Decoding thalamic afferent input using microcircuit spiking activity

Abbreviated title: Decoding thalamic input

Audrey J Sederberg
Department of Organismal Biology and Anatomy; Department of Neurobiology
University of Chicago

Stephanie E Palmer
Department of Organismal Biology and Anatomy; Committee on Computational Neuroscience
University of Chicago

Jason N MacLean
Department of Neurobiology; Committee on Computational Neuroscience
University of Chicago

Corresponding author:
Jason N MacLean, jmaclean@uchicago.edu

Number of figures: 6
Number of tables: 0
Number of multimedia files: 0

Number of words:
  Abstract: 144
  Introduction: 499
  Discussion: 1005

The authors declare no competing financial interests or other conflicts of interest.
Abstract

A behavioral response appropriate to a sensory stimulus depends on the collective activity of thousands of interconnected neurons. The vast majority of cortical connections arise from neighboring neurons, and thus understanding the cortical code requires characterizing information representation at the scale of the cortical microcircuit. Using two-photon calcium imaging we densely sampled the thalamically evoked response of hundreds of neurons spanning multiple layers and columns in thalamocortical slices of mouse somatosensory cortex. We then used a biologically plausible decoder to characterize the representation of two distinct thalamic inputs, at the level of the microcircuit, to reveal those aspects of the activity pattern that are likely relevant to downstream neurons. Our data suggest a sparse code, distributed across lamina, in which a small population of cells carries stimulus-relevant information. Furthermore, we find that, within this subset of neurons, decoder performance improves when noise correlations are taken into account.

Keywords: cortical coding, decoding, two-photon imaging

Introduction

A key step in the process of extracting relevant information from sensory inputs and transforming it into a behavioral output is performed by the local cortical microcircuit. Much work has been done to characterize the direct representation of feed-forward inputs to primary sensory cortex. Comparatively little is known of how local recurrent circuitry transforms sensory information in the first of multiple stages of cortico-cortical processing that ultimately lead to behavioral output. Perhaps the most relevant aspect of this representation of information is the degree to which inputs can be read out by subsequent processing stages. Systematically quantifying the decodability of circuit activity distributed across columns and layers allows us to begin to identify what parts of a response can be used by downstream processors and ultimately drive behavior.
This paper describes decoding capabilities of local populations of neurons in the mouse primary somatosensory barrel cortex. Although barrel cortex is somatotopically organized, with each barrel column mapping to a principal whisker, sensory information from each whisker is represented in surrounding columns as well (Drew and Feldman 2007, Kerr et al. 2007, Jadhav et al. 2009, Elstrott et al. 2014, Jouhanneau et al. 2014), indicating a distributed representation that would permit the rodent to integrate input from the full array of whiskers. Our study analyzes the decoding power of neuronal populations beyond the principal barrel columns.

Specifically, we stimulate two distinct sets of thalamic relay cells in a thalamocortical slice and test how stimulus identity can be decoded from the response of indirectly stimulated circuits, those outside the directly activated columns. Previously, it was shown that such thalamically evoked activity occurs in specific ensembles of neurons that are also co-active in spontaneous events (MacLean et al. 2005). The relationship between stimulus-evoked and spontaneously occurring activity in cortex has also been established in vivo (Tsodyks et al. 1999, Kenet et al. 2003, Fiser et al. 2004, Luczak et al. 2009, Luczak and MacLean 2012). We aim to uncover how information about thalamic inputs can be determined from these cortical activity patterns. Our ex vivo experiment allows us to monitor the spiking activity of a large, densely sampled population of neurons, of which many are recurrently connected (Song et al. 2005, Perin et al. 2011), across columns and layers with high temporal resolution. By focusing on activity outside of the principal columns, we deliberately emphasize that part of the code that is carried by recurrently generated activity in neocortex.

Serial recordings can be used to characterize information representation in neuronal populations, but this approach rests on the assumption that neuron-neuron correlations have a minimal role in sensory processing. In early sensory areas, such as retina, even weak neuron-neuron correlations shape the population response (Pillow et al. 2008; Scheidman et al. 2006; Schneidman et al. 2008; Shlens et al. 2006; Tkačik et al. 2010), though the degree to which this encoding structure affects decoding has only begun to be explored (Ganmor et al. 2011; Vidne
et al. 2011). In contrast to the retina, where processing is largely feed-forward, cortical circuits have a high degree of recurrent connectivity (Song et al. 2005), the effect of which is difficult to predict. Recurrent connections may serve to decorrelate cortical activity (Renart et al. 2010), such that cortical neurons represent independent information. Alternatively, recurrence may actively suppress noise in cortical activity patterns, enhancing the decodability of encoded information (Pouget et al. 1998). A difficulty of directly assessing the representation of stimulus information from sparsely sampled cortical populations is that cortical connectivity is itself sparse, although synaptic connections are more probable between spatially proximate neurons (Song et al. 2005; Fino and Yuste 2011; Ko et al. 2011; Lefort et al. 2009; Packer and Yuste 2011; Thomson 2002; Perin et al. 2011).

Decoding performance is quantified at the single-cell level across neuronal populations spanning multiple layers. We measure temporal aspects of the code and show that timing relative to the onset of population activity is sufficient to decode the stimulus identity. Finally, we quantified the representation built up by pairs and small groups of neurons, finding that noise correlations between neurons play a positive role in stimulus decoding. These data reveal insights into the cortical code defined at the level of the cortical microcircuit for somatosensation.

Materials and Methods

Preparation of slices for two-photon imaging.

Thalamocortical slices from somatosensory cortex (Agmon and Connors, 1991) were obtained from C57BL/6 strain mice (N = 18) of either sex on postnatal day 15-17 using methods as previously described (Sadovsky and Maclean, 2013; Sadovsky et al., 2011). Briefly, the brain was extracted and, after one minute in ice-cold solution (ACSF with NaCl replaced by sucrose; contents (in mM): 3 KCl, 26 NaHCO₃, 1 NaH₂PO₄, 0.5CaCl₂, 3.5MgSO₄, 25 dextrose, 123 sucrose), hemisected and blocked, then sliced with a Vibratome. Slices were placed in a 35°C incubation fluid (Incu-ACSF; contents contain the following, in mM: 123 NaCl, 3 KCl, 26 NaHCO₃, 1 NaH₂PO₄, 2 CaCl₂, 6 MgSO₄, 25 dextrose) for 30–45 min. Calcium dye loading was then
achieved by placing all slices into a small Petri dish containing 2 ml of Incu-ACSF, an aliquot of 
50 μg Fura-2AM (Invitrogen) in 13 μl DMSO and 2 μl of Pluronic F-127 (Invitrogen) and kept at 
35C for 25-27 minutes, with the shorter time used for the younger mice. Prior to 
experimentation, slices rested at room temperature in incubation fluid for a minimum of 40 
minutes. ACSF solutions were continuously aerated with 95% O₂/5% CO₂ gas. All procedures 
were performed in accordance and approved by the Institutional Animal Care and Use 
Committee at the University of Chicago.

