Neuron

Article

Recurrent cortical networks encode natural sensory statistics via sequence filtering

Highlights

- Local recurrent circuits in the visual cortex actively filter sequential inputs
- The strength of an incoming input response depends on prior context
- Natural sequences are selectively boosted, while others are suppressed
- This is a predictive function for local recurrent circuits

Authors

Ciana E. Deveau, Zhishang Zhou, Paul K. LaFosse, Yanting Deng, Saghar Mirbagheri, Nicholas Steinmetz, Mark H. Histed

Correspondence

zhishang.zhou@nih.gov (Z.Z.), mark.histed@nih.gov (M.H.H.)

In brief

The visual cortex receives a stream of high-dimensional sensory input. The role of dense local, recurrent cortical connections in shaping responses to these inputs has been unclear. Here, we show that sequence of inputs interact across time via recurrent interactions, so that natural sequences are boosted compared to irrelevant signals.

Deveau et al., 2026, Neuron *114*, 1489–1503
April 15, 2026 Published by Elsevier Inc.
https://doi.org/10.1016/j.neuron.2025.12.024

Neuron

CellPress

Article

Recurrent cortical networks encode natural sensory statistics via sequence filtering

Ciana E. Deveau,[1,2,3,6] Zhishang Zhou,[1,6,*] Paul K. LaFosse,[1,2,4] Yanting Deng,[1] Saghar Mirbagheri,[5] Nicholas Steinmetz,[5] and Mark H. Histed[1,7,*]

[1]Intramural Program, National Institute of Mental Health, National Institutes of Health, Bethesda, MD, USA
[2]NIH Graduate Partnership Program, Bethesda, MD, USA
[3]Department of Neuroscience, Brown University, Providence, RI, USA
[4]Neuroscience and Cognitive Science Program, University of Maryland, College Park, MD, USA
[5]Department of Biological Structure, University of Washington, Seattle, WA, USA
[6]These authors contributed equally
[7]Lead contact
*Correspondence: zhishang.zhou@nih.gov (Z.Z.), mark.histed@nih.gov (M.H.H.)
https://doi.org/10.1016/j.neuron.2025.12.024

SUMMARY

Recurrent neural networks can generate dynamics, but in the sensory cortex, it has been unclear if any dynamic processing is supported by the dense recurrent excitatory-excitatory network. Here, we show a role for recurrent connections in the mouse visual cortex: they support powerful dynamical computations, but by filtering sequences of input instead of generating sequences. Using two-photon optogenetics, we measure neural responses to natural images and play them back, finding that responses are boosted when inputs are played back during the correct movie dynamic context—when the preceding sequence corresponds to natural vision. This sequence selectivity depends on a network mechanism: earlier input patterns produce responses in other local neurons, which interact with later input patterns. We confirm this mechanism by designing sequences of inputs that are boosted or attenuated by the network. These data suggest that recurrent cortical connections perform predictive processing, encoding the statistics of the natural world in input-output transformations.

INTRODUCTION

A defining feature of the cerebral cortex is its extensive local recurrent connectivity. Excitatory cortical cells make up approximately 80% of the neurons in all cortical areas, and each excitatory cell receives hundreds or thousands of inputs.[1] A majority of those input synapses come from other local neurons, within a distance of a few hundred microns.[2–4]

However, in many cortical areas, the function of this extensive local recurrent network is not well understood. For example, the sensory functions of the primary visual cortex (V1) have often been explained by feedforward or feedback inputs alone, with little need for recurrent computations. Foundational work on V1 from Hubel and Wiesel[5,6] showed that feedforward summation of visual thalamic inputs can create the edge detection properties (i.e., orientation and direction tuning) that are a defining characteristic of early visual cortex. Later work using inactivation and intracellular recording[7–9] confirmed this, finding that even though recurrent inputs to V1 neurons were large, recurrent input did not substantially refine visual tuning, for example, not sharpening orientation responses.[10]

However, other work did find some contribution to visual function made by recurrent connections. Connectivity studies that combined *in vitro* and *in vivo* measurements showed neurons with similar orientation tuning are preferentially recurrently connected.[11–13] This observation suggested that the recurrent network could perform preferential amplification of feedforward inputs driving orientation tuning.[3,14] That prediction was borne out using two-photon holographic stimulation, which confirmed that cells with similar orientation sensitivity produced amplified responses when stimulated simultaneously.[15,16]

