Ensemble priming via competitive inhibition: local mechanisms of sensory context storage and deviance detection in the neocortical column

Ryan V. Thorpe[1], Christopher I. Moore[1,2], and Stephanie R. Jones[1,2]

[1]Department of Neuroscience, Brown University, Providence, RI, USA
[2]Carney Institute for Brain Science, Brown University, Providence, RI, USA

Abstract

The process by which neocortical neurons and circuits amplify their response to an unexpected change in stimulus, often referred to as deviance detection (DD), has long been thought to be the product of specialized cell types and/or routing between mesoscopic brain areas. Here, we explore a different theory, whereby DD emerges from local network-level interactions within a neocortical column. We propose that deviance-driven neural dynamics can emerge through interactions between ensembles of neurons that have a fundamental inhibitory motif: competitive inhibition between reciprocally connected ensembles under modulation from feed-forward selective (dis)inhibition. Using this framework, we were able to simulate a variety of phenomena pertaining to the experimentally observed shifts in neural tuning across neurons, time, and stimulus history. Anchoring our approach in a variety of experimentally observed phenomena, we used computation modeling in two types of neural networks of vastly different levels of biophysical detail to test hypotheses on emergent dynamics and explore the robustness of underlying connectivity parameters. With a number of corollary predictions that can be tested in future in vivo studies, we show that ensemble priming via competitive inhibition under modulation from selective (dis)inhibition acts as a local mechanism for sensory context storage and that DD does not require specialized input from other brain areas—a novel theoretical paradigm that resolves previously confounding aspects of sensory encoding and predictive processing in the neocortex.

Introduction

Neocortical circuits respond with distinct sensitivity to unexpected 'deviant' stimuli, an amplification mechanism widely regarded as crucial for sensory processing [1–7]. Among the most widespread phenomena associated with deviance detection (DD) is the observation that groups of neurons in the sensory neocortex tend to increase their average spike rates in response to deviant stimuli [8–10]. How specific neurons (comprising functional ensembles) shift their stimulus-evoked response profile on deviant trials, however, is an open ended question.

The concept of neural tuning—whereby a given neuron fires more or less to specific configurations of an external stimulus (e.g., the angle of an oriented bar in a visual scene) and appears to be statically 'tuned' to a subset of stimulus features [11, 12]—is useful for understanding broad categories of stimulus representations, but at face-value struggles to capture how a neuron's response shifts based on environmental context or stimulus history. At low stimulus magnitudes near perceptual threshold (i.e., within the dynamic range of a mouse's psychometric response), a deviant shift in stimulus amplitude (Figure 1a) modulates the salience of afferent input, producing a corresponding decrease (or increase) in the spike rate of neurons with simple tuning properties (e.g., as is typical in layer 4 of the primary somatosensory barrel cortex, L4; Figure 1b, exemplar single-unit response shown in pink and green). Unlike the characteristically 'feed-

forward' cells of L4, certain neurons such as those found commonly in L2/3 can behave with more complexity: a decrease (or increase) in sensory afferent drive corresponds to a decrease in the spike rate response of some neurons and increase in others, yet with no shift in the average (Figure 1c).

Recent evidence from Voigts *et al.* [13] suggests that deviance-driven complex tuning is unique to L2/3 (Figure 1c). Importantly, deviance-driven complex tuning in L2/3 emerges from local mechanisms at early latencies (i.e., < 100 ms) and is highly sensitive to stimulus-evoked cross-talk from another layer in the neocortex, namely, L6 [13]. With a deviant decrease (or increase) in vibrissa stimulus amplitude, putative excitatory regular spiking (RS) neurons in L2/3 shift their single-unit stimulus-evoked spike rate responses differently than in L4, exhibiting complex heterogeneous responses across the laminar population. One L2/3 RS might increase its spike rate response, while another might decrease its spike rate response on deviant trials (Figure 1c). Importantly, this occurs for both deviant increases and decreases in stimulus amplitude but not for baseline shifts in stimulus amplitude, indicating that the deviance-driven neural response somehow computes the magnitude of the derivative of sequential afferent drives rather than a tonic increase or decrease in the stimulus-evoked representation [13]. While the in vivo ensemble-level dynamics associated with deviance-driven complex tuning in L2/3 are currently unknown—as are its causal mechanisms—the primary goal of the present study is to elucidate them within the context of a unifying theory accounting for both the adaptation of neural tuning profiles (i.e., sharpening and/or broadening) [14–16], shifts in feature discrimination [14, 17, 18], and DD [8–10, 13]. Specifically, we sought to elucidate the underlying ensemble dynamics of deviance-driven complex tuning by characterizing how various inhibitory circuit motifs allow a deviant change in evoked afferent drive to selectively amplify the response of some neurons at the expense of others.

We propose that only a handful of network dynamic scenarios could possibly serve as a basis for deviance-driven complex tuning at the neural ensemble level (Figure 1d). Namely, neurons tuned to a stimulus can respond to deviant increases or decreases in stimulus amplitude with a more pronounced ordinal position (responder rank reinforcement; Figure 1d, left panel) or diminished ordinal position (responder rank weakening; Figure 1d, middle panel) among other stimulus-responsive units ranked by evoked spike rate magnitude. If a sharp tuning profile can give way to a weakened tuning profile, the ordinal position of various stimulus-responsive units can, in theory, invert completely compared to baseline to temporarily establish a sharp, yet reversed, tuning profile (responder rank reassignment; Figure 1d, right panel). We emphasize that in each of these cases, a deviant decrease in stimulus amplitude does not necessarily decrease the mean evoked response across neurons. Instead, it selectively amplifies the response of some neurons at the expense of others in a manner that maintains, if not increases, the net salience of the neocortical area's output. Each ensemble-level dynamic response intrinsically supports divisive normalization as a foundational computational principle, which we explore below from a dynamical systems perspective [19, 20].

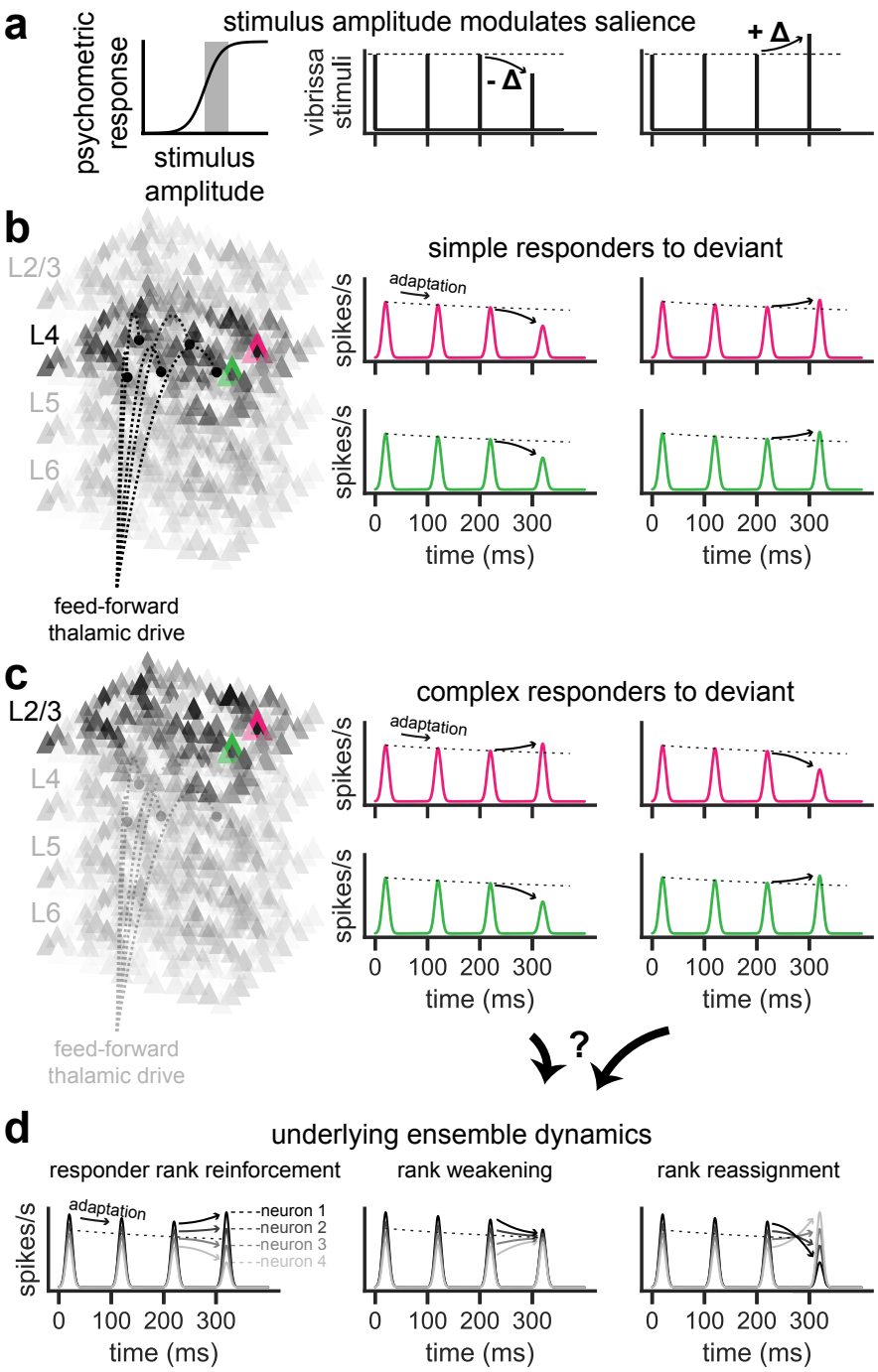

Figure 1: Deviance tuning emerges with a deviant shift in afferent drive in a subset of layer 2/3 neurons. (a) Experimental stimulation paradigm, where vibrissa stimuli with a deviant shift in amplitude lie within the dynamic range of a mouse's psychometric response. Deviant shifts in stimulus amplitude are expected to modulate salience in the same direction as the deviant (negative change, decrease; positive change, increase) for neurons with simple tuning properties. (b) Cartoon depiction of single-unit spike rate evoked responses in L4 of mouse barrel cortex: two example excitatory neurons (highlighted in pink and green) present as simple responders that shift their responses proportional to the change in stimulus amplitude on deviant trials [13]. (c) Same as in (b), except in L2/3 where complex responders emerge heterogeneously: one example neuron increases while the other decreases its spike rate response to a deviant change in stimulus amplitude [13]. (d) Hypothetical scenarios of possible ensemble dynamics underlying the transient shift in deviant tuning profiles of complex responders in L2/3. This study explores which of these unknown scenarios is most plausible and how they might emerge through recurrent dynamics between layers and cell types of a canonical neocortical column.

It's important to note that current theoretical paradigms for DD (e.g., predictive coding [2]) lack the ability to resolve the network mechanisms of deviance-driven complex tuning, which cannot be explained purely by a shift in top-down modulation from higher-order cortex, local stimulus-specific adaption (SSA; generally attributed to the cellular mechanism of synaptic depression), and/or a shift in thalamocortical routing [1, 5, 10, 21]. In order to facilitate selective amplification of a specific subset of neurons in a network at relatively short inter-stimulus intervals ($\leq$ 100 ms), such mechanisms would require a deviance-induced shift in the afferent pathway, whereby a distinct change in stimulus features evoke a novelty response in a newly targeted, non-adapted neuron ensemble. Since amplitude deviants hold the fundamental composition of the stimulus (i.e., the collection of features that define it) constant across standard and deviant stimulus trials, the stimulus' spatiotemporal map in the thalamus and cortex (theoretically) also remains constant. Deviance-driven phenomena of subtle amplitude changes, like complex tuning in L2/3, therefore don't fall under the purview of prior mechanistic models of DD and predictive processing.

We used simulations from two computational models of different complexity to test if and how distinct network configurations can leverage inhibition to prime and shape the response of neuron ensembles in service of DD. First, we used a discrete Wilson-Cowan neural mass model to test generalized principles of inhibitory dynamics between two arbitrary competing neural representations [22, 23]. We propose the network-level mechanisms of reciprocal inhibition between at least two competing ensembles of neurons (competitive inhibition) and selective (dis)inhibition as foundational to the emergence of experimentally observed deviance-driven neural dynamics. We then extended our theoretical results to a biophysically-detailed neocortical column model constrained by laminar structure, anatomical connectivity, and distinct neuron types with fast and slow excitation and inhibition [24].

Employing computational neural modeling and prior anatomical and physiological knowledge about neocortical column circuitry, we found that recurrent inhibition-driven network dynamics account for deviance-driven complex tuning in L2/3 and provide a handful of plausible connectivity motifs that facilitate the emergence of selective amplification of specific neural ensembles. With a number of corollary predictions that can be tested in future in vivo studies, we show that ensemble priming via competitive inhibition and modulated by selective (dis)inhibition acts as a local mechanism for sensory context storage and that DD does not require specialized input from other brain areas—a novel theoretical paradigm that resolves previously confounding aspects of sensory encoding and predictive processing in the neocortex.

