An anatomically constrained model of V1 simple cells predicts the coexistence of push-pull and broad inhibition.

M. Morgan Taylor¹, Diego Contreras¹, Alain Destexhe², Yves Frégnac², Jan Antolik³*

1. Dept of Neuroscience, University of Pennsylvania School of Medicine, Philadelphia, 19104, PA.
2. Paris Saclay University, Institute of Neuroscience (NeuroPSI-UNIC), CNRS, Gif sur Yvette, 91198, France.
3. Faculty of Mathematics and Physics, Charles University, Prague, 11800, Czechia.

* Corresponding author: antolik@ksvi.mff.cuni.cz

The authors declare no competing financial interests.

Number of pages: 45
Number of figures: 9
Number of words (abstract, introduction, discussion): 199, 677, 1801

Abbreviated title: A data-driven model of simple cells in cat V1

Abstract

The spatial organization and dynamic interactions between excitatory and inhibitory synaptic inputs that define the receptive field (RF) of simple cells in cat primary visual cortex (V1) still raise paradoxical issues: 1) stimulation of simple cells in V1 with drifting gratings supports a wiring schema of spatially segregated sets of excitatory and inhibitory inputs activated in an opponent way by stimulus contrast polarity; 2) in contrast, intracellular studies using flashed bars suggest that, while ON and OFF excitatory input are indeed segregated, inhibitory inputs span the entire RF irrespective of input contrast polarity. Here, we propose a biologically detailed computational model of simple cells embedded in a V1-like network that resolves this seeming contradiction. We varied parametrically the RF-correlation-based bias for excitatory and inhibitory synapses and found that a moderate bias of excitatory neurons to synapse onto other neurons with correlated receptive fields, and a weaker bias of inhibitory neurons to synapse onto other neurons with anti-correlated receptive fields can explain the conductance input, the postsynaptic membrane potential, and the spike train dynamics under both stimulation paradigms. This computational study shows that the same structural model can reproduce the functional diversity of visual processing observed during different visual contexts.
Acknowledgments

This work was supported by National Institutes of Health Grant R01EY027205 (D.C.), National Science Foundation Grant GRFP DGE-1321851 (M.M.T.), NEI Vision Training Grant T32 EY007035-38 (M.M.T., D.C.), CNRS (A.D., Y.F.), European Community (Human Brain Project, H2020-720270 and H2020-785907; A.D.), the French National Research Agency (ANR-Horizontal-V1; ANR-17-CE37-0006; Y.F.), through project Improvement of internationalization in the field of research and development (CZ.02.2.69/0.0/0.0/17_050/0008466) at Charles University (J.A.), Charles University Primus program No. 20/MED/006, France-US fellowship from the Chateaubriand Foundation (M.M.T), and the ANR 353 PARADOX, and the ICODE excellence network. M.M.T was hosted by the European Institute of Theoretical Neuroscience (EITN) in Paris. We would like to thank Madineh Sedigh-Sarvestani and Larry Palmer for valuable suggestions and insightful discussion.

Significance Statement

Identifying generic connectivity motives in cortical circuitry encoding for specific functions is crucial for understanding the computations implemented in cortex. Indirect evidence points to correlation-based biases in connectivity pattern in V1 of higher mammals, whereby excitatory and inhibitory neurons preferentially synapse onto neurons respectively with correlated and anti-correlated receptive fields. A recent intracellular study questions this “push-pull” hypothesis, failing to find spatial anti-correlation patterns between excitation and inhibition across the receptive field. We present here a spiking model of V1 that integrates relevant anatomical and physiological constraints, and shows that a more versatile motif of correlation-based connectivity with selectively tuned excitation and broadened inhibition is sufficient to account for the diversity of functional descriptions obtained for different classes of stimuli.
**Introduction**

Hubel & Wiesel (1962) hypothesized that orientation selectivity emerges due to specific alignment of feedforward excitatory ON/OFF inputs from lateral geniculate nucleus (LGN) onto contrast-matched RF sub-fields of layer 4 (L4) neurons in V1. Simultaneous recordings of LGN and V1 connected pairs provided support for this wiring rule (Toyama et al. 1981; Alonso et al., 2001; Sedigh-Sarvestani, 2017). The principles of intra-cortical connectivity are less clear. In cat L4 simple cells, intracellular recordings revealed the characteristic “push-pull” behavior, whereby presentation of sign-matched stimulus in RFs sub-field evokes depolarization (push), while stimulus of opposite polarity evokes hyperpolarization (pull) of the membrane potential (Ferster, 1986; Hirsch et al. 1998). These observations lead to the hypothesis of anti-correlated arrangement of excitation vs. inhibition across the RF space, whereby stimulation of a RF sub-field with sign-matched stimulus evokes predominantly excitation, whereas opposite stimulus polarity evokes predominantly inhibition (Troyer et al., 1998; Ferster and Miller, 2000; Hirsch, 2003;).

A way for such an anti-correlated interaction between excitation and inhibition across the RF to be implemented is for the connectivity in L4 cells itself to adhere to a “push-pull” schema: excitatory and inhibitory neurons preferentially connect other neurons with, respectively, correlated and anti-correlated RFs (Troyer et al., 1998; Anderson et al., 2000; Ferster and Miller, 2000; Hirsch, 2003). Such push-pull connectivity has been examined in numerous computational models, either on its own (Troyer et al., 1998; Ferster and Miller, 2000; Kremkow et al., 2016), or in combination with untuned inhibition (Lauritzen and Miller, 2003; Teich and Qian, 2006;).

Both extra- and intra-cellular studies support a correlation-based rule for excitatory connectivity showing that co-oriented pairs of V1 cells have a higher probability of connection (Michalski et al., 1983; Monier et al., 2003; Ko et al., 2011; Denman and Contreras, 2014; Lee et al., 2016; Wilson et al. 2016), and that excitatory neurons preferentially connect other neurons with correlated (in-phase) RFs (Cossell et al., 2015).
The anti-correlation rule for inhibition would predict maximal inhibitory synaptic inputs from inhibitory cells whose RF is co-tuned in orientation but in anti-phase. Such anti-phase arrangement of excitation and inhibition is supported by intracellular studies showing that drifting grating elicits excitatory and inhibitory inputs that are in anti-phase (Ferster, 1988; Anderson et al., 2000; Monier et al., 2003, 2008; Tan et al., 2011). Furthermore, inhibitory inputs on excitatory cells in ferret layer 2/3 are biased towards opposite direction of movement (Wilson et al. 2018).

However, a recent study revealed that, when the RFs of simple cells is mapped with optimally oriented flashed bars, inhibition is evoked broadly and is co-localized across the RF irrespective of stimulus polarity (Taylor et al., 2018), questioning the idea of a strict anti-phase arrangement of excitation and inhibition across the space of RF. Models utilizing push-pull connectivity (whether on its own or in combination with untuned inhibition) do not account for these findings, as they predict the inhibition to elicit spatially offset peaks across the RF field for bars of opposite polarity. Models relying solely on phase-unspecific inhibition (Ben-Yishai et al. 1995; Douglas et al. 1995; Carandini and Ringach 1997;) cannot explain the anti-phase relationship due to sinusoidal grating stimuli. Thus, we currently lack mechanistic explanation consistent with the key experimental findings on the stimulus evoked interaction between excitation and inhibition in V1.

To resolve this issue, we have integrated all the key experimental findings into a unified comprehensive model of cat V1 granular layer. Our model reproduces both the anti-phase behavior of excitation vs. inhibition for drifting gratings and the in-phase excitation but broad inhibition for flashed bars. Exploration of the connectivity parametric space of excitatory vs inhibitory tuning selectivity shows that a wide range of parametrizations induce neural dynamics comparable to those in cat simple cells. This suggests that elementary functional properties such as orientation tuning, contrast invariance, and push-pull organization can arise robustly across a variety of connectivity schemes. However, the underlying conductance dynamics change significantly across this parameter space, demonstrating that very diverse motifs of E/I interactions can underlie the same functional properties (see experimental evidence in (Baudot et al., 2013)).