In sum, many years of study of V1 responses have found a primary role of the area is to extract edges from the visual field[17,18] and also found that the recurrent network of V1 facilitates that function. Population-level recording studies confirm this idea, showing that much variation in V1 neurons' responses is well-described by receptive field estimates derived from simple edges.[19,20]

But this framework is not the whole story. V1 neuron responses during natural vision show significant variation that is not well explained by mere edge detection. That is, V1 response estimates derived from the tuning of the classical receptive field "center" do not come close to predicting all the variance of V1 neurons during natural vision. As Olshausen et al. put it,[21] "what is the other 85% of V1 doing [during natural vision]?" Up to 85% of the variability of V1 responses has not been explained

by orientation or direction-based receptive field models, both when that question was posed and at the present time.

One possibility is that many of the responses of V1 neurons are explained by non-visual factors. Over the past 15 years, a large number of studies have focused on non-visual responses in V1—responses influenced by the self-motion of the organism (e.g., running,[22–24] top-down inputs,[25,26] prediction,[27] or cross-modal responses to non-visual sensory input, such as tactile sensation or auditory stimuli).[28–31]

On the other hand, it may be that the remaining unexplained V1 function is due to visual properties that have not yet been uncovered. There are hints that unknown visual responses drive this missing variance. The first is that V1 responses to natural movies are stronger and sparser—statistically quite different—than predicted from models that use V1 receptive field (edge) properties alone,[32–34] implying some lawful rules of visual processing that have not yet been identified. Also suggesting there may be some V1 role in complex natural vision, neurons co-activated in natural vision are also more likely to be connected.[35] Large-scale recording methods show the responses of V1 neural populations to be high-dimensional, as if they describe a very large space of visual responses extending beyond classical oriented-edge receptive fields.[36] In fact, studies of V1 receptive fields generally observe a "non-classical surround"—a large region outside the classical receptive field that affects V1 neurons' responses, with hints of complexity[37] that have not been so far fully deciphered (reviewed in Olshausen and Field[38]).

One thing that could explain the missing variance in visual cortical responses is information about time. During natural vision, the visual input the organism receives changes as the visual world changes, both due to motion of objects and due to the organism's movements. That means that the natural dynamic stimuli that produce strong and sparse V1 responses contain complex motion of various sorts.

A major challenge to the idea that missing V1 variance is explained by temporal information is many past studies have found only simple modulation of neurons' activity with time. Responses to a natural image or orientation sequence are predicted by linear combinations of responses to single component stimuli,[19,39,40] and V1 neurons often do not have complex spatiotemporal receptive fields (STRFs) as have been seen in audition.[41–43] Receptive field models that explain V1 responses as well as currently possible are often—but not always[34,44,45]—spatiotemporally separable, so that the temporal kernel and spatial kernel of a V1 receptive field can be simply multiplied together to give its spatiotemporal response.[34,44,46]

We thus are left with a question about how the V1 recurrent network transforms spatiotemporally dynamic visual cortical responses. Much of what we can explain of V1 function can be explained by feedforward input and edge processing, but this does not explain V1 population responses to natural vision. What explains this missing variance in V1, and is that unknown non-classical function created by processing in the recurrent network?

Here, we show that V1 responses are modulated by temporal context. The recurrent network encodes features of natural vision, linking responses of neurons across time. We use two-photon holographic stimulation to simulate patterns of input derived from natural movies. We find the gain, or magnitude, of the response to a given feedforward input is affected by other neurons' activity in the recurrent network, showing that V1 responses to a visual stimulus vary in a predictable way based on temporal context. When a given patterned input is provided in the correct temporal context, preceded by the visual input that occurs in natural vision, the response to that patterned input is larger than when the same input pattern is delivered in a non-natural context. Our data may explain why past studies have found a large non-classical surround effect but have not been previously able to find lawful rules of how simple stimuli predict non-classical surround responses: the non-classical surround is too complex to fully explain by exploring the space of visual stimuli using recording methods. Here, by activating parts of known, identified natural stimuli using patterned stimulation, we show stimuli matching natural statistics specifically boost V1 responses, as populations of neurons in V1 responding to stimulus subparts influence each other through recurrent mechanisms.

