**Population-Level Neural Codes Are Robust to Single-Neuron Variability from a Multidimensional Coding Perspective**

**Highlights**
- Correlated trial-by-trial variability can impair the reliability of neural codes.
- Multidimensional codes of stimulus orientation are highly stable over weeks.
- Multidimensional correlations restrict variability to non-coding directions.
- Up to 50% of single-trial, single-neuron “noise” is predictable with correlations.

**In Brief**
Montijn et al. show that compared to independent codes, the accuracy of V1 population codes improves when using pairwise correlations but improves further when using multidimensional correlations. This may be achieved because high-dimensional correlations restrict trial-by-trial variability to directions that are perpendicular to the trajectories that encode stimulus features.

**Authors**
Jorrit S. Montijn, Guido T. Meijer, Carien S. Lansink, Cyriel M.A. Pennartz

**Correspondence**
j.s.montijn@uva.nl (J.S.M.), c.m.a.pennartz@uva.nl (C.M.A.P.)

**Article**
Montijn et al., 2016, Cell Reports 16, 2486–2498
August 30, 2016 © 2016 The Author(s).
http://dx.doi.org/10.1016/j.celrep.2016.07.065
Population-Level Neural Codes Are Robust to Single-Neuron Variability from a Multidimensional Coding Perspective

Jorrit S. Montijn,1,3,4,* Guido T. Meijer,1,3 Carien S. Lansink,1,2 and Cyriel M.A. Pennartz1,2,*

1Swammerdam Institute for Life Sciences, Center for Neuroscience, Faculty of Science, University of Amsterdam, 1098 XH Amsterdam, the Netherlands
2Research Priority Program Brain and Cognition, University of Amsterdam, 1098 XH Amsterdam, the Netherlands
3Co-first author
4Lead Contact
*Correspondence: j.s.montijn@uva.nl (J.S.M.), c.m.a.pennartz@uva.nl (C.M.A.P.)
http://dx.doi.org/10.1016/j.celrep.2016.07.065

SUMMARY

Sensory neurons are often tuned to particular stimulus features, but their responses to repeated presentation of the same stimulus can vary over subsequent trials. This presents a problem for understanding the functioning of the brain, because downstream neuronal populations ought to construct accurate stimulus representations, even upon singular exposure. To study how trial-by-trial fluctuations (i.e., noise) in activity influence cortical representations of sensory input, we performed chronic calcium imaging of GCaMP6-expressing populations in mouse V1. We observed that high-dimensional response correlations, i.e., dependencies in activation strength among multiple neurons, can be used to predict single-trial, single-neuron noise. These multidimensional correlations are structured such that variability in the response of single neurons is relatively harmless to population representations of visual stimuli. We propose that multidimensional coding may represent a canonical principle of cortical circuits, explaining why the apparent noisiness of neuronal responses is compatible with accurate neural representations of stimulus features.

INTRODUCTION

The presentation of stimulus features modulates the responses of single neurons in sensory cortex such that the outside world is represented in the activation pattern of neuronal populations. However, the activity of single neurons shows substantial variability in spike rate and timing across repeated presentations of the same stimulus (Faisal et al., 2008). This variability is often called neural noise and poses a problem: how can animals react quickly and reliably to sensory input when the stimulus representation would already be noisy at the first stage of cortical processing? It has been proposed that neural circuits solve this problem by combining information from multiple neurons into a population code. If the variability of neuronal responses were independent, higher precision of stimulus representation would be achieved by combining the responses of more neurons (Beck et al., 2008; Knill and Pouget, 2004; Ma et al., 2006). However, neurons are often correlated in the variability of their response to the same stimulus, which means that simple averaging is insufficient to achieve maximal precision (Averbeck and Lee, 2006; Hansen et al., 2012; Lee et al., 1998; Vinje and Gallant, 2000).

The interdependency between responses of pairs of neurons (i.e., noise correlations [NCs]) has been proposed to influence the amount of information that can be extracted from population codes in different ways, ranging from being beneficial to being mostly irrelevant or harmful (Averbeck et al., 2006; Cafaro and Rieke, 2010; Cohen and Kohn, 2011; Cohen and Maunsell, 2009; Ecker et al., 2011; Fiscella et al., 2015; Herrero et al., 2013; Montijn et al., 2014; Seriès et al., 2004). Another aspect that complicates the study of NCs is that these correlations can be heterogeneous in their size and effects on population codes, depending on factors such as the nature of the presented stimulus features, correlations between neurons other than the pair being studied, and differences in general arousal state (Chebli et al., 2016; Ince et al., 2013; Jazayeri and Movshon, 2006; Miller et al., 2014; Moreno-Bote et al., 2014; Pitkow et al., 2015; Schölvinck et al., 2015). Therefore, one of the most relevant challenges in neurophysiology is explaining how accurate sensory representations can be generated by neuronal populations in the face of instantaneous single-neuron response fluctuations.

