Mechanisms underlying reshuffling of visual responses by optogenetic stimulation in mice and monkeys

A. Sanzeni[1,2,3], A. Palmigiano[2,*], T.H. Nguyen[2,4,*], J. Luo[5], J.J. Nassi[5], J.H. Reynolds[5], M.H. Histed[6], K.D. Miller[2,7,†], and N. Brunel[3,8,†]

[1]Department of Computing Sciences, Bocconi University, Milan, Italy
[2]Center for Theoretical Neuroscience and Mortimer B Zuckerman Mind Brain Behavior Institute, Columbia University, New York, NY, USA
[3]Department of Neurobiology, Duke University, Durham, NC, USA
[4]Department of Physics, Columbia University, New York, NY, USA
[5]Systems Neurobiology Laboratory, The Salk Institute for Biological Studies, La Jolla, CA, USA
[6]National Institute of Mental Health Intramural Program, NIH, Bethesda, MD, USA
[7]Department of Neuroscience, Swartz Program in Theoretical Neuroscience, Kavli Institute for Brain Science, College of Physicians and Surgeons and Mortimer B. Zuckerman Mind Brain Behavior Institute, Columbia University, New York City, NY, USA
[8]Department of Physics, Duke University, Durham, NC, USA
[*]These authors made equal contributions.
[†]These authors made equal contributions.

November 8, 2022

Abstract

The ability to optogenetically perturb neural circuits opens an unprecedented window into mechanisms governing circuit function. We analyzed and theoretically modeled neuronal responses to visual and optogenetic inputs in mouse and monkey V1. In both species, optogenetic stimulation of excitatory neurons strongly modulated the activity of single neurons, yet had weak or no effects on the distribution of firing rates across the population. Thus, the optogenetic inputs reshuffled firing rates across the network. Key statistics of mouse and monkey responses lay on a continuum, with mice/monkeys occupying the low/high rate regions, respectively. We show that neuronal reshuffling emerges generically in randomly connected excitatory/inhibitory networks, provided the coupling strength (combination of recurrent coupling and external input) is sufficient that powerful inhibitory feedback cancels the mean optogenetic input. A more realistic model, distinguishing tuned visual vs. untuned optogenetic input in a structured network, reduces the coupling strength needed to explain reshuffling.

Introduction

Since their introduction nearly two decades ago, optogenetic methods have revolutionized neuroscience (Kim et al., 2017). In particular, optical stimulation of specific cell types, in conjunction with electrophysiological or optical recordings of neuronal activity, can be used to probe network operating regime and test predictions from theoretical models. Successful examples include testing the 'paradoxical effect'

in mouse cortex in inhibition-stabilized networks (Ozeki et al., 2009; Sanzeni et al., 2020; Tsodyks and Markram, 1997), and testing predictions of attractor models in both mouse cortex (Inagaki et al., 2019) and the fly central complex (Kim et al., 2017). Optogenetics, in conjunction with theoretical modeling, can also provide information about the structure of synaptic connectivity (*e.g.* Palmigiano et al., 2020). For instance, activation of single neurons in V1 suppresses other neurons, with a magnitude depending on the distance between their preferred orientations (Chettih and Harvey, 2019), consistent with a network with strong orientation specificity of excitatory-inhibitory (E-I) connectivity (Ko et al., 2011; Sadeh and Clopath, 2020).

To probe the operating regime of sensory cortex (Ahmadian and Miller, 2021) and the mechanisms that shape its responses to its inputs, we analyze electrophysiological recordings from both mouse (Histed, 2018) and monkey (Nassi et al., 2015) primary visual cortex (V1) during presentation of visual stimuli of various contrasts and/or optogenetic stimulation of E neurons. The optogenetic stimulation leads to a broad distribution of firing rate changes in single neurons, with some neurons being strongly excited, while others are strongly suppressed. The typical firing rate changes in single neurons are much larger than the changes in mean response of the population, which are only significant for lower visual contrasts in mice. In fact, the entire distribution of firing rates is largely unaffected by optogenetic stimulation (significant changes seen only at low visual contrast in mice), even though the variance of firing rate changes is comparable to the variance of rates prior to optogenetic stimulation. Thus, optogenetic stimulation of E neurons leads to a reshuffling of firing rates across the network, leaving invariant the whole distribution. We find a continuum from mice at low contrast to mice at high contrast to monkeys at low contrast to monkeys at high contrast, along which firing rates increase and the optogenetically-induced changes in the mean and distribution of firing rates shrink to zero.

We explore the mechanisms of such a reshuffling using theoretical analysis and modeling. We first examine randomly connected networks of E and I neurons. We find that such networks reproduce all the above mentioned features, provided the strengths of recurrent coupling and external input are strong enough, the network is inhibition dominated, and opsin-induced currents in E neurons are sufficiently heterogeneous and weaker in the mean than visually-induced currents. The continuity from mice to monkeys can be understood if mice are in a more weakly-coupled regime than monkeys, which can arise from weaker recurrent weights and/or weaker external inputs. We also investigate structured network models with connectivity dependent on spatial distance and orientation tuning, which can incorporate the difference between tuned visual and untuned optogenetic input. We find that reshuffling can arise in structured networks with weaker effective coupling and stronger optogenetic input than in the unstructured randomly connected model. These results provide new insights into the operating regime of sensory cortex and how this may differ across different mammalian species.

Results

Heterogeneous optogenetic modulation of visual responses in mice and monkeys

We analyzed recordings obtained in V1 of mice (data from Histed, 2018) and monkeys (data from Nassi et al., 2015). These experiments shared a number of common features (see *Methods* for details): visual stimuli were gratings of different contrasts ($[0-90]\%$ in mice and $[0-99]\%$ in monkeys); virally expressed opsins were used to optogenetically stimulate pyramidal cells. Both datasets consist of neural spiking responses recorded extracellularly in awake animals to combinations of visual and optogenetic stimuli. There are also a few differences between the recordings in (Histed, 2018) and (Nassi et al., 2015). In monkeys, visual stimuli were drifting sinusoidal gratings centered within the receptive field and matched

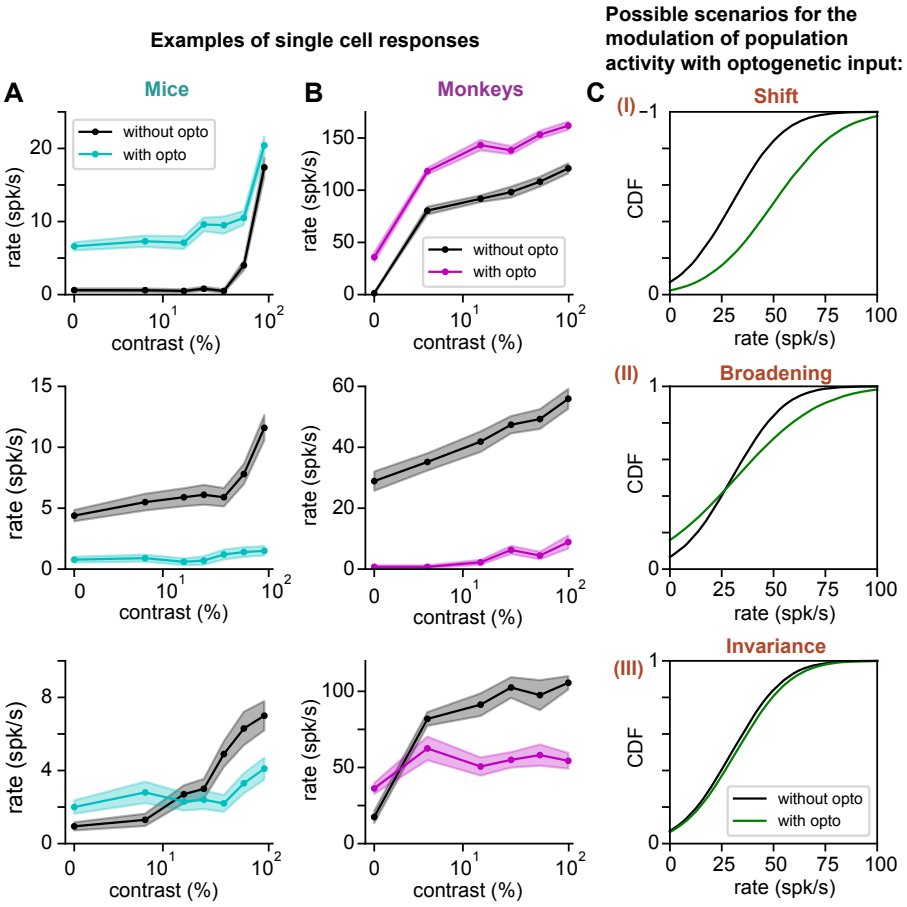

Figure 1: **Optogenetic modulation of visual responses in mice and monkeys: single neuron examples and potential scenarios at the population level.** (**A**) Examples of single neuron responses to visual stimuli in mice, without (black) and with (colored) optogenetic stimulation of E neurons (here and in the following figure, lines represent means, shaded regions are mean±sem). (**B**) As in **A** but in monkeys. In both species, some neurons strongly increase their rates at all contrasts (top row); others are strongly suppressed (center row); and others exhibit rate increase at low contrast, but suppression at high contrast (bottom row). Note that the same color code is used in all the following figures. (**C**) Possible scenarios of how optogenetic stimulation of E neurons might modify the distribution of firing rates in the network (see text).

to each recorded unit's preferred size, orientation, and spatial and temporal frequencies. In mice, on the other hand, multiple neurons were recorded simultaneously and there was no systematic relationship between the visual stimulus and preferences of the recorded neurons.

Both in mice and monkeys, as previously observed (Histed, 2018; Nassi et al., 2015), single neuron responses to visual stimuli were strongly modulated by optogenetic stimuli, in a highly heterogeneous fashion (Fig. 1A-B). Specifically, responses to visual stimuli were typically either enhanced by the laser at all contrasts, suppressed by the laser at all contrasts, or enhanced at low contrast and suppressed at high contrast (Fig. 1A-B); the fourth possibility, suppressed at low contrast and enhanced at high contrast, was not observed. These single neuron responses might underlie different scenarios at the population level (Fig. 1C). (i) Optogenetic stimulation could significantly increase the mean activity in the network, while broadening the distribution of single neuron rates. This scenario is what one would expect naively in experiments in which optogenetic stimulation targets E neurons. (ii) Optogenetic stimulation could broaden the distribution of rates, but a large number of suppressed cells could lead to a weak mean increase of the population activity. (iii) Optogenetic stimulation could affect only weakly the distribution of rates. Which of these scenarios better represents the effect of optogenetic stimulation on population activity? Do these effects change with contrast and species?

Optogenetic stimulation reshuffles visual responses in mice and monkeys

To answer the above mentioned questions, we analyzed the statistics of population responses to visual and optogenetic stimuli (Fig. 2). In both mice and monkeys, mean population activity increased monotonically with the contrast of visual stimuli, but firing rates in mice were significantly lower than in monkeys. The mean visual response in mice ranged from 6 ± 0.9 spk/s (mean±sem) at 0% contrast to 16.7 ± 2.0 spk/s at high contrast, while in monkeys it ranged from 20.2 ± 2.0 spk/s to 72.7 ± 6.7 spk/s (Fig. 2A). In both species, visual stimuli of fixed contrast generated a broad distribution of rates (Fig. 2B), and increasing contrast significantly shifted distributions toward higher rates (Fig. 2C–D). Surprisingly, optogenetic stimulation had much weaker effects at the population level (Fig. 2A–D). In mice, where all cells were recorded regardless of their feature preference, the mean and standard deviation of the population activity were significantly modulated by the optogenetic stimulation only for contrasts lower than 55% (Welch's t–test, p–value<0.05) and 90%(f–test, p–value<0.05), respectively. In monkeys, where cells were only recorded with stimuli matching their feature preferences, the mean population activity was not significantly modulated at any contrast (Welch's t–test, p–value>0.05) while the modulation of the standard deviation was significant only at contrasts lower than 12% (f–test, p–value<0.05). Even more strikingly, optogenetic stimuli weakly affected the whole distribution of rates in the network (Fig. 2D). In mice, the distribution of rates was significantly affected only at low contrast ($<$ 20%, KS test, p–value<0.05), for which optogenetic stimuli induced a weak but statistically significant shift of the distribution toward higher rates. In monkeys, the distribution of rates was not significantly affected by optogenetic stimulation at any contrast tested (KS test, p–value>0.05).

The weak optogenetic modulation observed at the population level was not due to small changes in single neuron firing. In fact, both in mice and monkeys, distributions of optogenetic modulations of single neuron firing rates were broad, with standard deviation comparable to or larger than the standard deviation of the rates without the optogenetic stimulus (Fig. 2E–F). Furthermore, unlike what was found in response to visual stimuli (Fig. 2C–D), the mean changes in rate due to optogenetic stimulation were characterized by a mean that was considerably smaller than the standard deviation of the rate changes (Fig. 2A, E). Specifically, combining all contrasts, changes in rates due to optogenetic input in mice and monkeys were 4.2 ± 8.9 spk/s (mean±std) and -0.7 ± 43 spk/s (mean±std), respectively. The distribution

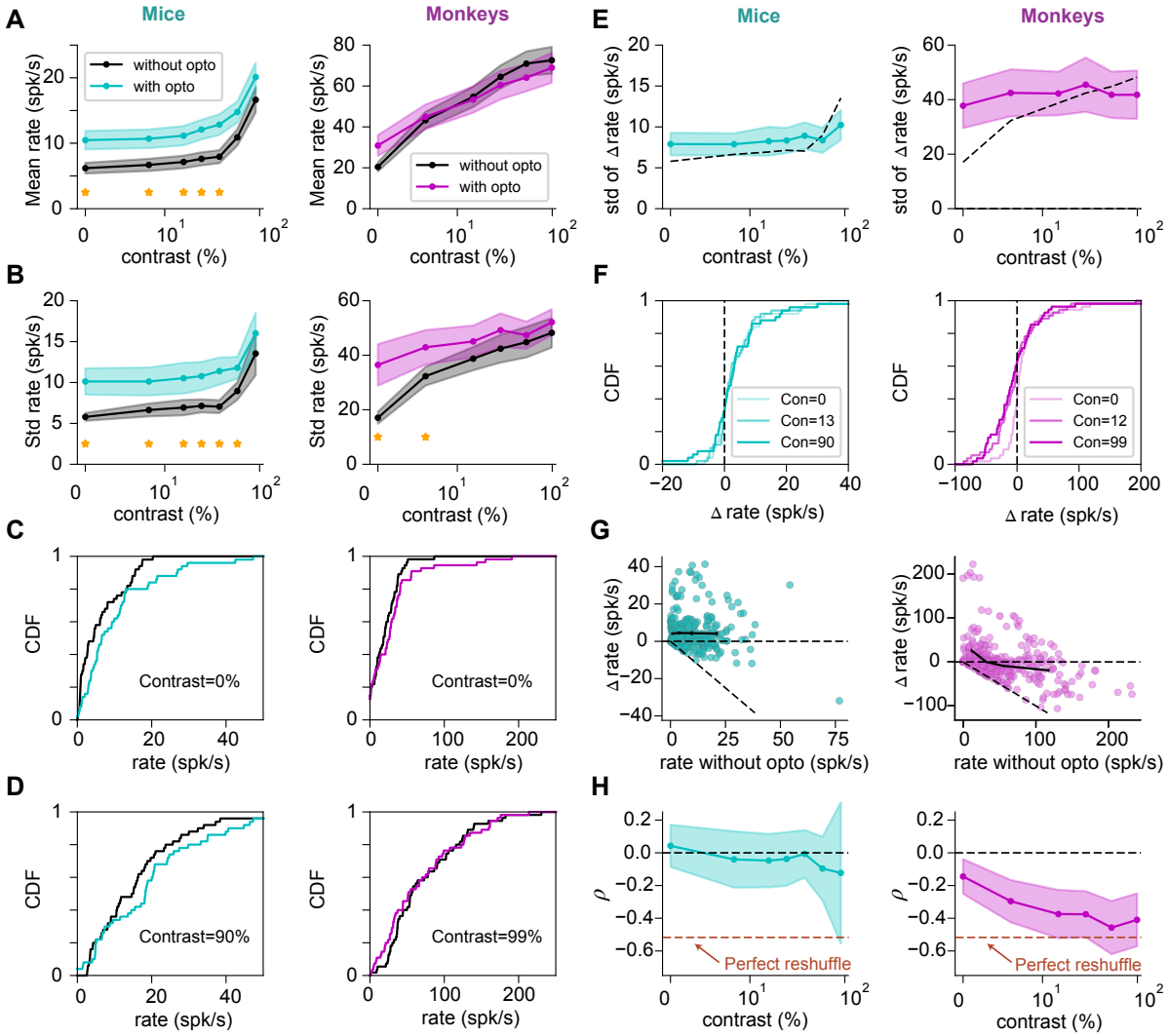

Figure 2: **Effects of optogenetic stimulation on population statistics in mice and monkeys.** (**A**) Population mean response to visual and optogenetic stimuli. Optogenetic modulation caused no significant change in mean rates except in mice at low contrast ($< 55\%$, Welch's t–test, p–value<0.05). (**B**) Standard deviation of the distribution of firing rates across the population. In both species, the standard deviation of firing rates is increased significantly by light only at low contrast (in mice, $< 90\%$; In monkeys, $< 12\%$; f–test, p–value<0.05). (**C-D**) Cumulative distribution function (CDF) of firing rates at low (**C**) and high (**D**) contrasts of visual stimuli, with (colored) and without (black) optogenetic stimuli. Optogenetic modulation was only significant for mice at low contrast (6% and 12%, KS test, p–value<0.05). (**E**) Standard deviation of the change in firing induced by optogenetic stimulation as a function of contrast of the visual stimulus (colored lines). In both species, the standard deviation of the change in firing rate was larger than the mean and of the same order of the standard deviation of rates without the laser (dashed black lines, same as black lines in B). (**F**) Cumulative distribution function of changes in firing rate (Δr) induced by optogenetic stimulation at three example contrasts (indicated by transparency levels). In both species, and at all contrasts, optogenetic stimuli generated a broad distribution of firing rate changes, with a large fraction of suppressed cells. (**G**) Change in firing rate induced by optogenetic stimulation as a function of baseline activity (i.e. activity without optogenetic stimuli). Each cell is represented once for every value of contrast of the visual stimuli. Continuous black lines represent mean Δr in quartiles of the distribution of rates without optogenetic stimuli. (**H**) Normalized covariance of baseline rates and change in rates induced by optogenetic stimuli ($\rho = \mathrm{Cov}(r, \Delta r)/\mathrm{Var}(\Delta r)$). This ratio decreases with contrast in both species and, in monkeys, it approaches the value -0.5 which, as described in the text, is required to obtain perfect invariance of the distribution of rates.

