Learning spatio-temporal V1 cells from diverse LGN inputs

Author Names and Affiliations: Marko A. Ruslim[1], Anthony N. Burkitt[1,2], Yanbo Lian[1*]

[1] Department of Biomedical Engineering, The University of Melbourne, Melbourne, VIC, 3010, Australia

[2] Greame Clark Institute of Biomedical Engineering, The University of Melbourne, Parkville, Victoria, 3010, Australia

***Correspondence:** yanbo.lian@unimelb.edu.au

Acknowledgements: This work received funding from the University of Melbourne via 2023 Early Career Researcher Grant, and the Australian Government via grant AUSMURIB000001 associated with ONR MURI grant N00014-19-1-2571.

Conflict of interest statement: The authors declare no competing financial interests.

Abstract

Since the Nobel-winning discovery of simple cells and complex cells in cat's primary visual cortex (V1) by Hubel and Wiesel, many experimental studies of the visual system of the brain have been conducted. Experimental data of V1 cells from animal recordings show spatio-temporal properties, namely that they display both spatial and temporal response properties. For spatial properties, each V1 cell responds to a specific feature (such as bars, blobs, etc.) in the visual space, which is called the receptive field of this cell. The receptive fields of different V1 cells are typically different in size, orientation, spatial frequencies, etc. Furthermore, V1 cells also display temporal properties, namely that the receptive fields of V1 cells can change over time, and there is a great variety of ways in which they change over time. For example, the population of V1 cells show a great diversity from monophasic response to biphasic response, and some V1 cells are selective to a preferred direction. However, given many computational learning models that explain how spatial properties of V1 cells can be learnt, how temporal properties emerge is still not well understood. In this paper, we use a simple learning model based on sparse coding to show that spatio-temporal V1 cells, such as biphasic and direction selective cell, can emerge via synaptic plasticity when diverse spatio-temporal LGN cells are used as upstream input to V1 cells. This work suggests that temporal along with spatial properties of V1 cells may simply come from a learning process that aims to encode upstream input with spatio-temporal properties, which greatly enhances our understanding of V1 cells.

1 Introduction

The properties of simple cells and complex cells in the primary visual cortex (V1) were described by Hubel and Wiesel in their landmark papers (Hubel and Wiesel, 1959, 1962, 1968). The properties of these cells can be characterized by their receptive fields (RFs). The visual RF of a neuron refers to the attributes of a visual stimuli that drive a neural response. For simple cells, their spatial RFs are spatially oriented, with alternating elongated ON and OFF subregions. Similarly, complex cells respond primarily to oriented edges and gratings, but unlike simple cells, they can display spatial phase invariance.

While the RF is often described solely in terms of spatial coordinates, the RF inherently involves both spatial and temporal dimensions. A RF is spatio-temporally separable if its three-dimensional structure, can

be written as the product of two independent functions: a spatial profile and a temporal profile. Otherwise, if a cell's RF cannot be broken down into independent spatial and temporal components, its RF is space-time inseparable. Although numerous types of spatio-temporal inseparability are possible, simple cells with inseparable RFs show a highly characteristic profile in which their RF shifts spatially over time. In experimental studies, these cells exhibit direction-selectivity (DeAngelis et al., 1995; De Valois et al., 2000). Non-directional V1 cells tend to have separable RFs that are temporally monophasic or biphasic.

Simple cells in V1 receive upstream input from the lateral geniculate nucleus (LGN). The spatial RFs of LGN neurons are unoriented and approximately circular with a center-surround structure. However, two different types of LGN cells display different temporal properties: Magnocellular LGN cells (M-cells) exhibit a biphasic temporal RF (i.e., their spatial RF will switch polarities over time), whereas parvocellular LGN cells (P-cells) are largely monophasic (i.e., their spatial RF will not change polarities over time). Due to these temporal properties, M-cells are sensitive to high temporal frequencies, tend to have transient responses and thus are able to detect quick changes in visual stimuli, while P-cells are sensitive to low temporal frequencies and tend to have more sustained responses (Nassi and Callaway, 2009).

Previous investigators have suggested that both M-cells and P-cells contribute to direction selectivity in V1 cells (De Valois et al., 2000; Saul et al., 2005). There are other various mechanisms that have been proposed to explain the direction selectivity of V1 simple cells. One mechanism is that orientation-selective inhibitory neurons that have spatially offset receptive fields provide delayed input to excitatory neurons which consequently leads to excitatory and inhibitory inputs that destructively interference in the non-preferred stimulus direction but not in the preferred stimulus direction (Freeman, 2021). Consequently, this asymmetry leads to direction selectivity. Another mechanism that could lead to direction selectivity is from the dynamic differences between ON and OFF LGN cells, namely differences in time delay and temporal kernel shape (Chariker et al., 2021). ON and OFF responses to sinusoidal gratings constructively or destructively interfere depending on stimulus direction. These models give the theory of the mechanism of direction selectivity, but how the direction selectivity emerges from a learning process is still largely unknown.

Computational learning models have shown how the spatial properties of V1 cells can be learnt from static natural images using the principle of sparse coding (Olshausen and Field, 1996). There are biologically

plausible models that can produce the diverse spatial properties of V1 cells (Zylberberg et al., 2011; Lian et al., 2019). These models however, fail to produce the temporal properties that V1 neurons possess. Spatio-temporal properties of place cells in the hippocampus is able to be learnt using a model based on sparse coding (Lian and Burkitt, 2022). In this paper, we also use a model based on sparse coding to show how the spatio-temporal properties of simple cells in V1 can be learnt. Namely, these diverse spatio-temporal properties include the spatial properties of Gabor-like and blob-like receptive fields, temporal properties of monophasic or biphasic profiles, and the space-time inseparable property of some V1 neurons that possess direction selectivity.

2 Methods

2.1 Sparse coding

Our learning model of spatio-temporal V1 cells is based on the principle of sparse coding. The original sparse coding model was introduced by Olshausen and Field (1996) who proposed that the sensory input to V1 simple cells can be represented by a small number of neurons. This model can be achieved by solving an optimization problem that reconstructs the input by a linear representation while also encouraging sparsity among model responses (Olshausen and Field, 1996, 1997).

Here, we implement the model based on the locally competitive algorithm that solves sparse coding efficiently (Rozell et al., 2008). The dynamics and learning rule of the model can be described by

$$\begin{aligned} \tau \dot{\mathbf{u}} &= -\mathbf{u} + \mathbf{A}^T \mathbf{r} + \mathbf{s} \\ \mathbf{r} &= \mathbf{x} - \mathbf{A}\mathbf{s} \\ s &= \max(\mathbf{u} - \lambda,\ 0), \end{aligned} \tag{1}$$

and

$$\Delta \mathbf{A} \propto \left\langle \mathbf{r}\mathbf{s}^T \right\rangle, \tag{2}$$

where τ is the time constant, $\mathbf{u}$ represents the membrane potentials of the model cells, $\mathbf{s}$ represents the non-negative firing response of the model cells, $\mathbf{r} = \mathbf{x}–\mathbf{A}\mathbf{s}$ is the reconstruction error, and $\langle\cdot\rangle$ denotes an

ensemble average.

