A recurrent circuit implements normalization, simulating the dynamics of V1 activity

David J. Heeger[a,b,1] and Klavdia O. Zemlianova[b]

[a]Department of Psychology, New York University, New York, NY 10003; and [b]Center for Neural Science, New York University, New York, NY 10003

The normalization model has been applied to explain neural activity in diverse neural systems including primary visual cortex (V1). The model's defining characteristic is that the response of each neuron is divided by a factor that includes a weighted sum of activity of a pool of neurons. Despite the success of the normalization model, there are three unresolved issues. 1) Experimental evidence supports the hypothesis that normalization in V1 operates via recurrent amplification, i.e., amplifying weak inputs more than strong inputs. It is unknown how normalization arises from recurrent amplification. 2) Experiments have demonstrated that normalization is weighted such that each weight specifies how one neuron contributes to another's normalization pool. It is unknown how weighted normalization arises from a recurrent circuit. 3) Neural activity in V1 exhibits complex dynamics, including gamma oscillations, linked to normalization. It is unknown how these dynamics emerge from normalization. Here, a family of recurrent circuit models is reported, each of which comprises coupled neural integrators to implement normalization via recurrent amplification with arbitrary normalization weights, some of which can recapitulate key experimental observations of the dynamics of neural activity in V1.

computational neuroscience | recurrent neural network | V1 | normalization | gamma oscillation

The normalization model was initially developed to explain stimulus-evoked responses of neurons in primary visual cortex (V1) (1–7) but has since been applied to explain neural activity and behavior in diverse cognitive processes and neural systems (see *SI Appendix* for references). The defining characteristic of normalization is that the response of each neuron is divided by a weighted sum of the activity of a pool of neurons (Fig. 1A). In V1, this normalization pool includes neurons selective for different visual stimulus features and spatial positions (i.e., receptive-field locations).

The normalization model mimics many well-documented physiological phenomena in V1 (Fig. 1 *B* and *C*) and their perceptual analogs (see *SI Appendix* for references). 1) Responses saturate (level off) when increasing the contrast of a preferred orientation test stimulus (e.g., a grating restricted to a neuron's receptive field [RF]) (Fig. 1B, blue curve). 2) Responses to a nonpreferred orientation are smaller than responses to the preferred orientation by a constant scale factor, saturating at the same contrast, not the same firing rate, for preferred and nonpreferred stimuli (Fig. 1B, orange vs. blue curves). 3) Responses to two or more stimuli presented together are much less than the linear sum of the individual responses: cross-orientation suppression when a mask stimulus (e.g., a grating of fixed contrast) that is orthogonal to the preferred orientation is superimposed with a preferred-orientation test stimulus (Fig. 1C, yellow vs. blue curves); and surround suppression when a mask stimulus is added in the region surrounding a neuron's RF. Different stimuli suppress responses by different amounts (see *SI Appendix* for references), suggesting that normalization is "tuned." The normalization weights specify the contribution of one neuron to another's normalization pool, determining the tuning.

Normalization has been shown to serve a number of functions in a variety of neural systems including automatic gain control (needed because of limited dynamic range), simplifying readout, conferring invariance with respect to one or more stimulus dimensions (e.g., contrast, odorant concentration), switching between averaging vs. winner-take-all, contributing to decorrelation and statistical independence of neural responses, stabilizing delay-period activity, and facilitating learning (see *SI Appendix* for references).

Neural activity in V1 exhibits complex dynamics linked to normalization. The rate of response increase following stimulus onset is typically faster than the decrease following stimulus offset (8). The rate of response increase is also stimulus dependent: faster for high-contrast stimuli and for stimuli in the center of the RF (8). The timing of response suppression depends on its strength (9). Temporal-frequency tuning depends on stimulus contrast, and simple-cell response phase depends on contrast (6, 10–13). Complex dynamics are evident also in the combined activity (e.g., as measured with local field potentials [LFPs]) of populations of neurons. LFPs exhibit gamma oscillations (∼30 to 80 Hz) that have been linked to normalization (14–16). Oscillation amplitude and frequency depend systematically on stimulus contrast, size, and spatial pattern (14, 15, 17–30).

The circuit mechanisms underlying normalization are not well understood. Experimental evidence supports the hypothesis that normalization operates via recurrent amplification, i.e., amplifying weak inputs more than strong inputs (31–34). The recurrent amplification hypothesis is also supported by anatomy: cortical circuits are dominated by recurrent connections (35–40). We have known since we first introduced the normalization model that it can be implemented in a recurrent circuit (4, 5). Since then,

Significance

A family of recurrent circuit models is proposed to explain the dynamics of neural activity in primary visual cortex (V1). Each of the models in this family exhibits steady-state output responses that are already known to fit a wide range of experimental data from diverse neural systems. These models can recapitulate the complex dynamics of V1 activity, including oscillations (so-called gamma oscillations, ∼30 to 80 Hz). This theoretical framework may also be used to explain key aspects of working memory and motor control. Consequently, the same circuit architecture is applicable to a variety of neural systems, and V1 can be used as a model system for understanding the neural computations in many brain areas.
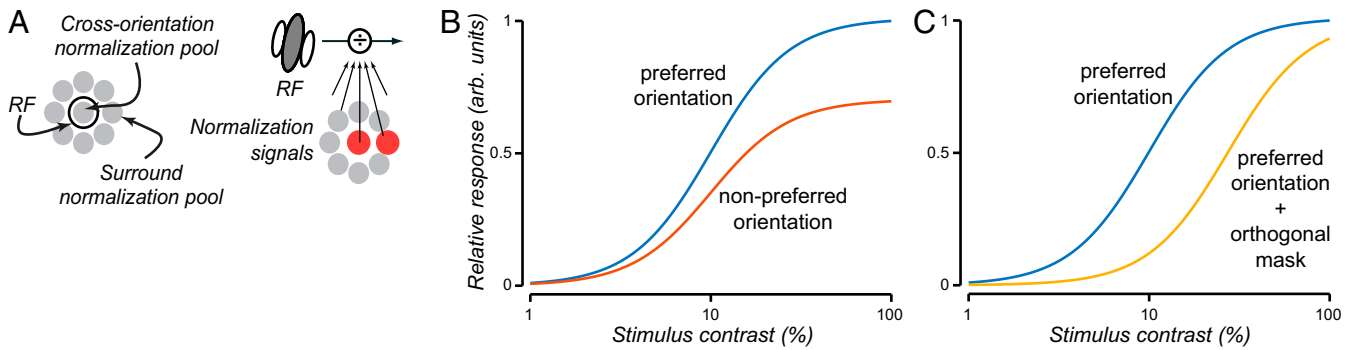
Fig. 1. Normalization model. (*A*) Conceptual diagram of normalization in which the neuron's response is suppressed by a weighted sum of activity of a pool of neurons. (*B*) Response saturation. Blue, responses to preferred stimulus orientation saturate at high contrasts. Orange, responses to nonpreferred orientation are a scaled-down copy of the responses to preferred orientation, saturating at the same contrast. (*C*) Cross-orientation suppression. Blue, preferred orientation. Yellow, superimposing an orthogonal stimulus (fixed contrast) suppresses responses to the preferred orientation. A similar result would be observed by adding a stimulus component that surrounds the preferred stimulus (surround suppression).

several hypotheses for the circuit mechanisms underlying normalization have been proposed, including shunting inhibition, synaptic depression, and inhibition-stabilized networks (6, 13, 41–46). See also refs. 47–49 for precursors to these circuit models. However, these models do not rely on recurrent amplification to achieve normalization and/or they do not exhibit complex dynamics (including gamma oscillations) linked to normalization (*Discussion*). Furthermore, these previous models only approximate weighted normalization; this has practical consequences for making experimentally testable predictions and for fitting data (*Discussion*).

Here, we introduce and characterize a family of dynamical systems that implement normalization with recurrent amplification. When the input drive is constant over time, each of the recurrent circuits in this family exhibits output responses that follow the normalization equation exactly, with arbitrary (nonnegative) normalization weights. Each model in this family is expressed as a coupled system of neural integrators, composed of two classes of neurons: principal cells and modulator cells. The key idea is that the amount of recurrent amplification in the principal cells depends inversely on the responses of the modulator cells. When the input is weak, the modulator cells have small responses and there is a large amount of recurrent amplification. When the input is strong, the modulator cell responses are large, which shuts down the recurrent amplification. The various models in this family of dynamical systems imply different circuits, some of which recapitulate the complex dynamics of V1 activity, including gamma oscillations. Although we focus on V1, this family of models is applicable to many neural systems (*Discussion*).

A preliminary version was posted on a preprint server (50). MATLAB code is available at hdl.handle.net/2451/61045 (51).

Results

Recurrent Normalization. Following our previous work (52, 53), responses of a population of V1 neurons are modeled as dynamical processes that evolve over time in a recurrent circuit (Fig. 2). The output firing rates of the principal cells depend on the sum of two terms: 1) input gain (Fig. 2, orange) multiplied by input drive (Fig. 2, blue), and 2) recurrent gain (Fig. 2, purple) multiplied by recurrent drive (Fig. 2, green). The input drive is a weighted sum of the responses of the population of input neurons, and the input gain is specified by a constant. These input neurons are presumed to be in the lateral geniculate nucleus (LGN) of the thalamus, which, in turn, receive their inputs from neurons in the retina. The recurrent drive is a weighted sum of principal cell responses, and the recurrent gain depends on the modulator cell responses ("modulator" refers to a multiplicative

computation, not to neuromodulators). Modulator cell responses depend on the principal cell responses (Fig. 2, purple).

There are two nested recurrent loops that oppose each other. 1) Recurrent drive: The recurrent drive is a weighted sum of the principal cell responses, and the principal cell responses depend on the recurrent drive (Fig. 2, green). 2) Recurrent gain: The recurrent gain depends inversely on the modulator cell responses, and the modulator cell responses depend on a sum of principal cell responses (Fig. 2, purple). The recurrent drive is
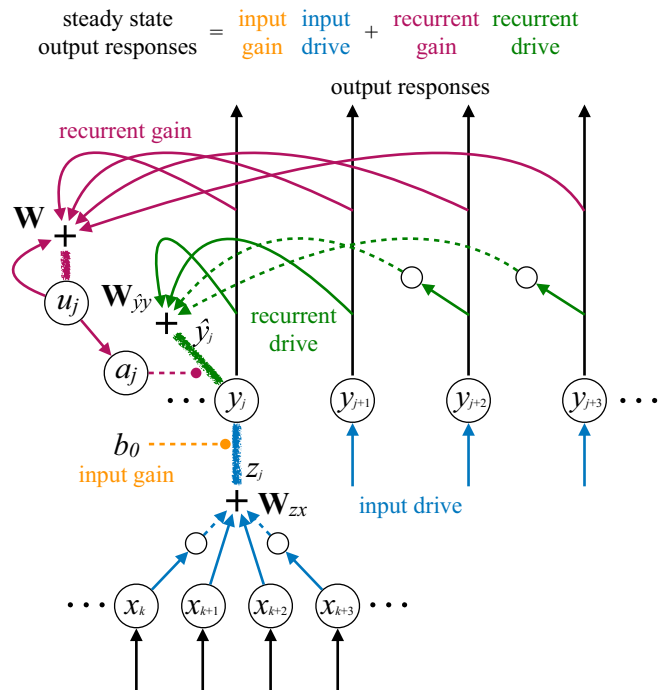
Fig. 2. Recurrent circuit. Orange, input gain. Blue, input drive. Purple, recurrent gain. Green, recurrent drive. The solid circles represent neurons. Different cell types: x_k, LGN inputs; y_j, principal cells; u_j and a_j, modulator cells; small circles, inhibitory interneurons. Thin lines with arrowheads, axons. Solid lines, excitatory connections. Dashed lines, inhibitory connections. Circle heads, modulatory (e.g., shunting) connections. Thick fuzzy lines without arrowheads, dendritic compartments. Synaptic weights: $\mathbf{W}_{zx}$, orientation-selective weights determine input drive; $\mathbf{W}_{yy}$, recurrent weights; $\mathbf{W}$, normalization weights. Dendritic computations (sum of synaptic currents): z_j, input drive; $\hat{y}_j$, recurrent drive.

