Choice is represented by ensembles of bursting neurons in the superficial layers of monkey V1

Veronika Koren [1,2,*], Ariana R. Andrei [3], Ming Hu[4], Valentin Dragoi [3], Klaus Obermayer [1,2]

1 Neural Information Processing Group, Institute of Software Engineering and Theoretical Computer Science, Technische Universität Berlin, Berlin, 10587, Germany **2** Bernstein Center for Computational Neuroscience Berlin, Germany **3** Department of Neurobiology and Anatomy, University of Texas Medical School, Houston, Texas, 77030, US **4** Baylor College of Medicine, Houston, TX 77030, US

[*] Correspondence: veronika.koren@ni.tu-berlin.de

Highlights

- The choice can be predicted from the spiking activity in the cortical column of V1 of the macaque.
- The representation of the choice relies on a partial overlap between the representation of the stimuli and the choice.
- Bursty neurons in the superficial layer of the cortex are the principal carriers of the choice signal.
- The choice-related information likely comes from a top-down source.

Abstract

Primary visual cortex (V1) is absolutely necessary for normal visual processing, but whether V1 encodes upcoming behavioral decisions based on visual information is an unresolved issue, with conflicting evidence. Further, no study so far has been able to predict choice from time-resolved spiking activity in V1. Here, we hypothesized that the choice cannot be decoded with classical decoding schemes due to the noise in incorrect trials, but it might be possible to decode with generalized learning. We trained the decoder in the presence of the information on both the stimulus class and the correct behavioral choice. The learned structure of population responses was then utilized to decode trials that differ in the choice alone. We show that with such generalized learning scheme, the choice can be successfully predicted from spiking activity of neural ensembles in V1 in single trials, relying on the partial overlap in the representation between the stimuli and the choice. In addition, we show that the representation of the choice is primarily carried by bursting neurons in the superficial layer of the cortex. We demonstrated how bursting of single neurons and noise correlations between neurons with similar decoding selectivity helps the encoding and transmission of the choice signal downstream. By studying the flow of V1-intrinsic activity across lamina, we show that flow of activity differs across stimuli but not across choices, suggesting a top-down origin of the choice signal.

Keywords: Visual cortex; cortical column; decoding; stimulus; choice; computation; representation; neural network; population code; information; discrimination

Introduction

In real-world scenarios, animals frequently make decisions based on the observation of their environment. In the context of such decision-making, it has been understood that encoding of sensory information and utilizing this information for guiding the behavior do not necessarily overlap [1]. So far, the primary visual cortex has been shown to be a site of relatively sophisticated sensory processing, from which other brain areas would read-out the sensory information in order to perform further computations, such as decision-making based on sensory cues. Corroborating this view has been the evidence of the choice probability in visual areas in the middle and higher in the cortical hierarchy and as "low" as V2, but not from V1 [2]. Recent study, however, has found the choice probability in V1 during a coarse orientation-discrimination task with artificial stimuli [3], arguing that, in order to observe choice probability, neurons must be topologically organized with respect to the task feature.

Representation of the behavioral choice in absence of sensory evidence has been extensively studied for three decades [4, 5, 6, 7, 8, 9]. Studies on choice probability have primarily addressed the time-averaged activity of single neurons and their relation to the behavioral choice of the animal, and single neurons have been shown to correlate with the choice behavior in multiple brain areas [4, 5, 7, 10, 8]. Even if it is possible to find neurons whose activity predicts the behavior with better-than-chance probability, it is not clear how would the brain isolate the activity of a single neuron from a vast number of cells that activate in parallel. Recent years have seen the revival and the development of the idea that behaviorally relevant variables are encoded by collective dynamics of neural ensembles, evolving in real time (see, e.g., [11, 12, 13, 14, 15]). While a major obstacle for testing theoretical predictions on experimental data is the incompleteness of datasets, recent analysis of a nearly complete networks in the retina demonstrates that a good description of neural activity has to take into account single neuron firing rates, the strength of pairwise interactions and a global modulation of the network activity ([16, 17]). Population codes seem to be appropriate to study neural responses in the auditory [18, 19], motor [20], prefrontal [21], and visual cortex [14].

In the present study, we ask if the behavioral choice can be predicted from the spiking activity of neural populations in V1, with a time-resolved code that reflects the instantaneous computation of the low-dimensional choice variable in animal's brain. Such a computation presumably utilizes spikes of neural populations and takes place in real time, in order to contributes to animal's behavior as it unfolds. Moreover, we use naturalistic stimuli that are close to the psychophysical threshold. The requirement of the time-resolved population code, and the use of naturalistic stimuli close to the psychophysical threshold have, to our not knowledge, not yet been shown to yield the prediction of the choice in V1. We show that the choice can be computed and successfully discriminated, not with classical decoding procedures but only when the neural representation is learned through the extension of classical decoding schemes.

In the experimental setup of the current study, animals sequentially view pairs of natural visual stimuli that can differ in their orientation. Animals then decide and report whether pairs were a "match" or "non-match". The choice of the animal is either compatible with the visualized stimulus class, leading to correct behavior, or not, leading to incorrect behavior (table 1). We can consider two classification problems, one where classification relies on the information on both the stimulus and the choice ($stimulus + choice$), and another, where the classification relies on the information about the choice alone (*choice*, table 1). Classical decoding procedure has been followed in our previous work, where we found that spike counts in V1 do allow the prediction of the binary variable $stimulus + choice$ ([22]). Read-out of the spike trains (without temporal averaging) is also successful in predicting the variable $stimulus + choice$ in single trials (see fig.2 in [23]).

In this study, we address the representation of the choice in absence of the information on the stimuli (representation of the variable *choice*, table 1). We find that the variable *choice* can be predicted through generalization of learning of the neural representation [24] from $stimulus + choice$ to *choice*. We implement such generalization (or transfer) of learning by training an optimal linear decoder in the context of $stimulus + choice$, and applying the structural features, extracted by the decoder, to construct a low-dimensional representation of the variable *choice*. If the representation in the context of $stimulus + choice$ exclusively relies on the difference in stimuli, learning to discriminate the variable $stimulus + choice$ cannot be generalized to the discrimination of

the variable *choice*. If, on the contrary, the representation of the choice and the stimulus are not independent, learning in the context of *stimulus* + *choice* can generalize to the representation of *choice*, due to the partial overlap in the information between the two classification problems. This suggests that, contrary to previous beliefs, upcoming behavioral choice can be decoded from the V1 spiking activity in a time-resolved population code.

Stimulus	Choice	Performance	Info. content
non-match	"different"	correct	stimulus+choice
match	"same"	correct	
non-match	"different"	correct	choice
non-match	"same"	incorrect	

Table 1. Informational content of two classification problems.

Our recordings yield the spiking activity of a small neural population, spanning the depth of a single cortical column. We model the synaptic current at the downstream neuron, receiving direct synaptic projections from the recorded population. In the cortex, synaptic inputs to a read-out neuron are not instantaneous events, but create a short depolarization of the membrane potential at the read-out neuron, followed by the decay towards the resting potential. This short period of integration of synaptic inputs is important, as it implies that temporal patterns of synaptic inputs can either help or hinder the formation of the choice signal. If synaptic inputs from neurons with the same choice preference are received within a short time window, this allows the formation of the choice signal. On the contrary, if spikes of neurons signaling opposite choices are received within a short time window, the information about the choice cancels out and the choice signal remains close to zero. The formation of the choice signal can rely on two dynamical mechanisms, bursting of single neurons and correlation of spike timing between neurons with the same preferred choice. Our results strongly suggests that the brain employs both of these mechanisms to encode the choice in real time.

In the present work, we show that the behavioral choice can be predicted from the spiking activity in the primary visual cortex of the macaque. We demonstrate how successful prediction of the choice relies on the partial similarity of representation of the stimuli and the choice, in the sense of [1]. We show that the choice information is almost exclusively present in the superficial layer of the cortex and primarily in bursty neurons. We show how bursting and correlated spiking of neurons with similar choice selectivity allow the choice signal to be formed. Finally, our results suggest that the spiking activity intrinsic to V1 does not underlie the computation of the choice by itself, but rather reflects a top-down projection from downstream brain areas.

Results

Behavioral choice can be discriminated from V1 spiking activity through generalized decoding

Two adult male macaques were trained on a delayed match-to-sample visual task. The subject visualized the target and the test stimuli, with a delay period in between (fig.1a). The target and the test stimulus were either identical (condition match) or not (condition non-match), with the only possible difference being the change in the orientation of the test stimulus. The task of the subject was to communicate its decision about the stimulus class ("same" or "different", see methods). The stimuli were complex naturalistic images in black and white, depicting an outdoor scene, and their identity changed from one recording session to another. The multiunit signal has been captured with linear arrays, inserted perpendicularly to the cortical surface and recording the activity across the cortical depth. After spike sorting, we obtained on the order of 10 units in each recording session, and the total of 160 units in 20 recording sessions. Example spike rasters for three

simultaneously recorded neurons, belonging to the superficial (supragranular or SG), middle (granular or G) and deep (infragranular or IG) cortical layer, are shown on fig.1b.

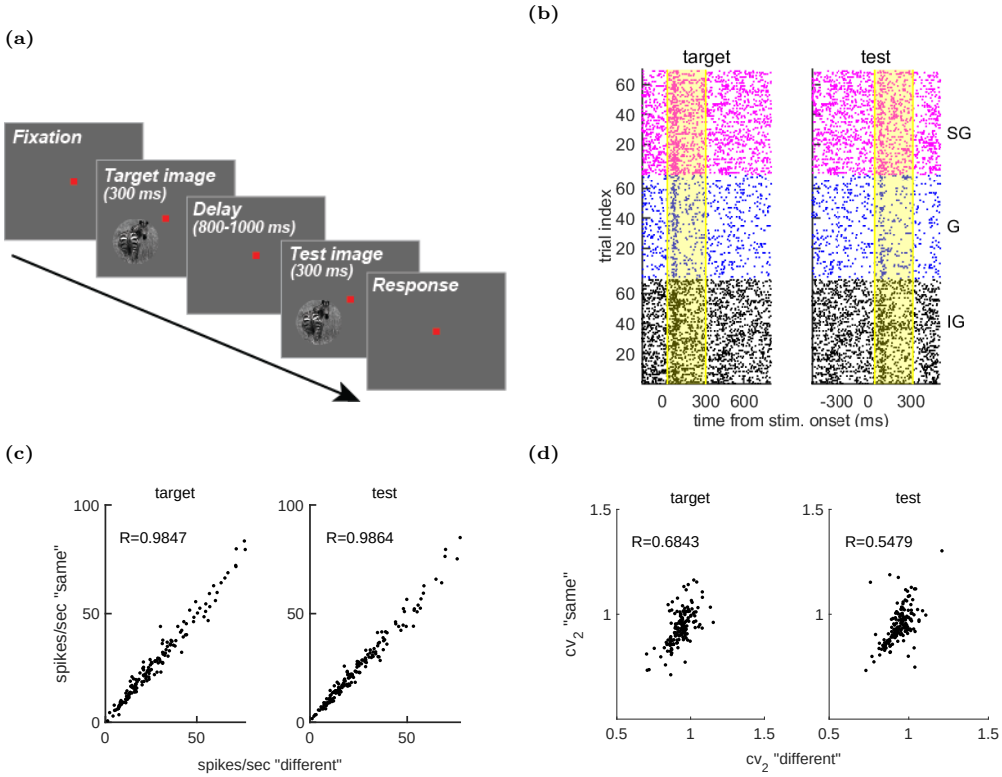

Figure 1. Experimental paradigm and single neuron statistics. (a) Experimental paradigm. The subject visualizes the target and the test stimulus, with a delay period in between. **(b)** Spike rasters of three representative single neurons for the behavioral choice "different". We show spike rasters of three neurons, one in superficial layer of the cortex (SG, magenta), one in the middle layer (G, blue) and one in the deep layer (IG, black). **(c)** Scatter plot of the mean firing rate for the choice "different" (x-axis) and for the choice "same" (y-axis). We used the time window [0,400] ms w.r.t. the stimulus onset. R is the Pearson correlation coefficient. **(d)** Same as in **(c)**, but showing the coefficient of variation (CV_2).

Single neuron statistics (fig.1c-d) and standard multivariate decoding methods from populations of simultaneously recorded neurons (fig.2) do not allow to predict animal's choice. Firing rates of single neurons vary strongly across units, but are very similar for the choice "same" and "different" (fig.1c). Coefficient of variation of single neurons (CV_2, see methods) varies around 1, which is the CV_2 of the Poisson process, and is strongly positively correlated across conditions (fig.1d). Prediction of the *choice* with a linear decoder utilizing parallel spike counts is not significantly different from chance (average prediction accuracy of 0.51; fig.2a, black crosses).

The failure of predicting the variable *choice* from spike counts demonstrates that spike counts of neural populations do not differ between the choice "same" and "different" in the absence of the difference in stimuli. The structure of spike counts, however, does differ between the two choices when the choices are both correct, since the variable $stimulus + choice$ is predicted better than chance with the same classification model (fig.2a, red crosses). Simple explanation of these results could imply that populations in V1 represent the stimulus class, but not the choice. An alternative explanation is that the choice is represented, but we cannot extract it by decoding conditions correct and incorrect non-match, due to the noise that comes along with incorrect behavior. We reason that the representation of the choice could rely on partial similarity of the representation of $stimulus + choice$ and $choice$.

