Size tuning of neural response variability in laminar circuits of macaque primary visual cortex

Lauri Nurminen[1,2], Maryam Bijanzadeh[1] and Alessandra Angelucci[1,*]

[1] Department of Ophthalmology and Visual Science, Moran Eye Institute, University of Utah, 65 Mario Capecchi Drive, Salt Lake City, UT 84132, USA.

[2] Present address: College of Optometry, University of Houston, 4401 Martin Luther King Boulevard, Houston, TX 77204-2020, USA

[*]Corresponding author's address:

65 Mario Capecchi Drive
Salt Lake City, UT 84132, USA
Tel: (801) 585 7489
Email: alessandra.angelucci@hsc.utah.edu

ABSTRACT

A defining feature of the cortex is its laminar organization, which is likely critical for cortical information processing. For example, visual stimuli of different size evoke distinct patterns of laminar activity. Visual information processing is also influenced by the response variability of individual neurons and the degree to which this variability is correlated among neurons. To elucidate laminar processing, we studied how neural response variability across the layers of macaque primary visual cortex is modulated by visual stimulus size. Our laminar recordings revealed that single neuron response variability and the shared variability among neurons are tuned for stimulus size, and this size-tuning is layer-dependent. In all layers, stimulation of the receptive field (RF) reduced single neuron variability, and the shared variability among neurons, relative to their pre-stimulus values. As the stimulus was enlarged beyond the RF, both single neuron and shared variability increased in supragranular layers, but either did not change or decreased in other layers. Surprisingly, we also found that small visual stimuli could increase variability relative to baseline values. Our results suggest multiple circuits and mechanisms as the source of variability in different layers and call for the development of new models of neural response variability.

INTRODUCTION

The cortex consists of six layers, each having distinct input/output relations and forming distinct intra-laminar circuits (Levitt and Lund, 2002a; Douglas and Martin, 2004; Nassi and Callaway, 2009). The distinct connectivity patterns of cortical layers suggests that they may serve different functions in cortical information processing and sensory perception (Constantinople and Bruno, 2013; Nandy et al., 2013; Bijanzadeh et al., 2018; Takahashi et al., 2020).

By measuring trial-averaged neural responses to visual stimuli, previous studies have shown that the response properties of neurons differ across cortical layers. For example, in the macaque primary visual cortex (V1) granular (G) layer neurons show broader orientation tuning (Ringach et al., 1997) and faster responses (Bijanzadeh et al., 2018) than neurons in supra (SG) or infragranular (IG) layers. Moreover, compared to other layers, the response of SG layer neurons to stimulation of their receptive field (RF) is more strongly suppressed by stimulation of the RF surround (Shushruth et al., 2009; Henry et al., 2013), and small versus large stimuli evoke distinct patterns of laminar activity (Bijanzadeh et al., 2018).

While studies based on measurements of trial-averaged neural responses have been foundational to our understanding of cortical layers, they have provided a limited view of cortical processing. Deviations from trial-averaged cortical responses (Tolhurst et al., 1983) have traditionally been interpreted as noise that impairs the fidelity of neural representations (Shadlen and Newsome, 1998; Moreno-Bote et al., 2014). However, reports that neural response variability is modulated by visual stimuli (e.g. Churchland et al., 2010) have led some to suggest that variability may, instead, play a role in sensory information processing (Festa et al., 2021). For example, computational studies have assigned a role for variability in perceptual inference (Orban et al., 2016; Henaff et al., 2020), and suggested that cortical layers may play distinct roles in perceptual inference (Bastos et al., 2012).

To understand how laminar processing is differentially affected by neural response variability, here we have used laminar recordings to investigate how variability is modulated by stimulus size across the layers of macaque V1. We have focused on size tuning of response variability, as this enables us to isolate the driving feedforward thalamic inputs to the RF, from the modulatory inputs arising from the RF-surround, thought to be mediated by intrinsic V1 and corticocortical circuits (Angelucci et al., 2002, 2017; Nurminen et al., 2018).

We found that in all layers, a stimulus matched to the RF size of the recorded neurons reduced cortical response variability compared to pre-stimulus baseline. However, modulation of variability by stimulation of the RF-surround was layer dependent. In SG layers, stimulation of the surround increased both single and shared neural response variability, relative to the variability measured for stimuli the size of the RF. In contrast, in G and IG layers, stimulation of the surround either had little effect or reduced response variability relative to its value measured during presentation of a stimulus matched to the RF size. Interestingly, we found that in a subset of neurons, small stimuli could increase variability compared to pre-stimulus baseline. Our results point to multiple sources of variability affecting cortical processing in a laminar-specific way, and call for new models of neural response variability.

RESULTS

We recorded visually evoked local field potential (LFP) and multi-unit spiking activity (MUA), using 24-channel linear electrode arrays (100µm electrode spacing) inserted perpendicularly to the surface of area V1 in two sufentanil-anesthetized macaque monkeys (see Methods). For accurate assignment of recorded responses to cortical layers, verticality of the electrode array was verified by the spatial overlap and similarity of orientation preference of the neurons' minimum response fields across the array, and confirmed by postmortem histology (as described in Bijanzadeh et al., 2018). Laminar boundaries were identified by current source density (CSD) analysis of LFP signals (Mitzdorf, 1985) averaged over all stimulus diameters used in this study. This allowed us to locate the granular (G) layer 4C as the site of the earliest current sink followed by a reversal to current source, the site of the reversal marking the bottom of the G layer; layers above and below G were defined as supragranular (SG) and infragranular (IG), respectively. We used spiking activity in response to the same stimulus, to identify the top and bottom of the cortex (Bijanzadeh et al., 2018).

To understand how stimulus size modulates cortical response variability across V1 layers, we measured Fano-factor and the shared variability among simultaneously recorded neurons as a function of grating diameter for 82 visually responsive multi-units. At the beginning of each penetration, we mapped the minimum-response fields of the recorded units (see Bijanzadeh et al., 2018 for details). We next presented grating stimuli, centered on the aggregate minimum response fields of the recorded units, to characterize the orientation, spatial frequency, and temporal frequency tuning of the recorded MUA. We then selected the stimulus parameters that maximized the response of as many simultaneously recorded units as possible. Using these optimized parameters, we ran size tuning experiments in which the diameter of a drifting grating stimulus was varied from 0.1° (0.2° in one penetration) to 26°.

Layer dependent modulation of neural response variability by stimulus size

Figure 1 shows size-tuning data for representative multi-units recorded in SG, G, and IG layers. For all three example units, Fano-factor decreased as the stimulus diameter was increased to fill the RF of the recorded neurons. However, for the SG layer unit (**Figure 1A**), increasing the stimulus diameter beyond the RF boundaries increased Fano-factor, relative to Fano-factor measured when the stimulus was matched to the size of the RF. In contrast, increasing the stimulus diameter beyond the RF boundaries did not affect Fano-factor for the G and IG layer units (**Fig. 1B-D**).

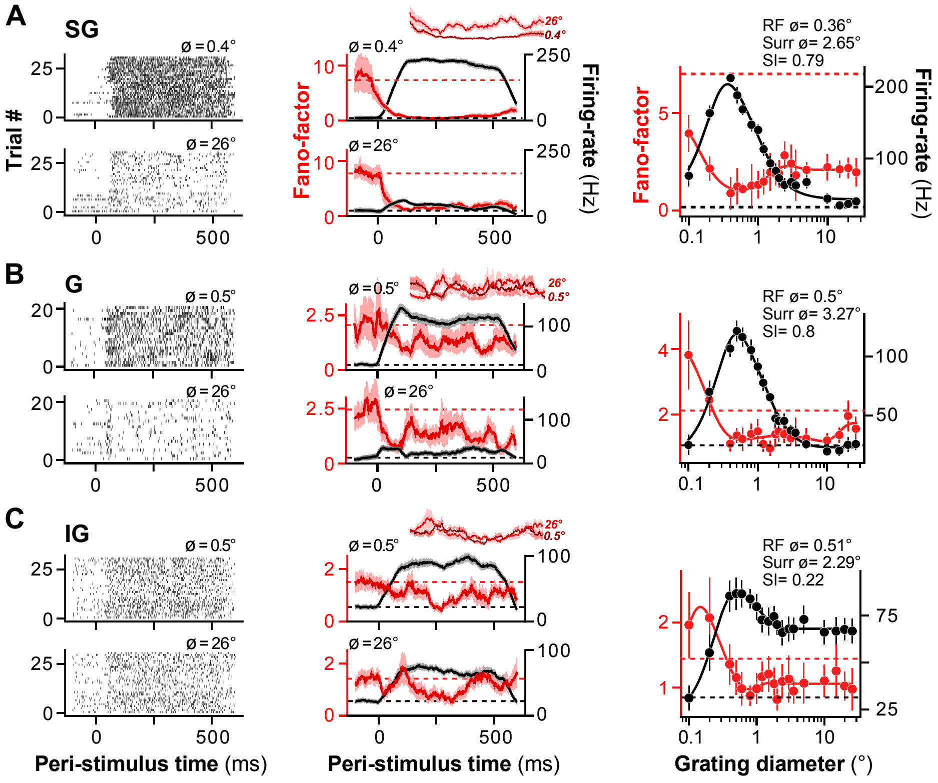

Figure 1. Size tuning of Fano-factor and mean firing rate in macaque V1: representative units. A) Representative supragranular (SG) layer unit. **Left:** MUA spike-rasters measured at two stimulus diameters, either a diameter equal to the RF diameter of the recorded multi-unit (Top), or a diameter of 26° (Bottom). **Middle**: Peri-stimulus time histograms (PSTHs) of Fano-factor (*red*) and mean firing-rate (*black*) computed in a 100 ms rectangular sliding window for the same two stimulus diameters. The shaded area represents the standard deviation (s.d.) of the bootstrapped Fano-factor distribution (for the Fano-factor curve) or the standard-error-of-the-mean (s.e.m., for the firing rate curve). *Inset*: Zoomed-in Fano-factor curves for the smaller (darker red) and larger (lighter red) stimulus diameters between 50 and 350 ms after stimulus onset. **Right**: Fano-factor (*red*) and firing-rate (*black*) averaged over 50-350 ms after stimulus onset and plotted against the stimulus diameter. *Solid lines*: fits to the data. *Dashed lines*: baseline Fano-factor (*red)* and firing rate (*black*), measured prior to stimulus onset. Error bars are: s.d. of the bootstrapped Fano-factor distribution (*red*) or s.e.m. (*black*). **B)** Representative granular (G) layer unit. **C)** Representative infragranular (IG) layer unit. Conventions in (B-C) are as in (A).

Similar results were observed for the population of recorded MUA. **Figure 2A** shows Fano-factor and mean firing-rate averaged over the population of multi-units in our sample. In SG (n=31), G (n=15) and IG (n=36) layers, increasing the stimulus diameter from 0.1 to a size equal to the aggregate RF diameter of the recorded cells progressively decreased Fano-factor and increased firing-rate. Fano-factor reached a minimum at the stimulus diameter matching, or slightly larger than that of the RF, the latter defined as the peak of the firing rate size-tuning curve.

Large gratings extending into the RF surround of V1 cells are known to suppress the mean spiking response evoked by a stimulus confined to the cells' RF, a phenomenon known as surround suppression (Sceniak et al., 2001; Angelucci et al., 2002; Cavanaugh et al., 2002; Levitt and Lund, 2002b; Shushruth et al., 2009; Angelucci and Shushruth, 2013). Consistent with these previous reports, as the stimulus size was increased beyond that of the RF diameter, firing rate decreased across all layers, and this suppression was strongest in the SG layers. However, increasing the stimulus diameter beyond the RF of the recorded units had different effects on Fano-factor in different layers. In SG layers, as the stimulus diameter was increased beyond that of the RF, Fano-factor significantly increased relative to its value when the stimulus matched the RF diameter (**Fig. 2A Left**; t-test stim. diam equals RF vs. stim diam equals 26°, p=0.002). In G and IG layers, instead, increasing the stimulus diameter beyond the aggregate RF did not significantly affect Fano-factor (**Fig. 2A Middle and Right**; G: t-test, p=0.80; IG: t-test, p=0.35).