Heeger and Zemlianova

www.manaraa.com

multiplied by the recurrent gain so that the modulators control the amount of recurrent amplification. Increasing the principal cell responses causes the modulator cells to increase their responses, which causes the amount of recurrent amplification to decrease. Therefore, as the activity of the principal cells increase, the first recurrent loop increases the amount of recurrent amplification, while the second loop decreases the amount of recurrent amplification. These two recurrent loops oppose each other such that the activity of the circuit may achieve a fixed point at which the neural activity is normalized. The responses at this fixed point typically exhibit one of two kinds of dynamics. If the modulator cells are sluggish, then the principal cells can exhibit an initial transient overshoot before achieving the fixed point. If instead the modulator cells have a short time constant and a delay, then the fixed point may be unstable and the responses may exhibit oscillations.

This circuit agrees with experimental results suggesting that normalization operates via recurrent amplification (31–34). The recurrent drive involves both excitation and inhibition (Fig. 2, green solid and dashed lines, respectively). The modulator cells control the amount of recurrent amplification (Fig. 2, purple line with circle head). Both excitatory and inhibitory recurrent signals are amplified by an amount that is controlled by the modulator cells (Fig. 2, purple).

The remainder of this subsection walks through the equations of the dynamical system corresponding to the circuit model in Fig. 2 (see *SI Appendix* for additional details). We then demonstrate that this model mimics experimental observations of the dynamics of neural activity. We present the model as a computational theory for the computations performed by neural circuits in V1 (but see Table 1 and *Discussion* for possible mechanisms).

Principal cell responses are as follows (see SI appendix of ref. 53 for a primer on neural integrators and Table 1 for mathematical symbols):

$$\tau_v \frac{dv_j}{dt} = -v_j + \left(\frac{b_0}{1+b_0}\right)z_j + \left(\frac{1}{1+a_j}\right)\hat{y}_j, \qquad [1]$$

$$y_j = \lfloor v_j \rfloor^2, \qquad [2]$$

$$\mathbf{z} = \mathbf{W}_{zx}\mathbf{x}, \qquad [3]$$

$$\hat{\mathbf{y}} = \mathbf{W}_{\hat{y}y}\sqrt{\mathbf{y}}. \qquad [4]$$

Vector $\mathbf{y} = (y_1, y_2, \ldots, y_j, \ldots, y_N)$ represents the firing rate responses of the principal cells, where the subscript j indexes different neurons in the population, with different RF centers, orientation preferences, and spatial and temporal phases. The underlying membrane potentials of the principal cells are represented by vector $\mathbf{v}$. Membrane potential of the jth principal cell v_j depends on a sum of two terms (Eq. 1): 1) input gain multiplied by input drive z_j and 2) recurrent gain multiplied by recurrent drive $\hat{y}_j$. The input drive z_j is a weighted sum of LGN inputs (Eq. 3 and Fig. 2, blue; see *SI Appendix* for details). The rows of the weight matrix $\mathbf{W}_{zx}$ determine the spatial RFs of the simple cells (*SI Appendix*, Fig. S1 *B–D* and see *SI Appendix* for details). The recurrent drive $\hat{y}_j$ is a weighted sum (with recurrent weights $\mathbf{W}_{\hat{y}y}$) of the square root of the principal cell responses y_j (Eq. 4 and Fig. 2, green; see *SI Appendix* for details). The input drive and the recurrent drive are each multiplied by a gain factor. The input gain is specified by a constant b_0. The recurrent gain depends on the responses of the modulator cells a_j, as detailed below. Half-squaring (half-wave rectification and squaring) in Eq. 2 is an expansive nonlinearity that approximates the transformation from the membrane potential of the principal cells to their firing rates. The square root in Eq. 4 is a compressive nonlinearity that approximates a transformation from firing rates to synaptic currents.

Table 1. Mathematical notation

Symbol	Description	Possible mechanism
$\mathbf{x} = (x_1, x_2, \ldots, x_i, \ldots, x_M)$	Inputs	Firing rates of LGN cells
$\mathbf{y} = (y_1, y_2, \ldots, y_j, \ldots, y_N)$	Principal cell responses	Firing rates of pyramidal cells
$\mathbf{v} = (v_1, v_2, \ldots, v_j, \ldots, v_N)$	Principal cell membrane potential (deviation from rest)	Input drive and recurrent drive computed in separate compartments of dendritic tree
$\mathbf{z} = (z_1, z_2, \ldots, z_j, \ldots, z_N)$	Input drive	Dendritic computation, sum of synaptic currents
$\hat{\mathbf{y}} = (\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_j, \ldots, \hat{y}_N)$	Recurrent drive	Dendritic computation, sum of synaptic currents
$\mathbf{W}_{zx}$	Input weight matrix ($N \times M$): each row corresponds to the spatial RF of one principal cell	Excitatory and inhibitory (i.e., positive and negative) synaptic weights
$\mathbf{W}_{\hat{y}y}$	Recurrent weight matrix: each row determines the recurrent drive for one principal cell	Excitatory and inhibitory (i.e., positive and negative) synaptic weights
$\mathbf{a} = (a_1, a_2, \ldots, a_j, \ldots, a_N)$	Modulator cell responses and recurrent gain	Firing rates of inhibitory neurons (proportional to membrane depolarization), each of which determines conductance of the dendritic compartment of a principal cell receiving that cell's recurrent drive
$\mathbf{u} = (u_1, u_2, \ldots, u_j, \ldots, u_N)$	Responses of second population of modulator cells	Firing rates of a type of excitatory neurons (proportional to membrane depolarization, above a spontaneous firing rate)
$\mathbf{W}$ and $w_{jk} \geq 0$	Normalization weight matrix $\mathbf{W}$ comprising normalization weights w_{jk}	Excitatory synaptic weights
τ_v, τ_a, τ_u	Intrinsic time constants of each of the corresponding cell classes	Membrane capacitance and conductance
$b_0 > 0$	Input gain (constant)	Conductance of the dendritic compartment of the principal cells receiving that cell's input drive
$\sigma > 0$	Contrast gain (constant)	Spontaneous firing rates of $\mathbf{u}$ modulator cells

Boldface lowercase letters denote vectors, and boldface uppercase letters denote matrices. The variables ($\mathbf{y}, \mathbf{v}, \hat{\mathbf{y}}, \mathbf{x}, \mathbf{z}, \mathbf{a}, \mathbf{u}$) are each functions of time, e.g., $\mathbf{y}(t)$, but we drop the explicit dependence on t to simplify the notation.

www.manaraa.com

Modulator cell responses are as follows:

$$\tau_a \frac{da_j}{dt} = -a_j + \sqrt{u_j} + a_j\sqrt{u_j}, \qquad [5]$$

$$\tau_u \frac{du_j}{dt} = -u_j + \sum_k w_{jk}y_k u_k + \left(\frac{\sigma b_0}{1+b_0}\right)^2. \qquad [6]$$

All variables in Eqs. 5 and 6 are constrained to be ≥0. Vectors **a** and **u** represent responses of the two types of modulator cells (firing rates proportional to membrane depolarization, i.e., without squaring unlike Eq. 2). The need for both classes of modulator cells is explained below (Variants of the Model). Modulator cell responses u_j represent a normalization pool, computed from the normalization weights w_{jk} and the principal cell responses y_j (Eq. 6 and Fig. 2, purple). Responses of the other population of modulator cells a_j are multiplied by the recurrent drive $\hat{y}_j$ (Eq. 1), thereby determining the recurrent gain and recurrent amplification. Responses a_j depend on responses u_j (Eq. 5), so that the recurrent amplification depends on the normalization pool.

When the input drive is constant over time, the model has a fixed point such that the neural activity is normalized:

$$\mathbf{y} = \frac{\lfloor \mathbf{z} \rfloor^2}{\sigma^2 + \mathbf{Wz}^2}, \qquad [7]$$

where the numerator is half-squared, and the quotient means element-by-element division. Indeed, the exact form of Eqs. 1–6 was designed so that it would achieve this fixed point. To derive Eq. 7, set the derivatives in Eqs. 1, 5, and 6 equal to 0 and simplify (*SI Appendix*). The values of w_{jk} in Eq. 6 are the normalization weights, i.e., the elements of **W** in Eq. 7. Variants of Eq. 7 (with various exponents) have been fit to a wide range of experimental data (see *SI Appendix* for references).

Simulated neural responses in the following figures are intended to exhibit qualitative aspects of neurophysiological phenomena, i.e., the models have not (yet) been optimized to replicate published data by tuning or fitting the model parameters (*SI Appendix*). We simulated responses to drifting sinusoidal gratings (or pairs of gratings) with various orientations, temporal frequencies, and contrasts. Responses to transient drifting gratings are more sustained than the responses to transient stationary gratings (54, 55). Unless otherwise stated, model parameters were as follows: $b_0 = 0.2$, $\sigma = 0.1$, $\tau_v = 1$ ms, $\tau_a = 2$ ms, and $\tau_u = 1$ ms. The normalization pool included all orientations (evenly weighted) at the center of a neuron's RF, and included only orientations near the preferred orientation at spatial locations surrounding the RF. Euler's forward method was used to compute Eqs. 1, 5, and 6 with time step $\Delta t = 0.1$ ms.

Recurrent Amplification, Effective Time Constant, Onset Transients, and Oscillations. The recurrent circuit model (expressed by Eqs. 1–6 and depicted in Fig. 2) mimics many features of the dynamics of V1 activity. We focus on response dynamics because the mean firing rates are given by Eqs. 7 and 8, which are already known to fit a wide range of experimental data (Fig. 1) (see *SI Appendix* for references).

