Specific excitatory connectivity for feature integration in mouse primary visual cortex

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

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

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

3. Department of Neurophysiology, Institute of Biology 2, RWTH Aachen University, Templergraben 55, 52062 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 61 267 17 75
Fax: +41 61 267 21 89
Email: dylan.muir@unibas.ch

The corresponding author is an early-career researcher.

In mouse primary visual cortex (V1), local excitatory connections are more prevalent, stronger and larger between neurons that share similar functional response features. However, the extent to which rules for local cortical connection specificity shape visual responses, as well as full details relating structure and function both remain unknown. We considered whether complex responses to plaid stimuli in mouse V1 could be explained by one of two alternative connectivity schemes: whether local connections are aligned with simple feedforward visual properties, or whether local connections group across feedforward visual properties. Using a combined experimental and computational approach, we found that responses to plaid stimuli in mouse V1 were best explained by a connectivity scheme which binds multiple feedforward visual properties. Our results show that feature binding can occur through a recurrent mechanism not requiring feedforward convergence; such a mechanism is consistent with visual responses in mouse V1. [150/150 words]

Much of our current understanding of local cortical connectivity is based on the presumption of randomness. Anatomical methods for estimating connection probabilities **(Binzegger et al. 2004; Braitenberg and Schüz 1991)** and techniques for using anatomical reconstructions to build models of cortical circuits **(Hill et al. 2012; Binzegger et al. 2009; Ramaswamy et al. 2012; Markram et al. 2015; Reimann et al. 2015)** are largely based on the assumption that connections between nearby neurons are made stochastically in proportion to the overlap between axonal and dendritic arborisations **(Peters 1979)**.

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 visual cortex, long-range intrinsic excitatory connections tend to connect regions of similar function **(Juliano et al. 1990; Malach et al. 1993; Bosking et al. 1997; Muir et al. 2011; Martin et al. 2014)**. Rodents exhibit a mapless "salt and pepper" arrangement of function across cortex **(Ohki et al. 2005)**, but non-random connectivity is nevertheless prevalent both

within and between cortical layers **(Yoshimura et al. 2005; Yoshimura and Callaway 2005; Perin et al. 2011; Kampa et al. 2006; Yu et al. 2012)**, and reflects similarities in functional properties **(Ko et al. 2011; Ko et al. 2013; Li et al. 2012; Cossell et al. 2015; Lee et al. 2016)** or projection targets **(Brown and Hestrin 2009; Morishima et al. 2011)**. Considerable non-random structure has also been described in patterns of anatomical connectivity across several species **(Sporns and Kötter 2004; Song et al. 2005)**.

Thus, specificity of cortical connections among excitatory neurons is an important feature of local circuitry, and is likely to be influential in determining the functional response properties of cortical neurons **(Cossell et al. 2015; Muir and Mrsic-Flogel 2015)**. However, the impact of specific excitatory connectivity on network representations of sensory inputs and information processing has not been addressed experimentally or through theory. It remains an open question how the arrangement of local recurrent connections affect cortical representations.

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 **(Cossell et al. 2015)**. This sea of weak synaptic inputs might be responsible for feature non-specific depolarisation **(Cossell et al. 2015)** or might permit plasticity of network function **(Song et al. 2005)**. However, another possibility is that weaker connections underly higher-order connectivity rules that have not yet been described.

For example, recent results show that responses to compound visual stimuli (e.g. plaid stimuli composed of two grating components) can be selective and highly complex in mouse V1 **(Muir et al. 2015)**. This could be explained by rules for local excitatory connections in cortex that are not simply tuned to feedforward response properties, but which specifically group neurons with different preferred

orientations. Alternatively, local connections might be aligned with feedforward response properties but be broadly tuned, such that many synapses are made between neurons with weakly similar responses.

We explored these two alternatives by simulating large networks with local connectivity rules that either aligned with feedforward response properties, or differed from feedforward responses. We then tested predictions from these models in mouse V1, by recording responses to grating, plaid and natural stimuli. We found that the complexity of plaid responses in mouse V1 was reproduced when local connections cut across feedforward response properties, by grouping neurons with different preferred orientations.

Our results 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

We considered that the configuration of local recurrent connections in cortex might differently process simple and compound visual stimuli. It is therefore important to understand 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 **(Juavinett and Callaway 2015; Muir et al. 2015)**. Most neurons in mouse V1 respond to plaid stimuli as a simple super-

imposition of their response to the two underlying grating components (i.e. "component cell" responses; **(Movshon et al. 1985)**). 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 **(Muir et al. 2015)**. Plaid stimuli are often constructed from a single choice of relative component angle (90º orthogonal gratings), leaving open the possibility that a richer set of plaid stimuli would help to classify neurons with these complex responses.

Accordingly, we probed mouse V1 with grating component stimuli composed of grating stimuli with 16 drift directions, and three full sets of plaid stimuli composed of 45º, 90º and 135º relative grating component orientations. We recorded responses from layer 2/3 neurons using two-photon imaging of animals with viral delivery of GCaMP6m (Fig. 1a–f; 8 animals, 8 sessions, 441/879 responsive/imaged neurons; 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 90º plaids (Fig. 1g; 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 CI unless stated otherwise). Responses to the full set of plaid stimuli were highly selective; significantly more selective than predicted by a component model (Fig. 1h; 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 90º plaids alone (Fig. 1h; median 90º 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 does not make responses to compound stimuli more comprehensible — instead, responses are more facilitatory and more selective. This suggests that using more plaid combinations gives a more accurate characterisation of the response properties of individual neurons.

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

Local excitatory connections in cortex are broadly tuned for preferred orientation

How are selective responses to plaid stimuli generated in V1? One possibility is that specific grating component representations are combined through local excitatory connectivity **(Muir et al. 2015)**. Synaptic connection probability in mouse V1 is enhanced by similarity of orientation preference **(Ko et al. 2011; Li et al. 2012; Lee et al. 2016)**, 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 **(Ko et al. 2011; Ko et al. 2013; Cossell et al. 2015)**.

We recorded responses to natural movie and drifting grating stimuli in populations of neurons in mouse V1 (5 animals, 129/391 responsive neurons with overlapping receptive fields/imaged neurons; see Fig. 7b–d). We found that neurons with high correlations to natural scenes, which are most likely to be connected in cortex **(Ko et al. 2011; Ko et al. 2013; Cossell et al. 2015)**, showed only a weak tendency to share similar orientation preferences (Fig. 2a–b; pairs with OSI > 0.3; p=0.39, Kuiper's test). This is consistent with earlier findings in cat area 17 (V1), which showed a poor relationship between responses to gratings and natural movies **(Martin and Schröder 2013)**.

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

We compared response correlations and preferred orientation between pairs of neurons which were known to be connected, from *in vivo*/*in vitro* characterisation of function and connectivity between neurons in mouse V1 (data from **(Cossell et al. 2015)** used with permission; 17 animals, 203 patched and imaged cells, 75 connections). 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. 2c; p=0.54, Kuiper's test). Also consistent with earlier findings **(Ko et al. 2011; Li et al. 2012)**, we observed a relationship between synaptic connectivity and similarity of orientation preference (Fig. 2d; p=0.045, Ansari-Bradley test). However, strongly connected pairs (strongest 50% of EPSPs over connected pairs) were not more similar in their preferred orientation than more weakly connected pairs (p=0.17, Ansari-Bradley test). 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 preferred orientation, but a majority of synaptic inputs are nevertheless formed between neurons with poorly- or un-correlated responses **(Cossell et al. 2015)**.

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. 2e; 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. 2d and elsewhere **(Ko et al. 2011; Cossell et al. 2015)** 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. 2f; 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.

The statistics of subnetwork connectivity shape cortical representations

We first explored the conceptual consequences of specific excitatory connectivity rules 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"). This small model consisted of four excitatory and one inhibitory neuron with homogenous connectivity, designed to be equivalent to a much larger model with stochastic synaptic connectivity (see Methods). As is expected to be the case for cortical circuitry, our model required inhibitory feedback to maintain stability (an inhibition-stabilised network or ISN; Figure 3—Figure supplement 1; **(Tsodyks et al. 1997; Ozeki et al. 2009)**; but see **(Atallah et al. 2012)**).