To see whether there is any similarity in representation between the two classification problems, we compare the structure of population responses for $stimulus + choice$ vs. $choice$. The structure of population responses is captured with population vectors, $\tilde{\mathbf{w}}^{S+C}$ and $\tilde{\mathbf{w}}^{C}$, and the angle between them (fig.2b) tells to what extent the structure of parallel spike counts is similar from one context to the other. If population vectors were independent, the angle between them would be orthogonal, $\alpha^{\text{independent}} = \pi/2$, while deviation from orthogonality shows the similarity of vectors $\tilde{\mathbf{w}}^{S+C}$ and $\tilde{\mathbf{w}}^{C}$. We find that, on average, the angle between the two weight vectors weakly

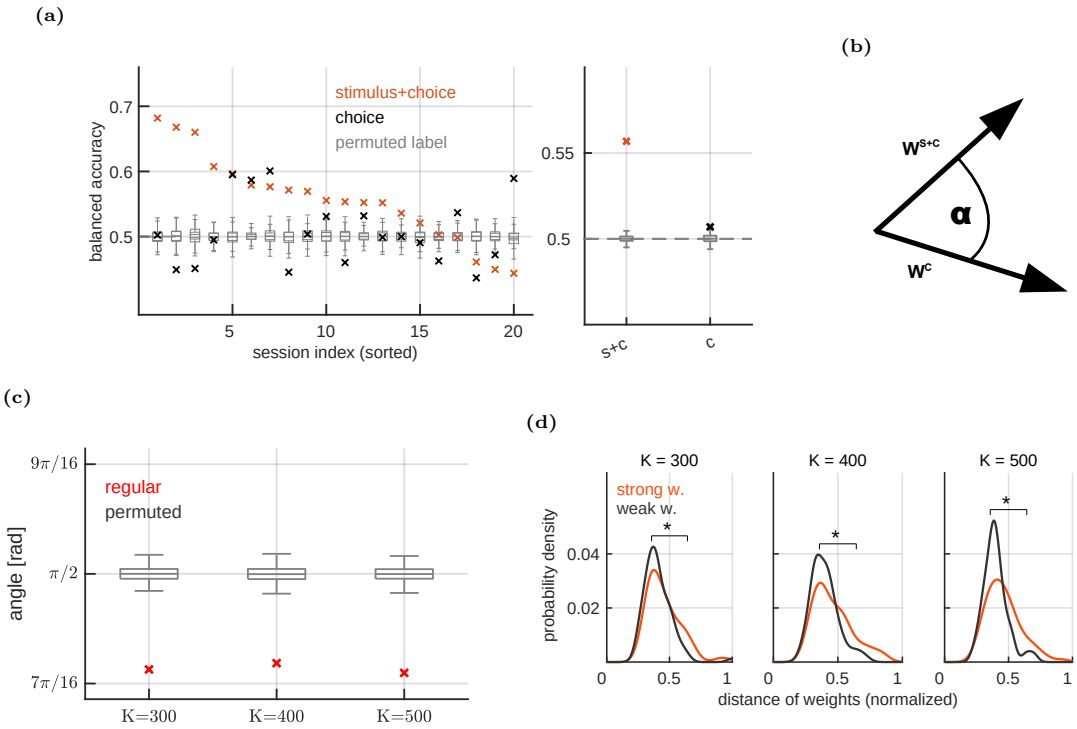

Figure 2. The structure of population responses in *stimulus + choice* **and in** *choice* **is weakly similar. (a)** Prediction accuracy of a linear classifier for the classification problem of *stimulus + choice* (red) and *choice* (black). Left plot shows results in recording sessions and right plot shows session-averaged results. Gray boxplots show the distribution of results of models with permuted class labels. **(b)** Schema of the angle between the population vectors in the context of *stimulus + choice* and *choice*. Population vectors are computed in each recording session. **(c)** Average angle between population vectors (red cross) and the distribution of same results for models with random weights (gray box-plot), for the length of the time window $K = 300$ (left), $K = 400$ (middle) and $K = 500$ (right). Parameter: $N_{perm} = 1000$. **(d)** Distance of weights between the two classification problems for neurons with strong (orange) and weak weights (black). We show results for the length of the time window $K = 300$ (left), $K = 400$ (middle) and $K = 500$ (right).

but significantly deviates from orthogonality (fig.2c), meaning that the two weight vectors are weakly similar. The similarity of vectors is robust to the length of the time window ($\alpha = [80.2,\ 80.8,\ 80.0]$ deg for the length of the time window $K = [300,\ 400,\ 500]$ ms, respectively) and is significant in all cases ($p < 0.001$, in all cases, permutation test).

The angle between the population vectors is a population-wide measure. We also compare decoding weights of single neurons between the contexts of *stimulus + choice* and *choice* by measuring the distance between decoding weight in the two classification problems for each neuron (see methods). Bigger distance means that the neuron responds more differently in the two informational contexts and is a proxy for the sensitivity of the single neuron to the change in the informational context. We divide neurons in those with strong and weak weights (see methods) and test whether the two groups differ in their sensitivity to the informational context. We find that neurons with strong weights are more sensitive to the change in the informational context than neurons with weak weights ($p = 0.0092$ for K=300, $p = 0.0008$ for K=400, $p = 0.0004$ for K=500, one-tailed t-test; fig.2d). As neurons are split according to other criteria such as the sign of the weight (positive or negative), burstiness of their spiking (bursty or non-bursty), and position in the cortical layer (SG, G and IG layer), there is no significant difference between groups (Supplementary Fig.S1).

Rather than with firing rates, the brain might encode behaviorally relevant variables dynamically, that is, with a signal that rapidly changes within the trial. We therefore utilize a read-out schema without temporal averaging, where spike trains of the neural population are weighted, summed up and convolved with an exponential filter (fig.3a; see also [23]). This models the post-synaptic current at the read-out neuron, where summing the activity across neurons is justified by the relevance of pooling of neuronal signals for perception ([25]), and the convolution models the dynamics of the post-synaptic potential. Decoding weights correspond to

modification of synaptic weights during the training phase of the experiment, and therefore stem from a slower process. For this reason, decoding weights are learned on distribution of spike counts, collected across trials (see methods).

To address the generalization of learning, we introduce several modifications to the classical decoding scheme (fig.3b). While training the classification model, the model finds the structure in the neural activity that is the best for distinguishing imposed categories. Within a classical validation scheme, learned structure is then tested on a hold-out set of trials, using the same input statistics and the same categories as the ones used for learning. With the present method, we deviate from the standard decoding procedure in two ways. First, we train the model on spike counts, while we test it on spike trains. Second, the training and the test set also differ in categories that they decode, since the training is done on the variable *stimulus* + *choice* while the test set is applied on the variable *choice*. The result of our decoding scheme is the choice signal, representing the synaptic current at a hypothetical read-out neuron.

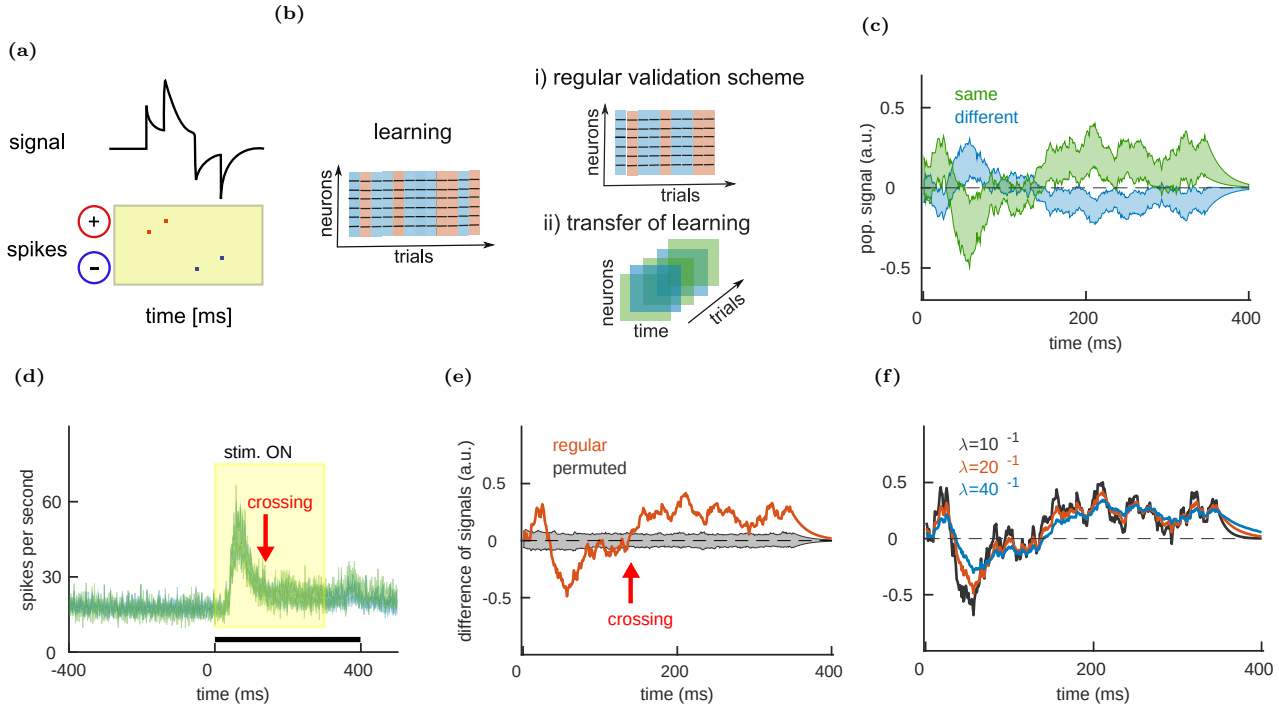

Figure 3. Behavioral choice can be discriminated with generalized learning of neural representation. (a) Schema of decoding by weighting spikes. Neurons with positive weights push the signal up and neurons with negative weights pull the signal down. **(b)** Decoding scheme. The learning utilizes samples of parallel spike trains (left). Following the regular validation scheme, the same type of data is used for validating the predictive power of the model on a hold-out set (top right). Here, we test the generalization (or transfer) of learning (bottom right), where the features of the learned classification model are applied on spike trains and on a different classification problem. **(c)** The choice signal for the choice "same" (green) and "different" (blue). We show the mean ± SEM for the variability across sessions. Parameter: $\lambda^{-1} = 20$. **(d)** Population PSTHs for choices "same" (green) and "different" (blue), with the mean ± SEM. The yellow window indicates the presence of the stimulus and the black bar indicates the time window, used for analysis. The red arrow marks the time of the zero crossing of the choice signal in **e**. **(e)** Session-averaged difference of the choice signals ("same" - "different") for the regular model (red) and the model with permuted class labels (gray). Parameters: $\lambda^{-1} = 20\ ms$, $N_{perm} = 1000$. **(f)** Difference of choice signals for three values of the time constant of the convolution.

Interestingly, we find that with generalized learning of the neural representation, the choice "same" vs. "different" can be successfully discriminated (fig.3c). The generalization of learning is necessary for predicting the choice. When the learning and the reconstruction step both utilize the variable *choice* (instead of learning on *stimulus*+*choice*), resulting signals are barely different from zero (Supplementary Fig.S2). Weighting spikes appears to be an essential step for discrimination, since the instantaneous firing rates for the two choices give highly overlapping signals (fig.3d). To asses the significance of the difference between choice signals, we compute the session-averaged difference of signals and compare it with the same result from models with random weights and randomly permuted class labels (see methods). The choice signals result significantly different during the

majority of the trial (fig.3e). The choice signals depend only weakly on the time scale of the convolution, with longer time scales resulting in a smoother signal (fig.3f). The discrimination is also robust to the change in the length of the time window that we use for decoding, as the choice can also be discriminated with time windows of 300 and 500 ms (Supplementary Fig.S3). We conclude that prediction of the choice from the spiking activity in V1 is possible with the generalized decoding scheme, and that learning of neural representation in the context of *stimulus* + *choice* successfully generalizes to the representation of the *choice* on a hold-out test set.

The difference of signals crosses the zero at about 140 ms after the stimulus onset (fig.3e, red arrow), implying a switch in decoded choice. We notice that this moment corresponds to the qualitative changes of the spiking activity from transient response after the stimulus onset to tonic firing thereafter (fig.3d, red arrow). The switch in decoded choice raises the possibility of conflicting information from feedforward and feedback inputs to the network. Moreover, we note that during the target time window, the information that is necessary for discrimination of matching from non-matching stimuli is not yet available to the visual system, and consistently, prediction of the choice behavior is not possible (Supplementary Fig.S4).

