

The spatial structure of correlated neuronal variability

Robert Rosenbaum^{1,2}, Matthew A Smith^{3–5}, Adam Kohn^{6,7}, Jonathan E Rubin^{5,8} & Brent Doiron^{5,8}

Shared neural variability is ubiquitous in cortical populations. While this variability is presumed to arise from overlapping synaptic input, its precise relationship to local circuit architecture remains unclear. We combine computational models and *in vivo* recordings to study the relationship between the spatial structure of connectivity and correlated variability in neural circuits. Extending the theory of networks with balanced excitation and inhibition, we find that spatially localized lateral projections promote weakly correlated spiking, but broader lateral projections produce a distinctive spatial correlation structure: nearby neuron pairs are positively correlated, pairs at intermediate distances are negatively correlated and distant pairs are weakly correlated. This non-monotonic dependence of correlation on distance is revealed in a new analysis of recordings from superficial layers of macaque primary visual cortex. Our findings show that incorporating distance-dependent connectivity improves the extent to which balanced network theory can explain correlated neural variability.

The spiking activity of cortical neurons is often characterized by their average response over a large number of trials, prompting a wealth of theoretical studies relating the structure of neuronal networks to their trial-averaged firing rate dynamics¹. However, trial averages do not capture the stochastic and irregular dynamics characteristic of cortical populations and the nervous system in general². Indeed, trial-to-trial fluctuations are central to contemporary theories of cortical computation^{3,4}. A deep mechanistic understanding of neuronal variability remains an open challenge.

Early theoretical studies deduced that variable spiking activity could arise through a balancing of strong, yet opposing, excitatory and inhibitory synaptic inputs^{5,6}. Expanding on this conjecture, van Vreeswijk and Sompolinsky⁷ showed that networks of recurrently coupled model neurons robustly create a state where strong excitation is approximately balanced by inhibition, creating a push-pull dynamic that generates irregular spiking activity. More recently, balanced networks have been implicated in theories of optimal coding⁸, working memory⁹ and stimulus tuning¹⁰. Numerous experimental studies have established that excitation is often approximately balanced by inhibition in cortical circuits^{11–17}. In sum, balanced networks provide a parsimonious model of the irregular spiking activity observed in cortical circuits.

Early balanced network models produced asynchronous activity through sparse connectivity^{7,18}. However, several experimental studies reveal that local cortical networks are densely connected, with connection probabilities between nearby neurons sometimes exceeding 40 percent^{19–22}. These data imply substantial overlap between local synaptic inputs, which could, in principle, synchronize cortical networks. However, counter to intuition, balanced networks with dense connectivity show weak spike train correlations²³. This ‘asynchronous

state’ results from the correlated excitatory (e) or inhibitory (i) afferents to neuron pairs being actively cancelled by a strong negative e–i correlation, establishing weak correlations even when connectivity is not sparse²³.

Consistent with the predicted asynchronous state, some multiunit extracellular recordings show noise correlations that are nearly zero on average²⁴. However, a majority of population recordings in cortex reveal comparatively large correlations^{25,26}. Several studies suggest that the magnitude of noise correlations is dependent on many factors²⁷, including arousal²⁸, attention²⁹, anesthetic state^{23,24,30,31} and cortical layer^{32,33}. Finally, while *in vivo* whole-cell recordings reveal strong positive e–e and i–i correlations coexisting with strong e–i correlations¹³, these correlation sources do not always perfectly cancel as predicted by some theoretical models²⁸. Taken together, these studies show that cortical circuits can exhibit both weak and moderate noise correlations, at odds with predictions from the current theory of balanced networks²³.

In this study, we generalize the theory of correlations in densely connected, balanced networks to include the widely observed dependence of synaptic connection probability on distance^{21,34}. We show that spatially broad recurrent projections disrupt the asynchronous state, producing a signature spatial correlation structure: nearby pairs of neurons are positively correlated on average, pairs at intermediate distances are negatively correlated and distant pairs are weakly correlated. These positive and negative correlations cancel so that the average correlation between pairs sampled randomly over a large range of distances is nearly zero. We uncover this non-monotonic dependence of correlation on distance in recordings from superficial layers of macaque primary visual cortex, but only after correcting for a latent source of shared fluctuations. Our findings decouple balanced excitation and inhibition from asynchronous network

¹Department of Applied and Computational Mathematics and Statistics, University of Notre Dame, Notre Dame, Indiana, USA. ²Interdisciplinary Center for Network Science and Applications, University of Notre Dame, Notre Dame, Indiana, USA. ³Department of Ophthalmology, University of Pittsburgh, Pittsburgh, Pennsylvania, USA. ⁴Fox Center for Vision Restoration, University of Pittsburgh, Pittsburgh, Pennsylvania, USA. ⁵Center for the Neural Basis of Cognition, Pittsburgh, Pennsylvania, USA. ⁶Department of Ophthalmology and Visual Sciences, Albert Einstein College of Medicine, Yeshiva University, Bronx, New York, USA. ⁷Dominick Purpura Department of Neuroscience, Albert Einstein College of Medicine, Yeshiva University, Bronx, New York, USA. ⁸Department of Mathematics, University of Pittsburgh, Pittsburgh, Pennsylvania, USA. Correspondence should be addressed to R.R. (robert.rosenbaum@nd.edu) or B.D. (bdoiron@pitt.edu).

Received 28 March; accepted 28 September; published online 31 October 2016; addendum published after print 25 May 2017; doi:10.1038/nn.4433

activity, greatly extending the applicability of balanced network theory to explaining cortical dynamics.

RESULTS

We consider a network of excitatory and inhibitory exponential integrate-and-fire model neurons. Neurons provide recurrent, lateral synaptic input to one another and receive feedforward synaptic input from a nonlocal presynaptic population. A detailed mathematical analysis of correlations in the limit of large network size is provided in **Supplementary Note 1** and **Supplementary Figures 1–3**. Below we provide an outline of these theoretical results and confirm their predictions using computer simulations. We first use a simplified network model to demonstrate how the asynchronous state considered in previous theoretical work²³ is broken by heterogeneous input correlations³⁵. We then consider a more realistic model where neurons belong to a continuous spatial domain and connection probability depends on distance.

Homogeneous input correlations are cancelled by balanced networks

To demonstrate the mechanisms affecting correlations in recurrent networks, we first simulated a simplified network of $N = 20,000$ neurons, half excitatory and half inhibitory, that all receive the same fluctuating feedforward input and are each connected with probability 0.25 (**Fig. 1a**). Despite the fact that neuron pairs share all of their feedforward input and 25% of their recurrent synaptic input on average, spiking activity was asynchronous, with an average pairwise spike count correlation of 6.8×10^{-5} (**Fig. 1b,c**).

This small average correlation is a defining characteristic of the asynchronous state. Mathematically, this state is realized when spike count covariances in the network satisfy²³

$$C_{SS} \sim O(1/N)$$

where C_{SS} denotes the average spike count covariance between pairs of neurons in the recurrent network, N is the number of neurons in the network and $\sim O(1/N)$ denotes asymptotic proportionality to $1/N$ for large N . Note that covariance and correlation scale identically with network size in balanced networks, so we discuss them interchangeably²³.

Our theoretical analysis proceeds by noting that spike count covariance is inherited from synaptic input covariance²⁶ and therefore the two scale similarly with N in the asynchronous state,

$$C_{II} \sim O(1/N)$$

where C_{II} denotes the average covariance between neurons' synaptic inputs.

Synaptic inputs can be decomposed into their feedforward and recurrent sources, $I = F + R$, so that neurons' input covariances decompose as

$$C_{II} = C_{FF} + C_{RR} + 2C_{RF}$$

where C_{FF} is the average covariance between neurons' feedforward input currents, C_{RR} between their recurrent inputs and C_{RF} between one neuron's recurrent and the other neuron's feedforward synaptic input. Recurrent synaptic input, R , is composed of positive contributions from lateral excitatory synaptic inputs and negative contributions from inhibitory ($R = e - i$).