Materials and Methods
The basic architecture of the model utilized in this study is derived from a recent data-driven model of V1 (Antolik et al., 2019), based on a consensus of the experimental electrophysiological and anatomical literature in cat area 17. However, here we have restricted simulations to the thalamo-cortical input recipient simple cells originating from layer 4 (the original model comprised also complex cells in supra-granular layers 2/3). Below we present a summarized description of the model used here and a full list of modifications from the original. We refer the reader to the parent model for further details (Antolik et al., 2019). The model has been implemented in the Mozaik framework (Antolík and Davison, 2013), while the NEST simulator (version 2.1.1; Gewaltig and Diesmann, 2007) was used as the back-end for all simulations described in this paper.

**V1 model**

The cortical model corresponds to a 1.2 x 1.2 mm patch of L4 of cat primary visual cortex centered at 5 degrees of visual field eccentricity (see Fig 2). It contains 3600 cortical neurons and ~4 million synapses which are driven by spikes representing LGN input. We simulated a population of excitatory neurons (corresponding to spiny stellate neurons in Layer 4) and one population of inhibitory neurons (representing all subtypes of inhibitory interneurons) in a 4:1 ratio (Gabbott and Somogyi, 1986; Beaulieu et al., 1992; Markram et al., 2004).

All neurons were modeled as single-compartment integrate and fire units. Specifically, we used the adaptive exponential (AdExp) integrate-and-fire model, which is computationally efficient, offers a broad range of firing dynamics, and offers more realistic membrane potential time courses than simpler integrate and fire schemes (Brette and Gerstner, 2005; Naud et al., 2008; Destexhe, 2009). We set the membrane resistance of all cortical neurons $R_m$ to 250MΩ (Monier et al., 2008). We set the membrane time constant of excitatory neurons to 20 ms, and of inhibitory neurons to 10 ms, close to values observed experimentally in cat V1 in vivo (Monier et al., 2008). The refractory period during which the membrane potential is held at -72 mV for all simulations was set to 2 ms and 0.5 ms for excitatory and inhibitory neurons, respectively. Overall these neural parameter differences between excitatory and inhibitory neurons reflect the experimentally observed greater excitability and higher maximum sustained firing rates of inhibitory neurons (McCormick et al. 1985; Contreras and Palmer, 2003). The excitatory and inhibitory reversal
potential $E_{\text{exc}}$ and $E_{\text{inh}}$ were set to 0 mV and -80 mV, respectively, in accordance with values observed experimentally (Monier et al., 2008). The threshold slope factor and the adaptation time constant of the AdExp model were set to 2.0 mV and 88 ms, respectively, for all neurons (Naud et al., 2008). We did not consider the sub-threshold and spike-triggered adaptation in this study, and thus the corresponding parameters of the AdExp model were set to zero.

Retino-thalamic pathway model

We did not explicitly model the retinal circuitry but instead used the retinal spike pattern, predicted by the center-surround model of receptive fields (RFs), to feed LGN neurons (Fig. 2). The centers of both ON and OFF LGN neuron RFs were uniformly randomly distributed in the visual space, with density of 100 neurons per square degree. Each LGN neuron had a spatiotemporal receptive field, with a difference-of-Gaussians spatial profile and a bi-phasic temporal profile defined by a difference-of-Gamma-functions. The exact spatial and temporal parameters have been adopted from Allen and Freeman (Allen and Freeman, 2006).

To obtain the spiking output of a given LGN neuron, the visual stimulus was sampled into 7 ms frames and convolved with the cell’s spatiotemporal receptive field. In addition, saturation of the LGN responses with respect to local contrast and luminance was modeled (Papaioannou and White, 1972; Bonin et al., 2005). For simplicity, the local luminance $ll$ was calculated as the mean of luminance values, and local contrast $lc$ as the standard deviation of the luminance values sampled within the RF of the given neuron. The response of the linear receptive field was separated into a DC (luminance) component $r_l$ and a contrast component $r_c$. The saturation of the two components is modeled with two Naka-Rushton functions: $\frac{r_l}{1 + \frac{ll}{\alpha}}$ and $\frac{r_c}{1 + \frac{lc}{\beta}}$ where $\alpha$ is the gain and $\beta$ is the saturation parameter of the corresponding component. The parameters $\alpha$ and $\beta$ were empirically adjusted to obtain luminance and contrast response curves whose saturation point and level are within the range of those observed experimentally (Papaioannou and White, 1972; Bonin et al., 2005).

The resulting luminance and contrast temporal traces were then summed and injected into integrate-and-fire (IAF) neurons as a current, inducing stimulus-dependent spiking responses. In
addition to the stimulus-dependent drive, neurons were also injected with white noise current. The magnitude and variance of this current was such that neurons fire at an average rate of 10 Hz in the no stimulus condition (Troyer et al., 1998). This artificially elicited spontaneous discharge was calibrated to reproduce experimentally observed spontaneous rates.

Thalamo-cortical model pathway

All cortical neurons in the model receive connections from the model LGN. For each neuron, the spatial pattern of thalamo-cortical connectivity was determined by a Gabor distribution, inducing the elementary RF properties in cortical neurons (Jones and Palmer, 1987a; Troyer et al., 1998) (see Fig. 2):

\begin{align*}
g(x, y, \ldots, \ldots) &= \exp\left(\frac{x'^2 + y'^2}{2} + \frac{i(2px' + y)}{\sigma^2}\right) \\
x' &= x \cos \psi + y \sin \psi \\
y' &= -x \sin \psi + y \cos \psi
\end{align*}

For individual neurons the orientation \(\theta\), phase \(\psi\), size \(\sigma\), spatial frequency \(\lambda\), and aspect ratio \(\gamma\) of the Gabor distribution were selected as follows. To induce functional organization in the model, we used an existing model of stimulus-dependent orientation map development (Antolík and Bednar, 2011) that utilizes Hebbian learning to compute a stabilized link map that conditions an orientation map. Such pre-computed orientation map, corresponding to the 1.2 x 1.2 mm of simulated cortical area, was overlaid onto the modeled cortical surface, thereby assigning each neuron an orientation preference \(\theta\). The phase \(\psi\) of the Gabor distribution was assigned randomly. For the sake of simplicity, the remaining parameters were set to constant values, matching the average of measurements in cat V1 RFs located in the para-foveal area (Jones and Palmer, 1987b). Specifically, the size \(\sigma\) was set to 0.15 degrees of visual field, the spatial frequency \(\lambda\) to 0.8 cycles per degree and the aspect ratio \(\gamma\) to 0.57 (Pei et al., 1994). For any given neuron, the thalamo-cortical connections were generated by overlaying its Gabor template over the model’s sheets of LGN ON and OFF neurons, then drawing the connections randomly (with replacement) from the resulting probability distribution over the LGN neurons. Each cortical neuron received between 60 and 180 (drawn uniformly within these bounds) thalamo-cortical synapses (da Costa and Martin, 2011).
Cortico-cortical connectivity

We modelled 1000 cortico-cortical synaptic inputs per modeled excitatory cell. Inhibitory neurons received 20% fewer synapses than excitatory neurons to account for their smaller size, but otherwise synapses were formed proportionally to the two cell type densities. The synapses were drawn probabilistically with replacement. The geometry of the cortico-cortical connectivity was determined based on two main principles: the connection probability falls off with increasing cortical distance between neurons (Budd and Kisvarday, 2001; Buzás et al., 2006; Stepanyants et al., 2009) (see Fig. 2), and connections are formed preferentially between neurons with similar functional properties (Ko et al., 2011). The two principles were each expressed as a connection-probability density function, then multiplied and re-normalized to obtain the final connection probability profiles, from which the actual cortico-cortical synapses were drawn. The following sections describe how the two probability density profiles of connectivity were obtained.

Spatial extent of local intra-cortical connectivity

The exact parameters of the spatial extent of the model local connectivity were established based on a re-analysis of data from cat published in Stepanyants (Stepanyants et al., 2008). The details of the analysis can be found in the publication of the parent model (Antolik et al. 2019), but briefly, the probability of potential connectivity data from Stepanyants et al. were averaged to obtain a 1D profile expressing probability of connectivity as a function of distance between two cortical neurons of specified type (excitatory or inhibitory).

To obtain a parametric representation of the distance connectivity profiles we fitted them with a zero mean hyperbolic distribution:

\[ pdf(x) = \exp(-\alpha \sqrt{x^2 + \theta^2}) \]

The resulting values of the parameters of the fitted hyperbolic distributions for all combinations of pre- and post- synaptic neuron types, which were used to generate the local connectivity distance dependent profiles in the model, can be found in Table 1.