Our results support the idea that the V1 recurrent network is performing complex computations over time, but not by generating sequential responses. Instead, the V1 network filters sequences of input. Natural vision produces a changing set of patterned input to the cortex, and either via experience or development, the cortical recurrent network has learned and boosts input sequences that correspond to natural vision. We confirm this in an artificial recurrent network, showing that a network selecting natural sequences of input is consistent with our data. Then, returning to experiments, we use two-photon stimulation to identify a recurrent mechanism for this sequence filtering, construct novel sequences of input patterns following this mechanism, and show that sequences designed to be boosted do indeed produce larger responses. This confirms a recurrent origin for the sequence filtering that can underlie natural visual processing in V1.

Together, our results show that V1 neural responses are indeed sensitive to temporal context, but in a way that is not easily described by single-cell response properties or receptive fields. Instead, populations of neurons affect other populations via recurrent influences. Because we find that neurons' gain is affected by the activity of other local neurons, the information for this boosting, or sequence amplification, process can be encoded in the large number of excitatory-excitatory synaptic weights in the V1 local network. These results provide one possible solution for what "the other 85%" of V1 is doing: beyond just edge extraction, the V1 recurrent network is sensitive to the complex temporal structure of the natural world.

RESULTS

Two-photon holographic stimulation (Figure 1E)[47,48] allows us to deliver direct inputs to populations of cells in the cortex and directly study the effects of local neural activation on other local neurons. We use a viral expression strategy[49] that co-expresses an opsin and a calcium indicator in excitatory neurons in layer 2/3 of mouse V1, stably over days to weeks, with minimal optical crosstalk between imaging laser and opsin (Figures 1A–1C; as we previously characterized[49]). This approach allows us to directly access the dense local recurrent circuitry between excitatory cells (Figure 1D). With two lasers, we simultaneously