Shared input fluctuations are visualized by averaging the inputs to several neurons, so that the unshared contributions average out (**Fig. 1d**). Overlapping inputs cause C_{FF} and C_{RR} to be positive (**Fig. 1d**). If feedforward input correlation is moderate, $C_{FF} \sim O(1)$, then recurrent input tracks the feedforward input so that C_{RF} is negative and nearly perfectly cancels the positive sources of correlations (i.e., $2C_{RF} = -(C_{FF} + C_{RR}) + O(1/N)$; see **Supplementary Note 1**). As a result, the covariance between the total synaptic inputs is weak, $C_{II} \sim O(1/N)$ (**Fig. 1d**). This cancellation arises naturally in

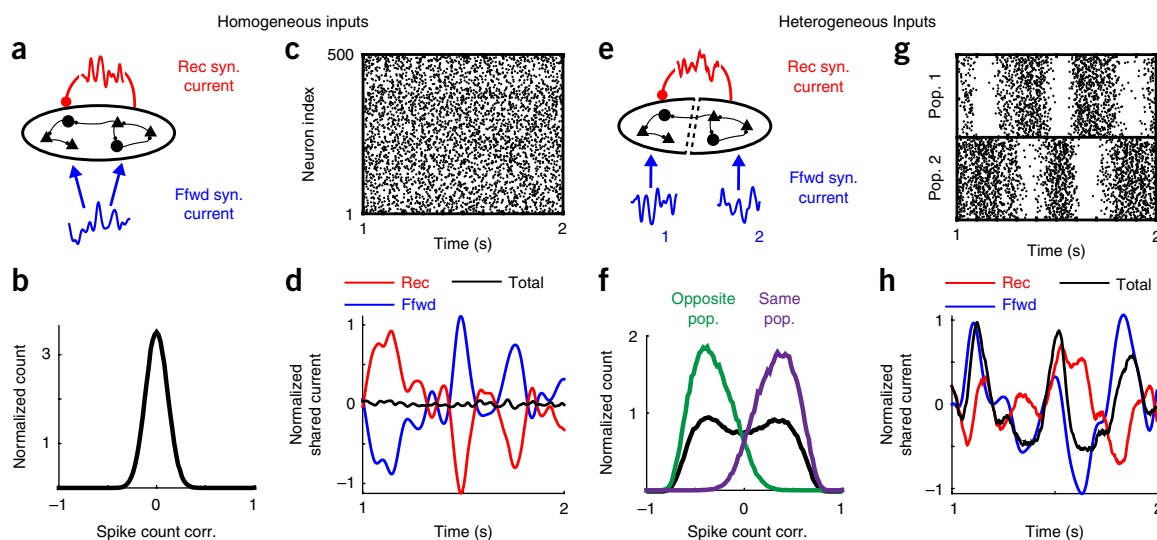


Figure 1 Heterogeneous feedforward input breaks asynchrony in balanced recurrent networks. **(a)** Homogeneous network diagram. A population of 20,000 recurrently connected excitatory and inhibitory neurons receives globally correlated feedforward input. **(b)** Normalized histogram of pairwise spike count correlations between 1,000 randomly selected neurons. All histograms are normalized by their integral. **(c)** Raster plot of 500 randomly chosen neurons plotted over 1 s. **(d)** Shared fluctuations in the feedforward (blue) and recurrent (red) synaptic inputs cancel so that shared fluctuations in the total synaptic currents (black) are weak. Curves were computed by averaging the synaptic input currents to 500 neurons, convolving with a Gaussian-shaped kernel ($\sigma = 15$ ms), subtracting the mean and dividing by the neurons' rheobase. **(e–h)** Same as **a–d** except neurons were separated into two populations with separate feedforward inputs. Currents in **h** are from neurons in population 2. Histograms in **f** show correlations from neuron pairs randomly selected from both populations (black), from the same population (purple) and from opposite populations (green). Rec, recurrent; ffwd, feedforward; syn, synaptic; corr., correlation; pop., population.

the balanced state and does not require a precise tuning of model parameters²³. Since spiking correlations are inherited from synaptic input covariance²⁶, this cancellation of input covariances leads to small, $O(1/N)$, spike count correlations.

Heterogeneous input correlations can disrupt the asynchronous state

To study the impact of heterogeneity on correlations in balanced networks, we modified the above model by dividing the neurons into two populations. Each population received a separate feedforward input (Fig. 1e). The two feedforward input sources were statistically identical but uncorrelated. Recurrent connectivity was not changed: neurons were randomly connected without respect to population membership (identically to Fig. 1a). This input heterogeneity dramatically changed the structure of correlations in the network. Pairs of neurons in the same population had strongly positive spike count correlations on average (0.34), while neuron pairs from opposite populations were negatively correlated with a nearly identical correlation magnitude (-0.34), and the average correlation between all pairs was nearly zero (4.2×10^{-4} ; Fig. 1f,g).

The mechanism responsible for this change in correlations can be understood by again separating the synaptic input covariance into its recurrent and feedforward sources, but generalizing the decomposition to account for neuron 'distance' to obtain

$$C_{II}(d) = C_{FF}(d) + C_{RR}(d) + 2C_{RF}(d) \quad (1)$$

Here $C_{II}(d)$ is the average covariance between input currents to pairs of neurons separated by distance d , where $d = 0$ for neurons in the same population and $d = 1$ for opposite population pairs, and similarly for the other terms. Feedforward input is only correlated between neurons in the same population, so $C_{FF}(0) > 0$, but $C_{FF}(1) = 0$.

In contrast, recurrent connections do not respect population membership and thus neither do the statistics of recurrent input,

$$C_{RR}(0) = C_{RR}(1) \text{ and } C_{RF}(0) = C_{RF}(1)$$

Since covariances $C_{RR}(d)$ and $C_{RF}(d)$ do not depend on d but $C_{FF}(d)$ does, cancellation cannot be achieved in equation (1) for both $d = 1$ and $d = 0$ simultaneously. In other words, the one 'copy' of shared recurrent synaptic input cannot cancel both versions of the feedforward synaptic inputs. The loss of cancellation causes the total synaptic current shared by neurons in the same population to inherit shared fluctuations from their feedforward inputs (Fig. 1h), giving rise to positive $O(1)$ correlations between same-population pairs. A competitive dynamic introduces negative correlations between neurons in opposite populations (Fig. 1f,g). A similar mechanism was considered in a recent theoretical study³⁵.

For illustrative purposes, we considered a simplified network model with discrete subpopulations, but correlations and connectivity in many cortical circuits depend on continuous quantities such as physical distance or tuning similarity^{19,21,36}. Next, we generalize these findings to more biologically realistic networks with connection probabilities that depend on neuron distance.

A spatially extended network model

We next considered a network of $N_e = 40,000$ excitatory and $N_i = 10,000$ inhibitory model neurons arranged on a square-shaped domain modeling a portion of a cortical layer. The neurons receive feedforward synaptic input from a separate layer of Poisson-spiking

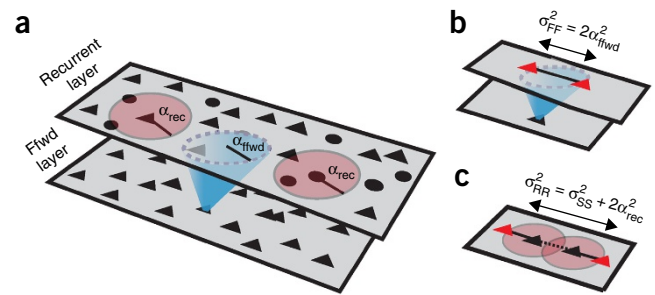


Figure 2 Correlation and projection widths in spatially extended networks. (a) Network schematic. Black triangles and circles represent excitatory and inhibitory neurons. Red disks indicate recurrent synaptic projections. Recurrent connection probability decays with distance with width parameter α_{rec} . Blue cone denotes feedforward (ffwd) synaptic projections from a separate layer, with width parameter α_{ffwd} . (b) Correlations introduced by overlapping feedforward input to neurons in the recurrent layer (shared blue input to red triangles) decay with distance twice as slowly as feedforward connection probability ($\sigma_{\text{FF}}^2 = 2\alpha_{\text{ffwd}}^2$). (c) The spatial width of correlations between two neurons' recurrent inputs (σ_{RR}^2 ; input from black triangles to red triangles) is equal to width of spike train correlations (σ_{SS}^2 ; dashed line) plus twice the width of recurrent projections ($2\alpha_{\text{rec}}^2$; solid lines).

neurons and are connected with a probability that decays with distance (Fig. 2). Specifically, the probability of a connection between two neurons in the recurrent network obeys

$$\text{Pr}(\text{connection}) \propto g(d; \alpha_{\text{rec}}^2)$$

where d is the distance between the neurons measured along the two-dimensional network, $g(d; \sigma^2) \propto \exp(-d^2/(2\sigma^2))$ is a Gaussian-shaped function and α_{rec} approximately represents the average length of a recurrent synaptic projection. Similarly, the probability of a synaptic projection from a neuron in the feedforward layer to a neuron in the recurrent layer decays with distance similarly to a Gaussian with width parameter α_{ffwd} , where distance is measured parallel to the cortical surface (Fig. 2). We next show that the asynchronous state is realized when $\alpha_{\text{rec}} < \alpha_{\text{ffwd}}$, then show that the asynchronous state cannot be realized when $\alpha_{\text{rec}} > \alpha_{\text{ffwd}}$.

The asynchronous state in spatially extended cortical circuits

As above, asynchronous spiking requires a cancellation between input covariances (cf. equation (1)), except that d now represents continuous instead of binary distance. Therefore, conditions on asynchrony require first an understanding of how input covariances depend on pairwise neuron distance.

