Adaptation of spontaneous activity in the developing visual cortex

Marina E. Wosniack¹,², Jan H. Kirchner¹,², Ling-Ya Chao¹, Nawal Zabouri³, Christian Lohmann³, Julijana Gjorgjieva¹,²

(1) Max Planck Institute for Brain Research, Frankfurt, Germany
(2) School of Life Sciences Weihenstephan, Technical University of Munich, Freising, Germany
(3) Netherlands Institute for Neuroscience, Amsterdam, Netherlands

*Corresponding author: gjorgjieva@brain.mpg.de

Abstract

Spontaneous activity drives the establishment of appropriate connectivity in different circuits during brain development. In the mouse primary visual cortex, two distinct patterns of spontaneous activity occur before vision onset: local low-synchronicity events originating in the retina, and global high-synchronicity events originating in the cortex. We sought to determine the contribution of these activity patterns to jointly organize network connectivity through different activity-dependent plasticity rules. We found that local events shape cortical input selectivity and topography, while global events have a homeostatic role regulating connection strength. To generate robust selectivity, we predicted that global events should adapt their amplitude to the history of preceding cortical activation, and confirmed by analyzing in vivo spontaneous cortical activity. This adaptation led to the sparsification of spontaneous activity on a slower timescale during development, demonstrating the remarkable capacity of the developing sensory cortex to acquire sensitivity to visual inputs after eye-opening.

Introduction

The impressive ability of the newborn brain to respond to its environment and generate coordinated output without any prior experience suggests that brain networks undergo substantial organization, tuning and coordination even as animals are still in the womb, driven by powerful developmental mechanisms. These broadly belong to two categories: activity-independent mechanisms, involving molecular guidance cues and chemoaffinity gradients which establish the initial coarse connectivity patterns at early developmental stages [1], and activity-dependent plasticity mechanisms which continue the refinement of this initially imprecise connectivity into functional circuits that can execute diverse behaviors in adulthood [2, 3, 4]. Non-random patterns of spontaneous activity drive these refinements, and act as training inputs to the immature circuits before the onset of sensory experience. Many neural circuits in the developing brain generate spontaneous activity, including the retina, hippocampus, cortex and spinal cord (reviewed in [5, 6]). This activity regulates a plethora of developmental processes such as neuronal migration, ion channel maturation and the establishment of precise connectivity [7, 8, 9], and perturbing this activity impairs different aspects of functional organization and axonal refinement [10, 11, 12]. These studies firmly demonstrate that spontaneous activity is necessary and instructive for the emergence of specific and distinct patterns of neuronal connectivity in the developing nervous system.

In the visual system, due to its accessibility, spontaneous activity and its role has been best characterized in the retina where it propagates as waves, and instructs the topographically ordered projection to the thalamus and superior colliculus [7]. This has inspired many computational models of topographic map development based on correlated patterns of activity (reviewed in [13, 14, 15]), including more recent approaches that simulate the entire temporal evolution of map formation from a combination of mechanisms [16, 17]. In the visual, and the sensory cortex more generally, spontaneous activity has been extensively characterized in isolated networks, either in cortical cultures or slices [18, 19, 20, 21]. Here, the spatiotemporal properties of spontaneous activity, including frequency, synchronicity, amplitude and spatial spread, depend on the studied region and developmental age [19, 20, 21, 22]. More recent in vivo recordings of
endogenous patterns of activity in the developing sensory cortex show that some spontaneous activity properties in the living animal are similar to those described in vitro [23, 24, 25]. These studies show that the generation and propagation of spontaneous activity in the intact cortex depend on input from different brain areas. For instance, activity from the sensory periphery substantially contributes to the observed activity patterns in the developing cortex, but there are other independent sources of activity within the cortex itself [24, 26, 27, 25]. Two-photon imaging of spontaneous activity in the in vivo mouse primary visual cortex before eye-opening (postnatal days, P8-10) has demonstrated that there are two independently occurring patterns of spontaneous activity with different sources and spatiotemporal characteristics. Peripheral events driven by retinal waves [28, 5] spread in the cortex as local events, engaging a relatively small number of the recorded neurons. In contrast, events intrinsic to the cortex that were unaffected by manipulation of retinal waves spread as global events, synchronously activating a large proportion of the recorded neurons [26].

We know relatively little about the information content of these local and global patterns of spontaneous cortical activity relevant for shaping local and brain-wide neural circuits. Specifically, it is unknown whether spontaneous activity from different sources affects distinct aspects of circuit organization, each providing an independent instructive signal, or if the two patterns cooperate to synergistically guide circuit organization. Therefore, using experimentally characterized properties of spontaneous activity in the visual cortex in vivo at P8-10, we developed a biologically plausible, yet analytically tractable, theoretical framework to determine the implications of this activity on normal circuit development with a focus on the topographic refinement of connectivity and the emergence of stable receptive fields.

We investigated the potential of two activity-dependent plasticity rules, the Hebbian covariance rule [29, 30, 31] and the Bienenstock-Cooper-Munro (BCM) rule [32], to generate stable and topographically refined receptive fields driven by L- and H-events. We postulated that peripheral L-events play a key role in topographically organizing receptive fields in the cortex, while H-events regulate connection strength homeostatically, operating in parallel to network refinements by L-events. We considered that H-events are ideally suited for this purpose because they maximally activate many neurons simultaneously, hence lack topographic information that can be used for synaptic refinement. Indeed, under the BCM rule, we found that these H-events dynamically slide the threshold between potentiation and depression, implementing the postulated homeostatic regulation. However, topography is disrupted since small L-events, which carry precise information for topographic connectivity refinements, mostly cause long-term depression in the synaptic weights. In comparison, under the Hebbian rule, we found that connectivity refinements are extremely sensitive to the properties of H-events. Therefore, we proposed that instead of regulating the threshold between potentiation and depression as in the BCM rule, H-events are themselves self-regulating. We implemented this self-regulation in the Hebbian rule through a fast adaptive mechanism that modulates H-event amplitude relative to ongoing activity in the network, effectively adjusting the homeostatic influence of H-events on connection strength. Re-examining the same spontaneous activity that inspired the model [26], we found a robust signature of adaptation in the amplitude of the cortically generated events. In our model, this adaptation led to the sparsification of cortical spontaneous activity over a prolonged timescale of development as observed experimentally in the visual and somatosensory cortices [33, 34], and previously unaccounted by other computational models. Therefore, our work proposes that global, cortically generated activity in the form of H-events rapidly adapts to ongoing network activity, supporting topographic organization of connectivity and maintaining synaptic strengths in an operating regime. The same adaptive mechanism eventually leads to the break up of large-scale spontaneous events, which acts to retain sensitivity of the cortex to incoming peripheral inputs as it transitions towards being visually-driven.

Results

A network model for connectivity refinements driven by spontaneous activity

Spontaneous activity in the mouse primary visual cortex before eye-opening exhibits two independently occurring patterns of spontaneous activity (Fig. 1A): (1) local low-synchronicity events involving 20-80% of the recorded neurons originating from the retina and propagated through the thalamus, called L-events and (2) global high-synchronicity events intrinsic to the cortex that activate nearly all cortical neurons (80-100%), called H-events [26]. This classification of spontaneous events into L- and H-events was recently validated both at the single-cell and population levels [35]. We first re-analyzed the extracted events from
Figure 1. Spontaneous activity patterns in early postnatal development. A. Two distinct patterns of spontaneous activity recorded in vivo in the visual cortex of young mice before eye-opening (P8-10). Blue shading denotes local spontaneous events (L-events; generated by retinal waves); orange shading denotes global spontaneous events (H-events). Activated neurons during each event are shown in red. B. Distributions of different event properties (amplitude, inter-event interval and event duration). Amplitude was measured as changes in fluorescence, relative to baseline, F/F0. C. Network schematic: thalamocortical connections are refined by spontaneous activity. The initially broad receptive fields with weak synapses evolve into a stable configuration with strong synapses organized topographically.

the original publication [26] and confirmed differences in specific features of these spontaneous events, as well as characterized novel aspects (Fig. 1B). In particular, L-events have lower amplitudes with smaller variance and inter-event intervals (IEI, the inverse of firing frequency) that follow an exponential-like distribution. H-events have a broader amplitude distribution with higher values and IEIs that follow a long-tail distribution. We found that L- and H-events have similar durations.

Next, we built a model that incorporates these two different patterns of spontaneous activity to investigate how they jointly refine network connectivity between the thalamus and the visual cortex (Fig. 1C). We used a one-dimensional feedforward network model – a microcircuit motivated by the small region of cortex imaged experimentally – composed of two layers, an input (presynaptic) layer corresponding to the sensory periphery (the thalamus) and a target (postsynaptic) layer corresponding to the primary visual cortex (Fig. 2A). Cortical activity \( v \) in the model is generated by two sources (Fig. 2B; Table 1): First, by L-events, \( u \), which activate neighboring thalamic cells and drive the cortex through the weight matrix, \( W \); and second, by H-events, \( v^{\text{spont}} \), activating the majority of the cortical cells. We used a rate-based unit with a membrane time constant \( \tau_m \) and linear activation function consistent with the coarse temporal structure of spontaneous activity during development, carrying information on the order of hundreds of milliseconds [13, 36, 3]:

\[
\tau_m \frac{dv(t)}{dt} = -v(t) + W(t)u(t) + v^{\text{spont}}(t).
\] (1)

To investigate the refinement of network connectivity during development, we studied the evolution of synaptic weights using rate-based plasticity rules operating over long timescales identified experimentally [37, 38]. First, we examined a classical Hebbian plasticity rule where coincident presynaptic thalamic activity and postsynaptic cortical activity in the form of L-events leads to synaptic potentiation. We postulated that H-events act homeostatically, maintaining synaptic weights in an operating regime by triggering global network depression involving the majority of synaptic weights in the absence of peripheral drive. Because they activate many neurons simultaneously, H-events lack the potential to drive topographical refinements. Their postulated homeostatic action resembles synaptic depression through downscaling, as observed in response to highly-correlated network activity, for instance upon blocking inhibition [39], or during slow-
wave sleep [40]. Therefore, we added a non-Hebbian term that depends only on the postsynaptic activity, with a proportionality constant which controls the relative amount of synaptic potentiation vs. depression. This differs from other Hebbian covariance plasticity rules for the generation of weight selectivity, which include non-Hebbian terms that depend on both pre- and postsynaptic activity [31, 41] and is mathematically related to models of heterosynaptic plasticity [42, 43, 44]. In this Hebbian rule, the change in synaptic weight between cortical neuron $j$ and thalamic neuron $i$ is given by:

$$
\tau_w \frac{dw_{ji}(t)}{dt} = v_j(t)(u_i(t) - \theta_u),
$$

where $\tau_w$ is the learning time constant and $\theta_u$ the proportionality constant in the non-Hebbian term, which we refer to as the ‘input threshold’. The activity time constant $\tau_m$ is much faster than the learning time constant, $\tau_m \ll \tau_w$, which allows us to separate timescales and to study how network activity on average affects learning (see Appendix 1).

Next, we investigated the BCM learning rule, commonly used in theoretical studies for its ability to generate selectivity in postsynaptic neurons which experience patterned inputs [32]. An important property of the BCM rule is its ability to homeostatically regulate the balance between potentiation and depression of all incoming inputs into a given neuron depending on the difference between the activity of that neuron and some target level. The change in synaptic weight between cortical neuron $j$ and thalamic neuron $i$ is given by:

$$
\tau_w \frac{dw_{ji}(t)}{dt} = v_j(t)u_i(t)(v_j(t) - \theta_v^i(t)),
$$

where

$$
\tau_\theta \frac{d\theta_v^i(t)}{dt} = -\theta_v^i(t) + \frac{v_j^2(t)}{v_0}
$$

describes how the sliding threshold $\theta_v^i(t)$ depends on postsynaptic activity, $v_0$ is the target rate of the cortical neurons and $\tau_\theta$ the sliding threshold time constant. According to this rule, synaptic weight change is Hebbian in that it requires coincident pre- and postsynaptic activity, as is only the case during L-events. H-events induce no direct plasticity in the network because of the absence of presynaptic activation, but they still trigger synaptic depression indirectly by increasing the sliding threshold between potentiation and depression.

Based on experimental measurements of the extent of thalamocortical connectivity at different developmental ages [45], we assumed that initial network connectivity was weak and all-to-all, such that each cortical neuron was innervated by all thalamic neurons. To account for the activity-independent stage of development guided by molecular guidance cues and chemoaffinity gradients [1, 46], a small bias was introduced to the initial weight matrix to generate a coarse topography in the network, where neighboring neurons in the thalamus project to neighboring neurons in the cortex and preserve spatial relationships (Fig. 2C, left). Following connectivity refinements through spontaneous activity and plasticity, a desired outcome is that the network achieves a stable topographic configuration (Fig. 2C, right) where each cortical neuron receives input only from a neighborhood of thalamic neurons. To evaluate the success of this process, we quantified two properties. First, the receptive field size defined as the average number of thalamic neurons that strongly innervate a cortical cell (Fig. 2D, 1). We normalized the receptive field size to the total number of thalamic cells, so that it ranges from 0 (no receptive field, all cortical cells decouple from the thalamus) to 1 (each cortical cell receives input from all the thalamic cells, all weights potentiate leading to no selectivity). We also quantified the topography of the final receptive field (Fig. 2D, 2 and Methods), which evaluates how well the initial bias is preserved in the final network connectivity. The topography ranges from 0 (all cortical neurons connect to the same set of thalamic inputs) to 1 (perfect topography relative to the initial bias). We note that the lack of initial connectivity bias did not disrupt connectivity refinements and receptive field formation, but could not on its own establish topography (Fig. S S1A). We also observed that in many cases cortical cells decouple from the thalamus.

