Vision and locomotion shape the interactions between neuron types in mouse visual cortex

Mario Dipoppa, Adam Ranson[1], Michael Krumin, Marius Pachitariu, Matteo Carandini[2], and Kenneth D Harris[2]

University College London, London EC1V 9EL, United Kingdom

[1] Present address: Neurosciences & Mental Health Research Institute, Cardiff University, Cardiff CF24 4HQ
[2] Co-senior author

Summary

In the mouse primary visual cortex (V1), sensory responses are shaped by behavioral factors such as locomotion. These factors are thought to control a disinhibitory circuit, whereby interneurons expressing vasoactive intestinal peptide (*Vip*) inhibit those expressing somatostatin (*Sst*), disinhibiting pyramidal cells (Pyr). We measured the effect of locomotion on these neurons and on interneurons expressing parvalbumin (*Pvalb*) in layer 2/3 of mouse V1, and found inconsistencies with the disinhibitory model. In the presence of large stimuli, locomotion increased *Sst* cell responses without suppressing *Vip* cells. In the presence of small stimuli, locomotion increased *Vip* cell responses without suppressing *Sst* cells. A circuit model could reproduce each cell type's activity from the measured activity of other cell types, but only if we allowed locomotion to increase feedforward synaptic weights while modulating recurrent weights. These results suggest that locomotion alters cortical function by changing effective synaptic connectivity, rather than only through disinhibition.

Introduction

Neocortical interneurons can be divided into genetically distinguishable types, which are arranged in specific functional circuits (Jiang et al., 2015; Kepecs and Fishell, 2014; Markram et al., 2015; Pfeffer et al., 2013; Tasic et al., 2016; Tremblay et al., 2016; Zeisel et al., 2015). There is great interest in understanding how these circuits map to specific features of cortical operation, and how different interneuron types influence each other and ultimately shape the activity of excitatory neurons.

In recent years, a "disinhibitory model" has been proposed as a canonical feature of many cortical areas. This model rests on converging evidence that interneuron-selective interneurons, such as those expressing vasoactive intestinal peptide (*Vip*), principally target somatostatin (*Sst*) interneurons. This arrangement has been observed anatomically in the hippocampus (Acsady et al., 1996a; Acsady et al., 1996b), and functionally in auditory cortex (Pi et al., 2013), medial prefrontal cortex (Pi et al., 2013), frontal association cortex (Garcia-Junco-Clemente et al., 2017) and visual cortex (Fu et al., 2014; Karnani et al., 2016a; Pfeffer et al., 2013). *Sst* neurons in turn inhibit most other cortical neuronal classes (with the exception of other *Sst* cells), including pyramidal (Pyr) cells (Jiang et al., 2015; Karnani et al., 2016b; Pfeffer et al., 2013). There is thus a potentially disinhibitory circuit involving *Vip* cells, *Sst* cells, and Pyr cells (Fu et al., 2014).

In sensory cortex, this disinhibitory circuit seems consistent with the modulation of sensory responses by behavioral state seen during arousal, whisking, and locomotion. In barrel cortex, for instance, whisking causes activity to increase in *Vip* cells, decrease in *Sst* cells, and increase in Pyr dendrites (Gentet et al., 2012; Lee et al., 2013). Furthermore, while the activity of Pyr cells correlates strongly with that of *Pvalb* cells, as predicted by the disinhibitory model it correlates weakly or negatively with the activity of *Sst* cells (Gentet et al., 2010; Gentet et al., 2012). In visual cortex, *Vip* cells are activated when animals run (Fu et al., 2014; Reimer et al., 2014), so a disinhibitory circuit might also explain the increases in visually-driven activity seen during locomotion (Ayaz et al., 2013; Erisken et al., 2014; Fu et al., 2014; Niell and Stryker, 2010).

The disinhibitory circuit, however, makes a key prediction that is unevenly supported by the data: that locomotion should decrease the activity of *Sst* neurons. One study confirmed the prediction, reporting that locomotion significantly (though mildly) reduced *Sst* neuron activity (Fu et al., 2014). Other studies, however, report the opposite effect or mixed effects (Pakan et al., 2016; Polack et al., 2013; Reimer et al., 2014).

Indirect evidence to support this prediction comes from visual cortex, where locomotion decreases the strength of a phenomenon, size tuning, that is thought to depend on *Sst* interneurons. Many V1 neurons exhibit size tuning: their responses are reduced when stimulus size increases beyond a preferred size. *Sst* interneurons are thought to contribute to size tuning because they integrate inputs from wide regions of cortex (Adesnik et al., 2012; Zhang et al., 2014) and because suppressing their activity reduces the strength of size tuning (Adesnik et al., 2012). If locomotion reduced *Sst* cell activity via inhibition from *Vip* cells, it should similarly reduce size tuning. This prediction is correct at least in the deep layers of V1 (Ayaz et al., 2013) and also in lateral geniculate nucleus (LGN, Erisken et al., 2014), perhaps because of cortical feedback. In both regions, locomotion decreases size tuning.

To test the predictions of the disinhibitory model, and more generally to understand the roles and relationships of different types of interneurons, we systematically investigated the effect of locomotion on visual cortical responses to stimuli of different sizes. We used two-photon microscopy to measure the activity of genetically-identified *Sst*, *Vip*, and *Pvalb* interneurons, together with pyramidal cells identified genetically or based on the skewness of their calcium traces.

Our results indicate that locomotor modulation of each cell class depends critically on the conditions of visual stimulation. Accordingly, subtle differences in experimental conditions can explain many of the apparent contradictions between previous studies. Our results, moreover, indicate multiple ways in which the effects of locomotion contradict the disinhibitory model. Modeling suggests that the complex interaction between locomotion, stimulus size, and cell class can be explained instead by a simple reweighting of feedforward vs. recurrent synapses.

Results

We used two-photon imaging to measure the activity of Pyr neurons and of *Sst*, *Vip,* and *Pvalb* interneurons in mouse V1 (Figure 1 and Supplementary Figure 1). Mice were head-fixed and free to run on an air-suspended ball (Niell and Stryker, 2010) while viewing a grating in a circular window whose size ranged between 5° and 60° of visual angle (Figure 1A_1). We selected for analysis only cells whose receptive field and orientation tuning matched the stimulus, and corrected for out-of-focus fluorescence (Supplementary Figure 2; Chen et al., 2013; Peron et al., 2015).

Recording the activity of identified cell classes

To identify neurons belonging to a specific class we used one of two approaches (Figure 1A, B). In the first approach (Figure 1A_2-A_4), GCaMP6m was virally expressed in all neurons, in mice where a specific cell class was also labeled with tdTomato. This approach allowed us to record the activity of identified interneurons and of many unlabeled neurons, which are likely to comprise mostly Pyr cells, along with interneurons of all classes except the labeled one. In the second approach, we expressed the calcium indicator exclusively in a chosen cell class, either by injecting a *Cre*-dependent GCaMP6m virus into an appropriate transgenic driver line (Figure 1B_1 and Supplementary Figure 2B_2) or via a triple-transgenic line that expressed GCaMP6s specifically in superficial layer Pyr cells.

Interneurons of all three classes fired much more frequently than pyramidal neurons (Figure 1C, D). Consistent with the sparse firing of superficial-layer pyramidal cells in mouse visual cortex (Niell and Stryker, 2010), the fluorescence traces of a typical identified pyramidal cell showed rare isolated calcium events (Figure 1C_1). Because events with strong fluorescence were rare in these traces, the distribution of fluorescence was highly skewed (Figure 1D_1). By contrast, the typical calcium fluorescence traces of identified *Pvalb, Vip,* and *Sst* interneurons showed frequent calcium events (Figure 1C_2-C_4), and the corresponding distributions of fluorescence showed little skewness (Figure

1D$_2$-D$_4$). Finally, many concurrently-measured unlabeled neurons showed sparse activity (example in Figure 1C$_5$) and high skewness (Figure 1D$_5$), resembling the identified Pyr cells (Figure 1C$_1$, D$_1$).

These differences in skewness were robust across the population, and allowed us to identify putative Pyr cells by their higher sparseness (Figure 1E). The precise value of the threshold for skewness made little difference to our results; we chose a conservative value of 2.7, that provided a small false-positive error rate (fraction of *Pvalb* neurons exceeding threshold: 24/1,511; *Vip*: 29/1,385; *Sst*: 91/537, Figure 1E$_2$-E$_4$), while still ensuring a majority of pyramidal cells exceeded threshold (2,598/4,949, Figure 1E$_1$), for all methods of GCaMP expression (Supplementary Figure 3). Unlabeled neurons exceeding this threshold value are therefore likely to be pyramidal neurons, while cells below the value could be any cell type, and were therefore excluded from further analysis (Figure 1E$_5$).

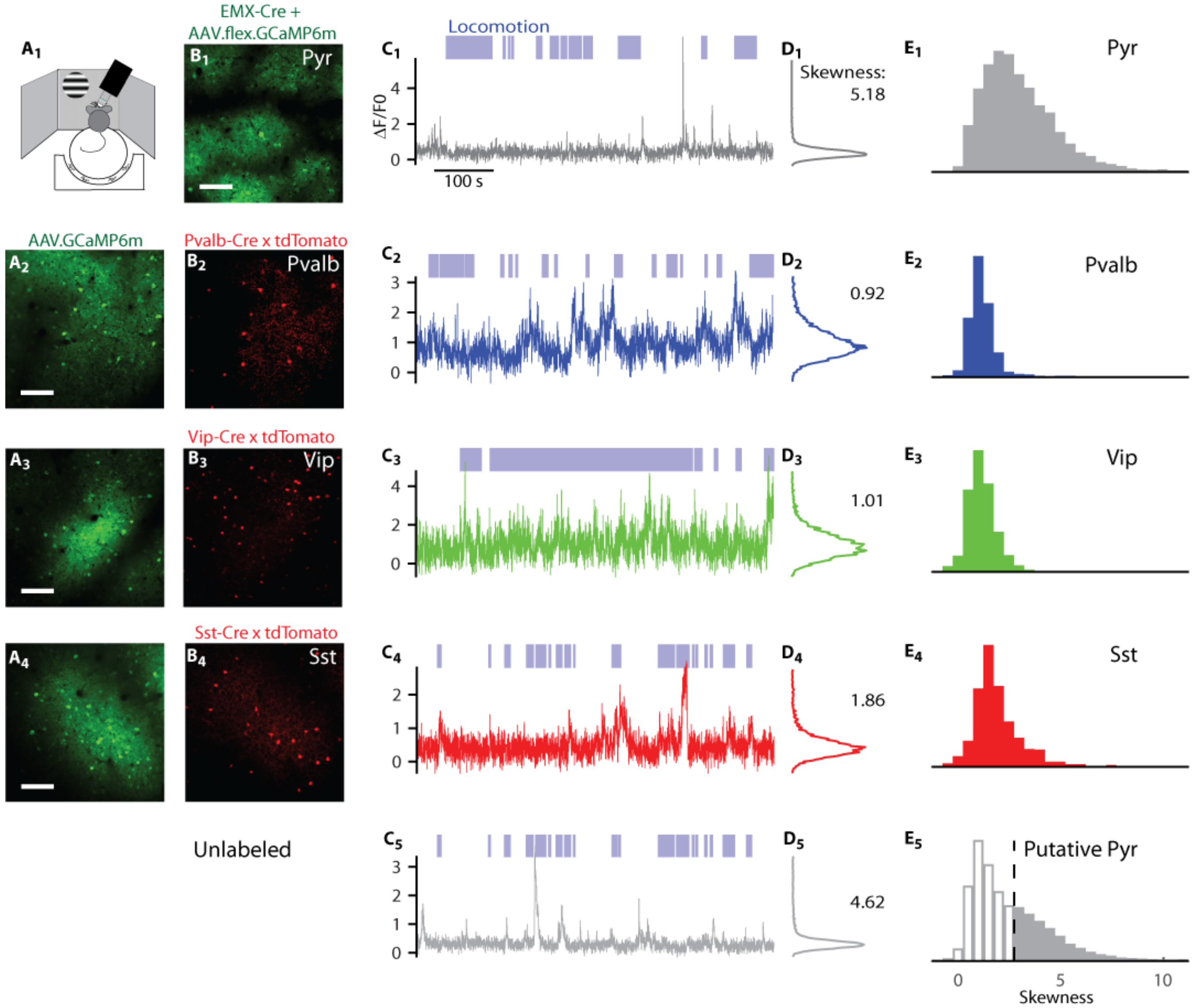

Figure 1. Recordings and classification of Pyr, *Pvalb*, *Vip*, and *Sst* cells in awake mouse V1.
A$_1$) Experimental setup showing the air-suspended ball surrounded by the three screens for stimulus presentation.
A$_2$-A$_4$) Green fluorescence from three mice expressing GCaMP6 following virus injection. Scale bars: 100 µm.
B$_1$) Green fluorescence from an *Emx1-Cre* mouse expressing GCaMP6m via virus injections.
B$_2$-B$_4$) Red fluorescence from the recordings in A$_2$-A$_4$, indicating tdTomato expression in *Pvalb* neurons (B$_2$), in *Vip* neurons (B$_3$), and in *Sst* neurons (B$_4$).
C) Normalized fluorescent traces from five representative neurons of the Pyr, *Pvalb*, *Vip*, *Sst*, and unlabeled types. The unlabeled neuron in C$_5$ was recorded simultaneously with the *Sst* example in C$_4$. Blue shading above axes represents periods of locomotion (> 1 cm/s).
D) Histogram of fluorescence values for those five example neurons. The number indicates the skewness of the distribution.
E) Distribution of skewness values over all Pyr, *Pvalb*, *Vip*, *Sst*, and unlabeled neurons. Unlabeled cells above a skewness threshold of 2.7 (dashed vertical line) are classified as putative Pyr (E$_5$).