Results

Deviance-driven complex tuning emerges with inhibitory competition and selective (dis)inhibition

Given the possible underlying ensemble dynamics that can account for the experimentally observed phenomenon of deviance-driven complex tuning in L2/3 (Figure 1d), we began this study by proposing the notion of competitive inhibition (CI) and selective (dis)inhibition (SI).

Consider two ensembles of neurons—each encoding a distinct feature or combination of features of the environment through the mean spike rate response across constituent neurons—that naturally receive different levels of afferent excitation evoked by an external stimulus. If these ensembles (or neural 'representations') reciprocally excite each other, evoked activity propagates from one ensemble to the next, resulting in mutual excitation (ME) that attenuates the relative difference between the ensembles' respective responses. With multiple repetitions of the same stimulus in a short time window, residual excitation accumulates (i.e., in the subthreshold currents of sparsely activated neural populations [25]) such that the evoked response

encoded by neurons across ensembles converges to some intrinsic equilibrium governed by the net excitation/inhibition (E/I) balance within each ensemble. If, on the other hand, the two ensembles reciprocally inhibit each other, evoked activity produces CI that augments the difference between each ensembles' response. With multiple repetitions of the same stimulus, CI drives distinct ensemble responses apart, aiding in feature discrimination that adapts with each repetition (i.e., adaptive sharpening of the neural tuning). The neural response across ensembles prevents itself from saturating, and the equilibrium state becomes markedly time and repetition-dependent as governed by the shifting E/I balance *between* ensembles. In this latter case, the neuronal representation across ensembles self-normalizes and from the perspective of a naive observer, appears to shift over time and space. Should evoked (dis)inhibition via a parallel feed-forward pathway (e.g., from another pair of ensembles or local neocortical layer) selectively dampen the activity of one representation more than the other in a stimulus-dependent manner (termed here as the process of selective [dis]inhibition, SI), the diverging effects of CI could be modulated to produce apparent shifts from the ensembles' baseline tuning that is sensitive to stimulus history and context.

We hypothesized that CI between at least two ensembles of neurons (Figure 2a,b) modulated by SI from another functional layer (Figure 2c) can account for one or more of the ensemble dynamics scenarios subserving deviance-driven complex tuning (Figure 1d). Importantly, such a system would need to produce complex tuning while also maintaining, if not increases, the total magnitude of the network's evoked output, a hallmark characteristic of DD [8–10]. We explored this idea in a discrete (space-clamped) Wilson-Cowan model of neural mass firing rates given by

$$\tau \frac{dx_i}{dt} = -x_i + \mathcal{F}_i \left(\sum_{j=1}^{N} w_{ij} x_j + I_i\left(t\right) \right) \tag{1}$$

where $\mathcal{F}_i$ is a sigmoid activation function for the ith unit (ensemble) that increases or decreases with endogenous recurrent and exogenous input $w_{ij}x_j$ and $I_i(t)$, respectively [22, 23]. Seeking to verify that the theoretic mechanisms of CI+SI are (at minimum) sufficient for producing deviance-driven complex tuning in a manner that maintains, if not increases, the total magnitude of the network's evoked output, we used the system defined in Equation 1 with $N = 4$ to establish a fundamental 4-dimensional network motif that boils our hypothesized mechanisms down to their simplest form (Figure 2c; note that CI without SI is shown for illustrative purposes in Figure 2b). As explored later in this study, the constraints on this motif can be loosened to represent general inhibitory ensemble interactions without CI and/or without SI.

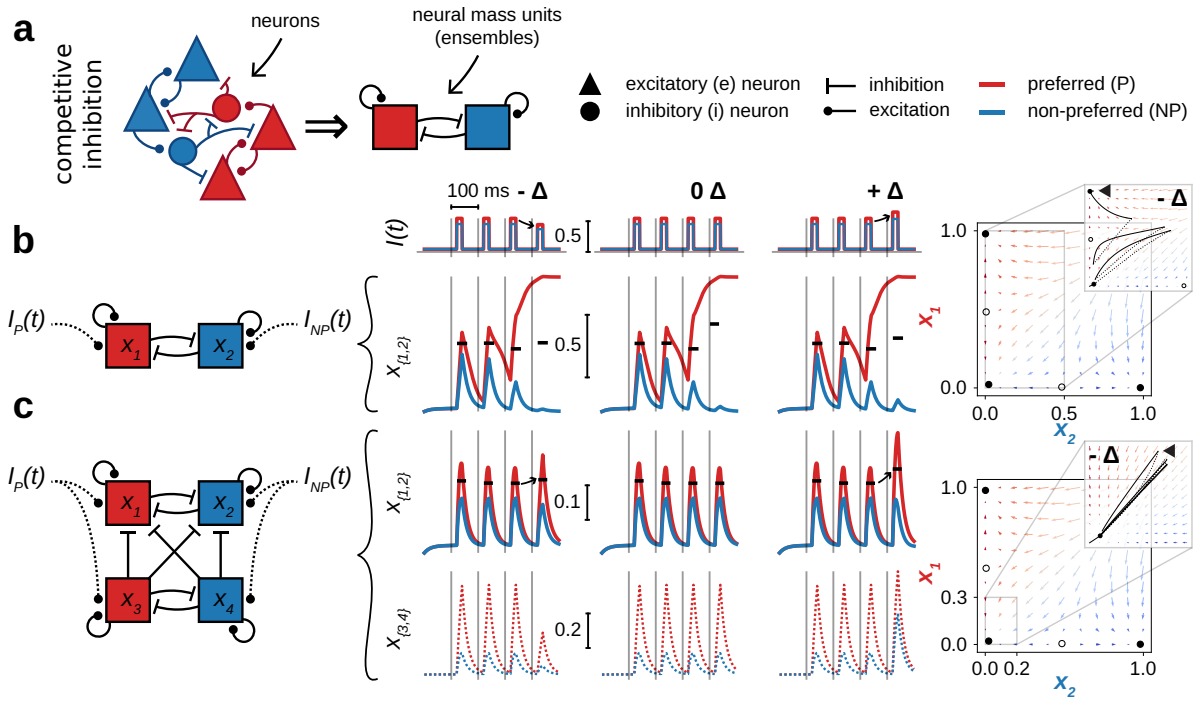

Figure 2: Competitive inhibition (CI) between neural representations modulated by selective (dis)inhibition (SI) from another functional neural layer facilitates net amplification of the deviance-driven response and deviance-driven complex tuning. (a) CI requires at least two neuron ensemble representations that reciprocally inhibit each other but generally allow excitation to spread among constituent neurons of each respective ensemble. One ensemble is maximally tuned to an arbitrary stimulus (preferred, red) while the other is sub-optimally tuned to the same stimulus (non-preferred, blue). (b) Simulation of diverging responses to repetitive sensory afferent drive within a two-dimensional model of CI. Simulations comprise a sequence of standard stimulus-evoked drives (top) followed by a negative deviant (middle-left), no deviant (middle-center), and a positive (middle-right) deviant. On the far right, the resting-state phase plane with stable and unstable fixed points (solid versus hollow dots, respectively) contains competing basins of attraction that a given simulation trajectory (inset) navigates between following each stimulus perturbation. (c) Same as in (b), except with a 4-dimensional network architecture that facilitates a deviance-driven mean spike rate increase and complex tuning in the upper units, x_1 and x_2. The connectivity motif enforces CI in two functional layers (upper and lower) and uni-directional cross-laminar inhibition from the lower to upper units. Evoked activity in the lower layer produces selective (dis)inhibition (SI) of evoked activity in the upper layer. For both negative and positive amplitude deviants, the model responds with a non-decreasing average response across upper units (black tick marks).

In the neural mass model defined in Equation 1, x_i, can be thought of as the proportion of constituent neurons of the ith ensemble that are firing at a given point in time. Each neural mass unit (ensemble) comprises a distinct neural representation that is either preferred (red) or non-preferred (blue) by exogenous stimulus-evoked drive. Exogenous excitatory drive is injected into the network through $I(t)$, which is a time-dependent step function designed to emulate the stimulus-evoked afferent input entering each ensemble of the neocortical column. As shown at the top of Figure 2b, $I(t)$ is slightly larger for red than blue (representing greater evoked drive targeting the preferred ensemble) when it is in its high state beginning at 20 ms post-stimulus, and then returns to a common low state at 40 ms post-stimulus in a sequence of four stimuli. On the fourth stimulus repetition, the magnitude decreases or increases across representations by

a scalar factor corresponding to a deviant decrease in stimulus strength. Note that the relative balance of $I_P : I_{NP}$ across standard and deviant afferent inputs remained constant: only the mean peak value across I_P and I_{NP} changed.

In addition to exogenous drive, the ith neural mass unit also receives input through recurrent connections with itself and other neighboring units, as determined by connectivity weights w_{ij} from the jth unit (Equation 1). The CI (Figure 2b) and CI+SI network (Figure 2c) motifs imply a few fundamental connectivity constraints. Connection weights constrained to be positive represent excitatory connections that propagate excitation throughout constituent neurons of an ensemble, whereas connection weights constrained to be negative represent inhibitory connections in service of CI between ensembles in a layer and/or SI between layers.

CI promotes diverging neural representations that augment with each presentation of stimulus-driven afferent drive (Figure 2b). Specifically, we see a diverging sequence of neural responses between two competing neural mass units (red versus blue) in response to a sequence of repetitive of drive, regardless of whether or not the final repetition in the sequence is deviant (Figure 2b, middle-left through middle-right). On the final stimulus, the dynamic state of the network has fallen into the upper attractor basin favoring the preferred representation, where it produces tonic spiking due to runaway excitation and minimal activity in the non-preferred representation.

With the addition of a modulatory layer that provides SI via uni-directional cross-laminar inhibition to the same units depicted in Figure 2c, the diverging effect of CI in the upper layer is dampened. When SI is decreased or increased based on the direction of the deviant (Figure 2c, middle-left and middle-right), the equilibrium of the network shifts, and the mean evoked response across upper layer units (black tick marks) not only refrains from decreasing on deviant trials, but increases for both negative (2c, middle-left) and positive deviants (2c, middle-right). Further, the network exhibits an increase in the difference between the red and blue evoked responses for both types of deviants compared to baseline, with red dominating. This is an example of responder rank reinforcement as introduced in Figure 1d.

To verify that the CI+SI motif is capable of responding uniquely to deviant stimuli that shift (or in our case, swap) which feature is preferred instead of decreasing the total intensity across all features, we modified network connectivity to either produce slowly converging or diverging evoked responses, and then simulated a feature-swap deviant (Figure S1). Given that such a scenario maintains a constant mean level of evoked afferent drive even for deviant stimuli, we expected it to amplify the history-dependent nature of the CI+SI motif so that the magnitude of the average deviance-driven response across representations increases depending on the number of prior stimuli of a particular feature preference. Indeed, this is exactly what we observed: for deviants that succeed 1, 2, 3, or 4 prior repetitive stimuli (Figure S1a,c), the magnitude of the average deviant response gets progressively larger (Figure S1b,d).

While other forms of deviant stimuli and deviance-driven complex tuning can emerge with a change in network parameters, this demonstrates the overarching utility of CI and SI: a local neocortical circuit (or other brain area for that matter) can maintain, if not amplify, the total deviance-driven response through a common connectivity motif responsible for deviance-driven complex tuning. The network motif is uniquely designed to selectively amplify the response of a specific neural subpopulation at the expense of another, which is a process that can shift dynamically in response to a deviant change in magnitude or feature preference in service of DD.

Inhibitory competition primes the network by facilitating adapting neural representations

We then sought to determine if and how CI and SI interact to support both adaptation and deviance-driven shifts in neural tuning. CI by nature augments the difference between competing ensemble representations, priming the network for a deviant response with each repetition of a stimulus. Yet, as demonstrated in Figure 2c, the effects of CI can be dampened or enhanced by SI based on the level of stimulus-evoked afferent drive. Numerous prior studies report that the stimulus tuning of neural representations shift over time, particularly when a given stimulus has been presented repetitively or sampled frequently by the peripheral nervous system within a short time window [14–16]. It has been suggested that adaptive sharpening (i.e., the narrowing of a neuron's tuning to external stimuli, similar to the ensemble-level phenomena of responder rank reinforcement) subserves the psychophysical phenomenon of enhanced feature discrimination due to stimulus repetition [16–18, 26–29]. While many cellular and circuit mechanism are likely involved in facilitating stimulus-specific adaptation [14, 16, 28, 30, 31], here we assumed that local inhibition plays a prominent role [32, 33].