2.2 Input

The input dataset used to train the model is a 2-minute video (1080×1920 pixels; frame rate 24 Hz) of natural scenes, which was taken in a forest walk (https://youtu.be/K-Vr2bSMU7o, with permission from the owner). For this study, this video is then trimmed to only keep the central 800×800 pixel region, similar to our previous study that learnt V1 complex cell properties (Lian et al., 2021).

2.2.1 Spatial filtering of LGN cells

Before the video is used for training, a pre-processing step is implemented to mimic the early visual processing done by LGN cells whose RFs have both spatial and temporal properties. The spatial property of LGN RFs is characterised by a divisively normalized difference-of-Gaussian filters (Tadmor and Tolhurst, 2000; Ratliff et al., 2010):

$$I(x, y) = \frac{I_0(x, y) - I_1(x, y)}{I_{\mathrm{d}}(x, y)}, \tag{3}$$

where $I(x, y)$ is the spatially filtered pixel intensity at point (x, y), and I_0, I_1 and I_d are the outputs from unit-normalized Gaussian filters of center (g_0), surround (g_1) and divisive normalization g_d, respectively. The standard deviations of the center, surround and divisive normalization filters are chosen to be 1, 1.5 and 1.5 pixels, respectively. Eq. 3 is applied to each video frame to perform spatial pre-processing. Additionally, temporal filtering is applied across video frames to mimic the temporal properties of LGN cells, as described below.

2.2.2 Temporal filtering of LGN cells

The profile of the LGN temporal filter is either monophasic or biphasic. The LGN neurons with monophasic temporal profiles exhibit low-pass temporal frequency tuning, whereas the neurons with biphasic profiles have bandpass tuning (DeAngelis et al., 1995). We use axiomatically determined temporal models of LGN cells that satisfy time-causality and show qualitative agreement with biological receptive fields (Lindeberg,

2021). The monophasic temporal filter is created by convolution of truncated exponential functions $h_{\text{exp}}(t)$:

$$h_{\text{exp}}(t) = \begin{cases} \frac{1}{\tau} \exp\{\frac{-t}{\tau}\}, & t \geq 0 \\ 0, & t < 0. \end{cases} \tag{4}$$

The convolution of $M = 7$ of these truncated exponential creates a temporally smooth monophasic temporal profile, $H_{\text{mono}}(t)$:

$$H_{\text{mono}}(t) = \underbrace{h_{\text{exp}} * h_{\text{exp}} * \cdots * h_{\text{exp}}(t)}_{M}. \tag{5}$$

Temporal derivatives of this temporal model also satisfy the assumptions of a normative theory of visual receptive fields (Lindeberg, 2021). The first temporal partial derivative of this temporal profile is used to define our biphasic temporal profile $H_{\text{bi}}(t)$:

$$H_{\text{bi}}(t) = \frac{\partial}{\partial t} H_{\text{mono}}(t). \tag{6}$$

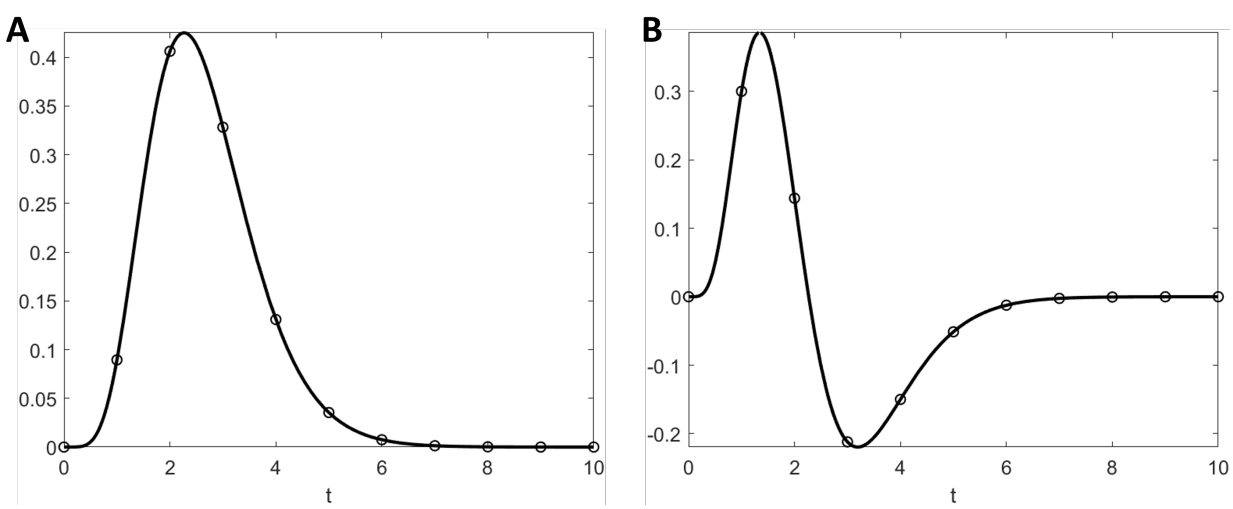

Figure 1: **LGN temporal filters**. (A) Monophasic profile. (B) Biphasic profile.

The monophasic and biphasic temporal filters of LGN used in this work are shown in Fig. 1. To apply these temporal filters to the video stimulus, τ was chosen to be 1, and the filters are applied to the current and the 10 preceding video frames, evaluated at $t = 0, 1, 2, ..., 10$.

In order to investigate how the temporal properties of LGN cells affects the temporal properties of learnt V1 cells, we define different scenarios where diverse LGN temporal filters are used to temporally filter the input that has been spatially filtered using Eq. 3.

Scenario 1a Modelled V1 cells only receive input from monophasic LGN cells. The LGN-V1 connection undergoes learning while the natural scenes video is presented. This scenario is designed to show that the spatial properties of V1 can be learnt, and also that their temporal properties can be inherited from LGN cells.

Scenario 1b Modelled V1 cells only receive input from biphasic LGN cells. Similar to Scenario 1a, the LGN-V1 connection undergoes learning.

Scenario 2 Modelled V1 cells receive input from both monophasic and biphasic LGN cells. The spatial and temporal properties of these LGN cells have been described previously. Similar to Scenario 1a & 1b, the LGN-V1 connection undergoes learning. This scenario is used to investigate whether diverse spatio-temporal properties of V1 cells can be learnt.

