RESEARCH ARTICLE

Functional triplet motifs underlie accurate predictions of single-trial responses in populations of tuned and untuned V1 neurons

Joseph B. Dechery[1], Jason N. MacLean[1,2]*

1 Committee on Computational Neuroscience, University of Chicago, Chicago, Illinois, United States of America, **2** Department of Neurobiology, University of Chicago, Chicago, Illinois, United States of America

* jmaclean@uchicago.edu

Abstract

Visual stimuli evoke activity in visual cortical neuronal populations. Neuronal activity can be selectively modulated by particular visual stimulus parameters, such as the direction of a moving bar of light, resulting in well-defined trial averaged tuning properties. However, given any single stimulus parameter, a large number of neurons in visual cortex remain unmodulated, and the role of this untuned population is not well understood. Here, we use two-photon calcium imaging to record, in an unbiased manner, from large populations of layer 2/3 excitatory neurons in mouse primary visual cortex to describe co-varying activity on single trials in neuronal populations consisting of both tuned and untuned neurons. Specifically, we summarize pairwise covariability with an asymmetric partial correlation coefficient, allowing us to analyze the resultant population correlation structure, or functional network, with graph theory. Using the graph neighbors of a neuron, we find that the local population, including both tuned and untuned neurons, are able to predict individual neuron activity on a moment to moment basis, while also recapitulating tuning properties of tuned neurons. Variance explained in total population activity scales with the number of neurons imaged, demonstrating larger sample sizes are required to fully capture local network interactions. We also find that a specific functional triplet motif in the graph results in the best predictions, suggesting a signature of informative correlations in these populations. In summary, we show that unbiased sampling of the local population can explain single trial response variability as well as trial-averaged tuning properties in V1, and the ability to predict responses is tied to the occurrence of a functional triplet motif.

Author summary

V1 populations have historically been characterized by single cell response properties and pairwise co-variability. Many cells, however, do not show obvious dependencies to a given stimulus or behavioral task, and have consequently gone unanalyzed. We densely record from large V1 populations to measure how trial-to-trial response variability relates to

these previously understudied neurons. We find that individual neurons, regardless of response properties, are inextricably dependent on the population in which they are embedded. Specifically, patterns of correlations between groups of neurons, allow us to predict moment to moment activity in individual neurons. Only by studying large, local, populations simultaneously were we able to find an emergent property of this information. These results imply that understanding how the visual system operates with substantial trial-to-trial variability will necessitate a network perspective that accounts for both visual stimuli and activity in the local population.

Introduction

In the visual system, decades of research have probed stimulus parameters that evoke responses in single neurons and these responses have been generally trial-averaged [1]. These response properties have revealed principles of functional organization in primary visual cortex (V1), such as orientation columns [2], and canonical computations, such as divisive normalization [3]. However, responses are variable across trials [4], making the relationship between perceptual stability and neuronal single-trial stimulus representations unclear [5]. The fluctuations of response strength are not independent across neurons, and this shared variability impacts population level representations of visual stimuli [6,7]. Neurons are highly interconnected and connection likelihood is biased toward spatially proximal neurons [8], suggesting that trial-to-trial response variability may be in part the manifestation of the state of the surrounding neuronal population [9,10]. Pairwise interactions within a population can shape information representation [11,12,13] and can be regulated by top-down influences [14]. Therefore, comprehensive descriptions of stimulus representations in primary sensory cortex require a network perspective. Here, we used two-photon imaging to record from large populations of L2/3 excitatory neurons in mouse V1 to study effects of local population activity on trial-to-trial variability.

Understanding the sources and consequences of response variability is necessary to extend theories of sensory computation from the average case to single trials. Perception and behavior take place in real time, after all, so variable responses must be taken into account to understand stimulus representations in cortex. Shared variability in neural responses is commonly quantified by the set of pairwise correlations between neurons, and the structure of these correlations can have constructive or destructive effects on stimulus encoding in populations of neurons [15,16], highlighting the importance of its characterization. Moreover, complex patterns of population activity in retina can be captured by taking into account only neuron firing rates and pairwise correlations [17]. Covariability can also be shaped by cognitive properties such as attention in order to improve perceptual acuity [14]. Whether trial-to-trial variability is harnessed to improve the fidelity of sensory representations, or accounted for when decoding from noisy signals, the properties of response variability have a large impact on neural function.

Research on the correlation structure of population activity is still incomplete, however, and can be meaningfully expanded by incorporating a more comprehensive sampling of the network [18]. V1 populations consist of neurons whose activity is not modulated by, or is untuned to, a given stimulus. It is still an open question how this subpopulation contributes to neuron correlations. Two-photon imaging results in a relatively unbiased sampling of spatially proximal neurons including both tuned and untuned subpopulations. Neurons unrelated to a behavioral task can help predict activity in neighboring neurons in hippocampal CA1 [9]. In

V1, it has been shown that untuned neurons can help to decode the orientation of drifting gratings [19]. We investigate how co-fluctuations in the activity within tuned and untuned neurons interacts with responses to drifting grating stimuli.

We characterize population activity and correlations between tuned and untuned subpopulations in order to understand the relationship between single-cell response properties and recurrent network dynamics. Traditional noise correlation analyses study covariability independent from stimulus-driven activity [20]. However, untuned neurons have no stimulus modulation, so we use an analogous partial-correlation based method that additionally accounts for population-wide covariability. Fluctuations common across a local population are an important determinant of single-trial responses in mouse V1 [21,22]. Capturing this additional variable allows us to study the correlations in the entire unbiased sampling of the population. Additionally, the partial correlation matrices are asymmetric and relatively sparse, and can thus be represented as a weighted, directed graph. Graph theory analysis is used to summarize structure in complex networks and can resolve emergent properties resulting from pairwise relationships. Connectivity patterns, or motifs, in graphs can be characterized within small groups of neurons [23] or across an entire population [24,25]. Motifs have proved to be impactful for understanding complex biological processes including transcription networks [26] and spike propagation [23]. More generally, motifs patterns impact information representation in complex systems [27,28] and have increasingly been a subject of interest among neuroscientific disciplines.

From the graph neighbors of a given neuron, we can accurately predict activity on single trials using a simple, linear model. The local population contains information sufficient to predict trial-to-trial variability and recapitulates average tuning properties. Furthermore, neurons that are well-modeled by the activity of their neighbors have specific signatures of functional connection motifs. Across the entire graph, the most predictive motifs are also the most prevalent, suggesting that this structure is responsible for the overall quality of reconstruction observed. The triplet motif that facilitates predictions of neural activity may have a broad impact on information representation in graphs. Notably, total variance explained in a field of view scales with the number of neurons imaged, suggesting larger sample sizes are required to fully capture local network interactions.

Results

Response properties of V1 populations

To study interactions and response variability in local, cortical populations (<800μm diameter imaging plane), we imaged L2/3 excitatory neurons (72–347 neurons; 25–33 Hz; n = 8 animals; 23 distinct fields of view; Fig 1A) in mouse V1 during presentation of drifting gratings (Fig 1B and 1C). Square-wave gratings at 12 directions were presented in pseudo-random order for 5 seconds each, interleaved with 3 seconds of mean-luminance matched grey screen. We designed grating stimuli with slightly longer durations than many studies [16,29,30] for two reasons: first, to allow for the slow decay of the calcium indicator to fall to baseline in order to remove any confounds from the previous grating, and second, to study the sustained response in the population, rather than a transient response to stimulus onset [31]. Mice were awake and allowed to freely run on a linear treadmill. The majority of neurons showed significantly increased activity to one or more gratings over grey screen (3023/4535). Of the responsive subpopulation, most neurons were significantly tuned to orientation or direction (2073/3023; 540/3023 respectively). Neuron tuning was measured by fitting an asymmetric circular Gaussian tuning curve to the trial-averaged mean fluorescence in each grating direction (Fig 1D). These numbers of tuned and untuned neurons are in line with other population studies of
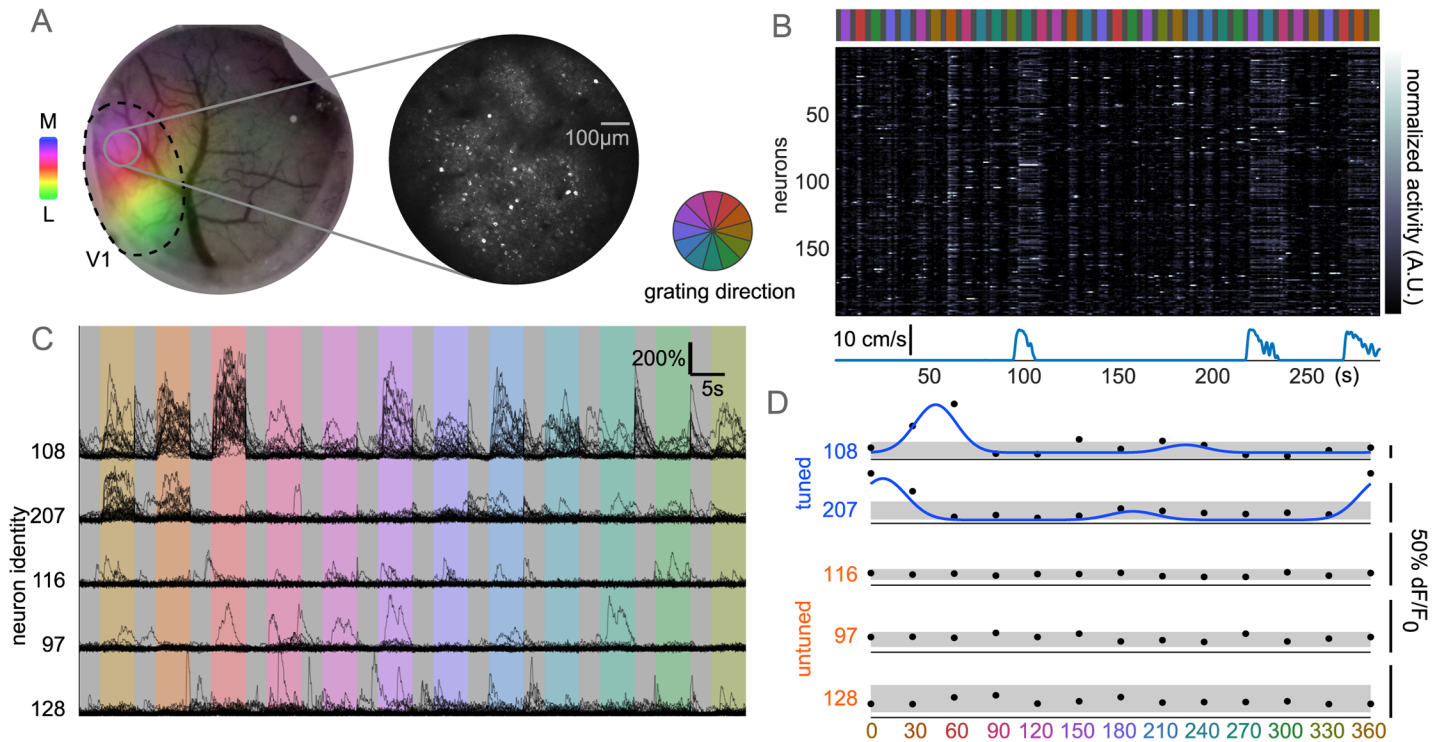
Fig 1. Categorizing response properties in V1 populations. (A) Left, anecdotal cranial window, V1 retinotopy mapped by intrinsic signal overlaid (colormap shows visual field azimuth, lateral to medial). Right, two-photon field of view L2/3 population from grey region on left. (B) One block of stimulus presentation (top; grey interleaved gratings) with simultaneous imaging data showing neuron normalized fluorescence over time (middle). Bouts of running (bottom) and gratings show coordinated increased activity across neurons. (C) Five example neurons' activity, black traces represent one trial, organized by grating direction. Grey period preceding gratings are kept together, with discontinuities at end of gratings. Top two neurons show increased activity to specific directions, others are untuned with substantial trial-to-trial variability. (D) Trial-averaged activity reveals classic tuning properties in a subpopulation of neurons. Tuned neurons fit with asymmetric circular Gaussian tuning curve (blue), untuned neurons show no increased activity relative to grey periods (grey bar mean+/-std average response).

https://doi.org/10.1371/journal.pcbi.1006153.g001

awake, mouse V1 [16,30]. In subsequent analyses, we pooled all visually responsive neurons without significant direction or orientation tuning into a class of 'untuned' neurons, differentiating two distinct subpopulations in V1 by their responsivity to drifting gratings.