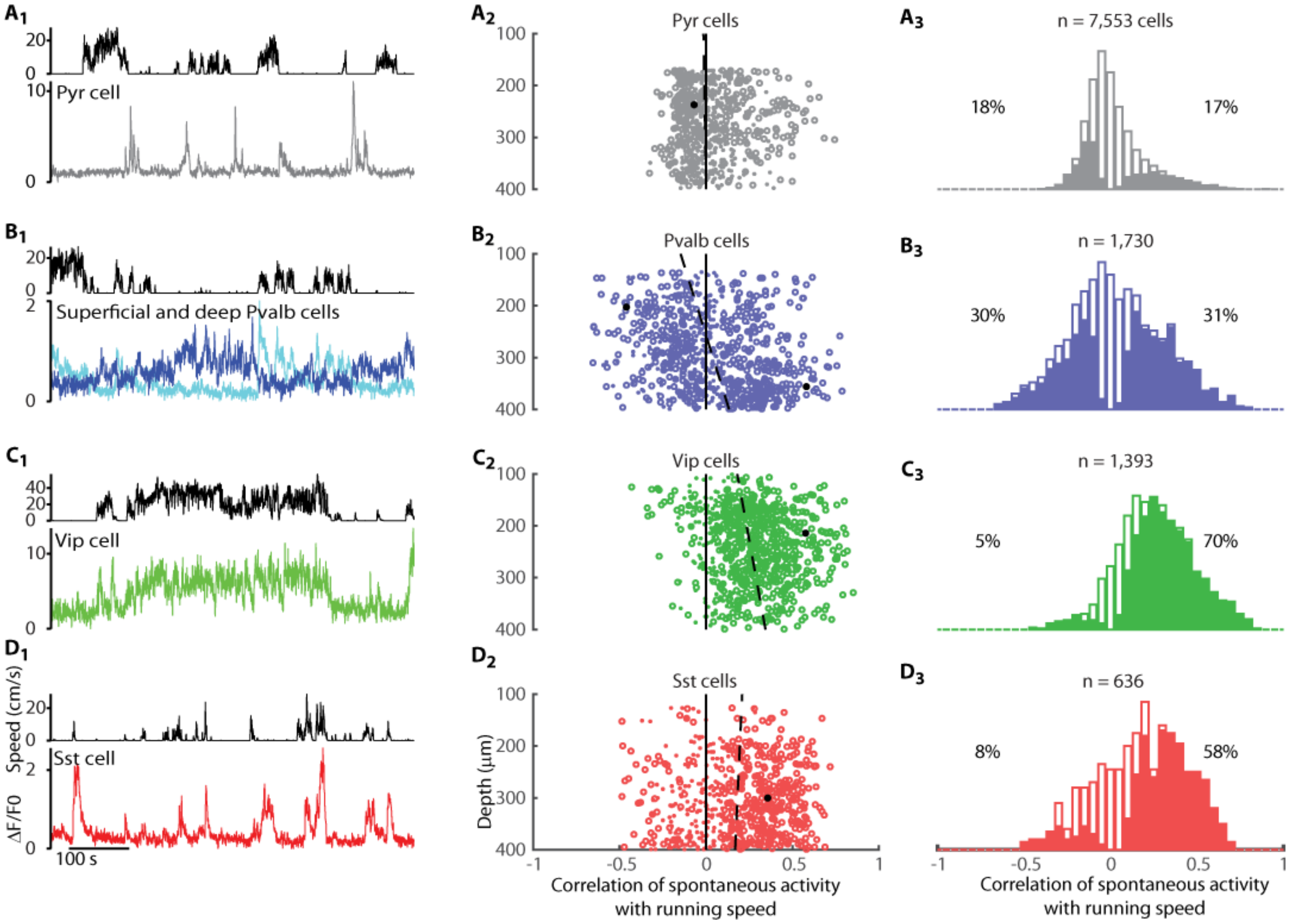

Figure 2. Correlates of running with neural activity in absence of visual stimuli (gray screen).

A_1) Fluorescence of representative L2/3 pyramidal neuron (bottom) and simultaneous running speed trace (top).

A_2) Correlation coefficient of recorded pyramidal cells with running speed, plotted vs. cell depth. Circles represent cells with significant correlations at $p<0.05$ (shuffle test); dots represent cells with insignificant correlations. For clearer visualization we plotted only a random subsample of 800 of genetically-identified Pyr cells. Dashed line represents fitted dependence of correlation vs. depth. Arrow indicates example cell shown in A_1.

A_3) Histogram of correlation coefficients of all pyramidal cells. Solid bars indicate significant correlations at $p<0.05$ (shuffle test). Values left and right of the histogram represent percentage of cells with a negative or positive correlation respectively.

B) Similar analysis for *Pvalb* neurons. The two traces in B_1 (top) show fluorescence traces of representative *Pvalb* cells of upper and lower L2/3 (blue and cyan respectively).

C) Similar analysis for *Vip* cells.

D) Similar analysis for *Sst* cells.

Effects of locomotion on baseline activity

As a first test of the disinhibitory model, we asked how locomotion affected baseline activity, measured during long periods of gray screen presentation (Figure 2). In *Sst* cells, these measurements revealed ways in which previous apparently conflicting reports could be reconciled. In other interneurons, these measurements showed strong effects on baseline activity, an unexpected dependence on cortical depth, and phenomena that run opposite to the predictions of the disinhibitory model.

Consistent with previous results (Niell and Stryker, 2010; Polack et al., 2013; Saleem et al., 2013), the effects of locomotion on the baseline activity of identified Pyr cells were weak and diverse (Figure 2A). The sparse baseline activity (measured in response to a gray screen) of a typical *Pyr* cell changed only weakly with running speed (Figure 2A1). Across Pyr cells, the average correlation of baseline activity and running speed was close to zero ($\rho_{gray} = 0.03 \pm 0.01$ S.E., n = 7,553 identified Pyr cells, Figure 2A3). Nevertheless, 35% of Pyr cells showed a significant positive or negative correlation with speed ($p < 0.05$ shuffle test), significantly more than the 5% that would be expected by chance ($p < 10^{-16}$, Fisher's combined probability test). Similar results were seen in the putative Pyr neurons identified by their sparse firing (in unconditional labeling): correlations were small on average ($\rho_{gray} = -0.01 \pm 0.01$; SE, n = 5,666, Supplementary Figure 4A,B), but 49% of the cells were significantly positively or negatively correlated with speed ($p < 0.05$), significantly more than expected by chance ($p < 10^{-16}$, Fisher's combined probability test). In agreement with previous results obtained in darkness (Fu et al., 2014), when the monitors were switched off, the mean correlation of Pyr activity with speed was again close to zero ($\rho_{dark} = 0.00 \pm 0.01$, SE, Supplementary Figure 5A_1,B_1). However, cells tended to show similar effects of locomotion when monitors were turned on vs. off (Pearson correlation between ρ_{gray} and ρ_{dark} was 0.33; $p < 10^{-71}$).

The effects of locomotion on the baseline activity of *Pvalb* interneurons were stronger and more varied, and also depended on cortical depth (Figure 2B). For example, in two typical *Pvalb* cells imaged simultaneously (Figure 2B_1), activity decreased with running speed in the more superficial cell (darker trace $= -0.47$, $p < 0.01$, shuffle test) and increased in the deeper cell (lighter trace, $\rho = 0.58$, $p < 0.01$ shuffle test). These results were typical of the population (n = 1,730), where the correlations were strong and depended significantly on depth (robust regression, $p < 10^{-21}$; Figure 2B_2), with high consistency across experiments ($p < 0.018$, t-test; Supplementary Figure 6). Among *Pvalb* cells in superficial L2/3 (depth <300 µm, n = 843) the average correlation with speed was slightly negative ($\rho_{gray} = -0.05 \pm 0.03$, SE), and this correlation was significantly negative in 36% of the cells and significantly positive only in 24% of the cells ($p<0.05$; shuffle test). The situation was reversed in deeper L2/3, where on average *Pvalb* cells showed a weak positive correlation with speed ($\rho_{gray} = 0.11 \pm 0.04$, SE, n = 831), with correlations significantly positive in 47% of cells and negative in only 18% of cells ($p < 0.05$; shuffle test). When pooling across depth, therefore, a wide variety of effects was seen (Figure 2B_3), echoing the wide and bimodal range of correlations observed in darkness (Fu et al., 2014). When the monitors were switched off, the mean correlation of *Pvalb* activity with speed was slightly negative ($\rho_{dark} = -0.14 \pm 0.01$, SE,Supplementary Figure 5A_2,B_2) with a positive correlation between ρ_{gray} and ρ_{dark} across cells (Pearson correlation 0.49; $p < 10^{-22}$).

Also as expected (Fu et al., 2014), the typical *Vip* cell increased baseline activity markedly with locomotion (Figure 2C_1), and the overall population showed almost exclusively positive correlations with running speed, with a mean correlation of $\rho_{gray} = 0.27 \pm 0.03$ (SE, n = 1,393). The correlation increased significantly with cortical depth (robust regression; $p < 10^{-10}$), an effect that was robust across experiments ($p < 0.01$, t-test; Supplementary Figure 6). We found similar results when we repeated these measurements with the monitors turned off ($\rho = 0.30 \pm 0.06$, SE, Supplementary Figure 5A_3,B_3) with a positive correlation between ρ_{gray} and ρ_{dark} across cells (Pearson correlation 0.54; $p < 10^{-48}$).

Contrary to the predictions of the disinhibitory circuit, however, locomotion generally increased the baseline activity of *Sst* interneurons (Figure 2D). The typical *Sst* cell increased its baseline activity markedly with locomotion (Figure 2D_1) and across the population the correlation of baseline activity with running speed was on average positive ($\rho_{gray} = 0.18 \pm 0.02$, SE, n = 636, Figure 2$D_3$), regardless of depth (robust regression, p = 0.39; Figure 2D_2).

These effects of locomotion on the baseline activity of *Sst* interneurons confirm some previous results (Pakan et al., 2016; Polack et al., 2013) but they appear to disagree with other measurements (Fu et al., 2014). Indeed, these results run opposite to the predictions of the disinhibitory circuit, which would have predicted a negative effect of locomotion on baseline activity of *Sst* cells.

To confirm the validity of these results, we first ensured that they were not due to background fluorescence that might originate from other cell classes. We repeated the measurements in *Sst-IRES-Cre* mice where we expressed the calcium indicator only in *Sst* cells by injecting a Cre-dependent GCaMP6m virus (Supplementary Figure 2B_2). At the time of the injection, the mice were already adult, thus excluding the off-target expression that might occur in cells expressing Cre only transiently during development (Hu et al., 2013). These experiments confirmed our results: the average correlation of baseline activity with running speed was positive (Supplementary Figure 2C-E) in all the locations where GCaMP6 was strongly expressed, be it cell bodies or neuropil. A pixel-by-pixel analysis further confirmed that these results did not depend on the method of ROI detection.

Previous studies have suggested that the modulation of spontaneous activity by locomotion can depend on the phase of the locomotion period, with stronger responses at locomotion onset (Vinck et al., 2015). However, for all cell types we found similar correlations between fluorescence and running speed after removing transition periods between locomotion and stationary periods from the analysis (Supplementary Figure 7).

We next asked if the disagreement on the effects of locomotion on the baseline activity of *Sst* cells could be due to different visual conditions used in different experiments (Pakan et al., 2016). The experiments of Fu et al. (2014) were performed in darkness, whereas in ours (Figure 2D) and those of Polack et al. (2013), the mouse faced a gray screen. We thus measured *Sst* baseline activity in darkness, and found that the effects of locomotion, albeit still diverse, were now overall negative ($\rho = -0.07 \pm 0.02$, SE; across experiments: p = 0.019, t-test, Supplementary Figure 5A_4,B_4,C). The same cell could show different modulation by locomotion depending on screen illumination (Supplementary Figure 5B_4): for example, of the cells showing significant modulation in both conditions, 26% showed $\rho_{dark} < 0$ and $\rho_{gray} > 0$. Not all cells, however, showed this diversity. On average, in fact, *Sst* cells showed a positive correlation between ρ_{gray} and ρ_{dark} (Pearson correlation 0.34; $p < 10^{-8}$).

These measurements, therefore, reconcile the apparent divergence of previous results (Fu et al., 2014; Pakan et al., 2016; Polack et al., 2013): the effect of locomotion on baseline activity of *Sst* cells is overall positive when mice view a gray screen, and mildly negative when mice are in darkness. The observations made with the gray screen, however, contradict the disinhibitory model.

Effects of locomotion on responses to gratings of different sizes

Having witnessed the importance of controlling visual stimulation when studying the effects of locomotion, we next asked how these effects depend on stimulus size. We measured responses to drifting gratings (relative to baseline activity measured before stimulus onset), and we focused on cells that were visually responsive (significant effect of stimulus size, $p < 0.05$, 1-way ANOVA) and whose RFs were centered within 10° of the stimulus center. To discount possible effects of eye movements (whose occurrence might change during locomotion), we only selected trials where the pupil pointed within 10° from the average position. Similar results were found when we removed this restriction (Supplementary Figure 8).

Identified L2/3 Pyr cells were selective for small stimuli (in agreement with Adesnik et al., 2012), and exhibited, on average, a mild effect of locomotion (Figure 3A). A typical Pyr neuron responded substantially more to a 5° stimulus than to a 60° stimulus, regardless of locomotion (Figure 3A_1). Accordingly, the neuron showed clear selectivity for smaller stimuli (Figure 3A_2), similar to that seen in the overall population of identified Pyr cells (n = 1,250; Figure 3A_3) where cells preferring large stimuli were rare (Supplementary Figure 9A). On average, locomotion slightly increased responses to both small stimuli ($p < 10^{-16}$, paired t-test across cells, $p < 0.01$, paired t-test across experiments) and large stimuli ($p < 10^{-10}$ and p = 0.02). However, this effect was diverse among cells, with locomotion significantly increasing or decreasing responses in 17% and 3% of Pyr cells respectively ($p < 0.05$; 2-way ANOVA, main effect of locomotion over 5° and 60° stimuli, Figure 3A_4). Many cells (17%) showed a significant interaction of locomotion and stimulus size ($p < 0.05$; 2-way ANOVA over 5° and 60° stimuli; Figure 3A_4; Supplementary Figure 10A,B_1). In these cells, locomotion changed the relative response to large and small stimuli, as seen previously in deeper layers (Ayaz et al., 2013). Similar results were found in putative pyramidal cells identified by the sparseness of their calcium traces (Supplementary Figure 4C,D).

