Brain-state mediated modulation of inter-laminar dependencies in visual cortex

Anirban Das[1,2,5,#], Alec G. Sheffield[3,#], Anirvan S. Nandy[2,4,†] and Monika P. Jadi[1,2,4,†,*]

[1]Department of Psychiatry, Yale University, New Haven, CT 06511

[2]Department of Neuroscience, Yale University, New Haven, CT 06511

[3]Interdepartmental Neuroscience Program, Yale University, New Haven, CT 06511

[4]Wu Tsai Institute, Yale University, New Haven, CT 06511

[5]Design and Patterning AI Group, Intel Corp., Hillsboro, Oregon 97124

[#]equal contribution

[†]senior author

[*]corresponding author (monika.jadi@yale.edu)

ABSTRACT

Adaptive information processing, comprised of local computations and their efficient routing, is crucial for flexible brain function. Spatial attention is a quintessential example of this adaptive process. It is critical for recognizing and interacting with behaviorally relevant objects in a cluttered environment[1]. Object recognition is mediated by ensembles of computational units distributed across the ventral visual hierarchy. How the deployment of spatial attention aids these hierarchical computations is unclear. Based on pairwise correlation analysis, two key mechanisms have been proposed: First is an improvement in the efficacy of unique information directed from one encoding stage to another, suggested by evidence along the visual hierarchy[2-6]. Based on the theoretical results that even weak correlated variability can substantially limit the encoding capacity of a neuronal pool[7], a second proposal is an improvement in the sensory information capacity of an encoding stage through a reduction in shared fluctuations[8,9]. However, pairwise analyses capture both unique and shared components of fluctuations, and therefore cannot

disambiguate the proposed mechanisms. To test these proposals, it is crucial to estimate the attentional modulation of unique information flow across and shared information within the stages of the visual hierarchy. We investigated this in the multi-stage laminar network of visual area V4, an area strongly modulated by attention[10-12]. Using network-based statistical modeling, we estimated the strength of inter-layer information flow by measuring statistical dependencies that reflect how the cortical layers *uniquely* drive each other's neural activity. We quantified their modulation across attention conditions (attend-in vs. attend-away) in a change detection task and found that deployment of attention indeed strengthened unique dependencies between the input and superficial layers. Using the partial information decomposition framework[13,14], we estimated modulation of shared dependencies and found that they are reduced within laminar populations, specifically the putative excitatory subpopulations. Surprisingly, we found a strengthening of unique dependencies within the laminar populations, a finding not previously predicted. Crucially, these modulation patterns were also observed across behavioral outcomes (hit vs. miss) that are thought to be mediated by endogenous state fluctuations[15-17]. By "decomposing" the modulation of dependency components and in combination with prior theoretical work[7], our results suggest the following computational model of optimal sensory states that are attained by either task demands or endogenous fluctuations in brain state: enhanced information flow between and improved information capacity within encoding stages.

43 **MAIN TEXT**

44 Visual cortex has a laminar organization and both sensory computations and information flow
45 patterns are layer specific, forming the building blocks of the ventral visual hierarchy[18-22]. We
46 hypothesized that the deployment of spatial attention strengthens unique inter-layer information
47 transfer between and weakens shared information within the input and superficial layers, both of
48 which are crucial nodes of feedforward information flow along the ventral visual hierarchy. To test
49 this, our primary goal was to quantify unique statistical spiking dependencies between the
50 populations of each layer, which requires characterizing the joint spiking activity of the laminar
51 ensemble. Cortical ensembles are highly interconnected, and this can be a source of both shared
52 and unique dependencies between any pair of ensemble members (Fig 1a, S1a-b). Pairwise
53 measures (such as correlations or univariate regression coefficients) reflect both unique and
54 shared sources of dependences (Fig 1b,c), whereas network-based measures (such as
55 regularized regression coefficients) primarily extract non-shared sources of dependencies (Fig
56 1d). Using the Partial Information Decomposition framework[13,14] (Fig 1e, see SI Methods), we
57 proceeded to use a combination of network-based and pairwise regression models to analyze the
58 neurophysiological data and "decompose" the modulation of these dependencies by spatial
59 attention. Probabilistic graphical models (PGM)[23] provide a powerful unbiased framework for
60 capturing unique dependency structure among variables in a large network. Dynamic Bayesian
61 Networks (DBN)[24-29] is a class of PGMs that describes conditional dependencies between
62 variables in a sparse graph structure. Fitting our data with DBNs allowed unbiased discovery of
63 multi-timescale directed and cyclical dependencies that are Granger causal[30], without making
64 assumptions about their specific nature (linear/nonlinear), direction, or latency (Fig 2a, Fig S2).
65 We estimated the probability of discovered dependencies (referred henceforth as weights of a
66 graph edge) and used it as a measure of their strength[31] (Fig 2a,b). We used a time-shuffled
67 estimate of the edge weights to determine the statistical significance of the discovered

68 dependencies (Fig 2a). We refer to this approach as multi-timescale weighted Dynamic Bayesian
69 Networks (MTwDBN). When tested on a ground truth synthetic network with recurrent
70 connections, the MTwDBN approach robustly recovered the dependency structure (Fig 2c-e), and
71 performed significantly better than regularized regression, especially when the populations were
72 sparsely sampled (Fig 2f) as is the case in neural recordings[32-34]. Additionally, MTwDBN
73 outperformed existing unweighted[29] or fixed-threshold weighted DBN methods[31] (Fig 2f). Finally,
74 in addition to dependency strength, edge weights in MTwDBN provided a more accurate estimate
75 of the dependency structure (Fig 2g).

76

77 We applied the MTwDBN analysis to laminar recordings in visual area V4, an area strongly
78 modulated by attention[8,11,35-37], of non-human primates performing an attention-demanding
79 orientation change detection task (Fig 3a,b; see Methods and Fig S3). We first analyzed the
80 ensemble of layer-wise populations (Fig 3c) and quantified the net attentional modulation of
81 dependencies across different timescales (Fig 3d). At longer timescales (> 60 ms lag)[33], while
82 attention weakened pairwise dependencies, in agreement with previous findings[38,39], we found a
83 strengthening of unique dependencies. This allowed us to infer (Fig 1e) that it is the shared
84 component of dependencies that is weakened by attention (Fig 3d, bottom), thereby providing
85 direct evidence for a previously hypothesized mechanism of perceptual improvement by
86 attention[38,39]. Unique dependencies between layers, specifically those between the input and
87 superficial layers, an important link in feedforward processing, were strengthened by attention
88 (Fig 3e). On the other hand, shared dependencies within the input and superficial layers were
89 weakened by attention (Fig 3f). Surprisingly, unique dependencies within the layers were
90 strengthened by attention at most timescales (Fig 3g). Taken together, these results provide the
91 first direct evidence that attention improves unique dependencies both within and across stages

of the laminar circuits of the ventral visual hierarchy, while weakening shared dependencies within these stages (Fig 3h).