**Electrophysiology**

Cortical activity for decoding analysis was evoked by thalamic stimulation. Two platinum iridium 
CE2C55 stimulating electrodes (Frederick Haer Co., Bowdoinham, ME) were placed in the 
ventral posteromedial nucleus (VPm). Stimulating electrodes were placed with at least one 
complete barreloid, identified under brightfield illumination, between them. Across slices, this 
separation was at least 90 microns (mean, 132 +/- 33 µm (SE)), along the dorsal-ventral axis. 
Afferents from the central, unstimulated barreloid were visually tracked to cortex to obtain the 
optimal imaging field of view. Correct placement of the stimulating electrodes was confirmed 
by a response in patch-clamped cell or, in the experiments in which electrophysiology was 
unavailable, by a population response from imaging data. Stimulation consisted of six 200-µs 
current pulses at 40 Hz with a minimal amplitude of 10-30µA and was applied in one of the two 
locations every 20 seconds. Stimulation amplitude was set to the minimum level that evoked a 
cortical response, determined by monitoring a patched cell in cortex. If no cell was patched, a 
brief movie consisting of 2-5 trials at each location was taken and the fluorescence traces 
visually checked for a population response to both stimuli.

Experimentation was performed in standard ACSF (containing the following, in mM: 123 NaCl, 3 
KCl, 26 NaHCO₃, 1 NaH₂PO₄, 2 CaCl₂, 2 MgSO₄, and 25 dextrose, which was continuously aerated 
with 95% O₂, 5% CO₂) flowing at 3 mL/minute. Whole-cell current-clamp recordings were made 
using Multiclamp 700B amplifiers (Molecular Devices). Cells selected for patch-clamp recording 
were mostly in layer 4, with a few in lower layer 3 (3 cells) or upper layer 5 (2 cells). Only cells
that were consistently responsive to both stimulation locations in an initial test period were monitored for the entirety of the experiment and included in the analysis.

**Two-photon imaging**

Rapid whole-field imaging of Fura-2AM loaded neurons was achieved using the Heuristically Optimal Path Scanning technique and microscopy setup as previously detailed (Sadovsky et al., 2011), allowing us to monitor action potential generation within individual neurons over each 6-minute movie. Dwell time was fixed at a value between 20 and 50 samples/cell/frame for each experiment.

Because of the importance of having highly stable neuronal recordings for each movie, regions of interest, corresponding to single cells, were extracted from the path scan by overlaying the average response along the path with the expected location of each putative cell. Within the expectation window, a cell mask was generated based on the smoothed (Gaussian filter, width of 5 samples, or 10% of on-cell samples) fluorescence profile. A cell mask was generated independently for each movie. To ensure recording stability, cell masks had to remain stationary across all movies, defined as better than 80% correlation between the smoothed cell masks of consecutive movies. Fluorescence traces for each cell were generated from the dot product between the cell mask and the path scan. In 3 of the 18 mice, multiple fields of view were used in recordings. Across a total of 22 fields of view, 3453 cells were identified. Spiking activity was inferred from fluorescence traces using a customized deconvolution method (Sadovsky and Maclean, 2013; Vogelstein et al., 2009). This gives a measure of the likelihood of a spiking event in each frame, ranging from 0 to 1. Based on hand-scoring a subset of data and observing that even small values (0.03) were indicative of probable spiking events, while large values (>0.2) corresponded to multi-spike events, we use this activity probability measure directly, rather than setting a threshold, as the input to all decoding analyses.

**Decoding analysis**
The decoder task is to identify which of the two thalamic locations was stimulated. We apply a linear decoder, implemented by a support vector machine (Boser et al. 1992; Cortes and Vapnik 1995) with a linear kernel (Matlab, The MathWorks Inc., Natick, MA). Perfect separating hyperplanes did not exist for most datasets, so soft margins were used, parameterized by regularization parameter C. To set C, we tested logarithmically spaced values from 0.1 to 10 on a subset of data; values from 0.3 to 3 produced comparable classification success rates. C was fixed at 1 for all subsequent analyses.

A minimum of 20 total trials was required for inclusion in the decoding analysis. A trial consisted of a stimulus delivered at a single location. The number of trials was limited by dye responsiveness. Four experiments had 24 trials; thirteen had between 30 and 50 trials; and five had 60 or more trials. Cross-validation is performed by dividing the trials into test (20% of trials) and training sets successively until all trials have been included in exactly one test set. Test sets are constrained to have equal number of trials for each stimulus. Decoding performance is the fraction of correctly identified test-set trials. We assess the quality of the decoder fit by comparing test set and training set performance. If the test set performance is higher than the training set performance by more than one trial, we conclude that there was too little data to fit the decoder and exclude the cell or cell set from decoding analysis. This occurs in <1% of all groups.

We decode activity pooled over a window of a fixed bin size starting a fixed time after the stimulus was delivered. For multi-cell decoders, the bin size and start time are the same for all cells. Because the number of cells prohibits exhaustive sampling of groups larger than 2, we randomly selected 2000 sets from each experiment for all sets of three or larger. The bin size and start time are varied over a range from one frame (30-100ms) to 500ms. For each decoder, the optimal window (bin size and start time) for decoding is identified from performance on the training set. The reported decoding performance of the cell is the test-set decoding performance over that window.
To determine whether a cell decodes above chance levels, we run the decoder on data with stimulus labels shuffled (250 shuffles) and compare the measured decoding performance to the distribution of stimulus-shuffled decoding performances. Cells that decode above chance levels are defined as those that score at or above the 95\textsuperscript{th} percentile of the shuffled distribution. With a small number of trials, it is possible that cells carrying a small amount of stimulus information would not be detected as decoding above chance levels. There was otherwise no relationship between the number of trials and decoder performance.

**Shuffling procedure for noise correlation analysis**

For the noise correlation analysis, we first identify the bin size and start time leading to the highest decoding performance for a given pair. For that combination of bin size and start time, we compose shuffled population responses for each stimulus by randomly drawing cell responses from different trials, thereby breaking trial-by-trial correlations between cells while keeping stimulus-driven correlation intact. The shuffled decoding performance is the mean decoding performance across 250 shuffles.