Responses are highly variable across trials

Single trial responses to gratings showed a high degree of variability, even in strongly tuned neurons, manifesting as occasional strong responses to null-directions and weak or absent responses to preferred directions (Figs 1C and 2A). Tuning curves described average response strength of tuned neurons well ($0.70+/-0.20$ R^2). Responses across single trials in tuned neurons, however, were not well-described by their tuning curves. The mean fluorescence for each direction only explained a small fraction of the total trial-to-trial variance which we calculated by subtracting mean fluorescence in each direction from fluorescence in each trial (Fig 2B). This finding matches earlier results in awake, mouse V1 [30]. We computed the distribution of response strength (mean fluorescence across a grey or grating presentation) within single trials, z-scoring to account for neurons with different activity levels. Single trial response distributions were skewed, with most responses weaker than the mean (Fig 2C). Tuned neurons showed slightly stronger responses during gratings, as compared to untuned neurons which had nearly identical response distributions in grey and grating trials. Tuned response distributions were strongly overlapping, however, consistent with the hypothesis that individual neuron activity is not solely driven by tuning properties.
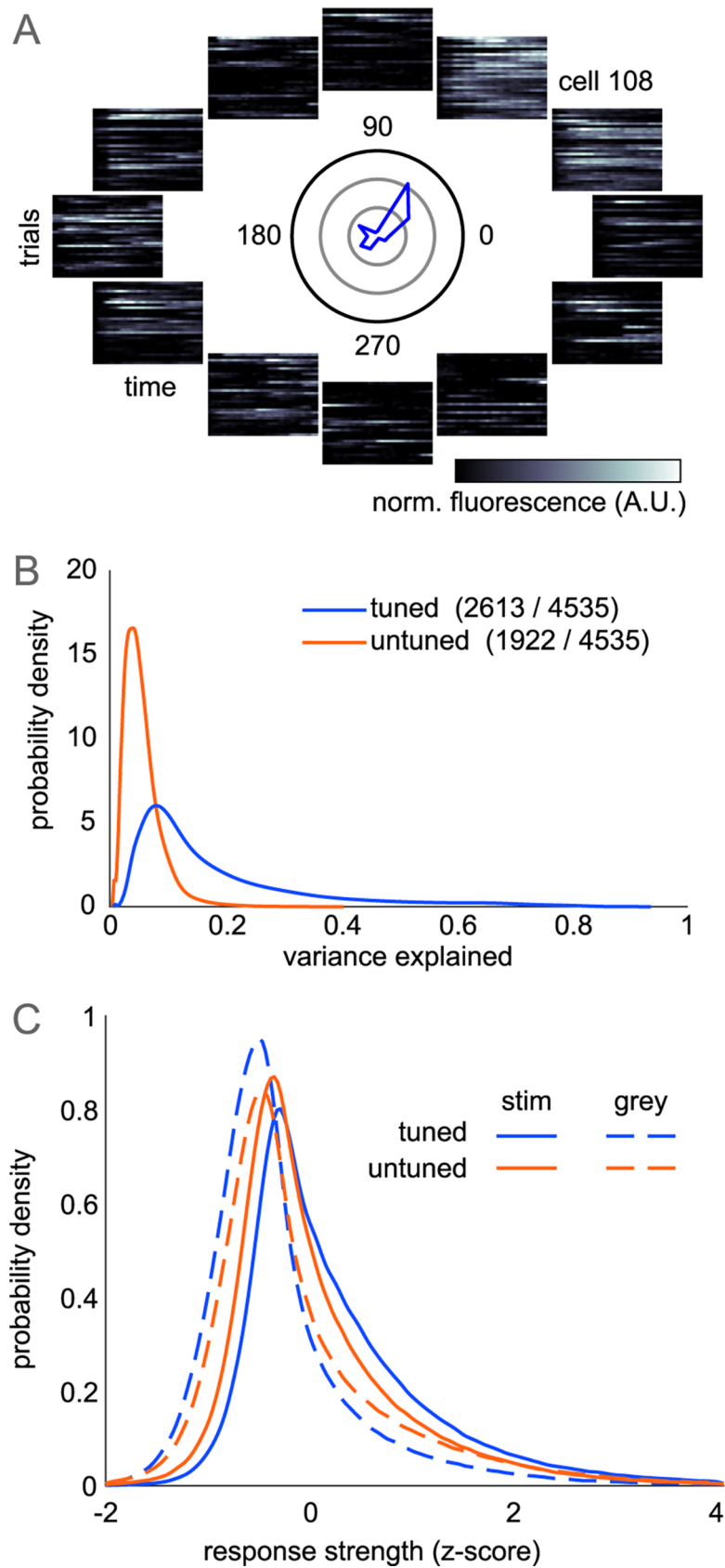
Fig 2. Single trial response variability. (A) Single trial responses of representative tuned neuron. Mean tuning curve in polar coordinates shown in center (rings = 100% dF/F$_0$). Each heatmap positioned by grating direction shows normalized time-varying activity for each trial. (B) Trial-averaged tuning responses do not explain trial-to-trial variability. Tuning average (blue curve in A) is subtracted from mean fluorescence in each trial to compute variance explained across trials. (C) Averaging across time, neurons show highly variable responses during grating and grey trials with a right-skewed distribution due to strong activity in a few trials. Tuned neurons show larger increase in response during gratings than untuned neurons.

To further describe population activity, we computed the time-varying activity during the presentation of a grating and its preceding and following grey presentations. For each trial, we removed neurons that were silent (defined as no fluorescence change 2*S.D. above baseline) and computed the time-varying z-scored fluorescence across neurons (Fig 3A). Despite the long duration of stimulus presentations, adaptation effects were minimal in these L2/3 neurons, as tuned neurons showed sustained activity throughout the 5 second stimulus presentation. Untuned neurons have weak modulation to the onset and offset of the stimulus and are equally active during the grey period. However, these effects are small in comparison to variability across trials, as indicated by strong overlap between the activity of the two subpopulations. In the awake animal, running speed is known to strongly influence spike rates [32], and we similarly observe that periods of high population activity are very likely to occur during periods of running (Fig 3B, anecdotally in 1B). However, mice did not preferentially run during grating or grey presentations (probability of running 9.7+/-7.7% during gratings; 10.0 +/-7.4% during greys; p = 0.278 paired t-test). While we used changes in fluorescence for all other analysis, we expanded our comparison of both subpopulations by estimating spike rates using a spike-inference from calcium fluorescence algorithm [33]. We found that untuned neurons exhibited an identical firing rate distribution to tuned neurons (Fig 3C). Therefore, differences between subpopulation dynamics cannot be explained by differences in firing-rates.

Correlation structure during drifting grating and grey screen trials

Untuned neurons are a large proportion of total neurons, exhibit similar spike rates to tuned neurons, and are likely to contribute to correlation structure in the population. To begin describing how dynamics are affected by stimuli in populations containing tuned and untuned neurons, we first analyzed pairwise correlations during grating and grey presentations. For each pair of neurons, we computed the correlation coefficient between the mean fluorescence (averaged over time) in either grating or grey trials. We did not remove signal-dependent responses, nor did we shuffle responses to eliminate simultaneous cofluctuations, therefore, these correlations are a combination of signal and noise correlations. Overall, within-subpopulation correlations are weak (0.014+/-0.027 tuned; 0.033+/-0.047 untuned), and beween-subpopulation activity is slightly anti-correlated (-0.019+/-0.017). Comparing mean pairwise correlations across all pairs according to their subpopulation, only within-subpopulation correlations are affected by grating stimuli, while between-subpopulation correlations are unchanged between stimulus and grey conditions (Fig 4A). Tuned neurons show a strong decrease in mean correlations during gratings, as seen in macaques [34]. Conversely, untuned neurons are more strongly correlated during gratings, and yet correlations between tuned and untuned neurons do not change in magnitude between stimulus and grey conditions. Untuned neuron activity is not directly modulated by the stimulus, so changes within this subpopulation most likely reflect changes in activity from the tuned subpopulation propagating through local synaptic connectivity. However, this occurs without a change in mean correlation between tuned and untuned neurons, bringing to question the mechanism involved. The pairwise
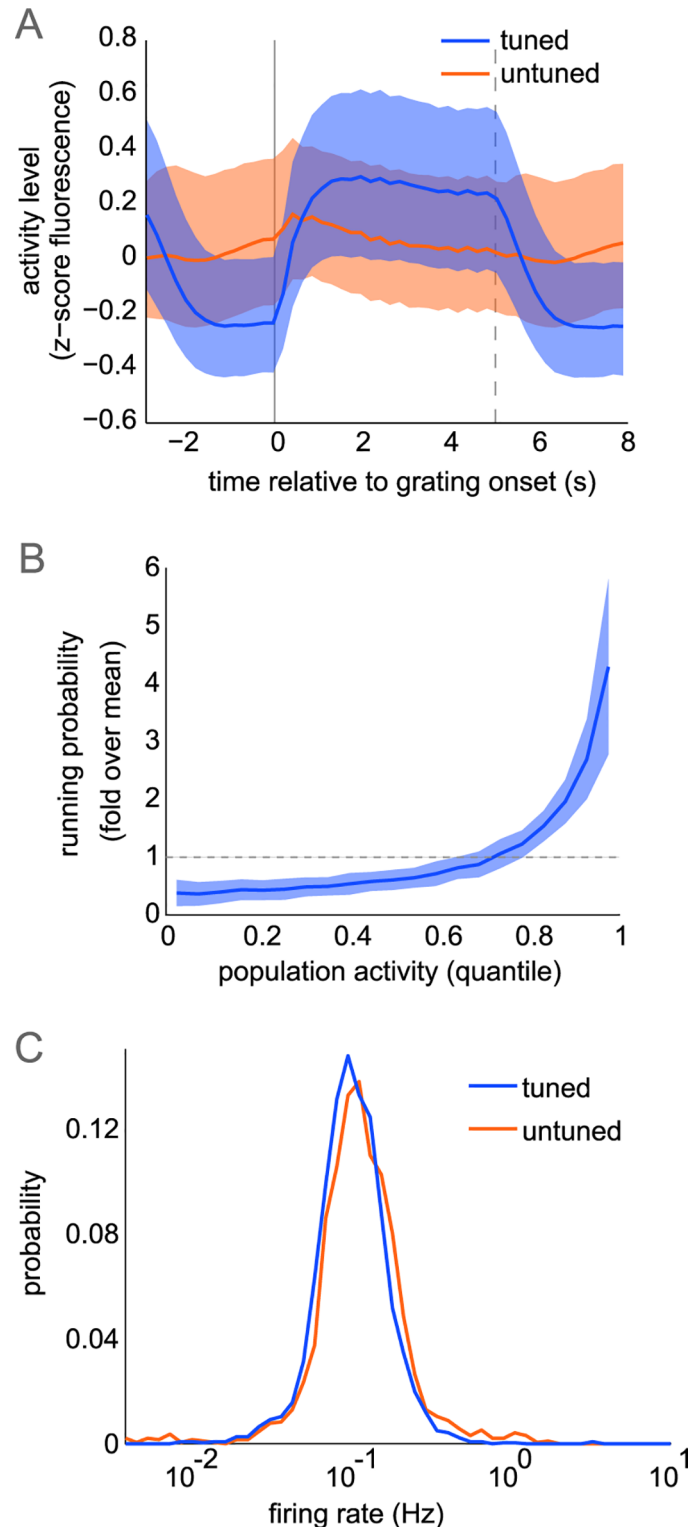
Fig 3. Population response properties. (A) Mean time-varying fluorescence of active neurons, pooled across datasets, during grating stimulus and preceding, following grey periods in tuned, untuned subpopulations (median+/-quartiles across pooled neurons). The 3-second grey period is shown before and after the grating response. Grating onset, offset shown in solid, dotted lines respectively. (B) Total population fluorescence split into equal quantiles to compute probability of running (speed>0.1cm/s). Plot shows mean+/-std across datasets (n = 23). (C) Inferred firing rate distributions computed across all imaging frames from spike trains inferred from fluorescence.

