A neural model for V1 that incorporates dendritic nonlinearities and back-propagating action potentials

Ilias Rentzeperis[1], Dario Prandi[2], and Marcelo Bertalmío[1]

[1]Spanish National Research Council, Spain

[2]Université Paris-Saclay and CNRS, France

Abstract

The groundbreaking work of Hubel and Wiesel has been instrumental in shaping our understanding of V1, leading to modeling neural responses as cascades of linear and nonlinear processes in what has come to be known as the "standard model" of vision. Under this formulation, however, some dendritic properties cannot be represented in a practical manner, while extensive evidence indicates that dendritic processes are an indispensable element of key neural behaviours. As a result, current V1 models fail to explain neural responses in a number of scenarios. In this work, we propose an implicit model for V1 that considers nonlinear dendritic integration and backpropagation of action potentials from the soma to the dendrites. This is a parsimonious scheme that minimizes an energy, allows for a better conceptual understanding of neural processes, and explains several neurophysiological phenomena that have challenged classical approaches.

1 Introduction

From recordings in the primary visual cortex (V1) of cat and monkey, Hubel and Wiesel categorized neurons into two types: simple cells, which respond maximally to oriented stimuli of a given contrast polarity and at a specific location, and complex cells, which show a degree of response invariance to position and polarity. On account of the neurons' laminar locations and response properties, they proposed a hierarchical model (HM), where simple cells receive thalamic input while complex cells receive input from simple cells [23, 24]. Based on this assumption, the response of the two types of cells were initially described mathematically in distinct ways: simple cells by a linear-nonlinear (LN) model where a linear filter is followed by a nonlinearity [33, 22], and complex cells by an energy model where the outputs of phase-shifted filters are rectified and then added up [1].

While the HM and the notion of cortical cells being either simple or complex have been instrumental in shaping our understanding of V1, both concepts have been challenged by experimental evidence: some studies show that there

are complex cells receiving direct thalamic input, and that there is a continuum from simple to complex cell behaviour [42, 9, 8].

To account for these findings, several current models assume that V1 cells may receive both thalamic and cortical inputs, and their responses are characterized by simple and complex subunits, each one expressed as a bank or a cascade of LN elements [29, 15, 48]. These models attribute complex-like behaviour to a greater relative weight for the complex subunits and a larger contribution of cortical input; however, neurophysiological data shows that some complex cells whose main input is thalamic maintain their complex behavior even when simple cells are silenced [32].

An important aspect of current modeling approaches is that most of them are data-driven; their main goal is obtaining better predictions and they involve a number of filters and several stages so formally they may be equivalent to a deep neural network [7, 13, 2, 27, 4]. Models of this type do not consider the synaptic characteristics of the underlying processes within the cell [18], behaving instead like a black-box [7]. In the past few years, especially with the prominence of deep learning approaches for neuroscience, the call for better conceptual frameworks has increased [35, 17, 43], as models whose mathematical form is based on physiological mechanisms can provide important advantages [7]: (1) they might allow us to gain insights into how specific processes contribute to the overall computation performed by neurons; and (2) they could disentangle complex interactions, revealing localized elements within the model that would otherwise be distributed over several subunits in current approaches.

Regarding these mechanisms, extensive evidence indicates that dendritic processes are an indispensable element of key neural behaviours [36]; they act as nonlinear and semi-independent integration modules, significantly enhancing neuronal computational capacity by executing intricate tasks, having a fundamental role in brain functions such as spatial navigation and perception [40], and even explaining complex cell responses without cortical input [31]. While nonlinear dendritic integration can be modeled as an LN cascade [39], a key property of dendritic processing is that dendrites receive backpropagating action potentials (bAP) initiated from the soma [45], allowing the cell output to interact with the input and turning the activation of the neuron into a complicated function that cannot be represented in a practical manner under an LN formulation, requiring differential equations instead [6, 26].

With all this in mind, here we propose a new mechanistic model for V1 cells. It can be seen as an extension of the classical formulation of the standard model, as it considers not only linear dendrites but also two different types of dendrites with nonlinearities, that operate in a dynamic and input-dependent manner because their processing is affected by bAP (**Fig. 1**). The proposed model provides a novel advance that is conceptually wide and moves the field forward, for two reasons: first, the most successful current approaches are data-driven, not mechanistic, and aside from the limitations they pose in terms of conceptual analysis (as mentioned above), they can't satisfactorily explain a number of physiological results that pose no problem for our model, like the appearance of complex cell behaviour when the input is only thalamic, the effect of stimulus statistics on simple/complex classification, or the effect of stimulus contrast on the phase invariance properties of complex cells; second, our model provides a very good compromise between complexity and amount of mechanistic components considered, while current mechanistic models face a trade-off between level of abstraction and complexity, with models of low complexity considering dendritic nonlinearities of different forms [46] but disregarding bAP and its effect on dendritic computations, while models of less abstraction level that do take into account bAP are much more complex [26], with lots of parameters whose fitting

may be unstable and non-unique [38].

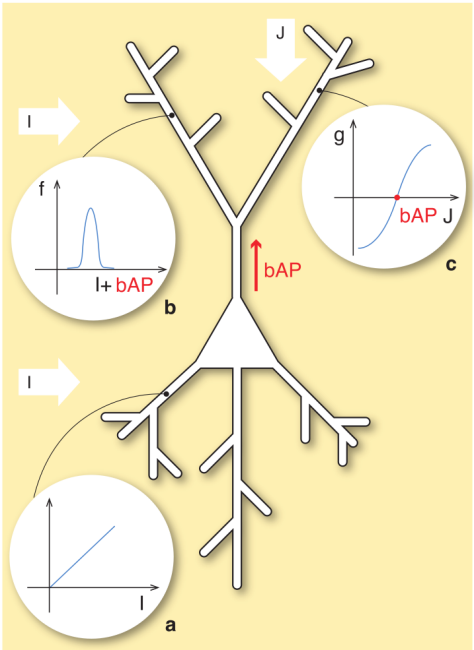

Figure 1: **Proposed model considers different types of dendrites, with nonlinearities and bAP. a**: Dendrites with a linear response to the input I. **b**: Active dendrites with a XOR nonlinearity. The nonlinear response f is large when either the input I or bAP are large, but not both. **c**: Active dendrites with a sigmoid nonlinearity. The nonlinear response g is affected by the value of the bAP and is applied to cortical inputs J from the same hypercolumn.

Results

Proposed model

Our model equation for the activity J at cortical location x with a preferred orientation k is:

$$J(x,k) = \alpha\Phi_k \cdot I(x) + \beta \sum_y w_1(x,y) f(I(y), J(x,k)) + \gamma \sum_j w_2(j,k) g(J(x,j), J(x,k)), \tag{1}$$

where each term corresponds to a different type of dendritic behaviour, weighted by the scalar values α, β, γ.

