Learning the Feedback Connections from V1 to LGN via Information Maximization

Anonymous Author(s)
Affiliation
Address
email

Abstract

The lateral geniculate nucleus (LGN) relay cells act as a gateway for transmitting visual information from retina to the primary visual cortex (V1). The activities of thalamic relay cells are modulated by feedback connections emanating from layer 6 of V1. While the receptive field (RF) properties of these early parts of the visual system are relatively well understood, the function, computational role, and details of the feedback network from V1 to LGN are not. Computational models of efficient coding have been successful in deriving RF properties of retinal ganglion and V1 simple cells by optimizing the Shannon information. Further, previous experimental results have suggested that the feedback increases the Shannon information. Motivated by this earlier work, we try to understand the function of the feedback as optimizing the feedforward information to cortex. We build a model that learns feedback weights by maximizing the feedforward Shannon information on naturalistic stimuli. Our model predicts the strength and sign of feedback from a V1 cell to all ON- and OFF-center LGN relay cells that are within or surrounding the V1 cell RF. We find a highly specific pattern of influence on ON and OFF-center LGN overlapping the V1 RF depending on whether they overlapped the ON or OFF zone of the V1 RF. In addition, we find general inhibitory feedback in the further surround, which sharpens the RFs and increases surround suppression in LGN relay cells. This is consistent with results of recent experiments exploring the impact of feedback on stimuli integration.

1 Introduction

The lateral geniculate nucleus (LGN) acts as a relay of information from the retina to the primary visual cortex (V1). This information flow is modulated by top-down feedback from layer 6 of V1 to LGN. While the input from the retina is the driver input for the LGN relay cells, the feedback is the modulating input for these cells [29].There is a large body of experimental work on the functionality of the feedback pathway from V1 to LGN. It has been observed that this feedback can play an important role in burst vs tonic firing of relay cells [28], increase the timing precision [1, 16], and affect the spatial specificity of LGN relay cells [12, 19].

Anatomical evidence indicates that the corticogeniculate feedback monosynaptically excites an LGN relay cell and disynaptically inhibits it through the thalamic reticular nucleus or the local inhibitory neuron in LGN [10]. Thus, the feedback can have both signs, it can be excitatory or inhibitory. Moreover, the experiments conducted on primates suggest that this feedback is pathway-specific, with parvocellular, magnocellular, and koniocellular LGN cells each receiving input from distinct classes of layer 6 cell [8]. In dual recordings of V1 and LGN, Tsumoto et al. observed mostly excitatory feedback if the V1 and LGN RFs were within 2.4° and inhibitory impact beyond that up to 3.1° [30].

Submitted to Information-Theoretic Principles in Cognitive Systems Workshop at the 36th Conference on Neural Information Processing Systems (NeurIPS 2022). Do not distribute.

These types of experimental findings have inspired the modeling of the feedback as an excitatory-center inhibitory-surround network [14]. A model proposed by Mobarhan et al., in which the feedback center and surround effects were fitted to experimental data on spatial integration in the dorsal LGN [24] also showed an excitatory-center inhibitory-surround. Dual recording experiments conducted by Want et al. [32, 33] suggest an even more specific pattern of connections. Specifically, in [32], the authors observe that LGN cells were significantly more likely to be impacted by feedback, either through facilitation or suppression, when their location was displaced in parallel or orthogonal directions relative to the preferred orientation of the V1 cells. In [33], by measuring the burst-to-tonic ratio in LGN, they established a phase-reversed pattern of connections for LGN cells that overlap the V1 RF.

There is no shortage of existing hypotheses about the functional role of feedback from V1 to LGN, probably starting with [15], who suggested the feedback could modify LGN responses iteratively to help the interpretation of the input. Feedback models were designed and discussed in the context of different visual tasks, such as line perception [27], visual grouping [13, 14], visual attention [6, 9, 22, 25], etc.

