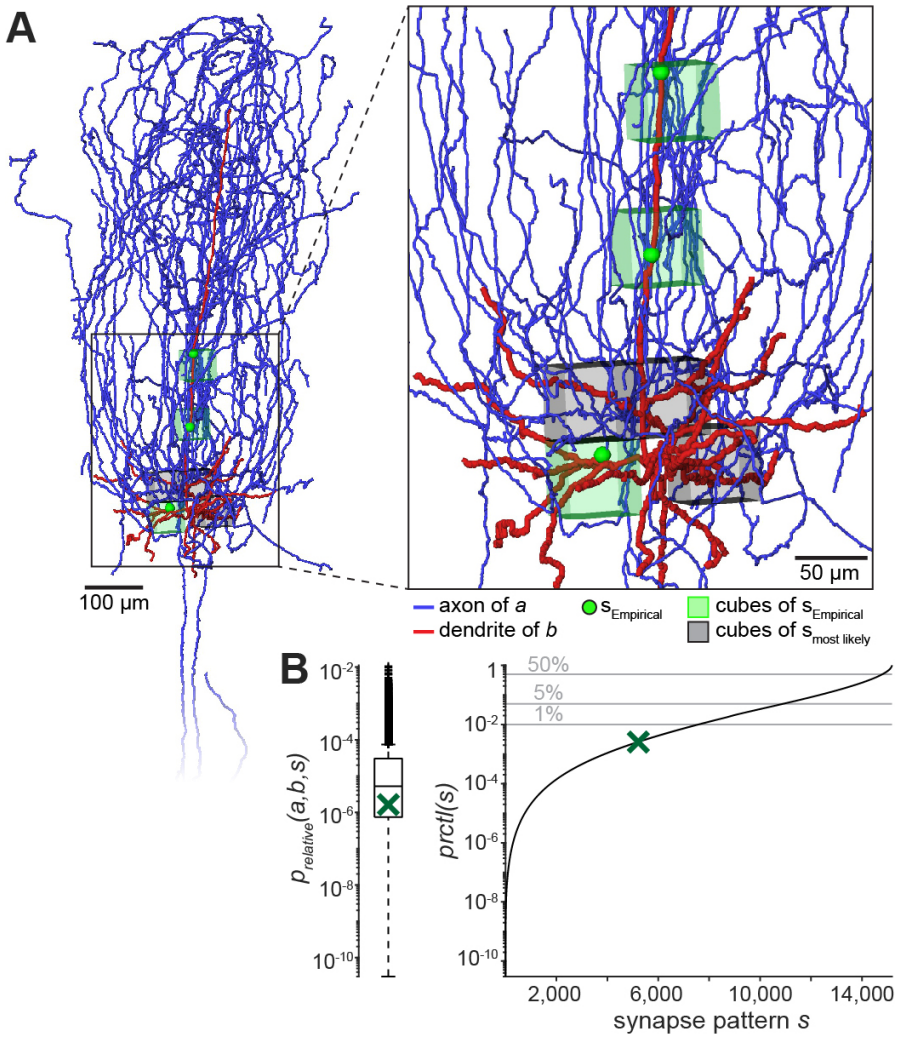

Figure 3.48. Example II: Empirical vs. predicted synapse pattern between *in vivo* labeled pair of L4sp. Same as in Figure 3.47 but for another exemplary pair. Here, $s_{Empirical}$, in terms of cube locations, does not match with $s_{most\ likely}$, the cubes predicted by the model to most likely contain three synapses (gray shaded boxes).

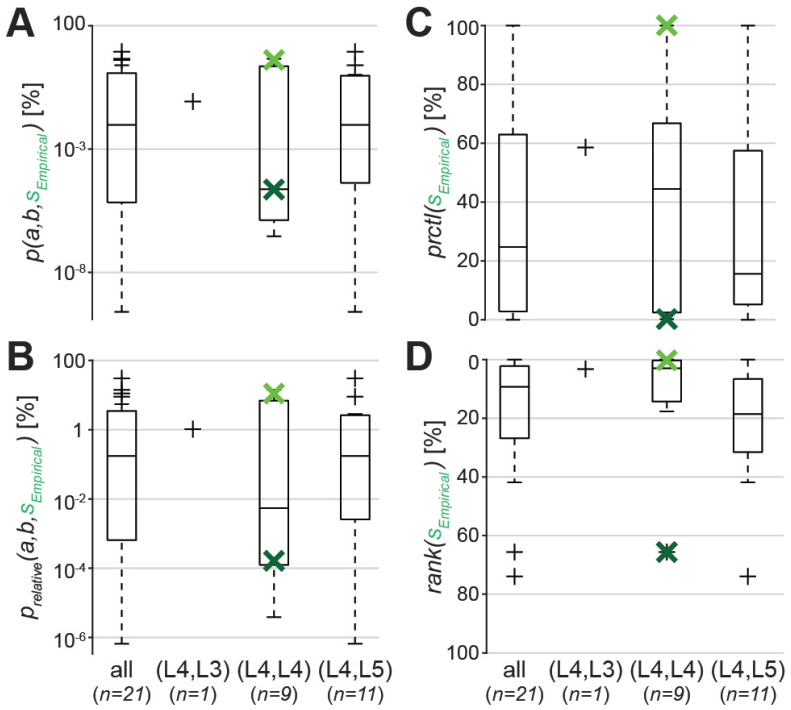

Figure 3.49. Empirical vs. predicted synapse patterns of *in vivo* labeled neuron pairs. (A) Absolute probability $p(a, b, s_{Empirical})$ of empirical synapse patterns in the model across all neuron pairs and layer-specific subgroups. Bright and dark green shaded crosses represent synapse patterns of Example I and II, respectively (see Figures 3.47 and 3.48). (B, C, D) same as in panel A but for relative probability $p_{relative}(a, b, s_{Empirical})$ (B), $prctl(s_{Empirical})$ (C), and $rank(s_{Empirical})$ (D; note the flipped y-axis). $p(a, b, s_{Empirical})$, $p_{relative}(a, b, s_{Empirical})$, and $prctl(s_{Empirical})$ values close to 100% and $rank(s_{Empirical})$ values close to 0% indicate consistency with model's predictions.

the model predicted three synapses to be most likely located in ($p(a, b, s_{Empirical}) = 2.3 \times 10^{-5}$ %, $p_{relative}(a, b, s_{Empirical}) = 1.6 \times 10^{-4}$ %, $prctl(s_{Empirical}) = 0.3\%$; see Figure 3.48). $s_{Empirical}$ is the 9,964th most likely synapse pattern of the 15,180 possible synapse patterns ($rank(s_{Empirical}) = 66\%$). While only one synapse is located on the basal dendrite and two on the apical trunk, the model predicts all three synapses to be most likely located in cubes innervated by the basal dendrites (see Figure 3.48).

The results of all neuron pairs ($n = 21$)[17] are shown in Figure 3.49.

[17] Two neuron pairs were excluded from the analysis; for details, see Methods Section 2.4.5.

The absolute probabilities of the empirical synapse patterns are generally low ($p(a, b, s_{Empirical}) = 1.12 \pm 2.21\%$, ranging from $2.51 \times 10^{-10}\,\%$ to $8.72\,\%$) indicating a low consistency with the model's predictions. Given the empirical number of synapses, the relative probabilities range from $6.62 \times 10^{-7}\,\%$ to $30.60\,\%$ ($p_{relative}(a, b, s_{Empirical}) = 3.69 \pm 7.41\%$). The majority of $prctl(s_{Empirical})$ values are larger than 5% ($n = 15$; $36 \pm 34\%$; ranging from $0.03\,\%$ to $100\,\%$). The empirical synapse patterns of six neuron pairs have $prctl(s_{Empirical})$ of less than 5%, and two of less than 1%, indicating that most of the empirical synapses are located in the cubes most likely to contain them. $rank(s_{Empirical})$ ranges from 0%, reflecting pairs in which $s_{Empirical}$ is located in the cubes predicted to most likely contain the respective number of synapses, to 74% ($17 \pm 21\%$). For half the neuron pairs their empirical synapses are located in the top 10% of most likely cube combinations. Only the synapse patterns of two neuron pairs are not among the top 42% of most likely cube combinations. Next, I assessed the impact of the empirical number of synapses on the model's predictions. The empirical $n(a, b)$ and $prctl(s_{Empirical})$ correlate negatively ($R = -0.48$, $p = 0.03$). This demonstrates that with increasing number of synapses between neuron pairs, the predicted cubes become less likely to contain the empirical synapses. This is due to the fact that the number of possible cube combinations $|\mathcal{S}|$ increases, resulting in a decrease of the relative probability of each combination. The relative rank of $s_{Empirical}$ is not significantly affected by the empirical number of synapses ($R = -0.18$, $p = 0.45$).

3.4 *BuildingBrains* web application

In collaboration with Philipp Harth, Ela Elsholz, Dr. Vincent J. Dercksen, and Prof. Dr. Hans-Christian Hege (Zuse Institute Berlin), we incorporated many tools to process and analyze reconstructed neuron morphologies into a web application called *BuildingBrains*[18] (see Figure 3.50). The respective tools were developed in the group of Dr. Marcel Oberlaender (for more details on the tools and functionality of *BuildingBrains*, see Methods Section 2.5). *BuildingBrains* allows uploading one's own neuron morphology, embedding the neuron morphology into the dense

[18] not publically available yet

structural model and predicting its distribution of impinging synapses based on the dense statistical connectome. In addition, *BuildingBrains* will serve as a repository for reconstructed neuron morphologies. Once made public, *BuildingBrains* will feature all *in vivo* labeled and reconstructed excitatory neuron morphologies ($n = 153$; Narayanan et al., 2017), the intracortical part of VPM axon morphologies ($n = 14$; Oberlaender, de Kock, et al., 2012; Oberlaender, Ramirez, & Bruno, 2012) and the inhibitory neurons of L1 ($n = 10$; Egger et al., 2015) that were used to generate the dense structural model. In addition, a more recent sample of eighteen L5PT neuron morphologies including their mapped subcortical target by Rojas-Piloni et al. (2017) will also be featured. All the neuron morphologies will be available for download. In the following, I describe an application example of *BuildingBrains*.

Application example: Constraining synaptic inputs onto neuron morphologies

When simulating neural activity of individual neuron morphologies, well-constraint synapse distributions are essential. Not only the location of each synapse along the dendrite of the neuron but also the origin of each synapse (i.e., cell type and location of presynaptic neuron) is desirable. Currently, empirical data providing cell type- and location-specific synapse distributions for an entire neocortical neuron morphology is not available (but see Schoonover et al., 2014). *BuildingBrains* provides one approach to tackle this challenge: A registered user can upload a reconstructed neuron morphology of rat vS1 that one wants to use for simulations. *BuildingBrains* embeds the registered version of the uploaded neuron morphology into the model of vS1 and predicts its synaptic input patterns based on the well-constraint and empirically validated dense statistical connectome (see Section 2.2; Figure 3.51). The location of each synapse along the dendrite of registered neuron morphology is predicted. In addition, the origin of each synapse is predicted, i.e., the cell type and soma location of the presynaptic neuron. The predicted synapse patterns can be downloaded and are tailored for network embedded simulations of the registered neuron morphology (e.g., in the NEURON simulation en-

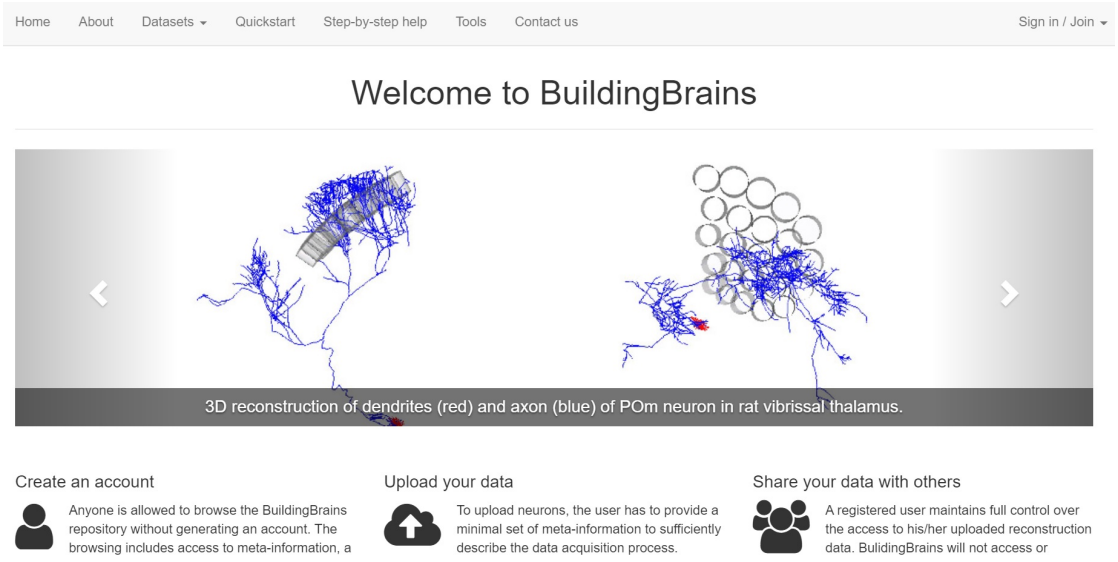

Figure 3.50. Cropped snapshot of the start page of the *BuildingBrains* web application. *BuildingBrains* applies a set of tools to a reconstructed neuron morphology, uploaded by the user: The uploaded neuron morphology is registered and placed into the geometrical model of vS1. Then *BuildingBrains* extracts morphological and topological parameters, automatically assigns the neuron to the most likely morphological cell type, and predicts synapse patterns onto the neuron morphology using the dense statistical connectome. Uploaded neuron morphologies and their structural information can be shared and explored by everyone, adding another feature to *BuildingBrains* as a repository of reconstructed neuron morphologies. Image taken from web application, accessed 16th January 2020.

vironment[19]). In combination with biophysical and functional data, the predicted synapse patterns were used in previous studies to simulate multi-compartmental neuron models embedded within the realistic model of vS1 (Egger et al., 2015, 2020).

3.5 *Cortex in Silico* web application

In collaboration with Philipp Harth, Dr. Vincent J. Dercksen, and Prof. Dr. Hans-Christian Hege (Zuse Institute Berlin), a web application of the dense structural model and the dense statistical connectome, called *Cortex in Silico*, was developed

[19] https://neuron.yale.edu/neuron/

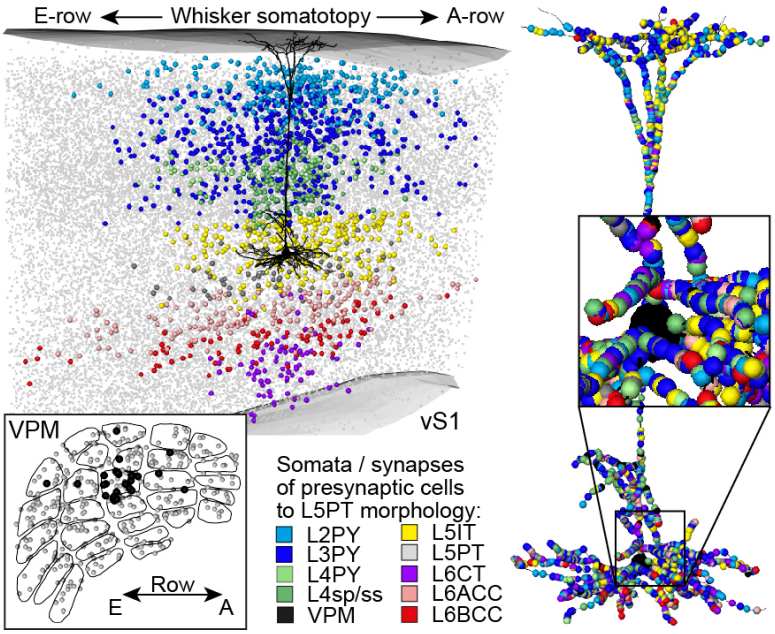

Figure 3.51. Example application of synaptic input mapper developed by Egger (2015) and available on the *BuildingBrains* web application. Reconstruction of *in vivo* labeled L5PT (black) embedded into the dense structural model of rat vS1. Gray markers represent all somata of excitatory neurons. Colored markers represent soma distribution of exemplary input neurons to the L5PT as predicted by the dense statistical connectome. Colored markers in right panels represent exemplary synapse distribution originating from the neurons shown in the left panels. Figure adapted from Egger et al. (2020).

(see Figure 3.52). The web application allows to replicate many of the results presented in this thesis (see Sections 3.1, 3.2, and 3.3) and in Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020). However, *Cortex in Silico* will not only serve as a resource of the models and their predictions, *Cortex in Silico* will also provide computational tools to explore and test principles of cortical structure and connectivity. In the following, I describe some application examples of *Cortex in Silico*. Note that the sections are in parts adapted from Udvary, Dercksen, Harth, Hege, and Oberlaender (2020) and that *Cortex in Silico* is an ongoing effort.

The different *in silico* experiments available in the web application allow mimicking different experimental scenarios. The subvolume experiment provides predictions for data accessible, for example, by saturated reconstruction in electron microscopic studies, the cellular experiment provides predictions for data acquired, for example, by *in vivo* or *in vitro* paired recordings, and the triplet motif experiment provides data acquired, for example, by triplet recordings. The in-degree correlation experiment provides predictions for data that is currently experimentally inaccessible at a large scale (i.e., numbers of synapses from two neuron populations onto a sample of individual neurons).

3.5.1 Application example I: Extracting structural information

Parts of the results of the dense structural model presented in Section 3.1, such as the layer-specific analysis of excitatory neurons or the subcellular analysis of the structural composition in small subvolumes, can be replicated in *Cortex in Silico*. To replicate the same data as reported in Figure 3.7, the user chooses the *in silico* subvolume experiment and applies the selection as shown in Figure 3.53A. The results of the *in silico* experiment are shown in a data table and visualized in multiple charts (see Figure 3.53B).

In addition, the structural estimates made accessible by *Cortex in Silico* are beneficial in multiple experimental scenarios: During blind recording or stimulation experiments only the location of the electrode is known, while knowledge about the structural composition of the targeted area is desired. *Cortex in Silico* allows the user to probe any targeted area of vS1 *in silico*, providing estimates of

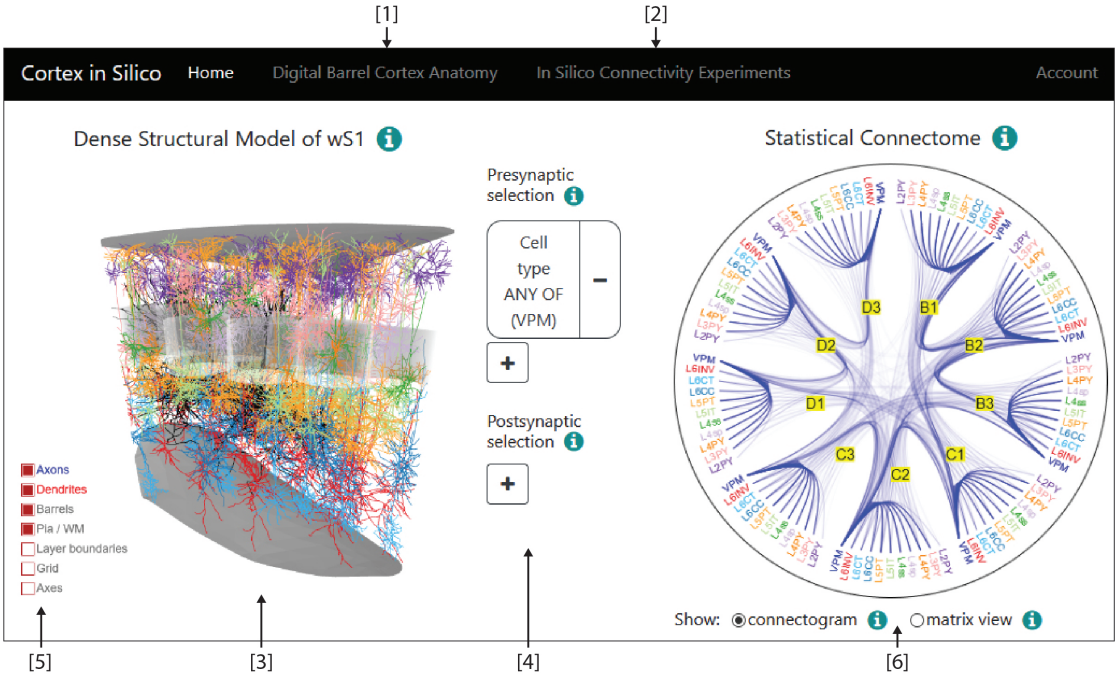

Figure 3.52. Start page of *Cortex in Silico* web application. [1,2] Tabs for the two main sections: [1] "Digital Barrel Cortex Anatomy" for exploring the dense structural model; [2] "In Silico Connectivity Experiments" for performing connectivity experiments *in silico*. [3] Interactive 3D visualization of dense structural model of vS1. Subset of axons and dendrites that match the user-defined selection of pre- and postsynaptic neurons [4] are shown, color-coded by cell type. Surfaces of pia, WM, and barrels rendered in gray. Here, the user restricted one's presynaptic selection to neurons located in VPM thalamus, while postsynaptic neurons were not restricted. [5] User-defined options for visualization of dense structural model. [6] Dense statistical connectome visualized as a "connectogram" for the user-defined selection of pre- and postsynaptic neurons [4] located in C2 barrel column and its eight surrounding barrel columns. User can change the visualization to a matrix representation. Image taken from web application, accessed 16th January 2020.

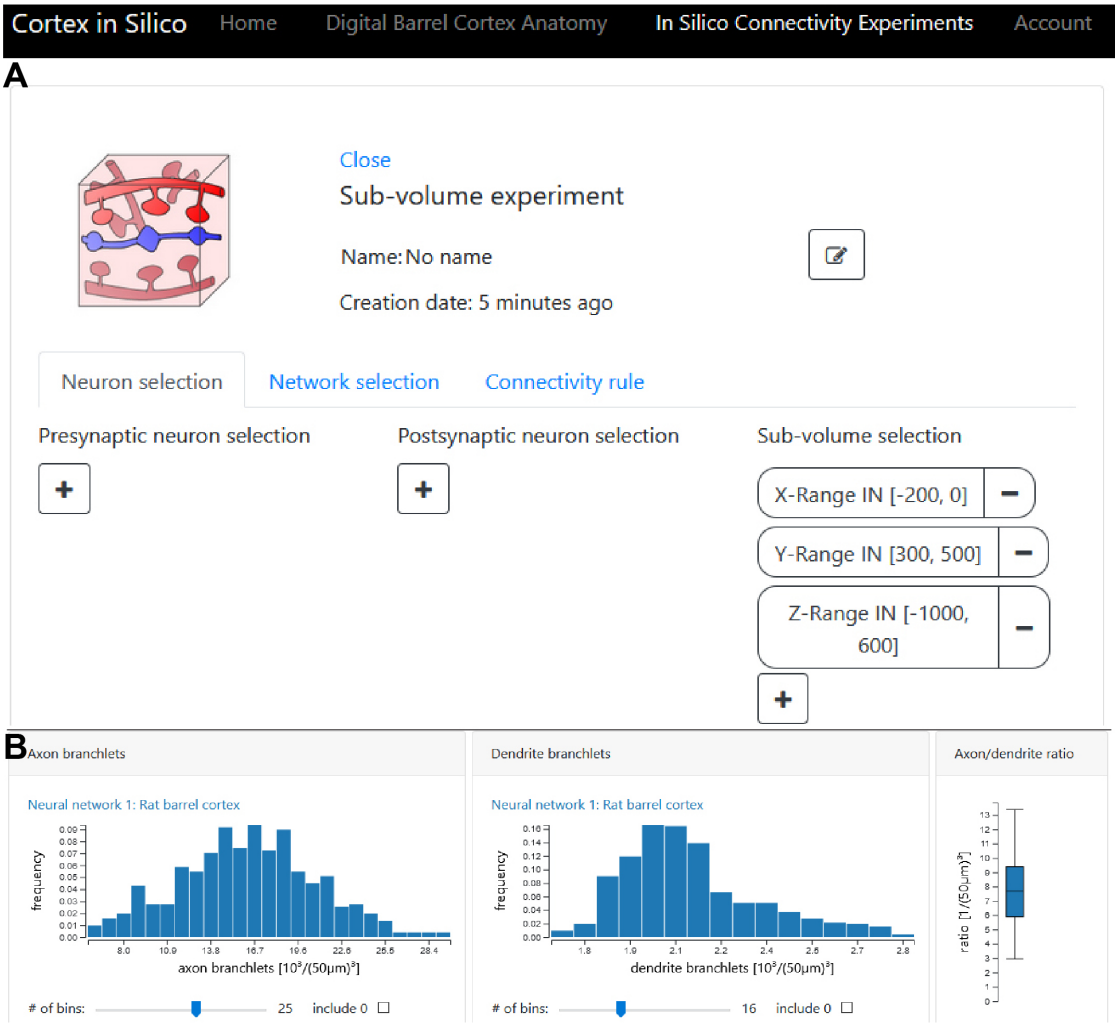

Figure 3.53. Cortex in Silico: Subvolume experiment. (A) Neuron selection to replicate parts of the results reported in Section 3.1. (B) Replication of Figure 3.7 in *Cortex in Silico*. Frequency of axon and dendrite branchlets (left, center) and the ratio between axon and dendrite branchlets per 50 µm cube (right) in the selected subvolume. Image taken from web application, accessed 16th January 2020.

its structural composition with respect to neuron distributions, cell types, axons, dendrites, boutons, and spines. The cell type-specific 3D distribution of boutons, for example, can help identify which regions stimulated neurons might target and thereby guide stimulation experiments. In contrast to blind recording or stimulation experiments, saturated reconstructions of electron microscopic data provide experimental access to the structural composition of small and truncated subvolumes of cortical tissue. Due to the truncation, the cell type or location of somata of innervating axons and dendrites are difficult to determine. The *in silico* subvolume experiment available in *Cortex in Silico* provides estimates of the cell types and soma locations of neurons that innervate any user-specified subvolume of vS1, helping to identify cellular identities of axon and dendrite branchlets.

