A Theory for the Emergence of Neocortical Network Architecture

Daniel Udvary[1], Philipp Harth[2], Jakob H. Macke[3], Hans-Christian Hege[2], Christiaan P.J. de Kock[4], Bert Sakmann[5], Marcel Oberlaender[1,*]

[1]Max Planck Group: In Silico Brain Sciences, Center of Advanced European Studies and Research (caesar), Ludwig-Erhard-Allee 2, 53175 Bonn, Germany. [2]Department of Visualization and Data Analysis, Zuse Institute Berlin, Takustraße 7, 14195 Berlin, Germany. [3]Machine Learning in Science, Tübingen University, Maria-von-Linden-Straße 6, 72076 Tübingen, Germany. [4]Department of Integrative Neurophysiology, Center for Neurogenomics and Cognitive Research, VU Amsterdam, De Boelelaan 1085, 1081 Amsterdam, the Netherlands. [5]Max Planck Institute of Neurobiology, Am Klopferspitz 18, 82152 Martinsried, Germany.

*** Editorial correspondence:** Max Planck Group: In Silico Brain Sciences, Center of Advanced European Studies and Research (caesar), Ludwig-Erhard-Allee 2, Bonn, 53175 Germany, marcel.oberlaender@caesar.de

Developmental programs that guide neurons and their neurites into specific subvolumes of the mammalian neocortex give rise to lifelong constraints for the formation of synaptic connections. To what degree do these constraints affect cortical wiring diagrams? Here we introduce an inverse modeling approach to show how cortical networks would appear if they were solely due to the spatial distributions of neurons and neurites. We find that neurite packing density and morphological diversity will inevitably translate into non-random pairwise and higher-order connectivity statistics. More importantly, we show that these non-random wiring properties are not arbitrary, but instead reflect the specific structural organization of the underlying neuropil. Our predictions are consistent with the empirically observed wiring specificity from subcellular to network scales. Thus, independent from learning and genetically encoded wiring rules, many of the properties that define the neocortex' characteristic network architecture may emerge as a result of neuron and neurite development.

Introduction

The structural organization of the mammalian neocortex is more complex than that of other biological tissues. Each cubic millimeter of the cortical neuropil contains hundreds of meters of dendritic and several kilometers of axonal path lengths (*1*). These neurites originate from hundreds of thousands of neurons with diverse structural, functional and/or genetic properties. Characteristic for the neocortex, this extremely dense and diversely structured neuropil is the result of genetically induced programs with different critical periods during embryonic and postnatal development **(Fig. 1A)**. Neurogenesis and radial migration (*2, 3*), in combination with several neurite growth mechanisms (*4, 5*) guide the neurons' cell bodies (somata), dendrites and axons into specific subvolumes of the cortical sheet (*6*). As a result, the cellular organization into cytoarchitectonic layers and vertical functional columns is highly specific for each area and species (*7, 8*). Depending on the cell type, dendrites and axons develop morphological properties that correlate with different features of the areas' specific laminar (*9*) and columnar layout (*10, 11*).

Once developed, soma distributions, as well as dendrite and axon morphologies remain largely stable throughout life (*12*). Thus, developmental programs that shape the specific structural composition of each cortical area essentially result in a set of lifelong constraints for where neurites – and of which neurons – can in principle form synaptic connections with one another. To what degree neuron and neurite distributions might contribute to the complex wiring architecture of cortical networks remains however unknown. Particular open questions are whether structural constraints for synapse formation are preserved across animals, and whether area-specific properties of cortical wiring diagrams – in the following referred to as "structural scaffoldings" – could thereby be a result of neuron and neurite development **(Fig. 1B)**.

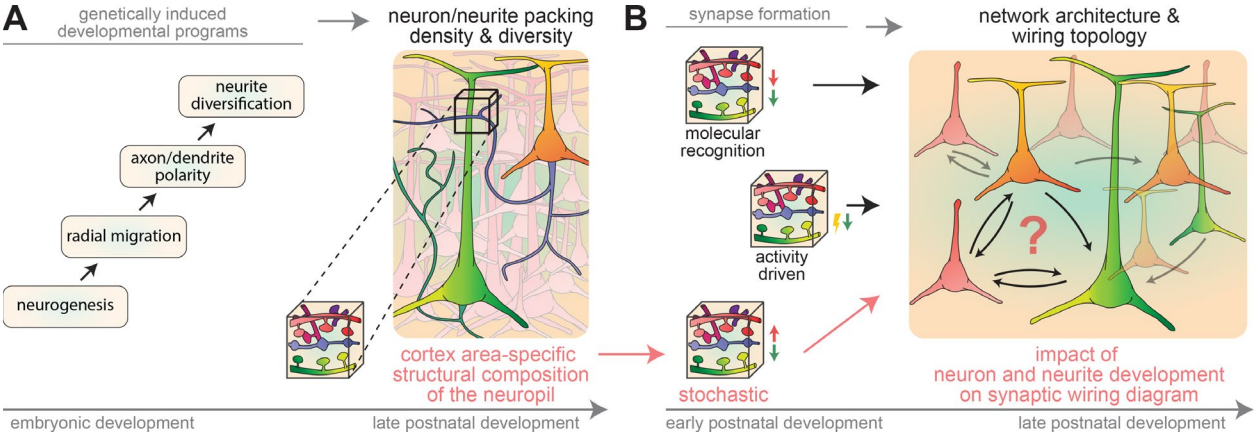

Figure 1. Potential impact of neuron development on cortical wiring. (A) Left: Schematic of cortex development (see also (*14*)). Right: By shaping the structural composition of the neocortex, genetically induced developmental programs provide constraints for where neurites – and of which neurons – can in principle form connections with one another. **(B)** Strategy for testing the impact of development on wiring. Reconstructing the structural composition allows computing how cortical wiring diagrams would appear in the absence of synapse formation mechanisms that rely on activity and cellular identity.

Here we quantitatively address these questions to reveal structural scaffoldings in the networks of the vibrissae-related part of the rat primary somatosensory cortex – i.e., the barrel cortex (*13*). For this, we introduce an inverse modeling approach that can predict all dense wiring diagrams – and their respective likelihoods – that could arise from neuron and neurite distributions. We find that whatever features may drive neuron and neurite development, structural scaffoldings with non-random properties will inevitably emerge in cortical wiring diagrams. These non-random properties are not arbitrary. Instead, we show that neurite packing density and morphological diversity translate into the shapes and correlations of connection probability distributions, and that the network's specific non-random topology reflects these parameters of the underlying pairwise statistics. The theory is consistent with wiring specificity observed empirically from subcellular to network scales, and provides quantitative predictions for future connectivity measurements. Thus, despite wiring mechanisms that rely on activity and/or cellular identity, emergent properties that characterize structural scaffoldings in cortical networks may largely persist throughout life.

Results

To reveal how neuron and neurite distributions affect cortical wiring diagrams, we had to overcome several major challenges. First, in parallel to neurite development (*4*), and to a lesser extent throughout life (*15, 16*), the neocortex is constantly remodeled via mechanisms that form, eliminate and replace synaptic connections in an activity dependent manner. Activity can also be

part of genetically induced developmental programs. For example, mechanisms that guide thalamocortical axons require spontaneous and periphery-driven activity (*17, 18*). Moreover, synapses can form based on genetically encoded wiring rules (*19, 20*), where molecular recognition between specific pre- and postsynaptic compartments results in connections that depend on subcellular, cellular and/or cell type identity (*21, 22*). This plethora of simultaneously active mechanisms involved in the generation and remodeling of connections constitutes a major obstacle when trying to infer the origin of wiring patterns that are observed empirically via *post hoc* reconstructions: which patterns reflect neuron and neurite distributions, which ones reflect learning or genetically defined cellular identity – or combinations thereof?

To overcome this obstacle, we developed an inverse modeling approach that allows predicting dense wiring diagrams from distributions of neurons and neurites, and comparing the predictions with available empirical connectivity data. We introduce the quantity *dense structural composition* (DSC) as the product of the numbers of pre- and postsynaptic structures that neurons $\boldsymbol{a}$ and $\boldsymbol{b}$ contribute to a subvolume $\hat{\boldsymbol{x}}$, relative to the total number of postsynaptic structures contributed by all neurons, here indexed with $\boldsymbol{N}$.

$$DSC_{(a,b,\hat{x})} = PRE_{(a,\hat{x})} \cdot \frac{POST_{(b,\hat{x})}}{\sum_N POST_{(N,\hat{x})}} \qquad \textbf{Equation (1)}$$

Based on this quantity, we formulate two assumptions about synapse formation mathematically. Any presynaptic structure can form a connection with any of the available postsynaptic structures. The probability p for the presence of n connections between neurons $\boldsymbol{a}$ and $\boldsymbol{b}$ within a subvolume $\hat{\boldsymbol{x}}$ is therefore given by a Poisson distribution with parameter n:

$$p_{(a,b,\hat{x},n)} = \frac{DSC_{(a,b,\hat{x})}{}^{n}}{n!} \cdot e^{-DSC_{(a,b,\hat{x})}} \qquad \textbf{Equation (2)}$$

The formation of a connection does not affect synapse formation elsewhere. Thus, the probability P that neurons $\boldsymbol{a}$ and $\boldsymbol{b}$ are connected by at least one synapse is given by:

$$P_{(a,b)} = 1 - e^{-\sum_{\hat{x}} DSC_{(a,b,\hat{x})}} = 1 - \prod_{\hat{x}} e^{-DSC_{(a,b,\hat{x})}} \qquad \textbf{Equation (3)}$$

where the index $\hat{\boldsymbol{x}}$ runs over all subvolumes that cover neurons $\boldsymbol{a}$ and $\boldsymbol{b}$. These equations do not reflect any particular wiring mechanisms at the molecular level. However, they capture stochastic (i.e., random) processes of synapse formation, elimination and replacement that are independent of (sub)cellular identity, including those shaped by competition or stochastic waves of activity as observed during development (*23*). These equations are also consistent with the synaptotropic hypothesis, which states that synaptic inputs control the fine-scale elaboration of dendritic and axonal arbors (*24*). Purposefully, these equations neglect any wiring mechanisms that rely on periphery-driven activity and/or genetically defined cellular identity. Parameterizing neuron and neurite distributions by the quantity DSC, followed by application of these equations, thereby allows calculating all wiring diagrams, as well as their respective likelihoods, that could originate from these specific underlying distributions. Our inverse modeling approach thus results in probability distributions of dense wiring diagrams – here referred to as 'statistical connectomes' (*25*) – that represent the structural scaffoldings for any given distributions of neurons and neurites.

Model captures the characteristic cellular organization of the barrel cortex

The second challenge was that electron microscopic approaches that could provide the distributions of neurons and neurites that are necessary for calculating statistical connectomes, remain limited to small volumes – currently to cubes with up to 100 µm edge lengths (*26*). However, developmental programs shape the cellular and morphological organization of the neocortex at scales of several hundreds of micrometers, and even millimeters (*11*). Thus, we decided to generate an anatomically detailed digital model of the neuron and neurite distributions

across the cortical volume that represents the 24 major facial whiskers along the animals' snout (A1-E4, α-δ). This model is based on anatomical data that we have systematically collected during the past two decades for the rat barrel cortex and primary thalamus of the whisker system **(Fig. 2A)** – the ventral posterior medial nucleus (VPM). To minimize sources that could potentially increase structural variability across animals, all data originated from rats of the same strain, housed in standard laboratory environments, and sacrificed at similar time points following the critical periods of neuron and neurite development.

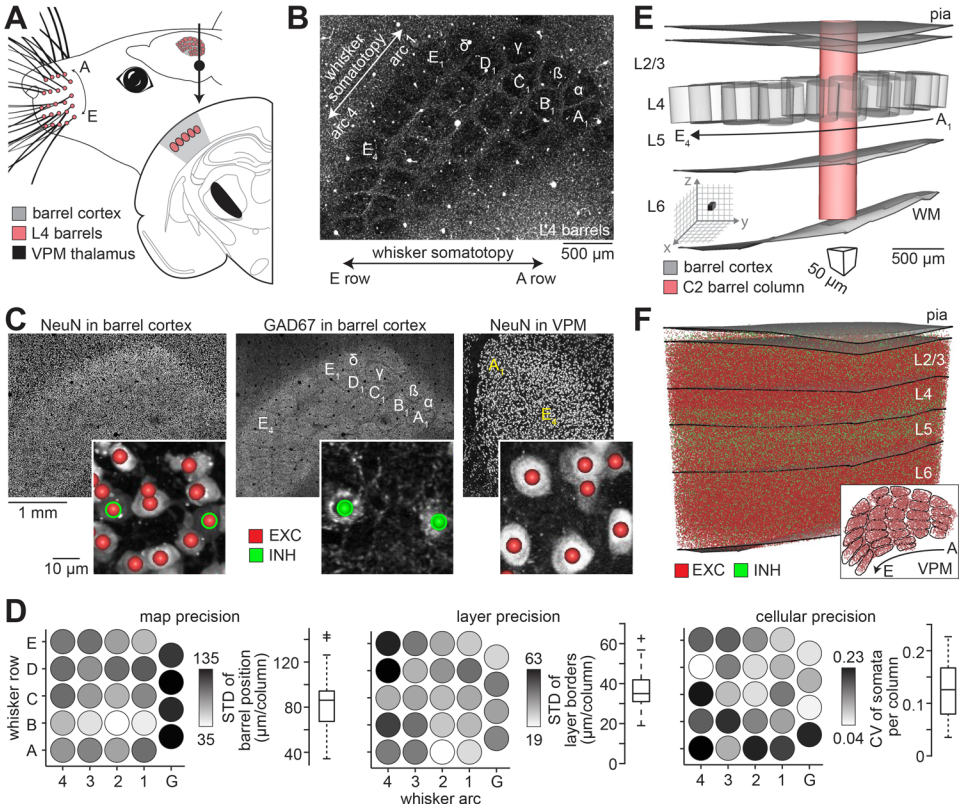

Figure 2. Cellular organization of rat barrel cortex. (A) The ventral posterior medial nucleus (VPM) in thalamus relays whisker input to the barrel cortex. **(B)** Whisker map of barrels in layer 4 (L4). Panel modified from Ref. (*27*). **(C)** Excitatory (EXC) and inhibitory (INH) somata across barrel cortex and VPM. Panel modified from Ref. (*28*). **(D)** Standard deviations (STDs, N=12 rats) of somatotopy (left), and STDs of cytoarchitecture (center) and coefficients of variations (CVs, N=4 rats) of neuron numbers (right). **(E)** Model of the barrel map, layers, pia and white matter at 50 µm resolution (WM). **(F)** Model of the cellular organization of barrel cortex and VPM (insert).