The first term, where Φ_k is a linear filter and I is the input, represents the filtering operation performed by linear dendrites. If the other, nonlinear terms have a negligible impact ($\beta \simeq \gamma \simeq 0$) then the model is no longer implicit, and becomes equivalent to the classical linear receptive field of simple cells.

In the second term the function f denotes an XOR dendritic nonlinearity, whose existence has recently been reported [21].We hypothesize that at each dendritic location y the XOR function is computed between the value of the input $I(y)$ and the bAP value given by the cell output $J(x,k)$. The linear weights w_1 denote the strength of the connection between each spatial location y and the cell location x.

In the third term the function g denotes a sigmoid nonlinearity that is applied to cortical inputs $J(x,j)$ from neurons in the same hypercolumn. We hypothesize that the bAP affects dendritic processing in this case by shifting the sigmoid nonlinearity by the value $J(x,k)$. The linear weights $w_2(j,k)$ denote the strength of the interactions

between units of preferred orientations j and k.

The model, closely related to the nonlinear summation model proposed in [6], is implicit, because the output $J(x,k)$ appears on both sides of the equation, and therefore finding its solution is not trivial. We have proven that the solution to Eq. 1 exists, minimizes an energy, and is the steady state of the following differential equation (Methods):

$$\frac{\partial J(x,k,t)}{\partial t} = -J(x,k,t) + \alpha\Phi_k \cdot I(x) + \beta \sum_y w_1(x,y) f(I(y), J(x,k,t)) + \gamma \sum_j w_2(j,k) g(J(x,j,t), J(x,k,t)) \quad (2)$$

Model instance chosen for experiments

To demonstrate the potential of the model we test its ability to predict a number of phenomena, observed in physiological experiments, that other models have difficulty in explaining. Very importantly, for each example all results are obtained with the same instance of the model, i.e. neither the structure of the model nor its parameters need to be adapted to the input, in contrast with other methods.

For the first term, and in consonance with classical formulations [14], we select the linear filter Φ_k to be a Gabor function with preferred orientation k.

For the second term, we take into account that XOR dendrites are apical [21] and that orientation tuning aligns with dendritic morphology in apical dendrites [49], and we select $w_1(x,y)$ to be a Gabor function with preferred orientation k, although with different spatial frequency and phase than Φ_k (Methods).

For the third term, and again following the classical view [5], we choose for $w_2(j,k)$ a function that decreases monotonically with increasing difference between the orientations j and k.

We consider the input I to be thalamic and assume that it comes from cells that are not orientation selective, as classically established for cat and primate [25], while the inputs $J(x,j)$ are cortical.

Finally, we normalize the coefficients, expressing them in the following way:

$$\alpha = \frac{1}{1+\lambda_1+\lambda_2}, \ \beta = \frac{\lambda_1}{1+\lambda_1+\lambda_2}, \ \gamma = -\frac{\lambda_2}{1+\lambda_1+\lambda_2},$$

with $\lambda_1 \geq 0$ regulating the impact of the first nonlinear term and $\lambda_2 \geq 0$ that of the second nonlinear term.

Model explains a continuum of simple to complex cell behaviour, even when input is only thalamic

Here we set $\lambda_2 = 0$ so that we can consider the input to be fully thalamic.

The first term of the model explains simple cell behaviour: setting $\lambda_1 = 0$, responses show orientation tuning but not phase invariance (**Fig. 2a**).

The second term of the model explains complex cell behaviour: for large values of λ_1 (so that the impact of the second term is predominant) the results show both orientation tuning and phase invariance (**Fig. 2b**).

We must remark that the model can explain complex cell responses for an input that is only thalamic, which is consistent with neurophysiology but a challenge for standard approaches.

These results do not hold for arbitrary choices of weight function w_1 or nonlinearity f: for example, if w_1 does not have an orientation preference then the result does not present orientation tuning (**Fig. 2d**), and if f is a sigmoid instead of an XOR nonlinearity then the result is not phase invariant (**Fig. 2e**).

As we increase the value of λ_1, and with it the contribution of the XOR dendrites, the model is able to explain a continuum of simple to complex cell behaviour (**Fig. 2c**).