2.3 Training

Random video patches that have been spatially and temporally pre-processed are the inputs to our model. Each frame of the video patch is of size 16×16 which results in 256 ON and 256 OFF LGN cells in the first layer. The number of model simple cells in the second layer is 256. Neural dynamics described by Eq. 1-1 were computed using the first-order Euler method. There are 50 integration steps with a time step of 1 ms.

The Hebbian learning rule in Eq. 2 is used to update the synaptic strengths. After each learning iteration, the columns of the weight matrix $\mathbf{A}$ are normalized to have an L2 norm of 1. Elements of this weight matrix $\mathbf{A}$ are initialized randomly from a Gaussian distribution and subsequently normalized before learning starts. For each learning iteration, 100 random 16×16 pixel regions of a video frame were presented.

The model was simulated in MATLAB (R2023b) and run on the high performance computing platform Spartan (Lafayette et al., 2016).

2.4 Recovering spatio-temporal receptive fields of learnt V1 cells

The method of spike-triggered average is used to estimate the receptive fields of V1 cells (Chichilnisky, 2001; Schwartz et al., 2006). We used white noise stimuli, $\mathbf{n}$, of dimension $16 \times 16 \times K$ where K is the number of temporal frames to store. This white noise was spatially and temporally pre-processed then presented to the model, and the firing rate of a simple cell is recorded. Its spatio-temporal receptive field, $\mathbf{F}$, can be estimated through the weighted average,

$$\mathbf{F} = \frac{\sum_{t=1}^{T} \mathbf{s}_t \mathbf{n}_t}{\sum_{t=1}^{T} \mathbf{s}_t}. \tag{7}$$

A way to visualise spatio-temporal receptive fields is by constructing an x-t plot (DeAngelis et al., 1995; De Valois et al., 2000). This plot summarises how one spatial dimension varies with time. The spatial dimension chosen for these plots is along the axis perpendicular to the preferred orientation of the cell. Intensities across the perpendicular spatial dimension were averaged.

2.5 Biphasic index

The biphasic index is a metric that quantifies the biphasic temporal characteristic of non-directional V1 cells (De Valois et al., 2000). It is defined as the ratio of the amplitude of the second temporal peak, P_2 to the amplitude of the first temporal peak, P_1 where these two peaks have reversed polarity:

$$\mathrm{BI} = \frac{P_2}{P_1}. \tag{8}$$

A low biphasic index close to zero suggests a temporally monophasic cell, and a high biphasic index indicates a very biphasic cell.

2.6 Direction tuning curve and directional index

The direction tuning curve is a response vs. direction plot that shows how the response of a neuron changes with different moving directions. To do this, we first search for the optimal spatial frequency of drifting sinusoidal grating that gives the maximal response of a model V1 cell. Then we present drifting sinusoidal grating with the optimal spatial frequency, at constant speed but of different moving directions (0 to 360°

with a step size of 15°).

To quantify a neuron's degree of directional selectivity, we used the directional index that was used in experimental studies (De Valois et al., 2000):

$$\mathrm{DI} = \left(1 - \frac{\mathrm{NP}}{\mathrm{P}}\right) \times 100\%, \tag{9}$$

where P is the response in the preferred direction (i.e., the peak of the direction tuning curve), and NP is the response in the non-preferred direction that is 180° apart. As in De Valois et al. (2000), we defined direction-selective cells to be those with a directional index greater than 70, and non-directional cells to be those with a directional index less than 70.

3 Results

3.1 Learnt V1 cells have spatial RFs and monophasic temporal property if LGN cells incorporate monophasic temporal filters

After learning in Scenario 1a, the synaptic weights between LGN input and model V1 cells shows good qualitative agreement with biological V1 receptive fields. We observe similar spatial structures such as oriented Gabor-like filters and non-oriented blobs. Similar to our previous study (Lian et al., 2019), we use the synaptic field to visualise the overall synaptic weights from ON and OFF LGN cells to V1, which is defined as

$$\mathbf{S}_{\mathrm{f}} = \mathbf{A}_{\mathrm{ON}} - \mathbf{A}_{\mathrm{OFF}}, \tag{10}$$

where $\mathbf{A}_{\mathrm{ON}}$ and $\mathbf{A}_{\mathrm{OFF}}$ represent the ON LGN to V1 connection and OFF LGN to V1 connection, respectively. As shown in Fig. 2A, the synaptic fields of V1 cells display simple-cell like features such as 2D Gabor filters and blobs.

However, the synaptic fields only show a static representation of a cell's receptive field while ignoring the temporal characteristics. Complete spatio-temporal receptive field profiles for representative model V1 cells are shown in Fig. 2B, where each row is a sequence of snapshots of a neuron's spatial receptive field at different time steps. These neurons show a clear monophasic temporal profile, where the response builds up

and then dies off over time. The peak appears at $t = 2$, which is inline with the LGN monophasic temporal profile shown in Fig. 1A.