For the model to capture the characteristic cellular organization of the barrel cortex, we reconstructed precise 3D maps of 'cortical barrel columns' **(Fig. 2B)** (*27*) and quantified the locations of all excitatory and inhibitory neuron somata for this volume (*28*). These data revealed relationships between the somatotopic map and cytoarchitecture **(Fig. 2C)** that result in neuron numbers per barrel column, and per layer therein, that are highly specific for each whisker. Variability of these whisker-specific features is, however, small across animals. The position of each barrel within the map, as well as layer borders within each barrel column, vary by less than 100 µm **(Fig. 2D)**, neuron densities across the barrel cortex by less than 5%. Given this precision of the columnar and laminar layout (±50 µm), we combined the reconstructions of the barrel maps with surface reconstructions of the layer borders, pia and white matter **(Fig. 2E)**, and subdivided the resultant volume into cubes with 50 µm edge length. Each cube was populated with the

measured average number of excitatory and inhibitory neuron somata, and assigned to one of the 24 barrel columns and one of the six layers therein **(Fig. 2F)** – or to the nearest column if located in the septum between barrels. The model captures 82% of the inter-animal-variability with respect to the whisker-specific cellular organization of the rat barrel cortex **(Fig. S1)**.

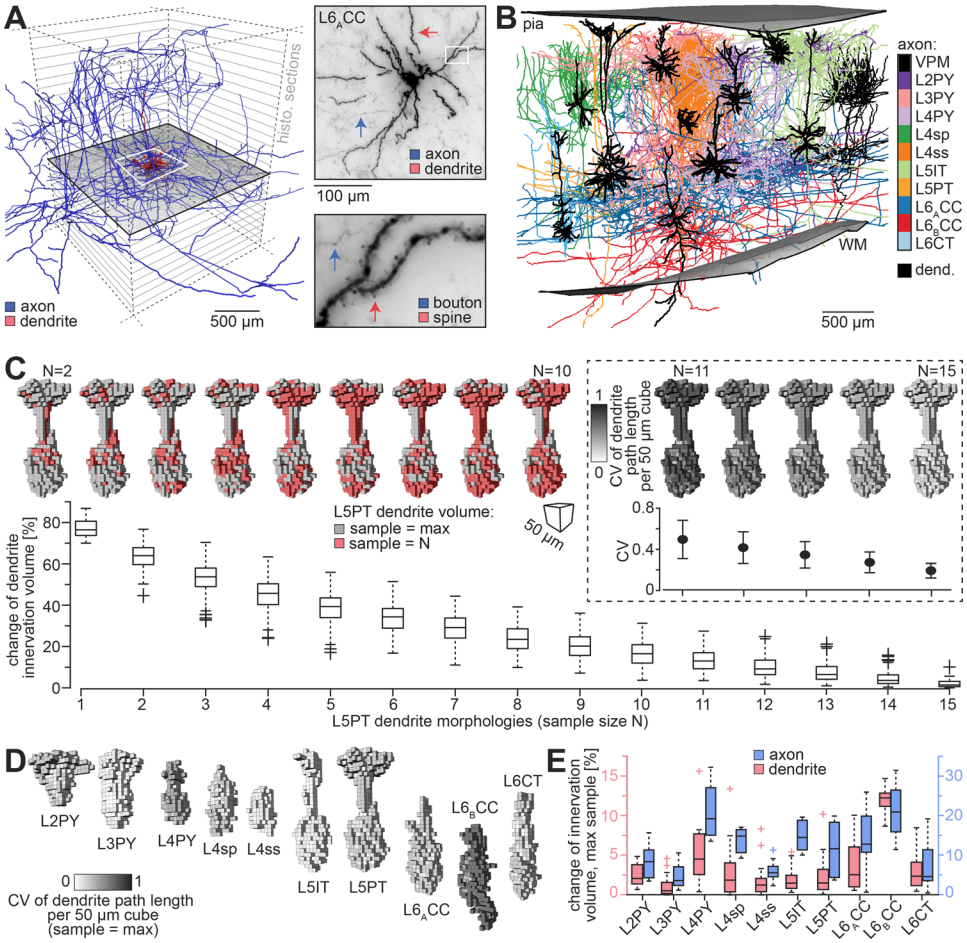

Figure 3. Cell type-specific morphologies in rat barrel cortex. (A) Left: Dendrites (red) and axon (blue) of exemplary *in vivo* labeled neuron reconstructed across sections from pia to WM. Top-right: Projection image of section containing the soma. Bottom-right: boutons and spines along biocytin labeled axon and dendrite. **(B)** Dendrites (black) and axons that represent the cortex' major excitatory cell types. Pyramidal neurons in layers 2-4: L2PY (n=16), L3PY (n=30), L4PY (n=7), star pyramids (L4sp, n=15), spiny stellates (L4ss, n=22); slender-tufted intratelencephalic (L5IT, n=18), thick-tufted pyramidal tract (L5PT, n=16), corticocortical neurons in upper ($L6_ACC$, n=14) and deep layer 6 ($L6_BCC$, n=5), corticothalamic (L6CT, n=13), and VPM (n=14). **(C)** Dendrite innervation volumes of L5PTs vs. number of reconstructed neurons. Insert shows CVs of dendrite densities within each 50 µm cube. **(D)** Innervation volumes and CVs of dendrite densities therein for the maximal sample of reconstructions per cell type. **(E)** Robustness of estimates for dendrite (red) and axon (blue) innervation volumes for each cell type.

For the model to capture the barrel cortex' cell type-specific morphological organization, we reconstructed a sample of *in vivo* labeled morphologies **(Fig. 3A)** that represents ~1% of the excitatory neurons located within a barrel column across layers 2 to 6 (n=154) (*11*). Similarly, we reconstructed a sample that represents ~5% of the neurons located within a VPM barreloid (n=14) (*29*). The sample includes morphologies for all major excitatory cell types of the neocortex **(Fig.**

3B). Complementing the morphological themes that are common across cortical areas and species (*9*), each cell type develops specific variations in its dendritic and axonal patterns that reflect different features of the barrel cortex' specific laminar and columnar layout. Most noticeably, neurons project the majority of their axons far beyond the dimensions of a single barrel column, either along the whisker row, arc or both. To investigate how representative these morphologies are, we calculated dendrite and axon innervation volumes for each cell type at the resolution of 50 µm, while increasing the number of reconstructed morphologies **(Fig. 3C)**. For example, the volume innervated by dendrites of layer 5 pyramidal tract neurons (L5PTs) will change by less than 3% even if additional morphologies are reconstructed. The dendritic path length within each 50 µm cube of this volume would change by less than 20%. Similar results are obtained for all cell types **(Fig. 3D)**. At the resolution of the model, our sample of morphologies thus captures 96% of the inter-animal-variability with respect to cell type-specific dendrite distributions, and 88% with respect to axon distributions **(Fig. 3E)**.

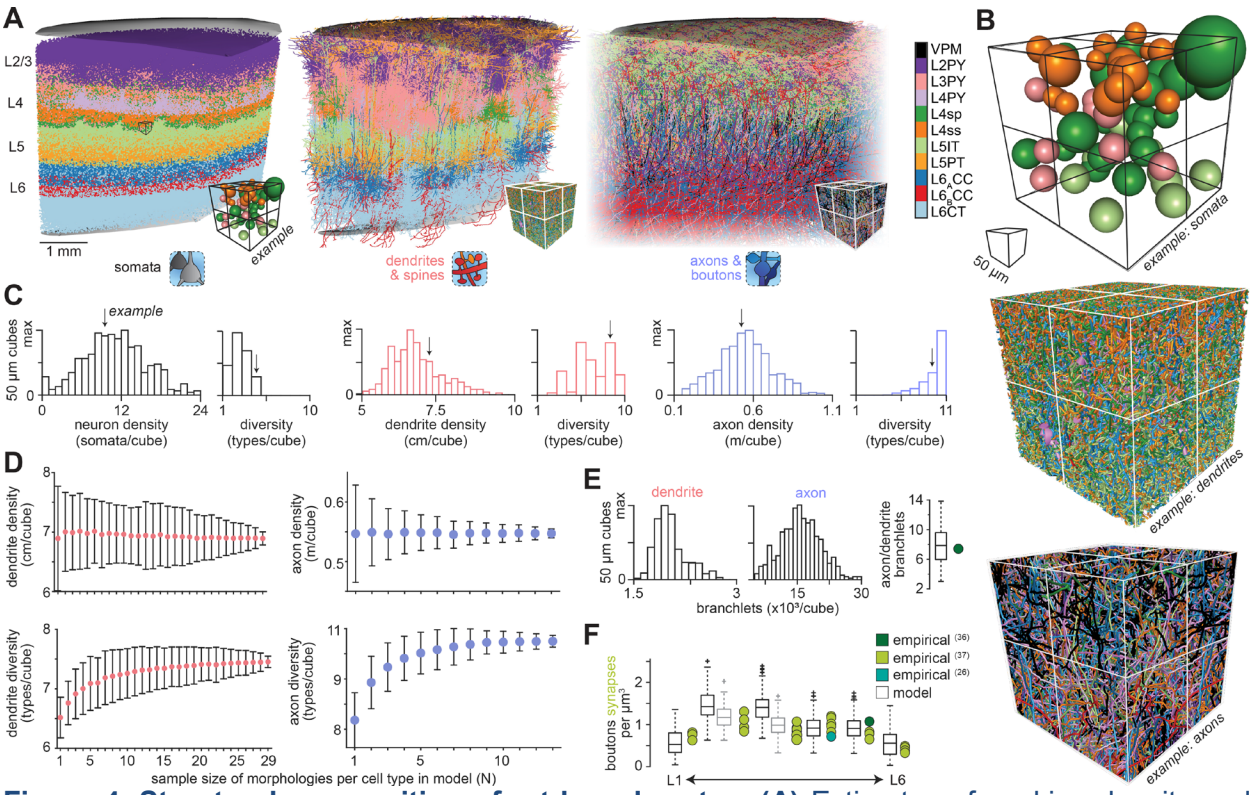

Figure 4. Structural composition of rat barrel cortex. (A) Estimates of packing density and diversity distributions of somata (left), dendrites (center), and axons (right) at a resolution of 50 µm cubes. **(B)** Zoom-ins to panel A show eight exemplary 50 µm cubes at the L4/5 border. Colors denote cell types and VPM axons (black). **(C)** Histograms of density distributions for somata, dendrites and axons per 50 µm cube. Diversity distributions reflect the numbers of cell types that neurons or neurites represent per cube (e.g. all types = 11). **(D)** Dendrite (left) and axon (right) packing densities, as well as cell type diversity of these neurites (bottom) within each 50 µm cube depending on the number of morphologies per cell type used to generate the model. **(E)** Predicted ratios of dendrite/axon branchlets per 50 µm cube vs. electron microscopy data (*36*). **(F)** Predicted bouton densities per layer vs. synapse densities measured via electron microscopy (*26, 36, 37*).

Model captures >90% of the dense structural organization preserved across animals

The third challenge was that proximity is only necessary, but not sufficient for the formation of synaptic connections. More specifically, a longstanding hypothesis states that the presence of

axons and dendrites within the same subvolume (i.e., axo-dendritic overlap) is predictive of synaptic connections between them (*30*). This hypothesis is commonly referred to as Peters' Rule (*31*) and it can be simply restated as: "proximity predicts connectivity". Even though the issue of how best to define Peters' Rule remains controversial (*32*), tests of this hypothesis at a number of spatial resolutions, and in several areas of the mammalian brain including the neocortex have failed to support it (*26, 33-36*). Proximity between two neurites hence does in general not provide sufficient information to account for properties of cortical wiring diagrams – neither at subcellular, cellular, nor at cell type levels. Consequently, the barrel cortex model needed to provide anatomically realistic and robust estimates for soma, dendrite and axon packing density distributions across the entire model volume to ensure that constraints for synapse formation reflect proximity between all neuronal structures that share a particular subvolume.