**Statistical tests**

**Shuffle test for pair separation significance**

Physical separation between neuron pairs is not an independent quantity across the population, so significance for the observed differences in pair separation between the jointly decoding pairs and the general class of above-chance decoder pairs was determined by performing a shuffle test. We computed all pairwise separations between above-chance decoder pairs. From this set, we drew random subsets matched to the number of jointly decoding pairs. We then compare the distribution of mean separations, calculated over 1000 random subsets, to the mean separation between jointly decoding pairs. The absolute difference is 5 \(\mu\text{m}\) (stimulus-locked) and 15 \(\mu\text{m}\) (population-locked), which is small relative to the mean separations (160 \(\mu\text{m}\)) but significant (stimulus-locked, \(p = 0.010\); population-locked, \(p < 0.001\)).
Computing the Kullback-Leibler (KL) divergence

To determine whether the best decoders were spatially distributed non-uniformly, we compared the distribution of distance from pia for the best decoders to the distribution of distance from pia of all recorded cells. We quantify this by computing the KL divergence:

\[ D_{KL}(p; q) = \int p(x) \log \frac{p(x)}{q(x)} dx \]

where \( p(x) \) is the probability of a ``best decoder'' being located at position \( x \) and \( q(x) \) is the probability of a recorded cell being located at position \( x \). We compute this using a nearest-neighbor estimate (Wang et al. 2006), which is an unbiased estimator of the divergence. To determine whether the measured \( D_{KL} \) was significant, we generated \( p^{shuf}(x) \) by drawing from all recorded cells (i.e. \( q(x) \)) using matched sample sizes. While \( D_{KL} \geq 0 \), the shuffle distribution of \( D_{KL} \) values is expected to bracket zero, and so will contain negative values. We therefore use a two-sided test for significance. If the measured \( D_{KL} \) fell within the 95% confidence interval (2.5 to 97.5 percentile) of 0, it was deemed not significantly different from zero. Quoted \( p \)-values are calculated from the percentile in the shuffle distribution; for example, the 90th percentile corresponds to a \( p \)-value of 0.2.

Temporal alignment for cortex-centric decoding

An alternative to aligning activity relative to stimulus onset is to align based on the onset of population activity, determined across the full set of cells for each individual trial. For each trial, we took the activity across all neurons and set \( t = 0 \) to the frame at which at least three neurons were co-active, which is two standard deviations above the background activity rate of 0.36 ± 1.24 active cells per frame, measured over the two-second period pre-stimulus. For population-locked alignment, a different start time relative to absolute start time is used in every trial. As with stimulus-locked activity, we performed the decoding analyses as described.

Simple model of Gaussian coding units

To set expectations for a decoder with a linear kernel, we demonstrate the performance of this decoder in a simple model that is commonly used to illustrate principles of population coding.
Consider an independent collection of $N$ Gaussian coding units, which each respond to two stimuli with mean $\mu = \pm 1$ and variance $\sigma^2$. The population response distribution is

$$P(r|\mu = \pm 1) = (2\pi\sigma^2)^{-1/2} \exp \left( -\frac{1}{2} (r - \mu)C^{-1}(r - \mu) \right)$$

where the covariance matrix is

$$C \equiv \frac{2}{\sqrt{2}}$$

In this simple case, the optimal linear decoder is a plane $P$ that is perpendicular to the vector $\mathbf{1}$. The success rate $S$ of this decoder is

$$S = \int_0^\infty d^N r P(r|\mu = 1) + \int_{-\infty}^0 d^N r P(r|\mu = 1)$$

The multidimensional integral is evaluated by rotating into the orthogonal coordinate system with $z$ axis along $\mathbf{1}$, so the plane $P$ is simply $z = 0$, The distance from $P$ to the mean response grows with $N^{1/2}$, while variance does not grow with $N$. The integrals in all dimensions except along $z$ evaluate to 1. The remaining integral along the $z$-axis is

$$S = \int_0^\infty (2\pi\sigma^2)^{-1/2} \exp \left( -\frac{(r - \sqrt{N})^2}{2\sigma^2} \right) dr$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\sqrt{N}}^\infty \exp \left( -\frac{r^2}{2\sigma^2} \right) dr$$

$$= 0.5 \left( 1 + \text{erf} \left( \sqrt{N/2\sigma^2} \right) \right).$$

This function is curve-down and saturates at 1: with each additional decoder, the improvement to the group decoding performance decreases.

We can extend this to the case of covariance matrices with identical off-diagonal elements $r$ (equivalent to the correlation coefficient between a pair of decoders). For example, if $N = 3$, this is:

$$C = \sigma^2 \begin{pmatrix} 1 & r & 2 \frac{2}{2} \\ r & 1 & r \\ r & r & 1 \end{pmatrix}.$$
This either flattens or elongates the spherical response distribution into an ellipsoidal distribution with one principal axis along the vector $\mathbf{1}$. The normalized eigenvectors of $C$ are $v_1 = N^{1/2} \phi^2 \mathbf{1}$ with eigenvalue $\lambda_1$ and the vectors spanning the subspace (of dimension $N-1$) orthogonal to $v_1$, with identical eigenvalues $\lambda_0$. The relationship between $\lambda_{0,1}$ and $r$ is:

$$\lambda_1 = 1 + (N-1)r \quad \lambda_0 = 1 - r$$

Because $\lambda_1 \geq 0$, there is a lower limit on how negative correlation can be: $r \geq -(N-1)^{-1}$. Intuitively, this comes from the fact that it is impossible to have an arbitrary number of perfectly anti-correlated variables.

The effect of correlation is to constrict or expand the responses along the direction of discrimination ($v_1$), depending on whether $\lambda_1 > 1 (r > 0)$ or $\lambda_1 < 1 (r < 0)$. The decoding success rate is:

$$S = \frac{1}{2} \left(1 + \text{erf} \left( \sqrt{\frac{N}{2\lambda_1 \sigma^2}} \right) \right)$$

This is in line with expectations: positive signal correlation (by assumption, since $\mu = +1$) and negative noise correlation ($r < 0$, or $\lambda_1 < 1$) increases the decoding success by shrinking the variance along the discrimination dimension.

