

Uncovering Potential Effects of Spontaneous Waves on Synaptic Development: The Visual System as a Model

Jennifer Crodelle^{1,*} and Wei P. Dai^{2,3,*†}

¹*Department of Mathematics and Statistics, Middlebury College, Middlebury, Vermont 05753, USA*

²*Research Institute of Intelligent Complex Systems, Fudan University, Shanghai, Shanghai 200433, China*

³*Shanghai Artificial Intelligence Laboratory, Shanghai, Shanghai 200232, China*



(Received 9 June 2025; accepted 16 January 2026; published 18 February 2026)

Spontaneous waves are ubiquitous during early brain development and are hypothesized to drive the development of receptive fields (RFs). Different stages of retinal waves, characterized by activity patterns in the retina and downstream areas, coincide with key developments along the visual pathway. Stage II waves, in particular, have been linked to a possible early synaptic pruning process; however, the mechanisms underlying their activity-dependent effects on thalamocortical synapses remain poorly understood. In this work, we develop a biologically constrained model of thalamocortical synapse formation driven by stage II retinal waves. Using this, alongside a reduced rate-based model, we propose mechanisms underlying the formation of distinct RF structures, including transitions from a disklike RF to a ringlike RF, modulated by learning parameters and initial synaptic connection strengths. We leverage the mechanisms uncovered through the reduced model to make two predictions: gap junctions promote local retinotopic refinement during stage II waves, and the resulting RF structure established during this stage controls the emergence of orientation selectivity later in development. Mechanisms identified here may aid in the understanding of developing synaptic structures across various cortical regions.

DOI: [10.1103/qr68-9vy9](https://doi.org/10.1103/qr68-9vy9)

I. INTRODUCTION

Patterned spontaneous activity during development is ubiquitous in the brain [1–4]. The retinal wave is one of such activity that propagates throughout the entire visual system before eye opening [3]. Its functional role in setting up synaptic structures along the visual pathway has been the subject of many studies [3,5–8], yet the exact mechanisms remain unclear. In the developing visual system of mice and ferrets, spontaneous waves in the retina occur from before birth (stage I) through the first postnatal week (stage II) to right before eye opening (stage III) [9]. These stages of waves are characterized by different mechanisms driving the activity and, as such, display different activity patterns. Stage I waves are mainly propagated through gap-junction connections between retinal ganglion cells (RGCs) [10,11]. Stage II waves, which begin around birth and last through the first postnatal week [9], are dependent upon acetylcholine release from starburst amacrine cells and are characterized by relatively slow wave speeds and restricted domain sizes [12]. Stage II waves have been implicated in instructing healthy receptive field refinement [6,13,14]. Stage III waves, which begin in the second

postnatal week and last until eye opening, depend upon glutamate release of retinal bipolar cells and are characterized by highly correlated firing among RGCs of the same sign (ON-ON or OFF-OFF), as well as distinct firing patterns across different RGC types (e.g., alpha, beta, gamma), potentially contributing to distinct wiring patterns across different RGCs [9,15]. In addition, due to the asynchronous and temporally delayed firing of ON and OFF cells during stage III waves, it is hypothesized that orientation selectivity of primary visual cortical (V1) cells is mostly set up during this stage of development [16].

The correlation between spontaneous retinal waves and circuit development has been studied for years [17,18], with the most recent evidence supporting activity-dependent plasticity as underlying circuit refinement during this time [19]. Abolishing or disrupting the retinal waves has been shown to disrupt circuit formation, though the interpretation of these results is often unclear due to other confounding factors [9]. Nonetheless, experiments have shown a clear causal link between nicotinic acetylcholine receptors, retinal waves, and the refinements of V1 circuitry [13]. Blocking stage II retinal waves results in large axonal projections and too-large receptive fields (RFs), as does NMDA-blockers in downstream neurons which disrupt the correlated firing needed for synaptic refinement and plasticity [5,6]. Additionally, in mice whose retinal waves have been blocked genetically, the RGC axonal arbors terminating in the dorsal lateral geniculate nucleus (dLGN) (and superior colliculus) have longer branch lengths that cover a larger area when compared to the wild-type mouse [7,11,20].

*These authors contributed equally to this work.

†Contact author: weidai@fudan.edu.cn

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

While experiments provide evidence that stage II waves play a role in synaptic refinement, the exact mechanisms underlying the reduction of RF size and the improved acuity of a single neuron remain elusive [16]. In particular, the relationship between characteristics of the retinal waves and the extent of synaptic pruning is unknown. In this work, we build a physiologically plausible computational model of the feedforward processing in developmental V1 driven by retinal waves to predict how the characteristics of the retinal waves may affect the learned structure of V1 RFs. Together with a rate-based reduced model analysis, we provide the underlying mechanisms to understand these pruning effects.

Beyond the pruning process, this understanding leads to an immediate demonstration on how simple periodic RF structures can arise under certain conditions in our model, which may aid in the understanding of the formation of more sophisticated periodic RFs or connectivity in different brain regions [21,22] through spontaneous waves [4,23]. As one application, we also extend our model to include gap junctions between V1 neurons during the stage II period [24–26] and show that electrical coupling during stage II waves may lead to the refinement of local retinotopy. Other than the retinotopic map, orientation selectivity is one of the most important common features in V1 across species [27–29] that emerges before eye-opening [9]. Unfortunately, little is known about how orientation preference is seeded without experience (likely during stage III [16,30]); however, foundational computational models have shown that orientation selectivity can emerge through Hebbian-like learning rules with correlated activity [17,18]. These models don't incorporate developmental characteristics during stage II that may affect the emergence of orientation selectivity. To shed light on this topic, we apply our model, connecting stages II and III, to first demonstrate that orientation selectivity can indeed be induced by randomly oriented stage III waves as suggested; and secondly, that its emergence under stage III retinal waves is critically affected by the pruning process during stage II waves. Through these applications, we highlight how the characteristics of spontaneous waves may influence different formations of synaptic structures.

II. RESULTS

We begin by constructing a biologically motivated model using relevant parameters from the mouse visual pathway to describe spontaneous stage II waves of retinal origin traveling through a local sheet of LGN cells, which in turn project onto cells in V1 with plastic synapses. The waves are considered as randomly oriented parallel wave fronts that originate distally from the local LGN sheet, motivating our approach to model them as oriented bars. LGN cells are organized on a 16×16 grid (corresponding to a 20×20 degree visual field in mouse), with ON and OFF cells co-located at each vertex [31] (which respond to light and dark stimuli, respectively), and output spikes in response to spatiotemporal activation from the waves through a linear-nonlinear-Poisson process. Individual V1 cells receive synaptic input from 95% of randomly selected LGN cells (though the choice of percentage is rather arbitrary, note that a reasonable range of sparsity do not affect our result qualitatively, see Fig. 11, in Appendix A) and are

modeled as adaptive exponential integrate-and-fire neurons [32]. The weights of the synaptic connections from LGN to V1 are plastic, following the triplet spike-timing dependent plasticity rule with parameters fit to visual cortex [33,34]. To start, V1 cells are uncoupled [25], essentially treated as random realizations or trials in the model. Gap junctions among V1 neurons are considered later in the paper. See Fig. 1(a) for a model schematic and Sec. V for details on the model equations and parameters.

To study the RF development of V1 neurons during stage II waves, in Figs. 2 and 3 in Sec. IIB, we present retinal waves to the local sheet of LGN cells and measure the RF of each V1 cell as the collection of potentiated synapses from LGN to V1. Note that we use RF, feedforward RF and feedforward LGN connectivity interchangeably in this work.

To achieve a in-depth understanding of the mechanisms underlying the model-predicted refinement of RFs, in Sec. IIC, we introduce a reduced firing rate-based version of this model and use it to identify the main mechanisms contributing to RF refinement across wave speed, width and more. Finally, we leverage the mechanistic understanding of RF refinement across several applications, including: (i) the emergence of a ringlike (and periodic) RF structure is dependent upon initial connectivity strength and the amount of long-term depression in the plasticity model (see Fig. 4 in Sec. IIC 1); (ii) the presence of gap junctions with RF refinement during stage II plays an important role in maintaining local retinotopy (see Fig. 5 in Sec. IID); and (iii) the size and shape of the RF after stage II waves affects the degree of orientation selectivity that arises during stage III waves (see Fig. 6 in Sec. IIE).

A. Stage II waves guide the synaptic pruning of LGN to V1 connections

We begin by demonstrating that, with parameters for wave speed and width chosen near the average for the range of experimentally determined values [35], our model simulations result in a refinement of the RF of V1 cells. The simulation begins with 1024 uncoupled V1 neurons receiving plastic excitatory input from a circular patch of the LGN sheet, mimicking the reachable pool of LGN cells through their local spread of axonal terminals; recall Fig. 1(a). In response to many presentations of stage II retinal waves, we record how the synaptic weights from LGN cells change over time and measure the RF as the radius of the collection of potentiated synapses (see Sec. V for more details on RF measures).

As expected, due to the simultaneous activation of ON and OFF LGN cells during stage II [36], the distribution of the connection weight difference between ON and OFF LGN cells to one V1 neuron is balanced on average over time; see Fig. 1(b). Synapses from LGN cells near the center of the local sheet potentiate over time while those in the surrounding regions decay, resulting in a disk-shaped RF structure for both ON and OFF LGN cells; see Fig. 1(c). Recording the weight of the ON [Fig. 1(d)] and OFF [Fig. 1(e)] synapses into one V1 cell across the radial axis and time (averaged over 24 trials) shows that synaptic weights at small radial distances, near the middle of the LGN pool, potentiate while those in the periphery decay. Averaging over time, we quantify this shrinking of the disklike RF using two measures: (1) connection-weighted

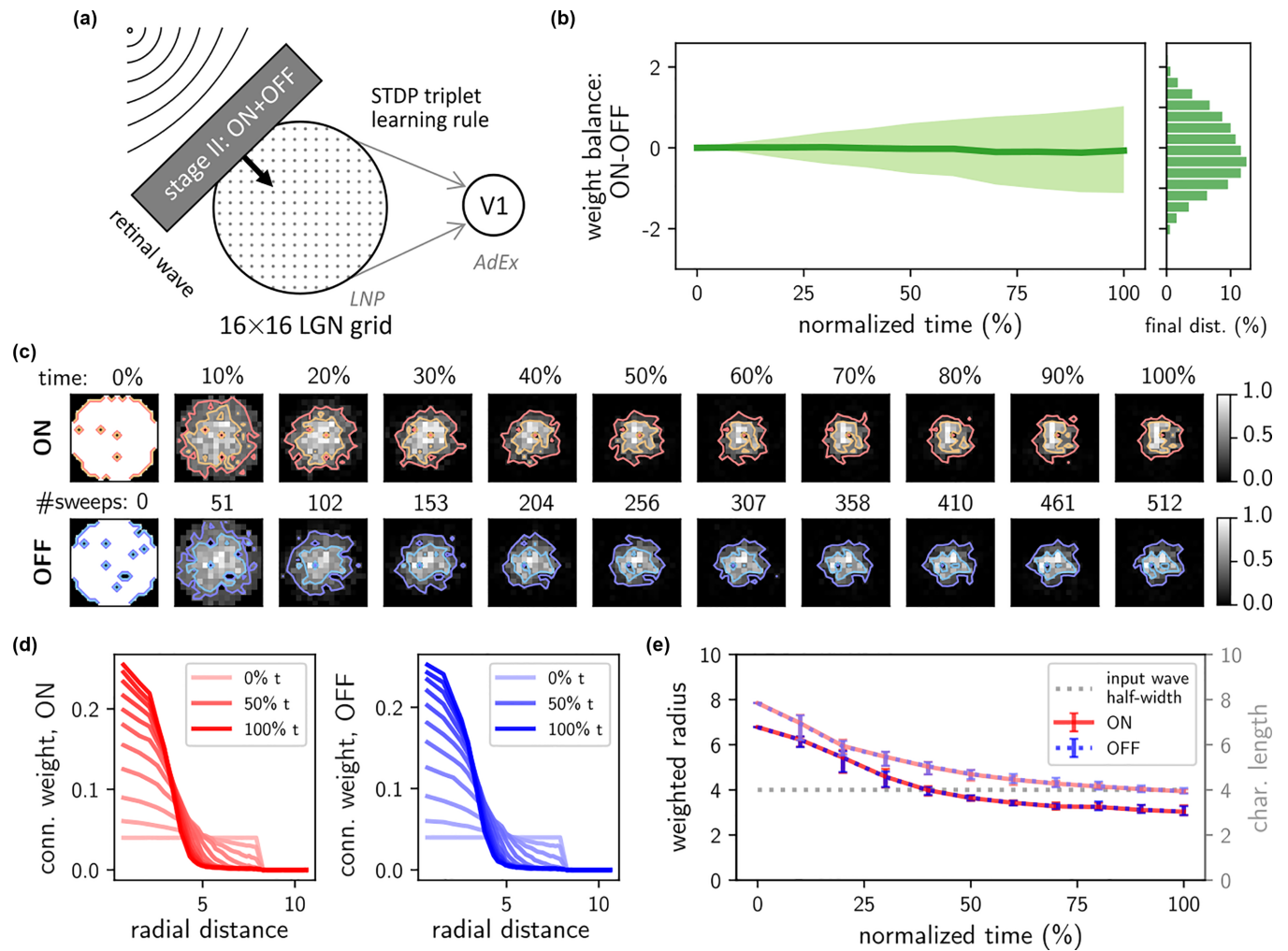


FIG. 1. Model setup and typical receptive field refinement. (a) Schematic of the stage II wave as a random parallel wave front that activates the LGN cells. Synapses from LGN to all V1 cells are plastic, following a triplet STDP learning rule. (b) Temporal structure (left) and final distribution (right) of the difference between the connection strength from ON and OFF LGN cells, averaged across all 1024 V1 neurons. Shading represents standard deviation. (c) The ON (top) and OFF (bottom) RF refinement of one sample V1 neuron over time shown by the spatiotemporal change of connection weights in grayscale (normalized by local maximum). The red (ON) and blue (OFF) solid-line contours enclose the potentiated connections while the orange (ON) and navy (OFF) contours trace the local half-maximum. (d) Average ON (left, red) and OFF (right, blue) connection weight distribution along the radial axis (averaged over polar angles) over time (saturating color) averaged across 24 trials. (e) (left y axis, saturated colors) Temporal evolution of radius-weighted ON (solid red) and OFF (dotted blue) LGN connection weights averaged across all 1024 V1 neurons and (right y axis, less saturated colors) the characteristic length averaged across all 1024 V1 neurons. The error bars show the standard deviation in the average over 24 trials. The gray dotted line is at the half-width of the input wave.

radius: the average distance to RF center weighted by connection strength; (2) characteristic length: the half-width of the contoured area as measured by the square root of the area of potentiated weights divided by two (see Sec. V for more details); see Fig. 1(f). We note here that synapses from ON and OFF LGN cells behave identically, consistent with experiments showing that ON-OFF segregation does not occur until near the end of stage II retinal waves [9]. As such, we refer to the RF of both ON and OFF cells as just “RFs” unless otherwise specified.

B. Effect of wave speed and width on V1 RF

Stage II waves in experiments on neonatal mice show a large variation in speed and width [35]. In the model, we vary

the speed and width of incoming retinal waves and characterize the effects on the RF structure.

The wave speed, v , is varied from 2 deg/s (half the default speed) to 16 deg/s (four times the default speed) while all other parameters remain the same. In this setup, the number of waves in a fixed time range increases with the wave speed, however, the total LGN activation, measured by the average number of LGN spikes in the period of the slowest wave remains mostly constant across wave speeds [see Fig. 8(a), in Appendix A].

