Excitatory and inhibitory subnetworks are equally selective during decision-making and emerge simultaneously during learning

Farzaneh Najafi[1], Gamaleldin F Elsayed[2], Robin Cao[3], Eftychios Pnevmatikakis[4], Peter E. Latham[3], John P Cunningham[5], Anne K Churchland[1,*]

[1]Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, [2]Columbia University, Neuroscience, New York, NY, [3]Gatsby Computational Neuroscience Unit, University College London, London, UK, [4]Flatiron Institute, New York, NY, [5]Columbia University, Statistics, New York, NY

[*]Correspondence: churchland@cshl.edu

Summary

Inhibitory neurons play a critical role in decision-making models and are often simplified as a single pool of non-selective neurons lacking connection specificity. This assumption is in keeping with observations in primary visual cortex: inhibitory neurons are broadly tuned in vivo, and show non-specific connectivity in slice. Selectivity of excitatory and inhibitory neurons within decision circuits is not known. We simultaneously measured their activity in the posterior parietal cortex of mice making multisensory decisions. Surprisingly, excitatory and inhibitory neurons were equally selective for the animal's choice, both at the single cell and population level. Further, excitatory and inhibitory populations exhibited similar changes in selectivity and temporal dynamics during the transition from novice to expert decision-making, paralleling behavioral improvements. These observations, combined with simulations, argue against models assuming non-selective inhibitory neurons. Instead, they argue for selective subnetworks of inhibitory and excitatory neurons that are shaped by experience to support expert decision-making.

Keywords

inhibitory neurons; excitatory neurons; decision-making models; learning; calcium imaging; posterior parietal cortex; decoding; stability; correlations; rate discrimination

Introduction

Theoretical models have been influential in interpreting behavioral and neural measurements during perceptual decision-making (Wang, 2002; Machens et al., 2005; Bogacz et al., 2006; Lim and Goldman, 2013; Rustichini and Padoa-Schioppa, 2015; Mi et al., 2017). Inhibition is an essential component of these models, usually provided by a single pool of non-selective inhibitory neurons receiving broad input from all excitatory neurons, independent of their functional selectivity (Deneve et al., 1999; Wang, 2002; Mi et al., 2017). Although widely accepted, these model architectures remain largely untested due to the challenges in identifying inhibitory cells reliably and in large numbers, in behaving animals.

The non-selective inhibition in modeling work is, presumably, motivated by empirical studies that examined the connectivity and tuning of inhibitory and excitatory neurons. The vast majority of these studies found broad tuning in inhibitory cells for visual stimulus features, such as orientation (Sohya et al., 2007; Niell and Stryker, 2008; Liu et al., 2009; Kerlin et al., 2010; Hofer et al., 2011; Atallah et al., 2012; Chen et al., 2013; Znamenskiy et al., 2018), spatial frequency (Niell and Stryker, 2008; Kerlin et al., 2010; Znamenskiy et al., 2018), and temporal frequency (Znamenskiy et al., 2018). The broad tuning in inhibitory neurons has been mostly attributed to their dense (Hofer et al., 2011; Packer and Yuste, 2011) and functionally unbiased inputs from the surrounding excitatory neurons (Kerlin et al., 2010; Hofer et al., 2011). This is in contrast to excitatory neurons, which show relatively sharp selectivity to stimulus features (Sohya et al., 2007; Niell and Stryker, 2008; Ch'ng and Reid, 2010; Kerlin et al., 2010; Hofer et al., 2011; Isaacson and Scanziani, 2011), reflecting their specific and non-random connectivity (Yoshimura et al., 2005; Ch'ng and Reid, 2010; Hofer et al., 2011; Ko et al., 2011; Cossell et al., 2015; Ringach et al., 2016).

The above studies have been chiefly performed in primary sensory areas; however, decision circuits include association areas, such as the posterior parietal cortex (PPC), which differ from early sensory areas in critical ways. In particular, there are marked differences in the relative abundance of distinct interneuron subtypes: association areas have more somatostatin (SOM)-positive interneurons, while early sensory areas have more parvalbumin (PV)-positive interneurons (Kim et al., 2017; Wang and Yang, 2018). The abundance of PV neurons in early sensory areas may be important for understanding the overall broad tuning observed in inhibitory neurons: PV interneurons are thought to have the broadest tuning, densest and least specific connectivity, compared to other interneuron subtypes (Wang et al., 2004; Wang and Yang, 2018). Furthermore, in contrast to primary sensory areas, association areas connect to multiple brain areas, and reflect higher order behavioral parameters, such as the animal's current choice and recent decision outcomes (Raposo et al., 2014; Morcos and Harvey, 2016; Hwang et al., 2017; Pho et al., 2018).

Therefore, it is unclear whether the findings from early sensory areas, broad tuning and relatively nonspecific connectivity of interneurons, are applicable to association areas involved in decision-making computations. Here, we attempted to bridge this gap by comparing the selectivity of inhibitory and excitatory neurons in PPC of mice during rate discrimination decisions. Our findings, supported by simulations, demonstrate that, in contrast to predictions from early sensory areas, excitatory and inhibitory neurons in PPC are equally choice-selective. Also,

during learning, the specificity of inhibitory neurons increased in parallel with excitatory neurons. These results argue that in decision areas, subnetworks of selective inhibitory neurons emerge during learning and are engaged during expert decisions.

Results

To test how excitatory and inhibitory neurons coordinate during decision-making, we measured neural activity in transgenic mice trained to report decisions about the repetition rate of a sequence of multisensory events by licking to a left or right waterspout (Figure 1A; Figure S1A). Trials consisted of simultaneous clicks and flashes, separated by Poisson distributed intervals so that the instantaneous rate fluctuated over a 1000 ms period (Brunton et al., 2013; Odoemene et al., 2017). Mice reported whether event rates were high or low compared to an abstract category boundary (16 Hz) that they learned from experience. Decisions depended strongly on stimulus rate: performance was at chance when the stimulus rate was at the category boundary, and was higher at rates further from the category boundary (Figure 1B). A logistic regression model demonstrated that choice depends on the current stimulus strength, previous choice outcome (Hwang et al., 2017), and the time elapsed since the previous trial (Figure S1B). We imaged excitatory and inhibitory neural activity by injecting a viral vector containing the calcium indicator GCaMP6f to layer 2/3 of mouse Posterior Parietal Cortex (PPC; 2mm posterior to Bregma, 1.7mm lateral to midline (Harvey et al., 2012; Funamizu et al., 2016; Goard et al., 2016; Morcos and Harvey, 2016; Hwang et al., 2017; Song et al., 2017)). Mice expressed the red fluorescent protein tdTomato transgenically in all GABAergic inhibitory neurons. We used a two-channel two-photon microscope to record the activity of all neurons, a subset of which were identified as inhibitory neurons (Figure 1C). This allowed us to measure the activity of excitatory and inhibitory populations in the same animal.

To detect neurons and extract calcium signals from imaging data, we leveraged an algorithm that simultaneously identifies neurons, de-noises the fluorescence signal and de-mixes signals from spatially overlapping components (Pnevmatikakis et al., 2016; Giovannucci et al., 2018) (Figure 1D middle). The algorithm also estimates spiking activity for each neuron, yielding, for each frame, a number that is related to the spiking activity during that frame (Figure 1D right). We refer to this number as "inferred spiking activity", acknowledging that estimating spikes from calcium signals is challenging (Chen et al., 2013). In particular, while higher inferred spiking activity within a single neuron indicates higher firing rates, comparison of firing rates across neurons is not possible with this method. Analyses were performed on inferred spiking activity. To identify inhibitory neurons, we used a method that we developed to correct for bleed-through from the green to the red channel (Methods). Next, we identified a subset of GCaMP6f-expressing neurons as inhibitory neurons based on the signal intensity on the red channel as well as the spatial correlation between red and green channels (Figure 1C right, cyan circles). Inhibitory neurons constituted 8.3% of the population, close to previous reports (Sahara et al., 2012) but slightly lower due to our desire to be conservative in assigning neurons to the inhibitory pool (Methods).

Confirming previous reports (Funamizu et al., 2016; Morcos and Harvey, 2016; Runyan et al., 2017), we observed that the activity of individual neurons peaked at time points that spanned the trial (Figure 1E,F). Diverse temporal dynamics were evident in both cell types (Figure 1E,F).

The inferred spiking activity was significantly different for inhibitory compared to excitatory neurons throughout the trial (Figure 1G; t-test, $p<0.001$). In the moments before the choice (97.1ms, average of 3 frames), this difference was clear (Figure 1H) and significant for all mice

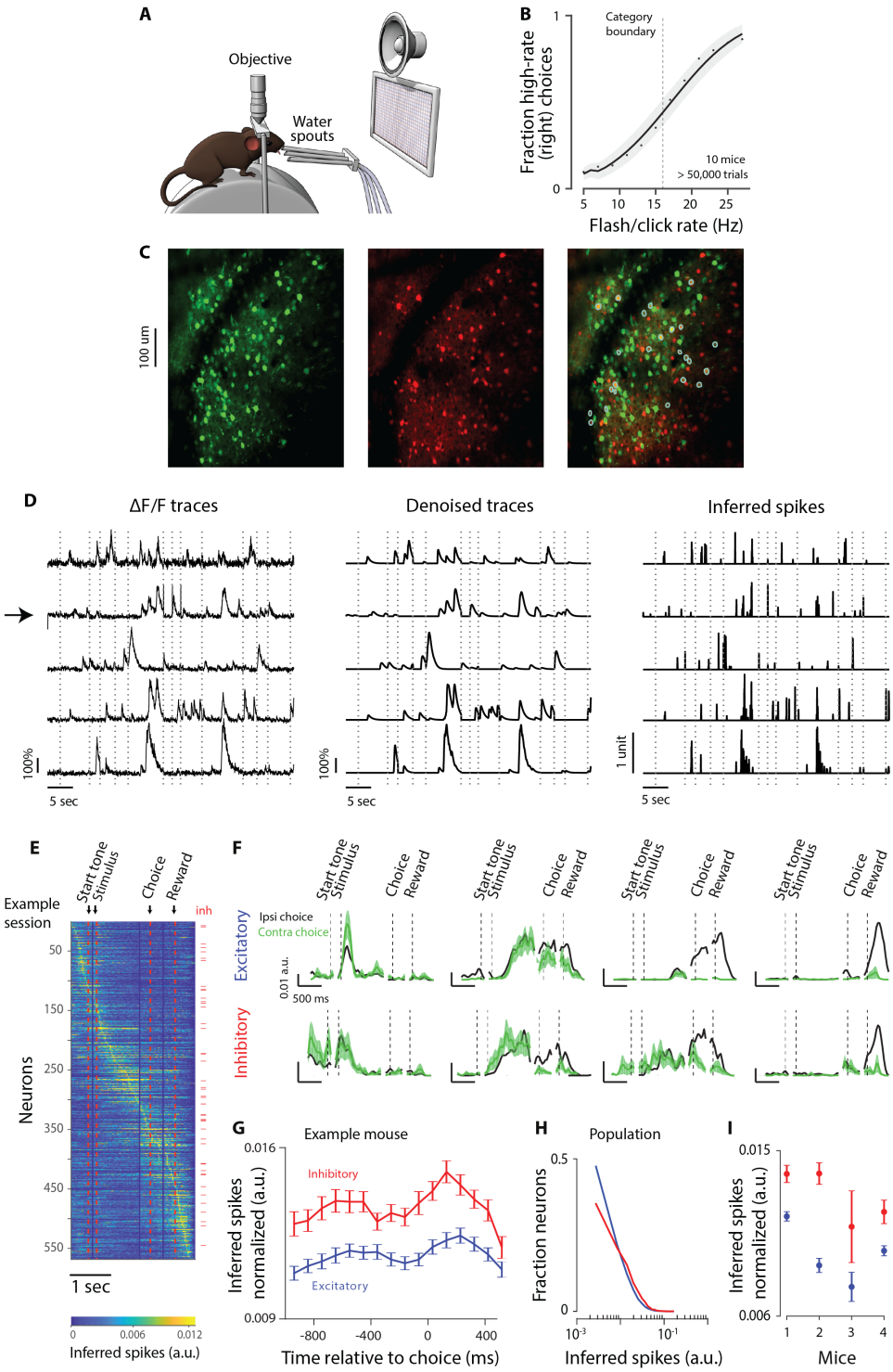

Figure 1. Simultaneous imaging of inhibitory and excitatory populations during decision-making to test decision-making models.
A. Behavioral apparatus in which a head-fixed mouse is atop a cylindrical wheel. Visual display and speaker present the multisensory stimulus. To initiate a trial, mice licked the middle waterspout. To report the decision about the stimulus rate, mice licked left/right spouts. Objective belongs to the 2-photon microscope used to image neural activity through a window implanted in the skull. **B.** Psychometric function showing the fraction of trials in which the mouse judged the stimulus as high rate as a function of stimulus rate. Dots: data, mean across 10 mice. Line: Logit regression model fit using glmfit.m; mean across mice. Shaded area: standard deviation of the fit across mice. Dashed vertical line: category boundary (16Hz). **C,** Average projection image of 10,000 frames. **Left:** green channel showing GCaMP6f expression. **Middle:** red channel showing tdTomato expression. **Right:** merge of left and middle. Cyan circles indicate GCaMP6f-expressing neurons that were identified as inhibitory. **D,** Five example neurons identified by the CNMF algorithm (arrow: inhibitory neuron). **Left:** raw $\Delta F/F$ traces. **Middle:** de-noised traces. **Right:** inferred spiking activity; normalized so that 1 unit corresponds to 100% $\Delta F/F$. Imaging was not performed during inter-trial intervals; traces from 13 consecutive trials are concatenated; dashed lines: trial onsets. **E,** Example session with 568 neurons. Each row shows the trial-averaged inferred spiking activity of a neuron (frame resolution: 32.4ms). Neurons are sorted according to the timing of their peak activity. Inhibitory neurons (n=45) are indicated by red ticks on the right. Red vertical lines mark trial events: initiation tone, stimulus onset, choice, and reward. Duration between events varied across trials, so traces were aligned for each trial event, and then concatenated. Vertical blue lines are due to concatenation. **F,** Trial-averaged inferred spiking activity of 4 excitatory (top) and 4 inhibitory (bottom) neurons, for ipsi (black) and contra-lateral (green) choices (mean +/- standard error; ~250 trials per session). **G,** Inferred spiking activity for excitatory (blue) and inhibitory (red) neurons during the course of a trial. Example mouse; mean +/- standard error across days (n=46). Each point corresponds to an average over trials and neurons. Inferred spiking activity was initially downsampled by averaging over three adjacent frames (Methods). Spiking activity was significantly higher for inhibitory neurons (t-test; $p<0.001$) at all times. **H,** Distribution of inferred spiking activity at time bin 0-97ms (averaged over the three frames before the choice) for all mice and all sessions (41,723 excitatory and 5,142 inhibitory neurons). **I,** Inferred spiking activity at time bin 0-97ms before the choice for each individual mouse (mean +/- standard error across days). Differences were significant for all subjects (t-test; $p<0.001$). In **E-I**, inferred spiking activity of each neuron is normalized to its maximum value during the session (Methods).

(Figure 1I). The probable differences in GCaMP expression levels and calcium buffering between excitatory and inhibitory neurons, as well as how spiking activity is inferred (Methods), precludes a direct estimate of the underlying firing rates (Kwan and Dan, 2012). However, the significant difference in the inferred spiking activity between excitatory and inhibitory neurons provides further evidence that we successfully identified two separate neural populations.

Individual excitatory and inhibitory neurons are similarly choice-selective

To assess the selectivity of individual excitatory and inhibitory neurons for the decision outcome, we performed receiver operating characteristic (ROC) analysis (Green and Swets, 1966) on single-neuron responses. For each neuron, at each time point, we calculated the area under the ROC curve (AUC) as a measure of the amount of overlap between the response distributions for ipsilateral vs. contralateral choices. A neuron was identified as "choice-selective" if its AUC value was significantly different ($p<0.05$) from a constructed shuffled distribution (Figure S2A; Methods), indicating that the neural activity was significantly different for ipsi- vs. contralateral choices (Figure 2A, shaded areas mark choice-selective neurons).

Approximately 13% of the neurons were significantly choice selective. Importantly, the fraction of choice-selective neurons was similar for both inhibitory and excitatory cells (Figure 2B; 15%

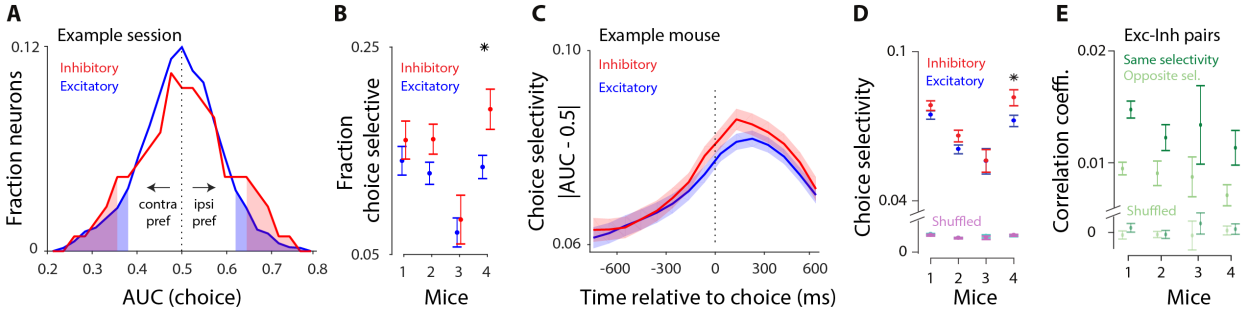

Figure 2. Single-cell and pairwise analyses argue for non-random connections between excitatory and inhibitory neurons.
A-D, Ideal observer analysis reveals the ability of individual neurons to distinguish left vs. right choices. In all panels, blue and red indicate excitatory and inhibitory neurons, respectively. **A,** Distribution of AUC values (area under the curve) of an ROC analysis for distinguishing choice from the activity of single neurons in an example session. Data correspond to the 97 ms window preceding the choice for 285 excitatory and 29 inhibitory neurons. Values large than 0.5 indicate neurons preferring the ipsi-lateral choice; values smaller than 0.5 indicated preference for the contralateral choice. Shaded areas mark significant AUC values (compared to a shuffle distribution). **B,** Fraction of excitatory and inhibitory neurons that are significantly choice-selective for each mouse: mean+/-standard error across days (n = 45, 48, 7, 35 sessions per mouse, measured 0-97 ms before the decision). Star (*) indicates significant difference between excitatory and inhibitory neurons (t-test; $p<0.05$). **C,** ROC analysis performed on 97 ms non-overlapping time windows. Time course of choice selectivity (absolute deviation of AUC from chance) shown for excitatory and inhibitory neurons in an example mouse; mean+/-standard error across days, n=45 sessions. **D,** Average choice selectivity of excitatory and inhibitory neurons from 0-97 ms before the choice is summarized for each mouse; mean+/-standard error across days. **E,** Noise correlations (Pearson's coefficient) for pairs of excitatory-inhibitory neurons with the same choice selectivity (dark green) or opposite choice selectivity (light green, i.e. one neuron prefers ipsi, and the other neuron prefers contralateral choice). "Shuffled" in (D,E) denotes quantities were computed using shuffled trial labels. Mean+/-standard error across days; 0-97 ms before the choice.

of inhibitory neurons and 12% of excitatory neurons; averaged across mice). Additionally, the time course and magnitude of choice-selectivity was similar between the two cell types (Figure 2C,D): choice selectivity gradually increased during the trial, peaking just after the animal reported its choice (Figure 2C). The existence of similarly strong choice selectivity in excitatory and inhibitory neurons was surprising, given the commonly accepted assumption of non-specific inhibition in theoretical studies (Deneve et al., 1999; Wang, 2002; Mi et al., 2017), as well as the empirical findings suggesting broad tuning and weakly specific connectivity in inhibitory neurons (Sohya et al., 2007; Niell and Stryker, 2008; Liu et al., 2009; Kerlin et al., 2010; Hofer et al., 2011; Packer and Yuste, 2011; Atallah et al., 2012; Chen et al., 2013). This is a first hint that specific functional subnetworks, preferring either ipsi- or contra-lateral choices, exist within the inhibitory population, just like the excitatory population (Yoshimura and Callaway, 2005; Znamenskiy et al., 2018).

