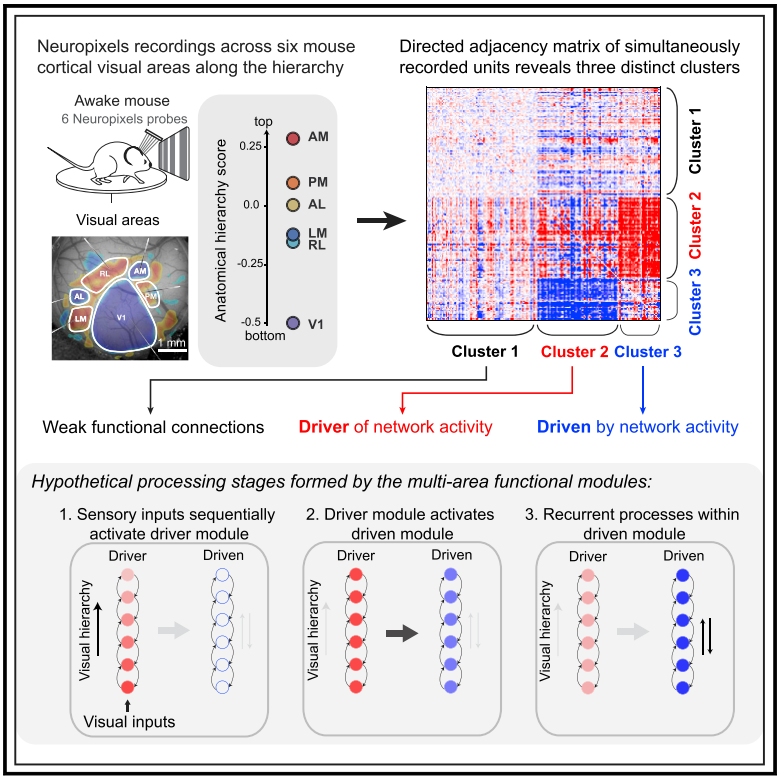
Neuron

Article

Multi-regional module-based signal transmission in mouse visual cortex

Graphical abstract



Highlights

- Cellular-scale functional network of six cortical areas along mouse visual hierarchy
- Multi-regional functional modules are revealed by clustering connectivity patterns
- Modules differ in anatomical distributions, network properties, and dynamics
- One module distributes feedforward signals and another supports recurrent processing

Authors

Xiaoxuan Jia, Joshua H. Siegle, Séverine Durand, Greggory Heller, Tamina K. Ramirez, Christof Koch, Shawn R. Olsen

Correspondence

jxiaoxuan@gmail.com (X.J.), shawno@alleninstitute.org (S.R.O.)

In brief

Jia et al. use simultaneous Neuropixels recordings from six levels of the mouse visual cortical hierarchy to generate cellular-scale, directed graphs of network interactions during sensory stimulation. Unsupervised clustering of functional connections uncovers distinct network modules that span hierarchical levels and engage in feedforward and recurrent processes.

Jia et al., 2022, Neuron *110*, 1–14
May 4, 2022 © 2022 The Author(s). Published by Elsevier Inc.
https://doi.org/10.1016/j.neuron.2022.01.027

CellPress

Please cite this article in press as: Jia et al., Multi-regional module-based signal transmission in mouse visual cortex, Neuron (2022), https://doi.org/10.1016/j.neuron.2022.01.027

Neuron

CellPress
OPEN ACCESS

Article

Multi-regional module-based signal transmission in mouse visual cortex

Xiaoxuan Jia,[1,2,*] Joshua H. Siegle,[1] Séverine Durand,[1] Greggory Heller,[1] Tamina K. Ramirez,[1] Christof Koch,[1] and Shawn R. Olsen[1,*]

[1]Allen Institute, Seattle, WA 98109, USA
[2]Lead contact
*Correspondence: jxiaoxuan@gmail.com (X.J.), shawno@alleninstitute.org (S.R.O.)
https://doi.org/10.1016/j.neuron.2022.01.027

SUMMARY

The visual cortex is hierarchically organized, yet the presence of extensive recurrent and parallel pathways make it challenging to decipher how signals flow between neuronal populations. Here, we tracked the flow of spiking activity recorded from six interconnected levels of the mouse visual hierarchy. By analyzing leading and lagging spike-timing relationships among pairs of simultaneously recorded neurons, we created a cellular-scale directed network graph. Using a module-detection algorithm to cluster neurons based on shared connectivity patterns, we uncovered two multi-regional communication modules distributed across the hierarchy. The direction of signal flow both between and within these modules, differences in layer and area distributions, and distinct temporal dynamics suggest that one module transmits feedforward sensory signals, whereas the other integrates inputs for recurrent processing. These results suggest that multi-regional functional modules may be a fundamental feature of organization beyond cortical areas that supports signal propagation across hierarchical recurrent networks.

INTRODUCTION

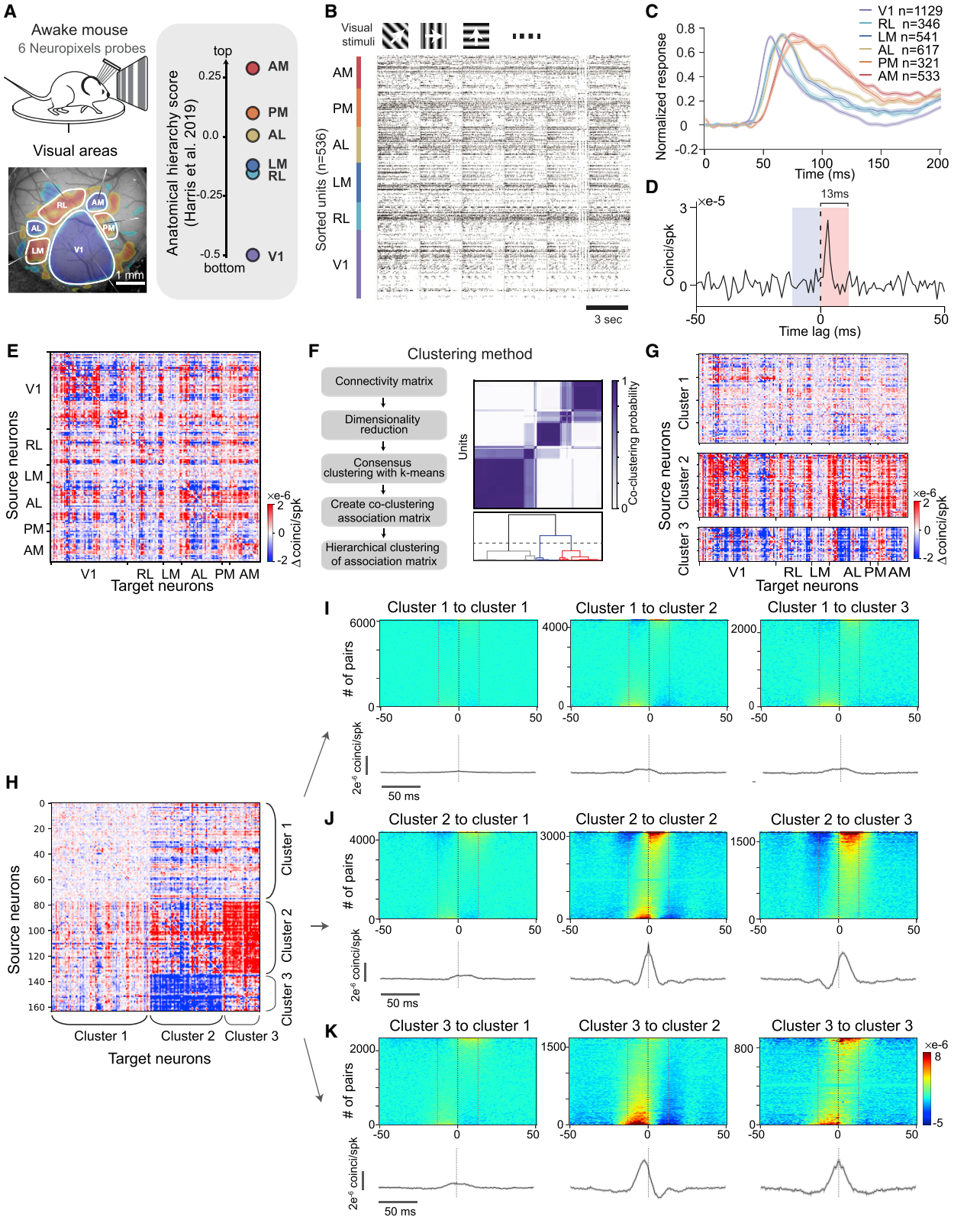
Information processing in the neocortex involves signal representation, transformation, and transmission between processing levels or modules (Perkel and Bullock, 1968). Sensory systems are organized in anatomical hierarchies (Felleman and Van Essen, 1991; Harris et al., 2019). Areas at different hierarchical levels are thought to correspond to distinct processing modules linked by feedforward and feedback projections (DiCarlo et al., 2012; Gilbert and Li, 2013; Hubel, 1988; Kumar et al., 2010). Whereas stimulus representation (such as single neuron tuning preference and population coding) has been extensively studied (Averbeck et al., 2006; Cunningham and Yu, 2014; Hubel and Wiesel, 1962; Mountcastle et al., 1963; Orban, 2008; Parker and Newsome, 1998), the principles of spike-based transmission between neuronal populations across a processing hierarchy are much less understood (Kohn et al., 2020; Kumar et al., 2010; Zavitz and Price, 2019).

Tracking signal propagation in distributed networks requires simultaneous measurement of large numbers of interacting neurons both within and between areas in behaving animals (Buzsáki, 2004; Kohn et al., 2020; Zavitz and Price, 2019). Previous work examining inter-area neuronal communication on millisecond timescales has revealed stimulus dependence and functional specificity of these interactions (Chen et al., 2017; Goldey et al., 2014; Jia et al., 2013; Reid and Alonso, 1995; Semedo et al., 2019; Zandvakili and Kohn, 2015). Other studies have used local field potentials (LFPs) as an indicator of aggregate population activity and found associations between different frequency bands with feedforward (gamma) versus feedback (beta) signaling (Bastos et al., 2015, 2018; van Kerkoerle et al., 2014; Wong et al., 2016). However, LFP measurements cannot resolve single neurons, limiting the ability of these experiments to monitor cellular-scale signal flow. Widefield calcium imaging techniques provide simultaneous access to many cortical regions (Musall et al., 2019; Salkoff et al., 2020; Wekselblatt et al., 2016). However, these signals have much slower dynamics that fail to capture signal transmission at millisecond timescales. Thus, a comprehensive view of network communication spanning multiple hierarchical levels of sensory processing is lacking because multi-area recordings with single-neuron and single-spike precision have only recently become more tractable, thanks to the introduction of compact, high-density implantable electrodes such as Neuropixels (Jun et al., 2017; Siegle et al., 2021; Steinmetz et al., 2021; Vesuna et al., 2020).

In our previous work, we built a recording platform using multiple Neuropixels probes to simultaneously measure hundreds of spiking neurons from up to eight levels of the mouse visual hierarchy (Siegle et al., 2021) (Figure 1A). By analyzing correlations between spiking units, we found that the average direction of signal flow during bottom-up sensory drive follows the anatomically-defined hierarchy (Harris et al., 2019; Siegle et al., 2021). Consistent with this, we observed longer response latencies at higher hierarchical levels. However, ample anatomical evidence demonstrates the existence of parallel pathways and recurrent connections between hierarchical levels of the cortex (Felleman and Van Essen, 1991; Gămănuţ et al., 2018; Harris et al., 2019; Markov et al., 2014a). In the mouse, primary visual cortex (V1) makes direct

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Please cite this article in press as: Jia et al., Multi-regional module-based signal transmission in mouse visual cortex, Neuron (2022), https://doi.org/10.1016/j.neuron.2022.01.027



(legend on next page)

Please cite this article in press as: Jia et al., Multi-regional module-based signal transmission in mouse visual cortex, Neuron (2022), https://doi.org/10.1016/j.neuron.2022.01.027

projections to all higher visual areas (Harris et al., 2019; Wang and Burkhalter, 2007), and even single cells can project in parallel to multiple areas via branching axons (Han et al., 2018). In addition, local subnetworks show clustered connectivity that could support the coactivation of groups (or assemblies) of neurons (Perin et al., 2011; Song et al., 2005; Yoshimura et al., 2005).

Given the complexity of local and inter-area connections, identifying the relevant signal transmission modules is challenging. The canonical interpretation is that each cortical area corresponds to a processing stage that performs local computations and sends output signals sequentially from one area to the next. However, this standard view is likely only a first-order characterization. Because of dense lateral and recurrent connectivity, it is possible that major communication hubs actually span multiple anatomical areas. Thus, a critical foundational step is to subdivide the system into functionally relevant processing modules at the multi-regional cellular level.

Here, we take such an approach by decomposing the cortical network into modules based on functional connectivity. Using multi-area recordings from six interconnected, hierarchically organized cortical areas (Siegle et al., 2021), we infer signal flow based on the statistics of leading-lagging spike timing between neurons. Treating each neuron as a node, we create an adjacency matrix with directed weights and then use an unsupervised clustering algorithm to uncover multi-regional sets of neurons based on their shared functional connectivity patterns. This method revealed two major visually engaged modules that span the cortical hierarchy. Several lines of evidence suggest that one module is mainly involved in the feedforward distribution of sensory information, while the other is more engaged in the recurrent integration of information.

RESULTS

We deployed up to six Neuropixels probes to record neurons from six areas of the mouse visual cortex. Each area has its own map of visual space (Garrett et al., 2014) and resides at a different hierarchical level as determined both anatomically (D'Souza et al., 2022; Harris et al., 2019) and physiologically (Siegle et al., 2021). Primary visual cortex (V1) is at the bottom of the hierarchy, followed by rostrolateral (RL), lateromedial (LM), anterolateral (AL), posteromedial (PM), and anteromedial (AM) areas (Figure 1A). Overall, recording sessions in this study yielded 632 ± 18 simultaneously recorded neurons per experiment (a.k.a. sorted "units" [Siegle et al., 2021]) distributed across cortical layers and areas (n = 19 mice, mean ± SEM) (Figure 1B). We used full-field drifting-grating stimuli to provide a strong bottom-up sensory input that evokes a large number of spikes recorded per unit time. Consistent with the known visual hierarchy in the mouse (Harris et al., 2019; Siegle et al., 2021), the mean response latency in each area followed a sequential progression (Figure 1C). However, all areas were coactive for substantial portions of the visual response, thereby providing opportunities for recurrent interactions. To facilitate functional connectivity analysis, neurons in our dataset were filtered by minimal firing rate (>2 spikes/s) and receptive field location (center of RF at least 10 degrees away from the edge of the monitor). After filtering, n = 3,487 units remained, which is 29% of total units recorded across all mice (see STAR Methods for details).

Identification of multi-regional functional modules

To characterize fast-timescale interactions relevant to signal transmission, we quantified spike timing cross-correlations between all pairs of neurons in each mouse using jitter-corrected cross-correlogram (CCG) analysis; this captures relative spike timing between two neurons within the jitter window (25 ms) but removes stimulus-locked signals and correlations longer than the jitter window (Jia et al., 2013; Smith and Kohn, 2008). For each neuronal pair, we determined their connection weight by computing the difference of the integral under the CCG curve in a 13 ms window (about half of the jitter window) before and after 0 time lag (Figure 1D), which is an indicator of the asymmetry of the CCG with short delays. As a result, the sign of this weight describes the "signal flow direction" (temporally leading or following) between pairs of neurons. Note that we did not select CCGs based on their shape (e.g., sharpness of the peak), so this measure likely includes both mono- and polysynaptic connections within the chosen time window. By computing the connection weight for all pairs, we produced a connectivity matrix describing the directed functional interactions of all simultaneously recorded neurons (Figure 1E). These functional connectivity matrices displayed non-random structure (Figure S1). Inspection indicated the existence of separable groups of neurons with similar patterns of functional connectivity. We next sought to algorithmically uncover these groups by clustering their connection profiles.

