Specific excitatory connectivity for feature integration in mouse primary visual cortex

Authors: Dylan R Muir[1,3], Patricia Molina-Luna[2], Morgane M Roth[1,2], Fritjof Helmchen[2] and Björn M Kampa[2,3,4]

1. Biozentrum, University of Basel, Klingelbergstrasse 50/70, 4056 Basel, Switzerland.

2. Laboratory of Neural Circuit Dynamics, Brain Research Institute, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland.

3. Department of Neurophysiology, Institute of Biology 2, RWTH Aachen University, Templergraben 55, 52062 Aachen, Germany.

4. JARA-BRAIN, 52074 Aachen, Germany.

Running title: Non-random cortical connectivity and feature integration

Major subject area: Neuroscience

Keywords: amplification; competition; decorrelation; cortical computation; mouse V1

Corresponding author: Dylan Muir, Biozentrum, University of Basel, Klingelbergstrasse 50/70, 4056 Basel, Switzerland.
Phone: +41 61 207 17 75
Fax: +41 61 207 21 89
Email: dylan.muir@unibas.ch

Abstract

In mouse visual cortex (V1) many neurons respond to overlapping grating stimuli (plaid stimuli) with highly selective and facilitatory responses, which are not simply predicted by responses to single gratings presented alone. This complexity is surprising since excitatory neurons in V1 are considered to be mainly tuned to single preferred orientations. We hypothesised that complex responses to plaid stimuli may arise as a consequence of rules for selective excitatory connectivity within the local network in the superficial layers of mouse V1. Although local excitatory connections have been found to be more prevalent and stronger between neurons that share similar functional response features, the extent to which local connectivity rules shape neuronal responses in V1 and the details of how network structure relates to functional responses remain unknown. Here we examined two possible alternative connectivity schemes: in the first case, *local connections are aligned with visual properties inherited from feedforward input* (a 'like-to-like' scheme specifically connecting neurons that share similar preferred orientations); in the second case, *local connections group neurons into excitatory subnetworks that combine and amplify multiple feedforward visual properties (a 'feature binding' scheme)*. By comparing predictions from large scale computational models with *in vivo* recordings of visual representations in mouse V1, we found that responses to plaid stimuli were best explained by a assuming 'feature binding' connectivity. Unlike under the 'like-to-like' scheme, selective amplification within feature-binding excitatory subnetworks replicated experimentally observed facilitatory responses to plaid stimuli; explained selective plaid responses not predicted by grating selectivity; and was consistent with broad anatomical selectivity observed in mouse V1. Our results show that visual feature binding can occur through local recurrent mechanisms without requiring feedforward convergence, and that such a mechanism is consistent with visual responses in mouse V1.

Author summary

The brain is a highly complex structure, with abundant connectivity between nearby neurons in the neocortex, the outermost and evolutionarily most recent part of the brain. Although the network architecture of the neocortex can appear disordered, connections between neurons seem to follow certain rules. These rules most likely determine how information flows through the neural circuits of the brain, but the relationship between particular connectivity rules and the function of the cortical network is not known. We built models of visual cortex in the mouse, assuming distinct rules for connectivity, and examined how the various rules changed the way the models responded to visual stimuli. We also recorded responses to visual stimuli of populations of neurons in anaesthetised mice, and compared these responses with our model predictions. We found that connections in neocortex probably follow a connectivity rule that groups together neurons that differ in simple visual properties, to build more complex representations of visual stimuli. This finding is surprising because primary visual cortex is assumed to support mainly simple visual representations. We show that including specific rules for non-random connectivity in models of the brain is crucial. Precise measurements of those rules in real brains is also essential.

Introduction

Much of our current understanding of local cortical connectivity in neuronal circuits of the neocortex is based on the presumption of randomness. Anatomical methods for estimating connection probabilities [1,2] and techniques for using anatomical reconstructions to build models of cortical circuits [3-7] are largely based on the assumption that connections between nearby neurons are made stochastically in proportion to the overlap between axonal and dendritic arborisations [8].

On the other hand, a wealth of evidence spanning many cortical areas and several species indicates that cortical connectivity is not entirely random. In species that display smooth functional maps in primary visual cortex (V1), such as cat and macaque monkey, long-range intrinsic excitatory connections tend to preferen-

tially connect regions of similar function [9-13]. Although rodents exhibit a mapless, "salt and pepper" representation of basic visual features across V1 [14], non-random connectivity is nonetheless prevalent both within and between cortical layers [15-20], reflecting similarities in functional properties [21-25] or projection targets [26-28].

The degree of specificity of cortical connections among excitatory neurons thus appears to be an important feature of local circuitry, which likely influences the functional response properties of cortical neurons [24,29]. In spite of its relevance, the impact of specific excitatory connectivity on network representations of sensory inputs and information processing has not been addressed experimentally or through theory; it therefore remains an open question how the arrangement of local recurrent connections affects cortical representations of visual stimuli. Despite multiple descriptions of specific connectivity in cortex, the rules underlying the configuration of these connections are not entirely clear. Whereas strong connections are more prevalent between neurons with similar receptive fields, the majority of synaptic connections are made between neurons with poorly-correlated receptive fields and poorly correlated responses [24]. This sea of weak synaptic inputs might be responsible for non-feature-specific depolarisation [24] or might permit plasticity of network function [20]. However, another possibility is that weaker connections underlie higher-order connectivity rules that have not yet been described. For example, we recently showed that neuronal responses to compound visual stimuli (e.g. plaid stimuli composed of two grating components) can be selective and highly complex in mouse V1 [30]. We proposed that local connectivity in V1 could be structured such that simple visual features were bound together through recurrent interactions [30]. But what is the exact relationship between connectivity rules and information processing in visual cortex? And how do the rules governing specific excitatory connectivity, which we know to exist in rodents [15,21,24], shape and inform visual representations?

Here we examined these questions by analysing the computational properties of cortical networks with defined rules for local connectivity. We simulated visual responses to grating and plaid stimuli in large networks with properties designed

to resemble the superficial layers of mouse V1, assuming distinct connectivity schemes. We then compared the response patterns and visual representations predicted by the network simulations with those recorded *in vivo* in mouse V1.

Specifically, we evaluated two broad classes of connectivity patterns, where specific local excitatory connectivity is defined according to the visual response properties of neurons (Fig. 1):

1. Strictly "like-to-like" connectivity, such that neurons with similar response properties defined by their feed-forward inputs to each neurons (e.g. orientation tuning of neurons in the superficial layers, arising from tuned input from layer 4) are grouped into subnetworks;
2. A form of "feature-binding" connectivity, such that excitatory neurons with differing feed-forward visual properties (e.g. distinct orientation preference) are grouped together.

**** FIGURE 1 NEAR HERE ****

These distinct rules give rise to radically different visual representations of plaid stimuli, both in terms of complexity of visual response selectivity of individual neurons and regarding facilitation versus suppression in response to these compound stimuli. We found that the complexity of plaid responses in mouse V1 was reproduced in our simulations when assuming the 'feature-binding' connectivity scheme, with local connections cutting across feedforward response properties, but not when assuming purely 'like-to-like' connections.

Our results therefore suggest that local excitatory connections within mouse V1 are formed with respect to complex or compound visual response properties, such that they do not necessarily align with simpler feedforward properties. This pattern of connectivity would allow subnetworks in V1 to detect particular configurations of visual stimuli, and might be used to tune visual cortex to the complex statistics of natural vision.

Results

Responses to plaid stimuli are selective and facilitatory in mouse V1

Under the assumption that the configuration of local recurrent connections in cortex might lead to differential processing of simple and compound visual stimuli, it is important to quantify the relationship between responses to grating and plaid stimuli in visual cortex. Recent reports have highlighted the facilitatory and selective nature of plaid responses in mouse primary visual cortex [30-32]. Most neurons in mouse V1 respond to plaid stimuli in accordance with a simple superimposition of their responses to the two underlying grating components (i.e. "component cell" responses; [33]). However, a significant proportion of neurons that are visually responsive, reliable and selective exhibit complex responses to plaid stimuli that are difficult to explain with respect to simple combinations of grating components [30]. Plaid stimuli are often constructed from a single choice of relative component angle (90° orthogonal gratings), leaving open the possibility that a richer set of plaid stimuli would help to classify neurons with these complex responses.

We therefore probed mouse V1 with grating component stimuli composed of grating stimuli with 16 drift directions, and three full sets of plaid stimuli composed of 45°, 90° and 135° relative grating component orientations. We recorded responses from layer 2/3 neurons in V1 using two-photon imaging of animals with viral delivery of GCaMP6m (Fig. 2a–f; 8 animals, 8 sessions, 441/879 responsive/imaged neurons; see Methods). We defined a modulation index (MI) to quantify the degree of facilitation or suppression elicited by plaid stimuli over grating stimuli, for single cortical neurons; large positive values for MI indicate strong facilitation in response to plaid stimuli, whereas large negative values indicate strong suppression (see Methods). Visual responses to the full set of plaid stimuli were dominated by facilitation, and were significantly more facilitatory than when considering only the set of 90° plaids (Fig. 2i; median modulation

index MI 0.098 ± [0.081 0.12] vs 0.011 ± [-0.0060 0.027]; $p < 1\times10^{-10}$, Wilcoxon rank-sum; all following values are reported as median ± 95% bootstrap confidence intervals unless stated otherwise).

**** FIGURE 2 NEAR HERE ****

The presence of stronger facilitation when comparing responses to the full set of plaid stimuli with responses to 90° plaids alone, is consistent with our earlier finding that some neurons in mouse V1 are highly selective for particular combinations of grating components [30]. Accordingly, we used a plaid selectivity index (PSI) to quantify how selective were the responses of single neurons over the set of plaid stimuli (see Methods). The PSI was defined in analogy to orientation or direction selectivity indices (OSI or DSI), such that values of PSI close to 1 indicate that a neuron responds to only a single plaid stimulus out of the set of presented plaid stimuli. Values of PSI close to 0 indicate that a neuron responds equally to all plaid stimuli. Responses to the full set of plaid stimuli were highly selective; significantly more selective than predicted by a component model generated using all plaid and grating stimuli (Fig. 2h; median PSI 0.38 ± [0.36 0.41] vs 0.30 ± [0.28 0.31]; $p < 1\times10^{-10}$, Wilcoxon rank-sum) and indeed significantly more selective than responses to the 90° plaids alone (Fig. 2h; median 90° PSI 0.25 ± [0.23 0.28]; $p < 1\times10^{-10}$ vs all plaids, Wilcoxon rank-sum).

Therefore, probing visual cortex with a dense set of plaid stimuli reveals richer visual responses than when probed with a set of only 90° plaids. Indeed, recent results suggest that using an expanded set of plaid stimuli evokes more pattern-cell responses in mouse V1 [32]. Consistent with this finding, our results show that using a dense set of plaids does not make responses to compound stimuli trivial to predict based on component responses. In addition, we found that visual responses were more facilitatory and more selective than when measured using 90° plaids alone.

Local excitatory connections in cortex are broadly selective for preferred orientation

How are selective responses to plaid stimuli generated in V1? As we suggested previously, one possibility is that specific grating component representations are combined through local excitatory connectivity [30]. Synaptic connection probability is enhanced by similarity of orientation preference [21,23,25], suggesting that local excitatory connections may group together neurons with common preferred orientations. Connection probability is even more strongly modulated by neuronal response correlations to natural movies; i.e., the likelihood for a synaptic connection is higher for neuronal pairs responding similarly to natural scenes [21,22,24].

If connections in mouse V1 were strictly governed by preferred orientation, then neurons with similar orientation preference should also predominately have similar responses to natural movies, and vice versa. We recorded visual responses populations of neurons labelled with the synthetic calcium indicator OGB in anaesthetized mouse V1 (5 animals, 129/391 responsive neurons with overlapping receptive fields/imaged neurons; Fig. 3a–c; see Methods). We used signal correlations to measure the similarity between the responses of pairs of neurons with identified receptive fields (Fig. 3a) to drifting grating (Fig. 3b) and natural movie (Fig. 3c) visual stimuli (see Methods).

We found that neuronal pairs with high signal correlations to natural scenes, which are most likely to be connected in cortex [21,22,24], showed only a weak tendency to share similar orientation preferences (Fig. 3d–e; pairs with OSI > 0.3; p=0.8, Kruskall-Wallis). This is consistent with earlier findings in cat area 17 (V1), which showed a poor relationship between responses to gratings and natural movies [34].

**** FIGURE 3 NEAR HERE ****

Similarly, under a like-to-like connectivity rule, synaptically connected neurons in mouse V1 should share both similar orientation preference and responsiveness to natural movies. We therefore compared response correlations and preferred orientations for pairs of mouse V1 neurons, which were known to be connected from

in vivo/in vitro characterisation of functional properties and connectivity (data from [24] used with permission; 17 animals, 203 patched and imaged cells, 75 connected pairs). Consistent with our results comparing responses to gratings and natural movies, connected pairs of cells with similar orientation preference were not more likely to share a high signal correlation to flashed natural scenes (Fig. 3f; p=0.54, Kruskall-Wallis). Also consistent with earlier findings [21,23], we observed a positive relationship between synaptic connectivity and similarity of orientation preference (Fig. 3g; p=0.045, Ansari-Bradley test). However, strongly connected pairs (strongest 50% of excitatory post-synaptic potentials — EPSPs — over connected pairs) were not more similar in their preferred orientation than the remaining pairs (p=0.17, Ansari-Bradley test vs weakest 50% of connected pairs). Connected pairs spanned a wide bandwidth of preferred orientations, with more than 20% of connections formed between neurons with orthogonal preferred orientations. Spatial correlation of receptive fields is a comparatively better predictor for synaptic connectivity than shared orientation preference, but a majority of synaptic inputs are nevertheless formed between neurons with poorly- or un-correlated responses [24]. We conclude that similarity in orientation preference only partially determines connection probability and strength between pairs of neurons in mouse V1.

This weak functional specificity for similar visual properties can be explained by two possible alternative connectivity rules. In the first scenario, local excitatory connections in cortex are aligned with feedforward visual properties, but with broad tuning (Fig. 1a; a "like-to-like" rule). As a consequence, all connections show an identical weak bias to be formed between neurons within similar tuning, and the average functional specificity reported in Fig. 3g and elsewhere [21,24] reflects the true connection rules between any pair of neurons in cortex.

Alternatively, local excitatory connections may be highly specifically tuned but follow rules that are not aligned with feedforward visual properties (Fig. 1b; a "feature-binding" rule). If measurements of functional specificity were made pair-wise and averaged across a large population, any specific tuning shared within groups of neurons would therefore be averaged away and appear as a sea of random connections. For example, subpopulations of excitatory neurons might share a small

set of feedforward visual properties; in this case, connections within a subpopulation could still be highly specific, but this specificity would not be detected through purely pairwise measurements.

Specific connectivity gives rise to amplification and competition

The dynamics of neuronal networks defined with particular connectivity rules remains generally unknown, although some results suggest that specific connectivity leads to reduced dimensionality of network activity patterns [35]. We therefore explored the relationship between specific connectivity and network dynamical properties in a non-linear, rate-based network model incorporating realistic estimates for recurrent excitatory and inhibitory connection strength in layers 2/3 of mouse V1 ("analytical model"; Fig. 4; see Methods). This small model consists of four excitatory and one inhibitory neuron with homogenous connectivity (Fig. 4a), designed to be equivalent to a much larger model with stochastic synaptic connectivity. As is suggested by estimations of strong excitatory feedback in cortex [4,36], our model required inhibitory feedback to maintain stability (an inhibition-stabilised network or ISN; Fig. S1; [37-40]; but see [41]).