**Spontaneous cortical H-events disrupt topographic connectivity refinement**

Both the Hebbian and the BCM learning rules are known to generate selectivity with patterned input stimuli [41, 32], and we confirmed that L-events can refine receptive fields in both scenarios (Fig. S1B, C). We found
that including H-events in the Hebbian rule requires that the input threshold $\theta_u$ and the inter-event interval for H-events $H_{int}$ be precisely balanced to generate selectivity and appropriately refined receptive fields (Fig. 3A, C-left). For a narrow range of $H_{int}$, weight selectivity emerges, but with some degree of decoupling between pre- and postsynaptic neurons (Fig. 3A, middle). Outside of this narrow functional range, individual cortical neurons are either non-selective (Fig. 3A, left) or decoupled from the thalamus (Fig. 3A, right). In comparison, including H-events in the BCM learning rule does not decouple pre- and postsynaptic neurons (Fig. 3B and Fig. S2) and selectivity can be generated over a wider range of H-inter-event-intervals $H_{int}$ and target rates $v_0$ for the BCM rule (Fig. 3C, right).

Despite this apparent advantage of the BCM rule, it generates receptive fields with much worse topography than the Hebbian learning rule (Fig. 3D). We found that the underlying reason for this worse topography of the BCM rule is the sign of synaptic change evoked by L-events of different sizes. In particular, small L-events generate postsynaptic cortical activity smaller than the sliding threshold and promote long-term synaptic depression (LTD), while large L-events generate cortical activity larger than the sliding threshold and promote long-term synaptic potentiation (LTP) (Fig. 3E,F). Therefore, the amount of information for connectivity refinements present in the small L-events is limited in the BCM learning rule resulting in poor topographic organization of receptive fields.

Taken together, our results confirm that H-events can operate in parallel to network refinements by L-events homeostatically regulating connection strength as postulated. However, the formation of receptive fields in the Hebbian rule is very sensitive to small changes in event properties (e.g., inter-event intervals), which are common throughout development [33]. In this case, H-events are disruptive and lead to the elimination of all thalamocortical synapses, effectively decoupling the cortex from the sensory periphery. In the BCM rule, including H-events prevents the decoupling of cortical cells from the periphery because the amount of LTD is dynamically regulated by the sliding threshold on cortical activity. However, L-events lose the ability to instruct topography because they generate LTP primarily when they are large.
Figure 3. Spontaneous cortical events disrupt receptive field refinement. A. Receptive fields generated by the Hebbian rule with input threshold $\theta_u = 0.4$ and increasing $H_{\text{int}}$. B. Receptive fields generated by the BCM rule with target rate $v_0 = 0.7$ and increasing $H_{\text{int}}$. C. Top: Average receptive field sizes obtained from 500 Monte Carlo simulations for combinations of $H_{\text{int}}$ and $\theta_u$ for the Hebbian rule (left) and $H_{\text{int}}$ and $v_0$ for the BCM rule (right). Bottom: Percentage of simulation outcomes classified as ‘selective’ when the average receptive field size is smaller than 1 and larger than 0, ‘non-selective’ when the average receptive field size is equal to 1, and ‘decoupled’ when the average receptive field size is 0 for the two rules. D. Topography of receptive fields classified as selective in C. Topography for the Hebbian rule: 0.78 ± 0.20, and for the BCM rule: 0.57 ± 0.25 (mean ± standard deviation). E. The response of a single cortical cell to L-events of different sizes (color) as a function of the sliding threshold for the BCM rule with $H_{\text{int}} = 3.5$ and $v_0 = 0.7$. The cell’s incoming synaptic weights from presynaptic thalamic neurons undergo LTP or LTD depending on L-event size. F. Probability of L-event size contributing to LTD (left) and LTP (right) for the BCM rule with the same parameters as in E.

Adaptive H-events achieve robust selectivity

Upon comparing the distinct outcomes of the Hebbian and BCM learning rules, we proposed that a mechanism that regulates the amount of LTD during H-events based on cortical activity, similar to the sliding threshold of the BCM rule, could be a potential biologically plausible solution to mitigate the decoupling of cortical cells in the Hebbian rule. Therefore, this mechanism combined with the Hebbian learning rule could lead to refined receptive fields that also have good topographic organization. Hence, we postulated that H-events adapt by assuming that during H-events cortical cells scale their amplitude to the average amplitude of the preceding recent events. In particular, for each cortical cell $j$ an activity trace $\eta_j$ integrates the cell’s firing rate $v_j$ over a timescale $\tau_\eta$ slower than the membrane time constant:

$$\tau_\eta \frac{d\eta_j(t)}{dt} = -\eta_j(t) + v_j(t).$$  \hspace{1cm} (5)$$

This activity trace $\eta_j$ then scales the intrinsic firing rate of the cortical cells during an H-event, $H_{\text{amp}} \rightarrow \eta_j H_{\text{amp}}$, making it dependent on its recent activity. The activity trace $\eta_j$ might biophysically be implemented.
through a calcium-dependent signaling pathway that is activated upon sufficient burst depolarization and that is able to modulate a cell’s excitability in the form of intrinsic plasticity [47, 48, 49]. A fast, activity-dependent mechanism that decreases single-neuron excitability following a prolonged period of high network activity has been identified in spinal motor neurons of neonatal mice [50]. However, there might be other ways to implement this adaptation (see Discussion).

Using adaptive H-events, we investigated the refinement of receptive fields in the network with the same Hebbian rule (Fig. 4A). In sharp contrast to the Hebbian rule with non-adaptive H-events (Fig. 3A), we observed that changing the average inter-event interval of H-events in a wider and more biologically realistic range (from the data, $H_{\text{int}} \sim 3L_{\text{int}}$) yields selectivity and appropriately refined receptive fields (Fig. 4A). Increasing $\theta_u$ or decreasing $H_{\text{int}}$ yields progressively smaller receptive fields, while mitigating cortical decoupling (Fig. 4B). The refined receptive fields also have a very good topography because L-events in the Hebbian learning rule carry nearest-neighbor information for the topographic refinements (Fig. 4C).

Next, we investigated how the proposed adaptive mechanism scales H-event amplitude by modulating the relative strength of H-events. For the Hebbian rule, we calculated the analytical solution for weight development with L- and H-events by reducing the dimension of the system to two: one being the average of the weights that potentiate and form the receptive field, $w_{R\text{F}}$, and the other being the average of the remaining weights, which we called ‘complementary’ to the receptive field, $w_C$ (Fig. 4D; see Appendix 1, Methods). We calculated the phase plane area of the reduced two-dimensional system with non-adaptive H-events (calculated as the proportion of initial conditions) that results in selectivity, potentiation or depression (Fig. 4D, bottom). We found that adaptively modulating the strength of H-events maximizes the area of the phase plane that results in selectivity (Fig. 4E, shaded region). The range of H-event strengths that maximizes the selective area for each input threshold in the reduced two-dimensional system can be related to the scaling of H-event amplitude in the simulations (Methods). In particular, the adaptation reliably shifts the H-event amplitude that would have occurred without adaptation, ‘non-adapted strength of H-events’, into the regime of amplitudes that maximizes selectivity, ‘adapted strength of H-events’ (Fig. 4E). Therefore, the adaptation of H-event amplitudes controls the selective refinement by peripheral L-events by modulating cortical depression by the adapted H-events.

**In vivo** spontaneous cortical activity shows a signature of adaptation

To determine whether spontaneous cortical activity contains a signature of our postulated adaptation mechanism of H-event amplitudes, we reanalyzed published *in vivo* two-photon Ca$^{2+}$ imaging data recorded in the visual cortex of young mice (P8-10) [26]. We found no evidence for long-term fluctuations in cortical excitability, making it unlikely that L- and H-event amplitudes are co-regulated by activity-independent mechanisms (Fig. S3). We identified L- and H-events based on previously established criteria, including participation rates, event amplitude and jitter [26]. For each detected H-event, we extracted all spontaneous (L- or H-) events that preceded this H-event combining several consecutive recordings in the same animal, up to a maximum time $\tau_{\text{max}} = 300$ s. We then scaled the amplitude of each previous event by an exponential kernel with a decay time constant of $\tau_{\text{decay}} = 1000$ s, which is sufficiently long to integrate many preceding spontaneous events (compared with the inter-event intervals in Fig. 1B), and averaged these scaled amplitudes to get an aggregate quantity over amplitude and frequency (see Methods).

We found that this aggregate amplitude of L- and H-events preceding a given H-event is significantly correlated ($r = 0.44, p < 10^{-10}$) to the amplitude of the selected H-event (Fig. 5B). Consequently, a strong (weak) H-event follows strong (weak) average preceding network activity (Fig. 5C), suggesting that cortical cells indeed adapt their spontaneous firing rates as a function of their previous activity levels. The correlations are robust to variations in the inclusion criteria, maximum time $\tau_{\text{max}}$ to integrate activity and the exponential decay time constant $\tau_{\text{decay}}$ (Fig. S4). This analysis suggests that the timescale of adaptation is reasonably fast, but at the same time sufficiently long to integrate preceding spontaneous events, matching the relevant timescales of information transmission in the developing cortex [51, 13].
Figure 4. Adaptive cortical events refine thalamocortical connectivity. A. Receptive field refinement with adaptive H-events and different H-inter-event-intervals, $H_{\text{int}}$. Top: $\theta_u = 0.5$; bottom: $\theta_u = 0.6$. B. Receptive field sizes from 500 Monte Carlo simulations for combinations of $H_{\text{int}}$ and $\theta_u$. Bottom: Percentage of simulation outcomes classified as ‘selective’ when the average receptive field size is smaller than 1 and larger than 0, ‘non-selective’ when the average receptive field size is equal to 1, and ‘decoupled’ when the average receptive field size is 0 for the two rules. C. Topography of receptive fields classified as selective in B using non-adaptive (0.78±0.20) and adaptive (0.78±0.19) H-events (mean ± standard deviation). D. Top: Reduction of the full weight dynamics into two dimensions. Two sets of weights were averaged: those which potentiate and form the receptive field, $w_{\text{RF}}$, and the complementary set of weights that depress, $w_C$. Bottom: Initial conditions in the reduced two-dimensional phase plane were classified into three outcomes: ‘selective’, ‘non-selective’ and ‘decoupled’. We sampled 2500 initial conditions which evolved according to Eq. A4 until the trajectories reached one of the selective fixed points, $(w_{\text{max}}, 0)$ and $(0, w_{\text{max}})$, or resulted in no selectivity either because both weights depressed to $(0, 0)$ or potentiated to $(w_{\text{max}}, w_{\text{max}})$. The normalized number of initial coordinates generating each region can be interpreted as the area of the phase plane that results in each outcome. E. Top: Normalized area of the phase plane of the reduced two-dimensional system that resulted in ‘selective’, ‘non-selective’ and ‘decoupled’ outcomes for $\theta_u = 0.53$ as a function of H-events strength. The darker shading indicates ranges of non-adapted H-event strength where the selectivity area is maximized. Bottom: The corresponding adapted strength of H-events was calculated in simulations with adaptive H-events, and plotted as a function of the nominal, non-adaptive strength of H-events. The range of adapted H-event strengths (bottom) corresponds to the range of non-adaptive values that maximize the selectivity area (top). Each point shows the average over 10 runs and the bars the standard deviation (which are very small).
Figure 5. Spontaneous events in developing cortex adapt to recent activity. A. Calcium trace of a representative recording with L- (blue) and H-events (orange) [26]. For each H-event, the amplitude of the spontaneous events (both L- and H-) that preceded it within a time window $\tau_{\text{max}} = 300 \text{s}$ was averaged with an exponential leak of decay time constant $\tau_{\text{decay}} = 1000 \text{s}$. B. Significant correlation between the amplitude of an H-event and the average amplitude of preceding events within $\tau_{\text{max}}$ ($N = 195$ events from 9 animals). Animals with fewer than 12 H-events preceded by activity within $\tau_{\text{max}}$ were excluded from this analysis (see Methods). The Pearson correlation coefficient is $r = 0.44, p < 10^{-10}, \text{CI} = (0.32, 0.54)$. Red line indicates regression line with 95% confidence bounds as shaded area. C. Schematic of the postulated adaptation: A weak (strong) H-event is more likely to be preceded by weak (strong) spontaneous events.