Figure 1. Identification of functional modules in multi-regional spike recordings

(A) Top: schematic of experimental setup. Bottom: Neuropixels recordings from six visual cortical areas. Retinotopic sign map is overlaid on vasculature image to guide area targeting. Right: replotting of anatomical hierarchy score from Harris et al. (2019); Figure 6E.

(B) Raster plot of 536 simultaneously recorded neurons during drifting-grating stimulation (six sequential trials are shown, each with 2 s grating presentation followed by 1 s gray period).

(C) Normalized peristimulus time histograms (PSTHs) in response to grating stimuli from six visual cortical areas.

(D) An example jitter-corrected CCG between a V1 neuron and an LM neuron with a sharp peak. The connection weight is the difference of the CCG amplitude averaged in a 13 ms window before and after 0 time lag, which reflects asymmetry (directionality) in the CCG.

(E) Directed cellular-scale connectivity matrix for one exemplar animal. Neurons are sorted by area and depth.

(F) Clustering procedure and result for the same data.

(G) Connectivity profiles of three functional clusters from the same mouse with source neurons organized by area and depth.

(H) Adjacency matrix organized by clusters indicating modular structure.

(I) Jitter-corrected CCGs from neurons in cluster 1 to neurons in each cluster for the same example mouse.

(J) Jitter-corrected CCGs from neurons in cluster 2 to each cluster.

(K) Jitter-corrected CCGs from neurons in cluster 3 to each cluster. Top row: jitter-corrected CCGs are ordered according to connection weight (high to low) from top to bottom. Bottom row: averaged CCGs are shown with error bars indicating standard error of the mean.

Please cite this article in press as: Jia et al., Multi-regional module-based signal transmission in mouse visual cortex, Neuron (2022), https://doi.org/10.1016/j.neuron.2022.01.027

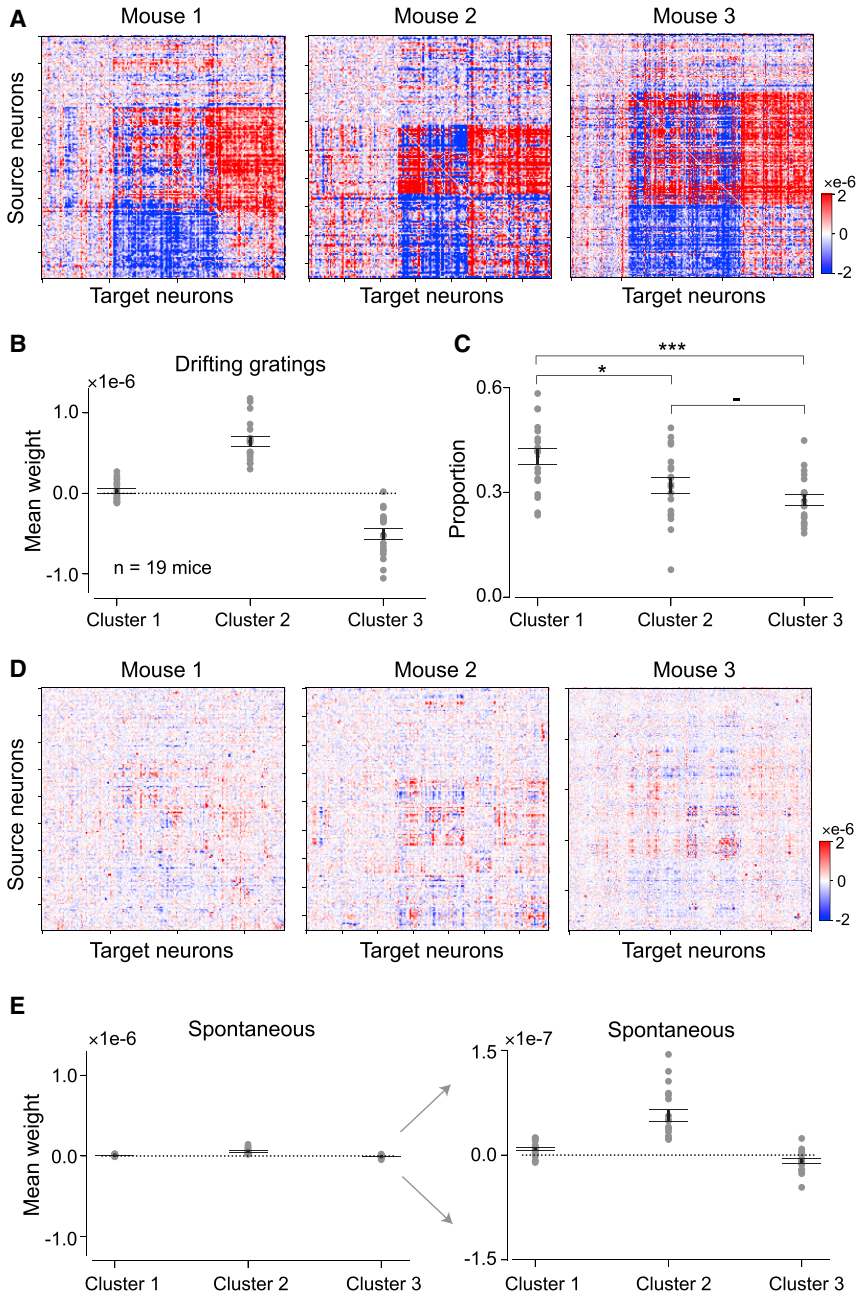


Figure 2. Consistent directed functional modules observed across mice

(A) Adjacency matrices for three example mice during grating stimuli (units grouped by cluster ID).

(B) Mean functional connection weight for each cluster during drifting-grating stimulation average across mice. Clusters 2 and 3 have values significantly different than 0 (Two-tailed one-sample t test to compare with 0, cluster 1: T = 1.1, p = 0.29; cluster 2: T = 10.9, p = 2.3e−9; cluster 3: T = −7.6, p = 4.7e−7).

(C) Proportion of units in each cluster in each mouse. Statistics are performed with Student paired t test (Bonferroni corrected). –, not significant; *, p < 0.05; ***, p < 0.001.

(D) Adjacency matrices for the same three mice during spontaneous activity (unit ordering is the same as in A).

(E) Mean weight during spontaneous period (gray screen). Two-tailed one-sample t test to compare with 0, cluster 1: T = 3.4, p = 0.0036; cluster 2: T = 6.8, p = 2.5e−6; cluster 3: T = −2.5, p = 0.024. Right is a zoomed-in view. Gray dots indicate individual mice. Data are represented as mean ± SEM across mice (n = 19 mice).

To systematically identify sets of source neurons with similar connections, we clustered the interaction matrix by treating connections from each source neuron to all neurons in the recorded network as features (Figures 1E and 1F; Figure S2). This procedure yielded three robust clusters of source neurons (Figure S3): cluster 1 had mostly weak connections; neurons in cluster 2 were dominated by strong positive connection weights, indicating that they tended to lead (or drive) network activity; and neurons in cluster 3 were dominated by strong negative connection weights, indicating that they tended to follow (or were driven by) activity in the network (Figures 1G and 1H). Visualizing the jitter-corrected CCGs between all neurons in each of the different clusters (Figures 1I–1K) showed that: (1) CCGs with cluster 1 units are typically very weak; (2) CCGs between cluster 2 and 3 are asymmetric, with cluster 2 units leading the activity of cluster 3 units; and (3) CCGs within either cluster 2 or cluster 3 have both leading and lagging pairs and are, on average, more balanced.

These three clusters were consistently identified in each mouse we examined, suggesting they are a core organizational feature of visual cortex (Figure 2A). Given the bias in positive versus negative connection weights from each cluster (and its functional implication of directionality), we refer to cluster 2 source neurons as the "driver module" and cluster 3 source neurons as the "driven module" (Figure 2B). Supporting the robustness of these clusters, we observed similar network modules using spectral clustering and bi-clustering algorithms (Pedregosa et al., 2011) (Figure S3). The relative proportions of driver and driven units are relatively similar (Figure 2C). These modules are readily apparent during periods of high-contrast drifting-grating stimulation (Figures 2A and 2B) but are less visible during spontaneous activity (gray screen; Figures 2D and 2E). However, we do observe a pattern of functional connectivity during the spontaneous condition that is similar to the drifting-grating condition when we expand the y-scale to reveal these weak connections (Figure 2E, right). This suggests that the network structure giving rise to the separation of driver and driven modules is likely present

Please cite this article in press as: Jia et al., Multi-regional module-based signal transmission in mouse visual cortex, Neuron (2022), https://doi.org/10.1016/j.neuron.2022.01.027

CellPress
OPEN ACCESS

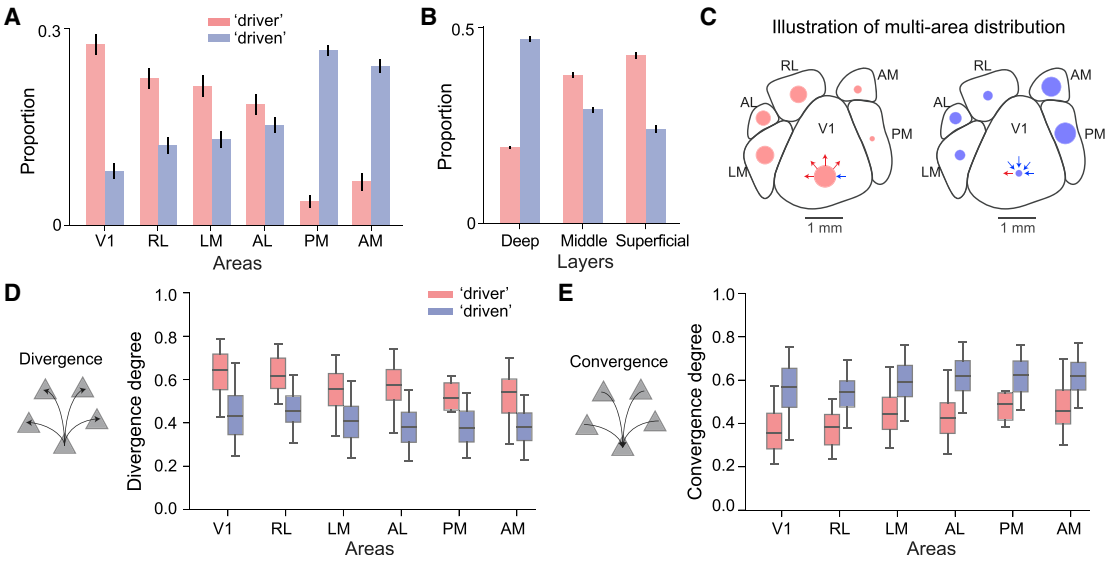


Figure 3. Area distribution, divergence, and convergence differ across modules

(A) Distribution of driver and driven modules across the visual cortical hierarchy (see STAR Methods).

(B) Cortical layer bias of the driver and driven modules for all mice. Error bars in (A) and (B) represent bootstrapped standard error (see STAR Methods).

(C) Illustration of the relative proportion of neurons (indicated by dot size) in driver (red dots) and driven (blue dots) modules across the visual cortical hierarchy. The arrows schematize the different pattern of convergence and divergence of the driver and driven modules.

(D) Boxplot of divergence degree distribution across areas for driver and driven modules. Box delimits first to third quartile, with the middle line indicating the median of the distribution. The whiskers correspond to 5% and 95% percentiles.

(E) Convergence degree distribution for driver and driven modules.

in the absence of visual stimulation, but is likely less visible because of an order of magnitude lower firing rate under spontaneous conditions (the averaged firing rate is 0.51 ± 0.0062 spike/s during spontaneous activity versus 4.7 ± 0.059 spike/s during drifting grating).

In which cortical areas and layers do neurons in these two modules reside? Interestingly, neurons in both the driver and driven modules were spatially distributed across all levels of the cortical hierarchy rather than being localized to specific regions (Figure 3A). Nonetheless, the proportion of neurons in the two modules showed area biases. Overall, driver neurons decreased along the hierarchy (Figure 3A; Spearman's correlation with each area's anatomical hierarchy score [see STAR Methods; Harris et al., 2019]: r = −0.94, p = 0.005), whereas driven neurons increased (Spearman's correlation = 0.94, p = 0.005 for driven module) (Figures S4A and S4B). Both driver and driven neurons were present in all cortical layers but also showed laminar biases. Driver neurons were enriched in the middle and superficial layers, whereas driven neurons were more common in deeper cortical layers (Figure 3B). Previous anatomical tracing experiments have shown that neurons mediating feedforward projections tend to originate in superficial layers (Felleman and Van Essen, 1991; Harris et al., 2019; Markov et al., 2014a) with the fraction of feedforward projecting neurons in superficial layers decreasing along the hierarchy (Barone et al., 2000; Markov et al., 2014b). In our analysis, the neurons in the driver module followed the same pattern, suggesting this module could be involved in feedforward processing, while the driven module might be more involved in recurrent processing.

Signal integration and distribution by different modules

Node convergence ("input" to one neuron from others) and divergence ("output" from one neuron to others) are two network properties that differentially support the integration versus distribution of signals (Tononi et al., 1998). The predominately positive connection weights of the driver neurons versus the mostly negative weights of the driven neurons suggest that driver neurons should, on average, have a higher divergence degree (more outward projections compared to inward projections). To quantify this, we computed the divergence degree of each neuron as the number of positive connections (outward projections from the neuron) relative to the size of the network. Likewise, convergence degree of a neuron was computed as the fraction of negative connections. In a randomly connected all-to-all network, the default value is 0.5 for divergence and convergence degree. Driver module neurons had higher divergence than driven neurons (Figure 3D; two-way ANOVA across areas, between modules F = 1029.5, p = 5.4e−184; among areas F = 35.4, p = 7.9e−35; interaction F = 3.6, p = 0.003). In contrast, the driven module neurons had higher convergence (Figure 3E; between modules F = 1029.5, p = 5.3e−184; among areas F = 35.3, p = 8.2e−35; interaction F = 3.6, p = 0.003). Thus, from a network perspective, neurons in the driver module are better positioned to distribute information, whereas neurons in the driven module are better positioned for signal integration. We also found significant correlations between divergence and convergence degree across the visual hierarchy (Figure S4). For example, the divergence degree of driver neurons gradually decreased along the visual hierarchy (Spearman's r = −0.94,

Please cite this article in press as: Jia et al., Multi-regional module-based signal transmission in mouse visual cortex, Neuron (2022), https://doi.org/10.1016/j.neuron.2022.01.027

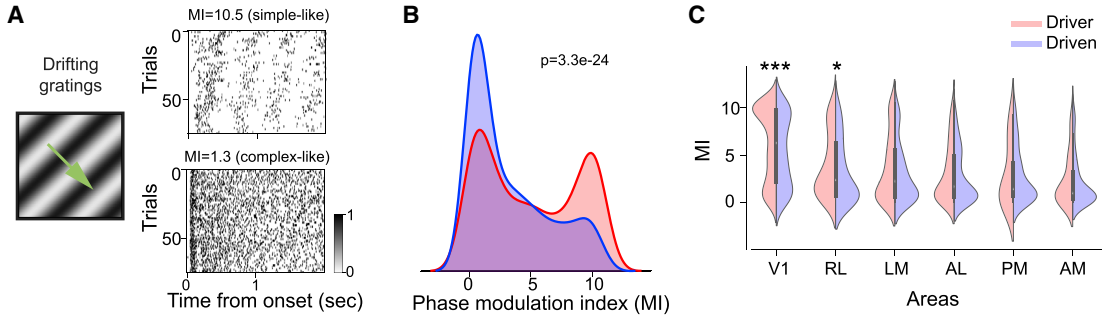


Figure 4. More complex-like neurons in the driven compared to the driver module

(A) Illustration of drifting-grating stimuli and the raster plot of two example units from driver and driven modules.