Pairwise dependencies might be important for neuronal populations that represent sensory information, but it has been hypothesized that the underlying structure of neural responses may be multidimensional—i.e., dependent on interactions among more than two neurons (Franke et al., 2016; Kanitscheider et al., 2015; Latham et al., 2003; Pasupathy and Connor, 2002; Pillow et al., 2008; Schneidman et al., 2006). A related computational problem in natural vision holds that many features are present simultaneously, and these features are thought to be represented with high fidelity by ensembles of neurons in early sensory areas (Saddeley et al., 1997; Eichhorn...
Such concurrent representation of multiple stimulus features is known to exist in higher visual and non-visual brain areas (DiCarlo et al., 2012; Sigala et al., 2008; Stokes et al., 2013), but it is unknown whether multidimensional coding is used in primary visual cortex (V1) to enable efficient representations of natural scenes. Moreover, long-term stability of these codes and their correlation structure is required for many aforementioned models of the cortex to be neurophysiologically plausible, and the extent of this temporal stability is as yet unknown.

We therefore set out to investigate two important factors in the interaction of correlated variability with population coding: (1) the stationarity of correlations over time and (2) the presence and potential use of higher-dimensional correlations in population coding. For instance, can responses of neuronal triplets be described well by the pairwise interactions of its members? If not, what might be the use of higher-dimensional interactions? We recorded long-term (>4 weeks) neuronal responses to drifting gratings and natural movies from the same populations of L2/3 (cortical layer 2/3) V1 neurons in awake mice using GCaMP6 calcium imaging (Chen et al., 2013). Our data show that neuronal responses are variable across trials but relatively stable across days. We observed that multidimensional correlations are of critical importance for the efficacy of population codes by restricting variability to those directions in neural space that are perpendicular to the axes coding for stimulus features. Moreover, these correlations can be used to predict up to half of the instantaneous noise in single-neuron activity. We conclude that much of the trial-by-trial fluctuation shown by individual neurons is not noise but might be functionally important for the neuronal population code of sensory input.

RESULTS

Neuronal Responses Are Variable across Trials, but Tuning Is Stable across Days

We performed chronic GCaMP6m imaging in V1 L2/3 of awake mice to study the variability in responses of neuronal populations to drifting gratings (n = 9 mice recorded long term, over 2–5 weeks) and natural movies (n = 4 mice recorded long term; n = 5 mice recorded short term, on a single day; Figure 1A). We found that the orientation tuning and responses to natural movies of many neurons were reliable over this period (Figures 1B–1F). However, some neurons showed weak or non-orientation-tuned responses across trials and recording sessions. Such neurons might be located at the edge of the focal plane in some recording sessions or might be non-responsive to visual stimulation by our drifting gratings. We therefore excluded these from further analyses and included only neurons that showed a non-random orientation preference across days (for long- and short-term recordings, respectively, on average 48% of 43–158 neurons and 55% of 130–181 neurons per mouse were consistently tuned to orientation across recording sessions; Figures S1 and S2). Although tuning to the orientation of drifting gratings was stable over days for many neurons, the responses of these stable neurons for subsequent trials of the same orientation were still variable (Figure 1D). We hypothesized that some of this variability might be related to the behavioral state of the animal. Because pupil size is positively correlated with arousal (Bradshaw, 1967; Coull et al., 2004; Vinck et al., 2015), we performed eye-tracking during neurophysiological recordings (Figure S3) and found that during trials in which the mouse’s pupil size was large, neuronal responses showed a higher variability (SD) across repetitions to the preferred orientation, as well as increased NCs, despite similar levels of mean activity (Figure 1E; Figure S3).

