Population coupling predicts the plasticity of stimulus responses in cortical circuits

Yann Sweeney
Department of Bioengineering
Imperial College London
y.sweeney@imperial.ac.uk

Claudia Clopath
Department of Bioengineering
Imperial College London
c.clopath@imperial.ac.uk

Abstract

Long-term imaging of sensory cortex reveals a diverse range of stimulus response stability: some neurons retain stimulus responses that are stable over days whereas other neurons have highly plastic stimulus responses. Using a recurrent network model, we explore whether this observation could be due to an underlying diversity in the synaptic plasticity of neurons. We find that, in a network with diverse learning rates, neurons with fast rates are more coupled to population activity than neurons with slow rates. This phenomenon, which we call a plasticity-coupling link, surprisingly predicts that neurons with high population coupling exhibit more long-term stimulus response variability than neurons with low population coupling. We substantiate this prediction using recordings from the Allen Brain Observatory which track the orientation preferences of 15,000 neurons in mouse visual cortex. In agreement with our model, a neuron's population coupling is correlated with the plasticity of its orientation preference. Finally, we show that high population coupling helps plastic neurons alter their stimulus preference during a simple perceptual learning task, but hinders the ability of stable neurons to provide an instructive signal for learning. This suggests a particular functional architecture: a stable 'backbone' of stimulus representation formed by neurons with slow synaptic plasticity and low population coupling, on top of which lies a flexible substrate of neurons with fast synaptic plasticity and high population coupling.

Keywords: Network models, Synaptic plasticity, Stimulus representation

Introduction

The brain encodes information about the external world via its neural activity. One aspect of such encoding is that neurons in sensory cortex often have a preferred stimulus which evokes a stronger response than other stimuli. These stimulus responses can change during learning or adaptation: if a particular stimulus feature is overexpressed within an environment, for example, more neurons will be recruited to encode this feature (Sengpiel et al., 1999). Advances in neural imaging techniques allow us to interrogate such changes by tracking stimulus responses of hundreds of neurons over many days in vivo (Andermann, 2010; Mank et al., 2008). These recordings reveal a substantial, and puzzling, variability in the long-term stability of responses in sensory cortex: some neurons retain highly stable preferences to specific stimuli, whereas the stimulus preference of other neurons change from day to day (Ranson, 2017; Clopath and Rose, 2017; Rose et al., 2016; Poort et al., 2015; Lütcke et al., 2013). The degree of stimulus response stability typically depends on brain region; whisking responses in mouse barrel cortex are highly plastic, whereas visual responses in mouse V1 are more stable but still exhibit fluctuations (Clopath and Rose, 2017; Lütcke et al., 2013). Moreover, it is possible to induce stimulus response plasticity through perturbations such as sensory deprivation (Rose et al., 2016), or to increase task-related stimulus response stability through rewarded learning (Poort et al., 2015).

Current theories which address the long-term variability of stimulus responses primarily ask how motor learning occurs with unstable representations (Driscoll et al., 2017; Ajemian et al., 2013; Rokni et al., 2007), or seek to explain it as a form of probabilistic sampling (Kappel et al., 2017, 2015). Although the stability of neural representation is correlated with firing rate in hippocampal place cells (Grosmark and Buzsaki, 2016) and in visual cortex (Ranson, 2017), it is not known how cellular or network properties influence a neuron's stimulus response stability (Clopath and Rose, 2017). We are therefore lacking a theory of why some neurons' stimulus responses are more stable than others, and how this affects perception and learning. By investigating how synaptic plasticity mediates stimulus response variability, we aim here to establish how this diversity of stimulus response stability emerges, and whether it is functionally relevant.

We propose that the observed diversity of stimulus response stability may be explained by a diversity of neurons' inherent plasticity (or learning rate) within a network. Consequently, we explore how diverse learning rates across neurons impact synaptic connectivity in a recurrent network model of mouse visual cortex. We find that neurons with fast learning rates exhibit more variability of their stimulus selectivity than neurons with slow learning rates. Intriguingly, we also find that fast neurons have higher population coupling, a measure of how correlated an individual neurons activity is with the rest of the population (Okun et al., 2015).

This unexpected plasticity-coupling link, in which more plastic neurons are also more coupled to the rest of the population, provides a mechanism for the diverse population coupling previously observed in sensory cortex (Okun et al., 2015). Moreover, the plasticity-coupling link predicts that neurons with high population coupling exhibit more long-term stimulus response variability than neurons with low population coupling. We substantiate this prediction with *in vivo* calcium imaging of mouse visual cortex from the Allen Brain Observatory (ABI, 2016), finding that a neuron is more likely to exhibit variability of its orientation preference if it has high population coupling.

Finally, we explore the functional implications of both diverse population coupling and diverse learning rates within our network model. We find that strong population coupling helps plastic neurons alter their stimulus preference during a simple perceptual learning task, but hinders the ability of stable neurons to provide an instructive signal for learning. The plasticity-coupling link exploits this dependence by ensuring that highly plastic neurons - the substrate for perceptual learning - are strongly coupled to the population, while less plastic neurons are weakly coupled and act as a stable 'backbone' of stimulus representation.

Results

2.1 A 'plasticity-coupling link' emerges in networks with diverse learning rates: fast neurons have higher population coupling than slow neurons

Our aim is to explore whether the diversity of stimulus response stability can be explained by a diversity of neurons' inherent plasticity (or learning rate) within a network. To this end, we

use network simulations to characterise the impact of diverse learning rates on recurrent synaptic connectivity in sensory cortex.

We first explore the impact of diverse learning rates in a simple, fully connected network of rate neurons (Figure 1A, Methods). Excitatory recurrent synapses in our network undergo Hebbian plasticity and synaptic scaling, while inhibitory synapses undergo homeostatic inhibitory plasticity (Vogels et al., 2011). Extending traditional models of Hebbian plasticity in which synaptic weight updates depend only on the correlation of pre- and post-synaptic activity, we introduce diversity by assigning either a fast or slow Hebbian learning rate (α) to individual neurons. The learning rate is expressed postsynaptically, such that the synaptic input weights onto neurons with a large α are more plastic than those with a small α (Equation 3).

Each neuron receives feedforward input from 1 of 4 possible visual stimuli representing gratings of different orientations, and independent noise. The Hebbian plasticity rule potentiates connections between neurons which share the same feedforward stimulus preference, due to their coactivity. This drives the emergence of strong bidirectional connections amongst stimulus-specific groups of neurons, while the remaining non-specific connections weaken (Figure 1B,C) (Ko et al., 2013; Clopath et al., 2010). Fast neurons develop these strong, specific connections sooner than slow neurons (Figure 1B, solid lines). However, the increased learning rate also leads to stronger synaptic weight fluctuations. These fluctuations occur both for synapses from neurons which share stimulus preference (specific connections) and for synapses from neurons which have different stimulus preference (non-specific connections). For slow neurons, in contrast, non-specific and specific connections tend towards either zero or the maximum synaptic weights respectively, remaining relatively stable after convergence (Figure 1B, black lines). This leads to connection specificity that is stronger and more stable compared with fast neurons (Figure 1D).

The observed dependence of connection specificity on learning rate is conserved if, instead of just two values of α representing either fast or slow neurons, we simulate plasticity in a network of neurons with a diverse range of α (Figure 1E). Increasing α predominantly drives an increase in non-specific connections rather than a decrease in specific connections. This leads to an overall increase in the amount of synaptic input amongst neurons with high α.

Population coupling is a recently characterised feature of neural activity which describes how correlated a neuron's activity is with the overall population activity, and which can be measured from calcium imaging recordings of neural activity (Okun et al., 2015). Since population coupling is correlated with the amount of local synaptic input in cortical networks (Okun et al., 2015), this measure could be a useful and experimentally observable proxy for the specificity of recurrent connectivity in our networks. We therefore investigate its suitability by measuring the population coupling of neurons in our network after synaptic plasticity (Methods). Interestingly, population coupling increases with learning rate, closely following the dependence of non-specific connectivity on α (Figure 1E, red points).

The dependence of a neuron's population coupling on its learning rate α, which we call a 'plasticity-coupling link', could provide a framework for relating the functional role of a neuron within a network to its dynamics. We therefore explore conditions necessary for this plasticity-coupling link by embedding a single plastic neuron within a static network and varying key model parameters (Figure S1). A strong plasticity-coupling link requires both moderate amounts of noise within the network and relatively slow synaptic scaling compared with Hebbian plasticity, in agreement with experimental data (Turrigiano et al., 1998). We next investigate whether this plasticity-coupling link is robustly observed in more biologically detailed networks.

2.2 Diverse population coupling emerges in cortical networks with diverse learning rates

As the plasticity-coupling link is robustly observed in a fully-connected small network with simple stimulus responses, we next investigate i) whether the plasticity-coupling link is also present in larger networks which more accurately represent the synaptic connectivity and stimulus response properties observed in mouse visual cortex, and ii) whether the diverse population coupling observed in sensory cortex emerges simply by introducing diverse learning rates (Okun et al., 2015).