If the decoders are synergistic, $S(N)$ would grow super-linearly in $N$. Finding a combination of $r$, $\sigma$, and $N$ that generates a synergistic curve $S(N)$ amounts to finding a value of $r$, $\sigma$, and $N$ for which:

$$S(N + 2) - S(N - 1) > S(N - 1) - S(N)$$

Although $N$ is discrete, $S(N)$ is well-defined for continuous $N$, so we solve

$$\frac{\partial^2 S}{\partial N^2} > 0$$
and look for solutions with integer values of $N$. The condition is

$$f''(N) - (f'(N))^2 - \frac{(f'(N))^3}{2f(N)} > 0$$

where

$$f(N) = \frac{N}{2\sigma^2(1 + (N - 1)r)}$$

This amounts to solving a quadratic equation in $N$ that is parameterized by $r$ and $\sigma$. Solving numerically, we find that the synergistic regime corresponds to unrealistically large negative correlations (generally $<-0.2$) and low single-cell decoding performances ($<0.67$).

Results

To characterize the representation of thalamic afferent input at the level of the cortical microcircuit, we uniformly and densely sampled thalamically evoked neuronal activity in slices of mouse somatosensory barrel neocortex using high speed multiphoton calcium imaging (Sadovsky et al. 2011; Vogelstein et al. 2009). We imaged an average of 157 +/- 60 SE (range =40-267) neurons within a single field of view at 9-50 Hz and simultaneously patch-clamped individual neurons ($n=21$, including 7 pairs), which provided 10 kHz sampling of recorded neurons (Fig. 1A-B). Cortical circuit activity was evoked using minimal electrical stimulation (Beierlein and Connors 2002), consisting of 6 current pulses (200 µs) at 40Hz, applied to one of two distinct locations in the ventral posteromedial (VPM) nucleus of thalamus (MacLean et al. 2005). Following the brief period of stimulation, a cortical response was triggered, characterized by an UP state in a single cell and a circuit event across the population. Stimulation locations were separated by at least 90 microns (mean, 132 +/- 33 µm (SE)), approximately equivalent to one complete barreloid in the VPM (Van Der Loos 1976), and the imaged field of view corresponded to an indirectly stimulated barrel. In vivo, activity triggered by the principal whisker is only briefly (<20ms) confined to the principal barrel before it spreads to adjacent barrels, and even within the principal barrel, a subset of cells respond more strongly
to neighboring whiskers (Ferezou et al. 2007). By centering the imaging in the indirectly
activated column, we specifically analyze the part of the code that arises from corticocortical
synaptic interactions.

We evaluated the representation of thalamic inputs by measuring the ability of cells and groups
of cells to decode stimulus location (Fig. 1D) using a linear decoder, i.e. a binary classifier with a
linear kernel. To establish performance expectations, we demonstrate how this decoder
performs for a population of neurons whose responses are characterized only by mean and
covariance. Such a simple model is a commonly used first approximation for characterizing the
role of correlation in a population code (Averbeck et al. 2006). We first consider the decoding
performance of a collection of independent, identical cells (Fig. 1D), with stimulus-dependent
responses characterized by a Gaussian distribution. In this model, decoding performance
increases sublinearly with each additional cell and the rate of improvement is determined by
the single-cell signal-to-noise ratio. Adding correlations between cells changes the multi-
neuronal decoding performance (Fig. 1E). In this case, negative correlation between cells
improves the decoding performance of a group over the independent condition, while positive
correlation decreases performance (Fig. 1E). From the model, it is expected that we will find the
largest gain in decoding performance derives from the first cell, with diminishing gains for each
subsequent cell if each neuron is independent. Further, the model indicates that we can expect
either an increase or decrease in decoding performance when we shuffle the data, depending
on whether noise correlations degrade or improve decoding accuracy.

**Single-cell decoding from patch-clamped cells**

Thalamically evoked activity was characterized at the single-cell level by prolonged
depolarizations, or UP states, in patch-clamped neurons and at the circuit level by a
multineuronal response in the imaged field (Fig. 1B). If patch-clamped neurons did not reliably
respond to the thalamic stimuli on every trial over an initial test period, then that neuron was
not monitored and another attempt was made to find a responsive cell. Overall, 21 patch-
clamped neurons, including 7 pairs, reliably responded to thalamic input and were used for decoding analysis.

To fit a decoder to spike train data, we counted spikes in a fixed peristimulus window of varying lengths (10 ms to 500 ms) and start times (0 to 500 ms) over the post-stimulus period (Fig. 2A). Chance-level performance was computed from the distribution of decoding performance from stimulus-shuffled data; cells that scored at or above the 95th percentile of the shuffle distribution were considered above chance. Given the simplicity of the experimental design and the fact that we biased our sample to reliably responsive neurons, we expected that single neurons could easily decode the two-stimulus task. However, only a small minority of cells (N = 3 of 21) decoded with near-perfect (>0.9) performance (Fig. 2B). Although most patch-clamped neurons decoded at levels exceeding chance (N = 18/21), the mean decoding performance was only 0.78 +/-0.11 (SE) (Fig. 2B).

Even for the best performing cells, there was a limit to the gains afforded by the high temporal resolution (10 kHz) provided by patch clamp recording. For instance, averaging over a window of 100 ms could be more informative than a shorter time window of 10 ms (Fig. 2C). The most informative temporal window within any one neuron was highly variable within the tested range (mean start time = 190 ms, SE = 170 ms; mean bin size 340 ms, SE 170 ms). Importantly, this timescale is accessible to 2-photon imaging of neuronal populations. That we rarely achieved perfect decoding using spike trains recorded from reliably responsive patch clamped neurons indicated the importance of the population for the representation of the thalamic input. To characterize how information is represented across the population, we evaluated the decoding performance in the imaged population.

Decoding with single cells from imaging
Of 3453 imaged neurons, 3326 (96%) had a detectable change in fluorescence within one second of the stimulus onset in at least one trial. We limited our decoding analysis to these neurons. Following thalamic stimulation, we found that the activity within this subset of neurons exhibited a range of reliability with some neurons being consistently active (559 neurons (16%) active in >50% of trials) and others responding in only a few trials (818 neurons (24%) in < 10% of trials). Across all imaging experiments (N = 22), the shortest latency to detectable activity in any one neuron was 49 ms +/- 20 ms and the subsequent propagation of activity through the local cortical circuits within each imaged field of view lasted 1 – 3 seconds (mean 2.1 +/- 0.9 SE) as previously reported (MacLean et al. 2005; Watson et al. 2008). The mean latency preceding population activity across trials in which there was a detectable response was 600 +/- 190 ms (mean+/-SE, N = 22 experiments). This is longer than the decoding timescale from the patch clamp recordings, but still consistent with intercolumnar activity propagation speeds (Wester and Contreras 2012). Thus, we were able to reliably evoke activity, primarily through indirect cortico-cortical connectivity with thalamic stimulation applied to one of two locations, and image cortical circuit activity.