Results show that the size of the disklike RF shrinks as wave speed increases. This occurs consistently across all V1 cells as exemplified in a sampled trial in Fig. 2(a). Nonetheless, the overall firing rate of the V1 cells increases with wave speed while the rising and saturating phases of the V1

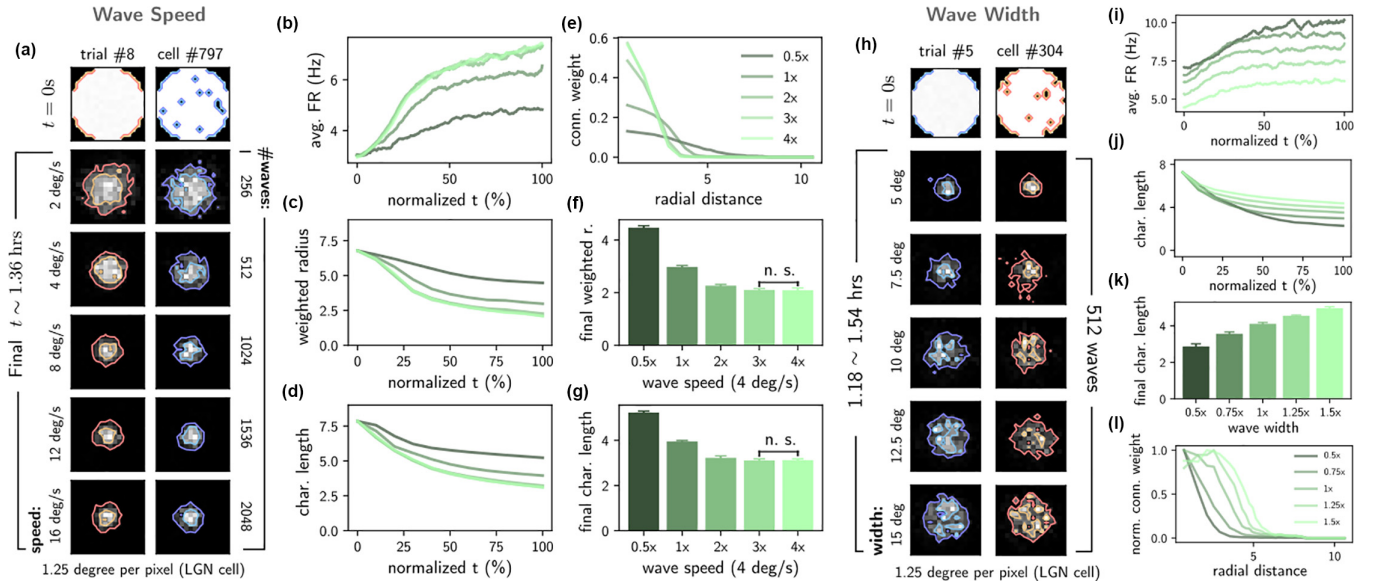


FIG. 2. Effects of varying wave speed and width. Panels (a)–(g) are for wave speed, and panels (h)–(l) for wave width. (a) The LGN-V1 normalized connection profiles (ON in red, OFF in blue) with different wave speeds for a population-averaged trial ($n = 1024$, left) and a single neuron example (right). The top row shows available connections at $t = 0$, while the remaining rows show the final weight distribution for increasing wave speeds. (b) Averaged V1 firing rates over time (81.7 minutes in total) across wave speeds. (c) Averaged connection-weighted radius of LGN connections over time across wave speeds. (d) Characteristic length over time across wave speeds. (e) Radial distribution of final connection weights (averaged over discrete polar angles) across wave speeds. (f) Weighted radius at the end of the simulation across wave speeds. (g) Characteristic length at the end of the simulation across wave speeds. (h)–(l) The same as panels (a)–(g) for wave speed, but for wave width and using the characteristic length measurement only. Error bars in panels (f), (g), and (k) denote the 10%–90% range of values from 18 trials. T test between $3\times$ and $4\times$ the typical wave speed (4 deg/s) in panels (d) and (g) gives p values of 0.4253 and 0.4966, respectively. Green lines and bars indicate averages across both ON and OFF subregions (or LGN cells) across all V1 neurons.

firing rate, indicating periods of learning, are matched across speeds; see Fig. 2(b). Over time, the connection-weighted radius of the disklike RF decreases for all wave speeds, illustrating a pruning process, with the greatest decrease occurring for the two fastest wave speeds; see Figs. 2(c) and 2(d). Not only does increasing wave speed result in a smaller disklike RF, but the connection weights at the center of the RF are larger for faster wave speeds, decreasing dramatically as the radial distance increases; see Fig. 2(e). These results are supported by measurements of the connection-weighted radius and characteristic length, which show a significant decrease as wave speed increases, although notably there is no significant difference between RFs resulting from the two fastest speeds; see Figs. 2(f) and 2(g).

The wave width, w , is varied from 4 LGN cells (half the default width) to 12 LGN cells (1.5 times the default width). The waves are presented at random orientations at the default wave speed (4 deg/s) and we measure the resulting RF structure.

As wave width increases, the size of the disklike RF increases, as measured across all V1 cells in an individual trial [Fig. 2(h), left column, and a typical single-neuron example on the right]. In this case, larger wave widths activate more LGN cells per unit time [see Fig. 8(b), in Appendix A]. To adjust for this change across changes in wave width, we decrease the initial connection strength of the LGN synapses directly in proportion to the average number of spikes induced by different wave widths to keep LGN activation the same [see

Fig. 8(b), and Figs. 8(c)–8(f), in Appendix A for results with constant initial connection strength]. The resulted rising and saturating phases of the firing rate of the V1 cell are matched across different widths; see Fig. 2(i). The characteristic length of the potentiated area decreases over time for all widths, again illustrating a pruning process regardless of wave width; see Fig. 2(j). In addition, the width of the final connection profile increases linearly with wave width; see Fig. 2(k). Consistently, the normalized connection weights as a function of the radial distance widens as the wave width increases; see Fig. 2(l).

C. Reduced rate-based model of V1 RF development

To understand the potential mechanisms underlying the full model-predicted changes in RF structure, we introduce a reduced firing-rate model. In this model, we collapse the LGN onto a one-dimensional space parameterized by x , along which retinal waves of width $2d$ may be presented from left to right or right to left; see Fig. 3(a). The reduced model comprises three differential equations for (1) the firing rate of the V1 cell, $r(t)$; (2) the average firing rate of the V1 cell, $\bar{r}(t)$; and (3) the connection weight from a position, x , in the one-dimensional LGN space to the V1 neuron, $w(x, t)$. Following previous approaches to deriving a BCM-like mean-field approximation to the triplet learning rule [34] and applying reasonable assumptions on the relative rate of change of the weight compared to the firing rate (see Appendix B for full

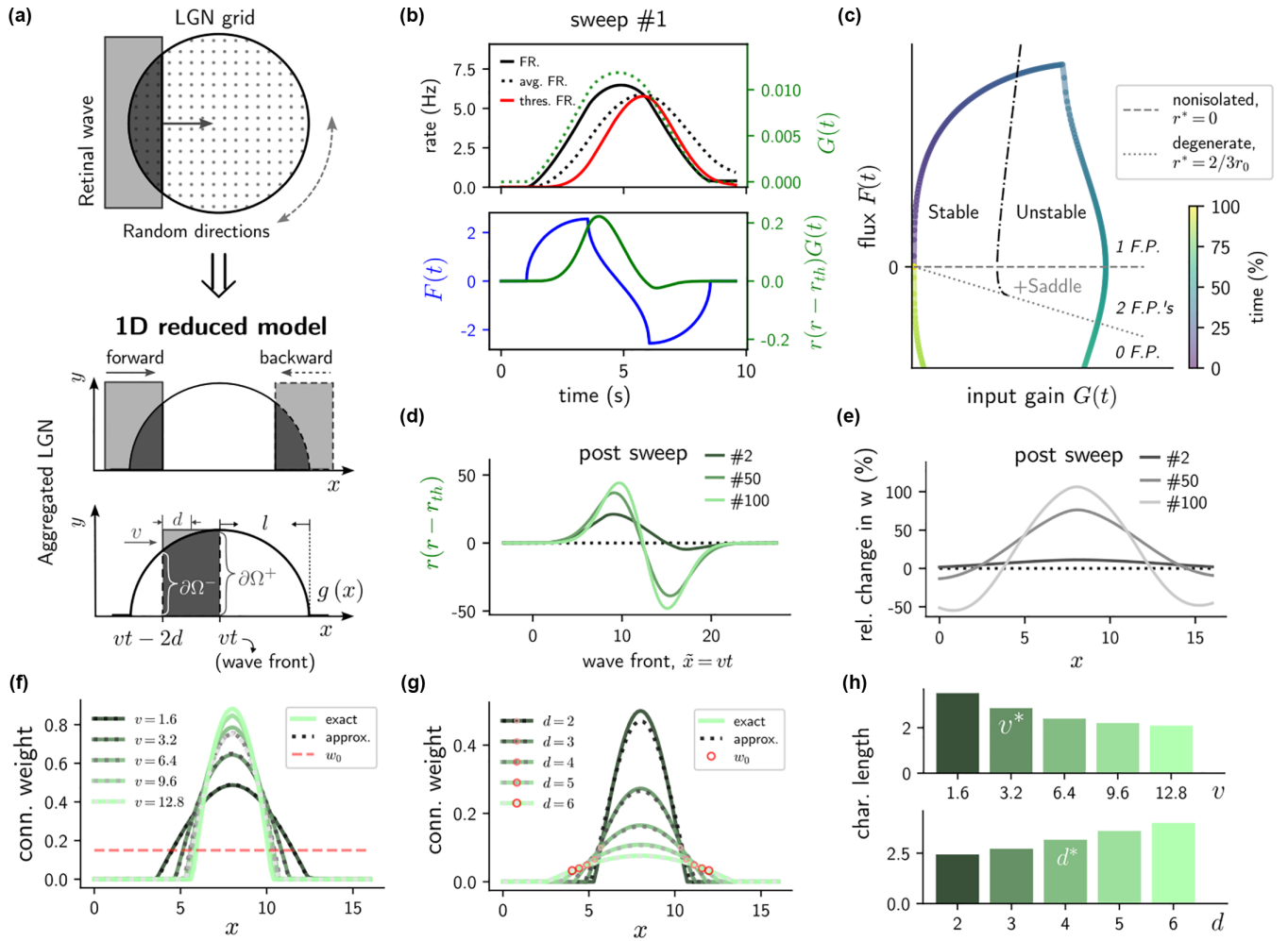


FIG. 3. Rate-based reduced model. (a) Schematic illustrating the 1D reduced model derived from the full simulation and the bottom panel accompanies Eq. (2). (b) Temporal evolution of the system during the first wave. The upper panel shows V1 firing rate (solid) and its average (dotted) in black, the threshold firing rate, r_{th} , in red. The input gain, $G(t)$, in dotted green is plotted against the right y axis. The lower panel shows the two terms on the right-hand-side of dr/dt in Eq. (1). (c) Phase diagram based on changes in $G(t)$ and $F(t)$ with parameters chosen to illustrate all possible regimes. Regimes of single, double and zero fixed points (FPs) are separated by a dashed gray line of nonisolated FPs, and a dotted gray line of degenerate FPs, in-between which lies the saddle points. The stable and unstable regimes are separated by the dash-dotted line. (d) Temporal change in $r(r - r_{th})$ after 2, 50, and 100 wave sweeps plotted against the wave front, $\tilde{x} = vt$. (e) Change in the LGN connection weights relative to their initial values in percentage across LGN space, x . (f) Spatial weight distribution of the reduced model are plotted for [50, 100, 200, 300, 400] sweeps with the increasing wave speeds, v , accordingly, making sure the activation durations are kept the same. The exact solutions are denoted by solid green lines and approximated solutions are denoted by dotted gray lines. Increment in the wave speeds are denoted by the increment in lightness. The red dashed line marks the initial connection weight, $w_0 = 0.15$. (g) Same as panel (f), but for change in the wave width, d , and different initial weights, $w_0 \simeq [0.08, 0.06, 0.05, 0.04, 0.035]$, are marked in red circles. (h) The characteristic length at the end of the simulation across wave speed (top) and width (bottom).

reduced model derivation), we consider the resulting two-dimensional ODE model for the firing rate and average firing rate as follows:

$$\begin{aligned} \frac{dr}{dt} &= r(r - r_{th})G(t) + F(t), \\ \tau \frac{d\bar{r}}{dt} &= -\bar{r} + r, \end{aligned} \quad (1)$$

where r_{th} is the threshold firing rate that determines the sign of the input contribution to firing rate.

The equation for the change in average firing rate, $\bar{r}(t)$, is straightforward. The contribution to the dynamics of the firing rate, $r(t)$, consists of two parts: the weight change due

to learning within an activated area of LGN, described by the input gain $G(t)$, and the weight change on the activation boundaries of the wave, described by the flux $F(t)$. Integrating the LGN input over the activated part of the pool results in the following equations for the input gain and flux:

$$\begin{aligned} G(t) &= g_0 A^+ \tau_+ \tau_{slow} \hat{r}^2 \int_{b(t)}^{u(t)} g(x) dx, \\ F(t) &= g_0 \hat{r} v [g(x)w(x)]_{\partial\Omega(t)}, \end{aligned} \quad (2)$$

where g_0 is the intrinsic gain of the neuron, v is the speed of a wave with half-width d traveling in a pool of LGN cells with radius l and firing rate $\hat{r}$, when activated. The flux depends on

evaluating the LGN activation curve, $g(x)$, and the synaptic weights, $w(x)$ at the boundary of the wave, $\partial\Omega$ defined by $u(t)$ and $b(t)$ (see details in Appendix B). Assuming synaptic weight changes are on a slower timescale than firing rate changes, we write the simplest version of the reduced model which updates the weights after the i -th presentation of a wave, $w_i(x)$, as follows:

$$w_i(x) = w_{i-1}(x) + \tau_+ \tau_{\text{slow}} A^+ \hat{r} \int_{x/v}^{(x+2d)/v} r(r - r_{\text{th}}) dt. \quad (3)$$

The parameters τ_+ and τ_{slow} come from the STDP triplet rule (see Sec. V). Here, the rate-based weight dynamics is BCM-like as it involves a threshold firing rate, $r_{\text{th}} = (\bar{r}/r_0)\bar{r}$, which upon crossing changes the sign of weight change [34,37] (r_0 is the target firing rate that determines the amplitude of synaptic depression in the triplet rule). The assumption to keep weights constant during the wave presentation allows for a more analytically tractable two-dimensional system of ODEs; however, throughout this work, we show that the results for the full system (which includes the partial differential equation for the weight change) qualitatively match those for the simplified two-dimensional model since the learning rate, A^+ , is small, 1×10^{-3} .

To illustrate the dynamics of the reduced model, we plot the system variables, r , $\bar{r}$, r_{th} , including the input gain $G(t)$ and flux $F(t)$, during the presentation of the first wave [“sweep #1”; see Fig. 3(b)]. Notice that the firing rate of the V1 cell is mainly driven by the change in flux, which is one magnitude stronger than the input gain multiplied by the change in weights [see right y axis of the bottom panel in Fig. 3(b)] due to the small learning rate. As the wave sweeps past the center, the flux switches sign and input gain reaches its peak, while the firing rate reaches its own peak slightly later.

The temporal phase diagram of the flux and input gain in Fig. 3(c) demonstrates all possible regimes that the system can achieve during the first wave. The temporal trace of $F(t)$ and $G(t)$ is symmetric since the weight is fixed until the end of the wave by assumption; however, the experienced regimes are asymmetric during the rise and decay of the response. See Fig. 7(a), in Appendix A for sample phase portraits over the first wave sweep.

As the number of wave sweeps increase (from 2, 50, to 100 sweeps, alternating forward and backward), the effect size of threshold crossing is amplified, as shown in Fig. 3(d). More importantly, the synaptic change, proportional to $r(r - r_{\text{th}})$, experiences a much more significant change in depression over potentiation, and the crossing happens closer to the center of the LGN pool. Consequently, the relative change of the synaptic weight is a bell-shaped function that sharpens as the number of sweeps increases; see Fig. 3(e).

Next, we mimic the simulation experiments for changing wave speed and width in the reduced model. The predominant effect of increasing wave speed is realized as a vertical stretching of the flux, which comes from the linear dependence of $F(t)$ on v in Eq. (2) and manifests in the connection-weight profile as a significant increase in peak height at the center of the LGN space and a shrinking of peak width; see Fig. 3(f).

Conversely, increasing the wave width, d most dominantly affects the LGN input gain since increasing wave width increases the difference between the bounds of integration for $G(t)$ in Eq. (2). This manifests as a decrease in the amplitude of the peak at the center of the LGN space and a spreading of the width, as shown in Fig. 3(g). For more details surrounding the mechanisms underlying the observed changes in connection-weight profile as wave speed and width are changed, see Figs. 7(b)–7(g), in Appendix A.