For this, we replaced each soma in the model with a morphology (*25*) from the sample of *in vivo* labeled neurons **(Fig. 4A)**. Morphologies were registered to model locations within ±50 µm of their 'true' soma depths, and by preserving their respective orientations within columns and relative path length distributions across layers (*25, 27*). Similarly, thalamocortical axons were placed by matching the respective numbers of neurons per VPM barreloid. This up-scaled model provides quantitative estimates for the dendritic and axonal path lengths that each neuron, depending on its soma location, cell type and morphology, can contribute to a particular cortical subvolume **(Fig. 4B)**. The predicted dense structural composition of the model depends strongly on the specific laminar and columnar location of each 50 µm cube **(Fig. 4C)**. However, dendrite and axon packing densities within each subvolume will not change by more than 8% and 7%, respectively – their cell type diversity by no more than 6% and 5% – even if the model is based on a larger sample of neuron morphologies **(Fig. 4D)**. At the resolution defined by the columnar and laminar layout of the barrel cortex (i.e., 50 µm), the estimated packing density distributions thus capture more than 90% of the structural composition for this cortical volume that is preserved across animals. Further, the estimated ratios of axon/dendrite branchlets per subvolume are consistent with empirical data **(Fig. 4E)** from dense electron microscopy reconstructions (*36*).

Model predicts all wiring diagrams that the barrel cortex' structure could in principle form

The model provides robust and realistic estimates of the cell type-specific soma, dendrite and axon packing density distributions across a volume that is large enough to capture the barrel cortex' specific cellular and morphological organization. To calculate the statistical connectome for this model, we parameterized its neuron and neurite packing density distributions in accordance with **Equation 1**. For this, we quantified boutons along the axons of the *in vivo* labeled morphologies with respect to cell type and target layer (*11*), and converted the axon (and dendrite) packing densities into bouton (and spine) densities (*25*). The resultant density distributions of synaptic structures were consistent with those of layer-specific synapse density measurements **(Fig. 4F)** reported from electron tomography (*37*).

Application of **Equations 2-3** hence resulted in a statistical connectome, which predicts the structural scaffoldings for the entire rat barrel cortex. More specifically, the statistical connectome provides the probabilities that neurons are connected to one another **(Fig. 5A)**, based on their respective contributions to the structural composition of the model volume. The anatomically realistic nature of the model allows grouping of neurons by their soma locations and/or cell types **(Fig. 5B)**. Thus, the statistical connectome provides connection probability distributions for arbitrarily defined subpopulations **(Fig. 5C)**, information about which neurons could be connected to one another **(Fig. 5D)**, as well as likelihoods where along the dendrites – i.e., in which of the 50 µm cubes – these connections could occur **(Fig. 5E)**. Even at this level of detail, describing the complexity of wiring in the barrel cortex requires more than 10^{16} connection probability values. To facilitate comprehensive analyses of this extensive dataset, we developed a web-based

framework, which provides online access to the statistical connectome, the barrel cortex model, as well as to the anatomical data and computational routines that were used to generate it (see **Materials and Methods**).

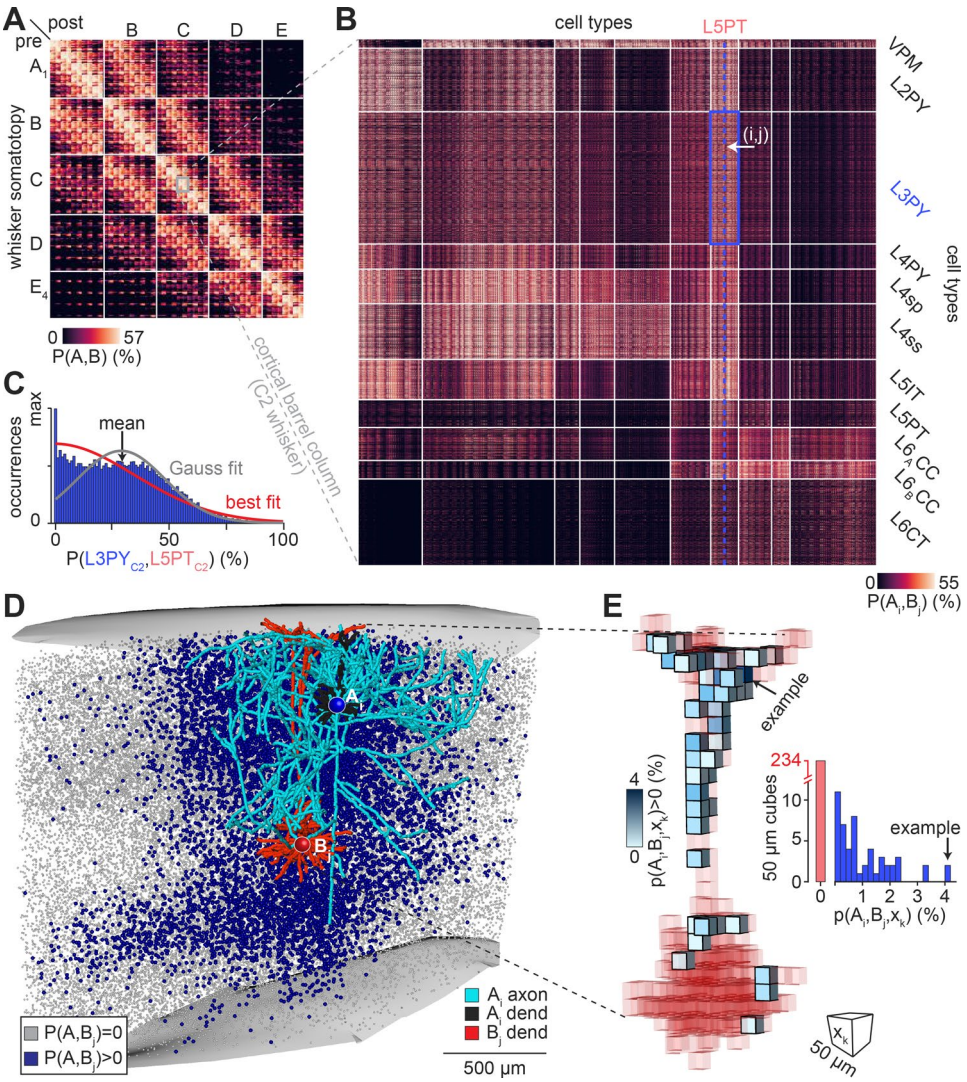

Figure 5. Statistical connectome of rat barrel cortex. (A) Matrix of connection probabilities between all excitatory neurons in the model sorted by the somatotopic location and cell type. **(B)** Zoom-in to panel A. **(C)** Connection probability distribution between L3PYs and L5PTs (blue rectangle in panel B). **(D)** Somata (blue; 50% shown) among all neurons (grey; 10% shown) that could connect to exemplary L5PT (blue dashed line in panel B). Dendrites (black) and axon (light blue) of one exemplary L3PY that could connect to the L5PT's dendrites (red). **(E)** Morphologies in the model are represented as density distributions (left), resulting in connection probabilities per pair (white arrow in panel B) and 50 µm cube (right).

Model predicts structural scaffoldings of non-random connectivity statistics

Here we restrict analyses to connections between excitatory populations (see **Materials and Methods** and **Discussion** for inhibitory neurons), and illustrate the results on neurons whose somata are located within layer 5 of the barrel column representing the C2 whisker. For all pairs of neurons from this exemplary volume **(Fig. 6A)**, we extracted the respective connection

probability values between them from the statistical connectome, and approximated the resultant connection probability distribution by the best fit with one of seven functions (Gaussian, Half-Normal, Exponential, Gamma, Binomial, zero-inflated Poisson, zero-inflated Negative Binomial). The shape of the distribution was further quantified by its mean, standard deviation (STD), coefficient of variation (CV) and skewness. Subdividing the selected neurons into different groups depending on their cell types, soma positions within the column or septum, inter-somatic distances, and combinations thereof, yields connection probability distributions that vary substantially **(Fig. 6B)** – i.e., the distributions have different means, CVs, and their shapes cannot be fitted consistently with one of the seven functions. These observations generalize to all possible groupings of neurons across layers and cell types **(Fig. S2)**. Irrespective of the grouping, the mean connection probabilities are, however, in general small (mean ± STD across n=110 groupings: 0.17 ± 0.11), whereas the CVs are large (1.18 ± 0.94). Less than 5% of the connection probability distributions are approximated best by a Gaussian – a common assumption for describing pairwise connectivity statistics between neurons.

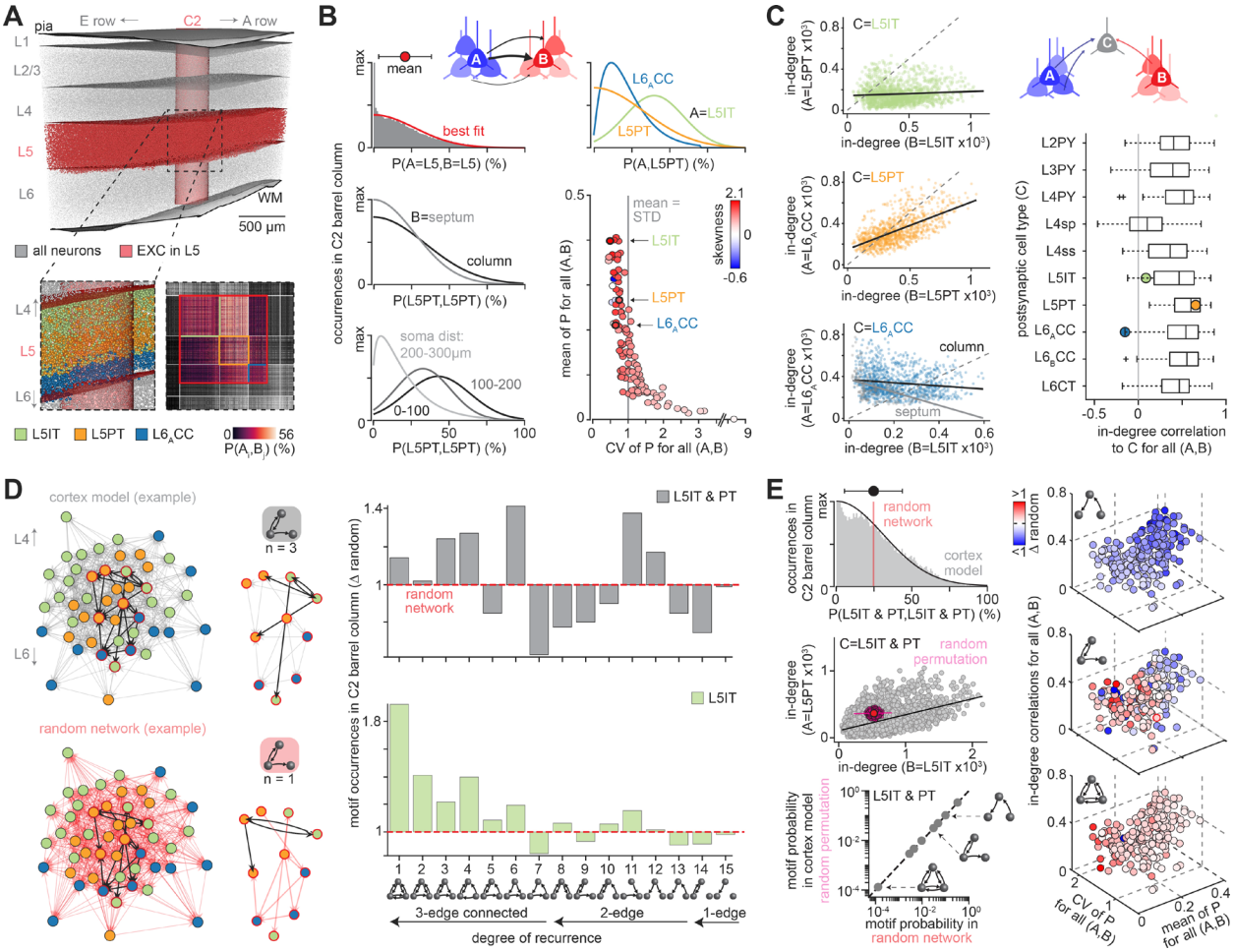

Figure 6. Structural scaffoldings of rat barrel cortex. (A) Exemplary selection of neurons from the model, comprising L5ITs, L5PTs and L6$_A$CCs. Zoom-in to the matrix in **Fig. 5A**, representing the selected neurons (bottom-right). **(B)** Connection probability distributions representing the best fits to connection probability histograms derived from the matrix in panel A for different groupings. Bottom-right: Means and CVs of connection probability distributions for all groupings (n=110). **(C)** Correlations between in-degrees for three of the groupings in panel B. Bottom-right: In-degree correlation coefficients for all groupings. **(D)** Wiring diagrams generated from the barrel cortex

model for 50 exemplary L5 neurons from panel A. The underlying pairwise statistics reflect either those predicted as structural scaffoldings in panels B/C (top-left) or by randomly connecting the neurons (bottom-left). Ratio between motif occurrences in the model and random network (1: equally abundant; >1: overrepresented in the model; <1: underrepresented). **(E)** First (top-left) and second order statistics (center-left) underlying the wiring diagram between L5ITs and L5PTs in the model and random network in panel D. Random permutation of connection probabilities in the barrel cortex model results in topologies that are indistinguishable from those of random networks (bottom-left). Right: Deviations in the occurrences for three exemplary motifs between the model and random networks vs. the means and CVs of connection probability distributions, and in-degree correlations for all groupings (see also **Fig. S6**).