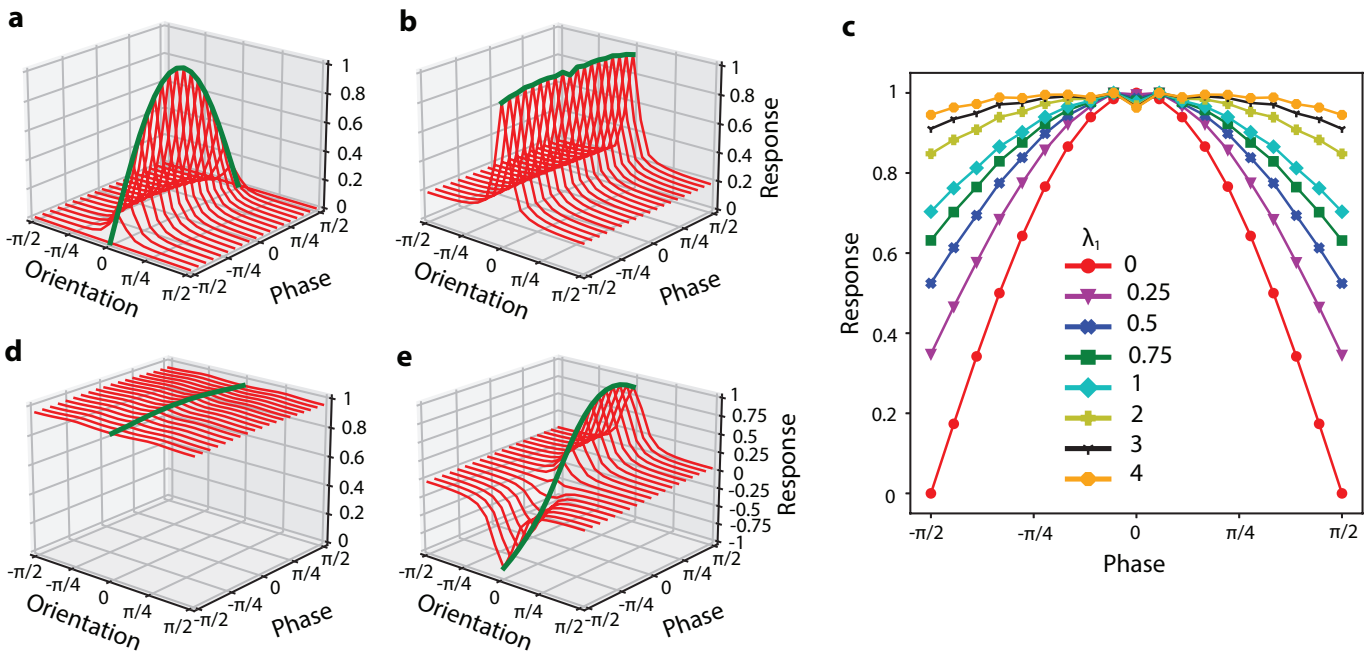

Figure 2: **Varying the contribution of XOR dendrites explains a continuum of simple to complex cell behaviour.** **a**: Without XOR dendrites ($\lambda_1 = 0$), the model explains simple cell behaviour: orientation tuning but no phase invariance. For a fixed phase, the plot shows the 2-D orientation tuning curve for stimuli of that phase. The green envelope shows the maximum responses of the model for stimuli of the same orientation (phase responses). **b**: When the contribution of XOR dendrites is predominant over the first term, the model explains complex cell behaviour: orientation tuning and phase invariance. **c**: Phase responses for different weights for the second term of the model, showing a gradual shift from simple to complex cell behaviour. **d**: The model loses its ability to explain complex cell behaviour if the linear filter w_1 in the second term is replaced by a non-oriented filter: the results now do not show orientation tuning. **e**: The model loses its ability to explain complex cell behaviour if the XOR nonlinearity is replaced by a sigmoid nonlinearity: the results now do not show phase invariance.

Model explains orientation tuning sharpening through cortical interactions

In the previous section we considered the case where the model was missing its third term, but it should be noted that the full model, with its third term present, is still able to represent a continuum of simple to complex cell responses that is obtained by varying the relative weight of the second term of the model (**Fig. 3a**). But now, the third term of the model allows it to explain the increase in orientation selectivity provided by cortical interactions, which is in agreement with physiology [44]: as the contribution of this term increases, the neural output becomes more sharply tuned (**Fig. 3b**). Conversely, removal of sigmoidal dendrites from our model prevents it from explaining orientation tuning sharpening.

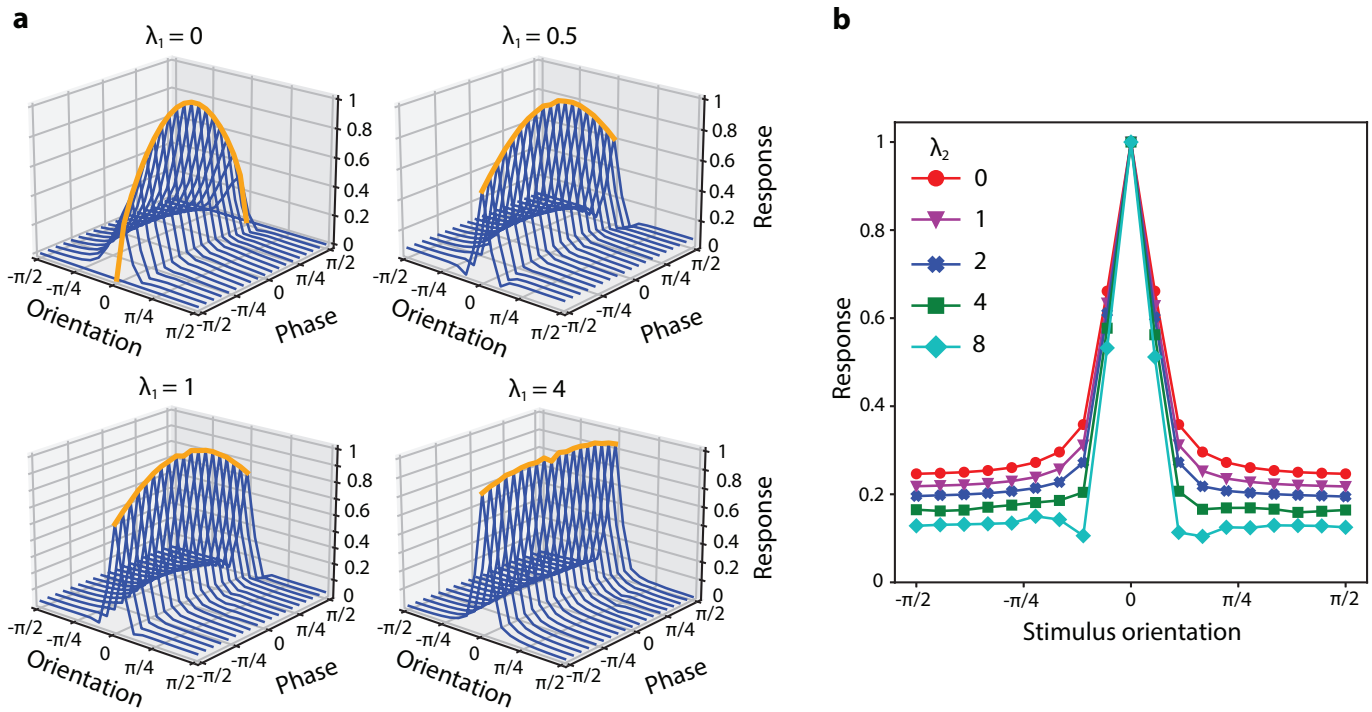

Figure 3: **The third term in the model allows to explain orientation tuning sharpening.** **a**: The model still produces a continuum from simple to complex cell responses with the addition of the third term: we fixed $\lambda_2 = 1$ and varied λ_1 as in Fig 2. **b**: Recurrent competition between units sharpens their orientation tuning curves. We set $\lambda_1 = 1$ and vary λ_2. Our tests consider 18 neural units, with orientation selectivities equally spaced between 0 and 2π that interact through the action of the third term of the model, where the linear weights w_2 are large for similar orientations and small for orientations that are very different (Methods). As λ_2 increases, the neural output becomes more sharply tuned.

Model explains that a neuron's classification as simple or complex is affected by input statistics

A puzzling phenomenon first reported in [16] is that the same cell has more simple-like behaviour for some inputs, like dense noise, but more complex-like for others, like sparse noise. Our model is able to explain these results (**Fig. 4a**), with a good qualitative correspondence to experimental data (**Fig. 4b**).

In [16] the phenomenon is explained with a model whose parameters change with the input stimulus, and it is conjectured that sparse stimuli recruit lateral interactions, which contribute complex-like nonlinearities. Our model, on the other hand, is able to remain constant for different inputs and works even if all inputs are thalamic and there are no cortical interactions, as in this experiment we have used fixed parameters and $\lambda_2 = 0$.

We point out that the presence of XOR dendrites in our model is essential for it to be able to explain this phenomenon, given that the representation of complex cell behaviour relies on them, as shown previously: when the term with XOR dendrites is missing ($\lambda_1 = 0$) the simpleness index is 1 for both types of input, indicating simple cell behaviour, and as λ_1 increases the cell tends towards a more complex-like behaviour (**Fig. 4a**), in agreement with what was observed before (**Fig. 2**).