Simulated responses to grating stimuli with various contrasts replicated experimental observations (Fig. 3). Response amplitudes of simulated V1 simple and complex cells were exactly equal to Eq. 8, saturating at high contrasts (Fig. 3 *A* and *E–G*). The responses of the modulator cells increased monotonically with contrast but did not saturate (Fig. 3*B*). Responses were amplified by 100× when contrast was low but by only ~1× when contrast was high (Fig. 3*C*). The effective time constant was correspondingly long for low contrasts but short for high
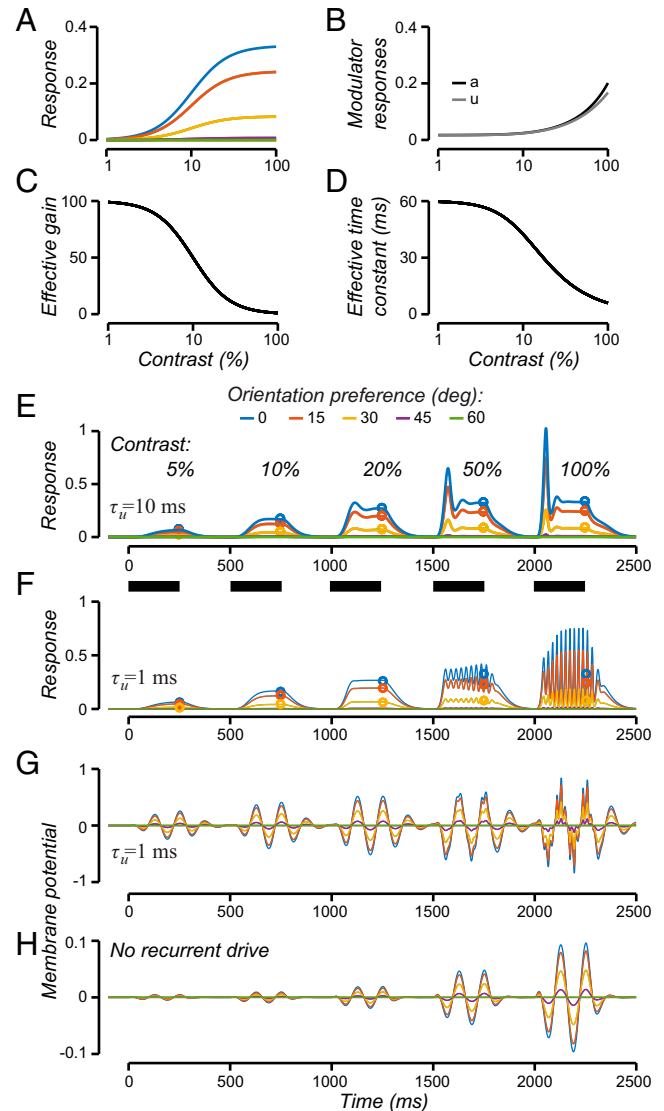

Fig. 3. Recurrent amplification, effective time constant, onset transients, and oscillations. (*A*) Response amplitudes of simulated simple and complex cells as a function of grating contrast. Different colors correspond to different orientation preferences. (*B*) Response amplitudes of the modulator cells: *a* (dark gray curve) and square root of *u* (light gray curve). (*C*) Effective gain (i.e., the ratio of *y* to *z²*). (*D*) Effective time constant. (*E*) Response dynamics of simulated complex cells for a sequence of stimulus contrasts. Thick horizontal bars, each stimulus presentation was 250 ms. Stimulus contrasts: 5%, 10%, 20%, 50%, and 100%. Different colors correspond to different orientation preferences. Open circles, steady-state response amplitudes (from *A*). Modulator cell time constant $\tau_u = 10$ ms. (*F*) Response dynamics of simulated complex cells for $\tau_u = 1$ ms. (*G*) Membrane potential responses (*v*) of simulated simple cells. Response modulation corresponds to stimulus temporal frequency (8 Hz). Different colors correspond to different orientation preferences, all with the same temporal phase. (*H*) Membrane potential responses of simulated simple cells, but with recurrent amplification disabled (note smaller *y* scale).

contrasts (Fig. 3*D*). Consequently, high-contrast stimuli evoked rapid increases in activity, whereas low-contrast stimuli evoked much slower and more gradual increases in activity before achieving steady state (Fig. 3*E*). The rate at which activity decreased following stimulus offset was different from the rate at which activity increased after lifting off from zero following stimulus onset (Fig. 3*E*). These results are similar to a variety of electrophysiological measurements (13, 28, 54–61).

Heeger and Zemlianova

www.manaraa.com

We can derive expressions for the effective gain and the effective time constant of the responses to generate experimentally testable predictions and for fitting data. The effective gain and effective time constant both decrease with increasing stimulus strength. Weak stimuli are strongly amplified (large effective gain) via the recurrent circuit, which takes a period of time (long effective time constant). Strong stimuli are weakly amplified (small effective gain), which happens more quickly (short effective time constant). The effective gain of each neuron in the circuit (the ratio of each element of $\mathbf{y}$ to each element of $\mathbf{z}^2$) depends on the input drive and the normalization weights: $\mathbf{Wz}^2$. The effective time constant depends on the effective gain (*SI Appendix*), so it too depends on $\mathbf{Wz}^2$. For stimuli composed of drifting sinusoidal gratings or pairs of gratings:

$$r \propto \frac{c_t^2}{\sigma^2 + c_t^2 + \beta c_m^2}, \qquad [8]$$

$$g = \frac{1}{\sigma^2 + c_t^2 + \beta c_m^2}, \qquad [9]$$

$$\tau = \tau_v \left(\frac{1 + b_0}{b_0} \right) \sqrt{g}, \qquad [10]$$

where r is the amplitude of a principal cell's response (e.g., the mean firing rate of a V1 complex cell), g is the effective gain of that neuron's responses, and τ is that neuron's effective time constant (see *SI Appendix* for derivations). The value of c_t is the contrast of a test grating (e.g., a preferred orientation grating restricted to the RF). The value of c_m is the contrast of a mask grating that by itself does not evoke a response. The value of $0 < \beta < 1$ depends on the normalization weights. Eqs. 8–10 follow from Eq. 7 because the input drive is a weighted sum of the input, i.e., z_j is proportional to contrast. From Eq. 8, it is evident that responses saturate (level off) when the test contrast is large ($\gg \sigma$), cross-orientation suppression results when a mask grating is superimposed that is orthogonal to the preferred orientation, and surround suppression results when a mask grating is added in the region surrounding the RF, all characteristics of visual neurophysiology (Fig. 1). From Eqs. 9 and 10, the effective gain

is large when stimulus contrast is zero ($g = 100$ for $c_t = c_m = 0$, $b_0 = 0.2$, $\sigma = 0.1$, and $\tau_v = 1$ ms), and the effective time constant is long ($\tau = 60$ ms for those parameters). However, the gain is small ($g \sim 1$) and the effective time constant is short ($\tau = 6$ ms) when contrast is high.

By changing one of the model parameters (specifically, the intrinsic time constant of the modulator cells τ_u), simulated responses to high-contrast stimuli exhibited either strong transients following stimulus onset (Fig. 3E, $\tau_u = 10$ ms) or stable, high-frequency ($\sim$40 to 50 Hz) oscillations (Fig. 3F, $\tau_u = 1$ ms), synchronized across neurons. Both of these phenomena—onset transients (54, 55, 62) and stable oscillations (14, 15, 17–30)—have been widely reported in experimental observations. Note, however, that the experimental evidence for gamma oscillations is based on LFP, electrocorticography, electroencephalogram, and magnetoencephalography measurements, each of which depend on the synchronized membrane potential fluctuations across a large population of neurons (63, 64); we would not expect oscillations to be evident in measurements of single-unit spiking (*SI Appendix*, Fig. S2, and see *SI Appendix*, text).

For some parameter regimes, the responses exhibited onset transients (Fig. 3E) followed by stable oscillations (Fig. 3F), but we have not systematically characterized the parameters that do so. The temporal filter that was used to simulate the responses of the LGN inputs (*SI Appendix*) attenuates the onset transients, without which there would typically be an initial transient overshoot.

In these simulations, the normalization weights were all equal, so the response transients and/or oscillations were perfectly synchronized across the population of neurons. Consequently, despite the complex dynamics, response ratios of neurons with different orientation preferences were maintained throughout each stimulus presentation, resembling some experimental results (55), and enabling an accurate readout of stimulus orientation at any time point. With unequal normalization weights, response ratios evolved over time with nonstationary readout, analogous to other experimental results (65). Furthermore, with unequal normalization weights, response ratios also depended on stimulus contrast so that the simulated neural responses did not exhibit perfectly contrast-invariant tuning curves.
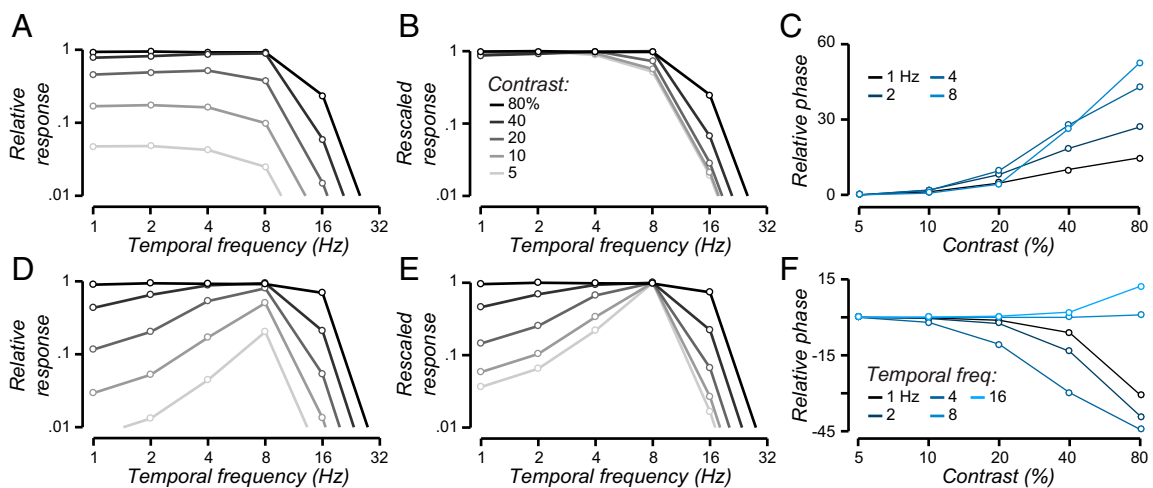

Fig. 4. Temporal-frequency tuning and phase advance depend on contrast. (*A–C*) Low-pass temporal-frequency tuning ($\omega = 0$ Hz). (*A*) Response amplitudes for each of several stimulus temporal frequencies. Different shades of gray correspond to different stimulus contrasts: 5%, 10%, 20%, 40%, and 80%. (*B*) Rescaled responses. The different curves (from *A*) are each rescaled to have the same maximum so that the shapes of the curves can be readily compared. Responses to high contrasts (darker curves) are elevated compared to the responses to low contrasts (lighter curves). (*C*) Phase advance. Response phase for each of several contrasts. Different colors correspond to different temporal frequencies: 1 2, 4, and 8 Hz. (*D–F*) Bandpass temporal-frequency tuning ($\omega = 8$ Hz).

