Characterization and modeling of primate cortical anatomy and activity

Inaugural-Dissertation

zur

Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität zu Köln

vorgelegt von

Aitor Morales-Gregorio
aus Gernika-Lumo

Universität zu Köln
August, 2022

Gutachter/in:	Prof. Dr. Sacha J. van Albada
Gutachter/in:	Prof. Dr. Martin P. Nawrot
Vorsitzender der Prüfungskommission:	Prof. Dr. Kei Ito
Tag der mündlichen Prüfung:	21.10.2022

The brain being indeed a machine, we must not hope to find its artifice through other ways than those which are used to find the artifice of the other machines. It thus remains to do what we would do for any other machine; I mean to dismantle it piece by piece and to consider what these can do separately and together.

— Nicolaus Steno, 1669[*]

At the risk of making a far-fetched comparison, I would defend this idea by saying that the cerebral cortex is like a garden full of an infinite number of trees – pyramidal cells – which, by careful cultivation, can produce more branches, push their roots deeper, and produce ever more varied and exquisite flowers and fruits.

— Santiago Ramón y Cajal, 1894[†]

We had dreams, Turing and I used to talk about the possibility of simulating entirely the human brain, could we really get a computer which would be the equivalent of the human brain or even a lot better? And it seemed easier then than it does now maybe. We both thought that this should be possible in not very long, in ten or 15 years. Such was not the case, it hasn't been done in thirty years.

— Claude Shannon, 1977[‡]

[*]Cobb (2020)
[†]Ramón y Cajal (1894)
[‡]Soni and Goodman (2017)

Abstract

Neuroscience is the study of the brain and all the complex mechanisms that make thought and cognition possible. The cerebral cortex is where some of the most complex cognitive processes are believed to occur. This work primarily focuses on the macaque, since it is a close relative to humans and a widely studied model animal. While experimental studies are limited to a few neurons and locations, computational models can compensate these limitations since they allow to study the entire system at will. However, there are many hurdles on the way to reliable and realistic brain models, some of which we addressed in this dissertation.

We identified some specific gaps in the knowledge that impede the creation of comprehensive brain models. These include: the lack of resting state extracellular neural recordings and its subsequent analysis, the lack of comprehensive neuron density estimates and their statistical distribution, and the lack of connectivity data within cortical areas. The aim of this dissertation is to address these gaps in the knowledge in order to construct comprehensive models of the macaque cortex at a neuronal level.

In this dissertation, we present high-resolution resting state data from macaque V1 and V4 areas, along with exhaustive quality controls and all the relevant metadata about the experiment. We then study the resting state data and show distinct structures in the population dynamics, which our analysis and simulations suggest could be modulated by feedback from V4 to V1. Moreover, we show that the distribution of neuron densities across and within the cortex of mammals is compatible with a lognormal distribution, which could easily emerge from a noisy cell division process. In addition, we present new measurements of neuron density in the macaque cortex, in an area and layer resolved manner. These measurements required a 3D reconstruction from histological slices and constitute, to the best of our knowledge, the first comprehensive data set of neuron densities in a single macaque. Finally, we present a method to estimate local microcircuit connectivity from resting state spiking activity, using single unit spiking statistics and the Wasserstein distance. We show that the activity is significantly different across the cortex and demonstrate the validity of our parameter estimation method using synthetic data.

In conclusion, this work provides activity and anatomical data for the neuroscience community, as well as several methods that will be applicable beyond the scope of this thesis. All in all, this work brings the field a small step closer to a comprehensive understanding of the cerebral cortex.

Zusammenfassung

Neurowissenschaften forscht das Gehirn und die komplexen Mechanismen, die Nachdenken und Erkenntnisse ermöglichen. Es wird allgemein angenommen, dass komplexe kognitive Prozesse in der Großhirnrinde stattfinden. Diese Arbeit fokussiert sich grundsätzlich auf Makaken, weil sie zu den Menschen eng verwandt sind und häufig erforscht werden. Experimentelle Studien sind oft auf geringe Neuronen und Gehirnareale begrenzt, aber numerische Modelle können diese Begrenzungen überwinden, indem sie die Erforshcung von dem ganzen System nach Bedarf ermöglichen. Zahlreiche Hindernisse erswchweren die Entwicklung zuverlässiger und bilogisch realistischer Gehirnmodelle. Diese Dissertation adressiert einige dieser Hindernisse.

Wir haben einige spezifische Wissenslücken entdeckt, die die Entwicklung vom umfassenden Gehirnmodellen verhindern. Diese Lücken sind: der Mangel an extrazellulären Neuronenaktivitätsaufnahmen im Ruhezustand und deren Analyse, der Mangel an umfassenden Neuronendichteschätzungen und deren statistische Verteilung, und der Mangel an Konnektivitätsdaten innerhalb von Gehirnarealen. Ziel dieser Arbeit ist es, diese Wissenslücken zu adressieren, um umfassende Modelle von Makakgroßhirnrinde auf Neuronaler ebene zu entwickeln.

In dieser Arbeit stellen wir hoch aufgelöste Ruhezustandsdaten von Makak V1 und V4 Arealen vor, zusammen mit vollständiger Qualitätskontrolle und alle die relevanten Metadaten über das Experiment. Danach werten wir die Ruhezustandsdaten aus und zeigen, dass es komplexe Strukturen in den Populationsdynamiken gibt. Mit unserer Analyse und Simulationen zeigen wir auch, dass diese Strukturen durch Feedbacksignale von V4 zu V1 entstehen könnten. Weiter untersuchen wir die Verteilung der Neuronendichte in der Großhirnrinde und zeigen, dass sie Lognormal erscheint und durch einen Rauschen behafteten Zellenspaltungsprozess entstehen könnte. Als letztes stellen wir eine neue Methode für die Schätzung der lokalen Konnektivität innerhalb eines Areales aus den Ruhezustandsdaten vor. Die experimentellen Daten lassen sich durch die Einzelneuronenstatistiken signifikant voneinander unterscheiden. Darüber hinaus bestätigen wir die Richtigkeit unserer Parameterschätzungsmethode mit synthetischen Daten.

Zusammenfassend stellt diese Arbeit allen Neurowissenschaftlern sowohl anatomische als auch Aktivitätsdaten zur Verfügung. Dabei wurden einige Methoden entwickelt, die über diese Arbeit hinaus nützlich sein können. Diese Dissertation bringt das Feld einem umfassenden Verstand der Großhirnrinde einen kleinen Schritt näher.

Acknowledgements

Conducting research means sometimes riding the tightrope above the abyss of the unknown, balancing the all encompassing doubts with exciting discoveries and insights. Luckily I did not have to walk this path alone, but had the support from countless people that I want to thank here.

My colleagues and friends at the INM-6 have made this endeavour a thoroughly enjoyable and fun experience with endless coffee breaks, lunch discussions and ample supply of niche memes. I thank especially Alessandra Stella and Anno Kurth, for countless discussions, hikes, dinners, car rides, beers, math, and more. Alex van Meegen for amazing discussions, better costumes, solving my (math) problems before I even knew I had them, great advice and co-mentoring. Robin Gutzen for your fantastic feedback, graphic design discussions, reproducibility and the always interesting food facts. Julia Sprenger for your support with navigating academia, endless supply of food, excellent python and overall joy of life. Tobias Schulte to Brinke for all the fun conversations and all our shared hobbies, in which you kept getting better than me in no time. Barna Zajzon for an out-of-the-blue collaboration that ended in nice board game meet-ups. Alex Kleinjohann for great coding sessions, teaching me proper array handling and constant bike lobbying. Rembrandt Bakker for introducing me to the fascinating world of neuroimaging, fun little anecdotes and bike facts. Shashwat Sridhar for the great conversations and reaching together for the stars. David Berling for the great ideas on neuron counting and contagious enthusiasm for neuroscience. Heather More for nice market lunches and fun conversations. Michael Dick for all the fun discussions, in spite of my irrational fear of the aberrant paella Hawaii. Simon Essink for the very nice conversations and fun mini hacking sessions. Jari Pronold for being a great office buddy, fantastic humor and fellow graph enthusiast. And to everybody else at the INM-6— Jasper, Kirsten, Hanjia, Agnes, Paulina, Moritz L, Pietro, Robin P, Dennis, Johanna, Junji etc.

Of course this thesis could only really happen thanks to the constant support and mentoring of my fantastic supervisor Sacha van Albada. Thank you for supporting me from the very beginning with all of my projects and scientific endeavours, for ensuring I did not drift too far from our objectives, for your very humane approach to supervision, and for all the opportunities for exchange and collaboration that you have enabled and encouraged.

I am very thankful to Sonja Grün and Markus Diesmann for providing me with the opportunity to conduct my research, and fostering a collaborative and open-minded research institute.

I was blessed to be part of some fantastic international collaborations. In particular I

thank: Xing Chen for your tireless efforts with the data, for teaching me all the details about macaque electrophysiology, for your patience, and for all the joyful discussions; Bjørg Kilavik for your straight to the point comments, much appreciated friendliness, passion for science and fantastic advice; as well as Georgia Gregoriou, Thomas Brochier and Pieter Roelfsema for providing me amazing growth opportunities, discussing science with me on equal footing and for the endless streams of brain activity data.

Finally on the personal side of things, thanks to my amazing partner, Jon, for your continued support (even proofreading!), patience, and making me smile even in the worst days; as well as ensuring the electric grid runs smoothly, quite nice of you. I am forever grateful to my family, who supports me to the fullest, even though the academia is still a big mystery for you sometimes. You taught me the value of knowledge and encouraged me to follow my curiosity so I could become an ”inventor of things that have not been invented”. And thanks for the ongoing support from my friends—Eneko, Jonan, Gontzal, Lorea, Irati, Oihana, Idoia, Josu, Eider, Lena, Hannah S & N, Marie, Luisa and many more.

Eskerrik asko bihotzez guztioi

Contents

Part I

Introduction

Aim

Neuroscience studies the interactions between billions of neurons via trillions of synapses and how those interactions give rise to sophisticated cognitive functions. Bridging the structure and function is crucial for understanding how the brain processes information and interacts with the environment.

An ultimate aim of neuroscience is to reach a holistic understanding of the brain, from the sensory input to the motor outputs and everything that happens in between. This gargantuan task is beyond the reach of any single scientist, thus we focus our efforts on the study of the cerebral cortex, where the most complex computations are believed to take place (Abeles, 1991; Kandel and Schwartz, 1981). This thesis further focuses on the study of the macaque, a close relative of humans and commonly studied model animal.

Much is known about how the different parts of the macaque cortex are organized and function (Abeles, 1991). However, crucial information is missing, such as precise estimates of anatomical structure, or reliable large-scale activity recordings. A holistic understanding of the macaque cortex can only be achieved if we have access to the full system, which is currently out of reach of even the most advanced experimental recording systems. Computational modeling approaches can compensate for the experimental limitations, since they enable ample experimentation and give access to all the variables of the system. However, models are limited by their assumptions taken in their construction, which always have to be critically analyzed when interpreting any results.

The construction of brain models requires consolidating vast knowledge into a concrete system:

> *Equations force a model to be precise, complete, and self-consistent, and they allow its full implications to be worked out. [...] Mathematical formulation of a model forces it to be self-consistent and, although self-consistency is not necessarily truth, self-inconsistency is certainly falsehood.*
>
> — Abbott (2008)

Building a precise, complete and self-consistent model can immediately expose any gaps in the available knowledge; either due to missing information on the structural level or due to its limited ability to exhibit certain realistic dynamics. In the process of constructing a large-scale model of the macaque cortex we identified several gaps in the available knowledge:

- Realistic spiking neuron models are often constructed to operate in a rest-like regime, without a particular task. However, neural electrophysiology data in the resting state is scarce, with most studies reporting only data for specific tasks.
- Standard analysis of neural activity has traditionally focused on small groups of neurons. However, such methods are ill suited to study large-scale recordings and simulations. New analysis tools are needed to study large spiking neuron data.
- There is no standard method for comparing experimental and simulated activity.
- The distribution of neuron densities across the cortex of mammals has not been statistically characterized.
- Neuron densities of the macaque cortex are not publicly available for all cortical areas in a layer-resolved manner.
- Intra-area connectivity of macaque neocortex has only been measured for a few cortical areas.
- Long-range inter-area connectivity for macaque neocortex is incomplete.

The aim of this thesis is to address these gaps in the available knowledge. Filling these gaps can facilitate the construction and analysis of large-scale neuron network models, and simultaneously improve the overall understanding of macaque brain structure and function.

Structure of this thesis

Based on the aforementioned gaps in the available knowledge, several scientific questions were identified:

1. What should we compare our brain simulations to? Can we bridge some of the gaps between experimental and theoretical neuroscience?
2. *After observing rich dynamics in the state space of the macaque visual cortex:* Does feedback modulate the state space of the visual cortex? Can brain simulations help us understand this modulation?
3. Do neuron densities follow a particular distribution across the cortex?
4. How many neurons are there in each area and layer of macaque cortex?
5. Can we estimate the connectivity within cortical areas from brain activity? How can we best compare simulated and experimental spiking neuron data?

The chapters in Part II of this thesis each answer one of the above questions, respectively. We elaborate on the response to these questions in the discussion (Part III).

Chapter 2 shows a novel discovery, that feedback can modulate the population activity and dimensionality in macaque V1. To do so, we analyzed the resting state data from Chapter 1 together with data from a different experiment (de Haan et al., 2018). We found distinct clusters in the activity state space of the macaque visual cortex. This observation could not be explained by the external input alone, and we indeed found that feedback from higher cortical areas was strongly correlated to the clusters in state space. This analysis used state-of-the-art methods for analyzing population level data, such as dimensionality reduction and spectral Granger causality. Using a balanced spiking neuron network (Brunel, 2000) we reproduced the state-space clusters by providing a feedback signal. The experimental analysis and simulations contribute to the understanding of the modulatory role of feedback in the visual cortex. This chapter contributes to the understanding of the population activity in the visual

cortex of the macaque, exposing some effects that future large-scale models should be able to reproduce.

Chapter 3 presents our discovery that neuron densities follow a lognormal distribution across the cortical areas of seven different mammalian species and within the cortical areas of the marmoset monkey. We show that a very simple cell division model with Gaussian noise naturally leads to a lognormal distribution of cell densities. In addition, including distributed cell proliferation times in the model could explain how the lognormal distribution appears simultaneously at different spatial scales. Thus, we uncovered a new organizational principle of the cortex that will be a valuable resource for future model building efforts.

Chapter 4 addresses a fundamental gap in knowledge: neuron densities for the macaque cortex have only been partially measured across different individuals and labs. Additionally, densities across cortical layers are only available for a handful of areas. We gathered openly available histological data, reconstructed the 3D brain from the slices, mapped the areas to a common atlas and estimated the neuron densities for all areas and layers. We applied our semi-automatic workflow to one individual—with three more still being processed—, providing a comprehensive measurement of neuron densities in the macaque. This work lays the groundwork for other researchers to apply similar methods in other animals.

Chapter 5 shows our approach to estimate intra-area connectivity (across layers) for several cortical areas. We devised a strategy to compare the activity from different areas and models to each other, based on single neuron statistics. Our comparison method reveals significant differences in the dynamics of the cortical areas during the resting state—in agreement with previous studies (Mochizuki et al., 2016). Then, we used an optimization algorithm which minimizes the difference between simulated and experimental activity, based on the single neuron statistics. Given our mechanistic models, we expect the estimated model parameters to approximate the underlying connectivity, thus providing estimates for intra-area connectivity.

In addition to the scientific questions answered in the main chapters, we explore further aspects of brain structure and modeling in the appendices (Part IV).

Appendix A outlines how to incorporate anatomical information into neuronal models. First, we list the most relevant methods for quantitative measurement of the brain anatomy. Then, we discuss how to incorporate those measurements into neuronal models. Finally, we also review different predictive methods and techniques to fill in the gaps where the

parameters are missing. All in all, this appendix provides insight into the relation between anatomy and modeling that was used throughout this thesis.

Appendix B is a short commentary on trans-thalamic connectivity, i.e. connections from cortex to thalamus and back to somewhere else in cortex. Previously, trans-thalamic connections had only been observed between similar area modalities (e.g. S1 to S2). In the commented paper the connections were found to link somatosensory and motor areas in mice, the first trans-thalamic connection across functionally distinct areas. We expand the discussion and explore the implications of such connections for brain function and modeling.

Background

In this section we provide the background information required to understand this thesis. The section is structured as follows: first, we discuss fundamental concepts of neuroscience; then, we discuss the most common methods for recording brain activity; and finally, we discuss computational models of neurons and networks.

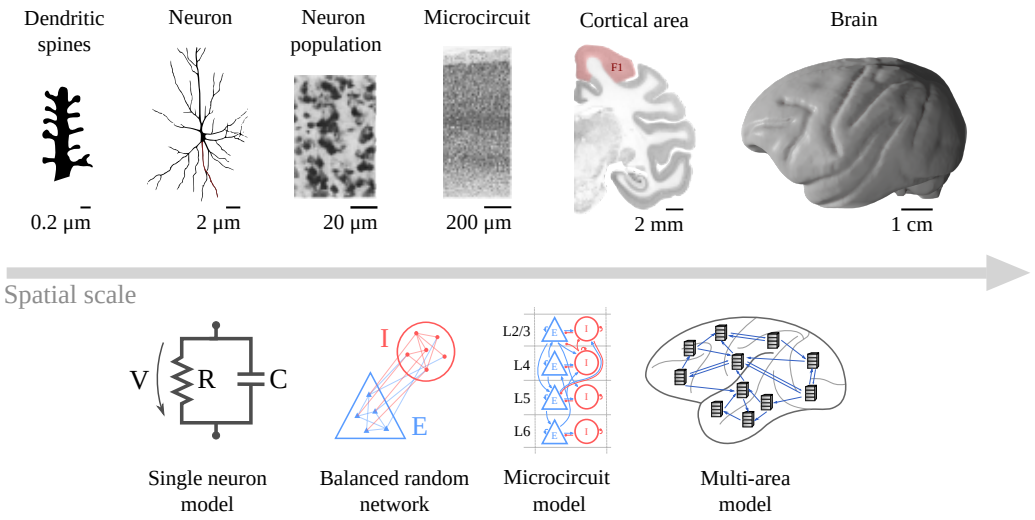
Fundamental principles

Nerve cells (*neurons*) are the fundamental components of the nervous system. Discrete electric pulses (*action potentials* or *spikes*) transmit information via anatomical connections (*synapses*, *gap junctions*) between the neurons. Action potentials are the orderly exchange of electrically charged ions between the intra- and extracellular space, via dedicated channels embedded in the semi-permeable membrane of neurons. The action potentials are generated at the cell body (*soma*) and travel along long protrusions of the cell (*axons*)—sometimes up to several meters—to reach their target. In the brain, neuronal axons target other neurons, to which they most often link via synapses onto the tree-like structures of neurons (*dendrites*). Upon receiving an action potential, specific chemicals (*neurotransmitters*) are released in the synapses. The neurotransmitters can increase (*excitation*) or reduce (*inhibition*) the postsynaptic membrane potential. The inputs received in the dendrites—in the form of changes of membrane potential—travel to the soma where they are aggregated. If a certain voltage is reached the neuron generates an action potential, i.e. *fires* a spike. Dale's law states that most neurons release a unique type of neurotransmitter, thus neurons can be grouped into two main categories: *excitatory* and *inhibitory* neurons. Synapses are not the only means of communication between neurons: some of them have their intracellular spaces directly connected via dedicated protein channels (*gap junctions*).

Neuron bodies are found in the outermost layer of the cerebrum, often called the cerebral *cortex* (latin for bark or rind) or *gray matter*, due to its color under the microscope. The internal part directly in contact with the cortex (*white matter*) is mainly composed of

Background Figure 1: Drawing of the cerebral cortex of a 15 day old infant by Santiago Ramón y Cajal, based on slices stained with the Golgi method. Credit: Legado Cajal. Instituto Cajal (CSIC), Madrid.

insulated axons (*myelinated*), which connect different parts of the cortex to each other and to the rest of the nervous system. Different cell types in the nervous tissue follow a certain arrangement, leading to distinct cellular compositions across different parts of the nervous system (*cytoarchitecture*, from the Greek root *cyto* meaning cell). The cortical gray matter in particular is organized in a layered structure from the outermost surface (*pia*) down to the white matter boundary. Certain regions of the cortex are specialized to certain functions—such as processing the visual inputs (visual cortex) or planning and executing motor outputs (motor cortex). Cortical areas often display distinct neuron types (e.g. *pyramidal*, *chandelier*, *Betz*...). Thus, the cortex is often subdivided along functional

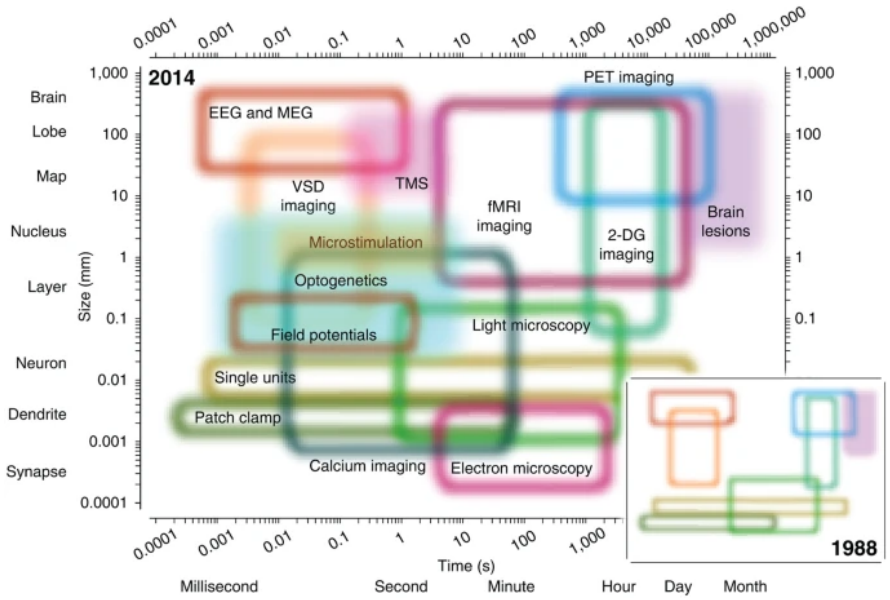


Background Figure 2: The brain at different spatial scales. Top showing different scales on a biological level. Dendritic spines re-drawn based on data by Bethe (1903); neuron re-drawn based on data by Golgi (1886); images for neuron population, microcircuit and cortical area derived from brainmaps.org (Mikula et al., 2007); brain rendered based on MRI atlas by Calabrese et al. (2015). Bottom shows different models used to represent the brain across scales.

and cytoarchitectural lines (*parcellation*). On slower time scales dedicated chemicals can modulate the activity of neurons (*neuromodulators*) as a means of communication.

The brain is a network (a *graph* in mathematical terms) where neurons are the nodes and the synapses between them are the edges. This corresponds to a *weighted* and *directed* graph, where the communication is delayed by the time it takes the spike to travel between the pre- and postsynaptic neurons. While connection strengths are prone to changes due to learning (*plasticity*) (Konorski, 1948; Hebb, 1949)—on both long and short timescales—, in this work we have limited ourselves to the study of static networks.

Brain networks maintain a stable level of activity via the balance between excitatory and inhibitory neuron activity (van Vreeswijk and Sompolinsky, 1996; Brunel, 2000; Haider, 2006), which has also been shown to enable rapid responses to external stimuli (van Vreeswijk and Sompolinsky, 1996).



Background Figure 3: Different techniques for measuring brain structure and function with their estimated spatial and temporal resolutions. Reproduced with permission from Sejnowski et al. (2014).

Measures of neuronal activity

The neuronal activity is the result of the transfer of charged ions, thus electric in nature. The changes in the electromagnetic field in the neuronal tissue can be recorded via different techniques with different spatio-temporal resolutions (Background Figure 3). In this section, we introduce a few of the most commonly used methods to measure neuronal activity.

Small dedicated electrodes can be used to measure the membrane potential intracellularly for a single neuron (*patch clamp*) or extracellularly for the aggregated activity from all surrounding neurons. This work focuses on the use of extracellular recordings, which can be processed to detect single neurons (*spike sorting*) (Quiroga et al., 2004), which we use in Chapter 5. The aggregate activity from extracellular recordings can also be studied at low (<200Hz) or high (>200Hz) frequencies, with the *local field potential* (LFP) or *multi-unit activity envelope* (MUAe) (Supèr and Roelfsema, 2005), respectively. We calculate and study the LFP and MUAe in Chapter 1 and Chapter 2.

The concentrations of certain ions can be measured in a time-dependent fashion using dedicated fluorescent proteins (Aria, 2020). Calcium imaging is the most widely used

method at the time of writing, although imaging of other ions such as sodium and potassium are under development (Lamy and Chatton, 2011; Rimmele and Chatton, 2014). Calcium is a by-product of synaptic activity, thus measuring it provides an indirect measurement of neuron activity. The light emitted by the dedicated proteins is recorded with an optical microscope (often a two-photon microscope) which sweeps over the region of interest periodically. The spatial and temporal resolution are therefore directly at odds, since larger spatial regions require a longer time to scan, reducing the temporal resolution.

All recording methods discussed until now are invasive, often requiring exposing the brain tissue via complex surgery. *Electrocorticography* (EcoG) is slightly less invasive, where the electrodes are placed on the dura—under the skull, but not in contact with the brain directly. Non-invasive methods such as *Electroencephalography* (EEG) or *Magnetoencephalography* (MEG) use larger electrodes placed on the scalp. These methods are less invasive, but do not have the ability to record from single neurons. Instead they provide an average of a larger area in the order of several millimeters.

Functional magnetic resonance imaging (fMRI) can measure the blood oxygen-level dependent (BOLD) signal, owing to the paramagnetic properties of hemoglobin. The fMRI can estimate the activity levels for the whole brain, under the assumption that high activity brain areas consume more oxygen. Given the non-invasive nature, availability at clinical facilities and its ability to record from the full brain at millimeter size voxels, the fMRI method is widely used to study the human brain.

Computational models

Measurement of brain activity and structure is complex and often done for different individuals under different conditions. Modeling provides a unifying framework to studying the brain, where all mechanisms and systems can in principle be studied simultaneously. Although models cannot yet capture all physical phenomena, one can instead focus on the mechanisms believed to be most relevant to the research question at hand.

Neuron models can be organized into two broad categories: phenomenological and biophysical. Phenomenological models capture the dynamics of the system without regard of the underlying mechanisms, while biophysical models attempt to include the underlying mechanisms (i.e. ion currents) that generate the correct dynamics. In general,

phenomenological models are built with fewer parameters and therefore are easier to scale up than biophysical models. However, even the most complex biophysical models rely on simplifications and approximations to capture the relevant dynamics as efficiently as possible.

One of the simplest phenomenological models for an action potential is the leaky integrate and fire (LIF) model, which is derived from the equivalent RC electrical circuit with one resistor and capacitor in parallel. In the LIF model the membrane potential V follows

$$\begin{cases} \tau_m \frac{dV}{dt} = -(V - v_{rest}) + R\,I, & V < v_{th} \\ V = v_{rest}, & V \geq v_{th} \end{cases} \tag{1}$$

where $\tau_m = RC$ is the membrane time constant, R is the membrane resistance, C is the membrane capacitance, I is the external current, v_{rest} is the resting membrane potential, and v_{th} is the firing threshold potential (Gerstner et al., 2014). When the threshold potential is reached the model is reset and a spike is emitted, while if the threshold is not reached and no input is received the membrane potential naturally tends to the resting potential. Spikes need to be accounted for by counting the number of threshold crossings. For its simplicity and versatility, this work uses the LIF model in Chapters 2 and 5.

Hodgkin and Huxley (HH) constructed their model based on experimental measurements of the dynamics of the squid giant axon (Hodgkin and Huxley, 1952; Gerstner et al., 2014). They found three main currents that determined the membrane potential: sodium, potassium and leak. They thus built their model based on these three currents and determined several variables for the opening and closing of the relevant gates and pumps on the membrane, based on their measurements. The model is described by a system of four differential equations describing the membrane potential V and three gating variables m, h and n. In spite of its historical importance, this thesis does not use the HH model.

Both the LIF and HH models are point neurons, i.e. they disregard the spatial extent of the neurons. This approximation is particularly useful when studying large groups of neurons interacting with each other. However, dendrites are complex and do not simply rely on the spikes received directly at the soma. Instead dendrites are believed to perform some nonlinear operations on the inputs, such that they can act as simple logic gates (Shepherd and Brayton, 1987). Dendrites can be modeled by defining different compartments for them. The membrane potential then has to be modeled for each compartment separately, while the

adjacent compartments can be coupled via the cable equation (Gerstner et al., 2014). Each compartment can be modeled using either the LIF, HH or some other neuron model. It is therefore easy to see how the computational cost increases dramatically, since a new set of differential equations needs to be solved per compartment. Some researchers study how to simplify the dendrite structure without losing its nonlinear properties (Wybo et al., 2021), such that the computational cost for simulation is reduced. In this work we only use point neurons, since we study relatively large systems with tens of thousands of neurons.

Neurons do not exist in a vacuum: they are connected to each other sending spikes to and receiving spikes from thousands of other neurons. Models can account for these interactions by delivering the spikes between neurons, using a prescribed connectivity. The spikes from a presynaptic neuron should reach the postsynaptic neuron after a short period of time (*delay*), and induce a (normally small) change to the postsynaptic membrane potential. Synaptic dynamics are diverse and depend on the interactions of neurotransmitters and neuromodulators with the ion gates at the synapses. However, they tend to follow a pattern of a peak followed by an exponential-like decay (Gerstner et al., 2014). The elicited postsynaptic current (PSC) is often modeled as an exponential function

$$I_{syn}(t) = w\, e^{-t/\tau_{syn}} \tag{2}$$

where w is the synaptic weight and τ_{syn} is the postsynaptic current time constant.

Graph theory provides the tools needed to study how the neurons connect to each other and how to build realistic networks of model neurons. One of the most renowned models for generating random graphs was proposed by Erdös and Rényi (Erdős and Rényi, 1959) back in 1959. However, this graphs lack clustering and scale-free properties observed in real world graphs (Girvan and Newman, 2002). The turn of the XXIth century brought a renewed interest in generative models for graphs, Watts and Strogatz (1998) proposed a model for clustered (small-world) networks, and Barabási and Albert (1999); Barabási and Oltvai (2004) proposed another model for scale-free networks. These models can be used in neuroscience to construct realistic networks of neurons.

At the largest scales simulating the whole brain at the single neuron resolution becomes prohibitively expensive. Thus, large-scale interactions are often studied using population models, i.e. the firing rate of entire populations becomes the simulated variable. Tools

from statistical physics allow the development of analytically tractable theories of population properties, where the effect of individual neurons is averaged out (Gerstner et al., 2014; Layer et al., 2022). This enables the reliable approximation of stationary first-order statistics (van Vreeswijk and Sompolinsky, 1996, 1998). Wilson and Cowan proposed a simple model of coupled ordinary differential equations to model the population rates in a time-dependent manner (Wilson and Cowan, 1972). The Wilson-Cowan model has been used to model the cortex to study different phenomena, such as resting state fluctuations (Deco et al., 2009), attractor dynamics (Deco and Jirsa, 2012), hierarchical timescales (Chaudhuri et al., 2015a), frequency dependent cortico-cortical communication (Mejias et al., 2016), or working memory (Froudist-Walsh et al., 2021).

The aim of building spiking neuron network models is often to replicate the structural and dynamical properties of the natural brain. For this aim, one needs to incorporate as much anatomical detail as possible into the models. We extensively discuss the most relevant anatomical variables for the construction of models in Appendix A. Biologically realistic models have been used to explore cortical activity at the level of a microcircuit (Douglas et al., 1989; Douglas and Martin, 2004). Notable examples include models of cat visual cortex (Douglas and Martin, 1991), a thalamocortical single column (Traub et al., 2005), macaque primary visual cortex (Zhu et al., 2009), a rodent cortical column for LFP estimation (Reimann et al., 2013), a local sensory cortical microcircuit (Potjans and Diesmann, 2014), rat somatosensory cortex (Markram et al., 2015), cat primary visual cortex (Antolík et al., 2018), layer 4 of mouse primary visual cortex (Arkhipov et al., 2018), mouse primary visual cortex (Billeh et al., 2020), macaque auditory cortex (Dura-Bernal et al., 2022a), and mouse primary motor cortex (Dura-Bernal et al., 2022b). Further biologically realistic models have aggregated such microcircuits into larger networks of point neurons, representing the visual cortex of the macaque (Schmidt et al., 2018a,b).

Part II

Publications and manuscripts

The eye and brain are not like a fax machine, nor are there little people looking at the images coming in.

Torsten Wiesel

1

Brain activity in the macaque visual cortex

Summary:

In this chapter we present the largest electrophysiological data set for non-human primates to date. The data were recorded during the resting state, which has rich dynamics and can help study the underlying structures and mechanisms of cortex. The data is accompanied by quality estimates, given strong visual stimuli. The receptive fields were also estimated based on a task with sweeping bars. The overall quality of the data was assessed using a novel approach for the detection of cross talk at high frequencies. We described the data acquisition and processing procedures, in compliance with FAIR principles. The full data, code and metadata are openly available to the entire scientific community.

scientific data

OPEN

DATA DESCRIPTOR

1024-channel electrophysiological recordings in macaque V1 and V4 during resting state

Xing Chen[1,8]✉, Aitor Morales-Gregorio[2,3,8], Julia Sprenger[2,4,5], Alexander Kleinjohann[2,4], Shashwat Sridhar[2,4], Sacha J. van Albada[2,3], Sonja Grün[2,4] & Pieter R. Roelfsema[1,6,7]

Co-variations in resting state activity are thought to arise from a variety of correlated inputs to neurons, such as bottom-up activity from lower areas, feedback from higher areas, recurrent processing in local circuits, and fluctuations in neuromodulatory systems. Most studies have examined resting state activity throughout the brain using MRI scans, or observed local co-variations in activity by recording from a small number of electrodes. We carried out electrophysiological recordings from over a thousand chronically implanted electrodes in the visual cortex of non-human primates, yielding a resting state dataset with unprecedentedly high channel counts and spatiotemporal resolution. Such signals could be used to observe brain waves across larger regions of cortex, offering a temporally detailed picture of brain activity. In this paper, we provide the dataset, describe the raw and processed data formats and data acquisition methods, and indicate how the data can be used to yield new insights into the 'background' activity that influences the processing of visual information in our brain.

Background & Summary

Using both depth electrode recording[1–6] and non-invasive brain imaging[7–13] techniques, a wealth of studies have shown that even in the absence of sensory input from the external environment, certain brain regions tend to share correlated patterns of neuronal activity, known as 'resting state correlations.' Such correlations have been observed across multiple sensory areas, such as auditory cortex[14], visual cortex[1,2,5,9,11,13,15,16], and somatosensory cortex[6,8,17,18]. They have also been observed in motor cortex[8,18] and in areas responsible for higher cognitive functions, such as the prefrontal cortex[8,18] and the parietal cortex[8,10].

Recent advances in ultra-high-density electrode fabrication and surgical implantation have spurred a surge in large-scale, multichannel recordings in rodents[19,20], including from multiple brain regions. However, ultra-high-channel-count electrophysiological recording techniques have yet to become widely adopted in non-human primates. Several challenges need to be addressed: electrode implantation requires access to the brain through a craniotomy (or several craniotomies) in the skull, limiting the number of recording sites and their spatial distribution. Existing probes with high channel counts, such as the Neuropixels probes from Imec (Belgium) were developed for mice and are often too fragile for chronic implantation in the primate brain[21], although more sturdy versions are under development. Presently available probes that are robust enough have relatively modest channel counts. Previous electrophysiological studies in non-human primates therefore usually involved the simultaneous implantation of up to dozens or, maximally, hundreds of electrodes in the brain[22].

In this study, we developed a novel neuronal recording system and implantable interface, to achieve chronic, high-resolution, large-spatial-scale recordings of neuronal activity in the visual cortex (V1 and V4) of two macaque monkeys[23]. These techniques allowed us to record neuronal activity across 1024 channels simultaneously, with extensive, high-density receptive field (RF) coverage across a large portion of the visual cortex (with

[1]Department of Vision & Cognition, Netherlands Institute for Neuroscience, Meibergdreef 47, 1105 BA, Amsterdam, Netherlands. [2]Institute of Neuroscience and Medicine (INM-6) and Institute for Advanced Simulation (IAS-6) and JARA Institute Brain Structure-Function Relationships (INM-10), Jülich Research Centre, Jülich, Germany. [3]Institute of Zoology, University of Cologne, Cologne, Germany. [4]Theoretical Systems Neurobiology, RWTH Aachen University, Aachen, Germany. [5]Institut de Neurosciences de La Timone, CNRS & Aix-Marseille University, Marseille, France. [6]Department of Integrative Neurophysiology, VU University, De Boelelaan 1085, 1081 HV, Amsterdam, Netherlands. [7]Department of Psychiatry, Academic Medical Center, Postbus 22660, 1100 DD, Amsterdam, Netherlands. [8]These authors contributed equally: Xing Chen, Aitor Morales-Gregorio. ✉e-mail: x.chen@nin.knaw.nl

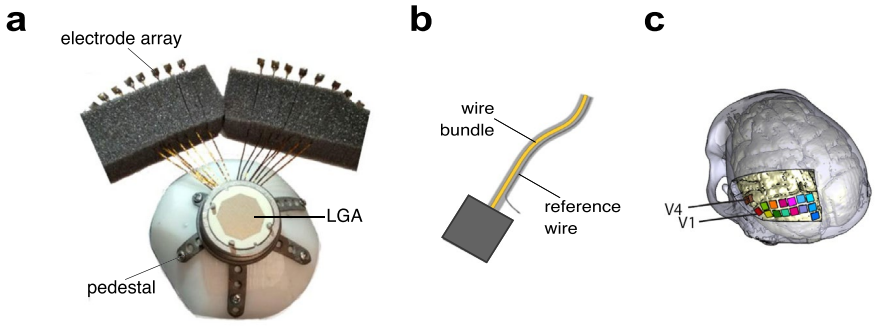


Fig. 1 (**a**) Photograph of the implant, consisting of a 1024-channel cranial pedestal connected to 16 Utah arrays. The base of the titanium pedestal was customized to fit precisely on the surface of the skull, as measured with a CT scan. (**b**) The references wires were located on alternating arrays (array numbers 1, 3, 5, 7, 9, 11, 13, and 15) and ran alongside the wire bundle before emerging several millimetres before the point of connection between the wire bundle and the array. (**c**) Location of implantation of every array in the visual cortex in monkey L. Two arrays were implanted in V4, and 14 arrays were implanted in V1, in each monkey.

overlap between V1 and V4 RFs), spanning the central 6–9 degrees of visual angle across one quadrant of the visual field. Our dataset[24] covers the full range of spectral components from raw signals sampled at 30 kHz to local field potentials (LFP, at 1–100 Hz) to multiunit spiking activity (MUA, at 500–9000 Hz).

We expect these resting state data to be of interest to neuroscientists in the fields of computational and systems neuroscience. Potential applications include correlation analyses, large-scale modelling[25], detection of activity waves[26], teaching material, and more. For example, the strength and anatomical distribution of co-variations in activity could shed light on the anatomical and functional connectivity between or within the areas under examination[27], including the retinotopic organisation of V1 and V4.

In the visual system, resting state correlations have been used to calculate functional correlations between brain regions in order to identify the borders of visual cortical areas such as V1, V2 and V3[15,28,29], and for the estimation of the retinotopic layout within individual visual areas[30–32]. Across visual areas, the lateral geniculate nucleus has been observed to exhibit higher levels of correlated activity with primary visual cortex than with higher-order visual areas, whereas activity in V2 and V3 is more closely correlated with V4 and hMT+[11]. Retinotopically corresponding locations across areas V1, V2 and V3 show increased functional connectivity, and a similar pattern has been observed for corresponding brain regions in the two hemispheres.

This dataset could further be used to compare electrophysiologically recorded neuronal activity to that obtained using non-invasive techniques. To give an example, we recently compared population RF estimates obtained with multiple-channel electrophysiology and fMRI-generated BOLD activity[33]. Indeed, MRI[8–10,15] and invasive electrophysiology[1–5,27,34–37] provide complementary approaches to examining correlations, including during resting state: fMRI offers a large-scale perspective, revealing the interplay between multiple brain areas and permitting the examination of entire resting state networks, via fluctuations in the MRI signal which have a relatively coarse spatial and temporal resolution. By contrast, electrophysiology yields direct recordings of neuronal activity from a smaller set of brain regions, but at a high spatial and temporal resolution.

The present dataset could also serve as a template for future publications of electrophysiology datasets, providing standardized methods and tools for the description, preparation, and organization of both data and metadata, thereby contributing to the present era of open data sharing and collaboration.

The data are available on the G-Node Infrastructure (GIN, https://gin.g-node.org/), an open-access data-sharing platform. The dataset version described in this publication can be found at https://gin.g-node.org/NIN/V1_V4_1024_electrode_resting_state_data. The dataset follows the FAIR principles[38], i.e. it is designed to be findable, accessible, interoperable, and reusable.

Methods

Our subjects were two male rhesus macaque monkeys (Macaca mulatta, monkeys A and L). Each animal received two cranial implants during two separate surgical procedures. The first of these implants was a customized, in-house-designed, 3D-printed head post for head fixation[39]. The head post was affixed to the setup to stabilize the head throughout the recordings. This ensured that the eye tracker captured the eye data (pupil diameter and position) accurately throughout the recordings (see Chen *et al*., 2017, for a detailed description of these methods). The second was a 1024-channel implant for the visual cortex, consisting of 16 Utah electrode arrays (Blackrock Microsystems) attached via 7-cm-long wire bundles to a customized, in-house-designed, 3D-printed pedestal (referred to in the rest of the manuscript as a '1024-channel pedestal,' Fig. 1)[23]. Each array contained an 8-by-8 grid with 64 iridium oxide electrodes. The length of each electrode shank was 1.5 mm and the spacing between adjacent shanks was 400 μm. The impedance of the electrodes at pre-implantation ranged from 6 to 12 kΩ (as measured by Blackrock Microsystems prior to lead attachment). Each electrode was connected to a contact pad on the Land-Grid-Array (LGA) interface of the pedestal. Reference wires were attached to every other array, and each reference wire served as the reference for two arrays, yielding eight reference wires in total. Each reference wire exited the wire bundle several millimetres before the point where the wire bundle met the array. The other end of the reference wire was connected to one of the contact pads on the LGA of the

pedestal (as was each of the electrodes), and referencing was performed by the Cereplex M headstage (i.e. the connections were hardwired such that the electrodes on each pair of arrays used the signal from the reference wire as their reference).

Surgeries. All experimental surgical procedures complied with the NIH Guide for Care and Use of Laboratory Animals (National Institutes of Health, Bethesda, Maryland), and were approved by the institutional animal care and use committee of the Royal Netherlands Academy of Arts and Sciences (approval number AVD-8010020171046). The subjects were 4 and 5 years old, and weighed 6.5 and 7.2 kg, respectively, at the time of head post implantation; and both were 7 years old, weighing 11.0 and 12.6 kg, respectively, during visual cortex implantation.

A course of antibiotics was started two days prior to each operation. We induced anaesthesia with intramuscularly administered ketamine (concentration of 7 mg/kg) and medetomidine (0.08 mg/kg). We administered 0.1 ml atropine (0.5 mg/ml) if the heart rate dropped below 75 bpm. The animal was placed on a heated mat to allow continuous regulation of body temperature. Eye ointment was applied to maintain hydration of the eyes. Xylocaine ointment was applied to the ear bars of the stereotaxic frame, and the animal's head was secured in the frame.

For the maintenance of anaesthesia, the animal was intubated and ventilated with 0.8–1% isoflurane (mixed with 60% O_2 and 40% air) and a catheter was inserted into a vein in the arm. During surgical implantation of the head post, we administered fentanyl at 0.005 mg/kg on indication, Ringer-glucose at 10–15 ml/kg/hour, and antibiotics intravenously. For surgical implantation of the electrode arrays, we additionally administered midazolam at 0.5 mg/kg (concentration 5 mg/ml) once per hour, and we administered dexamethasone at 0.25 mg/kg twice per hour, starting before opening of the skull until skull closure. ECG, heart rate, SpO_2, CO_2, temperature, muscle tone, respiration, and the response to pain stimuli were monitored continuously. The head was shaved and cleaned with chlorhexidine solution (Hibicet scrub) and iodine solution (5% iodine in water). For installation of cranial implants, a flap of skin was carefully detached from the skull over the desired implant location, reflected, and wrapped in damp cotton swabs to keep it moist. Methods of implanting the head post and arrays, and post-surgical recovery, are described separately in the following sections.

Head post: We sterilized the titanium head post by autoclaving it prior to surgery. It was placed on the skull and adjusted such that it fitted against the skull. We used 2-mm-diameter Ti cortex screws (DePuy Synthes, Amersfoort, Netherlands) to secure the head post to the bone. The wound margins were sutured together and an extra stitch was made to hold the skin closed around the base of the head post.

1024 channel implant: Before the surgery, we sterilized the implant (Fig. 1a) using gamma radiation. During the surgery the pedestal was placed on the skull and secured with bone screws. We made a craniotomy over the left hemisphere and opened the dura. We implanted 16 arrays of 64 electrodes each in the visual cortex (14 arrays in V1, and 2 in V4; Figs. 1c, 2a). The dura was sutured closed. We filled the space under the bone flap with Tissucol (Baxter) and placed the flap back while the Tissucol was still fluid. We secured the bone flap to the skull with Ti strips. The skin was pulled back around the pedestal and sutured closed.

Ten minutes before the end of the surgery, the ventilator was switched to stand-by mode, allowing spontaneous breathing. Upon conclusion of the procedure, the monkey was released from the stereotaxic frame. The isoflurane was switched off and an antagonist was administered intramuscularly (i.m.) (atipamezole 0.08 mg/kg), allowing the animal to wake up.

Recovery: Subjects were closely monitored following the operation and given several weeks to recover. We administered antibiotics (typically amoxicillin and clavulanic acid) for 10 days (in consultation with a veterinarian) and dexamethasone for five days in decreasing doses (from 0.7 mg/kg i.m. to 0.1 mg/kg i.m.). As analgesia, we initially used Temgesic, at two doses per day (0.003 mg/kg i.m.). After three days we switched to finadyne, once a day, for six days (1–2 mg/kg i.m.). The socially housed animals were housed solitarily during the first 8 to 9 days following surgery, after which social housing was resumed.

At the time of recording, the post-surgical implantation period was 2 and 4 years for the head posts, and 3 months and 1 year for the 1024-channel pedestal, for monkeys L and A, respectively, and the customized implants remained mechanically stable and well anchored to the skull throughout this period.

Datasets. In this study, we present 1) resting state data from the two monkeys, collected across three recording sessions per animal. In addition to this main dataset, we collected two supplementary datasets to allow further interpretation of the resting state data: 2) a dataset acquired during a visual fixation task, collected across three recording sessions per animal (on the same days as the acquisition of the resting state dataset), for quantification of the size and signal-to-noise ratio of the neuronal responses elicited by a visually presented checkerboard stimulus (Fig. 3a); and 3) a dataset acquired during a fixation task, collected across two recording sessions per animal, in which we presented moving light bars to map the receptive fields (RFs) of the neurons (Fig. 3b). Table 1 provides a list of all the recording sessions for the three datasets.

Resting state. For the resting state recordings, the monkey was seated, head-fixed, in a room next to the operator room, with the lights turned off. The room was silent during the recordings (although it was not acoustically isolated). Note that although the lights were off, the setup was not completely dark, due to the presence of small LED lights on our recording equipment, and a small amount of light coming under the door from the adjacent room. The monkey did not carry out a task and was allowed to stay awake or fall asleep at any point in time during the recording, and was free to shift its gaze and centre of attention. We recorded the pupil diameter and the eye camera also allowed us to determine whether the eyes were open or closed.

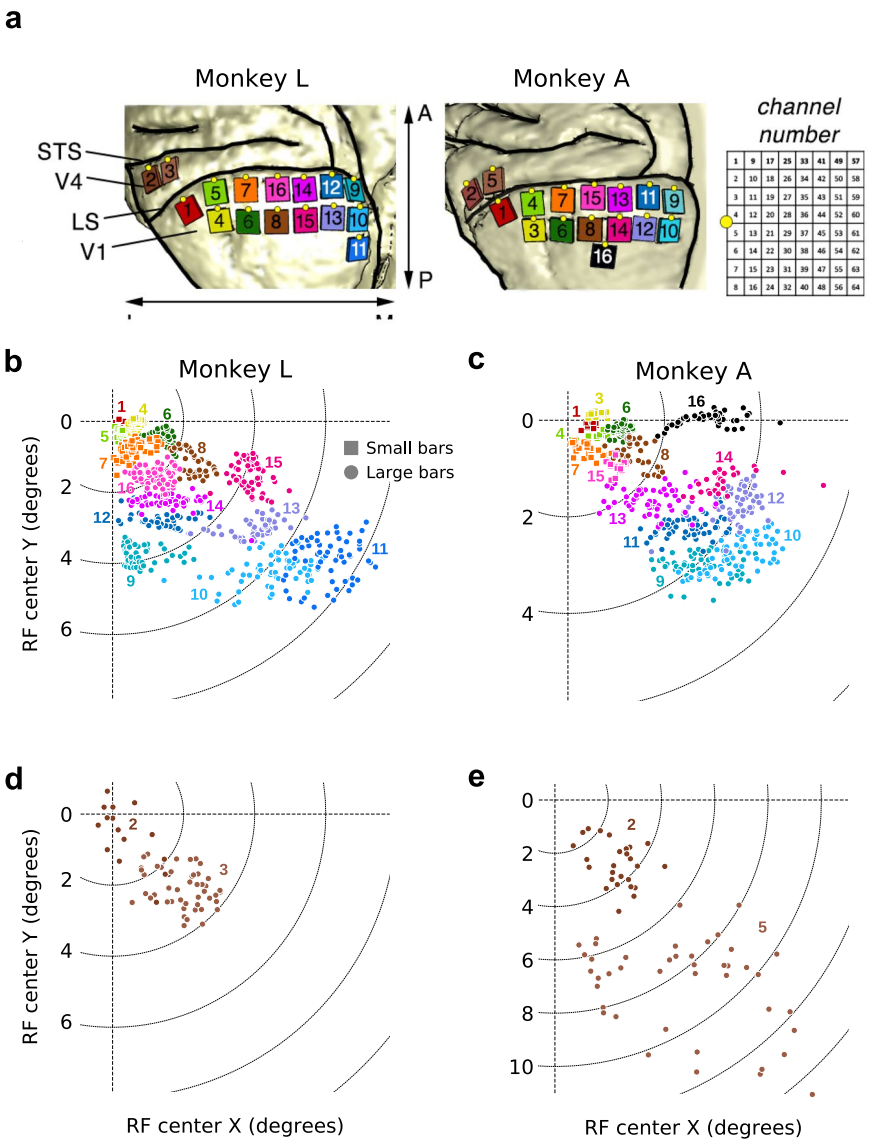


Fig. 2 (**a**) Numbering of the 16 arrays that were implanted in the visual cortex. LS: lunate sulcus. STS: superior temporal sulcus. The yellow dot on each array indicates the side on which the wire bundle exits the array. Right: Numbering of channels on each array, as viewed from the top of the array after implantation, rotated 90 degrees CCW relative to the left panel. (**b**–**e**) RF map, showing the coordinates of the V1 (**b**,**c**) and V4 (**d**,**e**) RF centres for channels with an SNR of more than 2 for each condition ($N = 893$ and 679 in monkeys L and A, respectively). Channels are colour-coded by array number, using the same colour code as in a. The receptive fields are located in the lower-right quadrant of the visual field.

Our aim was to provide a resting state dataset in which the signals were likely to be derived from the same or similar groups of neurons, allowing for pooling of data across the recording sessions. Therefore, the three resting state sessions were recorded within a short time span (across consecutive working days where possible).

Visually evoked activity. For each resting state dataset we also collected a dataset with visually evoked activity on the same day in order to provide an assessment of the quality of the neuronal signal on each channel that day. This dataset consisted of at least 30 trials in which the monkey viewed a grey screen (with a luminance of 14.8 cd/m^2) before a full-screen checkerboard stimulus was presented for 400 ms while the monkey maintained fixation on a dot located at the centre of the screen (Fig. 3a). The levels of visually evoked activity (relative to baseline activity) provided a measure of the quality of the neuronal signal obtained on each channel. We determined the 'signal-to-noise-ratio' (SNR) as the amplitude of the visually driven response divided by the standard deviation of activity in a time window before stimulus onset (see below for details). If desired, the SNR may be used during subsequent analyses to select only the channels from the corresponding resting state dataset that clearly have stimulus-evoked responses and to discard those that show poor or no signal. The size of the checkerboard squares was 1 degree of visual angle (dva), and the luminance values of the black and white squares were 0 and 92.1 cd/m^2, respectively.

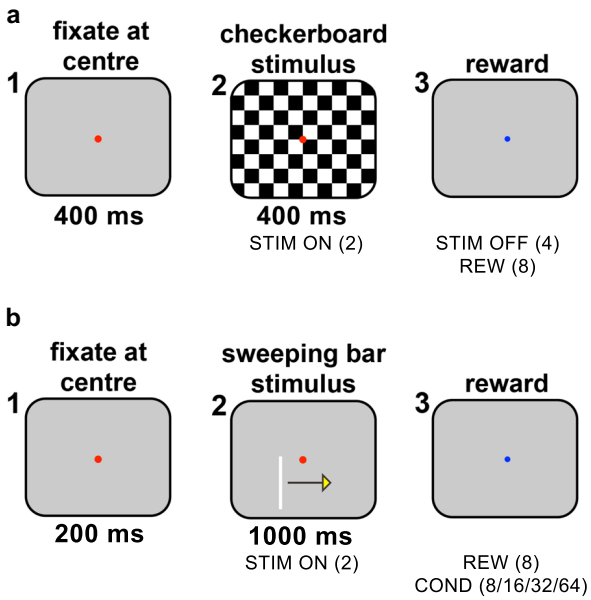


Fig. 3 (**a**) Illustration of the task used to measure visually evoked responses and calculate SNRs. 1: The monkey initiates the trial by fixating on a red spot at the centre of the screen. 2: After 400 ms, a checkerboard stimulus is presented. The monkey is required to maintain fixation throughout stimulus presentation, which lasts for 400 ms. 3: The monkey receives a reward upon stimulus offset. Event codes (as recorded in the.nev files) are shown between brackets; e.g. stimulus onset is encoded by the value '2.' (**b**) Illustration of the RF mapping task. 1: The monkey initiates the trial by fixating on a red spot at the centre of the screen. 2: After 200 ms, a bar stimulus is presented, which moves in one of the four cardinal directions (yellow arrow). The monkey is required to maintain fixation throughout stimulus presentation, which lasts for 1000 ms. 3: The monkey receives a reward upon stimulus offset.

Monkey	Task type		Recording day (dd/mm/yyyy)	Duration	Good channels (SNR > = 2)
A	Resting state		14/08/2019	32 min 34 s	—
			15/08/2019	38 min 17 s	—
			16/08/2019	42 min 0 s	—
	SNR		04/10/2018(*)	2 min 37 s	899
			14/08/2019	4 min 54 s	359
			15/08/2019	3 min 54 s	416
			16/08/2019	4 min 12 s	379
	RF	Large bars	28/08/2018	18 min 35 s	769(**)
		Small bars	29/08/2018	10 min 42 s	931(**)
L	Resting state		25/07/2017	22 min 42 s	—
			9/08/2017	22 min 0 s	—
			10/08/2017	21 min 37 s	—
	SNR		25/07/2017	1 min 36 s	981
			9/08/2017	1 min 35 s	977
			10/08/2017	1 min 58 s	992
	RF	Large bars	26/06/2017	33 min 43 s	957(**)
		Small bars	28/06/2017	31 min 47 s	821(**)

Table 1. Overview of datasets and sessions. Channel quality is based on the SNR. For each resting state session, an SNR session was collected on the same day to provide a measure of signal quality. *Extra SNR session, which was collected at an earlier period in time than the other SNR sessions, and which does not have a matching resting state session. **Channel quality in the RF datasets is considered to be good if the channel showed an SNR > = 2 for any sweeping bar direction.

In monkey A, the resting state data and matching visually evoked data were collected 1 year after surgical implantation, following the completion of other (unrelated) experiments. By this time, the number of channels with high SNR had decreased, compared to the number observed soon after surgery. To allow future users of the data to carry out analyses of visually evoked responses across close to 1024 channels (independently of the

Experiment control scripts		
Task	**Script name**	**Description**
Resting state	sync_pulse_resting_state.m	Sending of sync pulses to eight NSPs for post-hoc alignment of raw data.
SNR	runstim_CheckSNR.m	Presentation of full-screen checkerboard stimuli to elicit visually evoked responses.
RF	runstim_RF_barsweep_stimcondition1.m	Presentation of small sweeping bar stimuli to carry out RF mapping on channels where RFs were close to fixation.
	runstim_RF_barsweep_stimcondition2.m	Presentation of large sweeping bar stimuli to carry out RF mapping on channels where RFs were further from fixation.

Table 2. List of Matlab scripts used for experimental control and stimulus presentation, with script names and descriptions. The experimental control scripts are highly specific to the hardware and are not designed to run without the equipment. We provide the scripts for completeness.

resting state data), we provide an 'extra' dataset of visually evoked activity from monkey A, which was obtained 10 weeks after implantation. This 'early' SNR dataset was collected using an identical task design to that of the other SNR datasets, while providing a larger number of channels with high SNR. Note that this additional dataset is stand-alone and is not paired with a resting state session.

Receptive field mapping. The subjects viewed moving light bars that appeared at specific locations on the screen, allowing us to identify the RF location of the neurons recorded on each channel.

To characterize the receptive field properties on each channel, we recorded the responses evoked by white sweeping bar stimuli that moved in each of four possible directions (top to bottom; bottom to top; left to right; and right to left)[40]. RF size scales with eccentricity[41]: the farther away an RF is from the fixation, the larger its size. Neurons with small RFs respond best to small stimuli, whereas neurons with larger RFs show a more pronounced response to large stimuli.

Hence, the RF mapping task included two stimulus sets: 1) RFs of low eccentricity were mapped out using a small, thin, slow-moving bar (4 degrees of visual angle [dva] in length, 0.04 dva in thickness, moving at a rate of 4 dva per second) that was positioned close to the fixation spot. 2) RFs of higher eccentricity were mapped out using a long, thicker, faster-moving bar (20 dva in length, 0.19 dva in thickness, 20 dva/s) that was positioned farther from the fixation spot (see Fig. 3b). Stimulus presentation was controlled using custom-written Matlab scripts (Table 2) that were run on the stimulus control computer. The two types of visual stimuli elicited spatially and temporally well-defined neuronal responses, which allowed for the measurement of RFs closer to and farther away from fixation.

Data collection. Electrophysiological signals from V1 and V4 were recorded from 1024 channels distributed across 16 Utah Arrays (each consisting of 8×8 electrodes), at a sampling rate of 30 kHz (see Figs. 1c, 2a for their locations in the visual cortex), and further processed by equipment from Blackrock Microsystems (see Fig. 4 for a schematic overview of the setup). The neuronal signals were passively conducted via the LGA interface on the 1024-channel pedestal to an electronic interface board (EIB), i.e. an adapter with 32 36-channel Omnetics connectors, which in turn interfaced with eight 128-channel CerePlex M headstages. Each CerePlex M processed signals from two 64-channel Utah arrays, applying a 0.3–7500 Hz analog filter at unity gain (i.e. no signal amplification was carried out). The CerePlex M performed a 16-bit analog-to-digital conversion (ADC) with a sensitivity of 250 nV/bit. The digitized signal on each CerePlex M was sent to a 128-channel Digital Hub, i.e. each Digital Hub processed data coming from one CerePlex M, which in turn originated from two electrode arrays. The Digital Hub converted the digital signal into an optic-digital format, which was then sent via an optic-fibre cable to a 128-channel Neural Signal Processor (NSP) for further processing and storage. Each Digital Hub delivered the signal to a single NSP. There were eight NSPs and each NSP processed the data derived from two electrode arrays.

Control of the NSPs was carried out on two PCs (PC #1 and PC #2, running Windows 7 Professional) using the Blackrock Central Software Suite (version 6.5.4), with one instance of the software being run for each NSP, i.e. a total of eight instances of the software ran simultaneously during data acquisition. Each PC was connected to four NSPs, and four instances of the software were run on each PC. Each NSP stored the raw neuronal signals from 128 channels in a single raw data file (corresponding to channels 1 to 128), giving rise to a total of eight raw data files across the eight NSPs. The data recorded from the eight NSPs were temporally aligned as described in the section, 'Temporal alignment of raw data.'

Due to the high volume of data being processed and stored by each NSP, the onset of recording was controlled manually with a temporal offset of several seconds between NSPs. Before starting the recording on any given NSP, the operator checked to ensure that on-going recordings were running smoothly on the other NSPs, thereby avoiding buffer overflow and dropped packet issues due to system overload at the start of recording. Automatic updates were disabled to prevent unwanted disruptions during recording.

Eye tracking. During each recording, an infrared eye tracker (TREC ET-49B, version 1.2.8, Thomas Recording GmbH) was used to sample the eye position and pupil diameter for both the X- and Y-axes with a frame rate of 230 Hz, and the data were stored at a sampling rate of 30 kHz.

The eye tracking hardware was controlled by a dedicated PC (PC #4) using Eyetracer software (Thomas Recording), which forwarded the analog signals regarding the eye position and pupil diameter directly to NSP

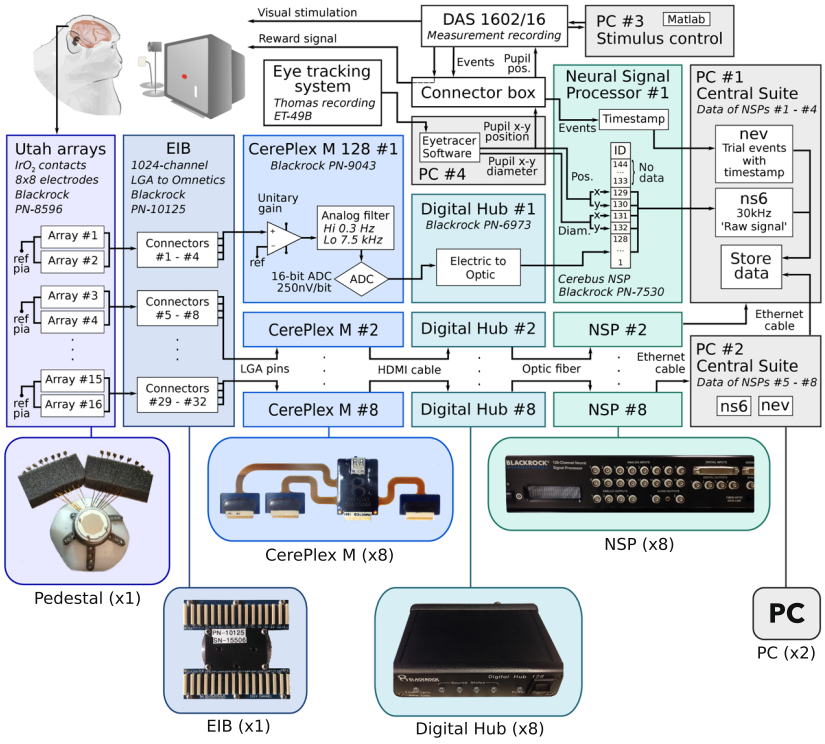


Fig. 4 Overview of devices and the total number of units (shown in parentheses) used to obtain, process and store data. Boxes represent individual devices/systems and arrows show the direction of signal transmission between the devices. Note that pairs of Utah arrays ultimately connect to a single neural signal processor (NSP), giving rise to a total of 8 parallel connection pathways. The apparatus and connections for the first connection pathway (i.e. for the first pair of Utah arrays) are shown in detail, while those for the other 7 pathways (for arrays 3 to 16) are depicted in condensed form. Pupil size and position are only recorded on NSP #1, and are not sent to the other 7 NSPs.

#1. X and Y eye positions were recorded on channels 129 and 130 of NSP #1, respectively, while X and Y pupil diameter were recorded on channels 131 and 132, respectively.

On all NSPs, the data collected on channels 1 to 128 comprised the neuronal signals. On NSP #1, channels 129 to 132 additionally contained the eye signals. Furthermore, analog synchronization signals were recorded on the NSP channel 144; these can be ignored as the synchronization pulses were also registered as digital events in the.nev files. All other NSP channels (133–143) did not record any data. We stored the raw data from the relevant channels (containing neural and eye signals) as specified in the configuration files that were loaded into the Blackrock Central Software Suite for each NSP.

Stimulus and reward timing, stimulus identity and experiment control. A stimulus control computer (PC #3) was used for the execution of task-related events and stimuli with high temporal precision during each of the task paradigms. The task-related event codes consisted of numbers that were sent out from PC #3, via a Data Acquisition and Control System (DAS) Multifunction Analog and Digital I/O board (DAS1602/16, Measurement Computing) through a splitter cable, to the digital input ports (16-bit DB37) on each of the NSPs. The corresponding channels on the digital input port of each NSP sampled the incoming signal at 30 kHz and were configured to detect when an incoming bit was set to a 'high' value on one of the pins. In our experiments, only the first 8 digital input pins (1 to 8) on the digital input ports of the NSPs were used, whereas the other 8 digital input pins (9 to 16) were disregarded. To encode a bit change initiated by the stimulus control computer, a 500-ms voltage pulse was sent on the desired pin. On each NSP, the event codes were recorded in the events file (.nev) as a sequence of numbers that ranged in value from 1 to 8. Note that since the DAS board used zero-based indexing, when instructions were sent from the Matlab script to the DAS board, the sequence of pin numbers specified in the Matlab script ranged from 0 to 7, instead of 1 to 8. Table 3 provides a list of the bit identities and their interpretations.

As the precise times at which recording was initiated or terminated varied across the eight NSPs, the duration of the raw data traces also varied slightly between NSPs. Hence, the common digital signal that was sent to all eight NSPs simultaneously via their digital input ports was used to precisely align the raw data traces between NSPs during data processing (described in the section, 'Temporal alignment of raw data').

During the resting state sessions, the digital signal consisted of a randomly generated sequence of numbers (ranging in value from 1 to 8), which were sent at 1-second intervals using a custom Matlab script that was run on the stimulus control computer. A list of experimental-control scripts is provided in Table 2. During the SNR

Bit	NEV encoding	Interpretation			
		Resting state	SNR	RF	
				Cond	Description
0	1	Sync pulse	—	—	—
1	2		Stimulus onset	—	Stimulus onset
2	4		Stimulus offset	—	Reward delivery
3	8		Reward delivery	1	Rightward sweeping bar
4	16		—	2	Upward sweeping bar
5	32			3	Leftward sweeping bar
6	64			4	Downward sweeping bar
7	128			—	—

Table 3. Relation between bit identity that is sent by the stimulus control computer, the event that is encoded in the events file (.nev) by the NSPs, and the trial-related event that occurred at the moment that the bit was set to 'high'. Note that only a single bit is activated at each point in time. The resulting decimal code is 2^*N*, where *N* is the identity of the active bit/pin.

and RF mapping tasks, a series of trial-related event codes were sent. Event codes were sent upon stimulus onset and offset and during reward delivery via the same system as that used to send sync pulses during the resting state. Additionally, during the RF task, the stimulus condition used on that particular trial (the direction of bar movement) was sent as an event code. PC #3 also received a copy of the X and Y eye position to check the gaze fixation and determine the success or failure of a trial. Instructions for fluid delivery were then sent to the reward system (Crist Instruments) (Table 3). A summary of the digital codes that were used for the three datasets is provided in Table 2. During post-hoc analysis of the raw data, trial-related events could be identified with high temporal precision (with 30-kHz resolution) and were used for precise temporal alignment of data across NSPs.

Data pre-processing. The datasets are comprised of temporally aligned raw data, as well as data that have been pre-processed to facilitate their usage. The pre-processing steps included the extraction of local field potential (LFP) signals and envelope multiunit activity (MUAe, which represents the aggregation of spiking activity across multiple units recorded via one electrode – details on how we computed MUAe are provided below)[40] from the raw recording traces, and a systematic registration of metadata. These steps were executed after the recording session and implemented into a Python workflow using the Snakemake workflow management system[42]. In addition to this fully integrated workflow, standalone Matlab pre-processing scripts are also included. The metadata integration is only provided based on Python. A full list of data-processing scripts is given in Online-only Table 1. See a schematic description of the data pre-processing workflow in Fig. 5.

Temporal alignment of raw data. The onset and offset of recording were not synchronous across NSPs. Hence, the raw neuronal data were temporally aligned across the files that were generated by the eight NSPs. Excess data at the beginning and end of each file that were not common to all eight NSPs were removed, yielding files of the same duration. Any channels that did not contain neuronal data, i.e. channels 133 or higher on NSP 1 and channel 129 or higher on NSPs 2 to 8, were also removed. The temporally aligned data were saved in the .nev and .ns6 formats. The unaligned raw data files are not provided in the data repository due to their large volume, but are available upon reasonable request. The lightweight events files (.nev) are provided in both their aligned and non-aligned form.

Eye signal processing. In many human studies on resting state activity, subjects are given blindfolds and asked to keep their eyes closed. For the resting state sessions in our subjects, we recorded the pupil size and included it in the dataset, instead of using blindfolds. The eye position and pupil diameter (channels 129 to 132 on NSP #1) were temporally aligned, labelled, and saved in .mat and .nix format (Fig. 5). The baseline value of the recordings containing the pupil diameter was not at 0 mV. Hence, this signal was corrected by subtracting its minimum value within the given session. Additionally, we identified whether the eyes were open or closed, i.e. eye closure. We down-sampled the signals to 1 Hz, to reduce noise and exclude short blinks, and combined the X and Y pupil diameter readings using the Euclidean norm. A low threshold was set and if the combined diameter signal fell below this threshold, we considered the eyes to be closed, otherwise they were considered to be open (Fig. 6). During the recordings, the subjects occasionally exhibited signs of sleepiness and their eyelids drooped for a while, before they closed their eyes completely. Their eyes would sometimes stay closed for minutes at a time. These epochs can be found in the eye data, as extended periods in which the pupil diameter is below threshold. Users may for instance select the time periods during which the monkeys' eyes were closed for a given duration for further analyses of the resting state. We provide both the full (30-kHz) and down-sampled (1-Hz) eye signals in the data repository.

Generation of MUAe and LFP signals from raw data. Following the temporal alignment of raw data across the NSPs, two commonly used types of neuronal signals were extracted from the data (Fig. 5): envelope multiunit activity (MUAe) and local field potentials (LFP).

To generate the MUAe, the raw data were filtered between 0.5 and 9 kHz. A full-wave rectification was performed on the filtered signal, followed by a low-pass filter of 200 Hz. Filtering was carried out using a

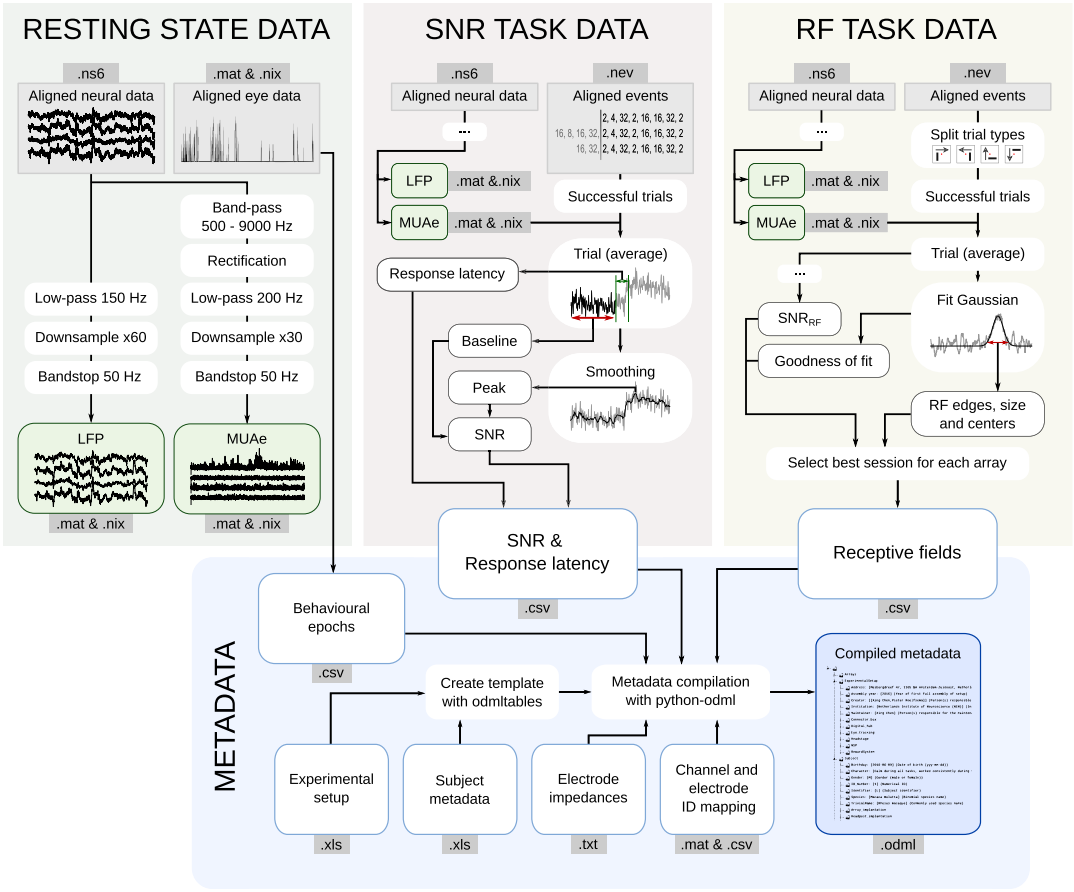


Fig. 5 Data pre-processing diagram. Top: Processing steps for the three datasets, leading to the output data and metadata files. Snippets of the full data are depicted for illustrative purposes. Data alignment precedes the processing steps shown here. Bottom: Integration of metadata into a hierarchical odML file. Metadata were both externally collected (recording apparatus, subject-specific metadata, etc.) and calculated from the recordings (eye signal epochs, RF, SNR). All metadata are integrated into a single odML file per session.

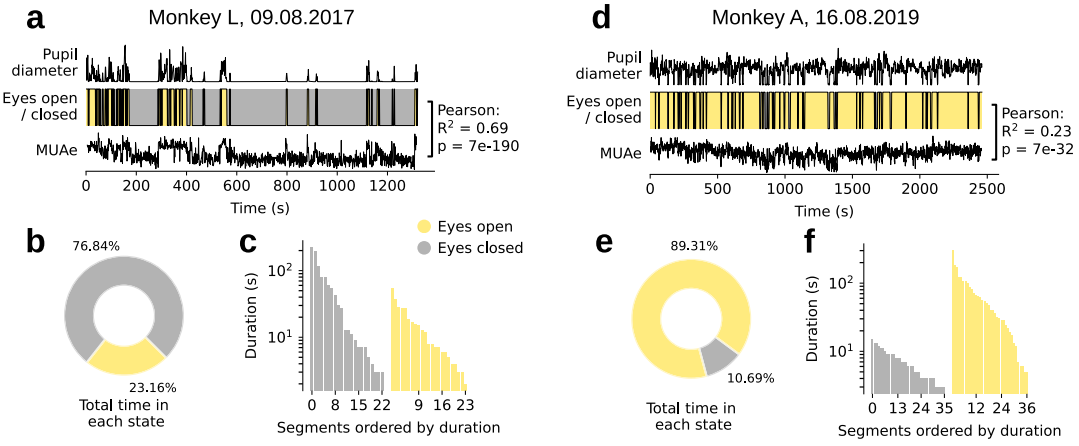


Fig. 6 Overview of pupil diameter during an example resting state session, for each monkey. (**a**,**d**) Traces showing the pupil diameter, the state of the eye ('open' or 'closed') and the mean MUAe from the highest-SNR electrode array. The Pearson correlation between the eye state and MUAe is also shown. (**b**,**e**) Percentage of time spent with eyes open and closed. (**c**,**f**) Bar plot of the duration of time segments that were spent in each state, ordered by duration (segments shorter than 100 ms are likely to be eye blinks and are not shown). Colour coding is identical for all panels.

Butterworth filter, of order 4. The data were down-sampled by a factor of 30, yielding for each original signal an MUAe signal with a sampling rate of 1 kHz.

To generate LFP signals, the raw data were low-pass filtered at 150 Hz (Butterworth filer, order 4) down-sampled to 500 Hz. The newly generated MUAe and LFP signals were saved in the .mat and .nix file format, where each file contains the data from one Utah array with 64 channels.