In spite of its simplicity, the decoding model has several components that may or may not be necessarily required for discrimination. In order to test the necessity of each component, we remove the information of a specific component and test the effect of such a perturbation on discrimination. We find that the sign of the decoding weight and the correct class label for spike trains (in the reconstruction step) are necessarily required for discrimination (fig.4a). Next, we test the necessity of correct spike timing for discrimination. As we randomly permute the spike timing across the entire time window of 400 ms, we get highly overlapping signals for conditions "same" and "different" (fig.4b, top). Interestingly, if the spike timing is permuted in the time window only before and after the switch in decoded choice (140 ms w.r.t. stimulus onset), the overlap is largely diminished (fig.4b, bottom), suggests different sources to the choice signal before and after the switch. Finally, we ask how does the temporal jitter in shorter time windows affect the read-out and find that the jitter makes the choice signals noisier, but still allows discrimination (fig.4c). The read-out of the choice is therefore robust to the perturbation of the spike timing on short time scales.

Noise correlations help the construction of the choice signal from spike.

The sign of the weights is the essential part of the decoding model, since neurons with the opposite sign of the weight have the opposite effect on the classification model, pulling the separation boundary in opposite directions (fig.4d). We separate the population with respect to the sign of the weight, getting subnetworks of neurons with positive (*plus*) and negative (*minus*) weights. As we compute the choice signal independently with each subnetwork (see methods), we find that the *plus* subnetwork discriminates the choice behavior better than the *minus* subnetwork (fig.5a). The choice signal of the *plus* subnetwork accounts for the zero crossing at around 140 ms after the stimulus onset (fig.5a, right). The *minus* subnetwork shows some discriminatory capacity at the very end of the trial (fig.5b, left), where the response goes in the opposite direction with respect to the signal of the *plus* subnetwork. Measuring the cross-correlation function of the signal of *plus* and *minus* subpopulations in single trials (see methods), we find that the choice signal of the two subnetworks is negatively correlated, consistently for different time constants of the convolution (fig.5c, top). This effect is strongly significant (fig.5c, bottom), and is a signature of competitive dynamics between the two subpopulations.

Better discrimination capacity of *plus* compared to *minus* neurons is not due to the difference in the number of neurons, since the number of neurons in each group is approximately balanced, and the distribution of weights is centered around zero ($p = 0.235$ for K=400, $p = 0.0842$ for K=500, t-test testing that the mean of the distribution is different from 0; fig.6a). We find there are two reasons for the signal of *plus* neurons to be more discriminative. First, the weights of *plus* neurons are more informative for classification of the stimuli. Separating neurons with positive and negative weights and decoding *stimulus* + *choice* separately with the information from each subnetwork (see methods), we find that the predictive power of neurons with positive weights is stronger compared to neurons with negative weights ($p < 10^{-12}$, two-tailed t-test; fig.6b). Second, patterns of spiking might be such as to help of hinder the accumulation of the choice signal. If neurons with

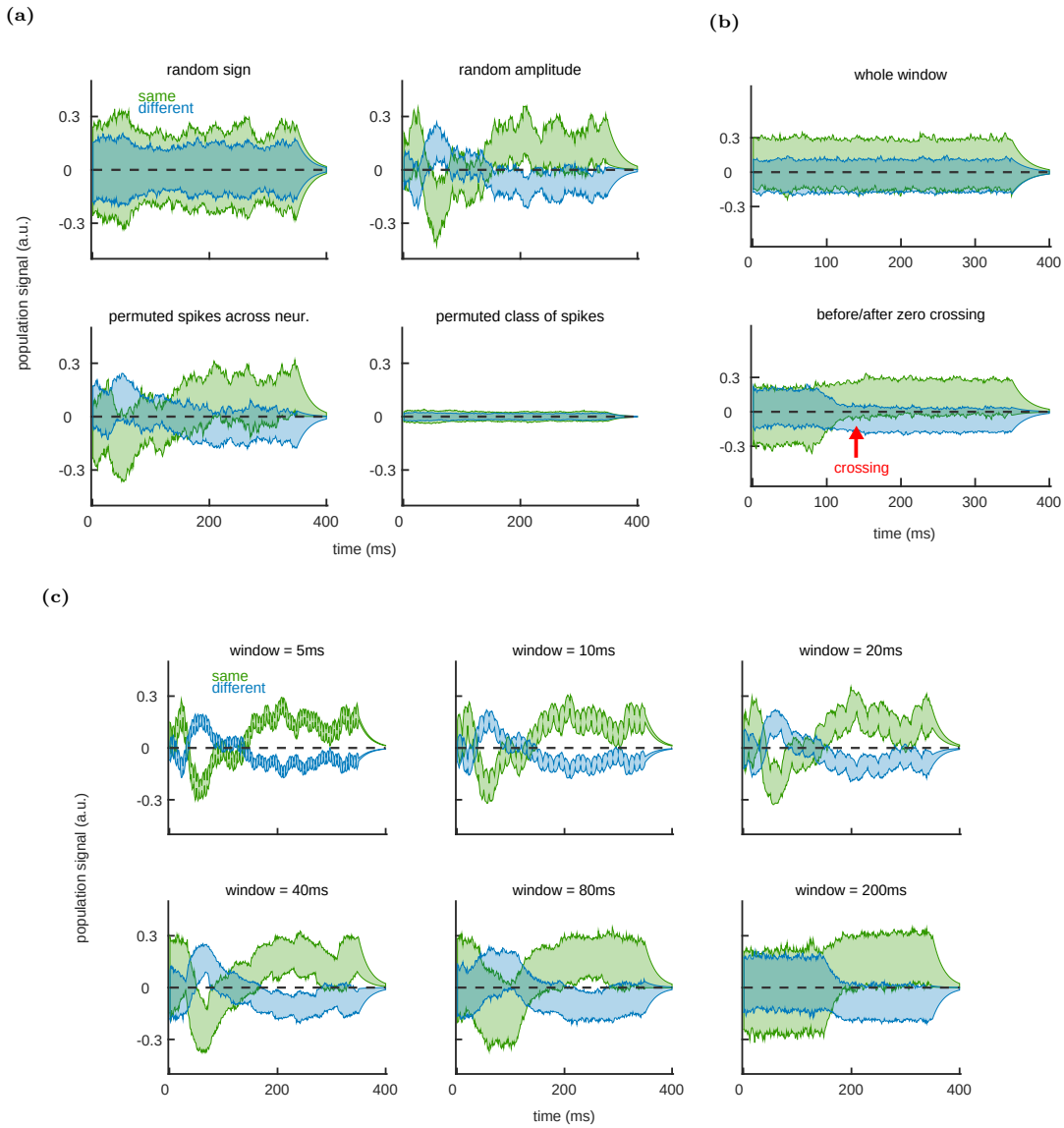

Figure 4. Correct sign of decoding weights is necessary for discrimination. (a) The choice signals with random sign of weights (top left), random amplitude of weights (top right), permuted spikes across neurons (bottom left) and permuted class labels of spike trains (bottom right). We plot the entire distribution of choice signals in conditions "same" (green) and "different" (blue), where "same" and "different" refer to the animal's choice on non-matching stimuli. **(b)** The choice signals with randomly permuted spike timing across the entire time window (top) and in two time windows, before and after the switch in the read-out of the choice (bottom). Red arrow marks the timing of the crossing (140 ms after the stimulus onset). **(c)** The choice signal with temporal jitter, for different length of the jittering window. Parameters for all plots: $\lambda^{-1} = 20\ ms$, $N_{perm} = 1000$.

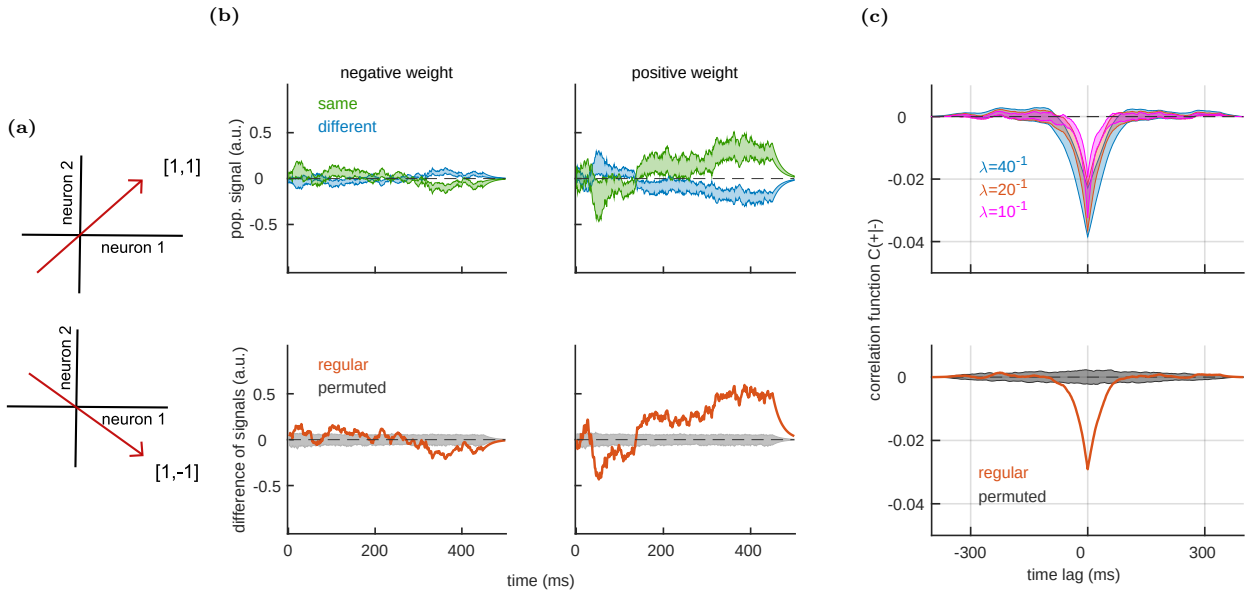

Figure 5. The choice signal of neurons with positive and negative weights is negatively correlated. (a) Schema of the effect of the sign of the weight on the orientation of the separation boundary (red arrow). In a toy model with 2 neurons, the upper schema shows the case where the weights of both neurons are positive, and the lower schema illustrates the case with positive and negative weight. **(b)** Top: the choice signal for neurons with negative (left) and positive weight (right). Bottom: Session-averaged difference of signals for the true model (red) and for models with permutation (gray area). **(c)** Top: Correlation function between the choice signal of *plus* and *minus* subnetworks for different values of the time constant of the convolution. Bottom: Session-averaged correlation function for the regular model (red) and the distribution of results for the model with permutation (gray area).

the same sign of the weight are synchronized and/or correlated on short time scales, this helps the choice signal to accumulate (fig.6c, top). On the contrary, if neurons with the opposite sign fire in alternation, the signal remains close to zero (fig.6c, bottom). Spike-time correlations and synchrony between neurons with the same sign therefore help the formation of the signal. Indeed, we find that neurons in the *plus* subnetwork are more strongly synchronized ($p < 0.001$, permutation test) and correlated than neurons within the *minus* subnetwork (fig.6d).

Besides the sign of the weight, the strength of the weight, $|\tilde{w}_n|$, is an important characteristics of the decoding model, since neurons with strong weight are those that contribute most to the classification model. We separate the population in neurons with strong and weak weights (see methods) and compute the choice signal with each group. We find that the subnetwork with strong weights is the one that carries the choice related signal, while the subnetwork with weak weights performs at chance and that the choice signal from the strong and weak subnetworks are positively correlated (Supplementary Fig.S5). As we compare spiking synchrony and correlations across the two groups, we find that neurons with strong weight have stronger synchrony and correlation than neurons with weak weights ($p = 0.0140$ permutation test, fig.6e). This result is an additional argument in favor of the hypothesis that correlations on short time scales are useful and helpful for decoding of the choice.