Neuronal cell-classes in the cortex contribute differentially to information processing[35,36,40]. To test if the above results hold when we allow discovery of cell-class specific dependencies, we next analyzed an ensemble of broad- and narrow-spiking layer-wise populations (Fig 3i) and quantified the net attentional modulation of dependencies across different timescales (Fig 3j). The pattern of net modulation of unique and shared dependencies in this ensemble largely mirrored that discovered in the layer-wise aggregated ensemble, specifically at the longer timescales. Same was the case for modulation of unique dependencies between layers and shared dependencies within layers (consistent across both animals) (Fig 3k,l). Theoretical predictions[7], including our hypotheses, regarding effects of shared fluctuations on information representation apply primarily to excitatory neurons. When we quantified the attention modulation of shared dependencies within layers in a cell-class specific manner, we found that the broad-spiking population (putative excitatory neurons) showed a robust weakening of shared dependencies within layers (consistent across animals) (Fig 3l). On the other hand, the narrow-spiking population (putative inhibitory neurons) showed a distinct pattern, one that was dominated by strengthening of shared dependencies within layers. Consistent with the findings in the aggregated ensemble, unique dependencies within the layers were strengthened by attention at most timescales (Fig 3m). Taken together, these results provide the first direct evidence that attention specifically weakens shared dependencies in the projection populations within encoding stages, in addition to robustly improving unique dependencies both within and across stages of the ventral visual hierarchy (Fig 3n).

To test if the pattern of inter- and intra-layer dependency modulation that we observed is a signature of brain states that are optimal for perceptual behavior, we analyzed the laminar

ensemble activity for a subset of trials within the attend-in condition in which the animal was equally likely to correctly detect (Hit) or fail to detect (Miss) the orientation change (Fig 4a,b). Controlling for task and stimulus conditions, these behavioral fluctuations are thought to arise from endogenous fluctuations such as attentional sampling and arousal changes[15-17,41,42]. The pattern of net modulation of unique and shared dependencies across behavioral outcome (Fig 4c) largely mirrored the pattern discovered across attentive states (Fig 3d) at longer timescales. Same was the case for modulation of unique dependencies between layers and dependencies within layers (Fig 4d), suggesting that the pattern of enhanced inter-layer unique dependencies and weakened intra-layer shared dependencies is a hallmark of brain states that are optimal for perceptual performance (Fig 4e). Combined with prior theoretical work[7], our results suggest a model of enhanced inter-layer communication and improved intra-layer information capacity during optimal states that are either imposed by task demands or are attained via endogenous fluctuations in brain state.

Our approach demonstrates that the information flow structure in a neural ensemble is robustly described by Dynamic Bayesian Network modeling that is adapted to include multiple lags spanning timescales of interest. This unbiased approach allows us to dive deeper into the causal functional interactions (beyond correlations) in a multivariate system with unknown nonlinear dependencies. Further, the results demonstrate that weighing dependencies using confidence measures provides a more accurate information flow structure that can be utilized to investigate how this structure is modulated by changes in brain and behavioral states. Finally, an integration of this approach with pairwise models allows us to dissect changes in distinct components of dependencies. Since temporally directed unique dependencies suggest causality and shared ones suggest common inputs, this allows us to gain insights into the underlying mechanisms. In contrast, prior correlational studies have quantified net dependency modulation [2,9,38]. While these

provided valuable characterization of the neural correlates of attention, the underlying network mechanisms remained hypothetical.

While approaches such as GLM have been useful in providing improved phenomenological models of individual neural responses to sensory stimulation in early sensory circuits[43], they are not ideal for dependency structure learning in a highly recurrent cortical ensemble. Their inherent assumption of the independence of predictor variables can discover spurious dependencies due to shared inputs, and result in a dense and likely inaccurate structure. Since structure learning in DBNs involves determining conditional independence by solving an optimization problem that penalizes density, our approach is ideal for generating a sparse and hence interpretable unique dependency structure in multivariate data.

This class of DBN models does not identify functional consequences of dependencies between populations, such as enhancement or suppression of target activity. Future work in this direction will further enhance the functional interpretability of discovered dependencies. Nonetheless, these models are effective in elucidating structure and modulation of information flow in a multivariate system such as the laminar cortical network.

In general, discovery of dependency structure using our approach is sensitive to the inclusion of relevant neural variables. It is indeed possible that a gross classification of subpopulations as broad and narrow in our study fuses distinct but relevant neural variables. On the other hand, our three and six population laminar DBNs yield consistent patterns of dependency modulations, suggesting an inclusion of relevant variables for the purposes of this study. It is important to note that while statistical dependencies discovered using our approach imply directed functional

connectivity, these do not necessarily imply direct anatomical connectivity: an edge in the structure could reflect indirect anatomical connectivity.

MTwDBN based models provide a quantitative description of information flow patterns in an ensemble. In this study, they provide the first structural description of the attention modulation of dependencies in cortical space and time in a compartmentalized network, such as the laminar cortical network. They allow us to quantify the unique contributions of activity history and network interactions to information processing in such a neural ensemble. We expect this framework to extend to neuronal ensembles in other parts of the nervous system, and to play an important role in revealing flexible information processing principles in the brain.

METHODS SUMMARY

Synthetic Data Dataset for Figure 1 was constructed by simulating a probabilistic spiking network with causal and shared connectivity among 8 neural variables**.** Dataset for evaluating the MTwDBN pipeline (Fig 2) was generated by simulating a network of stochastically spiking excitatory and inhibitory neurons.

Electrophysiological Data 16-channel laminar probes were used to record data from visual area V4 of 2 macaque monkeys as they performed an attention demanding task over 30 sessions total [35]. The probe recorded single units with overlapping receptive fields across the depth of the cortex. Single units were classified using peak-to-trough duration.

Behavior Monkeys were trained to perform an attention-demanding change detection task [35].

Fitting DBN graph structure was fitted to 200 ms of visual stimulus flashes. Data was prepared using 15 ms bins to fit a DBN with either 21 (3 laminar populations, 7 lags) or 42 nodes (3 lamina x 2 waveform types, 7 lags). Bootstrap methods were used to generate edge weights or confidence measures of edges discovered from the data.

192 **FIGURE CAPTIONS**

193 **Figure 1 Pairwise- and network-model parameters as a function of unique vs. shared**
194 **dependencies**. **a** A synthetic ensemble of eight neural variables with two kinds of dependencies
195 – unique or shared – between seven source variables (black) and one target variable (cyan). All
196 interactions are excitatory. Strength of dependencies is determined by model parameters P_{unique}
197 and P_{shared} (see SI Methods). **b** Normalized total mutual information, measured by uncertainty
198 coefficient, as a function of the sum of model parameters (P_{unique}, P_{shared}) that varied unique and
199 shared components of mutual information in a monotonic way (Fig S1c-d). **c** Coefficients of a
200 pairwise model (univariate logistic regression (UR)) as a function of P_{unique} and P_{shared}. White
201 arrow provides a visual guide for direction of highest change in coefficients. **d** Coefficients of a
202 network model (LASSO multivariate regularized regression (RR)) as a function of P_{unique} and
203 P_{shared}. White arrow provides a visual guide for direction of highest change in coefficients. **e**
204 Schema for utilizing pairwise and network methods for the estimation of total (black) and unique
205 (green) information modulation respectively, and to infer the modulation of shared information
206 (purple). Shaded blocks indicate indeterminate modulation direction of shared information in the
207 network.