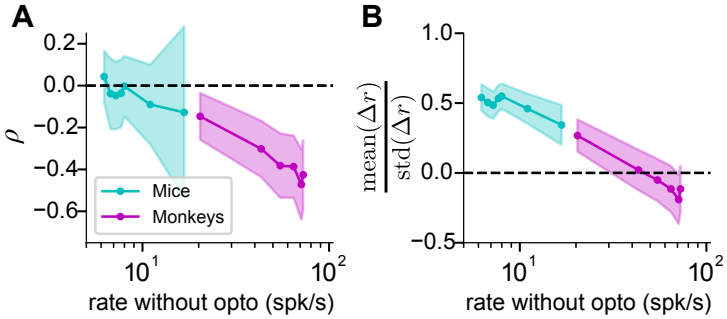

Figure 3: **Optogenetic responses in mice and monkeys as a function of average visual response.** (**A**) Normalized covariance (Eq. (1)) as a function of average visual response (different points correspond to different contrasts). (**B**) Ratio of mean over standard deviation of the change in rate, as a function of average visual response. Data in both animals seem to lie on the same underlying curves, with data in mice/monkeys covering the low/high rate regions, respectively.

of optogenetic responses was contrast invariant in mice (KS test, p-value>0.05). In monkeys, on the other hand, optogenetic responses observed at zero contrast followed a distribution that was significantly different (KS test, p-value<0.05) from those at higher contrasts, which were not significantly different from one another. Furthermore, a significant fraction of cells were suppressed (e.g. at high contrast, 16/55 in mice and 37/50 in monkeys, respectively).

In sum, the response statistics characterized here suggest that optogenetic stimuli *reshuffled* visual responses in mice and monkeys: they strongly modulated single neuron responses, but overall led to a weak modulation of population activity.

How can optogenetic stimulation leave the distribution of rates unchanged, while at the same time leading to large changes in firing rates of individual neurons (Fig. 2B)? If optogenetic responses were independent of baseline firing, as one might expect intuitively, optogenetic stimuli would significantly broaden the distribution of visual responses. The rate variance with ($\sigma^2_{r+\Delta r}$) or without (σ^2_r) optogenetic stimuli are related by $\sigma^2_{r+\Delta r} = \sigma^2_r + \sigma^2_{\Delta r} + 2\mathrm{Cov}\,(r, \Delta r)$ (where "Cov" is covariance). Thus, independence of r and Δr would imply the rate variance is increased by $\sigma^2_{\Delta r}$, which as we have seen (Fig. 2E) is comparable in size to σ^2_r. If the distribution of rates is not modified by optogenetic stimuli, then $\sigma^2_{r+\Delta r} = \sigma^2_r$, which requires a negative covariance with value $\mathrm{Cov}\,(r, \Delta r) = -0.5\sigma^2_{\Delta r}$.

The relationship between change in rate induced by laser Δr and rate without laser r is illustrated in Fig. 2G for both mice and monkeys. In mice, the two are weakly anticorrelated (combining all contrasts, the Pearson correlation coefficient was -0.06), consistent with a significant increase in the width of the rate distribution with optogenetic stimulation, while in monkeys, there is a strong anticorrelation between Δr and r (combining all contrasts, the Pearson correlation coefficient was -0.38) – neurons with higher (lower) baseline rates were more likely to be suppressed (excited) by optogenetic stimuli. To quantify the degree to which this anticorrelation leaves the distribution of firing rates invariant, we define a normalized covariance

$$\rho = \frac{\mathrm{Cov}\,(r, \Delta r)}{\sigma^2_{\Delta r}}; \tag{1}$$

so that ρ is equal to -0.5 when optogenetic stimuli leave invariant the distribution of rates. ρ is plotted as a function of visual contrast in Fig. 2H. In mice, the normalized covariance is never significantly different from zero, though there is a trend toward a negative covariance at the highest contrasts. In monkeys, it is significantly negative at all contrasts and close to -0.5 at high contrast, consistent with invariance of the distribution of rates and consequent reshuffling of rates by optogenetic stimulation.

The analysis shown so far shows common trends in mice and monkeys but also significant quantitative differences. We wondered whether these quantitative differences might at least partly be explained by the differences between the firing rates observed in the two animals. Therefore, we plotted dimensionless variables characterizing the effect of optogenetic stimuli on the statistics of network activity in the pres-

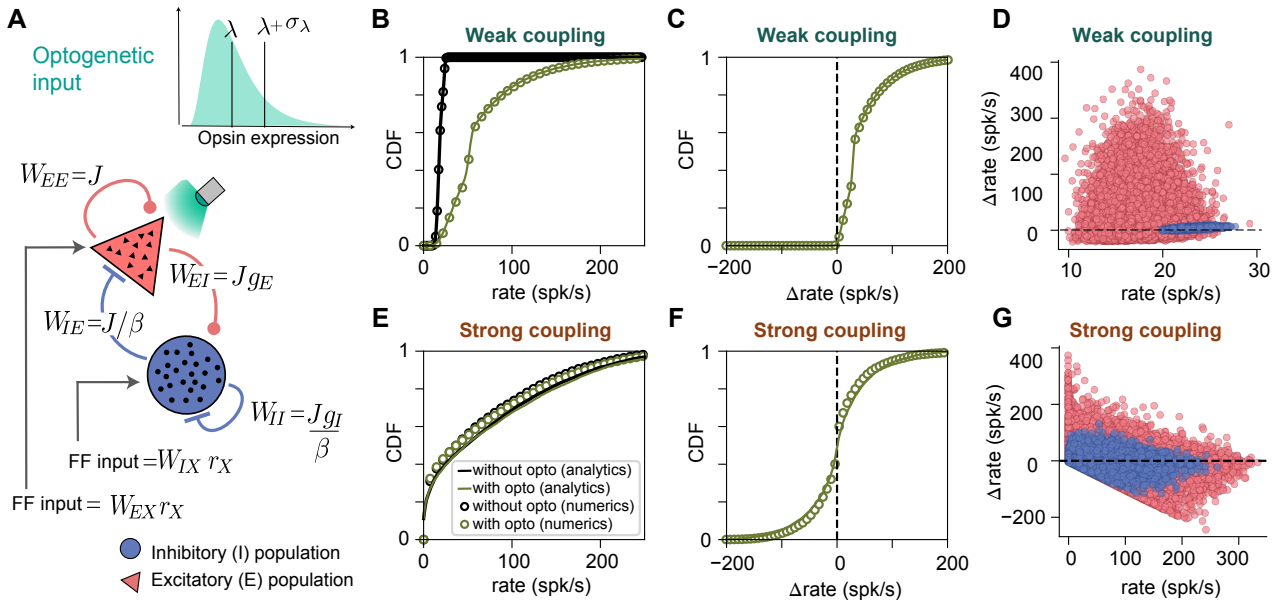

Figure 4: **Strong coupling leads to large fractions of suppressed cells, negative covariance and activity reshuffling in randomly connected E-I network models.** (**A**) Schematic of unstructured network model, with randomly connected E and I rate units. The network is driven by an external (visual) input, targeting both cell types, and by an optogenetic input targeting only E neurons. Visual inputs follow a Gaussian distribution, while opsin expression in E cells is distributed according to a lognormal distribution. (**B-D**) Responses to optogenetic stimulation of E cells in a weakly coupled network model. In all panels, dots represent numerical simulations while lines are mean field theory predictions. For weak coupling, optogenetic stimulation strongly affects the distribution of firing rates (**B**), with only a weak fraction of suppressed cells (**C**). Changes in firing rates are positively correlated with baseline activity (**D**, Pearson correlation coefficients are 0.1 and 0.6 for E and I cells, respectively). (**E-G**) As in panels **B-D**, but in a strongly coupled network model. For strong enough coupling, optogenetic stimulation does not significantly affect the distribution of rates (**E**). Firing rate changes are broadly distributed, with a mean around zero and a large fraction of suppressed cells (**F**). Firing rate changes are anticorrelated with baseline activity (**G**, Pearson correlation coefficients are -0.3 and -0.2 for E and I cells, respectively). Parameters: $J = 0.01$mV and 0.3mV for weak and strong coupling, respectively; for other parameters see table 1, *Methods*. In all panels, only cells whose response was higher than 1 spk/s in at least one of the two conditions are shown (100% for weak coupling, 43% for strong coupling.) Here and in all modeling figures, unless otherwise specified, results are shown for all cells, both E and I.

ence of a given visual stimulus as a function of the average firing rate induced by that visual stimulus. Remarkably, Figs. 3A–B show that curves traced by monkey data seem to be the continuation of those for mouse data, with mice occupying the low rate region and monkeys the high rate region of underlying common curves. In particular, in both species, the mean of Δr relative to its standard deviation decreases as a function of contrast until it becomes negligible at high firing rate. The normalized covariance ρ also decreases with firing rate, becoming closer to -0.5 at high rate.

Network reshuffling emerges in strongly coupled network models

To understand the mechanisms underlying reshuffling of visual responses by optogenetic stimuli, as well as the observed differences between mice and monkeys, we investigated the response properties of a mathematical model of cortical circuits. We analyzed a network of randomly connected E and I rate neurons (Harish and Hansel, 2015; Kadmon and Sompolinsky, 2015; Sompolinsky et al., 1988). In this unstructured model, the single neuron response function (firing rate vs input current) was taken to be

the f-I curve of leaky integrate-and-fire neurons driven by white noise (Amit and Brunel, 1997; Amit and Tsodyks, 1991; Siegert, 1951), which is supralinear at lower mean input and becomes sublinear at high input due to saturation, but the results presented here do not depend on this specific choice (see supplementary information, Fig. S1). Visual stimuli are described by external inputs to E and I units, whose strength is drawn from a Gaussian distribution (see details in *Methods*). Optogenetic inputs are described by an additional input to E cells, whose strength is drawn from a log-normal distribution.

Through numerical simulations and analytical calculations, we found that this network model reproduces the main features of the experimental data provided that recurrent connectivity and visual input are strong enough, and the distribution of opsin expression is heterogeneous enough. In particular, we first focus on the dependence of optogenetic responses on the network's effective synaptic coupling strength (referred to below as coupling strength), which is defined as the synaptic strength times the sensitivity of the postsynaptic cell to synaptic input (Ahmadian and Miller, 2021). This sensitivity is the cell's gain, i.e. the slope of its input/output function. In the supralinear portion of the input/output function, this increases with network activation and thus with the strength of visual input.

Example network simulations for weak and strong coupling are shown in Fig. 4B-G. In the absence of optogenetic stimulation, and consistent with previous theoretical studies (Amit and Brunel, 1997; Roxin et al., 2011; Sanzeni et al., 2022; van Vreeswijk and Sompolinsky, 1996), the distribution of rates is narrow for weak coupling and broad for strong coupling. Responses to optogenetic stimulation strongly depend on the strength of coupling. For weak coupling, optogenetic stimulation generates a significant shift of the distribution of rates (Fig. 4B). This leads to a broad distribution of optogenetic responses, characterized by a positive mean and a lack of significantly suppressed cells (Fig. 4C). Moreover, single neuron responses to optogenetic stimulation are positively correlated with baseline activity (Fig. 4D, Pearson correlation coefficients are 0.1 and 0.6 for E and I cells, respectively). This positive correlation follows from the supralinear shape of the transfer function at low rates. This effect is more pronounced in I cells because, with respect to E cells, they have a shorter membrane time constant and hence a larger supralinear region (Sanzeni et al., 2020). The response structure observed in weakly coupled networks matches naive expectations but strongly differs from experimental observations. When coupling is strong, on the other hand, optogenetic stimulation does not significantly modify the distribution of rates (Fig. 4E). It produces a broad distribution of rate changes, with a mean that is close to zero and a large fraction of suppressed cells (Fig. 4F). In this regime, single neuron responses to optogenetic stimulation are negatively correlated with baseline activity (Fig. 4G, Pearson correlation coefficient are -0.3 and -0.2 for E and I cells, respectively). These response properties are strikingly similar to experimental observations.

We now turn to a deeper investigation of the circuit mechanisms underlying these different aspects of the response of the network to optogenetic stimulation, using mean-field analysis.

Mechanisms shaping the response to optogenetic inputs in network models

We performed numerical and mean field analyses of the network (see details in *Methods*) to elucidate the circuit mechanisms underlying the simulation results highlighted in Fig. 4. These analyses show that, with strong recurrent connectivity, optogenetic input evokes only a weak change in average firing rate, provided the network is inhibition-dominated (Fig. 5). In particular, numerical simulations and mean field theory show that the average change in firing rate due to optogenetic stimulation decreases when (Fig. 5A-C, first row) the firing rate r_X of the neurons providing external inputs to the network increases, synaptic efficacies J become stronger, the strength of inhibition onto E neurons g_E becomes stronger, and/or the strength of inhibition onto I neurons g_I becomes weaker. These results can be understood analytically from the implicit equation defining the mean rates (Eq. (12), *Methods*). We linearize the equation about