We found that non-random connectivity, in the form of specific excitatory connections within subnetworks **(Yoshimura et al. 2005; Kampa et al. 2006)**, introduced selective amplification within subnetworks and competition between subnetworks (Fig. 3). Surprisingly, these computational mechanisms could be strongly expressed even when only relatively small proportions of synapses were made to be subnetwork-specific, but neither selective amplification nor competition were present without specific synaptic connectivity (Fig. 3c; Figure 3—Figure supplement 1).

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

We next 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 a "like-to-like" rule, neurons with similar preferred orientations were grouped into subnetworks (Fig. 4a–b). 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. 4a–b; "like-to-like"). Amplification occurred within subnetworks of neurons with the same preferred orientation, and competition between subnetworks with differing preferred orientation **(Douglas et al. 1994; Sadeh et al. 2015)**.

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

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. 4c–d). When this "feature-binding" connectivity rule was applied, neuronal responses to grating and plaid stimuli differed markedly (Fig. 4c vs d). 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. 4c; high ρ_g), the same two neurons would show decorrelated responses to a plaid stimulus (Fig. 4d; low ρ_p).

Functional differences in connectivity statistics are detectable in large networks

The results of our simulations of small networks suggest that rules for non-random local connectivity can have a profound influence on the pattern of network activation following an external stimulus. The question remains whether the

differences in representation induced by "like-to-like" compared with "feature-binding" connection rules will be detectable in large networks with realistic structure, and in visual cortex *in vivo*.

We therefore investigated how the type of connection specificity affects stimulus representations 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; **(Gabott and Somogyi 1986)**; 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 **(Ermentrout 1998)**; model neurons with linear-threshold dynamics can be directly translated into integrate-and-fire models with more complex dynamics **(Neftci et al. 2011; Neftci et al. 2013)**, and in addition form good approximations to conductance-based neuron models **(Shriki et al. 2003)**.

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 **(Ohki et al. 2005)**. 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 computed as for experimental responses; Fig. 7c, e).

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

**** 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. 5e–h), neurons showed decorrelation in response to plaid stimuli (Fig. 5h, R^2=0.13 between ρ_g and ρ_p).

Consistent with our analytical models, networks including random excitatory connectivity without any specificity did not give rise to decorrelation (Figure 5—Figure supplement 1d; R^2=0.72 between ρ_g and ρ_p). Inhibitory responses were untuned in our simulations (Fig. 5c, g), in agreement with experimental observations of poorly-tuned inhibition in mouse V1 **(Bock et al. 2011; Hofer et al. 2011; Kerlin et al. 2010; Liu et al. 2009)**.

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.

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

Our analytical network results show that in principle, whether local excitatory connections align with feedforward visual properties or span across feedforward visual properties, has a drastic effect on visual representations (Fig. 4). Our large-scale simulations show that these effects can be detected in large networks as differences in pairwise responses to simple and compound visual stimuli (Fig. 5). Responses to plaid stimuli in mouse V1 suggest that a stimulating with a denser sampling of compound stimulus space leads to a better characterisation of response selectivity (Fig. 1). 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.

Using two-photon calcium imaging, we recorded responses of populations of OGB-labelled neurons in mouse V1 to a set of contrast-modulated 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. 6; 5 animals, 5 sessions, 313 / 543 responsive / imaged neurons).

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

We found that consistent with previous reports **(Muir et al. 2015)**, 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. 6a, b; R^2=0.05 between ρ_g and ρ_p; OSI > 0.3). The degree of decorrelation we observed in mouse V1 was consistent with our "feature-binding" model, and considerably higher than predicted by the "like-to-like" model (Fig. 6d).

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 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. 6c, inset; see Methods). Our control data were generated from single-neuron, single-trial responses collected from mouse V1, and therefore included the same trial-to-trial variability as exhibited by cortex. This control analysis indicates that if neurons connected only according to a "like-to-like" rule, cortex would exhibit highly related grating and plaid responses (Fig. 6c; median $R^2 = 0.77 \pm [0.767\ 0.775]$ between ρ_g and ρ_p; n = 2000 bootstrap samples) which are very different to the decorrelated responses we observed experimentally ($p < 0.005$, Monte-Carlo test).

Importantly, this control analysis is not restricted to our "like-to-like" rule, but makes similar predictions of highly related 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 **(Muir et al. 2015)**. 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 our "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. 6e, f; 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" (F.B.) model (Fig. 6f; 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. 6f; $p < 0.001$ in both cases; two-tailed Fisher's exact test, $n_{L\text{-}to\text{-}L} = 729$, $n_{Rnd} = 729$).

Discussion

Whereas feedforward mechanisms for building receptive fields in visual cortex have been extensively studied, it is not well understood how receptive fields 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. 2e), or whether local excitatory connections

span across feedforward visual properties to group them ("feature-binding"; Fig. 2f). We found that highly selective and facilitatory responses to plaid stimuli observed in mouse V1 (Fig. 1, Fig. 6; **(Muir et al. 2015)**) are consistent with tuning of recurrent connections within small cohorts of neurons to particular plaid combinations. Moreover, responses in mouse V1 are inconsistent with a simple configuration of cortical connections strictly aligned with feedforward visual responses.

Amplification and competition; 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 and competition between subnetworks of excitatory neurons (Fig. 3). Several recent studies have demonstrated that visual input is amplified within the superficial layers of cortex **(Li et al. 2013; Lien and Scanziani 2013; Li et al. 2013)**, and recent results from motor cortex suggest competition between ensembles of neurons **(Zagha et al. 2015)**. Our modelling results indicated that some form non-random local excitatory connectivity is required for such amplification to occur through recurrent network interactions, under reasonable estimates of anatomical and physiological parameters for rodent cortex (Fig. 3; Figure 3—Figure supplement 1). 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 (Figure 6—Figure supplement 1). In our "feature-binding" model, the statistics of subnetwork connectivity were well described by plaid stimuli. As a result, plaid stimuli give rise to stronger amplification than single grating components alone, if the composition of the plaid matches the composition of a particular subnetwork. This leads to a facilitatory effect, where some neurons show stronger responses to plaid stimuli than to the grating com-

ponents underlying the plaid stimuli. Conversely, competition between subnetworks leads to weaker responses to some plaid stimuli, for neurons that "lose" the competition. Competition could therefore be one cortical mechanism underlying cross-orientation suppression in response to plaid stimulation.

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 (Figure 5—Figure supplement 1).

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 **(Kampa et al. 2006; Ko et al. 2011; Bock et al. 2011; Li et al. 2012; Ko et al. 2013; Cossell et al. 2015)**. Unfortunately, current technical limitations make it difficult to measure more complex statistical structures such as present under a "feature-binding" con-

nectivity 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 **(Perin et al. 2011)**, 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 20% of synapses are made within an excitatory subnetwork is sufficient to induce strong specific amplification and strong competition to the network, even though 80% of the synapses are made randomly (Fig. 3). 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 approached this question using a novel pre-synaptic labelling technique to functionally characterise the presynaptic inputs to single superficial-layer neurons in mouse V1 **(Wertz et al. 2015)**. Consistent with our results for preferred orientation (Fig. 2c, d), 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 maximally strong 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 **(Movshon et al. 1985; Rust et al. 2006)**? 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 **(Niell and Stryker 2008; Medini 2011)**, 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. 2a; Fig. 6a, b; **(Muir et al. 2015)**).

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 **(Mariño et al.**

2005), and consistent with experimental observations of weakly- or un-tuned inhibition in rodent visual cortex **(Liu et al. 2009; Kerlin et al. 2010; Bock et al. 2011; Hofer et al. 2011; Atallah et al. 2012)**. Although specific E↔I connectivity has been observed in rodent cortex **(Yoshimura and Callaway 2005)**, the majority of E↔I synapses are likely to be made non-specifically in line with the high convergence of E→I and I→E connections observed in cortex **(Fino and Yuste 2011; Bock et al. 2011; Hofer et al. 2011)**.

In our models, shared inhibition is crucial to mediate competition between excitatory subnetworks (Fig. 4; Fig. 3); 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 classes in cortex (e.g. somatostatin or vaso-intestinal peptide expressing neurons) have been shown to exhibit feature-selective responses **(Ma et al. 2010; Kerlin et al. 2010; Wilson et al. 2012)**. Recent computational work examined the influence of multiple inhibitory classes with different physiological and anatomical tuning properties in a model for rodent cortex, including a network with specific, orientation-tuned inhibitory connectivity, examining the effects on divisive and subtractive normalisation of network activity **(Litwin-Kumar et al. 2016)**. 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 played by 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.