If choice selectivity is preserved because excitatory and inhibitory neurons with the same choice preference tend to be preferentially connected, one prediction is that excitatory and inhibitory neurons with the same choice selectivity will be more strongly correlated with each other compared to excitatory and inhibitory neurons with the opposite choice selectivity. This was

indeed the case (Figure 2E: noise correlations, see Methods; Figure S2F: spontaneous correlations). The stronger correlation between neurons with the same choice selectivity was also evident in pairs consisting of only excitatory or only inhibitory neurons (Figure S2E,F), in keeping with previous observations in mouse V1 during passive viewing (Hofer et al., 2011; Ko et al., 2011; Cossell et al., 2015), as well as the prefrontal cortex in behaving monkeys (Constantinidis and Goldman-Rakic, 2002).

Finally, we assessed whether neurons reflected the animal's choice or the sensory stimulus, by comparing choice selectivity values computed on correct vs. error trials. For the majority of neurons, choice selectivity computed on correct trials was similar to that of error trials, resulting in a positive correlation of the two quantities across neurons (Figure S2C,D). Positive correlations indicate that most neurons primarily reflect the impending choice, as opposed to the sensory stimulus that informed it (Methods). Variability across mice in the strength of this correlation may indicate that the balance of sensory vs. choice signals within individual neurons varied across subjects (perhaps due to imaged subregions within the window, Figure S2D).

The animal's choice can be decoded with equal accuracy from both populations

While individual inhibitory neurons could distinguish the animal's choice about as well as excitatory ones, the overall choice selectivity in single neurons was small (Figure 2D). To further evaluate the discrimination ability of inhibitory neurons, we leveraged our ability to measure hundreds of neurons simultaneously. Specifically, we examined the ability of a linear classifier (support vector machine, SVM (Hofmann et al., 2008)) to predict the animal's choice from the single-trial population activity (cross-validated; L2 penalty; see Methods).

We first performed this analysis on all neurons imaged simultaneously in a single session (Figure 3A, left), training the classifier separately for every moment in the trial (97 ms bins). Classification accuracy gradually grew after stimulus onset and peaked at the time of the choice (Figure 3B, black). Performance was at chance on a shuffle control in which trials were randomly assigned as left or right choice (Figure 3B, shuffled). The ability of the entire population of PPC neurons to predict the animal's upcoming choice confirms previous observations (Funamizu et al., 2016; Goard et al., 2016; Morcos and Harvey, 2016; Driscoll et al., 2017). As with the single neuron data (Figure S2C,D), analysis of error trials indicated that the animal's choice, rather than the stimulus, was generally the key feature driving the high classification accuracy we observed (Figure S3A,B).

We then examined classifier accuracy for excitatory and inhibitory populations separately. For excitatory neurons, we subsampled the population so that the total number of neurons matched the number of inhibitory neurons in the same session (Figure 3A, middle). As expected, overall classification accuracy was reduced due to the smaller population size; although performance was still well above chance and the temporal dynamics were the same as when all neurons were included (Figure 3B, blue trace). Finally, we included all inhibitory neurons (Figure 3A, right). Surprisingly, the classification accuracy of inhibitory neurons did not systematically differ from that of excitatory neurons during the course of a trial (Figure 3B example session; red and blue traces overlap; Figure S3C: additional example sessions). Similar classification accuracy for excitatory and inhibitory populations was observed in all subjects (Figure 3C).

Our analysis may have obscured a difference between excitatory and inhibitory neurons because

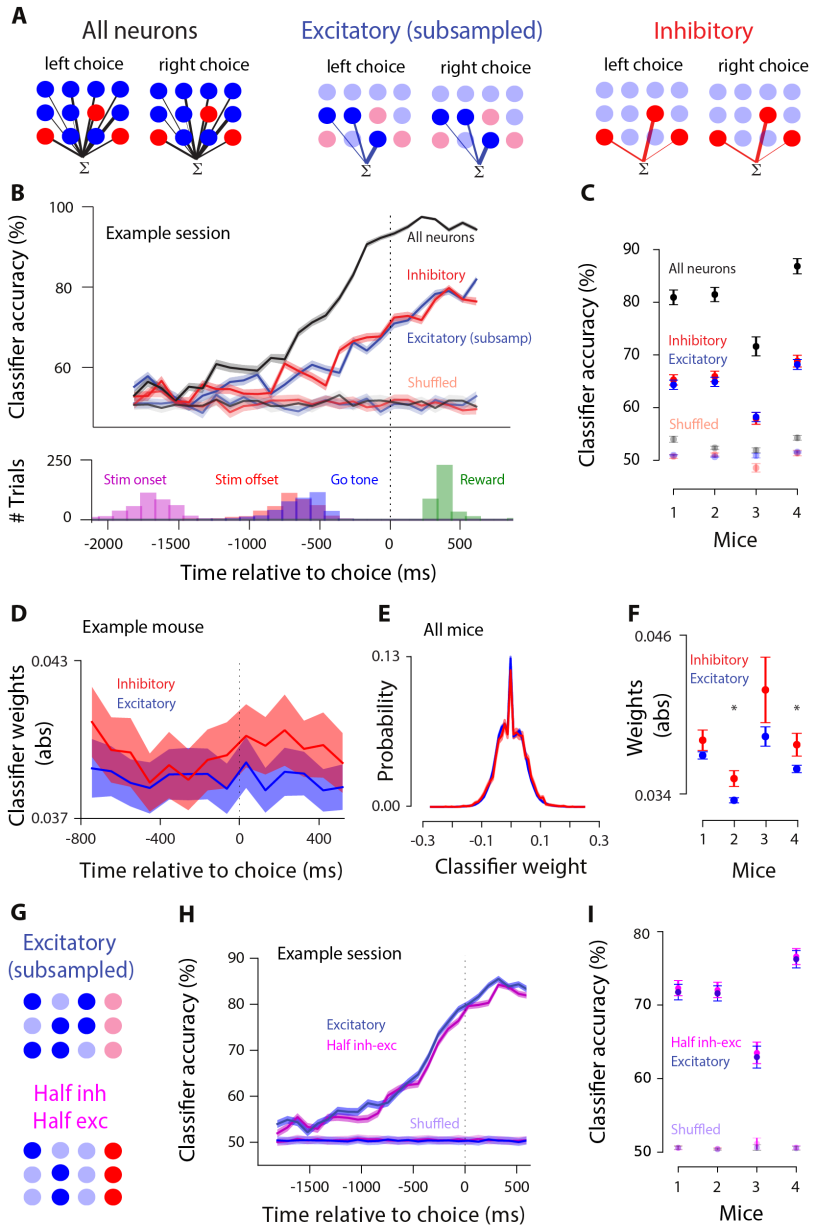

Figure 3. Linear classifiers can predict the animal's choice with equally high accuracy from the activity of either excitatory or inhibitory populations.
A, Schematic of decoding choice from the population activity of all neurons (left), only excitatory neurons (middle), subsampled to the same number as inhibitory neurons, and only inhibitory neurons (right). A linear SVM assigns weights of different magnitude (indicated by lines of different thickness) to each neuron in the population so that a weighted sum of population activity differs for trials preceding left vs. right choices. **B, Top:** classification accuracy of decoders trained on all neurons (black), subsampled excitatory neurons (blue), and inhibitory neurons (red) (cross-validated; decoders trained on every 97ms time bin; example session; mean+/-standard error across 50 cross-validated samples). Classification accuracy is lower for inhibitory or subsampled excitatory populations (red, blue) relative to all neurons (black) because of the smaller population size. Classifier

accuracy was similar for excitatory and inhibitory populations throughout the trial. Unsaturated lines show performance on shuffled trial labels. **Bottom:** distribution of stimulus onset, stimulus offset, go tone, and reward occurrence for the example session shown on the top. **C,** Classification accuracy during 0-97 ms before the choice for 4 animals on real (saturated) and shuffled (unsaturated) data. Mean+/-standard error across days per mouse. **D-F,** When all neurons were included in the decoder (panel A, left), excitatory and inhibitory neurons were assigned weights of similar magnitude. **D,** Absolute value of weights for excitatory and inhibitory neurons in the decoders trained on all neurons, at every moment in the trial; example mouse; mean+/-standard error across days. **E,** Distribution of classifier weights (decoder training time: 0-97 ms before the choice) are similar for excitatory and inhibitory neurons. Neurons from all mice pooled (42,019 excitatory and 5,172 inhibitory neurons). Shading reflects the standard error in each bin of the distribution. **F,** Absolute value of weights in the classifier trained from 0-97 ms before the choice for excitatory vs. inhibitory neurons, for each mouse. Mean+/-standard error across days. Star indicates $P<0.05$, t-test. **G,** Schematic of decoding choice from a population of subsampled excitatory neurons (top) vs. a population of the same size but including half inhibitory and half excitatory neurons (bottom). **H,** Classifier accuracy of populations including only excitatory (blue) or half inhibitory, half excitatory neurons (magenta); example session. Classifier trained at each moment in the trial. Traces show mean+/-standard error across 50 cross-validated samples. **I,** Summary of each mouse (mean+/-standard error across days) for the decoders trained from 0-97 ms before the choice.

it evaluated their performance separately, rather than considering how these neurons are leveraged collectively in a classifier that can take advantage of both cell types. To test this, we examined the classifier that was trained on all neurons (Figure 3A left; Figure 3B black), and compared the classifier weights assigned to excitatory vs. inhibitory neurons. If excitatory and inhibitory neurons contributed equally to the decoder, they should be assigned comparable weights by the classifier. This is indeed what we found: the weight magnitudes of excitatory and inhibitory neurons were matched for the entire course of the trial (Figure 3D; absolute value of weights). Also the distributions of weights were overlapping (Figure 3E: weights of all neurons of all mice. Figure 3F: absolute weights). The comparable classifier weights for excitatory and inhibitory neurons demonstrate that both cell types were similarly informative about the animal's upcoming choice.

Finally, we tested whether excitatory and inhibitory populations can be decoded more accurately from a mixed population. This could occur, for example, if the excitatory-inhibitory correlations were weak relative to excitatory-excitatory and inhibitory-inhibitory correlations (Panzeri et al., 1999; Averbeck et al., 2006; Moreno-Bote et al., 2014). To assess this, we trained the classifier on a population that included half excitatory and half inhibitory neurons (Figure 3G bottom), and compared its choice-prediction accuracy with the classifier that was trained on a population of the same size, but consisted only of excitatory neurons (Figure 3G top). We found similar classification accuracy for both decoders during the entire trial (Figure 3H,I), arguing that a mixed population offers no major advantage to decoding.

Implications for circuitry

These results would seem to rule out circuitry from traditional decision-making models, in which the inhibitory neurons are non-selective. This is because the average input to the inhibitory neurons is the same whether excitatory population E_1 or excitatory population E_2 is active (see Figure 4A, top panel). However, while the average input is the same, there are fluctuations in connection strength, leading to input fluctuations that can be informative. For instance, suppose that, because of the inherent randomness in neural circuits, an inhibitory neuron received more

connections from the excitatory neurons in population E_1 than those in population E_2. In that case, the firing rate of the inhibitory neuron would be slightly higher when population E_1 was active than when population E_2 was active. That difference in firing rate could potentially be exploited by a classifier to predict the choice of the animal.

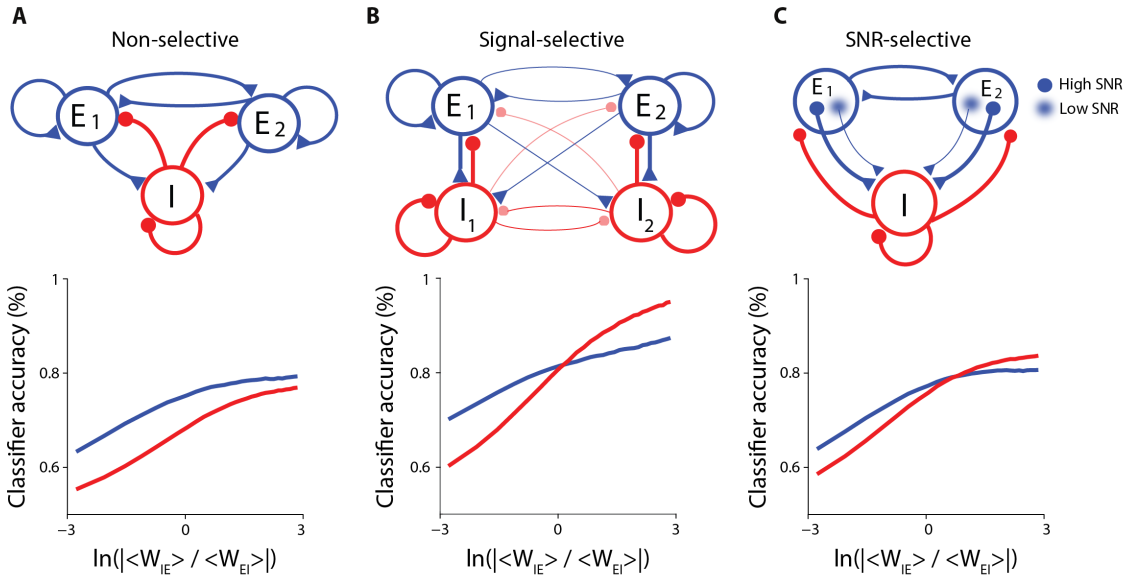

Figure 4. Selective connectivity between excitatory and inhibitory neurons allows for matched classification accuracy in the two populations.
A, Top: Non-selective decision-making model. E_1 and E_2 represent pools of excitatory neurons, each favoring a different choice. Both pools excite a single pool of non-selective inhibitory neurons (I), which, in turn, provides inhibition to both excitatory pools. Recurrent excitation drives persistent activity in the network, and inhibition allows for competition between the two choices. **Bottom**: Classification accuracy of excitatory (blue) and inhibitory (red) neurons as a function of the relative strength of excitatory-to-inhibitory vs. inhibitory-to-excitatory connections. For all values of this parameter, excitatory neurons had higher classification accuracy than inhibitory ones. This was true for all parameters tested (Methods and Supporting information, Figure S4; angle brackets denote averages over weights). **B, Top**: Selective decision-making model. I_1 and I_2 represent pools of inhibitory neurons that connect more strongly to E_1 and E_2, respectively, than to E_2 and E_1, and all cross-pool connections are weaker than within-pool connections. **Bottom**: Decoding accuracy depends on the relative strength of excitatory to inhibitory versus inhibitory to excitatory connections, with larger excitatory to inhibitory connections favoring inhibitory neurons. Cross-pool connectivity was 25% smaller than within-pool connectivity. **C, Top**: Selective decision-making model, except now inhibitory neurons connect more strongly to excitatory neurons with high signal to noise ratios. **Bottom**: Again, decoding accuracy depends on the relative strength of excitatory to inhibitory versus inhibitory to excitatory connections, with larger excitatory to inhibitory connections favoring inhibitory neurons. For all plots we used 50 excitatory and 50 inhibitory neurons out of a population containing 4000 excitatory and 1000 inhibitory neurons.

To test this quantitatively, we evaluated the classification accuracy of inhibitory neurons in a simulated non-selective circuit (Methods). For such a circuit, inhibitory neurons had lower classification accuracy than excitatory neurons (Figure 4A, bottom), inconsistent with our experimental results (Figure 3B,C). The classification accuracy of inhibitory neurons was low because even modest amounts of noise in the system were sufficient to overcome any informative input fluctuations. By contrast, if inhibitory neurons were connected preferentially to

one excitatory pool over the other, inhibitory and excitatory neurons could have matched classification accuracy (Figure 4B, bottom; Figure S4). Interestingly, a third circuit configuration likewise gave rise to inhibitory neurons with high classification accuracy (Figure 4C). In this configuration, inhibitory neurons were non-selective with respect to the excitatory pools, but were connected to the more informative excitatory neurons (those with a high signal-to-noise ratio; Figure 4C, top).

While it is not surprising that selective circuitry improves the decoding accuracy of inhibitory neurons, our results raise two questions. First, how can the inhibitory neurons have better decoding accuracy than the excitatory ones (Figure 4B,C, bottom; for part of the plot, red is above blue)? After all, in our model all information about the choice flows through the excitatory neurons. Second, why is the relative strength of the excitatory to inhibitory versus inhibitory to excitatory connections an important parameter (Figure 4, bottom; x-axis)? The answers are related. Increasing the strength of the excitatory to inhibitory connections increases the signal in the inhibitory neurons, and therefore effectively decreases the noise added to the inhibitory population (see Methods for details). This decrease in noise leads to improved decoding accuracy of both the excitatory and inhibitory populations, because the two populations are connected. However, the decrease in the noise added to the inhibitory neurons has a bigger effect on the inhibitory than the excitatory population; that's because noise directly affects the inhibitory neurons, but only indirectly, through the inhibitory to excitatory connections, affects the excitatory neurons. Thus, in all panels of figure 4, the classification accuracy increases faster for inhibitory neurons than excitatory ones as the excitatory to inhibitory connection strength increases.

Excitatory and inhibitory populations reflect the animals' choice with similar temporal stability of the classifier weights

If excitatory and inhibitory neurons are connected within subnetworks with frequent cross talk, the two populations should not only predict the animal's choice with similar accuracy, as shown above, but the readout weights (the weights assigned by the classifier) should exhibit similar temporal dynamics. To assess this, we quantified each population's stability: the extent to which a classifier trained at one moment could successfully classify neural activity as preceding left vs. right choice at different moments. If population-wide patterns of activity are similar over time (e.g., all neurons gradually increase their firing rates), classifiers trained at one moment can accurately classify neural activity at different moments. Excitatory and inhibitory populations might differ in this regard, with one population more stable than the other.

As the gap between testing and training time increased, a gradual drop occurred in the classifier accuracy, as expected (Figure 5A,B). This drop in accuracy occurred at a very similar rate for excitatory and inhibitory populations (Figure 5B). To quantify this, we determined the time window over which the classifier accuracy was within 2 standard deviations of the accuracy at the same training-testing time window (Figure 5C). This was indistinguishable for excitatory and inhibitory neurons (Figure 5D; Figure S5A). An alternate method for assessing stability, computing the angle between the weights of pairs of classifiers trained at different time windows, likewise suggested that excitatory and inhibitory populations are similarly stable (see Methods; Figure S5C).