208 **Figure 2 Multi-timescale & weighted Dynamic Bayesian Networks (DBN) based estimation**
209 **of unique dependencies in a sparsely sampled recurrent neuronal network. a** Analysis flow
210 for multi-timescale weighted DBN (MTwDBN) model fitting. **b** Edge weight of MTwDBN as a
211 function of connection weight in a 2-population simulated network using the pipeline in **a**. **c** Spiking
212 activity of 6 subpopulations in a simulated laminar network. Connectivity is visualized in the
213 overlaid schematic. **d** Directed dependencies (edge in the graph) in the simulated network in **c**,
214 estimated using MTwDBN fitting. **e** Summary graph of dependencies across all timescales from
215 **d**. **f** F-score (harmonic mean of precision and recall of dependency structure) as a function of %

216 of neural population observed. F-score was estimated for shuffle corrected weighted DAGs
217 (MTwDBN, green), weighted DAGs with fixed threshold (weightedFT, blue), unweighted DAGs
218 (red), or LASSO regression, an example of regularized regression (RR) models (black). Each
219 point represents the average of five separate runs, except 100 % (single run). Error bars indicate
220 standard deviation, some error bars are smaller than symbol size. **g** DBN decoder accuracy with
221 different sizes of MTwDBN DAGs. Decoders were trained to predict population activities using a
222 subsample of shuffle corrected edges (see SI Methods). Graph edges for the decoder were
223 sampled from the learned structure either in an unbiased fashion (black) or biased with the edge
224 weights (green).

225

226 **Figure 3 Modulation of dependencies in a V4 laminar network across attention conditions.**
227 **a** Experimental protocol: Paired Gabor stimuli with different contrasts (see Methods); one stimulus
228 was presented inside the receptive fields (RFs) of the recorded neurons and the other at an
229 equally eccentric location across the vertical meridian. Attention was cued either to the neurons'
230 RFs ("IN") or to the location in the contralateral visual hemifield ("AWAY"). Orientation of one of
231 the two stimuli changed at a random time. Monkey was rewarded for detecting the change by
232 making a saccade to the corresponding location. Task difficulty was controlled by the magnitude
233 of orientation change. **b** Recording approach: laminar recordings in visual area V4. **c** Neural
234 populations used for MTwDBN analysis. Current source density analysis identified different layers
235 (*superficial*, *input*, *deep*), and isolated single units were assigned to one of these layers (see SI
236 Methods). **d** *Top*: Average MTwDBN-based modulation (green) of all unique dependencies
237 between the laminar populations. Modulation of the same dependencies as estimated by logistic
238 regression (black). *Bottom*: PID framework-based estimated modulation sign of shared
239 dependencies (see schema in Fig 1e). Thicker line along the time axis indicates the timescales

240 of attentional modulation in prior studies[38,39]. Error bars indicate 95% confidence intervals. **e** Sign
241 of average modulation of unique dependencies between layers (bi-directionally). **f** Sign of average
242 modulation of shared dependencies within layers. **g** Sign of average modulation of unique
243 dependencies within layers. **h** Summary of dependency modulation pattern. **i** Neural populations
244 used for laminar MTwDBN analysis. Isolated single units were classified as broad and narrow
245 based on peak-to-trough duration in their average spike shape (see SI Methods). **j-n** Same as **d-**
246 **h**, but for populations shown in **i**. I: input layer; S: superficial layer; *M1*, *M2*: subject-wise
247 dependencies; *broad*, *narrow*: cell-class specific dependencies. See Fig S4 for modulation indices
248 in **e-g**, **k-n**.
249
250 **Figure 4 Modulation of dependencies in a V4 laminar network across behavioral outcomes**
251 **at perceptual threshold. a** Example session showing performance as a function of task difficulty.
252 (orange box: threshold orientation change at which the animal was equally likely to correctly
253 detect (Hit) or fail to detect (Miss) the change). **b** Laminar populations used for multi-lag analysis.
254 **c** Modulation magnitude (top) and sign (bottom) of all unique (green) and total (black) laminar
255 dependencies in **b** across Hits and Misses at perceptual threshold. Estimated modulation sign of
256 shared dependencies (bottom, see Fig 1e). **d** Modulation sign of between (BL) and within (WL)
257 layer dependencies. See Fig S5 for modulation indices. **e** Summary of laminar dependency
258 modulation pattern.
259
260 **Figure S1 Information decomposition in a multi-variate system**
261 **a** Partial information decomposition (PID) framework[13,14] based definition of types of information
262 that multiple sources can have about a target.

b Application of pairwise and network-based statistical models for approximate information decomposition in a multivariate system. Panel on the right illustrates interpretation of the modulation of these dependencies using the PID framework.

c, d Reduction in the proportion of total entropy (or information fraction) as a function of parameters (P_{unique}, P_{shared}) that control unique and shared information in the model. For the sake of computational efficiency, information fraction estimation as a function of P_{shared} (d) was performed using a subset of variables in the simulated network.

Figure S2 Additional MTwDBN performance metric for synthetic data

Recall, precision and F-score measures for weighted DAGs in Fig 2d as a function of number of lags.

Figure S3 Laminar recordings in area V4

a Stacked contour plot showing spatial receptive fields (RFs) along the laminar probe from an example session. The RFs are well aligned, indicating perpendicular penetration down a cortical column. Zero depth represents the center of the input layer as estimated with current source density (CSD) analysis. **b** CSD displayed as a colored map. The x-axis represents time from stimulus onset; the y-axis represents cortical depth. The CSD map has been spatially smoothed for visualization. **c** An example trial showing single-unit activity across the cortical depth in the attend-in condition. The time axis is referenced to the appearance of the fixation spot. Spikes (vertical ticks) in each recording channel come from the single unit with the highest spike rate in this trial. The gray boxes depict stimulus presentation epochs. In this particular trial, 8 sample stimuli with different contrasts were presented, followed by a target stimulus flash with an

orientation change that the monkey responded to correctly. Spike waveforms for two example narrow (cyan) and broad (orange) single units are shown.

Figure S4 Modulation of dependencies in a V4 laminar network across attention conditions. MTwDBN-based modulation (green) of all unique dependencies between the laminar populations. Modulation of the same dependencies as estimated by logistic regression (black). Open circles indicate no significant edges for a given time lag. Schematic at the top of every column depicts the populations used for identifying laminar dependency structure within (left) and between (right) cortical layers. *broad, narrow*: cell-class specific populations identified by spike shape (see SI Methods).

Figure S5 Modulation of dependencies in a V4 laminar network across behavioral outcomes at perceptual threshold. MTwDBN-based modulation (green) of all unique dependencies between the laminar populations. Modulation of the same dependencies as estimated by logistic regression (black). Schematic at the top of every column depicts the populations used for identifying laminar dependency structure within (left) and between (right) cortical layers.

AUTHOR CONTRIBUTIONS

AD and MPJ conceptualized the project. ASN collected the electrophysiological data. AD and MPJ generated synthetic data. AGS and AD designed the data analysis pipeline. MPJ supervised the project. AGS, AD, ASN and MPJ wrote the manuscript.

ACKNOWLEDGEMENTS

This research was supported by NIH R00 EY025026, NIH R21 MH126072 and SFARI 875855 to MPJ, NARSAD Young Investigator Grant, Ziegler Foundation Grant, Yale Orthwein Scholar Funds, NIH R21 MH126072 and SFARI 875855 to ASN, NSF GRFP fellowship to AGS and by NEI core grant for vision research P30 EY026878 to Yale University. We thank Steve Chang and Weikang Shi for helpful comments on the manuscript.