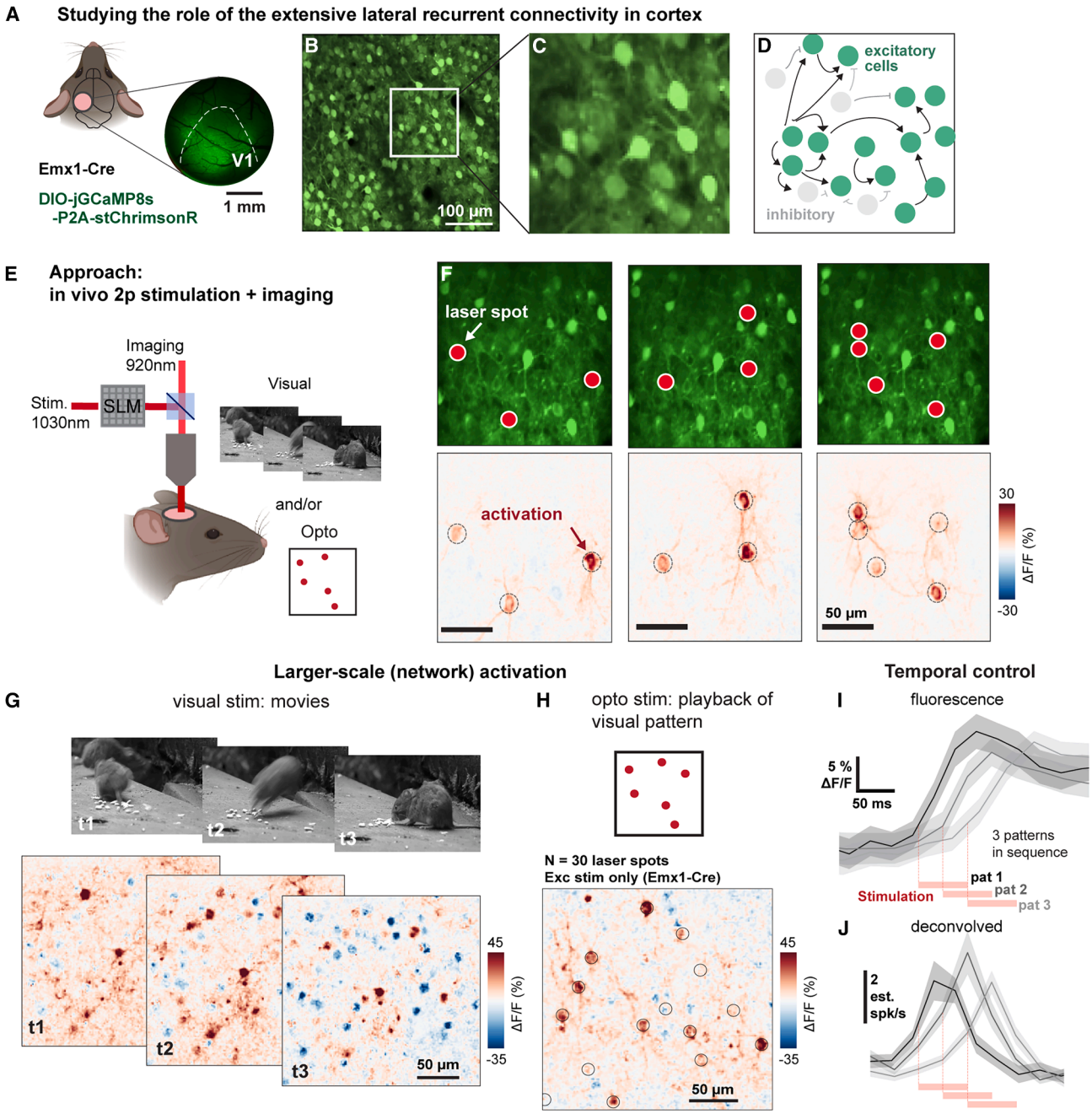

Figure 1. Two-photon holographic stimulation and imaging to probe the role of excitatory recurrent connections

(A) Window implant.

(B) Field of view (FOV): excitatory cells expressing GCaMP8s and stChrimsonR.

(C) Enlarged region of (B).

(D) Schematic of the ubiquitous recurrent connections between excitatory cortical neurons.

(E) Experiment schematic. Paired visual and patterned optogenetic stimuli. SLM: spatial light modulator.

(F) Example of spatial precision of 2p stimulation. Left to right: patterns with 3, 3, and 5 laser spots, each stimulated simultaneously.

(G) Network activation from natural movie input at three timepoints.

(H) Network activation, with non-target changes in non-stimulated neurons, also results from optogenetic patterns (black circles: laser spots, 30 total, a subset of the image shown here for visual clarity).

(I and J) Example of temporal precision of sequential 2p stimulation: 3 patterns, 15 cells each, 60 ms stim per pattern with 30 ms between patterns. (I) Average fluorescence traces across cells and trials, reflecting calcium responses of stimulated cells, and (J) corresponding deconvolved traces (STAR Methods). Even with some overlap, imaging can resolve distinct temporal peaks.

Error bars: SEM.

stimulate groups of neurons with single-cell resolution and measure responses to the optogenetic stimulation and/or visual stimuli (Figure 1E; details in STAR Methods). The optogenetic inputs are precise spatially (Figure 1F) and temporally, with one input pattern changing to the next within tens of milliseconds (Figures 1I and 1J; Figures S1B and S1C).

Two-photon optogenetics can mirror features of natural visual input by driving population activity in the cortical network. When animals are shown natural visual movies, many neurons in V1 respond (Figure 1G). Using two-photon stimulation, we can simulate these activity patterns by giving input to many neurons at once, activating the network through excitatory cell stimulation (Figure 1H; in this work, we target 15–30 cells,[50]). This all-optical approach allows us to measure sensory responses, select neurons that respond, and activate the excitatory patterns.

Sequential order affects V1 neurons' responses through non-target neural suppression

If recurrent effects influence V1 responses to temporally varying input, then changing the order of patterns of input might be expected to produce differences in neurons' responses. To determine if this is true, we chose patterns of neurons at random (three patterns: A, B, and C, each a fixed duration, 30 or 60 ms in different experiments, spanning a range of the temporal frequencies often seen in vision: e.g., frame times used in videos,[51] flicker fusion frequencies,[52,53] and photoreceptor dynamics.[54] We presented these patterns in two different orders (ABC or CBA; Figure 2A).