As a theoretic architectural constraint of neocortical circuitry, we specifically asked if the CI+SI network motif has a parameter regime where adaptive sharpening of neural tuning and deviance-driven complex tuning emerge, all while producing an average deviance-driven non-decreasing (DnD) response. It could be that deviance-driven complex tuning only emerges within a specific regime where adaptive sharpening cannot. Or, the average deviance-driven response might always decrease on negative deviant trials. Either way, a null result (where less than all desired phenomena are produced by the network) would invalid the CI+SI motif as a possible network architecture underlying DD.

To test if the CI+SI motif has a parameter regime where adaptive sharpening of neural tuning, deviance-driven complex tuning emerge, and a DnD response emerge, we performed robustness analysis that estimated the likelihood of observing specific emergent phenomena given a uniform parameter prior. Connectivity of the 4-dimensional CI+SI motif is determined by six parameters, so we sampled 1×10^6 random parameter values, each a 6-dimensional vector, to estimate the robustness of emergent dynamics across the parameter-space (Figure 3a).

To control for the possibility that our desired emergent phenomena (adaptive sharpening of neural tuning, deviance-driven complex tuning, and a DnD response) might exist in a parameter-space outside the fundamental constraints of CI and SI, we loosened the CI and SI constraints by allowing the cross-representation connections to take a negative (inhibitory) or positive (excitatory) value, and cross-laminar inhibitory connections to take a maximal value of 0. Thus, a randomly sampled parameter could produce a network with ME between ensemble representations and/or silenced cross-laminar SI.

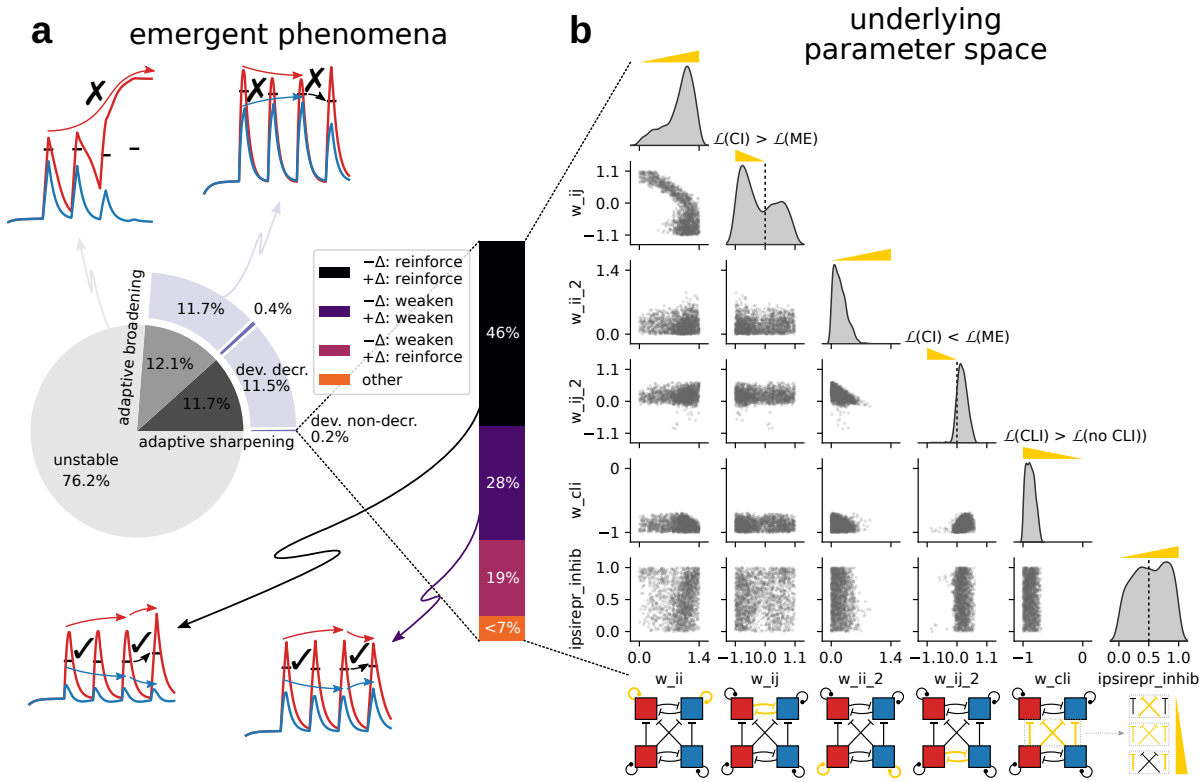

Figure 3: Adaptive sharpening and deviance-driven dynamics emerge from a small subset of the possible network parameter-space. (a) Pie plot showing the proportion of randomly sampled parameter configurations that produce desired emergent phenomena (i.e., adaptive sharping of a neural representation's tuning, deviance-driven complex tuning, and a DnD response) or undesired network phenomena (i.e., unstable runaway excitation, adaptive broadening, and/or a deviance-driven decreasing response). Samples producing all desired phenomena are further categorized by the form of deviance-driven complex tuning they produce, whose relative proportions are shown in the barplot (right). Example simulations with unstable runaway excitation (upper left), adaptive broadening with a deviance-driven decreasing response (upper right), adaptive sharpening with a DnD response and deviance-driven complex tuning (bottom left and right) are shown as inset plots. (b) Pairplot of the parameter values for each sample that produced all desired emergent phenomena: each point of a scatter plot denotes the 2-dimensional coordinate of a given parameter sample with the marginal kernel density estimate for each parameter shown on the diagonal. Each of the six network parameters represents the strength of a pair of connections constrained by the relaxed CI+SI motif (bottom). Yellow ramps indicate the direction of increasing strength for a fundamental CI+SI motif element mapped from its underlying network parameter. For example, increased CI is induced by decreasing `w_ij` or `w_ij_2` starting at zero (second and forth diagonal element from top).

We found that the network could indeed produce all three emergent phenomena, adaptive sharpening, deviance-driven complex tuning, and a DnD response; however, concurrence of such phenomena is rare and highly sensitive to network parameters. After discarding parameter configurations where the network produced a self-sustaining spike rate response through runaway excitation (termed here as 'unstable'), only $\sim 11.7\%$ of the total 10^6 samples produced adaptive sharpening (Figure 3a, dark gray). Of those, an even smaller percentage produced DnD responses and deviance-driven complex tuning (~ 0.2 of the total original number of samples; Figure 3a, outer layer of pie chart). Each parameter configuration that produced DnD responses also produced deviance-driven complex tuning, which we further categorized as either

responder rank reinforcement, weakening, or reassignment for negative and positive deviants as defined in Figure 1d (Figure 3a, right).

One form of deviance-driven complex tuning was more robust than the others (responder rank reinforcement for both negative and positive deviants) and therefore emerged more prominently from the network despite parameter variability, occupying $\sim 46\%$ of the valid parameter space. Despite this, the parameter-space responsible for producing the concurrent emergent phenomena of adaptive sharpening, deviance-driven complex tuning, and DnD responses was remarkably small. Through inspection of this underlying parameter-space, we were able to make a few inferences about the necessity of CI and SI (Figure 3b).

Emergent deviance tuning dynamics are highly sensitive to network connectivity, but competitive inhibition promotes robustness

As graphically depicted at the bottom of Figure 3b (yellow connections), the loosened CI+SI motif parameters comprise a positive (excitatory) recurrent connection weight for each of the upper layer units (`w_ii`; symmetric across units), a negative or positive connection weight between the upper layer units (`w_ij`) allowing for either CI or ME, two similar weight parameters for the lower layer (`w_ii_2` and `w_ij_2`), a negative (inhibitory) weight for the maximal cross-laminar inhibition (CLI) connection from the lower to upper layers that can have a maximal value of zero, and a $[0, 1]$ scalar value that indicates the fraction of total CLI getting routed through connection weights between member units of the same neural representation (`ipsirepr_inhib`). Note that connectivity symmetry is enforced between the preferred and non-preferred representations at all times, allowing the CI+SI motif to embody arbitrary anatomically-determined neural representations whose meaningful dynamics emerge from asymmetrical activation from thalamus or other brain area according to their stimulus, motor, or contextual mapping. Each randomly sampled parameter configuration that produced the aforementioned criteria of desired emergent phenomena is represented as a point in Figure 3b. A kernel density estimate was used to approximate the marginal distribution for each parameter (diagonal subplots).

The first inference we made about the underlying network parameters that promote adaptive sharpening and deviance-driven dynamics was that CI enhances robustness to parameter variability. The distribution of reciprocal connection weights in the upper layer (`w_ij` among the diagonal subplots of Figure 3b) has its most prominent peak in the negative range of values, indicating that random parameter configurations with CI have a greater likelihood of producing the desired emergent phenomena given a uniform prior than for ME ($\mathcal{L}(CI) > \mathcal{L}(ME)$). Note that increased magnitude of CI, along with other fundamental CI+SI motif elements like SI implemented through CLI, are indicated graphically for each parameter as a yellow ramp in Figure 3b. The lack of a prominent peak in favor of CI for the lower layer (`w_ij_2`; $\mathcal{L}(CI) < \mathcal{L}(ME)$) indicates that the adaptive effects of CI are undesirable for the ensemble responses in this layer, which makes sense given that the lower layer is the source of SI and plays a modulatory role that needs to encode stimulus intensity rather than the adaptive influence of stimulus repetition. The robustness with which the lower layer encodes raw stimulus intensity allows it to better modulate the inhibitory landscape of the upper layer in response to a deviant shift in intensity.

Upon observing that CI increases the robustness of emergent deviance-tuning dynamics midst connection variability, we characterized how sensitive such emergent dynamics were to time (specifically, the inter-stimulus interval [ISI]). Taking the connectivity parameter configurations sampled in Figure 3 that pass all criteria (i.e., exhibiting evoked responses with adaptive sharpening that were also DnD) for a 100 ms inter-stimulus interval (ISI), we extended or retracted the ISI over a range of values from 50-200 ms to quantify the proportion of connectivity configurations that retained stability and DnD status (Figure 4). Connectivity configurations originally classified according to their emergent deviance tuning forms in Figure 3a fell into one of three

major categories that each correspond to a unique region within parameter-space (Figure 4 inset with random parameter configurations plotted as points in 3-parameter dimensions): responder rank reinforcement for both negative and positive deviants (black), responder rank weakening for both negative and positive deviants (purple), and responder rank weakening (reinforcement) for negative (positive) deviants distinctively (pink). Of particular note, connectivity configurations with `w_ij` < 0 (i.e., CI instead of ME) predominantly led to responder rank reinforcement (black compared to pink and purple) and retained stability and DnD status across ISI values greater than 100 ms while other parameter configurations did not (Figure 4 solid and dotted traces).

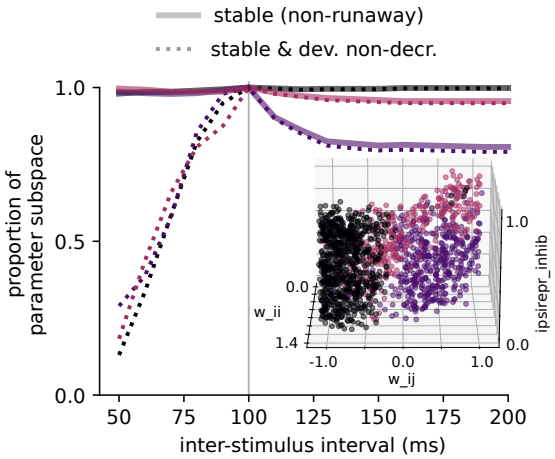

Figure 4: Deviance-driven dynamics are stable across a range of time intervals, particularly for networks with strong competitive inhibition. Taking the connectivity parameter configurations sampled in Figure 3 that pass all criteria (i.e., exhibiting evoked responses with adaptive sharpening and are deviance non-decreasing [DnD]) for a 100 ms inter-stimulus interval (ISI), we extended (or retracted) the ISI over a range of values from 50-200 ms to quantify the proportion of connectivity configurations that retained stability and DnD status. Connectivity configurations originally classified according to their emergent deviance tuning dynamics fall into one of three major categories that each correspond to a unique region within parameter-space (inset): responder rank reinforcement for both negative and positive deviants (black), responder rank weakening for both negative and positive deviants (purple), and responder rank weakening (reinforcement) for negative (positive) deviants distinctively (pink), as shown in Figure 3. Of particular note, connectivity configuration with `w_ij` < 0 (i.e., competitive inhibition, CI instead of mutual excitation, ME; black and pink) retained stability and DnD status across ISI values greater than 100 ms.

Deviance-driven tuning requires strong selective (dis)inhibition

The second inference we made about the underlying network parameters that promote adaptive sharpening and deviance-driven dynamics was that strong SI is crucial. Across all valid parameter configurations, the distribution of CLI values peaks near -1, at the lower end of its range, (`w_cli` in Figure 3b). In fact, no valid parameter samples exist near 0, indicating that having a network with SI of maximal strength not only increases the likelihood of producing the desired emergent phenomena given a uniform random prior of other parameters ($\mathcal{L}(CLI) > \mathcal{L}(no\ CLI)$), but a requirement for producing the desired emergent phenomena at all. We emphasize that representing SI as a uni-directional inhibitory connection between two layers in a network (i.e., CLI) is just one of many ways to implement SI in a model that are functionally equivalent. The important thing is that inhibition must come from a separate group

of neurons that share afferent input, which then modulates the E/I landscape of the group of neurons in which adaptive sharpening and deviance-driven complex tuning is measured.