As with the patch clamp data, we used a linear decoder over a variable peristimulus time window trained on the deconvolved fluorescence signal (Fig. 2D-E). For imaging datasets, overlapping decoding time windows had lengths from 20-50 ms to 500 ms and starting times ranging from 0 ms to 500 ms (Fig. 2F). We used the same criteria for decoding significance for the imaging data that we used for the electrophysiological data. Most neurons (N = 2341/3326, 70%) failed to decode above chance levels, performing below the 95th percentile in the stimulus-shuffled distribution. The remaining 985 cells, found across all 22 datasets, decoded the stimulus at higher than chance levels (Fig. 2G). Compared with the set of patch-clamp recorded cells, above-chance decoders were less common in the imaging dataset. However, patch-clamped cells were selected on the basis of stimulus responsiveness. Among the population of highly responsive cells (those active in >80% of trials, 119 cells total), most cells decoded above chance levels (N = 69 of 119, 58%) and the average decoding performance was 0.74 +/-0.09 (SE), which is comparable to the patch-clamped cell performance of 0.78 +/-0.11.
Among the imaged cells that decoded above chance, the optimal start time of the decoding window was distributed across the entire 500-ms range (mean, 280 +/- 180 ms (SE), N = 985), while the distribution of optimal window lengths tended toward the longest tested windows (mean, 460 +/- 130 ms, Fig. 2H-I). The rate of frame acquisition, which ranged from 9 to 50 Hz, will affect the start time and window length analysis somewhat. However, even among experiments with the fastest frame rates (>20 Hz), start times were 240 +/- 170 ms (N = 491), which corresponds to approximately five imaging frames. Increasing the frame rate further is therefore unlikely to reveal significantly earlier start times. The best decoders, those with decoding performance over 80%, had significantly earlier start times than other decoders (best decoders, mean, 190 +/- 150 ms, N = 64; Wilcoxon rank sum, p = 1.5e-5). Bin sizes for cells achieving the highest decoding performances were not significantly different from those of the rest of the decoders (mean 430 +/- 140 ms, N = 64; Fig. 2F). Thus, the most informative cells decoded stimulus location sooner than other decoders, by starting at an earlier post-stimulus time.

**Setting t = 0**

From trial to trial, there was a variable lag before the population became active. This latency was itself informative of the stimulus, but for a single cell to read out this information an additional “timekeeping” signal would be required. Whether such a signal is present can be debated, but it is clear that a single neuron receives a signal, in the form of synaptic input, indicative of activity in the surrounding population. We characterized whether neurons can decode stimulus location without explicit information about latency using instead timing relative to the population response.

For each trial, we considered activity across all neurons and set t = 0 to the frame at which at least three neurons were co-active (Fig. 3A), which was two standard deviations above the background activity rate. We found that nearly as many cells decode above chance levels as in the stimulus-aligned data (N = 942 of 3326; Fig. 3C). However, decoding performance was lower
than with stimulus-locked alignment. We found a decrease of 4 to 8 percentage points in decoding performance among neurons that exhibited a minimum of 70% (stimulus-locked) decoding performance (Fig. 3D). Additionally, among population-locked decoders, there were fewer very accurate cells (N = 45 with ≥80% decoding performance) than there were among stimulus-locked decoders (N = 63 with ≥ 80% performance).

Because the delay between stimulus and activity onset was eliminated by activity alignment, the best start times for above-chance decoders were earlier than for stimulus-locked alignment (population-locked mean, 250+/−7 ms (SEM) vs. 280+/−4 (SEM) ms for stimulus-locked; Wilcoxon rank sum, p = 3.9e-5; Fig. 3E and 2H). The optimal bin size was also smaller for decoders using population-locked alignment than stimulus-locked alignment (420+/− 5 ms vs. 460+/− 4 ms (SEM), Wilcoxon rank sum, p = 4.3e-5; Fig. 3F and 2I). For the best decoders, window lengths were not significantly different from those of the other above-chance decoders (430+/− 20 (best decoders) vs 420+/− 6 (rest), p = 0.82), but start times were earlier (160+/− 20 ms vs 237+/− 7 ms (SEM), p = 5.6e-5). Thus, temporal alignment to the population resulted in a small decrement in the quality, but not the number, of above-chance decoders and a shorter latency and time window for optimal decoding. In sum, these data suggest that the local population was capable of providing the information necessary to decode stimulus location and also tightened the temporal resolution necessary for optimal decoding.

**Anatomical distribution of decoding performance**

Imaged neurons were located between 125 µm and 825 µm from the pial surface, spanning from pia to lower layer 5. To determine whether there was a relationship between decoding performance and cortical depth, we calculated the fraction of cells that decoded above chance as a function of distance from pia by counting above-chance decoders and total sampled cells in a sliding window of width 50 µm. Cells decoding above chance comprised between 20-25% of sampled cells, regardless of layer (laminar assignment based on distance from pia: L2/3: 128-418µm; L4: 418-588 µm; L5, >588 µm; Lefort et al. 2009. Stimulus-locked decoding, Fig. 4A-B; mean +/- SE across animals: layer 2/3, 21 +/- 4%; layer 4, 27% +/- 5%; layer 5, 22 +/- 4%; for
population-locked decoders, Fig. 5A-B; mean +/- SE: layer 2/3, 22 +/- 4 %; layer 4, 27% +/- 5%; layer 5, 19 +/- 4%.) This range of decoding performances was within the variance across animals (+/- 5%). To determine whether high-performance decoders (Fig. 4B, 5B) were distributed differently from the sampled population, we computed the Kullback-Leibler divergence between the spatial distribution of high-performance decoders and the spatial distribution of all imaged cells. We found no significant difference between the distributions (stimulus-locked, d-KL = 0.54, p = 0.11; population-locked, d-KL = 0.33, p = 0.37, two-sided shuffle test; see Methods). Thus, we find that decoding performance at the single-cell level was not significantly associated with the cell’s distance from pia.

Decoding with larger groups

Moving beyond single cells, we characterized decoding performance as cells are added to the decoding pool. Decoding groups were constructed by drawing neurons from the small set of cells with greater than chance decoding rates, the core decoder population. Depending on the information contributed by each cell, decoding performance may saturate at the level of the most accurate single cell or increase rapidly to the performance ceiling (Fig. 1E). For both stimulus-locked and population-locked alignment, the average decoding performance of a group of core neurons increased with the size of the group, from 65% (+/- 1% SEM over N = 22 fields of view) for single neurons to 78% (+/- 2% SEM, N = 21) for groups of eight using stimulus-locked alignment and from 63% (+/- 1% SEM, N = 22) to 72% (+/- 2% SEM, N = 22) using population-locked alignment (Fig. 4C).