<table>
<thead>
<tr>
<th></th>
<th>( \alpha )</th>
<th>( \Theta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>L4 Exc</td>
<td>L4 Inh</td>
<td>L4 Exc</td>
</tr>
</tbody>
</table>
Table 1: The parameters of hyperbolic profiles of potential connectivity derived from connectivity data published in Stepanyants et al. (2008) used to generate the distance dependent profile of local connectivity in the model.

<table>
<thead>
<tr>
<th></th>
<th>Exc</th>
<th>Inh</th>
</tr>
</thead>
<tbody>
<tr>
<td>L4</td>
<td>0.0139</td>
<td>0.0126</td>
</tr>
<tr>
<td></td>
<td>0.0148</td>
<td>0.0119</td>
</tr>
<tr>
<td></td>
<td>207.7</td>
<td>237.5</td>
</tr>
<tr>
<td></td>
<td>191.8</td>
<td>256.4</td>
</tr>
</tbody>
</table>

Functionally specific connectivity

Recent studies of local connectivity in mice have shown that this iso-iso preference might be a general connectome feature independent of the species-specific presence of orientation preference maps, since, in spite of the “salt and pepper” V1 organization, local connections in rodent also showed a weak bias towards connecting neurons with similar receptive field properties (Cossell et al., 2015) or similar orientation (Denman and Contreras, 2014). Finally, the dominant anti-phase relationship between excitatory and inhibitory conductances in cat V1 simple cells (Anderson et al., 2000; Monier et al., 2008; Tan et al., 2011) has traditionally been interpreted as further evidence for functionally specific inputs (Anderson et al., 2000; Kremkow et al., 2016). Overall, these results point to a moderate tendency of excitatory neurons towards connecting nearby neurons of similar receptive field properties, although this bias increases somewhat for more distant post-synaptic neurons. In this study we will restrict our investigation to the local connectivity.

Due to the lack of clarity and specificity of experimental data on the functional connectivity in V1, we utilize previously hypothesized schemes of functional connectivity and adapt them to be compatible with the above experimental findings. Among cortical neurons we assume push-pull connectivity (Troyer et al., 1998) (see Fig. 2) of varying strength of the functional biases. For each pair of cortical neurons the correlation $c$ between their RFs was calculated. The connectivity likelihood for a given pair of neurons is given by

$$\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{(c-u)^2}{2\sigma^2}\right)$$

where $\mu = 1$ if the pre-synaptic neuron is excitatory or -1 if inhibitory. The exact width parameter $\sigma$, determining the degree of functional bias, was varied throughout the study (within the range from 0.6 to 4) and will be always specified for each virtual experiment described in the Results.
Synapses

Synaptic inputs were modeled as transient conductance changes, with exponential decay with time-constant $\tau_e = 1.1$ ms for excitatory synapses and constant $\tau_i = 1.9$ ms for inhibitory synapses. There exists relatively little data on the exact strength of synapses between different neural types and layers. The synaptic weights were thus selected to achieve an overall balance between excitation and inhibition that supports reasonable levels of both spontaneous and evoked activity, while being compatible with the limited physiological findings. Specifically, we have set each excitatory synapses onto inhibitory neuron to 2.4 nS, while all remaining synapses in the model were set to 1.4 nS. But note that the connection generation algorithm allows multiple synapses to form between pairs of neurons. We have also modeled synaptic depression for thalamo-cortical, and excitatory cortico-cortical synapses (Abbott et al., 1997) using the model of Markram (Markram et al., 1998). We did not model short-term plasticity for inhibitory synapses as it is not well studied (but see review in (Kriepke and Froemke, 2017)). For all excitatory synapses we assume parameters corresponding to moderate depression ($U=0.75$, $\tau_{rec} = 150$, $\tau_{psc} = 3.0$, and $\tau_{fac} = 0$) (Markram et al., 1998).

Delays

The delays in the feedforward thalamocortical pathway were drawn from a uniform distribution bounded between 1.4 and 2.4 ms. For all intra-cortical connectivity, a distance-dependent delay with propagation constant of 0.3 m/s (Bringuier et al., 1999) was used, corresponding to the slow propagation of action potentials along the intra-areal (lateral) un-myelinated processes. Furthermore, we have included a constant additive factor in all synaptic delays, specifically 1.4 ms to excitatory to excitatory synapses, 0.5 ms to excitatory to inhibitory synapses, 1.4 ms to inhibitory to excitatory synapses, and 1.4 ms to inhibitory to inhibitory synapses, in line with experimental observations (Ohana et al., 2012).

Visual stimuli

We used two types of visual stimuli: flashed stationary bars and drifting gratings. The bars were vertically oriented, matching the preferred orientation of the 30 excitatory cells whose responses are quantified in this manuscript, and either bright or dark contrast (100%) at varying locations across the visual field. Sinusoidally-varying contrast gratings (spatial frequency: 0.8 cycles per
degree; temporal frequency: 2 Hz; 100% contrast) were presented at various orientations, including optimal (0 degrees) and orthogonal (90 degrees) for the “recorded” cells.

Data analysis

Analysis of the model output was performed using customized code in MATLAB. Responses (Vm, gE, and gI) from multiple repeated trials were averaged for each neuron. We measured peak responses (V_m depolarization from rest or increase in synaptic conductance) as the average within a 50 ms window across all bar positions for each cell, allowing for some variability in the peak response times between bar positions. For visualization purposes, these responses were fit to a Gaussian distribution for each RF subregion. Visual space coordinates have been re-scaled and translated such, that 0 corresponded to the peak of the V_m response to bright and 1 to the peak V_m response to dark, such that the distance between the centers of adjacent V_m-derived RF subregions defined 1 receptive field unit or RFU.

In the experimental studies that form the basis of the present model (Monier et al, 2003; Taylor et al, 2018), excitatory and inhibitory synaptic conductances were estimated using current clamp. A full replay of such current clamp experiments would be extremely resource demanding in our large-scale detailed model, and in the context of the extensive parameter search that we performed. But, rather than estimating indirectly the underlying simulated synaptic conductances from voltage traces recorded under current clamp, the full network simulation allows to record them directly. To compare our virtual recordings with those estimated experimentally, we simulated current clamp responses to drifting sinusoidal gratings and estimated the underlying conductances from the voltage traces. On the whole, the estimated synaptic conductances (Fig 3A) are found similar to the underlying simulated model synaptic conductances (Fig 3B), but with somewhat smoother variations. The level of variability of the estimated conductance traces in the simulations matches that of experimental data in cat V1 (Fig 3C showing data from Baudot et al. (2013)). The key measure of our model (the temporal correlation between the excitatory and inhibitory conductance traces) remains the same, whether based on virtual recordings or experimental current clamp measurements (Fig 4D). We thus conclude, that the use of simulated conductance traces in this computational study is fully justified.
To quantify the selectivity to stimulus orientation, we calculated the half-width at half-height (HWHH) as in (Alitto, 2004), using a Gaussian fit to the orientation tuning data:

$$R(\phi) = \beta + \alpha \exp\left(\frac{\phi - \phi_{pref}}{2\sigma^2}\right)$$

where \( R \) is the spiking response of the given neuron to sinusoidal grating with orientation \( \phi \), \( \phi_{pref} \) is the preferred orientation of the given neuron, \( \sigma \) is the width of the tuning, \( \beta \) is the baseline activity and \( \alpha \) a scale factor. Neurons for which successful fit was not achieved (MSE > 30% of tuning curve variance) were excluded from further analysis. HWHH was then calculated as \( \sqrt{2 \ln(2)} \sigma \). Model parameter combinations (figure 6) where more the 20% of neurons had to be excluded in this way were excluded from orientation tuning analysis (black color in figure 6B).

