Transient inhibition to light explains stronger V1 responses to dark stimuli

David St-Amand & Curtis L. Baker Jr.*

McGill Vision Research Unit, Department of Ophthalmology & Visual Sciences, McGill University, Montreal, Quebec, Canada, H3G 1A4

***Corresponding author:**

Curtis L. Baker Jr.*

McGill Vision Research

MUHC Research Institute

1650 Cedar Avenue, L11-521

Email: curtis.baker@mcgill.ca

Tel: +1-514-934-1934 ext.34819

Fax: +1-514-843-1691

Conflict of interest: The authors declare no competing financial interests

Acknowledgements: We wish to thank Guangxing Li for contributions to software and technical support in experiments, and Philippe Nguyen and Amol Gharat for their important contributions to data collection and spike-sorting. Funded by Canadian Institutes of Health Research (CIHR) grant MOP-119498 to C.B.

Abstract

Neurons in the primary visual cortex (V1) receive excitation and inhibition from two different pathways processing lightness (ON) and darkness (OFF). V1 neurons overall respond more strongly to dark than light stimuli (Yeh, Xing and Shapley, 2010; Kremkow et al., 2014), consistent with a preponderance of darker regions in natural images (Ratliff et al., 2010), as well as human psychophysics (Buchner & Baumgartner, 2007). However, it has been unclear whether this "dark-dominance" is due to more excitation from the OFF pathway (Jin et al., 2008) or more inhibition from the ON pathway (Taylor et al., 2018). To understand the mechanisms behind dark-dominance, we record electrophysiological responses of individual simple-type V1 neurons to natural image stimuli and then train biologically inspired convolutional neural networks to predict the neurons' responses. Analyzing a sample of 74 neurons (in anesthetized, paralyzed cats) has revealed their responses to be more driven by dark than light stimuli, consistent with previous investigations (Yeh et al., 2010; Kremkow et al., 2013). We show this asymmetry to be predominantly due to slower inhibition to dark stimuli rather than by stronger excitation from the thalamocortical OFF pathway. Consistent with dark-dominant neurons having faster responses than light-dominant neurons (Komban et al., 2014), we find dark-dominance to solely occur in the early latencies of neurons' responses. Neurons that are strongly dark-dominated also tend to be less orientation selective. This novel approach gives us new insight into the dark-dominance

phenomenon and provides an avenue to address new questions about excitatory and inhibitory integration in cortical neurons.

Introduction

The early visual system is separated into two distinct pathways: an ON pathway which responds more to lighter parts of images and an OFF pathway which encodes darker image regions. Neurons in primary visual cortex (V1) combine inputs from these two pathways, but the nature of this integration is still poorly understood.

V1 neurons evidently receive asymmetrical inputs from the two pathways, since they are on average more responsive to dark than light stimuli (Jin et al., 2008; Yeh et al., 2009), especially at low spatial frequencies (Kremkow et al., 2014; Jansen et al., 2018) and shorter time latencies (Komban et al., 2014). This is presumably adaptive due to the preponderance of dark regions in natural images (Ratliff et al., 2010), which is also more pronounced at lower spatial frequencies (Cooper & Norcia, 2015). These asymmetries may influence human perception, since dark stimuli are processed faster and more reliably than light stimuli (Buchner & Baumgartner, 2007; Komban, Alonso & Zaidi, 2011).

This “dark-dominance” effect has been attributed to more OFF than ON excitatory inputs from the lateral geniculate nucleus (LGN) to the layer 4 of V1 (Jin et al., 2008). However, this does not explain why more dark-dominant neurons are found in layers 2/3 than in layer 4 (Yeh et al., 2009). This

discrepancy could be explained by stronger ON than OFF intracortical inhibition within V1 (Taylor et al., 2018). Hence, whether dark-dominance is mostly due to excitation to dark stimuli or inhibition to dark stimuli remains unclear. Here we develop a novel machine learning approach to disambiguate excitation from inhibition in extracellular recordings, which allows us to make quantitative inferences about how cortical neurons integrate ON and OFF inputs.

To better understand what drives V1 responses, we predict the responses of recorded neurons to natural images with a simple, biologically-inspired convolutional neural network. This neural network processes the natural images' light (ON) and dark (OFF) information in separate pathways. The first layer of each pathway consists of a convolution with a parametrized 2D gaussian, which represents the responses of LGN neurons (omitting the weaker surrounds; Croner & Kaplan, 1995). The second layer is a linear weighted sum of the excitatory or inhibitory contributions to the output of the pathway. These responses sum, to provide the model's output. From these estimated weights, we infer how much excitation and inhibition arises from each pathway, at every spatial location and temporal lag of a V1 simple cell's receptive field.

Using this approach, we find time to play an important role in the dark-dominance phenomenon in V1 neurons, which only occurs at the early response latencies. We show these stronger dark responses to be predominantly driven by a lack of inhibition to dark stimuli, which in turn causes neurons to be less orientation selective in their early responses. These findings suggest that slower

inhibition to dark than to light plays a crucial role in the dark-dominance found in primary visual cortex.

Methods

Animal preparation

Anesthesia in adult cats was induced by isoflurane-oxygen (3–5%) inhalation, followed by intravenous (iv) cannulation and bolus iv injection of propofol (5 mg/kg). Surgical anesthesia was maintained with supplemental doses of propofol. Glycopyrrolate (30 µg im) and dexamethasone (0.2 mg/kg iv or 1.8 mg im) were administered and a tracheal cannula or intubation tube was inserted. Throughout the surgery, body temperature was thermostatically maintained and heart rate was monitored (Vet/Ox Plus 4700).

The animal was then positioned in a stereotaxic apparatus and connected to a ventilator (Ugo Basile 6025). Cortical Area 17 was exposed by a craniotomy (P3/L1) and a small durotomy, and the cortical surface protected with 2% agarose capped with petroleum jelly. Local injections of bupivacaine (0.50%) were administered at all surgical sites. During recording, anesthesia was maintained by the infusion of propofol (5.3 $mg \cdot kg^{-1} \cdot h^{-1}$), and in addition, remifentanil (initial bolus injection, 1.25 $\mu g \cdot kg^{-1}$, then infusion, 3.7 $\mu g \cdot kg^{-1} \cdot h^{-1}$) and O_2/N_2O (30:70 ratio) delivered through the ventilator. Paralysis was produced

with a bolus iv injection of gallamine triethiodide (to effect), followed by infusion (10 $mg \cdot kg^{-1} \cdot h^{-1}$). Throughout subsequent recording, expired CO_2, EEG, ECG, body temperature, blood oxygen, heart rate, and airway pressure were monitored and maintained at appropriate levels. Intramuscular glycopyrrolate (16 μg) and dexamethasone (1.8 mg) were also administered daily.

Corneas were initially protected with topical carboxymethylcellulose (1%) and subsequently with neutral contact lenses. Spectacle lenses were selected with slit retinoscopy to produce emmetropia at 57 cm, and artificial pupils (2.5 mm) were provided. Topical phenylephrine hydrochloride (2.5%) and atropine sulfate (1%), or cyclopentolate (1.0 %) in later experiments, were administered daily.

All animal procedures were approved by the McGill University Animal Care Committee and are in accordance with the guidelines of the Canadian Council on Animal Care.

Extracellular recording

Recordings were performed using 32-channel silcon probes (NeuroNexus), in most cases polytrodes (A1x32-Poly2-5mm-50s-177) or occasionally linear arrays (A1x32-6mm-100-177), advanced with a stepping motor microdrive (M. Walsh Electronics, uD-800A). Raw electrophysiological signals were acquired with a Plexon Recorder (3 Hz to 8 kHz; sampling rate, 40 kHz), along with supplementary signals from a small photocell placed over one corner of the visual stimulus CRT, which were used for temporal registration of

stimuli and spikes, and to verify the absence of dropped frames. Spike waveforms were carefully classified from the recorded multichannel data into single units, using Spikesorter (Swindale & Spacek, 2014). Only clearly sorted units were used for further analysis.

Visual stimuli

Visual stimuli are presented on a gamma-corrected CRT monitor (NEC FP1350, 20 inches, 640x480 pixels, 150 Hz, 36 cd/m^2) at a viewing distance of 57 cm. Stimuli are produced by an Apple Macintosh computer (MacPro, 2.66 GHz, 6 GB, MacOSX ver. 10.6.8, NVIDIA GeForce GT 120) using custom software written in MATLAB (ver. 2012b) with the Psychophysics Toolbox (ver. 3.0.10; Pelli, 1997; Brainard, 1997; Kleiner et al., 2007). We select a channel having with good spike responses to hand-held bar stimuli, which we use to determine the dominant eye (with the non-dominant eye subsequently occluded), and to position the CRT monitor to be approximately centered around the population receptive field.