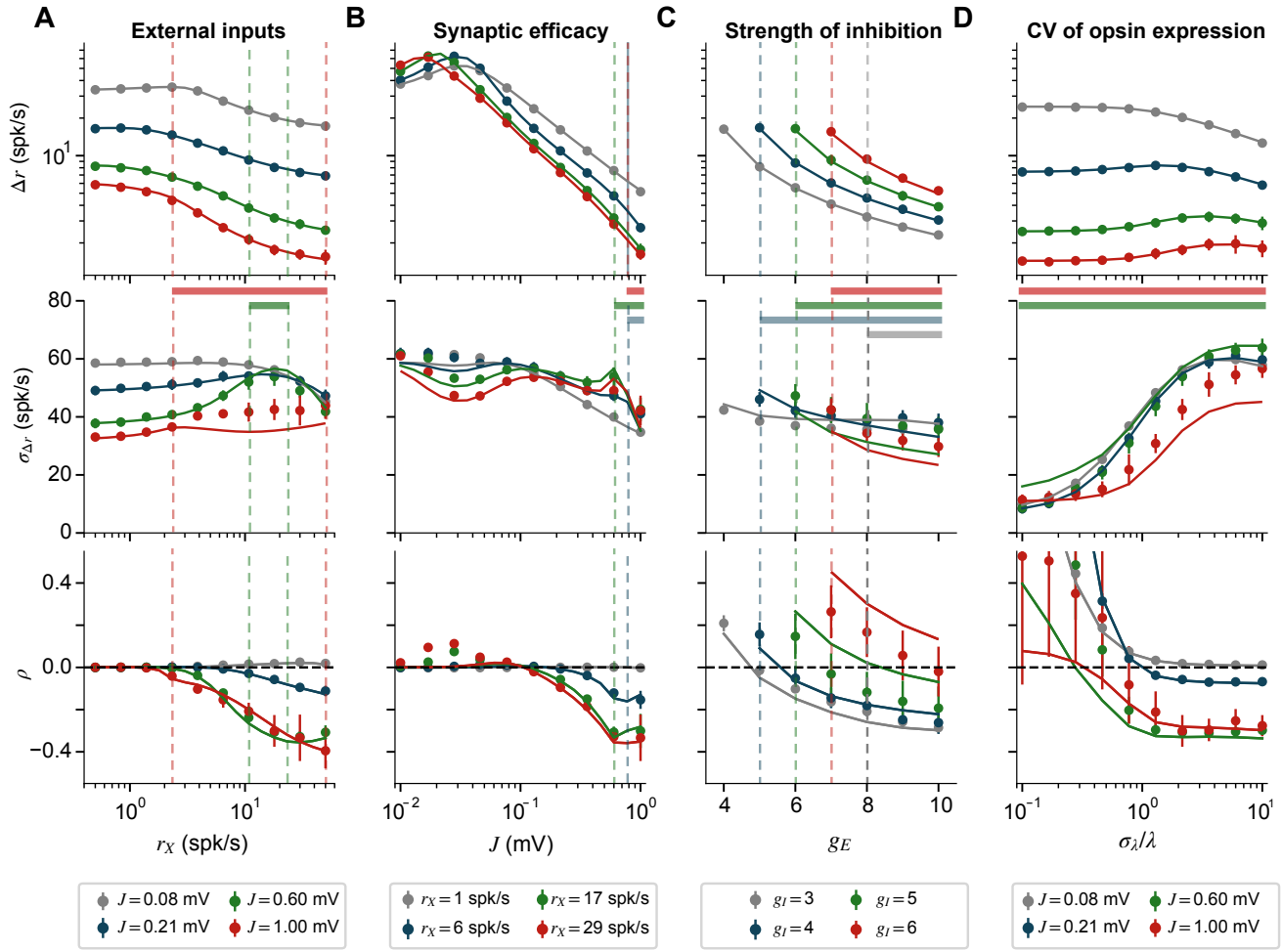

Figure 5: **Dependence of the statistics of network response to optogenetic input on network parameters and optogenetic input statistics.** (**A**) Statistics of network response to optogenetic stimuli as a function of external input r_X, for different values of recurrent synaptic efficacies (J, colors; J scales all weights, ({E,I,X}-to-{E,I}) where X is external input. Panels show (top to bottom): mean change in firing rate Δr; standard deviation of the change in firing rates $\sigma_{\Delta r}$; normalized covariance ρ between baseline rate and change in rates (Eq. 1). Colored bars under the first row indicate, for J indicated by color, values of r_x leading to chaotic dynamics. Filled circles with errorbars are mean±sem computed over 20 realizations of the network model. Continuous lines are mean field theory predictions. Increasing r_x generates smaller mean rate change and more negative covariance (except for the smallest J), while only mildly affecting the standard deviation of rate changes. Note that negative covariances can appear both in chaotic and non-chaotic networks (red and green curves, respectively, in lower panel). (**B**) Same as in **A** but as a function of J, for different values of r_X (colors). Increasing J significantly decreases mean rate change and response covariance, while affecting only mildly the standard deviation of rate changes. (**C**) Same as in **A** but as a function of the strength of inhibition to E cells g_E, for different values of the strength of inhibition to I cells g_I. Stronger inhibition to E cells and weaker inhibition to I cells both lead to lower mean rate changes and more negative covariance of rate change with baseline rate, while only mildly affecting the standard deviation of rate changes. Networks with $g_E < g_I$ are not analyzed because they are either unstable or multistable (Sanzeni et al., 2020). (**D**) Same as in **A** but as a function of the coefficient of variation (CV) of the distribution of opsin expression σ_λ/λ, for different values of J. The CV weakly affects the mean rate changes, but strongly affects the standard deviation of the rate changes and the response covariance. In particular, heterogeneous opsin expression ($\sigma_\lambda/\lambda > 1$) is required to generate broad distributions of responses and negative covariance. Simulation parameters are: $\sigma_\lambda/\lambda = 2$ in panels **A-C**; $g_E, g_I = 8, 3$ in panels **A,B,D**; $r_X = 10$spk/s and 18spk/s in panel **C** and **D**; $J = 0.6$mV in panel **C**.

the fixed point, thus assuming optogenetically induced rate changes in E and I firing rates, Δr_E and Δr_I, are small. In the limit of small mean optogenetic input λ, and neglecting the contributions coming from quenched disorder in the connectivity and in the opsin expression, we can solve the equations to find

$$\Delta r_E = \frac{g_I \lambda}{K J \tau_E \left(g_E - g_I\right)}, \quad \Delta r_I = \frac{\lambda}{\gamma K J \tau_E \left(g_E - g_I\right)} \tag{2}$$

where K and γK are the average number of E and I recurrent connections per neuron, while τ_A is the membrane time constant of population $A \in [E, I]$. Eq. (2) captures qualitatively the dependency on J and $g_{E,I}$ observed in simulations (Fig. 5A-C, first row).

Strong recurrent input can largely cancel the mean optogenetic input, but cannot cancel the cell-to-cell fluctuations in input induced by opsin expression heterogeneity (Renart et al., 2010; van Vreeswijk and Sompolinsky, 1998). This produces high response heterogeneity (Fig. 5, second row) that increases with opsin expression variability (σ_λ, Fig. 5D, second row), and only weakly depends on coupling strength (Fig. 5A–C, second row). Heterogeneous opsin expression and strong coupling also lead to a large number of suppressed cells (Figs. 4.)

The combination of strong coupling and a broad distribution of opsin expression naturally leads to a normalized covariance ρ closer to -0.5 (Fig. 5, third row) and to network reshuffling (Fig. 4E–G). The fact that ρ becomes negative is due to an interplay between single neuron nonlinearities and the strong inhibitory feedback canceling the mean network response to optogenetic input. Strong couplings leads to a broad distribution of baseline rates and therefore a large fraction of cells with very low baseline rates. Upon heterogeneous optogenetic stimulation, the non-negative nature of firing rates biases the response of cells with very low baseline rates toward excitation (see Fig. 4G). To compensate for this bias, and generate a close to zero mean change in rate, the circuit produces, on average, suppression in neurons with higher baseline activity (see Fig. 4G). This leads to a negative correlation between baseline activity and response to optogenetic input, i.e. a negative ρ. This negative correlation leads in turn to a negative correlation between baseline synaptic inputs and changes in inputs due to optogenetic drive (see Eq. (10).

Previous theoretical works have shown that the dynamics of rate networks becomes chaotic when recurrent connections are strong (Harish and Hansel, 2015; Kadmon and Sompolinsky, 2015; Sompolinsky et al., 1988). Consistent with these results, we observed the emergence of chaotic dynamics in our simulations as J and/or r_X increase (Fig. 5, lines under the first row). Note that in our simulations chaos is heterogeneous across cells, with a broad distribution of temporally-averaged firing rates that reflects the quenched disorder in the inputs and synaptic connectivity. Since negative covariance and neuronal reshuffling emerge at large contrast, we wondered if the appearance of these features is related to chaotic dynamics. To answer this question, we have systematically compared the regions of parameter space in which negative covariance and chaotic dynamics appear (supplementary information Fig. S2) but found that they only partially overlap. These results show that negative covariance and chaotic dynamics are two distinct phenomena that both emerge in networks of strongly coupled neurons.

Randomly connected network models capture experimental recordings in mice and monkeys

Optogenetic responses in mice and monkeys have qualitatively similar features, which we summarized above with the term "network reshuffling", but also important quantitative differences (Figs. 1 and 2). Compared to monkeys, optogenetic responses in mice had a larger mean over standard deviation ratio and weaker normalized covariance. However, when these quantities are plotted as a function of mean baseline population rate (in the absence of optogenetic stimuli), these differences seem to largely arise

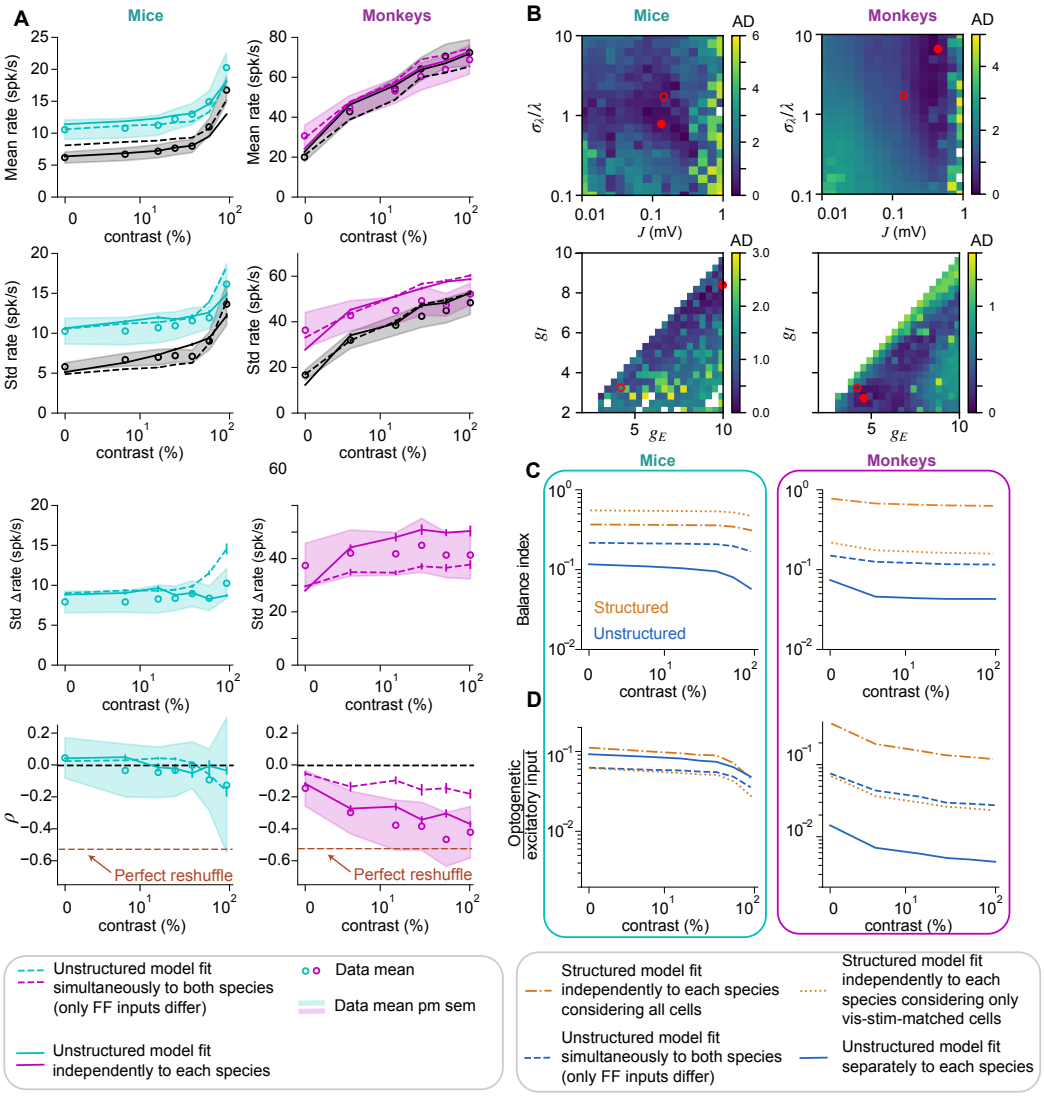

Figure 6: **Fitting mice and monkey data with the network model.** (**A**) Comparison between data (circles represent means, shaded regions are mean±sem) and network model with parameters that best fit data (dashed and continuous lines) in mice (left) and monkeys (right). Dashed lines represent best fits obtained with the same network model in mice and monkeys, but different external and optogenetic inputs. Continuous lines represent best fits obtained with different network models. Model plots are obtained from simulations, averaging over 20 realizations. Network models reproduce key features of experimental data, including: weak mean changes in firing rate; broad distribution of responses; and increasingly negative covariance with contrast of the visual stimulus. (**B**) Average discrepancy (AD, measured as $\log_{10}$ of the ratio between the actual and the minimum error) between best model predictions and experimental data for mice (left) and monkeys (right) obtained with models where two parameters are constrained to the values specified by both axis. Red circles indicate optimal parameters, used to generate panels **A** (open circles: same parameters in mice and monkeys; filled circles: different parameters in mice and monkeys). In both species, consistent with Fig. 5D, better fits are obtained for $\sigma_\lambda/\lambda > 1$, i.e. highly heterogeneous optogenetic expression, and for biologically realistic values of J, in the range of $[0.1 - 1]$mV. (**C**) Inferred values of the balance index (Ahmadian and Miller, 2021), i.e. the ratio between net input current (excitatory minus inhibitory) and the total excitatory input in response to visual stimuli. Fits in structured models (orange lines) are less strongly coupled that fits in the unstructured model. (**D**) Inferred values of the ratio between optogenetic input and the total visually-driven excitatory input. Both in mice and monkeys, the unstructured model (blue lines) requires a small mean optogenetic input, relative to visually-driven excitatory input, to capture experimental data. In structured models, data can be captured with a larger optogenetic to excitatory input ratio.

because of different ranges of baseline firing rates, possibly due to a difference in visual input strengths in the two species (Fig. 3A–B). Thus, the strength of external inputs might be the main factor explaining differences between mice and monkeys. To test this hypothesis, we investigated if a single unstructured network model, with only the strengths of external stimuli differing between the two species, can capture the responses observed in each.

We fitted our network model to recordings in mice and monkeys assuming common network parameters (see methods), but allowing for different external (r_X) and optogenetic (L) inputs. This model captures the main features of the recordings (Fig. 6A, dashed lines): weak mean optogenetic response; large standard deviation of optogenetic response; and increasingly negative normalized covariance with contrast of the visual stimulus. It also captures key differences between mice and monkeys: a lower visual response, and a larger ratio between mean and standard deviation of the optogenetic response in mice; and a more negative normalized covariance in monkeys. These results support the idea that the strength of external inputs is a key factor shaping differences between the two species.

At the same time, the model fails to capture quantitatively the mean optogenetic response and the amplitude of the negative normalized covariance in monkeys. These discrepancies suggest that differences in the network structure or the visual or optogenetic input distributions might be important to fully capture the observed response properties. To test this, we repeated the fitting procedure in unstructured models, allowing different connectivity and input parameters for mice and monkeys, and found that this model gives a better description of experimental recordings, including the strongly negative covariance seen in monkeys for high contrast (Fig. 6A, continuous lines; relative error decreases averaged over the quantities plotted in Fig. 6A were 22% in mice and 5% in monkeys). The best fit parameters are characterized by weaker overall coupling J, weaker inhibition (weaker overall $I \rightarrow$E strength, Jg_E, and weaker ratio of I→E to I→I strength, g_E/g_I), and less heterogeneous optogenetic expression (smaller σ_λ/λ) in mice than in monkeys.

Previous theoretical work (Ahmadian and Miller, 2021; Rubin et al., 2015; Sanzeni et al., 2020) has suggested that cortical circuits operate in a loosely balanced regime, characterized by the net input to cells (excitatory minus inhibitory) being comparable to its individual (excitatory or inhibitory) components. This can be quantified by the ratio of the two – net input over excitatory input – which is called the balance index (Ahmadian and Miller, 2021). In our models, using the parameters inferred from data, we found a smaller balance index in monkeys than in mice (Fig. 6C). Recordings in mice were consistent with the underlying network being loosely balanced at low and intermediate contrast (balance index greater than 0.1). For the model in which mice and monkeys were jointly fit, somewhat tighter balance was observed at high contrast in mice and for all contrasts in monkeys (Fig. 6C, dashed line), while monkeys showed very tight balance in the model in which they were independently fit (Figure 6C, filled line).

The unstructured model also requires the ratio of mean optogenetic input to visually evoked excitatory input to be small (~ 0.1), and to be very small ($\lesssim 0.01$) for the independent fit to monkeys (Figure 6D). The latter is related to tight balancing. The small balance index needed for the independent fit to monkeys indicates that the mean inhibitory recurrent input cancels most of the mean excitatory input (*e.g.* a balance index of 0.04, Figure 6C, means that all but 1/25 of the mean excitatory input is cancelled). In particular, both the visual and the optogenetic input are strongly cancelled, by similar factors. For the mean optogenetic response to be reduced by cancellation to near zero, while the mean visual response remains finite, the mean optogenetic input must be much smaller than the mean visual input. Thus, by both measures – the balance index, and the relative mean size of optogenetic vs. visual input – the nearly perfect reshuffling seen in monkeys at higher contrast (very small Δr and strongly negative normalized covariance ρ) appears to require tight balancing, which in turn means that coupling is very strong (Figure

5A,B; Ahmadian and Miller, 2021.