https://doi.org/10.1371/journal.pcbi.1006153.g003
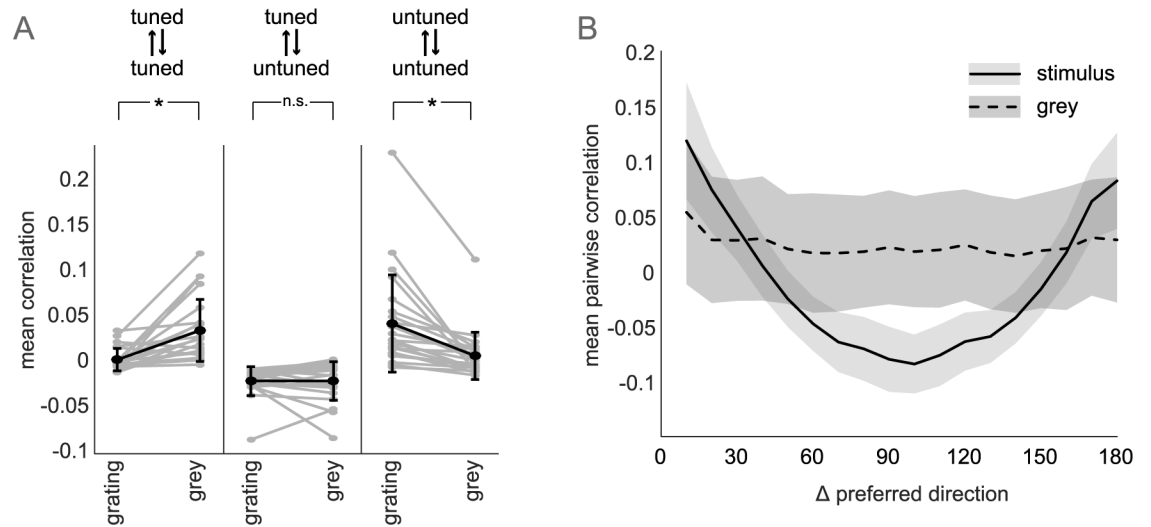
Fig 4. Correlations during grating and grey stimuli. (A) Mean activity correlation between stimulus conditions, within and between subpopulations (mean+/-std in black; individual datasets in grey). Significance determined with Tukey-Kramer corrected one-way ANOVA at alpha = 0.01 (tuned-tuned p = 0.004; tuned-untuned p = 0.999; untuned-untuned p = 0.001) (B) Correlations within grating and grey trials for tuned neurons (mean+/-variance pooling neurons with given difference in tuning over datasets).

https://doi.org/10.1371/journal.pcbi.1006153.g004

correlations in tuned neurons during the stimulus are a function of their preferred grating directions, as expected for signal correlations (Fig 4B). Similarly tuned neurons show strong correlations, while orthogonally tuned neurons show negative correlations. This structure is not present during activity in grey periods, however. This is surprising, because if local connectivity underlies correlations in the grey condition, one should expect the structure seen during gratings to remain in part because similarly tuned neurons are more likely to be connected [35].

Measuring trial-to-trial covariability

Overall correlations in populations including untuned neurons begins to reveal properties of local population activity, but in order to study the sources and structure of trial-to-trial shared variability, researchers attempt to remove the stimulus-dependent portion of responses leaving only variability, or 'noise' [20]. Correlated fluctuations between the remaining responses are therefore often called 'noise correlations.' However traditional noise correlation analysis was not appropriate in this case for the following three reasons: 1) the untuned neuronal subpopulation has no stimulus-dependent response; 2) in tuned neurons stimulus-driven response explains only a small portion of overall variability in tuned neurons (Fig 2B), making traditional noise-correlation analysis unsatisfactory for these neurons as well; 3) finally, V1 populations in mice, cat, and macaque are characterized by global cofluctuations common to every neuron [22,42,43] such as covariance driven by running (Figs 1B and 3B). Therefore in order to study pairwise noise correlations within and between subpopulations we used partial-correlation analysis that allowed us to account for stimulus-driven responses and population-wide co-fluctuations in both tuned and untuned neurons.

Visual stimuli were presented in 5-minute blocks. Each block of visual stimuli contained three repetitions of each direction in pseudo-random order and corresponding luminance matched grey periods. While the order of grating stimuli were pseudo random in each block the order was maintained between blocks. For each pair of neurons, the average activity across all remaining blocks, other than the block considered at that time, represented the stimulus-
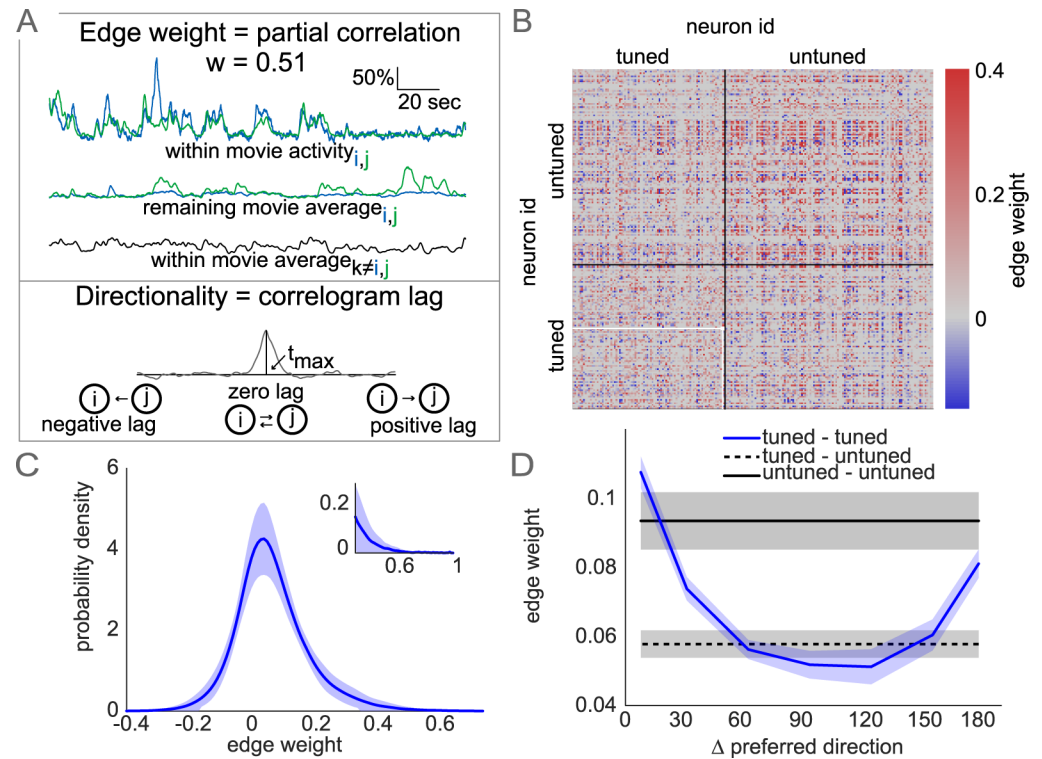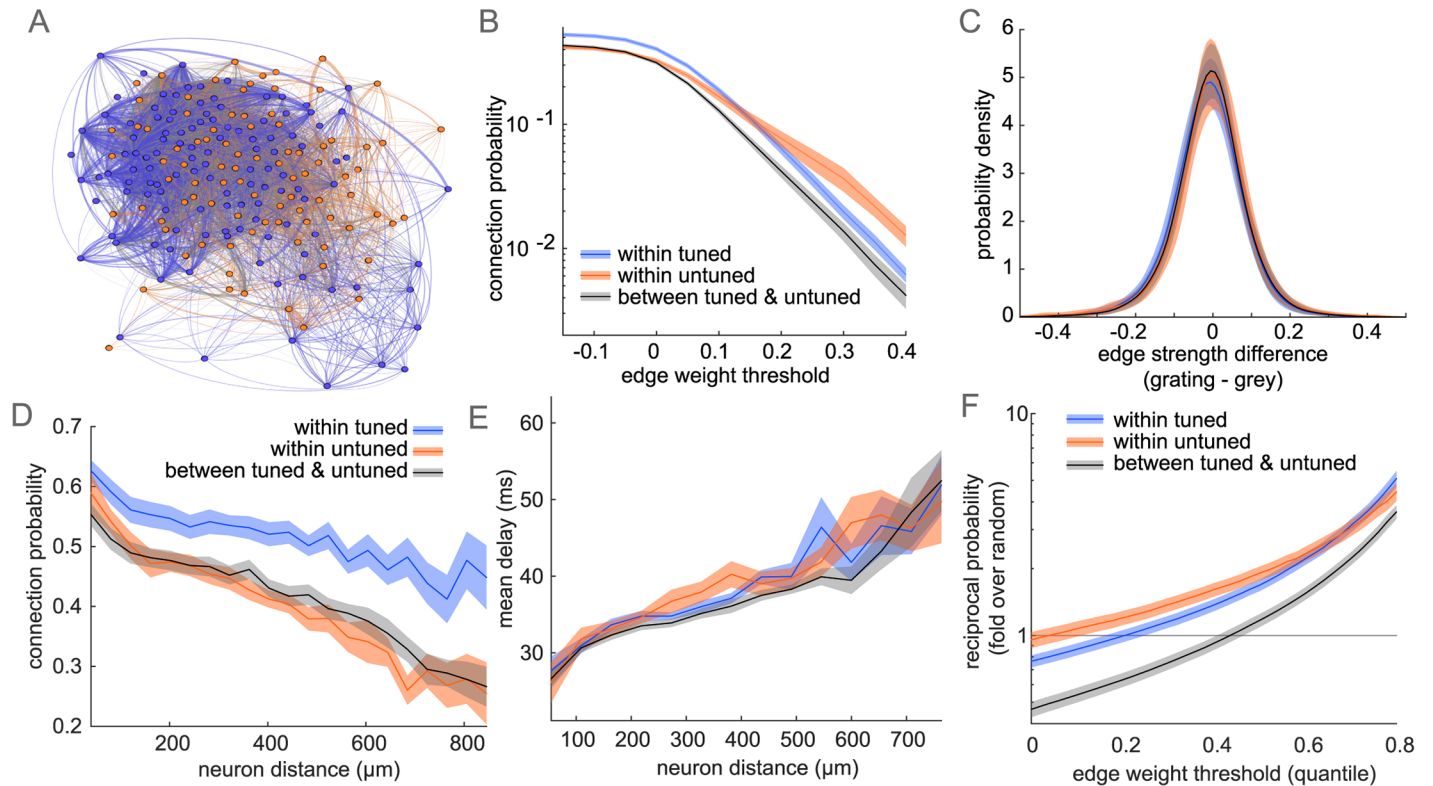
Fig 5. Partial correlation matrices representing trial-by-trial covariability. (A) Edge weight estimated by mean partial correlation. Fluorescence traces of a neuron pair (i, j) within one movie, and three variables accounted for in partial correlation: remaining-movie average activity of i, j and within-movie average activity of remaining population (top). Edge direction set by offset of peak (t_{max}) in neuron i, j correlogram (mean across movies). (B) Partial-correlation matrix sorted by subpopulation showing dense within- and between-population correlations. Neuron pair entry from (A) in white. (C) Kernel-density estimate distribution of all edge weights (mean+/-std across datasets). Inset shows right-tail of the strongest edges. (D) Tuned edge weights match overall stimulus correlations, despite accounting for stimulus response, and slightly weaker between-subpopulation edge weights (mean+/-sem over datasets).