Next, we quantified the number of connections that each neuron is predicted to receive from all other neurons in the exemplary volume. These 'in-degree' distributions are in general broad, and have shapes that depend on the grouping of neurons with respect to cell type, soma position, inter-somatic distance, and combinations thereof **(Fig. S3)**. Moreover, we found substantial correlations between in-degree distributions **(Fig. 6C)**. For example, the more connections L5PTs receive from one another, the more connections these neurons are predicted to receive from corticocortical neurons ($L6_ACCs$) that are located around the layer 5/6 border (*38*). In contrast, the more connections $L6_ACCs$ receive from intratelencephalic neurons (L5ITs), the less connections these neurons are predicted to receive from one another, in particular when their somata are located in the septum. These observations generalize to all groupings, where correlation coefficients (Pearson's R) between in-degree distributions range from −0.46 to 0.87 **(Fig. 6C)**. In-degree correlations are biased towards positive values (mean ± STD: 0.40 ± 0.27, n=550 groupings).

Finally, we investigated the topologies of networks that could arise from the statistical connectome. For this, we calculated the respective occurrences of the fifteen wiring patterns by which three neurons can be interconnected – commonly referred to as triplet motifs **(Fig. 6D)**. Depending on how the layer 5 neurons are grouped, the respective occurrence of each motif differs, but in general deviates from those of networks where neurons are interconnected randomly according to the mean of the underlying connection probability distribution. For example, compared to these random networks, unidirectionally connected chains of layer 5 neurons – the feed-forward motif – are predicted to be less frequent in the barrel cortex model (i.e., underrepresented). In contrast, bidirectionally connected loops – the fully recurrent motif – are overrepresented. Such deviations from the topology of random networks generalize to all groupings **(Fig. S4)**. Beyond triplet motifs, high recurrence characterizes the barrel cortex model, as overrepresentation increases with the number of bidirectional connections per motif, and with the number of neurons per motif **(Fig. S5)**. Conversely, underrepresentation of feed-forward motifs increases with the number of neurons per motif.

Both the random networks and those based on the barrel cortex model have the same mean connection probability values. However, variance in connection probability and degree distributions, as well as correlations, are absent in the random networks. To assess the impact of these differences on topology, we shuffled the connection probability values of the statistical connectome, which removes all correlations from the barrel cortex model, while maintaining the shapes of the connection probability distributions. The occurrences of motifs in the shuffled barrel cortex model are indistinguishable from those of random networks **(Fig. 6E)**. Thus, in the absence of correlations, the shapes of connection probability distributions have no impact on network topology, which will be random. The opposite is true in the presence of correlations. To illustrate this finding, we plotted how motif occurrences deviate from those in random networks against the connection probabilities' means and CVs, and in-degree correlation coefficients, for all groupings in the barrel cortex model **(Fig. S6)**. This analysis reveals that the shapes of connection probability

distributions in general affect the occurrences of motifs. Pairwise statistics can even affect network topology qualitatively, as motif occurrences transition between over- and underrepresentation, for example depending on the mean connection probability **(Fig. 6E)**.

Theory for the emergence of structural scaffoldings in cortical networks

We conceptualize the results of the statistical connectome analyses by formulating a general theory for how structural scaffoldings emerge from the underlying neuron and neurite distributions. First, the higher neurite packing densities within and across cortical subvolumes, the smaller the respective contributions of any particular neuron, and hence the smaller the probability that its neurites are connected to any one of the neurites in its vicinity. Thus, neurite packing density translates into the means of connection probability distributions, and thereby defines a networks' 'sparsity' **(Fig. 7A)**. Second, the higher the diversity of neurites within and across subvolumes – for example with respect to the cell types of the neurons that these neurites belong to – the broader the shapes of connection probability distributions when neurons are grouped by these cellular features. Thus, neurite diversity translates into the widths of connection probability distributions, and thereby defines a networks' 'heterogeneity'. Third, the more similar neurons contribute to the packing density and diversity distributions, the stronger are correlations between the resultant connection probability and/or degree distributions. Thus, neurite patterns that reflect similarities in the locations and morphologies of neurons translate into network correlations.

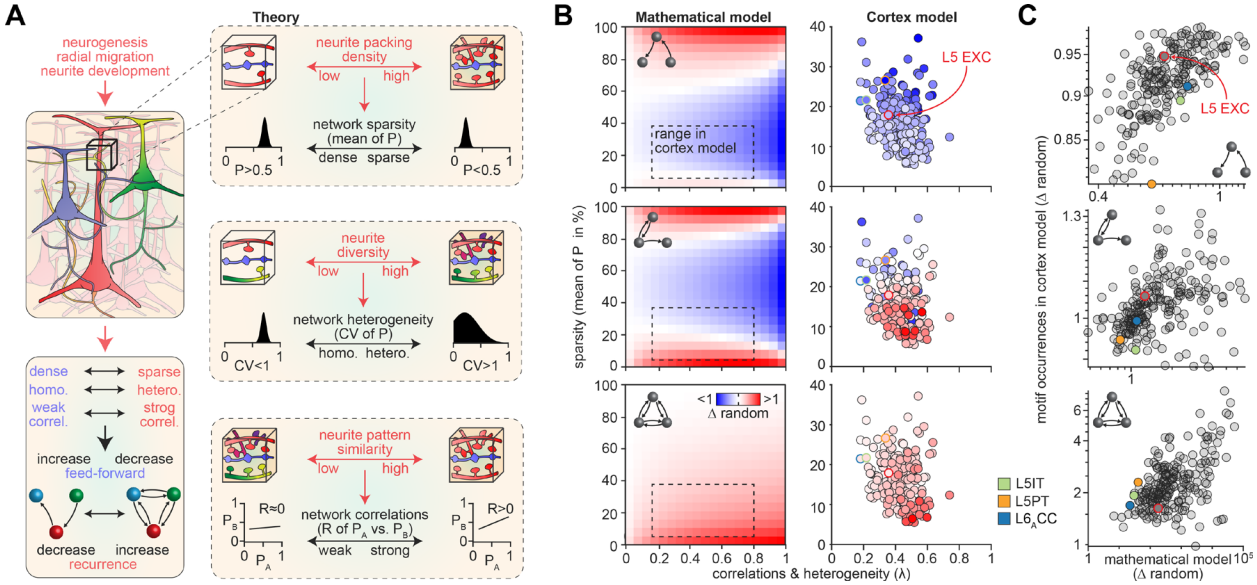

Figure 7. Theory for the emergence of structural scaffoldings. (A) Theory that conceptualizes the findings shown in **Fig. 6**. Top-left: By shaping the specific distributions of neurons and neurites within cortical areas, developmental programs give rise to structural constraints for synapse formation. Top-right: Neurite packing density defines a network's sparsity (means of connection probability distributions). Center-right: Morphological diversity of the neurites' cellular origins defines a network's heterogeneity (width of connection probability distributions). Bottom-right: Similarities between the neurons' locations and morphologies translate into patterns within neurite density and diversity distributions that define correlations in connection probability and degree distributions. Bottom-left: These properties of the underlying pairwise statistics define area-specific non-random topologies in wiring diagrams. For example, network architectures become increasingly recurrent or feed-forward with increasing or decreasing sparsity, heterogeneity and correlations. **(B)** Mathematical model shows for three exemplary triplet motifs (see also **Fig. S7**) that the shapes and correlations of the underlying connection probability distributions define the

specific non-random topology of networks (see **Materials and Methods** for a definition of λ). The dashed box reflects that range of sparsity and heterogeneity and correlations (i.e., λ) across all layer and/or cell type groupings in the barrel cortex model, as shown in the corresponding panels on the right. The colored circles represent the example groupings shown in **Fig. 6A**. **(C)** Occurrences of motifs for all groupings in the barrel cortex model vs. the respective predictions by the mathematical model.

In the **Materials and Methods** we provide mathematical evidence that these three properties of the underlying pairwise statistics define a network's topology. More specifically, in the presence of correlations, motif occurrences will deviate in general from those of random networks. These deviations are however not arbitrary **(Fig. S7)**, but instead reflect the degrees of sparsity and heterogeneity of the underlying pairwise connectivity statistics. For example, recurrent motifs become increasingly overrepresented the sparser and the more heterogeneous a network is interconnected **(Fig. 7B)**. Conversely, feed-forward motifs become increasingly overrepresented the denser and the more homogeneous the network is. The mathematical model was sufficient to account qualitatively for the non-random topology of the barrel cortex model **(Fig. 7C)**, validating that indeed sparsity, heterogeneity and correlations represent its defining sources.

Model predicts structural scaffoldings consistent with the available empirical data

We tested the theory by comparing the predicted properties of structural scaffoldings with the wealth of empirical data acquired for the barrel cortex during the past decades. First, we tested the degree of sparsity predicted to emerge from the structural composition at subcellular scales **(Fig. 8A)**. Across the model volume, only 1.1 ± 0.3% – and never more than 2% – of the neuron pairs whose dendrites and axons share the same 50 µm cube can be synaptically connected. This prediction is a direct consequence of the extremely high packing densities, which lead to a combination of axonal and dendritic branches that exceeds the number of synaptic connections in the corresponding volume by two orders of magnitude. Reflecting a general limit for the validity of Peters' Rule, this observation reveals one of the major reasons why empirically the vast majority of close-by neurites are found unconnected. However, if connected, the probability that an axon forms more than one synapse along the same dendrite is larger than zero, and the occurrences of such synaptic clusters per 50 µm cube are predicted to decrease with the number of synapses per cluster **(Fig. 8B)**. Recently, occurrences of clusters with two, three or four synaptic contacts between the same branches were reported from dense reconstructions of layer 4 in mouse barrel that are consistent, even quantitatively, with these predictions (*26*).

Next, we tested the degree of sparsity predicted to emerge from the structural composition at cellular scales. Pairwise connectivity between virtually all major neuronal populations of the barrel cortex have been assessed empirically. These studies employed somatic recordings for detecting postsynaptic potentials in response to action potentials that were induced in individual putatively presynaptic neurons either *in vivo* via sensory stimulation (e.g. (*39*)), or in acute brain slices *in vitro* via somatic current injections (e.g. (*40*)) or optical stimulations (e.g. (*41*)). Because the issue of space-clamping can hamper the detection of synaptic connections at distal dendrites (*44*), these empirical data are likely biased towards connections relatively close to the soma. Moreover, truncation of dendrites, and in particular of axons, is likely to introduce unsystematic biases when probing the fractions of connected neuron pairs *in vitro*. To compare the predictions of the statistical connectome with these empirical data, we therefore needed to mimic the respective experimental conditions of each empirical study **(Fig. 8C)**. For this, we grouped neurons in the barrel cortex model analogously to eighty-nine layer, cell type and/or distance dependent samplings, reported across a set of twenty-nine studies **(Tables S1-3)**. For *in vitro* studies, we generated slices from the model, truncated the neurites accordingly, sampled neurons in these virtual slices according to the respectively reported recording depths, and by restricting

connections to those along proximal dendrites **(Fig. S8)**. The means of the hence predicted connection probability distributions correlate significantly with the empirical data (R=0.75, $p<10^{-16}$). Even quantitatively, the predictions are remarkably consistent with the empirical data. About 2/3 of the empirically determined connectivity values deviate from the prediction by less than half a standard deviation of the respective connection probability distribution, 94% by less than one standard deviation. Only two studies, which reported unusually high connection probabilities, were inconsistent with the respective predictions **(Fig. 8D)**.