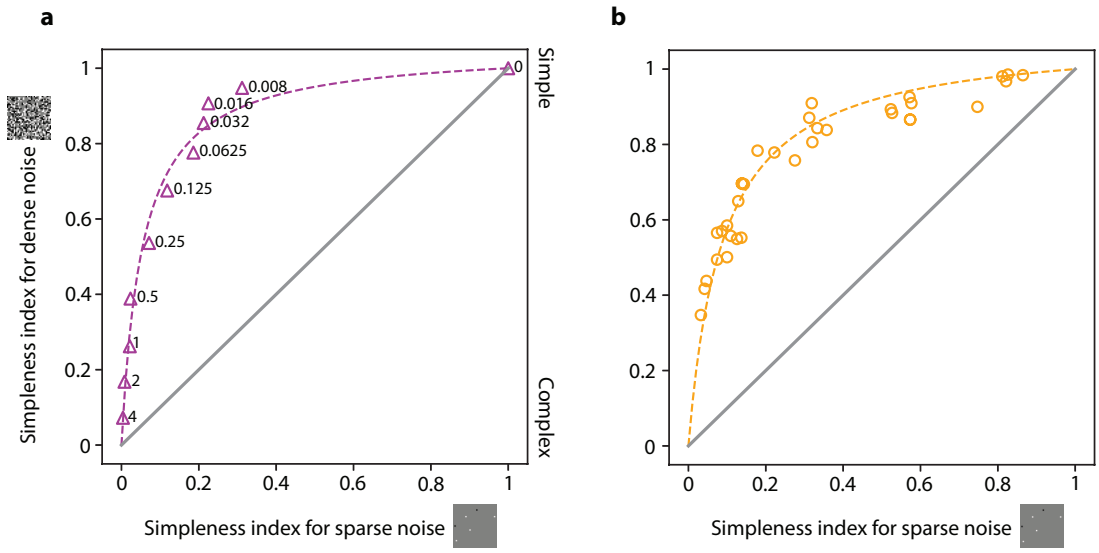

Figure 4: **The model, with fixed parameters, explains that simple/complex classification is affected by input statistics. a**: We consider the first two terms of the model (i.e. we take $\lambda_2 = 0$) and, in analogy with [16], we define a simpleness index as the contribution of the first term divided by the whole response. We generate random inputs of sparse and dense noise of dimension 33×33. Sparse noise stimuli were constructed by randomly interspersing either gray (same as the background), white, or black elements with equal probability, where the average number of elements that contributed to the noise (white or black) were 7.26 out of the total 1089 elements. Dense noise stimuli were generated by randomly interspersing either gray, white or black elements with an equal probability of 0.33. We consider a wide range of values for λ_1; for each of them and for each type of noise we generate 100 random inputs, apply our model to them and compute the associated simpleness indexes, which are then averaged to produce the final simpleness value. For all values of λ_1 (each depicted next to its associated data point), the model shows that the cell will have a more simple-like behaviour (larger simpleness value) for dense noise input than for sparse noise input. We can also see that for $\lambda_1 = 0$ the simpleness index is 1 for both types of input, indicating simple cell behaviour, and as λ_1 increases the cell tends towards a more complex-like behaviour, in agreement with what was observed in **Fig. 2**. In the study by [16], the experimental data points followed the arc of a continuous smooth curve defined as follows: $SI_{DN} = \frac{SI_{SN}}{SI_{SN}+(\frac{1}{\kappa})^2(1-SI_{SN})}$, where SI_{DN} and SI_{SN} are the simpleness indices for dense and sparse noise respectively and κ is a free parameter. Using the Levenberg-Marquardt algorithm we found the κ that provides the best fit (dashed curve) to the data points produced by our model. **b**: Experimental data adapted from fig. 3a in [16], (along with the corresponding best fit curve) which in that paper was explained using a model with input-dependent parameters. Each data point represents a different cell. Our model provides a very good qualitative match to this data, despite using fixed parameters.

Model explains that input contrast can affect phase invariance in complex cells

Extracellular recordings [12, 47] show that complex cell responses become sensitive to phase when the stimulus contrast decreases, a property that is more common on superficial cortical layers [30], that receive most of the thalamic input, and is not so prominent on deeper layers, where the input is mainly cortical, coming from cells with a preferred orientation; also, the contrast dependence of phase invariance virtually disappears for higher cortical areas [11].

In order to account for these properties, the proposed models have limitations: variations on the HM [47] are not fully consistent with neurophysiological data [30], models based on recurrent amplification [51] can't explain complex cell behaviour when the input is just thalamic, and data-driven models need to adapt the structure and number of their filters and the form of their nonlinearities to the different inputs [2]. Our proposed model, on the other hand, does not have these issues, and for thalamic input ($\lambda_2 = 0$) we can achieve a good qualitative correspondence with respect to experimental data with a fixed set of parameters and for a wide range of values of λ_1 (**Fig. 5a**). If the input I were cortical (simulated as thalamic input passed through an oriented filter), the model shows that the influence of input

contrast on phase invariance is reduced considerably, as expected (**Fig. 5b**). The proposed model can also explain the fact that phase invariance in mice complex cells is not too affected by signal contrast [52], as many thalamic cells in mice are orientation selective.

Again, our model relies on the presence of XOR dendrites in order to explain these complex cell properties.