Pvalb interneurons were similarly selective for stimulus size, but showed a stronger and overwhelmingly positive effect of locomotion (Figure 3B). A typical *Pvalb* interneuron responded strongly to small stimuli and more weakly to larger stimuli, and its responses markedly increased during locomotion (Figure 3B_1-B_2). These effects were highly consistent across *Pvalb* interneurons (n = 277, Figure 3B_3), with locomotion increasing firing rate in practically all cells (Figure 3B$_4$). This effect was seen in responses to both large stimuli ($p < 10^{-11}$, paired t-test across cells, $p < 0.01$, paired t-test across experiments) and small stimuli ($p < 10^{-13}$ and $p = 0.02$), with no significant interaction between stimulus size and locomotion ($p = 0.23$, 2-way ANOVA over 5° and 60° stimuli; Supplementary Figure 10A,B$_2$).

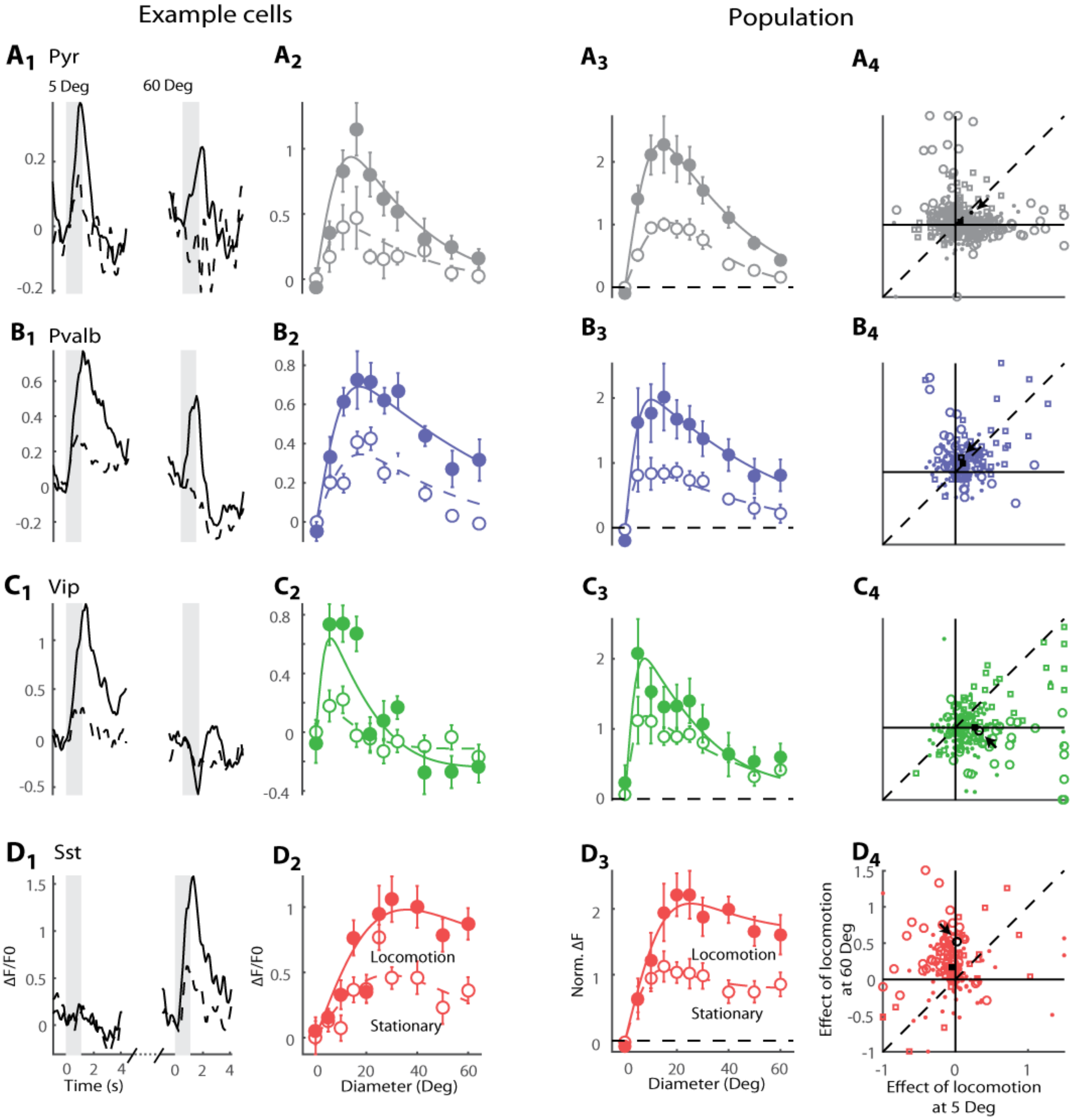

Figure 3. Locomotion increases visual responses for large drifting grating stimuli in *Sst* neurons, for small stimuli in *Vip* neurons, and for all stimulus sizes in *Pvalb* neurons.

A$_1$) Responses of a representative Pyr neuron. Black curves show trial-averaged response in the stationary (dashed line) and locomotion (continuous line) conditions. Panels show responses to stimuli of 5° (left) and 60° (right). Gray shaded regions indicate the 1 s stimulus presentation period.

A$_2$) Size tuning curve for this example cell. Solid line: locomotion; dashed line: stationary.

A$_3$) Size tuning curve, averaged over Pyr cells after normalization. Solid line: locomotion; dashed line: stationary.

A$_4$) Scatter plot showing change with locomotion of normalized responses to large stimuli (60°; y-axis) and to small stimuli (5°; x-axis). Circles represent cells whose responses have a significant interaction between size and locomotion (multi-way ANOVA over 5° and 60° stimuli only), squares represent cells that did not have a significant interaction but did have a significant effect of locomotion; dots represent cells with no significant effect of locomotion. Arrow marks example cell shown in A$_{1\text{-}2}$, square marks mean response.

B-D), similar analysis for *Pvalb*, *Vip*, and *Sst* neurons.

The responses of *Vip* interneurons (n = 239) were selective for stimulus size and increased with locomotion, but in contradiction with the disinhibitory circuit, this increase was generally restricted to responses to small stimuli (Figure 3C). A typical *Vip* interneuron responded most strongly to small stimuli during locomotion (Figure 3C$_1$-C$_2$). Similar results were seen across the population: *Vip* interneurons showed clear size tuning, and locomotion increased their responses to 5° stimuli ($p < 10^{-15}$, paired t-test across cells, p = 0.04, paired t-test across experiments) but not to 60° stimuli (p = 0.82 and p = 0.86, Figure 3C$_3$) with a significant interaction of size and locomotion ($p < 10^{-10}$, 2-way ANOVA over 5° and 60° stimuli; Figure 3C$_4$; Supplementary Figure 10A,B$_3$). Similarly to baseline activity, the effects of locomotion on visual responses grew with cortical depth (robust regression, $p < 10^{-3}$).

Sst interneurons tended to prefer large stimuli, if not the largest ones, and their responses – contradicting the disinhibitory circuit – increased with locomotion (Figure 3D). As observed by Adesnik et al. (2012), a typical *Sst* interneuron responded best to large stimuli, and even more so while the animal was running (Figure 3D$_1$, D$_2$). Similar results were seen across the population (n = 191, Figure 3D$_3$), with overall activity peaking for ~15° stimuli during stationary conditions, and ~25° stimuli during locomotion. *Sst* cells showed a significant interaction between stimulus size and locomotion ($p < 10^{-7}$, 2-way ANOVA over 5° and 60° stimuli), consistently across experiments (p < 0.01, t-test, Supplementary Figure 10A$_4$) and mice (p < 0.01, t-test, Supplementary Figure 10B$_4$). While locomotion did not significantly affect the responses to small stimuli (p = 0.06, paired t-test across cells, p = 0.65, paired t-test across experiments, Figure 3D$_4$), it strongly increased the responses to large stimuli ($p < 10^{-9}$ and p = 0.02), thus contradicting the disinhibitory circuit.

Some *Sst* cells, however, did show size tuning (Figure 3D, Supplementary Figure 9C,D). This observation is consistent with observations in anesthetized mice (Pecka et al., 2014), but differs from those of Adesnik et al. (2012) in awake mice. We reasoned that it may reflect high sensitivity of these cells to stimulus centering, and to investigate this possibility, we studied how size tuning varies with the distance between receptive field center and stimulus center. Regardless of cell class, when the stimulus was distant, the cells preferred the largest stimulus (Supplementary Figure 11). Size tuning emerged when stimulus distance decreased. For *Sst* cells, in particular, size tuning appeared when stimuli were < 20° away from the receptive field center (Supplementary Figure 11D).

For all cell types, we saw similar interactions of size tuning and locomotion when cells were recorded in the binocular and monocular regions of visual cortex (Supplementary Figure 12). Furthermore, the results did not depend on whether or not we deconvolved the calcium traces to estimate spike rates (Supplementary Figure 13).

In summary, our results for large stimuli contradict the disinhibitory model, which predicts that locomotion should increase firing in *Vip* cells and decrease it in *Sst* cells. Instead, we saw something closer to the opposite behavior: in the presence of large stimuli, locomotion does little to *Vip* cells and boosts the activity of *Sst* cells.

Correlations of interneuron activity with putative pyramidal activity

To further test the disinhibitory model, we examined the patterns of correlation between interneuron types and Pyr cells (Figure 4). To measure correlation, we relied on our ability to record simultaneously the activity of identified interneurons and of putative Pyr neurons (unlabeled sparse-firing cells, with skewness > 2.7, Figure 1). This analysis revealed dramatic differences between interneuron classes.

Consistent with the view that *Pvalb* interneurons track the activity of Pyr cells (Cruikshank et al., 2007; Isaacson and Scanziani, 2011; Okun and Lampl, 2008; Ozeki et al., 2009; Renart et al., 2010), we found strong positive correlations between *Pvalb* and putative Pyr populations (Figure 4A). In each experiment, we summed the activity of the *Pvalb* interneurons and compared it to the summed activity of the simultaneously recorded putative Pyr cells. Only cells whose receptive field lay within 10° from the stimulus center were considered. In a typical experiment the spontaneous correlations (ρ_0) measured in the presence of a gray screen were strongly positive, whether the mouse was stationary ($\rho_0 = 0.70$, Figure 4A$_1$) or running ($\rho_0 = 0.60$, Figure 4A$_2$). Similar results were seen across experiments ($\rho_0 = 0.60 \pm 0.04$, SE, during rest and 0.49 ± 0.03 during locomotion). Examining population signal correlations – i.e. the relationship of the mean summed responses of the *Pvalb* and Pyr cells across stimuli – also showed a strong, positive correlation in both stationary ($\rho_s = 0.95$ for the example in Figure 4A$_3$, 0.86 ± 0.04 across all experiments) and locomotion conditions ($\rho_s = 0.97$ for the example in Figure 4A$_4$, 0.91 ± 0.02 SE across all experiments). Similar results were found for population noise correlations (i.e. the relationship between trial-to-trial variability in summed activity of the *Pvalb* and Pyr populations; Supplementary Figure 14). These results were not biased by the exclusion of low-skewed Pyr cells, which did not pass our conservative skewness threshold: when we lowered the threshold (therefore including more Pyr cells but also some unlabeled inhibitory neurons), the correlations of putative Pyr cells with *Pvalb* cells remained high (Supplementary Figure 15A,B,C).

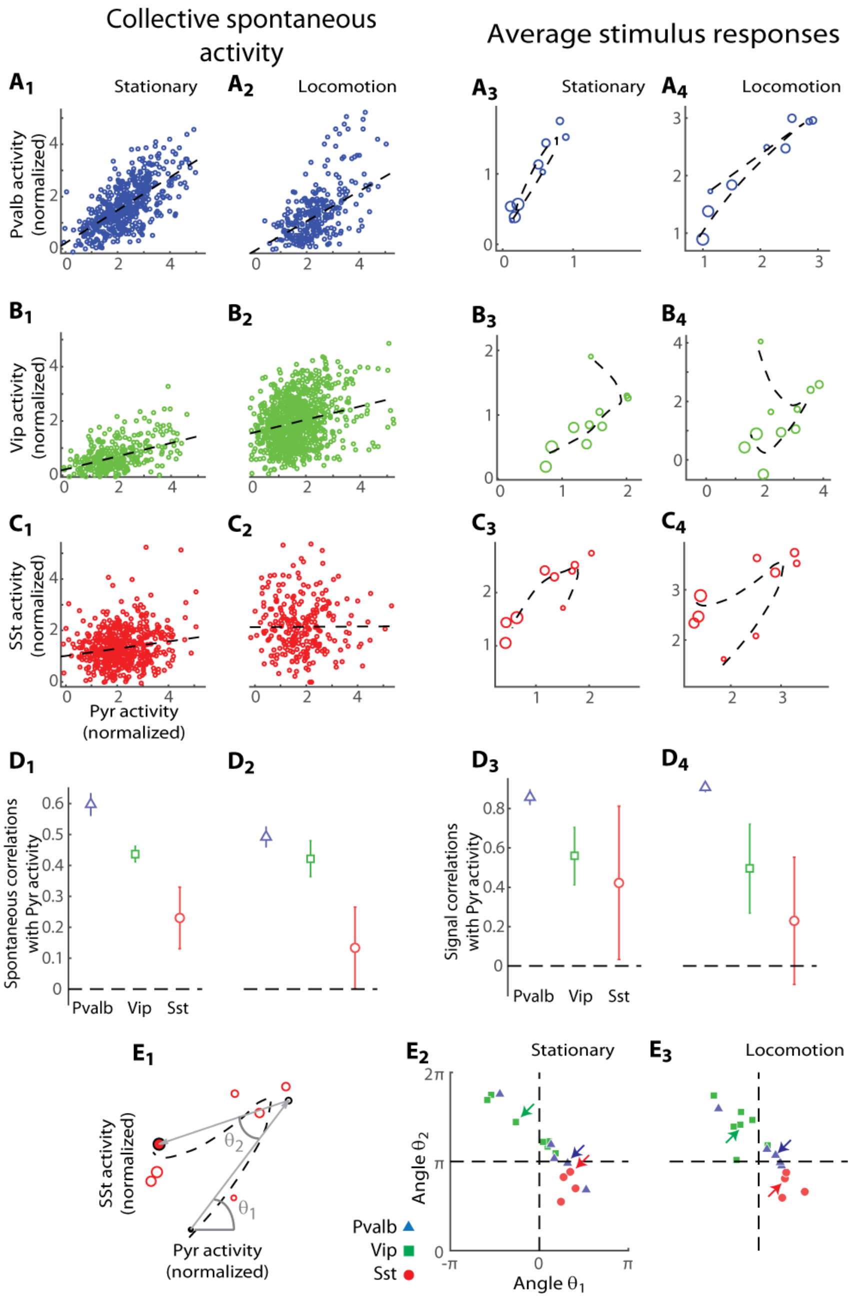

Figure 4. Pvalb cells are highly correlated with simultaneously recorded putative pyramidal cells, but *Vip* and *Sst* cells show nonlinear correlations.