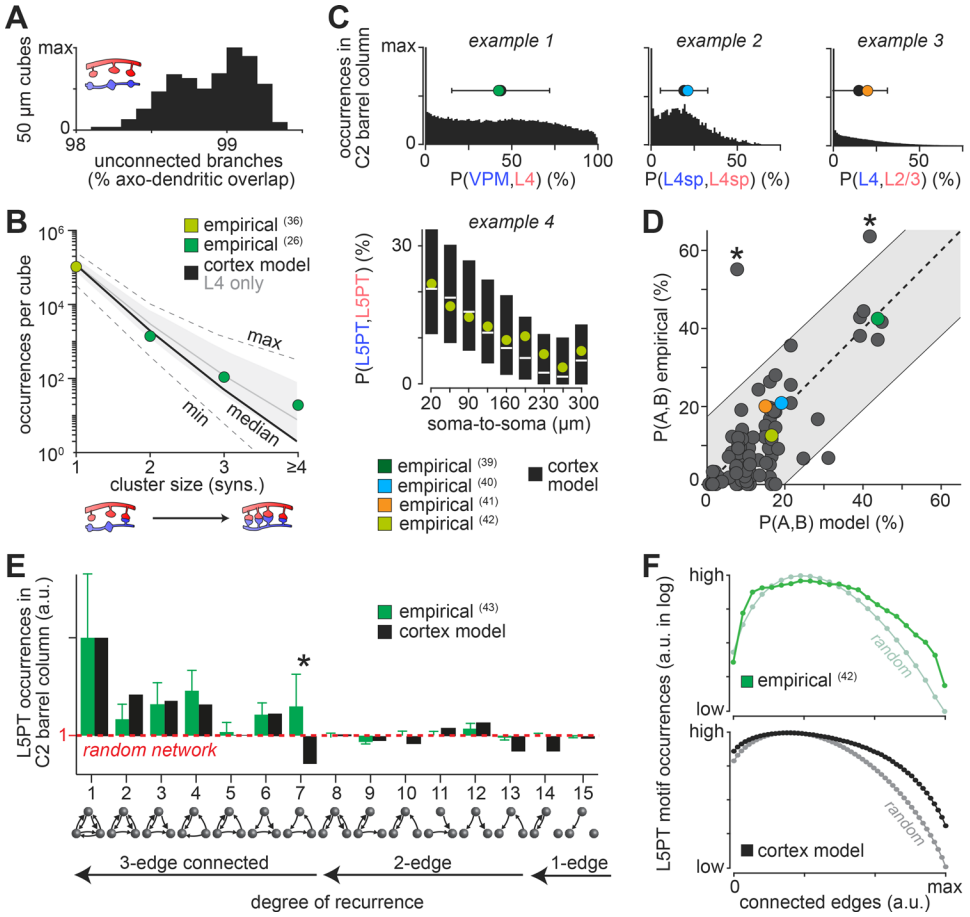

Figure 8. Structural scaffoldings vs. empirical data. (A) Axo-dendritic overlap (i.e., Peters' Rule) is insufficient to account for sparsity in the barrel cortex model at subcellular scales. **(B)** The occurrences of synaptic clusters in the barrel cortex model depends on the cluster size, which is consistent with electron microscopy data in L4 of the barrel cortex (*26, 36*). **(C)** Connection probability distributions predicted in the barrel cortex model for groupings that match those of four exemplary empirical studies (*39-42*). **(D)** Empirical connection probabilities for 89 layer and/or cell type groupings vs. those predicted in the barrel cortex model **(Tables S1-3)**. Grey-shaded area represents 95% prediction interval. Asterisks denote inconsistencies between empirical data and barrel cortex model. **(E)** Occurrences of triplet motifs between L5PTs in the barrel cortex model are consistent with empirical data (*43*). **(F)** Overrepresentation of L5PT motifs increases with the number of connected edges empirically (*42*) (top) and in the barrel cortex model (bottom).

We performed several additional analyses to solidify these observations. First, without emulating *in vitro* conditions, the model predictions remain correlated with the empirical data. However quantitatively, consistency decreases the more the experimental conditions of the respective studies are neglected – no emulation of space-clamping: R=0.72; no emulation of slicing: R=0.70;

no emulation of both: R=0.64. Second, to test whether consistency with the empirical data is due to large heterogeneity, and thereby emerges by chance, we performed random permutations of the eighty-nine predicted connection probabilities. Permutations yielded correlations with the empirical data that were in general not significant (R=0.00 ± 0.11; range: −0.38 to 0.51), and never as high as the one predicted to emerge from the structural composition (i.e., R=0.75). Third, we tested whether definitions of Peters' Rule at cellular scales (*35*) are sufficient to explain the model predictions **(Fig. S9)**. For example, axo-dendritic overlap between L5PTs and $L6_A$CCs is 1.44 times larger than that between $L6_A$CCs and L5PTs. Opposite to this relationship, the ratio between the means of the respectively predicted connection probability distributions is 0.72. This mismatch between overlap and connection probabilities generalizes to all groupings (F test that differences in axo-dendritic overlap results in the same differences of connection probabilities: F>23.2, degrees of freedom=1, $p<10^{-5}$). Thus, both at subcellular and cellular levels, definitions of Peters' Rule are insufficient to account for the properties of structural scaffoldings.

Finally, we tested properties predicted to emerge from the structural composition at network scales. Empirical assessments of network topology remain, however, scarce and limited to small populations of neurons. Empirical data for heterogeneity and correlations in structural connectivity are lacking altogether. Fortunately for L5PTs, the occurrences of all triplet motifs and their respective deviations from a random network were systematically assessed (*43*). The statistical connectome predicts motif occurrences for neurons of this cell type that are remarkably consistent with these empirical data **(Fig. 8E)**, with the notable exception of unidirectionally connected loops – the feed-forward loop motif (see Discussion). Moreover, probing the occurrences of motifs between up to eight L5PTs revealed that independent of their particular topology, motifs become increasingly overrepresented with increasing numbers of connected edges (*42*). This empirically determined relationship is qualitatively consistent with the model predictions **(Fig. 8F)**.

Discussion

We present a comprehensive theory for the emergence of complex network architectures in the mammalian neocortex. Our data reveal that simply because proximity is a necessary condition for synapse formation, structural constraints of neurite packing density and diversity provide themselves a major source for the neocortex' characteristic wiring properties from subcellular to network levels. Thus, whatever features may drive neuron migration and neurite guidance during development, non-random pairwise and higher-order connectivity statistics will inevitably emerge in the wiring diagrams of cortical networks, and thereby define structural scaffoldings that are specific for each area and species. Properties that characterize structural scaffoldings, as predicted here for the barrel cortex, can have significant qualitative impacts on cortical dynamics. Balance between excitation and inhibition depends for example critically on the network's heterogeneity and correlations (*45*). Consequently, to ensure robustness of cortical dynamics throughout life, wiring mechanisms that depend on activity and/or cellular identity may thus constantly remodel wiring in the neocortex, while maintaining the properties of its structural scaffoldings. Such homeostasis of structural scaffoldings provides a likely explanation to resolve what may otherwise seem paradoxical: the model in which any activity and/or cellular identity dependent wiring mechanisms were purposefully neglected, predicts connectivity for the barrel cortex that is consistent with empirical data from animals in which such mechanisms were however involved in generating and remodeling synaptic connections.

Structural scaffoldings beyond local excitatory connections

Our attempts to falsify the theory remain limited to local connections between excitatory populations and their innervation by long-range axons from primary thalamus. Recent studies

however indicate that the theory is likely to extend to other excitatory long-range pathways. For example, depending on their respective subcortical target areas, somata of L5PTs form sublayers in mouse motor (*46*) and rat barrel cortex (*47*). Moreover, these neurons have dendrite distributions across layers that correlate with their respective downstream targets (*47, 48*). Consequently, L5PTs contribute to the soma, dendrite and (potentially) axon packing density distributions in a target related manner, which may at least in part explain the target-related differences of connection probabilities between them (*35*). Recent advances in reconstructing individual neurons throughout the rodent brain may allow testing the general validity of our theory for long-range connectivity (*49*). It is less likely that the theory for structural scaffoldings generalizes to inhibitory connections. Even though morphologies of inhibitory neurons adapt to the specific columnar and laminar layout of a cortical area (*50, 51*), developmental programs that shape their neuron and neurite distributions are fundamentally different from those of excitatory neurons (*52, 53*). Inhibitory axons preferentially target specific cell types and/or even specific subcellular compartments (*54*). At electron microscopic resolution, structural properties can hence be predictive for excitatory, but not for inhibitory connections (*26*). Whether developmental programs that shape inhibitory neuron and neurite distributions contribute to structural scaffoldings needs however further investigation.

Evolutionary benefits of structural scaffoldings

In simpler organisms such as c. elegans, genomes have in principle the capacity to specify every connection between every neuron to the minutes detail (*55*). This is arguably not possible for the mammalian neocortex, even if the entire genome would solely encode cortical connections. Because of this 'genetic bottleneck' (*55*), it is believed that the genome may instead specify a set of rules, which provide blueprints for wiring up the nervous system during development. However, we show that the neocortex' characteristic network architecture does not necessarily rely on such explicitly encoded wiring rules. Instead, reflecting a result of genetically induced neuron and neurite development, wiring properties that emerge as structural scaffoldings are encoded implicitly in the genome – i.e., in the form of structural constraints that are consistent across animals. Moreover, periphery-driven activity can regulate guidance programs that shape the neocortex' structural composition (*56*). Accordingly, sensory experience potentially alters the properties of structural scaffoldings. Compared to explicitly encoding wiring diagrams – or sets of wiring rules, implicit encoding could thus facilitate adaptation of network architectures to the environment. The emergence of structural scaffoldings as a result of neuron and neurite development may hence represent an evolutionary strategy for the neocortex to ensure robust function, while providing flexibility to organisms for invading new ecological niches within relatively few generations (*57*).

Outlook

The present theory reveals general limitations for interpreting connectivity data. For empirical observations that are consistent with the properties of structural scaffoldings – no matter whether they violate Peters' Rule or deviate from a random network – it is not possible to infer unambiguously which wiring mechanisms may have formed these connections. Thus, disentangling the origins and relevance of cortical wiring patterns will become even more challenging, and will likely require combining structural with functional connectivity measurements. For example, estimating the strengths of connections revealed upper bounds for the fraction of synaptic clusters that could reflect Hebbian learning (*26*). Inconsistency with the present theory however allows concluding that such observations must reflect, at least in part, wiring mechanisms that rely on activity and/or cellular identity. For example, the particularly high density of excitatory synapses along specific dendritic compartments of L5PTs (*58*), as well as

overrepresentation of unidirectional loops between these neurons (*43*), cannot solely emerge from the structural composition **(Fig. 8E)**. Such observations that deviate qualitatively from the theory are hence prime candidates to reflect structural correlates of learning or genetically encoded wiring rules. To facilitate identification of such candidates, we developed ***CortexInSilico***. Beside access to all data reported here, this framework provides web-based interfaces to contribute empirical data for comparison with the theory, to formulate alternative wiring strategies mathematically, to compare predictions across strategies, and to test the predictions' effect on cortex function by downloading connectivity constraints for neuron (*38*) and network simulations (*45*).

Materials and Methods

All relevant data and codes are available from the authors. The model, including detailed documentation of all data and analyses will be publically available online via ***CortexInSilico***. We used custom-written routines in C++, Python, or MATLAB 2019b software (Mathworks, Natick, MA, USA) for analysis. Amira software (FEI) was used for visualization. Boxplots were generated with the Matlab built-in *boxplot* where the bottom and top of the box represents the 25th and 75th percentiles, and the line within the box the median. The lines extend to the adjacent values. Outliers are all values more than 1.5 times the interquartile range away from the top or bottom of the box.

NeuroNet: We used NeuroNet, a custom-designed extension package for Amira software, to model an entire neocortical area and to predict the area's structural scaffoldings. NeuroNet has been described in detail previously (*25*). Briefly, NeuroNet requires (i) a reference frame for a neocortical volume of interest (e.g., area-specific pial and white matter surfaces), (ii) the 3D soma density distribution within the volume, (iii) a sample of cell type-specific axon and dendrite morphologies and (iv) their respective distributions of pre- and postsynaptic structures along their axons and dendrites. NeuroNet generates a digital model of the respective volume by up-scaling the sample of neuron morphologies to all neuron somata in the volume, i.e., each neuron soma is represented by one axon and dendrite morphology from the sample of morphologies. Using **Equations 1-3** NeuroNet generates the 'statistical connectome', the likelihoods of all possible wiring diagrams between all neuron morphologies in the model. Hence, connectivity between neuron pairs is not binary. Instead, for each neuron pair NeuroNet outputs the pair's probabilities of forming *n* synapses in any subvolume as defined by the resolution limit of the reference frame. Note that based on these pair- and subvolume-wise probability statistics an unlimited amount of realizations of wiring diagrams can be generated.

Anatomical data: As input to NeuroNet we used averages of the empirically determined rat barrel cortex' geometry (Wistar, male/female, postnatal day 28; n=12 (*27*)) and 3D distributions of all excitatory and inhibitory neuron somata in rat barrel cortex and in the VPM (Wistar, male, postnatal day 28; n=4 for barrel cortex, n=3 for VPM (*28*)). The precision of the barrel map was defined as the standard deviation (STD) of the positions of the barrel top and bottom along the arc and row across twelve rats (*27*). The layer precision was defined as the mean of the STDs of each layer border across four rats (*28*). The cellular precision was defined as the coefficient of variation (CV) of the number of somata per barrel column across four rats (*28*). Given the precision of the barrel field map the resolution limit of the resultant digital model was defined by cubes of 50 µm edge length.

The model was populated with reconstructions of *in vivo* labeled excitatory neuron morphologies (Wistar, male/female, postnatal day 25-45; n=153 dendrite reconstructions, n=74 dendrite/axon reconstructions (*11*)) and the intracortical part of *in vivo* labeled VPM axon morphologies (Wistar, male/female, postnatal day 25-170; n=14 neurons (*29*)). For space filling purposes we

incorporated reconstructions of inhibitory neurons into the model (Wistar, male/female, postnatal day 18-35; 203 morphologies were labeled *in vitro* in layers (Ls) 2 to 6 (*59-61*), 10 were labeled *in vivo* in L1 (*62*)). *In vitro* labeled dendrite morphologies were curated by assuming radial symmetry of their axons and dendrites. Inhibitory neurons were not further analyzed and integrated using the same procedure as for the excitatory morphologies.

To test how representative the sample of excitatory axon and dendrite morphologies was, we performed the following analysis: First, the dendrite morphologies of each excitatory cell type registered to the D2 barrel column were aligned by their lateral soma position and their dendrite innervation volume per 50 µm cube was calculated. For each cell type, we determined all possible combinations of subsamples of dendrite morphologies. If there were more than 500 possible combinations for a given subsample size of morphologies, a random sample of 500 combinations was used. For each combination of morphologies the change of their total innervation volume to the volume innervated by all morphologies was calculated, i.e., one minus the ratio between the number of cubes innervated by the respective combination of morphologies and those innervated by all morphologies. Second, for each combination of morphologies their respective dendrite length contribution per cube was determined. For each subsample size of morphologies, the CV of the dendrite length per cube across all possible combinations of morphologies was calculated. We repeated the same analysis for the axon morphologies of each cell type but without aligning the axon morphologies by their somata.