Activity of bursty neurons and of neurons in the superficial layer is particularly informative about the choice

Besides correlations and synchrony, another way to help the accumulation of the signal might be bursting of single neurons. Bursting helps the accumulation of the signal for the same reason as correlated spiking of neurons with the same sign of decoding weight. We divide neurons into bursty and non-bursty, based on characteristics of the power spectrum of spike trains. The power spectrum of bursty neurons is characterized by a reduced power for middle range frequencies ([54], see fig.7a, middle right and right are examples of bursty neurons). We capture this effect by computing the normalized area under the power spectrum (see methods), and define

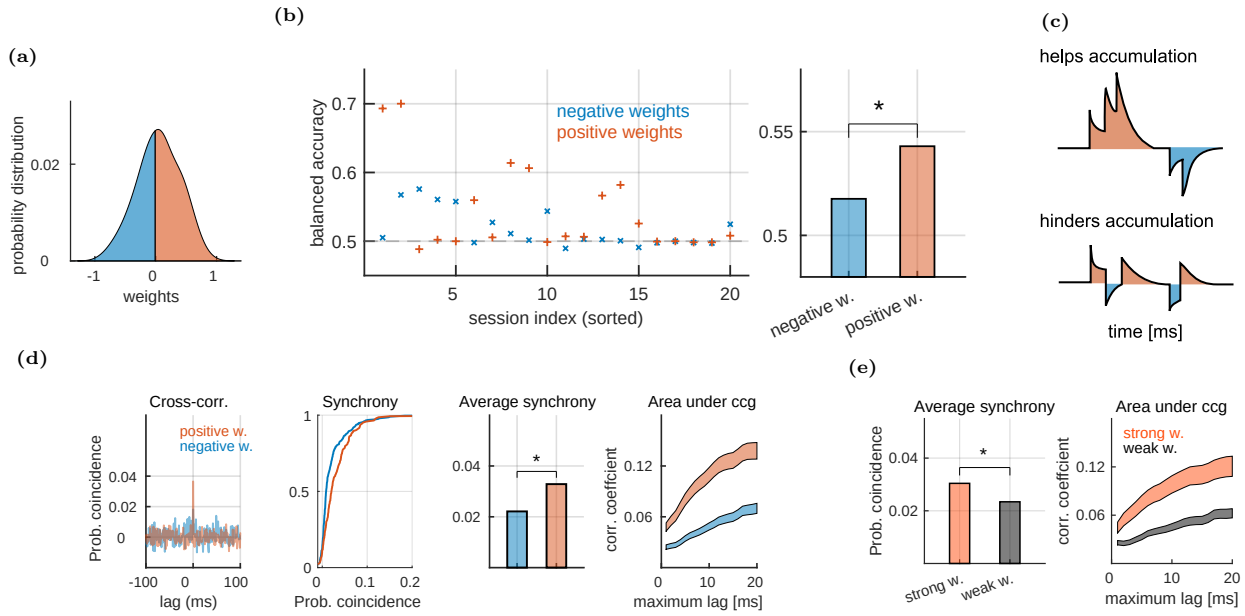

Figure 6. Noise correlations among neurons with positive weights help to build the choice signal. (a) Distribution of weights. Length of the time window is K=500 ms. **(b)** Prediction accuracy of the *stimulus* + *choice* for the *plus* (red) and *minus* subnetwork. Left plot shows results in sessions, ordered in the same way as in fig.2a, while the right plot shows session-averaged results. The difference in accuracy between bursty and non-bursty subnetworks is tested with the t-test on 100 instances of the balanced accuracy with random permutation, without repetition, of the class label (see methods). Parameters for all plots, when applicable: $\lambda^{-1} = 20\ ms$, $N_{perm} = 1000$. **(c)** Schema of spike patterns that either help or hinder the accumulation of the signal. Let's have a pattern of 5 spikes, three spikes of *plus* neurons and 2 spikes of *minus* neurons. The pattern where spikes of neurons with the same sign of the weight are grouped together makes the signal accumulate (top), while alternative firing of *plus* and *minus* neurons hinders the accumulation of the signal (bottom). **(d)** Synchrony and correlation within the group of *plus* (red) and *minus* (blue) neurons. Left: example cross-correlograms (*ccg*) for a pair of neurons with positive weights (red) and negative weights (blue). Middle left: empirical cumulative distribution functions of the *ccg* at zero lag, measuring pair-wise synchrony. The distribution shows all neuronal pairs, collected across recording sessions. Middle right: average synchrony between pairs of neurons with positive and negative weights. Right: Area under the *ccg* (also known as r^{ccg}), for the increasing length of the summation window. **(e)** Same as **(d)** middle right and right, but for pairwise interactions within the group of neurons with strong (orange) and weak (dark gray) weights. The separation in strong and weak weights is done w.r.t. the classification of the *stimulus* + *choice*, while the correlation and synchrony is computed for conditions corresponding to the variable *choice*.

bursty neurons as neurons that have the area under the power spectrum smaller than the lower bound of models with permuted spike timing (fig.7b). As we compute the choice signal of bursty and non-bursty subnetworks, we find that the bursty subnetwork is largely better at predicting the choice than the non-bursty subnetwork (fig.7c). Choice signals of bursty and non-bursty subnetworks are positively correlated (fig.7d).

Similarly to *plus* neurons, bursty neurons are better at predicting the choice because they have more informative weights and due to stronger synchrony and correlations than among non-bursty neurons. Bursty neurons tends to predict the variable *stimulus* + *choice* better than the non-bursty subnetwork ($p < 10^{-12}$, two-tailed t-test; fig.7e). Also, bursty neurons tend to have stronger decoding weight than non-bursty neurons, but only if their decoding weight is positive ($p = 0.012$ for neurons with positive weights, $p = 0.610$ for neurons with negative weight, permutation test; fig.7f). In addition, spiking synchrony and correlation on short time scales are stronger between bursty compared to non-bursty neurons ($p < 0.001$ for synchrony, permutation test; fig.7g).

At last, we divide neurons with respect to their location in three cortical layers. Cortical layers in V1 have been shown to importantly differ in their correlation structure [26], which might have direct implications on the population code. Using a method based on the current source density [23], we distinguish the superficial layer (supragranular, SG), the middle layer (granular, G) and the deep layer of the cortex (infragranular, IG). As we compute the choice signal in each layer, we find that the superficial layer is by far the most important for discriminating the choice (fig.8a). The superficial layer also reflects the crossing of the read-out at around

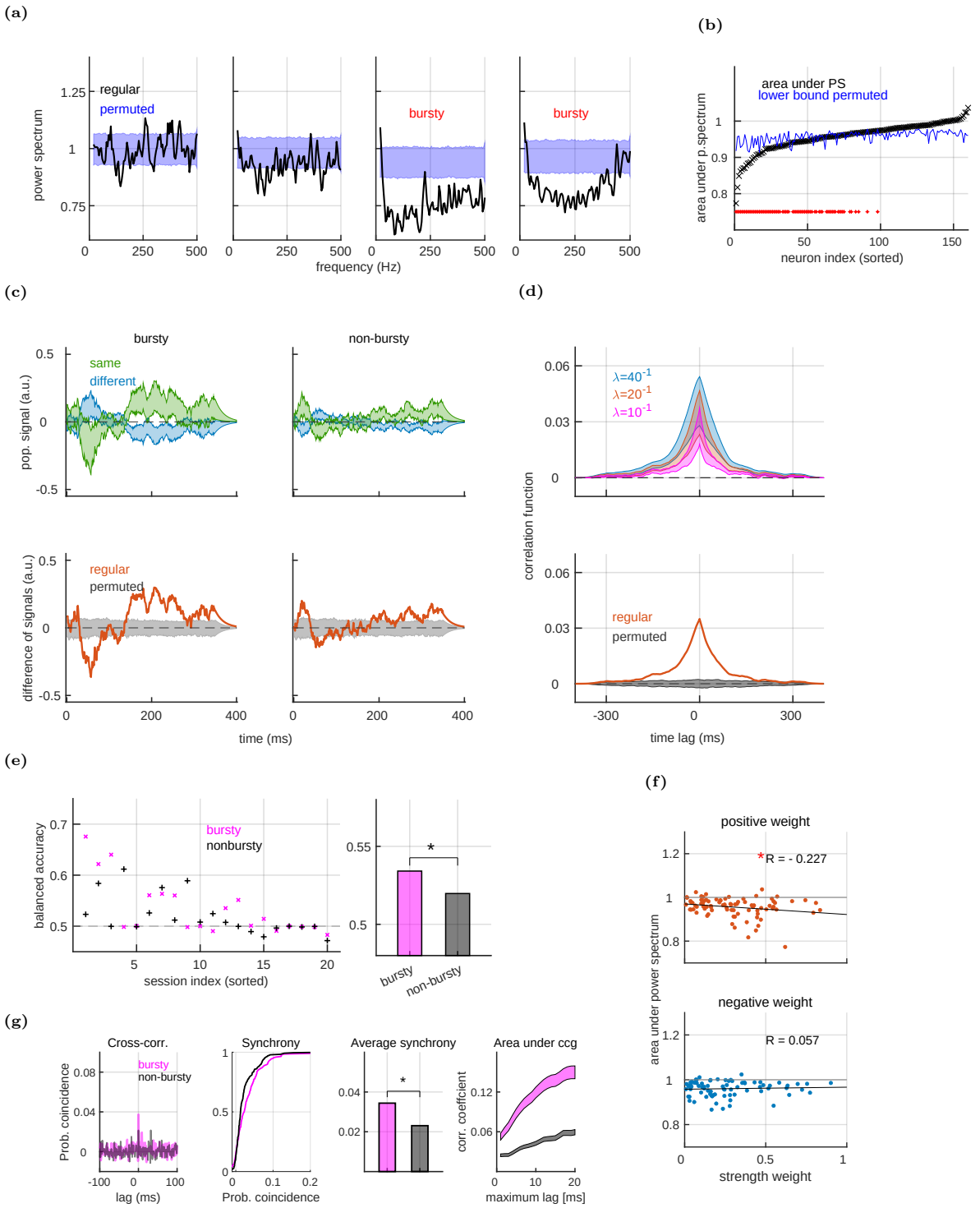

Figure 7. Bursty neurons discriminate better and earlier in the trial. (a) Power spectrum for 4 representative neurons from the same recording session. We show the power spectrum for the true model (black) and the distribution of power spectra for models with randomly permuted spike timing without repetition (blue). Bursty neurons are those with reduced power in lower and medium range frequencies (middle right and right). **(b)** Area under the power spectrum for all neurons, sorted from the lowest to the highest value (black), and the lower bound of the area under the power spectrum for models with permuted spike timing (blue). Red asterisk marks significance (permutation test, $\alpha = 0.05$). **(c)** Top: The choice signal for bursty (left) and non-bursty subnetworks (right) in recording sessions. Bottom: Session-averaged difference of signals for the true model (red) and for models with random permutation (gray area). Parameter: $\lambda^{-1} = 20\ ms$. **(d)** Top: Correlation function between the choice signal of bursty and non-bursty subnetworks. We show the mean $\pm$ SEM for the variability across recording sessions. Bottom: Session-averaged correlation function for the regular model (red) and the distribution of results for the model with permutation (gray area). **(e)** Balanced accuracy on decoding from parallel spike trains with the linear classifier (Support Vector Machine) for bursty (magenta) and non-bursty (black) subnetworks. The left plot shows results in recording sessions, ordered as on fig.2a. Right plot shows session-averaged results. **(f)** Scatter plot of the Area under the power spectrum and the Strength of the weight, for neurons with positive weights (top) and neurons with negative weights (bottom). We used trials corresponding to the variable $stimulus + choice$. R marks the Pearson correlation coefficient. **(g)** Same as in fig.6c, but testing interactions within the group of bursty (magenta) and non-bursty (black) neurons. Parameter for **(a)-(d)** and **(g)**: $N_{perm} = 1000$.

140 ms after the stimulus onset, seen in the overall response. Choice signals are weakly positively correlated across all pairs of layers (fig. 8b), consistently for all tested time constants of the convolution (fig.8b left), and the correlation is significant for short to intermediate time scales (fig.8b right).

Measuring pair-wise synchrony of spike timing with the cross-correlogram (fig.8c, left), we find that the synchrony is stronger in the middle and the superficial layer compared to the deep layer ($p < 0.001$ for both SG/IG and G/IG, permutation test), while the difference between the middle and the superficial layer is not significant ($p = 0.573$, fig.8c, middle plots). The strength of the synchrony in the middle and the superficial layer is therefore comparable, however, the superficial layer clearly shows stronger correlation of spike timing (fig.8c, right), measured as the area under the cross-correlogram. Besides strongest spike-time correlation, the fact that there is excess of bursty neurons and of neurons with positive weights in the superficial layer (fig.8d) explains why the superficial layer has stronger discriminatory capacity compared to other layers.