REFERENCES

1 Petersen, S. E. & Posner, M. I. The attention system of the human brain: 20 years after. *Annual review of neuroscience* **35**, 73-89, doi:10.1146/annurev-neuro-062111-150525 (2012).

2 Ruff, D. A. & Cohen, M. R. Attention Increases Spike Count Correlations between Visual Cortical Areas. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **36**, 7523-7534, doi:10.1523/JNEUROSCI.0610-16.2016 (2016).

3 Briggs, F., Mangun, G. R. & Usrey, W. M. Attention enhances synaptic efficacy and the signal-to-noise ratio in neural circuits. *Nature* **499**, 476-480, doi:10.1038/nature12276 (2013).

4 Womelsdorf, T. & Fries, P. The role of neuronal synchronization in selective attention. *Current opinion in neurobiology* **17**, 154-160, doi:10.1016/j.conb.2007.02.002 (2007).

5 Bosman, C. A. *et al.* Attentional stimulus selection through selective synchronization between monkey visual areas. *Neuron* **75**, 875-888, doi:10.1016/j.neuron.2012.06.037 (2012).

6 Gregoriou, G. G., Gotts, S. J., Zhou, H. & Desimone, R. High-frequency, long-range coupling between prefrontal and visual cortex during attention. *Science* **324**, 1207-1210, doi:10.1126/science.1171402 (2009).

7 Zohary, E., Shadlen, M. N. & Newsome, W. T. Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature* **370**, 140-143, doi:10.1038/370140a0 (1994).

8 Mitchell, J. F., Sundberg, K. A. & Reynolds, J. H. Spatial attention decorrelates intrinsic activity fluctuations in macaque area V4. *Neuron* **63**, 879-888, doi:10.1016/j.neuron.2009.09.013 (2009).

9 Cohen, M. R. & Maunsell, J. H. R. Attention improves performance primarily by reducing interneuronal correlations. *Nat Neurosci* **12**, 1594-1600, doi:10.1038/nn.2439 (2009).

10 Desimone, R. & Duncan, J. Neural mechanisms of selective visual attention. *Annual review of neuroscience* **18**, 193-222 (1995).

11 McAdams, C. & Maunsell, J. Effects of Attention on Orientation-Tuning Functions of Single Neurons in Macaque Cortical Area V4. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **19**, 431-441 (1999).
12 Reynolds, J., Pasternak, T. & Desimone, R. Attention increases sensitivity of V4 neurons. *Neuron* **26**, 703-714 (2000).
13 Wibral, M., Priesemann, V., Kay, J. W., Lizier, J. T. & Phillips, W. A. Partial information decomposition as a unified approach to the specification of neural goal functions. *Brain Cogn* **112**, 25-38, doi:10.1016/j.bandc.2015.09.004 (2017).
14 Williams, P. L. & Beer, R. D. Nonnegative decomposition of multivariate information. *arXiv preprint arXiv:1004.2515* (2010).
15 Fiebelkorn, I. C., Pinsk, M. A. & Kastner, S. in *Neuron* Vol. 99 842-853.e848 (2018).
16 Helfrich, R. F. *et al.* in *Neuron* Vol. 99 854-865.e855 (2018).
17 Fiebelkorn, I. C., Saalmann, Y. B. & Kastner, S. in *Current Biology* Vol. 23 2553-2558 (Elsevier, 2013).
18 Douglas, R. J. & Martin, K. A. C. Neuronal circuits of the neocortex. *Annual review of neuroscience* **27**, 419-451, doi:10.1146/annurev.neuro.27.070203.144152 (2004).
19 Douglas, R. J. & Martin, K. A. C. Recurrent neuronal circuits in the neocortex. *Current biology : CB* **17**, R496-500, doi:10.1016/j.cub.2007.04.024 (2007).
20 D'Souza, R. D. & Burkhalter, A. in *Front Neuroanat* Vol. 11 71 (2017).
21 Gattass, R., Galkin, T. W., Desimone, R. & Ungerleider, L. G. in *The Journal of comparative neurology* Vol. 522 1941-1965 (John Wiley & Sons, Ltd, 2014).
22 Ungerleider, L. G., Galkin, T. W., Desimone, R. & Gattass, R. Cortical connections of area V4 in the macaque. *Cerebral cortex (New York, NY : 1991)* **18**, 477-499, doi:10.1093/cercor/bhm061 (2008).
23 Koller, D. & Friedman, N. *Probabilistic graphical models: principles and techniques*. (MIT press, 2009).
24 Murphy, K. P. & Russell, S. Dynamic bayesian networks: representation, inference and learning. (2002).
25 Smith, V. A., Jarvis, E. D. & Hartemink, A. J. Evaluating functional network inference using simulations of complex biological systems. *Bioinformatics* **18**, S216-S224 (2002).
26 Zhang, L., Samaras, D., Alia-Klein, N., Volkow, N. & Goldstein, R. in *Advances in neural information processing systems.* 1593-1600.
27 Rajapakse, J. C. & Zhou, J. Learning effective brain connectivity with dynamic Bayesian networks. *Neuroimage* **37**, 749-760 (2007).
28 Smith, V. A., Yu, J., Smulders, T. V., Hartemink, A. J. & Jarvis, E. D. Computational inference of neural information flow networks. *PLoS computational biology* **2** (2006).
29 Eldawlatly, S., Zhou, Y., Jin, R. & Oweiss, K. G. On the use of dynamic Bayesian networks in reconstructing functional neuronal networks from spike train ensembles. *Neural Comput* **22**, 158-189, doi:10.1162/neco.2009.11-08-900 (2010).
30 Seth, A. K., Barrett, A. B. & Barnett, L. Granger causality analysis in neuroscience and neuroimaging. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **35**, 3293-3297, doi:10.1523/JNEUROSCI.4399-14.2015 (2015).
31 Friedman, N., Goldszmidt, M. & Wyner, A. Data analysis with Bayesian networks: A bootstrap approach. *arXiv preprint arXiv:1301.6695* (2013).

32 Levina, A. & Priesemann, V. Subsampling scaling. *Nat Commun* **8**, 15140, doi:10.1038/ncomms15140 (2017).
33 Dalgleish, H. W. *et al.* How many neurons are sufficient for perception of cortical activity? *Elife* **9**, doi:10.7554/eLife.58889 (2020).
34 Priesemann, V., Munk, M. H. & Wibral, M. Subsampling effects in neuronal avalanche distributions recorded in vivo. *BMC Neurosci* **10**, 40, doi:10.1186/1471-2202-10-40 (2009).
35 Nandy, A. S., Nassi, J. J. & Reynolds, J. H. Laminar Organization of Attentional Modulation in Macaque Visual Area V4. *Neuron* **93**, 235-246, doi:10.1016/j.neuron.2016.11.029 (2017).
36 Mitchell, J. F., Sundberg, K. A. & Reynolds, J. H. Differential attention-dependent response modulation across cell classes in macaque visual area V4. *Neuron* **55**, 131-141, doi:10.1016/j.neuron.2007.06.018 (2007).
37 Cohen, M. R. & Maunsell, J. H. R. Using neuronal populations to study the mechanisms underlying spatial and feature attention. *Neuron* **70**, 1192-1204, doi:10.1016/j.neuron.2011.04.029 (2011).
38 Mitchell, J. F., Sundberg, K. A. & Reynolds, J. H. in *Neuron* Vol. 63 879-888 (2009).
39 Cohen, M. R. & Maunsell, J. H. R. in *Nat. Neurosci.* Vol. 12 1594-1600 (2009).
40 Pettine, W. W., Steinmetz, N. A. & Moore, T. Laminar segregation of sensory coding and behavioral readout in macaque V4. *Proceedings of the National Academy of Sciences of the United States of America* **116**, 14749-14754, doi:10.1073/pnas.1819398116 (2019).
41 McGinley, M. J., David, S. V. & McCormick, D. A. in *Neuron* Vol. 87 179-192 (2015).
42 Landau, A. N. & Fries, P. Attention samples stimuli rhythmically. *Curr Biol* **22**, 1000-1004, doi:10.1016/j.cub.2012.03.054 (2012).
43 Pillow, J. W. *et al.* Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature* **454**, 995-999, doi:10.1038/nature07140 (2008).