(B) Distribution of phase modulation index (MI) for driver (red) and driven (blue) neurons.

(C) Mean MI of neurons belonging to different modules in different brain areas. Error bars represent SEM. Significance is tested with Mann Whitney U test. *, $p < 0.05$; ***, $p < 0.001$.

$p = 0.0048$; Pearson's $r = -0.84$, $p = 0.038$) (Figure S4C). This suggests that driver units in early visual areas are more capable of distributing signals compared to their peers in higher visual areas. This hierarchy correlation is not a simple outcome of our clustering because an area-label shuffled control analysis did not show a significant correlation with the hierarchy (ANOVA among areas $F = 0.60$, $p = 0.70$; Spearman's correlation $r = -0.09$, $p = 0.87$) but still retained a significant difference in convergence between the driver and driven neurons (ANOVA for divergence degree between modules: $F = 1553.7$, $p = 1.6e{-}254$; ANOVA for convergence degree between modules: $F = 311.7$, $p = 3.9e{-}65$).

The greater convergence to driven neurons suggests that they combine signals during visual processing. In the visual system, the convergence of multiple simple-like neurons that are modulated by drifting-grating phase can give rise to complex-like neurons with tolerance to different grating phases. Simple- and complex-like neurons can be quantified using a modulation index (MI) that describes the degree of modulation at the preferred temporal frequency of the neuron (Matteucci et al., 2019); simple-like cells show a large MI, while complex cells have a small MI (in the limit, close to zero). To test whether the driven module contains more complex-like neurons consistent with their dominant converging inputs relative to the driver module, we computed the MI for individual neurons and compared the two modules (Figures 4A and 4B). Overall, the driven population was more complex compared to the driver population ($p = 3.3e{-}24$; Mann-Whitney U test). Because there are more simple-like neurons in V1 (reflected by the bimodal distribution in Figure 4C), we also tested the MI between driver and driven modules when only comparing higher-visual areas. We found that driver neurons have significantly higher MI compared to driven neurons even after removing V1 ($p = 0.00011$; Mann-Whitney U test). This result is consistent with the driver neurons' higher divergence and suggests they are functionally deeper in the visual processing pathway.

Signal transmission between modules

Having defined two robust network modules, each spanning the visual hierarchy, we next sought to explore further the signal flow within and between these separable network clusters. To do so, we visualized and analyzed subnetworks of the directed weight matrix defined previously (Figures 1 and 2). We first focused on signal transmission between modules by examining the subnetworks defined by the connections from driver to driven neurons (Figure 5A, top) and from driven to driver neurons (Figure 5A, bottom). As expected, connections between these modules were largely unidirectional; that is, driver neurons from each source area (including top areas of the visual hierarchy, e.g., PM and AM) made output connections to driven neurons, and driven neurons in each source area received input connections from driver neurons (Figure 5B). The asymmetry of these connections indicates a largely unidirectional signal flow from the driver to driven module.

When determining functional modules, our clustering algorithm only considered patterns of connectivity and ignored their area labels. However, it is generally accepted that the hierarchical organization of cortical areas is relevant for signal propagation. Therefore, to quantify this signal flow from the perspective of individual cortical areas in each subnetwork, we define a metric called the "area in-out index," which describes the relative fraction of input versus output connections from a source area for a given subnetwork. An in-out index of 1 indicates all connections are inputs to the source area, and an index of -1 indicates all connections are outputs (see STAR Methods). For all areas of the driver-to-driven subnetwork, the in-out index was close to -1 (Figure 5C, top), indicating virtually all connections are outputs. In contrast, the in-out index was close to +1 in each area for the driven-to-driver subnetwork across mice (Figure 5C, bottom). These values are a consequence of the unidirectional signal flow from the driver to driven modules and are expected based on inspection of the weight matrix. However, a closer evaluation of the between module graphs (Figure 5B) indicated the layer distribution of connections is not equal across areas. To compare the anatomical layout of these connections, we overlayed the two subnetworks (driver-to-driven and driven-to-driver) for each source area; this showed a clear separation of inward and outward projections across the cortical depth (Figure 5D), with outward projections concentrated in middle/superficial layers of the source area and inward projections concentrated

Please cite this article in press as: Jia et al., Multi-regional module-based signal transmission in mouse visual cortex, Neuron (2022), https://doi.org/10.1016/j.neuron.2022.01.027

CellPress
OPEN ACCESS

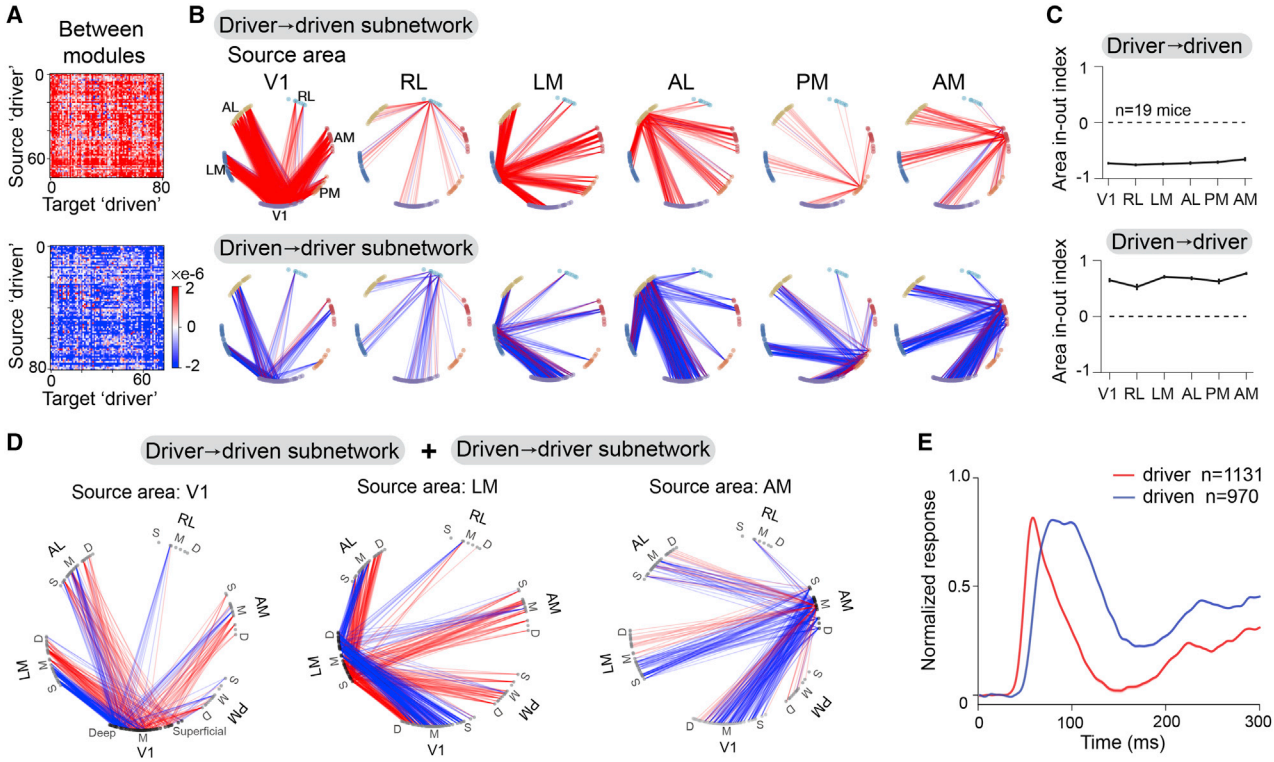


Figure 5. Unidirectional signal flow between distinct modules

(A) Directed connectivity matrix for between-module subnetworks (top: driver [n = 74] → driven [n = 81]; bottom: driven → driver) of one example mouse.

(B) Graph representation of strong connections (weight value $>10^{-6}$ for visualization purposes) between source area (labeled at the top) and all other areas for between module subnetwork in (A). Each node in the graph represents a recording site in cortex. Identity of visual areas is indicated by node color and arranged according to their anatomical location. Nodes in each area are arranged clockwise from superficial to deep layers. Line color indicates the directionality of connection (red, outward projections; blue, inward projections).

(C) Area in-out index averaged across mice (n = 19 mice) for between-module subnetworks.

(D) Graph representation of overlaid between-module subnetworks (driver-to-driven and driven-to-driver) for source areas V1, LM, and AM. Left: Directed graph representation for between-module subnetworks from source area V1 of one example mouse. Graph representation of strong connections (weight value $>2*10^{-6}$ for visualization purposes). Line color indicates the directionality of connection. Red indicates output from driver neurons in V1 → driven neurons in other areas; blue indicates input from driver neurons in other areas → driven neurons in V1. There is a clear separation of input and output from source area V1, and the separation is cortical depth dependent. Middle: Source area LM. Right: Source area AM.

(E) Population average of normalized PSTH for the two modules. Error bar represents SEM, which is too small to be visible.

in deep layers. These observations reflect the biases we found previously, in which driver neurons are abundant in superficial layers, whereas driven neurons are biased toward deep layers (Figure 3B).

The directional communication from driver to driven modules suggests that these two modules might be sequentially activated during visual stimulation. To test this, we quantified the stimulus response latency for each module and found that spikes in the driver module preceded the driven module (drive time to peak = 60.5 ± 0.3 ms versus 80.0 ± 0.6 ms for the driven module; Rank-sum test statistics = −29, p = 7.1e−186) (Figure 5E). We performed simulations demonstrating that the existence of brief timescale correlations between neurons in these modules does not necessarily entail a temporal offset between modules in the stimulus-triggered average response and vice versa (Figure S5). Thus, the directional communication implied by the between-module subnetworks is supported by the temporal progression of signals between the modules.

Signal transmission within modules

Whereas connectivity *between* modules was largely unidirectional, the connections *within* each module contained both directions, indicating more possibilities for recurrent processing (Figures 6A and 6B). This is supported by the averaged CCGs of the within-module networks, which showed a peak at 0 time delay (Figures 1J and 1K). To further explore the area specificity of the within-module communication structure, we analyzed the within-module subnetworks (Figure 6C). Interestingly, for the driver module, the area in-out index increased across the hierarchy: V1 had a negative in-out index and mostly made output connections to driver neurons in other areas; in contrast, driver neurons in AM received more inputs compared to outputs as indicated by a positive in-out index (Figure 6C, top; Spearman's correlation with hierarchy is 0.83, p = 0.04). Within the driven module, connections were more balanced: the in-out index was close to 0 for each area and did not significantly correlate with the anatomical hierarchy score (Spearman's correlation p = 0.21; Figure 6C,

Please cite this article in press as: Jia et al., Multi-regional module-based signal transmission in mouse visual cortex, Neuron (2022), https://doi.org/10.1016/j.neuron.2022.01.027

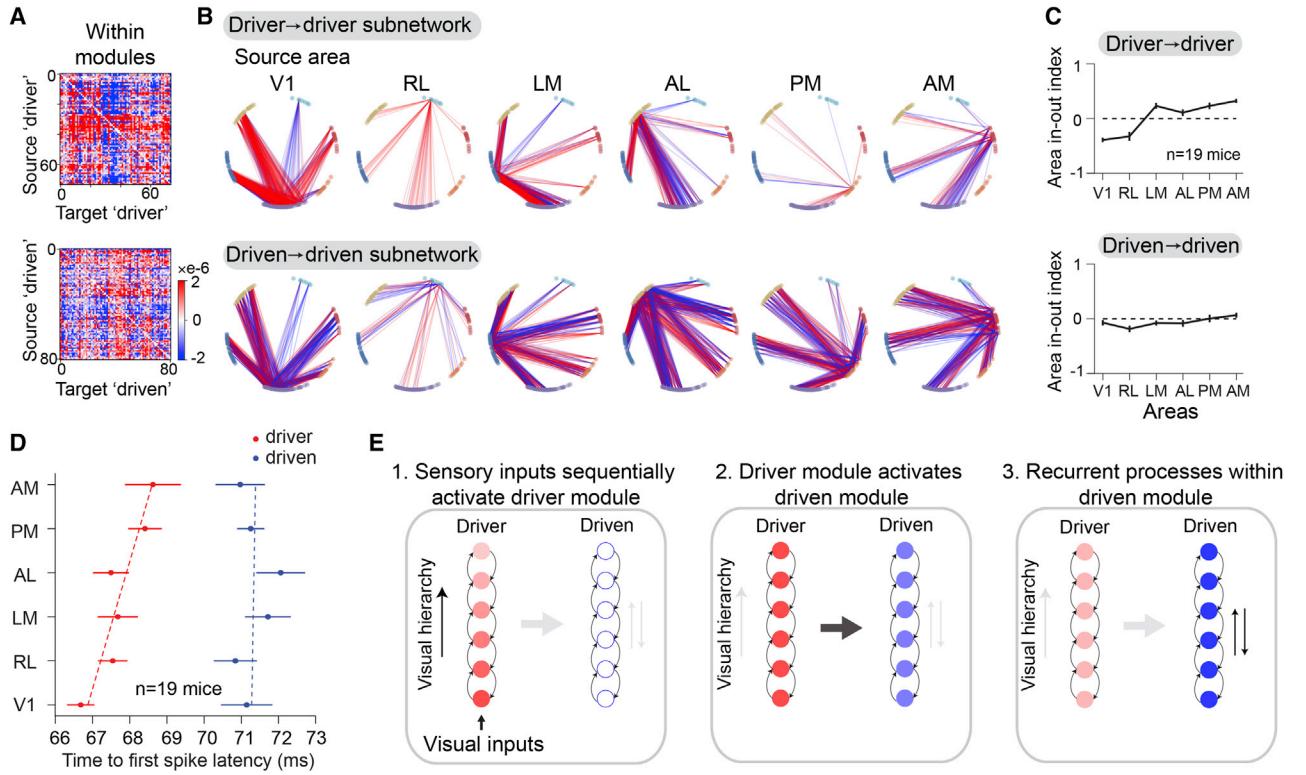


Figure 6. Signal flow within each module indicates a separation of feedforward and recurrent processes

(A) Directed connectivity matrix for within-module subnetworks (top: driver → driver [n = 74]; bottom: driven → driven [n = 81]) of the same example mouse as in Figure 5.

(B) Graph representation of strong connections (weight value $>10^{-6}$) between source area (labeled at the top) and all other areas for within module subnetwork in (A).

(C) In-out index averaged across all 19 mice for the within-module subnetworks.

(D) Response latency (time to first spike) for different modules across areas for all 19 mice.

(E) Diagram of the three temporal stages of information flow revealed by data. Intensity of color indicates activation level (darker is stronger).