A$_1$ and A$_2$) Summed activity of *Pvalb* population vs. Pyr population in the gray screen condition during stationary periods (A$_1$) and during locomotion (A$_2$). Each circle represents the simultaneous normalized value of the excitatory and the inhibitory populations at one time point. Dashed line indicates linear regression estimate of signal correlation.

A$_3$ and A$_4$) Average stimulus response of *Pvalb* population vs. average response of Pyr population in a typical experiment during stationary periods (A$_3$) and during locomotion (A$_4$). Each point represents a response to a stimulus, with larger circles representing larger stimuli. Dashed line represents non-linear interpolation of the Pvalb and Pyr size tuning obtained from their size-tuning curves.

B,C) Same as A, for *Vip* and *Sst* interneurons. Note the nonlinear signal correlations.

D$_1$ and D$_2$) Summary plots of spontaneous correlations during stationary periods (D$_1$) and during locomotion (D$_2$).

D$_3$ and D$_4$) Linear signal correlation (Pearson correlation coefficient) between the Pyr population and the three classes of interneurons averaged across all experiments during stationary periods (D$_3$) and during locomotion (D$_4$).

E$_1$) Same plot represented in C$_4$ illustrating the characteristic angles used to illustrate the nonlinear relationship between each interneuron class and Pyr mean visual responses. Circles with black contour indicate the minimum (5°, black filled) and maximum sizes (60°, red filled), and the cell's preferred size (white filled). θ_1 is the angle relative to horizontal of the line joining the response to 5° stimuli and the preferred stimulus for Pyr neurons. θ_2 is the angle between the latter line and the line joining the response preferred size to the response to 60°.

E$_2$ and E$_3$) Angle θ_2 vs. θ_1 as defined in E$_1$ for each experiment during stationary periods (E$_2$) and during locomotion (E$_3$). In E$_3$ arrows indicate examples in (A$_3$-C$_3$); in E$_4$ arrows indicate examples in (A$_4$-C$_4$).

Vip cells showed a markedly different behavior (Figure 4B). Their correlations with putative Pyr populations differed from those of *Pvalb* cells in two respects. First, while spontaneous and noise correlations tended to be positive ($\rho_0 = 0.43 \pm 0.03$, SE, during stationarity and 0.42 ± 0.06 during locomotion; $\rho_n = 0.33 \pm 0.07$, SE, during stationarity and 0.37 ± 0.02 during locomotion), they were weaker than those of *Pvalb* cells, at least during stationarity (Figure $4D_1$ and Supplementary Figure 14). Second, the relationship between the population mean responses of *Vip* and putative Pyr cells was clearly nonlinear (Figure $4B_{3,4}$). This nonlinearity reflects the different size tuning of *Vip* and Pyr cells, with *Vip* responses peaking at smaller stimulus sizes than Pyr responses (compare Figure $3C_3$ and A_3). It further suggests that, unlike for *Pvalb* cells, the sensory tuning of *Vip* cells during locomotion cannot be explained by a simple tracking of excitatory activity.

Sst interneurons showed yet a different sort of behavior, which depended strongly on locomotion (Figure 4C). The correlation of the *Sst* and Pyr populations was positive in stationary conditions ($\rho_0 = 0.25$ for the example in Figure $4C_1$, 0.23 ± 0.10 SE across experiments) but weak during locomotion ($\rho_0 = -0.01$ for the example in Figure $4C_2$, 0.13 ± 0.13 SE across experiments). Noise correlations were also positive (Supplementary Figure 14; $\rho_n = 0.39 \pm 0.11$ SE, stationary, $\rho_n = 0.30 \pm 0.04$ SE, locomotion). Signal correlations showed a non-linear character, but this differed to that of *Vip* cells (Figure $4C_{3,4}$).

To quantify the form of nonlinear relationship between each interneuron class's mean responses and those of putative Pyr cells, we parameterized their relationship using two angles θ_1 and θ_2 Figure $4E_1$). These angles were similar in stationary and locomotion conditions, and were consistent across datasets (Figure 4E); they were also robust to changes in the skewness threshold (Supplementary Figure 15D,E). The different form of nonlinear relationship between *Sst* and *Vip* cells result in a significant difference in θ_1 (stationary: $p = 0.026$, locomotion: $p < 0.01$, Watson's U^2 permutation test, N=1,000 permutations). Similarly, there was a significant difference of θ_2 between *Sst* and *Vip* cells (stationary: $p = 0.039$, locomotion: $p = 0.036$, Watsons U^2 permutation test, N=1,000 permutations).

Overall, thus, these results paint a more complex picture than would be expected from the disinhibitory model. The model predicts that Pyr activity should increase with *Vip* activity and decrease with *Sst* activity. This prediction is consistent with correlations measured during spontaneous activity in darkness, but not with the covariation of stimulus responses seen in the presence of medium to large stimuli, where both *Vip* and *Sst* cells decrease their activity with stimulus size. Visual stimulation, therefore, creates situations where the disinhibitory circuit cannot account for size tuning.

Recurrent network model

To explore how size tuning of the different neuron classes in layer 2/3 of area V1 could arise from cortical circuitry, we asked if their size tuning curves could be reproduced by a recurrent network model. We fit the model with a novel approach: we estimated the synaptic input parameters for each class of neurons to be those optimally predicting that class' average population sensory responses, with the average population activity of all other cell classes clamped to their measured values.

We modeled the activity of each cell class by a "neural field": a number that varied across the retinotopic cortical surface, representing the mean firing rate of all cells of that class at that location. The sensory response of cell class α was modeled as a function $f_\alpha^{(v)}(s, \mathbf{r})$ of stimulus size s and retinotopic position $\mathbf{r}$ (relative to stimulus center); dependence on $\mathbf{r}$ was assumed to be circularly symmetric. The superscript v indicates the locomotion condition ($v = 0$: rest, $v = 1$: locomotion). Synaptic connections were modeled as functions of retinotopic location: the connection strength to a cell of type α_1 at location $\boldsymbol{r}_1$ from a cell of type α_2 at location $\boldsymbol{r}_2$ decays as a two-dimensional Gaussian function of the distances between receptive field centers $G_{\alpha_1\alpha_2}(\|\boldsymbol{r}_1 - \boldsymbol{r}_2\|)$. The strength and spatial spread of these connections were fit by exhaustive search, minimizing the squared error between the predicted and actual rates summed over all stimulus sizes and offsets and locomotion conditions. Each cell integrates its inputs using a threshold-linear activation function. For each inhibitory connection, we tested both divisive and subtractive inhibition, with the choice made automatically for each inhibitory synapse type to minimize errors. The firing of dLGN inputs $h(s, \boldsymbol{r})$ was modeled using a ratio of Gaussians (Ayaz et al., 2013), with parameters (for both rest and locomotion conditions) fit to the data of Erisken et al. (2014).

The model accurately predicted the size-tuning curves of each cell class, for both centered and off-centered stimuli (Figure 5), but this success was predicated on certain conditions. Specifically, to predict the strong response of *Sst* cells to large, off-center stimuli (Figure 5D$_3$ and D$_5$), the model needed to include external excitatory input to these cells (e.g. from thalamus or other from cortical layers), because large stimuli elicited little response in Pyr cells (Figure 5A$_2$-A$_5$). A good fit was only obtained if this feedforward input to *Sst* cells had broad size tuning, as would be seen in thalamic neurons or perhaps in excitatory neurons of other cortical layers (Figure 5D$_1$). Moreover, to obtain the observed similarity in tuning of *Pvalb* and Pyr cells (Figure 5A,B), the model required these cell classes to have similar inhibition from *Sst* cells. Furthermore the model required Pyr neurons to lack *Vip* input. Finally, our parameter search only gave good results with divisive inhibition from *Sst* to *Vip* cells (Figure 5C$_1$): subtractive inhibition could not produce the observed sharp size-tuning favoring small stimuli (Figure 5C$_2$-C$_5$).

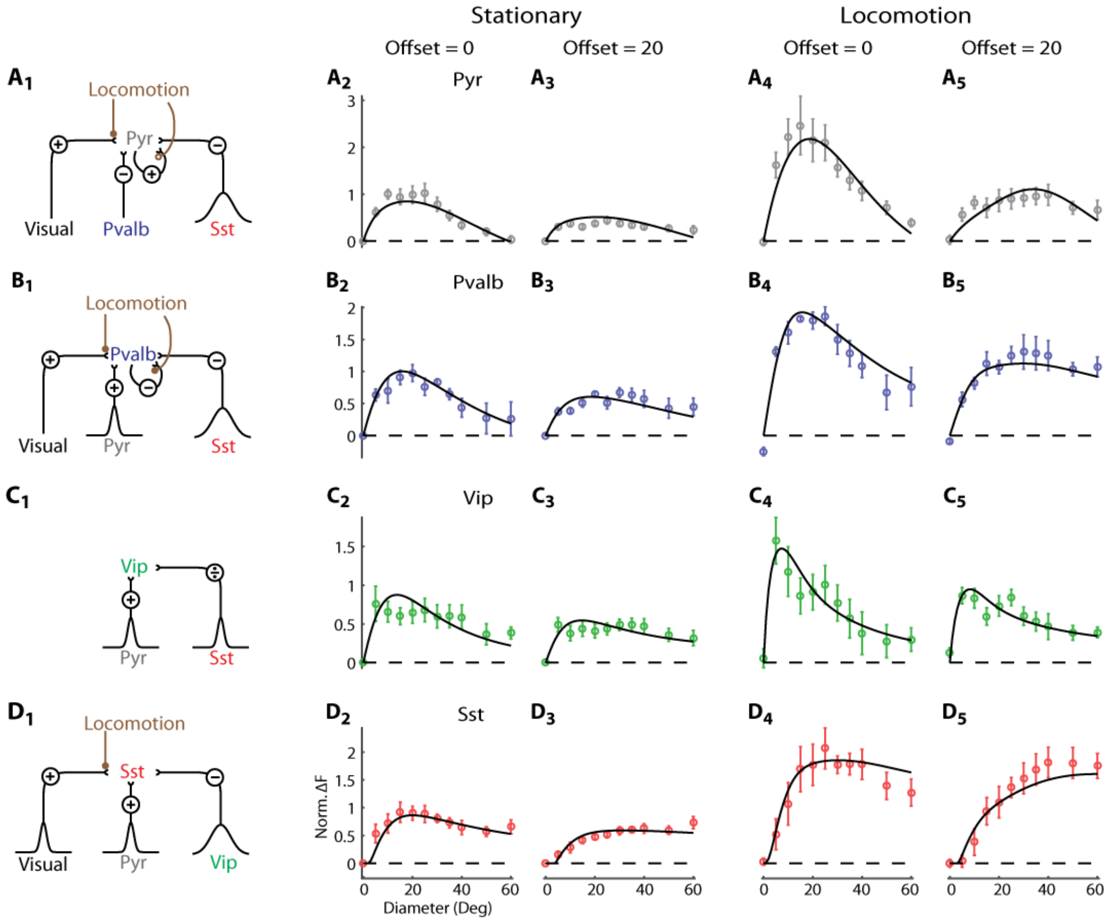

Figure 5. Computational model of size tuning in L2/3 of mouse V1. We searched the parameter space of neural field theories to find a model that can accurately reproduce our experimental results.

A$_1$) Input synapses received by Pyr cells, and their modulation by locomotion. Pyr cells receive feedforward visual excitatory inputs (e.g. from thalamus or other cortical layers) and subtractive inhibition from *Pvalb* and *Sst* cells. Locomotion increases the synaptic weights from external visual inputs to Pyr cells (filled circles) while decreasing recurrent Pyr-to-Pyr connections (empty circles). Pyr cells integrate from a broad pool of *Sst* cells (large Gaussian curve).

A$_{2\text{-}5}$) Model fit (black curves) of the size tuning data during stationary (A$_{2,3}$) and locomotion (A$_{4,5}$) periods, visualized for centered cells (offset = 0, A$_{2,4}$) and off-centered cells (offset = 20, A$_{3,5}$).

B$_1$) Input received by *Pvalb* cells. Locomotion boosts feedforward synaptic weights to Pvalb neurons while decreasing recurrent *Pvalb*-to-*Pvalb* synaptic weights.

B$_{2\text{-}5}$) Same as (A$_{2\text{-}5}$), for *Pvalb* cells.

C$_1$) Input received by *Vip* cells. *Vip* cells receive divisive inhibition from *Sst* cells; no modulation of these synapses by locomotion is required to obtain a good fit.

