

# Slow dynamics and high variability in balanced cortical networks with clustered connections

Ashok Litwin-Kumar<sup>1,2</sup> & Brent Doiron<sup>2,3</sup>

Anatomical studies demonstrate that excitatory connections in cortex are not uniformly distributed across a network but instead exhibit clustering into groups of highly connected neurons. The implications of clustering for cortical activity are unclear. We studied the effect of clustered excitatory connections on the dynamics of neuronal networks that exhibited high spike time variability owing to a balance between excitation and inhibition. Even modest clustering substantially changed the behavior of these networks, introducing slow dynamics during which clusters of neurons transiently increased or decreased their firing rate. Consequently, neurons exhibited both fast spiking variability and slow firing rate fluctuations. A simplified model shows how stimuli bias networks toward particular activity states, thereby reducing firing rate variability as observed experimentally in many cortical areas. Our model thus relates cortical architecture to the reported variability in spontaneous and evoked spiking activity.

Cortical neurons receive input with strong temporal fluctuations and produce spike trains with high variability<sup>1,2</sup>. Models of cortical networks often use balanced excitation and inhibition to account for this variability<sup>3–5</sup>. Consistent with these studies, balanced synaptic input has been measured in cortex<sup>6</sup> and is thought to substantially influence cortical processing<sup>7</sup>. However, along with spike emission variability, cortical neurons show firing rate fluctuations over long timescales<sup>8–10</sup>. These stochastic dynamics reflect large trial-to-trial variability in cortical responses<sup>11–14</sup>. Balanced cortical networks with simple uniform connection structures do not capture these dynamics<sup>4,15</sup>. Our study's central aim was to uncover which network features, beyond balanced architecture, are responsible for slow firing rate fluctuations in cortical networks.

Recent studies show that synaptic connections between excitatory neurons in cortex are clustered rather than uniform<sup>16,17</sup>. In visual cortex, clusters are related to processing, with highly connected neurons within a cortical column receiving similar visual input<sup>18,19</sup>. Clustering may also be related to activity level, with frequently firing neurons participating in clustered subnetworks<sup>20</sup>. In total, this suggests cortical wiring involving distinct, densely connected functional subnetworks.

Functional subnetworks are a popular architectural feature supporting attractor dynamics in network models<sup>21–25</sup>. Such models typically have neuronal subpopulations with increased local excitation, leading to a multitude of stable states in which particular subpopulations exhibit sustained firing. Although attractor models have been studied in the context of memory storage, their connection to fluctuations in spontaneous and stimulus-evoked cortical dynamics has not been examined.

We investigated the dynamical consequences of clustered excitatory connections in balanced networks. In these networks, specific neuronal clusters increase or decrease their firing rates over long

timescales, promoting high trial-to-trial variability. Stimuli that excite particular clusters bias activity in those clusters while recurrent inhibition suppresses others, placing the network collectively into particular activity states. Thus, stimulation reduces firing rate variability whereas Poisson-like spiking variability remains, consistent with recent cortical recordings<sup>9</sup>. Our results show that even small perturbations from uniform connection structures—here a rewiring of only 3% of excitatory connections—can substantially change balanced network dynamics.

## RESULTS

### Clustering of connections yields new dynamics

Many real-world networks exhibit nonuniform connectivity structures, often including clustering of connections between different units<sup>26</sup>. Indeed, clustering of synaptic connections between pyramidal neurons in cortex has been observed in many studies<sup>16,17,19</sup>. However, inferring dynamical consequences from these architectural properties is challenging<sup>27</sup>. In cortex, the random nature of synaptic connectivity and individual neuron physiology adds to this difficulty. We used a simple model of excitatory clustering in a recurrent network to investigate its consequences for neural activity.

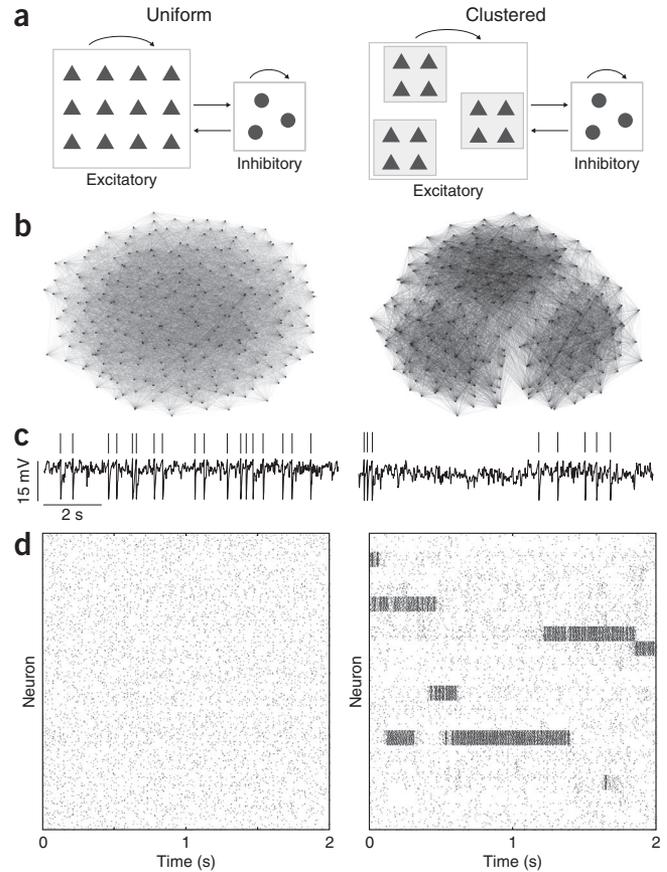
Our network consisted of 4,000 excitatory (E) and 1,000 inhibitory (I) model neurons. Connections involving inhibitory neurons were nonspecific, occurring with probability  $p^{EI} = p^{IE} = p^{II} = 0.5$  (refs. 28,29), where  $p^{xy}$  denotes the probability of a connection from a neuron in population  $y$  to a neuron in population  $x$ . Connections between excitatory neurons occurred with probability  $p^{EE} = 0.2$ . Synaptic dynamics were consistent with fast-acting excitatory and inhibitory neurotransmitters (see Online Methods).

We began by replicating the known asynchronous dynamics of uniform (non-clustered) balanced networks<sup>15</sup> (Fig. 1, left). Neuronal membrane potentials exhibited fluctuations owing to a balance

<sup>1</sup>Program for Neural Computation, Carnegie Mellon University and University of Pittsburgh, Pittsburgh, Pennsylvania, USA. <sup>2</sup>Center for the Neural Basis of Cognition, Pittsburgh, Pennsylvania, USA. <sup>3</sup>Department of Mathematics, University of Pittsburgh, Pittsburgh, Pennsylvania, USA. Correspondence should be addressed to B.D. (bdoiron@pitt.edu) or A.L.-K. (alk@cmu.edu).

Received 5 June; accepted 20 August; published online 23 September 2012; doi:10.1038/nn.3220

**Figure 1** Comparison between uniform and clustered network connectivity and dynamics. **(a)** Schematic of recurrent network, showing excitatory (triangles) and inhibitory (circles) neurons. Connections within and between populations were nonspecific for the uniform network. Shaded regions in the clustered network schematic indicate subpopulations with increased connection probability and strength for neurons belonging to the same cluster. The full network contained 4,000 excitatory neurons in 50 clusters of 80 neurons each and 1,000 inhibitory neurons. **(b)** Visualization of connectivity for a subpopulation of 240 excitatory neurons in three clusters. Nodes correspond to excitatory neurons; edges, synaptic connections. Nodes are positioned according to the Fruchterman-Reingold force algorithm (see Online Methods). Edge widths reflect synaptic strength. **(c)** Example voltage trace for an excitatory neuron. **(d)** Spike raster showing the spike times of a subpopulation of 1,600 excitatory neurons.



between excitation and inhibition<sup>4</sup> (**Fig. 1c**, left), and spiking activity was asynchronous (**Fig. 1d**, left). Our network models required no assumptions on the structure or timescale of external fluctuations, as variability arose solely from internal interactions<sup>4,5</sup>.

We next introduced clustered excitatory connections. The excitatory population was partitioned into clusters of 80 neurons each, and the connection probability  $p^{EE}$  was set to  $p_{in}^{EE}$  for neuron pairs in the same cluster or  $p_{out}^{EE}$  for neuron pairs in different clusters. The

ratio  $R^{EE} = \frac{p_{in}^{EE}}{p_{out}^{EE}}$  controlled neuronal clustering. Higher values of  $R^{EE}$

avored connections within a local cluster over nonlocal connections. The quantities  $p_{in}^{EE}$  and  $p_{out}^{EE}$  were chosen so that the connection probability between excitatory neurons remained 0.2 when averaged across all pairs. We also increased the synaptic strength for neurons in the same cluster<sup>17</sup>.