Heeger and Zemlianova

Disabling the recurrent amplification (i.e., simulating an experiment in which cortical spiking is shut down) attenuated the membrane potential response amplitudes by a factor of ~10× at high contrasts (Fig. 3H), while maintaining their orientation selectivity, resembling experimental results (66–69).

Temporal-Frequency Tuning and Phase Advance Depend on Contrast. Temporal-frequency tuning of both simple and complex cells depends on stimulus contrast, and simple-cell response phase depends on contrast (6, 10–13). It was previously proposed that these phenomena can be explained by a recurrent normalization model in which a neuron's conductance (and consequently its intrinsic time constant) depends on stimulus contrast (6, 13). Here, we hypothesize instead that the effective time constant depends on contrast because the amount of recurrent amplification in the circuit decreases with increasing contrast (Eqs. 9 and 10).

Simulated temporal-frequency tuning depended systematically on contrast, responding to a broader range of temporal frequencies at high contrasts (Fig. 4). Fig. 4 A and B plot results for a population of neurons with preferred temporal frequency $\omega = 0$ Hz, i.e., the recurrent drive in the model acted like a low-pass filter (*SI Appendix*). Increasing stimulus contrast increased the responsivity of the simulated neurons for high temporal frequencies. Fig. 4 D and E plot results for neurons with preferred temporal frequency $\omega = 8$ Hz, i.e., the recurrent drive in the model acted like a bandpass filter, matching the preferred temporal frequency of the simulated LGN inputs. In this case, increasing stimulus contrast increased the responsivity of the simulated neurons for both low and high temporal frequencies. For low contrasts, temporal-frequency tuning was bandpass with a relatively narrow bandwidth. Increasing stimulus contrast transformed the temporal frequency tuning from bandpass to low pass while nearly doubling the high temporal-frequency cutoff. This behavior arises in the model because the effective time constant depends on contrast: the effective gain decreases with increasing contrast (Eq. 9), and the effective time constant decreases with decreasing effective gain (Eq. 10). A shorter time constant corresponds to a broader bandwidth, raising the high temporal-frequency cutoff for a low-pass tuning curve, and raising both the low and high cutoffs for a bandpass tuning curve.

Response phase also depended systematically on contrast (Fig. 4 C and F). For simulated simple cells with low-pass temporal-frequency tuning, response phases advanced with increasing contrast, more so for higher temporal frequencies (Fig. 4C). For simulations with bandpass temporal-frequency tuning, response phases shifted in opposite directions for temporal frequencies above and below the preferred temporal frequency (Fig. 4F).

Results like those shown in Fig. 4 A–C have been observed experimentally (6, 10–13): increasing phase advance and increasing the high temporal-frequency cutoff with increasing contrast. The model predicts that the effects shown in Fig. 4 D–F may be evident for neurons with narrow temporal-frequency tuning, e.g., perhaps direction-selective neurons in layer 4b.

Response Dynamics Depend on Stimulus Location. The dynamics of V1 activity depends on whether a stimulus is placed in the center or flanks of a neuron's receptive field (8). Activity evoked by a small grating patch extends over a cortical region of several millimeters (depending on stimulus size, spatial frequency, and eccentricity). Following stimulus onset, responses rise simultaneously over the entire active region, but reach their peak more rapidly at the center. Furthermore, the rate of response increase following stimulus onset is faster for higher contrasts. Following stimulus offset, responses fall simultaneously at all locations, and the rate of response decrease is the same for all locations and all contrasts. It was previously proposed that these phenomena can

be explained by a recurrent normalization model in which a neuron's conductance (and consequently its intrinsic time constant) depends on the spatial distribution of stimulus contrasts, via the normalization weights (8). Here, we hypothesize instead that the effective time constant (as opposed to the intrinsic time constant) of each neuron depends on normalization weights.

Simulated responses recapitulated the experimentally measured spatiotemporal dynamics (Fig. 5). Responses lifted off simultaneously following stimulus onset but increased at a faster rate for RF locations centered on the stimulus (Fig. 5, darker colors) and for higher contrasts (Fig. 5, responses to second stimulus presentation at $t = 500$ ms). Recurrent amplification was weaker when the stimulus was presented closer to the center of a neuron's RF, and it was weaker for higher contrasts. Consequently, the effective gain was smaller (Eq. 9) and the time constant was shorter (Eq. 10) for these conditions. The effective time constant following stimulus offset was ~60 ms, regardless of what the stimulus had been (Eqs. 9 and 10 with $c_t = c_m = 0$, $b_0 = 0.2$, $\sigma = 0.1$, and $\tau_v = 1$ ms).

Oscillations Depend on Stimulus Contrast and Size. Simulated responses exhibited high-frequency oscillations (Fig. 3F). For grating stimuli, and in the absence of noise, these oscillations were evident only at high (>50%) contrasts, and the oscillation amplitudes (Fig. 6B) increased with stimulus size and contrast. Oscillation frequencies also increased with contrast. Response amplitudes, on the other hand, exhibited surround suppression so they were nonmonotonic with stimulus size at high contrasts (Fig. 6A, dark gray and black curves). The oscillations depended indirectly on stimulus temporal frequency because the input drive to each neuron depended on stimulus temporal frequency with respect to the neurons' preferred temporal frequency (Fig. 4). That is, a lower contrast grating with a temporal frequency at the peak of the tuning curve generated the same oscillations as a higher contrast with a temporal frequency on the flank of the tuning curve, such that the two stimuli evoked the same input drive amplitudes. However, the oscillations were otherwise (beyond the dependence on input drive amplitudes) independent of stimulus temporal frequency.

Oscillations were also evident for high-contrast plaid stimuli, composed of a pair of orthogonal gratings, but the oscillations
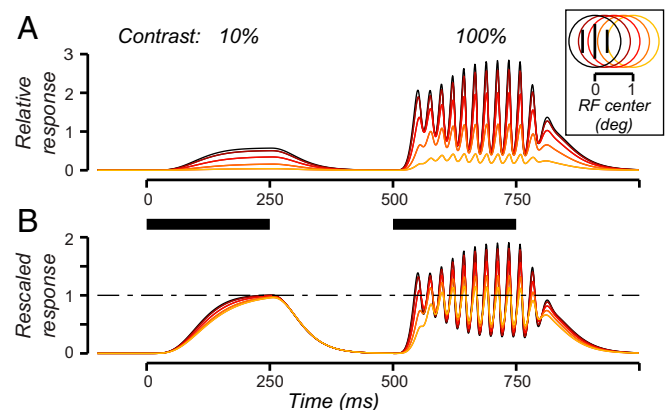

Fig. 5. Response dynamics depend on stimulus location. (*A*) Response dynamics of simulated complex cells for two contrasts (10% and 100%). Different colors correspond to different RF centers. Thick horizontal bars, stimulus presentations. *Inset*, spatial arrangement of stimulus and RF locations. (*B*) Rescaled responses. The different curves (from *A*) are each rescaled by the corresponding fixed point, so that the shapes of the curves can be readily compared. The rate of response increase following stimulus onset is faster for stimuli placed in the center of a receptive field (darker colors) and slower on the flanks of the receptive field (lighter colors).
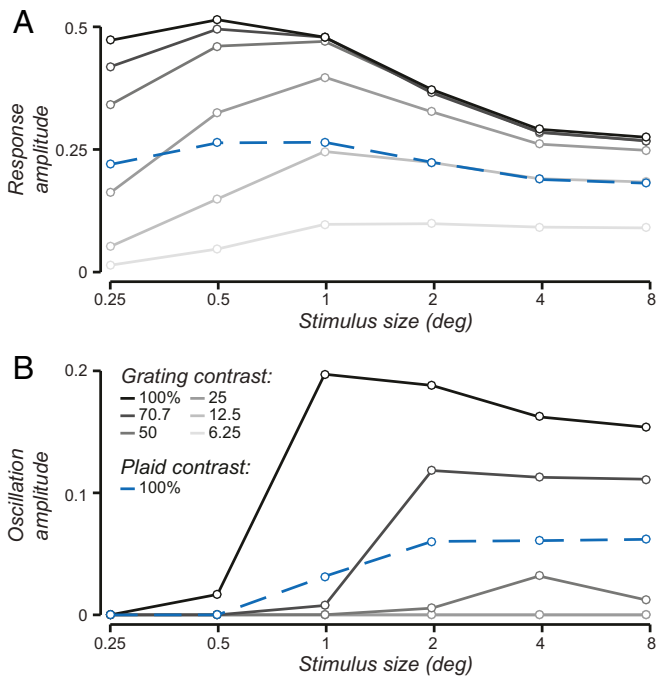
Fig. 6. Oscillations depend on stimulus contrast and size. (*A*) Surround suppression and cross-orientation suppression. Gray curves, response amplitudes to preferred-orientation gratings with each of several stimulus sizes. Different shades of gray correspond to different stimulus contrasts: 6.25%, 12.5%, 25%, 50%, and 100%. Dashed blue curve, 100% contrast plaid composed of a pair of orthogonal gratings (each 50% contrast). (*B*) Oscillation amplitudes (computed by summing the Fourier amplitudes of the responses between 30 and 200 Hz, ignoring the first 250 ms of the responses after stimulus onset). Oscillations are evident only for the highest contrasts and increase with both stimulus contrast and stimulus size.

generated by plaids were smaller in amplitude and lower in frequency than those generated by gratings of the same contrast. Responses to plaids exhibited cross-orientation suppression; the response evoked by a 50% contrast grating with a neuron's preferred orientation was suppressed by about a factor of 2 when an orthogonal mask grating (also 50% contrast) was superimposed (Fig. 6*A*, dashed blue curve vs. third to darkest gray curve). Oscillation amplitudes generated by 100% contrast plaids were about midway between those generated by 50% and 100% contrast gratings (Fig. 6*B*, dashed blue curve).

The oscillations depended on the strength of the normalization pool: specifically, the product of the normalization weights and the squared input drive $\mathbf{Wz}^2$. The normalization pool increased with contrast because the input drive $\mathbf{z}$ was proportional to contrast. The normalization pool increased with stimulus size because it comprised a weighted sum (with nonnegative weights) over space. The normalization pool was smaller for a 100% contrast plaid than a 100% contrast grating. If the normalization was untuned such that all of the normalization weights were 1, then the normalization pool for a 100% contrast plaid composed of two 50% contrast gratings ($\mathbf{Wz}^2 = 0.5^2 + 0.5^2$) would have been equal to that for a 70.7% contrast grating ($\mathbf{Wz}^2 = 0.707^2$). The simulated oscillations differed for these two stimulus conditions (Fig. 6*B*, dashed blue curve vs. second to darkest gray curve) because the normalization pool included only orientations near the preferred orientation at locations surrounding the RF.