Here we asked whether there could be an even simpler explanation of the feedback circuitry, the adaptation of the circuit for maximizing the LGN to V1 information flow. Originating in the work of [4, 2], this idea has gathered a lot of momentum. Experiments conducted by Mcclurkin et al. revealed that cooling V1 in awake monkeys reduced the average stimulus-related information transmitted by LGN neurons [23]. Further, principles of information maximization and efficient coding have been widely successful in understanding the structure of feedforward connections in the early visual system. These principles lead to normative models that learn V1-like receptive fields from natural images [5, 31, 26]. Similarly, the same principles, when combined with realistic noise levels present in the retina, yield models that reproduce the center-surround receptive fields of retinal ganglion cells [20].

We designed a novel type of normative model that employs the same information objective, in a circuitry with feedback. As forward connections, we use the known RF structure of LGN and cortex, the feedback connections are optimized by learning during naturalistic stimulation. The learned feedback weights show reciprocal excitation of LGN cells that are exciting the V1 cells, and inhibition of the LGN cells that inhibit the V1 cell. These weights also show a pattern of suppression for LGN cells that are retinotopically displaced relative to the V1 RF within 4 - 6 times the receptive field size. Testing the impact of this learned feedback on spatial integration in LGN points to an accentuation of surround suppression as a result of feedback which is also consistent with experimental results in [7].

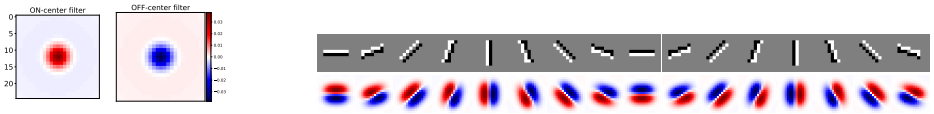
2 Model

We use rate code models of retina, LGN, and V1 to represent the feedforward pathway from retina to V1. Given a two-dimensional gray-scale input image X, the retinal output is calculated as $Y^p_{\text{retina}} = f^p_{\text{retina}}(W^p_{\text{retina}} * X)$ for $p \in \{\text{ON}, \text{OFF}\}$, where $W^{\text{ON}}_{\text{retina}}$ and $W^{\text{OFF}}_{\text{retina}}$ are the ON and OFF-center retinal ganglion cell linear filter weights and $*$ denotes the two-dimensional convolution operator. Each linear filter has an associated rectifying nonlinearity f^p_{retina} which is applied entry-wise to the two-dimensional convolution $W^p_{\text{retina}} * X$. The rectifying nonlinearity is given as

$$f(I) = \max(RI - v, 0), \tag{1}$$

where $R > 0$ and v are nonlinearity parameters. Note that the scaling factor R is required since the linear filter weights are fixed. The ON and OFF-center filters are shown in Figure 1A. The two linear filters are 27 by 27 pixel discrete approximations of a two-dimensional difference-of-Gaussians (DoG) filter given by $\pm\frac{1}{2\pi\sigma_1^2}e^{-(v_1^2+v_2^2)/(2\sigma_1^2)} \mp \frac{k}{2\pi\sigma_2^2}e^{-(v_1^2+v_2^2)/(2\sigma_2^2)}$. We set $\sigma_2/\sigma_1 = 7$ and $k = 0.9$ which are biologically plausible based on measurements in cat retina [11]. The convolution operation has a stride (spacing between neighboring cells) of $2\sigma_1$. This spacing is chosen based on reported measurements of the retina and computational work showing that information from natural images peaks at a receptive field spacing of $\sim 2\sigma_1$ [3].