Inhibitory competition and selective cross-laminar (dis)inhibition regulate deviance tuning in a spiking neocortical column network

Up until this point, we explored network dynamics underlying deviance-driven complex tuning in a simplified spike rate model that approximates ensembles of neurons as neural mass units. While such a model conveniently simulates smooth time-varying spike rate functions that emerge from a network governed by only six parameters, we sought to test if our results extend to spiking dynamics in a biophyscially-detailed neocortical column network. Specifically, we found earlier in this study that CI increases the robustness of a network's capacity for producing deviance-driven complex tuning and that selective (dis)inhibition is critical for such phenomena to emerge at all: are these network mechanisms still relevant in a random, sparsely connected network that has been constrained with biophysical realism and anatomical connectivity rules? We also wished to account for the potential influence of thalamocortical adaption, which is a prominent component of stimulus-evoked responses in the neocortex yet was absent from simulations in the neural mass model [28, 30, 32–34].

Here, we tested the hypothesis that at least one of the deviance-driven complex tuning scenarios of responder rank reinforcement, rank weakening, or rank reassignment (Figure 1d) can be produced within a biophysically-detailed neocortical column model constructed with CI between stimulus preferred and stimulus non-preferred ensembles and monosynaptic L6→L2/3 cross-laminar inhibition (CLI). Further, is deviance-driven complex tuning sensitive to ablation of the L6→L2/3 inhibitory connection, the anatomical pathway for selective (dis)inhibition. As elaborated on in Discussion, monosynaptic L6→L2/3 inhibition is implemented here as a hypothetical mechanism that warrants further experimental investigation. Our methods and results draw inspiration from prior work indicating a causal influence between stimulus-driven ensembles of L6 corticothalamic cells (CT) and the emergence of deviance-driven complex tuning in L2/3 [13], presumably through an established CLI pathway from L6 to the more superficial layers of neocortex [35, 36].

Using the Human Neocortical Neurosolver (HNN) modeling framework [24, 37], we simulated the process by which a neocortical column receives and processes repetitive and/or deviant stimulus-evoked exogenous perturbations from other brain regions such as the sensory thalamus (Figure 5). Constrained by prior experiments and neurophysiology data, the biophysically-detailed model contains neocortical layers 2/3, 5, and 6 and two distinct neuron ensembles in both L2/3 and L6 (preferred representation, red; non-preferred representation, blue) that receive different magnitudes of excitatory stimulus-evoked drive from exogenous sources (Figure 5a). Note that this model explicitly controls the activity of the granular layer (L4, the putative source of feed-forward thalamic input) by defining its mean evoked spike rate rather than building L4 neurons into the model. L5 comprises only one functional ensemble. Inputs to the network come in the form of excitatory pre-synaptic spikes that arrive in either (1) L4 and L6 from lemniscal thalamus (proximal drive), or (2) L2/3 from non-lemniscal thalamus or high-order neocortex (Figure 5b). Stimulus-evoked dynamics then propagate vertically (and laterally) throughout the network governed by biophysical constraints within individual neurons and local network connections between neurons (Figure 5b). For baseline activity, the network was provided with excitatory Poisson drive (i.e., spike time events sampled from a Poisson distribution) to the somas of each neuron to establish sparsely firing activity that matches average rates from previous recordings in rodent barrel cortex (see Materials and Methods and Figure S2) [38, 39].

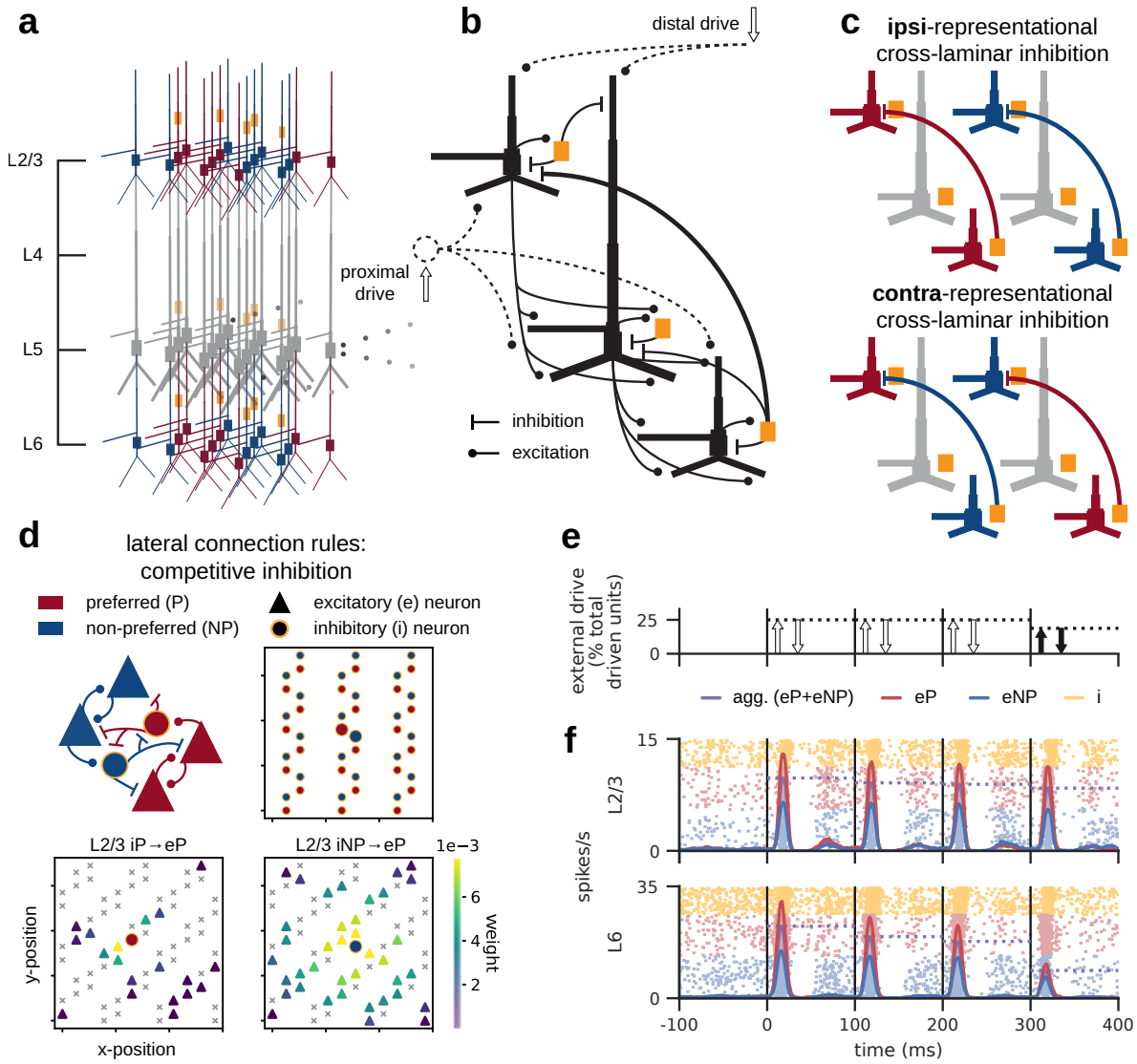

Figure 5: Competitive inhibition combined with cross-laminar selective (dis)inhibition can be extended to a spiking network model with biophysical realism and anatomical connectivity rules. (a) Overview of the model which contains neocortical layers 2/3, 5, and 6. L4 activity is not modeled, but rather controlled explicitly through proximal drive. Neurons in L2/3 and L6 are sorted into two distinct stimulus representations (preferred, red; non-preferred, blue) that make up functional ensembles and receive different magnitudes of excitatory stimulus-evoked drive from exogenous sources. (b) 'Proximal' drive (upward arrow) from lemniscal thalamus arrives at L4 and L6, while 'distal' drive (downward arrow) arrives at the uppermost extent of L2/3. Excitatory and inhibitory dynamics propagate throughout the network via local network connections. (c) L6→L2/3 cross-laminar inhibition (CLI) can have either ipsi-representational or contra-representational targets. Here, we constrained CLI to be 75% ipsi-representational. (d) Lateral connectivity between neurons within a given layer was parameterized for each neuron type using connection probability and distance-dependent maximal post-synaptic conductance. Competitive inhibition was implemented by increasing the probability and spatial conductance decay constant for cross-ensemble relative to within-ensemble inhibitory connections. For example, the connections from inhibitory interneurons to excitatory pyramidal neurons are more numerous and of greater strength between the red and blue representations than within a given representation. (e) The sequence of repetitive exogenous drives (proximal, then distal) that end with a negative amplitude deviant (solid black arrows). (f) Simulated aggregate spike raster and average spike rate in L2/3 and L6 across 40 random trials/networks.

Selective (dis)inhibition was implemented as L6→L2/3 CLI and can have either ipsi-representational or contra-representational targets (Figure 5c). For simplicity, we constrained CLI to be 75% ipsi-representational and 25% contra-representational; however, different configurations for this parameter could theoretically be used. Lateral connectivity rules between neurons were imposed by two parameters for each neuron type: connection probability and distance-dependent maximal post-synaptic conductance. CI was implemented by increasing the probability and spatial conductance decay constant for cross-ensemble inhibitory connections relative to cross-ensemble inhibitory connections (Figure 5d). For example, the connections from inhibitory interneurons to excitatory pyramidal neurons were more numerous and of greater strength between the preferred (red) and non-preferred (blue) representations than within a given representation.

A single simulated evoked response trial of the biophysically-detailed model was composed of a sequence of exogenous drives that represents the afferent and contextual exogenous inputs received from other brain areas evoked by four repetitive stimuli administered 100 ms apart. Building from prior work [24, 40–42], each repetition contained a proximal and distal drive at ∼12 ms and ∼35 ms post-stimulus, respectively, that were slightly weaker or stronger for the deviant (fourth) repetition (Figure 5e). As described in more detail in Materials and Methods, evoked drives were implemented as pre-synaptic spike events sampled at times from a Gaussian distribution, entering the cortex at a specific location (proximal or distal) and targeting specific post-synaptic cells. Evoked drive strength was modulated for deviant afferent drive by decreasing or increasing the number of total driven units from a baseline value for both proximal and distal drives. To account for the effect of thalamocortical adaptation thalamocortical adaptation, the synaptic strength (i.e., maximal post-synaptic conductance) of proximal drives attenuate at a rate of 8% per repetition. Importantly, the ratio of proximal drive strength between the two ensembles remained constant across the entire sequence of evoked drives. Note that for the portion of this study involving the biophsycially-detailed model, we focused primarily on the negative deviant case (i.e., a decrease in drive strength for the fourth repetition) since network dynamics that refrain from attenuating despite decreasing excitatory input are among the most challenging phenomena to imbue in a spiking neural network, particularly when simple disinhibition isn't at play.

Upon building the model with symmetrical CI between ensembles of neurons in L2/3 and L6 and added L6→L2/3 CLI that, we tested if the network could produce deviance-driven complex tuning in L2/3. We found that upon tuning the model to produce non-decreasing positive *and* negative deviance-evoked responses (averaged across all neurons in L2/3, stimulus-preferred and non-preferred), deviance-driven complex tuning indeed emerged.

In practice, model tuning involved a complex process that required us to modify network connection weights and the rate of thalamocortical adaptation within the constrains of symmetrical CI and selective (dis)inhibition via L6→L2/3 CLI, from the base values distributed with HNNcore [37] in a manner that maintained sparsity of baseline and evoked spike rates and sensitivity to recurrent stimulus-evoked drive (see Materials and Methods). Due to the stochastic nature of network connections and the spike times that perturb the network, a batch simulation of the final tuned model—for which spiking dynamics are shown in Figure 5f—consisted of 40 trials with random input spike times (Poisson-distributed for baseline drive, Gaussian-distributed for evoked) and random network connections (uniformly sampled among neuron pairs based on an average probability defined for cell type).