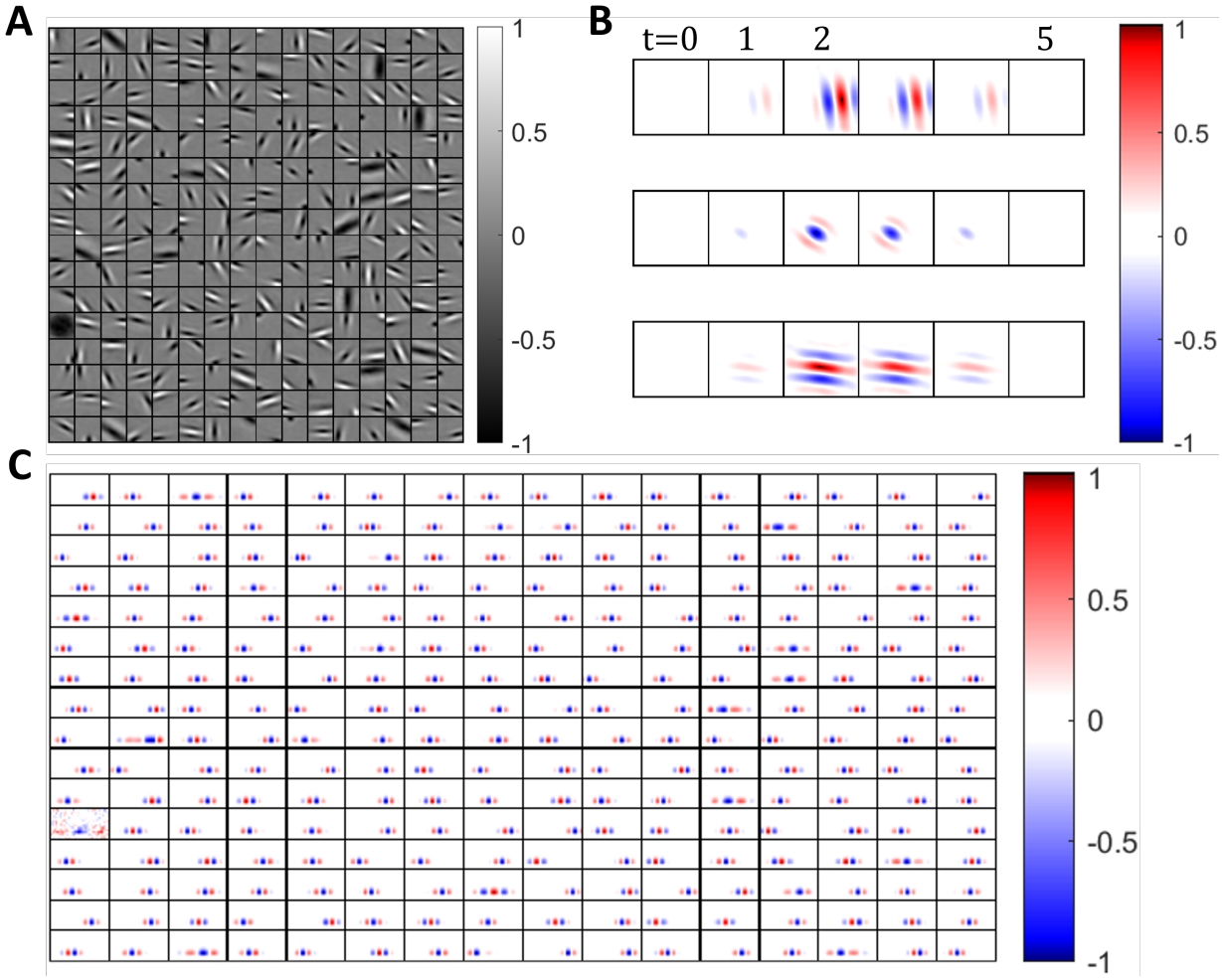

Figure 2: **Receptive fields of V1 neurons only receiving monophasic LGN inputs.** (A) Synaptic fields of all 256 cells. Values in each block are normalized to the range [-1 1]. (B) Dynamics of three example receptive field at snapshots of video frame number or time. Values in each row are normalized to the range [-1 1]. (C) x-t plot of the spatio-temporal RF of all 256 cells. Values in each block are normalized to the range [-1 1].

The spatio-temporal receptive fields of the whole population can be visualised by x-t plots (DeAngelis et al., 1995). The horizontal axis represents the spatial dimension perpendicular to that neuron's preferred orientation, and the vertical axis represents time. All these x-t plots have a characteristic monophasic temporal profile. This is expected because the model V1 neurons only receive LGN input from monophasic

cells. Additionally, the x-t plots show that these learnt simple cells have approximately space-time separable receptive fields.

3.2 Learnt V1 cells has spatial RFs and biphasic temporal property if LGN cells incorporate biphasic temporal filters

For Scenario 1b, where the input LGN neurons have a biphasic temporal profile, the V1 neurons are also able to learn the diverse spatial features of receptive fields, namely oriented Gabor-like filters and unoriented blobs. The synaptic fields in Fig 3A shows the spatial representation of the model cells' receptive field. Note that 195 out of 256 neurons were selected because not all the neuron's have learnt a receptive field. This could be because biphasic neurons exhibit transient responses and not sustained responses (De Valois et al., 2000). As a result, there are fewer instances for model V1 cells to learn stimulus features.

The time course of three example receptive fields are shown in Fig 3B. Unlike the monophasic responses in Scenario 1a which rise then fall, the responses in this simulation show a biphasic profile that reverses in polarity. For example, the top cell initially has a central dark OFF subregion surrounded by two bright ON subregions above and below. After $t > 2$, each subregion reverses polarity, so that the central region is now bright ON. This follows from the LGN's biphasic temporal kernel (Fig. 1B).

The spatio-temporal receptive field shown in Fig. 3C also highlights the biphasic property. Along the vertical axis, which represents time, the polarity reverses. All the model cells display a biphasic profile because the input LGN cells are purely biphasic as well.

3.3 Learnt V1 cells has spatial RFs and diverse temporal properties if LGN cells incorporate both monophasic and biphasic temporal filters

In Scenario 2, there are two LGN input populations. One population contains LGN neurons with a monophasic temporal profile, and the other population contains LGN neurons with a biphasic temporal profile. As a result, we produced two synaptic fields per model V1 cell (Fig. 4). The spatial structure of the two sets of synaptic fields look similar in terms of receptive field location and orientation. However, for some of the model V1 neurons, the synaptic fields from monophasic LGN inputs and biphasic LGN inputs have their