For a clustered network with  $R^{EE} = 2.5$ , visualizing the connectivity of a subset of excitatory neurons exposed clear divisions between distinct neuronal clusters (**Fig. 1a,b**, right). However, on average only 38 of 800 excitatory connections that these neurons received came from within their local cluster, and others were randomly distributed from outside their cluster. In clustered networks, neurons exhibited dynamic transitions between periods of higher and lower firing rate (**Fig. 1c,d**, right). These fluctuations reflected changes in the average firing rate of neurons in the same cluster and coexisted with randomness in the spike times of any individual neuron, yielding dynamics substantially different from those of the uniform network despite the small change in architecture.

### Spiking statistics for uniform and clustered networks

To quantify changes in network dynamics introduced by clustering, we calculated spike train statistics for excitatory neurons in uniform and clustered networks. The firing rate distribution was similar for both network types ( $3.3 \pm 4.1$  Hz for clustered and  $2.0 \pm 1.8$  Hz for uniform networks;  $\pm$  denotes s.d. across the population). Although neurons in the clustered network exhibited transient high activity states, both networks exhibited low, broadly distributed time-averaged firing rates, consistent with recordings from spontaneously active cortex<sup>30</sup> (**Fig. 2a**).

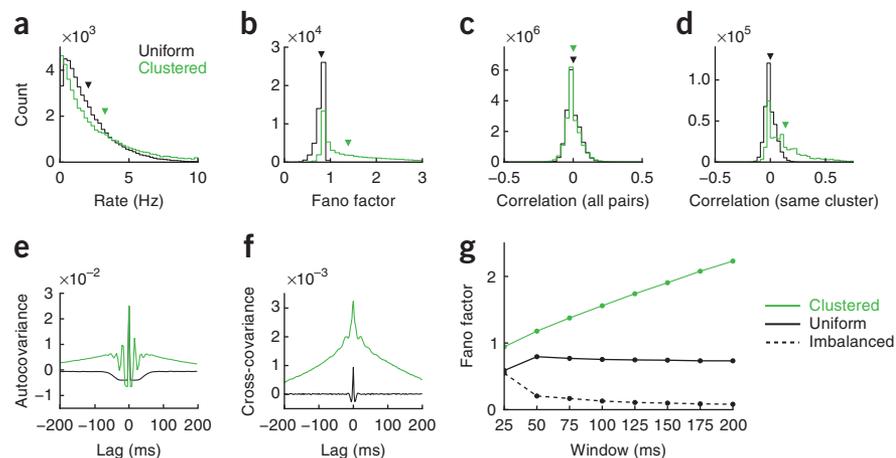
To measure spiking variability, we calculated the spike count Fano factor for each excitatory neuron, defined as the ratio of trial-to-trial variance to mean of the number of spikes a neuron emits in a fixed time window (here 100 ms; see Online Methods). For a Poisson process with a fixed firing rate, the Fano factor is 1, and Fano factors above 1 may be interpreted as evidence for variability in a neuron's underlying firing rate. High Fano factors before and during stimulus

application have been observed in many cortical systems<sup>1,9,31</sup>. In uniform networks, the mean Fano factor was 0.78 (s.d. 0.09; **Fig. 2b**). In clustered networks, it was substantially larger, with a mean of 1.4 (s.d. 0.7). This additional variability arose from firing rate fluctuations introduced by cluster state transitions from low to high activity (**Fig. 1b**). Because these transitions were stochastic, the configuration of highly active clusters for any given moment across trials was variable.

Cortical variability is also reflected in the joint activity of neuron pairs<sup>32</sup>. Excitatory neurons in both network types exhibited near-zero average spike count correlations ( $0.001 \pm 0.06$  for clustered and  $0.0005 \pm 0.05$  for uniform networks, **Fig. 2c**; see Online Methods), consistent with past theoretical studies of balanced networks<sup>15</sup>. However, differences emerged when the calculation of correlation was restricted to neurons in the same cluster. Clustered networks exhibited a distribution of correlation coefficients for intra-cluster pairs with a long positive tail and a mean of 0.13 (s.d. 0.18; **Fig. 2d**). The distribution's tail reflected the slow correlated fluctuations in average firing rate that neurons experienced. Because intra-cluster pairs only accounted for 2% of all excitatory neuron pairs, this tail was not visible in the distribution of randomly sampled correlation coefficients (**Fig. 2c**).

To analyze this slow spiking activity, we computed spike train autocovariance and cross-covariance functions for excitatory neurons (see Online Methods). Neurons in uniform networks exhibited no long timescales in their autocovariance or cross-covariance functions, whereas in clustered networks both exhibited long timescale decays (**Fig. 2e,f**). Both functions indicated a weak oscillatory tendency in the gamma frequency range due to interactions between excitatory and inhibitory populations<sup>33</sup>.

**Figure 2** Marginal and pairwise spiking statistics for neurons in clustered and uniform networks. Statistics for each network type were computed over 12 realizations for each of the two connectivities, with 9 trials for each realization of connectivity. (a) Histogram of excitatory neuron firing rates in clustered and uniform networks. Triangles mark mean value. (b) Histogram of excitatory neuron Fano factors computed over 100-ms windows. (c) Histogram of correlation coefficients computed over 50-ms windows for all excitatory neuron pairs. (d) Same as c, but only computed for neuron pairs belonging to the same cluster. The distribution of correlation coefficients for a random subset of pairs from the uniform network is shown, as the uniform network lacked clusters. (e) Average autocovariance function for excitatory neurons. (f) Average cross-covariance function for neuron pairs belonging to the same cluster. Correlation functions are normalized to firing rate. (g) Fano factors computed over different counting windows, comparing clustered, uniform and imbalanced networks. Imbalanced networks did not obey a balance between excitation and inhibition, and responses were mean-driven rather than fluctuation-driven (see Online Methods).



Finally, we explicitly related the system's long timescales to trial-to-trial variability by computing Fano factors as a function of time window. Compared to an imbalanced network in which responses were mean-driven, Fano factors for balanced networks were substantially larger (Fig. 2g), although they were still sub-Poisson owing to refractory effects (Fig. 2e), saturating for time windows greater than 100 ms. However, Fano factors increased for longer time windows in clustered networks. In total, clustered networks had slow firing rate fluctuations while retaining the fast spiking variability characteristic of balanced networks, making neuronal activity effectively a 'doubly stochastic' process<sup>9,10,34</sup>.

### Heterogeneity in cluster membership

In previous sections, we studied an idealized model in which all clusters were identically sized and neurons belonged to only one cluster. This connectivity led to discrete and unambiguous transitions between high and low activity states (Fig. 1). However, cortical assemblies may involve heterogeneity in size and overlap in membership. To determine whether our results hold for heterogeneous clustered networks, we generated a new network architecture in which cluster membership was random. Each excitatory neuron was assigned to two clusters out of 100 total, and connection probability and strength were increased for neuron pairs that shared membership in at least one cluster (Fig. 3a). Because of the random assignment of neurons to clusters, cluster sizes were heterogeneous (Fig. 3b).

Because neurons were not ordered by group membership, transitions were not immediately apparent in an unordered raster (Fig. 3c). However, when rows in the raster were sorted by cluster, slow transitions appeared (Fig. 3d). To compare homogeneous and heterogeneous clustered network dynamics, we calculated average firing rates for all neurons in individual clusters. Clusters in homogeneous networks exhibited discrete transitions between low and high activity states, with average firing rates of 50–60 Hz (Fig. 3e, top). In contrast, heterogeneous clustered networks exhibited more complicated temporal dynamics (Fig. 3e, bottom). Hence, networks with clustered architectures can generate complex firing rate patterns reminiscent of those seen in spontaneous cortical recordings<sup>13</sup>, although our study's primary results can be understood in the idealization of homogeneous clustered networks.

### Increased clustering lengthens rate fluctuation timescale

We have shown that clustered connections introduce new, slow dynamics not present in uniform networks (Fig. 1). We next show that this timescale is related to the density of clustering.

We examined three different clustered network connectivities with  $R^{EE}$  increasing from 2 to 3. As  $R^{EE}$  increased, neurons were more easily separated into distinct clusters (Fig. 4a) and clusters exhibited longer periods of high activity (Fig. 4b). To quantify the firing rate fluctuations' dependence on network connectivity, we calculated the average excitatory neuron Fano factor as a function of  $R^{EE}$  (Fig. 4c). Fano factors increased for sufficiently high clustering. The degree of clustering  $R^{EE}$  can be recast in terms of the percentage of excitatory connections that need to be rewired to generate a clustered network from a uniform one. This quantity was below 5% for the range studied (Fig. 4d). Thus only a small perturbation from uniform connectivity is required for substantial firing rate fluctuations.

### Dependence of dynamics on cluster and network size

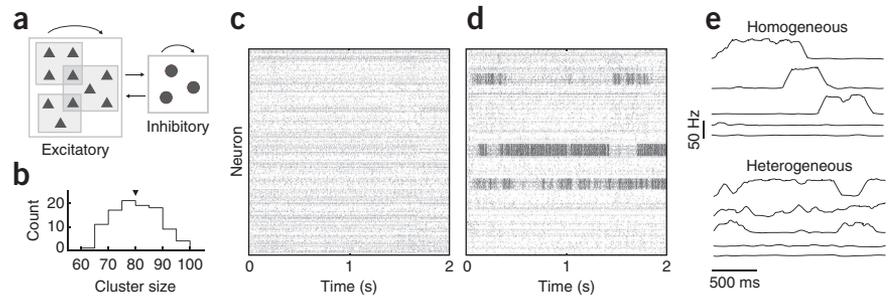
We next investigated whether changes in network dynamics were robust to different connectivity parameters. Above we saw that, above a critical level of clustering (approximately  $R^{EE} = 2.5$ ), the network exhibited high firing rate variability. However, for large  $R^{EE}$ , transitions occurred over very long timescales (Fig. 4b, right). Hence variability on biophysically relevant timescales should be maximized for a fixed level of clustering, a prediction we investigated by varying network parameters.