Therefore, results from the reduced model are consistent with results from the full simulation [compare Fig. 2(g) and 2(k) to Fig. 3(h)], while also providing intuition behind the main mechanisms leading to the observed trends.

1. Formation of a ringlike and periodic receptive field

Here we use our understanding of the reduced model to uncover two mechanisms that may underlie the transition from a disklike RF to a ringlike RF. From the reduced model, we know that the location of the stable fixed point, as well as its transition to an unstable saddle point, depends upon the location of the crossing of $r(t)$ and r_{th} ; recall Figs. 3(b) and 3(c). As w_0 increases, $r(t)$ increases and widens, shifting the intersection point to the left and lowering the height of the stable fixed point during the first sweep; see Fig. 4(a), top row. At the largest tested w_0 value, across numerous presentations of waves, the crossing is shifted enough to the left such that the fixed point goes from stable to unstable (saddle) twice, creating two small peaks in the firing rate curve; see Fig. 4(a), bottom row.

Due to the $r - r_{\text{th}}$ term in Eq. (3), the crossing of r and r_{th} also determines the difference between long-term potentiation (LTP) and depression (LTD) in the weights. Increasing the initial connection weight initially results in a wider disklike RF; see three darkest green curves in Fig. 4(b). Eventually, however, for large enough initial weight, LTD is induced in the center of the LGN, leaving peaks in the connection weight on the ends of the LGN; see three lighter curves in Fig. 4(b).

A similar trend is observed when the ratio of LTD to LTP, termed the LTD ratio, is increased, accomplished by scaling A^+ in Eqs. (2) and (3); see bottom row of Figs. 4(a) and 4(c). Note that increasing A^+ is identical to decreasing r_0 (see Sec. V) and recall that $r_{\text{th}} = (\bar{r}/r_0)\bar{r}$. Thus, increasing A^+ results in an increase in r_{th} and induces in a similar left-ward shift of the intersection of $r(t)$ and r_{th} as observed in the case of changing w_0 .

This result is replicated in the two-dimensional full simulation as well. Increasing w_0 and the LTD ratio results in the formation of a ringlike RF, with potentiation of the weights along the outskirts of the RF and depression in the center; see Figs. 4(d)–4(f).

These results lead to a model prediction: periodic RFs. Expanding upon the mechanism of crossings of the firing rate and its threshold leading to peaks in the connection weight profile, we increase the length of the LGN space in the reduced model to allow for the possibility of more than one peak in the connection weight profile. Indeed, by extending the LGN space from 16 to 24 in the reduced model, additional crossings of the firing rate and threshold firing rate occurs; see Fig. 4(g). Here, the crossing happens when the stable fixed

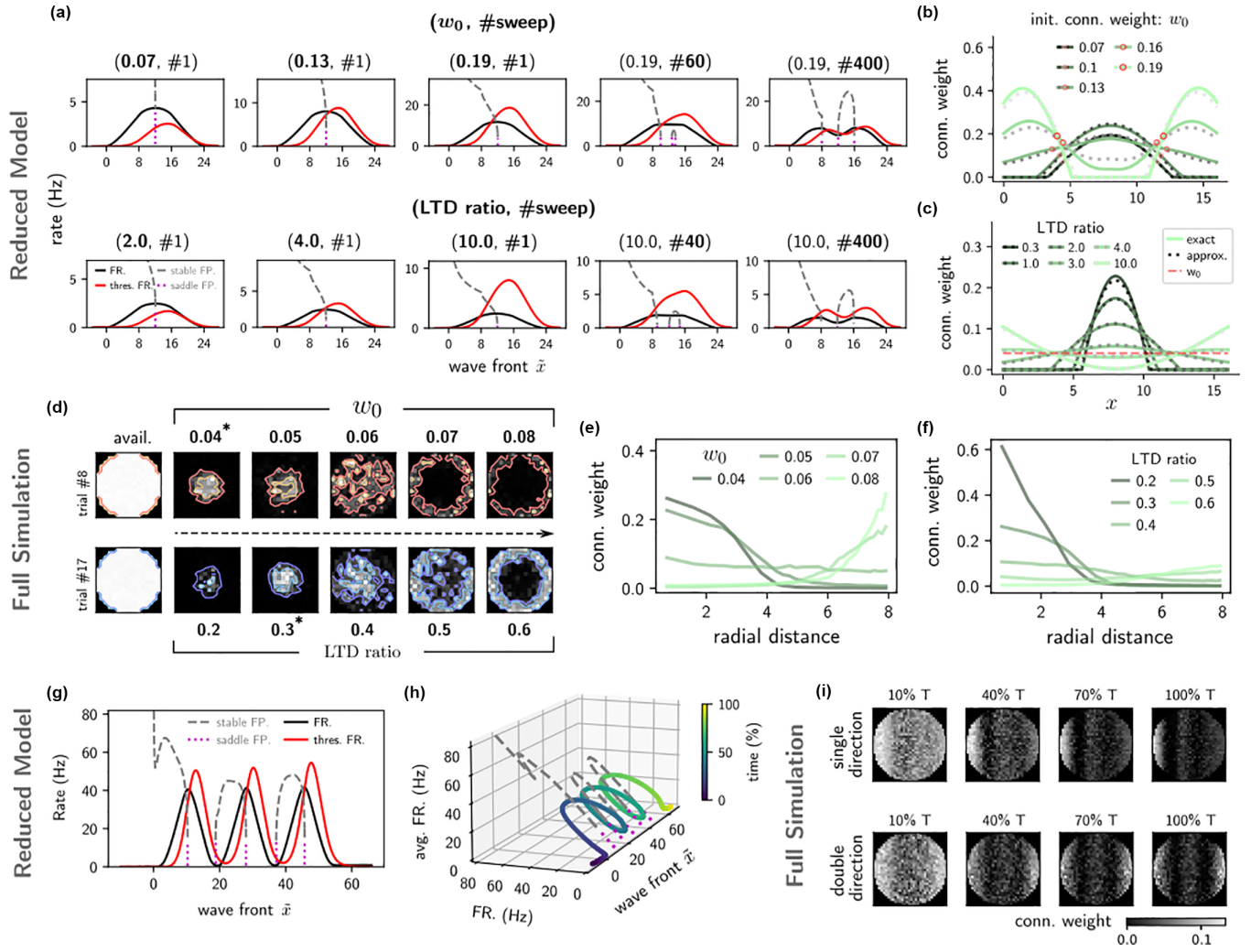


FIG. 4. Emergence of ringlike and periodic structures. (a) Average V1 firing rate (black curve) and threshold firing rate (red curve) plotted against wave front. In the top row, it is shown after the initial wave sweep across increasing initial weights (first three panels) and for a fixed high initial weight, experiencing increasing number of wave sweeps (last three panels). Gray dashed line shows the existence of a stable FP while pink dotted line shows the emergence of a saddle point. In the bottom row, the same is shown as in the top row but for LTD ratio. (b) Results from 400 sweeps of waves in alternating direction from the exact (solid green) and approximated (dotted gray) reduced model. (c) Same as panel (b), but for the LTD ratio. The initial connection weights are the same for different LTD ratios (dashed red). (d) Two trial-averaged samples from the full simulation as the initial weight increases (top row) and LTD ratio increases (bottom row). The asterisks denote the default parameters. The first column shows the available connections at $t = 0$. (e) Radial distribution of the final connection weights in the corresponding full simulations in panel (e). (f) The same as panel (e), but for LTD ratio. The radius of the LGN pool ($l = 8$) and wave width (half-width, $d = 4$) are standard for these ringlike structure. (g) Reduced model performance during the first wave sweep through a larger LGN space ($l = 24$), keeping the standard wave width ($d = 4$) and wave speed ($v = 4$ deg/s). (h) In the temporal phase plane (denoted by wave front, $\bar{x}$), average firing rate and firing rate spiral around the trace of the fixed points as they reemerge. (i) The connection weight profile in the full simulation showing periodic weight distribution over time (averaged over all neurons in a trial). LGN space is 32×32 . The top panel shows wave sweeps of single direction (left to right) while the bottom panel shows wave sweeps alternating in both directions horizontally. Parameters for reduced model(full simulation) for periodic structure: LTD ratio equals 0.3(0.3), $w_0 = 0.1(0.03)$

point annihilates the saddle node as dictated by the dynamics of the reduced model. The firing rate rises with the creation of the fixed points and quickly dips after the annihilation of the fixed points. In the 3D temporal phase space spanned by the firing rate and its average, we find the system exhibiting a coiled periodic structure [Fig. 4(h)]. These dynamics lead to the emergence of a third peak in the connection weight along the 1D LGN pool and verified in the original simulation; see Fig. 4(i) where vertical waves are presented in from left to

right in the top row and alternated in both directions in the bottom row.

D. Application: Gap junctions amend local continuity of the retinotopic map

So far, the V1 neurons in this model have been uncoupled. While cells in mouse V1 have yet to form cortical synapses during the developmental time of stage II waves, they are

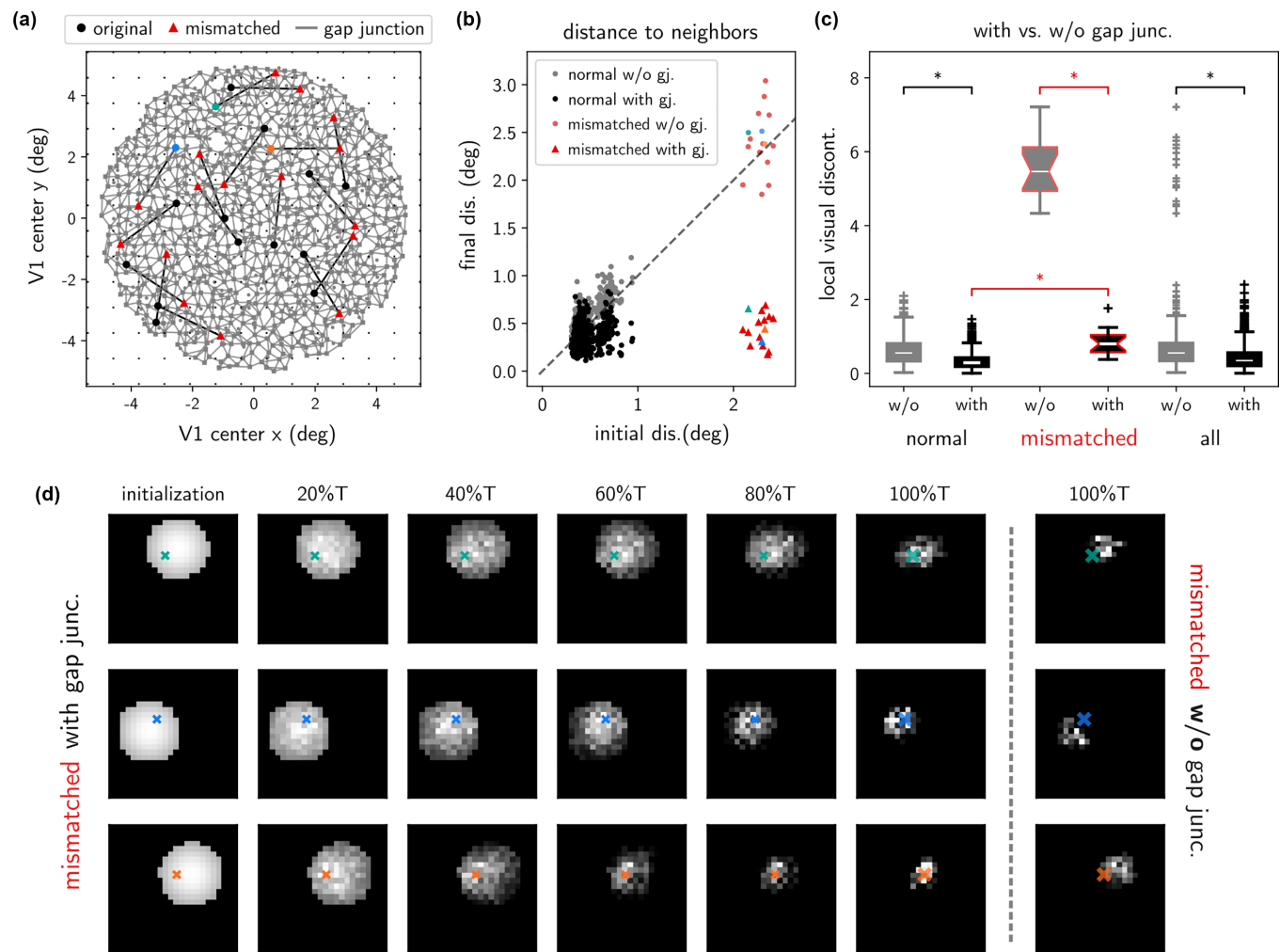


FIG. 5. Retinotopy refinement with gap junctions. (a) Schematic of the connectivity within V1 and its corresponding center in the 24×24 LGN grid. Gray squares correspond to V1 cells with normal LGN centers and gray lines connecting V1 neurons show gap junction connections to their nearest neighbors. Black dots mark the original RF center of V1 neurons with mismatched RF centers (red triangles). Colored dots match the corresponding sample neurons in panels (b) and (d). (b) Scatter plot of the RF center distance to neighbors before (x axis) and after refinement (y axis). The gray dots show simulation results without gap junctions, while black dots show results for a network with gap junctions. The dashed line denotes no change in the RF center distance to neighbors with refinement. (c) Local visual discontinuity (see Sec. V for definition) is calculated after refinement with and without gap junctions. V1 neurons are grouped into those with normal RF centers, those with mismatched RF centers, and the third group contains all V1 neurons. (d) Example refining process of V1 neurons with a mismatched RF center with gap junctions and without (right-most column). The cross marks the original RF center away from the center of the mismatched available LGN pool.

sparsely coupled with gap junctions [25]. Previous experimental and computational studies have shown that cells coupled by a gap junction during development tend to have more similar RFs at eye opening (in terms of preferred orientation) than those that were not coupled [38,39]. Here, we introduce gap junctions into the full simulation model and explore their effect combined with stage II waves on local retinotopic map refinement, expecting local gap-junctions during stage II waves will correct visually mismatched V1 cells through their nearest neighbors.

In the new setup shown in Fig. 5(a), the 1024 V1 cells now tile up a circular visual field with scattered visual positions and 75% of the V1 cells (gray squares) are allowed to be electrically coupled to their nearest neighbors with a proba-

bility of 75%, i.e., 4–6 gap junctions per neuron (depending on their distance to the boundary). While this is a higher coupling percentage than determined experimentally [40], we note that the results hold across changes in the gap junction connection probability (see Fig. 9 in Appendix A). The LGN pool is now extended to 24×24 such that each V1 cell samples from a circular subset of its local 16×16 subgrid of LGN cells that share the same center as the V1 cell. Additionally, we randomly select 16 gap-junction coupled V1 cells to sample from an LGN pool (red triangles) with a visual field center far from its supposed center (black circles). These V1 cells are referred to as “mismatched” cells.

We then measure the distance between the RF center of each V1 cell to the averaged RF center of its nearest neighbors

before and after the stage II waves [see Fig. 5(b)]. Consistent with previous studies, due to the synchrony induced by gap junctions, coupled V1 cells have RF centers that are closer in the visual field (black dots) compared to the same model in which gap junctions are blocked (gray dots), as shown in Fig. 5(b). In addition, the centers of the RFs belonging to the “mismatched” V1 cells (red triangles) were pulled much closer to the center of their gap-junction-coupled nearest neighbors than those that did not have a gap junction (pink dots).

Using a measure of local visual discontinuity (see Sec. V for details), we show that, after the stage II waves, the V1 network that includes gap junction couplings has a more continuous local map of RF centers (lower local discontinuity) than a network that did not have gap junctions; see first two box plots in Fig. 5(c). This difference is much more striking when considering a network with mismatched RF centers; see middle two box plots in Fig. 5(c). Although the RF centers of the refined mismatched neurons are still significantly different from the normal ones [red significance bar in Fig. 5(c)], the overall effect of including gap junctions is still clearly significant; see final two box plots in Figs. 5(c) and 5(d). These results suggest that gap junctions may play a role in smoothing the retinotopic map along with the stage II retinal waves.