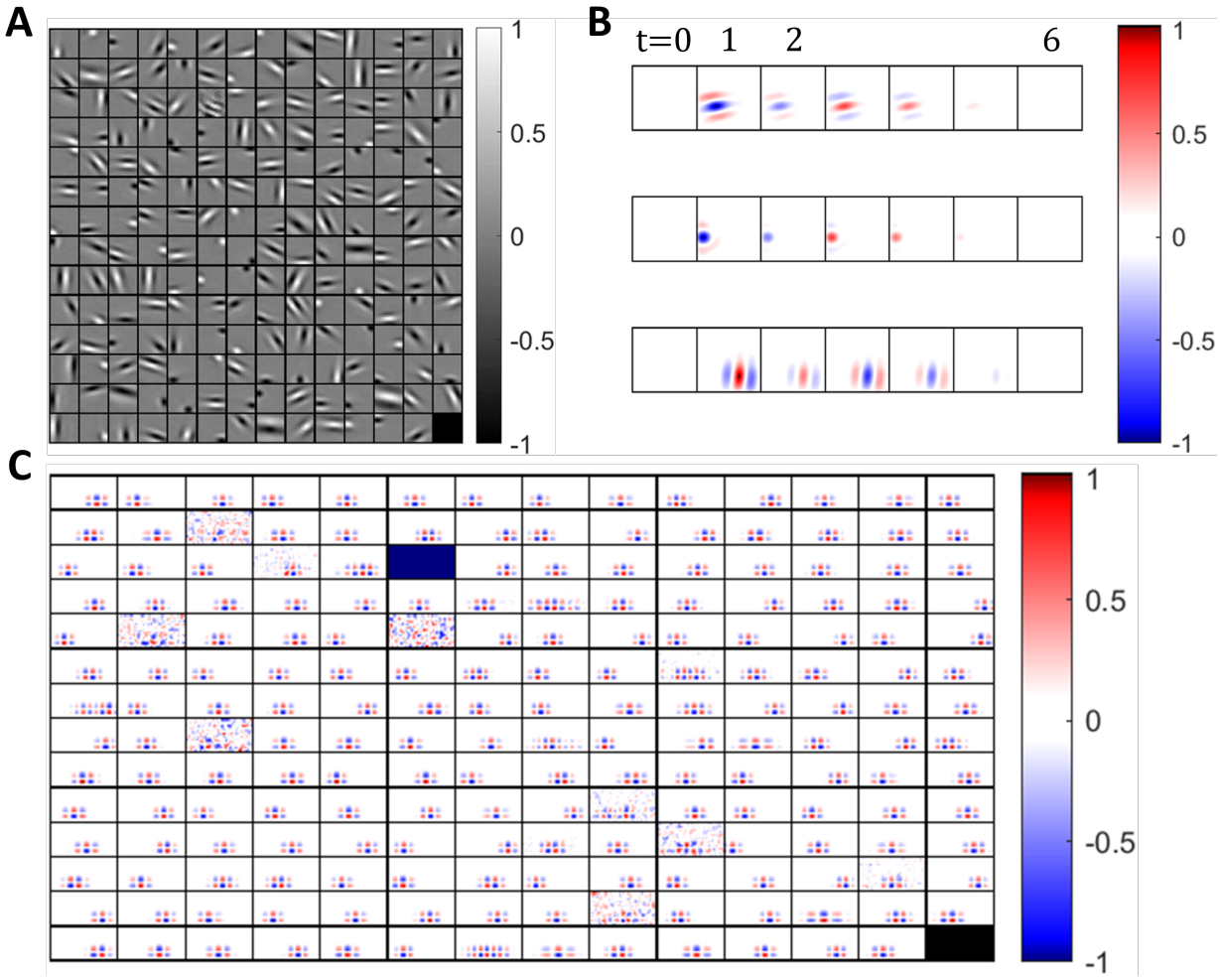

Figure 3: **Receptive fields of V1 neurons only receiving biphasic LGN inputs.** (A) Synaptic fields of 195 selected model V1 cells. (B) Dynamics of three example receptive field at snapshots of video frame number or time. (C) x-t plot of the spatio-temporal RF of the 195 selected cells.

spatial phase offset by some amount.

In the x-t plot (Fig. 4), a diverse range of temporal properties can be seen. There are some neurons that show a characteristic monophasic temporal profile and others that show a biphasic temporal filters. These neurons have separable spatial and temporal profiles. There are also neurons with receptive fields that are inseparable. In particular, their receptive field shifts in one direction over time. This can be seen by the slant in some of the x-t plots.

Direction tuning curves were generated for each model cell by presenting sinusoidal gratings at each neu-

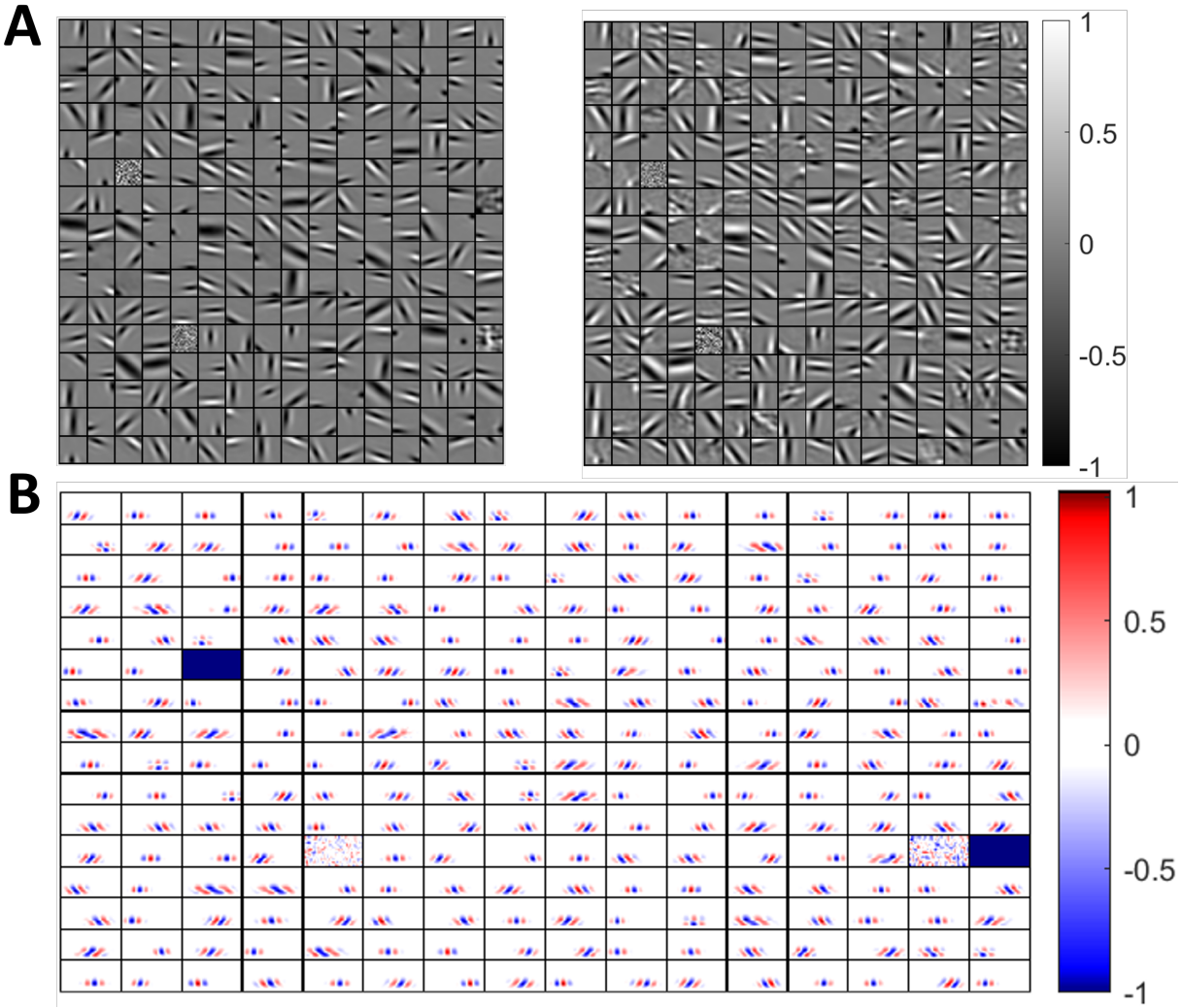

Figure 4: **Receptive fields of V1 neurons receiving monophasic and biphasic LGN inputs.** (A) Synaptic fields of all 256 model V1 cells from monophasic LGN inputs (left) and biphasic LGN inputs (right). (B) x-t plot of the spatio-temporal RF of all 256 cells.

ron's preferred spatial frequency and in all directions. Four representative direction tuning curves are shown in Fig. 5A. The curve in blue shows an orientation-selective cell, where the neuron primarily responds in one orientation or two directions that are separated by 180° with no preference for a particular direction. The curves in yellow and purple show a stronger selectivity for direction, because the peak for the preferred direction is greater than the peak for the non-preferred direction.