Modulating spontaneous activity properties makes different predictions for receptive field refinements

Our results make relevant predictions for the refinement of receptive fields upon manipulating spontaneous activity. H-event frequency can be experimentally reduced by a gap junction blocker (carbenoxolone) [26]. L-events can also be experimentally manipulated, for instance, by altering inhibitory signaling [35], or the properties of retinal waves which propagate as L-events into the cortex. While the effect of altered inhibitory signaling on receptive field refinements is still unknown, retinal wave manipulations evoke measurable defects in the retinotopic map refinement of downstream targets [52, 53, 12]. A most prominent example are $\beta 2$ knockout mice, which lack expression of the $\beta 2$ subunit of the nicotinic acetylcholine receptor ($\beta 2$-nAChR) that mediates spontaneous retinal waves in the first postnatal week. In these animals, retinal waves are consistently larger as characterized by the increased correlation with distance [54, 55, 56], in addition to other features. As a result, the receptive fields in the visual cortex in $\beta 2$-mutants are imprecise and less refined. Based on the analysis of our model (Fig. 4), we asked what implications the spatial structure of L-events might have on receptive field refinement.

We performed Monte Carlo simulations with a range of input thresholds $\theta_u$ and variable participation rates of thalamic neurons in L-events, using the Hebbian rule with adaptive H-events (Fig. 6A). Consistent with the imprecise receptive fields in the visual cortex in $\beta 2$-mutants [53], larger L-events in our model produce less refined i.e. larger receptive fields in the cortical network (Fig. 6B, C left). Smaller L-events also refine receptive fields with better topographic organization (Fig. 6C, right) and do not impair connectivity refinements. This result could be linked to experiments where the expression of $\beta 2$-nAChR is limited to the ganglion cell layer of the retina, resulting in smaller retinal waves than those in wild-type and undisturbed retinotopy in the superior colliculus [11], although the effects in the cortex are unknown. Therefore, we propose that manipulations that modulate the size of sensory activity from the periphery have a profound impact on the precision of receptive field refinement in downstream targets, making predictions to be tested experimentally.

Adaptive H-events promote the developmental event sparsification of cortical activity

Spontaneous activity is not static, but consists of distinct patterns that are dynamically regulated during development by ongoing activity-dependent plasticity which continuously reshapes network connectivity [33, 34, 57]. We next asked how our observed modifications in network connectivity that result in refined receptive fields based on fixed activity statistics during a single postnatal day might further modify spontaneous
Figure 6. Receptive field refinement depends on the properties of L-events. A. Receptive field sizes from 500 Monte Carlo simulations for different sizes of L-events where the minimum participation rate was 20% and the maximum participation rate was varied. The input threshold was taken from the range $0.3 \leq \theta_u \leq 0.7$, while the adaptive H-events had a fixed inter-event-interval $H_{\text{int}} = 3.5$. B. Individual receptive fields for different L-event maximum participation rates and $\theta_u = 0.50$. As the upper bound of the participation rate progressively increases from 40% to 80%, receptive fields get larger. C. Left: Receptive field sizes from A binned according to the maximum L-event size. Right: Corresponding topography of selective receptive fields for different sizes of L-events.

activity patterns on a much longer developmental timescale of several days. Therefore, we analyzed the activity of the simulated cortical neurons during the process of receptive field refinement in the presence of adaptive H-events. We detected ‘effective’ L- and H-events in the simulated cortical activity following an automatic event detection similar to the routine used to detect events in the experimental calcium traces (Figures 7A and B, see Methods for details). Given the sequence of these simulated effective L- and H-events, we reconstructed the probability distributions of their amplitudes, inter-event intervals and durations as in the data (Fig. 7B and Fig. S5A). Here we stress that the features of effective events differ from the features of ‘putative’ ones that were initially modeled according to the data (Table 1). During a putative H-event, the adaptation mechanism can decrease the amplitude of simulated cortical cells. Upon sufficient reduction in amplitude, such a putative H-event can have a reduced effective participation rate because only the subset of the cortical neurons that is activated above the detection threshold will be detected as active. Therefore, a putative H-event can be identified as an effective L-event.

What is the effect of this classification of putative H-events as effective L-events? The amount of putative H-events classified as effective L-events progressively increases as receptive fields are further refined, which represents developmental time in the model. Since the input threshold $\theta_u$ of the Hebbian learning rule is clearly related to receptive field size (Fig. 4B), we used $\theta_u$ as a proxy for time of development in the model: low $\theta_u$ corresponds to earlier developmental stages when receptive fields are large, while high $\theta_u$ corresponds to late developmental stages when receptive fields are refined. This assumption is also in line with the fact that the input resistance of neurons in V1 and S1 decreases during development [58, 34], so that the depolarizing current necessary to trigger an action potential increases with age.

We investigated the progressive evolution of spontaneous activity properties in our network model by varying $\theta_u$. At an early developmental stage in the model ($\theta_u = 0.45$), the relationship between effective event amplitude and participation rate (Fig. 7D, left) strongly resembles the data (Fig. 7C, reproduced from [26]). The amplitude of effective L-events is approximately half the amplitude of effective H-events. Moreover, we also observe a high proportion of events with large sizes (Fig. S5B), suggesting that in this network large
spontaneous events are very frequent. At an intermediate developmental stage in the model ($\theta_u = 0.50$), the resemblance between effective amplitude and participation rate to the data is less pronounced (Fig. 7D, middle), with lower average amplitudes and density of large events (Fig. S5B). Finally, at late developmental stages in the model ($\theta_u = 0.60$), the dependence of effective amplitude on participation rate is almost absent (Fig. 7D, right). Event amplitude is much lower and there are fewer effective events with more than 80% participation rate (Fig. S5C). Therefore, we observe a progressive event sparsification of the effective spontaneous events during ongoing development in our model, whereby spontaneous events become smaller in size, with fewer cells participating. This event sparsification also becomes apparent when examining the correlation matrices of simulated cortical activity for increasing input threshold values (Fig. 7E, F).

Interestingly, such event sparsification of spontaneous activity has been observed experimentally in the mouse barrel cortex during postnatal development from P4 to P26 [34] and in the visual cortex from P8 to P79 [33]. During this period, in the visual cortex, the size of spontaneous events decreases (Fig. 7G), the amplitude of the participating cells also decreases, while event frequency increases (Fig. 7H) [33]. This progressive event sparsification of cortical activity is generated by mechanisms intrinsic to the cortex, and does not seem to be sensory-driven [33, 34]. We found the same relationships in our model using $\theta_u$ as a proxy for developmental time (Fig. 7I, J).

In summary, our framework for activity-dependent plasticity and receptive field refinement between thalamus and cortex with adaptive H-events can tune the properties of cortical spontaneous activity and provide a substrate for the event sparsification of cortical activity during development on a much longer timescale than receptive field refinement. This sparsification has been found in different sensory cortices, including visual [33], somatosensory [34] and auditory [57], suggesting a general principle that underlies network refinement. However, the event sparsification we observe is different than sparse network activity implicated in sparse efficient coding, which interestingly seems to decrease during development [59, 60]. Our modeling predicts that cortical event sparsification is primarily due to the suppression of cortically-generated H-events in the Hebbian rule, which switches cortical sensitivity to input from the sensory periphery after the onset of sensory experience.

Discussion

We examined the information content of spontaneous activity for refining local microcircuit connectivity during early development. In contrast to classical works on activity-dependent refinements which used mathematically convenient formulations of spontaneous activity [61, 41], we used spontaneous activity patterns characterized in the mouse visual cortex in vivo before the onset of vision, which revealed its rich structure. Specifically, we explored the joint contribution of two distinct patterns of spontaneous activity recorded in the mouse visual cortex before the onset of vision, local (L-events) and global (H-events), on robustly establishing precise topographic receptive fields between the thalamus and the cortex in a rate-based network model with activity-dependent plasticity. We compared the emergence of selective and topographic receptive fields using two learning rules: the Hebbian and the BCM rule. Because of their spatially correlated activity, we proposed that peripherally generated L-events can establish topographically-refined receptive fields in the cortex, while H-events can regulate connection strength homeostatically. Although L-events successfully instructed topographic receptive field refinements in the Hebbian rule, naively including H-events provided too much depression, eliminating selectivity in the network despite fine-tuning. In contrast, in the BCM rule, H-events were indeed homeostatic, regulating the threshold between depression and potentiation. However, small L-events, which carry precise information for topographic connectivity refinements, mostly caused long-term depression in the synaptic weights disrupting topography. Inspired by the sliding threshold in the BCM rule, we proposed a similar adaptive mechanism operating at the single-cell level in the Hebbian rule, which regulates the amplitude of the cortically generated H-events according to the preceding average activity in the network to refine receptive fields with excellent topography. This mechanism adjusts the strength of H-events to homeostatically balance local increases and decrease in activity. Without any additional fine-tuning, this mechanism can also explain the long-term event sparsification of cortical activity as the circuit matures and starts responding to visual input. Therefore, we propose that L- and adaptive H-events cooperate to synergistically guide circuit organization of thalamocortical synapses during postnatal development.
Figure 7. Adaptive H-events promote sparsification of cortical activity during development. A. Identifying L- and H-events in mouse cortical activity [26]. Each activity trace represents an individual cortical cell. B. Left: We used a similar protocol to identify effective L- and H-events in the model. Right: Distributions of average amplitude and inter-event intervals extracted from the model. C. Amplitude vs. participation rate plot from the data [26]. The regression line for the amplitude vs. participation rate in H-events has a positive slope. D. Amplitude vs. participation rate plots from the model, for different values of \( \theta_u \). Inset: The regression line for the amplitude vs. participation rate in H-events has a slope that decreases with \( \theta_u \). E. Correlation between cortical neurons decreases as a function of the input threshold \( \theta_u \) in the model, a proxy for developmental time. F. Correlation matrices of simulated cortical neuron activity corresponding to D. G,H. Event sizes and the relationship between frequencies (open squares) and amplitudes (filled circles) of spontaneous events at different postnatal ages (data reproduced from [33]). I. Effective spontaneous event sizes as a function of the input threshold \( \theta_u \). J. Frequencies (squares) and amplitudes (circles) of effective L-events in the model at different input thresholds. For the data (model), the bars represent the standard error of the mean (standard deviation of 10 simulation runs).

The origin of cortical event amplitude adaptation After a re-examination of spontaneous activity recorded in the developing cortex in vivo between postnatal days 8 and 10 [26], we found evidence for the proposed H-event amplitude adaptation. This mechanism is sufficiently general in its formulation that it could be realized at the cellular, synaptic or network level. At the cellular level, the adaptation mechanism resembles intrinsic plasticity. Typically, intrinsic plasticity has been reported in response to long-term perturbations in activity or persistent changes in synaptic plasticity like LTP and LTD, where the intrinsic properties of single neurons are adjusted in an activity-dependent manner [48, 47]. During intrinsic plasticity,
neurons can alter the number and expression levels of ion channels to adjust their input-output function
by modifying their firing thresholds or response gains, which could represent the substrate for H-event
amplitude regulation. Our adaptation mechanism is consistent with fast intrinsic plasticity operating on
the timescale of several spontaneous events supported by many experimental studies. For instance, intrinsic
excitability of spine motoneurons is depressed after brief but sustained changes in spinal cord network
activity in neonatal mice [50]. Similarly, hippocampal pyramidal neurons also exhibit a rapid reduction of
intrinsic excitability in response to sustained depolarizations lasting up to several minutes [62]. In addition
to reduced excitability, in the developing auditory system, enhanced intrinsic excitability has been reported
in the cochlea followed reduced synaptic excitatory input from hair cells in a model of deafness, although
this change is slower than our proposed adaptation mechanism [63].

At the synaptic level, our adaptation mechanism can be implemented by synaptic scaling, a process
whereby neurons regulate their activity by scaling incoming synaptic strengths in response to perturbations
[64]. A second possibility is short-term depression, which appears to underlie the generation of spontaneous
activity episodes in the chick developing spinal cord [65, 66]. Similarly, release probability suppression
has been reported to strongly contribute to synaptic depression during weak activity at the calyx of Held
[67], which is more pronounced at immature synapses where morphological development renders synaptic
transmission less effective [68, 69]. This is also the case in the cortex, where short-term synaptic plasticity in
young animals is stronger [70]. Beyond chemical synapses, plasticity of gap junctions, which are particularly
prevalent in development [71], could also be a contributing mechanism that adapts overall network activity
[72, 73].

Finally, at the network level, the development of inhibition could be a substrate for amplitude adap-
tation of cortically generated events. The main inhibitory neurotransmitter, GABA, is thought to act as
a depolarizing neurotransmitter, excitatory in early postnatal days [74], although recent evidence argues
that GABAergic neurons have an inhibitory effect on the cortical network already in the second postnatal
week [75, 76, 77]. Thus, the local maturation of inhibitory neurons – of which there are several types
[78] – that gradually evolve to balance excitation and achieve E/I balance [79] could provide an alternative
implementation of the proposed H-event adaptation.