Structured connectivity leads to reshuffling with weaker coupling

The unstructured models studied to this point ignored the difference between tuned visual input, which specifically targets a set of cells with similar feature and spatial selectivity, and untuned optogenetic input, which targets cells in a given cortical region regardless of their selectivity. It similarly ignored the tuning of connectivity, which decreases with cortical distance (Angelucci et al., 2002; Rossi et al., 2020) and, at least in mice, with difference in feature preference (Cossell et al., 2015; Ko et al., 2011). We reasoned that a tuned stimulus would excite cells with matched tuning, which might preferentially excite one another and/or inhibit others; whereas an untuned stimulus would tend to excite and inhibit all cells equally, which could give a very small or even negative response. With such a mechanism, the network might not need such tight balance or small optogenetic input to explain reshuffling.

To test these ideas, we constructed alternative, structured models. We assumed that neuronal selectivity was determined by a single feature, preferred orientation. Connection strengths decayed as a product of Gaussian functions of the physical distance between neurons and of the difference in their preferred orientations. We used a spatially salt and pepper organization of preferred orientation in the mouse and a quasi-periodic map in the monkey, as reported experimentally (Hubel and Wiesel, 1977; Kaschube et al., 2010; Ohki et al., 2005, 2006). Visual input was also a product of Gaussian functions of the distance between receptive field center and stimulus location and of the difference between preferred orientation and stimulus orientation.

We fit these structured models to the data and found that the same quality of fit could be obtained as with the unstructured models fit independently to mice and monkeys (see Fig. S3). However, this fit for monkeys is obtained with a much larger balance index than the independently fit unstructured models: ~ 5 times larger if the index is computed only over cells whose preferred orientation matches that of the stimulus, (*visual-stimulus-matched* cells; Fig. 6C, orange dotted lines, see also Fig. S4) and even larger if computed over all cells (orange dash-dotted lines). Similarly, the structured models can match the monkey's nearly perfect reshuffling with a much larger ratio of mean optogenetic to visual input than the unstructured models (~ 10 times larger for visual-stimulus-matched cells, even larger for all cells; Fig. 6D, see also Fig. S4). As in the unstructured model (Fig. 4), the structured model requires sufficiently strong coupling for the mean optogenetic response to be suppressed (Fig. S5), although less strong than is required in the unstructured model (Fig. S4). In sum, the structured models can produce strong reshuffling with significantly weaker balance and stronger optogenetic input than the unstructured models.

Discussion

Summary of the results

We analyzed electrophysiological recordings of mice and monkey V1 while the animals viewed visual stimuli, with or without concurrent optogenetic stimulation of E neurons. In both species, and for all contrasts of visual stimuli, optogenetic stimuli generate a broad distribution of responses, whose mean is considerably smaller than the standard deviation. This large heterogeneity of optogenetic responses results, in particular, in large fractions of suppressed cells. The main quantitative differences between the two species are that, in monkeys, the firing-rate responses to visual and optogenetic stimulation, and the fraction of suppressed cells, are larger, while the mean optogenetic response is smaller. Despite

this strong optogenetic modulation of the activity of single neurons, the distribution of rates across the population is unchanged by optogenetic stimulation in monkeys and, for high visual contrast, in mice. This "reshuffling" of rates is specific to optogenetic stimuli; visual stimuli shift the distribution toward higher rates. Intriguingly, statistics of optogenetic responses across the two species form a single, continuous function of baseline (visually-induced) population activity, with mouse and monkey responses corresponding to lower and higher baseline activities, respectively.

Why does optogenetic stimulation drive large single-cell responses but little or no mean response? To address this, we first studied an unstructured, randomly connected network model of E and I neurons. We found that the "balancing" that occurs in sufficiently strongly coupled E/I networks (Ahmadian and Miller, 2021; van Vreeswijk and Sompolinsky, 1998) – the dynamic cancellation of external input by recurrent input – can produce this effect. Balancing suppresses the mean input, but not the variation in input across cells. Accordingly, when the heterogeneity of optogenetic input (*i.e.* of opsin expression) is comparable to or larger than the mean optogenetic input, the mean response is suppressed while individual cells show large responses. In this model, optogenetic input is distinguished from visual input in (1) targeting E, rather than both E and I, cells, (2) having a much smaller mean, and (3) having larger variability relative to mean. The small relative size of mean optogenetic input is a possible limitation of the model (further discussed below), and is required to have a substantial mean visual response with nearly zero mean optogenetic response, despite both being suppressed by balancing. Finally, for strong reshuffling – invariant firing rate distributions despite strong, heterogeneous single-cell responses – there must be a strong negative covariance between optogenetic and baseline responses. This arises because neurons with low baseline rates are biased by rectification to positive optogenetic responses, and the circuit's dynamic suppression of the mean response therefore causes suppression of neurons with higher baseline activity. With appropriate adjustment of parameters, this simple model could be well fit to the mouse and the monkey data. A second possible limitation of this model is that, to achieve strong reshuffling, the network balance of excitation and inhibition must be tighter than has been estimated from experiments (Ahmadian and Miller, 2021), also discussed below.

We then incorporated the fact that the visual input is tuned, *e.g.* it targets cells with preferred orientation near the stimulus orientation, while the optogenetic input targets all E cells. We reasoned that a tuned input might lead to stronger excitation and/or weaker inhibition than an untuned input (*e.g.* as in the model circuit of Troyer et al., 1998), allowing suppression of the mean response under less stringent conditions. Studying such a model with structured connectivity (feature- and distance-dependent), we found that it could achieve equally good fits to the data as the unstructured model, but with looser balancing in the reshuffled cells, and with larger optogenetic input relative to visual.

Theoretical implications: Coupling strength in cortex

Balanced network models (Ahmadian and Miller, 2021; Ahmadian et al., 2013; Amit and Brunel, 1997; Brunel, 2000; Renart et al., 2010; Roxin et al., 2011; Rubin et al., 2015; Sanzeni et al., 2020, 2022; van Vreeswijk and Sompolinsky, 1996, 1998) have been successful in reproducing several ubiquitous features of cortical activity, including highly irregular firing (Compte et al., 2003; Softky and Koch, 1993), excitation/inhibition balance (Haider et al., 2006), and broad distributions of firing rates (Buzsáki and Mizuseki, 2014; Hromádka et al., 2008; O'Connor et al., 2010). These properties can arise either in networks that are tightly balanced or loosely balanced (Ahmadian and Miller, 2021), meaning that the net input after cancellation of excitation and inhibition is respectively much smaller than, or comparable in size to, the cancelling factors (balance index roughly $\lesssim 0.1$ or $\gtrsim 0.1$, respectively). Tighter balance corresponds to stronger coupling. Here, we showed that such networks can reproduce the surprising effect of strong

reshuffling. This provides further evidence that recurrent connectivity in V1 is powerful, and stabilized by strong inhibition. A key property of the loosely balanced regime, as exemplified by the supralinear stabilized network (SSN) model (Ahmadian and Miller, 2021; Ahmadian et al., 2013; Rubin et al., 2015), is that network behavior depends on the network's activation level. In particular, in the region over which the neuronal input/output function is supralinear, increasing activation yields increasingly strong coupling. In our models, this accounts for the strong dependence of the results of optogenetic stimulation on the state of the network prior to stimulation, which in the experiments described here is controlled by the contrast of the visual stimulus.

The unstructured model illustrates basic mechanisms, but requires relatively tight balance (balance index (b.i.) of best fits at high contrast: monkey, ~ 0.03; mouse, ~ 0.05). The structured model shows similar mechanisms can operate with looser balance (b.i. of reshuffled cells in best fits, high contrast: monkey or mouse, ~ 0.2), consistent with estimates from experimental data of loose balance (Ahmadian and Miller, 2021). Note, however, that those estimates were based on cat and mouse V1; it is possible that monkey V1 has tighter balance, *e.g.* due to its higher firing rates (although extracellular recording may overestimate firing rates, Barth and Poulet, 2012).

An interesting feature of our analysis is that reshuffling and chaos often coexist (Engelken et al., 2020; Harish and Hansel, 2015; Kadmon and Sompolinsky, 2015; Sompolinsky et al., 1988). This makes sense since both require strong coupling. We note however that our results do not provide evidence of chaos in cortical circuits, since non-chaotic networks reproduce all the main features in the data.

Mice vs monkeys

Our work provides to our knowledge the first systematic comparison of how optogenetic stimuli modulate visual responses in mice and monkeys. It shows that baseline firing plays a key role in determining the statistics of neural responses, and suggests that the different activity levels, due to different strengths of external input, could underlie differences observed in the two species. When key statistics were plotted as a function of baseline rate, curves traced by mice and monkey data seem to lie on the same underlying curve, with mice occupying the low rate region while monkeys occupy the high rate region. Consistent with this observation, fitting data from both species with a network model with the same parameters, but different input strengths, gives a reasonable description of the data, though the quality of the fit improves when network parameters are allowed to differ between the two species.

A recent electron microscopy study performed a detailed comparison of reconstructed neurons and their synaptic connections in mouse and primate V1 (Wildenberg et al., 2021). Surprisingly, synaptic connectivity in primates was found to be much more sparse than in mice, with a larger ratio of excitatory to inhibitory synapses in mice. The larger ratio of excitatory to inhibitory synapses in mice seems consistent with the weaker inhibition we found in mice. However, the weaker anatomical coupling in monkeys than in mice seems at first sight at odds with our conclusions that monkeys appear to be in a more strongly coupled regime than mice. It is possible that the physiological strengths of the connections are stronger in monkeys, sufficiently so that monkey connections are equal or larger in overall strength to those in mice. In addition, strong coupling can arise from stronger activation levels as well as from stronger synaptic connections, and the former are larger in monkeys. Furthermore, both the unstructured and the structured models reproduce well the data in broad regions of parameter space (see, in Fig. 6B, the broad blue regions which indicate reasonable fits in a broad range of values of J, which scales both synaptic strengths and input strengths).

Running strongly enhances visual responses in mouse V1 (Niell and Stryker, 2010). Recent experiments found that, in V1 of a species of monkey (the marmoset), running does not increase the mean

response to visual stimuli, but strongly modulates single-neuron responses, with a large fraction of cells suppressed (Liska et al., 2022). This is similar to the difference we have found in mice vs. monkeys in responses to optogenetic stimuli (although the mouse mean response increase for running is considerably stronger), suggesting possible common mechanisms: the weaker running modulation in monkeys than in mice could, at least in part, be a consequence of stronger coupling in monkeys.

Relations with previous experimental works and computational consequences

The response properties of cortical networks to optogenetic inputs we have found here might also be relevant to understanding responses of neurons observed in other conditions. Large numbers of suppressed cells in response to optogenetic activation of pyramidal cells have been reported in experiments using one-photon stimulation in mouse visual cortex (O'Rawe et al., 2022), ferret visual cortex (Wang et al., 2022), and targeted two-photon stimulation of small ensembles in mouse barrel cortex (Dalgleish et al., 2020). Weak changes in mean population activity together with large changes in individual neuronal firing rates also characterize the response of cortex to some modulatory inputs (Sherman and Guillery, 1998), *e.g.* in behavioral modulations of activity by head movements (Bouvier et al., 2020) and running (Liska et al., 2022). Similar changes are seen after learning in experiments involving brain-machine interfaces (Engelhard et al., 2019).

Some recent works have reported significant changes in mean firing rate with optogenetic stimulation of E cells (Andrei et al., 2019; Chen et al., 2022) in monkeys. Many possible factors might contribute to differences from our results, including strength, timing, specificity, and spatial spread of optogenetic stimulation, visual stimulus strengths, match or lack thereof between the stimulus and neuronal preferences, recording methods, and behavioral paradigms used. In particular, Andrei et al. (2019) used very weak stimulus contrast, for which both our data and our models indicate weaker suppression, and used orientation-specific optogenetic stimulation which, like tuned visual input, may be able to drive mean activity. They could roughly match visual stimulus orientation, but not other stimulus features, to neuronal preferences. They also analyzed only cells with significant laser response, while we included all recorded cells in our analysis. Chen et al. (2022) used wide-field imaging to record the pooled activity of a large population of neurons with diverse tuning properties. Surprisingly, they generally found no mean visual response given the optogenetic stimulation, which might suggest their mean optogenetic input was very large compared to their mean visual input, the opposite of what we have concluded about our monkey experiments. Both experiments were done in the context of behavioral (detection or discrimination) tasks. Training mice to detect optogenetic stimuli can dramatically increase network responses to those stimuli (Akitake et al., 2022); it is conceivable that behavioral training could alter network responses in important ways.

Limitations and future directions

For the best fit match to the monkey at high contrast, the unstructured model had a ratio of optogenetic to excitatory visual input of about 5×10^{-3}. The structured model relaxed this requirement, but its ratio for visual-stimulus-matched cells was still small, about 2×10^{-2}. Are these plausible? In layer 4 of mouse V1, peak visually-evoked excitatory current was estimated to be 60-150pA across cells (Li et al., 2013; Lien and Scanziani, 2013). In cultured neurons, optogenetically-evoked currents are 100's to 1000's of pA (Mattis et al., 2011), but light is severely attenuated with transmission through cortex, by 50% over about 40 microns (Yona et al., 2016). Thus, it is difficult to assay whether the model's requirement is met *in vivo* in cortical cells, but this could be a problem for the present scenarios.

Though the current structured models succeed in explaining the experimental results with looser balance and stronger relative optogenetic input, we lack a detailed understanding of the exact mechanisms by which network structure affects reshuffling. It also remains uncertain whether different mechanisms or requirements might be revealed by considering other forms of structure, such as neurons with different preferred spatial phases (*e.g.* Kayser and Miller, 2002; Troyer et al., 1998) or preferred directions (Rossi et al., 2020). Future work will examine these questions.

Conclusions

The over 70 million years of evolution separating mice and monkeys (Siepel, 2009) have produced substantial differences in the structure of their cerebral cortices, e.g. in the presence of orientation and ocular dominance columns in primates (Hubel and Wiesel, 1977) but their absence in rodents (Dräger, 1974; Ohki et al., 2005; Van Hooser et al., 2005) the size of their receptive fields (Van Den Bergh et al., 2010), and their connection statistics (Wildenberg et al., 2021). Our analysis of optogenetic modulation of visual response in V1 shows that, despite these differences, mice and monkeys may share a common remarkable resistance of their cortex to optogenetic inputs at the population level, in spite of large changes in single neuron firing rates. At the same time, we find a key difference may be that monkeys are in a more strongly coupled regime of higher firing rates, with correspondingly stronger resistance. These results might more generally yield insight into circuit mechanisms underlying differences in the effects of sensory or driving vs. modulatory inputs on cortical responses.

Methods

Experimental methods

Mice

A detailed description of the experimental methods is given in (Histed, 2018). In brief, neurophysiological data from Emx1-Cre animals ($n = 4$) were collected. Animals kept on a monitored water schedule were given small drops of water ($\sim$1 μl) every 60–120 s during recording to keep them awake and alert. The visual stimulus, a Gabor patch with spatial frequency 0.1 cycle/deg and sigma 12.5 deg, were presented for 115 ms [full width at half maximum (FWHM) intensity], and successive visual stimuli were presented every 1 s. Optogenetic light pulses were delivered on alternating sets of 10 stimulus presentations (light onset 500 ms before first stimulus; offset 500 ms after end of last stimulus; total light pulse duration 10.2 s). A 1-s delay was added after each set of 10 stimulus presentations. Extracellular probes were 32-site silicon electrodes (Neuronexus, probe model A4x8). Channelrhodpsin-2 (ChR2) was expressed in E neurons (as described in Histed and Maunsell, 2014) using viral (AAV-EF1a-DIO-ChR2-mCherry, serotype 2 or 8; http://openoptogenetics.org) injections into the Emx1-Cre (Gorski et al., 2002; #5628, Jackson Laboratory) line. Virus (0.25–1.0 μl) was injected into a cortical site whose retinotopic location was identified by imaging autofluorescence responses to small visual stimuli. Light powers used for optogenetic stimulation were 500 μW/mm2 on the first recording session; in later sessions, dural thickening was visible and changes in firing rate were smaller, so power was increased (maximum 3 mW/mm2) to give mean spontaneous rate increases of approximately $\sim$5 spk/s in that recording session. Optogenetic light spot diameter was 400–700 μm (FWHM) as measured by imaging the delivered light on the cortical surface.

Spike waveforms were sorted after the experiment using OfflineSorter (Plexon, Inc.). Single units (SU) were identified as waveform clusters that showed clear and stable separation from noise and other clusters, unimodal width distributions, and interspike interval histograms consistent with cortical neuron absolute

and relative refractory periods. To compute neurons' visual responses, we counted spikes over a 175-ms period beginning 25 ms after stimulus onset. Zero-contrast responses were computed from the 175 ms period immediately preceding stimulus onset.