Signal-to-noise ratio (SNR). To quantify the signal quality of the recorded neuronal activity, the signal-to-noise ratio (SNR) for each channel was calculated based on the amount of visually evoked activity that was elicited upon presentation of a full-screen checkerboard stimulus, relative to baseline activity, across a minimum of 30 trials. We calculated the mean and standard deviation (SD) of the MUAe during the 300-ms time window prior to stimulus onset ($Mean_{spontaneous}$ and $SD_{spontaneous}$ of the baseline activity) for each trial. Next, trial-averaged MUAe data were smoothed with a moving average of 20 bins (i.e. at a sampling rate of 1 kHz, each bin comprised 20 ms), and we identified the peak level of activity elicited during stimulus presentation ($Peak_{stimulus\_evoked}$). The SNR was then calculated following Eq. 1:

$$SNR = \frac{Peak_{stimulus\_evoked} - Mean_{spontaneous}}{SD_{spontaneous}}, \quad (1)$$

A high SNR is indicative of a functional electrode that yields good-quality MUAe. Since the electrodes were located in the visual cortex, they were expected to show responses to visually presented stimuli. A low SNR value may be indicative of one of two situations: 1) For channels where the receptive fields overlap with or are located close to the fixation spot (close to the sulcus between V1 and V4), the presence of the fixation spot in the receptive field may elicit high levels of activity throughout the trial, including during the 'baseline activity' period that precedes stimulus onset. This would result in elevated levels of baseline activity and thereby decrease the SNR. 2) The quality of the signal recorded on that particular electrode may be poor due to factors such as electrode failure, connection failure, poor contact between the electrode and the neuronal tissue and/or excessive tissue gliosis around the electrode.

To select good channels for further analysis of MUAe, we recommend setting a threshold value for the SNR (e.g. 2 or higher) to include only the channels with an SNR value that is above the threshold in subsequent analyses. See Table 1 for the number of high-quality electrodes per session and Fig. 7a,c,e for the SNR values from an example session. The SNR values can be found in the metadata files for the corresponding session (Online-only Table 2).

Neuronal response latency. In addition to the SNR we estimate the neuronal response latency from the checkerboard stimulus task (i.e. the SNR task). We define the response latency as the time elapsed between stimulus onset and the first time that the trial-averaged MUAe signal is more than 2 times the $SD_{spontaneous}$ in 5 consecutive bins. The spontaneous activity period is defined as 300 ms prior to stimulus presentation, as in the SNR calculation. We require the activity to be above the threshold in several consecutive bins to ensure robustness against rapid noise fluctuations. Figure 7b,d,f shows the response latency for a sample session in each monkey. Our measurements are in agreement with previous reports of latency in the visual system[1]. The SNR and latency are calculated together and can be found in the same metadata files (Online-only Table 1).

Receptive field (RF) mapping. We estimated the RF of each electrode using sweeping bar stimuli. The average MUAe was calculated across trials with a given direction of bar motion. A Gaussian was fitted to this trace, and the onset and offset of the visually evoked response were calculated as the times on each trial that corresponded to the midpoint of the Gaussian minus and plus 1.65 times the standard deviation of the Gaussian, respectively. The vertical and horizontal boundaries of the RF on each channel were then calculated as the mean of two values: 1) the spatial location corresponding to the onset time of the response elicited by a bar moving in a particular direction, and 2) the spatial location corresponding to the offset time of the response elicited when the bar moved in the opposite direction[40]. The x- and y-coordinates of the RF centre were taken as the midpoints between the horizontal and vertical boundaries of the RF, respectively, and the RF size was calculated according to the equation:

$$D = \sqrt{(r - l)^2 + (t - b)^2}, \quad (2)$$

where D is the diameter of the RF, r and l are the x-coordinates of the right and left boundaries and t and b are the y-coordinates of the top and bottom boundaries.

The arrays with RFs located closest to the fixation spot (arrays 1 and 4 in monkey L, and 1, 3, 4, 6, 7, 8 and 15 in monkey A) were mapped using a small, thin, slowly moving bar and the other arrays were mapped using a large, thick, fast-moving bar. Note that during stimulus presentation, data were recorded from all the arrays, including from arrays that were not being mapped by the stimulus. Hence, the datasets obtained using the thick and thin bar stimuli were combined into a unified RF map and the remaining data was discarded. The combined RFs for each monkey can be found in the metadata repository (Online-only Table 2).

The RF maps depict the extent of spatial coverage across the visual field. We observed a clear retinotopic organization that matched the locations at which the arrays were implanted on the cortical surface (Fig. 2).

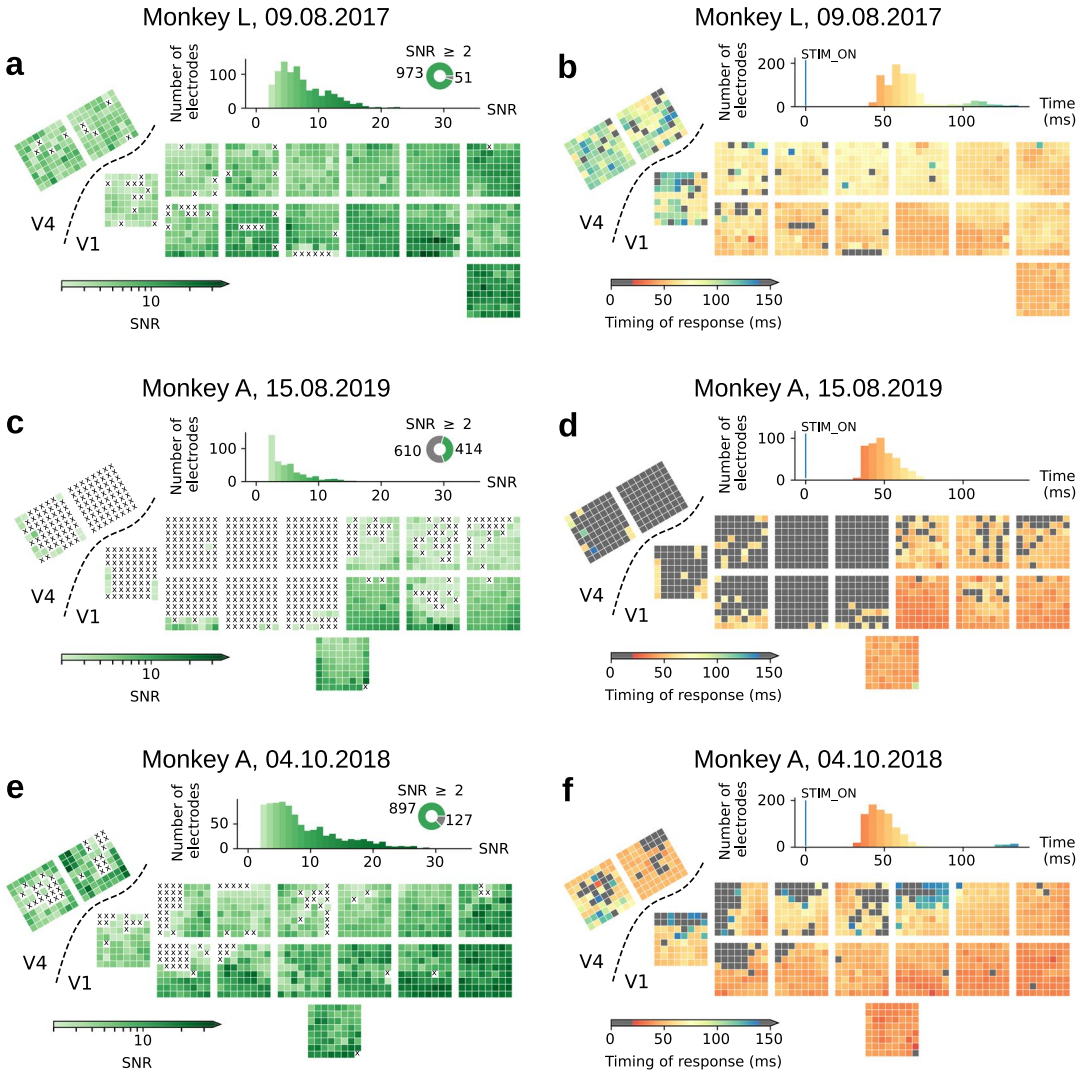


Fig. 7 (**a**,**c**,**e**) Channel signal-to-noise ratio (SNR) from an example session for each monkey, shown on a stylized schematic of the arrays on the cortex. Channels with SNR values < 2 are marked with an 'X'. Top: histogram showing the distribution of SNR values and pie plot showing the proportion of electrodes with SNR > 2. (**b**,**d**,**f**) Stimulus-evoked response timing of two sample SNR sessions, measured as the time at which the trial-averaged MUAe signal exceeds 2 SDs of the baseline (300 ms prior to stimulus onset) for at least 5 consecutive time steps (5 ms in the 1-kHz sampling of MUAe signals). Channels that exhibited a response too early (<20 ms) or too late (>150 ms), or had no response at all, are shown in gray (these channels often had SNR <2). Panels (**a**,**b**) show data from session L_SNR_090817 in monkey L; (**c**,**d**) from session A_SNR_150817 in monkey A; and (**e**,**f**) from session A_SNR_041018 in monkey A. Note that the resting state data in monkey A were recorded several months after surgical implantation of the electrodes, by which time the data quality had decreased (**c**,**d**). Hence, we provide an additional, stand-alone checkerboard stimulation dataset, session A_SNR_041018 (**e**,**f**) which was obtained at an earlier date (when the SNR was high on the majority of channels) and does not have a matching resting state session.

Data Records

Identification of array and channel number. As described in the Methods, of the 16 Utah arrays, 14 were implanted in V1 and 2 in V4. Figure 2a shows the location of implantation in the visual cortex for each of the 16 arrays. Each Utah array consisted of 64 electrodes, and each NSP recorded the signals obtained from two arrays, i.e. 128 channels (see the section, 'Data collection'). Each electrode was assigned a unique global identifier from 1 to 1024. To link the global identity of individual channels with the numbering within an array (out of 64) and the numbering within an NSP (out of 128 channels) we generated look-up tables (LUTs) for each monkey. Each row in the table represents a single electrode. For each electrode the global (out of 1024), within-NSP (out of 128) and within-array (out of 64) channel indices are indicated. Additionally, the NSP number (out of 8), array number (out of 16) and cortical area (V1 or V4) are specified. These tables allow the unique identification of the electrodes across indexing systems.

Raw and aligned neuronal data			
Format	**File naming convention**	**Number of files**	**Description**
.ns6	NSP*X*_aligned.ns6	8 per session	Temporally aligned raw neuronal data files. The length of the data segments is the same across NSPs.
.nev	NSP*X*_aligned.nev	8 per session	Temporally aligned event data files. The duration is the same as in the NS6 files.
	NSP*X*.nev	8 per session	Raw event data files. The duration is the same as in the NS6 files.
Eye signal data			
.mat	aligned_eye_data.mat	1 per resting state session	Eye position in horizontal (X) and vertical (Y) coordinates, and pupil diameter in horizontal (X) and vertical (Y) coordinates. Recorded with a sampling rate of 30 kHz, aligned with the neuronal data.
.nix	aligned_eye_data.nix	1 per resting state session	Eye position in horizontal (X) and vertical (Y) coordinates, and pupil diameter in horizontal (X) and vertical (Y) coordinates. Recorded with a sampling rate of 30 kHz, aligned with the neuronal data. The file includes all relevant metadata in the form of annotation dictionaries.
Processed neuronal data			
.mat	NSP*X*_array*Y*_MUAe.mat	16 per session	Temporally aligned MUAe neuronal data files, with a sampling rate of 1 kHz.
.nix	NSP*X*_array*Y*_MUAe.nix	16 per session	Temporally aligned MUAe neuronal data files, with a sampling rate of 1 kHz. The file includes all relevant metadata in the form of annotation dictionaries and event epochs.
.mat	NSP*X*_array*Y*_LFP.mat	16 per session	Temporally aligned LFP neuronal data files, with a sampling rate of 500 Hz.
.nix	NSP*X*_array*Y*_LFP.nix	16 per session	Temporally aligned LFP neuronal data files, with a sampling rate of 500 Hz. The file includes all relevant metadata in the form of annotation dictionaries and event epochs.

Table 4. Available neuronal and eye signal data files. In the naming conventions, '*X*' represents the NSP number (1 to 8) and '*Y*' represents the array number (1 to 16).

Description of file formats. All data can be found at this GIN repository (https://doi.org/10.12751/g-node.i20kyh)[24]. The raw, aligned data are provided in the proprietary Blackrock format, .ns6.

The pre-processed signals were stored as .mat and .nix (https://g-node.github.io/nix/) files; .nix files can be loaded using the Python Neo framework[43] (https://neuralensemble.org/neo/).

Basic metadata from the recording system are saved in the proprietary Blackrock formats,.nev and.ccf. Note that both .nix and .mat data files (listed in Table 4) contain basic metadata; however, for the complete metadata, the metadata files (listed in Online-only Table 2) should be used. All additional metadata files are provided in various machine- and human-readable formats, such as .txt, .xls, .csv and .mat. Metadata were diverse and originated from different sources, such as the experimental equipment, subject specifications, electrode identifiers, signal quality (SNR), receptive fields (RFs), etc. All metadata were organized into a single unified hierarchical structure, using the open metadata markup language (odML)[44] (https://g-node.github.io/python-odml/), a human- and machine-readable file format for reproducible metadata management in electrophysiology. The raw metadata were processed with odMLtables[45] (https://odmltables.readthedocs.io) and custom Python scripts. The generated metadata files are listed in Online-only Table 2, all of which are integrated into a single odML file per session.

Technical Validation

Impedance measurements. Post-implantation, electrode impedance was measured at 1 kHz using the Impedance Tester function in the Blackrock Central Software Suite. These measurements were carried out in the same month that the resting state data were collected, yielding one text file (.txt) per NSP.

These values were subsequently combined across the eight raw data files (one per NSP), yielding a single. csv file that contains impedance data across all 1024 channels. The impedances were also included in the hierarchically organized odML metadata files.

Eye closure validation. During the resting state sessions, eye pupil diameter and position were tracked using an infrared camera. The monkeys were head fixed throughout the recordings, and the pupil was within sight of the camera at all times, as verified by inspection of the camera feed by an experimenter. To identify the time points with eye closure, we set a threshold for the voltage obtained in the readings for pupil diameter. We further validated this method of threshold setting by comparing levels of cortical activity observed during eye closure and eye opening, and found that activity in the visual cortex was typically higher when the monkeys' eyes were open than when they were closed.

To carry out this validation, we identified the electrode array that yielded the highest signal-to-noise ratios across all 64 electrodes (monkey L: array 11; monkey A: array 10). We calculated the mean MUAe across electrodes on this array, as a measure of on-going neuronal activity. We observed a high correlation between activity levels and the status of the eye, indicating that eye opening was accompanied by an increase in V1 activity (Fig. 6a,d), and verifying the accuracy of the eye closure analysis.

Cross-talk removal. An additional analysis was performed on the resting state data in order to assess whether spurious correlations were present. Unexpectedly high correlations could originate from the induction of current via strong external electromagnetic radiation recorded by the electrodes (e.g. power line noise or

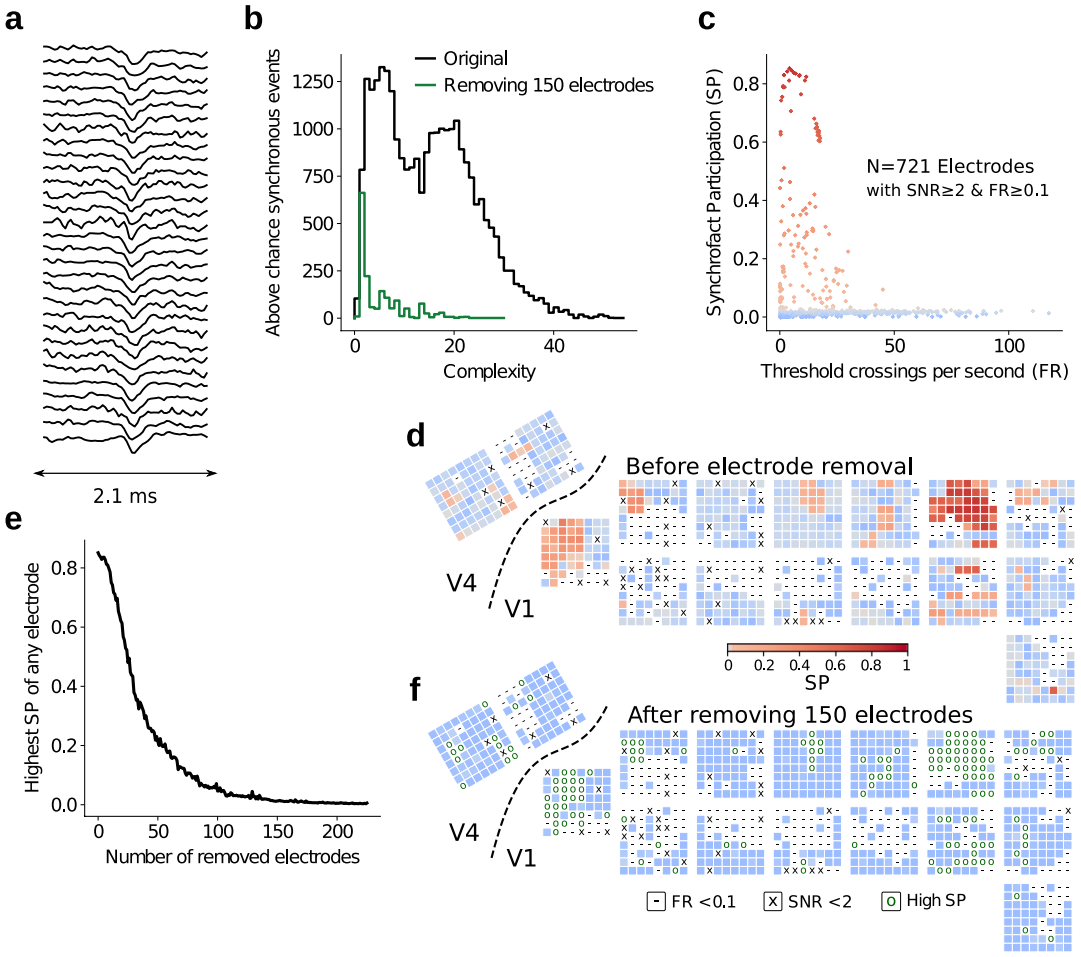


Fig. 8 Detection and removal of high-frequency synchronous events. All plots display data from a single resting state session (L_RS_250717), which was the session with the most cross talk. (**a**) Raw signal of a sample synchrofact (complexity = 30). (**b**) Above-chance part of the complexity histogram for the original data (black) and after removing 150 electrodes with the highest synchrofact participation (SP, green). (**c**) Scatterplot of SP of each electrode versus firing rate (FR, calculated as threshold crossings per second). Each point represents a single electrode. As shown by the absence of a positive correlation, the SP was not biased by the FR, due to our method of normalising the SP by the total number of synchronous events. (**d**) Synchrofact participation of each electrode in the original data (prior to removal of 150 electrodes with high SP). Crosses (X) indicate electrodes with SNR < 2 and hyphens (−) indicate electrodes with FR < 0.1; these electrodes were excluded from the cross-talk analysis. (**e**) For the remaining electrodes, those with a high SP were systematically removed, accompanied by a decrease in the largest SP that was observed across the remaining electrodes. (**f**) SP of each electrode after the removal of 150 electrodes with high SP; removed electrodes are indicated by a circle (O). (**c**) (**d**) and (**f**) use the same colour map.

telecommunications devices) or could be caused by electrical short circuits between two or more electrodes, i.e. cross-talk.

Cross-talk can arise when electrodes physically touch each other due to mechanical bending during or after surgery, or when currents arise between cables due to a breach in electrical insulation of channels at any point along the processing stream up to the conversion of data from analog to digital. We examined the data for spurious correlations, as they could potentially contaminate MUAe signals. Determining the precise source of such spurious correlations is beyond the scope of this publication.

In order to detect cross-talk artifacts in the data, we band-passed the raw signals at 250–9000 Hz. We then removed the first principal component for the channels that shared a common reference wire, roughly corresponding to the mean of the signals. Threshold crossing events were extracted as described by Quiroga *et al.*[46], with a threshold multiplier parameter of 5. Next, we counted synchronous threshold crossings at sampling resolution and in adjacent sampling bins (1/30 ms). The complexity of a synchronous event was defined as the number of near-simultaneous threshold crossings across electrodes[47]. These synchronous events could also occur in the data by chance. However, in some cases, the analog signals had the same shape across numerous electrodes, indicating that these synchronous events were likely artifacts (see Fig. 8a for an example). These non-random synchronous events are termed 'synchrofacts' (short for synchronous artifacts)[47].

Distinguishing synchrofacts from randomly occurring synchronous events is not trivial, due to the large number (up to hundreds of thousands) of synchronous events, hence we examined their complexity histogram, i.e. the number of detected synchronous events of a given complexity. We used a one-sided Monte Carlo permutation test to check whether the number of synchronous events of a given complexity was above chance level. We generated surrogate data (N=1000 surrogates), i.e. permutations, by uniformly dithering ($\pm$5 ms) the timing of the threshold crossings. The probability of obtaining a certain number of synchronous events by chance was estimated based on the distribution of events in the surrogate data. We found thousands of synchrofacts across complexity values ranging from 1 to >50 (Fig. 8b), far more than what would be expected by chance.

To distinguish synchrofacts from randomly occurring synchronous events, and to pinpoint the electrodes that were primarily responsible for the non-random events, complexity histograms were then calculated on an electrode-by-electrode basis. To provide a measure of the number of synchrofacts obtained per electrode, i.e. threshold crossings per second (Fig. 8c), we first tallied the number of above-chance threshold-crossing events for a given electrode. Electrodes with higher firing rates inevitably yield a larger number of randomly occurring synchronous events; to correct for this bias, we divided the number of above-chance events on each electrode by the total number of events seen on that electrode. We call this metric the 'synchrofact participation' (SP) of the electrode.

$$SP = \frac{\sum N_{aboveChance}}{N_{total}}, \tag{3}$$

where N denotes the number of synchronous events observed for a given electrode.

The SP takes a value between 0 and 1, indicating the proportion of synchronous events that were above chance for each electrode. Correcting by the total number of events leads to a measure of the synchrofacts per electrode that is not correlated to the firing rate. When mapping the cortical locations of the electrodes with high SP values, we found that they were grouped into several clusters (Fig. 8d). We detected synchronous artifacts in all three resting state sessions from monkey L. Resting state data from monkey A did not show large numbers of synchrofacts (likely due to the low firing rates obtained in those sessions).

The simplest approach to removing cross-talk from the data is to discard the electrodes with high SP from further analysis. We systematically removed electrodes with the highest SP one by one, and recalculated the SP and chance levels after the removal of each electrode. Our significance level was not adjusted for multiple comparisons, as this would lead to a high false negative rate and hinder the removal of electrodes with cross-talk from the dataset. We removed up to 250 electrodes (Fig. 8e), greatly reducing the levels of cross-talk in the data (Fig. 8f). Note that while the electrode removal process eliminates a large portion of the spurious correlations, it does not eliminate artifacts that occur sporadically or at low rates on a given electrode.

We provide a recommended list of electrodes to discard, and the order of their removal. The precise number of electrodes to be discarded can be adjusted as needed by data users, depending on the particular use case. Detailed plots depicting the complexity histograms and electrode SP for all the resting state sessions are available in our data repository.

A reference implementation for synchrofact detection was included in version 0.10.0 of the Electrophysiology Analysis Toolkit[26] (Elephant, RRID:SCR_003833, https://elephant.readthedocs.io). The full workflow used for systematic electrode removal was implemented with Python and can be found in the data respository.

Code availability

All scripts used for the processing of data and our preliminary analyses are available alongside the data at https://gin.g-node.org/NIN/V1_V4_1024_electrode_resting_state_data, in the 'code' folder. All experiment control scripts are listed in Table 2 and the data processing scripts are listed in Online-only Table 1.

Matlab version R2015b, Python version 3.7 and Snakemake version 5.8.1 were used. The only Matlab dependency was the NPMK toolbox (version 5.0, Blackrock Microsystems), a copy of which is included in the data repository.

Direct Python dependencies include neo 0.9.0, nixio 1.5.0 elephant 0.10.0, odml 1.4.5 and odmltables 1.0. A full list of all Python dependencies and the specific versions used can be found in the Python environment specifications, along with the Python scripts (Online-only Table 1).

Received: 18 May 2021; Accepted: 27 January 2022;
Published online: 11 March 2022

References

1. Nowak, L. G., Munk, M. H. J., James, A. C., Girard, P. & Bullier, J. Cross-Correlation Study of the Temporal Interactions Between Areas V1 and V2 of the Macaque Monkey. *J Neurophysiol* **81**, 1057–1074 (1999).
2. Leopold, D. A. Very Slow Activity Fluctuations in Monkey Visual Cortex: Implications for Functional Brain Imaging. *Cerebral Cortex* **13**, 422–433 (2003).
3. Bollimunta, A., Chen, Y., Schroeder, C. E. & Ding, M. Neuronal Mechanisms of Cortical Alpha Oscillations in Awake-Behaving Macaques. *Journal of Neuroscience* **28**, 9976–9988 (2008).
4. Buffalo, E. A., Fries, P., Landman, R., Buschman, T. J. & Desimone, R. Laminar differences in gamma and alpha coherence in the ventral stream. *PNAS* **108**, 11262–11267 (2011).
5. Maier, A., Adams, G. K., Aura, C. & Leopold, D. A. Distinct Superficial and Deep Laminar Domains of Activity in the Visual Cortex during Rest and Stimulation. *Front. Syst. Neurosci.* **4**, 31 (2010).
6. Wang, Z. *et al.* The Relationship of Anatomical and Functional Connectivity to Resting-State Connectivity in Primate Somatosensory Cortex. *Neuron* **78**, 1116–1126 (2013).
7. Raichle, M. E. & Mintun, M. A. Brain work and brain imaging. *Annu. Rev. Neurosci.* **29**, 449–476 (2006).

8. Thomas Yeo, B. T. *et al*. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J Neurophysiol* **106**, 1125–1165 (2011).
9. Heinzle, J., Kahnt, T. & Haynes, J.-D. Topographically specific functional connectivity between visual field maps in the human brain. *NeuroImage* **56**, 1426–1436 (2011).
10. Smith, S. M. *et al*. Temporally-independent functional modes of spontaneous brain activity. *PNAS* **109**, 3131–3136 (2012).
11. Genç, E., Schölvinck, M. L., Bergmann, J., Singer, W. & Kohler, A. Functional Connectivity Patterns of Visual Cortex Reflect its Anatomical Organization. *Cerebral Cortex* **26**, 3719–3731 (2016).
12. Raichle, M. E. & Gusnard, D. A. Intrinsic brain activity sets the stage for expression of motivated behavior. *J. Comp. Neurol.* **493**, 167–176 (2005).
13. Wang, K. *et al*. Spontaneous Activity Associated with Primary Visual Cortex: A Resting-State fMRI Study. *Cerebral Cortex* **18**, 697–704 (2008).
14. Nir, Y. *et al*. Interhemispheric correlations of slow spontaneous neuronal fluctuations revealed in human sensory cortex. *Nat Neurosci* **11**, 1100–1108 (2008).
15. Raemaekers, M. *et al*. Patterns of resting state connectivity in human primary visual cortical areas: A 7T fMRI study. *NeuroImage* **84**, 911–921 (2014).
16. Striem-Amit, E. *et al*. Functional connectivity of visual cortex in the blind follows retinotopic organization principles. *Brain* **138**, 1679–1695 (2015).
17. Hutchison, R. M., Womelsdorf, T., Gati, J. S., Everling, S. & Menon, R. S. Resting-state networks show dynamic functional connectivity in awake humans and anesthetized macaques. *Hum. Brain Mapp.* **34**, 2154–2177.
18. Contreras, D. & Steriade, M. Cellular basis of EEG slow rhythms: a study of dynamic corticothalamic relationships. *Journal of Neuroscience* **15**, 604–622 (1995).
19. Juavinett, A. L., Bekheet, G. & Churchland, A. K. Chronically implanted Neuropixels probes enable high-yield recordings in freely moving mice. *eLife* **8**, e47188 (2019).
20. Jun, J. J. *et al*. Fully integrated silicon probes for high-density recording of neural activity. *Nature* **551**, 232–236 (2017).
21. Steinmetz, N. A., Koch, C., Harris, K. D. & Carandini, M. Challenges and opportunities for large-scale electrophysiology with Neuropixels probes. *Neurotechnologies* **50**, 92–100 (2018).
22. Brochier, T. *et al*. Massively parallel recordings in macaque motor cortex during an instructed delayed reach-to-grasp task. *Sci Data* **5**, 180055 (2018).
23. Chen, X., Wang, F., Fernández, E. & Roelfsema, P. R. Shape perception via a high-channel-count neuroprosthesis in monkey visual cortex. *Science* **370**, 1191 (2020).
24. Chen, X. *et al*. 1024-channel electrophysiological recordings in macaque V1 and V4 during resting state. *G-Node* https://doi.org/10.12751/g-node.i20kyh (2021).
25. Schmidt, M. *et al*. A multi-scale layer-resolved spiking network model of resting-state dynamics in macaque visual cortical areas. *PLoS Comput Biol* **14**, e1006359 (2018).
26. Denker, M. *et al*. LFP beta amplitude is linked to mesoscopic spatio-temporal phase patterns. *Scientific Reports* **8**, 5200 (2018).
27. Roelfsema, P. R., Engel, A. K., König, P. & Singer, W. Visuomotor integration is associated with zero time-lag synchronization among cortical areas. *British Journal of Ophthalmology* **385**, 157–161 (1997).
28. Bock, A. S. *et al*. Resting-State Retinotopic Organization in the Absence of Retinal Input and Visual Experience. *The Journal of Neuroscience* **35**, 12366–12382 (2015).
29. Gravel, N. *et al*. Cortical connective field estimates from resting state fMRI activity. *Frontiers in Neuroscience* **8**, 1–10 (2014).
30. Hinds, O. P. *et al*. Accurate prediction of V1 location from cortical folds in a surface coordinate system. *NeuroImage* **39**, 1585–1599 (2008).
31. Benson, N. C. *et al*. The Retinotopic Organization of Striate Cortex Is Well Predicted by Surface Topology. *Current Biology* **22**, 2081–2085 (2012).
32. Benson, N. C., Butt, O. H., Brainard, D. H. & Aguirre, G. K. Correction of Distortion in Flattened Representations of the Cortical Surface Allows Prediction of V1-V3 Functional Organization from Anatomy. *PLoS Comput Biol* **10**, e1003538–9 (2014).
33. Klink, P. C., Chen, X., Vanduffel, W. & Roelfsema, P. R. Population receptive fields in nonhuman primates from whole-brain fMRI and large-scale neurophysiology in visual cortex. *eLife* **10**, e67304 (2021).
34. Roberts, M. J. *et al*. Robust Gamma Coherence between Macaque V1 and V2 by Dynamic Frequency Matching. *Neuron* **78**, 523–536 (2013).
35. Gray, C. M., Engel, A. K., König, P. & Singer, W. Stimulus-Dependent Neuronal Oscillations in Cat Visual Cortex: Receptive Field Properties and Feature Dependence. *European Journal of Neuroscience* **2**, 607–619 (1990).
36. Engel, A. K., König, P., Gray, C. M. & Singer, W. Stimulus-Dependent Neuronal Oscillations in Cat Visual Cortex: Inter-Columnar Interaction as Determined by Cross-Correlation Analysis. *European Journal of Neuroscience* **2**, 588–606 (1990).
37. Frien, A., Eckhorn, R., Bauer, R., Woelbern, T. & Kehr, H. Stimulus-specific fast oscillations at zero phase between visual areas V1 and V2 of awake monkey. *NeuroReport* **5**, 2273–2277 (1994).
38. Wilkinson, M. D. *et al*. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* **3**, 1–9 (2016).
39. Chen, X. *et al*. 3D printing and modelling of customized implants and surgical guides for non-human primates. *Journal of Neuroscience Methods* **286**, 38–55 (2017).
40. Supèr, H. & Roelfsema, P. R. in *Progress in Brain Research* **147**, 263–282 (Elsevier, 2005).
41. Tootell, R. B., Switkes, E., Silverman, M. S. & Hamilton, S. L. Functional anatomy of macaque striate cortex. II. Retinotopic organization. *J Neurosci* **8**, 1531 (1988).
42. Köster, J. & Rahmann, S. Snakemake–a scalable bioinformatics workflow engine. *Bioinformatics* **28**, 2520–2522 (2012).
43. Garcia, S. *et al*. Neo: an object model for handling electrophysiology data in multiple formats. *Front. Neuroinform.* **8**, 10 (2014).
44. Grewe, J., Wachtler, T. & Benda, J. A Bottom-up Approach to Data Annotation in Neurophysiology. *Front. Neuroinform.* **5**, 16–16 (2011).
45. Sprenger, J. *et al*. odMLtables: A User-Friendly Approach for Managing Metadata of Neurophysiological Experiments. *Front. Neuroinform.* **13**, 14563 (2019).
46. Quiroga, R. Q., Nadasdy, Z. & Ben-Shaul, Y. Unsupervised Spike Detection and Sorting with Wavelets and Superparamagnetic Clustering. *Neural Comput* **16**, 1661–1687 (2004).
47. Torre, E. *et al*. Synchronous Spike Patterns in Macaque Motor Cortex during an Instructed-Delay Reach-to-Grasp Task. *J Neurosci* **36**, 8329 (2016).

Acknowledgements

We thank Kor Brandsma and Anneke Ditewig for technical support; Matthew Self and Feng Wang for assistance during surgeries; Chris Klink for help with technical validation; John van Veldhuizen, Stephen Super, Joop Bos, Joost Brand, and Ruud van der Blom, for help with mechanical engineering; Cyril Voisard (Medicoat) for biocompatible coating of the implants; Florian Solzbacher, Marcus Gerhardt, Nick Halper, Stephen Hou, Rob Franklin, Saman Hagh-Gooie, Kian Torab, Sherman Wiebe, Charles Dryden, Vinh Ngo, William Yang, Greg

Palis, Mike Gruenhagen, and others at Blackrock Microsystems for scientific and technical collaborations; and Sebastian Lehmann for assistance with graphic design. Funding: This work was supported by NWO (STW grant number P15-42 'NESTOR' and Crossover grant number 17619 'INTENSE'), the European Union FP7 (ERC grant number 339490 'Cortic_al_gorithms'), the European Union Horizon 2020 Framework Programme for Research and Innovation under the Framework Partnership (HBP FPA agreement number 650003), the H2020 Framework Programme for Research and Innovation (Human Brain Project SGA2 grant number 785907, and HBP SGA3 grant number 945539), the European Union Horizon 2020 Future and Emerging Technologies (FET Open grant number 899287 'NeuraViPeR'), and the Deutsche Forschungsgemeinschaft (German Research Foundation, grant number 368482240/ RTG 2416).

Author contributions

X.C. and P.R. designed the study. X.C. set up the experimental setup, ran the experiments, and collected the data. X.C. and A.M.G. gathered the metadata and wrote the data processing scripts, and set up the data processing workflow. A.M.G. and J.S. set up the data repository. A.M.G. and A.K. designed and implemented the cross-talk removal. J.S., A.K. and S.S. supervised code implementation and best practices in data management. X.C., A.M.G. and J.S. wrote the manuscript, and A.K., S.S., S.v.A., S.G. and P.R. provided feedback. S.v.A. and S.G. supervised the data management. P.R. and X.C. supervised and coordinated the project.

Competing interests

P.R. and X.C. are co-founders and shareholders of a neurotechnology start-up, Phosphoenix (Netherlands) (https://phosphoenix.nl).

Additional information

Correspondence and requests for materials should be addressed to X.C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

To deal with a 14-dimensional space, visualize a 3-D space and say 'fourteen' to yourself very loudly. Everyone does it.

Geoffrey Hinton

2

State space analysis of visual cortex activity

Summary:

In this chapter we analyze the resting state data from the macaque visual cortex. Our analysis reveals that resting-state neural manifolds of macaque V1 are organized as two distinct high-dimensional clusters. We show that these clusters are primarily correlated with the behavior (eye closure) of the macaques and that the dimensionality of each of these clusters is significantly different, with higher dimensionality during the eyes-open periods. In addition, we use LFP coherence and Granger causality to estimate signatures of feedback from V4 and DP to V1 (in the beta range) and find that feedback signatures are significantly stronger during the eyes-open periods. Finally, we simulate a spiking neuron model under resting-state conditions and show that feedback signals can modulate the state space and dimensionality of our model: distinct clusters and dimensionality arise in the presence and absence of feedback. Taken together, the data analysis and simulations suggest that feedback signals actively modulate neural manifolds and dimensionality in the visual cortex of the macaque.

Feedback modulates neural manifolds and dimensionality in macaque V1 during resting state

Aitor Morales-Gregorio[1,2*], Anno C. Kurth[1,3], Junji Ito[1], Alexander Kleinjohann[1,4], Frédéric V. Barthélemy[1,5], Thomas Brochier[5], Sonja Grün[1,4] and Sacha J. van Albada[1,2]

[1]Institute of Neuroscience and Medicine (INM-6) and Institute for Advanced Simulation (IAS-6) and JARA-Institut Brain Structure-Function Relationships (INM-10), Jülich Research Centre, Jülich, Germany.
[2]Institute of Zoology, University of Cologne, Cologne, Germany.
[3]RWTH Aachen University, Aachen, Germany.
[4]Theoretical Systems Neurobiology, RWTH Aachen University, Aachen, Germany.
[5]Institut de Neurosciences de la Timone (INT), CNRS & Aix-Marseille Université, Marseille, France.

*Corresponding author. a.morales-gregorio@fz-juelich.de

Abstract

High-dimensional brain activity is often organised into lower-dimensional neural manifolds, which can represent a plethora of behavioural variables. In macaque V1, feedback from V4 is known to mediate attention for figure-ground segregation and contour integration. However, whether feedback signals can modulate neural manifolds in the brain is not known. To address this gap in the knowledge we study extracellular recordings of macaque areas V1, V4 and DP with unprecedented spatio-temporal resolution. We found multiple clusters in the V1 neural manifolds, which correlated strongly to eye closure (eyes-open/closed) and displayed distinct dimensionality. Moreover, we found feedback from V4 to V1 in the form of beta band spectral Granger causality, which was significantly stronger during the eyes-open periods. Finally, we reproduced the experimental results via numerical simulations of a balanced spiking neuron network. Taken together, our analyses and simulations show that feedback signals modulate the V1 activity. This finding suggests that feedback signals are key in the modulation of population activity across cortex.

Keywords: Neural manifolds, electrophysiology, macaque, visual cortex, feedback, resting state

Introduction

The brain can be described as a high-dimensional dynamical system capable of representing and processing a plethora of low-dimensional variables.

The dynamics of a population of neurons can be considered as a trajectory in a high-dimensional space, where each neuron represents one dimension; i.e., the state space of the neural system. Typically, the system does not visit all possible states in the state space, but rather remains confined to a small subset. This subset of state space is referred to as a neural manifold [1–4]—when the high-dimensional geometry is Euclidean everywhere—or sometimes as an attractor [5]—when the activity of the dynamical system tends to a certain part of the phase space for many different starting values. Neural manifolds have been shown to encode aspects such as decision-making in the prefrontal cortex of macaque [6], hand movement trajectories in the motor cortex of macaque [2, 3, 7], odour in the piriform cortex of mice [8], head direction in the anterodorsal thalamic nucleus of mice [5], and spatial position in the hippocampus of mice [9]. However, the study of neural manifolds in the visual cortex is, to the best of our knowledge, limited to mice [10, 11] and small samples of neurons in macaque [12].

Neural manifolds often have an intricate topology, which can be studied using methods borrowed from computational topology [5, 12, 13]. In addition to the topology, the number of uncorrelated covariates required to capture the variance in the state space is studied, the so-called dimensionality of a neural system [1, 10, 14–19]. Regardless of species and brain area, the dimensionality is drastically lower than the total number of recorded neurons [1], suggesting robust encoding of low-dimensional variables. Stringer et al. [10] showed that the dimensionality of the visual cortex in mice can vary dynamically to encode precise visual input, that the state space correlations obeyed a power law spectrum, and that the visual cortex neural manifolds are therefore not smooth.

Eye closure—whether the macaque had its eyes open or closed—is known to affect the activity in the visual cortex, even in darkness [20–24]. However, whether this behaviour influences the neural manifolds has not been shown.

The primary visual cortex (V1) is known to represent intricate details from the visual field both at a single neuron and population level [12, 25]. The visual system is hierarchical in nature, with information travelling from lower to higher areas (feedforward, FF) and vice-versa (feedback, FB), within known frequencies [26–28]. Feedback from V4 to V1 is known to mediate visual attention for figure-ground segregation and contour integration in macaque [29–32]. Recent evidence suggests that feedback signals can modulate neural manifold geometry and their dimensionality. On the one hand, Naumann et al. [33] showed *in silico* that feedback can rotate neural manifolds to modulate attention in a context-dependent manner. On the other hand, Dahmen et al. [34] show that recurrent connectivity motifs—not feedback explicitly—modulate

the dimensionality of the cortical activity. Thus, whether feedback signals modulate the neural manifold geometry and dimensionality *in vivo* remains to be shown.

Here, we study the neural manifolds of the primary visual cortex of the macaque (N=3) during resting state and its relation to the feedback signals from higher visual areas. We find that the neural manifolds of macaque V1 are organised as two distinct high-dimensional clusters, correlated with the behaviour (eye closure) of the macaques, and only partially correlated with higher activity levels. The dimensionality within each of these clusters is significantly different, with higher dimensionality found during the eyes-open periods than the eyes-closed periods. In addition, we estimated feedback from higher cortical areas to V1 and found that feedback signals (in the beta frequency range) are significantly stronger during the eyes-open periods, suggesting they play a role in modulating the neural manifolds. Finally, we simulate a spiking model under resting state conditions and show that feedback signals can modulate the state space of our model. Taken together, the data analysis and simulations show that feedback signals actively modulate neural manifolds, leading to two distinct clusters of the macaque visual cortex activity.

Results

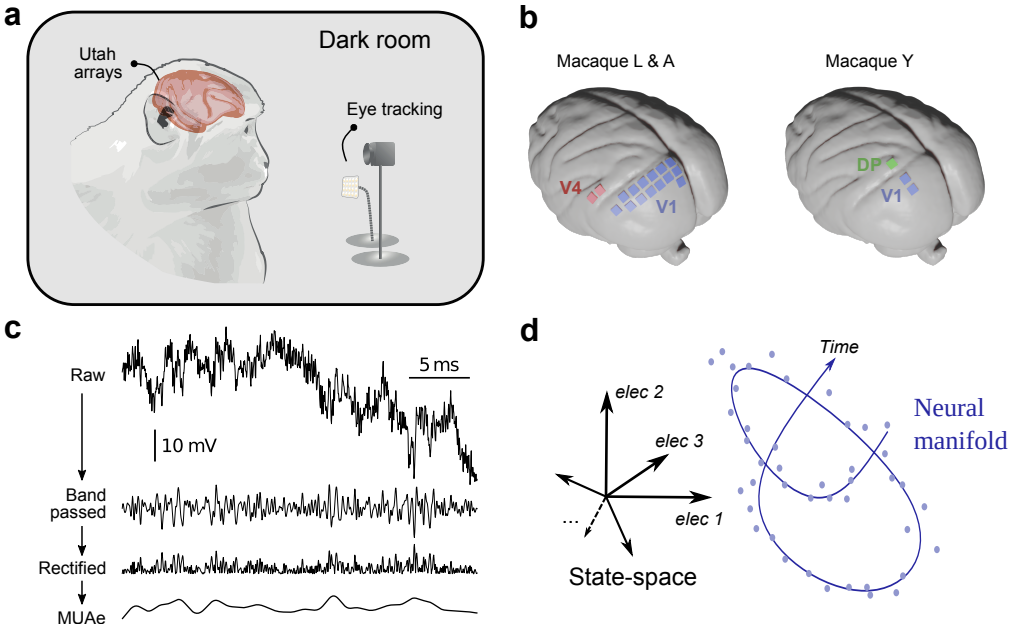


Fig. 1 Overview of the experiment and neural manifold construction. **a** Illustration of the experimental setup. **b** Approximate locations of array implants in both experiments. Exact placement of the arrays varies slightly for the subjects within a single experiment. **c** Steps used to obtain the multi-unit activity envelope (MUAe) [35], used in this study. Band pass filtering performed between 500 Hz and 9 kHz, rectified signal is low-passed at 200 Hz to obtain the MUAe. **d** Schematic representation of state space and a neural manifold. Note that time is implicit within the neural manifold.

To explore the activity in the visual cortex, the electrical potential from V1 cortex of three rhesus macaques was recorded. The experiments simultaneously recorded the activity from V1 and V4 (macaques L & A) [36] and from V1 and DP (macaque Y, see Figure 1b) [37]. The recordings were made in the resting state, i.e., the macaques sat head-fixed in a dark room and were not instructed to perform any particular task. In this state the macaques often showed signs of sleepiness and kept their eyes closed for long periods of time. The right eye—contralateral to the site of neural recording—was tracked using an infrared camera, allowing the identification of periods of open or closed eyes, linked to the state of awareness of the macaque. See methods Electrophysiological data from macaque L & A and Electrophysiological data from macaque Y for further details on the data acquisition and processing. The experimental setup and data processing are illustrated in Figure 1.

Two distinct high-dimensional clusters in V1 during resting state

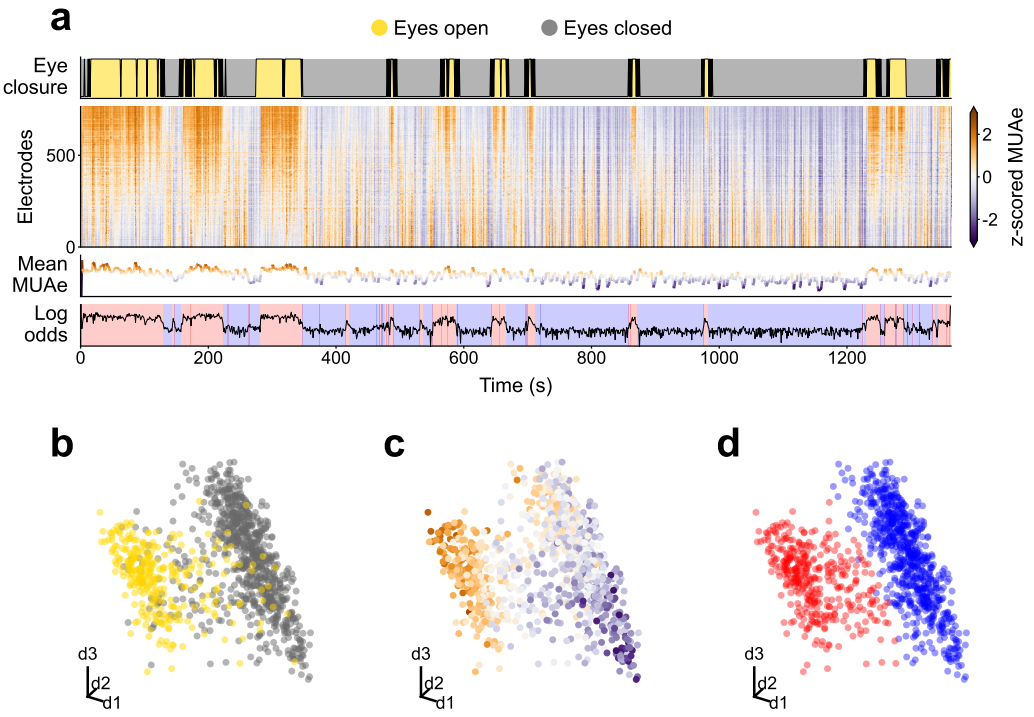


Fig. 2 Overview of the experimental data from session L_RS_250717. **a** From top to bottom: Time evolution of the eye signal; the z-scored MUAe signal for each electrode (electrodes ordered by their correlation to the eye signal); the mean z-scored MUAe at each time point; and the log odds overlaid with the most likely cluster identity (two clusters, Gaussian mixture model). **b, c, d** Three-dimensional PCA embedding of the MUAe neural manifold. Colours indicate eye closure (**b**), mean z-scored MUAe (**c**) and the clusters identified via the log odds of a Gaussian mixture (**d**). Each dot represents a different point in time. Outliers in the neural manifolds shown in **b-d** were excluded, see Outlier removal.

To explore the activity of the visual cortex, we constructed a high-dimensional (between 64 and 800 electrodes, see Table 1 for details) neural manifold for each area and macaque from the downsampled (1 Hz) multi-unit activity envelope (MUAe) [35], Figure 2a. We embedded the manifolds into a 3D space for visualisation using principal component analysis (PCA), Figure 2b-d—See supplementary videos for an animated visualisation of the 3D manifolds.

In V1, at least two distinct clusters are apparent in the 3D-embedded manifold, which we labelled via the log odds of a two-component Gaussian mixture model (see methods Neural manifolds and clustering and Outlier removal). The log odds represent the likelihood for a given data point to correspond to one cluster or the other. For the rest of this paper we consider points with

positive log odds to belong to cluster 0 (colour-coded in blue throughout the paper) and points with negative log odds to belong to cluster 1 (colour-coded in red throughout the paper).

To confirm that the two clusters in the lower-dimensional embedding are not an artefact of the dimensionality reduction, we estimated the homology groups of the high-dimensional manifold with persistent homology (see methods Topological data analysis). The persistence barcodes show that at least two clusters exist in the high-dimensional neural manifold (Figure S7). Thus, we confirmed that the clusters observed in 3D are inherent to the high-dimensional manifolds and not an artefact of the embedding.

For completeness, we also visualised the neural manifolds from V4 and DP (Figure S8, S9). In contrast to V1, the neural manifolds in areas V4 and DP do not appear to have two clear clusters.

Clusters are strongly correlated with eye closure, and less strongly correlated with activity levels

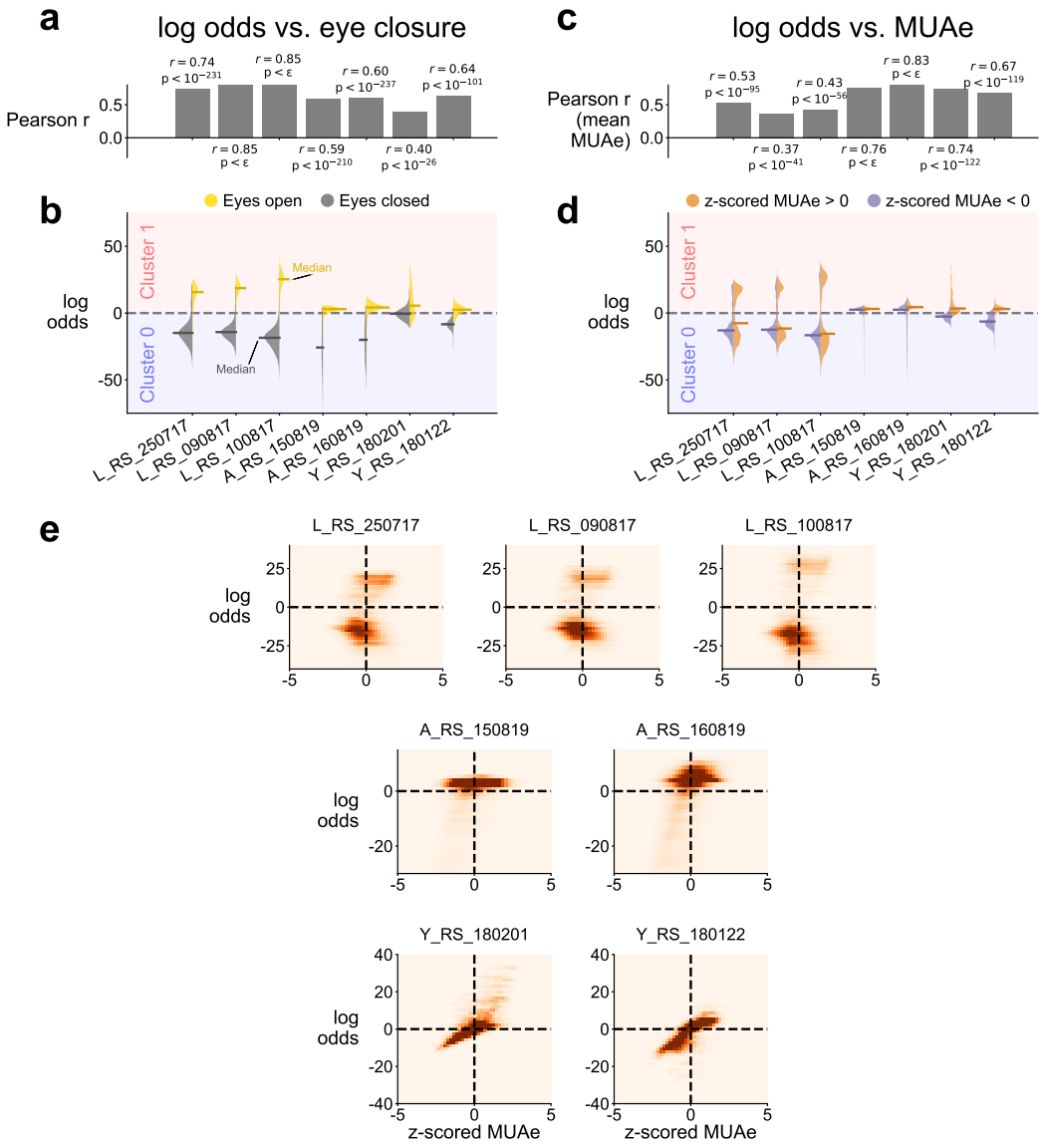


Fig. 3 Clusters are correlated with eye closure and mean activity. **a, c** Pearson correlation of eye closure (**a**) and mean MUAe (**c**) with log odds; r and p-value shown. **b, d** Violin plots of the distribution of the log odds across epochs, respectively distinguished according to eye closure (**b**) and z-scored MUAe (**d**). Note that **c** was calculated based on the mean MUAe across electrodes, while **d** shows the distribution of the full z-scored MUAe considering all electrodes separately. Horizontal bars indicate medians of the distributions. **e** 2D histograms of z-scored MUAe and log odds. Darker colour indicates higher occurrence.

While the activity of visual cortex is clearly modulated by visual input, whether and to what extent it is separately modulated by eye closure is unclear. Marking data points on the manifold with the eye closure state (Figure 2b) reveals that one cluster strongly relates to eyes-open periods, whereas the other cluster mainly related to eyes-closed periods.

To confirm the correlation between eye closure and cluster identity we tested the differences between the eyes-open and eyes-closed states using a twofold approach. First, we computed the Pearson correlation between the eye closure signal and the log odds, revealing a significantly higher than chance correlation in all sessions (Figure 3a). Second, we visualised the distribution of the log odds during the eyes-open and eyes-closed periods separately; showing a clear correspondence between the eye closure state and the sign of the log odds in most cases (Figure 3b). Taken together, the Pearson correlation and the log odds distributions demonstrate that the two neural clusters strongly correlate with eye closure.

The existence of two separate clusters could be trivially explained if the MUAe activity levels were significantly different between the eyes-open and eyes-closed periods, and the cluster identity simply reflected the population activity level. To examine this possibility, we tested whether higher-activity epochs uniquely correspond to one of the clusters in the data. We tested the correlation of the mean MUAe (across electrodes) with the log odds, yielding relatively high values (Figure 3c). However, the violin plots of the full data distribution—based on the z-scored MUAe shown in Figure 2a—show that there is no clear separation into two clusters (Figure 3d). Additionally, we visualised the 2D histograms of z-scored MUAe against log odds (Figure 3e). If the 2D histograms showed most occurrences only in the top right quadrant (high activity in cluster 1) and the bottom left quadrant (low activity in cluster 0) the clusters would be easily explained by increased activity rates alone. However, in all sessions we observe some occurrences in the top left quadrant (low activity in cluster 1) and/or the bottom right quadrant (high activity in cluster 0). Thus, although cluster identity (log odds) and activity levels (MUAe) are correlated to some extent, we conclude based on the violin plots and 2D histograms that the activity level alone does not fully explain the presence of the state-space clusters.

Finally, we also tested the eye closure preference in V4 and DP (Figure S10), using the same procedure as for V1. Although some correlation was observed between eye closure and log odds, the violins reveal no clear cluster separation. Thus, we find that the observed clustering is restricted to V1 and is not present in the other two areas.

In this section we have seen that the neural manifold clusters correlate with eye closure and mean MUAe, but not with full z-scored MUAe. Slightly higher activity levels may be expected during the eyes-open periods, due to the residual lights present in the otherwise dark room, which we observe for the mean MUAe (Figure 3c). However, the more nuanced non-averaged MUAe distribution does not show a clear separation into two clusters (Figure 3d).

Our analysis of the z-scored MUAe suggests that the neural manifold clusters do not exclusively depend on external stimuli.

Dimensionality differs between the clusters

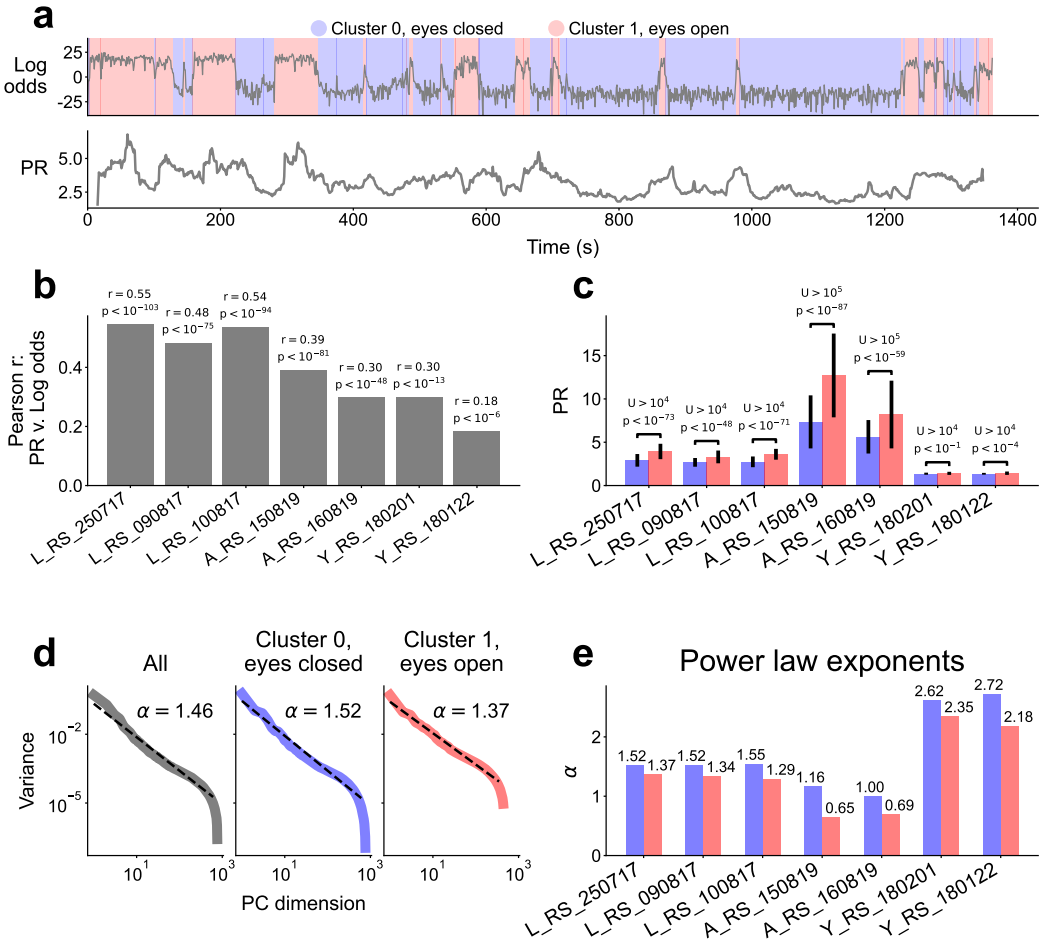


Fig. 4 **a** Decision function and participation ratio (PR) over time for session L_RS_250717. The PR was calculated on a sliding window of 30 s width. **b** Pearson correlation between decision function and PR. **c** Comparison of PR between clusters (Mann-Whitney U test). **d** Distribution of principal components and their explained variance on a log-log scale, respectively for the entire session (L_RS_250717) and for each cluster. Power law exponent α estimated over the ranges where the curves approximate a power law. **e** Comparison of power law exponents for the two clusters in all sessions. The eyes-open condition always had a smaller exponent, indicating a higher dimensionality.

To further understand the functional role and implications of the neural manifold clusters, we studied their dimensionality. For this goal, we used the

participation ratio (PR):

$$\mathrm{PR} = \frac{\left(\sum_i \lambda_i\right)^2}{\sum_i \lambda_i^2} \tag{1}$$

where λ_i are the eigenvalues of the covariance matrix [18, 34]. See Dimensionality for detailed methods.

We computed the time-varying PR from the z-scored MUAe signals, by calculating the PR for windows of 30 s width, see Figure 4a. Stronger MUAe activity is typically associated with higher variance, which may bias the results toward higher dimensionality. We avoided bias due to the varying activity level by normalising the data (z-scored) within each window. We found that there is a strong correlation between the log odds and the time-varying PR (Figure 4b) and compared the PR values between the two neural manifold clusters using a Mann-Whitney U test (Figure 4c). The correlation and tests show that the dimensionality is higher during the eyes-open periods consistently across data sets.

To further support this finding we also show the distribution of the variance explained by each of the principal components (PC) of the MUAe data, depicted on a log-log scale in Figure 4d. We fitted a power law to the PC variances and report the exponent α (Figure 4e). A higher α indicates faster decay of the curve, i.e., lower dimensionality. The power law exponents are in agreement with our sliding window approach: We observed higher dimensionality (higher PR) during eyes-open than during eyes-closed periods (Figure 4e).

Thus, the dimensionality varies over time, with higher dimensionality during eyes-open periods. Given that we normalised the data for this analysis, the effects from any residual external stimuli were dampened, yet a significant difference in dimensionality was observed. While the high dimensionality during eyes-open periods could still be a by-product of visual stimuli, the dark room makes this unlikely. Thus, we have compelling evidence that an internal mechanism is responsible for the observed modulation, which leads to both a difference in the state space geometry and a change in dimensionality.

Feedback is present in the form of beta-band spectral Granger causality

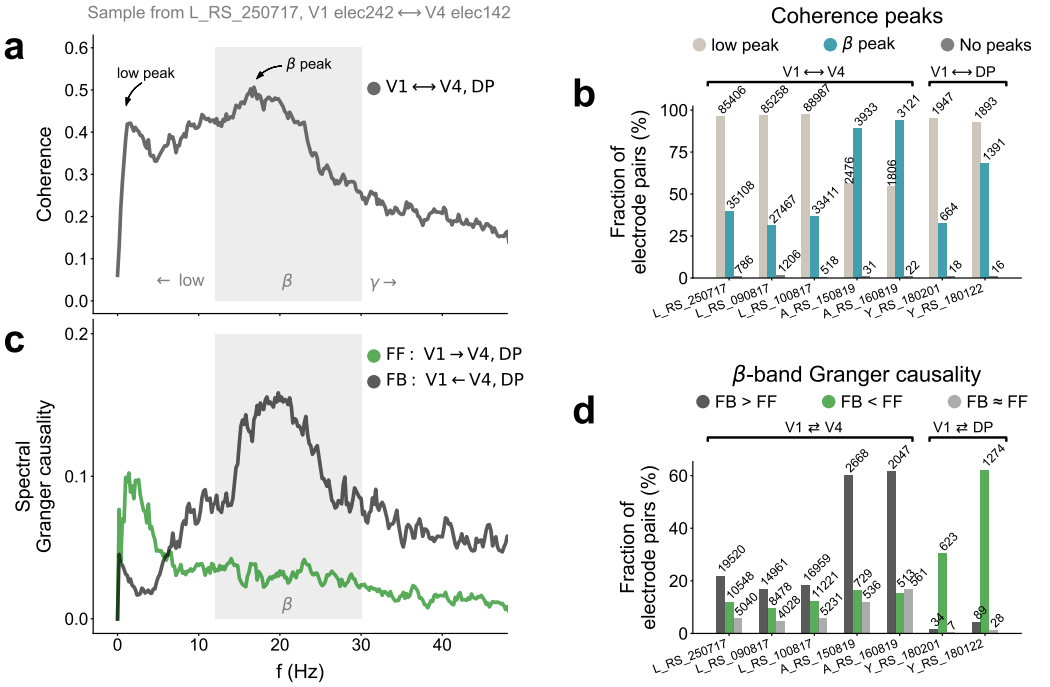


Fig. 5 Inter-area coherence and spectral Granger causality. **a** Representative sample of coherence between V1 and V4 (electrodes 242 and 142, respectively). Low-frequency and beta-band peaks indicated. **b** Quantification of coherence peaks across all sessions. A substantial portion of all electrode pairs displayed a beta peak. Note that the percentages for a given area can add up to more than 100% since the same electrode pair can have both a low-frequency and a beta peak. **c** Representative sample of spectral Granger causality; same electrodes as in panel **a**. **d** Quantification of beta-band spectral Granger causality for all sessions. Welch's t-test was used to determine whether feedback causality was greater than, less than, or roughly equal to feedforward causality. The test was only applied to those electrode pairs that showed a beta coherence peak. A large portion of V1 ⇄ V4 pairs show stronger causality in the feedback direction, while V1 ⇄ DP did not appear to have prominent feedback compared to feedforward causality.

In search of an internal mechanism that may modulate neural manifolds and dimensionality, we turned our attention to cortico-cortical interactions. Feedback communication has been shown to play an active role in contour integration in macaque V1 [29–32]. *In silico* experiments have shown that feedback can rotate neural manifolds to extract signals regardless of context [33] and that a learned feedback signal can modulate the synaptic weights in motor cortex during learning, essentially modulating the neural manifolds [38]. Feedback to V1 can be detected using spectral Granger causality, since it is known to occur within the beta frequency band (roughly 12-30 Hz) [27, 31].

To determine whether feedback is present in our data, we calculated the coherence and Granger causality between every pair of V1-V4 and V1-DP electrodes (see Coherence and Granger causality)—using the local field potential (LFP). We searched for peaks in the coherence, using an automatic method (see methods Peak detection), and found that beta peaks appear for a large fraction of electrode pairs (Figure 5a,b). Subsequently, we looked at the Granger causality of the electrode pairs that had beta peaks (sample shown in Figure 5c). We tested whether the beta-band feedback causality (from V4/DP to V1) was greater than the feedforward causality (from V1 to V4/DP) for each electrode using Welch's t-test, and conclude that feedback is stronger than feedforward for the majority of V1 $\rightleftarrows$ V4 electrode pairs (Figure 5d)—out of the ones with a coherence peak in the beta band. However, feedback causality exceeded feedforward causality for only very few electrode pairs for V1 $\rightleftarrows$ DP.

We further explored the spatial distribution of the overall causality strength $\mathcal{B}$, see Figure S11a–c. The electrodes with their receptive field (RF) around the fovea show stronger $\mathcal{B}$, in both the feedforward and the feedback direction, Figure S11d–f.

Thus, we found feedback from V4 to V1, in agreement with previous studies [27, 31]. Feedback signals from V4 to V1 are therefore possible candidates for the modulation of the neural manifolds and their dimensionality, which we explore further in the next section.

Stronger feedback during eyes-open periods

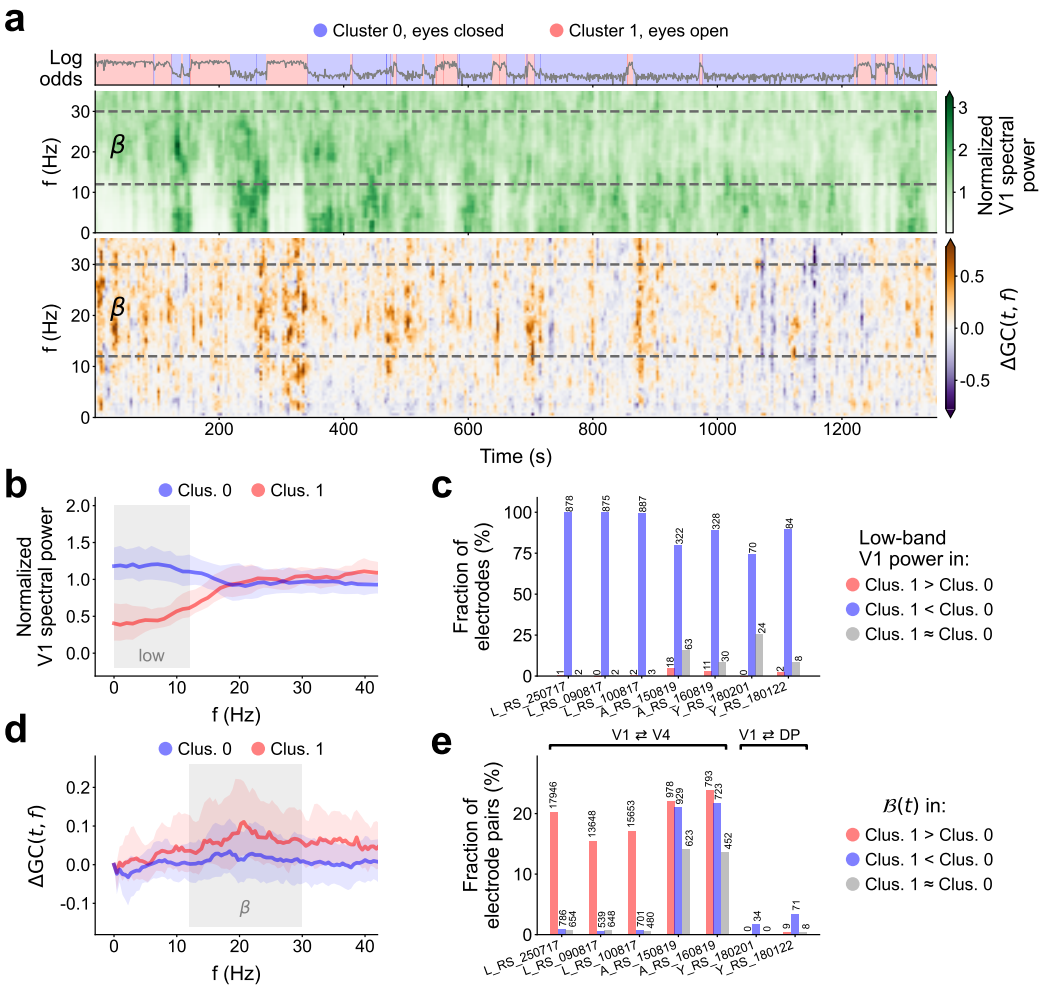


Fig. 6 Time-dependent power spectrum and spectral Granger causality reveal higher feedback in the eyes-open periods in spite of lower spectral power. **a** Time evolution of the log odds (top), the spectral power of a sample V1 electrode (middle), and the spectral Granger causality difference for a representative sample of V1-V4 electrodes (bottom). The sample electrodes were the same as in Figure 5. **b** Spectral power median (line) and 25th to 75th percentiles (shade) for each cluster for one sample V1 electrode. Low-frequency range highlighted. **c** Quantification of low-frequency spectral power in each cluster for all V1 electrodes in all sessions—using Welch's t-test. **d** Causality difference median (line) and 25th to 75th percentiles (shade) for each cluster for one sample V1-V4 electrode pair. Beta frequency range highlighted. **e** Quantification of beta-band causality difference $\mathcal{B}(t)$ over time (in each cluster) for all V1-V4 and V1-DP electrodes in all sessions—using Welch's t-test.

We have thus far shown that around 20–60% of V1-V4 electrode pairs displayed strong feedback from V4 to V1 (but not for V1-DP pairs) in the form of beta-band Granger causality. It has been suggested that V1–V4 interactions

are feedforward-dominated during visual stimulation and feedback-dominated with eyes open in the absence of stimuli [39]. This suggests that feedback signals should be present during eyes-open periods.

To elucidate the relation between feedback interactions and eye closure, we examined how the spectral power and beta-band Granger causality change over time, in relation to eye closure. To this end, we computed spectrograms of the LFPs and the beta-band causality difference $\mathcal{B}(t)$ at individual V1 electrodes using a 10-second sliding window (Figure 6a). Here, $\mathcal{B}(t)$ was calculated by subtracting the time-dependent feedforward spectral Granger causality from the feedback spectral Granger causality (Equation 8).

Visual inspection of the data from a representative electrode reveals that the spectral power in V1 decreases for low frequencies during eyes-open periods (Figure 6b), with the power modulations in the low-frequency band extending up to the lower beta band (roughly < 20 Hz). Thus, the V1 LFP power at frequencies up to about 20 Hz was generally lower during the eyes-open periods than during the eyes-closed periods. At the same time, the beta-band causality difference $\mathcal{B}(t)$ was higher during the eyes-open periods (shown in Figure 6d for a representative V1-V4 electrode pair), indicating stronger feedback from V4 to V1.

Next, we systematically quantified the difference in spectral power and causality across the two neural manifold clusters for all sessions and electrodes. To this end, we separated the spectra from each cluster using the log odds and tested for significant differences using Welch's t-test. First, we tested the difference in spectral power in the low-frequency band and found that the power is systematically higher during the eyes-closed periods (Figure 6c). Second, we tested the disparity in $\mathcal{B}(t)$ and found that the V1-V4 $\mathcal{B}(t)$ is systematically higher during the eyes-open periods. For completeness, we also tested the disparity in $\mathcal{B}(t)$ for V1-DP, revealing a slight tendency in the opposite direction.

The stronger feedback signals are not constant during the eyes-open periods, but rather appear to fluctuate (Figure 6a), which might be associated with visual stimulation or object recognition. Since any visual stimuli present despite the darkened room would be expected to be located in particular parts of the visual field, we further tested $\mathcal{B}(t)$ against the gaze direction for all electrode pairs (Figure S12). Figure S12a,b depict the signals from one sample electrode, whereas Figure S12c shows the distribution of correlation coefficients for all electrode pairs. Finally, we show the overall beta-band causality $\mathcal{B}$ against the $\mathcal{B}(t)$-gaze signal correlation (Figure S12d). This test checks whether there is a second-order effect, by assessing whether stronger feedback in a certain electrode pair was also associated with a strong gaze to feedback correlation. However, we could not observe any clear trend, thus suggesting no relation between gaze direction and feedback strength. All in all, the time-varying feedback signal does not depend on gaze direction, suggesting that the visual scene does not affect the observed feedback signals.

In conclusion, the time-dependent spectral analysis reveals large variations of power and causality. On the one hand, the spectral power at low frequencies increases during eyes-closed periods, a known phenomenon [40]. On the other hand, the V4-to-V1 feedback is strongest during the eyes-open periods. The time-varying feedback did not strongly correlate with gaze direction or change, suggesting no relation between feedback and the visual scene. These results suggest that feedback from V4 to V1 modulates V1 activity, activating a different neural manifold cluster of increased dimensionality. To further confirm this finding, we used numerical simulations, in which we can control the feedback signals and observe whether they have the hypothesised modulatory effect.

Simulation of a balanced spiking neural network with feedback reproduces manifold clustering

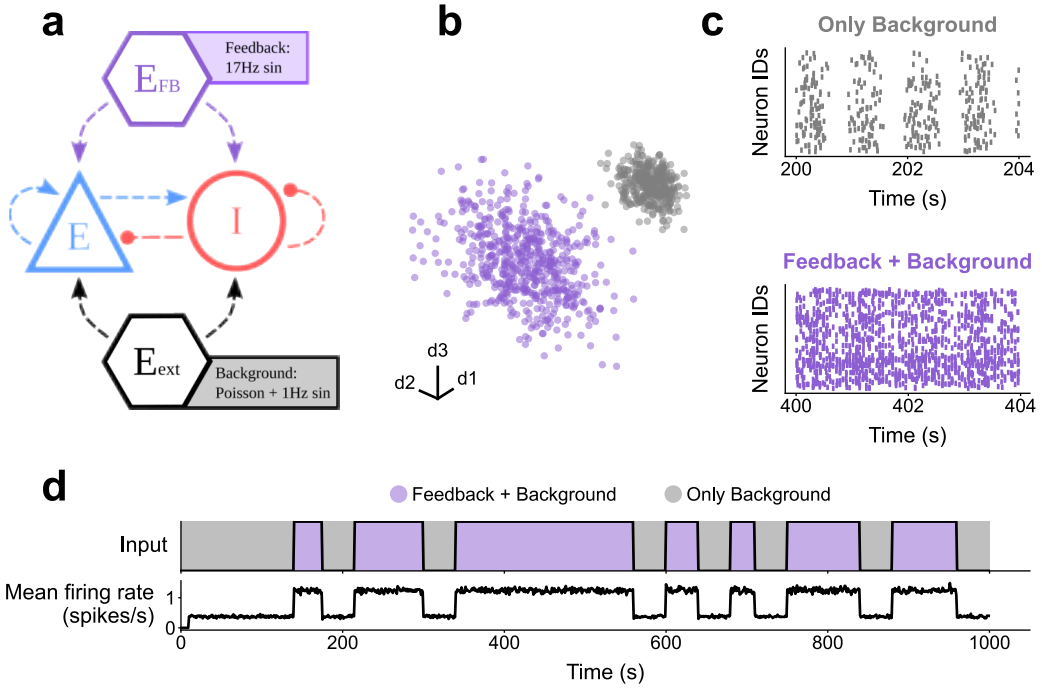


Fig. 7 Simulation of a balanced spiking neural network with feedback. **a** Diagram of balanced random spiking neural network. Background input is provided constantly, feedback signals were provided intermittently. **b** Three-dimensional PCA embedding of the model firing rate time histogram over all neurons (binsize = 1 s). Colours indicate whether feedback was provided. **c** Raster plot snippets of the spiking activity for the different epochs. **d** Time evolution of feedback signal and mean firing rate.