Figure 2B shows Fano-factor estimated at 4 different stimulus diameters individually for each multi-unit (values were extracted from functions fit to the data; see Methods), and then averaged over the units. Consistent with the population size-tuning curves, this analysis also showed a laminar dependence of the impact of surround stimulation on Fano-factor. In SG layers, Fano-factor was significantly higher for the 26° diameter stimulus than for the stimulus matching

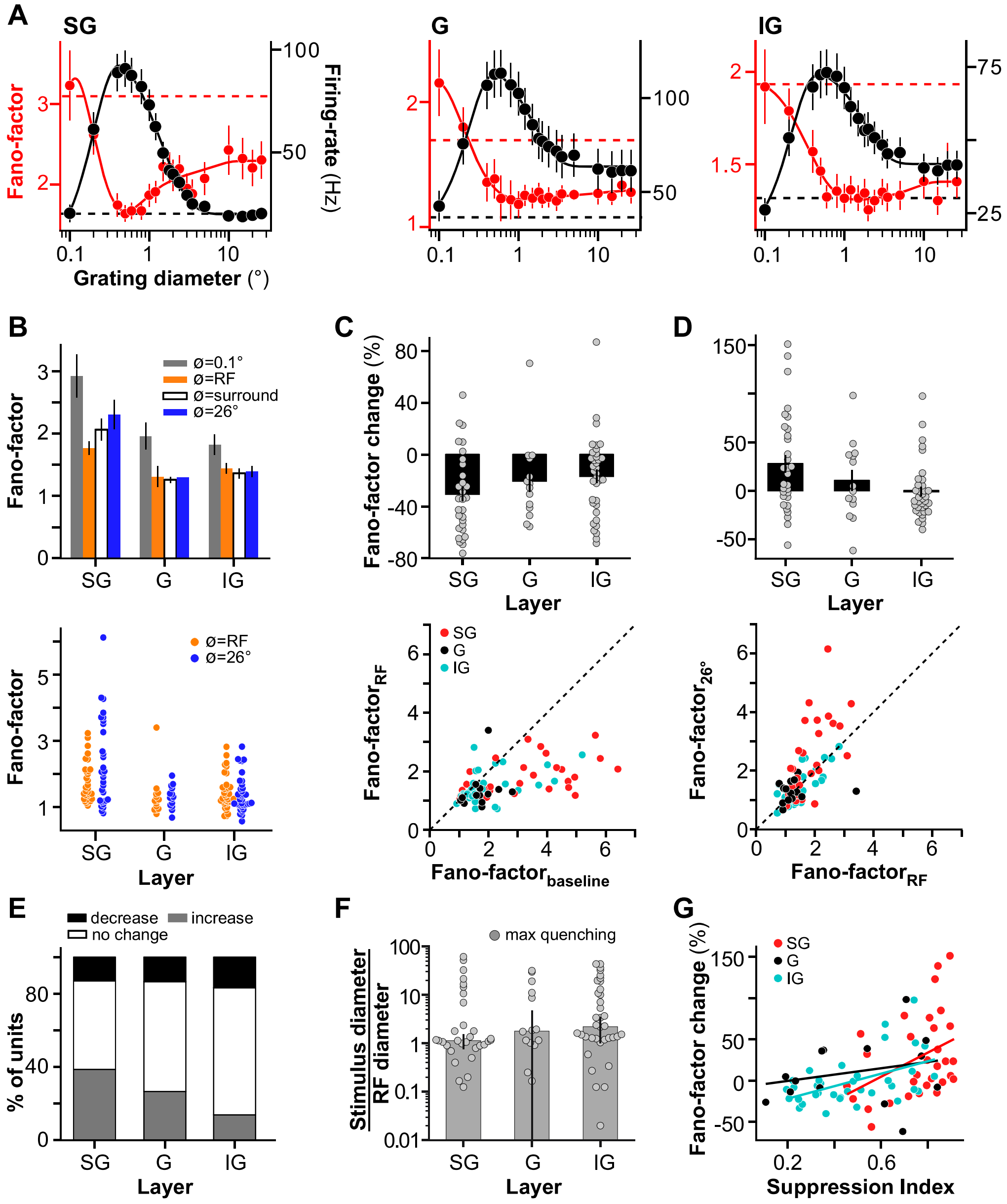

Figure 2. Size tuning of Fano-factor and mean firing rate: population data. A) Average Fano-factor (*red*) and mean firing-rate (*black*) as a function of stimulus diameter for the population of SG (**Left**; n=31), G (**Middle**; n=15) and IG (**Right**; n=36) layer units. *Dashed lines*: average baseline Fano-factor (*red)* and firing rate (*black*); error bars: s.e.m. **B) Top:** Fano-factor values averaged over 82 units at four different stimulus diameters (0.1°, a diameter equal to the RF diameter, a diameter equal to the RF-surround diameter (see Methods for definition), and 26°). Error bars: s.e.m. **Bottom:** Fano-factor values for individual multi-units in SG, G and IG layers at two different stimulus diameters (as indicated). **C) Top:** Mean percent change in Fano-factor relative to baseline induced by a stimulus matched in size to the RF diameter, for the different layers. *Dots:* Individual data points. Error bars: s.e.m**.** **Bottom:** Scatter plot of Fano-factor during pre-stimulus baseline vs during presentation of a stimulus matched to the RF diameter. Different colored dots indicate units in different layers. **D) Top:** Mean percent change in Fano-factor induced by a 26° diameter stimulus relative to the Fano-factor value evoked by a stimulus matched to the RF diameter. **Bottom:** Scatter plot of Fano-factor for presentation of stimuli of two different sizes (a diameter equal to that of the RF vs. a diameter of 26°). Other conventions as in panel C). **E)** Percent of multi-units in each layer for which stimulation of the RF significantly decreased variability (*black*), did not affect variability (*white*) or increased variability (*gray*). **F)** Median stimulus diameter at the largest decrease in Fano-factor (or max quenching), normalized to the RF diameter of the recorded units, for different layers. Error bars: s.d. of the bootstrapped distributions. *Dots:* individual cell data. **G)** Scatter plot of percent change in Fano-factor evoked by the largest surround stimulus (26° diameter) relative to the Fano-factor evoked by a stimulus matched to the RF diameter vs. suppression index (see Methods). Color dots identify units in different layers, as indicated. Lines are regression lines fitted to the individual layer data.

the RF diameter (mean ± s.e.m.: 2.30±0.23 vs. 1.77±0.11, t-test, p=0.005). However, in G and IG layers, Fano-factor did not differ significantly for these two size stimuli (G: 1.30±0.09 vs. 1.31±0.17, t-test, p=0.96; IG: 1.40±0.09 vs. 1.44±0.09, t-test, p=0.42).

The top panel in **Figure 2C** plots for each layer the percent change in Fano-factor, relative to baseline, evoked by a stimulus equal in size to the RF diameter of the recorded units. The bottom panel shows a scatter plot of Fano-factor at baseline vs during presentation of a stimulus matched to the RF size. In all layers, presentation of visual stimuli in the RF reduced Fano-factor relative to baseline; there was no statistically significant difference across layers in the percent change in Fano-factor (one-way ANOVA, main effect of layer on percent change in Fano-factor, p=0.20, n=81; mean ± s.e.m. SG: -30.50±5.63%, n=31; G: -20.38±8.54%, n=14; IG: -16.77±5.16%, n=36). The top panel in **Figure 2D** plots for each layer the percent change in Fano-factor evoked by a 26° stimulus relative to a stimulus matched to the RF diameter, and the bottom panel shows a scatter plot of Fano-factor values at these two stimulus diameters. As for the previous analysis, the impact of surround stimulation on Fano-factor was layer dependent (one-way ANOVA, main effect of layer on percent change in Fano-factor, p=0.02, n=81). In G and IG layers, there was no statistically significant change in Fano-factor (mean ± s.e.m. G: 10±10.8%, n=14, t-test μ≠0, p=0.35; IG: -0.98±4.91%, n=36, t-test μ=0, p=0.84). In contrast, in SG layers there was a strong percent increase in Fano-factor as the stimulus involved the RF surround (SG: 27.82±9.15%, n=31, t-test μ=0, p=0.004). One unit with extreme change in Fano-factor (564%), caused by dividing with a number close to zero, was removed from the analyses presented in **Figure 2C,D,G**.

A unit-by-unit analysis revealed that stimulation of the RF surround affected the variability of V1 neurons in three distinct ways. For the majority of the units (59.7%, n=49), surround stimulation did not significantly affect variability, as determined by bootstrapping (see Methods). Compared to Fano-factor measured when the stimulus was confined to the RF, a stimulus in the RF surround statistically significantly increased Fano-factor in 25.6% of the units (n=21) and decreased it in 14.6% of the units (n=12). These three distinct effects were found in all layers, but in different proportions (**Figure 2E)**. In SG layers (n=31), surround stimulation increased variability in 38% of the units, decreased it in 13% of the units and did not have a statistically significant impact on variability in 49% of the units. In G layers (n=15), surround stimulation increased variability in 26% of the units, decreased it in 13% of the units and did not have a statistically significant impact on variability in 61% of the units. In IG layers (n=36), surround stimulation increased variability in 14% of the units, decreased it in 17% of the units and did not have a statistically significant impact on variability in 69% of the units. Consistent with these results, the stimulus diameter at the lowest Fano-factor value (or max quenching) was equal or close the RF diameter in SG layers (median±s.d. of the bootstrapped median distribution 1.14±0.35), but larger than the RF diameter in G (1.80±2.98) and IG layers (2.22±1.20; **Fig. 2F**).

We found that surround suppression was stronger in units in which Fano-factor was increased by surround stimulation (one-way ANOVA, strength of surround suppression conditioned on whether RF surround increased, decreased or had no effect on Fano-factor relative to RF stimulation, p=0.001, n=82; **Fig. 2G**). For the units in which surround stimulation increased Fano-factor relative to RF stimulation (n=21), the strength of surround suppression was 74.1±2.91%, while it averaged 54.7±3.47% for the units in which surround stimulation did not affect variability (n=49), and 51.8±5.63% for the units in which surround stimulation reduced variability (n=12). Moreover, in SG and IG, but not in G, layers, there was a statistically significant correlation between the strength of surround suppression and the percent change in Fano-factor caused by surround stimulation (SG: r=0.39, p=0.028; G: r=0.24, p=0.40; IG: r=0.50, p=0.002; Pearson correlation; **Fig. 2G**). In all layers, statistically significant increases in variability (as determined by bootstrapping; see Methods) induced by surround stimulation had larger magnitude than decreases in variability (increase vs. decrease SG: 103±18.9% vs -36.5%±6.10%; G: 68.1±14.4% vs -40.1±20%; IG: 91.6±31.0% vs -36.5±6.09% independent samples t-test pooled over layers and computed over the absolute value of the Fano-factor change induced by RF-surround, p = 0.002).

To rule out that the changes in Fano-factor with stimulus size are trivially related to changes in firing rate, we performed a "mean-matched" analysis (Mitchell et al., 2009; Churchland et al., 2010) (see Supplementary Methods, Supplementary Results, and **Supplementary Fig. 1**). This analysis showed that changes in firing rate were not the cause of stimulus-size dependent changes in Fano-factor.