Fig. 5B shows the time course of the receptive field of the neurons in Fig. 5. The top row corresponds to one

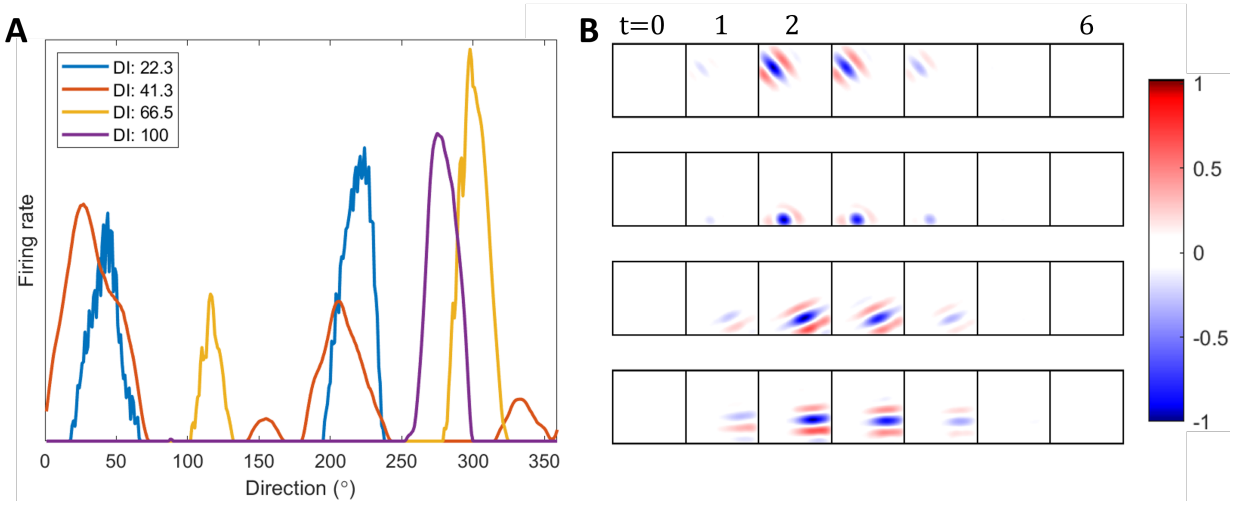

Figure 5: **Direction tuning curves and receptive field dynamics.** (A) Direction tuning curves of four selected model V1 neurons. (B) Dynamics of the four selected receptive field at snapshots of video frame number or time.

orientation-selective cell, with the blue direction tuning curve and a directional index of 22.3. This neuron prefers stimuli oriented at $45°$ which is visible by its receptive field shape. The bottom cell corresponds to one direction-selective cell, with the purple direction tuning curve and a directional index of 100. This neuron prefers stimuli moving in one direction only and does not respond to stimuli moving in the opposite direction. From our simulations with both monophasic and biphasic input LGN cells, we show model V1 cells with diverse spatial and temporal properties. In particular, there are Gabor-like and blob-like spatial features, and there are neurons that are temporally monophasic, biphasic and direction-selective. From our simulations using these LGN filters, there are 188 direction-selective cells and 65 non-directional cells.

Fig. 6 shows the histograms of directional index of experimental V1 cells and model V1 cells. As seen in Fig. 6A, there is a wide distribution of directional index in experimental data (De Valois et al., 2000). However, most model V1 cells are direction selective with high directional indices, though there are also model V1 cells with diverse directional indices (Fig. 6B).

The biphasic index was calculated for each model non-directional cell (Fig. 7). Similar to experimental data, we observe that the majority of non-directional cells are monophasic that have a very low biphasic index. In the data from De Valois et al. (2000), they found a diverse range of biphasic indices as well as a distinct small population of biphasic cells. However, we do not see this in our simulations. This could be due to the

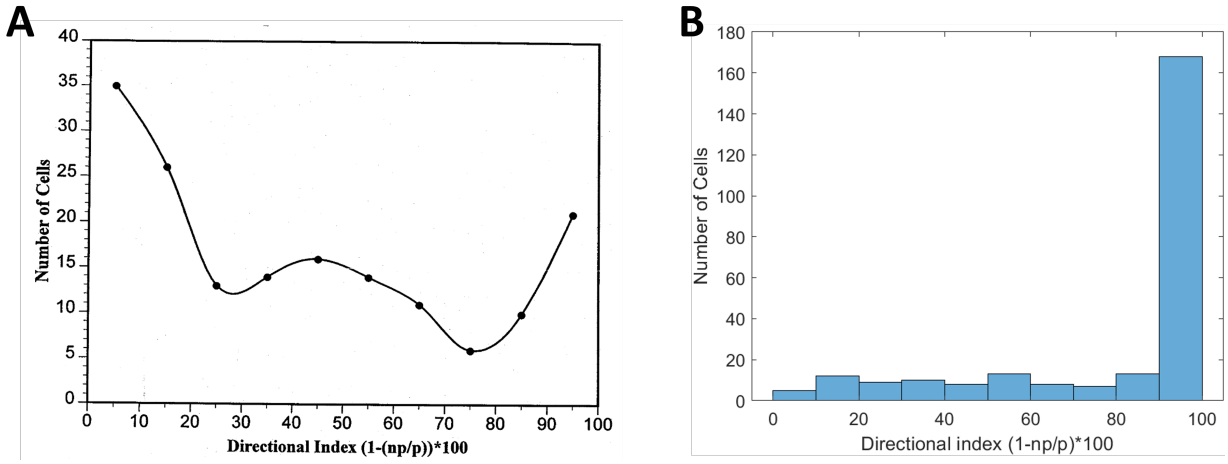

Figure 6: **Distribution of directional index of V1 population.** (A) Experimental data in macaque monkey (De Valois et al., 2000). (B) Model data after learning.

fact that the simulation in Scenario 2 only contains two types of LGN filters. LGN cells, however, display diverse temporal response properties, including lagged and non-lagged (DeAngelis et al., 1995), different time to peak responses and also different biphasic indices (De Valois et al., 2000).