dependent responses capturing tuning properties, when present. Additionally, we accounted for the mean within-block population-wide activity of all remaining neurons. Consequently, we were able to compute a partial correlation coefficient in each block between the activity of every pair of neurons, controlling for stimulus responses and population co-activity (Fig 5A). The mean partial-correlation across blocks is taken as the final correlation strength and entered as an edge weight into the functional connectivity matrix. We also added directionality to the partial-correlation by examining the mean cross-correlogram across blocks for the neuron pair. If the peak value occurs at lag 0 (i.e. within the same imaging frame), the edge was bidirectional, otherwise the edge was in the direction of positive lag. Lags greater than 500ms were thrown out and correlations set to zero. This resulted in a functional network described by a directed weight matrix. Though we interpret these partial-correlations as equivalent to noise correlations, the correlation matrices are different from traditional noise correlations in two important ways. First, many pairs of neuron correlations are exactly zero (51.7+/-7.2%), and second, non-zero correlations are asymmetric (Fig 5B). This allows us to analyze these matrices from a graph-theoretic perspective representing the functional partial correlations as a weighted, directed graph. Graph representations of pairwise edges allow us to analyze population-wide statistical features of the correlation structure. Overall, partial-correlation strengths, synonymous with edge weights, were long-tailed, centered slightly above zero (Fig 5C), similar to noise-correlations observed elsewhere [36]. As expected, tuned edge weights were similar to

Fig 6. Directed graph structure of partial-correlation matrices. (A) Example graph representation of matrix in (5B), nodes colored by subpopulation, edges colored by connection type as in (B-F), edge thickness represents absolute-value correlation strength. Blue nodes and edges indicate tuned neurons and their interconnections, orange untuned neurons and edges and grey indicate edges between tuned and untuned neurons. (B) Probability of connection by connection type for matrices with edges weaker than an increasing threshold set to zero. (C) Difference in edge strength between matrices using exclusively grating or grey frames. (D) Connection probability as a function of distance between neurons. (E) Mean lag of edges over neuron distance (F) Probability of a bidirectional edge at increasing thresholds, normalized by bidirectional probability assuming independence of edges. Plots in (B-F) show mean+/-sem across datasets.

https://doi.org/10.1371/journal.pcbi.1006153.g006

signal correlations [36,37], with similarly tuned neurons having larger edge weights on average (Fig 5D).

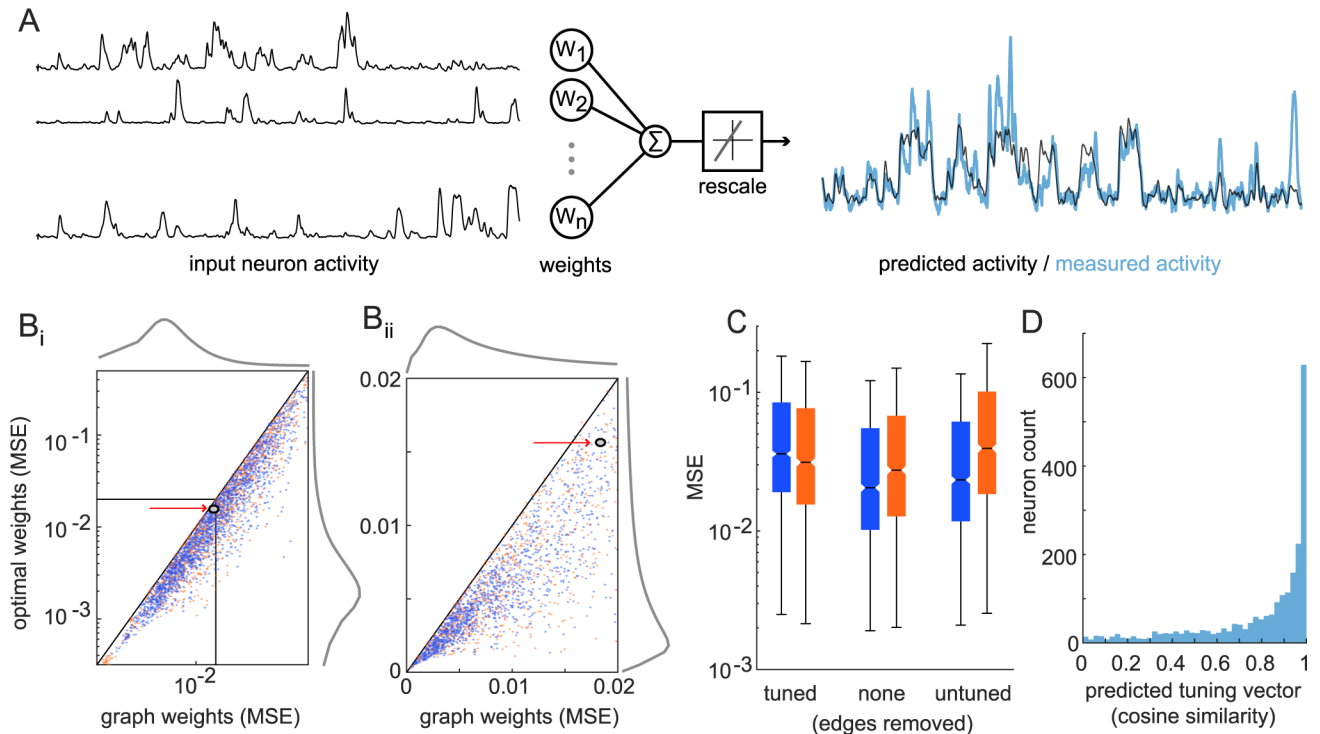## Graph structure of partial-correlations

We next analyzed the partial-correlations within and between tuned and untuned subpopulations. The graphs exhibit dense correlations with varying strengths among subpopulations (Fig 6A). To analyze biases in edge strengths, we thresholded the matrices at increasing values, setting all edges below each threshold to zero. Among all edges, within-tuned connections are more likely, while within-untuned and between-subpopulations are less common (Fig 6B; within-tuned 54.2+/-8.6%, within-untuned 44.1+/-6.0% between 43.9+/-6.2%). Between-subpopulation connections remain the least likely at higher thresholds, but among the strongest edges, within-untuned connections are the most likely. We then recomputed partial-correlation matrices, exclusively using frames during the grey condition or during grating condition to see how correlation strengths were affected, despite controlling for mean stimulus-dependent activity. The two resulting matrices did not have significantly different edge strength (stimulus 0.12+/-.02; grey 0.12+/-.03; p = 0.75 one-way ANOVA). However, probability of connection was higher within grey frames (grating 45.3+/-6.5%; grey 58.0+/-8.2%; p = 5.8*10^-7 one-way ANOVA), indicating a higher degree of interconnectivity in the population in the absence of stimuli. When analyzing magnitude of edge strength differences between grey and grating matrices, we found

that all connection types changed similarly with a mean near zero (Fig 6C; within-tuned -0.011 +/-0.097; within-untuned -0.005+/-0.103; between -0.008+/-0.093).

Since population dynamics are partly constrained by synaptic connectivity [38], we evaluated whether there was a spatial component to the weight matrices. Edge probability fell monotonically with distance between neurons, similar to traditional noise correlations [36] and synaptic connections [8,39]. Notably, the decay in connection probability is slower within tuned neurons compared to within untuned neurons (Fig 6D). Between-population decay lies in the middle. If the spatial structure of untuned correlations is exclusively driven by local connectivity, then bottom-up sensory drive is the most likely source for the longer range of functional correlations among tuned neurons. Furthermore, the mean lag (delay of the cross-correlogram peak used to determine edge directionality) is greater over longer distances and accumulated evenly across subpopulation connection type (Fig 6E). Assuming a linear change in lag over space, these data suggest the speed of functional correlations in this preparation is roughly 25 mm/s. Despite allowing for directional edges, nearly half of all edges were bidirectional (42.5+/-16.8%). As a function of edge strength, bidirectional edges are more prominent within-subpopulations, and are strongly biased toward the strongest weights (Fig 6F). Thresholding at increasing edge strengths sparsifies the matrices, so we normalized the bidirectional edge counts by the probability of bidirectional edges assuming connections are placed randomly. Among all edges, bidirectional edges occur less often than random. The strongest edges, however, are roughly 5 times more likely to be bidirectional than random. Overall, zero-lag connections are less frequent between tuned and untuned neurons, suggesting a transmission or propagation of information, rather than simultaneous representation of the same information between subpopulations.