We evaluated cluster firing rate variability as a function of  $R^{EE}$ , for different network sizes  $N$  and cluster sizes  $C$  (Fig. 4e). For each combination of  $N$  and  $C$ , there existed a critical value of  $R^{EE}$  above which firing rate variability increased. However, variability eventually decreased for higher  $R^{EE}$ , when the transition timescale was longer than the simulation time. This was in contrast to the dependence of the Fano factor on clustering (Fig. 4c), as Fano factors were computed across trials with different initial conditions whereas firing rate variability was computed over time.

The shape of these curves depended both on  $N$  and  $C$ . As  $C$  increased, the firing rate variability peak location and height decreased because more excitatory connections were present and cluster transition times diminished. However, the peak location increased as  $N$  grew because, in balanced networks, synaptic strength scales inversely with network

**Figure 3** Heterogeneous clustered network.

(a) Schematic of connectivity for heterogeneous network. Each excitatory neuron was assigned randomly to two clusters (schematic shows only a subset of neurons and clusters). Connection probability and strength was increased for any neuron pair that shared membership in at least one cluster. (b) Histogram of the number of neurons per cluster for the network shown in the following panels. Cluster size was heterogeneous because membership was assigned randomly. For this network, clusters ranged in size from 63 to 99 neurons. (c) Spike raster showing the spike times of 1,600 excitatory neurons. As cluster membership was chosen randomly, the neurons are not ordered by cluster. (d) Same as c, but row indices reordered so that blocks of rows correspond to clusters. Some rows are repeated, as neurons belonged to more than one cluster. (e) Example average firing rates for different clusters from the homogeneous clustered network studied in **Figures 1 and 2** and the heterogeneous network. Rates were estimated from the number of spikes emitted by all neurons in a cluster in a sliding window of 100 ms.



size (see Online Methods). Because a peak occurred for each combination of  $N$  and  $C$ , our results are robust to changes in network parameters. We further explore network scaling in the next section.

### Theoretical analysis

Our results indicate that network architecture, more so than individual neuron dynamics, gives rise to cluster state transitions. We therefore use a simplified binary neuron model to relate architecture to transitions. Similar analyses have been performed for memory networks<sup>15,24,35</sup>, but the timescale of spontaneous transitions between different states has not been examined.

Our model simplifies the state of neuron  $i$ , denoted  $s_i$ , to either inactive ( $s_i = 0$ ) or active ( $s_i = 1$ ). The input to neuron  $i$ , denoted  $I_i$ , is a weighted sum of the activities of the  $K$  excitatory and  $K$  inhibitory neurons from which it receives input plus a constant mean bias:  $I_i = \sum_j J_{ij} s_j + \mu_i$ , where  $\mu_i$  is the bias and  $J_{ij}$  is the strength of a synaptic connection from neuron  $j$  to neuron  $i$ . If  $I_i$  is positive, then the neuron is active; otherwise, it is inactive.

We began by studying a network whose connectivity was uniform except for a single cluster of  $C$  excitatory neurons. A neuron in the cluster receives  $K_{in}^{EE} = p_{in}^{EE} C$  synaptic inputs from other neurons in the cluster. The mean and variance of the total input to such a neuron are

$$\langle I_{in}^E \rangle = J_{in}^{EE} K_{in}^{EE} \langle s_{in}^E \rangle + J_{out}^{EE} K_{out}^{EE} \langle s_{out}^E \rangle + J^{EI} K \langle s^I \rangle + \mu^E \quad (1)$$

$$\text{Var}(I_{in}^E) = (J_{in}^{EE})^2 K_{in}^{EE} \langle s_{in}^E \rangle + (J_{out}^{EE})^2 K_{out}^{EE} \langle s_{out}^E \rangle + (J^{EI})^2 K \langle s^I \rangle \quad (2)$$

Here  $J^{xy}$  denotes the strength of synaptic connections from neurons in population  $y$  to those in population  $x$ , and  $\langle s^x \rangle$  corresponds to the average activity of neurons in population  $x$ . Subscripts “in” and “out” refer to excitatory neurons within the cluster and outside the cluster, respectively. Because  $\text{Var}(I_{in}^E) \sim J^2 K$ , we need  $J \sim \frac{1}{\sqrt{K}}$  for  $\text{Var}(I_{in}^E)$  to be  $O(1)$ . As a consequence,  $\langle I_{in}^E \rangle$  will diverge as  $\sqrt{K}$  unless excitation and inhibition balance<sup>4</sup>.

Under these conditions, we analyzed the effect of local cluster input. The first term in equation (1) is the average recurrent excitation from other neurons in the cluster:

$$I_{rec} = J_{in}^{EE} K_{in}^{EE} \langle s_{in}^E \rangle \quad (3)$$

For  $I_{rec}$  to be  $O(1)$  as  $K$  becomes large, we need  $K_{in}^{EE}$  to be inversely proportional to  $J_{in}^{EE}$ . Because  $J \sim \frac{1}{\sqrt{K}}$ , we therefore require

$$K_{in}^{EE} \sim \sqrt{K} \quad (4)$$

In other words, for a balanced network, local cluster inputs may be much smaller in number (by a factor of  $\sqrt{K}$ ) than total inputs  $K$  but still have a substantial effect. Indeed, as noted previously, neurons received on average only 38 of 800 total excitatory inputs from within their local cluster in our integrate-and-fire simulations. This sensitivity to rewiring is a consequence of the balance between excitation and inhibition; small deviations in balance due to local cluster activity strongly affect neurons’ firing rates.

### Transitions between low and high activity states

With our simplified model, we next studied the dynamics of  $\langle s_{in}^E \rangle$ , the average firing rate of neurons in the cluster. We determined the potential energy landscape  $U$  that governed this quantity (see Online Methods). Minima in  $U$  correspond to stable values of  $\langle s_{in}^E \rangle$ . For small  $R^{EE}$ , the only minimum was at a low firing rate, so no transitions could occur (**Fig. 5a**). When  $R^{EE}$  increased sufficiently, a second minimum appeared, and  $\langle s_{in}^E \rangle$  was therefore bistable, as is common in networks with local recurrent excitation<sup>21–24,35</sup>. Furthermore, because the cluster size was small, dynamic fluctuations due to finite size effects caused transitions between minima. Deeper potential wells corresponded to more attractive activity states, as quantified by average time spent in a well (**Fig. 5b**).

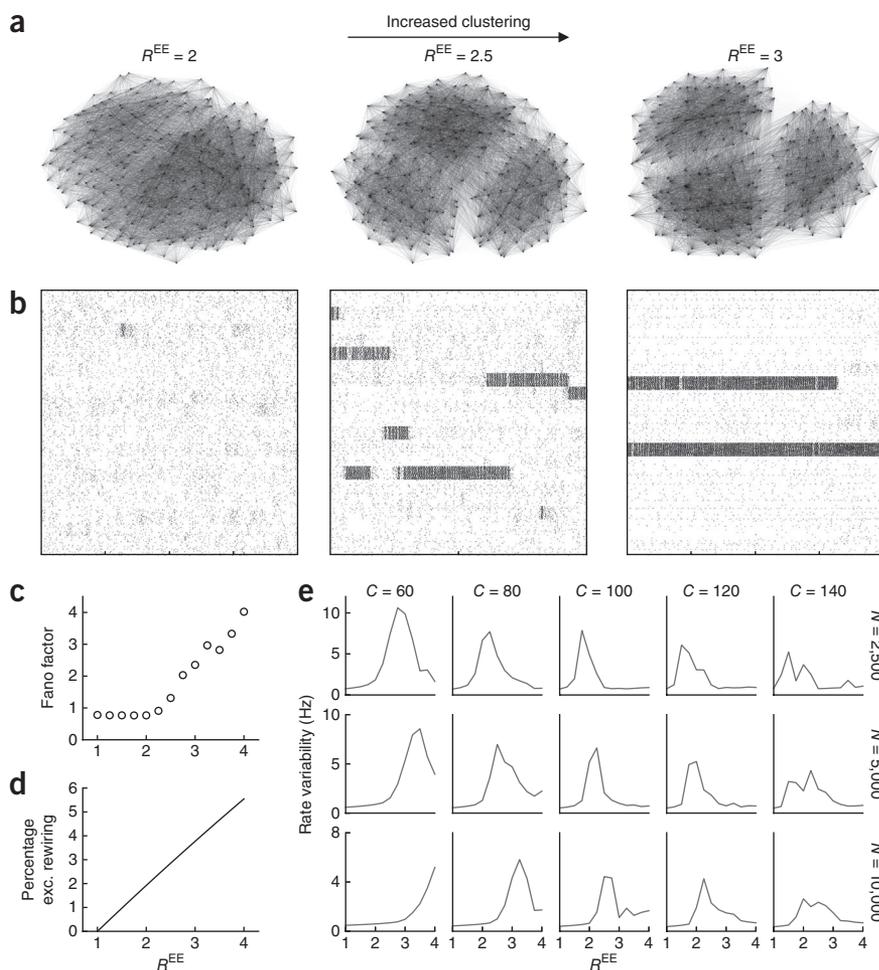
Transitions between low and high activity states introduced long-timescale variability to the cluster’s activity (**Fig. 5c**, top). Although transitions between wells were a source of variability in our model, we emphasize that they were not the only source of variability. Balanced excitation and inhibition ensured that neurons were bombarded with many synaptic inputs that produced high input variability relative to input mean, even when conditioned on their cluster being in one state (**Fig. 5c**, bottom). Model neurons were therefore doubly stochastic, with spike time variability because of this broad input distribution and firing rate variability because of transitions between wells.