Developmental sparsification of cortical activity  On a longer timescale than receptive field refine-
ment, we demonstrated that the adaptation of H-event amplitude can also bring about the event sparsification
of cortical activity, as global, cortically generated H-events are attenuated and become more localized. The
notion of ‘sparse neural activity’ has received significant attention in experimental and theoretical studies of
sensory processing in the cortex, including differing definitions and implementations [80, 81, 59, 82, 83]. In
particular, sparse activity in the mature cortex has been argued to be important for the efficient coding of
sensory inputs of different modalities [82, 80]. Hence, the developmental process of receptive field refinement
might be expected to produce sparser network activity over time. However, experiments directly testing this
idea have found no, or even opposite, evidence for the developmental emergence of efficient sparse coding
[59, 60]. In the context of our work, our notion of sparsification is fairly simplistic: it refers to an overall
sparsification of network events (fewer active cells per event) where cortically-generated H-events become
replaced with peripherally-driven L-events, following [26]. Given that our data pertain to developmental
spontaneous activity before eye-opening, in complete absence of stimulation, it is not straightforward to
relate our event sparsification to the sparse efficient coding hypothesis.

Assumptions in the model  Our model was based on the assumption that L- and H-events have distinct
roles during the development of the visual system. Retinal waves, the source of L-events, carry information
downstream about the position and function of individual retinal ganglion cells [55], hence they are ideally
suited to serve as ‘training patterns’ to enable activity-dependent refinements based on spatiotemporal
correlations [84, 2, 4]. Since all cells are maximally active during H-events, these patterns likely do not
carry much information that can be used for activity-dependent refinement of connectivity. In contrast, we
assumed that H-events homeostatically control synaptic weights, operating in parallel to network refinements
by L-events. Indeed, highly correlated network activity can cause homeostatic down-regulation of synaptic
weights via a process known as synaptic scaling [39]. The homeostatic role of H-events is also consistent
with synaptic downscaling driven by slow waves during sleep, a specific form of synchronous network activity
Since during development sleep patterns are not yet regular, we reasoned that refinement (by L-events) and homeostasis (by H-events) occur simultaneously instead of being separated into wake and sleep states.

We focused on the role of spontaneous activity in driving receptive field refinements, rather than study how spontaneous activity is generated. While the statistical properties of spontaneous activity in the developing cortex are well-characterized, the cellular and network mechanisms generating this activity remain elusive. In particular, while H-event generation has been shown to rely on gap junctions [26, 71], which recurrently connect developing cortical cells, not much is known about how the size of cortical events is modulated and how an L-event is prevented from spreading and turning into an H-event. It is likely that cortical inhibition plays a critical role in localizing cortical activity and shaping receptive field refinements [86, 35], for instance, through the plasticity of inhibitory connections by regulating E/I balance [79]. As new experiments are revealing more information about the cellular and synaptic mechanisms that generate spatiotemporally patterned spontaneous activity [87], a full model of the generation and the effect of spontaneous activity might soon be feasible.

We found that two learning rules can generate selectivity and receptive field refinement, the Hebbian rule with adaptive H-events, and the BCM rule. Both rules have an adaptive component: in the BCM rule it is the threshold between potentiation and depression that slides as a function of postsynaptic activity, while in the Hebbian rule it is the adaptive amplitude of H-events, while the rule itself is fixed. Although experimental data contains signatures of the stereotypical activity dependence of the BCM rule [88, 89], whether a sliding threshold exists is still debated. Moreover, the timescale over which the threshold slides to prevent unbounded synaptic growth needs to be much faster than experimentally found [90]. Our proposed H-event amplitude adaptation operates on the fast timescale of several spontaneous events found experimentally [26, 62, 50]. Hence, together with the better topography and the resulting event sparsification as a function of developmental stage that the Hebbian rule with adaptive H-events generates, we propose it as the more likely plasticity mechanism to refine receptive fields in the developing visual cortex.

Finally, we have focused here on the traditional view that molecular gradients set up a coarse map that activity-dependent mechanisms then refine [14]. In our model, this was implemented as a weak bias in the initial connectivity, which did not affect our results regarding the refinement of receptive fields. Both activity and molecular gradients may work together in interesting ways to refine receptive fields [91], and future work should include both aspects.

**Predictions of the model**  Our model makes several experimentally testable predictions. First, we showed that changing the frequency of H-events can affect the size of the resulting receptive fields under both the BCM (Fig. 3) and the Hebbian rule with adaptive H-events (Fig. 4). The frequency of H-events can be experimentally manipulated using optogenetics or pharmacology. For instance, gap junction blocker (carbenoxolone) has been shown to specifically reduce the frequency of H-events [26], hence in that scenario our results predict broader receptive fields.

Additionally, L-events can also be experimentally manipulated. Recently, reduced inhibitory signaling by suppressing somatostatin-positive interneurons have has been shown to increase the size of L-events in the developing visual cortex [35]. With the effect of altered inhibitory signaling on receptive field refinements still unknown, our work predicts larger receptive fields and worse topography upon reduction of inhibition. L-events can also be experimentally manipulated by changing the properties of retinal waves, which can significantly affect retinotopic map refinement of downstream targets [52, 53, 12]. Indeed, β2 knockout mice discussed earlier have larger retinal waves and less refined receptive fields in the visual cortex [54, 55, 56]. If we assume that these larger retinal waves manifest as larger L-events in the visual cortex following [26], then these experimental observations are in agreement with our model results.

Third, our model predicts that as a result of receptive field refinement during development, network events sparsify as global, cortically generated events are attenuated and become more localized. Interestingly, the properties of spontaneous activity measured experimentally in different sensory cortices [33, 57, 92, 93, 94, 34] and in the olfactory bulb [87] change following a very similar timeline during development as predicted in our model. However, in many of these studies activity has not been segregated into peripherally driven L-events and cortically generated H-events. Therefore, our model predicts that the frequency of L-events would increase while the frequency of H-events would decrease over development.

Finally, we propose that for a Hebbian rule to drive developmental refinements of receptive fields using
spontaneous activity L- and H-event patterns recorded in vivo [26], H-events need to adapt to ongoing network activity. Whether a fast adaptation mechanism like the one we propose operates in the cortex requires prolonged and detailed activity recordings in vivo, which are within reach of modern technology [2, 95, 25]. Our framework also predicts that manipulations that affect overall activity levels of the network, such as activity reduction by eye enucleation, would correspondingly affect the amplitude of ongoing H-events.

Conclusion  In summary, we studied the refinement of receptive fields in a developing cortex network model constrained by realistic patterns of experimentally recorded spontaneous activity. We proposed that adaptation of the amplitude of cortically generated spontaneous events achieves this refinement without additional assumptions on the type of plasticity in the network. Our model further predicts how cortical networks could transition from supporting highly-synchronous activity modules in early development, to sparser peripherally driven activity suppressing local amplification which could be useful for preventing hyper-excitability and epilepsy in adulthood, while enhancing the processing of sensory stimuli.

Methods and Materials

Network model
We studied a feedforward, rate-based network with two one-dimensional layers, one of \( N_u \) thalamic neurons (u) and the other of \( N_v \) cortical neurons (v), with periodic boundary conditions in each layer to avoid edge effects. The initial connectivity in the network was all-to-all with uniformly distributed weights in the range \( w_{\text{init}} = [a, b] \). In addition, a topographic bias was introduced by modifying the initially random connectivity matrix to have strongest connections between neurons at the matched topographic location, and which decay with a Gaussian profile with increasing distance (Fig. 2C), with amplitude \( s \) and spread \( \sigma_s \). During evolution of the weights, soft bounds were applied on the interval \( [0, w_{\text{max}}] \). We studied weight evolution under two activity-dependent learning rules: the Hebbian (Eq. 2) and the BCM (Eq. 3) rules. Table 1 lists all parameters. Sample codes can be found at github.com/comp-neural-circuits/LH-events.

Generation of L- and H-events
We modeled two types of spontaneous events in the thalamic (L-events) and the cortical (H-events) layer of our model [26]. During L-events, the firing rates of a fraction (\( L_{\text{pct}} \)) of neighboring thalamic neurons were set to \( L_{\text{amp}} = 1 \) during a period \( L_{\text{dur}} \) and were otherwise 0. Similarly, during H-events, the firing rates of a fraction (\( H_{\text{pct}} \)) of cortical cells were set to \( H_{\text{amp}} \) during time \( H_{\text{dur}} \). As a result, cortical neuron activity was composed of H-events and L-events transmitted from the thalamus. For each H-event, \( H_{\text{amp}} \) was independently sampled from a Gaussian distribution with mean \( H_{\text{amp}} \) and standard deviation \( H_{\text{amp}}/3 \). The inter-event intervals were \( L_{\text{int}} \) and \( H_{\text{int}} \) sampled from experimentally characterized distributions in [26] (Table 1). We note that in the experiments, both L- and H-events were characterized in the primary visual cortex; in our model, we assume that L-events are generated in the retina and subsequently propagated through the thalamus to the cortex, where they manifest with the experimentally reported characteristics (see Fig. 7B for example). This interpretation is supported by experimental evidence [26], but we cannot exclude the possibility that the retina also generates H-events, or that L-events are generated in the cortex.

Reduction of the weight dynamics to two dimensions
To reduce to reduce the full weight dynamics to a two-dimensional system, we averaged all the \( n \) weights belonging to the receptive field that are predicted to potentiate along the initial topological bias, as \( w_{\text{RF}} \), and all the \( N_u - n \) remaining weights, which we call complementary to the receptive field, as \( w_{\text{C}} \). When all weights behaved the same, we arbitrarily split them into two groups of the same size. Details about the classification of weights as \( w_{\text{RF}} \) or \( w_{\text{C}} \) can be found in the Appendix.
| Name   | Value/Distribution | Description |
|--------|--------------------|-------------|
| Network |                    |             |
| $N_u$  | 50                 | number of thalamic neurons |
| $N_v$  | 50                 | number of cortical neurons |
| $T$    | 50000              | simulation length [s] |
| Weights |                    |             |
| $w_{\text{ini}}$ | U(0.15,0.25)   | range of initial weights (U: uniform dist.) |
| $s$    | 0.05               | amplitude of Gaussian bias |
| $\sigma_s$ | 4                  | spread of Gaussian bias |
| $w_{\text{max}}$ | 0.5                | weight saturation limit |
| L-events |                  |             |
| $L_{\text{amp}}$ | 1.0               | amplitude (equivalently, binary neuron) |
| $L_{\text{pct}}$ | U(20%,80%)   | percentage of thalamic cells activated |
| $L_{\text{dur}}$ | $N'(0.15,0.015)$ | mean duration [s] ($N$: Gaussian dist.) |
| $L_{\text{int}}$ | Exp(1.5)       | mean inter-event interval [s] (Exp: exponential dist.) |
| H-events |                  |             |
| $H_{\text{amp}}$ | $N'(6, 2)$      | amplitude |
| $H_{\text{pct}}$ | U(80%,100%)   | percentage of cortical cells activated |
| $H_{\text{dur}}$ | $N'(0.15,0.015)$ | mean duration [s] |
| $H_{\text{int}}$ | Gamma(3,5, 1.0) | mean inter-event interval [s] (Gamma: Gamma dist.) |
| Time constants |            |             |
| $\tau_m$ | 0.01             | membrane time constant [s] |
| $\tau_w$ | 500               | weight-change time constant for Hebbian rule [s] |
| $\tau_w$ | 1000              | weight-change time constant for BCM rule [s] |
| $\tau_\theta$ | 20               | output threshold time constant for BCM rule [s] |
| $\tau_\eta$ | 1                | adaptation time constant [s] |

Table 1. List of parameters used in the model unless stated otherwise.

Computing the strength of simulated H-events

To relate the reduced two-dimensional phase planes to the simulation results, we wrote down the steady state activity of neuron $j$ Eq. 1, which contains the rate gain from H-events relative to L-events, $\langle R_H \rangle$ (also called ‘Strength of H-events’ in Fig. 4E):

$$\langle R_H \rangle = \frac{\langle L_{\text{int}} \rangle \langle H_{\text{amp}} \rangle \langle H_{\text{dur}} \rangle}{\langle L_{\text{amp}} \rangle \langle H_{\text{int}} \rangle \langle H_{\text{amp}} \rangle} = \frac{L_{\text{int}}}{H_{\text{int}}} \langle H_{\text{amp}} \rangle$$

since $\langle L_{\text{dur}} \rangle = \langle H_{\text{dur}} \rangle$ and $\langle L_{\text{amp}} \rangle = 1$.

In the absence of adaptive H-events, for a fixed set of values for $H_{\text{amp}}$ and $L_{\text{int}}$ (as in Table 1) and a chosen $\langle R_H \rangle$ which we called ‘Non-adapted strength of H-events’ in Fig. 4E, we used Eq. 6 to find the $H_{\text{int}}$ value that satisfies the equation. Next, we ran simulations with the same $H_{\text{int}}$ and $L_{\text{int}}$ parameters, but adaptive $H_{\text{amp}}$. We fixed the inter-event intervals of both L- and H-events to their mean values $L_{\text{int}}$ and $H_{\text{int}}$ instead of sampling them from distributions in Fig. 4E. Then we numerically estimated the average amplitude of H-events with adaptation which we called ‘Adapted strength of H-events’ in Fig. 4E at the end of the simulation (final 5% of the simulation time) when the dynamics were stationary.

Receptive field statistics

The following receptive field statistics were used to quantify properties of the weight matrix $W$ after the developing weights became stable.