**Results**

**Simple cell characteristics and classification criteria**

In order to evaluate the physiological fidelity of our model, we first established a set of criteria for which it has to match experimental data. Simple cells in V1 L4 have many well-characterized features in response to visual stimuli, and while many of these have been demonstrated in a previous version of this model (Antolik et al., 2019), here we focused on responses to drifting gratings and to flashed bars (Fig. 1A). When stimulated with drifting gratings, simple cells respond with maximal Vm depolarization and spike output to a preferred orientation, and this orientation selectivity is invariant to contrast (Fig. 1B top). The Vm and firing rate are temporally modulated in phase with the grating (Fig. 1A, B bottom). Optimally oriented bright bars flashed in ON subregions evoke depolarization and an increase in firing rate, while dark bars in the same ON subregion evoke hyperpolarization and an absence of spikes (Fig. 1C). This characteristic behavior in simple cell RFs, observed at the Vm and spike levels (Fig. 1C, bottom) are classically interpreted as the landmarks of “push-pull” connectivity. The measure of a negative (anti-correlated) spatial correlation of the Vm response to light vs dark bar stimuli is used here to characterize the “simpleness” of V1 receptive fields.

When it comes to synaptic conductances under these stimulation protocols, the literature has produced seemingly conflicting results. In response to drifting gratings at the optimal orientation, the current-clamp observation of spatial anti-correlation between the temporal patterns of
dominant excitation and inhibition driven by opponent contrast have led some authors to infer that
the “push-pull” organization observed for Vm and spikes extends into the conductance domain
and implies that gE and gI are separated in space, or spatially anti-correlated (Fig 1E). This spatial
segregation predicts that stimulation of a RF subregion with signed matched flashed bar (bright
bar on ON subregion or dark bar on OFF subregion) will evoke mostly gE, while the opposite
contrast bar will evoke mostly gI. However, experimental evidence has suggested, that when
measured with flashed bars, the spatial footprint of gI is broad across the entire RF, with co-
localized peaks for the two stimulus polarities (Taylor et al 2018; Fig. 1F; see also figure 2 in
Borg-Graham et al., 1998) which is in agreement with voltage-clamp measurements of spatial
overlap of excitation and inhibition in simple cell RFs (Borg-Graham et al., 1998; Monier et al.,
2008). This suggests that while gE responses to bright versus dark bars across the RF are anti-
correlated over space, gI responses to bright versus dark bars have a positive correlation. We used
these observations to guide our exploration of connectivity parameter space in the model,
described next.

Model construction
We modified a model of the thalamo-cortical LGN-L4 V1 circuit (Antolik et al., 2019), whose
details are described in the Methods and schematized in Figure 2. Briefly, the model contains
center-surround LGN cells that are connected to a network of excitatory and inhibitory neurons in
L4, all of which have simple RFs. Feedforward connections from the LGN follow RF correlation
rules (Hubel and Wiesel, 1962; Alonso et al., 2001; Sedigh-Sarvestani et al., 2017; Fig. 2D), while
local connections between excitatory and inhibitory cells were determined based on a combination
of distance between the cell pair (Fig. 2A) and their RF correlations (Fig. 2C,E). The connectivity
shown in this first iteration of the model, that best reproduces the experimental data, can be
described as a push-pull like organization: connections from excitatory cells to other neurons were
more likely if the RFs were highly correlated, and much less likely if they were anti-correlated
(Fig. 2C and 2E, red line), while connections from inhibitory cells were only mildly biased towards
anti-correlated RFs (Fig. 2C and 2E, blue line). Later in the article, we further explore a subset of
this intracortical connectivity parameter space.
We characterized the behavior of this network in response to two sets of visual stimuli: flashed stationary bright and dark bars at the neurons’ preferred orientation, at different locations across the RF, and full-field drifting gratings at multiple orientations. We recorded the spiking output, Vm, gE, and gI from 30 excitatory cells in L4 in response to these stimuli.

**Broad inhibition in response to flashed bars**

We first tested the model’s response to flashed bars of bright or dark contrast presented across the width of simple cell RFs, at their preferred orientation. Vm and spike responses to a single bright or dark bar presented within the ON subregion of an example cell in the network (Fig. 4A, left) show the expected push-pull relationship: a bright bar evoked depolarization and spikes, while a dark bar evoked hyperpolarization and spiking suppression. The underlying synaptic conductances also matched experimental data, with both stimuli evoking large magnitudes of gI increase, but only the bright bar evoking a significant increase in gE (Fig. 4A, right).

We plotted the peak evoked values of Vm, gE, and gI to each bar position (peak values indicated by circles in Fig. 4A) over space in Figure 4B. The Vm responses to bright versus dark bars over space were anti-correlated (Fig. 4B, top; Vm spatial correlation = -0.95), reflecting the cell’s RF and following the push-pull characterization of simple cells. The gE responses also had a negative correlation (Fig. 4B, middle; gE spatial correlation = -0.23). The gI responses, however, produced a broad footprint in response to both bright and dark stimuli, resulting in positive correlation over space (Fig. 4B, bottom; gI spatial correlation = 0.93). The correlations were performed on the peak evoked values in response to bright and dark for each position, not on the Gaussian fits to the data, which are shown in the figure for clarity.

The example cell was a representative sample of the population of 30 excitatory cells analyzed in the simulation. The distributions of spatial correlation values for Vm, gE, and gI are plotted in Figure 4C for both the cells in this simulation (black) and for the experimental dataset (purple) from Taylor et al. (2018). The spatial correlation values for Vm in the simulation are significantly more negative than those from the data (Fig. 4C, top; simulation median Vm spatial correlation = -0.93, p = 1.7e-6, Wilcoxon signed rank test; experiment median Vm spatial correlation = -0.49, p = 1.4e-3, Wilcoxon signed rank test). The stronger anti-correlation in the simulation arises from
strong hyperpolarizing responses to bars flashed in RF subregion of opposite polarity, which reflects the more depolarized Vrest of the cells in the simulation relative to the experimental neurons. In any case, the anti-correlated Vm responses to bright vs dark bars show that the Vm of neurons in L4 behave in a push-pull manner typical of simple cells in L4.

The peak gE responses to bright and dark bars were spatially anti-correlated, while the peak gI responses were strongly spatially correlated (Fig. 4C; simulation median gE spatial correlation = -.35, p = 4.7e-6, Wilcoxon signed rank test; simulation median gI spatial correlation = +.81, p = 1.6e-6, Wilcoxon signed rank test). The simulation matches the data very well for the spatial correlation values of gE and gI (experiment median gE spatial correlation = -.43, p = 0.017, Wilcoxon signed rank test; experiment median gI spatial correlation = +.71, p = 2.9e-4, Wilcoxon signed rank test; calculated from data in Taylor et al., 2018).

In summary, despite only a weak bias towards antiphase connectivity between inhibitory and excitatory neurons, this network produced spatially restricted excitation and broad inhibition in response to flashed bars across the RF, consistent with experimental findings (Taylor et al., 2018).

**Antiphase inhibition in response to drifting gratings**

Having replicated the experimental findings of spatially broad inhibition evoked by flashed bars across the width of simple cell RFs, we asked whether this network would also produce anti-phase (temporally anti-correlated) modulation of excitatory and inhibitory conductances in response to drifting sinusoidal gratings. A drifting grating at the preferred orientation produced modulated spike output and Vm (Fig. 5A), whereby spikes occur at the peaks of the modulated Vm trace, when the RF subregions are aligned with the matching luminance regions of the drifting sinusoidal grating stimulus. The conductance responses to the preferred orientation grating were noisy but did appear to be modulated in temporal antiphase (Fig. 5B, top). By smoothing the conductance traces, we revealed a clear anti-correlated relationship between gE and gI (Fig. 5B, bottom), with a correlation value of -0.74. The population of cells analyzed in the simulation showed a strongly negative distribution of temporal correlations, with a median at -0.61 (Fig. 5C; n=30 cells, p = 5.3e-4, Wilcoxon signed rank test).
In summary, a strong bias of excitatory connectivity and a weak bias of inhibitory connectivity among anti-correlated neurons in L4 produced a strong anti-phase relationship between excitation and inhibition, consistent with experimental findings (Ferster, 1988; Anderson et al., 2000; Monier et al., 2003; Tan et al., 2011).