What controls the timescale of these transitions? For deep enough wells and assuming that the effective noise driving average cluster activity is Gaussian and white, the transition time out of a well of depth  $\Delta U$  will scale as

$$T \propto e^{\frac{k\Delta U}{D}} \quad (5)$$

where  $k$  is a constant and  $D$  is the amplitude of the effective diffusion driving the fluctuations in mean activity<sup>36</sup> (see Online Methods).  $D$  depends both on system size, with larger clusters exhibiting diminished fluctuations owing to averaging, and on correlations between neurons, which amplify fluctuations in mean activity.

As increased clustering increases the depth of the high activity state’s well (**Fig. 5a**), we expect highly clustered networks to take longer to



**Figure 4** Effect of increased clustered connection probability. **(a)** Visualization of connectivity for a subset of 240 neurons in three different networks. The network for  $R^{EE} = 2.5$  is identical to the network shown in **Figure 1**, right. **(b)** Spike raster for a subset of 1,600 excitatory neurons from the three different network configurations. As clustering increased, the timescale of cluster-activity state transitions lengthened. **(c)** Average Fano factor computed over 100-ms windows among excitatory neurons as a function of  $R^{EE}$ . **(d)** Average percentage of excitatory (exc.) connections rewired as a function of  $R^{EE}$ . Percentage rewired was defined by calculating the average number of connections an excitatory neuron received from other neurons in its cluster, subtracting the average number of such connections it would have received if  $R^{EE} = 1$  and dividing by the total number of connections. **(e)** Dependence of firing rate variability on clustering  $R^{EE}$ , for different values of cluster size  $C$  and total network size  $N$ . The firing rate variability for each cluster was estimated by computing the population firing rate for the  $C$  neurons in the cluster over 2 s, in 20 windows of 100 ms each. This was then averaged over all clusters and realizations.

recurrent network, this is not the case. Recurrent inhibition mediates cluster–cluster interaction, limiting the number of clusters simultaneously in the high activity state and thus promoting transitions. We examine this more closely as we next consider the relationship between stimulation and clustered network dynamics.

transition out of this state. To verify this prediction, we measured the average transition time out of the high activity state for networks with different degrees of clustering and plotted the result as a function of  $\Delta U$  (**Fig. 5d**). The logarithm of  $T$  increased linearly with  $\Delta U$ , as predicted by equation (5). In total, architectural clustering controls  $U$  and hence the presence of bistability and the attractiveness of high activity states.

We next sought to understand how clusters interact. So far, our analysis had focused on a single cluster's dynamics, neglecting its influence on other clusters or the possibility of other clusters changing state. Indeed, if every cluster was independent and all were governed by the potential in **Figure 5b**, we would expect most clusters to remain in the high activity state and average activity to be unrealistically high. In the fully

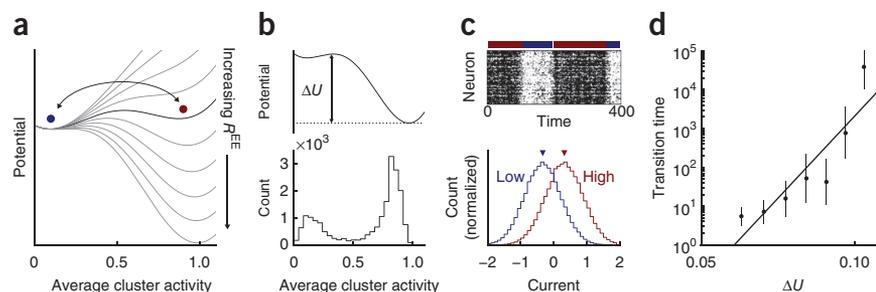
### Biasing activity state with stimulation

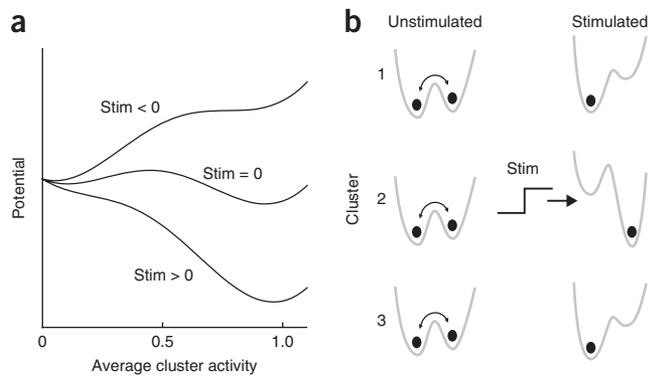
The relationship between spontaneous and evoked trial-to-trial variability has recently attracted attention<sup>9,37</sup>. Throughout the cortex, trial-to-trial spiking variability is reduced in driven conditions, when compared to that measured in spontaneous states<sup>9</sup>. Despite its generality, the mechanisms responsible for this variability reduction are undetermined. With our theory, we examined stimulation's effects on variability in clustered networks.

We calculated the potential  $U$  for different stimulus values (**Fig. 6a**). Negative stimuli made the high activity state less attractive, while positive stimuli increased its attractiveness. This, along with previous observations about recurrent inhibition, leads to a prediction about

**Figure 5** Emergence of bistability in simplified

model. **(a)** Potential governing dynamics of average cluster activity  $\langle s_{in}^E \rangle$ , for different values of clustering  $R^{EE}$ . As clustering increases beyond a critical value, bistability emerges (black curve) and switches (arrow) can occur between the two stable states. **(b)** Potential (top) and corresponding histogram for average cluster activity (bottom). **(c)** Top: example activity raster for the parameters in **b**. Transitions are evident between the high activity (red) and low activity (blue) states. Bottom: histogram of input currents for neurons in the cluster conditioned on the cluster being in the low or high activity state (defined as average activity less than or greater than 0.5, respectively) and normalized in peak height. **(d)** Average timescale of transitions from high activity to low activity state as a function of well depth  $\Delta U$ . Each point corresponds to an average over four networks simulated for 200,000 time steps. The line corresponds to a linear regression of well depth against the logarithm of the transition time (coefficient of determination  $R^2 = 0.89$ ).





**Figure 6** Effects of stimulation. **(a)** Potentials for  $R^{EE} = 2$  and different values of stimulus (stim) applied to the cluster. A stimulus can bias activity toward the low or high activity state depending on its sign. **(b)** Consequences of stimulation and recurrent inhibition. Left: when all clusters are symmetric, any cluster can transition to the high activity state. Right: when cluster 2 receives a stimulus to bias it to the high activity state, other clusters are suppressed by recurrent inhibition. Hence the network remains in one particular activity configuration.

network dynamics under stimulation. Without stimulation, clusters are symmetric and any cluster may transition between states (**Fig. 6b**, left). When a fraction of clusters are stimulated, this symmetry is broken and stimulated clusters are more likely to have high activity. Recurrent inhibition discourages other clusters from transitioning to high activity states, forcing the entire network into a single configuration and thus suppressing long-timescale variability (**Fig. 6b**, right). We next investigate this prediction, returning to a full spiking network.

### Effect of stimulation on spiking variability

Because our spiking model exhibited high trial-to-trial variability consistent with cortical data, we investigated the effect of stimulation, using a simple depolarizing stimulus applied to 5 of the 50 clusters for 400 ms. During spontaneous activity, clusters exhibited transitions. When the stimulus was applied, stimulated clusters became highly active, while unstimulated clusters were suppressed (**Fig. 7a,b**). After stimulus application ceased, the system eventually relaxed to spontaneous dynamics.