All of these results are commensurate with experimental observations that oscillation amplitudes and frequencies depend systematically on stimulus contrast, size, and spatial pattern (14, 15, 17–30), and that oscillations are linked to normalization (14–16). Like the simulation results, oscillation amplitudes in V1

increase with stimulus contrast and size, oscillation frequencies increase with stimulus contrast, and oscillation amplitudes are smaller for plaids than for gratings (and even smaller for stimuli composed of multiple components, also predicted by the model).

Using the current model configuration, simulated oscillation frequencies increased with stimulus size, however, unlike experimental measurements that decrease with stimulus size (14, 15, 26). Previous models have tackled this problem by incorporating a mechanism that pools over large spatial regions and provides excitatory feedback to the principal cells (26, 70). The current family of models may, likewise, be extended by enhancing the recurrent drive with an additional weighted sum over a larger region of the visual field (53). We have verified that doing so may explain the observed decrease in oscillation frequency with increasing stimulus size.

Phase Space Trajectories and Bifurcation Analysis. Oscillations emerged for some parameter regimes of the model, not others, and oscillations in the gamma frequency band corresponded to restricted ranges of those parameter regimes. A bifurcation analysis was performed to determine the ranges of parameter values for which oscillations occur and to determine the corresponding oscillation frequencies.

We analyzed a reduced version of the model in which each of the variables was a scalar instead of a vector (*SI Appendix*, Eq. S37), i.e., one neuron of each of the three types (*y*, *a*, and *u*) instead of a population of neurons with different RF centers and orientation preferences. We characterized the dynamics of the model as a function of the input drive (*z*), the intrinsic time constants (τ_v, τ_a, and τ_u), and the input gain (b_0). In this reduced model, the input was a step at time $t = 0$ and maintained a constant value thereafter.

The model exhibited distinct behaviors with boundaries (state transitions) between them (Fig. 7). When the input drive was small, the fixed point was stable (i.e., an attractor) and simulated responses (*y*) achieved steady state with no oscillations (Fig. 7*A*, green point; Fig. 7*B*). When the input drive was large, the fixed point was unstable with a stable limit cycle and responses exhibited stable oscillations (Fig. 7*A*, orange point and dotted gray curves; Fig. 7*D*). For a middle range of input drives, the fixed point was a spiral attractor and responses exhibited oscillations transiently before achieving steady state (Fig. 7*A*, yellow point; Fig. 7*C*). The steady-state responses increased monotonically with input drive until the bifurcation, at which point the responses exhibited stable oscillations around the fixed point and no longer achieved a steady state (Fig. 7*A*, intersection of solid black, dashed black, and dotted gray curves).

The input drive that induced a bifurcation depended systematically on model parameters (Fig. 7*E*). Each panel of Fig. 7*E* depicts a two-dimensional (2D) bifurcation diagram, i.e., a 2D slice through the space of model parameters. Each panel indicates the input drives for which bifurcations occurred (solid black curves) for different values of τ_u, and the different panels correspond to different values of τ_v and τ_a. Also indicated are the oscillation frequencies (gray scale) when the model exhibited stable oscillations or zero (white) otherwise.

Variants of the Model. The dynamical system expressed by Eqs. 1–6 is but one example from a family of circuit models of normalization, each of which implements normalization via recurrent amplification (see *SI Appendix* for several examples of alternative models from this family). Some of these various models exhibit qualitatively different dynamics such that measurements of the dynamics of neural activity in V1 may be used to distinguish between the alternatives. Each of the various models in this family imply different circuits, such that they may be distinguished experimentally using cell type-specific indicators.
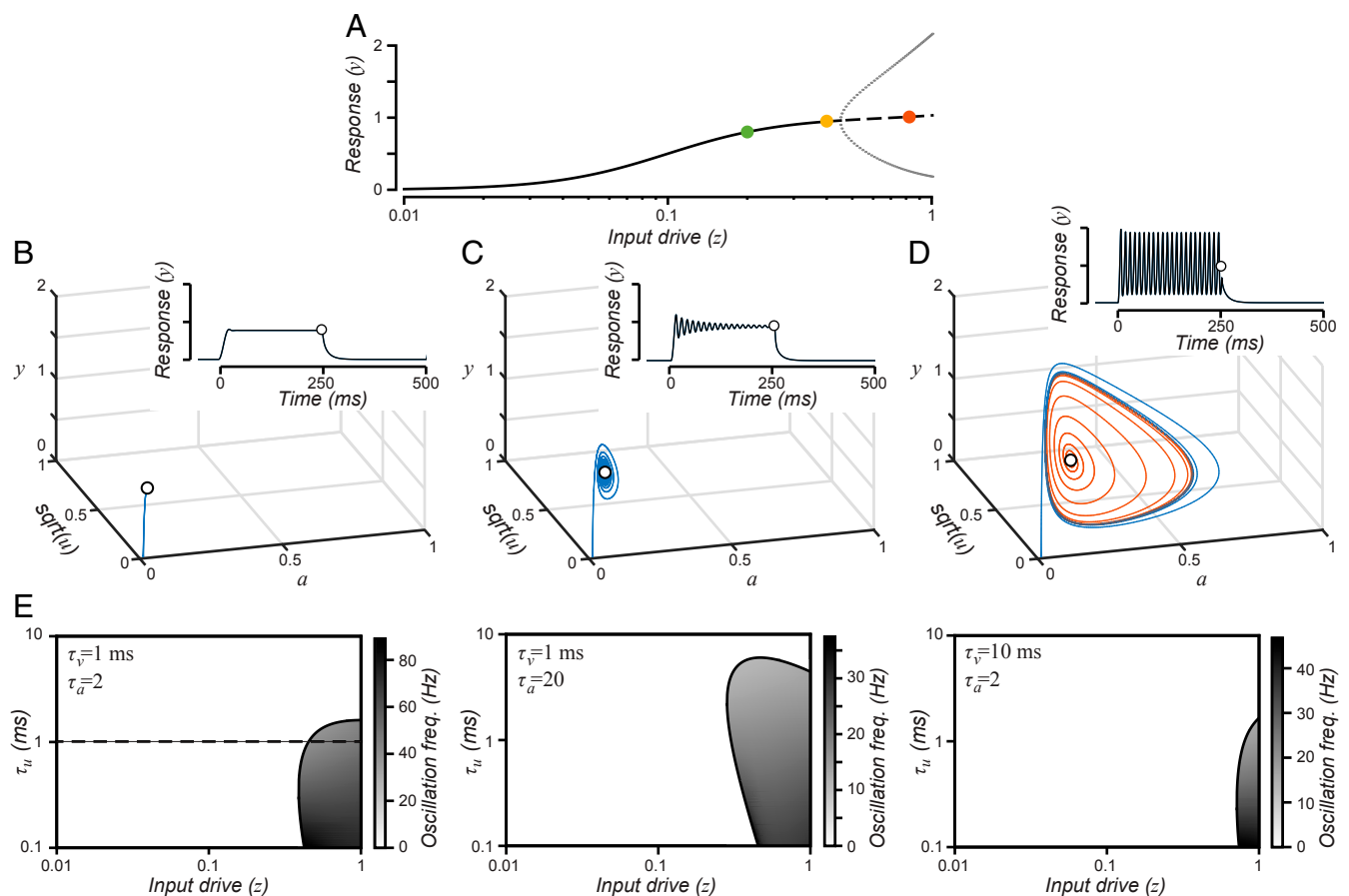
Heeger and Zemlianova

www.manaraa.com

Fig. 7. Phase space trajectories and bifurcation analysis. (*A*) Principal cell responses (*y*) as a function of input drive. Solid curve, stable fixed point (i.e., attractor). Dashed curve, unstable fixed point. Dotted gray curves, maximum and minimum values of *y* during stable oscillations. Green, yellow, and orange dots correspond to phase space trajectories in *B*–*D*, respectively. (*B*) Phase space trajectory convergent to a stable fixed point, corresponding to the green dot in *A* ($z = 0.2$). Blue curve, trajectory of responses of the three neurons (*y*, *a*, and *u*) starting from rest (zero). Open circle indicates the fixed point. *Inset*, response time course of principal cell *y* (open circle, fixed point). (*C*) Phase space trajectory convergent to a spiral attractor, corresponding to the yellow dot in *A* ($z = 0.4$). (*D*) Phase space trajectories for an unstable fixed point with stable limit cycle, corresponding to the orange dot in *A* ($z = 0.8$). Blue curve, trajectory of responses of the three neurons starting from rest (zero). Orange curve, trajectory of responses of the three neurons for an initial condition that is a slight (1%) perturbation from the fixed point. The blue curve spirals clockwise into the limit cycle, and the orange curve spirals clockwise outward from near the fixed point to the limit cycle. The responses traverse a curved manifold (shaped like a saddle or potato chip). (*E*) Bifurcation analysis as a function of input drive (*z*) and time constant parameters (τ_v, τ_a, and τ_u). Other model parameters: $b_0 = 0.2$, $\sigma = 0.1$. Responses in the shaded regions exhibit stable oscillations (unstable fixed points with stable limit cycles). Grayscale indicates oscillation frequency. The dashed line corresponds to *A*.

For example, one of these variants can be ruled out as a plausible model of V1 activity because it does not exhibit dynamics commensurate with V1 activity. This variant (*SI Appendix*, Eq. **S35**) is a simpler circuit with only two types of neurons, a principal cell and a single type of modulator cell instead of two. The circuit has a stable fixed point such that the primary neurons achieve steady-state responses given by the normalization equation (Eqs. **7** and **8**). We have been able to prove mathematically, over a very broad range of parameter values, that the fixed point is stable over the full range of input drives. That is, there is no parameter regime in which the responses exhibit stable oscillations (*SI Appendix*).

Discussion

We developed a family of circuit models of normalization. The key idea is that normalization operates via recurrent amplification, amplifying weak inputs more than strong inputs (31–34). The modulator cells determine the recurrent gain, thereby controlling the amount of recurrent amplification. Each of the models in this family exhibits output responses with a fixed point that follows the normalization equation (Eqs. **7** and **8**) exactly,

for arbitrary (nonnegative) normalization weights. The normalization equation is already known to fit a wide range of experimental data (see *SI Appendix* for references).

These models mimic experimental observations of V1 dynamics linked to normalization: onset transients (Fig. 3*E*), the contrast dependence of the rate of response increase following stimulus onset and response decrease following stimulus offset (Figs. 3 and 5), and the contrast dependence of temporal-frequency tuning and phase advance (Fig. 4). Furthermore, for some models in this family, the fixed point is unstable for large, high-contrast grating stimuli, and responses exhibit oscillations (Figs. 3*F* and 7). The oscillations emerge because of the recurrent circuitry, depending on the strength of the normalization pool, thereby offering an explanation for the link between gamma oscillations and normalization (14–16). Despite the complex dynamics, ratios of the simulated responses across neurons with different stimulus preferences may be maintained throughout each stimulus presentation, enabling an accurate readout of stimulus orientation (or other stimulus parameters) at any time point following the onset of the responses.