Exploring parameter space of intracortical connectivity

The simulation shown in Figures 4&5 had specific values determining the RF-correlation-based likelihood of connectivity between any pair of neurons in L4 (Fig. 2E; $\sigma_E = 1, \sigma_I = 2$). As shown above, this simulation reproduced the temporal and spatial conductance dynamics shown in previous experiments from different labs. We next assessed the importance of these intracortical connectivity rules in determining the behavior of the network, by exploring the parameter space of $\sigma_E$ and $\sigma_I$. We independently varied $\sigma_E$ and $\sigma_I$ from 0.6 to 4.0, with lower values representing less selective connections and vice versa (Fig. 6A). Throughout Figure 6, 2D-plots illustrate inhibitory connectivity in the abscissa, and excitatory connectivity in the ordinates. The sigmoidal scale functions of probability of connection as a function of RF correlation for the four extreme points in parameter space are illustrated on the right (Fig. 6A, a-d). We note that this parameter space explores the classical “push-pull” connectivity scheme, which at its extreme would be characterized by the lower left portion (Fig. 6A, example ‘c’). We explored from this extreme to almost no bias, independently for excitatory and inhibitory connections.

We first explored the role of connection selectivity on the spiking and Vm characteristics of L4 neurons in response to drifting gratings and flashed bars (Fig. 6B). We quantified the median orientation tuning width measured as half-width at half height (HWHH), of the neurons’ spiking response to drifting gratings (Fig. 6B, left), the change in HWHH between low- and high-contrast gratings (Fig. 6B, center), and the spatial correlation of the Vm responses to bright and dark bars at the neurons’ preferred orientation (Fig. 6B, right).

Overall, we found only a small effect, with the exception of a section of parameter space where excitatory and especially inhibitory connections were highly selective. This was reflected in a less physiologically realistic behavior: for example, orientation tuning curves of most neurons could not be well fit by a Gaussian function (see Fig 6B, black squares). Immediately surrounding this...
area in parameter space, the neurons exhibited large contrast dependent changes in orientation
tuning (Fig. 6B, center, yellow squares), indicating that highly selective inhibitory connections
preclude contrast-invariant orientation tuning. This portion of parameter space also featured
neurons with less anti-correlated Vm in response to flashed bars (Fig. 6B, right). Closer inspection
of population dynamics of models in this subdomain revealed unstable, often oscillatory behavior
explaining the breakdown of elementary functional features reported in Figure 6B. In summary,
we conclude that, with the exception of highly selective connection schemes, a relatively wide
range of connectivity parameters are sufficient to reproduce experimentally determined spiking
and Vm characteristics of simple cells in L4 of V1.

We next examined the differential effect of the selectivity of the excitatory and inhibitory
connectivity profiles on the relationship between excitatory and inhibitory input conductances. In
contrast to spiking and Vm responses (Fig. 6B), the temporal and spatial correlations of synaptic
conductances showed strong variations (Fig. 6C). However, gE and gI responses to drifting
gratings remained robustly anti-correlated even when the specificity of inhibitory connectivity was
decreased, approaching zero or positive values for extremely broad tuning values (Fig. 6C left, Fig
7AB). This striking observation occurred despite the decreasing magnitude of the phase dependent
modulation of the inhibitory synaptic input (Fig 7AB). Thus, even weakly biased inhibitory
connectivity can give rise to experimentally observed levels of temporal “push-pull” between
excitatory and inhibitory conductances in response to drifting gratings. We also observed a
breakdown of the anti-correlation between the gE and gI when selectivity of the inhibitory
connections was very high (Fig. 6C, bottom edge). But it is important to point out that in this region
of the parameter space, the model dynamics were unstable, and thus the correlation measure of the
push-pull was dominated by large stimulus-independent run-away events characterized by
concomitant rise and fall of excitation and inhibition throughout the cortical network.

Increasing the selectivity of inhibitory connections thus decreased the spatial correlations between
gI responses to bright versus dark bars (Fig. 6C right), and at the same time decreased the temporal
anti-correlation between excitatory and inhibitory conductances in response to gratings (Fig. 6C
left). Crucially, the temporal anti-correlation of gE and gI to gratings broke down only at very
broad selectivity of inhibitory connectivity (σI>~3), while the inhibition evoked by bars became
broad and overlapping already at moderate inhibitory connection selectivity values ($\sigma_I > 1.2$). Thus, one can find regions of the parameter space where the model behaves in line with experimental evidence for both drifting grating and flashed bars (Fig 6D).

The question remains, how can there exist model parametrizations where selectivity of inhibitory connections is sufficiently broad to induce overlapping inhibition across space in response to flashed bars, but sufficiently specific to induce the temporally anti-correlated excitation and inhibition observed in response to drifting gratings? We hypothesize that the explanation lies in the difference of temporal dynamics of the input drive respectively fed by drifting grating vs. flashed bar stimulus. A flashed bar is an abrupt step stimulus that evokes strong initial response (onset) followed by rapid decline to a lower sustained level (adaptation). This justifies why, as most experimenters do, Taylor et al. (2018) focused their analysis only on the early 50ms of the response unaffected by the adaptation. The early response is, however, dominated by the stimulus onset dynamics, characterized by concomitant rise of excitation and inhibition (Fig 2 in Borg-Graham et al, 1998). Such a short flashed stimulation protocol does not allow for enough time for the cortical-network to self-stabilize into a dynamic regime dictated by the biases in the functional circuitry, unless the biases are very strong, as revealed in our parameter search (Fig 6). This non-stationarity hypothesis is supported by the experimental evidence showing sharpening of orientation tuning of spike response over time from stimulus onset (Ringach et al. 1997, Schummers et al. 2007). This is in contrast to steady-state response to drifting grating stimuli which engages the cortical network in a steady slowly changing fashion, that reduces adaption over period of seconds. This delayed and long integration time allows the network dynamics to settle in a dynamic regime dictated by its connectivity biases, making even weak biases in the cortical connectivity robustly functionally impacting on both excitatory and inhibitory input conductances. Incidentally, the on-set dynamics with concomitant increase in excitation and inhibition are also present for the drifting grating stimuli, both in our model (e.g. see Fig 5B, or Fig 5&10 in Antolik et al. 2019) or in cat (Fig 5 in Monier et al. 2008). But due to the short duration of the on-set dynamics, this has little influence on the measure of correlation between excitation and inhibition in response to the long stimulation periods covering several cycles of drifting gratings.
The parameter search also revealed that the selectivity of excitatory and inhibitory connections did not act independently, as is evidenced by a moderate slope in the correlation levels between $g_I$ responses to bright versus dark bars (Fig 6C, right). This is illustrated in the two model parametrizations shown in figure 7CD, and marked by corresponding letters in figure 6D. Both parametrizations correspond to models with the same inhibitory but different excitatory selectivity. Of the two models, the model with more selective excitation showed lower correlations between $g_I$ to bright versus dark bars (Fig 7C), compared to the model with broader excitatory selectivity (Fig 7D). This can be explained by the fact that the excitatory selectivity parameter dictates also the selectivity of connections from excitatory to inhibitory neurons, which, however, influence the strength of the phase dependent modulation of the output of inhibitory neurons.

An important property of the model that emerged from the parameter search was the higher sensitivity to changes in inhibition than excitation. Figure 6B shows that the transition of the model dynamics to unstable happened mostly along the bottom edge of the plots, indicating that smaller change in selectivity of inhibitory than excitatory connectivity can push the model into the unstable state. Similarly, the temporal correlation between $g_E$ and $g_I$ (Fig 6C, left panel) change more rapidly along the inhibitory selectivity axis. Such greater sensitivity to changes in inhibition are likely due to the thalamo-cortical excitatory connections imposing strong bias on excitation in the cortex. On the other hand, there are no thalamo-cortical inhibitory connections. This means that the same relative change to the excitatory cortical circuitry has lesser relative impact on the overall excitation in cortex, then equal change to inhibitory cortical circuitry has on the overall inhibitory cortical dynamics. This hypothesis is consistent with the finding that the correlation (across RF space) between synaptic conductance evoked by ON and OFF bar stimuli exhibits greater magnitude of change when excitation selectivity varies (Fig 6C, middle) than when inhibition selectivity varies (Fig 6C, right).

Taken together, the parameter search showed that experimentally observed simple cell characteristics of spiking, $V_m$, and synaptic conductances in response to both drifting gratings and flashed bars arise when excitatory connections are relatively selective (pre- and post-synaptic neurons have correlated RFs) and when inhibitory connections are only mildly selective (pre- and
post-synaptic neurons tend to have slightly anti-correlated RFs). This area of the parameter space is illustrated in Figure 6D.