Error bars in (C) and (D) represent SEM.

bottom). Thus, the signal flow between driver neurons, but not between driven neurons, is correlated with the visual hierarchy; this result is not explained by the biased distribution in driver neurons across areas because the proportion of driven neurons is also biased across areas (Figures S4A and S4B).

These within-module patterns of connectivity suggest that driver module neurons relay feedforward signals to driver neurons in other areas along the hierarchy, whereas bi-directional connections within the driven module are positioned to mediate recurrent interactions between areas. Consistent with this observation, the mean visually evoked response latency within the driver module systematically increased across area hierarchy (Figure 6D; correlation with mouse anatomical hierarchy score from Harris et al. [2019]: Spearman's r = 0.83, p = 0.04). In contrast, visual latencies of neurons in the driven module were delayed relative to the driver module but did not show an organized progression across the hierarchy (Spearman's r = 0.14, p = 0.79).

Together, these results suggest a working model in which these two modules support distinct stages of signal propagation: one module transmits feedforward signals about external stimuli along the hierarchy; the other integrates and processes recurrent signals given inputs from the driver module (summarized in Figure 6E).

Temporal dispersion of population spiking differs between modules

Modeling studies of feedforward networks have investigated how spiking signals are propagated across modules of sequential processing (Kumar et al., 2010; Reyes, 2003; Vogels and Abbott, 2005). Depending on various network parameters, including connection density and synaptic weights, successive stages can propagate synchronous activity (e.g., a synfire chain; Abeles, 1991) or asynchronous fluctuations in firing rate (rate-coded signals). *In vitro* experiments with cultured networks report decreased within-module synchrony of spiking activity as signals are relayed across sequential processing modules (Barral et al., 2019). However, evidence from networks in behaving animals is rare (Zandvakili and Kohn, 2015). The distributed functional modules we describe here could represent sequential stages of signal processing. In this context, we compared the temporal dispersion of the first stimulus-evoked spike on a trial-by-trial basis for the population of neurons in either the driver or driven modules (Figures 7A–7C). The driver population had less temporal dispersion in response to stimulus onset compared with the driven module (Figure 7C; the onset temporal dispersion for the driver module is 14.0 ± 3.9 ms and

Please cite this article in press as: Jia et al., Multi-regional module-based signal transmission in mouse visual cortex, Neuron (2022), https://doi.org/10.1016/j.neuron.2022.01.027

CellPress
OPEN ACCESS

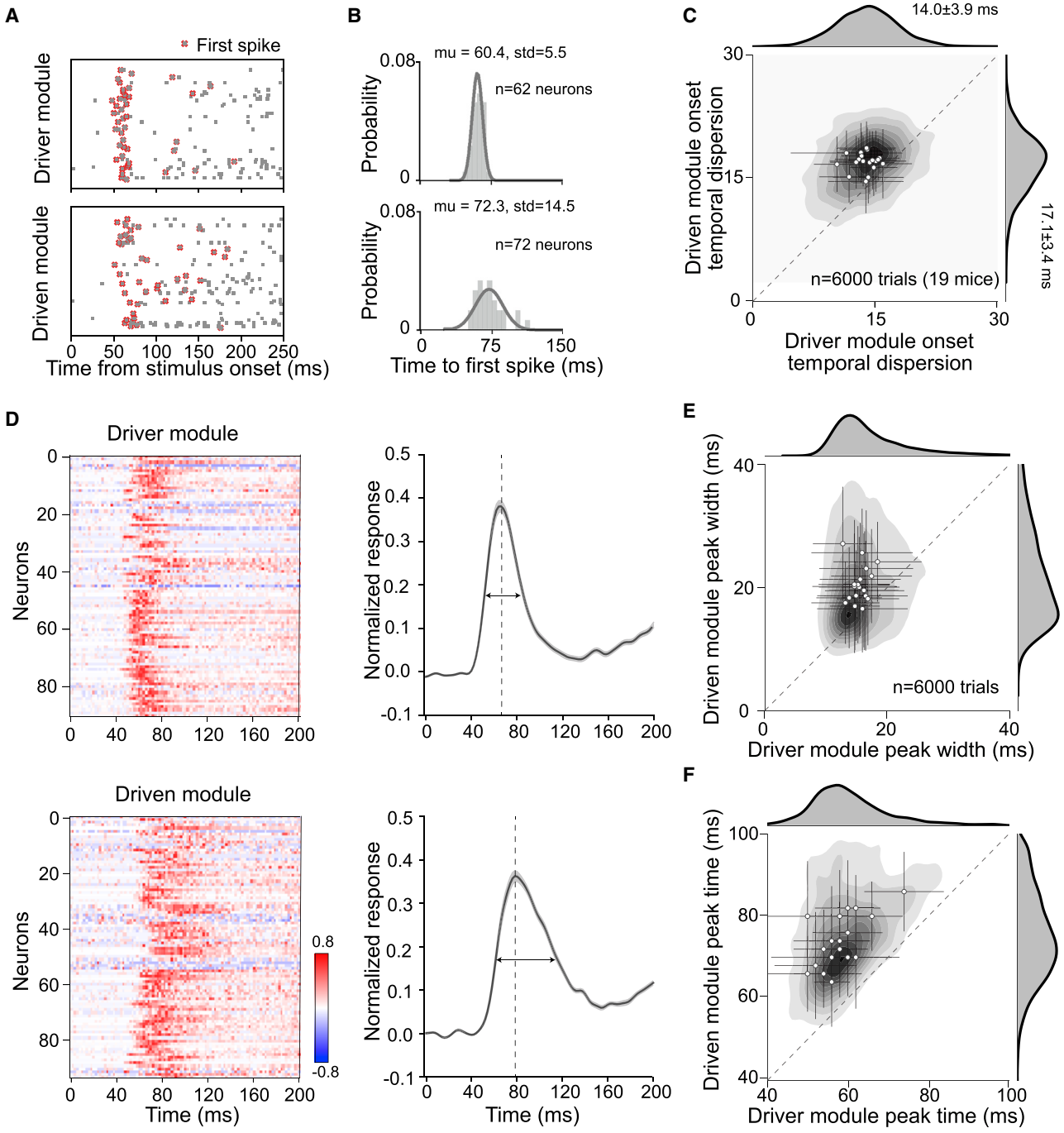


Figure 7. Temporal dispersion of population spiking differs between modules

(A) Raster plots for simultaneously recorded neurons from the driver and driven modules for a single trial from an example mouse. Each row is the spike times of one neuron. Red cross highlights the first stimulus-evoked spike.

(B) First spike latency distribution of each module from the trial shown in (A). The spread quantifies module onset-latency dispersion.

(C) Joint plot of module onset temporal dispersion comparing driver and driven module across trials (n = 6,000 trials from 19 mice).

(D) Left: normalized PSTHs of driver neurons (n = 91) and driven neurons (n = 94) from an example mouse. Right: the corresponding normalized population response from the same mouse (driver: peak at 66 ms with spread of 15.24 ms; driven: peak at 66 ms with spread of 15.24 ms; n = 75 trials).

(E) Joint scatterplot of the population response peak spread across all trials.

(F) Joint scatterplot of population response peak latency across all trials. White dots represent mean of individual mouse with standard deviation across trials.

Please cite this article in press as: Jia et al., Multi-regional module-based signal transmission in mouse visual cortex, Neuron (2022), https://doi.org/10.1016/j.neuron.2022.01.027

for the driven module is 17.1 ± 3.4 ms, n = 6,000 trials combined across mice, Student's two-tailed t test statistics T = 43.6, p < 0.001). Interestingly, the onset dispersion of the single trial population response was weakly correlated between the driver and driven modules (Pearson r = 0.26, p < 0.001, n = 6,000) across trials, which indicates the general level of population jittering could change across trials or fluctuate across time (e.g., based on brain state). We also computed the temporal spread of the trial-wise population response (Figures 7D–F) and found that the driver module was more compact (Figure 7E; Student's two-tailed t test statistics T = −31.4, p = 3.3e−211, n = 6,000 trials across 19 mice) and occurred earlier (Figure 7F; T = −17.4, p = 1.4e−67) than the driven module. These results show that the transmission between the driver and driven module is associated with increased temporal dispersion of spiking and reduced stimulus onset synchronization. This is consistent with the concept of increasing recurrent interactions deeper into the processing chain (Goris et al., 2014; Lu et al., 2001).

DISCUSSION

Separate sets of neurons for distributing and integrating signals

Our results provide a multi-area perspective on signal flow in the mouse visual cortex. Contrary to a simple feedforward model in which each hierarchical level sequentially transmits signals point-to-point up the chain of areas, we provide evidence for module-based signal propagation that simultaneously involves neurons at multiple levels of the hierarchy. The cluster of neurons we have denoted as the *driver* module are engaged earlier in the sensory processing stream, while *driven* neurons are recruited later; this suggests these two modules might represent distinct processing stages: one module is distributing feedforward sensory signals while the other is responsible for recurrent integration. This conclusion is supported by various differences between these two modules, including their functional interactions, layer and area distributions, network convergence and divergence, and within-module temporal coordination.

We made several physiological and computational assumptions during experiments and analysis. First, we selected neurons that have an average firing rate larger than a threshold of 2 spikes/s when driven by drifting-grating stimuli; this constrains our analysis to units that are stimulus activated. Second, because anatomical connections are retinotopic in the visual cortical areas, to maximize the number of measurable functional connections between areas, we selected units with receptive field centers at least 10 degrees away from the edge of the monitor. This selection criterion ensures that the neurons we study have visual receptive field locations that are stimulated by the visual stimulus and facilitates overlap between receptive fields for functional connectivity estimates. For validation, we analyzed the full functional network of all recorded units before applying any firing rate or receptive field criteria and found that the units that do not pass our selection criteria resulted in very low connectivity weights and thus fell into cluster 1, which was dominated by weak connections; therefore, these selection criteria are unlikely to influence our conclusions regarding the driver (cluster 2) and driven (cluster 3) modules. Third, when constructing the adjacency matrix from functional connectivity estimated from the jitter-corrected CCG, we did not base our analysis on the shape of the CCG (such as a sharp peak profile to select for putative monosynaptic connections). Instead, we used a permissive definition of directed functional weight (asymmetry of CCG around zero time offset) within a temporal window of ± 13 ms that likely includes polysynaptic effects. Fourth, we only studied the functional network using drifting-grating stimuli (presented at different orientations with many repeats) and gray screen periods (spontaneous); these visual inputs do not have a rich spatiotemporal structure that is present in more ecological visual input. Finally, our recordings are made in mice that are passively viewing the visual stimuli without being engaged in a task. This likely biases the functional network to reflect bottom-up sensory-driven signals rather than top-down signals from higher-order cortical regions. Indeed, during spontaneous conditions when no stimulus was shown to the mouse, the connectivity weights between the modules were substantially reduced (Figure 2), which is consistent with our previous findings (Siegle et al., 2021). Our results and findings should be interpreted with these caveats and constraints in mind.

Despite this, our study reveals a novel organizational principle in the mouse visual cortex in which subsets of neurons distributed across hierarchical levels share functional connectivity and participate in multi-area signal transmission modules. Independent measurements ranging from the composition of simple-like neurons to anatomical distributions and temporal dynamics are all consistent with the different processing stages of these modules, suggesting that these multi-area modules are a fundamental organization for visual sensory processing in the mouse. Our general approach can be applied to other brain regions and sensory systems to test the generalization of this organizational principle.

A method for identifying multi-regional functional modules

Community detection methods have been used in human brain imaging studies to identify distributed functional modules at the millimeter scale (Betzel et al., 2018; Power et al., 2011; Sporns and Betzel, 2016), but these techniques have rarely been applied to cellular resolution networks of the cortex (Billeh et al., 2014; Kiani et al., 2015). The clustering algorithm we developed in this study provides a new approach for identifying functional modules in spiking networks and could be helpful for dissecting network substructure in a variety of systems and contexts. Our method for identifying functional modules relied on first quantifying fast timescale functional connectivity in the recorded network (Smith and Kohn, 2008). We define a directed rather than undirected adjacency matrix of interactions between all simultaneously recorded neurons. Our unsupervised clustering method then identifies distinct neuronal populations in the network based on their shared patterns of inferred input and output connectivity. By allowing our algorithm to cluster neurons from all recorded cortical areas together, we uncovered modules that do not map directly onto single cortical areas or layers, although there are area and layer biases. Thus, our approach has the potential for identifying functionally relevant and interacting subpopulations of neurons that are not

Please cite this article in press as: Jia et al., Multi-regional module-based signal transmission in mouse visual cortex, Neuron (2022), https://doi.org/10.1016/j.neuron.2022.01.027

CellPress
OPEN ACCESS

constrained by standard anatomical parcellation schemes. This is critical, because many behavioral and cognitive operations are likely mediated by such distributed neuronal ensembles (Buzsáki, 2010; Hebb, 1949).

Because excitatory neurons are primarily responsible for inter-area projections (Harris and Shepherd, 2015), we performed a control analysis in which we removed putative inhibitory cells from the functional networks and recomputed the various metrics for the driver and driven modules. Based on spike waveform duration (Jia et al., 2019; Mitchell et al., 2007), we identified ~19% of all units as fast-spiking (FS) (putative inhibitory neurons) as opposed to regular spiking (RS) (putative excitatory cells). FS units were present in both driver and driven modules. We found that the anatomical distribution and properties of the RS-only network were not different from the full network that included both RS and FS units (Figure S6).

Comparison with previous work

Previous work has proposed a canonical model of communication between cortical layers and areas based on the laminar-specific feedforward and feedback connections (Bastos et al., 2012; Callaway, 1998; Douglas and Martin, 2004; Felleman and Van Essen, 1991; Harris et al., 2019; Markov et al., 2014b). Specifically, within a single column, inputs are received in the middle (primary sensory area) or superficial layers (higher cortical areas) and then propagated from superficial layers to deep layers. Between cortical areas, feedforward connections predominately originate from superficial layers, whereas feedback connections are mostly from deep layers. We find the driver module neurons are concentrated in the middle and superficial layers with their relative proportion gradually decreasing along the visual hierarchy, consistent with the anatomical distribution of feedforward projecting neurons (Felleman and Van Essen, 1991; Harris et al., 2019; Markov et al., 2014a). Driven neurons are more likely to be found in deep layers, receiving functional inputs from driver neurons in the same and other areas and forming recurrent connections with driven neurons in other areas. This distribution is similar to the anatomically defined feedback neurons. However, it is important to keep in mind that the anatomical connectivity provides the scaffold that constrains the functional network, but the functional network is stimulus and brain state dependent, which can change over time. Therefore, even though we observed similarity between the anatomical distribution of the functionally defined driver and driven modules and the distribution of neurons participating in the feedforward and feedback processes, it could be a consequence of the sensory-driven state of the functional network and may not always hold true under different behavioral and cognitive contexts.

Several theories of cortical processing, including predictive coding (Friston, 2005) and Ullman's counter-stream hypothesis (Ullman, 1995), posit segregated feedforward and feedback circuits. The functional modules we describe could have implications for these theories. Beyond feedforward and feedback circuits, the idea of separate sets of neurons involved in distributing versus integrating signals has not been sufficiently explored in a sensory processing hierarchy, where the dominant framework is that of sequential feedforward layers of processing (Riesenhuber and Poggio, 1999). Our findings could have interesting computational implications for the network architecture of cortex. Rather than building a hierarchical network with sequential stages that map onto areas, these models can include intercalated groups of neurons that participate in distinct subnetworks with different computational properties.