We find the sequential order of these population input patterns does indeed affect neurons' responses (Figures 2B–2E). The first pattern produces a greater response than the same pattern when it is delivered later in the sequence (Figures 2B–2E). This sequential context dependence occurs in the cells stimulated as part of the A, B, or C patterns. On average, the cells stimulated first in the sequence show stronger responses (Figures 2C–2E; the median of the A and C distributions differ in Figure 2E). Large effects are seen in individual neurons, which are strongly modulated by sequential position (e.g., there is significant variance in the distributions in Figure 2E: 37% of A neurons and 45% of C neurons are significantly different from zero, Kolmogorov-Smirnov (K-S) test $p < 0.05$).

To understand whether prior patterns of input to other cells could change the response of neurons to a particular input, we analyzed responses in the B cells (Figure 2F). The B input pattern is preceded by one pattern in each sequential order: the A pattern in the ABC sequence and the C pattern in the reverse sequence. Thus, if there is a specific effect of sequential context on individual neurons, the firing of some B cells to the B input should be modulated also by the pattern that precedes it. Indeed, we found that many of these cells were significantly modulated by sequential context (21% of B cells $p < 0.05$, Mann-Whitney U test; Figures 2G–2L). Even though the B input was always the same, some B cell responses were larger in the ABC sequence (Figures 2G–2I) and some were larger in the CBA sequence, as the pattern that preceded the B pattern was changed (Figures 2J–2L; variance differences via Anderson-Darling (A-D) test; Figures 2N and 2O). This variation in response was not dependent on small differences in distance between each B cell and the previously stimulated A or C cell (Figure S2), implying these effects are not due to laser crosstalk between patterns. Further, stimulation of each pattern individually yielded a similar distribution of excitation and suppression in the non-stimulated neurons (Figures 2P and 2Q). Thus, these effects are not driven by changes in average responses caused by different patterns but are due to changes in individual cell responses based on the specific preceding pattern.

In sum, stimulation with sequences of random population input patterns shows that V1 neurons' responses to fixed inputs are highly dependent on sequential context. Responses to an input pattern are affected by prior patterns of input to different neurons. This suggests that the visual cortex recurrent network may perform a time-based computation for sensory stimuli. To understand this, we next investigated how sequential modulation relates to natural vision.

Patterns corresponding to natural visual inputs produce larger responses

To study sequential modulation in the visual context, we turned to dynamic natural visual inputs. One of the most common kinds of sequential input that the visual cortex receives comes in response to natural motion. Because of the size and scatter of cortical receptive fields,[55,56] and because axons carrying inputs contact many cells, the cortical network receives a changing set of patterned inputs during dynamic vision that are scattered nearly randomly across areas the size of our fields of view.

In principle, the cortical network could store, in its recurrent connections, information about these patterns and thus the temporal structure of the natural visual world. This could allow the network to preferentially respond to sequences of input corresponding to natural vision. In that case, naturally occurring sequences would be expected to produce different responses in the network than sequences that do not correspond to natural vision. This could occur by earlier parts of a dynamic natural stimulus influencing responses to input arriving at later times, as we observed with random input.

We examined this by using two-photon stimulation to mimic responses from one part of a natural movie while we changed the prior visual context. We first measured responses to a single frame taken from a movie (Figure 3A; Figure S3D). We constructed a two-photon input pattern from the cells activated by that single frame. We then showed animals the movie and played back the response to the frame with two-photon stimulation, replacing the visual frame with the optogenetic input pattern either at the correct time in the same movie (Figures 3A and 3B: matched context) or in a different movie such that the preceding movie frames were not matched to the stimulation pattern (Figures 3A and 3E: unmatched context).