Receptive field size

The receptive field of a cortical neuron is the group of weights from thalamic cells for which $w_{ij} > w_{\text{max}}/5$. The lower threshold was chosen to make the measurement robust to small fluctuations around 0, which are present because of the soft bounds. Mathematically, we compute the receptive field size of cortical neuron $j$ as:

$$R_F(w_j) = \frac{1}{N_u} \sum_{i=1}^{N_u} I_i$$

(7)
with the \( \mathbf{I} \) vector given by:

\[
\mathbf{I}_i = \begin{cases} 
1, & w_{ij} > w_{\text{max}}/5 \\
0, & \text{otherwise}.
\end{cases}
\]  (8)

The normalized receptive field ranges from 0 corresponding to a total decoupling of the cortical cell from the input layer, to 1 corresponding to no selectivity due to the potentiation of all weights from the input layer to that neuron. To compute the average receptive field size of the network, we include only the cortical neurons (\( N^* \)) that have not decoupled:

\[
R_F(W) = \frac{1}{N^*} \sum_{j=1}^{N^*} R_F(w_j).
\]  (9)

If all the cortical cells have decoupled from the thalamus, we set \( R_F(W) = 0 \).

**Topography \( T \)**

The topography of the network is a measure of how much of the initially weak biased topography is preserved in the final receptive field. Due to our biased initial conditions, neighboring thalamic cells are expected to project to neighboring cortical cells, yielding a diagonal weight matrix. For each cortical neuron, we calculated how far the center of its receptive field is from the ideal diagonal. Mathematically, for each row \( j \) of \( W \), we determined the center of the receptive field \( c_j \) and calculated the smallest distance (while considering periodic boundary conditions) between the receptive field center and the diagonal element \( j \). Then, we summed all the squared distances and calculated the average error of the topography:

\[
\xi = \frac{1}{N^u} \sum_{j=1}^{N^u} |c_j - j|^2.
\]  (10)

To normalize the topography, we compared \( \xi \) to the topography error \( \Xi \) of a column receptive field (Fig. S1A) where the centers of all cortical receptive fields were the same, \( c_j = c \) (\( c \) a constant). For such a column receptive field, \( \Xi = \frac{N^2}{12} \). Therefore, we define the topography score \( T \) as:

\[
T = 1 - \frac{\xi}{\Xi}.
\]  (11)

The topography will be close to 1 if the weight matrix is perfectly diagonal, and 0 if the final receptive field is a column (\( \xi = \Xi \)).

**Proportion of cortical decoupling \( D \)**

To quantify the cortical decoupling we use Eq. 8 to compute the fraction of decoupled neurons divided by the number of neurons, \( \frac{1}{N^u} \sum_{j=1}^{N^u} \prod_{i=1}^{N^u} (1 - \mathbf{I}_{ij}) \). If the decoupling is 0, no cortical neuron has decoupled from the thalamus, while decoupling of 1 means that all the cortical neurons are decoupled from the thalamus.

**Quantifying adaptation in the data**

We first investigated if fluctuations in the activity across recordings could generate significant correlations. We analyzed consecutive recordings (each \( \sim5 \) mins long) in the same animal of which we had between 3 and 14 in all 26 animals (separated by \( \sim5 \) mins due to experimental constraints on data collection) to identify possible fluctuations on a longer timescale. We found that the average amplitude of all (L and H) events is not significantly different across consecutive recordings of the same animal (Fig. S3A, one-way ANOVA tests, \( p > 0.05 \) in 23 out of 26 animals). Across different animals and ages, individual event amplitudes remained uncorrelated between successive recordings at this timescale, which we confirmed by plotting the difference in event amplitude as a function of the time between recordings (Fig. S3B), Kruskal-Wallis test, \( p > 0.05 \).
For our reanalysis of the spontaneous events (Fig. 5), we only included events that recruited at least 20% of the cells in the imaging field of view following [26]. We computed the average amplitude of all events that occurred within a time window $\tau_{max}$ before an H-event (consecutive recordings were concatenated) and compared it to the amplitude of the H-event. We excluded animals that had fewer than 12 H-events preceded by spontaneous activity within the time window $\tau_{max}$ (9 animals remained after exclusion). Next, we computed the correlation coefficient of the relationship between H-event amplitude and the average amplitude of preceding activity within $\tau_{max}$ with a leaky accumulator of time constant $\tau_{\text{decay}}$. To estimate the 95% confidence interval, we performed a bootstrap analysis in which we generated 1000 bootstrap datasets by drawing without replacement from the valid pairs of H-event amplitudes and average amplitude of preceding activity. We repeated this analysis with different thresholds for excluding data (Fig. S4A,B), different values of the time window $\tau_{max}$ within which events are averaged (Fig. S4C) and for different decay time constants $\tau_{\text{decay}}$ (Fig. S4D). All data and analysis code can be found at github.com/comp-neural-circuits/LH-events.

Spontaneous events identification in the model

To quantify the properties of spontaneous activity in the cortical layer of our model, we used the time series of activity of all the simulated cortical neurons (after weight stabilization is achieved) sampled in a high time resolution (0.01 s, Fig. 7B). We defined a global activity threshold $\nu = v_{\text{max}}/r$, where $v_{\text{max}}$ is the highest amplitude among the cortical cells in the recording and $r$ is a fixed scaling constant ($r = 8$ for all recordings). For each cortical cell $j$, we labeled the intervals where the cell was active (1) or inactive (0) based on:

$$x_j(t) = \begin{cases} 
1, & \text{if } v_j(t) \geq \nu, \\
0, & \text{otherwise}.
\end{cases}$$

(12)

We then used the trace $X(t) = \sum_{j=1}^{N} x_j(t)$ to define the number of active cortical cells at each time step $t$. If at time step $t = \tau$ the proportion of active cells is between 20-80% of the total number $N_v$, then we label $X(\tau)$ as an effective L-event, otherwise if the proportion of active cells is higher than 80%, we label $X(\tau)$ as an effective H-event. The duration of each event was then the number of continuous time steps classified as either L- or H-event, and the inter-event interval was calculated based on the number of time steps between two events of the same type. For each identified event, we averaged the amplitude of the active cells to obtain the amplitude vs. participation rate relationship.

Acknowledgments

This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (Grant agreement No. 804824 to JG). This work was further supported by the Max Planck Society (MW, JHK, LYC, JG), a NARSAD Young Investigator Grant from the Brain and Behavior Research Foundation (JG), a Capes-Humboldt Research fellowship (MW), the Smart Start joint training program in computational neuroscience (JHK), and by grants of the Netherlands Organization for Scientific Research (NWO, ALW Open Program grants, no. 819.02.017, 822.02.006 and ALWOP.216; ALW Vici, no. 865.12.001) and the "Stichting Vrienden van het Herseninstituut" (NZ, CL). We thank Stephen Eglen for discussions and ideas in the initial stages of the project.

References

1. D. A. Feldheim and D. D. O’Leary. Visual map development: bidirectional signaling, bifunctional guidance molecules, and competition. Cold Spring Harb Perspect Biol, 2:a001768, 2010.

2. J. B. Ackman and M. C. Crair. Role of emergent neural activity in visual map development. Curr Opin Neurobiol, 24:166–175, 2014.

3. L.M.A. Richter and J. Gjorgjieva. Understanding neural circuit development through theory and models. Curr Opin Neurobiol, 46:39–47, 2017.
4. A. Thompson, A. Gribizis, C. Chen, and M. C. Crair. Activity-dependent development of visual receptive fields. *Curr Opin Neurobiol*, 42:136–143, 2017.

5. A. G. Blankenship and M. B. Feller. Mechanisms underlying spontaneous patterned activity in developing neural circuits. *Nat Rev Neurosci*, 11:18–29, 2010.

6. H. C. Wang and D. E. Bergles. Spontaneous activity in the developing auditory system. *Cell Tissue Res*, 361:65–75, 2015.

7. A. D. Huberman, M. B. Feller, and B. Chapman. Mechanisms underlying development of visual maps and receptive fields. *Ann Rev Neurosci*, 31:479–509, 2008.

8. W. J. Moody and M. M. Bosma. Ion channel development, spontaneous activity, and activity-dependent development in nerve and muscle cells. *Physiol Rev*, 85:883–941, 2005.

9. L. A. Kirkby, G. S. Sack, A. Firl, and M. B. Feller. A role for correlated spontaneous activity in the assembly of neural circuits. *Neuron*, 80:1129–1144, 2013.

10. J. Cang, M. Kaneko, J. Yamada, G. Woods, M. P. Stryker, and D. A. Feldheim. Ephrin-As guide the formation of functional maps in the visual system. *Neuron*, 94:577–589, 2005.

11. H. P. Xu, M. Furman, Y. S. Mineur, H. Chen, S. L. King, D. Zenisek, Z. J. Zhou, D. A. Butts, N. Tian, M. R. Picciotto, and M. C. Crair. An instructive role for patterned spontaneous retinal activity in mouse visual map development. *Neuron*, 70:1115–1127, 2011.

12. T. J. Burbridge, H.-P. Xu, J. B. Ackman, X. Ge, Y. Zhang, M.-J. Ye, Z. J. Zhou, J. Xu, A. Contractor, and M. C. Crair. Visual circuit development requires patterned activity mediated by retinal acetylcholine receptors. *Neuron*, 84:1049–1064, 2014.

13. J. Gjorgjieva, T. Toyoizumi, and S. J. Eglen. Burst-time-dependent plasticity robustly guides ON/OFF segregation in the lateral geniculate nucleus. *PLoS Comp Biol*, 5:e1000618, 2009.

14. G. Goodhill and J. Xu. The development of retinotectal maps: A review of models based on molecular gradients. *Network: Computation in Neural Systems*, 16:5–34, 2005.

15. G. J. Goodhill. Contributions of theoretical modeling to the understanding of neural map development. *Neuron*, 56:301–311, 2007.

16. F. Grimbert and J. Cang. New model of retinocollicular mapping predicts the mechanisms of axonal competition and explains the role of reverse molecular signaling during development. *J Neurosci*, 32:9755–9768, 2012.

17. K. B. Godfrey and N. V. Swindale. Modeling development in retinal afferents: Retinotopy, segregation, and Ephrinx/Epha mutants. *PloS ONE*, 9:e104670, 2014.

18. H. J. Luhmann, A. Sinning, J.-W. Yang, V. Reyes-Puerta, M. C. Stüttgen, S. Kirischuk, and W. Kilb. Spontaneous neuronal activity in developing neocortical networks: From single cells to large-scale interactions. *Front Neural Circuits*, 10:40:doi: 10.3389/fncir.2016.00040, 2016.

19. O. Garaschuk, J. Linn, J. Eilers, and A. Konnerth. Large-scale oscillatory calcium waves in the immature cortex. *Nat Neurosci*, 3:452–459, 2000.

20. J. W. Lischalk, C. R. Easton, and W. J. Moody. Bilaterally propagating waves of spontaneous activity arising from discrete pacemakers in the neonatal mouse cerebral cortex. *Devel Neurobiol*, 69:407–414, 2009.

21. 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. *Devel Neurobiol*, 70:679–92, 2010.
22. S. Namiki, H. Norimoto, C. Kobayashi, K. Nakatani, N. Matsuki, and Y. Ikegaya. Layer iii neurons control synchronized waves in the immature cerebral cortex. *J Neurosci*, 33:987–1001, 2013.

23. H. Adelsberger, O. Garaschuk, and A. Konnerth. Cortical calcium waves in resting newborn mice. *Nat Neurosci*, 8:988–990, 2005.

24. J. B. Ackman, T. H Burbridge, and M. C. Crair. Retinal waves coordinate patterned activity throughout the developing visual system. *Nature*, 490:219–225, 2012.

25. 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*, 104:1–13, 2019.

26. 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*, 22:253–258, 2012.

27. I. L. Hanganu, Y. Ben Ari, and R. Khazipov. Retinal waves trigger spindle bursts in the neonatal rat visual cortex. *J Neurosci*, 26:6728–6736, 2006.

28. 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*, 272:1182–1187, 1996.

29. K. D. Miller, J. B. Keller, and M. P. Stryker. Ocular dominance column development: Analysis and simulation. *Science*, 245:605–615, 1989.

30. K. D. Miller. Models of activity-dependent neural development. *Seminars in Neuroscience*, 4:61–73, 1992.

31. C. W. Lee, S. J. Eglen, and R. O. L. Wong. Segregation of ON and OFF retinogeniculate connectivity directed by patterned spontaneous activity. *J Neurophysiol*, 88:2311–2321, 2002.

32. E. L. Bienenstock, L. N. Cooper, and P. W. Munro. Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. *J Neurosci*, 2:32–48, 1982.

33. N. L. Rochefort, O. Garaschuk, R.-I. Milos, M. Narushima, N. Marandi, B. Pichler, Y. Kovalechuk, and A. Konnerth. Sparsification of neuronal activity in the visual cortex at eye-opening. *Proc Natl Acad USA*, 106:15049–15054, 2009.

34. P. Golshani, J. T. Goncalves, S. Khoshkhoo, R. Mostany, S. Smirnakis, and C. Portera-Cailliau. Internally mediated developmental desynchronization of neocortical network activity. *J Neurosci*, 29:10890–10899, 2009.