To quantify reductions in variability, we calculated the mean-matched Fano factor over 100-ms windows as a function of time. Mean matching controlled for stimulus-induced changes in firing rate, ensuring that observed changes in Fano factor were due to a change in underlying rate variability<sup>9</sup>. For clustered networks, Fano factors were above 1 during spontaneous activity but dropped slightly below 1 during stimulation (**Fig. 7c**). Similar results were obtained for heterogeneous clustered networks (**Supplementary Fig. 1**), clustered networks with more complex stimulus tuning (**Supplementary Fig. 2**) and clustered networks with ring or feed-forward structure (**Supplementary Fig. 3**). In contrast, uniform networks exhibited no noticeable decrease in variability when the stimulus was applied (**Fig. 7c**). The Fano factor restricted to

either stimulated or unstimulated neurons also exhibited a reduction in Fano factor only for clustered networks (**Fig. 7d**). These results are consistent with cortical recordings in which neurons whose preference did not match a presented stimulus nevertheless showed a stimulus-induced drop in variability<sup>9</sup>, a feature not present in other models of cortical variability<sup>25</sup>.

### Variability of inhibitory neurons

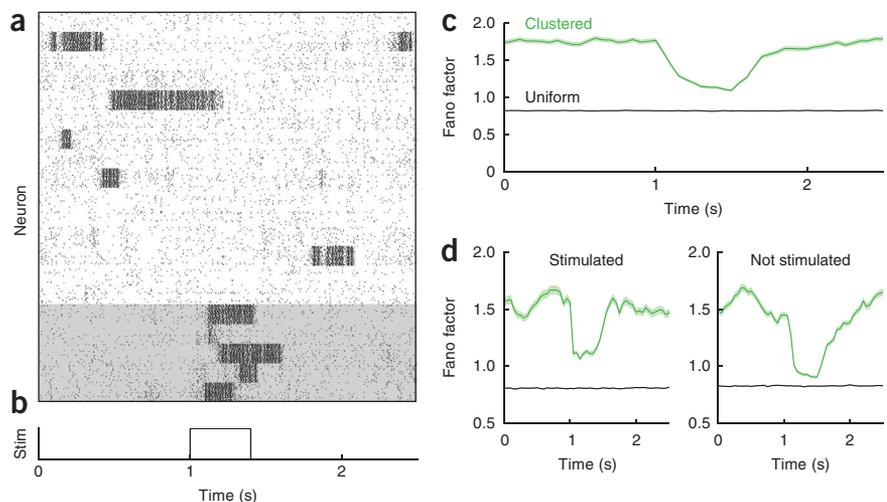
To this point we have been exploring excitatory neuron spiking variability, but inhibitory neuron variability may also be stimulus dependent. In visual area V4, a reduction in inhibitory neuron Fano factor with stimulation has been reported<sup>38</sup>, but firing rates increased substantially during stimulation, and no rate matching was performed to control for this modulation. We show that inhibitory neurons in clustered networks may also exhibit decreased Fano factors due to stimulation, an effect that depends on the specificity of their afferent connections from excitatory neurons.

In the above model, the connections involving inhibitory neurons were uniform. Anatomical evidence suggests that inhibitory innervation of pyramidal neurons is nonspecific<sup>29</sup>, but there is some evidence for selectivity in excitatory input to certain classes of inhibitory interneurons<sup>39</sup>. We therefore studied the dependence of inhibitory neuron variability on excitatory-to-inhibitory input specificity.

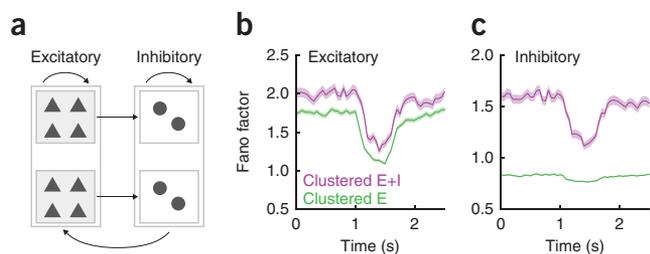
We compared the clustered networks presented previously to ones in which inhibitory neurons were more likely to receive input from excitatory neurons in a particular cluster (**Fig. 8a**). In clustered E+I networks, we assigned inhibitory neurons to 50 clusters of 20 neurons each. Excitatory-to-inhibitory connections occurred with probability  $p_{in}^{IE}$  if both neurons belonged to the  $i^{\text{th}}$  cluster of their respective popu-

lation and otherwise with probability  $p_{out}^{IE}$ .  $R^{IE} = \frac{p_{in}^{IE}}{p_{out}^{IE}}$  was set to 1.5.

Excitatory neurons in both networks had high variability and exhibited a reduction in Fano factor with stimulus (**Fig. 8b**). However, inhibitory neuron Fano factors in clustered E+I networks increased



**Figure 7** Effect of stimulation on spiking variability and dynamics. **(a)** Spike raster showing the activity of 1,600 excitatory neurons on one trial. Neurons in the shaded region received the stimulus when applied. **(b)** Time course of stimulus, a depolarizing step of current that lasted 400 ms. The stimulus caused the clusters that received it (**a**, shaded region) to transition into a high activity state. **(c)** Fano factor computed over 100-ms windows as a function of time for excitatory neurons. For comparison, the average Fano factor for uniform networks receiving an identical stimulus is shown. Only clustered networks exhibited a noticeable decrease in variability when stimulated. Shaded regions denote 95% confidence intervals. **(d)** Left: same as **c**, but restricted to neurons in stimulated clusters (corresponding to **a**, shaded region). Right: same as left, but restricted to neurons not in stimulated clusters. Fano factors were computed using mean-matching techniques presented in ref. 9.



**Figure 8** Effect of stimulation on inhibitory neuron spiking variability. (a) Schematic of network connectivity for the clustered E+I network. Excitatory projections were cluster-specific, whereas inhibitory projections were uniform. (b) Mean-matched Fano factor computed over 100-ms windows as a function of time for excitatory neurons in the clustered E+I network and the clustered E network (Fig. 7). Shaded region denotes 95% confidence interval. (c) Same as b but for inhibitory neurons.

substantially compared to the clustered excitatory network, with a correspondingly greater reduction with stimulation (Fig. 8c). Thus, excitatory input specificity can drive stimulus-induced variability reduction in inhibitory neurons. Studies have shown that the specificity of excitatory input to different inhibitory cell classes varies<sup>39</sup>, with fast-spiking interneurons receiving more specific input than adapting interneurons. We predict that adapting interneurons should therefore exhibit a smaller effect of stimulus on variability.

## DISCUSSION

We have shown that clustering of excitatory connections<sup>16,17,19</sup> substantially changes balanced network dynamics. Small perturbations of connectivity that introduced excitatory clustering yielded dynamical phenomena observed in cortex experimentally but not present in simple uniform networks: slow firing rate fluctuations in spontaneous conditions and stimulus-induced reductions in trial-to-trial rate variability with spike time variability remaining intact<sup>9</sup>.

### Anatomical realism and attractor dynamics

The network connectivity in this study was motivated by anatomical evidence for clustering of connections between pyramidal neurons in cortex. Bidirectional and clustered three-neuron connection motifs occur with frequencies significantly above chance in the visual system<sup>16</sup>. These subnetworks also receive similar feedforward input, suggesting that clustering may be related to functional maps<sup>18</sup>, an idea with recent *in vivo* support<sup>19</sup>. Inhibitory projections were not similarly clustered in our model, consistent with studies showing that inhibitory neurons connect densely and nonspecifically to pyramidal neurons<sup>29</sup> and that inhibitory neuron connectivity is less tuning specific than that of pyramidal neurons<sup>40</sup>.

The dynamics we investigated are reminiscent of persistent state activity, often studied in attractor networks with highly segregated subnetworks<sup>22,25</sup>. Spatially dependent connection probabilities can also lead to persistent activity when excitation is more local than inhibition<sup>41</sup>. Our result for clustered networks, although it involves excitation to a local subset of neurons, assumes neither a spatial distribution of connections nor an excessive reorganization from an unstructured network. Similar results are obtained in networks with spatially dependent connection structures, as long as an underlying multistability is present (Supplementary Fig. 3).

### Robustness and dependence of cluster size

Our results depended on cluster size, and our parameter choices are consistent with anatomical evidence for fine-scale clusters of tens, rather than thousands, of neurons<sup>17</sup>. Large-scale, biophysically realistic

network models have shown clusters of this size emerging with spike time-dependent plasticity<sup>42</sup>. For a variety of cluster and network sizes, high firing rate variability emerged in our model (Fig. 4). Because larger clusters have smaller fluctuations in mean firing rate owing to averaging, very large clusters were unlikely to exhibit frequent transitions and also tended to suppress smaller clusters (Fig. 3). However, large-scale correlated fluctuations, adaptation or other mechanisms not modeled in our study could promote transitions even for very large clusters.

We also note that our model exhibits sharp transitions between attractor states with large differences in firing rate (Fig. 3e). These shifts are consistent with rapid state transitions observed in behaving animals<sup>43</sup>. Although average firing rates in our model are consistent with spontaneous data (Fig. 2a), activity within active clusters is more regular than typically recorded in spontaneous activity. This observation has been made in previous studies of bistability in balanced networks<sup>21–23</sup>. Exactly matching higher order spike train statistics is difficult, yet the agreement for time-averaged firing rate and spike train variability is promising. As such, clustered networks provide a simplified framework for studying multistability and the spiking variability it produces.

