Strong and localized coupling controls dimensionality of neural activity across brain areas

David Dahmen[1,*], Stefano Recanatesi[2,*], Xiaoxuan Jia[3], Gabriel K. Ocker[3,4], Luke Campagnola[3], Tim Jarsky[3], Stephanie Seeman[3], Moritz Helias[1,5,+], and Eric Shea-Brown[2,3,+]

[1]Institute of Neuroscience and Medicine (INM-6 and INM-10) and Institute for Advanced Simulation (IAS-6), Jülich Research Centre, Jülich, Germany
[2]University of Washington Center for Computational Neuroscience and Swartz Center for Theoretical Neuroscience, Seattle, WA, USA
[3]Allen Institute, Seattle, WA, USA
[4]Department of Mathematics and Statistics, Boston University, MA, USA
[5]Department of Physics, Faculty 1, RWTH Aachen University, Aachen, Germany
[*]These authors share first authorship
[+]These authors share senior authorship

Neurons in the brain interact through a complex network of recurrent connections, but how complex is their activity? This is measured by dimensionality: the number of degrees of freedom in network activity. We use a novel type of network theory to identify a regime where network connectivity constrains and controls dimensionality – the strongly coupled *sensitive regime*. Through new analysis of high-density Neuropixels recordings, we find strong evidence that areas across the mouse brain operate in this sensitive regime. We then study connectivity within brain areas via analysis of a massive, newly released synaptic physiology dataset from both mouse and human. We quantify the building blocks of this connectivity in terms of network motifs, and find that cell-type specific synaptic motifs that impact the sensitive regime are prevalent in both mouse and human brains. Thus, we propose that local circuitry strongly constrains and controls the dimensionality of brain networks.

Introduction

A hallmark of neural circuits is their highly complex recurrent connectivity. Complexity of dynamics can be measured by the number of collective degrees of freedom across the network – the dimensionality of network activity. Dimensionality is tightly linked to the computational properties of neural systems. Signal classification, for example, generally benefits from network operations that expand the dimensionality of the incoming signals that are to be categorized (1–5). Generalization of classification or other network computations to new signals, however, benefits from operations that suitably compress inputs into lower-dimensional activity patterns (6, 7). Dimensionality has been quantified in diverse neural systems (8). In different experimental settings and protocols, studies have emphasized the relatively high (5, 9, 10) or low (11, 12) dimensionality of recordings. Moreover, the dimensionality of neural dynamics can change over time (13), across the processing hierarchy (14), or during learning (7, 15, 16). Taken together, these findings underscore the importance of dimensionality as a flexible property of network activity, and one that will depend on the nature of the computation performed in a circuit.

A fundamental question is therefore the mechanism by which a network's structure – its connectivity – controls the dimensionality of its dynamics (17–21). Here, we address this question for networks in the widely studied balanced state believed to characterize cortex (22, 23), where weakly correlated activity (24) results from the dominance of inhibitory recurrent feedback (25). We develop a theory that reduces vastly complex connectivity patterns to a single number, the *recurrency* R, which measures the effective strength of recurrent coupling and directly determines dimensionality via a fixed, one-to-one relationship (Fig. 1a). We find that strong coupling leads to not only low dimensionality, but also to its sensitive dependence to changes in R. Moreover, we quantify how a highly tractable and localized feature of network connectivity – connectivity motifs among pairs and triplets of cells (21, 26–34) – can strongly modulate R and hence dimensionality in the strongly coupled regime (Fig. 1b).

We then present evidence that the brain operates in this strongly coupled and sensitive regime, through a novel analysis of recent large scale electrophysiology recordings (35, 36) that use Neuropixels probes to record from more than 30,000 neurons. We find systematic trends in the resulting dimensionality across brain areas and across cortical layers (Fig. 1b). We then analyze vast newly released synaptic physiology datasets (37, 38), in which synaptic connections among more than 22,000 pairs of neurons were probed, and we find that the connectivity motifs implicated by our theory as regulating R are strongly present in both mouse and human brain. We quantify their predicted contribution to dimensionality across brain areas and layers, and show how previously established patterns of cell-type specific modulation and adaptation can have a new effect: to further regulate connectivity motifs and hence dimensionality across time and brain state.

Electrophysiology recordings display signatures of strongly recurrent dynamics across brain areas

Do brain networks operate in a strongly recurrent regime?

Recent theoretical work has developed a robust way to assess the strength of recurrent coupling based on activity

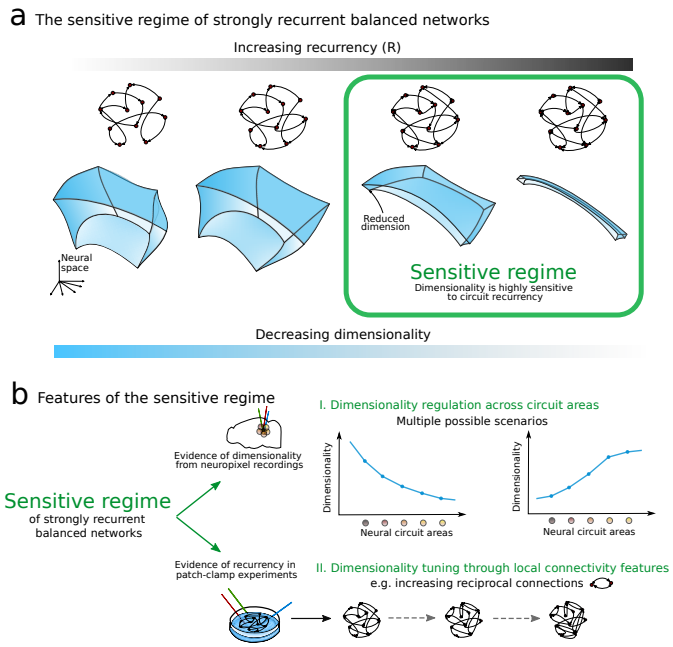

Fig. 1. Sensitive regime of balanced networks. a) We find that dimensionality decreases as recurrency increases in balanced networks. Electrophysiology and synaptic physiology datasets point to neural networks operating in a strongly coupled sensitive regime where network connectivity is strongly recurrent and the dimensionality is highly sensitive to the strength of recurrency. b) Computational advantages of the sensitive regime: I) Top: brain circuits along functional pathways can regulate the dimensionality of their responses. II) Bottom: the dimensionality can be tightly controlled by local synaptic features such as reciprocal connections.

measurements from neural circuits (18). We start by using this method, previously applied only to a single brain area (macaque motor cortex), to analyze large-scale neural activity data recorded across multiple regions of the mouse brain. These data were recorded by the Allen Institute for Brain Science using recently developed, high density Neuropixels probes, Fig. 2b (for details see (35, 36)). We analyzed 32,043 neurons across 15 brain areas (Table S1), recorded during sessions lasting on average more than 3 hours (cf. sample of 2 minutes of recorded activity, Fig. 2c). We focused on periods where either no stimulus was presented to the animal (spontaneous condition) or where drifting gratings were displayed (evoked condition), cf. Methods and Fig. S1.

The method builds on the assumption that neural networks of cortical and subcortical circuits operate in the so-called balanced state (22, 23), where excitatory and inhibitory synaptic currents dynamically cancel each other to produce fluctuation-driven asynchronous irregular spiking activity (22, 24). In this regime the strength of the recurrent coupling, theoretically corresponding to the radius R of the effective connectivity spectrum underlying the neural dynamics (Fig. 3b), can be assessed by measuring the statistics of the network's variability (18)

$$R = \sqrt{1 - \sqrt{1/(1+Ns^2)}}. \qquad \textbf{(1)}$$

Specifically, $s = \frac{\delta c}{\bar{a}} = \mathcal{O}(1/\sqrt{N})$ is the ratio between the standard deviation δc of cross-covariances and the average auto-covariance $\bar{a}$, Fig. S2a. While the radius R is indepen-

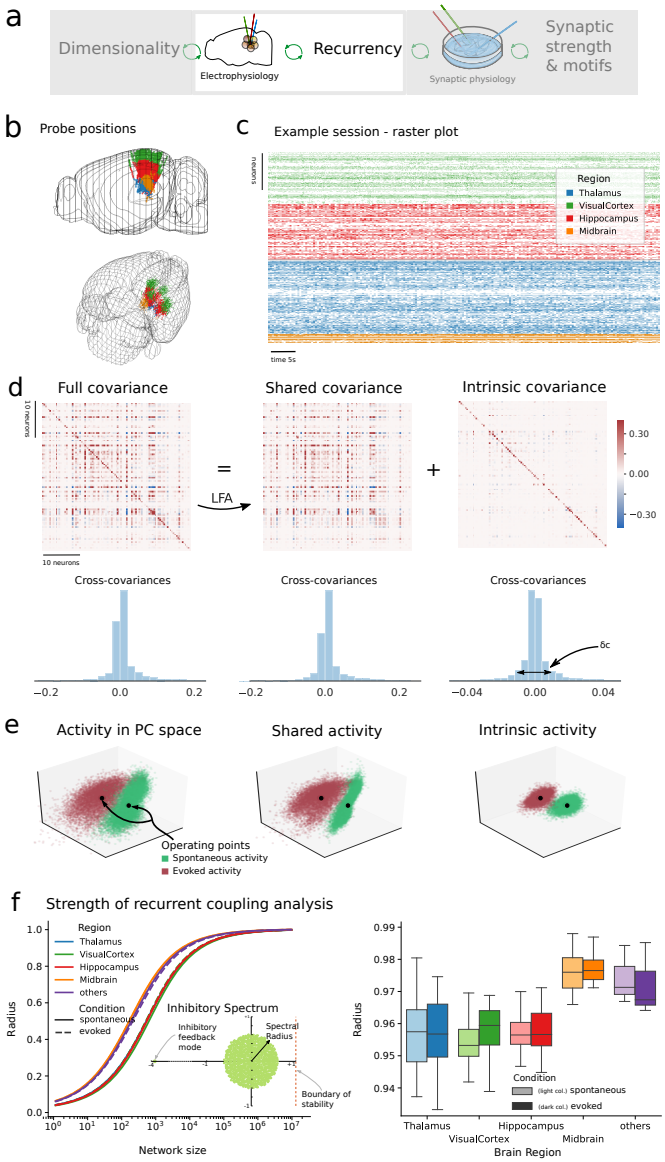

Fig. 2. Estimation of recurrent coupling strength in electrophysiology recordings. a) Figure focus: network recurrency inferred via electrophysiology recordings. b) Sites of Neuropixels recordings colored by brain region. c) Raster plot example of Neuropixels recordings for one experimental session (session id=715093703). d) Top panels: Schematic of Latent Factor Analysis decomposition of the full covariance into shared and intrinsic covariances (cf. Fig. S3a). Bottom panels: distribution of cross-covariances for the three matrices. e) Neural activity of example session for spontaneous (green) and evoked (red) condition in the coordinate axes given by the top Principal Components (PC) determined across both conditions. Evoked condition corresponds to drifting grating stimuli with 75 repeats per stimulus orientation. The three panels represent respectively the total, shared and intrinsic activity (cf. Methods, Fig. S3). Operating points are defined as the average activity per condition. f) Strength of recurrent coupling. Left: Average spectral radius inferred from neural data, as a function of the network size. Inset: eigenvalue spectrum of connectivity matrix. Right: inferred spectral radius for network size 10^6 across conditions and brain regions.

dent of network size, its inference from measured covariance statistics s using Eq. (1) requires additional knowledge of the underlying network size N. Therefore, in the following we employ N as a free parameter and use anatomical data on neuron densities as constraints to obtain a lower bound on R. Importantly, the above theoretical result for R relies on the internally generated *intrinsic* variability s, which is due to the reverberation of ongoing fluctuations through the network (cf. the histogram of intrinsic cross-covariances in Fig. 2d). In electrophysiology recordings there is, however, typically a second contribution to covariances due to *shared* variability across neurons that is often linked to input signals to the network or behavioral low-rank components of the activity (10). Assessing the statistics of intrinsic variability from electrophysiology recordings is therefore challenging. Here we build a two-step procedure to estimate s, and thus R.

First, to separate intrinsic and shared variability we worked under a linear assumption for the network dynamics around each network's state, or *operating point* (Fig. 2e, cf. Methods). In this case the shared and intrinsic variability contributions independently influence the covariance matrix C of neural activity (Fig. 2d):

$$C = C_{\text{shared}} + C_{\text{intrinsic}}. \tag{2}$$

To identify shared sources of variability in the neural activity we exploited a cross-validated Latent Factor Analysis (LFA) procedure (17) that yields the number of shared factors across the neural populations and allows us to factor out their contribution to network activity (Figs. S3a to S3e). We validated this procedure both via simulations in our network model with known ground truth (Fig. S4) and by using cross-validated Principal Component Analysis in place of LFA (Fig. S6).

Second, to address biases in the estimates of s due to finite sampling (13) (cf. Methods) we carried out a theoretically grounded subsampling procedure on the shared and intrinsic activity components to fit the dependence of cross-covariances on the number of samples. This yielded an unbiased estimate of s, Figs. S3f to S3g.

A key fact is that the resulting estimated statistic of cross-covariances s depends on the operating point, identified by the activity profile of the neural population and other neural properties (e.g. adaptation mechanisms, gain modulation etc.). Henceforth the recurrent coupling strength R is not a static property of a neural circuit, but it depends on the underlying experimental condition, as illustrated by the different operating points in Fig. 2e. We thus inferred the recurrent coupling strength R for each brain region and experimental condition separately. For spontaneous activity we thereby controlled robustness of results with respect to different behaviorally related stationary periods using a Hidden Markov Model (Fig. S8), and for the evoked condition we confirmed consistency of our results across orientations of the drifting gratings (Fig. S7).