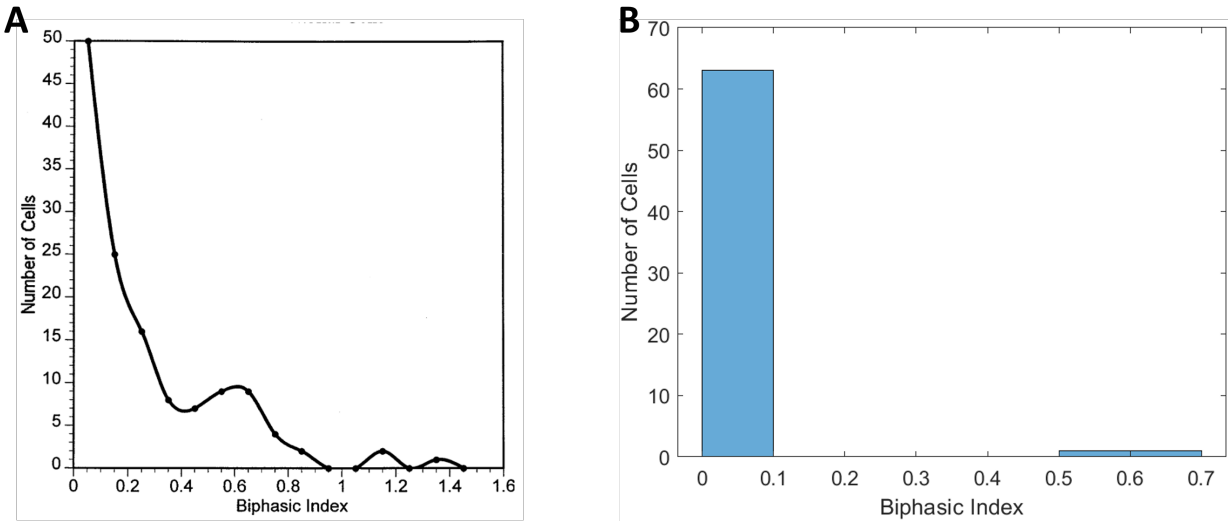

Figure 7: **Distribution of biphasic index from the population of non-directional V1 cells.** (A) Experimental data in macaque monkey (De Valois et al., 2000). (B) Model data after learning.

3.4 Temporal properties of learnt V1 cells depend on the exact profile of LGN temporal filters

The properties of the model V1 cells appear to depend on the properties of the input LGN cells. Slightly different temporal profiles were applied to the LGN cells (Fig. 8A). The values were chosen such that the

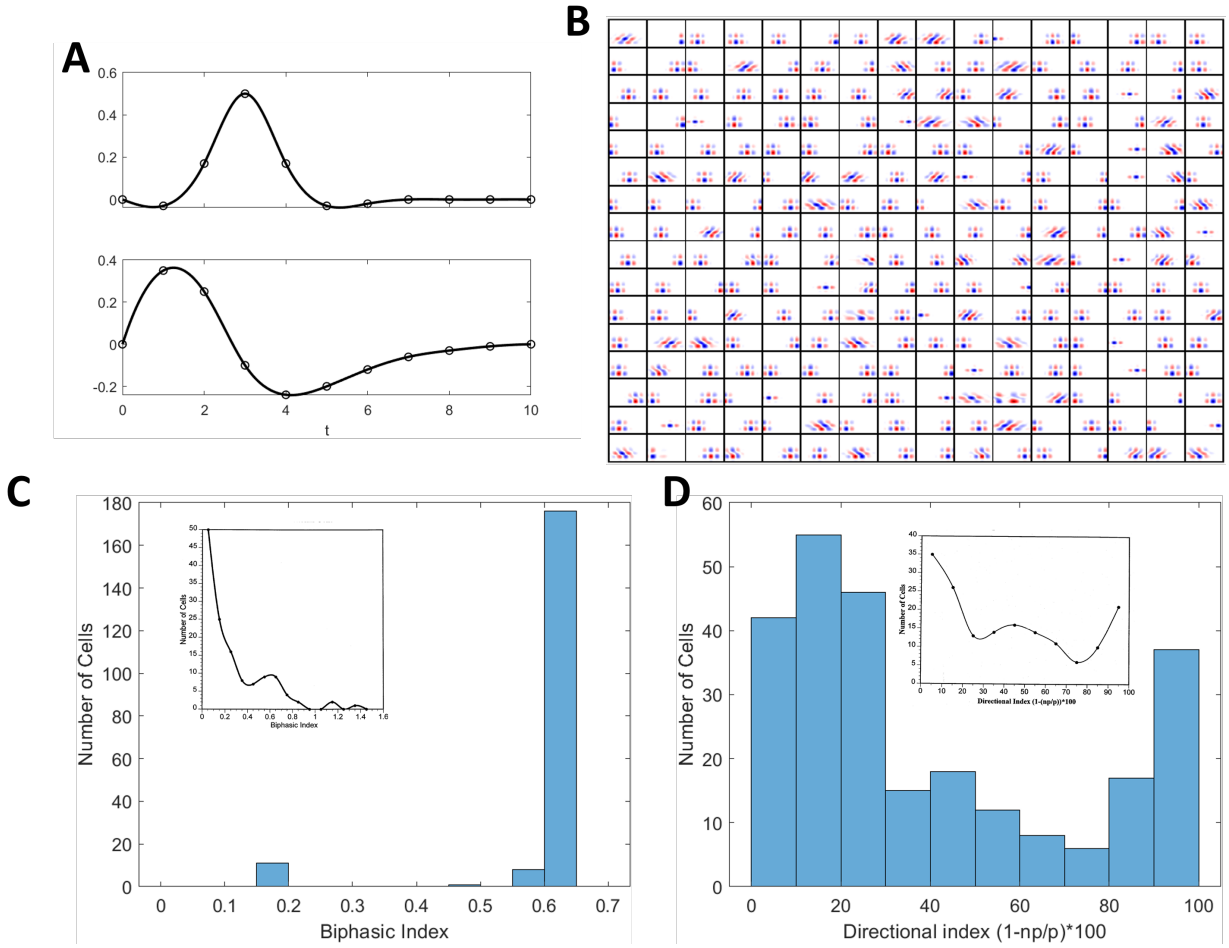

Figure 8: **Different LGN temporal filters lead to different V1 population statistics.** (A) Chosen LGN temporal filters for monophasic (top) and biphasic (bottom) cells. (B) x-t plot of learnt model V1 cells. (C) Distribution of biphasic index of model non-directional cells compared with experimental data (inset) from De Valois et al. (2000). (D) Distribution of directional index of model V1 cells compared with experimental data (inset) from De Valois et al. (2000).

filters had the characteristic shape of a monophasic and biphasic profile and also so that the area under the curves are roughly equal. The resulting receptive fields that are learnt by the model V1 cells are shown as an x-t plot in Fig. 8B. Diverse spatial and temporal features can be seen, such as monophasic, biphasic and

direction-selective properties. However, unlike scenario 2, in this simulation, the population statistics have changed. For instance, instead of a large monophasic population, there is a large biphasic population out of the non-directional cells (Fig. 8C). Additionally, the number of model cells from this simulation that were directionally selective was fewer than in scenario 2. The distribution of directional index follows a similar trend to the directional index measured experimentally (Fig. 8D). By modelling the diverse properties of LGN temporally filters, it may be possible to more closely match experimental distributions of biphasic index and directional index.