Overlapping feedforward synaptic projections introduce $O(1)$ correlations between the feedforward inputs to neuron pairs. Since nearby pairs share more feedforward inputs, these correlations are distance dependent. Specifically, synaptic divergence causes feedforward input correlations to be $O(1)$ and twice as broad as synaptic projection widths (Fig. 2b),

$$C_{\text{FF}}(d) \propto g(d; \alpha_{\text{ffwd}}^2).$$

The fact that $C_{\text{FF}}(d) \sim O(1)$ at first seems to preclude the possibility of an asynchronous state because $C_{\text{FF}}(d)$ is one component of $C_{II}(d)$ in equation (1) and the asynchronous state requires $C_{II}(d) \sim O(1/N)$. However, the asynchronous state is realized under a cancellation between positive (C_{FF} and C_{RR}) and negative (C_{RF}) sources of correlations in equation (1). Cancellation at all distances requires

that all correlation sources have the same shape (Supplementary Note 1), meaning that

$$C_{RF}(d), C_{RR}(d) \propto g(d; \alpha_{ffwd}^2) \quad (2)$$

The implications of this requirement on the spatial profile of spiking correlations are clarified by noting that recurrent synaptic input is generated by spike trains in the recurrent network. Synaptic divergence causes the correlations between neurons' recurrent synaptic inputs to be broader in space than the correlations between spike trains according to (Fig. 2c):

$$\sigma_{RR}^2 = \sigma_{SS}^2 + \alpha_{rec}^2$$

Here σ_{RR} is the width of correlations between neurons' recurrent synaptic input currents and σ_{SS} is the width of spike train correlations in the recurrent network. In general, we use α to denote the widths of synaptic projections and σ to denote the widths of correlations.

Correlations between recurrent inputs are constrained by the cancellation required in the asynchronous state. Specifically, equation (2) requires that the width of correlations between recurrent synaptic inputs satisfy

$$\sigma_{RR}^2 = 2\alpha_{ffwd}^2$$

Combining the two expressions for σ_{RR} above yields

$$\sigma_{SS}^2 = 2(\alpha_{ffwd}^2 - \alpha_{rec}^2) \quad (3)$$

The existence of a real solution to equation (3) requires that $\alpha_{ffwd} > \alpha_{rec}$; in other words, the spatial width of the recurrent projections must be narrower than the width of feedforward projections for the asynchronous state to exist. Further, equation (3) implies that $\sigma_{SS}^2 > \sigma_{FF}^2 - \alpha_{rec}^2$, so that spike train correlations are spatially narrower than correlations between feedforward input currents. Thus, recurrent dynamics actively sharpen the spatial profile of correlations in the asynchronous state (compare to the sharpening of tuning curves in previous work³⁷).

To test these theoretical findings, we performed network simulations with feedforward synaptic projections broader than recurrent projections (Fig. 3a). The simulations confirmed that $C_{RR}(d)$ and $C_{RF}(d)$ decayed similarly with distance to $C_{FF}(d)$ (Fig. 3b). This allowed a cancellation between positive and negative sources of correlations, so that correlations between neurons' total synaptic currents and between their spike trains were weak over all distances (Fig. 3b–e). Despite their small average, spike count correlations had a larger standard deviation (s.d. = 0.11; Fig. 3d), consistent with results for nonspatial networks²³ (Fig. 1b). Neurons in the network receive strong excitation that is canceled by strong inhibition on average (Fig. 3f and Supplementary Fig. 4a,b), confirming that the network maintains a balanced state. Correlations computed from simulations agreed with closed-form mathematical predictions (Fig. 3e; see Supplementary Note 1 for equations). Additional simulations confirmed that mean correlations decay toward zero at increasing network size (Supplementary Note 2 and Supplementary Fig. 4c,d).

Broad lateral connections produce a signature spatial correlation structure

As noted above, the cancellation between positive and negative correlations necessary for the asynchronous state cannot be realized when recurrent projections are broader than feedforward ($\alpha_{rec} > \alpha_{ffwd}$)

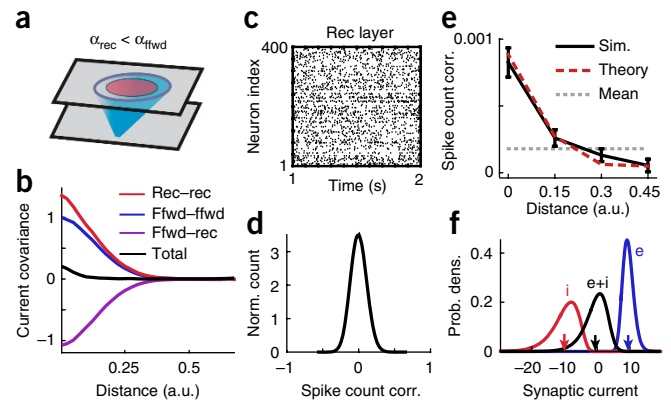


Figure 3 The asynchronous state in a spatially extended network model. (a) Network schematic. As in Figure 2 with recurrent projections narrower than feedforward projections ($\alpha_{rec} = 0.5\alpha_{ffwd}$). (b) Average covariance between different sources of synaptic currents to excitatory neuron pairs as a function of distance. Positive covariance between neurons' feedforward input currents (blue) and between their recurrent input currents (red) cancel with negative covariance between one neuron's feedforward and the other neuron's recurrent input (purple) to produce weak covariance between their total input (black). Curves were computed from input currents to 400 randomly selected excitatory neurons and were normalized by the peak feedforward input covariance. (c) Spike rasters of the 400 excitatory neurons comprising the center 20×20 square of neurons in the recurrent layer. (d) Normalized histogram of pairwise spike count correlation between 5,000 randomly selected neurons. (e) Mean spike count correlation between neurons (\pm s.e.m.) as a function of their distance. Solid black curve computed from the correlations between 5,000 randomly sampled neurons. Dashed red curve is from mathematical calculations (see Supplementary Note 1). Gray dashed line shows mean across all sampled pairs. (f) Distribution of excitatory (blue), inhibitory (red) and total (black) synaptic currents across the membranes of 400 randomly selected excitatory neurons, measured in units of the neurons' rheobase. Arrows indicate mean values. Rec, recurrent; ffwd, feedforward; corr., correlation; sim., simulation; prob. dens., probability density; e, excitatory; i, inhibitory.

because equation (3) cannot be solved in this case. Instead, neuron pairs inherit correlations from overlapping feedforward inputs so that $C_{SS}(d) \sim O(1)$.

We confirmed this prediction by numerical simulations identical to those discussed above, but with recurrent projections broader than feedforward (Fig. 4a). As predicted, recurrent input correlations were too spatially broad to cancel with the more sharply decaying feedforward correlations (Fig. 4b), so that the total input correlation between nearby neurons was large (Fig. 4b; compare to Fig. 3b). This effect introduced moderately strong correlations between nearby spike trains (Fig. 4c–e) that did not decay to zero at increasing network size (Supplementary Note 2 and Supplementary Fig. 4e,f). Nevertheless, the network maintained excitatory–inhibitory balance (Fig. 4f and Supplementary Fig. 4a,b).

Since recurrent inputs must cancel feedforward inputs in balanced networks, $C_{RF}(d)$ is negative (Figs. 1, 3 and 4 and Supplementary Note 1). Moreover, broad recurrent projections cause $C_{RF}(d)$ to decay slowly with distance (Fig. 4b). Through equation (1), this imparts a non-monotonicity in the dependence of $C_{II}(d)$ on d (Fig. 4b), and spike count correlations inherit this non-monotonic shape (Fig. 4e).

Following the same argument made for the homogeneous network, the spike count correlations averaged over neuron pairs at all distances is $O(1/N)$ (Fig. 4e and Supplementary Note 1). However, as noted above, the average correlation over each distance cannot be $O(1/N)$. Hence, there must be a cancellation between positive and negative

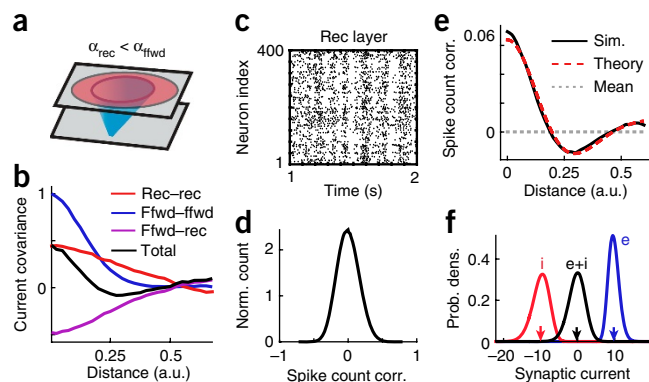


Figure 4 Broad recurrent projections lead to a correlated balanced state. (a–f) Same as **Figure 3a–e** except recurrent projections were changed to be broader than feedforward projections ($\alpha_{\text{rec}} = 2.5\alpha_{\text{fwd}}$). This change prevents the recurrent network from canceling positive feedforward input correlations (b), resulting in population-wide spike count correlations with increased s.d. (c,d) and with positive correlations between nearby neurons but negative correlations between more distant neurons (e). Nevertheless, the network maintains balance (f).

correlations at different distances. As in **Figure 1e–h**, a competitive dynamic causes nearby neurons to be positively correlated and more distant neurons to be negatively correlated. This competitive dynamic does not extend beyond the reach of recurrent projections, so sufficiently distant neurons are weakly correlated. Hence, correlation decreases and then increases with distance. This non-monotonicity can be explained more precisely using a mathematical theory of correlation transfer (**Fig. 4e** and **Supplementary Note 1**). The heterogeneity of positive and negative correlations at different distances increases the standard deviation of pairwise correlations, but only modestly (s.d. = 0.16, **Fig. 4d**; compare to **Fig. 3d**).