We explore this in a network of 250 excitatory neurons with randomly generated Gabor receptive fields. This network has been shown to reproduce receptive field correlations and synaptic weight statistics that are observed in mouse visual cortex (Watanabe et al., 2016; Cossell et al., 2015) (see

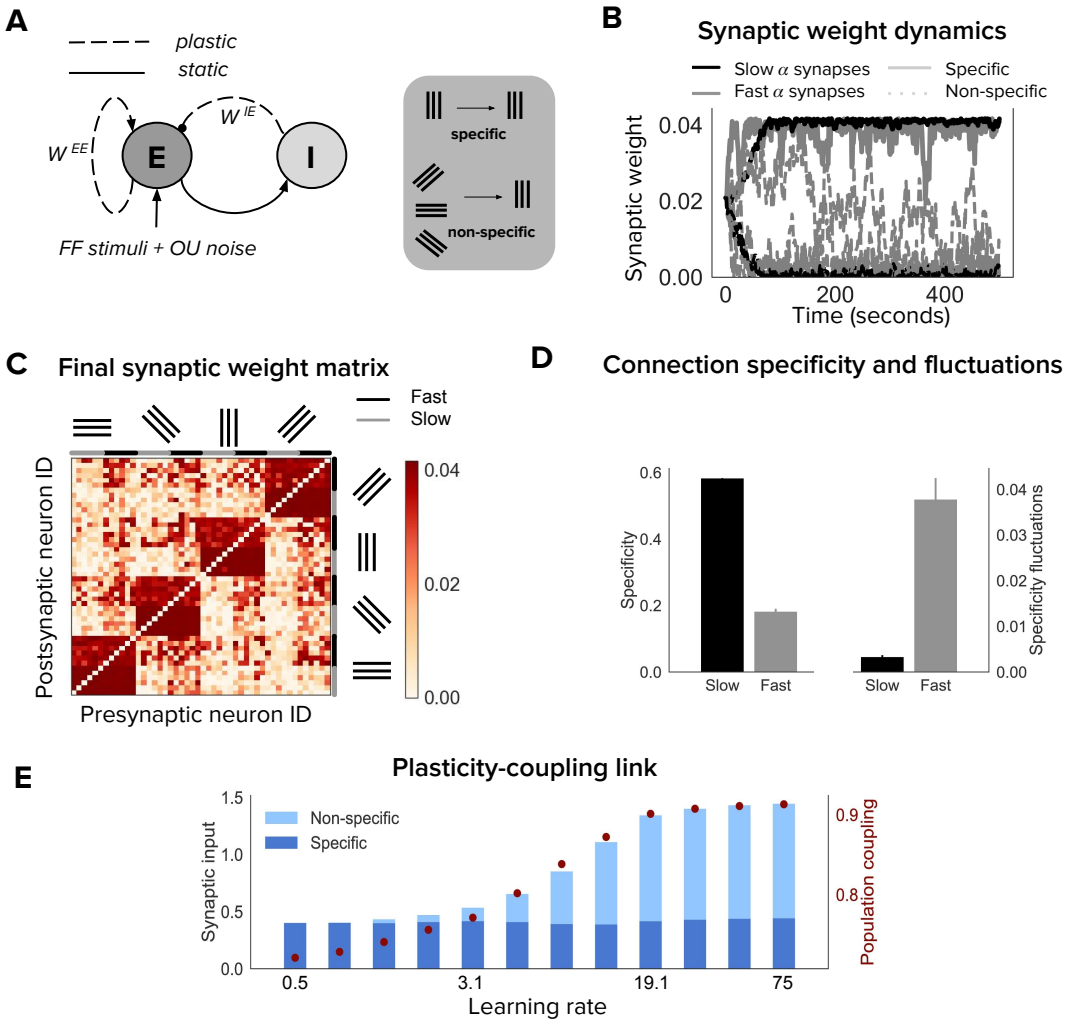

Figure 1: **Neurons with fast learning rates develop more non-specific connections, and higher population coupling, than neurons with slow learning rates (A)** Connection diagram of the recurrent network model with excitatory (E) and inhibitory (I) neurons. Dashed lines denote plastic synapses and solid lines denote static synapses. **(B)** Synaptic weight dynamics during presentation of random sequences of stimuli to the network. Synaptic inputs onto slow neurons ($\alpha = 1$, gray) and onto fast neurons ($\alpha = 5$, black). Synapses between neurons which share the same feedforward stimulus preference (specific) have solid lines, and synapses between neurons which have different feedforward stimulus preference (non-specific) have dashed lines. **(C)** Excitatory synaptic weight matrix of the recurrent network after synaptic plasticity. Neuron IDs are organised by feedforward stimulus preference. For each of the 4 stimulus groups the first 6 neurons are slow ($\alpha = 1$) and the next 6 neurons are fast ($\alpha = 5$). **(D)** Connection specificity (ratio of specific to non-specific synaptic input strength) after synaptic plasticity for slow and fast neurons (left), and the standard deviation over time of the connection specificity for slow and fast neurons (right). **(E)** Amount of non-specific (light blue) and specific (dark blue) synaptic input for neurons in a network with diverse learning rates, as the learning rate of the postsynaptic neuron is varied along a logarithmic scale. Population coupling of neurons with different learning rates (red points).

Methods section; Receptive-field based network model). We compare networks in which there is a uniform α across all neurons to networks with diverse α.

Both networks with uniform α and networks with diverse α develop strong synaptic connections between neurons with similar receptive fields. There is, however, a broader range of summed synaptic inputs in diverse networks, when compared with uniform networks (Figure 2A). This occurs because the total excitatory synaptic input onto a neuron covaries with α in the diverse network (Figure 2B).

In agreement with our previous observations, the population coupling of a neuron is determined by its total excitatory synaptic input (Figure 2C, blue line. r=0.29, p<1e-5, Spearman correlation). Diverse learning rates within a cortical network indeed lead to a broad distribution of population coupling, as observed by Okun et al. (2015) (Figure 2D, blue). Although the network with uniform α also exhibits some heterogeneity of population coupling, in this network a neuron's population coupling is not correlated with the amount of synaptic input it receives (Figure 2C,D, green. p=0.52, Spearman correlation). This contradicts experiments which demonstrate a correlation between synaptic input and population coupling (Okun et al., 2015).

We next investigate the long-term variability of stimulus selectivity within both networks by measuring the fluctuations of neuronal stimulus selectivity throughout a period of synaptic plasticity (Methods). We find that the magnitude of these fluctuations is independent of population coupling in the uniform network (p=0.4, Spearman correlation), but is correlated with population coupling in the diverse network (r=0.18,p=1e-5, Spearman correlation Figure 2E).

We then characterise the dependence of population coupling and stimulus selectivity on the amplitude of external input noise, again for networks with either uniform or diverse α (Figure 2F). As the majority of excitatory synaptic input received by neurons in visual cortex is recurrent, we simulate a regime with relatively weak feedforward stimulus-related input and high noise for Figure 2A-E (Cossell et al., 2015; Lin et al., 2015). This results in a broader distribution of population coupling and weaker stimulus selectivity for networks with diverse α, compared to networks with uniform α (Figure 2F). The dynamics of cortical activity observed *in vivo* are therefore more closely captured by networks with diverse α, compared to networks with uniform α.

Overall, these simulations show that the plasticity-coupling link observed in our small network model is robust in a larger network with receptive field properties and neuronal responses similar to mouse visual cortex. Networks with diverse α exhibit a broader range of population coupling than networks with uniform α. Moreover, diverse learning rates introduce a correlation between a neuron's population coupling and its total excitatory synaptic input, in agreement with experimental observations (Okun et al., 2015). Taken together, diverse learning rates provide a parsimonious explanation for the diverse population coupling observed in sensory cortical networks.

2.3 Experimental validation: population coupling is correlated with stimulus response variability in vivo

We have demonstrated that the population coupling of a neuron in a recurrent network model depends on its inherent plasticity. This plasticity-coupling link predicts a correlation between a neuron's population coupling and the variability of its stimulus selectivity. We now test this prediction using 2-photon calcium imaging of visual cortex in awake adult mice (Methods). The data we analyse is publicly available and was collected by the Allen Brain Institute ABI (2016). Mice passively viewed drifting or static gratings, interleaved with natural movies, while the simultaneous responses of $\sim 15,000$ excitatory neurons from 64 animals were recorded (Figure 3A). We measure the population coupling of each neuron over the entire recording session, and the preferred orientation of each neuron during the first 10 minutes and last 10 minutes of the experiment (Methods, Figure 3B). We then compare these two measurements of orientation preference to identify whether the preferred orientation of some neurons vary over the course of the experiment.

There is a broad distribution of population coupling, in agreement with previous observations (Figure 3C) (Sedigh-Sarvestani et al., 2017; Okun et al., 2015). Roughly 60% of neurons express variability of their preferred orientation between the beginning and the end of the experiment. The distribution of changes in preferred orientation ($\Delta \mathrm{ORI}_{\mathrm{pref}}$) is highly skewed towards smaller magnitudes. (Figure 3D).