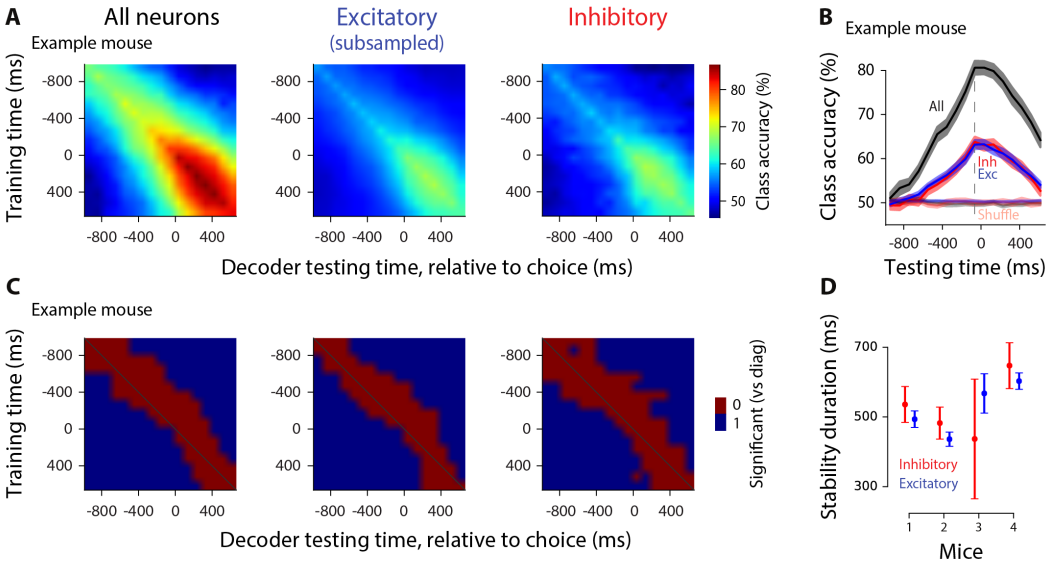

Figure 5. Classifiers, whether trained on excitatory or inhibitory neurons, show comparable stability during decision formation.
Cross-temporal generalization of choice decoders. **A,** Classification accuracy of decoders for each pair of training and testing time points, using the population activity of all neurons (left), subsampled excitatory neurons (middle), or inhibitory neurons (right). Diagonal: same training, testing time (same as in Figure 3). Example mouse, mean across 45 sessions. **B,** Example classification accuracy traces showing how classifiers trained at 0-97 ms before choice generalize to other times in the trial. Excitatory and inhibitory neurons show the same time course of generalization. Same mouse as in (A), mean+/-standard error across days **C,** Decoders are stable in a short window outside their training time. Red indicates stability: classification accuracy of a decoder tested at a time different from its training time is within 2 standard deviations of the decoder tested at the same time as the training time. Example mouse; mean across days. **D,** Summary of stability duration for the decoder trained from 0-97 ms before the choice, using inhibitory neurons (red) or subsampled excitatory neurons (blue), for each mouse. Mean+/-standard error across days, per mouse.

Correlated variability limits classification accuracy and is higher among inhibitory neurons

Our results thus far demonstrate that neural activity in both excitatory and inhibitory neurons reflects an animal's impending choice (Figure 3B,C), and that there are significant noise correlations among neurons in PPC (Figure 2E). However, the analyses thus far do not demonstrate how this noise affects the ability to decode neural activity overall, or for excitatory and inhibitory populations separately. Examining the effect of noise is essential because noise correlations can limit or enhance the ability to decode population activity depending on their structure (Panzeri et al., 1999; Averbeck et al., 2006). Fortunately, our dataset includes simultaneous activity from hundreds of neurons and is therefore especially well-suited to assess noise correlations: correlations can have a large effect at the population level even when their effect at the level of neuron pairs is small (Averbeck et al., 2006; Moreno-Bote et al., 2014).

253 To examine how noise correlations affected classification accuracy, we sorted neurons based on
254 their individual choice selectivity, added them one by one to the population (from highest to
255 lowest choice selectivity defined as |AUC-0.5|), and measured classification accuracy as a
256 function of population size. Classification accuracy improved initially as more neurons were
257 included in the decoder, but quickly saturated (Figure 6A black; 0-97 ms before the choice). To
258 understand why classification accuracy saturates, we tested the effect of noise correlations on
259 classification accuracy. Specifically, we created "pseudo populations", in which each neuron in
260 the population was taken from a different trial (Figure 6A gray). This removed noise correlations
261 because those are shared across neurons within a single trial. Higher classification accuracy in

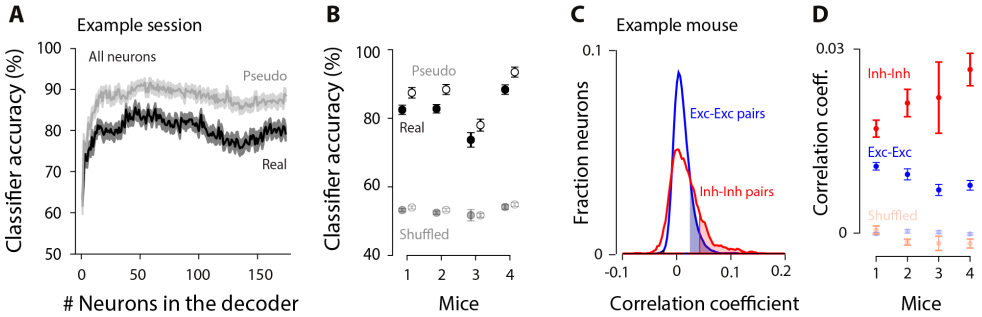

Figure 6. Noise correlations limit the efficacy of pooling and are higher among inhibitory neurons.
A, Classification accuracy for an example session (at time window 0-97 ms before the choice) on neural ensembles of increasingly larger size, with the most choice-selective neurons added first. Mean+/-standard error across 50 cross-validated samples. Gray: classification accuracy for pseudo-populations, in which noise correlations were removed by shuffling. Black: real populations. Both cell types were included ("All neurons"). **B,** Summary for each mouse; points show mean+/-standard error across days. Values were computed for the largest neuronal ensemble (the max value on the horizontal axis in (A)). **C,** Example mouse: distribution of noise correlations (Pearson's correlation coefficients, 0-97 ms before the choice) for excitatory neurons (blue; n=11867) and inhibitory neurons (red; n=1583). Shaded areas indicate significant quantities compared to a shuffled control: trial orders were shuffled for each neuron to remove noise correlations. **D,** Summary of noise correlation coefficients for each mouse; mean+/-standard error across days.

262 pseudo populations compared to real populations indicates the presence of noise that overlaps
263 with signal, limiting information (Panzeri et al., 1999; Averbeck et al., 2006; Averbeck and Lee,
264 2006; Moreno-Bote et al., 2014). This is what we observed (Figure 6A, gray trace above black
265 trace). Across all mice, removing noise correlations resulted in a consistent increase in
266 classification accuracy for the full population (Figure 6B; filled vs. open circles). This
267 establishes that noise correlations limit population decoding in PPC.

268 We next compared the strength of pairwise noise correlations for excitatory vs. inhibitory
269 neurons. Inhibitory pairs had significantly higher noise correlations compared to excitatory pairs
270 (Figure 6D). Similar results were observed when we measured correlations during the
271 spontaneous activity before the stimulus was presented (Figure S6B), ensuring that correlations
272 did not arise from shared sensory-driven responses. Importantly, we obtained the same results
273 even when we restricted the analysis to those inhibitory and excitatory neurons that had the same
274 inferred spiking activity (Figure S6C,D). This was done because the higher spiking activity of
275 inhibitory neurons (Figure 1G-I) could potentially muddle the comparison of pairwise noise
276 correlations between excitatory and inhibitory neurons. Finally, similar to previous reports

(Hofer et al., 2011; Khan et al., 2018), we found intermediate correlations for pairs consisting of one inhibitory neuron and one excitatory neuron (Figure S6).

These findings align with previous studies in sensory areas that have demonstrated stronger correlations among inhibitory neurons (Hofer et al., 2011; Khan et al., 2018) due to their dense connectivity (Galarreta and Hestrin, 1999; Packer and Yuste, 2011; Kwan and Dan, 2012). The difference we observed between excitatory and inhibitory neurons argues that this feature of early sensory circuits is shared by decision-making areas. Further, this clear difference between excitatory and inhibitory neurons, like the difference in inferred spiking (Fig. 1G-I), confirms that we successfully measured two distinct populations.

Selectivity increases in parallel in inhibitory and excitatory populations during learning

Our observations thus far argue that excitatory and inhibitory neurons form selective subnetworks (Yoshimura and Callaway, 2005; Znamenskiy et al., 2018). To assess whether the emergence of these subnetworks is experience-dependent, and if it varies between inhibitory and excitatory populations, we measured neural activity as animals transitioned from novice to expert decision-makers (3 mice; 35-48 sessions; Figure S7A). We trained a linear classifier for each training session, and for each moment in the trial. This allowed us to compare the dynamics of the choice signal in excitatory and inhibitory populations over the course of learning.

Classification accuracy of the choice decoder increased consistently as animals became experts in decision-making (Figure 7A left; Figure 7D black), leading to a strong correlation between the classifier performance and the animal's performance across training days (Figure 7B left). The population representation of the choice signal also became more prompt: the choice signal appeared progressively earlier in the trial as the animals became experts. Initially, classification accuracy was high only after the choice (Figure 7A, black arrow). As the animals gained experience, high classification accuracy occurred progressively earlier in the trial, eventually long before the choice (Figure 7A, gray arrow). This resulted in a negative correlation between the animal's performance and the onset of super-threshold choice accuracy across training days (Figure 7C left; Figure 7E black).

Importantly, the dynamics of the choice decoder changed in parallel in both excitatory and inhibitory populations as a result of training: the choice signal emerged at the same time in both populations, and its magnitude and timing was matched for the two cell types throughout learning (Figure 7A-C, middle, right; Figure 7D-E, blue, red). These finding indicate that learning induces the simultaneous emergence of choice-specific subpopulations in PPC.

Notably, the animal's licking or running behavior could not explain the learning-induced changes in the magnitude of classification accuracy (Figure S8). The center-spout licks that preceded the left vs. right choices were overall similar during the course of learning (Figure S8A), and did not differ in early vs. late training days (Figure S8B).

The similarity in lick movements for early vs. late sessions stands in contrast to the changes in the choice classifier for early vs. late sessions (Figure 7). We also assessed animals' running behavior during the course of learning (Figure S8C,D). In some sessions, the running distance differed preceding left vs. right choices (Figure S8C). Nonetheless, when we restricted our analysis to days in which the running distance was indistinguishable for the two choices (0-97

ms before the choice, t-test, P>0.05), we were able to accurately classify the animal's choice using neural activity (Figure S8D). These observations provide reassurance that the population activity does not entirely reflect preparation of licking and running movements and argues instead that the population activity reflects the animal's stimulus-informed choice.

Finally, we studied how cofluctuations changed over the course of training. Pairwise

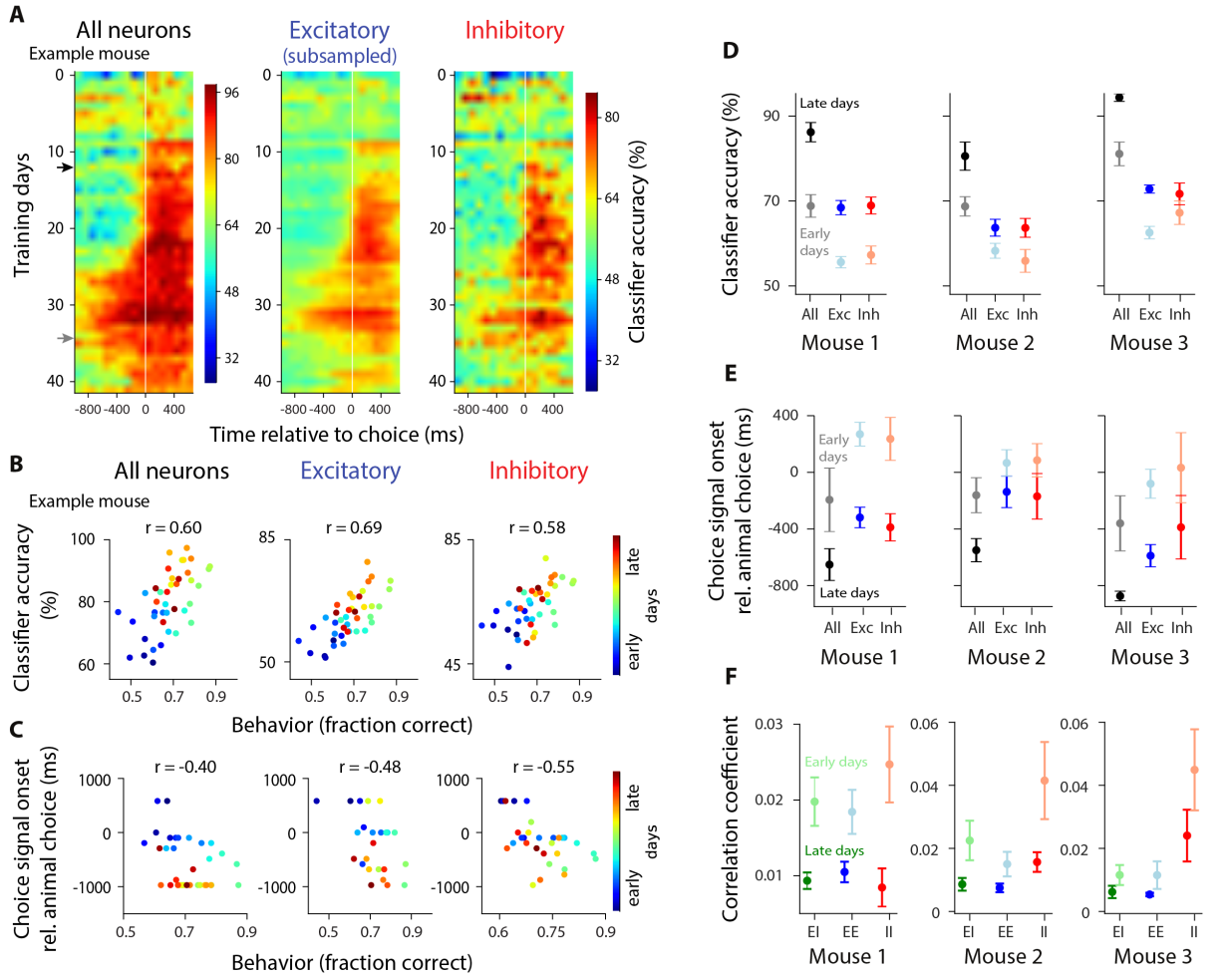

Figure 7. Learning leads to increased magnitude and advanced timing of the choice signal, as well as reduced noise correlations, in both excitatory and inhibitory populations.
A, Decoder accuracy is shown for each training session, for all neurons (left), subsampled excitatory (middle), and inhibitory neurons (right). White vertical line: choice onset. Each row: average across cross-validation samples; example mouse. Colorbar of the inhibitory plot applies to the excitatory plot too. **B,** Scatter plot of classifier accuracy at 0-97 ms before the choice vs. behavioral performance (fraction correct on easy trials), including all training days. r is Pearson correlation coefficient (p<0.001 in all plots); same example mouse as in (A). Correlations for behavior vs. classification accuracy for all neurons, excitatory and inhibitory: 0.55, 0.35, 0.32 in mouse 2; 0.57, 0.63, 0.32 in mouse 3. Correlations for behavior vs. choice-signal onset for all neurons, excitatory and inhibitory: -0.60, -0.34, -0.38, in mouse 2; -0.60, -0.27, -0.28 in mouse 3. All values: p<0.05 **C,** Same as (B), except showing the onset of choice signal, i.e. the first moment in the trial that classifier accuracy was above chance (ms, relative to choice onset) vs. behavioral performance. **D,** Summary of each mouse, showing classification accuracy averaged across early (dim colors) vs. late (dark colors) training days. **E,** Same as (D), but showing choice signal onset (ms). **F,** Same as (D), but showing pairwise noise correlation coefficients. Early days were the first few training days in which the animal's performance was lower than the 20th percentile of animal's performance across all days. Late days included the last training days in which the animal's behavioral performance was above the 80th percentile of performance across all days.

correlations in neural activity were overall higher in early training days, when mice were novices, compared to late training days, as they approached expert behavior (Figure 7F, unsaturated colors above saturated colors). This effect was observed for all combinations of neural pairs (Figure 7F, green: excitatory-inhibitory; blue: excitatory-excitatory; red: inhibitory-inhibitory). These findings are in agreement with previous reports suggesting that learning results in reduced noise correlations (Gu et al., 2011; Jeanne et al., 2013; Khan et al., 2018; Ni et al., 2018), enhancing information that is encoded in neural populations (Zohary et al., 1994). To test if the learning-induced increase in classification accuracy (Figure 7A,B,D) was entirely a consequence of the reduction in noise correlations (Figure 7F), we studied how classification accuracy of pseudo populations (Figure 6A,B), which lack noise correlations, changed with training. Interestingly, we still observed a significant increase in the classification accuracy of pseudo populations as a result of training (Figure S9A,B). Therefore, the reduction in noise correlations cannot alone account for the improved classification accuracy that occurs during learning. Instead, it suggests that choice selectivity of individual neurons changes with learning. Indeed, the fraction of choice-selective neurons increased ~3-fold, in both excitatory and inhibitory cell types, as a result of training (Figure S9C), contributing to the improved classification accuracy at the ensemble level.

Discussion

Despite a wealth of studies assessing selectivity of inhibitory neurons in response to sensory features, little is known about the specificity of inhibitory neurons in decision-making. This represents a critical gap in knowledge, and has left untested key features of decision-making models relying on inhibitory neurons. To close this gap, we simultaneously measured excitatory and inhibitory populations during perceptual decisions about multisensory stimuli.

We demonstrated that excitatory and inhibitory neurons predict the animal's impending choice with equal fidelity. This result, along with simulations and analysis, suggests that specific functional subnetworks exist within inhibitory populations, just like excitatory populations. This implies targeted connectivity between excitatory and inhibitory neurons (Lim and Goldman, 2013) and argues against models of decision-making in which inhibitory neurons receive completely nonspecific input from excitatory populations (Figure 4A). Our findings are in support of an alternative circuit architecture with functionally specific subnetworks within excitatory and inhibitory populations that are selective and reciprocally connected (Figure 7). This architecture is supported on theoretical grounds because of stability (Lim and Goldman, 2013; Znamenskiy et al., 2018), and has experimental support in the rodent primary visual area (Yoshimura and Callaway, 2005; Znamenskiy et al., 2018), the primate prefrontal cortex (Rao et al., 1999), and the goldfish oculomotor integrator (Aksay et al., 2007).

The equal selectivity for choice that we observed in excitatory and inhibitory populations is surprising given that the majority of studies have reported weaker stimulus tuning (Sohya et al., 2007; Niell and Stryker, 2008; Kerlin et al., 2010; Hofer et al., 2011; Znamenskiy et al., 2018)(but see (Runyan et al., 2010)) and dense connectivity (Hofer et al., 2011; Packer and Yuste, 2011; Znamenskiy et al., 2018) for interneurons. Two differences between our study and previous ones may explain why we saw equal selectivity in excitatory and inhibitory populations.