We aimed to better quantify the observation that variability can be high across trials mere seconds apart, whereas tuning similarity is stable across days (Figure S4). We therefore calculated NC and signal correlation (SC) matrices for each recording session. This splits the neuronal responses into a signal component that encodes grating orientation (Figures 2A and 2B) and a noise component that reflects trial-by-trial fluctuations (Figures 2C and 2D). The stability of the population response can be approximated by calculating the Pearson correlation between pairs of SC matrices recorded during different sessions. Analysis of SC versus inter-recording time in days (Figure 2B) yielded three main results. First, when two sessions were recorded on the same day (inter-recording time was 0 days), the correlation between these recordings was relatively high (r = 0.52 ± 0.088, mean ± SD) but clearly far from identical (r = 1.0). This means that pairwise neuronal responses to the same stimuli across repetitions are largely defined by short-term fluctuations that occur on the order of minutes to hours. Second, longer intervals, on the order of days, decrease correlations at a slower pace. Even after an interval of a month, the correlations were well above zero (one-sample t test across four recording pairs with largest time intervals, mean interval 29.8 days, p < 0.01; Figure 2B). After fitting the data with exponential decay functions, we found that SC half-lives were similar across animals (mean ± SEM half-life across mice, 37.3 ± 10.8 days). Third, when repeating these analyses for across-recording similarity in NCs, the effects were comparable in form and magnitude. We found that r was 0.48 ± 0.097 (mean ± SD) for recordings made on the same day and that the mean ± SEM of half-life across mice was 41.1 ± 6.0 days (Figures 2D and 2E). These results suggest that correlation structures are relatively stable over time, showing a slow decay that is in line with previous reports of multi-day stability of orientation tuning (Chen et al., 2013; Lütcke et al., 2013; Mank et al., 2008). However, these results also indicate that there are large fluctuations in pairwise NCs and SCs on the order of minutes to hours (cf. Figures S4 and S5).

Next, we asked whether higher-dimensional representations would be similarly stable over days. We calculated a multidimensional population code similarity, based on the distance between neural representations of the same orientation on two different recording days (the number of dimensions in this approach equals the number of neurons; see Experimental Procedures). The similarity has a maximum of 1.0 and is normalized by the average trial-by-trial variability in multidimensional representations of the same orientation. A value of 0.0 indicates that representations of the same orientation are as distant as the mean trial-by-trial variability within those recordings. First, we performed this procedure for pairs of neurons, each time averaging across 100 randomly selected groups. After fitting the data with...
an exponential decay function, we found that this metric's similarity was 0.51 ± 0.04 (mean ± SD) for recordings on the same day and that the mean ± SEM of the half-life across mice was 62.0 ± 9.2 days (Figure 3A). Next, this procedure was repeated with different group sizes (triplets, quadruplets, etc.), yielding a half-life for each dimensionality. Analysis of half-lives as a function of dimensionality showed that high-dimensional representations of visual stimulus orientation are more temporally stable than pairwise representations by the same populations (Figure 3B). We observed a consistent within-animal effect, in which high-dimensional codes were more stable than pairwise codes (t test of high-dimensional half-lives normalized to pairwise; p < 0.001; Figure 3C), although raw half-lives were quite variable across animals (mean ± SEM of half-life decay times; at pairwise: 62.0 ± 9.2 days; at maximum dimensionality: 74.4 ± 10.3 days).

Neuronal Populations Encode Stimulus Orientation in Higher-Dimensional Space

Earlier work has demonstrated that non-zero spike-count correlations between pairs of neurons can lead to higher-dimensional network states that are hard to predict from pairwise correlations (Schneidman et al., 2006). Therefore, the multidimensional correlation structure of a neuronal population could in principle contain more information than might be apparent from the responses of pairs of neurons. It is unknown whether neurons in mouse visual cortex encode stimuli in a lower-dimensional (e.g., pairwise) way or whether stimuli are represented in a multidimensional response space that cannot be inferred from lower-dimensional statistical interdependencies. To investigate the potential of multidimensional population coding, we created a Mahalanobis-distance-based decoder that assumes multivariate Gaussian responses and can be used for any number of dimensions (i.e., a variant of a quadratic discriminant analysis; Figure 4A; see Figure S6 for analysis of this assumption).

We used Mahalanobis space because this normalizes all variability, even across multiple dimensions and regardless of its direction (see Experimental Procedures). This means that Mahalanobis space automatically incorporates multidimensional correlations. The variability normalization also means that decision boundaries between stimulus classes are always linear in
Mahalanobis space (but not necessarily in non-normalized response space; e.g., cf. Figures 4A and 4E; De Maesschalck et al., 2000). This simplifies the neural computations necessary to optimally extract information from a population code, assuming that neural circuits can perform response normalization (Carandini et al., 1997; Lee and Maunsell, 2009; Montijn et al., 2012; Reynolds and Heeger, 2009; Ringach, 2010) and decorrelation (Wiechert et al., 2010).