Future directions

The anatomical substrate for the functional modules we identified here is yet to be determined. In the mouse visual cortex, single V1 neurons often make branched projections to multiple higher-order visual areas (Han et al., 2018). If these projections were module specific, this anatomical connectivity could underlie the functional network structure we observe. There is evidence that feedforward and feedback circuits are mediated by separate sets of neurons (Berezovskii et al., 2011; Markov et al., 2014b) and that the rules of local connectivity between V1 neurons depend on the areas to which they project (Kim et al., 2018); these observations of anatomical compartmentalization could help explain the division between the driver and driven modules. Ultimately, causal perturbations will be necessary to confirm the functional relevance of these two modules and to work out their exact mechanisms of signal transmission. Because the modules we describe are highly distributed across different cortical areas and layers, precise labeling and manipulation of neurons in the modules will be technically challenging. Cell type- and layer-specific genetic tools such as Cre lines are currently insufficient to study these modules. Photo-stimulation approaches that allow random access activation of both distributed and intermingled cells could be useful for this purpose (Marshel et al., 2019). This method would require that the driver and driven modules be identified prior to photo-stimulation. In principle this could be done using large-scale two-photon calcium imaging to image activity in distributed populations of neurons across visual cortical areas.

The modules we described must be constrained by anatomy; yet there are elaborate anatomical projections within and between brain regions that are likely to engage in the flexible and dynamic routing of information in a task- and state-dependent manner (Olshausen et al., 1993). Future studies can use more complex stimuli to probe how modular network structures depend on different types of inputs, especially naturalistic visual stimuli such as spatiotemporal movies and visuomotor feedback (e.g., virtual reality) (Saleem et al., 2018). In addition, behavioral states and cognitive task conditions might also restructure the functional networks in the mouse visual cortex. It remains to be determined whether such top-down signals take advantage of the bottom-up functional structure we observe here, perhaps by modulating the gain of transmission between the driver and driven modules. It is also unknown whether or how top-down signals might reconfigure circuits to engage separate subsets of neurons than those comprising the driver and driven modules we observe during passive bottom-up stimulation. Finally, the anatomical hierarchy continues well beyond visual cortex (Harris et al., 2019). More complete sampling of this network will be necessary to determine the extent to which driver and driven modules extend throughout the cortical and thalamic system of the mouse, and more generally, whether similar functional modules exist in the brains of other species.

Please cite this article in press as: Jia et al., Multi-regional module-based signal transmission in mouse visual cortex, Neuron (2022), https://doi.org/10.1016/j.neuron.2022.01.027

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Mice
- METHOD DETAILS
  - Data collection
  - Surgical methods
  - Intrinsic signal imaging
  - Habituation
  - Electrophysiology experiments
  - Cortical area targeting
  - Visual stimulus
  - Spike sorting
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Dataset
  - Quantification and statistical analysis
  - Visual receptive fields
  - Peristimulus time histogram (PSTH)
  - Functional connectivity
  - Clustering
  - Modulation index
  - Module distribution
  - Graph creation
  - Divergence and convergence degree
  - Temporal dynamics analysis
  - Area in-out index
  - Layer definition
  - Simulations to test mathematical relationship between PSTH shape and CCG sharp peak

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.neuron.2022.01.027.

ACKNOWLEDGMENTS

We thank the Allen Institute founder, Paul G. Allen, for his vision, encouragement, and support. We thank Transgenic Colony Management for mouse breeding and Laboratory Animal Services for mouse import and wellness care. We thank the Neurosurgery and Behavior Team for surgical procedures and habituation. We thank Shiella Caldejon for running intrinsic signal imaging experiments and Rusty Nicovich and Kiet Ngo for collecting optical projection tomography data. We thank Yazan Billeh, Uygar Sumbul, and Daniel Denman for helpful discussions, and Daniel Denman, Hannah Choi, Marina Garrett, Gabe Ocker, and Adam Kohn for helpful feedback on the manuscript.

AUTHOR CONTRIBUTIONS

Conceptualization, X.J., J.H.S., and S.R.O.; investigation, validation, and methodology, X.J., J.H.S., S.D., G.H., and T.K.R.; formal analyses, X.J.; visualization, X.J., S.R.O.; original draft written by X.J. and S.R.O. with input and editing from J.H.S. and C.K.; All co-authors reviewed the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: May 21, 2021
Revised: November 20, 2021
Accepted: January 22, 2022
Published: February 9, 2022

REFERENCES

Abeles, M. (1991). Corticonics: Neural Circuits of the Cerebral Cortex (New York: Cambridge University Press).

Averbeck, B.B., Latham, P.E., and Pouget, A. (2006). Neural correlations, population coding and computation. Nat. Rev. Neurosci. *7*, 358–366.

Barone, P., Batardiere, A., Knoblauch, K., and Kennedy, H. (2000). Laminar distribution of neurons in extrastriate areas projecting to visual areas V1 and V4 correlates with the hierarchical rank and indicates the operation of a distance rule. J. Neurosci. *20*, 3263–3281.

Barral, J., Wang, X.J., and Reyes, A.D. (2019). Propagation of temporal and rate signals in cultured multilayer networks. Nat. Commun. *10*, 3969.

Bastos, A.M.M., Usrey, W.M.M., Adams, R.A.A., Mangun, G.R.R., Fries, P., and Friston, K.J.J. (2012). Canonical microcircuits for predictive coding. Neuron *76*, 695–711.

Bastos, A.M., Vezoli, J., Bosman, C.A., Schoffelen, J.M., Oostenveld, R., Dowdall, J.R., De Weerd, P., Kennedy, H., and Fries, P. (2015). Visual areas exert feedforward and feedback influences through distinct frequency channels. Neuron *85*, 390–401.

Bastos, A.M., Loonis, R., Kornblith, S., Lundqvist, M., and Miller, E.K. (2018). Laminar recordings in frontal cortex suggest distinct layers for maintenance and control of working memory. Proc. Natl. Acad. Sci. USA *115*, 1117–1122.

Berezovskii, V.K., Nassi, J.J., and Born, R.T. (2011). Segregation of feedforward and feedback projections in mouse visual cortex. J. Comp. Neurol. *519*, 3672–3683.

Betzel, R.F., Medaglia, J.D., and Bassett, D.S. (2018). Diversity of meso-scale architecture in human and non-human connectomes. Nat. Commun. *9*, 346.

Billeh, Y.N., Schaub, M.T., Anastassiou, C.A., Barahona, M., and Koch, C. (2014). Revealing cell assemblies at multiple levels of granularity. J. Neurosci. Methods *236*, 92–106.

Buzsáki, G. (2004). Large-scale recording of neuronal ensembles. Nat. Neurosci. *7*, 446–451.

Buzsáki, G. (2010). Neural syntax: cell assemblies, synapsembles, and readers. Neuron *68*, 362–385.

Callaway, E.M. (1998). Local circuits in primary visual cortex of the macaque monkey. Annu. Rev. Neurosci. *21*, 47–74.

Chen, R., Wang, F., Liang, H., and Li, W. (2017). Synergistic Processing of Visual Contours across Cortical Layers in V1 and V2. Neuron *96*, 1388–1402.e4.

Chung, J.E., Magland, J.F., Barnett, A.H., Tolosa, V.M., Tooker, A.C., Lee, K.Y., Shah, K.G., Felix, S.H., Frank, L.M., and Greengard, L.F. (2017). A Fully Automated Approach to Spike Sorting. Neuron *95*, 1381–1394.e6.

Cunningham, J.P., and Yu, B.M. (2014). Dimensionality reduction for large-scale neural recordings. Nat. Neurosci. *17*, 1500–1509.

DiCarlo, J.J., Zoccolan, D., and Rust, N.C. (2012). How does the brain solve visual object recognition? Neuron *73*, 415–434.

Douglas, R.J., and Martin, K.A.C. (2004). Neuronal circuits of the neocortex. Annu. Rev. Neurosci. *27*, 419–451.

D'Souza, R.D., Wang, Q., Ji, W., Meier, A.M., Kennedy, H., Knoblauch, K., and Burkhalter, A. (2022). Hierarchical and nonhierarchical features of the mouse visual cortical network. Nat Commun *13*. https://doi.org/10.1038/s41467-022-28035-y.

Felleman, D.J., and Van Essen, D.C. (1991). Distributed hierarchical processing in the primate cerebral cortex. Cereb. Cortex *1*, 1–47.

Please cite this article in press as: Jia et al., Multi-regional module-based signal transmission in mouse visual cortex, Neuron (2022), https://doi.org/10.1016/j.neuron.2022.01.027

CellPress
OPEN ACCESS

Fornito, A., Zalesky, A., and Bullmore, E.T. (2016). Fundamentals of Brain Network Analysis (Elsevier Inc.).

Friston, K. (2005). A theory of cortical responses. Philos. Trans. R. Soc. B Biol. Sci. *360*, 815–836.

Gămănuț, R., Kennedy, H., Toroczkai, Z., Ercsey-Ravasz, M., Van Essen, D.C., Knoblauch, K., and Burkhalter, A. (2018). The Mouse Cortical Connectome, Characterized by an Ultra-Dense Cortical Graph, Maintains Specificity by Distinct Connectivity Profiles. Neuron *97*, 698–715.e10.

Garrett, M.E., Nauhaus, I., Marshel, J.H., and Callaway, E.M. (2014). Topography and areal organization of mouse visual cortex. J. Neurosci. *34*, 12587–12600.

Gilbert, C.D., and Li, W. (2013). Top-down influences on visual processing. Nat. Rev. Neurosci. *14*, 350–363.

Goldey, G.J., Roumis, D.K., Glickfeld, L.L., Kerlin, A.M., Reid, R.C., Bonin, V., Schafer, D.P., and Andermann, M.L. (2014). Removable cranial windows for long-term imaging in awake mice. Nat. Protoc. *9*, 2515–2538.

Goris, R.L.T., Movshon, J.A., and Simoncelli, E.P. (2014). Partitioning neuronal variability. Nat. Neurosci. *17*, 858–865.

Hagberg, A., Schult, D., and Swart, P. (2008). Exploring Network Structure, Dynamics, and Function using Networkx. Report from the U.S. Department of Energy Office of Scientific and Technical Information. https://www.osti.gov/biblio/960616.

Han, Y., Kebschull, J.M., Campbell, R.A.A., Cowan, D., Imhof, F., Zador, A.M., and Mrsic-Flogel, T.D. (2018). The logic of single-cell projections from visual cortex. Nature *556*, 51–56.

Harris, C.R., Millman, K.J., van der Walt, S.J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N.J., et al. (2020). Array programming with NumPy. Nature *585*, 357–362.

Harris, K.D., and Shepherd, G.M.G.G. (2015). The neocortical circuit: themes and variations. Nat. Neurosci. *18*, 170–181.

Harris, J.A., Mihalas, S., Hirokawa, K.E., Whitesell, J.D., Choi, H., Bernard, A., Bohn, P., Caldejon, S., Casal, L., Cho, A., et al. (2019). Hierarchical organization of cortical and thalamic connectivity. Nature *575*, 195–202.

Harrison, M.T., and Geman, S. (2009). A rate and history-preserving resampling algorithm for neural spike trains. Neural Comput. *21*, 1244–1258.

Hebb, D.O. (1949). The Organization of Behavior; A Neuropsychological Theory, *63*, p. 633.

Hill, D.N., Mehta, S.B., and Kleinfeld, D. (2011). Quality metrics to accompany spike sorting of extracellular signals. J. Neurosci. *31*, 8699–8705.

Hubel, D.H. (1988). Eye, Brain and Vision, volume 22 of Scientific American Library. (Sci. Am. Press.).

Hubel, D.H., and Wiesel, T.N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. J. Physiol. *160*, 106–154.

Jia, X., Tanabe, S., and Kohn, A. (2013). γ and the coordination of spiking activity in early visual cortex. Neuron *77*, 762–774.

Hunter, J.D. (2007). Matplotlib: A 2D graphics environment. Computing in science & engineering *9*, 90–95.

Jia, X., Siegle, J.H., Bennett, C., Gale, S.D., Denman, D.J., Koch, C., and Olsen, S.R. (2019). High-density extracellular probes reveal dendritic backpropagation and facilitate neuron classification. J. Neurophysiol. *121*, 1831–1847.

Jun, J.J., Steinmetz, N.A., Siegle, J.H., Denman, D.J., Bauza, M., Barbarits, B., Lee, A.K., Anastassiou, C.A., Andrei, A., Aydın, Ç., et al. (2017). Fully integrated silicon probes for high-density recording of neural activity. Nature *551*, 232–236.

Kiani, R., Cueva, C.J., Reppas, J.B., Peixoto, D., Ryu, S.I., and Newsome, W.T. (2015). Natural grouping of neural responses reveals spatially segregated clusters in prearcuate cortex. Neuron *85*, 1359–1373.

Kim, M.H., Znamenskiy, P., Iacaruso, M.F., and Mrsic-Flogel, T.D. (2018). Segregated Subnetworks of Intracortical Projection Neurons in Primary Visual Cortex. Neuron *100*, 1313–1321.e6.

Kohn, A., Jasper, A.I., Semedo, J.D., Gokcen, E., Machens, C.K., and Yu, B.M. (2020). Principles of Corticocortical Communication: Proposed Schemes and Design Considerations. Trends Neurosci. *43*, 725–737.

Kumar, A., Rotter, S., and Aertsen, A. (2010). Spiking activity propagation in neuronal networks: reconciling different perspectives on neural coding. Nat. Rev. Neurosci. *11*, 615–627.

Lu, T., Liang, L., and Wang, X. (2001). Temporal and rate representations of time-varying signals in the auditory cortex of awake primates. Nat. Neurosci. *4*, 1131–1138.

Markov, N.T., Ercsey-Ravasz, M.M., Ribeiro Gomes, A.R., Lamy, C., Magrou, L., Vezoli, J., Misery, P., Falchier, A., Quilodran, R., Gariel, M.A., et al. (2014a). A weighted and directed interareal connectivity matrix for macaque cerebral cortex. Cereb. Cortex *24*, 17–36.

Markov, N.T., Vezoli, J., Chameau, P., Falchier, A., Quilodran, R., Huissoud, C., Lamy, C., Misery, P., Giroud, P., Ullman, S., et al. (2014b). Anatomy of hierarchy: Feedforward and feedback pathways in macaque visual cortex (John Wiley & Sons, Ltd).

Marshel, J.H., Kim, Y.S., Machado, T.A., Quirin, S., Benson, B., Kadmon, J., Raja, C., Chibukhchyan, A., Ramakrishnan, C., Inoue, M., et al. (2019). Cortical layer-specific critical dynamics triggering perception. Science *365*, eaaw5202.

Matteucci, G., Bellacosa Marotti, R., Riggi, M., Rosselli, F.B., and Zoccolan, D. (2019). Nonlinear processing of shape information in rat lateral extrastriate cortex. J. Neurosci. *39*, 1649–1670.

McKinney, W. (2010). Data structures for statistical computing in python. Proceedings of the 9th Python in Science Conference, 51–56.

Mitchell, J.F., Sundberg, K.A., and Reynolds, J.H. (2007). Differential attention-dependent response modulation across cell classes in macaque visual area V4. Neuron *55*, 131–141.

Monti, S., Tamayo, P., Mesirov, J., Golub, T., Sebastiani, P., Kohane, I.S., and Ramoni, M.F. (2003). Consensus Clustering: A Resampling-Based Method for Class Discovery and Visualization of Gene Expression Microarray Data. Machine Learning *52*, 91–118.