The outcome of our analysis was that the spectral radius across all regions and conditions was predicted to be at least $R = 0.95$ on a scale from 0 to 1, with 1 marking the threshold to coupling so strong that activity would be come linearly unstable. This bound corresponded to values of the network size $N \geq 10^6$ (Fig. 2f). As recent experiments report a cell density across the mouse cortex to fall in between $0.48 \cdot 10^5$ cells$/mm^3$ in orbital cortex to $1.55 \cdot 10^5$ cells$/mm^3$ in visual cortex (39), these results are consistent with neural activity, across all brain regions analyzed, being generated by neural networks operating in a strongly recurrent regime.

Sensitive control of dimensionality

We now show that the strongly recurrent regime found in the previous section corresponds to the most fundamental metric of the complexity of neural activity – its dimensionality – being under sensitive control of the recurrent coupling strength R.

To show this, we study the participation ratio D_{PR}, a widely used measure of linear dimensionality which accounts for the extent to which neural responses are spread along different axes directions; in many often-encountered settings D_{PR} corresponds to the number of principal components required to capture roughly 80% of a signal's variability (11) (Fig. S2b). D_{PR} is given by

$$D_{\text{PR}}(C) = \frac{\left(\sum_i \lambda_i\right)^2}{\sum_i \lambda_i^2}, \tag{3}$$

where λ_i is the eigenvalue associated with the $i-$th principal component. This measure can be rewritten in terms of the statistics of the covariance matrix (13) (Fig. S2a) and, in large balanced networks of size N, its leading contribution comes from the relative dispersion s of intrinsic cross-covariances across neurons $D_{\text{PR}} = \frac{N}{1+Ns^2}$ (cf. Suppl. Mat.). Combined with Eq. (1), this yields a one-to-one relation between the dimensionality of intrinsic covariances and the spectral radius so that $D_{\text{PR}}(C_{\text{intrinsic}}) = N(1-R^2)^2$, Fig. 3b (cf. Suppl. Mat. for a formal derivation and Fig. S10 for validation in complex nonlinear spiking networks; for an alternative derivation see (21)).

The relationship between D_{PR} and R shows that the dimensionality of the network smoothly decreases with increasing spectral radius towards $R = 1$, which is the coupling level at which the network becomes (linearly) unstable, Fig. 3b. In strongly recurrent regimes, where $R \lessapprox 1$, the network's dimensionality is tightly constrained, being substantially smaller than its number of neurons. Furthermore, in this regime the relative change in dimensionality with respect to R (Fig. 3c)

$$\frac{\delta D_{\text{PR}}}{D_{\text{PR}}} = \frac{dD_{\text{PR}}}{dR}\frac{1}{D_{\text{PR}}} = \frac{4R}{R^2-1} \tag{4}$$

is greatest. Thus, networks with strong recurrent coupling achieve sensitive control of their dimensionality as a function of their coupling strength.

In sum, the high values of R estimated above throughout the regions of mouse brain (Fig. 2f), together with the decreasing relationship between dimensionality and spectral radius R, suggest that the dimensionality for these brain regions will

be low – and hence in a regime where it is under sensitive control. We confirm this next.

To compare recordings with different numbers of neurons we developed a theoretically unbiased extrapolation of the dimensionality as a function of the number of neurons recorded N_{rec} (cf. Methods and Fig. S11). This enabled us to extrapolate the estimates of dimensionality up to realistic values of the size of local circuits (i.e., $N = \mathcal{O}(10^6)$). It also allowed a systematic comparison between the dimensionality of the intrinsic covariance and of the full covariance. By construction, in balanced networks, this extrapolation saturates at $D_{\mathrm{PR}} = 1/s^2$ for the dimensionality of intrinsic covariance, while it is a function of multiple moments of the covariance statistics for the dimensionality of the full covariance (cf. Methods).

Applying the procedure above to the Allen Institute Neuropixels data showed that the extrapolated dimensionality of the full covariance (cf. Fig. 3d) saturated for network sizes $N = \mathcal{O}(10^5)$, at values on the order of ~ 100 dimensions. The dimensionality of intrinsic covariances saturated at higher values of several hundreds of dimensions, Fig. 3e. These two estimates can be taken as a lower and upper bound, respectively, of the dimensionality of the network's activity, which thus appears to be consistently described by a few hundred dimensions – across all brain regions. This is a small number when compared to the number of neurons in the network: indeed, for a network of 10^5 neurons this corresponds to a dimensionality of less than 1% of its size.

Before moving to a more detailed analysis and interpretation of dimensionality within brain areas, we confirm that our techniques reproduce two established effects. The first is that stimuli are known to reduce the dimensionality of responses in cortical activity (13, 40). We found that in the evoked condition, versus the spontaneous, the dimensionality of activity in visual cortex was indeed significantly lower (t-test p-value$< 10^{-3}$, Fig. 3d). The second is a recent estimate of the dimensionality of cortical responses to visual stimuli (9) which we link to the framework developed here (cf. Fig. S9).

We concluded that neural networks in regions across the brain, and across experimental conditions, operate in a regime where their dimensionality is under sensitive control by the strength of recurrent coupling R. We can interpret this as the ability to flexibly set the number of modes that can participate in a computation, a feature which may play a substantial functional role across the brain.

Dimensionality across the visual hierarchy and cortical layers

Does dimensionality of neural responses relate to information processing in neural circuits? If so, the ability of specific brain areas – operating in the strongly coupled, sensitive regime described above – to modulate this dimensionality would acquire a functional role.

Several studies in deep and recurrent artificial neural networks have highlighted how dimensionality modulation (compression and expansion) in neural representations across

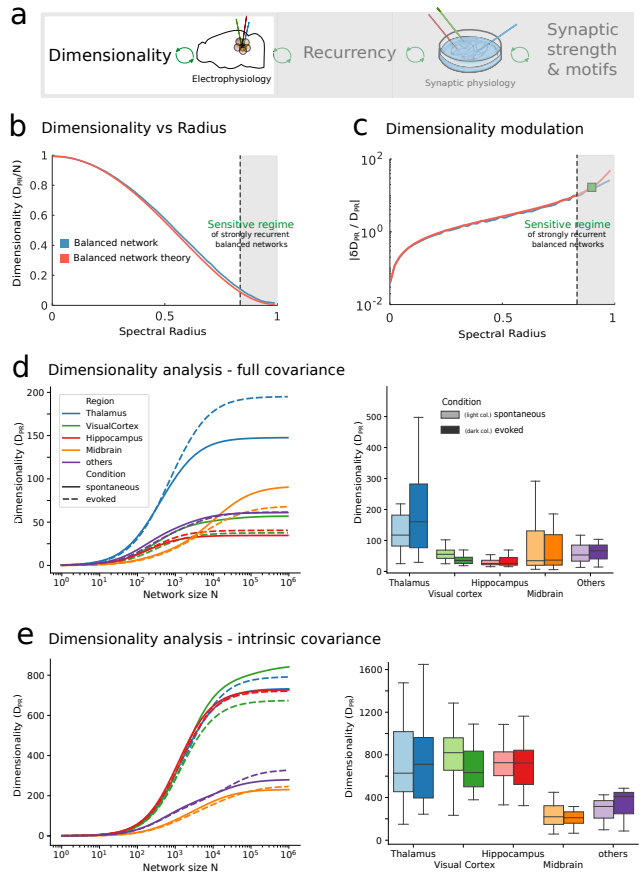

Fig. 3. Dimensionality across brain regions. a) Figure focus: Dimensionality inferred via electrophysiology recordings. b) Normalized dimensionality D_{PR}/N for a balanced network as a function of the spectral radius R. Blue: simulation, red: theory. Shaded area highlights sensitive regime consistent with electrophysiology recordings. c) Relative modulation of dimensionality as a function of the spectral radius, Eq. (4). Blue and red curves overlap. d) Left: Dimensionality extrapolation as a function of network size for the full covariance before applying Latent Factor Analysis (LFA). Lines represent mean values and error bars are shown in Fig. S5. Right: Dimensionality based on full covariance across brain regions for a network of size $N = 10^6$ neurons. Boxes capture lower and upper interquartile range of the variability across experimental recording sessions. e) Left: Dimensionality extrapolation based on intrinsic covariances upon applying LFA. Right: Dimensionality based on intrinsic covariances for a network of size $N = 10^6$ neurons. The number of latent modes individuated by the cross-validated LFA analysis was consistently low ($N_{\mathrm{factors}} \leq 16$) (Fig. S6b).

network layers (14, 19) and stages of learning (7) have functional roles in information processing. We next compute dimensionality on a finer scale than for the regions studied above – here for areas that subdivide those regions – to test this idea in data from diverse neural circuits. We focus first on the dimensionality of the full covariance, and then on the intrinsic dimensionality.

Specifically, we first studied the full activity of areas across the visual functional hierarchy (36). Analyzing the full covariance for the Neuropixels electrophysiology data revealed a trend of dimensionality expansion from primary visual to higher visual cortical areas (t-test max p-value across binary comparisons $< 10^{-15}$), Figs. 4a to 4c and Figs. S12a to S12b. Such a trend is consistent with the hypothesis that the visual stream performs a stimulus-dependent dimensionality expansion, akin to the trend described in artificial neural networks and often explained in terms of feature expansion of the input,

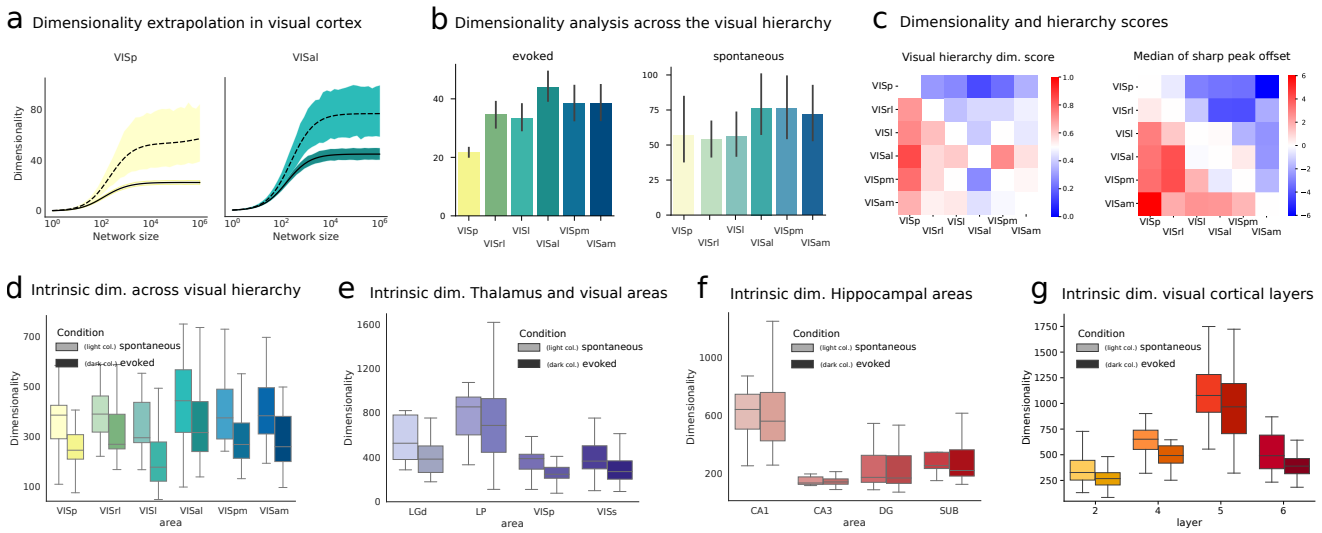

Fig. 4. Dimensionality analysis across brain areas and cortical layers. a) Dimensionality extrapolation as a function of number of neurons in the network based on the full covariance for primary visual cortex and downstream area VIS al (see (36)). b) Dimensionality analysis of the full covariance across the visual hierarchy (extrapolated for 10^6 neurons). Left panel: dimensionality increase along the visual hierarchy in the evoked condition. Right panel: dimensionality for the spontaneous condition, non-significant increase. c) Left: Ranked dimensionality analysis of full covariance across areas of the visual hierarchy for the evoked condition. Box color shows fraction of recordings (on a total number of recording sessions times the number of stimulus orientations for drifting gratings) where the dimensionality of areas reported in the rows is higher than the dimensionality of areas reported in the columns. Right: quantification of the visual hierarchy for each area. Combined median of functional delay for each pair of cortical areas. Box color reflects the shift from zero in milliseconds in the cross-correlogram between neurons from two areas. Analysis and plot reproduced from (36) with permission of the authors. d) Dimensionality across the visual hierarchy based on the intrinsic covariance. e) Dimensionality of intrinsic covariances for thalamic and visual brain areas. VISs corresponds to the activity of all secondary visual areas. f) Dimensionality of intrinsic covariances for hippocampal brain areas. g) Dimensionality of intrinsic covariances across visual cortical layers. For supplementary plots regarding the dimensionality of full and intrinsic covariances see Fig. S12.

Figs. S12a to S12b (14, 19, 42, 43). These results underscore the functional value of both dimensionality mechanisms and the visual hierarchy per se (36).

The dimensionality of intrinsic covariances was consistent with the hypothesis of visual cortical circuits being in the strongly recurrent regime, where dimensionality is under sensitive control. While the same trend of increasing dimensionality across the visual cortex hierarchy was not present for intrinsic covariances (Fig. 4d t-test max p-value=0.77 across binary comparisons), there were robust trends from thalamic to primary and secondary visual cortical areas (LGd and LP to VisP and VISs, Fig. 4e, t-test max p-value< 0.013) and across hippocampal areas (CA1, CA3, DG, SUB), Fig. 4f, t-test p-value CA1 vs CA3 $<10^{-15}$), suggesting robust differences in their intrinsic connectivity. Overall, areas considered to be possible input regions to broader circuits (LGd, LP and CA1) displayed a high dimensionality corresponding to a less recurrent, and potentially more feed forward, circuit, when compared with their visual cortex and hippocampal counterparts. The area CA3 in particular, known to have strong recurrent connections (44), appeared to have the lowest dimensionality of intrinsic covariances in line with such assumption.