Using this decoder, we performed a bootstrapping procedure of orientation decoding for all dimensionalities, in which the dimensionality is defined as the number of cells within one randomly drawn group of neurons. We integrated information ranging from 1 to 100 groups (i.e., samples) of neurons of different sizes (i.e., dimensionalities), ranging from 2 to 14 neurons per group (see Experimental Procedures; Figure 4B). This analysis showed that the decoder’s performance saturated ~60–80 randomly drawn groups of neurons for all dimensionalities and all mice, so we performed all further analyses with the integration of 100 random groups for each dimensionality.

To quantify the effect of experimentally measured correlations on neuronal population codes, we compared the decoder’s performance on the real data to its performance on shuffled datasets, in which trials were randomized across repetitions of the same stimulus type. This procedure preserves stimulus tuning but destroys NC structures and therefore allows the identification of effects that are only due to NCs of a certain dimensionality. For independent decoding (i.e., dimensionality 1) there was no difference between shuffled and non-shuffled datasets (orientation decoding accuracy, mean ± SEM across animals: 49.5% ± 4.1% for shuffled, 50.5% ± 3.6% for non-shuffled; errors corrected over shuffled: 1.2% ± 2.0%; paired t test, p = 0.336), but the decoding of stimulus orientation gradually improved based on the higher-dimensional structure present in the recorded data (orientation decoding accuracy at dimensionality 15: 44.7% ± 4.2% for shuffled, 53.2% ± 4.0% for non-shuffled; errors corrected over shuffled: 15.8% ± 2.0%; paired t test, p < 0.001; Figure 4C).
We fitted a half-logistic growth function to the observed performance across dimensionalities 2–15 and calculated the asymptotic performance that would hypothetically be reached (see Experimental Procedures). Asymptotic performance showed an even larger difference between shuffled and non-shuffled performance than independent and pairwise performance, suggesting that V1 neuronal populations encode unique information in high-dimensional space that cannot be inferred using lower-dimensional representations (Figure 4D).

This effect is not due to simply taking information from more neurons; if this were the case, decoding performance for any dimensionality (especially the low-dimensional ones) would not saturate /C24 60–80 random groups (Figure 4B); more importantly, there would be no difference between shuffled and non-shuffled decoding performance (Figure 4C). This within-dimensionality saturation effect, combined with the across-dimensionality increase in performance, shows that additional information on stimulus orientation is encoded in higher-dimensional neuronal response space that is not present in a lower-dimensional space. Moreover, the effect was present when using a range of time windows, showing that multidimensional coding is not dependent on particular epochs during stimulus processing (Figure S6L).

**Multidimensional Response Variability Is Structured to Reduce Impairment of Orientation Coding**

As with any classification problem, variability orthogonal to a decision boundary increases the likelihood that a stimulus is misclassified, while variability parallel to decision boundaries is irrelevant for the classification of a stimulus pair (Figure 4E). We therefore tested whether neuronal circuits might be more variable along dimensions that are not relevant for coding primary stimulus features (i.e., orientation for V1; Figure 4F). In a higher-dimensional space with N neurons, all instances of a stimulus feature (i.e., orientation) can be captured within a single curve. In this case, a line tangential to the orientation coding curve exists for each orientation. Movement along this line corresponds to a shift in the encoded stimulus orientation, while all N – 1 other possible directions are irrelevant for encoding orientation of that stimulus. These other dimensions may then be used to represent other stimulus features, such as contrast and spatial frequency, or for modulatory effects, such as attention, arousal, or other factors, without interfering with the encoding of orientation (Figure 4G).

We tested the non-uniformity of variability in higher-dimensional space by calculating the across-repetition variability orthogonal and parallel to decision boundaries for all pairs of adjacent stimulus orientations (see Experimental Procedures). As before, we compared this variability to the shuffled variability, in which correlation structures are destroyed. Our data show that variability is higher parallel than orthogonal to decision boundaries across all animals, suggesting that variability occurs more in directions that do not impair orientation coding than in those that do (paired t test, n = 9, orthogonal versus parallel variability, p < 0.005; Figure 4H; Figures S7A–S7D).