To answer whether feedback signals can modulate the activity of V1, we simulated a balanced spiking neural network akin to [41] with strong interactions

[42]. We provided background input and feedback at a similar frequency and a similar temporal profile as in the experimental data.

Our model consists of two populations of excitatory and inhibitory leaky integrate-and-fire (LIF) neurons, recurrently connected (Figure 7a). See Spiking neural network simulations for a full description of the model and its parameters. We provided background input at all times in the form of Poisson noise (800 spikes/s) and an oscillating inhomogeneous Poisson process (1 Hz sine wave with 500 spikes/s amplitude). Additionally, we provided a feedback signal on top of the background input at varying time intervals, see Figure 7d. The feedback was another inhomogeneous Poisson process oscillating in the beta range (17 Hz sinusoidal with 500 spikes/s amplitude).

We used the firing rate of the LIF neurons, downsampled to 1 Hz, to construct the neural manifold. This firing rate signal is comparable to the MUAe used in previous sections [35]. A 3D PCA embedding of the neural manifold is shown in Figure 7b. The manifold displays two distinct clusters, corresponding to the times with or without feedback. The simulated activity oscillated at a low frequency when no feedback was present and turned into an asynchronous irregular state when the feedback was provided, Figure 7c. Additionally, the mean firing rate increased as a consequence of the feedback, Figure 7d.

Thus, our simulation reproduced the changes in neural manifolds that we also observed in the experimental data. These findings support the role of feedback as a modulator of V1 activity. Taken together, the data analysis and simulations show that feedback signals from V4 to V1 in the beta frequency range can modulate the activity of V1, suggesting that the feedback plays a crucial role in the appearance of multiple clusters in the neural manifolds.

Discussion

The aim of this study was to determine whether cortico-cortical feedback modulates neural manifolds and dimensionality in visual cortex. In this study, we provide the first *in vivo* evidence for two distinct neural manifold clusters in the resting-state activity of macaque V1, where feedback from V4 to V1 may switch the activity to the higher-dimensional cluster. Evidence for the existence of clusters in the neural manifolds and feedback comes from electrophysiological recordings in macaque (N=3), from two different experimental labs. The effects of feedback on the neural manifolds were confirmed using a computational model of spiking neurons.

We show that distinct clusters are present in the neural manifolds via PCA (Figure 2, S1-S6) and confirmed their existence in high dimensions via persistent homology (Figure S7). We further show that the clusters are correlated with eye closure (Figure 3). In V4 and DP we find no neural manifold clusters (Figure S8-S10).

Neural dimensionality varies depending on the cluster identity (Figure 4). Our measured dimensionality is in agreement with previous reports in the visual cortex [10, 15]. Previous work has also shown higher dimensionality in the primary motor cortex during eyes-open than eyes-closed periods [43], analogous to our findings for the visual cortex. Thus, eye closure might lead to changes in dimensionality across the cortex, not just in the visual areas.

We hypothesised that feedback from other cortical areas could be the modulatory mechanism responsible for the changes observed in the neural manifolds and dimensionality of V1. Indeed, we found that feedback, in terms of Granger causality, is present in all of our recordings (Figure 5), particularly from V4 to V1. We also found the feedback to vary over time, with increased presence during the eyes-open periods, Figure 6. Previous studies found that cortico-cortical feedback signals are predominantly present in the beta (12–30 Hz) frequency band, while feedforward signals are present in the delta/theta (< 8 Hz) and gamma (> 30 Hz) bands [27, 44], in agreement with our findings. In our analysis we did not find any gamma band causality (Figure 5), since our recordings were from the deep cortical layers (mostly layer 5) and gamma oscillations are known to be absent from layer 5 of the visual cortex [26]. Further studies found that feedback from V2 and V3 to V1 also contributes to surround suppression [45]. Our findings are consistent with the work by Semedo et al. [39], who suggested that feedforward activity dominates during visual stimulation and feedback dominates in the absence of visual stimuli with the eyes open at all times. In the present study, it was not possible to test directly whether the V4-to-V1 feedback is responsible for the modulation. Future studies could inactivate the V4 $\rightarrow$ V1 pathway, e.g., via reducing the temperature of V4 with the 'cryoloop' technology [45, 46].

Numerical simulations confirm the ability of feedback signals to modulate neural manifolds (Figure 7). We provided feedback signals in the form of sinusoidal oscillating inhomogeneous Poisson processes, and found that these could induce a second manifold cluster.

Therefore, our simulations suggest that the feedback alone is sufficient to cause the observed qualitative changes in V1 activity, although some other mechanisms might be involved in regulating and controlling this process. Other possible mechanisms are neuromodulation or adaptation of recurrent connectivity via plasticity. Recurrent connectivity, in the form of cell-type-specific motifs, has been shown to constrain and control the dimensionality of brain networks [34]. However, this mechanism would still require some sort of behavioural information about the eye closure to be activated at the right times. Thus, besides modulating the activity of V1, feedback could also trigger other mechanisms for a more refined modulation.

Given the link between V1 activity and behaviour we explore the possible communication pathways. There are three main candidates: the visual stimulus or absence thereof from the retina to V1; the proprioception of eyelid muscles via the somatosensory cortex; and the voluntary motor commands for eye closure. On the one hand, visual stimuli are transmitted from the retina to V1 via the lateral geniculate nucleus (LGN). The absence of stimuli could be the reason for the observed changes in the V1 activity. However, the macaques in our experiments had very little to no visual inputs since they were sitting in a dark room. Additionally, we found no consistent difference in MUAe activity levels between the eyes-open and eyes-closed clusters (see results section 5). On the other hand, proprioception of the eyelid could inform the cortex when the eyes are closed and trigger the activity changes in V1. Proprioceptive neurons in the eyelid project via the oculomotor nerve to the midbrain (possibly to the superior colliculus) [47], eventually entering the cortex via the somatosensory area (S1) [48]. From S1 the signal could find its way to V1 via several cortico-cortical pathways, potentially including neurons in V4. Alternatively, voluntary eyelid closure initiated by the motor cortex and eye movements controlled by the frontal eye field may be communicated to the visual cortex via cortico-cortical connections or the superior colliculus. Given that V4 is part of the fronto-parietal network, the proprioceptive and eye movement signals could easily reach V4, which then finally modulates the V1 activity.

Recent work suggests that communication between V1 and V4—in the presence of visual stimuli—occurs via a dedicated subspace [49]. Our work did not focus on the communication of sensory stimuli between areas, but instead focused on the changes observed within V1. Additionally, our data were recorded in the resting state, where we expect communication between areas to be reduced compared to visual stimulation tasks. Nevertheless, our findings suggest that V4-to-V1 interactions shape the activity of V1, and future work could explore whether, and if so, how such interactions also alter the communication subspace.

Given the complex mechanisms that seem to be involved to ensure that V1 population activity depends on eye closure, there must be some purpose to or benefit from doing so. First of all, if the eyes are closed then no visual stimuli are expected, so V1 firing rates can drop to a minimum in order to save energy. When the eyes are open, a higher dimensionality might also be

advantageous since visual stimuli also have a high dimensionality [10], so that a high-dimensional state in V1 could facilitate processing visual stimuli.

In summary, we found two clusters in the state space of macaque V1 activity during the resting state. The clusters were correlated with eye closure, but not consistently related to activity levels. We also found higher dimensionality for the eyes-open than for the eyes-closed periods. We hypothesised that the clusters are modulated by feedback from V4 in a time-dependent fashion, and showed that such feedback is present in the form of beta-band spectral Granger causality. Finally, we reproduced our experimental observations in a spiking neuron network, which confirms the modulatory role of feedback. We have thus for the first time provided *in vivo* evidence for the modulation of neural manifolds and dimensionality by cortico-cortical feedback. This finding provides a new perspective on the interplay between cortico-cortical feedback and the structure of cortical activity; and underlines the importance of considering eye closure in the analysis of resting-state signals.

Methods

Electrophysiological data from macaque L & A

We used publicly available [36] neural activity recorded from the neocortex of rhesus macaques (N=2) during rest and a visual task. The macaques were implanted with 16 Utah arrays (Blackrock microsystems), two of them in visual area V4 and the rest in the primary visual cortex (V1), with a total of 1024 electrodes. The recording system recorded the electric potential at each electrode with a sampling rate of 30 kHz. A full description of the experimental setup and the data collection and preprocessing has already been published [36]; here we only provide the details relevant to this study.

Three resting-state (RS) sessions were recorded per macaque, during which the subjects did not have to perform any particular task and sat in a quiet dark room. Pupil position and diameter data were collected using an infrared camera in order to determine the direction of gaze and eye closure of the macaques. On the same days as the RS recordings a visual response task was also performed. The visual response data were used to calculate the signal-to-noise ratio (SNR) of each electrode and all electrodes with an SNR lower than 2 were excluded from further analysis. Additionally, we excluded up to 100 electrodes that contributed to high-frequency cross-talk in each session, as reported in the original data publication. The sessions, duration and number of electrodes per subject are listed in Table 1.

The raw neural data were processed into the multi-unit activity envelope (MUAe) signal and local field potential (LFP). To obtain MUAe data, the raw data were high pass filtered at 500 Hz, rectified, low-pass filtered at 200 Hz and downsampled to 1 kHz. Finally, the 50, 100 and 150 Hz components were removed with a band-stop filter in order to remove the European electric grid noise and its main harmonics. During the recordings, the macaque's head was held in position with a custom-made headstage. To obtain the LFP data, the raw data was low pass filtered at 250 Hz, downsampled to 500Hz and a band-stop filter was applied to remove the European electric grid noise (50, 100 and 150 Hz).

The MUAe and LFP data for each array were already provided by the original authors in the open-source .nix format, which uses python-neo data structures to hierarchically organize and annotate electrophysiological data and metadata. The metadata, such as the cross-talk removal or the positions of the arrays in the cortex, were provided in the .odml machine- and human-readable format, which could be easily incorporated into the python analysis scripts.

Electrophysiological data from macaque Y

In addition to the published data from macaques L & A, we also used an unpublished data set from one additional rhesus macaque (N=1). Neural activity was recorded during rest and during a visuomotor integration task. The

Table 1 Summary of subjects and recordings included in this study.

Subject	Session	Duration (s)	Areas	Contacts
L	L_RS_250717	1363	V1 V4	765 116
L	L_RS_090817	1321	V1 V4	761 116
L	L_RS_100817	1298	V1 V4	774 118
A	A_RS_150819	2278	V1 V4	402 11
A	A_RS_160819	2441	V1 V4	369 9
Y	Y_RS_180122	906	V1 DP	64 36
Y	Y_RS_180201	699	V1 DP	64 36

recording apparatus is described elsewhere [37]. The macaque was implanted with five Utah arrays (Blackrock microsystems), two of them in the primary visual cortex (V1), one in dorsal prelunate cortex (area DP), one in area 7A and one in the motor cortex (M1/PMd). In this study we only included the 6x6 electrode arrays from V1 (two arrays) and DP (one array), for a total of 108 electrodes. The recording system recorded the electric potential at each electrode with a sampling rate of 30 kHz.

Two resting-state (RS) sessions were recorded, during which the macaque did not have to perform any particular task and sat in a quiet dark room. Pupil position and diameter data were collected using an infrared camera in order to determine the direction of gaze and eye closure of the macaque. See Table 1 for an overview of the sessions used in this study.

MUAe and LFP signals were computed using the same procedure as for the other data sets.

Neural manifolds and clustering

The MUAe data were downsampled to 1 Hz and arranged into a single array, which constituted a very high-dimensional manifold, with between 50 and 900 recording locations per session.

In order to visualise the data, we used a standard dimensionality reduction technique (principal component analysis, PCA) to reduce the neural manifold

to 3D. The clusters observed in the RS sessions were labelled using a two-component Gaussian mixture model on the 3D embedding. The clustering method provides the log odds, i.e., the chance that any given point belongs to one cluster or the other. The log odds captures the multi-cluster structure of the manifold in a single time series and can be considered equivalent to the cluster identities.

Outlier removal

The neural manifolds in our analysis are a collection of time points scattered across the state space. In the data some time points appear very distant from all other points, which we associate with noise and we therefore seek to remove them. To identify the outliers we used a procedure similar to the one used by Chaudhuri et al. [5]. First, we calculated the distance matrix of all points to each other, and took the 1st percentile value from the distance distribution, D_1. We then estimated the number of neighbours that each point had within D_1 distance, and finally discarded the 20 percent of points with the fewest neighbours.

Topological data analysis

We used persistent homology to confirm that the lower-dimensional structures that we observed in the 3D embedding of the neural manifolds are in fact topological features of the data and not just an artefact of the dimensionality reduction. Before computing the persistence barcodes we embedded the data into a 10D subspace using the isomap technique [50]. The method aims at approximately preserving the geodesic distance between data points (that is the shortest path between two points on the neural manifold) and thus is suited for reducing the dimensionality of the data when applying a topological data analysis. The analysis on the 10D data showed qualitatively equivalent results to the full-dimensional data, while requiring a much shorter computation time.

To calculate the persistence barcodes of the Vietoris-Rips complex of the neural manifold we used an efficient open-source implementation (Ripser[1]). Briefly, the algorithm successively inflates balls with radius r around each point of the manifold. If k points have a pairwise distance smaller than r (that is, for all pairs of points both points are contained in the ball of the other point), they form a $(k-1)$D simplex. Thus, the neural manifold gives rise to a simplicial complex (a collection of simplices of potentially different dimension) the topological features of which represent the topology of the neural manifold and can be extracted computationally. As r is increased, many short-lived features appear by chance. If the manifold has complex topological structures, they should continuously appear as the radius of the balls grows for a large range of r. We computed the persistence barcode for the first three homology groups H_0, H_1 and H_2. Homology groups are topological invariants that capture topological features of a given dimension of the neural manifold. The

[1] https://pypi.org/project/ripser/

long-lasting bars in the n-th persistence barcode correspond to the number of independent generators β_n of the respective homology group H_n. For low dimensionalities, they can be interpreted intuitively: β_0 is the number of connected components, β_1 the number of 1D holes, β_2 the number of enclosed 2D voids. Throughout all plots of this paper we display the top 1% longest-lasting barcodes for each homology group.

Dimensionality

We used two different approaches to study the dimensionality of the neural data.

First, we compute the time-varying participation ratio (PR, Equation 1) from the covariance matrix. We take a 30 s sliding window with a 1 s offset over the MUAe data and compute the PR for each window separately. Higher activity leads to higher variance; thus, we normalised the data within each window via z-scoring to minimise this effect. The PR does not require setting an arbitrary threshold. From the time-varying PR we measured the correlation between the log odds and the PR, and the PR distribution in each cluster.

Second, we computed the eigenvalue distribution of the neural data for the entire session as well as within each cluster. Once again we normalised the data after sampling each cluster. The distribution appeared to follow a power law, in agreement with previous studies [10]. We used a linear regression in log-log space to fit a power law to our data, where the slope of the linear fit in the log-log plot corresponds to the exponent α of the power law.

Coherence and Granger causality

To study the communication between cortical areas we rely on the coherence and Granger causality.

Coherence is the quantification of linear correlations in the frequency domain. Such that

$$C_{xy}(f) = \frac{|S_{xy}(f)|^2}{S_{xx}(f)S_{yy}(f)} \tag{2}$$

where C_{xy} is the frequency (f) dependent coherence between two signals x and y, $S_{xy}(f)$ is the cross-spectral density, and $S_{xx}(f)$ and $S_{yy}(f)$ are the auto-spectral densities.

In order to assess the directionality of frequency dependent interactions between the areas we applied spectral Granger causality analysis to the LFP recordings [51]. We first computed the cross-spectral matrix $S(f)$ with the multitaper method. To this end, we subdivided the chosen signal pairs into 10 s long segments. These were processed individually with 3 Slepian tapers and averaged in the end. This yielded the cross-spectrum. The segments had an overlap of 50%. Next, we decomposed the cross-spectrum into the covariance matrix Σ and the transfer function $H(f)$ with the Wilson spectral matrix factorisation [52], obtaining the matrix equation

$$S = H(f)\Sigma H^{\dagger}(f). \tag{3}$$

With these factors, one is able to obtain a version of directional functional connectivity between the first and second signals via

$$\mathrm{GC}_{x\to y}(f) = \frac{S_{xx}(f)}{\tilde{H}_{xx}(f)\Sigma_{xx}\tilde{H}^{\dagger}_{xx}(f)} \tag{4}$$

where $\tilde{H}_{xx}(f) = H_{xx}(f) + \Sigma_{xy}/\Sigma_{xx}H_{xy}(f)$ and mutatis mutandis for the influence of the second onto the first signal. The analysis was performed for all pairs of channels between the areas that exhibited a peak in the coherence in the β band 12 Hz $< f <$ 30 Hz.

We quantify the beta-band Granger causality strength as

$$\mathcal{B} = \sum_{f=12\ \mathrm{Hz}}^{f=30\ \mathrm{Hz}} \mathrm{GC}_{x\leftarrow y}(f) - \mathrm{GC}_{x\to y}(f). \tag{5}$$

We also analysed the time-varying spectral Granger causality. For this aim we used 10 s windows and moved them across the data with 1 s steps, for a final time resolution of 1 Hz. We calculated the spectral Granger causality for each window separately. The initial and final 5 s were discarded to avoid disruptions at the boundaries. So the time-varying causality spectrogram is

$$\mathrm{GC}_{x\to y}(t,f) = \mathrm{GC}_{x\to y}(f)\Big|_{t_0}^{t_1}, \ \ldots\ , \mathrm{GC}_{x\to y}(f)\Big|_{t_{n-1}}^{t_n} \tag{6}$$

and mutatis mutandis for the $y \to x$ direction.

Finally, we also define the time-varying Granger causality difference

$$\Delta\mathrm{GC}(t,f) = \mathrm{GC}_{x\leftarrow y}(t,f) - \mathrm{GC}_{x\to y}(t,f), \tag{7}$$

which summed over the beta band we call

$$\mathcal{B}(t) = \sum_{f=12\ \mathrm{Hz}}^{f=30\ \mathrm{Hz}} \Delta\mathrm{GC}(t,f). \tag{8}$$

Note that

$$\mathcal{B} \neq \frac{\sum^{N}\mathcal{B}(t)}{N} \tag{9}$$

due to the nonlinearities in the Granger causality calculation.

Both the coherence and spectral Granger causality were implemented in the Electrophysiology Analysis Toolkit (Elephant) [53].

Peak detection

To detect coherence peaks we used a standard peak detection algorithm for 1D arrays using wavelet transforms. We computed the continuous wavelet transform (cwt) for wavelets with widths from 10 Hz to 100 Hz (at 0.1 Hz steps), using a Ricker wavelet—i.e., a Mexican hat. Next, we searched for ridge lines in the cwt—peaks across different wavelet lengths—following standard criteria [54]. Finally, the ridge lines were filtered based on their total length, gaps and signal-to-noise ratio. The resulting ridge lines (if any) were considered as peaks in the coherence.

The detected peaks tended to be broad, since our parameter choice intentionally rejected narrow peaks. We chose this configuration in favour of robustness and to minimise false positives. Nevertheless, peaks were detected for a majority of electrode pairs.

Table 2 Peak detection algorithm parameters.

CWT peak detection parameters		
widths	$10 - 100$ Hz	Width range for CWT matrix.
width_step	0.1 Hz	Step between widths.
wavelet	Ricker	Wavelet used for convolution.
max_distances	widths / 4	Criterion to consider ridge lines connected.
gap_thresh	10 Hz	Ridge lines farther apart will not be connected.
min_length	225	Minimum length of ridge line.
min_snr	1	Percentile within the same ridge line.
noise_perc	10	Percentile below which to consider noise.

Spiking neural network simulations

To investigate the hypothesis that feedback in the β-band induces a change in dimensionality, we conducted a spiking neural network simulation. The network consists of $10,000$ excitatory and $2,500$ inhibitory leaky integrate-and-fire neurons with exponential post-synaptic currents. Two neurons are randomly connected with a connection probability of $p = 0.1$. The spike transmission delay is randomly sampled following a log-normal distribution. During a simulation, we distinguish two alternating states corresponding to the eyes-open and eyes-closed periods during the recordings. In the first state (background state), the input consists of spike trains sampled from an inhomogeneous Poisson process with a baseline rate of ν_{bg} Hz that is modulated with a 1 Hz sinusoidal oscillation. In the second state (feedback state), the network additionally receives input spike trains from inhomogeneous Poisson processes whose rates are oscillating at 17 Hz. The first state represents the eyes-closed, the second the eyes-open. Both input regimes provide independent input to each neuron, based on the same rate profiles. During the simulation, we recorded the spiking activity of 200 excitatory and 75 inhibitory neurons and determined the dimensionality as well as the power-law exponent of the spike trains as described above with a bin size of 1 s. See Table 3, Table 4 for

Table 3 General model description

Model Summary	
Populations	two populations, one excitatory, one inhibitory
Connectivity	random connectivity
Neuron model	leaky integrate-and-fire model
Synapse model	exponential postsynaptic current
Input	independent spike trains from inhomogeneous Poisson processes with given rate $r(t)$
Neuron and synapse model	
Subthreshold dynamics	$\frac{dV}{dt} = -\frac{V}{\tau_m} + \frac{I_{syn}(t)}{C_m}$, $I_{syn}(t) = Je^{-(t-t^*-d)/\tau_{syn}} H(t-t^*-d)$
Spiking	If $V(t-) < \theta$ and $V(t+) \geq \theta$, 1. Set $t^* = t$ and $V(t) = V_0$, and 2. Emit spike with time stamp t^*.
Connectivity	
Type	pairwise Bernoulli, i.e., for each pair of neurons generate a synapse with probability p
Weights	fixed source and target population specific weights
Delays	log-normally distributed delays for excitatory and inhibitory neurons
Input	
Rate	$r(t) = \max(0, \nu_{\text{base}} + \nu_{\text{amp}} \cdot \sin(2\pi f \cdot t))$

a full description of the network. For the simulations we used NEST (version 3.3) [55].

Table 4 Simulation parameters.

Population Parameters		
N_{ex}	10,000	number of excitatory neurons
N_{in}	2,500	number of inhibitory neurons
Connectivity Parameters		
p	0.1	connection probability
Neuron parameters		
τ_{m}	20ms	membrane time constant
τ_{r}	2ms	absolute refractory period
τ_{syn}	2ms	postsynaptic current time constant
C_{m}	1pF	membrane capacity
V_{m}	0mV	resting potential
E_{L}	0mV	membrane capacity
V_{reset}	0mV	reset membrane potential
V_{th}	20mV	threshold
Stimulus parameters: Background		
ν_{base}	800 spikes/s	base line rate
ν_{amp}	500 spikes/s	amplitude
f	1 Hz	sinusoidal oscillation frequency
Stimulus parameters: Feedback		
ν_{base}	0 spikes/s	base line rate
ν_{amp}	500 spikes/s	amplitude
f	17 Hz	sinusoidal oscillation frequency
Synapse parameters		
J_{EE}	0.5 mV	synaptic efficacy excitatory to excitatory
J_{IE}	0.75 mV	synaptic efficacy excitatory to inhibitory
g	4	relative inhibitory synaptic efficacy
Delay parameters		
μ_{ex}	1.5 ms	mean of lognormal distribution for excitatory connections
μ_{in}	0.75 ms	mean of lognormal distribution for inhibitory connections
$\sigma_{\text{ex,in}}$	0.5 ms	standard deviation of lognormal distribution for all connections

Author contributions

AMG, JI and AK conceptualised the study. AMG curated and processed the data from macaques L and A. AKJ curated and processed the data from macaque Y, with minor inputs from AMG. AMG and AK performed the dimensionality analysis, statistical testing and Granger causality analysis. AK performed the spiking neuron network simulations and AMG analyzed the results. AMG created all figures, with feedback from all other authors. AMG wrote the initial manuscript, all other authors provided feedback. JI, SG, SvA supervised the project. SG and SvA procured the funding.

Acknowledgements

This project received funding from the DFG Priority Program (SPP 2041 "Computational Connectomics") [S.J.van Albada: AL 2041/1-1]; the EU's Horizon 2020 Framework Grant Agreement No. 945539 (Human Brain Project

SGA3); the DFG (RTG 2416 "MultiSenses-MultiScales"); and the CNRS Multidisciplinary Exploratory Projects initiative.

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

[1] Gao, P., Trautmann, E., Yu, B., Santhanam, G., Ryu, S., Shenoy, K., Ganguli, S.: A theory of multineuronal dimensionality, dynamics and measurement. bioRxiv (2017). https://doi.org/10.1101/214262

[2] Gallego, J.A., Perich, M.G., Miller, L.E., Solla, S.A.: Neural Manifolds for the Control of Movement. Neuron **94**(5), 978–984 (2017). https://doi.org/10.1016/j.neuron.2017.05.025

[3] Gallego, J.A., Perich, M.G., Naufel, S.N., Ethier, C., Solla, S.A., Miller, L.E.: Cortical population activity within a preserved neural manifold underlies multiple motor behaviors. Nature Communications **9**(1), 1–13 (2018). https://doi.org/10.1038/s41467-018-06560-z

[4] Feulner, B., Clopath, C.: Neural manifold under plasticity in a goal driven learning behaviour. PLOS Computational Biology **17**(2), 1008621 (2021). https://doi.org/10.1371/journal.pcbi.1008621

[5] Chaudhuri, R., Gerçek, B., Pandey, B., Peyrache, A., Fiete, I.: The intrinsic attractor manifold and population dynamics of a canonical cognitive circuit across waking and sleep. Nature Neuroscience **22**(9), 1512–1520 (2019). https://doi.org/10.1038/s41593-019-0460-x

[6] Mante, V., Sussillo, D., Shenoy, K.V., Newsome, W.T.: Context-dependent computation by recurrent dynamics in prefrontal cortex. Nature **503**(7474), 78–84 (2013). https://doi.org/10.1038/nature12742

[7] Churchland, M.M., Cunningham, J.P., Kaufman, M.T., Foster, J.D., Nuyujukian, P., Ryu, S.I., Shenoy, K.V.: Neural population dynamics during reaching. Nature **487**(7405), 51–56 (2012). https://doi.org/10.1038/nature11129

[8] Wu, A., Pashkovski, S., Datta, S.R., Pillow, J.W.: Learning a latent manifold of odor representations from neural responses in piriform cortex. In: Bengio, S., Wallach, H., Larochelle, H., Grauman, K., Cesa-Bianchi, N., Garnett, R. (eds.) Advances in Neural Information Processing Systems, vol. 31 (2018)

[9] Gardner, R.J., Hermansen, E., Pachitariu, M., Burak, Y., Baas, N.A., Dunn, B.A., Moser, M.-B., Moser, E.I.: Toroidal topology of population activity in grid cells. Nature **602**(7895), 123–128 (2022). https://doi.org/10.1038/s41586-021-04268-7

[10] Stringer, C., Pachitariu, M., Steinmetz, N., Carandini, M., Harris, K.D.: High-dimensional geometry of population responses in visual cortex. Nature **571**(7765), 361–365 (2019). https://doi.org/10.1038/s41586-019-1346-5

[11] Froudarakis, E., Cohen, U., Diamantaki, M., Walker, E.Y., Reimer, J., Berens, P., Sompolinsky, H., Tolias, A.S.: Object manifold geometry across the mouse cortical visual hierarchy. bioRxiv (2020). https://doi.org/10.1101/2020.08.20.258798

[12] Singh, G., Memoli, F., Ishkhanov, T., Sapiro, G., Carlsson, G., Ringach, D.L.: Topological analysis of population activity in visual cortex. Journal of Vision **8**(8), 11–11 (2008). https://doi.org/10.1167/8.8.11

[13] Ghrist, R.: Barcodes: The persistent topology of data. Bulletin of the American Mathematical Society **45**(1), 61–75 (2008). https://doi.org/10.1090/S0273-0979-07-01191-3

[14] Cunningham, J.P., Yu, B.M.: Dimensionality reduction for large-scale neural recordings. Nature Neuroscience **17**(11), 1500–1509 (2014). https://doi.org/10.1038/nn.3776

[15] Gao, P., Ganguli, S.: On simplicity and complexity in the brave new world of large-scale neuroscience. Current Opinion in Neurobiology **32**, 148–155 (2015). https://doi.org/10.1016/j.conb.2015.04.003

[16] Pang, R., Lansdell, B.J., Fairhall, A.L.: Dimensionality reduction in neuroscience. Current Biology **26**(14), 656–660 (2016). https://doi.org/10.1016/j.cub.2016.05.029

[17] Williamson, R.C., Cowley, B.R., Litwin-Kumar, A., Doiron, B., Kohn, A., Smith, M.A., Yu, B.M.: Scaling Properties of Dimensionality Reduction for Neural Populations and Network Models. PLOS Computational Biology **12**(12), 1005141 (2016). https://doi.org/10.1371/journal.pcbi.1005141

[18] Mazzucato, L., Fontanini, A., La Camera, G.: Stimuli Reduce the Dimensionality of Cortical Activity. Frontiers in Systems Neuroscience **10** (2016). https://doi.org/10.3389/fnsys.2016.00011

[19] Avitan, L., Stringer, C.: Not so spontaneous: Multi-dimensional representations of behaviors and context in sensory areas. Neuron, 0896627322005888 (2022). https://doi.org/10.1016/j.neuron.2022.06.019

[20] Marx, E., Stephan, T., Nolte, A., Deutschländer, A., Seelos, K.C., Dieterich, M., Brandt, T.: Eye closure in darkness animates sensory systems. NeuroImage **19**(3), 924–934 (2003). https://doi.org/10.1016/S1053-8119(03)00150-2

[21] Marx, E., Deutschländer, A., Stephan, T., Dieterich, M., Wiesmann, M., Brandt, T.: Eyes open and eyes closed as rest conditions: impact on brain activation patterns. NeuroImage **21**(4), 1818–1824 (2004). https://doi.org/10.1016/j.neuroimage.2003.12.026

[22] Boytsova, Y.A., Danko, S.G.: EEG differences between resting states with eyes open and closed in darkness. Human Physiology **36**(3), 367–369 (2010). https://doi.org/10.1134/S0362119710030199

[23] McAvoy, M., Larson-Prior, L., Ludwikow, M., Zhang, D., Snyder, A.Z., Gusnard, D.L., Raichle, M.E., d'Avossa, G.: Dissociated mean and functional connectivity BOLD signals in visual cortex during eyes closed and fixation. Journal of Neurophysiology **108**(9), 2363–2372 (2012). https://doi.org/10.1152/jn.00900.2011

[24] Brodoehl, S., Klingner, C.M., Witte, O.W.: Eye closure enhances dark night perceptions. Scientific Reports **5**(1), 10515 (2015). https://doi.org/10.1038/srep10515

[25] Chen, X., Wang, F., Fernandez, E., Roelfsema, P.R.: Shape perception via a high-channel-count neuroprosthesis in monkey visual cortex. Science **370**(6521), 1191–1196 (2020). https://doi.org/10.1126/science.abd7435

[26] van Kerkoerle, T., Self, M.W., Dagnino, B., Gariel-Mathis, M.-A., Poort, J., van der Togt, C., Roelfsema, P.R.: Alpha and gamma oscillations characterize feedback and feedforward processing in monkey visual cortex. Proceedings of the National Academy of Sciences **111**(40), 14332–14341 (2014). https://doi.org/10.1073/pnas.1402773111

[27] Bastos, A., Vezoli, J., Bosman, C., Schoffelen, J.-M., Oostenveld, R., Dowdall, J., De Weerd, P., Kennedy, H., Fries, P.: Visual Areas Exert Feedforward and Feedback Influences through Distinct Frequency Channels. Neuron **85**(2), 390–401 (2015). https://doi.org/10.1016/j.neuron.2014.12.018

[28] Klink, P.C., Dagnino, B., Gariel-Mathis, M.-A., Roelfsema, P.R.: Distinct Feedforward and Feedback Effects of Microstimulation in Visual Cortex Reveal Neural Mechanisms of Texture Segregation. Neuron **95**(1), 209–2203 (2017). https://doi.org/10.1016/j.neuron.2017.05.033

[29] Poort, J., Raudies, F., Wannig, A., Lamme, V.F., Neumann, H., Roelfsema, P.: The Role of Attention in Figure-Ground Segregation in Areas V1 and V4 of the Visual Cortex. Neuron **75**(1), 143–156 (2012). https://doi.org/10.1016/j.neuron.2012.04.032

[30] Self, M., van Kerkoerle, T., Supèr, H., Roelfsema, P.: Distinct Roles of the Cortical Layers of Area V1 in Figure-Ground Segregation. Current Biology **23**(21), 2121–2129 (2013). https://doi.org/10.1016/j.cub.2013.09.013

[31] van Kerkoerle, T., Self, M.W., Roelfsema, P.R.: Layer-specificity in the effects of attention and working memory on activity in primary visual cortex. Nature Communications **8**(1), 13804 (2017). https://doi.org/10.1038/ncomms13804

[32] Liang, H., Gong, X., Chen, M., Yan, Y., Li, W., Gilbert, C.D.: Interactions between feedback and lateral connections in the primary visual cortex. Proceedings of the National Academy of Sciences **114**(32), 8637–8642 (2017). https://doi.org/10.1073/pnas.1706183114

[33] Naumann, L.B., Keijser, J., Sprekeler, H.: Invariant neural subspaces maintained by feedback modulation. eLife **11**, 76096 (2022). https://doi.org/10.7554/eLife.76096

[34] Dahmen, D., Recanatesi, S., Jia, X., Ocker, G.K., Campagnola, L., Jarsky, T., Seeman, S., Helias, M., Shea-Brown, E.: Strong and localized recurrence controls dimensionality of neural activity across brain areas. bioRxiv (2020). https://doi.org/10.1101/2020.11.02.365072

[35] Supèr, H., Roelfsema, P.R.: Chronic multiunit recordings in behaving animals: advantages and limitations. Progress in Brain Research **147**, 263–282 (2005). https://doi.org/10.1016/S0079-6123(04)47020-4

[36] Chen, X., Morales-Gregorio, A., Sprenger, J., Kleinjohann, A., Sridhar, S., van Albada, S.J., Grün, S., Roelfsema, P.R.: 1024-channel electrophysiological recordings in macaque V1 and V4 during resting state. Scientific Data **9**(1), 77 (2022). https://doi.org/10.1038/s41597-022-01180-1

[37] de Haan, M.J., Brochier, T., Grün, S., Riehle, A., Barthélemy, F.V.: Real-time visuomotor behavior and electrophysiology recording setup for use with humans and monkeys. Journal of Neurophysiology **120**(2), 539–552 (2018). https://doi.org/10.1152/jn.00262.2017

[38] Feulner, B., Clopath, C.: Neural manifold under plasticity in a goal driven learning behaviour. PLOS Computational Biology **17**(2), 1008621 (2021). https://doi.org/10.1371/journal.pcbi.1008621

[39] Semedo, J.D., Jasper, A.I., Zandvakili, A., Krishna, A., Aschner, A., Machens, C.K., Kohn, A., Yu, B.M.: Feedforward and feedback interactions between visual cortical areas use different population activity patterns. Nature Communications **13**(1), 1099 (2022). https://doi.org/10.1038/s41467-022-28552-w

[40] Lewine, J.D., Orrison, W.W.: Clinical Electroencephalography and Event-Related Potentials. In: Functional Brain Imaging, pp. 327–368. Elsevier, North Holland (1995). https://doi.org/10.1016/B978-0-8151-6509-5.50012-6

[41] Brunel, N.: Dynamics of Sparsely Connected Networks of Excitatory and Inhibitory Spiking Neurons. Journal of Computational Neuroscience **8**(3), 183–208 (2000). https://doi.org/10.1023/A:1008925309027

[42] Ostojic, S.: Two types of asynchronous activity in networks of excitatory and inhibitory spiking neurons. Nature Neuroscience **17**(4), 594–600 (2014). https://doi.org/10.1038/nn.3658

[43] Dabrowska, P.A., Voges, N., von Papen, M., Ito, J., Dahmen, D., Riehle, A., Brochier, T., Grün, S.: On the Complexity of Resting State Spiking Activity in Monkey Motor Cortex. Cerebral Cortex Communications **2**(3), 033 (2021). https://doi.org/10.1093/texcom/tgab033

[44] Vezoli, J., Vinck, M., Bosman, C.A., Bastos, A.M., Lewis, C.M., Kennedy, H., Fries, P.: Brain rhythms define distinct interaction networks with differential dependence on anatomy. Neuron **109**(23), 3862–38785 (2021). https://doi.org/10.1016/j.neuron.2021.09.052

[45] Nassi, J.J., Lomber, S.G., Born, R.T.: Corticocortical Feedback Contributes to Surround Suppression in V1 of the Alert Primate. Journal of Neuroscience **33**(19), 8504–8517 (2013). https://doi.org/10.1523/JNEUROSCI.5124-12.2013

[46] Lomber, S.G., Payne, B.R., Horel, J.A.: The cryoloop: an adaptable reversible cooling deactivation method for behavioral or electrophysiological assessment of neural function. Journal of Neuroscience Methods **86**(2), 179–194 (1999). https://doi.org/10.1016/S0165-0270(98)00165-4

[47] Porter, J.D., Guthrie, B.L., Sparks, D.L.: Innervation of monkey extraocular muscles: Localization of sensory and motor neurons by retrograde transport of horseradish peroxidase. The Journal of Comparative Neurology **218**(2), 208–219 (1983). https://doi.org/10.1002/cne.902180208

[48] Balslev, D., Albert, N.B., Miall, C.: Eye muscle proprioception is represented bilaterally in the sensorimotor cortex. Human Brain Mapping **32**(4), 624–631 (2011). https://doi.org/10.1002/hbm.21050

[49] Semedo, J.D., Zandvakili, A., Machens, C.K., Yu, B.M., Kohn, A.: Cortical Areas Interact through a Communication Subspace. Neuron **102**(1), 249–2594 (2019). https://doi.org/10.1016/j.neuron.2019.01.026

[50] Tenenbaum, J.B., Silva, V.d., Langford, J.C.: A Global Geometric Framework for Nonlinear Dimensionality Reduction. Science **290**(5500), 2319–2323 (2000). https://doi.org/10.1126/science.290.5500.2319

[51] Dhamala, M., Rangarajan, G., Ding, M.: Estimating Granger Causality from Fourier and Wavelet Transforms of Time Series Data. Physical Review Letters **100**(1), 018701 (2008). https://doi.org/10.1103/PhysRevLett.100.018701

[52] Wilson, G.T.: The Factorization of Matricial Spectral Densities. SIAM Journal on Applied Mathematics **23**(4), 420–426 (1972). https://doi.org/10.1137/0123044

[53] Denker, M., Köhler, C., Morales-Gregorio, A., Kleinjohann, A., Bouss, P., Stella, A., Jurkus, R., Kramer, M., Dabrowska, P., Kern, M., Kurth, A.C., Gutzen, R., Porrmann, F., Pilz, S.: Elephant 0.11.1. Zenodo (2022). https://doi.org/10.5281/ZENODO.6470226

[54] Du, P., Kibbe, W.A., Lin, S.M.: Improved peak detection in mass spectrum by incorporating continuous wavelet transform-based pattern matching. Bioinformatics **22**(17), 2059–2065 (2006). https://doi.org/10.1093/bioinformatics/btl355

[55] Spreizer, S., Mitchell, J., Jordan, J., Wybo, W., Kurth, A., Vennemo, S.B., Pronold, J., Trensch, G., Benelhedi, M.A., Terhorst, D., Eppler, J.M., Mørk, H., Linssen, C., Senk, J., Lober, M., Morrison, A., Graber, S., Kunkel, S., Gutzen, R., Plesser, H.E.: NEST 3.3. Zenodo (2022). https://doi.org/10.5281/ZENODO.6368023

Extended data figures

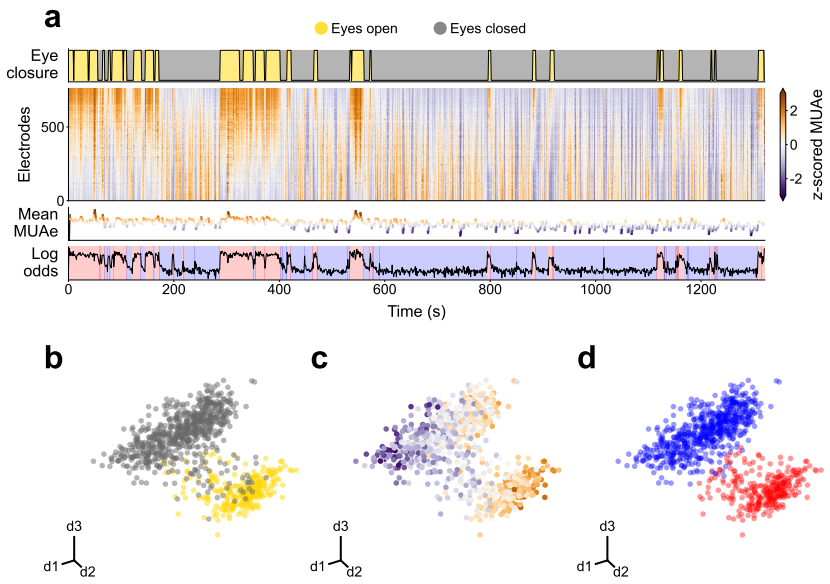


Fig. S1 Overview of the experimental data from session L_RS_090817. **a** Time evolution of signals. **b, c, d** Three dimensional PCA embedding of the MUAe neural manifold..

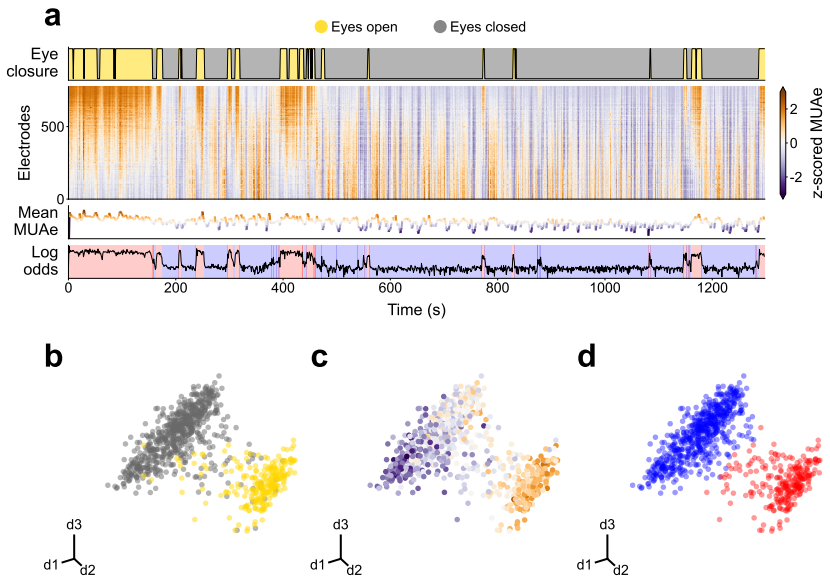


Fig. S2 Overview of the experimental data from session L_RS_100817. **a** Time evolution of signals. **b, c, d** Three dimensional PCA embedding of the MUAe neural manifold.

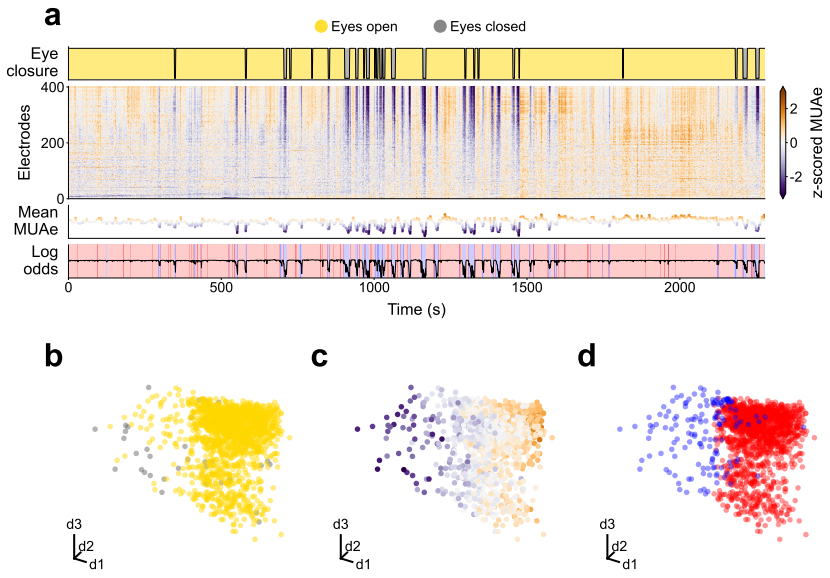


Fig. S3 Overview of the experimental data from session A_RS_150819. **a** Time evolution of signals. **b, c, d** Three dimensional PCA embedding of the MUAe neural manifold.

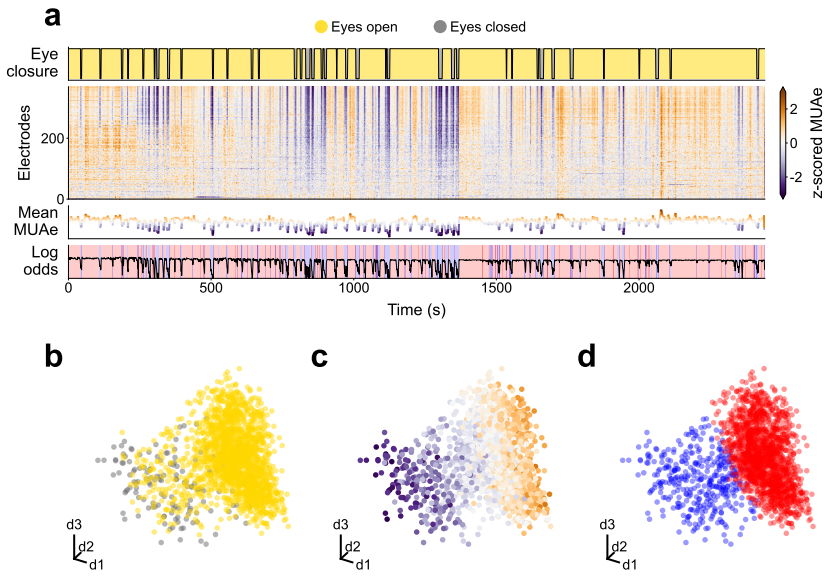


Fig. S4 Overview of the experimental data from session A_RS_160819. **a** Time evolution of signals. **b, c, d** Three dimensional PCA embedding of the MUAe neural manifold.

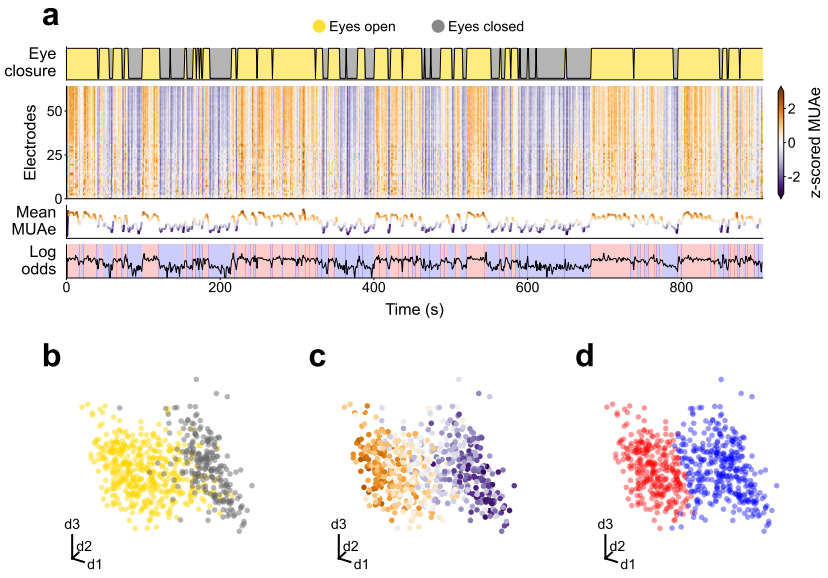


Fig. S5 Overview of the experimental data from session Y_RS_180122. **a** Time evolution of signals. **b, c, d** Three dimensional PCA embedding of the MUAe neural manifold.

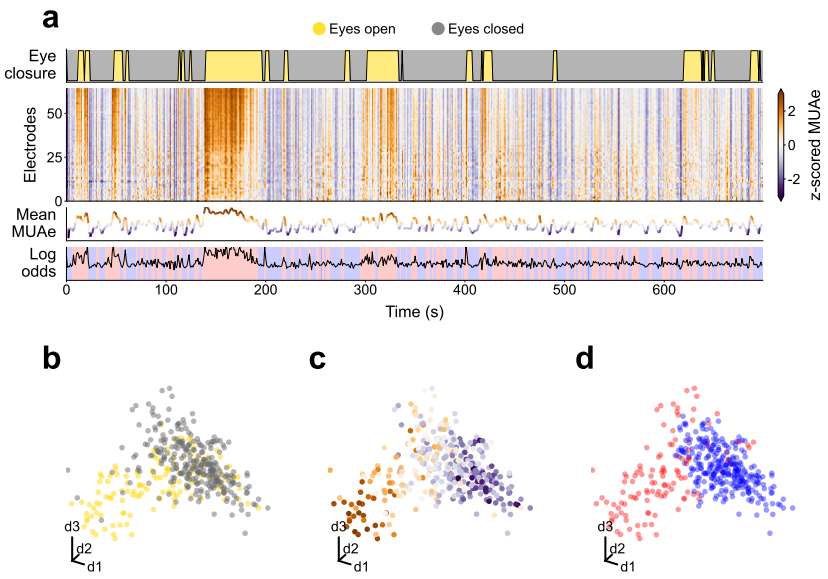


Fig. S6 Overview of the experimental data from session Y_RS_180122. **a** Time evolution of signals. **b, c, d** Three dimensional PCA embedding of the MUAe neural manifold.

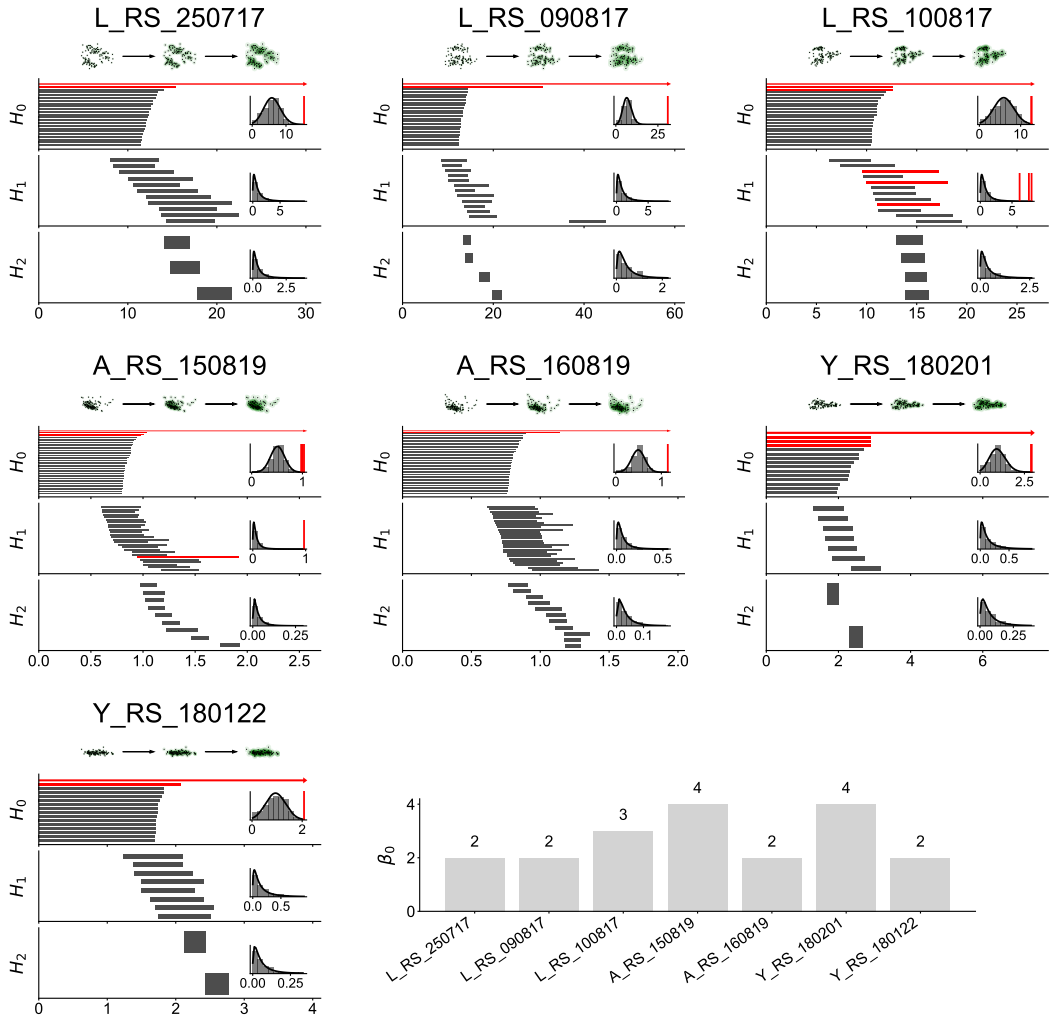


Fig. S7 Persistence homology of the high-dimensional manifolds show the presence of at least two clusters. Each panel shows data for one session. For each panel, (Top) Sample clouds with a green radius around them. These correspond to the radius used in the persistent homology process. (Main plots) Persistence barcodes of the Vietoris-Rips complex of the 10D neural manifolds, for all sessions. (Inset plots) Distribution of barcode length with a fitted lognormal distribution, long barcodes coloured red (determined ad-hoc). Bottom right bar plot shows number of clusters found in each session.

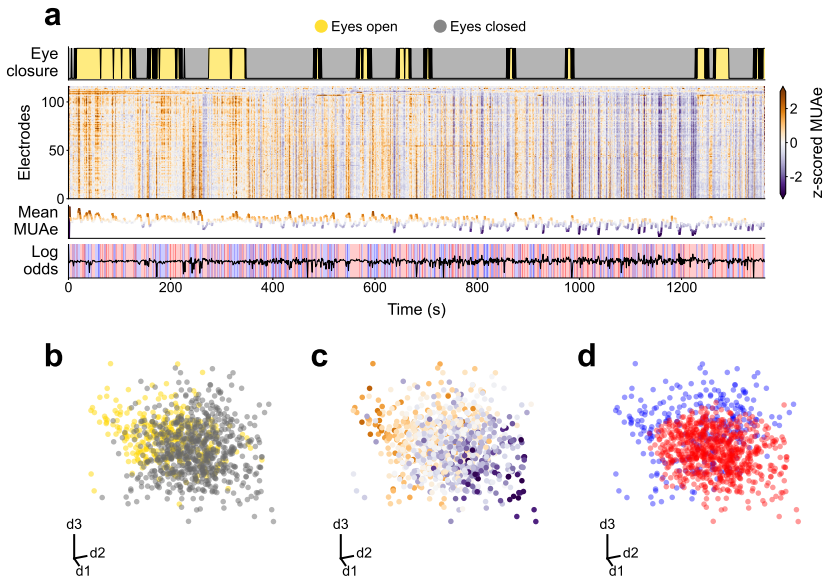


Fig. S8 V4 activity from session L_RS_250717 does not show distinct clusters in its neural manifold. **a** Time evolution of signals. **b, c, d** Three dimensional PCA embedding of the MUAe neural manifold.

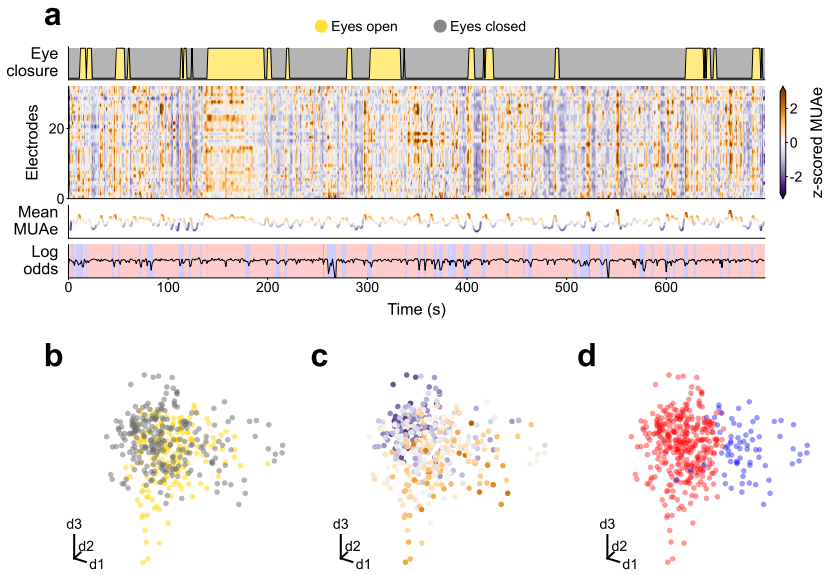


Fig. S9 DP activity from session Y_RS_180201 does not show distinct clusters in its neural manifold. **a** Time evolution of signals. **b, c, d** Three dimensional PCA embedding of the MUAe neural manifold.

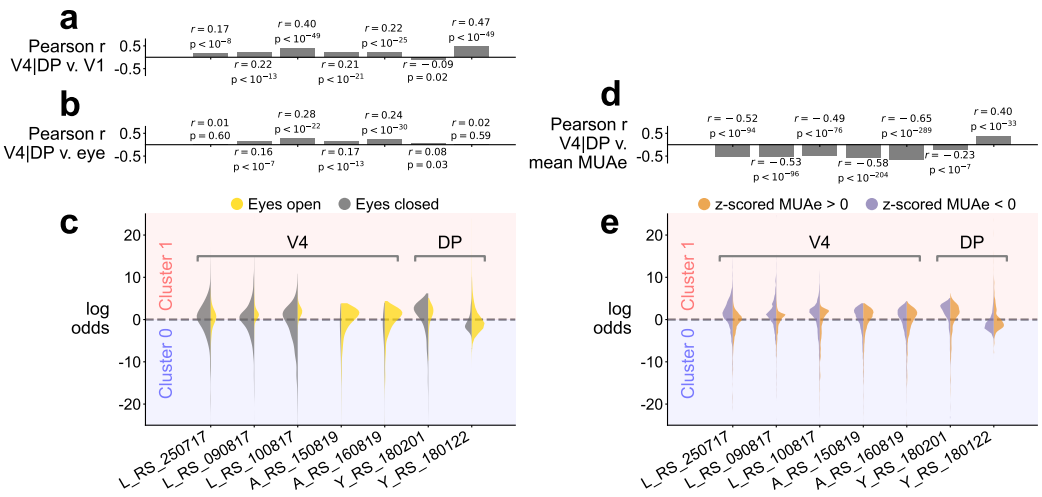


Fig. S10 V4 and DP manifold log odds are not strongly correlated with V1, eye closure nor MUAe. **a** Log odds in V4 and DP do not strongly correlate with log odds in V1. **b** Log odds in V4 and DP do not strongly correlate with eye closure. **c** Violin plots of V4 and DP eye closure show no separation along different clusters. **d** Log odds in V4 and DP do not strongly correlate with mean MUAe. **e** Violin plots of V4 and DP MUAe activity show no separation along different clusters.

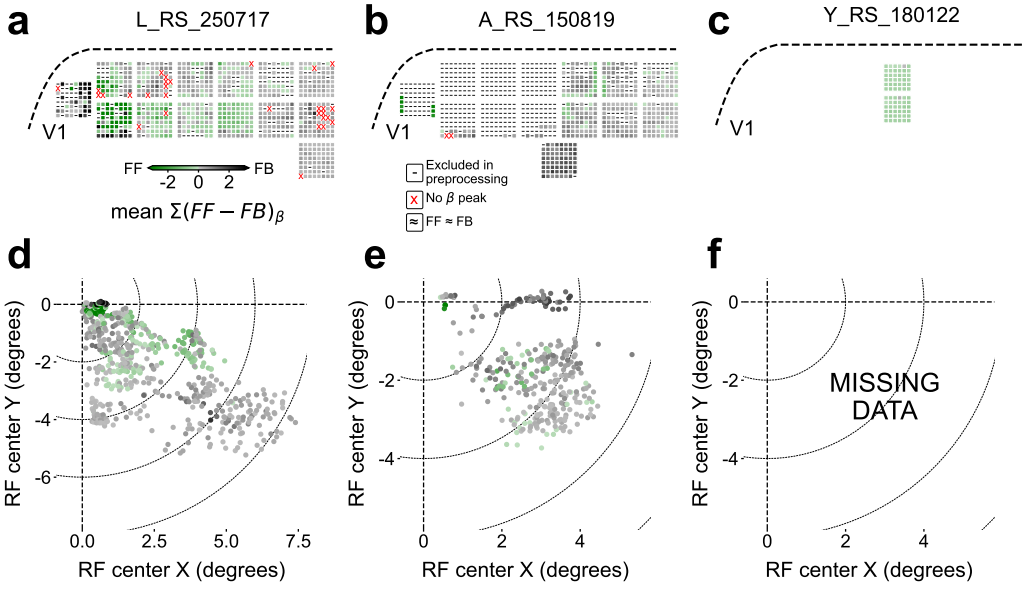


Fig. S11 Spatial distribution of Granger causality strength per electrode. One representative session shown per macaque. **a, b, c** Schematic representation of the electrode locations overlaid with the beta causality strength $\mathcal{B}$. **d, e, f** Receptive field (RF) map overlaid with $\mathcal{B}$. Note that stronger $\mathcal{B}$ is found around the foveal region. Receptive field data for macaque Y is missing for now but will be incorporated as soon as it is available.

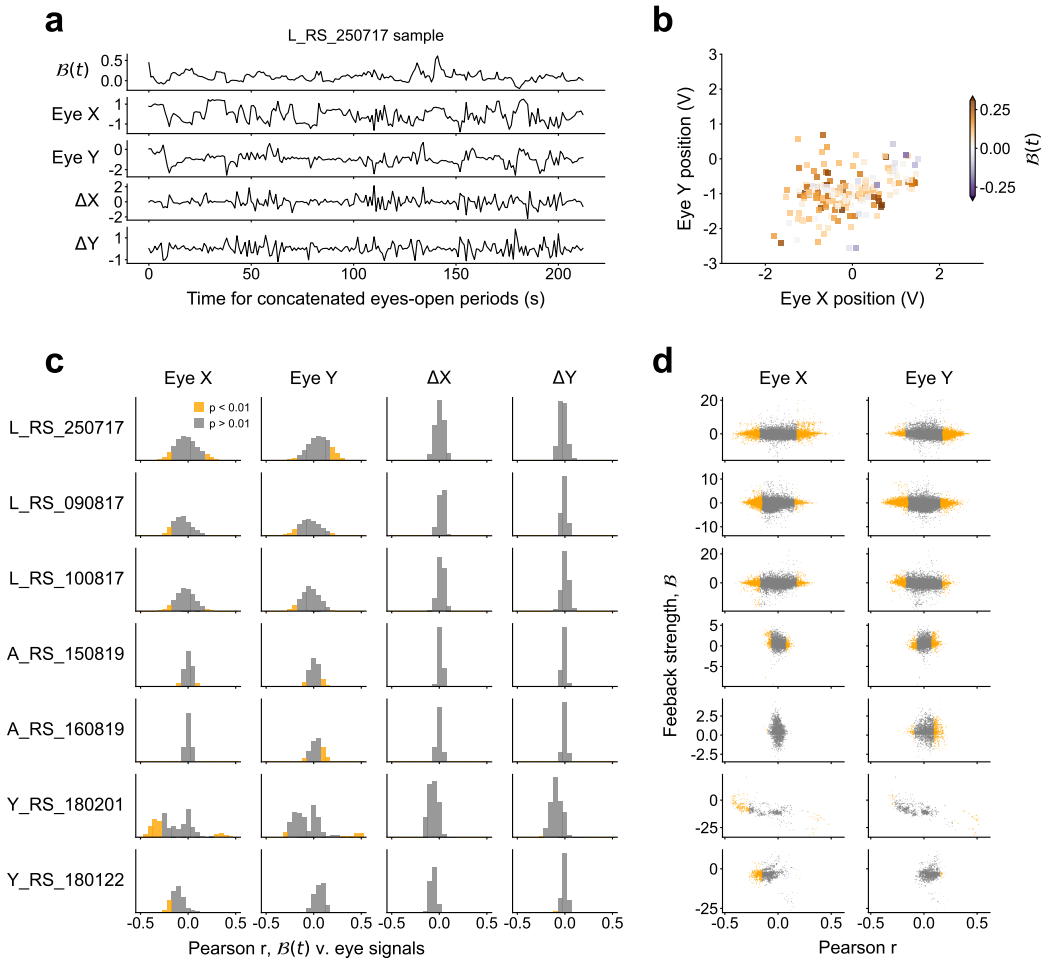


Fig. S12 Feedback signals are not correlated with gaze direction. **a** Sample traces of the mean beta-band Granger causality difference, gaze direction (Eye X, Eye Y), and gaze direction derivative (ΔX, ΔY). **b** Sample beta causality difference over the gaze locations. Higher feedback is not concentrated in particular regions. **c** Histograms of Pearson correlation coefficients between time-dependent causality difference and gaze signals, computed for all electrode pairs in all sessions. Significant ($p < 0.01$ two-tailed) part of histograms shown in orange. Gaze direction derivatives show no significant correlations. Note that we did not correct for multiple testing, since reducing the p-value threshold would simply reinforce our finding that no strong correlation was present between the gaze and the feedback. **d** Scatter plot of the summed time-independent causality difference against the correlation with gaze direction. There is no clear relation between $\mathcal{B}(t)$-gaze correlation and causality strength.

At many physiological and anatomical levels in the brain, the distribution of numerous parameters is strongly skewed with a heavy tail.

Buzsáki and Mizuseki (2014)

3

Lognormal distribution of neuron densities

Summary:

In this chapter we study the neuron density distribution in the mammalian neocortex. We show that neuron densities are compatible with a lognormal distribution across cortical areas in several mammalian species, a previously unknown feature. We find that this holds true for uniformly sampled regions across cortex of several species; even when accounting for several confounding variables, such as staining types or the presence of glia. The lognormal distribution also holds true within cortical areas of the marmoset. Additionally, we provide a simple model of cell division, which in the presence of Gaussian noise leads to lognormal distributions of neuron densities and, given distributed proliferation times, can also account for the coexistence of lognormal neuron density distributions both within and across the cortical areas. Finally, we compared the lognormal distribution with other statistical models and found that the lognormal distribution performs equally or better than all other models.

Ubiquitous lognormal distribution of neuron densities across mammalian cerebral cortex

Aitor Morales-Gregorio,[1,2,*,†] Alexander van Meegen,[1,2,*] Sacha J. van Albada[1,2]

[1]Institute of Neuroscience and Medicine (INM-6) and Institute for Advanced Simulation (IAS-6) and JARA-Institut Brain Structure-Function Relationships (INM-10), Jülich Research Centre, Jülich, Germany

[2]Institute of Zoology, University of Cologne, Cologne, Germany

∗These authors contributed equally to this work.

†To whom correspondence should be addressed; e-mail: a.morales-gregorio@fz-juelich.de

Abstract

Numbers of neurons and their spatial variation are fundamental organizational features of the brain. Despite the large corpus of data available in the literature, the statistical distributions of neuron densities within and across brain areas remain largely uncharacterized. Here, we show that neuron densities are compatible with a lognormal distribution across cortical areas in several mammalian species. We find that this also holds true for uniformly sampled regions across cortex as well as within cortical areas. Our findings uncover a new organizational principle of cortical cytoarchitecture. The ubiquitous lognormal distribution of neuron densities adds to a long list of lognormal variables in the brain.

Introduction

Neurons are not uniformly distributed across the cerebral cortex; their density varies strongly across areas and layers [1]. The neuron density directly affects short-range as well as long-range neuronal connectivity [2, 3]. Elucidating the distribution of neuron densities across the brain therefore provides insight into its connectivity structure and, ultimately, cognitive function. Additionally, statistical distributions are essential for the construction of computational models, which rely on predictive relationships and organizational principles where the experimental data are missing [4, 5]. Previous quantitative studies have provided reliable estimates for cell densities across the cerebral cortex of rodents [6, 7, 8], non-human primates [8, 9, 10, 11, 12, 13], large carnivores [14], and humans [15, 1]. However, to the best of our knowledge, the univariate distribution of neuron densities across and within cortical areas has not yet been statistically characterized. Instead, most studies focus on qualitative and quantitative comparisons across species, areas, or cortical layers. Capturing the entire distribution is necessary because long-tailed, highly skewed distributions are prevalent in the brain [16] and invalidate the intuition—guided by the central limit theorem—that the vast majority of values are in a small region of a few standard deviations around the mean.

Here, we for the first time characterize the distribution of neuron densities ρ across mammalian cerebral cortex. Based on the sample histograms (Figure 1) we hypothesize that ρ follows a lognormal distribution, similar to many other neuroanatomical and physiological variables such as synaptic strengths, axonal widths, and cortico-cortical connection densities [16, 17, 18]. Using neuron density data from mouse (*Mus musculus*), marmoset (*Callithrix jacchus*), macaque (*Macaca mulatta*), human (*Homo sapiens*), galago (*Otolemur garnettii*), owl monkey (*Aotus nancymaae*), and baboon (*Papio cynocephalus anubis*) we confirm this hypothesis for the given species (see Cell density data for a detailed description of the data). Going

beyond the distribution across cortical areas, we furthermore show that neuron densities within most areas of marmoset cortex are also compatible with a lognormal distribution. Moreover, we show that the lognormal distribution can emerge during neurogenesis from a simple cell division model with variability. Finally, we compare with several other distributions and find that none outperform the lognormal distribution as a model of the data within and across cortex.

Results

To test for lognormality, we take the natural logarithm, $\ln(\rho)$, which converts lognormally distributed samples into normally distributed samples (Figure 1B). We then test for normality of $\ln(\rho)$ using the Shapiro-Wilk (SW) test, the most powerful among a number of commonly applied normality tests [19]. Large outliers ($|\text{z-scored } \ln(\rho)| \geq 3$; marked with a red cross in Figure 1C) were excluded from the normality test. The removed outliers are area $\mathrm{V1}$ in macaque and marmoset, which have densities far outside the range for all other areas in both species, and area APir in marmoset, which has a noticeably distinct cytoarchitecture with respect to the rest of the cerebral cortex [9]. We denote different data sets for the same species with subscript indices (see Cell density data). The SW test concludes that the normality hypothesis of $\ln(\rho)$ cannot be rejected for mouse, marmoset, macaque$_1$, human, galago$_1$, owl monkey, and baboon (see Figure 1B). For the data sets macaque$_2$ and galago$_2$ the normality hypothesis is rejected ($p < 0.05$); however, in these data sets, the densities were sampled neither uniformly nor based on a cytoarchitectonic parcellation. The normality hypothesis for the distribution across cytoarchitectonic areas is further supported by Figure 1C, which shows that the relation between theoretical quantiles and ordered samples is almost perfectly linear except for macaque$_2$ and galago$_2$. Next, we test the z-scored $\ln(\rho)$ from the different species and data sets against each other and find that they are pairwise statistically indistinguishable ($\alpha = 0.05$ level; two-sample two-sided Kolmogorov-Smirnov test, see Figure S1 for full test results).

Additionally, we control for cell types in the distributions of the mouse, galago$_1$, owl monkey, and baboon data. In the mouse data, different types of neurons and glia were labeled with specific genetic markers and their respective densities were reported separately for all cell types [7]. In the galago$_1$, owl monkey, and baboon data sets, the total numbers of cells and neurons were reported separately [11]. We show that all subtypes of neurons in the mouse are compatible with a lognormal distribution (Figure S2; SW test on $\ln(\rho)$, $p > 0.05$) while glia are not—with the notable exception of oligodendrocytes. When neurons and glia are pooled together (Figure S2C,D), the distribution of $\ln(\rho)$ still passes the SW normality test, likely due to the distribution being dominated by the neurons. Similar observations are made in the baboon data, where the glia do not pass the lognormality test, but the neurons do. In the cases of galago$_1$ and owl monkey both the neurons and glia pass the lognormality test (Figure S2), which may, however, be partly due to the small number of density samples (N=12 in both cases). Thus, the mouse and baboon data—with large samples sizes (N=42 and N=142, respectively)—suggest that it is the neuron densities that follow a lognormal distribution but not necessarily the glia densities.

Furthermore, we also perform a control test on the different types of staining—Nissl and NeuN—using the macaque$_1$ data. The staining methods differ in their treatment of glia: NeuN tends to label neuronal cell bodies only while Nissl indiscriminately labels both neurons and glia. We show that regardless of staining type the cell densities pass the lognormality test (Figure S3; SW test on $\ln(\rho)$ with $p > 0.05$), suggesting that counting some glia in the cell densities does not confound our analysis of the macaque$_1$ data.

Taken together, the normality test, the quantiles plots, the pairwise tests, the cell type comparison, and the staining method comparison provide compelling evidence that the logarithmized neuron densities are normally distributed across cytoarchitectonic areas. This also holds for uniformly sampled neuron densities (baboon) but not for a sampling that is neither uniform

nor based on a cytoarchitectonic parcellation (macaque$_2$, galago$_2$). Thus, the neuron densities are consistent with a lognormal distribution across the different cortical areas, as long as sampling is not irregular.

To investigate whether the lognormal distribution holds within cortical areas, we leverage numerical estimates of neuron density in marmoset [9]. Neurons were counted within $150 \times 150\ \mu m$ counting frames for four strips per cortical area, all originating from the same subject. The neuron densities within the counting frames ρ_s are the within-area samples; their sample distributions in three representative areas (MIP, V2, and V3; Figure 2A) again suggest a lognormal distribution. As before, we test for lognormality by testing $\ln(\rho_s)$ for normality with the SW-test (for full test results see Table S2). At significance level $\alpha = 0.05$, the normality hypothesis is not rejected for 86 out of 116 areas; whereas at $\alpha = 0.001$, this is the case for 112 out of 116 areas (Figure 2B,C). Thus, regardless of the precise significance threshold, the lognormality hypothesis cannot be rejected within most cortical areas in the marmoset cortex.

This finding raises the question how the intricate process of neurogenesis [20] culminates in lognormally distributed neuron densities in almost all areas. A simple model shows that there is no need for a specific regulatory mechanism: assuming that the proliferation of the neural progenitor cells is governed by a noisy rate

$$\lambda(t) = \mu(t) + \xi(t), \tag{1}$$

where $\mu(t)$ denotes the mean rate and $\xi(t)$ is a zero-mean Gaussian process, the resulting population of progenitor cells, and eventually neurons, is lognormally distributed (see Model of progenitor cell division with variability). Thus, the lognormal neuron density distribution within areas could be a hallmark of a cell division process with variability. The model furthermore predicts that the mean and variance of $\ln(\rho)$ increase with proliferation time. Since the proliferation time varies up to twofold between areas [20], mean and variance of $\ln(\rho)$ are correlated

across areas according to the model—indeed, they are significantly correlated in the marmoset data (Pearson $r = 0.32,\ p < 10^{-3}$, Figure S4).

To complement the statistical hypothesis tests on the logarithmic densities, we compared the lognormal model with six other statistical distributions based on the relative likelihood (see Statistical model comparison). We included statistical distributions with support in $\mathbb{R}^+$ since neuron densities cannot be negative: lognormal, truncated normal, inverse normal, gamma, inverse gamma, Lévy, and Weibull. Of those distributions the lognormal, inverse normal, and inverse gamma stand out as the distributions with the highest relative likelihoods, both across the entire cortex and within cortical areas (Figure S5A, Figure S6A). A visual inspection of the fitted distribution reveals that the lognormal, inverse normal, and inverse gamma produce virtually indistinguishable probability densities (Figure S5B, Figure S6C); indeed, the relative likelihoods of the three models are above 0.05 in all cases. This suggests that the data could theoretically be distributed according to either the lognormal, inverse normal, or inverse gamma distribution. However, out of these, the lognormal distribution could arise from a simple model of cell division (equation (1))—while no interpretable mechanisms leading to inverse normal or inverse gamma distributions are known in this context. Thus, the similar likelihood and a simple biophysical explanation together argue for a lognormal rather than an inverse normal or inverse gamma distribution of neuron densities.