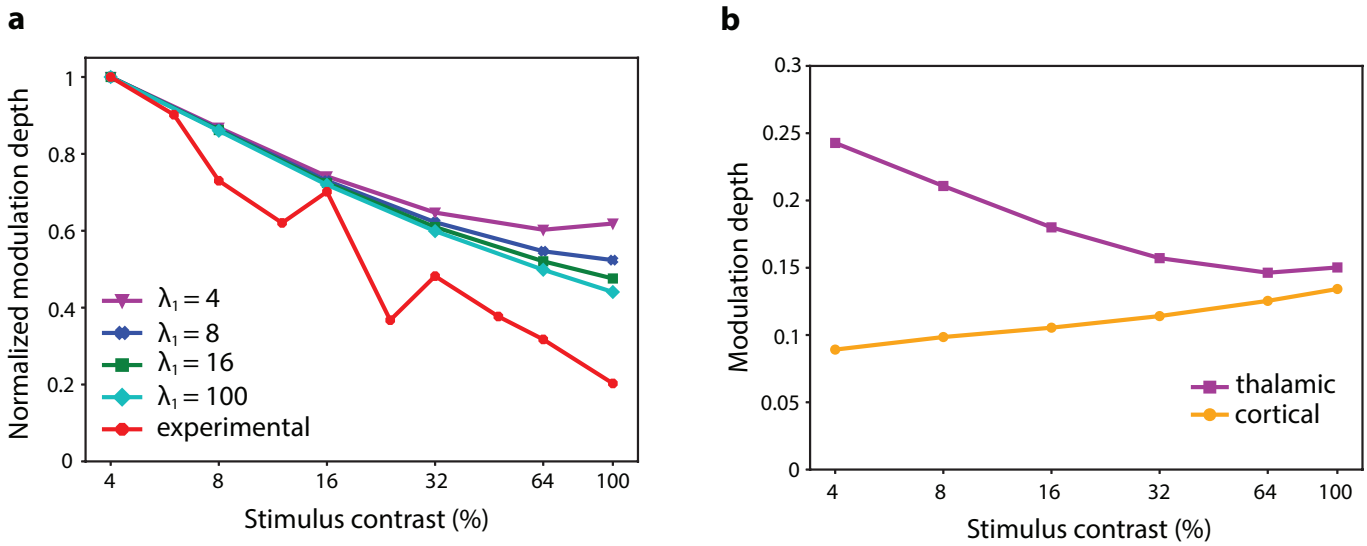

Figure 5: **The model, with fixed parameters, explains that stimulus contrast can affect phase invariance.** **a**: Normalized modulation depth as a function of stimulus contrast, for different values of λ_1 (we ignore the third term of the model, i.e. we consider $\lambda_2 = 0$). Modulation depth is typically defined as the change between the maximum and the minimum in the phase response profile, relative to the maximum, so a complex cell with perfect phase invariance would have a modulation depth of 0, and the larger the modulation depth the more sensitive the cell is to phase changes; here we normalize the result by the modulation depth value obtained for the minimum stimulus contrast considered (4%). The plot also shows experimental data adapted from fig. 3c in [30] and, as we can see, our model provides a good qualitative correspondence to it, with modulation depth decreasing linearly as the stimulus contrast increases logarithmically. **b**: We set $\lambda_1 = 4$ and consider the cases where the input in the second term is assumed to be thalamic (purple plot) or cortical (orange plot; cortical input from a simple cell is simulated by filtering thalamic stimulus with a Gabor function with the same orientation preference as Φ_k). As the figure shows, the model can also explain the phenomenon by which complex cells which receive cortical but not thalamic input have a phase response which is less affected by changes in input contrast [11].

Discussion

We have put forward a formulation for V1 cell responses that can be seen as a contribution in the search of a *consensus model* for V1 [28], as it combines ideas from hierarchical, parallel and recurrent models while relying on a common ground circuitry. The proposed model is mechanistic, relying on dendritic nonlinearities and bAP for its nonlinear behavior, and hence we see it as biologically justified, whereas many current models for complex cells assume nonlinear processes that do not seem to have a biological counterpart [27].

Current data-driven neural models involve a number of filters and several stages and may be formally equivalent to a deep neural network. Our mechanistic model is more parsimonious, with only three filters in a single stage, but despite its reduced complexity it is able to make a number of predictions which are supported by physiological results, while these experimental data challenge current approaches. For the experiments in this paper we have used a model instance with hand-picked (although biologically plausible) filters and nonlinearities, but it is easy to see that all these

elements can be considered model parameters and learned so as to fit some set of experimental data, making the model data-driven.

Our model also offers several conceptual advantages. First, a better understanding of neural processes, through a formulation that is simple but also flexible enough to be able to explain very different phenomena using a reduced and fixed set of parameters that do not need to change with the input. Second, modularity and adaptability, because new terms can be added representing other types of dendrites, and the elements of each term (filters, nonlinearities) can be updated when new experimental data becomes available. Third, our proposed model (and its potential extensions) may assist in the investigation of bAP and its roles; given the significant difficulties associated with this line of research, it's been suggested that an optimal approach is a synergistic loop involving experiments and computational models [20], and in this work we have followed this route by proposing a model that hypothesizes how bAP may affect different types of nonlinear dendrites, and then observing how the predictions yielded by the model align with experimental outcomes.

The fact that our proposed model minimizes an energy points to a connection with the efficient representation principle [3] and the free energy principle [19], and the orientation tuning effect of the third term of the model relates it directly with approaches for sparse coding [41, 37, 34]. Interestingly, while our model can represent a continuum of simple to complex cell behaviour, efficient coding implemented in a biologically plausible way is able to learn simple but not complex cell properties [27]. Also of note, the gradient descent equation for the energy associated to our model (the differential equation providing the output of the model, Eq. 2) can be seen as a generalization of the classical Wilson-Cowan formulation [50, 10], now with different types of nonlinearities that are input dependent.

Methods

Energy formulation

We recall that the implicit model considered in this paper is the following,

$$J(x,k) = \alpha \Phi_k \cdot I(x) + \beta \sum_y w_1(x,y) f(I(y), J(x,k)) + \gamma \sum_j w_2(j,k) g(J(x,j), J(x,k)) \tag{3}$$

The function f is a XOR nonlinearity that we model as $f(I(y), J(x,k)) = \sigma_1(J(x,k) - I(y))$ where $\sigma_1(s)$ is an even positive function with $\sigma_1(s) \to 1$ as $s \to +\infty$ and $\sigma_1(s) = 0$. This choice yields a minimal response (~ 0) when $J(x,k) \sim I(y)$, which increases as the difference between the two inputs increases (see Figure 1). The nonlinearity g is considered to be an odd bounded function with bias $J(x,k)$, i.e., $g(a,b) = \sigma_2(a-b)$ where σ_2 is bounded and $\sigma_2(s) = -\sigma_2(s)$. Finally, we assume that the filters Φ_k are L^2-normalized (i.e., $\|\Phi_k\|_2 = 1$), that $w_1(x,y) = w_1(x-y)$, and that $w_2(j,k) = |\langle \Phi_j, \Phi_k \rangle|$.

Solutions to Eq. 3 are steady states of the following differential equation:

$$\frac{\partial J(x,k,t)}{\partial t} = -J(x,k,t) + \alpha \Phi_k \cdot I(x) + \beta \sum_y w_1(x,y) f(I(y), J(x,k,t)) + \gamma \sum_j w_2(j,k) g(J(x,j,t), J(x,k,t)) \tag{4}$$

Let us denote by Σ_1 and Σ_2 the primitives of σ_1 and σ_2, respectively. E.g., $\Sigma_1(s) = \int_0^s \sigma_1(\tau)\, d\tau$. Under these assumptions we have the following.

Proposition 1. *Eq. 4 is the gradient descent equation corresponding to the following energy :*

$$E(J) = \sum_k \left[\frac{1}{2} \|J(\cdot,k) - \alpha \Phi_k \cdot I\|_2^2 - \beta \sum_{x,y} w_1(x,y) \Sigma_1(J(x,k) - I(y)) \right] - \frac{\gamma}{2} \sum_x \sum_{j,k} w_2(j,k) \Sigma_2(J(x,j) - J(x,k)). \quad (5)$$

Namely, for any position x, orientation k, and time t, it holds

$$\frac{\partial J(x,k,t)}{\partial t} = -\frac{\partial E(J(\cdot,\cdot,t))}{\partial J(x,k)}. \quad (6)$$

Moreover, the energy E admits a global minimum.