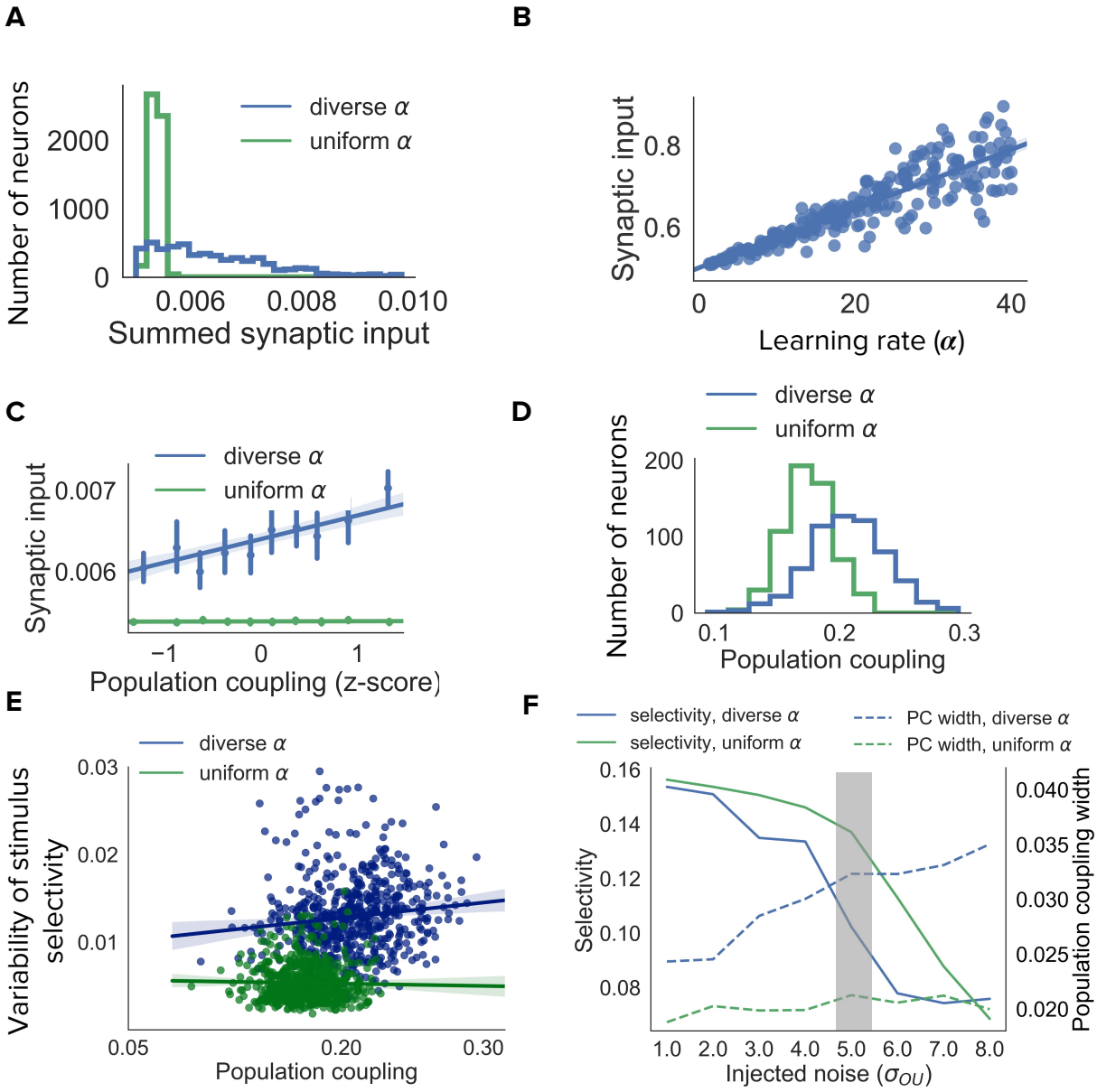

Figure 2: **Diverse population coupling from diverse learning rates in a cortical network model** **(A)** Distribution of summed synaptic input onto each neuron in networks with diverse learning rates (blue), and networks with uniform learning rates (green) **(B)** Mean recurrent excitatory synaptic input received by a neuron correlates with its learning rate, α. **(C)** The population coupling of a neuron is correlated with the amount of recurrent synaptic input it receives for the network with diverse learning rates (blue), as opposed to the network with uniform learning rates (green).**(D)** Diverse population coupling occurs in our recurrent network model. The population coupling distribution is wider for networks with diverse learning rates (blue) compared to networks with uniform learning rates (green, p<1e-5, Levene test). **(E)** The variability of stimulus selectivity is correlated with population coupling in networks with diverse learning rates (blue, r=0.18, p=1e-5, Spearman correlation), but not in networks with uniform learning rates (green, p=0.4, Spearman correlation). **(F)** Dependence of network properties on the amplitude of injected noise (σ_{OU}). Stimulus selectivity decreases with increasing σ_{OU} for networks with both diverse and uniform learning rates (blue and green lines, respectively). The distribution of population coupling broadens with increasing noise for networks with diverse learning rates, but not for networks with uniform learning rates (blue and green dashed lines, respectively). Panel **A-E** use $\sigma_{\mathrm{OU}} = 5.0$ (shaded gray area).

Population coupling is weakly but significantly correlated with the average change in preferred orientation, when pooling together neurons across all experiments (Figure 3E, r=0.04, p<1e-4, Spearman correlation). We characterise this dependence for each experiment by comparing the population coupling of neurons with variable orientation preferences (those with $\Delta ORI_{pref} > 0$) versus those with stable orientation preference. While there is substantial variability of the strength of the effect, the majority of experiments show a trend in which neurons with plastic orientation preferences have a higher mean population coupling than those with stable orientation preferences (Figure 3F, p<0.001, t-test). As the mean activity level of a neuron could conceivably determine its stimulus preference stability (Ranson, 2017; Grosmark and Buzsaki, 2016), we tested this and found no dependence of the tendency of a individual neuron to change stimulus preference on its average calcium fluorescence (p=0.17, Spearman correlation).

(Okun et al., 2015) did not observe any correlations between population coupling of a neuron and its orientation selectivity. In contrast, our network model predicts that neurons with high population coupling are less selective than neurons with low population coupling. We tested whether there was this predicted dependence between population coupling and orientation selectivity in these data. We indeed found a weak anti-correlation between population coupling and orientation selectivity index (Figure S2, r=-0.05,p<1e-6, Pearson correlation).

2.4 Diverse learning rates maintain both a stable backbone and a flexible substrate of stimulus representation

Our analysis thus far explored the impact of diverse rates of plasticity on synaptic connectivity. We established a link between diverse population coupling and diverse stimulus response variability, both of which are observed in sensory cortex. We now explore the functional implications of diverse population coupling and learning rates within recurrent networks. In order to simplify our analysis we consider both forms of diversity in isolation.

The presence of diverse rates of plasticity in a network suggests a dichotomy of roles: less plastic neurons could form stable stimulus representations while more plastic neurons could allow flexible representation. This could, for example, be beneficial during perceptual learning. We test this hypothesis by simulating an extended period of perceptual learning in our small network model (Methods, Figure 4A). We do this using a simple paradigm in which a randomly chosen feedforward stimuli is associated with an increased external input. This external input could be mediated by a reward, or some other top-down signal. Hebbian plasticity potentiates the recurrent synaptic connections from neurons which are tuned to the stimulus onto all neurons. This increases the selectivity of all neurons to the associated stimulus (Figure 4A).

We evaluate the ability of our network to continually learn these stimulus associations in the case where α is slow for all neurons, α is fast for all neurons, or where there is diverse α (both slow and fast) for each feedforward stimulus group.

We find that a network with only fast α quickly learns the stimulus associations (Figure 4A, bottom). However, repeated associations with neurons that do not share feedforward stimuli cause the specificity of recurrent connectivity to decrease, thus degrading the representation of feedforward stimuli (Figure 4A, top). Although neurons still form associations with the feedforward stimulus, this is because we keep the feedforward stimulus weights fixed; one can imagine that this feedforward selectivity may also degrade if these weights were plastic. Conversely, the network with only slow α retains a stable representation of the feedforward stimuli but performs poorly in representing the associated stimulus (Figure 4A). The network with diverse α overcomes these issues by having fast neurons which flexibly learn stimulus associations and slow neurons which maintain a 'backbone' of stimulus representation (see diagram, Figure 4B).

2.5 The plasticity-coupling link enables efficient perceptual learning

Having demonstrated the advantage of diverse learning rates within a network for perceptual learning, we now ask whether diverse population coupling has any impact on a network's performance in this task. Given the plasticity-coupling link, we are particularly interested in whether the impact of population coupling on performance is dependent on a neuron's rate of plasticity. To investigate this, we choose the extreme case in which there is a single neuron with plastic synaptic inputs embedded in an otherwise static recurrent network. Since all other synapses in the network are static (see diagram,

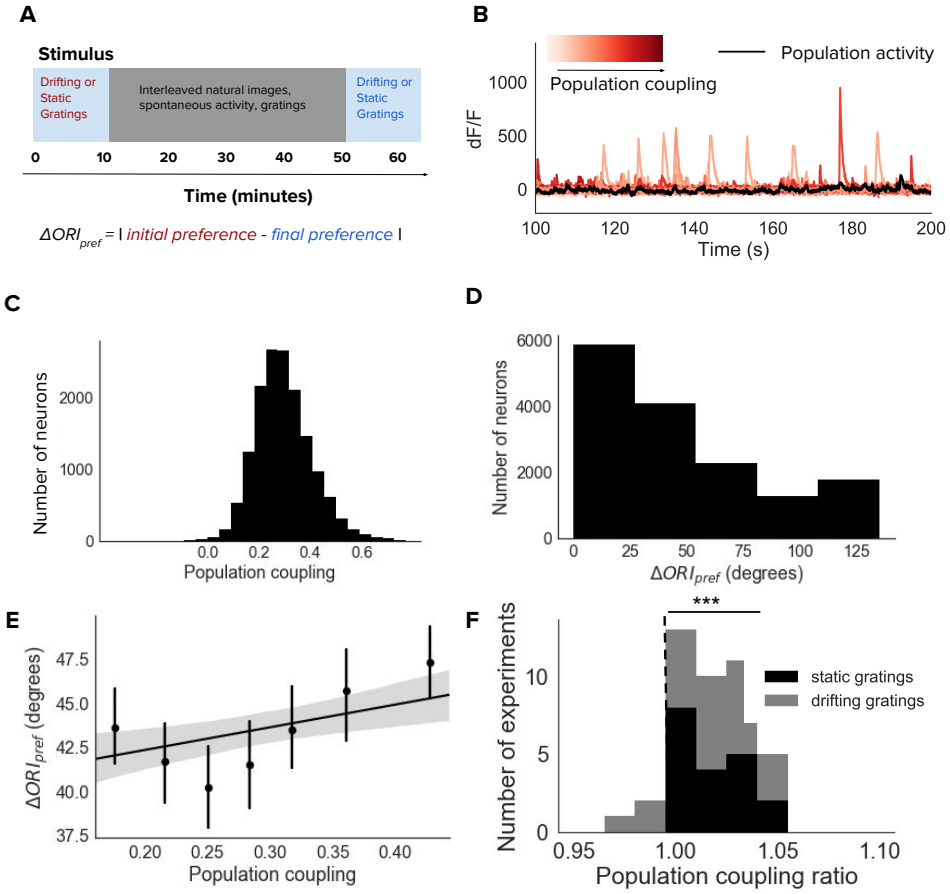

Figure 3: **Population coupling is correlated with stimulus response variability in mouse visual cortex *in vivo*** **(A)** Diagram of stimulus and data analysis protocol **(B)** dF/F calcium fluorescence traces of neurons (red traces) in an example experiment from the Allen Brain Observatory. Mean activity is shown in black, and the population coupling of each neuron is indicated by its colour changing from light to dark red with increased population coupling. **(C)** The population coupling distribution of all neurons across all experiments (64 experiments, 15,281 neurons). **(D)** Absolute difference in preferred orientation ($\Delta\text{ORI}_{\text{pref}}$) between the beginning and the end of the recording session. The distribution of $\Delta\text{ORI}_{\text{pref}}$ is for all neurons across all experiments (both drifting gratings and static gratings). **(E)** $\Delta\text{ORI}_{\text{pref}}$ is correlated with population coupling (r=0.04, p<1e-4). Data shown for all neurons across all experiments, binned by population coupling. Linear regression fit for all datapoints (shaded gray area indicates 95% confidence interval). **(F)** Ratios of the mean population coupling of neurons that change their preferred orientation ($\Delta\text{ORI}_{\text{pref}} > 0$) versus mean population coupling of neurons that conserve their preferred orientation ($\Delta\text{ORI}_{\text{pref}} = 0$), for each individual static grating (black) or drifting grating (gray) experiment. Dashed vertical line indicates expected value if a neuron's orientation preference variability is not dependent on its population coupling (*** p<0.001, one sample t-test).