Finally we considered whether different cortical layers could carry out similar functional roles in expanding or reducing the dimensionality of neural representations. We found that layers 2 and 5 had respectively the lowest and highest dimensionality (t-test max p-value across conditions $< 10^{-11}$). Intriguingly, this result is consistent with the hypothesis that layer 2 performs computation through strongly recurrent circuitry (45), Fig. 4g.

These systematic trends across brain areas and layers, in both the full and intrinsic dimensionality, suggest that the modulation of dimensionality across brain networks can be associated with functional information processing. The robust trends we described for intrinsic dimensionality reveal the potential for local circuitry to tune this dimensionality, a topic to which we turn in more detail next.

Local tuning of the global recurrent coupling strength R

We next asked how balanced neural networks can regulate their overall recurrent coupling strength R and hence their dimensionality. While the majority of previous works focused on how global properties of the recurrent anatomical connectivity affect the spectral radius R (18, 46, 47), we here focus on the impact of local connectivity motifs on the radius and the dimensionality. These motifs are statistics of the neural connectivity W that involve *pairs* of connections (see Methods), and are the fundamental local building blocks of networks.

Second order motifs appear in four types: reciprocal, divergent, convergent, and chain motifs (Fig. 5b), together with the variance of neural connections already present in purely random models (46). These motifs have been shown to play important roles determining neuron-to-neuron correlations and allied circuit dynamics (26–28, 32, 48–53) and emerge from learning rules consistent with biological STDP mechanisms (54, 55).

We here develop a general theory that takes full account of all second order motifs. Our approach generalizes previous work on reciprocal connections (56) or on networks of purely

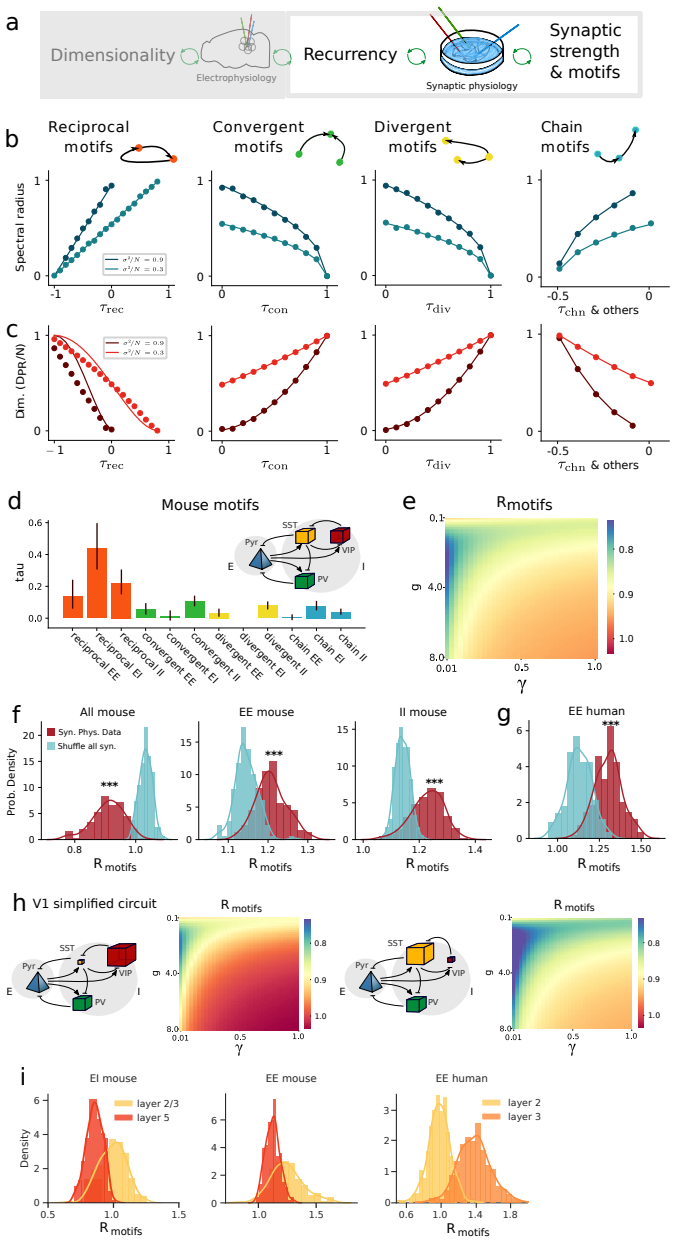

Fig. 5. Motif analysis in synaptic physiology datasets. a) Figure focus: network recurrency inferred via synaptic physiology datasets. b) Theoretical dependence of spectral radius on motif abundances. c) Theoretical dependence of dimensionality on motif abundances. d) Motif abundances in mouse V1. Inset: Simplified V1 circuit diagram with only prevalent connections (41). e) Inferred radius as a function of relative strength g of inhibitory and excitatory synapses and ratio γ of inhibitory to excitatory population size. f) Estimation of R_{motifs} from mouse data (500 bootstraps based on random subsets of 80% of sessions). Shuffle of synapses within each experimental session preserving EI synapse type (shuffle EI syn.). Effect of all EI motifs ($g = 4$ and $\gamma = 0.25$, red point in panel e), of only EE motifs, or only II motifs (cf. Fig. S14). g) Same as f for EE motifs in human dataset. h) Effect of VIP regulation on R_{motifs}. Left: SST is inactive and R_{motifs} is computed over the displayed circuit involving PV and VIP. Right: SST is active and VIP is inhibited resulting in PV and SST balancing the activity of the Pyr population. i) Layer-wise estimation of R_{motifs} for EI balanced motifs, EE-only motifs in mouse and EE-only motifs in human. Same bootstrapping as in f.

excitatory neurons (51), and is consistent with recent results for single-population networks that were independently obtained by powerful but distinct theoretical tools based on the full covariance spectrum (21). Our analysis yields a novel compact analytical quantity:

$$R = \sigma \frac{1 - \tau_{\text{div}} - \tau_{\text{con}} - 2\tau_{\text{chn}} + \tau_{\text{rec}}}{\sqrt{1 - \tau_{\text{div}} - \tau_{\text{con}}}} \,, \qquad \textbf{(5)}$$

where τ_{rec}, τ_{chn}, τ_{div}, τ_{con} denote correlation coefficients between pairs of synapses that capture the abundance of reciprocal, chain, divergent, and convergent motifs, respectively (cf. Methods and Suppl. Mat.) and σ captures the overall synaptic strength. This formula describes how the spectral radius R is affected by increasing or decreasing the prevalence of second order motifs (Fig. 5a) and thus links the modulation of auto- and cross-covariances and the dimensionality of neural responses across the global network to the statistics of local circuit connectivity, as shown in Figs. 5b to 5c. While Eq. (5) is exact for the simplest type of balanced networks, which are networks of inhibitory neurons whose recurrent interactions balance the excitatory external input, we show that it generalizes to standard models of balanced excitatory-inhibitory networks (57). Here, σ and τ combine the corresponding statistics of the excitatory and inhibitory subpopulations (cf. Fig. S13 and Suppl. Mat.). This direct link between quantifiable, local connectivity statistics and the global network property R opened the door to novel functional analyses of very large-scale synaptic physiology datasets in both mouse and human, as we describe next.

Cortical circuits in mouse and human employ local synaptic motifs to modulate their recurrent coupling

We analyzed newly released synaptic physiology datasets from both mouse and human cortex (37, 38) to assess the involvement of synaptic motifs in modulating the recurrent coupling strength. This synaptic physiology dataset was based on simultaneous in-vitro recordings of 3-to-8 cell groups (cf. Methods) and consisted of $1,368$ identified synapses from mouse primary visual cortex (out of more than $22,000$ potential connections that were tested) and 363 synapses from human cortex. The recurrent coupling strength R defined by Eq. (5) has an overall scaling term, σ, and a motif contribution term given by $R_{\text{motifs}} = R/\sigma$ which encapsulates whether the overall motif structure increases ($R_{\text{motifs}} > 1$) or decreases ($R_{\text{motifs}} < 1$) the spectral radius R. We begin by assessing the probability of occurrence of individual motifs and hence estimating R_{motifs}, cf. Methods. The relationship Eq. (5) defines a specific hypothesis for empirical motif statistics that modulate global circuit dimensionality: if they combine to produce $R_{\text{motifs}} > 1$ then they are tuned to reduce dimensionality, and vice-versa for $R_{\text{motifs}} < 1$.

Beginning with the mouse data, we calculated the statistics of individual motifs, separating those for excitatory (E) and inhibitory (I) synapses (EE, EI, II, Fig. 5d), and found many motifs to be significantly present. We then combined these to compute R_{motifs}. This requires two parameters: one regulating the overall ratio of inhibitory to excitatory neurons (γ), and another the relative strength of the inhibitory synapses

(g) (cf. Suppl. Mat.). We found that $R_{\mathrm{motifs}} < 1$ across all choices of these parameters (Fig. 5e). Many different motifs, differently involving E and I cell types, combine to produce this value of R_{motifs}. To study how this occurs, we separated the contribution to R_{motifs} from motifs within the excitatory population (EE type only) by assuming that other motifs occur at chance level. Interestingly, the EE motifs operating alone produced the opposite trend, increasing the radius $R_{\mathrm{motifs}}^{EE\ \mathrm{only}} > 1$ (Fig. 5f center, one-sided t-test p-value$< 10^{-20}$). The same was true for motifs within the inhibitory population $R_{\mathrm{motifs}}^{II\ \mathrm{only}} > 1$ (Fig. 5f right, one-sided t-test p-value p-value$< 10^{-20}$), and for motifs within the excitatory population in human cortical circuits (Fig. 5g). We further confirmed that this effect is also predicted for previously published data on excitatory connections in rat visual cortex (30) (cf. Methods). The increased recurrent coupling strengths within both the excitatory and inhibitory populations underscore the prominent role of EI motifs, specifically reciprocal EI motifs, in decreasing and potentially regulating the overall recurrent coupling strength to be $R_{\mathrm{motifs}} < 1$ (Fig. 5d, Fig. 5e, Fig. S13). Overall, the distinct roles of motifs among E and I cells types in regulating R_{motifs} point to ways that the radius, and hence dimension, may be controlled dynamically in neural circuits, as we explore next.

One pathway for this control is via cell types, which subdivide E and I populations (Fig. 5h) and are separately identified in the synaptic physiology dataset which we analyze. As Table S2e shows, reciprocal EI motifs were prevalent when the inhibitory interneuron was a somatostatin cell (SST) or a parvalbumin cell (PV), but not a VIP cell. Recent findings have shown that VIP interneurons (58, 59) are important regulators of cortical functions, are modulated by arousal and movement (60), and are recruited by reinforcement signals (61). We thus hypothesized that VIP interneurons could dial the recurrent coupling of cortical circuits by exerting disinhibitory control (the regulation of the overall activity by inhibiting a second population of inhibitory interneurons) over the SST neural population (41, 58) without being directly involved in regulating the balanced regime, given the almost absent recurrent connectivity within the VIP population and sparse connectivity with pyramidal cells (Fig. 5h). Under the simplest form of this hypothesis, there exists a mutual antagonism between VIP and SST populations that results in only one of these populations being active at a time, and we therefore derived the values of R_{motifs} for either case. We found that activation of the VIP pathway substantially increased R_{motifs} (Fig. 5h). This shows how VIP interneurons, which themselves may collect top-down signals from higher cortical areas, could selectively tune the dimensionality of local cortical activity. This adds another channel for control of information processing in cortical circuits to existing hypotheses on how VIP neurons regulate gain (59). Furthermore, we predicted a similar trend in increasing R_{motifs} to follow from short-term synaptic plasticity (STP) in modulating cell-type specific connections upon stimulus onset, although a detailed analysis awaits future investigation (Figs. S15a to S15c). Finally, we note that, our results were robust to inclusion of estimates of the relative synaptic strength of cell type specific connections (Fig. S15d) and the cell type specific prevalence of the three inhibitory subpopulations (Fig. S15d) (see Methods).

In sum, we asked whether the experimentally derived structure of cortical networks – quantified by their motifs – enables the tuning of recurrence strength R and hence dimensionality. We found that the answer is yes, and that the VIP disinhibitory pathway and STP modulation both provide examples of how motifs are likely to play a substantial role in this tuning. As we reviewed above, high dimensional activity can retain stimulus details, while lower dimensional activity can promote robust and general downstream decoding. Taken together, this points to new functional roles for modulatory and adaptive mechanisms known to take effect across time during stimulus processing and to be engaged across brain states (for example in an aroused versus passive states).
Finally, we find evidence for synaptic motifs in the cross-layer differences in the dimensionality of cortical activity identified above (Fig. 4g). There, activity in mouse cortex layer 2 showed lower dimensionality, corresponding to an increased overall recurrent coupling strength R compared to layers 4, 5 or 6. Intriguingly, the corresponding motif contribution R_{motifs} to this coupling strength was significantly stronger for layer 2 than for layer 4 or 5. Moreover, a similar result held true when performing the analysis on the human dataset for excitatory connections in layers 2 and 3 (Fig. 5i and Fig. S16).