In summary, when recurrent projections are spatially narrower than feedforward projections ($\alpha_{\text{rec}} < \alpha_{\text{fwd}}$, as in **Fig. 3**), correlations are weak between pairs of neurons at all distances. When recurrent projections are broader than feedforward ($\alpha_{\text{rec}} > \alpha_{\text{fwd}}$, as in **Fig. 4**), nearby neurons are positively correlated, neurons at moderate distances are negatively correlated and distant neurons are weakly correlated. Moreover, the average correlation between pairs of neurons sampled randomly at all distances is small. The non-monotonic dependence of correlation on distance is a distinct signature of correlations arising from broad recurrent projections. We next investigated whether this correlation structure predicted by our theory is present in cortical recordings.

Spatial correlation structure in a visual cortical circuit

We next asked whether our theoretical characterization of correlations in spatially extended networks can explain correlations in a cortical circuit. Layers 2/3 and layer 4C of macaque primary visual cortex (L2/3 and L4C) provide an ideal circuit for testing our predictions. Pairs of neurons in L2/3 exhibit moderately large noise correlations that decay with distance, but neurons in L4C, which are a primary source of interlaminar input to L2/3, exhibit extremely weak pairwise noise correlations^{32,33} (**Fig. 5**, with data from previous studies^{33,36}).

Neurons in L4C receive much of their feedforward input from thalamic projections, which form spatially broad synaptic fields, around 1 mm in diameter, but lateral projections within macaque L4C form narrower, sub-millimeter synaptic fields³⁴. Thus, our theoretical prediction that correlations are weak when $\alpha_{\text{fwd}} > \alpha_{\text{rec}}$ is consistent with the weak pairwise correlations observed between L4C neurons *in vivo* (as in **Fig. 3**).

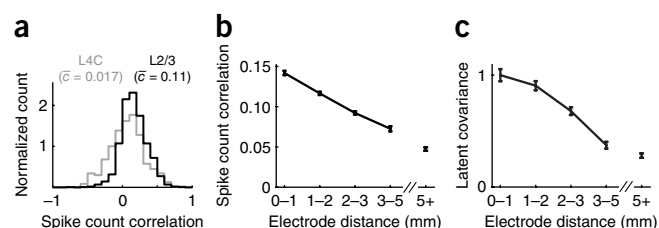


Figure 5 Dependence of correlations on layer and distance in macaque V1. (a) Histogram of pairwise correlations between neurons in superficial (putative L2/3, black) and middle (putative L4C, gray) layers of macaque primary visual cortex. Legends give average correlation. (b) Average pairwise correlation between putative L2/3 neurons (\pm s.e.m.) as a function of the distance between the electrodes on which the neurons were recorded. (c) Average latent covariance between putative L2/3 neurons (\pm s.e.m.). Data points in c were normalized by the peak at 0–1 mm.

Interlaminar projections from L4C to L2/3 have a similar sub-millimeter width to excitatory intralaminar projections within L4C, and lateral projections from inhibitory basket cells in L2/3 form sub-millimeter synaptic fields similar to those in L4C³⁴. Excitatory neurons in L2/3, however, form long-range lateral synaptic projections with synaptic fields spanning several millimeters³⁴. Our theoretical results can be generalized to this setting, where inhibitory and excitatory projections have different spatial profiles (**Supplementary Note 1**). This extension predicts the same correlation structure reported in **Figure 4**. However, correlations measured in L2/3 are positive on average over a broad range of distances³⁶ (**Fig. 5b**), in disagreement with this prediction.

We hypothesized that this inconsistency could be explained by recent studies showing that much of the correlated variability measured in L2/3 arises from a low-dimensional shared source of latent variability^{30,31,38–40}. We conjectured that this shared variability increases pairwise correlations in L2/3 at all distances, thereby ‘washing out’ the negative correlations predicted by our theory. To search for low-dimensional variability in our data, we used Gaussian process factor analysis^{31,41} (GPFA), a statistical algorithm that extracts shared fluctuations from a population of spike trains (see Online Methods). Applying this algorithm to our L2/3 recordings revealed a source of one-dimensional covariability that decays with distance (**Fig. 5c**). This distance dependence implies that nearby neurons are affected similarly by the latent variable.

To test whether one-dimensional latent variability explains the discrepancy between our theoretical predictions and data, we built a two-layer network model representing a 10 mm by 10 mm square of cortex (**Fig. 6a**). The first layer, representing L4C, was similar to the model in **Figure 3**, with the profile of feedforward and recurrent projections chosen to match experimentally constrained thalamic and lateral projection widths³⁴. The second layer, representing L2/3, was similar to **Figure 4**, with feedforward synaptic input from excitatory neurons in the L4C model and recurrent projection widths also chosen to match anatomical measurements. To capture latent variability in L2/3, the feedforward synaptic input to each neuron in the second layer was modulated by a time-varying, multiplicative gain modulation. We chose a multiplicative source of variability to be consistent with the properties of low-dimensional variability previously reported in macaque V1 (ref. 31), but an additive source of latent variability would produce similar overall results. The gain modulation contributes an $O(\sqrt{N})$ source of covariance to the feedforward inputs that the recurrent network cannot cancel²³. To capture the distance dependence of latent variability (**Fig. 5c**), the magnitude of the gain modulation was heterogeneous across the network in such a way that nearby neurons received similar modulations and more distant neurons received less similar modulations.

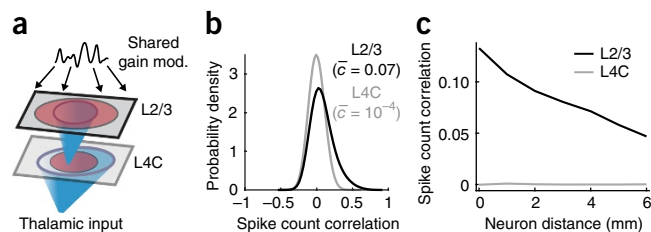


Figure 6 Dependence of correlations on layer and distance in a spatially extended, multilayer network model. (a) Network schematic. Thalamic input to L4C is broader than recurrent projections within L4C. Projections from L4C to L2/3 are narrower than recurrent excitatory (but not inhibitory) projections within L2/3. Neurons in L2/3 also receive a shared gain modulation (mod.). (b) Histograms of pairwise correlations between randomly selected neurons in each layer. (c) Average pairwise correlation between neurons in each layer as a function of the distance between the neurons.

Simulations of this two-layer model revealed that correlations between neurons in L4C were extremely small on average (Fig. 6b,c), consistent with our theoretical predictions (Fig. 3) and consistent with *in vivo* recordings (Fig. 5). Correlations in the model L2/3 layer were moderately large and positive over all distances (Fig. 6b,c), comparable to those in *in vivo* recordings (Fig. 5).

Thus, our model recovers the coarse structure of correlations in L4C and L2/3. However, our explanation of positive correlations in L2/3 is unsatisfying because the addition of globally shared variability destroys the distinct non-monotonic relationship between correlation and distance predicted by our theory (compare Fig. 4e to Fig. 6b). We next asked whether this structure could be recovered by filtering out globally shared variability. To accomplish this, we computed the residual correlation matrix estimated by GPFA. Residual correlations approximate the spike count correlations with the contribution from low-dimensional variability removed³¹.

Residual correlations computed between the simulated L2/3 spike trains exhibited the predicted non-monotonic dependence on distance, corroborating the ability of the GPFA algorithm to extract low-dimensional variability and leave the structure of residual correlations intact. We next computed the mean residual correlation in macaque L2/3 as a function of electrode distance. In doing so, we observed the same non-monotonic dependence of residual correlation on distance predicted by our theory (Fig. 7b; further statistical analysis in Supplementary Note 3 and Supplementary Fig. 5).