Figure 4C), we focus on how synaptic inputs onto the plastic neuron evolve during learning. We adjust the population coupling (PC) of either the single plastic neuron ($PC_{plastic}$) or the static neurons (PC_{static}), and measure the ability of the plastic neuron to learn a stimulus association (Methods). We simulate perceptual learning by turning on an extra external input to all neurons in the network whenever the associated stimulus (red) is presented to the network (Figure 4C). We judge the plastic neuron to have learned the association if the synaptic weight from the presynaptic neuron selective to the associated stimulus becomes stronger than the weight from the presynaptic neuron (blue). We find that strongly coupling the plastic neuron to the population improves performance, while strongly coupling the static neurons to the population impairs performance (Figure 4D).

We can understand this by considering that, for learning to occur, synaptic potentiation must happen between the static neuron corresponding to the associated stimulus (red) and the plastic neuron. Increasing the plastic neuron's coupling to the rest of the population amplifies the correlation between the pre- and post- synaptic neuron when the associated stimulus is present, since the entire population receives an extra external input. On the other hand, strong coupling amongst the presynaptic static neurons decreases their stimulus selectivity, since they will be more co-active regardless of the stimulus identity. This corrupts the signal during stimulus association. These two effects combine, such that the new stimulus association is learned only when there is low population coupling amongst static neurons (PC_{static}) and high population coupling for the plastic neuron ($PC_{plastic}$) (Figure 4D, labelled Φ). In order to enhance perceptual learning with diverse learning rates, plastic neurons should therefore be more coupled to the rest of the population than stable neurons. Correlated diversity of population coupling and plasticity helps achieve this (Figure 2E), ensuring that neurons best suited to the necessary stimulus representation remain stable, while neurons best suited to learning stimulus associations remain flexible. The plasticity-coupling link therefore efficiently exploits the functional advantages conferred by both diverse learning rates and diverse population coupling.

2.6 Diverse learning rates lead to networks with improved stimulus coding capabilities

Until now we have considered the effect of population coupling on a network's ability to learn stimulus associations. We are also interested in the impact of population coupling on a task that does not involve synaptic plasticity, since the differences in non-specific connectivity alone may affect a neuron's computational capability. We choose stimulus decoding as a simple example, and measure performance at decoding pairs of stimuli in a static network, after it has gone through a period of synaptic plasticity (Methods). We compare three different network types; one which has been developed while it had only slow α, one developed with only fast α, and one developed with diverse α (Figure 4E). In a network with only slow α, and therefore low population coupling, stimulus decoding performs relatively well when there are high levels of noise in the input. Networks with only fast α perform relatively well when there are low levels of noise. A network with diverse α seems to advantageously combine both of these properties, so that its performance is high across the entire range of input strength and noise levels.

Discussion

We have studied the impact of diverse learning rates in a recurrent network model of visual cortex. Intriguingly, a plasticity-coupling link emerges in networks with diverse learning rates, in which neurons with fast learning rates are more coupled to population activity than neurons with slow learning rates. We substantiated a key prediction of our plasticity-coupling link with *in vivo* calcium imaging of mouse visual cortex from the Allen Brain Observatory (ABI, 2016), finding that a neuron is more likely to exhibit stimulus preference variability if it has high population coupling. Based on our findings we propose that the plasticity-coupling link efficiently combines stable and flexible stimulus representation.

3.7 Stability and plasticity of stimulus responses

The architecture of a plastic substrate of neurons on top of a stable 'backbone' (Figure 4B) has been hypothesised before, and there is some compelling experimental evidence for this proposal (Grosmark and Buzsaki, 2016; Clopath and Rose, 2017; Rose et al., 2016; Panas et al., 2015). In particular, tracking of hippocampal cell assemblies reveal subsets of either plastic, highly active neurons or rigid, less active neurons (Grosmark and Buzsaki, 2016). Likewise, a statistical-mechanical analysis of

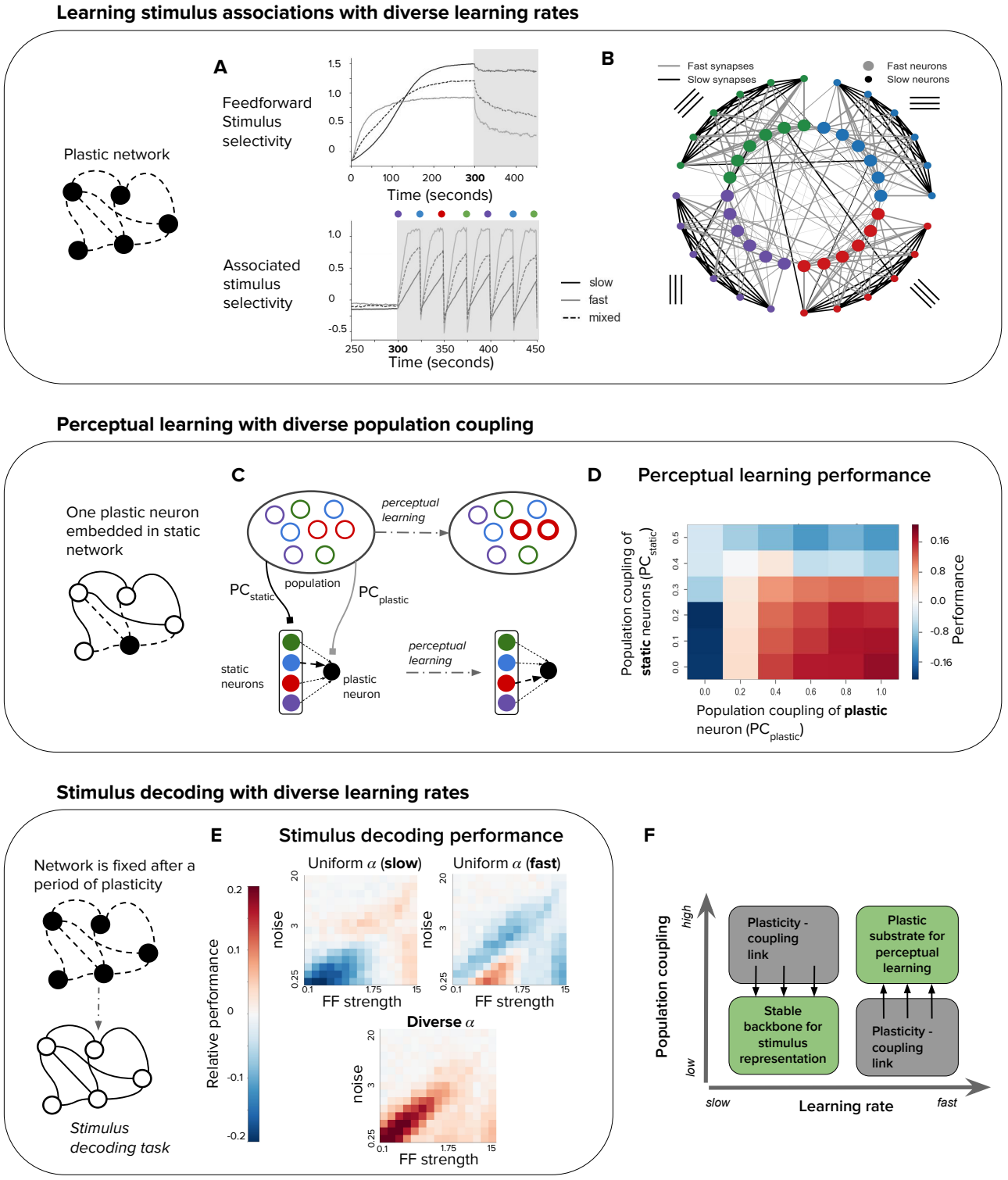

Figure 4: **Plasticity-coupling link enables both robust stimulus representation and a flexible substrate for perceptual learning (A)** The evolution of mean selectivity to the feedforward stimulus (top) and a stimulus associated with an additional external input (bottom) in networks composed either entirely of neurons with slow α (black), fast α (gray), or a mix of both fast and slow α (dashed black). Shaded gray region indicates when the additional external input is present, and the coloured circles indicate the stimulus the external input is associated with at that time (this switches every 25 seconds) **(B)** Synaptic connectivity after plasticity for a network of neurons with slow (small circles) or fast (large circles) learning rates. Neurons in the network receive input selective to 1 of 4 possible stimuli (colour denotes stimulus preference). Synaptic inputs onto fast neurons and slow neurons are coloured gray and black respectively. The spatial organisation of neurons is for visualisation purposes only.