Conclusion

We showed that neural networks across the mouse brain operate in a strongly recurrent regime. A feature of this regime is the ability of neural circuits to sensitively modulate the relative dimensionality of their activity patterns via their recurrent coupling strength. This has potentially important consequences for computation. Indeed, our analyses of large scale Neuropixels recordings from areas within the thalamus and hippocampus, and across layers in cortex, showed systematic trends in this dimensionality. Our theory links these findings to clear predictions for recurrent coupling strength in these areas: a higher dimensionality suggests a lower recurrent coupling strength and vice-versa. Moreover, we showed that the critical circuit features that determine a circuit's recurrent coupling strength R – and hence the dimensionality of its activity patterns – are not just its overall synaptic strength, but also a tractable set of local synaptic motifs. We use theoretical tools to quantify the effect of these motifs via a compact index R_{motifs}. This provides a concrete target quantity that can, as we show, be readily obtained from emerging, large-scale synaptic connectivity datasets and used to check predictions about the role of synaptic structure in controlling dimensionality. Thus theory and brain-wide experimental analyses converge to provide new evidence for an

intriguing concept: that the connectivity of brain networks exert global control over their activity in a highly tractable manner, via the building blocks of their local circuitry.

Acknowledgments

We thank Krešimir Josić, Brent Doiron, Luca Mazzucato, Stefan Mihalas, Nicholas Steinmetz, Yu Hu, and Leenoy Meshulam for helpful discussions. S.R. was supported by a Swartz Fellowship in Theoretical Neuroscience at the University of Washington, and by NIH BRAIN Grant R01EB026908, and E.S.B. by NIH R01EB026908 and NSF DMS Grant 1514743. D.D. and M.H. were supported by the HGF young investigator's group VH-NG-1028, the European Union's Horizon2020 research and innovation program under Grant agreements No. 785907 (Human Brain Project SGA2) and No.945539 (Human Brain Project SGA3), and funded under the Excellence Strategy of the Federal Government and the Länder (G:(DE-82)EXS-PF-JARA-SDS005). We thank the Allen Institute for Brain Science founder, Paul G. Allen, for his vision, encouragement, and support.

Bibliography

1. T. M. Cover, *IEEE Transactions on Electronic Computers* **EC-14**, 326 (1965).
2. V. N. Vapnik, *Statistical learning theory* (1998).
3. N. A. Cayco-Gajic, C. Clopath, R. A. Silver, *Nature Communications* **8**, 1116 (2017).
4. A. Litwin-Kumar, K. D. Harris, R. Axel, H. Sompolinsky, L. F. Abbott, *Neuron* **93**, 1153 (2017).
5. M. Rigotti, *et al.*, *Nature* **497**, 585 (2013). Number: 7451 Publisher: Nature Publishing Group.
6. S. Fusi, E. K. Miller, M. Rigotti, *Current opinion in neurobiology* **37**, 66 (2016).
7. M. Farrell, S. Recanatesi, G. Lajoie, E. Shea-Brown, *bioRxiv* p. 564476.
8. J. P. Cunningham, B. M. Yu, *Nature Neuroscience* **17**, 1500 (2014).
9. C. Stringer, M. Pachitariu, N. Steinmetz, M. Carandini, K. D. Harris, *Nature* **571**, 361 (2019).
10. C. Stringer, *et al.*, *Science* **364** (2019).
11. P. Gao, *et al.*, *BioRxiv* p. 214262 (2017).
12. J. A. Gallego, M. G. Perich, L. E. Miller, S. A. Solla, *Neuron* **94**, 978 (2017).
13. L. Mazzucato, A. Fontanini, G. La Camera, *Frontiers in Systems Neuroscience* **10** (2016).
14. U. Cohen, S. Chung, D. D. Lee, H. Sompolinsky, *Nature communications* **11**, 1 (2020).
15. S. Recanatesi, *et al.*, *bioRxiv* p. 471987 (2019).
16. M. e. a. Stern, In the footsteps of learning: Changes in network dynamics and dimensionality with task acquisition. (2020). COSYNE Conference Abstract.
17. R. C. Williamson, *et al.*, *PLOS Computational Biology* **12**, 1 (2016).
18. D. Dahmen, S. Grün, M. Diesmann, M. Helias, *Proceedings of the National Academy of Sciences* **116**, 13051 (2019).
19. S. Recanatesi, *et al.*, *arXiv preprint arXiv:1906.00443* .
20. C. Huang, *et al.*, *Neuron* **101**, 337 (2019).
21. Y. Hu, H. Sompolinsky, *bioRxiv* (2020).
22. C. van Vreeswijk, H. Sompolinsky, *Science (New York, N.Y.)* **274**, 1724 (1996).
23. A. S. Ecker, *et al.*, *Science* **327**, 584 (2010). WOS:000274020500039.
24. A. Renart, *et al.*, *Science* **327**, 587 (2010). WOS:000274020500040.
25. T. Tetzlaff, M. Helias, G. T. Einevoll, M. Diesmann, *Plos Computational Biology* **8**, e1002596 (2012). WOS:000308553500003.
26. Y. Hu, J. Trousdale, K. Josic, E. Shea-Brown, *Physical Review E* **89**, 032802 (2014). WOS:000332672200015.
27. V. Pernice, B. Staude, S. Cardanobile, S. Rotter, *Plos Computational Biology* **7**, e1002059 (2011). WOS:000291015800032.
28. J. Trousdale, Y. Hu, E. Shea-Brown, K. Josic, *Plos Computational Biology* **8**, e1002408 (2012). WOS:000302244000016.
29. Y. Hu, J. Trousdale, K. Josic, E. Shea-Brown, *Journal of Statistical Mechanics-Theory and Experiment* p. P03012 (2013). WOS:000316056900012.
30. S. Song, P. J. Sjostrom, M. Reigl, S. Nelson, D. B. Chklovskii, *Plos Biology* **3**, 507 (2005). WOS:000227984000018.
31. R. Perin, T. K. Berger, H. Markram, *Proceedings of the National Academy of Sciences* p. 201016051 (2011).
32. L. Zhao, B. Beverlin, T. Netoff, D. Q. Nykamp, *Frontiers in Computational Neuroscience* **5** (2011).
33. O. Sporns, *Discovering the human connectome* (MIT press).
34. F. Mastrogiuseppe, S. Ostojic, *Plos Computational Biology* **13**, e1005498 (2017). WOS:000402542900036.
35. V. C. . . S. dev documentation, https://allensdk.readthedocs.io/en/v2.2.0/visual_coding_neuropixels.
36. J. H. Siegle, *et al.*, *Nature* pp. 1–7 (2021).
37. L. Campagnola, *et al.*, *bioRxiv* (2021).
38. S. P. A. documentation, https://portal.brain-map.org/explore/connectivity/synaptic-physiology.
39. D. Keller, C. Erö, H. Markram, *Frontiers in neuroanatomy* **12**, 83 (2018).
40. M. M. Churchland, *et al.*, *Nature neuroscience* **13**, 369 (2010).
41. D. J. Millman, *et al.*, *Elife* **9**, e55130 (2020).
42. J. Shi, E. Shea-Brown, M. Buice, *Advances in Neural Information Processing Systems 32*, H. Wallach, *et al.*, eds. (Curran Associates, Inc., 2019), pp. 5764–5774.
43. E. Froudarakis, *et al.*, *bioRxiv* (2020).
44. M. P. Witter, *Learning & memory* **14**, 705 (2007).
45. S. Peron, *et al.*, *Nature* **579**, 256 (2020).
46. H. Sompolinsky, A. Crisanti, H.-J. Sommers, *Physical review letters* **61**, 259 (1988).
47. J. Aljadeff, D. Renfrew, M. Vegué, T. O. Sharpee, *Physical Review E* .
48. N. Brunel, *Nature Neuroscience* **19**, 749 (2016).
49. D. Martí, N. Brunel, S. Ostojic, *Physical Review E* **97**, 062314 (2018). Publisher: American Physical Society.
50. G. K. Ocker, *et al.*, *Current Opinion in Neurobiology* **46**, 109 (2017).
51. S. Recanatesi, G. K. Ocker, M. A. Buice, E. Shea-Brown, *PLoS computational biology* **15**, e1006446.
52. D. Zhang, C. Zhang, A. Stepanyants, *Journal of Neuroscience* **39**, 6888. Publisher: Soc Neuroscience.
53. Y. Hu, J. Trousdale, K. Josić, E. Shea-Brown, *Journal of Statistical Mechanics: Theory and Experiment* **2013**, P03012 (2013).
54. M. Gilson, A. N. Burkitt, D. B. Grayden, D. A. Thomas, J. L. van Hemmen, *Biological Cybernetics* **101**, 427 (2009). WOS:000272176000008.
55. G. K. Ocker, A. Litwin-Kumar, B. Doiron, *Plos Computational Biology* **11**, e1004458 (2015). WOS:000360824500049.
56. H. Sommers, A. Crisanti, H. Sompolinsky, Y. Stein, *Physical Review Letters* **60**, 1895 (1988).
57. N. Brunel, *Journal of Computational Neuroscience* **8**, 183 (2000). WOS:000087725300001.
58. M. M. Karnani, *et al.*, *Journal of Neuroscience* **36**, 3471 (2016).
59. K. A. Ferguson, J. A. Cardin, *Nature Reviews Neuroscience* **21**, 80 (2020).
60. Y. Fu, *et al.*, *Cell* **156**, 1139 (2014).
61. H.-J. Pi, *et al.*, *Nature* **503**, 521 (2013).

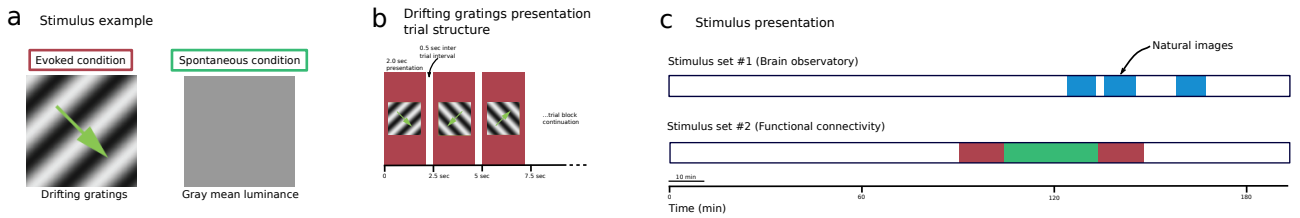

Fig. S1. Statistics of stimuli. a) Example of screen displayed to the animal during the evoked condition with drifting-grating stimuli (left) and the spontaneous condition (gray screen, right). b) Drifting gratings were presented with four different orientations ($0^\circ, 45^\circ, 90^\circ, 135^\circ$) and one temporal frequency (2Hz) for the duration of 2 seconds with 75 repeats each (0.5 seconds of inter-trial interval). Spontaneous activity was recorded for the duration of 30 minutes. The statistics of sessions is reported in Table S1. c) Position of the two experimental conditions (red: evoked condition, green: spontaneous condition) in the recordings for the two stimulus trains ("brain observatory" and "functional connectivity"). Also natural image presentations are shown in blue, as analyzed in Fig. S12. In all experimental sessions neural recordings were performed with Neuropixels probes.

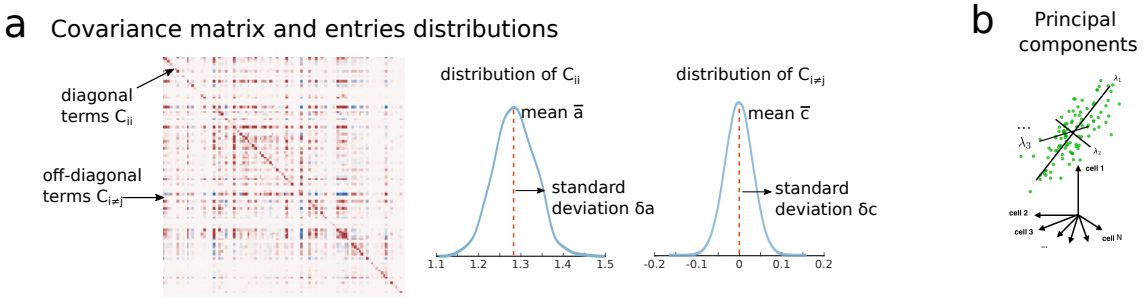

Fig. S2. Statistics of covariances in balanced networks. a) Example of covariance matrix and statistics for a balanced network. The covariance of the neural activity is computed by analyzing the binned neural activity. For each neuron spikes are binned into windows of 100ms. Then the covariance is computed resulting in a matrix where the number of rows and columns corresponds to the number of neurons. The two distributions on the right capture respectively the statistics of diagonal (auto-covariances) and off-diagonal (cross-covariances) entries of the covariance matrix. Of specific importance for our study are the standard deviation of cross-covariances δc and the mean of auto-covariances $\bar{a}$ as their ratio, termed s in our study, is in one-to-one correspondence with the spectral radius R. b) Example of data distribution where each green dot corresponds to a single binned vector of spike counts. The geometrical interpretation for such a point cloud is that the axes of major variability are determined by the eigenvectors of the covariance matrix. Eigenvalues, when normalized to their sum, determine the amount of variance of the data in the direction identified by the eigenvector.