These models are examples of oscillatory recurrent gated neural integrator circuits (ORGaNICs) (52, 53). ORGaNICs are a generalization of and a biophysically plausible implementation of long short-term memory units (LSTMs), a class of artificial recurrent neural networks (71) that have been used in machine learning applications (e.g., refs. 72–75). ORGaNICs may be used to explain the complex dynamics of delay-period activity during a working memory task, and how information is manipulated (as well as maintained) during a delay period (53). When applied to motor systems, these circuits convert spatial patterns of premotor activity to temporal profiles of motor control activity: Different spatial patterns of premotor activity evoke different motor control dynamics (53). ORGaNICs are also capable of prediction over time (52). The modulators in ORGaNICs perform multiple functions: normalization, controlling working-memory maintenance and manipulation, controlling pattern generators, gated integration/updating, time warping, reset, controlling the effective time constant, controlling the relative contributions of bottom-up vs. top-down connections, and weighting sensory evidence (likelihood) and internal model (prior) for inference and multisensory integration (52, 53, 76).

Here, we demonstrate that this same family of circuit models can simulate the dynamics of neural activity in V1. Consequently, this theoretical framework is applicable to diverse cognitive processes and neural systems, and we can use V1 as a model system for understanding neural computations and circuits in many brain areas.

Gamma Oscillations. Narrow-band gamma oscillations have been proposed to play a functional role in stimulus feature binding, attention, and/or synchronizing neuronal activity to enhance signal transmission and communication between brain areas (see *SI Appendix* for references). These speculations have been met with considerable skepticism (18, 21, 23, 26, 27, 77–81), in part because oscillation amplitude depends strongly on stimulus conditions (14, 15, 17–30), incommensurate with perception.

Gamma oscillations in the current family of models emerge from the nonlinear dynamics of the recurrent circuit. Synchronized spiking was not required to generate gamma oscillations. Gamma oscillations were generated for a restricted subset of stimulus conditions, depending on the strength of the normalization pool. Consequently, oscillation amplitude was strongest for large, high-contrast gratings, and weaker (or nonexistent) for other spatial patterns and low contrasts, similar to experimental results (14, 15, 17–30).

Long-wavelength stimuli have been found to generate particularly large-amplitude gamma oscillations (29, 82). It should be straightforward to extend the current family of models to account for these results by including red-green and blue-yellow color-opponent channels (83–85) in the LGN input, and by setting the normalization weights to be large for the red-green channel.

The current theoretical framework is most similar to bifurcation-based models of gamma oscillations (86, 87), as opposed to the so-called pyramidal-interneuron gamma (PING) and interneuron gamma (ING) mechanisms for producing gamma oscillations (see *SI Appendix* for references and details). Unlike any of the previous models of gamma oscillations, we designed the current family of models to perform a function (normalization), and gamma oscillations emerged as a by-product.

Failures and Extensions. Stable oscillations were observed in the simulation results for input drives (i.e., contrasts) above a threshold level (above the bifurcation), but narrow-band gamma power has been observed experimentally to change gradually with continuous parametric variation in stimulus parameters (14, 18, 25, 26). All of the simulation results reported above were

performed in the absence of noise. With noise added to the input drive, we observed activity in the gamma-frequency range, even for weak inputs below the bifurcation (*SI Appendix*, Fig. S2 and see *SI Appendix*). This suggests that gamma-band activity may be induced by broadband noise in neural activity (70, 80); the noise spectrum is shaped by recurrent normalization to exhibit a resonant peak in the gamma-frequency range.

The effective time constants of the principal cells in our simulations ranged from 6 to 60 ms, which is within a reasonable range for in vivo cortical neurons, but the values of the intrinsic time constants (1 to 2 ms) were ~10 times shorter than experimental measurements of intrinsic time constants (88, 89). Increasing the values of the time constant parameters would make the responses sluggish and decrease the oscillation frequencies (Fig. 7). This is a challenge for any model that relies heavily on recurrent amplification because the recurrence takes time (multiples of the time constant). See *SI Appendix* for details.

Attention is associated with both increases in the gain of visually evoked responses and increases in gamma oscillations (see *SI Appendix* for references). The constant input gain parameter b_0 may be replaced by a variable vector $\mathbf{b}$ in Eq. 1 (while keeping the constant b_0 in Eq. 6), in combination with normalization, to model the effects of attention on sensory responses. The elements of $\mathbf{b}$ determine the relative attentional gain for each neuron in the circuit (i.e., with different RF centers and different orientation preferences). Doing so would yield steady-state output responses that are already known to fit experimental measurements of response gain changes with attention (e.g., ref. 90). This would also affect the dynamics of the responses and may be used to explain the ostensible link between attention and gamma oscillations (15, 78).

Mechanisms. We have presented a computational theory for what computations are performed by neural circuits in V1, not how they are implemented. However, we can speculate about the underlying mechanisms:

The circuit (Fig. 2) comprises an excitatory principal cell y_j, an inhibitory modulator cell a_j, and another excitatory cell type u_j that makes local recurrent connections. Each type of neuron performs a different dendritic computation (Eqs. 1, 5, and 6).

The circuit also includes inhibitory interneurons (Fig. 2, small circles) to invert the sign of the responses, corresponding to negative weights in the synaptic weight matrices $\mathbf{W}_{zx}$ and $\mathbf{W}_{yy}$. These inhibitory neurons need not be one-to-one with their excitatory inputs as drawn in the figure. Rather, each may compute a weighted sum of their inputs to contribute the terms in Eqs. 3 and 4 with negative weights.

The responses of the principal cells (Eq. 1) may be implemented with a simplified biophysical (equivalent electrical circuit) model of a pyramidal cell (52, 53), in which the two terms of Eq. 1 are computed in separate dendritic compartments. The conductance of the first compartment determines the input gain and the synaptic current in that compartment is the input drive. The conductance and synaptic current in the second compartment correspond, respectively, to the recurrent gain and recurrent drive.

The input drive is computed with positive and negative synaptic weights, i.e., both feedforward excitation and feedforward inhibition (Fig. 2, blue solid and dashed lines, respectively).

The recurrent drive also involves both excitation and inhibition (Fig. 2, green solid and dashed lines, respectively; *SI Appendix*, Eq. S7), presumably via lateral connections within V1. These excitatory and inhibitory recurrent signals are both amplified by an amount that is controlled by the modulator cells, consistent with the experimental observation that surround

Heeger and Zemlianova

www.manaraa.com

suppression involves a decrease in both recurrent excitatory and recurrent inhibitory conductances (42).

Some principal cells may share the same modulators (e.g., principal cells with the same RF and orientation preference but with different temporal phases; see *SI Appendix* for details), suggesting a much larger number of principal cells than modulator cells.

The squaring nonlinearity (Eq. **2**) may be approximated with a high threshold in combination with neural noise (91–93).

The square roots in Eqs. **4** and **5** may be approximated by synaptic depression, which acts as a compressive nonlinearity because the probability of neurotransmitter release is lower at higher firing rates. Alternatively, the square roots in Eqs. **4** and **5** may be replaced by adding another cell type in the circuit (*SI Appendix*, Eq. **S39**).

The modulator cells with firing rates a_j may act via shunting (52, 53), i.e., increasing conductance by a balanced increase in excitation and inhibition without changing the total synaptic current (6, 94).

The modulator cells may correspond to parvalbumin-expressing (PV) (95, 96) and/or somatostatin-expressing (SOM) (97) inhibitory neurons. The modulator cells are expected to have large RFs and broad orientation selectivity (reflecting properties of the normalization pool), consistent with the response properties of SOM and PV neurons, respectively. Furthermore, PV neurons form a local recurrent circuit with excitatory cells, receiving inputs from excitatory cells (98–100), and targeting nearby excitatory cells (101, 102). The modulator cell responses may depend in part on loops through higher visual cortical areas and/or thalamocortical loops (see *SI Appendix* for details and references).

Modulator cell responses a_j depend on a product of a_j with the square root of u_j (Eq. **5**). This product may be computed via NMDA receptors with synaptic current approximately proportional to the product of the presynaptic and postsynaptic firing rates. It may instead be computed with a synaptic current from u_j and an intrinsic voltage-sensitive ion channel (89) such that conductance is inversely proportional to membrane depolarization a_j (noting that firing rates a_j are proportional to membrane depolarization).

For the summation over $w_{jk} y_k u_k$ in Eq. **6**, each term may be computed in separate dendritic compartments.

Some of the effects of cross-orientation suppression may be due to feedforward (not recurrent) mechanisms, and the simulations here incorrectly ignored the fact that some of the normalization is inherited from the LGN inputs (see *SI Appendix* for details and references). Regardless, evidence suggests that cortical circuits make an important contribution to cross-orientation suppression (103), and there is consensus that some of the effects of normalization (e.g., surround suppression) are computed with cortical circuits.

Comparison with Previous Models. The current theoretical framework is superior to both the original recurrent normalization model and alternative recurrent models of normalization (4–6, 13, 41–46). First, none of the previous models converge exactly to the normalization equation (Eqs. **7** and **8**) for arbitrary normalization weights. Although they may approximate weighted normalization, the extent to which the previous recurrent models fit the full range of experimental data is unknown. The current family of recurrent circuit models has a mathematically tractable solution that equals weighted normalization. This has practical consequences, enabling us to derive closed-form expressions

(Eqs. **8–10**; see also *SI Appendix*) for making experimentally testable predictions and for fitting data. Second, the current theoretical framework, unlike previous models, mimics the dynamics of V1 activity. Third, most of the previous models do not rely on recurrent amplification to achieve normalization. Fourth, the current theoretical framework is applicable to diverse cognitive processes and neural systems, e.g., working memory and motor control (52, 53), enabling us to use V1 as a model system for understanding the neural computations and circuits in many brain areas. Fifth, by virtue of being a generalization of LSTMs, the current theoretical framework can solve relatively sophisticated tasks.

The current family of models is most similar to the inhibition stabilized network (ISN) (42) and the stabilized supralinear network (SSN) (45), but there are also crucial differences. All of these models include recurrent excitation that would be unstable if inhibition was absent or held fixed. All of them also include inhibitory stabilization, but the stabilizing inhibition in the current model is modulatory (multiplicative), unlike ISN and SSN in which inhibition is subtractive. Inhibitory stabilization, by itself, does not explain the phenomena associated with normalization. A linear recurrent model, which does not exhibit any of the nonlinear effects associated with normalization, may be stabilized by inhibition, i.e., it would be unstable if inhibition were removed or held fixed (45, 104). Normalization phenomena arise in the SSN model from a combination of amplification and inhibitory stabilization. SSN (45) and also earlier models (4–6, 13) amplify weak inputs more than strong inputs due to a power law relationship (e.g., half-squaring) between membrane depolarization and firing rate (91–93, 105). Removing the power function from SSN yields a linear model that is qualitatively different, in which responses increase in proportion to contrast (45). The current family of models also includes half-squaring, but it is not critical for normalization. Removing the squaring yields qualitatively similar phenomena; for example, the contrast-response function would be proportional to $c/(c + \sigma)$ rather than $c^2/(c^2 + \sigma^2)$. Instead normalization in the current models relies on recurrent amplification via the product of recurrent gain and recurrent drive.