Visual stimuli are ensembles of 375 natural images taken from the McGill Calibrated Colour Image Database (Olmos & Kingdom, 2004), cropped to 480x480, converted to monochrome 8-bit integers - as in Talebi & Baker (2021), but with a higher RMS contrast. We randomly present each ensemble at 75 images per second (i.e. every 13.33 ms) in short movies of 5 seconds each. We have separated the ensembles into three sets, to evaluate predictive performance independently from overfitting. The training set has 20 movies which are presented 5 times each, while the validation and testing sets each

have 5 movies which are presented 20 times each. The validation and testing sets are presented more often to provide less noisy estimates of the fitted model's predictive performance.

Model architecture

To better understand differences between the ON and OFF pathways, we employ a model architecture inspired from visual circuitry (Figure 1), which is optimized to predict a recorded cortical neuron's mean spiking responses to the natural image ensembles. We model LGN receptive fields as parametrized 2D isotropic gaussians, acting convolutionally on the stimulus images. The antagonistic surrounds are neglected, so there is only a pair of gaussian width parameters, for the ON and OFF pathways, to be estimated. The connections between the gaussian operators and the model cortical neuron are a pair of linear weighted sums of rectified responses of the gaussian-operators, across a series of time lags. Each of these linear weighted sums acts like a "dense layer" in machine learning, but note that there is *not* a subsequent rectification. The output of each dense layer might be thought of as a presynaptic membrane potential contribution, from its respective ON or OFF pathway.

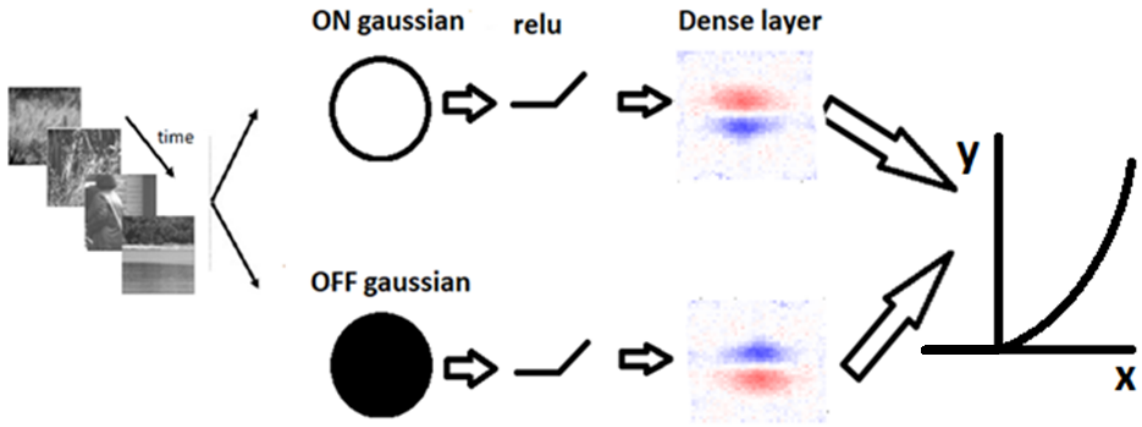

Figure 1. Model architecture for responses of a cortical neuron to visual stimuli such as natural images. Light and dark image regions are encoded as rectified responses of convolution with positive and negative spatial gaussians, respectively. Linear weighted sums are separately are taken for each each pathway and summed, followed by a half-power pointwise nonlinearity. A machine learning algorithm estimates the sizes and amplitudes of the parameterized gaussian operators, and the two sets of dense layer weights, for each of a series of time lags.

The inputs to the model are the pixel luminances of the natural image stimuli (cropped, as described below), which are spatially downsampled to 40x40 pixels. The mean of the inputs is centered at zero by subtracting the overall mean across all images within an ensemble. To model the neuron's temporal processing, the inputs to the regression (see below) are composed of the preceding 7 images, each of which were presented for 13.33 ms. The model output is the neuron's response, with spike times collected into time bins of 13.33 ms each. (duration of each stimulus image frame).

The stimulus images are convolved with a pair of parametrized 2D gaussian filters (with positive or negative polarity, for the ON and OFF pathways, respectively), each followed with a half-wave rectification (ReLU). The 2D gaussians represent receptive fields of LGN neurons in which the weaker surrounds (Croner & Kaplan, 1995) are neglected, as follows:

$$g(h,v,p,t) = \frac{\alpha_p}{\sqrt{2\pi\sigma_p^2}} e^{-\left(\frac{h^2+v^2}{\sigma_p^2}\right)} \quad (1)$$

with *h* and *v* the horizontal and vertical distances between a pixel and the center of the gaussian, respectively, and σ representing the standard deviation (i.e. width) and α the amplitude (i.e. the height) of the 2D gaussian. The σ and α parameters are estimated separately for each pathway (*p*). For the ON pathway to respond selectively to light information, α is always positive (and negative in the OFF pathway). The convolution of the 2D gaussian with the inputs are half-wave rectified (ReLU) spiking responses of LGN neurons (Persi et al., 2011):

$$c(i, j, t, p, k) = \max(0, \sum_{h=-5}^{6} \sum_{v=-5}^{6} x_{i+h,j+v,k,t} * g(h,v,p,t)) \quad (2)$$

where *i* and *j* are the horizontal and vertical coordinates of the center of the 12x12 2D gaussian, *x* is the luminance of a specific downsampled pixel, *t* is the number of bins between the shown image and the recorded response (latency), and *k* is the time bin of the neuron's response. The convolution with the 2D gaussians is implemented with zero-padding and a "stride" of 1. Due to the first rectification, the ON pathway encodes luminances above (lighter than) the mean, and the OFF pathway luminances below (darker than) the mean.

For each of the ON or OFF pathways, the model then takes a linear weighted sum of the convolution outputs from the respective rectified gaussians, with each weight notionally representing the excitatory or inhibitory inputs from an array of LGN cells to the cortical neuron. The sum of responses from these dense layers is followed by a rectified power law output non-linearity, which forms the final output of the model and the prediction of the neuron's mean spiking response:

$$\hat{y}_{pre}(k) = \max(0, \sum_{p=1}^{2} \sum_{t=1}^{7} \sum_{i=1}^{40} \sum_{j=1}^{40} c(i,j,p,t,k) * w_{i,j,p,t}) \quad (3)$$

$$\hat{y}(k) = a\, \hat{y}_{pre}^{b}(k) \quad (4)$$

where *w* represents the dense layer weights, $\hat{y}(k)$ is the prediction of a neuron's response for the kth time bin, *b* is the estimated exponent of the rectified power nonlinearity, and *a* is a scale (gain) factor.

Optimization and regularization

To characterize a neuron's receptive field, we find the model parameters which minimize the difference between its recorded responses and the responses predicted by the model, which requires fitting a total of 2x40x40x7 = 22,400 dense layer weights, and 2 parameters for the 2D gaussians. To minimize over-fitting due to the large number of parameters, we employ L2-regularization by penalizing the squared amplitude of the weights (Hoel & Kennard, 1970),

implemented by minimizing a loss function in which the first term is the squared error of the model prediction and the second term the regularization penalty:

$$\sum_{k=1}^{n}(y_k - \hat{y}_k)^2 + \lambda \sum_{p=1}^{2} \sum_{t=1}^{7} \sum_{i=1}^{40} \sum_{j=1}^{40} \beta_{i,j,t,p}^2 \quad (5)$$

where y_k is the neuron's recorded response, $\hat{y}_k$ the model's predicted response for the kth time bin, $\beta_{i,j,t,p}$ the dense layer weights (for the *i*-th On/Off stream, *t*-th time lag, and *i,j*-th dense layer position) and λ the L2-regularization hyperparameter. Based on pilot results from a representative subset of neurons, the hyperparameter λ is set to 5 x 10^{-6} in a first pass and 2 x 10^{-6} in a second pass (see below, re three-pass training procedure). In the third pass, we train the model with different λ values of [1,2,4,8,16] x 10^{-6}, and choose the λ value giving the best model performance on the validation dataset for each neuron.

This loss function is minimized using the Adam optimization algorithm (Kingma & Ba, 2014) with mini-batch gradient descent (Li et al., 2014). To further reduce overfitting, we apply dropout during training to both the convolutional and final dense layers with a probability of 50% (Srivasta et al., 2014).

The data is separated into training, validation, and test sets, corresponding to the three sets of stimulus movies. The model parameters are fit to the training set using a mini-batch size of 100 stimulus-response pairs. As an additional regularization measure, early stopping is implemented by saving a neuron's model parameters at the point in training when it reaches a peak of performance for predicting the validation dataset, and training stops if there is no improvement over the subsequent 50 epochs. The reported performance of the

model is evaluated on a third, separate test set to prevent bias in estimated predictive performance.