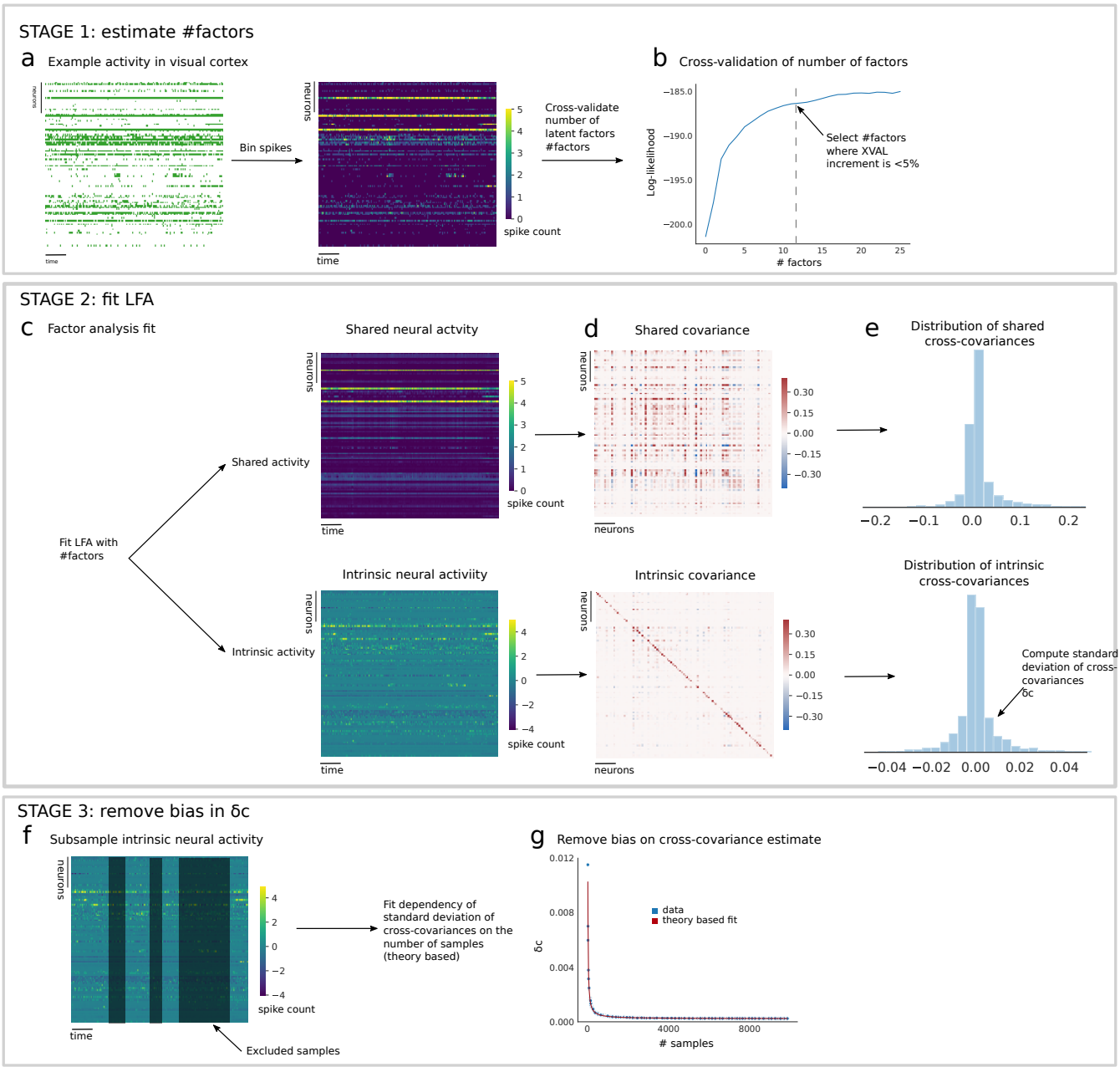

Fig. S3. Method for estimating shared and intrinsic activity statistics. a) Bin spikes of each recorded neuron in each experimental session into 100ms windows to obtain spike-count vectors. b) Perform 5 fold cross-validated latent factor analyses (LFA) with increasing number of factors (from 1 to 25). Determine log-likelihood as a function of the number of factors. Select number of factors (#factors) as the minimal one that yields a relative increment of the cross-validation curve lower than 5%. c) Perform LFA with #factors to split total activity into shared and intrinsic activity components. d) Compute covariance of shared and intrinsic components. e) Compute distribution of covariances for shared and intrinsic components. f) Subsample intrinsic neural activity and compute standard deviation of cross-covariances δc as a function of samples (#bins) used. g) Fit dependence of δc on the number of bins to estimate and remove bias induced by the finite duration of the recording (see Methods). This procedure yielded an unbiased estimate of the statistics of intrinsic auto- and cross-covariances.

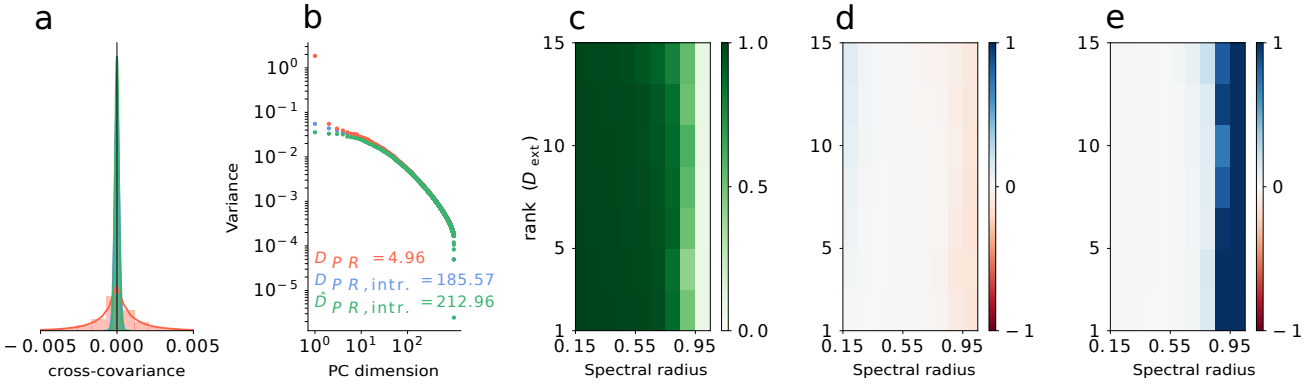

Fig. S4. Latent Factor Analysis applied to simulation data. LFA yields an upper bound on the dimensionality of intrinsic network activity in linear rate networks. Correlated external inputs can strongly alter covariances within the network, but LFA faithfully extracts intrinsic covariances over a wide range of spectral radii and ranks of external input covariances. a) A rank-one input, for example, can lead to broader distributions of covariances (red) with respect to intrinsically generated covariances (blue). Applying LFA, inferred intrinsic covariances (green) recover features of ground-truth intrinsic covariances well. b) Covariance spectra corresponding to distributions in a. c) Pearson correlation coefficient between ground-truth and inferred intrinsic covariances for different spectral radii and ranks of external input covariances. d) Relative error of inferred spectral radius. e) Relative error of inferred intrinsic dimensionality. For large spectral radii $R \lessapprox 1$, LFA wrongly subtracts low-dimensional components of intrinsic covariances, which, however, consistently leads to conservative results, i.e. underestimated spectral radii and overestimated intrinsic dimensionalities. Note that color scales in panels d and e are cut at the value 1 for better visibility. Details on the model system and the LFA application are shown in Suppl. Mat. S3.4.

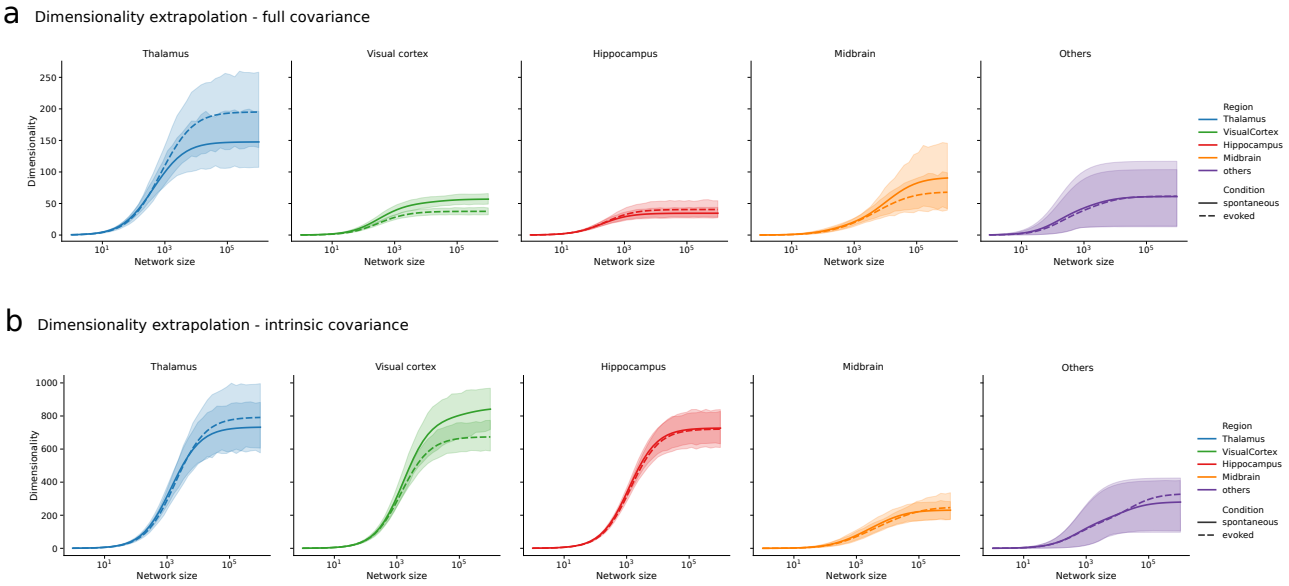

Fig. S5. Dimensionality extrapolation analysis across regions. a) Extrapolated dimensionality based on the full covariance for individual brain regions as a function of the network size, see Methods. For each region we computed the unbiased covariance matrix statistics (Fig. S3), and based on these statistics we computed the dimensionality as a function of the network size (the number of neurons in that particular region). This was performed for spontaneous (solid line) and evoked activity (dashed line). Shaded areas indicate confidence interval across sessions. Interestingly the estimates in most panels are not significantly different across conditions suggesting that a robust estimate of the dimensionality was achieved as the two conditions are based on independent statistics. The only significant difference between conditions is in visual cortex where in the evoked condition the dimensionality is significantly lower (t-test p-value$< 10^{-3}$) in line with previous findings (13, 40). b) Same as in panel a with the extrapolation here based on the statistics of intrinsic covariances. Comparing the extrapolated values of panel a and panel b allows to establish a respectively lower and upper bound on the dimensionality of neural activity in each area. Henceforth a bound on the value of recurrencies can be similarly derived due to the one-to-one relationship between recurrency and dimensionality based on intrinsic covariances.

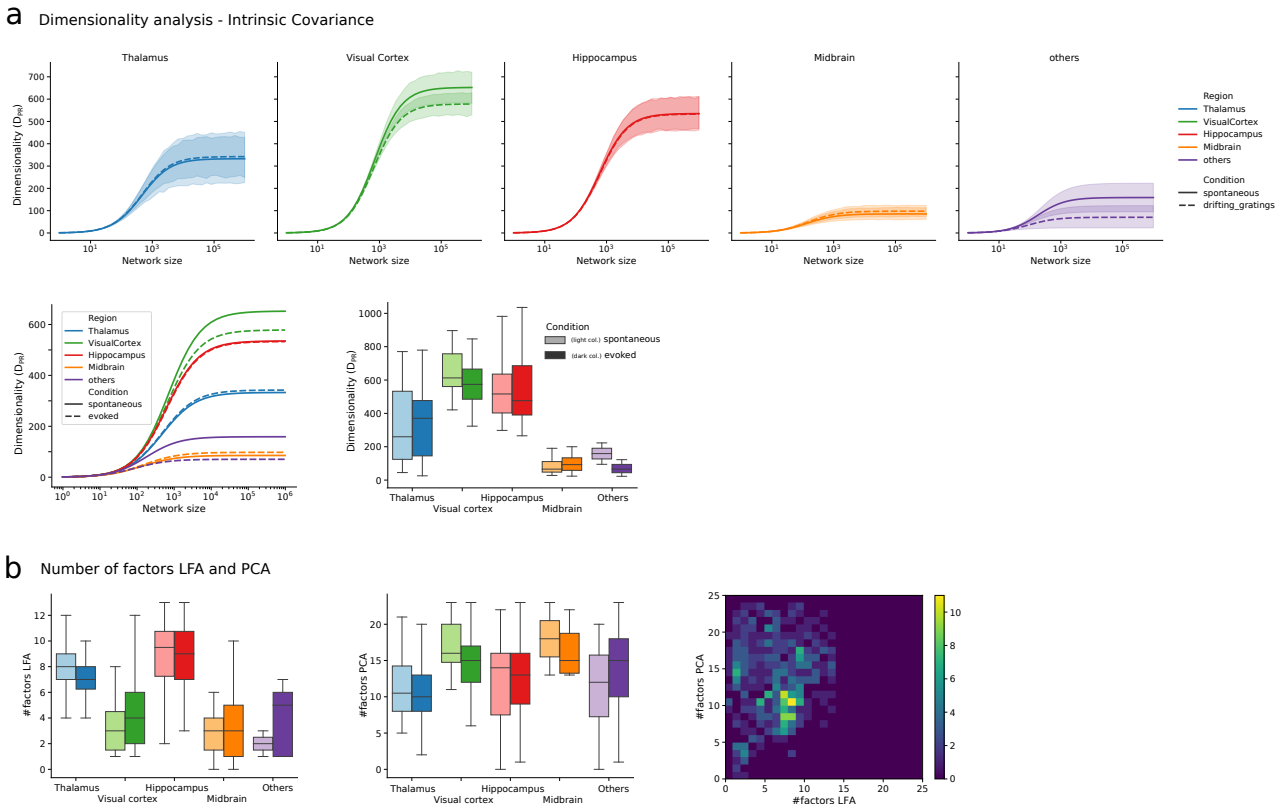

Fig. S6. Dimensionality analysis across regions with cross-validated Principal Component Analysis (PCA). In this figure we present the plots of Fig. 3 obtained by using a cross-validated PCA in place of a cross-validated LFA in the method used to estimate the shared and intrinsic activity components (Fig. S3). The method is applied identically, except that the LFA step is replaced by a cross-validated PCA. Similarly neural activity is then projected onto the number of principal components identified by the method to obtain the shared activity (projected onto components) and intrinsic activity (activity not captured by selected components). a) Top: dimensionality of intrinsic covariance extrapolation across brain regions and conditions. Same as plots shown in Fig. S5b. Bottom left: Average of the dimensionality extrapolation in each top panel to ease comparison across regions. Same as Fig. 3d left. Bottom right: dimensionality of intrinsic activity across regions and conditions visualized as a box plot (the box displays average and interquartiles). The statistic within the box corresponds to 50% of the statistics, therefore boxes are larger in size than more common standard deviations error bars. b) Left: number of factors extracted by the cross-validated LFA technique across regions (cf. Fig. S3). Middle: number of factors extracted by the cross-validated PCA technique across regions. Right: comparison between number of factors extracted by cross-validated LFA and cross-validated PCA. There is a significant correlation (p-value 0.037) between the number of factors extracted with the two methods across experimental sessions. This statistics is of relevance as it can be considered an estimate of the dimensionality of neural activity on its own. The number of cross-validated factors (for LFA) and components (for PCA) are indeed often taken as a measure of dimensionality. These plots suggest that according to these measures the dimensionality of neural activity across regions is on the order of 10-20 factors or components.