Monkeys

A detailed description of the experimental methods is given in (Nassi et al., 2015). In brief, two adult male rhesus macaques (Macaca mulatta) were each implanted with a custom titanium head post and silicone-based artificial dura recording chamber over V1. We injected a VSVg-pseudotyped lentivirus carrying the C1V1-EYFP gene behind the 1.3 kb CaMKIIα promoter (lenti-CaMKIIα-C1V1E162T-ts-EYFP; titer 3 $\times$ 1010 TU/ml) into a single location in V1 in each of the two monkeys (monkeys A and M) while they were anesthetized and secured in a stereotactic frame. Injections of the same viral construct were made into V1 of one additional monkey in order to assess specificity of viral expression to E neurons. In a second, distant location in V1 of monkey A, we injected an adeno-associated virus carrying the eArch3.0-EYFP gene behind the 1.3 kb CamKIIα promoter (AAV5-CamKIIα-eArch3.0-ts-EYFP; titer 4 $\times$ 1012 VP/ml).

Monkeys were alert and head restrained during all experiments. Single- and multi-unit waveforms were recorded using thin glass-coated tungsten electrodes (Alpha Omega, Engineering Inc., Nazareth, Israel). After isolating a single-unit or multi-unit cluster, we first assessed sensitivity to optogenetic stimulation. We randomly interleaved different stimulation intensities. Stimulation on each trial was continuous and lasted for 200 ms. Each condition was repeated at least five times. For a subset of light-sensitive units, we proceeded to measure responses to simultaneous optogenetic stimulation and visual stimulus presentation. The presented visual stimuli consisted of circular patches of drifting sinusoidal gratings whose size was matched to each recorded unit's preferred size (median diameter = 0.63deg), of mean luminance matching the surround (42 cd-m2) at the optimal orientation and spatial and temporal frequencies. Each stimulus condition was presented at least five times. For simultaneous optogenetic stimulation and visual stimulus presentation, we varied the contrast of the presented gratings in log steps (0%, 6%, 12%, 25%, 50%, and 99%) and stimulated with four different intensities including zero. Visual contrast and stimulation intensity were randomly interleaved. Optogenetic stimulation onset occurred simultaneously with the onset of the visual stimulus and lasted 200ms. The visual stimulus was always presented for at least 200ms, but in some cases lasted 300ms. Response was measured as the mean response over the interval 100-200ms after stimulus onset. Here, we show results only for two optogenetic intensities, zero and the largest intensity, but similar results hold also for intermediate intensities.

Mathematical Methods

Network Models

For all models, the single neuron response function (firing rate vs input current relationship) was taken to be the f-I curve of leaky integrate-and-fire neurons driven by white noise (Amit and Brunel, 1997; Amit and Tsodyks, 1991; Sanzeni et al., 2020; Siegert, 1951). Specifically, the firing of a neuron belonging to population $A \in [E, I]$ in response to an input x, is given by

$$r_A = \phi_A(x) = \left[\tau_{rp} + \tau_A\sqrt{\pi}\int_{u_{min,A}}^{u_{max,A}} e^{u^2}\left(1+\mathrm{erf}(u)\right)\,du\right]^{-1}, \tag{3}$$

with

$$u_{max,A} = \frac{\theta - x}{\sigma_A}, \quad u_{min,A} = \frac{V_r - x}{\sigma_A}. \tag{4}$$

In Eq. 3, τ_{rp} indicates the single neuron refractory period so that $1/\tau_{rp}$ is the maximal single neuron firing rate; τ_A is the membrane time constant; σ_A is a parameter controlling the smoothness of the transfer function. Throughout the paper, we assumed the values $\tau_{rp} = 2$ms, $\theta = 20$mV, $V_r = 10$mV, $\tau_{E,I} = 20, 10$ms and $\sigma_{E,I} = 10$mV.

Indicating with r_A^i and μ_A^i the rate and the total input (recurrent + external) of the i-th cell in population $A \in [E, I]$, activity in the network evolves in time according to the equation

$$\tau_A \frac{dr_A^i}{dt} = -r_A^i + \phi_A\left(\mu_A^i\right) , \quad \mu_A^i = \sum_{B\in[E,I,X]} \sum_{j=1}^{N_B} c_{AB}^{ij} W_{AB}^{ij} \tau_A r_B^j \tag{5}$$

Here c_{AB}^{ij} is the adjacency matrix (equal to 1 or 0 according to whether or not there is a connection from cell j of type B to cell i of type A) and N_B is the number of neurons in population B, where (X) indicates the externally driven input, *i.e.* visual input. We always use $N_X = N_E = 4\,N_I$.

Throughout the paper, we assumed uniform W's, $W_{AB}^{ij} = W_{AB}$, where the weights W_{AB} are parameterized as follows:

$$\begin{aligned} W_{EE} = J\,, \quad W_{EI} = Jg_E\,, \quad W_{EX} = (G_I g_E N_I/N_E - G_E)W_{EE} \\ W_{IE} = J/\beta\,, \quad W_{II} = Jg_I/\beta\,, \quad W_{IX} = (G_I g_I N_I/N_E - G_E)W_{IE}\,. \end{aligned} \tag{6}$$

These definitions were chosen so that the parameters G_I and G_E give the gain of the E and I population in the strong coupling limit (see (Sanzeni et al., 2020)).

With optogenetic input, the network dynamics becomes

$$\tau_A \frac{d(r_A^i + \Delta r_A^i)}{dt} = -r_A^i - \Delta r_A^i + \phi_A\left(\mu_A^i + \Delta\mu_A^i + \lambda_A^i\right) , \quad \Delta\mu_A^i = \sum_{B\in[E,I,X]} \sum_{j=1}^{N_B} c_{AB}^{ij} W_{AB}^{ij} \tau_A \Delta r_B^j\,, \tag{7}$$

where λ_A^i represents the optogenetic input of the i-th neuron in the network, which varies according to the opsin expression in that neuron, while Δr_A^i and $\Delta\mu_A^i$ indicate the change in rate and in the recurrent inputs produced by optogenetic stimuli.

Unstructured Models.

We investigated large randomly connected networks of E and I rate neurons. In the theoretical analysis of the model (mean field theory described below and simulations shown in Figures 4 and 5 and in supplementary information), the network connectivity was assumed to follow an Erdös–Rényi statistics, i.e. each element of the adjacency matrix c_{AB}^{ij} was 1 with probability $p = 0.1$ and 0 otherwise. The mean and variance of the number of pre-synaptic neurons from population B per cell are given by $K_B = pN_B$ and $\sigma_{K_B}^2 = (1-p)K_B$, respectively. We define $K \equiv K_E = pN_E$, and use the notation $\gamma = K_I/K_E$. In fitting the model to data (Figure 6), we considered more general statistics by generating c_{AB}^{ij} as follows: for each row i, a random number, K_i (Gaussian distributed, with mean K_B and standard deviation $\sigma_{K_B} = CV_K \times K_B$) of elements were set to 1, with autapses not allowed, and the other elements were set to 0 (where CV_K is a constant that is identical for $B = E$, I, or X).

Structured Models.

For the structured models, we place cells onto evenly-spaced discrete locations on a 25×25 periodic grid, so that there are $N_G = 625$ grid points, with 8 E and 2 I cells per grid location. The grid is a square of side $dl = 1$mm. For the monkey, we introduce a smooth aperiodic orientation map using a method similar to that of Kaschube et al. 2010 by summing 500 plane waves of the form $\exp\left[i(k\cos(j\pi/500)x + k\sin(j\pi/500)y + \phi_j)\right]$ for $j = 0, 1, ..., 499$. The wavenumber k is chosen to be $6\pi/dl$ such that the wavelength is $dl/3$, and the phase ϕ_j is chosen as a random angle for each j. The

preferred orientation at a point in the grid is calculated as half the phase of the sum of plane waves at that point, such that the orientations are between 0° and 180°. For the mouse, we generate a salt-and-pepper orientation map by randomly shuffling a smooth orientation map generated as described above. We let $\theta(\mathbf{x})$ be the preferred orientation at two-dimensional grid position $\mathbf{x}$.

The probability of a connection from a cell at position $\mathbf{y}$ of type A, $A \in E, I$, to any cell at position $\mathbf{x}$ is computed as follows. We let $d(\mathbf{x}, \mathbf{y})$ be the shortest distance across the periodic grid between $\mathbf{x}$ and $\mathbf{y}$, and $d\theta(\theta(\mathbf{x}), \theta(\mathbf{y}))$ be the shortest distance around the 180° circle of orientations between $\theta(\mathbf{x})$ and $\theta(\mathbf{y})$. We let n_A be the number of cells of type A at a single grid position ($n_E = 8$, $n_I = 2$). We then form the functions $\tilde{g}_l(\mathbf{x}, \mathbf{y}, A) = e^{-d(\mathbf{x},\mathbf{y})^2/(2S^2_{l,A})}$ and $\tilde{g}_{ori}(\mathbf{x}, \mathbf{y}, A) = e^{-d\theta(\theta(\mathbf{x}),\theta(\mathbf{y}))^2/(2S^2_{ori,A})}$, and normalize each by its mean to obtain the product

$$g(\mathbf{x}, \mathbf{y}, A) = \frac{\tilde{g}_l(\mathbf{x}, \mathbf{y}, A)}{\sum_{x'y'} \tilde{g}_l(\mathbf{x}', \mathbf{y}', A)/N_G^2} \frac{\tilde{g}_{ori}(\mathbf{x}, \mathbf{y}, A)}{\sum_{x'y'} \tilde{g}_{ori}(\mathbf{x}', \mathbf{y}', A)/N_G^2} .$$

We then choose the number of connections $K^i_A(\mathbf{x})$ for the ith cell at $\mathbf{x}$ from a Gaussian distribution with mean $K_A(\mathbf{x}) = pn_A \sum_{\mathbf{y}} g(\mathbf{x}, \mathbf{y}, A)$ and standard deviation $CV_K \times K_A(\mathbf{x})$, where p is a constant that is chosen such that the mean number of excitatory connections is similar to that of the unstructured model (for excitatory connections, ≈ 525 in our mouse models, ≈ 508 in our monkey models). We choose a total of $K^i_A(\mathbf{x})$ presynaptic neurons (excluding autapses) by a weighted random sample without replacement, with weights $g(\mathbf{x}, \mathbf{y}, A)/\left(\sum_{y'} n_A g(\mathbf{x}, \mathbf{y}, A)\right)$ (choice function from python package numpy.random; if there is a potential autapse, it is removed from the sum in the denominator). The visual input to a cell of type A at $\mathbf{x}$ from a stimulus centered at $\mathbf{y}$ with orientation θ_{stim} is $W_{AX} K_X r_X e^{-d(\mathbf{x},\mathbf{y})^2/(2S^2_{stim})} e^{-d\theta(\theta(\mathbf{x}),\theta_{stim})^2/(2S^2_{tune})}$. Here, $K_X = 0.1N_X$ is the number of visual inputs to a cell and r_X their firing rate. We set the input to have zero variance for simplicity. When we refer to visual-stimulus-matched cells, we mean cells with orientation tuning within 22.5° and receptive field within $0.6dl$ of the visual stimulus.

Network Simulations

Simulations of the network dynamics were performed in python. For each parameter set, we generated one realization of the network structure (adjacency matrices c^{ij}_{AB}, firing of the external population r^i_X (generated from a Gaussian distribution with mean r_X and standard deviation $\sigma_X = 0.2\, r_X$), opsin expression λ^i_A (generated from a lognormal distribution with mean λ and standard deviation σ_λ) and simulated the network dynamics given by Eqs. (5) and (7) using the explicit Runge-Kutta method of order 5 (implemented with the function solve_ivp of the python package scipy.integrate). The network dynamics was run for a total simulated time of $200\tau_E = 4$s for the unstructured model and $300\tau_E = 6$s for the structured model and computed rates were stored every $\tau_I/3 = 3.33$ms. The firing rate of a cell was obtained by measuring its average rate over the simulation time (excluding the initial $10\tau_E = 0.2$s for the unstructured model and $100\tau_E = 2$s for the structured one.) To measure optogenetic responses, we simulated the same realization of network structure twice, once with $\lambda^i_A \neq 0$ and once with $\lambda^i_A = 0$. Parameters for all simulations are given in Tables 1 (unstructured models) and 2 (structured models).

Fitting Procedure

Unstructured Model. Predictions of the model depend on a list of unknown parameters (indicated in what follows as Θ) related to: connectivity (J, K, β, g_E, g_I, CV_K); external inputs (G_E, G_I and one value of r_X per each contrast); and optogenetic inputs (λ, σ_λ). To fit these parameters to experimental data, we assumed $G_I = 2G_E$, consistent with the fact that in vivo, I firing rates are typically about twice E firing rates (e.g. (Sanzeni et al., 2020)). This choice is motivated by the fact that experimental recordings

do not distinguish cell types; this limitation makes our model under-constrained by the data (since the same average firing could be obtained by an infinite combination of firing rates of E and I cells). Without loss of generality, we assumed $G_E = 1$ (any other choice can be obtained by rescaling r_X) and $K = 500$ (any other choice can be obtained by rescaling $J \sim 1/\sqrt{K}$). Moreover, we took $\sigma_X = 0$, and controlled variability in the external input by varying CV_K.

With the above assumptions, model parameters were inferred as follows. First, we measured optogenetic responses in a large number (770195) of network simulations; each one obtained assuming model parameters generated randomly and uniformly in the intervals

$$\begin{aligned} g_E \in [3,7]\,, \quad g_I \in [2, g_E - 0.5]\,, \quad \log_{10}(\beta) \in [-1,1]\,, \\ \log_{10}(\sigma_\lambda/\lambda) \in [-1,1]\,, \quad \log_{10}(J) \in [-5,-3]\,, \quad \log_{10}(r_X) \in [-0.3,1.7]\,, \\ \log_{10}(\lambda) \in [-0.3,1.7]\,, \quad \log_{10}(CV_K) \in [-3.5,-0.5]. \end{aligned} \tag{8}$$

Starting from the results of these simulations, we trained a multilayer perceptron (MLPRegressor from python package sklearn.neural_network (Pedregosa et al., 2011); specifics were activation=ReLU, hidden_layer_sizes=(100, 150, 50),)) to predict the mapping $f(\Theta)$ between model parameters (Θ) and moments of the optogenetic response (r, Δr, σ_r, $\sigma_{\Delta r}$, $\mathrm{Cov}(r, \Delta r)$). The learned multilayer perceptron was able to predict the outcome of simulations with high accuracy (with 20% of held-out simulations, the crossvalidated performance had $R^2 = 0.924$). We then fitted model parameters Θ to data by minimizing the squared error between predictions given by $f(\Theta)$ and moments of the optogenetic responses measured in experiments (least_squares function from python package scipy.optimize). In the fits, we excluded simulation parameters leading to strongly chaotic dynamics by imposing an additional penalty in the least squared error for networks whose predicted autocorrelation time was smaller than 1 second. Specifically, all moments of the network response for which the predicted autocorrelation time was smaller than 1 second were set to 10^{10}; this prevented the least square optimization algorithm from exploring regions of the parameter space that lead to strong chaotic dynamics.

Structured Model. In order to fit the structured model, we allowed G_I to vary since the spatial structure often causes the network to be more inhibition dominated than the unstructured model. As for the unstructured model, we measured optogenetic responses in a large number (484848 with salt-and-pepper maps and 508719 with smooth maps) of network simulations; we generated parameters randomly and uniformly in the following intervals

$$\begin{aligned} G_I \in [1,2]\,, \quad \log_{10}(\sigma_\lambda/\lambda) \in [-0.5,1.5]\,, \\ S_{l,E} \in [0,1]\,, \quad S_{l,I} \in [(2/3)S_{l,E}, 2S_{l,E}]\,, \\ S_{ori,I} \in [20,60]\,, \quad S_{tune} \in [10,40]. \end{aligned} \tag{9}$$

In addition, we excluded parameters where the parameterization from Eq. (6) would give negative mean external input to any population. We fixed $S_{ori,E} = 30°$ and $S_{stim} = 0.5dl$. All other parameters not listed above are generated with the same intervals as for the unstructured model. For each simulation, we also generate a different random orientation map.