First, we measured neural activity in PPC where the proportion of interneuron subtypes differ from V1; in particular, early sensory areas are more enriched in PV interneurons relative to SOM and VIP neurons, whereas it's the opposite in association areas (Kim et al., 2017; Wang and Yang, 2018). Moreover, interneuron subtypes vary in their specificity of connections (Pfeffer et al., 2013); for instance, PV interneurons are suggested to have broader tuning than SOM and VIP cells (Wang et al., 2004). Therefore, the strong selectivity that we found in all GABAergic interneurons in PPC may not contradict the broad selectivity observed in studies largely performed on PV interneurons in V1. Future studies that measure the selectivity of distinct interneuron populations during decision-making in V1 vs. PPC will be helpful. Here, we measured all GABAergic interneurons instead of individual interneuron subtypes; this was because of the technical challenges in reliably identifying more than two cell types in a single animal, and because of the importance of simultaneously measuring the activity of excitatory and inhibitory neurons within the same subject. Had we lacked within-animal measurements, our ability to compare excitatory vs. inhibitory neurons would have been compromised by animal-to-animal variability.

Second, analyzing neural activity in the context of decision-making naturally led us to make different comparisons than those carried out in previous work. For example, we measured selectivity for a binary choice instead of tuning for a continuous sensory stimulus (e.g., orientation). Further, we measured activity in response to an abstract stimulus, the meaning of which was learned gradually by the animal. This may recruit circuits that differ from those supporting sensory processing in naive mice. Finally, we used stochastically fluctuating multisensory stimuli, which have not been evaluated in mouse V1. Future studies that examine the tuning of V1 neurons to the sensory stimulus used here will determine if V1 inhibitory neurons will be as sharply tuned as excitatory neurons to the stimulus. This is a possibility: the tuning strength of interneurons can vary substantially for different stimulus features. For instance, PV neurons in V1 have particularly poor tuning to the orientation of visual stimuli, while their temporal-frequency tuning is considerably stronger (Znamenskiy et al., 2018).

Our long-term monitoring of neural activity within the same subjects provide a critical new insight into decision-making circuitry by demonstrating how acquiring expertise modulates the activity of excitatory and inhibitory neurons in PPC. We observed that learning induced an increase in the number of choice-selective neurons and a decrease in noise correlations, indicating plasticity and reorganization of connections. As a result, population responses preceding the two choices became progressively more distinct with training. Importantly, these changes occurred in parallel in both excitatory and inhibitory cells. Our findings are partially in agreement with those in V1, in which learning improves tuning to sensory stimuli in excitatory (Schoups et al., 2001; Poort et al., 2015; Khan et al., 2018) and some inhibitory subtypes (Khan et al., 2018). However, in V1 excitatory neurons have stronger tuning to sensory stimuli early in training (Khan et al., 2018); in contrast, the magnitude of choice selectivity in PPC was the same for both cell types throughout training in our study (Figure 6). Primate studies have likewise observed that perceptual learning changes the selectivity of neurons (Freedman and Assad, 2006; Law and Gold, 2008; Viswanathan and Nieder, 2015) and reduces noise correlations (Gu et al., 2011; Ni et al., 2018).

Finally, we demonstrated that the learning-induced changes in PPC selectivity were closely associated with the changes in animal performance, in keeping with primate studies of decision-making (Law and Gold, 2008). This, together with our finding that population activity does not purely reflect movements (Figure S8), further corroborates the suggested role for PPC in mapping sensation to action (Law and Gold, 2008; Raposo et al., 2014; Pho et al., 2018). Future experiments using causal manipulations will reveal whether the increased choice selectivity we observed in PPC originates there or is inherited from elsewhere in the brain.

By simultaneously measuring neural activity in parietal cortex we have provided evidence that excitatory and inhibitory populations are equally choice-selective, and that these ensembles emerge in parallel as mice become skilled decision-makers. The existence of functional subnetworks of inhibitory neurons in parietal cortex argues against non-specific connections with excitatory neurons, at least in decision circuits. Moving forward, these features can be incorporated into decision-making models, and their impact on key model outputs, like reaction time distributions and firing rate predictions, can be evaluated. Future studies will shed light on how microcircuits of inhibitory and excitatory neurons may vary across areas in their selectivity and specificity of connections, and will reveal the circuit architectures that allow for equally selective inhibitory and excitatory neurons.

Methods

Imaging and behavioral dataset

Our simultaneous imaging and decision-making dataset includes 135 sessions from 4 mice (45, 48, 7, and 35 sessions per mouse). Median number of trials per session is 213, 253, 264, and 222, for each mouse. On average, 480 neurons were imaged per session, out of which ~40 neurons were inhibitory and ~330 were excitatory. Approximately 100 neurons per session were not classified as either excitatory or inhibitory since they did not meet our strict cell-type classification criteria (see below). In 3 of the mice, the same group of neurons was imaged throughout learning (35-48 training days).

Mice and surgical procedure

Gad2-IRES-CRE(Taniguchi et al., 2011) mice were crossed with Rosa-CAG-LSL-tdTomato-WPRE (aka Ai14(Madisen et al., 2010)) to create mice in which all GABAergic inhibitory neurons were labeled. Adult mice (~2-month old) were used in the experiments. Meloxicam (analgesic), dexamethasone (anti-inflammatory) and Baytril (enrofloxacin; anti-biotic) were injected 30min before surgery. Using a biopsy punch, a circular craniotomy (diameter: 3mm) was made over the left PPC (stereotaxic coordinates: 2 mm posterior, 1.7 mm lateral of bregma(Harvey et al., 2012) under isoflurane (~5%) anesthesia. Pipettes (10-20 um in diameter, cut at an angle to provide a beveled tip) were front-filled with AAV9-Synapsin-GCaMP6f (U Penn, Vector Core Facility) diluted 2X in PBS (Phosphae-buffered saline). The pipette was slowly advanced into the brain (Narishige MO-8 hydraulic micro-manipulator) to make ~3 injections of 50nL, slowly over an interval of ~5-10min, by applying air pressure using a syringe. Injections ware made near the center of craniotomy at a depth of 250-350um below the dura. A glass plug consisting of a 5mm coverslip attached to a 3mm coverslip (using IR-curable optical bond, Norland) was used to cover the craniotomy window. Vetbond, followed by metabond, was

used to seal the window. All surgical and behavioral procedures conformed to the guidelines established by the National Institutes of Health and were approved by the Institutional Animal Care and Use Committee of Cold Spring Harbor Laboratory.

Imaging

We used a 2-photon Moveable Objective Microscope with resonant scanning at approximately 30 frames per second (Sutter Instruments, San Francisco, CA). A 16X, 0.8 NA Nikon objective lens was used to focus light on fields of view of size 512x512 pixels (~575 µm x ~575 µm). A Ti:sapphire laser (Coherent) delivered excitation light at 930nm (average power: 20-70 mW). Red (ET670/50m) and green (ET 525/50m) filters (Chroma Technologies) were used to collect red and green emission light. The microscope was controlled by Mscan (Sutter). In mice in which chronic imaging was performed during learning, the same plane was identified on consecutive days using both coarse alignment, based on superficial blood vessels, as well as fine alignment, using reference images of the red channel (tdTomato expression channel) at multiple magnification levels. For each trial, imaging was started 500ms before the trial-initiation tone, and continued 500ms after reward or time-out.

Decision-making behavior

Mice were gradually water restricted over the course of a week, and were weighed daily. Mice harvested at least 1 mL of water per behavioral/imaging session, and completed 100-500 trials per session. After approximately one week of habituation to the behavioral setup, 15-30 training days were required to achieve 75% correct choice. Animal training took place in a sound isolation chamber. The stimulus in all trials was multisensory, consisting of a series of simultaneous auditory clicks and visual flashes, occurring with Poisson statistics(Brunton et al., 2013; Odoemene et al., 2017). Stimulus duration was 1000ms. Each pulse was 5ms; minimum interval between pulses was 32ms, and maximum interval was 250ms. The pulse rate ranged from 5 to 27 Hz. The category boundary for marking high-rate and low-rate stimuli was 16Hz, at which animals were rewarded randomly on either side. The highest stimulus rates used here are known to elicit reliable, steady state flicker responses in retinal ERG in mice(Krishna et al., 2002; Tanimoto et al., 2015).

Mice were on top of a cylindrical wheel and a rotary encoder was used to measure their running speed. Trials started with a 50ms initiation tone (Figure S1A). Mice had 5sec to initiate a trial by licking the center waterspout, after which the multisensory stimulus was played for 1 second. If mice again licked the center waterspout, they received 0.5µL water on the center spout, and a 50ms go cue was immediately played. Animals had to report a choice by licking to the left or right waterspout within 2sec. Mice were required to confirm their choice by licking the same waterspout one more time within 300ms after the initial lick. If the choice was correct, mice received 2-4µL water on the corresponding waterspout. An incorrect choice was punished with a 2sec time-out. The experimenter-imposed inter-trial intervals (ITI) were drawn from a truncated exponential distribution, with minimum, maximum, and lambda equal to 1sec, 5sec, and 0.3sec, respectively. However, the actual ITIs could be much longer depending on when the animal initiates the next trial. Bcontrol (Raposo et al., 2014) with a Matlab interface was used to deliver trial events (stimulus, reward, etc) and collect data.

Logistic regression model of behavior

A modified version of the logistic regression model in(Busse et al., 2011) was used to assess the extent to which the animal's choice depends on the strength of sensory evidence, i.e. how far the stimulus rate is from the category boundary at 16Hz, the previous choice outcome (success or failure) and ITI, i.e. the time interval between the previous choice and the current stimulus onset (Figure S1B).

$$p = \frac{1}{1+e^{-z}} \qquad \textit{eq. 1}$$

$$z = \beta_0 + (\beta_{r1} R_1 + \beta_{r2} R_2 + \beta_{r3} R_3 + \beta_{r4} R_4 + \beta_{r5} R_5 + \beta_{r6} R_6) + (\beta_{s1} S_1 + \beta_{s2} S_2) + (\beta_{f1} F_1 + \beta_{f2} F_2)$$

where p is the probability of choosing the left choice, and z is the decision variable. R, S and F are vectors of indicator variables; each element corresponds to 1 trial. Stimulus strength (R) was divided into 6 bins (R_1 to R_6). Previous success (S) was divided into 2 bins (S_1 to S_2): success after a long ITI (> 7sec) and success after a short ITI (< 7sec). Previous failure (F) was divided into 2 bins (F_1 to F_2): failure after a long and short ITI. For instance, if a trial had stimulus strength 3Hz, and was preceded by a success choice with ITI 5sec, then R_2 and S_1 would be set to 1 and all other R, S and F parameters to 0 (Figure S1B).

For each session the scalar coefficients β_0, β_{r1} to β_{r6}, β_{s1}, β_{s2}, β_{f1}, and β_{f2} were fitted using Matlab glmfit.m. Figure S1B left shows β_{r1} to β_{r6}. Figure S1B middle shows β_{s1} and β_{s2}, and Figure S1B right shows β_{f1} and β_{f2}.

ROI (region of interest) extraction and deconvolution

The recorded movies from all trials were concatenated and corrected for motion artifacts by cross-correlation using DFT registration (Guizar-Sicairos et al., 2008). Subsequently, active ROIs (sources) were extracted using the CNMF algorithm(Pnevmatikakis et al., 2016) as implemented in the CaImAn package (Giovannucci et al., 2018) in MATLAB. The traces of the identified neurons were ΔF/F normalized and then deconvolved by adapting the FOOPSI deconvolution algorithm (Vogelstein et al., 2010; Pnevmatikakis et al., 2016) to a multi-trial setup. Each value of Foopsi deconvolution represents spiking activity at each frame for a given neuron. We have referred to the deconvolved values as "inferred spiking activity" throughout the paper. The deconvolved values do not represent absolute firing rates, so they cannot be compared across neurons. However, for a particular neuron, higher inferred spiking activity means higher firing rate.

The adaptation of the FOOPSI for multi-trial setup involved the following steps. For each component, the activity trace over all the trials was used to determine the time constants of the calcium indicator dynamics as in (Pnevmatikakis et al., 2016). Then the neural activity during each trial was deconvolved separately using the estimated time constant and a zero baseline (since the traces were ΔF/F normalized). A difference of exponentials was used to simulate the rise and decay of the indicator.

Neuropil Contamination removal

The CNMF algorithm demixes the activity of overlapping neurons. It takes into account background neuropil activity by modeling it as a low rank spatiotemporal matrix (Pnevmatikakis

et al., 2016). In this study a rank two matrix was used to capture the neuropil activity. To evaluate its efficacy we compared the traces obtained from CNMF to the traces from a "manual" method similar to(Chen et al., 2013) (Figure S10): the set of spatial footprints (shapes) extracted from CNMF were binarized by thresholding each component at the 0.2x its maximum value level. The binary masks were then used to average the raw data and obtain an activity trace that also included neuropil effects. To estimate the background signal, an annulus around the binary mask was constructed with minimum distance 3 pixels from the binary mask and width 7 pixels (Figure S10A). The average of the raw data over the annulus defined the background trace, which was then subtracted from the activity trace. The resulting trace was then compared with the CNMF estimated temporal trace for this activity. The comparison showed a very high degree of similarity between the two traces (Figure S10; example component; r=0.96), with the differences between the components being attributed to noise and not neuropil related events. Note that this "manual" approach is only applicable in the case when the annulus does not overlap with any other detected sources. These results demonstrate the ability of the CNMF framework to properly capture neuropil contamination and remove it from the detected calcium traces.

ROI inclusion criteria

We excluded poor-quality ROIs identified by the CNMF algorithm based on a combination of criteria: 1) size of the spatial component, 2) decay time constant, 3) correlation of the spatial component with the raw ROI image built by averaging spiking frames, 4) correlation of the temporal component with the raw activity trace, and 5) the probability of fluorescence traces maintaining values above an estimated signal-to-noise level for the expected duration of a calcium transient(Giovannucci et al., 2018) (GCaMP6f, frame rate: 30Hz). A final manual inspection was performed on the selected ROIs to validate their shape and trace quality.

Identification of inhibitory neurons

We used a two-step method to identify inhibitory neurons. First, we corrected for bleed-through from green to red channel by considering the following regression model,

$$\boldsymbol{r}_i = \beta_i 1_p + s\boldsymbol{g}_i + \epsilon \qquad \textit{eq. 2}$$

where i denotes ROI, $\boldsymbol{r}_i \in \mathbb{R}^p$ and $\boldsymbol{g}_i \in \mathbb{R}^p$ are vectors with each element corresponding to the intensity of one pixel averaged across time for the red and green channels, respectively $\forall i \in \{1,..,N\}$ (N: total number of ROIs; p: the number of pixels in ROI i). $\beta_i \in \mathbb{R}$ is an offset, and s is the parameter that tells us how much of the green channel bleeds through to the red one. $1_p \in \mathbb{R}^p$ is a vector whose components are all 1.

It is the parameter s that we are interested in. To find s, we define a cost function, C,

$$\boldsymbol{C} = \sum_i \left|\boldsymbol{r}_i - \beta_i 1_p + s\boldsymbol{g}_i\right|_2^2 \qquad \textit{eq. 3}$$

and minimize it with respect to s and all the β_i. The value of s at the minimum reflects the fraction of bleed-through from the green to the red channel. That value, denoted s*, is then used to compute the bleedthrough-corrected image of the red-channel, denoted I via the expression

565 $I = R - s^* G$ *eq. 4*

566 where R and G are the time-averaged images of the red and green channels, respectively.

567 Once the bleedthrough-corrected image, I, was computed, we used it to identify inhibitory
568 neurons using two measures:

569 1) A measure of local contrast, by computing, on the red channel, the average pixel intensity
570 inside each ROI mask relative to its immediate surrounding mask (width=3 pixels). Given the
571 distribution of contrast levels, we used two threshold levels, T_E and T_I, defined, respectively, as
572 the 80th and 90th percentiles of the local contrast measures of all ROIs. ROIs whose contrast
573 measure fell above T_I were identified as inhibitory neurons. ROIs whose contrast measure fell
574 below $T_E$were identified as excitatory neurons, and ROIs with the contrast measure in between
575 T_E and T_I were not classified as either group ("unsure" class).

576 2) In addition to a measure of local contrast, we computed for each ROI the correlation between
577 the spatial component (ROI image on the green channel) and the corresponding area on the red
578 channel. High correlation values indicate that the ROI on the green channel has a high signal on
579 the red channel too; hence the ROI is an inhibitory neuron. We used this correlation measure to
580 further refine the neuron classes computed from the local contrast measure (i.e. measure 1
581 above). ROIs that were identified as inhibitory based on their local contrast (measure 1) but had
582 low red-green channel correlation (measure 2), were reset as "unsure" neurons. Similarly, ROIs
583 that were classified as excitatory (based on their local contrast) but had high red-green channel
584 correlation were reclassified as unsure. Unsure ROIs were included in the analysis of all-neuron
585 populations (Figure 3A left); but were excluded from the analysis of excitatory only or inhibitory
586 only populations (Figure 3A middle, right). Finally, we manually inspected the ROIs identified
587 as inhibitory to confirm their validity. This method resulted in 8.3% inhibitory neurons, which is
588 in reasonable agreement with previous studies, although slightly lower than some reports(Sahara
589 et al., 2012). This was expected given our strict method for identifying inhibitory neurons.

590 **General analysis procedures**

591 All analyses were performed on inferred spiking activity. Traces were down-sampled, so each
592 bin was the non-overlapping moving average of 3 frames (97.1 ms). Inferred spiking activity for
593 each neuron was normalized so the max spiking activity for each neuron equaled 1. The trace of
594 each trial was aligned to the time of the choice (i.e. the time of the 1st lick to either of the side
595 waterspouts after the go tone). Two-tailed t-test was performed for testing statistical significance.
596 Summary figures including all mice were performed on the time bin preceding the choice, i.e. 0-
597 97 ms before choice onset. All reported correlations are Pearson's coefficients. Analyses were
598 performed in Python and Matlab.

599 **ROC analysis**

600 The area under the ROC curve (AUC) was used to measure the choice preference of single
601 neurons. Choice selectivity was defined as the absolute deviation of AUC from chance level
602 (0.5). To identify significantly choice-selective neurons, for each neuron we performed ROC on
603 shuffled trial labels (i.e. left and right choices were randomly assigned to each trial). This
604 procedure was repeated 50 times to create a distribution of shuffled AUC values for each neuron

605 (Figure S2A, “shuffled”). A neuron’s choice selectivity was considered to be significant if the
606 probability of the actual AUC (Figure S2A, “real”) being drawn from the shuffled AUC
607 distribution was less than 0.05. Time points from 0–97 ms before the decision were used to
608 compute the fraction of choice-selective neurons (Figure 2B; Figure S9C).