Curves from particular groups have a diversity of shapes, however. A substantial fraction of groups of four or more neurons achieved near-perfect decoding performances (i.e. > 90% correct): 7% (+/- 3% SEM, N = 22 fields of view) of groups of four above-chance decoders and 14% (+/- 5% SEM, N = 21 fields of view) of groups of eight (Fig. 4D). With population-locked alignment, fewer individual cells achieved near-perfect decoding compared to stimulus-locked alignment, and only 2% (+/- 1% SEM, N = 21 fields of view) of four-cell groups of above-chance decoders achieved near-perfect decoding performance (Fig. 4D). To determine whether this
difference between stimulus-locked and population-locked alignment derived from differences in the single-cell performance or differences in how information was combined across neurons, we computed the relative decoding performance of the group, the ratio of group decoding performance to the performance of the best cell in the group (Fig. 4E). Relative decoding performance for both temporal alignments is significantly larger than one, indicating that in both scenarios groups do improve over the best-cell performance by integrating information from the other members of the decoding pool. The relative performance of small groups was slightly higher with stimulus-locked decoding than with population-locked decoding (groups of 4: 5.4% +/- 0.6% (SEM) for stimulus-locked groups, vs. 2.7% +/- 0.8% (SEM) for population-locked, N = 22 slices; ranksum tests: 2 neurons, p = 0.042; 3 neurons, p = 0.053; 4 neurons, p = 0.023; 5 neurons, p = 0.020; 6 or more neurons, p > 0.05). However, we did not find a consistent significant difference. Thus, the lower incidence of near-perfect decoder groups for population-locked alignment appears to be a consequence of drawing from a less informative pool of single cells than the stimulus-locked decoding groups did, not from a difference in how decodable information is built up within the group.

We also analyzed whether decoding performance depended on the location of the group members. Using published laminar boundaries for somatosensory cortex (Lefort et al. 2009), we classified cells by lamina and measured the decoding performance of groups confined to a single layer. We find no significant difference between any pair of layers for any size of group (all p > 0.05, ranksum with Bonferroni correction; Fig. 4F-H and Fig. 5C-E).

Role of noise correlation in decoding

As suggested by the model (Fig. 1E), group decoding performance typically increased with added cells and the overall performance level was strongly dependent on the decodable information carried by the most accurate single cell. Using the imaging data, we examined the role correlation plays in representation of stimulus information by dissecting the performance of pairs of cells. We illustrate the strong potential effect of noise correlation with an example.
from our data in Figure 6A. These two cells displayed strong positive noise correlation that
allowed for decoding of the stimulus identity. We found no significant difference in the pair-
decoding performance of the stimulus-locked versus population-locked responses (Fig. 6B).

Because we were able to image neurons immediately adjacent to one another, we did not
expect that every neuron would carry independent information, and so not every set of two
cells will decode significantly better than one (as in Fig. 1E). The gain in decoding performance
from adding a second cell was significantly greater than zero on average (Fig. 6C; mean 0.033,
sem 0.001, pairs drawn from N = 985 cells with above-chance performance), though a small
fraction (14% stimulus-locked; 20% population-locked) of pairs had negative gains. This
apparent decrement in decoding performance was used as a measure of the noise around the
no-gain condition (Fig. 6D). We identified jointly decoding pairs – i.e. pairs in which decoding
performance substantially improved when using the joint activity pattern – to be those above a
threshold set by comparing with a zero-mean distribution of equal variance (Fig. 6D, gray
distribution). The same threshold (gain > 0.09) was used for stimulus-locked and population-
locked alignment.

Among jointly decoding pairs, each neuron contributed a “piece” of information. To determine
the role of correlated activity in decoding the stimulus, we compared the pair decoding
performance to the trial-shuffled performance. If cells operate independently, then shuffling
will have no effect on decoding performance. Shuffling will increase decoding performance if
trial-by-trial correlations are detrimental, while it will decrease if correlated variability served as
an additional information channel (Fig. 1E, 6D). We found that shuffling tended to decrease
decoding performance, indicating a positive effect of correlation (Fig. 6D; mean 0.028, SEM
0.002, pairs drawn from N=419 unique cells, p <1e-10, signed rank), for both stimulus-locked
and population-locked alignment (Fig. 6E; for population-locked alignment, mean correlation
effect, +0.034, SEM 0.002, pairs drawn from N=373 unique cells, p <1e-10, signed rank). These
pairs were slightly closer in space than other pairs (160 µm +/- 90 µm (SE) vs. 165 µm +/- 90 µm
(SE), \( p < 0.01 \), shuffle test). We also analyzed peristimulus lag, intercellular difference in lag, and fraction of trials active, but found no systematic difference.

Discussion

The classical approach to characterizing the cortical code has been to record the responses of single cells and map out such properties as receptive fields and tuning curves. While informative, this approach has proven difficult to extend to a coding scheme utilized by local cortical microcircuit comprised of interconnected neurons. Considering the strong influence of connectivity on cortical information processing, investigation at the level of the microcircuit is necessary to begin to generate a general framework for the input-output function of the brain.

Here, we have presented a careful quantification of decoding of thalamic inputs in large cortical populations extending across layers and columns. We focus on decoding, rather than a generalized measurement of encoded information, which permits us to characterize what information is accessible to downstream neurons via a plausible biological mechanism based on the linear decoder (Seung and Sompolinsky 1993). Additionally, a decoding approach has the advantage of higher robustness to data limitations (Averbeck et al. 2006) than directly measuring encoded information.

While the decoding task is simple, it is not trivial: The majority of single cells do not decode stimulus information at levels exceeding chance. By contrast, cells selected for reliable stimulus response, including both those imaged and recorded intracellularly, exhibited a much higher decoding performance, reflecting how strongly selection bias can affect measured decoding performance. These data would suggest that there is a subset of neurons that are the key to encoding each ‘piece’ of information in cortex. This is consistent with the finding that only a specific subset of cortical neurons is reliably responsive to whisker stimulation (Yassin et al. 2010) and with sparse coding theories in barrel cortex (Crochet et al. 2011; Ganguli and
Sompolinsky 2012; O’Connor et al. 2010). While categorical differences in stimulus-specific responsiveness exist, the majority of decoders responded to both stimuli and much of the decoding power of a cell derives from a difference in lag in stimulus-driven activity, which is consistent with studies showing a latency code in barrel cortex (Bale and Petersen 2009; Petersen et al. 2001). Previous experiments, including single-unit recordings, have shown that the latency to first spike \textit{in vivo} is on the order of tens of milliseconds (Swadlow 1995, Drew and Feldman 2007, Jadhav et al. 2009). When we broadly sample a cortical population outside of the primary column we find that the first spikes of the population response appear within 50 ms of stimulus onset. Regardless of the exact latency, we found that decoding and discrimination of the stimulus requires further computation – namely, a comparison between the responses – and the best decoders discriminated the stimulus within 200ms of onset. Over a longer time window (up to 500 ms), many additional cells were capable of decoding stimulus identity at levels exceeding chance. An important future direction will be to conduct a similar analysis of cortical circuit activity driven by temporally extended stimuli that fluctuate on naturalistic timescales reflecting the characteristics of whisking behavior.