**** FIGURE 4 NEAR HERE ****

Non-random connectivity, in the form of specific excitatory connections within subnetworks (Fig. 4b; SNs; [15,18]), introduced selective amplification within subnetworks and competition between subnetworks (Fig. 4c). Surprisingly, these computational mechanisms were strongly expressed even when only a minority of synapses (s around 20%) were made to be subnetwork-specific (Fig. 4c; Fig. S1). Specific connectivity rules resulted in functional grouping of sets of excitatory neurons (Fig. 4b), permitting the network to operate in a soft winner-take-all regime [42,43].

Neither competition nor amplification was present under parameters designed to approximate random connectivity in mouse V1 (Fig. 4a, c; Fig. S1). This is not because the network architecture was incapable of expressing competition, but because the recurrent excitatory connections were insufficiently strong under

assumptions of random stochastic connectivity. We conclude that specific excitatory connectivity strongly promotes amplification and competition in neuronal responses.

Selective amplification under like-to-like and feature-binding connectivity rules

Amplification in the network with specific connectivity is *selective* (Fig. 4b–c): neurons within a subnetwork recurrently support each other's activity, while neurons in different subnetworks compete. Therefore, which sets of neurons will be amplified or will compete during visual processing depends strongly on the precise rules used to group neurons into subnetworks. We therefore examined the impact of "like-to-like" and "feature-binding" rules on responses in our analytical model. The excitatory network was partitioned into two subnetworks; connections within a subnetwork corresponded to selective local excitatory connectivity within rodent V1. Under the "like-to-like" rule, neurons with similar orientation preferences were grouped into subnetworks (Fig. 4d). We tested the response of this network architecture to simulated grating and plaid stimuli, by injecting currents into neurons according to the similarity between the orientation preference of each neuron and the orientation content of a stimulus. Under the "like-to-like" rule, responses of pairs of neurons to simple grating stimuli and more complex plaid stimuli were highly similar (Fig. 4d). Amplification occurred within subnetworks of neurons with the same preferred orientation, and competition between subnetworks with differing preferred orientation [42,44] (visible by complete suppression of response of neurons in lower traces of Fig. 4d).

Alternatively, we configured the network such that the rules for local excitatory connectivity did not align with feedforward visual properties (a "feature-binding" rule). We configured subnetworks by grouping neurons showing preference for either of two specific orientations (Fig. 4e). When this "feature-binding" connectivity rule was applied, neuronal responses to grating and plaid stimuli differed markedly (cf. top vs bottom panels of Fig. 4e). Selective amplification was now arrayed within populations of neurons spanning differing orientation preferences,

and competition occurred between subnetworks with different compound feature preferences. Importantly, a "feature-binding" rule implies that neurons with the same preferred orientation could exist in competing subnetworks. While their responses to a simple grating of the preferred orientation would be similar and correlated (Fig. 4e; indicated by a high response correlation measured over grating responses ρ_g), the same two neurons would show decorrelated responses to a plaid stimulus (Fig. 4e; indicated by a low response correlation measured over plaid responses ρ_p). We conclude that changes in pairwise response similarity, provoked by varying the inputs to a network, can provide information about the connectivity rules present in the network.

Feature-binding connectivity leads to facilitation and decorrelation in large networks

The results of our simulations of small networks suggest that rules for specific local connectivity can modify the correlation of activity between two neurons in a network, depending on the input to the network. The question follows how connectivity rules shape distributed representations of visual stimuli, examined across a large network and over a broad set of stimuli.

We therefore simulated the presentation of grating and plaid visual stimuli in a large-scale non-linear, rate-based model of the superficial layers of mouse V1, consisting of 80,000 neurons (of which approximately 20% were inhibitory; [45,46]; see Table 1 for all parameters used in these models). Non-spiking linear-threshold neuron models provide a good approximation to the F–I curves of adapted cortical neurons [47]; model neurons with linear-threshold dynamics can be directly translated into integrate-and-fire models with more complex dynamics [48,49], and in addition form good approximations to conductance-based neuron models [50].

Our model included realistic estimates for connection strength and connection sparsity in mouse V1, and a random salt-and-pepper arrangement of orientation preference as reported for rodent V1 [14]. We defined connection rules for sparse stochastic connectivity based primarily on overlap of dendritic and axonal fields,

modulated by connectivity rules designed to test the difference between "like-to-like" and "feature-binding" schemes. We quantified response similarity between pairs of neurons as suggested by the results of the small network simulations: by measuring response similarity over a set of grating stimuli (ρ_g), and separately over a set of plaid stimuli (ρ_p; see Methods).

In the network that implemented a "like-to-like" connection rule for recurrent excitatory connectivity (Fig. 5a–b), pairs of neurons showed similar responses to both grating and plaid stimuli (Fig. 5b; R^2=0.83 between ρ_g and ρ_p), in agreement with the analytical "like-to-like" model of Fig. 4d.

**** FIGURE 5 NEAR HERE ****

However, in the network that implemented a "feature-binding" connection rule, where in addition to spatial proximity and similarity in preferred orientation sub-networks were defined to group neurons of two distinct preferred orientations (Fig. 5c–d), neurons showed decorrelation in response to plaid stimuli (Fig. 5d, R^2=0.13 between ρ_g and ρ_p), in agreement with the analytical "feature-binding" model of Fig. 4e. Different configurations of local recurrent excitatory connectivity produced by "like-to-like" or "feature-binding" rules can therefore be detected in large networks, by comparing responses to simple and compound stimuli.

Consistent with our analytical models, networks including only random excitatory connectivity without any specificity did not give rise to decorrelation (Fig. S2b; R^2=0.72 between ρ_g and ρ_p). This shows that decorrelation between plaid and grating responses in our models does not arise simply due to random connectivity, but requires the active mechanism of selective amplification through feature-binding subnetwork connectivity.

Inhibitory responses were untuned in our simulations (blue traces in Fig. 5a, c), in agreement with experimental observations of poorly-tuned inhibition in mouse V1 [46,51-53].

Visual responses in mouse V1 are consistent with "feature-binding" connection rules

Our analytical network results show that in principle the configuration of local excitatory connectivity, whether aligned with or spanning across feedforward visual properties, has a strong effect on visual representations (Fig. 4). Our large-scale simulations show that these effects can be detected in large networks as differences in the pairwise correlations of responses to simple and compound visual stimuli (Fig. 5). We therefore aimed to test which connectivity scheme is more likely to be present in visual cortex, by examining responses of neurons in mouse V1.

Using two-photon calcium imaging, we recorded responses of populations of OGB-labelled neurons in mouse V1 to a set of contrast-oscillating oriented grating stimuli over a range of orientations, as well as the responses to the set of plaid stimuli composed of every possible pair-wise combination of the oriented grating stimuli (Fig. 6a; 5 animals, 5 sessions, 313/543 responsive/imaged neurons; see Methods). Responses to plaid stimuli in mouse V1 suggest that stimulating with a denser sampling of compound stimulus space leads to a better characterisation of response selectivity (Fig. 2). Accordingly, we probed responses in mouse V1 under stimuli analogous to those used in the model simulations, with a dense coverage of plaid combinations over a set of finely-varying grating orientations.

**** FIGURE 6 NEAR HERE ****

We found that consistent with our earlier findings examining 90° drifting plaid stimuli [30], responses to grating stimuli did not well predict responses to plaid stimuli. Pairs of neurons with similar preferred orientation but with highly differing responses to plaid stimuli were common (Fig. 6b–c; R^2=0.05 between ρ_g and ρ_p; OSI > 0.3). The degree of decorrelation we observed in mouse V1 was considerably higher than predicted by the "like-to-like" model, and was more consistent with the "feature-binding" model (Fig. 6e).

Decorrelation induced by plaid responses and the lack of a relationship between grating and plaid responses in mouse V1 were not a result of unreliable or noisy responses *in vivo*. We included in our analysis only neurons that were highly reliable, and responded significantly more strongly than the surrounding neuropil (see Methods). As a further control, we used experimentally recorded responses to grating stimuli to generate synthetic plaid responses for mouse V1 that would result from a cortex with like-to-like subnetwork connectivity (Fig. 6d, inset; see Methods). Our control data were generated from single-trial responses of single V1 neurons, and therefore included the same trial-to-trial variability exhibited by cortex. This control analysis indicates that a "like-to-like" rule among V1 neurons would result in a higher correlation of grating and plaid responses than experimentally observed (Fig. 6d; median $R^2 = 0.77 \pm [0.767\ 0.775]$ between ρ_g and ρ_p; $n = 2000$ bootstrap samples; compared with $R^2 = 0.05$ for experimental results; $p < 0.005$, Monte-Carlo test).

Importantly, this control analysis is not restricted to our "like-to-like" rule, but makes similar predictions of highly correlated grating and plaid responses for any arbitrary model that combines grating components to produce a plaid response, as long as that rule is identical for every neuron in the network [30]. This is because if a single consistently-applied rule exists, then any pair of neurons with similar grating responses (high ρ_g) will also exhibit similar plaid responses (high ρ_p). In contrast, neurons that are connected within the "feature-binding" model combine different sets of grating components, depending on which subnetwork the neurons are members of.

Neurons in mouse V1 exhibited a wide range of facilitatory and suppressive responses to plaid stimuli, roughly equally split between facilitation and suppression (Fig. 6f–g; 45% vs 42%; $MI > 0.05$ and $MI < -0.05$). The proportion of facilitating and suppressing neurons in mouse V1 was similar to that exhibited by responsive neurons in our "feature-binding" model (Fig. 6g; V1 versus F.B., $p = 0.17$; two-tailed Fisher's exact test, $n_{V1} = 313$, $n_{F.B.} = 809$). In contrast, neither the "like-to-like" model (L-to-L) nor a model with random non-specific connectivity (Rnd) exhibited significant facilitation in responsive neurons, and both were significantly different from the distribution of facilitation and suppression in mouse V1

(Fig. 6g; $p < 0.001$ in both cases; two-tailed Fisher's exact test, $n_{L\text{-to-}L} = 729$, $n_{Rnd} = 729$). The wide range of facilitatory and suppressive responses observed in mouse V1 is more consistent with a feature-binding rule for local connectivity, compared with a like-to-like rule or random non-specific connectivity.

Discussion

Whereas feedforward mechanisms for building response properties in visual networks have been extensively studied, it is not well understood how visual responses are shaped by local recurrent connections. We hypothesised that the configuration of local recurrent cortical connectivity shapes responses to visual stimuli in mouse V1, and examined two alternative scenarios for local connection rules: essentially, whether local excitatory connections are made in accordance with feedforward visual properties ("like-to-like"; Fig. 1a), or whether local excitatory connections span across feedforward visual properties to group them ("feature-binding"; Fig. 1b). We found that highly selective and facilitatory responses to plaid stimuli observed in mouse V1 (Fig. 2, Fig. 6; [30]) are consistent with tuning of recurrent connections within small cohorts of neurons to particular combinations of preferred orientations. Moreover, responses in mouse V1 are inconsistent with a simple configuration of cortical connections strictly aligned with feedforward visual responses.

Amplification and competition might underlie facilitation and suppression

Our theoretical analysis and simulation results demonstrate that non-random excitatory connectivity affects the computational properties of a cortical network by introducing amplification of responses within subnetworks of excitatory neurons, and competition in responses between subnetworks (Fig. 4a–c). Several recent studies have demonstrated that visual input is amplified within the superficial layers of cortex [54-56], and recent results from motor cortex suggest competition between ensembles of neurons [57]. Our modelling results indicated that some form of non-random local excitatory connectivity is required for such amplifica-

tion to occur through recurrent network interactions, with reasonable assumptions for anatomical and physiological parameters for rodent cortex (Fig. 4a–c; Fig. S1). This still leaves in question whether the *particular configuration* of non-random excitatory connectivity plays a role.

Our simulation results showed that the effects of amplification and competition on cortical responses are tuned to the statistics of local connectivity. This implies that stimuli matching the statistics of a subnetwork will undergo stronger amplification than non-matching stimuli (Fig. 7). In our "feature-binding" model, the statistics of subnetwork connectivity were similar to plaid stimuli composed of two grating components. As a result, plaid stimuli gave rise to stronger amplification than single grating components alone, when the composition of the plaid matched the composition of connectivity within a particular subnetwork. This led to a facilitatory effect, where some neurons responded more strongly to plaid stimuli than to the grating components underlying the plaid stimuli. Conversely, competition between subnetworks led to weaker responses to some plaid stimuli, for neurons that "lost" the competition. Competition could therefore be one cortical mechanism underlying cross-orientation suppression in response to plaid stimulation.

*** FIGURE 7 NEAR HERE ***

In contrast, suppression in the "like-to-like" and "random" models occur because the energy in the stimulus is spread across two grating components, and is not combined by the network to form strong plaid selectivity. In the "like-to-like" model, competition occurs between representations of the two oriented grating components of the plaid, causing additional suppression. The presence of amplified, strongly facilitating plaid responses in mouse V1 is therefore consistent with the existence of subnetworks representing the conjunction of differently-oriented edges.

Detecting feature-binding connectivity rules in cortex

We found that the precise rules that determine local connections among neurons in cortex can strongly affect the representation of visual stimuli. The "feature-binding" rule we examined embodies the simplest second-order relationship between connectivity and preferred orientation, and was chosen for this reason. We cannot rule out more complicated connectivity rules as being present in mouse V1, but we have shown that a simple "like-to-like" rule cannot explain responses to plaid visual stimuli. Random, non-specific connections were also unable to explain complex plaid responses in mouse V1 (Fig. S2).

How can the detailed statistics of "feature-binding" rules be measured in cortex? Existing experimental techniques have been used to measure only first-order statistical relationships between function and cortical connectivity [18,21-24,51]. Unfortunately, current technical limitations make it difficult to measure more complex statistical structures such as present under a "feature-binding" connectivity rule. Simultaneous whole-cell recordings are typically possible from only a small numbers of neurons, thus sparsely testing connectivity within a small cohort. Even if simultaneous recordings of up to 12 neurons are used [17], identifying and quantifying higher-order statistics in the local connectivity pattern is limited by the low probability of finding connected excitatory neurons in cortex.

In addition, our results highlight that small changes in the statistics of local connectivity can have drastic effects on computation and visual coding. Introducing a small degree of specificity, such that a minority of synapses are made within an excitatory subnetwork, is sufficient to induce strong specific amplification and strong competition to the network, even though a majority of the synapses are made randomly (Fig. 4a–c). Under our "feature-binding" model 68% of synapses are made randomly; approximately 21% are made under a "like-to-like" rule and the remaining 11% are used to bind visual features. Clearly, detecting the small proportion of synapses required to implement feature binding in V1 will be difficult, using random anatomical sampling techniques.

A recent study functionally characterised the presynaptic inputs to single superficial-layer neurons in mouse V1, using a novel pre-synaptic labelling technique [58]. Consistent with our results for preferred orientation (Fig. 3f, g), they found that presynaptic inputs were similarly tuned as target neurons but over a wide bandwidth. The majority of synaptically connected networks were tuned for multiple orientation preferences across cortical layers, similar to the feature-binding networks in our study.

We implemented an alternative approach, by inferring the presence of higher-order connectivity statistics from population responses in cortex. This technique could be expanded experimentally, by presenting a parameterised battery of simple and complex stimuli. Stimuli close to the configuration of local connectivity rules would lead to maximal facilitation and competition within the cortical network. Importantly, our results strongly suggest that simple stimuli alone are insufficient to accurately characterise neuronal response properties in visual cortex.

Building plaid responses from convergence of simple feedforward inputs

Could the complexity of plaid texture responses in mouse V1 be explained by convergence of differently tuned feedforward inputs from layer 4 onto single layer 2/3 neurons, similar to the proposed generation of pattern-selective responses in primate MT [33,59]? Building plaid responses in this way would imply that layer 2/3 neurons would respond to multiple grating orientations, since they would receive approximately equal inputs from at least two oriented components. However, layer 4 and layer 2/3 neurons are similarly tuned to orientation in rodent V1 [60,61], in conflict with this feedforward hypothesis.