35. A. H. Leighton, G. J. Houwen, J. E. Cheyne, P. P. Maldonado, F. De Winter, and C. Lohmann. Control of spontaneous activity patterns by inhibitory signaling in the developing visual cortex. *biorXiv*, 2020.

36. D. A. Butts and P. O. Kanold. The applicability of spike time dependent plasticity to development. *Front Syn Neurosci*, 2:30:1–9, 2010.

37. 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*, 5:0651–0661, 2007.

38. J. Winnubst, J. E. Cheyne, D. Niculescu, and C. Lohmann. Spontaneous activity drives local synaptic plasticity in vivo. *Neuron*, 87:399–410, 2015.

39. G. G. Turrigiano and S. B. Nelson. Homeostatic plasticity in the developing nervous system. *Nat Rev Neurosci*, 5:97–107, 2004.

40. G. Tononi and C. Cirelli. Sleep function and synaptic homeostasis. *Sleep Med Rev*, 10:49–62, 2006.

41. D. J. C. MacKay and K. D. Miller. Analysis of Linsker’s application of Hebbian rules to linear networks. *Network*, 1:257–297, 1990.
42. M. Chistiakova, N. M. Bannon, M. Bazhenov, and M. Volgushev. Heterosynaptic plasticity multiple mechanisms and multiple roles. *Neuroscientist*, 20:483–498, 2014.

43. G. S. Lynch, T. Dunwiddie, and V. Gribkoff. Heterosynaptic depression: A postsynaptic correlate of long-term potentiation. *Nature*, 266:737–739, 1977.

44. F. Zenke, E. J. Agnes, and W. Gerstner. Diverse synaptic plasticity mechanisms orchestrated to form and retrieve memories in spiking neural networks. *Nat Commun*, 6:6922, 2015.

45. G. López-Bendito. Development of the thalamocortical interactions: Past, present and future. *Neuroscience*, 385:67–74, 2018.

46. G. J. Goodhill. Can molecular gradients wire the brain? *Trends Neurosci*, 39:202–211, 2016.

47. N. S. Desai, L. C. Rutherford, and G. G. Turrigiano. Plasticity in the intrinsic excitability of cortical pyramidal neurons. *Nat Neurosci*, 2:515–520, 1999.

48. G. Daoudal and D. Debanne. Long-term plasticity of intrinsic excitability: learning rules and mechanisms. *Learn Mem*, 10:456–465, 2003.

49. N. Tien and D. Kerschensteiner. Homeostatic plasticity in neural development. *Neural Development*, 13, 2018.

50. J. Lombardo, J. Sun, and M. A. Harrington. Rapid activity-dependent modulation of the intrinsic excitability through up-regulation of kcnq/kv7 channel function in neonatal spinal motoneurons. *PLoS ONE*, 13:e0193948, 2018.

51. D. A. Butts and D. S. Rokhsar. The information content of spontaneous retinal waves. *J Neurosci*, 21(3):961–973, 2001.

52. 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 β2 subunit of the nicotinic acetylcholine receptor. *Neuron*, 40:1161–1172, 2003.

53. 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*, 48:797–809, 2005.

54. C. Sun, D. K. Warland, J. M. Ballesteros, D. van der List, and L. M Chalupa. Retinal waves in mice lacking the β2 subunit of the nicotinic acetylcholine receptor. *Proc Natl Acad Sci USA*, 105:13638–13643, 2008.

55. 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*, 64:200–212, 2009.

56. C. S. Cutts and S. J. Eglen. Detecting pairwise correlations in spike trains: An objective comparison of methods and application to the study of retinal waves. *J Neurosci*, 34:14288–14303, 2014.

57. C. G. Frye and J. N. MacLean. Spontaneous activations follow a common developmental course across primary sensory areas in mouse neocortex. *Jornal of Neurophysiology*, 116:431–437, 2016.

58. S. J. Etherington and S. R. Williams. Postnatal development of intrinsic and synaptic properties transforms signaling in the layer 5 excitatory neural network of the visual cortex. *J Neurosci*, 31:9526–9537, 2011.

59. P. Berkes, B. White, and J. Fiser. No evidence for active sparsification in the visual cortex. In Y. Bengio, D. Schuurmans, J. D. Lafferty, C. K. I. Williams, and A. Culotta, editors, *Advances in Neural Information Processing Systems 22*, pages 108–116. Curran Associates, Inc., 2009.

60. P. Berkes, G. Orbán, M. Lengyel, and J. Fiser. Spontaneous cortical activity reveals hallmarks of an optimal internal model of the environment. *Science*, 331:83–87, 2011.
61. D. J. Willshaw and C. von der Malsburg. How patterned neural connections can be set up by self-organization. Proc R Soc Lond Ser B, 194:431–445, 1976.

62. A. Sánchez-Aguilera, J. L. Sánchez-Alonso, M. A. Vicente-Torres, and A. Colino. A novel short-term plasticity of intrinsic excitability in the hippocampal ca1 pyramidal cells. The Journal of Physiology, 592:2845–2864, 2014.

63. T. A. Babola, S. Li, A. Gribizis, B. J. Lee, J. B. Issa, H. C. Wang, M. C. Crair, and D. E. Bergles. Homeostatic control of spontaneous activity in the developing auditory system. Neuron, 99:511–524, 2018.

64. G. G. Turrigiano, K. R. Leslie, N. S. Desai, L. C. Rutherford, and S. B. Nelson. Activity-dependent scaling of quantal amplitude in neocortical neurons. Nature, 391:892–895, 1998.

65. J. Tabak, J. Rinzel, and M. J. O’Donovan. The role of activity-dependent network depression in the expression and self-regulation of spontaneous activity in the developing spinal cord. J Neurosci, 21:8966–8978, 2001.

66. J. Tabak, M. Mascagni, and R. Bertram. Mechanism for the universal pattern of activity in developing neuronal networks. J Neurophysiol, 103:2208–2221, 2010.

67. J. Xu and L. G. Wu. The decrease in the presynaptic calcium current is a major cause of short-term depression at a calyx-type synapse. Neuron, 46:633–645, 2005.

68. R. Renden, H. Taschenberger, N. Puente, D. A. Rusakov, R. Duvoisin, L.-Y. Wang, K. P. Lehre, and H. von Gersdorff. Glutamate transporter studies reveal the pruning of metabotropic glutamate receptors and absence of AMPA receptor desensitization at mature calyx of Held synapses. Journal of Neuroscience, 25:8482–8497, 2005.

69. Y. Nakamura and T. Takahashi. Developmental changes in potassium currents at the rat calyx of Held presynaptic terminal. Journal of Physiology, 581:1101–1112, 2007.

70. A.-M. M. Oswald and A. D. Reyes. Maturation of intrinsic and synaptic properties of layer 2/3 pyramidal neurons in mouse auditory cortex. Journal of Neurophysiology, 99:2998–3008, 2008.

71. D. Niculescu and C. Lohmann. Gap junctions in developing thalamic and neocortical neuronal networks. Cerebral Cortex, 24:3097–3106, 2014.

72. R. Cachope, K. Mackie, A. Triller, J. O’Brien, and A. E. Pereda. Potentiation of electrical and chemical synaptic transmission mediated by endocannabinoids. Neuron, 56:1034–1047, 2007.

73. J. Haas, B. Zavala, and C. E. Landisman. Activity-dependent long-term depression of electrical synapses. Science, 334:389–393, 2011.

74. Y. Ben-Ari. Excitatory actions of GABA during development: The nature of the nurture. Nature Reviews Neuroscience, 3:728–739, 2002.

75. Y. Murata and M. T. Colonnese. GABAergic interneurons excite neonatal hippocampus in vivo. Sci Adv, 6:eaba1430, 2020.

76. K. Kirmse, M. Kummer, Y. Kovalchuk, O. W. Witte, O. Garaschuk, and K. Holthoff. GABA depolarizes immature neurons and inhibits network activity in the neonatal neocortex in vivo. Nature Comm, 6:7750, 2015.

77. G. Valeeva, T. Tressard, M. Mukhtarov, A. Baude, and R. Khazipov. An optogenetic approach for investigation of excitatory and inhibitory network GABA actions in mice expressing channelrhodopsin-2 in GABAergic neurons. The Journal of Neuroscience, 36:5961–5973, 2016.

78. R. Tremblay, S. Lee, and B. Rudy. GABAergic interneurons in the neocortex: From cellular properties to circuits. Neuron, 91:260–292, 2016.
79. A. L. Dorn, K. Yuan, A. J. Barker, C. E. Schreiner, and R. C. Froemke. Developmental sensory experience balances cortical excitation and inhibition. *Nature*, 465(7300):932–936, 2010.

80. D. J. Field. What is the goal of sensory coding? *Neural Comput*, 6:559–601, 1994.

81. B. Willmore and D. J. Tolhurst. Characterizing the sparseness of neural codes. *Network: Comput Neural Syst*, pages 255–270, 2001.

82. B. A. Olshausen and D. J. Field. Sparse coding of sensory inputs. *Current Opinion in Neurobiology*, 14:481–487, 2004.

83. J. Zyllberg and M. R. DeWeese. Sparse coding models can exhibit decreasing sparseness while learning sparse codes for natural images. *PLOS Comput Biol*, 9:e1003182, 2013.

84. H. Ko, S. B. Hofer, B. Pichler, K. A. Buchanan, P. J. Sjöström, and T. D. Mrsic-Flogel. Functional specificity of local synaptic connections in neocortical networks. *Nature*, 473:87–90, 2011.

85. V. V. Vyazovskiy, C. Cirelli, M. Pfister-Genskow, U. Faraguna, and G. Tononi. Molecular and electrophysiological evidence for net synaptic potentiation in wake and depression in sleep. *Nat Neurosci*, 11:200–208, 2008.

86. K. C. Wood, J. M. Blackwell, and M. N. Geffen. Cortical inhibitory interneurons control sensory processing. *Current Opinion in Neurobiology*, 46:200–207, 2017.

87. S. Fujimoto, M. N. Leiwe, R. Sakaguchi, Y. Muroyama, R. Kobayakawa, K. Kobayakawa, T. Saito, and T. Imai. Spontaneous activity generated within the olfactory bulb establishes the discrete wiring of mitral cell dendrites. *biorXiv*, 2019.

88. A. Kirkwood, M. G. Rioult, and M. F. Bear. Experience-dependent modification of synaptic plasticity in visual cortex. *Nature*, 381:526–528, 1996.

89. P. J. Sjöström, G. G. Turrigiano, and S. B. Nelson. Rate, timing and cooperativity jointly determine cortical synaptic plasticity. *Neuron*, 32:1149–1164, 2001.

90. F. Zenke, W. Gerstner, and S. Ganguli. The temporal paradox of hebbian learning and homeostatic plasticity. *Curr Op in Neurobiol*, 43:166–176, 2017.

91. H. Naoki. Revisiting chemoaffinity theory: Chemotactic implementation of topographic axonal projection. *PLoS computational biology*, 13:e1005702, 2017.

92. G. B. Smith, A. Sederberg, Y. M. Elyada, S. D. van Hooser, M. Kaschube, and D. Fitzpatrick. The development of cortical circuits for motion discrimination. *Nat Neurosci*, 18:252–261, 2015.

93. K. Ikezoe, H. Tamura, F. Kimura, and I. Fujita. Decorrelation of sensory-evoked neuronal responses in rat barrel cortex during postnatal development. *Neuroscience Research*, 73:312–320, 2012.

94. J. Shen and M. T. Colonnese. Development of activity in mouse visual cortex. *J Neurosci*, 36:12259–12275, 2016.

95. N. Ji. Adaptive optical fluorescence microscopy. *Nature Methods*, 14:374–380, 2017.

96. R. M. Gray. Toeplitz and circulant matrices: A review. *Foundations and Trends in Communications and Information Theory*, 2:155–239, 2006.