### Models of cortical variability and co-variability

Recently, it has been shown that uniform, balanced networks evolve to an asynchronous state in which correlations are negligible even when connections are dense<sup>15</sup>, consistent with some experimental data<sup>44</sup>. The distribution of correlation coefficients in clustered networks has a near-zero mean, similar to that in these studies (Fig. 2c). However, neuron pairs within the same cluster do exhibit nonzero correlations (Fig. 2d). Assuming that clusters are related to functional processing, this observation is consistent with experimental studies showing that correlations are enhanced for similarly tuned neurons<sup>8,32</sup> and predicts that measured correlations will depend on the subset of neurons to which recording is restricted.

Other modeling studies have examined networks exhibiting high variability from conductance modulations or other biophysical properties<sup>5,45</sup>, but long-timescale firing rate variability (Fig. 2g) has not been addressed. Our minimal model captures the distributions of firing rates<sup>30</sup> (Fig. 2a), pairwise spiking correlations<sup>32,44</sup> (Fig. 2c,d,f) and rate variability<sup>9</sup> (Fig. 2b,e) characteristic of spontaneous, unanesthetized cortex. However, dual whole-cell recordings from spontaneously firing excitatory neurons in the somatosensory cortex of awake rodents show prominent, low-frequency correlated membrane potential fluctuations<sup>46</sup>. Our model does not capture this correlation, nor does any other spiking model (to our knowledge) lacking externally imposed variability. Imaging studies suggest that the correlated fluctuations are likely due to traveling waves between motor and sensory cortex<sup>47</sup>, and spatial effects would need to be introduced into a model to properly account for these dynamics (Supplementary Fig. 3). Nonetheless, this membrane potential correlation does not transfer to synchronous excitatory neuron spiking<sup>46</sup>, presumably owing to sparse firing in the spontaneous state. Understanding the source of these fluctuations is an important topic for future study.

### Relationship between spontaneous and evoked activity

Patterns of neuronal activity evoked by sensory stimulation are thought to be a subset of those sampled during spontaneous activity<sup>12,14</sup>. Clustered network dynamics are consistent with this idea, with activity states sampled stochastically in spontaneous conditions and biased toward particular states with stimuli. The relationship between strong local connectivity and increased attractiveness of high-activity states (Fig. 4) suggests that Hebbian plasticity may wire together stimulated clusters so that they are sampled more frequently. Thus, the spontaneous state's statistics would reflect past stimulus input statistics, as reported in visual cortex<sup>48</sup>.

Observations of reductions in single-neuron trial-to-trial variability from spontaneous to evoked conditions complement the above data<sup>9</sup>. In describing these results, spiking activity is typically characterized as a doubly stochastic process, in which variability in spike emission and rate dynamics can be separated. While this simplification is attractive, the two are not easily segregated in mechanistic models. Stimulus-induced reductions in variability have been studied in firing rate models<sup>37</sup>; however, these models implicitly separate firing rate and spiking variability. In contrast, a recent attractor-based study<sup>25</sup> also modeled variability reduction but assumed doubly stochastic external inputs. We studied generation of both rate and spike time variability by internal network interactions, providing mechanistic plausibility for the intuition presented in ref. 9.

Finally, our work begins to link cortical architecture, variability and computation. There is growing evidence that probabilistic inference is central to neural coding<sup>10,49</sup>. Rich, stochastic spontaneous dynamics may reflect sampling that supports probabilistic inference<sup>50</sup>. Merging attractor theories of computation with probabilistic coding schemes is an important direction for systems neuroscience, for which a clear theory of neural variability in cortical circuits will be needed.

## METHODS

Methods and any associated references are available in the online version of the paper.

Note: Supplementary information is available in the online version of the paper.

## ACKNOWLEDGMENTS

We thank R. da Silveira and H. Sompolinsky for discussions that helped shape an early version of this work. We also thank B. Yu and J. de la Rocha for comments. B.D. was supported by the US National Science Foundation (NSF DMS-1121784) and a Sloan research fellowship. A.L.-K. was supported by the US National Defense Science & Engineering Graduate Fellowship program.

## AUTHOR CONTRIBUTIONS

A.L.-K. and B.D. conceived the study and wrote the manuscript. A.L.-K. performed the simulations and analyzed the data.

## COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at <http://www.nature.com/dofinder/10.1038/nn.3220>.

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

- Britten, K.H., Shadlen, M.N., Newsome, W.T. & Movshon, J.A. Responses of neurons in macaque MT to stochastic motion signals. *Vis. Neurosci.* **10**, 1157–1169 (1993).
- London, M., Roth, A., Beeren, L., Hausser, M. & Latham, P.E. Sensitivity to perturbations in vivo implies high noise and suggests rate coding in cortex. *Nature* **466**, 123–127 (2010).
- Shadlen, M.N. & Newsome, W.T. Noise, neural codes and cortical organization. *Curr. Opin. Neurobiol.* **4**, 569–579 (1994).
- van Vreeswijk, C. & Sompolinsky, H. Chaotic balanced state in a model of cortical circuits. *Neural Comput.* **10**, 1321–1371 (1998).
- Vogels, T.P. & Abbott, L.F. Signal propagation and logic gating in networks of integrate-and-fire neurons. *J. Neurosci.* **25**, 10786–10795 (2005).
- Shu, Y., Hasenstaub, A. & McCormick, D.A. Turning on and off recurrent balanced cortical activity. *Nature* **423**, 288–293 (2003).
- Destexhe, A., Rudolph, M. & Pare, D. The high-conductance state of neocortical neurons in vivo. *Nat. Rev. Neurosci.* **4**, 739–751 (2003).
- Kohn, A. & Smith, M.A. Stimulus dependence of neuronal correlation in primary visual cortex of the macaque. *J. Neurosci.* **25**, 3661–3673 (2005).
- Churchland, M.M. *et al.* Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nat. Neurosci.* **13**, 369–378 (2010).
- Churchland, A.K. *et al.* Variance as a signature of neural computations during decision making. *Neuron* **69**, 818–831 (2011).
- Arieli, A., Sterkin, A., Grinvald, A. & Aertsen, A. Dynamics of ongoing activity: Explanation of the large variability in evoked cortical responses. *Science* **273**, 1868–1871 (1996).
- Tsodyks, M., Kenet, T., Grinvald, A. & Arieli, A. Linking spontaneous activity of single cortical neurons and the underlying functional architecture. *Science* **286**, 1943–1946 (1999).
- Churchland, M.M. *et al.* Neural variability in premotor cortex provides a signature of motor preparation. *J. Neurosci.* **26**, 3697–3712 (2006).
- Luczak, A., Bartho, P. & Harris, K.D. Spontaneous events outline the realm of possible sensory responses in neocortical populations. *Neuron* **62**, 413–425 (2009).
- Renart, A. *et al.* The asynchronous state in cortical circuits. *Science* **327**, 587–590 (2010).
- Song, S., Sjöström, P.J., Reigl, M., Nelson, S. & Chklovskii, D.B. Highly nonrandom features of synaptic connectivity in local cortical circuits. *PLoS Biol.* **3**, e68 (2005).
- Perin, R., Berger, T.K. & Markram, H. A synaptic organizing principle for cortical neuronal groups. *Proc. Natl. Acad. Sci. USA* **108**, 5419–5424 (2011).
- Yoshimura, Y., Dantzker, J.L.M. & Callaway, E.M. Excitatory cortical neurons form fine-scale functional networks. *Nature* **433**, 868–873 (2005).
- Ko, H. *et al.* Functional specificity of local synaptic connections in neocortical networks. *Nature* **473**, 87–91 (2011).
- Yassin, L. *et al.* An embedded subnetwork of highly active neurons in the neocortex. *Neuron* **68**, 1043–1050 (2010).
- 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).
- Wang, X.-J. Probabilistic decision making by slow reverberation in cortical circuits. *Neuron* **36**, 955–968 (2002).
- Renart, A., Moreno-Bote, R., Wang, X.-J. & Parga, N. Mean-driven and fluctuation-driven persistence in recurrent networks. *Neural Comput.* **19**, 1–46 (2007).
- Roudi, Y. & Latham, P.E. A balanced memory network. *PLoS Comput. Biol.* **3**, e141 (2007).
- Deco, G. & Hugues, E. Neural network mechanisms underlying stimulus driven variability reduction. *PLoS Comput. Biol.* **8**, e1002395 (2012).
- Albert, R. & Barabási, A.-L. Statistical mechanics of complex networks. *Rev. Mod. Phys.* **74**, 47–97 (2002).
- Boccaletti, S., Latora, V., Moreno, Y., Chavez, M. & Hwang, D.-U. Complex networks: structure and dynamics. *Phys. Rep.* **424**, 175–308 (2006).
- Oswald, A.-M.M., Doiron, B., Rinzel, J. & Reyes, A.D. Spatial profile and differential recruitment of GABAergic modulate oscillatory activity in auditory cortex. *J. Neurosci.* **29**, 10321–10334 (2009).
- Fino, E. & Yuste, R. Dense inhibitory connectivity in neocortex. *Neuron* **69**, 1188–1203 (2011).
- Hromádka, T., Deweese, M.R. & Zador, A.M. Sparse representation of sounds in the unanesthetized auditory cortex. *PLoS Biol.* **6**, e16 (2008).
- Tolhurst, D.J., Movshon, J.A. & Dean, A.F. The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision Res.* **23**, 775–785 (1983).
- Cohen, M.R. & Kohn, A. Measuring and interpreting neuronal correlations. *Nat. Neurosci.* **14**, 811–819 (2011).
- Brunel, N. & Wang, X.-J. What determines the frequency of fast network oscillations with irregular neural discharges? I. Synaptic dynamics and excitation-inhibition balance. *J. Neurophysiol.* **90**, 415–430 (2003).
- Teich, M.C., Heneghan, C., Lowen, S.B., Ozaki, T. & Kaplan, E. Fractal character of the neural spike train in the visual system of the cat. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* **14**, 529–546 (1997).
- Sompolinsky, H. & van Vreeswijk, C. Irregular activity in large networks of neurons. in *Les Houches Lectures LXXX: Methods and Models in Neurophysics* (eds. Chow, C.C., Gutkin, B., Hansel, D., Meunier, C. & Dalibard, J.) 341–402 (Elsevier, London, 2005).
- Hänggi, P., Talkner, P. & Borkovec, M. Reaction-rate theory: fifty years after Kramers. *Rev. Mod. Phys.* **62**, 251–341 (1990).
- Rajan, K., Abbott, L.F. & Sompolinsky, H. Stimulus-dependent suppression of chaos in recurrent neural networks. *Phys. Rev. E* **82**, 011903 (2010).
- Mitchell, J.F., Sundberg, K.A. & Reynolds, J.H. Differential attention-dependent response modulation across cell classes in macaque visual area V4. *Neuron* **55**, 131–141 (2007).
- Yoshimura, Y. & Callaway, E.M. Fine-scale specificity of cortical networks depends on inhibitory cell type and connectivity. *Nat. Neurosci.* **8**, 1552–1559 (2005).
- Hofer, S.B. *et al.* Differential connectivity and response dynamics of excitatory and inhibitory neurons in visual cortex. *Nat. Neurosci.* **14**, 1045–1052 (2011).
- Coombes, S. Waves, bumps, and patterns in neural field theories. *Biol. Cybern.* **93**, 91–108 (2005).
- Izhikevich, E.M., Gally, J.A. & Edelman, G.M. Spike-timing dynamics of neuronal groups. *Cereb. Cortex* **14**, 933–944 (2004).
- Abeles, M. *et al.* Cortical activity flips among quasi-stationary states. *Proc. Natl. Acad. Sci. USA* **92**, 8616–8620 (1995).
- Ecker, A.S. *et al.* Decorrelated neuronal firing in cortical microcircuits. *Science* **327**, 584–587 (2010).
- Morrison, A., Aertsen, A. & Diesmann, M. Spike-timing-dependent plasticity in balanced random networks. *Neural Comput.* **19**, 1437–1467 (2007).
- Gentet, L.J., Avermann, M., Matyas, F., Staiger, J.F. & Petersen, C.C.H. Membrane potential dynamics of GABAergic neurons in the barrel cortex of behaving mice. *Neuron* **65**, 422–435 (2010).
- Ferezou, I. *et al.* Spatiotemporal dynamics of cortical sensorimotor integration in behaving mice. *Neuron* **56**, 907–923 (2007).
- Han, F., Caporale, N. & Dan, Y. Reverberation of recent visual experience in spontaneous cortical waves. *Neuron* **60**, 321–327 (2008).
- Ma, W.J., Beck, J.M., Latham, P.E. & Pouget, A. Bayesian inference with probabilistic population codes. *Nat. Neurosci.* **9**, 1432–1438 (2006).
- Buesing, L., Bill, J., Nessler, B. & Maass, W. Neural dynamics as sampling: a model for stochastic computation in recurrent networks of spiking neurons. *PLoS Comput. Biol.* **7**, e1002211 (2011).