In summary, combining theoretical analysis and computer simulations of a multilayer network reveals a parsimonious model of the sources of shared variability in a visual cortical circuit *in vivo*. Under this model, positive correlations introduced by shared thalamic inputs to L4C neurons are actively canceled by negative correlations arising from recurrent circuitry so that pairs of L4C neurons at all distances exhibit weak average spike count correlations^{32,33}. Correlations between neurons in L2/3 are introduced by overlapping feedforward inputs from L4C and a low-dimensional source of variability. Correlations arising from overlapping L4C projections are filtered by recurrent circuitry in L2/3 to promote a non-monotonic dependence of correlation on distance. This non-monotonic correlation structure is washed out by low-dimensional latent variability, but can be recovered using GPFA to estimate and remove this variability.

DISCUSSION

Previous theoretical work on spatially homogeneous balanced networks with dense connectivity shows that they produce very

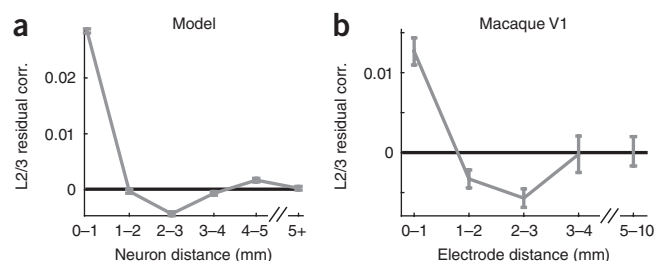


Figure 7 Residual correlations in macaque V1 and in a model. (a) Residual correlation (corr.) between neurons within the model L2/3 network as a function of distance. (b) Residual correlations between putative L2/3 neurons in macaque primary cortex (same data as Fig. 5b). Residual correlation approximates spike count correlations after a single source of shared latent variability is removed. Both plots show mean \pm s.e.m. Correlations decreased in the first two bins ($P < 10^{-12}$; unpaired *t*-test), increased from the third to fourth bin ($P = 0.019$) and from the third to the fifth bin ($P = 0.0067$).

weak spike train correlations²³. We have generalized this theory to account for heterogeneous inputs and distance-dependent connection probability. In this framework we have made two notable discoveries.

First, in agreement with the original findings, when lateral synaptic projections are spatially narrower than incoming feedforward projections, correlations are extremely weak on average at all distances. This theoretical finding can explain the weak pairwise correlations observed between neurons in middle layers of macaque primary visual cortex^{32,33}. However, correlations measured in cortical recordings are not always weak²⁵. Second, networks with broader lateral than feedforward projections produce correlations that do not decay to zero at increasing network size.

In previous studies of balanced networks with spatially homogeneous or clustered connectivity^{23,42}, the asynchrony condition $C_{SS} \sim O(1/N)$ is satisfied and population averaged pairwise correlations vanish in the large network limit. In contrast, spatially extended networks with broad lateral projections violate the asynchrony condition, and consequently the expected pairwise correlations at a specific distance do not vanish. Nonetheless, mean excitatory and inhibitory currents balance and firing rates are moderate even when the asynchrony condition is violated (Supplementary Fig. 4a,b and Supplementary Note 1). This represents a novel solution for balanced networks that, for the first time, formally decouples network-wide asynchrony from excitatory–inhibitory balance.

We focused on the dependence of correlations on distance, but correlations also depend on tuning similarity. Partitioning L2/3 neuron pairs by tuning similarity reveals that correlations are strongest between similarly tuned neurons³⁶ (Supplementary Fig. 6a,b). Modifying our computational model to capture tuning-dependent correlations produced a non-monotonic dependence of residual correlation on tuning similarity in some parameter regimes, but the relevant parameters have not been measured experimentally (Supplementary Fig. 6c–e and Supplementary Note 4). Nevertheless, the modified theory could explain negative correlations previously observed in computer simulations of networks with tuning-specific connectivity³² and the finding that negative correlations are more frequent between disparately tuned neurons in V1 (ref. 43).

As with nearly any computational model, many of the parameters used in our simulations may not reflect their corresponding values in specific cortical areas of specific species. However, our theoretical analysis does not depend on the precise values of these

parameters. Our finding that the asynchronous state requires $\alpha_{\text{rec}} < \alpha_{\text{fwd}}$ is a fundamental property of networks with balanced excitation and inhibition.

We used a simplified model of a visual cortical circuit. In reality, pyramidal neurons in V1 form both local and long-range projections, connection probability in primate V1 depends on both distance and tuning similarity, and these dimensions are coupled⁴⁴. Moreover, connectivity properties of inhibitory neurons depend on their subtype⁴⁵. We modeled unidirectional connections from L4 to L2/3, but L4 also receives indirect feedback from L2/3 through deeper cortical layers. Spike trains in our model feedforward layer were modeled by homogeneous Poisson processes, in contrast to the oscillatory firing rates evoked by drifting grating stimuli in the data we analyzed. Our model can be extended to account for these additional features without affecting our overall conclusions.

Our findings have important implications for the interpretation of correlations in neural recordings. The average (residual) correlation between cell pairs sampled across a large range of distances could be extremely small, even when nearby pairs are positively correlated with moderate magnitude (Figs. 4 and 7). Hence, subtracting low-dimensional latent variability and partitioning neuron pairs by distance can reveal correlation structure that would otherwise not be apparent. A previous study³¹ computed residual correlations as a function of distance in primate V1, but did not report a non-monotonic dependence. While we cannot be certain why their findings differ from ours, the accurate estimation of residual correlations with GPFA depends on the amount of data used to estimate shared variability. Our data are well-suited for this purpose, as they contain over 800 pairs of units per recording on average.

There is a long history of computational models of cortical circuits that consider either networks with spatially dependent coupling¹ or balanced excitation and inhibition in spatially homogeneous networks^{7,23}. Only recently has the spatial structure of cortical connectivity been included in networks with balanced excitation and inhibition^{9,37,46}, and guiding theoretical principles are lacking. Our theory has taken this spatial structure into account and produced two core predictions for cortical circuits with long range lateral connections: first, nearby neurons exhibit significant positive correlations; second, the dependence of pairwise correlation on pairwise distance is non-monotonic. These predictions are clearly falsifiable and hence represent strong tests of our theory. The superficial layers of visual cortex have long-range lateral connections³⁴, making them a suitable test bed for our theory of correlations. After accounting for a source of global variability, both of our predictions were verified from population recordings in macaque V1 (Fig. 7). Further, a similar noise correlation structure has been reported in recordings from mouse V1 (ref. 47). The successful validation of our predictions marks our theory as a promising framework for studying the structure of neural variability in cortical circuits. Nevertheless, there are many aspects of cortical dynamics that remain unexplained by balanced networks, such as infrequent yet large membrane fluctuations during spontaneous dynamics^{15,48}. Capturing these dynamics in cortical models with balanced architectures remains an open challenge.

METHODS

Methods, including statements of data availability and any associated accession codes and references, are available in the [online version of the paper](#).

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We are grateful to T.S. Lee and A. Movshon for research support. This work was supported by National Science Foundation grants NSF-DMS-1517828 (R.R.), NSF-DMS-1313225 (B.D.), NSF-DMS-1517082 (B.D.), NSF-DMS-1612913 (J.E.R.), NSF-DMS-1516288 (J.E.R.) and NSF-DMS-1312508 (J.E.R.); National Institute of Health grants R01NS070865 (B.D., J.E.R.), CRCNS-R01DC015139 (B.D.), R01EY016774 (A.K.), R01EY022928 (M.A.S.) and P30EY008098 (M.A.S.); two grants from the Simons Foundation collaboration on the global brain (SCGB#325293MC;BD, B.D. and 364994AK, A.K.); by the Eye and Ear Foundation of Pittsburgh (M.A.S.); and by Research to Prevent Blindness (A.K., M.A.S.).

AUTHOR CONTRIBUTIONS

R.R. and B.D. conceived the project; R.R. performed the simulations, data analysis and mathematical calculations; M.A.S. and A.K. provided the experimental data. J.E.R. and B.D. supervised the project. All authors contributed to writing the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>.