Discussion

In conclusion, we show that neuron densities are compatible with a lognormal distribution across cortical areas in multiple mammalian cortices and within most cortical areas of the marmoset, uncovering a previously unexplored organizational principle of cerebral cortex. Furthermore, we propose a simple model, based on a cell division process of the progenitor cells with variability, that accounts for the emerging lognormal distributions within areas. Lastly, we show

that none of an extensive list of statistical models outperform the lognormal distribution. Our results are in agreement with the observation that surprisingly many characteristics of the brain follow lognormal distributions [16]. Moreover, this analysis highlights the importance of characterizing the statistical distributions of brain data because simple summary statistics—such as the mean or standard deviation—lack nuance and are not necessarily a good representation of the underlying distribution.

The distributions of neuron and cell densities in general depend on the underlying spatial sampling. We found that neuron densities follow a lognormal distribution within cytoarchitectonically defined areas, across such areas, and when averaged within small parcels uniformly sampled across cortex, but not when sampled in a highly non-uniform manner not following cytoarchitectonic boundaries. The observation of lognormality both within and across cytoarchitectonic areas as well as across small uniformly sized parcels suggests an interesting topic for further research: uncovering whether the neuron densities obey an invariance principle across scales.

In principle, cortex-wide organizational structures might be by-products of development or evolution that serve no computational function [21]—but the fact that we observe the same organizational principle for several species and across most cortical areas suggests that the lognormal distribution serves some purpose. Heterogeneous neuron densities could assist computation through their association with heterogeneity in other properties such as connectivity and neuronal time constants [4, 22]; indeed, such heterogeneity is known to be a valuable asset for neural computation [23, 24]. Alternatively, localized concentration of neurons in certain areas and regions could also serve a metabolic purpose [25], because centralization supports more efficient energy usage. This is particularly relevant since approximately half of the brain's energy consumption is used to support the communication between neurons [26]. Also from the perspective of cortical hierarchies it makes sense to have few areas with high neuron densities

and many areas with lower neuron densities: Low-density areas contain neurons with large dendritic trees [27] receiving convergent inputs from many neurons in high-density areas lower in the hierarchy. The neurons with extensive dendritic trees in higher areas are involved in different, area-specific abstractions of the low-level sensory information [28, 29]. There is probably not a single factor that leads to lognormal neuron densities in the cortex; further research will be needed to refine our findings and uncover the functional implications.

References

[1] C. von Economo, G. N. Koskinas, L. C. Triarhou, *Atlas of Cytoarchitectonics of the Adult Human Cerebral Cortex* (Karger, 2008).

[2] V. Braitenberg, A. Schüz, *Anatomy of the Cortex: Statistics and Geometry* (Springer-Verlag, Berlin, Heidelberg, New York, 1991).

[3] M. Ercsey-Ravasz, *et al.*, *Neuron* **80**, 184 (2013).

[4] C. C. Hilgetag, S. F. Beul, S. J. van Albada, A. Goulas **3**, 905 (2019).

[5] S. J. van Albada, *et al.*, *arXiv* (2020).

[6] S. Herculano-Houzel, C. Watson, G. Paxinos, *Frontiers in Neuroanatomy* **7** (2013).

[7] C. Erö, M.-O. Gewaltig, D. Keller, H. Markram, *Frontiers in Neuroinformatics* **12**, 84 (2018).

[8] C. J. Charvet, D. J. Cahalane, B. L. Finlay, *Cerebral Cortex* **25**, 147 (2015).

[9] N. Atapour, *et al.*, *Cerebral cortex* **29**, 3836 (2019).

[10] S. F. Beul, C. C. Hilgetag, *NeuroImage* **189**, 777 (2019).

[11] C. E. Collins, D. C. Airey, N. A. Young, D. B. Leitch, J. H. Kaas, *Proceedings of the National Academy of Sciences* **107**, 15927 (2010).

[12] C. E. Collins, *et al.*, *Proceedings of the National Academy of Sciences* **113**, 740 (2016).

[13] E. C. Turner, *et al.*, *Brain, Behavior and Evolution* **88**, 1 (2016).

[14] D. Jardim-Messeder, *et al.*, *Frontiers in Neuroanatomy* **11**, 118 (2017).

[15] C. S. von Bartheld, J. Bahney, S. Herculano-Houzel, *Journal of Comparative Neurology* **524**, 3865 (2016).

[16] G. Buzsáki, K. Mizuseki, *Nature Reviews Neuroscience* **15**, 264 (2014).

[17] N. T. Markov, *et al.*, *Cerebral Cortex* **24**, 17 (2014).

[18] P. A. Robinson, X. Gao, Y. Han, *Biological Cybernetics* **115**, 121 (2021).

[19] N. M. Razali, B. W. Yap, *Journal of Statistical Modeling and Analytics* **2**, 21 (2011).

[20] P. Rakic, *Nature Reviews Neuroscience* **3**, 65 (2002).

[21] A. G. Otopalik, A. C. Sutton, M. Banghart, E. Marder, *eLife* **6**, e23508 (2017).

[22] W. Rall, *Biophysical Journal* **9**, 1483 (1969).

[23] R. Duarte, A. Morrison, *PLOS Computational Biology* **15**, e1006781 (2019).

[24] N. Perez-Nieves, V. C. H. Leung, P. L. Dragotti, D. F. M. Goodman, *Nature Communications* **12**, 5791 (2021).

[25] M. Bélanger, I. Allaman, P. Magistretti, *Cell Metabolism* **14**, 724 (2011).

[26] S. B. Laughlin, T. J. Sejnowski, *Science* **301**, 1870 (2003).

[27] G. N. Elston, M. Rosa, *Cerebral Cortex* **8**, 278 (1998).

[28] S. Kumar, K. E. Stephan, J. D. Warren, K. J. Friston, T. D. Griffiths, *PLoS Computational Biology* **3**, e100 (2007).

[29] S. L. Brincat, M. Siegel, C. von Nicolai, E. K. Miller, *Proceedings of the National Academy of Sciences* **115** (2018).

[30] E. S. Lein, *et al.*, *Nature* **445**, 168 (2007).

[31] H. W. Dong, *The Allen reference atlas: A digital color brain atlas of the C57Bl/6J male mouse.* (John Wiley & Sons inc., 2008).

[32] G. Paxinos, C. R. R. Watson, M. Petrides, M. G. Rosa, H. Tokuno, *The Marmoset Brain in Stereotaxic Coordinates* (2012).

[33] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry* (North Holland, 2007), third edn.

[34] C. A. Braumann, *Mathematical Biosciences* **206**, 81 (2007).

Acknowledgements

We thank Günther Palm for useful discussions, Alexandre René for useful discussions and help with the Bayesian model comparison, Anno Kurth for discussions about geometric Brownian motion, and Jon Martinez Corral for proofreading an early draft. **Funding:** This work was supported by the European Union Horizon 2020 Framework Programme for Research and Innovation (Human Brain Project SGA2 grant number 785907 and HBP SGA3 grant number 945539) and the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Priority Program (SPP 2041 ”Computational Connectomics”) [S.J. van Albada: AL 2041/1-1 and

2041/1-2] and Open Access Publication Costs grant 491111487. **Author contributions:** Conceptualization AMG, AvM, SvA; Data curation AMG; Formal Analysis AMG, AvM; Funding acquisition SvA; Investigation AMG, AvM, SvA; Methodology AMG, AvM, SvA; Project administration SvA; Resources SvA; Software AMG, AvM; Supervision SvA; Validation AvM; Visualization AMG; Writing - original draft AMG, AvM; Writing – review & editing AMG, AvM, SvA. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** This work produced no new data and instead relied on a corpus of neuron density data available from the literature, which we gratefully acknowledge; see Cell density data for a detailed description.

Figures

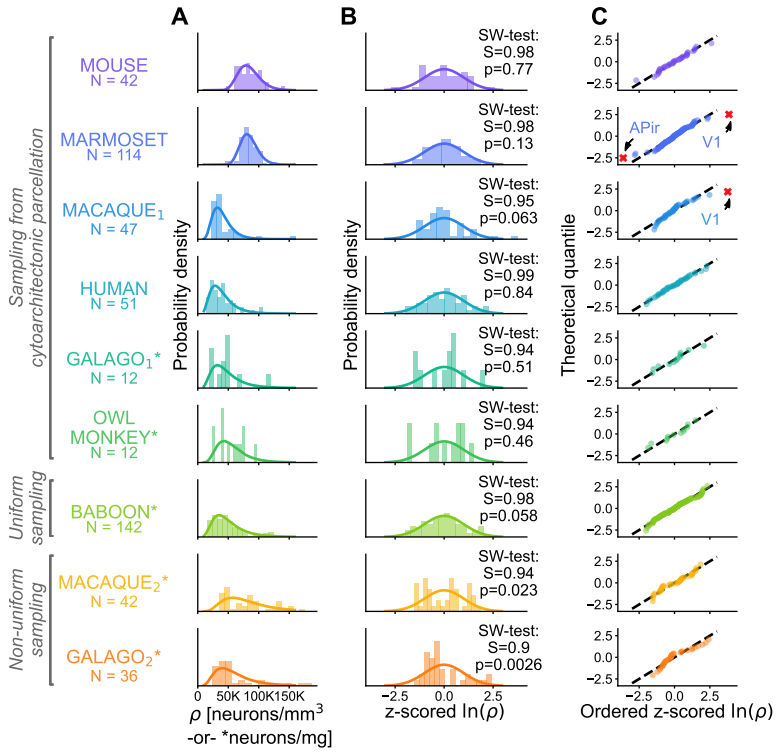


Figure 1: Neuron and cell densities ρ follow a lognormal distribution across cortical areas for multiple species. **A** Histogram of ρ (bars) and probability density function of a fitted lognormal distribution (line). **B** Z-scored $\ln(\rho)$ histogram (bars), standard normal distribution (line), and result of the Shapiro-Wilk normality test. **C** Probability plot of z-scored $\ln(\rho)$. Discarded outliers marked with a red cross.

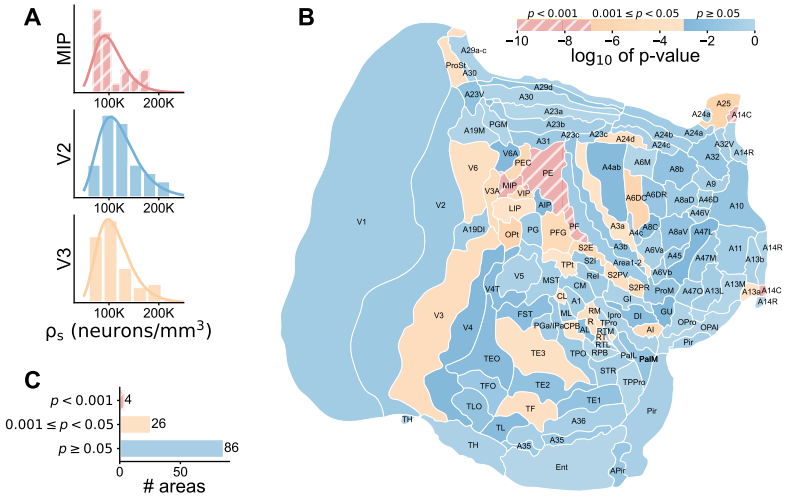


Figure 2: Neuron densities ρ_s follow a lognormal distribution within most areas of marmoset cortex. **A** Sample histograms of ρ_s and fitted lognormal distributions for three areas representing different degrees of lognormality. **B** Log_{10} of p-value of Shapiro-Wilk normality test of $\ln(\rho_s)$ on a flattened representation of the marmoset cortex [9]. **C** Number of areas with p-values in the given significance ranges.

Materials and methods

Cell density data

Estimates of neuron density for the available cortical areas across the mouse (*Mus musculus*), marmoset (*Callithrix jacchus*), macaque (*Macaca mulatta*), human (*Homo sapiens*), galago (*Otolemur garnettii*), owl monkey (*Aotus nancymaae*), and baboon (*Papio cynocephalus anubis*) cerebral cortex were used in this study.

In the cases of mouse, marmoset, macaque$_1$, human, galago$_1$, and owl monkey the data were sampled from standard cytoarchitectonic parcellations; abbreviated names for all areas are listed in Table S1. Note that we use subscript indices to distinguish between different data sets on the same model animal, e.g. macaque$_1$ and macaque$_2$.

Neuron density estimates for the mouse were published in [7], and were measured from Nissl-body-stained slices, where genetic markers were used to distinguish between cell types. The data were provided in the Allen Brain Atlas parcellation [30, 31].

Neuron density estimates for the marmoset cortex were published in [9], and were measured from NeuN-stained slices. The data were provided in the Paxinos parcellation [32]. Neuron densities within each counting frame used in the original publication [9] (their Figure S1) were obtained via personal communication with Nafiseh Atapour, Piotr Majka, and Marcello G. Rosa.

The neuron density estimates in the first macaque data set, macaque$_1$, were previously published in visual form in [10], and were obtained from both Nissl-body- and NeuN-stained brain slices. Counts based on Nissl-body staining were scaled according to a linear relationship with the counts from NeuN staining obtained from selected areas where both types of data were available [10]. The data follow the M132 parcellation [17] and numerical values were provided by Sarah F. Beul and Claus C. Hilgetag via personal communication.

Cell density estimates for the human cortex were previously published in [1], and were

measured from Nissl-body-stained brain slices. The human data therefore most likely reflect combined neuron and glia densities. The data were provided in the von Economo parcellation [1].

Cell and neuron density estimates for galago$_{1\&2}$, owl monkey, baboon, and macaque$_2$ were previously published in [11], and were measured using the isotropic fractionator method. The data are sampled from common parcellation schemes in galago$_1$ and owl monkey, approximately equal-size samples in the baboon, and irregular non-uniform samples in macaque$_2$ and galago$_2$.

Statistical model comparison

In order to assess which model is most compatible with the data, we compared the relative likelihood of different distributions against each other. We included an extensive list of distributions with support on $\mathbb{R}^+$, estimated the distributions' parameters using maximum likelihood, and calculated the Akaike Information Criterion (AIC)

$$AIC = 2k - 2\ln\mathcal{L} \tag{2}$$

where k is the number of estimated parameters of the model and $\mathcal{L}$ is the estimated maximum likelihood. We further compare the models using the relative likelihood ($\mathcal{L}_r$)

$$\mathcal{L}_r = e^{(AIC_{\min} - AIC_i)/2} \tag{3}$$

where $AIC_{\min}$ is the minimum AIC across all models and AIC_i is the AIC for the ith model. Note that the relative likelihood is equal to the relative likelihood if the number of estimated parameters is the same in both models. The relative likelihood indicates the probability that, from among the tested models, the ith model most strongly limits the information loss. We take a significance threshold of $\alpha = 0.05$ on the relative likelihood to determine whether a model is significantly worse than the best possible model.

Model of progenitor cell division with variability

We assume that the proliferation of the neural progenitor cells is governed by a noisy rate

$$\lambda(t) = \mu_{\text{rate}} + \sigma_{\text{rate}}\xi(t), \tag{4}$$

where μ_{rate} denotes the mean rate, $\xi(t)$ is a zero-mean Gaussian white noise process, and σ_{rate} controls the strength of the noise. During proliferation, we assume that the population size of the progenitor cells grows exponentially with rate λ, i.e., it obeys $\frac{d}{dt}N = \lambda N$. Dividing by a reference volume and inserting equation (4), we obtain a stochastic differential equation (SDE) for the density of progenitor cells ρ:

$$\frac{d}{dt}\rho = (\mu_{\text{rate}} + \sigma_{\text{rate}}\xi(t))\rho \tag{5}$$

We here use the Stratonovich interpretation, i.e., we assume that the noise process has a small but finite correlation time before taking the white-noise limit [33].

Working in the Stratonovich interpretation, we can transform the SDE to $\frac{d}{dt}\ln\rho = \mu_{\text{rate}} + \sigma_{\text{rate}}\xi(t)$ with the solution [34]

$$\ln\rho(t) = \ln\rho_0 + \mu_{\text{rate}}t + \sigma_{\text{rate}}\int_0^t \xi(s)ds. \tag{6}$$

Since $\xi(t)$ is Gaussian and equation (6) is linear, $\ln\rho(t)$ is Gaussian and hence $\rho(t)$ is log-normally distributed. The parameters of this lognormal distribution are $\mu(t) = \langle\ln\rho(t)\rangle$ and $\sigma^2(t) = \langle\Delta(\ln\rho(t))^2\rangle$. Using equation (6), $\langle\xi(s)\rangle = 0$, and $\langle\xi(s)\xi(s')\rangle = \delta(s-s')$, we obtain [34]

$$\mu(t) = \ln\rho_0 + \mu_{\text{rate}}t \qquad \text{and} \qquad \sigma^2(t) = \sigma^2_{\text{rate}}t. \tag{7}$$

Thus, the neuron densities resulting from the model of cell division with variability, equation (5), are lognormally distributed with parameters $\mu(t)$ and $\sigma^2(t)$ specified in equation (7). In particular, equation (7) predicts that both parameters increase with the proliferation time t.

The model can be generalized while still leading to a lognormal distribution of neuron densities: 1) The mean rate can be time-dependent, $\mu_{\text{rate}} = \mu_{\text{rate}}(t)$. 2) The noise process can be an arbitrary zero-mean (a non-zero mean can always be incorporated into $\mu_{\text{rate}}(t)$) Gaussian process with correlation function $C_\xi(t, t')$. Both generalizations allow one to incorporate a time dependence of mean and noise strength during the proliferation. Assuming an absence of correlation between noise and neuron density prior to $t = 0$, the above steps lead to the generalized solution

$$\ln \rho(t) = \ln \rho_0 + \int_0^t \mu_{\text{rate}}(s)ds + \int_0^t \xi(s)ds. \tag{8}$$

Here, $\ln \rho(t)$ is still a Gaussian process, because it is a linear transformation of the Gaussian process $\xi(t)$. Due to the marginalization property of Gaussian processes, $\ln \rho(t)$ is normally distributed for any fixed time t with parameters

$$\mu(t) = \ln \rho_0 + \int_0^t \mu_{\text{rate}}(s)ds \qquad \text{and} \qquad \sigma^2(t) = \int_0^t \int_0^t C_\xi(s, s')dsds'. \tag{9}$$

Thus, $\rho(t)$ is lognormally distributed with parameters $\mu(t)$ and $\sigma^2(t)$ specified in equation (9). Note that in equation (9), in contrast to equation (7), $\mu(t)$ and $\sigma^2(t)$ do not necessarily grow linearly with time but may exhibit a more intricate temporal dependence. Nonetheless, equation (9) predicts that $\mu(t)$ and $\sigma^2(t)$ are related through the proliferation time.

Supplementary tables

Table S1: Cortical areas included in this study.

Species	Area abbreviations
Mouse	FRP, MOp, MOs, SSp, SS-n, SSp-bfd, SSp-ll, SSp-m, SSp-ul, SS-tr, SSs, VISC, AUDd, AUDp, AUDpo, AUDv, VISal, VISam, VISl, VISp, VISpl, VISpm, ACAd, ACAv, ACAv, ACAv, ORBl, ORBm, ORBvl, AId, AIp, AIv, RSPagl, RSPd, RSPv, AONd, AONe, AONl, AONm, AONpv, TTd, TTv
Marmoset	A10, A9, A46V, A46D, A8aD, A8b, A8aV, A47L, A47M, A45, A47O, ProM, A11, A13b, A13a, A13L, A13M, OPAl, OPro, Gu, A32, A32V, A14R, A14C, A25, A24a, A24b, A24c, A24d, A6DR, A6Vb, A6Va, A8C, A6M, A6DC, A4c, A4ab, PaIM, AI, PaIL, DI, GI, IPro, TPro, S2PR, A3a, S2PV, A3b, S2I, S2E, A1-2, AuRTL, AuRT, AuRPB, AuRTM, AuR, AuRM, AuAL, AuA1, AuCM, AuCPB, AuML, AuCL, TPPro, STR, TE1, TPO, ReI, TE2, PGa-IPa, TPt, TE3, TEO, Pir, APir, Ent, A36, A35, TF, TL, TH, TLO, TFO, A23c, A23a, A29d, A30, A23b, A29a-c, A23V, ProSt, PF, PE, PFG, A31, AIP, PG, PEC, VIP, LIP, PGM, V6A, OPt, MIP, MST, FST, V5, V4T, A19M, V3A, V4, V6, A19DI, V3, V2, V1
Macaque_1	2, 5, 9, 10, 11, 12, 13, 14, 23, 25, 32, 24a, 24c, 24d, 46d, 46v, 7A, 7B, 7m, 8B, 8l, 8m, 8r, 9-46d, 9-46v, DP, ENTO, F1, F2, F3, F4, F5, F6, F7, LIP, MT, OPAI, OPRO, PERI, STPi, TEad, TEav, TEO, TH-TF, V1, V2, V3A, V4
Human	FA, FB, FC, FCBm, FD, FDΔ, FDt, FE, FF, FG, FH, FJ, FK, FL, FM, FN, LA1, LA2, LC1, LC2, LC3, LD, LE1, LE2, IA, IB, OA, OB, OC, PA, PB1, PB2, PC, PD, PE, PF, PG, PH, HA, HB, HC, HD, HE, HF, TA, TB, TC, TD, TE, TF, TG
Galago_1 & Owl Monkey	V1, V2, dV3, vV3, S1, M1, A1, MT, premotor, DL, Remain Ctx, Surr Ctx

Table S2: Results of the Shapiro-Wilk test for normality of $\ln(\rho_s)$ in marmoset cortical areas. Values rounded to two significant digits.

Area	S	p-value	Area	S	p-value	Area	S	p-value
V1	0.97	0.39	AI	0.95	0.0043	TH	0.97	0.66
A10	0.95	0.19	PaIL	0.95	0.33	TLO	0.96	0.18
A9	0.98	0.51	DI	0.97	0.098	TFO	0.97	0.26
A46V	0.98	0.56	GI	0.97	0.67	A23c	0.97	0.36
A46D	0.98	0.49	Ipro	0.97	0.66	A23a	0.99	0.98
A8aD	0.97	0.34	TPro	0.97	0.77	A29d	0.95	0.21
A8b	0.96	0.16	S2PR	0.92	0.006	A30	0.98	0.73
A8aV	0.96	0.17	A3a	0.95	0.04	A23b	0.97	0.45
A47L	0.96	0.052	S2PV	0.93	0.014	A29a-c	0.97	0.70
A47M	0.97	0.30	A3b	0.96	0.20	A23V	0.96	0.15
A45	0.96	0.18	S2I	0.97	0.33	ProSt	0.93	0.018
A47O	0.98	0.70	S2E	0.94	0.0046	PF	0.94	0.00083
ProM	0.97	0.21	Area1-2	0.97	0.37	PE	0.94	0.00065
A11	0.97	0.41	AuRTL	0.97	0.40	PFG	0.92	0.0046
A13b	0.96	0.58	AuRT	0.97	0.031	A31	0.97	0.31
A13a	0.91	0.048	AuRPB	0.98	0.89	AIP	0.96	0.063
A13L	0.97	0.45	AuRTM	0.97	0.73	PG	0.99	0.37
A13M	0.99	0.97	AuR	0.98	0.0093	PEC	0.91	0.0032
OPAl	0.99	0.99	AuRM	0.9	0.017	VIP	0.92	0.0044
OPro	0.98	0.75	AuAL	0.94	0.12	LIP	0.95	0.042
GU	0.95	0.058	AuA1	0.98	0.48	PGM	0.98	0.78
A32	0.97	0.20	AuCM	0.97	0.33	V6A	0.95	0.068
A32V	0.96	0.51	AuCPB	0.93	0.037	OPt	0.91	0.0015
A14R	0.98	0.77	AuML	0.97	0.44	MIP	0.9	0.00091
A14C	0.79	5.5e-06	AuCL	0.94	0.045	MST	0.98	0.53
A25	0.89	0.022	TPPro	0.98	0.91	FST	0.95	0.10
A24a	0.96	0.35	STR	0.96	0.44	V5	0.98	0.68
A24b	0.97	0.41	TE1	0.96	0.17	V4T	0.95	0.082
A24c	0.97	0.54	TPO	0.97	0.31	A19M	0.98	0.80
A24d	0.92	0.017	ReI	0.95	0.40	V3A	0.91	0.006
A6DR	0.97	0.23	TE2	0.96	0.15	V4	0.97	0.064
A6Vb	0.97	0.32	PGa/IPa	0.97	0.45	V6	0.96	0.017
A6Va	0.98	0.56	TPt	0.94	0.033	A19DI	0.95	0.074
A8C	0.95	0.055	TE3	0.93	0.026	V3	0.95	0.0076
A6M	0.99	0.98	TEO	0.95	0.087	V2	0.96	0.29
A6DC	0.91	0.002	A36	0.98	0.54	Ent	0.99	0.99
A4c	0.97	0.43	A35	0.97	0.31	APir	0.94	0.24
A4ab	0.96	0.076	TF	0.96	0.021	Pir	0.97	0.53
PaIM	0.93	0.20	TL	0.98	0.084			

Supplementary figures

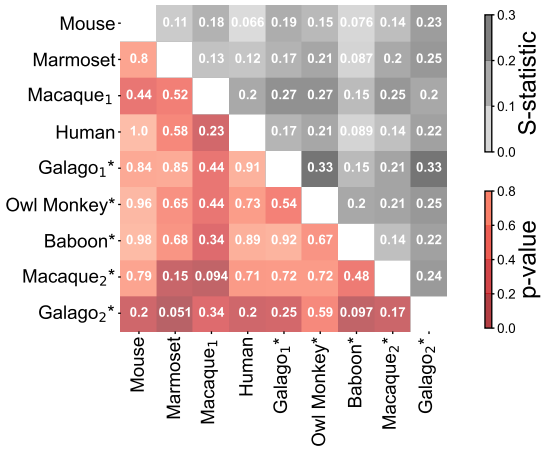


Figure S1: The z-scored log neuron density distributions of the four species are statistically indistinguishable at the 0.05 level based on pairwise Kolmogorov-Smirnov two-sample two-sided tests. P-values and S-statistics displayed below and above the diagonal, respectively.

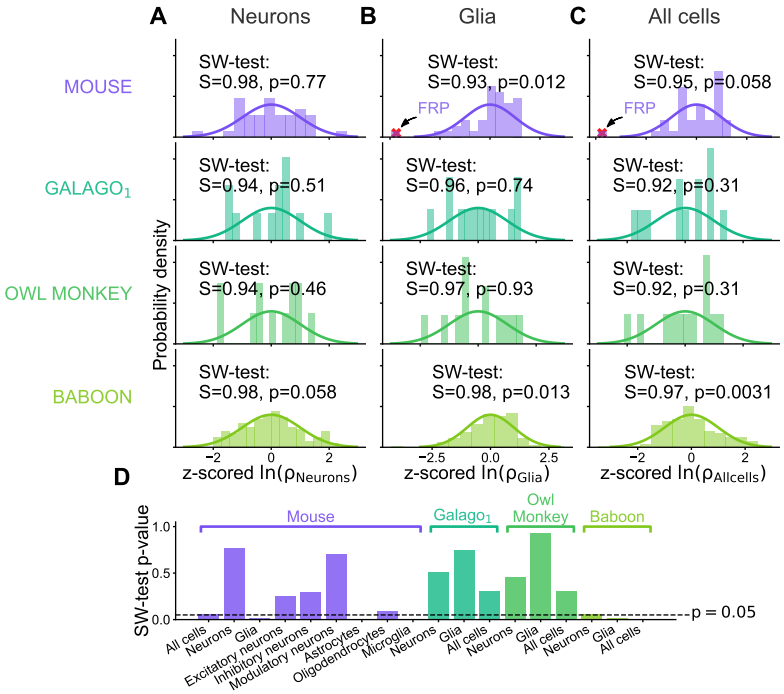


Figure S2: Comparison of neuron and glia lognormality. **A**–**C** Histogram of z-scored log density and result of Shapiro-Wilk test for neurons (**A**), glia (**B**), and all cells combined (**C**). **D** Barplot of p-values resulting from Shapiro-Wilk normality test for all cell types. Panel **A** is equivalent to the data shown in Figure 1.

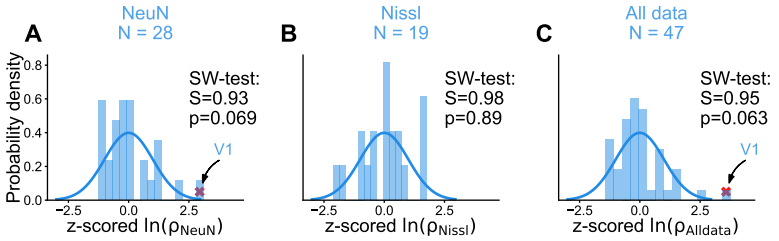


Figure S3: Lognormality of cell densities from different staining types in macaque cortex based on the macaque$_1$ data set. **A**-**C** Histogram of z-scored log density and result of Shapiro-Wilk test for NeuN staining only (**A**), Nissl staining only (**B**) and all measurements combined (**C**). The Nissl data were scaled down based on the linear relationship with the NeuN data [10]. Red crosses indicate outliers ($|\text{z-scored } \ln(\rho)| \geq 3$, which were excluded from the test. Panel **C** is equivalent to the data shown in Figure 1.

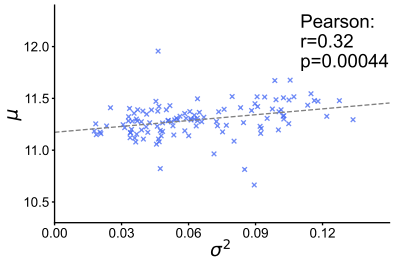


Figure S4: Neuron densities in the marmoset are compatible with our model of progenitor cell division with variability. μ and σ^2 are the mean and variance of $\ln(\rho)$, respectively; and are significantly correlated with each other, as predicted by the model.

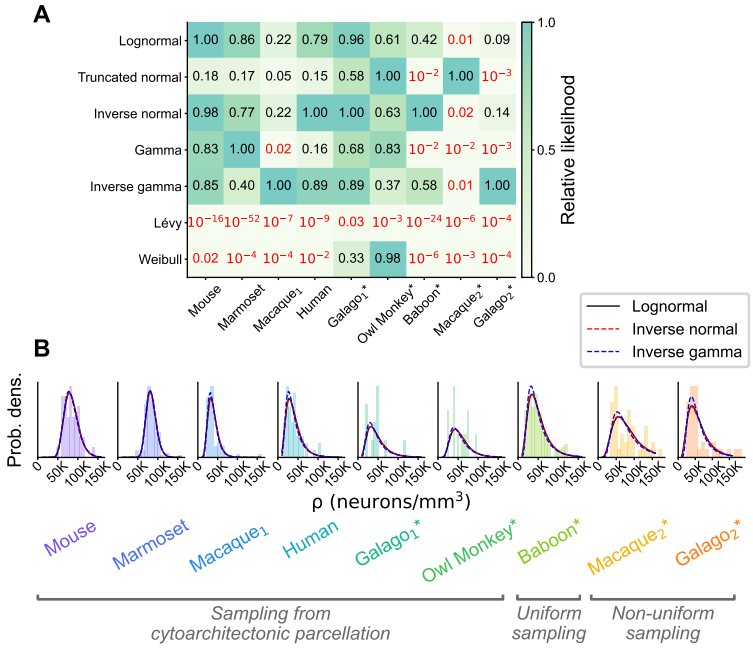


Figure S5: Statistical model comparison across the entire cortex of different animals. **A** Relative likelihood for seven compatible statistical models for all available area-level neuron density data sets; numerical values indicated for each model and animal. The red color indicates a relative likelihood < 0.05 with respect to the model with the highest likelihood. **B** The three best statistical models (according to the relative likelihood) fitted to the neuron density histograms in each animal; the three models produce visually nearly indistinguishable fits.

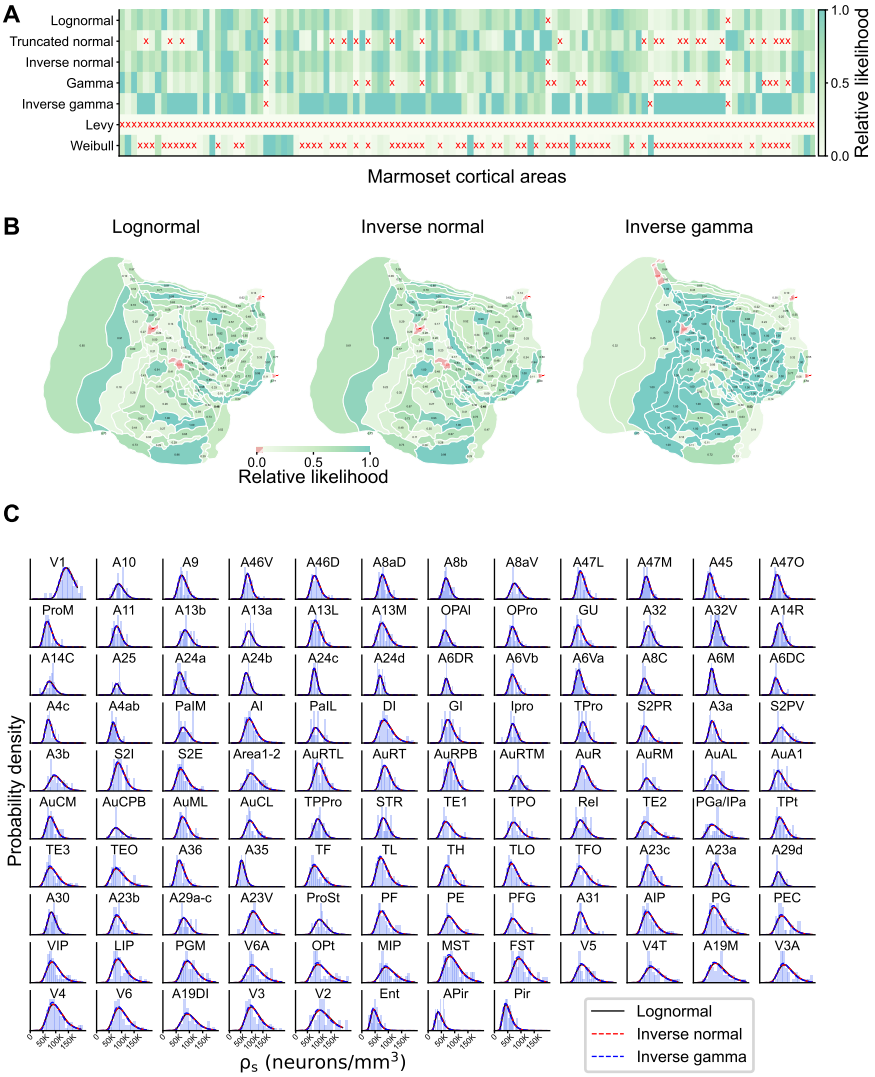


Figure S6: Statistical model comparison within the marmoset cortical areas. **A** Relative likelihood for seven compatible statistical models for all areas of the marmoset; a red cross (x) indicates a relative likelihood < 0.05 with respect to the model with the highest likelihood. **B** Spatial distribution of relative likelihood for the three best statistical models. **C** The three best statistical models fitted to the neuron density histograms in each area of marmoset cortex; the three models produce visually nearly indistinguishable fits.

The macroscopic constructional principles of the vertebrate brain, especially those of primates, have a common organisational layout and differ only in the relative proportions of certain regions.

Glees (1988)

4

Quantification of neuron density and white matter distance in macaque neocortex

Summary:

In this chapter we estimated the neuron density and white matter across the neocortex of one macaque (*macaca mulatta*). Improving on previous methods, we reconstructed a 3D volume from Nissl-stained histological slices of the macaque brain and estimated the cortical area locations on a standard parcellation. We then used automated counting techniques to estimate the number of neurons found within each area and validated our findings against previous reports. Finally, we also measured the distance between all cortical areas through the white matter. Our findings fill a gap in the knowledge, since reliable quantitative measurements of neuron density and white matter distance were missing for large parts of the cerebral cortex of the macaque.

Quantification of layer-resolved neuron density and white matter distance in the macaque neocortex

Aitor Morales-Gregorio[1,2*], Rembrandt Bakker[1,3] and Sacha J. van Albada[1,2]

[1]Institute of Neuroscience and Medicine (INM-6) and Institute for Advanced Simulation (IAS-6) and JARA-Institut Brain Structure-Function Relationships (INM-10), Jülich Research Centre, Jülich, Germany.
[2]Institute of Zoology, University of Cologne, Cologne, Germany.
[3]Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Nijmegen, Netherlands.

*Corresponding author. a.morales-gregorio@fz-juelich.de

Abstract

Neuron density and white matter distance are crucial parameters to understand the organization and structure of the cerebral cortex. Neuron densities have been quantitatively reported with great detail for a plethora of mammals over the decades. However, no comprehensive dataset of area- and layer-resolved macaque neuron densities across cortex has been published to date. The existing data is a compilation of different studies, with neuron densities recorded from different individuals, using different staining methods. Here, we address this gap in the knowledge by estimating the area- and layer-resolved neuron densities across the macaque cerebral cortex within the same subject. For this aim, we reconstructed the brain volume of one macaque from Nissl-body stained slices and matched it to an standard cortical parcellation. We then extracted counting strips and used an automated workflow to detect neurons in the images. Neuron densities were estimated after correcting for tissue shrinkage and stereological errors, and validated against previous reports of macaque neuron density and their distributions. Finally, we also provide new measurements of cortico-cortical distance through the white matter. In conclusion, we provide a comprehensive account of the macaque cortical structure, including neuron density distribution and white matter distance.

Keywords: Neuron density, macaque, cortex, white matter distance, brain structure

Introduction

The macaque is a widely studied model animal in neuroscience. Studies using histological data have measured relevant structural parameters such as neuron densities [1], long-range connectivity [2, 3], and receptor types [4] across large swathes of the macaque brain. However, there is a large uncertainty regarding the organization of the macaque brain due to the limited number of studied subjects, the high inter-subject variability, missing measurements for several parts of the brain and the overall difficulty of holistically studying such a complex system. Additionally, the lack of consensus on a cortical parcellation further complicates the compilation of the already existing anatomical data [5].

Neuron densities across the cerebral cortex are of particular interest for modeling [6] and comparative studies [7]. Neuron and cell density have been quantitatively estimated in rodents [8–10], non-human primates [1, 10–14], large carnivores [15], and humans [16, 17]. In particular, comprehensive layer- and area-resolved neuron density estimates exist for mouse [9] and marmoset [11]. However, to the best of our knowledge, no comprehensive map of layer-resolved macaque neuron densities across cortex has been reported to date. The most complete dataset [1] is a collation of several subjects and measurements, not a comprehensive analysis of a single subject. Given the large variability across subjects it is imperative to estimate neuron densities within the same subject.

To close this gap in the knowledge, we for the first time estimate the neuron densities across all cortical areas for a single macaque (*Macaca mulatta*). In order to do so we used a new semi-automatic workflow to reconstruct the full brain volume of a macaque from histological slice data. We used open tools [18, 19] and extended existing methods (PoSSum) [20] to create a 3D reconstruction. The reconstructed volume was then matched to a standard atlas [3], and counting strips were extracted for all cortical areas. Finally, we estimated the neuron densities by counting neurons within standardized counting frames, using a state-of-the-art machine learning based classifier [21]. Additionally, we also measure the distance through the white matter for all cortical area pairs using a dedicated heuristic. This new dataset will be a valuable resource for neuroscientists studying the macaque as a model animal.

Materials and methods

Histology and microscopy

To estimate the neuron density across areas and layers in the macaque cortex we used data from the Allen Institute NIH Blueprint Non-Human Primate (NHP) Atlas [22]. The data consists of histological slices for one macaque (*Macaca mulatta*).

The *ex vivo* brain was embedded in a gelatin block, frozen and cut into 50 μm thick coronal sections. Every 5$^{\text{th}}$ section was stained with cresyl violet (Nissl-body stain) and scanned at 1 μm/pixel plane resolution with a

ScanScope® scanners (Aperio Technologies, Inc.; Vista, CA) microscope. The resulting dataset consists of 276 high resolution images of Nissl-body stained cortical slices. See the technical white paper for a more detailed protocol [23].

Volume reconstruction

We reconstructed an anisotropic 3D volume from the histological slices from one rhesus macaque.

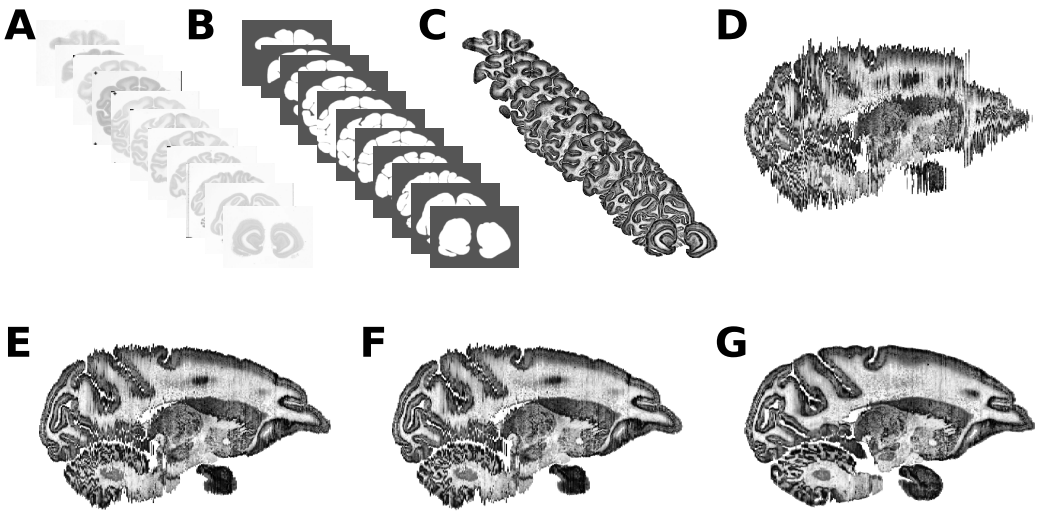


Fig. 1 Image processing workflow for 3D volume reconstruction. **A** Low resolution images of Nissl stained brain slices. **B** Manually created masks, with special attention to the delineation of sulci. **C** Masked and equalized images. **D** Lateral slice of 3D volume, which was created by aligning centers of mass of the coronal slices. **E** Result of pairwise registration of the stained slices to a reference MRI image. **F** Result after sequential and coarse-to-fine alignment. **G** Final reconstructed volume after deformable sequential alignment.

Image pre-processing

First, we removed the image background. In order to do so, we applied a Gaussian filter ($\sigma = 5$) and estimated the background color gradient between the image corners (100×100 pixel, excluding dark colors to avoid artifacts, pixel intensity < 30). We subtracted the estimated background gradient from the images and applied a threshold (pixel intensity > 248) to obtain an automatically estimated rough binary mask. Next, the binary mask was dilated (radius = 20 pixel) and subsequently eroded (radius = 10 pixel), note that dilation increased the black background (smaller mask) while erosion reduced it (larger mask). The dilation-erosion step aims to remove small artifacts such as thin filaments and protrusions. Then, we searched for small specks (surface $<$ 400 pixel) and holes (surface $<$ 2000 pixel) in the image and

removed them by filling them with the opposite color. Finally, we manually processed the automatically estimated masks, to exclude any remaining artifacts (Figure 1B). The manual mask processing was a time intensive task.

Second, we processed the images to have uniform color across slices. We applied Contrast Limited Adaptive Histogram Equalization (CLAHE) [24] to the raw images (clip limit = 1, 20×20 tiles), followed by an adaptive histogram equalization (clip limit = 0, 8×8 tiles). Then, we applied a Gaussian filter ($\sigma = 5$). Finally, the previously estimated masks were used to remove the background from the images (Figure 1C).

Volume reconstruction with PoSSum

An initial 3D volume was created by aligning the pre-processed images along their centers of mass (Figure 1D). In order to correctly align the slices in 3D we used PoSSum [20], which we updated to Python 3.x. We used the following steps from the PoSSum toolbox: 1) affine pairwise alignment to a reference volume (Figure 1E), 2) sequential affine registration, 3) coarse-to-fine registration (Figure 1F) and 4) sequential deformable registration (Figure 1G). The reference volume was a high resolution Magnetic Resonance Image (MRI) that belongs to the Calabrese et al. [25] atlas.

Atlas registration

A semi-automatic workflow was used to register the reconstructed volume to the M132 91 area atlas [3]. We first registered the M132 atlas to the same reference volume as the histological slices using the same steps as used for the 3D reconstruction, excluding the pre-processing. This produced a high quality registration between the atlas and the reference volume. The atlas was then nonlinearly registered to the reconstructed histological volume, to ensure maximum correspondence.

Finally, we used the inverse transformation from PoSSum to convert the registered atlas slices from the reference volume back to the original histological slices. A sample overlay between one histological slice and the corresponding atlas shown in Figure 2B. This step allows us to reliably identify the cortical areas within the original histological coronal slices.

Cortical sample extraction

We extracted gray matter neuron counting samples (micrographs) for all areas from the high resolution images using a simple graphical widget that we implemented in Python. The samples were selected according to their position (as far as possible from area boundaries), orientation (as perpendicular as possible to the gray matter surface), absence of artifacts (such as air bubbles, folds or other anomalies) and absence of blood vessels. We also avoided regions of high curvature, although this was not always possible due to the location of some areas. The most rostral and caudal slices were avoided, since they are far from

perpendicular to the gray matter. The counting strips were 1 mm wide and spanned the entire depth of the gray matter, extending into the white matter.

During the sample extraction process we also labelled the cortical layers for each sample. We marked the boundaries of the Pia, layers 2/3, 4, 5, 6 and the boundary to the white matter. The marking of layers was based on standard criteria, such as the presence of a granular layer and the cell sizes.

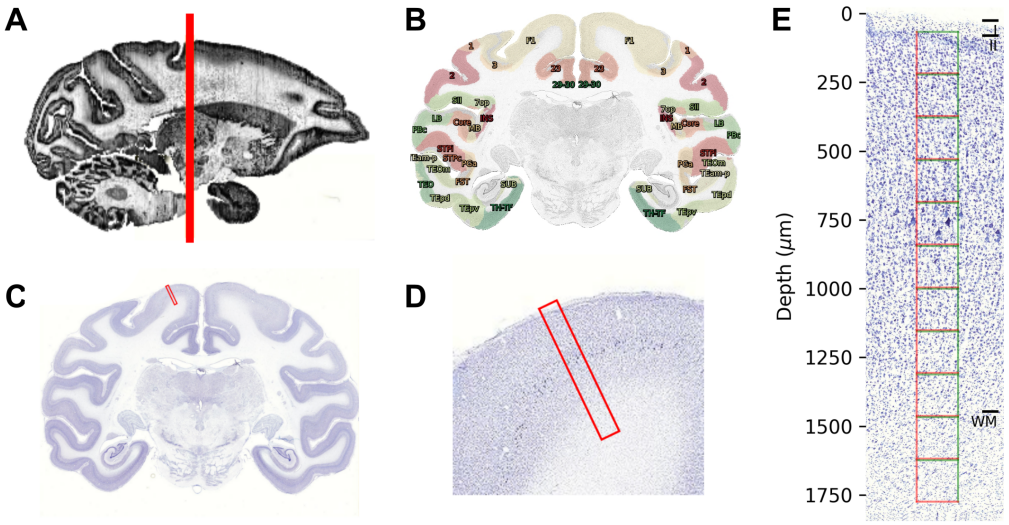


Fig. 2 Sample extraction for cell counting. **A** Lateral slice of the 3D reconstructed volume, red line indicates position of coronal slice for the rest of the figure. **B** Atlas overlayed on a sample slice. **C** Location of sample in the cortical slice. **D** Zoomed in sample. **E** Cortical strip (micrograph) for neuron counting with overlayed counting boxes.

Shrinkage estimation

In order to estimate the tissue shrinkage during the histological staining process we compared the gray matter thickness from our samples to high-quality estimates published elsewhere [26]. We measured the gray matter thickness as the distance between the bottom of layer 1 and the white matter boundary, manually labelled in our counting samples. We estimated the ratio between the gray matter thickness from our study and the thicknesses from Wagstyl et al. [26] for all cortical areas, and took the median value across areas as our shrinkage factor.

$$\text{Shrinkage factor} = \frac{\text{Gray matter thickness } (\textit{This study})}{\text{Gray matter thickness } (\textit{Wagstyl et al. 2015})} \approx 0.56 \quad (1)$$

Thus, the *in vitro* gray matter thickness was around 56% of the original thickness in the *in vivo* brain.

Neuron density estimation

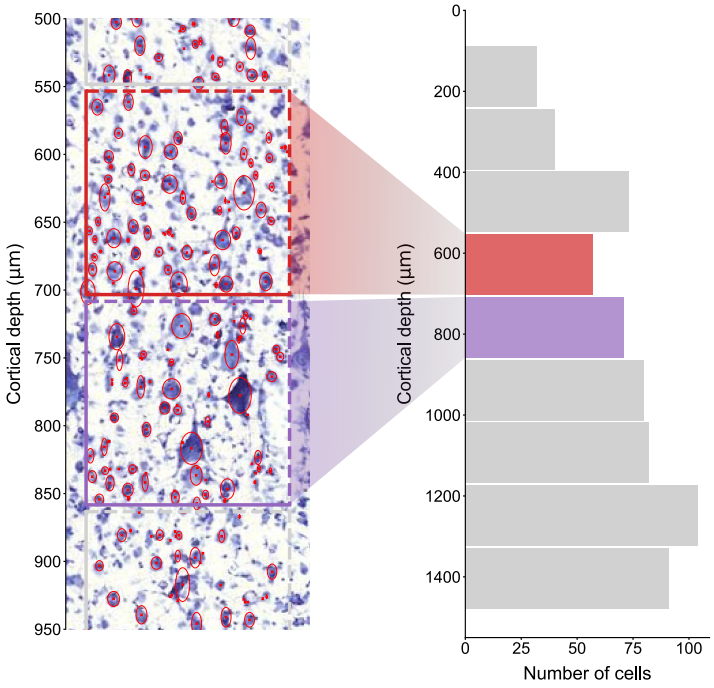


Fig. 3 Example cell counting output for area F1. Left shows zoomed in cortical strip, where each red "x" is the center of a neuron detected using ilastik. Right shows the number of neurons found in each counting box for one cortical strip.

The high resolution cortical sample images were segmented using ilastik [21], a machine learning tool for interactive image classification, segmentation and analysis. Ilastik outputs the segmentation of neurons, alongside their position, shape and other properties. The software learns image segmentation by letting the user label some sample images, which it then applies to all other images. Based on standard criteria [16, 27–29], we only marked cells with a clear nucleus for counting. Smaller cells without a clear nucleus were excluded from the counting procedure, this includes most glia, but potentially also some small neurons.

We estimated the neuron density by using the neuron outlines from the automated segmentation of our images. Square counting frames (150 μm $\times$ 150 μm in-plane, 50 μm thickness and 5 μm spacing between adjacent counting frames) were aligned across the cortical layers. The top and right edges were exclusion lines, while the bottom and left edges were inclusion lines. The uppermost counting frame was aligned to the interface between layers 1 and 2. Some of the frames extended well beyond the gray matter, into the white matter, where some cells could still be found. If a fraction of the counting frame was within the white matter, then the number of neurons was multiplied by the gray matter fraction (specific to each sample), to avoid counting

white matter neurons. Counting frames that were more than 50% within the white matter were excluded from the neuron density estimation.

Assuming that shrinkage was isotropic and uniform across all slices, and given the previously estimated shrinkage factor of 0.56, we estimate that the volume of one counting box *in vivo* was

$$\mathrm{V}_{\mathrm{box}} = 50 \cdot (150/0.56)^2 \ \mu\mathrm{m}^3 \approx 3.6 \cdot 10^6 \ \mu\mathrm{m}^3,$$

by applying the shrinkage factor only to the in-plane direction.

Stereological correction

Counting neurons in tissue slices tends to overestimate the neuron numbers due to the presence of neuron fragments within the counting volume [27, 29]. This is due to some neurons being cut during the slicing procedure and therefore present across multiple adjacent slices. Our study did not include adjacent slices, they were much further apart (250 μm) than the largest neuron bodies in the cortex. In spite of the absence of any overlap, our counting volumes would still overestimate neuron densities, since cell fractions could be counted as full neurons. Previously, relatively thin (10-20 μm) subsequent slices were used to correct for this bias [30–32], by some relation between the neuron diameter and slice thickness. Our slices were relatively thick (50 μm), but still not immune from oversampling.

In order to correct for oversampling due to sliced neurons we applied the following correction to our estimates

$$\rho = \frac{h}{h+d} \cdot \tilde{\rho} \tag{2}$$

where ρ is the corrected neuron density within one counting box, $\tilde{\rho}$ is the uncorrected neuron density, $h = 50$ μm is the slice thickness, and d is the median neuron diameter within a counting box. Our approximation assumes that the cells within one counting box have roughly the same diameter and that cell bodies are spherical. Note that for $h \gg d$, $h/(h+d) \to 1$ and thus $\rho \approx \tilde{\rho}$. Only in the cases where d is relatively large does the correction have a significant effect, e.g. in our case a hypothetical large neuron with $d = 30$ μm would lead to $\rho = 0.625 \cdot \tilde{\rho}$.

White matter distance

We measured the distance between all cortical areas via the white matter. This measure aims to mimic the paths that axons would take to establish connections between distant cortical areas within one hemisphere. Finding the optimal path around a certain set of obstacles is commonly known as motion planning or path planning [33]. This problem is known to be solvable in polynomial time for 2-dimensional landscapes, but NP-complete in 3 or more

dimensions [34]. We devised a simple heuristic to calculate the shortest paths between the centers of mass of each pair of cortical areas, see Algorithm 1.

Algorithm 1 Pseudocode of white matter distance estimation heuristic.

```
for each area pair (source, target) do
    C ← white matter, source and target
    G ← nearest_neighbour_graph(C)
    source_center ← center_of_mass(source)
    target_center ← center_of_mass(target)
    valid_path ← False
    while valid_path == False do
        path ← dijkstra_shortest_path(G, source_center, target_center)
        if all edges ∈ C and all nodes ∈ white matter then
            valid_path ← True
        else
            G ← remove_invalid_edges(G)
        end if
    end whilereturn path
end for
```

Our heuristic starts by defining the three-dimensional volume $\mathcal{C}$, through which the shortest path has to be found. $\mathcal{C}$ includes the volume corresponding to the the white matter as well as the volumes of the two areas for which the distance is calculated. Once $\mathcal{C}$ is determined we build a weighted undirected graph G with N nodes representing each voxel of the volume. We connect each node in G to its m nearest neighbours, yielding a total of $mN/2$ edges, which are weighted by their euclidean length. We connect our graph beyond a simple mesh ($m > 4$) in order to avoid coarse paths that would overestimate real distance. However, the run-time and memory usage increase significantly as $m \to N$, thus we keep $m \ll N$. We settled for $m = 15$ for obtaining the results in this work.

The distances were measured from the centers of mass of each area. If the center of mass of a given area lay outside the volume of the area, we took the closest point in the volume to the center of mass. The shortest path between the centers of mass of the two areas was determined using Dijkstra's algorithm on the previously constructed graph G. A path was considered valid if no edge belonging to that path: 1) traveled outside of $\mathcal{C}$ or 2) connected directly from the source to the target cortical area (without passing through the white matter). If either of this criteria was met, an edge would be considered invalid, since the path would violate the definition of the WM distance. All invalid edges in the shortest path were removed from G and a new shortest path was calculated. This process was then repeated until a shortest path was found where all edges were valid.

Once the shortest path consisting exclusively of valid edges was found, it was smoothed using spline interpolation. The smoothing step improves the final WM estimate, since it reduces distance overestimation due to the discretization, especially for longer paths. The length of the smooth spline was the final WM distance.

Our heuristic is certainly not the most efficient path planning algorithm. Nevertheless, it is well suited to measure the distances given our constraints. The algorithm was executed on a high performance computing (HPC) system and parallelized to estimate the distance between every pair of areas (91 areas $\rightarrow$ 4095 area pairs).

Results

We measured the neuron density across cortical areas and layers of the macaque cortex. First, we reconstructed a 3D volume from the histological slices (see methods Volume reconstruction), in order to perform accurate atlas registration (see methods Atlas registration). Then, we extracted four counting strips per area (two from each hemisphere), based on the atlas registration but also on common landmarks and cytoarchitectural patterns (see methods Cortical sample extraction. Finally, we estimated the number of neurons within fixed counting frames (150 μm $\times$ 150 μm in-plane and 50 μm thickness) using a machine learning tool trained on manually segmented sample data (see methods Neuron density estimation). The measurements were corrected for tissue shrinkage (see methods Shrinkage estimation) and stereological errors due to sliced neurons (see methods Stereological correction).

Shrinkage

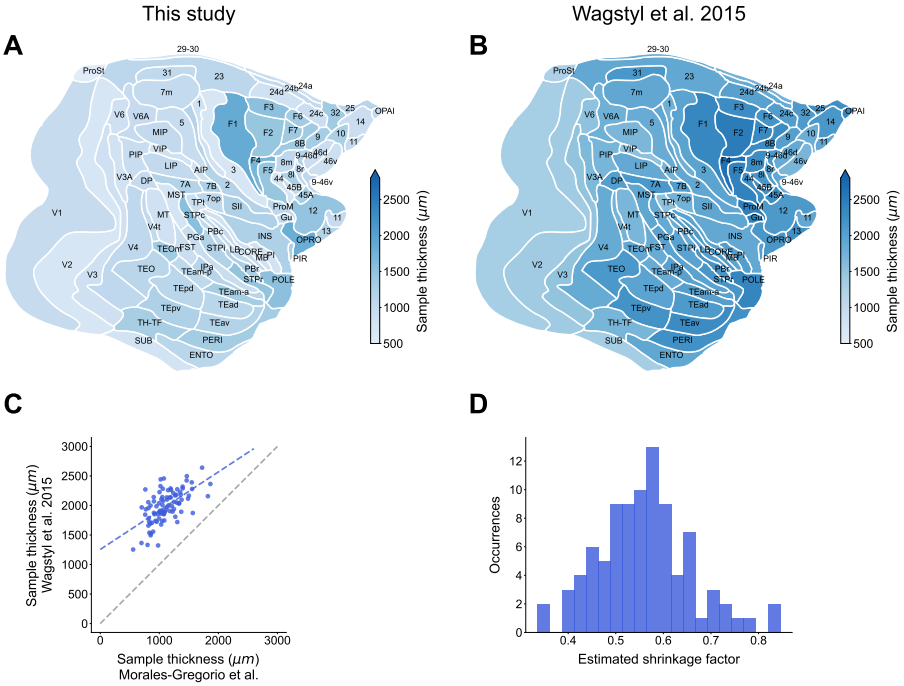


Fig. 4 Sample thickness from our study compared to previous estimates [26]. **A** Median gray matter thickness per area estimated from the counting strips in our study, without shrinkage correction. Flattened cortical surface was based on the M132 atlas from [3]. **B** Median gray matter thickness per area estimated from *in vivo* imaging in a previous study [26]. **C** Comparison of our (This study) estimated thickness with the previous study (Wagstyl et al.). **D** Distribution of estimated shrinkage factor.

The brain tissue inevitably shrinks during the histological processing. To estimate the shrinkage levels and correct for this volumetric bias, we compare our estimates with previous measurements [26] (see methods Shrinkage estimation). We compared the gray matter thickness from our cortical strips to previous measurements (Figure 4A-C) and found that the gray matter thickness of our strips is systematically smaller, which was expected due to the tissue shrinkage. We estimate the tissue shrinkage to be around 0.55 (median of the distribution from Equation 1, see Figure 4D).

Neuron densities

Using an automatic image segmentation (see methods Neuron density estimation), we estimated the number of neurons within each counting frame. We then corrected for shrinkage and stereological errors (see methods Shrinkage estimation and Stereological correction). Finally, we averaged the neuron densities within each counting strip and across counting strips to determine the mean neuron density per area.

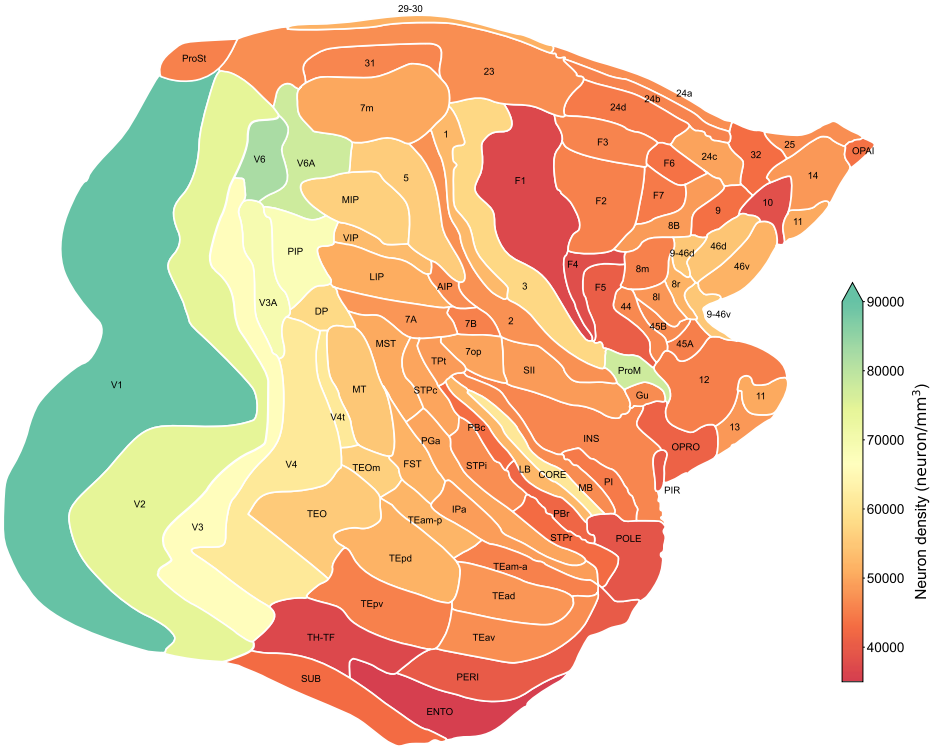


Fig. 5 Preliminary estimates of cell densities for all areas of the macaque neocortex. Flattened cortical surface was re-drawn based on the M132 atlas from [3].

We found that the largest neuron density is in the visual cortex and that the lowest neuron densities are found in the temporal and motor cortex, in agreement with previous knowledge about the neuron densities in the macaque [1].

Validation of neuron densities

In order to validate our neuron density estimates we compare them to previously published data [1] and assess whether they follow a lognormal distribution, a known feature of neuron densities across and within cortical areas [7].

First, we compare our estimates with previous ones [1], Figure 6A. It appears that our method is slightly biased with respect to the previously reported values, since it overestimates neuron densities in low-density areas and underestimates it in high-density areas. In spite of this bias, the overall trend is conserved. Second, we assess whether the neuron density distribution is compatible with a lognormal distribution. We first test the distribution of the mean density across areas (Shapiro-Wilk test for normality on $\ln(\rho)$, Figure 6B,C) and we then test the distribution within each cortical area (Figure 6D). The test results suggest that the estimated neuron density distribution is not compatible with a lognormal distribution across cortical areas

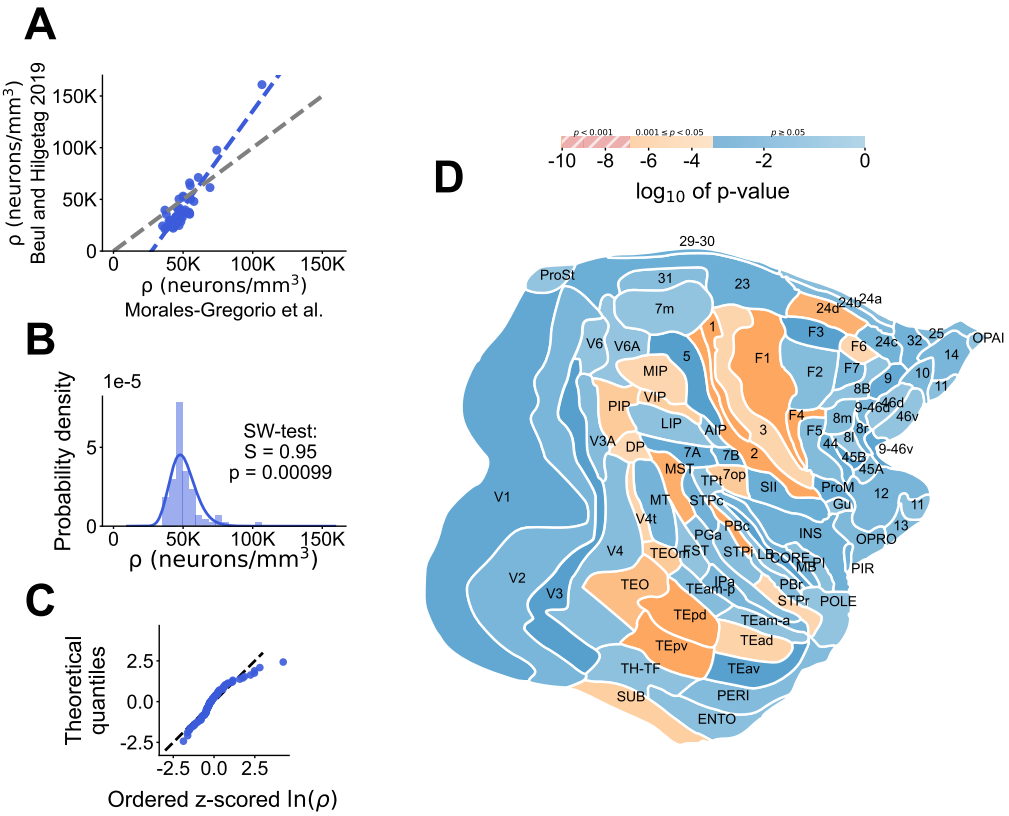


Fig. 6 Validation of neuron density estimates in the macaque cortex. **A** Comparison of our estimated neuron densities per area with a previous study [1]. **B** The distribution of neuron densities does not appear to be lognormal, we test for normality of $\ln(\rho)$ with a Shapiro-Wilk test. **C** Probability plot of z-scored $\ln(\rho)$. **D** Log_{10} of p-value of Shapiro-Wilk normality test of $\ln(\rho_s)$ on a flattened representation of the macaque cortex.

($p \leq 0.05$), in contrast to previous reports of neuron density distribution in the mammalian cortex [7]. Within the cortical areas the lognormal distribution seems to be prevalent, in agreement with the distributions of within-area neuron densities in the marmoset that have been shown to be compatible with lognormal distributions [7]. The validation suggests that our method introduces certain bias with respect to the previous reports. Further improvements to the neuron counting method will be needed, possibly using an additional step to remove the bias.

White matter distance

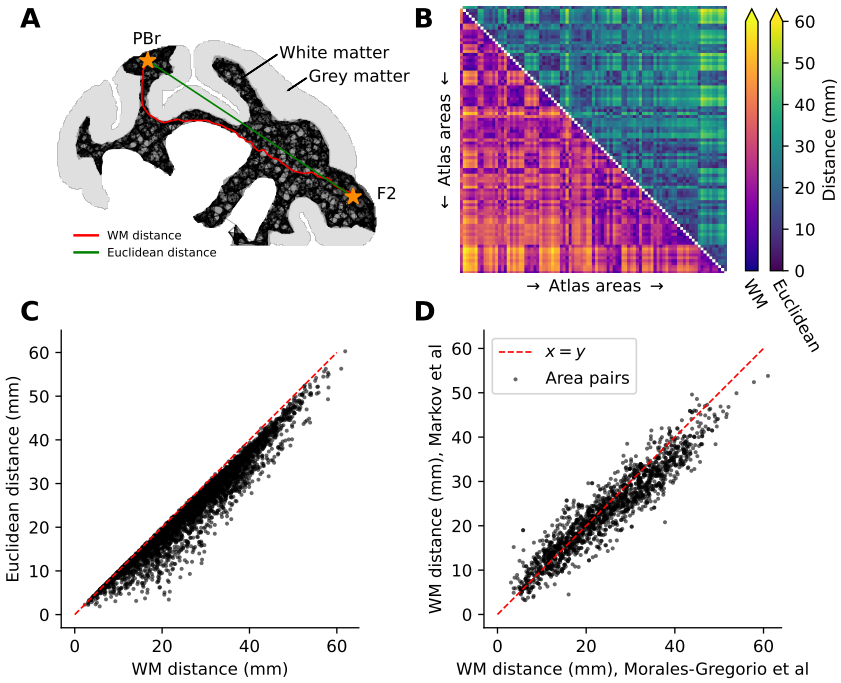


Fig. 7 White matter distance measurement and validation. **A** Schematic representation of the WM distance on a 2D slice of the atlas (the real distance is measured in 3D). Euclidean distance also shown. **B** WM (lower triangular) and euclidean (upper triangular) measured distances. **C** Comparison of WM and euclidean distance, showing that WM $\geq$ euclidean distance. **D** Comparison of WM distance measured in this study and previous measurements [3].

The white matter (WM) distance between all area pairs was measured using a custom heuristic (see methods White matter distance). The WM distance was measured using the M132 parcellation [3], which was registered to the same reference as the reconstructed volume [25] (see methods Atlas registration). The main reason for using the Calabrese et al. reference volume was that the white matter had already been labelled and that we also used the same reference to reconstruct the volume from histological slices.

Briefly, we measured the distance between the centers of mass of all areas to each other through the white matter. This measure approximates the length a given axon would need to have in order to establish long-range synapses across the cortex. Our method creates a mesh including the white matter and the involved cortical areas and searches for the shortest path that does not take any *shortcuts* outside the tissue. (Figure 7A). We present the measured distances as well as the corresponding euclidean distances between all area pairs (Figure 7B). The white matter distance is by definition longer or equal to the euclidean distance, which we observe in our measurements to validate the results (Figure 7C). Previous estimates of the white matter distance exist for a subset of areas in the macaque monkey [3]. We compared our measurements

to the previously reported values (Figure 7D) and show a clear correspondence trend between the two datasets. The observed differences between the two datasets might be explained by the fact that we used a different reference space for our measurements and that the method from Markov et al. for measuring the distances was very different from ours. In their study, Markov et al. reported that the Map3D software was used to define a trajectory between centers of gravity in a 3D reconstruction of the macaque cortex.

Discussion

In this study we have measured layer- and area-resolved neuron densities across the macaque cortex, large parts of which were previously unknown. In order to achieve our goal, we created an anisotropic 3D volume of the macaque brain from Nissl-stained coronal slices [22], a workflow with potential uses beyond this project. Additionally, we also estimated the tissue shrinkage from the gray matter thickness, in order to provide reliable density data. The estimated neuron densities were validated against existing data [1] and known neuron distributions [7], which revealed the need for further refinement of the methods. Finally, we also estimated the white matter distance between all area pairs, a relevant structural parameter to characterize the macaque brain.

In order to accurately map the parcellation to the histological slices we constructed a 3D volume from the coronal slices. This volume is valuable on its own as a research object, since it enables further studies in brain structure and organization beyond the cortical neuron densities. Long-standing projects aim to reconstruct an isotropic volume from histology for a human brain, known as BigBrain [35–37]. The BigBrain is a valuable resource for the exploration and study of the human brain, and relies on years of careful curation and management to construct. The volume we created in our study is a first step towards a BigBrain for the macaque. The resolution of the volume is limited in the anterior-posterior axis (250 μm), but has a high in-plane resolution (1 μm) that enables the study of cell-body distributions across the whole brain. Our approach used lower-resolution images for the reconstruction, since we only needed the brain volume for atlas registration. Future work could use the transformations from the low-resolution reconstruction and apply them to the high-resolution data for further investigation of the macaque brain.

Neuron density is a crucial variable for characterizing the neural tissue, since neurons are the fundamental units of the nervous system. Reliable and accurate neuron densities can be used to construct realistic spiking neuron models [38–40] and uncover the underlying structural organization of the brain. Our measurements correctly captured several known features of neuron densities: primary visual cortex has the highest neuron density in the primate cortex; the motor and temporal cortices have some of the lowest neuron densities; and neuron density changes following several gradients e.g. the visual hierarchy. However, our measurements did not accurately match previous quantitative values [1], and failed to display a lognormal distribution across cortical areas

[7]. We attribute these shortcomings to the inherent noisy nature of the cell segmentation process. The task of identifying neurons is further complicated by the Nissl-staining, since it indiscriminately labels both neurons and glia. Thus, further improvements to our classifier will be required and possibly a cross-validation using manually estimated densities. Additionally, our analysis was limited to a single macaque, and thus does not capture variability across subjects. To address this issue, we are already working on processing further histological data from other openly available histological data of macaque brains [41, 42].

Another parameter needed for multi-area cortical models is the white matter distance between all area pairs, which we therefore also measured. Our measurements were in agreement with previously reported values in the same parcellation [3]. Our approach uses a simple heuristic that is reproducible and is not biased by human error. The heuristic we used was not very computationally efficient since finding geodesic distances is a well-known hard problem [33], but our heuristic is designed to avoid overestimation of distance due to small sample steps. Besides the construction of models, the white matter distance has been shown to be a strong predictor of connectivity strength [43] and is relevant—together with the axon diameter—for estimating transmission delay [44].

In summary, we have constructed a 3D brain volume from histological data, estimated layer- and area-resolved neuron density and white matter distance across the macaque cortex. We for the first time report neuron densities for all cortical areas of a single macaque, including areas for which no previous quantitative data was available. Additionally, we also for the first time report the white matter distance between all cortical area pairs, based on a reproducible heuristic. Our measurements provide a comprehensive quantitative description of the macaque cortex with unprecedented detail. We used the same parcellation as in previous tract-tracing studies, such that our measurements are fully interoperable with the valuable structural connectivity data [3]. The methods developed in this study can be applied to other species' brains, in order to measure structural quantitative data and increase our understanding on brain structure across different evolutionary paths.