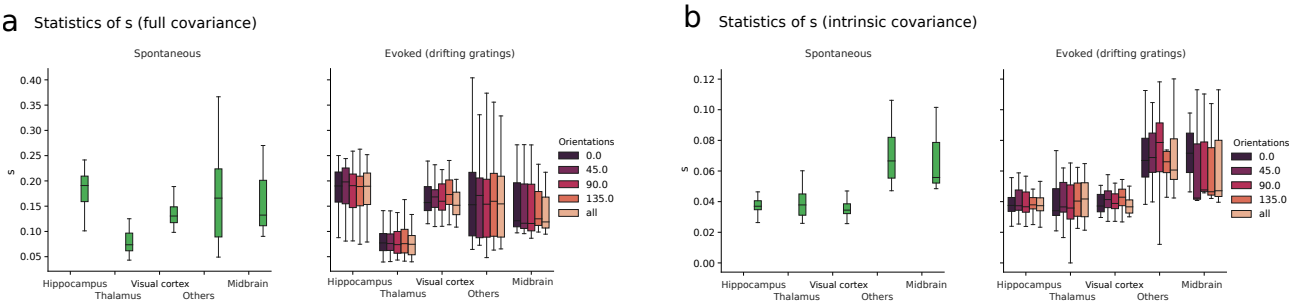

Fig. S7. Analysis of covariance statistics in the evoked condition. In order to test the robustness of our methods in extracting the statistics of cross-covariances we verified that such statistics were consistent by splitting the data into smaller sets. These sets were naturally selected by the fact that the evoked condition studied here corresponded to drifting gratings with multiple orientations. Therefore studying statistics across orientations allowed us to verify the robustness of our metrics. a) Quantity s (defined as the ratio between the standard deviation of cross-covariances and the average of autocovariances) for the full covariance of the activity. This is the central object for our analysis. Left: spontaneous condition across regions displayed for comparison to the evoked condition. Right: evoked condition across regions for different stimulus orientations and for all stimulus orientations analyzed together. Results in the main figures (Figs. 2 to 3) are computed, for each session, as the average across the four different orientations presented here. b) Same as panel a for the intrinsic covariance. These statistics allow us to verify the robustness of our analysis across orientations for the evoked condition and also verify that the 'all' condition, where orientations are considered all together, did not qualitatively differ from individual orientations.

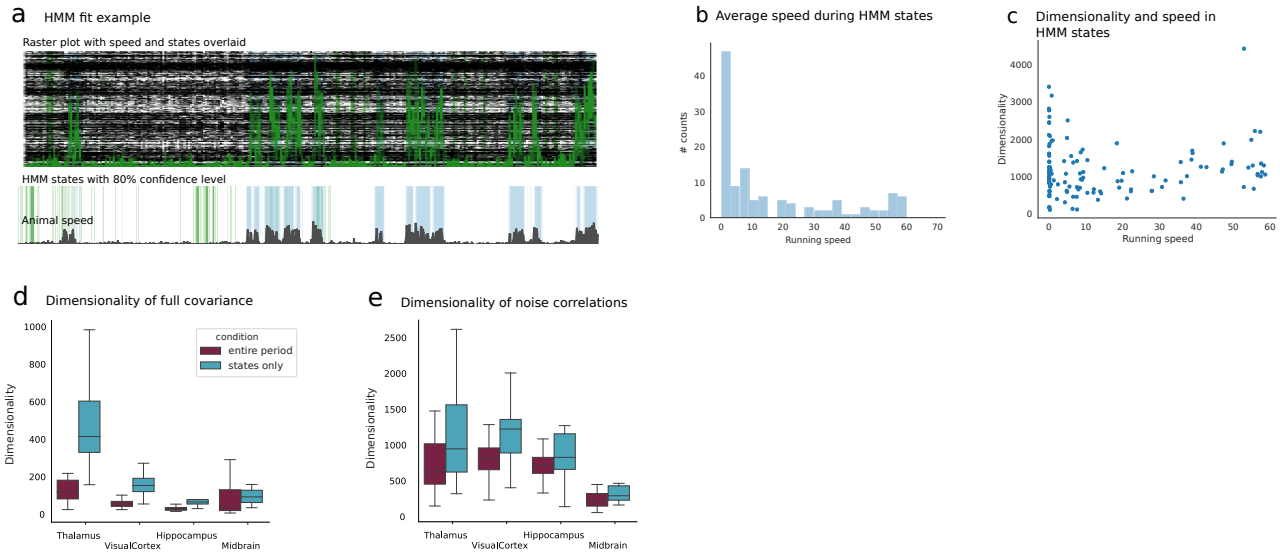

Fig. S8. Hidden Markov Model (HMM) analysis of spontaneous activity. Here we verify the robustness of the dimensionality and recurrency estimation in the spontaneous activity. During spontaneous activity neural activity is known to be strongly influenced by the movement of the animal and other factors (e.g. arousal). Therefore we verified that our estimates of dimensionality were not specifically driven by such factors. To achieve such control in an unsupervised way we analysed spontaneous activity by detecting, in the interval of 30 minutes corresponding to spontaneous activity, periods where the neural activity was well captured by the same latent state. This hypothesis, that the activity is driven by a specific latent state, corresponds to the hypothesis underlying a Hidden Markov Model (HMM). Henceforth we analyzed spontaneous activity with a HMM. This algorithm (see Methods) returned a parsing of spontaneous periods of activity into multiple underlying states by means of an unsupervised detection algorithm. In each state, the assumption underlying the algorithm is that the average neural activity is steady and therefore fluctuations around such average activity can be considered as intrinsically generated. a) Example raster plot during one of the sessions with overlaid running speed of the animal (green). The HMM extracts different states whose appearance strongly correlates with the animal movement (e.g. blue state). b) We noticed that some states were driven by the speed of the animal. This could be visualized by the average speed of the animal in individual HMM states across all sessions. Multiple states appear to correspond to either static or moving conditions. c) For the neural activity during all bins where a specific state appeared we computed the dimensionality of the neural activity for that specific state. The plot shows the dimensionality of so obtained HMM state dependent neural activity versus the average speed in each state. There is no substantial correlation between the running speed and dimensionality. d) Dimensionality of the full covariance across regions computed for the entire period of the spontaneous condition (entire period, purple) or for each individual state and averaged across all states and sessions in each brain region (states only, blue). The dimensionality computed over the entire period is systematically lower as the HMM extracts periods where the neural activity tends to be stationary thus limiting the effect of shared variability modes on the dimensionality of the full covariance. This suggests that the lower dimensionality obtained by the computation over the entire period is driven by the components that account for the switch between multiple states. In the computation we retained states which accounted for at least 2 minutes of the entire period. In areas which are elsewhere defined as "other" there were no states that passed this threshold, therefore data are not shown. e) Dimensionality of intrinsic covariances computed across regions for the entire period (purple) or during states only (blue). The dimensionality computed in the two ways appears more in agreement than in panel d suggesting that the LFA analysis (Fig. S3) successfully extracts components of shared variability coherently throughout the period in a way comparable to what is achieved by parsing the neural activity with an HMM. Therefore these analyses confirm the robustness of our estimation method validating the ability of the LFA analysis to capture sources of shared variability in the spontaneous condition.

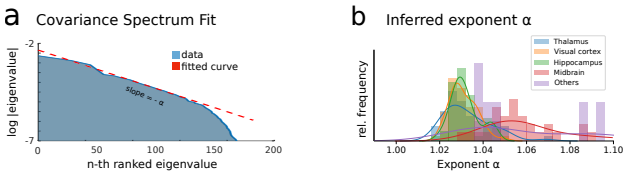

Fig. S9. Analysis of dimensionality via covariance eigenvalues spectrum. An alternative method of dimensionality estimation has been recently proposed in (9). We verified that our analysis is consistent with the results identified by this alternative technique. a) This technique is based on estimating the exponent α of a power law fit to the normalized full spectrum of the covariance eigenvalues. The figure shows an example estimation of the critical exponent α from the normalized spectrum of the full covariance matrix. b) The exponent α can also be extracted through the dimensionality (see Methods) by assuming the power law decay of the full spectrum. The distribution across all sessions for each brain area are here shown. The values obtained between 1.02 and 1.08 are consistent with the ones reported in (9). This allows to validate our techniques and their relation to established findings in the literature.

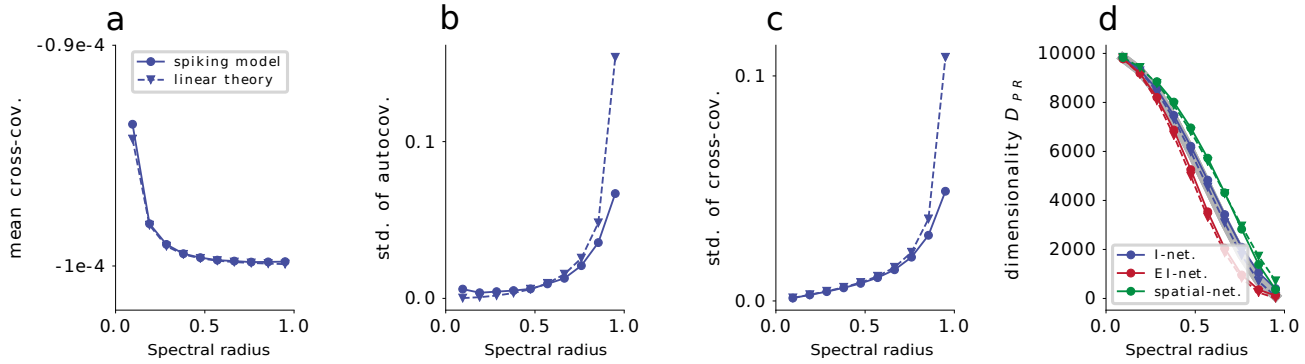

Fig. S10. Linear-response theory for dimensionality of spiking network dynamics. The theory for linear rate models predicts the dimensionality for more complex network topologies and spiking neuron dynamics. The participation ratio D_{PR} can be expressed in terms of three quantities (see Methods): a) mean cross-covariances, b) standard deviation of autocovariances, c) standard deviation of cross-covariances. All quantities are rescaled by mean autocovariances. Solid lines indicate simulations of homogeneous inhibitory networks with leaky integrate-and-fire neuron models and dashed lines show analytical predictions using linear response theory (18). d) The matching theoretical predictions for the statistics of covariances in homogeneous inhibitory networks yield an accurate prediction for the dimensionality (grey), not only for simulations of homogeneous single population inhibitory networks (blue), but also for homogeneous two-population excitatory-inhibitory networks (red) and spatially organized single-population inhibitory networks (green). The neuron and network models and parameters are as described in (18).

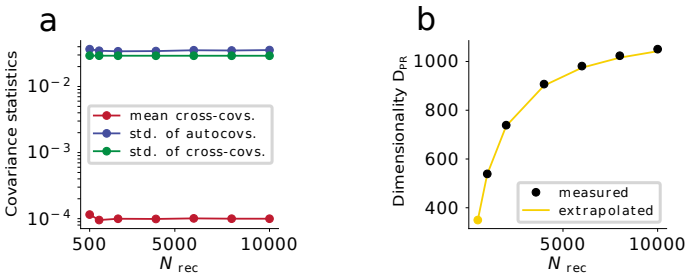

Fig. S11. Dimensionality extrapolation based on unbiased subsampling of covariances. a) Statistics of covariances based on activity from N_{rec} recorded neurons of a homogeneous inhibitory network of $N = 10000$ leaky integrate-and-fire neurons. The participation ratio D_{PR} can be expressed in terms of three quantities: mean cross-covariances (red) as well as standard deviations of auto- (blue) and cross-covariances (green), each rescaled by mean autocovariances (see Methods). The covariance statistics being independent of the number of recorded neurons N_{rec} allows for an unbiased dimensionality extrapolation based on subsampled covariances. b) Measured dimensionality (black dots) based on statistics of covariances from different numbers N_{rec} of recorded neurons as well as extrapolated dimensionality (yellow curve) based on covariance statistics of $N_{\mathrm{rec}} = 500$ neurons (corresponding to yellow dot). The extrapolated dimensionality fits the true dimensionality well. Statistics of covariances are averaged over 100 random subsamplings for each N_{rec} to avoid statistical fluctuations. Same homogeneous inhibitory network as in Fig. S10.