In our model each LGN relay cell is driven by a single RGC and the LGN output is $Y^p_{\text{LGN}} = f^p_{\text{LGN}}(Y^p_{\text{retina}})$ for $p \in \{\text{ON}, \text{OFF}\}$, where f^p_{LGN} are rectifying nonlinearities of type given in (1). The LGN output then serves as input to V1. To create the V1 output, we form Hubel and Wiesel V1 simple cells [17] by assigning equal positive weights to 8 ON LGN and 8 OFF-center LGN cells along 8



(A) (B)

Figure 1: **Feedforward and feedback models.** A: ON and OFF Difference of Gaussian filters used in the model of retina. The filters parameters are: $(\sigma_1, \sigma_2, k) = (2, 14, 0.9)$. B: The V1 weights and the corresponding receptive fields when the weights are applied to ON and OFF-center LGN cells.

different directions. These 8 filters alongside their phased-reversed version form the V1 filterbank of size 16 (see Figure 1B). The V1 output is given by $Y^k_{V_1} = f^k_{\text{V}_1}(\sum_{p\in\{\text{ON,OFF}\}} W^{k,p}_{V1} * Y^p_{\text{LGN}})$ for $k \in \{0, 1, \ldots, 15\}$, where $f^k_{\text{V}_1}$ is applied entry-wise to the output of the convolution.

2.1 Feedback model

The majority of the deriving axons from LGN relay cells target cells in layer 4C of the cortex, while the feedback to the LGN emanates from layer 6. Hence a complete model of the feedback should include the mechanisms that give rise to layer 6 cells' receptive fields requiring access to many unknown parameters related to the micro-circuitry in a V1 column. To avoid these complications, we model the feedback as direct feedback from the same V1 cells that are being driven by LGN.

We allow a V1 cell to feedback on LGN relay cells directly overlapping the V1 cell RF or within a retinotopic neighborhood of the V1 cell RF. Associated with each one of the 16 V1 filter types, there are two sets of two-dimensional feedback weights which determine the feedback impact on ON and OFF-center LGN cells. These weights are denoted by $W^{k,\text{ON}}_{\text{fb}}$ and $W^{k,\text{OFF}}_{\text{fb}}$ for $k \in \{0, 1, 2, ..., 15\}$. The V1 activation after applying the feedback will be the solution to the following equation:

$$Y^k_{\text{V1}} = f^k_{\text{V}_1}\left(\sum_{p\in\{\text{ON},OFF\}} W^{k,p}_{\text{V}_1} * f^p_{LGN}\left(Y^p_{\text{retina}} + \sum_{l=0}^{15} W^{l,p}_{\text{fb}} * Y^l_{\text{V1}}\right)\right) \text{ for } k \in \{0, 1, ..., 15\}.$$

Figure 3A depicts a block diagram of the model. To numerically estimate Y_{V1}, we unroll this feedback loop in 3 steps, a standard technique used for recurrent neural networks.

3 Learning the feedback weights

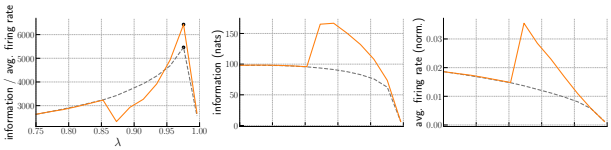
To learn the feedback kernel and nonlinearity parameters in the feedforward model, we use an information objective. Since we are considering a noiseless channel from retina to V1, the mutual information between the input image patches and the V1 output is given by the differential entropy of the V1 output. Differential entropy increases when the firing rates are scaled up. Therefore, our optimization objective, similar to the objective used in [20], has a penalty term that penalizes the average firing rate. We calculate an upper bound on the differential entropy of a 2×2 patch of V1 cells, where each location is covered by 16 types of V1 cells. Let Y_{patch} be the vectorized response vector of these 64 cells. We denote the differential entropy of Y_{patch} by $H(Y_{\text{patch}})$ and note that $H(Y_{\text{patch}}) \leq \frac{1}{2}\log\det\left(\Sigma_{Y_{\text{patch}}}\right) + const.$, where $\Sigma_{Y_{\text{patch}}}$ is the empirical covariance of Y_{patch}. We maximize a penalized version of this upper bound on the information:

$$(1-\lambda)\log\det\left(\Sigma_{Y_{\text{patch}}}\right) - \lambda \mathbf{1}^T \bar{Y}_{\text{patch}}, \tag{2}$$