3.5.2 Application example II: Extracting predicted connectivity

Many of the results of the dense statistical connectome presented in Sections 3.2 and 3.3, such as comparing empirical and predicted connection probabilities or triplet motif spectra, can be replicated in *Cortex in Silico*. To replicate one of the comparisons presented in Figure 3.33A (center panel; *in vitro* paired recording by Sun et al., 2006), the user chooses the *in silico* cellular experiment, emulates a slice, and applies the selection as reported in Table A10 and shown in Figure 3.54AB. In contrast to the results presented in Section 3.3.2 and Table A10, where the predicted connection probabilities were aggregated over all ten *in silico* slices, *Cortex in Silico* currently allows probing only one *in silico* slice at a time. The results of the *in silico* experiment are shown in the web application as data table and visualized as two charts (see Figure 3.54C).

In addition to the results and analysis presented in this thesis, *Cortex in Silico* can be valuable in the following scenarios:

Constraining networks for simulating neural activity: Simulations of neural activity rely on well-constraint connected networks of neurons. Currently, however, not one empirically determined network for an entire neocortical area

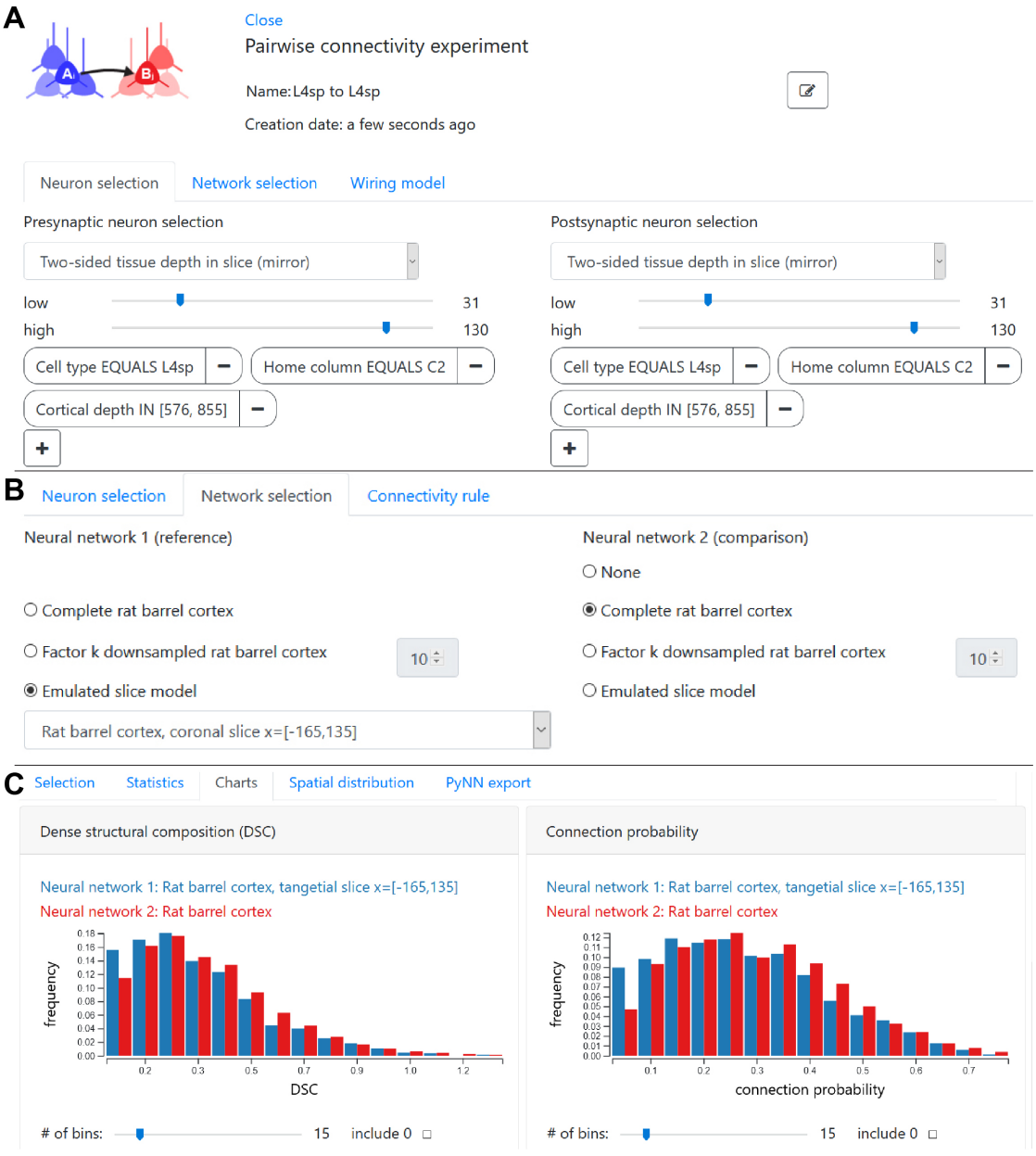

Figure 3.54. Cortex in Silico: Cellular experiment to replicate empirical connection probability between pairs of L4sp *in vitro* reported by Sun et al. (2006). (A) Neuron selection for replication. (B) Network selection for replication: Network 1 emulates a slice experiment (left), network 2 (right) emulates an *in vivo* experiment. *Caption continued on next page.*

Figure 3.54 — continued from previous page. (C) Results of cellular experiment: Histograms of predicted DSC(A, B) (left) and $P(A, B)$ (right), *in vitro* prediction in blue, *in vivo* prediction in red. Sun et al. (2006) reported a measured connection probability between pairs of L4sp of 24%. Images taken from web application, accessed 3rd July 2020.

is available. *Cortex in Silico* overcomes this challenge by providing access to the predictions of the validated dense statistical connectome (see Section 3.3.2) or any other user-specified synapse formation strategy. Specifically, the user can access and download predicted pair-wise connection probabilities and synapse distributions[20]. Furthermore, predicted networks are available in a standardized format suitable for `PyNN` (Davison et al., 2009). These predictions can be used to constrain connectivity in networks as well as synapse distributions onto individual neurons (e.g., for simulating activity in multi-compartment models of neurons) as demonstrated by Egger et al. (2015); Landau, Egger, Dercksen, Oberlaender, and Sompolinsky (2016); Egger et al. (2020).

Assessing the impact of slicing on connectivity: The majority of reported connection probabilities are inferred from functional paired recordings in tissue slices. Since slicing impacts connectivity in an unsystematic way (as shown in Section 3.2.1), inferring *in vivo* connection probabilities from *in vitro* ones is challenging and depends on many factors. By replicating the *in vitro* experiment in *Cortex in Silico* and comparing the empirical and/or predicted *in vitro* connectivity to the predicted *in vivo* connectivity, *Cortex in Silico* provides an estimate of the impact of slicing on connectivity. One example is illustrated in Figures 3.54.

Assessing the contribution of cell types to connectivity: In many experiments the morphology and thus the cell type of neurons involved in the empirically determined connectivity is unknown. For example, in most paired or multi-electrode experiments morphologies of recorded neurons are not retrieved. Neurons are only identified by the laminar position of their somata, resulting in layer-specific connectivity. In reconstructions of electron microscopic data the cell

[20] see also other web application *BuildingBrains* described in Section 3.4

type and location of somata of synaptically connected axon and dendrite branchlets are unknown. By replicating these experiments in *Cortex in Silico*, the contribution of different neuron populations to the empirically determined connectivity can be probed.

Assessing functional recordings: Functional studies investigating electrophysiological responses of neurons revealed correlated and heterogeneous responses (e.g., Cohen-Kashi Malina, Mohar, Rappaport, & Lampl, 2016). These responses might be the result of specific network topologies. By performing in-degree correlation or triplet motif experiments in *Cortex in Silico*, the underlying network structure can be probed. For example, correlated responses between neurons might arise from correlated in-degrees.

3.5.3 Application example III: Investigating principles of wiring

All the connectivity results presented so far were the result of the synapse formation strategy formulated in Section 2.2 (i.e., see Equations 2.2 and 2.3). *Cortex in Silico* offers the possibility to change the synapse formation strategy and thereby to investigate principles of wiring. Different strategies of synapse formation can be compared with each other and to empirical data. In Figures 3.55 and 3.56, two arbitrary synapse formation strategies (blue) were formulated and compared to the connectivity predicted by the strategy formulated in Section 2.2 (i.e., see Equations 2.2 and 2.3).

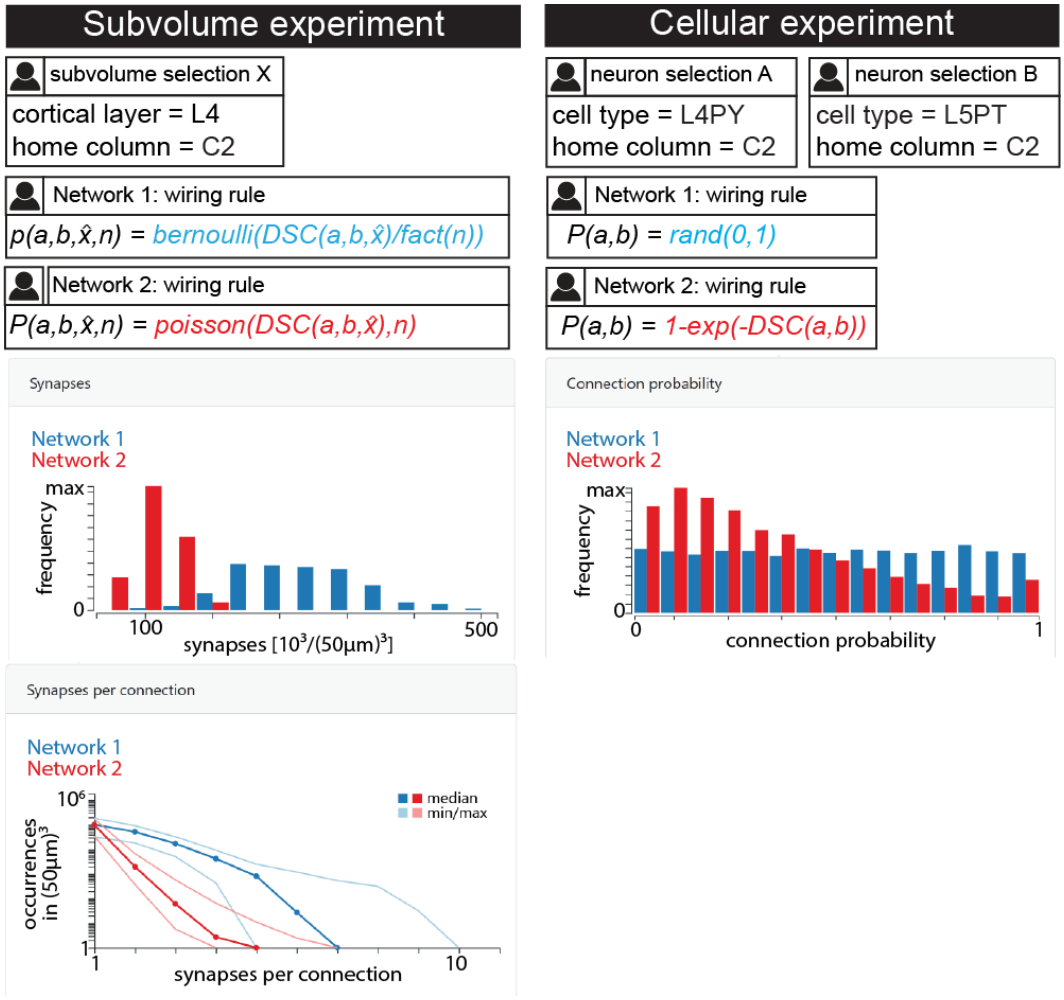

Figure 3.55. Cortex in Silico: Changing the synapse formation strategy. Left: Comparing two synapse formation strategies for an *in silico* subvolume experiment in L4 of the C2 barrel column. Strategy following Poisson distribution in red and Bernoulli distribution in blue. Charts show number of synapses per 50 µm cube and the number of synapses per connection. Right: Connection probabilities in *in silico* cellular experiment between L4PY and L5PT in the C2 barrel column for two different synapse formation strategies. Chart shows the histogram of the connection probabilities. Images modified from web application, accessed 16th January 2020.

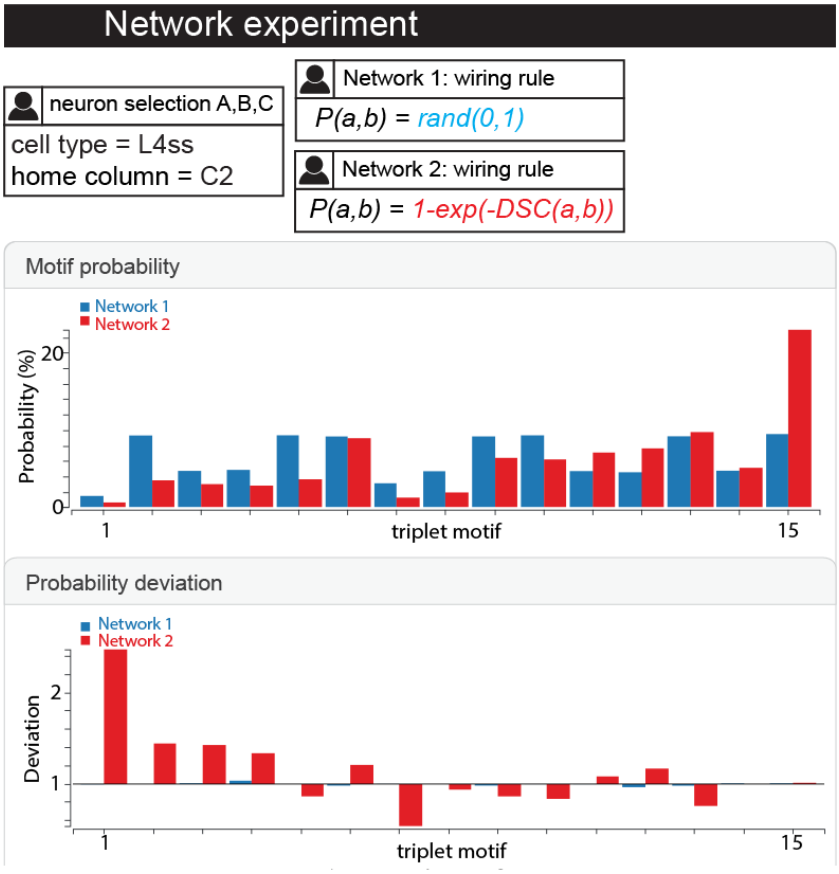

Figure 3.56. Cortex in Silico: Changing the synapse formation strategy in a network experiment. Comparison between two synapse formation strategies with respect to triplet motif probabilities and deviations. Synapse formation strategy in network 1 (blue) mimics a random network. Network 2 uses the strategy presented in this thesis. Images modified from web application, accessed 16th January 2020.

4 | Discussion

In this thesis, I presented an anatomically detailed, reverse engineered model of the characteristic structural composition of the neuropil of an entire neocortical area. The model provided realistic and robust estimates of the structural properties of neuropil preserved across animals. I used the model to predict how wiring diagrams of the modeled area would look like if they were solely shaped by the estimated distribution of neurons and neurites, in the absence of recognition-dependent strategies of synapse formation. Although only shaped by the structural composition of the neuropil, the predicted networks are sparse, heterogeneous, correlated, and structured unlike random networks. A systematic comparison with empirical wiring properties revealed a high degree of consistency with the model's predictions. Hence, many wiring properties could emerge implicitly through the constraints imposed by the outcome of neurite and neuron development. The entire model was incorporated into the web application *Cortex in Silico*, allowing the scientific community and public to interact with the model, access its predictions, and even modify the strategy of synapse formation.

In the following, I first discuss the structural model and its results with respect to the goals stated in the Introduction (see Section 1.4). I then briefly relate the model to other approaches and discuss the model's limitations. After that, I discuss the following questions before closing with a conclusion and outlook:

- How can complex networks emerge from the constraints imposed by the structure of the underlying neuropil?
- What do wiring properties represent that are consistent with the model's predictions?

- What do wiring properties represent that are inconsistent with the model's predictions?
- What about wiring properties involving inhibitory neurons?

4.1 A digital model of the structural composition of the neuropil

To explore the impact of the structural composition of the neuropil on wiring, I first required access to the structural composition of the neuropil preserved across animals. In an ideal scenario, this data could be provided by multiple dense reconstructions of the neuropil using an electron microscopic approach and identifying the structural properties preserved across animals. However, state-of-the-art dense electron microscopic reconstructions are limited to small volumes (e.g., Motta et al., 2019). Due to these current shortcomings, I used an empirically constraint model of rat vS1 (Egger et al., 2014). Based on the interanimal variability of cytoarchitectonic features (see Figure 3.2), the model's resolution limit is defined by cubes of 50 µm edge length. In the context of this thesis, the model should (i) capture the structural principles characteristic for the respective neocortical area and provide (ii) robust and (iii) realistic estimates of the structural composition of the neuropil that then can be used to explore how they constrain wiring properties. In the following, I focus on the model with respect to excitatory neurons.

Does the model capture the structural principles characteristic for rat vS1? The fact that the model is constraint by empirical data obtained from rats of the same strain and developmental stage and housed in standard laboratory environments (Egger et al., 2012; Meyer et al., 2013; Oberlaender, de Kock, et al., 2012; Narayanan et al., 2015) in combination with the remarkable low interanimal variability of these empirical constraints (see Figures 3.2 and A1) allows it to capture structural principles of rat vS1 preserved across rats. For example, the model captures the average of area-specific geometrical properties (e.g., curvature of the pial surface, whisker-specific locations, volumes, and orientations of each barrel; Egger et al., 2012) and the average 3D distributions of neuron somata

including area-specific properties (e.g., differences in the number of neurons across cortical depth and barrel columns; Meyer et al., 2013).

The sample of *in vivo* labeled excitatory neuron morphologies (Oberlaender, de Kock, et al., 2012; Narayanan et al., 2015) can be considered representative since it comprises about 1% of all excitatory neurons of one barrel column including multiple morphologies for each of the major excitatory cell types (Harris & Shepherd, 2015). For the most complex dendritic cell type, the L5PT, a second set of eighteen reconstructed L5PT dendrite morphologies had been obtained more recently (Rojas-Piloni et al., 2017). In terms of morphological and topological features, this set was indistinguishable from the sixteen L5PT dendrites used within this thesis (see Figure 3B of Rojas-Piloni et al., 2017) indicating that the sample of L5PT dendrites is not only representative with respect to their average structure but also their biological variability. The assessment of how representative the morphologies for each cell type are revealed that the current sample captures nearly all the variability of the cell type-specific dendrite distributions and much of the variability of the cell type-specific axon distributions (see Figure 3.3). However, the variability of two cell types, L4PY and $L6_BCC$, was considerably higher compared to the other cell types. The reason for this is that both cell types are each only represented by four axon morphologies, the lowest number of cell type-specific morphologies in the sample. Both cell types are also the least abundant ones in the vS1 model (see Table 3.1), representing only about ∼8% of all neurons. $L6_BCC$ further comprise highly diverse dendrite and axon morphologies (Narayanan et al., 2015), that are often grouped together with $L6_ACC$ (e.g., Z.-W. Zhang & Deschênes, 1997; Harris & Shepherd, 2015). The lack of captured variability by these two cell types might, therefore, have a negligible impact on the overall area- and cell type-specific organizational principles of axons and dendrites.

Does the model provide robust estimates of the structural composition of rat vS1? The model is populated by upscaling the sparse but representative sample of morphologies to the entire neuron population. I find that the upscaling captures more than 90% of the neurite density and diversity within the resolution limit of the model that are preserved across animals (see Figure 3.5). This result indicates that the model provides a robust estimate of neuron packing densities

and diversities. What has not been assessed so far is how robust these estimates are with respect to each cell type and the sample of inhibitory neurons.

Does the model provide realistic estimates of the structural composition of rat vS1? I tested whether the model provides realistic estimates of the structural composition of vS1 within its resolution limit. The estimated ratio of axon and dendrite branches and bouton densities are in line with the few instances of available empirical data on the electron microscopic level (see Figure 3.7; Kasthuri et al., 2015; Motta et al., 2019; Santuy et al., 2018). These measurements were not used to constrain the model. More empirical data to validate the structural composition at the electron microscopic level is currently not available. Given that the model is empirically well constrained and validated on the subcellular scale, I argue that the model's estimates provide a realistic characterization of the excitatory neuron and neurite distributions within vS1 expected across animals (but see Section 4.3 for model limitations).

4.1.1 Exploring principles of cortical organization using the structural model

Beyond providing constraints for synapse formation, the estimated structural composition of the model presents a unique opportunity in various scenarios. Within this thesis, I utilized the structural model to investigate organizational principles of excitatory and inhibitory neurons. The analysis of excitatory neurons that was published as a review article (Narayanan et al., 2017) revealed that each cortical layer comprises somata of multiple cell types, demonstrating that layers cannot be considered proxies for cell types (see Figure 3.8). Therefore, when grouping neurons by layer one needs to be aware that functionally and morphologically distinct cell types are pooled together, potentially complicating the interpretation of the results. The dense structural model can hence act as a guide to help interpret layer-specific measurements or manipulations with respect to the underlying cell types. Nonetheless, laminar landmarks, such as the center or borders of a layer, are related to the cell type-specific dendrite and axon projection patterns (see Figure 3.9) and therefore potentially act as structural references to guide neurites during

development (Shimogori & Grove, 2005; Sanes & Yamagata, 2009; M. E. Williams, de Wit, & Ghosh, 2010).

Like excitatory neurons whose neurites are organized with respect to their soma densities and thus layers, neurites of inhibitory neurons seem to be organized with respect to the inhibitory soma densities, specifically to two zones of high inhibitory neuron density located between layers 2 and 3 and layers 4 and 5 (see Figure 3.12; Meyer et al., 2011). This can be mainly attributed to the fact that inhibitory neurons generally have more localized axons and dendrites whose projection patterns follow their overall soma distribution. For a detailed discussion about inhibitory neurons, please refer to Section 4.3 and 4.7. In addition, the structural estimates can be used to advance our understandings of functional phenomena, such as the fast and broad cortical output of L5PTs (Egger et al., 2020). In Egger et al. (2020), the structural model did not only provide constraints for simulating sensory-evoked signal flow (see Figure 3.51) but also allowed estimating the temporal delay of an action potential traveling along VPM axons before reaching different neuron populations (see Figure 3.10). The estimated temporal delay based on path length differences could explain the differences in response latencies between the populations.

4.2 Relationship to other approaches

In contrast to previous studies that reported large-scale network models (e.g., Blue Brain Project; Markram et al., 2015; Reimann et al., 2019 or by the Allen Brain Institute; Billeh et al., 2020), predicted wiring patterns in the model are not constrained by connectivity measurements. Instead, wiring patterns represent probabilistic predictions of the model based on the herein formulated synapse formation strategy. The model, however, is not limited to the synapse formation strategy. In *Cortex in Silico*, the web application of the model, the user can specify different synapse formation strategies and compare their impact on network architecture. Other features that taken together make the model unique are (i) its strong reliance on assessed empirical data than on assumptions, (ii) its volume, i.e., the model covers an entire neocortical area, (iii) its ability to capture area-specific geometrical, cellular, and morphological details that are observed across

animals, and (iv) the range of its predictions across scales from pair-wise wiring properties between neuron branches within 50 µm cubes to network motifs between neurons across entire vS1. Other current large-scale models miss many area-specific properties due to their reliance on *in vitro* reconstructions or on assumptions such as uniformly distributed somata or stereotypical columns (e.g., Markram, 2006; Markram et al., 2015). A detailed discussion of the model compared to other approaches can be found in Egger et al. (2014).

Studies investigating how structure impacts wiring: Within this thesis, I presented the first systematic analysis of the impact of the structural composition of the neuropil on wiring properties of an entire neocortical area. Unlike other approaches that assessed the impact of pairs of neuron morphologies on their wiring in isolation (e.g., Brown & Hestrin, 2009; Petreanu et al., 2009) (for review, see Rees et al., 2017), the presented approach considers synaptic structures of all neurons present in the same overlapping subvolumes. For example, Brown and Hestrin (2009) argued that observed differences in the connection probabilities between different L5 populations cannot fully be accounted for by their sparse overlap and instead depend on their long-range target. Petreanu et al. (2009) argued that overlap cannot explain subcellular input patterns to neocortical neurons (e.g., see Figure 3.38A). Here, I demonstrated that these two observed phenomena — as well as many others besides them — can, however, emerge from the dense structural composition of the neuropil. In line with Brown and Hestrin (2009), I further demonstrated that differences in sparse axodendritic overlap also do not fully explain differences in the predicted connection probabilities as derived from the structural composition of the neuropil (see Figure 3.18).

4.3 Limitations

Many of the model's limitations do not represent limitations of the model itself but limitations in data availability. Once more data (e.g., morphological reconstructions of particular neuron populations) is available, these components can be integrated into the model overcoming the previous limitations. However, not only the model comes with limitations but so do the empirical wiring properties used to

assess the model. In the last part of this section, I discuss the approach-depending limitations of empirical wiring properties derived from experimental data.

4.3.1 Limitations of the model with respect to inhibitory neurons

The characterization and integration of the diverse inhibitory neurons into the model is an ongoing effort in the group of Dr. Marcel Oberlaender. So far, only the population of inhibitory neurons in L1 is as well characterized, constraint, and integrated as the excitatory neurons (Egger et al., 2015). Unlike for the excitatory neurons, a quantitative assessment of how representative and robust the sample of inhibitory neuron morphologies is has yet to be performed. Although the sample spans the entire cortical sheet, represents nearly 10% of all inhibitory neurons of one barrel column, and the morphologies resemble most previously suggested inhibitory cell types (Udvary, 2014), the sample and its integration in the model comes with limitations.