Development of non-random connectivity is likely to embed natural scene statistics from visual experience

In visual cortex of primates, carnivores and rodents, orientation tuning develops before postnatal eye opening and in the absence of visual experience **(White and Fitzpatrick 2007; Rochefort et al. 2011)**. Local recurrent connections develop after the onset of visual experience and maintain their plasticity into adulthood **(White and Fitzpatrick 2007; Galuske and Singer 1996; Luhmann et al. 1986; Luhmann et al. 1990; Katz and Callaway 1992; Miller et al. 2001)**. Statistical correlations in natural scenes might therefore lead to wiring of subnetworks under an activity-dependent mechanism such as spike-time dependent plasticity (STDP) **(Kampa et al. 2007; Markram et al. 2012; Clopath et al. 2010; Litwin-Kumar and Doiron 2014; Sadeh et al. 2015)**. Along these lines, recent examinations of the development of specific excitatory connections after eye opening found that similarities in feedforward input were progressively encoded in specific excitatory connections **(Ko et al. 2013)**.

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 **(Pecka et al. 2014)**. 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. This network architecture would then permit cortex to combine relevant components of the complex and ambiguous natural environment, and recover an interpretation that is consistent with learned statistical properties of the sensory world.

In pre-frontal cortex, compound or mixed selectivity of single neurons to combinations of task-related responses occurs often, and facilitates efficient decoding of arbitrary decision-related variables **(Rigotti et al. 2013; Raposo et al. 2014)**. 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 Zurich 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 **(Kampa et al. 2011; Roth et al. 2012; Cossell et al. 2015; Muir et al. 2015)**.

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 iso-

tonic 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 (Fig. 7a; **(Stosiek et al. 2003)**). Sulforhodamine (SR–101; Invitrogen) was applied topically to the pial surface, to provide labelling of the astrocytic network **(Nimmerjahn et al. 2004)**. 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 **(Muir and Kampa 2015)**.

*** FIGURE 7 NEAR HERE ***

Preparation and imaging with GCaMP6 Adult male mice (P75–P90) were initially anesthetized 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 **(Chen et al. 2013)** (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 several days after implantation, animals were handled to reduce stress during experiments.

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 **(Cossell et al. 2015)**. 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.; Fig. 7a). 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 **(Kleiner et al. 2007; Muir and Kampa 2015)**.

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. 7b). 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. 7c). 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. 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). Full-field natural movies consisted of a 43 s continuous sequence with three segments (example calcium response shown in Fig. 7d). 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. 7e). 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-oscillating 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 (Muir and Kampa 2015). 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. 7b). 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. 7d.

Additional responses were recorded while stimulating with contrast-oscillating grating and plaid stimuli. Since stationary stimuli were used, signals were less robust than for drifting high-contrast grating stimuli. An example single-trial calcium response is shown in Fig. 7e.

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. 7c, 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. 7e, 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. 7d, 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. 2c) was measured as the linear correlation between the vector of responses of a single neuron to a set of 1800 natural stimuli.

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 **(Muir et al. 2015)**. The PSI of a neuron ranges 0..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 **(Muir et al. 2015)**. The MI of a neuron ranges −1..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. 3, Figure 3—Figure supplement 1, Figure 6—Figure supplement 1). 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

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

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 $g_j \cdot n_{i,j} \cdot \alpha_j$, 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 **(Binzegger et al. 2009)**. Excitatory and inhibitory neurons were assigned output gains of 0.066 spikes/pC **(Ahmed et al. 1998)**.

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, Fig. 3). 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 excitat-

ory 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 and Fig. 3. 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 ($I_i(t) = 0$).

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 **(Hahnloser 1998; Muir and Cook 2014)**. Either of these modes is undesirable for cortex.

To determine whether the parameter regimes place the network in an inhibition-stabilised (ISN) regime, we performed an eigenvalue analysis of the system in which all inhibitory connections were removed (i.e. $w_I = 0$). To be in an ISN regime, either one eigenvalue of the corresponding Jacobian J^E of the excitatory-only network or the system trace $\mathrm{Tr}(J^E)$ must be positive, but the system including inhibitory feedback must be stable.

The presence and strength of competition in Fig. 3–2 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. 3a). Negative net currents correspond to competition between the stimulated and recorded excitatory neurons (shown as shading in Figure 3—Figure supplement 1).

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 **(Schüz and Palm 1989)**.

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; **(Binzegger et al. 2004; Schüz and Palm 1989)**).

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 v and u, computed over the surface of a 2D torus. In our models, each neuron had a Gaussian dendritic field of $\rho_{\mathrm{d}}=75\,\mu\mathrm{m}$ (approximate width of 300 μm; **(Hellwig 2000)**); and axonal field of $\rho_{\mathrm{a,e}}=290\,\mu\mathrm{m}$ for excitatory neurons (approximate width of 1100 μm; **(Boucsein et al. 2011; Holmgren et al. 2003; Hellwig 2000)**) and $\rho_{\mathrm{a,i}}=100\,\mu\mathrm{m}$ for inhibitory neurons (approximate width of 400 μm; **(Binzegger et al. 2007)**).

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 **(Peters 1979; Braitenberg and Schüz 1991)**. This was estimated by computing the integrated product of axonal and dendritic fields

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

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 match-

ing 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

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

$p_{ori} = \text{vonmises}(\theta_i, \theta_j, \kappa)$; p_{Peters} is the connection probability under non-specific Peters' rule connectivity, defined above; and s_1 is the proportional strength of specificity [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 $\mathbf{z}_j = \exp(-\mathrm{i}\zeta_j)$ for each neuron in the model, with uniformly-distributed orientations ζ_j ranging $[-\pi, \pi)$. Here "i" represents the complex number $\sqrt{-1}$; k ranges $1..N_S$, where N_S is the number of subnetworks in the model; q ranges $1..\vartheta$, 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 $\rho = 75\,\mu\text{m}$ (approximate width 300 μm). 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 func-

tional 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 the 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

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

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

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

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

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

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 **(Niell and Stryker 2008; Medini 2011)**. Each excitatory neuron was assigned an orientation tuning curve based on a von Mises function (a circular, Gaussian-like function), with a randomly chosen preferred orientation θ_i and a common input tuning curve width $\kappa = 4$. $\text{vonmises}(\cdot)$ is the non-normalised von Mises function with values [0..1], given by

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

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

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

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

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

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 determined 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* **(Ko et al. 2011)**. 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 Callaway suggest that on the order of $N = 5$–6 subnetworks exist in layers 2/3 of rodent cortex **(Yoshimura et al. 2005)**. 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 for providing the data analysed in Fig. 2. We are grateful to M.A. Penny and the attendees of the Capo Caccia workshop for helpful discussions.

Funding

This work was supported by the Velux Stiftung (grant number 787 to D.R.M.); the Novartis Foundation (grants to D.R.M. and B.M.K.); the Swiss National Science Foundation (grant number 31–120480 to B.M.K.); the European Commission FP7 program (grant BrainScales 269921 to F.H. and B.M.K.); and by the Convergent Science Network (fellowships to D.R.M).

Competing interests

The auhors declare no competing interests.

References

Ahmed B, Anderson JC, Douglas RJ, Martin KA, Whitteridge D. 1998. Estimates of the Net Excitatory Currents Evoked by Visual Stimulation of Identified Neurons in Cat Visual Cortex. *Cerebral Cortex* 8:462–476.

Atallah BV, Bruns W, Carandini M, Scanziani M. 2012. Parvalbumin-expressing Interneurons Linearly Transform Cortical Responses to Visual Stimuli. *Neuron* 73:159–170.

Binzegger T, Douglas R, Martin KAC. 2009. Topology and Dynamics of the Canonical Circuit of Cat V1. *Neural Networks* 22:1071–1078.

Binzegger T, Douglas RJ, Martin KAC. 2004. A Quantitative Map of the Circuit of Cat Primary Cortex. *Journal of Neuroscience* 24:8441–8453.

Binzegger T, Douglas RJ, Martin KAC. 2007. Stereotypical Bouton Clustering of Individual Neurons in Cat Primary Visual Cortex. *Journal of Neuroscience* 27:12242–12254.