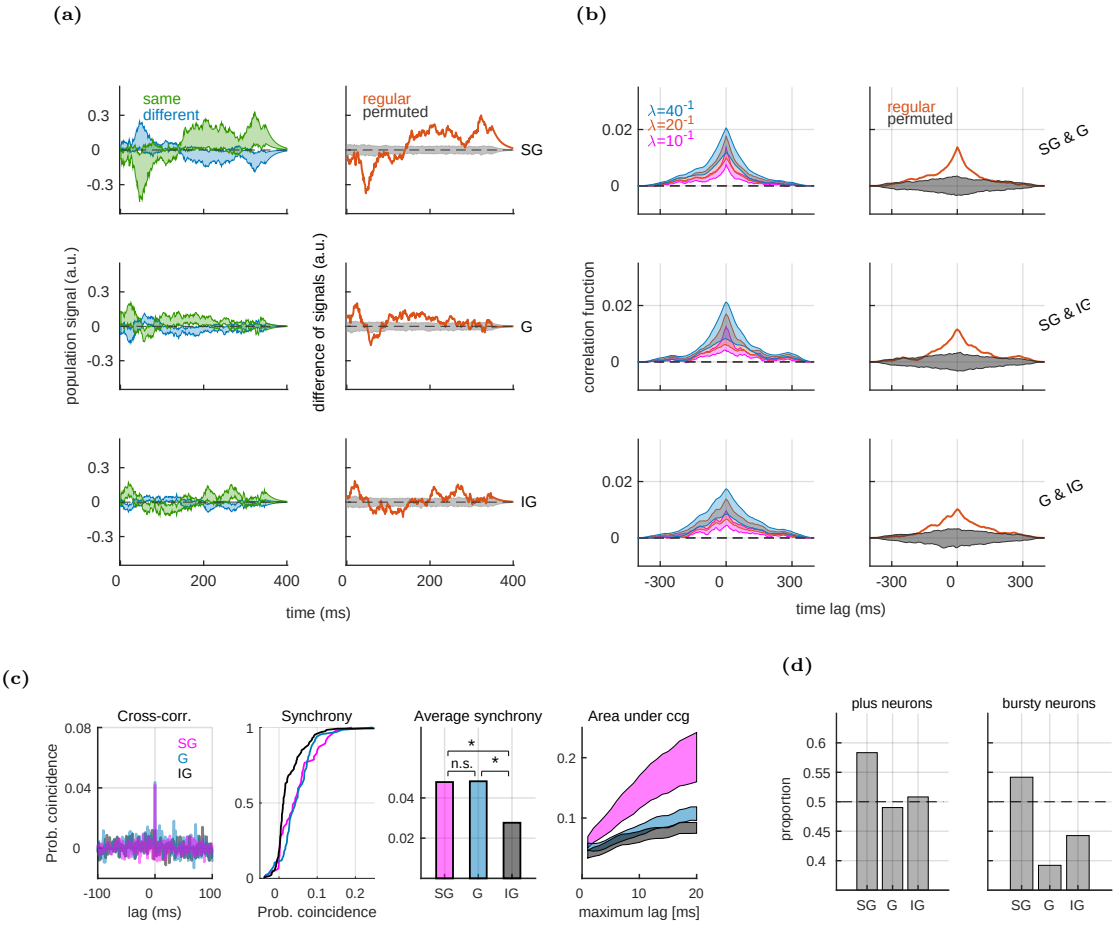

Figure 8. Spiking activity in the superficial layer of the cortex is the most discriminative of the choice behavior. **(a)** Left: The choice signal in the SG (top), G (middle) and IG layer (bottom), showing the mean ± SEM across recording sessions. Right: Session-averaged difference of choice signals for the true model (red) and the model with permuted class labels (gray). Parameters: $\lambda^{-1} = 20\ ms$, $N_{perm} = 1000$. **(b)** Left: Correlation function of choice signals across pairs of layers. We show the mean ± SEM across recording sessions and plot results for three values of the time constant of the convolution. Right: Session-averaged correlation function for the regular model (red) and for models with permutation (gray area). We use the time constant $\lambda^{-1} = 20\ ms$ and $N_{perm} = 1000$ random permutations. **(d)** The proportion of neurons neurons with positive weight(left) and of bursty neurons (right) in cortical layers. The proportion is computed in each layer individually (numbers do not sum to 1 across layers, see methods).

The flow of activity across cortical layers differs between stimulus classes rather than between choices

The measure of cross-correlation evaluates the interaction of spike timing for pairs of neurons, excluding the trial invariant spiking. The cross-correlogram, however, cannot separate signals generated internally within the network from signals that drive the network from the outside, since the extrinsic sources are likely not entirely trial invariant. To investigate more closely the origin of the choice signal in V1, we apply the method that allows to distinguish the flow of activity that is intrinsic to V1 from the external input. The non-linear method of convergent cross-mapping excludes external sources from intrinsically generated interactions by construction [27, 28, 30]. The cross-mapping is a non-linear method of state-space reconstruction [29], and previous work on spiking data *in vitro* has demonstrated that the cross-mapping coefficient (CMC) for a pair of neurons is proportional to the synaptic interaction between the two neurons [31]. Here, the goal of applying the method of cross-mapping is to study interactions that are intrinsic to the V1 network, excluding the influence of the shared input on the interaction, an analysis that is helpful for teasing apart the sources of the choice signal.

First, we were interested in finding out whether the strength of V1-specific interactions differs from one layer to another. To this purpose, we measured the CMC for pairs of neurons within layers, and compared results across layers. We found that the CMC is stronger within the superficial layer compared to other layers ($p = 0.0062$, Kruskal-Wallis test on all layers, $p = 0.0123$ for SG/G and $p = 0.0089$ for SG/IG, multiple comparison test with Tukey-Kramer correction; fig.9a). Note that we used trials from all the three conditions, since the same effect was present in all of them.

Next, we measured the flow of activity across layers by computing the CMCs for pairs of neurons located in different layers (see methods), and build a directed graph of the activity flow across V1 lamina (fig.9b-d). In all conditions, the strength of links is heterogeneous ($p < 10^{-4}$ in "match", $p < 10^{-3}$ in "correct non-match" and in 'incorrect non-match", Kruskal-Wallis test), and the directionality of interactions is highly relevant. In particular, the link from deep to superficial layer is much stronger than the link from superficial to deep layer ($p = 0.0040$, multiple comparison test). With matching stimuli, the strongest link goes from the middle to the superficial layer (fig.9b), while with non-matching stimuli, the strongest link goes from the deep to the superficial layer (fig.9c-d).

The superficial layer seems to act as an activity hub, as it integrates the V1-intrinsic activity across the cortical column. In all conditions, the strongest among incoming and outgoing connections is the incoming link to the superficial layer (fig.9e-g). In condition with matching stimuli, the strongest outgoing link originates in the middle layer, while in both conditions with non-matching stimuli, the strongest outgoing link originates in the deep layer (fig.9e-g).

While directed graph markedly differs across stimuli, it is rather similar between correct and incorrect behavior (compare fig.9b-d). Comparing "match" to "non-match", they visibly differ with respect to the strength of the incoming and outgoing links to the middle layer (fig.9e-g, middle). To better grasp this difference, we evaluated the total flow of activity (see methods). We find that the flow at superficial and middle layer is stronger in "match" compared to "correct non-match" ($p < 10^{-7}$, Kruskall-Wallis including all layers, $p = 0.0119$ in SG, $p = 0.001$ in G; fig.9h), while other comparisons are not significant (Supplementary Fig.S6). Matching stimuli strongly activate the path between the middle and the superficial layer, less strongly activated by non-matching stimuli.

All together, we find the strongest flow of activity within the superficial layer, and show that the difference in stimuli, rather than the difference in choice, is reflected in V1-intrinsic flow of activity across layers. This excludes the possibility that the choice signal stems from interactions between lamina, and they also exclude the bottom-up origin of the choice signal. If V1 received markedly different bottom-up drive relative to the choice, this difference would be inherited by the V1-intrinsic activity flow, contrary to the evidence just exposed. If the choice is not reflected in V1-intrinsic flow across layers, it therefore has to either be generated within the superficial layer, or come from a top-down source.

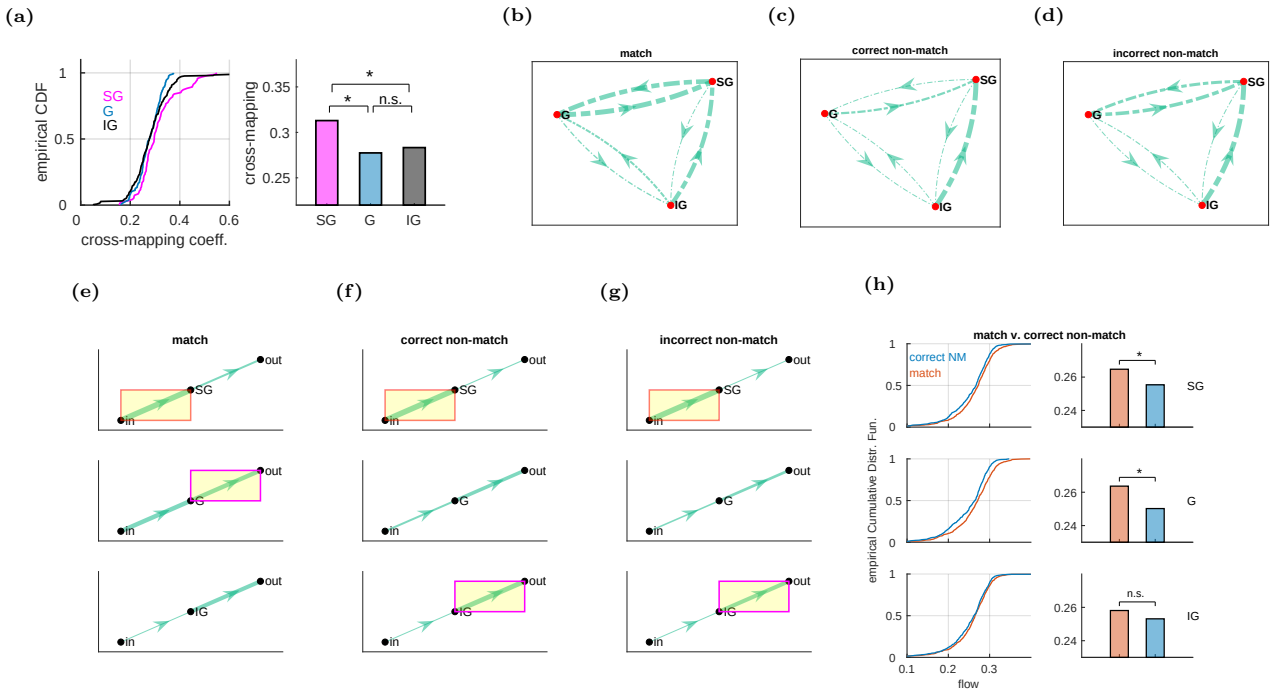

Figure 9. Flow of activity across cortical lamina (a) Cross mapping within cortical layers. Left: Empirical cumulative distribution function of the distribution of cross-mapping coefficients for pairs of neurons within the SG (magenta), G (blue) and IG (gray) layers. Cross-mapping coefficients are computed in single trials, averaged across trials, and collected across neurons from all recording sessions. Right: Average of distributions on the left. The cross-mapping coefficient evaluates the strength and directionality of pair-wise interaction between neurons. The cross-mapping coefficient ranges between 0 and 1, with 0 indicating no flow of information between the two neurons, and 1 indicating that the neuron 2 is driven exclusively by neuron 1 and does not receive any other inputs, except the shared input. **(b-d)** Directed graphs where the thickness of links correspond to the normalized strength of the average cross-mapping coefficient. We show the graph in condition "match", "correct non-match" and "incorrect non-match". **(e-g)** The normalized strength of incoming and outgoing connections at the SG (top), G (middle) and IG layer (bottom). Yellow rectangles show the strongest link for the incoming connections (red border) and for the outgoing connections (magenta border). We show results in conditions "match", "correct non-match" and "incorrect non-match" . **(h)** The flow of activity at each layer, comparing conditions "match" (red) and "correct non-match" (blue). Left plots show the empirical Cumulative Distribution Function, while bar plots on the right are averages of distributions on the left. For all plots, we used the following parameters: maximal delay of 50 ms, 51 nearest neighbors, duration of the time window K=350 ms, starting at 140 ms after the stimulus onset (after the zero crossing of the difference of signals); spike trains in single trials were smoothed with a Gaussian filter with variance of 4 ms^{-2}.

Discussion

Previous work has suggested that the structure of noise correlations in V1 is primarily determined by the feedback from downstream brain areas [32]. In line with these results, we showed that the spiking synchrony and noise correlations on short time scale are stronger between neurons that are also the carriers of the choice signal. We have shown stronger spike-timing noise correlations among neurons with strong weights, among bursty neurons, and among neurons in the superficial cortical layers. Stronger correlations between neurons that are decisive for discrimination suggest that these neurons are driven by common input [33], most likely from a top-down source [32].

Further on, we demonstrated how noise correlations between neurons with similar decoding selectivity might be used by the brain to build the choice signal that is strong enough to be propagated downstream. Noise correlations between neurons with similar selectivity are known to decrease the quantity of information in a population code [34], and they do so also in the present dataset [22]. Nevertheless, the informativeness of the population code is not the only important property of neural responses, it is also the reliability of transmission that has to be ensured if a spiking signal is to reach the neural structures downstream [35]. The spiking signal is, generally speaking, a temporally sparse signal, and correlations among units with similar selectivity might be instrumental to make the signal rise from the noisy background and influence the activity of the read-out

neuron. Note that in the present dataset, there is no difference in the firing rate between neurons with positive and negative weights, and if the spiking activity of neurons with the same sign of the weight was entirely uncorrelated, the read-out would result in a noisy oscillation around zero, precluding the formation of any signal. Present results therefore suggest a simple but useful mechanism that leverages the structure of spike-timing correlations on short time scales to build the choice signal and help discrimination. Besides synchrony and correlations, bursting of single neurons is another dynamical mechanisms that can increase the reliability of transmission to downstream areas, by counteracting the synaptic failure [36], or by driving synchronization in the Gamma frequency, as suggested by a recent study [37].

The feedforward vs. feedback sources of choice information in the sensory areas is a matter of ongoing discussion [38]. It has been proposed that the choice signal in the cortical networks might be conveyed through feed-forward projections [14], feed-back projections [9, 33, 3], or be due to local network dynamics (see [39, 40] for modeling studies). One piece of evidence from the present analysis that allows to discuss about the sources of the choice signal is the switch in decoded choice at about 140 ms after the stimulus onset, which coincides with the qualitative change of the population response from the strong transient at the stimulus onset to tonic firing thereafter. At the stimulus onset, the neural activity is expected to be mainly driven by the feed-forward input, while feed-back and local projections are expected to have have stronger effect (respectively to the feed-forward ones) during the second half of the trial [41]. If the representation in the context of *choice* would be only due to the feed-forward drive throughout the trial, we would expect that it remains consistent throughout the trial. Also, we would expect to decode the *choice* already in the middle layer, which is not supported by present results. We propose that the change in decoded choice is caused by the feed-back taking over the feed-forward drive, where the switch in decoded choice would be due to incongruent signaling between the bottom-up and the top-down pathways.