609 **Decoding population activity**

610 A linear SVM (Python sklearn package) was trained on each bin of the population activity in
611 each session (non-overlapping 97ms time bins). To break any dependencies on the sequence of
612 trials, we shuffled the order of trials for the entire population. To avoid bias in favor of one
613 choice over the other, we matched the number of left- and right-choice trials used for classifier
614 training. L2 regularization was used to avoid over-fitting. 10-fold cross validation was performed
615 by leaving out a random 10% subset of trials to test the classifier performance, and using the
616 remaining trials for training the classifier. This procedure was repeated 50 times. A range of
617 regularization values was tested, and the one that gave the smallest error on the validation dataset
618 was chosen as the optimal regularization parameter. Classifier accuracy was computed as the
619 percentage of testing trials in which the animal’s choice was accurately predicted by the
620 classifier, and summarized as the average across the 50 repetitions of trial subsampling. A
621 minimum of 10 correct trials per choice was required in order to run the SVM on a session.
622 Inferred spiking activity of each neuron was z-scored before running the SVM.

623 When comparing classification accuracy for excitatory vs. inhibitory neurons, the excitatory
624 population was randomly sub-sampled to match the population size of inhibitory neurons to
625 enable a fair comparison (Figure 3, blue vs. red). To compare the distribution of weights in the
626 all-neuron classifier (Figure 3 black), the weight vector for each session was normalized to unity
627 length (Figure 3D-F).

628 **Stability**

629 To test the stability of the population code, decoders were trained and tested at different time
630 bins (Kimmel et al., 2016) (Figure 5). To avoid the potential effects of auto-correlation, we
631 performed cross validation not only across time bins, but also across trials. In other words, even
632 though the procedure was cross validated by testing the classifier at a time different from the
633 training time, we added another level of cross-validation by testing on a subset of trials that were
634 not used for training. This strict method allowed our measure of stability duration to be free of
635 auto-correlation effects.

636 As an alternative measure of stability, the angle between pairs of classifiers that were trained at
637 different moments in the trial was computed (Figure S5C). Small angles indicate alignment,
638 hence stability, of the classifiers. Large angles indicate misalignment, i.e. instability of the
639 classifiers.

640 **Noise correlations**

641 To estimate noise correlations at the population level, the order of trials was shuffled for each
642 neuron independently. Shuffling was done within the trials of each choice, hence retaining the
643 choice signal, while de-correlating the population activity to remove noise correlations. Then we
644 classified population activity in advance of left vs. right choice (at time bin 0–97 ms before the

645 choice) using the de-correlated population activity. This procedure was performed on neural
646 ensembles of increasingly larger size, with the most selective neurons (|AUC-0.5|) added first
647 (Figure 6A). To summarize how noise correlations affected classification accuracy in the
648 population (Figure 6B), we computed, for the largest neural ensemble (Figure 6A, max value on
649 the horizontal axis), the change in classifier accuracy in the de-correlated data ("pseudo
650 populations") vs. the original data. This analysis was only performed for the entire population;
651 the small number of inhibitory neurons in each session prevented a meaningful comparison of
652 classification accuracy on real vs. pseudo populations.

653 To measure pairwise noise correlations, we subtracted the trial-averaged response to a particular
654 choice from the response of single trials of that choice. This allowed removing the effect of
655 choice on neural responses. The remaining variability in trial-by-trial responses can be attributed
656 to noise correlations, measured as the Pearson correlation coefficient for neuron pairs. We also
657 measured noise correlations using the spontaneous activity defined as the neural responses in 0-
658 97 ms preceding the trial initiation tone (Figure S2F; Figure S6B). We computed the pairwise
659 correlation coefficient (Pearson) for a given neuron with each other neuron within an ensemble
660 (e.g., excitatory neurons). The resulting coefficients were then averaged to generate a single
661 correlation value for that neuron. This was repeated for all neurons within the ensemble (Figure
662 6C).

663 To compute pairwise correlations on excitatory and inhibitory neurons with the same inferred
664 spiking activity (Figure S6C,D), we computed the median inferred spiking activity across trials
665 for individual excitatory and inhibitory neurons in a session. The medians were then divided into
666 50 bins. The firing-rate bin that included the maximum number of inhibitory neurons was
667 identified ("max bin"); inhibitory and excitatory neurons whose firing rate was within this "max
668 bin" were used for the analysis. The firing rates were matched for these neurons because their
669 median firing rate was within the same small bin of firing rates. Pairwise correlations were then
670 computed as above.

671 **Learning analysis**

672 In 3 of the mice, the same field of view was imaged each session during learning. This was
673 achieved in two ways. First, the vasculature allowed a coarse alignment of the imaging location
674 from day to day. Second, the image from the red channel was used for a finer alignment. Overall,
675 most neurons were stably present across sessions (Figure S7). This suggests that we likely
676 measured activity from a very similar population each day. Importantly, however, our
677 conclusions do not rely on this assumption: our measures and findings focus on learning-related
678 changes in the PPC population overall, as opposed to tracking changes in single neurons. To
679 assess how population activity changed over learning, we evaluated classification accuracy each
680 day, training a new decoder for each session. This approach allowed us to compute the best
681 decoding accuracy for each session.

682 "Early days" (Figure 7; Figure S9) included the initial training days in which the animal's
683 performance, defined as the fraction of correct choices on easy trials, was lower than the 20th
684 percentile of performance across all days. "Late days" (Figure 7; Figure S9) included the last
685 training days in which the animal's behavioral performance was above the 80th percentile of
686 performance across all days.

To measure the timing of decision-related activity (Figure 7C,E), we identified all sessions in which classifier accuracy was significantly different than the shuffle (t-test, $p<0.05$) over a window of significance that was at least 500 ms long. We defined the "choice signal onset" (Figure 7C,E) as the trial time corresponding to the first moment of that window. Sessions in which the 500 ms window of significance was present are included in Figure 7C. The number of points (and hence the relationship between session number and color in Figure 7C) differs slightly across the three groups. This is because on some sessions, the window of significance was present in one group but not another. For example, in a session the population including all neurons might have a 500 ms window of significance, hence it will contribute a point to Figure 7C left, while the population with only inhibitory neurons might be only transiently significant for <500ms, hence it will be absent from Figure 7C right.

Circuit analysis

We considered a linearized rate network of the form

$$\frac{d\mathbf{v}_E}{dt} = -\mathbf{v}_E + \mathbf{W}_{EE} \cdot \mathbf{v}_E - \mathbf{W}_{EI} \cdot \mathbf{v}_I + \mathbf{h}_s + \boldsymbol{\xi}_E$$
$$\frac{d\mathbf{v}_I}{dt} = -\mathbf{v}_I + \mathbf{W}_{IE} \cdot \mathbf{v}_E - \mathbf{W}_{II} \cdot \mathbf{v}_I + \boldsymbol{\xi}_I$$

where E and I refer to the excitatory and inhibitory populations, respectively, $\mathbf{v}_E$ and $\mathbf{v}_I$ are vectors of firing rates ($\mathbf{v}_E = \nu_{E1}, \nu_{e2}, \ldots$, and similarly for $\mathbf{v}_I$), $\mathbf{W}_{\boldsymbol{EE}}$, $\mathbf{W}_{\boldsymbol{EI}}$, $\mathbf{W}_{\boldsymbol{IE}}$ and $\mathbf{W}_{\boldsymbol{II}}$ are the connectivity matrices ($\mathbf{W}_{\boldsymbol{EI}}$ indicates connection from inhibitory to excitatory neuron). $\mathbf{h}_s$ is the input, with s either 1 or 2 (corresponding to left and right licks), and $\boldsymbol{\xi}$ is trial to trial noise, taken to be zero mean and Gaussian, with covariance matrices

$$\langle \boldsymbol{\xi}_E \boldsymbol{\xi}_E \rangle = \boldsymbol{\Sigma}_{EE}$$
$$\langle \boldsymbol{\xi}_I \boldsymbol{\xi}_I \rangle = \boldsymbol{\Sigma}_{II}.$$

For the input we'll assume that about half the elements of h_s are h_0 for the rightward choice and $-h_0$for the leftward choice, and the rest are $-h_0$ for the rightward choice and h_0 for the leftward choice. We used $h_0 = 0.1$ (see Table 1). The noise covariance is diagonal but non-identity, with diagonal elements distributed as

$$\sqrt{\Sigma_{EE,ii}} \sim \mathrm{Unif}\left(\sigma - \frac{\delta}{2}, \sigma + \frac{\delta}{2}\right)$$
$$\sqrt{\Sigma_{II,ii}} \sim \mathrm{Unif}\left(\sigma - \frac{\delta}{2}, \sigma + \frac{\delta}{2}\right).$$

The goal is to determine the value of s (that is, determine whether $\mathbf{h}_1$ or $\mathbf{h}_2$ was present) given the activity of a subset of the neurons from either the excitatory or inhibitory populations. We'll work in steady state, for which

$$\boldsymbol{\nu}_E = \mathbf{W}_{EE} \cdot \boldsymbol{\nu}_E - \mathbf{W}_{EI} \cdot \boldsymbol{\nu}_I + \mathbf{h}_s + \boldsymbol{\xi}_x$$
$$\boldsymbol{\nu}_I = \mathbf{W}_{IE} \cdot \boldsymbol{\nu}_E - \mathbf{W}_{II} \cdot \boldsymbol{\nu}_I + \boldsymbol{\xi}_y \,.$$

Solving for $\boldsymbol{\nu}_E$ and $\boldsymbol{\nu}_I$ yields

$$\boldsymbol{\nu}_E = \mathbf{J}_E \cdot (\mathbf{h}_c + \boldsymbol{\xi}_E - \widetilde{\mathbf{W}}_{EI} \cdot \boldsymbol{\xi}_I)$$
$$\boldsymbol{\nu}_I = \mathbf{J}_I \cdot (\boldsymbol{\xi}_I + \widetilde{\mathbf{W}}_{IE}(\mathbf{h}_c + \boldsymbol{\xi}_E)$$

where

$$\mathbf{J}_E \equiv (\mathbf{I} - \mathbf{W}_{EE} + \widetilde{\mathbf{W}}_{EI} \cdot \mathbf{W}_{IE})^{-1}$$
$$\mathbf{J}_I \equiv (\mathbf{I} + \mathbf{W}_{II} + \widetilde{\mathbf{W}}_{IE} \cdot \mathbf{W}_{EI})^{-1}$$
$$\widetilde{\mathbf{W}}_{EI} \equiv \mathbf{W}_{EI}(\mathbf{I} + \mathbf{W}_{II})^{-1}$$
$$\widetilde{\mathbf{W}}_{IE} \equiv \mathbf{W}_{IE}(\mathbf{I} - \mathbf{W}_{EE})^{-1} \,,$$

and $\mathbf{I}$ is the identity matrix. We are interested in the decoding accuracy of a sub- population of neurons. For that we'll use a matrix $\mathbf{D}_n$ that picks out n components of whatever it's operating on. So, for instance, $\mathbf{D}_n \cdot \boldsymbol{\nu}_E$ is an n-dimensional vector with components equal to n of the components of $\boldsymbol{\nu}_E$.

For a linear and Gaussian model such as ours, in which the covariance is independent of s, we need two quantities to compute the performance of an optimal decoder: the difference in the means of the subsampled populations when $\mathbf{h}_1$ versus $\mathbf{h}_2$ are present, and covariance matrix of the subsampled populations. The difference in means are given by

$$\Delta\langle \mathbf{D}_n \cdot \boldsymbol{\nu}_E \rangle = \mathbf{D}_n \cdot \mathbf{J}_E \cdot \Delta\mathbf{h}$$
$$\Delta\langle \mathbf{D}_n \cdot \boldsymbol{\nu}_I \rangle = \mathbf{D}_n \cdot \mathbf{J}_I \cdot \widetilde{\mathbf{W}}_{IE} \cdot \Delta\mathbf{h}$$

where $\Delta\mathbf{h}$ is the difference between the two inputs,

$$\Delta\mathbf{h} \equiv \mathbf{h}_1 - \mathbf{h}_2 \,.$$

The covariances are given by

$$\mathrm{Cov}[\mathbf{D}_n \cdot \boldsymbol{\nu}_E] = \mathbf{D}_n \cdot \mathbf{J}_E \cdot \left[\boldsymbol{\Sigma}_{EE} + \widetilde{\mathbf{W}}_{EI} \cdot \boldsymbol{\Sigma}_{II} \cdot \widetilde{\mathbf{W}}_{EI}^T\right] \cdot \mathbf{J}_E^T \cdot \mathbf{D}_n^T$$
$$\mathrm{Cov}[\mathbf{D}_n \cdot \boldsymbol{\nu}_I] = \mathbf{D}_n \cdot \mathbf{J}_I \cdot \left[\boldsymbol{\Sigma}_{II} + \widetilde{\mathbf{W}}_{IE} \cdot \boldsymbol{\Sigma}_{EE} \cdot \widetilde{\mathbf{W}}_{IE}^T\right] \cdot \mathbf{J}_I^T \cdot \mathbf{D}_n^T$$

where T denotes transpose. Combining the mean and covariance gives us the signal to noise ratio,

$$(S/N)_E = \Delta\mathbf{h} \cdot \mathbf{J}_E^T \cdot \mathbf{D}_n^T \cdot \left(\mathbf{D}_n \cdot \mathbf{J}_E \cdot \left[\boldsymbol{\Sigma}_{EE} + \widetilde{\mathbf{W}}_{EI} \cdot \boldsymbol{\Sigma}_{II} \cdot \widetilde{\mathbf{W}}_{EI}^T\right] \cdot \mathbf{J}_E^T \cdot \mathbf{D}_n^T\right)^{-1} \cdot \mathbf{D}_n \cdot \mathbf{J}_E \cdot \Delta\mathbf{h}$$
$$(S/N)_I = \Delta\mathbf{h} \cdot \widetilde{\mathbf{W}}_{IE}^T \cdot \mathbf{J}_I^T \cdot \mathbf{D}_n^T \cdot \left(\mathbf{D}_n \cdot \mathbf{J}_I \cdot \left[\boldsymbol{\Sigma}_{II} + \widetilde{\mathbf{W}}_{IE} \cdot \boldsymbol{\Sigma}_{EE} \cdot \widetilde{\mathbf{W}}_{IE}^T\right] \cdot \mathbf{J}_I^T \cdot \mathbf{D}_n^T\right)^{-1} \cdot \mathbf{D}_n \cdot \mathbf{J}_I \cdot \widetilde{\mathbf{W}}_{IE} \cdot \Delta\mathbf{h} \,.$$

The performance of an optimal decoder is then given by

$$\text{fraction correct} = \Phi\left(\frac{\sqrt{S/N}}{\sqrt{2}}\right)$$

where Φ is the cumulative normal function. All of our analysis is based on this expression. Differences in fraction correct depend only on differences in the connectivity matrices, which we describe next.

Connectivity matrices

We consider three connectivity structures: completely non-selective, signal-selective, and signal-to-noise selective (corresponding to Figures 4A, 4B and 4C, respectively). In all cases the connectivity is sparse (the connection probability between any two neurons is 0.1). What differs is the connection strength when neurons are connected. We describe below how the connection strength is chosen for our three connectivity structures.

Non-selective. The connectivity matrices have the especially simple form

$$\mathbf{W}_{\alpha\beta,ij} = \begin{cases} \frac{w_{\alpha\beta}}{\sqrt{cN}} & \text{with probability } c \\ 0 & \text{otherwise} \end{cases}$$

where $\alpha, \beta \in \{E, I\}$, $N(\equiv N_E + N_I)$ is the total number of neurons, and $w_{\alpha\beta}$ are parameters (see Table 1).

Signal-selective. We divide the neurons into two sets of excitatory and inhibitory sub-populations, as in Figure 4B. The connection strengths are still given by the above expression, but now α and β acquire subscripts that specify which population they are in: $\alpha, \beta \in \{E_1, E_2, I_1, I_2\}$, with E_1 and I_1 referring to population 1 and E_2 and I_2 to population 2. The within-population connection strengths are the same as for the non-selective population ($w_{\alpha_i\beta_i} = w_{\alpha\beta}, i = 1, 2$), but the across-population connection strengths are smaller by a factor of ρ,

$$\frac{w_{\alpha_i,\beta_j}}{w_{\alpha_i,\beta_i}} = \rho$$

for $i = 1$ and $j = 2$ or vice-versa. The value of ρ determines how selective the sub-populations are: $\rho = 0$ corresponds to completely selective sub-populations while $\rho = 1$ corresponds to the completely non-selective case.

SNR- selective. We choose the connectivity as in the non-selective case, and then change synaptic strength so that the inhibitory neurons receive stronger connections from the excitatory neurons with high signal to noise ratios. To do that, we first rank excitatory units in order of ascending signal to noise ratio (by using $\mathbf{D}_1$ in the expression for $(S/N)_E$ in the previous section). We then make the substitution

$$W_{IE,ij} \to W_{IE,ij} \left(\frac{r_j}{N_E}\right)^4$$

where r_j is the rank of excitatory j in the order of ascending signal to noise ratio and, recall, N_E is the number of excitatory neurons. This downweights projections from low signal to noise ratio excitatory neurons and upweights connections from high signal to noise ratio neurons. Finally,

all elements are scaled to ensure that the average connection strength from the excitatory to the inhibitory network is the same as before the substitution.

Simulation details

The simulation parameters are given in Table 1. In addition, there are a number of relevant details, the most important of which is related to the input, $\mathbf{h}_s$. As mentioned in the previous section, about half the elements of $\mathbf{h}_s$ are h_0 for the rightward choice and $-h_0$ for the leftward choice, and the rest are h_0 for the leftward choice $-h_0$ for the rightward choice. This is strictly true for the completely non-selective and signal to noise selective connectivity; for the signal selective connectivity, we use $\mathbf{h}_{s,i} = h_0$ for the rightward choice and $-h_0$ for the leftward choice when excitatory neuron i is in population 1, and $\mathbf{h}_{s,i} = h_0$ for the leftward choice and $-h_0$ for the rightward choice when excitatory neuron i is in population 2. In either case, however, this introduces a stochastic element: for the completely non-selective and signal to noise selective connectivities, there is randomness in both the input and the circuit; for the signal selective connectivity, there is randomness in the circuit. In the former case, we can eliminate the randomness in the connectivity by averaging over the input, as follows.

Because the components of $\Delta\mathbf{h}$ are independent, we have

$$\langle \Delta h_{s,i} \Delta h_{s,j} \rangle = \delta_{ij} \langle \Delta h_{s,i}^2 \rangle$$

where δ_{ij} is the Kronecker delta ($\delta_{ij} = 1$ if $i = j$ and zero otherwise). Because $\Delta h_{s,i}$ is either $+h_0$ or $-h_0$, we have

$$\langle \Delta\mathbf{h} \Delta\mathbf{h} \rangle = 4h_0^2 \mathbf{I}$$

where $\mathbf{I}$ is the identity matrix. Thus, when we average the signal to noise ratios over $\Delta\mathbf{h}$, the expressions simplify slightly,

$$\frac{\langle (S/N)_E \rangle}{4h_0^2} = \mathrm{trace}\left\{ \left(\mathbf{D}_n \cdot \mathbf{J}_E \cdot \left[\mathbf{\Sigma}_{EE} + \widetilde{\mathbf{W}}_{EI} \cdot \mathbf{\Sigma}_{II} \cdot \widetilde{\mathbf{W}}_{EI}^T \right] \cdot \mathbf{J}_E^T \cdot \mathbf{D}_n^T \right)^{-1} \cdot \mathbf{D}_n \cdot \mathbf{J}_E \cdot \mathbf{J}_E^T \cdot \mathbf{D}_n^T \right\}$$

$$\frac{\langle (S/N)_I \rangle}{4h_0^2} = \mathrm{trace}\left\{ \left(\mathbf{D}_n \cdot \mathbf{J}_I \cdot \left[\mathbf{\Sigma}_{II} + \widetilde{\mathbf{W}}_{IE} \cdot \mathbf{\Sigma}_{EE} \cdot \widetilde{\mathbf{W}}_{IE}^T \right] \cdot \mathbf{J}_I^T \cdot \mathbf{D}_n^T \right)^{-1} \cdot \mathbf{D}_n \cdot \mathbf{J}_I \cdot \widetilde{\mathbf{W}}_{IE} \cdot \widetilde{\mathbf{W}}_{IE}^T \cdot \mathbf{J}_I^T \cdot \mathbf{D}_n^T \right\}.$$

To avoid having to numerically average over input, we used these expression when computing decoding accuracy for the completely non-selective and signal to noise selective connectivity. That left us with some randomness associated with the networks (as connectivity is chosen randomly), but that turned out to produce only small fluctuations, so each data point in Figures 4A and 4C was from a single network. For the signal selective connectivity (Figure 4B), the network realization turned out to matter, so we averaged over 25 networks, and for each of them we did a further averaging over 100 random picks of the 50 neurons from which we decoded.