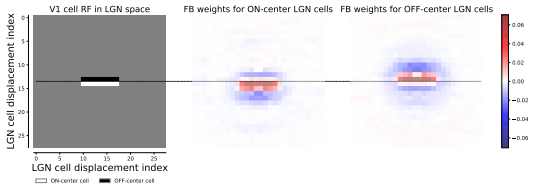
where $\lambda \in (0, 1)$ is the firing rate penalty parameter and $\bar{Y}_{\text{patch}}$ is the empirical mean of Y_{patch}.

We perform gradient ascent on the upper bound given in (2) to find the feedback kernels and the nonlinearity parameters. The gradients are calculated using automatic differentiation available in the PyTorch library. We use step size of 10^{-4} and run the optimization for 2500 full batch iterations [1]. The parameter λ is varied in $[0.775, 1)$. For each value of λ, we solved two optimization problems:

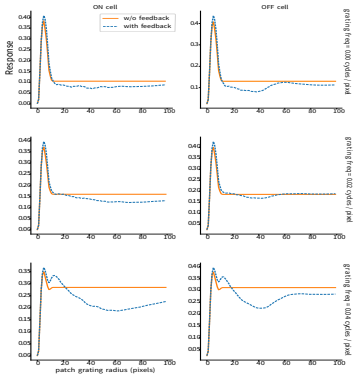
[1] The code and natural images used to produce the results of this paper are available at `https://anonymous.4open.science/r/learning-v1-lgn-feedback-9DBE`.



(A) **Information and firing rates vs. the firing rate penalty parameter λ.**



(B) **The learned feedback weights.**



(C) **Area summation curves with and without feedback.**

Figure 2: A The information normalized by average firing rate, information, and average firing rate of V1 as for $\lambda \in [0.75, 1)$ with and without feedback. B The learned feedback weights for $\lambda = 0.988$ which achieves the maximum information per spike. The feedback weights for ON-center (OFF-center) LGN cells determine the impact of the feedback from the V1 cell on ON-center (OFF-center) LGN cells. C The response of an ON and an OFF LGN cell as a function of the size of the patch sinusoidal grating for three different spatial frequencies with and without feedback.

one only over feedforward model nonlinearity parameters without any feedback, and one over both the feedback and feedforward model parameters.

Figure 2A shows the information, information per spike, and the average firing rate when λ ranges from 0.75 to 1.The information per spike with feedback achieves its maximum of 7154 for $\lambda = 0.988$. This is greater than the maximum achieved without feedback by approximately $\%14$. The learned feedback weights for $\lambda = 0.988$ show a pattern of connection that is aligned with the V1 RF (Figure 2B). In other words, feedback excites the ON(OFF)-center LGN cells that overlap the ON(OFF) region of the V1 receptive field. The feedback weights have a small inhibitory impact on the phase reversed cells. Those are the ON(OFF)-center cells on the OFF(ON) regions of the V1 receptive field. Another signature of the feedback pattern is a general inhibition of the LGN cells in a neighborhood of the V1 cell RF elongated in the preferred orientation of the cell. This can explain the impact of this feedback on stimuli integration, which is explored in section 3.1.

3.1 Impact of feedback on RF properties

To investigate the impact of the feedback on LGN RF properties, we computed the area summation curve for ON and OFF-center LGN cells with and without feedback. These curves are computed by measuring the peak response of the LGN cell to a windowed moving sinusoidal grating as the size of the window increases from zero to a few multiples of the LGN RF size. Figure 2C shows the area summation curves for spatial freq = 0 (light spot for ON-center cell, dark spot for OFF-center cell), 0.02 and 0.04 (cycles / pixel). In all three cases, feedback increases the peak firing rate for small windows and decreases the firing rate for large window sizes. This is in agreement with experimental results in [7], which show a similar feedback effect on stimuli integration in LGN.