**Response to flashed bright or dark squares**

The Hirsch et al. (1998) experiments have provided grounding support to the push-pull concept, by showing that the suppression (pull) evoked by stimulus flashed in the RF subfield of mismatched polarity is due to increase of inhibition, rather than just a withdrawal of excitation. As expected, our model reproduces the characteristic push-pull response to flashed squares (Hirsch et al, 1998), just as it does to flashed bar stimulus (Taylor et al, 2018). Our simulations show Vm depolarizations for flashed point-like stimuli whose contrast (light/dark) matches the polarity of the RF subregion (Fig 8AB), whereas opponent hyperpolarization is induced by a stimulus of the opposite contrast in the same RF location (Fig 8CD).

Whereas evoked hyperpolarization grows with the intensity of intracellularly injected positive current (Fig 8E), its amplitude diminishes (Fig 8F) and eventually reverts into depolarization (Fig 8G) for negative currents, when reaching the “apparent” reversal potential of dominant inhibition, in line with Hirsch et al. (1998). When measuring the membrane resistance during the visually evoked hyperpolarization in response to sign-mismatched flashed square stimulus, we observe roughly a doubling of the membrane conductance (Fig 8H), also in agreement with Hirsch et al. (1998). Finally, while excitatory conductance input is present only in response to flashed point-stimulus with contrast matching the polarity of the RF subfield, inhibition is present irrespective of the RF subfield polarity throughout the whole extent of the RF, just as predicted by the flashed bar experiments in Troyer et al. (2018). This shows that our model predictions account for the experimental observations in both stimulation paradigms, reported independently by Hirsch et al. (1998) and Troyer et al. (2018), respectively.

**Relationship between the spike and input orientation tuning of conductances**

To elucidate the generation of orientation tuning in the model we investigated the relationship between the orientation tuning of the neuron’s spiking response and of its excitatory and inhibitory input conductances. First, let us point out that in our model, the preset push-pull bias imposed in intra-cortical connectivity (see Fig 2A) implies that both excitatory and inhibitory synaptic inputs
originating predominantly from neurons of similar orientation preference. Thus, we would expect orientation tuning of neuron’s spike response and both excitatory and inhibitory inputs to coincide. While it is generally admitted that neurons in V1 receive inputs preferentially from excitatory and inhibitory neurons that are co-tuned in the orientation domain (Michalski et al., 1983; Smith and Kohn, 2008; Ko et al., 2011; Denman and Contreras, 2014; Lee et al., 2016), more direct voltage clamp measurements reveal a diversity of the relative tuning preferences of the neuron’s excitatory and inhibitory inputs (Monier et al., 2003; Monier et al., 2008), much larger than that assumed in the spike-based literature or inferred by current clamp experiments (Anderson et al., 2000).

To validate our model against these experimental findings, we applied the analysis from Monier et al., 2003 to the responses to drifting sinusoidal gratings. Figure 9 shows the results from our model (Fig 9A-C) as well as replotted analogous data from Monier et al., 2003 (Fig 9DE). The response magnitude of the Vm of a simple cell to a drifting grating can be measured by the DC component of the subthreshold signals (F0 component), and by the magnitude of the modulation at the frequency of the drift (F1 component).

As Figure 9A shows, while the tuning of the F0 component of both excitatory and inhibitory inputs was similar to that of the spiking response. Paradoxically, a significant number of neurons exhibited different tuning of input conductance and spiking response. For such neurons, excitation and inhibition showed similar tuning. The same behavior has been identified in sub-group of neurons reported in the Monier et al. 2003 (Fig 9D, points along the identity axis). On the other hand, if we examine the F1 component of the conductance signals, we see a different pattern: there are few neurons for which the orientation preference of the excitatory conductance differs from that of spiking response, but in contrast, the preference of orientation tuning of the inhibitory inputs varies widely and is independent of the excitatory preference. Again, such behavior has been noticed in a second sub-group of neurons reported in Monier et al. 2003 (Fig 9D, points along y-axis). Finally, we find that the mean tuning width of the excitatory and inhibitory inputs in model neurons is remarkably similar (Fig 9C) and in a very good agreement with Monier et al. 2003 (Fig 9E), although we do observe lower overall variability of the tuning width in comparison to the experimental data. Such lower variability could reflect the greater regularity that our model construction process induces in comparison to the biological V1 circuitry.
We should note that there are, however, some differences between the present computational study and Monier et al. 2003. First, Monier et al. used moving bars rather than drifting sinusoidal gratings, which precludes the separate analysis of the F1 and F0 components in their study. Second, their study was not restricted to layer 4 simple cells, and we cannot exclude the existence of systematic tuning differences between the different layers and functional cell types. Regardless, the presented analysis on sub-threshold orientation tuning genesis shows a close fit between our simulations and these experimental data, which strengthens the validity of the working hypothesis of our model.

**Discussion**

We have presented a biologically constrained model of the thalamo-cortical visual circuit that replicates two sets of seemingly contradictory experimental findings. The model employs correlation-based local connectivity among simple V1 cells, in which excitatory projections are strongly biased towards cells with spatially correlated RFs, whereas inhibitory projections are weakly biased towards cells with spatially anti-correlated RFs. This connectivity regime generates broad inhibition evoked by flashed bars across the RF (Taylor et al., 2018) as well as anti-phase modulation of excitation and inhibition in response to drifting gratings (Anderson et al., 2000; Monier et al., 2008; Tan et al., 2011), resolving the seeming contradiction in the experimental data.

Generally, these findings support the view that V1 simple cell connectivity in cat V1 does not adhere to a strict “push-pull” schema. It is often assumed that antiphase modulation of gE and gI in response to drifting gratings requires an underlying connectivity rule in which the inhibitory inputs to a simple cell must arise from an inhibitory cell co-tuned in orientation but antiphase in the spatial dimension (Troyer et al., 1998; Anderson et al., 2000). This idea was recently challenged by experimental findings showing that inhibition could be triggered from flashed bar stimuli across simple cell RF, independent of phase, suggesting a largely indiscriminate scheme for inhibitory connectivity not biased by RF correlations (Monier et al., 2008; Taylor et al., 2018). The simulations presented here support a compromise between the two strict connectivity schemes; the parameter search (Fig. 6) showed that neither a strictly push-pull nor a fully indiscriminate inhibitory connectivity scheme is compatible with the experimental findings.
The tendency of excitatory neurons synapsing onto neurons with correlated functional properties has been well established (Michalski et al., 1983; Monier et al., 2003; Smith and Kohn, 2008; Ko et al., 2011; Denman and Contreras, 2014; Lee et al., 2016; Gerard-Mercier et al., 2016), supporting the ‘correlated’ part of the push-pull like scheme. However, evidence on connectivity rules of inhibitory neurons, especially on the specific implication of the push-pull like scheme that the inhibitory neurons should target preferentially neurons with anti-correlated functional properties, was lacking. Importantly, a series of recent studies from the Fitzpatrick lab gave support to this notion, showing that inhibitory neurons in layer 2/3 themselves receive selective inputs from their potential pre-synaptic pool (Wilson et al., 2017), and preferentially target other inhibitory neurons with opposite direction preference (Wilson et al., 2018). Here we focus on the relationship between the phase of RFs of pairs of neurons and the likelihood of a connection between them, while Wilson et al., 2018 investigated the direction preference, not phase. But if we extend a spatial RF kernel into the temporal domain, two neurons preferring the same orientation but opposite direction preference will be perfectly anti-correlated. The Wilson et al., 2018 study thus offers important indirect evidence supporting the existence of the anti-correlated inhibition part of the push-pull like scheme as hypothesized here (see also example of directionally anti-correlated inhibition in Fig 4 (cell 11) in Monier et al, 2003).

In the present study we assume the putative correlation-based canonical circuit only for simple cells receiving direct thalamic input. A wider repertoire of gE/gI configurations were observed electrophysiologically in cells with simple-like spike-based behavior, but unidentified layer of origin (Monier et al., 2003; 2008; Baudot et al., 2013). Interestingly, our results show that despite the regularities imposed by our correlation based circuit scheme, the model also generates gE/gI configurations that are not straightforwardly predicted from the correlation-based connectivity scheme (see Fig 9), and crucially, specific relationships between the relative gE/gI tunings emerge that are in a remarkable accordance with the experimental data (see Fig 9).