The other relevant parameter is the choice of parameter to vary: in Figure 4, the x-axis is the ratio of the average connection strength from excitatory to inhibitory neurons to the average connection strength from inhibitory to excitatory neurons. This was chosen because it turned out to be the connectivity parameter with the largest effect on decoding accuracy. That in turn is

because it turns out to be equivalent to the input noise to the inhibitory population. To see why, make the substitution

$$\mathbf{W}_{IE} \to \gamma \mathbf{W}_{IE}$$
$$\mathbf{W}_{EI} \to \gamma^{-1} \mathbf{W}_{EI}.$$

By letting $\boldsymbol{\nu}_I \to \gamma^{-1}\boldsymbol{\nu}_I$, we see that this is formally equivalent to letting $\boldsymbol{\xi}_E \to \gamma^{-1}\boldsymbol{\xi}_E$, which in turn corresponds to letting $\boldsymbol{\Sigma}_{II} \to \gamma^{-2}\boldsymbol{\Sigma}_{II}$. Thus the x-axis in Figure 4 can be thought of as the axis of decreasing input noise to the inhibitory neurons.

Table 1. Parameters used in simulations

σ	1.25	noise level
δ	0.75	breadth of noise level distribution
w_{EE}	0.25	excitatory → excitatory coupling
w_{II}	-2	inhibitory → inhibitory coupling
w_{IE}	0.87	excitatory → inhibitory coupling
w_{EI}	-0.87	inhibitory → excitatory coupling
c	0.1	connection probability
N_E	4000	number of excitatory neurons
N_I	1000	number of inhibitory neurons
n	50	number of readout neurons
h_0	0.1	input strength
ρ	0.75	selectivity index

Data and code availability

All the data used in the paper are publicly available on CSHL repository: http://repository.cshl.edu/36980/

Code for data processing and analysis is publicly available on github: https://github.com/farznaj/imaging_decisionMaking_exc_inh

Author Contributions

Conceptualization and Writing: FN and AKC. Experiments and Analysis: FN. Decoding methodology and common-slope regression model: GFE, JPC and FN. Circuit modeling: RC and PEL. Spike-inference methodology: EAP. Funding Acquisition, Resources and Supervision: AKC.

Acknowledgements

We thank Hien Nguyen for help with training mice, Matt Kaufman, Kachi Odoemene, Fred Marbach for technical assistance and thoughtful conversations. We thank Andrea Giovannucci for help with ROI inclusion criteria #5. We thank Ashley Juavinett, Simon Musall, and Sashank Pisupati for helpful discussions and feedback on early versions of the manuscript. The work was supported by the Simons Collaboration on the Global Brain, ONR MURI, the Klingenstein-Simons Foundation, the Pew Charitable Trust and the Gatsby Charitable Foundation.

Supplemental Figures

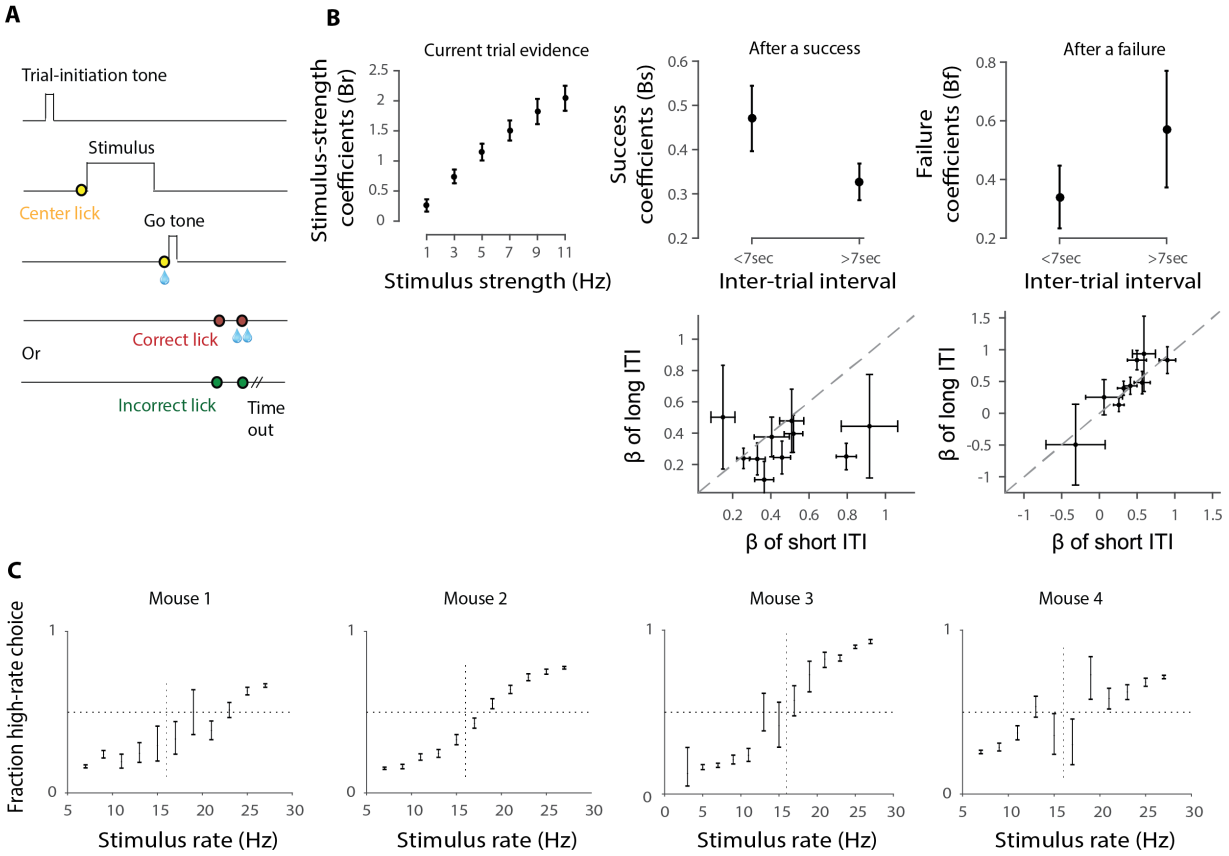

Figure S1. Related to Figure 1. Perceptual decisions about stimulus rate reflect current evidence, previous trial's outcome, and the time passed since the previous trial.

A**,** Trial structure. In each trial, first a brief tone is presented to indicate to the animal to initiate the trial ("trial-initiation tone"). Once the animal licks to the center waterspout (row 2: yellow circle), the stimulus is presented for 1 sec. At the end of the stimulus, the animal is required to lick again in the center (row 3: yellow circle). This will result in: 1) a small water reward in the center, 2) a "go tone" that indicates to the animal to make its choice. If the animal licks to the correct side (row 4, 1st red circle), and confirms this lick (row 4, 2nd red circle), it will receive water as a reward. If the animal licks to the wrong side (last row, 1st green circle), and confirms this lick (last row, 2nd green circle), there will be a time-out, i.e. longer time before the next trial can start. **B**, A logistic regression model was used to assess the extent to which the animal's choice depends on stimulus strength (how far the stimulus rate is from the category boundary at 16Hz), previous choice outcome, and the time interval since the previous choice. Stimulus strength was divided into 6 bins (**left**); previous success was divided into 2 bins: success after a long ITI and success after a short ITI (**middle**); previous failure was also divided into 2 bins: failure after a long ITI and failure after a short ITI (**right**). Plots in top row show β averaged across animals (same 10 animals as in Fig. 1c). Error bars: S.E.M across subjects. **Top left**: strength of the sensory evidence affects the animal's choices: the stronger the evidence, the higher the impact. **Top middle**: Success of a previous trial also affects animal's decision; the effect is stronger when the previous success occurs after a short ITI (<7sec). **Top Right**: Same but for previous incorrect trials; the effect of ITI after a failure was not significant. Plots in **bottom** row show success (left) and failure (right) β for individual mice. Error bars: S.E.M returned from glmfit.m in Matlab. **C,** Behavioral performance of the four mice in which we imaged excitatory and inhibitory activity during decision-making. In mice 1, 2, and 4, imaging was performed throughout learning by tracking the same group of neurons. Plots reflect data from all sessions. Errors bars: Wilson Binomial Confidence Interval.

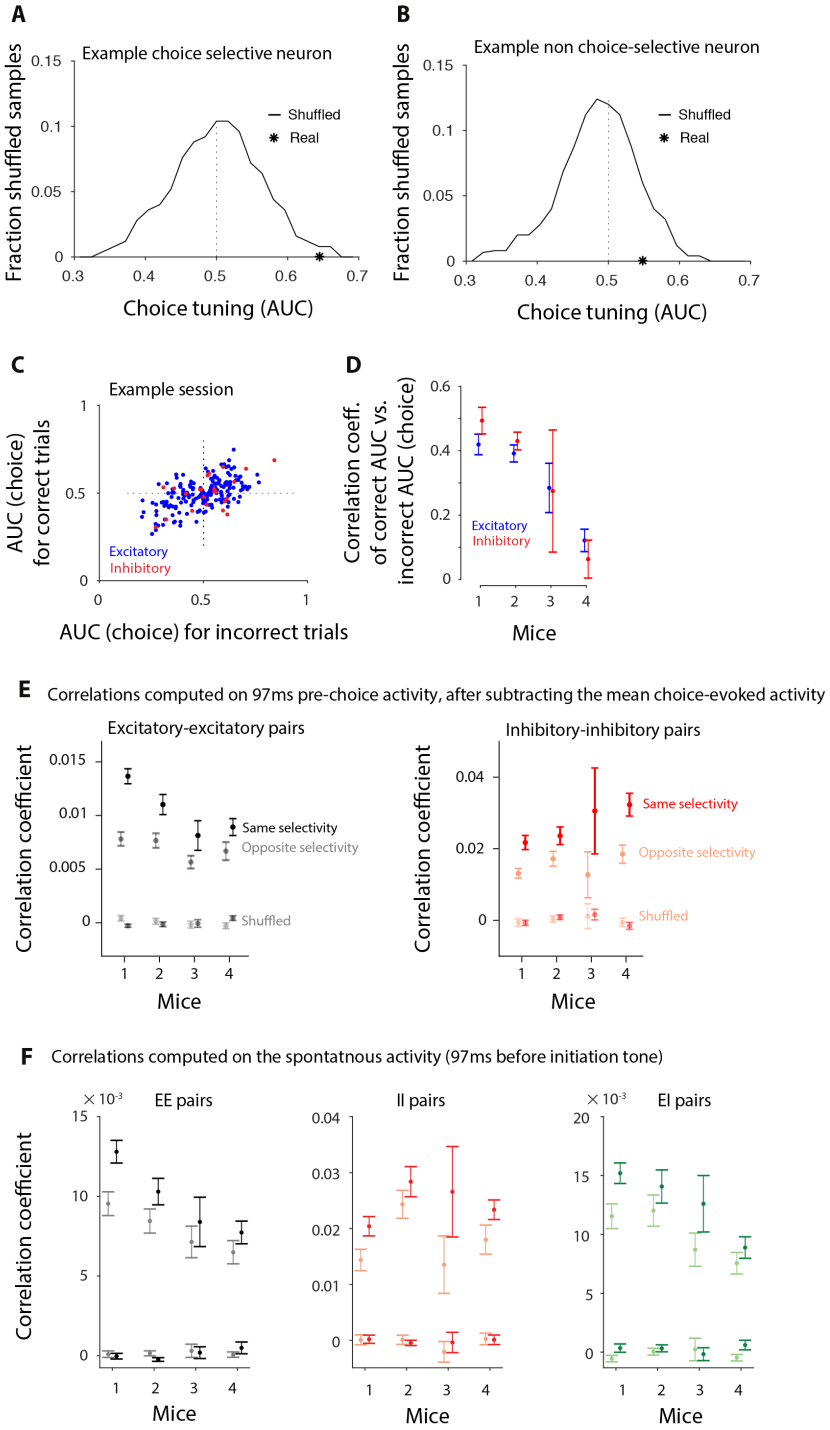
A
Example choice selective neuron
Fraction shuffled samples
Shuffled
Real
Choice tuning (AUC)
B
Example non choice-selective neuron
Fraction shuffled samples
Shuffled
Real
Choice tuning (AUC)
C
Example session
AUC (choice) for correct trials
Excitatory
Inhibitory
AUC (choice) for incorrect trials
D
Correlation coeff. of correct AUC vs. incorrect AUC (choice)
Excitatory
Inhibitory
Mice
E
Correlations computed on 97ms pre-choice activity, after subtracting the mean choice-evoked activity
Excitatory-excitatory pairs
Correlation coefficient
Same selectivity
Opposite selectivity
Shuffled
Mice
Inhibitory-inhibitory pairs
Correlation coefficient
Same selectivity
Opposite selectivity
Shuffled
Mice
F
Correlations computed on the spontatnous activity (97ms before initiation tone)
EE pairs
Correlation coefficient
Mice
II pairs
Mice
EI pairs
Mice

Figure S2. Related to Figure 2. Single neuron measures reveal similar choice selectivity in excitatory and inhibitory neurons.

A,B, Example neurons to illustrate the method for assessing significant choice selectivity in individual neurons. In both panels, the solid line shows the distribution of values for the area under the ROC curve (AUC) generated by 50 different trial shuffles in which trials were randomly assigned to a left vs. right choice. Star indicates the estimated AUC value of the neuron. Significance was assessed from the probability of observing the estimated AUC in the shuffle distribution. When probabilities were <0.05, neurons were considered choice selective. Only the neuron in (a) has significant choice selectivity. **C,** Choice selectivity was computed on correct trials (vertical axis) as well as error trials (horizontal axis), and was correlated between the two conditions. Data is from a single session, each point shows an individual neuron whose cell type is indicated by its color. The positive correlation indicates that choice selectivity was overall similar on correct and error trials (Pearsons' correlation coefficient, excitatory neurons: r=0.58; p<0.001; inhibitory neurons: r=0.55, p=0.007). The small number of points in quadrants 2 and 4 indicate less frequent neurons that showed opposite selectivity on correct vs. error trials. **D,** Summary of correlation coefficient of AUC on correct vs. AUC on incorrect trials, mean across sessions for each animal. Error bars: S.E.M. across sessions. The weaker correlation in mouse 4 indicates that this animal had a mixture of cells selective for the stimulus and cells selective for the choice. Note that although the center of the imaging window was identical in all animals, the imaging location within the window of this animal was slightly posterior to the others. The enrichment of cells selective for the stimulus, in this mouse compared to other mice, may reflect that the region we imaged in mouse 4 was closer to primary visual cortex. **E,** Noise correlations for pairs of excitatory neurons (left) and pairs of inhibitory neurons (red). Noise correlations 0-97 ms before the choice. Signal correlations were removed by subtracting off the across-trial mean response from each trial, separately done for trials of each choice. Noise correlations were overall much stronger for inhibitory neurons (note different scaling on vertical axes for left vs. right plots). For both cell types, correlations were stronger when cells had the same choice selectivity (saturated colors) vs. opposite choice selectivity (unsaturated colors). Unsaturated points are for shuffled trials in which the trial order is scrambled for each neuron independently. Mean +/- standard error across sessions. Correlations for excitatory-inhibitory pairs are shown in Fig. 2E. **F,** Noise correlations for three combinations of cell types: excitatory-excitatory (left), inhibitory-inhibitory (middle) and excitatory-inhibitory (right). Signal correlations were not present because the data are from the inter-trial interval (0-97 ms before the initiation tone) and therefore only reflect spontaneous activity. Color conventions are the same as in (E).

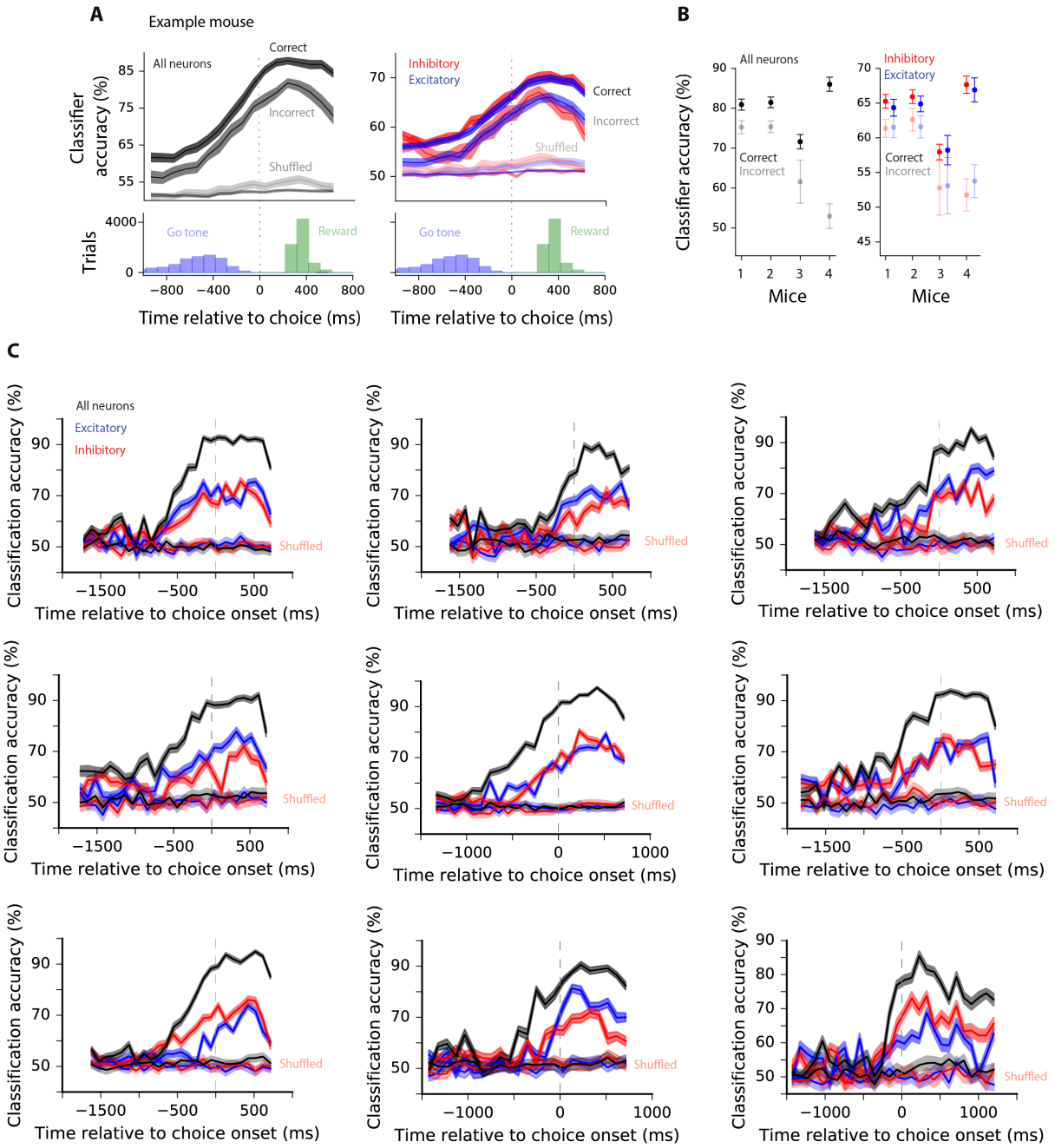
A
Example mouse
All neurons
Correct
Incorrect
Shuffled
Classifier accuracy (%)
Trials
Go tone
Reward
Time relative to choice (ms)
Inhibitory
Excitatory
B
Mice
C
Classification accuracy (%)
Time relative to choice onset (ms)

Figure S3. Related to Figure 3. Classification accuracy on error trials varies across mice but is consistently similar for excitatory and inhibitory populations.