E. Application: Role of stage II waves in seeding orientation preference

Another functional implication of changes in RF structure during stage II waves involves the subsequent refining that may occur during stage III. Stage III waves, i.e., glutamatergic retinal waves, are faster and more spatially localized than stage II waves and, most notably, recruit ON and OFF retinal cells in succession rather than simultaneously as in stage II waves [36] [illustrated in Fig. 6(a)]. It is thought that stage III waves contribute to the formation of orientation preference [16,30], a trait of V1 cells where their ON and OFF LGN inputs are spatially offset and aligned to a particular orientation angle [41].

In this section, we demonstrate the formation of orientation preference during stage III wave and how it depends upon the resulted RF relayed from stage II waves. To achieve this, we manually paused the stage II simulation (with the default parameter set) after 50% or 100% of the full simulation time, which results in a flat and broad disklike RF (50%) and tall and narrow disklike RF (100%), respectively [see Fig. 6(b)]. Note that similarly shaped RFs may also arise from changes in stage II wave speed and width (recall Fig. 2). Then, we replace stage II waves with randomly oriented stage III waves that repeatedly sweeping through LGN cells, activating first ON, and then OFF cells during each wave; the same triplet STDP learning rule is applied. See Sec. V for detailed setup.

Our simulation shows that if the initialized RF relayed from the stage II has a moderated size and strength (denoted as broad) the resulted RF induced by stage III waves will acquire a stable orientation preference [see Fig. 6(d), top row]. In this simulation, although the same random sequence of oriented waves is used to stimulate a batch of 1024 neurons, each neuron still acquires its own orientation preference due to the 95% randomly chosen LGN connection. Orientation

biases are seeded by this randomness, and then fixated by the accumulated weight changes resulted from the waves with corresponding orientations [see Fig. 6(e), top row]. The triplet STDP rule always lead the connection to the extremes (i.e., either the max or min weight allowed), just like the pairwise STDP rule that eventually breaks the balance between LTP and LTD and leads to a bimodal weight distribution [42].

However, if the relayed RF has a smaller RF with stronger connection strengths (denoted as narrow), then the V1 cells can hardly form any orientation preference at the end of stage III, except for ON-dominated iso-oriented RFs [as shown in Fig. 6(d), bottom row]. Due to the stronger connections, stage III waves always lead to more potentiation for the active ON synapses while the OFF synapses near the exiting direction of the wave are depressed, (explained earlier in Sec. II C). As the waves reoccur for each orientation, the potentiated ON synapses now activate the postsynaptic V1 cells more easily, therefore pushing the front of this potentiation toward the incoming direction of the wave (at all orientations), leaving almost all the OFF synapses depressed. Such a process is irreversible since the counteracting depression onto these ON synapses are mediated by the competing OFF synapses, which are weaker in strength and small in number due to the smaller RF, as exemplified in Fig. 6(e), bottom row.

Thus, under the assumption of ON-OFF balance (see a discussion on the topic in Sec. III F 1), our model predicts that when the stage II period persists for too long, or the LGN synapses are pruned too much even with stronger connection strength, the resulted RF will be dominated by the potentiated ON synapses obscuring any initial bias in orientation.

III. DISCUSSION

In this work, we constructed a spike-timing plasticity model of the LGN-V1 pathway during an early developmental phase in which stage II retinal waves induce ON-OFF synchronized spontaneous activity in the LGN. In this model, we explored the effects of changing wave speed and wave width on the synaptic refinement process underlying the development of disklike feed-forward RFs in V1, and contrasted it with a ringlike RF shape when increasing initial connection strength or LTD to LTP ratio. We further clarified the mechanism underlying these effects through the analysis of a reduced rate-based model. With the reduced model, we uncovered the possibility of periodic RF development with the same conditions that lead to ringlike RFs when there is a larger pool of available feed-forward connections, and verified the results in the full spiking model. Adding known gap junctions to the V1 network, we showed that gap junctions in the presence of stage II retinal waves can help recover local retinotopy even when some V1 neurons' initial visual field centers are mismatched with the local LGN cell. Finally, extending the developmental period into stage III waves, we found that V1 neurons with broader circular RFs at the end of stage II were more likely to develop orientation selectivity at eye opening, illustrating the importance of stage II RF refinement. This study clearly identifies the potential mechanisms affecting the feedforward synaptic structure from LGN to V1 during the first two postnatal weeks of development. These results may be valuable in understanding the developmental processes

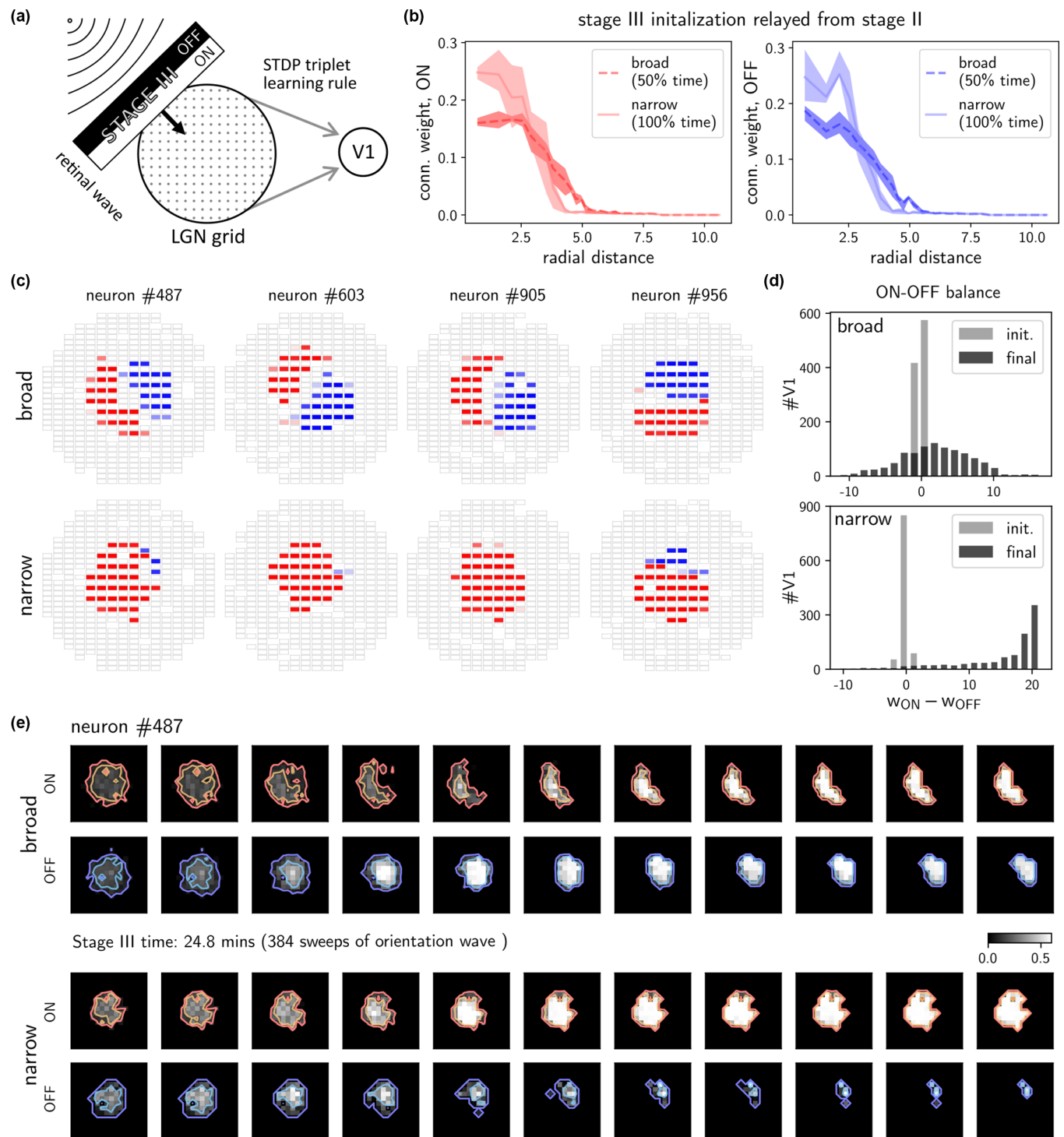


FIG. 6. Influence of refining during stage II waves on formation of orientation selectivity after stage III waves. Randomly picked trial from standard parameter of stage II retinal waves. (a) Schematic of the stage III wave presented to the LGN. (b) Final distribution of connection weights from ON (left, red) and OFF (right, blue) LGN cells after 50% (broad) and 100% (narrow) of the default number of presentations of stage II waves plotted against radial distance. (c) Distribution of difference between ON and OFF LGN cells for all V1 cells. (d) In each box the synaptic weight trajectory is shown for the LGN cells whose synapses potentiated from the initial strength for simulations in which the size of the RF after stage II is small (top row) and large (bottom row). Red curves correspond to ON LGN cells and blue to OFF cells (right). histogram of the difference in connections strength from ON and OFF cells for the small (top) and large (bottom) starting RFs. (e) Time dynamics for the synaptic weights from LGN to a sample V1 cell for small (top two rows) and large (bottom two rows) starting RFs. Lighter colors correspond to stronger weights.

underlying the formation of synaptic pathways in other cortical areas and during different developmental time periods.

A. Synaptic pruning and potentiation during development

Stage II waves are driven by acetylcholine release from starburst amacrine cells [12]. Experiments in the superior colliculus (SC) [5], dorsal LGN [6,7] and primary visual cortex [13] show that removing the acetylcholine receptor abolishes the retinal waves [11,20] and results in much coarser retinotopic maps after development. The much larger axonal spread of RGCs found in these experiments indicates a synaptic pruning process may be missing in the absence of stage II retinal waves. Additionally, a very recent study specifically showed that retinal waves instruct RGC axonal pruning on the periphery and growth in the center [19]. From the perspective of the postsynaptic neuron (V1 cells in our study, SC in experiments), these results are consistent with our model prediction: that stage II waves typically prune synapses in the surround and potentiate those in the center. Although experiments specifically measuring the thalamocortical axonal spread during development are missing, other results showing the appearance of defected retinotopic maps in dLGN and V1 without stage II waves [6,7,13] suggest that synaptic pruning is crucial for retinotopic map refinement in the retina-dLGN-V1 pathway.

B. Wave speed and width

Our model results show that size variability in V1 receptive fields [29,43] may be at least partially explained by the variability in wave speed and width of stage II retinal waves [35]. While previous experiments have achieved simultaneous changes in the speed and width of the retinal waves using cyclic adenosine monophosphate (cAMP) [44], they did not investigate the resulting effects in cortex. A subsequent modeling study of ocular competition in dLGN with cAMP implied a RF refining effect, which is similar to ours, through a Hebbian burst-based learning rule for retinogeniculate synapses [8]. However, the underlying mechanisms, are obscured by the simultaneous consideration of changes in width and speed (determined by cAMP), which further influences the level of activation and the time window for learning. Therefore, to verify the results (or predictions) in our model experimentally, well-controlled manipulation of wave properties and synaptic dynamics are needed, such as spatiotemporally controlled release of acetylcholine in the retina.

C. Learning rules

Spike-time-dependent learning rules and its variants have been found in various adult sensory cortices [45–47] and during development [48] ever since its first discovery in the hippocampus [49]. In the visual cortex specifically, the triplet learning rule [33] that we applied here has been shown to fit well with experimental results [47]. In addition, the modeling study on a burst-based learning rule found in the retinogeniculate synapses during stage II waves, suggests that modified STDP rules with longer time windows and involving multiple spikes, like the triplet rule, can better explain the observed plasticity over naive STDP rules [8] (but see [50]). This work observed less RF refinement with cAMP application, which

increases the size, speed and frequency of the waves [44]. Those results are closely related to our observation on RF refinement, albeit ours is in V1, the next stage in the early visual pathway. With more carefully controlled numerical experiments, here we demonstrated that a triplet STDP learning rule, reducible to BCM when assumed Poisson [34,37], can refine single neuron RF as well as the local retinotopy, and clarified how the triplet learning rule leads to the various effects with the reduced model analysis.

We also introduced a type of homeostasis that restricts the sum of incoming synaptic weights to one neuron to stay close to its initial value on the timescale of interwave interval, similar to synaptic scaling [51] (for details see Sec. V). This type of homeostasis encourages competition between synapses, effectively reducing the number of waves needed for observing the RF refinement effect (see Fig. 10 in Appendix A).

D. Periodic connectivity

Periodicity in connectivity may arise both within cortical areas or across brain regions. For example, Layer (L)2/3 neurons in primary visual cortices that have orientation columns horizontally connect to neurons across several hyper-columns with similar orientation preference or the same ocularity in a periodic fashion through long-range horizontal connections (LHCs) [52,53]. Additionally, grid cells in the medial entorhinal cortex (mEC) have hexagonal RF patterns that map out a subject's movement in a two-dimensional space [21], the same structure has now been found in rat S1 and V2 area [54,55]. The mechanisms underlying these periodic connections remain elusive. Some experiments from L2/3 V1 neurons suggest that LHCs are independent of retinal activity [56–58], while other studies in primary auditory cortex demonstrate an activity-dependent re-wiring of axons [59]. In terms of the periodic RF in mEC, studies suggest that the hexagonal RF may result from periodic connections set up during development [60,61]; however, periodic spontaneous activity has also been found in L3 of the developing entorhinal cortex [23], suggesting a different mechanism. The emergence of periodicity in the RFs in our model relies only on the interaction between an STDP learning rule and retinal wave properties, supporting the case in which periodic connections arise in an activity-dependent way.

A theoretical mathematical model investigating the interaction between one-dimensional traveling waves and pattern formation found that activity-dependent plasticity can inform periodic RFs [50]. However, it is hard to decouple the frequency of multiple incoming waves and the resulting periodicity of the RF in their model. And they focus more on demonstrating the effects of different STDP rules and noise in the presence of multiple waves than a mechanistic explanation. In contrast, our model uses biologically relevant connection profiles, spontaneous retinal wave parameters, and a triplet STDP rule verified for the visual cortex to extract the mechanisms underlying the formation of periodic RFs from decoupled waves of spontaneous activity.

E. Retinotopic refinement and gap junctions

Independent of retinal waves, spontaneous activity arises in the cortex during development [1,4]. A recent simplified

one-dimensional theoretical modeling work of retinotopic refinement [62] characterizes the interaction of local (L, of thalamic origin) and global (H, endogenous to cortex) events based on experimental data [63] and a Hebbian learning rule; the model postulates that the L-events carry retinotopic information while H-events regulate cortical connectivity. Previous theoretical modeling studies (as reviewed by Swindale 1996 [64]) have shown that the retinotopy can emerge with plastic chemical synapses between output neurons. However, local recurrent chemical synapses are mostly absent in the mouse cortex during stage II waves [25] and cortical activity is mediated through gap junctions and silent synapses among subplate neurons [26]. Therefore, we postulate (and verified in our model) that stage II retinal waves are able to refine the local retinotopic map by encouraging nearby gap-junction connected neurons to share similar thalamocortical input, even when a few neurons have initially mismatched connections.

1. Gap junctions in V1 during development

In a more detailed biological picture, thalamocortical synapses are first formed in the subplate and connections to L4 neurons are then guided by these subplate neurons through correlated spiking activity [26,65,66]. During the developmental period coincident with stage II retinal waves (P0-P7), excitatory neurons derived from the same progenitor cell are coupled by gap junctions rather than chemical synapses [25,26,67]. While the influential role of gap junctions during synaptogenesis in the next developmental stage (when gap junctions are effectively gone during stage III waves) has been deeply explored by experiments and models [38–40], little is known about its role during stage II retinal waves.

A very recent experiment on the superior colliculus of the adult mouse has shown that postsynaptic cells have an imprecise retinotopy despite a precise RGC axonal tiling [68], consistent with a study on local retinotopy of mouse V1 [69]. These results support the existence of spatially biased random sampling of incoming connections on the single-neuron level in the developing cortex. In our work, we propose that a gap-junction-coupled network, which promotes correlated activity in the nearest neighbors, could help mitigate such bias in single neurons, without which the resulting retinotopy could be even less precise.