First, while the samples of L2-6 inhibitory morphologies were acquired under similar experimental conditions with only minor differences, such as slicing thickness, the morphologies were labeled *in vitro* (Koelbl et al., 2013; Helmstaedter et al., 2009a; Arzt et al., 2018). To compensate for the truncation of axons and dendrites, the morphologies were "curated" assuming radially symmetry of the neuron arbors in the horizontal (i.e., tangential) plane. However, inhibitory axons in vS1 can have asymmetric horizontal projection patterns, that in addition, project beyond the dimensions of one barrel column (e.g., Slabik, 2018). Therefore, the assumptions that most of the axon organization is preserved within a slice or that the neuron's arbor is radial symmetry cannot be justified for all inhibitory neurons in vS1.

Second, the sample of inhibitory neurons in L4 integrated into the model is exclusively fast-spiking (Koelbl et al., 2013). Non-fast-spiking inhibitory neurons that constitute about 20 % to 40 % of inhibitory neurons in L4 of which some have distinct morphologies (Emmenegger et al., 2018) are currently not represented in the model.

Third, the sample of inhibitory neurons was not systematically probed with re-

spect to their molecular identities. Hence, inhibitory neurons in the model cannot be grouped by their molecular identities, and thus, a direct comparison to empirical wiring properties of molecularly identified inhibitory neurons is currently not possible.

Fourth, the lack of data available with respect to pre- and postsynaptic densities of inhibitory synapses hampers constraining those densities in the model. Postsynaptic densities for inhibitory synapses are assumed to be proportional to the surface area of the targeted neuron, a structural property not available for all neurons in the sample of morphologies. The surface areas were therefore estimated based on a sample of neurons with reconstructed surfaces (for details, see Egger, 2015). The assumption that bouton densities along inhibitory axons of L2-6 neurons are constant across the entire cortical depth is at odds with reports of cell type-specific inhibitory bouton densities (Markram et al., 2004). Cell type-specific bouton densities might explain part of the inconsistency between estimated and measured inhibitory synapse densities across cortical depth (see Figure 3.14).

4.3.2 Other limitations and caveats of the model

Long-range inputs: Currently, long-range inputs to the model of rat vS1 are limited to inputs from VPM thalamus. Other long-range inputs, for example those from the posterior medial nucleus of the thalamus, motor cortex, or secondary somatosensory cortex (see C. C. H. Petersen, 2019 for an overview), are currently lacking but can be incorporated in the future. Based on the finding that the empirical distribution of synapse densities matches the estimated bouton densities (see Figure 3.7), the missing long-range inputs represent, however, only a small fraction of the overall synapses present in rat vS1.

Long-range targets: The vast majority of long-range targets of the neurons in the model are unknown. Hence, it is currently not possible to test how long-range targets relate to their predicted intracortical wiring patterns (for example as reported by Brown & Hestrin, 2009) unless neurons with different long-range targets can be unambiguously differentiated in the model (e.g., L5ITs vs. L5PTs). The group of Dr. Marcel Oberlaender has already obtained a sample of L5PT dendrites

with mapped long-range subcortical targets (Rojas-Piloni et al., 2017). However, their intracortical axon projection patterns were not recovered. Therefore, it remains to be seen whether and how predicted wiring patterns vary depending on their long-range subcortical target. Slight structural differences in the distribution of L5PT dendrites and soma depth (Rojas-Piloni et al., 2017) suggest the possibility that at least with respect to the herein formulated synapse formation strategy synaptic inputs might vary differently depending on their subcortical target.

Representation of axon morphologies: Although axon morphologies in the vS1 model are duplicated, they are not transformed to their assigned soma position[1]. This procedure, while preserving specific projection patterns and orientations, carries caveats: First, effects related to the soma position are not necessarily fully captured by the model. In cases where I investigated soma position-related effects (e.g., connection probability vs. intersomatic distance), the original soma position of the axon was used, hence making the presynaptic component sparse. Second, the model does not capture any dependencies between individual axons and dendrites within cell types of a single barrel column (e.g., the larger the dendrite, the larger the axon). As a consequence, for example, cell type-specific in- and out-degrees are uncorrelated and doublet motifs are as frequent as in a random network, conflicting with empirically determined frequencies of doublet motifs (e.g., Song et al., 2005) (see Figure 3.44). Whether these effects remain when using the original soma positions and axodendritic morphologies, needs to be investigated in the future.

Wiring statistics, individual wiring patterns, and underlying synapse formation mechanisms: Since the model provides constraints for synapse formation that are preserved from animal to animal within a certain resolution limit, the model's strength is in predicting wiring statistics that are presumably preserved from animal to animal as well. Comparing the model's predictions to individual wiring patterns between individual neuron pairs (see Sections 2.4.5 and 3.3.5), therefore, has to be interpreted with huge caution. On the subcellular

[1] Axon morphologies are randomly assigned to a soma position with a transformed dendrite of the same cell type and barrel column (for details, see Methods Section 2.1.1)

scale, for example, the positioning of the 50 µm cubes might be crucial and explain parts of the failure to predict 3D putative synaptic distributions (see Figure 3.49). It is further important to point out that the model cannot provide any information about the underlying mechanisms that have shaped a particular neocortical synapse. Such an endeavor might be anyway extremely challenging due to the existence of many convoluted, potentially codependent biological and cellular mechanisms (Sanes & Yamagata, 2009). A recent preprint showed that even in a tractable model system, such as the *Drosophilia*, multiple mechanisms underlie circuit assembly in the nerve cord (Valdes-Aleman et al., 2019).

Animal model and age range: The empirical data used to build the model of vS1 as well as most of the measurements used for assessing the model's predictions come from relatively young animals[2] that although having passed the critical periods of vS1 development (Erzurumlu & Gaspar, 2012) have not reached sexual maturity (Sengupta, 2013). Whereas most evidence indicates that both structure and wiring statistics, such as connection probabilities (e.g., Fino & Yuste, 2011; Packer et al., 2013) remain stable throughout the remaining animal's life, two recent studies argue that connection probabilities can be different in sensory areas of animals older than 60 days (Jiang et al., 2015; Scala et al., 2019). While further investigations are necessary, it is important to acknowledge that the model and its predictions capture only the result of development, specifically of young Wistar albino rats raised in laboratory environments[3]. Multiple models built for rats of different developmental stages and reared conditions might help in elucidating the impact of development, aging, and even environment on the structure of the neuropil and its predicted wiring properties.

4.3.3 Limitations of empirical wiring properties

Deriving wiring properties from extremely challenging and labor extensive experiments comes with different approach-depending limitations. Most empirical

[2] postnatal day 25-45 for most empirical constraints; as young as 14 days for some connectivity measurements, e.g., Perin et al., 2011; C. C. Petersen & Sakmann, 2000.

[3] e.g., cellular composition within sensory areas varies between laboratory and wild rats (Krubitzer, Campi, & Cooke, 2011)

wiring properties reflect partial, in few cases, even distorted information about the actual underlying wiring diagram (discussed in, e.g., Thomson & Lamy, 2007; S. R. Williams & Mitchell, 2008; Stepanyants, Martinez, Ferecskó, & Kisvárday, 2009; Barth et al., 2016). The three most crucial limitations are that most connectivity statistics are inferred in (i) slice preparations using (ii) functional recordings and (iii) often without recovering the sampled neuron morphologies (e.g., Thomson et al., 2002; Lefort et al., 2009; Perin et al., 2011; Jiang et al., 2015). So far, only a few studies have combined functional recordings with anatomical mappings of synaptic contacts (e.g., Markram et al., 1997; Feldmeyer et al., 1999, 2006) or inferred connectivity from functional recordings *in vivo* with post hoc neuron reconstructions (e.g., Bruno & Sakmann, 2006; Constantinople & Bruno, 2013).

As demonstrated within this thesis, pair-wise connectivity statistics are unsystematically distorted by slice preparations (see Figure 3.20), a finding in line with previous reports (e.g., Stepanyants et al., 2009). This finding hence indicates that approaches that use simple scaling factors to compensate for slicing (e.g., Jiang et al., 2016) are generally not justified.

Functional recordings come with additional limitations, most notably space clamp errors: Connections established at the distal parts of the dendrite or by weak synapses evoke postsynaptic potentials that when reaching the soma might not be detected anymore by the recording electrode and therefore missed (S. R. Williams & Mitchell, 2008). Another limitation of many functional recordings is that the morphologies of most neurons are not recovered. Hence, different, spatially intermingling morphological cell types are pooled together. As shown within this thesis, wiring statistics are affected by how neuron populations are sampled (e.g., see Figure 3.16). Extrapolating such results to other neuron populations, cortical areas, or larger volumes, should therefore be examined critically (e.g., Reimann et al., 2019).

When comparing the empirical data to the model, I took these experimental limitations into account by (i) mimicking the truncation introduced by slicing, (ii) by ignoring predicted connections established at the apical dendrite (see Figure 3.36), and by (iii) sampling the neurons in a similar manner as in the experiment. However, these limitations, together with the fact that wiring patterns are heterogeneous, have to be kept in mind when evaluating the model's predictions.

4.4 How can complex networks emerge from the constraints imposed by the structure of the underlying neuropil?

The findings presented in this thesis demonstrate that the dense structural composition of vS1 imposes structural constraints for synapse formation that by itself can result in network architectures defined by sparse, skewed, and heterogeneous wiring (see Figure 3.16), correlated connectivity statistics (see Figure 3.25), and motif spectra that are unlike those in random networks (see Figure 3.26). Similar network architectures can also arise from learning mechanisms (e.g., D. Zhang et al., 2019). Using a mathematical model developed in collaboration with Prof. Dr. Jakob Macke (University Tübingen) (see Figures 3.31 and 4.1), any sparse, heterogeneous, and correlated network results in the emergence of network motifs whose frequencies differ from those expected in a random network. For example, recurrent connections are more frequent than in a random network. These properties, of which all except for correlations have been anatomically observed in the neocortex, can hence be related to structural properties of the neuropil as estimated in the model (see Figure 4.1).

Sparsity: In the model, sparse connectivity between neurons (i.e., connection probability $< 50\%$) arises from the sparse connectivity at the subcellular level, where the vast majority of close-by branchlets remain unconnected ($> 98\%$; see Figure 3.32). This sparse wiring, in turn, is the result of the structural constraints found in the densely packed neuropil: One 50 µm cube is densely packed with on average approximately 0.5 m of axon and 7 cm of dendrite path length (see Figure 3.6). The denser the neuropil, the more pre- and postsynaptic structures are available, and thus, the larger the combinatorial space of possible connections between pre- and postsynaptic structures becomes. Since the model assumes that each pre- and postsynaptic structure forms only one connection, sparse wiring is the result. In a scenario of a less densely packed neuropil (i.e., less available synaptic structures) or in a scenario where multiple synapses are formed with a

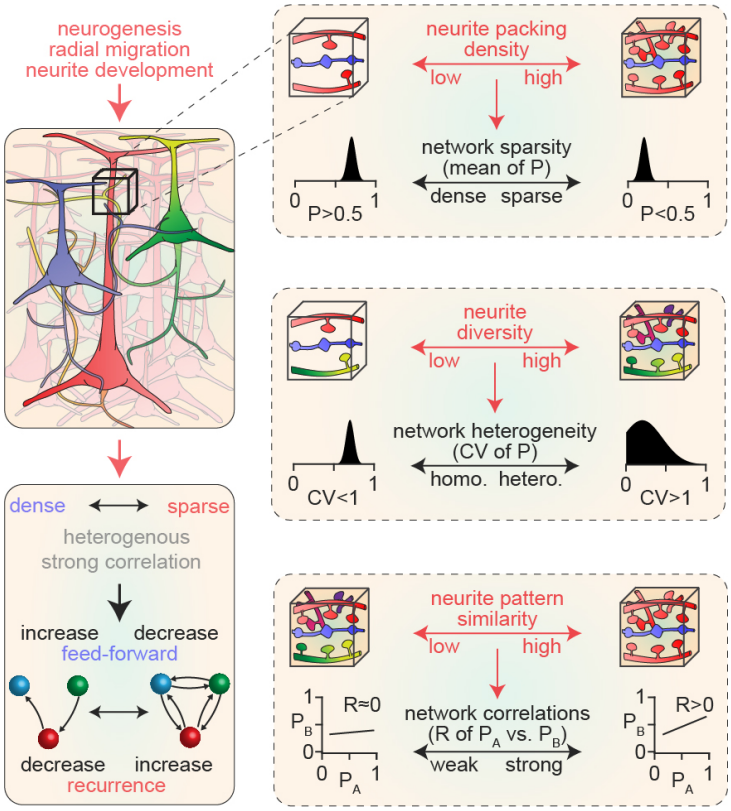

Figure 4.1. Schematic for how structural properties of the neuropil — as established during development — affect properties of the network, such as the means and CVs of connection probability distributions. The predicted network sparsity emerges from the neurite packing density (top panel), the network heterogeneity emerges from the neurite diversity (center panel), and the network correlations emerge from the similarity between the neurite patterns (bottom panel). These structural properties and their derived wiring properties (sparse, heterogeneous, correlated) give rise to nonrandom network topologies, such as the relative increase and decrease in the frequency of the recurrent and feed-forward network motif, respectively, compared to a random network (lower left panel). A network with different properties (e.g., dense wiring) alters the expected frequencies of various motifs. For more details, see main text. Figure adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020).

single pre- or postsynaptic structure[4], a more densely wired network would be predicted to emerge.

Empirical evidence for sparse connectivity: Neurons in the neocortex are sparsely connected, e.g., the majority of neurites in close proximity remain unconnected (e.g., Kasthuri et al., 2015; Motta et al., 2019) and connection probabilities between surrounding neurons are generally below 50% (e.g., Lefort et al., 2009).

Heterogeneity: The predicted heterogeneous connectivity (i.e., CV of the connection probabilities is close or larger than 1; see Figure 3.16) emerges from the diversely packed neuropil. Axons and dendrites of uniquely shaped and positioned neurons, even within the same morphologically defined cell type (Narayanan et al., 2015), intermingle in space (see Figure 3.6), each locally forming synapses in a stochastic manner. The result is that each neuron in the model has a unique wiring pattern, that when grouped together, results in heterogeneous connectivity. In a scenario with a less diverse neuropil, in which subvolumes are, for example, only innervated by the same two neurons or cell type populations, the heterogeneity of the predicted network would drop. In line with these interpretations, a recent modeling approach found that morphological diversity increases network heterogeneity and ultimately, alters higher-order network topologies (Nolte, Gal, Markram, & Reimann, 2020).

Empirical evidence for heterogeneous wiring: Wiring between neurons in the neocortex has been shown to vary substantially, e.g., the number of thalamic synapses varies between neocortical neurons of similar cortical depth (cat primary visual cortex: da Costa & Martin, 2011; rat and mouse vS1: Furuta, Deschênes, & Kaneko, 2011; Schoonover et al., 2014; Motta et al., 2019). As a side note, the distributions underlying the heterogeneous wiring in the model are often skewed, a property also observed in the brain (Buzsáki & Mizuseki, 2014)

Correlations: The interplay between different structural features gives rise to correlated connectivity in the model, summarized in Figure 4.1 by the term neurite pattern similarity. Intuitively, if presynaptic neurons innervate similar cortical

[4]The measured average number of synapses per spine or bouton is close to one in the somatosensory cortex (Santuy et al., 2018)

subvolumes with respect to the targeted population, their wiring patterns are predicted to be similar, and thus, potentially correlated. This is, for example, the case for the in-degree correlations of $L6_ACC$ and L5PT onto L5PTs. There, the innervation volume and density (i.e., DSC) between $L6_ACC$ and L5PT axon projections and L5PT dendrites are similar: Both populations innervate L5PT basal dendrites more densely than their apical dendrites (see Figure 3.25A). Structural variations within the postsynaptic population, such as varying soma depth and dendrite length (Narayanan et al., 2015), will hence affect the number of predicted synapses per presynaptic population in similar ways. In this particular example, dendrite length and soma depth of L5PTs are both positively correlated with the number of synapses they are predicted to receive from $L6_ACC$ and L5PT ($R > 0.32$). In contrast, in the case of the uncorrelated in-degree correlations of L5PT and L5IT onto L5ITs, the degree of innervation onto L5ITs by L5ITs is much stronger than by L5PTs, especially along the apical dendrites that L5PTs barely innervate. Structural variations within L5ITs, therefore, affect the number of predicted synapses in different ways. In this particular case, there are no common correlations between soma depth and dendrite length of L5ITs and the number of synapses they are predicted to receive from the two populations. While the in-degree correlations predicted by the model are biased towards positive, some in-degree correlations are negative (see Figure 3.25BC). The negative correlation in case of the in-degrees of $L6_ACC$ and L5IT onto $L6_ACC$ located in the septal region between the cortical columns emerges from the particular structural arrangement between the neurons: Since $L6_ACC$ dendrites terminate in the upper part of L4, the input domain is restricted to the granular and infragranular layers. There, L5IT axons are less dense and more restricted to regions located within the barrel column. Hence, their inputs are narrowly distributed within the cortical column, while $L6_ACC$ inputs extend beyond the horizontal dimensions of the cortical column. This results in the prediction that while inputs from L5ITs decrease with the horizontal distance of the postsynaptic $L6_ACC$ to the barrel column ($R = -0.39$), inputs form $L6_ACC$ are unaffected ($R = 0.06$).

Empirical evidence for correlated wiring: Anatomical evidence for correlated connectivity is currently not available due to the lack of appropriate methods. However, observed correlated activity might be one indication of correlated synap-

tic inputs (e.g., Smith & Kohn, 2008; Xue, Atallah, & Scanziani, 2014).

In combination, these network properties contribute to the nonrandom network topologies predicted by the model, another property commonly observed in brain networks (e.g., Watts & Strogatz, 1998; Milo et al., 2002; Song et al., 2005). The mathematical model demonstrates what impact the different network properties have on the frequency of each triplet motif. A denser network, for example, decreases the frequency of the recurrent motif while increasing the frequency of the feed-forward motif, even shifting it in a new regime where the motif becomes more abundant than in a random network (see Figures 3.31 and A6). Dense networks in the neocortex have only been reported for inhibitory neurons (e.g., Fino & Yuste, 2011). However, network motifs involving neocortical inhibitory neurons were not probed systematically so far. In the cerebellum, feed-forward motifs between inhibitory neurons with a maximal connection probability of 42% are also underrepresented (Rieubland et al., 2014). Therefore, it remains an open question if and how motif spectra of more densely connected inhibitory neurons differ from those of excitatory neurons and to those predicted by the model (see Figure A7).

4.5 What do wiring properties represent that are consistent with the model's predictions?

The majority of predicted wiring properties from the subcellular, cellular, to the network level, are remarkably consistent with empirical data. This, however, does not mean that they originate from the stochastic synapse formation strategy formulated within this thesis. In fact, no conclusion about their underlying synapse formation strategy can be made. It is equally likely that wiring properties consistent with the model's predictions are, for example, due to recognition-dependent synapse formation strategies induced by activity and/or explicitly encoded in the genome (see Figure 4.2A and Sanes & Yamagata, 2009). However, what can be concluded from this finding is that any of the consistently predicted wiring properties does not necessarily require a recognition-dependent synapse formation strategies to emerge.

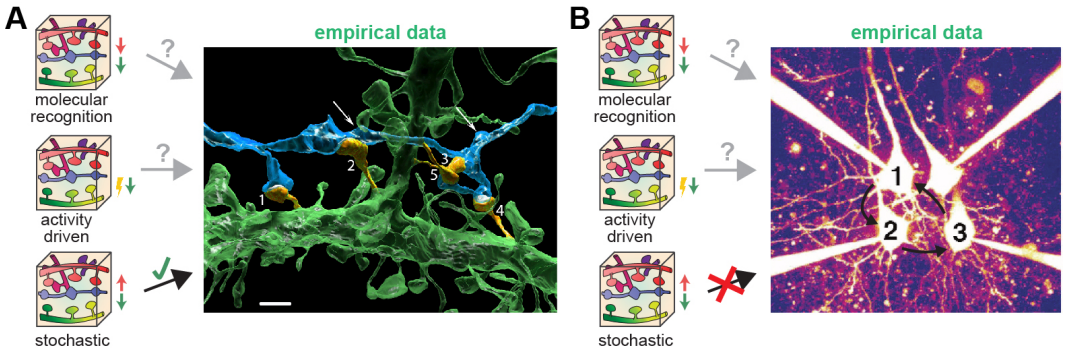

Figure 4.2. Consistently vs. inconsistently predicted wiring properties. (A) Consistency between empirical and predicted wiring properties demonstrates that a stochastic strategy of synapse formation cannot be ruled out to contribute to the observed property, e.g., multiple synapses between two neuron branchlets (Kasthuri et al., 2015). However, other strategies are equally likely to be involved. Right image taken from Kasthuri et al. (2015). (B) Inconsistency between empirical and predicted wiring properties demonstrates that the observed wiring property, e.g., overrepresentation of feed-forward loops (Song et al., 2005), cannot solely emerge from the herein formulated stochastic synapse formation strategy. Right image modified from Song et al. (2005).

In particular activity-driven mechanisms inspired by the work of Hebb (1949) but also molecular cues are speculated to give rise to clustered connectivity observed across scales (e.g., Song et al., 2005; Perin et al., 2011; Kasthuri et al., 2015; Lee et al., 2016). Kasthuri et al. (2015) speculated that activity-driven recognition-dependent strategies could underlie their observation of multiple synapses between the same pair of neurite branches, while the majority of nearby branch pairs remain unconnected. While my approach cannot give any clue about the origins of that particular observed synapse pattern, the pure frequency of observing such synapse patterns, however, should not be considered as evidence for an activity-dependent synapse formation — as the same frequency is expected in a network that is only shaped by its structural composition (see Figure 3.32). Likewise, network topologies deviating from those of a random network were speculated, for example, to underlie learning processes (Song & Abbott, 2001; Song et al., 2005; Rieubland et al., 2014; Ravid Tannenbaum & Burak, 2016; D. Zhang et al., 2019) or be genetically encoded (Perin et al., 2011). As discussed by Rieubland et al. (2014) and in line with the presented findings, nonrandom network topologies

cannot only emerge from genetically encoded rules or activity-dependent plasticity mechanisms but implicitly through variations in the structural constraints within the neuropil (see Figure 3.43). A recent preprint by Gal et al. (2019) supports these findings showing that nonrandom network topologies can emerge from the asymmetric shapes of neuron morphologies and their relative spatial positions. Compared to encoding wiring patterns explicitly in the genome, an implicit wiring strategy, such as the one described here, allows passing through the "genomic bottleneck", i.e., the genome's insufficient capacity to precisely store every single connection (Zador, 2019).

Innate wiring scaffoldings: Although the origin of these wiring properties remains unknown, the remarkable consistency of many of the model's predictions indicates that whatever developmental mechanisms shaped the structural composition of the neuropil, specifically the guidance of neurites into particular subvolumes, provides strong constraints for the network architecture of vS1. Since the network architecture is derived from structural properties preserved across animals, the network's properties potentially represent those preserved across animals as well. Because these wiring properties can significantly impact network function and dynamics (e.g., Landau et al., 2016; Nolte et al., 2020) or even animal behavior (Linneweber et al., 2020), the consistently predicted wiring properties could represent "innate wiring scaffoldings" that guarantee proper cortex function. The plethora of plasticity mechanisms (Feldman, 2009) that alter the circuitry during adulthood (e.g., by forming or eliminating synapses) may, therefore, only remodel the wiring diagram in such a way that most of the wiring properties (i.e., innate wiring scaffoldings) are maintained. This kind of homeostatic regulation of wiring properties may act to maintain core circuit function throughout an animal's life while at the same time encoding extrinsic information induced by learning or experience.

Take, for example, the comparison between predicted and putative synaptic contact distributions obtained by Dr. David A. Slabik (see Section 3.3.5; Slabik, 2018). The model's consistency is higher on the cellular level (i.e., number of synapses) than on the subcellular level (i.e., number and position of synapses) (see Figures 3.46 and 3.49). One could speculate that this discrepancy in consistency

on the subcellular versus cellular level could be due to plasticity mechanisms that alter synapse locations on the postsynaptic dendrite while largely maintaining the number of synapses for the respective connection (i.e., their innate wiring scaffoldings).

In summary, although the origins of consistently predicted wiring properties remain unknown, they, in principle, can emerge from the structural constraints imposed by the neuropil without requiring recognition-dependent synapse formation strategies. The consistently predicted wiring properties might represent innate wiring scaffoldings, implicitly encoded through guidance mechanisms that shaped the neuropil.

4.6 What do wiring properties represent that are not consistent with the model's predictions?

Multiple empirically determined wiring properties are inconsistently predicted by the model. Most notably the number of synapses per connection on the cellular level (e.g., Markram et al., 1997; Feldmeyer, Lübke, Silver, & Sakmann, 2002; Lübke, Roth, Feldmeyer, & Sakmann, 2003) and the relative frequency of individual network motifs such as the feedforward loop (Song et al., 2005). Hence, the question arises what do these wiring properties represent? The most obvious answer to this question is that these properties represent wiring properties that cannot emerge from the herein formulated stochastic synapse formation strategies (see Figure 4.2B). Consequently, these properties might reflect wiring induced by other strategies. These strategies could also be related to structural properties present in the neuropil, that were not part of the herein formulated strategy (e.g., angle of axodendritic approach; Balaskas et al., 2019) or they depend on molecular cues or activity between the pre- and postsynaptic neurons (Sanes & Yamagata, 2009). Precisely, activity-dependent strategies are speculated to represent learning or memory processes (Bailey & Kandel, 1993; Chklovskii et al., 2004; Fox & Wong, 2005; S. J. Barnes & Finnerty, 2010; Albieri et al., 2015; Bennett et al., 2018).