Incongruent top-down and bottom-up signals in the context of *choice* would suggest congruent signaling of top-down and bottom-up sources in the context of $stimulus + choice$. In the latter context, the behavior of the animal is correct for both choices, which would imply consistent representation of the choice throughout the trial. In our previous work, we have shown that this is indeed the case [23]. Another result that suggests the existence of both bottom-up and top-down sources is related to the spike timing. Permuting the spike timing across the entire time window has destroyed the discriminability of the choice signal, while limiting the window of permutation to the time before and after the switch in decoded choice, the discriminability is partially preserved. This supports the hypothesis of different origins of the signal before and after the switch in decoded choice. Precise timing of the feedback in visual areas has been shown to be crucial for perception [42] and computation in recurrent networks [43], and our current results point out it is also important for signaling of the behavioral choice.

The third piece of evidence in favor of the importance of feed-back projections for the choice signal is the condition-specific divergence of the choice signals for neurons with positive and negative weights towards the end of the trial. This effect can be explained by a top-down input that selectively drives neurons with positive weights for the decision "same" and neurons with negative weights for the decision "different". It has been shown that such a a context-dependent computation is feasible in a recurrent network [44, 45], and it could be implemented in the prefrontal cortex. Precise long range microcircuit-to-microcircuit connections between the prefrontal cortex and V1, recently put in evidence [46], could then feed back the choice signal to V1. Current results showed the importance of bursting neurons for the signaling of the behavioral choice. It appears likely that bursty neurons described here correspond to the recently reported class of bursting excitatory neurons in the V1 of the macaque, suggested to have an important role in encoding and transmission of cortical signals [37]. Alternatively, a top-down choice signal could drive bursty neurons in the superficial layer through dendritic feedback inhibition, as suggested by a theoretical study [47].

Finally, we studied V1-intrinsic flow of activity across lamina, and were able to exclude the bottom-up and intra-laminar origin of the choice signal. Interestingly, the flow of activity differs more between matching and non-matching stimuli (variable $stimulus + choice$) than between correct and incorrect behavior (variable *choice*), a rather clear indication that the flow of activity across layers of V1 primarily reflects the stimulus and

not the choice or the correctness of the choice. This leaves two possibilities, either the choice is computed within the superficial layer, or it is computed downstream and is merely reflected in the activity of the superficial layer. While answering this question is beyond the scope of this study, we speculate that the local origin of the choice within the superficial layer of V1 is unlikely. If the choice was computed in the superficial layer, we would expect to observe an important difference in the outgoing flow of activity from the superficial to middle and deep layer. This is contrary to evidence, since the outgoing flow of the superficial layer is very similar between conditions correct and incorrect non-match. However, since the outgoing flow from the superficial layer of V1 downstream is not part of the dataset, we cannot entirely exclude the active role of the superficial layer in the generation of the choice signal.

Recent study in monkey middle temporal (MT) area has pointed out that the stimulus and the choice related information lie in partially distinct subspaces, with a substantial component of the choice information being in a subspace orthogonal to the stimulus [49]. This is similar to our findings, since we found that the population vectors corresponding to the $stimulus + choice$ and $choice$ are only weakly similar, and are in fact closer to orthogonality than alignment. In the present data, population vectors deviate for about 10 deg from orthogonality, and therefore deviate for 80 deg from alignment. In [49], authors suggest that the benefit of partial misalignment of the stimulus and the choice might be the co-existence of two representations and a non-corrupting feedback. Our study, on the other hand, points out the benefit or the partial alignment of the information on the choice and the stimuli. In fact, the partial alignment allows the representation of the choice in V1, and representing the choice in multiple sensory areas might serve the purpose of reverberating and amplifying the choice-related signal within the cortex [32]. Such reverberation could implement a form of working memory by sending the information about a relevant variable across cortical areas, and as such maintain the relevant information in the cortical circuitry.

The question remains, why is the generalized decoding scheme necessary for discrimination of the choice? The main reason might be additional sources of noise that are associated with incorrect behavior. Likely, sensory noise is stronger in incorrect trials [50]. Moreover, the behavioral task consisted in the comparison of the target and the test stimulus, and as such required the information on the first stimulus to be stored in the cortex. Transmission of the signal from V1 downstream and then back to V1 is prone to failure and is likely compromised in incorrect trials [51]. Recent analysis on the overlapping dataset has indeed demonstrated that the spiking synchrony and the information transfer between V1 and V4 is stronger in correct compared to incorrect trials, suggesting that the transmission of the signal from V1 downstream is compromised in incorrect trials [52, 53].

Methods

Methods availability

Code and dataset will be freely available in a pubic GitHub repository.

Experimental model and subject details

Ethics statement All experiments were conducted in accordance with protocols approved by The Animal Welfare Committee (AWC) and the Institutional Animal Care and Use Committee (IACUC) for McGovern Medical School at The University of Texas Health Science Center at Houston (UTHealth), and met or exceeded the standards proposed by the National Institutes of Health's Guide for the Care and Use of Laboratory Animals.

Method details

Animal subjects and the Experimental setup Two male rhesus macaques (Macaca mulatta; M1, 7 years old, 15kg; M2, 11 years old, 13kg) were used in this study. Animals were trained on a delayed match-to-sample task with visual stimuli. One trial comprised the fixation within the fixation area, automatically triggering the

appearance of the target and the test stimuli. The target and the test stimuli were shown for 300 ms each, with a variable delay period in between (800 to 1000 ms). The identity of stimuli, complex naturalistic images depicting an outdoor scene, changed from one recording session to another. Within the same session, the target and the test stimuli could be either the same (condition "match") or the test stimulu was rotated w.r.t. the target (condition "non-match"). The difference in orientation was such as to keep the performance stable at around 70 percent correct on non-matching stimuli, and ranged between 3 and 10 degrees. Animal subjects were trained to communicate their choice by holding the bar for the choice "different" and releasing the bar for "same". For the trial to be valid, the response had to be communicated between 200 and 1200 ms after the offset of the test stimulus.

Recordings Multi-unit signal was measured with laminar arrays comprising 16 channels with 100 $\mu-$ meter spacing between adjacent contacts. Electrodes captured neuronal activity across the cortical depth, from superficial to deep layers. The multi-unit signal was spike sorted and we analyzed all cells that responded to the stimulus with at least a 4-fold increase of the firing rate with respect to the baseline. The data was collected in 20 recording sessions, which gave 160 neurons. We analyzed 3 conditions, "correct match" (CM), "correct non-match" (CNM) and "incorrect non-match" (INM), where "correct/incorrect" refers to the behavioral performance and "match/non-match" refers to the stimulus class. In condition "incorrect match", there were not enough trials to perform the analysis. Below, table 2 reports the summary statistics on the number of trials in each condition.

Condition	Nb. trials min.	Nb. trials max.	Nb trials average
CM	51	290	118
CNM	42	230	107
INM	21	98	39

Table 2. The minimal, maximal, and average number of trials across recording sessions, for each condition.

Quantification and statistical analysis

The analysis was done with Matlab, Mathworks. The spike train of a single neuron n in trial j is defined as a binary vector of zeros and ones,

$$o_{n,j}(t_k) = \begin{cases} 1, & \text{if neuron } n \text{ in trial } j \text{ spikes during the } k\text{-th millisecond} \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

where $n = 1, ..., N$ is the neural index, $k = 1, ..., K$, is the time index with step of 1 millisecond, and $j = 1, .., J_1, J_1 + 1, ..., J_2, J_2 + 1, ...J$ is the trial index. Trials were collected in conditions CM ($j = 1, ..., J_1$), CNM ($j = J_1 + 1, ..., J_2$) and INM ($j = J_2 + 1, ..., J$).

The coefficient of variation for the neuron n in trial j is defined as follows,

$$CV2_{n,j} = 2\frac{|ISI_{i+1} - ISI_i|}{ISI_{i+1} + ISI_i}$$

where ISI is the inter-spike interval with index $i = 1, ..., N_{int}$. We report trial-averaged results, distinguishing trials from conditions CNM (decision "same") and INM (decision "different").

$$\begin{aligned} CV2_n^{\text{different}} &= \frac{1}{J_2 - J_1} \sum_{j=J_1+1}^{J_2} CV2_{n,j} \\ CV2_n^{\text{same}} &= \frac{1}{J - J_2} \sum_{j=J_2+1}^{J} CV2_{n,j} \end{aligned} \tag{2}$$

Learning of structural features from parallel spike counts

We computed spike counts in the time window of $[0, K]$ ms with respect to the onset of the stimulus (target or test), $s_{n,j} = \sum_{t_k=1}^{K} o_{n,j}(t_k)$. Spike counts were z-scored:

$$\tilde{s}_{n,j} = \frac{s_{n,j} - \langle s_{n,j} \rangle_j}{\sqrt{\mathrm{Var}_j\ (s_{n,j})}}, \tag{3}$$

where $\langle s_{n,j} \rangle_j$ and $\mathrm{Var}_j\ (s_{n,j})$ are the mean and the variance across trials from all conditions.

The linear classification model (linear Support Vector Machine, SVM) is computed on conditions CM and CNM, which differ in both the stimulus class and the choice. The input to the linear SVM are z-scored spike counts of N neurons that have beer recorded in parallel. From feature weights of the linear SVM, we calculate the population vector that contains a decoding weight for each neuron. These weights extract the structure of spike counts. We then use the population vector to compute the choice signal in conditions CNM and INM, which only differ in the choice. The choice signal is computed from parallel spike trains, without averaging across time, and in single trials.

The training set comprises all trials from condition CM and half of the trials from condition CNM (trials with index $j = 1, ..., J_1, J_1 + 1, ..., \frac{J_2 - J_1}{2}$). The reconstruction of the choice signal is computed on a hold-out set, utilizing the remaining half of the trials from condition CNM and all trials from condition INM (trials with index $j = \frac{J_2 - J_1}{2} + 1, ..., J_2, J_2 + 1, ..., J$). The split of trials in training and reconstruction set in condition CNM is cross-validated with Monte-Carlo method, using N_{cv} random splits. The split of trials in training and reconstruction set is non-overlapping, such that no trials that have been used for training appear in the reconstruction set. All reported results are averaged across cross-validations.

Estimation of the population vector In the N-dimensional space of inputs, one sample is the vector of z-scored spike counts of N simultaneously recorded neurons in trial j, $\mathbf{s}_j = [\tilde{s}_{1,j}, \tilde{s}_{2,j}, ..., \tilde{s}_{N,j}]^T$. Linear SVM searches for an $N-1$-dimensional plane (a hyperplane) that optimally separates points in conditions CNM and CM. The hyperplane is defined as follows,

$$H_0 : \mathbf{w}^T \mathbf{s}_j + b = w_1 s_{1,j} + w_2 s_{2,j} + ... + w_N s_{N,j} + b = 0 \tag{4}$$

where $\mathbf{w}$ is the vector of feature weights and b is the offset of the hyperplane from the origin. On each side of H_0, we can define a hyperplane that verifies the following:

$$\begin{aligned} H_1 : \mathbf{w}^T \mathbf{s}_j + b = -1 \\ H_2 : \mathbf{w}^T \mathbf{s}_j + b = 1 \end{aligned} \tag{5}$$

If the problem is linearly separable, all training samples verify the following the inequality,

$$\begin{aligned} \mathbf{w}^T \mathbf{s}_j + b \leq -1 \text{ if } y_j = -1 \\ \mathbf{w}^T \mathbf{s}_j + b \geq 1 \text{ if } y_j = 1 \end{aligned} \tag{6}$$

where $y_j \in \{-1, 1\}$ is the class label ($y_j = -1$ in condition CNM and $y_j = 1$ in condition CM). Training the linear SVM consists in maximizing the number of correctly classified samples and, at the same time, minimizing the distance between H_1 and H_2, which can be expressed with the Lagrangian.

$$L_p = \frac{1}{2} \mathbf{w}^T \mathbf{w} - \sum_{j=1}^{\frac{J_2 - J_1}{2}} \lambda_j [y_j (\mathbf{w}^T \mathbf{s}_j + b) - 1] \tag{7}$$

The first term on the right hand side ensures the maximal distance between hyperplanes H_1 and H_2, and the second term on the right ensures correct classification. As the derivative of the Lagrangian with respect to $\mathbf{w}$

is set to zero, we get the expression for the vector of weights.

$$\mathbf{w} = \sum_{j=1}^{\frac{J_2-J_1}{2}} \lambda_j y_j \mathbf{s}_j \tag{8}$$

Since $\lambda_j \neq 0$ only for trials that define the margin (points that lie on $H1$ or on $H2$), the weight vector only depends on support vectors,

$$\mathbf{w} = \sum_{q=1}^{Q} \lambda_q y_q \mathbf{v}_q \tag{9}$$

where $q = 1, 2, ..., Q$ are the support vectors, with $Q < \frac{J_2-J_1}{2}$. The weight vector is normalized, $\tilde{\mathbf{w}} = \mathbf{w}/||\mathbf{w}||$, with $||\mathbf{w}|| = (w_1^2 + ... + w_N^2)^{\frac{1}{2}}$. We refer to the vector $\tilde{\mathbf{w}} = [\tilde{w}_1, \tilde{w}_2, ..., \tilde{w}_n, ..., \tilde{w}_N]^T$ as the population vector and to the $n-$th entry of $\tilde{\mathbf{w}}$, $\tilde{w}_n$, as the decoding weight of the neuron n.