If V1 responses were dependent on the dynamic context of natural vision, neurons should produce different responses to the same, fixed input pattern when it is presented in different contexts. We subtracted the respective visual response from the visual plus optogenetic response to give the two-photon incremental response in each context (Figures 3B–3H). This reveals a context-dependent effect, where the input pattern produced a larger response at the appropriate time in the visual

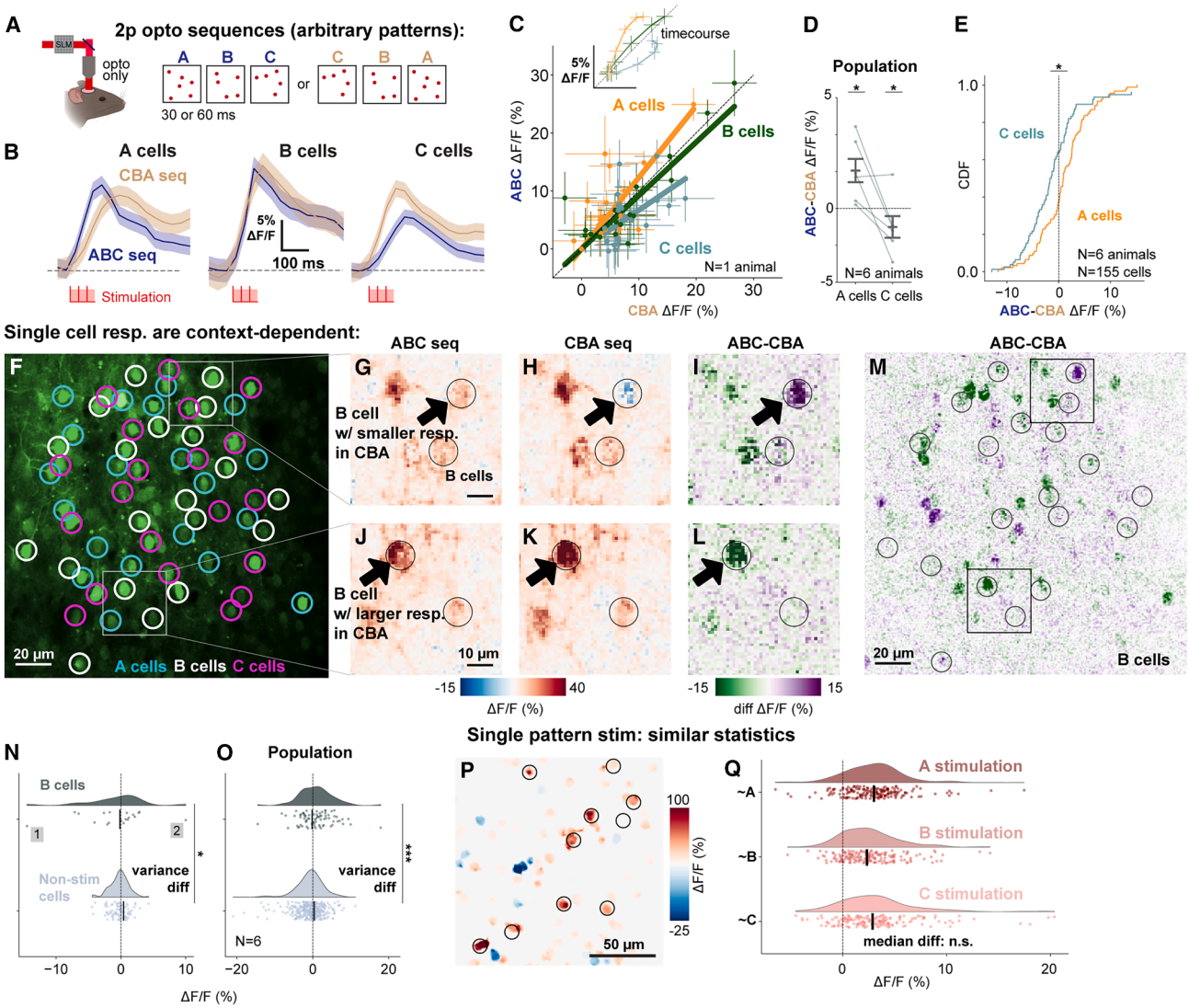

Figure 2. Patterns of input produce different responses in the V1 network depending on the sequential context

(A) Experiment schematic: stimulation of three patterns (N = 30 cells per pattern, randomly selected; each pattern 30 ms). Two sequences: forward (ABC) and reverse (CBA).