Amplification of cortical response variability by small stimuli

It has been previously reported that the onset of a visual stimulus reduces cortical response variability relative to pre-stimulus baseline (Churchland et al., 2010). However, previous studies used relatively large stimuli, and the impact of stimulus size on response variability has not been explored. Previous experimental studies (Ichida et al., 2007) have shown, and several models of cortical dynamics predicted, that when the cortex is weakly driven (for example by a small stimulus), the cortical state is dominated by excitation, whereas it is dominated by inhibition when the cortex is strongly driven, e.g. by a large stimulus (Schwabe et al., 2006, 2010; Rubin et al., 2015; Hennequin et al., 2018). In an excitation-dominated cortical state, stochastic supralinear stabilized networks predict amplification of response variability relative to pre-stimulus baseline (Hennequin et al., 2018). To test this model's prediction, we examined the impact of small stimuli on response variability.

Figure 3A shows the response of one example IG layer multi-unit to gratings of 0.1° or 1° in diameter, respectively, centered on its RF. Both of these stimuli evoked firing-rates higher than the pre-stimulus baseline firing-rate (**Fig. 3A,** left and middle). In contrast, changes in Fano-factor after stimulus onset depended on stimulus size: Fano-factor decreased after presentation of a 1° stimulus, but increased after presentation of a 0.1° stimulus (**Fig. 3A**, right).

Amplification of variability for small stimuli was seen also at the population level. **Figure 3B** compares Fano-factor evoked by a 0.1° diameter grating with that evoked by a grating of diameter equal to the RF diameter of the recorded multi-units, normalized to the pre-stimulus baseline, and averaged over the population of SG (n=31, left), G (n=15, middle) and IG (n=36, right) units. Presentation of the small stimulus significantly increased Fano-factor relative to pre-stimulus baseline in G and IG ($p<0.05$, one-sample t-test, n=15 and 36, respectively), but not SG ($p=0.14$, n=31), layers. Consistent with previous studies (Churchland et al., 2010), in all layers, the larger stimulus decreased Fano-factor relative to baseline. There was no obvious difference in firing-rates across the layers that could have explained the increase in Fano-factor in G and IG layers for smaller stimuli, but not in SG layers (**Fig. 3B**).

To provide a better understanding of variability amplification across layers, we performed a unit-by-unit analysis. This revealed significant variability amplification for small stimuli in all layers. We included in this analysis only units showing statistically significant increases or decreases (see Methods) in Fano-factor relative to baseline for at least one data-point. While all units in our sample showed statistically significant stimulus-evoked decreases in Fano-factor, 67%

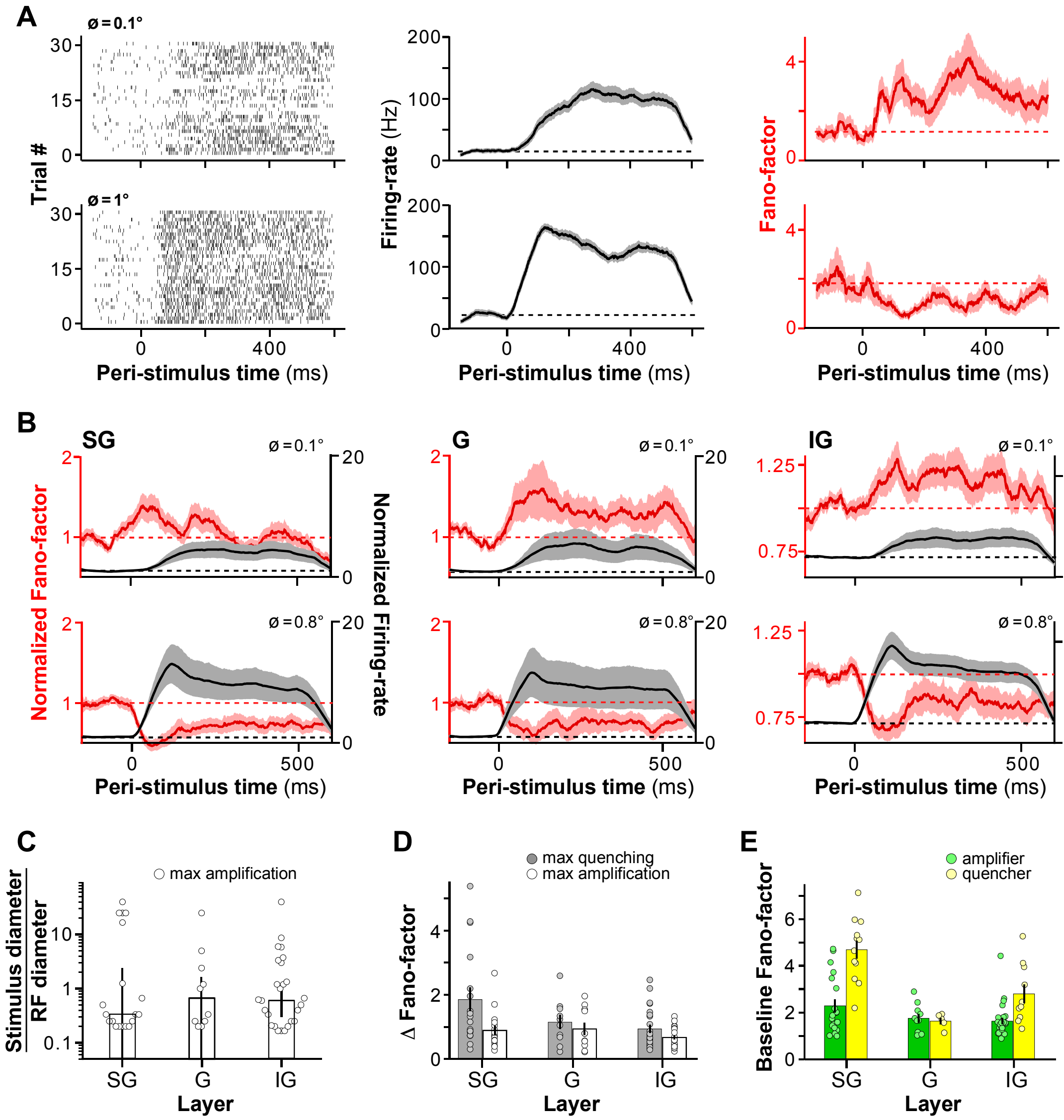

Figure 3. Amplification of cortical response variability by small visual stimuli. A) An example unit showing stimulus-evoked increases in firing rate and Fano-factor for a small (0.1°) grating diameter (**Top**), but a decrease in Fano-factor for a larger stimulus matching the RF diameter (1°; **Bottom**). **Left:** Spike rasters. **Middle:** PSTHs of firing-rate computed in a 100ms sliding window. *Shaded gray area* here and in B) indicates the s.e.m. computed over trials. **Right:** Fano-factor computer over 100ms sliding window. *Shaded red area* here and in B) is the s.d. of the Fano-factor distribution bootstrapped over trials. **B)** Population-averaged time course of Fano-factor (*red*) and firing-rate (*black*) in SG **(Left),** G **(Middle),** and IG **(Right)** layers computed at two stimulus diameters (**Top:** 0.1°, **Bottom:** 0.8°). Both the Fano-factor and firing rate were normalized to the pre-stimulus baseline of each unit before averaging. **C)** Median stimulus diameter evoking the largest magnitude increase in Fano-factor, normalized to the RF diameter of the recorded units, for different layers. Error bars: s.d. of the bootstrapped distributions. *Dots* here and in (D-E)*:* individual cell data. **D)** Median difference in Fano-factor (Fano-factor at the stimulus diameter causing the largest change in Fano-factor minus the baseline Fano-factor) ± s.d. of the bootstrapped distributions at max quenching (*gray*) and max amplification (*white*) for different layers. **E)** Mean baseline Fano-factor for amplifier (*green*) and quencher (*yellow*) units. Error bars: s.e.m.

of these units (55/82) also showed statistically significant increases in Fano-factor for presentation of small stimuli. The proportion of units showing both increases and decreases in stimulus-evoked Fano-factors was fairly constant across layers (**SG,** 61%; **G,** 67%; **IG,** 72%).

On average, the largest stimulus-evoked increase in Fano-factor for the cells that showed variability amplification relative to baseline was observed when the stimulus diameter was smaller than the RF of the recorded units (**Fig. 3C**; median stimulus diameter normalized to the RF diameter at the largest increase in Fano-factor relative to baseline, ±s.d. of the bootstrapped distribution: **SG,** 0.33±1.20; **G,** 0.67±0.52; **IG,** 0.63±0.26). We found that stimulus-evoked increases in Fano-factor were smaller in magnitude than stimulus-evoked decreases in Fano-factor. The difference between the magnitude of maximum variability quenching and magnitude of maximum variability amplification was statistically significant in SG and IG layers (**Fig. 3D**; mean±s.e.m. quenching vs. amplification: **SG,** 1.85±0.37 vs. 0.89±0.14, t-test p=0.01; **IG,** 0.93±0.12 vs. 0.67±0.06, p=0.03), but was not statistically significant in G layers (**G**, 1.14±0.21vs. 0.93±0.19, p=0.23). For this analysis, we removed outlier data points that were at least 2.5 absolute median deviations above or below the median.

Across the entire population, the units showing variability amplification for small stimuli (here termed "amplifier") had a significantly lower baseline Fano-factor than units in which a stimulus always reduced variability (termed "quencher") (mean baseline Fano-factor ± s.e.m.: 3.44±0.33 for quencher units vs. 1.90±0.13 for amplifier units, t-test p=0.000001). However, this varied by layer (**Fig. 3E**); in SG and IG layers, baseline Fano-factor was significantly higher in the quencher units than in the amplifier units (mean baseline Fano-factor±s.e.m.: **SG,** quencher 4.70±0.39, n=12, vs. 2.30±0.29, n=19, t-test p=0.00023; **IG,** quencher 2.82±0.40, n=10, vs. amplifier 1.65±0.14, n=26, t-test p=0.0013). Instead, in the G layer, baseline Fano-factor did not differ significantly between these two groups (quencher 1.65±0.14, n=5, vs. amplifier 1.77±0.19, n=10, t-test p=0.68). Moreover, baseline firing rate was significantly lower in the amplifier units compared to the quencher units (mean±s.e.m. baseline firing-rate: 4.1±0.4Hz vs. 6.0±0.7Hz, t-test $p=2^{-12}$). Importantly, however, all amplifier units also showed variability quenching at larger stimuli. This suggests that a floor effect, due to low baseline firing-rates, cannot explain the variability amplification in our data (see Discussion).

Layer dependent size-tuning of network variability

The results presented above indicate that the response variability of individual cortical neurons is modulated by stimulus size. However, the impact of neural response variability on visual processing also depends on how strongly the variability is shared across neurons (Shadlen and Newsome, 1998; Bair et al., 2001). To determine the impact of stimulus size on shared variability, we exploited the covariance of simultaneous recordings obtained with electrode arrays.