Proof. We start by providing and argument for Eq. 6. If $\gamma = 0$ this equation follows from Eq. 4 and the direct computation of the gradient of E. Let us denote by E_2 the factor of γ in the expression for E, namely

$$E_2(J) = \frac{1}{2} \sum_x \sum_{j,k} w_2(j,k) \Sigma_2(J(x,j) - J(x,k)).$$

Then, we have

$$\frac{\partial E_2(J)}{\partial J(z,h)} = \frac{1}{2} \sum_k w_2(h,k) \sigma_2(J(z,h) - J(z,k)) - \frac{1}{2} \sum_j w_2(j,h) \sigma_2(J(z,j) - J(z,h)) = \sum_k w_2(h,k) \sigma_2(J(z,h) - J(z,k)). \quad (7)$$

Here, we used the fact that $w_2(j,h) = w_2(h,j)$ and that $\sigma_2(-s) = -\sigma_2(s)$. By linearity of the differential, this completes the proof of Eq. 6.

To complete the proof of the statement, we are left to show the existence of a global minimum. Since J belongs to a finite dimensional space (real-valued functions from the couples (x,k) of discretized positions and orientations), this follows from a lower bound of the form

$$E(J) \geq \sum_k \|J(x,k)\|^2 - a \sqrt{\sum_k \|J(x,k)\|^2} - b, \quad (8)$$

for some $a, b \in \mathbb{R}$.

To prove Eq. 8, we start by observing that $\|\Phi_k \cdot I\|_2 \leq \|I\|_2$ due to the normalization $\|\Phi_k\|_2 = 1$. Hence, letting $K \in \mathbb{N}$ be the total number of orientations, by the Cauchy-Schwarz inequality we obtain

$$\frac{1}{2} \sum_k \|J(\cdot,k) - \alpha \Phi_k \cdot I\|_2^2 \geq \sum_k \|J(x,k)\|^2 - 2\alpha \|I\|_2 \sqrt{\sum_k \|J(x,k)\|^2} - K\alpha^2 \|I\|^2. \quad (9)$$

On the other hand, thanks to the fact that $\Sigma_1(s) \leq \|\sigma_1\|_\infty |s|$ and $\Sigma_2(s) \leq \|\sigma_2\|_\infty |s|$, for any β and γ there exists constants $c_1, c_2 \in \mathbb{R}$ that depend on $\|w_1\|_\infty$, $\|w_2\|_\infty$, and the total number of positions N and orientations K, such

that

$$-\beta \sum_k \sum_{x,y} w_1(x,y)\Sigma_1(J(x,k) - I(y)) - \gamma \sum_x \sum_{j,k} w_2(j,k)\Sigma_2(J(x,j) - J(x,k)) \geq c_1 \sqrt{\sum_k \|J(x,k)\|^2} + c_2\|I\|_2. \quad (10)$$

The lower bound in Eq. 8 then follows by putting together Eq. 9 and Eq. 10. □

The following result insures that the numerical solution of the differential equation Eq. 4 yields pertinent results.

Proposition 2. *There exists an explicit set A, that is invariant for the flow of Eq. 4. Moreover, all solutions of Eq. 3 are contained in A.*

Proof. Consider the change of variables $H(x,k) = J(x,k) - \alpha\Phi_k \cdot I(x)$. Eq. 4 then reads

$$\begin{aligned} \frac{\partial H(x,k,t)}{\partial t} &= -H(x,k,t) + \beta \sum_y w_1(x,y)\,\sigma_1(H(x,k,t) - \tilde{I}(y)) + \gamma \sum_j w_2(j,k)\,\sigma_2(H(x,j,t) - H(x,k,t)) \\ &= -H(x,k,t) + V(H(\cdot,\cdot,t)) \end{aligned} \quad (11)$$

Here, the vector field V is defined by the last equality, and we let $\tilde{I} = I - \alpha\Phi_k \cdot I$. We have that

$$|V(H)| \leq R := \beta\|\sigma_1\|_\infty\|w_1\|_1 + \gamma\|\sigma_2\|_\infty\|w_2\|_1. \quad (12)$$

In particular, since solutions of Eq. 3 satisfy $H(x,k) = V(H(x,k))$, this implies that they must belong to the ball $B_R(0)$. Taking into account the change of variable, this proves the second part of the statement for $A := B_R(I - \tilde{I})$.

To complete the proof of the statement, it suffices to differentiate $\|H(\cdot,\cdot,t)\|_2^2$ with respect to time. Indeed, using Eq. 11, we have

$$\frac{1}{2}\frac{\partial\|H(\cdot,\cdot,t)\|_2^2}{\partial t} = -\|H(\cdot,\cdot,t)\|_2^2 - H(\cdot,\cdot,t)\cdot V(H(\cdot,\cdot,t)) \geq -\|H(\cdot,\cdot,t)\|_2\left(\|H(\cdot,\cdot,t)\|_2 - R\right). \quad (13)$$

Since the quantity on the right-hand side is null for $\|H(\cdot,\cdot,t)\| = R$ and negative for $H(\cdot,\cdot,t) \in B_R(0)$, this shows that $B_R(0)$ is invariant w.r.t. the dynamics Eq. 11. As a consequence, the previously defined set $A = B_R(I - \tilde{I})$ is invariant w.r.t. the dynamics Eq. 4. □

Experiments

For all our experiments we have considered an instance of our model where:

$$\alpha = \frac{1}{1+\lambda_1+\lambda_2}, \quad \beta = \frac{\lambda_1}{1+\lambda_1+\lambda_2}, \quad \gamma = -\frac{\lambda_2}{1+\lambda_1+\lambda_2}. \quad (14)$$

For the reduced model (without the recurrent term), Φ_k is a Gabor filter positioned at the center of an $N \times N$ frame of gray background with $N = 257$ pixels. The grating has a preferred orientation equal to zero and a period that is one fifth the size of the frame ($L = N/5 = 51.4$ pixels; thus the spatial frequency is $SF = 5/N$ cycles per

pixel). It is windowed by a Gaussian envelope with aspect ratio $\gamma = 0.125$ and standard deviation $\text{std} = L/8 = 6.425$ pixels.

The kernel $w_1(x, y)$ takes at y the value of a Gabor filter centered at x, which has the same orientation as Φ_k but is phase shifted by -90 degrees. The windowing Gaussian envelope in this case has a standard deviation that is 3 times larger than the one of Φ_k (std $= 19.275$).