In addition, if responses to complex stimuli were built by feedforward combination of simple grating components, then the response of a neuron to the set of grating stimuli would directly predict the plaid response of that neuron. This would then imply that two neurons with similar responses to plaid stimuli must

have similar responses to grating stimuli. However we found this not to be the case; two neurons with similar responses to grating components often respond differently to plaid textures or to natural scenes (Fig. 3d; Fig. 6a, b; [30]).

Computational role of inhibitory connectivity and physiology

Non-specific connectivity between excitatory and inhibitory neurons, as assumed in our simulation models, is consistent with the concept that inhibitory neurons simply integrate neuronal responses in the surrounding population [62], and is also consistent with experimental observations of weakly tuned or untuned inhibition in rodent visual cortex [41,46,51-53]. Although specific E↔I connectivity has been observed in rodent cortex [16,28], the majority of E↔I synapses are likely to be made functionally non-specifically in line with the high convergence of E→I and I→E connections observed in cortex [51,52,63].

In our models, shared inhibition is crucial to mediate competition between excitatory subnetworks (Fig. 4); inhibition is untuned because excitatory inputs to the inhibitory population are pooled across subnetworks. Poorly tuned inhibition, as expressed by the dominant class of cortical inhibitory neurons (parvalbumin expressing neurons), therefore plays an important computational role and is not merely a stabilising force in cortex.

Other inhibitory neuron classes in cortex (e.g. somatostatin or vaso-intestinal peptide expressing neurons) have been shown to exhibit feature-selective responses [46,64,65]. Recent computational work examined the influence of multiple inhibitory neuron classes with different physiological and anatomical tuning properties in a model for rodent cortex [66]. They examined the role of inhibitory connectivity on divisive and subtractive normalisation of network activity in a network with specific, orientation-tuned inhibitory connectivity. They found that specific inhibitory feedback could lead to divisive normalisation of network activity, while non-specific inhibitory feedback could lead to subtractive normalisation.

However, the computational role of specific inhibition is likely to rest on the precise rules for connectivity expressed between excitatory and inhibitory neurons. If the rules for E↔E and E↔I connections align, then a specific inhibitory population could act as a break on excitation within a subnetwork, and could allow more specific anatomical connectivity to persist while maintaining the balance between excitation and inhibition in cortex. The functional profile of this balancing pool would be highly tuned, and be similar to that of the excitatory neurons in the subnetwork, suggesting a physiological signature of specific inhibitory feedback that could be sought experimentally. Alternatively, if E↔I connection rules result in counter-tuned specificity, these connections would act to strengthen competition between subnetworks.

Feature binding to detect higher-order visual statistics

In visual cortex of primates, carnivores and rodents, orientation tuning develops before postnatal eye opening and in the absence of visual experience [67,68]. Local recurrent connections develop after the onset of visual experience and maintain their plasticity into adulthood [67,69-73]. Statistical correlations in natural scenes might therefore lead to wiring of subnetworks under an activity-dependent mechanism such as spike-time dependent plasticity (STDP) [74-78]. Along these lines, examinations of the development of specific excitatory connections after eye opening found that similarities in feedforward input were progressively encoded in specific excitatory connections [22].

We expect that, as the specificity of lateral connections forms during development, the emergence of compound feature selectivity will gradually occur after the onset of sensory experience. This hypothesis is consistent with experience-dependent development of modulatory effects due to natural visual stimulation outside of the classical receptive field, as has been observed in mouse V1 [79]. While a complete factorial combination of all possible features occurring in natural vision is clearly not possible, presumably the most prominent statistical features of cortical activity patterns would be prioritised for embedding through recurrent excitatory connectivity. At the same time, competition induced by non-

specific shared inhibition will encourage the separation of neurons into subnetworks. In our interpretation, single subnetworks would embed learned relationships between external stimulus features into functional ensembles in cortex, such that they could be recovered by the competitive mechanisms we have detailed.

In pre-frontal cortex, compound or mixed selectivity of single neurons to combinations of task-related responses, facilitating efficient decoding of arbitrary decision-related variables, has been found in several studies [80,81]. Binding feedforward cortical inputs into compound representations, as occurs in our "feature-binding" model, is therefore a useful computational process with general applicability. We propose that specific local excitatory connectivity is a general circuit mechanism for shaping information processing in cortical networks.

Materials and Methods

In-vivo calcium imaging

Experimental procedures followed institutional guidelines and were approved by the Cantonal Veterinary Office in Zürich or the UK Home Office. Procedures for urethane anaesthesia, craniotomies, bulk loading of the calcium indicator, as well as for *in vivo* two-photon calcium imaging and *in vitro* recording of synaptic connection strength were the same as described previously [24,30,82,83].

Preparation and imaging with OGB Male and female three-month old wild type C7BL/6 mice were sedated with chlorprothixene (10 mg/ml in Ringer solution; 0.01 ml per 20 g by weight) then anaesthetised with urethane (10% in isotonic saline; initial dose 0.1 ml per 20 g by weight; supplemented as required to maintain anaesthesia). The body temperature of anaesthetised animals was monitored and controlled using a heating pad and rectal thermometer. Atropine was given to reduce secretions (0.16 ml per 20 g by weight). Intrinsic optical imaging was used to locate primary visual cortex, and a craniotomy was made over V1.

We performed bulk loading of the synthetic calcium indicator Oregon Green-BAPTA–1 (OGB–1; Invitrogen). Several acute injections of OGB–1–AM were made under visual guidance into the visual cortex [84]. Sulforhodamine (SR–101; Invitrogen) was applied topically to the pial surface, to provide labelling of the astrocytic network [85]. Time-series stacks recording activity in layer 2/3 cortical neurons were acquired at a 4–10 Hz frame rate with a custom-built microscope equipped with a 40× objective (LUMPlanFl/IR, NA 0.8; Olympus) and an 80 MHz pulsed Ti:Sapphire excitation laser (MaiTai HP; Spectra Physics, Newport). Acquisition of calcium transients was performed using custom-written software in LabView (National Instruments), and analysis was performed using the open-source FocusStack toolbox [86].

Preparation and imaging with GCaMP6 Adult male mice (P75–P90) were initially anaesthetized with 4–5% isoflurane in O_2 and maintained on 1.5–2% during the surgical procedure. The primary visual cortex (V1) was localized using intrinsic

imaging. Briefly, the skull above the estimated location of V1 was thinned and we illuminated the cortical surface with 630 nm LED light, presented drifting gratings for 5 s, and collected reflectance images through a 4× objective with a CCD camera (Toshiba TELI CS3960DCL).

A craniotomy of 3–4 mm was opened above the region of strongest intrinsic signal response, which we assumed to be centred over V1. We then injected the genetically encoded calcium indicator GCaMP6m [87] (AAV1.Syn.GCaMP6m.WPRE.SV40; UPenn) around 250 µm below the cortical surface to target superficial layer neurons. 2–3 injections were made in a single animal and a volume of approximately 200 nl was injected at each location. The craniotomy was sealed with a glass window and a metal post for head fixation was implanted on the skull with dental acrylic, contralateral to the cranial window.

For imaging, animals were anaesthetised with isoflurane at 4% for induction, then head fixed. Isoflurane concentration was lowered to 0.5–0.75% during the experiment. We maintained the animal's body temperature at 37°C using a rectal thermometer probe and a heating pad placed under the animal. Silicon oil was applied to the eyes to keep them moist.

In vivo/in vitro characterisation of function and connectivity Methods for obtaining visual responses *in vivo* and measuring synaptic connectivity *in vitro* are described in [24]. Briefly, young C75/BL6 mice (P22–26) were anaesthetised (fentanyl, midazolam and medetomidine) and injected with OGB calcium indicators, lightly anaesthetised with isoflurane (0.3–0.5%) and head fixed. Two-photon imaging of calcium responses was used to record the response of neurons to a sequence of natural images (1800 individual images). After *in vivo* imaging experiments, simultaneous whole-cell recordings of up to six neurons at a time were performed *in vitro*. Evoked spikes and recorded EPSPs were used to identify synaptically connected pairs of neurons.

Visual stimulation

Visual stimuli for receptive field characterisation, drifting gratings and plaids and masked natural movies were provided by an LCD monitor (52.5 × 29.5 cm; BenQ) placed 10–11 cm from the eye of the animal and covering approximately 135 × 107 visual degrees (v.d.). The monitor was calibrated to have a linear intensity response curve. Contrast-oscillating grating and plaid stimuli were presented on an LCD monitor (15.2 × 9.1 cm; Xenarc) placed 9 cm from the eye of the animal and covering 80 × 54 v.d. The same screen was used for stimulus presentation during intrinsic imaging to locate visual cortex and during two-photon imaging. The open-source StimServer toolbox was used to generate and present visual stimuli via the Psychtoolbox package [86,88].

Stimuli for receptive field characterisation comprised a 5 × 5 array of masked high contrast drifting gratings (15 v.d. wide; overlapping by 40%; 9 v.d. per cycle; 1 Hz drift rate; 0.5 Hz rotation rate) presented for 2 s each in random order, separated by a blank screen of 2 s duration, with 50% luminance (example calcium response shown in Fig. 3a). Full-field high-contrast drifting gratings (33.33 v.d. per cycle; 1 Hz drift rate) were presented drifting in one of 8 directions for 2 s each in random order, separated by a 6 s period of blank screen with 50% luminance (example calcium response shown in Fig. 3b). Full-field 50% contrast drifting gratings (25 v.d. per cycle; 1 Hz drift rate) were presented drifting in one of 16 directions for 1 s each in random order (calcium responses shown in Fig. 2). Full-field drifting plaid stimuli were constructed additively from 50% contrast grating components (25 v.d. per cycle; 1 Hz drift rate; 1 s duration; Fig. 2). Full-field natural movies consisted of a 43 s continuous sequence with three segments (example calcium response shown in Fig. 3c). Full-field contrast-oscillating gratings and plaid stimuli were composed of bars of 8 v.d. width which oscillated at 2 Hz between black and white on a 50% grey background, and with a spatial frequency of 20 v.d./cycle (example calcium response shown in Fig. 6a). On each subsequent oscillation cycle the bars locations shifted phase by 180°. Static gratings were used to avoid introducing a movement component into the stimulus. A base orientation for the gratings of either horizontal or vertical was chosen, and five orientations spanning ±40 deg. around the base orientation were used. Contrast-oscillat-

ing plaids were composed of every possible combination of the five oscillating grating stimuli, giving 5 grating and 10 plaid stimuli for each experiment. A single trial consisted of a blank period (50% luminance screen) presented for 10 s, as well as presentations of each of the gratings and plaids for 5 s each, preceded by 5 s of a blank 50% luminance screen, all presented in random order.

Analysis of calcium transients

Analysis of two-photon calcium imaging data was conducted in Matlab using the open-source FocusStack toolbox [86]. During acquisition, individual two-photon imaging trials were visually inspected for Z-axis shifts of the focal plane. Affected trials were discarded, and the focal plane was manually shifted to align with previous trials before acquisition continued. Frames recorded from a single region were composed into stacks, and spatially registered with the first frame in the stack to correct lateral shifts caused by movement of the animal. Only pixels for which data was available for every frame in the stack were included for analysis. A background fluorescence region was selected in the imaged area, such as the interior of a blood vessel, and the spatial average of this region was subtracted from each frame in the stack. The baseline fluorescence distribution for each pixel was estimated by finding the mean and standard deviation of pixel values during the 10 s blank periods, separately for each trial. Regions of interest (ROIs) were selected either manually, or by performing low-pass filtering of the OGB (green) and sulforhodamine (red) channels, subtracting red from green and finding the local peaks of the resulting image.

A general threshold for responsivity was computed to ensure that ROIs considered responsive were not simply due to neuropil activity. The responses of all pixels outside any ROI were collected (defined as "neuropil"), and the Z-scores of the mean $\Delta F/F_0$ responses during single visual stimulus presentations were computed per pixel, against the 10 s baseline period. A threshold for single-trial responses to be deemed significant (z_{trial}) was set by finding the Z-score which would include only 1% of neuropil responses ($\alpha = 1\%$). A similar threshold was set for comparison against the strongest response of an ROI, averaged over all trials

(z_{max}). These thresholds always exceeded 3, implying that single-trial responses included for further analysis were at least 3 standard deviations higher than the neuropil response. Note that this approach does not attempt to subtract neuropil activity, but ensures that any ROI used for analysis responds to visual stimuli with calcium transients that can not be explained by neuropil contamination alone.

The response of a ROI to a stimulus was found on a trial-by-trial basis by first computing the spatial average of the pixels in a ROI for each frame. The mean of the frames during the blank period preceding each trial was subtracted and used to normalise responses ($\Delta F/F_0$), and the mean $\Delta F/F_0$ of the frames during the trial was computed. The standard deviation for the baseline of a neuron was estimated over all $\Delta F/F_0$ frames from the long baseline period and the pre-trial blank periods. ROIs were included for further analysis if the ROI was visually responsive according to trial Z-scores (maximum response > z_{max}) and reliable (trial response > z_{trial} for more than half of the trials). The response of a neuron to a stimulus was taken as the average of all single-trial $\Delta F/F_0$ responses.

Receptive fields of neurons recorded under natural movie and drifting grating stimulation were characterised by presenting small, masked high-contrast drifting gratings from a 5 × 5 array, in random order (see above; Fig. 3a). A receptive field for each neuron was estimated by a Gaussian mixture model, composed of circularly symmetric Gaussian fields ($\rho = 7.5$ v.d.) placed at each stimulus location and weighted by the response of the neuron to the grating stimulus at that location. The centre of the receptive field was taken as the peak of the final Gaussian mixture. Neurons were included for further analysis if the centre of their receptive field lay within a 7.5 v.d. circle placed at the centre of the natural movie visual stimulus. Example single-trial and trial-averaged calcium responses to natural movie stimuli are shown in Fig. 3c.

Response similarity measures and response metrics

The similarity in response between two neurons was measured independently for grating and plaid stimuli. The set of grating responses for each neuron were composed into vectors $R1_g$ and $R2_g$. Similarity in grating response was then given

by the Pearson's correlation coefficient between $R1_g$ and $R2_g$: $\rho_g = \mathrm{corr}\left(R1_g, R2_g\right)$ (see Fig. 3b, inset). The similarity in response to plaid stimuli was computed analogously over the sets of plaid responses $R1_p$ and $R2_p$: $\rho_p = \mathrm{corr}\left(R1_p, R2_p\right)$ (see Fig. 6a, inset). Similarity was only measured between neurons recorded in the same imaging site.

The similarity between neurons in their responses to movie stimuli (ρ_{m}) was measured by computing the signal correlation as follows. The calcium response traces for a pair of neurons were averaged over trials. The initial 1 s segment of the traces following the onset of a movie segment were excluded from analysis, to reduce the effect of transient signals in response to visual stimulus onset on analysed responses. The Pearson's correlation coefficient was then calculated between the resulting pair of traces (ρ_{m}; see Fig. 3c, inset). Note that correlations introduced through neuropil contamination were not corrected for, with the result that the mean signal correlation is positive rather than zero. For this reason we used thresholds for "high" correlations based on percentiles of the correlation distribution, rather than an absolute correlation value.

The similarity between neurons in their responses to flashed natural stimuli (ρ_{Ca}; Fig. 3f) was measured as the linear correlation between the vector of responses of a single neuron to a set of 1800 natural stimuli [24].