Predictions. The real value of this family of recurrent circuit models of normalization rests on whether it can push the field forward by making quantitative and testable predictions, leading to new experiments that may reveal novel phenomena. Some of these predictions are as follows:

We predict that the effective time constant is contrast dependent (Eqs. **9** and **10**); high-contrast stimuli are integrated over much briefer periods of time (by a factor of ∼10×) than low-contrast stimuli. A functional advantage of doing so is to increase the signal-to-noise ratio (SNR) of responses evoked by a low-contrast stimulus. Low contrasts evoke weak input drives with correspondingly low SNRs. Integrating these inputs over a long period of time (i.e., with a low-pass filter or local average) increases the SNR of the neural representation. This hypothesized difference in dynamics could be tested either electrophysiologically or psychophysically.

We hypothesize a link between effective gain and effective time constant: effective time constant should increase with the square root of effective gain (Eq. **10**). This is analogous to the previous shunting inhibition model of normalization (6, 13), but the prediction of that model was that both the gain and time constant change with intrinsic conductance, whereas the effective gain and time constant in the current family of models is a network effect, emerging from the recurrent amplification in the circuit. This hypothesized link may be tested either electrophysiologically or psychophysically (106).

www.manaraa.com

The link between effective gain and effective time constant is further constrained by the value of the input gain parameter b_0 (Eq. **10**). The input gain of neurons in layer 4C (the input layer) may be estimated from intracellular measurements of membrane potential fluctuations with and without disabling cortical spikes (e.g., via optogenetics) as simulated in Fig. 3 *G* and *H*. The input gain may also be manipulated with attention (e.g., ref. 90).

We predict a link between the intrinsic time constants and oscillation frequencies (Fig. 7*E*). In our simulations, oscillation frequency depended systematically on the values of the intrinsic time constants (τ_v, τ_a, and τ_u), and the input gain (b_0). An experimental test of this prediction would involve manipulating the intrinsic time constant (i.e., the conductance) of a particular cell type in the circuit and/or manipulating the input gain with attention.

The effects shown in Fig. 4 *D–F* (increasing responsivity of both low and high temporal frequencies with increasing contrast, and shifting response phases in opposite directions for temporal frequencies above and below the preferred temporal frequency) may be evident for neurons with narrow temporal frequency tuning, e.g., perhaps direction-selective neurons in layer 4b.

Data Availability. There are no data underlying this work.

1. D. J. Heeger, E. H. Adelson, Nonlinear model of cat striate cortex. *Optics News* **15**, A-42 (1989).
2. D. G. Albrecht, W. S. Geisler, Motion selectivity and the contrast-response function of simple cells in the visual cortex. *Vis. Neurosci.* **7**, 531–546 (1991).
3. D. J. Heeger, ""Nonlinear model of neural responses in cat visual cortex"" in *Computational Models of Visual Processing*, M. S. Landy, J. A. Movshon, Eds. (MIT Press, Cambridge, MA, 1991), pp. 119–133.
4. D. J. Heeger, Normalization of cell responses in cat striate cortex. *Vis. Neurosci.* **9**, 181–197 (1992).
5. D. J. Heeger, Modeling simple-cell direction selectivity with normalized, half-squared, linear operators. *J. Neurophysiol.* **70**, 1885–1898 (1993).
6. M. Carandini, D. J. Heeger, Summation and division by neurons in primate visual cortex. *Science* **264**, 1333–1336 (1994).
7. M. Carandini, D. J. Heeger, Normalization as a canonical neural computation. *Nat. Rev. Neurosci.* **13**, 51–62 (2011).
8. Y. F. Sit, Y. Chen, W. S. Geisler, R. Miikkulainen, E. Seidemann, Complex dynamics of V1 population responses explained by a simple gain-control model. *Neuron* **64**, 943–956 (2009).
9. W. Bair, J. R. Cavanaugh, J. A. Movshon, Time course and time-distance relationships for surround suppression in macaque V1 neurons. *J. Neurosci.* **23**, 7690–7701 (2003).
10. R. A. Holub, M. Morton-Gibson, Response of visual cortical neurons of the cat to moving sinusoidal gratings: Response-contrast functions and spatiotemporal interactions. *J. Neurophysiol.* **46**, 1244–1259 (1981).
11. A. F. Dean, D. J. Tolhurst, Factors influencing the temporal phase of response to bar and grating stimuli for simple cells in the cat striate cortex. *Exp. Brain Res.* **62**, 143–151 (1986).
12. D. G. Albrecht, Visual cortex neurons in monkey and cat: Effect of contrast on the spatial and temporal phase transfer functions. *Vis. Neurosci.* **12**, 1191–1210 (1995).
13. M. Carandini, D. J. Heeger, J. A. Movshon, Linearity and normalization in simple cells of the macaque primary visual cortex. *J. Neurosci.* **17**, 8621–8644 (1997).
14. M. A. Gieselmann, A. Thiele, Comparison of spatial integration and surround suppression characteristics in spiking activity and the local field potential in macaque V1. *Eur. J. Neurosci.* **28**, 447–459 (2008).
15. S. Ray, A. M. Ni, J. H. Maunsell, Strength of gamma rhythm depends on normalization. *PLoS Biol.* **11**, e1001477 (2013).
16. D. Hermes, D. Petridou, K. N. Kay, J. Winawer, An image-computable model for the stimulus selectivity of gamma oscillations. *eLife* **8**, e47035 (2019).
17. C. Kayser, R. F. Salazar, P. Konig, Responses to natural scenes in cat V1. *J. Neurophysiol.* **90**, 1910–1920 (2003).
18. J. A. Henrie, R. Shapley, LFP power spectra in V1 cortex: The graded effect of stimulus contrast. *J. Neurophysiol.* **94**, 479–490 (2005).
19. Z. Zhou, M. R. Bernard, A. B. Bonds, Deconstruction of spatial integrity in visual stimulus detected by modulation of synchronized activity in cat visual cortex. *J. Neurosci.* **28**, 3759–3768 (2008).
20. B. Lima, W. Singer, N. H. Chen, S. Neuenschwander, Synchronization dynamics in response to plaid stimuli in monkey V1. *Cereb. Cortex* **20**, 1556–1573 (2010).
21. S. Ray, J. H. Maunsell, Differences in gamma frequencies across visual cortex restrict their possible use in computation. *Neuron* **67**, 885–896 (2010).
22. M. J. Bartolo *et al.*, Stimulus-induced dissociation of neuronal firing rates and local field potential gamma power and its relationship to the resonance blood oxygen level-dependent signal in macaque primary visual cortex. *Eur. J. Neurosci.* **34**, 1857–1870 (2011).
23. X. Jia, A. Kohn, Gamma rhythms in the brain. *PLoS Biol.* **9**, e1001045 (2011).
24. X. Jia, A. M. Smith, A. Kohn, Stimulus selectivity and spatial coherence of gamma components of the local field potential. *J. Neurosci.* **31**, 9390–9403 (2011).
25. S. Ray, J. H. Maunsell, Different origins of gamma rhythm and high-gamma activity in macaque visual cortex. *PLoS Biol.* **9**, e1000610 (2011).
26. X. Jia, D. Xing, A. Kohn, No consistent relationship between gamma power and peak frequency in macaque primary visual cortex. *J. Neurosci.* **33**, 17–25 (2013).
27. D. Hermes, K. J. Miller, B. A. Wandell, J. Winawer, Stimulus dependence of gamma oscillations in human visual cortex. *Cereb. Cortex* **25**, 2951–2959 (2015).
28. M. W. Self *et al.*, The effects of context and attention on spiking activity in human early visual cortex. *PLoS Biol.* **14**, e1002420 (2016).
29. E. Bartoli *et al.*, Functionally distinct gamma range activity revealed by stimulus tuning in human visual cortex. *Curr. Biol.* **29**, 3345–3358.e7 (2019).
30. N. M. Brunet, P. Fries, Human visual cortical gamma reflects natural image structure. *Neuroimage* **200**, 635–643 (2019).
31. H. Adesnik, M. Scanziani, Lateral competition for cortical space by layer-specific horizontal circuits. *Nature* **464**, 1155–1160 (2010).
32. T. K. Sato, B. Haider, M. Häusser, M. Carandini, An excitatory basis for divisive normalization in visual cortex. *Nat. Neurosci.* **19**, 568–570 (2016).
33. H. Adesnik, Synaptic mechanisms of feature coding in the visual cortex of awake mice. *Neuron* **95**, 1147–1159.e4 (2017).
34. K. A. Bolding, K. M. Franks, Recurrent cortical circuits implement concentration-invariant odor coding. *Science* **361**, eaat6904 (2018).
35. R. J. Douglas, C. Koch, M. Mahowald, K. A. Martin, H. H. Suarez, Recurrent excitation in neocortical circuits. *Science* **269**, 981–985 (1995).
36. E. M. Callaway, Local circuits in primary visual cortex of the macaque monkey. *Annu. Rev. Neurosci.* **21**, 47–74 (1998).
37. Y. Yoshimura, J. L. Dantzker, E. M. Callaway, Excitatory cortical neurons form fine-scale functional networks. *Nature* **433**, 868–873 (2005).
38. R. J. Douglas, K. A. Martin, Recurrent neuronal circuits in the neocortex. *Curr. Biol.* **17**, R496–R500 (2007).
39. C. Kapfer, L. L. Glickfeld, B. V. Atallah, M. Scanziani, Supralinear increase of recurrent inhibition during sparse activity in the somatosensory cortex. *Nat. Neurosci.* **10**, 743–753 (2007).
40. S. Peron *et al.*, Recurrent interactions in local cortical circuits. *Nature* **579**, 256–259 (2020).
41. M. Carandini, D. J. Heeger, W. Senn, A synaptic explanation of suppression in visual cortex. *J. Neurosci.* **22**, 10053–10065 (2002).
42. H. Ozeki, I. M. Finn, E. S. Schaffer, K. D. Miller, D. Ferster, Inhibitory stabilization of the cortical network underlies visual surround suppression. *Neuron* **62**, 578–592 (2009).
43. T. Brosch, H. Neumann, Computing with a canonical neural circuits model with pool normalization and modulating feedback. *Neural Comput.* **26**, 2735–2789 (2014).
44. K. Louie, T. LoFaro, R. Webb, P. W. Glimcher, Dynamic divisive normalization predicts time-varying value coding in decision-related circuits. *J. Neurosci.* **34**, 16046–16057 (2014).
45. D. B. Rubin, S. D. Van Hooser, K. D. Miller, The stabilized supralinear network: A unifying circuit motif underlying multi-input integration in sensory cortex. *Neuron* **85**, 402–417 (2015).
46. E. Koch, J. Jin, J. M. Alonso, Q. Zaidi, Functional implications of orientation maps in primary visual cortex. *Nat. Commun.* **7**, 13529 (2016).
47. G. Sperling, M. M. Sondhi, Model for visual luminance discrimination and flicker detection. *J. Opt. Soc. Am.* **58**, 1133–1145 (1968).
48. S. Grossberg, Contour enhancement, short-term memory, and constancies in reverberating neural networks. *Stud. Appl. Math.* **52**, 217–257 (1973).
49. S. Grossberg, Adaptive pattern classification and universal recoding: I. Parallel development and coding of neural feature detectors. *Biol. Cybern.* **23**, 121–134 (1976).
50. D. J. Heeger, K. O. Zemlianova, Dynamic normalization. *bioRxiv*:10.1101/2020.03.22.002634 (2020). Posted 25 March 2020.
51. D. J. Heeger, K. O. Zemlianova, Supplemental material for "Dynamic Normalization" (NYU Faculty Digital Archive, 2020). https://archive.nyu.edu/handle/2451/61045. Accessed 25 March 2020.
52. D. J. Heeger, W. E. Mackey, ORGaNICs: A theory of working memory in brains and machines. arXiv:1803.06288 (2018). Posted 16 March 2018.
53. D. J. Heeger, W. E. Mackey, Oscillatory recurrent gated neural integrator circuits (ORGaNICs), a unifying theoretical framework for neural dynamics. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 22783–22794 (2019).
54. J. R. Müller, A. B. Metha, J. Krauskopf, P. Lennie, Information conveyed by onset transients in responses of striate cortical neurons. *J. Neurosci.* **21**, 6978–6990 (2001).
55. D. G. Albrecht, W. S. Geisler, R. A. Frazor, A. M. Crane, Visual cortex neurons of monkeys and cats: Temporal dynamics of the contrast response function. *J. Neurophysiol.* **88**, 888–913 (2002).
56. D. J. Tolhurst, N. S. Walker, I. D. Thompson, A. F. Dean, Non-linearities of temporal summation in neurones in area 17 of the cat. *Exp. Brain Res.* **38**, 431–435 (1980).