Note that our model effectively considers subplate neurons and immature L4 neurons as the same population, some of which is coupled by gap junctions. In reality, vertically aligned sister cells are the only neurons in this time period that have been shown to have gap junction connections [40]. We have shown that our result—that gap junction connections between neurons during stage II retinal waves serve to refine and correct the local topographic map—still holds across different percentages of gap-junction coupling (see Fig. 9 in Appendix A), even when that coupling is reduced to only one or two connections per neuron (as in the developing cortex).

F. Orientation preference and stage III retinal waves

In contrast to ON-OFF synchronous stage II retinal waves, stage III waves are now known to be ON-OFF asynchronous in the RGCs and LGN cells, where the OFF cells are activated right after the neighboring ON cells within each wave [36].

Experimental evidence suggests that orientation selectivity is formed by spontaneous activity before eye-opening (around P14), since broad tuning of V1 neurons can be observed at eye-opening [38,70]. The asynchronous activation during stage III waves makes it a perfect candidate for early development of orientation preference in the primary visual cortex [16,30], and our preliminary modeling of the stage III waves with triplet STDP rule indeed supports this hypothesis. As the STDP triplet rule can be reduced to a BCM-like rule in a rate-model [34], our result is also consistent with an earlier study that shows the segregation of ON-OFF subregion is dependent on BCM rule operating in a linear region using natural images [71].

In addition, our model shows that the condition for the emergence of orientation selectivity also depends on the feed-forward connection profile (or RF) initialized during stage II retinal waves. A narrow bell-shaped RF with little room for nonoverlapped ON and OFF LGN connections can make it very hard to develop orientation selectivity, demonstrating the importance of the width and speed of individual stage II waves for proper RF development in later stages. It is worth noting that our model explores the relative timing of stage II and stage III waves and stimulates the LGN with stage III waves until a stable orientation preference emerges. It may be the case that eye opening occurs earlier in our simulation time and the RFs at that time would have much less orientation selectivity, as indicated by experiments [38]. Nevertheless, these results may help explain the high variability in the selectivity of V1 neurons.

1. ON-OFF subregion imbalance

Our model would predict that neurons with narrow RF at the end of stage II will eventually acquire a ON-dominated RF at the end of stage III. However, this prediction relies on the assumption that the ON and OFF inputs are balanced in each V1 cell. We note that extensive experimental studies have shown that OFF RGC and OFF LGN cells are more active than ON cells [72,73], due to their differences in the numbers and anatomical pathways [72,74]. For simplicity, we drop this asymmetry between the ON and OFF synapses in our model, whereas in reality, the intrinsic bias toward OFF cells may give more room for the emergence of orientation preference in the visual system. Note that the triplet learning rule biases toward enhancing the weights for whichever cells are activated first in the wave [ON cells in the case for mouse [36], and see Fig. 6(c), bottom panels]. This suggests that our model can tolerate an extended range of parameters in the stage II retinal waves [35].

IV. CONCLUSION

In summary, we have used two levels of mathematical modeling to identify and demonstrate potential mechanisms underlying the refined RFs of V1 cells as a function of different characteristics of the retinal waves. Our model predicts that periodic RFs may arise due to changes in long-term depression or initial connection strengths, gap junctions in V1 during the first postnatal week may play a role in increasing local visual continuity, and variability in the shape and size of the RF at the end of stage II influences the emergence of

orientation preference of V1 cells during stage III retinal waves (before eye-opening). The effects introduced in our model are robust under a wide range of parameters, therefore our model are not only useful in understanding activity-dependent development of V1 RFs, but can also be used for predictions across different sensory cortices during development, given parameters under different settings.

V. MODELS AND METHODS

A. Setup of the developmental model

This model describes feed-forward learning of V1 receptive fields in the developmental stage, driven by successive random spontaneous stage II waves originating in RGCs and relayed by LGN cells to V1. We model 1024 local V1 neurons that share the same visual field with a standard 16×16 grid of LGN cells. Each V1 neuron is postsynaptic to 80% of the LGN cells, randomly chosen, in the grid initially. The network we consider is feed-forward only, i.e., V1 lacks intracortical chemical synapses during this stage [75]. They do have gap-junction connections, which we add to the model later [24].

1. Stage II waves

Many more retinogeniculate synapses (around 30) per thalamocortical projecting cell exist in dLGN during development than a mature one (around 10, but functionally 1–3) [76]. This indicates an abundance in the local activation of dLGN from retinal waves. Nevertheless, given the fact that the receptive fields of dLGN neurons' are still largely center-surround like the RGCs despite the higher convergence ratios in reality [77]. In addition, we'd like to keep our model simple while being reasonably realistic. Thus, we assume that LGN activity is directly proportional to the amplitude of the retinal waves, i.e., we consider a simplified 1:1 linear mapping from retina to LGN. Focusing on a local population of LGN cells, we model the effect of the retinal waves as drifting sinusoidal bars coming from random directions, mimicking parallel wave fronts.

This is a valid assumption when the distance from the origin of these waves is large compared to the size of the local population. According to the thorough investigation of retinal wave characteristics by Maccione *et al.*, in 2014 [35], stage II retinal waves travel around $120 \mu\text{m/s}$. In our model, this translates to around 4 degrees/s or 3.2 units of distance between the nearest LGN cells per second. Stage II waves have directional bias dependent on different quadrants but are still largely random [14] so we only apply an uniform distribution of wave direction in our model. Assuming balanced ON-OFF coverage of a local visual field, for simplicity, we construct two grids with overlapping vertices of the same size for ON and OFF LGN cells. The standard stage II wave width is 10 degrees, or 8 in the unit of interval between LGN cells.

2. Stage III waves

Our modeling of the stage III waves follows the simplicity of stage II waves as oriented drifting bars activating the LGN cells, but with OFF cell activation lagging behind the ON cells in each sweep instead of overlapping with each other [see Fig. 6(a)], though the total width of the wave is still set

TABLE I. Default parameter set for the retinal waves.

	Stage II waves	Stage III waves
Wave speed (deg/s)	4	5
Wave width (deg)	10	5
Exposed LGN cells (diameter)	16	16
Interwave interval (s)	6	3
On-Off segregation	False	True

as 10 degrees. However to reflect other difference in the experimental observations, the speed of stage III wave is faster, 5 degree/s, and each sweep is repeated three times [35,36]. Also see Table I for a table of the standard parameters.

3. Models of V1 neurons and LGN cells

In the full simulations, V1 neurons are modeled by conductance-based adaptive exponential integrate-and-fire models [32]:

$$\begin{aligned}
 C_m \frac{dV}{dt} &= -g_L V_L + g_L \exp \left[-\frac{V - V_T}{\Delta_T} \right] + g_E (V - V_E) \\
 &\quad - Q + g_{\text{gap}} \sum_i (V - V_i), \\
 \tau_d \frac{dg_E}{dt} &= -g_E + h_E, \\
 \tau_r \frac{dh_E}{dt} &= -h_E + \sum_j w_j \sum_k \delta(t - t_k^{(j)}), \\
 \tau_Q \frac{dQ}{dt} &= a(V - V_L) - Q,
 \end{aligned} \tag{4}$$

where V is the membrane potential and Q is the adaptive current. When V surpasses the firing threshold, $V_T = -50 \text{ mV}$, it exponentially increases until it reaches 0 mV , whereupon it gets reset to $V_r = -65 \text{ mV}$. At the time of reset, Q is increased by $b = 2.5 \text{ pA}$. The synaptic conductance, g_E , comes from the feed-forward spike train $t_k^{(j)}$, containing k many spikes from connected LGN cells indexed by j . For those simulations that contain gap junctions, $g_{\text{gap}} = 0.75 \text{ nS}$, otherwise it is set to 0. Also see Table II.

LGN cells are modeled by a linear-nonlinear-Poisson (LNP) process with parameters extracted from experiments on mouse LGN [78]. When modeling the developmental process in the absence of visual experience (during first two postnatal weeks when eyes are closed), the linear convolution in the LNP is directly replaced by a normalized retinal wave amplitude, $I \in [0, 1]$. The nonlinear (sigmoid) gain function for the LGN firing rate, $\hat{r}$, takes the form of a generalized logistic function:

$$\hat{r}(I) = A + \frac{B}{1 + \exp[K(c_{50} - I)]}, \tag{5}$$

where parameters A and B can be found by solving $\hat{r}(0) = 3 \text{ Hz}$ (spontaneous firing rate) and $\hat{r}(1) = 60 \text{ Hz}$ (maximum firing rate). The strongest gain occurs for $K = 3$, and $c_{50} = 0.25$ marks the input strength that results in half of the maximum output. From this process, a time-dependent Poisson spike train is generated.

TABLE II. Parameters for synaptic plasticity.

Parameter	A^+	r_{LTD}	τ_+	τ_-	τ_{slow}	r_0	τ	τ_h
Value	0.003	0.3	17 ms	34 ms	114 ms	1000 ms	6 Hz	2500 ms

4. Learning rule of the LGN-V1 connections

We apply the STDP triplet rule as the learning rule in our model, which is suitable for describing plasticity in the visual cortex [33]. By employing one presynaptic tracer variable ζ_- and two postsynaptic tracer variables ζ_+ and ζ_{slow} , the weight change Δw , can be expressed as the following:

$$\Delta w = \begin{cases} A^+ \zeta_+(t) \zeta_{\text{slow}}(t - \epsilon), & \text{if } t = t_{\text{post}}, \\ A^-(t) \zeta_-(t), & \text{if } t = t_{\text{pre}}, \\ 0, & \text{otherwise.} \end{cases} \quad (6)$$

The synaptic tracer variables ζ_k , ($k = -, +, \text{slow}$) increment by 1 at the time of a spike in their corresponding presynaptic neuron t_{pre} or postsynaptic neuron t_{post} , and otherwise decay exponentially. Additionally, we add a homeostatic fast-rate detector introduced by Zenke *et al.*, 2013 [34] to balance LTP and LTD by setting a target firing rate r_0 for V1 neurons as follows:

$$A^- = \frac{\tau_+ \tau_{\text{slow}} [\bar{r}(t)]^2}{\tau_- r_0} A^+, \quad (7)$$

where the average firing rate of the V1 neuron is $\bar{r} = \frac{1}{\tau} \sum_i \exp(-\frac{t-t_i}{\tau})$. Parameters for the STDP rule can be found in Table III. See also Fig. 12, in Appendix A for the effects when the timescales are varied relative to the cycle of waves.

5. Optional homeostasis

In the full simulations we applied an optional slow homeostasis onto the synapses to enhance the competition between LTP and LTD and shorten the time for simulation apart from the homeostatic fast-rate detector. In the mechanistic analysis of reduced model, we intentionally turn off this homeostatic change to clearly demonstrate the effects, except Figs. 3(f) and 3(g) where it is retained to compare with the full simulation. The slow homeostasis gradually relax the total synaptic weight of a V1 neuron back to its initial value on the timescale of 2.5 s, comparable with the interval between two stage II waves. The dynamics of the total synaptic weight follows a simple ODE:

$$\tau_h \frac{ds}{dt} = -(s - s_0), \quad (8)$$

where s is the total synaptic weight for each neuron, and s_0 is its initialized value. For each time step, we add the total change in synaptic weights $\sum_i \Delta w_i$ from LTD and LTP to s and evolve the dynamics to get the total change Δs due to this homeostasis, and then redistribute the change equally to each individual synapse. For detailed comparison of this homeostatic effect, see Fig. 10 in Appendix A.

B. Receptive field analysis

In this work, the difference between ON and OFF LGN cells are largely ignored because ON-OFF segregation is not present during stage II retinal waves [9]. Therefore, we derive metrics that characterize the size and overall shape of the feed-forward connections by which we define the RF, ignoring ON and OFF designations. The metric we call “weighted radius” is used to quantify the spread of the disklike RF for each V1 neuron and is defined as follows:

$$R = \sum_{k,l} w_{kl} \left| \bar{x}_{kl} - \frac{1}{N} \sum_{i,j} w_{ij} \bar{x}_{ij} \right|, \quad (9)$$

where w_{ij} is the weight of the connection from the LGN cell at location $\bar{x}_{ij} = (i, j)$ in the visual field. To quantify the actual amount of connections with enhanced strength in the same unit as weighted radius, we take the square root of the area of increased connection strengths (compared to their initial value) and divide it by two, called the characteristic length of the RF. The characteristic length for a one-dimensional RF, as in the reduced model, is defined as the width of the weight distribution at the initial connection strength [i.e., the difference between the red dots in Figs. 3(f) and 3(g)] divided by two.

In addition to the two metrics above, we construct the radial distribution of connection weights by averaging connection weights along the radial axis to give the most intuitive description of RF regardless of the disklike or ringlike structure.

To address the retinotopic continuity of the visual field coverage by the RFs of neighboring neurons in the gap junction study, we define a local discontinuity metric. Assuming the V1 neurons are densely packed in a 2D cortical surface (Poisson disk sampling algorithm is used to generate the physical positions), we quantify “local” by the optimal distance between those densely packed neurons, h . Then we can obtain

TABLE III. Parameters for exponential integrate-and-fire neuron.

Voltage	Values (mV)	Timescale	Values (ms)	Conductance	Values (nS)
V_L	-65.0	C_m/g_L	20	a	0.2
V_T	-50.0	τ_d	3	g_{gap}^a	0.75
V_θ	-20.0	τ_r	1		
V_E	0	τ_Q	15		
Δ_T	1.5				

^aNote that g_{gap} equals zero except in Sec. IID.

a metric that describes the smoothness in the local retinotopic mapping, D for each neuron i as follows:

$$D_i = \left| \vec{p}_i - \frac{1}{m} \sum_{j \in \{|\vec{q}_j - \vec{q}_i| < d\}} \vec{p}_j \right| / h, \quad (10)$$

where $\vec{p}$ denotes the position vector for each neuron's RF center in the visual field and $\vec{q}$ denotes the physical location for each neuron in the cortical surface. The "local" visual center is summed over the m neighboring neurons within distance h . Values above 1 indicate considerable discontinuity at the corresponding visual position.

ACKNOWLEDGMENTS

The authors want to acknowledge Prof. David W. McLaughlin for frequent discussions on the topic which improved the quality of this work. This work is funded by the Science and Technology Innovation 2030—Brain Science and Brain-Inspired Intelligence Project No. 2021ZD0201300 (W.P.D.); the National Natural Science Foundation of China Grant No. 12201125 (W.P.D.); the Science and Tech-

nology Committee of Shanghai Municipality Grant No. 22YF1403300 (W.P.D.); and an Institutional Development Award (IDeA) from the National Institute of General Medical Sciences of the National Institutes of Health under Grant No. P20GM103449 (J.C.).

DATA AVAILABILITY

The data that support the findings of this article are openly available [79,80]; embargo periods may apply.