The Orientation Selectivity Index (OSI) of a neuron was estimated using the formula $OSI = \left[\max\left(R_g\right) - \min\left(R_g\right)\right] / \mathrm{sum}\left(R_g\right)$, where R_g is the set of responses of a single neuron to the set of grating stimuli.

The Plaid Selectivity Index (PSI) of a neuron, describing how selective a neuron is over a set of plaid stimuli, was calculated using the formula $PSI = 1 - \left[-1 + \sum_j R_{p,j} / \max\left(R_p\right)\right] / \left[\#\left(R_p\right) - 1\right]$, where $\#\left(R_p\right)$ is the number of stimuli in R_p [30]. The PSI of a neuron ranges from 0 to 1, where a value of 1 indicates a highly selective response, where a neuron responds to only a single stimulus; a value of 0 indicates equal, nonselective responses to all stimuli.

A plaid Modulation Index (MI), describing the degree of facilitation or suppression of a neuron in response to plaid stimuli, was calculated using the formula $MI = \left[\max\left(R_p\right) - \max\left(R_g\right)\right] / \left[\max\left(R_p\right) + \max\left(R_g\right)\right]$, where R_p is the set of responses of a single neuron to the set of plaid stimuli [30]. The MI of a neuron ranges from –1 to 1. Values of MI < 0 indicate stronger responses to grating stimuli compared with plaid stimuli; values of MI > 0 indicate stronger responses to plaid stimuli. A value of MI = –1 indicates that a neuron responds only to grating stimuli; a value of MI = 1 indicates that a neuron responses to only plaid stimuli.

The proportion of facilitating and suppressing neurons was compared between mouse V1 and model responses using two-tailed Fisher's exact tests. The population of responsive neurons was divided into three groups: facilitating (MI > 0.05); suppressing (MI < –0.05); and non-modulated (–0.05 <= MI <= 0.05). These categories were arranged into three 2 × 3 contingency tables, with each table tested to compare facilitation and suppression between mouse V1 and one model.

Generation of V1 control responses

We used single-cell, single-trial responses to oscillating contrast grating stimuli to explore whether we could distinguish between correlated and decorrelated responses to plaid stimuli, given experimental variability and noise. For each cell in the experimentally-recorded data set, we used the set of grating responses R_g to generate plaid responses R_p for the same cell, under the assumption that the response to a plaid was linearly related to the sum of the responses to the two grating components. For each plaid, we randomly selected a single-trial response for each of the grating components of the plaid. The predicted single-trial plaid response was the sum of the two grating responses. We generated 100 bootstrap samples for each experimental population, with each sample consisting of the same number of trials and neurons as the experimental population. We then quantified the relationship between grating and plaid responses as described for the experimental data.

Models of mouse V1

We designed a model of the superficial layers of mouse primary visual cortex, to explore the effect of different connectivity rules on information processing within the cortex. A simple version of this model, comprising only five neurons with mean-field connectivity, was used for analytical exploration ("analytical model"; Fig. 4, Fig. S1, Fig. 7). A large-scale version, comprising 80,000 neurons with sparse connectivity, was used for direct comparison with experimental results (Fig. 5–6). A full list of parameters for both models is given in Table 1.

Common model dynamics Individual excitatory neurons (approximating layer 2/3 pyramidal cells) and inhibitory neurons (approximating layer 2/3 basket cells) were modelled as linear-threshold units, with equal time constants and thresholds set to zero. The dynamics of each rate-coded neuron in the large- and small-scale models was governed by the differential equation

$$\tau_i \cdot \dot{x}_i = -x_i + \sum_j^{N_N} w_{i,j} \left[x_j - \beta_j \right]^+ + I_i(t) + \sigma_i \cdot \zeta_i(t), \tag{1}$$

where τ_i is the time constant of neuron *i*; x_i is the instantaneous current being injected into neuron *i*; $[\ ]^+$ denotes the linear-threshold transfer function $[x]^+ = \max(x, 0)$; β_j is the activation threshold of neuron *j*; $I_i(t)$ is the stimulus input current provided to neuron *i* at time *t*; $\sigma_i \cdot \zeta_i(t)$ is a white noise process included to approximate the barrage of spontaneous E- and I-PSPs experienced by cortical neurons; and N_N is the total number of neurons in the model. The total directed connection strength between two neurons j and i is given in Eq. (1) by $w_{i,j} = g_j \cdot n_{i,j} \cdot \alpha_j$, where g_j is the charge injected by a synapse from neuron *j* to neuron *i* and $n_{i,j}$ is the number of synapses made by neuron *j* onto neuron *i*; α_j is the gain of neuron *j*.

Synaptic input Synapses were modelled as constant current sources that injected an amount of charge per second related to the average firing rate of the presynaptic neuron, modulated by the synaptic release probability. Single excitatory synapses were assigned a weight of 0.01 pC/spike/synapse; single inhibitory synapses were considered to be 10 times stronger [4]. Excitatory and inhibitory neurons were assigned output gains of 0.066 spikes/pC [89].

Analytical model To explore the basic stability and computational consequences of non-random excitatory connectivity, we built a small five-node model (four excitatory and one inhibitory neuron; Fig. 4). Connections within this model were defined to approximate the average expected connectivity between populations of neurons in layers 2/3 of mouse V1. Excitatory neurons were grouped into two subnetworks, and a proportion s of synapses from each excitatory neurons was reserved to be made within the same subnetwork. When $s=0$, E↔E synapses were considered to be made randomly, such that each connection in the small model approximated the average total connection strength expected in mouse V1. When $s=1$, all E↔E synapses were considered to be specific within the same subnetwork, such that no synapses were made between excitatory neurons in different subnetworks. Connections to and from the inhibitory node were considered to be made randomly in every case. The resulting weight matrix for this network is therefore given by

$$\mathrm{W}=\begin{bmatrix} a & a & b & b & -w_{ie} \\ a & a & b & b & -w_{ie} \\ b & b & a & a & -w_{ie} \\ b & b & a & a & -w_{ie} \\ w_{ei} & w_{ei} & w_{ei} & w_{ei} & -w_I \cdot f_I \end{bmatrix}, \text{where} \tag{2}$$

$a=w_S/2+w_N/4$ is the excitatory weight between neurons in the same subnetwork; $b=w_N/4$ is the excitatory weight between neurons in different subnetworks; $w_{ie}=w_I\cdot(1-f_I)/4$ is the nonspecific inhibitory to excitatory feedback weight; $w_{ei}=w_E\cdot f_I$ is the nonspecific excitatory to inhibitory weight; $w_S=w_E\cdot(1-f_I)\cdot s$ is the specific weight component, $w_N=w_E\cdot(1-f_I)\cdot(1-s)$ is the nonspecific weight component, w_E is the total synaptic weight from a single excitatory neuron, w_I is the total synaptic weight from a single inhibitory neuron and $f_I=1/5$ is the proportion of inhibitory neurons. Preferred orientations for each excitatory neuron are indicated in Fig. 4. When a stimulus matched the preferred orientation of a neuron, a constant input current was injected ($I_i(t)=\iota$); when a stimulus did not match the preferred orientation, no input current was provided to that neuron ($I_i(t)=0$). When simulating the analytical model, the input current $\iota=1$.

Measuring stability and competition To determine network stability in the analytical model, we performed an eigenvalue analysis of the system Jacobian, given by J = (W–I)./T, where W is the system weight matrix as given above, I is the identity matrix, T is the matrix composed of time constants for each post-synaptic neuron corresponding to elements in W and A./B indicates element-wise division between matrices A and B. The network was considered stable if all eigenvalues of J as well as the trace of the Jacobian Tr(J) were non-positive. The non-linear dynamical system was linearised about the fixed point where all neurons are active; if this fixed point is unstable then the system operates in either a hard winner-take-all mode if a different partition is stable, or is globally unstable [90,91]. Either of these modes is undesirable for cortex.

For a network to be in an ISN regime, the excitatory portion of the network must be unstable in the absence of inhibition, and inhibition must be strong enough in the full network to balance excitation. To determine whether the parameter regimes place the network in an inhibition-stabilised (ISN) regime, we therefore performed an eigenvalue analysis of the system in which all inhibitory connections were removed (i.e. $w_I = 0$). Either one eigenvalue of the corresponding Jacobian J^E of the excitatory-only network or the system trace $Tr(J^E)$ was required to be positive, but the system including inhibitory feedback was required to be stable.

The presence and strength of competition in Fig. 4 was determined by injecting current into a single excitatory neuron and recording the net current received by an excitatory neuron in the opposite subnetwork at the network fixed point (see Fig. 4a). Negative net currents correspond to competition between the stimulated and recorded excitatory neurons (shown as shading in Fig. S1).

Large-scale model To construct the large-scale simulation model of mouse V1, 80,000 linear-threshold neurons were each assigned a random location $\mathbf{u}_i \in \mathbb{T}^2$ where $\mathbb{T}$ defines the surface of a virtual torus of size 2.2×2.2 mm. Excitatory and inhibitory neurons were placed with relative densities appropriate for layers 2 and 3 of mouse cortex [92].

To determine patterns of synaptic connectivity, we calculated for each neuron the probability distribution of forming a synaptic connection with all other neurons in the model. A fixed number of synapses was drawn from this distribution; the number was chosen as an estimate of the number of synapses formed with other superficial layer neurons in rodent cortex (8142 from each excitatory and 8566 from each inhibitory neuron; [1,92]). Since a simulation with the full density of cortical neurons was computationally infeasible, the size of the simulations was scaled to 10% of estimated cortical density. The sparsity of local synaptic connectivity was maintained by also scaling the number of synapses made by each neuron, while maintaining the total synaptic conductance formed by each neuron.

Axonal and dendritic densities for each neuron were described by a two-dimensional Gaussian field

$$\mathcal{G}(\mathbf{v},\mathbf{u}_i,\rho_i)=\exp\left(\frac{-\|\mathbf{v},\mathbf{u}_i\|^2}{2\rho_i^{\,2}}\right), \tag{3}$$

where ρ_i is a field dispersion parameter associated with neuron i and $\|\mathbf{v},\mathbf{u}\|$ is the Euclidean distance between $\mathbf{v}$ and $\mathbf{u}$, computed over the surface of a 2D torus. In our models, each neuron had a Gaussian dendritic field of $\rho_{\mathrm{d}}=75\,\mu\mathrm{m}$ (corresponding to an approximate width of $4\rho=300\,\mu\mathrm{m}$; [93]); and axonal field of $\rho_{\mathrm{a,e}}=290\,\mu\mathrm{m}$ for excitatory neurons (width $1100\,\mu\mathrm{m}$; [93-95]) and $\rho_{\mathrm{a,i}}=100\,\mu\mathrm{m}$ for inhibitory neurons (width $400\,\mu\mathrm{m}$; [96]).

Anatomical connectivity rule Our default rule for forming synapses was based on Peters' Rule, in that the probability of forming a synapse was proportional to the overlap between axonal and dendritic fields [2,8]. This was estimated by computing the integrated product of axonal and dendritic fields over a torus $\mathbb{T}$:

$$p_{Peters}=\left[\!\left[\iint_{\mathbb{T}}\mathcal{G}(\mathbf{v},\mathbf{u}_i,\rho_{d,i})\mathcal{G}(\mathbf{v},\mathbf{u}_j,\rho_{a,j})\mathrm{d}\mathbf{v}\right]\!\right], \tag{4}$$

where p_{Peters} is the probability of forming a single synapse between neurons i and j, and the notation $[\![\cdots]\!]$ indicates that the expression between the double brackets is normalised to form a probability density function, such that if summed across all possible target neurons the total will be equal to 1.

Like-to-like connectivity rule We investigated two rules for anatomical specificity in intra-cortical excitatory recurrent connections. The first such rule corresponds to the case where local recurrent connectivity is aligned with matching feedforward visual properties (preferred orientation, in our case). We therefore assumed that the probability of forming a synapse is modulated by the similarity in preferred orientation between two excitatory neurons ("Like-to-Like" rule; see Fig. 5a). The probability of connection between two neurons was proportional to

$$p_{conn} \propto p_{Peters}\left(s_1 [\![p_{ori}]\!] + (1 - s_1)\right), \text{ where} \tag{5}$$

$p_{ori} = \text{vonmises}\left(\theta_i, \theta_j, \kappa\right)$; p_{Peters} is the connection probability under non-specific Peters' rule connectivity, defined above; and s_1 is the proportional strength of specificity $s_1 \in [0,1]$. If $s_1 = 0$ then Eq. (5) becomes equivalent to Peters' rule. When $s_1 = 1$ then the probability of connecting orthogonally tuned neurons is zero.

Feature-binding connectivity rule The second rule for anatomical connection specificity corresponds to the case where local recurrent connectivity is not aligned with feedforward visual properties. Instead, it was designed to explore binding of simple visual features ("Feature-Binding" specificity; see Fig. 5e). Under this rule, a subnetwork combined neurons with a number ϑ of different orientation preferences. The preferred orientations used to compose a subnetwork in the Feature-Binding specificity model were chosen from periodic filtered noise fields.

Each noise field $\mathbf{Z}_{k,q}$ was built by generating a unit-magnitude complex number $\mathbf{z}_j = \exp\left(-\mathrm{i}\zeta_j\right)$ for each neuron in the model, with uniformly-distributed orientations $\zeta_j \in [-\pi, \pi)$. Here "i" represents the complex number $\sqrt{-1}$; $k \in [1, N_S]$, where N_S is the number of subnetworks in the model; $q \in [1, \vartheta]$, where ϑ is the number of preferred orientations per subnetworks. In our models described in this paper, $N_S = 6$ and $\vartheta = 2$.

A field $\mathbf{Z}_{k,q}$ was defined by placing each $\mathbf{z}_j$ at the location $\mathbf{u}_j$ of the corresponding neuron. Each complex field $\mathbf{Z}_{k,q}$ was spatially filtered by convolving with a Gaussian field $\mathcal{G}_\rho$ on a torus, with a spatial std. dev. of $\rho = 75\,\mu\text{m}$ (width $300\,\mu\text{m}$). The angles from the resulting field of complex numbers was used as one orientation component for one subnetwork, at each point in simulated space. The composition of each subnetwork therefore changed smoothly across cortical space, so

that nearby neurons in the same subnetwork had similar functional selectivity. Therefore, $\angle\left(\mathbf{Z}\circ\mathcal{G}_{\rho}\right)$ defines a $N_S\times\vartheta$ matrix of numbers where each element determines one preferred orientation component of the corresponding subnetwork.

Neurons were assigned to one of N_S subnetworks, according to the maximum similarity between a neuron's preferred orientation and the orientation composition of the set of subnetworks at the location of the neuron's soma. The similarity between a neuron's preferred orientation and a subnetwork orientation was computed using a von Mises function with width parameter κ_2, such that the membership probability was proportional to

$$p_m\left(k,\theta_i\right)\propto\left[\!\left[\max\left[\text{vonmises}\left(\theta_i,\theta_{k,1},\kappa_2\right),\text{vonmises}\left(\theta_i,\theta_{k,2},\kappa_2\right)\right]\right]\!\right], \tag{6}$$

where k is the index of an SSN consisting of preferred orientations $\theta_{k,1}$ and $\theta_{k,2}$; θ_i is the preferred orientation of a neuron under consideration; and the expression within the double brackets $[\![\cdots]\!]$ was normalised to be a valid probability density function over k. A neuron was assigned membership of an SSN according to the formula

$$M(i)=\arg\max_k\left(p_m\left(k,\theta_i\right)\right), \tag{7}$$

where $M(i)$ gives the index of the SSN of which neuron i is a member.