Modeling trial-to-trial variability from local population activity

Because trial-averaged tuning poorly captures trial-to-trial responses we asked whether information in the local population could better explain V1 trial-to-trial response variability. Pairwise correlations have been shown to capture a significant portion of the complexity in population activity [17]. We tested whether the activity of a neuron could be modeled from the activity of its neighbors that had a non-zero correlation. Since correlation coefficients capture the linear relationship between neuron coactivity, we used a simple linear combination of the input neuron activity with partial-correlation coefficients (edge strengths) as weights. This model gave a time-varying prediction of the fluorescence of a given neuron, which was then rescaled by an offset and a gain to account for different numbers of input neurons (Fig 7A). In many cases, this model resulted in highly accurate predictions of activity. Mean squared error of the reconstruction was small, and often near optimal compared to weights estimated by regression (Fig 7B). Tuned neurons were slightly better-modeled on average than untuned neurons (tuned MSE 0.014 median, 0.037 inter-quartile range; untuned MSE 0.017 median, 0.046 inter-quartile range), possibly because of the additional stimulus-dependent information captured by other tuned inputs. The ratio of MSE to mean-squared fluorescence from the true trace, subtracted from one, gives the percent variance explained of the model for each neuron. The optimal reconstructions from regression explained 77.2+/-15.2% of the variance in activity across all neurons (n = 4531). Using partial-correlation coefficients as weights performed near this upper-bound, at 65.8+/-17.1%. We tested the robustness of the partial-correlation based predictions modeling 5 minutes of fluorescence change, corresponding to one block, using a functional network built using 45 min, corresponding to the nine other blocks, of non-overlapping recordings to recompute the weights. We note that one block, i.e. five minutes of stimulus epochs, corresponded to three stimuli at each of the 12 directions and interleaved grey

Fig 7. Modeling time-varying fluorescence from partial-correlation graphs. (A) Illustration showing linear combination of incoming-edge neighbors fluorescence traces and rescaling to predict a given neuron's activity. (Bi) Reconstruction quality across tuned and untuned neurons, quantified by mean-squared error (MSE) of reconstruction, compared to MSE using optimal weights (marginals in grey; arrows show neuron from (A)) (Bii) zoom in on neurons with best reconstruction. (C) Increase in MSE when selectively removing within- or between-subpopulation connections (boxplots show median, quartiles, with whiskers extending to most extreme value within 1.5*inter-quartile range outside quartiles). (D) Similarity of mean tuning vector between tuned neuron and reconstructed average activity.

periods. We then tested our predictions on the left-out 5 minute dataset (repeated to leave out each dataset once). On this cross-validation procedure, the average performance on the left out epoch reached 55.4+/-18.9% variance explained.

We selectively removed either tuned or untuned input neurons to evaluate the relative contributions to model predictions. The accuracy of prediction decreased more when removing within-subpopulation inputs as compared to removing between-subpopulation inputs, but the decrease is small and distributions across neurons strongly overlap (Fig 7C). Within-subpopulation inputs are more frequent, however, so removing a larger fraction of inputs should be expected to have a larger impact on our ability to model neuronal fluorescence. On a single neuron basis, correlations between tuned and untuned neurons both contribute to predicting time-varying activity. We next asked whether our model based on local population activity could also predict trial-averaged tuning properties. For tuned neurons, we used the modeled fluorescence traces to recompute the mean fluorescence in each grating direction. The average responses in direction space were added together to obtain a mean tuning vector. The modeled fluorescence had very similar tuning properties to the data, as measured by the cosine similarity between model and data tuning vectors (Fig 7D). The modeled activity was constructed from a neuron's input edges; we also asked if a neuron's outgoing edges were related to tuning properties, which may indicate neurons that are good at decoding the stimulus (strong tuning) were also good at decoding activity in its neighboring neurons (strong outgoing edges). However, there was nearly zero correlation between the tuning strength of tuned neurons and its mean outgoing connection strength ($r = 0.04$, $p = 0.02$). Local population activity therefore contains information to capture

trial-to-trial variability as well as trial-average stimulus response properties, but a neuron's dependence on the stimulus does not affect its local population correlations.

Pairwise and motif contributions to model performance

To investigate how individual neurons inputs contribute to reconstruction of activity, we selectively removed input neurons based on their edge strength and measured the increase in reconstruction error (MSE), normalized by total mean squared fluorescence in the neuron. As expected, the strongest edges contribute the most to activity reconstruction with the strongest 25% of weights containing over half of the reconstruction capability, whereas removing half of the weakest edges had no discernible effect (Fig 8A). Interestingly, randomly removing edges does not linearly reduce prediction performance. This suggests a level of redundancy in the predictive information of input neuron activity. Accounting for the cumulative weight removed, we still found worse performance when removing the strongest edges first, and removing half of the total weight using only the weakest edges has minimal effect on reconstruction error (Fig 8B). Normalizing for the total weight removed reveals a nearly-linear increase in reconstruction error when removing the strongest weights, suggesting that these neurons may hold independent information from the remaining input pool.

To address the possibility of synergistic or redundant information between input neurons, we analyzed connections between triplets of neurons termed 'motifs' in graph theory literature. Triplet connection motifs are built up from pairwise connectivity and collectively represent higher-order connectivity patterns that cannot be captured by individual edges, and can have strong implications for computation and information propagation within graphs [27]. We looked at the clustering of triplet motifs for each type of triangle that can be formed with directed edges (Fig 8C) [40]. Clustering is a measure of how many motifs are present among a neuron's neighbors given its input and output connections. In comparison to Erdos-Reyni graphs, which have uniform, low levels of clustering across motifs, cycles of edges in the data are less clustered on average, with all other motifs showing elevated clustering. The middle-man motif shows the strongest clustering. Similar results have been found in activity generated by simulated and ex vivo neural networks, although fan-in clustering was higher than middle-man [23]. To address the possibility that triplet motifs are responsible for explaining more of the neuronal response than pairwise edges alone, we analyzed the relationship between motif clustering in single neurons and the variance explained by the predicted activity. Because the magnitude of clustering is different across motifs, we first z-scored clustering coefficients, then computed the mean across neurons with different levels of variance explained. Neurons with the best reconstruction showed higher clustering of middle-man motifs and lower cycle clustering, relative to other neurons in the population (Fig 8D). Total clustering, as well as fan-in and fan-out, had weak, positive correlation with variance explained. Together, these relationships map directly onto the overall prevalence of the graph motifs, suggesting that the graph structure has an important function in representing population information.

Because the motif with the highest mean clustering was also most indicative of model performance, we hypothesized that the partial-correlation structure might underlie our ability to predict neuron activity from its local population. Interestingly, across fields of view, the total variance explained in the population increased with number of neurons imaged (Fig 8E; r = 0.58). To compute population variance explained, we took the sum of all squared prediction errors (residuals) across neurons in a field of view, divided by the sum of the squared population fluorescence activity, and subtracted from one. This improvement of variance explained across the population with more recorded neurons suggests that, in addition to motifs, total neurons sampled in the population determines our ability to measure a neurons' single trial dependence on its local population.
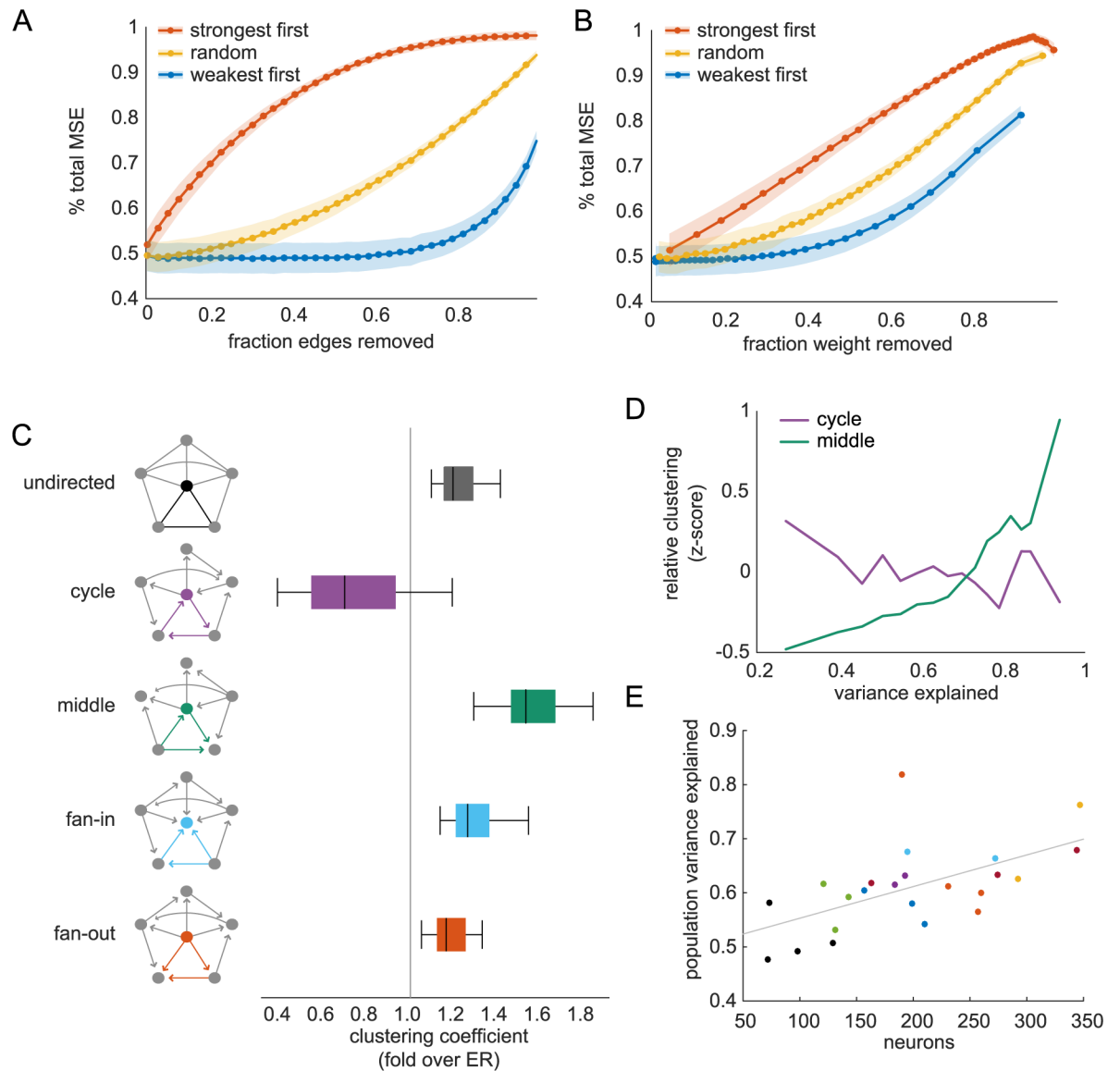
Fig 8. Pairwise and triplet motif contribution to model performance. (A) Increase in mean-squared error (MSE) by number of edges removed, normalized by neuron total MSE (mean+/-variance across all neurons). (B) Increase in MSE by cumulative weight removed (mean+/-variance). (C) Clustering of triplet motifs in graphs (illustrated in left column) for each type of directed triplet motif. Box-plots as in Fig 7. Mean clustering was normalized by sparsity-matched Erdos-Reyni graphs. (D) Mean clustering coefficient, z-scored by mean/std for each motif type in (C), at different levels of model prediction showing neurons with best reconstruction had highest middle-man clustering. (E) Total population variance explained for each dataset (colored by mouse identity) shows better performance with more neurons imaged (linear trend in grey).