C$_{2\text{-}5}$) Same as (A$_{2\text{-}5}$), for *Vip* cells.

D$_1$) Input received by *Sst* cells. In addition to inputs from Pyr and *Vip* cells, *Sst* cells receive a feedforward input that we propose to originate from thalamus or other cortical layers. As with Pyr and *Pvalb* cells, the strengths of these synapses are boosted by locomotion.

D$_{2\text{-}5}$) Same as (A$_{2\text{-}5}$), for *Sst* cells.

We next asked what modifications of the model parameters could explain the effects of locomotion on the sensory responses of each cell type. Modeling the locomotor modulation of size tuning required that locomotion change synaptic strengths: if we kept the input synapses of Pyr cells at the strengths fit to stationary data, while clamping the activity of other cell types to their values during rest, we obtained a poor prediction of Pyr cells' size tuning during locomotion (not shown). To capture the effects of locomotion on each cell type, we therefore searched for all possible ways that locomotion could modulate the thalamocortical and recurrent synaptic weights of each class (Supplementary Figure 16).

For Pyr cells, we could prove analytically that the fit quality depended only on the strength of their "effective connections" (see Experimental Procedures). These effective connections depended on several network parameters, including the strengths of external, recurrent excitatory, *Pvalb*, and *Sst* inputs onto Pyr neurons, and take into account the amplification caused by recurrent excitation and *Pvalb* inhibition (Douglas et al., 1995). Thus, while the model fit identified unique values of the effective connections, these values could in turn be achieved through multiple possible strengths of the synaptic parameters (Prinz et al., 2004).

Capturing the modulation of Pyr tuning by locomotion required an increase in effective connection from external inputs to Pyr cells, and a decrease in effective connection of *Sst to* Pyr cells. Together, this produced the observed strong increase in responses of centered cells to medium-sized stimuli, together with a milder increase in response to larger stimuli (Figure 5A_2 vs. A_4). However, this change in effective connection strength did not require a weakening of the physical *Sst*->Pyr connections: the same change in effective connection could also be achieved by an increase in external excitation together with a decrease in recurrent excitation (Supplementary Figure 16A_2). Intriguingly, these two effects are precisely those observed *in vitro* during cholinergic modulation of cortical synapses (Gil et al., 1997).

Consistent with the close correlations of *Pvalb* and Pyr cells, modeling the observed effects of locomotion on *Pvalb* size tuning required similar modulations to those required for Pyr cells: a decrease in effective inhibition from *Sst* cells. Again, however, this did not necessarily require a weakening of *Sst->Pvalb* synapses, as the same effect could be obtained via strengthening of *Pvalb->Pvalb* connections and of the external input. Producing the observed effects of locomotion on *Vip* cell tuning required no further changes in effective connection: locomotion only increased the responses of centered *Vip* cells to small stimuli (Figure 5C_2 vs. C_4), and this increase could be readily provided by increased activity of local Pyr cells. Finally, correct modulation of *Sst* firing required boosting of the external excitatory inputs responsible for their responses to large stimuli (Figure 5D), but did not require any change to their inhibitory inputs.

In summary, we were able to capture the effects of locomotion on all cell types through a reweighting of feedforward and recurrent connections: an increase in external excitatory input to all cell types, a decrease in recurrent excitation between Pyr cells, and an increase in recurrent inhibition between *Pvalb* cells.

Discussion

We have shown that locomotion does not simply increase or decrease the activity of a particular cell class: its effects depend on the precise sensory conditions, and even on cortical depth. Its effect on sensory responses was an increase in all cell classes, but the increase varied with cell type and stimulus, being largest in *Sst* responses to large stimuli and in *Vip* responses to small stimuli. The effects of locomotion on baseline neural activity (as assessed by gray screen viewing) were more complex: locomotion increased activity in most *Sst* and *Vip* neurons, and had diverse effects on *Pvalb* and Pyr cells, suppressing most *Pvalb* cells in superficial L2/3 and increasing activity in deeper *Pvalb* cells.

Our results further indicate that apparent discrepancies in the literature may in fact reflect the subtly different methods used in different experiments. Specifically, two studies on the spontaneous activity of *Sst* cells reported opposite effects of locomotion: increased activity (Polack et al., 2013) and decreased activity (Fu et al., 2014). When we replicated their experimental conditions (gray screen for the first study, complete darkness for the second), we replicated both observations in a common set of neurons. These results reconcile the apparent contradiction between these studies (see also Pakan et al., 2016).

The results also reinforce the importance of correcting for out-of-focus fluorescence in 2-photon calcium imaging. Indeed, without correcting for this confound, one would observe an artefactual negative correlation of fluorescence with running speed, particularly in image regions with weak GCaMP expression (Supplementary Figure 2).

We had set out to test a specific, simple "disinhibitory" hypothesis: that during locomotion, increased *Vip* activity would inhibit *Sst* cells, and thus increase the activity of pyramidal cells (Fu et al., 2014). Some of our results were consistent with the hypothesis. For instance, we confirmed that locomotion increases *Vip* responses. Other results, instead, could not be explained by disinhibition alone. The hypothesis predicts that increased *Vip* activity should suppress *Sst* firing, and yet we found that *Sst* firing increased during locomotion, as long as visual inputs were present (even just a blank gray screen).

Another area where previous models need to be updated concerns the mechanisms of size tuning. Adesnik et al (2012) proposed that an important contribution to size tuning comes from the activity of *Sst* cells. They reported these cells as firing mostly during locomotion and showing negligible size tuning. Our experiments confirmed that *Sst* cells fire more during locomotion, but replicated the finding about size tuning only when stimuli were poorly centered on the receptive field. When they were centered, instead, *Sst* neurons displayed marked preferences for intermediate stimulus sizes. In light of this finding, the role of *Sst* neurons in size tuning needs to be re-evaluated. Size tuning is likely to reflect a more complex interaction between the different cell classes.

To understand what form this interaction might take, we built a network model and searched for the parameters that caused it to produce accurate fits to our measurements. Previous studies that proposed a network circuit to explain surround suppression in V1 had not included different cell types (Ozeki et al., 2009; Rubin et al., 2015) or had not validated the predictions with quantitative fits (Litwin-Kumar et al., 2016). Our analysis cast further doubt on the possibility that a disinhibitory mechanism is sufficient to explain changes in size tuning. Indeed, had a disinhibitory mechanism been sufficient, the measured locomotion modulation of interneuron tuning – with no changes in synaptic connection strengths – would have been sufficient to cause the observed modulation in pyramidal cell tuning. Instead, the model required locomotion to change effective connections between neuronal classes, i.e. the effect of one class on another after taking into account amplification through recurrent excitation and inhibition.

These specific changes in effective connections could in turn be instantiated through multiple possible modulations of physical synaptic strengths. The activity produced by a neural circuit is not always sufficient to constrain its underlying connectivity (Prinz et al., 2004); for the current model, we could mathematically prove that multiple underlying connectivity patterns yield identical sensory responses. Nevertheless, the parameter space consistent with our experimental observations favored one particularly attractive possibility, where locomotion would increase external excitatory input to all cell types, decrease recurrent excitation between Pyr cells, and increase recurrent inhibition between *Pvalb* cells. The first two of these are known effects of cholinergic modulation on cortical circuits (Gil et al., 1997).

The model required specific conditions to produce an accurate quantitative fit to our measurements, which we therefore consider experimental predictions. First, the model required that *Sst* cells receive a feedforward sensory input: i.e. an excitatory input conveying visual input other than from local Pyr cells. This was required because off-center *Sst* cells fire strongly to large stimuli, but neither centered nor off-center Pyr cells fire strongly enough to large stimuli to provide sufficient input. Whether *Sst* cells receive direct thalamic inputs is controversial (Cruikshank et al., 2010; Lee et al., 2013; Tan et al., 2008). Even without direct thalamic afferents however, such an input could be conveyed to superficial *Sst* cells via other cortical layers. Interestingly, the optimal model parameters required that the external inputs *Sst* cells receive be spatially diffuse, as would be expected if this input had experienced an additional round of divergence through Pyr cells of other layers. A second condition required by our model was that *Vip* cells receive divisive, rather than subtractive inhibition from *Sst* cells. We are not aware of any data on this question, and therefore consider it to be a model prediction.

In contrast to the complex relationship of *Vip* and *Sst* neurons to summed Pyr activity, the activity of *Pvalb* interneurons was in all cases tightly locked to the local excitatory population. These data are therefore consistent with a primary role for *Pvalb* cells of stabilize the activity of the local circuit via tracking summed excitatory firing (Cruikshank et al., 2007; Ozeki et al., 2009; Renart et al., 2010), rather than directly sculpting receptive field shapes. In our model, we were able to reproduce the similar tuning of *Pvalb* and Pyr cells only if they received similar inputs from other inhibitory classes. Consistent with *in vitro* observations (Pfeffer et al., 2013), the model required that both Pyr and *Pvalb* cells received inhibition from *Sst*, but not *Vip* cells.

What computational benefit might the visual cortex derive from these interactions of locomotion, stimulus size, and cell type? One can consider two hypotheses for why locomotion should modulate visual cortical activity. In the first hypothesis, boosting of activity in visual cortex in running animals serves as a form of modality-specific attention: because vision is particularly important during navigation, locomotion increases V1 activity while suppressing the activity of competing sensory systems. This possibility is supported by the fact that locomotion appears to decrease activity in other cortical regions (Schneider et al., 2014; Zhou et al., 2014), but it cannot explain why locomotion has such diverse effects on V1 pyramidal cells, boosting some while suppressing others, and increasing responses only to certain sensory stimuli (Keller et al., 2012; Saleem et al., 2013). The second hypothesis holds that locomotor modulation of V1 is not simply a matter of gain control, but a complex and diverse neuronal integration of visual and locomotor input. Indeed, work in virtual reality environments has suggested that, rather than simply making visual responses larger or smaller, locomotor modulation may underlie more complex computations, such as integrating the animal's own movement with that of the world (Keller et al., 2012; Saleem et al., 2013). Diverse interaction of sensory stimuli and locomotion in different interneuron classes might form a key mechanism behind this integration.

Experimental procedures

All experimental procedures were conducted according to the UK Animals Scientific Procedures Act (1986). Experiments were performed at University College London under personal and project licenses released by the Home Office following appropriate ethics review.

Mice

Experiments in which an interneuron class was labelled with tdTomato and recorded together with other cells were conducted in double-transgenic mice obtained by crossing Gt(ROSA)26Sor<tm14(CAG-tdTomato)Hze> reporters (Madisen et al., 2010) with appropriate drivers: *Pvalb*<tm1(cre)Arbr> (Hippenmeyer et al., 2005) (2 males, 3 females), *Vip*<tm1(cre)Zjh> (Taniguchi et al., 2011) (3 males, 2 females), *Sst*<tm2.1(cre)Zjh> (Taniguchi et al., 2011) (2 males, 1 female), and *Gad*2<tm2(cre)Zjh> (Taniguchi et al., 2011) (2 females). Experiments in which indicator was expressed uniquely in one neuron class were conducted in single transgenic mice: *Emx1*-IRES(cre) (n=1), *Pvalb*<tm1(cre)Zjh> (n=1), *Vip*<tm1(cre)Zjh> (n=1), *Sst*<tm2.1(cre)Zjh> (n=3), referred to as *Vip-Cre* and *Sst*-Cre respectively. Experiments in which pyramidal cells were labelled exclusively were conducted in *CamK2a*-tTA; Ai94(TITL-GCaMP6s); *Rasgrf2*-2A-dCre triple transgenic mice (n=3) (Madisen et al., 2015). Mice were used for experiments at adult postnatal ages (P54-110).

Animal preparation and virus injection

The surgeries were performed in adult mice (P35–P76) in a stereotaxic frame and under isoflurane anesthesia (5% for induction, 0.5-1% during the surgery). During the surgery we implanted a head-plate for later head-fixation, made a craniotomy with a cranial window implant for optical access, and, on relevant experiments, performed virus injections, all during the same surgical procedure. In experiments where an interneuron class was recorded together with other cells, mice were injected with an unconditional GCaMP6m virus, AAV1.Syn.GCaMP6m.WPRE.SV40 (referred to as non-flex.GCaMP6m). In experiments where an interneuron class was labelled by unique expression, mice were injected with AAV1.Syn.Flex.GCaMP6m.WPRE.SV40 (flex.GCaMP6m) and AAV2/1.CAG.FLEX.tdTomato.WPRE.bGH (flex.tdTomato); all viruses were acquired from University of Pennsylvania Viral Vector Core. In both cases, viruses were injected with a beveled micropipette using a Nanoject II injector (Drummond Scientific Company, Broomall, PA 1) attached to a stereotaxic micromanipulator. One to three boli of 100-200 nl virus (2.23×10^{12} GC/ml for non-flex.GCaMP6m; 2.71×10^{12} for flex.GCaMP6m) were slowly (23 nl/min) injected unilaterally into monocular V1 (Wagor et al., 1980), 2.1-3.3 mm laterally and 3.5-4.0mm posteriorly from Bregma and at a depth of L2/3 (200-400 μm).

Intrinsic Imaging

Prior to performing calcium imaging experiments, we performed intrinsic imaging of the optically accessible cortex to confirm the location of V1 within the cranial window (Supplementary Figure 1A,B). The intrinsic imaging was performed in all mice (n = 22) about 7-14 days after the surgery. We illuminated the cortex through the epi-illumination path using a high-power LED (central wavelength: 560 nm, M565L3, Thorlabs, Ely, UK), and acquired images at 5 Hz at 1024 x 1024 pixels using a CMOS camera (MV-D1024E-160; Photonfocus, Lachen, Switzerland) combined with a microscope objective (4x, NA: 0.13, FN: 26.5, UPLFLN, Olympus, Tokyo, Japan). To prevent the light contamination from the computer monitors we optically shielded the recording chamber with a custom black cone surrounding the objective.