Funding

This project received funding from the DFG Priority Program (SPP 2041 "Computational Connectomics") [S.J.van Albada: AL 2041/1-1]; the EU's Horizon 2020 Framework Grant Agreement No. 785907 (HBP-SGA2); the DFG (RTG 2416 "MultiSenses-MultiScales");

References

[1] Beul, S.F., Hilgetag, C.C.: Neuron density fundamentally relates to architecture and connectivity of the primate cerebral cortex. NeuroImage **189**, 777–792 (2019). https://doi.org/10.1016/j.neuroimage.2019.01.010

[2] Bakker, R., Wachtler, T., Diesmann, M.: CoCoMac 2.0 and the future of tract-tracing databases. Frontiers in Neuroinformatics **6** (2012). https://doi.org/10.3389/fninf.2012.00030

[3] Markov, N.T., Ercsey-Ravasz, M., Ribeiro Gomes, A.R., Lamy, C., Magrou, L., Vezoli, J., Misery, P., Falchier, A., Quilodran, R., Gariel, M.A., Sallet, J., Gamanut, R., Huissoud, C., Clavagnier, S., Giroud, P., Sappey-Marinier, D., Barone, P., Dehay, C., Toroczkai, Z., Knoblauch, K., Van Essen, D.C., Kennedy, H.: A weighted and directed interareal connectivity matrix for macaque cerebral cortex. Cerebral Cortex (2012). https://doi.org/10.1093/cercor/bhs270

[4] Froudist-Walsh, S., Xu, T., Niu, M., Rapan, L., Margulies, D.S., Zilles, K., Wang, X.-J., Palomero-Gallagher, N.: Gradients of receptor expression in the macaque cortex. bioRxiv (2021). https://doi.org/10.1101/2021.02.22.432173

[5] Kötter, R., Wanke, E.: Mapping brains without coordinates. Philosophical Transactions of the Royal Society B: Biological Sciences **360**(1456), 751–766 (2005). https://doi.org/10.1098/rstb.2005.1625

[6] van Albada, S.J., Morales-Gregorio, A., Dickscheid, T., Goulas, A., Bakker, R., Bludau, S., Palm, G., Hilgetag, C.-C., Diesmann, M.: Bringing Anatomical Information into Neuronal Network Models. In: Giugliano, M., Negrello, M., Linaro, D. (eds.) Computational Modelling of the Brain vol. 1359, pp. 201–234. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-89439-9_9. Series Title: Advances in Experimental Medicine and Biology

[7] Morales-Gregorio, A., van Meegen, A., van Albada, S.J.: Ubiquitous lognormal distribution of neuron densities across mammalian cerebral cortex. bioRxiv (2022). https://doi.org/10.1101/2022.03.17.480842

[8] Herculano-Houzel, S., Watson, C., Paxinos, G.: Distribution of neurons in functional areas of the mouse cerebral cortex reveals quantitatively different cortical zones. Frontiers in Neuroanatomy **7** (2013). https://doi.org/10.3389/fnana.2013.00035

[9] Erö, C., Gewaltig, M.-O., Keller, D., Markram, H.: A Cell Atlas for the Mouse Brain. Frontiers in Neuroinformatics **12**, 84 (2018). https://doi.org/10.3389/fninf.2018.00084

[10] Charvet, C.J., Cahalane, D.J., Finlay, B.L.: Systematic, Cross-Cortex Variation in Neuron Numbers in Rodents and Primates. Cerebral Cortex **25**(1), 147–160 (2015). https://doi.org/10.1093/cercor/bht214

[11] Atapour, N., Majka, P., Wolkowicz, I.H., Malamanova, D., Worthy, K.H., Rosa, M.G.P.: Neuronal Distribution Across the Cerebral Cortex of the Marmoset Monkey (Callithrix jacchus). Cerebral cortex **29**(9), 3836–3863 (2019). https://doi.org/10.1093/cercor/bhy263

[12] Collins, C.E., Airey, D.C., Young, N.A., Leitch, D.B., Kaas, J.H.: Neuron densities vary across and within cortical areas in primates. Proceedings of the National Academy of Sciences **107**(36), 15927–15932 (2010). https://doi.org/10.1073/pnas.1010356107

[13] Collins, C.E., Turner, E.C., Sawyer, E.K., Reed, J.L., Young, N.A., Flaherty, D.K., Kaas, J.H.: Cortical cell and neuron density estimates in one chimpanzee hemisphere. Proceedings of the National Academy of Sciences **113**(3), 740–745 (2016). https://doi.org/10.1073/pnas.1524208113

[14] Turner, E.C., Young, N.A., Reed, J.L., Collins, C.E., Flaherty, D.K., Gabi, M., Kaas, J.H.: Distributions of Cells and Neurons across the Cortical Sheet in Old World Macaques. Brain, Behavior and Evolution **88**(1), 1–13 (2016). https://doi.org/10.1159/000446762

[15] Jardim-Messeder, D., Lambert, K., Noctor, S., Pestana, F.M., de Castro Leal, M.E., Bertelsen, M.F., Alagaili, A.N., Mohammad, O.B., Manger, P.R., Herculano-Houzel, S.: Dogs Have the Most Neurons, Though Not the Largest Brain: Trade-Off between Body Mass and Number of Neurons in the Cerebral Cortex of Large Carnivoran Species. Frontiers in Neuroanatomy **11**, 118 (2017). https://doi.org/10.3389/fnana.2017.00118

[16] von Bartheld, C.S., Bahney, J., Herculano-Houzel, S.: The search for true numbers of neurons and glial cells in the human brain: A review of 150 years of cell counting: Quantifying neurons and glia in human brain. Journal of Comparative Neurology **524**(18), 3865–3895 (2016).

https://doi.org/10.1002/cne.24040

[17] von Economo, C., Koskinas, G.N., Triarhou, L.C.: Atlas of Cytoarchitectonics of the Adult Human Cerebral Cortex. Karger, Basel, Switzerland (2008)

[18] Bakker, R., Tiesinga, P., Kötter, R.: The Scalable Brain Atlas: Instant Web-Based Access to Public Brain Atlases and Related Content. Neuroinformatics **13**(3), 353–366 (2015). https://doi.org/10.1007/s12021-014-9258-x

[19] Microdraw, https://microdraw.pasteur.fr/. Online (2021)

[20] Majka, P., Chaplin, T.A., Yu, H.-H., Tolpygo, A., Mitra, P.P., Wójcik, D.K., Rosa, M.G.P.: Towards a comprehensive atlas of cortical connections in a primate brain: Mapping tracer injection studies of the common marmoset into a reference digital template. Journal of Comparative Neurology **524**(11), 2161–2181 (2016). https://doi.org/10.1002/cne.24023

[21] Berg, S., Kutra, D., Kroeger, T., Straehle, C.N., Kausler, B.X., Haubold, C., Schiegg, M., Ales, J., Beier, T., Rudy, M., Eren, K., Cervantes, J.I., Xu, B., Beuttenmueller, F., Wolny, A., Zhang, C., Koethe, U., Hamprecht, F.A., Kreshuk, A.: ilastik: interactive machine learning for (bio)image analysis. Nature Methods (2019). https://doi.org/10.1038/s41592-019-0582-9

[22] © 2009 Allen Institute for Brain Science: NIH Blueprint Non-Human Primate (NHP) Atlas. Available from: www.blueprintnhpatlas.org

[23] © 2009 Allen Institute for Brain Science: Technical white paper: MRI and histological reference series of postnatal brain development in rhesus macaque (2014)

[24] Bradski, G.: The OpenCV Library. Dr. Dobb's Journal of Software Tools (2000)

[25] Calabrese, E., Badea, A., Coe, C.L., Lubach, G.R., Shi, Y., Styner, M.A., Johnson, G.A.: A diffusion tensor MRI atlas of the postmortem rhesus macaque brain. NeuroImage **117**, 408–416 (2015). https://doi.org/10.1016/j.neuroimage.2015.05.072

[26] Wagstyl, K., Ronan, L., Goodyer, I.M., Fletcher, P.C.: Cortical thickness gradients in structural hierarchies. NeuroImage **111**, 241–250 (2015). https://doi.org/10.1016/j.neuroimage.2015.02.036

[27] Schmitz, C., Hof, P.R.: Design-based stereology in neuroscience. Neuroscience **130**(4), 813–831 (2005). https://doi.org/10.1016/j.neuroscience.2004.08.050

[28] Guillery, R.W., August, B.K.: Doubt and certainty in counting. Progress in Brain Research **135**, 25–42 (2002). https://doi.org/10.1016/S0079-6123(02)35005-2

[29] West, M.J.: Stereological methods for estimating the total number of neurons and synapses: issues of precision and bias. Trends in Neurosciences **22**(2), 51–61 (1999). https://doi.org/10.1016/S0166-2236(98)01362-9

[30] Abercrombie, M.: Estimation of nuclear population from microtome sections. The Anatomical Record **94**(2), 239–247 (1946). https://doi.org/10.1002/ar.1090940210

[31] Ebbesson, S.O.E., Tang, D.: A method for estimating the number of cells in histological sections. Journal of the Royal Microscopical Society **84**(4), 449–464 (1965). https://doi.org/10.1111/j.1365-2818.1965.tb02146.x

[32] Clarke, R.: A comparative analysis of methods of estimating the size of cell populations from microtome sections. Journal of the Royal Microscopical Society **88**(2), 189–203 (1968). https://doi.org/10.1111/j.1365-2818.1968.tb00607.x

[33] Latombe, J.-C.: Robot Motion Planning. Springer, New York, NY (2012)

[34] Canny, J., Reif, J.: New lower bound techniques for robot motion planning problems. In: 28th Annual Symposium on Foundations of Computer Science (sfcs 1987), pp. 49–60. IEEE, Los Angeles, CA, USA (1987). https://doi.org/10.1109/SFCS.1987.42

[35] Amunts, K., Lepage, C., Borgeat, L., Mohlberg, H., Dickscheid, T., Rousseau, M.-, Bludau, S., Bazin, P.-L., Lewis, L.B., Oros-Peusquens, A.-M., Shah, N.J., Lippert, T., Zilles, K., Evans, A.C.: BigBrain: An Ultrahigh-Resolution 3D Human Brain Model. Science **340**(6139), 1472–1475 (2013). https://doi.org/10.1126/science.1235381

[36] Wagstyl, K., Lepage, C., Bludau, S., Zilles, K., Fletcher, P.C., Amunts, K., Evans, A.C.: Mapping Cortical Laminar Structure in the 3D BigBrain. Cerebral Cortex **28**(7), 2551–2562 (2018). https://doi.org/10.1093/cercor/bhy074

[37] Wagstyl, K., Larocque, S., Cucurull, G., Lepage, C., Cohen, J.P., Bludau, S., Palomero-Gallagher, N., Lewis, L.B., Funck, T., Spitzer, H., Dickscheid, T., Fletcher, P.C., Romero, A., Zilles, K., Amunts, K., Bengio, Y., Evans, A.C.: BigBrain 3D atlas of cortical layers: Cortical and laminar thickness gradients diverge in sensory and motor cortices. PLOS Biology **18**(4), 3000678 (2020). https://doi.org/10.1371/journal.pbio.3000678

[38] Potjans, T.C., Diesmann, M.: The Cell-Type Specific Cortical Microcircuit: Relating Structure and Activity in a Full-Scale Spiking Network Model. Cerebral Cortex **24**(3), 785–806 (2014). https://doi.org/10.1093/cercor/bhs358

[39] Schmidt, M., Bakker, R., Shen, K., Bezgin, G., Diesmann, M., van Albada, S.J.: A multi-scale layer-resolved spiking network model of resting-state dynamics in macaque visual cortical areas. PLOS Computational Biology **14**(10), 1006359 (2018). https://doi.org/10.1371/journal.pcbi.1006359

[40] Schmidt, M., Bakker, R., Hilgetag, C.C., Diesmann, M., van Albada, S.J.: Multi-scale account of the network structure of macaque visual cortex. Brain Structure and Function **223**(3), 1409–1435 (2018). https://doi.org/10.1007/s00429-017-1554-4

[41] Mikula, S., Trotts, I., Stone, J.M., Jones, E.G.: Internet-enabled high-resolution brain mapping and virtual microscopy. NeuroImage **35**(1), 9–15 (2007). https://doi.org/10.1016/j.neuroimage.2006.11.053

[42] Jones, E.G., Stone, J.M., Karten, H.J.: High-resolution digital brain atlases: a Hubble telescope for the brain: A Hubble telescope for the brain. Annals of the New York Academy of Sciences **1225**(S1), 147–159 (2011). https://doi.org/10.1111/j.1749-6632.2011.06009.x

[43] Ercsey-Ravasz, M., Markov, N.T., Lamy, C., Van Essen, D.C., Knoblauch, K., Toroczkai, Z., Kennedy, H.: A Predictive Network Model of Cerebral Cortical Connectivity Based on a Distance Rule. Neuron **80**(1), 184–197 (2013). https://doi.org/10.1016/j.neuron.2013.07.036

[44] More, H.L., Hutchinson, J.R., Collins, D.F., Weber, D.J., Aung, S.K.H., Donelan, J.M.: Scaling of sensorimotor control in terrestrial mammals. Proceedings of the Royal Society B: Biological Sciences **277**(1700), 3563–3568 (2010). https://doi.org/10.1098/rspb.2010.0898

Similar and functional network behavior can result from widely differing combinations of intrinsic and synaptic properties.

Astrid Prinz, Dirk Bucher and Eve Marder
(Prinz et al., 2004)

5

Activity-driven estimation of local connectivity

Summary:

In this chapter we present a method to estimate the connectivity of a cortical microcircuit based on resting state spiking activity. We first introduce a method to compare the spiking activity from different sources using a multi-dimensional cloud of summary statistics. We show that the multi-dimensional summary statistics are unique to a certain cortical area and can thus be used to estimate similarity between simulated and experimental data. Finally, we provide a proof of concept of our parameter estimating methods by applying our optimization algorithm to synthetic data, which correctly estimated the connectivity parameters of the model. Further work will apply these methods to larger more biologically realistic cortical microcircuit models, thus providing estimates for the connectivity within cortical areas.

Activity-driven microconnectome estimation of macaque visuomotor cortices

Aitor Morales-Gregorio[1,2*], Robin Gutzen[1,3], Paulina A. Dąbrowska[1,3], Alper Yegenoglu[4,5], Sandra Diaz-Pier[4], Sarah Palmis[6], Sofia Paneri[7,8], Alexandre René[1,9], Panagiotas Sapountzis[7], Markus Diesmann[1,10,11], Sonja Grün[1,12], Johanna Senk[1], Georgia G. Gregoriou[7,8], Bjørg E. Kilavik[6] and Sacha J. van Albada[1,2]

[1]Institute of Neuroscience and Medicine (INM-6) and Institute for Advanced Simulation (IAS-6) and JARA-Institut Brain Structure-Function Relationships (INM-10), Jülich Research Centre, Jülich, Germany.
[2]Institute of Zoology, University of Cologne, Cologne, Germany.
[3]RWTH Aachen University, Aachen, Germany.
[4]SimLab Neuroscience, Jülich Supercomputing Centre (JSC), Institute for Advanced Simulation, JARA, Jülich Research Centre, Jülich, Germany.
[5]Institute of Geometry and Applied Mathematics, Department of Mathematics, RWTH Aachen University, Aachen, Germany.
[6]Institut de Neurosciences de la Timone (INT), CNRS & Aix-Marseille Université, Marseille, France.
[7]Institute of Applied and Computational Mathematics, Foundation for Research and Technology Hellas, Heraklion, Greece.
[8]Department of Basic Sciences, Medical School, University of Crete, Heraklion, Greece.
[9]Department of Physics, University of Ottawa, Ottawa, Canada.
[10]Department of Psychiatry, Psychotherapy and Psychosomatics, School of Medicine, RWTH Aachen University, Aachen, Germany.
[11]Department of Physics, Faculty 1, RWTH Aachen University, Aachen, Germany.
[12]Theoretical Systems Neurobiology, RWTH Aachen University, Aachen, Germany.

*Corresponding author. a.morales-gregorio@fz-juelich.de

Keywords: Spiking data, macaque, cortex, microcircuit, connectivity, optimization, spiking neuron networks

Abstract

The connectivity structure is essential to understand the activity and function of neuronal networks in the central nervous system. Long-range connectivity in the macaque cortex has been thoroughly studied. However, little is known about the microscale connectivity within cortical areas and across cortical layers—"the microconnectome"—outside of the early sensory and primary motor areas. Here, we introduce a novel method to estimate the microconnectome from neuronal spiking activity. We measured single-neuron statistics from the resting-state activity of several cortical areas (V1, V4, M1, PMd, dlPFC) in macaque monkeys, describing their firing rates, irregularity, and correlations. Our measurements show significant differences in the spiking activity of each cortical area, which we hypothesize arise from structural differences in the microconnectome within these areas. Our method to estimate the microconnectome uses an optimization algorithm that maximizes the similarity between the experimental spiking data and biologically realistic models. We demonstrate our method on synthetic data generated by a small balanced spiking neuron network model. Further work will extend the method to microcircuit models with the goal of estimating the microconnectome across cortical areas in the macaque.

Introduction

Connectivity is the backbone of any neuronal network. It determines how the information is passed between neurons and processed to eventually generate complex cognitive functions. While much is known about the cortico-cortical long-range connectivity in the macaque from tract-tracing studies [1–3], the connectivity within the cortical areas remains understudied. Current estimates of the within-area connectivity rely on a combination of measurements from different animals and areas [4, 5]. While the cortical areas differ in their activity and functionality, they still rely on a similar basic columnar architecture. The cortical microcircuit is often considered as a canonical building block of the cortex [6]. However, little is known about the microcircuit connectivity (microconnectome) outside of early sensory and motor areas. Estimating such local connectivity often relies on paired recordings [7–10] or glutamate uncaging [11–17], which are both tedious processes and sample from relatively few neurons at a time. Here, we propose an activity-driven method to estimate the microconnectome from *in vivo* data, using spiking neuron models and maximizing the similarity of single-neuron statistics between simulations and experiments.

There are several methods to estimate connectivity from activity data. Correlation-based methods [18–22] can estimate the functional connectivity, which under ideal stationary conditions resembles the structural connectivity [23–25]. However, simulation studies have shown that spurious correlations lead to systematic errors in the inference of structural connectivity from pairwise correlations [26].

Another option is to fit the observed activity to a mechanistic model which captures the relevant aspects of the studied network. Optimal model parameters for the mechanistic models can be determined via a brute force approach [27, 28], evolutionary optimization methods [29–31], or plasticity rules [32]. When the models are analytically tractable and a computationally inexpensive likelihood function can be derived, the parameter distributions can be inferred relatively fast [33–36]. Machine learning approaches can also be used to estimate the likelihood function when no analytical solution is available, thus estimating parameter distributions from simulations [37, 38]. In either case, one does not expect to find unique parameter sets, but rather a range of possible values, because disparate network parameters are known to be able to produce the same dynamics [27].

The neuronal dynamics are significantly different across cortical areas, but surprisingly consistent across animal species for the same area during rest-like periods [39]. These differences are for instance observable in the inter-spike interval (ISI) variability [39, 40] and the autocorrelation decay, i.e., the intrinsic timescale [41–43]. There are two options for how this variability in neuronal dynamics could emerge: 1) differences in the received activity from outside the cortical area, or 2) differences in the internal structure (connectivity and/or neuron properties). Therefore, defining parameters of the external input and the internal connectivity probabilities as free variables in the optimization workflow may yield some explanatory insights into the differences between cortical areas.

Here, we present a method to estimate the parameters for mechanistic models based on single-neuron statistics. First, we gathered extracellular electrophysiological data from across the cerebral cortex of several macaques ($N = 5$, *Macaca mulatta*). We spike-sorted the data and measured the single-neuron statistics for 10-second slices within each area. We show that the multi-dimensional summary statistics are significantly different across areas and layers, when comparing all experiments. We then present a custom optimization algorithm—a combination of random search, gradient descent and genetic algorithms—, which we use to estimate anatomical parameters from the multi-dimensional summary statistics. We demonstrate that the method can correctly estimate the connectivity parameters of a small balanced spiking neuron network, from the multi-dimensional summary statistics of spiking activity alone. Further work will include adapting the methods to work for larger models and estimating the connectivity parameters from the experimental data.

Results

A parameter estimation method for spiking neuron models

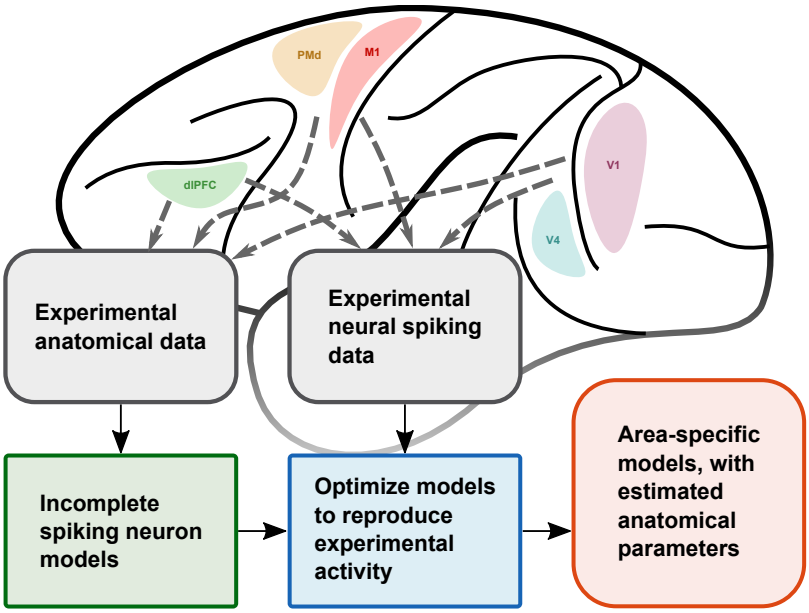


Fig. 1 Graphical abstract. Anatomical and activity data from multiple cortical areas are gathered, with which we construct population-resolved spiking neuron models. We use an optimization algorithm to estimate the connectivity parameters of the models that lead to the maximum similarity between the simulated and experimental neuronal activity.

We here present a method to estimate model parameters from spiking neuron data. We use a simple randomized search optimization algorithm to explore the parameter space (see Methods Workflow overview). Since spike train data are difficult to compare directly, we compare the following single-neuron statistics between the target activity and the simulated activity: firing rate (FR), local variability (LvR), and cross-correlation (CC_{avg}, CC_{std}), which we assemble into a multi-dimensional cloud. We found the multi-dimensional summary statistics to be informative, without being redundant.

Multi-dimensional summary statistics of spiking activity

In order to compare the spiking activity of different cortical areas, we first gathered resting-state spike train data from several sources (see Methods Electrophysiological data collection in resting state). Figure 2A shows an overview of the brain locations and layers of each experimental recording. For our analysis, we sliced the spike train data into 10 s slices, with samples shown in Figure 2B. We display the total length of the data for each experiment in Figure S11. We then calculated several single-neuron statistics (see Methods

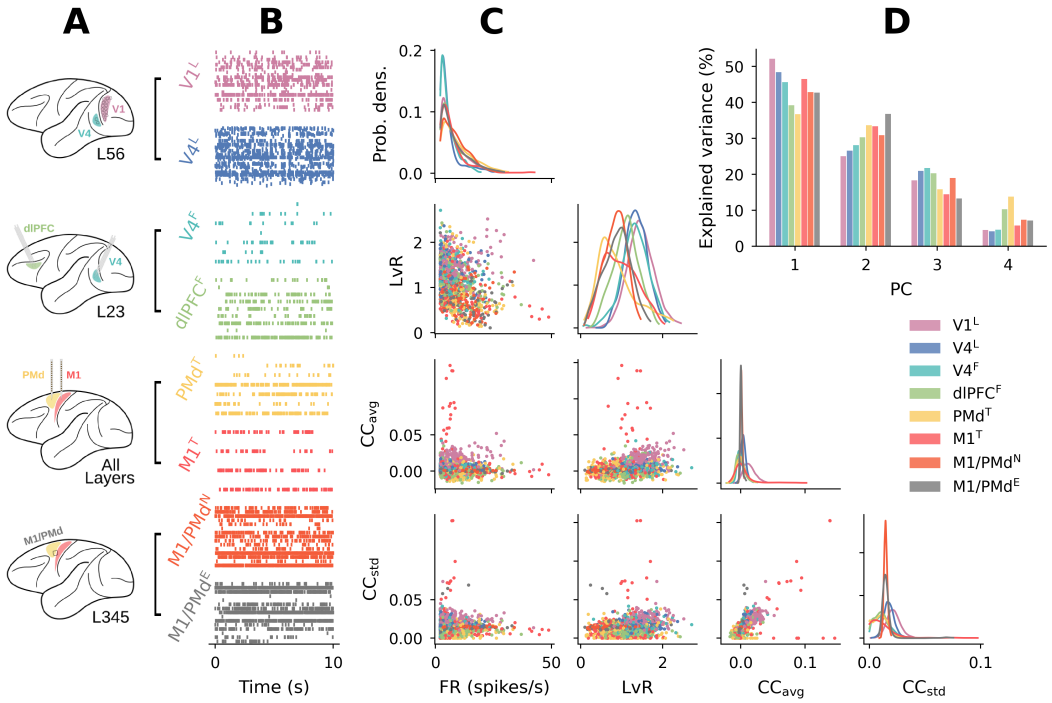


Fig. 2 Sample of available experimental data and their summary statistics. **A)** Schematic representation of the data recording location. **B)** Sample recordings of simultaneous spike trains from each dataset for a 10 s window, superscripts indicate the subject name. **C)** Summary statistics of the single unit spike trains. Each point in the scatter plot corresponds to a 10 s spike train of a single neuron. **D)** Variance explained by the first four principal components (PC) of the multi-dimensional summary statistics.

Summary statistics of spiking neuronal data) for each data slice and assembled them into a multi-dimensional summary statistics point cloud. Figure 2C depicts a pairwise plot (off-diagonal) of the four single-neuron statistics we calculated for each session, and their univariate distribution (diagonal). We also illustrated the multi-dimensional summary statistics for each animal, area, and layer (L23, L345 or L56[1]) in Figure S1-S10. We demonstrate that the single-unit statistics included are all relevant since they each explain at least 5-10% of the variance in the data set (Figure 2D). We also measured the power spectral density from the spike time histograms (Figure S12), which is not included in the summary statistics due to its different dimensionality.

In addition to the multi-dimensional summary statistics from the experimental data we also illustrated the multi-dimensional summary statistics for the two spiking neuron models used in this study, see Methods Spiking neuron models for details. The multi-dimensional summary statistics of the small balanced spiking neuron network appear to span a narrower range (Figure S13), whereas those of the microcircuit model span a broader range that is more similar to the experimental data (Figure S14).

[1]Note that throughout this paper L23 = L2 + L3, L345 = L3 + L4 + L5, and L56 = L5 + L6

Multi-dimensional summary statistics differ across cortical areas

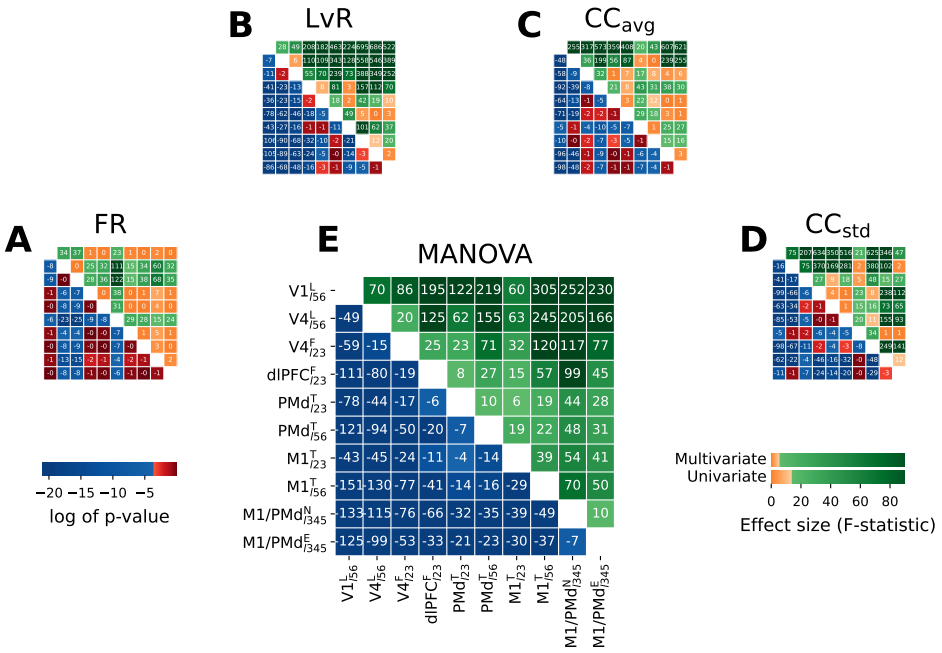


Fig. 3 Uni- and multivariate pairwise test results over the summary statistics. We test the null hypothesis that two or more groups have the same population mean. In all panels, lower triangular entries show the logarithm of the p-values and the upper triangular part shows the F-statistic. **A-D)** show the univariate pairwise analysis of variance (ANOVA) tests. **E)** shows the multivariate pairwise analysis of variance (MANOVA) tests. Significance levels ($\alpha = 0.05$) are corrected for multiple testing following the Bonferroni correction. Note that the area ticks refer to the lower triangular entries, while the upper triangular part is mirrored along the diagonal.

A preliminary look at the multi-dimensional summary statistics (Figure 2C) from the different brain areas and experiments reveals some noticeable differences. To quantify the differences between the multi-dimensional summary statistics we test whether their means are significantly different using two statistical approaches. First, we perform a pairwise univariate analysis of variance (ANOVA) for each of our four statistics (Figure 3A–D). Second, we perform pairwise multivariate ANOVA (MANOVA) tests (Figure 3E), where all the single-neuron statistics are considered in combination. For the tests we used a fixed number of points from the multi-dimensional summary statistics ($N = 350$, randomly sampled), since the tests are robust to the normality assumption if the samples are large (generally $N > 30$) and robust to the assumption of equal (co-)variance (homoscedasticity) if the sample sizes are equal [44]. The significance level $\alpha = 0.05$ was Bonferroni corrected for multiple testing ($k = 47$ experiment pairs $\cdot$ 5 test types $= 235$), such that $\alpha = 0.05/k = 2.13 \cdot 10^{-4}$.

The univariate tests (Figure 3A–D) reveal that no single summary statistic can distinguish the activity from all areas. However, when all the summary statistics are aggregated and tested as a whole with the MANOVA test (Figure 3E) all area pairs become statistically different from each other. The multi-dimensional summary statistics are thus area-specific, which we can exploit to estimate area-specific parameters.

Fitness is robust to time slice length

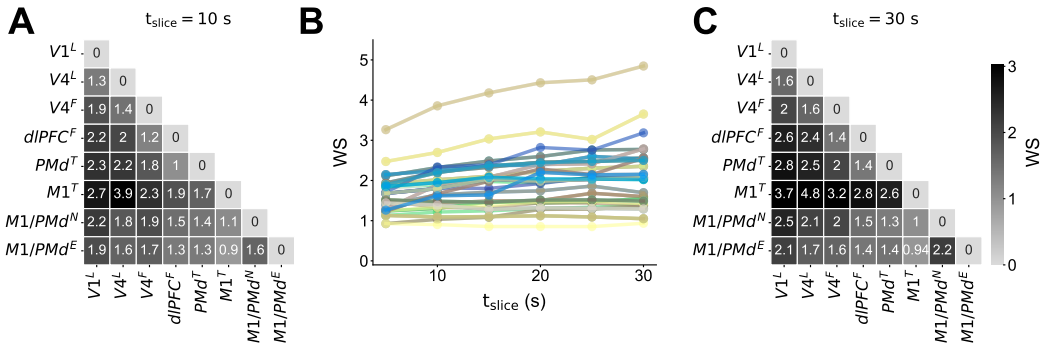


Fig. 4 The Wasserstein distance (WS) measured using different time slice length. **A, C)** WS between all multi-dimensional summary statistics clouds for t_{slice} = 10 s and t_{slice} = 30 s, using the same color scale for both panels. **B)** WS as a function of t_{slice}. Each line represents one pair of recordings, colors are arbitrary.

In order to quantify the differences between the multi-dimensional summary statistics we use the Wasserstein distance (WS), see Methods Multidimensional fitness function for details. A low WS indicates a high similarity between the multi-dimensional summary statistics, with WS = 0 indicating an exact overlap.

The WS depends on the spiking statistics of the underlying data, and on the length of the selected data segments. Long segments contain more spikes and can provide more accurate statistics. However, transient effects on longer time scales can influence the output and lead to unrealistic spiking statistics. Thus, the length of the data segments t_{slice} is a crucial parameter. Here, we measured the WS between pairs of the experimental datasets for different values of t_{slice} (Figure 4). In most cases, the WS appears to slightly increase as t_{slice} becomes longer (Figure 4B). This could be a consequence of having fewer data points as t_{slice} increases. Nevertheless, the relative ordering seems to be fairly conserved between t_{slice} = 10 s and t_{slice} = 30 s, Figure 4A,C; and we are mostly interested in the relative WS values not the absolute ones. Therefore, this comparison suggests that the WS is robust to changes in t_{slice}.

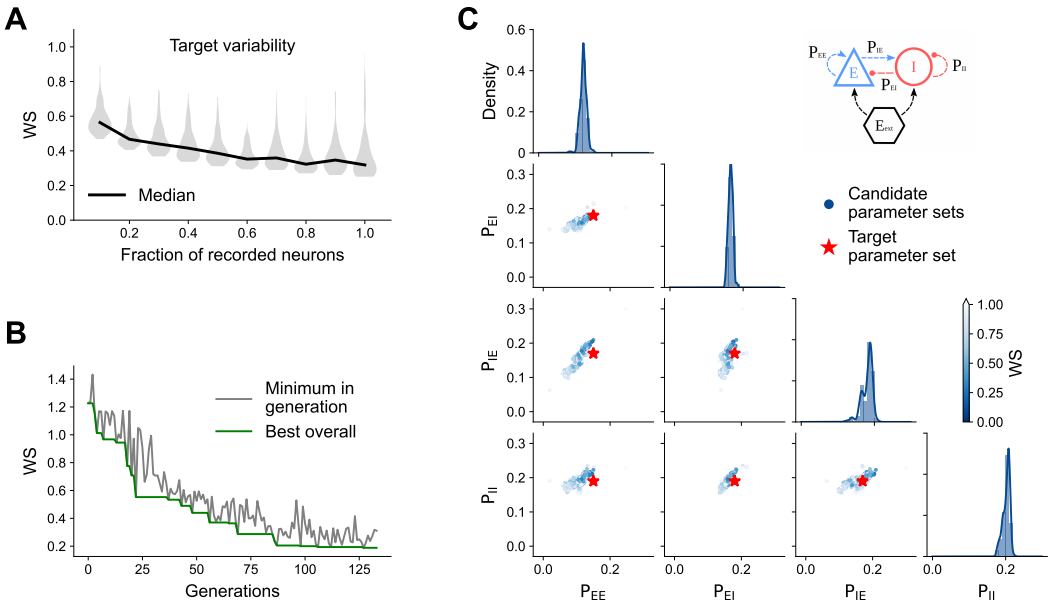


Fig. 5 Proof of concept of optimization method using WS between the multi-dimensional summary statistics. **A** WS variability of the target when recording only a certain fraction of the neurons in the model. Distribution of WS measured from 20 simulations with the same target parameters, but different randomization. **B** Progress of the optimization algorithm. Lowest WS shown overall and within each generation. **C** Pairplot of estimated parameter sets. Off-diagonal plots show the parameters of the estimated models. Each point represents one model, the color indicates the WS for that model, with respect to the target. Diagonal plots show the histogram of estimated parameters for models with WS < 0.5. Inset on top right shows a schematic representation of the fitted model.

Method verification with synthetic data from a small model

To estimate the underlying model parameters from activity data we minimize the WS (see Methods Multidimensional fitness function) between simulation and experimental recordings and search the parameter space using an optimization algorithm (see Methods Optimization algorithm). However, we first need to verify that the optimization method is effective in finding the correct parameter values.

In order to demonstrate that our methods can effectively recover connectivity parameters from multi-dimensional summary statistics alone, we tested them against synthetic data. We generated synthetic data by simulating a simple small balanced spiking neuron network ($N = 12500$, see Methods Small balanced spiking neuron network model). We estimated the multi-dimensional summary statistics from the synthetic data and used it as our target. The objective of this synthetic experiment was for our optimization algorithm to estimate the connectivity parameters (P_{EE}, P_{EI}, P_{IE} and P_{II}) used for the original simulation, based on the multi-dimensional summary statistics alone. See Figure S13 for the multi-dimensional summary statistics of the synthetic data.

First, we quantified the variability within the target model. Since the simulations rely on randomized input, changes to the randomization inevitably lead to differences in the multi-dimensional summary statistics. We simulated the small balanced spiking neuron network using different realisations ($N = 20$ simulations, all parameters equal except the random seed) and measured the WS between all the simulations (Figure 5A). Since all the neurons in the model are statistically equivalent, we also tested the effect of subsampling neurons from the model, and found that it only slightly increases the WS variability (Figure 5A). A low sensitivity to subsampling is desirable, since the experimental data represent only a small fraction of the total number of neurons in a cortical area.

Second, we ensured that the optimization landscape (parameter space) is smooth. A non-smooth parameter space would be an indication of numerical instabilities in the model, the fitness function, or the single unit statistics. Therefore, we performed parameter scans of the small model in the ranges $P_{xy} \in [0.05, 0.20]$ with a 0.01 resolution, corresponding to 16 possible values per parameter. Since we were searching the values of four parameters (P_{EE}, P_{EI}, P_{IE} and P_{II}), the full parameter scan required $N = 16^4 = 65536$ simulations. We found that the parameter space is indeed smooth (Figure S15), suggesting that the optimization should be able to find the global minimum and that no numerical instabilities are present.

Finally, we executed our optimization algorithm for the small balanced spiking neuron network using synthetic data. The optimization converged in about 100 generations (Figure 5B), i.e., $N \approx 12800$ simulations were needed to find the solution, since there were 128 simulations per generation. The estimated parameters with WS < 1 were all found in the vicinity of the target parameters (Figure 5C). Given the variability of the WS, the multi-dimensional summary statistics of all those parameter combinations were equivalent to the target simulation.

Thus we conclude that our optimization methods can successfully estimate the connectivity parameters of a small model based only on the multi-dimensional summary statistics of one target simulation.

Discussion

The aim of this study was to estimate connectivity parameters of the cortical microcircuit from activity data alone. We presented a novel method for model-to-data comparison, using a multi-dimensional single-neuron statistic cloud. We showed that the multi-dimensional summary statistics varies across the cortex, and is robust to data duration lengths. Additionally, we presented a novel optimization approach that minimizes the distance between some specific target spiking activity and some spiking neural network model. Finally, we provided a proof of concept of our optimization method using synthetic data. Future work will see this approach extended to larger more biologically

realistic models, which we will then use to estimate local cortical connectivity parameters.

Our findings suggest rich and varied multi-dimensional summary spiking statistics across different cortical areas. This result is in agreement with previous reports of variability across the cortex [39]. Our spiking statistics did not include the autocorrelation decay timescale due to instabilities arising from the numerical integration. Timescales are well known to vary across cortex at both the individual neuron and population levels [42]. Therefore, future work should also explore the timescales, possibly as a *post hoc* validation.

Our optimization approach was specially designed to balance exploration and exploitation in our case. We initially used evolutionary algorithms akin to [30, 31]. However, we found that when fitting the connection probability between neuron populations, the crossover steps would move the optimizer away from the global optimum. Our interpretation was that the crossover shifting one parameter pushed the model into a completely different regime, which is mostly not the case in genetics, where a single base pair swap often causes only marginal differences. Thus, we removed the crossover step and substituted it with random sampling.

This work has so far demonstrated a proof of concept with one small balanced piking neuron network. There were four free parameters, making the parameter space relatively small, since we could even do brute-force scans. The parameter space for the more biologically realistic microcircuit model is several orders of magnitude larger, since there are (at least) 64 free parameters. Doing a parameter scan for the microcircuit model with the same resolution as before would require $N = 16^{64} \approx 10^{77}$, several orders of magnitude over the estimated total number of atoms of planet Earth ($\sim 10^{50}$). This combinatorial explosion also raises doubts about the effectiveness of the random search algorithm, since it relies on randomly finding the regions close to the global optimum. The parameter space will have to be constrained, for example by imposing a strict E-I balance on the connectivity, or requiring that the total number of synapses lie in a specific range. Furthermore, initial explorations of the parameter space could be performed using computationally cheaper mean-field models [45].

In summary, we presented a novel approach to microconnectome estimation based on multi-dimensional clouds of single-neuron statistics. We found that the multi-dimensional summary statistics vary significantly between cortical areas, suggesting it can be used to uniquely characterize each cortical area. Thus, we exploit the differences in the spiking statistics with our optimization approach, which we have shown to work in a synthetic experiment. Further work will be needed to extend the methods to more biologically realistic models, which could then be used to estimate the connectivity parameters from the experimental data.

Materials and methods

Workflow overview

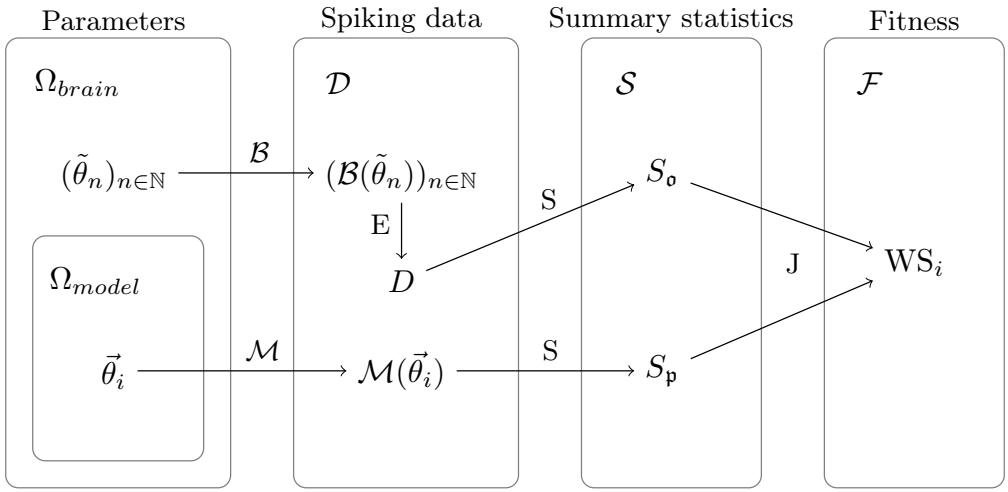


Fig. 6 Schematic illustration of the different statistical value spaces on each level of abstraction. Starting from the underlying parameter spaces in both the model simulations (Ω_{model}) and the cortical areas (Ω_{brain}), both of which produce certain spiking data (in $\mathcal{D}$). Since the spiking data cannot be directly compared we measure some summary statistics (in $\mathcal{S}$), which can then be compared by a difference measure (WS) in the space of a fitness function ($\mathcal{F}$).

Assessing the similarity between experimental and simulated data is a non-trivial procedure that can take many different forms. Our approach is based on the joint distance between the statistics of the data. We schematically show the steps we take from parameters to the fitness function in Figure 6.

Take a set $(\tilde{\theta}_n)_{n\in\mathbb{N}}$ of all the possible parameters that describe the activity in the central nervous system. This would include anything that affects neural activity, such as neuroanatomy, chemical concentrations or environment temperatures. The brain function $\mathcal{B}$, is a highly nonlinear function $\mathcal{B} : \Omega_{brain} \to \mathcal{D}$, mapping from parameter to data space, leading to the time series $(\mathcal{B}(\tilde{\theta}_n))_{n\in\mathbb{N}}$. We will specifically deal with spiking data from the brain, following the assumption that information is encoded within this signal. Current technology does not enable simultaneously recording from all neurons in the brain, thus we need to work with a subset of the brain activity $D \subset (\mathcal{B}(\tilde{\theta}_n))_{n\in\mathbb{N}}$ which we can record using extracellular electrodes E. See a detailed description of the data D in the Methods Electrophysiological data collection in resting state.

On the modeling side, we need to constrain the wide range of parameters that describes the brain activity to the subset of the most relevant parameters $\theta_i \subset (\tilde{\theta}_n)_{n\in\mathbb{N}}$. In our case, we chose to study the neuroanatomy of cortex at the single neuron level, such as the neuron counts, connectivity and single neuron membrane properties. In this work we focus only on the connectivity parameters. The studied parameters are unavoidably linked to the model $\mathcal{M}$ of choice. For a given set of parameters the model will create a time series

$\mathcal{M}(\theta_i)$ of spiking activity. We provide a detailed account of the chosen models M in the Methods Spiking neuron models.

Our objective is to assess whether our choice of model parameters generates a similar spiking activity to recorded experimental activity, $\mathcal{M}(\theta_i) \approx D$. Given the high dimensionality of the data (hundreds or up to thousands of neurons) this is a non-trivial task. We therefore extract summary statistics from the data, in the form of single-unit level summary statistics. We call this function S and it yields two clouds of points $S_{\mathfrak{o}}$ (observation) and $S_{\mathfrak{p}}$ (prediction), where each dot in the cloud corresponds to the summary statistics of a single spike train. See Methods Summary statistics of spiking neuronal data for a detailed account of the used summary statistics.

Finally, we can compare the multi-dimensional summary statistics to each other. Using a judge function J we can estimate a certain fitness WS_i, which is a measure of how similar the multi-dimensional summary statistics are. We used the Wasserstein distance (WS) as a fitness function, see Methods Multidimensional fitness function for details.

Electrophysiological data collection in resting state

All data was collected from macaque monkeys (*Macaca Mulatta*, N = 5) in the resting state. The macaques were sitting in a dim lit room, with no particular task, while the continuous activity from the cortex was recorded. The relevant behavior in each experiment was also tracked using either videos or eye tracking systems.

The recordings were made using different devices thus resulting in data for different layers. Throughout this document we refer to the combined cortical layers L2 and L3 as L23. Consequently, L5 + L6 = L56, and L3 + L4 + L5 = L345.

Table 1 Summary of subjects and recordings included in this study. Name identifier of the subject, cortical area and layers from which the data were recorded, number of electrode contacts in each area and number of recording sessions are indicated.

Name	Sessions	Areas	Layers	Contacts	Source
T	7	M1	All	24	Novel from
	11	PMd	All	24	Kilavik lab
F	59	dlPFC	L23	4	Novel from
	59	V4	L23	4	Gregoriou lab
L	1	V1	L56	896	Spike sorted
	1	V4	L56	128	Chen et al 2021
N	2	M1/PMd	L345	96	Brochier lab
E	2	M1/PMd	L345	96	Brochier lab

Data from macaque T

The data from macaque T was recorded from premotor cortex (PMd) and primary motor cortex (M1) (N=1 subject, N=11 $\tilde{1}$5 min sessions). Acute recordings were made with laminar probes (Plexon and Alpha Omega, 24 contacts, 100 and 200 µm pitch). The laminar probes enabled recording from across the cortical layers. The 200 µm pitch probes could record all layers simultaneously, while the 100 µm pitch probes did not span the entirety of the cortical gray matter. The motor cortex gray matter is known to be approximately 3.5 mm thick [46], with the superficial and deep layers roughly split in half. A guard zone was applied around the middle of the probes of 0.5 mm, where the contacts were excluded. The rest were identified as either superficial (L23) or deep layers (L56). Preliminary analysis of this data has already been published [47].

The raw data was spike sorted offline. Spike sorting identified 5-13 clean single units per probe and session.

In addition to the spiking data, surface Electromyography (EMG) of the contralateral deltoid muscle, the heart rate with an ear clip and a video of the macaque behaviour were recorded. In all the behavioural videos the screen LEDs were used to send a 1 s long blink every minute that can be used to realign video with the neural recordings. We performed a video-based segmentation into behavioural epochs: eyes-open, eyes-closed and movement periods. We excluded the movement periods from our analysis, since they are associated with high motor cortex activity and variability [48]. Including the movement periods would bias the spiking statistics of the motor cortex to include non resting state dynamics.

Data from macaque F

The data from macaque F was recorded from visual area V4 and dorsolateral prefrontal cortex (dlPFC) (N=1 subject, N=59 5 min sessions). Acute recordings were made with up to 4 simultaneous Plexon electrodes, recording from the superficial layers (L2/3) during resting state with.

The eye pupil was tracked for behavioural segmentation into eyes-open and eyes-closed epochs. Spike sorting identified 4-10 clean single units per area and session.

Data from macaque L

The data from macaque L was recorded from visual areas V1 and V4 (N=1 subject, N=1 $\tilde{2}$0 min session). Chronic recordings were made using 16 8x8 electrode Utah arrays (Blackrock microsystems), 2 of them in visual area V4 and the rest in the primary visual cortex (V1), with a total of 1024 electrodes. The electrodes were 1.5 mm long, thus the recordings were made from the deep layers L5 and L6. A full description of the experimental setup and the data collection and preprocessing has already been published [49].

Pupil position and diameter data were collected using an infrared camera in order to determine the direction of gaze and eye closure of the macaques. In addition to the resting state recordings, a visual response task was also performed. The visual response data were used to calculate the signal-to-noise ratio (SNR) of each electrode and all electrodes with an SNR lower than 2 were excluded from further analysis. Additionally, we excluded up to 100 electrodes that contributed to high frequency cross-talk in each session, as reported in the original data publication [49].

The raw data were spike-sorted using a semi-automatic workflow with Spyking Circus—a free, open-source, spike-sorting software written entirely in Python [50]. Extensive description of the methods of this algorithm can be found in their publication, as well as in the online documentation of Spyking Circus[2].

After the automatic sorting the waveform clusters were manually merged and labelled as single-unit activity, multi-unit activity, or noise. Only single-unit activity (SUA) spike trains were included in this study. The waveform signal-to-noise ratio (wfSNR) was calculated for all SUA, and those with an wfSNR < 2 or electrode SNR < 2 (from the visual response task) were excluded from the analysis.

Data from macaques N & E

The data from macaques N & E was recorded from the interface between premotor (PMd) and primary motor (M1) cortex (N=2 subjects, N=2 15-20 min sessions per subject). The recordings were chronic using 16 8x8 electrode Utah arrays (Blackrock microsystems), 2 of them in visual area V4 and the rest in the primary visual cortex (V1)

A full description of the experimental setup and the data collection and preprocessing—for the accompanying experiment in the same subject—has already been published [51]. An extensive analysis of the resting state data has also been published [48].

In addition to the registration of brain activity, the monkey's behavior was video recorded and synchronized with the electrophysiology recording.

Spiking neuron models

Two different spiking neuron models were used in this study, a small balanced spiking neuron network [52, 53] and a cortical microcircuit model [5]. Both of which were simulated using NEST 3.2 [54].

Small balanced spiking neuron network model

We first used a small balanced spiking neuron network model [52, 53] to demonstrate the effectiveness of our methods. The model consisted of two populations one excitatory and one inhibitory connected to each other. An overview of

[2] spyking-circus.readthedocs.io

the model can be found in Table S1, and all model parameters are listed in Table S2.

For the synthetic data we used the following connectivity matrix:

$$P_0 = \begin{bmatrix} P_{EE} & P_{EI} \\ P_{IE} & P_{II} \end{bmatrix} = \begin{bmatrix} 0.15 & 0.18 \\ 0.17 & 0.19 \end{bmatrix},$$

which show the connection probability between the given populations, e.g. P_{EI} is the probability of $I \rightarrow E$ connections.

Microcircuit model

We used the cortical microcircuit model described in [5] for a more realistic model of the cortex. The model includes all layers (L23, L4, L5, L6) under a square of 1 mm^2 cortical surface. Each layer consists of two populations of point neurons, one excitatory and one inhibitory, thus a total of 8 populations. The total number of neurons was based on layer-resolved stereological neuron estimates. All layers receive a constant background input with Poisson noise.

An overview of the model can be found in Table S3, and the model parameters are listed in Table S4 and Table S5.

For the synthetic data we used the following connectivity matrix:

$$P_0 = \begin{bmatrix} 0.1009 & 0.1689 & 0.0437 & 0.0818 & 0.0323 & 0 & 0.0076 & 0 \\ 0.1346 & 0.1371 & 0.0316 & 0.0515 & 0.0755 & 0 & 0.0042 & 0 \\ 0.0077 & 0.0059 & 0.0497 & 0.1350 & 0.0067 & 0.0003 & 0.0453 & 0 \\ 0.0691 & 0.0029 & 0.0794 & 0.1597 & 0.0033 & 0 & 0.1057 & 0 \\ 0.1004 & 0.0622 & 0.0505 & 0.0057 & 0.0831 & 0.3726 & 0.0204 & 0 \\ 0.0548 & 0.0269 & 0.0257 & 0.0022 & 0.0600 & 0.3158 & 0.0086 & 0 \\ 0.0156 & 0.0066 & 0.0211 & 0.0166 & 0.0572 & 0.0197 & 0.0396 & 0.2252 \\ 0.0364 & 0.0010 & 0.0034 & 0.0005 & 0.0277 & 0.0080 & 0.0658 & 0.1443 \end{bmatrix}$$

which was previously derived from anatomical studies [4, 5].

Summary statistics of spiking neuronal data

The experimental and simulated spike trains are sliced into 10 second samples. We compute a number of summary statistics for each one of these data samples and construct a multi-dimensional cloud of spiking neuron statistics. This multi-dimensional summary statistics are used to characterize a given area and to assess the similarities between experiments and simulations.

We use the following summary statistics of spike trains:

- Mean firing rate (FR), given as the total number of spikes divided by the duration of the recording sample.

- Revised local variability of the inter-spike interval (LvR) [39],

$$\mathrm{LvR} = \frac{3}{N-1} \sum_{i=1}^{N-1} \left(1 - \frac{4 I_i I_{i+1}}{(I_i + I_{i+1})^2}\right) \left(1 + \frac{4R}{I_i + I_{i+1}}\right) \tag{1}$$

  where N is the total number of spikes, I_i is the ith inter-spike interval, and $R = 5$ ms is the refractoriness constant. A value of LVR $= 0$ corresponds to perfectly rhythmic spiking, whereas LVR $= 1$ corresponds to Poisson point process.
- The average of the Pearson cross-correlation (binwidth = 2 ms) with all other simultaneously recorded neurons ($\mathrm{CC_{avg}}$). This metric quantifies whether the measured neuron is tuned to many of the other neuron (i.e. choroist) or if it is spiking independently (i.e. soloist).
- The standard deviation of the Pearson cross-correlation with all other simultaneously recorded neurons ($\mathrm{CC_{std}}$).

Further summary statistics were considered but not used in our analysis because they were strongly correlated to other statistics. If strongly correlated statistics were introduced, some spiking properties would be over-represented in our fitness function, since all statistics are given the same weight in the final distance. For example the coefficient of variation (Cv, Cv2) [55] and local variability (Lv) [40] of the inter-spike intervals are excluded due to their similarity to the LvR. Introducing all Cv, Cv2, Lv and LvR would significantly increase the representation of the inter-spike interval variation, without a large increase in the explained variance. Another statistic that was considered was the spike triggered population response (stPR) [56], but not included due to its similarity with the $\mathrm{CC_{avg}}$.

All summary statistics are dependent on, and may be biased by, the number of spikes in the given spike train. For example, the LvR requires a minimum of three spikes in total, or it cannot be computed. To ensure reliable estimations of the summary statistics, we set a threshold of FR ≥ 2 spikes/s for any given spike train to be included in the analysis, both in the experimental and simulated data. Furthermore, all the considered statistics assume stationarity of the data, thus the chosen length of spike trains must not bee too long. We therefore choose a spike train sample length ($\mathrm{t_{slice}}$) of 10 seconds.

All metrics were computed using their implementation in the elephant toolbox [57], within the NetworkUnit reproducible testing framework [58].

Multidimensional fitness function

Selection of a fitness function

In order to assess the fitness of a certain parameter set we need to compare the multi-dimensional single neuron statistics. Our problem thus reduces to computing a distance between multi-dimensional probability distributions. Such problems have already been explored in the context of GANs, where the

Wasserstein distance [59], also known as the earth mover's distance, was found to have many desirable properties:

1. agnostic about the underlying statistical distribution
2. jointly evaluates multivariate distributions, thus incorporating the covariance structure
3. can compare samples of different sizes
4. is a true distance: symmetric and positive definite
5. is extendable to higher dimensions
6. is numerically robust on point-distributions (no integration needed)

Not all distribution similarity measures share these properties. For example, the Kullback-Leibler divergence violates properties 4 and 6. The Jensen-Shannon entropy violates property 6.

Wasserstein distance

We used the Wasserstein distance (WS), defined as an optimal transport problem between the observations ($S_{\mathfrak{o}}$, the target multi-dimensional summary statistics) and the predictions ($S_{\mathfrak{p}}$, the multi-dimensional summary statistics from the candidate model). Each multi-dimensional summary statistics cloud is an $N \times M$ matrix of N number of samples (spike train slices) and M number of summary statistics. The observations and predictions must have the same number of summary statistics $M_{\mathfrak{o}} = M_{\mathfrak{p}}$, but can have different sample sizes $N_{\mathfrak{o}}$ and $N_{\mathfrak{p}}$.

First, we normalized (z-scored) the statistics cloud with respect to the observations $S_{\mathfrak{o}}$ across samples, for both $S_{\mathfrak{o}}$ and $S_{\mathfrak{p}}$. This normalization ensured that the measured distances remain comparable across many different predictions $S_{\mathfrak{p}}$, since they are all normalized to the same range. The normalization step also ensured that all metrics M were equally weighted for the distance measurement, without normalization the firing rates would affect the distance by several orders of magnitude more than the correlation statistics ($\mathrm{CC}_{\mathrm{avg}}$, $\mathrm{CC}_{\mathrm{std}}$).

Second, we assigned equal weights (mass) $w_{\mathfrak{o}}$, $w_{\mathfrak{p}}$ to each neuron within each statistics cloud, such that $\sum^{N_{\mathfrak{o}}} w_{\mathfrak{o}} = \sum^{N_{\mathfrak{p}}} w_{\mathfrak{p}} = 1$. Thus, the same total mass was assigned to both $S_{\mathfrak{o}}$ and $S_{\mathfrak{p}}$ and thus statistics clouds with different sample sizes could be compared.

Finally, we searched for the optimal transport of mass between $S_{\mathfrak{o}}$ and $S_{\mathfrak{p}}$, by finding the graph configuration that minimizes the work required to transport all the weights:

$$\min \sum_{i}^{S_{\mathfrak{o}}} \sum_{j}^{S_{\mathfrak{p}}} w_{i,j} \cdot d_{i,j} \tag{2}$$

where $w_{i,j}$ is the weight transported between points i and j, and $d_{i,j}$ is the Euclidean distance between them. In order to find the optimal graph configuration we used a simplex algorithm implemented in OpenCV for this purpose [60].

The Wasserstein distance (WS) is the work normalized by the transported weight, which we defined to $\sum^{N} w = 1$, thus $\sum_{i}^{S_o} \sum_{j}^{S_p} w_{i,j} = 1$, and the WS is therefore

$$\mathrm{WS} = \sum_{i}^{S_o} \sum_{j}^{S_p} w_{i,j} \cdot d_{i,j} \tag{3}$$

Multi-objective vs single-objective optimization

An alternative approach to a multi-dimensional distance is a multi-objective optimization, where the multi-dimensional distribution is separated into univariate distributions. Each distributions is tested separately and a Pareto front of all objectives is sought. However, toy tests with this approach struggled to fit all summary statistics simultaneously and would lead to inconsistent results. We encountered three problems with a multi-objective approach:

1. Logical "Or": different summary statistics fitted separately and not simultaneously.
2. Ignored covariance structure due to univariate testing.
3. Sensitive to the choice of statistical test or distance (e.g. Kolmogorov-Smirnov, t-test).

Ultimately we found that single-objective optimization with a multi-dimensional distance was more robust than a multi-objective optimization.

Optimization algorithm

We used a simple optimization algorithm to explore the parameter space of our models maximizing the similarity between the target observations and the model predictions. As described in the previous section, we used the WS as a fitness function, which the optimizer had to minimize.

For our optimization we devised a random search algorithm, see pseudocode 1. All the hyperparameters of the algorithm are listed and briefly described in Table 2.

Algorithm 1 Pseudocode of random search algorithm

```
population ← generate random individuals
for generation ≤ N do
    population.fitness ← evaluate(population)
    survivors ← select(all individuals)
    mutants ← mutate(survivors)
    newcomers ← generate random individuals
    population ← newcomers and mutants
end for
```

First, our optimizer generated an initial population of individuals $\vec{\theta}_0$, $\vec{\theta}_1$, ..., $\vec{\theta}_N$, of length $N_{\text{population}}$. Then, the individuals were evaluated:

1. A spiking neuron simulation was run with the corresponding parameters $\vec{\theta}_i$,
2. the multi-dimensional summary statistics S_i were calculated from the simulated spike trains,
3. and the WS was calculated between the S_i and the pre-computed target multi-dimensional summary statistics S_o.

The target multi-dimensional summary statistics S_o were pre-computed to avoid unnecessary calculations of summary statistics at every evaluation step.

The evaluation step was parallelized in a high-performance computing (HPC) system. Thus, this step took only as long as the slowest individual in the generation. It is to be noted that the computational cost of spiking model simulations scales up with the total number of synapses, thus some parameter combinations with unrealistically many synapses would require a much longer simulation time. To avoid this problem we set a time limit for the evaluation step (tuned to each model); if the time limit was reached, the individual would get the worst possible fitness.

After the evaluation step the $p_{\text{survival}} \cdot N_{\text{population}}$ with the best fitness (lowest WS) were selected, and the rest of the models discarded. We call the remaining individuals the survivors.

To the survivors we included a small fraction of the best ever individuals, by selecting $p_{\text{from\_best}} \cdot N_{\text{population}}$ out the a hall of fame with the N_{best} models with the highest score of the whole optimization. In the early generations this step is superfluous. However, some models can by random chance drift away from a good fitness region towards a worse one and including some of the best models back helps ensure that the algorithm exploits the good fitness areas.

Then, we mutated the survivors (including those selected from the hall of fame) using a Gaussian mutation with the mean equal to the parameter values and the standard deviation fixed to a certain value σ_{mutation} (see Table 2. In some cases, with a probability of p_{gradient}, the mutation was not random but rather followed the natural gradient, calculated from the rest of the nearest individuals in parameter space (including previously discarded ones). We

estimate the natural gradient as follows:

$$\vec{g} \approx \frac{1}{N_{\text{nearest}}} \sum_{j=1}^{N_{\text{nearest}}} (\text{WS}_j - \text{WS}_0) \frac{\vec{\theta}_j - \vec{\theta}_0}{\sqrt{\vec{\theta}_j \cdot \vec{\theta}_0}} \tag{4}$$

where $\vec{g}$ is the gradient vector, $\vec{\theta}_0$ is the current individual with fitness WS_0 (i.e. the point in parameter space for which the gradient is being estimated), and $\vec{\theta}_j$ are the nearest neighbours (N_{nearest}) to $\vec{\theta}_0$ in the parameter space.

Finally, the population is completed with new random individuals, such that the new population has again a length of $N_{\text{population}}$. The new population then undergoes the same evaluation, selection and mutation steps until a certain prescribed number of iterations $N_{\text{iterations}}$ is reached.

Table 2 Random search optimization parameters

Optimization description		
Individual generation		Latin hypercube sampling for $\theta \in [0.0, 0.20]$
Fitness		Wasserstein distance between the target and individual statistics clouds.
Selection		Fraction of individuals ($\text{p}_{\text{survival}}$) with the best fitness
Mutation		Gaussian with $\mu = \theta$ and standard deviation σ_{mutation}. Mutations sometimes follow the gradient $\vec{g}$ and sometimes are random
Gradient		Natural gradient estimated from nearest neighbours
Optimization parameters		
	123	Seed for random numbers
$N_{\text{population}}$	128	Number of models per generation
$N_{\text{iterations}}$	250	Maximum number of total iterations
σ_{mutation}	0.01	Standard deviation for the Gaussian mutation.
$\text{p}_{\text{survival}}$	0.5	Fraction of the population that will not be discarded before the next generation.
N_{best}	40	Number of models in the hall of fame
$\text{p}_{\text{from\_best}}$	0.1	Fraction of the new population derived from the hall of fame.
$\text{p}_{\text{gradient}}$	0.5	Probability of mutations following the gradient, otherwise a random direction is chosen.
N_{nearest}	1000	Number of nearest individuals to use for the gradient estimation

Funding

This project received funding from the DFG Priority Program (SPP 2041 "Computational Connectomics") [S.J.van Albada: AL 2041/1-1]; the EU's Horizon 2020 Framework Grant Agreement No. 785907 (HBP-SGA2); the DFG (RTG 2416 "MultiSenses-MultiScales"); the CNRS Multidisciplinary Exploratory Projects initiative; and the FLAG-ERA grant [PrimCorNet].

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

[1] Felleman, D.J., Van Essen, D.C.: Distributed Hierarchical Processing in the Primate Cerebral Cortex (1991)

[2] Bakker, R., Wachtler, T., Diesmann, M.: CoCoMac 2.0 and the future of tract-tracing databases. Frontiers in Neuroinformatics **6** (2012). https://doi.org/10.3389/fninf.2012.00030

[3] Markov, N.T., Ercsey-Ravasz, M., Ribeiro Gomes, A.R., Lamy, C., Magrou, L., Vezoli, J., Misery, P., Falchier, A., Quilodran, R., Gariel, M.A., Sallet, J., Gamanut, R., Huissoud, C., Clavagnier, S., Giroud, P., Sappey-Marinier, D., Barone, P., Dehay, C., Toroczkai, Z., Knoblauch, K., Van Essen, D.C., Kennedy, H.: A weighted and directed interareal connectivity matrix for macaque cerebral cortex. Cerebral Cortex **24**(1), 17–36 (2012). https://doi.org/10.1093/cercor/bhs270

[4] Binzegger, T.: A Quantitative Map of the Circuit of Cat Primary Visual Cortex. Journal of Neuroscience **24**(39), 8441–8453 (2004). https://doi.org/10.1523/JNEUROSCI.1400-04.2004

[5] Potjans, T.C., Diesmann, M.: The Cell-Type Specific Cortical Microcircuit: Relating Structure and Activity in a Full-Scale Spiking Network Model. Cerebral Cortex **24**(3), 785–806 (2014). https://doi.org/10.1093/cercor/bhs358

[6] Douglas, R.J., Martin, K.A.C., Whitteridge, D.: A Canonical Microcircuit for Neocortex. Neural Computation **1**(4), 480–488 (1989). https://doi.org/10.1162/neco.1989.1.4.480

[7] Thomson, A.M.: 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**(9), 936–953 (2002). https://doi.org/10.1093/cercor/12.9.936

[8] Song, S., Sjöström, P.J., Reigl, M., Nelson, S., Chklovskii, D.B.: Highly Nonrandom Features of Synaptic Connectivity in Local Cortical Circuits. PLoS Biology **3**(3), 68 (2005). https://doi.org/10.1371/journal.pbio.0030068

[9] Perin, R., Berger, T.K., Markram, H.: A synaptic organizing principle for cortical neuronal groups. Proceedings of the National Academy of Sciences **108**(13), 5419–5424 (2011). https://doi.org/10.1073/pnas.1016051108

[10] Kodandaramaiah, S.B., Flores, F.J., Holst, G.L., Singer, A.C., Han, X., Brown, E.N., Boyden, E.S., Forest, C.R.: Multi-neuron intracellular recording in vivo via interacting autopatching robots. eLife **7**, 24656 (2018). https://doi.org/10.7554/eLife.24656

[11] Callaway, E.M., Katz, L.C.: Photostimulation using caged glutamate reveals functional circuitry in living brain slices. Proceedings of the National Academy of Sciences **90**(16), 7661–7665 (1993). https://doi.org/10.1073/pnas.90.16.7661

[12] Nikolenko, V., Poskanzer, K.E., Yuste, R.: Two-photon photostimulation and imaging of neural circuits. Nature Methods **4**(11), 943–950 (2007). https://doi.org/10.1038/nmeth1105

[13] Weiler, N., Wood, L., Yu, J., Solla, S.A., Shepherd, G.M.G.: Top-down laminar organization of the excitatory network in motor cortex. Nature Neuroscience **11**(3), 360–366 (2008). https://doi.org/10.1038/nn2049

[14] Anderson, C.T., Sheets, P.L., Kiritani, T., Shepherd, G.M.G.: Sublayer-specific microcircuits of corticospinal and corticostriatal neurons in motor cortex. Nature Neuroscience **13**(6), 739–744 (2010). https://doi.org/10.1038/nn.2538

[15] Noguchi, J., Nagaoka, A., Watanabe, S., Ellis-Davies, G.C.R., Kitamura, K., Kano, M., Matsuzaki, M., Kasai, H.: *In vivo* two-photon uncaging of glutamate revealing the structure-function relationships of dendritic spines in the neocortex of adult mice: *In vivo* two-photon glutamate uncaging. The Journal of Physiology **589**(10), 2447–2457 (2011). https://doi.org/10.1113/jphysiol.2011.207100

[16] Yamawaki, N., Borges, K., Suter, B.A., Harris, K.D., Shepherd, G.M.G.: A genuine layer 4 in motor cortex with prototypical synaptic circuit connectivity. eLife **3**, 05422 (2014). https://doi.org/10.7554/eLife.05422

[17] Yamawaki, N., Shepherd, G.M.G.: Synaptic Circuit Organization of Motor Corticothalamic Neurons. Journal of Neuroscience **35**(5), 2293–2307 (2015). https://doi.org/10.1523/JNEUROSCI.4023-14.2015

[18] Berger, D., Warren, D., Normann, R., Arieli, A., Grün, S.: Spatially organized spike correlation in cat visual cortex. Neurocomputing **70**(10-12), 2112–2116 (2007). https://doi.org/10.1016/j.neucom.2006.10.141

[19] Dann, B., Michaels, J.A., Schaffelhofer, S., Scherberger, H.: Uniting functional network topology and oscillations in the fronto-parietal single unit network of behaving primates. eLife **5**, 15719 (2016). https://doi.org/10.7554/eLife.15719

[20] English, D.F., McKenzie, S., Evans, T., Kim, K., Yoon, E., Buzsáki, G.: Pyramidal Cell-Interneuron Circuit Architecture and Dynamics in Hippocampal Networks. Neuron **96**(2), 505–5207 (2017). https://doi.org/10.1016/j.neuron.2017.09.033

[21] Pastore, V.P., Massobrio, P., Godjoski, A., Martinoia, S.: Identification of excitatory-inhibitory links and network topology in large-scale neuronal assemblies from multi-electrode recordings. PLOS Computational Biology **14**(8), 1006381 (2018). https://doi.org/10.1371/journal.pcbi.1006381

[22] Kobayashi, R., Kurita, S., Kurth, A., Kitano, K., Mizuseki, K., Diesmann, M., Richmond, B.J., Shinomoto, S.: Reconstructing neuronal circuitry from parallel spike trains. Nature Communications **10**(1), 4468 (2019). https://doi.org/10.1038/s41467-019-12225-2

[23] Ostojic, S., Brunel, N., Hakim, V.: How Connectivity, Background Activity, and Synaptic Properties Shape the Cross-Correlation between Spike Trains. Journal of Neuroscience **29**(33), 10234–10253 (2009). https://doi.org/10.1523/JNEUROSCI.1275-09.2009

[24] Bullmore, E., Sporns, O.: Complex brain networks: graph theoretical analysis of structural and functional systems. Nature Reviews Neuroscience **10**(3), 186–198 (2009). https://doi.org/10.1038/nrn2575

[25] Eickhoff, S.B., Jbabdi, S., Caspers, S., Laird, A.R., Fox, P.T., Zilles, K., Behrens, T.E.J.: Anatomical and Functional Connectivity of Cytoarchitectonic Areas within the Human Parietal Operculum. Journal of Neuroscience **30**(18), 6409–6421 (2010). https://doi.org/10.1523/JNEUROSCI.5664-09.2010

[26] Das, A., Fiete, I.R.: Systematic errors in connectivity inferred from activity in strongly recurrent networks. Nature Neuroscience **23**(10), 1286–1296 (2020). https://doi.org/10.1038/s41593-020-0699-2

[27] Prinz, A.A., Bucher, D., Marder, E.: Similar network activity from disparate circuit

parameters. Nature Neuroscience **7**(12), 1345–1352 (2004). https://doi.org/10.1038/nn1352

[28] Stringer, C., Pachitariu, M., Steinmetz, N.A., Okun, M., Bartho, P., Harris, K.D., Sahani, M., Lesica, N.A.: Inhibitory control of correlated intrinsic variability in cortical networks. eLife **5**, 19695 (2016). https://doi.org/10.7554/eLife.19695

[29] Druckmann, S.: A novel multiple objective optimization framework for constraining conductance-based neuron models by experimental data. Frontiers in Neuroscience **1**(1), 7–18 (2007). https://doi.org/10.3389/neuro.01.1.1.001.2007

[30] Rossant, C.: Automatic fitting of spiking neuron models to electrophysiological recordings. Frontiers in Neuroinformatics **4** (2010). https://doi.org/10.3389/neuro.11.002.2010

[31] Carlson, K.D., Nageswaran, J.M., Dutt, N., Krichmar, J.L.: An efficient automated parameter tuning framework for spiking neural networks. Frontiers in Neuroscience **8** (2014). https://doi.org/10.3389/fnins.2014.00010

[32] Diaz-Pier, S., Naveau, M., Butz-Ostendorf, M., Morrison, A.: Automatic Generation of Connectivity for Large-Scale Neuronal Network Models through Structural Plasticity. Frontiers in Neuroanatomy **10** (2016). https://doi.org/10.3389/fnana.2016.00057

[33] Paninski, L., Pillow, J.W., Simoncelli, E.P.: Maximum Likelihood Estimation of a Stochastic Integrate-and-Fire Neural Encoding Model. Neural Computation **16**(12), 2533–2561 (2004). https://doi.org/10.1162/0899766042321797

[34] Pillow, J.W.: Prediction and Decoding of Retinal Ganglion Cell Responses with a Probabilistic Spiking Model. Journal of Neuroscience **25**(47), 11003–11013 (2005). https://doi.org/10.1523/JNEUROSCI.3305-05.2005

[35] Ladenbauer, J., McKenzie, S., English, D.F., Hagens, O., Ostojic, S.: Inferring and validating mechanistic models of neural microcircuits based on spike-train data. Nature Communications **10**(1), 4933 (2019). https://doi.org/10.1038/s41467-019-12572-0

[36] René, A., Longtin, A., Macke, J.H.: Inference of a mesoscopic population model from population spike trains (2019). https://doi.org/10.48550/ARXIV.1910.01618. Publisher: arXiv Version Number: 2

[37] Bittner, S.R., Palmigiano, A., Piet, A.T., Duan, C.A., Brody, C.D., Miller, K.D., Cunningham, J.P.: Interrogating theoretical models of neural computation with emergent property inference. bioRxiv (2019). https://doi.org/10.1101/837567

[38] Gonçalves, P.J., Lueckmann, J.-M., Deistler, M., Nonnenmacher, M., Öcal, K., Bassetto, G., Chintaluri, C., Podlaski, W.F., Haddad, S.A., Vogels, T.P., Greenberg, D.S., Macke, J.H.: Training deep neural density estimators to identify mechanistic models of neural dynamics. eLife **9**, 56261 (2020). https://doi.org/10.7554/eLife.56261

[39] Mochizuki, Y., Onaga, T., Shimazaki, H., Shimokawa, T., Tsubo, Y., Kimura, R., Saiki, A., Sakai, Y., Isomura, Y., Fujisawa, S., Shibata, K.-i., Hirai, D., Furuta, T., Kaneko, T., Takahashi, S., Nakazono, T., Ishino, S., Sakurai, Y., Kitsukawa, T., Lee, J.W., Lee, H., Jung, M.W., Babul, C., Maldonado, P.E., Takahashi, K., Arce-McShane, F.I., Ross, C.F., Sessle, B.J., Hatsopoulos, N.G., Brochier, T., Riehle, A., Chorley, P., Grun, S., Nishijo, H., Ichihara-Takeda, S., Funahashi, S., Shima, K., Mushiake, H., Yamane, Y., Tamura, H., Fujita, I., Inaba, N., Kawano, K., Kurkin, S., Fukushima, K., Kurata, K., Taira, M., Tsutsui, K.-I., Ogawa, T., Komatsu, H., Koida, K., Toyama, K., Richmond, B.J., Shinomoto, S.: Similarity in Neuronal Firing Regimes across Mammalian Species. Journal of Neuroscience **36**(21), 5736–5747 (2016). https://doi.org/10.1523/JNEUROSCI.0230-16.2016

[40] Shinomoto, S., Shima, K., Tanji, J.: Differences in Spiking Patterns Among Cortical Neurons. Neural Computation **15**(12), 2823–2842 (2003). https://doi.org/10.1162/089976603322518759

[41] Murray, J.D., Bernacchia, A., Freedman, D.J., Romo, R., Wallis, J.D., Cai, X., Padoa-Schioppa, C., Pasternak, T., Seo, H., Lee, D., Wang, X.-J.: A hierarchy of intrinsic timescales across primate cortex. Nature Neuroscience **17**(12), 1661–1663 (2014). https://doi.org/10.