Figure 1

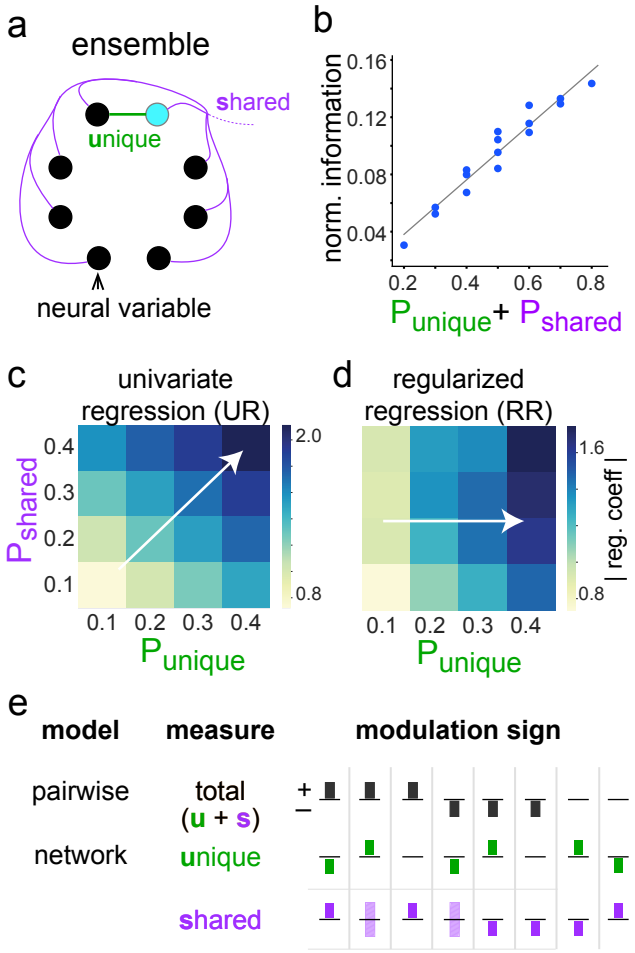

Figure 1 Pairwise- and network-model parameters as a function of unique vs. shared dependencies. a A synthetic ensemble of eight neural variables with two kinds of dependencies – unique or shared – between seven source variables (black) and one target variable (cyan). All interactions are excitatory. Strength of dependencies is determined by model parameters P_{unique} and P_{shared} (see SI Methods). **b** Normalized total mutual information, measured by uncertainty coefficient, as a function of the sum of model parameters (P_{unique} ,P_{shared}) that varied unique and shared components of mutual information in a monotonic way (Fig S1c-d). **c** Coefficients of a pairwise model (univariate logistic regression (UR)) as a function of P_{unique} and P_{shared}. White arrow provides a visual guide for direction of highest change in coefficients. **d** Coefficients of a network model (LASSO multivariate regularized regression (RR)) as a function of P_{unique} and P_{shared}. White arrow provides a visual guide for direction of highest change in coefficients. **e** Schema for utilizing pairwise and network methods for the estimation of total (black) and unique (green) information modulation respectively, and to infer the modulation of shared information (purple). Shaded blocks indicate indeterminate modulation direction of shared information in the network.

Figure 2

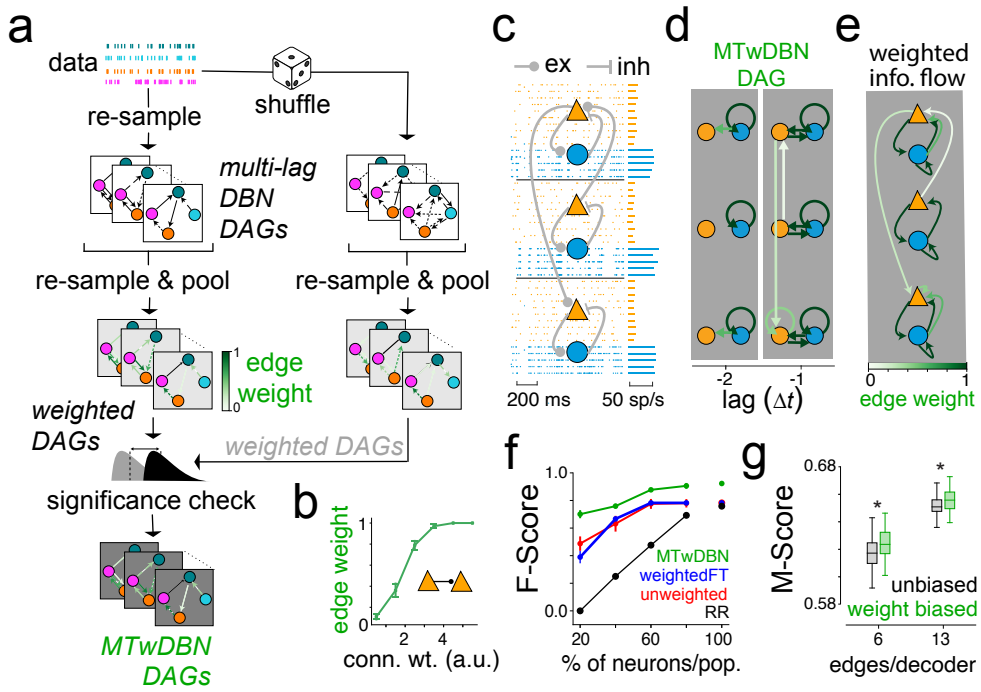

Figure 2 Multi-timescale & weighted Dynamic Bayesian Networks (DBN) based estimation of unique dependencies in a sparsely sampled recurrent neuronal network. **a** Analysis flow for multi-timescale weighted DBN (MTwDBN) model fitting. **b** Edge weight of MTwDBN as a function of connection weight in a 2-population simulated network using the pipeline in **a**. **c** Spiking activity of 6 subpopulations in a simulated laminar network. Connectivity is visualized in the overlaid schematic. **d** Directed dependencies (edge in the graph) in the simulated network in **c**, estimated using MTwDBN fitting. **e** Summary graph of dependencies across all timescales from **d**. **f** F-score (harmonic mean of precision and recall of dependency structure) as a function of % of neural population observed. F-score was estimated for shuffle corrected weighted DAGs (MTwDBN, green), weighted DAGs with fixed threshold (weightedFT, blue), unweighted DAGs (red), or LASSO regression, an example of regularized regression (RR) models (black). Each point represents the average of five separate runs, except 100 % (single run). Error bars indicate standard deviation, some error bars are smaller than symbol size. **g** DBN decoder accuracy with different sizes of MTwDBN DAGs. Decoders were trained to predict population activities using a subsample of shuffle corrected edges (see SI Methods). Graph edges for the decoder were sampled from the learned structure either in an unbiased fashion (black) or biased with the edge weights (green).