Bock DD, Lee W-CA, Kerlin AM, Andermann ML, Hood G, Wetzel AW, Yurgenson S, Soucy ER, Kim HS, Reid RC. 2011. Network Anatomy and *in Vivo* Physiology of Visual Cortical Neurons. *Nature* 471:177–182.

Bosking WH, Zhang Y, Schofield B, Fitzpatrick D. 1997. Orientation Selectivity and the Arrangement of Horizontal Connections in Tree Shrew Striate Cortex. *Journal of Neuroscience* 17:2112–2127.

Boucsein C, Nawrot MP, Schnepel P, Aertsen A. 2011. Beyond the Cortical Column: Abundance and Physiology of Horizontal Connections Imply a Strong Role for Inputs From the Surround. *Front Neurosci* 5:32.

Braitenberg, Schüz. 1991. *Anatomy of the Cortex: Statistics and Geometry. Anatomy of the Cortex: Statistics and Geometry.*

Brown SP, Hestrin S. 2009. Intracortical Circuits of Pyramidal Neurons Reflect Their Long-range Axonal Targets. *Nature* 457:1133–1136.

Chen T-W, Wardill TJ, Sun Y, Pulver SR, Renninger SL, Baohan A, Schreiter ER, Kerr RA, Orger MB, Jayaraman V. 2013. Ultrasensitive Fluorescent Proteins for Imaging Neuronal Activity. *Nature* 499:295–300.

Clopath C, Büsing L, Vasilaki E, Gerstner W. 2010. Connectivity Reflects Coding: A Model of Voltage-based STDP with Homeostasis. *Nat Neurosci* 13:344–352.

Cossell, Iacaruso MF, Muir DR, Houlton R, Sader EN, Ko H, Hofer SB, Mrsic-FLogel. 2015. Functional Organization of Excitatory Synaptic Strength in Primary Visual Cortex. *Nature* 518:399–403.

Douglas RJ, Mahowald MA, Martin KAC. 1994. Hybrid Analog-digital Architectures for Neuromorphic Systems. *IEEE International Conference on Neural Networks* 3:1848–1853.

Ermentrout B. 1998. Linearization of F-I Curves by Adaptation. *Neural Computation* 10:1721–1729.

Fino E, Yuste R. 2011. Dense Inhibitory Connectvity in Neocortex. *Neuron* 69:1188–1203.

Gabott PLA, Somogyi P. 1986. Quantitative Distribution of GABA-immunoreactive Neurons in the Visual Cortex (area 17) of the Cat. *Experimental Brain Research* 61:323–331.

Galuske RA, Singer W. 1996. The Origin and Topography of Long-range Intrinsic Projections in Cat Visual Cortex: A Developmental Study. *Cerebral Cortex* 6:417–430.

Hahnloser RHR. 1998. On the Piecewise Analysis of Networks of Linear Threshold Neurons. *Neural Networks* 11:691–697.

Hellwig B. 2000. A Quantitative Analysis of the Local Connectivity Between Pyramidal Neurons in Layers 2/3 of the Rat Visual Cortex. *Biological Cybernetics* 82:111–121.

Hill SL, Wang Y, Riachi I, Schürmann F, Markram H. 2012. Statistical Connectivity Provides a Sufficient Foundation for Specific Functional Connectivity in Neocortical Neural Microcircuits. *Proceedings of the National Academy of Sciences* 109:E2885–E2894.

Hofer SB, Ko H, Pichler B, Vogelstein J, Ros H, Zeng H, Lein E, Lesica NA, Mrsic-Flogel TD. 2011. Differential Connectivity and Response Dynamics of Excitatory and Inhibitory Neurons in Visual Cortex. *Nat Neurosci* 14:1045–1052.

Holmgren C, Harkany T, Svennenfors B, Zilberter Y. 2003. Pyramidal Cell Communication Within Local Networks in Layer 2/3 of Rat Neocortex. *The Journal of Physiology* 551:139–153.

Juavinett AL, Callaway EM. 2015. Pattern and Component Motion Responses in Mouse Visual Cortical Areas. *Curr Biol* 25:1759–1764.

Juliano SL, Friedman DP, Eslin DE. 1990. Corticocortical Connections Predict Patches of Stimulus-evoked Metabolic Activity in Monkey Somatosensory Cortex. *Journal of Comparative Neurology* 298:23–39.

Kampa, Roth, Göbel, Helmchen. 2011. Representation of Visual Scenes by Local Neuronal Populations in Layer 2/3 of Mouse Visual Cortex. *Frontiers in Neural Circuits* 5 (18).

Kampa BM, Letzkus JJ, Stuart GJ. 2006. Cortical Feed-forward Networks for Binding Different Streams of Sensory Information. *Nat Neurosci* 9:1472–1473.

Kampa BM, Letzkus JJ, Stuart GJ. 2007. Dendritic Mechanisms Controlling Spike-timing-dependent Synaptic Plasticity. *Trends Neurosci* 30:456–463.

Katz LC, Callaway EM. 1992. Development of Local Circuits in Mammalian Visual Cortex. *Annual Review of Neuroscience* 15:31–56.

Kerlin AM, Andermann ML, Berezovskii VK, Reid RC. 2010. Broadly Tuned Response Properties of Diverse Inhibitory Neuron Subtypes in Mouse Visual Cortex. *Neuron* 67:858–871.

Kleiner M, Brainard D, Pelli D, Ingling A, Murray R, Broussard C. 2007. What's New in Psychtoolbox-3. *Perception* 36:1–1.

Ko H, Cossell L, Baragli C, Antolik J, Clopath C, Hofer SB, Mrsic-Flogel TD. 2013. The Emergence of Functional Microcircuits in Visual Cortex. *Nature* 496:96-100.

Ko H, Hofer SB, Pichler B, Buchanan KA, Sjöström PJ, Mrsic-Flogel TD. 2011. Functional Specificity of Local Synaptic Connections in Neocortical Networks. *Nature* 473:87–91.

Lee WA, Bonin V, Reed M, Graham BJ, Hood G, Glattfelder K, Reid RC. 2016. Anatomy and Function of An Excitatory Network in the Visual Cortex. *Nature.*

Lien AD, Scanziani M. 2013. Tuned Thalamic Excitation Is Amplified by Visual Cortical Circuits. *Nature Neuroscience* 16:1315–1323.

Li L-Y, Li Y-T, Zhou M, Tao HW, Zhang LI. 2013. Intracortical Multiplication of Thalamocortical Signals in Mouse Auditory Cortex. *Nature Neuroscience* 16:1179–1181.

Litwin-Kumar A, Doiron B. 2014. Formation and Maintenance of Neuronal Assemblies Through Synaptic Plasticity. *Nat Commun* 5:5319.

Litwin-Kumar A, Rosenbaum R, Doiron B. 2016. Inhibitory Stabilization and Visual Coding in Cortical Circuits with Multiple Interneuron Subtypes. *J Neurophysiol (In Press).*

Liu B-H, Li P, Li Y-T, Sun YJ, Yanagawa Y, Obata K, Zhang LI, Tao HW. 2009. Visual Receptive Field Structure of Cortical Inhibitory Neurons Revealed by Two-Photon Imaging Guided Recording. *The Journal of Neuroscience* 29:10520–10532.

Li Y, Lu H, Cheng P-L, Ge S, Xu H, Shi S-H, Dan Y. 2012. Clonally Related Visual Cortical Neurons Show Similar Stimulus Feature Selectivity. *Nature* 486:118–121.

Li Y-T, Ibrahim LA, Liu B-H, Zhang LI, Tao HW. 2013. Linear Transformation of Thalamocortical Input by Intracortical Excitation. *Nature Neuroscience* 16:1324–1330.

Luhmann HJ, Millán LM, Singer W. 1986. Development of Horizontal Intrinsic Connections in Cat Striate Cortex. *Experimental Brain Research* 63:443–448.

Luhmann HJ, Singer W, Martinez-Millán L. 1990. Horizontal Interactions in Cat Striate Cortex: I. Anatomical Substrate and Postnatal Development. *European Journal of Neuroscience* 2:344–357.

Malach R, Amir Y, Harel M, Grinvald A. 1993. 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* 90:10469–10473.