A, Data from an example animal (48 sessions). **Top**: Classification accuracy of decoders trained on all neurons (left), subsampled excitatory neurons (right, blue trace), and inhibitory neurons (right, red trace). In all cases, classifiers were trained on correct trials; however they were tested either on correct (dark lines: "Correct") or incorrect (dim lines: "Incorrect") trials. Unsaturated lines show performance on shuffled trials. Classification accuracy was high, indicating that population activity primarily represents the animal's choice, yet it differs at least slightly for correct and incorrect trials. This reduction was similar for excitatory and inhibitory neurons (blue are red traces are overlapping in the right panel). **Bottom**: As in Fig 3b, histograms reflect the across-trial distribution of go tones and reward delivery. These are variable because the data are aligned to the animal's choice, which was under the animal's control and thus occurred at a variable time relative to go-tone or reward. Left and right panels are the same and are duplicated to facilitate alignment to each corresponding plot above. **B,** Summary across all mice for all neurons (left) and excitatory and inhibitory neurons separately (right). Classifier performance on correct (dark colors) and incorrect (dim colors) trials is shown. Mouse 4 had the largest difference in classification accuracy for correct vs. error trials. As with the single-neuron analysis, this difference likely reflects that the imaging region was slightly posterior within the window for this animal (see Figure 2SD, legend). Importantly, for all mice, the change in classification accuracy was quite similar for excitatory and inhibitory neurons (right), indicating that both populations reflect choice vs. stimulus to a comparable degree. **C**, Classification accuracy of the choice decoder at each moment in the trial for 9 example sessions. Dashed lines: choice onset. Black: all neurons included in the decoder; blue: subsampled excitatory neurons; red: inhibitory neurons; dim colors: shuffled control. In most sessions, inhibitory and subsampled excitatory populations have comparable classification accuracy.

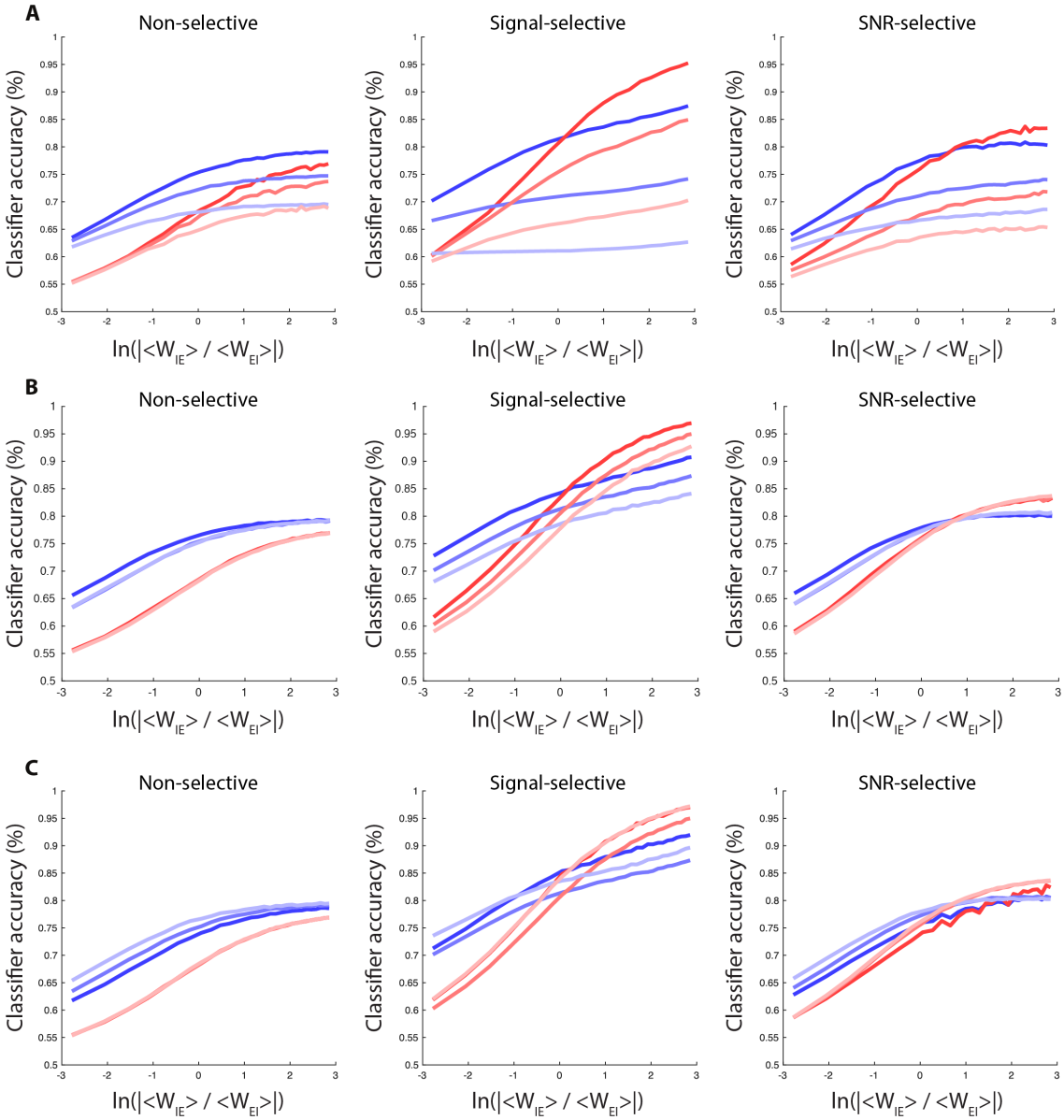

Figure S4. Related to Figure 4. Figure 4. Selective connectivity between excitatory and inhibitory neurons allows for matched classification accuracy in the two populations.

Decoding accuracy versus three parameters. **A,** Differential correlations. $\mathbf{\Sigma}_{EE} \rightarrow \mathbf{\Sigma}_{EE} + \epsilon \Delta \mathbf{h} \Delta \mathbf{h}$. Dark to light hues: $\epsilon = 0, 17.78, 56.23$. **B,** Excitatory to excitatory connections. Dark to light hues: $w_{EE} =$ 0.35, 0.3, 0.25 (default). **C,** Dark to light hues: $w_{II} =$ -2.4, -2.0 (default), and -1.6.

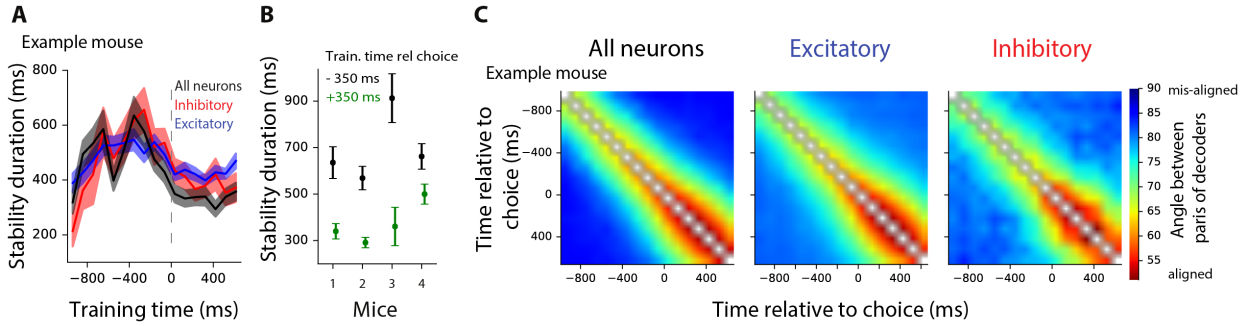

Figure S5. Related to Figure 5. Additional analyses provide more evidence for similar stability of the choice decoder in excitatory and inhibitory populations.

A**,** In an example mouse, population activity that predicts the animal's choice is similarly stable for excitatory and inhibitory neurons during the course of a trial. The vertical axis shows the stability duration for decoders trained at different times during the trial. Stability duration is defined as the width of the testing window over which decoder accuracy is not statistically different (red regions of Fig. 4C) from that obtained by using the same training and testing times (diagonal of Fig. 4A). Error bars: S.E.M. across sessions. Summary data for all mice at training time 0-97 ms before choice (dashed line) are shown in Fig. 4d. **B,** Stability duration of the all-neuron decoder (black in panel a) is compared for decoders trained 350ms before the choice (black), and 350ms after the choice (green). Population stability was lower after the choice than before the choice. This may be due to additional events, e.g. reward delivery and repeated licking, which follow the choice. **C**, Another measurement of stability likewise suggests similar temporal stability for excitatory and inhibitory populations. Stability was assessed by measuring the angle between pairs of decoders trained at different points in the trial. If a similar pattern of population activity represents choice from moment t_1 to moment t_2, the choice classifiers trained at these times will be aligned, i.e. the angle between the two classifiers will be small. The colors indicate the angle between pairs of decoders trained at different moments in the trial. Small angles (hot colors) indicate alignment of choice decoders; hence stable activity patterns across neurons. Left: all neurons; middle: excitatory neurons (subsampled to match the number of inhibitory neurons); right: inhibitory neurons. As with our other method (Fig. 4), the time course of stability was similar for excitatory and inhibitory neurons.

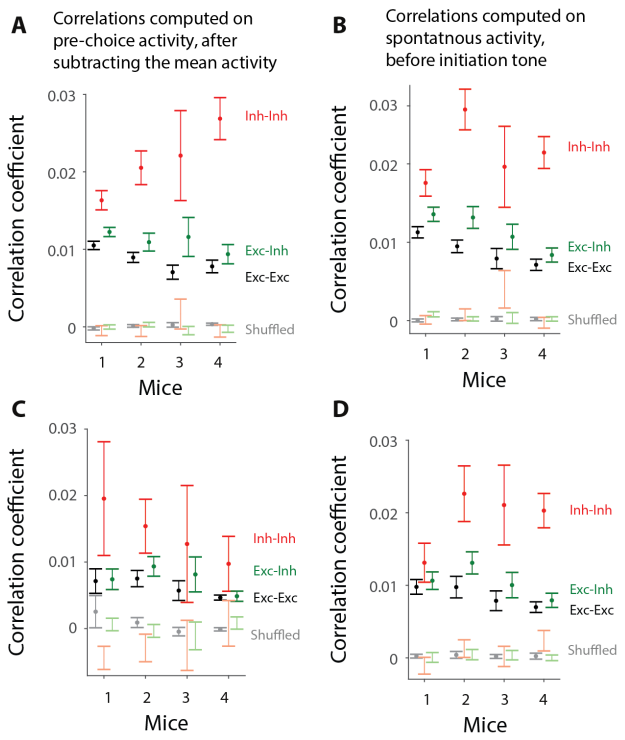

Figure S6. Related to Figure 6. Higher noise correlations among inhibitory neurons.
A, Noise correlations for pairs of excitatory neurons (black), pairs of inhibitory neurons (red) or excitatory-inhibitory pairs (green) during the time period 0-97 ms before the choice. Signal correlations were removed by subtracting off the mean response from each trial (as in Extended Data Fig 2E). Noise correlations were overall much stronger for inhibitory-inhibitory pairs, and had intermediate values for excitatory-inhibitory pairs. **B,** Same as (A) but for the time period 0-97 ms before the trial initiation tone. As in Extended Data Fig. 2F, signal correlations were not present because there is no stimulus and all activity is spontaneous. **C,D,** same as in A,B, except correlations were computed only on those excitatory and inhibitory neurons with the same median spiking activity (Methods).

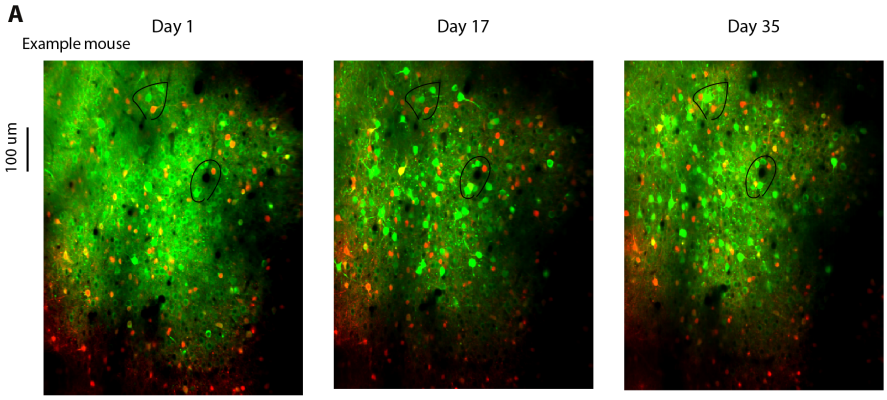

Figure S7. Related to Figure 7. The same field of view was imaged during learning.
A, Field of view from three example sessions of a mouse: 1st days of imaging (left), a middle imaging session (middle), and last day of imaging (right). Left to right panels span 60 days, out of which 35 days were imaged. Black circles mark example areas that can be easily matched among the sessions. Each panel is an average projection image of all the frames imaged in the session. Green and red (bleedthrough corrected) images were merged.

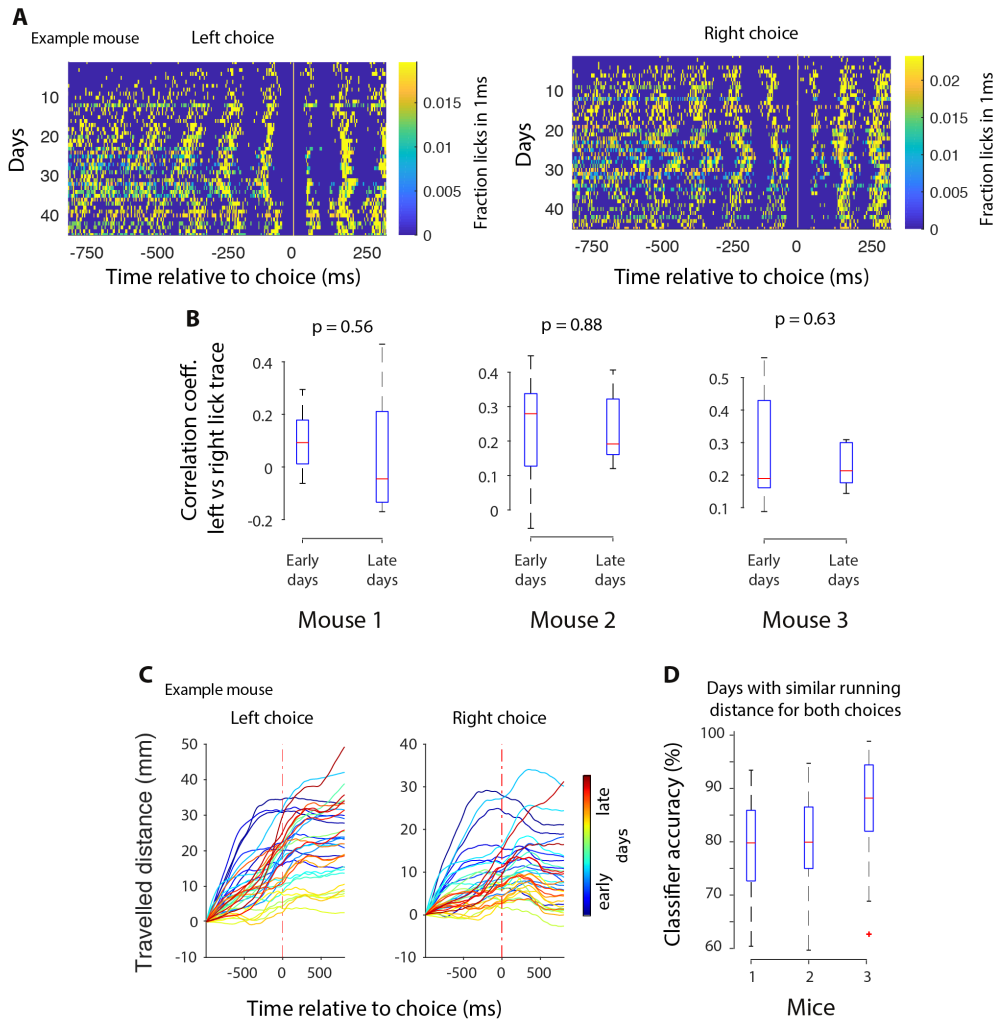

Figure S8. Related to Figure 7. Further analysis of learning-induced changes in the population activity: changes in licking and running movements are unlikely to account for improved classifier accuracy during learning.

A, Licking was similar in advance of high rate vs. low rate choices, both early and late in training. Licks that occur before the choice are to the center waterspout, and licks that occur after the choice are to the side waterspouts; example mouse. **B,** Each plot shows the Pearson's correlation coefficient between licking patterns for left and right choices, calculated 250ms before the choice. These correlations were typically similar for early vs. late training days, indicating that animal's licking pattern preceding left vs. right choices did not change drastically over the course learning. **C,** Distance that the animal travelled during the decision (as measured by the rotary encoder on the running wheel) was similar in advance of left vs. right choices; example mouse; each line represents a session (cold colors: early sessions; hot colors: late sessions). **D,** Classifier accuracy (0-97 ms before the choice) of the full population was high even when the analysis was restricted to sessions in which the distance travelled was not significantly different (t-test, P>0.05; time 0-97 ms before the choice) for left vs. right choices. This analysis was necessary because for some mice in some sessions, there were idiosyncratic differences between the distances travelled in advance of left vs. right choices. In (B) and (D), median (red horizontal line), inter-quartile range (blue box), and the entire range of data (dashed black lines) are shown.