Three-pass training procedure

Because V1 receptive fields usually only occupy a small subset of the displayed visual stimulus images, it would be detrimental to optimize each neuron's model based on the full extent of the images. Doing so would entail a very high overparameterization, or a loss of spatial resolution due to excessive downsampling of the stimulus images, in either case yielding poorer predictive performance. To address this issue, we use a three-pass training procedure, with each pass improving the spatial resolution of the receptive field estimate. In the first-pass, we optimize the model parameters using the full 480x480 stimulus images downsampled to 40x40. We then manually designate a square cropping window that encloses an area slightly larger than the apparent receptive field. Next, we crop the image within that window, and rescale the image within it to 40x40. This cropped image is then used to re-train the model in the second pass, where we repeat the procedure, but with the cropped image. In the third pass, we adjust the cropping window based on the model estimate obtained in the second pass. This third pass provides much higher accuracy in identifying the boundaries of the receptive field, and gives us the final model fits that we use for the remaining analysis. This three-pass training procedure allows us to characterize a neuron's receptive field with high resolution and substantially increases predictive performance.

Output non-linearity

As a neuron's input increases, its firing rate can be modeled as a rectified power law (Heeger, 1991; Anzai et al, 1999; Persi et al., 2011). It has proven problematic to simultaneously estimate the power law exponent with the other parameters using backpropagation. This problem is most likely due in part to the "exploding gradient" problem (Pascanu, Mikolov & Bengio, 2012). To resolve this issue, we initially set a power-law exponent value of unity, and wait 100 epochs to get a rough estimate of the other parameter values. We then pause the model optimization, to fit the two parameters of the output nonlinearity to the predicted vs. measured neuron responses - and then resume full model parameter optimization, keeping the output nonlinearity parameters fixed.

To address the heavily uneven distribution of the measured data, we bin the predicted responses into 100 bins of 75 responses each and compute the mean measured response for each bin - a modification of the method used by Anzai et al (1999). We then fit a scaling factor '*a*' and an exponent '*b*' (Eq. 10) to minimize the difference between the binned predicted responses $\hat{y}$ and the measured spike rates *y*, using python scipy's '*optimize.curve_fit'*.

Estimating excitation and inhibition

The spatiotemporal properties of each of the ON and OFF pathways in the fitted model depend on both the dense layer weights and estimated 2D gaussians (which may differ in amplitude and width for the ON and OFF

pathways). To incorporate both in our analysis, for each of the ON and OFF pathways we convolve the 2D gaussian with the corresponding dense weights, to produce a "reconstructed" receptive field, which represents the neuron's response to either light or dark stimuli. This procedure results in a 40x40x7 spatio-temporal filter for each pathway, which is indicative of how the model neuron responds to light and dark stimuli. For further analyses we estimate the overall amount of excitation and inhibition from the filter for each pathway and time lag, by taking the sum of all positive or negative values in either the ON or OFF reconstructed receptive field. This procedure provides an inference of the total amount of ON excitation, ON inhibition, OFF excitation and OFF inhibition contributing to each neuron's response.

$$ON_{excit}(t) = \sum_{j=1}^{40} \sum_{i=1}^{40} \sum_{h=-5}^{6} \sum_{v=-5}^{6} \max\left(0, w_{i+h,j+v,1,t} * g(h,v,1,t)\right) \quad (6)$$

$$OFF_{excit}(t) = \sum_{j=1}^{40} \sum_{i=1}^{40} \sum_{h=-5}^{6} \sum_{v=-5}^{6} \max\left(0, w_{i+h,j+v,2,t} * -g(h,v,2,t)\right) \quad (7)$$

$$ON_{inhib}(t) = -\sum_{j=1}^{40} \sum_{i=1}^{40} \sum_{h=-5}^{6} \sum_{v=-5}^{6} \min\left(0, w_{i+h,j+v,1,t} * g(h,v,1,t)\right) \quad (8)$$

$$OFF_{inhib}(t) = -\sum_{j=1}^{40} \sum_{i=1}^{40} \sum_{h=-5}^{6} \sum_{v=-5}^{6} \min\left(0, w_{i+h,j+v,2,t} * -g(h,v,2,t)\right) \quad (9)$$

$$Light(t) = ON_{excit}(t) + OFF_{inhib}(t) \quad (10)$$

$$Dark(t) = OFF_{excit}(t) + ON_{inhib}(t) \quad (11)$$

Light/dark balance

To better understand how much individual neurons are light- or dark-dominated, we designed a light-dark balance index (LDB) to indicate the relative influence of a neuron's light and dark weights:

$$LDB(t) = (Light(t) - Dark(t)) / (Light(t) + Dark(t)) \quad (12)$$

This index varies from -1.0 to 1.0. Positive LDB values indicate a neuron is light-dominated, and negative values that it is dark-dominated.

Excitation/Inhibition balance

The excitation/inhibition balance (EIB) index is similar to LDB, but contrasts excitation with inhibition instead of light with dark:

$$EIB(t) = (Excit(t) - Inhib(t))/(Excit(t) + Inhib(t)) \quad (13)$$

where

$$Excit(t) = ON_{excit}(t) + OFF_{inhib}(t) \quad (14)$$

$$Inhib(t) = ON_{inhib}(t) + OFF_{inhib}(t) \quad (15)$$

This index varies from -1 to 1. Positive EIB values indicate a neuron's response reflects relatively stronger excitation, and negative EIB values stronger inhibition.

Orientation selectivity

We simulated the measurement of orientation selectivity to sinewave grating stimuli. We computed the fitted models' responses to sinewave gratings at each of 36 different orientations (with increments of 5 degrees), 56 different

spatial frequencies (equally spaced from 0.0667 to 0.143 cycles per image) and 36 different phases (increments of 5 degrees). These responses were used to compute the orientation selectivity of each neuron using a vector summation method (Bonhoeffer et al. 1995):

$$OB = \frac{(a^2 + b^2)^{1/2}}{\sum_{i=0}^{N-1} R(x_i)} \tag{16}$$

with

$$a = \sum_{i=0}^{N-1} R(x_i)\, cos\,(2x_i) \tag{17}$$

$$b = \sum_{i=0}^{N-1} R(x_i)\, sin\,(2x_i) \tag{18}$$

where N is the number of sinewave gratings x_i, and $R(x_i)$ represents the simulated responses. The orientation bias index, *OB*, was computed separately for each time lag in every neuron.

Results

By fitting a simple neural network model (Figure 1) to responses from individual neurons, we have estimated 2D gaussians and 3D spatiotemporal filters (dense layers) separately for ON and OFF inputs, as well as a power law output nonlinearity, for each neuron. Figure 2 shows these estimated model parameters for four example neurons, which all had peak responses at the 13-27 and 27-40 ms latencies. As we observed more generally, the early ON and OFF Gaussian filters were in each case about the same size, but opposite in polarity.

And for each neuron, the spatiotemporal filters (dense layers) were largely similar, both spatially and temporally, but opposite in polarity.

Many of the neurons had Gabor-like receptive fields that would be orientation selective (Hubel & Wiesel, 1962), like the one shown in Figure 2A. At the 27-40 ms latency, this neuron has balanced light and dark responses for both the ON (left) and OFF (right) pathways (LDB = 0.04). This balance does not occur at the 13-27 ms latency, where the neuron responds more strongly to dark stimuli (LDB = -0.2). This bias is due to the OFF pathway having stronger excitation (red) than inhibition (blue), with the ON pathway being balanced.

Another neuron (Figure 2B) is also orientation selective, balanced (LDB = 0.02) at the 27-40 ms latency, and exhibits a bias toward dark responses (LDB = -0.3) at the 13-27 ms latency. However, for this neuron the 27-40 ms latency is imbalanced due to both the ON and OFF pathways, with the ON pathway having weaker excitation and the OFF pathway having weaker inhibition.

The neuron shown in Figure 2C differs from the previous examples in that it has low orientation selectivity due to its isotropic receptive field, which has a dark center and an opposite-polarity surround. At the 13-27 ms latency, this neuron is dark-dominant (LDB = -0.34) due to its weaker surround, especially in the OFF pathway. Contrary to the above two example neurons, at the 27-40 ms latency this neuron is not balanced but light-dominant, due to stronger inhibition than excitation in the OFF pathway.