Mariño J, Schummers J, Lyon DC, Schwabe L, Beck O, Wiesing P, Obermayer K, Sur M. 2005. Invariant Computations in Local Cortical Networks with Balanced Excitation and Inhibition. *Nature Neuroscience* 8:194–201.

Markram H, Gerstner W, Sjöström J. 2012. Spike-Timing-Dependent Plasticity: A Comprehensive Overview. *Frontiers in Synaptic Neuroscience* 4.

Markram H, Muller E, Ramaswamy S, Reimann MW, Abdellah M, Sanchez CA, Ailamaki A, Alonso-Nanclares L, Antille N, Arsever S, Kahou GA, Berger TK, Bilgili A, Buncic N, Chalimourda A, Chindemi G, Courcol JD, Delalondre F, Delattre V, Druckmann S, Dumusc R, Dynes J, Eilemann S, Gal E, Gevaert ME, Ghobril JP, Gidon A, Graham JW, Gupta A, Haenel V, Hay E, Heinis T, Hernando JB, Hines M, Kanari L, Keller D, Kenyon J, Khazen G, Kim Y, King JG, Kisvarday Z, Kumbhar P, Lasserre S, Le Bé JV, Magalhães BR, Merchán-Pérez A, Meystre J, Morrice BR, Muller J, Muñoz-Céspedes A, Muralidhar S, Muthurasa K, Nachbaur D, Newton TH, Nolte M, Ovcharenko A, Palacios J, Pastor L, Perin R, Ranjan R, Riachi I, Rodríguez JR, Riquelme JL, Rössert C, Sfyrakis K, Shi Y, Shillcock JC, Silberberg G, Silva R, Tauheed F, Telefont M, Toledo-Rodriguez M, Tränkler T, Van Geit W, Díaz JV, Walker R, Wang Y, Zaninetta SM, DeFelipe J, Hill SL, Segev I, Schürmann F. 2015. Reconstruction and Simulation of Neocortical Microcircuitry. *Cell* 163:456–492.

Martin KA, Roth S, Rusch ES. 2014. Superficial Layer Pyramidal Cells Communicate Heterogeneously Between Multiple Functional Domains of Cat Primary Visual Cortex. *Nat Commun* 5:5252.

Martin KAC, Schröder S. 2013. Functional Heterogeneity in Neighboring Neurons of Cat Primary Visual Cortex in Response to Both Artificial and Natural Stimuli. *Journal of Neuroscience* 33:7325–7344.

Ma WP, Liu BH, Li YT, Huang ZJ, Zhang LI, Tao HW. 2010. Visual Representations by Cortical Somatostatin Inhibitory Neurons--selective but with Weak and Delayed Responses. *J Neurosci* 30:14371–14379.

Medini P. 2011. Cell-type-specific Sub- and Suprathreshold Receptive Fields of Layer 4 and Layer 2/3 Pyramids in Rat Primary Visual Cortex. *Neuroscience* 190:112–126.

Miller B, Blake NM, Erinjeri JP, Reistad CE, Sexton T, Admire P, Woolsey TA. 2001. Postnatal Growth of Intrinsic Connections in Mouse Barrel Cortex. *The Journal of Comparative Neurology* 436:17–31.

Morishima M, Morita K, Kubota Y, Kawaguchi Y. 2011. Highly Differentiated Projection-Specific Cortical Subnetworks. *The Journal of Neuroscience* 31:10380–10391.

Movshon J, Adelson E, Gizzi M, Newsome W. 1985. The Analysis of Moving Visual Patterns. In *Study Week on Pattern Recognition Mechanisms. Pontificiae Academiae Scientiarum Scripta Varia*, editor. Carlos Chagas, Ricardo Gattass and Charles Gross, 54:117–151:117–151.

Muir DR, Cook M. 2014. Anatomical Constraints on Lateral Competition in Columnar Cortical Architectures. *Neural Computation* 26:1624–1666.

Muir DR, Da Costa NMA, Girardin C, Naaman S, Omer DB, Ruesch E, Grinvald A, Martin KA, Douglas RJ. 2011. Embedding of Cortical Representations by the Superficial Patch System. *Cerebral Cortex* 21:2244–2260.

Muir DR, Kampa BM. 2015. FocusStack and StimServer: A New Open Source MATLAB Toolchain for Visual Stimulation and Analysis of Two-photon Calcium Neuronal Imaging Data. *Frontiers in Neuroinformatics* 8:85.

Muir DR, Mrsic-Flogel. 2015. Eigenspectrum Bounds for Semirandom Matrices with Modular and Spatial Structure for Neural Networks. *Physical Review E* 91:042808.

Muir DR, Roth MM, Helmchen F, Kampa BM. 2015. Model-based Analysis of Pattern Motion Processing in Mouse Primary Visual Cortex. *Front Neural Circuits* 9:38.

Neftci E, Binas J, Rutishauser U, Chicca E, Indiveri G, Douglas RJ. 2013. Synthesizing Cognition in Neuromorphic Electronic Systems. *Proceedings of the National Academy of Sciences* 110:E3468–E3476.

Neftci E, Chicca E, Indiveri G, Douglas R. 2011. A Systematic Method for Configuring VLSI Networks of Spiking Neurons. *Neural Computation* 23:2457–2497.

Niell CM, Stryker MP. 2008. Highly Selective Receptive Fields in Mouse Visual Cortex. *The Journal of Neuroscience* 28:7520–7536.

Nimmerjahn A, Kirchhoff F, Kerr JN, Helmchen F. 2004. Sulforhodamine 101 As a Specific Marker of Astroglia in the Neocortex in Vivo. *Nat Methods* 1:31–37.

Ohki K, Chung S, Ch'ng YH, Kara P, Reid RC. 2005. Functional Imaging with Cellular Resolution Reveals Precise Micro-architecture in Visual Cortex. *Nature* 433:597–603.

Ozeki H, Finn IM, Schaffer ES, Miller KD, Ferster D. 2009. Inhibitory Stabilization of the Cortical Network Underlies Visual Surround Suppression. *Neuron* 62:578–592.

Pecka M, Han Y, Sader E, Mrsic-Flogel TD. 2014. Experience-Dependent Specialization of Receptive Field Surround for Selective Coding of Natural Scenes. *Neuron*.

Perin R, Berger TK, Markram H. 2011. A Synaptic Organizing Principle for Cortical Neuronal Groups. *Proceedings of the National Academy of Sciences* 108:5419–5424.

Peters A. 1979. Thalamic Input to the Cerebral Cortex. *Trends in Neurosciences* 2:183–185.

Ramaswamy S, Hill SL, King JG, Schürmann F, Wang Y, Markram H. 2012. Intrinsic Morphological Diversity of Thick-tufted Layer 5 Pyramidal Neurons Ensures Robust and Invariant Properties of in Silico Synaptic Connections. *The Journal of Physiology* 590:737–752.

Raposo D, Kaufman MT, Churchland AK. 2014. A Category-free Neural Population Supports Evolving Demands During Decision-making. *Nature Neuroscience* 17:1784–1792.

Reimann MW, King JG, Muller EB, Ramaswamy S, Markram H. 2015. An Algorithm to Predict the Connectome of Neural Microcircuits. *Front Comput Neurosci* 9:120.

Rigotti M, Barak O, Warden MR, Wang X, Daw ND, Miller EK, Fusi S. 2013. The Importance of Mixed Selectivity in Complex Cognitive Tasks. *Nature* 497:585–590.

Rochefort NL, Narushima M, Grienberger C, Marandi N, Hill DN, Konnerth A. 2011. Development of Direction Selectivity in Mouse Cortical Neurons. *Neuron* 71:425–432.

Roth MM, Helmchen F, Kampa BM. 2012. Distinct Functional Properties of Primary and Posteromedial Visual Area of Mouse Neocortex. *J Neurosci* 32:9716–9726.

Rust NC, Mante V, Simoncelli EP, Movshon JA. 2006. How MT Cells Analyze the Motion of Visual Patterns. *Nat Neurosci* 9:1421–1431.

Sadeh S, Clopath C, Rotter S. 2015. Emergence of Functional Specificity in Balanced Networks with Synaptic Plasticity. *PLoS Comput Biol* 11:e1004307.

Sadeh S, Clopath C, Rotter S. 2015. Processing of Feature Selectivity in Cortical Networks with Specific Connectivity. *PLOS ONE* 10:e0127547.