However, as discussed in Section 4.3.3, empirical wiring patterns come with limitations that might explain some of the discrepancies. In the following, I discuss connection probabilities, number of synapses per connection, and network motifs with respect to their respective experimental limitations.

Connection Probabilities: One empirical connection probability was entirely outside of the predicted range across all conditions (see Figure 3.33): Thomson et al. (2002) reported a connection probability from L3 to L5 excitatory neurons of 55% derived by probing 29 neuron pairs functionally. Such a high connection probability is in stark contrast to more recent empirical connection probabilities from L3 to L5A and L5B excitatory neurons also derived functionally but from more than three times as many tested pairs (P(L3,L5A) = 6%, n = 87 tested pairs, and P(L3,L5B) = 12%, n = 164 tested pairs; Lefort et al., 2009). These connection probabilities are more in line with the model's predictions. The reason for the discrepancy can be many-sided, such as differences in sample size, in slice thickness leading to different degrees of truncation of neurites (450 µm to 500 µm vs. 300 µm), in animal age and species (young adult rats without disclosing the age range vs. mice, 18 to 21 days old), as well as the usage of sharp electrodes by Thomson et al. (2002) whose recordings are nosier (Li, Soffe, & Roberts, 2004; Thomson & Lamy, 2007) than of whole cell patch electrodes used by Lefort et al. (2009). It is, therefore, unclear to what degree the reported connection probability for L3 to L5 excitatory neurons reflects their actual wiring *in vivo*. Note that Thomson et al. (2002) reported four other layer-specific connection probabilities that are consistent with their respective predictions and with connection probabilities reported in other studies.

Number of synapses per connection: In contrast to connection probabilities, the observed average number of synapses between connected neuron pairs is inconsistent with the majority of the model's predictions (see Figure 3.39). Empirical data shows that if neuron pairs are connected, they are connected on average not by one synapse — as predicted by the model — but by multiple synapses (e.g., Markram et al., 1997; Feldmeyer et al., 2002; Lübke et al., 2003). The inconsistency with respect to the average number of synapses per connected pair (e.g.,

Schoonover et al., 2014 vs. model n(VPM,L4): 2.5 vs. 1.8 synapses per connection) and the consistency with respect to connection probabilities (e.g., Bruno & Sakmann, 2006 vs. model P(VPM,L4): 43% vs. 44%) is at first surprising since both wiring properties are not independent of each other: In theory, doubling the number of synapses per connected pair within the framework of the model results in halving the respective average connection probability given that the pre- and postsynaptic densities do not change. How can this discrepancy be resolved?

First, the discrepancy could be due to experimental limitations such as in identifying only strongly connected neuron pairs, i.e., those connected by multiple synapses evoking a detectable postsynaptic potential at the soma. The preliminary data by Dr. David A. Slabik indicates that the distribution of the number of putative synapses per connection could be less bimodal than inferred from functional recordings (Slabik, 2018). His reported putative synapse numbers on the cell-to-cell level are consistent with the model's predictions (see Figure 3.46).

Second, the discrepancy could indicate violations of one of the model's assumptions, such as the independence of synapse formation (Fares & Stepanyants, 2009). However, for measurements involving VPM synapses the violation of another assumption resolves parts of the discrepancy. While the model assumes that each bouton forms exactly one synapse, recent empirical data shows that within vS1 a single bouton along axons from VPM thalamus can form up to four synapses with an average 1.6 synapses per bouton (Rodriguez-Moreno et al., 2018). Taking this finding into account could hence explain the ∼1.4 factor difference between the empirical and predicted number of synapses per VPM to L4 connection (Schoonover et al., 2014). However, for the remaining measurements, the discrepancy cannot be resolved at the moment.

Network motifs: The empirically derived network motif spectra (Song et al., 2005; Kampa et al., 2006; Perin et al., 2011) come with the same limitations as other functionally derived connectivity statistics. Despite these limitations, it is remarkable that, while quantitatively the degree of deviations does not match, the direction and relative degree of most deviations are consistent with their respective predictions in the model ($R \geq 0.64$; see Figure 3.43). Among the few inconsistent predictions is the direction of the deviation of the feedforward loop

(motif ID 7). While measurements revealed that the motif was overrepresented[5], the model predicted the motif to be underrepresented (i.e., less frequent than in a random network). Since the directions of all other motifs match, experimental limitations are rather unlikely to explain the discrepancy. The motif could, therefore, represent a structural imprint of a synapse formation strategy other than the herein formulated one. Although having to be interpreted with caution, the quantitative mismatch between the predicted and measured motif deviations across different motif sizes (see Figures 3.43, 3.44, and 3.45) suggests an involvement of other synapse formation strategies. These could, for example, act on top of the predicted wiring scaffoldings.

How to identify structural imprints of different synapse formation strategies: The findings show that if trying to detect structural imprints, for example of activity-driven synapse formation strategies, it is more adequate to probe higher-order connectivity patterns than first-order connectivity statistics, such as connection probabilities. Network topologies but also correlations or variations in first-order connectivity statistics (e.g., shape of the distribution) hence represent suitable observables that might be shaped by learning, experience, or genetically predefined wiring rules. For example, a narrow connection probability distribution between highly diverse neuron morphology, strong negative in-degree correlations, or an underrepresented fully recurrent network motif could be an indication of the involvement of a recognition-dependent synapse formation strategy. On the subcellular level, compartment specificity (e.g., Schoonover et al., 2014; Karimi et al., 2020) may most likely reflect a different synapse formation strategy. Technological advances in electron microscopy and cell labeling and recording techniques will make more wiring properties, especially higher-order statistics, observable and thus, testable in the future.

In summary, inconsistently predicted wiring properties indicate origins other than the formulated stochastic strategy. The strategy and the model hence can be considered a null hypothesis for testing what properties could be due to recognition-dependent synapse formation strategies relying, for example, on activity or genet-

[5] when correcting for doublet motifs

ically specified rules.

4.7 Wiring involving inhibitory neurons

Given the previously discussed limitations of the inhibitory sample (see Discussion Section 4.3.1) and the large diversity of the inhibitory population (Cadwell et al., 2016; Feldmeyer et al., 2018) the predictions of the model with respect to inhibitory neurons need to be considered with caution. A common wiring theme involving inhibitory neurons is their wiring specificity, e.g., preferential targeting of specific neuron populations (e.g., Zhou, Rickmann, Hafner, & Staiger, 2017) or even subcellular domains (e.g., DeFelipe, 1997; Somogyi et al., 1998; Motta et al., 2019; Karimi et al., 2020). These studies indicate that the distribution of inhibitory axons and dendrites within the neuropil might have little impact on their wiring properties. For example, Motta et al. (2019) showed that in contrast to excitatory connectivity, structural properties were not predictive of inhibitory connectivity on the electron microscopic level.

It therefore comes as a surprise that there are wiring properties involving inhibitory neurons that are within the predicted ranges of the model. One such wiring property largely consistent with empirical data is synapse densities along dendrites and somata, with the exception of inhibitory synapse densities on inhibitory neurons (see Figure 3.40). The reported inhibitory synapse densities by White and Rock (1981) and Hioki et al. (2013) are much larger than predicted and than reported in other studies (White et al., 1984; Keller & White, 1987). These discrepancies are likely due to different sampling of dendrite branches and neuron populations: White and Rock (1981) probed more proximal dendrite branches, while White et al. (1984) and Keller and White (1987) probed more distal dendrite branches; and Hioki et al. (2013) probed only a subpopulation of molecularly identified inhibitory neurons.

Another predicted wiring property that is consistent with empirical data is the average connection probabilities involving inhibitory neurons (see Figure 3.42). The reason for the consistency, although lower than for the excitatory neurons, could be that the shape and positioning of inhibitory axons and dendrites within the neuropil contribute to their pair-wise connectivity statistics, a conclusion in

agreement with Fino and Yuste (2011) and Packer et al. (2013). However, it has to be emphasized that for the comparison I did not mimic the respective experimental conditions; instead, I only sampled from inhibitory neurons of the same layer as in the respective experiments, thereby grouping highly diverse inhibitory morphologies together (e.g., Koelbl et al., 2013; Muñoz, Tremblay, & Rudy, 2014; Nigro, Hashikawa-Yamasaki, & Rudy, 2018). A more precise grouping of morphologically or molecularly identified inhibitory neurons in combination with taking into account the respective experimental conditions will change the consistency between predicted and empirical connectivity statistics. Therefore, it requires further investigation of how morphologies of inhibitory neurons, specifically those of particular inhibitory cell types, impact wiring properties. It is expected that wiring properties, such as the preferential subcellular targeting of particular inhibitory types (e.g., Somogyi, 1977; Zhou et al., 2017), is beyond what would be expected from the herein formulated synapse formation strategy.

4.8 Conclusion & Outlook

The major result presented in this thesis is that it is possible to reverse engineer an anatomically realistic model that captures the characteristic structural composition of one neocortical area as well as many properties of its network architecture. Specifically, I found that the structural constraints for synapse formation imposed through neurons and neurites can give rise to specialized innate wiring scaffoldings of which many properties are consistent with empirical observations. This finding demonstrates that many wiring properties, such as sparse wiring or nonrandom network topologies, do not necessarily reflect recognition-dependent synapse formation strategies. However, some observed wiring properties, such as the frequency of certain motifs or wiring properties involving inhibitory neurons, deviate from the model's predictions, raising questions about their underlying synapse formation strategies. A more likely source of these wiring properties might hence be a recognition-dependent synapse formation strategy. Thus, the model can act as a null hypothesis for recognition-dependent synapse formation strategies rendering the model a unique starting point to identify wiring correlates of sensory experience, learning, or memory.

Future work can address the current limitations of the model and incorporate more structural constraints. For example, currently the group of Dr. Marcel Oberlaender in collaboration with Prof. Dr. Dirk Feldmeyer (Research Center Jülich and RWTH Aachen University) is acquiring and incorporating a more representative sample of inhibitory neurons (e.g., including the non-fast-spiking L4 inhibitory neuron morphologies by Emmenegger et al., 2018). In addition, the 3D distributions of molecularly identified inhibitory neurons within rat vS1 are quantified and first *in vivo* labeled and reconstructed inhibitory neurons are characterized with respect to their morphological and molecular properties and sensory-evoked responses. These ongoing efforts will help to better constrain the inhibitory component of the model and advance our knowledge of the role of inhibitory neurons within neocortex (Feldmeyer et al., 2018; Roux & Buzsáki, 2015).

Furthermore, future work can utilize the unique opportunity provided by the current model and its respective web application *Cortex in Silico*, such as exploring yet unanswered questions about the relationship between neuron structure, synapse formation strategies, and network architecture. For example, the herein presented approach enables future endeavors to quantitatively compare wiring patterns across different stages of development, aging, or neuropathology. Currently, it would be, for instance, possible to investigate the impact of experience-induced remodeling of thalamic axons in rat vS1 onto predicted wiring distributions (Oberlaender, Ramirez, & Bruno, 2012).

The wiring of neocortical networks is a pivotal constraint for cortical function and signal flow. Due to the model's realistic wiring predictions and the current lack of an entire empirical wiring diagram, the model allows studying neural activity in an realistically constraint network and investigating the impact of both the structure of the neuropil and wiring on cortex function. The model (or previous versions thereof) has been used so far in three studies to elucidate functional principles (Egger et al., 2015; Landau et al., 2016; Egger et al., 2020), for example, by providing constraints for the synaptic inputs to full compartmental models of L5PT neurons.

In the future, the model, in combination with functional data, might be able to answer questions about the function of vS1 (Stüttgen & Schwarz, 2018). Notably, the approach and model is, in principle, neither limited to vS1 nor the stochastic

synapse formation strategy. The presented web application *Cortex in Silico* provides the opportunity to formulate, compare, and test different synapse formation strategies against empirical data. In the future, *Cortex in Silico* will enable the user to perform inferences over differently constraint strategies by minimizing the deviation to the empirical data. The herein presented model and web application will, therefore, provide a foundation for exploring the links between strategies of synapse formation, distribution of neurons and neurites, network structure, and their relevance for cortex function.

References

Ahmed, B., Anderson, J. C., Martin, K. A., & Nelson, J. C. (1997, April). Map of the synapses onto layer 4 basket cells of the primary visual cortex of the cat. *The Journal of comparative neurology*, *380*, 230–242.

Albieri, G., Barnes, S. J., de Celis Alonso, B., Cheetham, C. E. J., Edwards, C. E., Lowe, A. S., ... Finnerty, G. T. (2015, September). Rapid bidirectional reorganization of cortical microcircuits. *Cerebral cortex*, *25*, 3025–3035. doi: 10.1093/cercor/bhu098

Alvarez, V. A., & Sabatini, B. L. (2007). Anatomical and physiological plasticity of dendritic spines. *Annual review of neuroscience*, *30*, 79–97.

Arzt, M., Sakmann, B., & Meyer, H. S. (2018, August). Anatomical correlates of local, translaminar, and transcolumnar inhibition by layer 6 gabaergic interneurons in somatosensory cortex. *Cerebral cortex*, *28*, 2763–2774. doi: 10.1093/cercor/bhx156

Ascoli, G. A., Alonso-Nanclares, L., Anderson, S. A., Barrionuevo, G., Benavides-Piccione, R., Burkhalter, A., ... Yuste, R. (2008, July). Petilla terminology: nomenclature of features of gabaergic interneurons of the cerebral cortex. *Nature Reviews Neuroscience*, *9*, 557–568. doi: 10.1038/nrn2402

Avermann, M., Tomm, C., Mateo, C., Gerstner, W., & Petersen, C. C. H. (2012, June). Microcircuits of excitatory and inhibitory neurons in layer 2/3 of mouse barrel cortex. *Journal of neurophysiology*, *107*, 3116–3134. doi: 10.1152/jn.00917.2011

Bailey, C. H., & Kandel, E. R. (1993). Structural changes accompanying memory storage. *Annual review of physiology*, *55*, 397–426. doi: 10.1146/annurev.ph.55.030193.002145

Balaskas, N., Abbott, L., Jessell, T. M., & Ng, D. (2019). Positional strategies for connection specificity and synaptic organization in spinal sensory-motor circuits. *Neuron*.

Bannister, A. P., & Thomson, A. M. (2007, September). Dynamic properties of excitatory synaptic connections involving layer 4 pyramidal cells in adult rat and cat neocortex. *Cerebral cortex*, *17*, 2190–2203. doi: 10.1093/cercor/bhl126

Barabási, D. L., & Barabási, A.-L. (2020, February). A genetic model of the connectome. *Neuron*, *105*, 435–445.e5. doi: 10.1016/j.neuron.2019.10.031

Barnes, A. P., & Polleux, F. (2009). Establishment of axon-dendrite polarity in developing neurons. *Annual review of neuroscience*, *32*, 347–381.

Barnes, S. J., & Finnerty, G. T. (2010, April). Sensory experience and cortical rewiring. *The Neuroscientist*, *16*, 186–198. doi: 10.1177/1073858409343961

Barth, A., Burkhalter, A., Callaway, E. M., Connors, B. W., Cauli, B., DeFelipe, J., ... Yuste, R. (2016, September). Comment on "principles of connectivity among morphologically defined cell types in adult neocortex". *Science*, *353*, 1108. doi: 10.1126/science.aaf5663

Beierlein, M., & Connors, B. W. (2002, October). Short-term dynamics of thalamocortical and intracortical synapses onto layer 6 neurons in neocortex. *Journal of neurophysiology*, *88*, 1924–1932. doi: 10.1152/jn.2002.88.4.1924

Beierlein, M., Gibson, J. R., & Connors, B. W. (2003, November). Two dynamically distinct inhibitory networks in layer 4 of the neocortex. *Journal of neurophysiology*, *90*, 2987–3000. doi: 10.1152/jn.00283.2003

Bennett, S. H., Kirby, A. J., & Finnerty, G. T. (2018). Rewiring the connectome: evidence and effects. *Neuroscience & Biobehavioral Reviews*, *88*, 51–62.

Berry, K. P., & Nedivi, E. (2017, September). Spine dynamics: Are they all the same? *Neuron*, *96*, 43–55. doi: 10.1016/j.neuron.2017.08.008

Billeh, Y. N., Cai, B., Gratiy, S. L., Dai, K., Iyer, R., Gouwens, N. W., ... others (2020). Systematic integration of structural and functional data into multi-scale models of mouse primary visual cortex. *Neuron*.

Bosking, W. H., Zhang, Y., Schofield, B., & Fitzpatrick, D. (1997). Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. *Journal of neuroscience*, *17*(6), 2112–2127.

Boucsein, C., Nawrot, M. P., Schnepel, P., & Aertsen, A. (2011). Beyond the cortical column: abundance and physiology of horizontal connections imply a strong role for inputs from the surround. *Frontiers in neuroscience*, *5*, 32. doi: 10.3389/fnins.2011.00032

Bourgeois, J.-P., & Rakic, P. (1993). Changes of synaptic density in the primary visual cortex of the macaque monkey from fetal to adult stage. *Journal of neuroscience*, *13*(7), 2801–2820.

Braitenberg, V., & Schüz, A. (1991). Peters' rule and white's exceptions. In *Anatomy of the cortex: Statistics and geometry* (pp. 109–112). Berlin, Heidelberg: Springer Berlin Heidelberg. doi: 10.1007/978-3-662-02728-8_21

Braitenberg, V., & Schüz, A. (1998). *Cortex: Statistics and geometry of neuronal connectivity*. Springer.

Brown, S. P., & Hestrin, S. (2009). Intracortical circuits of pyramidal neurons reflect their long-range axonal targets. *Nature*, *457*(7233), 1133.

Bruno, R. M., & Sakmann, B. (2006). Cortex is driven by weak but synchronously active thalamocortical synapses. *Science*, *312*(5780), 1622–1627.

Bruno, R. M., & Simons, D. J. (2002, December). Feedforward mechanisms of excitatory and inhibitory cortical receptive fields. *Journal of neuroscience*, *22*, 10966–10975.

Buzsáki, G., & Mizuseki, K. (2014, April). The log-dynamic brain: how skewed distributions affect network operations. *Nature Reviews Neuroscience*, *15*, 264–278. doi: 10.1038/nrn3687

Cadwell, C. R., Palasantza, A., Jiang, X., Berens, P., Deng, Q., Yilmaz, M., ... Tolias, A. S. (2016, February). Electrophysiological, transcriptomic and morphologic profiling of single neurons using patch-seq. *Nature biotechnology*, *34*, 199–203. doi: 10.1038/nbt.3445

Cantallops, I., & Cline, H. T. (2008, May). Rapid activity-dependent delivery of the neurotrophic protein cpg15 to the axon surface of neurons in intact xenopus tadpoles. *Developmental neurobiology*, *68*, 744–759. doi: 10.1002/dneu.20529

Chen, J. L., Villa, K. L., Cha, J. W., So, P. T. C., Kubota, Y., & Nedivi, E. (2012, April). Clustered dynamics of inhibitory synapses and dendritic spines in the adult neocortex. *Neuron*, *74*, 361–373. doi: 10.1016/j.neuron.2012.02.030

Cheng, P.-l., & Poo, M.-m. (2012). Early events in axon/dendrite polarization. *Annual review of neuroscience*, *35*, 181–201.

Chklovskii, D. B., Mel, B. W., & Svoboda, K. (2004, October). Cortical rewiring and information storage. *Nature*, *431*, 782–788. doi: 10.1038/nature03012

Cline, H., & Haas, K. (2008). The regulation of dendritic arbor development and plasticity by glutamatergic synaptic input: a review of the synaptotrophic hypothesis. *The Journal of physiology*, *586*(6), 1509–1517.

Cohen-Kashi Malina, K., Mohar, B., Rappaport, A. N., & Lampl, I. (2016, September). Local and thalamic origins of correlated ongoing and sensory-evoked cortical activities. *Nature communications*, *7*, 12740. doi: 10.1038/ncomms12740

Constantinople, C. M., & Bruno, R. M. (2013, June). Deep cortical layers are activated directly by thalamus. *Science*, *340*, 1591–1594. doi: 10.1126/science.1236425

Crandall, S. R., Patrick, S. L., Cruikshank, S. J., & Connors, B. W. (2017). Infrabarrels are layer 6 circuit modules in the barrel cortex that link long-range inputs and outputs. *Cell reports*, *21*(11), 3065–3078.

Cuntz, H., Forstner, F., Borst, A., & Häusser, M. (2010, August). One rule to grow them all: a general theory of neuronal branching and its practical application. *PLoS computational biology*, *6*. doi: 10.1371/journal.pcbi.1000877

da Costa, N. M., & Martin, K. A. C. (2011, February). How thalamus connects to spiny stellate cells in the cat's visual cortex. *Journal of neuroscience*, *31*, 2925–2937. doi: 10.1523/JNEUROSCI.5961-10.2011

Davison, A. P., Brüderle, D., Eppler, J., Kremkow, J., Muller, E., Pecevski, D., ... Yger, P. (2009). Pynn: A common interface for neuronal network simulators. *Frontiers in neuroinformatics*, *2*, 11. doi: 10.3389/neuro.11.011.2008

DeFelipe, J. (1997, December). Types of neurons, synaptic connections and chemical characteristics of cells immunoreactive for calbindin-d28k, parvalbumin and calretinin in the neocortex. *Journal of chemical neuroanatomy*, *14*, 1–19. doi: 10.1016/s0891-0618(97)10013-8

DeFelipe, J. (2015). The anatomical problem posed by brain complexity and size: a potential solution. *Frontiers in neuroanatomy*, *9*, 104. doi: 10.3389/fnana.2015.00104

DeFelipe, J., López-Cruz, P. L., Benavides-Piccione, R., Bielza, C., Larrañaga, P., Anderson, S., ... Ascoli, G. A. (2013, March). New insights into the classification and nomenclature of cortical gabaergic interneurons. *Nature Reviews Neuroscience*, *14*, 202–216. doi: 10.1038/nrn3444

De Paola, V., Holtmaat, A., Knott, G., Song, S., Wilbrecht, L., Caroni, P., & Svoboda, K. (2006, March). Cell type-specific structural plasticity of axonal branches and boutons in the adult neocortex. *Neuron*, *49*, 861–875. doi: 10.1016/j.neuron.2006.02.017

De Wit, J., & Ghosh, A. (2016). Specification of synaptic connectivity by cell surface interactions. *Nature Reviews Neuroscience*, *17*(1), 4.

Eavri, R., Shepherd, J., Welsh, C. A., Flanders, G. H., Bear, M. F., & Nedivi, E. (2018). Interneuron simplification and loss of structural plasticity as markers of aging-related functional decline. *Journal of neuroscience*, *38*(39), 8421–8432.

Egger, R. (2015). *Simulation of sensory-evoked signal flow in anatomically realistic models of neural networks* (Unpublished doctoral dissertation). Eberhard-Karls-Universität Tübingen, Germany.

Egger, R., Dercksen, V. J., Udvary, D., Hege, H.-C., & Oberlaender, M. (2014). Generation of dense statistical connectomes from sparse morphological data. *Frontiers in neuroanatomy*, *8*, 129. doi: 10.3389/fnana.2014.00129

Egger, R., Narayanan, R. T., Guest, J. M., Bast, A., Udvary, D., Messore, L. F., ... Oberlaender, M. (2020). Cortical output is gated by horizontally projecting neurons in the deep layers. *Neuron*.

Egger, R., Narayanan, R. T., Helmstaedter, M., de Kock, C. P., & Oberlaender, M. (2012). 3d reconstruction and standardization of the rat vibrissal cortex for precise registration of single neuron morphology. *PLoS computational biology*, *8*(12), e1002837.

Egger, R., Schmitt, A. C., Wallace, D. J., Sakmann, B., Oberlaender, M., & Kerr, J. N. D. (2015, November). Robustness of sensory-evoked excitation is increased by inhibitory inputs to distal apical tuft dendrites. *Proceedings of the National Academy of Sciences of the United States of America*, *112*, 14072–14077. doi: 10.1073/pnas.1518773112

Elston, G. N., DeFelipe, J., Arellano, J. I., Gonzilez-Albo, M. C., & Rosa, M. G.

(1999, September). Variation in the spatial relationship between parvalbumin immunoreactive interneurones and pyramidal neurones in rat somatosensory cortex. *Neuroreport*, *10*, 2975–2979. doi: 10.1097/00001756-199909290-00019

Elston, G. N., Pow, D. V., & Calford, M. B. (1997). Neuronal composition and morphology in layer iv of two vibrissal barrel subfields of rat cortex. *Cerebral cortex*, *7*, 422–431. doi: 10.1093/cercor/7.5.422

Emmenegger, V., Qi, G., Wang, H., & Feldmeyer, D. (2018, April). Morphological and functional characterization of non-fast-spiking gabaergic interneurons in layer 4 microcircuitry of rat barrel cortex. *Cerebral cortex*, *28*, 1439–1457. doi: 10.1093/cercor/bhx352

Emoto, K. (2011). Dendrite remodeling in development and disease. *Development, growth & differentiation*, *53*(3), 277–286.

Erzurumlu, R. S., & Gaspar, P. (2012). Development and critical period plasticity of the barrel cortex. *European Journal of neuroscience*, *35*(10), 1540–1553.

Espinosa, J. S., Wheeler, D. G., Tsien, R. W., & Luo, L. (2009). Uncoupling dendrite growth and patterning: single-cell knockout analysis of nmda receptor 2b. *Neuron*, *62*(2), 205–217.