1038/nn.3862

[42] Chaudhuri, R., Knoblauch, K., Gariel, M.-A., Kennedy, H., Wang, X.-J.: A Large-Scale Circuit Mechanism for Hierarchical Dynamical Processing in the Primate Cortex. Neuron **88**(2), 419–431 (2015). https://doi.org/10.1016/j.neuron.2015.09.008

[43] Runyan, C.A., Piasini, E., Panzeri, S., Harvey, C.D.: Distinct timescales of population coding across cortex. Nature **548**(7665), 92–96 (2017). https://doi.org/10.1038/nature23020

[44] Lindman, H.R.: Analysis of Variance in Experimental Design. Springer, New York (1992)

[45] Layer, M., Senk, J., Essink, S., van Meegen, A., Bos, H., Helias, M.: NNMT: Mean-Field Based Analysis Tools for Neuronal Network Models. Frontiers in Neuroinformatics **16**, 835657 (2022). https://doi.org/10.3389/fninf.2022.835657

[46] Koo, B.-B., Schettler, S.P., Murray, D.E., Lee, J.-M., Killiany, R.J., Rosene, D.L., Kim, D.-S., Ronen, I.: Age-related effects on cortical thickness patterns of the Rhesus monkey brain. Neurobiology of Aging **33**(1), 200–2320031 (2012). https://doi.org/10.1016/j.neurobiolaging.2010.07.010

[47] Kilavik, B.E.: Directional selectivity across macaque motor cortical layers during reach planning and execution. Technical report (2018)

[48] Dąbrowska, P.A., Voges, N., von Papen, M., Ito, J., Dahmen, D., Riehle, A., Brochier, T., Grün, S.: On the Complexity of Resting State Spiking Activity in Monkey Motor Cortex. Cerebral Cortex Communications **2**(3), 033 (2021). https://doi.org/10.1093/texcom/tgab033

[49] Chen, X., Morales-Gregorio, A., Sprenger, J., Kleinjohann, A., Sridhar, S., van Albada, S.J., Grün, S., Roelfsema, P.R.: 1024-channel electrophysiological recordings in macaque V1 and V4 during resting state. Scientific Data **9**(1), 77 (2022). https://doi.org/10.1038/s41597-022-01180-1

[50] Yger, P., Spampinato, G.L., Esposito, E., Lefebvre, B., Deny, S., Gardella, C., Stimberg, M., Jetter, F., Zeck, G., Picaud, S., Duebel, J., Marre, O.: A spike sorting toolbox for up to thousands of electrodes validated with ground truth recordings in vitro and in vivo. eLife **7**, 34518 (2018). https://doi.org/10.7554/eLife.34518

[51] Brochier, T., Zehl, L., Hao, Y., Duret, M., Sprenger, J., Denker, M., Grün, S., Riehle, A.: Massively parallel recordings in macaque motor cortex during an instructed delayed reach-to-grasp task. Scientific Data **5**(1), 180055 (2018). https://doi.org/10.1038/sdata.2018.55

[52] Brunel, N.: Dynamics of Sparsely Connected Networks of Excitatory and Inhibitory Spiking Neurons. Journal of Computational Neuroscience **8**, 183–208 (2000)

[53] Ostojic, S.: Two types of asynchronous activity in networks of excitatory and inhibitory spiking neurons. Nature Neuroscience **17**(4), 594–600 (2014). https://doi.org/10.1038/nn.3658

[54] de Schepper, R., Eppler, J.M., Kurth, A., Nagendra Babu, P., Deepu, R., Spreizer, S., Trensch, G., Pronold, J., Vennemo, S.B., Graber, S., Morales-Gregorio, A., Linssen, C., Benelhedi, M.A., Mørk, H., Morrison, A., Terhorst, D., Mitchell, J., Diaz, S., Kitayama, I., Enan, M., Kamiji, N.L., Plesser, H.E.: NEST 3.2. Zenodo (2022). https://doi.org/10.5281/ZENODO.5886893

[55] Holt, G.R., Softky, W.R., Koch, C., Douglas, R.J.: Comparison of discharge variability in vitro and in vivo in cat visual cortex neurons. Journal of Neurophysiology **75**(5), 1806–1814 (1996). https://doi.org/10.1152/jn.1996.75.5.1806

[56] Okun, M., Steinmetz, N.A., Cossell, L., Iacaruso, M.F., Ko, H., Barthó, P., Moore, T., Hofer, S.B., Mrsic-Flogel, T.D., Carandini, M., Harris, K.D.: Diverse coupling of neurons to populations in sensory cortex. Nature **521**(7553), 511–515 (2015). https://doi.org/10.1038/nature14273

[57] Denker, M., Köhler, C., Morales-Gregorio, A., Kleinjohann, A., Bouss, P., Stella, A., Jurkus,

R., Kramer, M., Dąbrowska, P., Kern, M., Kurth, A.C., Gutzen, R., Porrmann, F., Pilz, S.: Elephant 0.11.1. Zenodo (2022). https://doi.org/10.5281/ZENODO.1186602

[58] Gutzen, R., von Papen, M., Trensch, G., Quaglio, P., Grün, S., Denker, M.: Reproducible Neural Network Simulations: Statistical Methods for Model Validation on the Level of Network Activity Data. Frontiers in Neuroinformatics **12**, 90 (2018). https://doi.org/10.3389/fninf.2018.00090

[59] Arjovsky, M., Chintala, S., Bottou, L.: Wasserstein GAN (2017). https://doi.org/10.48550/ARXIV.1701.07875

[60] Bradski, G.: The OpenCV Library. Dr. Dobb's Journal of Software Tools (2000)

Supplementary figures

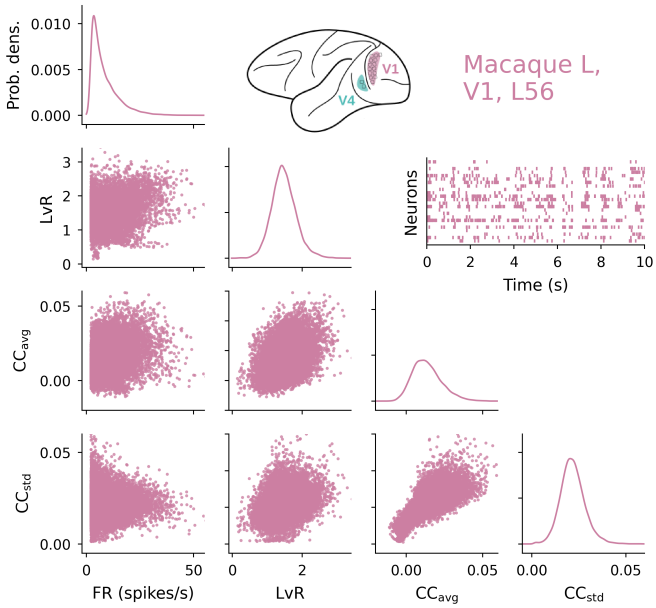


Fig. S1 Multi-dimensional summary statistics for macaque L, area V1, L56. Sample spike trains shown.

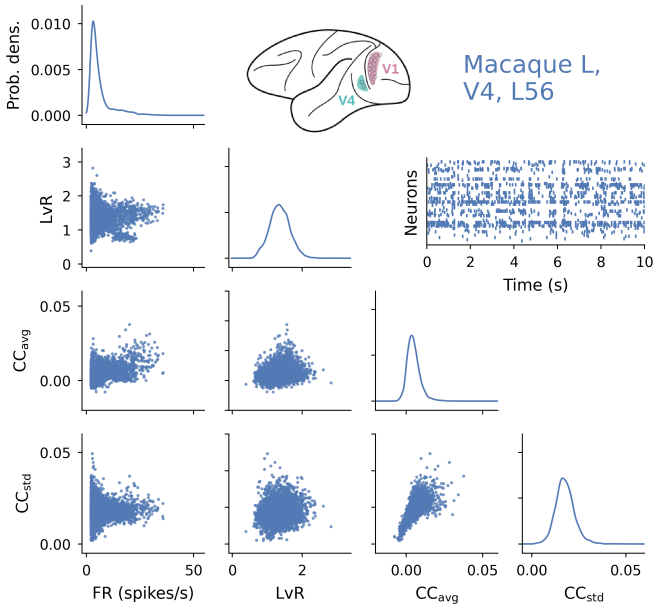


Fig. S2 Multi-dimensional summary statistics for macaque L, area V4, L56. Sample spike trains shown.

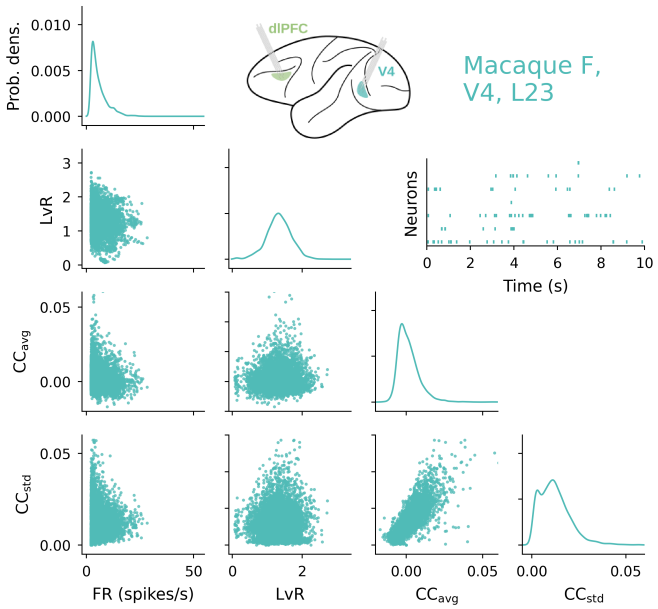


Fig. S3 Multi-dimensional summary statistics for macaque F, area V4, L23. Sample spike trains shown.

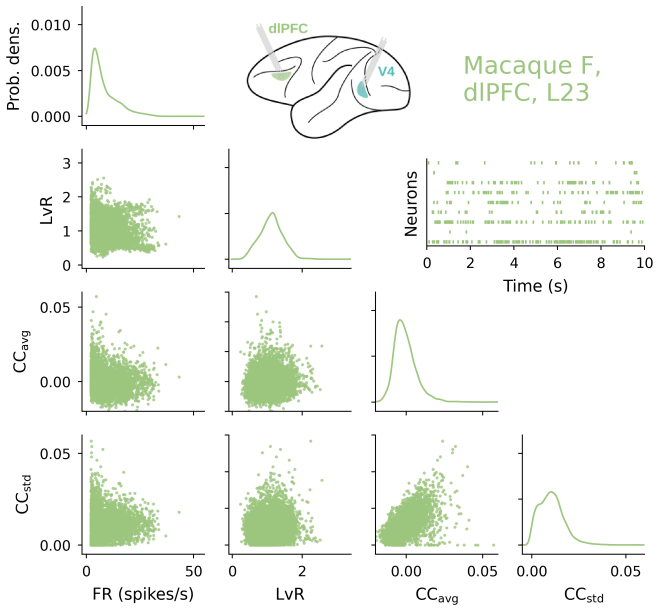


Fig. S4 Multi-dimensional summary statistics for macaque F, area dlPFC, L23. Sample spike trains shown.

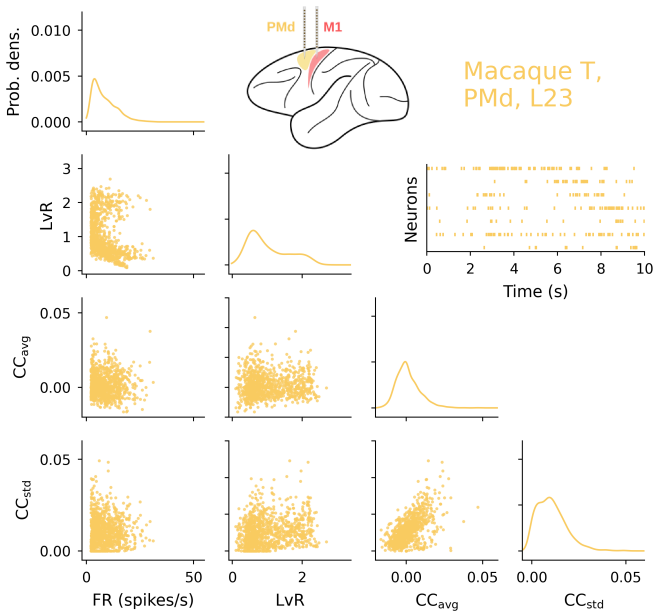


Fig. S5 Multi-dimensional summary statistics for macaque T, area PMd, L23. Sample spike trains shown.

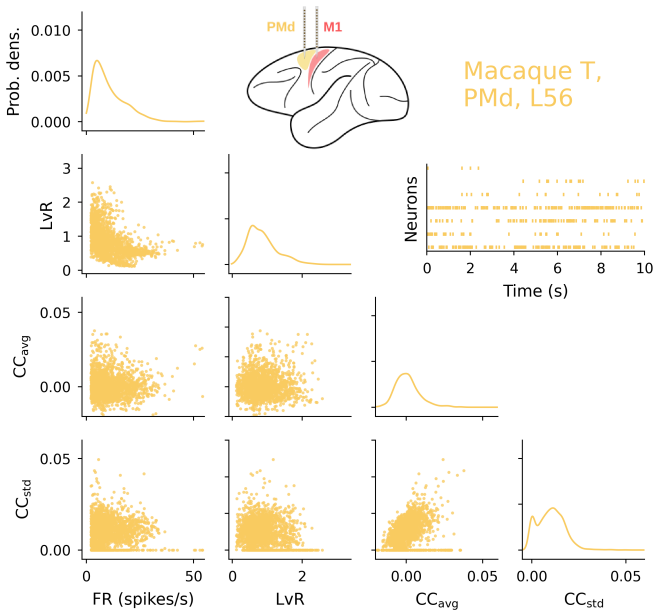


Fig. S6 Multi-dimensional summary statistics for macaque T, area PMd, L56. Sample spike trains shown.

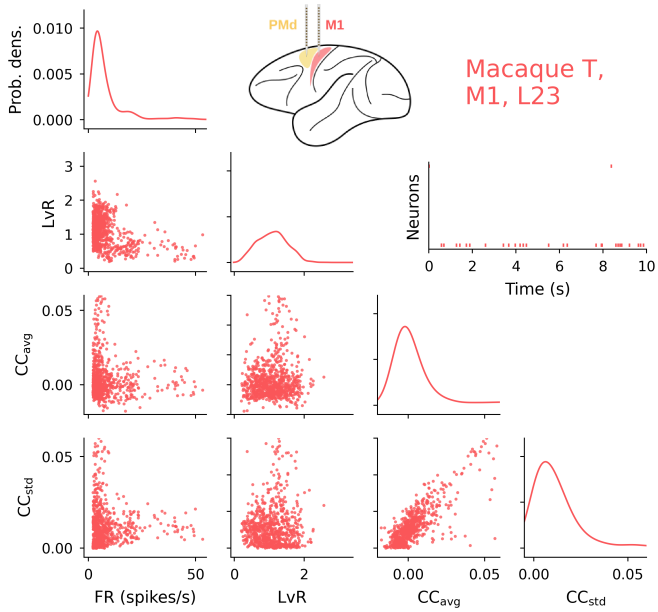


Fig. S7 Multi-dimensional summary statistics for macaque T, area M1, L23. Sample spike trains shown.

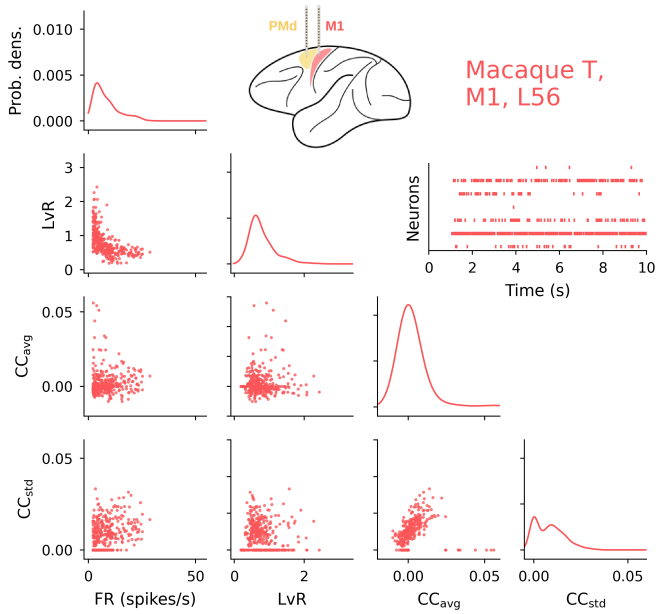


Fig. S8 Multi-dimensional summary statistics for macaque T, area M1, L56. Sample spike trains shown.

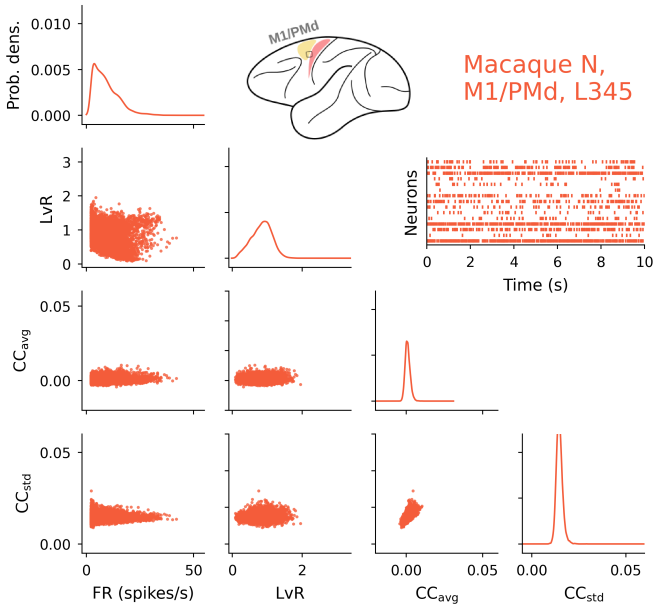


Fig. S9 Multi-dimensional summary statistics for macaque N, area M1/PMd, L345. Sample spike trains shown.

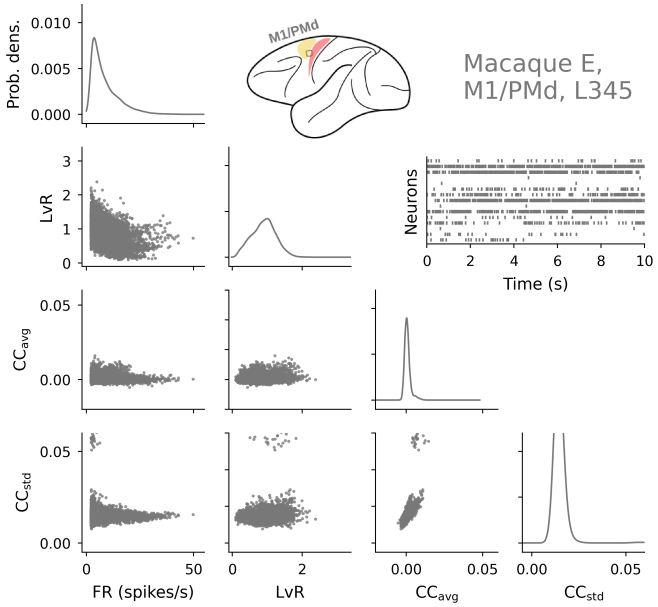


Fig. S10 Multi-dimensional summary statistics for macaque E, area M1/PMd, L345. Sample spike trains shown.

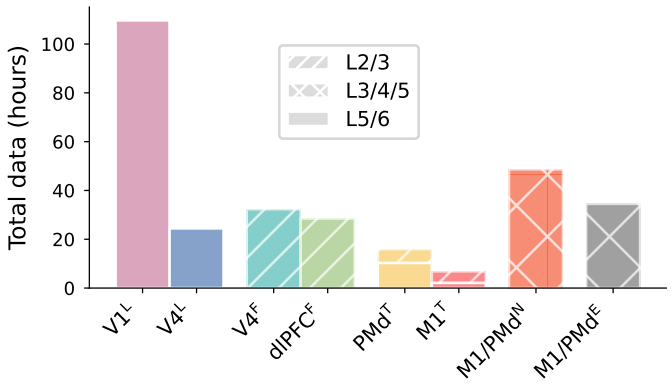


Fig. S11 Total amount of available data as the sum of all data slices, given in hours.

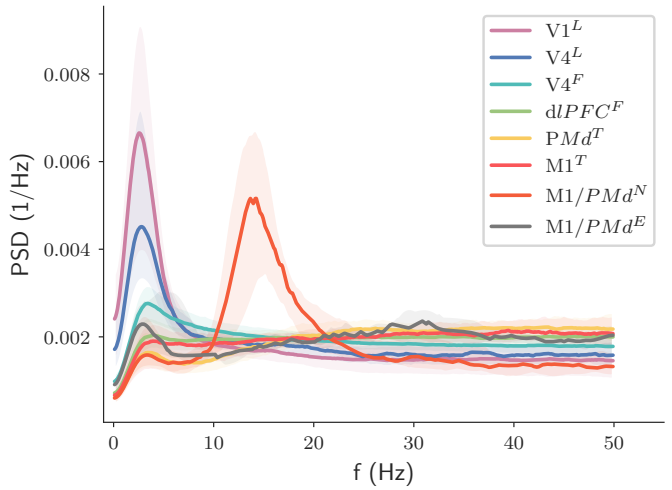


Fig. S12 Power spectral density (PSD) of the spike time histogram (bin size 10 ms) within the groups for t_{slice} = 20 s data slices. Median for each area (solid line) and the 5-95% percentiles (shading) shown.

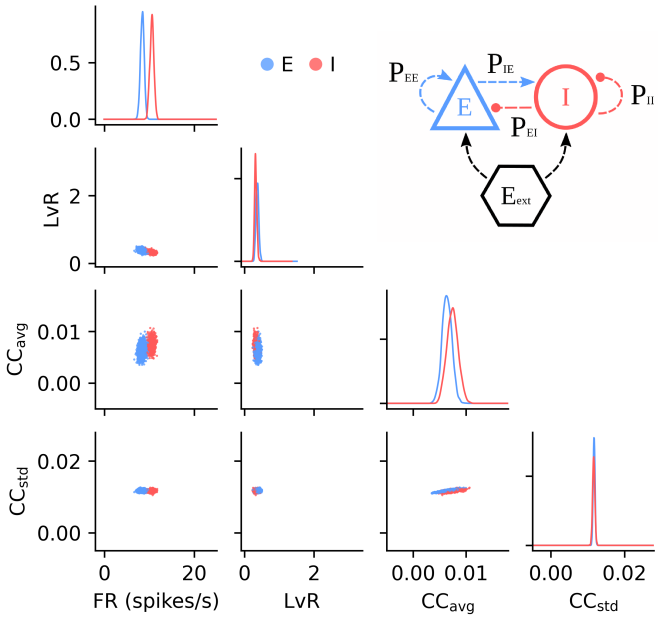


Fig. S13 Multi-dimensional summary statistics of the small balanced spiking neuron network with the target parameters.

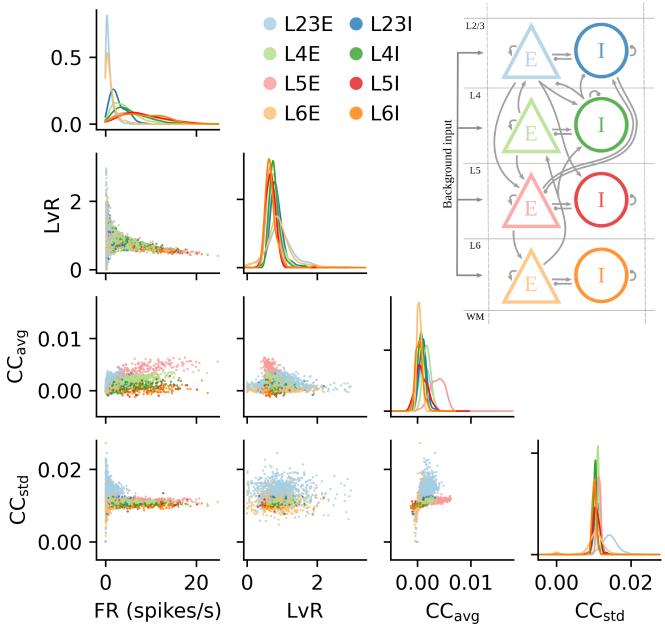


Fig. S14 Multi-dimensional summary statistics of the microcircuit model with the target parameters.

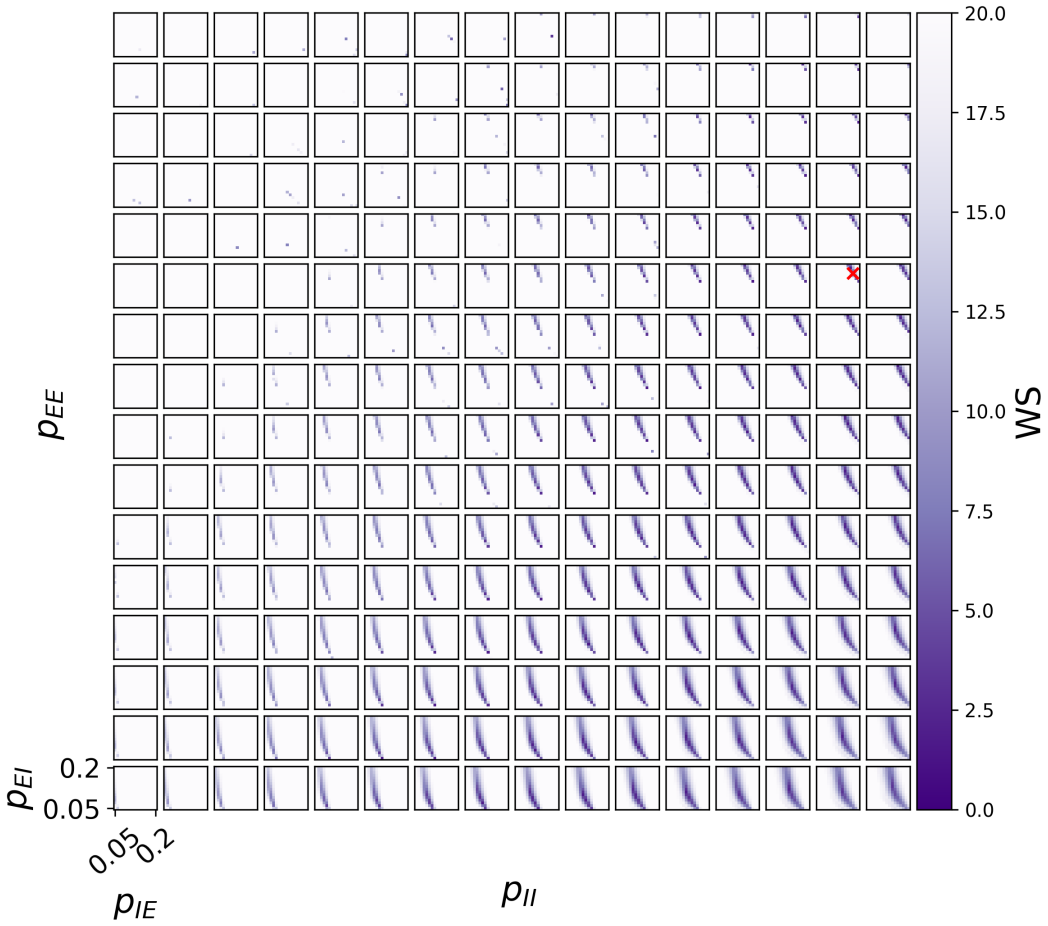


Fig. S15 Parameter scan of the Brunel model. WS fitness shown with respect to the reference model (indicated with a red cross) from all other connecivity parameter combinations. Parameters scanned at a 0.01 resolution in the $[0.05, 0.20]$ range. Panels indicate the P_{EE} and P_{II} parameters, increasing from left to right and frome the bottom to the top, respectively. P_{EI} and P_{IE} are represented within each panel.

Supplementary tables

Table S1 General model description of the small balanced spiking neuron model.

Model Summary	
Populations	two populations, one excitatory, one inhibitory
Connectivity	random connectivity
Neuron model	leaky integrate-and-fire model
Synapse model	exponential postsynaptic current
Input	independent spike trains from inhomogeneous Poisson processes with given rate ν
Neuron and synapse model	
Subthreshold dynamics	$\frac{dV}{dt} = -\frac{V}{\tau_m} + \frac{I_{syn}(t)}{C_m}$, $I_{syn}(t) = Je^{-(t-t^*-d)/\tau_{syn}} H(t-t^*-d)$
Spiking	If $V(t-) < \theta$ and $V(t+) \geq \theta$, 1. Set $t^* = t$ and $V(t) = V_0$, and 2. Emit spike with time stamp t^*.
Connectivity	
Type	pairwise Bernoulli, i.e. for each pair of neurons generate a synapse with probability p
Weights	fixed source and target population specific weights
Delays	log-normally distributed delays for excitatory and inhibitory neurons
Input	
Rate	Poisson noise with constant rate ν

Table S2 Simulation parameters of the small balanced spiking neuron model.

Population Parameters		
N_E	10000	number of excitatory neurons
N_I	2500	number of inhibitory neurons
Connectivity Parameters		
$\mathrm{P_{EE}}$	0.15	connection probability from $E \to E$
$\mathrm{P_{EI}}$	0.18	connection probability from $I \to E$
$\mathrm{P_{IE}}$	0.17	connection probability from $E \to I$
$\mathrm{P_{II}}$	0.19	connection probability from $I \to I$
Neuron parameters		
τ_{m}	20 ms	membrane time constant
τ_{r}	2 ms	absolute refractory period
τ_{syn}	2 ms	postsynaptic current time constant
C_{m}	1 pF	membrane capacity
V_{m}	0 mV	resting potential
E_{L}	0 mV	membrane capacity
V_{reset}	0 mV	reset membrane potential
V_{th}	20 mv	threshold
Stimulus parameters		
ν	spikes/s	base line rate
Synapse parameters		
J_{EE}	0.5	synaptic efficacy excitatory to excitatory
J_{IE}	0.75	synaptic efficacy excitatory to inhibitory
g	4	relative inhibitory synaptic efficacy
Delay parameters		
μ_E	0.29	mean of underlying normal distribution for excitatory connections
σ_E	0.22	standard deviation of underlying normal distribution for excitatory connections
μ_I	-0.40	mean of underlying normal distribution for inhibitory connections
σ_I	0.22	standard deviation of underlying normal distribution for inhibitory connections

Table S3 General model description of the microcircuit model.

Model Summary	
Populations	8 populations, one excitatory and one inhibitory for each layer L23, L4, L5 and L6
Connectivity	random connectivity
Neuron model	leaky integrate-and-fire model
Synapse model	exponential postsynaptic current
Input	independent spike trains from inhomogeneous Poisson processes with given rate ν
Neuron and synapse model	
Subthreshold dynamics	$\frac{dV}{dt} = -\frac{V}{\tau_m} + \frac{I_{syn}(t)}{C_m}$, $I_{syn}(t) = Je^{-(t-t^*-d)/\tau_{syn}} H(t-t^*-d)$
Spiking	If $V(t-) < \theta$ and $V(t+) \geq \theta$, 1. Set $t^* = t$ and $V(t) = V_0$, and 2. Emit spike with time stamp t^*.
Connectivity	
Type	pairwise Bernoulli, i.e. for each pair of neurons generate a synapse with probability p
Weights	fixed source and target population specific weights
Delays	log-normally distributed delays for excitatory and inhibitory neurons
Input	
Rate	Poisson noise with constant rate ν

Table S4 Population parameters of the microcircuit model.

Population Parameters								
L23E	L23I	L4E	L4I	L5E	L5I	L6E	L6I	population names
20683	5834	21915	5479	4850	1065	14395	2948	N number of neurons
1600	1500	2100	1900	2000	1900	2900	2100	K external synapses

Table S5 Neuron parameters of the microcircuit model.

Neuron parameters		
τ_{m}	20 ms	membrane time constant
τ_{r}	2 ms	absolute refractory period
τ_{syn}	2 ms	postsynaptic current time constant
C_{m}	1 pF	membrane capacity
V_{m}	0 mV	resting potential
E_{L}	0 mV	membrane capacity
V_{reset}	0 mV	reset membrane potential
V_{th}	20 mv	threshold
Stimulus parameters		
ν	8 spikes/s	base line rate
Synapse parameters		
J_{EE}	0.5	synaptic efficacy excitatory to excitatory
J_{IE}	0.75	synaptic efficacy excitatory to inhibitory
g	4	relative inhibitory synaptic efficacy
Delay parameters		
D_E	1.5 ± 0.75 ms	mean and std of underlying normal distribution for excitatory connections
D_I	0.75 ± 0.375 ms	mean and std of underlying normal distribution for inhibitory connections

Part III

Discussion

Knowledge on the structure and activity of the cerebral cortex is essential to understand its functions and underlying mechanisms, which we can do with the help of neuronal models. From small groups of neurons to brain-wide interactions, the underlying connectivity and cytoarchitecture determine which activity states are even possible. While the possible structural configurations and activity landscape are vast, incorporating as much information as possible into our brain models is bound to reduce the problem down to its fundamental components. The available information on the cortical structure is very diverse and does not necessarily reveal a unique and elegant blueprint of the brain. As the computational neuroscientist Grace Lindsay put it:

> *Biology took whatever route it needed to create functioning organisms, without regard to how understandable any part of it would be. It should be no surprise, then, to find that the brain is a mere hodgepodge of different components and mechanisms. That's all it needs to be to function.*
>
> — Lindsay (2021)

Thus, there is no guarantee that a unified grand theory of the brain can be achieved, akin to the elegant grand theories in physics. Therefore, we focus on describing the brain structure wherever possible such that the collective knowledge can be sensibly incorporated into unified models. These models will have to span multiple scales, from microscopic neurons, to mesoscopic brain areas and macroscopic brains. As it stands, if we want to build a realistic model, we have to put in all the bits an pieces from the bottom up.

This thesis contributes to the understanding of the brain by addressing some fundamental gaps in the knowledge. We identified several gaps (see Aim), such as the lack of open large-scale cortical recordings in macaque; the incomplete map of neuron densities and their distributions across the cortex; or precise estimates of intra-areal connectivity. These gaps are significant hurdles for constructing realistic brain models. These hurdles need to be addressed before informative models can be built, which will enable a more nuanced insight into the mechanisms and interactions that make up the brain.

The following sections discuss the results from the chapters in this thesis and how they answer the scientific questions that we formulated back in Part I. The discussion here focuses on an abstract level, for concrete discussion on specific results and methods see the discussion section within each individual chapter.

Brain activity in the macaque visual cortex

In Chapter 1 we addressed the question:

> What should we compare our brain simulations to? Can we bridge some of the gaps between experimental and theoretical neuroscience?

We can compare brain simulations to extracellular activity recordings. To answer the question, we have published a large new data set of macaque visual cortex (V1 and V4) extracellular recordings in the resting state. This data is freely and openly available to anyone willing to investigate it. This dataset was published in collaboration with the Netherlands institute for Neuroscience, where the recordings were made.

The data consist of 16 8x8 Utah arrays (Blackrock microsystems) for a total of 1024 electrodes. 14 arrays were implanted in V1, spanning about a fourth of the receptive field and 2 arrays were implanted in V4. Data sets without adequate and consistent metadata have very little value. Providing the relevant metadata to a dataset is essential for its reusability. Thus, we provided not only a huge dataset, but also all the relevant metadata about the experimental setup, experimental conditions, electrode and array locations, data alignment, electrode quality, receptive fields etc.

The dataset complies with the FAIR (findable, accessible, interoperable and reusable) principles for data sharing. The data can be found in a dedicated repository (Chen et al., 2021), it can be easily accessed both programatically and using the graphical user interface. Data and metadata can be operated in both Python and Matlab which run on all operating systems, and they can be easily reused thanks to the extensive metadata and documentation. The data were published under a Creative Commons Attribution 4.0 International Public License (CC-BY).

Most current biologically realistic models are not tuned to perform any particular task, instead they are used to investigate the structure of the nervous system and the dynamics that it generates. In this context, the traditional electrophysiology experiments, where the

macaques perform certain very specific tasks, are not very useful to generalist modelers. Neurons in the brain are also active in the absence of any particular task or stimulus. It is therefore essential to understand this underlying activity on top of which the all other complex dynamics occur.

In conclusion, we provide a new resource that can be used by the neuroscience community to study, validate, test and explore the brain activity. We adhere to the highest data sharing standards, with rich metadata and interoperable data formats. Modelers in general can now compare their brain simulations to this data, therefore slightly narrowing the gap between experimenters and modelers.

State space analysis of visual cortex activity

In Chapter 2 we addressed the question:

> Does feedback modulate the state space of the visual cortex? Can brain simulations help us understand this modulation?

We analyzed the data from Chapter 1, which suggests that feedback from V4 to V1 plays a role in modulating the state-space and the dimensionality of the V1 activity. This finding was verified with simulations in which we showed that feedback-like input could reproduce the experimental results. The interplay of experimental data and modeling helped us reach a more robust result that would not have been possible from the experimental data alone.

We studied the resting state data in V1, V4 and DP areas and found that the activity in the state space tends to occur within two separate state-space regions in V1 (but not in V4 or DP). We called these regions the clusters of the neural manifold and showed that they are strongly correlated to the behavioral state of the macaques; whether the macaque had its eyes open or closed. Within the clusters we also found that the intrinsic dimensionality—measured as a fraction of explained variance—is significantly higher during the eyes-open periods. Since the macaques sat in a dark room with no visual inputs, the differences in manifolds and dimensionality appeared to be modulated by some internal mechanism. We found that feedback from V4 to V1 was particularly stronger during the eyes-open periods. Finally, we observed similar changes in a computational model of spiking neurons with feedback-like input.

To the best of our knowledge, this is the first report suggesting cortico-cortical feedback could modulate neural manifolds *in vivo*. Nevertheless, previous evidence exists *in silico* for context dependent input discrimination (Naumann et al., 2022). We focused on the possible sources and explanations for this modulation in the discussion of Chapter 2.

Given the robustness of our observations we expect some benefit from such a state-space modulation. The eyes-closed state had higher low frequency power (alpha-band among

others, as is typically expected). If the low frequency oscillations are internally generated in V1 or as an interplay with thalamus, then the feedback from V4 during the eyes-open periods seems to disrupt them, leading to an asynchronous irregular state with a flatter spectrum. The mechanisms generating the alpha oscillations could still be present, but the sole presence of further higher frequency input can push the network into this state, which is exactly what we observed in our simulations.

But why force an asynchronous irregular state during eyes-open periods if there is no visual input? Maybe this state is better suited to rapidly respond to visual input, which can unexpectedly appear while the eyes are open, similar to what was shown by van Vreeswijk and Sompolinsky (1996). This idea is supported by the fact that the dimensionality of visual input also follows a power law decay, thus matching this power law in V1 may facilitate the information processing (Stringer et al., 2019). Further dedicated experiments would be needed to test whether the asynchronous irregular state is better suited to processing visual inputs than an oscillatory one.

In conclusion, we found two very distinct states in the V1 activity during the resting state. For the first time, we provide evidence that supports a modulatory role of $\mathrm{V4} \rightarrow \mathrm{V1}$ feedback. This findings are, to the best of our knowledge, the first *in vivo* evidence suggesting such a modulatory mechanism on a population level. Computational models enabled a cost-effective validation of our hypothesis, supporting our findings and encouraging further investigation of this phenomenon.

Lognormal distribution of neuron densities

In Chapter 3 we addressed the question:

> Do neuron densities follow a particular distribution across the cortex?

We demonstrated that neuron densities across cortical areas are compatible with a lognormal distribution. We show that this is also the case within the cortical areas of the marmoset cortex. In addition, we demonstrated that the lognormal distribution was not outperformed by any other statistical distribution with similar characteristics. We finally present a simple noisy cell division model showing the emergence of the lognormal neuron density. Our findings are in agreement for previous reports that numerous brain variables follow a long-tailed, potentially lognormal, distribution (Buzsáki and Mizuseki, 2014).

Our simple neurodevelopmental model predicted lognormal neuron densities for either across or within areas, but not both of them simultaneously. This could be overcome by incorporating the proliferation times of the neuron progenitor cells, because it can vary up to twofold across areas (Rakic, 2002). The progenitor cells also undergo a specification process, in which they turn into neurons, glia or more progenitor cells (Rakic, 2009; Cadwell et al., 2019). This process is bound to significantly affect the final distribution of neurons and glia, which we know to be lognormal-like. Some developmental models have already been proposed to incorporate these effects (Picco et al., 2018). Further work is thus needed to expand our model and incorporate all these effects, such that it can explain the simultaneous emergence of within and across cortical area lognormal distribution of neuron densities. Preliminary results with an extended model have shown some promise, but further work will be needed to ensure that the model predictions match the experimental observations.

In conclusion, we showed that the neuron densities are compatible with a lognormal distribution, within and across cortical areas. This finding can be used to validate further results, predict missing values in models or validate neurodevelopmental hypotheses.

Quantification of neuron density and white matter distance in macaque neocortex

In Chapter 4 we addressed the question:

How many neurons are there in each area and layer of macaque cortex?

We have answered the question by measuring the number of neurons within the macaque cortex with a semi-automatic workflow. Our measurements were based on nissl-stained cortical slices, which we reconstructed into a 3D volume and mapped to a standard parcellation. We then extracted counting strips and used an automatic classifier to count the number of neurons within stereological counting boxes. We validated our estimates against previous data (Beul et al., 2017) and by testing whether the neuron density was lognormal across and within cortical areas, as we saw in Chapter 3. In addition, we also measured the white matter distance between all cortical areas, extending a previously incomplete data set (Markov et al., 2012).

These measurements fill in a crucial gap in the knowledge, since the neuron densities per area and layer were previously only partially known in the macaque cortex. The number of neurons is a crucial parameter of spiking neuron models, thus measurements within a single individual will enable the construction of more reliable models. Furthermore, we expanded on the methods from previous studies in the marmoset (Atapour et al., 2019), making the neuron density estimation process easier than before, since we automatized some crucial steps. Thus, we expect our methods to be valuable to neuroscientists studying other animal species as well. Future work will still be needed to estimate the neuron densities for more than a single subject, at the time of writing data is openly available for three more subjects.

In conclusion, we measured a comprehensive data set of neuron densities and white matter distance in the macaque cortex, which will be valuable for modelers and anatomists alike.

Activity-driven estimation of local connectivity

In Chapter 5 we addressed the question:

> Can we estimate the connectivity within cortical areas from brain activity? How can we best compare simulated and experimental spiking neuron data?

In order to answer these questions we implemented an optimization workflow to estimate cortical connectivity parameters from brain activity. We first estimated summary statistics from the spiking data, which we assembled into a multi-dimensional cloud. The multi-dimensional summary statistics are unique to each area, suggesting that they are a fingerprint of the underlying cortical structure. Furthermore, we showed that we can indeed estimate the connectivity from a synthetic experiment, but only for a small model (Brunel, 2000). Further work is therefore still needed since we did not yet apply the parameter estimation method to the more biologically realistic microcircuit model (Potjans and Diesmann, 2014).

The major difference between our approach and previous parameter estimation methods is in the cost function. We decided to use spiking neuron statistics, which we assembled into a multi-dimensional summary statistics cloud, and then compared them to each other using the Wasserstein distance. In contrast, previous studies have focused on oscillatory activity (Prinz et al., 2004), multi-objective optimization of single neuron statistics (Druckmann, 2007), exact coincidences in spike times (Rossant, 2010; Ladenbauer et al., 2019), or aggregating the result from independently testing single neuron statistics (Carlson et al., 2014; Stringer et al., 2016). On the one hand, our approach incorporates the information from the co-variations between the different single neuron statistics, which all other approached ignore. On the other hand, we need all of the single neuron statistics to have the same dimensionality (in our case simple scalars) in order to assemble them into one

multi-dimensional cloud. Metrics like the power spectrum, correlation matrix or inter-spike interval distribution thus need to be reduced to scalars, which has the potential to introduce numerical errors.

Our model comparison approach can be applied to compare any two combination of simulations or experiments. For example, we used the multi-dimensional summary statistics and Wasserstein distance to compare experimental data from different areas. Another interesting application of this comparison scheme is comparing simulations when they are performed by different hardware—CPU, GPU or neuromorphic (SpiNNaker, Loihi etc.)—or software—NEST, NEURON, PyNN etc.—, thus the multi-dimensional summary statistics can help ensure that the outcome simulations are equivalent. Previous comparisons between CPU and SpiNNaker relied on exact spiking times (van Albada et al., 2018), which are very strict requirements considering how fundamentally different the underlying hardware is. Future work can use our comparison approach—implemented in NetworkUnit (Gutzen et al., 2018)—to estimate whether two simulations are statistically equivalent, with less concerns about minor numerical differences.

Finally, we also showed that our approach can be used to estimate spiking neuronal network parameters. We tested our methods using a small model (Brunel, 2000). However, we faced major scalability issues, since the parameter space grows by many orders of magnitude as the model size increases. In order to scale our methods up we will therefore need to find suitable regions of the parameter space *ad hoc*, e.g., by ensuring E-I balance or requiring that the total number of synapses lie in a specific range.

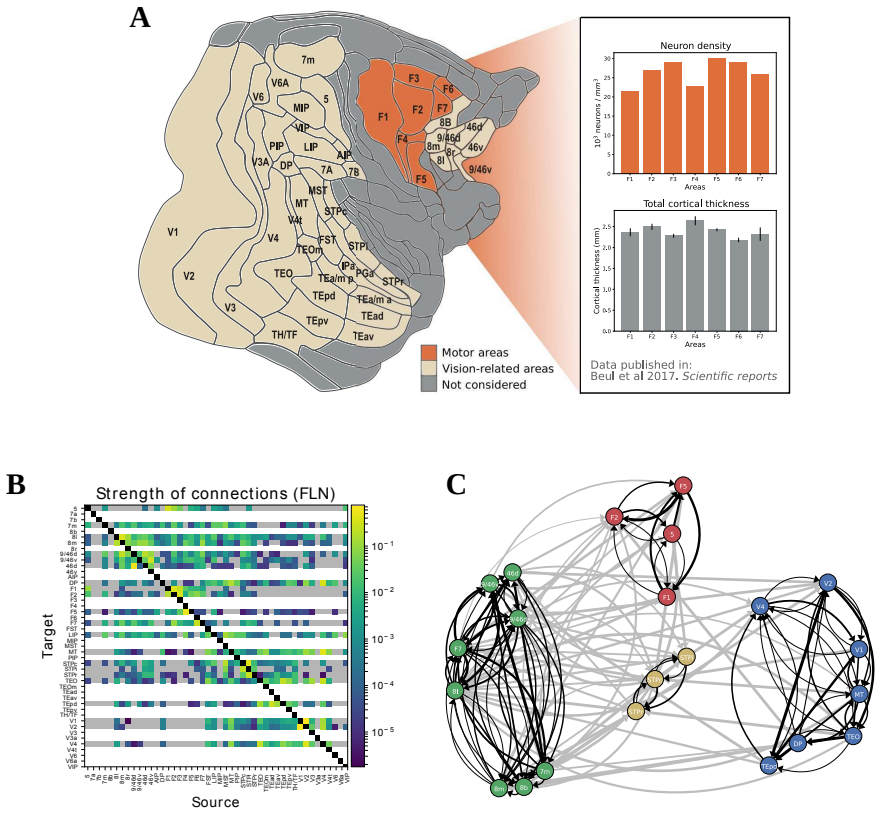
In conclusion, we provided a new method for the statistical comparison of spiking neuron data, which we used as a cost function in a parameter estimation problem. We showed that our methods can successfully estimate the connectivity of a small balanced spiking neuron model. Future work will focus on scaling to more biologically realistic models and testing whether the predicted connectivity is consistent with direct measurements of connectivity.

Outlook

This thesis has addressed some crucial aspects that will enable the future construction of biologically realistic large-scale neuronal network models. Such models will have to be validated against neuronal activity, which we have published and investigated for the visual cortex of the macaque (Chapter 1 and Chapter 2). Additionally, to construct realistic models one needs to know how the neurons are distributed across the cortex and how many of them there are in each cortical area and layer (Chapter 3 and Chapter 4). Another crucial aspect of cortical networks is the connectivity, which we will estimate from spiking activity for individual cortical areas (Chapter 5). Our methods for connectivity parameter estimation within the microcircuit could be applied to computationally inexpensive population models (Wilson and Cowan, 1972; Chaudhuri et al., 2015b; Mejias et al., 2016; Froudist-Walsh et al., 2021), which can then inform large-scale spiking neuron models.

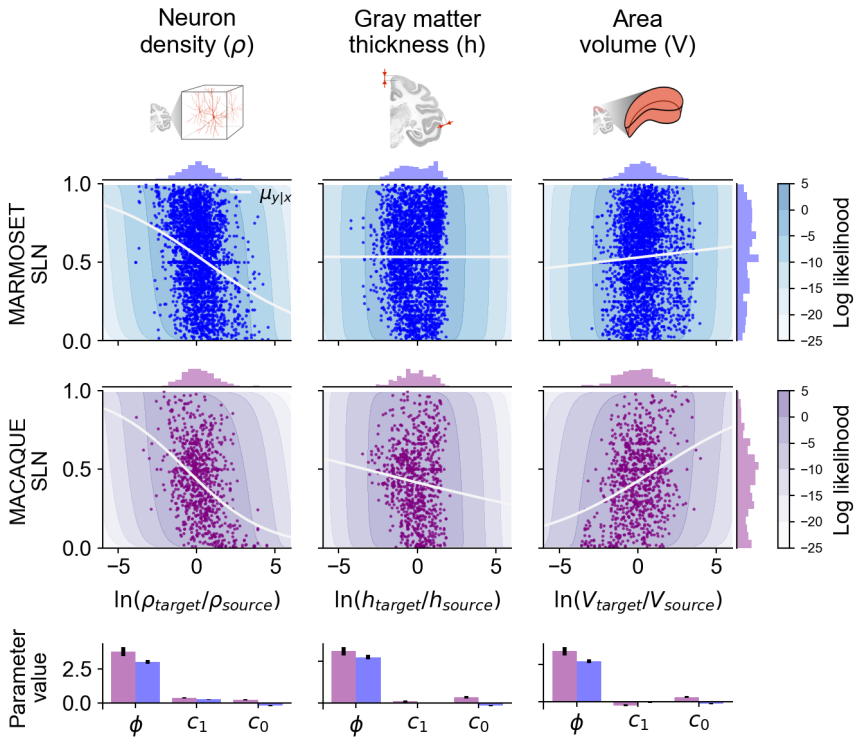
Previous efforts have already constructed a large-scale spiking neuron model of the macaque visual cortices (Schmidt et al., 2018a,b). This studies relied on long-range cortico-cortical connectivity from tract-tracing studies in macaque (Markov et al., 2012, 2014). The original 29 area data set of cortico-cortical connectivity was extended with the LIP area (Chaudhuri et al., 2015b) and later to 40 injection sites in (Froudist-Walsh et al., 2021). The missing connectivity can be predicted from the white matter distance between areas (which we estimated in Chapter 4), using the exponential distance rule (Ercsey-Ravasz et al., 2013), or cytoarchitectural types (Beul et al., 2017).

As a next step in large-scale modeling, we aim to construct a large-scale visuomotor model of macaque cortex (Outlook Figure 1) by extending the previous models (Schmidt et al., 2018a,b). For this aim, the advancements of this thesis will be instrumental, since the model will be able to incorporate reliable neuron density, connectivity and activity data. The available neuron density and connectivity data for marmoset also enables constructing marmoset specific models (Majka et al., 2016, 2018; Atapour et al., 2019), which could be constructed based on the same modeling framework.



Outlook Figure 1: Outlook on visuomotor multi-area model. **A** Overview of the areas that will be included in the model, with the motor cortex areas and their neuron density and grey matter thickness highlighted. **B** Cortico-cortical connection strengths from tract-tracing studies for the areas that will be included in the model (Markov et al., 2012). **C** Graph clustering using the heat equation.

One fundamental aspect of the long-range connectivity is the layer specificity of the connections. Markov et al. (2014) reported which fraction of presynaptic neurons was found above the granular layers in the source areas, which they call the supragranular labelled neurons (SLN). The SLN have been shown to relate to the hierarchy of the cortical areas, and to neuron density ratios (Markov et al., 2014; Schmidt et al., 2018a). Thus, such predictions were used to estimate the layer connection patterns in the previous large-scale models (Schmidt et al., 2018a,b). However, the bivariate distribution of SLN and neuron density ratios is skewed and heteroscedastic, thus classical regression models are unable to provide reliable predictions.



Outlook Figure 2: Structural predictors of laminar connectivity patterns in macaque (*Macaca mulatta*) and marmoset (*Callithrix jacchus*) neocortex. Fraction of supragranular labelled neurons (SLN) plotted against the log ratio of neuron density, gray matter thickness and area volume. The predicted log-likelihood is colour coded in the background and the conditional mean ($\mu_{y|x}$) of a fitted Bayesian beta-binomial model shown. Histogram of the marginals shown along the corresponding axis. Bottom row shows the parameters of the Bayesian beta-binomial model.

To overcome the limitations of skewed and heteroscedastic data, we developed a bivariate maximum likelihood regression Bayesian model, which provides a better statistical model for future connectivity predictions. We applied our new model to anatomical data from macaque monkey (*Macaca mulatta*) (Markov et al., 2012, 2014) and for the first time to the marmoset monkey (*Callithrix jacchus*) (Majka et al., 2016, 2018; Atapour et al., 2019). As a preliminary analysis, we studied the neuron density, gray matter thickness and area volume ratios (Outlook Figure 2). Further work in this direction will explore other structural variables, such as the white matter distance (Chapter 4), or the density of different neuronal receptors (Froudist-Walsh et al., 2021). Furthermore, the Bayesian model allows fitting to

higher dimensional data, thus enabling the simultaneous study of several structural variables in relation to the SLN. We expect this line of work will provide insight into the relation between brain connectivity and organization.

Future models constructed based on the connectivity predictions and using the anatomical data presented in this thesis can be validated against the resting state data (Chapter 1 and Chapter 5). The validation scheme could rely on the same procedure already presented in Chapter 5, comparing the multi-dimensional summary statistics of different areas against each other simultaneously. Constructing models that can successfully reproduce the resting state dynamics of the macaque cortex would provide a baseline upon which further hypothesis and complex tasks could be simulated and tested.

In conclusion, the work in this thesis constitutes an incremental improvement to the field of computational neuroscience. Our contributions will enable further work in modeling, data analysis and elsewhere. Especially given the attention to data sharing and open access that will ensure maximum outreach and dissemination of these findings. While our work has mainly focused on the macaque monkey, our methods can be applied to a range of other species; such as marmoset monkeys or humans. Thus, this work contributes to the overall understanding and characterization of the cortex; while the methods and data tested and analyzed will be valuable beyond this work.

Part IV

Appendices

A

From anatomy to models

Summary:

This appendix contains a book chapter, that describes different types of anatomical data and the techniques used to measure them; as well as pointing to existing datasets of interest. We include both direct measurement methods and indirect estimates, based on predictive relationships. This book chapter is an essential guide for computational neuroscientists aiming to construct biologically realistic neuronal models.

Bringing Anatomical Information into Neuronal Network Models

S.J. van Albada[1,2], A. Morales-Gregorio[1,3], T. Dickscheid[4], A. Goulas[5], R. Bakker[1,6], S. Bludau[4], G. Palm[1], C.-C. Hilgetag[5,7], and M. Diesmann[1,8,9]

[1]Institute of Neuroscience and Medicine (INM-6) Computational and Systems Neuroscience, Institute for Advanced Simulation (IAS-6) Theoretical Neuroscience, and JARA-Institut Brain Structure-Function Relationships (INM-10), Jülich Research Centre, Jülich, Germany
[2]Institute of Zoology, Faculty of Mathematics and Natural Sciences, University of Cologne, Germany
[3]RWTH Aachen University, Aachen, Germany
[4]Institute of Neuroscience and Medicine (INM-1) Structural and Functional Organisation of the Brain, Jülich Research Centre, Jülich, Germany
[5]Institute of Computational Neuroscience, University Medical Center Eppendorf, Hamburg, Germany
[6]Department of Neuroinformatics, Donders Centre for Neuroscience, Radboud University, Nijmegen, the Netherlands
[7]Department of Health Sciences, Boston University, Boston, USA
[8]Department of Psychiatry, Psychotherapy and Psychosomatics, School of Medicine, RWTH Aachen University, Aachen, Germany
[9]Department of Physics, Faculty 1, RWTH Aachen University, Aachen, Germany

Abstract For constructing neuronal network models computational neuroscientists have access to wide-ranging anatomical data that nevertheless tend to cover only a fraction of the parameters to be determined. Finding and interpreting the most relevant data, estimating missing values, and combining the data and estimates from various sources into a coherent whole is a daunting task. With this chapter we aim to provide guidance to modelers by describing the main types of anatomical data that may be useful for informing neuronal network models. We further discuss aspects of the underlying experimental techniques relevant to the interpretation of the data, list particularly comprehensive data sets, and describe methods for filling in the gaps in the experimental data. Such methods of 'predictive connectomics' estimate connectivity where the data are lacking based on statistical relationships with known quantities. It is instructive, and in certain cases necessary, to use organizational principles that link the plethora of data within a unifying framework where regularities of brain structure can be exploited to inform computational models. In addition, we touch upon the most prominent features of brain organization that are likely to influence predicted neuronal network dynamics, with a focus on the mammalian cerebral cortex. Given the still existing need for modelers to navigate a complex data landscape full of holes and stumbling blocks, it is vital that the field of neuroanatomy is moving toward increasingly systematic data collection, representation, and publication.

1 Introduction

Some of the defining characteristics of a neuronal network model are the size of the neuronal populations and the connectivity between the neurons. To determine these properties, the modeler has access to information in multiple forms and based on various experimental methods, where the completeness of the data varies widely across species and brain areas. For instance, the connectivity data for the nervous system of the nematode (roundworm) *C. Elegans* are nearly complete and have enabled full connectomes to be derived with minimal extrapolation from the data [1]. These graphs encode all connections between all of the neurons of the male and hermaphrodite worms. However, the 302 neurons of the hermaphrodite and the 385 neurons of the male worm pale in comparison to larger brains such as the human brain with its roughly 86 billion neurons and trillions of connections. Here, and for most species, measuring a full connectome is still far from feasible in terms of technical and computational effort. For this reason, the anatomical data often need to be complemented with statistical estimates in order to define complete network models of the brain. Filling in the gaps in the known connectivity in this way may be referred to as *predictive connectomics*. The corresponding predictions have to be validated in some way, for instance by leaving out part of the known anatomical data and determining how well these are reproduced by the statistical estimates.

Understanding the human brain is often considered the holy grail of neuroscience, not least because of the hope of finding novel cures and therapies for brain diseases. However, due to its size and enormous complexity, it can be helpful on the way to this goal to investigate simpler, more tractable brains of other species. Eric Kandel took this approach in his famous studies on the sea slug *Aplysia* [2], and it is a guiding thought behind the OpenWorm project on modeling *C. Elegans*. Furthermore, data obtained with invasive methods are, for obvious reasons, much more abundant for non-human brains. Of course, understanding the brains of species besides humans can be seen as a valuable aim in itself—for improving the well-being of animals, for inspiring industrial applications, or as an intellectual pursuit, like cosmology or paleontology, which enriches us culturally even if it has no direct practical application. And, as it is with all basic sciences, one never knows what innovations the knowledge gained may inspire many years into the future. For these reasons, we do not restrict ourselves to the human brain, but also consider various other species. However, we focus on mammalian brains, which exhibit qualitative similarity to the human brain and may therefore teach us most about our own brains. Non-human primate brains deserve particular attention, as they are closest to the human brain in terms of anatomy and function. Although extensive differences in detailed organization remain [3,4,5] the anatomical similarities and evolutionary path give hope that universal principles can be discovered extending to the human brain. Furthermore, the chapter has an emphasis on our study object of choice—the cerebral cortex.

To limit the scope of the chapter, we also restrict ourselves to anatomical properties relevant for networks of point neurons or neural populations, neglect-

ing most aspects of detailed neuron morphology and placement of synapses on the dendritic tree and axonal arborizations. The anatomical characteristics entering into the definition of such neural network models can be classified into brain morphology, cytoarchitecture, and structural connectivity. Brain morphology describes geometric macroanatomical properties, for instance the thickness of the cerebral cortex and its layers, or the curvature. Cytoarchitecture refers to the composition of brain regions in terms of the sizes, shapes, and densities of neurons. Structural connectivity refers to properties of the synaptic connections between neurons, including numbers of synapses between a given pair of neurons, or the probability for neurons from two given populations to be connected.

The type and level of detail of anatomical information that is required depends on the type and aim of the modeling study. A population model, describing only the aggregate activity of entire populations of neurons, does not require the connectivity to be resolved at the level of individual neurons, nor is it generally necessary to know the number of neurons in each population for such models. For models resolving individual neurons, in some cases it may be of interest to incorporate detailed connectivity patterns, while sometimes population-level connection probabilities suffice. The difference lies in the questions that the different types of models allow one to address. In one approach, the modeler tries to derive as realistic a connectivity matrix as possible, in the hope of obtaining the best possible predictions of dynamics and information processing on the anatomical substrate. Here, it always needs to be kept in mind that more detail does not necessarily mean better predictions: adding more parameters can actually reduce the predictive power of a model, for instance when these parameters are not sufficiently constrained [6,7]. However, if this approach is successful, it in principle allows the effects of detailed physiological parameter changes on network dynamics to be predicted (somewhat akin to weather forecasts), which may ultimately find clinical applications. In a contrasting modeling approach, connectivity features are abstracted and the influence of these abstract features (e.g., small-worldness, clustering, hierarchical organization, etc.) on graph theoretical, dynamical, or functional properties of the network are investigated. This approach places less emphasis on strict biological realism and attempts to provide a more conceptual understanding of the links between brain anatomy, dynamics, and function. In practice there is a continuum of approaches between these two extremes. For instance, models may incorporate biologically realistic features at an intermediate level of detail (e.g., population-specific connection probabilities without detailed connectivity at the single-neuron level) in order to simultaneously enable conceptual scientific conclusions and a degree of validation of these conclusions by direct model comparisons with experimental data.

Formulating and parametrizing neuronal network models is still often a painstaking effort, where the researcher digs through a vast literature to collect the relevant parameter values, from disparate experimental methods and labs. This systematization of the available knowledge into a common framework forms a central part of computational modeling work, and allows future researchers to continue at the next level of complexity. It is also highly specific to the modeling

problem and data modalities at hand, so that we cannot give one-size-fits-all advice on how to deal with and interpret anatomical data to develop network models. However, we can provide general guidance regarding what to look out for in the various data modalities, and how to incorporate the corresponding data into models. Furthermore, data are increasingly collected in systematic databases, which make the modeler's life easier by offering comprehensive data obtained with the same experimental methods, often even from the same lab. Most promising for facilitating this process are recent multilevel brain atlases, which aggregate both macro- and microstructural information into systematic anatomical reference frameworks.

In this chapter, we provide an overview of the types of anatomical information that can be used to define biological neural network models, point to available resources and databases, and describe methods for predicting connectivity and validating the predictions. The text considers physiological properties only where they relate directly to anatomy. This overview is intended as an aid for computational neuroscientists to develop accurate models of biological neuronal networks.

2 Brain morphology and cytoarchitecture

In this section, we describe the main types of information on the morphology and cytoarchitecture of brain regions, and corresponding resources available to modelers. We start by providing a brief introduction to brain atlases, which systematize information on these anatomical properties. Next, we treat the morphological property of cortical and laminar thicknesses in more detail. We then go into the determination of neural population sizes and the location of neurons within brain regions, and close with a short discussion of the use of morphology and cytoarchitecture in computational models. We do not distinguish between cell types within regions, as this would substantially extend the scope of the chapter, and, especially in the context of network models that do not resolve neural compartments, more directly concerns chemical and electrophysiological instead of anatomical properties.

2.1 Brain atlases

Brain atlases are a tool for defining brain areas and aggregating regional descriptions of the brain in a consistent anatomical framework. A brain atlas typically consists of a template space, a set of maps or a parcellation, and a taxonomy, which provides the names and mutual relationships of those regions.

The template space of a brain atlas is typically represented by one or multiple scans of a brain, which provide an anatomical description of an underlying standardized coordinate space. Depending on the task at hand, different template spaces are used. A classical template space for the human brain is Talairach space [8], which assumes that the relative distances between brain regions are

preserved between individuals, and defines a rescalable grid accordingly. Talairach coordinates are still in wide use in functional neuroimaging. Today, it is more common to use one of the MNI templates defined by the Montreal Neurological Institute [9,10], which include single- and multi-subject averages of MRI scans as volumetric standard spaces. While the MNI templates define standard spaces at millimeter resolution, the BigBrain offers a brain model of a single subject based on a three-dimensional reconstruction from 7,400 histological sections, at an isotropic resolution of 20 μm [11]. As the tissue sections were stained for cell bodies, this model provides the most detailed three-dimensional reference of human cytoarchitecture available today. Ongoing research addresses the three-dimensional cellular-level reconstruction of brains at 1 μm resolution, which poses considerable techical challenges for human brains due to their size and topological complexity [12].

Brain maps and parcellations assign brain regions to coordinates of a template space. In case of a standard whole-brain parcellation, each voxel has a unique region index, and the assigned regions do not overlap. In case of probabilistic maps, however, each coordinate is assigned a probability to belong to any of the regions, resulting in a set of overlapping maps to define the atlas. Parcellations are based on different modalities of brain organization, including cytoarchitecture (e.g. [13]), chemoarchitecture (spatial distribution patterns of molecules like specific neurotransmitter receptors, e.g. [14]), structural connectivity (patterns of connectivity with other brain regions as defined by axonal connections, e.g. [15,16]), functional connectivity (spatial co-activation patterns under different cognitive conditions (e.g. [17]), anatomical landmarks, or a combination of such features in the case of multimodal parcellations [18,19,20].

The gold standard of brain parcellations is based on cytoarchitecture as measured in histological sections. The early Brodmann atlas of the cerebral cortex of humans and other primates uses such a cytoarchitectonic parcellation [21]. Some years later, von Economo and Koskinas developed an atlas [22] with a more comprehensive characterization of the cortical layers, and taking into account cortical folding by describing cytoarchitecture orthogonal to the cortical surface. However, the bases of these pioneering works remain collections of separate brain slices, thereby lacking coverage of the full three-dimensional anatomical space, as well as of the variability across subjects. Recent work in probabilistic cytoarchitectonic mapping addresses the latter challenge by aggregating microscopic maps from ten different subjects in MNI space [13]. Furthermore, different groups are working on full three-dimensional, microscopic resolution maps of cytoarchitectonic areas [23] and cortical layers [24] in the BigBrain model, giving access to region- and layer specific measures of, e.g., cell densities and laminar thickness.

In connectivity-based parcellation, voxels with similar connection properties are grouped together [15]. An example of an atlas using connectivity-based parcellation is the human Brainnetome Atlas [16], which takes the Desikan-Killiany atlas based on cortical folds (the sulci and gyri) [25] as its starting point. The Brainnetome atlas has the advantage for modeling studies that data on functional connectivity, a term used in neuroscience for activity correlations, is freely

available in the same parcellation, allowing straightforward testing of model predictions on network dynamics.

The Allen Institute has published multiatlases of the developing[1] and adult human brain [26,27], mapping cytoarchitecture, gene expression, and for the adult brain also connectivity as measured with diffusion tensor imaging (DTI), a magnetic resonance imaging method that detects axon tracts. This multimodality, where different types of data are represented in the same template space and parcellation, is useful for modelers, not only because of the richness of the data, but also as mapping data from different sources between template spaces and parcellations introduces inevitable errors.

The macaque, as a close relative of humans, is an important model organism, for which several atlases have been created. These include the atlas of Markov et al. (2014) [28] with the so-called M132 parcellation of 91 cortical areas, and a whole-brain atlas by Calabrese et al. (2015) [29] based on DTI. Another commonly studied species is the mouse, for which state-of-the-art atlases of gene expression data [30], cytoarchitecture as measured with Nissl staining, which stains nucleic acids and thereby cell bodies of both neurons and glia, and mesoscopic connectivity obtained by anterograde viral tracing [31,32] are provided by the Allen Institute. Paxinos and Franklin provide the other most commonly used mouse brain atlas [33], which recent work combines with the Allen Institute coordinate framework [34].

Several online resources exist for browsing brain atlases. The Scalable Brain Atlas provides web-based access to a collection of atlases for the human brain and for a number of other mammals, including macaque, mouse, and rat [35]. The Human Brain Project provides online services for interactive exploration of atlases for the mouse, rat, and human brain through the EBRAINS infrastructure[2]. The human brain atlas is a multilevel framework based on probabilistic atlases of human cytoarchitecture, and includes links with maps of fiber bundles and functional activity, as well as a representation of the microscopic scale in the form of the BigBrain model with maps of cortical layers and cytoarchitectonic maps at full microscopic resolution [36].

2.2 Cortical and laminar thicknesses

The geometrical properties of the global and regional morphology of the brain have obvious relevance for brain models that explicitly represent space, but can also be important for estimating connectivity and numbers of neurons in non-spatial models. These properties include coordinates of region boundaries, spatial extents of brain regions, and properties of regional substructures such as thicknesses of cortical layers. Coordinates and spatial extents of brain regions are captured by atlases as described in the previous section. Another geometric property that is often of interest is the thickness of cortex and its layers.

[1] BrainSpan Atlas of the Developing Human Brain (2011) http://brainspan.org. Funded by ARRA Awards 1RC2MH089921-01, 1RC2MH090047-01, and 1RC2MH089929-01.

[2] https://ebrains.eu/services/atlases/brain-atlases

Cortical and laminar thicknesses can be either determined directly from histology of brain slices, or using structural MRI. When the MRI scans have sufficiently high resolution, these methods yield comparable results [37,38,39,40], but both methods have their own drawbacks. Brain slices generally represent sparse samples, are difficult to obtain precisely perpendicularly to the cortical sheet, and are subject to shrinkage, which has to be controlled for. Furthermore, identification of layers and the boundary between gray and white matter is still often performed manually, although automatic procedures are under development [24,41]. Structural MRI can cover the entire cortex and at least the gray/white matter boundary tends to be segmented using computer algorithms, but it has a lower resolution in the section plane than microscopy of brain slices, the exact resolution depending on the strength of the scanner and the scanning protocol. Von Economo provides laminar and total cortical thicknesses for all areas of human cortex based on 25 μm sections [42]. More recently, cortical and laminar thicknesses (the thicknesses of the individual cortical layers) have been identified in the BigBrain, forming a state-of-the-art, comprehensive dataset on human cortex [24,43]. The gray and white matter volumes and surfaces, along with the layer surfaces, are freely available[3] and can be explored interactively in the EBRAINS human brain atlas viewer. Alvarez et al. (2019) [44] determined the thicknesses of 25 human visual areas from 700 μm resolution MRI data from the Human Connectome Project, also making the quantitative area-averaged data freely available. Calabrese et al. (2015) [29] derived macaque cortical thicknesses from MRI scans at 75 μm resolution, available as an image file. Hilgetag et al. (2016) [45] provide total cortical thicknesses for 22 vision-related cortical areas of the macaque monkey, determined from brain slices sampled every $150-200\mu$m throughout the region of interest. At least in the vision-related areas of macaque cortex, total cortical thickness correlates inversely with neuron density, so that a statistical fit allows the thicknesses of the remaining vision-related areas to be estimated [46]. Correspondingly, cortical thickness varies systematically along the anterior-posterior axis in primates [47]. Rough estimates of the laminar thicknesses of macaque vision-related areas based on a survey of micrographs (microscopic images) have been published [46]. Comprehensive data on cortical thicknesses of other species are sparse, especially in a form that is directly usable by modelers. Methods for extracting cortical thicknesses from MRI in rodents are under development [48,49].

2.3 Numbers of neurons

Another basic property of brain circuits is their numbers of neurons, which can be determined from the size of brain regions and their neuron density. Over the years, different methods of counting cells have been used [50,51]. When total cell counts are of interest and their precise distribution across space is less important, tissue can simply be homogenized and the numbers of cell nuclei suspended in a fluid can be counted in samples under a microscope. The isotropic fractionator

[3] ftp://bigbrain.loris.ca/

is a version of such a homogenization and direct counting method [52]. The term 'fractionator' refers to a uniform random sampling scheme which divides samples into 'fractions' or counting boxes, enabling a statistical estimate of total cell counts to be obtained by considering only some fractions [53].

Stereological methods are a more involved class of methods that determine three-dimensional properties from two-dimensional sections through the tissue. The advantage of these methods is that the cells are counted in their real three-dimensional environment (depending on the section thickness) and thus spatial and area-specific values can be collected, e.g. cell densities in a single cortical lamina. Beside the fact that most stereological methods are quite labor- and time-intensive, the problem arises that the same cell may appear in two or more sections but should only be counted once. The disector addresses this issue by considering pairs of adjacent sections and only counting the cells that are present in the second but not the first section, effectively counting only the 'tops' [54]. The success of this approach depends on being able to recognize if features in the adjacent sections belong to the same cell, and on effectively correcting for large structures that extend across more than two sections. The optical fractionator combines the aforementioned uniform sampling method (the 'fractionator') with optical disection, in which objective lenses with a high numerical aperture are used to focus through the tissue to identify individual cells. A guarding zone above and below the inspected volume prevents multiple counting of truncated structures.

For cell bodies to be identified under the microscope, they are first dyed. Two commonly used methods are the aforementioned Nissl staining, and antibody staining of the protein NeuN that is present in the nuclei of most vertebrate neurons but not in glia [55]. Another technique dying both neurons and glia is silver staining [56], used for instance in the BigBrain model.

A number of comprehensive data sets on cell and neuron counts are available, although estimates can vary quite a bit across studies [57]. Overall numbers of neuronal and non-neuronal cells have been estimated for the brain as a whole, and for its major components like the cerebral cortex and the cerebellum, for a large number of species[4] [58,59,60,61,62]. In most cases, these cell numbers were acquired using the above-described techniques based on homogenized tissue. The von Economo atlas contains cell densities for human cortex with areal and laminar resolution, as determined with Nissl staining [42]. Because the Nissl technique stains both neurons and glia, which can, however, be distinguished based on morphology, it is not entirely clear whether glia are included in these cell densities. Furthermore, the cell numbers were measured without modern stereological approaches and without characterizing inter-individual variability. Modern high-performance computing methods are being applied for image registration of two-dimensional cortical and subcortical images to determine three-dimensional cell distributions [12] (figure 1), laying the foundation for future quantitative data sets representing an update and refinement with respect to the von Economo study. Collins et al. [63] provide cortical area-specific neuron

[4] https://en.wikipedia.org/wiki/List_of_animals_by_number_of_neurons

densities for the non-human primates galago, owl monkey, macaque, and baboon as determined with the isotropic fractionator. So-called cortical types or architectural types characterize the neuron density and laminar differentiation of primate cortical areas in a discretized manner, and thereby enable rough neuron density estimates where these have not been directly measured [64,65,46,66,67]. Herculano-Houzel et al. (2013) [68] measured neuron and cell counts and densities for the areas of mouse isocortex. Keller et al. (2018) [69] systematically reviewed region-specific neuron and glial densities throughout the mouse brain. Structures that have been characterized in detail also include the somatosensory areas of rat cortex and thalamus [70,71]. Despite many more data having been published, a large number of species-specific brain region compositions are still unknown, especially for subcortical regions. Scaling laws across species enable numbers of neurons to be estimated based on structural properties like brain and regional mass and volume [72,58,59,60,61,62].

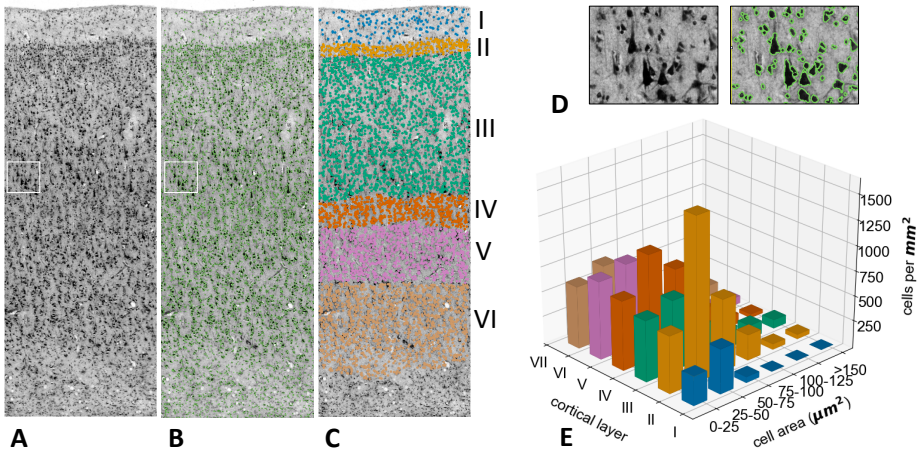


Figure 1. Extraction of layer-specific cell density estimates from microscopic scans of histological sections stained for cell bodies. **A.** Cortical patch of a scan. **B.** Example result of automatic instance segmentation of cell bodies using state-of-the-art image analysis (E. Upschulte, Institute of Neuroscience and Medicine, Forschungszentrum Jülich). **C.** Centroids of detected cell bodies, colored by cortical layer. **D.** Zoom into the local region of interest indicated by the white rectangle in Panels A and B. **E.** Two-dimensional histogram showing the number of cells in each layer, grouped by area of the cell body as segmented in the image.

Neuron counts or densities may not always be available in the particular parcellation chosen by the modeler. A mapping between parcellations may be performed by determining the overlaps between areas in different parcellations, for which the parcellations have to be in the same reference space. A large number of methods for registering images to the same reference space using nonlinear

deformations have been developed [73,74,75]. For macaque atlases registered to the so-called F99 surface, a tool provided alongside the CoCoMac database on macaque brain connectivity[5] [76] calculates the absolute and relative overlaps between cortical areas. The data in the new parcellation can then be computed as a weighted sum over the contributions from the areas in the original parcellation. However, this method entails the assumption that the anatomical data for each given area are representative of that area as a whole, and neglects inhomogeneities within areas. It should further be noted that criteria for area definitions, such as their cytoarchitecture or connectivity, are likely to provide information beyond this purely spatial approach. Nonlinear image registration techniques can take such factors into account, or alternatively, a coordinate-independent mapping can be performed [77]. No perfect solution for mapping anatomical data between parcellations exists, but in general, the more criteria are considered, the better the mapping.