In a representative example batch simulation (consisting of 40 random trials) of the final tuned model (Figure 5e,f), notice that the number of cells targeted by the evoked drive (empty arrows; Figure 5e) remains constant across each stimulus repetition until the end, at which point a deviant decrease ($-\Delta$) occurs (solid black arrows). Note that we designed deviant afferent drives to maintain the same ratio of post-synaptic targets between ensembles (preferred to

non-preferred) as for the non-deviant drives, ensuring that any emergent shifts induced by deviance were the product of intrinsic network dynamics rather than an external shift in, e.g., thalamic tuning. The mean single-unit spike rate across simulation trials and neurons increased following each proximal and distal drive in L2/3 (red and blue traces), but only for proximal drive in L6 due to the fact that L6 lacks direct distal input (Figure 5f). For both L2/3 and L6, the mean evoked response in the preferred ensemble (red) always peaked higher than for the non-preferred ensemble (blue), reflecting the difference in drive strength received by each ensemble. The mean peak response across ensembles (horizontal purple dotted line) slightly decreases with each repetition due too the modeled effect of thalamocortical adaption, which decreases the synaptic strength of afferent drive with each stimulus repetition. On the final repetition of proximal and distal afferent drives, mean L6 peak activity decreases significantly with the deviant change in proximal drive, but L2/3 evoked activity merely keeps pace with thalamocortical adaptation (Figure 6a, purple).

We tested for the presence of deviance-driven complex tuning by comparing the difference between preferred (eP) and non-preferred (nP) evoked peak responses on deviant versus repetition-matched non-deviant (control) trials (Figure 6a,b). For a negative deviant where the afferent drive to the network decreases, the average neural response aggregated across all excitatory neurons in L2/3 (purple) doesn't significantly differ from that of the non-deviant control (Figure 6a; two-tailed Wilcoxon signed-rank test across trials/networks, P-value=0.217). When grouped according to their respective ensembles (red vs. blue), however, the response of excitatory neurons in the preferred ensemble (eP) were visibly higher than that of the non-preferred ensemble (eNP) for both the deviant and control, as expected given differing levels of evoked drive. Deviance-driven complex tuning emerged in the model as an increased difference between the eP and eNP peak responses (i.e., responder rank reinforcement; Figure 6b; upper-tailed Wilcoxon signed-rank test, P-value=0.011).

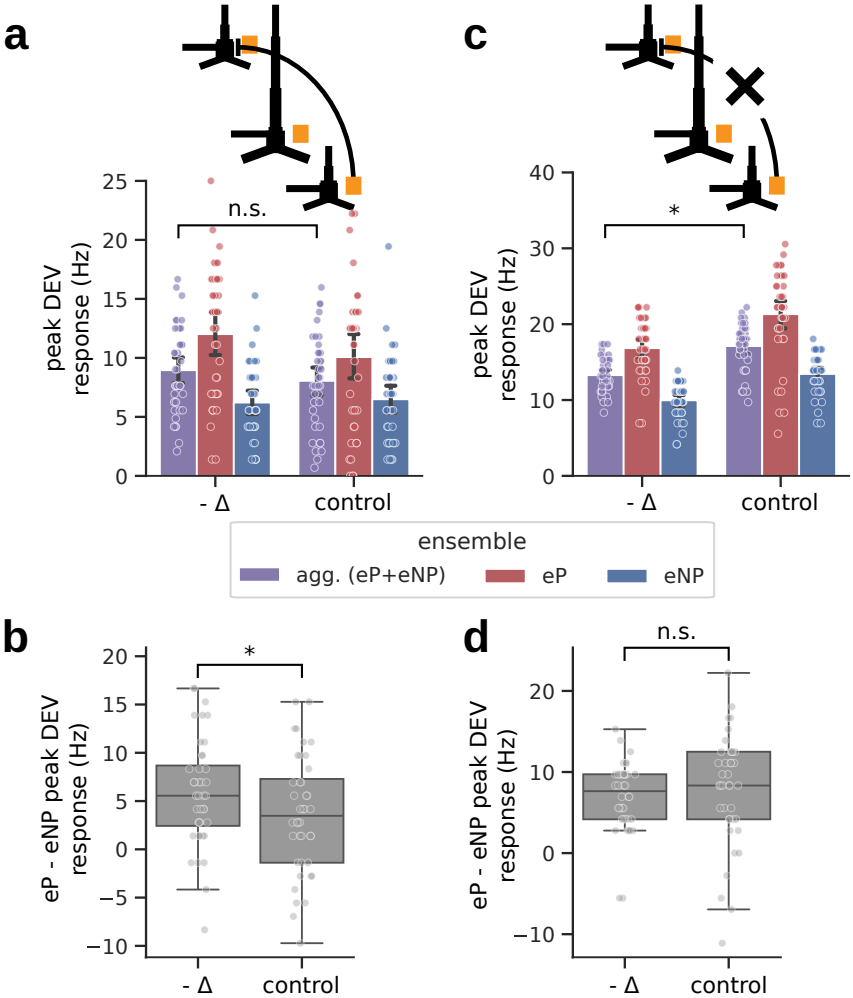

Figure 6: Despite the added constraints of biophysical realism and anatomical connectivity rules, deviance-driven complex tuning emerges with competitive inhibition and cross-laminar selective (dis)inhibition in a spiking neocortical column model. Consistent with dynamics of the neural mass model, selective (dis)inhibition is essential. (a) Mean deviant evoked response across n=40 random trials/networks with 95% confidence interval estimated via bootstrap over 1000 permutations. For a negative deviant ($-\Delta$) where the afferent drive to the network decreases, the average neural response aggregated across all excitatory neurons in L2/3 (purple) doesn't decrease compared to the non-deviant control. (b) Box plot showing the quartiles of the difference between preferred (eP) and non-preferred (nP) evoked responses across paired samples [red vs. blue in (a)], each pair from a given random network. The difference in evoked response is larger for the negative deviant compared to control indicating significant deviance-driven tuning, specifically, responder rank reinforcement. (c) With ablation of the CLI connection, the decrease in afferent drive correspondingly decreases the mean deviant evoked response across all ensembles compared to control. (d) Similar to (c), except where ablation of the CLI connection correspondingly removes deviance-driven tuning. *P-value<0.05.

Further, we verified that the emergence of deviance-driven complex tuning was sensitive to removal of selective (dis)inhibition by ablating the CLI connection in the model (Figure 6b,c). With ablation of the CLI connection (Figure 6c), the decrease in afferent drive correspondingly decreases the aggregate deviant evoked response compared to control (P-value$=4.19\times10^{-7}$), demonstrating that a fundamental condition of DD (maintaining, if not increasing, the net salience of the neocortical area's output) was lost. Significant deviance-driven complex tuning

was also lost (Figure 6d; P-value=0.916). Consistent with our theoretic results from the simple neural mass model earlier in this study, CLI ablation didn't significantly change baseline activity in L2/3 (data not shown), but rather, shifted how the two different ensembles responded to stimulus-evoked drive and each other.

It's important to note that simulations with this particular tuning of the biophysically-detailed model do not rule out the possibility that the other forms of deviance-driven complex tuning (i.e., responder rank weakening and rank reassignment as depicted in Figure 1d) may indeed be achievable with different network configurations (e.g., with a different balance of ipsi- versus contra-representational CLI or with a different level of within-ensemble excitatory synaptic strength). Here, we demonstrate that despite the added constraints of biophysical realism and anatomical connectivity rules, deviance-driven complex tuning emerges with CI and cross-laminar selective (dis)inhibition in a spiking neocortical column model. Consistent with dynamics of the neural mass model and in support of our overarching hypothesis, selective (dis)inhibition is essential.

Discussion

In this study we introduce and explore a novel theory that the neural dynamics subserving DD in the early-latency phase of a stimulus-evoked response are the product of competing neural representations, each representation encoding specific features of the stimulus and/or contextual environment. The network mechanisms of this process we term CI and SI, which together imply a specialized network connectivity motif that promotes a handful of emergent neural phenomena previously unaccounted for from a dynamical systems perspective.

CI is the process by which at least two neural representations embodied by distinct (but not necessarily mutually exclusive) ensembles of neurons reciprocally inhibit each other and augment the difference in stimulus tuning with each repetition of the same stimulus. SI is the process by which evoked afferent drive modulates the speed and extent of CI via feed-forward inhibition, establishing a novel E/I landscape for the competing ensembles to navigate whenever the raw intensity of a stimulus (or its relative composition of features) deviate. We first show that deviance-driven complex tuning emerges with CI and SI in a 4-dimensional neural mass model, demonstrating sufficiency of the underlying CI+SI motif for facilitating the most fundamental criteria of DD: selective amplification of the deviance-driven response from a subset of neurons at the expense of others without decreasing the net output of the network.

It has been widely accepted that receptive fields shift from one context to the next based on stimulus history, which for predictive processing has typically been associated with top-down or bottom-up modulation from other brain regions relative to the cortical hierarchy (e.g., as in hierarchical predictive coding [2, 4, 43, 44]). Inspired by prior work that demonstrates fast time-scale inhibition can dynamically shape functional neural representations before and after network learning [39, 45–48], we establish that recurrent CI allows for recent historical context (i.e., the basis of a sensory prediction) to be computed and stored locally through a stimulus' representational evolution in state space, as induced by repetitive afferent drive. Such a mechanism relies on two features of functional cell-to-cell connectivity: 1) a high degree of interconnectedness among excitatory neurons of similar stimulus preference that 2) innervate fast-spiking (e.g., parvalbumin-positive, PV) interneurons with post-synaptic targets of a different stimulus preference. While the former condition receives strong support from in vivo calcium imaging combined with cell-matched in vitro slice recordings in L2/3 of primary visual neocortex (V1) [49, 50], the latter does not [48]. However, the experimentally observed PV functional organization may enforce CI dynamically when under the influence of other interneuron cell types common to L2/3, such as vasoactive intestinal peptide-expressing (VIP)

and somatostatin-expressing (SST) interneurons that are known to modulate functional connectivity. VIP have been shown to interact with SST to help regulate the gain of neural circuits in service of enhanced stimulus specificity without compromising network stability [51]. It's also important to note that placing the CI+SI motif within the context of ensembles rather than single neurons promotes robustness despite variance in single neuron activation, which is particularly notable in the sparse stimulus-driven representations of barrel cortex [39].

We further show that CI can prime the network by facilitating adapting neural representations (specifically, adaptive sharpening), a previously established phenomenon of neural tuning that had until now lacked theoretical backing for shared circuitry with DD [28]. We also establish that while SI is necessary for deviance-driven complex tuning, CI is not. Rather, CI enhances robustness to connectivity and temporal variation for the co-emergence of adaptive sharpening and deviance-driven complex tuning. This latter result is likely due to the existence of a highly sensitive parameter regime where intrinsic differences in afferent excitation are able to provide enough repetition-driven adaptation that SI can still instigate a deviance-driven shift in tuning. While this matter requires further investigation, our cumulative results demonstrate that the CI+SI motif is at minimum necessary for the robustness of deviance-driven shifts in neural tuning given intrinsic variance contained in most (if not all) real-life neural circuits and processes.

Finally, we extend the CI+SI motif to a spiking neocortical column network constrained with biophysical realism and anatomical connectivity rules. We show that deviance-driven complex tuning emerges from spiking interactions in and between ensembles of distinct neurons (as opposed to hypothetical neural mass units of aggregate spike rates), and that deviance-driven complex tuning is particularly sensitive to SI, as expected from results obtained using the neural mass model. While we make a few assumptions about the true neocortical elements underlying our proposed CI+SI motif, these assumptions in no way compromise the functional significance of our theory. In particular, we assume that SI is mediated by monosynaptic L6→L2/3 CLI. While there is experimental evidence for a causal link between specific ensembles of stimulus-driven L6 CT and the emergence of deviance-driven complex tuning [13], as well as an established cross-laminar interneuron in L6 that provides inhibitory gain modulation onto the more superficial layers of the neorcortex [35, 36], it could be that L6 only influences L2/3 through polysynaptic connections.

L6 is compelling as a potential source of SI for a number of other reasons given the unique stimulus tuning properties of L6 CT and their anatomical connectivity within thalamocortical circuitry. L6 CT exhibit distinctly sharp tuning profiles to external stimuli [52], allowing them to encode specific features of the sensory or contextual environment in a robust manner as sparse ensembles. They also receive direct projections from lemniscal thalamus and send direct projections to both lemniscal and non-lemniscal/high-order thalamus, giving CT a crucial role in mediating thalamocortical feedback and modulating contextual input that enters the superficial layers via non-lemniscal thalamus [52, 53].

DD is generally associated with a pronounced increase in evoked spike rates across neurons of an experimentally measured population [8–10]. Here, we focused on early-latency (< 100 ms) evoked response phenomena associated with DD that involves the selective amplification of the response of some neurons over others, especially highlighting the case where the mean response across neurons remains constant despite a deviant decrease in afferent drive. This is exemplified through the emergence of deviance-driven complex tuning as described in prior research [13] and depicted in Figure 1. Our hypothesized mechanisms of CI modulated by SI intrinsically produce ensemble-level evoked responses that shift over time and space/neurons, implying that the conjunctive (i.e., aggregate) representation across ensembles transiently shifts as well. From the perspective of an observer who is naive to the complete anatomical connectivity of a random

sample of neurons and the ensembles they belong to (as is the case in most electrophysiology experiments), the early-latency deviant response is predicted to correspond to a notable shift in stimulus decoding accuracy relative to baseline. The shift in stimulus encoding can either increase or decrease depending on the specific form of deviance-driven complex tuning: responder rank reinforcement corresponds to increased decoding accuracy, whereas responder rank weakening or reassignment corresponds to decreased decoding accuracy. Based on the results of this study, responder rank reinforcement is the most robust form of deviance-driven complex tuning and emerges exclusively from networks with CI and strong SI. Future in vivo experiments can test this theoretical prediction, along with whether functional ensembles of neurons classified according to their baseline tuning profile shift on deviant trials and whether that shift critically depends on cross-ensemble inhibition defined in the CI+SI motif.