Retinotopic mapping from intrinsic imaging

To obtain retinotopic maps from intrinsic imaging we used the methods described in Pisauro et al. (2013). Briefly, we first removed global fluctuations from the signal, which are not stimulus driven. The residual signal reflects the retinotopic, stimulus-evoked responses. Visual stimuli were periodic drifting and flickering bars (Kalatsky and Stryker, 2003). Flickering bars (flicker frequency 2 Hz) drifted (speed = 0.8 deg/s) across –135° to 45° of the horizontal visual field (with bars oriented vertically) and –45° to 45°of the vertical visual field (with bars oriented horizontally) for 3 cycles. We calculated retinotopic maps using the method described in Kalatsky and Stryker (2003). Retinotopic contours (Supplementary Figure 1B) where obtained after removal of artefactual extreme values (e.g. red regions in the top and bottom left corners of Supplementary Figure 1A) and replacing the removed values by values interpolated using sum of normalized Gaussian functions with standard deviation of 20 μm centered on non-artefactual pixels. Consistent with a location in V1, the imaged regions were within an area of diameter at least 2 mm where the gradient of vertical retinotopy was aligned from anterior to posterior (lower to higher values of elevation) and the gradient of the horizontal retinotopy was aligned from medial to lateral (temporal to central).

Visual stimuli

Stimuli were horizontal gratings drifting downward, presented in a location adjusted to match the center of GCaMP expression, on one of two screens that together spanned -45° to +135° of the horizontal visual field and ±42.5° of the vertical visual field (left and central screens in **Figure 1A**$_1$). During gray screen presentation, the screens were set to a steady gray level equal to the background of all the stimuli presented for visual responses protocols. Gratings had a duration of 1-2 s temporal frequency of 2 Hz and spatial frequency of 0.15 cycles/deg.

Eye-tracking movie acquisition and analysis

For eye tracking we used a collimated infrared LED (SLS-0208-B, λ_{peak} = 850nm; controller: SLC-AA02-US; Mightex Systems, Toronto, Canada) to illuminate the eye contralateral to the recording site. Videos of eye position were captured at 30 Hz with a monochromatic camera (DMK 21BU04.H, The Imaging Source, Bremen, Germany) equipped with a zoom lens (MVL7000; Navitar, Rochester, NY), and positioned at approximately 50° azimuth and 50° elevation relative to the center of the mouse' field of view. Contamination light from the monitors and the imaging laser was rejected using an optical band-pass filter (700-900nm) positioned in front of the camera objective (long-pass 092/52x0.75, The Imaging Source, Bremen, Germany; short-pass FES0900, Thorlabs, Ely UK).

To calibrate pupil displacement relative to the mouse visual field, we recorded additional movies at the end of each experiment while the mouse was still in exactly the same position as during the experiment. The eye was illuminated sequentially from a grid of known locations, the reflections were captured by the camera, and then this reflected grid was used to map the pupil displacement in pixels to pupil displacement in degrees of visual field.

Movie processing was performed offline using custom code written in MATLAB (Mathworks, Natick, MA) on a frame-by-frame basis. Briefly, each frame was mildly spatially low-pass filtered to reduce noise, then the pupil contour was detected by a level-crossing edge detector, and finally the position and the area of the pupil were calculated from the ellipse fit to the pupil contour. The output of the algorithm was visually inspected, and adjustments to the parameters (e.g. spatial filter strength, or level-crossing threshold) were made if necessary.

In vivo calcium imaging

Experiments were performed 16-34 days after virus injection (P54-110). We used a commercial two-photon microscope (B-scope, ThorLabs, Ely UK), with an acquisition frame rate of about 30Hz (at 512 by 512 pixels, corresponding to a rate of 4-6 Hz per plane), which was later interpolated to a frequency of 10 Hz, common to all planes. Recordings were performed in the area where expression was strongest. In most recordings (n = 16) this location was in the monocular zone (MZ, horizontal visual field preference > 30°)(Wagor et al., 1980). Other recordings (n = 11) were performed in the callosal binocular zone (CBZ, n= 4, 0-15°)(Wang and Burkhalter, 2007) and others (n = 7) in the acallosal binocular zone (ABZ, 15-30°). We observed no difference in results between recordings in monocular and binocular zones (Supplementary Figure 12).

Calcium data processing

Raw calcium movies were analyzed with Suite2p, which performs several processing stages (Pachitariu, 2016). First, Suite2p registers the movies to account for brain motion, then clusters neighboring pixels with similar timecourses into regions of interest (ROIs). ROIs were manually curated in the Suite2p GUI, to distinguish somata from dendritic processes based on their morphology. Cells expressing tdTomato were identified semi-automatically using an algorithm based on their average fluorescence in the red channel. For spike deconvolution from the Calcium traces, we used the default method in Suite2p (Pachitariu, 2016). Whether we performed performing spike deconvolution or analysed raw calcium signals made no difference to our results (Supplementary Figure 13).

Pixel maps of calcium data

To confirm the correlation of running speed and fluorescence independent of ROI detection, we computed correlation maps (Supplementary Figure 2C), showing for each pixel the Pearson correlation between the activity of the pixel and the running speed (c.f. Freeman et al. (2014)). Prior to correlation, the activity of each pixel was smoothed by convolving with a spatial Gaussian with standard deviation equal to 1.5 pixels, and a temporal Hamming window of 1 s width.

The correlation of baseline fluorescence with running speed varied across the field of view. In regions where GCaMP expression level was high, baseline fluorescence correlated positively with running speed, likely indicating an increase in axonal and dendritic activity in locomoting animals. However, in areas where GCaMP fluorescence was weak, the correlation of the background with running speed was negative, likely indicating that in absence of GCaMP the signal is dominated by increased hemodynamic filtering of the light due to stronger blood flow during running (Huo et al., 2015). To ensure this did not affect our results, we removed background fluorescence from the detected fluorescence of recorded neurons (see below).

Background fluorescence correction

With two-photon GCaMP imaging, an important concern is that out-of-focus fluorescence can contaminate the signal ascribed to particular neurons; this is of particular concern in situations where the surrounding GCaMP-labelled neuropil may itself show modulation by stimuli or behaviors such as locomotion. In order to correct out-of-focus contamination, we adopted the method of Peron et al. (2015). A “neuropil mask” was defined as the region up to 35 μm from the ROI border, excluding pixels corresponding to other detected cells (Supplementary Figure 2A$_1$), and the fluorescence signal in this mask region was subtracted from that of the cell soma, weighted by a correction factor α_{exp} that was determined separately for each experiment.

To determine the correction factor, we estimated the linear relationship specifying the lowest possible somatic fluorescence compatible with any value of fluorescence in the neuropil mask (Supplementary Figure

2A$_2$). To do so, for each cell i we binned the neuropil signals $N_i(t)$ into 20 intervals, and for each one estimated the 5th percentile of the raw somatic fluorescence $F_i(t)$. We computed α_i by linear regression, which accurately matched the lower envelope of the scatterplot of neuropil vs. somatic fluorescence (Supplementary Figure 2A$_2$). This method gave consistent results for sparse firing cells, but not always for densely firing cells for which a correlation of cellular activity with the neuropil signal could lead to misestimated slopes, as densely firing cells might only rarely exhibit baseline fluorescence. We therefore computed the correction factor α_{exp} for each experiment by averaging α_i over cells with high skewness (> 4). The corrected fluorescence was computed as $F(t) - \alpha_{exp}N(t)$ (Supplementary Figure 2A$_3$). In experiments where only interneurons (and thus low skewed cells) expressed GCaMP6, we used as a correction factor an average from all the other experiments equal to $\langle\alpha_{exp}\rangle = 0.82$.

Analysis of neural activity

The average fluorescence response to each stimulus was defined by $\Delta F/F_0 = (F-F_0)/F_0$, where F is the average raw calcium signal during the first second of the stimulus presentation, and F_0 is the global minimum of the fluorescence trace filtered with a Hamming window of duration 0.5 s. The correlation of neural activity with locomotion speed during gray screen presentation was assessed by the Pearson correlation coefficient between the calcium signal and the locomotion speed trace, on an interpolated timebase of 10 Hz, smoothed (5 points) and decimated (1 Hz). To ascertain the significance of this correlation we used a shuffling method, in which the speed trace was randomly circularly shifted relative to the fluorescence trace 1,000 times – this was necessary because serial correlation in the time series of fluorescence and speed rendered successive samples statistically dependent.

The size of a cell's response to a stimulus was defined by the difference of $\Delta F/F_0$ between the first 1 s of the stimulus period, and the 1 s of baseline activity prior to stimulus presentation. We defined a neuron to have significant size tuning if it passed in at least one of the two locomotion conditions (rest or running) a one-way ANOVA test ($p < 0.05$) comparing the mean visual responses to different stimuli.

To measure each cell's retinotopic location, in the majority of datasets (n = 24) receptive fields were obtained from responses to sparse, uncorrelated noise. The screen was divided into squares of 5 by 5 degrees, and each square was independently turned on and off randomly at a 5Hz overall rate. At any time, 5% of all squares were on. Each cell's response to each square was obtained using stimulus-triggered averaging of the non-neuropil corrected trace. The RFs were smoothed in space and their peak was identified as the preferred spatial position. In a subset of early experiments (n = 3), sparse noise was not presented, and RFs were assessed with flickering vertical or horizontal bars appearing in different locations; we verified in a further n = 4 datasets that the two measures of a cell's receptive field were consistent.

When computing size tuning curves, we normalized the calcium activity (Figure 3, column 3 and Supplementary Figure 12) or the spike rate (Supplementary Figure 13) in the following way: for each cell, the response ΔF_0 to the "blank condition" (i.e. a stimulus of contrast 0) during stationary periods was subtracted from the raw cell response ΔF (1 s during stimulation minus 1 s of baseline activity): $\Delta F \to \Delta F - \Delta F_0$. Then, for each recording we computed the average $\langle\Delta F\rangle_{all\ cells}$ over all selected cells whose distance of the receptive field from the stimulus center r was < 20°. Finally we divided the average response of either the centered cells (r < 10°) or the off-centered cells (r > 10°, r < 20°) by the maximum value of all the cells combined: $\langle\Delta F\rangle_{cent.}/\max(\langle\Delta F\rangle_{all\ cells})$ or $\langle\Delta F\rangle_{off-cent.}/\max(\langle\Delta F\rangle_{all\ cells})$. We then averaged these values across all recording sessions.

When computing the interaction between locomotion and size (Figure 3, column 4 and Supplementary Figure 10A,B) we normalized the responses of the individual cells for each experiment in the following way: after subtracting the average response to the blank during stationary periods $\Delta F \to \Delta F - \Delta F_0$, we divided the responses by the average minimal calcium trace across cells $\langle F_0\rangle_{cells}$ within the same experiment or mouse (Supplementary Figure 10A,B).

Curve fitting

We fitted the size tuning curves of Figure 3 and Supplementary Figures 4, 8, 9, 12, and 13 by least squares with the following function family: $f(s) = R[\mathrm{erf}(s/\sigma_1) -$

$k \cdot \mathrm{erf}(s/\sigma_2)]$ where $\mathrm{erf}(x)$ corresponds to the error function and s is the size of the stimulus. The free parameters of the function are R, k, σ_1 and σ_2. To estimate nonlinear signal correlation curves (Figure 4, columns 3 and 4), we first smoothed the responses to large sizes (s >20°) for each population with a moving average method with span 25°. Then we smoothed again the responses for all sizes with a moving average method with span 20°. Finally we interpolated the values between the measured size with a shape-preserving piecewise cubic interpolation.

Size tuning maps

To compute how size tuning depends on stimulus centering, we computed two-dimensional maps illustrating how each cell class' average activity depends on stimulus diameter s, and the offset of the receptive field center from the stimulus center r_i (Supplementary Figure 11). To do so, we first computed for each cell i a normalized tuning curve $n_{i,v}(s)$, where v represents locomotion condition. Dependence on r_i was estimated by smoothing: a two-dimensional map was made for each cell as an outer product: $f_{i,v}(s,r) = n_{i,v}(s) g_i(r)$, where $g_i(r) = e^{-\frac{(r-r_i)^2}{2\sigma_y^2}}$ is a Gaussian centered at the offset value r_i of width $\sigma_y = 5°$. Then we summed the maps belonging to one recording session j and divided by the sum of all the Gaussians centered at different offsets: $m_{j,v}(s,r) = \sum_{i \in j} f_{i,v}(s,r) / \sum_{i \in j} g_i(s,r)$, and normalized this value for each experiment by the value at stationary, 0 offset and diameter 10° $f_{j,v}(s,r) = m_{j,v}(s,r)/m_{j,v=1}(s = 10, r = 0)$. Finally, for each cell class we obtained the size-tuning offset maps by averaging across experiments: $\langle f_{j,v}(s,r) \rangle_j$.

Inter-population correlation analysis

To compute spontaneous correlations (Figure 4, columns 1 and 2), we first normalized the deconvolved spike trace S of each cell *i* over time *t*: $n(i,t) = S(i,t)/max_t(S)$. Then, for each experiment and each class *c* (interneurons or putative Pyr cells) we computed the average population rate across cells i belonging to class c: $R_c(t) = \langle n(i,t) \rangle_{\{i\} \in c}$. We then smoothed $R_c(t)$ with a moving average method with span of 1 s and then decimated the sampling rate to 1 point every 1 s. To make the plots of different experiments visually comparable we normalized these responses: $K_c(t) = (R_c(t) - R_c^0)/\sigma_t(R_c)$. Where R_c^0 is the 1st percentile of R_c and $\sigma_t(R_c)$ the standard deviation of $R_c(t)$ across time.

To compute signal correlations (Figure 4, columns 3 and 4), for each experiment and each cell class *c* we first computed the average population response for each stimulus size *s* and locomotion condition *v* (*v = 0* stationary, *v = 1* running) by averaging over all cells *i* belonging to that class: $\Delta F_c(s,v) = \langle \Delta f(i,s,v) \rangle_{\{i\} \in c}$. To make the plots of different experiments visually comparable we normalized the responses: $\Delta R_c(s,v) = \Delta F_c(s,v) / \sigma_{s,v}(\Delta F_c)$. Finally we subtracted the blank response during the resting condition: $\Delta K_c(s,v) = \Delta R_c(s,v) - \Delta R_c(s=0, v=0)$.