97. K. D. Miller and D. J. C. MacKay. The role of constraints in Hebbian learning. *Neural Comp*, 6:100–126, 1994.
Appendix

The weight dynamics under the Hebbian rule

Since synaptic plasticity operates on a much slower time scale than the response dynamics of the output neuron, we make a steady state assumption and write Eq. 1 for neuron $j$ as:

$$v_j(t) = u^T(t)w_j(t) + \langle R_H \rangle(t), \quad (A1)$$

where $w_j$ is the vector of the elements in the row of the weight matrix $W$, i.e. the vector of weights from all thalamic neurons into cortical neuron $j$. $\langle R_H \rangle(t)$ is the rate gain from H-events relative to L-events, which depends on the duration, amplitude, and inter-event intervals of both L- and H-events. Specifically, $\langle R_H \rangle$ is proportional to $\langle H_{\text{amp}} \rangle$, $\langle H_{\text{dur}} \rangle$, $\langle L_{\text{int}} \rangle$ and inversely proportional to $\langle H_{\text{int}} \rangle$, $\langle L_{\text{amp}} \rangle$ and $\langle L_{\text{dur}} \rangle$ (where $\langle \rangle$ denotes an ensemble average over the activity patterns), such that:

$$\langle R_H \rangle = \frac{\langle L_{\text{int}} \rangle \langle H_{\text{amp}} \rangle \langle H_{\text{dur}} \rangle}{\langle H_{\text{int}} \rangle \langle L_{\text{amp}} \rangle \langle L_{\text{dur}} \rangle}. \quad (A2)$$

Rewriting Eq. 2 in vector form:

$$\tau_u \frac{d w_j(t)}{dt} = \langle v_j(t) (u(t) - \theta_u) \rangle. \quad (A3)$$

Inserting Eq. A1 into the weight dynamics from Eq. A3 yields (vector dependence on $t$ is omitted for clarity):

$$\tau_u \frac{d w_j}{dt} = \langle (u^T w + R_H) (u - \theta_u) \rangle = C w_j - \langle \theta_u - \langle u \rangle \rangle \langle u^T \rangle w_j - \langle R_H \rangle \theta_u = C'(\theta_u) w_j - \langle R_H \rangle \theta_u, \quad (A4)$$

where $C = Q - \langle u \rangle^2$ is the input covariance matrix, $Q = \langle uu^T \rangle$ is the input correlation matrix and we have defined $C'(\theta_u) = C - \langle u \rangle (\theta_u - \langle u \rangle) = Q - \langle u \rangle \theta_u$ to be the ‘modified covariance matrix’ of the learning rule. We write $\langle u \rangle$ to denote the vector with repeated element $\langle u \rangle$, which is the mean normalized size of an L-event (e.g., for L-events engaging 20–80% of input neurons, $\langle u \rangle = 0.5$). We also used the fact that the occurrence of L- and H-events is uncorrelated [26], such that $\langle R_H u \rangle = 0$.

The input correlation matrix

The spontaneous L-events in our network model have useful mathematical properties that allowed us to derive an analytical expression for the elements of the input correlation matrix $Q$. The locality and periodicity of L-events, as well as the long-term averaging of their stationary dynamics generate a symmetric and circulant matrix $Q$ [96]. In a circulant matrix, each row is rotated one element to the right relative to the preceding row, and thus $Q$ can be completely defined by a vector.

Let $\ell$ be the fixed size of an L-event between 0 and $N_u$. By computing the correlations among all the possible L-events of size $\ell$, we can write the elements of the vector $q_{\ell k} = \langle \ell q_{1}, q_{2}, \ldots, q_{N_u} \rangle$, which completely defines $Q$, as:

$$q_{\ell k} = \frac{\ell - \min (N_u - k - 1, k - 1)}{N_u}, \quad (A5)$$

where $\min(a, b)$ returns the minimum of $a$ and $b$. Since we wanted to explore the dependence of refinement on the size of L-events, defined by the minimum ($\ell_{\text{min}}$) and maximum ($\ell_{\text{max}}$) number of cells they activate in the input layer, we average over the size of a given event:

$$q_k = \frac{1}{\ell_{\text{max}} - \ell_{\text{min}}} \sum_{\ell=\ell_{\text{min}}}^{\ell_{\text{max}}} \ell - \min (N_u - k - 1, k - 1) \min (N_u - \ell, \ell) \frac{1}{N_u}. \quad (A6)$$

Finally, using the emergent symmetry of $Q$, we can simplify the elements of this vector as follows:

$$q_k = q - \min (k - 1, N_u - k + 1) \delta_q, \quad \text{for all} \quad k = 1, 2, \ldots, N_u, \quad (A7)$$

with $q = (\ell_{\text{max}} + \ell_{\text{min}})/2N_u$ and $\delta_q = 1/N_u$.

Since $C(\theta_u)$ and $C'(\theta_u)$ are defined in terms of $Q$ minus a constant, both are circulant matrices as well.
Fixed point of the weight dynamics

The fixed point of the weight dynamics under L- and H-events is obtained by setting Eq. A4 to zero and solving the resulting equation for $w^*$:

$$w^* = (R_{H})\theta_u C'(\theta_u)^{-1}n,$$  \hspace{1cm} (A8)

where $n$ is a vector of ones. To study the nature of the fixed point $w^*$, we need to investigate the eigenvalues of the circulant matrix $C'(\theta_u)^{-1}$ (the inverse of a circulant matrix is also a circulant matrix). A circulant matrix has the property that its eigenvectors can be written in terms of roots of unity. The eigenvalues are real and can be written as the discrete Fourier transforms of any row of the matrix [96]. In particular, one of the eigenvalues is the sum of the elements of any row of the matrix, with a corresponding eigenvector that is a constant. We call this special eigenvalue the 'row-sum eigenvalue'. All eigenvalues except for the row-sum are non-negative.

We can write the fixed point as $w^* = (\lambda^*)^{-1}\theta_u(R_{H})n$, where $\lambda^*$ is the row-sum eigenvalue of $C'(\theta_u)$. It is clear that if only L-events are present, $(R_{H}) = 0$, the fixed point is always at the origin, $w^* = 0$. Consequently, adding H-events in the cortical layer moves the location of the fixed point along the diagonal of the phase plane, i.e. equally in all directions. The fixed point will be positive (negative) when $\lambda^*$ is positive (negative). Furthermore, it will be an unstable fixed point for $\lambda^* > 0$, since all the eigenvalues of the dynamical system are positive. The fixed point is a saddle node when $\lambda^* < 0$ since the other eigenvalues are positive.

Calculation of receptive field size

The Hebbian rule on its own does not have a mechanism to prevent the weights to grow infinitely large or negative. Thus, to generate receptive fields we imposed a lower bound at 0 and an upper bound at $w_{\text{max}}$. This enabled us to calculate the size of the receptive field, $n$, as a function of L-event properties and the input threshold, $\theta_u$ in the absence of H-events. Because the cortical cells are not recurrently connected, here we examine the weight vector onto a single cortical neuron. Assuming the dynamical system has reached steady state, to get a receptive field of size $n < N_u$ for this cortical neuron, we can write the fixed point weight vector as $w^* = (w_{\text{max}}, \ldots, w_{\text{max}}, 0, \ldots, 0)^\tau$, where $n$ of the weights have reached the upper bound $w_{\text{max}}$, while the remaining weights the lower bound 0. The specific weight identity reaching the upper bound is not important, as long as they are topographically near each other. To achieve this fixed point:

$$\dot{w}_i\big|_{w=w^*} = \begin{cases} 
0, & 1 \leq i \leq n \\
< 0, & i > n.
\end{cases}$$  \hspace{1cm} (A9)

Using the structure of our circulant correlation matrix (Eq. A7), we can rewrite Eq. A4 (note that in the absence of H-events $(R_{H}) = 0$) as:

$$\tau_w\begin{bmatrix}
\dot{w}_1 \\
\vdots \\
\dot{w}_n \\
\dot{w}_{n+1}
\end{bmatrix} = \begin{bmatrix}
q_1 - \langle u \rangle \theta_u & \ldots & q_n - \langle u \rangle \theta_u & q_{n+1} - \langle u \rangle \theta_u & \ldots & q_{N_u} - \langle u \rangle \theta_u \\
\vdots & & \vdots & \vdots & & \vdots \\
q_n - \langle u \rangle \theta_u & \ldots & q_1 - \langle u \rangle \theta_u & q_2 - \langle u \rangle \theta_u & \ldots & q_{n+1} - \langle u \rangle \theta_u \\
q_{n+1} - \langle u \rangle \theta_u & \ldots & q_2 - \langle u \rangle \theta_u & q_1 - \langle u \rangle \theta_u & \ldots & q_{n+2} - \langle u \rangle \theta_u \\
\vdots & & \vdots & \vdots & & \vdots \\
q_2 - \langle u \rangle \theta_u & \ldots & q_{n+1} - \langle u \rangle \theta_u & q_{n+2} - \langle u \rangle \theta_u & \ldots & q_1 - \langle u \rangle \theta_u \\
\vdots & & \vdots & \vdots & & \vdots \\
q_n - \langle u \rangle \theta_u & \ldots & q_{N_u} - \langle u \rangle \theta_u & 0 & \ldots & 0
\end{bmatrix}\begin{bmatrix}
w_{\text{max}} \\
w_{\text{max}} \\
w_{\text{max}} \\
w_{\text{max}} \\
w_{\text{max}} \\
w_{\text{max}} \\
w_{\text{max}}
\end{bmatrix}.$$  \hspace{1cm} (A10)
Now, we study the conditions to guarantee that Eq. A9 is satisfied. Computing each equation explicitly, we obtain:

\[
\begin{align*}
\tau_u \dot{w}_1 &= (q_1 + \ldots + q_n)w_{\text{max}} - n\langle u \rangle \theta_u w_{\text{max}} \\
&\vdots \\
\tau_u \dot{w}_n &= (q_n + \ldots + q_1)w_{\text{max}} - n\langle u \rangle \theta_u w_{\text{max}} \\
\tau_u \dot{w}_{n+1} &= (q_{n+1} + \ldots + q_2)w_{\text{max}} - n\langle u \rangle \theta_u w_{\text{max}} \\
&\vdots \\
\tau_u \dot{w}_{N_u} &= (q_2 + \ldots + q_{n+1})w_{\text{max}} - n\langle u \rangle \theta_u w_{\text{max}}.
\end{align*}
\] (A11)

And finally, using Eq. A7 and after some manipulations this can be written as:

\[
\tau_u \dot{w}_i = \left\{ nq - \left[ \frac{i(i-1)}{2} + \frac{(n-i)(n-i+1)}{2} \right] \delta_q - n\langle u \rangle \theta_u \right\} w_{\text{max}}, \quad i = 1, 2, \ldots, N_u.
\] (A12)

To determine the input threshold \( \theta_u \) in Eq. A12 that yields a receptive field of size \( n \), we assume that \( \dot{w}_i|_{w=w_{\text{max}}} = 0 \) for \( i = n \); this implies that \( \dot{w}_i|_{w=w_{\text{max}}} > 0 \) for \( i < n \) and \( \dot{w}_i|_{w=w_{\text{max}}} < 0 \) for \( i > n \). We write \( \theta_u \) as a linear combination of \( q \) and \( \delta_q \):

\[
\theta_u = \alpha q + \beta \delta_q.
\] (A13)

Using this ansatz in Eq. A12 and setting it to zero for \( i = n \) (since \( w_{\text{max}} > 0 \)), we find that \( \alpha = 1/\langle u \rangle \) and \( \beta = -(n-1)/(2\langle u \rangle) \). Therefore, only in the presence of L-events, we derive the input threshold at which the resulting receptive field size is \( n \):

\[
\theta_u^n = \frac{q - (n-1)\delta_q}{\langle u \rangle}.
\] (A14)

Plugging this threshold into Eq. A12 for all \( i \), we get a quadratic polynomial in \( i \):

\[
\tau_u \dot{w}_i = (-i^2 + (n+1)i - n) \delta_q w_{\text{max}}, \quad i = 1, 2, \ldots, N_u.
\] (A15)

Since \( \delta_q, w_{\text{max}} > 0 \), indeed \( 1 \leq i \leq n \) results in \( \dot{w}_i \geq 0 \) while \( i > n \) yields \( \dot{w}_i < 0 \), thus, satisfying Eq. A9.

In the absence of H-events, we computed the average size of receptive fields using Eq. A14 for a range of input thresholds \( \theta_u \) and maximum participation rates of L-events, while keeping the minimum participation rate at 20% (Fig. A1A, contour lines). We verified our analytical predictions with Monte Carlo simulations for the same range of parameters (Fig. A1A). We confirmed that the size of L-events, which depends on the range of participation rates, has a direct impact on receptive field size, with larger L-events resulting in larger receptive fields for a fixed input threshold (Fig. A1B). Low input thresholds generate refined receptive fields only if the size of spontaneous events is small.

The eigenspace of \( C'(\theta_u) \)

To gain intuition for the weight dynamics in Eq. A4, we first investigated the eigenspace of \( C'(\theta_u) \), the vector space spanned by the eigenvectors of \( C'(\theta_u) \). Specifically, we focused on the conditions that enabled the robust formation of cortical receptive fields.

Using the fact that \( Q \) and \( C'(\theta_u) \) are circulant matrices (Fig. A2A, B), we identified two input thresholds, \( \theta^* \) and \( \theta^{**} \), that define three dynamical regions that the row-sum eigenvalue of \( C'(\theta_u) \), \( \lambda^* \), can occupy (Fig. A2C). To obtain the first critical input threshold \( \theta^* \) that characterizes the transition from region (i) to region (ii), we set, for any row \( j \), the row-sum eigenvalue to the largest (fixed) eigenvalue of \( C'(\theta_u) \):

\[
\sum_{i=1}^{N_u} Q_{ji} - N_u \theta^* \langle u \rangle = \lambda_1,
\] (A16)

and for the input statistics of experimentally measured L-events (Table 1), we obtained \( \theta^* = 0.414 \). Similarly, the transition from region (ii) to region (iii) is achieved when the row-sum eigenvalue is set to zero:

\[
\sum_{i=1}^{N_u} Q_{ji} - N_u \theta^{**} \langle u \rangle = 0,
\] (A17)
and the second critical input threshold is obtained as $\theta^{**} = 0.564$.