The pattern we've seen so far is that neurons are dark-dominant at the 13-27 ms latency and balanced or light-dominant at the 27-40 ms latency. However,

this is not true for all neurons. For example, the neuron in Figure 2D is light-dominant at the 13-27 ms latency, due to its Gaussian-like receptive field with a light-responsive center and a weak surround. Similar to previous results, this neuron is balanced at the 27-40 ms latency. Although neurons vary a lot in their dark and light dominance at different latencies, there is a tendency for neurons to be dark-dominant at the 13-27 ms latency and balanced or light-dominant at the 27-40 ms latency, as will be shown below.

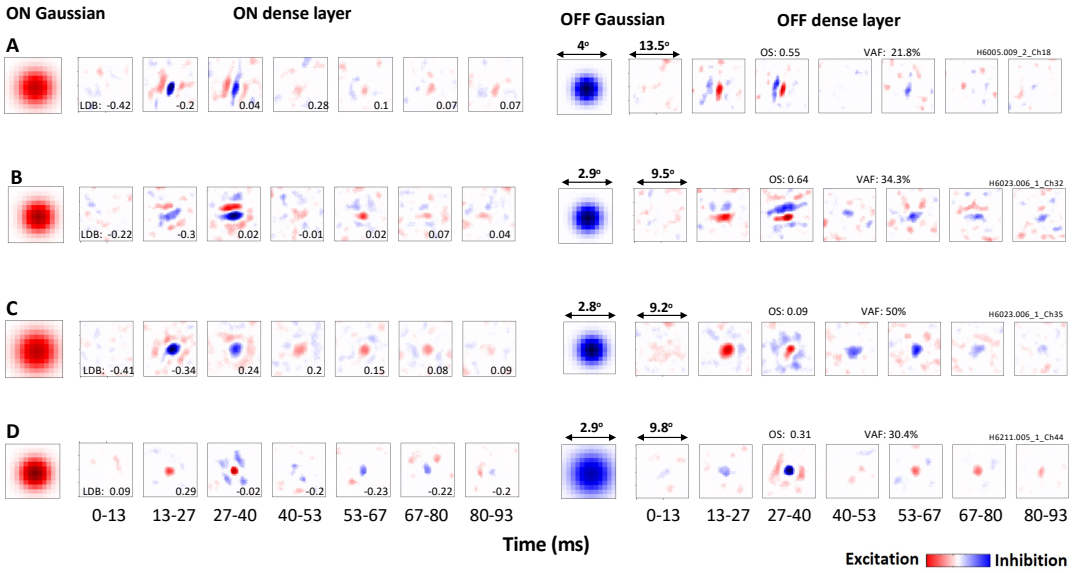

Figure 2. Gaussian filters and spatiotemporal dense layers estimated for four example neurons, one in each row. Elements of ON pathway on the left and OFF pathway on the right. Each dense layer is shown at a series of latencies ranging from 0-13 ms to 80-93 ms, with neurons being most responsive at the 13-27 ms and 27-40 ms latencies. Positive values (excitation) in red and negative values (inhibition). Orientation selectivity (OS) and variance accounted for (VAF) indicated for each neuron, and light-dark balance (LDB) values for each latency. (A, B) Neurons that respond more strongly to dark stimuli (LDB < 0) at the 13-27 ms latency, and become more balanced (LDB ~ zero) at the 27-40 ms latency. (C) Neuron that is also dark-dominant at the 13-

27 ms latency but responds more strongly to light at the 27-40 ms latency. (D) Neuron that is instead light-dominant at the 13-27 ms latency, and balanced at the 27-40 ms latency.

Population responses

To investigate whether dark-dominance is more due to ON inhibition or OFF excitation across the population of V1 neurons, we computed the sums of the four types of inputs for each neuron's optimal time lag (see Methods). Neurons had an optimal time latency of either 0-13.3 (38 neurons), 13.3-26.7 ms (35 neurons) or 26.7-40 ms (1 neuron). As described in the Methods, we estimated the overall amount of excitation and inhibition from the ON and OFF pathways, and also used these values to calculate an index of light-vs-dark dominance, LDB. We classified each neuron as dark-dominated (LDB < 0) or light-dominated (LDB > 0 depending on whether it was more responsive to dark (OFF excitation and ON inhibition) or light (ON excitation and OFF inhibition) at its optimal time latency. Across our population of 74 neurons, we found 46 neurons (62.16%) to be dark-dominated (LDB<0) and 28 neurons (37.84%) to be light-dominated (LDB>0) at their optimal latencies, similar to Yeh et al. (2009). The neurons in our sample had a wide range of LDB values (Figure 3A; minimum = -0.62, maximum = 0.57, median = -0.078), but were on average dark-dominated, with an average LDB of -0.094 (Figure 3A; $t = -3.49$, $df = 73$, $p = 0.00081$).

To better understand why cortical neurons are on average more responsive to dark than light stimuli, we next compare the four types of inputs

(Figure 3B; statistics in Table 1). ON pathway inhibition is the strongest type of input on average, and is significantly stronger than the other three (Table 1). In contrast, OFF inhibition is on average the weakest type of input, and is significantly weaker than ON inhibition and OFF excitation (but not significantly weaker than ON excitation; Table 1). The overall dark-dominance effect we observe seems to be due to this strong imbalance between ON and OFF inhibition: while inhibition is on average 37.95% stronger from the ON than from the OFF pathway (Figure 3C; Table 1), there is no significant difference between excitation from the ON and OFF pathways (Figure 3D; Table 1). The difference between ON and OFF inhibition (Figure 3C; Table 1) is significantly stronger than the difference between ON and OFF excitation (Figure 3D; Table 1). In addition, whether a given neuron is light- or dark-dominated is strongly related to whether ON inhibition exceeds OFF inhibition (Figure 3C, red points above 1:1 line vs. blue points below). However, the imbalance of ON vs. OFF excitation poorly predicts whether a neuron is light or dark-dominant (Figure 3D). Overall, these results suggest the dark-dominance effect to be more driven by differences in ON/OFF inhibition than by differences in ON/OFF excitation.

	ON inhibition	OFF excitation	OFF inhibition
ON excitation	t = -4.21 p = 7.07×10^{-5}*	t = -0.809 p = 0.421	t = 1.9 p = 0.0581
ON inhibition	X	t = 3.03	t = 3.95

		p = 0.00340*	p = 0.000178*
OFF excitation	X	X	t = 2.42 p = 0.0178

Table 1. Pairwise t-tests between each type of input at the optimal latency. Degrees of freedom, 73 for all tests.

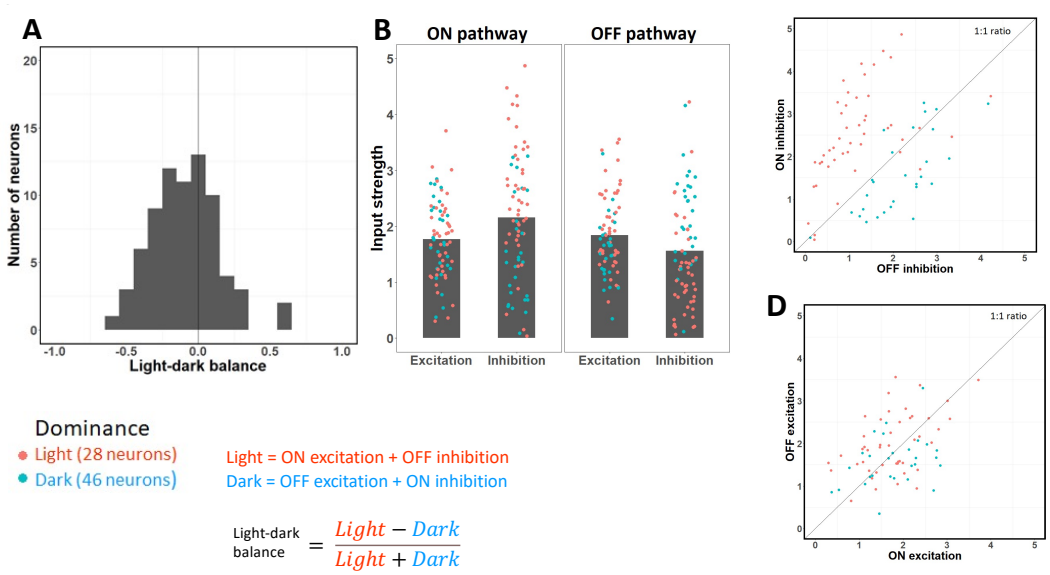

Figure 3. Strengths of excitation and inhibition from the ON and OFF pathways at each neuron's optimal time lag. (A) Distribution of light-dark balance (LDB) index values for each neuron at its optimal latency. This index is on average negative, which indicates neurons respond more strongly to dark than light stimuli. (B) Strength of excitation and inhibition across the ON and OFF pathways for each neuron, with average values shown as gray bars. Note ON inhibition is the strongest input on average. Blue dots represent light-dominant neurons and orange dots represent dark-dominant neurons. (C) Scatterplot of ON vs. OFF inhibition, for each of the 74 neurons. Most neurons have stronger ON inhibition, and whether ON or OFF inhibition is stronger is correlated with light and dark-dominance. (D) same as (C) but for ON and OFF excitation.