We created multiple barrel cortex models where subsamples of excitatory morphologies were used for up-scaling, i.e., models based on only one morphology per cell type, two morphologies per cell type, and so on. Specifically, we determined all possible combinations of subsamples of morphologies per cell type. If there were more than 500 possible combinations for a given number of morphologies per cell type, a random sample of 500 combinations was used. Up-scaled versions of morphologies that were not in the selected combination were removed from the model. To compensate for the removed morphologies, the remaining morphologies in the model were duplicated to match the overall number of neurons in the model in a cell type and column-specific manner. If the number of morphologies per cell type equaled or exceeded the number of morphologies of a particular cell type, no morphologies of the respective cell type were removed.

Across L2 to L6 of the C2 barrel column we determined the mean, SD, and CV of the length density and diversity of axon and dendrites within each 50 µm cube depending on the number of morphologies per cell type used for up-scaling in the model. The resultant barrel cortex model used here comprised 477,537 excitatory and 69,810 inhibitory neurons in barrel cortex, and 6,225 in the VPM. For each 50 µm cube in a subvolume spanning L2 to L6 of the C2 barrel column, we calculated the number of axon and dendrite branchlets and compared their ratio to empirical data for validation (*36*). The color maps for visualizing the distribution of somata, dendrites and axons are adapted from www.ColorBrewer2.org, by Cynthia A. Brewer, Penn State.

Conversion of morphologies into pre- and postsynaptic densities: Axons and dendrites are not represented by their respective trajectories. Instead, morphologies are converted into densities per 50 µm cube of pre- and postsynaptic structures. As reported in (*25*), the densities of presynaptic structures (i.e., axonal boutons) were derived by multiplying the axon length that each neuron contributes to a particular 50 µm cube with the number of boutons per length, as measured for the respective cell type and target layer (*11*). We compared the resultant presynaptic density distributions to empirical data: We calculated the number of excitatory boutons per 50 µm cube for the D2 barrel column, and grouped the 50 µm cubes by their respective laminar positions within L1 to L6. Layer borders were defined by the excitatory somata density profile along the cortical depth, as reported previously (*28*). The border between L2 and L3 was defined by the inhibitory somata density profile along the cortical depth, as reported previously (*63, 64*). Due to the consistency of the presynaptic density distribution with empirical data and the difficulties in

measuring densities of postsynaptic structures (i.e., spines), we scaled the total number of postsynaptic structures in the barrel cortex model to match the empirically validated number of presynaptic structures along the cortical depth. Specifically, the number of postsynaptic structures along excitatory dendrites (i.e., spines) was derived by assuming that spine densities are proportional to dendritic length. The number of postsynaptic structures along inhibitory dendrites and somata was derived by assuming proportionality to their respective surface areas. The derived density of postsynaptic structures for excitatory neurons ranged from 1.04 to 1.68 spines per µm dendritic length, depending on the subcellular compartment and cell type. The derived density of postsynaptic structures for inhibitory neurons was 0.74 per µm² of dendritic or somatic surface. The derived densities are consistent with empirical spine density measurements (*65, 66*) and with empirical synapse density measurements on inhibitory somata (*67, 68*).

Statistical connectome

We applied **Equations 1-3** to the neuron and neurite distributions of the barrel cortex model yielding a barrel cortex' statistical connectome. In **Equation 1,** $\sum_N POST_{(N,\hat{x})}$ refers to the total number of postsynaptic structures contributed to subvolume $\hat{x}$ by both, excitatory and inhibitory, neurons. All analysis was performed on the predicted pairwise connection probability statistics between excitatory neurons unless otherwise noted. Realizations of wiring diagrams derived from these statistics were only used for illustration. The color map for the matrix representation was mapped on the respective percentiles of the connection probability distribution; the color map for the zoom-in of the matrix representation was limited to 95% of the connection probability values. We computed the mean, STD, CV, and Pearson mode skewness (mode was defined as the most frequent connection probability value when rounded four digits to the right of the decimal point) of the connection probabilities between all cell type groupings of the C2 barrel column. We determined which of seven functions (Gaussian, Half-Normal, Exponential, Gamma, Binomial, zero-inflated Poisson, zero-inflated Negative Binomial) best approximated the underlying connection probability distribution. For this, we converted the distribution of the connection probabilities into probability mass functions p_{pmf}. For discrete mass functions (i.e., zero-inflated Negative Binomial, zero-inflated Poisson, and Binomial), we mapped each bin (from 0% to 100%) of p_{pmf} on discrete values from 0 to 100. In the case of the gamma function, all zero connection probabilities were replaced with epsilon (2^{-52}). Each function was fitted by maximum likelihood estimation (MLE) either using the Matlab built-in *fitdist* or in case of the zero-inflated models using the Python libraries statsmodels 0.10.1 and NumPy 1.16.5. The best-fitting function was the one resulting in the lowest area difference between the fits and p_{pmf} or the mapped p_{pmf}.

We assessed correlations between neurons by calculating the quantity $DSC(A,B)$ between all presynaptic neurons A and VPM thalamus and all postsynaptic cells B in the C2 barrel column:

$$DSC(A,B) = \sum_{\hat{x}} DSC_{(A,B,\hat{x})} \qquad \textbf{Equation (4)}$$

We grouped all pre- and postsynaptic neurons by their cell type identity, and summed the $DSC(A,B)$ values across each presynaptic population. This resulted in the mean number of connections (i.e., in-degree) each postsynaptic neuron B receives from this population. We computed a linear regression fit and Pearson's linear correlation coefficient between the in-degrees of two different presynaptic populations onto all neurons per postsynaptic cell type. We repeated this computation for all possible combinations of presynaptic cell types, resulting in the distribution of 55 correlation coefficients per postsynaptic cell type.

We used the Matlab built-in *digraph* to illustrate one possible realization of the statistical connectome for a network of 50 neurons located within L5 of the C2 barrel column as a graph. Edges between each neuron pair in the model were realized based on their predicted connection probability. In the random graph, edges were realized based on the mean connection probabilities

across all 50 neurons. To analyze network topologies, we calculated the (occurrence) probability of each of the 15 motifs for a given set of 10,000 randomly selected neuron triplets. Each neuron triplet is represented only once per set. The neuron triplets were sampled such that each neuron was from one of the ten cell types or a neuron from L2/3, L4, L5, or L6, and located closest to the C2 barrel column. We repeated this step 10 times and calculated the mean of each motif probability. We compared the motif probabilities predicted by the statistical connectome with those expected in a random network. First, we calculated the mean connection probabilities for each of the six edges between all three neuron populations. Second, we used these mean connection probabilities to compute the probability of each motif. Third, the predicted motif (occurrence) probabilities were divided by their respective expected (occurrence) probabilities in the random network. We computed the deviation of motif occurrences of all 15 motifs for all 210 cell type-specific triplet combinations with at least two different cell types. For each triplet combination, we calculated the mean and CV of their connection probability distribution across all six edges and their mean correlation of in-degrees (i.e., mean across all in-degree correlation coefficients involving the cell types of the respective triplet combination).

We extended our analysis to motifs between more than three neurons. We computed the probabilities of two exemplary motifs for up to ten neurons located closest to the C2 barrel column. The first motif represented 'full recurrence', i.e., all neurons are connected via bi-directional edges. We randomly sampled 10,000,000 sets of neurons per motif size (i.e., 3 to 10) from the model, and computed the probability of the fully bi-directional motif, respectively. The respective probabilities in the random network were computed based on the mean connection probability across all neurons of the sample. The second motif represented a feed-forward chain of neurons (i.e., neurons except for two have exactly one incoming and one outgoing unidirectional edge, and the two remaining neurons have either only one incoming or one outgoing unidirectional edge). We randomly sampled 10,000 sets of neurons per motif size (i.e., 3 to 10) from the model, computed the probability of 1,000 randomly sampled synaptic chain motif per neuron set, respectively.

Comparison with empirical data

The probabilities that two neurons a and b form zero, one, two, three and (at least) four synapses was calculated for all excitatory pairs and 50 µm cubes of the C2 barrel column, and multiplied by the number of neuron pairs that innervate each cube, respectively. The predicted median, minimum and maximum synapses per neuron pair were compared with saturated reconstructions (*26, 36*) that determined occurrences of the respective number of synapses between the same branches empirically, by scaling the respectively reconstructed volumes (1.5×10^{-6} mm^3 and 5×10^{-4} mm^3) to a 50 µm cube. We repeated the analysis for 50 µm cubes located in L4 of the C2 barrel column.

We compared 89 empirical measurements of connection probabilities from 29 studies (*35, 39-43, 69-91*) with the predicted mean connection probability by mimicking the respective experimental conditions and neuron groupings. To emulate the respective experimental conditions in the model, we created ten virtual slices of 300 µm thickness through the model. Each slice contained either the entire C2 barrel column or parts of it. The slices were shifted by 20 µm with respect to one another along the rostro-caudal axis. Neurites of all neurons whose somata were located within a slice were truncated, i.e., branches were cut at their intersection with the slice surface, and branches that became disconnected from the soma were removed from the model. We computed the connection probabilities between each neuron pair in the virtual slices as defined by **Equations 1-3** with the quantity DSC being the contribution of pre- and postsynaptic structures by the truncated neuron pairs' morphologies with respect to the total number of postsynaptic structures contributed by all neurons. We grouped the neurons as described in the respective studies (*35, 39-43, 69-91*) (**Table S1**); i.e., by their laminar soma location and – if reported – by

their cell type. Layer borders were defined as reported previously (*28*). The borders between L2 and L3, L5A and L5B, and L6A and L6B were set to the center location of L2/3, L5 and L6, respectively. If the recording depth was not reported, we restricted the comparison to neuron pairs within the mean reported range of recording depths (31 µm to 130 µm). To assess the predictions, we computed the Pearson's linear correlation coefficient between the empirical and predicted connection probabilities and the 95% confidence bounds for new observations based on a linear regression with no intercept using the Matlab built-ins *fitlm* and *predict*. We performed a random permutation test on the correlation coefficient by shuffling the empirical and the predicted connection probabilities and re-computing their correlation coefficient. We repeated this step 100,000 times.

We tested to what degree connections from excitatory onto inhibitory neurons affect structural scaffoldings between excitatory populations. For this, we removed the inhibitory neurons from the model, thus altering the quantity DSC (**Equation 1**) and the derived connection probabilities (**Equations 2-3**). The removal had no qualitative impact on the structural scaffoldings between excitatory neurons. The resultant correlation coefficient between the predicted and the empirical average connection probabilities was the same as with inhibitory neurons (R=0.75). We compared the model predictions with two empirical studies that performed connectivity measurements *in vitro* as a function of inter-somatic distance (*42, 69*) (**Tables S2-3**). Here, we grouped neurons additionally by their inter-somatic distance along the lateral axis (i.e., the axis running parallel to the slicing surface).

We tested Peters' Rule at cellular scales. As defined previously (*35*) we aligned all neuron morphologies registered to the D2 barrel column by their somata in the horizontal plane (i.e., preserving their cortical depth). We transformed each axon and dendrite morphology in a 3D axon or dendrite length density with a resolution of 50 µm cubes and multiplied each axon length density with each dendrite length density. For each neuron pair we summed the resulting axo-dendritic overlap across all cubes. We calculated the mean and STD across all cell-to-cell overlaps for each cell type combination. To compare these values to the respective connection probabilities predicted by the barrel cortex model we performed the following analysis: First, we calculated the correlation coefficient between the mean of the connection probability and of the axo-dendritic overlap. Second, we calculated the ratios between all axo-dendritic overlaps across all cell type combinations and the ratios between all connection probabilities across all cell type combinations. We assessed whether these ratios are linearly related with a slope of one, i.e., whether an increase of the axo-dendritic overlap between two cell type combinations is reflected in the same increase of the respective connection probabilities. Specifically, we performed 100 trials of 1,000 randomly sampled ratios, fitted a linear model with an intercept and tested whether the fitted slope was significantly different from a slope of one using the Matlab built-in *coefTest*.

We compared the predicted deviations of motif occurrences across 10 sets of 10,000 L5PT triplets with empirical observations (*43*). We compared the motif probabilities across the number of edges in motifs of eight neurons in the barrel cortex model and a random network to empirical observations (*42*). Therefore, we randomly sampled 10,000 sets of eight L5PTs. For each set of neurons and each number of possible edges (ranging from 0 to 56 possible edges), we computed the number of possible edge combinations (e.g., 1 possible combination of 0 or 56 edges, but 3.6×10^{10} possible combinations of 10 edges). If the number of edge combinations was less than 5,000, we iterated over all possible combinations. If the number of combinations was larger than 5,000, we randomly generated 5,000 motifs that matched the number of edges. We calculated the respective (occurrence) probability of each edge motif in the barrel cortex model and a random network constraint by the predicted mean connection probability between L5PT neurons.