4 Discussion

4.1 Summary

In this paper, we use the sparse coding model to demonstrate that the temporal properties of V1 cells can be learnt simultaneously with spatial properties given that upstream LGN cells have both spatial and temporal properties. Though V1 cells display diverse temporal properties such as biphasic response or direction selectivity, previous learning models of V1 cells primarily focus on the spatial properties. Our work shows that the principle of sparse coding can be used to learn spatio-temporal V1 cells from spatio-temporal LGN input, which suggests that both spatial and temporal properties of V1 cell can emerge as a result of cortical processing of upstream input. Combined with our previous work that demonstrates that spatio-temporal hippocampal place cells can be learnt from upstream spatio-temporal grid cells in the medial entorhinal cortex (Lian and Burkitt, 2022), it suggests that the processing of spatio-temporal upstream input may also account for the development of cell properties in other brain areas.

4.2 Discrepancies between model and experimental data

After learning, model V1 cells can display Gabor-like spatial receptive fields as well as diverse temporal properties such as monophasic response, biphasic response and directional selectivity. However, there are still discrepancies between model and experimental data (e.g., see Fig. 7 & 8). We suspect the differences may come from a more diverse LGN temporal properties in the real brain neural circuits. In this work, we incorporate both monophasic and biphasic temporal filters of LGN cells, but LGN temporal filters with the

same type have exactly the same profile, which may lead to less diverse temporal input to V1. Therefore, we infer that the model should learn more diverse temporal properties of V1 cells given heterogeneous LGN cells with a variety of temporal profiles.

4.3 Implementing the model in biologically plausible neural circuits

The sparse coding model used in this paper is implemented using locally competitive algorithm (Rozell et al., 2008), which is also used in other works (Zhu and Rozell, 2013; Lian and Burkitt, 2021, 2022; Lian et al., 2023). Though some components of the model are not biologically plausible, the essence of this paper is to demonstrate that the principle of sparse coding can be used to account for the development of spatio-temporal properties V1 cells. Moreover, designing a biologically plausible neural circuits of implementing the principle of sparse coding is an interesting and important research question by itself, which is also our ongoing research.

References

Chariker L, Shapley R, Hawken M, Young LS (2021) A theory of direction selectivity for macaque primary visual cortex. *Proc. Natl. Acad. Sci.* 118:e2105062118.

Chichilnisky EJ (2001) A simple white noise analysis of neuronal light responses. *Network (Bristol, England)* 12:199–213.

De Valois RL, Cottaris NP, Mahon LE, Elfar SD, Wilson JA (2000) Spatial and temporal receptive fields of geniculate and cortical cells and directional selectivity. *Vision Res.* 40:3685–3702.

DeAngelis GC, Ohzawa I, Freeman RD (1995) Receptive-field dynamics in the central visual pathways. *Trends Neurosci.* 18:451–458.

Freeman AW (2021) A model for the origin of motion direction selectivity in visual cortex. *J. Neurosci.* 41:89–102.

Hubel DH, Wiesel TN (1959) Receptive fields of single neurones in the cat's striate cortex. *J. Physiol.* 148:574–591.

Hubel DH, Wiesel TN (1962) Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.* 160:106–154.

Hubel DH, Wiesel TN (1968) Receptive fields and functional architecture of monkey striate cortex. *J. Physiol.* 195:215–243.

Lafayette L, Sauter G, Vu L, Meade B (2016) Spartan performance and flexibility: An hpc-cloud chimera. *OpenStack Summit, Barcelona* 27.

Lian Y, Almasi A, Grayden DB, Kameneva T, Burkitt AN, Meffin H (2021) Learning receptive field properties of complex cells in V1. *PLoS Comput. Biol.* 17:e1007957.

Lian Y, Burkitt AN (2021) Learning an efficient hippocampal place map from entorhinal inputs using non-negative sparse coding. *eNeuro* 8:1–19.

Lian Y, Burkitt AN (2022) Learning spatiotemporal properties of hippocampal place cells. *eNeuro* 9.

Lian Y, Grayden DB, Kameneva T, Meffin H, Burkitt AN (2019) Toward a biologically plausible model of LGN-V1 pathways based on efficient coding. *Front. Neural Circuits* 13:13.

Lian Y, Williams S, Alexander AS, Hasselmo ME, Burkitt AN (2023) Learning the vector coding of egocentric boundary cells from visual data. *J. Neurosci.* .

Lindeberg T (2021) Normative theory of visual receptive fields. *Heliyon* 7.

Nassi JJ, Callaway EM (2009) Parallel processing strategies of the primate visual system. *Nat. Rev. Neurosci.* 10:360–372.

Olshausen BA, Field DJ (1996) Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature* 381:607–609.

Olshausen BA, Field DJ (1997) Sparse coding with an overcomplete basis set: A strategy employed by V1? *Vision Res.* 37:3311–3325.

Ratliff CP, Borghuis BG, Kao YH, Sterling P, Balasubramanian V (2010) Retina is structured to process an excess of darkness in natural scenes. *Proc. Natl. Acad. Sci.* 107:17368–17373.

Rozell CJ, Johnson DH, Baraniuk RG, Olshausen BA (2008) Sparse coding via thresholding and local competition in neural circuits. *Neural Comput.* 20:2526–2563.

Saul AB, Carras PL, Humphrey AL (2005) Temporal Properties of Inputs to Direction-Selective Neurons in Monkey V1. *J. Neurophysiol.* 94:282–294.

Schwartz O, Pillow JW, Rust NC, Simoncelli EP (2006) Spike-triggered neural characterization. *J. Vis.* 6:13.

Tadmor Y, Tolhurst DJ (2000) Calculating the contrasts that retinal ganglion cells and LGN neurones encounter in natural scenes. *Vision Res.* 40:3145–3157.

Zhu M, Rozell CJ (2013) Visual nonclassical receptive field effects emerge from sparse coding in a dynamical system. *PLoS Comput. Biol.* 9:e1003191.

Zylberberg J, Murphy J, DeWeese M (2011) A sparse coding model with synaptically local plasticity and spiking neurons can account for the diverse shapes of V1 simple cell receptive fields. *PLoS Comput. Biol.* 7:e1002250.