Unlike the result for inhibition in (C), note that ON and OFF excitation have relatively similar strength on average.

Time dynamics

Since responses to dark stimuli have previously been found to have shorter latencies than responses to light stimuli (Komban et al., 2014), we suspected the above results might vary as a function of response latency. The dependence of light-dark balance is shown for each of the measured time lags in Figure 4A, with data points for each sampled neuron, and gray bars indicating their averages. The dark-dominance effect is especially predominant at the 0-13.3 ms (one sample t-test; $t = -7.7$, $df = 73$, $p = 5.6 \times 10^{-11}$) and 13.3-26.7 ms latencies ($t = -3.9$, $df = 73$, p-value = 0.00021). The dark-dominance effect disappears at the 26.7-40 ms latency, with slightly stronger average responses to light than dark ($t = 2.48$, $df = 73$, $p = 0.015$). At the longer latencies, there was no significant average light- or dark-dominance ($p > 0.15$. These findings suggest that while V1 neurons are on average biased towards dark responses in their short latencies, the somewhat longer response latencies are instead slightly biased towards light responses.

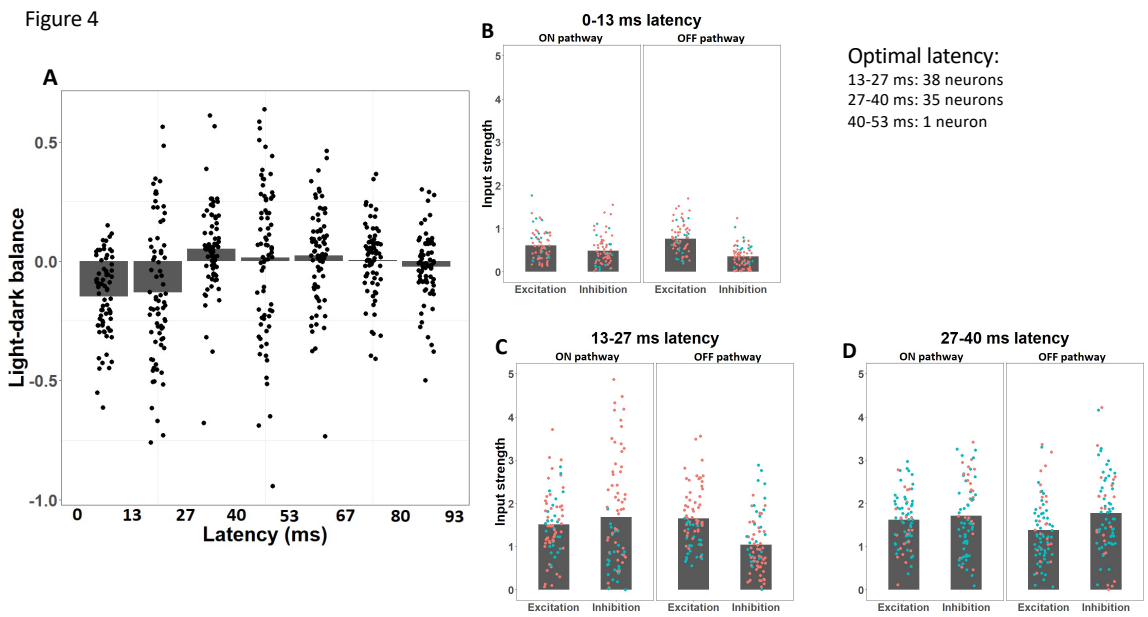

Figure 4: Light-dark balance index and strength of excitation/inhibition of ON and OFF pathways for all neurons, across different latencies. (A) Light-dark balance index values, shown as bar graph of average values for each latency, with superimposed data points for individual neurons. The 0-13.3 and 13.3-26.7 ms latencies exhibit dark-dominance, the 26.7-40 ms latency shows light-dominance and the later latencies are relatively balanced. (B) Excitation and inhibition from the ON and OFF pathways for each neuron at the 0-13.3 ms latency. OFF excitation is stronger than ON excitation on average, and inhibition is significantly weaker than excitation at this latency. (C) Same as (B) but for the 13.3-26.7 ms latency - note the relatively balanced values on average, except for OFF inhibition which is significantly weaker than the other three types of input. (D) Same as (B,C,D) but for the 26.7-40 ms latency - note the significantly weaker OFF excitation on average compared to the other three types of input.

To understand why neurons are dark-dominated in their early latencies, we investigate how the strength of each input type varies as a function of time. As we can see in Figure 4B, at a latency of 0-13.3 ms OFF excitation is the strongest input on average - it is 25.5% stronger than ON excitation (t = 5.5, df =

73, $p = 5.95 \times 10^{-7}$). Inhibition is significantly weaker than excitation, both in the ON ($t = 4.35$, $df = 73$, $p = 4.36 \times 10^{-5}$) and OFF ($t = 13.1$, $df = 73$, $p < 2.2 \times 10^{-16}$) pathways. However, this discrepancy is stronger in the OFF than in the ON pathway ($t = 7.51$, $df = 73$, $p = 1.17 \times 10^{-10}$). ON Inhibition is on average 37.9% stronger from the ON than from the OFF pathway ($t = 5.35$, $df = 73$, $p = 9.76 \times 10^{-7}$), thereby contributing to stronger dark responses. These results are consistent with findings from Jin et al. (2008), who demonstrated stronger OFF than ON excitation at short latencies to be an important mechanism contributing to the dark-dominance phenomenon.

Responses at the 13.3-26.7 ms latency are also stronger to dark (Figure 4C), but for a different reason. OFF excitation is not significantly stronger than ON excitation at the 13.3-26.7 ms latency ($t = 1.65$, $df = 73$, $p = 0.103$). Instead, dark-dominance at this latency seems to be due to weaker OFF inhibition compared to the other three types of inputs. Inhibition from the OFF pathway is on average 38.3% weaker than inhibition from the ON pathway ($t = 4.58$, $df = 73$, $p = 1.88 \times 10^{-5}$). OFF inhibition is also on average 31.2% weaker than ON excitation ($t = 7.67$, $df = 73$, $p = 5.77 \times 10^{-11}$) and on average 36.9% weaker than OFF excitation ($t = 5.84$, $df = 73$, $p = 1.37 \times 10^{-7}$). No other pair of inputs are significantly different from each other ($p < 0.05$) at the 13.3-26.7 ms latency, further strengthening the idea that the imbalance between light and dark responses at this latency is due to weaker OFF inhibition.

Contrary to the results for the previous two latencies, the 26.7-40 ms latency is slightly light-dominated, with OFF excitation being weaker than the

other types of inputs (Figure 4D). OFF excitation is 15% weaker than ON excitation ($t = 2.57$, $df = 73$, $p = 0.121$), 22.4% weaker than OFF inhibition ($t = 5.37$, $df = 73$, $p = 9.12 \times 10^{-7}$) and 19.7% weaker than ON inhibition ($t = 3.8$, $df = 73$, $p = 0.0003$). No other pair of inputs are significantly different from each other ($p < 0.05$) at this latency, and thus weaker OFF excitation seems to be the main reason why responses at the 26.7-40 ms latency are light-dominated.

To further understand the time dynamics of dark and light responses, we next analyze how the strength of each type of input changes across time lags (Figure 5), in particular the two latencies with the strongest responses, 13.3-26.7 ms and 26.7-40 ms, which the above analysis showed to have the greatest differential effects. Consistent with the above results, OFF inhibition is 41.4% weaker at 13.3-26.7 ms than at 26.7-40 ms (Figure 5A; $t = 8.14$, $df = 73$, $p = 7.63 \times 10^{-12}$). OFF excitation is 16.6% weaker at 26.7-40 ms than at 13.3-26.7 ms (Figure 5B; $t = 2.11$, $df = 73$, $p = 0.0381$). There are no significant differences between the 13.3-26.7 ms and 26.7-40 ms latencies for both ON inhibition (Figure 5C; $t = 0.197$, $df = 73$, $p = 0.845$) and ON excitation (Figure 5D; $t = 1.15$, $df = 73$, $p = 0.256$). These findings support the previous results, where dark-dominance at the 13.3-26.7 ms is due to weaker OFF inhibition and light-dominance at the 26.7-40 ms latency is due to weaker OFF excitation.