Materials and Methods

Neural mass model simulations

The Wilson-Cowan neural mass model given by Equation 1 contained up to four units with a sigmoid activation function of the form

$$\begin{aligned} \mathcal{F}_i\left(y\right) &= \frac{1}{1+e^{-\gamma_i(y-0.5)}} \\ \gamma_i &= \begin{cases} 8, & i \in \{1,2\} \\ 15, & i \in \{3,4\} \end{cases} \end{aligned} \tag{2}$$

The response of each unit thus had a dynamic range in $(0,1)$, 0 representing a quiescent ensemble (i.e., with a spike rate of 0), 1 representing a maximally active ensemble (i.e., with a saturated peak spike rate). The steepness parameter γ was symmetrical across the stimulus-preferred and non-preferred representations and fixed at a smaller value for units in the upper layer (units 1 and 2) so that the dynamic ranges of their activation functions were greater, or slower acting, than for units of the lower layer (units 3 and 4). The time-varying injected current parameter $I(t)$ took a value of 0.01 at baseline and 0.55 and 0.45 for a duration of 20 ms during evoked drive for the preferred and non-preferred units, respectively. The magnitude of evoked drive increased or decreased by a multiple of 0.2 for positive or negative deviants, respectively. We simulated network dynamics using the Runge-Kutta method (RK4) with integration time step $\Delta t = 0.1$ ms, time constant $\tau = 20$ ms, and initial conditions $x_i = 0 \; \forall \; i \in \{1,2,3,4\}$ (Equation 1).

For the initial proof-of-principle of emergent dynamics examined further in this study, the specific connection weights for the model simulated in Figure 2c,d were selected using an optimization routine that targeted the mean deviance-driven response across units in the upper layer. Pairs of connection weights constrained by symmetry between the stimulus-preferred and non-preferred representations, as defined in the fundamental CI+SI motif, were allowed to vary during optimization, but `ipsirepr_inhib` was held fixed at 0.75. We first set parameter bounds that limited runaway excitation/saturation across units (`w_ii` $\in (0.1, 1)$, `w_ij` $\in (-1.1, -0.2)$, `w_ii_2` $\in (0.1, 1)$, `w_ij_2` $\in (-1.1, -0.2)$, and `w_cli` $\in (-1, -0.1)$, see Figure 3b for a graphical depiction of these parameters) and then ran a gradient decent optimization routine (`SciPy`'s L-BFGS-B implementation) over those ranges to minimize loss L given by

$$L = \begin{cases} \bar{x}_{-\Delta} \cdot \bar{x}_{+\Delta}, & \text{if} \quad \bar{x}_{-\Delta} < 0 \quad \text{and} \quad \bar{x}_{+\Delta} < 0 \\ -\bar{x}_{-\Delta} \cdot \bar{x}_{+\Delta}, & \text{otherwise} \end{cases} \tag{3}$$

where $\bar{x}_{-\Delta}$ and $\bar{x}_{+\Delta}$ denote the mean peak firing rates across units in the upper layer (x_1 and x_2) for the negative and positive evoked responses, respectively, minus that of the non-deviant evoked response. Note that loss L in Equation 3 is greater than zero iff either $\bar{x}_{-\Delta}$ or $\bar{x}_{+\Delta}$ fall below 0, allowing the optimization routine to minimize the extent to which a deviant evoked response decreases relative to a non-deviant response. The 2-dimensional model simulation of CI without SI shown in Figure 2b uses the same optimized parameters as that in Figure 2b, except with the upper and lower units disconnected.

While our goal was to observe if/how deviance-driven complex tuning might emerge from the model, it's important to note that we never explicitly forced the model to produce any of the specific forms of deviance-driven complex tuning (responder rank reinforcement, weakening, or reassignment, as depicted in Figure 1d). Instead, we optimized the model to produce a mean evoked firing rate that increases (or at least doesn't decrease) with deviant afferent drive regardless of the direction of the deviant change in stimulus amplitude, which is a fundamental aspect of DD and constraint on emergent dynamics. We had no prior guarantees that a non-decreasing mean deviance-driven response across ensemble representations could be achieved by this system given our CI+SI network motif; however, we reasoned that if the average rate didn't decrease on a deviant, the underlying system would be forced to compensate by amplifying the response of one, and only one, of the constituent units in the upper layer. If a positive deviant were to increase the evoked response of both units, the system would be overwhelmingly governed by afferent excitation and result in a decrease in the mean evoked response on a negative deviant. If, on the other hand, a positive deviant were to decrease the evoked response of both units, the system would be overwhelmingly governed by cross-laminar inhibition and result in an increase in the the mean evoked response on a negative deviant (i.e., from acute disinhibition).

Biophysically-detailed model simulations

All simulations of the biophysically-detaild model were run using HNN-core [37], a modular and expandable Python implementation of HNN [24]. HNN is based on a pre-constructed and pre-tuned model of a laminated neocortical circuit under thalamocortical and cortico-cortical influences. While creation of HNN was originally inspired by the desire to interpret the cellular and microcircuit origin of human magnetoencephalography (MEG) and electroencephalography (EEG) signals by relating the primary electrical currents (i.e., current source dipoles) in pyramidal neuron dendrites to a variety of commonly measured neurophysiology data types including local network spiking activity, local field potential (LFP), and laminar current-source density (CSD), it was built using canonical features of neocortical dynamics and circuitry conserved across rodent, non-human primate, and human anatomy, thus allowing cross-species inference [24, 40, 41]. Here we used the HNN model to validate our theoretical results obtained in a neural mass model of ensemble spike rates to spiking dynamics between various neuron types and layers as constrained by the anatomy of a neocortical column.

Aside from stochastic baseline drive that is provided as 20 Hz Poisson pre-synaptic spike events with a 300 ms burn-in, activation of the HNN network is achieved by defining patterns of action potentials representing activity in exogenous brain areas (e.g., thalamus or higher order cortex) that drive the local network through layer specific patterns of excitatory synaptic inputs. Synaptic inputs that drive the network target the dendrites of excitatory pyramidal cells and the somas of inhibitory basket cells. One pathway of input simulates perturbations ascending from the peripheral nervous system, through the lemniscal thalamus to the granular layer (L4), and finally propagating to the proximal dendrites of the pyramidal neurons in L2/3 and L5, as well as the distal dendrites of pyramidal neurons in L6 [54]. These inputs are referred to as proximal drive. The other pathway of input comes from the non-lemniscal thalamus or high-order cortex and synapses in L2/3, targeting the distal dendrites of the pyramidal neurons. These inputs

are referred to as distal drive [24].

It is important to acknowledge that there are numerous types of neocortical column models with varying levels of detail; however, this model was selected for this study because it balances biophysical detail with simplifying assumptions that make manipulations of within-layer and between-layer connectivity tractable and simulations computationally efficient compared to other large-scale models. Namely, there are a handful of cell types, each of which is composed of up to a nine compartments governed by Hodgkin-Huxley cable equations implemented on the backend in NEURON [55]. By default, the HNN model has two primary layers (L2/3 and L5), however, we added L6 (comprised of two cell types, pyramidal cells and inhibitory basket cells) for the present study. Neurons within a given layer were placed at regular intervals in the laminar plane such that we could modulate the degree of lateral and vertical connectivity through three parameters: maximal synaptic conductance, horizontal spatial decay (i.e., the rate at which the synaptic conductance decays with horizontal distance between two neurons), and connection probability (see Figure 5 for a graphical depiction of the network connectivity rules). The position of neurons belonging to the preferred (P) and non-preferred (NP) representations were interleaved across a hexagonal 2D mesh within each layer to ensure that excitation within and competitive inhibition between ensembles was symmetrical despite the 1:3 proportion of inhibitory-to-excitatory neurons.

Consistent with prior recordings that observed that L4 and L6 neurons receive direct afferent projections from sensory thalamus [56] and that stimulus-driven activation of local inhibition generally proceeds excitation in distinct representational ensembles [39], synaptic delays of proximal drive in our model were chosen to activate L6 apical dendrites first, followed by the activation of L2/3 inhibitory interneurons and then excitatory pyramidal neurons. Stimulus-driven activity in L2/3 was therefore shaped primarily by the strength of proximal drive propagating from L4 combined with polysynaptic activity routed through L6. Given the purpose of this model and the scope of phenomena we wished to observe, L4 and L5 were assumed to primarily act as relay nodes within the column. This does not diminish the fact that L4 and L5 are known to play complex and crucial roles in other types of cortico-cortical and thalamocortical processing, such as in in the dynamic shaping of spatial receptive fields in L4 neurons [12, 25, 32] and dendritic integration in L5 pyramidal cells [57, 58].

Biophysically-detailed model parameter tuning

Tuning of the biophysically-detailed model involved an iterative procedure until the target baseline (Figure S2) and evoked spike rates were achieved. Due to the highly non-linear and cost-intensive nature of running simulations in this model, all tuning was conducted manually to produce resting state spiking dynamics consistent with in vivo mouse recordings from prior studies [38, 39]. The procedure was as follows.

1. Fix cell positions, morphology and biophysics.
2. Connect Poisson drive to the soma of each cell independently.
3. In the disconnected network, tune Poisson drive synaptic weight parameters to achieve the disconnected (intermediate) average baseline firing rate for each cell type, which is a fraction of the final target average firing rate for each cell type.
4. With the Poisson drive synaptic weight parameters fixed, provisionally fix the local network connectivity parameters (this includes a pairwise connection probability for each connection type and spatial synaptic weight decay).
5. Instantiate the random network connections with a seeded psuedorandom number generator.

6. Tune maximal conductance (synaptic strength) parameters for each connection type in the local network:
    - 6.1 Achieve the target connected (final) average baseline firing rate for each cell type [38, 39].
    - 6.2 Achieve limited spatiotemporal spike correlations by visual inspection of 2 seconds of simulated baseline activity.
    - 6.3 If the network cannot accommodate these achievements with a manual parameter search, return to (4).
7. With local network connectivity parameters provisionally fixed, tune maximal conductance (synaptic strength) and thalamocortical adaptation parameters for each type of evoked drive (proximal and distal):
    - 7.1 Achieve reasonable evoked spike rates that don't decrease their peak magnitude (averaged across ensembles) for both positive *and* negative deviants.
    - 7.2 Achieve non-saturating evoked spike rates in each neural population.
    - 7.3 Achieve evoked spike rates that appear to be sensitive to prior evoked activity.
    - 7.4 If the network cannot accommodate these achievements with a manual parameter search, return to (6).

The final tuned parameters and source code for running simulations of this model are publicly available and can be accessed at `https://github.com/rythorpe/hnn-core/tree/L6_model`.