APPENDIX A: SUPPLEMENTAL TABLES AND FIGURES

The phase diagram over time and mechanistic details of the contributing components in the reduced models are shown in Fig. 7. Consistent pruning effect of increasing wave width while keeping the same initial connection weights are shown in Fig. 8. The effect of visual continuity of stage II retinal waves on electrically-coupled network under different connection probability are shown in Fig. 9. The detailed effect of homeostasis and learning rate on the progress of RF development is shown in Fig. 10. The pruning effect of stage II retinal wave is stable across a wide range of sparsity of LGN to V1

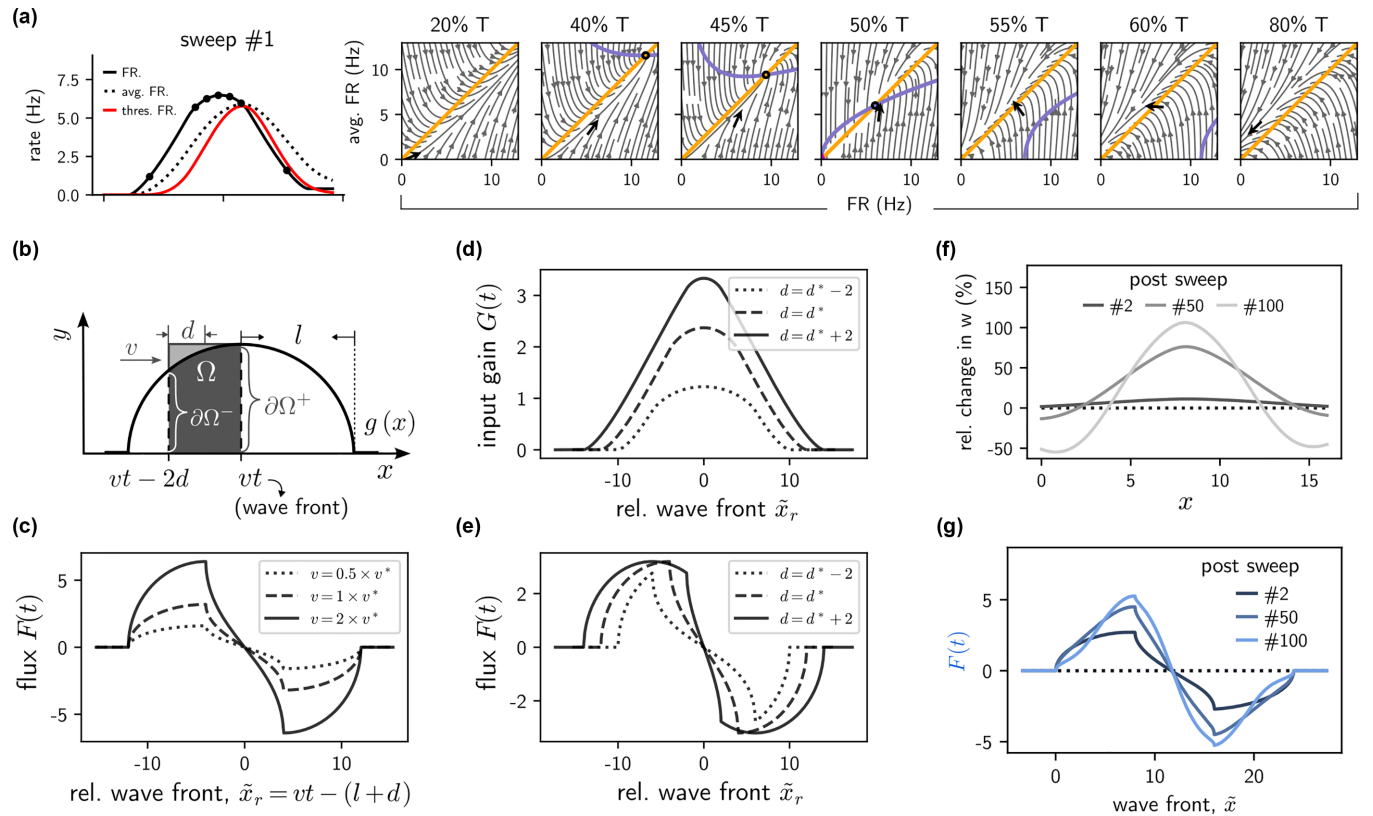


FIG. 7. Mechanistic details. (a) Temporal evolution of the reduced model during the first wave. The left panel shows V1 firing rate (solid) and its average (dotted) in black, the threshold firing rate, r_{th} , in red. The input gain, $G(t)$, in dotted green is plotted against the right y axis. Phase portraits of $r(t)$ vs. $\bar{r}(t)$ at seven time points during the first wave corresponding to solid circles in the left panel. The $\bar{r}$ -nullclines are in orange and the r -nullclines are in purple. The black arrows mark the vectors $(r, \bar{r})$ at the corresponding time points in the first panel. (b) Schematic illustrating a wave with speed v , half-width d , and boundary $\partial\Omega$ traveling through a one-dimensional continuous LGN space with half-length l that goes with Eq. (2). (c) The flux $F(t)$ plotted against relative wave front over time (for alignment at the center) across wave speeds. (d), (e) Input gain $G(t)$ and flux $F(t)$, respectively, plotted against relative wave front across wave widths. (f) Spatial distribution of LGN connection weights across different numbers of intermediate wave sweeps [data of Fig. 3(e)]. (g) The flux $F(t)$ across different weight distributions given in panel (f).

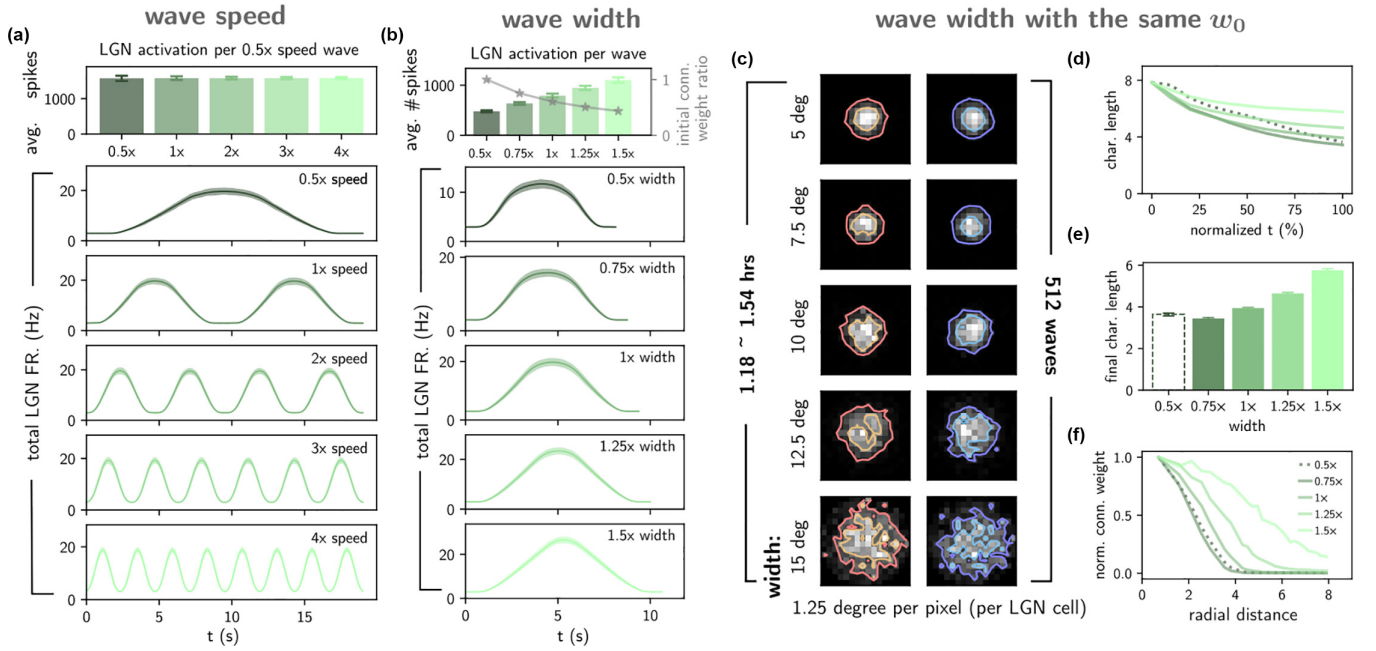


FIG. 8. The effect of different wave widths but with the same connection strength. (a) The LGN activation per 0.5x speed wave, measured as the number of LGN spikes per normalized time range, remains the same (T test p value = 0.415) across wave speeds (speed is colored with increasingly saturated green). Lower panels show the firing rate of all connected LGN cells, averaged over retinal waves. This setup that pairs with Fig. 2(a). (b) Same as panel (a) but the top panel shows the total LGN spikes per wave across wave widths with the initial LGN-V1 connection strength, normalized to the smallest wave width (gray stars). This setup pairs with Fig. 2(h). (c) Two population-averaged LGN-V1 connection profile at the end of stage II with different wave width from a sample trial, labeled with half wave-width in degrees. (d) Temporal evolution of population averaged characteristic length (CL) over 10 trials with different wave widths. (e) Errorbars of CL collected from 10 trials over different wave width at the end of stage II. (f) Radial distribution of connection strength averaged from the 10 trials over different wave widths at the end of stage II. In this setup, we use $w = 0.3$ for all wave widths, resulting in an insufficient V1 activation for the wave with 5-degree half-width ($0.5\times$ the standard, denoted by dotted lines/contours) such that it does not follow the trend as in the main text. A stronger connection strength will lead to ring structure in the 15-degree wave.

connections, as shown in Fig. 11. Varying the time constant beyond interwave interval makes it impossible to obtain any stable RF as shown in Fig. 12.

APPENDIX B: DERIVATION OF THE REDUCED RATE-BASED MODEL FOR STAGE II SIMULATION

For further analysis on the process of stage II wave influence on the thalamocortical connections, we approximate the full simulation of Eq. (6) with a rate-based model using a similar set of steps as has been done previously [34,37]. First, we reduce the two-dimensional, discrete pool of LGN cells to a one-dimensional, continuous distribution of LGN cells along the x dimension. We assume that increasing x is in the direction of an incoming wave [recall Fig. 3(a)]. Then, assuming a synapse exists as each x location in the pool, we can write the following equation describing the change in the synaptic weight $\hat{w}(x, t)$ at time t and location x under the triplet rule [37]:

$$\frac{\partial \hat{w}(x, t)}{\partial t} = A^+ z_+(x, t) z_{\text{slow}}(t - \epsilon) S^{\text{V1}}(t) - A^-(t) z_-(t) S^{\text{LGN}}(x, t), \quad (\text{B1})$$

where $S^{\text{LGN}}(x, t) = \sum_i \delta(t - t_i^x)$ is the spike train with spike times t_i^x for the LGN cell at location x . The spike train for the

V1 neuron, $S^{\text{V1}}(t)$, is similarly defined. The presynaptic tracer variable $z_+(x, t)$ satisfies the differential equation

$$\frac{\partial z_+(x, t)}{\partial t} = -\frac{z_+(x, t)}{\tau_+} + S^{\text{LGN}}(x, t). \quad (\text{B2})$$

The postsynaptic tracer variables, $z_-(t)$ and $z_{\text{slow}}(t)$, are similarly defined. The LTP and LTD learning rates, A^+ and A^- , respectively, are related by a fast-rate detector in the following equation [34]

$$A^-(t) = \frac{\tau_+ \tau_{\text{slow}} [\bar{r}(t)]^2}{\tau_- r_0} A^+, \quad (\text{B3})$$

where r_0 is the target firing rate and $\bar{r}(t)$ is the average firing rate of the V1 cell, defined by the following differential equation:

$$\tau \frac{d\bar{r}}{dt} = -\bar{r} + S^{\text{V1}}(t). \quad (\text{B4})$$

We take the expectation over many spike trains and arrive at a rate-based approximation for V1 firing

$$\tau \frac{d\bar{r}(t)}{dt} = -\bar{r} + r. \quad (\text{B5})$$

Then, plug in for $A^-(t)$ using Eq. (B3) and take expectation over many pairs of LGN and V1 spike trains we write

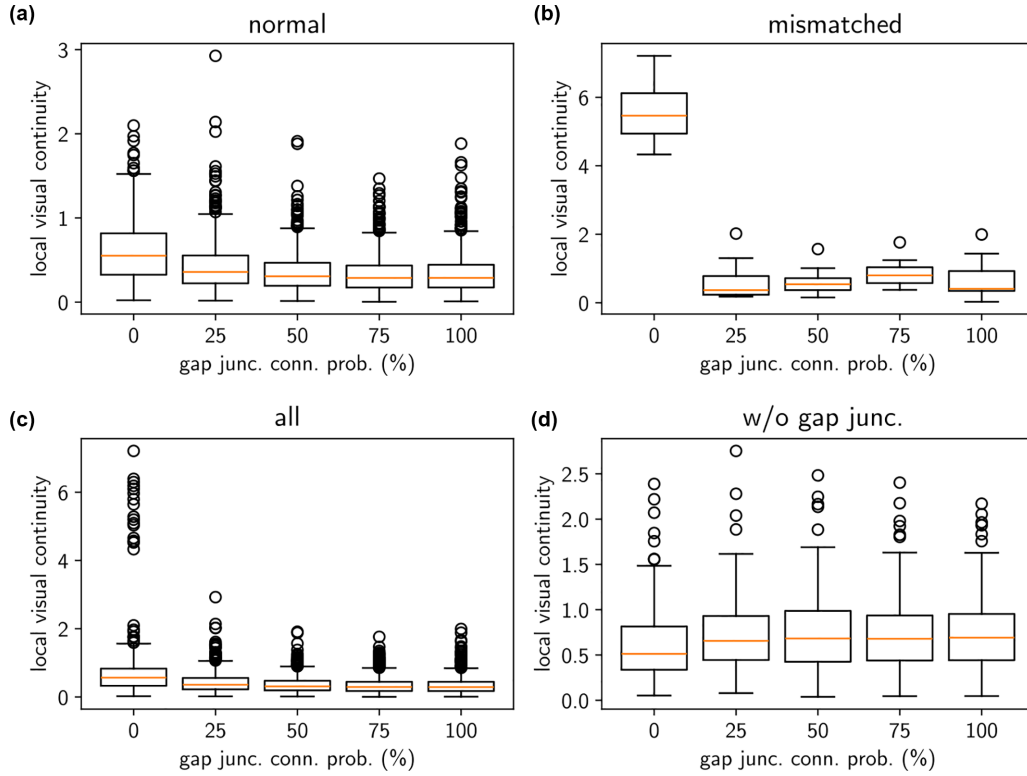


FIG. 9. Visual continuity for different gap junctions probability. (a) Local visual continuity calculated for a gap-junction-coupled network without mismatched LGN connections. (b), (c), (d) Same metric as in panel (a), but calculated for mismatched neurons, all neurons, and neurons that not eligible to be gap-junction-coupled in the mismatched setup, respectively. Note that percentages are for the gap-junction-eligible neurons only; i.e., 100% coupling probability leaves 25% of neurons that do not couple with gap junctions.

down the rate-based approximation for the change in synaptic weight as

$$\begin{aligned} \left\langle \frac{\partial \hat{w}}{\partial t} \right\rangle &= A^+ \tau_+ r^{\text{LGN}}(x, t) \tau_{\text{slow}} [r(t)]^2 \\ &\quad - \left(\frac{\tau_+ \tau_{\text{slow}} [\bar{r}(t)]^2}{\tau_- r_0} A^+ \right) \tau_- r(t) r^{\text{LGN}}(x, t) \\ &= A^+ \tau_+ \tau_{\text{slow}} r^{\text{LGN}}(x, t) r(t) \left[r(t) - \frac{[\bar{r}(t)]^2}{r_0} \right], \end{aligned}$$

where $r(t)$ is the firing rate of the V1 cell at time t and $r^{\text{LGN}}(x, t)$ is the firing rate of the continuous LGN at location x and time t . We define the rate-based approximation of connection weight in our model, $w(x, t)$ to be the integral of the approximated change in average, $\langle \partial \hat{w} / \partial t \rangle$, from $t = 0$ plus some initial value $w(x, 0)$. The weight dynamics is BCM-like as it involves a threshold which upon crossing changes the sign of weight change.

Now, we define the firing rate of the V1 cell as the integration of weighted LGN firing over the area activated by the retinal wave, as follows:

$$r(t) = g_0 \hat{r} \int_{\max(0, vt-2d)}^{\min(2l, vt)} g(x) w(x, t) dx, \quad (\text{B6})$$

where g_0 is the intrinsic gain of the V1 neuron, l is the radius of the pool of LGN cells, and d is the half-width of the wave in units of number of LGN cells [recall Fig. 3(a), bottom

panel]. The function $g(x) = 2\sqrt{l^2 - (x-l)^2}$ gives the length of a chord in the circular pool of LGN synapses and represents the number of LGN synapses in the vertical direction that are activated at each location of x assuming that they all respond identically to the wave.