Schüz A, Palm G. 1989. Density of Neurons and Synapses in the Cerebral Cortex of the Mouse. *Journal of Comparative Neurology* 286:442–455.

Shriki O, Hansel D, Sompolinsky H. 2003. Rate Models for Conductance-Based Cortical Neuronal Networks. *Neural Computation* 15:1809–1841.

Song S, Sjöström PJ, Reigl M, Nelson S, Chklovskii DB. 2005. Highly Nonrandom Features of Synaptic Local Connectivity in Local Cortical Circuits. *Public Library of Science Biology* 3:0507–0519.

Sporns O, Kötter R. 2004. Motifs in Brain Networks. *PLoS Biology* 2:e369.

Stosiek C, Garaschuk O, Holthoff K, Konnerth A. 2003. In Vivo Two-photon Calcium Imaging of Neuronal Networks. *Proceedings of the National Academy of Sciences* 100:7319–7324.

Tsodyks MV, Skaggs WE, Sejnowski TJ, McNaughton BL. 1997. Paradoxical Effects of External Modulation of Inhibitory Interneurons. *J Neurosci* 17:4382–4388.

Wertz A, Trenholm S, Yonehara K, Hillier D, Raics Z, Leinweber M, Szalay G, Ghanem A, Keller G, Rózsa B, Conzelmann KK, Roska B. 2015. Single-cell-initiated Monosynaptic Tracing Reveals Layer-specific Cortical Network Modules. *Science* 349:70–74.

White LE, Fitzpatrick D. 2007. Vision and Cortical Map Development. *Neuron* 56:327–338.

Wilson NR, Runyan CA, Wang FL, Sur M, Wilson NR, Runyan CA, Wang FL, Sur M. 2012. Division and Subtraction by Distinct Cortical Inhibitory Networks in Vivo. *Nature*.

Yoshimura Y, Callaway EM. 2005. Fine-scale Specificity of Cortical Networks Depends on Inhibitory Cell Type and Connectivity. *Nature Neuroscience* 8:1552–1559.

Yoshimura Y, Dantzker JLM, Callaway EM. 2005. Excitatory Cortical Neurons Form Fine-scale Functional Networks. *Nature* 433:868–873.

Yu YC, He S, Chen S, Fu Y, Brown KN, Yao XH, Ma J, Gao KP, Sosinsky GE, Huang K, Shi SH. 2012. Preferential Electrical Coupling Regulates Neocortical Lineage-dependent Microcircuit Assembly. *Nature* 486:113–117.

Zagha E, Ge X, McCormick DA. 2015. Competing Neural Ensembles in Motor Cortex Gate Goal-Directed Motor Output. *Neuron* 88:565–577.

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.	σ_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	

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 1

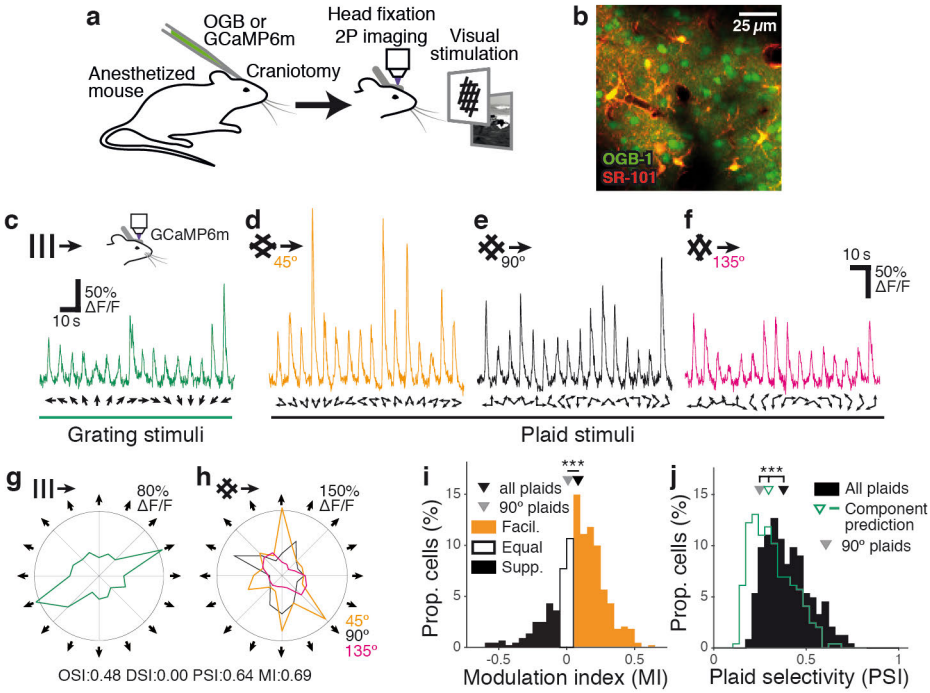

Figure 1: Plaid responses are facilitatory and selective in mouse V1. (**a, b**) Two-photon calcium imaging of visual responses in mouse V1. (**c**) Trial-averaged responses of a single neuron in mouse V1 to grating and (**d–f**) plaid stimuli of varying relative component orientations. **g–h** Response 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 are facilitatory (64% with MI > 0.05). In contrast, when only 90° plaids are presented neurons are more evenly split (39% with MI > 0.05; Supplementary Fig. 1). **j** Responses to three 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 (Muir and Kampa 2015). *** $p < 1\times10^{-10}$, Wilcoxon rank-sum test.

Figure 1 —Figure supplement 1

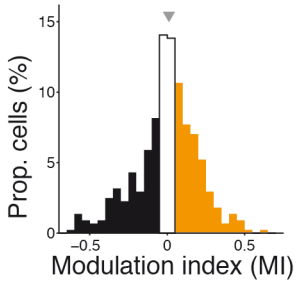

Figure 1 — Figure supplement 1: Facilitation and suppression under 90° plaid stimuli. Distribution of MI in response to drifting grating and 90° plaid stimuli, in mouse V1. Compare with Fig. 2g.

Figure 2

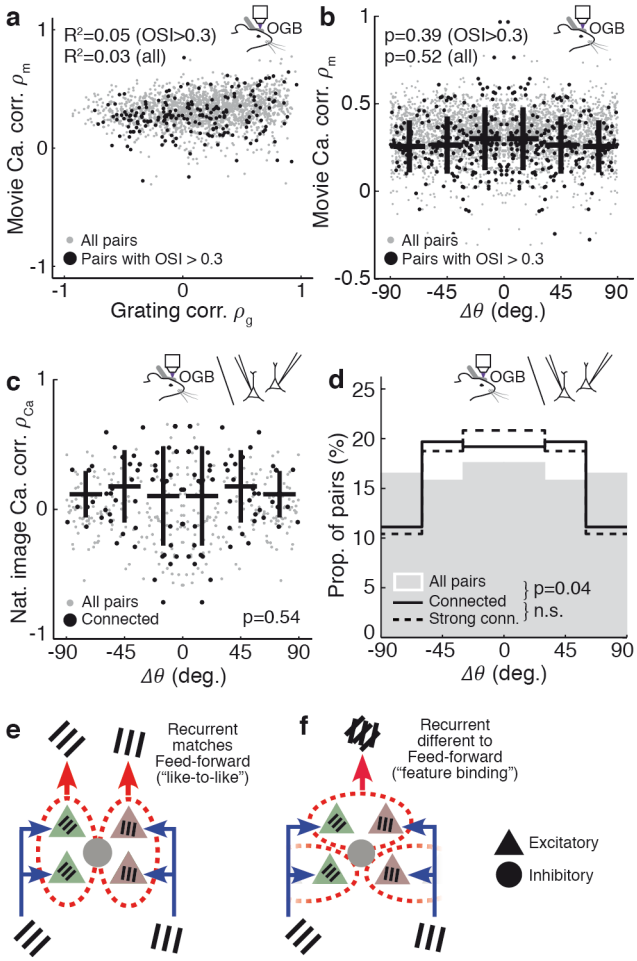

Figure 2: Connected neurons span a wide range of preferred orientations. a Pairs of neurons with high signal correlations to natural movies (ρ_m), which predicts high connection probability (Ko et al. 2011), can have similar or dissimilar grating responses. Pairs of neurons with similar orientation preference are not more likely to have high ρ_m (**b**) or high signal correlation to flashed natural scenes (**c**) than pairs with dissimilar orientation preference. **d** Connected pairs are slightly more likely to share similar orientation preferences than unconnected pairs (Cossell et al. 2015), but nevertheless span almost arbitrary orientation differences (≈20% of pairs with close to orthogonal orientation preference). **e** Connectivity scheme where local recurrent excitatory connections are matched to the feedforward visual preferences of the connected neurons ("like-to-like"). **f** Connectivity scheme where local recurrent excitatory connections are different from the feedforward visual preferences of connected neurons ("feature binding"). b, c Kuiper tests; d Ansari-Bradley test. n.s.: $p > 0.05$. Strong connections: strongest 50% of connected pairs, measured by EPSP amplitude.