Cells in the brain may not have independent access to stimulus onset times and must, therefore, compute time zero from locally available measures. We reanalyzed the data using the timing of activity relative to the population onset instead of the stimulus. While it is not clear what exactly sets the clock in cortex, information about population activity is locally available to an individual cell as total synaptic drive, making it a plausible candidate variable. Some of the high-performance decoders lost decoding power when population activity was used for temporal alignment. Among the decoders that remained, the maximally informative window for stimulus discrimination immediately followed population onset. For a single cell, the difference in lags between the two stimuli reflects a different position of that cell in multi-neuronal spatiotemporal sequence triggered by the stimulus. In other words, the stimulus-dependent pattern of circuit activity sets a stimulus-dependent lag at the level of the single cell, so this lag may be used to determine the stimulus identity that may in turn reflect stimulus
location. The partial loss of decoding power when absolute temporal information is removed is consistent with previous reports (Arabzadeh et al. 2006; Panzeri and Diamond 2010).

The role of noise correlation in the cortical code is a topic of much theoretical speculation (Averbeck et al. 2006, Abbott and Dayan 1999; Ecker et al. 2011; Latham and Nirenberg 2005; Schneidman et al. 2003) and little experimental consensus (Cohen and Kohn 2011; Nirenberg and Latham 2003; Osborne et al. 2008; Smith and Kohn 2008; Ecker et al. 2010; Hansen et al. 2012; Snyder et al. 2014). On the basis of multi-unit recordings, noise correlations have not been found to increase the information carried by populations of neurons. However, cells in these recordings are typically separated by hundreds of microns (Adibi et al. 2014; Ince et al. 2013) significantly diminishing the likelihood of synaptic connectivity between units. In contrast, here we do find a significant effect of noise correlation in decoding, among a specific set of pairs: those that decode better than either constituent cell. Such pairs are rare among all recorded pairs (< 1% of 306,456), but comprise about 12% of the pairs drawn from the core population of above-chance decoders. It is not possible to know a priori which subset of cells form the ‘core’ decoding population, and so discovering these pairs requires dense sampling of a sufficiently large population.

Finally, we evaluated decoding performance as we built up groups of up to eight above-chance decoders. Regardless of how we temporally aligned the data, there is an increase in decoding performance with each cell added to the group. However, improvements saturate quickly: the third cell in the group improves the performance much less than the second cell did, the fourth less than the third, and so on. On average, combining more than five cells no longer improved the group performance. This saturation is partially due to the growing fraction of groups that have achieved near-perfect performance, and, along those lines, a natural next step is to make the decoding task more challenging by increasing the complexity of the stimulus space. By this logic, the 84% of groups that don’t reach near-perfect decoding are simply lacking some stimulus information. Alternatively, this saturation may be a limitation of the type of decoder we used. Our simple model demonstrated that, typically, decoding performance would saturate
as cells are added to the pool. Other decoders, which permit XOR-type logic, might display supra-linear scaling with N, indicating a set of neurons that are individually uninformative about the stimulus but have a highly informative joint activity pattern. In this case, the helpful effect of correlation we observed among some pairs might be a suggestion of truly synergistic decoding at the population level.

Acknowledgements

We thank the MacLean lab for comments on the manuscript. This work was supported by the DANA Foundation (AJS; JNM), the National Science Foundation CAREER Award 0952686 (JNM), and the Mary-Rita Angelo Fellowship (AJS)

References


Beierlein M, Connors BW. Short-term dynamics of thalamocortical and intracortical synapses


Luczak A, MacLean JN. Default activity patterns at the neocortical microcircuit level. *Front Integr Neurosci* 6 (30), 2012.


O’Connor DH, Peron SP, Huber D, Svoboda K. Neural activity in barrel cortex underlying

Packer AM, Yuste R. Dense, unspecific connectivity of neocortical parvalbumin-positive

Panzeri S, Diamond ME. Information carried by population spike times in the whisker sensory
cortex can be decoded without knowledge of stimulus time. *Front Syn Neuro* 2: 1-17, 2010.

Perin R, Berger TK, Markram H. A synaptic organizing principle for cortical neuronal groups.

Petersen RS, Panzeri S, Diamond ME. Population coding of stimulus location in rat

Pillow JW, Shlens J, Paninski L, Sher A, Litke AM, Chichilnisky EJ, Simoncelli EP. Spatio-
temporal correlations and visual signalling in a complete neuronal population. *Nature* 454:

Pouget A, Zhang K, Deneve S, Latham PE. Statistically efficient estimation using population


Sadovsky AJ, Maclean JN. Scaling of topologically similar functional modules defines mouse

Sadovsky AJ, Kruskal PB, Kimmel JM, Ostmeyer J, Neubauer FB, Maclean JN. Heuristically
optimal path scanning for high-speed multiphoton circuit imaging. *J Neurophysiol* 106:

Schneidman E, Berry MJ, Segev R, Bialek W. Weak pairwise correlations imply strongly

Schneidman E, Bialek W, Berry M. Synergy, redundancy, and independence in population

Schneidman E, Puchalla JL, Segev R, Harris RA, Bialek W, Berry MJ. Synergy from Silence in a

Seung HS, Sompolinsky H. Simple models for reading neuronal population codes. *PNAS* 90:


Figure Legends

Fig. 1: Experiment and overview of decoding analysis. A: Thalamocortical slices including S1BF. Stimulating electrodes are placed in separate locations in thalamus (orange, stim a and blue, stim b), and cortical cells in barrel field (circle) are recorded using two-photon calcium imaging. For some experiments (N = 7) electrophysiology is also obtained from two cortical cells, indicated by patch pipettes over cortex. B: Imaged field of view. Cells are loaded with fluorescent dye Fura-2AM. C: Sample whole-cell patch clamp recordings, population activity, and average population activity for responses to the two stimuli. Arrows mark time of stimulus delivery. Cells are sorted by onset time for stimulus a. Activation order is different for stimulus b. D: Overview of decoder task. Stimulus location (a or b) is decoded from the activity of cells or groups of cells. We demonstrate the decoder performance in a simple model in which the stimulus-dependent response (orange, stim a and blue, stim b) is Gaussian-distributed. E: Decoding performance of a group of neurons depends on the number of cells in the group and on the correlations between cells. Pairwise correlation is illustrated by response density plots, with the 95%-confidence region outlined. Negative correlation (black line) between cells improves the group performance relative to the independent (gray line) case, while positive correlation degrades group performance (light gray line).