Figure 3

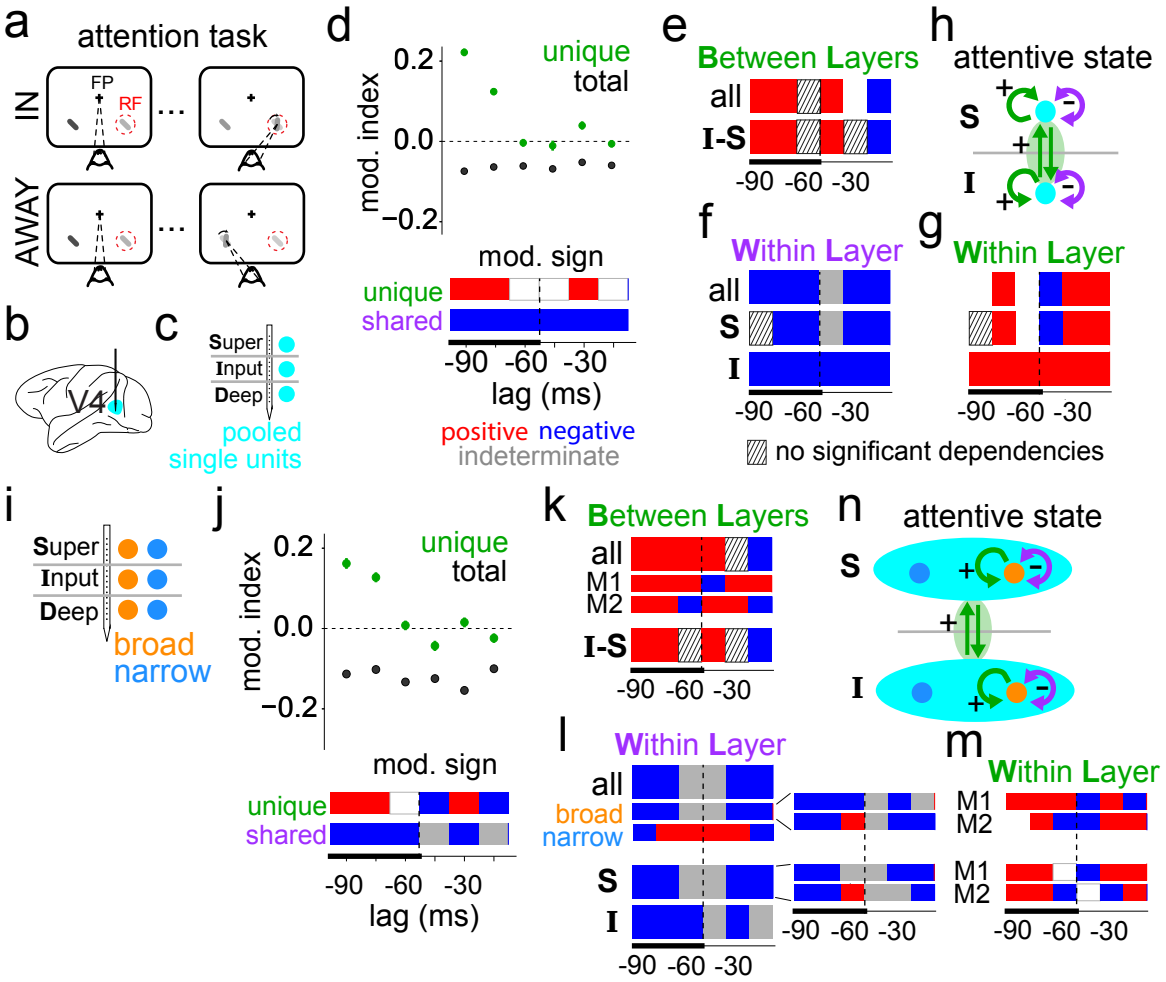

Figure 3 Modulation of dependencies in a V4 laminar network across attention conditions. **a** Experimental protocol: Paired Gabor stimuli with different contrasts (see Methods); one stimulus was presented inside the receptive fields (RFs) of the recorded neurons and the other at an equally eccentric location across the vertical meridian. Attention was cued either to the neurons' RFs ("IN") or to the location in the contralateral visual hemifield ("AWAY"). Orientation of one of the two stimuli changed at a random time. Monkey was rewarded for detecting the change by making a saccade to the corresponding location. Task difficulty was controlled by the magnitude of orientation change. **b** Recording approach: laminar recordings in visual area V4. **c** Neural populations used for MTwDBN analysis. Current source density analysis identified different layers (*superficial*, *input*, *deep*), and isolated single units were assigned to one of these layers (see SI Methods). **d** *Top*: Average MTwDBN-based modulation (green) of all unique dependencies between the laminar populations. Modulation of the same dependencies as estimated by logistic regression (black). *Bottom*: PID framework-based estimated modulation sign of shared dependencies (see schema in Fig 1e). Thicker line along the time axis indicates the timescales of attentional modulation in prior studies[38,39]. Error bars indicate 95% confidence intervals. **e** Sign of average modulation of unique dependencies between layers (bi-directionally). **f** Sign of average modulation of shared dependencies within layers. **g** Sign of average modulation of unique dependencies within layers. **h** Summary of dependency modulation pattern. **i** Neural populations used for laminar MTwDBN analysis. Isolated single units were classified as broad and narrow based on peak-to-trough duration in their average spike shape (see SI Methods). **j-n** Same as **d-h**, but for populations shown in **i**. I: input layer; S: superficial layer; *M1*, *M2*: subject-wise dependencies; *broad*, *narrow*: cell-class specific dependencies. See Fig S4 for modulation indices in **e-g**, **k-n**.

Figure 4

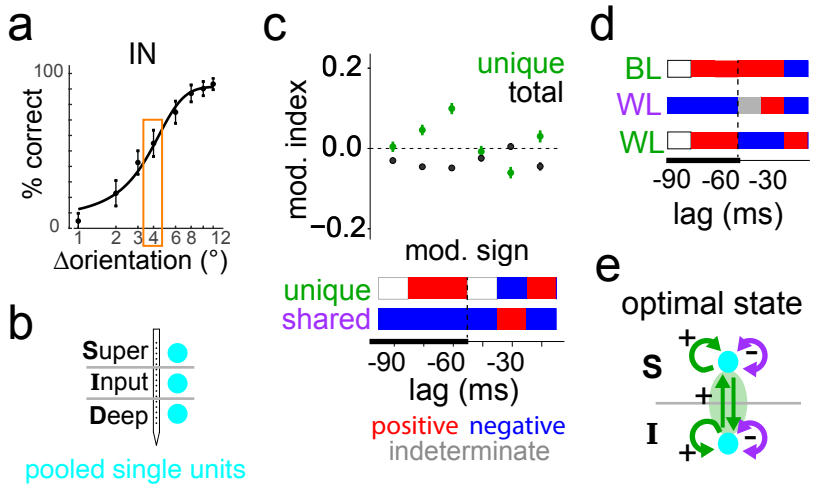

Figure 4 Modulation of dependencies in a V4 laminar network across behavioral outcomes at perceptual threshold. **a** Example session showing performance as a function of task difficulty. (orange box: threshold orientation change at which the animal was equally likely to correctly detect (Hit) or fail to detect (Miss) the change). **b** Laminar populations used for multi-lag analysis. **c** Modulation magnitude (top) and sign (bottom) of all unique (green) and total (black) laminar dependencies in **b** across Hits and Misses at perceptual threshold. Estimated modulation sign of shared dependencies (bottom, see Fig 1e). **d** Modulation sign of between (BL) and within (WL) layer dependencies. See Fig S5 for modulation indices. **e** Summary of laminar dependency modulation pattern.