Instead of training a multilayer perceptron to predict the mapping between model parameters and moments as in the unstructured model, we looked for simulations that already had a small squared error between the simulated predictions and the moments from the experimental data at any contrast. For these models we then resimulated them, averaging moments across four different orientation maps, and lastly if the moments still robustly fit the data after averaging then we calculated the moments at various r_X in order to find a mapping between r_X and contrast that minimized the squared error (least_squares

Parameter	Symbol	Units	Figure 4 (weak, strong)	Figure 5	Figure 6 (comb., mice, monk.)
Mean E connections per neuron	K		2000	500	500
Coeff. of Variation of In-Degree	CV_K		(1-p)/$\sqrt{K}$	(1-p)/$\sqrt{K}$	(0.015, 0.020, 0.007)
Connection probability	p		0.1	0.1	0.1
Number of excitatory neurons	N_E		K/p	K/p	K/p
Ratio of I to E Population Size	γ		0.25	0.25	0.25
E time constant	τ_E	ms	20	20	20
I time constant	τ_I	ms	10	10	10
Synaptic efficacy	J	mV	(0.01, 0.3)	(**C**) 0.6	(0.14, 0.13, 0.44)
Ratio of I-to-E over E-to-E	g_E		8	(**A,B,D**) 8	(4.2, 9.97, 4.58)
Ratio of I-to-I over E-to-I	g_I		3	(**A,B,D**) 3	(3.24, 8.4, 2.73)
Ratio of E-to-E over E-to-I	β		1	1	(0.17, 0.33, 0.81)
Gain of the E population	G_E		from Eq. 6	1	1
Gain of the I population	G_I		from Eq. 6	2	2
Strength of FF input to E	W_{EX}		3	from Eq. 6	from Eq. 6
Strength of FF input to I	W_{IX}		(0.5,2.5)	from Eq. 6	from Eq. 6
Mean FF input rates	r_X	spk/s	20	(**C**) 10, (**D**) 17.97	([13.5 - 23.4] & [30.9 - 77.6], [7.24 - 13.8], [14.45 - 47.86])
Std FF input rates	σ_X	spk/s	$0.2r_X$	$0.2r_X$	0
Mean optogenetic input strength	λ	mV	20	20	(1.9 & 5.5, 4.2, 2.5)
Std / mean of opsin expression	σ_λ/λ		1	(**A,B**) 2.15, (**C**) 1	(1.69, 0.79, 6.6)

Table 1: List of network parameters for unstructured models for Figures 4-6. Here and in Table 2 "FF input" means feedforward input, *i.e.* external (visual) input.

function from python package scipy.optimize).

Mean Field Theory for the Unstructured Model

For $N_B \gg 1$ in the unstructured model, μ_A^i and $\Delta\mu_A^i$ in Eqs. (5) and (7) are well approximated by correlated random Gaussian variables with

$$\begin{aligned}
\mu_A &= \tau_A \sum_{B\in[E,I,X]} W_{AB}K_B r_B\,, \quad \sigma_{\mu_A}^2 = \tau_A^2 \sum_{B\in[E,I,X]} W_{AB}^2 K_B \left[\sigma_{r_B}^2 + (1-p)r_B^2\right]\,, \\
\Delta\mu_A &= \tau_A \sum_{B\in[E,I]} W_{AB}K_B \Delta r_B\,, \quad \sigma_{\Delta\mu_A}^2 = \tau_A^2 \sum_{B\in[E,I]} W_{AB}^2 K_B \left[\sigma_{\Delta r_B}^2 + (1-p)\Delta r_B^2\right]\,, \\
\mathrm{Cov}(\mu_A, \Delta\mu_A) &= \tau_A^2 \sum_{B\in[E,I]} W_{AB}^2 K_B \left[\mathrm{Cov}(r_B, \Delta r_B) + (1-p) r_B \Delta r_B\right]\,.
\end{aligned} \tag{10}$$

In Eq. (10), μ_A ($\Delta\mu_A$) and σ_{μ_A} ($\sigma_{\Delta\mu_A}$) are the mean and variance of μ_A^i ($\Delta\mu_A^i$). The covariance

Parameter	Symbol	Units	Mouse	Monkey
Mean FF E connections per neuron	K_X		500	500
Coeff. of Variation of In-Degree	CV_K		0.0398	0.0004
Mean recurrent connection probability	p		0.09	0.09
Number of grid locations	N_l		25×25	25×25
Number of E neurons per grid location	N_E		8	8
Ratio of I to E Population Size	γ		0.25	0.25
E time constant	τ_E	ms	20	20
I time constant	τ_I	ms	10	10
Synaptic efficacy	J	mV	0.123	0.525
Ratio of I-to-E over E-to-E	g_E		5.56	4.61
Ratio of I-to-I over E-to-I	g_I		2.16	2.49
Ratio of E-to-E over E-to-I	β		0.99	1.40
Gain of the E population	G_E		1	1
Gain of the I population	G_I		1.92	1.70
Strength of FF input to E	W_{EX}		from Eq. 6	from Eq. 6
Strength of FF input to I	W_{IX}		from Eq. 6	from Eq. 6
Mean FF input rates	r_X	spk/s	[7.52 - 15.54]	[5.33 - 20.12]
Std FF input rates	σ_X	spk/s	0	0
Mean optogenetic input strength	λ	mV	1.72	7.62
Std / mean of opsin expression	σ_λ/λ		2.55	4.55
Spatial grid length	dl	mm	1	1
Spatial width of FF input	S_{stim}	mm	0.5	0.5
Tuning width of FF input	S_{tune}	deg°	31.14	34.40
Spatial width of recurrent E	$S_{l,E}$	mm	0.028	0.833
Spatial width of recurrent I	$S_{l,I}$	mm	0.031	0.589
Tuning width of recurrent E	$S_{ori,E}$	deg°	30	30
Tuning width of recurrent I	$S_{ori,I}$	deg°	47.28	38.76

Table 2: List of network parameters of best fit structured models for Figures 6 and S3.

$\mathrm{Cov}(\mu_A, \Delta\mu_A)$ is computed starting from

$$\mu_A^i + \Delta\mu_A^i = \sum_{B\in[E,I,X]} \tau_A W_{AB} \sum_{j=1}^{N_B} c_{AB}^{ij} \left(r_B^j + \Delta r_B^j \right) , \tag{11}$$

applying the identity $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X,Y)$ to both sides and using the central limit theorem. Note that optogenetic input affects firing differently than external and recurrent inputs. In fact, firing depends on the whole distribution of λ_A^i but only on the first two moments of external and recurrent inputs.

If the network dynamics settle in a stable fixed point, the statistics of currents and rates in the network are related by

$$r_A^i = \phi_A \left[\mu_A^i \right] , \quad r_A^i + \Delta r_A^i = \phi_A \left[\lambda_A^i + \mu_A^i + \Delta\mu_A^i \right] , \tag{12}$$

and moments of the rates can be computed by self-consistently solving a set of ten equations

$$\begin{cases} r_A = \int_{-\infty}^{\infty} d\mu_A^i P(\mu_A^i)\phi(\mu_A^i)\,, \\ \sigma_{r_A}^2 = \int_{-\infty}^{\infty} d\mu_A^i P(\mu_A^i)\phi(\mu_A^i)^2 - r_A^2\,, \\ \Delta r_A = \int_{-\infty}^{\infty} d\mu_A^i d\Delta\mu_A^i d\lambda_A^i P(\mu_A^i, \Delta\mu_A^i) P(\lambda_A^i) \left[\Delta\phi_A^i\right]\,, \\ \sigma_{\Delta r_A}^2 = \int_{-\infty}^{\infty} d\mu_A^i d\Delta\mu_A^i d\lambda_A^i P(\mu_A^i, \Delta\mu_A^i) P(\lambda_A^i) \left[\Delta\phi_A^i\right]^2 - \Delta r_A^2\,, \\ \mathrm{Cov}(r_A, \Delta r_A) = \int_{-\infty}^{\infty} d\mu_A^i d\Delta\mu_A^i d\lambda_A^i P(\mu_A^i, \Delta\mu_A^i) P(\lambda_A^i) \left[\phi(\mu_A^i) - r_A\right] \left[\Delta\phi_A^i - \Delta r_A\right]\,, \end{cases} \tag{13}$$

with $\Delta\phi_A^i = \phi(\mu_A^i + \Delta\mu_A^i + \lambda_A^i) - \phi(\mu_A^i)$.

If the fixed point of the network dynamics is unstable, rates and currents become time-dependent. One can again use a Gaussian ansatz, in which the total inputs at different times ($\mu_A^i(s)$ and $\mu_A^i(t)$) are jointly Gaussian with means $\mu_A(s), \mu_A(t)$ and covariance $\sigma_{\mu_A}^2(s,t)$. Analogously, rates at different times ($r_A^i(s)$ and $r_A^i(t)$) are jointly Gaussian with means $r_A(s), r_A(t)$ and covariance $\sigma_{r_A}^2(s,t)$. At any given pair of time points t and s, the moments of rates and currents are related by Eq. (10). The temporal evolution of the moments can be computed using standard techniques (Sompolinsky et al., 1988). In particular, indicating with $C_{r_A}(s,t) = N_A^{-1}\sum_{i=1}^{N_A} r_A^i(s) r_A^i(t)$ the autocorrelation function of the rates of population A at times s and t, we can define the autocovariance function of the rates as

$$\sigma_{r_A}^2(s,t) = C_{r_A}(s,t) - r_A(s) r_A(t)\,. \tag{14}$$

By multiplying Eq. (7) at times s and t, averaging over cells in population A, and rearranging terms we see that $C_{r_A}(s,t)$ evolves in time according to

$$\left(\tau_A \frac{d}{ds} + 1\right)\left(\tau_A \frac{d}{dt} + 1\right) C_{r_A}(s,t) = C_{\phi_A[\mu_A]}(s,t)\,.$$

We can repeat the above steps to derive equations for the evolution in time of all five order parameters per population of neurons (mean, autocovariance, and cross-covariance of rates without and with optogenetic stimuli). Defining

$$M_{\phi_A[\mu_A]}(t) = N_A^{-1} \sum_{i=1}^{N_A} \phi_A\left[\mu_A^i(t)\right]$$

and

$$M_{\phi_A[\tilde{\mu}_A]}(t) = N_A^{-1} \sum_{i=1}^{N_A} \phi_A\left[\mu_A^i(t) + \Delta\mu_A^i(t) + \lambda_A^i\right]\,,$$

the order parameters evolve in time according to

$$\begin{aligned} \left(\tau_A \frac{d}{dt} + 1\right) r_A &= M_{\phi_A[\mu_A]}(t)\,, \quad \left(\tau_A \frac{d}{dt} + 1\right) \Delta r_A = M_{\phi_A[\tilde{\mu}_A]}(t) - M_{\phi_A[\mu_A]}(t)\,, \\ \left(\tau_A \frac{d}{ds} + 1\right)\left(\tau_A \frac{d}{dt} + 1\right) C_{r_A}(s,t) &= C_{\phi_A[\mu_A]}(s,t)\,, \\ \left(\tau_A \frac{d}{ds} + 1\right)\left(\tau_A \frac{d}{dt} + 1\right) C_{\Delta r_A}(s,t) &= C_{\phi_A[\tilde{\mu}_A]}(s,t) - \mathrm{Corr}(\phi_A\left[\mu_A\right], \phi_A\left[\tilde{\mu}_A\right])(s,t) \\ &\quad - \mathrm{Corr}(\phi_A\left[\tilde{\mu}_A\right], \phi_A\left[\mu_A\right])(s,t) + C_{\phi_A[\mu_A]}(s,t)\,, \\ \left(\tau_A \frac{d}{ds} + 1\right)\left(\tau_A \frac{d}{dt} + 1\right) \mathrm{Corr}(r_A, \Delta r_A)(s,t) &= \mathrm{Corr}(\phi_A\left[\mu_A\right], \phi_A\left[\tilde{\mu}_A\right])(s,t) - C_{\phi_A[\mu_A]}(s,t)\,. \end{aligned} \tag{15}$$

Numerical Solution of Mean Field Equations. We solve for the stationary statistics of the chaotic network by explicitly integrating Eq. (15) in time until the order parameters converge to a steady state. We start with initial values for $r_A(0)$ and $\Delta r_A(0)$ as well as initial functions for $C_{r_A}(s,0) = C_{r_A}(0,t)$, $C_{\Delta r_A}(s,0) = C_{r_A}(0,t)$, and $\mathrm{Corr}(r_A, \Delta r_A)(s,0) = \mathrm{Corr}(r_A, \Delta r_A)(0,t)$. The means $r_A(t)$ and $\Delta r_A(t)$ are integrated with the usual explicit RK4 method, but the autocovariance and cross-covariance functions are integrated on a grid with an explicit Euler-like method derived using forward finite differences. For example, when integrating $C_{r_A}(s,t)$ we can express $C_{r_A}(s+dt, t+dt)$ in terms of $r_A(s)$, $r_A(t)$, $C_{r_A}(s,s)$, $C_{r_A}(t,t)$, $C_{r_A}(s,t)$, $C_{r_A}(s{+}dt,t)$, and $C_{r_A}(s,t{+}dt)$. Assuming that we have solved for the order parameters at all time points $(s \leq t_0, t \leq t_0)$, we can explicitly solve for the order parameters at time points $(s \leq t_0 + dt, t_0 + dt)$ and $(t_0 + dt, t \leq t_0 + dt)$. Furthermore, we can ignore portions of the grid for $|s-t| > T$ for some T that is much longer than the width of the steady state autocorrelation function in order to reduce the computation time required for the integration.

A useful simplification of the equations governing the evolution of the order parameters is to calculate quantities such as $C_{\phi_A[\mu_A]}(s,t)$ by setting the mean and variance of μ_A^i at times s and t equal to the mean and variance at time $\max(s,t)$. While this change increases the required integration time for the order parameters to reach the steady state, it reduces the number of dependent variables needed to calculate quantities like $C_{\phi_A[\mu_A]}(s,t)$ from five to three. With only three dependent variables, it becomes feasible to pre-calculate these autocorrelation functions on a 3D grid and interpolate between the pre-computed points in order to accelerate the integration. In addition, we can precompute quantities such as $M_{\phi_A[\mu_A]}(t)$, which only depends on the mean and variance of the inputs, on a 2D grid and interpolate them as well.

The cross-correlation function $\mathrm{Corr}(\phi_A[\mu_A], \phi_A[\tilde{\mu}_A])(s,t)$ cannot be pre-computed and interpolated since the means and variances of the cross-correlated variables are not equal. However, since we have already interpolated $M_{\phi_A[\mu_A]}(t)$ and $M_{\phi_A[\tilde{\mu}_A]}(t)$ we can efficiently calculate the cross-correlation function from a single Gaussian integral.

Indicating with $M_{\phi_A[\mu_A]}\left[\mu_A(t), \sigma^2_{\mu_A}(t,t)\right]$ and $M_{\phi_A[\tilde{\mu}_A]}\left[\mu_A(t) + \Delta\mu_A(t), \sigma^2_{\mu_A+\Delta\mu_A}(t,t)\right]$ the precomputed mean rates as a function of the means and variances of the inputs, we can express the cross-correlation function as

$$\begin{aligned}\mathrm{Corr}(\phi_A[\mu_A], &\phi_A[\tilde{\mu}_A])(s,t) = \\ &\int_{-\infty}^{\infty} Dx M_{\phi_A[\mu_A]}\left[\mu_A(s) + \mathrm{sign}(\eta)\, x\sqrt{|\eta|}\,\sigma_{\mu_A}(s,s), (1-|\eta|)\sigma^2_{\mu_A}(s,s)\right] \\ &\quad \times M_{\phi_A[\tilde{\mu}_A]}\left[\mu_A(t) + \Delta\mu_A(t) + x\sqrt{|\eta|}\sigma_{\mu_A+\Delta\mu_A}(t,t), (1-|\eta|)\sigma^2_{\mu_A+\Delta\mu_A}(t,t)\right],\end{aligned} \tag{16}$$

with the Pearson correlation coefficient η defined as

$$\eta = \frac{\mathrm{Cov}(\mu_A, \mu_A + \Delta\mu_A)(s,t)}{\sigma_{\mu_A}(s,s)\sigma_{\mu_A+\Delta\mu_A}(t,t)}\ .$$

Acknowledgments

This work was supported by the NIMH Intramural Research Program and by NIH BRAIN U01 NS108683 (to N.B., M.H.H., K.D.M. and J.H.R.). A.S., A.P., T.H.N. and K.D.M. were partially supported by the Gatsby Charitable Foundation GAT3708 and by NSF Grant DBI-1707398. A.P. and K.D.M. were partially supported by NIH BRAIN U19 NS107613. A.P. would like to acknowledge the support of the Swartz Foundation Fellowship for Theory in Neuroscience 2019-4 and 2020-6. T.H.N. would like to acknowledge the support by NIH grant 5R01EY029999. J.J.N was funded, in part, by NIH grant R01-EY021827. J. Luo was supported, in part, by the Fiona and Sanjay Jha Chair in Neuroscience. This work used the computational resources of the Duke Compute Cluster.

References

Ahmadian, Y. and K. D. Miller (2021). What is the dynamical regime of cerebral cortex? *Neuron 109*, 3373–3391.

Ahmadian, Y., D. B. Rubin, and K. D. Miller (2013). Analysis of the stabilized supralinear network. *Neural Computation 25*(8), 1994–2037.

Akitake, B., H. M. Douglas, P. K. LaFosse, C. E. Deveau, A. J. Li, L. N. Ryan, S. P. Duffy, Z. Zhou, Y. Deng, and M. H. Histed (2022). Amplified cortical neural responses as animals learn to use novel activity patterns. *bioRxiv*.

Amit, D. J. and N. Brunel (1997). Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. *Cerebral Cortex 7*, 237–252.