A question remains as to the source of this wider repertory of configurations if they cannot be linked directly to the underlying connectivity scheme in our model. We hypothesize that the probabilistic nature of connection generation utilized in the model is likely to induce small random biases in individual neurons, that are further amplified by the recurrent cortical dynamics. Another
likely contribution, already proposed by Monier et al., 2003, is local neighborhood heterogeneity
in the orientation maps, whereby a neuron, depending on its location in the map (close to pinwheel
singularity vs. in the center of an iso-orientation domain), could recruit patterns of excitatory and
inhibitory inputs of varying orientation preference. Due to local averaging, the direct impact of
these factors is likely small, but it is further amplified by the recurrent interactions of the cortical
network. In this study, the size of the model, spanning approximately a single hyper-column, and
the need to record only neurons in the central area of the model to avoid border effects (see
Methods), did not allow to collect sufficiently diverse data to test this hypothesis.

The majority of the model parameters were set based on existing experimental data. But some,
where relevant data or quantification are missing, could not be fully constrained. Among these are
the functional biases of the local connections between excitatory and inhibitory cells, which, as
extensively discussed above, constitute key elements for the results of the model. However, for
these two parameters, we have explored extensively the parameter space, determining the range
where observed data could be reproduced (Fig 6). It is unfortunately computationally infeasible to
perform a systematic exploration of all remaining, not fully constrained, model parameters. But it
is important to point out that all model parameters are further indirectly constrained by the battery
of functional tests which the model has been already validated against in the parent manuscript
(Antolik et al. 2019). However, our empirical experience with the model identified several
parameters, to which the model dynamics are more sensitive, such as the relative contribution of
the feed-forward vs. intra-cortical excitatory drive to cells, or generally parameters that influence
the overall balance of excitation vs. inhibition.

Since the discovery of poorly orientation tuned inhibitory neurons in the granular layer of V1 (3
reconstructed smooth cells in Hirsch et al., 2003), untuned inhibition has been welcomed in
computational models. Its use enabled simulation of contrast invariant orientation tuning and low-
pass temporal frequency tuning in simple cells (Lauritzen and Miller, 2003). But later studies
(Cardin et al., 2008; Nowak et al., 2007), that investigated this issue in 55 fast spiking cells (of
unidentified layer origin), failed to identify inhibitory neurons that lack orientation tuning. Cardin
et al. (2008) found that when measured as HWHH, fast-spiking neurons were on average only
moderately more broadly tuned than excitatory ones. They did not observe any untuned fast
spiking cells, and only one with orientation tuning greater than 100 degrees. Consistently, Nowak et al. (2007), report only 2 (out of 19) inhibitory neurons with RURA above 50%, and none about 70%. They also show that the orientation tuning width of fast-spiking cells forms a continuum across the population. The present model does produce some poorly tuned inhibitory cells (e.g. in the model condition marked with * in figure 6D, 16.6% of inhibitory cells exhibit HWHH of orientation tuning of spike response greater than 45º while only 2.5% of excitatory cells do so), that contribute to the moderately selective inhibition that has been identified here as a prerequisite for explaining out the experimental data. Note however, when pooling all layer recordings, voltage clamp conductance decomposition and non-linear STC analysis of intracellular responses reveal a similar orientation tuning width in the presumed excitatory and inhibitory sources of simple cells (Figure 6B in (Monier et al., 2003), Figure 8B in (Fournier et al., 2014)), which is also reproduced in our model (see Fig 9C). Thus, while our model is consistent with the statistical evidence of some poorly tuned inhibitory neuron in V1, a separate class of untuned inhibitory cells was not added. The important finding is that, even without this built-in tuning constraint, the model did generate contrast-invariant orientation tuning, as experimentally reported and quantified in figure 6B.

Other attempts have been made to constrain neural modeling of V1 with a wealth of biological data. At the structural level, large-scale integrative approaches have been gaining traction recently. In mouse, several studies explored extensive morphologically detailed circuit reconstructions of sensory cortex (Markram et al., 2015; Hawrylycz et al., 2016; Arkhipov et al., 2018). In higher mammals, number of less anatomically detailed studies investigate visually evoked sub-threshold network dynamics in biologically realistic spiking networks (Tao et al., 2006; Wielaard & Sajda, 2006; Rangan et al., 2009; Chariker et al., 2016). But rarely did models tried to reconcile structure and function assessed from the microscopic to the mesoscopic scales in a way to confront predictions with the stimulus-dependency of visual responses reported in V1 in vivo.

It is important to also identify limitations of our model. The presented model focuses on simple cells receiving direct thalamic input. While it is well established that the majority of simple cells are confined to layer 4 and layer 6 in cat V1, it should be acknowledged here that the key experimental studies constraining this computational work (Anderson et al., 2000; Monier et al., 2003; 2008; Tan et al., 2011; Taylor et al., 2018) did not provide the exact layer origin of all the
recorded simple cells. However, since the exact laminar location is of lesser importance here than
the “simple” push-pull behavior of the spiking RF, it remains fair, in our view, to accept that the
present model best applies to the organization of excitation and inhibition in simple cells in a
virtual thalamo-cortical recipient layer (which may collapse layer 4 and layer 6 in cat V1). This
model has still other pending structural unknowns. The topological organization of the ON/OFF
thalamic pathway should be incorporated into the model in future, since it can have interesting
implications for organization of excitation and inhibition in relation to the neuron’s RF and explain,
for instance, the OFF bias of inhibition shown by Taylor et al. (2018). Finally, a more refined
biophysical dissection should be engineered, concerning the subcomponents of the excitatory and
inhibitory synaptic composite drive due to the various receptor types (NMDA, AMPA, GABA_A,
GABA_B), which may be recruited differentially by the different visual input regimes.

In summary, this study shows that explicit multiscale integration of the structural and biophysical
properties of the underlying neural substrate can be a powerful approach for accounting for the
diversity of electrophysiological measurements of V1 physiology. Our model demonstrates that in
complex dynamical systems such as visual cortex, using restricted results from a limited number
of experimental conditions (and visual contexts) is insufficient to make far-reaching, generalized
conclusions on the synaptic nature of the underlying connectivity. Only systematic, comprehensive
computational models of the primary visual cortex, that incorporate a broad range of established
constraints from different experimental designs, can lead to reconciliation of the often seemingly
contradictory diversity of experimental findings, and in turn to an accurate characterization of the
system under study.

References


Alitto HJ (2004) Influence of Contrast on Orientation and Temporal Frequency Tuning in Ferret


Figure 1. Schematics of V1 response characteristics to bars and gratings: (A) Left, example of a 2D receptive field (RF) of a simple cell; Middle, types of stimuli used in our simulations: upper panel, grating, lower panel, bars. Right, Vm and conductance responses. (B) Schematized spiking, Vm and conductance response profiles to drifting gratings. Left, the FR and Vm responses to a grating at the cell’s preferred orientation are modulated in time. Right, excitatory (gE, red) and inhibitory (gI, blue) conductance have been shown to be modulated in temporal antiphase in response to this stimulus. (C) Spiking and Vm responses to flashed bright or dark bars at the cell’s preferred orientation. Top, the 1D RF, calculated from the cell’s response to bright minus the response to dark, reveals the three RF subunits (ON, OFF, ON), also shown in 2D (left inset, with overlaid bar positions in dotted lines). Middle, spiking responses to bright (grey thick line) or dark (black thin line) bars at different positions across the cell’s RF in visual space. Bottom, Vm responses to flashed bars have a similar pattern, and are spatially anti-correlated. (D) Conductance responses over space to flashed bars, as shown by recent experimental results (Taylor et al., 2018). While the gE responses over space...
are anti-correlated, the gI responses are broad and highly spatially correlated. (E) Conductance responses over space to flashed bars, as predicted based on the anti-correlated conductance responses to gratings. In this scheme, gE and gI responses to bright and dark bars are spatially anti-correlated.
Figure 2. Schematic of LGN-V1 model architecture. (A) Cortical lateral connectivity. Functional specificity of local connections is not shown and connection ranges are not to scale. (B) Extent of modeled visual field and example of receptive fields (RFs) of one ON- and one OFF-center LGN relay neuron. The model is retinotopically organized, and the extent of the modeled visual field is larger than the corresponding visuotopic area of modeled cortex in order to prevent clipping of LGN RFs. (C) Local connectivity cortex follows a biased push-pull organization: excitatory connections are biased towards correlated RFs, while inhibitory connections are mildly biased towards anti-correlated RFs (see E). (D) Afferent RFs of cortical neurons are formed by sampling synapses from a probability distribution defined by a Gabor function overlaid on the ON and OFF LGN sheets, where positive parts of the Gabor function are overlaid on ON and negative on OFF sheets. The ON regions of RFs are indicated by white color, while OFF regions by black. (E) Connection bias factor as a function of RF correlation between pre- and post-synaptic cells. Excitatory cells (red) are much more likely to connect to other cells with similar RFs, while inhibitory cells (blue) are slightly more likely to connect to other cells with anti-correlated RFs (see Figure 6 in Monier et al., 2003). Later in the paper, we explore a parameter space of excitatory and inhibitory connection specificity by varying these values of $\sigma_E$ and $\sigma_I$. 
Figure 3. *Comparison of true and estimated mean synaptic excitatory and inhibitory conductance in model neurons.*