Figure 3

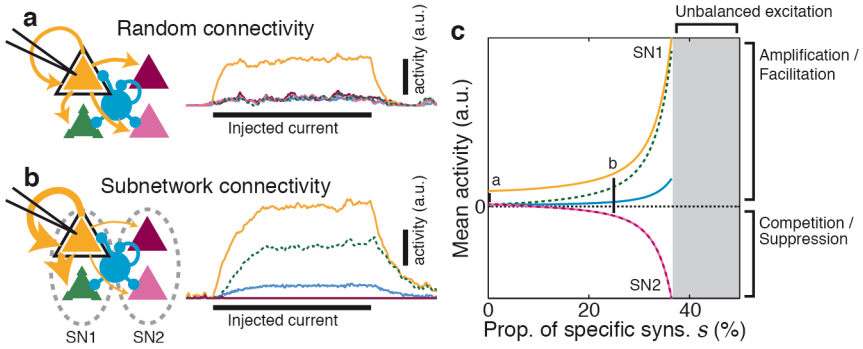

Figure 3: Non-random connectivity underlies amplification and competition. **a** In a simple model for random connectivity in mouse V1, injecting current into a single neuron (black outline) leads to non-specific activation of other excitatory (triangle) and inhibitory neurons (circle). Traces show the 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 Supplementary Fig. 3). Values of s used in panels a–b indicated on plot. a.u.: arbitrary units.

Figure 3 — Figure supplement 1

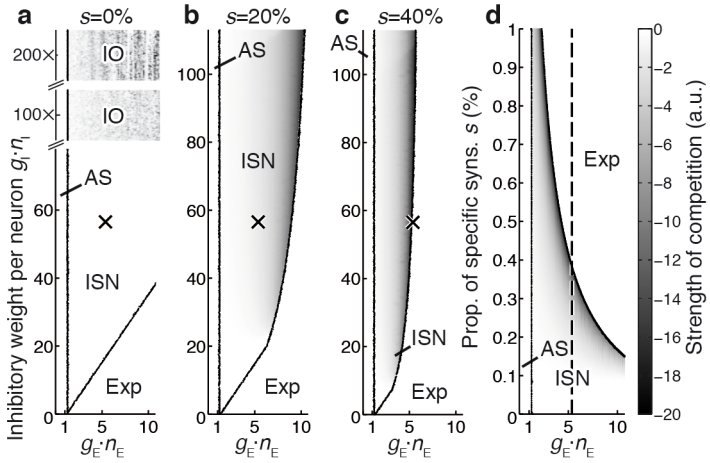

Figure 3 — Figure supplement 1: Estimated parameters for cortex place it in an Inhibition-Stabilised Network (ISN) regime, with competition provided by specific excitatory connectivity. a The network stability regimes in the parameter space defined by total inhibitory weight gI·nI and total excitatory weight gE·nE 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; (Tsodyks et al. 1997)), 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 gE·nE, with nominal value chosen for gI·nI (crosses in a–c). Nominal value for gE·nE is indicated by a dashed line. Both excitatory strength gE·nE and the proportion of specific synapses s affect network stability and the strength of competition. Abbreviations: gI,E: Synaptic strength per inhibitory or excitatory synapse; nI,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.

Figure 4

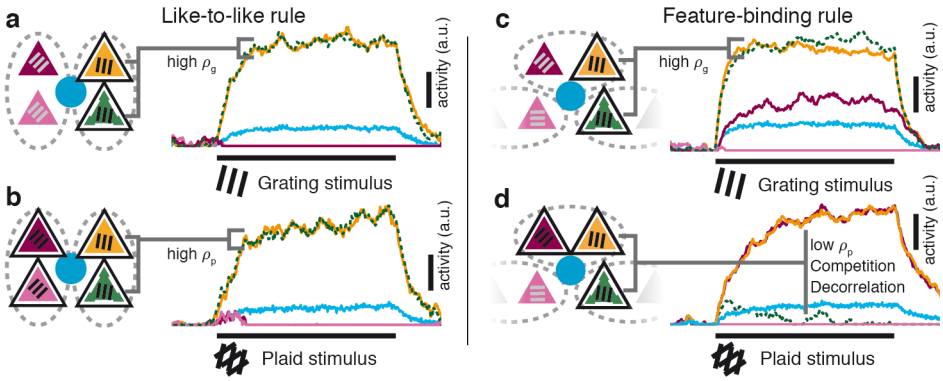

Figure 4: Rules for excitatory connectivity influence stimulus representations. a–b When local recurrent excitatory connections match the feedforward visual properties of connected neurons ("like-to-like"), grating responses (a) and plaid responses (b) are highly similar (high ρ_g & ρ_p). **c–d** In 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 (c, high ρ_g) can have dissimilar plaid responses (d, low ρ_p), reflecting decorrelation of these responses caused by competition. Black outlines: stimulated neurons. Grating labels: preferred orientation of that neuron. Other conventions as in Fig. 3e, f.

Figure 5

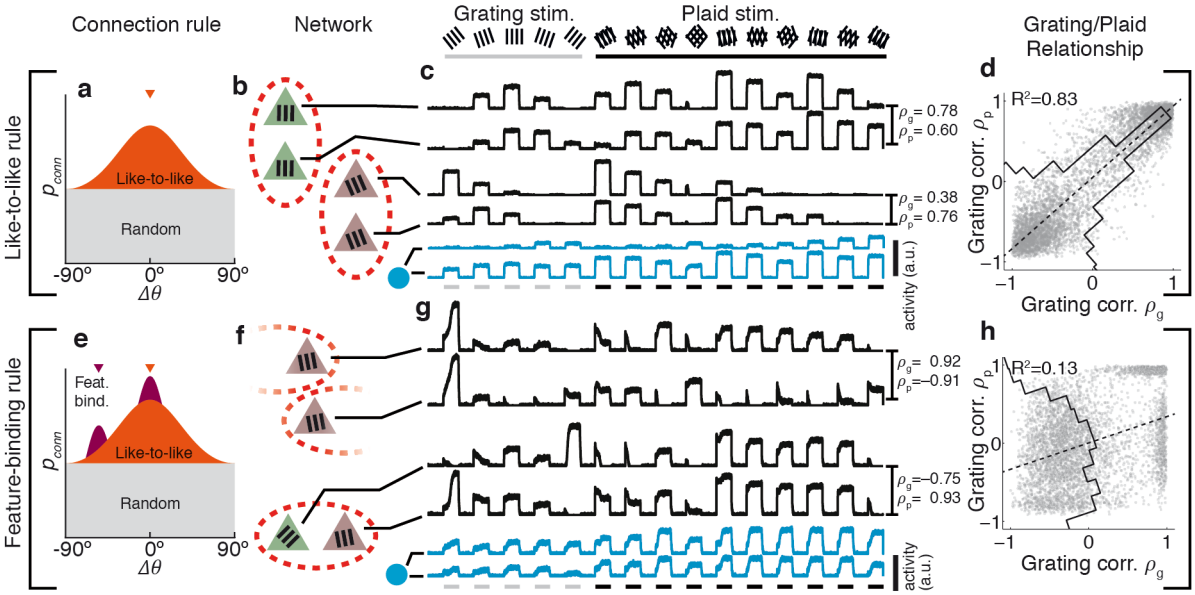

Figure 5: Rules for excitatory connectivity determine response correlation and decorrelation in a model of mouse V1. a–d In a large-scale network simulation incorporating like-to-like selective excitatory connectivity (connectivity rule and network schematic shown in a–b), responses of pairs of neurons to grating and plaid stimuli are always similar (c–d; 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 clustered around the diagonal (black line in d; high R^2). **e–h** When in addition to like-to-like connectivity, subnetworks also group neurons with several preferred orientations (e–f), then pairs of neurons with similar preferred orientations (g, high ρ_g) can respond differently to plaid stimuli (g, low ρ_p), and vice versa. **h** 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 h), indicating poor predictability between grating and plaid responses. Inhibitory responses are poorly tuned in both models (blue traces in c & g). Pips in a and e indicate example preferred orientations of a single subnetwork. Conventions in (b, f) as in Fig. 3e, f