Representation of the choice signal as a read-out of parallel spike trains

In the following, we describe the reconstruction part. For further information about the method, see [23].

Consider the vector of spike trains of N simultaneously recorded neurons.

$$\mathbf{o}_j(t_k) = [o_{1,j}(t_k), o_{2,j}(t_k)..., o_{N,j}(t_k)]^T \tag{10}$$

The choice signal is a projection of the vector of spike trains on the vector of weights,

$$x_j(t_k) = F(\tilde{\mathbf{w}}^T \mathbf{o}_j(t_k)) - \langle F(\tilde{\mathbf{w}}^T \mathbf{o}_j(t_k)) \rangle_j \tag{11}$$

where $\langle F(\tilde{\mathbf{w}}^T \mathbf{o}_j(t_k)) \rangle_j$ is the trial average, utilizing all trials of the reconstruction set. By subtracting the trial average, we compute the deviation of the signal from the mean. We argue that the deviation of the signal from its mean, rather than the absolute value of the signal, might be biologically relevant. As the transfer function, $F(y(t_k))$, we use a convolution with an exponential kernel,

$$F(y) = \sum_{\tau \in T} y(t_k - \tau) u(\tau) \tag{12}$$

with $u(\tau) = exp(-\lambda\tau),\ \tau \in T$, with support $T = \{-100, ..., 100\}$ ms. Convolution with an exponential kernel models the causal effect of the presynaptic spike on the neural membrane of the read-out neuron. Notice that $x_j(t_k)$ is a time-resolved, low-dimensional representation of parallel spike trains in single trials.

To test the discriminability of conditions CNM and INM, we average the choice signal across trials, distinguishing conditions CNM (decision "different") and INM (decision "same"),

$$\begin{aligned} x^{\text{same}}(t_k) &= \frac{2}{J_2 - J_1} \sum_{j=J'}^{J_2} x_j(t_k) \\ x^{\text{different}}(t_k) &= \frac{1}{J - J_2} \sum_{j=J_2+1}^{J} x_j(t_k) \end{aligned} \tag{13}$$

where $J' = \lceil \frac{1}{2}(J_1 + J_2) \rceil + 1$, and $\lceil z \rceil$ is the ceiling function . The significance of the difference between choice signals for the decision "same" and "different" is evaluated with the permutation test. We compute the difference of the choice signals in every recording session,

$$\Delta x(t_k) = x^{\text{same}}(t_k) - x^{\text{different}}(t_k) \tag{14}$$

and average across sessions. We then rank $\Delta x(t_k)$ among $\Delta x_p^{perm}(t_k)$, where the latter has been computed with random weights, and the class labels in the validation set have been randomly permuted. Random weights were

drawn from the uniform distribution with the same range as the regular weights. The permutation procedure is repeated N_{nperm}-times and gives a distribution of results for each time step. When the result of the true model, $\Delta x(t_k)$, appears outside of the distribution of results of the null model, $\Delta x_p^{perm}(t_k)$ for $p = 1, 2, ..., N_{perm}$, we consider that signals $x^{\text{same}}(t_k)$ and $x^{\text{diff}}(t_k)$ have been successfully discriminated.

Criteria for division into subpopulations

Sign of the weight We separate the population with respect to the sign of the decoding weight, distinguishing neurons with positive weight ($\tilde{w}_n > 0$, *plus* neurons) and negative weight ($\tilde{w}_n < 0$, *minus* neurons).

Strength of the weight We distinguish neurons with strong and weak weights by ranking the absolute value of the weight, $|\tilde{w}_n|$, among the distribution of weights of models with permuted class labels. The strength of the weight of a single neuron n is a scalar, and the same result for models with permuted class labels is a distribution of N_{perm} values, where N_{perm} is the number of random permutations of class labels. If the strength of the weight of the neuron n is ranked within the first 25 % of weights from the null model, we assume that the neuron n has strong weight, and assume it has weak weight otherwise.

Burstiness We distinguish bursty and non-bursty neurons with a simple method based on the power spectrum of spike trains. It has been shown that bursty neurons have decreased power spectrum in low and middle frequency ranges [54]. We compute the power spectrum of spike trains for every single neuron, using multiplication of the spike train with Slepian tapers to increase the reliability of the estimation [55]. We use 5 Slepian tapers and the time window of $K = 400$ ms. Power spectra are normalized with neuron's firing rate. As a reference, homogeneous Poisson process has a flat power spectrum of 1, and typically, the power spectrum of a bursty neuron is lower than 1 for low and middle frequency ranges. We compute the normalized sum of the power spectrum for frequencies between 10 and 200 Hz. Frequencies below 10 Hz are discarded, since the power spectrum at those frequencies cannot be accurately estimated due to the short time window K. Significance of the sum under the power spectrum is estimated with the permutation test. The null model is computed by randomly permuting, without repetition, the time index of spike trains $N_{perm}-$ times. We rank the sum under the power spectrum of the regular model among the same results for models with permuted time index. We assume that the neuron n is bursty if the normalized sum under its power spectrum is significantly below the same result of models with permuted spike timing ($\alpha = 0.05$). If the neuron does not fulfill this criterion, it is assumed to be non-bursty.

Cortical layers We distinguish three cortical layers, the superficial (supragranular, SG), the middle (granular, G) and the deep layer (infragranular, IG, [26]). The method for determining cortical layers utilizes the covariance matrix of the current source density (see methods and fig.6 in [23]).

The choice signal in subpopulations

Sign-specific choice signal is computed by removing the information from the spike train of neurons with the opposite sign. We remove the information by randomly permuting class labels "same" and "different" in the reconstruction step. When the class label in the reconstruction step is permuted, we get a random association between the weight and the class label of the spike train, resulting in a signal that is close to zero at all times. As an example, the choice signal of *plus* neurons is computed with the spike train $\mathbf{o}_{j,p}^{perm,+}(t_k)$, where the label of the spike train is correct for *plus* neurons, and random (i.e., correct or incorrect with equal probability) for *minus* neurons.

$$x_{j,p}^{+}(t_k) = F(\tilde{\mathbf{w}}^T \mathbf{o}_j^{perm,+}(t_k)) - \langle F(\tilde{\mathbf{w}}^T \mathbf{o}_j^{perm,+}(t_k)) \rangle_j \tag{15}$$

Similarly, the signal of *minus* neurons is computed by utilizing the spike train $\mathbf{o}_{j,p}^{perm,-}(t_k)$, where the label of the spike train is correct for *minus* neurons, and random for *plus* neurons.

$$x_{j,p}^{-}(t_k) = F(\tilde{\mathbf{w}}^T \mathbf{o}_j^{perm,-}(t_k)) - \langle F(\tilde{\mathbf{w}}^T \mathbf{o}_j^{perm,-}(t_k))\rangle_j \tag{16}$$

Random permutation is repeated $N_{perm}-$times, with $p = 1, 2, ..., N_{perm}$ random permutations, without repetition, of the order of trials. As before, we average each of the signals across trials, distinguishing conditions "same" and "different".

$$\begin{aligned} x_p^{+,\text{same}}(t_k) &= \frac{2}{J_2 - J_1} \sum_{j=J'}^{J_2} x_{j,p}^{+}(t_k) \\ x_p^{+,\text{different}}(t_k) &= \frac{1}{J - J_2} \sum_{j=J_2+1}^{J} x_{j,p}^{+}(t_k) \end{aligned} \tag{17}$$

Same follows for *minus* neurons. The signal is then averaged across permutations, getting $x^{+,\text{same}}(t_k)$ and $x^{+,\text{different}}(t_k)$ as the signal for *plus* subnetwork and $x^{-,\text{same}}(t_k)$ and $x^{-,\text{different}}(t_k)$ as the signal for the *minus* subnetwork. The significance is evaluated with the permutation test. The test statistic is the sign-specific difference of signals in conditions "same" and "different".

$$\begin{aligned} \Delta x^{+}(t_k) &= x^{+,\text{same}}(t_k) - x^{+,\text{different}}(t_k) \\ \Delta x^{-}(t_k) &= x^{-,\text{same}}(t_k) - x^{-,\text{different}}(t_k) \end{aligned} \tag{18}$$

The null model is computed with the random permutation of class labels 1) when training the classification model and 2) in the reconstruction step. In addition, we use a random assignment to the class of *plus* and *minus* neurons by randomly permuting neural indexes.

The same methods is used to compute the choice signal for subnetworks of neurons with strong and weak weights, for bursty and non-bursty subnetworks and in cortical layers.

Correlation function between the choice signals

The correlation function between the choice signals of *plus* and *minus* subnetworks in trial j and for the permutation instance p is defined as follows:

$$R_{j,p}^{+-}(\tau) = \begin{cases} \sum_{k=0}^{K-\tau-1} x_{j,p}^{+}(t_k) x_{j,p}^{-}(t_k + \tau) & \text{for } \tau \geq 0 \\ R_j^{-+}(-\tau) & \text{for } \tau < 0 \end{cases} \tag{19}$$

with time lag $\tau = 1, 2, ..., 2K - 1$. The correlation function is normalized with autocorrelation functions at zero time lag,

$$\tilde{R}_{j,p}^{+-}(\tau) = \frac{R_{j,p}^{+-}(\tau)}{R_{j,p}^{--}(0) R_{j,p}^{++}(0)} \tag{20}$$

where R^{++} (R^{--}) is the autocorrelation function for *plus* (*minus*) neurons.

$$R_{j,p}^{++}(\tau) = \begin{cases} \sum_{k=0}^{K-\tau-1} x_{j,p}^{+}(t_k) x_{j,p}^{+}(t_k + \tau), & \text{for } \tau \geq 0 \\ R_{j,p}^{++}(-\tau) & \text{for } \tau < 0 \end{cases} \tag{21}$$

The correlation function is computed in single trials and then averaged across trials and across permutations. Since there was no difference in the correlation across conditions, we used all trials from the reconstruction set (conditions CNM and INM, trials with index $j = J_1 + 1, ..., J$).

$$\tilde{R}^{+-}(\tau) = \frac{1}{N_{perm}(J - J_1)} \sum_{p=1}^{N_{perm}} \sum_{j=J_1+1}^{J} \tilde{R}^{+-}_{j,p}(\tau) \tag{22}$$

The significance of the correlation function is estimated with the permutation test. We compute the choice signal with random weights and random class label of spike trains. In addition, we use a random assignment to the group of *plus* and *minus* neurons as we compute the correlation function. The same method is used to compute the correlation function between subnetworks with strong and weak weights, and between bursty and non-bursty subnetworks.

Correlation function of choice signals in cortical layers

Similarly, we compute the cross-correlation of choice signals between pairs of cortical layers. The correlation function is computed for the choice signals from two cortical layers,

$$R^{c_1 c_2}_{j,p}(\tau) = \begin{cases} \sum_{k=0}^{K-\tau-1} x^{c_1}_{j,p}(t_k) x^{c_2}_{j,p}(t_k + \tau), & \text{for } \tau \geq 0 \\ \\ R^{c_2 c_1}_{j,p}(-\tau), & \text{for } \tau < 0 \end{cases} \tag{23}$$

with $(c_1, c_2) \in \{(SG, G), (SG, IG), (G, IG)\}$. The rest of the procedure is the same as for *plus* and *minus* neurons. The significance of results is evaluated with the permutation test, where signals are computed with random weights and random class of the spike train. During the computation of the correlation function, the null model uses random assignment to one of the three cortical layers.

Comparison of population vectors across informational contexts of *stimulus+choice* and *choice*

Similarity of population vectors We compare the population vectors in the context of *stimulus+choice* and in the context of *choice*. We train the linear SVM and compute the population vector in each of the two contexts, getting two vectors, $\tilde{\mathbf{w}}^{S+C}$ and $\tilde{\mathbf{w}}^{C}$, in each recording session. Decoding model in the context of *stimulus+choice* utilizes trials from conditions CM and CNM, while decoding model in the context of *choice* utilizes trials from conditions CNM and INM. The number of trials is imbalanced across conditions (namely, there are less trials in condition INM compared to CM), and such imbalance can affect the population vector. We balance the number of trials with the bootstrap method. In each recording session, we find the number of trials of the condition with most trials, and randomly sample, with repetition, the same number of trials from the two other distributions. All reported results are averaged across bootstraps.

We measure the similarity of the two population vectors by computing the angle between them,

$$\alpha = \arccos(\tilde{\mathbf{w}}^{S+C} \cdot \tilde{\mathbf{w}}^{C}) \tag{24}$$

where $(\cdot)$ is the dot product between the two vectors. Notice that, since vectors are normalized, we have that $||\tilde{\mathbf{w}}^{S+C}|| = ||\tilde{\mathbf{w}}^{C}|| = 1$. If the vectors $\tilde{\mathbf{w}}^{S+C}$ and $\tilde{\mathbf{w}}^{C}$ are similar, they point in similar direction, and the angle between them is small. If, conversely, the two vectors are pointing in random directions between 0 and π, the angle between them is, on average, orthogonal (the average is across bootstrapped samples). The significance of the angle is evaluated with the permutation test, using the test statistics of the angle, averaged across recording sessions. To construct the null model, we draw random vectors from the uniform distribution that have the same range as the true population vectors, and compute the angle between the two random vectors, $\alpha_p = \arccos(\tilde{\mathbf{w}}^{S+C}_p \cdot \tilde{\mathbf{w}}^{C}_p)$, $p = 1, ..., N_{perm}$. The p-value is computed by ranking the angle of the true model

among the distribution of angles of the model with random weights. The test is significant if $p < {}^{0.05}/_{N_{test}}$, where the division with the number of tests implements the Bonferroni correction for multiple testing.