3.2 Phase-aligned vs phase-reversed

Our model suggests a phase-aligned feedback arrangement for the LGN cells directly overlapping the V1 RF consistent with results in [7, 24]. In [33], by measuring the burst-to-tonic ratios for LGN cells after feedback deactivation, the authors deduce a phase-reversed arrangement for LGN cells that directly overlap the V1 cell RF. In [18], A computational model based on predictive coding has been suggested to account for this. In a recent work, Lian et al propose a model of V1-LGN pathway based on efficient coding that learns the feedback weights using an anti-Hebbian rule [21]. This makes the feedback weights converge to the negative of feedforward weights, resulting in a phase-reversed

pattern. However, there is overwhelming evidence that feedback impact goes beyond the receptive field of the V1 cell [30, 32]. We believe more comprehensive experiments are required to resolve these conflicting results.

References

[1] I. M. Andolina, H. E. Jones, W. Wang, and A. M. Sillito. Corticothalamic feedback enhances stimulus response precision in the visual system. *Proceedings of the National Academy of Sciences*, 104(5):1685–1690, 2007.

[2] F. Attneave. Some informational aspects of visual perception. *Psychological review*, 61(3):183, 1954.

[3] V. Balasubramanian and P. Sterling. Receptive fields and functional architecture in the retina. *The Journal of physiology*, 587(12):2753–2767, 2009.

[4] H. B. Barlow et al. Possible principles underlying the transformation of sensory messages. *Sensory communication*, 1(01), 1961.

[5] A. J. Bell and T. J. Sejnowski. The "independent components" of natural scenes are edge filters. *Vision research*, 37(23):3327–3338, 1997.

[6] J. Bickle, M. Bernstein, M. Heatley, C. Worley, and S. Stiehl. A functional hypothesis for lgn-v1-trn connectivities suggested by computer simulation. *Journal of computational neuroscience*, 6(3):251–261, 1999.

[7] G. Born, F. A. Schneider-Soupiadis, S. Erisken, A. Vaiceliunaite, C. L. Lao, M. H. Mobarhan, M. A. Spacek, G. T. Einevoll, and L. Busse. Corticothalamic feedback sculpts visual spatial integration in mouse thalamus. *Nature neuroscience*, 24(12):1711–1720, 2021.

[8] F. Briggs and W. M. Usrey. Parallel processing in the corticogeniculate pathway of the macaque monkey. *Neuron*, 62(1):135–146, 2009.

[9] F. Crick. Function of the thalamic reticular complex: the searchlight hypothesis. *Proceedings of the National Academy of Sciences*, 81(14):4586–4590, 1984.

[10] J. Cudeiro and A. M. Sillito. Looking back: corticothalamic feedback and early visual processing. *Trends in neurosciences*, 29(6):298–306, 2006.

[11] C. Enroth-Cugell and J. G. Robson. The contrast sensitivity of retinal ganglion cells of the cat. *The Journal of physiology*, 187(3):517–552, 1966.

[12] D. Eyding, J. D. Macklis, U. Neubacher, K. Funke, and F. Wörgötter. Selective elimination of corticogeniculate feedback abolishes the electroencephalogram dependence of primary visual cortical receptive fields and reduces their spatial specificity. *Journal of Neuroscience*, 23(18):7021–7033, 2003.

[13] S. Grossberg and E. Mingolla. Neural dynamics of perceptual grouping: Textures, boundaries, and emergent segmentations. In *The adaptive brain II*, pages 143–210. Elsevier, 1987.

[14] S. Grossberg and J. R. Williamson. A neural model of how horizontal and interlaminar connections of visual cortex develop into adult circuits that carry out perceptual grouping and learning. *Cerebral cortex*, 11(1):37–58, 2001.