Figure 4: **(C-D)** Investigating the impact of population coupling on perceptual learning. **(C)** Coupling of either the plastic neuron or static neurons to the population is set by adjusting $PC_{plastic}$ and $PC_{plastic}$ respectively. Perceptual learning is simulated through an additional external input whenever the preferred stimulus of the red neurons is present. This leads to the predominant synaptic weight onto the plastic neuron (black) switching from the neuron with the same original preferred stimulus (blue) to the neuron with the associated preferred stimulus (red). **(D)** Amount of perceptual learning which occurs at the plastic neuron, as the population coupling of either the plastic neuron ($PC_{plastic}$, x-axis) or static (PC_{static}, y-axis) is varied. Perceptual learning is quantified by the ratio of the red synaptic weight (associated stimulus) to the blue synaptic weight (original preferred stimulus of the plastic neuron) after plasticity. Red regions (Φ) indicate successful perceptual learning, and occur only when $PC_{plastic}$ is high and PC_{static} is low. **(E)** Relative stimulus decoding performance of fixed recurrent networks after a period of plasticity in order to develop the network. Networks were developed using either entirely neurons with fast learning rates, slow learning rates, or a 50/50 mix of both learning rates. The feedforward stimulus strength (x-axes) and noise (y-axes) were varied along a logarithmic scale. **(F)** Illustration of the synergistic effect of the plasticity-coupling link on perceptual learning. The plasticity-coupling link ensures that slow neurons have low population coupling and fast neurons have high population coupling, which panel **D** demonstrates is necessary for perceptual learning.

network activity in hippocampal cell cultures identified both neurons which are highly active and contribute predominantly to network stability, and neurons which exhibit more long-term activity fluctuations without compromising overall network stability (Panas et al., 2015). In primary visual cortex - which we model - neurons exhibit characteristic fluctuations of their stimulus selectivity during baseline measurements, but nonetheless tend to retain their preferred stimulus following recovery from sensory deprivation (Rose et al., 2016). This provides evidence for a stable 'backbone' of recurrent connectivity which is resistant to sensory perturbations (Clopath and Rose, 2017). Ranson (2017) investigated the stability of locomotion-dependent modulation of visual responses across 14 days and, in contrast to Grosmark and Buzsaki (2016), found that highly responsive neurons exhibited reasonably stable stimulus preference while weakly responsive neurons exhibit plastic stimulus preference. However, these experiments tracked different stimulus features - and over longer timescales - when compared with our study. Moreover, our inclusion of a homeostatic inhibitory plasticity rule that precisely controls excitatory firing rate precludes us from making predictions about the dependence of a neurons average firing rate and its propensity for stimulus preference plasticity (Vogels et al., 2011). Similar links between plasticity and population dynamics could emerge in other experiments that chronically image cortical network activity (Driscoll et al., 2017; Singh et al., 2015; Peron et al., 2015)

Since the majority of experiments which track stimulus preference evolution do so during visual discrimination paradigms, it is likely that top-down influences such as attention or reward modulation play significant roles in their observed dynamics (Pakan et al., 2018; Caras and Sanes, 2017; Poort et al., 2015; Schoups et al., 2001). An exception is Goltstein et al. (2013), in which stimulus preference is measured in the anaesthetised state, meaning that top-down inputs are likely to be absent. Likewise, Ranson (2017) tracked stimulus response stability during passive viewing, similar to the experimental setup of the data we analyse (ABI, 2016). As well as top-down modulation, further features missing from our network model are a realistic inhibitory circuitry (Tremblay et al., 2016; Letzkus et al., 2015), and incorporating changes in network dynamics which occur during sleep (Grosmark and Buzsaki, 2016; Singh et al., 2015), both of which are widely viewed to play an important role in regulating the plasticity of neural representation.

3.8 A plasticity-coupling link *in vivo*

Our analysis of *in vivo* calcium imaging substantiates a key prediction of our network model by observing a correlation between the stimulus preference plasticity of a neuron and its population coupling (Figure 3D). Note that this relationship does not arise in our receptive-field network model with uniform learning rates (Figure 2E), so it is not a trivial consequence of any network model that exhibits diverse population coupling. Although the correlations we measured are quite small, this variability reflects what is observed in our network model (Figure 2E), and is not surprising given that there are likely many unobserved factors - aside from population coupling - which contribute to the dynamics of a neuron's observed stimulus response. Indeed, our network with uniform learning rates demonstrates significant stimulus response variability (Figure 2E, green), but crucially does not

capture the correlation between this variability and population coupling which we observe both *in vivo* and in the network with diverse learning rates.

An advantage of the Allen Brain Observatory is the large amount of data and easily replicable data processing pipeline which allows us to build upon previous work investigating population coupling in the same dataset (Sedigh-Sarvestani et al., 2017). Since the population coupling of a neuron is correlated across brain states, and is only weakly dependent on stimulus type and mean fluorescence, we believe that it provides a good measure of a neurons functional integration within the local network, and would - according to our model - therefore provide a reasonable estimate of its propensity for perceptual learning (Figure 4D) (Sedigh-Sarvestani et al., 2017; Okun et al., 2015). In agreement with our network model, and in contrast with observations from (Okun et al., 2015), population coupling is anti-correlated with orientation selectivity in the Allen Brain Observatory dataset (Figure S2). The disparity between these two experiments could be due to different experimental conditions, or the effect may not have been previously observed due to the smaller number of neurons used previously (n=431 in (Okun et al., 2015)). Moreover, our observation that the changes in stimulus preferences ($\Delta \mathrm{ORI}_{\mathrm{pref}}$) are often non-zero but skewed towards small absolute values (Figure 3C) are in agreement with the hypothesis that stimulus preference is a slowly drifting property (Rose et al., 2016). Unfortunately, the experimental protocol limits us to directly comparing stimulus preference at only two timepoints; the beginning and end of a 62 minute imaging session (Figure 3A). Nonetheless, significant changes in synaptic efficacies can be expressed within this time (Meyer et al., 2014). We hope that these findings will stimulate further experiments that allow us to more precisely test for the presence of a plasticity-coupling link across longer timepoints, and during learning.

3.9 Population coupling and neuron function

Our network model provides a parsimonious explanation for the diverse population coupling recently observed in sensory cortex (Okun et al., 2015). Population coupling is dependent on the amount of recurrent synaptic input a neurons receives, in agreement with experimental data (Figure 2C). Note that this dependence is not present in networks with uniform learning rates, even though they too exhibit diverse population coupling. Moreover, the width of the population coupling distribution increases as the recurrent network approaches a dynamic regime dominated by high noise and diverse selectivity, typical in cortical networks (Figure 2F). These findings suggest that different population couplings may simply be a feature of varying learning rates and does not necessarily mean (although we cannot exclude it) that the observed diversity reflects entirely different cell classes. Furthermore, one can imagine alternative mechanisms that lead to diverse population coupling in recurrent networks, such as imposing heterogeneous targets for the number of synaptic inputs received by each neuron. Investigating such alternative mechanisms was outside the scope of our study, but would provide an interesting avenue for further theoretical research.

The proposed plasticity-coupling link presents a counterintuitive interpretation of the role of 'soloists' and 'choristers' originally described by Okun et al. (2015). While one may naively suppose that the weakly coupled 'soloists' are suited to undergo plasticity during learning, we propose that it is in fact the strongly coupled 'choristers' with a more plastic representation.

The functional impact of population coupling on learning is crucial: in order to enhance perceptual learning, plastic neurons in recurrent networks should be more coupled to the rest of the population than stable neurons (Figure 4D,F). We find that high population coupling helps plastic neuron change their stimulus preference towards an associated stimulus, but hinders the ability of stable neurons to provide an instructive signal for learning. Correlated diversity of population coupling and learning rate therefore enables both robust stimulus representation (low α, PC) and a flexible substrate suitable for perceptual learning (high α, PC). Strikingly, this relationship is precisely what the predicted plasticity-coupling link ensures (Figure 4E). Moreover, a recent theoretical study of sensory decoding proposed that untuned neurons contribute to decoding when they are correlated with tuned neurons Zylberberg (2017). Again, this is the relationship predicted by our model, since plastic neurons are less tuned than rigid neurons and are more strongly coupled to the population (Figure 1D,E).

3.10 Previous theoretical work

There are many previous theoretical explorations of how diversity in the synaptic plasticity of individual neurons affects learning. A recent study proposes a conceptually similar mechanism for

modulating the stability or flexibility of memory formation, by implementing either symmetric or asymmetric STDP learning rules (Park et al., 2017). Diversity in synaptic learning rates was also explored within the traditional machine learning framework, whereby fast weights store temporary memories of recent events, compared with slow weights which capture regularities in input structure (Ba et al., 2016). Our work is related to previous approaches for overcoming catastrophic forgetting, which is often observed in neural networks during learning (Grossberg, 1987; Carptenter and Grossberg, 1987; McClelland et al., 1995; Fusi et al., 2005; Roxin and Fusi, 2013; Benna and Fusi, 2016). These approaches typically involve partitioning memories across timescales by implementing either synaptic states with different timescales, or neural architectures with different timescales. Here, we intead based our approach on experimental observations that suggest diverse learning rates within a sensory cortical network (Ranson, 2017; Clopath and Rose, 2017; Rose et al., 2016; Poort et al., 2015; Lütcke et al., 2013). Finally, individual synaptic updates in our model are defined by the learning rate of the postsynaptic neuron (Equation 3). Further work could explore whether our observed outcomes change if updates are instead dependent on the learning rate of the presynaptic neuron.

The plasticity-coupling link's impact on perceptual learning suggests a dichotomy of roles amongst neurons in a network, tied to a particular functional architecture: a stable 'backbone' of stimulus representation formed by neurons with slow synaptic plasticity and low population coupling, on top of which lies a flexible substrate of neurons with fast synaptic plasticity and high population coupling. Diverse learning rates naturally enable this architecture, and offer a compelling candidate mechanism for mediating both forms of diversity - population coupling and stimulus response stability - recently observed in cortical networks. Finally, the plasticity-coupling link provides neuroscientists with a means to assess the tendency of particular neurons to influence future learning: those which are highly coupled to population activity are most likely to express plasticity. Ongoing advances in chronic multi-neuron calcium imaging, alongside neuron-specific optogenetic stimulation, will allow us to further probe and exploit these possibilities.

Methods

Our network model simulations were written in python with numpy and scipy.

4.11 Neuron model

For both the fully connected and the receptive field networks we use a simple firing rate neuron model, given by the transfer function $g(x)$ defined below, and as used previously by Rajan et al. (2010); Hennequin et al. (2014).

$$\begin{aligned} g(x) &= 0 \text{ if } x < 0 \\ &= (r_{\max} - r_0)\tanh\big(x/(r_{\max} - r_0)\big) \text{ if } x \geq 0. \end{aligned} \tag{1}$$

This leads to firing rates with a baseline of r_0 and a maximum of $r_{\max}$. Following Rajan et al. (2010), the firing rates y_i of neuron i driven by external input H_i in a network are described below.

$$\frac{dy_i}{dt} = -y_i + \sum_{j=1}^{N} W_{ij} g(y_j) + H_i \,, \tag{2}$$

where W_{ij} is the weight of the synaptic connection from neuron j to neuron i.