Amit, D. J. and M. V. Tsodyks (1991). Quantitative study of attractor neural network retrieving at low spike rates I: Substrate – spikes, rates and neuronal gain. *Network 2*, 259–274.

Andrei, A. R., S. Pojoga, R. Janz, and V. Dragoi (2019). Integration of cortical population signals for visual perception. *Nat. Commun. 10*(1), 1–13.

Angelucci, A., J. B. Levitt, E. J. S. Walton, J.-M. Hupe, J. Bullier, and J. S. Lund (2002). Circuits for local and global signal integration in primary visual cortex. *J Neurosci 22*, 8633–8646.

Barth, A. L. and J. F. Poulet (2012). Experimental evidence for sparse firing in the neocortex. *Trends Neurosci. 35*, 345–355.

Bouvier, G., Y. Senzai, and M. Scanziani (2020). Head movements control the activity of primary visual cortex in a luminance-dependent manner. *Neuron 108*, 500–511.

Brunel, N. (2000). Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *J. Comput. Neurosci. 8*, 183–208.

Buzsáki, G. and K. Mizuseki (2014, 04). The log-dynamic brain: how skewed distributions affect network operations. *Nature reviews. Neuroscience 15*(4), 264–278.

Chen, S. C. Y., G. Benvenuti, Y. Chen, S. Kumar, C. Ramakrishnan, K. Deisseroth, W. S. Geisler, and E. Seidemann (2022). Similar neural and perceptual masking effects of low-power optogenetic stimulation in primate V1. *Elife 11*, 1–21.

Chettih, S. N. and C. D. Harvey (2019). Single-neuron perturbations reveal feature-specific competition in V1. *Nature 567*, 334–340.

Compte, A., C. Constantinidis, J. Tegnér, S. Raghavachari, M. Chafee, P. S. Goldman-Rakic, and X.-J. Wang (2003). Temporally irregular mnemonic persistent activity in prefrontal neurons of monkeys during a delayed response task. *J. Neurophysiol. 90*, 3441–3454.

Cossell, L., M. F. Iacaruso, D. R. Muir, R. Houlton, E. N. Sader, H. Ko, S. B. Hofer, and T. D. Mrsic-Flogel (2015). Functional organization of excitatory synaptic strength in primary visual cortex. *Nature 518*, 399–403.

Dalgleish, H. W., L. E. Russell, A. M. Packer, A. Roth, O. M. Gauld, F. Greenstreet, E. J. Thompson, and M. Häusser (2020). How many neurons are sufficient for perception of cortical activity? *eLife 9*, e58889.

Dräger, U. C. (1974). Autoradiography of tritiated proline and fucose transported transneuronally from the eye to the visual cortex in pigmented and albino mice. *Brain Res. 82*(2), 284–292.

Engelhard, B., R. Darshan, N. Ozeri-Engelhard, Z. Israel, U. Werner-Reiss, D. Hansel, H. Bergman, and E. Vaadia (2019). Neuronal activity and learning in local cortical networks are modulated by the action-perception state. *bioRxiv*.

Engelken, R., F. Wolf, and L. F. Abbott (2020). Lyapunov spectra of chaotic recurrent neural networks. *arXiv preprint, arXiv:2006.02427*.

Haider, B., A. Duque, A. R. Hasenstaub, and D. A. McCormick (2006). Neocortical network activity in vivo is generated through a dynamic balance of excitation and inhibition. *J. Neurosci. 26*, 4535–4545.

Harish, O. and D. Hansel (2015). Asynchronous Rate Chaos in Spiking Neuronal Circuits. *PLoS Comput Biol 11*, e1004266.

Histed, M. H. (2018). Feedforward inhibition allows input summation to vary in recurrent cortical networks. *eNeuro 5*(1).

Hromádka, T., M. R. DeWeese, and A. M. Zador (2008, 01). Sparse representation of sounds in the unanesthetized auditory cortex. *PLOS Biology 6*(1), 1–14.

Hubel, D. and T. Wiesel (1977). Ferrier lecture. functional architecture of macaque monkey visual cortex. *Proc R Soc Lond B Biol Sci. 198*, 1–59.

Inagaki, H. K., L. Fontolan, S. Romani, and K. Svoboda (2019). Discrete attractor dynamics underlies persistent activity in the frontal cortex. *Nature 566*, 212–217.

Kadmon, J. and H. Sompolinsky (2015). Transition to chaos in random neuronal networks. *Phys. Rev. X 5*, 041030.

Kaschube, M., M. Schnabel, S. Löwel, D. M. Coppola, L. E. White, and F. Wolf (2010). Universality in the evolution of orientation columns in the visual cortex. *Science 330*(6007), 1113–1116.

Kayser, A. S. and K. D. Miller (2002). Opponent inhibition: A developmental model of layer 4 of the neocortical circuit. *Neuron 33*, 131–142.

Kim, C. K., A. Adhikari, and K. Deisseroth (2017). Integration of optogenetics with complementary methodologies in systems neuroscience. *Nat Rev Neurosci 18*, 222–235.

Kim, S. S., H. Rouault, S. Druckmann, and V. Jayaraman (2017). Ring attractor dynamics in the Drosophila central brain. *Science 356*, 849–853.

Ko, H., S. B. Hofer, B. Pichler, K. A. Buchanan, P. J. Sjostrom, and T. D. Mrsic-Flogel (2011). Functional specificity of local synaptic connections in neocortical networks. *Nature 473*, 87–91.

Li, Y. T., L. A. Ibrahim, B. H. Liu, L. I. Zhang, and H. W. Tao (2013). Linear transformation of thalamocortical input by intracortical excitation. *Nat. Neurosci. 16*, 1324–1330.

Lien, A. D. and M. Scanziani (2013). Tuned thalamic excitation is amplified by visual cortical circuits. *Nat. Neurosci. 16*, 1315–1323.

Liska, J. P., D. P. Rowley, T. T. K. Nguyen, J.-O. Muthmann, D. A. Butts, J. L. Yates, and A. C. Huk (2022). Running modulates primate and rodent visual cortex via common mechanism but quantitatively distinct implementation. *bioRxiv*.

Mattis, J., K. M. Tye, E. A. Ferenczi, C. Ramakrishnan, D. J. O'Shea, R. Prakash, L. A. Gunaydin, M. Hyun, L. E. Fenno, V. Gradinaru, O. Yizhar, and K. Deisseroth (2011). Principles for applying optogenetic tools derived from direct comparative analysis of microbial opsins. *Nat Methods 9*, 159–172.

Nassi, J. J., M. C. Avery, A. H. Cetin, A. W. Roe, and J. H. Reynolds (2015). Optogenetic activation of normalization in alert macaque visual cortex. *Neuron 86*(6), 1504–1517.

Niell, C. M. and M. P. Stryker (2010). Modulation of visual responses by behavioral state in mouse visual cortex. *Neuron 65*(4), 472 – 479.

O'Connor, D. H., S. P. Peron, D. Huber, and K. Svoboda (2010). Neural activity in barrel cortex underlying vibrissa-based object localization in mice. *Neuron 67*(6), 1048 – 1061.

Ohki, K., S. Chung, Y. H. Ch'ng, P. Kara, and R. C. Reid (2005). Functional imaging with cellular resolution reveals precise micro-architecture in visual cortex. *Nature 433*(7026), 597–603.

Ohki, K., S. Chung, P. Kara, M. Hübener, T. Bonhoeffer, and R. C. Reid (2006). Highly ordered arrangement of single neurons in orientation pinwheels. *Nature 442*(7105), 925–928.

Ozeki, H., I. M. Finn, E. S. Schaffer, K. D. Miller, and D. Ferster (2009). Inhibitory stabilization of the cortical network underlies visual surround suppression. *Neuron 62*, 578–592.

O'Rawe, J. F., Z. Zhou, A. J. Li, P. K. LaFosse, H. C. Goldbach, and M. H. Histed (2022). Cortical suppression from excitatory input due to heterogeneous recurrent connectivity. *bioRxiv*.

Palmigiano, A., F. Fumarola, D. P. Mossing, N. Kraynyukova, H. Adesnik, and K. D. Miller (2020). Structure and variability of optogenetic responses identify the operating regime of cortex. *bioRxiv*, 2020.11.11.378729.

Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research 12*, 2825–2830.

Renart, A., J. de la Rocha, P. Bartho, L. Hollender, N. Parga, A. Reyes, and K. D. Harris (2010). The asynchronous state in cortical circuits. *Science 327*, 587–590.

Rossi, L. F., K. D. Harris, and M. Carandini (2020). Spatial connectivity matches direction selectivity in visual cortex. *Nature 588*, 648–652.

Roxin, A., N. Brunel, D. Hansel, G. Mongillo, and C. van Vreeswijk (2011). On the distribution of firing rates in networks of cortical neurons. *Journal of Neuroscience 31*(45), 16217–16226.

Rubin, D. B., S. D. Van Hooser, and K. D. Miller (2015). The stabilized supralinear network: a unifying circuit motif underlying multi-input integration in sensory cortex. *Neuron 85*, 402–417.

Sadeh, S. and C. Clopath (2020). Theory of neuronal perturbome in cortical networks. *Proc Natl Acad Sci U S A 117*, 26966–26976.

Sanzeni, A., B. Akitake, H. C. Goldbach, C. E. Leedy, N. Brunel, and M. H. Histed (2020). Inhibition stabilization is a widespread property of cortical networks. *Elife 9*.

Sanzeni, A., M. H. Histed, and N. Brunel (2020). Response nonlinearities in networks of spiking neurons. *PLoS Comput Biol 16*, e1008165.

Sanzeni, A., M. H. Histed, and N. Brunel (2022). Emergence of irregular activity in networks of strongly coupled conductance-based neurons. *Phys. Rev. X 12*, 011044.

Sherman, S. M. and R. W. Guillery (1998). On the actions that one nerve cell can have on another: distinguishing "drivers" from "modulators". *Proc Natl Acad Sci U S A 95*, 7121–7126.

Siegert, A. J. F. (1951). On the first passage time probability problem. *Phys. Rev. 81*, 617–623.

Siepel, A. (2009). Phylogenomics of primates and their ancestral populations. *Genome Res. 19*(11), 1929–1941.

Softky, W. R. and C. Koch (1993). The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. *J. Neurosci. 13*, 334–350.

Sompolinsky, H., A. Crisanti, and H. J. Sommers (1988). Chaos in random neural networks. *Phys. Rev. Lett. 61*, 259–262.

Troyer, T. W., A. E. Krukowski, N. J. Priebe, and K. D. Miller (1998). Contrast-invariant orientation tuning in cat visual cortex: thalamocortical input tuning and correlation-based intracortical connectivity. *J. Neurosci. 18*, 5908–5927.

Tsodyks, M. V. and H. Markram (1997). The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proc. Natl. Acad. Sci. USA 94*, 719–723.

Van Den Bergh, G., B. Zhang, L. Arckens, and Y. M. Chino (2010). Receptive-field properties of V1 and V2 neurons in mice and macaque monkeys. *J. Comp. Neurol. 518*(11), 2051–2070.

Van Hooser, S. D., J. A. Heimel, S. Chung, S. B. Nelson, and L. J. Toth (2005). Orientation selectivity without orientation maps in visual cortex of a highly visual mammal. *J. Neurosci. 25*, 19–28.

van Vreeswijk, C. and H. Sompolinsky (1996). Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science 274*, 1724–1726.

van Vreeswijk, C. and H. Sompolinsky (1998). Chaotic balanced state in a model of cortical circuits. *Neural Computation 10*, 1321–1371.

Wang, S., A. Palmigiano, K. D. Miller, and S. D. Van Hooser (2022). Targeted cortical stimulation reveals principles of cortical contextual interactions. *bioRxiv*.

Wildenberg, G. A., M. R. Rosen, J. Lundell, D. Paukner, D. J. Freedman, and N. Kasthuri (2021). Primate neuronal connections are sparse in cortex as compared to mouse. *Cell Rep 36*, 109709.

Yona, G., N. Meitav, I. Kahn, and S. Shoham (2016). Realistic Numerical and Analytical Modeling of Light Scattering in Brain Tissue for Optogenetic Applications(1,2,3). *eNeuro 3*(1).

Supplementary Information

Theoretical results are preserved with different transfer functions

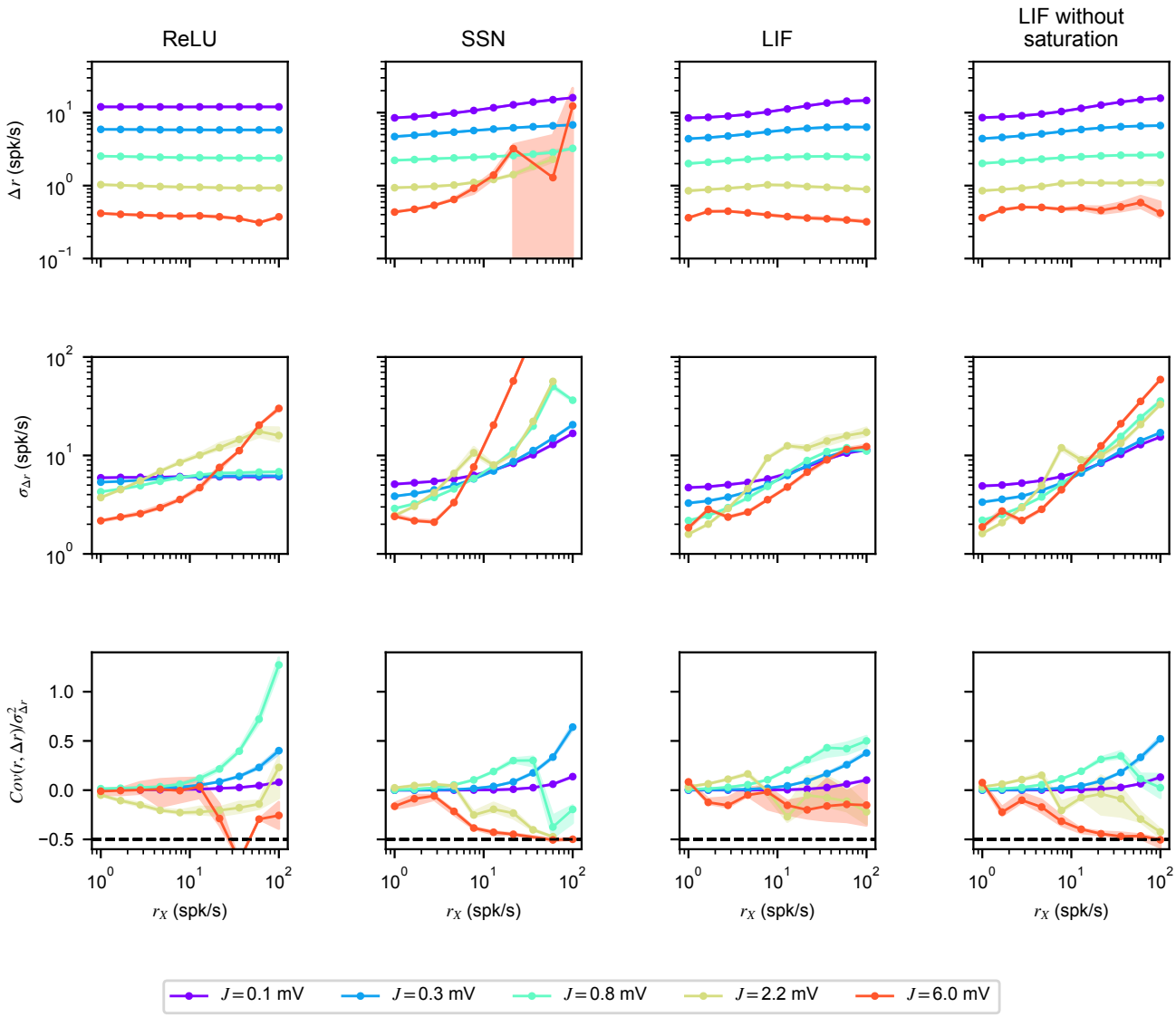

Figure S1: **Neuronal reshuffling emerges in network models with different single transfer functions.** Optogenetic responses obtained in numerical simulations of inhibitory networks (dynamics given determined by Eqs. (5) and (7), without E cells and with optogenetic stimuli targeting I cells) as a function of J (colors) and r_X, for different choices of the single neuron transfer function ϕ_A (titles). Plots are obtained averaging over 20 realizations of the network models (dots represent means, shaded regions are mean±sem). Transfer function used were: $\phi(x) = a\,[x]_+$ (ReLU, $a = 1.5$ mV^{-1}); $\phi(x) = a[x]_+^2$ (SSN, $a = 0.05$ mV^{-2}); $\phi(x)$ as in Eq. (3) with τ_{rp} =2ms (LIF) or $\tau_{rp} = 0$ (LIF without saturation).