Moreover, the linear trend had no discernible plateau, so representing the local population may require recording from more than 300 neurons simultaneously.

Discussion

Summary and correspondence with previous findings

We sought to describe the interrelationships within local populations of V1 neurons, including tuned and untuned neurons, as they relate to single-trial responses to grating stimuli. We used

two-photon imaging to record from L2/3 excitatory populations constitutively expressing calcium indicator GCaMP6s. Neurons with similar response properties showed stronger co-variability on average, but across the entire population there was a broad distribution of correlations driven, presumably, by a confluence of sensory drive and activity in the local population. The functional correlations in the recorded populations were sufficient to predict activity in individual neurons, far surpassing predictions from tuning characteristics alone. The dependence on local population activity reinforces theories of layer 2/3 acting under strong modulation with sparse activity and weaker dependence on sensory drive than layer 4 [41]. We summarized the structure of correlations as directed, sparse matrices in order to analyze population dynamics from a graph theoretic perspective. We demonstrate that a simple population model capable of predicting single-trial neural responses is also able to accurately predict trial-averaged tuning responses, a key feature of V1 function. We found that the prevalence a specific triplet connectivity motif, built up from pairwise correlations, corresponded with our ability to predict activity on single-trials. This result could not have been observed from only studying pairwise correlations and motivates the continued use of graph theory to study neural population dynamics.

The single-neuron response properties in our data replicate imaging and electrophysiological results in awake recordings of V1 including proportion of tuned neurons [29,30,36], and variance explained by the mean tuning curve [30]. Both tuned and untuned neurons have low firing rates on average, though estimation of firing rates from calcium imaging is not always straightforward. We note that single neuron properties, including firing rate and trial-to-trial variance, can change substantially between animal models and in different states of anesthesia [42,43].

Population structure of pairwise correlations

We found that the magnitude of signal correlations between tuned and untuned subpopulations do not change between grey and grating stimulus conditions, while within subpopulation correlations do change. Perhaps ubiquitous changes would be more expected, or exclusive changes among neurons modulated by the stimulus. Within subpopulation correlations change in opposite directions, however, and could serve as a mechanism to balance changes between the subpopulations.

We found that the spatial organization of the network is a strong determinant of correlation structure with correlations decaying over distance, consistent with paired patch clamp recordings [8] and the correlation structure of activity in isolated preparations [44]. Correlation matrices were computed as an asymmetric partial correlation coefficient, accounting for stimulus and population effects, to allow the incorporation of untuned neurons into traditional noise correlation analyses. While this approach differs from standard noise correlation estimates, the magnitude of correlations and dependence on tuning similarity replicate previous results measuring noise correlations [36]. For this reason, we interpret the partial correlations as measuring trial-to-trial covariability, and as being broadly equivalent to noise correlations. Tuned neurons, which combine feedforward sensory inputs with recurrent inputs, show a slower spatial decay of correlational values than untuned neurons, the latter of which presumably are driven more exclusively by recurrent, local inputs. These data show than subdividing the population by their response to grating stimuli can differentiate rates of spatial decay within the network.

Noise correlations are hypothesized to be driven in part by shared synaptic input [45]. Spatial dependencies of feedforward and recurrent connectivity can qualitatively change the spiking activity in network models [46]. The relatively small diameter of the fields of view imaged here (<1mm vs 10mm) does not allow us to differentiate between the two modes of activity predicted by these models. These data suggest, however, that functional recurrent correlations

have less spatial extent than feed-forward functional correlations. To our knowledge, analogous synaptic connectivity estimates do not exist for mouse V1.

In addition to spatial structure of correlations, we found an increase in mean temporal delay of correlations over distance. This delay spans roughly 20–50 ms in our field of view and is therefore likely to reflect timescales of functional correlations rather than monosynaptic transmission delays. We have previously found that functional correlations are indicative of synaptic connections, in some cases, when considering synaptic integration rather than synaptic delays [23]. The implied speed of this delay accumulation is much faster than propagating LFP waves observed in macaque M1 [47] after normalizing for total cortical surface area [48]. The propagation of beta-oscillations and temporal delays of functional correlations likely have different underlying mechanisms, which may explain the different speeds.

Prediction of single-trial responses from pairwise correlations

We were able to extract a substantial amount of predictive information from the local population using a simple, linear model. Other approaches have successfully predicted single-trial responses from ongoing population activity at a larger spatial scale, averaging over many neurons in a population [49]. In contrast, we use a large, unbiased sampling of the local population with single cell resolution to predict variability on single trials. Alternative models incorporating known neuronal nonlinearities [50], or more sophisticated or biologically-relevant predictive models [51], may have improved performance. Consequently, the increase in total variance explained with population size may show different trends with alternative models. Nevertheless, we chose this modeling approach to maintain ease of interpretation and utilize the linear correlation coefficients estimated from the data in a straightforward manner. The linear models had good performance, and allowed us to remove input neurons without laboriously re-training the models. This straightforward modeling approach may not capture all of the information present in the local population, but its performance sets a lower bound on predictive information in the population. A small subset of neurons have high mean-squared error of predictions, and we speculate that we have not sampled enough of the population to predict these neurons given that total variance explained scales with population size. Neuron activity was only predicted from the in-degree neurons, rather than the entire population, reducing the number of parameters by 52+/-22% across neurons, though in either case, parameters scale by order N. These results substantiate previous models showing that the collection of pairwise relationships in large populations can explain complex activity patterns [17]. The predictions of single-trial activity that are obtained from local activity are, on one hand, a description of inter-neuronal dependencies in recurrent networks, and on the other hand, are capable of recapitulating trial-averaged tuning properties, extending these dependencies to stimulus encoding. We further explored this idea by identifying a specific motif of pairwise correlations underlying accurate predictions of neuron activity on single-trials.

Functional consequences of triplet motifs

The middle-man motif underlying the most accurate predictions is a specific pattern of pairwise correlations and represents a higher order feature of covariability. Local populations have been shown to contain such high-order correlations, but were not seen at larger spatial scales [52,53]; further experimentation is necessary to test whether functional triplet motifs occur across thousands of microns. The finding that middle-man motifs underlie the best predictions of activity may be initially surprising from the connection pattern of the motif. Compared to the fan-in motif, which has two input connections, the middle-man has one input and one output connection, and acts to route connections from its input to its target. However,

the motifs were quantified using the clustering coefficient, which normalizes motif count by the total number of possible motifs (i.e. high fan-in clustering doesn't correspond to high in-degree). The functional significance of any given triangle motif clustering is unknown and is likely dependent on the underlying system represented by the graph. In a neural network, middle-man clustering may indicate a hub-like property common in the brain [54], having a combination of convergent and divergent correlations. The cycle motif similarly has a combination of convergence and divergence, yet its clustering has a negative effect on prediction accuracy. The difference in the two motifs is the direction of connections between neighbors. In these networks, it may be that cycle clustering reflects recurrent, redundant correlations reducing our ability to predict activity from the population. Conversely, the middle-man motif is isometric to fan-in and fan-out motifs, and could allow for transfer of information between motifs, and in turn increase predictive power. Finally, this result demonstrates that in addition to providing insights into synaptic mechanisms underlying dynamics [23], network science can also provide insights into predictions of single trial neural responses as we have demonstrated here.

V1 is the first stage in which visual information is encoded in densely recurrent cortical networks. Thus, in order to understand activity patterns in V1, one must take into account visual drive as well as local network activity. We have provided a quantitative comparison of the relative influences of these two factors in awake, ambulating mice. Local network effects dominate on single trials, highlighting the importance of investigating cortical computation from a population perspective in order to understand how information is encoded in single trials. Populations of neurons exhibit emergent properties beyond the sum of their individual neurons and connections, and we use the analytic framework of graph theory to begin unraveling this emergent structure.

Materials and methods

Animals and surgery

All procedures were performed in accordance and approved by the Institutional Animal Care and Use Committee at the University of Chicago. Data was collected from C57BL/6J mice of either sex (n = 4 female; 4 male) expressing transgene Tg(Thy1-GCaMP6s)GP4.12Dkim (Jackson Laboratory) between ages P84 –P191. After induction of anesthesia with isoflurane (induction at 4%, maintenance at 1–1.5%), a 3mm diameter cranial window was implanted above V1 by stereotaxic coordinates and cemented in place alongside a custom titanium headbar. Mice recovered for at least 8 days before intrinsic signal imaging to identify V1 followed by two-photon data collection.

Intrinsic signal V1 identification

Boundaries of V1 were identified by intrinsic signal imaging post-surgery [55] (Fig 1A, left]. Mice were anesthetized with isoflurane and head-clamped under a CCD camera (Qimaging Retiga-SRV). A vertical white-bar stimulus (100% contrast, 0.125Hz) was repeatedly presented on an LED monitor (AOC G2460) approximately 20cm from the contralateral eye while capturing cortical reflectance under 625nm illumination. The retinotopic mapping of V1 was then estimated at each pixel from the phase of peak reflectance driven by increases in activity-dependent blood flow.

Data collection

Two-photon imaging was collected from awake, head-fixed mice on a linear treadmill. Running speed was measured with a rotary encoder attached to the treadmill axle. A L2/3 field of view (roughly 800µm diameter) in V1 was identified with galvanometer-mirror raster

scanning (Cambridge Technologies; 6215H). Once a suitable field of view was found, raster scans (1Hz) were continuously acquired for roughly 10 minutes alongside visual stimulation. Neurons (n = 72–347 per field of view) were then automatically identified using custom image processing software for imaging during visual stimulation using Heuristically Optimal Path Scanning [56] at 25–33 Hz). All imaging was performed at 910nm (Coherent; Chameleon Ultra) with a 20X 1.1NA Olympus objective and GaAsP PMT (Hamamatsu; H10770A-40). Field of view size was estimated by fitting circles to a single raster scan of 15µm fluorescent microbeads and used for each dataset, though true field of view size may vary up to 8% across datasets from realignment of laser beam path.