The probability of connection between two neurons under the feature-binding model is therefore given by

$$p_{conn}\propto(1-s_2)p_{Peters}\left(s_1\cdot[\![p_{ori}]\!]+1-s_1\right)+s_2[\![b_{SSN}\cdot p_{Peters}]\!], \tag{8}$$

where parameter s_1 determines the relative contribution of Non-Specific versus orientation-tuned Like-to-Like specificity as in Eq. (5); s_2 determines the relative contribution of Feature-Binding specificity; $p_{ori}=\text{vonmises}\left(\theta_i,\theta_j,\kappa_1\right)$ as in Eq. (5); and b_{SSN} is a value equal to 1 when the two neurons fall within the same subnetwork; that is

$$b_{SSN}=\begin{cases}1\,\text{iff}\,M(i)=M(j)\\ 0\ \text{otherwise}\end{cases} \tag{9}$$

Network input Input was provided to the network as a simulation of orientation-tuned projections from layer 4 to layers 2/3 [60,61]. Each excitatory neuron was assigned an orientation tuning curve based on a von Mises function (a circu-

lar, Gaussian-like function), with a randomly chosen preferred orientation θ_i and a common input tuning curve width $\kappa = 4$. vonmises$(\cdot)$ is the non-normalised von Mises function with vonmises$(\cdot) \in [0,1]$, given by

$$\text{vonmises}(t,\theta,\kappa) = \exp[\kappa \cos 2(t-\theta)]. \tag{10}$$

Current was injected into each simulated neuron proportional to the orientation tuning curve of that neuron, according to the orientation content of the stimulus:

$$I_i(t) \propto \frac{A(t)}{N_N} \text{vonmises}\left(\theta_g(t), \theta_i, \kappa_i\right), \tag{11}$$

where $A(t)$ is the amplitude of the stimulus at time *t*; $\theta_g(t)$ is the orientation of a grating stimulus at time *t*; θ_i is the preferred orientation of neuron *i*; κ_i is the tuning curve width of neuron *i*; N_N is the total number of neurons in the network. The input to the network is normalised such that the total current injected into the network is equal to $A(t)$. For a simulated plaid stimulus composed of the two component orientations θ_{g1} and θ_{g2}, input to a neuron was the linear average of input associated with each grating component, given by

$$I_i(t) \propto \frac{A(t)}{2N_N}\left(\text{vonmises}\left(\theta_{g1}, \theta_i, \kappa_i\right) + \text{vonmises}\left(\theta_{g2}, \theta_i, \kappa_i\right)\right). \tag{12}$$

Both grating and plaid stimuli were considered to cover the full visual field. Tuned input currents were injected only into excitatory neurons, because we wanted to investigate the effect of excitatory recurrence on cortical information processing. Providing untuned feedforward input to inhibitory neurons can produce the illusion of competition between excitatory neurons, merely due to the thresholding effect of feedforward inhibitory input shared between those neurons.

Inclusion of experimental response variability We simulated large-scale networks as described above, and obtained responses to simulated visual stimuli. In order to mimic the response variability due to experimental conditions, such as recording noise and intrinsic neuronal response variability, we introduced a random component to the model responses.

For each presented stimulus i (e.g. a grating of a given orientation), we obtained a set S_i of single-trial responses $r_{i,j}$ for a single neuron such that $r_{i,j} \in S_i$, and the trial-averaged response $\overline{r_i} = \sum_{j=1..T} r_{i,j}/T$, where T is the number of trials collected for that stimulus. Over the full set of stimuli for a given neuron, we determ-

ined the maximum trial-averaged response $\overline{r}_{\max} = \max_i \overline{r}_i$. We then measured the standard deviation σ over the collection of all single-trial responses over all stimuli for a given neuron normalised by $\overline{r}_{\max}$, such that $\sigma = \mathrm{std}\left(\bigcup_i S_i / \overline{r}_{max}\right)$. The estimated experimental variability $\hat{\sigma}$ was defined as the median σ over all recorded neurons.

A similar procedure in reverse was applied to model-simulated visual responses, to mimic experimental variability. Activity of single neurons in response to a simulated stimulus i was interpreted as the mean response $\overline{r}_i$, with $\overline{r}_{\max}$ defined as above. Single-trial model responses were then generated as $r_{i,j} = \overline{r}_i + N(0, \hat{\sigma} \cdot \overline{r}_{\max})$, where $N(\mu, \sigma)$ generates a single normally-distributed random variate with mean μ and standard deviation σ. Twelve trials were generated for each stimulus (i.e. T=12), and single-trial responses were then analysed as described for experimentally recorded responses.

Estimation of parameters for connection rules Ko and colleagues characterised functional specificity in mouse primary visual cortex, by recording in slice from pairs of neurons that were functionally characterised *in vivo* [21]. We fit our function p_{conn} (Eq. (5)) to their measurements of the probability of connection between neurons tuned for orientation, giving estimates for both κ_1 and s_1 ($\hat{\kappa}_1 = 0.5$; $\hat{s}_1 = 0.45$). These parameters correspond to fairly weak functional specificity. We found that in the Like-to-Like specificity model, in order to have an appreciable network effect we had to increase the strength of functional specificity to $s_1 = 0.8$ (with $\kappa_1 = 0.5$). The connectivity measurements of Yoshimura and colleagues suggest that on the order of $N = 5$–6 subnetworks exist in layers 2/3 of rodent cortex [15]. For the Feature-Binding specificity model, we took the parameters $s_1 = 0.45$, $s_2 = 0.225$, $\kappa_1 = 0.5$, $\kappa_2 = 4$, $N = 6$, $\vartheta = 2$.

Statistical methods

We used a sample size commensurate with those used in the field, and sufficient for statistical analysis of our observations. No explicit sample size computation was performed.

Two-sided, non-parametric statistical tests were used unless stated otherwise in the text.

Acknowledgements

The authors gratefully acknowledge T Mrsic-Flogel, L Cossell and MF Iacaruso for providing the data analysed in Fig. 3. We are grateful to MA Penny and the attendees of the CapoCaccia workshop for helpful discussions.

Funding

This work was supported by the Velux Stiftung (grant number 787 to DRM); the Novartis Foundation (grants to DRM and BMK); the Swiss National Science Foundation (grant number 31–120480 to BMK); the European Commission FP7 program (grant BrainScales 269921 to FH and BMK); and by the Convergent Science Network (fellowships to DRM). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests

The authors declare no competing interests.

References

1. Binzegger T, Douglas RJ, Martin KAC. A quantitative map of the circuit of cat primary cortex. Journal of Neuroscience 2004;24(39):8441-53.

2. Braitenberg, Schüz. Anatomy of the cortex: Statistics and geometry. Springer-Verlag; 1991.

3. Hill SL, Wang Y, Riachi I, Schürmann F, Markram H. Statistical connectivity provides a sufficient foundation for specific functional connectivity in neocortical neural microcircuits. Proc Natl Acad Sci U S A 2012, Oct;109(42):E2885-94.

4. Binzegger T, Douglas R, Martin KAC. Topology and dynamics of the canonical circuit of cat V1. Neural Networks 2009;22:1071-8.

5. Ramaswamy S, Hill SL, King JG, Schürmann F, Wang Y, Markram H. Intrinsic morphological diversity of thick-tufted layer 5 pyramidal neurons ensures robust and invariant properties of in silico synaptic connections. The Journal of Physiology 2012, Feb 2;590(4):737-52.

6. Markram H, Muller E, Ramaswamy S, Reimann MW, Abdellah M, Sanchez CA, et al. Reconstruction and simulation of neocortical microcircuitry. Cell 2015, Oct 8;163(2):456-92.

7. Reimann MW, King JG, Muller EB, Ramaswamy S, Markram H. An algorithm to predict the connectome of neural microcircuits. Front Comput Neurosci 2015;9:120.

8. Peters A. Thalamic input to the cerebral cortex. Trends Neurosci 1979;2:183-5.

9. Juliano SL, Friedman DP, Eslin DE. Corticocortical connections predict patches of stimulus-evoked metabolic activity in monkey somatosensory cortex. Journal of Comparative Neurology 1990;298:23-39.

10. Malach R, Amir Y, Harel M, Grinvald A. Relationship between intrinsic connections and functional architecture revealed by optical imaging and in vivo targeted biocytin injections in primate striate cortex. Proc. Natl. Acad. Sci. USA 1993, Nov;90:10469-73.

11. Bosking WH, Zhang Y, Schofield B, Fitzpatrick D. Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. Journal of Neuroscience 1997, Mar 15;17(6):2112-27.

12. Muir DR, Da Costa NMA, Girardin C, Naaman S, Omer DB, Ruesch E, et al. Embedding of cortical representations by the superficial patch system. Cerebral Cortex 2011, Mar 7;21(10):2244-60.

13. Martin KA, Roth S, Rusch ES. Superficial layer pyramidal cells communicate heterogeneously between multiple functional domains of cat primary visual cortex. Nat Commun 2014;5:5252.

14. Ohki K, Chung S, Ch'ng YH, Kara P, Reid RC. Functional imaging with cellular resolution reveals precise micro-architecture in visual cortex. Nature 2005;433:597-603.

15. Yoshimura Y, Dantzker JLM, Callaway EM. Excitatory cortical neurons form fine-scale functional networks. Nature 2005, Feb 24;433:868-73.

16. Yoshimura Y, Callaway EM. Fine-scale specificity of cortical networks depends on inhibitory cell type and connectivity. Nat Neurosci 2005;8(11):1552-9.

17. Perin R, Berger TK, Markram H. A synaptic organizing principle for cortical neuronal groups. Proc Natl Acad Sci U S A 2011, Mar 3;108(13):5419-24.

18. Kampa BM, Letzkus JJ, Stuart GJ. Cortical feed-forward networks for binding different streams of sensory information. Nat Neurosci 2006, Dec;9(12):1472-3.

19. Yu YC, He S, Chen S, Fu Y, Brown KN, Yao XH, et al. Preferential electrical coupling regulates neocortical lineage-dependent microcircuit assembly. Nature 2012, Jun 7;486(7401):113-7.

20. Song S, Sjöström PJ, Reigl M, Nelson S, Chklovskii DB. Highly nonrandom features of synaptic local connectivity in local cortical circuits. Public Library of Science Biology 2005, Mar;3(3):0507-19.

21. Ko H, Hofer SB, Pichler B, Buchanan KA, Sjöström PJ, Mrsic-Flogel TD. Functional specificity of local synaptic connections in neocortical networks. Nature 2011;473:87-91.

22. Ko H, Cossell L, Baragli C, Antolik J, Clopath C, Hofer SB, Mrsic-Flogel TD. The emergence of functional microcircuits in visual cortex. Nature 2013, Apr 3;496(7443):96-100.

23. Li Y, Lu H, Cheng P-L, Ge S, Xu H, Shi S-H, Dan Y. Clonally related visual cortical neurons show similar stimulus feature selectivity. Nature 2012, May 5;486(7401):118-21.

24. Cossell L, Iacaruso MF, Muir DR, Houlton R, Sader EN, Ko H, et al. Functional organization of excitatory synaptic strength in primary visual cortex. Nature 2015, Feb 4;518(7539):399-403.

25. Lee WA, Bonin V, Reed M, Graham BJ, Hood G, Glattfelder K, Reid RC. Anatomy and function of an excitatory network in the visual cortex. Nature 2016, Mar 28.

26. Brown SP, Hestrin S. Intracortical circuits of pyramidal neurons reflect their long-range axonal targets. Nature 2009, Feb 26;457(7233):1133-6.

27. Morishima M, Morita K, Kubota Y, Kawaguchi Y. Highly differentiated projection-specific cortical subnetworks. J Neurosci 2011, Jul 7;31(28):10380-91.

28. Bopp R, Maçarico da Costa N, Kampa BM, Martin KA, Roth MM. Pyramidal cells make specific connections onto smooth (gabaergic) neurons in mouse visual cortex. PLoS Biol 2014, Aug;12(8):e1001932.

29. Muir DR, Mrsic-Flogel T. Eigenspectrum bounds for semirandom matrices with modular and spatial structure for neural networks. Physical Review E 2015, Apr 24;91(4):042808.

30. Muir DR, Roth MM, Helmchen F, Kampa BM. Model-based analysis of pattern motion processing in mouse primary visual cortex. Front Neural Circuits 2015;9:38.

31. Juavinett AL, Callaway EM. Pattern and component motion responses in mouse visual cortical areas. Curr Biol 2015, Jun 29;25(13):1759-64.

32. Palagina G, Meyer JF, Smirnakis SM. Complex visual motion representation in mouse area V1. J Neurosci 2017, Jan 4;37(1):164-83.

33. Movshon J, Adelson E, Gizzi M, Newsome W. The analysis of moving visual patterns. In: Chagas, Gattass, Gross, editors. Study Week on Pattern Recognition Mechanisms. Pontificiae Academiae Scientiarum Scripta Varia. Vatican City: Vatican Press; 1985. p. 117-51.

34. Martin KAC, Schröder S. Functional heterogeneity in neighboring neurons of cat primary visual cortex in response to both artificial and natural stimuli. Journal of Neuroscience 2013, Apr 24;33(17):7325-44.

35. Aljadeff J, Renfrew D, Vegué M, Sharpee TO. Low-dimensional dynamics of structured random networks. Phys Rev E 2016, Feb;93(2):022302.

36. Lefort S, Tomm C, Floyd Sarria JC, Petersen CC. The excitatory neuronal network of the C2 barrel column in mouse primary somatosensory cortex. Neuron 2009, Jan 29;61(2):301-16.

37. Amit DJ, Brunel N. Learning internal representations in an attractor neural network with analogue neurons. Network: Computation in Neural Systems 1995;6(3):359-88.

38. van Vreeswijk C, Sompolinsky H. Chaos in neuronal networks with balanced excitatory and inhibitory activity. Science 1996, Dec 6;274(5293):1724-6.

39. Tsodyks MV, Skaggs WE, Sejnowski TJ, McNaughton BL. Paradoxical effects of external modulation of inhibitory interneurons. J Neurosci 1997, Jun 1;17(11):4382-8.

40. Ozeki H, Finn IM, Schaffer ES, Miller KD, Ferster D. Inhibitory stabilization of the cortical network underlies visual surround suppression. Neuron 2009, May 28;62(4):578-92.

41. Atallah BV, Bruns W, Carandini M, Scanziani M. Parvalbumin-expressing interneurons linearly transform cortical responses to visual stimuli. Neuron 2012, Jan 12;73(1):159-70.

42. Douglas RJ, Mahowald MA, Martin KAC. Hybrid analog-digital architectures for neuromorphic systems. IEEE International Conference on Neural Networks 1994;3:1848-53.

43. Ben-Yishai R, Bar-Or RL, Sompolinsky H. Theory of orientation tuning in visual cortex. Proc Natl Acad Sci U S A 1995, Apr;92:3844-8.

44. Sadeh S, Clopath C, Rotter S. Processing of feature selectivity in cortical networks with specific connectivity. PLoS One 2015, Jun 17;10(6):e0127547.

45. Gabott PLA, Somogyi P. Quantitative distribution of gaba-immunoreactive neurons in the visual cortex (area 17) of the cat. Experimental Brain Research 1986;61:323-31.

46. Kerlin AM, Andermann ML, Berezovskii VK, Reid RC. Broadly tuned response properties of diverse inhibitory neuron subtypes in mouse visual cortex. Neuron 2010, Sep;67:858-71.

47. Ermentrout B. Linearization of F-I curves by adaptation. Neural Computation 1998;10:1721-9.

48. Neftci E, Chicca E, Indiveri G, Douglas R. A systematic method for configuring VLSI networks of spiking neurons. Neural Computation 2011;23(10):2457-97.

49. Neftci E, Binas J, Rutishauser U, Chicca E, Indiveri G, Douglas RJ. Synthesizing cognition in neuromorphic electronic systems. Proceedings of the National Academy of Sciences 2013;110(37):E3468-76.