Fares, T., & Stepanyants, A. (2009, September). Cooperative synapse formation in the neocortex. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 16463–16468. doi: 10.1073/pnas.0813265106

Fauth, M., & Tetzlaff, C. (2016). Opposing effects of neuronal activity on structural plasticity. *Frontiers in neuroanatomy*, *10*, 75. doi: 10.3389/fnana.2016.00075

Favero, M., Busetto, G., & Cangiano, A. (2012). Spike timing plays a key role in synapse elimination at the neuromuscular junction. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(25), E1667–E1675.

Feldman, D. E. (2009). Synaptic mechanisms for plasticity in neocortex. *Annual review of neuroscience*, *32*, 33–55. doi: 10.1146/annurev.neuro.051508.135516

Feldman, D. E., & Brecht, M. (2005). Map plasticity in somatosensory cortex. *Science*, *310*(5749), 810–815.

Feldmeyer, D., Brecht, M., Helmchen, F., Petersen, C. C. H., Poulet, J. F. A.,

Staiger, J. F., ... Schwarz, C. (2013, April). Barrel cortex function. *Progress in neurobiology*, *103*, 3–27. doi: 10.1016/j.pneurobio.2012.11.002

Feldmeyer, D., Egger, V., Lubke, J., & Sakmann, B. (1999, November). Reliable synaptic connections between pairs of excitatory layer 4 neurones within a single 'barrel' of developing rat somatosensory cortex. *The Journal of physiology*, *521 Pt 1*, 169–190. doi: 10.1111/j.1469-7793.1999.00169.x

Feldmeyer, D., Lübke, J., & Sakmann, B. (2006, September). Efficacy and connectivity of intracolumnar pairs of layer 2/3 pyramidal cells in the barrel cortex of juvenile rats. *The Journal of physiology*, *575*, 583–602. doi: 10.1113/jphysiol.2006.105106

Feldmeyer, D., Lübke, J., Silver, R. A., & Sakmann, B. (2002, February). Synaptic connections between layer 4 spiny neurone-layer 2/3 pyramidal cell pairs in juvenile rat barrel cortex: physiology and anatomy of interlaminar signalling within a cortical column. *The Journal of physiology*, *538*, 803–822. doi: 10.1113/jphysiol.2001.012959

Feldmeyer, D., Qi, G., Emmenegger, V., & Staiger, J. F. (2018). Inhibitory interneurons and their circuit motifs in the many layers of the barrel cortex. *Neuroscience*, *368*, 132–151.

Feldmeyer, D., Roth, A., & Sakmann, B. (2005, March). Monosynaptic connections between pairs of spiny stellate cells in layer 4 and pyramidal cells in layer 5a indicate that lemniscal and paralemniscal afferent pathways converge in the infragranular somatosensory cortex. *Journal of neuroscience*, *25*, 3423–3431. doi: 10.1523/JNEUROSCI.5227-04.2005

Fino, E., & Yuste, R. (2011, March). Dense inhibitory connectivity in neocortex. *Neuron*, *69*, 1188–1203. doi: 10.1016/j.neuron.2011.02.025

Fishell, G., & Kepecs, A. (2019). Interneuron types as attractors and controllers. *Annual review of neuroscience*, *43*.

Flores, C. E., & Méndez, P. (2014). Shaping inhibition: activity dependent structural plasticity of gabaergic synapses. *Frontiers in cellular neuroscience*, *8*, 327. doi: 10.3389/fncel.2014.00327

Fox, K. (2002). Anatomical pathways and molecular mechanisms for plasticity in the barrel cortex. *Neuroscience*, *111*(4), 799–814.

Fox, K., & Wong, R. O. (2005). A comparison of experience-dependent plasticity

in the visual and somatosensory systems. *Neuron*, *48*(3), 465–477.

Frick, A., Feldmeyer, D., Helmstaedter, M., & Sakmann, B. (2008, February). Monosynaptic connections between pairs of l5a pyramidal neurons in columns of juvenile rat somatosensory cortex. *Cerebral cortex*, *18*, 397–406. doi: 10.1093/cercor/bhm074

Furuta, T., Deschênes, M., & Kaneko, T. (2011, April). Anisotropic distribution of thalamocortical boutons in barrels. *Journal of neuroscience*, *31*, 6432–6439. doi: 10.1523/JNEUROSCI.6154-10.2011

Gabernet, L., Jadhav, S. P., Feldman, D. E., Carandini, M., & Scanziani, M. (2005, October). Somatosensory integration controlled by dynamic thalamocortical feed-forward inhibition. *Neuron*, *48*, 315–327. doi: 10.1016/j.neuron.2005.09.022

Gal, E., Perin, R., Markram, H., London, M., & Segev, I. (2019). Neuron geometry underlies a universal local architecture in neuronal networks. *bioRxiv*. doi: 10.1101/656058

Gao, P., Sultan, K. T., Zhang, X.-J., & Shi, S.-H. (2013, July). Lineage-dependent circuit assembly in the neocortex. *Development (Cambridge, England)*, *140*, 2645–2655. doi: 10.1242/dev.087668

Gibson, J. R., Beierlein, M., & Connors, B. W. (1999, November). Two networks of electrically coupled inhibitory neurons in neocortex. *Nature*, *402*, 75–79. doi: 10.1038/47035

Gouwens, N. W., Sorensen, S. A., Berg, J., Lee, C., Jarsky, T., Ting, J., ... others (2019). Classification of electrophysiological and morphological neuron types in the mouse visual cortex. *Nature neuroscience*, *22*(7), 1182–1195.

Harris, K. D., & Shepherd, G. M. G. (2015, February). The neocortical circuit: themes and variations. *Nature neuroscience*, *18*, 170–181. doi: 10.1038/nn.3917

Hassan, B. A., & Hiesinger, P. R. (2015, October). Beyond molecular codes: Simple rules to wire complex brains. *Cell*, *163*, 285–291. doi: 10.1016/j.cell.2015.09.031

Hebb, D. O. (1949). *The organization of behavior: a neuropsychological theory*. J. Wiley.

Helmstaedter, M., Sakmann, B., & Feldmeyer, D. (2009a). L2/3 interneuron

groups defined by multiparameter analysis of axonal projection, dendritic geometry, and electrical excitability. *Cerebral Cortex*, *19*(4), 951–962.

Helmstaedter, M., Sakmann, B., & Feldmeyer, D. (2009b). Neuronal correlates of local, lateral, and translaminar inhibition with reference to cortical columns. *Cerebral Cortex*, *19*(4), 926–937.

Helmstaedter, M., Sakmann, B., & Feldmeyer, D. (2009c, April). The relation between dendritic geometry, electrical excitability, and axonal projections of l2/3 interneurons in rat barrel cortex. *Cerebral Cortex*, *19*(4), 938–950. doi: 10.1093/cercor/bhn138

Helmstaedter, M., Staiger, J. F., Sakmann, B., & Feldmeyer, D. (2008, August). Efficient recruitment of layer 2/3 interneurons by layer 4 input in single columns of rat somatosensory cortex. *Journal of neuroscience*, *28*, 8273–8284. doi: 10.1523/JNEUROSCI.5701-07.2008

Hioki, H., Okamoto, S., Konno, M., Kameda, H., Sohn, J., Kuramoto, E., ... Kaneko, T. (2013, January). Cell type-specific inhibitory inputs to dendritic and somatic compartments of parvalbumin-expressing neocortical interneuron. *Journal of neuroscience*, *33*, 544–555. doi: 10.1523/JNEUROSCI.2255-12.2013

Hofer, S. B., Ko, H., Pichler, B., Vogelstein, J., Ros, H., Zeng, H., ... Mrsic-Flogel, T. D. (2011, July). Differential connectivity and response dynamics of excitatory and inhibitory neurons in visual cortex. *Nature neuroscience*, *14*, 1045–1052. doi: 10.1038/nn.2876

Holmgren, C., Harkany, T., Svennenfors, B., & Zilberter, Y. (2003, August). Pyramidal cell communication within local networks in layer 2/3 of rat neocortex. *The Journal of physiology*, *551*, 139–153. doi: 10.1113/jphysiol.2003.044784

Holtmaat, A. J., Trachtenberg, J. T., Wilbrecht, L., Shepherd, G. M., Zhang, X., Knott, G. W., & Svoboda, K. (2005). Transient and persistent dendritic spines in the neocortex in vivo. *Neuron*, *45*(2), 279–291.

Hong, W., Mosca, T. J., & Luo, L. (2012). Teneurins instruct synaptic partner matching in an olfactory map. *Nature*, *484*(7393), 201–207.

Hooks, B. M., & Chen, C. (2006, October). Distinct roles for spontaneous and visual activity in remodeling of the retinogeniculate synapse. *Neuron*, *52*, 281–291. doi: 10.1016/j.neuron.2006.07.007

Hubel, D. H., & Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *The Journal of physiology*, *148*(3), 574–591.

Inan, M., & Anderson, S. A. (2014). The chandelier cell, form and function. *Current opinion in neurobiology*, *26*, 142–148.

Inoue, T., & Imoto, K. (2006, October). Feedforward inhibitory connections from multiple thalamic cells to multiple regular-spiking cells in layer 4 of the somatosensory cortex. *Journal of neurophysiology*, *96*, 1746–1754. doi: 10.1152/jn.00301.2006

Ishino, Y., Yetman, M. J., Sossi, S. M., Steinecke, A., Hayano, Y., & Taniguchi, H. (2017). Regional cellular environment shapes phenotypic variations of hippocampal and neocortical chandelier cells. *Journal of neuroscience*, *37*(41), 9901–9916.

Jabaudon, D. (2017). Fate and freedom in developing neocortical circuits. *Nature communications*, *8*(1), 1–9.

Jiang, X., Shen, S., Cadwell, C. R., Berens, P., Sinz, F., Ecker, A. S., ... Tolias, A. S. (2015, November). Principles of connectivity among morphologically defined cell types in adult neocortex. *Science*, *350*, aac9462. doi: 10.1126/science.aac9462

Jiang, X., Shen, S., Sinz, F., Reimer, J., Cadwell, C. R., Berens, P., ... others (2016). Response to comment on "principles of connectivity among morphologically defined cell types in adult neocortex". *Science*, *353*(6304), 1108–1108.

Jouhanneau, J.-S., Kremkow, J., Dorrn, A. L., & Poulet, J. F. A. (2015, December). In vivo monosynaptic excitatory transmission between layer 2 cortical pyramidal neurons. *Cell reports*, *13*, 2098–2106. doi: 10.1016/j.celrep.2015.11.011

Jouhanneau, J.-S., Kremkow, J., & Poulet, J. F. A. (2018, April). Single synaptic inputs drive high-precision action potentials in parvalbumin expressing gabaergic cortical neurons in vivo. *Nature communications*, *9*, 1540. doi: 10.1038/s41467-018-03995-2

Kaas, J. H. (2011). Neocortex in early mammals and its subsequent variations. *Annals of the New York Academy of Sciences*, *1225*.

Kameda, H., Hioki, H., Tanaka, Y. H., Tanaka, T., Sohn, J., Sonomura, T., ...

Kaneko, T. (2012, March). Parvalbumin-producing cortical interneurons receive inhibitory inputs on proximal portions and cortical excitatory inputs on distal dendrites. *The European journal of neuroscience*, *35*, 838–854. doi: 10.1111/j.1460-9568.2012.08027.x

Kampa, B. M., Letzkus, J. J., & Stuart, G. J. (2006, December). Cortical feed-forward networks for binding different streams of sensory information. *Nature neuroscience*, *9*, 1472–1473. doi: 10.1038/nn1798

Kapfer, C., Glickfeld, L. L., Atallah, B. V., & Scanziani, M. (2007, June). Supralinear increase of recurrent inhibition during sparse activity in the somatosensory cortex. *Nature neuroscience*, *10*, 743–753. doi: 10.1038/nn1909

Karimi, A., Odenthal, J., Drawitsch, F., Boergens, K. M., & Helmstaedter, M. (2020, February). Cell-type specific innervation of cortical pyramidal cells at their apical dendrites. *eLife*, *9*. doi: 10.7554/eLife.46876

Karube, F., Kubota, Y., & Kawaguchi, Y. (2004). Axon branching and synaptic bouton phenotypes in gabaergic nonpyramidal cell subtypes. *Journal of neuroscience*, *24*(12), 2853–2865.

Kasthuri, N., Hayworth, K. J., Berger, D. R., Schalek, R. L., Conchello, J. A., Knowles-Barley, S., ... Lichtman, J. W. (2015, July). Saturated reconstruction of a volume of neocortex. *Cell*, *162*, 648–661. doi: 10.1016/j.cell.2015.06.054

Katz, L. C., & Shatz, C. J. (1996, November). Synaptic activity and the construction of cortical circuits. *Science*, *274*, 1133–1138. doi: 10.1126/science.274.5290.1133

Kawaguchi, Y., Karube, F., & Kubota, Y. (2006, May). Dendritic branch typing and spine expression patterns in cortical nonpyramidal cells. *Cerebral cortex*, *16*, 696–711. doi: 10.1093/cercor/bhj015

Keller, A., & White, E. L. (1987, August). Synaptic organization of gabaergic neurons in the mouse smi cortex. *The Journal of comparative neurology*, *262*, 1–12. doi: 10.1002/cne.902620102

Kepecs, A., & Fishell, G. (2014). Interneuron cell types are fit to function. *Nature*, *505*(7483), 318–326.

Knott, G. W., Quairiaux, C., Genoud, C., & Welker, E. (2002, April). Formation of dendritic spines with gabaergic synapses induced by whisker stimulation

in adult mice. *Neuron*, *34*, 265–273. doi: 10.1016/s0896-6273(02)00663-3

Koelbl, C., Helmstaedter, M., Lübke, J., & Feldmeyer, D. (2013). A barrel-related interneuron in layer 4 of rat somatosensory cortex with a high intrabarrel connectivity. *Cerebral Cortex*, bht263.

Krieger, P., Kuner, T., & Sakmann, B. (2007, October). Synaptic connections between layer 5b pyramidal neurons in mouse somatosensory cortex are independent of apical dendrite bundling. *Journal of neuroscience*, *27*, 11473–11482. doi: 10.1523/JNEUROSCI.1182-07.2007

Kriegstein, A. R., & Noctor, S. C. (2004). Patterns of neuronal migration in the embryonic cortex. *Trends in neurosciences*, *27*(7), 392–399.

Krishnaswamy, A., Yamagata, M., Duan, X., Hong, Y. K., & Sanes, J. R. (2015). Sidekick 2 directs formation of a retinal circuit that detects differential motion. *Nature*, *524*(7566), 466–470.

Krubitzer, L., Campi, K. L., & Cooke, D. F. (2011). All rodents are not the same: a modern synthesis of cortical organization. *Brain, behavior and evolution*, *78*(1), 51–93.

Kubota, Y. (2014, June). Untangling gabaergic wiring in the cortical microcircuit. *Current opinion in neurobiology*, *26*, 7–14. doi: 10.1016/j.conb.2013.10.003

Kwan, K. Y., Šestan, N., & Anton, E. (2012). Transcriptional co-regulation of neuronal migration and laminar identity in the neocortex. *Development*, *139*(9), 1535–1546.

Landau, I. D., Egger, R., Dercksen, V. J., Oberlaender, M., & Sompolinsky, H. (2016, December). The impact of structural heterogeneity on excitation-inhibition balance in cortical networks. *Neuron*, *92*, 1106–1121. doi: 10.1016/j.neuron.2016.10.027

Lang, S., Dercksen, V. J., Sakmann, B., & Oberlaender, M. (2011). Simulation of signal flow in 3d reconstructions of an anatomically realistic neural network in rat vibrissal cortex. *Neural Networks*, *24*(9), 998–1011.

Langley, J. N. (1895, July). Note on regeneration of prae-ganglionic fibres of the sympathetic. *The Journal of physiology*, *18*, 280–284. doi: 10.1113/jphysiol.1895.sp000566

Lanuza, M. A., Tomàs, J., Garcia, N., Cilleros-Mañé, V., Just-Borràs, L., & Tomàs, M. (2018). Axonal competition and synapse elimination during neuromus-

cular junction development. *Current Opinion in Physiology*, *4*, 25–31.

Larkman, A. U. (1991, April). Dendritic morphology of pyramidal neurones of the visual cortex of the rat: Iii. spine distributions. *The Journal of comparative neurology*, *306*, 332–343. doi: 10.1002/cne.903060209

Lee, W.-C. A., Bonin, V., Reed, M., Graham, B. J., Hood, G., Glattfelder, K., & Reid, R. C. (2016, April). Anatomy and function of an excitatory network in the visual cortex. *Nature*, *532*, 370–374. doi: 10.1038/nature17192

Lefebvre, J. L., Sanes, J. R., & Kay, J. N. (2015). Development of dendritic form and function. *Annual review of cell and developmental biology*, *31*, 741–777.

Lefort, S., Tomm, C., Floyd Sarria, J.-C., & Petersen, C. C. H. (2009, January). The excitatory neuronal network of the c2 barrel column in mouse primary somatosensory cortex. *Neuron*, *61*, 301–316. doi: 10.1016/j.neuron.2008.12.020

Li, W.-C., Cooke, T., Sautois, B., Soffe, S. R., Borisyuk, R., & Roberts, A. (2007). Axon and dendrite geography predict the specificity of synaptic connections in a functioning spinal cord network. *Neural Development*, *2*(1), 17.

Li, W.-C., Soffe, S. R., & Roberts, A. (2004). A direct comparison of whole cell patch and sharp electrodes by simultaneous recording from single spinal neurons in frog tadpoles. *Journal of neurophysiology*, *92*(1), 380–386.

Linneweber, G. A., Andriatsilavo, M., Dutta, S. B., Bengochea, M., Hellbruegge, L., Liu, G., ... others (2020). A neurodevelopmental origin of behavioral individuality in the drosophila visual system. *Science*, *367*(6482), 1112–1119.

Lübke, J., Roth, A., Feldmeyer, D., & Sakmann, B. (2003, October). Morphometric analysis of the columnar innervation domain of neurons connecting layer 4 and layer 2/3 of juvenile rat barrel cortex. *Cerebral cortex*, *13*, 1051–1063. doi: 10.1093/cercor/13.10.1051

Macke, J. H., Opper, M., & Bethge, M. (2011). Common input explains higher-order correlations and entropy in a simple model of neural population activity. *Physical Review Letters*, *106*(20), 208102.

Magee, J. C., & Grienberger, C. (2020, February). Synaptic plasticity forms and functions. *Annual review of neuroscience*. doi: 10.1146/annurev-neuro-090919-022842

Marik, S. A., Yamahachi, H., McManus, J. N., Szabo, G., & Gilbert, C. D. (2010). Axonal dynamics of excitatory and inhibitory neurons in somatosensory cortex. *PLoS biology*, *8*(6).

Markram, H. (2006). The blue brain project. *Nature Reviews Neuroscience*, *7*(2), 153–160.

Markram, H., Lübke, J., Frotscher, M., Roth, A., & Sakmann, B. (1997, April). Physiology and anatomy of synaptic connections between thick tufted pyramidal neurones in the developing rat neocortex. *The Journal of physiology*, *500 (Pt 2)*, 409–440. doi: 10.1113/jphysiol.1997.sp022031

Markram, H., Muller, E., Ramaswamy, S., Reimann, M. W., Abdellah, M., Sanchez, C. A., ... others (2015). Reconstruction and simulation of neocortical microcircuitry. *Cell*, *163*(2), 456–492.

Markram, H., Toledo-Rodriguez, M., Wang, Y., Gupta, A., Silberberg, G., & Wu, C. (2004). Interneurons of the neocortical inhibitory system. *Nature Reviews Neuroscience*, *5*(10), 793.

Mercer, A., West, D. C., Morris, O. T., Kirchhecker, S., Kerkhoff, J. E., & Thomson, A. M. (2005). Excitatory connections made by presynaptic cortico-cortical pyramidal cells in layer 6 of the neocortex. *Cerebral cortex*, *15*(10), 1485–1496.

Meyer, H. S., Egger, R., Guest, J. M., Foerster, R., Reissl, S., & Oberlaender, M. (2013, November). Cellular organization of cortical barrel columns is whisker-specific. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(47), 19113–8. doi: 10.1073/pnas.1312691110

Meyer, H. S., Schwarz, D., Wimmer, V. C., Schmitt, A. C., Kerr, J. N. D., Sakmann, B., & Helmstaedter, M. (2011, October). Inhibitory interneurons in a cortical column form hot zones of inhibition in layers 2 and 5a. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(40), 16807–12. doi: 10.1073/pnas.1113648108

Meyer, H. S., Wimmer, V. C., Oberlaender, M., De Kock, C. P., Sakmann, B., & Helmstaedter, M. (2010). Number and laminar distribution of neurons in a thalamocortical projection column of rat vibrissal cortex. *Cerebral cortex*, *20*(10), 2277–2286.

Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., & Alon, U. (2002,

October). Network motifs: simple building blocks of complex networks. *Science*, *298*, 824–827. doi: 10.1126/science.298.5594.824

Miner, D., & Triesch, J. (2016, February). Plasticity-driven self-organization under topological constraints accounts for non-random features of cortical synaptic wiring. *PLoS computational biology*, *12*, e1004759. doi: 10.1371/journal.pcbi.1004759

Mishchenko, Y., Hu, T., Spacek, J., Mendenhall, J., Harris, K. M., & Chklovskii, D. B. (2010). Ultrastructural analysis of hippocampal neuropil from the connectomics perspective. *Neuron*, *67*(6), 1009–1020.

Motta, A., Berning, M., Boergens, K. M., Staffler, B., Beining, M., Loomba, S., ... Helmstaedter, M. (2019, November). Dense connectomic reconstruction in layer 4 of the somatosensory cortex. *Science*, *366*. doi: 10.1126/science.aay3134

Mountcastle, V. B. (1957). Modality and topographic properties of single neurons of cat's somatic sensory cortex. *Journal of neurophysiology*, *20*(4), 408–434.

Muñoz, W., Tremblay, R., & Rudy, B. (2014). Channelrhodopsin-assisted patching: in vivo electrophysiological recording of genetically and morphologically identified neurons throughout the brain. *Cell reports*, *9*(6), 2304.

Narayanan, R. T., Egger, R., Johnson, A. S., Mansvelder, H. D., Sakmann, B., de Kock, C. P. J., & Oberlaender, M. (2015, apr). Beyond columnar organization: Cell type- and target layer-specific principles of horizontal axon projection patterns in rat vibrissal cortex. *Cerebral Cortex*, *25*(11), 4450-4468. doi: 10.1093/cercor/bhv053

Narayanan, R. T., Udvary, D., & Oberlaender, M. (2017, oct). Cell type-specific structural organization of the six layers in rat barrel cortex. *Frontiers in neuroanatomy*, *11*. doi: 10.3389/fnana.2017.00091

Nedivi, E. (1999, October). Molecular analysis of developmental plasticity in neocortex. *Journal of neurobiology*, *41*, 135–147.

Niell, C. M., Meyer, M. P., & Smith, S. J. (2004). In vivo imaging of synapse formation on a growing dendritic arbor. *Nature neuroscience*, *7*(3), 254–260.

Nigro, M. J., Hashikawa-Yamasaki, Y., & Rudy, B. (2018, February). Diversity and connectivity of layer 5 somatostatin-expressing interneurons in the mouse barrel cortex. *Journal of neuroscience*, *38*, 1622–1633. doi:

10.1523/JNEUROSCI.2415-17.2017

Nolte, M., Gal, E., Markram, H., & Reimann, M. W. (2020). Impact of higher order network structure on emergent cortical activity. *Network Neuroscience*, *4*(1), 292–314.

Oberlaender, M., Boudewijns, Z. S., Kleele, T., Mansvelder, H. D., Sakmann, B., & de Kock, C. P. (2011). Three-dimensional axon morphologies of individual layer 5 neurons indicate cell type-specific intracortical pathways for whisker motion and touch. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(10), 4188–4193.

Oberlaender, M., de Kock, C. P., Bruno, R. M., Ramirez, A., Meyer, H. S., Dercksen, V. J., ... Sakmann, B. (2012). Cell type–specific three-dimensional structure of thalamocortical circuits in a column of rat vibrissal cortex. *Cerebral Cortex*, *22*(10), 2375–2391.

Oberlaender, M., Ramirez, A., & Bruno, R. M. (2012, May). Sensory experience restructures thalamocortical axons during adulthood. *Neuron*, *74*, 648–655. doi: 10.1016/j.neuron.2012.03.022

Otsuka, T., & Kawaguchi, Y. (2009, August). Cortical inhibitory cell types differentially form intralaminar and interlaminar subnetworks with excitatory neurons. *Journal of neuroscience*, *29*, 10533–10540. doi: 10.1523/JNEUROSCI.2219-09.2009

Packer, A. M., McConnell, D. J., Fino, E., & Yuste, R. (2013, December). Axo-dendritic overlap and laminar projection can explain interneuron connectivity to pyramidal cells. *Cerebral cortex*, *23*, 2790–2802. doi: 10.1093/cercor/bhs210

Packer, A. M., & Yuste, R. (2011, September). Dense, unspecific connectivity of neocortical parvalbumin-positive interneurons: a canonical microcircuit for inhibition? *Journal of neuroscience*, *31*, 13260–13271. doi: 10.1523/JNEUROSCI.3131-11.2011

Pala, A., & Petersen, C. C. H. (2015, January). In vivo measurement of cell-type-specific synaptic connectivity and synaptic transmission in layer 2/3 mouse barrel cortex. *Neuron*, *85*, 68–75. doi: 10.1016/j.neuron.2014.11.025

Perin, R., Berger, T. K., & Markram, H. (2011, March). A synaptic organizing principle for cortical neuronal groups. *Proceedings of the National Academy*

of Sciences of the United States of America, *108*, 5419–5424. doi: 10.1073/pnas.1016051108

Peters, A., & Feldman, M. L. (1976, February). The projection of the lateral geniculate nucleus to area 17 of the rat cerebral cortex. i. general description. *Journal of neurocytology*, *5*, 63–84. doi: 10.1007/bf01176183

Petersen, C. C. (2007). The functional organization of the barrel cortex. *Neuron*, *56*(2), 339–355.