## ONLINE METHODS

**Spiking network simulations.** Neurons were modeled as leaky integrate-and-fire units whose voltages obey

$$\dot{V} = \frac{1}{\tau}(\mu - V) + I_{\text{syn}}$$

When neurons reach a threshold  $V_{\text{th}} = 1$ , a spike is emitted and they are reset to  $V_{\text{re}} = 0$  for an absolute refractory period of 5 ms. The membrane time constant  $\tau$  was 15 ms and 10 ms for excitatory and inhibitory neurons, respectively. The bias  $\mu$  was chosen according to a uniform random distribution between 1.1 and 1.2 for excitatory neurons and between 1 and 1.05 for inhibitory neurons. Although these values are suprathreshold, balanced dynamics ensured that the mean membrane potentials were subthreshold<sup>4</sup>. In **Figure 1**, the nondimensionalized voltages were transformed so that  $V_{\text{th}} = -50$  mV and  $V_{\text{re}} = -65$  mV.

Synapses between neurons were modeled as differences of exponentials, and the total synaptic input to neuron  $i$  in population  $x$  was:

$$I_{i,\text{syn}}^x(t) = \sum_{jy} J_{ij}^{xy} F^y * s_j^y(t)$$

where  $x, y \in \{E, I\}$  denote excitatory or inhibitory populations of  $N^E = 4,000$  and  $N^I = 1,000$  neurons each,  $J_{ij}^{xy}$  is the strength of synaptic connections from neuron  $j$  in population  $y$  to neuron  $i$  in population  $x$ ,  $F^y(t)$  is the synaptic filter for projections from neurons in population  $y$ ,  $*$  denotes convolution and  $s_j^y(t)$  is the spike train of neuron  $j$  in population  $y$ , a series of  $\delta$ -functions at time points where the neuron emits a spike.

$$F^y(t) = \frac{1}{\tau_2 - \tau_1} \left( e^{-t/\tau_2} - e^{-t/\tau_1} \right)$$

with  $\tau_2 = 3$  ms for excitatory synapses and 2 ms for inhibitory synapses, and  $\tau_1 = 1$  ms. Connection probabilities  $p^{xy}$  from neurons in population  $y$  to  $x$  were  $p^{EI} = p^{IE} = p^{II} = 0.5$ ;  $p^{EE}$  was on average 0.2 for all networks but depended on clustering as described in the main text. If a connection from neuron  $j$  in population  $y$  to neuron  $i$  in population  $x$  existed,  $J_{ij}^{xy} = J^{xy}$  (unless  $x, y = E$  and neurons were in the same cluster; then connection strength was multiplied by 1.9 for homogeneous clustered networks and 1.8 for heterogeneous clustered networks); otherwise  $J_{ij}^{xy} = 0$ . Synaptic strengths were  $J^{EE} = 0.024$ ,  $J^{EI} = -0.045$ ,  $J^{IE} = 0.014$  and  $J^{II} = -0.057$ . These parameters, multiplied by 15 mV, would give the deflection of the membrane potential of the postsynaptic target, neglecting leak, in our dimensionalized units. When clusters of excitatory neurons were stimulated (**Fig. 7**), stimulation was accomplished by increasing  $\mu$  for neurons in those clusters by 0.07. Simulations were performed using Euler integration with a time step of 0.1 ms.

To generate an imbalanced network (**Fig. 2**, dashed curve), we assumed weak synapses by reducing  $J^{EE}$  and  $J^{IE}$  by a factor of 20 and chose  $\mu^E$  uniformly between 1.05 and 1.15. Consequently, excitatory neurons were mean-driven rather than fluctuation-driven.

Finally, to vary the size of the network (**Fig. 4e**), we kept all network parameters the same except that we scaled the connection strengths  $J^{xy}$  by a factor proportional to  $1/\sqrt{N}$ , so that the  $J^{xy}$  values were multiplied by  $\sqrt{2}$  for the network with  $N = 2,500$  and  $1/\sqrt{2}$  for  $N = 10,000$ .

**Visualizing and measuring clustering.** Visualizations of connectivity for subsets of the excitatory networks in **Figures 1** and **4** were created using NetworkX (<http://networkx.lanl.gov/>). The layout of nodes was performed using the Fruchterman-Reingold force-directed algorithm<sup>51</sup>, which places connected nodes nearby and yields layouts that respect the clustering of the network. For the purposes of these analyses, connections were taken to be undirected and unweighted.

**Spike train statistics.** Spike train statistics were computed for excitatory neurons. We denote the spike times of neuron  $i$  as  $\{t_{i1}, t_{i2}, t_{i3}, \dots\}$ . We can then define neuron  $i$ 's spike train:  $y_i(t) = \sum_k \delta(t - t_{ik})$ . The number of spikes emitted by the neuron between times  $t$  and  $t + \Delta t$  is

$$N_i(t, t + \Delta t) = \int_t^{t+\Delta t} y_i(t') dt'$$

The firing rate of a neuron over an interval  $(t, t + \Delta t)$  is defined as

$$r_i(t, t + \Delta t) = \frac{1}{\Delta t} N_i(t, t + \Delta t)$$

For the networks studied, firing rates and other statistics for the spontaneous state were calculated with  $t = 1.5$  s to prevent effects due to initial conditions and  $\Delta t = 1.5$  s.