50. Shriki O, Hansel D, Sompolinsky H. Rate models for conductance-based cortical neuronal networks. Neural Computation 2003, Aug;15(8):1809-41.

51. Bock DD, Lee W-CA, Kerlin AM, Andermann ML, Hood G, Wetzel AW, et al. Network anatomy and *in vivo* physiology of visual cortical neurons. Nature 2011, Mar;471:177-82.

52. Hofer SB, Ko H, Pichler B, Vogelstein J, Ros H, Zeng H, et al. Differential connectivity and response dynamics of excitatory and inhibitory neurons in visual cortex. Nat Neurosci 2011, Aug;14(8):1045-52.

53. Liu B-H, Li P, Li Y-T, Sun YJ, Yanagawa Y, Obata K, et al. Visual receptive field structure of cortical inhibitory neurons revealed by two-photon imaging guided recording. J Neurosci 2009, Aug;29(34):10520-32.

54. Li L-Y, Li Y-T, Zhou M, Tao HW, Zhang LI. Intracortical multiplication of thalamocortical signals in mouse auditory cortex. Nat Neurosci 2013, Aug 11;16(9):1179-81.

55. Lien AD, Scanziani M. Tuned thalamic excitation is amplified by visual cortical circuits. Nat Neurosci 2013, Aug 11;16(9):1315-23.

56. Li Y-T, Ibrahim LA, Liu B-H, Zhang LI, Tao HW. Linear transformation of thalamocortical input by intracortical excitation. Nat Neurosci 2013, Aug 11;16(9):1324-30.

57. Zagha E, Ge X, McCormick DA. Competing neural ensembles in motor cortex gate goal-directed motor output. Neuron 2015, Nov 4;88(3):565-77.

58. Wertz A, Trenholm S, Yonehara K, Hillier D, Raics Z, Leinweber M, et al. Single-cell-initiated monosynaptic tracing reveals layer-specific cortical network modules. Science 2015, Jul 3;349(6243):70-4.

59. Rust NC, Mante V, Simoncelli EP, Movshon JA. How MT cells analyze the motion of visual patterns. Nat Neurosci 2006, Nov;9(11):1421-31.

60. Niell CM, Stryker MP. Highly selective receptive fields in mouse visual cortex. J Neurosci 2008, Jul;28(30):7520-36.

61. Medini P. Cell-type-specific sub- and suprathreshold receptive fields of layer 4 and layer 2/3 pyramids in rat primary visual cortex. Neuroscience 2011, Sep 8;190:112-26.

62. Mariño J, Schummers J, Lyon DC, Schwabe L, Beck O, Wiesing P, et al. Invariant computations in local cortical networks with balanced excitation and inhibition. Nat Neurosci 2005;8(2):194-201.

63. Fino E, Yuste R. Dense inhibitory connectvity in neocortex. Neuron 2011;69:1188-203.

64. Ma WP, Liu BH, Li YT, Huang ZJ, Zhang LI, Tao HW. Visual representations by cortical somatostatin inhibitory neurons--selective but with weak and delayed responses. J Neurosci 2010, Oct 27;30(43):14371-9.

65. Wilson NR, Runyan CA, Wang FL, Sur M, Wilson NR, Runyan CA, et al. Division and subtraction by distinct cortical inhibitory networks in vivo. Nature 2012, Aug 8.

66. Litwin-Kumar A, Rosenbaum R, Doiron B. Inhibitory stabilization and visual coding in cortical circuits with multiple interneuron subtypes. J Neurophysiol 2016, Jan 6;115:1399-409.

67. White LE, Fitzpatrick D. Vision and cortical map development. Neuron 2007;56:327-38.

68. Rochefort NL, Narushima M, Grienberger C, Marandi N, Hill DN, Konnerth A. Development of direction selectivity in mouse cortical neurons. Neuron 2011, Aug 11;71(3):425-32.

69. Galuske RA, Singer W. The origin and topography of long-range intrinsic projections in cat visual cortex: A developmental study. Cerebral Cortex 1996;6:417-30.

70. Luhmann HJ, Millán LM, Singer W. Development of horizontal intrinsic connections in cat striate cortex. Experimental Brain Research 1986;63:443-8.

71. Luhmann HJ, Singer W, Martinez-Millán L. Horizontal interactions in cat striate cortex: I. Anatomical substrate and postnatal development. European Journal of Neuroscience 1990;2:344-57.

72. Katz LC, Callaway EM. Development of local circuits in mammalian visual cortex. Annu Rev Neurosci 1992;15:31-56.

73. Miller B, Blake NM, Erinjeri JP, Reistad CE, Sexton T, Admire P, Woolsey TA. Postnatal growth of intrinsic connections in mouse barrel cortex. The Journal of Comparative Neurology 2001, Jul 17;436(1):17-31.

74. Kampa BM, Letzkus JJ, Stuart GJ. Dendritic mechanisms controlling spike-timing-dependent synaptic plasticity. Trends Neurosci 2007, Sep;30(9):456-63.

75. Markram H, Gerstner W, Sjöström J. Spike-Timing-Dependent plasticity: A comprehensive overview. Frontiers in Synaptic Neuroscience 2012;4.

76. Clopath C, Büsing L, Vasilaki E, Gerstner W. Connectivity reflects coding: A model of voltage-based STDP with homeostasis. Nat Neurosci 2010, Mar;13(3):344-52.

77. Litwin-Kumar A, Doiron B. Formation and maintenance of neuronal assemblies through synaptic plasticity. Nat Commun 2014;5:5319.

78. Sadeh S, Clopath C, Rotter S. Emergence of functional specificity in balanced networks with synaptic plasticity. PLoS Comput Biol 2015, Jun;11(6):e1004307.

79. Pecka M, Han Y, Sader E, Mrsic-Flogel TD. Experience-Dependent specialization of receptive field surround for selective coding of natural scenes. Neuron 2014.

80. Rigotti M, Barak O, Warden MR, Wang X, Daw ND, Miller EK, Fusi S. The importance of mixed selectivity in complex cognitive tasks. Nature 2013, May 19;497(7451):585-90.

81. Raposo D, Kaufman MT, Churchland AK. A category-free neural population supports evolving demands during decision-making. Nat Neurosci 2014, Nov 10;17(12):1784-92.

82. Kampa, Roth, Göbel, Helmchen. Representation of visual scenes by local neuronal populations in layer 2/3 of mouse visual cortex. Front Neural Circuits 2011, Nov 23;5 (18).

83. Roth MM, Helmchen F, Kampa BM. Distinct functional properties of primary and posteromedial visual area of mouse neocortex. J Neurosci 2012, Jul 11;32(28):9716-26.

84. Stosiek C, Garaschuk O, Holthoff K, Konnerth A. In vivo two-photon calcium imaging of neuronal networks. Proc Natl Acad Sci U S A 2003, Jun 6;100(12):7319-24.

85. Nimmerjahn A, Kirchhoff F, Kerr JN, Helmchen F. Sulforhodamine 101 as a specific marker of astroglia in the neocortex in vivo. Nat Methods 2004, Oct;1(1):31-7.

86. Muir DR, Kampa BM. FocusStack and stimserver: A new open source MATLAB toolchain for visual stimulation and analysis of two-photon calcium neuronal imaging data. Frontiers in Neuroinformatics 2015, Jan 20;8:85.

87. Chen T-W, Wardill TJ, Sun Y, Pulver SR, Renninger SL, Baohan A, et al. Ultrasensitive fluorescent proteins for imaging neuronal activity. Nature 2013;499(7458):295-300.

88. Kleiner M, Brainard D, Pelli D, Ingling A, Murray R, Broussard C. What's new in psychtoolbox-3. Perception 2007;36(14):1-.

89. Ahmed B, Anderson JC, Douglas RJ, Martin KA, Whitteridge D. Estimates of the net excitatory currents evoked by visual stimulation of identified neurons in cat visual cortex. Cerebral Cortex 1998;8:462-76.

90. Hahnloser RHR. On the piecewise analysis of networks of linear threshold neurons. Neural Networks 1998;11:691-7.

91. Muir DR, Cook M. Anatomical constraints on lateral competition in columnar cortical architectures. Neural Computation 2014, Aug;26(8):1624-66.

92. Schüz A, Palm G. Density of neurons and synapses in the cerebral cortex of the mouse. Journal of Comparative Neurology 1989;286(4):442-55.

93. Hellwig B. A quantitative analysis of the local connectivity between pyramidal neurons in layers 2/3 of the rat visual cortex. Biol Cybern 2000;82(2):111-21.

94. Boucsein C, Nawrot MP, Schnepel P, Aertsen A. Beyond the cortical column: Abundance and physiology of horizontal connections imply a strong role for inputs from the surround. Front Neurosci 2011;5:32.

95. Holmgren C, Harkany T, Svennenfors B, Zilberter Y. Pyramidal cell communication within local networks in layer 2/3 of rat neocortex. The Journal of Physiology 2003, Aug;551(1):139-53.

96. Binzegger T, Douglas RJ, Martin KAC. Stereotypical bouton clustering of individual neurons in cat primary visual cortex. Journal of Neuroscience 2007, Nov 7;27(45):12242-54.

Tables

Table 1: Summary of nominal model parameters and model variables. Abbreviations: Exc: Excitatory; Inh: Inhibitory; Prop: proportion.

Parameter	Description	Nominal value
τ_i	Lumped neuron time constant for neuron *i*	10 ms
g_j	Nominal charge injected by synapses from neuron *j*	Exc.: 0.01 pC/spike/synapse Inh.:10×0.01 pC/spike/synapse
α_j	Nominal output gain of neuron *j*	0.066 spikes/pC
$n_{i,j}$	Number of synapses made from neuron *j* to neuron *i*	
β_j	Threshold of neuron *j*	Zero
$\sigma_i \cdot \zeta_i(t)$	Noise current injected into neuron *i*. Wiener process with std. dev. σ_i after 1 sec.	$\sigma_i = 5$ mA
N_N	Number of neurons in simulation	80,000 (10% of cortical density)
Prop. inh.	Proportion of inhibitory neurons	18%
	Dimensions of simulated torus space	2.2×2.2 mm
S_i	Nominal number of synapses made by neuron *i* (within superficial layers only)	Exc.: 8142 Inh.: 8566
$\sigma_{d,i}$	Std. Dev. of Gaussian dendritic field of neuron *i*	75μm (approx. width 300μm)
$\sigma_{a,i}$	Std. dev. of Gaussian axonal field of neuron *i*	Exc.: 290μm (approx. width 1100μm) Inh.: 100μm (approx. width 400μm)
κ_i	Input orientation tuning width parameter for neuron *i*	4
s_1	Degree of like-to-like modulation of anatomical connection probability	
s_2	Degree of feature-binding modulation of connection probability	
κ_1	Orientation tuning of like-to-like connection probability	
κ_2	Orientation tuning of subnetwork membership probability	
N_S	Number of subnetworks that exist at a point in cortex	
ϑ	Number of preferred orientations bound in an subnetwork	

Table 2: Parameter values used to specify large-scale network models.

Network configuration	Parameter values
Random connectivity model	$s_1 = 0,\ s_2 = 0$
Like-to-like specificity model	$s_1 = 0.8,\ s_2 = 0,\ \kappa_1 = 0.5$
Feature-binding specificity model	$s_1 = 0.1,\ s_2 = 0.25,\ \kappa_1 = 0.5,\ \kappa_2 = 4,\ N_S = 6,\ \vartheta = 2$

Figure captions

Figure 1: Like-to-like and feature-binding rules for local recurrent connectivity. a Connectivity scheme where local recurrent excitatory connections (between neurons grouped by dashed ovals) are matched to the feedforward visual preferences of the connected neurons ("like-to-like" over orientation preference, indicated by grating icons). **b** Connectivity scheme where local recurrent excitatory connections are different from the feedforward visual preferences of connected neurons ("feature-binding"). Exc.: excitatory.

Figure 2: Plaid responses are facilitatory and selective in mouse V1. (**a, b**) Two-photon calcium imaging of visual responses in mouse V1. (**b**) Average of 6 imaging frames. (**c**) Trial-averaged responses (7 trials) of a single neuron in mouse V1 to grating and (**d–f**) plaid stimuli of varying relative component orientations. **g–h** Response tuning of the same neuron in c–f. Neurons can be highly tuned to oriented gratings, and also highly selective for particular combinations of grating components. **i** When three sets of plaid stimuli with varying relative component angles are presented, the majority of neurons have facilitating responses to plaids (64% with MI>0.05). In response to 90º plaids alone however, neurons are more evenly split between facilitation and suppression (39% with MI>0.05; dashed). **j** Responses to combined plaid sets are significantly more selective than responses to 90º plaids alone, and significantly more selective than predicted plaid responses under a component response model [86]. *** $p<1\times10^{-10}$, Wilcoxon rank-sum test. OSI: orientation selectivity index; DSI: direction selectivity index; PSI: plaid selectivity index; MI: modulation index; facil.: facilitating; supp.: suppressing; prop.: proportion.

Figure 3: Connected neurons span a wide range of preferred orientations in mouse V1**. a** Characterisation of receptive field location using sparse drifting/rotating grating stimuli. Single-trial OGB calcium responses (black); presentation time of optimal stimulus and suboptimal stimulus indicated (black and grey bars). Right inset: estimated RF location for the same neuron. **b** Single-trial OGB calcium response to drifting grating stimuli (black); presentation of optimal stimulus orientation indicated above, all stimulus presentation times indicated below. Right inset: calculation of grating response similarity ρ_g between two neurons. **c** Single-trial (grey) and trial-averaged OGB calcium response (black) to natural movie stimuli. Vertical lines indicate timing of movie sequence onset. Right inset: calculation of movie response similarity (ρ_m), using signal correlations over trial-averaged responses from two neurons. **d** Pairs of neurons with high signal correlations to natural movies (ρ_m), which predicts a high probability of connection [21], can have similar or dissimilar grating responses. Pairs of neurons with similar orientation preference are not more likely to have high ρ_m (**e**) or high signal correlation to flashed natural scenes ρ_{Ca} (**f**) than pairs with dissimilar orientation preference. **g** Connected pairs are slightly more likely to share similar orientation preferences than unconnected pairs [21,24], but nevertheless span almost arbitrary orientation differences (≈20% of pairs with close to orthogonal orientation preference). d–e: *in vivo*

two-photon calcium imaging; f–g: *in vivo* calcium imaging coupled with *in vitro* simultaneous patching to detect connected pairs; data from [24]. e–f: Kruskal-Wallis tests; g: Ansari-Bradley test. n.s.: $p > 0.05$. Strong connections: strongest 50% of connected pairs, measured by EPSP amplitude. Corr: correlation; conn.: connection.

Figure 4: Rules for excitatory connectivity influence stimulus representations, and underlie amplification and competition. a In a simple model for random connectivity in mouse V1, injecting current into a single neuron (black outline) leads to non-specific activation of other excitatory (triangle) and inhibitory neurons (circle). Traces show the instantaneous firing rate of each neuron. **b** When the model is partitioned into subnetworks (SN1 & 2; dashed ovals), injecting current into a single neuron gives rise to an amplified response within the same subnetwork and suppresses activity in the non-driven subnetwork. **c** The degree of amplification and suppression depends directly on the proportion of excitatory synapses *s* restricted to be made within a subnetwork (see Fig. S1). Values of *s* used in panels **a–b** indicated on plot. **d** When local recurrent excitatory connections match the feedforward visual properties of connected neurons ("like-to-like"), grating responses (top) and plaid responses (bottom) are highly similar (high ρ_g & ρ_p). **e** In contrast, when local recurrent connections are different from the feedforward visual properties — in this case, grouping two different preferred orientations ("feature-binding") — then neurons with similar grating responses (top, high ρ_g) can have dissimilar plaid responses (bottom, low ρ_p), reflecting decorrelation of these responses caused by competition. Black outlines: stimulated neurons. Grating labels: preferred orientation of that neuron. Dashed ovals: neurons grouped by specific excitatory connectivity. a.u.: arbitrary units; prop.: proportion; syns.: synapses. Other conventions as in Fig. 1.