Mountcastle, V.B., Poggio, G.F., and Werner, G. (1963). The relation of thalamic cell response to peripheral stimuli varied over an intensive continuum. J. Neurophysiol. *26*, 807–834.

Musall, S., Kaufman, M.T., Juavinett, A.L., Gluf, S., and Churchland, A.K. (2019). Single-trial neural dynamics are dominated by richly varied movements. Nat. Neurosci. *22*, 1677–1686.

Olshausen, B.A., Anderson, C.H., and Van Essen, D.C. (1993). A neurobiological model of visual attention and invariant pattern recognition based on dynamic routing of information. J. Neurosci. *13*, 4700–4719.

Orban, G.A. (2008). Higher order visual processing in macaque extrastriate cortex. Physiol. Rev. *88*, 59–89.

Pachitariu, M., Steinmetz, N.A., Kadir, S.N., Carandini, M., and Harris, K.D. (2016). Fast and accurate spike sorting of high-channel count probes with KiloSort. Adv. Neural Inf. Process. Syst. *29*, 4448–4456.

Parker, A.J., and Newsome, W.T. (1998). Sense and the single neuron: probing the physiology of perception. Annu. Rev. Neurosci. *21*, 227–277.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al. (2011). Scikit-learn: Machine learning in Python. J. Mach. Learn. Res. *12*, 2825–2830.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al. (2011). Scikit-learn: Machine learning in Python. Journal of machine learning research *12*, 2825–2830.

Peirce, J.W. (2007). PsychoPy-Psychophysics software in Python. J. Neurosci. Methods *162*, 8–13.

Perin, R., Berger, T.K., and Markram, H. (2011). A synaptic organizing principle for cortical neuronal groups. Proc. Natl. Acad. Sci. USA *108*, 5419–5424.

Perkel, D.H., and Bullock, T. (1968). Neural Coding: A Report based on an NRP Work Session. Neurosci. Research Prog. Bull. *6*, 220–349.

Please cite this article in press as: Jia et al., Multi-regional module-based signal transmission in mouse visual cortex, Neuron (2022), https://doi.org/10.1016/j.neuron.2022.01.027

Pham, D.T., Dimov, S.S., and Nguyen, C.D. (2005). Selection of K in K-means clustering. Proc. Inst. Mech. Eng. Part C. J. Mech. Eng. Sci. *219*, 103–119.

Power, J.D., Cohen, A.L., Nelson, S.M., Wig, G.S., Barnes, K.A., Church, J.A., Vogel, A.C., Laumann, T.O., Miezin, F.M., Schlaggar, B.L., and Petersen, S.E. (2011). Functional network organization of the human brain. Neuron *72*, 665–678.

Reid, R.C., and Alonso, J.M. (1995). Specificity of monosynaptic connections from thalamus to visual cortex. Nature *378*, 281–284.

Reyes, A.D. (2003). Synchrony-dependent propagation of firing rate in iteratively constructed networks in vitro. Nat. Neurosci. *6*, 593–599.

Riesenhuber, M., and Poggio, T. (1999). Hierarchical models of object recognition in cortex. Nat. Neurosci. *2*, 1019–1025.

Saleem, A.B., Diamanti, E.M., Fournier, J., Harris, K.D., and Carandini, M. (2018). Coherent encoding of subjective spatial position in visual cortex and hippocampus. Nature *562*, 124–127.

Salkoff, D.B., Zagha, E., McCarthy, E., and McCormick, D.A. (2020). Movement and Performance Explain Widespread Cortical Activity in a Visual Detection Task. Cereb. Cortex *30*, 421–437.

Seabold, S., and Perktold, J. (2010). Statsmodels: Econometric and Statistical Modeling with Python.

Semedo, J.D., Zandvakili, A., Machens, C.K., Yu, B.M., and Kohn, A. (2019). Cortical Areas Interact through a Communication Subspace. Neuron *102*, 249–259.e4.

Siegle, J.H., López, A.C., Patel, Y.A., Abramov, K., Ohayon, S., and Voigts, J. (2017). Open Ephys: an open-source, plugin-based platform for multichannel electrophysiology. J. Neural Eng. *14*, 045003.

Siegle, J.H., Jia, X., Durand, S., Gale, S., Bennett, C., Graddis, N., Heller, G., Ramirez, T.K., Choi, H., Luviano, J.A., et al. (2021). Survey of spiking in the mouse visual system reveals functional hierarchy. Nature *592*, 86–92.

Smith, M.A., and Kohn, A. (2008). Spatial and temporal scales of neuronal correlation in primary visual cortex. J. Neurosci. *28*, 12591–12603.

Song, S., Sjöström, P.J., Reigl, M., Nelson, S., and Chklovskii, D.B. (2005). Highly nonrandom features of synaptic connectivity in local cortical circuits. PLoS Biol. *3*, e68.

Spielman, D.A. (2007). Spectral Graph Theory and its Applications. 48th Annual IEEE Symposium on Foundations of Computer Science (FOCS'07), 29–38.

Sporns, O., and Betzel, R.F. (2016). Modular Brain Networks. Annu. Rev. Psychol. *67*, 613–640.

Steinmetz, N.A., Aydin, C., Lebedeva, A., Okun, M., Pachitariu, M., Bauza, M., Beau, M., Bhagat, J., Böhm, C., Broux, M., et al. (2021). Neuropixels 2.0: A miniaturized high-density probe for stable, long-term brain recordings. Science *372*, eabf4588.

Stoelzel, C.R., Bereshpolova, Y., and Swadlow, H.A. (2009). Stability of thalamocortical synaptic transmission across awake brain states. J. Neurosci. *29*, 6851–6859.

Stringer, C., Pachitariu, M., Steinmetz, N., Reddy, C.B., Carandini, M., and Harris, K.D. (2019). Spontaneous behaviors drive multidimensional, brainwide activity. Science *364*, 255.

Tibshirani, R., Walther, G., and Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. J. R. Stat. Soc. Ser. B. Stat. Methodol. *63*, 411–423.

Tononi, G., Edelman, G.M., and Sporns, O. (1998). Complexity and coherency: integrating information in the brain. Trends Cogn. Sci. *2*, 474–484.

Ullman, S. (1995). Sequence seeking and counter streams: a computational model for bidirectional information flow in the visual cortex. Cereb. Cortex *5*, 1–11.

van Kerkoerle, T., Self, M.W., Dagnino, B., Gariel-Mathis, M.-A., Poort, J., van der Togt, C., and Roelfsema, P.R. (2014). Alpha and gamma oscillations characterize feedback and feedforward processing in monkey visual cortex. Proc. Natl. Acad. Sci. USA *111*, 14332–14341.

Vesuna, S., Kauvar, I.V., Richman, E., Gore, F., Oskotsky, T., Sava-Segal, C., Luo, L., Malenka, R.C., Henderson, J.M., Nuyujukian, P., et al. (2020). Deep posteromedial cortical rhythm in dissociation. Nature *586*, 87–94.

Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., et al.; SciPy 1.0 Contributors (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat. Methods *17*, 261–272.

Vogels, T.P., and Abbott, L.F. (2005). Signal propagation and logic gating in networks of integrate-and-fire neurons. J. Neurosci. *25*, 10786–10795.

Wang, Q., and Burkhalter, A. (2007). Area map of mouse visual cortex. J. Comp. Neurol. *502*, 339–357.

Wekselblatt, J.B., Flister, E.D., Piscopo, D.M., and Niell, C.M. (2016). Large-scale imaging of cortical dynamics during sensory perception and behavior. J. Neurophysiol. *115*, 2852–2866.

Wong, Y.T., Fabiszak, M.M., Novikov, Y., Daw, N.D., and Pesaran, B. (2016). Coherent neuronal ensembles are rapidly recruited when making a look-reach decision. Nat. Neurosci. *19*, 327–334.

Yoshimura, Y., Dantzker, J.L.M., and Callaway, E.M. (2005). Excitatory cortical neurons form fine-scale functional networks. Nature *433*, 868–873.

Zandvakili, A., and Kohn, A. (2015). Coordinated Neuronal Activity Enhances Corticocortical Communication. Neuron *87*, 827–839.

Zavitz, E., and Price, N.S.C. (2019). Understanding Sensory Information Processing Through Simultaneous Multi-area Population Recordings. Front. Neural Circuits *12*, 115.

Please cite this article in press as: Jia et al., Multi-regional module-based signal transmission in mouse visual cortex, Neuron (2022), https://doi.org/10.1016/j.neuron.2022.01.027

CellPress
OPEN ACCESS

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Passive viewing dataset collected with Neuropixels probes	Allen Institute (https://allensdk.readthedocs.io/en/latest/visual_coding_neuropixels.html)	2019 release
NWB files containing all data used in this study	Zenodo	https://doi.org/10.5281/zenodo.5701556
Experimental models: Organisms/strains		
Mouse: Pvalb-IRES-Cre;Ai32	Jackson Laboratory	N/A
Mouse: Sst-IRES-Cre;Ai32	Jackson Laboratory	N/A
Mouse: Vip-IRES-Cre;Ai32	Jackson Laboratory	N/A
Mouse: C57BL/6J	Jackson Laboratory	N/A
Software and algorithms		
Python	Python Software Foundation	3.6
AllenSDK	https://github.com/AllenInstitute/AllenSDK	2.11.2
Anaconda	https://www.anaconda.com/	N/A
SciPy	https://scipy.org/ Virtanen et al., 2020	1.4.1
seaborn	https://seaborn.pydata.org/	0.10.0
matplotlib	https://matplotlib.org/ Hunter, 2007	3.2.1
NumPy	https://numpy.org/ Harris et al., 2020	1.18.1
networkx	https://networkx.org/	2.4
jupyterlab	https://jupyter.org/	2.1.1
Pandas	https://pandas.pydata.org/ McKinney, 2010	1.0.3
scikit-learn	https://scikit-learn.org/stable/ Pedregosa et al., 2011	0.22.2
statsmodels	https://www.statsmodels.org/stable/index.html	0.10.2
Open Ephys GUI	open-ephys/plugin-GUI open-ephys-plugins/neuropixels-3a open-ephys-plugins/neuropixels-PXI	N/A
Kilosort2	https://github.com/MouseLand/Kilosort/releases/tag/v2.0 Stringer et al., 2019	2.0
pyNWB	https://pynwb.readthedocs.io/en/stable/	2.0
AIBSOPT	https://github.com/alleninstitute/AIBSOPT	N/A
Custom analysis code	https://github.com/jiaxx/modular_network	https://doi.org/10.5281/zenodo.5889134
Other		
Neuropixels probes	https://www.neuropixels.org/	Neuropixels 3a and Neuropixels 1.0

RESOURCE AVAILABILITY

Lead contact

Further information and resources should be directed to and will be fulfilled by the lead contact, Xiaoxuan Jia (jxiaoxuan@gmail.com).

Materials availability

This study did not generate new unique reagents.

Data and code availability

The majority of the data in this study (13 of 19 experiments) were publicly released as an open dataset on the Allen Institute website in October 2019, and are available via the AllenSDK (https://allensdk.readthedocs.io/en/latest/visual_coding_neuropixels.html). All of the raw data from 19 mice are available on Zenodo (https://doi.org/10.5281/zenodo.5701556) in the form of NWB files. Analysis codes are publicly available on GitHub at https://github.com/jiaxx/modular_network (https://doi.org/10.5281/zenodo.5889134).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mice

Mice were maintained in the Allen Institute animal facility and used in accordance with protocols approved by the Allen Institute's Institutional Animal Care and Use Committee. Four mouse genotypes were used: wild-type C57BL/6J (Jackson Laboratories) (n = 11) or Pvalb-IRES-Cre (n = 1), Vip-IRES-Cre (n = 2), and Sst-IRES-Cre (n = 5) mice bred in-house and crossed with an Ai32 channelrhodopsin reporter line. Following surgery, all mice were single-housed and maintained on a reverse 12 h light cycle. All experiments were performed during the dark cycle.

METHOD DETAILS

Data collection

Experimental data collection followed the procedures described in Siegle et al. (2021). A summary of these methods is provided below.

Surgical methods

All surgical methods used here are the same as Siegle et al. (2021). Briefly, each animal was implanted with a titanium headframe that permits head fixation in a reproducible configuration. To implant the headframe, mice were first anesthetized with 5% isoflurane (1-3 min) and placed in a stereotaxic frame (Model #1900, Kopf). Isoflurane was maintained at 1.5%–2.5% for surgery. Body temperature was held at 37.5°C. Carprofen was administered for pain management (5-10 mg/kg, S.C.) and atropine was used to suppress bronchial secretions and regulate heart rhythm (0.02-0.05 mg/kg, S.C.). The headframe was placed on the skull and fixed in place with white C&B Metabond (Parkell). Once the Metabond was dry, the mouse was placed in a custom clamp to position the skull at a rotated angle of 20°, to facilitate the creation of a craniotomy over visual cortex. A circular piece of skull 5 mm in diameter was removed, and a durotomy was performed. The brain was covered by a 5 mm diameter circular glass coverslip, with a 1 mm lip extending over the intact skull. The bottom of the coverslip was coated with a layer of silicone to reduce adhesion to the brain surface. At the end of the procedure, but prior to recovery from anesthesia, the mouse was transferred to a photo-documentation station to capture a spatially registered image of the cranial window.

Neuropixels recordings were performed at least four weeks after the initial surgery. The cranial coverslip was removed and replaced with an insertion window containing holes aligned to six cortical visual areas. First, the mouse was anesthetized with isoflurane (3%–5% induction and 1.5% maintenance, 100% O_2) and eyes were protected with ocular lubricant (I Drop, VetPLUS). Body temperature was maintained at 37.5°C (TC-1000 temperature controller, CWE, Incorporated). The cranial window was gently removed to expose the brain. An insertion window with holes for probe penetration was then placed in the headframe well and sealed with Metabond. An agarose mixture was injected under the window. The mixture consisted of 0.4 g high EEO Agarose (Sigma-Aldrich), 0.42 g Certified Low-Melt Agarose (Bio Rad), and 20.5 mL ACSF (135.0 mM NaCl, 5.4 mM KCl, 1.0 mM MgCl2, 1.8 mM CaCl2, 5.0 mM HEPES). This mixture was optimized to be firm enough to stabilize the brain with minimal probe drift, but pliable enough to permit the probes to pass through without bending. A layer of silicone oil (30,000 cSt, Aldrich) was added over the holes in the insertion window to prevent the agarose from drying. A 3D-printed plastic cap was screwed into the headframe well to keep out cage debris. At the end of this procedure, mice were returned to their home cages for 1-2 h prior to the Neuropixels recording session.

Intrinsic signal imaging

Intrinsic signal imaging (ISI) was performed approximately 15 days after the initial surgery and 25 days before the experiment. ISI was used to generate retinotopic maps of the visual field (or, in this case, the coordinate position on the stimulus monitor) to locations in each cortical area (Figure 1A). The maps were used to delineate functionally defined visual area boundaries to target Neuropixels probes to retinotopically defined locations in primary and higher order visual cortical areas (Garrett et al., 2014).

Habituation

Mice underwent two weeks of habituation in sound-attenuated training boxes containing a headframe holder, running wheel, and stimulus monitor. Each mouse was handled by the same operator during the 2-week period. In the first week, the operator introduced them to the running wheel, and head-fixed them with progressively longer durations each day. In the second week, mice were exposed to visual stimuli for 10 to 50 min per day. For two prior to recording mice underwent habituation sessions of 75 min and 100 min on the recording rig, in which they viewed a shortened version of the stimulus set shown during the experiment.