Visual presentation

Drifting grating stimuli were presented on an ASUS VG248QE, 20cm from the contralateral eye at 60Hz; 60cd/m2. The mean luminance was measured and gamma correction was performed and confirmed using a luminance meter. Square-wave gratings were shown at 80% contrast, 2Hz, 0.04 cyc/deg at 12 evenly spaced directions. Gratings were presented for 5s, interleaved with 3s mean-luminance grey screen. Three repetitions of each orientation were presented in a pseudo-random order, resulting in a roughly 5min stimulus movie. The grating order was preserved between movie presentations, and mice were shown 8–11 repetitions of the movie (24–33 repetitions of each direction).

Data acquisition and pre-processing

Stimulus presentation was monitored with a photodiode (Thorlabs) and synchronized with running speed and imaging frames at 2kHz. For each neuron, baseline fluorescence was estimated from raw fluorescence by thresholding to eliminate spike-induced fluorescence transients and smoothed with a 4th-order, 81-point Savitzky-Golay filter. Fluorescence time-series were then normalized to percent change from baseline (dF/F_0) using this time-varying baseline. Spike inference from fluorescence traces was performed using the OASIS algorithm [33] implemented using software made freely available (github.com/j-friedrich/OASIS). Inference outputs probability of spiking at each time point. As commonly done [57], probabilities were thresholded to obtain a binary spike train.

Response properties and tuning classification

Neurons were classified as visually responsive if the mean response to any grating was significantly greater than the mean response across all grey periods by Dunnett-corrected one-way ANOVA (alpha = 0.01). In these analyses, each trial response is the mean fluorescence across the entire grating presentation (5000ms), or the last half of the grey presentation (1500ms) to allow for fluorescence from grating responses to decay. Responsive neurons were then tested for statistically significant orientation- or direction-tuned responses according to the trial vectors in orientation or direction spaces ([58] for more detailed methods]. For significantly tuned responses, tuning curves were then fit with an asymmetric-circular Gaussian to significantly tuned neurons. Tuning curve parameters (baseline, tuning width, peak amplitudes, and preferred direction) were fit repeatedly using randomized initial conditions. The parameter set that minimized mean-squared error was maintained.

Signal correlations

For each neuron, the mean responses in each trial (using the same time windows as for tuning classification) for either stimulus or grey trials were used as response vectors. Typically, these

response vectors had 360 elements (12 directions, 30 trials each). The correlation coefficient between each pair of these vectors was used to compute a pairwise correlation matrix for the grating and grey conditions. We did not shuffle responses, therefore these values measure the combination of signal and noise correlations.

Partial correlation matrices

For each pair of neurons, pairwise-correlation was computed as the mean partial correlation between their fluorescence across movies while accounting for three variables. This was computed with a built-in MATLAB function (partialcorr.m). We computed these correlations on time-varying traces of activity, rather than time-averaged trial activity as done in previous correlation analysis. Controlling for a single variable can be computed with the following equation, where $r_{x,y}$ is the correlation coefficient between time-varying fluorescence traces $x(t)$, $y(t)$, and $r_{x,y|z}$ denotes the partial correlation between $x(t)$ and $y(t)$, controlling for $z(t)$:

$$r_{x,y \lor z} = \frac{r_{x,y} - r_{x,z} r_{y,z}}{\sqrt{1 - r_{x,z}^2} \sqrt{1 - r_{y,z}^2}}$$

Here, $x(t)$ is the fluorescence trace of the 'input' neuron, $y(t)$ is the fluorescence trace of the 'output' neuron, and $z(t)$ represents the three control fluorescence traces. Controlling for more than one variable can be achieved by successive iterations of this procedure. This was computed for each movie, or 5-minute block of grating presentations that was repeated in each experiment. The first two control variables are the mean response of the two neurons in all other movies, accounting for the stimulus-driven responses. This is similar to normalizing by the mean response as in traditional noise correlation estimates. The third control variable is the within-movie mean fluorescence of all other neurons and was included to control for population-wide covariability, for example running speed effects. Furthermore, the cross-correlogram between the two neurons' fluorescence traces, averaged across movies, was used to compute directionality of the correlation. The time-lag of the cross-correlogram global maximum determined the direction and lag of the edge. If the lag was zero, the correlation was bidirectional. If the lag was greater than 500ms (roughly 14 imaging frames), no edge was included.

Graph analysis

The partial-correlation matrix could equivalently be analyzed as a directed, weighted graph. Open source software (Gephi) was used for visualization, with node layout determined by the Yifan-Hu algorithm and tuned by hand. Edge weights less than 0.05 were set to zero for visualization clarity. Erdos-Reyni (ER) null graphs were generated for each graph to match the mean connection probability. The mean directed clustering coefficient across nodes was calculated across 50 ER graphs and averaged for comparison with data. Clustering coefficients were computed with binary matrices (nonzero weights set to one).

Modeling neuron fluorescence

To model neuron responses, we used a linear weighting of the fluorescence of every in-degree using the weights in the partial-correlation matrix. At each time point, a weighted sum was calculated, resulting in a time-varying predicted fluorescence trace. Because different numbers of input neurons and varying calcium transient amplitudes, the modeled trace was then fit with a linear offset and a gain to minimize mean-squared error with the true fluorescence trace. These two parameters were not changed when input neurons were removed (7C, 8A, 8B). For

tuned neurons, we also recomputed the trial-averaged tuning response of modeled activity. The fluorescence over a grating presentation was averaged, then trials were averaged over directions to obtain a mean direction-response. The sum of these vectors in direction space gave the model-estimated tuning vector, and the cosine similarity with the data-derived tuning vector was used to quantify the reconstruction of the tuning properties. Cosine similarity was computed in direction or orientation space according to each neuron's tuning properties. To compute total population variance explained, modeled traces were subtracted from the data traces to obtain residuals, and the ratio of total sum of squares across neurons were subtracted from one as

$$1 - \frac{\sum_i \sum_t (X_i(t) - \tilde{X}_i(t))^2}{\sum_i \sum_t (X_i(t))^2}$$

where $X_i(t)$ is the time-varying fluorescence trace of neuron i, and $X_i(t)$ *tilde* is the predicted trace.

Optimal weight estimation

Optimal weights for all incoming edges were computed for each neuron by LASSO regression as

$$\min_{\beta, \beta_0} \left(\frac{1}{2N} \sum_{i=1}^{N} (y_i - \beta_0 - x_i^T \beta)^2 + \lambda \sum_{j=1}^{p} |\beta_j| \right)$$

For weights β, and offset β_0, with modeled neuron fluorescence as y and input neuron activity as x. Weight estimation and 5-fold cross validation to estimate MSE standard error was performed with MATLAB R2016a implementation. The maximum regularization parameter (λ) whose mean-squared error did not exceed the standard error of the minimum MSE was used to find the set of optimal weights.

Acknowledgments

Author Contributions

Conceptualization: Joseph B. Dechery, Jason N. MacLean.

Data curation: Joseph B. Dechery.

Formal analysis: Joseph B. Dechery.

Funding acquisition: Jason N. MacLean.

Investigation: Joseph B. Dechery, Jason N. MacLean.

Methodology: Joseph B. Dechery, Jason N. MacLean.

Software: Joseph B. Dechery.

Supervision: Jason N. MacLean.

Visualization: Joseph B. Dechery.

Writing – original draft: Joseph B. Dechery.

Writing – review & editing: Joseph B. Dechery, Jason N. MacLean.

References

1. Hubel DH, Wiesel TN. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. The Journal of physiology. 1962 Jan 1; 160(1):106–54.

2. Lund JS, Angelucci A, Bressloff PC. Anatomical substrates for functional columns in macaque monkey primary visual cortex. Cerebral Cortex. 2003 Jan 1; 13(1):15–24. PMID: 12466211

3. Heeger DJ. Normalization of cell responses in cat striate cortex. Visual neuroscience. 1992 Aug; 9 (2):181–97. PMID: 1504027

4. Shadlen MN, Newsome WT. The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. Journal of neuroscience. 1998 May 15; 18(10):3870–96. PMID: 9570816

5. Tolhurst DJ, Movshon JA, Dean AF. The statistical reliability of signals in single neurons in cat and monkey visual cortex. Vision research. 1983 Dec 31; 23(8):775–85. PMID: 6623937

6. Seriès P, Latham PE, Pouget A. Tuning curve sharpening for orientation selectivity: coding efficiency and the impact of correlations. Nature neuroscience. 2004 Oct 1; 7(10).

7. Zohary E, Shadlen MN, Newsome WT. Correlated neuronal discharge rate and its implications for psychophysical performance. Nature. 1994 Jul 14; 370(6485):140–3. https://doi.org/10.1038/370140a0 PMID: 8022482

8. Perin R, Berger TK, Markram H. A synaptic organizing principle for cortical neuronal groups. Proceedings of the National Academy of Sciences. 2011 Mar 29; 108(13):5419–24.

9. Meshulam L, Gauthier JL, Brody CD, Tank DW, Bialek W. Collective behavior of place and non-place neurons in the hippocampal network. arXiv preprint arXiv:1612.08935. 2016 Dec 28.

10. Buonomano DV, Maass W. State-dependent computations: spatiotemporal processing in cortical networks. Nature Reviews Neuroscience. 2009 Feb 1; 10(2):113–25. https://doi.org/10.1038/nrn2558 PMID: 19145235

11. Moreno-Bote R, Beck J, Kanitscheider I, Pitkow X, Latham P, Pouget A. Information-limiting correlations. Nature neuroscience. 2014 Oct 1; 17(10):1410–7. https://doi.org/10.1038/nn.3807 PMID: 25195105

12. Averbeck BB, Latham PE, Pouget A. Neural correlations, population coding and computation. Nature reviews neuroscience. 2006 May 1; 7(5):358–66. https://doi.org/10.1038/nrn1888 PMID: 16760916

13. Ecker AS, Berens P, Tolias AS, Bethge M. The effect of noise correlations in populations of diversely tuned neurons. Journal of Neuroscience. 2011 Oct 5; 31(40):14272–83. https://doi.org/10.1523/JNEUROSCI.2539-11.2011 PMID: 21976512

14. Cohen MR, Maunsell JH. Attention improves performance primarily by reducing interneuronal correlations. Nature neuroscience. 2009 Dec 1; 12(12):1594–600. https://doi.org/10.1038/nn.2439 PMID: 19915566

15. Hu Y, Zylberberg J, Shea-Brown E. The sign rule and beyond: boundary effects, flexibility, and noise correlations in neural population codes. PLoS computational biology. 2014 Feb 27; 10(2):e1003469. https://doi.org/10.1371/journal.pcbi.1003469 PMID: 24586128

16. Montijn JS, Meijer GT, Lansink CS, Pennartz CM. Population-level neural codes are robust to single-neuron variability from a multidimensional coding perspective. Cell reports. 2016 Aug 30; 16(9):2486–98. https://doi.org/10.1016/j.celrep.2016.07.065 PMID: 27545876

17. Schneidman E, Berry MJ, Segev R, Bialek W. Weak pairwise correlations imply strongly correlated network states in a neural population. Nature. 2006 Apr 20; 440(7087):1007–12. https://doi.org/10.1038/nature04701 PMID: 16625187

18. Olshausen BA, Field DJ. What is the other 85 percent of V1 doing. van Hemmen L., & Sejnowski T. (Eds.). 2006; 23:182–211.