Petersen, C. C., & Sakmann, B. (2000, October). The excitatory neuronal network of rat layer 4 barrel cortex. *Journal of neuroscience*, *20*, 7579–7586.

Petersen, C. C. H. (2019, September). Sensorimotor processing in the rodent barrel cortex. *Nature Reviews Neuroscience*, *20*, 533–546. doi: 10.1038/s41583-019-0200-y

Petreanu, L., Mao, T., Sternson, S. M., & Svoboda, K. (2009, February). The subcellular organization of neocortical excitatory connections. *Nature*, *457*, 1142–1145. doi: 10.1038/nature07709

Porter, J. T., Johnson, C. K., & Agmon, A. (2001, April). Diverse types of interneurons generate thalamus-evoked feedforward inhibition in the mouse barrel cortex. *Journal of neuroscience*, *21*, 2699–2710.

Qi, G., & Feldmeyer, D. (2016, April). Dendritic target region-specific formation of synapses between excitatory layer 4 neurons and layer 6 pyramidal cells. *Cerebral cortex*, *26*, 1569–1579. doi: 10.1093/cercor/bhu334

Qiao, Q., Ma, L., Li, W., Tsai, J.-W., Yang, G., & Gan, W.-B. (2016, March). Long-term stability of axonal boutons in the mouse barrel cortex. *Developmental neurobiology*, *76*, 252–261. doi: 10.1002/dneu.22311

Quattrocolo, G., Fishell, G., & Petros, T. J. (2017). Heterotopic transplantations reveal environmental influences on interneuron diversity and maturation. *Cell reports*, *21*(3), 721–731.

Rakic, P. (2009). Evolution of the neocortex: a perspective from developmental biology. *Nature Reviews Neuroscience*, *10*(10), 724–735.

Rash, B. G., & Grove, E. A. (2006). Area and layer patterning in the developing cerebral cortex. *Current opinion in neurobiology*, *16*(1), 25–34.

Ravid Tannenbaum, N., & Burak, Y. (2016, August). Shaping neural circuits by high order synaptic interactions. *PLoS computational biology*, *12*, e1005056.

doi: 10.1371/journal.pcbi.1005056

Rees, C. L., Moradi, K., & Ascoli, G. A. (2017). Weighing the evidence in peters rule: does neuronal morphology predict connectivity? *Trends in neurosciences*, *40*(2), 63–71.

Reimann, M. W., Gevaert, M., Shi, Y., Lu, H., Markram, H., & Muller, E. (2019). A null model of the mouse whole-neocortex micro-connectome. *Nature communications*, *10*(1), 1–16.

Rieubland, S., Roth, A., & Häusser, M. (2014, February). Structured connectivity in cerebellar inhibitory networks. *Neuron*, *81*, 913–929. doi: 10.1016/j .neuron.2013.12.029

Rodriguez-Moreno, J., Rollenhagen, A., Arlandis, J., Santuy, A., Merchan-Pérez, A., DeFelipe, J., ... Clasca, F. (2018). Quantitative 3d ultrastructure of thalamocortical synapses from the "lemniscal" ventral posteromedial nucleus in mouse barrel cortex. *Cerebral Cortex*, *28*(9), 3159–3175.

Rojas-Piloni, G., Guest, J. M., Egger, R., Johnson, A. S., Sakmann, B., & Oberlaender, M. (2017, October). Relationships between structure, in vivo function and long-range axonal target of cortical pyramidal tract neurons. *Nature communications*, *8*, 870. doi: 10.1038/s41467-017-00971-0

Roux, L., & Buzsáki, G. (2015, January). Tasks for inhibitory interneurons in intact brain circuits. *Neuropharmacology*, *88*, 10–23. doi: 10.1016/ j.neuropharm.2014.09.011

Salami, M., Itami, C., Tsumoto, T., & Kimura, F. (2003). Change of conduction velocity by regional myelination yields constant latency irrespective of distance between thalamus and cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(10), 6174–6179.

Sanes, J. R., & Lichtman, J. W. (1999). Development of the vertebrate neuromuscular junction. *Annual review of neuroscience*, *22*(1), 389–442.

Sanes, J. R., & Yamagata, M. (2009). Many paths to synaptic specificity. *Annual Review of Cell and Developmental*, *25*, 161–195.

Santuy, A., Rodriguez, J. R., DeFelipe, J., & Merchan-Perez, A. (2018, January). Volume electron microscopy of the distribution of synapses in the neuropil of the juvenile rat somatosensory cortex. *Brain structure & function*, *223*, 77–90. doi: 10.1007/s00429-017-1470-7

Scala, F., Kobak, D., Shan, S., Bernaerts, Y., Laturnus, S., Cadwell, C. R., ... others (2019). Layer 4 of mouse neocortex differs in cell types and circuit organization between sensory areas. *Nature communications*, *10*(1), 1–12.

Schoonover, C. E., Tapia, J.-C., Schilling, V. C., Wimmer, V., Blazeski, R., Zhang, W., ... Bruno, R. M. (2014, May). Comparative strength and dendritic organization of thalamocortical and corticocortical synapses onto excitatory layer 4 neurons. *Journal of neuroscience*, *34*, 6746–6758. doi: 10.1523/JNEUROSCI.0305-14.2014

Schröter, M., Paulsen, O., & Bullmore, E. T. (2017, March). Micro-connectomics: probing the organization of neuronal networks at the cellular scale. *Nature reviews. Neuroscience*, *18*, 131–146. doi: 10.1038/nrn.2016.182

Sengupta, P. (2013). The laboratory rat: relating its age with human's. *International journal of preventive medicine*, *4*(6), 624.

Shen, K., & Scheiffele, P. (2010). Genetics and cell biology of building specific synaptic connectivity. *Annual review of neuroscience*, *33*, 473–507. doi: 10.1146/annurev.neuro.051508.135302

Shigematsu, N., Ueta, Y., Mohamed, A. A., Hatada, S., Fukuda, T., Kubota, Y., & Kawaguchi, Y. (2016, June). Selective thalamic innervation of rat frontal cortical neurons. *Cerebral cortex*, *26*, 2689–2704. doi: 10.1093/cercor/bhv124

Shimogori, T., & Grove, E. A. (2005). Fibroblast growth factor 8 regulates neocortical guidance of area-specific thalamic innervation. *Journal of neuroscience*, *25*(28), 6550–6560.

Silberberg, G., & Markram, H. (2007, March). Disynaptic inhibition between neocortical pyramidal cells mediated by martinotti cells. *Neuron*, *53*, 735–746. doi: 10.1016/j.neuron.2007.02.012

Simi, A., & Studer, M. (2018, December). Developmental genetic programs and activity-dependent mechanisms instruct neocortical area mapping. *Current opinion in neurobiology*, *53*, 96–102. doi: 10.1016/j.conb.2018.06.007

Slabik, D. A. (2018). *Putative synaptische verbindungen zwischen kortikalen nervenzellen in vivo* (Unpublished doctoral dissertation). Eberhard-Karls-Universität Tübingen, Germany.

Smith, M. A., & Kohn, A. (2008, November). Spatial and temporal scales of

neuronal correlation in primary visual cortex. *Journal of neuroscience*, *28*, 12591–12603. doi: 10.1523/JNEUROSCI.2929-08.2008

Somogyi, P. (1977). A specific 'axo-axonal' interneuron in the visual cortex of the rat. *Brain Res*, *136*(2), 345–350.

Somogyi, P., Tamás, G., Lujan, R., & Buhl, E. H. (1998, May). Salient features of synaptic organisation in the cerebral cortex. *Brain research reviews*, *26*, 113–135. doi: 10.1016/s0165-0173(97)00061-1

Song, S., & Abbott, L. F. (2001, October). Cortical development and remapping through spike timing-dependent plasticity. *Neuron*, *32*, 339–350. doi: 10 .1016/s0896-6273(01)00451-2

Song, S., Sjöström, P. J., Reigl, M., Nelson, S., & Chklovskii, D. B. (2005, March). Highly nonrandom features of synaptic connectivity in local cortical circuits. *PLoS biology*, *3*, e68. doi: 10.1371/journal.pbio.0030068

Sperry, R. W. (1963). Chemoaffinity in the orderly growth of nerve fiber patterns and connections. *Proceedings of the National Academy of Sciences of the United States of America*, *50*(4), 703–710.

Staiger, J. F., Möck, M., Proenneke, A., & Witte, M. (2015). What types of neocortical gabaergic neurons do really exist? *e-Neuroforum*, *6*(2), 49–56.

Staiger, J. F., Zilles, K., & Freund, T. F. (1996, November). Distribution of gabaergic elements postsynaptic to ventroposteromedial thalamic projections in layer iv of rat barrel cortex. *The European journal of neuroscience*, *8*, 2273–2285. doi: 10.1111/j.1460-9568.1996.tb01191.x

Stalling, D., Westerhoff, M., & Hege, H.-C. (2005). Amira: A highly interactive system for visual data analysis. *The visualization handbook*.

Stepanyants, A., Martinez, L. M., Ferecskó, A. S., & Kisvárday, Z. F. (2009). The fractions of short-and long-range connections in the visual cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(9), 3555–3560.

Stoeckli, E. T. (2018). Understanding axon guidance: are we nearly there yet? *Development*, *145*(10), dev151415.

Stüttgen, M. C., & Schwarz, C. (2018). Barrel cortex: What is it good for? *Neuroscience*, *368*, 3–16.

Südhof, T. C. (2018, October). Towards an understanding of synapse formation.

Neuron, *100*, 276–293. doi: 10.1016/j.neuron.2018.09.040

Sun, Q.-Q., Huguenard, J. R., & Prince, D. A. (2006, January). Barrel cortex microcircuits: thalamocortical feedforward inhibition in spiny stellate cells is mediated by a small number of fast-spiking interneurons. *Journal of neuroscience*, *26*, 1219–1230. doi: 10.1523/JNEUROSCI.4727-04.2006

Takahashi, N. (2019, April). Synaptic topography - converging connections and emerging function. *Neuroscience research*, *141*, 29–35. doi: 10.1016/j.neures.2018.11.001

Tarusawa, E., Sanbo, M., Okayama, A., Miyashita, T., Kitsukawa, T., Hirayama, T., ... others (2016). Establishment of high reciprocal connectivity between clonal cortical neurons is regulated by the dnmt3b dna methyltransferase and clustered protocadherins. *BMC biology*, *14*(1), 103.

Thomson, A. M., & Lamy, C. (2007). Functional maps of neocortical local circuitry. *Frontiers in neuroscience*, *1*, 2.

Thomson, A. M., West, D. C., Hahn, J., & Deuchars, J. (1996, October). Single axon ipsps elicited in pyramidal cells by three classes of interneurones in slices of rat neocortex. *The Journal of physiology*, *496 (Pt 1)*, 81–102. doi: 10.1113/jphysiol.1996.sp021667

Thomson, A. M., West, D. C., Wang, Y., & Bannister, A. P. (2002, September). Synaptic connections and small circuits involving excitatory and inhibitory neurons in layers 2-5 of adult rat and cat neocortex: triple intracellular recordings and biocytin labelling in vitro. *Cerebral cortex*, *12*, 936–953. doi: 10.1093/cercor/12.9.936

Trachtenberg, J. T., Chen, B. E., Knott, G. W., Feng, G., Sanes, J. R., Welker, E., & Svoboda, K. (2002). Long-term in vivo imaging of experience-dependent synaptic plasticity in adult cortex. *Nature*, *420*(6917), 788–794.

Udvary, D. (2013). *Generation of 3d inhibitory dendrite/axon morphologies from truncated reconstructions in vitro.* (Unpublished laboratory report). Graduate School of Neural Information Processing, Eberhard-Karls-Universität Tübingen, Germany.

Udvary, D. (2014). *Types of axonal inhibitory interneurons in rat vibrissal cortex* (Unpublished master's thesis). Graduate School of Neural Information Processing, Eberhard-Karls-Universität Tübingen, Germany.

Udvary, D., Dercksen, V. J., Harth, P., Hege, H.-C., de Kock, C. P. J., Sakmann, B., & Oberlaender, M. (2020). *A theory for the origins of structural scaffoldings in synaptic wiring diagrams of the neocortex.* (Manuscript in preparation)

Udvary, D., Dercksen, V. J., Harth, P., Hege, H.-C., & Oberlaender, M. (2020). *Barrel cortex in silico.* (Manuscript in preparation)

Udvary, D., Egger, R., Meyer, H. S., Guest, J. M., Sakman, B., Feldmeyer, D., & Oberlaender, M. (2016). *Transfer of thalamocortical input by projection type-specific inhibitory axon pathways in a column of rat vibrissal cortex.* (Unpublished manuscript)

Valdes-Aleman, J., Fetter, R. D., Sales, E. C., Doe, C. Q., Landgraf, M., Cardona, A., & Zlatic, M. (2019). Synaptic specificity is collectively determined by partner identity, location and activity. *bioRxiv*. doi: 10.1101/697763

Van der Loos, H., & Woolsey, T. A. (1973). Somatosensory cortex: structural alterations following early injury to sense organs. *Science*, *179*(4071), 395–398.

van der Zee, E. A. (2015, March). Synapses, spines and kinases in mammalian learning and memory, and the impact of aging. *Neuroscience and biobehavioral reviews*, *50*, 77–85. doi: 10.1016/j.neubiorev.2014.06.012

van Kerkoerle, T., Marik, S. A., Meyer Zum Alten Borgloh, S., & Gilbert, C. D. (2018, October). Axonal plasticity associated with perceptual learning in adult macaque primary visual cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *115*, 10464–10469. doi: 10.1073/pnas.1812932115

Vaughn, J. E., Barber, R. P., & Sims, T. J. (1988). Dendritic development and preferential growth into synaptogenic fields: A quantitative study of golgi-impregnated spinal motor neurons. *Synapse*, *2*(1), 69–78.

Walker, F., Möck, M., Feyerabend, M., Guy, J., Wagener, R. J., Schubert, D., ... Witte, M. (2016, November). Parvalbumin- and vasoactive intestinal polypeptide-expressing neocortical interneurons impose differential inhibition on martinotti cells. *Nature communications*, *7*, 13664. doi: 10.1038/ncomms13664

Wamsley, B., & Fishell, G. (2017). Genetic and activity-dependent mechanisms

underlying interneuron diversity. *Nature Reviews Neuroscience*, *18*(5), 299.

Wang, Y., Gupta, A., Toledo-Rodriguez, M., Wu, C. Z., & Markram, H. (2002, April). Anatomical, physiological, molecular and circuit properties of nest basket cells in the developing somatosensory cortex. *Cerebral cortex*, *12*, 395–410. doi: 10.1093/cercor/12.4.395

Ward, A., Hong, W., Favaloro, V., & Luo, L. (2015). Toll receptors instruct axon and dendrite targeting and participate in synaptic partner matching in a drosophila olfactory circuit. *Neuron*, *85*(5), 1013–1028.

Watts, D., & Strogatz, S. (1998). An undirected, unweighted network representing the topology of the western states power grid of the united states. *Nature*, *393*, 440–442.

White, E. L., Benshalom, G., & Hersch, S. M. (1984, November). Thalamocortical and other synapses involving nonspiny multipolar cells of mouse smi cortex. *The Journal of comparative neurology*, *229*, 311–320. doi: 10.1002/cne.902290303

White, E. L., & Rock, M. P. (1981, January). A comparison of thalamocortical and other synaptic inputs to dendrites of two non-spiny neurons in a single barrel of mouse smi cortex. *The Journal of comparative neurology*, *195*, 265–277. doi: 10.1002/cne.901950207

Williams, M. E., de Wit, J., & Ghosh, A. (2010, October). Molecular mechanisms of synaptic specificity in developing neural circuits. *Neuron*, *68*, 9–18. doi: 10.1016/j.neuron.2010.09.007

Williams, S. R., & Mitchell, S. J. (2008). Direct measurement of somatic voltage clamp errors in central neurons. *Nature neuroscience*, *11*(7), 790.

Willshaw, D. J. (1981, May). The establishment and the subsequent elimination of polyneuronal innervation of developing muscle: theoretical considerations. *Proceedings of the Royal Society of London. Series B, Biological sciences*, *212*, 233–252. doi: 10.1098/rspb.1981.0036

Xue, M., Atallah, B. V., & Scanziani, M. (2014, July). Equalizing excitation-inhibition ratios across visual cortical neurons. *Nature*, *511*, 596–600. doi: 10.1038/nature13321

Yang, J.-W., Kilb, W., Kirischuk, S., Unichenko, P., Stüttgen, M. C., & Luhmann, H. J. (2018). Development of the whisker-to-barrel cortex system. *Current*

opinion in neurobiology, *53*, 29–34.

Yogev, S., & Shen, K. (2014). Cellular and molecular mechanisms of synaptic specificity. *Annual review of cell and developmental biology*, *30*, 417–437.

Yoshimura, Y., & Callaway, E. M. (2005, November). Fine-scale specificity of cortical networks depends on inhibitory cell type and connectivity. *Nature neuroscience*, *8*, 1552–1559. doi: 10.1038/nn1565

Yoshimura, Y., Dantzker, J. L. M., & Callaway, E. M. (2005, February). Excitatory cortical neurons form fine-scale functional networks. *Nature*, *433*, 868–873. doi: 10.1038/nature03252

Yu, Y.-C., Bultje, R. S., Wang, X., & Shi, S.-H. (2009, March). Specific synapses develop preferentially among sister excitatory neurons in the neocortex. *Nature*, *458*, 501–504. doi: 10.1038/nature07722

Zador, A. M. (2019, August). A critique of pure learning and what artificial neural networks can learn from animal brains. *Nature communications*, *10*, 3770. doi: 10.1038/s41467-019-11786-6

Zhang, D., Zhang, C., & Stepanyants, A. (2019). Robust associative learning is sufficient to explain the structural and dynamical properties of local cortical circuits. *Journal of Neuroscience*, *39*(35), 6888–6904.

Zhang, Z.-W., & Deschênes, M. (1997). Intracortical axonal projections of lamina vi cells of the primary somatosensory cortex in the rat: a single-cell labeling study. *Journal of neuroscience*, *17*(16), 6365–6379.

Zhou, X., Rickmann, M., Hafner, G., & Staiger, J. F. (2017, November). Subcellular targeting of vip boutons in mouse barrel cortex is layer-dependent and not restricted to interneurons. *Cerebral cortex*, *27*, 5353–5368. doi: 10.1093/cercor/bhx220

Zipursky, S. L., & Sanes, J. R. (2010, October). Chemoaffinity revisited: dscams, protocadherins, and neural circuit assembly. *Cell*, *143*, 343–353. doi: 10.1016/j.cell.2010.10.009

Appendix A

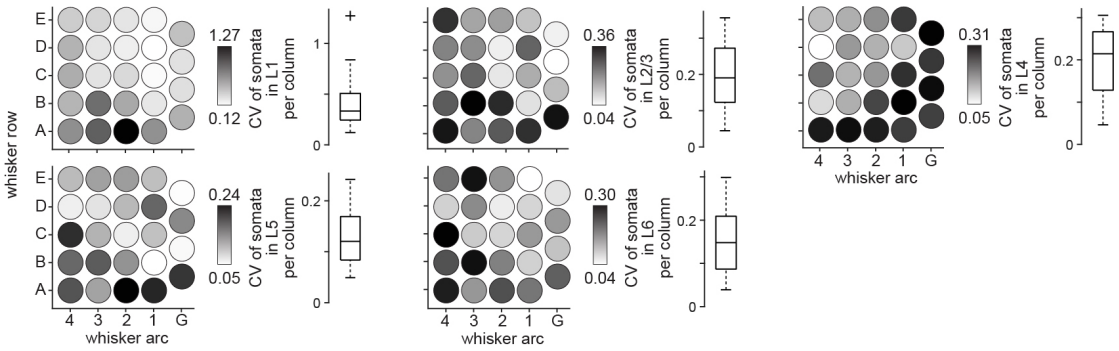

Figure A1. Interanimal variability of somata per barrel column and layer in vS1. Panel represents reanalysis of data reported in Meyer et al. (2013). Figure adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020)

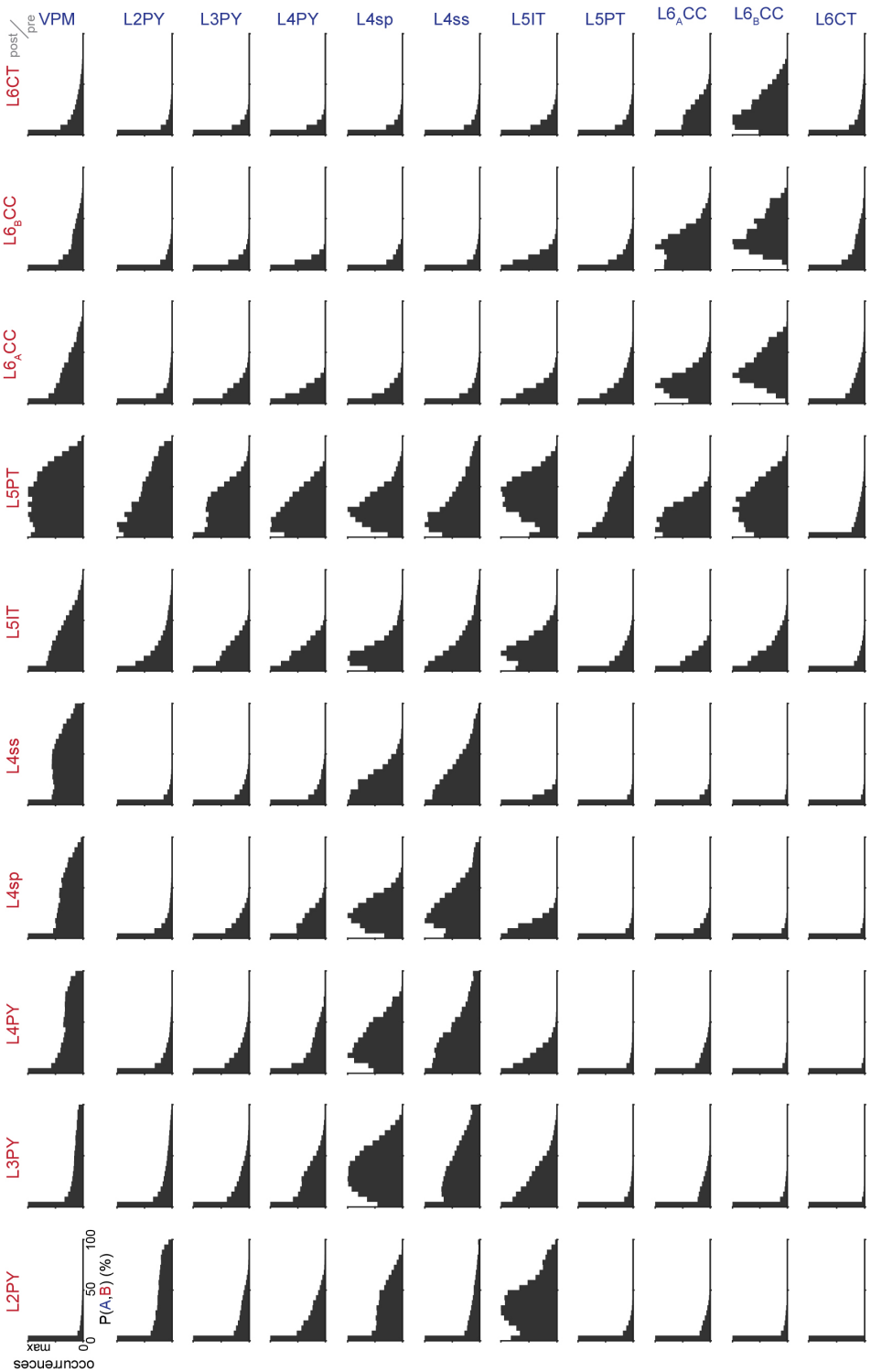

Figure A2. Cell type-specific connection probability distributions in C2 barrel column. Y-axis normalized to maximum number of occurrences, x-axis ranges from 0% to 100% connection probability, bin size of 5%

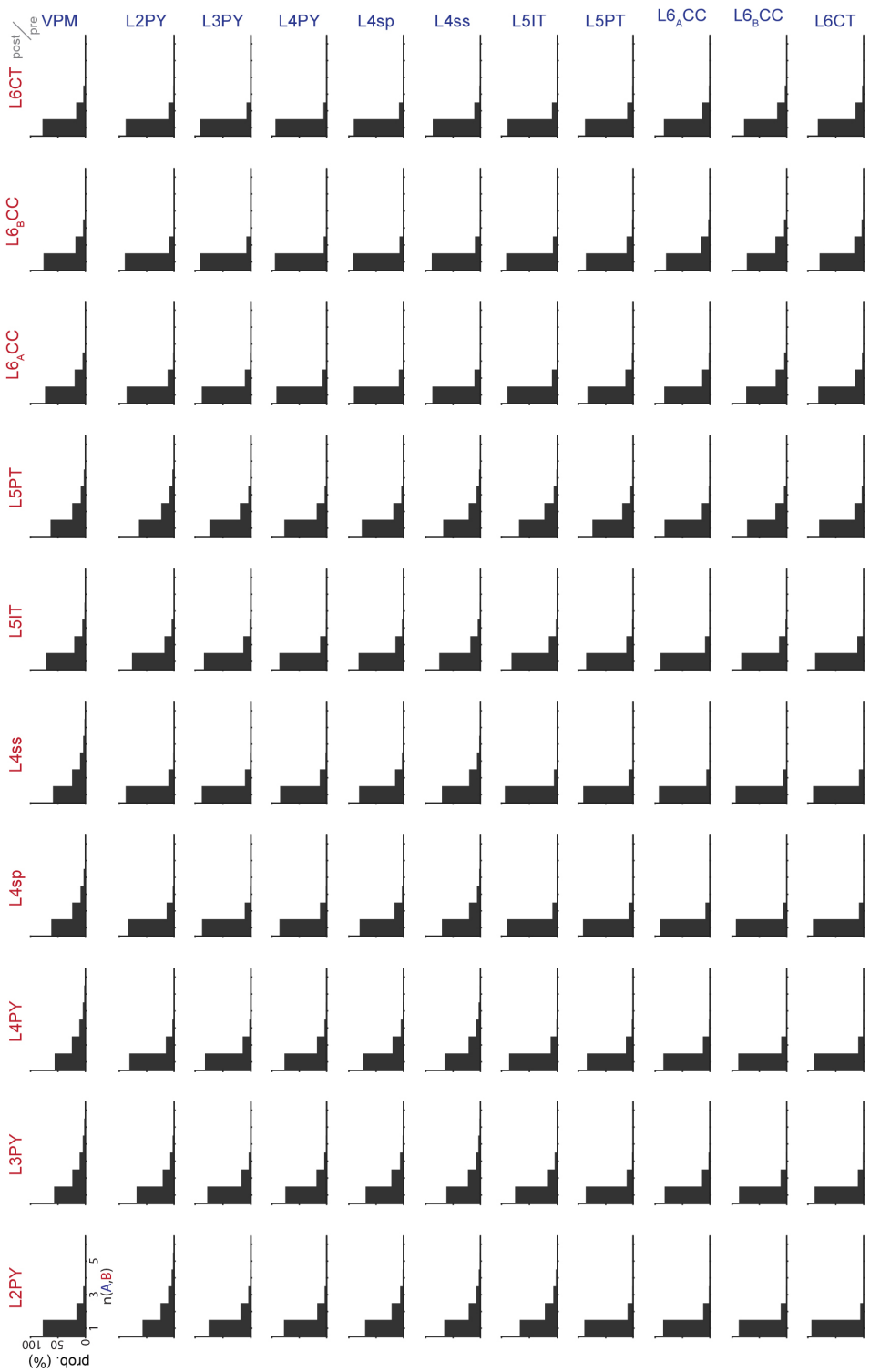

Figure A3. Cell type-specific distributions of synapses per connected neuron pair in C2 barrel column. Y-axis ranges from 0% to 100%, x-axis ranges from 1 to 5 synapses.