The raster plots in **Figure 4A-C** show the spiking activity of simultaneously recorded neurons in a single example penetration spanning all layers (SG, n= 4 units, **Fig. 4A**; G, n=3, **Fig.**

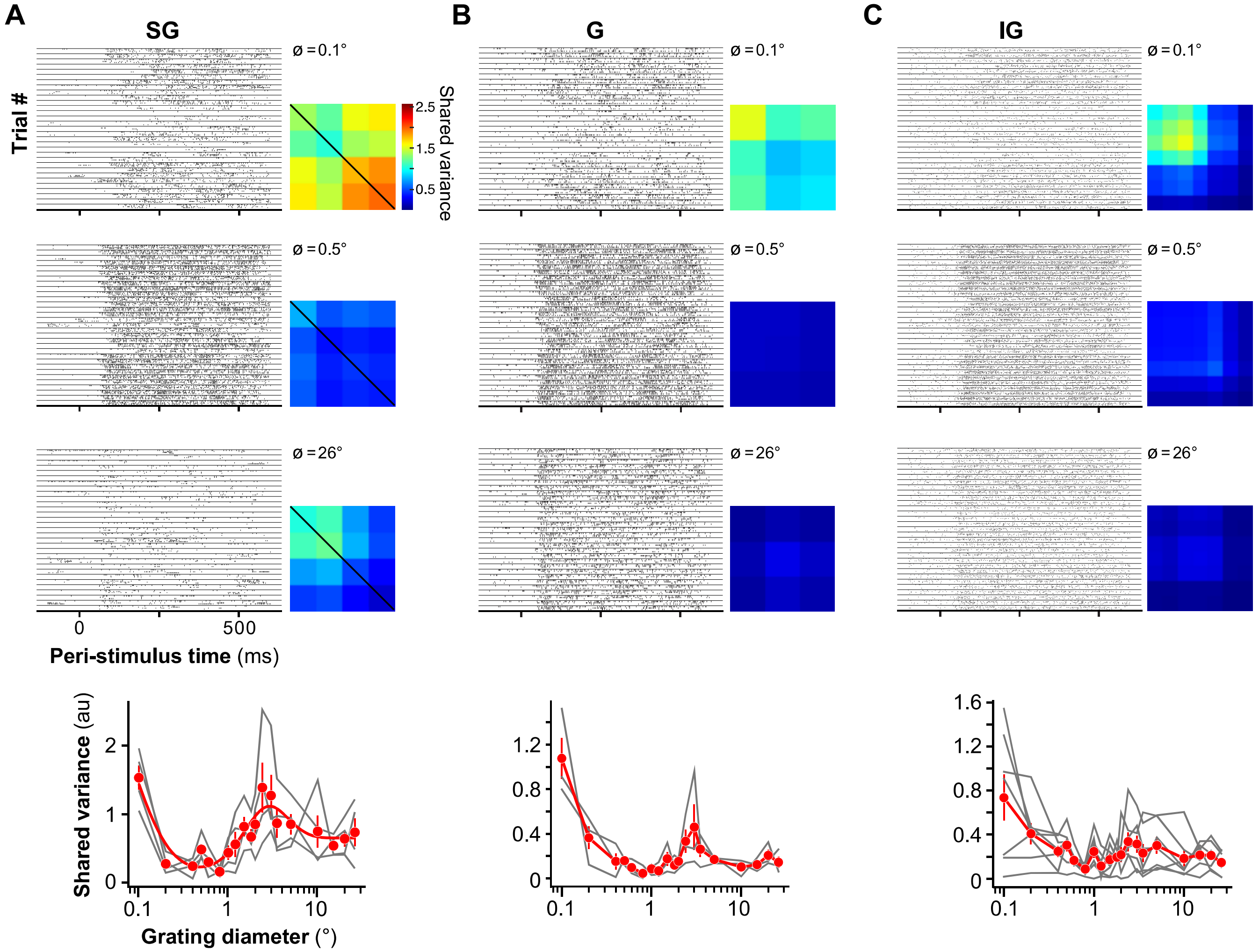

Figure 4. Size tuning of shared variance across V1 layers: an example penetration. A) Left: raster plots showing the spike times of four simultaneously recorded SG neurons, across several trials, in response to 0.1° (Top), 0.5° (Middle), and 26° (Bottom) diameter gratings. The responses of all 4 neurons in a single trial are shown between two consecutive horizontal lines. Horizontal lines separate different trials. **Right:** Network covariance matrices estimated with a single-factor factor analysis for each of the same 3 different stimulus diameters. The diagonal of the network covariance matrix holds the shared variance for each recorded unit. **Bottom:** Shared variance as a function of stimulus diameter. The red markers show mean±s.e.m. of the shared variance computed over the SG neuron population recorded in this example penetration (n=4). The gray curves show the data for the individual 4 units. **B-C)** same as in A), but for G (n=3) and IG (n=7) layer units.

was silent. However, the population responses appeared less coordinated following presentation of the 0.5° diameter grating. The covariation in population responses to presentation of the 26° diameter grating, instead, appeared to be layer dependent. In all layers, responses to a 26° stimulus were reduced compared to those to a 0.5° diameter grating. However, compared to the responses to the 0.5° stimulus, neural activity in response to the 26° stimulus appeared more strongly coordinated in SG layers, but remained relatively uncoordinated in G and IG layers.

To quantify these observations, we used factor analysis (see Methods). This allowed us to isolate the network variability (i.e. the variability that is shared across neurons) from the single neuron spiking variability (i.e. the variability that is private to each recorded unit) (Churchland et al., 2010). Factor analysis decomposes the measured covariance matrix into a low-rank network covariance matrix and a diagonal matrix that holds the private variances for each unit. We modeled the network covariance matrix with a single factor, and took the diagonal of the network covariance matrix as the shared variability (Churchland et al., 2010). A separate single-factor model was learned for each layer, stimulus condition and penetration. The right panels in **Figures 4A-C** show the network covariance matrices for the same units and three stimulus diameters used for the raster plots. In all layers, shared variability was highest when the smallest of the three stimuli was presented, and dropped to near zero in response to the 0.5° stimulus (0.1° vs. 0.5° mean±s.e.m.: SG, 1.53±0.18 vs. 0.48±0.08, n=4; G, 1.07±0.18 vs. 0.16±0.02, n=3; IG, 0.74±0.21 vs. 0.30±0.06, n=7). The impact of larger stimuli on shared variability, instead, depended on layer. Compared to shared variability in response to the 0.5° stimulus, shared variability in response to the 26° stimulus increased in SG layers (0.5° vs. 26° mean±s.e.m.: 0.48±0.08 vs. 0.74±0.20, n=4), did not change in G layers (0.16±0.02 vs. 0.14±0.05, n=3), and decreased in IG layers (0.30±0.06 vs. 0.14±0.04, n=7). The bottom panel of **Figure 4A-C** shows, for the same example units, shared variability as a function of stimulus diameter for all diameters used in our study (0.1-26°).

Figure 5A shows mean firing rate and mean shared variance as a function of stimulus size, computed separately over the entire population of multi-units across all penetrations, in SG, G and IG layers. Shared variance was tuned for stimulus size in a manner that resembled the size tuning of Fano-factor (compare with **Fig. 2A**). In all layers, increasing the stimulus diameter from 0.1 to a size equal to the aggregate RF diameter of the recorded cells progressively increased firing-rate but decreased shared variance (**Fig. 5A-B**). Shared variance also decreased relative to baseline for a stimulus matched to the RF diameter (mean percent change ±s.e.m, t-test % change < 0: SG, -

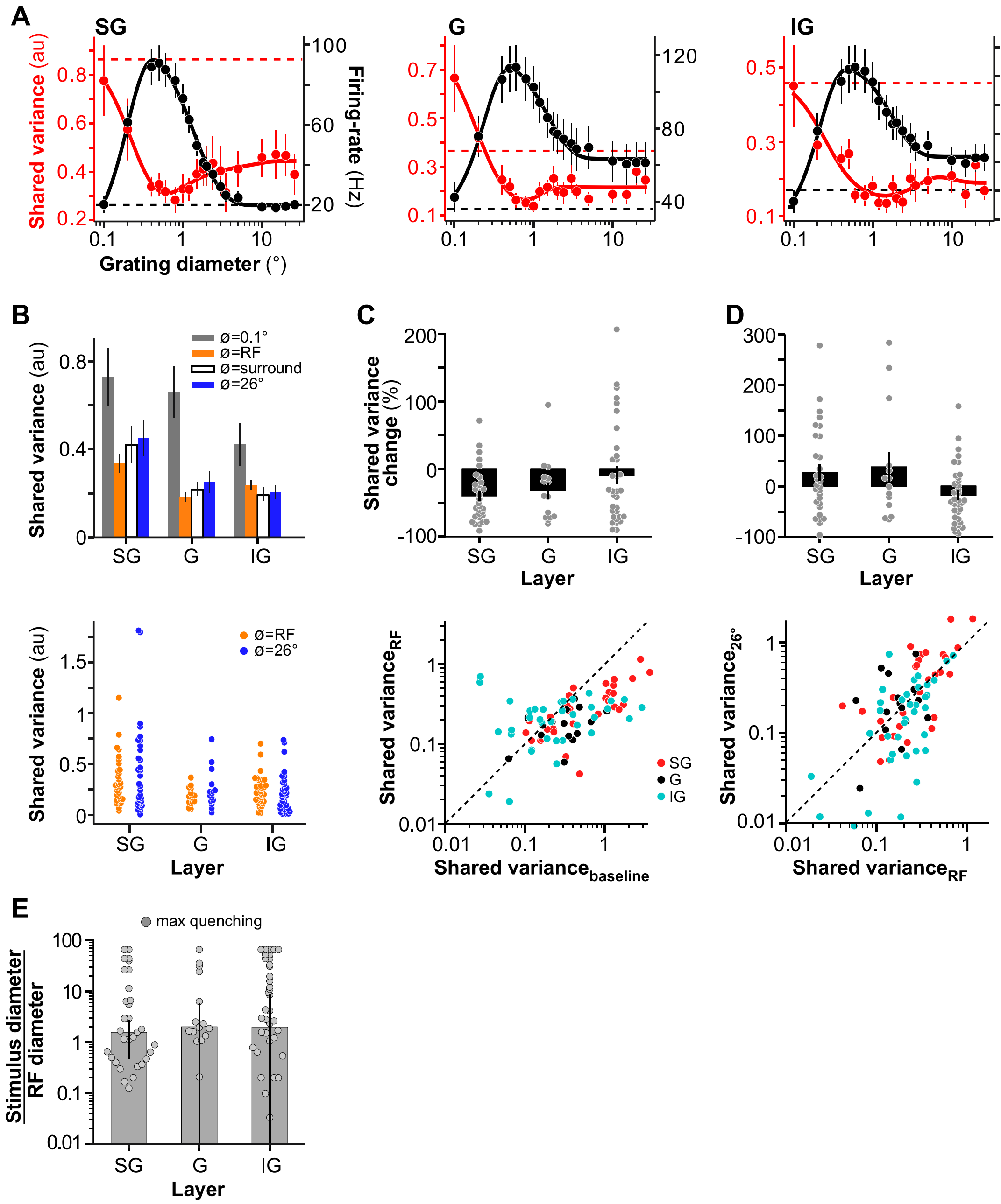

Figure 5. Size tuning of shared variance across V1 layers: population data. A) Mean firing rate (*black*) and mean shared variance (*red*) as a function of stimulus diameter, averaged over the population of recorded units separately for the different layers (from left to right: SG, n=31 units; G, n=15; IG, n=36). **B) Top:** shared variance averaged over the population at four different stimulus diameters (as indicated); shared variance values at specific stimulus sizes were extracted from functions fitted to the size-tuning data (see Methods). **Bottom:** single data points for the data in the top panel at the indicated two stimulus diameters. **C) Top:** Mean percent change in shared variance relative to baseline induced by a stimulus matched in size to the RF diameter for the different layers. *Dots here and in (D):* Individual data points. Error bars: s.e.m. **Bottom:** Scatter plot of shared variance during pre-stimulus baseline vs. during presentation of a stimulus matched to the RF diameter. Different colored dots indicate units in different layers. **D) Top:** Mean percent change in shared variance induced by a 26° diameter stimulus relative to the shared variance evoked by a stimulus matched to the RF diameter, for different layers. **Bottom:** Scatter plot of shared variance for presentation of stimuli of two different sizes (a diameter equal to that of the RF vs. a diameter of 26°). Other conventions as in panel C). **E)** Median stimulus diameter at the largest decrease in shared variance, normalized to the RF diameter of the recorded units, for different layers. Error bars: s.d. of the bootstrapped distributions. *Dots:* individual cell data.