19. Zylberberg J. Untuned but not irrelevant: A role for untuned neurons in sensory information coding. bioRxiv. 2017 Jan 1:134379.

20. Cohen MR, Kohn A. Measuring and interpreting neuronal correlations. Nature neuroscience. 2011 Jul 1; 14(7):811–9. https://doi.org/10.1038/nn.2842 PMID: 21709677

21. Okun M, Steinmetz NA, Cossell L, Iacaruso MF, Ko H, Barthó P et al. Diverse coupling of neurons to populations in sensory cortex. Nature. 2015 May 28; 521(7553):511–5. https://doi.org/10.1038/nature14273 PMID: 25849776

22. Lin IC, Okun M, Carandini M, Harris KD. The nature of shared cortical variability. Neuron. 2015 Aug 5; 87(3):644–56. https://doi.org/10.1016/j.neuron.2015.06.035 PMID: 26212710

23. Chambers B, MacLean JN. Higher-order synaptic interactions coordinate dynamics in recurrent networks. PLoS computational biology. 2016 Aug 19; 12(8):e1005078. https://doi.org/10.1371/journal.pcbi.1005078 PMID: 27542093

24. Nigam S, Shimono M, Ito S, Yeh FC, Timme N, Myroshnychenko M et al. Rich-club organization in effective connectivity among cortical neurons. Journal of Neuroscience. 2016 Jan 20; 36(3):670–84. https://doi.org/10.1523/JNEUROSCI.2177-15.2016 PMID: 26791200

25. Watts DJ, Strogatz SH. Collective dynamics of 'small-world' networks. nature. 1998 Jun 4; 393 (6684):440–2. https://doi.org/10.1038/30918 PMID: 9623998

26. Alon U. Network motifs: theory and experimental approaches. Nature Reviews Genetics. 2007 Jun 1; 8 (6):450–61. https://doi.org/10.1038/nrg2102 PMID: 17510665

27. Timme N, Alford W, Flecker B, Beggs JM. Synergy, redundancy, and multivariate information measures: an experimentalist's perspective. Journal of computational neuroscience. 2014 Apr 1; 36 (2):119–40. https://doi.org/10.1007/s10827-013-0458-4 PMID: 23820856

28. Yu S, Yang H, Nakahara H, Santos GS, Nikolić D, Plenz D. Higher-order interactions characterized in cortical activity. Journal of neuroscience. 2011 Nov 30; 31(48):17514–26. https://doi.org/10.1523/JNEUROSCI.3127-11.2011 PMID: 22131413

29. Durand S, Iyer R, Mizuseki K, de Vries S, Mihalas S, Reid RC. A comparison of visual response properties in the lateral geniculate nucleus and primary visual cortex of awake and anesthetized mice. Journal of Neuroscience. 2016 Nov 30; 36(48):12144–56. https://doi.org/10.1523/JNEUROSCI.1741-16.2016 PMID: 27903724

30. Reimer J, Froudarakis E, Cadwell CR, Yatsenko D, Denfield GH, Tolias AS. Pupil fluctuations track fast switching of cortical states during quiet wakefulness. Neuron. 2014 Oct 22; 84(2):355–62. https://doi.org/10.1016/j.neuron.2014.09.033 PMID: 25374359

31. McLelland D, Baker PM, Ahmed B, Bair W. Neuronal responses during and after the presentation of static visual stimuli in macaque primary visual cortex. Journal of Neuroscience. 2010 Sep 22; 30 (38):12619–31. https://doi.org/10.1523/JNEUROSCI.0815-10.2010 PMID: 20861368

32. Niell CM, Stryker MP. Modulation of visual responses by behavioral state in mouse visual cortex. Neuron. 2010 Feb 25; 65(4):472–9. https://doi.org/10.1016/j.neuron.2010.01.033 PMID: 20188652

33. Friedrich J, Zhou P, Paninski L. Fast online deconvolution of calcium imaging data. PLoS computational biology. 2017 Mar 14; 13(3):e1005423. https://doi.org/10.1371/journal.pcbi.1005423 PMID: 28291787

34. Smith MA, Kohn A. Spatial and temporal scales of neuronal correlation in primary visual cortex. Journal of Neuroscience. 2008 Nov 26; 28(48):12591–603. https://doi.org/10.1523/JNEUROSCI.2929-08.2008 PMID: 19036953

35. Ko H, Hofer SB, Pichler B, Buchanan KA, Sjöström PJ, Mrsic-Flogel TD. Functional specificity of local synaptic connections in neocortical networks. Nature. 2011 May 5; 473(7345):87–91. https://doi.org/10.1038/nature09880 PMID: 21478872

36. Montijn JS, Vinck M, Pennartz CM. Population coding in mouse visual cortex: response reliability and dissociability of stimulus tuning and noise correlation. Frontiers in computational neuroscience. 2014; 8.

37. Cohen MR, Newsome WT. Context-dependent changes in functional circuitry in visual area MT. Neuron. 2008 Oct 9; 60(1):162–73. https://doi.org/10.1016/j.neuron.2008.08.007 PMID: 18940596

38. Luczak A, MacLean JN. Default activity patterns at the neocortical microcircuit level. Frontiers in integrative neuroscience. 2012; 6.

39. Oswald AM, Doiron B, Rinzel J, Reyes AD. Spatial profile and differential recruitment of GABAB modulate oscillatory activity in auditory cortex. Journal of Neuroscience. 2009 Aug 19; 29(33):10321–34. https://doi.org/10.1523/JNEUROSCI.1703-09.2009 PMID: 19692606

40. Fagiolo G. Clustering in complex directed networks. Physical Review E. 2007 Aug 16; 76(2):026107.

41. Petersen CC, Crochet S. Synaptic computation and sensory processing in neocortical layer 2/3. Neuron. 2013 Apr 10; 78(1):28–48. https://doi.org/10.1016/j.neuron.2013.03.020 PMID: 23583106

42. Ecker AS, Berens P, Cotton RJ, Subramaniyan M, Denfield GH, Cadwell CR et al. State dependence of noise correlations in macaque primary visual cortex. Neuron. 2014 Apr 2; 82(1):235–48. https://doi.org/10.1016/j.neuron.2014.02.006 PMID: 24698278

43. Schölvinck ML, Saleem AB, Benucci A, Harris KD, Carandini M. Cortical state determines global variability and correlations in visual cortex. Journal of Neuroscience. 2015 Jan 7; 35(1):170–8. https://doi.org/10.1523/JNEUROSCI.4994-13.2015 PMID: 25568112

44. Sadovsky AJ, MacLean JN. Scaling of topologically similar functional modules defines mouse primary auditory and somatosensory microcircuitry. Journal of Neuroscience. 2013 Aug 28; 33(35):14048–60. https://doi.org/10.1523/JNEUROSCI.1977-13.2013 PMID: 23986241

45. Doiron B, Litwin-Kumar A, Rosenbaum R, Ocker GK, Josić K. The mechanics of state-dependent neural correlations. Nature neuroscience. 2016 Mar 1; 19(3):383–93. https://doi.org/10.1038/nn.4242 PMID: 26906505

46. Rosenbaum R, Smith MA, Kohn A, Rubin JE, Doiron B. The spatial structure of correlated neuronal variability. Nature neuroscience. 2016 Oct 31.

47. Rubino D, Robbins KA, Hatsopoulos NG. Propagating waves mediate information transfer in the motor cortex. Nature neuroscience. 2006 Dec 1; 9(12):1549–57. https://doi.org/10.1038/nn1802 PMID: 17115042

48. Rakic P. Evolution of the neocortex: a perspective from developmental biology. Nature Reviews Neuroscience. 2009 Oct 1; 10(10):724–35. https://doi.org/10.1038/nrn2719 PMID: 19763105

49. Arieli A, Sterkin A, Grinvald A, Aertsen AD. Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses. Science. 1996 Sep 27; 273(5283):1868–71. PMID: 8791593

50. Pillow JW, Shlens J, Paninski L, Sher A, Litke AM, Chichilnisky EJ et al. Spatio-temporal correlations and visual signalling in a complete neuronal population. Nature. 2008 Aug 21; 454(7207):995–9. https://doi.org/10.1038/nature07140 PMID: 18650810

51. Truccolo W, Hochberg LR, Donoghue JP. Collective dynamics in human and monkey sensorimotor cortex: predicting single neuron spikes. Nature neuroscience. 2010 Jan 1; 13(1):105–11. https://doi.org/10.1038/nn.2455 PMID: 19966837

52. Ohiorhenuan IE, Mechler F, Purpura KP, Schmid AM, Hu Q, Victor JD. Sparse coding and high-order correlations in fine-scale cortical networks. Nature. 2010 Jul; 466(7306):617. https://doi.org/10.1038/nature09178 PMID: 20601940

53. Ohiorhenuan IE, Victor JD. Information-geometric measure of 3-neuron firing patterns characterizes scale-dependence in cortical networks. Journal of computational neuroscience. 2011 Feb 1; 30(1):125–41. https://doi.org/10.1007/s10827-010-0257-0 PMID: 20635129

54. van den Heuvel MP, Sporns O. Network hubs in the human brain. Trends in cognitive sciences. 2013 Dec 31; 17(12):683–96. https://doi.org/10.1016/j.tics.2013.09.012 PMID: 24231140

55. Kalatsky VA, Stryker MP. New paradigm for optical imaging: temporally encoded maps of intrinsic signal. Neuron. 2003 May 22; 38(4):529–45. PMID: 12765606

56. Sadovsky AJ, Kruskal PB, Kimmel JM, Ostmeyer J, Neubauer FB, MacLean JN. Heuristically optimal path scanning for high-speed multiphoton circuit imaging. Journal of neurophysiology. 2011 Sep 1; 106 (3):1591–8. https://doi.org/10.1152/jn.00334.2011 PMID: 21715667

57. Vogelstein JT, Packer AM, Machado TA, Sippy T, Babadi B, Yuste R et al. Fast nonnegative deconvolution for spike train inference from population calcium imaging. Journal of neurophysiology. 2010 Dec 1; 104(6):3691–704. https://doi.org/10.1152/jn.01073.2009 PMID: 20554834

58. Mazurek M, Kager M, Van Hooser SD. Robust quantification of orientation selectivity and direction selectivity. Frontiers in neural circuits. 2014; 8.