In region (i), $0 < \theta_u < \theta^*$ and $\lambda^* > 0$ is the dominant (largest) eigenvalue of $C'(\theta_u)$. The eigenvector corresponding to it is a constant (Fig. A2D, (i)). This predicts that all synaptic weights to a given cortical cell will potentiate preventing the formation of a localized receptive field. Since all eigenvalues are non-negative, the fixed point is an unstable node of the linear dynamical system (Eq. A4).

In region (ii), $\theta^* < \theta_u < \theta^{**}$ and $\lambda^* > 0$. However, $\lambda^*$ is no longer the dominant eigenvalue. In this setting, there is a pair of dominant eigenvalues with the corresponding eigenvectors taking the form of out-of-sync sine waves with positive and negative elements. The sign of these elements predicts that some weights will potentiate while others depress, thus enabling the formation of receptive fields (Fig. A2D, (ii)). All eigenvalues in this second region remain non-negative and the fixed point is still an unstable node of the dynamical system.

Finally, in region (iii), $\theta^{**} < \theta_u < 1$ and $\lambda^* < 0$. While the dominant eigenvectors are similar to those in region (ii), enabling the formation of localized receptive fields (Fig. A2D, (iii)), the dynamics of the dynamical system are different because the fixed point is now a saddle node.

**Analysis of the weight dynamics in two dimensions**

We reduced the dimension of the weight dynamics by defining two distinct sets of weights. The first set, $w_{RF}$, corresponds to the $n$ weights which correspond to the topographically biased locations of the receptive field. The complementary set $w_C$ contains the remaining weights. To classify the weights into $w_{RF}$ and $w_C$, first we solved Eq. A4 with the biased initial condition and limited the sum to the dominant eigenvalues and respective eigenvectors. In the case of selectivity, the eigenvectors have half the elements positive and half negative. Therefore, if the receptive field size is $n < N_u/2$ we needed to subsample the potentiating weights to achieve the smaller receptive field size. We did it by keeping only the $n$ largest positive elements in $w_{RF}$ and moving the remaining ones to $w_C$. If $n > N_u/2$, we downsampled $w_C$ by moving the $n - N_u/2$ less-negative weights to $w_{RF}$. Due to the topographically biased initial conditions, $w_{RF}$ always contained the weights potentiating along the diagonal $w_{RF} = w_C$.

We then regularly sampled initial conditions in $[0, w_{\text{max}}] \times [0, w_{\text{max}}]$. For a given receptive field size $n$, we set $\theta_u = \theta^{*}$ according to Eq. A14. If $w_{RF}(0) > w_C(0)$, $w_{RF}$ contains the $n$ weights that form the receptive field by potentiating to the upper bound $w_{\text{max}}$, while $w_C$ contains the remaining $N_u - n$ weights that depress to 0. Similarly, if $w_{RF}(0) < w_C(0)$, $w_C$ contains the $n$ weights that potentiate to the upper bound, while $w_{RF}$ contains the remaining $N_u - n$ weights that depress to 0. At each initial condition, we solved the weight evolution (Eq. A4) for each weight, and averaged the weights in $w_{RF}$ and $w_C$ for a small time interval to
obtain the direction of the phase plane arrows. We computed the evolution trajectory by solving Eq. A4 with a topographically biased initial condition.

**Analytical solution of the two-dimensional weight dynamics with only L-events in the Hebbian rule**

We first examined the weight dynamics in the reduced two-dimensional phase plane $w_{RF} \times w_C$ for only L-events ($\langle R_H \rangle = 0$ in Eq. A4). The phase plane is symmetric about the diagonal $w_{RF} = w_C$ due to the symmetry of the dominant eigenvectors (Fig. A3A). As predicted by the eigenvectors (Fig. A2D), in region (i) both $w_{RF}$ and $w_C$ converge to the upper bound and the fixed point, which is located in the origin, is an unstable node (Fig. A3A, left). Therefore, all weights potentiate and no receptive field can be formed. In regions (ii) and (iii), the eigenvectors predict the formation of receptive fields with $w_{RF} \to w_{max}$ and $w_C \to 0$, respectively (Fig. A3A, middle and right), though the dynamics are different in each case because the origin is an unstable node or a saddle node, respectively.

Our analytical predictions of the reduced two-dimensional system with only L-events were confirmed in numerical simulations of the full $N$-dimensional system. In particular, in region (i) all weights potentiate with each cortical cell receiving input from all thalamic inputs, such that no receptive field forms (Fig. A3B, left). In regions (ii) and (iii), receptive fields form with good topography (Fig. A3B, middle and right). Therefore, consistent with the analytical prediction of receptive field size (Fig. A1), higher input thresholds resulted in smaller receptive fields.

Thus, the Hebbian rule can generate receptive fields of size depending on the input threshold $\theta_u$ in the presence of only L-events originating from the sensory periphery. This result is in agreement with previous findings of the emergence of other aspects of development including topographic maps [61] and ocular dominance [29, 30] or other selectivity [41, 97, 31] in the presence of correlated activity in the input layer of similar feedforward networks. We find that when the only input to the cortex are peripheral L-events, intrinsic properties of the learning rule, such as the threshold between potentiation and depression, control receptive field refinement.
Figure A3. Peripheral L-events generate robust receptive field refinement. A. The reduced two-dimensional weight dynamics in the phase plane with the same $\theta_u$ as Fig. A2B and D. For each plane, the red trajectory depicts the weight evolution from an initial condition where $w_{RF}(0) > w_C(0)$ until the weights’ upper bound ($w_{\text{max}} = 0.5$). Left: $w_C \rightarrow w_{\text{max}}$ and $w_{RF} \rightarrow w_{\text{max}}$, resulting in no selectivity. Middle and right: $w_C \rightarrow 0$ and $w_{RF} \rightarrow w_{\text{max}}$, resulting in selectivity and receptive field refinement. B. Simulation results for the same input thresholds of Fig. A2B, D for the full 50-dimensional system.

Analytical solution of the two-dimensional weight dynamics with L- and H-events in the Hebbian rule

We also studied how the addition of H-events affects network refinements. To investigate the role of H-events in a systematic way, we repeated our analytical study of the weight-dynamics from Eq. A4, but with $\langle R_H \rangle \neq 0$. In the reduced two-dimensional phase plane, $w_{RF} \times w_C$, including spontaneous events in the cortical layer moves the fixed point of Eq. A4 away from the origin to the coordinates $w_{RF} = w_C = (\lambda^*)^{-1} \theta_u(\langle R_H \rangle)$. Nevertheless, the different dynamical regimes reported in Fig. A2C continue to be valid. In region (i), the addition of cortical events moves the unstable node away from the origin and into the first quadrant (Fig. A4A top). As a result, a small region of selectivity emerges in the plane in which initial conditions generate refined receptive fields (Fig. A4A, top middle). Therefore, the addition of H-events enables the emergence of weight selectivity but through a different mechanism than the one obtained with only L-events. Rather than modulating the learning rule through the input threshold, changing the H-event statistics through the $\langle R_H \rangle$ parameter can generate different receptive field sizes for a fixed input threshold. However, the strength of H-events has to be fine-tuned to generate refined receptive fields. Within a small range of $\langle R_H \rangle$ the network transitions from no-selectivity (where all weights potentiate, Fig. A4A, top left) to complete decoupling (where all weights depress, Fig. A4A, top right).

To relate the reduced two-dimensional phase planes to the simulation results, we used Eq. A2 to obtain $\langle R_H \rangle$ by taking into account the simulation parameters in Table 1. We next verified the predictions of the reduced two-dimensional system in numerical simulations of the full network with H-events (Fig. A4A bottom). To capture the gradual increase of $\langle R_H \rangle$ as in the reduced two-dimensional system, we decreased the average inter-event interval between H-events, $H_{\text{int}}$. As before, only a narrow range of $H_{\text{int}}$ leads to refined receptive fields, albeit with some degree of decoupling (Fig. A4A, bottom middle). Outside of this range, individual cortical neurons are either non-selective (Fig. A4A, bottom left) or nearly completely decoupled from the thalamus (Fig. A4A, bottom right). In regions (ii) and (iii), the fixed point moves from the origin to the first and third quadrants, respectively (Fig. A4B and Fig. A4C). In both cases, only very weak H-events can sustain finite receptive fields because the high input threshold value already provides sufficient depression to the network. We confirmed our analytical results in regions (ii) and (iii) with numerical simulations of the full network (Fig. A4B, bottom and Fig. A4C, bottom).
Figure A4. Spontaneous cortical H-events disrupt receptive field refinement. A. Top: Phase planes of the reduced two-dimensional system for input threshold $\theta_u = 0.4$ (region i) and increasing strength of cortical events $\langle R_H \rangle$ with an example trajectory (red). Selectivity can only be observed for fine-tuned $\langle R_H \rangle$. The fixed point (open circle), an unstable node, has moved to the first quadrant. Bottom: Simulations of receptive field development with the same parameters where $H_{int}$ was progressively reduced (this is the same set of parameters as shown in Fig. 3A). B. Top: Phase planes for $\theta_u = 0.52$ (region ii) and increasing $\langle R_H \rangle$ with an example trajectory (red). The unstable node (open circle) moves from the origin to the first quadrant as $\langle R_H \rangle$ increases. Bottom: Simulations with the same parameters where $H_{int}$ was progressively reduced. C. Top: Phase planes for $\theta_u = 0.6$ (region iii) show the transition from selective receptive fields to cortical decoupling in response to increasing $\langle R_H \rangle$. The fixed point (half-open circle), now a saddle node because $\lambda^* < 0$, has moved away from the origin to the third quadrant. Bottom: Simulations with the same parameters with very infrequent H-events where $H_{int}$ was progressively reduced.
Supplementary Figures

Figure S1. Peripheral L-events generate robust receptive field refinement. A. Without the weak bias along the diagonal, selectivity is achieved so that each cortical cell is connected to a few thalamic inputs; however, the same set of input neurons innervates all cortical neurons. Receptive fields were generated with the Hebbian rule, but similar outcomes are obtained with the BCM rule. B. Receptive fields generated by the Hebbian rule with varying input threshold $\theta_u$. C. Receptive fields generated by the BCM rule with varying target output target rate $v_0$. 

Figure S2. Decoupling of cortical cells is a frequent outcome of H-events in the Hebbian rule A. Proportion of receptive fields that decouple obtained from 500 Monte Carlo simulations for combinations of $H_{int}$ and $\theta_u$ for the Hebbian rule (left) and $H_{int}$ and $v_0$ for the BCM rule (right). B. Average proportion of decoupling of simulation outcomes classified as ‘selective’ in A.
Figure S3. Fluctuations in cortical activity cannot generate correlation. A. Spontaneous events have the same mean amplitudes across consecutive recordings (each ~ 5 mins long) in the same animal (3-14 recordings in 26 animals separated by < 5 mins due to experimental constraints on data collection). Each plot corresponds to one animal where each boxplot contains the event amplitudes (both L- and H-events) from each recording. P-values of one-way ANOVA tests of the event amplitudes across recordings are shown for each animal. B. Absolute difference in the mean event amplitude as a function of time between recordings. Mean values of each recording difference shown as red squares. The values follow the same distribution (Kruskal-Wallis test p = 0.41). The slope of the linear regression fit (red) is very close to 0, confirming that no clear trend of increasing/decreasing activity exists across recordings. C. Number of cells per recording, grouped by different postnatal ages.
Figure S4. Robustness of correlation under variations in the inclusion criteria, $\tau_{\text{max}}$ and $\tau_{\text{decay}}$. 
A. Correlation coefficient between H-event amplitude and leaky average preceding activity for an increasing lower bound in the number of events necessary to include an animal in the analysis ($\tau_{\text{max}} = 300$ s and $\tau_{\text{decay}} = 1000$ s fixed). 95% confidence interval (shaded area) computed through bootstrap regression (Methods). B. Corresponding number of animals satisfying the lower bound criteria in A. C. Correlation coefficient between H-event amplitude and leaky average preceding activity for varying windows of integration $\tau_{\text{max}}$ (lower bound to include events fixed in 12, $\tau_{\text{decay}} = 1000$ s). D. Correlation coefficient as a function of the exponential decay time constant $\tau_{\text{decay}}$ (lower bound to include events fixed in 12, $\tau_{\text{max}} = 300$ s). The arrows indicate the parameters of the analysis in Fig. 5, namely a minimal of 12 events per animal, $\tau_{\text{max}} = 300$ s, $\tau_{\text{decay}} = 1000$ s.

Figure S5. Properties of effective spontaneous events: duration, distribution of event sizes and H-event frequency and amplitude as a function of the input threshold. A. The durations of effective L- and H-events in the data are normally distributed with the same parameters. B. Top: Sample traces of cortical activity corresponding to Fig. 7D. Bottom: Distribution of spontaneous event sizes for the different input thresholds in the model as Fig. 7D. C. Frequencies (squares) and amplitudes (circles) of H-events in the model at different input thresholds (abscissa). H-events are suppressed both in amplitude and frequency at higher input threshold $\theta_u$. Bars represent the standard deviation of 10 simulation runs.