±s.e.m. t-test % change < 0: -18.52±10.68, p=0.046). Consistent with these results, on average the stimulus diameter at maximum variability quenching (i.e. the lowest shared variance value) was close to the RF diameter for SG layers, but larger than the RF diameter in IG layers (**Fig. 5E**; median stimulus diameter at the lowest shared variance normalized to the RF diameter, ±s.d. of the bootstrapped median distribution: SG, 1.56±1.22; G, 1.94±3.81; IG, 1.93±3.83).

To rule out changes in firing rate as the main cause of changes in shared variance with stimulus size, we computed the average shared variance following the same mean-matching procedure performed for Fano-factor. This analysis was consistent with the results shown in **Fig. 5** (see Supplementary Methods and Supplementary Results and **Supplementary Fig. 2**).

DISCUSSION

Using linear multi-electrode array recordings, we have studied how neural response variability is modulated by stimulus size across the layers of macaque V1. We found that both single neuron response variability and the shared variability among neurons are size-tuned, and this tuning is layer dependent. In all layers, variability declined as a stimulus was progressively increased in size from 0.1° to the diameter of the RF. However, as the stimulus was enlarged beyond the RF, variability changed in a layer dependent manner. In SG layers, surround stimuli increased both single neuron and shared variability (relative to their value for a stimulus matched to the RF diameter), but did not change them or reduced them in G and IG layers. Given the hypothesized influence of variability on visual information processing and the encoding of sensory inputs, these laminar differences suggest that the different layers employ different strategies for coding large stimuli. Moreover, given known laminar differences in connectivity, the laminar specific effects of stimulus size on variability observed in our study suggest different underlying circuit mechanisms.

Theoretical work has shown that correlated variability can be detrimental for sensory processing (Abbott and Dayan, 1999; Averbeck et al., 2006). Consistent with this idea, a number of top-down modulations thought to improve sensory processing and perception tend to reduce spike-count variability (Fano-factor) and correlated variability. For example, attention directed towards a visual stimulus reduces correlated variability as well as Fano-factor in primate areas V4 (Cohen and Maunsell, 2009; Mitchell et al., 2009) and V1 (Herrero et al., 2013), and decorrelation is induced by perceptual learning in area MSTd (Gu et al., 2011) or surround suppression in V1 (Snyder et al., 2014). Our results on the effects of surround stimulation in IG layers are consistent with a previous study (Snyder et al., 2014) showing that surround suppression reduces correlated variability (although we used factor analysis as a measure of shared variability, unlike this previous study, which instead measured correlations). However, we have additionally shown that the effect of surround suppression on both Fano-factor and shared variability depends on cortical layer. Specifically, stimulation of the surround increased Fano-factor and shared variability in SG layers, relative to stimuli matched to the RF size, but decreased them in IG layers. It is unlikely that the increased variability in SG layers induced by presentation of large stimuli is detrimental for visual processing. The impact of surround modulation on visual processing ultimately depends on the way neuronal responses are readout, which may vary with cortical layer. It is also possible, that variability-increase in SG layers and decrease in IG layers induced by surround stimulation both facilitate encoding and perception. Indeed, theoretical and experimental work has indicated that not all correlations impede encoding (Averbeck et al., 2006), that the strength of variability and correlations depend on stimulus, cognitive factors, cortical layers and area (Hansen et al., 2012; Smith et al., 2013; Ruff and Cohen, 2016b, a), and that some form of correlations can facilitate, rather than impede, perception (Ruff and Cohen, 2014; Haefner et al., 2016). Thus, additional studies are necessary to determine how these laminar-specific modulations of variability affect perception and behavioral performance.

It is often assumed that the trial-to-trial variability of neural responses follows Poisson statistics (Simoncelli et al., 2004). Under Poisson statistics, variability does not depend on firing-rate; rather, Fano-factor remains constant (=1) regardless of mean firing-rate. In contrast, we found that, across all layers, firing-rate increased, and Fano-factor decreased as the stimulus diameter was increased from 0.1° to a diameter equal to that of the RF. These results are inconsistent with the Poisson model of neural response statistics. An extension of the Poisson model, the modulated Poisson model (Goris et al., 2014), augments the Poisson model by a stochastic gain variable. With this addition, the model captures overdispersion (Fano-factor > 1) of neural responses in a physiologically and statistically meaningful way. The modulated Poisson model predicts that Fano-factor increases as firing-rate increases (Goris et al., 2014). However, our data shows that for stimuli within the RF, Fano-factor decreases as mean firing-rate increases, which is inconsistent with the modulated Poisson model. Our results using small stimuli resemble those of Solomon and

Coen-Cagli (2019) who showed that, in macaque V1, Fano-factor decreased as stimulus contrast (and firing-rate) increased. These authors also concluded that their data is incompatible with the modulated Poisson model, and suggested that neural response statistics in macaque V1 are better captured by stochastic normalization models.

In G and IG layers, stimulation of the RF surround suppressed firing-rates, but did not significantly affect Fano-factors (but note that the mean-matched analysis revealed a significant decrease in Fano-factor in both layers for larger stimuli). Although we consistently observed Fano-factors above 1, the decoupling of Fano-factor and firing-rate for larger stimuli is roughly consistent with the Poisson model of neural response variability. A previous study measured Fano-factor in macaque V1 as a function of the diameter of a natural image patch and found that Fano-factor was not affected by stimulation of the surround (Festa et al., 2021). Although these authors did not report the laminar origin of their recordings, the latter were likely from the G layer, as they were performed using 1-mm shank length Utah arrays. While our G and IG layer results are consistent with the study by Festa et al. (modulo the mean-matched data), here we have additionally shown that surround stimulation has a different effect on Fano-factor in the SG layers.

The study by Festa et al. (2021), mentioned above, was designed to test these authors' own sampling-based model of probabilistic inference. These class of models are based on the idea of perception as probabilistic inference (Knill and Pouget, 2004); the type of probabilistic inference models favored by Festa et al. (2021) view spikes as representing samples from probability distributions, and neural response variability as the uncertainty of the inferences (Hoyer and Hyvärinen, 2003; Fiser et al., 2010; Orban et al., 2016; Echeveste et al., 2020; Festa et al., 2021). The model by Festa et al. (2021) predicted a decrease in Fano-factor induced by surround stimulation relative to its value for stimulation of the RF. These authors' neural recording results did not confirm this model prediction, as surround stimulation was found not to affect Fano-Factor (see above). Consistent with the model of Festa et al. (2021), instead, here we found that about 15% of cells across all layers (more numerous in IG layers) showed decreases in Fano-factor when their RF surround was stimulated. This points to the intriguing possibility that in all layers, a small, but significant, proportion of neurons serves to perform probabilistic inference.

The laminar differences in variability found in our study suggest different underlying circuit mechanisms. One plausible hypothesis is that laminar-specific inhibitory circuits underlie the different effects of surround stimulation on variability. This hypothesis is based on the assumption that inhibition plays an important role in modulating cortical response variability, as postulated by several models (Stringer et al., 2016; Hennequin et al., 2018; Huang et al., 2019), and the well-established laminar differences in the distribution of inhibitory neuron types. In macaque sensory-motor cortex, somatostatin-positive inhibitory interneurons predominate in SG layers (Hendry et al., 1984). In mouse cortex, these interneuron types subtractively control the gain of their target neurons (Sturgill and Isaacson, 2015), by hyperpolarizing the dendrites of pyramidal cells (Markram et al., 2004; Pouille et al., 2013), and mediate surround suppression (Adesnik et al., 2012). This subtractive inhibition counteracts the excitatory feedforward drive (Pouille et al., 2013), thus, ultimately affecting neural responses in a manner resembling reduced feedforward input. Because reduced feedforward input to the cortex increases neural response variability (e.g. Churchland et al., 2010; Festa et al., 2021; and this study), activation of somatostatin neurons ultimately would lead to increased neural response variability. This hypothesis predicts that neurons in which variability is increased by surround stimulation, which are present in all layers but dominate in the SG layers, are those in which surround suppression is mediated by somatostatin cells. Alternatively, reduced feedforward drive induced by surround stimulation, leading to increased variability, may result from withdrawal of feedforward excitation from surround-suppressed excitatory neurons in the thalamus or cortex itself, a mechanism that is consistent with a recent model of neural response variability (Bressloff, 2019). This hypothesis predicts that neurons in which variability is increased by surround stimulation are those inheriting surround suppression from other suppressed excitatory neurons. In contrast, neurons for which surround stimulation does not affect or decreases variability (more numerous in G and IG layers) may be surround suppressed via different circuit mechanisms; for example, via inhibitory cells that track the activity of excitatory cells (such as parvalbumin interneurons). This mechanism quenches variability in a stochastic inhibition-stabilized network model (Hennequin et al., 2018).

Quenching of cortical response variability by stimulus onset is considered to be a universal property of the cortex (Churchland et al., 2010). In line with this idea, we showed that presenting a stimulus most commonly quenched variability compared to pre-stimulus baseline. However, in addition to variability quenching, we found that, for a substantial fraction of cells in all layers, small stimuli amplified variability relative to pre-stimulus baseline. The use of very small stimuli and a cell-by-cell analysis were the key differences between our study and previous studies that failed to observe amplification of variability by small stimuli.

Amplification of cortical response variability by small visual stimuli relative to pre-stimulus baseline was predicted by a supralinear stabilized network model of cortical response variability (Hennequin et al., 2018). This model predicts variability amplification when the stimulus-evoked response is of comparable magnitude to spontaneous activity. In contrast, in our data, variability amplification was observed also when neural responses were significantly above the spontaneous baseline. Thus, the prediction of supralinear stabilized network models of cortical response variability are not in quantitative agreement with the results of this study. Variability amplification can trivially arise in units with close to zero baseline firing-rates, as it is often the case for anesthetized primate V1, because the variance of a spike-train with zero mean is necessarily zero and can only increase as the firing-rate increases. In our dataset, units that showed variability amplification had lower baseline firing rates and Fano-factor values, but also showed variability quenching for larger stimuli. Thus, a floor effect due to low baseline firing-rates cannot explain the variability amplification in our data.

A number of different dynamical models have been proposed to explain various aspects of stimulus-dependent variability. In models with multi-stable dynamics, response variability arises from the stochastic wandering across the cortex of spontaneously-formed tuning curves or bumps (Ponce-Alvarez et al., 2013; Bressloff, 2019; Huang et al., 2019). In these models, increased stimulus drive reduces wandering of the activity patterns, locking the stimulus-driven bump in place, and as a consequence quenching variability. Recently, one such model explicitly predicted an increase in variability by surround stimulation (Bressloff, 2019). This prediction is consistent with our results in SG layers, but not in G and IG layers, although this model captures well the experimentally-observed differences in the magnitude of variability across cortical layers in the spontaneous state (Smith et al., 2013). In a different class of models, instead, variability results from fluctuations about a single, stimulus-driven attractor in a stochastic stabilized supralinear network (Hennequin et al., 2018). In these models, when stimulus drive increases, the balanced network causes an increase in inhibition which leads to reduced variability. Thus, in these models, variability quenching results from increased inhibition, as opposed to the multiple-attractor models described above in which variability quenching results from increased excitation. Although the effects of surround suppression on variability have not been explicitly studied in stabilized supralinear network models, they would seem consistent with our results in G and IG layers, i.e. a reduction or saturation of variability by surround stimulation, but not in SG layers.

