**Figure 1. Two-photon imaging of population activity in primary visual cortex in awake behaving mice**

(A) Illustration of the experimental setup. Mice were head fixed to a two-photon microscope, and were allowed to run on a treadmill. Visual stimuli of drifting gratings were presented on a screen to the monocular side of the mice. (B)A representative field of view with detected ROIs. Scale bar represents xxx. (C) Representative raster plot of extracted spikes during spontaneous activity or under visual stimuli. Mouse was shown with horizontal and vertical drifting gratings in an alternative fashion. Light red and blue vertical stripes indicate the time of horizontal and vertical stimuli, respectively. Scale bar represents 100 frames. (D) Schematic of generating the dissimilarity matrix for finding neuronal ensembles. Binary spike vectors for each frame were normalized using TF-IDF, and the cosine distance between every pair of normalized vectors were calculated, which was further used as the input to SVD. (E) Example of visual ensembles identified using SVD, from the experiment shown in C. Red cells represent the ensemble of horizontal visual stimuli; blue cells represent the ensemble of vertical visual stimuli. (F) Activities of visual and spontaneous ensembles. Cells in the two visual ensembles tend to be co-active during their corresponding visual stimuli, whereas cells in spontaneous ensembles exhibit more irregular patterns.

**Figure 2. Identification of ensembles using conditional random field model**

(A) Illustration of conditional random field model. Shaded nodes (x) represent the observed binary spiking state of the neurons. White nodes (y) represent true states of the neurons, and are connected by edges that indicate their mutual dependencies; node potentials are defined over the two possible states of each node, and edge potentials are defined over the four possible states of each existing edge, depending on the state of the two nodes it connects. The probability distribution of the network over all possible states can therefore be factorized using the node and edge potentials. (B) Illustration of defining ensembles by identifying group of neurons that are likely to be co-active. The constructed graph is simplified by only keeping the edge potentials of 1-1 state, and a subgraph with the maximum edge potentials is found. The constituent nodes in this subgraph are defined as an ensemble. (C) Illustration of defining multiple ensembles by stimulus association. An additional stimulus-associated neuron (xs) is added for each visual stimulus, and the observation of this stimulus-associated neuron is set to 1 when the corresponding visual stimulus is present. Visual ensembles are defined as the nodes that are connected with the corresponding stimulus nodes and have a positive 1-1 edge potential.(D) An example of two visual ensembles identified using the method described in (C). Red and blue squares represent the latent node associated with horizontal and vertical visual stimuli, correspondingly. Red and blue circles represent the identified ensemble of horizontal stimuli and vertical stimuli, correspondingly.

**Figure 3. k-clique community properties**

(A) Illustration of k-clique structures and the corresponding adjacency matrices when k = 3, 4, 5. (B) Illustration of a 3-clique community. Nodes of the same color belong to the same 3-clique community; nodes with multiple colors are shared between communities indicated by the color. (C) Number of k-cliques in graphs constructed with spiking data and with random data. There are significantly more k-cliques when k<5 in the real data-based graphs, indicating a non-random structure. (D) Number of k-clique communities in graphs constructed with spiking data and with random data. There are significantly more k-cliques for all tested k values in the real data-based graphs, indicating a more localized structure. (E) Distribution of 3-clique community size in random and real data. The heavy tail in the distribution from real data indicates the existence of generally larger communities.

**Figure 4. Graph properties of CRF models and correlation-based models**

(A) An examples of graphs constructed with CRF model and with pairwise correlations (CC). In the latter case, pairwise Pearson correlations between frames were calculated, and the threshold is determined by 5% significance level of correlation values of shuffled data. Here node size is proportional to the node degree, and the edge color represents the synchrony edge potential (ф11+ ф00- ф01- ф10). (B) Cells participated in at least one community in the above CRF and CC models are shown here. (C-I) The complementary cumulative distribution of node degrees (C), local clustering coefficients (D), community size (E), community degree (F), community overlap (G), and community membership (H) in CC and CRF model. Both models show more dispersed distributions than random models, while the properties of CRF model differ from CC model.

**Figure 5. Predicting visual stimuli with identified ensembles**

(A) Examples of ensembles identified using SVD with TF-IDF normalization, stimulus-associated group in CRF graphs, co-active group in CRF graphs, community membership and eigenvector centrality in both CRF and CC graphs, and high OSI cells. Red and blue circles represent the ensemble cells of horizontal and vertical visual stimuli, respectively; black circles represents the cells shared by the two ensembles. (B) Percentage of overlap between high OSI cells and ensembles identified by the other methods. Percentage is calculated by number of cells belong to both high OSI group and ensemble group divided by the number of cells belonging to high OSI group or ensemble group. (C)Percentage of overlap between CRF stimulus-associated group and ensembles identified by the other methods, calculated similarly as above. (D) Raster plot of prediction and frame cosine similarities calculated with different models. Identified ensembles were represented by a binary vector over all neurons, and the entries corresponding to the ensemble members were set to 1, while the rest were set to 0. Cosine similarities between ensemble vectors and frame activity vectors were calculated, and a threshold was determined by 3 times the standard deviation of baseline noise. The cosine similarity between two frame activity vectors depicts the angle between two vectors in the high-dimensional space; orthogonal angles indicate that the active neurons in the two frames are mostly different, while small angles indicate that the active neurons are mostly the same. Frames that were significantly similar compared with the threshold were taken as stimulus-on frames. Black bars in the raster plot indicate the predicted stimulus-on frames. Light red and blue vertical stripes indicate the time of horizontal and vertical stimuli, respectively. (E) Cosine similarities of frames with stimuli that match or not match the model. SVD, CRF co-active group, and CRF stimulus-associated group show significantly higher similarities with matching frames than non-matching frames. (F-H) Accuracy (F), precision (G) and recall (H) of predictions from different models. SVD and CRF stimulus-associated group show better performance than other models.

**Figure 6. Ensembles are composed of groups of neurons that are specific to the stimulus**

(A) Examples of identified ensemble cells, 10% subsample of ensemble cells, randomly sampled cells of 20% ensemble cell number, and randomly sample cells of 90% ensemble cell number. Their corresponding prediction raster plot and frame cosine similarities are shown on the right. (B) Mean cosine similarity of randomly down-sampled or up-sampled ensemble groups, in frames with matching or non-matching stimulus. (C-D) Accuracy (C), precision (D) and recall (E) of predictions from randomly down-sampled or up-sampled ensemble groups. (F) Mean cosine similarity of randomly sampled cells, in frames with matching or non-matching stimulus. (G-I) Accuracy (G), precision (H) and recall (I) of predictions from randomly down-sampled or up-sampled ensemble groups.