[15] E. Harth. Visual perception: A dynamic theory. *Biological Cybernetics*, 22(3):169–180, 1976.

[16] J. M. Hasse and F. Briggs. Corticogeniculate feedback sharpens the temporal precision and spatial resolution of visual signals in the ferret. *Proceedings of the National Academy of Sciences*, 114(30):E6222–E6230, 2017.

[17] D. H. Hubel and T. N. Wiesel. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of physiology*, 160(1):106–154, 1962.

[18] J. F. Jehee and D. H. Ballard. Predictive feedback can account for biphasic responses in the lateral geniculate nucleus. *PLoS computational biology*, 5(5):e1000373, 2009.

[19] H. E. Jones, I. M. Andolina, B. Ahmed, S. D. Shipp, J. T. Clements, K. L. Grieve, J. Cudeiro, T. E. Salt, and A. M. Sillito. Differential feedback modulation of center and surround mechanisms in parvocellular cells in the visual thalamus. *Journal of Neuroscience*, 32(45):15946–15951, 2012.

[20] Y. Karklin and E. P. Simoncelli. Efficient coding of natural images with a population of noisy linear-nonlinear neurons. In *Advances in neural information processing systems*, pages 999–1007, 2011.

[21] Y. Lian, D. B. Grayden, T. Kameneva, H. Meffin, and A. N. Burkitt. Toward a biologically plausible model of lgn-v1 pathways based on efficient coding. *Frontiers in neural circuits*, 13:13, 2019.

[22] K. McAlonan, J. Cavanaugh, and R. H. Wurtz. Guarding the gateway to cortex with attention in visual thalamus. *Nature*, 456(7220):391–394, 2008.

[23] J. W. McClurkin, L. M. Optican, and B. J. Richmond. Cortical feedback increases visual information transmitted by monkey parvocellular lateral geniculate nucleus neurons. *Visual neuroscience*, 11(3):601–617, 1994.

[24] M. H. Mobarhan, G. Halnes, T. Hafting, M. Fyhn, G. T. Einevoll, et al. Firing-rate based network modeling of the dlgn circuit: Effects of cortical feedback on spatiotemporal response properties of relay cells. *PLoS computational biology*, 14(5):e1006156, 2018.

[25] V. L. Mock, K. L. Luke, J. R. Hembrook-Short, and F. Briggs. Dynamic communication of attention signals between the lgn and v1. *Journal of neurophysiology*, 120(4):1625–1639, 2018.

[26] B. A. Olshausen and D. J. Field. Natural image statistics and efficient coding. *Network: computation in neural systems*, 7(2):333–339, 1996.

[27] P. Sastry, S. Shah, S. Singh, and K. Unnikrishnan. Role of feedback in mammalian vision: a new hypothesis and a computational model. *Vision Research*, 39(1):131–148, 1999.

[28] S. M. Sherman. Tonic and burst firing: dual modes of thalamocortical relay. *Trends in neurosciences*, 24(2):122–126, 2001.

[29] S. M. Sherman and R. Guillery. The role of the thalamus in the flow of information to the cortex. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 357(1428):1695–1708, 2002.

[30] T. Tsumoto, O. Creutzfeldt, and C. Legendy. Functional organization of the corticofugal system from visual cortex to lateral geniculate nucleus in the cat. *Experimental Brain Research*, 32(3):345–364, 1978.