Figure 5 —Figure supplement 1

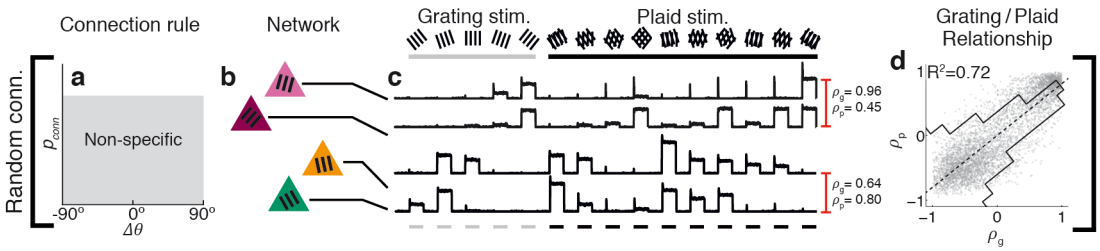

Figure 5 — Figure supplement 1: Grating and plaid responses are highly correlated in a model with random connectivity. a Under the non-specific connectivity model, the probability of forming a synapse with a target neuron is uniform over difference in preferred orientation. Neurons form synapses only according to spatial proximity. **b** Two example pairs of neurons are shown, and their responses to a set of grating and plaid stimuli (**c**). **d** Neurons with similar responses to grating stimuli (high ρ_g) always have similar responses to plaid stimuli (high ρ_p), and vice versa.

Figure 6

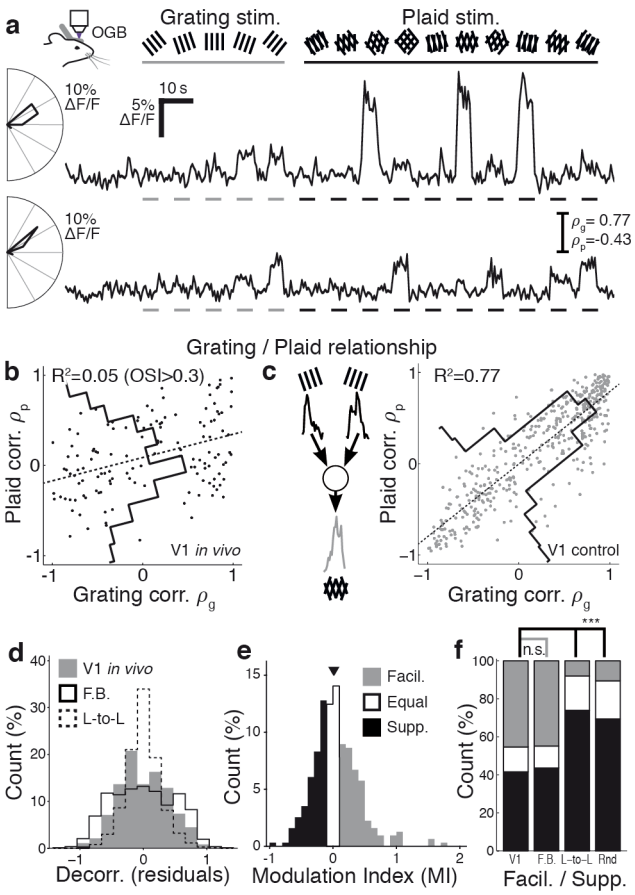

Figure 6: Responses to contrast-oscillating plaid and grating stimuli in mouse V1 suggest feature-binding connection rules. a Trial-averaged responses of a pair of neurons from a single imaging site, with similar preferred orientations (polar plots; high ρ_g) but with dissimilar responses to plaid stimuli (low ρ_p). **b** 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). **c** 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 b. **d** 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). **e** Responses to in V1 are split between facilitating and suppressing (45% MI > 0.05; 42% MI < -0.05). **f** 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.

Figure 6 — Figure supplement 1

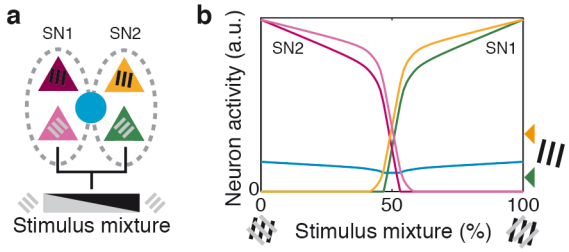

Figure 6 — Figure supplement 1: Non-random connectivity supports auto-associative behaviour. In a simple model with two subnetworks (**a**), presenting a linear graduated mixture between the ideal stimuli for the two subnetworks (**b**) results in competition and switching between network representations. 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%.

Figure 7

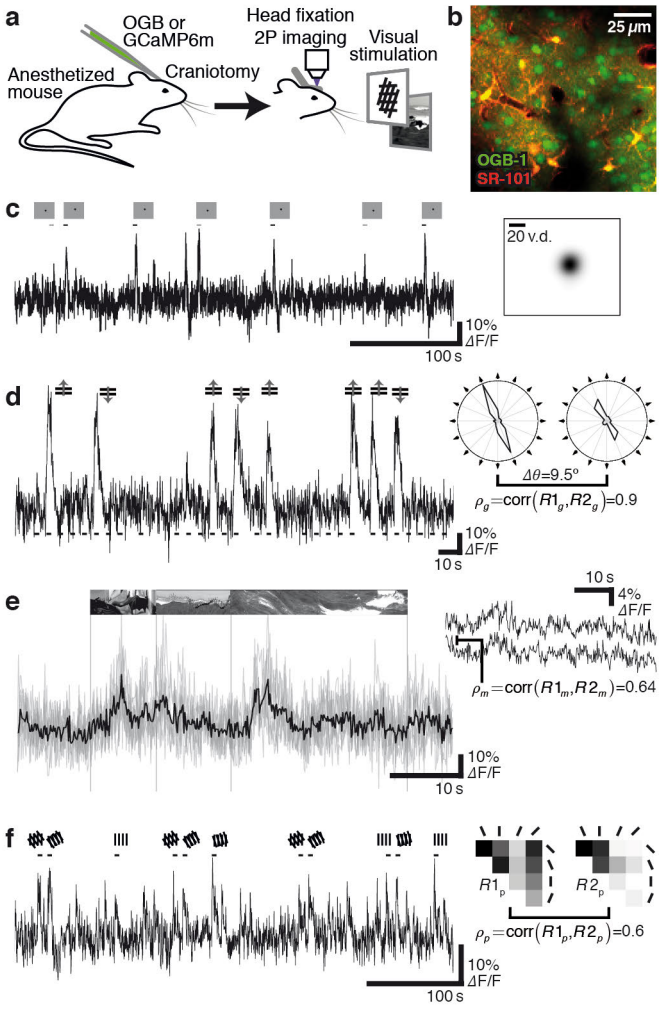

Figure 7: Measurement of visually-evoked responses and response similarity in mouse V1. a Imaging was performed in anaesthetised mice; visual stimuli were presented on a screen in front of the animal (see Methods). **b** Characterisation of receptive field location using sparse drifting/rotating grating stimuli. Single-trial calcium responses (grey); presentation time of optimal stimulus and sub-optimal stimulus indicated (black and grey). Inset: estimated RF location for the same neuron. **c** Single-trial calcium response to drifting grating stimuli (grey); presentation of optimal stimulus orientation indicated above, all stimulus presentation times indicated below. Inset: calculation of grating response similarity ρ_g between two neurons. **d** Single-trial (grey) and trial-averaged calcium response (black) to natural movie stimuli. Vertical lines indicate timing of movie sequence onset. Inset: calculation of movie response similarity (ρ_m), using signal correlations over trial-averaged responses from two neurons (coloured traces). **e** Single-trial calcium response to contrast-oscillating grating and plaid stimuli; presentation time of stimuli evoking strong responses indicated. Inset: measurement of plaid response similarity ρ_p between two neurons.