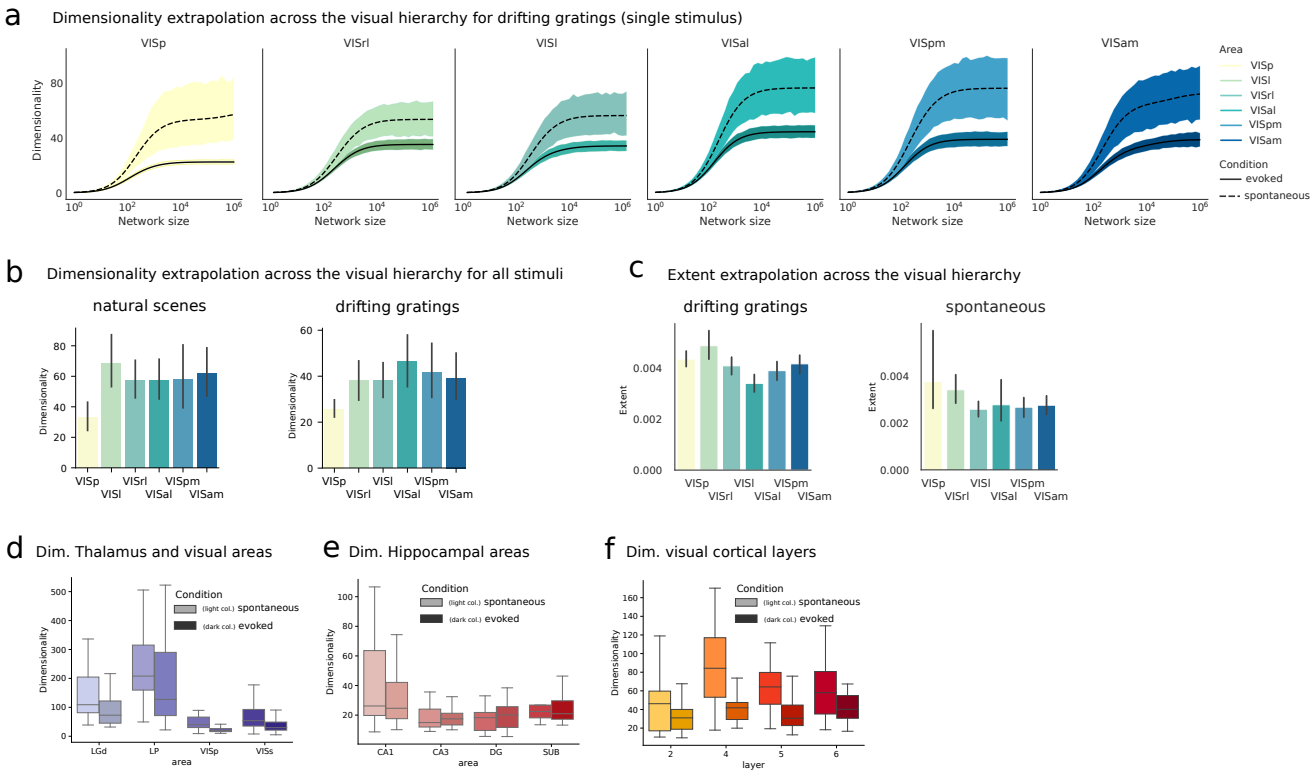

Fig. S12. Dimensionality analysis across brain areas. Here we provide supplementary plots to the analyses shown in Fig. 4. a) Dimensionality extrapolation as a function of number of neurons in the network (network size) based on the full covariance across areas involved in the visual hierarchy (36). In all areas the evoked condition is significantly lower dimensional than the spontaneous condition (max p-value across areas $< 10^{-15}$). b) Dimensionality of evoked responses for the full covariance across all stimuli in the drifting gratings (i.e. across all orientations as for condition "all" in Fig. S7) and natural scenes (cf. Methods and Fig. S1). Significant increase in dimensionality from primary to higher visual cortical areas was found in both conditions. c) The increase in dimensionality across the visual hierarchy has important connections with the processing of information in artificial deep neural networks classifying images (14, 19). To provide further data for such comparison we analyze the measure of extent of neural representations across the hierarchy given by $E(C) = \sum_i \lambda_i^2 / \left(\sum_i \lambda_i\right)$. This measure exploits the same statistics of the dimensionality Eq. (3) and can thus be extrapolated with identical techniques, but identifies the volume of the point cloud of data rather than its dimensionality. Here we plot the average extent across orientations for the evoked drifting gratings condition and the average extent in the spontaneous condition (same as Fig. 4b). The significant decrease in extent for neural representations, together with the increase in dimensionality across all conditions suggests an increased separability of neural representations along the visual hierarchy: point-clouds of data become smaller (lower extent) but are arranged in a higher dimensional fashion (higher dimensionality). d) Dimensionality of the full covariance across thalamic and visual areas. Secondary visual areas are here considered together (VISs): neurons from any secondary cortical area are grouped together in "area" VISs for analyses reporting VISs. e) Dimensionality based on the full covariance for hippocampal areas. f) Dimensionality based on the full covariances across visual cortical layers.

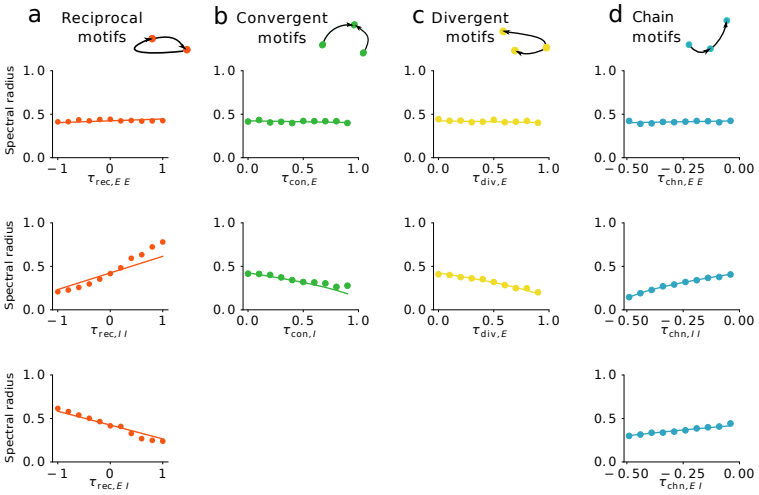

Fig. S13. Numerical validation of theory for spectral radius in excitatory-inhibitory networks with population-specific motif abundances. Spectral radius as a function of abundances of excitatory and inhibitory a) recurrent motifs, b) convergent motifs, c) divergent motifs, and d) chain motifs. Only one parameter is varied at a time with all other motif abundances set to zero. Theoretical predictions are given by solid curves, numerical evaluations of single network realizations are shown by colored markers. Cf. Suppl. Mat. for network construction algorithms. The approximate theory correctly predicts the influence of changes in abundance of each type of motif on the spectral radius. In particular, an over-representation of reciprocal EE and II motifs increases the spectral radius while an over-representation of reciprocal EI motifs decreases the spectral radius. These effects are the dominant contributions to the results obtained from the synaptic physiology dataset.

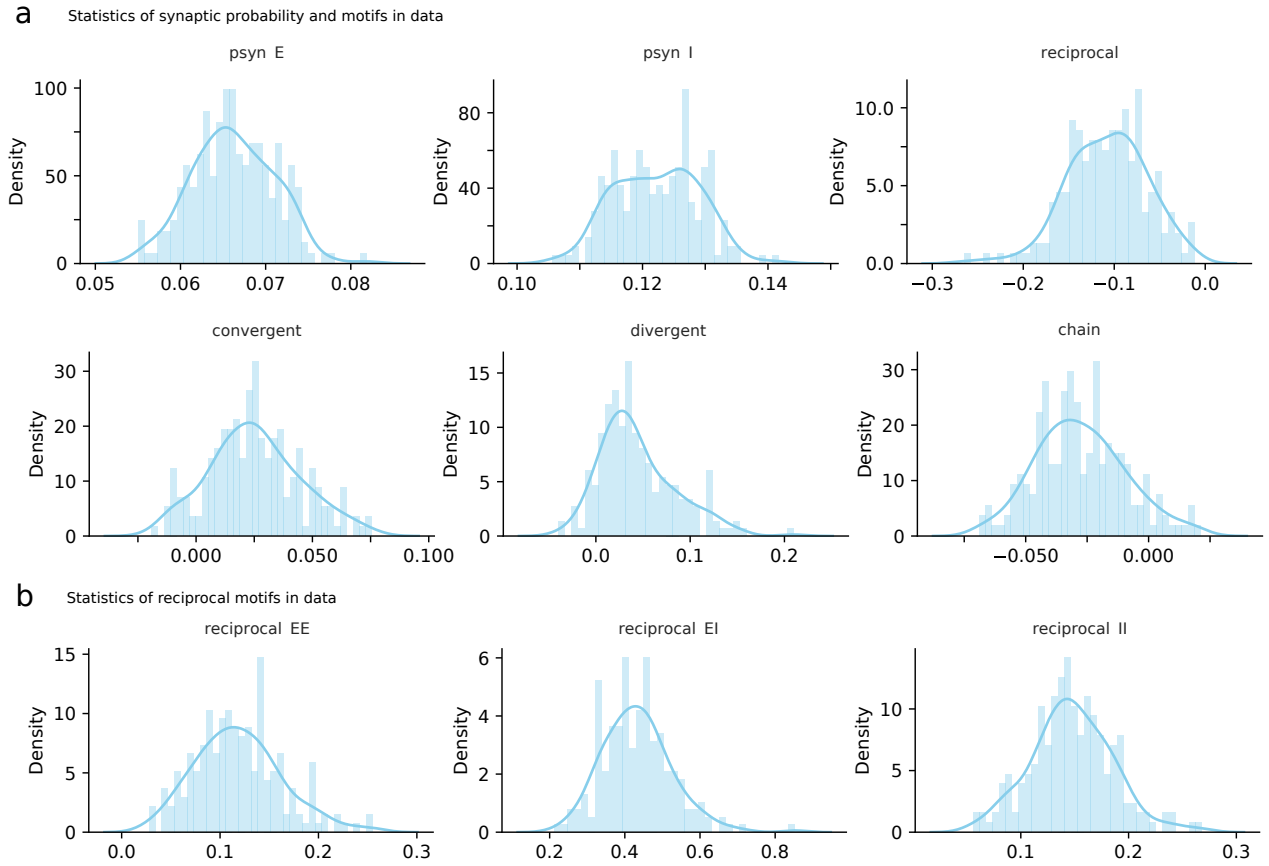

Fig. S14. Analysis of motifs in synaptic connectivity dataset. a) Statistics of synaptic probability (psyn) and τ values corresponding to each motif for the bootstrap analysis of mouse synaptic physiology data. The statistics reported in each panel corresponds to 500 bootstraps where in each bootstrap 80% random experimental sessions were analyzed. The plots allow us to verify the variability in the inference of synaptic probabilities and motif values. b) Same as in panel a for reciprocal motifs with synaptic type dependence (EE, EI, II corresponding respectively to the two synapses being excitatory, of mixed type or inhibitory). These supplementary plots utilize the same data as shown in Fig. 5f.

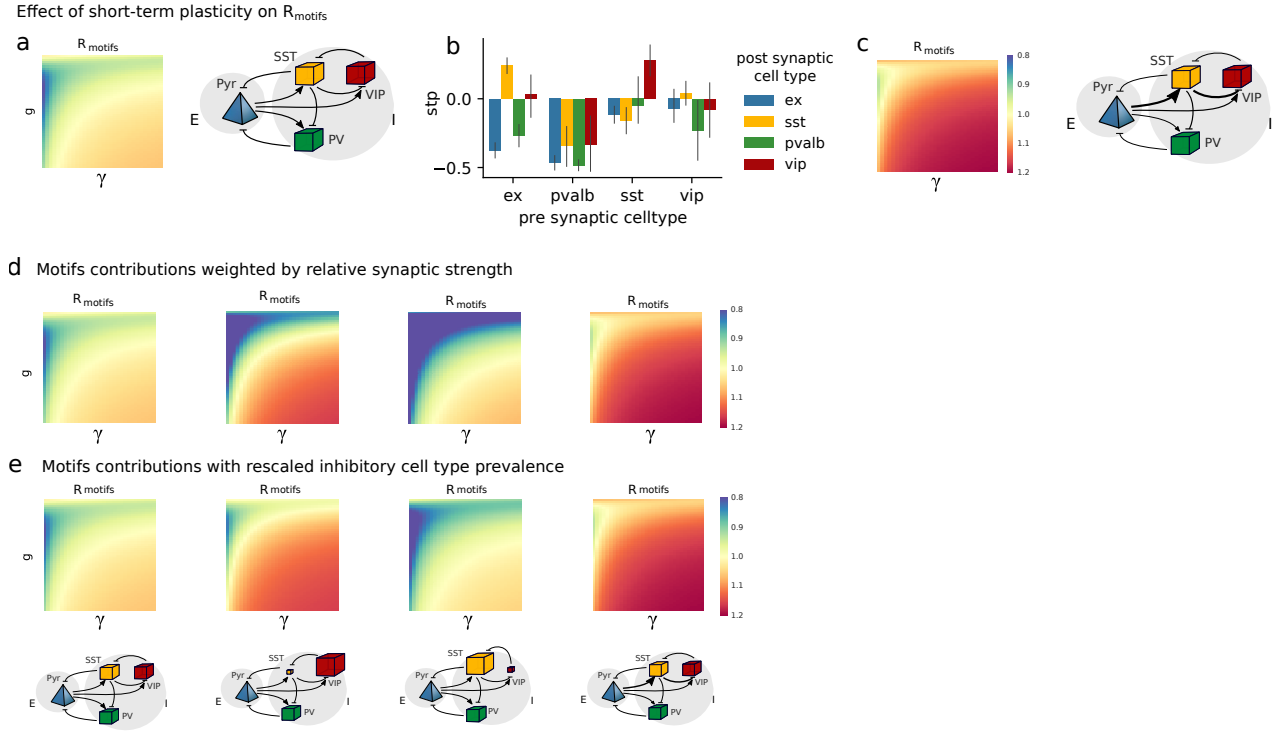

Fig. S15. Effect of short-term synaptic plasticity (STP) on recurrent coupling strength. We captured the effect of STP on the recurrency R_{motifs} with a dedicated analysis (see Methods). a) Values of R_{motifs} as a function of parameters g and γ. This plot is a reproduction of Fig. 5h for ease of comparison. b) Values of STP metric as a function of pre- and post-synaptic cell type as computed in the Allen synaptic physiology dataset ((37)). Values of STP above zero indicate a facilitating synapse (that increases its strength with stimulus repetitions) while STP values lower than zero indicate a depressing synapse (that decreases its strength with stimulus repetition). c) Values of R_{motifs} as a function of parameters g and γ upon considering the effective trend of STP modulation on individual motifs. This analysis points to the fact that R_{motifs} is modulated by STP effects, in particular, upon adaptation of the synapses STP appears to increase R_{motifs}. This would correspond to an increase of recurrency and a decrease in dimensionality. This is in agreement with findings suggesting that stimuli suppress the dimensionality of the representation as we found in visual cortex (cf. Fig. 4d, (13)). d) Control that accounting for the relative synaptic strength of synapses among different cell types did not affect our findings. We included in our analysis the average relative values of synaptic strength computed between populations of interneurons (see Methods). Each panel has a corresponding circuit cartoon at the bottom of the figure (from left to right: regular analysis same as Fig. S15a, analysis with silenced SST population same as Fig. 5f, analysis with silenced VIP population same as Fig. 5g, analysis with STP effect same as Fig. S15c). e) We verified that the statistics collected during the experiment would reflect the expected statistics relative to the difference prevalence of the three populations of interneurons in the visual cortex. We did so by accounting for such bias in our analysis and recomputing all values of R_{motifs}, see Methods. The order of the panels follows the same prescriptions as in panel d.