Acknowledgments

We thank Scott Susi, Christopher Deister, and Jakob Voigts for helpful discussion on nuances of neocortical and thalamocortical anatomy, realistic spiking behavior in selected cell types, as well as the pitfalls of various simplifying assumptions within the biophysically-detailed model used in this study. We also thank Nicholas Tolley for guidance on implementing parameter sweeps via high-performance computing and providing crucial insight and discussion regarding avenues for statistical inference. Model simulations were conducted using computational resources and services at the Center for Computation and Visualization, Brown University. This study was supported by the National Institute of Neurological Disorders and Stroke (R01NS108414, U24NS129945), and the National Institute of Mental Health (T32MH115895) of the National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Supplementary Material

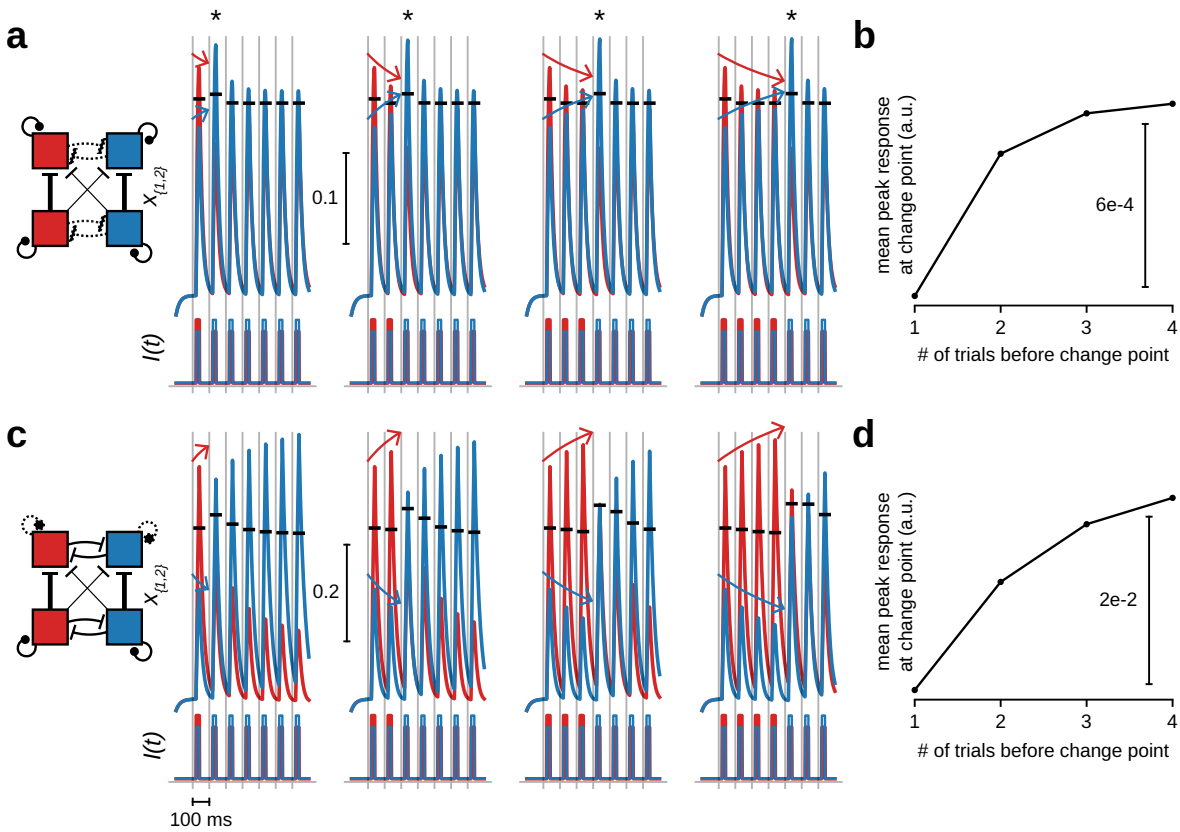

Figure S1: Competitive inhibition within the upper layer leads to ensemble priming, a network mechanism for context storage and history dependence. With pronounced within-layer inhibition (a) or within-unit excitation (c) that causes evoked responses between the upper layer units (i.e., x_1 and x_2) to either converge (adaptive broadening) or diverge (adaptive sharpening) with more stimulus repetitions of a particular feature preference, the average deviance-induced evoked response on a change point (indicated by an asterisk) increases (b,d). Note that in this stimulus paradigm, the relative balance of salient stimulus (or contextual) features serving as inputs to the system can be repeated and stored over time (e.g., red stimulus component > blue stimulus component or vice versa), as indicated by the history-dependent shift in deviant response magnitude.

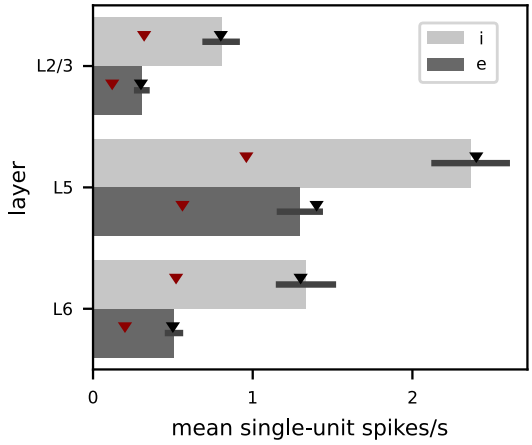

Figure S2: Target (triangles) and achieved (bars) mean baseline spike rates used for tuning the baseline activity of the biophyscially-detailed model. Poisson drive parameters were first tuned to match a fraction (0.4, red) of the final target baseline spike rates (black) in a disconnected network. Then the network connectivity was instantiated and tuned to increase the mean baseline spike rates from their intermediate to final target values (from top to bottom, L2/3-L6, 0.8, 0.3, 2.4, 1.4, 1.3, and 0.5 Hz). Tuning was considered complete once the final target fell within one standard deviation of the simulated mean (S.E., black error bars).

References

1. Keller, G. B. & Mrsic-Flogel, T. D. Predictive Processing: A Canonical Cortical Computation. *Neuron* **100.** Publisher: Elsevier, 424–435. ISSN: 0896-6273. https://www.cell.com/neuron/abstract/S0896-6273(18)30857-2 (2024) (Oct. 24, 2018).
2. Rao, R. P. N. & Ballard, D. H. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience* **2.** Publisher: Nature Publishing Group, 79–87. ISSN: 1546-1726. https://www.nature.com/articles/nn0199_79 (2024) (Jan. 1999).
3. Arnal, L. H. & Giraud, A.-L. Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences* **16,** 390–398. ISSN: 1364-6613. http://www.sciencedirect.com/science/article/pii/S1364661312001210 (2021) (July 1, 2012).
4. Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P. & Friston, K. J. Canonical Microcircuits for Predictive Coding. *Neuron* **76,** 695–711. ISSN: 0896-6273. http://www.sciencedirect.com/science/article/pii/S0896627312009592 (2021) (Nov. 21, 2012).
5. Carbajal, G. V. & Malmierca, M. S. The Neuronal Basis of Predictive Coding Along the Auditory Pathway: From the Subcortical Roots to Cortical Deviance Detection. *Trends in Hearing* **22.** Publisher: SAGE Publications Inc, 2331216518784822. ISSN: 2331-2165. https://doi.org/10.1177/2331216518784822 (2021) (Jan. 1, 2018).
6. Parras, G. G., Nieto-Diego, J., Carbajal, G. V., Valdés-Baizabal, C., Escera, C. & Malmierca, M. S. Neurons along the auditory pathway exhibit a hierarchical organization of prediction error. *Nature Communications* **8.** Bandiera_abtest: a Cc_license_type: cc_by Cg_type: Nature Research Journals Number: 1 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Cognitive neuroscience;Cortex;Midbrain;Sensory processing;Thalamus Subject_term_id: cognitive-neuroscience;cortex;midbrain;sensory-processing;thalamus, 2148. ISSN: 2041-1723. https://www.nature.com/articles/s41467-017-02038-6 (2021) (Dec. 15, 2017).

7. Marvan, T. & Phillips, W. A. Cellular mechanisms of cooperative context-sensitive predictive inference. *Current Research in Neurobiology* **6,** 100129. ISSN: 2665-945X. https://www.sciencedirect.com/science/article/pii/S2665945X24000068 (2024) (Jan. 1, 2024).
8. Hamm, J. P., Shymkiv, Y., Han, S., Yang, W. & Yuste, R. Cortical ensembles selective for context. *Proceedings of the National Academy of Sciences* **118.** Publisher: Proceedings of the National Academy of Sciences, e2026179118. https://www.pnas.org/doi/10.1073/pnas.2026179118 (2023) (Apr. 6, 2021).
9. Gallimore, C. G., Ricci, D. A. & Hamm, J. P. Spatiotemporal dynamics across visual cortical laminae support a predictive coding framework for interpreting mismatch responses. *Cerebral Cortex,* bhad215. ISSN: 1047-3211. https://doi.org/10.1093/cercor/bhad215 (2023) (June 12, 2023).
10. Musall, S., Haiss, F., Weber, B. & von der Behrens, W. Deviant Processing in the Primary Somatosensory Cortex. *Cerebral Cortex* **27,** 863–876. ISSN: 1047-3211. https://doi.org/10.1093/cercor/bhv283 (2021) (Jan. 1, 2017).
11. Hubel, D. H. & Wiesel, T. N. Receptive fields of single neurones in the cat's striate cortex. *The Journal of Physiology* **148.** _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1113/jphysiol.1959.sp0 574–591. ISSN: 1469-7793. https://onlinelibrary.wiley.com/doi/abs/10.1113/jphysiol.1959.sp006308 (2024) (1959).
12. Somers, D. C., Nelson, S. B. & Sur, M. An emergent model of orientation selectivity in cat visual cortical simple cells. *Journal of Neuroscience* **15.** Publisher: Society for Neuroscience Section: Articles, 5448–5465. ISSN: 0270-6474, 1529-2401. https://www.jneurosci.org/content/15/8/5448 (2024) (Aug. 1, 1995).
13. Voigts, J., Deister, C. A. & Moore, C. I. Layer 6 ensembles can selectively regulate the behavioral impact and layer-specific representation of sensory deviants. *eLife* **9** (eds Colgin, L. L., Huguenard, J. R., Stanley, G. B. & Paz, J. T.) Publisher: eLife Sciences Publications, Ltd, e48957. ISSN: 2050-084X. https://doi.org/10.7554/eLife.48957 (2021) (Dec. 2, 2020).
14. Dragoi, V., Sharma, J., Miller, E. K. & Sur, M. Dynamics of neuronal sensitivity in visual cortex and local feature discrimination. *Nature Neuroscience* **5.** Publisher: Nature Publishing Group, 883–891. ISSN: 1546-1726. https://www.nature.com/articles/nn900 (2024) (Sept. 2002).
15. Grill-Spector, K., Henson, R. & Martin, A. Repetition and the brain: neural models of stimulus-specific effects. *Trends in Cognitive Sciences* **10.** Publisher: Elsevier, 14–23. ISSN: 1364-6613, 1879-307X. https://www.cell.com/trends/cognitive-sciences/abstract/S1364-6613(05)00323-2 (2024) (Jan. 1, 2006).
16. Rideaux, R., West, R. K., Rangelov, D. & Mattingley, J. B. Distinct early and late neural mechanisms regulate feature-specific sensory adaptation in the human visual system. *Proceedings of the National Academy of Sciences* **120.** Publisher: Proceedings of the National Academy of Sciences, e2216192120. https://www.pnas.org/doi/10.1073/pnas.2216192120 (2024) (Feb. 7, 2023).
17. Phinney, R. E., Bowd, C. & Patterson, R. Direction-selective Coding of Stereoscopic (Cyclopean) Motion. *Vision Research* **37,** 865–869. ISSN: 0042-6989. https://www.sciencedirect.com/science/article/pii/S0042698996002441 (2024) (Apr. 1, 1997).
18. Clifford, C. W. G., Wyatt, A. M., Arnold, D. H., Smith, S. T. & Wenderoth, P. Orthogonal adaptation improves orientation discrimination. *Vision Research* **41,** 151–159. ISSN: 0042-6989. https://www.sciencedirect.com/science/article/pii/S0042698900002480 (2024) (Jan. 1, 2001).
19. Heeger, D. J. Normalization of cell responses in cat striate cortex. *Visual Neuroscience* **9,** 181–197. ISSN: 1469-8714, 0952-5238. https://www.cambridge.org/core/journals/visual-neuroscience/article/abs/normalization-of-cell-responses-in-cat-

striate-cortex/0851FEE8DEE00514E1A432123E703643#access-block (2024) (Aug. 1992).