Chaos is not required for negative covariance to appear

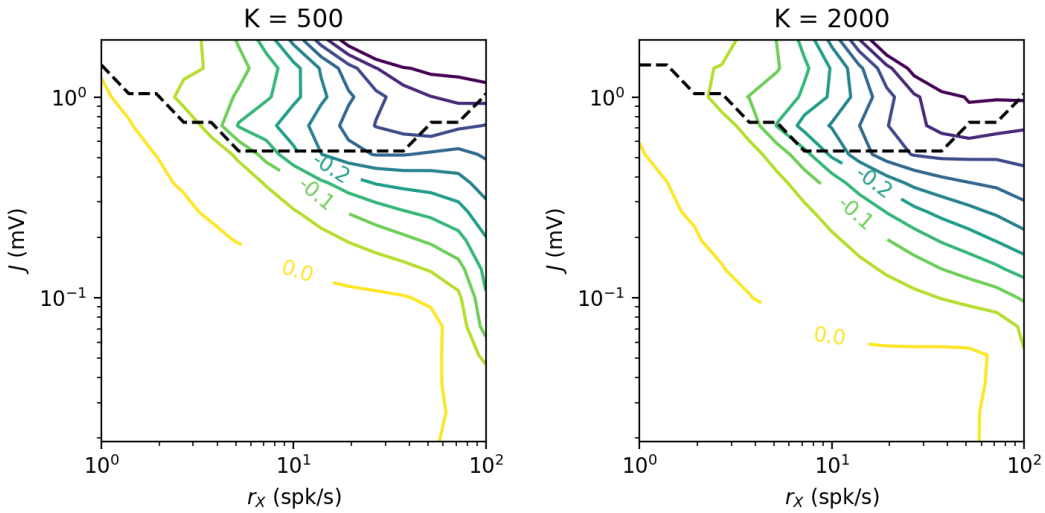

Figure S2: **Phase diagram of chaos and negative covariance.** Colored lines indicate contours of constant normalized covariance between 0 and -0.5 as a function of the external input r_X and recurrent synaptic efficacies J. The dashed line indicates the transition to chaos; Below the dashed line the baseline system is stationary, and above the line the network is chaotic. In order to investigate the effect of finite network size, the left plot is calculated with $K = 500$ while the right plot is calculated with $K = 2000$. Increasing K tends to increase the magnitude of the normalized covariance while decreasing the region of chaos. All values are calculated from mean field theory. Network parameters are identical to those used in Figure 5 panels **A-B**.

Structured Connectivity Reshuffles Visual-Stimulus-Matched Rates Even with Large Optogenetic Input

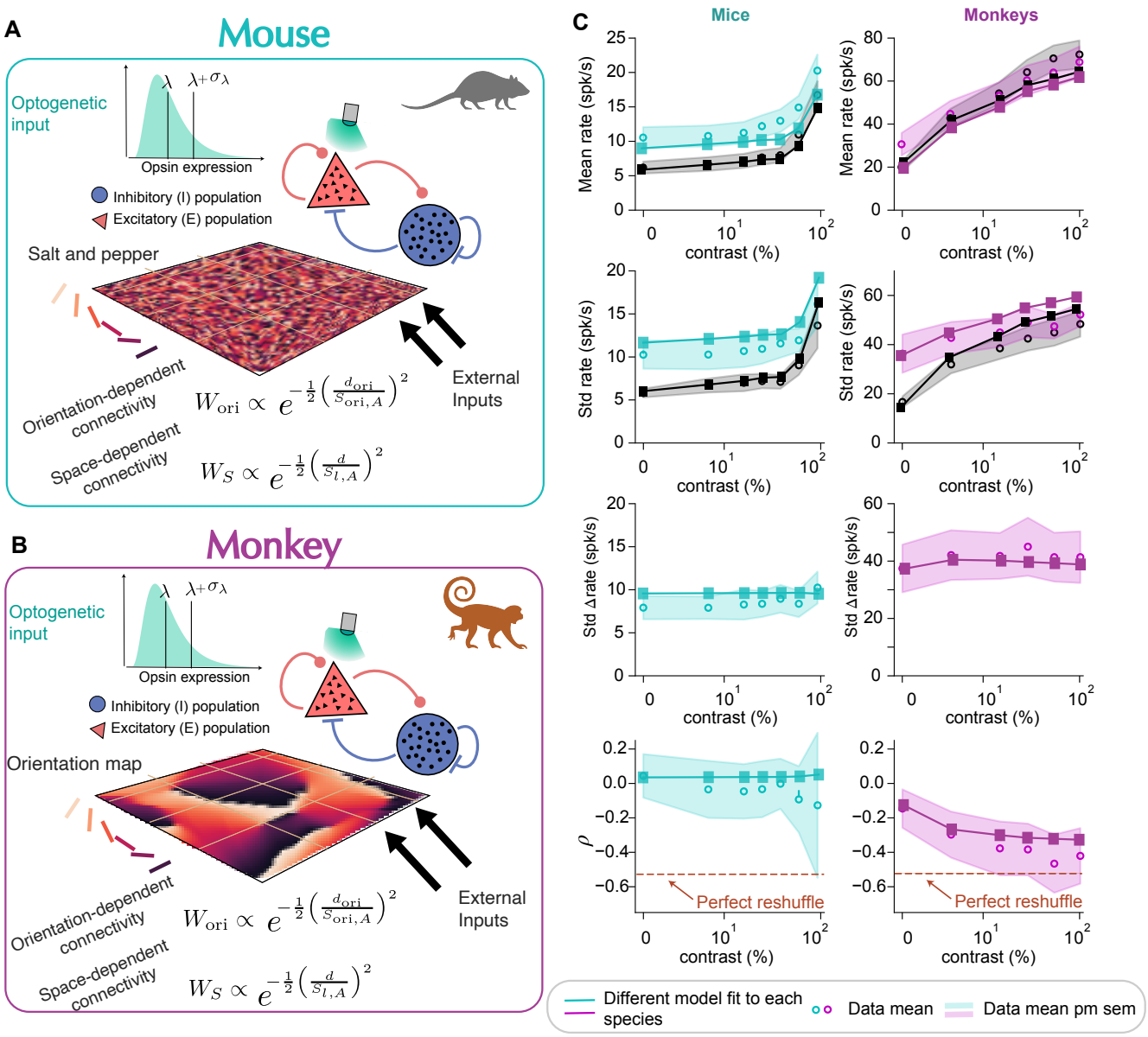

Figure S3: **Structured models capture the data.** **A** Structure of the mouse model. The model consists of a grid of 25x25 spatial locations. Each spatial location contains a total of 10 neurons, 8 excitatory and 2 inhibitory, to match the size of the unstructured networks. The orientation preference of each cell has a salt-and-pepper arrangement in the 2D space, the external inputs are a Gaussian function of the difference between their preferred orientation and the orientation of the stimulus. The connectivity between two cells is proportional to the product of two Gaussian functions, one of the difference between the preferred orientation of the cells , W_{ori}, and another of their spatial location, W_S. As in the structure-less model described Fig. 4, external inputs target both inhibitory and excitatory cells while optogenetic inputs target only excitatory cells. **B** Structure of the monkey model. Idem (A) but the spatial arrangement of the orientation preference is given by an aperiodic map. Parameters of these fits can be found on Table 2. **C** Comparison between data (circles represent means, shaded regions are mean±sem) and network models (full lines with squares) as described in A for the mouse and in B for the monkey. In the mouse, all cells were taken into account in order to perform the averages, as was done in the data. For the monkey, only visual-stimulus-matched cells were used to fit the data.

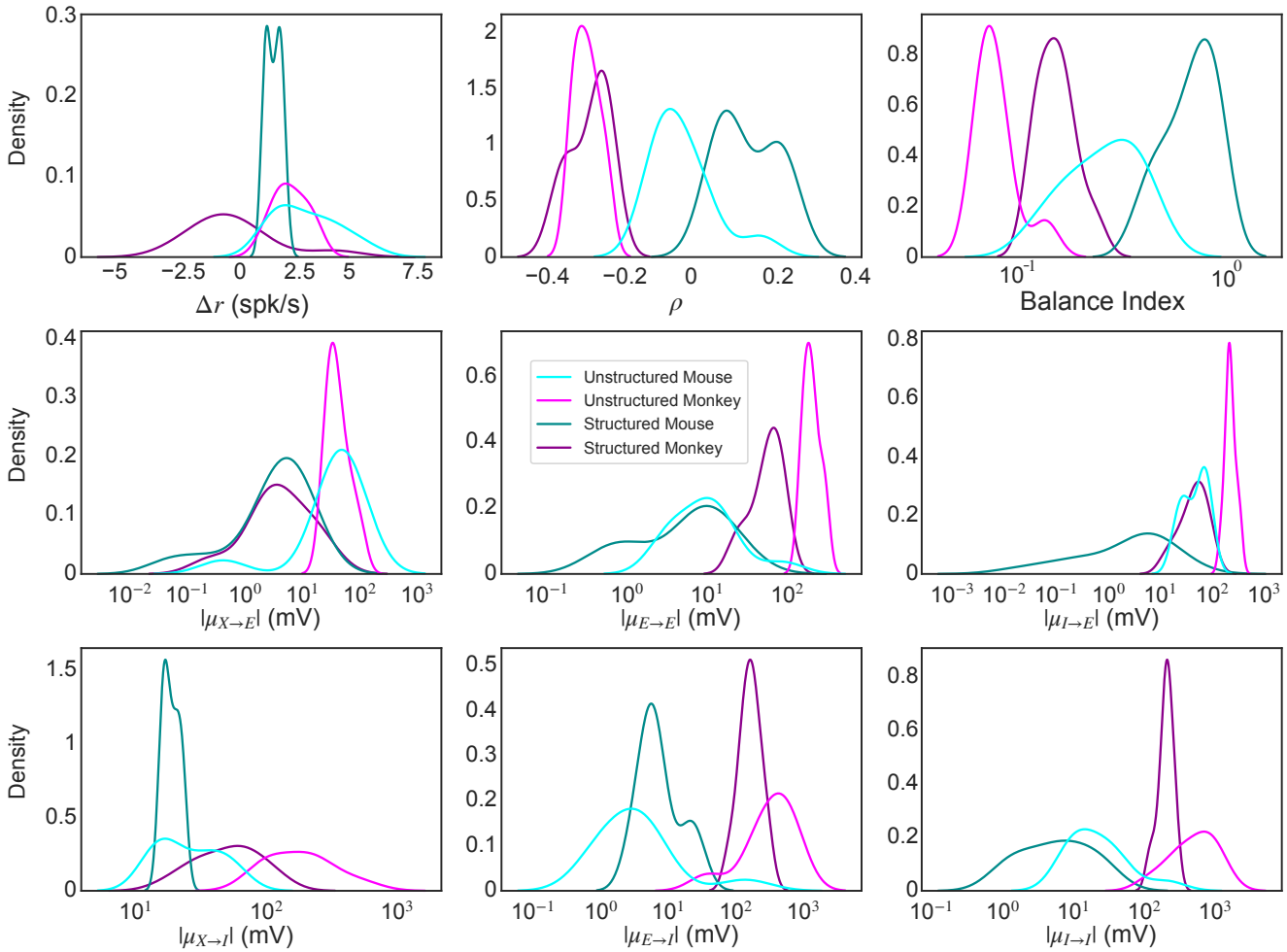

Figure S4: **Unstructured models require tighter balance to fit the data.** Kernel density estimation plots of Δr, normalized covariance ρ, balance index, and absolute mean inputs $\mu_{A\to B}$ from inputs of type A to cells of type B, for unstructured and structured models that best fit experimental data for each animal at the highest contrast. All quantities are calculated from the entire network except for structured monkey models, where we restrict our analysis to only visual-stimulus-matched cells as we did for the fitting procedure. Only structured models with quasi-periodic maps (fit to experimental results for the monkey) are found to have negative Δr, suggesting an alternative mechanism for the small change in mean response compared to unstructured networks. Structured networks tend to have looser balance (higher balance index) and smaller absolute mean inputs compared to unstructured networks fit to the same animal data (except for the $E \to I$ input for the mouse). Models were selected by taking all previously simulated optogenetic responses and repeating the fitting procedure used for the structured models in order to find model parameters that had the least squared error between simulated responses and experimental data at all contrasts. We then chose cutoffs for the squared error to select about 10 of the best models for each case, and plotted kernel density estimations using results at the highest contrast.

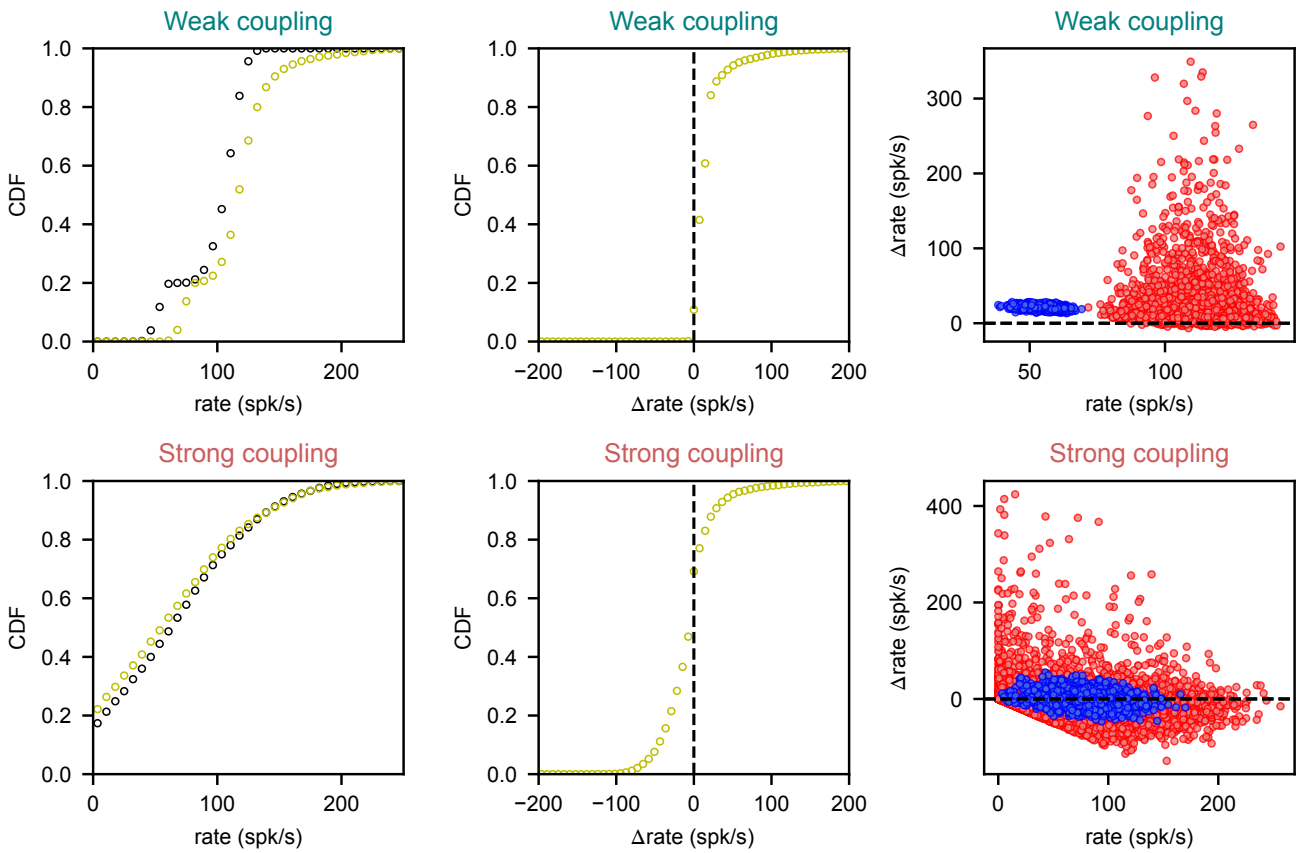

Figure S5: **Strong coupling is also necessary for reshuffling in structured networks.** Panels are arranged as in Figure 4 **B-G**, but instead show the firing rates of visual-stimulus-matched cells in a structured network with an quasi-periodic map. With weak coupling, optogenetic stimulation tends to excite most cells whereas with strong coupling, the distribution of rates is not significantly altered by optogenetic input. The normalized covariances ρ (for E, I, and E+I cells, respectively) are -0.03, -0.64, and -0.03 for the weakly coupled network and -0.28, -0.44, and -0.29 for the strongly coupled network. The Pearson correlation coefficients are -0.08, -0.32, and -0.04 for the weakly coupled network and -0.22, -0.26, and -0.22 for the strongly coupled network. All parameters except the synaptic efficacy J are taken from the best structured fit to the monkey data at the highest contrast (see Table 2). In the top row, the best fit synaptic efficacy is decreased by a factor of six, whereas in the bottom row the synaptic efficacy is taken from the best fit.