We also computed the Fano factor  $F_i(t, t + \Delta t)$  for neuron  $i$  by evaluating

$$F_i(t, t + \Delta t) = \frac{\text{Var}(N_i(t, t + \Delta t))}{\langle N_i(t, t + \Delta t) \rangle}$$

where the expectations are over repeated trials of the same network with random initial conditions. When computing the Fano factor as a function of time relative to stimulus onset, we computed the mean-matched Fano factor described in ref. 9 to control for changes in firing rate. Fano factors were computed over 100-ms windows.

We computed correlation coefficients for the spike counts of neuron pairs. The correlation between neurons  $i$  and  $j$  was given by

$$\rho_{ij} = \frac{\text{Cov}(N_i(t, t + \Delta t), N_j(t, t + \Delta t))}{\sqrt{\text{Var}(N_i(t, t + \Delta t)) \text{Var}(N_j(t, t + \Delta t))}}$$

where the covariances and variances were computed over overlapping windows within each trial and then averaged across trials.

Finally, we computed the mean-subtracted spike train autocorrelation and cross-correlation functions  $C_{ii}(T)$  and  $C_{i \neq j}(T)$ . These were defined as

$$C_{ij}(T) = \int_t^{t+\Delta t} y_i(t') y_j(t' - T) dt' - r_i(t, t + \Delta t) r_j(t, t + \Delta t)$$

To estimate this quantity, spike trains were discretized with a time step of  $\delta t = 2$  ms, so that  $y_i(t) \rightarrow y_i(t_n)$  with  $t_n = n\delta t$  and  $y_i(t_n) = N_i(t_n, t_n + \delta t)$ . Then  $C_{ij}(T) = \sum_n y_i(t_n) y_j(t_n - T) - r_i(t, t + \Delta t) r_j(t, t + \Delta t)$  with  $T = \{\dots, -\delta t, 0, \delta t, \dots\}$  and  $n$  such that  $t_n \in (t, t + \Delta t)$ .

**Analysis of simplified network model.** For our binary network model, the activity of neuron  $i$  in population  $x$  was given by  $s_i^x(t) = \Theta(I_i^x(t)) \in \{0, 1\}$ , where  $I_i^x(t)$  is the input to the neuron and  $\Theta$  is the Heaviside step function. For simplicity, excitatory and inhibitory populations both consisted of  $N^E = N^I = 4,000$  neurons with an average connection probability of 0.2 for all connection types, so that neurons received  $K = 800$  connections on average from each population. The input was given by

$$I_i^x(t) = \sum_{jy} J_{ij}^{xy} s_j^y(t) + \mu^x$$

where  $\mu^E = 0.075\sqrt{K}$  and  $\mu^I = 0.05\sqrt{K}$ . The presence of synaptic connections, including the effects of clustering, was determined in the same way as in the spiking networks, although for our initial analysis only one cluster was generated in the network.

In balanced networks, it is assumed that the synaptic strengths scale as  $J \sim 1/\sqrt{K}$  so that the input variance remains constant as  $K$  grows<sup>4</sup>. In our network, the synaptic strengths were  $J^{EE} = 1/\sqrt{K}$  (unless neurons belonged to the same cluster, in which case connection strength was multiplied by 1.5),  $J^{EI} = -1.2/\sqrt{K}$ ,  $J^{IE} = -1/\sqrt{K}$  and  $J^{II} = -1/\sqrt{K}$ . Simulations were performed using discrete time steps, on each of which  $N^E + N^I$  neurons were chosen randomly and updated asynchronously.

For uniform networks, the average activity of neurons in each population,  $\langle s^E \rangle$  and  $\langle s^I \rangle$ , can be calculated in the infinite  $K$  limit<sup>4</sup>:

$$\langle s^E \rangle = \frac{J^{II}\mu^E - J^{EI}\mu^I}{\sqrt{K}(J^{EI} - J^{II})}, \quad \langle s^I \rangle = \frac{\mu^E - \mu^I}{K(J^{EI} - J^{II})} \quad (6)$$

Furthermore, by assuming that each neuron is independent and is active randomly and sparsely, the input variances can be calculated:

$$\begin{aligned} \text{Var}(I^E) &= (J^{EE})^2 K \langle s^E \rangle + (J^{EI})^2 K \langle s^I \rangle \\ \text{Var}(I^I) &= (J^{IE})^2 K \langle s^E \rangle + (J^{II})^2 K \langle s^I \rangle \end{aligned}$$



Assuming the input distribution is Gaussian, we have

$$\langle s^x \rangle = \frac{1}{2} \operatorname{erfc} \frac{-\langle I^x \rangle}{\sqrt{2 \operatorname{Var}(I^x)}}$$

which allows us to infer the mean input  $\langle I^x \rangle = -\sqrt{2 \operatorname{Var}(I^x)} \operatorname{erfc}^{-1}(2 \langle s^x \rangle)$ .

To analyze the effect of clustering, we calculate the deviation from the mean input in a uniform network induced by the presence of clustered connections. This approach has been taken in previous studies of memory networks<sup>23,24,35</sup>. In a uniform network, the mean input is given by

$$\langle I^E \rangle = J^{EE} K \langle s^E \rangle + J^{EI} K \langle s^I \rangle + \mu^E$$

For a clustered network, we assume that the mean firing rate of neurons outside the cluster is given by  $\langle s_{\text{out}}^E \rangle = \langle s^E \rangle$ , whereas inside the cluster  $\langle s_{\text{in}}^E \rangle$  may vary. Hence, the mean input to neurons in the cluster is

$$\langle I_{\text{in}}^E \rangle = J_{\text{in}}^{EE} K_{\text{in}}^{EE} \langle s_{\text{in}}^E \rangle + J_{\text{out}}^{EE} K_{\text{out}}^{EE} \langle s^E \rangle + J^{EI} K \langle s^I \rangle + \mu^E \quad (7)$$

where  $K_{\text{in}}^{EE}$  denotes the average number of connections from inside the cluster. With  $\mu \sim \sqrt{K}$ , the final three terms in equation (7) are  $O(\sqrt{K})$  but are balanced by the recurrent network dynamics, ensuring that their contribution does not diverge as  $K$  becomes large.

We next calculate  $\Delta I_{\text{in}}^E = \langle I_{\text{in}}^E \rangle - \langle I^E \rangle$ :

$$\Delta I_{\text{in}}^E = K_{\text{in}}^{EE} (J_{\text{in}}^{EE} \langle s_{\text{in}}^E \rangle - J_{\text{out}}^{EE} \langle s^E \rangle)$$

For cluster activity to have an effect on the mean input a neuron receives, this difference must be  $O(1)$ , which implies that  $K_{\text{in}}^{EE} J_{\text{in}}^{EE} \sim O(1)$ . As  $J_{\text{in}}^{EE} \sim 1/\sqrt{K}$ ,

this implies that  $K_{\text{in}}^{EE} \sim \sqrt{K}$ , so that the number of local cluster inputs scales as the square root of the total number of connections. In particular, for large  $K$ ,  $K_{\text{in}}^{EE} \ll K$ .

We next analyze the time dynamics of the average cluster activity. The dynamics will follow<sup>52</sup>

$$\left\langle \dot{s}_{\text{in}}^E \right\rangle = -\langle s_{\text{in}}^E \rangle + \frac{1}{2} \operatorname{erfc} \frac{-\langle I^E \rangle + \Delta I_{\text{in}}^E}{\sqrt{2 \operatorname{Var}(I^E)}} \equiv -\frac{d}{d \langle s_{\text{in}}^E \rangle} U(\langle s_{\text{in}}^E \rangle)$$

where we have defined a potential  $U$  as the integral of  $\left\langle \dot{s}_{\text{in}}^E \right\rangle$  with respect to  $\langle s_{\text{in}}^E \rangle$ . Minima of this potential correspond to self-consistent solutions of average cluster activity. This potential can be obtained using the equations defined

above (for **Fig. 5b**, empirically measured values of  $\langle s^E \rangle$  and  $\langle s^I \rangle$  were used

in the calculations, to correct for deviations from equation (6) due to finite size effects).

Assuming that the fluctuations in  $\langle s_{\text{in}}^E \rangle$  can be modeled as Gaussian white noise and are small relative to well depth  $\Delta U$ , the average transition time  $T$  out of a potential well should follow the Kramers formula<sup>36</sup>  $T \propto e^{k \Delta U / D}$ , where  $D$  is the diffusion coefficient governing the fluctuations of the mean activity. While we can calculate  $\Delta U$  using our theory above, the effective  $D$  depends on system size and correlations and cannot be easily evaluated. Nonetheless, we find a relationship between  $\Delta U$  and  $T$  consistent with the Kramers formula (**Fig. 5d**).

51. Fruchterman, T.M.J. & Reingold, E.M. Graph drawing by force-directed placement. *Softw. Pract. Exp.* **21**, 1129–1164 (1991).

52. Glauber, R. Time-dependent statistics of the Ising model. *J. Math. Phys.* **4**, 294–307 (1963).