According to the previous section, we consider $f(I(y), J(x, k, t)) = \sigma_1(I(y) - J(x, k, t))$ where $\sigma_1(z) = \frac{|z|^p}{|z|^p+q}$ with $p = 1$, and $q = 0.01$; $w_2(j, k)$ is given by the absolute value of the inner product of the filters Φ_j and Φ_k; $g(J(x, j, t), J(x, k, t)) = \sigma_2(J(x, j, t) - J(x, k, t))$, where the sigmoid σ_2 is taken as $\sigma_2(x) = \tanh(\frac{x}{2})$.

We numerically implement the differential equation Eq. 4 on a time interval $[0, T]$ via a forward Euler scheme with a time-discretization of $\Delta t_n \leq 0.1$, $n \in \{1, \ldots, 19\}$. The time discretization Δt_n decays according to a time-based decay schedule: $\Delta t_n = 0.1/(1 + (n - 1)d)$ where $d = 0.02$. This yields $T = \sum_{n=1}^{19} \Delta t_n \simeq 1.62$. The initial condition is fixed at $J(x, k, 0) = 0$.

It is physiologically plausible to assume a delay in the onset of the bAP and the cortical interaction [45, 44], so we initially consider $\beta = 0$ and $\gamma = 0$, and activate them sequentially at times $T_1 = \Delta t_1 + \Delta t_2$ and $T_2 = \sum_{n=1}^{8} \Delta t_n$. More precisely, we consider

$$\alpha(t) = \frac{1}{1 + \lambda_1 + \lambda_2}, \quad \beta(t) = \begin{cases} 0 & \text{if } t \leq T_1, \\ -\frac{\lambda_1}{1+\lambda_1+\lambda_2} & \text{otherwise,} \end{cases} \quad \gamma(t) = \begin{cases} 0 & \text{if } t \leq T_2, \\ -\frac{\lambda_2}{1+\lambda_1+\lambda_2} & \text{otherwise.} \end{cases} \tag{15}$$

Effectively, this can be interpreted as the numerical implementation of Eq. 4, with the parameters defined in Eq. 14, on the time interval $[T_2, T]$ with the pre-processed initial condition $J(x, T_2)$. Thresholding (i.e., not fixing the final time T, but stopping when $\|J(\cdot, \cdot, T + \Delta t_n) - J(\cdot, \cdot, T)\| < \varepsilon$ for some $\varepsilon > 0$) provides results yielding the same qualitative conclusions, but we have found that running the algorithm a fixed number of iterations provides a good trade-off between accuracy and speed.

References

[1] Edward H Adelson and James R Bergen. Spatiotemporal energy models for the perception of motion. *Josa a*, 2(2):284–299, 1985.

[2] Ali Almasi, Hamish Meffin, Shaun L Cloherty, Yan Wong, Molis Yunzab, and Michael R Ibbotson. Mechanisms of feature selectivity and invariance in primary visual cortex. *Cerebral Cortex*, 30(9):5067–5087, 2020.

[3] Horace Barlow. Possible principles underlying the transformation of sensory messages. *Sensory communication*, 1:217–234, 1961.

[4] Felix Bartsch, Bruce G Cumming, and Daniel A Butts. Model-based characterization of the selectivity of neurons in primary visual cortex. *Journal of Neurophysiology*, 128(2):350–363, 2022.

[5] Rani Ben-Yishai, R Lev Bar-Or, and Haim Sompolinsky. Theory of orientation tuning in visual cortex. *Proceedings of the National Academy of Sciences*, 92(9):3844–3848, 1995.

[6] Marcelo Bertalmío, Alex Gomez-Villa, Adrián Martín, Javier Vazquez-Corral, David Kane, and Jesús Malo. Evidence for the intrinsically nonlinear nature of receptive fields in vision. *Scientific reports*, 10(1):16277, 2020.

[7] Daniel A Butts. Data-driven approaches to understanding visual neuron activity. *Annual review of vision science*, 5:451–477, 2019.

[8] Matteo Carandini, Jonathan B Demb, Valerio Mante, David J Tolhurst, Yang Dan, Bruno A Olshausen, Jack L Gallant, and Nicole C Rust. Do we know what the early visual system does? *Journal of Neuroscience*, 25(46):10577–10597, 2005.

[9] Xiaodong Chen, Feng Han, Mu-ming Poo, and Yang Dan. Excitatory and suppressive receptive field subunits in awake monkey primary visual cortex (V1). *Proceedings of the National Academy of Sciences*, 104(48):19120–19125, 2007.

[10] Carson C Chow and Yahya Karimipanah. Before and beyond the wilson–cowan equations. *Journal of neurophysiology*, 123(5):1645–1656, 2020.

[11] Shaun L Cloherty and Michael R Ibbotson. Contrast-dependent phase sensitivity in V1 but not v2 of macaque visual cortex. *Journal of neurophysiology*, 113(2):434–444, 2015.

[12] Nathan A Crowder, Joshua Van Kleef, Bogdan Dreher, and Michael R Ibbotson. Complex cells increase their phase sensitivity at low contrasts and following adaptation. *Journal of neurophysiology*, 98(3):1155–1166, 2007.

[13] Saskia EJ de Vries, Jerome Lecoq, Michael A Buice, Peter A Groblewski, Gabriel K Ocker, Michael Oliver, David Feng, Nicholas Cain, Peter Ledochowitsch, Daniel Millman, et al. A large-scale, standardized physiological survey reveals higher order coding throughout the mouse visual cortex. *BioRxiv*, page 359513, 2018.

[14] David J Field and David J Tolhurst. The structure and symmetry of simple-cell receptive-field profiles in the cat's visual cortex. *Proceedings of the Royal society of London. Series B. Biological sciences*, 228(1253):379–400, 1986.

[15] Julien Fournier, Cyril Monier, Manuel Levy, Olivier Marre, Katalin Sári, Zoltán F Kisvárday, and Yves Frégnac. Hidden complexity of synaptic receptive fields in cat V1. *Journal of Neuroscience*, 34(16):5515–5528, 2014.

[16] Julien Fournier, Cyril Monier, Marc Pananceau, and Yves Frégnac. Adaptation of the simple or complex nature of V1 receptive fields to visual statistics. *Nature neuroscience*, 14(8):1053–1060, 2011.

[17] Yves Frégnac. Big data and the industrialization of neuroscience: A safe roadmap for understanding the brain? *Science*, 358(6362):470–477, 2017.

[18] Yves Frégnac and Brice Bathellier. Cortical correlates of low-level perception: from neural circuits to percepts. *Neuron*, 88(1):110–126, 2015.

[19] Karl Friston. The free-energy principle: a unified brain theory? *Nature reviews neuroscience*, 11(2):127–138, 2010.

[20] Sonia Gasparini and Michele Migliore. Action potential back-propagation. In *Encyclopedia of computational neuroscience*, pages 151–156. Springer, 2022.