Figure 5

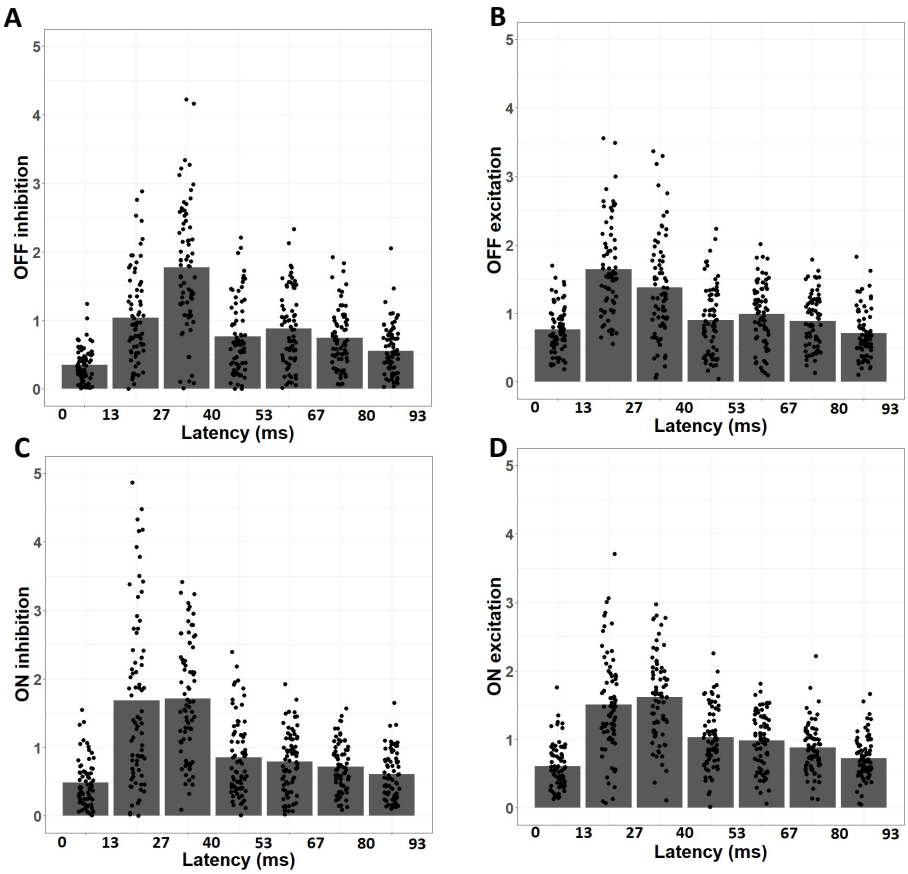

Figure 5. Temporal dependence of contributions from ON / OFF excitation and inhibition. (A) Bar graph of average OFF inhibition strength across time lags, with data points indicating values for individual neurons. Note OFF inhibition is weaker at the 13.3-26.7 ms latency than at the 26.7-40 ms latency. (B) Same as (A) but for OFF excitation - note it is weaker at the 26.7-40 ms latency than at 13.3-26.7 ms. (C) Same as (A-B) but for ON inhibition. (D) Same as (A-C) but for OFF excitation. Note that both have similar strength at the 13.3-26.7 and 26.7-40 ms latencies.

Orientation selectivity

Previous studies demonstrated that V1 neurons are less orientation selective in their early responses (Ringach, Hawken, Shapley, 1997; Shapley et al, 2003). Since we have found early latencies to respond more strongly to dark stimuli, we wondered how the degree of dark-dominance might relate to orientation selectivity. To infer orientation selectivity, we simulated the responses of the neurons' fitted models to static sinewave grating stimuli with a series of orientations, spatial frequencies, and phases. For each latency, we select the sinewave gratings with the best phase and spatial frequency for each orientation. We then use the model's responses of these sinewave gratings to measure orientation selectivity, using a conventional vector summation method (see Methods). Orientation selectivity for each neuron across latencies can be seen in Figure 6A. We find orientation selectivity to increase as the time latency increases, up to the 26.7-40 ms latency. Orientation selectivity is significantly higher at 13.3-26.7 ms than at 0-13.3 ms (paired t-test; $t = 2.2$, $df = 73$, $p = 0.033$), and also significantly higher at 26.7-40 ms than at 13.3-26.7 ms ($t = 6.1$, $df = 73$, $p = 5.2 \times 10^{-8}$).

We next investigated the relationship between orientation selectivity and light-dark balance at each neuron's optimal latency, which can be seen in Figure 6B. Neurons having very little orientation selectivity tend to be dark dominant (LDB << 0) or (rarely) light-dominant (LDB >> 0), while those that are more orientation selective are more often relatively light-dark balanced (LDB ~ 0). This apparent relationship is confirmed statistically: there is a significant negative

relationship (r = -0.45) between orientation selectivity and absolute values of LDB (t = -4.4, df = 72, p = 4 x 10^{-5}). These results suggest a response bias towards dark stimuli might reduce a neuron's orientation selectivity (Figure 6B), especially at the 0-13.3 and 13.3-26.7 ms latencies (Figure 6A).

Figure 6

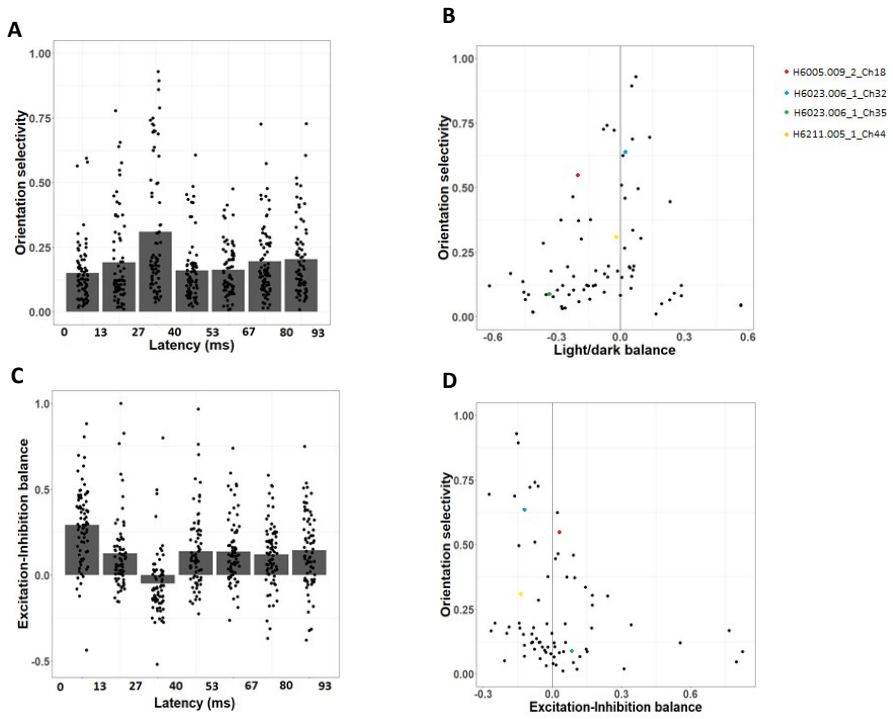

Figure 6. Changes in orientation selectivity and excitation-inhibition balance across latencies. (A) Average orientation selectivity peaks at the 26.7-40 ms latency, and is relatively low at the 0-13.3 and 13.3-26.7 ms latencies. (B) Relationship between orientation selectivity (ordinate) and light-dark balance (abscissa). Neurons with greater orientation selectivity tend to be more balanced. (C) Excitation-Inhibition balance (EIB) index as a function of latency. Excitation is stronger than inhibition at the 0-13.3 and 13.3-26.7 ms latencies, while excitation and inhibition are relatively balanced at the 26.7-40 ms latency. (D) Relationship between orientation selectivity (ordinate) and EIB (abscissa). Neurons with stronger excitation than inhibition tend to be less orientation selective.

Another possible explanation for weaker orientation selectivity at early latencies (Ringach et al., 1997; Shapley et al., 2003) is faster excitation than inhibition. Figure 6C shows the relative amount of excitation vs. inhibition (EIB index; see Methods) at each latency. Excitation is stronger than inhibition at the 0-13.3 ms ($t = 11.3$, $df = 73$, $p < 2.2 \times 10^{-16}$) and 13.3-26.7 ms latencies ($t = 5.1$, $df = 73$, $p = 2.8 \times 10^{-6}$), while inhibition is slightly stronger than excitation at the 26.7-40 ms latency ($t = -2.3$, $df = 73$, $p = 0.025$). Inhibition becomes relatively stronger over time, with higher EIB values for the 0-13.3 than for the 13.3-26.7 ms latency ($t = 4.9$, $df = 73$, $p = 5.73 \times 10^{-6}$). EIB values are also higher for the 13.3-26.7 than for the 26.7-40 ms latency ($t = 7.44$, $df = 73$, $p = 1.58 \times 10^{-10}$). Also consistent with Ringach et al. (1997) and Shapley et al. (2003), we find a negative correlation of $r = -0.26$ between orientation selectivity and EIB (Figure 6D). Overall, these results suggest that both dark-dominance and stronger excitation contribute to weaker orientation selectivity at early latencies. However, another interpretation might be that weaker OFF inhibition is responsible for stronger dark responses, weaker overall inhibition and weaker orientation selectivity at early latencies (see Discussion).