**Single-Trial Natural Scene Decoding Is More Reliable when Using High-Dimensional Correlations**

To test whether our previous observations apply to natural scenes that contain higher-order statistical visual features not present in drifting gratings, we presented movies from the BBC’s Earthflight (Winged Planet)–Condor Flight School (Figure 5A) to nine mice (see Supplemental Experimental Procedures). These movies consisted of four distinct scenes that elicited reliable time-locked responses across repetitions over several weeks (Figure 5B). Similarly to orientation decoding, we found a saturation of accuracy for scene decoding /C24 60–80 randomly drawn groups of neurons (Figure 5C) and performed all further analyses integrating information from 100 random groups of neurons.
The improvement in decoding accuracy relative to the shuffle control also increased for natural scenes as a function of dimensionality, confirming the results for oriented gratings (Figure 5D). However, for natural scenes, we found that low dimensionalities (up to $\sqrt{24}$) showed worse decoding performance for unshuffled than for shuffled datasets (scene decoding accuracy for dimensionality 1, mean ± SEM across animals: 84.0% ± 6.7% for shuffled, 81.5% ± 6.2% for non-shuffled; paired t test, p < 0.05), suggesting that lower-dimensional correlations impair natural scene coding (e.g., they might reflect common input noise; see also Abbott and Dayan, 1999; Ecker et al., 2010; Kohn and Smith, 2005; Sompolinsky et al., 2001; Zohary et al., 1994). Higher dimensionalities did not show this effect (scene decoding accuracy for dimensionality 1, mean ± SEM across animals: 88.4% ± 37.6%; paired t test, p < 0.05).
dimensionality 40: 79.4% ± 6.4% for shuffled, 80.7% ± 6.4% for non-shuffled; errors corrected over shuffled: 10.3% ± 3.8%; paired t test, $p < 0.05$), suggesting more reliable scene decoding when taking into account higher-dimensional correlations (Figure 5E). As for drifting gratings, we found that for distinct natural scenes, the variability orthogonal to decision boundaries was lower than it was parallel to these boundaries (paired t test, $p < 0.05$; Figure 5F). Therefore, the neural code representing the outside world should also differ over time within, not just between, natural scenes. The multidimensional neural representation of the 20-s-long stimuli can therefore be visualized as a curved trajectory through neural space (Figure 6A). The farther apart two points are on this curve, the better segregated are their corresponding neural representations. One way to study the reliability of such representations over repetitions of the same sequence is to perform a cross-validated decoding procedure on all time points (single data acquisition frames) and trial repetitions. This yields a confusion matrix that can be used to visualize the decoder’s performance (Figure 6B). The mean squared errors (MSEs) of this example animal’s confusion matrices suggest an improved accuracy with higher dimensionality (independent, MSE = 131 ms, pairwise; MSE = 126 ms; maximum dimensionality, MSE = 99 ms; Figures 6B–6D).

However, the MSE is sensitive to outliers, so we proceeded to quantify the decoder’s performance with the temporal inaccuracy as full width at half maximum (FWHM) of the temporal uncertainty around the actual frame. We compared this uncertainty...
for unshuffled datasets to shuffled datasets and, as before, found that while independent and pairwise performance was better with shuffled than with unshuffled datasets, maximum-dimensional performance was higher with the correlation structures intact (paired t tests across animals, n = 9; independent-pairwise, p < 0.005; independent-asymptote, p < 0.005; pairwise-asymptote, p < 0.005; Figure 6E). We conclude that high-dimensional correlations enhance population code accuracy for drifting gratings, as well as scene-based and instantaneous time-based representation of natural movies. The increased temporal decoding accuracy using higher population response dimensions suggests that stimulus classes (i.e., frames in the natural movie) are well separated, but this does not necessarily mean that noise is predictable.

We hypothesized that random fluctuations in the stimulus-driven response at the level of single neurons (i.e., neural noise) might be predictable when higher-dimensional interrelations among neurons are taken into account. We therefore calculated for each trial and neuron the most likely variability in dF/F0 activity based on the activity of all other neurons at that point in time (see Experimental Procedures; Figures 7A–7E). When we predicted this instantaneous neural noise for different dimensionalities, we found that pairwise correlations could be used to explain about 5% of the trial-by-trial variability in neuronal responses (Figures 7F–7H). However, at maximal dimensionalities, almost half (~45%) of all instantaneous trial-by-trial neural noise was predictable (Figures 7F–7H). This value is likely a lower bound, because most factors that could influence the predictability (such as measurement noise) would decrease its value. We therefore conclude that almost half of all observed trial-by-trial fluctuations in single-neuron responses do not constitute noise, in the sense that noise would be random and unpredictable. Single-neuron fluctuations are strongly correlated in higher-dimensional space to the whole of the local neuronal population in which the neuron is embedded.