www.manaraa.com

57. A. F. Dean, D. J. Tolhurst, On the distinctness of simple and complex cells in the visual cortex of the cat. *J. Physiol.* **344**, 305–325 (1983).

58. W. Bair, J. R. Cavanaugh, M. A. Smith, J. A. Movshon, The timing of response onset and offset in macaque visual neurons. *J. Neurosci.* **22**, 3189–3205 (2002).

59. M. A. Smith, W. Bair, J. A. Movshon, Dynamics of suppression in macaque primary visual cortex. *J. Neurosci.* **26**, 4826–4834 (2006).

60. S. P. Burns, D. Xing, R. M. Shapley, Comparisons of the dynamics of local field potential and multiunit activity signals in macaque visual cortex. *J. Neurosci.* **30**, 13739–13749 (2010).

61. J. Zhou, N. C. Benson, K. N. Kay, J. Winawer, Compressive temporal summation in human visual cortex. *J. Neurosci.* **38**, 691–709 (2018).

62. J. R. Müller, A. B. Metha, J. Krauskopf, P. Lennie, Rapid adaptation in visual cortex to the structure of images. *Science* **285**, 1405–1408 (1999).

63. G. Buzsáki, C. A. Anastassiou, C. Koch, The origin of extracellular fields and currents–EEG, ECoG, LFP and spikes. *Nat. Rev. Neurosci.* **13**, 407–420 (2012).

64. E. R. Kupers, N. C. Benson, J. Winawer, A visual encoding model links magneto-encephalography signals to neural synchrony in human cortex. *bioRxiv*:10.1101/2020.04.19.049197 (2020). Posted 10 April 2020.

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

66. D. Ferster, S. Chung, H. S. Wheat, Orientation selectivity of synaptic input from lateral geniculate nucleus to simple cells of cat visual cortex. *Nature* **380**, 249–252 (1996).

67. S. Chung, D. Ferster, Strength and orientation tuning of the thalamic input to simple cells revealed by electrically evoked cortical suppression. *Neuron* **20**, 1177–1189 (1998).

68. K. Reinhold, A. D. Lien, M. Scanziani, Distinct recurrent versus afferent dynamics in cortical visual processing. *Nat. Neurosci.* **18**, 1789–1797 (2015).

69. A. D. Lien, M. Scanziani, Cortical direction selectivity emerges at convergence of thalamic synapses. *Nature* **558**, 80–86 (2018).

70. K. Kang, M. Shelley, J. A. Henrie, R. Shapley, LFP spectral peaks in V1 cortex: Network resonance and cortico-cortical feedback. *J. Comput. Neurosci.* **29**, 495–507 (2010).

71. S. Hochreiter, J. Schmidhuber, Long short-term memory. *Neural Comput.* **9**, 1735–1780 (1997).

72. A. Graves, Generating sequences with recurrent neural networks. arXiv:1308.0850 (2013).

73. A. Graves, A.-r. Mohamed, G. Hinton, ""Speech recognition with deep recurrent neural networks"" in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, (Curran Associates, Red Hook, NY, 2013), pp. 6645–6649.

74. K. Cho *et al*, Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv*:1406.1078 (2014). Posted 3 June 2014.

75. I. Sutskever, O. Vinyals, Q. V. Le, ""Sequence to sequence learning with neural networks"" in *Advances in Neural Information Processing Systems*, Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, K. Q. Weinberger, Eds. (Neural Information Processing Systems Foundation, San Diego, CA, 2014), pp. 3104–3112.

76. D. J. Heeger, Theory of cortical function. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 1773–1782 (2017).

77. S. P. Burns, D. Xing, M. J. Shelley, R. M. Shapley, Searching for autocoherence in the cortical network with a time-frequency analysis of the local field potential. *J. Neurosci.* **30**, 4033–4047 (2010).

78. M. Chalk *et al*., Attention reduces stimulus-driven gamma frequency oscillations and spike field coherence in V1. *Neuron* **66**, 114–125 (2010).

79. S. P. Burns, D. Xing, R. M. Shapley, Is gamma-band activity in the local field potential of V1 cortex a "clock" or filtered noise? *J. Neurosci.* **31**, 9658–9664 (2011).

80. D. Xing *et al*., Stochastic generation of gamma-band activity in primary visual cortex of awake and anesthetized monkeys. *J. Neurosci.* **32**, 13873–80a (2012).

81. D. Hermes, K. J. Miller, B. A. Wandell, J. Winawer, Gamma oscillations in visual cortex: The stimulus matters. *Trends Cognit. Sci.* **19**, 57–58 (2015).

82. V. Shirhatti, S. Ray, Long-wavelength (reddish) hues induce unusually large gamma oscillations in the primate primary visual cortex. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 4489–4494 (2018).

83. L. M. Hurvich, D. Jameson, An opponent-process theory of color vision. *Psychol. Rev.* **64**, 384–404 (1957).

84. A. M. Derrington, J. Krauskopf, P. Lennie, Chromatic mechanisms in lateral geniculate nucleus of macaque. *J. Physiol.* **357**, 241–265 (1984).

85. P. K. Kaiser, R. M. Boynton, *Human Color Vision*, (Optical Society of America, Washington, D.C, 1996).

86. N. Brunel, X. J. Wang, What determines the frequency of fast network oscillations with irregular neural discharges? I. Synaptic dynamics and excitation-inhibition balance. *J. Neurophysiol.* **90**, 415–430 (2003).

87. S. Keeley, Á. Byrne, A. Fenton, J. Rinzel, Firing rate models for gamma oscillations. *J. Neurophysiol.* **121**, 2181–2190 (2019).

88. A. Y. Tan, Y. Chen, B. Scholl, E. Seidemann, N. J. Priebe, Sensory stimulation shifts visual cortex from synchronous to asynchronous states. *Nature* **509**, 226–229 (2014).

89. B. Li, B. N. Routh, D. Johnston, E. Seidemann, N. J. Priebe, Voltage-gated intrinsic conductances shape the input-output relationship of cortical neurons in behaving primate V1. *Neuron* **107**, 185–196.e4 (2020).

90. J. H. Reynolds, D. J. Heeger, The normalization model of attention. *Neuron* **61**, 168–185 (2009).

91. J. S. Anderson, I. Lampl, D. C. Gillespie, D. Ferster, The contribution of noise to contrast invariance of orientation tuning in cat visual cortex. *Science* **290**, 1968–1972 (2000).

92. M. Carandini, D. Ferster, Membrane potential and firing rate in cat primary visual cortex. *J. Neurosci.* **20**, 470–484 (2000).

93. K. D. Miller, T. W. Troyer, Neural noise can explain expansive, power-law nonlinearities in neural response functions. *J. Neurophysiol.* **87**, 653–659 (2002).

94. F. S. Chance, L. F. Abbott, A. D. Reyes, Gain modulation from background synaptic input. *Neuron* **35**, 773–782 (2002).

95. B. V. Atallah, W. Bruns, M. Carandini, M. Scanziani, Parvalbumin-expressing interneurons linearly transform cortical responses to visual stimuli. *Neuron* **73**, 159–170 (2012).

96. A. Sanzeni *et al*., Inhibition stabilization is a widespread property of cortical networks. *eLife* **9**, e54875 (2020).

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

98. K. Sohya, K. Kameyama, Y. Yanagawa, K. Obata, T. Tsumoto, GABAergic neurons are less selective to stimulus orientation than excitatory neurons in layer II/III of visual cortex, as revealed by in vivo functional Ca^{2+} imaging in transgenic mice. *J. Neurosci.* **27**, 2145–2149 (2007).

99. A. M. Kerlin, M. L. Andermann, V. K. Berezovskii, R. C. Reid, Broadly tuned response properties of diverse inhibitory neuron subtypes in mouse visual cortex. *Neuron* **67**, 858–871 (2010).

100. D. D. Bock *et al*., Network anatomy and in vivo physiology of visual cortical neurons. *Nature* **471**, 177–182 (2011).

101. E. Fino, R. Yuste, Dense inhibitory connectivity in neocortex. *Neuron* **69**, 1188–1203 (2011).

102. A. M. Packer, R. Yuste, Dense, unspecific connectivity of neocortical parvalbumin-positive interneurons: A canonical microcircuit for inhibition? *J. Neurosci.* **31**, 13260–13271 (2011).

103. J. J. Nassi, M. C. Avery, A. H. Cetin, A. W. Roe, J. H. Reynolds, Optogenetic activation of normalization in alert macaque visual cortex. *Neuron* **86**, 1504–1517 (2015).

104. M. V. Tsodyks, W. E. Skaggs, T. J. Sejnowski, B. L. McNaughton, Paradoxical effects of external modulation of inhibitory interneurons. *J. Neurosci.* **17**, 4382–4388 (1997).

105. D. J. Heeger, Half-squaring in responses of cat striate cells. *Vis. Neurosci.* **9**, 427–443 (1992).

106. Y. Petrov, M. Carandini, S. McKee, Two distinct mechanisms of suppression in human vision. *J. Neurosci.* **25**, 8704–8707 (2005).

NEUROSCIENCE

PSYCHOLOGICAL AND COGNITIVE SCIENCES

www.manaraa.com