Next, we take the time derivative of Eq. (B6). To make the computation simpler, let the upper limit of integration be $u(t) = \min(2l, vt)$, the bottom limit be $b(t) = \max(0, vt - 2d)$ and suppose $Q(x, t) = \int g(x) w(x, t) dx$ is the anti-derivative. Then, we can compute the derivative of the integral on the right-hand side of Eq. (B6) as follows:

$$\frac{d}{dt} \left(\int_{b(t)}^{u(t)} g(x) w(x, t) dx \right) = \frac{d}{dt} [Q(u(t), t) - Q(b(t), t)]. \quad (\text{B7})$$

By the chain rule, the time derivative of $Q(x(t), t)$ gives

$$\begin{aligned} \frac{dQ(x(t), t)}{dt} &= \frac{\partial Q(x, t)}{\partial x} \frac{dx}{dt} + \frac{\partial Q(x, t)}{\partial t} \\ &= g(x) w(x, t) \frac{dx}{dt} + \int g(x) \frac{\partial w(x, t)}{\partial t} dx. \end{aligned}$$

We again consider a mean-field approximation of the firing rate by taking the expectation over many spikes to write

$$\frac{dr}{dt} = g_0 \hat{r} \left[g(x) w(x, t) \frac{dx}{dt} \Big|_{x=b(t)}^{x=u(t)} + \frac{\partial w}{\partial t} \int_{b(t)}^{u(t)} g(x) dx \right], \quad (\text{B8})$$

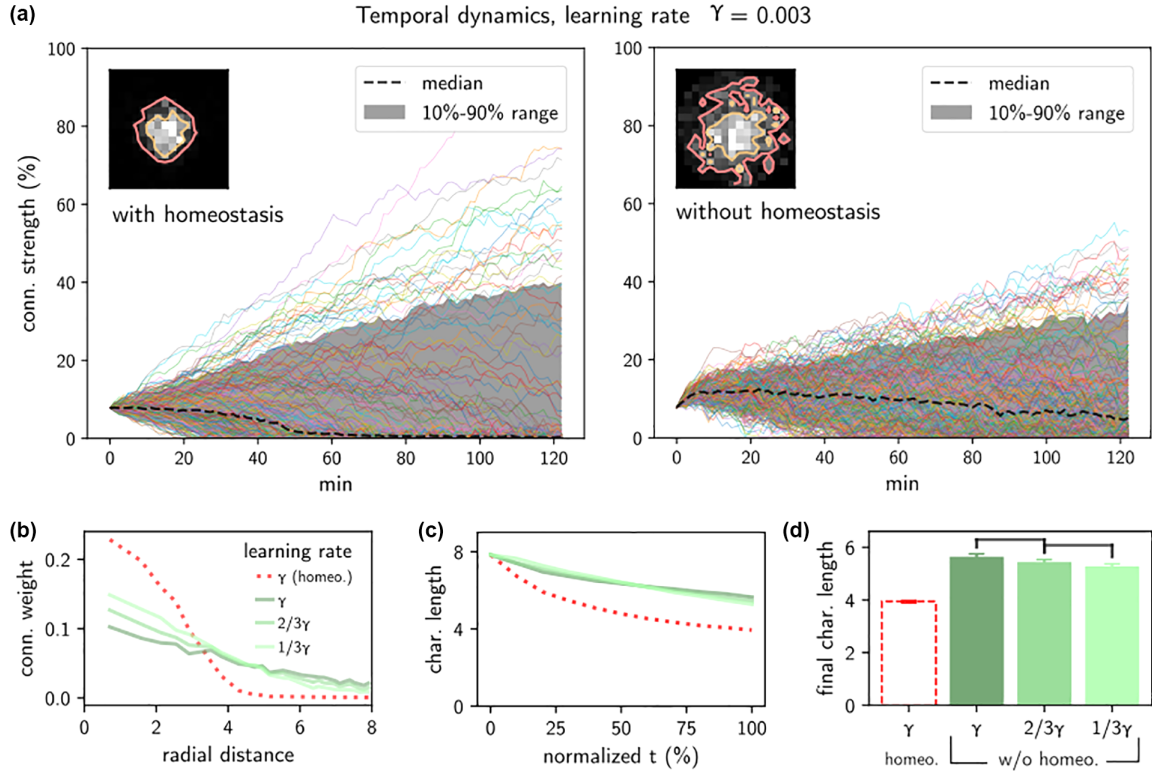


FIG. 10. Effect of global homeostasis and learning rate on synaptic pruning. (a) Temporal dynamics of the connection strengths from LGN to one sample V1 neuron for the full simulation with (left panel) vs without (right panel) global homeostasis, with standard learning rate $\gamma = 0.003$. The insets illustrate the final connection profile. (b) Radial distribution of connection strengths averaged from 5 trials for 3 different learning rates without global homeostasis (shades of green). Note that the simulation time is extended correspondingly (e.g., three times for $1/3$ learning rate). Standard learning rate with global homeostasis is plotted in red for comparison. (c) Same as panel (b), but for the characteristic length over time. (d) Bar plot of the final characteristic length in b; T-test performed between different learning rates without global homeostasis (T value = 0.0195 between γ and $2/3\gamma$; T value = 0.0247 between $2/3\gamma$ and $1/3\gamma$). Global Homeostasis enhances the competition between synapses, resulting in a faster pruning and a stronger nonlinear structure in the spatial profile.

noting the assumption that the firing rate of the LGN neural field is the constant $r^{\text{LGN}}(x, t) = \hat{r}$ when activated by a retinal wave (inside the bounds of integration) and vanishes otherwise allows us to pull the $\partial w / \partial t$ term outside the integral.

We make a final simplifying assumption to reduce this model to a two-dimensional system of $r(t)$ and $\bar{r}(t)$. To do this, we assume that the learning rate, A^+ , is small enough such that the weight, $w(x, t)$, during a wave can be well-approximated by the weight at the start of the wave (i.e., for the i th wave we have $w_i(x) := w(x, t = iT)$ where T denotes the duration of the wave). We update the weight between each wave using Eq. (B6) and integrating over the time that the wave passes through each x -location:

$$w_i(x) = w_{i-1}(x) + \tau_+ \tau_3 A^+ \hat{r} \int_{x/v}^{(x+2d)/v} r \left(r - \frac{\bar{r}}{r_0} \bar{r} \right) dt. \quad (\text{B9})$$

Substituting in the expression for $\partial w / \partial t$ found in Eq. (B6) into Eq. (B8), we arrive at the equation for the change in firing

rate of the V1 cell for the i th wave as follows:

$$\begin{aligned} \frac{dr}{dt} &= r(t) \left(r(t) - \frac{\bar{r}(t)}{r_0} \bar{r}(t) \right) G(t) + F(t), \\ G(t) &= g_0 A^+ \tau_+ \tau_{\text{slow}} \hat{r}^2 \int_{b(t)}^{u(t)} g(x) dx, \\ F(t) &\approx g_0 \hat{r} g(x) w_i(x) \frac{dx}{dt} \bigg|_{b(t)}^{u(t)}. \end{aligned} \quad (\text{B10})$$

Together with Eqs. (B5) and (B9), we have arrived at the two-dimensional reduced model used in the text. We note that the 3D system does not qualitatively differ from the 2D system by comparing the simulations' results.

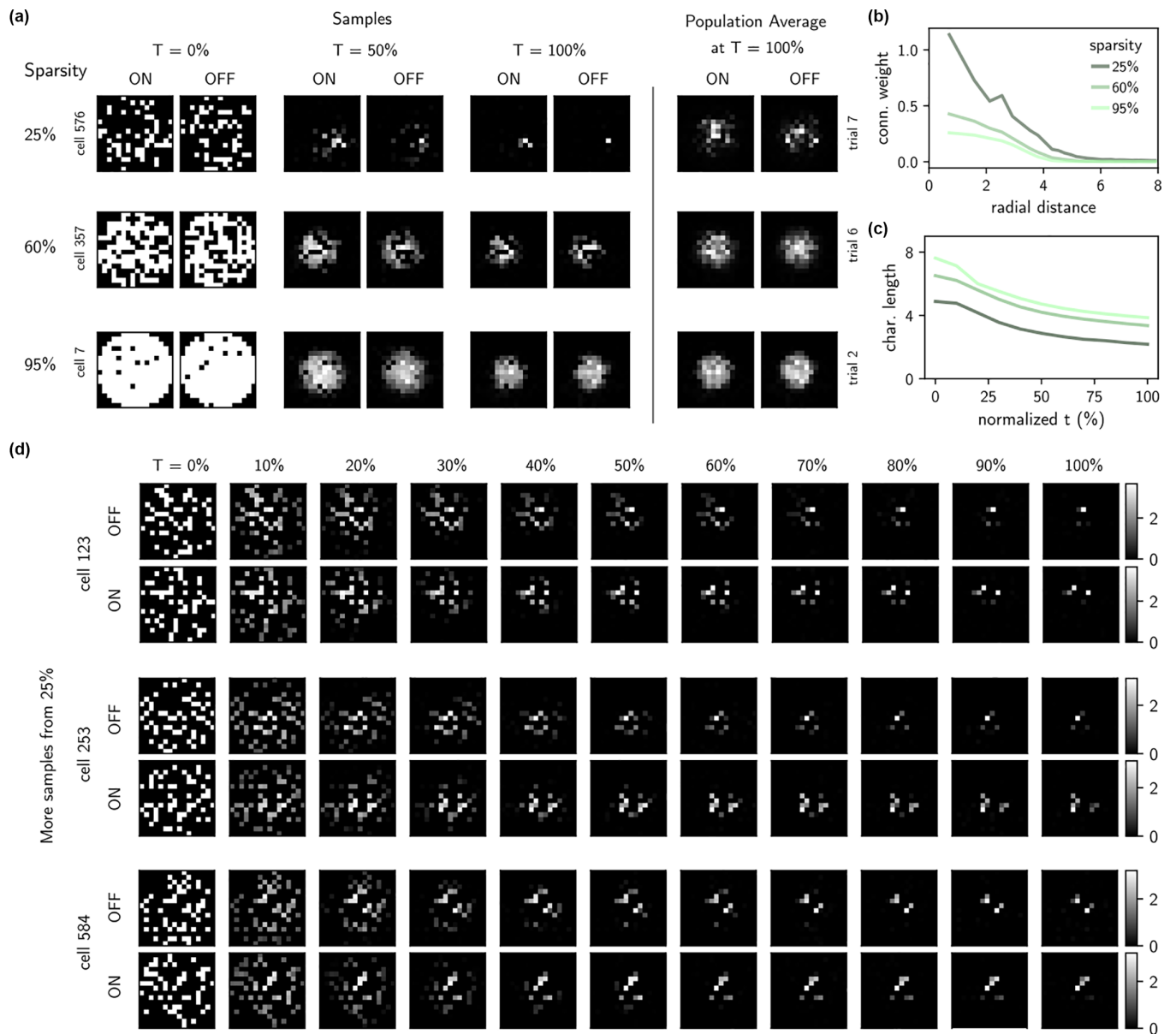


FIG. 11. Effect of varying the sparsity of LGN-V1 connection. Setup with sparser initial connectivity is compensated by proportionally higher initial connection strengths ($\times 1/\text{sparsity}$, i.e., 0.152, 0.0633, 0.04). This allows the model to learn with a reasonable number of spikes [as in panel (b)]. All other parameters remain the same (standard wave width and speed). (a) Examples for three levels of sparsity (25%, 60%, 95%) in our model. Sampled neurons are shown at three different time points, population averages for some trial are shown in right-most column. Compared with a denser initial connectivity, the sparser connectivity results in a more structured receptive field, although this is less obvious in the population average. (b) Radial distribution of connection weights over distance to the RF center for the three level of sparsity (lighter green for denser connectivity). The off-zero bump in the 25% radial distribution exemplifies the randomness caused by the sparser connectivity. (c) Temporal evolution of the characteristic length. Sparser connectivity inevitably produces smaller values of characteristic lengths [less so by weighted radius, shown in the last column of panel (a)] but the trend of its decrease is similar for different sparsity levels. (d) More examples for 25% sparsity to showcase the temporal evolution of different connectivity structures.

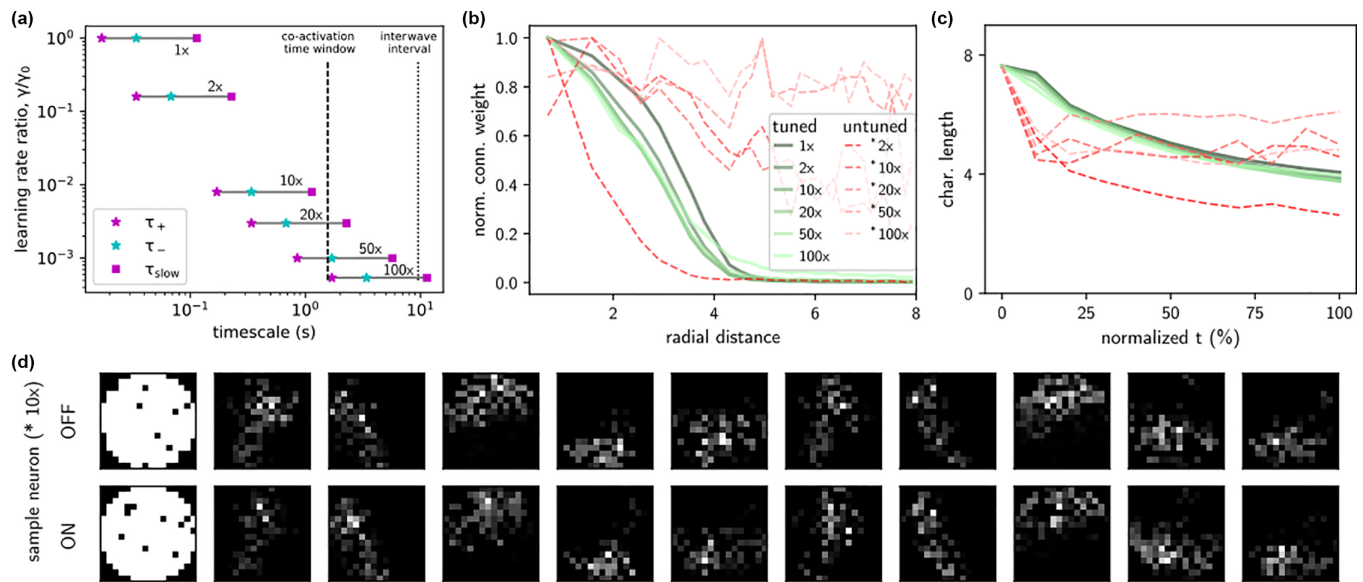


FIG. 12. Effect of varying the time constants of the triplet learning rule. In this experiment, the three time constants of the synaptic tracer variables that govern the triplet learning rule are increased by multiples of 2, 10, 20, 50, and 100. As the three timescales are varied, we decrease the LTD ratio proportionally (divided by the same ratio) so that the amplitude of LTD and LTP stay balanced despite changes in the timescale. The learning rates are also tuned to compensate for the increase in spike triplets that occur in the larger time windows. (a) The ratio of tuned learning rates with respect to the reference learning rate ($\gamma_0 = 0.0025$) is decreasing against the increasing time constants, so that the increasing number of spike triplets involved in the longer time constant case are accounted for. The dashed line indicates the co-activation time window of LGN and the retinal wave (1.563 s). The dotted line indicates the interwave interval (9.571 s). Beyond the interwave interval, we find it impossible to retain the pruning effect. (b) The radial distribution of the normalized learned connection weights, with tuned learning rates shown in shades of green and the untuned ones in shades of red. The synaptic pruning effect still exists for the untuned case with the lower multiples (e.g., $2\times$), but quickly disappears for higher multiples. Notice that the tuned $100\times$ radial distribution begins to have nonzero surrounding weights. (c) The temporal evolution of the characteristic length of the experiment from panel (b). Notice that only the $2\times$ untuned case with a reasonably higher learning rate shows the consistent effect of synaptic pruning, while other untuned ones fluctuate randomly. The tuned cases all behave similarly. (d) A neuron sample from an untuned $10\times$ trial is shown here. Its connectivity pattern changes violently from wave to wave as the untuned learning rate is too high.