(B) Average responses (across cells) of A (left), B (middle), and C (right) cells (N = 1 expt; N = 30 cells per panel).

(C) Points: neurons, average response across trials, and same data as in (B). Inset: time course of average response across cells for each pattern (begins 5 frames prior to stimulation, and each point is one imaging frame).

(D) Population (N = 6 expts). Gray: ABC-CBA responses of A (left) and C (right) cells, averaged over cells. t test, A cells (N = 79): $p < 0.0001$, C cells (N = 76): $p = 0.001$.

(E) ABC-CBA responses for A cells (N = 79) and C cells (N = 76) (Kolmogorov-Smirnov [K-S] test, $p < 0.001$).

(F) Green: bicistronic GCaMP8s and stChrimsonR expression.

(G–L) B cell (black circles) responses to stimulation change based on prior stim. (G–I) Example B cell with stronger activation in the ABC sequence. (J–L) Example stronger in CBA.

(M) Larger view of G-I. Image matches FOV in (F). Black squares: regions in (G–I).

(N) Variance in cell responses. x axis: B cell response difference between ABC and CBA patterns, non-stimulated neurons (bottom: gray, cells not stimulated in any A, B, or C pattern), and larger variability for B cells (Anderson-Darling [A-D] test for variance diff, N = 19 B cells, N = 95 non-stim cells, expt shown in F–M, two-tailed $p < 0.01$).

(O) Same as N, across N = 6 expts (A-D test, N = 81 B cells, N = 339 non-stim cells, $p < 0.001$).

(P and Q) Modulation is not due to response differences to single pattern stim: A, B, and C patterns have similar statistics (data from B–E; Mood's median test, N = 3 stim, $p = 0.15$).

Error bars: SEM.