- Ermentrout, B. Neural networks as spatio-temporal pattern-forming systems. *Rep. Prog. Phys.* **61**, 353 (1998).
- Faisal, A.A., Selen, L.P. & Wolpert, D.M. Noise in the nervous system. *Nat. Rev. Neurosci.* **9**, 292–303 (2008).
- Pouget, A., Beck, J.M., Ma, W.J. & Latham, P.E. Probabilistic brains: knowns and unknowns. *Nat. Neurosci.* **16**, 1170–1178 (2013).
- Shamir, M. Emerging principles of population coding: in search for the neural code. *Curr. Opin. Neurobiol.* **25**, 140–148 (2014).
- Softky, W.R. & Koch, C. The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. *J. Neurosci.* **13**, 334–350 (1993).
- Shadlen, M.N. & Newsome, W.T. Noise, neural codes and cortical organization. *Curr. Opin. Neurobiol.* **4**, 569–579 (1994).
- van Vreeswijk, C. & Sompolinsky, H. Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science* **274**, 1724–1726 (1996).
- Denève, S. & Machens, C.K. Efficient codes and balanced networks. *Nat. Neurosci.* **19**, 375–382 (2016).
- Lim, S. & Goldman, M.S. Balanced cortical microcircuitry for spatial working memory based on corrective feedback control. *J. Neurosci.* **34**, 6790–6806 (2014).
- Hansel, D. & van Vreeswijk, C. The mechanism of orientation selectivity in primary visual cortex without a functional map. *J. Neurosci.* **32**, 4049–4064 (2012).
- Shu, Y., Hasenstaub, A. & McCormick, D.A. Turning on and off recurrent balanced cortical activity. *Nature* **423**, 288–293 (2003).
- Haider, B., Duque, A., Hasenstaub, A.R. & McCormick, D.A. Neocortical network activity in vivo is generated through a dynamic balance of excitation and inhibition. *J. Neurosci.* **26**, 4535–4545 (2006).
- Okun, M. & Lampl, I. Instantaneous correlation of excitation and inhibition during ongoing and sensory-evoked activities. *Nat. Neurosci.* **11**, 535–537 (2008).
- Dorn, A.L., Yuan, K., Barker, A.J., Schreiner, C.E. & Froemke, R.C. Developmental sensory experience balances cortical excitation and inhibition. *Nature* **465**, 932–936 (2010).
- Graupner, M. & Reyes, A.D. Synaptic input correlations leading to membrane potential decorrelation of spontaneous activity in cortex. *J. Neurosci.* **33**, 15075–15085 (2013).
- Zhou, M. *et al.* Scaling down of balanced excitation and inhibition by active behavioral states in auditory cortex. *Nat. Neurosci.* **17**, 841–850 (2014).
- Xue, M., Atallah, B.V. & Scanziani, M. Equalizing excitation-inhibition ratios across visual cortical neurons. *Nature* **511**, 596–600 (2014).
- Amit, D.J. & Brunel, N. Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. *Cereb. Cortex* **7**, 237–252 (1997).
- Ko, H. *et al.* Functional specificity of local synaptic connections in neocortical networks. *Nature* **473**, 87–91 (2011).
- Fino, E. & Yuste, R. Dense inhibitory connectivity in neocortex. *Neuron* **69**, 1188–1203 (2011).
- Levy, R.B. & Reyes, A.D. Spatial profile of excitatory and inhibitory synaptic connectivity in mouse primary auditory cortex. *J. Neurosci.* **32**, 5609–5619 (2012).
- Oswald, A.M., Doiron, B., Rinzel, J. & Reyes, A.D. Spatial profile and differential recruitment of GABA_B modulate oscillatory activity in auditory cortex. *J. Neurosci.* **29**, 10321–10334 (2009).
- Renart, A. *et al.* The asynchronous state in cortical circuits. *Science* **327**, 587–590 (2010).
- Ecker, A.S. *et al.* Decorrelated neuronal firing in cortical microcircuits. *Science* **327**, 584–587 (2010).
- Cohen, M.R. & Kohn, A. Measuring and interpreting neuronal correlations. *Nat. Neurosci.* **14**, 811–819 (2011).

26. Doiron, B., Litwin-Kumar, R., Rosenbaum, R., Ocker, G.K. & Josić, K. The mechanics of state-dependent neural correlations. *Nat. Neurosci.* **19**, 383–393 (2016).
27. Kohn, A., Zandvakili, A. & Smith, M.A. Correlations and brain states: from electrophysiology to functional imaging. *Curr. Opin. Neurobiol.* **19**, 434–438 (2009).
28. Poulet, J.F. & Petersen, C.C. Internal brain state regulates membrane potential synchrony in barrel cortex of behaving mice. *Nature* **454**, 881–885 (2008).
29. Cohen, M.R. & Maunsell, J.H. Attention improves performance primarily by reducing interneuronal correlations. *Nat. Neurosci.* **12**, 1594–1600 (2009).
30. Mochol, G., Hermoso-Mendizabal, A., Sakata, S., Harris, K.D. & de la Rocha, J. Stochastic transitions into silence cause noise correlations in cortical circuits. *Proc. Natl. Acad. Sci. USA* **112**, 3529–3534 (2015).
31. Ecker, A.S. *et al.* State dependence of noise correlations in macaque primary visual cortex. *Neuron* **82**, 235–248 (2014).
32. Hansen, B.J., Chelaru, M.I. & Dragoi, V. Correlated variability in laminar cortical circuits. *Neuron* **76**, 590–602 (2012).
33. Smith, M.A., Jia, X., Zandvakili, A. & Kohn, A. Laminar dependence of neuronal correlations in visual cortex. *J. Neurophysiol.* **109**, 940–947 (2013).
34. Lund, J.S., Angelucci, A. & Bressloff, P.C. Anatomical substrates for functional columns in macaque monkey primary visual cortex. *Cereb. Cortex* **13**, 15–24 (2003).
35. Wimmer, K. *et al.* The dynamics of sensory integration in a hierarchical network explains choice probabilities in MT. *Nat. Commun.* **6**, 6177 (2015).
36. Smith, M.A. & Kohn, A. Spatial and temporal scales of neuronal correlation in primary visual cortex. *J. Neurosci.* **28**, 12591–12603 (2008).
37. Rosenbaum, R. & Doiron, B. Balanced networks of spiking neurons with spatially dependent recurrent connections. *Phys. Rev. X* **4**, 021039 (2014).
38. Yu, B., Kohn, A. & Smith, M.A. Estimating shared firing rate fluctuations in neural populations. *Soc. Neurosci. abstr.* 483.18 (2011).
39. Lin, I.C., Okun, M., Carandini, M. & Harris, K.D. The nature of shared cortical variability. *Neuron* **87**, 644–656 (2015).
40. Schölvinc, M.L., Saleem, A.B., Benucci, A., Harris, K.D. & Carandini, M. Cortical state determines global variability and correlations in visual cortex. *J. Neurosci.* **35**, 170–178 (2015).
41. Yu, B.M. *et al.* Gaussian-process factor analysis for low-dimensional single-trial analysis of neural population activity. *J. Neurophysiol.* **102**, 614–635 (2009).
42. Litwin-Kumar, A. & Doiron, B. Slow dynamics and high variability in balanced cortical networks with clustered connections. *Nat. Neurosci.* **15**, 1498–1505 (2012).
43. Chelaru, M.I. & Dragoi, V. Negative correlations in visual cortical networks. *Cereb. Cortex* **26**, 246–256 (2016).
44. Bosking, W.H., Zhang, Y., Schofield, B. & Fitzpatrick, D. Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. *J. Neurosci.* **17**, 2112–2127 (1997).
45. Pfeffer, C.K., Xue, M., He, M., Huang, Z.J. & Scanziani, M. Inhibition of inhibition in visual cortex: the logic of connections between molecularly distinct interneurons. *Nat. Neurosci.* **16**, 1068–1076 (2013).
46. Kriener, B., Helias, M., Rotter, S., Diesmann, M. & Einevoll, G.T. How pattern formation in ring networks of excitatory and inhibitory spiking neurons depends on the input current regime. *Front. Comput. Neurosci.* **7**, 187 (2014).
47. Rikhye, R.V. & Sur, M. Spatial correlations in natural scenes modulate response reliability in mouse visual cortex. *J. Neurosci.* **35**, 14661–14680 (2015).
48. Tan, A.Y., Chen, Y., Scholl, B., Seidemann, E. & Priebe, N.J. Sensory stimulation shifts visual cortex from synchronous to asynchronous states. *Nature* **509**, 226–229 (2014).

ONLINE METHODS

Description of computational model. We modeled a square of cortex with N neurons, N_e of which are excitatory and N_i inhibitory. The membrane potential of neuron j from the excitatory ($a = e$) or inhibitory ($a = i$) population obeyed exponential integrate-and-fire (EIF) dynamics,

$$C_m \frac{dV_j^a}{dt} = I_L(V_j^a) + f(V_j^a) + I_j^a(t)$$

Each time that V_j^a exceeds a threshold at V_{th} , the neuron spikes and the membrane potential is held for a refractory period τ_{ref} , then reset to a fixed value V_{re} . The leak current is given by

$$I_L(V) = -g_L(V - E_L)$$

and a spike-generating current is defined by

$$f(V) = g_L \Delta_T \exp[(V - V_T)/\Delta_T]$$

For excitatory neurons, $\tau_m = C_m/g_L = 15$ ms, $E_L = -60$ mV, $V_T = -50$ mV, $V_{th} = -10$ mV, $\Delta_T = 2$ mV, $V_{re} = -65$ mV and $\tau_{ref} = 1.5$ ms. Inhibitory neurons were the same except $\tau_m = 10$ ms, $\Delta_T = 0.5$ mV and $\tau_{ref} = 0.5$ ms.