To compute noise correlations (Supplementary Figure 14, columns 1 and 2), for each experiment and each cell class *c* we first computed the average population response for each trial *t* by averaging over all cells *i* belonging to that class: $\Delta F_c(t) = \langle \Delta f(i,t) \rangle_{\{i\} \in c}$. Then for each stimulus and locomotion condition we subtracted the mean response from the related trials: $\Delta N_c(t \in \{s,v\}) = \Delta F_c(t \in \{s,v\}) - \langle \Delta F_c \rangle_{t \in \{s,v\}}$. Finally, to make the plots of different experiments visually comparable, we normalized the responses by z-scoring over all trials: $\Delta Z_c(t) = (\Delta N_c(t) - \langle \Delta N_c \rangle_t)/\sigma_t(\Delta N_c)$.

When measuring signal and noise correlations and for both interneurons and putative Pyr neurons, we selected cells whose receptive field center was <10° away from the stimulus center. For Figure 4 we selected putative Pyr cells as unlabeled (non tdTomato) neurons whose skewness was > 2.7. In a control analysis (Supplementary Figure 15) we show that the value of the skewness threshold makes little difference to these results. A skewness value of 0 corresponds to the case where we selected all unlabeled cells as putative Pyr cells.

Correlation of running modulation with depth

To determine whether running modulation of a given cell class varied significantly with cortical depth, we computed $\rho(gray)$ for each cell as the Pearson correlation of that cell's neuropil-corrected fluorescence (without spike deconvolution) and running speed. We then assessed a significant relationship of $\rho(gray)$

with depth using robust regression (bisquare-weighting).

Computational model

We asked whether we could predict the mean size tuning of each cell class using a neural field theory model. In this model, the mean firing rate of cells of type α is captured by a function $f_\alpha(s, \mathbf{r})$, where s represents the stimulus size, and $\mathbf{r}$ represents position on the cortical surface, measured in retinotopic coordinates. We model the external excitatory input arriving at point r (e.g. from thalamus or other cortical layers) by a function $h(s, \mathbf{r})$, again in retinotopic coordinates.

We denote the experimentally measured responses of cell class α by $f_\alpha^{(v)}(s, \mathbf{r})$, where v denotes locomotion condition ($v = 0$: stationary, $v = 1$: running). We assume that responses are circularly symmetrical, i.e. that responses depend on $\mathbf{r}$ only through the radial distance of the receptive field center from the stimulus center, r. The response of each cell class $f_\alpha^{(v)}(s, \mathbf{r})$ is modeled by the following equations:

$$\begin{cases} f_E = R[w_{EH}h + w_{EE}(G_{EE} * f_E) - w_{EP}(G_{EP} * f_P);\ w_{ES}(G_{ES} * f_S)] \\ f_P = R[w_{PH}h + w_{PE}(G_{PE} * f_E) - w_{PP}(G_{PP} * f_P);\ w_{PS}(G_{PS} * f_S)] \\ f_S = R[w_{SH}(G_{SH} * h) + w_{SE}(G_{SE} * f_E);\ w_{SV}(G_{SV} * f_V)] \\ f_V = R[w_{VE}(G_{VE} * f_E);\ w_{VS}(G_{VS} * f_S)] \end{cases}$$

Only synaptic connections demonstrated *in vitro* (Pfeffer et al., 2013) are included in this equation; however, adding other potential synapses (*Vip*->Pyr) did not improve the fit (Supplementary Figure 16).

For each postsynaptic cell class we tested different combination of subtractive and divisive inhibition from *Sst* and *Vip* cells:

$$R(x; y) = \begin{cases} \lfloor x - y \rfloor_+ & \mathit{subtractive} \\ \left\lfloor \dfrac{x}{1 + y} \right\rfloor_+ & \mathit{divisive} \end{cases}$$

Here, f_E, f_P f_S, and f_V reflect the visual responses of the Pyr, *Pvalb*, *Sst*, and *Vip* cells respectively; $\lfloor z \rfloor_+$ is the positive part of z; $w_{\alpha\beta}$ are the peak synaptic weights between the presynaptic cell class β and the postsynaptic cell class α (which can in principle depend on running condition v); $G_{\alpha\beta}$ is a two-dimensional Gaussian function defined by $G_{\alpha\beta}(\mathbf{z}) = \exp[-|\mathbf{z}|^2/(2\sigma_{\alpha\beta}^2)]/(2\pi\sigma_{\alpha\beta}^2)$, with radius $\sigma_{\alpha\beta}$ that can depend on the pre- and post-sysnpatic cell type; and $*$ represents convolution over retinotopic space: $[G_{\alpha\beta} * f_\beta](s, \mathbf{r}) = \iint G_{\alpha\beta}(\mathbf{r} - \mathbf{r}') f_\beta(s, \mathbf{r}') d^2\mathbf{r}'$.

The equations describing the activity of the cell classes can be simplified with the following assumptions:

- $\forall (v, s, r), f_E > 0, f_P > 0$ (Pyr and *Pvalb* cells are not suppressed by stimuli, as seen in the data).
- The recurrent connections of Pyr and *Pvalb* neurons, and the connection from *Pvalb* to Pyr are local: $G_{EE}(r) = G_{EP}(r) = G_{PP}(r) = \delta(r)$ where $\delta(r)$ is the Dirac delta function.
- $f_P \approx \mu f_E$ (i.e. Pvalb activity closely tracks Pyr activity, as seen in the data).

We can then rewrite the equations as:

$$\begin{cases} f_E = R[w_{EH}h;\ w_{ES}(G_{ES} * f_S)]/(1 - w_{EE} + \mu w_{EP}) \\ f_P = R[w_{PH}h + w_{PE}(G_{PE} * f_E);\ w_{PS}(G_{PS} * f_S)]/(1 + w_{PP}) \\ f_S = R[w_{SH}(G_{SH} * h) + w_{SE}(G_{SE} * f_E);\ w_{SV}(G_{SV} * f_V)] \\ f_V = R[w_{VE}(G_{VE} * f_E);\ w_{VS}(G_{VS} * f_S)] \end{cases}$$

We can further simplify this equation as:

$$\begin{cases} f_E = R[\widetilde{w}_{EH}h; \widetilde{w}_{ES}(G_{ES} * f_S)] \\ f_P = R[\widetilde{w}_{PH}h + \widetilde{w}_{PE}(G_{PE} * f_E); \widetilde{w}_{PS}(G_{PS} * f_S)] \\ f_S = R[w_{SH}(G_{SH} * h) + w_{SE}(G_{SE} * f_E);\ w_{SV}(G_{SV} * f_V)] \\ f_V = R[w_{VE}(G_{VE} * f_E);\ w_{VS}(G_{VS} * f_S)] \end{cases} \tag{1}$$

where $\widetilde{w}_{EH}$ and $\widetilde{w}_{ES}$ are the "effective connections" from external inputs and *Sst* neurons to excitatory cells, given by $w_{EH}/(1-w_{EE}+\mu w_{EP})$ and $w_{ES}/(1-w_{EE}+\mu w_{EP})$ while $\widetilde{w}_{PH}$, $\widetilde{w}_{PE}$, and $\widetilde{w}_{PS}$ represent the effective weights onto *Pvalb* cells, equal to $w_{PH}/(1+w_{PP})$, $w_{PE}/(1+w_{PP})$, and $w_{PS}/(1+w_{PP})$.

Estimation of thalamic input

We estimated the average external inputs $h(s,0)$ from the experimentally-measured responses of centered ($r=0$) thalamic cells using the data in Erisken et al. (2014), for stationary and locomotion periods separately. We first fit the size tuning curve of each cell (n = 21) with a ratio of Gaussians $h_i^{(v)}(s) = b_i\,[\mathrm{erf}(s_i/m_i)]^2/\{1+b_i[\mathrm{erf}(s_i/m_i)]^2\}$ as in Erisken et al. (2014). We then normalized each $h_i^{(v)}(s)$ by its maximum across s and v, and estimated the mean visual responses by averaging across all cells. To fit off-center responses, we extended this approach using a Ratio-of-Gaussians model (Ayaz et al., 2013): $h(s,r) = a_1\,u(s,r,\sigma_1)/[1+a_2 u(s,r,\sigma_2)]$ with $u(s,r,\sigma) = \mathrm{erf}[(s+r)/\sigma] + \mathrm{sign}(s-r)\mathrm{erf}(|s-r|/\sigma)$. The estimated parameters are during the stationary periods were $a_1 = 1.2$, $a_2 = 1.9$, $r_1 = 36.7$, and $r_2 = 33.9$ while during locomotion $a_1 = 0.5$, $a_2 = 0.4$, $\sigma_1 = 24.7°$, and $\sigma_2 = 10.0°$.

Estimation of presynaptic inputs

To fit the model, we clamped the firing rate functions $f_\alpha^{(v)}(s,\mathbf{r})$ in equation (1) to their experimentally-measured values, and fit synaptic parameters to reduce the discrepancy between the right and left sides. To do so required extending our experimental data to continuous functions of s and r. For retinotopic positions $r < r_L = 33^{\mathrm{o}}$, $f(s,r)$ was fitted from the data with a difference of Gaussians function: $f(s,r) = R_r\left[\mathrm{erf}\left(s/\sigma_{1,r}\right) - k_r\cdot\mathrm{erf}\left(s/\sigma_{2,s}\right)\right]$. Because our data for cells off-center by more than 33° were sparse, we extrapolated the values for $r > r_L$ with a decaying exponential approximation: $f(s,r) \approx a_s e^{-(r-r_m)/b_s}$, where r_m is the offset value that maximizes the response of that cell class, and the parameters a_s and b_s were estimated for each stimulus size to minimize squared error in the range $r_m \le r \le r_L$. Finally, the values of b_s were interpolated beyond the presented stimulus sizes with a third degree polynomial function.

Parameter estimation

To estimate the parameters we minimized an objective function equal to the normalized mean-square error, plus additional penalty terms to favor simpler models:

$$Err = \frac{\langle\left[\hat{f}_\alpha^{(v)}(s,r) - f_\alpha^{(v)}(s,r)\right]^2\rangle_{s,r,v}}{\langle\mathrm{var}_x\left[f_{\alpha,x}^{(v)}(s,r)\right]\rangle_{s,r,v}} + \lambda_1 n_{\Delta w} + \lambda_2\left\|\frac{\sigma_{\alpha\beta}}{\sigma_L}\right\|_2^2 + \lambda_3 R_\alpha^2$$

Here, $f_\alpha^{(v)}$ represents the measured firing rate, and $\hat{f}_\alpha^{(v)}$ represents the right hand size of equation (1). The normalization factor of $1/\langle\mathrm{var}_x\left[f_{\alpha,x}^{(v)}(s,r)\right]\rangle_{s,r,v}$ ensures that conditions with high intertrial variability do not overly influence the objective function; the normalized error can also be interpreted as the log-likelihood of the model fit under a Gaussian distribution estimated from all experiments ($\mathrm{var}_x\left[f_{\alpha,x}^{(v)}(s,r)\right]$ denotes the variance of normalized visual responses for each value of s, r and v over all experiments x).

The last three terms represent regularization parameters. The first regularization term controls the number of synaptic strengths that are allowed to change with locomotion, $n_{\Delta w}$; for the current analysis we used a value of $\lambda_1 = 0.1$. The second regularization term controls the spatial distribution of synaptic weights; we used parameters $\sigma_L = 40°$, $\lambda_2 = 0.01$. The final term R_α, with $\lambda_3 = 0.2$, represents a factor to add biologically motivated constraints in the *Pvalb* and *Sst* equations.

We assume that Sst cells receive most of their input from a local area, therefore

$$R_S = \textstyle\sum_{v=1,2} w_{SH}^{(v)}/\left(w_{SE}^{(v)} + w_{SV}^{(v)}\right).$$

The activities of Pyr and Pvalb are very similar, so to avoid f_E dominating the f_P equation we have

$$R_P = \textstyle\sum_{v=1,2} w_{PE}^{(v)}/\left(w_{PH}^{(v)} + w_{PS}^{(v)}\right).$$

To determine the optimal parameters of the model we first performed an exhaustive search over the extent of the spatial integration $\sigma_{\alpha\beta}$ parameters of all the presynaptic cell classes for each postsynaptic cell class. For each combination of $\{\sigma_{\alpha\beta_1}, \sigma_{\alpha\beta_2,\ldots}\}$ we then found the optimal synaptic strength parameters $w_{\alpha\beta}^{(v)}$ using a combination of the trust region reflective and Levenberg-Marquardt algorithms (MATLAB), after 50 random initialization of the initial parameters $w_{\alpha\beta}^{(v)}$. We then chose the values of $\sigma_{\alpha\beta}$ and $w_{\alpha\beta}^{(v)}$ minimizing Err.

To fit the way locomotion affects synaptic strengths, we found the optimal of several possibilities for each synaptic type: 1) equal weights in locomotion and stationary $w_{\alpha\beta}^{(0)} = w_{\alpha\beta}^{(1)}$; 2) fixing all but one of the synaptic weights $w_{\alpha\beta}^{(v)}$ (therefore testing one model for each synaptic connection) 3) allowing all the synaptic weights to change with locomotion. The final model was chosen as that minimizing the penalized error function Err (Supplementary Figure 16A-D_1).

Author contributions

M.D. performed the experiments, analyzed the data and implemented the model; M.D., A.R. and M.P. performed the surgeries; A.R. and M.K. built the two-photon set-up; M.D, M.K. and M.P. implemented the retinotopic mapping software; K.D.H. and M.P. implemented the ROI detection algorithms; M.D., A.R., M.C. and K.D.H. designed the experiments; M.D. and K.D.H. designed the model; M.D., M.C. and K.D.H. wrote the paper with inputs from all authors.