Please cite this article in press as: Jia et al., Multi-regional module-based signal transmission in mouse visual cortex, Neuron (2022), https://doi.org/10.1016/j.neuron.2022.01.027

CellPress
OPEN ACCESS

Electrophysiology experiments

Neural recordings used Neuropixels probes (Jun et al., 2017). Each probe has 960 recording sites, a subset of these (374 for "Neuropixels 3a" or 383 for "Neuropixels 1.0") can be configured for recording at any given time. The electrodes sites closest to the tip were used. The sites form a checkerboard pattern on a 70 mm wide x 10 mm long shank. Signals from each recording site are split into a spike band (30 kHz sampling rate, 500 Hz highpass filter) and an LFP band (2.5 kHz sampling rate, 1000 Hz lowpass filter).

The experimental rig was designed for six Neuropixels probes to be inserted into the brain approximately perpendicular to the surface of visual cortex (Siegle et al., 2021). Each probe was mounted on a 3-axis micromanipulator (New Scale Technologies, Victor, NY), which were in turn mounted on a solid aluminum plate (the probe cartridge). The mouse was placed on the running wheel and fixed to the headframe clamp. The tip of each probe was aligned to target the desired retinotopic region in each area. Brightfield photo-documentation images were taken with the probes fully retracted, after the probes reached the brain surface, and again after the probes were fully inserted. An IR dichroic mirror was placed in front of the right eye to allow an eyetracking camera to operate without interference from the visual stimulus. A black curtain was then lowered over the front of the rig, placing the mice in complete darkness except for the visual stimulus monitor.

Neuropixels data was acquired at 30 kHz (spike band) and 2.5 kHz (LFP band) using the Open Ephys GUI (Siegle et al., 2017). Gain settings of 500x and 250x were used for the spike band and LFP band, respectively. Each probe was either connected to a dedicated FPGA streaming data over Ethernet (Neuropixels 3a) or a PXIe card inside a National Instruments chassis (Neuropixels 1.0). Raw neural data was streamed to a compressed format for archiving which was extracted prior to analysis.

Cortical area targeting

To confirm the identity of the cortical visual areas, images of the probes taken during the experiment were compared to images of the brain surface vasculature taken during the ISI session (see above). Vasculature patterns were used to overlay the visual area map on an image of the brain surface with the probes inserted (Figure 1A). To maximize measurable functional connectivity across areas, we targeted the center of gaze in all areas (except for RL, which targeted the center of mass because of geometry) with overlapping receptive fields (RF) guided by a retinotopic map. Targeting was validated by mapping receptive fields of all sorted units with small Gabor patches presented at different locations on the screen (see below). All analysis was restricted to neurons with well-defined receptive fields within the screen boundaries.

Visual stimulus

Visual stimuli were generated with custom scripts based on PsychoPy (Peirce, 2007) and were displayed using an ASUS PA248Q LCD monitor, with 1920 × 1200 pixels (21.93 in wide, 60 Hz refresh rate). Stimuli were presented monocularly, and the monitor was positioned 15 cm from the mouse's right eye and spanned 120° x 95° of visual space prior to stimulus warping. Each monitor was gamma corrected and had a mean luminance of 50 cd/ m^2. To account for the close viewing angle of the mouse, a spherical warping was applied to all stimuli to ensure that the apparent size, speed, and spatial frequency were constant across the monitor as seen from the mouse's perspective.

Visual stimuli for receptive fields (RFs)

Receptive field location was mapped with small Gabor patches. This mapping stimulus consisting of 2 Hz, 0.04 cycles per degree drifting gratings (3 directions: 0°, 45°, 90°) with a 20° circular mask. Gabor patches were randomly presented at one of 81 locations on the screen (9 × 9 grid with 10° spacing) for 250 ms at a time, with no blank interval.

Visual stimuli for current source density (CSD)

Current source density for layer estimation used the full-field flash stimuli (a series of dark or light full field image with luminance = 100 cd/ m^2) lasting 250 ms each and separated by a 1.75 s inter-trial interval.

Visual stimuli for functional connectivity

Functional connectivity during the stimulus-driven condition was measured using drifting grating stimuli, which were presented at 4 directions (0°, 45°, 90°, 135°), with temporal frequency equal to 2 cycle/sec and contrast equal to 0.8. In each trial, the grating is presented for 2 s followed by 1 s gray screen. Each condition was presented for 75-100 trials.

Spike sorting

Prior to spike sorting, the spike-band data passed through 4 steps: DC offset removal, median subtraction, filtering, and whitening. First, the median value of each channel was subtracted to center the signals around zero. Next, the median across channels was subtracted to remove common-mode noise. The median-subtracted data file is the input to the Kilosort2 MATLAB package (https://github.com/MouseLand/Kilosort), which applies a 150 Hz high-pass filter, followed by whitening in blocks of 32 channels. The filtered, whitened data is saved to a separate file for the spike sorting step.

Kilosort2 was used to identify spike times and assign spikes to individual units (Stringer et al., 2019). Kilosort2 attempts to model the complete dataset as a sum of spike "templates." The shape and locations of each template is iteratively refined until the data can be accurately reconstructed from a set of N templates at M spike times, with each individual template scaled by an amplitude, a. A critical feature of Kilosort2 is that it allows templates to change their shape over time, to account for the motion of neurons relative to the probe over the course of the experiment. Stabilizing the brain using an agarose-filled plastic window virtually eliminated

Please cite this article in press as: Jia et al., Multi-regional module-based signal transmission in mouse visual cortex, Neuron (2022), https://doi.org/10.1016/j.neuron.2022.01.027

probe motion associated with animal running, but slow drift of the probe over ~3 h experiments is still observed. Kilosort2 is able to accurately track units as they move along the probe axis, eliminating the need for the manual merging step that was required with the original version of Kilosort (Pachitariu et al., 2016). The spike-sorting step runs in approximately real time (~3 h per session) using a dual-processor Intel 4-core, 2.6 GHz workstation with an NVIDIA GTX 1070 GPU. We used the default parameters in Kilosort2, with an initial threshold of 12, and a final-pass threshold of 8.

The Kilosort2 algorithm will occasionally fit a template to the residual left behind after another template has been subtracted from the original data, resulting in double-counted spikes. This can create the appearance of an artificially high number of ISI violations for one unit or artificially high zero-time-lag synchrony between nearby units. To eliminate the possibility that this artificial synchrony will contaminate data analysis, the outputs of Kilosort2 are post-processed to remove spikes with peak times within 5 samples (0.16 ms) and peak waveforms within 5 channels (~50 microns).

Kilosort2 generates templates of a fixed length (2 ms) that matches the time course of an extracellularly detected spike waveform. However, there are no constraints on template shape, which means that the algorithm often fits templates to voltage fluctuations with characteristics that could not physically result from the current flow associated with an action potential. The units associated with these templates are considered "noise," and are automatically filtered out based on 3 criteria: spread (single channel, or > 25 channels), shape (no peak and trough, based on wavelet decomposition), or multiple spatial peaks (waveforms are non-localized along the probe axis).

Following the spike sorting step, data for each session was uploaded to the Allen Institute Laboratory Information Management System (LIMS). Each dataset was run through the same series of processing steps using a set of project-specific workflows (AllenSDK v1.0.2) in order to generate NeurodataWithoutBorders (NWB) files used for further analysis.

QUANTIFICATION AND STATISTICAL ANALYSIS

Dataset

In total, units from 19 mice were included in our functional connectivity analysis. Spike sorting, quality control, and preprocessing steps followed the same procedures as Siegle et al. (2021). 13 out of 19 of these datasets were previously released on the Allen Institute website via the AllenSDK (https://github.com/AllenInstitute/AllenSDK). All of the raw data from 19 mice are available on Zenodo (https://doi.org/10.5281/zenodo.5701556) in the form of Neurodata Without Borders (NWB) files. On average, 632 ± 18 sorted cortical units were simultaneously recorded in each mouse. We set a firing rate threshold to select units for functional connectivity analysis. Firing rate (FR) was defined as the average number of spikes in a window from 50 ms to 500 ms after the onset of the drifting gratings stimulus. Only units with mean FR > 2 spikes/second were used for pairwise cross-correlogram (CCG) calculation, which resulted in an average of 356 ± 7 units in each mouse (n = 6773 units in total). Because functional connectivity varies with receptive field position (Jia et al., 2013), we further constrained the dataset to include units with receptive field centers at least 10 degree away from the edge of the monitor (see Visual receptive fields below. After filtering by receptive field location, we ended up with 184 ± 8 per mouse used for the final clustering procedure (n = 3487 units in total). After applying clustering on the functional connectivity matrix constrained by both FR and RF location in each mouse, the total numbers of units belonging to each cluster were: n_cluster1 = 1386, n_cluster2 = 1131, n_cluster3 = 970.

Quantification and statistical analysis

All analyses were performed in Python. The main analysis packages used in this paper were Scipy (Virtanen et al., 2020), scikit-learn (Pedregosa et al., 2011), statsmodels (Seabold and Perktold, 2010), and networkx (Hagberg et al., 2008). Error bars, unless otherwise specified, were computed as standard error of the mean. When comparing the difference between two independent variables, if their distribution is Gaussian like (normality test), we used Student's t test; if their distribution is non-Gaussian, we used a rank sum test. When testing whether a distribution is significantly different from 0, we used a one-sample t test. When comparing variables between modules across cortical areas, we used two-way analysis of variance (ANOVA) to assess both the main effect between modules and whether there is any interaction across areas. When comparing similarity to the previously established anatomical visual hierarchy in mouse (Harris et al., 2019), we calculated the correlation between our measured variable (e.g., first spike latency) and the previously calculated anatomical hierarchy score (V1: −0.50, RL: −0.14, LM: −0.13, AL: 0.00, PM: 0.12, AM: 0.29), using Spearman's correlation to estimate the rank order significance. Statistical details and p-values can be found in the Results section or figure legends.

Visual receptive fields

Receptive fields were mapped with Gabor patches (20 degree each; 3 different orientations (0, 45, 90), temporal frequency = 2 cyc/s, spatial frequency = 0.04 cyc/deg) shown randomly at 81 different locations (9 × 9 grid, 10° separation between pixel centers) with gray background on a 120° x 95° monitor (1920 × 1200 pixels, 21.93 inches wide, 60 Hz refresh rate). The receptive field map (RF) for one unit is defined as the mean 2D histogram of spike counts at each of 81 locations, each pixel covers a 10° x 10° square. The receptive field was then thresholded at 20% of maximum response to remove potential noisy pixels. Then, a 2D Gaussian

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\left[\frac{(x-\mu_x)^2}{2\sigma_x^2} + \frac{(y-\mu_y)^2}{2\sigma_y^2}\right]}$$

was fit to the thresholded visual receptive map to estimate the center of the receptive field location.

Peristimulus time histogram (PSTH)

To visualize the temporal dynamics of a neuronal population (Figure 1, Figure 3, Figure 4, and Figure 7), the activity of each neuron was binned at 1 ms, averaged across trials (n = 75), smoothed with a Gaussian filter with standard deviation of 3 ms, baseline subtracted (baseline period from 0 to 0.03 s relative to stimulus onset), and normalized by dividing the maximum of the response between 0 to 1.5 s after stimulus onset. The normalized PSTHs of individual neuron were averaged within a neuronal population; the error bars indicate standard error of the mean across neurons.

Functional connectivity

We analyzed functional interactions between pairs of simultaneously recorded neurons by calculating the spike train cross-correlogram (CCG) (Jia et al., 2013; Smith and Kohn, 2008; Zandvakili and Kohn, 2015). For a pair of neurons with spike train x_1 and x_2, the CCG is defined as:

$$\mathrm{CCG}(\tau) = \frac{\frac{1}{M}\sum_{i=1}^{M}\sum_{t=1}^{N} x_1^i(t)x_2^i(t+\tau)}{\theta(\tau)\sqrt{\lambda_1\lambda_2}}$$

where M is the number of trials, N is the number of bins in the trial, x_1^i and x_2^i are the spike trains of the two units on trial i, τ is the time lag relative to reference spikes, and λ_1 and λ_2 are the mean firing rates of the two units. The CCG is essentially a sliding dot product between two spike trains. $\theta(\tau)$ is the triangular function which corrects for the overlapping time bins caused by the sliding window. To correct for firing rate dependency, we normalized the CCG by the geometric mean spike rate. An individually normalized CCG is computed separately for each drifting grating orientation and averaged across orientations to obtain the CCG for each pair of units.

The jitter-corrected CCG was created by subtracting the expected value of CCGs produced from a resampled version of the original dataset with spike times randomly perturbed (jittered) within the jitter window (Harrison and Geman, 2009; Smith and Kohn, 2008). Given two discretized spike trains, X and Y, of two neurons with dimension $T \times N$, where $X_{t,n}$ is the number of spikes in bin t (width = 1 ms) from trial n for spike train X, and $Y_{t,n}$ is the number of spikes in bin t (width = 1 ms) from trial n for spike train Y. The PSTH of X is represented as

$$PSTH_t(X) := \frac{1}{N}\sum_{n=1}^{N} X_{t,n}$$

After binning the spike trains by summing the number of spikes in each jitter window L for each trial n, we obtained a new matrix X^L with $s = T/L$ bins and N trials. This X^L matrix was normalized by dividing each element by $E(X^L_{s,n}) := \frac{1}{N}\sum_{n=1}^{N} X^L_{s,n}$ to obtain X^{L-norm}, which is a $S \times N$ matrix. Each bin in X^{L-norm} is then replicated L times to transfer the X^{L-norm} from $S \times N$ back to $T \times N$ matrix. We then calculated the PSTH- jitter mean of the original spike train X as:

$$E(X^L_{t,n}|X) := PSTH_t(X)X^{L-norm}_{t,n}$$

This $E(X^L|X)$ represents the probability distribution of the resampled spike train, which is a $T \times N$ matrix. This was directly used to calculate the jittered CCG:

$$CCG_{jittered} = \left(E(X^L|X) * E(Y^L|Y)\right)$$

The correction term $(CCG_{jittered})$ is the true expected value which reflects the average over all possible resamples of the original dataset. $CCG_{jittered}$ is normalized by the geometric mean rate before subtracting from $CCG_{original}$. The analytical formula used to create a probability distribution of resampled spikes was provided in Harrison and Geman (2009). This method disrupts temporal correlations within the jitter window, while maintaining the number of spikes in each jitter window and the shape of the PSTH averaged across trials.

$$\mathrm{CCG}_{\mathrm{jitter\_corrected}} = \mathrm{CCG}_{\mathrm{original}} - \mathrm{CCG}_{\mathrm{jittered}}$$

For our measurement, a 25 ms jitter window was chosen based on previous studies (Jia et al., 2013; Zandvakili and Kohn, 2015). This jitter-correction method removes both the stimulus-locked component of the response, as well as slow fluctuations larger than the jitter window. The remaining fast timescale correlation is more likely to be related to signal propagation between two neurons. Therefore, the jitter-corrected CCG reflects temporal correlations between a pair of neurons within the jitter-window (25ms).