- [1] R. Khazipov and H. J. Luhmann, Early patterns of electrical activity in the developing cerebral cortex of humans and rodents, *Trends Neurosci.* **29**, 414 (2006).
- [2] J. Conhaim, E. R. Cedarbaum, M. Barahimi, J. G. Moore, M. I. Becker, H. Gleiss, C. Kohl, and W. J. Moody, Bimodal septal and cortical triggering and complex propagation patterns of spontaneous waves of activity in the developing mouse cerebral cortex, *Dev. Neurobiol.* **70**, 679 (2010).
- [3] J. B. Ackman, T. J. Burbridge, and M. C. Crair, Retinal waves coordinate patterned activity throughout the developing visual system, *Nature (London)* **490**, 219 (2012).
- [4] F. J. Martini, T. Guillamón-Vivancos, V. Moreno-Juan, M. Valdeolmillos, and G. López-Bendito, Spontaneous activity in developing thalamic and cortical sensory networks, *Neuron* **109**, 2519 (2021).
- [5] T. McLaughlin, C. L. Torborg, M. B. Feller, and D. D. M. O'Leary, Retinotopic map refinement requires spontaneous retinal waves during a brief critical period of development, *Neuron* **40**, 1147 (2003).
- [6] M. S. Grubb, F. M. Rossi, J.-P. Changeux, and I. D. Thompson, Abnormal functional organization in the dorsal lateral geniculate nucleus of mice lacking the $\beta 2$ subunit of the nicotinic acetylcholine receptor, *Neuron* **40**, 1161 (2003).
- [7] C. Pfeifferberger, J. Yamada, and D. A. Feldheim, Ephrin-As and patterned retinal activity act together in the development of topographic maps in the primary visual system, *J. Neurosci.* **26**, 12873 (2006).
- [8] D. A. Butts, P. O. Kanold, and C. J. Shatz, A burst-based "Hebbian" learning rule at retinogeniculate synapses links retinal waves to activity-dependent refinement, *PLoS Biol.* **5**, e61 (2007).
- [9] A. D. Huberman, M. B. Feller, and B. Chapman, Mechanisms underlying development of visual maps and receptive fields, *Annu. Rev. Neurosci.* **31**, 479 (2008).
- [10] M. M. Syed, S. Lee, S. He, and Z. J. Zhou, Spontaneous waves in the ventricular zone of developing mammalian retina, *J. Neurophysiol.* **91**, 1999 (2004).
- [11] A. Bansal, J. H. Singer, B. J. Hwang, W. Xu, A. Beaudet, and M. B. Feller, Mice lacking specific nicotinic acetylcholine receptor subunits exhibit dramatically altered spontaneous activity patterns and reveal a limited role for retinal waves in forming ON and OFF circuits in the inner retina, *J. Neurosci.* **20**, 7672 (2000).
- [12] M. B. Feller, D. P. Wellis, D. Stellwagen, F. S. Werblin, and C. J. Shatz, Requirement for cholinergic synaptic transmission

- in the propagation of spontaneous retinal waves, *Science* **272**, 1182 (1996).
- [13] J. Cang, R. C. Rentería, M. Kaneko, X. Liu, D. R. Copenhagen, and M. P. Stryker, Development of precise maps in visual cortex requires patterned spontaneous activity in the retina, *Neuron* **48**, 797 (2005).
 - [14] B. K. Stafford, A. Sher, A. M. Litke, and D. A. Feldheim, Spatial-temporal patterns of retinal waves underlying activity-dependent refinement of retinofugal projections, *Neuron* **64**, 200 (2009).
 - [15] L. C. Liets, B. A. Olshausen, G.-Y. Wang, and L. M. Chalupa, Spontaneous activity of morphologically identified ganglion cells in the developing ferret retina, *J. Neurosci.* **23**, 7343 (2003).
 - [16] A. Thompson, A. Gribizis, C. Chen, and M. C. Crair, Activity-dependent development of visual receptive fields, *Curr. Opin. Neurobiol.* **42**, 136 (2017).
 - [17] R. Linsker, From basic network principles to neural architecture: Emergence of orientation-selective cells, *Proc. Natl. Acad. Sci. USA* **83**, 8390 (1986).
 - [18] C. von der Malsburg, Self-organization of orientation sensitive cells in the striate cortex, *Kybernetik* **14**, 85 (1973).
 - [19] N. Matsumoto, D. Barson, L. Liang, and M. C. Crair, Hebbian instruction of axonal connectivity by endogenous correlated spontaneous activity, *Science* **385**, eadh7814 (2024).
 - [20] O. S. Dhande, E. W. Hua, E. Guh, J. Yeh, S. Bhatt, Y. Zhang, E. S. Ruthazer, M. B. Feller, and M. C. Crair, Development of single retinofugal axon arbors in normal and $\beta 2$ knock-out mice, *J. Neurosci.* **31**, 3384 (2011).
 - [21] E. I. Moser, E. Kropff, and M.-B. Moser, Place cells, grid cells, and the brain's spatial representation system, *Annu. Rev. Neurosci.* **31**, 69 (2008).
 - [22] S. Löwel and W. Singer, Selection of intrinsic horizontal connections in the visual cortex by correlated neuronal activity, *Science* **255**, 209 (1992).
 - [23] M. G. Sheroziya, O. von Bohlen und Halbach, K. Unsicker, and A. V. Egorov, Spontaneous bursting activity in the developing entorhinal cortex, *J. Neurosci.* **29**, 12131 (2009).
 - [24] Y. Li, H. Lu, P.-L. Cheng, S. Ge, H. Xu, S.-H. Shi, and Y. Dan, Clonally related visual cortical neurons show similar stimulus feature selectivity, *Nature (London)* **486**, 118 (2012).
 - [25] Y.-C. Yu, R. S. Bultje, X. Wang, and S.-H. Shi, Specific synapses develop preferentially among sister excitatory neurons in the neocortex, *Nature (London)* **458**, 501 (2009).
 - [26] Z. Molnár, H. J. Luhmann, and P. O. Kanold, Transient cortical circuits match spontaneous and sensory-driven activity during development, *Science* **370**, eabb2153 (2020).
 - [27] D. L. Ringach, R. M. Shapley, and M. J. Hawken, Orientation selectivity in macaque V1: Diversity and laminar dependence, *J. Neurosci.* **22**, 5639 (2002).
 - [28] B. Chapman and M. P. Stryker, Development of orientation selectivity in ferret visual cortex and effects of deprivation, *J. Neurosci.* **13**, 5251 (1993).
 - [29] C. M. Niell and M. P. Stryker, Highly selective receptive fields in mouse visual cortex, *J. Neurosci.* **28**, 7520 (2008).
 - [30] D. Kerschensteiner, Glutamatergic retinal waves, *Front. Neural Circuits* **10**, 38 (2016).
 - [31] R. Clay Reid and J.-M. Alonso, Specificity of monosynaptic connections from thalamus to visual cortex, *Nature (London)* **378**, 281 (1995).
 - [32] R. Brette and W. Gerstner, Adaptive exponential integrate-and-fire model as an effective description of neuronal activity, *J. Neurophysiol.* **94**, 3637 (2005).
 - [33] J.-P. Pfister, Triplets of spikes in a model of spike timing-dependent plasticity, *J. Neurosci.* **26**, 9673 (2006).
 - [34] F. Zenke, G. Hennequin, and W. Gerstner, Synaptic plasticity in neural networks needs homeostasis with a fast rate detector, *PLoS Comput. Biol.* **9**, e1003330 (2013).
 - [35] A. Maccione, M. H. Hennig, M. Gandolfo, O. Muthmann, J. van Coppenhagen, S. J. Eglén, L. Berdondini, and E. Sernagor, Following the ontogeny of retinal waves: Pan-retinal recordings of population dynamics in the neonatal mouse, *J. Physiol.* **592**, 1545 (2014).
 - [36] D. Kerschensteiner and R. O. L. Wong, A precisely timed asynchronous pattern of ON and OFF retinal ganglion cell activity during propagation of retinal waves, *Neuron* **58**, 851 (2008).
 - [37] J. Gjorgjieva, C. Clopath, J. Audet, and J.-P. Pfister, A triplet spike-timing-dependent plasticity model generalizes the Bienenstock–Cooper–Munro rule to higher-order spatiotemporal correlations, *Proc. Natl. Acad. Sci. USA* **108**, 19383 (2011).
 - [38] H. Ko, L. Cossell, C. Baragli, J. Antolik, C. Clopath, S. B. Hofer, and T. D. Mrsic-Flogel, The emergence of functional microcircuits in visual cortex, *Nature (London)* **496**, 96 (2013).
 - [39] J. Crodelle and D. W. McLaughlin, Modeling the role of gap junctions between excitatory neurons in the developing visual cortex, *PLoS Comput. Biol.* **17**, e1007915 (2021).
 - [40] Y.-C. Yu, S. He, S. Chen, Y. Fu, K. N. Brown, X.-H. Yao, J. Ma, K. P. Gao, G. E. Sosinsky, K. Huang, and S.-H. Shi, Preferential electrical coupling regulates neocortical lineage-dependent microcircuit assembly, *Nature (London)* **486**, 113 (2012).
 - [41] D. H. Hubel and T. N. Wiesel, Receptive fields of single neurones in the cat's striate cortex, *J. Physiol.* **148**, 574 (1959).
 - [42] A. J. Watt and N. S. Desai, Homeostatic plasticity and STDP: Keeping a neuron's cool in a fluctuating world, *Front. Syn. Neurosci.* **2**, 5 (2010).
 - [43] G. Van den Bergh, B. Zhang, L. Arckens, and Y. M. Chino, Receptive-field properties of V1 and V2 neurons in mice and macaque monkeys, *J. Comp. Neurol.* **518**, 2051 (2010).
 - [44] D. Stellwagen, C. J. Shatz, and M. B. Feller, Dynamics of retinal waves are controlled by cyclic AMP, *Neuron* **24**, 673 (1999).
 - [45] D. E. Feldman, Timing-based LTP and LTD at vertical inputs to layer II/III pyramidal cells in rat barrel cortex, *Neuron* **27**, 45 (2000).
 - [46] J. W. H. Schnupp, T. M. Hall, R. F. Kokelaar, and B. Ahmed, Plasticity of temporal pattern codes for vocalization stimuli in primary auditory cortex, *J. Neurosci.* **26**, 4785 (2006).
 - [47] P. J. Sjöström, G. G. Turrigiano, and S. B. Nelson, Rate, timing, and cooperativity jointly determine cortical synaptic plasticity, *Neuron* **32**, 1149 (2001).
 - [48] R. S. Larsen, D. Rao, P. B. Manis, and B. D. Philpot, STDP in the developing sensory neocortex, *Front. Synaptic Neurosci.* **2**, (2010).
 - [49] G. Q. Bi and M. M. Poo, Synaptic modifications in cultured hippocampal neurons: Dependence on spike timing, synaptic strength, and postsynaptic cell type, *J. Neurosci.* **18**, 10464 (1998).
 - [50] J. E. M. Bennett and W. Bair, Refinement and pattern formation in neural circuits by the interaction of traveling waves

- with spike-timing dependent plasticity, *PLoS Comput. Biol.* **11**, e1004422 (2015).
- [51] T. Keck, G. B. Keller, R. I. Jacobsen, U. T. Eysel, T. Bonhoeffer, and M. Hübener, Synaptic scaling and homeostatic plasticity in the mouse visual cortex *in vivo*, *Neuron* **80**, 327 (2013).
- [52] L. C. Sincich and G. G. Blasdel, Oriented axon projections in primary visual cortex of the monkey, *J. Neurosci.* **21**, 4416 (2001).
- [53] J. S. Lund, A. Angelucci, and P. C. Bressloff, Anatomical substrates for functional columns in macaque monkey primary visual cortex, *Cereb. Cortex* **13**, 15 (2003).
- [54] X. Long and S.-J. Zhang, A novel somatosensory spatial navigation system outside the hippocampal formation, *Cell Res.* **31**, 649 (2021).
- [55] X. Long, B. Deng, J. Cai, Z. S. Chen, and S.-J. Zhang, A compact spatial map in V2 visual cortex, *bioRxiv* (2021).
- [56] E. S. Ruthazer and M. P. Stryker, The role of activity in the development of long-range horizontal connections in area 17 of the ferret, *J. Neurosci.* **16**, 7253 (1996).
- [57] G. B. Smith, B. Hein, D. E. Whitney, D. Fitzpatrick, and M. Kaschube, Distributed network interactions and their emergence in developing neocortex, *Nat. Neurosci.* **21**, 1600 (2018).
- [58] A. Gribizis, X. Ge, T. L. Daigle, J. B. Ackman, H. Zeng, D. Lee, and M. C. Crair, Visual cortex gains independence from peripheral drive before eye opening, *Neuron* **104**, 711 (2019).
- [59] J. Sharma, A. Angelucci, and M. Sur, Induction of visual orientation modules in auditory cortex, *Nature (London)* **404**, 841 (2000).
- [60] R. K. Naumann, S. Ray, S. Prokop, L. Las, F. L. Heppner, and M. Brecht, Conserved size and periodicity of pyramidal patches in layer 2 of medial/caudal entorhinal cortex, *J. Comp. Neurol.* **524**, 783 (2016).
- [61] T. Wang, F. Yang, Z. Wang, B. Zhang, W. Wang, and F. Liu, Modularization of grid cells constrained by the pyramidal patch lattice, *iScience* **24**, 102301 (2021).
- [62] M. E. Wosniack, J. H. Kirchner, L.-Y. Chao, N. Zabouri, C. Lohmann, and J. Gjorgjieva, Adaptation of spontaneous activity in the developing visual cortex, *eLife* **10**, e61619 (2021).
- [63] F. Siegel, J. A. Heimel, J. Peters, and C. Lohmann, Peripheral and central inputs shape network dynamics in the developing visual cortex *in vivo*, *Curr. Biol.* **22**, 253 (2012).
- [64] N. V. Swindale, The development of topography in the visual cortex: A review of models, *Netw. Comput. Neural Syst.* **7**, 161 (1996).
- [65] K. Herrmann, A. Antonini, and C. J. Shatz, Ultrastructural evidence for synaptic interactions between thalamocortical axons and subplate neurons, *Eur. J. Neurosci.* **6**, 1729 (1994).
- [66] F. Sengpiel and P. C. Kind, The role of activity in development of the visual system, *Curr. Biol.* **12**, R818 (2002).
- [67] M. T. Colonnese, J. Shen, and Y. Murata, Uncorrelated neural firing in mouse visual cortex during spontaneous retinal waves, *Front. Cell. Neurosci.* **11**, 289 (2017).
- [68] D. Molotkov, L. Ferrarese, T. Boissonnet, and H. Asari, Topographic axonal projection at single-cell precision supports local retinotopy in the mouse superior colliculus, *Nat. Commun.* **14**, 7418 (2023).
- [69] S. L. Smith and M. Häusser, Parallel processing of visual space by neighboring neurons in mouse visual cortex, *Nat. Neurosci.* **13**, 1144 (2010).
- [70] J. L. Hoy and C. M. Niell, Layer-specific refinement of visual cortex function after eye opening in the awake mouse, *J. Neurosci.* **35**, 3370 (2015).
- [71] A. B. Lee, B. Blais, H. Z. Shouval, and L. N. Cooper, Statistics of lateral geniculate nucleus (LGN) activity determine the segregation of ON/OFF subfields for simple cells in visual cortex, *Proc. Natl. Acad. Sci. USA* **97**, 12875 (2000).
- [72] C. P. Ratliff, B. G. Borghuis, Y.-H. Kao, P. Sterling, and V. Balasubramanian, Retina is structured to process an excess of darkness in natural scenes, *Proc. Natl. Acad. Sci. USA* **107**, 17368 (2010).
- [73] Y. Jiang, G. Purushothaman, and V. A. Casagrande, The functional asymmetry of ON and OFF channels in the perception of contrast, *J. Neurophysiol.* **114**, 2816 (2015).
- [74] T. Ichinose and S. Habib, On and off signaling pathways in the retina and the visual system, *Front. Ophthalmol.* **2**, 989002 (2022).
- [75] M. Li, Z. Cui, Y. Niu, B. Liu, W. Fan, D. Yu, and J. Deng, Synaptogenesis in the developing mouse visual cortex, *Brain Res. Bull.* **81**, 107 (2010).
- [76] E. Y. Litvina and C. Chen, Functional convergence at the retinogeniculate synapse, *Neuron* **96**, 330 (2017).
- [77] S. M. Sherman, The thalamus is more than just a relay, *Curr. Opin. Neurobiol.* **17**, 417 (2007).
- [78] M. S. Grubb and I. D. Thompson, Quantitative characterization of visual response properties in the mouse dorsal lateral geniculate nucleus, *J. Neurophysiol.* **90**, 3594 (2003).
- [79] W. P. Dai, Code repository for the reduced model, GitHub (2026), https://github.com/g13/spont_approx.
- [80] W. P. Dai, Code repository for the full model, GitHub (2026), <https://github.com/g13/patchV1>.