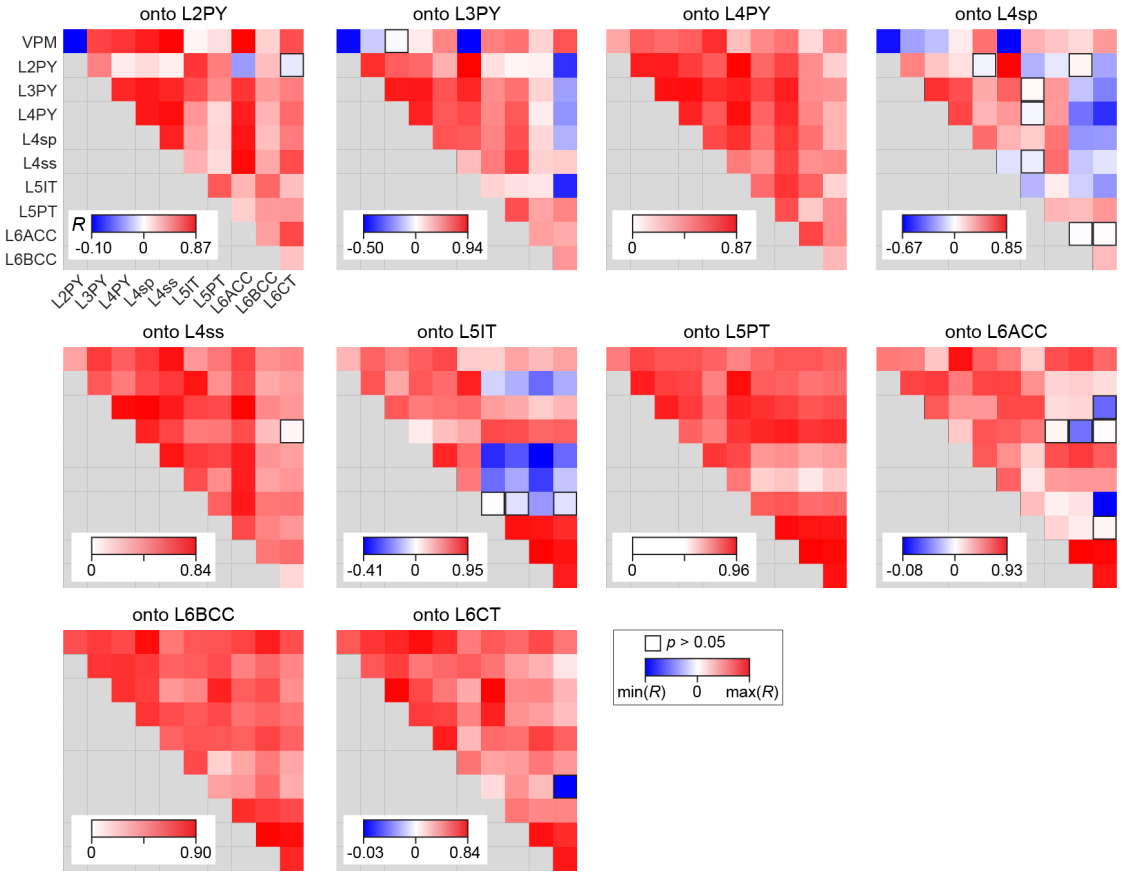

Figure A4. In-degree correlations between all possible cell type groupings. Each matrix represents one postsynaptic cell type in one barrel column and the corresponding in-degree correlations across its postsynaptic cell types. Negative R shaded in blue, positive R shaded in red, color map normalized to range of R plus $R = 0$ for each matrix, and non-significant correlations marked by black square ($p > 0.05$).

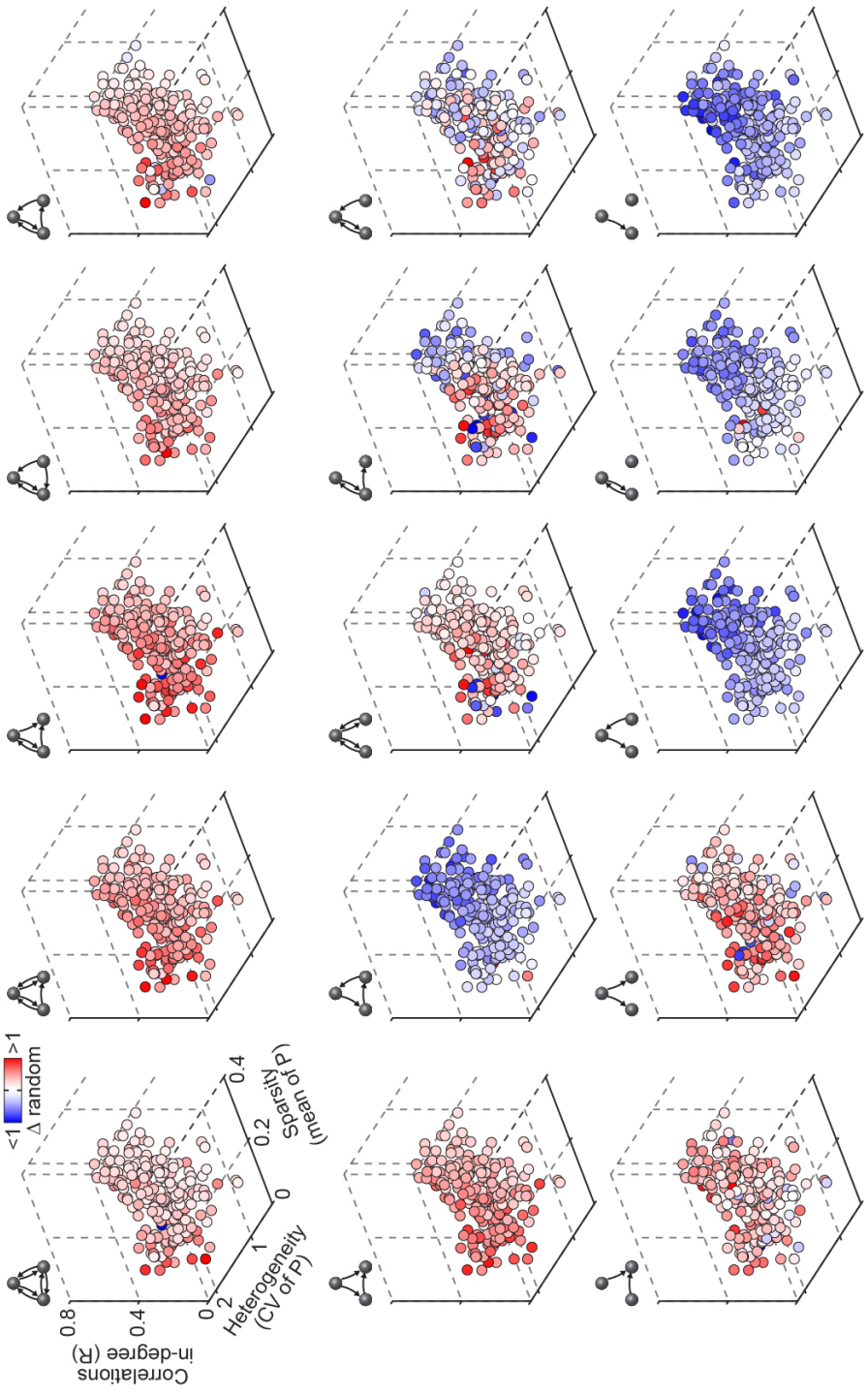

Figure A5. Wiring properties vs. triplet motif deviation. Deviation of each triplet motif as a function of its underlying average in-degree correlations, heterogeneity (CV of connection probabilities), and sparsity (mean of connection probabilities).

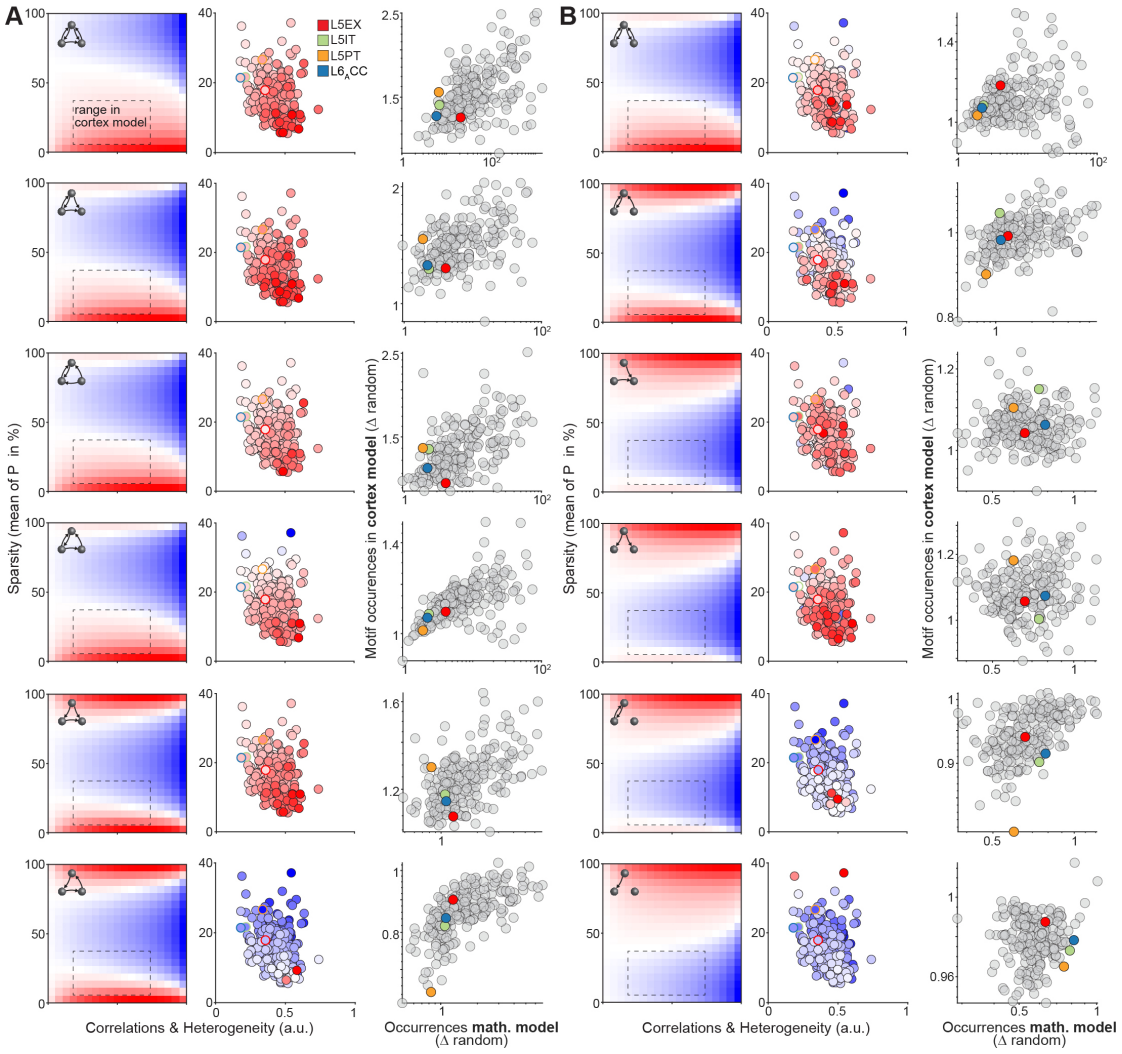

Figure A6. Wiring properties shape triplet motif deviations. Same as in Figure 3.31 but for remaining triplet motifs. (A) Left column: Relationship in mathematical model between sparsity, heterogeneity, and correlation, and the triplet motif deviation. Dashed box represents predicted range from vS1 model. Center column: Respective distribution in vS1 model for all cell type-specific groupings. Right column: Deviation of each triplet motif predicted by vS1 model vs. mathematical model. (B) same as in panel A but for remaining motifs. Figure adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020).

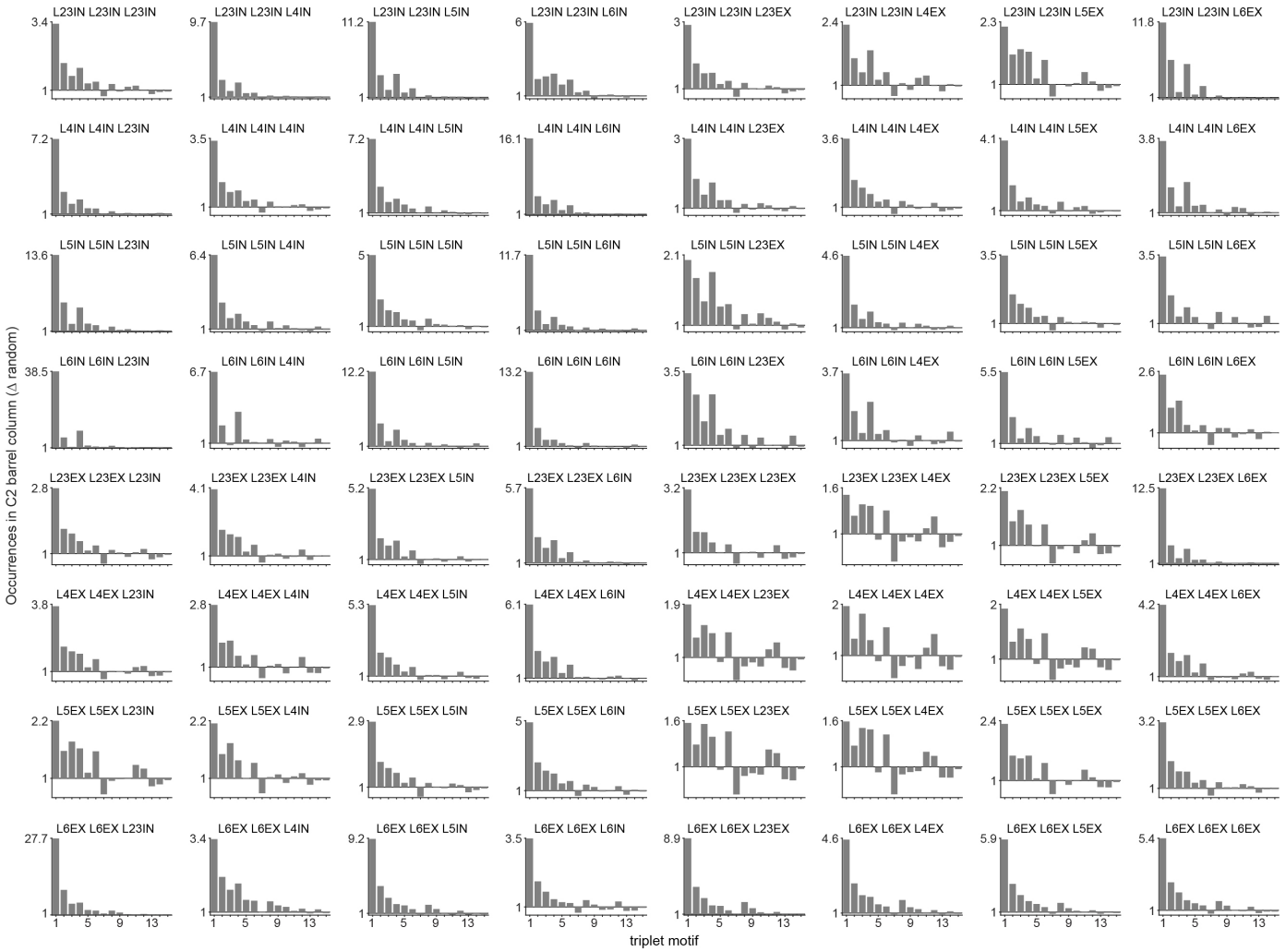

Figure A7. Layer-specific triplet motif spectra for excitatory and inhibitory neurons in C2 barrel column.

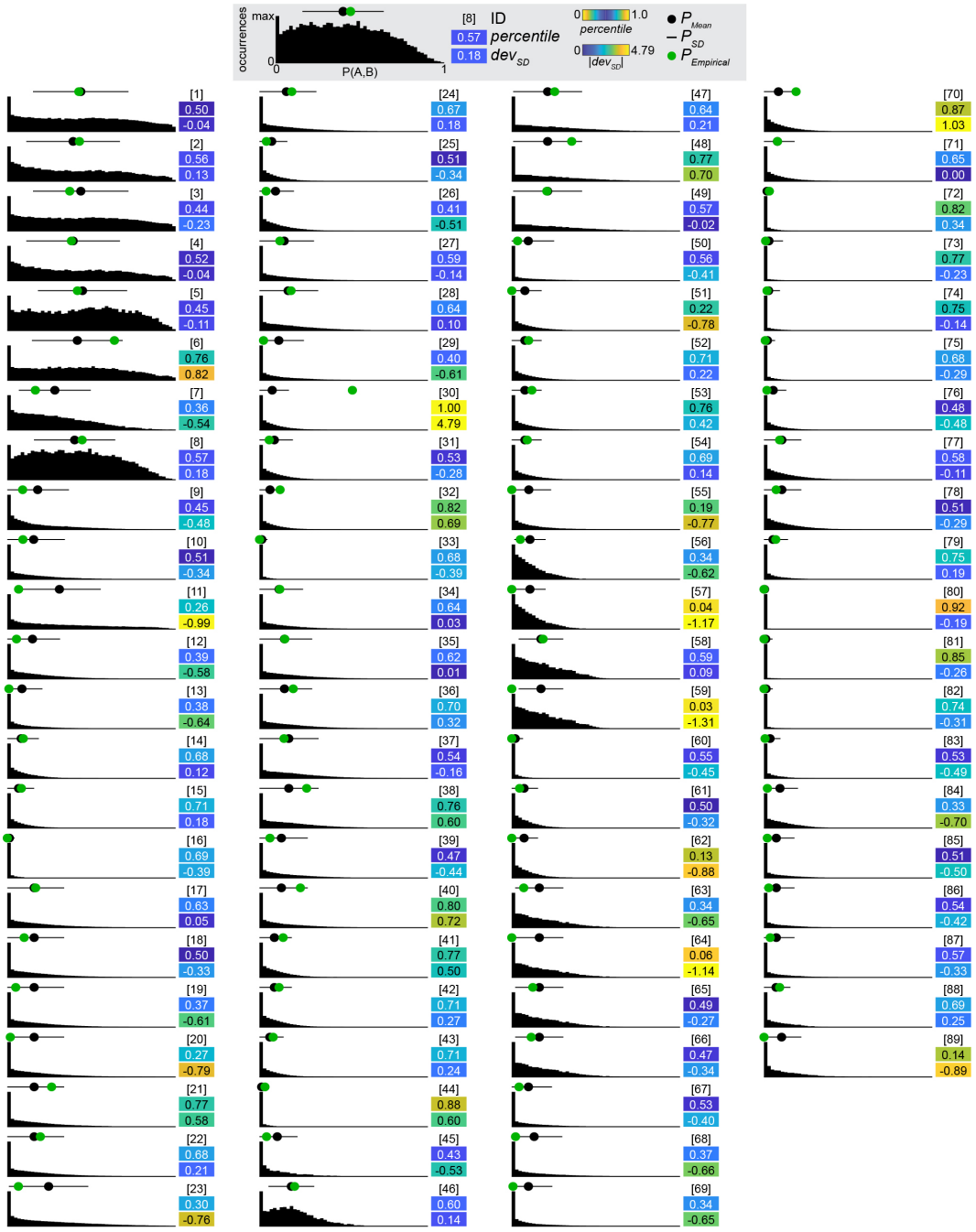

Figure A8. Comparison between 89 empirical and predicted connection probabilities between excitatory neurons. For experimental details, see respective ID in Table A10.

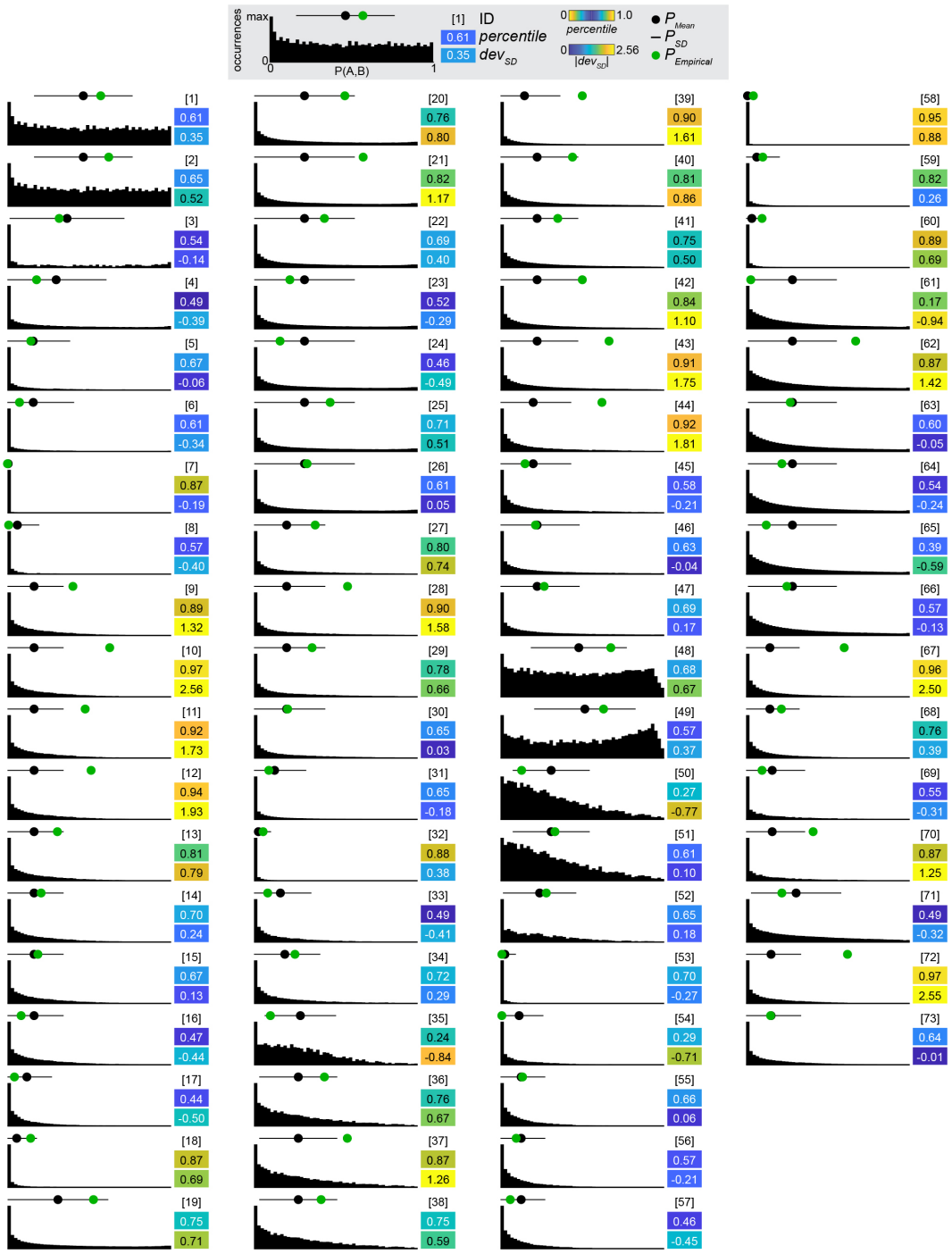

Figure A9. Comparison between 73 empirical and predicted connection probabilities involving inhibitory neurons. For experimental details, see respective ID in Table A11.