In summary, existing models of cortical response variability are either inconsistent, or only party consistent with our results, capturing the effects on variability of stimulus size we have observed for some but not all layers. We suspect that both kinds of mechanisms may occur, depending on particular cortical operating conditions, and the specific layer. Therefore, our results call for the extension of these existing models or the development of new models that can capture the laminar differences in the stimulus-dependent modulation of cortical response variability we have observed in our study.

MATERIALS AND METHODS

Experimental model

Linear array recordings were made in the parafoveal representation (4-8° eccentricity) of V1 in two anesthetized adult macaque monkeys (*Macaca Fascicularis, 1 male, 1 female*, 3-4 kg). Here we report recordings from a total of 82 contacts from 5 array penetrations. All experimental procedures were in accordance with protocols approved by the University of Utah Institutional Animal Care and Use Committee and with NIH guidelines.

Surgery

The surgical procedures are described in detail in our previous study (Bijanzadeh et al., 2018). Briefly, anesthesia was induced with ketamine (10 mg/kg, i.m.). An intravenous catheter and endotracheal tube were inserted, the head fixed in a stereotaxic apparatus, and the animal was artificially ventilated with a 70:30 mixture of O_2 and N_20. End-tidal CO_2, blood O_2 saturation, electrocardiogram, blood pressure, lung pressure, and body temperature were monitored continuously. A small craniotomy and durotomy were performed over the opercular region of V1 and a PVC chamber was glued to the skull surrounding the craniotomy and filled with agar and silicon oil to prevent cortical pulsation and dehydration, respectively. On completion of the surgery, and after a stable plane of anesthesia was reached, the animal was paralyzed with vecuronium bromide (0.3 mg/kg/h, i.v.), to prevent eye movements. Recordings were performed under continuous infusion of sufentanil citrate anesthesia (4-12 µg/kg/h). The pupils were dilated with topical atropine, and the corneas were protected with gas-permeable contact lenses. The eyes were refracted using corrective lenses, and the foveae were plotted on a tangent screen using a reverse ophthalmoscope, and periodically remapped throughout the experiment.

Electrophysiological recordings

To record the activity of V1 neurons across cortical layers, 24-channel linear arrays (V-Probe, Plexon, Dallas, Texas, 100 µm contact spacing and 20 µm contact diameter) were inserted into area V1, perpendicular to the pial surface to a depth of 2.0-2.2 mm. A custom-made guide tube provided mechanical stability to the array. To facilitate post-mortem visualization of the lesion tracks, the probes were coated with DiI (Molecular Probes, Eugene, OR) prior to insertion. We recorded extracellularly multiunit spiking activity (MUA) and local field potentials (LFP). The signals were amplified, digitized, and sampled at 30 kHz using a 128- system (Cerebus,16-bit A-D, Blackrock Microsystems, Salt Lake City, UT).

Multi-unit selection

All analysis was performed on MUA. MUA was detected by bandpass filtering continuous voltage traces and thresholding the filtered trace at 4 times the background noise standard deviation, estimated as the median of the continuous recording divided by 0.6745 (Quiroga et al., 2004). The analyses were done only on multi-units in which the most strongly driving stimulus evoked at least 3 spikes above the spontaneous activity (count window 50-350ms after the stimulus onset). Moreover, only multi-units in which the response was tuned for stimulus size were analyzed. Whether a unit was statistically significantly tuned for stimulus size was determined by performing ANOVA on the stimulus evoked spike counts. The units in which the effect of stimulus size was statistically significant (one-way ANOVA $p < 0.05$) were consider size-tuned. In addition, only those units which showed at least 5% surround suppression, defined as percent reduction in spike count from peak evoked by a 26° diameter stimulus, were included in the final analysis.

Visual stimuli

Visual stimuli were generated using Matlab (Mathworks Inc., Natick, MA; RRID:SCR_001622) and presented on a calibrated CRT monitor (Sony, GDM-C520K, 600x800 pixels, 100Hz frame rate, mean luminance 45.7cd/m^2, at 57cm viewing distance), and their timing was controlled using the ViSaGe system (Cambridge Research Systems, Cambridge, UK; RRID:SCR_000749). All stimuli were displayed for 500 ms, followed by 750 ms interstimulus interval.

We quantitatively mapped the minimum response field (mRF) of units across contacts by flashing a 0.5° black square stimulus over a 3x3° visual field area. The aggregate mRF of the column was defined as the visual field region in which the square stimulus evoked a mean response (+2 s.d. of the stimulus evoked response) that was > 2 s.d. above mean spontaneous activity, and the geometric center of this region was taken as the multi-units aggregate RF center. All subsequent stimuli were centered on this field. We then determined orientation, eye dominance, spatial and temporal frequency preferences of cells across contacts using 1-1.5° diameter drifting sinusoidal grating patches of 100% contrast presented monocularly. Subsequent stimuli were presented at the optimal parameters for most units across the column. We measured size tuning across the column using 100% contrast drifting grating patches of increasing size (0.1-26°) centered over the aggregate mRF of the column. To monitor eye movements, the RFs were remapped by hand approximately every 10-20 minutes and stimuli re-centered on the RF if necessary. To ensure that the array was positioned orthogonal to the cortical surface, we used as criteria the vertical alignment of the mapped mRFs at each contact, and the similarity in the orientation tuning curves across contacts. If RFs were misaligned across contacts, the array was retracted and repositioned.

Quantification and Statistical Analysis

Current Source Density (CSD) analysis

We used CSD responses to small stimuli flashed inside the RFs to identify laminar borders (as detailed in the Results). CSD analysis was applied to the band-pass filtered (1-100Hz) and trial averaged LFP using the kernel CSD toolbox (kCSD_Matlab) (Potworowski et al., 2012). CSD was calculated as the second spatial derivative of the LFP signal. To estimate CSD across layers, we interpolated the CSD every 10μm. The CSD was baseline corrected (Z-scored). In particular, we normalized the CSD of each profile to the s.d. of the baseline (defined as 200ms prior to stimulus onset) after subtraction of the baseline mean (see Bijanzadeh et al. 2018 for details).

Fano-factor

To quantify trial-to-trial variability, we computed Fano-factor by dividing the spike-count variance by the mean spike-count over trials. A small constant (0.0000001) was added to the mean spike-count to avoid dividing by zero. During the course of developing the analysis, we also used a method in which spike-count variance was plotted against mean spike-count computed over trials in 100 ms non-overlapping bins, and by fitting to the variance-to-mean curves a line so that the intersection of the line and the y-axis was constrained to be zero and the slope of the line was taken as the Fano-factor. All findings of the study were replicated using both methods, but we chose the direct division for convenience as it allows for more efficient bootstrapping of errors. All of our analyses were performed between 50 to 450 ms after stimulus onset, except for the pre-stimulus baseline that was computed from -400 to 0 ms before stimulus onset.

To determine the significance of the different effects of surround stimulation on Fano-factor (the data in **Fig. 2E**), we re-sampled Fano-factors 3000 times with replacement from the distributions measured at the RF size and at 26° stimulus diameter. The means of these two distributions were replaced with a common mean (mean of means), and a bootstrapped distribution of Fano-factor difference was generated by subtracting the values in each re-sampled distribution. If Fano-factor measured at 26° stimulus diameter was larger (smaller) than Fano-factor measured at the RF size, and this difference was above the 95th (below the 5th) percentile of the bootstrapped distribution of Fano-factor difference, we concluded that stimulation of the RF surround increased (decreased) Fano-factor relative to the RF-only. All other results were interpreted as surround stimulation having no effect on Fano-factor.

To determine whether a stimulus caused statistically significant increase or decrease in Fano-factor relative to baseline (the analyses presented in **Fig. 3**), the distribution of the difference between Fano-factor and baseline at each stimulus size was resampled with replacement 3000 times. The mean of this distribution was set to zero. If the Fano-factor measured at a given stimulus diameter was higher (smaller) than the 95th (5th) percentile of this distribution, we concluded that the stimulus significantly increased (decreased) Fano-factor relative to baseline.

For details on the mean-matched Fano-factor analysis see Supplementary Methods.

Function fitting and receptive field size estimation

To estimate the size of the RF center and surround for each unit, we measured size tuning as described above and plotted the mean firing rate of the unit against stimulus diameter; we, then, fitted these data with ratio-of-Gaussians functions (Cavanaugh et al., 2002). The Fano-factor data was fitted with two ratio-of-Gaussians functions that were summed. These two ratio-of-Gaussians functions had independent parameters. The parameters were optimized by minimizing the squared difference between the function and the data. For firing-rate, the minimization was performed with the Levenberg-Marquardt algorithm as implemented in SciPy (Virtanen et al., 2020). The parameters of the function were constrained to be positive, including zero. For Fano-factor, the parameters of the function were fitted with the basinhopping algorithm as implemented in Scipy. As that the sum of two ratio-of-Gaussians function was overfitting the data, we constrained the parameters to be always positive with an upper bound between 1 and 100, depending on the parameter. With these constraints, the fitted functions were always smooth. Two ratio-of-Gaussians functions were also fitted to the shared variance data.

From the fitted functions, the size of the RF center was taken to be the stimulus diameter at which the function peaked. The size of the surround was taken to be the smallest stimulus diameter, larger than the RF size, at which the slope of the fitted size-tuning function was at least 10% higher than the slope at the RF size. The slope was computed at all stimulus sizes between the RF size and 26°.

Factor Analysis

We used factor analysis to decompose the trial-to-trial spike-count covariance matrix into private (single neuron spiking variability) and shared (network) components. Factor-analysis was separately performed for each penetration, stimulus condition, and layer. A 300 ms-window was used. Given that neural response variability is low-dimensional in the visual cortex (Huang et al., 2019), and that our columnar recordings recover a subspace of the full-dimensional response space, we used just one factor to model the covariances. The covariance matrices were modeled as the product of the factor loading matrix and its transpose, plus a diagonal matrix containing the variances that are private to each unit. The matrix of factor loadings and the diagonal private variance matrix were estimated with the Gaussian-process factor analysis toolbox of Yu et al. (2009). As an estimate of the shared variance for each unit, we used the diagonal components of the matrix that results from multiplying the factor loading matrix with its own transpose.

ACKNOWLDEGMENTS

We thank Kesi Sainsbury for technical assistance, and Drs. Frederick Federer and Sam Merlin for help with experiments. Supported by grants from the National Institute of Health to A.A. (R01

EY026812, R01 EY019743, R01 EY031959, BRAIN U01 NS099702) and to L.N. (K99/R00 EY029374), the National Science Foundation (IOS 1755431) and the Mary Boesche endowed Professorship, to A.A., a Postdoctoral Fellowship to L.N. from the Ella and Georg Ehrnrooth Foundation, an unrestricted grant from Research to Prevent Blindness, Inc. and a core grant from the National Institute of Health (EY014800) to the Department of Ophthalmology, University of Utah.

COMPETING INTERESTS

The authors declare no competing interests.

REFERENCES

Abbott LF, Dayan P (1999) The effect of correlated variability on the accuracy of a population code. Neural Comput 11:91-101.

Adesnik H, Bruns W, Taniguchi H, Huang ZJ, Scanziani M (2012) A neural circuit for spatial summation in visual cortex. Nature 490:226-231.

Angelucci A, Shushruth S (2013) Beyond the classical receptive field: surround modulation in primary visual cortex. In: The new visual neurosciences (Chalupa LM, Werner JS, eds), pp 425-444. Cambridge: MIT press.

Angelucci A, Levitt JB, Walton E, Hupé JM, Bullier J, Lund JS (2002) Circuits for local and global signal integration in primary visual cortex. J Neurosci 22:8633-8646.