Synaptic input currents were defined by

$$C_m^{-1} I_j^a(t) = F_j^a(t) + R_j^a(t)$$

where $F^a(t)$ is the feedforward input and $R^a(t)$ the recurrent input to neuron j in population $a = e, i$. The feedforward input was modeled differently for different figures, as described below. The recurrent input was defined by

$$R_j^a(t) = \sum_{b=e,i} \sum_{k=1}^{N_b} \frac{J_{jk}^{ab}}{\sqrt{N}} \sum_n \eta_b(t - t_n^{b,k}) \quad (4)$$

where $t_n^{b,k}$ is the n th spike time of neuron k in population $b = e, i$. The $1/\sqrt{N}$ scaling of synaptic weights is a defining feature of the balanced network formalism and captures the balance between excitatory and inhibitory currents as well as intrinsically generated temporal variability for large N (ref. 7). Each term J_{jk}^{ab} represents the synaptic weight from presynaptic neuron k in population b to postsynaptic neuron j in population a . For all simulations, we modeled synaptic kinetics using

$$\eta_b(t) = \exp(-t/\tau_b)/\tau_b \text{ for } t > 0$$

where $\tau_e = 6$ ms and $\tau_i = 5$ ms. All networks were 'dense' in the sense that connection probabilities are $O(1)$ (ref. 23).

For the model in **Figure 1**, there were $N = 20,000$ neurons, half of which were excitatory and half inhibitory. For each (presynaptic) neuron, we randomly and uniformly chose 2,500 excitatory and 2,500 inhibitory postsynaptic neurons in the network. Postsynaptic neurons were sampled with replacement, so that a single presynaptic neuron could make multiple contacts with a postsynaptic neuron. The synaptic weight of each connection depended on the pre- and postsynaptic neuron types (excitatory or inhibitory). Specifically,

$$J_{jk}^{ab} = (\text{number of contacts}) \times j_{ab} \quad (5)$$

where $j_{ee} = 12.5$ mV, $j_{ie} = 20$ mV and $j_{ii} = j_{ei} = -50$ mV. Note that synaptic weights were scaled by $\sqrt{N} = 141$ in equation (4), so that the actual synaptic weight of each contact was on the order of 0.1 mV.

For **Figure 1a–d**, the feedforward input to each neuron was given by the sum of an input bias and a smoothly varying signal,

$$F_j^a(t) = \sqrt{N} m_a + \sigma_s s(t)$$

Here, $s(t)$ is a shared source of smooth, unbiased Gaussian noise defined by its auto-covariance function,

$$\text{cov}(s(t), s(t + \tau)) = \exp(-\tau^2/2\tau_s^2),$$

$\tau_s = 40$ ms sets the correlation timescale and $\sigma_s = 0.1$ mV/ms scales the magnitude of the fluctuations. The terms $m_e = 0.015$ mV/ms and $m_i = 0.01$ mV/ms introduce a static bias to the input current. The model for **Figure 1e–h** was identical, except that two independent realizations, $s_1(t)$ and $s_2(t)$, of the shared input were generated. Half of the neurons received $s_1(t)$ and the other half received $s_2(t)$. Firing rates for **Figure 1a–d** were 7.6 Hz on average for excitatory neurons and 3.8 Hz

for inhibitory neurons. For **Figure 1e–h**, average firing rates were 7.4 Hz for excitatory and 3.8 Hz for inhibitory neurons.

To model the spatially extended recurrent network in **Figures 3** and **4**, we arranged $N_e = 40,000$ excitatory and $N_i = 10,000$ inhibitory EIF model neurons on a uniform grid covering a two-dimensional square domain. The feedforward layer was modeled by a population of $N_F = 5,625$ excitatory Poisson-spiking neurons on a uniform grid covering a square that is parallel to the recurrent network. Feedforward input to the recurrent layer was defined by

$$F_j^a(t) = \sum_{k=1}^{N_F} \frac{J_{jk}^{aF}}{\sqrt{N}} \sum_n \eta_e(t - t_n^{F,k})$$

where $t_n^{F,k}$ is the n th spike time of neuron k in the feedforward layer. Each spike train in the feedforward layer was modeled as independent Poisson processes with rate $r_F = 5$ Hz.

To simplify calculations, we measured distance in units of the side-length of the square domain. In these units, the domain is represented as the unit square, $\Gamma = [0, 1] \times [0, 1]$. Neurons were connected randomly and the probability that two neurons were connected depended on their distance measured periodically on Γ . The precise algorithm for generating connections is described in **Supplementary Note 5**. This algorithm assures that the expected number of synaptic contacts from a presynaptic neuron at coordinates $\mathbf{y} = (y_1, y_2)$ in population b to a postsynaptic neuron at $\mathbf{x} = (x_1, x_2)$ in population a is given by

$$p_{ab}(\mathbf{x} - \mathbf{y}) = \frac{K_{ab}^{\text{out}}}{N_a} g(x_1 - y_1; \alpha_b) g(x_2 - y_2; \alpha_b)$$

where $g(u; \alpha)$ is a wrapped Gaussian distribution³⁷. Out-degrees were $K_{ee}^{\text{out}} = K_{ei}^{\text{out}} = 2,000$, $K_{ie}^{\text{out}} = K_{ii}^{\text{out}} = 500$, $K_{eF}^{\text{out}} = 10,000$ and $K_{iF}^{\text{out}} = 800$. It follows that the network-wide average number of synaptic inputs to excitatory and inhibitory neurons in the recurrent network was $K = 3,715$. Synaptic weights were determined by equation (5) where $j_{ee} = 40$ mV, $j_{ie} = 120$ mV, $j_{ei} = j_{ii} = -400$ mV, and $j_{eF} = j_{iF} = 120$ mV. Note again that these terms were divided by $\sqrt{N} \approx 224$ as indicated in equation (4), so that the actual synaptic weights were between 0.18 mV and 1.8 mV.

Excitatory and inhibitory recurrent projection widths were $\alpha_{rec} = \alpha_e = \alpha_i = 0.05$ for **Figure 3** and $\alpha_{rec} = 0.25$ for **Figure 4**. Feedforward connection widths were $\alpha_{fwd} = 0.1$ in both figures. For the simulations in **Figure 3**, average firing rates were 3.9 Hz for excitatory and 6.2 Hz for inhibitory neurons. For the simulations in **Figure 4**, average rates were 4.0 Hz for excitatory and 6.1 Hz for inhibitory neurons.

The first layer (L4C) in the model in **Figure 6** was identical to the model in **Figure 3** except that $\alpha_{fwd} = 0.1$, $\alpha_e = 0.05$ and $\alpha_i = 0.03$. The length units used in **Figures 6** and **7a** were determined by interpreting the network domain, Γ , as a 10 mm by 10 mm square. Thus, in physical dimensions, $\alpha_{fwd} = 1$ mm, $\alpha_e = 0.5$ mm and $\alpha_i = 0.3$ mm. Average firing rates in the L4C layer were 3.7 Hz for excitatory and 6.1 Hz for inhibitory neurons.

Connectivity in the second layer (L2/3) in **Figure 6** was identical to that in the L4C layer except that $\alpha_e = 0.15$ (or 1.5 mm), $\alpha_{fwd} = 0.05$ (or 0.5 mm). The spike times from equation (6) for the L2/3 layer were given by the spike times of neurons in the L4C layer, so that $N_F = 50,000$. The magnitude of feedforward connectivity was also modified by setting $K_{eF} = 1,406$, $K_{iF} = 113$ and $j_{eF} = j_{iF} = 220$ mV. A shared gain modulation was implemented by altering feedforward input currents according to

$$F_j^a(t) \leftarrow F_j^a(t) [1 + w_L(\mathbf{x}) L(t)]$$

The shared gain modulation, $L(t)$, is a realization of unbiased Gaussian noise defined by its auto-covariance function

$$\text{cov}(L(t), L(t + \tau)) = \exp(-\tau^2/2\tau_L^2)$$

with correlation timescale $\tau_L = 40$ ms. The dimensionless weight factor $w_L(\mathbf{x})$ depended on the coordinates, $\mathbf{x} = (x_1, x_2) \in \Gamma$, of the neuron and was given by

$$w_L(x_1, x_2) = 0.5 g(x_1 - c; \sigma_L) g(x_2 - c; \sigma_L)$$

where $c = 0.5$ and $\sigma_L = 0.25$ (in physical dimensions, $c = 5$ mm and $\sigma_L = 2.5$ mm) so that neurons near the center of the network receive a more intense gain modulation and the strength of the modulation decays slowly with distance from the center. This imparted a long-range distance dependence to the correlations (Fig. 6c), similar to that observed in the cortical recordings (Fig. 5b). Average firing rates in the L2/3 layer were 5.7 Hz for excitatory and 9.1 Hz for inhibitory neurons.