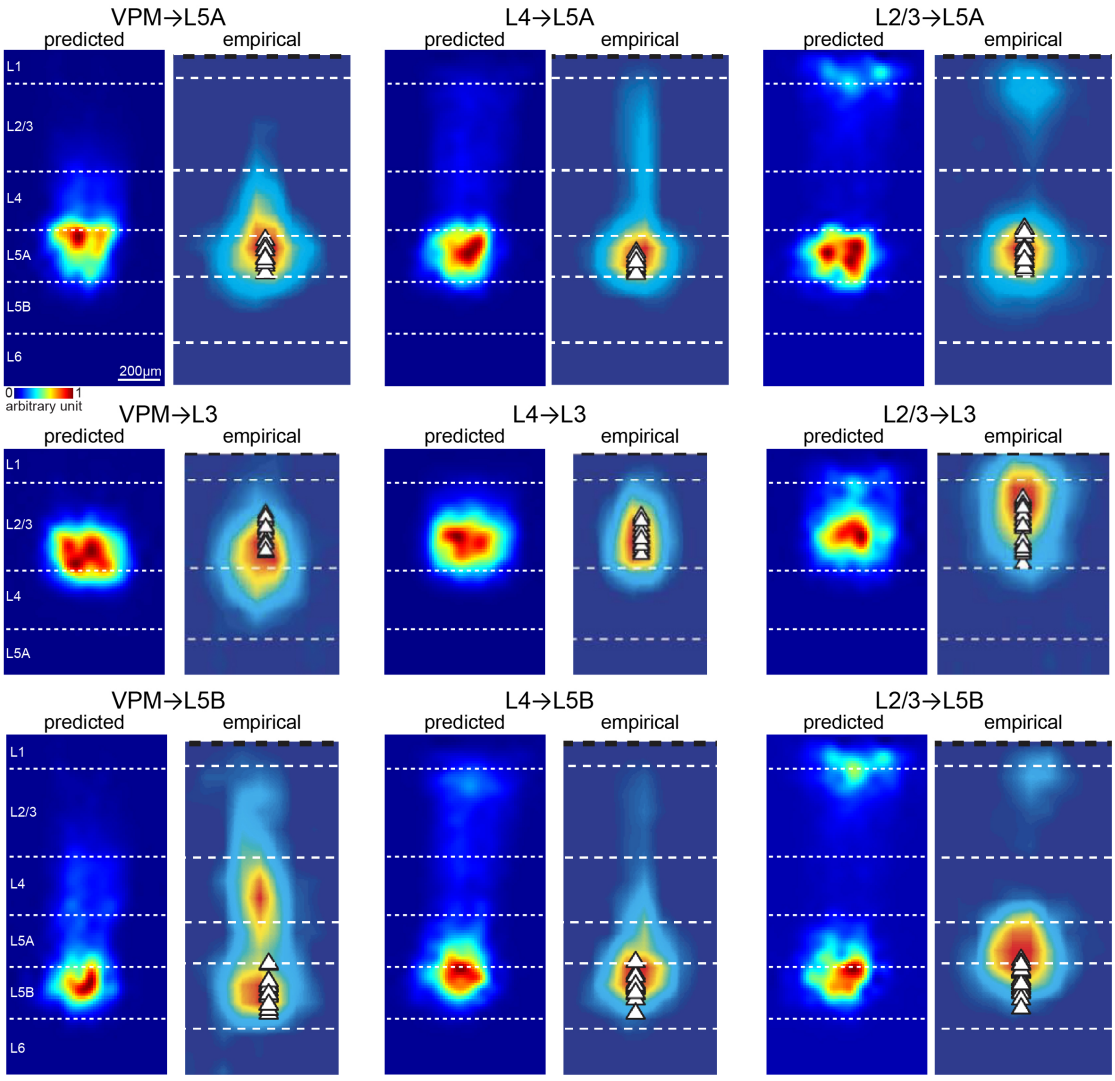

Figure A10. Predicted (left densities) vs. empirical (right densities; adapted from Petreanu et al., 2009) subcellular distribution of synapses between different neuron populations. Empirical densities represent amplitude of excitatory postsynaptic current evoked by photostimulation (Petreanu et al., 2009). Densities normalized to 1. The predicted densities represent DSC (i.e., average synapse) densities.

Table A1. Estimates of the amount of dendrites/axons within vS1, contributed by INs (excluding L1 INs) and excitatory neurons with somata located within an average C2 barrel column (i.e., soma column). Table adapted from Udvary et al. (2016).

Parameter	Local $n = 1,034$	SupraDesc $n = 150$	GranSym $n = 181$	GranAsc $n = 200$	InfraAsc $n = 421$	IN $n = 1,986$	EX $n = 14,266$
Dendrite length (m)	4.06	0.54	0.66	0.72	1.47	7.46	98.09
Dendrite volume (mm^3)	0.41	0.15	0.21	0.21	0.25	0.52	0.57
POSTs ($\times 10^6$)	6.52	0.77	1.21	1.23	3.17	12.97	127.90
Axon length (m)	48.68	7.21	9.34	15.92	29.91	111.06	859.02
Axon volume (mm^3)	0.62	0.19	0.51	0.52	0.68	0.87	1.45
Boutons ($\times 10^6$)	9.74	1.44	1.87	3.18	5.98	22.21	216.45

Table A2. Predicted connection probabilities vs. axodendritic overlap in 2D and 3D for different neuron groupings across all cell type combinations. Table adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020).

overlap in	intersomatic distance (µm)	R	n	slope	F	p	n
2D	-	0.91	110	0.89	132.9	5.87×10^{-29}	5995
	0 to 100	0.90	100	0.89	34.4	6.08×10^{-9}	4950
	100 to 200	0.94	100	0.89	163.6	8.31×10^{-35}	4950
	200 to 300	0.91	100	0.89	41.8	1.58×10^{-10}	4950
	0 to 200	0.93	100	0.88	155.3	3.07×10^{-33}	4950
	0 to 300	0.94	100	0.87	230.7	4.88×10^{-47}	4950
	100 to 300	0.94	100	0.88	143.5	5.50×10^{-31}	4950
3D	-	0.88	110	0.90	23.2	1.66×10^{-6}	5995
	0 to 100	0.94	100	1.00	0.0	9.84×10^{-1}	4950
	100 to 200	0.87	100	0.89	23.9	1.21×10^{-6}	4950
	200 to 300	0.79	100	0.81	46.1	1.90×10^{-11}	4950
	0 to 200	0.91	100	0.91	30.9	3.52×10^{-8}	4950
	0 to 300	0.89	100	0.88	40.1	3.65×10^{-10}	4950
	100 to 300	0.85	100	0.86	42.5	1.11×10^{-10}	4950

Intersomatic distance denotes range of lateral intersomatic distance of neurons pairs used for calculating P(A,B). Dash denotes no restriction in intersomatic distances (i.e., all neuron pairs located closest to the C2 barrel column were used). R: correlation coefficient between P(A,B) and overlap(A,B); slope: mean slope of linear regression models between ratios of all combinations of overlaps and of connection probabilities across 100 trials each with 1000 samples; F: minimum F test statistic for a linear hypothesis test on a slope of 1 across 100 trials; p: corresponding p value for F test. For details, see Method Section 2.3.1 on page 41.

Table A3. Absolute and relative (in brackets) number of neuron pairs connected by zero to four or more synapses in 50 µm cubes. Statistics across 512 cubes.

	neuron pairs connected by n synapses per cube			
n	mean (%)	SD (%)	min. (%)	max. (%)
0	1.1×10^7 (98.89)	3.1×10^6 (0.26)	2.6×10^6 (98.13)	1.9×10^7 (99.34)
1	1.2×10^5 (1.08)	4.1×10^4 (0.25)	3.3×10^4 (0.65)	2.5×10^5 (1.81)
2	2.6×10^3 (2.31×10^{-2})	1.6×10^3 (1.28×10^{-2})	3.9×10^2 (6.17×10^{-3})	1×10^4 (7.38×10^{-2})
3	1×10^2 (9.13×10^{-4})	1.3×10^2 (1.04×10^{-3})	6 (7.16×10^{-5})	1.3×10^3 (8.18×10^{-3})
4	7.2 (6.26×10^{-5})	18 (1.38×10^{-4})	0 (9.19×10^{-7})	2.4×10^2 (1.56×10^{-3})
5≥	0.86 (7.54×10^{-6})	4.2 (3.06×10^{-5})	0 (1.24×10^{-8})	72 (4.58×10^{-4})

Table A4. Absolute and relative (in brackets) number of branch pairs connected by zero to four or more synapses in 50 µm cubes. Statistics across 512 cubes.

	branch pairs connected by n synapses per cube			
n	mean (%)	SD (%)	min. (%)	max. (%)
0	3.2×10^7 (99.61)	1×10^7 (0.07)	8.1×10^6 (99.34)	6.6×10^7 (99.76)
1	1.2×10^5 (0.39)	4.3×10^4 (0.07)	3.3×10^4 (0.24)	2.7×10^5 (0.65)
2	1×10^3 (3.14×10^{-3})	5.4×10^2 (1.19×10^{-3})	2.1×10^2 (1.18×10^{-3})	3.9×10^3 (9.04×10^{-3})
3	16 (4.67×10^{-5})	21 (4.69×10^{-5})	2 (8.14×10^{-6})	2.5×10^2 (5.24×10^{-4})
4	0.44 (1.45×10^{-6})	2.2 (4.73×10^{-6})	0 (5.38×10^{-8})	35 (7.32×10^{-5})
5≥	0.029 (1.06×10^{-7})	0.38 (7.91×10^{-7})	0 (3.90×10^{-10})	7 (1.39×10^{-5})

Table A5. Empirical vs. predicted location-specific connection probabilities of L5PTs (Perin et al., 2011). Table adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020)

intersomatic distance (µm)	P(A,B) empirical (%)	median (%)	P(A,B) model 25^{th} percentile (%)	75^{th} percentile (%)
18	21.8	20.7	10.9	33.4
53	16.9	18.7	9.0	30.1
88	14.5	15.7	7.3	26.7
123	12.5	11.1	4.6	22.5
158	9.5	7.8	2.1	19.1
193	10.3	5.6	1.1	14.5
228	6.5	2.4	0.0	10.7
263	3.6	1.6	0.0	9.9
298	7.1	5.1	0.0	12.8

Table A6. Empirical vs. predicted location-specific connection probabilities in L2/3 (Avermann et al., 2012). Table adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020)

intersomatic distance (µm)	P(A,B) empirical (%)	median (%)	P(A,B) model 25^{th} percentile (%)	75^{th} percentile (%)
20	21.4	18.6	6.2	35.7
60	17.3	16.8	4.8	32.1
100	12.8	12.5	2.8	26.3
140	14.2	8.6	1.0	21.0

Table A7. Empirical vs. predicted synapse distributions along dendrites in L5 (Petreanu et al., 2009). Table adapted from Udvary, Dercksen, Harth, Hege, de Kock, et al. (2020)

Presynaptic layer	Postsynaptic layer	Correlation coefficient R	p value
L4	L5B	0.81	2.35×10^{-9}
L2/3	L5B	0.83	1.73×10^{-10}
VPM	L5B	0.87	7.17×10^{-11}
L4	L5A	0.93	4.21×10^{-13}
L2/3	L5A	0.93	1.19×10^{-13}
VPM	L5A	0.94	2.29×10^{-13}

Table A8. Empirical vs. predicted number of synapses per excitatory connection.

	Grouping		Empirical			Model		
	Presynaptic	Postsynaptic	n(A,B)			n(A,B)		
Ref.	A	B	range		mean	range		mean
[1]	L2/3	L2/3	2	4	2.8	1	5	1.47
[2]	L4	L2/3	4	5	4.5	1	5	1.47
[3]	L4ss	L2/3	4	6	4.8	1	6	1.63
[4]	L4ss	L4ss	2	5	3.4	1	5	1.44
[5]	L5A	L5A	1	6	3.5	1	3	1.25
[6]	L5	L5	4	8	5.5	1	3	1.26
[7]	VPM	L4	1	6	2.5	1	6	1.80

[1] Feldmeyer et al. (2006) [2] Feldmeyer et al. (2002) [3] Lübke et al. (2003) [4] Feldmeyer et al. (1999) [5] Frick, Feldmeyer, Helmstaedter, and Sakmann (2008) [6] Markram et al. (1997) [7] Schoonover et al. (2014)

Table A9. Empirical vs. predicted numbers of synapses per connection involving inhibitory neurons.

	Grouping		Empirical			Model		
	Presynaptic	Postsynaptic	n(A,B)			n(A,B)		
Ref.	A	B	range		mean	range		mean
[1]	L4IN	L4ss	2	6	3.5	1*	6*	1.9*
[1]	L4BIn	L4ss	2	6	3.7	1*	7*	2.0*
[1]	L4ss	L4IN	2	3	2.1	1	4	1.4
[2]	L4ss	L2/3IN	1	3	2.3	1	4	1.4
[3]	L5IN[a]	L5EX	9	15	12.0	1	6	1.8
[3]	L5EX	L5IN[a]	6†	11†	8.6	1	3	1.1

[1] Koelbl et al. (2013) [2] Helmstaedter et al. (2008) [3] Silberberg and Markram (2007) [a] empirical data for subgroup * predicted by emulating slice † estimates from reported SD (8.6 ± 2.1).

Table A10. Comparison of connection probabilities between excitatory neurons.

	Grouping				Empirical			Model		Comparison	
	Presynaptic A		Postsynaptic B					P(A,B)	P(A,B)		
							P(A,B)	MEAN	SD		*prctl*
ID	layer	type	layer	type	*in vivo*	*n*	(%)	(%)	(%)	dev_{SD}	(%)
1[1]	VPM		L4	n/s	yes	40	43	44	28	-0.04	50
2[2]					yes	62	37			-0.23	44
3[1*]					yes	14	43	39	28	0.13	56
4[2*]					yes	21	38			-0.04	52
5[1]				L4sp	yes	24	42	45	26	-0.11	45
6[1]				L4ss	yes	11	64	42	27	0.82	76
7[3]			L5	L5IT	yes	18	17	28	21	-0.54	36
8[3]				L5PT	yes	9	44	40	24	0.18	57
9[3]			L6	n/s	yes	11	9	18	19	-0.48	45
10[4]	L2	n/s	L2	n/s	no	950	9	16	19	-0.34	51
11[5]					yes	878	7	31	25	-0.99	26
12[4]			L3	n/s	no	183	5	15	16	-0.58	39
13[4]			L4	n/s	no	208	1	9	12	-0.64	38
14[4]			L5A	n/s	no	211	9	8	10	0.12	68
15[4]			L5B	n/s	no	108	8	7	9	0.18	71

Continued on next page

Table A10 – continued from previous page

16[4]			L6	n/s	no	50	0	1	3	-0.39	69
17[6]	L2/3	n/s	L2/3	n/s	no	95	17	16	18	0.05	63
18[7]					no	-	10			-0.33	50
19[8]					no	542	5			-0.61	37
20[9]					no	112	2			-0.79	27
21[10]					no	247	26			0.58	77
22[11]					no	112	20			0.21	68
23[12]					no	235	19			0.18	67
24[13]					yes	774	7	25	24	-0.76	30
25[9]			L5	L5IT	no	98	4	7	9	-0.34	51
26[9]				L5PT	no	51	4	9	11	-0.51	41
27[4]	L3	n/s	L2	n/s	no	182	12	15	18	-0.14	59
28[4]			L3	n/s	no	513	19	17	18	0.1	64
29[4]			L4	n/s	no	170	2	11	15	-0.61	40
30[10]			L5	n/s	no	29	55	8	10	4.79	100
31[4]			L5A	n/s	no	87	6	9	11	-0.28	53
32[4]			L5B	n/s	no	164	12	6	9	0.69	82
33[4]			L6	n/s	no	61	0	1	3	-0.39	68
34[4]	L4	n/s	L2	n/s	no	208	12	12	14	0.03	64
35[14]			L2/3	n/s	no	64	15	15	17	0.01	62

Continued on next page

Table A10 – continued from previous page

36[11]					no	50	20			0.32	70
37[4]			L3	n/s	no	172	15	17	18	-0.16	54
38[10]					no	25	28			0.6	76
39[15]			L4	n/s	no	89	6	13	16	-0.44	47
40[4]					no	1046	24			0.72	80
41[16]			L5A	n/s	no	-	14	9	10	0.5	77
42[4]					no	276	12			0.27	71
43[4]			L5B	n/s	no	136	8	6	8	0.24	71
44[4]			L6	n/s	no	93	3	1	3	0.6	88
45[17]		L4PY	L4	L4PY	no	528	4	11	12	-0.53	43
46[18]		L4sp		L4sp	no	24	21	19	14	0.14	60
47[19]		L4ss		L4ss	no	94	26	21	20	0.21	64
48[20]					no	146	36			0.7	77
49[18]					no	24	21			-0.02	57
50[10]	L5	n/s	L3	n/s	no	29	3	10	15	-0.41	56
51[9]			L5	n/s	no	150	0	8	10	-0.78	22
52[21]					no	500	10			0.22	71
53[22]					no	1450	12			0.42	76
54[10]					no	163	9			0.14	69
55[9]		L5IT	L2/3	n/s	no	98	0	10	13	-0.77	19
56[23]			L5	L5IT	no	118	5	11	9	-0.62	34

Continued on next page

Table A10 – continued from previous page

57[9]					no	66	0			-1.17	4
58[23]				L5PT	no	86	19	17	13	0.09	59
59[9]					no	36	0			-1.31	3
60[9]		L5PT	L2/3	n/s	no	51	0	2	5	-0.45	55
61[23]			L5	L5IT	no	86	5	7	8	-0.32	50
62[9]					no	36	0			-0.88	13
63[23]				L5PT	no	225	7	16	14	-0.65	34
64[9]					no	12	0			-1.14	6
65[24]					no	3446	13			-0.27	49
66[25]					no	8050	12			-0.34	47
67[4]	L5A	n/s	L2	n/s	no	209	4	10	14	-0.4	53
68[4]			L3	n/s	no	89	2	13	17	-0.66	37
69[4]			L4	n/s	no	275	1	10	14	-0.65	34
70[4]			L5A	n/s	no	934	19	9	10	1.03	87
71[4]			L5B	n/s	no	175	8	8	10	0	65
72[4]			L6	n/s	no	158	3	2	4	0.34	82
73[4]	L5B	n/s	L2	n/s	no	104	1	3	8	-0.23	77
74[4]			L3	n/s	no	167	2	3	7	-0.14	75
75[4]			L4	n/s	no	137	1	2	4	-0.29	68
76[4]			L5A	n/s	no	174	2	5	8	-0.48	48

Continued on next page

Table A10 – continued from previous page

77[26]			L5B	n/s	no	269	9	11	12	-0.11	58
78[4]					no	555	7			-0.29	51
79[4]			L6	n/s	no	100	7	5	9	0.19	75
80[4]	L6	n/s	L2	n/s	no	50	0	0	2	-0.19	92
81[4]			L3	n/s	no	64	0	1	4	-0.26	85
82[4]			L4	n/s	no	94	0	1	4	-0.31	74
83[4]			L5A	n/s	no	160	1	4	6	-0.49	53
84[4]			L5B	n/s	no	100	2	9	11	-0.7	33
85[27]			L6	n/s	no	102	2	7	11	-0.5	51
86[4]					no	532	3			-0.42	54
87[28]					no	27	4			-0.33	57
88[29]	L6A	$L6_ACC$	L6A	L6CT	no	43	9	7	9	0.25	69
89[29]	L6A	L6CT	L6A	$L6_ACC$	no	40	0	10	12	-0.89	14

[1] Bruno and Sakmann (2006) [2] Bruno and Simons (2002) [3] Constantinople and Bruno (2013) [4] Lefort et al. (2009) [5] Jouhanneau et al. (2015) [6] Avermann et al. (2012) [7] Feldmeyer et al. (2006) [8] Holmgren, Harkany, Svennenfors, and Zilberter (2003) [9] Jiang et al. (2015) [10] Thomson et al. (2002) [11] Yoshimura et al. (2005) [12] Hofer et al. (2011) [13] Jouhanneau, Kremkow, and Poulet (2018) [14] Feldmeyer et al. (2002) [15] Beierlein, Gibson, and Connors (2003) [16] Feldmeyer, Roth, and Sakmann (2005) [17] Bannister and Thomson (2007) [18] Sun et al. (2006) [19] Feldmeyer et al. (1999) [20] C. C. Petersen and Sakmann (2000) [21] Markram et al. (1997) [22] Silberberg and Markram (2007) [23] Brown and Hestrin (2009) [24] Perin et al. (2011) [25] Song et al. (2005) [26] Krieger, Kuner, and Sakmann (2007) [27] Beierlein and Connors (2002) [28] Mercer et al. (2005) [29] Crandall, Patrick, Cruikshank, and Connors (2017) * denotes connectivity measurements of neurons located between barrel columns (i.e., within the septum). *n* denotes the number of pairs tested.

Table A11. Comparison of connection probabilities involving inhibitory neurons.

	Grouping		Empirical	Model		Comparison	
ID	Presynaptic A	Postsynaptic B	P(A,B) (%)	P(A,B) MEAN (%)	P(A,B) SD (%)	dev_{SD}	*prctl* (%)
1[1]	VPM	L4IN	57	46	30	0.35	61
2[2b]			62			0.52	65
3[3]	L1IN	L1IN	32	36	35	-0.14	54
4[3]		L2/3EX	18	30	31	-0.39	49
5[3]		L2/3IN	14	16	23	-0.06	67
6[3]		L5EX	7	16	25	-0.34	61
7[3]		L5IN	0	0	2	-0.19	87
8[3]	L2/3EX	L1IN	1	6	13	-0.4	57
9[4]		L2/3IN	40	16	18	1.32	89
10[5b]			63			2.56	97
11[6b]			48			1.73	92
12[7b]			51			1.93	94
13[7b]			31			0.79	81
14[8]			21			0.24	70
15[9b]			19			0.13	67
16[3]			8			-0.44	47
17[3]		L5IN	4	12	15	-0.5	44
18[3]	L2/3IN	L1IN	14	6	12	0.69	87
19[4]		L2/3EX	53	31	31	0.71	75
20[10a]			55			0.8	76
21[6a]			67			1.17	82
22[11a]			43			0.4	69
23[12a]			22			-0.29	52
24[8]			16			-0.49	46
25[9a]			47			0.51	71

Continued on next page

Table A11 – continued from previous page

26^{3}			32			0.05	61
27^{4}		L2/3IN	37	20	24	0.74	80
28^{13ab}			57			1.58	90
29^{13ab}			35			0.66	78
30^{3}			21			0.03	65
31^{3}		L5EX	9	13	19	-0.18	65
32^{3}		L5IN	6	3	8	0.38	88
33^{8}	L3EX	L4IN	8	16	19	-0.41	49
34^{8}	L3IN	L3IN	25	19	22	0.29	72
35^{8}	L4EX	L3IN	10	28	22	-0.84	24
36^{14b}		L4IN	43	27	24	0.67	76
37^{14b}			57			1.26	87
38^{15}			41			0.59	75
39^{8}	L4IN	L3EX	50	15	22	1.61	90
40^{14a}		L4EX	44	22	25	0.86	81
41^{14a}			35			0.5	75
42^{16a}			50			1.1	84
43^{17a}			66			1.75	91
44^{15ab}		L4IN	62	20	23	1.81	92
45^{15ab}			15			-0.21	58
46^{18a}		L4sp/ss	21	22	26	-0.04	63
47^{18a}			27			0.17	69
48^{19}		L4ss	67	48	29	0.67	68
49^{19a}	L4BIn	L4ss	63	52	31	0.37	57
50^{18b}	L4sp/ss	L4IN	13	31	24	-0.77	27
51^{18b}			33			0.1	61
52^{20}	L4ss	L2/3IN	28	24	22	0.18	65
53^{3}	L5EX	L1IN	1	3	7	-0.27	70
54^{3}		L2/3IN	1	11	15	-0.71	29
55^{21b}		L5IN	13	13	15	0.06	66

Continued on next page

Table A11 – continued from previous page

56[8]			10			-0.21	57
57[3]			6			-0.45	46
58[3]	L5IN	L1IN	4	1	4	0.88	95
59[3]		L2/3EX	10	6	14	0.26	82
60[3]		L2/3IN	10	4	9	0.69	89
61[21a]		L5EX	3	28	27	-0.94	17
62[11a]			67			1.42	87
63[22a]			27			-0.05	60
64[12a]			22			-0.24	54
65[8]			12			-0.59	39
66[3]			25			-0.13	57
67[8]		L5IN	60	15	18	2.5	96
68[3]			22			0.39	76
69[23b]	L6EX	L6IN	10	16	20	-0.31	55
70[15]			41			1.25	87
71[12a]	L6IN	L6EX	22	31	28	-0.32	49
72[15ab]		L6IN	62	15	18	2.55	97
73[15ab]			15			-0.01	64

[1] Porter, Johnson, and Agmon (2001) [2] Bruno and Simons (2002) [3] Jiang et al. (2015) (average across all subgroups in respective layers) [4] Avermann et al. (2012) [5] Holmgren et al. (2003) [6] Kapfer, Glickfeld, Atallah, and Scanziani (2007) [7] Pala and Petersen (2015) [8] Thomson et al. (2002) [9] Yoshimura and Callaway (2005) [10] Fino and Yuste (2011) [11] Packer and Yuste (2011) [12] Thomson, West, Hahn, and Deuchars (1996) [13] Walker et al. (2016) [14] Beierlein et al. (2003) [15] Gibson, Beierlein, and Connors (1999) [16] Gabernet, Jadhav, Feldman, Carandini, and Scanziani (2005) [17] Inoue and Imoto (2006) [18] Sun et al. (2006) [19] Koelbl et al. (2013) [20] Helmstaedter et al. (2008) [21] Otsuka and Kawaguchi (2009) [22] Silberberg and Markram (2007) [23] Beierlein and Connors (2002) [a] empirical data for presynaptic subsample [b] empirical data for postsynaptic subsample.