(A-C) the mean over 10 trial of excitatory (red) and inhibitory (blue) conductance trace in response to 2 second presentation of sinusoidal grating of optimal orientation, drifting at 2Hz. (A) conductance estimated through simulated current clamp method in three representative model simple cells. (B) true conductance in the same three model neurons as in (A). (C) an example of excitatory and inhibitory conductance estimated through current clamp in a simple cell recorded in cat V1 (adopted from Baudot et al. (2013)). (D) Pearson correlation between estimated excitatory and inhibitory conductance for 30 model neurons (abscissa) plotted against the same measure derived from the true recorded conductance of the same neurons (ordinate).
Figure 4. Model neurons’ responses to flashed bars (A) Vm (center; 4 example single trials overlaid, top rows, and average ± SD, bottom row) and conductance (right, average ± SD) responses of an example simple cell from the model to a bright bar flashed in its ON subregion (top row) and a dark bar flashed in the same position (middle row). The cell fired in response to the bright bar in the ON subregion, and this response was underlied by a combination of gE and gI (circles indicate maximum response). A dark bar in the ON subregion evoked hyperpolarization of the Vm, underlied by minimal gE and a strong gI response. (B) Average Vm, gE, and gI responses to flashed bright and dark bars across space, for the same example cell. Circles represent raw values at each location, and lines show a Gaussian best-fit, for illustration only. Spatial correlation values between bright and dark responses (right) were calculated from the data points. (C) Histograms of spatial correlations from n=30 cells from the simulation (black) and from n=17 cells from a previously published dataset (pink, Taylor et al., 2018). Vm and gE responses to bright versus dark bars are negatively correlated, while gI responses are positively correlated. Arrowheads indicate medians of each distribution.
Figure 5. Model neurons’ responses to drifting gratings (A) Vm and spiking responses over time from an example cell (same cell as in Figure 3) to drifting gratings at its preferred orientation (top, single trials; middle, average [without spikes] ± SD; bottom, spike rasters from 10 trials). (B) The same cell’s conductance responses over time to the same grating. Top, average gE (red) and gI (blue) from 10 trials. Bottom, smoothed conductances (see Methods) clearly illustrate a negative temporal correlation (correlation coefficient = -0.74) between gE and gI. (C) Histogram of temporal correlations from n=30 cells from the simulation. As in the example cell, the majority of cells show temporal anti-correlation between gE and gI to the drifting grating. Arrowhead indicates the median correlation from this distribution.
Figure 6. Spiking, Vm, and conductance characterizations across a parameter space of intra-cortical connectivity schemes (A) Illustration of the parameter space explored in this figure. The sigmoidal function parameters $\sigma_{E}$ (for excitatory connections) and $\sigma_{I}$ (for inhibitory connections) were varied independently between 0.6 (highly selective)
and 4 (virtually unselective) values. Throughout this figure, selectivity of excitatory connections grows to the left (of the x-axis) and that of inhibitory connections grows to the bottom (of the y-axis) of each matrix. RF correlation-based connectivity curves for the four extreme combinations (corners of parameter space, a-d) are shown in the right panel (see also Fig 2). (B) Variation of spiking and Vm characteristics across explored parameter space. Left matrix, the median half-width at half-height (HWHH, degrees) of the spiking responses to high-contrast drifting gratings at varying orientations. Center, the median change in HWHH between the tuning curves in response to gratings at high versus low contrast. Right, the median spatial correlation of Vm responses to bright versus dark bars (as in Fig. 4) at the cells’ preferred orientation. Aside from the lower/lower-left region of parameter space (corresponding with highly selective inhibitory connectivity), these metrics did not vary substantially across the connectivity parameter space. Black squares indicate parametrizations where orientation tuning could not be well fit with Gaussian curve. (C) Variation of conductance characteristics across explored parameter space. Left, the median temporal correlation of gE and gI in response to high-contrast drifting gratings at the cells’ preferred orientation (as in Fig. 5C). Center, the median spatial correlation of gE responses to bright versus dark bars flashed in different locations across each cell’s RF (as in Fig. 4). Right, the median spatial correlation of gI responses to bright versus dark bars flashed in different locations across each cell’s RF (as in Fig. 4). (D) Illustration of the region of parameter space that yields simulations with values of spiking, Vm, and conductance response characteristics that are consistent with experimental findings. This region of parameter space corresponds with relatively selective excitatory connections (0.6<σE<2.0) and much less selective inhibitory connections (σI>1.2). Note that the simulation illustrated in Figures 4,5 is shown within this parameter space with an asterisk (*).
Figure 7. Push-pull measures for different model parametrizations. (A-B) panels refer to parametrizations marked with corresponding letter in Fig 5D. Each panel shows the average $g_E$ (top left) and $g_I$ (middle left) responses to flashed bars across space for an example cell; $g_E$ and $g_I$ responses to drifting grating stimulus (bottom left) for the same example cell; histograms of spatial correlations between $g_E$ responses to dark and bright bars (top right), between $g_I$ responses to dark and bright bars (top middle) and between $g_E$ and $g_I$ responses to drifting sinusoidal grating stimulus. All plotting conventions match corresponding plots in figures 4 and 5.
Figure 8. Response of model neurons to flashed square stimuli under different current injection levels. The four columns of the figure show data from 4 representative excitatory cortical model neurons. (A,C) receptive field maps (light for ON-subfield, dark for OFF-subfield) with the position of the contrast-matched square stimulus (A) and contrast-mismatched square (C) stimulus overlaid on the stimulation grid. Grid spacing was 0.3°. (B and D): Top traces, conductance changes (gE, red; gI, blue). Bottom traces, membrane potential (black line). (B) response to contrast-matched stimulus. (D) response to contrast-mismatched stimulus. (E-G) voltage response to contrast-mismatched stimulus during +0.1nA (E), -0.1nA (F), and -0.2nA (G) current injection. (H) Total synaptic conductance change from baseline calculated following the same procedure as in Hirsch et al. (1998). Stimulus presentation is indicated by black bar under each voltage trace.
Figure 9. Tuning properties of the excitatory and inhibitory input conductances in excitatory simple cells. (A-B) Scatter plot of preferred orientations of $g_E$ (abscissa) and $g_I$ (ordinates) input conductances, in response to the drifting sinusoidal grating stimuli. The preferred orientation of the synaptic input is plotted relative to the orientation preference of the spiking response for each model cell. The ISO-condition corresponds to the bottom left corner of each scatter plot, where $g_E$, $g_I$ and spike output preferences are co-aligned. In A, the F0 component was used to compute the orientation preference, whereas in B, we used the F1 component of the input conductance. (C) The half-width at half-height (HWHH) of the excitatory vs. inhibitory input conductance tuning curves. (D) As in A and B, but taken from the in-vivo data from simple V1 cells in cat (Monier et al., 2003, Fig 5 bottom left panel). Since the test stimulus was a moving bar, the separation into F1 and F0 components was not possible. We have excluded from the plot cells shown in (Monier et al., 2003) which had low selectivity (see caption of Fig. 5 in Monier et al., 2003) since, for such cells, the preferred direction preference cannot be determined reliably. (E) as in C but data from cat (Monier et al., 2003, Fig. 5 top left panel). In order, to make the Monier et al 2003 data directly comparable to the present study, we have collapsed, in D and E, opposite directions corresponding to the same orientation within the (iso, cross) range. All model data correspond to the condition marked with * in Fig 6D.