Angelucci A, Bijanzadeh M, Nurminen L, Federer F, Merlin S, Bressloff PC (2017) Circuits and mechanisms for surround modulation in visual cortex. Ann Rev Neurosci 40:425-451. doi:doi: 10.1146/annurev-neuro-072116-031418

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

Bair W, Zohary E, Newsome WT (2001) Correlated firing in macaque visual area MT: time scales and relationship to behavior. J Neurosci 21:1676-1697.

Bastos AM, Usrey WM, Adams RA, Mangun GR, Fries P, Friston KJ (2012) Canonical microcircuits for predictive coding. Neuron 76:695-711.

Bijanzadeh M, Nurminen L, Merlin S, Clark AM, Angelucci A (2018) Distinct Laminar Processing of Local and Global Context in Primate Primary Visual Cortex. Neuron 100:259-274 e254. doi:10.1016/j.neuron.2018.08.020

Bressloff PC (2019) Stochastic neural field model of stimulus-dependent variability in cortical neurons. PLoS Comput Biol 15:e1006755.

Cavanaugh JR, Bair W, Movshon JA (2002) Nature and interaction of signals from the receptive field center and surround in macaque V1 neurons. J Neurophysiol 88:2530-2546.

Churchland MM et al. (2010) Stimulus onset quenches neural variability: a widespread cortical phenomenon. Nat Neurosci 13:369-378. doi:10.1038/nn.2501

Coen-Cagli R, Solomon SS (2019) Relating Divisive Normalization to Neuronal Response Variability. J Neurosci 39:7344-7356. doi:10.1523/JNEUROSCI.0126-19.2019

Cohen MR, Maunsell JH (2009) Attention improves performance primarily by reducing interneuronal correlations. Nat Neurosci 12:1594-1600.

Constantinople CM, Bruno RM (2013) Deep cortical layers are activated directly by thalamus. Science 340:1591-1594.

Douglas RJ, Martin KA (2004) Neuronal circuits of the neocortex. Annu Rev Neurosci 27:419-451. doi:10.1146/annurev.neuro.27.070203.144152

Echeveste R, Aitchison L, Hennequin G, Lengyel M (2020) Cortical-like dynamics in recurrent circuits optimized for sampling-based probabilistic inference. Nat Neurosci 23:1138-1149. doi:10.1038/s41593-020-0671-1

Festa D, Aschner A, Davila A, Kohn A, Coen-Cagli R (2021) Neuronal variability reflects probabilistic inference tuned to natural image statistics. Nature Communications 12:1-11. doi:https://doi.org/10.1101/2020.06.17.142182

Fiser J, Berkes P, Orban G, Lengyel M (2010) Statistically optimal perception and learning: from behavior to neural representations. Trends Cogn Sci 14:119-130. doi:10.1016/j.tics.2010.01.003

Goris RL, Movshon JA, Simoncelli EP (2014) Partitioning neuronal variability. Nat Neurosci 17:858-865. doi:10.1038/nn.3711

Gu Y, Liu S, Fetsch CR, Yang Y, Fok S, Sunkara A, DeAngelis GC, Angelaki DE (2011) Perceptual learning reduces interneuronal correlations in macaque visual cortex. Neuron 71:750-761.

Haefner RM, Berkes P, Fiser J (2016) Perceptual Decision-Making as Probabilistic Inference by Neural Sampling. Neuron 90:649-660. doi:10.1016/j.neuron.2016.03.020

Hansen BJ, Chelaru MI, Dragoi V (2012) Correlated variability in laminar cortical circuits. Neuron 76:590-602.

Henaff OJ, Boundy-Singer ZM, Meding K, Ziemba CM, Goris RLT (2020) Representation of visual uncertainty through neural gain variability. Nat Commun 11:2513. doi:10.1038/s41467-020-15533-0

Hendry SH, Jones EG, Emson PC (1984) Morphology, distribution, and synaptic relations of somatostatin- and neuropeptide Y-immunoreactive neurons in rat and monkey neocortex. J Neurosci 4:2497-2517.

Hennequin G, Ahmadian Y, Rubin DB, Lengyel M, Miller KD (2018) The Dynamical Regime of Sensory Cortex: Stable Dynamics around a Single Stimulus-Tuned Attractor Account for Patterns of Noise Variability. Neuron 98:846-860 e845. doi:10.1016/j.neuron.2018.04.017

Henry CA, Joshi S, Xing D, Shapley RM, Hawken MJ (2013) Functional characterization of the extraclassical receptive field in macaque V1: contrast, orientation, and temporal dynamics. J Neurosci 33:6230-6242. doi:10.1523/JNEUROSCI.4155-12.2013

Herrero JL, Gieselmann MA, Sanayei M, Thiele A (2013) Attention-induced variance and noise correlation reduction in macaque V1 is mediated by NMDA receptors. Neuron 78:729-739.

Hoyer P, Hyvärinen A (2003) Interpreting neural response variability as Monte Carlo sampling of the posterior. In: Advances in Neural Information Processing Systems (Becker S, ed), pp 277-284: MIT Press.

Huang C, Ruff DA, Pyle R, Rosenbaum R, Cohen MR, Doiron B (2019) Circuit Models of Low-Dimensional Shared Variability in Cortical Networks. Neuron 101:337-348 e334. doi:10.1016/j.neuron.2018.11.034

Ichida JM, Schwabe L, Bressloff PC, Angelucci A (2007) Response facilitation from the "suppressive" receptive field surround of macaque V1 neurons. J Neurophysiol 98:2168-2181. doi:10.1152/jn.00298.2007

Knill DC, Pouget A (2004) The Bayesian brain: the role of uncertainty in neural coding and computation. Trends Neurosci 27:712-719. doi:10.1016/j.tins.2004.10.007

Levitt JB, Lund JS (2002a) Intrinsic connections in mammalian cerebral cortex. In: Cortical areas unity and diversity (Miller R, Schuz A, eds), pp 133-154: CRC Press.

Levitt JB, Lund JS (2002b) The spatial extent over which neurons in macaque striate cortex pool visual signals. Vis Neurosci 19:439-452.

Markram H, Toledo-Rodriguez M, Wang Y, Gupta A, Silberberg G, Wu C (2004) Interneurons of the neocortical inhibitory system. Nat Rev Neurosci 5:793-807. doi:10.1038/nrn1519

Mitchell JF, Sundberg KA, Reynolds JH (2009) Spatial attention decorrelates intrinsic activity fluctuations in macaque area V4. Neuron 63:879-888.

Mitzdorf U (1985) Current source-density method and application in cat cerebral cortex: investigation of evoked potentials and EEG phenomena. Physiol Rev 65:37-100.

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

Nandy AS, Sharpee TO, Reynolds JH, Mitchell JF (2013) The fine structure of shape tuning in area V4. Neuron 78:1102-1115. doi:10.1016/j.neuron.2013.04.016

Nassi JJ, Callaway EM (2009) Parallel processing strategies of the primate visual system. Nat Rev Neurosci 10:360-372. doi:10.1038/nrn2619

Nurminen L, Merlin S, Bijanzadeh M, Federer F, Angelucci A (2018) Top-down feedback controls spatial summation and response amplitude in primate visual cortex. Nature Commun 9:2281.

Orban G, Berkes P, Fiser J, Lengyel M (2016) Neural Variability and Sampling-Based Probabilistic Representations in the Visual Cortex. Neuron 92:530-543. doi:10.1016/j.neuron.2016.09.038

Ponce-Alvarez A, Thiele A, Albright TD, Stoner GR, Deco G (2013) Stimulus-dependent variability and noise correlations in cortical MT neurons. Proc Natl Acad Sci U S A 110:13162-13167. doi:10.1073/pnas.1300098110

Potworowski J, Jakuczun W, Leski S, Wojcik D (2012) Kernel current source density method. Neural Comput 24:541-575.

Pouille F, Watkinson O, Scanziani M, Trevelyan AJ (2013) The contribution of synaptic location to inhibitory gain control in pyramidal cells. Physiol Rep 1:e00067. doi:10.1002/phy2.67

Quiroga RQ, Nadasdy Z, Ben-Shaul Y (2004) Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. Neural Comput 16:1661-1687. doi:10.1162/089976604774201631

Ringach DL, Hawken MJ, Shapley R (1997) Dynamics of orientation tuning in macaque primary visual cortex. Nature 387:281-284. doi:10.1038/387281a0

Rubin DB, Van Hooser SD, Miller KD (2015) The stabilized supralinear network: a unifying circuit motif underlying multi-input integration in sensory cortex. Neuron 85:402-417.

Ruff DA, Cohen MR (2014) Attention can either increase or decrease spike count correlations in visual cortex. Nat Neurosci 17:1591-1597.

Ruff DA, Cohen MR (2016a) Stimulus Dependence of Correlated Variability across Cortical Areas. J Neurosci 36:7546-7556. doi:10.1523/JNEUROSCI.0504-16.2016

Ruff DA, Cohen MR (2016b) Attention Increases Spike Count Correlations between Visual Cortical Areas. J Neurosci 36:7523-7534. doi:10.1523/JNEUROSCI.0610-16.2016

Sceniak MP, Hawken MJ, Shapley RM (2001) Visual spatial characterization of macaque V1 neurons. J Neurophysiol 85:1873-1887.

Schwabe L, Obermayer K, Angelucci A, Bressloff PC (2006) The role of feedback in shaping the extra-classical receptive field of cortical neurons: a recurrent network model. J Neurosci 26:9117-9129. doi:10.1523/JNEUROSCI.1253-06.2006

Schwabe L, Ichida JM, Shushruth S, Mangapathy P, Angelucci A (2010) Contrast-dependence of surround suppression in Macaque V1: experimental testing of a recurrent network model. Neuroimage 52:777-792. doi:10.1016/j.neuroimage.2010.01.032

Shadlen MN, Newsome WT (1998) The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. J Neurosci 18:3870-3896.

Shushruth S, Ichida JM, Levitt JB, Angelucci A (2009) Comparison of spatial summation properties of neurons in macaque V1 and V2. J Neurophysiol 102:2069-2083.

Simoncelli EP, Paninski L, Pillow J, Schwartz O (2004) Characterization of neural responses with stochastic stimuli. In: The New Cognitive Neurosciences (Gazzaniga M, ed): MIT Press.

Smith MA, Jia X, Zandvakili A, Kohn A (2013) Laminar dependence of neuronal correlations in visual cortex. Journal of neurophysiology 109:940-947. doi:10.1152/jn.00846.2012

Snyder AC, Morais MJ, Kohn A, Smith MA (2014) Correlations in V1 are reduced by stimulation outside the receptive field. J Neurosci 34:11222-11227.