Distance of decoding weights While the similarity of population vectors is a population-wise measure, we are also interested in understanding how different are weights of single neurons in the context of *stimulus+choice* and in the context of *choice*. For each single neuron, we compute the distance between its decoding weight in the two contexts.

$$d_n^{S+C,C} = |\tilde{w}_n^{S+C} - \tilde{w}_n^{C}| \tag{25}$$

Results are collected across recording sessions. We then split neurons in groups according to a specific criterion (sign of the weight, strength of the weight, burstiness, layers) and test the difference between groups with a two-tailed t-test.

Convergent cross-mapping

We use the version of convergent cross-mapping with random projection that ensures that eventual oscillation of the neural activity does not deteriorate the estimation [30]. The cross-mapping coefficient (CMC) is evaluated for smoothed spike trains of neurons that were recorded simultaneously, and is measured in single trials. The measure is repeated across trials and averaged across trials. Since the cross-mapping method is direction-selective, we get two coefficients for each pair of neurons, one from neuron 1 to neuron 2 and the other going in the opposite direction.

To compare the strength of synaptic interaction from one layer to another, we compute the cross-mapping coefficient (CMC) for each pair of neurons in the same layer, and compare distributions of CMCs across layers. To estimate the flow of information from one layer to another, we compute CMC coefficients for neuronal pairs in different layers (e.g., one neuron in the middle layer and the other neuron in the superficial layer, etc..). Averaging across homonymous pairs, we get an estimate of the average connectivity from one layer to another (directed graph). The strength of incoming (outgoing) connections is measured as the sum of the two incoming (outgoing) links at each layer. For example, the superficial layer receives connections from the middle and the deep layer, and the incoming connection to the superficial layer is the sum of the two incoming links. Finally, the flow of activity through a layer is estimated by summing the incoming and the outgoing connections to the layer, $flow = incoming + outgoing$.

Acknowledgements

This work was supported by Deutsche Forschungsgemeinschaft, grant GRK 1589/2.

Declaration of interest

The authors declare no competing interests.

References

1. Panzeri, S., Harvey, C. D., Piasini, E., Latham, P. E., & Fellin, T. (2017). Cracking the neural code for sensory perception by combining statistics, intervention, and behavior. Neuron, 93(3), 491-507.

2. Nienborg, H., & Cumming, B. G. (2006). Macaque V2 neurons, but not V1 neurons, show choice-related activity. Journal of Neuroscience, 26(37), 9567-9578.

3. Nienborg, H., & Cumming, B. G. (2014). Decision-related activity in sensory neurons may depend on the columnar architecture of cerebral cortex. Journal of Neuroscience, 34(10), 3579-3585.

4. Shadlen, M. N., & Newsome, W. T. (1996). Motion perception: seeing and deciding. Proceedings of the national academy of sciences, 93(2), 628-633.

5. Britten, K. H., Newsome, W. T., Shadlen, M. N., Celebrini, S., & Movshon, J. A. (1996). A relationship between behavioral choice and the visual responses of neurons in macaque MT. Visual neuroscience, 13(1), 87-100.

6. Shadlen, M. N., Britten, K. H., Newsome, W. T., & Movshon, J. A. (1996). A computational analysis of the relationship between neuronal and behavioral responses to visual motion. Journal of Neuroscience, 16(4), 1486-1510.

7. Logothetis, N. K., & Schall, J. D. (1989). Neuronal correlates of subjective visual perception. Science, 245(4919), 761-763.

8. Roitman, J. D., & Shadlen, M. N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. Journal of neuroscience, 22(21), 9475-9489.

9. Nienborg, H., & Cumming, B. G. (2009). Decision-related activity in sensory neurons reflects more than a neuron's causal effect. Nature, 459(7243), 89.

10. Uka, T., & DeAngelis, G. C. (2004). Contribution of area MT to stereoscopic depth perception: choice-related response modulations reflect task strategy. Neuron, 42(2), 297-310.

11. Gao, P., & Ganguli, S. (2015). On simplicity and complexity in the brave new world of large-scale neuroscience. Current opinion in neurobiology, 32, 148-155.

12. Saxena, S., & Cunningham, J. P. (2019). Towards the neural population doctrine. Current opinion in neurobiology, 55, 103-111.

13. Boerlin, M., Machens, C. K., & Denève, S. (2013). Predictive coding of dynamical variables in balanced spiking networks. PLoS computational biology, 9(11), e1003258.

14. Ni, A. M., Ruff, D. A., Alberts, J. J., Symmonds, J., & Cohen, M. R. (2018). Learning and attention reveal a general relationship between population activity and behavior. Science, 359(6374), 463-465.

15. Koren, V., & Denève, S. (2017). Computational Account of Spontaneous Activity as a Signature of Predictive Coding. PLoS computational biology, 13(1), e1005355.

16. Tkačik, G., Marre, O., Amodei, D., Schneidman, E., Bialek, W., & Berry II, M. J. (2014). Searching for collective behavior in a large network of sensory neurons. PLoS computational biology, 10(1), e1003408.

17. Tkačik, G., Mora, T., Marre, O., Amodei, D., Palmer, S. E., Berry, M. J., & Bialek, W. (2015). Thermodynamics and signatures of criticality in a network of neurons. Proceedings of the National Academy of Sciences, 112(37), 11508-11513.

18. Harris, K. D., Bartho, P., Chadderton, P., Curto, C., de la Rocha, J., Hollender, L., ... & Sakata, S. (2011). How do neurons work together? Lessons from auditory cortex. Hearing research, 271(1-2), 37-53.

19. Luczak, A., Barthó, P., & Harris, K. D. (2009). Spontaneous events outline the realm of possible sensory responses in neocortical populations. Neuron, 62(3), 413-425.

20. Churchland, M. M., Cunningham, J. P., Kaufman, M. T., Foster, J. D., Nuyujukian, P., Ryu, S. I., & Shenoy, K. V. (2012). Neural population dynamics during reaching. Nature, 487(7405), 51.

21. Rigotti, M., Barak, O., Warden, M. R., Wang, X. J., Daw, N. D., Miller, E. K., & Fusi, S. (2013). The importance of mixed selectivity in complex cognitive tasks. Nature, 497(7451), 585.

22. Veronika Koren, Ariana R. Andrei, Ming Hu, Valentin Dragoi, Klaus Obermayer bioRxiv 645135; https://doi.org/10.1101/645135, *preprint*

23. Koren V, Andrei AR, Hu M, Dragoi V, Obermayer K (2019) Reading-out task variables as a low-dimensional reconstruction of neural spike trains in single trials. PLoS ONE 14(10): e0222649.

24. Bernardi, S., Benna, M. K., Rigotti, M., Munuera, J., Fusi, S., & Salzman, C. D. (2018). The geometry of abstraction in hippocampus and pre-frontal cortex. bioRxiv, 408633.

25. Andrei, A. R., Pojoga, S., Janz, R., & Dragoi, V. (2019). Integration of cortical population signals for visual perception. Nature communications, 10(1), 1-13.

26. Hansen, B. J., Chelaru, M. I., & Dragoi, V. (2012). Correlated variability in laminar cortical circuits. Neuron, 76(3), 590-602.

27. Sugihara, G., & May, R. M. (1990). Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. Nature, 344(6268), 734-741.

28. Sugihara, G., May, R., Ye, H., Hsieh, C. H., Deyle, E., Fogarty, M., & Munch, S. (2012). Detecting causality in complex ecosystems. science, 338(6106), 496-500.

29. Sauer, T., Yorke, J. A., & Casdagli, M. (1991). Embedology. Journal of statistical Physics, 65(3-4), 579-616.

30. Tajima, S., Yanagawa, T., Fujii, N., & Toyoizumi, T. (2015). Untangling brain-wide dynamics in consciousness by cross-embedding. PLoS computational biology, 11(11), e1004537.

31. Tajima, S., Mita, T., Bakkum, D. J., Takahashi, H., & Toyoizumi, T. (2017). Locally embedded presages of global network bursts. Proceedings of the National Academy of Sciences, 114(36), 9517-9522.

32. Bondy, A. G., Haefner, R. M., & Cumming, B. G. (2018). Feedback determines the structure of correlated variability in primary visual cortex. Nature neuroscience, 21(4), 598-606.

33. Nienborg, H., R. Cohen, M., & Cumming, B. G. (2012). Decision-related activity in sensory neurons: correlations among neurons and with behavior. Annual review of neuroscience, 35, 463-483.

34. Averbeck, B. B., Latham, P. E., & Pouget, A. (2006). Neural correlations, population coding and computation. Nature reviews neuroscience, 7(5), 358-366.

35. Zylberberg, J., Pouget, A., Latham, P. E., & Shea-Brown, E. (2017). Robust information propagation through noisy neural circuits. PLoS computational biology, 13(4), e1005497.

36. Lisman, J. E. (1997). Bursts as a unit of neural information: making unreliable synapses reliable. Trends in neurosciences, 20(1)

37. Onorato, I., Neuenschwander, S., Hoy, J., Lima, B., Rocha, K. S., Broggini, A. C., Uran, C., Spyropoulos, G., Klon-Lipok, J., Womelsdorf, T. et al. (2020). A distinct class of bursting neurons with strong gamma synchronization and stimulus selectivity in monkey V1. Neuron, 105(1), 180-197.

38. Cumming, B. G., & Nienborg, H. (2016). Feedforward and feedback sources of choice probability in neural population responses. Current opinion in neurobiology, 37, 126-132.

39. Wang, X. J. (2002). Probabilistic decision making by slow reverberation in cortical circuits. Neuron, 36(5), 955-968.

40. Wang, X. J. (2008). Decision making in recurrent neuronal circuits. Neuron, 60(2), 215-234.

41. Callaway, E. M. (2004). Feedforward, feedback and inhibitory connections in primate visual cortex. Neural Networks, 17(5-6), 625-632.

42. Silvanto, J., Cowey, A., Lavie, N., & Walsh, V. (2005). Striate cortex (V1) activity gates awareness of motion. Nature neuroscience, 8(2), 143.

43. Kietzmann, T. C., Spoerer, C. J., Sörensen, L. K., Cichy, R. M., Hauk, O., & Kriegeskorte, N. (2019). Recurrence is required to capture the representational dynamics of the human visual system. Proceedings of the National Academy of Sciences, 116(43), 21854-21863.

44. Engel, T. A., & Wang, X. J. (2011). Same or different? A neural circuit mechanism of similarity-based pattern match decision making. Journal of Neuroscience, 31(19), 6982-6996.

45. Mante, V., Sussillo, D., Shenoy, K. V., & Newsome, W. T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. Nature, 503(7474), 78.

46. Ren, S. Q., Li, Z., Lin, S., Bergami, M., & Shi, S. H. (2019). Precise Long-Range Microcircuit-to-Microcircuit Communication Connects the Frontal and Sensory Cortices in the Mammalian Brain. Neuron, 104(2), 385-401.

47. Naud, R., & Sprekeler, H. (2018). Sparse bursts optimize information transmission in a multiplexed neural code. Proceedings of the National Academy of Sciences, 115(27), E6329-E6338.

48. Ledochowitsch, P., Yazdan-Shahmorad, A., Bouchard, K. E., Diaz-Botia, C., Hanson, T. L., He, J. W., ... & Schreiner, C. E. (2015). Strategies for optical control and simultaneous electrical readout of extended cortical circuits. Journal of neuroscience methods, 256, 220-231.

49. Zhao, Y., Yates, J. L., Levi, A. J., Huk, A. C., & Park, I. M. (2020). Stimulus-choice (mis) alignment in primate area MT. PLoS computational biology, 16(5), e1007614.

50. Brunton, B. W., Botvinick, M. M., & Brody, C. D. (2013). Rats and humans can optimally accumulate evidence for decision-making. Science, 340(6128), 95-98.

51. Kaufman, M. T., & Churchland, A. K. (2013). Cognitive neuroscience: sensory noise drives bad decisions. Nature, 496(7444), 172-173.

52. Shahidi, N., Andrei, A. R., Hu, M., & Dragoi, V. (2019). High-order coordination of cortical spiking activity modulates perceptual accuracy. Nature neuroscience, 22(7), 1148.

53. Nigam, S., Pojoga, S., & Dragoi, V. (2019). Synergistic Coding of Visual Information in Columnar Networks. Neuron, 104(2), 402-411.

54. Compte, A., Constantinidis, C., Tegnér, J., Raghavachari, S., Chafee, M. V., Goldman-Rakic, P. S., & Wang, X. J. (2003). Temporally irregular mnemonic persistent activity in prefrontal neurons of monkeys during a delayed response task. Journal of neurophysiology, 90(5), 3441-3454.

55. Pesaran, B., Pezaris, J. S., Sahani, M., Mitra, P. P., & Andersen, R. A. (2002). Temporal structure in neuronal activity during working memory in macaque parietal cortex. Nature neuroscience, 5(8), 805.