2.4 Local variations in cytoarchitecture

Even within brain regions, cell densities are not constant but display local variations. An example of known spatial organization of neuron positions are so-called cortical minicolumns, also known as microcolumns, arrangements of on the order of 100 neurons perpendicular to the cortical surface, across the cortical layers. Cortical macrocolumns or hypercolumns are millimeter-scale structures containing thousands or tens of thousands of neurons with similar response properties in one or a few coding dimensions, for instance ocular dominance or position in the visual field. Cortical macrocolumns are particularly pronounced in the barrel cortex of rodents, which encodes whisker movements. In barrel cortex, the 'barrels' are cylindrical structures in layer IV containing neurons that respond preferentially to a particular whisker and have response properties and connectivity distinct from the interbarrel regions.

Various data on variations in neuron density within brain regions are available. Probably the most comprehensive data set of three-dimensional cell distributions is the Allen Mouse Brain Atlas, which contains both neurons and glia [57]. Spatial gradients in retinal cell densities have been well characterized [78,79,80,81,82,83], and those in thalamus to a lesser extent (e.g., [84]). The vertical distribution of cells in several cortical areas has also been characterized at a spatial resolution beyond that of cortical layers [85,86,87].

Studies resolving small cortical patches provide a sense of the variability of neuron density across the cortical sheet within primate cortical areas [63,88]. Furthermore, many studies have subdivided brain regions into discrete components with different cellular compositions, e.g., [89,90,91,92].

2.5 Use of morphology and cytoarchitecture in models

While most neural network models specify their architecture using concepts such as areas and layers, in some cases the neurons are simply assigned positions in

[5] http://cocomac.g-node.org/services/f99_region_overlap.php

continuous three-dimensional space and the connectivity is specified without reference to such concepts (e.g. [93]). In the conceptual approach, different connectomes may be obtained depending on the chosen parcellation. The particular choice of parcellation for instance affects topological properties of the corresponding connectomes [94,95]. Apart from this 'gerrymandering' issue, when predictive connectomics is used to fill in gaps in connectivity data with the conceptual approach, the choice of parcellation may influence the results. The findings of [94,95] for instance imply that incomplete connectomes completed via topological rules could differ depending on the parcellation. In view of the variability induced by differences between parcellations, there is something to be said for the continuum approach when the data allow it. Interpretation of the network dynamics in terms of region-specific activity may then be done in a post-hoc manner, flexibly with regard to the region definitions.

In spatially extended models, the neurons may be placed on a regular grid, with some jittering, at random positions, or at precise coordinates in space. Here, artificial symmetries in the network dynamics due to grid-like placement of neurons, which may arise for instance when the connectivity and delays are directly determined by the distances between neurons, should be avoided. Besides informing connectivity, the positions can be important for predicting signals with spatial dependence, like the local field potential (LFP), electroencephalogram (EEG) or magnetoencephalogram (MEG).

Precise region shapes are so far hardly used in computational modeling. Rather, the relatively rare network models that take into account three-dimensional structure tend to restrict themselves to simple geometric shapes like cubes or cylinders. An available but not yet widely used tool enables three-dimensional region volumes to be modeled through a combination of deformable two-dimensional sheets, where atlas data or histological images can support the modeling process via integration with the software Blender [96]. In an example application, the three-dimensional shape of the hippocampus was shown to substantially affect the connectivity between neurons predicted based on their distance. Accurate representations of volume transmission effects such as ephaptic coupling (non-synaptic communication via electrical fields or ions) [97], as well as the prediction of meso- and macroscopic signals like the LFP, EEG, and MEG also rely on the spatial distribution of neurons and thus benefit from measured three-dimensional brain morphology [98,99,100].

On the scale of local microcircuits on the order of a millimeter, spatial variations in cortical and laminar thicknesses across the cortical sheet within each area are limited and are generally ignored in computational models. Cortical and laminar thicknesses are then straightforwardly incorporated by scaling the numbers of neurons accordingly, and sometimes by distributing the neurons across cortical depth. In future, as resources become available for modeling extended cortical regions in detail, continuous variations in cortical and laminar thicknesses may be incorporated.

It is also not yet common for computational models to take into account continuous variations in neuron density within brain regions. However, a number

of models already divide regions into discrete subdivisions with different cellular compositions, e.g., [101]. The organization of cortex into minicolumns and macrocolums has been incorporated for instance in models of attractor memory [102,103] motivated by a functional interpretation. In future, increasingly realistic placement of neurons in models may yield more sophisticated predictions of spatially resolved brain signals and of network dynamics, through associated properties like distance-dependent connectivity.

3 Structural connectivity

Neurons in the brain exchange chemical signals via synapses, and in some cases are in more direct contact via so-called gap junctions. Although gap junctions are probably important for some phenomena (e.g. [104]), we here focus on the former, much more numerous type of connections, the synapses. The huge number of synapses in mammalian brains has so far precluded mapping all of them individually, although efforts are underway towards dense reconstruction of the mouse brain [105]. However, various methods exist for measuring neuronal connectivity, at scales ranging from individual synapses to entire axon bundles between areas. While some models distinguish individual synapses and thus need information at this level, other models lump synapses together, so that aggregated connectivity information suffices.

This section provides an overview over available types of information on neuronal network connectivity, along with resources and databases that can be used for constructing neuronal network models. We describe connectivity information according to the major experimental methods: microscopy, paired recordings, glutamate uncaging, axonal tracing, and diffusion magnetic resonance imaging (diffusion MRI), of which the most commonly used form is diffusion tensor imaging (DTI).

3.1 Microscopy

The oldest and lowest-resolution form of microscopy is light microscopy, providing a magnification factor of up to about 1,000. Neuron reconstructions from light microscopy of adjacent tissue slices allow rough estimates of connectivity based on the proximity of pre- and postsynaptic neural processes (cf. section 4.1). Following this approach, Binzegger et al. (2004) [106] derived a population-level local connectivity map for cat primary visual cortex. However, as detailed in section 4.1, predicting connectivity based on proximity has its drawbacks, which should be kept in mind when interpreting the resulting connectomes. Furthermore, tissue slicing cuts off dendrites and axons, which may extend over millimeters and more, so that assessing medium- to long-range connectivity requires extensive three-dimensional reconstructions. A method that facilitates such reconstructions is block-face tomography, in which scanning of the surface of a tissue block is alternated with the removal of thin slices from the surface [107].

Two-photon microscopy is a sub-micron resolution imaging technique that uses laser irradiation of tissue to elicit fluorescence through two-photon excitation of molecules [108]. A high-throughput block-face tomography pipeline has enabled the reconstruction of the full morphologies of 1,000 projection neurons in the mouse brain at a resolution of $0.3 \times 0.3 \times 1\,\mu\text{m}^3$, the MouseLight data set of Janelia Research Campus [109,110]. A viewer for the MouseLight morphologies is available[6]. A finding that stands out from this data set is the remarkable variability in projection patterns, each neuron projecting to a different subset of target regions for the given source region.

At nanometer spatial scales, electron microscopy enables the identification of individual synapses and the precise shape and size of the presynaptic and postsynaptic elements, even down to individual synaptic vesicles. This method is extremely labor-intensive, but heroic efforts have nevertheless led for instance to estimates of synapse density in different areas of human cortex [111,112] , a volume reconstruction of the entire Drosophila (fruit fly) brain [113], the morphological reconstruction of 1,009 neurons in a microcircuit of rat somatosensory cortex [71], and full reconstructions of $1{,}500\,\mu\text{m}^3$ [114] and more recently $> 5 \times 10^5\,\mu\text{m}^3$ [115] of mouse cortical tissue. A noteworthy finding from these studies is that the presence of synapses is not perfectly determined by the close proximity of axons and dendrites (appositions). For instance, an apposition is far more likely to predict an actual synaptic contact for pairs of neurons that also form synapses elsewhere on the axon and dendrite [114]. Such a rule will tend to lead to a long-tailed distribution of the multiplicity of synapses between pairs of neurons.

Synapses may look asymmetric or symmetric under the microscope, where asymmetric synapses have a pronounced postsynaptic density and are predominantly excitatory, while symmetric synapses have roughly equally thick pre- and postsynaptic densities and tend to be inhibitory. Both the size of synapses and their location on dendrites are informative about their effective strength in terms of postsynaptic potentials evoked at the soma [116,117,118,119]. Furthermore, synapse locations on dendrites can tell us something about their interaction with other synapses; however, these complex interactions are not captured by point neuron or population models. Axonal varicosities or boutons are swellings along axons (boutons en passant) or at axon terminals (terminal boutons) that host synapses, and which are detectable through all microscopic methods mentioned here. Even when the synapses themselves are not directly imaged, boutons may be taken as evidence for synapses, with the caveats that some synapses are not established on boutons, and individual boutons may contain different numbers of synapses [120].

In summary, microscopy is useful for estimating connectivity based on appositions, reliable estimates of numbers of synapses in a given volume, detailed connectivity features such as the multiplicity of synapses between pairs of neurons, and correlative information on synaptic efficacy.

[6] https://neuroinformatics.nl/HBP/mouselight-viewer/

3.2 Paired recordings

In paired recordings, electrodes are used to simultaneously stimulate one cell and measure the response in another cell, either *in vitro* or *in vivo*. Stimulation may be performed extracellularly, intracellularly with sharp electrodes, or via patch clamp; recordings normally use one of the latter two techniques. This method sums up the contributions from potentially multiple synapses between the pair of neurons, which should be kept in mind when incorporating the corresponding synaptic strengths into models. Where anatomy-based methods can have the drawback that they do not provide conclusive evidence for physiologically active synapses, paired recordings identify functional synapses. However, existing connections may be missed depending on the experimental protocol, for instance due to axons and dendrites being cut off during slice preparation. Each pair of neurons should also be tested multiple times, because in individual trials, axonal or synaptic transmission failures may occur, or the postsynaptic potential may be too small to be detectable among the noise [121]. Paired recordings may be biased toward neurons that are easier to patch or insert an electrode into, for instance larger cells. Especially *in vivo*, where the network exhibits background activity, responses may in principle be caused by activation of neurons other than the one that is stimulated. Responses are judged to be monosynaptic based on a short, consistent response latency, usually of a few tenths of milliseconds [122,123].

Most paired recordings are highly local, with a distance no greater than 100 μm between the somas of the pre- and postsynaptic cells. They provide the modeler with connection probabilities in terms of the fraction of pairs of neurons that have at least one synapse between them. For interpreting these connection probabilities, it is important to take into account the spatial range of the recordings, as connection probability is generally distance-dependent. The measurements represent a spatial average over this distance-dependent connectivity, which is in mathematical terms a double sum (which may in continuum approximation be represented by an integral) over the positions of the source and target neurons.

Paired recordings show that, on the scale of local microcircuits up to 200 μm from the presynaptic soma, bidirectional connections between pyramidal neurons in cortical layer V occur significantly more often than would be expected by chance [124]. In some studies, researchers have recorded from multiple neurons simultaneously [125,126,127,128]. Simultaneous recordings from respectively four [126] and twelve [127] rat cortical neurons confirm the overrepresentation of bidirectional connections regardless of the distance from the soma. This type of analysis has also revealed that motifs with clustered connections among three or more neurons are more common in the cerebral cortex than would be predicted based on pairwise connection probabilities alone [126,127] (cf. section 4.4).

3.3 Glutamate uncaging

Similarly to paired recordings, glutamate uncaging generates action potentials in presynaptic neurons and records the response in postsynaptic neurons con-

nected to them. Usually, the method is applied to slice preparations and neurons are recorded intracellularly, but *in vivo* application and extracellular recordings are also possible. First, a compound consisting of glutamate bound to another molecule is introduced, for instance by bathing a brain slice in a solution with the caged glutamate. Then glutamate is released by photolysis of the compound through focal light stimulation, causing action potentials in neurons with their soma close to the stimulation site. Brain slices are generally scanned systematically, generating for each given target neuron a grid-like map of response amplitudes for each stimulated location.

Originally, glutamate was uncaged using ultraviolet light [129], but due to light scattering and a large uncaging area, this stimulated multiple neurons, making the results harder to interpret. Two-photon stimulation, in which photolysis is triggered by the absorption of two photons, enables individual neurons and even individual dendritic spines to be stimulated [130,131]. As with paired recordings, an issue is that it cannot be known with certainty whether the responses are monosynaptic or emerge due to sequential activation of two or more neurons, but short-latency responses time-locked to presynaptic action potentials in the absence of background activity reliably indicate monosynaptic connections. Another issue is that the uncaged glutamate may directly influence the recorded neuron, so that stimulations that lead to short-latency responses with excessive amplitudes have to be excluded from analysis. Furthermore, the same caveats as for paired recordings apply with regard to distance dependence of connectivity, and potential cutting of dendrites and axons during slice preparation.

Purely based on glutamate uncaging response maps, it is not possible to directly derive a neuron-level connectivity map, because it is unknown how many different presynaptic neurons are activated across stimulation sites. However, by combining glutamate uncaging with imaging of the neurons, the connectivity between neurons can be determined [130]. In the absence of such direct imaging, the number of source neurons eliciting a given glutamate uncaging response can be estimated by dividing by the unitary synaptic strength (the PSP or PSC size due to a single presynaptic neuron), if an independent estimate for the latter is available. If one in addition makes an assumption about the average number of sites from which a given presynaptic neuron is activated, which depends on the resolution of the stimulation grid, this yields an estimate of the number of neurons impinging on a given postsynaptic cell. Typically, action potentials can be elicited in a given neuron from a handful of sites [132,133]. Finally, one can derive a connection probability by dividing by the approximate number of neurons in the stimulated volume. Clearly, many assumptions and approximations are involved in such derivations, so that it is currently still difficult to reliably determine the connectivity of neural network models from glutamate uncaging data. However, in some cases, data obtained by this method are the best available for a given brain region, in which case one may proceed via such assumptions [134].

3.4 Axonal tracing

The technique of axonal or neuroanatomical tracing entails injecting a tracer, which can be a molecule or virus, which is taken up by neurons and transported toward cell bodies or axon terminals. In anterograde tracing, the tracer is transported in the forward direction toward the synapses, while in retrograde tracing, it is transported in the backward direction from axons toward the cell bodies of the sending neurons. In practice, most tracers are to some extent both anterograde and retrograde, but one transport direction dominates [135]. Detection of the tracer happens in one of multiple ways: the tracer may itself be fluorescent, it may be radioactively tagged or conjugated with a dye or enzymatically active probe, or it may be detected via antibody binding [136]. Axonal tracing is generally performed in the living brain, after which the animal is sacrificed to detect where the tracer has ended up, but some substances also enable tracing in postmortem tissue and therefore even in the human brain, albeit over limited distances [137,138,139]. The method is well suited to characterizing medium-to-long-range connections such as those between cortical areas. A number of tracers, especially certain viral tracers, are transneuronal, crossing synapses and tracking polysynaptic pathways [140]. Furthermore, it is possible to perform double or even triple labeling to visualize the participation of neurons in two or more connection pathways [141]. Double labeling with retrograde tracers for instance suggests that the vast majority of cortico-cortical projection neurons in macaque visual cortex send connections either in the feedforward direction or in the feedback direction, not both, with respect to the hierarchy of visual areas [28].

Tracer injections typically cover a millimeter-scale area, so that multiple axons are traced at the same time, not individual ones. Because of the local spreading of the tracer, axonal tracing does not provide reliable information about the region immediately surrounding the injection site. An important drawback of the method is that only up to a few injections can be performed in each animal, so that data have to be combined across many animals to obtain a complete connectivity graph. This introduces inevitable inaccuracies due to inter-individual differences. Because tracers are taken up by neurons indiscriminately, conventional tracing does not allow the specific connections of separate subpopulations of neurons to be identified, let alone of individual neurons. However, over the past decades a number of viral tracing methods have been developed that trace specific molecularly marked neuronal subpopulations [136]. A modern technique uniquely labeling neurons with random RNA sequences enables high-throughput mapping of projections at the level of individual source neurons [142].

While axonal tracing traditionally only gave qualitative information about connectivity, for instance describing staining as sparse, moderate, or dense, more recently a number of groups have gone through the painstaking effort of counting the numbers of labeled cells in retrograde tracing experiments. A notable quantitative tracing data set characterizes the connectivity between a large number of areas in macaque cortex in terms of overall fractions of labeled neurons (FLN) and fractions of supragranular labeled neurons (SLN) in all source areas project-

ing to each injected target area [28,143]. SLN relates to the hierarchy of vision-related cortical areas, as feedforward projections tend to emanate from layer II/III and thus have a high SLN, while feedback projections emanate preferentially from infragranular layers and have a low SLN. A similarly comprehensive resource of quantitative retrograde tracing data is available for the marmoset neocortex [144,145].

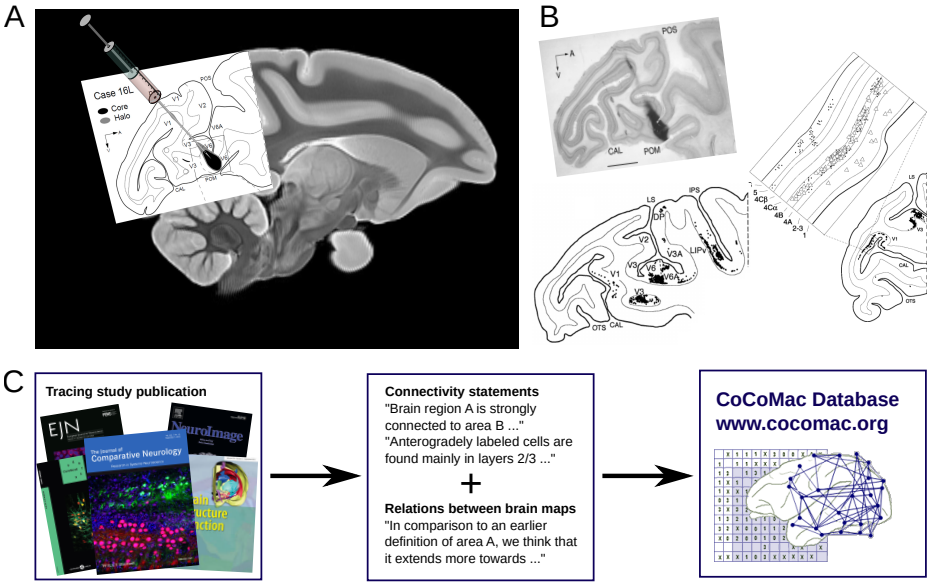


Figure 2. Workflow from tracing experiment to entry in the CoCoMac database. **A.** A tracing study is performed to study a particular part of the brain, by injecting a tracer substance into a target area. Shown is Case 16L from Galletti et al. [146], here registered to the macaque atlas of Calabrese et al. [29] via the Scalable Brain Atlas [35]. **B.** Tracer is picked up by axons, and depending on the substance it is either transported anterogradely towards the axon terminals, or retrogradely to the cell bodies, or both. After sacrificing the animal, a careful investigation of labeled cell bodies and/or axon terminals across the brain is carried out, sometimes including layer-specific quantitative data. **C.** After the results have been written up and subject to peer review, collators from the CoCoMac database take out statements on connectivity and the definitions of brain areas.

The CoCoMac database, which stands for Collation of Connectivity data on the Macaque brain [76,147], contains both anterograde and retrograde tracing data from a large number of published studies, especially for the cerebral cortex, in part with laminar resolution. Figure 2 illustrates the prerequisites for creating such a database. Another collation effort [148] has reconstructed the area-level structural connectome of the cat from qualitative axonal tracing data. The Allen Institute provides an anterograde tracing data set encompassing hundreds of

injections throughout the mouse brain [149]. A comprehensive characterization of laminar target patterns of connections between cortical areas in primate is missing to date.

Axonal tracing is a reliable method for identifying actual connection pathways, and often serves as the ground truth for evaluating diffusion tensor imaging results (cf. section 3.5). However, the fact that connectomes based on tracing data are a composite of connectivity in many individuals warrants special caution in their interpretation. The average or union of the connections in many brains in all likelihood does not accurately represent the connectivity of any individual brain.

3.5 Diffusion tensor imaging (DTI)

Diffusion tensor imaging (DTI) is a form of diffusion MRI or diffusion-weighted imaging (DWI), which measures the local rate of water diffusion at a resolution of typically a few millimeters. DTI detects anisotropies in the diffusion of water by using several different orientations of the magnetic field gradients to obtain information about the directionality of the diffusion in each voxel [150]. Since the diffusion is greater along than perpendicular to myelinated axons, the method enables the main local orientation of axonal fiber tracts to be identified. The paths of the fiber tracts maximally consistent with the local orientations are reconstructed using so-called tractography. The density of these 'streamlines' is a measure of connectivity between distant brain regions, and can for instance be summed within cortical areas to obtain an area-level cortical connectivity map. DTI is non-invasive and can reveal the connectivity of the whole brain at once. However, apart from possible directional specificity introduced by the choice of seed points for tractography, the connectivity provided by DTI is symmetric, as it can resolve the orientation but not the direction of fiber tracts. While most cortical inter-area projections are reciprocal with positively correlated connection density in the two directions [143,151,145], a proportion of connections is asymmetric, and these asymmetries are hereby missed. Such asymmetries are likely to be important for the dynamics predicted from neuronal network models [152]. Further drawbacks of DTI are its lack of laminar resolution and its inability to distinguish fibers with different orientations in the same voxel, such as crossing or touching ('kissing') fibers. Local tractographic errors due to kissing or crossing fibers add up over distance, limiting the reliability of the resulting connectivity maps, especially giving many false positives for long-distance connections [153].

The Human Connectome Project provides high-resolution preprocessed human diffusion MRI data for > 1100 subjects. Tractography was performed on an earlier, smaller data set from the Human Connectome Project and the resulting connectome was made available via the Brainnetome Atlas [16]. Prominent DTI connectomes for the macaque and mouse brains were published by Duke University [29,154].

As yet, there is no straightforward way to derive fully reliable and accurate connectomes from DTI. The same holds more generally for all the types of connectivity information we have discussed. All experimental connectivity data

have 'gaps': they only cover a certain spatial scale, they represent a subsample or lack precision at the given scale, or additional information is required to turn the experimental values into numbers of synapses. For this reason, methods are needed for filling in the gaps in the data in order to fully specify network models. This is the topic of the next section.

4 Predictive connectomics

Where the experimental connectivity data have gaps, we can try to fill these in using statistical estimates based on relationships of the known connectivity with properties such as cytoarchitecture or distance between brain regions. We refer to this approach as 'predictive connectomics'. Such statistical estimates still tend to have a high degree of uncertainty associated with them, but if we want to fully define a network model, there is no way around making certain assumptions and approximations. From another perspective, the statements of predictive connectomics represent formalized hypotheses for further anatomical studies. The spatial and temporal organization of neurodevelopment simultaneously explains many empirical relationships between connectivity and other structural properties of the brain. In the present section, we discuss the major heuristics for predicting connectivity, including Peters' rule, architectural principles, and methods based on distance and network topology, and describe how developmental origins form a common denominator for many of these heuristics. Finally, we touch upon the inference of structural connectivity from activity data.

4.1 Peters' rule

Peters' rule postulates that proximity between neurites (i.e. presynaptic axons and postsynaptic dendrites) can predict neuronal connectivity. It was originally proposed by Peters and Feldman (1976) [155] for the projections from the lateral geniculate nucleus to the visual cortex of the rat. The term 'Peters' rule' was later coined by Braitenberg and Schüz (1991) [156], who also generalized this idea beyond the particular case studied by Peters and Feldman. The rule has since been widely used by researchers. Over time its application has varied. Rees et al. (2017) [157] reviewed the relevant literature and distinguished between three conceptually different usages of the rule, which correspond to increasing level of detail (illustrated in figure 3):

1. Population level. In the original formulation, the rule was applied as a predictor of connectivity between populations of neurons of the same type. Consider a group of neurons A (for example in the thalamus) projecting to a region containing another group B (for example pyramidal cells in visual cortex), where all neurons within the groups are of the same type. According to the original rule, the number of synapses between A and B is correlated with the spatial overlap of presynaptic axons of population A and postsynaptic dendrites of population B.

2. Single-neuron level. Extending the example from the previous point, take two neurons a_i and b_j from populations A and B, respectively. In this formulation, the probability p_{ij} for a connection between a_i and b_j to exist is proportional to the spatial proximity between their respective pre- and postsynaptic arbors.
3. Subcellular level. At the subcellular level, Peters' rule has been used to link the number of axonal-dendritic appositions to the number of synapses, regardless of cell types.

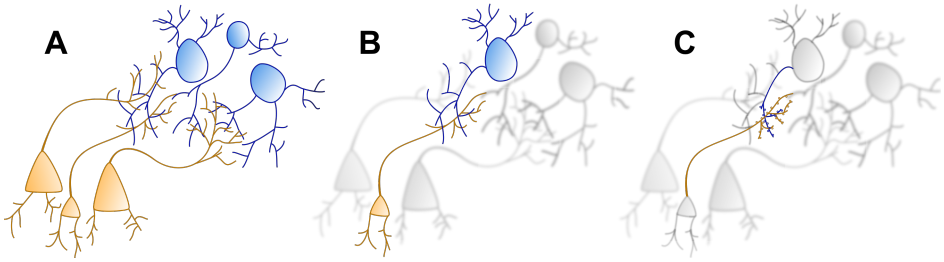


Figure 3. Illustration of the different levels of detail in the usage of Peters' rule, as described in [157]. **A.** Population level, **B.** Single-neuron level and **C.** Subcellular level.

Peters' rule is not universal and has been shown to hold for certain cases and fail in others, for all levels of detail. Section 3.1 describes an exception to Peters' rule at the subcellular level, which probably carries over to the single-neuron level as well: an apposition is more likely to predict a synapse if other synapses are present on the same neurites [114]. Other studies have provided evidence both in favor of and against the heuristic at the subcellular level [158,159,160,115]. Neurite proximity is undeniably a necessary condition for the formation of synapses, but in general not sufficient to explain it, for instance as activity-dependent plasticity may support preferential connectivity between neurons with similar response properties. Nevertheless, Peters' rule is a decent heuristic at the population level, with the main caveat that some cell types do not connect to each other even if they come into close proximity [106,157]. Thus, the rule may be fruitfully applied at the population level as long as such cell-type-specific absence of connections is taken into account.

4.2 Architectural principles

The cytoarchitecture and laminar composition of cortical areas are predictive of their connectivity, as first noted for frontal areas of macaque cortex [161,64]. In particular, architecturally more similar areas are more likely to be connected, and if they are connected, the connection density tends to be higher [162,151,45,163]. However, while architectural similarity reliably predicts the existence and ab-

sence of connections, connection densities are better explained by inter-area distances (cf. section 4.3) [162]. The characterization of areal architecture in terms of laminar differentiation was systematized using the notion of architectural types, which also consider the thickness of layer IV [65]. Areas with low architectural type have low neuron density, a thin or absent layer IV, and indistinct lamination. Areas with high architectural type have high neuron density, a thick layer IV, and distinct lamination. The progression from low to high architectural types roughly corresponds to the inverse of cortical hierarchies, down from limbic to early sensory areas. Instead of using architectural types, which discretize what is in fact a continuum of structural features across areas [164], one may use neuron density as a continuous explanatory variable. However, compared to neural density differences, architectural type differences are a better predictor of the existence and absence of connections between macaque visual areas [45].

Besides correlating with the existence or absence of connections and with connection density, architectural differences are informative of laminar projection patterns. Cytoarchitectonic difference is the only consistent predictor that explains the majority of the variance in laminar source patterns when compared with other candidate explanatory variables such as rostrocaudal distance [165]. Areas with more distinctive layers and higher neuron density tend to send projections from their upper (supragranular) layers to areas with less distinctive layers and lower neuron density. Reversely, projections from the latter to the former type of areas tend to emanate from the lower (infragranular) layers. These patterns seem to generalize across species, having already been demonstrated for cat, marmoset, and macaque [166]. Since laminar origin patterns are correlated with laminar termination patterns, for instance supragranular projections tend to target the granular layer IV [167], also termination patterns can be in part inferred from architectural similarity [151,46]. However, as the majority of layer-resolved axonal tracing data is retrograde, origin patterns have been more extensively studied than termination patterns. For human cortex, laminar origin and termination patterns of inter-area projections are still mostly unknown. For modeling purposes, the relationships between laminar patterns and cytoarchitectural differences between areas that have been observed in different mammalian species may be used to assign laminar patterns to human connectomes (figure 4).

Cortical thickness similarity has also been investigated as an explanatory variable for inter-area connectivity. Areas with more similar thickness are more likely to be connected, although this relationship does not hold consistently [45]. Thickness differences also relate to laminar patterns: projections from thinner to thicker areas tend to have a more supralaminar origin [163]. The fact that cortical thickness is somewhat predictive of connectivity fits with the observation that cortical thickness correlates negatively with neuron density [163,46]. However, compared to cortical thicknesses, architectural types and neuron densities are more systematically related to connectional features, indicating that cytoarchitecture is at the heart of the relation between cortical thickness and connectivity. More commonly, thickness similarity has been characterized in the sense of co-variation across subjects, areas with positively co-varying thicknesses

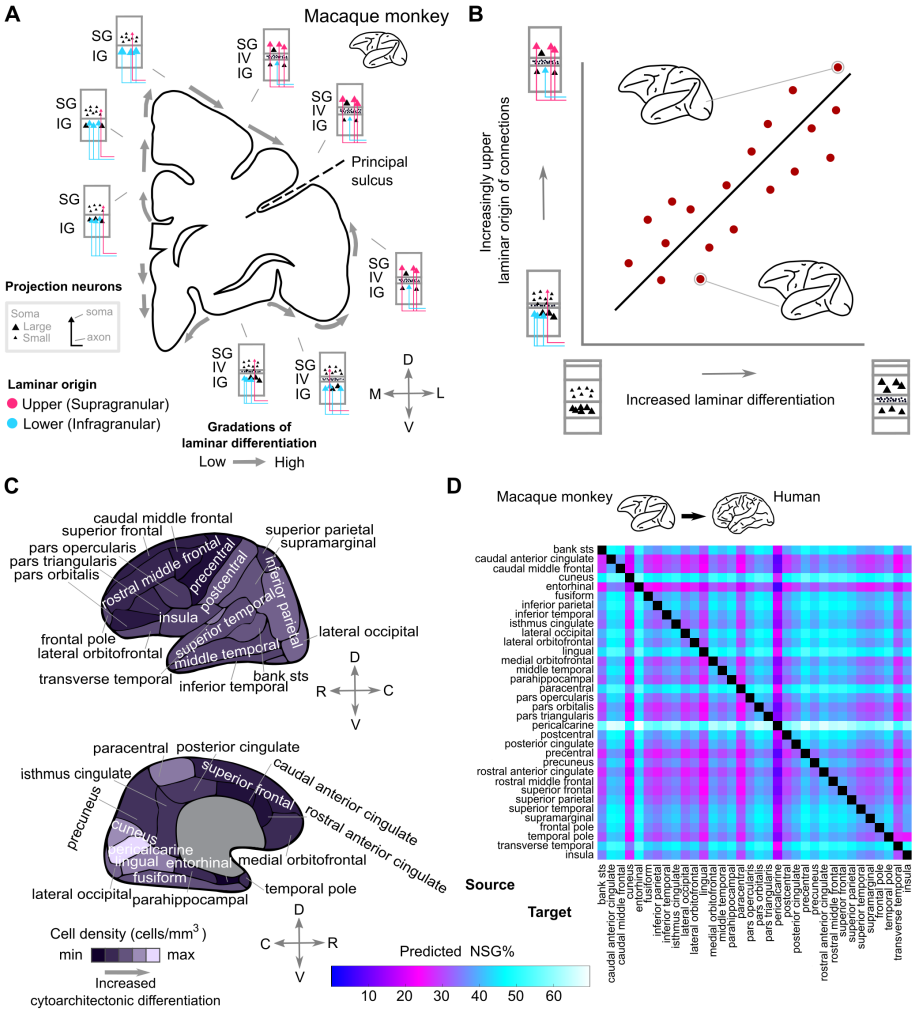


Figure 4. Laminar origin of connections, cytoarchitecture, and predictive connectomics. **A.** Laminar origin of connections shifts from lower to upper layers across the cortical sheet of the macaque monkey. **B.** Schematic illustration of the quantitative relation between the cytoarchitecture of cortical areas and the laminar origin of their connections to other areas. The transition from less to more laminar differentiation (horizontal axis), associated also with increased neural density, is accompanied by a transition of predominantly lower to upper laminar origin of connections (vertical axis). **C.** Cell densities of human cortical areas based on von Economo and Koskinas, 1925 [22]. Top, lateral view and bottom, medial view of the right hemisphere. **D.** A monkey-to-human prediction of laminar origin of connections ($NSG\%$, relative number of supragranular neurons) between all pairs of cortical areas based on human cell densities (Panel C) and the relationship between cytoarchitecture and the laminar origin of connections (Panel B). Panel A based on a drawing from [168]. Panels C and D reproduced from [165].

across subjects being more likely to be connected [169,170,171]. However, also this correlation is far from perfect, and a large percentage of regions have co-varying thickness without being connected [170].

4.3 Distance dependence

Both for connectivity between neurons within a given brain region and for that between brain regions, shorter connections are more likely or more numerous than longer ones. This rule makes sense considering the material and energetic cost of wiring and the space taken up by axons and axon bundles. Nevertheless, non-random long-range connections between specific regions exist, which are in part explained by spatiotemporal patterns of brain development (cf. section 4.5). Locally within cortical areas, connection probability of both excitatory and inhibitory neurons falls off approximately exponentially with intersomatic distance with a space constant around $150 - 300\ \mu$m [126,127,172,173]. Besides these local connections, pyramidal cells establish patchy connectivity at distances on the scale of millimeters [174].

Similarly to local connectivity, projections between cortical areas follow an 'exponential distance rule' in which the lengths of axons are exponentially distributed and the probability for a neuron to send a projection between cortical areas thus falls off exponentially with distance [175]. This exponential distance rule at the level of individual neurons translates into an exponential decay in connection density at the level of areas as well [46]. Given the connectivity between cortical areas, the spatial arrangement of areas in the brain to a good approximation minimizes the total wiring length [176,177,175]. In a study of the connectivity between macaque cortical areas [143], the combination of the log ratio of neuron densities and Euclidian distance between areas provided the best statistical predictions of the existence of connections [163]. All in all, physical distance constitutes a useful explanatory variable for the existence and density of both local and long-range connectivity.

4.4 Connectome topology

So far we have considered connectivity predictions based on the properties of pairs of network nodes (neurons or areas). It is possible to go beyond pairwise properties and look at patterns of three or more nodes to infer connectivity. According to the homophily principle—described in social network theory as 'the tendency to choose as friends those similar to oneself' [178]—nodes with common neighbors are more likely to be themselves connected [179,165]. This property is for instance displayed by so-called small-world networks, in which a combination of many short-range and a few long-range connections enables any node to be reached via a small number of hops through the network. The homophily principle holds sway both at the single-neuron level and at the level of brain regions, in both vertebrate and invertebrate brains [179].

In local cortical circuits, certain connection motifs—patterns of connectivity in small groups of nodes—between three or more neurons are overrepresented

with respect to random graphs defined by pairwise connection probabilities alone [126,127]. In a study of groups of up to twelve neurons, the probability of a connection between a pair of neurons was found to increase linearly with the number of common neighbors. Through this expression of the homophily principle, cortical neurons cluster into small-world networks [127]. Furthermore, like-to-like connectivity between neurons with similar functional specificity, e.g., neurons in primary visual cortex having similar orientation preference or responding to the same type of visual stimuli [180], is an important ingredient of the local network topology [181].

At the level of brain regions, Jouve et al. [182] noticed that directly connected areas in macaque vision-related cortex have far more indirect connections between them than do unconnected areas. The author defined an index of connectivity that captures the fraction of shared first-order intermediate nodes between any two areas (Figure 5A). They found that this metric is related to the existence or absence of connections in macaque visual cortex, and used this to infer the connectivity of area pairs for which no tracing data were available. As pointed out in the study, the given indirect connectivity index cannot predict all connections accurately, but nevertheless exposes an underlying principle in the structure of the primate connectome.

We computed the index of indirect connectivity and the triadic motif counts on the tract-tracing data from macaque [143,183] and marmoset [145] monkeys, using the subgraphs without unknown connections. This analysis reveals that the motif counts, relative to random graphs defined by pairwise connection probabilities alone, have a similar structure in both primates, as shown previously [184] (Figure 5B). We also see that the index of connectivity has a large overlap for areas with and without a direct connection in both primates (Figure 5D). However, extreme values (> 0.8 and < 0.3) reliably distinguish existing connections from non-existing ones.

A combination of spatial proximity and homophily accounts for many topological characteristics of human cortical networks such as degree, clustering, and betweenness centrality distributions [185,186]. Chen et al. (2020) [187] found that adding cytoarchitectonic similarity to distance dependence and topological constraints resulted in even better predictions when applied to the macaque cortical connectome. These findings place local topology, and especially homophilic attachment, in the list of overarching properties governing neural network structure.

4.5 Neurodevelopmental underpinnings of connectivity heuristics

Many of the aforementioned connectivity heuristics can be brought together in a common developmental framework. The spatiotemporal ontogeny of the brain provides simultaneous explanations for distance-dependent connectivity, the preferential connectivity between cytoarchitectonically similar areas, and aspects of the network topology of the brain [179]. It also accounts for deviations from a simple decay in connection probability with distance. For instance,

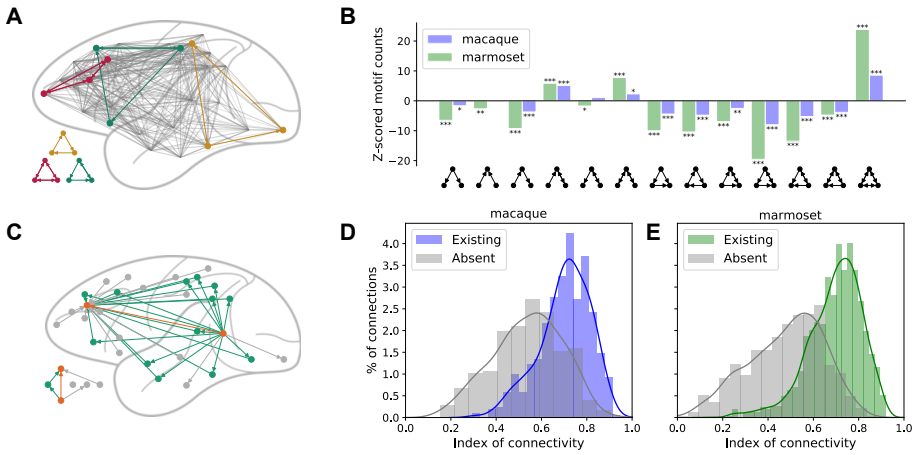


Figure 5. Illustration of topological connectivity features of macaque and marmoset cortical graphs. **A.** Schematic depiction of motifs in the area-level macaque cortico-cortical connectivity. **B.** Z-score of the motif counts for all connected triads in the macaque and marmoset. Motif counts are normalized by the mean and standard deviation of the motif counts from 1,000 random graphs with the same connection probability as the experimental data in each case; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. **C.** Schematic depiction of the area-level index of connectivity as described in [182]. Shared neighbors (green nodes) contribute to the prediction of a direct connection (orange), while non-shared neighbors (gray nodes) make a direct connection less likely. **D,E.** Distribution of the index of connectivity for existing and absent cortico-cortical connections in macaque (D) and marmoset (E).

changes in the parameters of the distance-dependent connectivity during development can yield a small-world network structure with multiple clusters [188]. Limbic cortical areas, of low architectural type, develop earlier and over a shorter period than areas of high laminar differentiation such as primary visual cortex. This rapid development not only underlies the less distinct lamination and low neuron density of limbic areas, but also gives these areas a longer time window for connecting to other regions, thus supporting their coordinating role [189]. The importance of spatial embedding and heterochronicity—the existence of a sequence of developmental time windows—for brain wiring were demonstrated for species ranging from the fruit fly to the mouse, rat, macaque monkey, and human [190,179]. Thus, taking into account spatiotemporal gradients of brain development can help predict more realistic connectomes regardless of the species under investigation.

4.6 Reconstructing connectivity from activity

So far we have focused on predictive relations derived from the anatomical features of the nervous tissue. However, anatomical information is often costly to

obtain or requires invasive methods and is therefore often not available for all the different brain regions. An alternative approach is to derive neural network structure from activity data. While promising results in this direction have been obtained, this approach suffers from the drawbacks that widely different network parameters can lead to closely similar activity [191] and that the external input to the network modulates the link between structure and activity [192].

When relating activity to connectivity, we need to distinguish a few different terms. Besides structural connectivity, the topic of this chapter, there are two types of activity-dependent 'connectivity': so-called functional connectivity, and effective connectivity. Functional connectivity is symmetric between source and target nodes, and describes correlations between their activity. It is often used in the context of functional imaging studies to characterize the interactions between brain regions. Effective connectivity is a directed measure, describing the minimal graph that would be needed to account for the observed interactions between nodes [192]. In a stricter mathematical sense, one can define effective connectivity as the product of the structural connectivity and effective synaptic weights that depend on the activity level of the target nodes and quantify their susceptibility to increased input [193]. The same structural substrate can support different functional and effective connectivities depending on the external drive and the network state. When inferring structural connectivity from activity data, the lines between the different types of connectivity can be somewhat blurred, but it is useful to keep in mind the distinctions.

We have already discussed two physiological methods that help estimate structural connectivity at the microscopic scale: paired recordings (section 3.2) and glutamate uncaging (section 3.3). These methods provide reliable connectivity data, but are constrained to small numbers of neurons. Parallel electrophysiological recordings of up to hundreds of individual neurons are now possible for instance with Utah arrays or Neuropixels probes [194,195], and functional magnetic resonance imaging enables recording whole-brain activity, resolved into ever smaller voxels [196,197].

A number of methods have been proposed for inferring the underlying connectivity from these large-scale activity data. Time-lagged correlations between the spike trains of pairs of neurons are informative about the direction of the information flow and have been shown to be linked to the structural connectivity [198]. A few studies have used this fact to reconstruct network connectivity from parallel spike train cross-correlation histograms [199,200,201]. Pairwise correlations are shaped not only by direct connections between neurons, but also by indirect connections, the electrophysiological properties of the individual neurons, transmission delays, and the external drive to the network [202,203,204]. Given certain conditions such as stationarity and knowledge of the single-neuron electrophysiology, the structural connectivity can in principle be uniquely reconstructed from the pairwise correlation functions; that is, one can compute and thereby take into account the influence of the indirect connections and shared input [205,204,193]. In practice, biological neural networks do not fulfill ideal con-

ditions and experiments do not fully provide the required information, setting a ceiling on the accuracy of structural connectivity inferred from correlations.

Going beyond pairwise correlations, Casadiego et al. [206] propose a method for inferring synaptic connections from the dependence of inter-spike intervals on cross-spike intervals, i.e. intervals between the spike times of different neurons. The method can successfully distinguish excitatory and inhibitory synapses, as validated with point neuron network simulations. Networks exhibiting phase-locked activity may not sufficiently explore the dynamical landscape to enable all synapses to be reconstructed. In such cases it can help to expose the network to different external driving conditions [207]. Similarly using only knowledge of the spiking activity and not requiring membrane potential traces, Zaytsev et al. [208] infer the connectivity of simulated networks of a thousand neurons using maximum likelihood estimation of a generalized linear model of the spiking activity. Such methods based on generalized linear models can work well when the activity of all neurons is recorded [209], but, like for any connectivity reconstruction method, undersampling is expected to diminish their performance.

Fitting the observed activity to a dynamical network model can be a complex and computationally intensive procedure. Structural connectivity parameters are sought that optimize a score or cost function based on some features of interest. In simulation-based methods, optimal parameter combinations can be searched via brute force [191,210], stochastic optimization techniques such as evolutionary methods [211,212,213], or plasticity rules [214]. Likelihood-based methods do not require costly simulations [215,216,217,218] and under some conditions allow straightforward optimization via gradient ascent or simplex methods. However, estimating the analytical likelihood function is a challenging task for complex models. Machine learning methods are starting to be developed that can overcome this issue and estimate parameter distributions given emergent dynamical properties of modeled networks [219,220].

All in all, establishing unequivocal links between structural connectivity and neural activity remains a major challenge in neuroscience, and structural connectivity estimates from population recordings should generally be interpreted with caution.

5 Validation of predicted connectivity

The most direct way of validating connectivity predictions is of course experimental confirmation. Barring the ideal situation where this is possible, we have a few options at our disposal for putting predictions to the test. In this context, different types of predictions exist: sometimes, a full connectome is generated, while sometimes merely statistical regularities in connectivity data are obtained. For the case of full connectomes, we can further distinguish generative models that do not directly rely on connectivity data, for instance based on distance, cytoarchitectonics, and topological constraints; and cases where gaps in connectivity data are filled in.

Where the result of the prediction is a full connectome, one can compare with experimentally obtained connectomes either edge-wise or based on graph properties such as degree distributions, clustering, modularity, characteristic path length, small worldness, or betweenness centrality [185,186]. The choice of properties to compare is nontrivial and depends on their presumed importance with regard to the scientific question. Ideally, the fitness of the generative model is quantified using a likelihood function, but where this is difficult, other objective functions may be defined [221].

In case of statistical fits to connectivity data, we can check the robustness of the predictions by determining confidence intervals for the fit parameters. When no straightforward expressions for these are available, bootstrapping provides a solution in which random data samples are drawn with replacement and the statistic of interest is computed for each sample [222]. A similar strategy can be applied when filling gaps in connectomes: leaving out part of the known data and either determining how well the predictions fit to the left-out data, or again computing graph properties and assessing their variability. Alternatively, we can add noise to the underlying data on the order of the uncertainty in the data. Depending on the case, 'uncertainty' in this context can for instance include experimental noise, inter-individual and inter-species variability, or uncertainty due to mapping between parcellations. Since it is in practice difficult to determine the size of the uncertainty, one can add different levels of noise to the estimated model parameters and check whether the predictions hold true even for relatively high noise levels.

Another route for testing the plausibility of connectivity predictions is to build corresponding network models, perform dynamical simulations, and compare the resulting activity with experimental activity data. Software tools supporting the systematic comparison between simulated and experimental activity data are available for both single neurons and networks of neurons [223]. This method is complicated by the fact that not only the connectivity but also the dynamical properties of the nodes (neurons or populations of neurons), the transmission delays, and the external drive contribute to the network dynamics. However, depending on the dynamical regime, network dynamics can be fairly robust to electrophysiological properties of the individual nodes [224]. The parameter space can be explored systematically via parameter scans, or in a more targeted manner via stochastic optimization. If at least some parameter settings for the nodes, delays, and external drive, consistent with biological data, can be found for which the predicted connectivity yields realistic activity, this provides some degree of validation. Stronger support is provided if the experimental activity data are no longer successfully reproduced upon changing the connectivity. Ultimately, neural network models should be consistent with both anatomical and electrophysiological properties of the brain.

6 Concluding remarks

Data on brain anatomy are increasingly made available as systematic, quantitative data sets, facilitating their use in neuronal network models. Inspired by seminal works like those of von Economo [22] and Braitenberg and Schüz [156], modern anatomists recognize the importance of systematization and quantification for informing analyses and models. Historically, much anatomical data was made available only in the natural language text of publications. On the example of tracing studies, the creators of the CoCoMac database [76] recognized the need to bring these data into a machine readable format and to create a framework for systematically mapping the parcellations mentioned in the text to different parcellations of choice when constructing connectivity maps. The modern, systematic way of publishing data is most prominently represented by large-scale initiatives like the Allen Institute for Brain Science, Janelia Research Campus, the Human Connectome Project, the Japanese Brain/MINDS project, and the European Human Brain Project. Nevertheless, there is sometimes still a disconnect between experimentalists and computational neuroscientists in terms of the formats in which the data are published. Anatomical data are still often made available as image files which require additional processing before they can flow into models, in formats specific to the discipline. An illustrative anecdote is that in 2018 Schmidt et al. [46] still obtained cortical thickness from micrographs by measuring with a ruler the distance between layer markers. One reason why modelers generally cannot use image data directly is that they tend to work with concepts like definite cortical areas and layers, rather than in a spatial continuum. These categorical concepts constitute strong hypotheses that help to reduce and interpret the data. Tables of area or laminar averages are then more useful than images. If the data are offered as images, at least scripts and documentation should be published alongside the data to enable the relevant quantities to be potentially more easily extracted. The latter approach retains flexibility with respect to particular parcellations and is future-proof as algorithms of feature extraction improve and concepts of brain organization may change over time.

We have described methods ranging from microscopy to diffusion magnetic resonance imaging for measuring connectivity. However, this list is not exhaustive and novel techniques are continuously developed. A modern technique is polarized light imaging (PLI), which measures fiber orientations in brain slices using the birefringence properties of myelin [225,226]. Three-dimensional reconstructions enable fiber tracts to be followed through the brain at a resolution of some tens of micrometers. Axons entering the white matter can be visualized with an in-plane pixel size down to the micrometer scale. An add-on to PLI, also based on transmitting polarized light through histological sections, is Diattenuation Imaging, which provides complementary information on tissue composition [227]. These methods promise new ways of determining the connectivity of neural network models.

Also in the field of predictive connectomics, our treatment of methods has not been exhaustive. Besides predictions based on the proximity of neural pro-

cesses or cell bodies, cytoarchitecture, topological constraints, and neural network activity, it is for instance possible to generate connectomes based on gene expression data [228,229,230]. Another possibility we have only briefly alluded to is a normative approach, in which the connectome is in some sense assumed to be optimal, and the implications of this assumption for connectivity are investigated [231,232]. As in so many fields of science, machine learning methods and artifical neural network models provide a promising new avenue for identifying regularities in data that help to predict connectivity.

As we have seen, connectomes for neural network models are subject to a variety of uncertainties. Each experimental method carries with it measurement errors, data from multiple individuals tend to be needed to fully specify a connectome, and in many cases the best available estimates even come from different species. We have largely skipped over the vast and difficult topic of mapping data between species. In many cases, the sobering truth is that this cannot be done in a fully principled manner. All types of uncertainties, whether due to experimental methods, individual differences, or interspecies differences, lead to uncertainties in predicted model dynamics. We have described some ways of verifying the robustness of network models to these uncertainties.

Brain models based on these statistical rules are necessarily models of an average brain. This limits their explanatory power. Not only in humans but also in other species, macroscopic features of brain dynamics, like dominant frequencies and functional connectivity, vary from individual to individual [233,234,235]. When the deviation of simulated brain activity from experimental data is of the same order as the inter-individual variability, there is nothing left to explain for this type of model. Schmidt et al. [236] illustrate this situation for the prediction of functional connectivity between areas on the basis of a spiking network model. Such observations challenge the research strategy to aggregate data from different species and individuals to arrive at a statistical model of brain structure. Progress may eventually only be possible by further constraining generic connectivity rules by anatomical data obtained from the individual delivering the brain activity data to be predicted [237].

Ultimately, the statistical descriptions we apply to summarize brain organization are not the rules by which brains are built in nature. The rules mathematically formalize the limits of our knowledge on the structure of individual brains. And using these rules is to date just the most efficient way of instantiating large-scale neuronal networks in a computer by a fully parallel process [238]. In nature brains are pre-shaped by evolution and further formed by growth rules in continuous interaction with the environment. Eventually we need to understand and formalize these more fundamental rules to grow artificial individual brains in a computer. This implies the existence of a sufficiently accurate model of the environment. Averages over such model instances then in turn need to be consistent with our former statistical descriptions of brain structure.

Nevertheless, the major short-term challenge consists in the construction of brain models encompassing different brain components, as already alluded to in the introduction of this chapter. With a few notable exceptions, until today

models of neuronal networks are usually constructed by a single researcher, often a PhD student, or small research groups. It seems likely that we have hit a complexity barrier and for this reason the complexity of the majority of models has not increased much over past decade. In order to integrate the heterogeneity of different brain areas and their multi-level hierarchical organization into a brain model will require that we learn to use models of brain components created by other researchers as building blocks.

International large-scale projects like EBRAINS have started to create the ICT infrastructure enabling the sharing and reuse of data and model components, as well as the simulation of multi-scale models and their environments. The hope is that using these infrastructures fosters the required culture of sharing and collaboration in neuroscience.

Acknowledgments

Supported by the European Union's Horizon 2020 Framework Programme for Research and Innovation under Specific Grant Agreements No. 785907 and 945539 (Human Brain Project SGA2, SGA3) and Priority Program 2041 (SPP 2041) "Computational Connectomics" of the German Research Foundation (DFG).

References

1. Cook, S.J., Jarrell, T.A., Brittin, C.A., Wang, Y., Bloniarz, A.E., Yakovlev, M.A., Nguyen, K.C., Tang, L.T.H., Bayer, E.A., Duerr, J.S., et al.: Whole-animal connectomes of both Caenorhabditis elegans sexes. Nature **571**(7763) (2019) 63–71
2. Kandel, E.R.: In Search of Memory: The Emergence of a New Science of Mind. WW Norton & Company, New York (2007)
3. Orban, G.A., Van Essen, D., Vanduffel, W.: Comparative mapping of higher visual areas in monkeys and humans. Trends Cogn. Sci. **8**(7) (2004) 315–324
4. Sereno, M.I., Tootell, R.B.: From monkeys to humans: what do we now know about brain homologies? Curr. Opin. Neurobiol. **15**(2) (2005) 135–144
5. Hutchison, R.M., Everling, S.: Monkey in the middle: why non-human primates are needed to bridge the gap in resting-state investigations. Front. Neuroanat. **6** (2012) 29
6. Jolivet, R., Schürmann, F., Berger, T.K., Naud, R., Gerstner, W., Roth, A.: The quantitative single-neuron modeling competition. Biol. Cybern. **99**(4-5) (2008) 417–426
7. Teeter, C., Iyer, R., Menon, V., Gouwens, N., Feng, D., Berg, J., Szafer, A., Cain, N., Zeng, H., Hawrylycz, M., Koch, C., Mihalas, S.: Generalized leaky integrate-and-fire models classify multiple neuron types. Nat. Commun. **9** (2018) 709
8. Talairach, J., Tournoux, P.: Co-planar stereotaxic atlas of the human brain. Thieme, Stuttgart (1988)
9. Lancaster, J.L., Tordesillas-Gutiérrez, D., Martinez, M., Salinas, F., Evans, A., Zilles, K., Mazziotta, J.C., Fox, P.T.: Bias between MNI and Talairach coordinates analyzed using the ICBM-152 brain template. Human Brain Mapping **28**(11) (2007) 1194–1205

10. Laird, A.R., Robinson, J.L., McMillan, K.M., Tordesillas-Gutiérrez, D., Moran, S.T., Gonzales, S.M., Ray, K.L., Franklin, C., Glahn, D.C., Fox, P.T., Lancaster, J.: Comparison of the disparity between Talairach and MNI coordinates in functional neuroimaging data: validation of the Lancaster transform. NeuroImage **51**(2) (2010) 677–683
11. Amunts, K., Lepage, C., Borgeat, L., Mohlberg, H., Dickscheid, T., Rousseau, M.É., Bludau, S., Bazin, P.L., Lewis, L.B., Oros-Peusquens, A.M., et al.: BigBrain: an ultrahigh-resolution 3D human brain model. Science **340**(6139) (2013) 1472–1475
12. Dickscheid, T., Haas, S., Bludau, S., Glock, P., Huysegoms, M., Amunts, K.: Towards 3D reconstruction of neuronal cell distributions from histological human brain sections. In Grandinetti, L., Joubert, G.R., Michielsen, K., eds.: Future Trends of HPC in a Disruptive Scenario. IOS Press (2019) 223–238
13. Amunts, K., Mohlberg, H., Bludau, S., Zilles, K.: Julich-Brain: a 3D probabilistic atlas of the human brain's cytoarchitecture. Science (First Release) (2020) eabb4588
14. Zilles, K., Palomero-Gallagher, N., Schleicher, A.: Transmitter receptors and functional anatomy of the cerebral cortex. J. Anat. **205**(6) (2004) 417–432
15. Eickhoff, S.B., Thirion, B., Varoquaux, G., Bzdok, D.: Connectivity-based parcellation: Critique and implications. Human Brain Mapping **36**(12) (2015) 4771–4792
16. Fan, L., Li, H., Zhuo, J., Zhang, Y., Wang, J., Chen, L., Yang, Z., Chu, C., Xie, S., Laird, A.R., et al.: The human Brainnetome Atlas: a new brain atlas based on connectional architecture. Cereb. Cortex **26**(8) (2016) 3508–3526
17. Gordon, E.M., Laumann, T.O., Adeyemo, B., Huckins, J.F., Kelley, W.M., Petersen, S.E.: Generation and evaluation of a cortical area parcellation from resting-state correlations. Cereb. Cortex **26**(1) (2016) 288–303
18. Bohland, J.W., Bokil, H., Allen, C.B., Mitra, P.P.: The brain atlas concordance problem: quantitative comparison of anatomical parcellations. PLOS One **4**(9) (2009)
19. Arslan, S., Ktena, S.I., Makropoulos, A., Robinson, E.C., Rueckert, D., Parisot, S.: Human brain mapping: A systematic comparison of parcellation methods for the human cerebral cortex. NeuroImage **170** (2018) 5–30
20. Van Essen, D.C., Glasser, M.F.: Parcellating cerebral cortex: How invasive animal studies inform noninvasive mapmaking in humans. Neuron **99**(4) (2018) 640–663
21. Brodmann, K.: Vergleichende Lokalisationslehre der Großhirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues. Johann Ambrosius Barth, Leipzig (1909)
22. von Economo, C.F., Koskinas, G.N.: Die Cytoarchitektonik der Hirnrinde des erwachsenen Menschen. J. Springer (1925)
23. Schiffer, C., Spitzer, H., Kiwitz, K., Amunts, K., Dickscheid, T.: Deep learning speeds up gapless cytoarchitectonic mapping in the human brain. `https://ww5.aievolution.com/hbm1901/index.cfm?do=abs.viewAbs&abs=1535` (2019) OHBM conference abstract.
24. Wagstyl, K., Lepage, C., Bludau, S., Zilles, K., Fletcher, P.C., Amunts, K., Evans, A.C.: Mapping cortical laminar structure in the 3D BigBrain. Cereb. Cortex **28**(7) (2018) 2551–2562
25. Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., et al.: An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. NeuroImage **31**(3) (2006) 968–980

26. Shen, E.H., Overly, C.C., Jones, A.R.: The Allen Human Brain Atlas: comprehensive gene expression mapping of the human brain. Trends Neurosci. **35**(12) (2012) 711–714
27. Sunkin, S.M., Ng, L., Lau, C., Dolbeare, T., Gilbert, T.L., Thompson, C.L., Hawrylycz, M., Dang, C.: Allen Brain Atlas: an integrated spatio-temporal portal for exploring the central nervous system. Nucleic Acids Research **41**(D1) (2012) D996–D1008
28. Markov, N.T., Vezoli, J., Chameau, P., Falchier, A., Quilodran, R., Huissoud, C., Lamy, C., Misery, P., Giroud, P., Ullman, S., Barone, P., Dehay, C., Knoblauch, K., Kennedy, H.: Anatomy of hierarchy: Feedforward and feedback pathways in macaque visual cortex. J. Compar. Neurol. **522**(1) (2014) 225–259
29. Calabrese, E., Badea, A., Coe, C.L., Lubach, G.R., Shi, Y., Styner, M.A., Johnson, G.A.: A diffusion tensor MRI atlas of the postmortem rhesus macaque brain. NeuroImage **117** (2015) 408–416
30. Lein, E.S., Hawrylycz, M.J., Ao, N., Ayres, M., Bensinger, A., Bernard, A., Boe, A.F., Boguski, M.S., Brockway, K.S., Byrnes, E.J., et al.: Genome-wide atlas of gene expression in the adult mouse brain. Nature **445**(7124) (2007) 168–176
31. Dong, H.W.: The Allen reference atlas: A digital color brain atlas of the C57Bl/6J male mouse. John Wiley & Sons Inc (2008)
32. Kuan, L., Li, Y., Lau, C., Feng, D., Bernard, A., Sunkin, S.M., Zeng, H., Dang, C., Hawrylycz, M., Ng, L.: Neuroinformatics of the Allen Mouse Brain Connectivity Atlas. Methods **73** (2015) 4–17
33. Paxinos, G., Franklin, K.B.: Paxinos and Franklin's The Mouse Brain in Stereotaxic Coordinates. 5 edn. Academic Press (2019)
34. Chon, U., Vanselow, D.J., Cheng, K.C., Kim, Y.: Enhanced and unified anatomical labeling for a common mouse brain atlas. Nature Communications **10**(1) (2019) 1–12
35. Bakker, R., Tiesinga, P., Kötter, R.: The Scalable Brain Atlas: Instant web-based access to public brain atlases and related content. Neuroinformatics **13** (2015) 353–366
36. Schiffer, C., Kiwitz, K., Amunts, K., Dickscheid, T.: Ultrahigh resolution 3D cytoarchitectonic map of Area hOc1 (V1, 17, CalcS) created by a Deep-Learning assisted workflow [Data set]. Human Brain Project Neuroinformatics Platform (2019) DOI: 10.25493/DGEZ-Q93.
37. Fischl, B., Dale, A.M.: Measuring the thickness of the human cerebral cortex from magnetic resonance images. Proc. Natl. Acad. Sci. USA **97**(20) (September 2000) 11050–11055
38. Lüsebrink, F., Wollrab, A., Speck, O.: Cortical thickness determination of the human brain using high resolution 3 T and 7 T MRI data. NeuroImage **70** (2013) 122–131
39. Cardinale, F., Chinnici, G., Bramerio, M., Mai, R., Sartori, I., Cossu, M., Russo, G.L., Castana, L., Colombo, N., Caborni, C., et al.: Validation of FreeSurfer-estimated brain cortical thickness: comparison with histologic measurements. Neuroinformatics **12**(4) (2014) 535–542
40. Wagstyl, K., Lerch, J.P.: Cortical thickness. In: Brain Morphometry. Springer (2018) 35–49
41. Li, D., Zavaglia, M., Wang, G., Xie, H., Hu, Y., Werner, R., Guan, J.S., Hilgetag, C.C.: Discrimination of the hierarchical structure of cortical layers in 2-photon microscopy data by combined unsupervised and supervised machine learning. Scientific Reports **9**(1) (2019) 1–16

42. Von Economo, C.: Cellular Structure of the Human Cerebral Cortex. Karger Medical and Scientific Publishers (2009) Translated and edited by L.C. Triarhou.
43. Wagstyl, K., Larocque, S., Cucurull, G., Lepage, C., Cohen, J.P., Bludau, S., Palomero-Gallagher, N., Lewis, L.B., Funck, T., Spitzer, H., et al.: BigBrain 3D atlas of cortical layers: Cortical and laminar thickness gradients diverge in sensory and motor cortices. PLOS Biol. **18**(4) (2020) e3000678
44. Alvarez, I., Parker, A.J., Bridge, H.: Normative cerebral cortical thickness for human visual areas. NeuroImage **201** (2019) 116057
45. Hilgetag, C.C., Medalla, M., Beul, S.F., Barbas, H.: The primate connectome in context: Principles of connections of the cortical visual system. NeuroImage **134** (2016) 685–702
46. Schmidt, M., Bakker, R., Hilgetag, C.C., Diesmann, M., van Albada, S.J.: Multi-scale account of the network structure of macaque visual cortex. Brain Struct. Func. **223**(3) (April 2018) 1409–1435
47. Cahalane, D., Charvet, C., Finlay, B.: Systematic, balancing gradients in neuron density and number across the primate isocortex. Front. Neuroanat. **6** (2012) 28
48. Pagani, M., Damiano, M., Galbusera, A., Tsaftaris, S.A., Gozzi, A.: Semi-automated registration-based anatomical labelling, voxel based morphometry and cortical thickness mapping of the mouse brain. J. Neurosci. Methods **267** (2016) 62–73
49. Feo, R., Giove, F.: Towards an efficient segmentation of small rodents brain: a short critical review. J. Neurosci. Methods (2019) 82–29
50. West, M.J.: New stereological methods for counting neurons. Neurobiol. Aging **14**(4) (1993) 275–285
51. Miller, D.J., Balaram, P., Young, N.A., Kaas, J.H.: Three counting methods agree on cell and neuron number in chimpanzee primary visual cortex. Front. Neuroanat. **8** (2014) 36
52. Herculano-Houzel, S., Lent, R.: Isotropic fractionator: a simple, rapid method for the quantification of total cell and neuron numbers in the brain. J. Neurosci. **25**(10) (2005) 2518–2521
53. West, M., Slomianka, L., Gundersen, H.J.G.: Unbiased stereological estimation of the total number of neurons in the subdivisions of the rat hippocampus using the optical fractionator. The Anatomical Record **231**(4) (1991) 482–497
54. Sterio, D.: The unbiased estimation of number and sizes of arbitrary particles using the disector. J. Microsc. **134**(2) (1984) 127–136
55. Mullen, R.J., Buck, C.R., Smith, A.M.: NeuN, a neuronal specific nuclear protein in vertebrates. Development **116**(1) (1992) 201–211
56. Merker, B.: Silver staining of cell bodies by means of physical development. J. Neurosci. Methods **9**(3) (1983) 235–241
57. Erö, C., Gewaltig, M.O., Keller, D., Markram, H.: A cell atlas for the mouse brain. Front. Neuroinform. **12** (2018) 84
58. Herculano-Houzel, S., Mota, B., Lent, R.: Cellular scaling rules for rodent brains. Proc. Natl. Acad. Sci. USA **103**(32) (2006) 12138–12143
59. Azevedo, F.A.C., Carvalho, L.R.B., Grinberg, L.T., Farfel, J.M., Ferretti, R.E.L., Leite, R.E.P., Filho, W.J., Lent, R., Herculano-Houzel, S.: Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. J. Compar. Neurol. **513**(5) (April 2009) 532–541
60. Herculano-Houzel, S.: The human brain in numbers: a linearly scaled-up primate brain. Front. Hum. Neurosci. **3** (2009) 31
61. Sarko, D.K., Catania, K.C., Leitch, D.B., Kaas, J.H., Herculano-Houzel, S.: Cellular scaling rules of insectivore brains. Front. Neuroanat. **3** (2009) 8

62. Herculano-Houzel, S.: The remarkable, yet not extraordinary, human brain as a scaled-up primate brain and its associated cost. Proceedings of the National Academy of Sciences **109**(Supplement 1) (June 2012) 10661–10668
63. Collins, C.E., Airey, D.C., Young, N.A., Leitch, D.B., Kaas, J.H.: Neuron densities vary across and within cortical areas in primates. Proc. Natl. Acad. Sci. USA **107**(36) (September 2010) 15927–15932
64. Barbas, H., Rempel-Clower, N.: Cortical structure predicts the pattern of corticocortical connections. Cereb. Cortex **7**(7) (1997) 635–646
65. Dombrowski, S., Hilgetag, C., Barbas, H.: Quantitative architecture distinguishes prefrontal cortical systems in the rhesus monkey. Cereb. Cortex **11**(10) (2001) 975–988
66. García-Cabezas, M.Á., Zikopoulos, B., Barbas, H.: The Structural Model: a theory linking connections, plasticity, pathology, development and evolution of the cerebral cortex. Brain Struct. Func. **224**(3) (2019) 985–1008
67. Hilgetag, C.C., Beul, S.F., van Albada, S.J., Goulas, A.: An architectonic type principle integrates macroscopic cortico-cortical connections with intrinsic cortical circuits of the primate brain. Netw. Neurosci. **3**(4) (2019) 905–923
68. Herculano-Houzel, S., Watson, C., Paxinos, G.: Distribution of neurons in functional areas of the mouse cerebral cortex reveals quantitatively different cortical zones. Front. Neuroanat. **7** (2013) 35
69. Keller, D., Erö, C., Markram, H.: Cell densities in the mouse brain: a systematic review. Front. Neuroanat. **12** (2018) 83
70. Meyer, H.S., Wimmer, V.C., Oberlaender, M., de Kock, C.P., Sakmann, B., Helmstaedter, M.: Number and laminar distribution of neurons in a thalamocortical projection column of rat vibrissal cortex. Cereb. Cortex **20**(10) (2010) 2277–2286
71. Markram, H., Muller, E., Ramaswamy, S., Reimann, M.W., Abdellah, M., Sanchez, C.A., Ailamaki, A., Alonso-Nanclares, L., Antille, N., Arsever, S., Kahou, G.A.A., Berger, T.K., Bilgili, A., Buncic, N., Chalimourda, A., Chindemi, G., Courcol, J.D., Delalondre, F., Delattre, V., Druckmann, S., Dumusc, R., Dynes, J., Eilemann, S., Gal, E., Gevaert, M.E., Ghobril, J.P., Gidon, A., Graham, J.W., Gupta, A., Haenel, V., Hay, E., Heinis, T., Hernando, J.B., Hines, M., Kanari, L., Keller, D., Kenyon, J., Khazen, G., Kim, Y., King, J.G., Kisvarday, Z., Kumbhar, P., Lasserre, S., Bé, J.V.L., Magalhães, B.R., Merchán-Pérez, A., Meystre, J., Morrice, B.R., Muller, J., Muñoz-Céspedes, A., Muralidhar, S., Muthurasa, K., Nachbaur, D., Newton, T.H., Nolte, M., Ovcharenko, A., Palacios, J., Pastor, L., Perin, R., Ranjan, R., Riachi, I., Rodríguez, J.R., Riquelme, J.L., Rössert, C., Sfyrakis, K., Shi, Y., Shillcock, J.C., Silberberg, G., Silva, R., Tauheed, F., Telefont, M., Toledo-Rodriguez, M., Tränkler, T., Geit, W.V., Díaz, J.V., Walker, R., Wang, Y., Zaninetta, S.M., DeFelipe, J., Hill, S.L., Segev, I., Schürmann, F.: Reconstruction and simulation of neocortical microcircuitry. Cell **163**(2) (October 2015) 456–492
72. Braitenberg, V.: Brain size and number of neurons: an exercise in synthetic neuroanatomy. J. Comput. Neurosci. **10**(1) (2001) 71–77
73. Dale, A.M., Fischl, B., Sereno, M.I.: Cortical surface-based analysis: I. Segmentation and surface reconstruction. NeuroImage **9**(2) (1999) 179–194
74. Fischl, B., Sereno, M.I., Dale, A.M.: Cortical surface-based analysis: II: inflation, flattening, and a surface-based coordinate system. NeuroImage **9**(2) (1999) 195–207
75. Thompson, P.M., Toga, A.W.: A framework for computational anatomy. Computing and Visualization in Science **5**(1) (2002) 13–34

76. Stephan, K., Kamper, L., Bozkurt, A., Burns, G., Young, M., Kötter, R.: Advanced database methodology for the collation of connectivity data on the macaque brain (CoCoMac). Phil. Trans. R. Soc. B **356** (2001) 1159–1186
77. Stephan, K.E., Zilles, K., Kötter, R.: Coordinate–independent mapping of structural and functional data by objective relational transformation (ort). Phil. Trans. R. Soc. B **355**(1393) (2000) 37–54
78. Dräger, U., Olsen, J.F.: Ganglion cell distribution in the retina of the mouse. Investigative Ophthalmology & Visual Science **20**(3) (1981) 285–293
79. Stone, J., Rapaport, D.H., Williams, R.W., Chalupa, L.: Uniformity of cell distribution in the ganglion cell layer of prenatal cat retina: implications for mechanisms of retinal development. Developmental Brain Research **2**(2) (1981) 231–242
80. Curcio, C.A., Allen, K.A.: Topography of ganglion cells in human retina. J. Compar. Neurol. **300**(1) (1990) 5–25
81. Wässle, H., Grünert, U., Martin, P.R., Boycotts, B.B.: Immunocytochemical characterization and spatial distribution of midget bipolar cells in the macaque monkey retina. Vision Research **34**(5) (1994) 561–579
82. Euler, T., Wässle, H.: Immunocytochemical identification of cone bipolar cells in the rat retina. J. Compar. Neurol. **361**(3) (1995) 461–478
83. Shand, J., Chin, S.M., Harman, A.M., Moore, S., Collin, S.P.: Variability in the location of the retinal ganglion cell area centralis is correlated with ontogenetic changes in feeding behavior in the black bream, Acanthopagrus butcheri (Sparidae, Teleostei). Brain, Behavior and Evolution **55**(4) (2000) 176–190
84. Ahmad, A., Spear, P.D.: Effects of aging on the size, density, and number of rhesus monkey lateral geniculate neurons. J. Compar. Neurol. **334**(4) (1993) 631–643
85. Mitra, N.: Quantitative analysis of cell types in mammalian neo-cortex. J. Anat. **89**(Pt 4) (1955) 467–483
86. Sloper, J., Hiorns, R., Powell, T.P.S.: A qualitative and quantitative electron microscopic study of the neurons in the primate motor and somatic sensory cortices. Phil. Trans. R. Soc. B **285**(1006) (1979) 141–171
87. Cozzi, B., De Giorgio, A., Peruffo, A., Montelli, S., Panin, M., Bombardi, C., Grandis, A., Pirone, A., Zambenedetti, P., Corain, L., Granato, A.: The laminar organization of the motor cortex in monodactylous mammals: a comparative assessment based on horse, chimpanzee, and macaque. Brain Struct. Func. **222**(6) (2017) 2743–2757
88. Turner, E.C., Young, N.A., Reed, J.L., Collins, C.E., Flaherty, D.K., Gabi, M., Kaas, J.H.: Distributions of cells and neurons across the cortical sheet in Old World macaques. Brain, Behavior and Evolution **88**(1) (2016) 1–13
89. McDonald, A.J.: Cytoarchitecture of the central amygdaloid nucleus of the rat. J. Compar. Neurol. **208**(4) (1982) 401–418
90. Stepniewska, I., Kaas, J.H.: Architectonic subdivisions of the inferior pulvinar in New World and Old World monkeys. Vis. Neurosci. **14**(6) (1997) 1043–1060
91. Voogd, J., Glickstein, M.: The anatomy of the cerebellum. Trends Cogn. Sci. **2**(9) (1998) 307–313
92. Duvernoy, H.M.: The human hippocampus: functional anatomy, vascularization and serial sections with MRI. Springer (2005)
93. Schumann, T., Erő, C., Gewaltig, M.O., Delalondre, F.J.: Towards simulating data-driven brain models at the point neuron level on petascale computers. In Di Napoli, E., Hermanns, M.A., Iliev, H., Lintermann, A., Peyser, A., eds.: High-Performance Scientific Computing: First JARA-HPC Symposium, JHPCS 2016, Aachen, Germany, October 4–5, 2016, Revised Selected Papers. Volume 10164., Springer (2017) 160–169

94. Romero-Garcia, R., Atienza, M., Clemmensen, L.H., Cantero, J.L.: Effects of network resolution on topological properties of human neocortex. NeuroImage **59**(4) (2012) 3522–3532
95. de Reus, M.A., Van den Heuvel, M.P.: The parcellation-based connectome: limitations and extensions. NeuroImage **80** (2013) 397–404
96. Pyka, M., Klatt, S., Cheng, S.: Parametric anatomical modeling: a method for modeling the anatomical layout of neurons and their projections. Front. Neuroanat. **8** (2014) 91
97. Anastassiou, C.A., Koch, C.: Ephaptic coupling to endogenous electric field activity: why bother? Curr. Opin. Neurobiol. **31** (2015) 95–103
98. Jirsa, V.K., Jantzen, K.J., Fuchs, A., Kelso, J.S.: Neural field dynamics on the folded three-dimensional cortical sheet and its forward EEG and MEG. In: Biennial International Conference on Information Processing in Medical Imaging, Springer (2001) 286–299
99. Hagen, E., Dahmen, D., Stavrinou, M.L., Lindén, H., Tetzlaff, T., van Albada, S.J., Grün, S., Diesmann, M., Einevoll, G.T.: Hybrid scheme for modeling local field potentials from point-neuron networks. Cereb. Cortex **26**(12) (October 2016) 4461–4496
100. Hagen, E., Næss, S., Ness, T.V., Einevoll, G.T.: Multimodal modeling of neural network activity: computing LFP, ECoG, EEG, and MEG signals with LFPy 2.0. Front. Neuroinform. **12** (2018) 92
101. Casali, S., Marenzi, E., Medini, K.C., Casellato, C., D'Angelo, E.: Reconstruction and simulation of a scaffold model of the cerebellar network. Front. Neuroinform. **13** (2019) 37
102. Lundqvist, M., Rehn, M., Djurfeldt, M., Lansner, A.: Attractor dynamics in a modular network model of neocortex. Network: Comput. Neural Systems **17**(3) (January 2006) 253–276
103. Johansson, C., Lansner, A.: Imposing biological constraints onto an abstract neocortical attractor network model. Neural Comput. **19** (2007) 1871–1896
104. Traub, R.D., Kopell, N., Bibbig, A., Buhl, E.H., LeBeau, F.E., Whittington, M.A.: Gap junctions between interneuron dendrites can enhance synchrony of gamma oscillations in distributed networks. J. Neurosci. **21**(23) (2001) 9478–9486
105. DeWeerdt, S.: How to map the brain. Nature **571**(7766) (2019) S6
106. Binzegger, T., Douglas, R.J., Martin, K.A.C.: A quantitative map of the circuit of cat primary visual cortex. J. Neurosci. **39**(24) (2004) 8441–8453
107. Denk, W., Horstmann, H.: Serial block-face scanning electron microscopy to reconstruct three-dimensional tissue nanostructure. PLOS Biol. **2**(11) (2004)
108. Denk, W., Strickler, J.H., Webb, W.W.: Two-photon laser scanning microscopy. Science **248** (1990) 73–76
109. Economo, M.N., Clack, N.G., Lavis, L.D., Gerfen, C.R., Svoboda, K., Myers, E.W., Chandrashekar, J.: A platform for brain-wide imaging and reconstruction of individual neurons. eLife **5** (2016) e10566
110. Winnubst, J., Bas, E., Ferreira, T.A., Wu, Z., Economo, M.N., Edson, P., Arthur, B.J., Bruns, C., Rokicki, K., Schauder, D., et al.: Reconstruction of 1,000 projection neurons reveals new cell types and organization of long-range connectivity in the mouse brain. Cell **179**(1) (2019) 268–281
111. Alonso-Nanclares, L., Gonzalez-Soriano, J., Rodriguez, J., DeFelipe, J.: Gender differences in human cortical synaptic density. Proc. Natl. Acad. Sci. USA **105**(38) (2008) 14615–14619