20. Carandini, M. & Heeger, D. J. Normalization as a canonical neural computation. *Nature Reviews Neuroscience* **13.** Publisher: Nature Publishing Group, 51–62. ISSN: 1471-0048. https://www.nature.com/articles/nrn3136 (2024) (Jan. 2012).
21. Ross, J. M. & Hamm, J. P. Cortical Microcircuit Mechanisms of Mismatch Negativity and Its Underlying Subcomponents. *Frontiers in Neural Circuits* **14.** ISSN: 1662-5110. https://www.frontiersin.org/articles/10.3389/fncir.2020.00013 (2023) (2020).
22. Wilson, H. R. & Cowan, J. D. Excitatory and Inhibitory Interactions in Localized Populations of Model Neurons. *Biophysical Journal* **12,** 1–24. ISSN: 00063495. https://linkinghub.elsevier.com/retrieve/pii/S0006349572860685 (2024) (Jan. 1972).
23. Wilson, H. R. & Cowan, J. D. A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik* **13,** 55–80. ISSN: 1432-0770. https://doi.org/10.1007/BF00288786 (2024) (Sept. 1, 1973).
24. Neymotin, S. A., Daniels, D. S., Caldwell, B., McDougal, R. A., Carnevale, N. T., Jas, M., Moore, C. I., Hines, M. L., Hämäläinen, M. & Jones, S. R. Human Neocortical Neurosolver (HNN), a new software tool for interpreting the cellular and network origin of human MEG/EEG data. *eLife* **9** (eds Ivry, R. B., Stolk, A., Stolk, A. & Dalal, S. S.) Publisher: eLife Sciences Publications, Ltd, e51214. ISSN: 2050-084X. https://doi.org/10.7554/eLife.51214 (2020) (Jan. 22, 2020).
25. Moore, C. I. & Nelson, S. B. Spatio-Temporal Subthreshold Receptive Fields in the Vibrissa Representation of Rat Primary Somatosensory Cortex. *Journal of Neurophysiology* **80.** Publisher: American Physiological Society, 2882–2892. ISSN: 0022-3077. https://journals.physiology.org/doi/full/10.1152/jn.1998.80.6.2882 (2024) (Dec. 1998).
26. Vierck, C. J. & Jones, M. B. Influences of low and high frequency oscillation upon spatio-tactile resolution. *Physiology & Behavior* **5,** 1431–1435. ISSN: 0031-9384. https://www.sciencedirect.com/science/article/pii/0031938470901320 (2024) (Dec. 1, 1970).
27. Tannan, V., Whitsel, B. L. & Tommerdahl, M. A. Vibrotactile adaptation enhances spatial localization. *Brain Research* **1102,** 109–116. ISSN: 0006-8993. https://www.sciencedirect.com/science/article/pii/S0006899306014028 (2024) (Aug. 2, 2006).
28. Ollerenshaw, D. R., Zheng, H. J. V., Millard, D. C., Wang, Q. & Stanley, G. B. The Adaptive Trade-Off between Detection and Discrimination in Cortical Representations and Behavior. *Neuron* **81,** 1152–1164. ISSN: 0896-6273. https://www.sciencedirect.com/science/article/pii/S0896627314000567 (2024) (Mar. 5, 2014).
29. Wiggs, C. L. & Martin, A. Properties and mechanisms of perceptual priming. *Current Opinion in Neurobiology* **8,** 227–233. ISSN: 0959-4388. https://www.sciencedirect.com/science/article/pii/S095943889880144X (2024) (Apr. 1, 1998).
30. Chung, S., Li, X. & Nelson, S. B. Short-Term Depression at Thalamocortical Synapses Contributes to Rapid Adaptation of Cortical Sensory Responses In Vivo. *Neuron* **34.** Publisher: Elsevier, 437–446. ISSN: 0896-6273. https://www.cell.com/neuron/abstract/S0896-6273(02)00659-1 (2024) (Apr. 25, 2002).
31. Desimone, R. Neural mechanisms for visual memory and their role in attention. *Proceedings of the National Academy of Sciences* **93.** Publisher: Proceedings of the National Academy of Sciences, 13494–13499. https://www.pnas.org/doi/10.1073/pnas.93.24.13494 (2024) (Nov. 26, 1996).
32. Moore, C. I., Nelson, S. B. & Sur, M. Dynamics of neuronal processing in rat somatosensory cortex. *Trends in Neurosciences* **22,** 513–520. ISSN: 01662236. https://linkinghub.elsevier.com/retrieve/pii/S0166223699014526 (2021) (Nov. 1999).
33. Moore, C. I. Frequency-Dependent Processing in the Vibrissa Sensory System. *Journal of Neurophysiology* **91.** Publisher: American Physiological Society, 2390–2399. ISSN: 0022-

3077. https://journals.physiology.org/doi/full/10.1152/jn.00925.2003 (2024) (June 2004).

34. Garabedian, C. E., Jones, S. R., Merzenich, M. M., Dale, A. & Moore, C. I. Band-Pass Response Properties of Rat SI Neurons. *Journal of Neurophysiology* **90.** Publisher: American Physiological Society, 1379–1391. ISSN: 0022-3077. https://journals.physiology.org/doi/full/10.1152/jn.01158.2002 (2022) (Sept. 2003).
35. Olsen, S. R., Bortone, D. S., Adesnik, H. & Scanziani, M. Gain control by layer six in cortical circuits of vision. *Nature* **483.** Number: 7387 Publisher: Nature Publishing Group, 47–52. ISSN: 1476-4687. https://www.nature.com/articles/nature10835 (2021) (Mar. 2012).
36. Bortone, D. S., Olsen, S. R. & Scanziani, M. Translaminar Inhibitory Cells Recruited by Layer 6 Corticothalamic Neurons Suppress Visual Cortex. *Neuron* **82,** 474–485. ISSN: 0896-6273. https://www.sciencedirect.com/science/article/pii/S0896627314001524 (2021) (Apr. 16, 2014).
37. Jas, M., Thorpe, R., Tolley, N., Bailey, C., Brandt, S., Caldwell, B., Cheng, H., Daniels, D., Pujol, C. F., Khalil, M., Kanekar, S., Kohl, C., Kolozsvári, O., Lankinen, K., Loi, K., Neymotin, S., Partani, R., Pelah, M., Rockhill, A., Sherif, M., Hamalainen, M. & Jones, S. HNN-core: A Python software for cellular and circuit-level interpretation of human MEG/EEG. *Journal of Open Source Software* **8,** 5848. ISSN: 2475-9066. https://joss.theoj.org/papers/10.21105/joss.05848 (2024) (Dec. 15, 2023).
38. De Kock, C. P. J., Bruno, R. M., Spors, H. & Sakmann, B. Layer- and cell-type-specific suprathreshold stimulus representation in rat primary somatosensory cortex. *The Journal of Physiology* **581.** _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1113/jphysiol.2006.124321, 139–154. ISSN: 1469-7793. https://onlinelibrary.wiley.com/doi/abs/10.1113/jphysiol.2006.124321 (2023) (2007).
39. Reyes-Puerta, V., Sun, J.-J., Kim, S., Kilb, W. & Luhmann, H. J. Laminar and Columnar Structure of Sensory-Evoked Multineuronal Spike Sequences in Adult Rat Barrel Cortex In Vivo. *Cerebral Cortex* **25,** 2001–2021. ISSN: 1047-3211. https://doi.org/10.1093/cercor/bhu007 (2023) (Aug. 1, 2015).
40. Jones, S. R., Pritchett, D. L., Stufflebeam, S. M., Hamalainen, M. & Moore, C. I. Neural Correlates of Tactile Detection: A Combined Magnetoencephalography and Biophysically Based Computational Modeling Study. *Journal of Neuroscience* **27,** 10751–10764. ISSN: 0270-6474, 1529-2401. http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.0482-07.2007 (2020) (Oct. 3, 2007).
41. Jones, S. R., Pritchett, D. L., Sikora, M. A., Stufflebeam, S. M., Hämäläinen, M. & Moore, C. I. Quantitative Analysis and Biophysically Realistic Neural Modeling of the MEG Mu Rhythm: Rhythmogenesis and Modulation of Sensory-Evoked Responses. *Journal of Neurophysiology* **102.** Publisher: American Physiological Society, 3554–3572. ISSN: 0022-3077. https://journals.physiology.org/doi/full/10.1152/jn.00535.2009 (2020) (Oct. 7, 2009).
42. Thorpe, R. V., Black, C. J., Borton, D. A., Hu, L., Saab, C. Y. & Jones, S. R. Distinct neocortical mechanisms underlie human SI responses to median nerve and laser-evoked peripheral activation. *Imaging Neuroscience* **2,** 1–29. ISSN: 2837-6056. https://doi.org/10.1162/imag_a_00095 (2024) (Feb. 22, 2024).
43. Friston, K. A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences* **360.** Publisher: Royal Society, 815–836. https://royalsocietypublishing.org/doi/full/10.1098/rstb.2005.1622 (2021) (Apr. 29, 2005).
44. Auksztulewicz, R. & Friston, K. Repetition suppression and its contextual determinants in predictive coding. *Cortex. Special Issue:Repetition suppression-an integrative view* **80,** 125–140. ISSN: 0010-9452. https://www.sciencedirect.com/science/article/pii/S0010945216000101 (2021) (July 1, 2016).

45. Sanzeni, A., Palmigiano, A., Nguyen, T. H., Luo, J., Nassi, J. J., Reynolds, J. H., Histed, M. H., Miller, K. D. & Brunel, N. Mechanisms underlying reshuffling of visual responses by optogenetic stimulation in mice and monkeys. *Neuron* **111,** 4102–4115.e9. ISSN: 0896-6273. https://www.sciencedirect.com/science/article/pii/S0896627323007067 (2024) (Dec. 20, 2023).
46. Khan, A. G., Poort, J., Chadwick, A., Blot, A., Sahani, M., Mrsic-Flogel, T. D. & Hofer, S. B. Distinct learning-induced changes in stimulus selectivity and interactions of GABAergic interneuron classes in visual cortex. *Nature Neuroscience* **21.** Publisher: Nature Publishing Group, 851–859. ISSN: 1546-1726. https://www.nature.com/articles/s41593-018-0143-z (2024) (June 2018).
47. Machens, C. K., Romo, R. & Brody, C. D. Flexible Control of Mutual Inhibition: A Neural Model of Two-Interval Discrimination. *Science.* Publisher: American Association for the Advancement of Science. https://www.science.org/doi/10.1126/science.1104171 (2024) (Feb. 18, 2005).
48. Znamenskiy, P., Kim, M.-H., Muir, D. R., Iacaruso, M. F., Hofer, S. B. & Mrsic-Flogel, T. D. Functional specificity of recurrent inhibition in visual cortex. *Neuron* **112,** 991–1000.e8. ISSN: 08966273. https://linkinghub.elsevier.com/retrieve/pii/S0896627323009728 (2024) (Mar. 2024).
49. Ko, H., Hofer, S. B., Pichler, B., Buchanan, K. A., Sjöström, P. J. & Mrsic-Flogel, T. D. Functional specificity of local synaptic connections in neocortical networks. *Nature* **473.** Publisher: Nature Publishing Group, 87–91. ISSN: 1476-4687. https://www.nature.com/articles/nature09880 (2024) (May 2011).
50. Cossell, L., Iacaruso, M. F., Muir, D. R., Houlton, R., Sader, E. N., Ko, H., Hofer, S. B. & Mrsic-Flogel, T. D. Functional organization of excitatory synaptic strength in primary visual cortex. *Nature* **518.** Publisher: Nature Publishing Group, 399–403. ISSN: 1476-4687. https://www.nature.com/articles/nature14182 (2024) (Feb. 2015).
51. Millman, D. J., Ocker, G. K., Caldejon, S., Kato, I., Larkin, J. D., Lee, E. K., Luviano, J., Nayan, C., Nguyen, T. V., North, K., Seid, S., White, C., Lecoq, J., Reid, C., Buice, M. A. & de Vries, S. E. VIP interneurons in mouse primary visual cortex selectively enhance responses to weak but specific stimuli. *eLife* **9** (eds Vinck, M. & Wassum, K. M.) Publisher: eLife Sciences Publications, Ltd, e55130. ISSN: 2050-084X. https://doi.org/10.7554/eLife.55130 (2021) (Oct. 27, 2020).
52. Crandall, S. R., Patrick, S. L., Cruikshank, S. J. & Connors, B. W. Infrabarrels Are Layer 6 Circuit Modules in the Barrel Cortex that Link Long-Range Inputs and Outputs. *Cell Reports* **21,** 3065–3078. ISSN: 2211-1247. https://www.sciencedirect.com/science/article/pii/S2211124717316868 (2023) (Dec. 12, 2017).
53. Crandall, S. R., Cruikshank, S. J. & Connors, B. W. A Corticothalamic Switch: Controlling the Thalamus with Dynamic Synapses. *Neuron* **86.** Publisher: Elsevier, 768–782. ISSN: 0896-6273. https://www.cell.com/neuron/abstract/S0896-6273(15)00264-0 (2021) (May 6, 2015).
54. Oberlaender, M., de Kock, C. P. J., Bruno, R. M., Ramirez, A., Meyer, H. S., Dercksen, V. J., Helmstaedter, M. & Sakmann, B. Cell Type–Specific Three-Dimensional Structure of Thalamocortical Circuits in a Column of Rat Vibrissal Cortex. *Cerebral Cortex* **22,** 2375–2391. ISSN: 1047-3211. https://doi.org/10.1093/cercor/bhr317 (2023) (Oct. 1, 2012).
55. Hines, M. L. & Carnevale, N. T. The NEURON simulation environment. *Neural computation* **9.** Publisher: MIT Press, 1179–1209 (1997).
56. Constantinople, C. M. & Bruno, R. M. Deep Cortical Layers Are Activated Directly by Thalamus. *Science* **340.** Publisher: American Association for the Advancement of Science, 1591–1594. https://www.science.org/doi/full/10.1126/science.1236425 (2023) (June 28, 2013).

57. Aru, J., Suzuki, M. & Larkum, M. E. Cellular Mechanisms of Conscious Processing. *Trends in Cognitive Sciences* **0.** Publisher: Elsevier. ISSN: 1364-6613, 1879-307X. https://www.cell.com/trends/cognitive-sciences/abstract/S1364-6613(20)30175-3 (2020) (Aug. 24, 2020).
58. Takahashi, N., Ebner, C., Sigl-Glöckner, J., Moberg, S., Nierwetberg, S. & Larkum, M. E. Active dendritic currents gate descending cortical outputs in perception. *Nature Neuroscience* **23.** Number: 10 Publisher: Nature Publishing Group, 1277–1285. ISSN: 1546-1726. https://www.nature.com/articles/s41593-020-0677-8 (2023) (Oct. 2020).