Figure 5: Rules for excitatory connectivity determine response correlation and decorrelation in a model of mouse V1. a–b In a large-scale network simulation incorporating like-to-like selective excitatory connectivity (connectivity rule and network schematic shown at left), responses of pairs of neurons to grating and plaid stimuli are always similar (**b**; similar ρ_g & ρ_p, high R^2). Traces: instantaneous firing rates for single example excitatory (black) and inhibitory (blue) neurons. Responses to grating stimuli are highly predictive of plaid responses; distribution of ρ_g versus ρ_p is therefore clustered around the diagonal (black line in **b**; high R^2). **c–d** When in addition to like-to-like connectivity, subnetworks also group neurons with several preferred orientations, then pairs of neurons with similar preferred orientations can respond differently to plaid stimuli, and vice versa (see response traces). **d** Competition due to feature-binding connectivity leads to decorrelation of the population response (low R^2). The distribution of ρ_g versus ρ_p is broad (black line in **d**), indicating poor predictability between grating and plaid responses. Inhibitory responses are broadly tuned in both models (blue traces in **a** & **b**). Pips in connectivity diagram in c indicate example preferred orientations of a single subnetwork. Conventions as in Fig. 1. Stim.: stimuli; a.u.: arbitrary units; corr.: correlation; feat. bind.: feature binding.

Figure 6: Responses to contrast-oscillating plaid and grating stimuli in mouse V1 suggest feature-binding connection rules. a Single-trial OGB calcium response to contrast-oscillating grating and plaid stimuli; presentation time of stimuli evoking strong responses indicated above trace. Right inset: measurement of plaid response similarity ρ_p between two neurons. **b** Trial-averaged responses (8 trials) of a pair of neurons from a single imaging site, with similar preferred orientations (polar plots at left; high ρ_g) but with dissimilar responses to plaid stimuli (low ρ_p). **c** Responses to grating and plaid stimuli are poorly related in orientation-tuned neurons in mouse V1 (Broad distribution of ρ_g versus ρ_p residuals—black line, low R^2). **d** Control data that includes experimental noise and response variability, obtained by resampling experimental responses and assuming a like-to-like connectivity rule (inset; see Methods), predicts a strong relationship between grating and plaid representations (high R^2) and is easily distinguished from observed V1 responses in c. **e** Decorrelation in mouse V1 is similar to the "feature-binding" model (F.B.), and much broader than the "like-to-like" model (L-to-L). **f** Responses to plaid stimuli in V1 are split between facilitating and suppressing (45% MI > 0.05; 42% MI < –0.05). **g** The distribution of facilitating (Facil.; MI > 0.05) and suppressing (Supp., MI < –0.05) responses is similar between mouse V1 and the "feature-binding" model (F.B.; p = 0.17, Fisher's exact test). The "like-to-like" and random non-specific (Rnd) connectivity models produced predominately suppressing responses. ***p < 0.001. $n_{V1} = 313$; $n_{F.B.} = 804$; $n_{L\text{-}to\text{-}L} = 729$; $n_{Rnd} = 729$; significantly responsive neurons with OSI > 0.3. Stim: stimuli; corr.: correlation; decorr.: decorrelation.

Figure 7: Non-random connectivity supports autoassociative behaviour. In a simple model with two subnetworks (**a**), presenting a linear graduated mixture between the ideal stimuli for the two subnetworks (**b**) results in competition and switching between network representations. When the stimulus is ideal for one subnetwork (mixture = 0% or 100%), then strong amplification of the network response occurs (compare with response of SN1 to a single grating component; arrowheads at right of b). When an approximately even mixture is presented (above and below 50%), the network switches rapidly from one representation to the other. Proportion of specific excitatory synapses $s = 25\%$. Dashed ovals: neurons grouped by specific excitatory connectivity. Other conventions as in Fig. 1. a.u.: arbitrary units.

Supplementary Figure 1: Estimated parameters for cortex place it in an Inhibition-Stabilised Network (ISN) regime, with competition provided by specific excitatory connectivity. a The network stability regimes in the parameter space defined by total inhibitory weight $g_I \cdot n_I$ and total excitatory weight $g_E \cdot n_E$ for a random network (proportion of specific synapses $s = 0\%$). Nominal parameter estimates for rodent cortex (cross) place the network in a regime that requires inhibitory feedback for stability (an ISN; [39]), but which does not lead

to competition between excitatory neurons. Inhibition must be unrealistically strengthened to obtain competition (100× and 200× estimates for rodent cortex; top of panel; shading indicates competition). However, overly-strong inhibition leads to inhibition-driven oscillations (IO). **b** When the proportion of specific synapses s is raised to 20%, nominal parameters for rodent cortex permit competition (shading indicates strength of competition; see Methods). Note that the maximum excitatory strength permitted while maintaining network stability is reduced. **c** When $s = 40\%$, nominal parameters for rodent cortex become unstable (cross is just outside stable region). **d** Network stability regimes for the parameter space defined by s and $g_E \cdot n_E$, with nominal value chosen for $g_I \cdot n_I$ (crosses in a–c). Nominal value for $g_E \cdot n_E$ is indicated by a dashed line. Both excitatory strength $g_E \cdot n_E$ and the proportion of specific synapses s affect network stability and the strength of competition. Abbreviations: $g_{I,E}$: Synaptic strength per inhibitory or excitatory synapse; $n_{I,E}$: Number of synapses made by each inhibitory or excitatory neuron; AS: Intrinsically stable network, stable in the absence of inhibition; ISN: Inhibition-Stabilised Network, requiring inhibitory feedback for stability; Exp: Runaway activity due to exponentially divergent unstable fixed point; IO: Oscillatory activity due to strong inhibition. a.u.: arbitrary units.

Supplementary Figure 2: Grating and plaid responses are highly correlated in a model with random connectivity. **a** Under the non-specific connectivity model, synapses between pairs of neurons are formed without regard to functional response similarity of the neurons. Neurons form synapses stochastically, according to spatial proximity. Two example pairs of neurons are shown, and their responses to a set of grating and plaid stimuli. **b** Neurons with similar responses to grating stimuli (high ρ_g) always have similar responses to plaid stimuli (high ρ_p), and vice versa. Conn.: connectivity; stim.: stimuli.

Figure 1

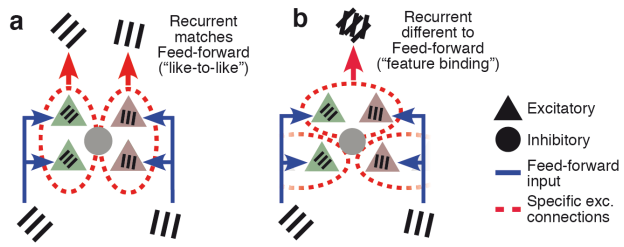


Figure 1: Like-to-like and feature-binding rules for local recurrent connectivity. **a** Connectivity scheme where local recurrent excitatory connections (between neurons grouped by dashed ovals) are matched to the feedforward visual preferences of the connected neurons ("like-to-like" over orientation preference, indicated by grating icons). **b** Connectivity scheme where local recurrent excitatory connections are different from the feedforward visual preferences of connected neurons ("feature-binding"). Exc.: excitatory.

Figure 2

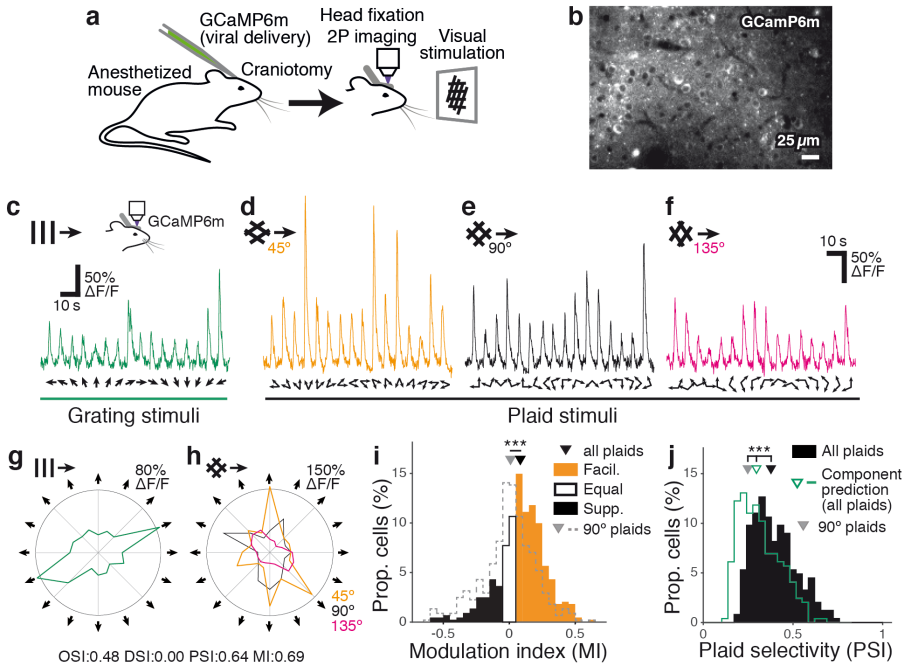


Figure 2: Plaid responses are facilitatory and selective in mouse V1. (a, b) Two-photon calcium imaging of visual responses in mouse V1. (**b**) Average of 6 imaging frames. (**c**) Trial-averaged responses (7 trials) of a single neuron in mouse V1 to grating and (**d–f**) plaid stimuli of varying relative component orientations. **g–h** Response tuning of the same neuron in c–f. Neurons can be highly tuned to oriented gratings, and also highly selective for particular combinations of grating components. **i** When three sets of plaid stimuli with varying relative component angles are presented, the majority of neurons have facilitating responses to plaids (64% with MI > 0.05). In response to 90° plaids alone however, neurons are more evenly split between facilitation and suppression (39% with MI > 0.05; dashed). **j** Responses to combined plaid sets are significantly more selective than responses to 90° plaids alone, and significantly more selective than predicted plaid responses under a component response model [86]. *** $p < 1 \times 10-10$, Wilcoxon rank-sum test. OSI: orientation selectivity index; DSI: direction selectivity index; PSI: plaid selectivity index; MI: modulation index; facil.: facilitating; supp.: suppressing; prop.: proportion.

Figure 3

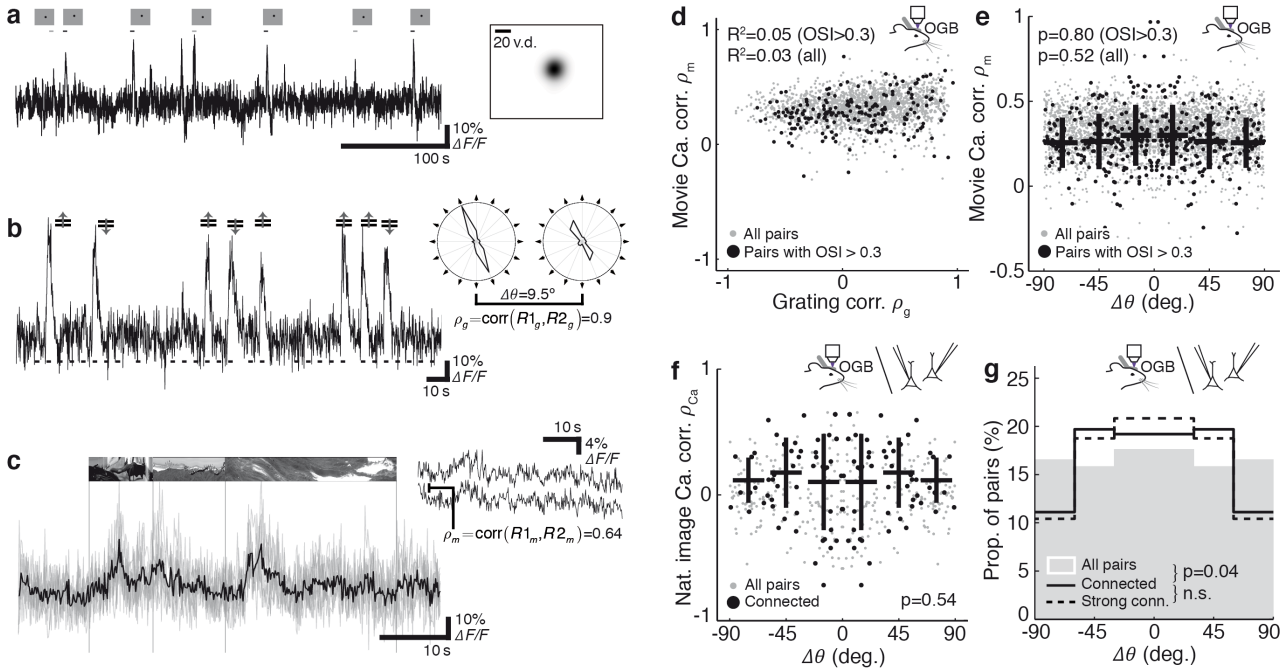


Figure 3: Connected neurons span a wide range of preferred orientations in mouse V1. a Characterisation of receptive field location using sparse drifting/rotating grating stimuli. Single-trial OGB calcium responses (black); presentation time of optimal stimulus and sub-optimal stimulus indicated (black and grey bars). Right inset: estimated RF location for the same neuron. **b** Single-trial OGB calcium response to drifting grating stimuli (black); presentation of optimal stimulus orientation indicated above, all stimulus presentation times indicated below. Right inset: calculation of grating response similarity ρ_g between two neurons. **c** Single-trial (grey) and trial-averaged OGB calcium response (black) to natural movie stimuli. Vertical lines indicate timing of movie sequence onset. Right inset: calculation of movie response similarity (ρ_m), using signal correlations over trial-averaged responses from two neurons. **d** Pairs of neurons with high signal correlations to natural movies (ρ_m), which predicts a high probability of connection [21], can have similar or dissimilar grating responses. Pairs of neurons with similar orientation preference are not more likely to have high ρ_m (**e**) or high signal correlation to flashed natural scenes ρ_{Ca} (**f**) than pairs with dissimilar orientation preference. **g** Connected pairs are slightly more likely to share similar orientation preferences than unconnected pairs [21,24], but nevertheless span almost arbitrary orientation differences (≈20% of pairs with close to orthogonal orientation preference). d–e: in vivo two-photon calcium imaging; f–g: in vivo calcium imaging coupled with in vitro simultaneous patching to detect connected pairs; data from [24]. e–f: Kruskal-Wallis tests; g: Ansari-Bradley test. n.s.: $p > 0.05$. Strong connections: strongest 50% of connected pairs, measured by EPSP amplitude. Corr: correlation; conn.: connection.