**DISCUSSION**

We found that neuronal responses are variable across trials but that the statistics of neuronal responses (i.e., orientation preference, structure of correlations, and multidimensional population codes) are relatively stable across days (Figures 1, 2, and 3). The relatively long-term stability of the population code makes it neurophysiologically plausible that neuronal populations may take into account higher-dimensional neural response interdependencies in the inputs they receive. We hypothesized that information about stimuli might be encoded not only within pairwise response relations between neurons but also within higher dimensions of population codes. An analysis of multidimensional responses showed that V1 populations encode stimulus orientation or the identity of natural scenes more reliably in higher dimensions and that the shape of correlational dependencies...
with dimensionalities of 10–20 is structured so as to reduce response variability that might otherwise impair orientation coding (Figures 4, 5, and 6). Moreover, we found that neural noise apparent at the single-neuron level becomes quite predictable when analyzed through higher-dimensional population codes (Figure 7). While single neurons might appear noisy, trial-by-trial fluctuations are thus relatively harmless and predictable when viewed from a population perspective.

One important potential confound for these results is non-stationarity of responses across multiple recording days. The exponential decay of SCs and NCs we observed (Figure 2) could theoretically be caused by a slow decrease in signal-to-noise ratio (SNR) across days. A decrease in SNR across time, for example, due to GCaMP overexpression, would bias correlation values toward zero as time progresses. However, if this were the case, we would observe a relation between time after recording start and inter-recording correlations: two sessions recorded in the first week of the experiment would show a higher correlation than two sessions recorded in its last week. As can be seen in Figures 2B and 2D, this was not the case. The dark blue (near the start of the experiment) and dark red points (near the end of the experiment) are intermixed, showing that SNR reduction over time...
cannot explain the observed exponential decay. Moreover, a more exhaustive analysis based on neuropil contamination confirmed that our results were robust and not influenced by neuropil signals (Figure S5).

A non-stationarity in responses across days may also confound the decoding and noise prediction results. The predictability of noise in neuronal responses might be high only because of systematic changes in neuronal responses across days. To address this issue, we recorded data across several weeks, as well as on a single day. We tested whether noise predictability was different for across-days datasets and within-day datasets, but we found no difference between the two (two-sample t test, \( p = 0.915 \), not significant [NS]). This argues against the noise prediction being dependent on slow, long-term changes in population responses. In contrast, it shows that noise predictability is robust in the face of these slow changes and that fluctuations in neuronal activity can be predicted over short timescales (i.e., several hours; pooling recordings from a single day) and long timescales (i.e., weeks; pooling recordings from different days).

Finally, an important confound may be the undersampling of response distributions when using high dimensionalities. Most of our analyses showed a saturation of effects at high dimensionalities. However, the absence of further increases in performance above these ranges does not necessarily mean that correlations of higher dimensionalities do not enhance population coding efficacy; rather, it means that we were likely confronted with undersampling of multidimensional response distributions, given the size of the datasets available. To minimize this potential confound, we always compared shuffled to non-shuffled datasets, because bias due to undersampling should be equal for both (see also Figures S6I–S6K for a comparison using greedy classifiers). Nonetheless, the scope of our results should only be taken to pertain to L2/3 neuronal populations of a limited size (≈20–40) in mouse V1.

Probabilistic codes have been proposed in previous work to enable optimal cue combination, to use metabolic energy efficiently, and to provide a framework for integrating learning experiences (Deneve et al., 1999; Knill and Pouget, 2004; Ma et al., 2006; Beck et al., 2008; Denève and Machens, 2016). Our observations are not mutually exclusive with a Bayesian interpretation but instead provide a multidimensional interpretation of probabilistic population codes. Variability in neuronal activity may translate into increased distance of the instantaneous population response to the multidimensional curve that represents the continuum of stimulus orientations (Figure 4G). In other words, low certainty about stimulus features might translate into increased distance parallel to decision boundaries. The classical Bayesian interpretation of population codes is then the projection of this multidimensional representation onto a one-dimensional firing rate axis, where each neuron is a single point. However, this multidimensional coding framework is currently a hypothesis; to test this proposal, future research will have to be performed to assess the dependence of parallel distance on stimulus reliability.

Our data show that variability orthogonal to decision boundaries (impairing stimulus discrimination) is lower than it is parallel to these boundaries (Figures 4H and 5F). We showed this to be the case for the orientation of drifting gratings and for natural scenes, but in addition to orientation (and the non-specific bulk of features present in natural scenes), many other well-defined features, such as spatial and temporal frequency, may show this non-uniform variability. For an ensemble of N neurons, there are as many orthogonal directions in which stimulus features could be encoded independently without interference. In the case of V1, this could mean not only that orientation representations are less variable in multidimensional space than expected from a random distribution but also that all canonical stimulus features encoded by V1, such as contrast and temporal frequency, may be encoded in multidimensional space in a way that reduces variability along these coding curves.