4.12 Modelling synaptic plasticity with diverse learning rates

We use a simple Hebbian learning rule with homeostatic synaptic scaling to model synaptic plasticity of recurrent excitatory to excitatory (E-E) synapses (Gerstner and Kistler, 2002),

$$\frac{dW_{ij}^{EE}}{dt} = \alpha_i y_i y_j - \zeta\big(\sum_{k=1}^{N_E} W_{ik}^{EE} - W_{\text{total}}^{EE}\big) \tag{3}$$

where α_i is the learning rate of the postsynaptic neuron and y_j and y_i are the activities of the pre- and postsynaptic neuron respectively. ζ is the time constant of synaptic scaling, and W_{total}^{EE} is the target amount of total recurrent synaptic input which each neuron can receive under the synaptic scaling rule.

This form of excitatory plasticity introduces competition amongst presynaptic synaptic weights and leads to the development of stimulus selectivity, as discussed in Ko et al. (2013). We use a homeostatic rule to model inhibitory synaptic plasticity of recurrent inhibitory to excitatory (I-E) weights (Vogels et al., 2011),

$$\frac{d\mathrm{W}_{ij}^{IE}}{dt} = \eta y_j (y_i - y_0) \,, \tag{4}$$

where y_0 is the homeostatic target firing rate, η is the learning rate, and W_{ij}^{IE} is the weight of the synaptic connection from inhibitory neuron j to excitatory neuron i.

Excitatory weights are bounded so that their values lie between 0 and $w_{\max}$, and inhibitory weights are bounded so that they lie between $-w_{\text{max-inh}}$ and 0.

While including two homeostatic mechanisms in our network model may seem redundant, they play different regulatory roles. Inhibitory plasticity largely controls the balance of excitation and inhibition received by a neuron, ensuring that it operates within its dynamic range. Synaptic scaling ensures that the total amount of recurrent excitation in the network is kept fixed as we vary its external input, while also introducing competition between presynaptic weights so that stimulus selectivity emerges. The synergistic effect of including multiple forms of plasticity has been widely studied in theoretical studies (Zenke et al., 2015; Litwin-Kumar and Doiron, 2014; Triesch, 2007; Clopath et al., 2016).

Note that the speed of all learning rates α, ζ, and η are artificially increased in order to reduce the computational times resources required to simulate our network model. The timescales of synaptic plasticity in our network models are in the order of hundreds of seconds, while synaptic plasticity

during perceptual learning occurs over the course of days in vivo. This increased learning rate does not qualitatively affect our results, as there is a sufficient separation of timescales between synaptic plasticity and network dynamics, and is a standard assumption in network models of synaptic plasticity (Zenke et al., 2015; Litwin-Kumar and Doiron, 2014).

4.13 Fully connected network model

The fully connected network consists of N_E excitatory neurons and a global inhibitory neuron ($N_I = 1$). The dynamics of both inhibitory (I) and excitatory (E) neurons are described by Equation 1 and Equation 2. There is dense all-to-all synaptic connectivity in the E-E, E-I and I-E populations, and no I-I connectivity. Self-connections, or autapses, are not permitted in this network. W in Equation 2 is a square matrix with $(N_E + N_I)^2$ elements.

For Figure 1 we use a network with 48 excitatory neurons, and 4 input stimuli. Each neuron i has a preferred stimulus $\theta_{\mathrm{i}}^{\mathrm{pref}}$, such that there are 12 neurons corresponding to each input stimulus. Each neuron receives its' preferred stimulus input H_{stim}, and an independent noise source generated by an Ornstein-Uhlenbeck process, OU, with a mean of 0, variance of σ_{OU} and correlation time τ_{OU}. The external input H_i to a neuron i is therefore given by

$$H_i(t) = \delta(\theta_{\mathrm{i}}^{\mathrm{pref}} - \theta_{\mathrm{input}}(t))H_{\mathrm{stim}} + \mathrm{OU}_i(t)\ . \tag{5}$$

For Figure 1B-D, these input groups are further divided so that there are 6 slow neurons (with $\alpha_i = \alpha_{\mathrm{s}}$) and 6 fast neurons (with $\alpha_i = 5\alpha_{\mathrm{s}}$) per group. For Figure 1E, each input group of 12 neurons contains a single neuron corresponding to each of the 12 learning rates. The learning rates are logarithmically spaced between $0.5\alpha_{\mathrm{s}}$ and $75\alpha_{\mathrm{s}}$.

All excitatory-to-inhibitory synapses are uniformly initialised with weights $W_{\mathrm{init}}^{\mathrm{EE}}$, excitatory-to-inhibitory synapses with weights $W_{\mathrm{init}}^{\mathrm{EI}}$, and inhibitory-to-excitatory synapses with weights $W_{\mathrm{init}}^{\mathrm{IE}}$. We simulate the evolution of synaptic weights during visually evoked activity by sequentially presenting the network with a randomly chosen stimulus from the 4 input stimuli. Each stimulus is presented for 500 ms. The total simulation time is 500 seconds, and synaptic weights are updated at each timestep with the learning rules given by Equation 3 and Equation 4.

For Figure 1D, connection specificity is defined as the average ratio of specific to non-specific excitatory synaptic weights received by neurons. Synaptic inputs from neurons in the same input stimulus group as the postsynaptic neuron are specific (i.e. they share the same feedforward stimulus preference), while all other synaptic inputs are non-specific. Specificity fluctuations are defined as the standard deviation of the connection specificity over time, where specificity is sampled every second from 200 to 500 seconds.

Table 1: **Simulation Parameters**

H_{stim}	8	r_0	1.0	r_{max}	20.0	dt	0.05 ms
α_{s}	2.0x10^{-6} Hz	ζ	2x10^{-4} Hz	η	1.0x10^{-5} Hz	y_0	5
w_{max}	0.042	$w_{\text{max-inh}}$	50	W_{total}^{EE}	0.75		
σ_{OU}	1	τ_{OU}	10 ms	$W_{\mathrm{init}}^{\mathrm{EE}}$	$0.5w_{\mathrm{max}}$	$W_{\mathrm{init}}^{\mathrm{IE}}$	0.2

4.14 Measuring population coupling

As introduced by Okun et al. (2015), the population coupling PC_i of a neuron i is measured by calculating the Pearson correlation coefficient of each neurons' activity x_i with the average activity of the rest of the population;

$$\mathrm{PC}_i = \mathrm{corr}\big(x_i, \frac{1}{N-1}\sum_{j\neq i}^{N} x_j\big)\ . \tag{6}$$

Synaptic weights are kept fixed during the population coupling measurement, while external input is as in Equation 5. We measured population coupling using 250 seconds of activity.

4.15 Receptive-field based network model

For Figure 2, we adapt a previously developed model of receptive field properties in mouse visual cortex (Watanabe et al., 2016). We add neuronal dynamics and, beginning with uniform connectivity, simulate synaptic plasticity as visual stimuli are presented to the network. This model is constructed by assigning receptive fields to each excitatory neuron from a 2D Gabor function,

$$\begin{aligned} \mathrm{RF}(x', y') &= \mathrm{A}\ \exp\Big(\frac{-x'^2}{2\sigma_x^2} - \frac{-y'^2}{2\sigma_y^2}\Big)\cos(2\pi f x' + \phi) \\ x' &= x\ \cos\theta - y\ \sin\theta \\ y' &= x\ \sin\theta + y\ \cos\theta \end{aligned} \tag{7}$$

where A is the amplitude, σ_x and σ_y are the standard deviations of the Gaussian, θ is the orientation, f is the frequency and ϕ is the phase of the receptive field. A network of 250 excitatory neurons with receptive fields is randomly generated from Equation 7, with $f = 2$, $\sigma_x = \sigma_y = 0.5$, $\phi \sim (0, 2\pi)$, $\theta \sim \{\pi/4, \pi/2, 3\pi/4, ..., 2\pi\}$. As in the previous network, there is a single inhibitory neuron which all excitatory neurons project to, and receive inhibition from.

Neurons are rate-based and have similar dynamics as in the simple network model (Equation 1,Equation 2). Synaptic plasticity is also governed by the same learning rules (Equation 3,Equation 4). Inputs are presented to the network in the form of 2D images, and the input to each neuron i for a given image I_{ext} is determined by the pixel-wise dot product of that image with the neurons' receptive field RF_i, in addition to an independent noise term for each neuron given by an Ornstein-Uhlenbeck process;

$$H_i(t) = I_{\text{ext}} \cdot \mathrm{RF}_i + \mathrm{OU}_i(t)\ . \tag{8}$$

All excitatory-to-inhibitory synapses are uniformly initialised with weights $W^{\text{EE}}_{\text{init-RF}}$ and inhibitory-to-excitatory synapses with weights $W^{\text{IE}}_{\text{init-RF}}$. We simulate the evolution of synaptic weights during visually-evoked activity by sequentially presenting the network with randomly chosen bars of different orientations . Each image is presented for 500 ms. The total simulation time is 500 seconds, and synaptic weights are updated at each timestep. All results in Figure 2 are pooled from 15 independent network instances, with 250 excitatory neurons in each network instance.

We define the selectivity of each neuron as $\bar{w}_{\text{specific}} - \bar{w}_{\text{non-specific}}$, where w_{specific} are the synaptic weights from neurons which share the same receptive field orientation and $w_{\text{non-specific}}$ are the synaptic weights from neurons which have a different receptive field orientation.

4.16 The Allen Brain Observatory: 2-photon calcium imaging of visual responses *in vivo*

We use data from the Allen Brain Observatory, a publicly available and curated survey of neural activity in adult mouse visual cortex. A comprehensive description of the experimental methods, data acquisition and data analyses are available as white papers published by the Allen Brain Institute (ABI, 2016).

Briefly, GCaMP6F was expressed in forebrain excitatory neurons of transgenic mice line Ai93. Cranial surgery was performed to insert a window between p37-p63, followed by 2 weeks of habituation to the experiment setting. Mice were head-fixed on top of a rotating disk and could walk freely. 2-photon imaging experiments were conducted as the mouse passively viewed the stimulus protocol on a screen. The stimulus protocols included in our analysis consisted of either i) 10 minutes of drifting gratings, followed by 42 minutes of interleaved natural movies, drifting gratings and spontaneous activity, followed by another 10 minutes of drifting gratings, or ii) 8 minutes of static gratings, followed by 45 minutes of interleaved natural movies, static gratings and spontaneous activity, followed by 9 minutes of static gratings (Figure 3A, see white paper for further details). Drifting gratings were presented at 8 uniformly separated directions and at 5 different temporal frequencies. Static gratings were presented at 6 uniformly separated orientations separated, 5 different spatial frequencies and 4 different phases. 112 imaging experiments were initially included in our analysis.