All simulations and numerical computations were performed on a MacBook Pro running OS X 10.9.5 with a 2.3 GHz Intel Core i7 processor. All simulations were written in a combination of C and Matlab (Matlab R 2014b, MathWorks). The differential equations defining the neuron model were solved using a simple forward Euler method with time step 0.1 ms and all simulations were run for a duration of 22 s.

Experimental methods. Anesthesia was induced with ketamine (10 mg/kg) and maintained during preparatory surgery with isoflurane (1.0–2.0% in 95% O₂). Anesthesia during recordings was maintained with sufentanil citrate (6–18 $\mu\text{g}\cdot\text{kg}^{-1}\cdot\text{h}^{-1}$). Vecuronium or pancuronium bromide (0.1–0.15 $\text{mg}\cdot\text{kg}^{-1}\cdot\text{h}^{-1}$) was used to suppress eye movements. Physiological signs were monitored to ensure adequate anesthesia and animal well-being. Vital signs (EEG, ECG, blood pressure, end-tidal P_{CO_2} , temperature, and airway pressure) were monitored continuously. We used supplementary lenses to bring the retinal image into focus. At the end of the recording session, animals were killed and tissue was processed histologically to verify recording locations. All procedures were approved by the Institutional Animal Care and Use Committee of the Albert Einstein College of Medicine.

The pairwise correlations used to make the distributions in Figure 5a were computed using the same methods as those reported previously³³ and included 1,613 L2/3 pairs and 469 L4 pairs. The data were recorded from nine anesthetized, adult male macaque monkeys (*Macaca fascicularis*). Recording was made with a group of five to seven linearly arranged (305 μm spacing) platinum-tungsten electrodes or tetrodes (Thomas Recording), inserted normally to the cortical surface. The electrodes were advanced together through cortex, sampling in 200- μm intervals until all electrodes had exited into white matter. Location of the middle layer sites involved a number of criteria, including nominal depth, histology, and CSD analysis, as detailed previously³³.

Data for distance-dependent correlations reported in Figures 5b,c and 7b were from recordings described previously³⁶. The data were recorded from seven hemispheres of four adult male macaque monkeys (*M. fascicularis*). There was one experimental group of normal animals. The array consisted of a 10×10 grid of silicon microelectrodes (1 mm in length) spaced 400 μm apart. We inserted the array partially into cortex, resulting in recordings confined mostly to layers 2–3. In two cases, we recorded simultaneously with a group of seven linearly arranged (2 mm extent) platinum-tungsten microelectrodes or tetrodes (Thomas Recording), positioned so that the nearest electrode was ~5 mm anterior to one edge of the multi-electrode array. In this configuration, the distances between these electrodes and the array ranged from ~5 to 10 mm. Neuronal receptive fields in all data sets were within 5° of the fovea. Waveform segments that exceeded a user-defined threshold were sorted offline (Plexon Offline Sorter). We quantified sort quality using the signal-to-noise ratio (SNR) of each candidate unit⁴⁹, keeping units with an SNR of at least 2.3 (for the data in Fig. 5a) or at least 2.75 (for Figs. 5b and 7b). In both data sets, we eliminated neurons for which the best grating stimulus did not evoke a response of at least 2 spikes/s. Changing the SNR or responsivity threshold did not qualitatively change any of the results described herein.

Visual stimuli were generated with EXPO and displayed on a linearized CRT monitor (mean luminance 40 cd/m^2) with a resolution of 1,024 by 768 pixels and a refresh rate of 100 Hz. Stimuli were presented in a circular aperture surrounded by a gray field of average luminance. We mapped the spatial receptive field (RF) of units by presenting small (0.6°) drifting gratings at a range of spatial positions. We then centered our stimuli on the aggregate RF of the recorded units.

Stimuli were viewed binocularly and presented for 1.28 s, separated by 1.5 s intervals of isoluminant gray screen (except in one penetration, for which the interval was 10 s). We presented full-contrast drifting sinusoidal gratings at 8 or 12 orientations spaced equally (22.5° or 30° increments). The spatial frequency (1–1.3 cpd) and temporal frequency (3 or 6.25 Hz) values were chosen to correspond to the typical preference of parafoveal V1 neurons⁵⁰. The position and size (3.9–10°) of the grating were sufficient to cover the receptive fields of all the neurons. Stimulus

orientation was block randomized, each stimulus orientation was repeated 100–200 times per recording session and each of the eight recording sessions yielded 21–68 units (210–2,278 pairs) that met our SNR and responsivity thresholds. This yielded a total of 318 units and 6,907 simultaneously recorded pairs.

Statistical methods. To compute spike count correlations for Figures 1, 3, 4 and 6 we first randomly sampled a subset of the excitatory neurons in the recurrent network. The number of neurons sampled is indicated in each figure caption. Neurons with firing rates less than 1 Hz were excluded from the correlation analysis. Pearson correlation coefficients were computed by counting spikes over 250 ms windows and then computing the Pearson correlation coefficients between all pairs. The first 2 s of each simulation was excluded from the correlation analysis.

To obtain the residual correlations in Figure 7a, we randomly sampled 500 neurons with firing rates above 2 Hz from the L2/3 layer simulations. After omitting the first 2 s of the simulation, spike counts were computed using 50-ms bins, then partitioned into a sequence of 250-ms windows. To each window, we directly applied the GPFA algorithm used in a previous study³¹ and downloaded from <http://toliaslab.org/publications/ecker-et-al-2014/>. This is a modified version of the algorithm introduced in ref. 41. Briefly, GPFA extracts shared fluctuations from a population of spike trains by fitting spike count covariance matrices to a model of the form

$$\text{cov}(\mathbf{y}) = \mathbf{c}\mathbf{c}^T + \mathbf{Q}$$

Here, \mathbf{c} is a column vector of coefficients quantifying the contribution of latent variability to each neuron's spiking activity. Latent variability is modeled by a one-dimensional Gaussian stochastic process. Hence, $\mathbf{c}\mathbf{c}^T$ is a rank-one matrix of latent covariances and \mathbf{Q} quantifies the residual covariability not captured by the latent variable. Latent covariance in Figure 5c was computed by averaging the entries of $\mathbf{c}\mathbf{c}^T$ computed using GPFA. Residual correlations were computed by using the "window" option (so that spike count covariances were computed over 250-ms windows), then converting the resulting covariance matrix, \mathbf{Q} , to a correlation matrix using the `corrcoef` command in Matlab (MathWorks).

Spike count correlations in Figure 5b were computed across trials, counting spikes over 250-ms time bins and subtracting the trial-averaged rate of each unit over each stimulus orientation. Hence, all reported spike count correlations represent correlated trial-to-trial variability (noise correlations) with stimulus correlations removed. Latent covariance in Figure 5c and residual correlation in Figure 7b were computed from the mean-subtracted spike counts using 50-ms time bins and 250-ms window sizes (treating each 250-ms window as a separate trial), then using the same GPFA algorithms and procedures that were used for the simulated data in Figure 7a (see description above). The three P -values reported in Figure 7b were computed using one-sided, unpaired t -tests with $t = 8.1$, 2.1 and 2.5 and degrees of freedom 3,937, 2,404 and 2,508, respectively.

The four error bars in Figure 3e were computed from $n = 880,057$, 2,646,645, 4,420,219 and 4,550,579 pairs respectively. The five error bars in Figure 5b were computed from $n = 1,688$, 2,251, 1,948, 458 and 562 pairs, respectively, and the same for Figures 5c and 7b. The six error bars in Figure 7a were computed from $n = 3,885$, 11,696, 19,493, 27,521, 35,336 and 26,819 pairs, respectively. No statistical methods were used to predetermine sample sizes, but our sample sizes are the same as those reported in previous publications^{33,36}. Where t -tests were used, data distribution was assumed to be normal but this was not formally tested. Data collection and analysis were performed blind to the conditions of the experiments. No animals were excluded from the analysis.

Code availability. Computer code for all simulations and analysis of the resulting data is included in **Supplementary Software**.

Data availability. The data that support the findings of this study are available from the corresponding author upon request.

A **Supplementary Methods Checklist** is available.

49. Kelly, R.C. *et al.* Comparison of recordings from microelectrode arrays and single electrodes in the visual cortex. *J. Neurosci.* **27**, 261–264 (2007).

50. De Valois, R.L., Albrecht, D.G. & Thorell, L.G. Spatial frequency selectivity of cells in macaque visual cortex. *Vision Res.* **22**, 545–559 (1982).

Addendum: The spatial structure of correlated neuronal variability

Robert Rosenbaum, Matthew A Smith, Adam Kohn, Jonathan E Rubin & Brent Doiron

Nat. Neurosci. **20**, 107–114 (2017); published online 31 October 2016; addendum published after print 25 May 2017

A more complete analysis suite has been added to the Supplementary Software.