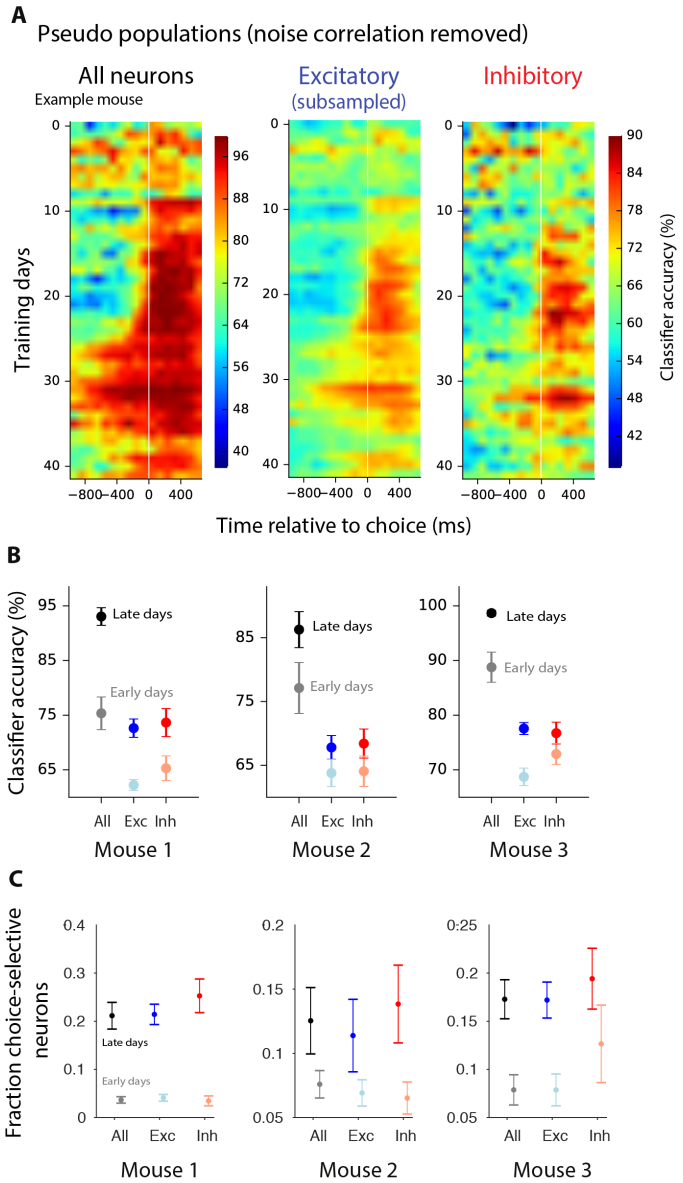

Figure S9. Related to Figure 7. Further analysis of learning-induced changes in the population activity: the reduction in noise correlations is insufficient to account for the improved classification accuracy during learning. Instead, the improvement can be explained by an increase in the fraction of significantly choice-selective neurons. A, Classification accuracy for each training session (average of cross-validation samples), for all neurons (left), subsampled excitatory (middle), and inhibitory neurons (right); example mouse. White vertical line: choice onset. This format is the same as Fig. 6a, but here the noise correlations are removed by making pseudo populations. **B,** Summary of each mouse, showing classification accuracy averaged across early (unsaturated colors) vs. late (saturated colors) training days. As in (A), data are based on pseudo-populations in which the noise correlations are removed. The learning-induced improvement in the classifier accuracy in pseudo populations indicates that reduced noise correlations (Figure 6F) cannot solely account for the enhanced classifier accuracy during learning (Figure 6A). **C,** Fraction of choice-selective neurons increases as a result of training; average across early (dim colors) and late (dark colors) training days; time points 0-97 ms before the choice.

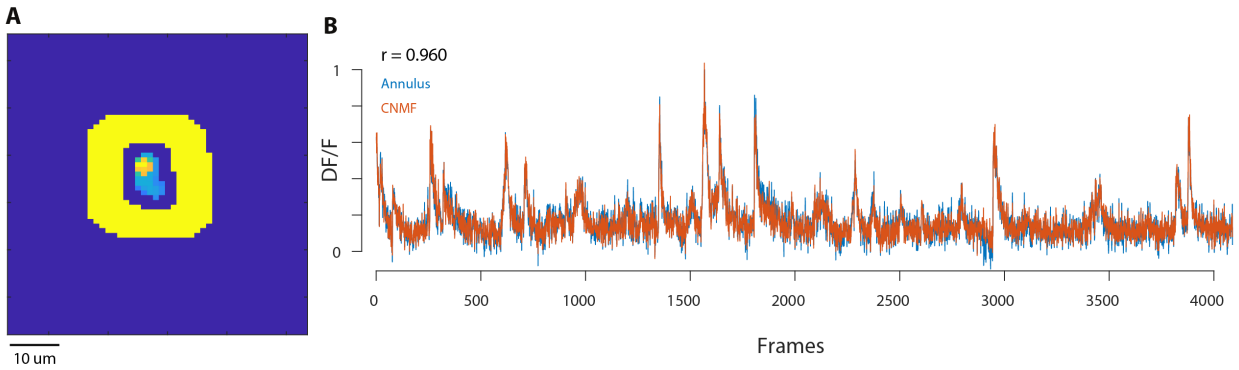

Figure S10. Related to Methods section "Neuropil Contamination removal". Removing neuropil contamination with CNMF or manually using an annulus leads to the same results.

A**,** An example spatial component in the FOV and its surrounding annulus (yellow). **B,** ΔF/F trace for the same component obtained by manually subtracting the neuropil activity averaged over the annulus region (blue trace) or by using the output of the CNMF processing pipeline (red trace). The two traces look nearly identical as also demonstrated by their high correlation coefficient (r = 0.96; the traces are not denoised). These results demonstrate the ability of the CNMF framework to properly capture neuropil contamination and remove it from the detected calcium traces.

References

Aksay E, Olasagasti I, Mensh BD, Baker R, Goldman MS, Tank DW (2007) Functional dissection of circuitry in a neural integrator. Nat Neurosci 10:494-504.

Atallah BV, Bruns W, Carandini M, Scanziani M (2012) Parvalbumin-expressing interneurons linearly transform cortical responses to visual stimuli. Neuron 73:159-170.

Averbeck BB, Lee D (2006) Effects of noise correlations on information encoding and decoding. J Neurophysiol 95:3633-3644.

Averbeck BB, Latham PE, Pouget A (2006) Neural correlations, population coding and computation. Nat Rev Neurosci 7:358-366.

Bogacz R, Brown E, Moehlis J, Holmes P, Cohen JD (2006) The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. Psychol Rev 113:700-765.

Brunton BW, Botvinick MM, Brody CD (2013) Rats and humans can optimally accumulate evidence for decision-making. Science 340:95-98.

Busse L, Ayaz A, Dhruv NT, Katzner S, Saleem AB, Scholvinck ML, Zaharia AD, Carandini M (2011) The detection of visual contrast in the behaving mouse. J Neurosci 31:11351-11361.

Ch'ng YH, Reid RC (2010) Cellular imaging of visual cortex reveals the spatial and functional organization of spontaneous activity. Front Integr Neurosci 4.

Chen TW, Wardill TJ, Sun Y, Pulver SR, Renninger SL, Baohan A, Schreiter ER, Kerr RA, Orger MB, Jayaraman V, Looger LL, Svoboda K, Kim DS (2013) Ultrasensitive fluorescent proteins for imaging neuronal activity. Nature 499:295-300.

Constantinidis C, Goldman-Rakic PS (2002) Correlated discharges among putative pyramidal neurons and interneurons in the primate prefrontal cortex. J Neurophysiol 88:3487-3497.

Cossell L, Iacaruso MF, Muir DR, Houlton R, Sader EN, Ko H, Hofer SB, Mrsic-Flogel TD (2015) Functional organization of excitatory synaptic strength in primary visual cortex. Nature 518:399-403.

Deneve S, Latham PE, Pouget A (1999) Reading population codes: a neural implementation of ideal observers. Nat Neurosci 2:740-745.

Driscoll LN, Pettit NL, Minderer M, Chettih SN, Harvey CD (2017) Dynamic Reorganization of Neuronal Activity Patterns in Parietal Cortex. Cell 170:986-999 e916.

Freedman DJ, Assad JA (2006) Experience-dependent representation of visual categories in parietal cortex. Nature 443:85-88.

Funamizu A, Kuhn B, Doya K (2016) Neural substrate of dynamic Bayesian inference in the cerebral cortex. Nat Neurosci 19:1682-1689.

Galarreta M, Hestrin S (1999) A network of fast-spiking cells in the neocortex connected by electrical synapses. Nature 402:72-75.

Giovannucci A, Friedrich J, Gunn P, Kalfon J, Koay SA, Taxidis J, Najafi F, Gauthier JL, Zhou P, Tank DW, Chklovskii DB, Pnevmatikakis EA (2018) CaImAn: An open source tool for scalable Calcium Imaging data Analysis. bioRxiv.

Goard MJ, Pho GN, Woodson J, Sur M (2016) Distinct roles of visual, parietal, and frontal motor cortices in memory-guided sensorimotor decisions. Elife 5.

Green DM, Swets JA (1966) Signal detection theory and psychophysics. New York: Wiley.

Gu Y, Liu S, Fetsch CR, Yang Y, Fok S, Sunkara A, DeAngelis GC, Angelaki DE (2011) Perceptual learning reduces interneuronal correlations in macaque visual cortex. Neuron 71:750-761.
Guizar-Sicairos M, Thurman ST, Fienup JR (2008) Efficient subpixel image registration algorithms. Opt Lett 33:156-158.
Harvey CD, Coen P, Tank DW (2012) Choice-specific sequences in parietal cortex during a virtual-navigation decision task. Nature 484:62-68.
Hofer SB, Ko H, Pichler B, Vogelstein J, Ros H, Zeng H, Lein E, Lesica NA, Mrsic-Flogel TD (2011) Differential connectivity and response dynamics of excitatory and inhibitory neurons in visual cortex. Nat Neurosci 14:1045-1052.
Hofmann T, Scholkopf B, Smola AJ (2008) Kernel methods in machine learning. Ann Statist 36:1171-1220.
Hwang EJ, Dahlen JE, Mukundan M, Komiyama T (2017) History-based action selection bias in posterior parietal cortex. Nat Commun 8:1242.
Isaacson JS, Scanziani M (2011) How inhibition shapes cortical activity. Neuron 72:231-243.
Jeanne JM, Sharpee TO, Gentner TQ (2013) Associative learning enhances population coding by inverting interneuronal correlation patterns. Neuron 78:352-363.
Kerlin AM, Andermann ML, Berezovskii VK, Reid RC (2010) Broadly tuned response properties of diverse inhibitory neuron subtypes in mouse visual cortex. Neuron 67:858-871.
Khan AG, Poort J, Chadwick A, Blot A, Sahani M, Mrsic-Flogel TD, Hofer SB (2018) Distinct learning-induced changes in stimulus selectivity and interactions of GABAergic interneuron classes in visual cortex. Nature Neuroscience.
Kim Y, Yang GR, Pradhan K, Venkataraju KU, Bota M, Garcia Del Molino LC, Fitzgerald G, Ram K, He M, Levine JM, Mitra P, Huang ZJ, Wang XJ, Osten P (2017) Brain-wide Maps Reveal Stereotyped Cell-Type-Based Cortical Architecture and Subcortical Sexual Dimorphism. Cell 171:456-469 e422.
Kimmel D, Elsayed GF, Cunningham JP, Rangel A, Newsome WT (2016) Encoding of value and choice as separable, dynamic neural dimensions in orbitofrontal cortex. Cosyne.
Ko H, Hofer SB, Pichler B, Buchanan KA, Sjostrom PJ, Mrsic-Flogel TD (2011) Functional specificity of local synaptic connections in neocortical networks. Nature 473:87-91.
Krishna VR, Alexander KR, Peachey NS (2002) Temporal properties of the mouse cone electroretinogram. J Neurophysiol 87:42-48.
Kwan AC, Dan Y (2012) Dissection of cortical microcircuits by single-neuron stimulation in vivo. Curr Biol 22:1459-1467.
Law CT, Gold JI (2008) Neural correlates of perceptual learning in a sensory-motor, but not a sensory, cortical area. Nat Neurosci 11:505-513.
Lim S, Goldman MS (2013) Balanced cortical microcircuitry for maintaining information in working memory. Nat Neurosci 16:1306-1314.
Liu BH, Li P, Li YT, Sun YJ, Yanagawa Y, Obata K, Zhang LI, Tao HW (2009) Visual receptive field structure of cortical inhibitory neurons revealed by two-photon imaging guided recording. J Neurosci 29:10520-10532.
Machens CK, Romo R, Brody CD (2005) Flexible control of mutual inhibition: a neural model of two-interval discrimination. Science 307:1121-1124.

Madisen L, Zwingman TA, Sunkin SM, Oh SW, Zariwala HA, Gu H, Ng LL, Palmiter RD, Hawrylycz MJ, Jones AR, Lein ES, Zeng H (2010) A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. Nat Neurosci 13:133-140.

Mi Y, Katkov M, Tsodyks M (2017) Synaptic Correlates of Working Memory Capacity. Neuron 93:323-330.

Morcos AS, Harvey CD (2016) History-dependent variability in population dynamics during evidence accumulation in cortex. Nat Neurosci 19:1672-1681.

Moreno-Bote R, Beck J, Kanitscheider I, Pitkow X, Latham P, Pouget A (2014) Information-limiting correlations. Nat Neurosci 17:1410-1417.

Ni AM, Ruff DA, Alberts JJ, Symmonds J, Cohen MR (2018) Learning and attention reveal a general relationship between population activity and behavior. Science 359:463-465.

Niell CM, Stryker MP (2008) Highly selective receptive fields in mouse visual cortex. J Neurosci 28:7520-7536.

Odoemene O, Pisupati S, Nguyen H, Churchland AK (2017) Visual evidence accumulation guides decision-making in unrestrained mice. bioRxiv.

Packer AM, Yuste R (2011) Dense, unspecific connectivity of neocortical parvalbumin-positive interneurons: a canonical microcircuit for inhibition? J Neurosci 31:13260-13271.

Panzeri S, Schultz SR, Treves A, Rolls ET (1999) Correlations and the encoding of information in the nervous system. Proc Biol Sci 266:1001-1012.

Pfeffer CK, Xue M, He M, Huang ZJ, Scanziani M (2013) Inhibition of inhibition in visual cortex: the logic of connections between molecularly distinct interneurons. Nat Neurosci 16:1068-1076.

Pho GN, Goard MJ, Woodson J, Crawford B, Sur M (2018) Task-dependent representations of stimulus and choice in mouse parietal cortex. Nat Commun 9:2596.

Pnevmatikakis EA, Soudry D, Gao Y, Machado TA, Merel J, Pfau D, Reardon T, Mu Y, Lacefield C, Yang W, Ahrens M, Bruno R, Jessell TM, Peterka DS, Yuste R, Paninski L (2016) Simultaneous Denoising, Deconvolution, and Demixing of Calcium Imaging Data. Neuron 89:285-299.

Poort J, Khan AG, Pachitariu M, Nemri A, Orsolic I, Krupic J, Bauza M, Sahani M, Keller GB, Mrsic-Flogel TD, Hofer SB (2015) Learning Enhances Sensory and Multiple Non-sensory Representations in Primary Visual Cortex. Neuron 86:1478-1490.

Rao SG, Williams GV, Goldman-Rakic PS (1999) Isodirectional tuning of adjacent interneurons and pyramidal cells during working memory: evidence for microcolumnar organization in PFC. J Neurophysiol 81:1903-1916.

Raposo D, Kaufman MT, Churchland AK (2014) A category-free neural population supports evolving demands during decision-making. Nat Neurosci 17:1784-1792.

Ringach DL, Mineault PJ, Tring E, Olivas ND, Garcia-Junco-Clemente P, Trachtenberg JT (2016) Spatial clustering of tuning in mouse primary visual cortex. Nat Commun 7:12270.

Runyan CA, Piasini E, Panzeri S, Harvey CD (2017) Distinct timescales of population coding across cortex. Nature 548:92-96.

Runyan CA, Schummers J, Van Wart A, Kuhlman SJ, Wilson NR, Huang ZJ, Sur M (2010) Response features of parvalbumin-expressing interneurons suggest precise roles for subtypes of inhibition in visual cortex. Neuron 67:847-857.
Rustichini A, Padoa-Schioppa C (2015) A neuro-computational model of economic decisions. J Neurophysiol 114:1382-1398.
Sahara S, Yanagawa Y, O'Leary DD, Stevens CF (2012) The fraction of cortical GABAergic neurons is constant from near the start of cortical neurogenesis to adulthood. J Neurosci 32:4755-4761.
Schoups A, Vogels R, Qian N, Orban G (2001) Practising orientation identification improves orientation coding in V1 neurons. Nature 412:549-553.
Sohya K, Kameyama K, Yanagawa Y, Obata K, Tsumoto T (2007) GABAergic neurons are less selective to stimulus orientation than excitatory neurons in layer II/III of visual cortex, as revealed by in vivo functional Ca2+ imaging in transgenic mice. J Neurosci 27:2145-2149.
Song YH, Kim JH, Jeong HW, Choi I, Jeong D, Kim K, Lee SH (2017) A Neural Circuit for Auditory Dominance over Visual Perception. Neuron 93:940-954 e946.
Taniguchi H, He M, Wu P, Kim S, Paik R, Sugino K, Kvitsiani D, Fu Y, Lu J, Lin Y, Miyoshi G, Shima Y, Fishell G, Nelson SB, Huang ZJ (2011) A resource of Cre driver lines for genetic targeting of GABAergic neurons in cerebral cortex. Neuron 71:995-1013.
Tanimoto N, Sothilingam V, Kondo M, Biel M, Humphries P, Seeliger MW (2015) Electroretinographic assessment of rod- and cone-mediated bipolar cell pathways using flicker stimuli in mice. Sci Rep 5:10731.
Viswanathan P, Nieder A (2015) Differential impact of behavioral relevance on quantity coding in primate frontal and parietal neurons. Curr Biol 25:1259-1269.
Vogelstein JT, Packer AM, Machado TA, Sippy T, Babadi B, Yuste R, Paninski L (2010) Fast nonnegative deconvolution for spike train inference from population calcium imaging. J Neurophysiol 104:3691-3704.
Wang XJ (2002) Probabilistic decision making by slow reverberation in cortical circuits. Neuron 36:955-968.
Wang XJ, Yang GR (2018) A disinhibitory circuit motif and flexible information routing in the brain. Curr Opin Neurobiol 49:75-83.
Wang XJ, Tegner J, Constantinidis C, Goldman-Rakic PS (2004) Division of labor among distinct subtypes of inhibitory neurons in a cortical microcircuit of working memory. Proc Natl Acad Sci U S A 101:1368-1373.
Yoshimura Y, Callaway EM (2005) Fine-scale specificity of cortical networks depends on inhibitory cell type and connectivity. Nat Neurosci 8:1552-1559.
Yoshimura Y, Dantzker JL, Callaway EM (2005) Excitatory cortical neurons form fine-scale functional networks. Nature 433:868-873.
Znamenskiy P, Kim M-H, Muir DR, Iacaruso MF, Hofer SB, Mrsic-Flogel TD (2018) Functional selectivity and specific connectivity of inhibitory neurons in primary visual cortex. bioRxiv.
Zohary E, Shadlen MN, Newsome WT (1994) Correlated neuronal discharge rate and its implications for psychophysical performance. Nature 370:140-143.