4.17 Measuring population coupling and stimulus response variability *in vivo*

The Allen Brain Institute API provides functions which allow us to extract fluorescence traces and measure average stimulus response properties of individual cells during an experiment (ABI,

2016). Motion correction, ROI detection and segmentation, and the removal of neuropil fluorescence artefacts are automatically performed using the API. We customised scripts within this API so that we could measure population coupling and stimulus response properties under specific conditions and timeframes.

We measure population coupling similarly to the network model analysis (Equation 6), but where a neuron's activity is represented by calcium fluorescence dF/F. To ensure a reliable estimate of the population activity when calculating population coupling, we exclude any experiments in which less than 50 neurons were recorded. We also exclude any experiments in which the population couplings are not sufficiently consistent when using either half (randomly chosen) of the neurons to estimate population activity ($r^2 < 0.8$, linear regression). This reduces the number of experiments in our analysis from 112 to 64, for a total of 15,281 neurons.

We measure the preferred orientation of each neuron during both the first presentation of static or drifting gratings ($\text{ORI}_{\text{pref-1}}$), and the final presentation of static or drifting gratings ($\text{ORI}_{\text{pref-2}}$) (Figure 3A). The preferred orientation is defined as the grating that evoked the largest mean response across all trials. Note that each experiment contains only either static or drifting gratings, so there is no overlap between these two conditions. The absolute difference in preferred orientation is calculated as: $\Delta\text{ORI}_{\text{pref}} = |\text{ORI}_{\text{pref-1}} - \text{ORI}_{\text{pref-2}}|$.

4.18 Simulating perceptual learning

For the perceptual learning simulation in Figure 4A, the input H_i to each neuron is simulated as before, but with an additional term which is active whenever the stimulus associated with the additional external input is present (Equation 9). We first simulate synaptic plasticity without any stimulus associations for 300 seconds (i.e. with $H_{\text{associated}} = 0$), and then simulate perceptual learning (with $H_{\text{associated}} = 10$). The identity of the associated stimulus is changed every 25 seconds to simulate continual learning.

$$H_i(t) = \delta(\theta_{\text{i}}^{\text{pref}} - \theta_{\text{input}}(t))H_{\text{stim}} + \delta(\theta^{\text{associated}} - \theta_{\text{input}}(t))H_{\text{associated}} + \text{OU}_i(t) \tag{9}$$

Feedforward stimulus selectivity is defined as $\frac{\bar{w}_{\text{specific}}}{\bar{w}_{\text{non-specific}}} - 1$, where w_{specific} are the synaptic weights from neurons which share the same feedforward stimulus preference and $w_{\text{non-specific}}$ are the synaptic weights from neurons which have a different feedforward stimulus preference. Likewise, associated stimulus selectivity is defined as $\frac{\bar{w}_{\text{associated}}}{\bar{w}_{\text{non-associated}}} - 1$, where $w_{\text{associated}}$ are the synaptic weights from neurons whose feedforward stimulus preference is the associated stimulus, and $w_{\text{non-associated}}$ are the synaptic weights from other neurons.

4.19 Single plastic neuron embedded in a static network

In order to systematically investigate the effect of population coupling on perceptual learning, we must keep the population coupling of both the static and plastic population fixed throughout the experiment. Since changes in the synaptic weights connecting both these populations will alter their population coupling, we overcome this by making these particular synapses functionally silent. That is, while their synaptic weight is updated depending on pre- and post-synaptic activity as before (Equation 3), these synapses do not contribute when calculating the activity of the static and plastic neurons. The activities of the static and plastic neurons (y^{static} and y^{plastic}) are therefore only determined by their external input and the activities of the population (y^{pop}, Figure 4C), meaning that their population coupling can be systematically varied:

$$\begin{aligned}
\frac{dy_i^{\text{pop}}}{dt} &= -y_i^{\text{pop}} + \sum_{j=i}^{N} W_{ij} g(y_j^{\text{pop}}) + H_i \,, \\
\frac{dy_i^{\text{plastic}}}{dt} &= -y_i^{\text{plastic}} + \text{PC}_{\text{plastic}} \sum_{j=i}^{N} W_{ij} g(y_j^{\text{pop}}) + H_i \,, \\
\frac{dy_i^{\text{static}}}{dt} &= -y_i^{\text{static}} + \text{PC}_{\text{static}} \sum_{j=i}^{N} W_{ij} g(y_j^{\text{pop}}) + H_i \,.
\end{aligned} \tag{10}$$

W_{ij} is fixed for the duration of the simulation, while the synaptic weights from the static to the plastic population are updated as below;

$$\frac{dW_{ij}^{plastic}}{dt} = \alpha y_i^{\text{static}} y_j^{\text{plastic}} - \zeta\Big(\sum_{k=1}^{N_{\text{plastic}}} W_{kj}^{plastic} - W_{\text{total}}^{EE}\Big)\,. \tag{11}$$

As before, the input H_i has an additional term which is active whenever the stimulus associated with the additional external input is present (Equation 9, i.e. whenever the red stimulus is being presented). We first simulate synaptic plasticity without any stimulus associations for 500 seconds (i.e. with $H_{\text{associated}} = 0$), and then simulate perceptual learning (with $H_{\text{associated}} = 10$) for 100 seconds. Perceptual learning is quantified by the ratio of the red synaptic weight (associated stimulus) to the blue synaptic weight (original preferred stimulus of the plastic neuron) after plasticity

4.20 Measuring stimulus decoding performance

We train a perceptron to decode the stimulus identity from the individual activity of all neurons in the network, using the scikit-learn python package. The average activity of each neuron across a 500 ms sampling period are used as inputs during training. For Figure 4E, performance at decoding pairs of stimuli simultaneously presented to the network is shown. Relative deviation from the average performance of a perceptron trained to decode pairs of stimuli (28 possible pairs from 8 stimuli) over all 3 network types is calculated. The relative deviations for each network type from the average across all networks types are shown (Figure 4E).

4.21 Code availability

Code will be made publicly available on github and modeldb, and can be made available to reviewers.

Acknowledgements

We are grateful to Tobias Rose and Alex Cayco-Gajic for their invaluable scientific feedback and comments on the manuscript. This work was funded by the Wellcome Trust and the BBSRC.

References

(2016). Allen Brain Observatory. *http://observatory.brain-map.org/visualcoding/*, (June):1–24.

Ajemian, R., D'Ausilio, A., Moorman, H., and Bizzi, E. (2013). A theory for how sensorimotor skills are learned and retained in noisy and nonstationary neural circuits. *Proceedings of the National Academy of Sciences of the United States of America*, 110(52):E5078–87.

Andermann, M. L. (2010). Chronic cellular imaging of mouse visual cortex during operant behavior and passive viewing. *Frontiers in Cellullar Neuroscience*, 4(March):1–16.

Ba, J., Hinton, G., Mnih, V., Leibo, J. Z., and Ionescu, C. (2016). Using Fast Weights to Attend to the Recent Past. pages 1–9.

Benna, M. K. and Fusi, S. (2016). Computational principles of synaptic memory consolidation. *Nature neuroscience*, 19(12):1697.

Caras, M. L. and Sanes, D. H. (2017). Top-down modulation of sensory cortex gates perceptual learning. *Proceedings of the National Academy of Sciences*, 114(37):201712305.

Carptenter, G. A. and Grossberg, S. (1987). Discovering Order in Chaos: Stable Self-Organization of Neural Recognition Codes. *Annals of the New York Academy of Sciences*, 504(1):33–51.

Clopath, C., B"using, L., Vasilaki, E., and Gerstner, W. (2010). Connectivity reflects coding: a model of voltage-based {STDP} with homeostasis. *Nature neuroscience*, 13(3):344–352.

Clopath, C. and Rose, T. (2017). Variance and Invariance of Neuronal Long-term Representations. *Philosophical Transactions of the Royal Society B: Biological Sciences*.

Clopath, C., Vogels, T. P., Froemke, R. C., and Sprekeler, H. (2016). Receptive field formation by interacting excitatory and inhibitory synaptic plasticity. Technical report.

Cossell, L., Iacaruso, M. F., Muir, D. R., Houlton, R., Sader, E. N., Ko, H., Hofer, S. B., and Mrsic-flogel, T. D. (2015). Functional organization of excitatory synaptic strength in primary visual cortex. *Nature*, 518(00):1–5.

Driscoll, L. N., Pettit, N. L., Minderer, M., Chettih, S. N., and Harvey, C. D. (2017). Dynamic Reorganization of Neuronal Activity Patterns in Parietal Cortex. *Cell*, pages 986–999.

Fusi, S., Drew, P. J., and Abbott, L. F. (2005). Cascade models of synaptically stored memories. *Neuron*, 45(4):599–611.

Gerstner, W. and Kistler, W. M. (2002). Mathematical formulations of Hebbian learning. *Biological cybernetics*, 87(5-6):404–415.

Goltstein, P. M., Coffey, E. B. J., Roelfsema, P. R., and Pennartz, C. M. A. (2013). In Vivo Two-Photon Ca2+ Imaging Reveals Selective Reward Effects on Stimulus-Specific Assemblies in Mouse Visual Cortex. *Journal of Neuroscience*, 33(28):11540–11555.

Grosmark, A. D. and Buzsaki, G. (2016). Diversity in neural firing dynamics supports both rigid and learned hippocampal sequences. *Science*, 351(6280):1440–1443.

Grossberg, S. (1987). Processing of expected and unexpected events during conditioning and attention: A psychophysiological theory. *Advances in Psychology*, 42(C):181–237.

Hennequin, G., Vogels, T. P., and Gerstner, W. (2014). Optimal control of transient dynamics in balanced networks supports generation of complex movements. *Neuron*, 82(6):1394–1406.

Kappel, D., Habenschuss, S., Legenstein, R., and Maass, W. (2015). Network Plasticity as Bayesian Inference. *{PLoS} computational biology*, 11(11):e1004485.

Kappel, D., Legenstein, R., Habenschuss, S., Hsieh, M., and Maass, W. (2017). Reward-based stochastic self-configuration of neural circuits. pages 1–37.