We then calculated the directed connection weight by subtracting the sum of (−13 to 0) ms of the CCG from the sum of (0 to 13) ms of the jitter-corrected CCG (Figure 1D). The 13ms window was defined as half of the 25 ms jitter window we used, and also

Please cite this article in press as: Jia et al., Multi-regional module-based signal transmission in mouse visual cortex, Neuron (2022), https://doi.org/10.1016/j.neuron.2022.01.027

because real functional delay between neurons in mouse occur on the timescale of milliseconds to tens of milliseconds (Siegle et al., 2021). The resulting value indicates the strength and the sign indicates the directionality of the functional connection between a pair of neurons. Computing this for all pairs of neurons produced a directional, cellular-resolution connectivity matrix for each mouse (Figure 1E).

Clustering

Non-randomness

We first tested whether there was modular structure (non-randomness) in the measured connectivity matrix by computing the graph spectrum (based on spectral graph theory (Spielman, 2007). The eigenvalues of a graph are defined as the eigenvalues of its adjacency matrix (Fornito et al., 2016). The set of eigenvalues of a graph forms a graph spectrum. The randomness of the matrix is quantified by comparing the graph spectrum of the original connectivity matrix with its shuffled connectivity matrix, where the x and y axis are shuffled independently, and a randomly generated connectivity matrix with the same size. We found that the graph spectrum of the original matrix showed significantly higher explained variance by the top eigenvalues than both the shuffled matrix and the random matrix, suggesting that the measured connectivity matrix has non-random structure (Figure S1).

Defining the number of clusters

The number of clusters was determined using several complementary methods (Figure S3A):

1. The Elbow method estimates the percentage of variance explained for a given number of k. The number of cluster $\widehat{k}$ is estimated at the point when the curve turns into a plateau. The following measure represents the sum of within-cluster distances (pairwise distances) between all points in a given cluster C_k containing n_k points:

$$D_k = \sum_{x_i \in C_k} \sum_{x_j \in C_k} \|x_i - x_j\|^2 = 2n_k \sum_{x_j \in C_k} \|x_i - \mu_k\|^2$$

Adding the normalized within-cluster sum-of-squares gives a measure of the compactness of our clustering, or the pooled within-cluster sum of squares around the cluster means:

$$W_k = \sum_{k=1}^{K} \frac{1}{2n_k} D_k$$

W_k increases monotonically with number of clusters k. The number of clusters is chosen at the point where the marginal gain drops (or the point slope change most dramatically), the 'elbow'.

2. Gap statistics (Tibshirani et al., 2001) seeks to standardize the comparison of log W_k with a null reference distribution of the data, i.e., a distribution with no obvious clustering. The estimate for the optimal number of clusters K is the value for which log W_k falls the farthest below this reference curve. This information is contained in the following formula for the gap statistic:

$$Gap_n(k) = E_n^*\{\log W_k\} - \log W_k,$$

Where E_n^* denotes the expectation under a sample of size n from the reference distribution. The estimate $\widehat{k}$ will be the value maximizing $Gap_n(k)$ after we take the sampling distribution into account.

3. Clustering density estimates the data distribution density for a given k by calculating a density function $f(k)$ (Pham et al., 2005). The value of $f(k)$ is the ratio of the real distortion to the estimated distortion. When the data are uniformly distributed, the value of $f(k)$ is 1. When there are areas of concentration in the data distribution, the value of $f(k)$ decreases. Therefore, the number of $\widehat{k}$ clusters is determined by finding the minimum value of $f(k)$.

Combining the estimation of $\widehat{k}$ using the above three methods, we determined the optimal number of clusters to be 3.

Method for clustering

In order to find neurons that have correlated connectivity patterns to the rest of the network, we clustered the directed connectivity matrix by treating the weights from each source neuron to all target neurons as features (Figure 1F and Figure S2). To reduce noise, we projected the connectivity features into a lower dimensional space with principal component analysis (PCA), only keeping the top principal components that explained 80% of total variance. We then applied a consensus clustering method (Monti et al., 2003) with k-means to obtain robust clusters that are not biased by random initial conditions. First, we constructed a co-clustering association matrix by running k-means with different initial conditions 100 times (reached stable co-clustering). Each entry in the matrix represents the probability of two units belonging to the same cluster. Then, we clustered the association matrix with hierarchical clustering to determine the cluster labels. The number of clusters was determined using methods described in the previous section.

Comparing different clustering methods

Our consensus clustering was based on *k*-means clustering methods, which measures the compactness of points based on features in the reduced PCA space (see above). We compared this clustering method with two other clustering methods to detect modular structure in the adjacency matrix: the spectral clustering method (sklearn.cluster.SpectralClustering) and bi-clustering method (sklearn.cluster.SpectralBiclustering).

Spectral clustering determines the clusters based on the connectivity of data points: points that are connected or immediately next to each other are placed in the same cluster. In spectral clustering, the data points are treated as nodes of a graph, and the clustering is treated as a graph partitioning problem. The nodes are then mapped to a low-dimensional space that can be easily segregated to form clusters. The spectral clustering is carried out in 3 steps: 1. Compute a similarity graph (*k*-nearest neighbors). 2. Project the data onto a low-dimensional space (compute Graph Laplacian, and eigenvalues and eigenvector for L). 3. Create clusters (based on the eigenvector corresponding to the 2^{nd} eigenvalue to assign values to each node, then split the nodes with *k*-means for the given number of clusters).

Biclustering (or block clustering) is a method to simultaneously cluster the rows and columns of a matrix. For a *m* (sample) by *n* (feature) matrix, the algorithm generates biclusters, which are a subset of rows that exhibit similar connectivity pattern across a subset of columns.

The results of consensus clustering, spectral clustering, and biclustering of the functional connectivity matrix are shown in (Figure S3C). The three methods showed relatively consistent clustering results (Figure S3D) in detecting units that belong to the three clusters dominated by different weight pattern. Therefore, our clustering findings are general and do not depend on the specific clustering method we used.

Cluster quality

We used two methods, which were previously used to evaluate spike sorting cluster quality (Siegle et al., 2021), to quantify neuronal population cluster quality given different number of clusters (Figure S3B). The *d*-prime (*d′*) was calculated using Fisher's linear discriminant analysis to find the line of maximum separation in PC space (Hill et al., 2011). *d′* indicates the unbiased separability of the cluster of interest from all other clusters. The higher the value, the more distinguishable are the clusters. Hit-rate was calculated with nearest-neighbors method (n_neighbors = 3), which is a non-parametric estimate of exemplar contamination in each cluster. For each unit belonging to the cluster of interest, the three nearest units in principal-component space are identified. The "hit rate" is defined as the fraction of these units that belong to the cluster of interest. This metric is based on the "isolation" metric from (Chung et al., 2017). The higher the value, the less contamination in each cluster.

Modulation index

To estimate how spiking activity of each unit is modulated by the temporal frequency of the drifting grating stimulus, we calculated the stimulus modulation index (MI) (Matteucci et al., 2019; Siegle et al., 2021), which is defined as:

$$\left| \frac{PS(f_{pref}) - \langle PS \rangle_f}{\sqrt{\langle PS^2 \rangle_f - \langle PS \rangle_f^2}} \right|$$

PS indicates the power spectral density of the peristimulus time histogram (PSTH). The power spectrum was computed using Welch's method on the 10 ms-binned PSTH. $\langle PS \rangle_f$ denotes the averaged power over all frequencies; $PS(f_{pref})$ is the power at the preferred temporal frequency of the unit (f_{pref}). MI reflects the difference in power of the visually evoked response at a unit's preferred stimulus frequency versus the average power spectrum. MI > 3 corresponds to strong modulation of spiking at the stimulus frequency (indicative of simple-cell–like responses), whereas smaller MI values indicate less modulation by stimulus temporal frequency (indicative of complex-cell–like responses) (Matteucci et al., 2019).

Module distribution

To estimate the distribution of units in each module across areas, we first combined all units across mice to increase the total number of units in each area and minimize recording bias. Then, since we observe a sampling bias across areas, we randomly subsampled units in each area to match the number of units in each area. After that, we estimated the distribution of driver neurons by dividing the number of driver units in each area by the total number of driver units after subsampling. The same calculation is applied to estimate the distribution of driven units. The proportion of one module across areas sums to 1. The final result was a bootstrapped mean (sampling with replacement to match the number of units in each area; n_boot = 100). Error bars represent the bootstrapped standard deviation across samples. Results are only shown for the 'driver' and 'driven' modules. No systematic area bias was observed for cluster 1 (the cluster with non-significant connection) units (result not shown).

The distribution of each neuronal module across layers was quantified by first dividing units into superficial, middle, and deep layers according to the location of layer 4 estimated from the CSD. To estimate the distribution of units in each module across depth, we first combined all units across mice to increase the number of units and minimize sampling bias across mice. Then, since there is depth dependency of number of units, we randomly subsampled units in each depth to match the number of units across superficial, middle, and deep layers. After that, we estimated the distribution of driver units across depth by dividing the number of driver units at

Please cite this article in press as: Jia et al., Multi-regional module-based signal transmission in mouse visual cortex, Neuron (2022), https://doi.org/10.1016/j.neuron.2022.01.027

each depth by the total number of driver units. The same is done for driven units. The proportion of one module across depth sums to 1. Error bars represent bootstrapped standard deviation across samples (n_boot = 100).

Graph creation

To create graph visualizations (Figures 5B, 5D, and 6B), we first condensed our single-unit connectivity matrix to a single-recording-site connectivity matrix by combining units with peak channels on the same electrode. Then, we treat each site as a node in the graph. For an intuitive representation, nodes belonging to the same cortical area are close by and arranged clockwise from superficial to deep layer. The location of each area is determined by the top-down view of the physical locations of visual areas on the left hemisphere (Figure 1A). The edges of the graph represent connections between sites, with red lines indicating projections from the source unit (positive weight) and blue lines indicating projections back to the source unit (negative weight).

Divergence and convergence degree

Divergence degree is similar in concept to the outdegree of a graph. It is defined as the proportion of significant positive connections (weight $> 10^{-6}$) from a source neuron to the rest of the network (N neurons). $C_{i,+}$ represents the number of positive connections from neuron i to the network.

$$divergence = \frac{C_{i,+}}{N-1}$$

Convergence degree is similar in concept to the indegree of a graph. It is defined as the proportion of significant negative connections (weight $< -10^{-6}$) to source neuron i from the rest of the network.

$$convergence = \frac{C_{i,-}}{N-1}$$

Temporal dynamics analysis

Response latency

Two different measurements were used to estimate response latency. First, the peak response latency was defined as the time when a neuron's response reached its first peak after stimulus onset. Second, the time to first spike was estimated in each trial by looking for the time of the first spike 30 ms after stimulus onset. If no spike was detected within 250 ms after stimulus onset, that trial was not included. The overall latency for each unit was defined as the mean time to first spike across trials.

Population onset temporal dispersion

We used the spread of time-to-first spike (temporal dispersion) for all neurons within a module for a single trial in each mouse as an indicator of population response onset synchronization. After obtaining the distribution of time to first spike of all neurons in the population for a given trial in each mouse, we calculated the median of the distribution ($t_{\_median}$). The spread was calculated by fitting a Gaussian to this time-to-first-spike distribution within a window of ± 40 ms of $t_{\_median}$:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2},$$

where x is the spike time relative to stimulus onset, μ is an estimate of the average time-to-first-spike and σ is an estimate of the spread of the time-to-first spike distribution for one trial.

Population response spread of the first peak

To quantify spike the response spread of a neuronal ensemble, we estimated the width of the population PSTH for each trial. The PSTH was calculated with 2 ms bins, and convolved with a Gaussian kernel of width 5 ms. The properties of the first peak were estimated using scipy.signal.find_peaks. The peak width represents the half-width at half maximum of the peak, while peak height represents the maximum of the peak. The spike spread of a neuronal population is an important parameter for quantifying how signals are transmitted through a feedforward network.

Area in-out index

To quantify the proportion of projections out from a source area relative to inputs back into the source area, we defined the area in-out index as

$$index = \frac{C_{in} - C_{out}}{C_{in} + C_{out}}$$

where C_{out} is the number of connections from source to other areas in the given network and C_{in} is the number of connections from other areas to source area. This index reflects the asymmetry of in-and-out degree of a source area. When the value is close to -1, the source area is dominated by outward projections (positive weights). When the value is close to 1, the source area is

dominated by inward projections (negative weights). When the value is 0, the source area has balanced outward and inward connections.

Layer definition

We estimated the depth of the middle layer of cortex by first calculating the current source density (CSD) using simultaneously recorded local field potentials (LFP). The CSD was computed using the method in (Stoelzel et al., 2009), using the LFP within 250 ms after stimulus onset. First, we calculated the average evoked (stimulus locked) local field potential at each recording site. Next, we duplicated the uppermost and lowermost field traces and smoothed these signals across sites

$$\overline{\varphi}(r) = \frac{1}{4}(\varphi(r + h) + 2\varphi(r) + \varphi(r - h))$$

where φ is the field potentials, r is the coordinate perpendicular to the layers, h is the spatial sampling interval (40 mm in our case). Then, we calculated the second spatial derivative

$$D = \frac{1}{h^2}(\overline{\varphi}(r + h) - 2\overline{\varphi}(r) + \overline{\varphi}(r - h)).$$

In the resulting CSD map, current sinks are indicated by downward deflections and sources by upward deflections. To facilitate visualization, we smoothed the CSD with 2D Gaussian kernels ($\sigma_x = 1$; $\sigma_x = 2$). To find the middle layer, we defined the first sink within 100 ms after stimulus onset as the input layer (center channel) by searching for the local maximum on the CSD map (first sink), followed by source.

We used the middle layer estimation for the calculation of layer distribution bias of 'driver' and 'driven' modules. We partitioned the cortical layers into three layers: middle layer (center channel ± 8 channels, which is ± 40ĕm), superficial layer (channels above middle layer), and deep layers (channels below middle layer and above white matter).

Simulations to test mathematical relationship between PSTH shape and CCG sharp peak

Because we observed that the functional connectivity defined 'driver' module responded earlier than the 'driven' module (Figure 5E), we wondered whether the brief timescale relationship of 'driver' leading 'driven' was a consequence of the general latency reflected in averaged PSTH. Even though our jitter-correction method should have removed stimulus-locked components and the observed directionality should only reflect brief-timescale signal transmission, we still wanted to rule out the possibility that the observed asymmetry in the CCG is merely a reflection of the trial-averaged PSTH latency.

We used a simple simulation to carry out positive and negative controls (Figure S5). The negative control tested whether two neurons with correlated, but temporally offset, PSTH traces will necessarily show significant peaks in their jitter-corrected CCG (25 ms jitter window). The positive control tested whether two neurons with uncorrelated PSTH traces can produce a significant peak in their CCG if we artificially introduce millisecond-timescale correlations. The mathematical expression of the tests is formulated as follows:

Given two PSTH traces: $\lambda_1(t)$ and $\lambda_2(t)$, we simulated *Poisson* spike trains:

$$x_1(t) \sim Pois(\lambda_1(t))$$

$$x_2(t) \sim Pois(\lambda_2(t)),$$

over time T for 100 repeats, where the PSTHs of the two simulated spike trains (X_1 and X_2) matched the shape of $\lambda_1(t)$ and $\lambda_2(t)$. Synchronized spikes were introduced to $x_1(t)$ and $x_2(t)$ only for the positive control. CCGs before and after jitter correction were calculated between X_1 and X_2. We found that brief timescale correlations between two neurons (identified by significant peaks in the CCG) do not depend on the shape and relative timing of their PSTHs, but—as expected—reflect only their fine-timescale temporal relationship.