Mathematical foundation of the theory for the emergence of structural scaffoldings

We investigated how properties of the mean and STD of the connection probability distributions in the absence of correlations relate to their motif occurrences. Suppose that K connection probabilities p_i are drawn from any generating distribution $Q(p|\mu,\sigma)$, which is parameterized by its mean μ (i.e., sparsity) and variance σ (i.e., heterogeneity). Each of the K connections x_i is drawn independently with probability $P(x_i = 1|p_i) = p_i$. The probability of observing a fully recurrent motif is accordingly $P(x_1 = x_2 = \cdots\ x_K = 1|p_1, p_2, \ldots p_K) = \prod_{i=1}^{K} p_i$. The occurrence of such motifs in a network is hence the expected value of the underlying distribution $Q(p|\mu,\sigma)$:

$$P(motif) = E_Q\left(\prod_{i=1}^{K} p_i\right) = \prod_{i=1}^{K} E_Q(p_i) = \mu^K \qquad \textbf{Equation (5)}$$

Thus, in the absence of correlations, motif occurrences are independent from the network's heterogeneity and only depend on the mean of the underlying distribution.

Now suppose that the connection probabilities between any two edges are correlated. We therefore developed a simplified mathematical model of a network with correlated connectivity that allows investigating how properties of the connection probability distributions, such as their mean, STD, and correlation, relate to their motif occurrences. The model is closely related to a model studied in (*92*). In the following, $\mathcal{N}(\mu,\lambda)$ denotes a Gaussian distribution with mean μ and variance λ, $\varphi(t,\mu,\lambda)$ denotes the respective Gaussian probability density function evaluated at *t*, $\Phi(s,\mu,\lambda)$ denotes the respective cumulative probability density function evaluated at *s,* and the complementary cumulative probability density function is defined by $L(s,\mu,\lambda) = 1 - \Phi(s,\mu,\lambda)$. For simplicity, the model assumes that whether there is an *i*-th edge between two nodes (denoted by X_i = 1, otherwise X_i = 0) is the result of a combination of only one 'private' source T_i, and one 'shared' source *S*. The bigger the shared source *S* is relative to the private one T_i, the more correlated the resultant connection probabilities are. Note the model can be easily generalized by incorporating more shared sources (e.g. per cell type combination). The mathematical model has two parameters: λ, bounded between 0 and 1, and representing the magnitude of the shared source, and thus the degree of correlation in the sources. As we will demonstrate in the following, λ also determines the heterogeneity of connection probabilities - the bigger λ is, the more heterogeneous the connection probabilities are. The second parameter, γ, represents the degree of connectivity - the greater γ is, the bigger connection probabilities are.

We define that the *i*-th edge exists (i.e., X_i = 1) whenever the joint input of T_i and *S*, denoted by Z_i, is larger than 0:

$$X_i = 1 \text{ whenever } Z_i > 0 \text{ where} \qquad \textbf{Equation (6)}$$

$$Z_i = \gamma + \sqrt{\lambda}S + \sqrt{\eta}T_i$$

where

$$\eta = 1 - \lambda\ ,\ S \sim \mathcal{N}(0,1),\ T_i \sim \mathcal{N}(0,1)$$

$$Z_i \sim \mathcal{N}(\gamma, \eta + \lambda) = \mathcal{N}(\gamma, 1)$$

Thus, $cov(Z_i, Z_j) = \lambda$. If $\lambda = 1$, X_i is only determined by the shared source *S*, while if $\lambda = 0$, X_i is only determined by the private source T_i. Given this mathematical model, the connection probability p_i for each edge X_i is given by:

$$p_i(S) = P(X_i = 1|S) = L(0, \gamma + \sqrt{\lambda}S, \eta) \qquad \textbf{Equation (7)}$$

And we likewise get

$$\mu = E_s(p_i) = L(0, \gamma, 1) \qquad \textbf{Equation (8)}$$

$$\begin{aligned} \sigma^2 &= Var_s(p_i) \\ &= \int_{-\infty}^{\infty} P(X_i = 1|s)^2 \varphi(s, 0,1) ds - \mu^2 \\ &= \int_{-\infty}^{\infty} L(0, \gamma + \sqrt{\lambda}s, \eta)^2 \varphi(s, 0,1) ds - \mu^2 \end{aligned} \qquad \textbf{Equation (9)}$$

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

$$\begin{aligned} cov(X_i, X_j) &= E(X_i, X_j) - \mu^2 = P(X_i = X_j = 1) - \mu^2 \\ &= \int_{-\infty}^{\infty} P(X_i = 1|s) P(X_j = 1|s) \varphi(s, 0,1) ds - \mu^2 \\ &= \int_{-\infty}^{\infty} L(0, \gamma + \sqrt{\lambda}s, \eta)^2 \varphi(s, 0,1) ds - \mu^2 \\ &= Var(p_i) \end{aligned} \qquad \textbf{Equation (10)}$$

Thus, in the simplified mathematical model the covariance of the connections is equal to the variance of connection probabilities – the more strongly the connection probabilities vary, the more strongly the connections themselves are correlated. Hence, the parameter λ represents a measure of both, the degree of correlation and heterogeneity. To assess the impact of λ and the mean connection probability μ onto motif occurrences and deviations, the probability that k out of K connections of a motif are realized is given by:

$$P(|X| = k) = \int_{-\infty}^{\infty} \binom{K}{k} L(0, \gamma + \sqrt{\lambda}s, \eta)^k \Phi(0, \gamma + \sqrt{\lambda}s, \eta)^{K-k} \varphi(s, 0,1) ds \qquad \textbf{Equation (11)}$$

This mathematical model was implemented as a numerical simulation in Matlab. We iterated over 250 γ-values ranging from -2 to 2, and over 250 λ-values ranging from 0 to 1. Per combination of γ and λ values, 10 trials, each with 100,000 random samples, were generated. For each trial, the mean and variance across the connection probabilities p_i, (i.e., μ and σ²), the probability of each triplet motif $P(|X| = k)$ with K=6 (i.e., maximal number of edges in a triplet), and the respective probability expected in random network based solely on μ was calculated.

The details of the mathematical model for one trial are below:

Algorithm 1: Mathematical model of correlated connectivity

Input: simulator with degree of correlations and heterogeneity λ and degree of connectivity γ.

randomly initialize shared source $S \sim \mathcal{N}(0,1)$.

$\eta := 1 - \lambda$

$p_i := L(0, \gamma + \sqrt{\lambda}S, \eta)$ // p_i is a vector of connection probabilities

$\mu := E(p_i)$ // E denotes the expected value

$\sigma^2 := Var(p_i)$ // Var denotes the variance

$K := 6$ // maximal number of edges in a triplet

for k = 0 to K do

$P(k) \coloneqq \binom{K}{k} E(p_i{}^k (1-p_i)^{K-k})$ // probability of triplet motif with k edges

$P_{random}(k) \coloneqq \binom{K}{k} \mu^k (1-\mu)^{K-k}$ // probability of triplet motif with k edges in // random network

return P, P_{random}, μ, σ^2

The deviation of motif occurrences from a random network of each triplet motif was the ratio between the means of the motif probabilities across all trials. The deviations were mapped on a grid spanned by 20 μ-values (i.e., sparsity) and 20 λ-values (i.e., correlations & heterogeneity) and visualized by a log-space color map. Each of the 220 cell type-specific triplet combinations was mapped into the grid space. Specifically, for each combination its respective λ-value was inferred based on the variance and mean of the connection probabilities of each combination and a lookup table of μ, σ², and λ-values as determined by the numerical simulation.

CortexInSilico

The web-based interface *CortexInSilico* provides access to all data and analysis routines reported here, including the anatomical data used to generate the statistical connectome model. The interface allows the user to group neurons arbitrarily and to access their aggregated summary statistics as visualizations and for download. The user can contribute empirical connectivity measurements and compare them to the theory's predictions. To explore the impact of alternative wiring strategies onto wiring diagrams, **Equations 2-3** can be modified within the web-based interface.

References

1. V. Braitenberg, A. Schüz, *Cortex: statistics and geometry of neuronal connectivity.*, (Springer, 1998).
2. A. R. Kriegstein, S. C. Noctor, Patterns of neuronal migration in the embryonic cortex. *Trends Neurosci* **27**, 392-399 (2004).
3. R. Ayala, T. Shu, L. H. Tsai, Trekking across the brain: the journey of neuronal migration. *Cell* **128**, 29-43 (2007).
4. J. L. Lefebvre, J. R. Sanes, J. N. Kay, Development of dendritic form and function. *Annu Rev Cell Dev Biol* **31**, 741-777 (2015).
5. E. T. Stoeckli, Understanding axon guidance: are we nearly there yet? *Development* **145**, (2018).
6. B. G. Rash, E. A. Grove, Area and layer patterning in the developing cerebral cortex. *Current opinion in neurobiology* **16**, 25-34 (2006).
7. A. Simi, M. Studer, Developmental genetic programs and activity-dependent mechanisms instruct neocortical area mapping. *Curr Opin Neurobiol* **53**, 96-102 (2018).
8. J. W. Yang *et al.*, Development of the whisker-to-barrel cortex system. *Curr Opin Neurobiol* **53**, 29-34 (2018).
9. K. D. Harris, G. M. Shepherd, The neocortical circuit: themes and variations. *Nature Neuroscience* **18**, 170-181 (2015).
10. W. H. Bosking, Y. Zhang, B. Schofield, D. Fitzpatrick, Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. *Journal of Neuroscience* **17**, 2112-2127 (1997).

11. R. T. Narayanan *et al.*, Beyond Columnar Organization: Cell Type- and Target Layer-Specific Principles of Horizontal Axon Projection Patterns in Rat Vibrissal Cortex. *Cereb Cortex* **25**, 4450-4468 (2015).
12. S. J. Barnes, G. T. Finnerty, Sensory experience and cortical rewiring. *Neuroscientist* **16**, 186-198 (2010).
13. D. Feldmeyer *et al.*, Barrel cortex function. *Prog Neurobiol* **103**, 3-27 (2013).
14. K. Y. Kwan, N. Šestan, E. Anton, Transcriptional co-regulation of neuronal migration and laminar identity in the neocortex. *Development* **139**, 1535-1546 (2012).
15. K. P. Berry, E. Nedivi, Spine Dynamics: Are They All the Same? *Neuron* **96**, 43-55 (2017).
16. S. H. Bennett, A. J. Kirby, G. T. Finnerty, Rewiring the connectome: Evidence and effects. *Neurosci Biobehav Rev* **88**, 51-62 (2018).
17. A. Antonini, M. P. Stryker, Rapid remodeling of axonal arbors in the visual cortex. *Science* **260**, 1819-1821 (1993).
18. A. Antonini, M. Fagiolini, M. P. Stryker, Anatomical correlates of functional plasticity in mouse visual cortex. *Journal of Neuroscience* **19**, 4388-4406 (1999).
19. J. R. Sanes, M. Yamagata, Many paths to synaptic specificity. *Annu Rev Cell Dev Biol* **25**, 161-195 (2009).
20. S. Yogev, K. Shen, Cellular and molecular mechanisms of synaptic specificity. *Annu Rev Cell Dev Biol* **30**, 417-437 (2014).
21. J. de Wit, A. Ghosh, Specification of synaptic connectivity by cell surface interactions. *Nat Rev Neurosci* **17**, 22-35 (2016).
22. S. L. Zipursky, J. R. Sanes, Chemoaffinity revisited: dscams, protocadherins, and neural circuit assembly. *Cell* **143**, 343-353 (2010).
23. B. A. Hassan, P. R. Hiesinger, Beyond Molecular Codes: Simple Rules to Wire Complex Brains. *Cell* **163**, 285-291 (2015).
24. J. E. Vaughn, Fine structure of synaptogenesis in the vertebrate central nervous system. *Synapse* **3**, 255-285 (1989).
25. R. Egger, V. J. Dercksen, D. Udvary, H. C. Hege, M. Oberlaender, Generation of dense statistical connectomes from sparse morphological data. *Front Neuroanat* **8**, 129 (2014).
26. A. Motta *et al.*, Dense connectomic reconstruction in layer 4 of the somatosensory cortex. *Science* **366**, (2019).
27. R. Egger, R. T. Narayanan, M. Helmstaedter, C. P. de Kock, M. Oberlaender, 3D reconstruction and standardization of the rat vibrissal cortex for precise registration of single neuron morphology. *PLoS Comput Biol* **8**, e1002837 (2012).
28. H. S. Meyer *et al.*, Cellular organization of cortical barrel columns is whisker-specific. *Proc Natl Acad Sci U S A* **110**, 19113-19118 (2013).
29. M. Oberlaender, A. Ramirez, R. M. Bruno, Sensory experience restructures thalamocortical axons during adulthood. *Neuron* **74**, 648-655 (2012).
30. A. Peters, M. L. Feldman, The projection of the lateral geniculate nucleus to area 17 of the rat cerebral cortex. I. General description. *J Neurocytol* **5**, 63-84 (1976).
31. V. Braitenberg, A. Schüz, in *Anatomy of the cortex*. (Springer, 1991), pp. 109-112.
32. C. L. Rees, K. Moradi, G. A. Ascoli, Weighing the Evidence in Peters' Rule: Does Neuronal Morphology Predict Connectivity? *Trends Neurosci* **40**, 63-71 (2017).
33. Y. Mishchenko *et al.*, Ultrastructural analysis of hippocampal neuropil from the connectomics perspective. *Neuron* **67**, 1009-1020 (2010).
34. K. L. Briggman, M. Helmstaedter, W. Denk, Wiring specificity in the direction-selectivity circuit of the retina. *Nature* **471**, 183-188 (2011).
35. S. P. Brown, S. Hestrin, Intracortical circuits of pyramidal neurons reflect their long-range axonal targets. *Nature* **457**, 1133-1136 (2009).
36. N. Kasthuri *et al.*, Saturated Reconstruction of a Volume of Neocortex. *Cell* **162**, 648-661 (2015).