Ko, H., Cossell, L., Baragli, C., Antolik, J., Clopath, C., Hofer, S. B., and Mrsic-Flogel, T. D. (2013). The emergence of functional microcircuits in visual cortex. *Nature*, 496(7443):96–100.

Letzkus, J. J., Wolff, S. B. E., and Lüthi, A. (2015). Disinhibition, a Circuit Mechanism for Associative Learning and Memory. *Neuron*, 88(2):264–276.

Lin, I.-C., Okun, M., Carandini, M., and Harris, K. D. (2015). The Nature of Shared Cortical Variability. *Neuron*, 87(3):1–13.

Litwin-Kumar, A. and Doiron, B. (2014). Formation and maintenance of neuronal assemblies through synaptic plasticity. *Nature communications*, 5(May):5319.

Lütcke, H., Margolis, D. J., and Helmchen, F. (2013). Steady or changing? Long-term monitoring of neuronal population activity. *Trends in Neurosciences*, 36(7):375–384.

Mank, M., Santos, A. F., Direnberger, S., Mrsic-Flogel, T. D., Hofer, S. B., Stein, V., Hendel, T., Reiff, D. F., Levelt, C., Borst, A., Bonhoeffer, T., Hübener, M., and Griesbeck, O. (2008). A genetically encoded calcium indicator for chronic in vivo two-photon imaging. *Nature Methods*, 5(9):805–811.

McClelland, J. L., McNaughton, B. L., and O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, 102(3):419–457.

Meyer, D., Bonhoeffer, T., and Scheuss, V. (2014). Balance and stability of synaptic structures during synaptic plasticity. *Neuron*, 82(2):430–443.

Okun, M., Steinmetz, N. A., Cossell, L., Iacaruso, M. F., Ko, H., Barthó, P., Moore, T., Hofer, S. B., Mrsic-Flogel, T. D., Carandini, M., and Harris, K. D. (2015). Diverse coupling of neurons to populations in sensory cortex. *Nature*.

Pakan, J. M., Francioni, V., and Rochefort, N. L. (2018). Action and learning shape the activity of neuronal circuits in the visual cortex. *Current opinion in neurobiology*, 52:88–97.

Panas, D., Amin, H., Maccione, a., Muthmann, O., van Rossum, M., Berdondini, L., and Hennig, M. H. (2015). Sloppiness in Spontaneously Active Neuronal Networks. *Journal of Neuroscience*, 35(22):8480–8492.

Park, Y., Choi, W., and Paik, S.-B. (2017). Symmetry of learning rate in synaptic plasticity modulates formation of flexible and stable memories. *Scientific Reports*, 7(1):5671.

Peron, S. P., Freeman, J., Iyer, V., Guo, C., and Svoboda, K. (2015). A Cellular Resolution Map of Barrel Cortex Activity during Tactile Behavior. *Neuron*, 86(3):783–799.

Poort, J., Khan, A. G., Pachitariu, M., Nemri, A., Orsolic, I., Krupic, J., Bauza, M., Sahani, M., Keller, G. B., Mrsic-Flogel, T. D., and Hofer, S. B. (2015). Learning Enhances Sensory and Multiple Non-sensory Representations in Primary Visual Cortex. *Neuron*, 86(6):1478–1490.

Rajan, K., Abbott, L., and Sompolinsky, H. (2010). Stimulus-dependent suppression of chaos in recurrent neural networks. *Physical Review E*, 82.

Ranson, A. (2017). Stability and Plasticity of Contextual Modulation in the Mouse Visual Cortex. *Cell Reports*, 18(4):840–848.

Rokni, U., Richardson, A. G., Bizzi, E., and Seung, H. S. (2007). Motor Learning with Unstable Neural Representations. *Neuron*, 54(4):653–666.

Rose, T., Jaepel, J., Hübener, M., and Bonhoeffer, T. (2016). Cell-specific restoration of stimulus preference after monocular deprivation in the visual cortex. *Science (New York, N.Y.)*, 352(6291):1319–22.

Roxin, A. and Fusi, S. (2013). Efficient Partitioning of Memory Systems and Its Importance for Memory Consolidation. *PLoS Computational Biology*, 9(7).

Schoups, A., Vogels, R., Qian, N., and Orban, G. (2001). Practising orientation identification improves orientation coding in V1 neurons. *Nature*, 412(6846):549–553.

Sedigh-Sarvestani, M., Nolte, M., and Mardoum, P. (2017). Population coupling in the mouse visual cortex. In *Computational and systems neuroscience (Cosyne)*, pages Poster III–44.

Sengpiel, F., Stawinski, P., and Bonhoeffer, T. (1999). Influence of experience on orientation maps in cat visual cortex. *Nature neuroscience*, 2(8):727.

Singh, A., Peyrache, A., and Humphries, M. D. (2015). Plasticity of a neural dictionary in prefrontal cortex.

Tremblay, R., Lee, S., and Rudy, B. (2016). GABAergic Interneurons in the Neocortex: From Cellular Properties to Circuits. *Neuron*, 91(2):260–292.

Triesch, J. (2007). Synergies between intrinsic and synaptic plasticity mechanisms. *Neural computation*, 19(4):885–909.

Turrigiano, G. G., Leslie, K. R., Desai, N. S., Rutherford, L. C., and Nelson, S. B. (1998). Activity-dependent scaling of quantal amplitude in neocortical neurons. *Nature*, 391(6670):892–896.

Vogels, T. P., Sprekeler, H., Zenke, F., Clopath, C., and Gerstner, W. (2011). Inhibitory plasticity balances excitation and inhibition in sensory pathways and memory networks. *Science {(New} York, {N.Y.)}*, 334(6062):1569–1573.

Watanabe, K., Teramae, J.-n., and Wakamiya, N. (2016). *Inferred Duality of Synaptic Connectivity in Local Cortical Circuit with Receptive Field Correlation*, pages 115–122. Springer International Publishing, Cham.

Zenke, F., Agnes, E. J., and Gerstner, W. (2015). Diverse synaptic plasticity mechanisms orchestrated to form and retrieve memories in spiking neural networks. *Nature Communications*, 6:6922.

Zylberberg, J. (2017). Untuned but not irrelevant: A role for untuned neurons in sensory information coding. *bioRxiv*, pages 1–18.

Supplementary materials

Methods for Figure S1

We simulate a network of 1 postsynaptic neuron and 10 presynaptic neurons. 3 of the presynaptic neurons share the same stimulus preference as the postsynaptic neuron and the remaining have different preferred stimuli. We then identify the parameter regime in which the coupling of the single plastic neuron to the rest of the population is correlated to its learning rate. To do so, we measure the population coupling of the postsynaptic neuron for a range of different learning rates from $0.5\alpha_s$ to $10\alpha_s$, using a separate network instantiation for each value of α. We then estimate the slope of the relationship between population coupling and α using linear regression, across a range of values for the synaptic scaling rate (ζ) and noise magnitude (σ_{OU}).

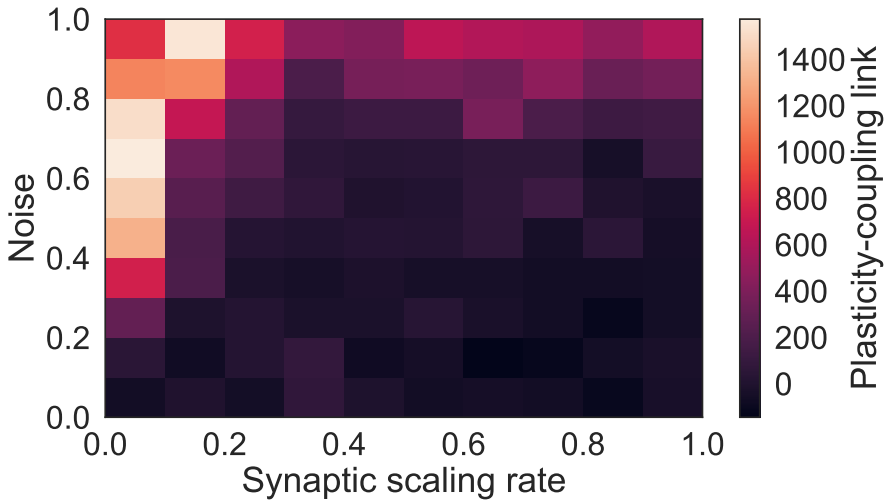

Figure S1: **Plasticity-coupling link requires moderate noise and slow synaptic scaling**. The slope of the linear dependence between α and population coupling within a network with diverse α, for different values of the synaptic scaling rate (ζ) and injected noise (σ_{OU}). The presence of noise introduces transient correlations across the network which lead to fluctuations of both specific and non-specific synaptic weights. This ensures that non-specific synaptic weights do not all tend towards zero. Likewise, if synaptic scaling is too fast compared with Hebbian plasticity - contrary to experimentally observed timescales (Turrigiano et al., 1998) - then only specific synapses, which share highly correlated inputs, can sustain strong weights while non-specific synapses tend towards zero.

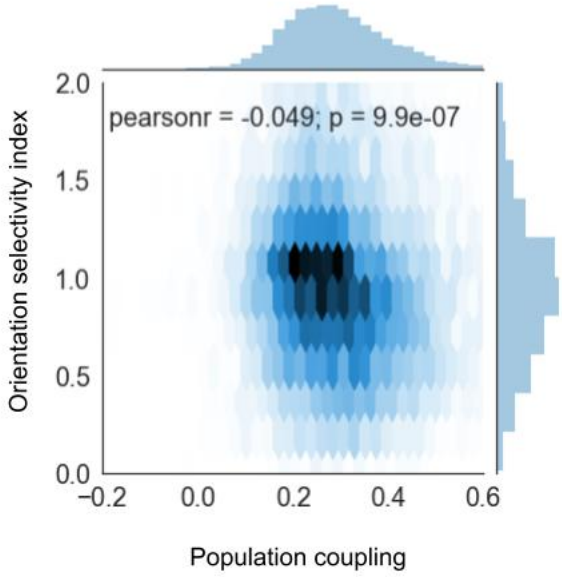

Figure S2: **Orientation selectivity index is anti-correlated with population coupling**.