Figure S1

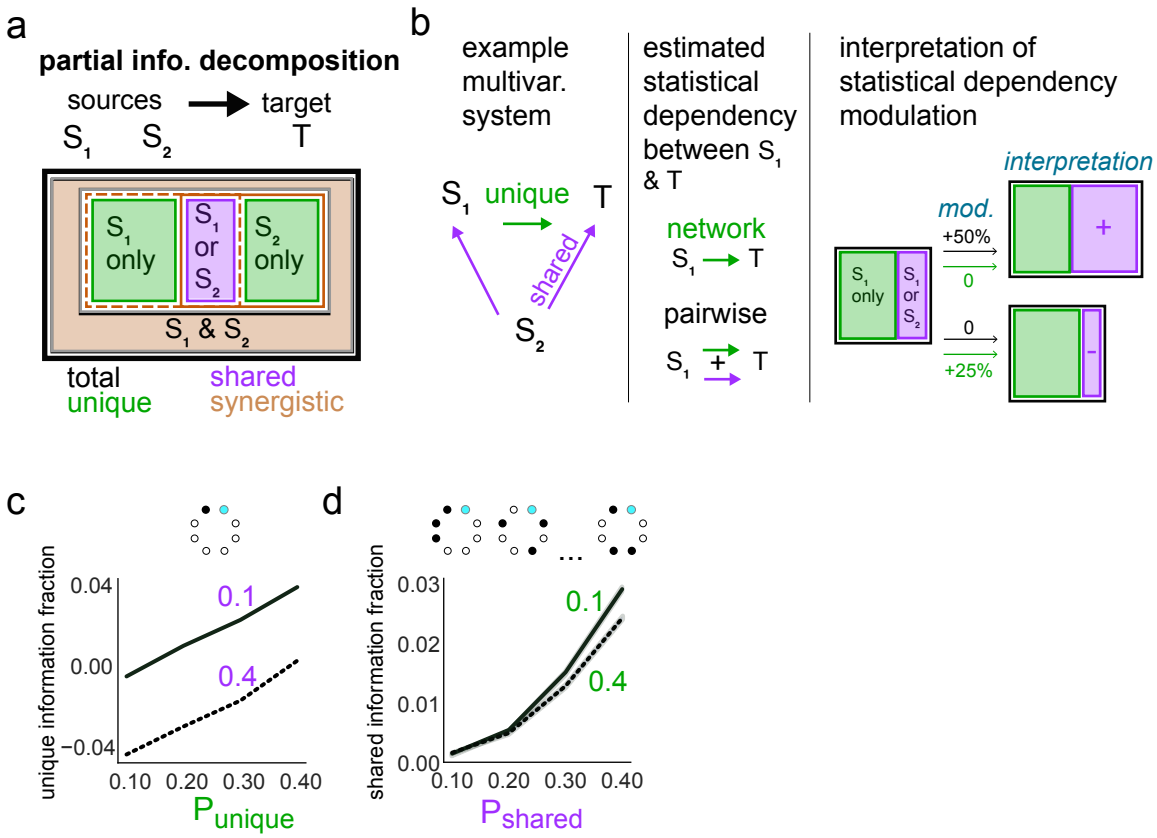

Figure S1 Information decomposition in a multi-variate system

a Partial information decomposition (PID) framework[13,14] based definition of types of information that multiple sources can have about a target.

b Application of pairwise and network-based statistical models for approximate information decomposition in a multivariate system. Panel on the right illustrates interpretation of the modulation of these dependencies using the PID framework.

c, d Reduction in the proportion of total entropy (or information fraction) as a function of parameters (P_{unique},P_{shared}) that control unique and shared information in the model. For the sake of computational efficiency, information fraction estimation as a function of P_{shared} (**d**) was performed using a subset of variables in the simulated network.

Figure S2

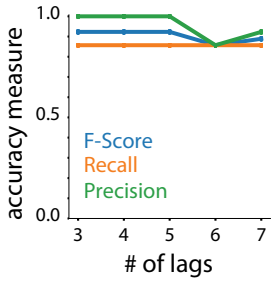

Figure S2 Additional MTwDBN performance metric for synthetic data
Recall, precision and F-score measures for weighted DAGs in Fig 2d as a function of number of lags.