37. A. Santuy, J. R. Rodriguez, J. DeFelipe, A. Merchan-Perez, Volume electron microscopy of the distribution of synapses in the neuropil of the juvenile rat somatosensory cortex. *Brain Struct Funct* **223**, 77-90 (2018).
38. R. Egger *et al.*, Cortical Output Is Gated by Horizontally Projecting Neurons in the Deep Layers. *Neuron* **105**, 122-137 e128 (2020).
39. R. M. Bruno, B. Sakmann, Cortex is driven by weak but synchronously active thalamocortical synapses. *Science* **312**, 1622-1627 (2006).
40. Q. Q. Sun, J. R. Huguenard, D. A. Prince, Barrel cortex microcircuits: thalamocortical feedforward inhibition in spiny stellate cells is mediated by a small number of fast-spiking interneurons. *Journal of Neuroscience* **26**, 1219-1230 (2006).
41. Y. Yoshimura, J. L. Dantzker, E. M. Callaway, Excitatory cortical neurons form fine-scale functional networks. *Nature* **433**, 868-873 (2005).
42. R. Perin, T. K. Berger, H. Markram, A synaptic organizing principle for cortical neuronal groups. *Proc Natl Acad Sci U S A* **108**, 5419-5424 (2011).
43. S. Song, P. J. Sjostrom, M. Reigl, S. Nelson, D. B. Chklovskii, Highly nonrandom features of synaptic connectivity in local cortical circuits. *Plos Biol* **3**, 507-519 (2005).
44. S. R. Williams, S. J. Mitchell, Direct measurement of somatic voltage clamp errors in central neurons. *Nature Neuroscience* **11**, 790 (2008).
45. I. D. Landau, R. Egger, V. J. Dercksen, M. Oberlaender, H. Sompolinsky, The Impact of Structural Heterogeneity on Excitation-Inhibition Balance in Cortical Networks. *Neuron* **92**, 1106-1121 (2016).
46. M. N. Economo *et al.*, Distinct descending motor cortex pathways and their roles in movement. *Nature* **563**, 79-84 (2018).
47. G. Rojas-Piloni *et al.*, Relationships between structure, in vivo function and long-range axonal target of cortical pyramidal tract neurons. *Nature Communications* **8**, 1-11 (2017).
48. C. Guo *et al.*, Single-axon level morphological analysis of corticofugal projection neurons in mouse barrel field. *Scientific reports* **7**, 1-9 (2017).
49. J. Winnubst *et al.*, Reconstruction of 1,000 projection neurons reveals new cell types and organization of long-range connectivity in the mouse brain. *Cell* **179**, 268-281. e213 (2019).
50. G. Quattrocolo, G. Fishell, T. J. Petros, Heterotopic transplantations reveal environmental influences on interneuron diversity and maturation. *Cell Reports* **21**, 721-731 (2017).
51. Y. Ishino *et al.*, Regional cellular environment shapes phenotypic variations of hippocampal and neocortical chandelier cells. *Journal of Neuroscience* **37**, 9901-9916 (2017).
52. S. A. Anderson, D. D. Eisenstat, L. Shi, J. L. Rubenstein, Interneuron migration from basal forebrain to neocortex: dependence on Dlx genes. *Science* **278**, 474-476 (1997).
53. L. Lim, D. Mi, A. Llorca, O. Marin, Development and Functional Diversification of Cortical Interneurons. *Neuron* **100**, 294-313 (2018).
54. Y. Kubota, Untangling GABAergic wiring in the cortical microcircuit. *Current opinion in neurobiology* **26**, 7-14 (2014).
55. A. M. Zador, A critique of pure learning and what artificial neural networks can learn from animal brains. *Nature Communications* **10**, 3770 (2019).
56. D. E. Feldman, M. Brecht, Map plasticity in somatosensory cortex. *Science* **310**, 810-815 (2005).
57. K. C. Catania, Behavioral pieces of neuroethological puzzles. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol* **203**, 677-689 (2017).
58. A. Karimi, J. Odenthal, F. Drawitsch, K. M. Boergens, M. Helmstaedter, Cell-type specific innervation of cortical pyramidal cells at their apical dendrites. *Elife* **9**, e46876 (2020).
59. M. Arzt, B. Sakmann, H. S. Meyer, Anatomical Correlates of Local, Translaminar, and Transcolumnar Inhibition by Layer 6 GABAergic Interneurons in Somatosensory Cortex. *Cereb Cortex* **28**, 2763-2774 (2018).

60. M. Helmstaedter, B. Sakmann, D. Feldmeyer, L2/3 interneuron groups defined by multiparameter analysis of axonal projection, dendritic geometry, and electrical excitability. *Cereb Cortex* **19**, 951-962 (2009).
61. C. Koelbl, M. Helmstaedter, J. Lubke, D. Feldmeyer, A barrel-related interneuron in layer 4 of rat somatosensory cortex with a high intrabarrel connectivity. *Cereb Cortex* **25**, 713-725 (2015).
62. R. Egger *et al.*, Robustness of sensory-evoked excitation is increased by inhibitory inputs to distal apical tuft dendrites. *Proc Natl Acad Sci U S A* **112**, 14072-14077 (2015).
63. H. S. Meyer *et al.*, Inhibitory interneurons in a cortical column form hot zones of inhibition in layers 2 and 5A. *Proc Natl Acad Sci U S A* **108**, 16807-16812 (2011).
64. R. T. Narayanan, D. Udvary, M. Oberlaender, Cell Type-Specific Structural Organization of the Six Layers in Rat Barrel Cortex. *Front Neuroanat* **11**, 91 (2017).
65. A. U. Larkman, Dendritic morphology of pyramidal neurones of the visual cortex of the rat: III. Spine distributions. *J Comp Neurol* **306**, 332-343 (1991).
66. Y. Kawaguchi, F. Karube, Y. Kubota, Dendritic branch typing and spine expression patterns in cortical nonpyramidal cells. *Cereb Cortex* **16**, 696-711 (2006).
67. A. Keller, E. L. White, Synaptic organization of GABAergic neurons in the mouse SmI cortex. *J Comp Neurol* **262**, 1-12 (1987).
68. B. Ahmed, J. C. Anderson, K. A. Martin, J. C. Nelson, Map of the synapses onto layer 4 basket cells of the primary visual cortex of the cat. *J Comp Neurol* **380**, 230-242 (1997).
69. M. Avermann, C. Tomm, C. Mateo, W. Gerstner, C. C. Petersen, Microcircuits of excitatory and inhibitory neurons in layer 2/3 of mouse barrel cortex. *Journal of Neurophysiology* **107**, 3116-3134 (2012).
70. A. P. Bannister, A. M. Thomson, Dynamic properties of excitatory synaptic connections involving layer 4 pyramidal cells in adult rat and cat neocortex. *Cereb Cortex* **17**, 2190-2203 (2007).
71. M. Beierlein, B. W. Connors, Short-term dynamics of thalamocortical and intracortical synapses onto layer 6 neurons in neocortex. *Journal of Neurophysiology* **88**, 1924-1932 (2002).
72. M. Beierlein, J. R. Gibson, B. W. Connors, Two dynamically distinct inhibitory networks in layer 4 of the neocortex. *Journal of Neurophysiology* **90**, 2987-3000 (2003).
73. R. M. Bruno, D. J. Simons, Feedforward mechanisms of excitatory and inhibitory cortical receptive fields. *Journal of Neuroscience* **22**, 10966-10975 (2002).
74. C. M. Constantinople, R. M. Bruno, Deep cortical layers are activated directly by thalamus. *Science* **340**, 1591-1594 (2013).
75. S. R. Crandall, S. L. Patrick, S. J. Cruikshank, B. W. Connors, Infrabarrels Are Layer 6 Circuit Modules in the Barrel Cortex that Link Long-Range Inputs and Outputs. *Cell Reports* **21**, 3065-3078 (2017).
76. D. Feldmeyer, V. Egger, J. Lubke, B. Sakmann, Reliable synaptic connections between pairs of excitatory layer 4 neurones within a single 'barrel' of developing rat somatosensory cortex. *Journal of Physiology* **521 Pt 1**, 169-190 (1999).
77. D. Feldmeyer, J. Lubke, B. Sakmann, Efficacy and connectivity of intracolumnar pairs of layer 2/3 pyramidal cells in the barrel cortex of juvenile rats. *Journal of Physiology* **575**, 583-602 (2006).
78. D. Feldmeyer, J. Lubke, R. A. Silver, B. Sakmann, Synaptic connections between layer 4 spiny neurone-layer 2/3 pyramidal cell pairs in juvenile rat barrel cortex: physiology and anatomy of interlaminar signalling within a cortical column. *Journal of Physiology* **538**, 803-822 (2002).
79. D. Feldmeyer, A. Roth, B. Sakmann, Monosynaptic connections between pairs of spiny stellate cells in layer 4 and pyramidal cells in layer 5A indicate that lemniscal and

paralemniscal afferent pathways converge in the infragranular somatosensory cortex. *Journal of Neuroscience* **25**, 3423-3431 (2005).
80. S. B. Hofer *et al.*, Differential connectivity and response dynamics of excitatory and inhibitory neurons in visual cortex. *Nature Neuroscience* **14**, 1045-U1146 (2011).
81. C. Holmgren, T. Harkany, B. Svennenfors, Y. Zilberter, Pyramidal cell communication within local networks in layer 2/3 of rat neocortex. *Journal of Physiology* **551**, 139-153 (2003).
82. X. L. Jiang *et al.*, Principles of connectivity among morphologically defined cell types in adult neocortex. *Science* **350**, (2015).
83. J. S. Jouhanneau, J. Kremkow, A. L. Dorrn, J. F. Poulet, In Vivo Monosynaptic Excitatory Transmission between Layer 2 Cortical Pyramidal Neurons. *Cell Reports* **13**, 2098-2106 (2015).
84. J. S. Jouhanneau, J. Kremkow, J. F. A. Poulet, Single synaptic inputs drive high-precision action potentials in parvalbumin expressing GABA-ergic cortical neurons in vivo. *Nature Communications* **9**, 1540 (2018).
85. P. Krieger, T. Kuner, B. Sakmann, Synaptic connections between layer 5B pyramidal neurons in mouse somatosensory cortex are independent of apical dendrite bundling. *Journal of Neuroscience* **27**, 11473-11482 (2007).
86. S. Lefort, C. Tomm, J. C. F. Sarria, C. C. H. Petersen, The Excitatory Neuronal Network of the C2 Barrel Column in Mouse Primary Somatosensory Cortex. *Neuron* **61**, 301-316 (2009).
87. H. Markram, J. Lubke, M. Frotscher, A. Roth, B. Sakmann, Physiology and anatomy of synaptic connections between thick tufted pyramidal neurones in the developing rat neocortex. *Journal of Physiology* **500**, 409-440 (1997).
88. A. Mercer *et al.*, Excitatory connections made by presynaptic cortico-cortical pyramidal cells in layer 6 of the neocortex. *Cereb Cortex* **15**, 1485-1496 (2005).
89. C. C. H. Petersen, B. Sakmann, The excitatory neuronal network of rat layer 4 barrel cortex. *Journal of Neuroscience* **20**, 7579-7586 (2000).
90. G. Silberberg, H. Markram, Disynaptic inhibition between neocortical pyramidal cells mediated by martinotti cells. *Neuron* **53**, 735-746 (2007).
91. A. M. Thomson, D. C. West, Y. Wang, A. P. Bannister, 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. *Cereb Cortex* **12**, 936-953 (2002).
92. J. H. Macke, M. Opper, M. Bethge, Common input explains higher-order correlations and entropy in a simple model of neural population activity. *Physical Review Letters* **106**, 208102 (2011).

Acknowledgments

We thank Kevin Briggman for discussions and advice; Vincent J. Dercksen and Robert Egger for developing early versions of the online model; and Haim Sompolinsky, and Peter L. Strick for comments on the manuscript. Funding was provided from the Center of Advanced European Studies and Research, the Max Planck Institute for Biological Cybernetics, the Center for Neurogenomics and Cognitive Research, the European Research Council under the European Union's Horizon 2020 research and innovation program (grant agreement 633428; to M.O.), the German Federal Ministry of Education and Research (grants BMBF/FKZ 01GQ1002 and 01IS18052; to M.O. and J.M.), and the Deutsche Forschungsgemeinschaft (SFB 1089 and SPP 2041 to M.O. and J.M.; and EXC number 2064/1, 39072764 to J.M.); **Author contributions:** M.O. conceived and designed the study. D.U. developed the model and analysis routines. P.H. and H.H. developed the online model. J.M. developed the mathematical model. C.K. and B.S. provided data. D.U. and M.O. analyzed the data, and wrote the paper with help from all authors. **Declaration of Interests:** The authors declare no competing interests.