Discussion

Estimating excitation and inhibition from machine learning

In this study we use a machine learning algorithm based on known neurobiology, which makes our model better suited for system identification of V1 neurons. Like the LGN, the model is separated into ON and OFF pathways, each composed of Gaussian linear filters followed by half-wave rectifiers. The weaker surrounds of LGN neurons are omitted to help the model converge on a set of parameter values. Using this method, we can distinguish between excitation and inhibition to light and dark stimuli to study the dark-dominance effect. Even though our model architecture is inspired from known neurobiology, the excitation and inhibition we estimate does not necessarily reflect direct LGN inputs. This is especially true for inhibition, as there are no direct inhibitory inputs from the LGN to V1 (Ferster & Lindström 1983, Martin & Whitteridge 1984). Inhibition to V1 neurons arise from inhibitory interneurons, who either relay laterogeniculate inputs or receive inputs from other V1 neurons (Isaacson & Scanziani, 2011). Although V1 neurons directly receive laterogeniculate excitation, there is also intracortical excitation within V1 (Teich & Quian, 2013). The circuitry within V1 is very complex, and our model architecture is definitely an oversimplification of the inputs the neurons receive. For this reason, we interpret the estimated excitation and inhibition as a measure of how a neuron's response varies as a function of light and dark stimuli. This interpretation comes from how the model architecture separately processes information above (ON) and below (OFF) the mean luminance level, which means it is well-suited to measure the asymmetry

between responses to light and dark stimuli. For example, a neuron might increase its response when there is dark in a part of its receptive field, but not decrease its response when there is light in the same region (see Figure 2A). Making this distinction is possible thanks to our model architecture combined with the use of rich stimuli such as natural images.

Time dynamics of dark-dominance

Primary visual cortex neurons have previously been shown to respond more strongly to dark than light stimuli (Jin et al., 2008; Yeh et al., 2009), with responses to dark stimuli having 3-6 ms shorter latencies than responses to light stimuli (Jin et al., 2011). However, the mechanisms and properties of this dark-dominance effect are still poorly understood. Here we use a machine learning approach combined with natural image stimuli to infer the excitatory and inhibitory inputs to V1 neurons. We replicate the dark-dominance effect, stronger OFF than ON excitation at very short latencies (0-13.3 ms; Jin et al., 2008), and weaker orientation selectivity at shorter latencies (Shapley et al., 2003).

A novel finding of this study is how the dark-dominance changes as a function of latency. Even though we replicate the dark-dominance effect at the 0-13.3 and 13.3-26.7 latencies, we instead find a slight light-dominance at the 13.3-26.7 latency. We were able to find this relationship between latency and dark-dominance because we estimate and analyze light and dark responses at every latency for each neuron. Other studies focus on each neuron's optimal latency

(Yeh et al., 2009), which still clearly shows the dark-dominance effect (Figure 3) but ignores the effect of latency on the strength of dark responses. Because of this, dark responses were thought to be stronger in general, whereas we find that effect to be specific to the earlier latencies.

This relationship between dark-dominance and latency is not so surprising, considering dark-dominant V1 neurons have previously been found to respond 3-6 ms faster than light-dominant neurons (Komban et al., 2014). These faster dark responses in V1 have been attributed to faster OFF than ON LGN responses (Jin et al., 2011). While we do find the 0-13.3 ms latency to be dark-dominant due to stronger OFF than ON excitation (Figure 4B), most neurons have poor responses at this latency. The dark-dominance effect is most salient at the 13.3-26.7 ms latency, where stronger responses are due to weaker inhibition to dark stimuli (Figure 4C). These results are consistent with findings from Taylor et al. (2018), who found intracortical inhibition to be stronger for light than for dark stimuli. Therefore, the dark-dominance results at each neuron's optimal time lag from Yeh et al. (2009) and Mazade et al. (2021) are mostly due to weaker inhibition to dark stimuli rather than stronger excitation to dark stimuli.

These findings help explain why dark-dominance is strongest in layer 2/3 of primate V1 (Yeh et al., 2009). If dark-dominance were almost only due to stronger lateral geniculate excitation from the OFF pathway, we would expect dark-dominance to be at least just as strong in layer IV than in the other layers, since this is where most of LGN neurons synapse. While layer IV does show dark-dominance with two-thirds of its neurons being dark-dominant, this effect is

much stronger in layers 2/3 where almost every neuron is dark-dominant (Yeh et al., 2009). This laminar difference might be due to pyramidal neurons in layers 2/3 receiving extensive inhibition as has been shown in the mouse (Kätzel et al., 2011), with inhibition being stronger to light than dark stimuli (Taylor et al., 2018).

Relationship with orientation selectivity

This study also brings a new perspective on the intracortical mechanisms of orientation selectivity, and helps explain why V1 neurons are less orientation-selective in their early time lags (Shapley et al., 2003). It was previously proposed for this phenomenon to be due to the intracortical inhibition needing to go through an inhibitory interneuron and therefore being slower than laterogeniculate excitation (Shapley et al., 2003), and we do find inhibition strength to be positively correlated with orientation selectivity (Figure 6C). We also find neurons with balanced light/dark responses to be more orientation selective (Figure 6A). Consistent with these results, the earliest measured latency (0-13.3 ms) has on average poor orientation selectivity, weak inhibition and is stronger responses to dark stimuli (Figures 4A and 4B). The 13.3-26.7 latency also has poor orientation selectivity and is dark-dominant. However, inhibition to light is just as strong as excitation at this latency, while inhibition to dark is weaker to all three other types of inputs (Figure 4C). Therefore, it seems slower inhibition to dark than to light stimuli causes the 13.3-26.7 latency to be both dark-dominant and less orientation selective.

In conclusion, we used a machine learning approach combined with natural images to infer the excitation and inhibition to light and dark stimuli of the

primary visual cortex. First, we first found stronger dark than light responses to be more strongly due to inhibition from light stimuli than to excitation from dark stimuli. Secondly, we showed the dark-dominance effect to only occur at the early latencies of V1 responses. Lastly, our findings suggest dark-dominance might be related to lower orientation selectivity in the early latencies. This research brings us one step closer to understanding the intricate information processing achieved by the primary visual cortex.

References

Anzai, A., Ohzawa, I., & Freeman, R. D. (1999). Neural mechanisms for processing binocular information II. Complex cells. J Neurophysiol, *82(2),* 909-924.

Bonhoeffer, T., Kim, D. S., Malonek, D., Shoham, D., & Grinvald, A. (1995). Optical imaging of the layout of functional domains in area 17 and across the area 17/18 border in cat visual cortex. *European Journal of Neuroscience*, *7*(9), 1973-1988.

Brainard, D. H. (1997) The Psychophysics Toolbox, *Spatial Vision 10*:433-436.

Buchner, A., & Baumgartner, N. (2007). Text–background polarity affects performance irrespective of ambient illumination and colour contrast. *Ergonomics*, *50*(7), 1036-1063.

Cooper, E. A., & Norcia, A. M. (2015). Predicting cortical dark/bright asymmetries from natural image statistics and early visual transforms. *PLoS Comput Biol*, *11*(5), e1004268.

Croner, L. J., & Kaplan, E. (1995). Receptive fields of P and M ganglion cells across the primate retina. *Vision Research*, *35*(1), 7-24.

Ferster, D., & Miller, K. D. (2000). Neural mechanisms of orientation selectivity in the visual cortex. *Annual Review of Neuroscience*, *23*(1), 441-471.

Ferster, D., & Lindström, S. (1983). An intracellular analysis of geniculo-cortical connectivity in area 17 of the cat. *The Journal of Physiology*, *342*(1), 181-215.

Heeger, D. J. (1991). Nonlinear model of neural responses in cat visual cortex. In: Computational Models of Visual Processing, 119-133.

Hirsch, J. A., Martinez, L. M., Pillai, C., Alonso, J. M., Wang, Q., & Sommer, F. T. (2003). Functionally distinct inhibitory neurons at the first stage of visual cortical processing. *Nature Neuroscience*, *6*(12), 1300-1308.

Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, *12*(1), 55-67.

Isaacson, J. S., & Scanziani, M. (2011). How inhibition shapes cortical activity. *Neuron*, *72*(2), 231-243.

Jansen, M., Jin, J., Li, X., Lashgari, R., Kremkow, J., Bereshpolova, Y., ... & Alonso, J. M. (2019). Cortical balance between ON and OFF visual responses is modulated by the spatial properties of the visual stimulus. *Cerebral Cortex*, *29*(1), 336-355.

Jin, J. Z., Weng, C., Yeh, C. I., Gordon, J. A., Ruthazer, E. S., Stryker, M. P., ... & Alonso, J. M. (2008). On and off domains of geniculate afferents in cat primary visual cortex. *Nature Neuroscience*, *11*(1), 88-94.

Jin, J., Wang, Y., Lashgari, R., Swadlow, H. A., & Alonso, J. M. (2011). Faster thalamocortical processing for dark than light visual targets. *Journal of Neuroscience*, *31*(48), 17471-17479.

Kätzel, D., Zemelman, B. V., Buetfering, C., Wölfel, M., & Miesenböck, G. (2011). The columnar and laminar organization of inhibitory connections to neocortical excitatory cells. *Nature Neuroscience*, *14*(1), 100-107.

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Kleiner M, Brainard D, Pelli D, 2007, "What's new in Psychtoolbox-3?" Perception 36 ECVP Abstract Supplement.

Komban, S. J., Alonso, J. M., & Zaidi, Q. (2011). Darks are processed faster than lights. *Journal of Neuroscience*, *31*(23), 8654-8658.

Komban, S. J., Kremkow, J., Jin, J., Wang, Y., Lashgari, R., Li, X., ... & Alonso, J. M. (2014). Neuronal and perceptual differences in the temporal processing of darks and lights. *Neuron*, *82*(1), 224-234.

Kremkow, J., Jin, J., Komban, S. J., Wang, Y., Lashgari, R., Li, X., ... & Alonso, J. M. (2014). Neuronal nonlinearity explains greater visual spatial resolution for darks than lights. *Proceedings of the National Academy of Sciences*, *111*(8), 3170-3175.

Li, G., Yang, Y., Liang, Z., Xia, J., & Zhou, Y. (2008). GABA-mediated inhibition correlates with orientation selectivity in primary visual cortex of cat. *Neuroscience*, *155*(3), 914-922.

Martin, K. A., & Whitteridge, D. (1984). Form, function and intracortical projections of spiny neurones in the striate visual cortex of the cat. *The Journal of Physiology*, *353*(1), 463-504.

Mazade, R., Jin, J., Pons, C., & Alonso, J. M. (2019). Functional specialization of ON and OFF cortical pathways for global-slow and local-fast vision. *Cell Reports*, *27*(10), 2881-2894.

Mechler, F., & Ringach, D. L. (2002). On the classification of simple and complex cells. *Vision Research*, *42*(8), 1017-1033.

Molnar, A., Hsueh, H. A., Roska, B., & Werblin, F. S. (2009). Crossover inhibition in the retina: circuitry that compensates for nonlinear rectifying synaptic transmission. *Journal of Computational Neuroscience*, *27*(3), 569.

Monier, C., Chavane, F., Baudot, P., Graham, L. J., & Frégnac, Y. (2003). Orientation and direction selectivity of synaptic inputs in visual cortical neurons: a diversity of combinations produces spike tuning. *Neuron*, *37*(4), 663-680.

Montero, V. M. (1986). The interneuronal nature of GABAergic neurons in the lateral geniculate nucleus of the rhesus monkey: a combined HRP and GABA-immunocytochemical study. *Experimental Brain Research*, *64*(3), 615-622.

Nelson, R., & Connaughton, V. (2012). Bipolar cell pathways in the vertebrate retina. In *Webvision: The Organization of the Retina and Visual System [https://www.ncbi.nlm.nih.gov/books/NBK11521/]*. University of Utah Health Sciences Center.

Olmos, A., & Kingdom, F. A. (2004). A biologically inspired algorithm for the recovery of shading and reflectance images. *Perception*, *33*(12), 1463-1473.

Pascanu, R., Mikolov, T., & Bengio, Y. (2012). Understanding the exploding gradient problem. *CoRR, abs/1211.5063*, *2*, 417.

Pelli, D. G. (1997) The VideoToolbox software for visual psychophysics: Transforming numbers into movies, *Spatial Vision 10*:437-442.

Persi, E., Hansel, D., Nowak, L., Barone, P., & Van Vreeswijk, C. (2011). Power-law input-output transfer functions explain the contrast-response and tuning properties of neurons in visual cortex. *PLoS Comput Biol*, *7*(2), e1001078.

Hu, H., Gan, J., & Jonas, P. (2014). Fast-spiking, parvalbumin+ GABAergic interneurons: From cellular design to microcircuit function. *Science*, *345*(6196).

Ratliff, C. P., Borghuis, B. G., Kao, Y. H., Sterling, P., & Balasubramanian, V. (2010). Retina is structured to process an excess of darkness in natural scenes. *Proceedings of the National Academy of Sciences*, *107*(40), 17368-17373.

Reiner, A., & Levitz, J. (2018). Glutamatergic signaling in the central nervous system: ionotropic and metabotropic receptors in concert. *Neuron*, *98*(6), 1080-1098.

Ringach, D. L., Hawken, M. J., & Shapley, R. (1997). Dynamics of orientation tuning in macaque primary visual cortex. *Nature*, *387*(6630), 281-284.

Ringach, D. L., Sapiro, G., & Shapley, R. (1997). A subspace reverse-correlation technique for the study of visual neurons. *Vision Research*, *37*(17), 2455-2464.

Sclar, G., & Freeman, R. D. (1982). Orientation selectivity in the cat's striate cortex is invariant with stimulus contrast. *Experimental Brain Research*, *46*(3), 457-461.

Shapley, R., Hawken, M., & Ringach, D. L. (2003). Dynamics of orientation selectivity in the primary visual cortex and the importance of cortical inhibition. *Neuron*, *38*(5), 689-699.

Skottun, B. C., Bradley, A., Sclar, G., Ohzawa, I., & Freeman, R. D. (1987). The effects of contrast on visual orientation and spatial frequency discrimination: a comparison of single cells and behavior. *Journal of Neurophysiology*, *57*(3), 773-786.

Skottun, B. C., De Valois, R. L., Grosof, D. H., Movshon, J. A., Albrecht, D. G., & Bonds, A. B. (1991). Classifying simple and complex cells on the basis of response modulation. *Vision Research*, *31*(7-8), 1078-1086.

Somers, D. C., Nelson, S. B., & Sur, M. (1995). An emergent model of orientation selectivity in cat visual cortical simple cells. *Journal of Neuroscience*, *15*(8), 5448-5465.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, *15*(1), 1929-1958.

Swindale, N. V., & Spacek, M. A. (2014). Spike sorting for polytrodes: a divide and conquer approach. *Frontiers in Systems Neuroscience*, *8*, 6.

Talebi, V., & Baker, C. L. (2012). Natural versus synthetic stimuli for estimating receptive field models: a comparison of predictive robustness. *Journal of Neuroscience*, *32*(5), 1560-1576.

Taylor, M. M., Sedigh-Sarvestani, M., Vigeland, L., Palmer, L. A., & Contreras, D. (2018). Inhibition in simple cell receptive fields is broad and OFF-subregion biased. *Journal of Neuroscience*, *38*(3), 595-612.

Troyer, T. W., Krukowski, A. E., Priebe, N. J., & Miller, K. D. (1998). Contrast-invariant orientation tuning in cat visual cortex: thalamocortical input tuning and correlation-based intracortical connectivity. *Journal of Neuroscience*, *18*(15), 5908-5927.

Tucker, T. R., & Fitzpatrick, D. (2006). Luminance-evoked inhibition in primary visual cortex: a transient veto of simultaneous and ongoing response. *Journal of Neuroscience*, *26*(52), 13537-13547.

van Versendaal, D., & Levelt, C. N. (2016). Inhibitory interneurons in visual cortical plasticity. *Cellular and Molecular Life Sciences*, *73*(19), 3677-3691.

Wood, K. C., Blackwell, J. M., & Geffen, M. N. (2017). Cortical inhibitory interneurons control sensory processing. *Current Opinion in Neurobiology*, *46*, 200-207.

Yeh, C. I., Xing, D., & Shapley, R. M. (2009). “Black” responses dominate macaque primary visual cortex v1. *Journal of Neuroscience*, *29*(38), 11753-11760.