Figure 4

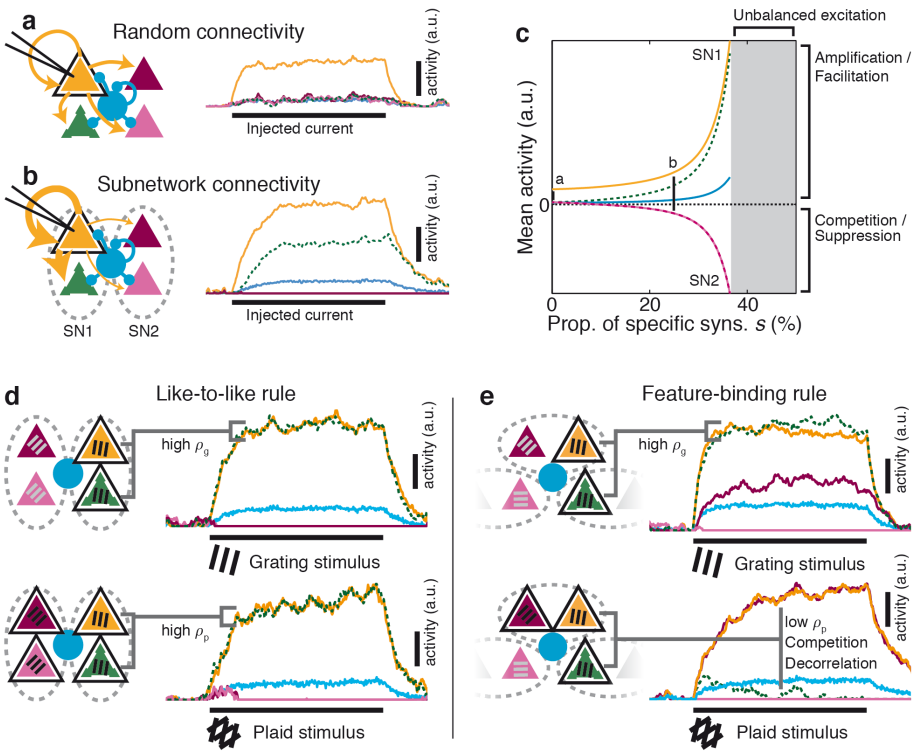


Figure 4: Rules for excitatory connectivity influence stimulus representations, and underlie amplification and competition. a In a simple model for random connectivity in mouse V1, injecting current into a single neuron (black outline) leads to non-specific activation of other excitatory (triangle) and inhibitory neurons (circle). Traces show the instantaneous firing rate of each neuron. **b** When the model is partitioned into subnetworks (SN1 & 2; dashed ovals), injecting current into a single neuron gives rise to an amplified response within the same subnetwork and suppresses activity in the non-driven subnetwork. **c** The degree of amplification and suppression depends directly on the proportion of excitatory synapses *s* restricted to be made within a subnetwork (see Fig. S1). Values of *s* used in panels a–b indicated on plot. **d** When local recurrent excitatory connections match the feedforward visual properties of connected neurons ("like-to-like"), grating responses (top) and plaid responses (bottom) are highly similar (high ρ_g & ρ_p). **e** In contrast, when local recurrent connections are different from the feedforward visual properties — in this case, grouping two different preferred orientations ("feature-binding") — then neurons with similar grating responses (top, high ρ_g) can have dissimilar plaid responses (bottom, low ρ_p), reflecting decorrelation of these responses caused by competition. Black outlines: stimulated neurons. Grating labels: preferred orientation of that neuron. Dashed ovals: neurons grouped by specific excitatory connectivity. a.u.: arbitrary units; prop.: proportion; syns.: synapses. Other conventions as in Fig. 1.

Figure 5

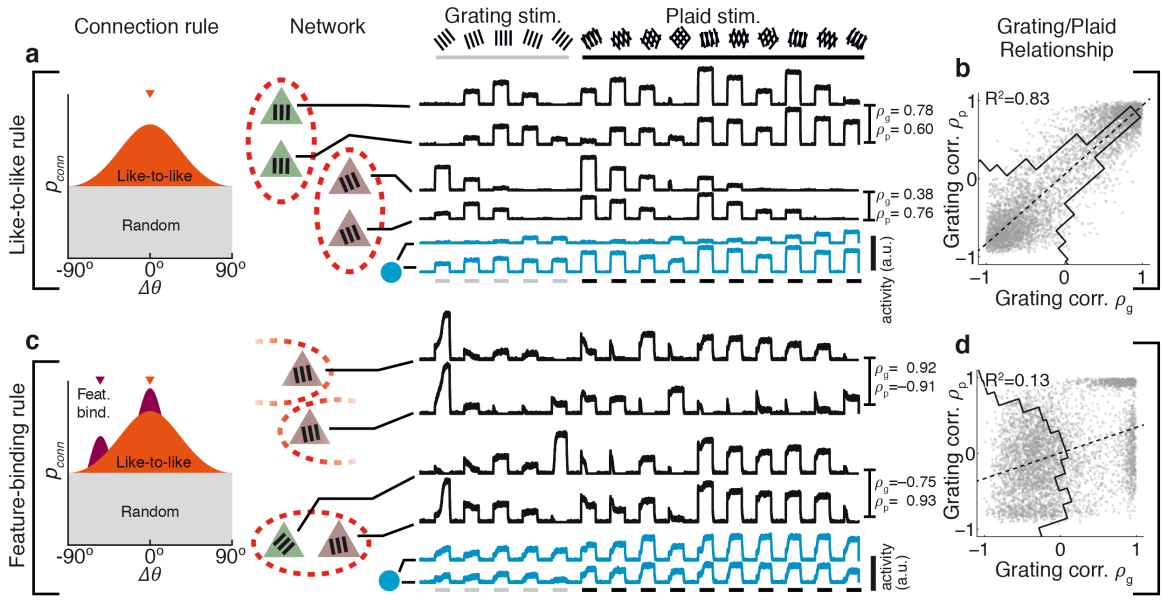


Figure 5: Rules for excitatory connectivity determine response correlation and decorrelation in a model of mouse V1. a–b In a large-scale network simulation incorporating like-to-like selective excitatory connectivity (connectivity rule and network schematic shown at left), responses of pairs of neurons to grating and plaid stimuli are always similar (**b**; similar ρ_g & ρ_p, high R^2). Traces: instantaneous firing rates for single example excitatory (black) and inhibitory (blue) neurons. Responses to grating stimuli are highly predictive of plaid responses; distribution of ρ_g versus ρ_p is therefore clustered around the diagonal (black line in b; high R2). **c–d** When in addition to like-to-like connectivity, subnetworks also group neurons with several preferred orientations, then pairs of neurons with similar preferred orientations can respond differently to plaid stimuli, and vice versa (see response traces). **d** Competition due to feature-binding connectivity leads to decorrelation of the population response (low R^2). The distribution of ρ_g versus ρ_p is broad (black line in d), indicating poor predictability between grating and plaid responses. Inhibitory responses are broadly tuned in both models (blue traces in a & b). Pips in connectivity diagram in c indicate example preferred orientations of a single subnetwork. Conventions as in Fig. 1. Stim.: stimuli; a.u.: arbitrary units; corr.: correlation; feat. bind.: feature binding.

Figure 6

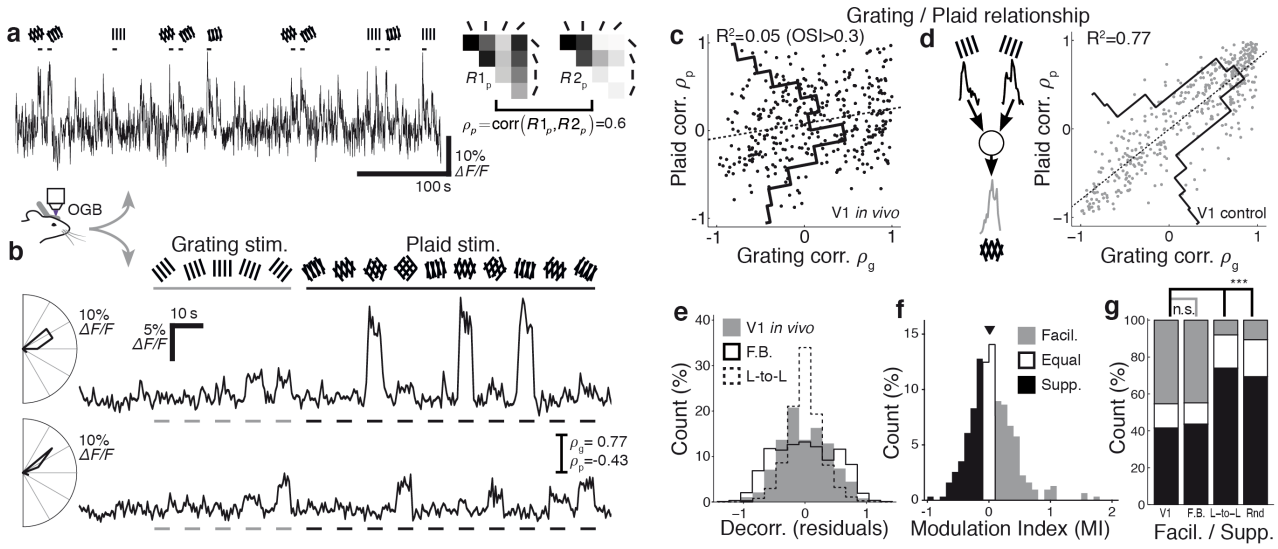


Figure 6: Responses to contrast-oscillating plaid and grating stimuli in mouse V1 suggest feature-binding connection rules. a Single-trial OGB calcium response to contrast-oscillating grating and plaid stimuli; presentation time of stimuli evoking strong responses indicated above trace. Right inset: measurement of plaid response similarity ρ_p between two neurons. **b** Trial-averaged responses (8 trials) of a pair of neurons from a single imaging site, with similar preferred orientations (polar plots at left; high ρ_g) but with dissimilar responses to plaid stimuli (low ρ_p). c Responses to grating and plaid stimuli are poorly related in orientation-tuned neurons in mouse V1 (Broad distribution of ρ_g versus ρ_p residuals — black line, low R2). **d** Control data that includes experimental noise and response variability, obtained by resampling experimental responses and assuming a like-to-like connectivity rule (inset; see Methods), predicts a strong relationship between grating and plaid representations (high R2) and is easily distinguished from observed V1 responses in c. **e** Decorrelation in mouse V1 is similar to the "feature-binding" model (F.B.), and much broader than the "like-to-like" model (L-to-L). **f** Responses to plaid stimuli in V1 are split between facilitating and suppressing (45% MI > 0.05; 42% MI < –0.05). **g** The distribution of facilitating (Facil.; MI > 0.05) and suppressing (Supp., MI < –0.05) responses is similar between mouse V1 and the "feature-binding" model (F.B.; p = 0.17, Fisher's exact test). The "like-to-like" and random non-specific (Rnd) connectivity models produced predominately suppressing responses. ***p < 0.001. nV1 = 313; nF.B. = 804; nL-to-L = 729; nRnd = 729; significantly responsive neurons with OSI > 0.3. Stim: stimuli; corr.: correlation; decorr.: decorrelation.

Figure 7

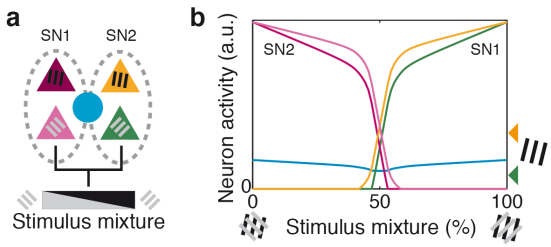
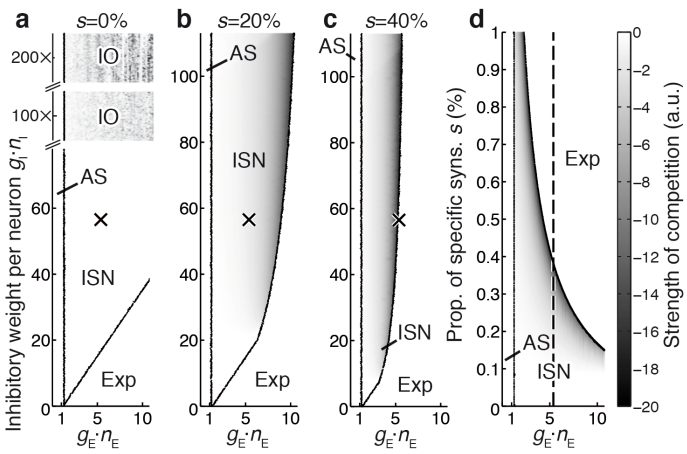


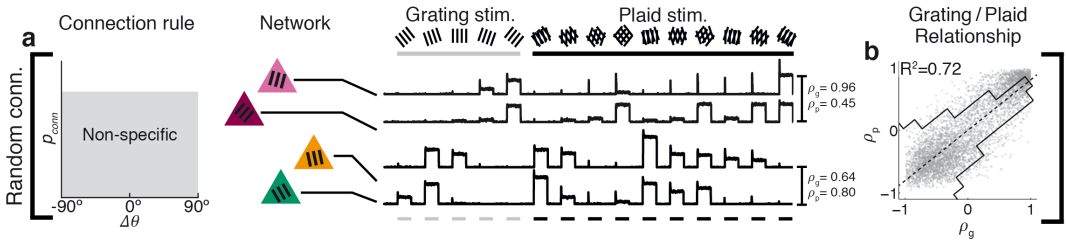
Figure 7: Non-random connectivity supports autoassociative behaviour. In a simple model with two subnetworks (**a**), presenting a linear graduated mixture between the ideal stimuli for the two subnetworks (**b**) results in competition and switching between network representations. When the stimulus is ideal for one subnetwork (mixture = 0% or 100%), then strong amplification of the network response occurs (compare with response of SN1 to a single grating component; arrowheads at right of b). When an approximately even mixture is presented (above and below 50%), the network switches rapidly from one representation to the other. Proportion of specific excitatory synapses s = 25%. Dashed ovals: neurons grouped by specific excitatory connectivity. Other conventions as in Fig. 1. a.u.: arbitrary units.

Supp. Figure 1



Supplementary Figure 1: Estimated parameters for cortex place it in an Inhibition-Stabilised Network (ISN) regime, with competition provided by specific excitatory connectivity. a The network stability regimes in the parameter space defined by total inhibitory weight $g_I \cdot n_I$ and total excitatory weight $g_E \cdot n_E$ for a random network (proportion of specific synapses s = 0%). Nominal parameter estimates for rodent cortex (cross) place the network in a regime that requires inhibitory feedback for stability (an ISN; [39]), but which does not lead to competition between excitatory neurons. Inhibition must be unrealistically strengthened to obtain competition (100× and 200× estimates for rodent cortex; top of panel; shading indicates competition). However, overly-strong inhibition leads to inhibition-driven oscillations (IO). **b** When the proportion of specific synapses s is raised to 20%, nominal parameters for rodent cortex permit competition (shading indicates strength of competition). Note that the maximum excitatory strength permitted while maintaining network stability is reduced. **c** When s = 40%, nominal parameters for rodent cortex become unstable (cross is just inside unstable region). **d** Network stability regimes for the parameter space defined by s and $g_E \cdot n_E$, with nominal value chosen for $g_I \cdot n_I$ (crosses in a–c). Nominal value for $g_E \cdot n_E$ is indicated by a dashed line. Both excitatory strength $g_E \cdot n_E$ and the proportion of specific synapses s affect network stability and the strength of competition. Abbreviations: $g_{I,E}$: Synaptic strength per inhibitory or excitatory synapse; $n_{I,E}$: Number of synapses made by each inhibitory or excitatory neuron; AS: Intrinsically stable network, stable in the absence of inhibition; ISN: Inhibition-Stabilised Network, requiring inhibitory feedback for stability; Exp: Runaway activity due to exponentially divergent unstable fixed point; IO: Oscillatory activity due to strong inhibition.

Supp. Figure 2



Supplementary Figure 2: Grating and plaid responses are highly correlated in a model with random connectivity. a Under the non-specific connectivity model, synapses between pairs of neurons are formed without regard to functional response similarity of the neurons. Neurons form synapses stochastically, according to spatial proximity. Two example pairs of neurons are shown, and their responses to a set of grating and plaid stimuli. **b** Neurons with similar responses to grating stimuli (high ρ_g) always have similar responses to plaid stimuli (high ρ_p), and vice versa.