[21] Albert Gidon, Timothy Adam Zolnik, Pawel Fidzinski, Felix Bolduan, Athanasia Papoutsi, Panayiota Poirazi, Martin Holtkamp, Imre Vida, and Matthew Evan Larkum. Dendritic action potentials and computation in human layer 2/3 cortical neurons. *Science*, 367(6473):83–87, 2020.

[22] David J Heeger. Half-squaring in responses of cat striate cells. *Visual neuroscience*, 9(5):427–443, 1992.

[23] David H. Hubel and Torsten N. Wiesel. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of physiology*, 160(1):106, 1962. Publisher: Wiley-Blackwell.

[24] David H Hubel and Torsten N Wiesel. Receptive fields and functional architecture of monkey striate cortex. *The Journal of physiology*, 195(1):215–243, 1968.

[25] David Hunter Hubel and Torsten Nils Wiesel. Ferrier lecture-functional architecture of macaque monkey visual cortex. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 198(1130):1–59, 1977.

[26] Matthew E Larkum. Are dendrites conceptually useful? *Neuroscience*, 489:4–14, 2022.

[27] Yanbo Lian, Ali Almasi, David B Grayden, Tatiana Kameneva, Anthony N Burkitt, and Hamish Meffin. Learning receptive field properties of complex cells in V1. *PLoS computational biology*, 17(3):e1007957, 2021.

[28] Luis M Martinez and Jose-Manuel Alonso. Complex receptive fields in primary visual cortex. *The neuroscientist*, 9(5):317–331, 2003.

[29] James M McFarland, Yuwei Cui, and Daniel A Butts. Inferring nonlinear neuronal computation based on physiologically plausible inputs. *PLoS computational biology*, 9(7):e1003143, 2013.

[30] Hamish Meffin, Markus Antero Hietanen, Shaun L Cloherty, and Michael Robert Ibbotson. Spatial phase sensitivity of complex cells in primary visual cortex depends on stimulus contrast. *Journal of neurophysiology*, 114(6):3326–3338, 2015.

[31] Bartlett W Mel, Daniel L Ruderman, and Kevin A Archie. Translation-invariant orientation tuning in visual "complex" cells could derive from intradendritic computations. *Journal of Neuroscience*, 18(11):4325–4334, 1998.

[32] Marc Mignard and Joseph G Malpeli. Paths of information flow through visual cortex. *Science*, 251(4998):1249–1251, 1991.

[33] J Anthony Movshon, Ian D Thompson, and David J Tolhurst. Spatial summation in the receptive fields of simple cells in the cat's striate cortex. *The Journal of physiology*, 283(1):53–77, 1978.

[34] Ian Antón Oldenburg, William D Hendricks, Gregory Handy, Kiarash Shamardani, Hayley A Bounds, Brent Doiron, and Hillel Adesnik. The logic of recurrent circuits in the primary visual cortex. *Nature Neuroscience*, pages 1–11, 2024.

[35] Bruno A Olshausen. 20 years of learning about vision: Questions answered, questions unanswered, and questions not yet asked. In *20 Years of Computational Neuroscience*, pages 243–270. Springer, 2013.

[36] Michalis Pagkalos, Roman Makarov, and Panayiota Poirazi. Leveraging dendritic properties to advance machine learning and neuro-inspired computing. *Current opinion in neurobiology*, 85:102853, 2024.

[37] Dylan M Paiton, Charles G Frye, Sheng Y Lundquist, Joel D Bowen, Ryan Zarcone, and Bruno A Olshausen. Selectivity and robustness of sparse coding networks. *Journal of vision*, 20(12):10–10, 2020.

[38] Konstantinos-Evangelos Petousakis, Anthi A Apostolopoulou, and Panayiota Poirazi. The impact of hodgkin–huxley models on dendritic research. *The Journal of physiology*, 601(15):3091–3102, 2023.

[39] Panayiota Poirazi, Terrence Brannon, and Bartlett W Mel. Pyramidal neuron as two-layer neural network. *Neuron*, 37(6):989–999, 2003.

[40] Panayiota Poirazi and Athanasia Papoutsi. Illuminating dendritic function with computational models. *Nature Reviews Neuroscience*, 21(6):303–321, 2020.

[41] Christopher J Rozell, Don H Johnson, Richard G Baraniuk, and Bruno A Olshausen. Sparse coding via thresholding and local competition in neural circuits. *Neural computation*, 20(10):2526–2563, 2008.

[42] Nicole C Rust, Odelia Schwartz, J Anthony Movshon, and Eero P Simoncelli. Spatiotemporal elements of macaque V1 receptive fields. *Neuron*, 46(6):945–956, 2005.

[43] Andrew Saxe, Stephanie Nelli, and Christopher Summerfield. If deep learning is the answer, what is the question? *Nature Reviews Neuroscience*, 22(1):55–67, 2021.

[44] Robert Shapley, Michael Hawken, and Dario L Ringach. Dynamics of orientation selectivity in the primary visual cortex and the importance of cortical inhibition. *Neuron*, 38(5):689–699, 2003.

[45] Greg J Stuart and Nelson Spruston. Dendritic integration: 60 years of progress. *Nature neuroscience*, 18(12):1713–1721, 2015.

[46] Alexandra Tzilivaki, George Kastellakis, and Panayiota Poirazi. Challenging the point neuron dogma: Fs basket cells as 2-stage nonlinear integrators. *Nature communications*, 10(1):3664, 2019.

[47] Joshua P Van Kleef, Shaun L Cloherty, and Michael R Ibbotson. Complex cell receptive fields: evidence for a hierarchical mechanism. *The Journal of physiology*, 588(18):3457–3470, 2010.

[48] Brett Vintch, J Anthony Movshon, and Eero P Simoncelli. A convolutional subunit model for neuronal responses in macaque V1. *Journal of Neuroscience*, 35(44):14829–14841, 2015.

[49] Simon Weiler, Drago Guggiana Nilo, Tobias Bonhoeffer, Mark Hübener, Tobias Rose, and Volker Scheuss. Orientation and direction tuning align with dendritic morphology and spatial connectivity in mouse visual cortex. *Current Biology*, 32(8):1743–1753, 2022.

[50] Hugh R Wilson and Jack D Cowan. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical journal*, 12(1):1–24, 1972.

[51] Molis Yunzab, Veronica Choi, Hamish Meffin, Shaun L Cloherty, Nicholas J Priebe, and Michael R Ibbotson. Synaptic basis for contrast-dependent shifts in functional identity in mouse V1. *Eneuro*, 6(2), 2019.

[52] Molis Yunzab, Shaun L Cloherty, and Michael R Ibbotson. Comparison of contrast-dependent phase sensitivity in primary visual cortex of mouse, cat and macaque. *NeuroReport*, 30(14):960, 2019.