This multidimensional orthogonal coding principle may be extended to other sensory areas, for example, the auditory cortex, where the properties of pitch and spatial location may be encoded in a way that minimizes variability along their coding curves. The orthogonal-coding principle proposed here is distinct from other models that rely on multidimensional attractor points or lines (Latham et al., 2003). Attractor points are stable nodes in multidimensional space, and intermediate states are by definition transient and unstable. Our alternative coding scheme predicts that such intermediate states around the trajectory representing a continuous feature (e.g., grating orientation) are meaningful by encoding the relative unreliability of information about the feature in question. If the multidimensional orthogonal coding principle is a canonical coding feature in cortical circuits across the brain, it may help explain why previous reports on the effects of pairwise NCs have been so heterogeneous and why single-neuron responses to repeated presentations of the same stimuli seem noisy.

**EXPERIMENTAL PROCEDURES**

**Experimental Protocols and Data Preprocessing**

Detailed information on experimental protocols and data preprocessing is available in the Supplemental Experimental Procedures. In short, we used 14 C57BL/6 wild-type mice in this study; nine mice were recorded over several weeks, five within a single day. All experimental procedures were conducted with approval of the animal ethics committee of the University of Amsterdam. Two-photon calcium imaging was performed with a 512 × 512 pixel frame size at a sampling frequency of 12.7 Hz on a modified Leica SP5 confocal system with a wavelength of 880–910 nm to excite GCaMP6 molecules after virally induced expression in L2/3 of mouse V1. Visual stimuli were either bidirectionally moving square-wave drifting gratings (60 retinal degrees, 0.05 cycles/degree, 1-Hz temporal frequency, 3-s duration with direction reversal after 1.5 s, 5-s inter-trial interval [ITI]), or natural movies presented at 25 Hz that consisted of four scenes taken from the BBC’s Earth-flight (Winged Planet)–Condor Flight School. Each presentation lasted 20 s and was repeated 31 times per recording session; no interval was present between repetitions. We discarded the first of these 31 repetitions to avoid onset effects. All stimuli were presented on a 15-inch thin-film transistor (TFT) screen with a refresh rate of 60 Hz positioned 18 cm from the mouse’s eye, which was controlled by MATLAB using the PsychToolbox extension (Brainard, 1997; Pelli, 1997). Data were x-y registered and checked for movement artifacts along the z axis, and region of interests were semi-automatically selected using custom MATLAB software. Mice were awake during all recordings, and the visually stimulated eye was monitored with an infrared camera. When necessary, t test p values were corrected for multiple comparisons using a false discovery rate (FDR) procedure (Benjamini and Hochberg, 1995). Pairwise NCs and SCs were calculated as described previously (Montijn et al., 2014).
Multidimensional Population Code Stability

In addition to pairwise interactions, we tested the stability of multidimensional neuronal representations of oriented gratings across days (Figure 3). We calculated, for the same stimulus orientation, the Euclidean distance in multidimensional space between population responses on different recording days as follows. For each trial t, the population response can be represented as a multidimensional vector \( r_t = [r_1, \ldots, r_N] \), where \( r_i \) is neuronal activity in dF/F0 (mean over frames during stimulus presentation per neuron) and \( N \) is the number of neurons. The average population response \( \mu_t \) for orientation \( q \) is the mean over all repetitions for that orientation. For each orientation, we normalized the Euclidean distance \( d_{t,1:2} \) between \( \mu_t \) for two recording days \( t = 1 \) and \( t = 2 \) by the mean distance across repetitions within those recordings; \( \bar{d}_{t,1:2} \) and \( \bar{d}_{t,2:2} \):

\[
\bar{d}_{t,1:2} = \frac{d_{t,1:2}}{(\bar{d}_{t,1} + \bar{d}_{t,2})/2}
\]

(1)

A value of 1.0 indicates that the distance in population response between days is the same as the variability in responses within 1 day. We averaged the distance across all orientations to get one value per recording pair. To be able to fit an exponential decay function to this distance, we transformed it into similarity metric \( s_{MD} \):

\[
s_{MD,1:2} = 1 - d_{t,1:2}.
\]

(2)

For a similarity of 1.0 the population response distance is 0.0 (i.e., the population responses are identical), and for a similarity of 0.0 the variability is equal within and across days. To study different dimensionalities, we subsampled the total population 100 times and averaged the mean Euclidian distances obtained from the neuronal groups of the same size (e.g., for dimensionality 3, we took the mean Euclidian distance across 100 triplets of neurons).