Figure S3

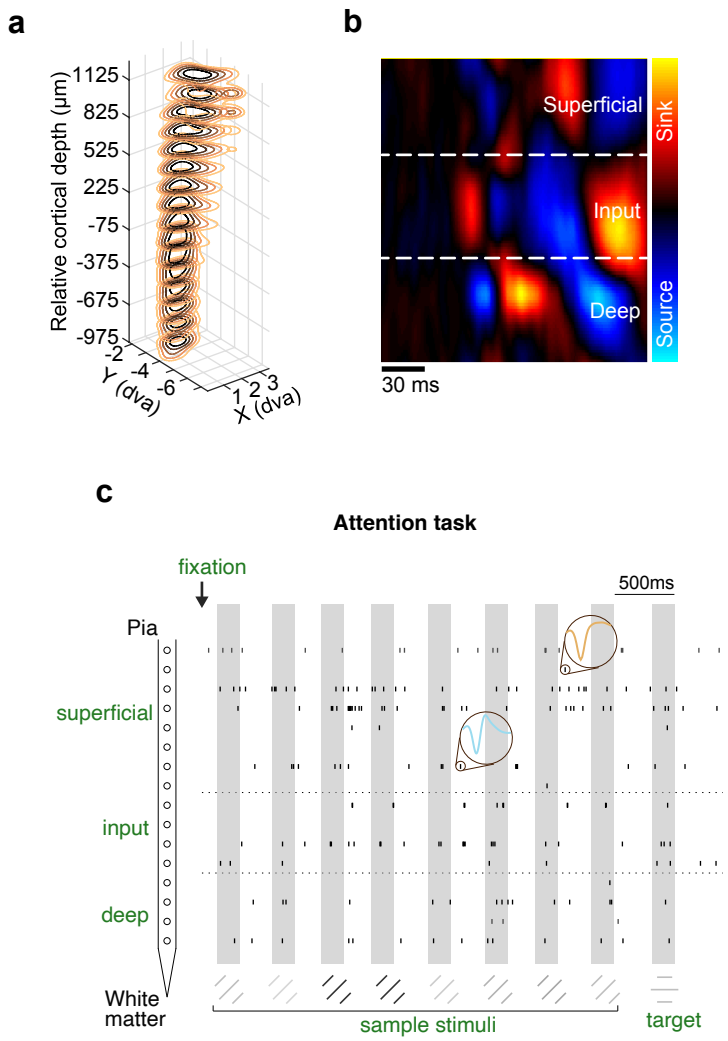

Figure S3 Laminar recordings in area V4

a Stacked contour plot showing spatial receptive fields (RFs) along the laminar probe from an example session. The RFs are well aligned, indicating perpendicular penetration down a cortical column. Zero depth represents the center of the input layer as estimated with current source density (CSD) analysis. **b** CSD displayed as a colored map. The x-axis represents time from stimulus onset; the y-axis represents cortical depth. The CSD map has been spatially smoothed for visualization. **c** An example trial showing single-unit activity across the cortical depth in the attend-in condition. The time axis is referenced to the appearance of the fixation spot. Spikes (vertical ticks) in each recording channel come from the single unit with the highest spike rate in this trial. The gray boxes depict stimulus presentation epochs. In this particular trial, 8 sample stimuli with different contrasts were presented, followed by a target stimulus flash with an orientation change that the monkey responded to correctly. Spike waveforms for two example narrow (cyan) and broad (orange) single units are shown.

Figure S4

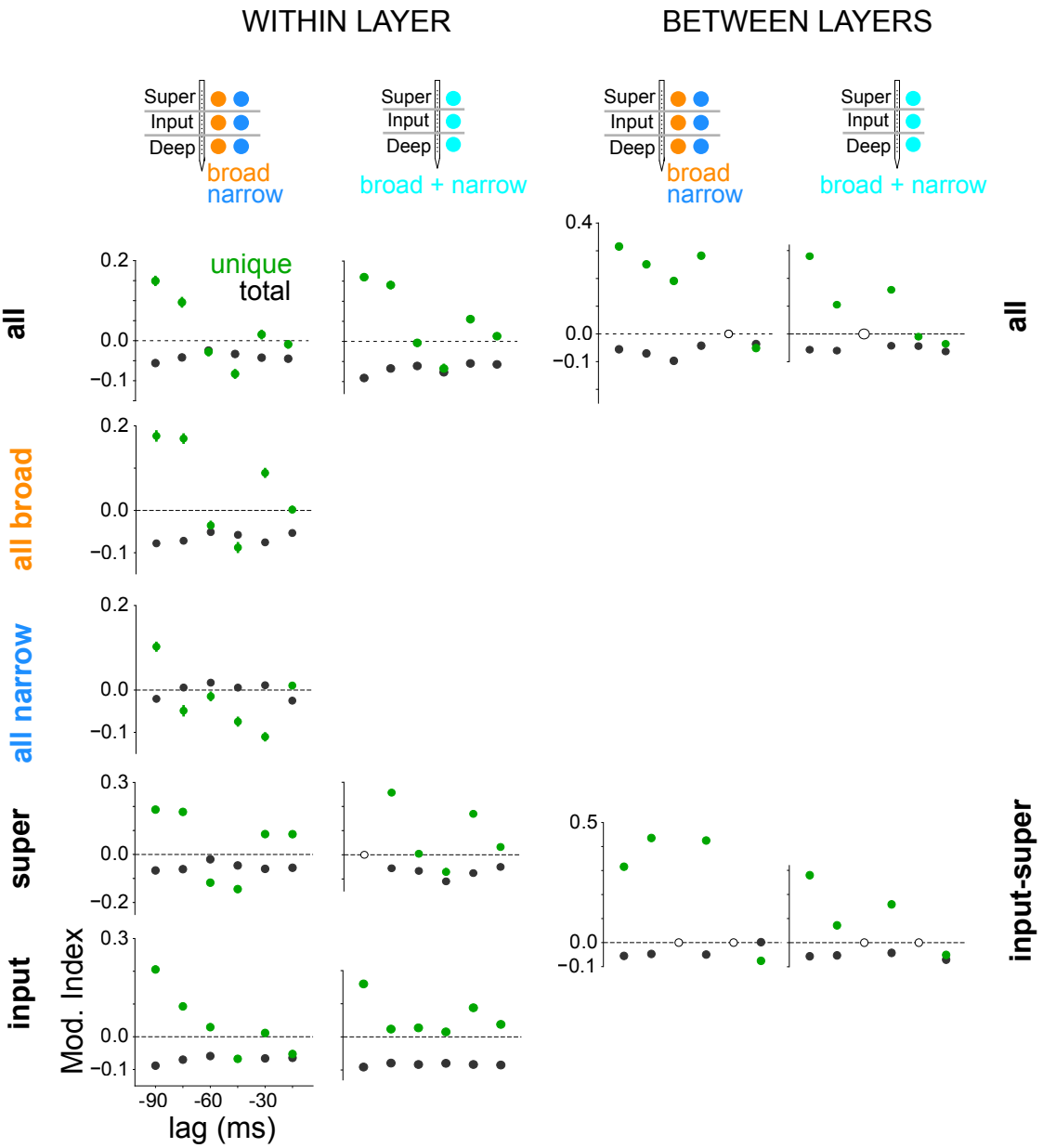

Figure S4 Modulation of dependencies in a V4 laminar network across attention conditions. MTwDBN-based modulation (green) of all unique dependencies between the laminar populations. Modulation of the same dependencies as estimated by logistic regression (black). Open circles indicate no significant edges for a given time lag. Schematic at the top of every column depicts the populations used for identifying laminar dependency structure within (left) and between (right) cortical layers. *broad*, *narrow*: cell-class specific populations identified by spike shape (see SI Methods).

Figure S5

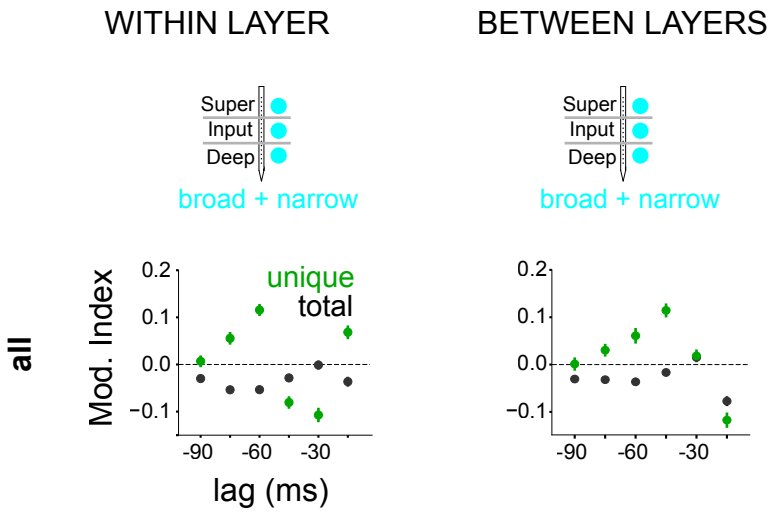

Figure S5 Modulation of dependencies in a V4 laminar network across behavioral outcomes at perceptual threshold. MTwDBN-based modulation (green) of all unique dependencies between the laminar populations. Modulation of the same dependencies as estimated by logistic regression (black). Schematic at the top of every column depicts the populations used for identifying laminar dependency structure within (left) and between (right) cortical layers.