Stringer C, Pachitariu M, Steinmetz NA, Okun M, Bartho P, Harris KD, Sahani M, Lesica NA (2016) Inhibitory control of correlated intrinsic variability in cortical networks. Elife 5. doi:10.7554/eLife.19695

Sturgill JF, Isaacson JS (2015) Somatostatin cells regulate sensory response fidelity via subtractive inhibition in olfactory cortex. Nat Neurosci 18:531-535. doi:10.1038/nn.3971

Takahashi N, Ebner C, Sigl-Glockner J, Moberg S, Nierwetberg S, Larkum ME (2020) Active dendritic currents gate descending cortical outputs in perception. Nat Neurosci 23:1277-1285. doi:10.1038/s41593-020-0677-8

Tolhurst DJ, Movshon JA, Dean AF (1983) The statistical reliability of signals in single neurons in cat and monkey visual cortex. Vision Res 23:775-785. doi:10.1016/0042-6989(83)90200-6

Virtanen P et al. (2020) SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat Methods 17:261-272. doi:10.1038/s41592-019-0686-2

Yu BM, Cunningham JP, Santhanam G, Ryu SI, Shenoy KV, Sahani M (2009) Gaussian-process factor analysis for low-dimensional single-trial analysis of neural population activity. J Neurophysiol 102:614-635. doi:10.1152/jn.90941.2008

FIGURE LEGENDS

Figure 1. Size tuning of Fano-factor and mean firing rate in macaque V1: representative units. A) Representative supragranular (SG) layer unit. **Left:** MUA spike-rasters measured at two stimulus diameters, either a diameter equal to the RF diameter of the recorded multi-unit (Top), or a diameter of 26° (Bottom). **Middle**: Peri-stimulus time histograms (PSTHs) of Fano-factor (*red*) and mean firing-rate (*black*) computed in a 100 ms rectangular sliding window for the same two stimulus diameters. The shaded area represents the standard deviation (s.d.) of the bootstrapped Fano-factor distribution (for the Fano-factor curve) or the standard-error-of-the-mean (s.e.m., for the firing rate curve). *Inset***:** Zoomed-in Fano-factor curves for the smaller (darker red) and larger (lighter red) stimulus diameters between 50 and 350 ms after stimulus onset. **Right**: Fano-factor (*red*) and firing-rate (*black*) averaged over 50-350 ms after stimulus onset and plotted against the stimulus diameter. *Solid lines*: fits to the data. *Dashed lines*: baseline Fano-factor (*red)* and firing rate (*black*), measured prior to stimulus onset. Error bars are: s.d. of the bootstrapped Fano-factor distribution (*red*) or s.e.m. (*black*). **B)** Representative granular (G) layer unit. **C)** Representative infragranular (IG) layer unit. Conventions in (B-C) are as in (A).

Figure 2. Size tuning of Fano-factor and mean firing rate: population data. A) Average Fano-factor (*red*) and mean firing-rate (*black*) as a function of stimulus diameter for the population of SG (**Left**; n=31), G (**Middle**; n=15) and IG (**Right**; n=36) layer units. *Dashed lines*: average baseline Fano-factor (*red)* and firing rate (*black*); error bars: s.e.m. **B) Top:** Fano-factor values averaged over 82 units at four different stimulus diameters (0.1°, a diameter equal to the RF diameter, a diameter equal to the RF-surround diameter (see Methods for definition), and 26°). Error bars: s.e.m. **Bottom:** Fano-factor values for individual multi-units in SG, G and IG layers at two different stimulus diameters (as indicated). **C) Top:** Mean percent change in Fano-factor relative to baseline induced by a stimulus matched in size to the RF diameter, for the different layers. *Dots:* Individual data points. Error bars: s.e.m**. Bottom:** Scatter plot of Fano-factor during pre-stimulus baseline vs during presentation of a stimulus matched to the RF diameter. Different colored dots indicate units in different layers. **D) Top:** Mean percent change in Fano-factor induced by a 26° diameter stimulus relative to the Fano-factor value evoked by a stimulus matched to the RF diameter. **Bottom:** Scatter plot of Fano-factor for presentation of stimuli of two different sizes

(a diameter equal to that of the RF vs. a diameter of 26°). Other conventions as in panel C). **E)** Percent of multi-units in each layer for which stimulation of the RF significantly decreased variability (*black*), did not affect variability (*white*) or increased variability (*gray*). **F)** Median stimulus diameter at the largest decrease in Fano-factor (or max quenching), normalized to the RF diameter of the recorded units, for different layers. Error bars: s.d. of the bootstrapped distributions. *Dots:* individual cell data. **G)** Scatter plot of percent change in Fano-factor evoked by the largest surround stimulus (26° diameter) relative to the Fano-factor evoked by a stimulus matched to the RF diameter vs. suppression index (see Methods). Color dots identify units in different layers, as indicated. Lines are regression lines fitted to the individual layer data.

Figure 3. Amplification of cortical response variability by small visual stimuli. A) An example unit showing stimulus-evoked increases in firing rate and Fano-factor for a small (0.1°) grating diameter (**Top**), but a decrease in Fano-factor for a larger stimulus matching the RF diameter (1°; **Bottom**). **Left:** Spike rasters. **Middle:** PSTHs of firing-rate computed in a 100ms sliding window. *Shaded gray area* here and in B) indicates the s.e.m. computed over trials. **Right:** Fano-factor computer over 100ms sliding window. *Shaded red area* here and in B) is the s.d. of the Fano-factor distribution bootstrapped over trials. **B)** Population-averaged time course of Fano-factor (*red*) and firing-rate (*black*) in SG **(Left),** G **(Middle),** and IG **(Right)** layers computed at two stimulus diameters (**Top:** 0.1°, **Bottom:** 0.8°). Both the Fano-factor and firing rate were normalized to the pre-stimulus baseline of each unit before averaging. **C)** Median stimulus diameter evoking the largest magnitude increase in Fano-factor, normalized to the RF diameter of the recorded units, for different layers. Error bars: s.d. of the bootstrapped distributions. *Dots* here and in (D-E)*:* individual cell data. **D)** Median difference in Fano-factor (Fano-factor at the stimulus diameter causing the largest change in Fano-factor minus the baseline Fano-factor) ± s.d. of the bootstrapped distributions at max quenching (*gray*) and max amplification (*white*) for different layers. **E)** Mean baseline Fano-factor for amplifier (*green*) and quencher (*yellow*) units. Error bars: s.e.m.

Figure 4. Size tuning of shared variance across V1 layers: an example penetration. A) Left: raster plots showing the spike times of four simultaneously recorded SG neurons, across several trials, in response to 0.1° (Top), 0.5° (Middle), and 26° (Bottom) diameter gratings. The responses of all 4 neurons in a single trial are shown between two consecutive horizontal lines. Horizontal

lines separate different trials. **Right:** Network covariance matrices estimated with a single-factor factor analysis for each of the same 3 different stimulus diameters. The diagonal of the network covariance matrix holds the shared variance for each recorded unit. **Bottom:** Shared variance as a function of stimulus diameter. The red markers show mean±s.e.m. of the shared variance computed over the SG neuron population recorded in this example penetration (n=4). The gray curves show the data for the individual 4 units. **B-C)** same as in A), but for G (n=3) and IG (n=7) layer units.

Figure 5. Size tuning of shared variance across V1 layers: population data. A) Mean firing rate (*black*) and mean shared variance (*red*) as a function of stimulus diameter, averaged over the population of recorded units separately for the different layers (from left to right: SG, n=31 units; G, n=15; IG, n=36). **B) Top:** shared variance averaged over the population at four different stimulus diameters (as indicated); shared variance values at specific stimulus sizes were extracted from functions fitted to the size-tuning data (see Methods). **Bottom:** single data points for the data in the top panel at the indicated two stimulus diameters. **C) Top:** Mean percent change in shared variance relative to baseline induced by a stimulus matched in size to the RF diameter for the different layers. *Dots here and in (D):* Individual data points. Error bars: s.e.m. **Bottom:** Scatter plot of shared variance during pre-stimulus baseline vs. during presentation of a stimulus matched to the RF diameter. Different colored dots indicate units in different layers. **D) Top:** Mean percent change in shared variance induced by a 26° diameter stimulus relative to the shared variance evoked by a stimulus matched to the RF diameter, for different layers. **Bottom:** Scatter plot of shared variance for presentation of stimuli of two different sizes (a diameter equal to that of the RF vs. a diameter of 26°). Other conventions as in panel C). **E)** Median stimulus diameter at the largest decrease in shared variance, normalized to the RF diameter of the recorded units, for different layers. Error bars: s.d. of the bootstrapped distributions. *Dots:* individual cell data.

SUPPLEMENTARY MATERIAL

SUPPLEMENTARY METHODS

Mean-matched Fano-factor analysis

To ensure that the stimulus size-dependent modulation of variability was not a simple consequence of firing-rate modulation, we performed a mean-matched analysis (Mitchell et al., 2009; Churchland et al., 2010). First, the firing-rates of the recorded neurons were binned separately for the conditions to be compared. Second, neurons were randomly removed from the bin, until the bin had an equal number of neurons for both stimulus conditions. This procedure was repeated for all bins until the number of neurons in each bin was equal across the mean-matched conditions. Finally, fano-factors for the remaining neurons were computed as described for the main analysis.

Mean-matched Factor analysis

To mean-match the shared variability data, we first estimated shared variability for all neurons in the same way as described in the Results and Methods of the manuscript. Next, mean-matching of firing-rates was performed in the same way as described above for the Fano-factor analysis. Finally, for the units that survived mean-matching, shared variability was averaged and plotted in **Supplementary Fig. 2**.

SUPPLEMENTARY RESULTS

Mean-matched Fano-factor analysis

In all layers, the mean matched Fano-factor analysis showed a statistically significant decrease in Fano-factor for a stimulus diameter (0.4°) evoking a higher mean spike count, relative to a 0.2° diameter stimulus (mean±s.d of the mean matched Fano-factor PSTH at 0.2° vs 0.4° stimulus diameter: SG, 2.30±0.62 vs 1.78±0.33, $p < 0.0001$; G, 1.49±0.36 vs 1.31±0.23, $p < 0.0001$; IG, 1.85±0.26 vs 1.711±0.29, $p < 0.0001$; first and third column in **Supplementary Fig.1**). Note that for this comparison we selected slightly different stimulus diameters (0.2° and 0.4°, respectively) than for the analyses in **Figs. 2B-C** (0.1° and the RF diameter, respectively), to ensure a greater overlap in the mean-spike count distributions between the two stimulus diameters being compared, which is necessary for the mean-matched analysis.

For the SG layers, the mean-matched Fano-factor showed a significant increase for a stimulus diameter covering the RF-surround (26°) relative to a 0.4° stimulus inside the RF (mean±s.d. of the mean matched Fano-factor PSTH at 0.4° vs 26° stimulus diameter: 1.88±0.43 vs. 3.24±0.53, p<0.0001; second and fourth columns in **Supplementary Fig. 1A**). In contrast, mean-matched Fano-factors for these two stimulus diameters showed a statistically significant decrease in G and IG layers (mean±s.d. of the mean matched Fano-factor PSTH at 0.4° vs 26° stimulus diameter: G, 1.43±0.29 vs. 1.14±0.37; IG, 1.76±0.27 vs 1.20±0.16, p<0.0001; second and fourth columns in **Supplementary Fig. 1B-C**, respectively). Again, the choice of stimulus diameters for this mean-matched comparison was slightly different from that used for the analyses in **Fig. 2B,D**, in order to ensure a greater overlap in the mean-spike count distributions. Note that due to mean-matching against a different stimulus size, the mean Fano-factor for the 0.4° diameter stimulus is slightly different depending on whether it was matched against a 0.2° or a 26° diameter stimulus.

Mean-matched factor analysis

Mean-matched shared variance data showed the same trends as the non-mean-matched data. However, the differences in the mean-matched data were not statistically significant, likely because the mean-matching procedure reduces sample size. Thus, we cannot fully rule out the possibility that changes in firing-rate caused the changes in shared variance. Regardless, in the mean-matched data, shared variance decreased in all layers as the stimulus diameter was increased from 0.2 to 0.4° (mean±s.e.m: SG, 0.42±0.17 vs 0.32±0.08; G, 0.54±0.08 vs. 0.31±0.16; IG, 0.35±0.08 vs. 0.28±0.05). Instead, shared variance increased in SG and G layers as the stimulus diameter was increased from 0.4 to 26° (mean±s.e.m: SG, 0.32±0.12 vs 0.57±0.14; G, 0.17±0.06 vs. 0.28±0.02;). In IG layers, the shared variance decreased as the stimulus diameter was increased from 0.4 to 26° (mean±s.e.m: IG: 0.17±0.06 vs. 0.16±0.05).