**Supplementary Figure S1. Cortical ensembles as a representation of multidimensional population vectors obtained with two-photon calcium imaging**

(A) Schematic representation of active neurons at different frames. Black dots represent active neurons at different times (left). Binary raster plot representing the overall network activity of observed neurons (right). Population vectors capture the coordinated activity of a given neuronal ensemble. (B) Population vectors can be understood as a multidimensional array in which clusters of population vectors taken from different times define network states.

**Supplementary Figure S2. Hidden nodes does not affect graph properties**

(A) CRF graphs of baseline model (no hidden nodes) and hidden node model, trained with the same experiment. Edge color represents the strength of ; node color represents the rank of cells. Node size represents the node degree. (B) Graph densities, (C) edge potentials, (D) node degrees, (E) clustering coefficients and (F) centrality do not differ significantly between the two models.

**Supplementary Figure S3. Comparison of prediction performance with CRF, SVD ensembles and high OSI cells**

(A) Examples of core neurons identified using SVD (green) and CRFs (orange). Circles represent core neurons of horizontal (left) and vertical (right) visual stimuli, respectively. Neurons shared between CRFs and SVD methods are represented by green and orange dots. (B) Percentage from the total population size representing core neurons for SVD and CRFs methods. (C) Percentage of shared neurons identified by SVD and CRFs methods. Percentage is calculated by number of cells that belong to both methods divided by the total number of unique cells in both methods. (D) Cosine similarity between population vectors that belong to given visual stimuli (color) compared to population vectors from different visual stimuli (black). (E) Accuracy, (F) precision and (G) recall of predictions for each visual stimuli using both methods.

**Supplementary Figure S4. Minimum dataset size to obtain stable performance**

(A) Similarity, (B) accuracy, (C) precision and (D) recall of CRF ensemble prediction reach stable with a minimum of 200 frames training data. The data used in this experiment contains 101 neurons.

**Supplementary Figure S5. CRF ensembles are able to predict multiple stimuli**

(A) An example of constructed CRF graph from the Paul Allen dataset, with four visual stimulus (squares). Edge color indicates the strength of inferred connections; node size indicates the node degrees. (B) Prediction of visual stimulus in withheld test data, using the full model. Background stripe color represents the stimulus type. (C) Accuracy, (D) precision and (E) recall of prediction with CRF model trained with temporal frequency 1Hz, on dataset of temporal frequency 2, 4, 8, and 15Hz.