Acknowledgments

We thank Charu Reddy for transgenic mouse breeding and maintenance, surgeries, and technical support, Chris Burgess for building the data acquisition software set-up, Luigi Federico Rossi for assistance during surgeries, Miles Wells for assistance in mice habituation, Andrea Pisauro for implementing the intrinsic imaginc retinotopic mapping software, Sylvia Schroeder for implementing the calcium retinotopic mapping software. We thank Boris S. Gutkin for constructive discussion. For the recorded size tuning curves of the dLGN cells in vivo we acknowledge Laura Busse's lab. For the use of GCaMP6f we acknowledge Vivek Jayaraman, Rex A. Kerr, Douglas S. Kim, Loren L. Looger, Karel Svoboda from the GENIE Project, Janelia Farm Research Campus, Howard Hughes Medical Institute. This work was supported by the Wellcome Trust. MD was supported by a Marie Curie Intra-European Fellowship for Career Development, KDH by EPSRC. MC holds the GlaxoSmithKline / Fight for Sight Chair in Visual Neuroscience.

References

Acsady, L., Arabadzisz, D., and Freund, T.F. (1996a). Correlated morphological and neurochemical features identify different subsets of vasoactive intestinal polypeptide-immunoreactive interneurons in rat hippocampus. Neuroscience 73, 299-315.

Acsady, L., Gorcs, T.J., and Freund, T.F. (1996b). Different populations of vasoactive intestinal polypeptide-immunoreactive interneurons are specialized to control pyramidal cells or interneurons in the hippocampus. Neuroscience 73, 317-334.

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

Ayaz, A., Saleem, A.B., Scholvinck, M.L., and Carandini, M. (2013). Locomotion controls spatial integration in mouse visual cortex. Current biology : CB 23, 890-894.

Chen, T.W., Wardill, T.J., Sun, Y., Pulver, S.R., Renninger, S.L., Baohan, A., Schreiter, E.R., Kerr, R.A., Orger, M.B., Jayaraman, V., *et al.* (2013). Ultrasensitive fluorescent proteins for imaging neuronal activity. Nature 499, 295-300.

Cruikshank, S.J., Lewis, T.J., and Connors, B.W. (2007). Synaptic basis for intense thalamocortical activation of feedforward inhibitory cells in neocortex. Nature neuroscience 10, 462-468.

Cruikshank, S.J., Urabe, H., Nurmikko, A.V., and Connors, B.W. (2010). Pathway-specific feedforward circuits between thalamus and neocortex revealed by selective optical stimulation of axons. Neuron 65, 230-245.

Douglas, R.J., Koch, C., Mahowald, M., Martin, K.A.C., and Suarez, H.H. (1995). Recurrent Excitation in Neocortical Circuits. Science 269, 981-985.

Erisken, S., Vaiceliunaite, A., Jurjut, O., Fiorini, M., Katzner, S., and Busse, L. (2014). Effects of locomotion extend throughout the mouse early visual system. Current biology : CB 24, 2899-2907.

Freeman, J., Vladimirov, N., Kawashima, T., Mu, Y., Sofroniew, N.J., Bennett, D.V., Rosen, J., Yang, C.T., Looger, L.L., and Ahrens, M.B. (2014). Mapping brain activity at scale with cluster computing. Nature methods 11, 941-950.

Fu, Y., Tucciarone, J.M., Espinosa, J.S., Sheng, N., Darcy, D.P., Nicoll, R.A., Huang, Z.J., and Stryker, M.P. (2014). A cortical circuit for gain control by behavioral state. Cell 156, 1139-1152.

Garcia-Junco-Clemente, P., Ikrar, T., Tring, E., Xu, X.M., Ringach, D.L., and Trachtenberg, J.T. (2017). An inhibitory pull-push circuit in frontal cortex. Nature neuroscience 20, 389-392.

Gentet, L.J., Avermann, M., Matyas, F., Staiger, J.F., and Petersen, C.C. (2010). Membrane potential dynamics of GABAergic neurons in the barrel cortex of behaving mice. Neuron 65, 422-435.

Gentet, L.J., Kremer, Y., Taniguchi, H., Huang, Z.J., Staiger, J.F., and Petersen, C.C. (2012). Unique functional properties of somatostatin-expressing GABAergic neurons in mouse barrel cortex. Nature neuroscience 15, 607-612.

Gil, Z., Connors, B.W., and Amitai, Y. (1997). Differential regulation of neocortical synapses by neuromodulators and activity. Neuron 19, 679-686.

Hippenmeyer, S., Vrieseling, E., Sigrist, M., Portmann, T., Laengle, C., Ladle, D.R., and Arber, S. (2005). A developmental switch in the response of DRG neurons to ETS transcription factor signaling. PLoS biology 3, e159.

Hu, H., Cavendish, J.Z., and Agmon, A. (2013). Not all that glitters is gold: off-target recombination in the somatostatin-IRES-Cre mouse line labels a subset of fast-spiking interneurons. Frontiers in neural circuits 7, 195.

Huo, B.X., Gao, Y.R., and Drew, P.J. (2015). Quantitative separation of arterial and venous cerebral blood volume increases during voluntary locomotion. Neuroimage 105, 369-379.

Isaacson, J.S., and Scanziani, M. (2011). How inhibition shapes cortical activity. Neuron 72, 231-243.

Jiang, X., Shen, S., Cadwell, C.R., Berens, P., Sinz, F., Ecker, A.S., Patel, S., and Tolias, A.S. (2015). Principles of connectivity among morphologically defined cell types in adult neocortex. Science 350, aac9462.

Kalatsky, V.A., and Stryker, M.P. (2003). New paradigm for optical imaging: temporally encoded maps of intrinsic signal. Neuron 38, 529-545.

Karnani, M.M., Jackson, J., Ayzenshtat, I., Hamzehei Sichani, A., Manoocheri, K., Kim, S., and Yuste, R. (2016a). Opening Holes in the Blanket of Inhibition: Localized Lateral Disinhibition by VIP Interneurons. The Journal of neuroscience : the official journal of the Society for Neuroscience 36, 3471-3480.

Karnani, M.M., Jackson, J., Ayzenshtat, I., Tucciarone, J., Manoocheri, K., Snider, W.G., and Yuste, R. (2016b). Cooperative Subnetworks of Molecularly Similar Interneurons in Mouse Neocortex. Neuron 90, 86-100.

Keller, G.B., Bonhoeffer, T., and Hubener, M. (2012). Sensorimotor mismatch signals in primary visual cortex of the behaving mouse. Neuron 74, 809-815.

Kepecs, A., and Fishell, G. (2014). Interneuron cell types are fit to function. Nature 505, 318-326.

Lee, S., Kruglikov, I., Huang, Z.J., Fishell, G., and Rudy, B. (2013). A disinhibitory circuit mediates motor integration in the somatosensory cortex. Nature neuroscience 16, 1662-1670.

Litwin-Kumar, A., Rosenbaum, R., and Doiron, B. (2016). Inhibitory stabilization and visual coding in cortical circuits with multiple interneuron subtypes. Journal of neurophysiology 115, 1399-1409.

Madisen, L., Garner, A.R., Shimaoka, D., Chuong, A.S., Klapoetke, N.C., Li, L., van der Bourg, A., Niino, Y., Egolf, L., Monetti, C.*, et al.* (2015). Transgenic mice for intersectional targeting of neural sensors and effectors with high specificity and performance. Neuron 85, 942-958.

Madisen, L., Zwingman, T.A., Sunkin, S.M., Oh, S.W., Zariwala, H.A., Gu, H., Ng, L.L., Palmiter, R.D., Hawrylycz, M.J., Jones, A.R.*, et al.* (2010). A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. Nature neuroscience 13, 133-140.

Markram, H., Muller, E., Ramaswamy, S., Reimann, M.W., Abdellah, M., Sanchez, C.A., Ailamaki, A., Alonso-Nanclares, L., Antille, N., Arsever, S.*, et al.* (2015). Reconstruction and Simulation of Neocortical Microcircuitry. Cell 163, 456-492.

Niell, C.M., and Stryker, M.P. (2010). Modulation of visual responses by behavioral state in mouse visual cortex. Neuron 65, 472-479.

Okun, M., and Lampl, I. (2008). Instantaneous correlation of excitation and inhibition during ongoing and sensory-evoked activities. Nature neuroscience 11, 535-537.

Ozeki, H., Finn, I.M., Schaffer, E.S., Miller, K.D., and Ferster, D. (2009). Inhibitory stabilization of the cortical network underlies visual surround suppression. Neuron 62, 578-592.

Pachitariu, M., Stringer, C., Schröder, S., Dipoppa, M., Rossi, L. F., Carandini, M., & Harris, K. D. (2016). Suite2p: beyond 10,000 neurons with standard two-photon microscopy. bioRxiv, 061507.

Pakan, J.M., Lowe, S.C., Dylda, E., Keemink, S.W., Currie, S.P., Coutts, C.A., and Rochefort, N.L. (2016). Behavioral-state modulation of inhibition is context-dependent and cell type specific in mouse visual cortex. eLife 5.

Pecka, M., Han, Y., Sader, E., and Mrsic-Flogel, T.D. (2014). Experience-dependent specialization of receptive field surround for selective coding of natural scenes. Neuron 84, 457-469.

Peron, S.P., Freeman, J., Iyer, V., Guo, C., and Svoboda, K. (2015). A Cellular Resolution Map of Barrel Cortex Activity during Tactile Behavior. Neuron 86, 783-799.

Pfeffer, C.K., Xue, M., He, M., Huang, Z.J., and Scanziani, M. (2013). Inhibition of inhibition in visual cortex: the logic of connections between molecularly distinct interneurons. Nature neuroscience 16, 1068-1076.

Pi, H.J., Hangya, B., Kvitsiani, D., Sanders, J.I., Huang, Z.J., and Kepecs, A. (2013). Cortical interneurons that specialize in disinhibitory control. Nature 503, 521-524.

Pisauro, M.A., Dhruv, N.T., Carandini, M., and Benucci, A. (2013). Fast hemodynamic responses in the visual cortex of the awake mouse. The Journal of neuroscience : the official journal of the Society for Neuroscience 33, 18343-18351.

Polack, P.O., Friedman, J., and Golshani, P. (2013). Cellular mechanisms of brain state-dependent gain modulation in visual cortex. Nature neuroscience 16, 1331-1339.

Prinz, A.A., Bucher, D., and Marder, E. (2004). Similar network activity from disparate circuit parameters. Nature neuroscience 7, 1345-1352.

Reimer, J., Froudarakis, E., Cadwell, C.R., Yatsenko, D., Denfield, G.H., and Tolias, A.S. (2014). Pupil Fluctuations Track Fast Switching of Cortical States during Quiet Wakefulness. Neuron 84, 355-362.

Renart, A., de la Rocha, J., Bartho, P., Hollender, L., Parga, N., Reyes, A., and Harris, K.D. (2010). The asynchronous state in cortical circuits. Science 327, 587-590.

Rubin, D.B., Van Hooser, S.D., and Miller, K.D. (2015). The Stabilized Supralinear Network: A Unifying Circuit Motif Underlying Multi-Input Integration in Sensory Cortex. Neuron 85, 402-417.

Saleem, A.B., Ayaz, A., Jeffery, K.J., Harris, K.D., and Carandini, M. (2013). Integration of visual motion and locomotion in mouse visual cortex. Nature neuroscience 16, 1864-1869.

Schneider, D.M., Nelson, A., and Mooney, R. (2014). A synaptic and circuit basis for corollary discharge in the auditory cortex. Nature 513, 189-194.

Tan, Z.J., Hu, H., Huang, Z.J., and Agmon, A. (2008). Robust but delayed thalamocortical activation of dendritic-targeting inhibitory interneurons. Proceedings of the National Academy of Sciences of the United States of America 105, 2187-2192.

Taniguchi, H., He, M., Wu, P., Kim, S., Paik, R., Sugino, K., Kvitsiani, D., Fu, Y., Lu, J., Lin, Y., *et al.* (2011). A resource of Cre driver lines for genetic targeting of GABAergic neurons in cerebral cortex. Neuron 71, 995-1013.

Tasic, B., Menon, V., Nguyen, T.N., Kim, T.K., Jarsky, T., Yao, Z., Levi, B., Gray, L.T., Sorensen, S.A., Dolbeare, T., *et al.* (2016). Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. Nature neuroscience 19, 335-346.

Tremblay, R., Lee, S., and Rudy, B. (2016). GABAergic Interneurons in the Neocortex: From Cellular Properties to Circuits. Neuron 91, 260-292.

Vinck, M., Batista-Brito, R., Knoblich, U., and Cardin, J.A. (2015). Arousal and locomotion make distinct contributions to cortical activity patterns and visual encoding. Neuron 86, 740-754.

Wagor, E., Mangini, N.J., and Pearlman, A.L. (1980). Retinotopic organization of striate and extrastriate visual cortex in the mouse. The Journal of comparative neurology 193, 187-202.

Wang, Q., and Burkhalter, A. (2007). Area map of mouse visual cortex. The Journal of comparative neurology 502, 339-357.

Zeisel, A., Munoz-Manchado, A.B., Codeluppi, S., Lonnerberg, P., La Manno, G., Jureus, A., Marques, S., Munguba, H., He, L., Betsholtz, C., *et al.* (2015). Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. Science 347, 1138-1142.

Zhang, S., Xu, M., Kamigaki, T., Hoang Do, J.P., Chang, W.C., Jenvay, S., Miyamichi, K., Luo, L., and Dan, Y. (2014). Selective attention. Long-range and local circuits for top-down modulation of visual cortex processing. Science 345, 660-665.

Zhou, M., Liang, F., Xiong, X.R., Li, L., Li, H., Xiao, Z., Tao, H.W., and Zhang, L.I. (2014). Scaling down of balanced excitation and inhibition by active behavioral states in auditory cortex. Nature neuroscience 17, 841-850.