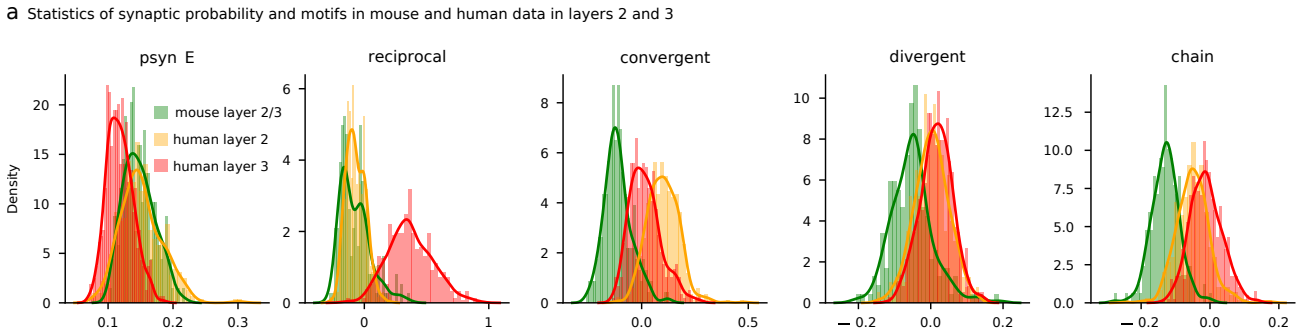

Fig. S16. Motifs in layers 2 and 3 of mouse and human cortex. In this analysis we compare the statistics of motifs in layer 2 and 3 for both mouse and human. These layers are the only ones that have a sufficient data in both datasets allowing for a direct comparison. a) Statistics of synaptic probability (psyn) and taus corresponding to each motif for the bootstrap analysis of mouse and human synaptic physiology data. The statistics reported in each panel corresponds to 500 bootstraps where in each bootstrap 80% random experimental sessions were analyzed. These supplementary plots utilize the same data as shown in Fig. 5i. It is remarkable how the distributions show similarities and differences in a motif specific way: In the case of chain and divergent motifs, the values for layer 2 and 3 for the human dataset are much in agreement, while for reciprocal connections they are markedly different.

Session id	Type	Thalamus	Vis.Cortex	Hippocampus	Midbrain	others	LGd	LP	VISp	VISl	VISam	VISpm	VISrl	VISal	VIS	CA1	CA3	DG	ProS	SUB	APN	lays1	lays2	lays4	lays5	lays6
715093703	brain observatory 1.1	297	258	185	26	117	82	139	60	42	30	50	76			145	14	26			26		42	51	41	14
719161530	brain observatory 1.1	193	166	178	176	19	71	28	52	40	37	18	10	9		108	14	14		42	176		15	23	57	19
721123822	brain observatory 1.1	72	214	122	15			69	41	27	39		46	37	24	71	10	41			10		28	62	50	9
732592105	brain observatory 1.1		365			459			110	40		66	64	85												
737581020	brain observatory 1.1		142			410			40	22		13	67													
739448407	brain observatory 1.1		278			347			19	36	56		70		97											
742951821	brain observatory 1.1		339			554			33	52		23	79	44	108											
743475441	brain observatory 1.1	207	185	60	34		42	131	45	19	53		68			20	5	12	1	22	34		38	33	58	11
744228101	brain observatory 1.1	185	275	165	20		1	84	35	26			26	41	147	112	9	10		34	20	1	15	20	49	8
746083955	brain observatory 1.1	65	213	12		292	34		14	17		115	35	32		7	5							13	17	2
750332458	brain observatory 1.1	4	295	38		558	4		63	38	60	19	44	71		16	15	7					4	7	16	11
750749662	brain observatory 1.1	325	287	149			82	142	52	20	64	64	41	46		113	10	26					30	51	111	43
751348571	brain observatory 1.1	331	279	177	40		55	137	49	27	71	37	95			78	16	37		46	37		41	67	93	29
754312389	brain observatory 1.1	28	275	197		1	27		102	14	64	22	41	32		135	15	47					26	52	71	24
754829445	brain observatory 1.1	429	192	158	45		90	170	92		47	21	32			123	12	23			45		12	23	40	25
755434585	brain observatory 1.1	118	361	121	17	25	44	27	75	39	94	62	49	42		45	8	15	53		14	4	55	33	122	23
756029989	brain observatory 1.1	168	318	168	27		60	27	51	30	72	90	24	51		113	10	37		8	21		62	46	117	42
757216464	brain observatory 1.1	305	352	215	55		6	114	85	53	64		37	53	60	112	13	42	7	41	47		31	63	94	19
757970808	brain observatory 1.1	184	304	218	48		45	58	80		70		46	54		130	4	37	47				26	46	77	21
758798717	brain observatory 1.1	115	199	181	68			39	47		49	23		80		116		40	1	24	53		33	22	68	29
759883607	brain observatory 1.1	94	205	137	38		23		58		55	48		44		82	3	15	2	35	30		11	25	91	20
760345702	brain observatory 1.1	64	274	148	5		37	17	72	49	55	44		54		91	30	27				2	24	36	106	34
760693773	brain observatory 1.1		406			420			88		70	54	66	128												
761418226	brain observatory 1.1	103	373	197	6	12	41	44	36			41	111	185		145	11	41			6		55	91	156	35
762120172	brain observatory 1.1		332			385			84	74	33			91	50											
762602078	brain observatory 1.1	32	303	138	25			27	75		52		60		116	101	5	6		26	19		22	27	49	14
763673393	brain observatory 1.1	87	278	193	50	6	77	3	73	36	74		44		51	131	32	30			48		18	35	78	23
766640955	functional connectivity	215	244	238	85		76	101	52	34	56		37	65		163	16	59			64		29	42	84	37
767871931	functional connectivity	119	342	195	49			83	101		52	64	33	46		100	19	32	1	43	41		22	41	100	32
768515987	functional connectivity	169	291	189	76		48	21	72	59	49		50	61		96	23	25	45		57	1	30	53	106	29
771160300	functional connectivity	49	450	369	1	8	6	43	85		87	58		108	21	269	42	29	26	3	1		59	42	110	42
771990200	functional connectivity	155	224	141	24		60	31	54		43	37	19	29		66	6	29	40		24		14	22	81	11
773418906	brain observatory 1.1	86	259	188	7				37		17		39	166		72	15	50	17	34	6		45	52	92	33
774875821	functional connectivity	31	279	264	27				69	53	45		40	72		190		58	16		27	1	33	40	118	18
778240327	functional connectivity	59	387	309	10		2	51	85	62	77	68	13	82		240	24	41	4		10	1	41	47	154	59
778998620	functional connectivity	113	364	224	57	12		77	75	74	58	45	51	61		130	40	14		40	55		30	64	134	61
779839471	functional connectivity	118	454	246	31		73	12	126		72	84	26		146	133	22	48	43		28		28	43	95	16
781842082	functional connectivity	128	355	165	58		11	2	83		58		42	172		112	18	27		2	51		25	68	132	47
786091066	functional connectivity	28	432	218	9				82		73	65	47	165		86	3	47	28	54	9		57	67	171	55
787025148	functional connectivity	76	272	295	21			38	68		66		11	54	8	185	42	18	17	33	20		19	33	58	21
789848216	functional connectivity	78	125	162	2	32		16	14	43	30		24	14		100	25	37			2		20	45	41	5
791319847	brain observatory 1.1	24	316	192		4	8	9	93	56	49	17	58	43		85	33	32	1	41			33	65	86	39
793224716	functional connectivity	134	313	246	61	15		101	80				39	171	23	151	27	27		41	56		26	65	78	41
794812542	functional connectivity	103	520	292	38			20	108	100	109	65	89	49		135	26	13	71	47	17		79	71	193	69
797828357	brain observatory 1.1	74	297	152	39	2		67	85	58	67	52	11	24		49	29	30		44	39		42	35	104	31
798911424	brain observatory 1.1	68	443	254	26			65	94	78	135		47	89		134	21	31	9	59	25		59	65	177	48
799864342	brain observatory 1.1	127	225	188	11		70	51	75	40	29		29	52		109	16	20	36	7	11		23	32	66	29
816200189	functional connectivity	100	297	183	17			55	61	65	51	76			44	120	8	20		35	13		48	36	72	36
819186360	functional connectivity	118	251	115	23		34	78		72	52	20	107			65	17	28		5	23		27	34	152	38
819701982	functional connectivity	128	242	182	15		31	44		66	78		98			125	6	27	15	9	9	1	30	73	99	39
821695405	functional connectivity	49	254	136		26		48	34	71	64	59		26		55	16	38	27				37	17	126	40
829720705	functional connectivity		371	143					52	78	77		75	89		79	18	46					50	53	174	42
831882777	functional connectivity		411	129	12	56			65	73	63	72	92	46		107		19	2		8		58	57	126	32
835479236	functional connectivity	29	318	160			21	8	79	76			87	76		103	9	33	15			3	34	81	80	41
839068429	functional connectivity	73	352	294	23		9	64	85	47	42	60	44		74	172	22	60	4	36	22		35	31	96	31
839557629	functional connectivity	36	219	193	1		36		53	48	38	38		42		107	11	48	27		1		21	26	97	22
840012044	functional connectivity	189	275	249	29			98	46	45	48	53	83			162	15	72			29		39	55	105	30
847657808	functional connectivity	55	346	389	66			53	50	76	65	55	45	55		184	5	64	97	39	63	1	45	68	143	39

Table S1. Statistics of Neuropixels recordings. The table shows the number of neurons recorded across sessions indicating the type of recording (stimulus set Fig. S1) and grouping the neurons into three categories: brain regions (light blue background), brain areas (dark blue background) and visual cortical layers (gray background). Sessions are sorted by the session id for reference in the dataset. In our analyses we utilized recorded activity of a specific brain area or region whenever the number of neurons recorded was 20 or higher.

a Motifs statistics EI mouse

Type	Count	Total
reciprocal EE	31	1119
reciprocal EI	179	2246
reciprocal II	131	4316
convergent EE	108	5810
convergent EI	81	11022
convergent II	270	21718
divergent EE	114	11321
divergent EI	0	0
divergent II	387	27229
chain EE	70	11620
chain EI	322	22044
chain II	360	43436

b Motifs statistics mouse

Type	Count	Total
reciprocal	351	8044
convergent	487	40455
divergent	535	40455
chain	789	80910

c Motifs statistics EI human

Type	Count	Total
reciprocal EE	28	530
reciprocal EI	25	851
reciprocal II	9	568
convergent EE	86	2908
convergent EI	28	4536
convergent II	17	2951
divergent EE	85	5176
divergent EI	0	0
divergent II	50	5219
chain EE	102	5816
chain EI	70	9072
chain II	31	5902

d Motifs statistics human

Type	Count	Total
reciprocal	62	2093
convergent	134	11136
divergent	140	11136
chain	209	22272

e Cell-type specific motifs statistics mouse

Type	Vip - Count	Vip - Total	Sst - Count	Sst - Total	Pv - Count	Pv - Total
reciprocal EI	1	361	16	534	28	408
reciprocal II	0	338	5	685	43	444
convergent EI	1	1245	7	1902	15	1345
convergent II	3	1159	30	2336	45	1411
divergent EI	0	0	0	0	0	0
divergent II	10	3852	50	5509	113	4075
chain EI	1	2486	32	3731	52	2626
chain II	0	2317	9	4595	93	2743

Table S2. Statistics of synaptic motifs. Here we report the statistics of detected (count) motifs per motif type versus the maximum number of possible motifs that could be detected if all synapses were present (total). The ratio of the two values (count/total) indicates the probability of occurrence of a specific motif. (a) Statistics of synaptic motifs in the mouse dataset per motif type. (b) Statistics of synaptic motifs in the mouse dataset independently of the cell type being excitatory or inhibitory. (c) Same as (a) for human synaptic physiology dataset. (d) Same as (b) for human synaptic physiology dataset. (e) Statistics of motifs in the mouse synaptic physiology dataset for each of the inhibitory populations of interneurons.