Mahalanobis Analysis of Group-Wise Multidimensional Activity Decoder

To investigate neuronal response interdependencies within groups larger than pairs, we constructed a multidimensional decoder that classifies trials by minimizing multidimensional Mahalanobis distances to stimulus classes, in which each neuron’s activity represents a dimension. Because the Mahalanobis distance normalizes multidimensional variability and covariability, the optimal interclass separation boundaries are linear hyperplanes in Mahalanobis space (but not necessarily in normal response space; cf. Figures 4A and 4E; De Maeschalck et al., 2000). We calculated the Mahalanobis distance \( D_0(t) \) between the population activity \( r_t \) of each trial \( t \) (as earlier) and the mean population activity \( \mu_t \) for stimulus class \( q \), in which each dimension represents a single neuron in the population:

\[
D_q(t) = \sqrt{\left( r_t - \mu_q \right)^T S_q^{-1} \left( r_t - \mu_q \right)}.
\]

(3)

Here, \( ^T \) indicates the vector transpose and \( S_q^{-1} \) is the inverse of the covariance matrix over all neurons for stimulus \( q \). We repeated this procedure for all stimulus classes (e.g., eight gratings orientations), yielding Mahalanobis distances for all combinations of trials and stimulus types. The decoded stimulus type for a trial \( t \) is the class \( q \) with the lowest Mahalanobis distance \( D(t) \). We used a leave-one-repetition-out (cross-validation) procedure, in which the class means and covariance matrix were calculated with exclusion of the to-be-decoded repetition block consisting of all stimulus types (i.e., grating orientations or natural movie scenes or frames).

When the neuronal group size (number of dimensions) grows, the number of possible combinations grows beyond exponentially. We therefore performed a decoding procedure for all neuronal group sizes while integrating information from different numbers of groups by taking the sum of Mahalanobis distances to all stimulus classes over neuronal groups (Figure 4B). We observed a saturation of performance of ~60–80 groups of neurons, indicating that adding more random groups of neurons will no longer increase the amount of information that can be extracted. For further analyses, we therefore used the decoder’s performance when integrating up to 100 random groups of neurons and assumed this performance reflects the maximal amount of information that can be extracted (e.g., Figures 4C and 4D).

Multidimensional Response Variability and Noise Prediction in Natural Movies

To assess multidimensional neuronal variability, we calculated the distance in neural space orthogonal and parallel to decision boundaries between adjacent orientations (\( \theta_1, \theta_2 \); Equations S9 and S10). The variability is expressed as SD (units of dF/F0) across repetitions of the same orientation. To control for biases across animals and across dimensionalities, we normalized the raw variability by the variability obtained after shuffling.

We performed a leave-one-repetition-out cross-validated prediction of across-repetition neuronal response noise of the same movie frame to ascertain the noisiness of neuronal responses. To do so, we fitted a multivariate Gaussian of varying dimensionality (i.e., number of neurons) and used as readout the most likely response of neuron \( N \), given the neuronal response vector of neurons \( \{1 \ldots (N-1)\} \) during that repetition for the fitted multivariate Gaussian (also see the Supplemental Experimental Procedures).

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures and seven figures and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2016.07.065.

AUTHOR CONTRIBUTIONS

J.S.M. and G.T.M. performed the experiments and analyzed the data. G.T.M. developed the viral injection protocols for GCaMP6 expression. J.S.M. and C.M.A.P. designed the experiments and analyses. J.S.M. wrote the paper. J.S.M., G.T.M., C.S.L., and C.M.A.P. discussed the results and contributed to the final manuscript.

ACKNOWLEDGMENTS

We thank Sophie Denève for helpful comments and scientific discussions, Pietro Girotti and Quentin Perenoud for helpful suggestions on how to improve the manuscript, and Gwylan Scheeren for assistance with developing the eye-tracking procedure. For the use of GCaMP6, we thank Vivek Jayaraman, Rex Kerr, Douglas Kim, Loren Looger, and Karel Svoboda from the GENIE Project, Janelia Research Campus, Howard Hughes Medical Institute. This work was supported by a Netherlands Organization for Scientific Research-Excellence grant for the Brain & Cognition Project (433-09-208) and the Human Brain Project (HBP SGA1) to C.M.A.P. and by a Netherlands Organization for Scientific Research Veni grant (863.11.010) to C.S.L.

Received: November 26, 2015
Revised: April 21, 2016
Accepted: July 25, 2016
Published: August 18, 2016

REFERENCES


2496 Cell Reports 16, 2486–2498, August 30, 2016