112. Alonso-Nanclares, L., Kastanauskaite, A., Rodriguez, J.R., Gonzalez-Soriano, J., DeFelipe, J.: A stereological study of synapse number in the epileptic human hippocampus. Front. Neuroanat. **5** (2011) 8
113. Zheng, Z., Lauritzen, J.S., Perlman, E., Robinson, C.G., Nichols, M., Milkie, D., Torrens, O., Price, J., Fisher, C.B., Sharifi, N., et al.: A complete electron microscopy volume of the brain of adult drosophila melanogaster. Cell **174**(3) (2018) 730–743
114. Kasthuri, N., Hayworth, K.J., Berger, D.R., Schalek, R.L., Conchello, J.A., Knowles-Barley, S., Lee, D., Vázquez-Reina, A., Kaynig, V., Jones, T.R., et al.: Saturated reconstruction of a volume of neocortex. Cell **162**(3) (July 2015) 648–661
115. Motta, A., Berning, M., Boergens, K.M., Staffler, B., Beining, M., Loomba, S., Hennig, P., Wissler, H., Helmstaedter, M.: Dense connectomic reconstruction in layer 4 of the somatosensory cortex. Science **366**(6469) (2019) eaay3134
116. Spruston, N., Jaffe, D.B.: Dendritic attenuation of synaptic potentials and currents: the role of passive membrane properties. Trends Neurosci. **17** (1994) 161–166
117. Murthy, V.N., Schikorski, T., Stevens, C.F., Zhu, Y.: Inactivity produces increases in neurotransmitter release and synapse size. Neuron **32**(4) (2001) 673–682
118. Harris, K.M., Fiala, J.C., Ostroff, L.: Structural changes at dendritic spine synapses during long-term potentiation. Phil. Trans. R. Soc. B **358**(1432) (2003) 745–748
119. Kwon, T., Sakamoto, M., Peterka, D.S., Yuste, R.: Attenuation of synaptic potentials in dendritic spines. Cell Reports **20**(5) (2017) 1100–1110
120. Rodriguez-Moreno, J., Porrero, C., Rollenhagen, A., Rubio-Teves, M., Casas-Torremocha, D., Alonso-Nanclares, L., Yakoubi, R., Santuy, A., Merchan-Pérez, A., DeFelipe, J., Lübke, J.: Area-specific synapse structure in branched posterior nucleus axons reveals a new level of complexity in thalamocortical networks. J. Neurosci. **40**(13) (2020) 2663–2679
121. Debanne, D., Boudkkazi, S., Campanac, E., Cudmore, R.H., Giraud, P., Fronzaroli-Molinieres, L., Carlier, E., Caillard, O.: Paired-recordings from synaptically coupled cortical and hippocampal neurons in acute and cultured brain slices. Nat. Protoc. **3**(10) (2008) 1559
122. Berry, M., Pentreath, V.: Criteria for distinguishing between monosynaptic and polysynaptic transmission. Brain Res. **105**(1) (1976) 1–20
123. Sedigh-Sarvestani, M., Vigeland, L., Fernandez-Lamo, I., Taylor, M.M., Palmer, L.A., Contreras, D.: Intracellular, in vivo, dynamics of thalamocortical synapses in visual cortex. J. Neurosci. **37**(21) (2017) 5250–5262
124. Markram, H., Lübke, J., Frotscher, M., Roth, A., Sakmann, B.: Physiology and anatomy of synaptic connections between thick tufted pyramidal neurons in the developing rat neocortex. J. Physiol. (Lond) **500**(2) (1997) 409–440
125. Thomson, A.M., West, D.C., Wang, Y., Bannister, A.P.: Synaptic connections and small circuits involving excitatory and inhibitory neurons in layer 2-5 of adult rat and cat neocortex: Triple intracellular recordings and biocytin labelling in vitro. Cereb. Cortex **12**(9) (September 2002) 936–953
126. Song, S., Sjöström, P., Reigl, M., Nelson, S., Chklovskii, D.: Highly nonrandom features of synaptic connectivity in local cortical circuits. PLOS Biol. **3**(3) (2005) e68
127. Perin, R., Berger, T.K., Markram, H.: A synaptic organizing principle for cortical neuronal groups. Proc. Natl. Acad. Sci. USA **108**(13) (March 2011) 5419–5424

128. Kodandaramaiah, S.B., Flores, F.J., Holst, G.L., Singer, A.C., Han, X., Brown, E.N., Boyden, E.S., Forest, C.R.: Multi-neuron intracellular recording in vivo via interacting autopatching robots. eLife **7** (2018) e24656
129. Callaway, E.M., Katz, L.C.: Photostimulation using caged glutamate reveals functional circuitry in living brain slices. Proc. Natl. Acad. Sci. USA **90**(16) (August 1993) 7661–7665
130. Nikolenko, V., Poskanzer, K.E., Yuste, R.: Two-photon photostimulation and imaging of neural circuits. Nat. Methods **4**(11) (2007) 943–950
131. Noguchi, J., Nagaoka, A., Watanabe, S., Ellis-Davies, G.C., Kitamura, K., Kano, M., Matsuzaki, M., Kasai, H.: In vivo two-photon uncaging of glutamate revealing the structure–function relationships of dendritic spines in the neocortex of adult mice. J. Physiol. (Lond.) **589**(10) (2011) 2447–2457
132. Dantzker, J.L., Callaway, E.M.: Laminar sources of synaptic input to cortical inhibitory interneurons and pyramidal neurons. Nat Neurosci **3**(7) (2000) 701–707
133. Schubert, D., Kötter, R., Zilles, K., Luhmann, H.J., Staiger, J.F.: Cell type-specific circuits of cortical layer IV spiny neurons. J. Neurosci. **23**(7) (Apr 2003) 2961–2970
134. Hooks, B.M., Hires, S.A., Zhang, Y.X., Huber, D., Petreanu, L., Svoboda, K., Shepherd, G.M.G.: Laminar analysis of excitatory local circuits in vibrissal motor and sensory cortical areas. PLOS Biol. **9**(1) (2011) e1000572
135. Lanciego, J.L., Wouterlood, F.G.: A half century of experimental neuroanatomical tracing. J. Chem. Neuroanat. **42**(3) (2011) 157–183
136. Saleeba, C., Dempsey, B.R., Le, S., Goodchild, A.K., McMullan, S.: A student's guide to neural circuit tracing. Front. Neurosci. **13** (2019) 897
137. Galuske, R., Schlote, W., Bratzke, H., Singer, W.: Interhemispheric asymmetries of the modular structure in human temporal cortex. Science **5486**(289) (2000) 1946–1949
138. Tardif, E., Clarke, S.: Intrinsic connectivity of human auditory areas: a tracing study with DiI. Eur. J. Neurosci. **13**(5) (2001) 1045–1050
139. Seehaus, A.K., Roebroeck, A., Chiry, O., Kim, D.S., Ronen, I., Bratzke, H., Goebel, R., Galuske, R.A.: Histological validation of DW-MRI tractography in human postmortem tissue. Cereb. Cortex **23**(2) (2013) 442–450
140. Kuypers, H., Ugolini, G.: Viruses as transneuronal tracers. Trends Neurosci. **13**(2) (1990) 71–75
141. Köbbert, C., Apps, R., Bechmann, I., Lanciego, J.L., Mey, J., Thanos, S.: Current concepts in neuroanatomical tracing. Prog. Neurobiol. **62**(4) (2000) 327–351
142. Chen, X., Sun, Y.C., Zhan, H., Kebschull, J.M., Fischer, S., Matho, K., Huang, Z.J., Gillis, J., Zador, A.M.: High-throughput mapping of long-range neuronal projection using in situ sequencing. Cell **179**(3) (2019) 772–786
143. Markov, N.T., Ercsey-Ravasz, M.M., Ribeiro Gomes, A.R., Lamy, C., Magrou, L., Vezoli, J., Misery, P., Falchier, A., Quilodran, R., Gariel, M.A., Sallet, J., Gamanut, R., Huissoud, C., Clavagnier, S., Giroud, P., Sappey-Marinier, D., Barone, P., Dehay, C., Toroczkai, Z., Knoblauch, K., Van Essen, D.C., Kennedy, H.: A weighted and directed interareal connectivity matrix for macaque cerebral cortex. Cereb. Cortex **24**(1) (2014) 17–36
144. Majka, P., Chaplin, T.A., Yu, H.H., Tolpygo, A., Mitra, P.P., Wójcik, D.K., Rosa, M.G.: Towards a comprehensive atlas of cortical connections in a primate brain: Mapping tracer injection studies of the common marmoset into a reference digital template. J. Compar. Neurol. **524**(11) (2016) 2161–2181

145. Majka, P., Bai, S., Bakola, S., Bednarek, S., Chan, J.M., Jermakow, N., Passarelli, L., Reser, D.H., Theodoni, P., Worthy, K.H., et al.: Open access resource for cellular-resolution analyses of corticocortical connectivity in the marmoset monkey. Nat. Commun. **11**(1) (2020) 1–14
146. Galletti, C., Gamberini, M., Kutz, D.F., Fattori, P., Luppino, G., Matelli, M.: The cortical connections of area V6: an occipito-parietal network processing visual information. Eur. J. Neurosci. **13**(8) (2001) 1572–1588
147. Bakker, R., Thomas, W., Diesmann, M.: CoCoMac 2.0 and the future of tract-tracing databases. Front. Neuroinform. **6** (2012) 30
148. Scannell, J., Blakemore, C., Young, M.: Analysis of connectivity in the cat cerebral cortex. J. Neurosci. **15**(2) (1995) 1463–1483
149. Oh, S.W., Harris, J.A., Ng, L., Winslow, B., Cain, N., Mihalas, S., Wang, Q., Lau, C., Kuan, L., Henry, A.M., et al.: A mesoscale connectome of the mouse brain. Nature **508**(7495) (2014) 207–214
150. Basser, P.J., Mattiello, J., LeBihan, D.: MR diffusion tensor spectroscopy and imaging. Biophys. J. **66**(1) (1994) 259–267
151. Beul, S.F., Grant, S., Hilgetag, C.C.: A predictive model of the cat cortical connectome based on cytoarchitecture and distance. Brain Struct. Func. **220**(6) (2015) 3167–3184
152. Knock, S., McIntosh, A., Sporns, O., Kötter, R., Hagmann, P., Jirsa, V.: The effects of physiologically plausible connectivity structure on local and global dynamics in large scale brain models. J. Neurosci. Methods **1**(183) (2009) 86–94
153. Maier-Hein, K.H., Neher, P.F., Houde, J.C., Côté, M.A., Garyfallidis, E., Zhong, J., Chamberland, M., Yeh, F.C., Lin, Y.C., Ji, Q., et al.: The challenge of mapping the human connectome based on diffusion tractography. Nat. Commun. **8** (2017) 1349
154. Calabrese, E., Badea, A., Cofer, G., Qi, Y., Johnson, G.A.: A diffusion MRI tractography connectome of the mouse brain and comparison with neuronal tracer data. Cereb. Cortex (2015) bhv121
155. Peters, A., Feldman, M.L.: The projection of the lateral geniculate nucleus to area 17 of the rat cerebral cortex. I. General description. J. Neurocytol. **5**(1) (1976) 63–84
156. Braitenberg, V., Schüz, A.: Anatomy of the Cortex: Statistics and Geometry. Springer-Verlag, Berlin, Heidelberg, New York (1991)
157. Rees, C.L., Moradi, K., Ascoli, G.A.: Weighing the evidence in Peters' rule: Does neuronal morphology predict connectivity? Trends Neurosci. **40**(2) (February 2017) 63–71
158. Packer, A.M., McConnell, D.J., Fino, E., Yuste, R.: Axo-dendritic overlap and laminar projection can explain interneuron connectivity to pyramidal cells. Cereb. Cortex **23**(12) (2013) 2790–2802
159. Merchán-Pérez, A., Rodríguez, J.R., González, S., Robles, V., DeFelipe, J., Larrañaga, P., Bielza, C.: Three-dimensional spatial distribution of synapses in the neocortex: a dual-beam electron microscopy study. Cereb. Cortex **24**(6) (2014) 1579–1588
160. Lee, W.C.A., Bonin, V., Reed, M., Graham, B.J., Hood, G., Glattfelder, K., Reid, R.C.: Anatomy and function of an excitatory network in the visual cortex. Nature **532**(7599) (2016) 370–374
161. Barbas, H.: Pattern in the laminar origin of corticocortical connections. J. Comp. Neurol. **252**(3) (1986) 415–422

162. Hilgetag, C.C., Grant, S.: Cytoarchitectural differences are a key determinant of laminar projection origins in the visual cortex. NeuroImage **51**(3) (2010) 1006–1017
163. Beul, S.F., Barbas, H., Hilgetag, C.C.: A predictive structural model of the primate connectome. Sci. Rep. **7**(43176) (2017) 1–12
164. von Economo, C.F., Van Bogaert, L.: L'architecture cellulaire normale de l'écorce cérébrale. Paris, Masson et Cie (1927)
165. Goulas, A., Majka, P., Rosa, M.G., Hilgetag, C.C.: A blueprint of mammalian cortical connectomes. PLOS Biol. **17**(3) (2019) e2005346
166. Goulas, A., Zilles, K., Hilgetag, C.C.: Cortical gradients and laminar projections in mammals. Trends Neurosci. **41**(11) (2018) 775–788
167. Felleman, D.J., Van Essen, D.C.: Distributed hierarchical processing in the primate cerebral cortex. Cereb. Cortex **1** (1991) 1–47
168. Sanides, F.: Functional architecture of motor and sensory cortices in primates in the light of a new concept of neocortex evolution. In Noback, C., Montagna, W., eds.: The Primate Brain: Advances in Primatology. New York, Appleton-Century-Crofts Educational Division/Meredith Corporation (1970) 137–208
169. Lerch, J.P., Worsley, K., Shaw, W.P., Greenstein, D.K., Lenroot, R.K., Giedd, J., Evans, A.C.: Mapping anatomical correlations across cerebral cortex (MACACC) using cortical thickness from MRI. NeuroImage **31**(3) (2006) 993–1003
170. Gong, G., He, Y., Chen, Z.J., Evans, A.C.: Convergence and divergence of thickness correlations with diffusion connections across the human cerebral cortex. NeuroImage **59**(2) (2012) 1239–1248
171. Alexander-Bloch, A., Giedd, J.N., Bullmore, E.: Imaging structural co-variance between human brain regions. Nat. Rev. Neurosci. **14**(5) (2013) 322–336
172. Packer, A.M., Yuste, R.: Dense, unspecific connectivity of neocortical parvalbumin-positive interneurons: A canonical microcircuit for inhibition? J. Neurosci. **31**(37) (September 2011) 13260–13271
173. Levy, R.B., Reyes, A.D.: Spatial profile of excitatory and inhibitory synaptic connectivity in mouse primary auditory cortex. J. Neurosci. **32**(16) (April 2012) 5609–5619
174. Voges, N., Schüz, A., Aertsen, A., Rotter, S.: A modeler's view on the spatial structure of horizontal cortical connectivity in the neocortex. submitted (September 2008)
175. Ercsey-Ravasz, M., Markov, N.T., Lamy, C., Essen, D.C.V., Knoblauch, K., Toroczkai, Z., Kennedy, H.: A predictive network model of cerebral cortical connectivity based on a distance rule. Neuron **80**(1) (2013) 184–197
176. Young, M.P.: Objective analysis of the topological organization of the primate cortical visual system. Nature **358**(6382) (1992) 152–155
177. Klyachko, V.A., Stevens, C.F.: Connectivity optimization and the positioning of cortical areas. Proc. Natl. Acad. Sci. USA **100**(13) (2003) 7937–7941
178. Granovetter, M.: The strength of weak ties: A network theory revisited. Sociological Theory (1983) 201–233
179. Goulas, A., Betzel, R.F., Hilgetag, C.C.: Spatiotemporal ontogeny of brain wiring. Sci. Adv. **5**(6) (2019) eaav9694
180. Ko, H., Hofer, S.B., Pichler, B., Buchanan, K.A., Sjöström, P.J., Mrsic-Flogel, T.D.: Functional specificity of local synaptic connections in neocortical networks. Nature **473**(7345) (May 2011) 87–91
181. Billeh, Y.N., Cai, B., Gratiy, S.L., Dai, K., Iyer, R., Gouwens, N.W., Abbasi-Asl, R., Jia, X., Siegle, J.H., Olsen, S.R., et al.: Systematic integration of structural and

functional data into multi-scale models of mouse primary visual cortex. Neuron (2020)
182. Jouve, B., Rosenstiehl, P., Imbert, M.: A mathematical approach to the connectivity between the cortical visual areas of the macaque monkey. Cereb. Cortex **8**(1) (1998) 28–39
183. Mejias, J.F., Murray, J.D., Kennedy, H., Wang, X.J.: Feedforward and feedback frequency-dependent interactions in a large-scale laminar network of the primate cortex. Sci. Adv. **2**(11) (2016) e1601335
184. Theodoni, P., Majka, P., Reser, D.H., Wójcik, D.K., Rosa, M.G., Wang, X.J.: Structural attributes and principles of the neocortical connectome in the marmoset monkey. bioRxiv (2020)
185. Vértes, P.E., Alexander-Bloch, A.F., Gogtay, N., Giedd, J.N., Rapoport, J.L., Bullmore, E.T.: Simple models of human brain functional networks. Proc. Natl. Acad. Sci. USA **109**(15) (2012) 5868–5873
186. Betzel, R.F., Avena-Koenigsberger, A., Goñi, J., He, Y., De Reus, M.A., Griffa, A., Vértes, P.E., Mišic, B., Thiran, J.P., Hagmann, P., et al.: Generative models of the human connectome. NeuroImage **124** (2016) 1054–1064
187. Chen, Y., Zhang, Z.K., He, Y., Zhou, C.: A large-scale high-density weighted structural connectome of the macaque brain acquired by predicting missing links. Cereb. Cortex (2020) bhaa060
188. Nisbach, F., Kaiser, M.: Developmental time windows for spatial growth generate multiple-cluster small-world networks. Eur. Phys. J. B **58**(2) (2007) 185–191
189. Barbas, H., García-Cabezas, M.Á.: How the prefrontal executive got its stripes. Curr. Opin. Neurobiol. **40** (2016) 125–134
190. Bayer, S.A., Altman, J.: Directions in neurogenetic gradients and patterns of anatomical connections in the telencephalon. Prog. Neurobiol. **29**(1) (1987) 57–106
191. Prinz, A.A., Bucher, D., Marder, E.: Similar network activity from disparate circuit parameters. Nat. Neurosci. **7** (2004) 1345–1352
192. Aertsen, A.M.H.J., Gerstein, G.L., Habib, M.K., Palm, G.: Dynamics of Neuronal Firing Correlation: Modulation of 'Effective Connectivity'. J. Neurophysiol. **61**(5) (1989) 900–917
193. van Albada, S.J., Helias, M., Diesmann, M.: Scalability of asynchronous networks is limited by one-to-one mapping between effective connectivity and correlations. PLOS Comput. Biol. **11**(9) (2015) e1004490
194. Maynard, E.M., Nordhausen, C.T., Normann, R.A.: The Utah intracortical electrode array: A recording structure for potential brain-computer interfaces. EEG Clin. Neurophysiol. **102**(3) (March 1997) 228–239
195. Jun, J.J., Steinmetz, N.A., et al.: Fully integrated silicon probes for high-density recording of neural activity. Nature **551** (2017) 232–236
196. Zimmermann, J., Goebel, R., De Martino, F., Van de Moortele, P.F., Feinberg, D., Adriany, G., Chaimow, D., Shmuel, A., Uğurbil, K., Yacoub, E.: Mapping the organization of axis of motion selective features in human area mt using high-field fmri. PLOS One **6**(12) (2011) e28716
197. De Martino, F., Zimmermann, J., Muckli, L., Ugurbil, K., Yacoub, E., Goebel, R.: Cortical depth dependent functional responses in humans at 7T: improved specificity with 3D GRASE. PLOS One **8**(3) (2013) e60514
198. Ostojic, S., Brunel, N., Hakim, V.: How connectivity, background activity, and synaptic properties shape the cross-correlation between spike trains. J. Neurosci. **29**(33) (2009) 10234–10253

199. English, D.F., McKenzie, S., Evans, T., Kim, K., Yoon, E., Buzsáki, G.: Pyramidal cell-interneuron circuit architecture and dynamics in hippocampal networks. Neuron **96**(2) (October 2017) 505–520.e7
200. Pastore, V.P., Massobrio, P., Godjoski, A., Martinoia, S.: Identification of excitatory-inhibitory links and network topology in large-scale neuronal assemblies from multi-electrode recordings. PLOS Comput. Biol. **14**(8) (08 2018) 1–25
201. Kobayashi, R., Kurita, S., Kurth, A., Kitano, K., Mizuseki, K., Diesmann, M., Richmond, B.J., Shinomoto, S.: Reconstructing neuronal circuitry from parallel spike trains. Nat. Commun. **10**(1) (2019) 1–13
202. Cohen, M.R., Kohn, A.: Measuring and interpreting neuronal correlations. Nat. Rev. Neurosci. **14**(7) (July 2011) 811–819
203. Helias, M., Tetzlaff, T., Diesmann, M.: Echoes in correlated neural systems. New J. Phys. **15** (2013) 023002
204. Helias, M., Tetzlaff, T., Diesmann, M.: The correlation structure of local cortical networks intrinsically results from recurrent dynamics. PLOS Comput. Biol. **10**(1) (2014) e1003428
205. Grytskyy, D., Helias, M., Diesmann, M.: Reconstruction of network connectivity in the irregular firing regime. In: Proceedings 10th Göttingen Meeting of the German Neuroscience Society. (2013) 1192–1193
206. Casadiego, J., Maoutsa, D., Timme, M.: Inferring network connectivity from event timing patterns. Phys. Rev. Lett. **121** (Aug 2018) 054101
207. van Bussel, F., Kriener, B., Timme, M.: Inferring synaptic connectivity from spatio-temporal spike patterns. Front. Comput. Neurosci. **5**(3) (2011) doi: 10.3389/fncom.2011.00003
208. Zaytsev, Y., Morrison, A., Deger, M.: Reconstruction of recurrent synaptic connectivity of thousands of neurons from simulated spiking activity. J. Comput. Neurosci. **39** (2015) 77–103
209. Gerhard, F., Kispersky, T., Gutierrez, G.J., Marder, E., Kramer, M., Eden, U.: Successful reconstruction of a physiological circuit with known connectivity from spiking activity alone. PLOS Comput. Biol. **9**(7) (2013) e1003138
210. Stringer, C., Pachitariu, M., Steinmetz, N.A., Okun, M., Bartho, P., Harris, K.D., Sahani, M., Lesica, N.A.: Inhibitory control of correlated intrinsic variability in cortical networks. eLife **5** (dec 2016) e19695
211. Druckmann, S., Banitt, Y., Gidon, A.A., Schürmann, F., Markram, H., Segev, I.: A novel multiple objective optimization framework for constraining conductance-based neuron models by experimental data. Front. Neurosci. **1**(1) (2007) 7–18 doi:10.3389/neuro.01/1.1.001.2007.
212. Rossant, C., Goodman, D.F., Platkiewicz, J., Brette, R.: Automatic fitting of spiking neuron models to electrophysiological recordings. Front. Neuroinform. **4** (2010)
213. Carlson, K., Nageswaran, J., Dutt, N., Krichmar, J.: An efficient automated parameter tuning framework for spiking neural networks. Front. Neurosci. **8** (2014) 10
214. Diaz-Pier, S., Naveau, M., Butz-Ostendorf, M., Morrison, A.: Automatic generation of connectivity for large-scale neuronal network models through structural plasticity. Front. Neuroanat. **10** (2016) 57
215. Paninski, L.: Maximum likelihood estimation of cascade point-process neural encoding models. Network: Comput. Neural Systems **15**(4) (2004) 243–262
216. Pillow, J.W., Paninski, L., Uzzell, V.J., Simoncelli, E.P., Chichilnisky, E.J.: Prediction and decoding of retinal ganglion cell responses with a probabilistic spiking model. J. Neurosci. **25**(47) (2005) 11003–11013

217. Ladenbauer, J., McKenzie, S., English, D., et al.: Inferring and validating mechanistic models of neural microcircuits based on spike-train data. Nat. Commun. **10**(4933) (2019)
218. René, A., Longtin, A., Macke, J.H.: Inference of a mesoscopic population model from population spike trains. Neural Comput. **32**(8) (2020) 1448–1498
219. Bittner, S.R., Palmigiano, A., Piet, A.T., Duan, C.A., Brody, C.D., Miller, K.D., Cunningham, J.P.: Interrogating theoretical models of neural computation with deep inference. bioRxiv (2019)
220. Gonçalves, P.J., Lueckmann, J.M., Deistler, M., Nonnenmacher, M., Öcal, K., Bassetto, G., Chintaluri, C., Podlaski, W.F., Haddad, S.A., Vogels, T.P., Greenberg, D.S., Macke, J.H.: Training deep neural density estimators to identify mechanistic models of neural dynamics. bioRxiv (2020)
221. Betzel, R.F., Bassett, D.S.: Generative models for network neuroscience: prospects and promise. J. R. Soc. Interface **14**(136) (2017) 20170623
222. Mooney, C.Z., Duval, R.D.: Bootstrapping: A nonparametric approach to statistical inference. Number 95. Sage (1993)
223. Gutzen, R., von Papen, M., Trensch, G., Quaglio, P., Grün, S., Denker, M.: Reproducible neural network simulations: Statistical methods for model validation on the level of network activity data. Front. Neuroinform. **12** (2018) 90
224. Sahasranamam, A., Vlachos, I., Aertsen, A., Kumar, A.: Dynamical state of the network determines the efficacy of single neuron properties in shaping the network activity. Sci. Rep. **6**(1) (2016) 1–16
225. Larsen, L., Griffin, L.D., Gräßel, D., Witte, O.W., Axer, H.: Polarized light imaging of white matter architecture. Microsc. Res. Techniq. **70**(10) (2007) 851–863
226. Axer, M., Grässel, D., Kleiner, M., Dammers, J., Dickscheid, T., Reckfort, J., Hütz, T., Eiben, B., Pietrzyk, U., Zilles, K., et al.: High-resolution fiber tract reconstruction in the human brain by means of three-dimensional polarized light imaging. Front. Neuroinform. **5**(34) (2011)
227. Menzel, M., Axer, M., Amunts, K., De Raedt, H., Michielsen, K.: Diattenuation Imaging reveals different brain tissue properties. Sci. Rep. **9**(1) (2019) 1–12
228. Fornito, A., Arnatkevičiūtė, A., Fulcher, B.D.: Bridging the gap between connectome and transcriptome. Trends Cogn. Sci. **23**(1) (2019) 34–50
229. Barabási, D.L., Barabási, A.L.: A genetic model of the connectome. Neuron **105**(3) (2020) 435–445
230. Timonidis, N., Bakker, R., Tiesinga, P.: Prediction of a cell-class-specific mouse mesoconnectome using gene expression data. Neuroinformatics (2020) 1–16
231. Chklovskii, D.B.: Synaptic connectivity and neuronal morphology: Two sides of the same coin. Neuron **43** (2004) 609–617
232. Samu, D., Seth, A.K., Nowotny, T.: Influence of wiring cost on the large-scale architecture of human cortical connectivity. PLOS Comput. Biol. **10**(4) (2014) e1003557
233. Van Albada, S.J., Rennie, C.J., Robinson, P.A.: Variability of model-free and model-based quantitative measures of EEG. J. Integr. Neurosci. **6**(02) (2007) 279–307
234. Gordon, E.M., Laumann, T.O., Adeyemo, B., Petersen, S.E.: Individual variability of the system-level organization of the human brain. Cereb. Cortex **27**(1) (2017) 386–399
235. Xu, T., Sturgeon, D., Ramirez, J.S., Froudist-Walsh, S., Margulies, D.S., Schroeder, C.E., Fair, D.A., Milham, M.P.: Interindividual variability of func-

tional connectivity in awake and anesthetized rhesus macaque monkeys. Biol. Psychiatry Cogn. Neurosci. Neuroimaging **4**(6) (2019) 543–553
236. Schmidt, M., Bakker, R., Shen, K., Bezgin, G., Diesmann, M., van Albada, S.J.: A multi-scale layer-resolved spiking network model of resting-state dynamics in macaque visual cortical areas. PLOS Comput. Biol. **14**(10) (2018) e1006359
237. Proix, T., Bartolomei, F., Guye, M., Jirsa, V.K.: Individual brain structure and modelling predict seizure propagation. Brain **140**(3) (2017) 641–654
238. Morrison, A., Mehring, C., Geisel, T., Aertsen, A., Diesmann, M.: Advancing the boundaries of high connectivity network simulation with distributed computing. Neural Comput. **17**(8) (2005) 1776–1801

B

Trans-thalamic connections

Summary:

It has been shown that cortico-cortical communication can occur via intermediate thalamic neurons in mice. In this journal club article we explore the potential implications of such trans-thalamic connections and ellaborate on its potential implications, including the possibility that such connections support cortico-cortical communication in primates.

Journal Club

Editor's Note: These short reviews of recent *JNeurosci* articles, written exclusively by students or postdoctoral fellows, summarize the important findings of the paper and provide additional insight and commentary. If the authors of the highlighted article have written a response to the Journal Club, the response can be found by viewing the Journal Club at www.jneurosci.org. For more information on the format, review process, and purpose of Journal Club articles, please see http://www.jneurosci.org/content/jneurosci-journal-club.

Trans-thalamic Pathways: Strong Candidates for Supporting Communication between Functionally Distinct Cortical Areas

Barna Zajzon[1,2] **and Aitor Morales-Gregorio**[1,3]

[1]Institute of Neuroscience and Medicine (INM-6), Computational and Systems Neuroscience and Institute for Advanced Simulation (IAS-6), Theoretical Neuroscience and JARA-Institut Brain Structure-Function Relationships (INM-10) 52428, [2]Department of Psychiatry, Psychotherapy and Psychosomatics, RWTH Aachen University, Germany 52074, and [3]RWTH Aachen University, Germany 52074

Review of Mo and Sherman.

Introduction

The thalamus was long considered a passive relay of sensory information with little or no active role in higher cognitive functions. However, mounting evidence suggests that thalamic nuclei form complex loops with the cortex and are involved in a myriad of cognitive processes, including attention and working memory (Ward, 2013). Although first-order thalamic nuclei (e.g., lateral geniculate nucleus) play a key role in the transmission of ascending sensory input to the cortex, higher-order nuclei (e.g., pulvinar or mediodorsal nucleus) are believed to be involved in sustaining and modulating communication within and between cortical regions (Guillery, 1995). Understanding the functional role of such nuclei in a mechanistic manner requires, in addition to behavioral experiments, a detailed anatomical and physiological mapping of the thalamocortical circuitry.

Thalamic cells in higher-order nuclei have been shown to act as an intermediary between cortical areas, providing a cortico-thalamo-cortical pathway that augments direct communication (Sherman and Guillery, 2013). These circuits are known to originate in layer 5 (L5) of cortex, but their target layers have not been exhaustively studied. They are typically arranged in parallel to strong feedforward corticocortical projections, which in mice usually originate in cortical layers L2/3 and L5 and project to most layers in the target area (e.g., S1→ M1; Porter and White, 1983). Such trans-thalamic connections involve class 1 or "driver" glutamatergic synapses (Sherman and Guillery, 2011), which produce large, depressing postsynaptic currents, activate mainly ionotropic receptors, and have large and small boutons targeting proximal dendrites. Because of their high probability of neurotransmitter release, class 1 terminals are thought to be efficient information carriers and can reliably elicit thalamic action potentials (Rovó et al., 2012).

Instead of just passively relaying signals from one cortical area to another, the thalamus might manipulate information arriving from L5 in a context-dependent manner. Specifically, thalamic nuclei have been suggested to dynamically construct task-relevant functional circuits (Nakajima and Halassa, 2017), as well as change the effective connectivity between cortical regions through targeted gain modulation (Jaramillo et al., 2019). However, anatomical evidence of such trans-thalamic pathways was previously limited to projections between primary and secondary auditory, visual, and somatosensory cortices (Theyel et al., 2010; Sherman and Guillery, 2013). Given that most higher-order nuclei receive connections from and project to multiple cortical areas, a natural question is whether the existence of parallel routes is a general organizing principle, possibly linking functionally distinct cortical areas.

To answer this question, Mo and Sherman (2019) used trans-synaptic viral tracing in mouse slice preparations to demonstrate the existence of an indirect pathway between the primary somatosensory cortex (S1) and the primary motor cortex (M1) through the thalamic posterior medial (POm) nucleus, a higher-order nucleus in the rodent whisker system that is actively involved in processing sensorimotor information (Krieger and Groh, 2015). Specifically, the authors found M1-projecting POm neurons receiving input from S1 (45.3%) and S2 (26.3%). Although neurons projecting to POm were more abundant in L6 of S1, the

Received March 22, 2019; revised June 14, 2019; accepted July 5, 2019.

This work was supported by the European Union's Horizon 2020 Framework Programme for Research and Innovation (Grant 785907, Human Brain Project SGA2), the Priority Program (SPP 2041 "Computational Connectomics") of the Deutsche Forschungsgemeinschaft (S.J. van Albada: AL 2041/1-1), and the Helmholtz Alliance through the Initiative and Networking Fund of the Helmholtz Association. We thank Drs. Renato Duarte and Sacha van Albada for supportive comments and constructive suggestions.

The authors declare no competing financial interests.

Correspondence should be addressed to Barna Zajzon at b.zajzon@fz-juelich.de.

https://doi.org/10.1523/JNEUROSCI.0656-19.2019

hypothesized projections from L5 of S1 were also found. This establishes the presence of a previously unknown S1 L5→POm→M1 pathway involving direct monosynaptic connections.

To determine the synaptic properties of the circuit, first the inputs to POm from L5 of S1 were targeted for optogenetic and electrical stimulation. Whole-cell patch recordings of M1-projecting POm neurons revealed large postsynaptic currents and strong paired-pulse depression after both axonal and dendritic stimulation of S1 L5 inputs to POm cells, which were eliminated after blocking ionotropic glutamate receptors. Additionally, bouton size analysis revealed the presence of large and small terminals. As noted above, these characteristics are indicative of class 1 synapses (Sherman and Guillery, 2013). Because the synapses from POm to M1 displayed similar features, the authors concluded that the S1 L5→POm→M1 pathway includes exclusively class 1 synapses. These results are consistent with previous reports (Reichova and Sherman, 2004) and suggest a robust trans-thalamic pathway through POm with synapses well suited for the successful relay of somatosensory information (Sherman and Guillery, 2011).

It should be noted that the net effect of S1 L5 activity on M1 via the POm remains somewhat unclear because the optogenetic stimulation did not exclusively target these projections. Indeed, only one-third of the M1-projecting POm neurons responded to stimulating the inputs from S1 L5 (see Figure 3 in Mo and Sherman, 2019), suggesting that S1 is not the main source of input to many of the POm neurons that project to M1. Furthermore, as the authors point out, even though most cells across all layers in M1 responded strongly to POm stimulation (see Figure 4 in Mo and Sherman, 2019), this activity was mediated by a larger set of POm cells than those receiving input from S1 L5. Nevertheless, the fact that S1 L5→POm connections are relatively strong despite their sparsity, with unitary postsynaptic potentials from L5 axons able to elicit action potentials in POm cells (Mease et al., 2016c), supports the case for the S1 L5→POm→M1 pathway being an efficient trans-thalamic communication channel.

A similar path between S1 and S2 through the POm was deemed critical for activation of S2 by stimulation of barrel field (S1 L5B) *in vitro* (Theyel et al., 2010). In that study, information transfer from S1 to S2 continued after permanent disruption of the direct corticocortical projections, and was only interrupted by chemically induced thalamic inhibition. A similar experimental protocol could be applied to investigate the efficacy and the functional role of the S1 L5→POm→M1 pathway in sensorimotor processing. Such an experiment would be valuable given that single-whisker-evoked responses in M1 were shown to depend on S1 activation (Ferezou et al., 2007).

Although the exact nature of information transmitted along the corticocortical and trans-thalamic S1 to M1 projections is unknown, Mo and Sherman (2019) suggest that these two routes carry different content because they originate from non-overlapping populations in S1 (Petrof et al., 2012). Furthermore, the corticothalamic cells in L5 projecting to POm also branch to subcortical motor centers (e.g., the brainstem; Krieger and Groh, 2015), indicating that these might transmit efference copies of motor-related instructions originating in S1 and forwarded to higher cortical centers (Sherman and Guillery, 2013). This hypothesis is supported by the fact that S1 was shown to participate in motor control and directly drive whisker retraction even after M1 inactivation (Matyas et al., 2010), suggesting that whisking-related instructions originate not only in M1 but also in S1.

However, information routed through the highly state-dependent POm is bound to undergo significant modulation as the nucleus can dynamically reconfigure its circuits depending on the arousal level (Sobolewski et al., 2015). During high alertness, POm combines not only descending cortical input, but also sensory information ascending via the paralemniscal pathway through the spinal trigeminal nucleus (mainly SP5i; Krieger and Groh, 2015). These inputs converge onto single thalamic cells, leading to timing-sensitive nonlinear responses driven by coincident L5 and SP5i input, much like an "AND-gate" (Groh et al., 2014). Such convergence zones are spatially restricted, highlighting possible functional subdivisions within the nucleus (Mease et al., 2016b). Given that Mo and Sherman (2019) did not find any spatial localization of M1-projecting neurons in POm (their Figure 3*G*), these could overlap with the aforementioned convergence zones. If confirmed, then the trans-thalamic route might enable the integration of sensory input and cortical output in a time- and behavior-dependent manner en route to M1.

This pathway and its potential role in the temporally precise integration of sensory information become more interesting when one considers that transmission through the POm is potentially contingent on motor activity in M1 (Urbain and Deschênes, 2007). At rest, whisker responses in POm are gated through inhibitory projections from the zona incerta (ZI), an area also innervated by projections from M1. During active palpation, however, corticofugal messages from M1 modulating vibrissa motion can inhibit parts of ZI, thus creating a window of disinhibition in POm and enabling information flow. At the same time, POm neurons might relay signals back to S1, because some cells project to multiple sensory and motor cortical areas (Ohno et al., 2012). Reciprocal projections between POm and L5 of S1 have been suggested to participate in a cortico-thalamo-cortical loop that amplifies and temporally sustains certain relevant sensory content, similar to a vibrissal attention mechanism (Mease et al., 2016a). The POm and the newly mapped pathway are therefore strategically positioned to play a major role in sensorimotor processing, fusing both cortical and sensory input under top-down modulation from higher motor centers.

The findings of Mo and Sherman (2019) support the possibility that trans-thalamic pathways bridging functionally distinct cortical areas are a common theme in the mammalian brain. Although it is unclear whether such connections exist in other species, some tract-tracing studies in primates provide supportive evidence. In macaques, a potential route related to oculomotor processing might involve the mediodorsal nucleus, which receives input from area 7m (medial posterior parietal cortex) and projects to area 8 (the frontal eye fields; Leichnetz, 2001). This putative pathway is also parallel to strong feedforward corticocortical connections (Markov et al., 2014), akin to the circuits observed in mice (Sherman and Guillery, 2013). In addition to mediating information flow from sensory to higher cortical areas, thalamic nuclei may also play an important role in executive control of motor behavior. McFarland and Haber (2002) proposed that direct signaling from prefrontal and higher motor centers to primary motor areas is complemented by a series of feedforward projections via the thalamus. This raises the possibility that thalamic pathways are involved in regulating information flow in both directions of the cortical hierarchy.

In conclusion, Mo and Sherman (2019) have confirmed the existence of a strong S1 L5→POm→M1 pathway involving

class 1 glutamatergic synapses, which is the first trans-thalamic connection revealed between functionally distinct cortical regions. Although this supports the case for such circuits being an integral link in cortical communication, further evidence is necessary to establish their generality beyond the sensorimotor areas and across species. Whereas the functional implications of this particular pathway remain unclear, it might play a key part in motor control, mediate sensory attention, or act as an integrator of peripheral and cortical signals. Narrowing down these potential roles is a challenging and exciting prospect that will require additional experimental work, including both quantitative thalamic connectivity data and targeted behavioral experiments on sensorimotor processing.

References

Ferezou I, Haiss F, Gentet LJ, Aronoff R, Weber B, Petersen CC (2007) Spatiotemporal dynamics of cortical sensorimotor integration in behaving mice. Neuron 56:907–923.

Groh A, Bokor H, Mease RA, Plattner VM, Hangya B, Stroh A, Deschênes M, Acsády L (2014) Convergence of cortical and sensory driver inputs on single thalamocortical cells. Cereb Cortex 24:3167–3179.

Guillery RW (1995) Anatomical evidence concerning the role of the thalamus in corticocortical communication: a brief review. J Anat 187:583–592.

Jaramillo J, Mejias JF, Wang XJ (2019) Engagement of pulvino-cortical feedforward and feedback pathways in cognitive computations. Neuron 101:321–336.e9.

Krieger P, Groh A (2015) Sensorimotor integration in the whisker system. New York: Springer.

Leichnetz GR (2001) Connections of the medial posterior parietal cortex (area 7m) in the monkey. Anatomical Record 263:215–236.

Markov NT, Ercsey-Ravasz MM, Ribeiro Gomes AR, Lamy C, Magrou L, Vezoli J, Misery P, Falchier A, Quilodran R, Gariel MA, Sallet J, Gamanut R, Huissoud C, Clavagnier S, Giroud P, Sappey-Marinier D, Barone P, Dehay C, Toroczkai Z, Knoblauch K, Van Essen DC, Kennedy H (2014) A weighted and directed interareal connectivity matrix for macaque cerebral cortex. Cereb Cortex 24:17–36.

Matyas F, Sreenivasan V, Marbach F, Wacongne C, Barsy B, Mateo C, Aronoff R, Petersen CC (2010) Motor control by sensory cortex. Science 330:1240–1243.

McFarland NR, Haber SN (2002) Thalamic relay nuclei of the basal ganglia form both reciprocal and nonreciprocal cortical connections, linking multiple frontal cortical areas. J Neurosci 22:8117–8132.

Mease RA, Metz M, Groh A (2016a) Cortical sensory responses are enhanced by the higher-order thalamus. Cell Rep 14:208–215.

Mease RA, Sumser A, Sakmann B, Groh A (2016b) Cortical dependence of whisker responses in posterior medial thalamus in vivo. Cereb Cortex 26:3534–3543.

Mease RA, Sumser A, Sakmann B, Groh A (2016c) Corticothalamic spike transfer via the L5B-POm pathway in vivo. Cereb Cortex 26:3461–3475.

Mo C, Sherman SM (2019) A sensorimotor pathway via higher-order thalamus. J Neurosci 39:692–704.

Nakajima M, Halassa MM (2017) Thalamic control of functional cortical connectivity. Curr Opin Neurobiol 44:127–131.

Ohno S, Kuramoto E, Furuta T, Hioki H, Tanaka YR, Fujiyama F, Sonomura T, Uemura M, Sugiyama K, Kaneko T (2012) A morphological analysis of thalamocortical axon fibers of rat posterior thalamic nuclei: a single neuron tracing study with viral vectors. Cereb Cortex 22:2840–2857.

Petrof I, Viaene AN, Sherman SM (2012) Two populations of corticothalamic and interareal corticocortical cells in the subgranular layers of the mouse primary sensory cortices. J Comp Neurol 520:1678–1686.

Porter LL, White EL (1983) Afferent and efferent pathways of the vibrissal region of primary motor cortex in the mouse. J Comp Neurol 214:279–289.

Reichova I, Sherman SM (2004) Somatosensory corticothalamic projections: distinguishing drivers from modulators. J Neurophysiol 92:2185–2197.

Rovó Z, Ulbert I, Acsády L (2012) Drivers of the primate thalamus. J Neurosci 32:17894–17908.

Sherman SM, Guillery RW (2013) Functional connections of cortical areas: a new view from the thalamus. Cambridge, MA: MIT.

Sherman SM, Guillery RW (2011) Distinct functions for direct and transthalamic corticocortical connections. J Neurophysiol 106:1068–1077.

Sobolewski A, Kublik E, Swiejkowski DA, Kamiński J, Wróbel A (2015) Alertness opens the effective flow of sensory information through rat thalamic posterior nucleus. Eur J Neurosci 41:1321–1331.

Theyel BB, Llano DA, Sherman SM (2010) The corticothalamocortical circuit drives higher-order cortex in the mouse. Nat Neurosci 13:84–88.

Urbain N, Deschênes M (2007) Motor cortex gates vibrissal responses in a thalamocortical projection pathway. Neuron 56:714–725.

Ward LM (2013) The thalamus: gateway to the mind. Wiley Interdiscip Rev Cogn Sci 4:609–622.

References

Abbott, L. F. (2008). Theoretical neuroscience rising. *Neuron*, 60(3):489–495.

Abeles, M. (1991). *Corticonics*. Cambridge University Press, Cambridge, England.

Antolík, J., Monier, C., Frégnac, Y., and Davison, A. P. (2018). A comprehensive data-driven model of cat primary visual cortex. preprint, Neuroscience.

Aria, M. M. (2020). Calcium imaging and optical electrophysiology. In *Electrophysiology Measurements for Studying Neural Interfaces*, pages 105–141. Elsevier.

Arkhipov, A., Gouwens, N. W., Billeh, Y. N., Gratiy, S., Iyer, R., Wei, Z., Xu, Z., Abbasi-Asl, R., Berg, J., Buice, M., Cain, N., da Costa, N., de Vries, S., Denman, D., Durand, S., Feng, D., Jarsky, T., Lecoq, J., Lee, B., Li, L., Mihalas, S., Ocker, G. K., Olsen, S. R., Reid, R. C., Soler-Llavina, G., Sorensen, S. A., Wang, Q., Waters, J., Scanziani, M., and Koch, C. (2018). Visual physiology of the layer 4 cortical circuit in silico. *PLOS Computational Biology*, 14(11):e1006535.

Atapour, N., Majka, P., Wolkowicz, I. H., Malamanova, D., Worthy, K. H., and Rosa, M. G. (2019). Neuronal Distribution Across the Cerebral Cortex of the Marmoset Monkey (Callithrix jacchus). *Cerebral cortex*, 29(9):3836–3863.

Barabási, A.-L. and Albert, R. (1999). Emergence of Scaling in Random Networks. *Science*, 286(5439):509–512.

Barabási, A.-L. and Oltvai, Z. N. (2004). Network biology: understanding the cell's functional organization. *Nature Reviews Genetics*, 5(2):101–113.

Bethe, A. T. J. (1903). *Allgemeine Anatomie und Physiologie des Nervensystems*. Leipzig, Verlag von Georg Thieme.

Beul, S. F., Barbas, H., and Hilgetag, C. C. (2017). A Predictive Structural Model of the Primate Connectome. *Scientific Reports*, 7(May 2016):1–12.

Billeh, Y. N., Cai, B., Gratiy, S. L., Dai, K., Iyer, R., Gouwens, N. W., Abbasi-Asl, R., Jia, X., Siegle, J. H., Olsen, S. R., Koch, C., Mihalas, S., and Arkhipov, A. (2020). Systematic Integration of Structural and Functional Data into Multi-scale Models of Mouse Primary Visual Cortex. *Neuron*, 106(3):388–403.e18.

Brunel, N. (2000). Dynamics of Sparsely Connected Networks of Excitatory and Inhibitory Spiking Neurons. *Journal of Computational Neuroscience*, 8(3):183–208.

Buzsáki, G. and Mizuseki, K. (2014). The log-dynamic brain: how skewed distributions affect network operations. *Nature Reviews Neuroscience*, 15(4):264–278.

Cadwell, C. R., Bhaduri, A., Mostajo-Radji, M. A., Keefe, M. G., and Nowakowski, T. J. (2019). Development and arealization of the cerebral cortex. *Neuron*, 103(6):980–1004.

Calabrese, E., Badea, A., Coe, C. L., Lubach, G. R., Shi, Y., Styner, M. A., and Johnson, G. A. (2015). A diffusion tensor MRI atlas of the postmortem rhesus macaque brain. *NeuroImage*, 117:408–416.

Carlson, K. D., Nageswaran, J. M., Dutt, N., and Krichmar, J. L. (2014). An efficient automated parameter tuning framework for spiking neural networks. *Frontiers in Neuroscience*, 8.

Chaudhuri, R., Knoblauch, K., Gariel, M.-A., Kennedy, H., and Wang, X.-J. (2015a). A Large-Scale Circuit Mechanism for Hierarchical Dynamical Processing in the Primate Cortex. *Neuron*, 88(2):419–431.

Chaudhuri, R., Knoblauch, K., Gariel, M. A., Kennedy, H., and Wang, X. J. (2015b). A Large-Scale Circuit Mechanism for Hierarchical Dynamical Processing in the Primate Cortex. *Neuron*, 88(2):419–431.

Chen, X., Morales-Gregorio, A., Sprenger, J., Kleinjohann, A., Sridhar, S., van Albada, S., Grün, S., and Roelfsema, P. (2021). 1024-channel electrophysiological recordings in macaque v1 and v4 during resting state. *G-Node. https://doi.org/10.12751/g-node.i20kyh*.

Cobb, M. (2020). *The Idea of the Brain: A History*. Profile.

de Haan, M. J., Brochier, T., Grün, S., Riehle, A., and Barthélemy, F. V. (2018). Real-time visuomotor behavior and electrophysiology recording setup for use with humans and monkeys. *Journal of Neurophysiology*, 120(2):539–552.

Deco, G., Jirsa, V., McIntosh, A. R., Sporns, O., and Kötter, R. (2009). Key role of coupling, delay, and noise in resting brain fluctuations. *Proceedings of the National Academy of Sciences*, 106(25):10302–10307.

Deco, G. and Jirsa, V. K. (2012). Ongoing Cortical Activity at Rest: Criticality, Multistability, and Ghost Attractors. *Journal of Neuroscience*, 32(10):3366–3375.

Douglas, R. J. and Martin, K. A. (1991). A functional microcircuit for cat visual cortex. *The Journal of Physiology*, 440(1):735–769.

Douglas, R. J. and Martin, K. A. (2004). NEURONAL CIRCUITS OF THE NEOCORTEX. *Annual Review of Neuroscience*, 27(1):419–451.

Douglas, R. J., Martin, K. A., and Whitteridge, D. (1989). A Canonical Microcircuit for Neocortex. *Neural Computation*, 1(4):480–488.

Druckmann, S. (2007). A novel multiple objective optimization framework for constraining conductance-based neuron models by experimental data. *Frontiers in Neuroscience*, 1(1):7–18.

Dura-Bernal, S., Griffith, E. Y., Barczak, A., O'Connell, M. N., McGinnis, T., Schroeder, C. E., Lytton, W. W., Lakatos, P., and Neymotin, S. A. (2022a). Data-driven multiscale model of macaque auditory thalamocortical circuits reproduces in vivo dynamics. preprint, Neuroscience.

Dura-Bernal, S., Neymotin, S. A., Suter, B. A., Dacre, J., Schiemann, J., Duguid, I., Shepherd, G. M., and Lytton, W. W. (2022b). Multiscale model of primary motor cortex circuits predicts in vivo cell type-specific, behavioral state-dependent dynamics. preprint, Neuroscience.

Ercsey-Ravasz, M., Markov, N. T., Lamy, C., Van Essen, D. C., Knoblauch, K., Toroczkai, Z., and Kennedy, H. (2013). A Predictive Network Model of Cerebral Cortical Connectivity Based on a Distance Rule. *Neuron*, 80(1):184–197.

Erdős, P. and Rényi, A. (1959). On random graphs. *Publicationes Mathematicae*, 6:290–297.

Froudist-Walsh, S., Bliss, D. P., Ding, X., Rapan, L., Niu, M., Knoblauch, K., Zilles, K., Kennedy, H., Palomero-Gallagher, N., and Wang, X.-J. (2021). A dopamine gradient controls access to distributed working memory in the large-scale monkey cortex. *Neuron*, 109(21):3500–3520.e13.

Gerstner, W., Kistler, W. M., Naud, R., and Paninski, L. (2014). *Neuronal Dynamics. From Single Neurons to Networks and Models of Cognition*. Cambridge University Press.

Girvan, M. and Newman, M. E. J. (2002). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, 99(12):7821–7826.

Glees, P. (1988). *The human brain*. Cambridge University Press.

Golgi, C. (1886). *Sulla fina anatomia degli organi centrali del sistema nervoso: studi*. Ulrico Hoepli.

Gutzen, R., von Papen, M., Trensch, G., Quaglio, P., Grün, S., and Denker, M. (2018). Reproducible Neural Network Simulations: Statistical Methods for Model Validation on the Level of Network Activity Data. *Frontiers in Neuroinformatics*, 12:90.

Haider, B. (2006). Neocortical Network Activity In Vivo Is Generated through a Dynamic Balance of Excitation and Inhibition. *Journal of Neuroscience*, 26(17):4535–4545.

Hebb, D. O. (1949). *The Organization of Behavior: A Neuropsychological Theory*. New York: Wiley and Sons.

Hodgkin, A. L. and Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology*, 117(4):500–544.

Kandel, E. R. and Schwartz, J. H. (1981). *Principles of Neural Science*. Elsevier/North- Holland, New York.

Konorski, J. (1948). *Conditioned reflexes and neuron organization*. Cambridge University Press.

Ladenbauer, J., McKenzie, S., English, D., et al. (2019). Inferring and validating mechanistic models of neural microcircuits based on spike-train data. *Nature Communications*, 10(4933).

Lamy, C. M. and Chatton, J.-Y. (2011). Optical probing of sodium dynamics in neurons and astrocytes. *NeuroImage*, 58(2):572–578.

Layer, M., Senk, J., Essink, S., van Meegen, A., Bos, H., and Helias, M. (2022). NNMT: Mean-Field Based Analysis Tools for Neuronal Network Models. *Frontiers in Neuroinformatics*, 16:835657.

Lindsay, G. (2021). *Models of the mind*. Bloomsbury Sigma, London, England.

Majka, P., Chaplin, T. A., Yu, H. H., Tolpygo, A., Mitra, P. P., Wójcik, D. K., and Rosa, M. G. (2016). Towards a comprehensive atlas of cortical connections in a primate brain: Mapping tracer injection studies of the common marmoset into a reference digital template. *Journal of Comparative Neurology*, 524(11):2161–2181.

Majka, P., Wolkowicz, I. H., Malamanova, D., Worthy, K. H., and Rosa, M. G. P. (2018). Neuronal Distribution Across the Cerebral Cortex of the Marmoset Monkey (Callithrix jacchus). pages 1–28.

Markov, N. T., Ercsey-Ravasz, M., Ribeiro Gomes, A. R., Lamy, C., Magrou, L., Vezoli, J., Misery, P., Falchier, A., Quilodran, R., Gariel, M. A., Sallet, J., Gamanut, R., Huissoud, C., Clavagnier, S., Giroud, P., Sappey-Marinier, D., Barone, P., Dehay, C., Toroczkai, Z., Knoblauch, K., Van Essen, D. C., and Kennedy, H. (2012). A weighted and directed interareal connectivity matrix for macaque cerebral cortex. *Cerebral Cortex*, 24(1):17–36.

Markov, N. T., Vezoli, J., Chameau, P., Falchier, A., Quilodran, R., Huissoud, C., Lamy, C., Misery, P., Giroud, P., Ullman, S., Barone, P., Dehay, C., Knoblauch, K., and Kennedy, H. (2014). Anatomy of hierarchy: Feedforward and feedback pathways in macaque visual cortex. *Journal of Comparative Neurology*, 522(1):225–259.

Markram, H., Muller, E., Ramaswamy, S., Reimann, M., Abdellah, M., Sanchez, C., Ailamaki, A., Alonso-Nanclares, L., Antille, N., Arsever, S., Kahou, G., Berger, T., Bilgili, A., Buncic, N., Chalimourda, A., Chindemi, G., Courcol, J.-D., Delalondre, F., Delattre, V., Druckmann, S., Dumusc, R., Dynes, J., Eilemann, S., Gal, E., Gevaert, M., Ghobril, J.-P., Gidon, A., Graham, J., Gupta, A., Haenel, V., Hay, E., Heinis, T., Hernando, J., Hines, M., Kanari, L., Keller, D., Kenyon, J., Khazen, G., Kim, Y., King, J., Kisvarday, Z., Kumbhar, P., Lasserre, S., Le Bé, J.-V., Magalhães, B., Merchán-Pérez, A., Meystre, J., Morrice, B., Muller, J., Muñoz-Céspedes, A., Muralidhar, S., Muthurasa, K., Nachbaur, D., Newton, T., Nolte, M., Ovcharenko, A., Palacios, J., Pastor, L., Perin, R., Ranjan, R., Riachi, I., Rodríguez, J.-R., Riquelme, J., Rössert, C., Sfyrakis, K., Shi, Y., Shillcock, J., Silberberg, G., Silva, R., Tauheed, F., Telefont, M., Toledo-Rodriguez, M., Tränkler, T., Van Geit, W., Díaz, J., Walker, R., Wang, Y., Zaninetta, S., DeFelipe, J., Hill, S., Segev, I., and Schürmann, F. (2015). Reconstruction and Simulation of Neocortical Microcircuitry. *Cell*, 163(2):456–492.

Mejias, J. F., Murray, J. D., Kennedy, H., and Wang, X.-J. (2016). Feedforward and feedback frequency-dependent interactions in a large-scale laminar network of the primate cortex. *Science Advances*, 2(11):e1601335.

Mikula, S., Trotts, I., Stone, J. M., and Jones, E. G. (2007). Internet-enabled high-resolution brain mapping and virtual microscopy. *NeuroImage*, 35(1):9–15.

Mochizuki, Y., Onaga, T., Shimazaki, H., Shimokawa, T., Tsubo, Y., Kimura, R., Saiki, A., Sakai, Y., Isomura, Y., Fujisawa, S., Shibata, K.-i., Hirai, D., Furuta, T., Kaneko, T., Takahashi, S., Nakazono, T., Ishino, S., Sakurai, Y., Kitsukawa, T., Lee, J. W., Lee, H., Jung, M. W., Babul, C., Maldonado, P. E., Takahashi, K., Arce-McShane, F. I., Ross, C. F., Sessle, B. J., Hatsopoulos, N. G., Brochier, T., Riehle, A., Chorley, P., Grün, S., Nishijo, H., Ichihara-Takeda, S., Funahashi, S., Shima, K., Mushiake, H., Yamane, Y., Tamura, H., Fujita, I., Inaba, N., Kawano, K., Kurkin, S., Fukushima, K., Kurata, K., Taira, M., Tsutsui, K.-I., Ogawa, T., Komatsu, H., Koida, K., Toyama, K., Richmond, B. J., and Shinomoto, S. (2016). Similarity in neuronal firing regimes across mammalian species. *Journal of Neuroscience*, 36(21):5736–5747.

Naumann, L. B., Keijser, J., and Sprekeler, H. (2022). Invariant neural subspaces maintained by feedback modulation. *eLife*, 11:e76096.

Picco, N., García-Moreno, F., Maini, P. K., Woolley, T. E., and Molnár, Z. (2018). Mathematical Modeling of Cortical Neurogenesis Reveals that the Founder Population does not Necessarily Scale with Neurogenic Output. *Cerebral Cortex*, 28(7):2540–2550.

Potjans, T. C. and Diesmann, M. (2014). The Cell-Type Specific Cortical Microcircuit: Relating Structure and Activity in a Full-Scale Spiking Network Model. *Cerebral Cortex*, 24(3):785–806.

Prinz, A. A., Bucher, D., and Marder, E. (2004). Similar network activity from disparate circuit parameters. *Nature Neuroscience*, 7(12):1345–1352.

Quiroga, R. Q., Nadasdy, Z., and Ben-Shaul, Y. (2004). Unsupervised Spike Detection and Sorting with Wavelets and Superparamagnetic Clustering. *Neural Computation*, 16(8):1661–1687.

Rakic, P. (2002). Neurogenesis in adult primate neocortex: an evaluation of the evidence. *Nature Reviews Neuroscience*, 3(1):65–71.

Rakic, P. (2009). Evolution of the neocortex: a perspective from developmental biology. *Nature Reviews Neuroscience*, 10(10):724–735.

Ramón y Cajal, S. (1894). La fine structure des centres nerveux. *Proceedings of the Royal Society of London*, 55(331-335):444–468.

Reimann, M., Anastassiou, C., Perin, R., Hill, S. L., Markram, H., and Koch, C. (2013). A Biophysically Detailed Model of Neocortical Local Field Potentials Predicts the Critical Role of Active Membrane Currents. *Neuron*, 79(2):375–390.

Rimmele, T. S. and Chatton, J.-Y. (2014). A Novel Optical Intracellular Imaging Approach for Potassium Dynamics in Astrocytes. *PLoS ONE*, 9(10):e109243.

Rossant, C. (2010). Automatic fitting of spiking neuron models to electrophysiological recordings. *Frontiers in Neuroinformatics*, 4.

Schmidt, M., Bakker, R., Hilgetag, C. C., Diesmann, M., and van Albada, S. J. (2018a). Multi-scale account of the network structure of macaque visual cortex. *Brain Structure and Function*, 223(3):1409–1435.

Schmidt, M., Bakker, R., Shen, K., Bezgin, G., Diesmann, M., and van Albada, S. J. (2018b). A multi-scale layer-resolved spiking network model of resting-state dynamics in macaque visual cortical areas. *PLOS Computational Biology*, 14(10):e1006359.

Sejnowski, T. J., Churchland, P. S., and Movshon, J. A. (2014). Putting big data to good use in neuroscience. *Nature Neuroscience*, 17(11):1440–1441.

Shepherd, G. M. and Brayton, R. K. (1987). Logic operations are properties of computer-simulated interactions between excitable dendritic spines. *Neuroscience*, 21(1):151–165.

Soni, J. and Goodman, R. (2017). *A Mind at Play: How Claude Shannon Invented the Information Age*. Simon & Schuster.

Stringer, C., Pachitariu, M., Steinmetz, N., Carandini, M., and Harris, K. D. (2019). High-dimensional geometry of population responses in visual cortex. *Nature*, 571(7765):361–365.

Stringer, C., Pachitariu, M., Steinmetz, N. A., Okun, M., Bartho, P., Harris, K. D., Sahani, M., and Lesica, N. A. (2016). Inhibitory control of correlated intrinsic variability in cortical networks. *eLife*, 5:e19695.

Supèr, H. and Roelfsema, P. R. (2005). Chronic multiunit recordings in behaving animals: advantages and limitations. *Progress in Brain Research*, 147:263–282.

Traub, R. D., Contreras, D., Cunningham, M. O., Murray, H., LeBeau, F. E. N., Roopun, A., Bibbig, A., Wilent, W. B., Higley, M. J., and Whittington, M. A. (2005). Single-Column Thalamocortical Network Model Exhibiting Gamma Oscillations, Sleep Spindles, and Epileptogenic Bursts. *Journal of Neurophysiology*, 93(4):2194–2232.

van Albada, S. J., Rowley, A. G., Senk, J., Hopkins, M., Schmidt, M., Stokes, A. B., Lester, D. R., Diesmann, M., and Furber, S. B. (2018). Performance Comparison of the Digital Neuromorphic Hardware SpiNNaker and the Neural Network Simulation Software NEST for a Full-Scale Cortical Microcircuit Model. *Frontiers in Neuroscience*, 12:291.

van Vreeswijk, C. and Sompolinsky, H. (1996). Chaos in Neuronal Networks with Balanced Excitatory and Inhibitory Activity. *Science*, 274(5293):1724–1726.

van Vreeswijk, C. and Sompolinsky, H. (1998). Chaotic Balanced State in a Model of Cortical Circuits. *Neural Computation*, 10(6):1321–1371.

Watts, D. J. and Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440–442.

Wilson, H. R. and Cowan, J. D. (1972). Excitatory and Inhibitory Interactions in Localized Populations of Model Neurons. *Biophysical Journal*, 12(1):1–24.

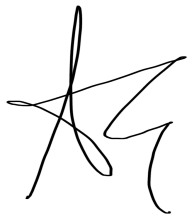
Wybo, W. A., Jordan, J., Ellenberger, B., Marti Mengual, U., Nevian, T., and Senn, W. (2021). Data-driven reduction of dendritic morphologies with preserved dendro-somatic responses. *eLife*, 10:e60936.

Zhu, W., Shelley, M., and Shapley, R. (2009). A neuronal network model of primary visual cortex explains spatial frequency selectivity. *Journal of Computational Neuroscience*, 26(2):271–287.

Offizielle Erklärung

Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation selbstständig und ohne die Benutzung anderer als der angegebenen Hilfsmittel und Literatur angefertigt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten und nicht veröffentlichten Werken dem Wortlaut oder dem Sinn nach entnommen wurden, sind als solche kenntlich gemacht. Ich versichere an Eides statt, dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie - abgesehen von unten angegebenen Teilpublikationen und eingebundenen Artikeln und Manuskripten - noch nicht veröffentlicht worden ist sowie, dass ich eine Veröffentlichung der Dissertation vor Abschluss der Promotion nicht ohne Genehmigung des Promotionsausschusses vornehmen werde. Die Bestimmungen dieser Ordnung sind mir bekannt. Darüber hinaus erkläre ich hiermit, dass ich die Ordnung zur Sicherung guter wissenschaftlicher Praxis und zum Umgang mit wissenschaftlichem Fehlverhalten der Universität zu Köln gelesen und sie bei der Durchführung der Dissertation zugrundeliegenden Arbeiten und der schriftlich verfassten Dissertation beachtet habe und verpflichte mich hiermit, die dort genannten Vorgaben bei allen wissenschaftlichen Tätigkeiten zu beachten und umzusetzen. Ich versichere, dass die eingereichte elektronische Fassung der eingereichten Druckfassung vollständig entspricht.

Teilpublikationen: Siehe Publications.



———————————

Jülich 2022 Aitor Morales-Gregorio

Publications

Chapter 1:
1024-channel electrophysiological recordings in macaque V1 and V4 during resting state

Xing Chen*, Aitor Morales-Gregorio*, Julia Sprenger, Alexander Kleinjohann, Shashwat Sridhar, Sacha J. van Albada, Sonja Grün and Pieter R. Roelfsema

***These authors contributed equally to this work**

published in *Scientific Data* (2022)

Chapter 2:
Feedback modulates neural manifolds and dimensionality in macaque V1 during resting state

Aitor Morales-Gregorio, Anno C. Kurth, Junji Ito, Alexander Kleinjohann, Frédéric V. Barthélemy, Thomas Brochier, Sonja Grün and Sacha J. van Albada

In preparation

Chapter 3:
Ubiquitous lognormal distribution of neuron densities across mammalian cerebral cortex

Aitor Morales-Gregorio*, Alexander van Meegen* and Sacha J. van Albada

***These authors contributed equally to this work**

published in *bioRxiv* (2022)

Chapter 4:
Quantification of layer-resolved neuron density and white matter distance in the macaque neocortex

Aitor Morales-Gregorio, Rembrandt Bakker and Sacha J. van Albada

In preparation

Chapter 5:
Activity-driven microconnectome estimation of macaque visuomotor cortices

Aitor Morales-Gregorio, Robin Gutzen, Paulina A. Dąbrowska, Alper Yegenoglu, Sandra Diaz-Pier, Sarah Palmis, Sofia Paneri, Alexandre René, Panagiotis Sapountzis, Markus Diesmann, Sonja Grün, Johanna Senk, Georgia G. Gregoriou, Bjørg E. Kilavik and Sacha J. van Albada

In preparation

Appendix A:
Bringing Anatomical Information into Neuronal Network Models

Sacha J. van Albada, Aitor Morales-Gregorio, Timo Dickscheid, Alexandros Goulas, Rembrandt Bakker, Sebastian Bludau, Günther Palm, Claus C. Hilgetag and Markus Diesmann

published in *Advances in Experimental Medicine and Biology* (2022) and *arXiv* (2020)

Appendix B:
Trans-thalamic Pathways: Strong Candidates for Supporting Communication between Functionally Distinct Cortical Areas

Barna Zajzon and Aitor Morales-Gregorio

published in *The Journal of Neuroscience* (2019)

Author's contribution to publications

Aitor Morales-Gregorio (**AMG**) collaborated with all of the following people (N=31, in alphabetical order) in the work that constitutes this dissertation:

Rembrandt Bakker (RB), Frédéric V. Barthélemy (FB), Sebastian Bludau (SB), Thomas Brochier (TB), Xing Chen (XC), Paulina A. Dąbrowska (PD), Sandra Diaz-Pier (SD), Timo Dickscheid (TD), Markus Diesmann (MD), Claus C. Hilgetag (CH), Alexandros Goulas (AG), Georgia G. Gregoriou (GG), Sonja Grün (SG), Robin Gutzen (RG), Junji Ito (JI), Bjørg E. Kilavik (BK), Alexander Kleinjohann (AKJ), Anno C. Kurth (AK), Günther Palm (GP), Sarah Palmis (SaP), Sofia Paneri (SoP), Alexandre René (AR), Pieter R. Roelfsema (PR), Panagiotis Sapountzis (PS), Johanna Senk (JSe), Julia Sprenger (JS), Shashwat Sridhar (SSr), Sacha J. van Albada (SvA), Alexander van Meegen (AvM), Alper Yegenoglu (AY), and Barna Zajzon (BZ).

The author deeply appreciates the help and support from all co-authors, without them this thesis would not have been possible.

Chapter 1:

AMG shares the first authorship together with XC. XC and PR conceptualized the project. XC set up the experimental apparatus, recorded the data, and contributed standalone matlab processing scripts for the raw data. **AMG** and XC gathered the metadata. **AMG** curated the metadata into a single descriptive file. **AMG** implemented a reproducible python workflow to processes the raw data, which was reviewed by JS, SSr and XC. **AMG** and JS set up the data sharing repository. **AMG** and AKJ wrote the novel cross-talk detection workflow. XC created Figures 1 and 3. **AMG** created Figures 2, 4, 5, 6, 7, and 8. XC wrote the initial manuscript. **AMG** contributed to writing all parts of the manuscript. The manuscript was reviewed by all authors. SG, SvA and PR supervised the project and procured the funding.

Although not reflected in the manuscript **AMG** and SSr also spike sorted the data with a semi-automatic workflow. A part that had to be left out due to varying quality between the two macaques included in the study.

Chapter 2:

AMG is the first author. **AMG**, JI and AK conceptualized the study. **AMG** curated and processed the data from macaques L and A. AKJ curated and processed the data from macaque Y, with minor inputs from **AMG**. **AMG** and AK performed the dimensionality analysis, statistical testing and Granger causality analysis. AK performed the spiking neuron network simulations and **AMG** analyzed the results. **AMG** created all figures, with feedback from all other authors. **AMG** wrote the initial manuscript, all other authors provided feedback. JI, SG, SvA supervised the project. SG and SvA procured the funding.

Chapter 3:

AMG shares the first authorship together with AvM. **AMG** collected, curated and analyzed the data, and created all the figures. AvM developed the model and focused on the statistical tests, with feedback from **AMG** and SvA. **AMG** and AvM wrote the first draft of this manuscript, and it was jointly revised by all authors. SvA supervised the project and procured the funding.

Chapter 4:

AMG is the first author. **AMG** conceptualized the study. **AMG** curated the raw data. **AMG** and RB pre-processed the raw data. **AMG** manually processed the masks and constructed the initial volume. RB developed the workflow and reconstructed the 3D volume. **AMG** implemented the widget to extract the samples and extracted the counting samples. **AMG** trained the neuron counting software and analyzed the results. **AMG** conceptualized and implemented the white matter distance measuring heuristic. **AMG** produced all plots and wrote the initial manuscript. RB and SvA reviewed the manuscript and supervised the project. SvA procured the funding.

Chapter 5:

AMG is the first author. **AMG**, BK and SvA conceptualized the study. SaP and BK collected and spike sorted the resting state data from macaque T. SoP, PS and GG collected and spike sorted the resting state data from macaque F. **AMG** spike sorted the data from macaque L. **AMG** and PD performed behavioral segmentation for macaque T. **AMG** curated the data, with help from PD. **AMG** and RG developed and implemented the statistical comparison framework. **AMG**, AY, SD and PD implemented the optimization algorithm, and AR provided useful feedback. **AMG** adapted the spiking neuron models into the optimization framework, with assistance from AR and JSe. **AMG** run all the data analysis, optimization and produced all the plots. **AMG** wrote the initial manuscript with inputs from PD, RG and AR. All authors provided feedback on the manuscript, plot design and several poster presentations. BK, GG, SG, MD and SvA procured the funding. **AMG** and SvA coordinated the project. SvA supervised the project.

Appendix A:

AMG is the second author. SvA conceptualized the study. SvA wrote the bulk of the initial manuscript. **AMG** wrote major contributions to the predictive connectomics and validation parts. TD and SB provided Figure 1 and contributed to the atlasing sections in the manuscript. RB created Figure 2. **AMG** created Figures 3 and 5. AG provided Figure 4. MD wrote parts of the discussion. All authors contributed to the final version of the manuscript text. SvA and MD procured the funding.

Appendix B:

AMG is the second author. BZ conceptualized the study. Both authors wrote and reviewed the manuscript. SvA provided feedback on the manuscript, but the article type did not allow for senior researchers as co-authors.

Author's contribution to open software

Neo

Neo is a Python package for working with electrophysiology data in Python.

`https://github.com/NeuralEnsemble/python-neo`

The author contributed functions for the processing and handling of analog signals, reported bugs and actively engaged with the community. The Neo package was extensively used in Chapter 1.

Elephant - Electrophysiology Analysis Toolkit

Elephant package analyses all sorts of neurophysiological data: spike trains, LFP, analog signals.

`https://github.com/NeuralEnsemble/elephant`

The author contributed functions for the calculation of spike variability (LvR), autocorrelation timescale, synchronous artifact detection, spectral tools for coherence and causality and more. The author contributed to documentation, bug reports and memory efficiency of the software, as well as actively engaging with the community. The Elephant package was extensively used in Chapter 1, Chapter 2 and Chapter 5.

The Neural Simulation Tool - NEST

NEST is a simulator for spiking neural network models that focuses on the dynamics, size and structure of neural systems rather than on the exact morphology of individual neurons.

`https://github.com/nest/nest-simulator`

The author contributed to the documentation and consistency of the Kernel parameters. The author engaged with the community and assisted the writing of some of the examples. The NEST package was extensively used in Chapter 2 and Chapter 5.

NetworkUnit

A SciUnit library for validation testing of spiking networks.

`https://github.com/INM-6/NetworkUnit/tree/v0.2`

The author actively contributed to all aspects of development for the 0.2 version. The author wrote functions, included functionalities, improved code style, improved efficiency, implemented caching, parallelization and more. The NetworkUnit package was extensively used in Chapter 5.

SciUnit

A Test-Driven Framework for Formally Validating Scientific Models Against Data

`https://github.com/scidash/sciunit`

The author contributed to improve the package by including caching options, for their use within NetworkUnit. The SciUnit package is the parent framework for NetworkUnit and therefore extensively used in Chapter 5.

PoSSum3

The Possum framework is a set of scripts and workflows for reconstructing three-dimensional images based on series of serial sections of different modalities and quality.

`https://github.com/pmajka/poSSum`

`https://github.com/morales-gregorio/poSSum3/tree/poSSum3`

The author updated the PoSSum framework from Python2 to Python3 (hence PoSSum3), improving the code base and enabling its use with modern software. The PoSSum3 framework was extensively used in Chapter 4.