Fig. 2: Decoding stimulus location with single neurons. A: Spikes recorded from a patch-clamped cell in response to stimulus a (top) and stimulus b (bottom). Arrows indicate time at which the stimulus was delivered. B: Histogram of decoding performances from patch-clamped cells performing above chance (black) and at chance (gray). Chance-level
performance is determined by comparison with the stimulus-shuffled performance
distribution. For details, see main text. C: For the cell shown in A, decoding performance of
linear decoder fit to activity averaged over a bin size of 0.1s (solid black) or 0.01 s (dashed
black). Decoding performance was higher using the 0.1-s bin size. D: Fluorescence trace
from a single neuron with three detected events (orange bars). E: Histogram of activity for a
single cell. Activity values are obtained by averaging the deconvolved fluorescence trace
over a window 450 ms wide starting a time 390 ms after the stimulus onset. This cell had a
higher response to stimulus a (orange) than to stimulus b (blue). F: For the cell shown in D
and E, performance for multiple decoders over a range of post-stim times and bin sizes. G:
Distribution of highest decoding performance found across all averaging windows for each
imaged cell. Black bars indicate cells with decoding performances significantly above
chance levels. H: Distribution of optimal start times among all cells with decoding
performance above chance and less than 0.8 (gray). Bars (green) superimposed show a
histogram for the best decoders, with performances greater than 0.8. High decoding
performances are associated with earlier start times. I: Distribution of optimal bin sizes
among all cells with decoding performance above chance and less than 0.8 (gray). Bars
(green) superimposed show a histogram for the best decoders, with performances greater
than 0.8.

Fig. 3: Decoding with single cells using population activity for temporal alignment. A: Multi-
trial population activity raster. Solid horizontal lines separate single-trial responses of a
subset of 40 cells from the population. Stimulus was delivered at t = 0; color indicates
stimulus identity. Trials are sorted by stimulus identity, but were interleaved in the
experiment. For population-locked decoding, activity is aligned based on the onset of
population activity (vertical black lines) on a trial-by-trial basis. This is generally different
from the stimulus delivery time. B: Distribution of stimulus-specific latency to population
activity across all trials. C: Decoding performance of single cells using population-locked
timing, with above-chance cells in black and at-chance cells in gray. D: Difference in
decoding performance between stimulus- and population-locked responses. Decoding
performance typically decrease by 0.04 to 0.08 with population alignment. E: Optimal start
time distributions for above-chance cells (gray). Bars (green) superimposed show a
histogram for the best decoders, with performances greater than 0.8. F: Optimal bin size
distributions for above-chance cells. Optimal start times and bin sizes are earlier and
shorter with population-locked timing than stimulus-locked timing (cf. Fig. 2H, I).

**Fig. 4: Distribution of stimulus-locked decoders across lamina.** A: Fraction of recorded neurons
that decoded above chance (stimulus-locked) as a function of distance from pia. Shaded
region marks +/- 1 SE computed across experiments. B: Locations and decoding
performance of the best decoders (>80% success). Distribution is not significantly different
from the overall distribution of sampled cells. Bars at right show approximate laminar
boundaries. C: Decoding performance for groups drawn from the core population of
above-chance cells. Performance increases with the number of cells. D: Fraction of each
group of cells that decodes with 90% or higher decoding performance. E: Group decoding
performance relative to the best individual cell performance within the group. F-H: For
stimulus-locked decoding, versions of C-E with groups composed of cells located within a
single layer. See Fig. 5 for population-locked decoders. No significant differences were
found between any layers on any of these measures.

**Fig. 5: Distribution of population-locked decoders across lamina** A: Fraction of neurons that
decoded above chance (population-locked) as a function of distance from pia. Shaded
region marks +/- 1 SE computed across experiments. B: Locations and decoding
performance of the best decoders (>80% success). Distribution is not significantly different
from the overall distribution of sampled cells. Bar at right shows approximate laminar
boundaries. C-E: Decoding with groups composed of cells located within a single layer. C:
Decoding performance for groups drawn from the core population of above-chance cells. D:
Fraction of each set that decodes with 90% or higher decoding performance. E: Group
decoding performance relative to the best individual cell performance within the group. No
significant differences were found between any layers on any of these measures.

**Fig. 6: Role of noise correlation in decoding performance.** A: Plot of pairwise activity, with cell 1 and cell 2 responses on x- and y-axes, and histograms (as in Fig. 2E) of single-neuron activity off axis. For this cell pair, the activity pattern across the two cells could be decoded with 90% success. Individual cells were only 60-65% successful. B: Decoding performance of pairs drawn from the core above-chance population using stimulus-locked (top) and population-locked (bottom) timing. C: Histogram of performance gain from the first and second cell using stimulus-locked (top) and population-locked (bottom) timing. First-cell gain is the difference between single-cell performance and chance (0.5), and second-cell gain is the difference between pair performance and single cell performance (see text). As expected from the model (Fig. 1E), the first-cell gain is generally higher than the second-cell gain. D: Histogram of effect of noise correlation on decoding performance, measured as the difference between true decoding performance and trial-shuffled decoding performance. Positive values indicate an improvement in decoding when correlations are taken into account. Pairs with significant second-cell gains in decoding performance showed on average a positive effect of correlation. Filled gray outline shows zero-mean normal distribution with matched variance, a null model for no overall effect of noise correlation. E: Box plot of correlation effect for stimulus-locked decoders (as in D) and population-locked decoders. Both show a significant positive effect of correlation.
A

B

C

D

E

[Diagram A: Schematic of a brain region with stimuli labeled as `stim a` and `stim b`.

Diagram B: Image of neural activity with scale bar indicating 200 µm.

Diagram C: Time series plots showing `stim a` and `stim b` responses, with detected events and population average.

Diagram D: Decoder task involving `stim a` and `stim b`.

Diagram E: Graphs showing decoding performance with different group sizes and correlation coefficients (`r` values).]
Fraction of cells that decode above chance

Distance from pia (μm)

Performance (best decoders)

Decoding performance

Near-perfect decoding fraction

Performance relative to best cell
Fraction of cells that decode above chance

Distance from pia (μm)

Best decoders performance

Group size

Decoding performance

Fractional improvement of group over best cell

Population-locked