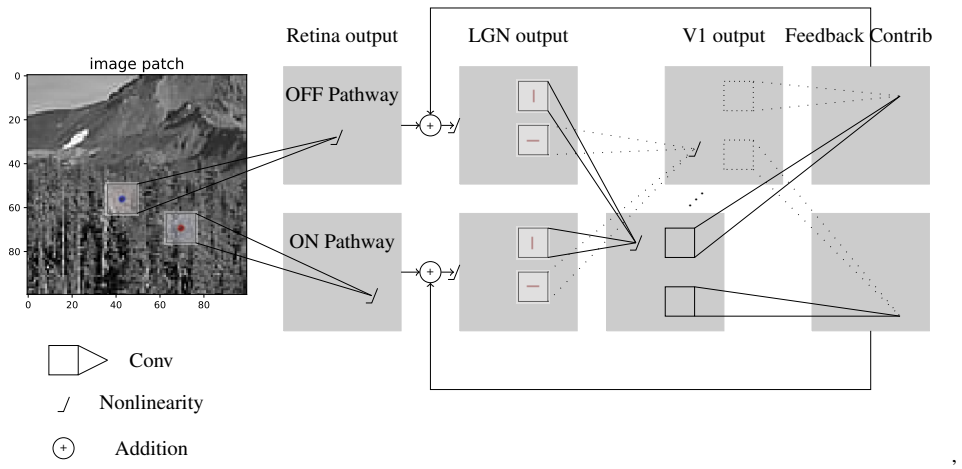
[31] J. H. Van Hateren and A. van der Schaaf. Independent component filters of natural images compared with simple cells in primary visual cortex. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 265(1394):359–366, 1998.

[32] W. Wang, I. M. Andolina, Y. Lu, H. E. Jones, and A. M. Sillito. Focal gain control of thalamic visual receptive fields by layer 6 corticothalamic feedback. *Cerebral cortex*, 28(1):267–280, 2018.

[33] W. Wang, H. E. Jones, I. M. Andolina, T. E. Salt, and A. M. Sillito. Functional alignment of feedback effects from visual cortex to thalamus. *Nature neuroscience*, 9(10):1330, 2006.

A Appendix

A.1 Supplementary figures



(A)

Figure 3: **Block diagram of the model.**

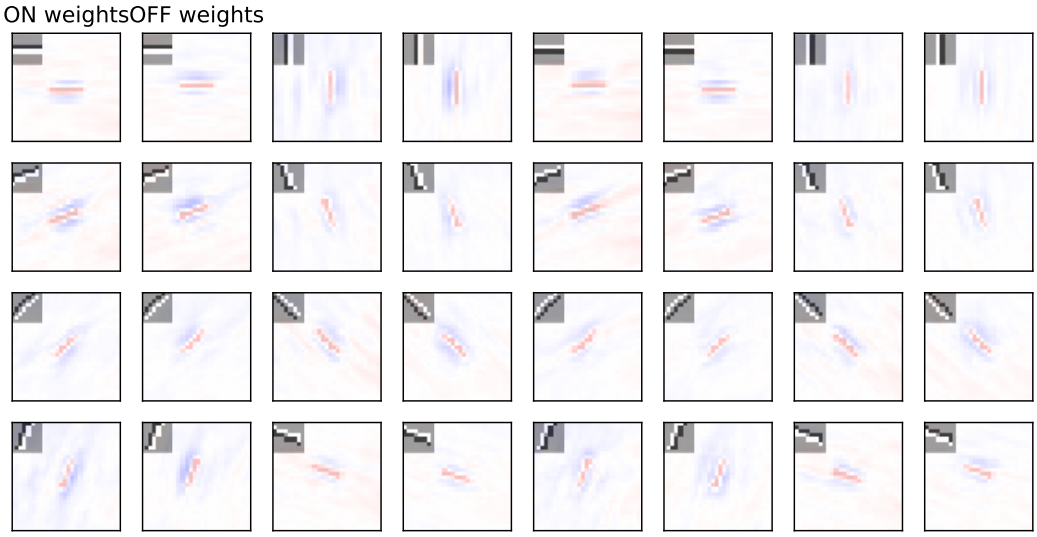


Figure 4: **All the learned feedback weights.** The two sets of weights for all the 16 types of V1 filters, where the V1 filter in LGN space is shown in the upper left corner for reference.