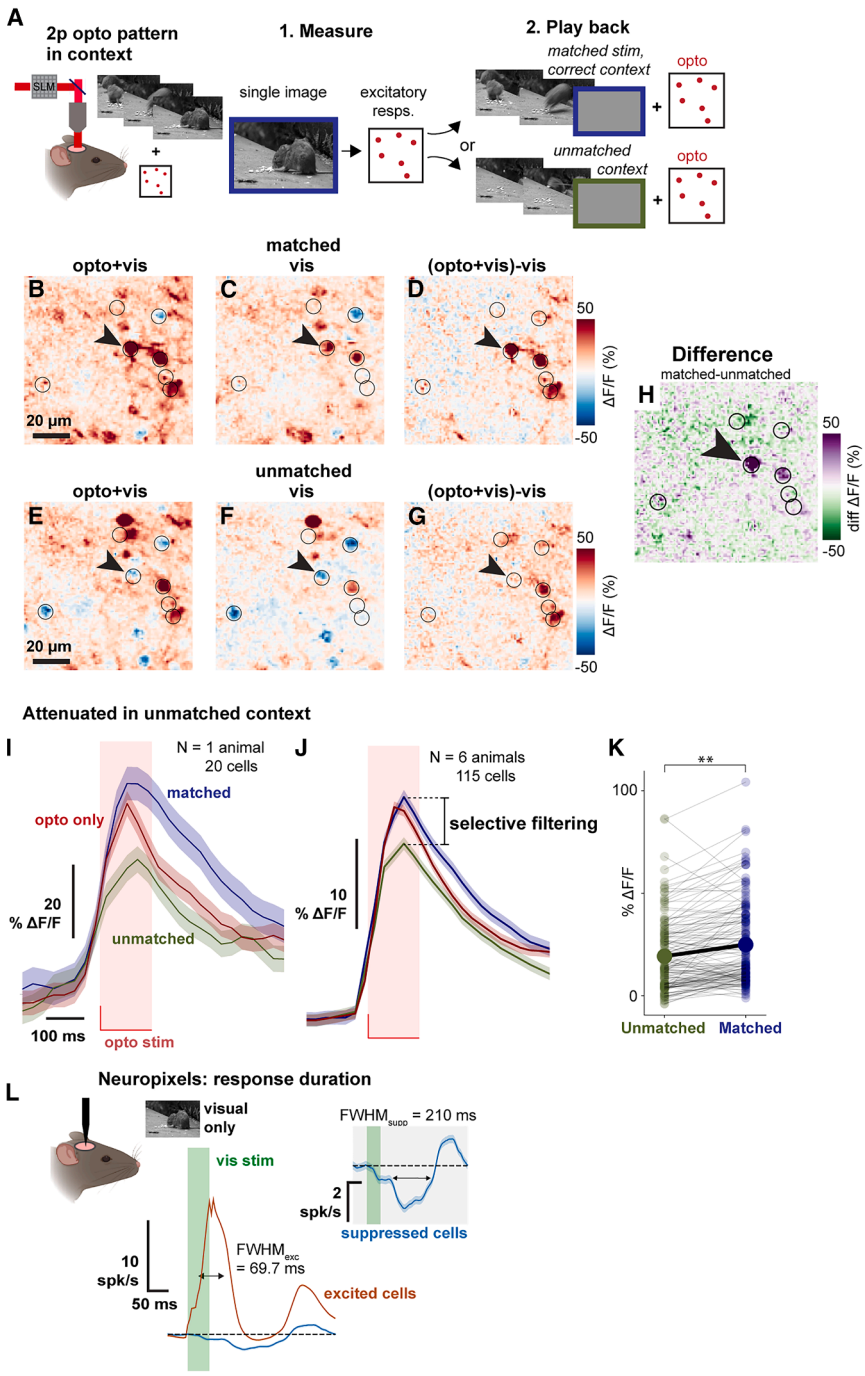

Figure 3. Sequence amplification for visual input, observed by simulating responses to a single frame of a movie

(A) Experiment schematic: responses to a single frame played back using stimulation at the correct time in the same movie (matched context) or in a different movie (unmatched). Three temporal locations were tested in two movies (STAR Methods).

(B–G) Example responses to combined opto and visual, visual alone, and isolated opto. Matched (B–D) and unmatched (E–G) contexts. (D and G) Responses in either context: subtracting vis from opto + vis combined. Black circles: optogenetic targets.

(H) Example larger matched response (arrow), difference of (D) and (G).

(I) Mean responses (across cells) in matched (blue), unmatched (green), and opto-only (red) trials ($N = 1$ expt, pattern duration: 120 ms).

(J) Population time courses, $N = 6$ animals.

(K) Points: cell responses in matched (blue) and unmatched (green) contexts from (J) (Mann-Whitney U test, $N = 115$ cells, $p < 0.01$).

(L and M) Electrophysiological responses to a flashed natural image ($N = 5$ animals, $N = 11$ expts).

(L) Suppressed cells (N = 3,569 cells, blue). (M) Excited cells (N = 14,528 cells, orange; error bar, SEM, hidden by mean). FWHM: full width at half max.

Error bars: SEM.

movie (matched context) than when presented in a different visual context (the unmatched context; Figures 3D and 3G–3K).

The average response in the matched case is similar to the response in the opto-only case (i.e., blue and red lines are similar in Figures 3I–3K; see also Figure S3A). It is not larger than the opto-only case, as might be predicted by simple amplification with no overall, population-average changes based on context. However, it could be that some patterns (in the matched case) are indeed selectively amplified, but the overall network response is smaller during complex natural vision than when animals view simple unchanging stimuli (gray screen: opto only). That is, the cortical state could change during natural vision, scaling both matched and unmatched responses down so that the matched response is similar to the opto-only response. Thus, the key observation here is not similarity to the opto-only response but the differential response in the matched and non-matched contexts.

An immediate question is about mechanism: how might responses to one pattern of input influence responses to later input? For a recurrent network interaction to modify responses to later inputs, it would seem that responses from one frame should be sustained for some time. Then, one input pattern could produce non-target responses in other local neurons, and a later input could fall on these non-target neurons while their responses were changed by the previous input, allowing the two inputs to interact with each other. To determine if this kind of sustained response was present, we turned to Neuropixels electrophysiological recordings, due to their fidelity in measuring the times of spikes. We recorded V1 neuron responses to flashed single image frames (as in Figure 3A) and measured the duration of their non-target responses. We sorted neurons into groups based on whether they showed activated or suppressed responses and found that both activated and suppressed responses were

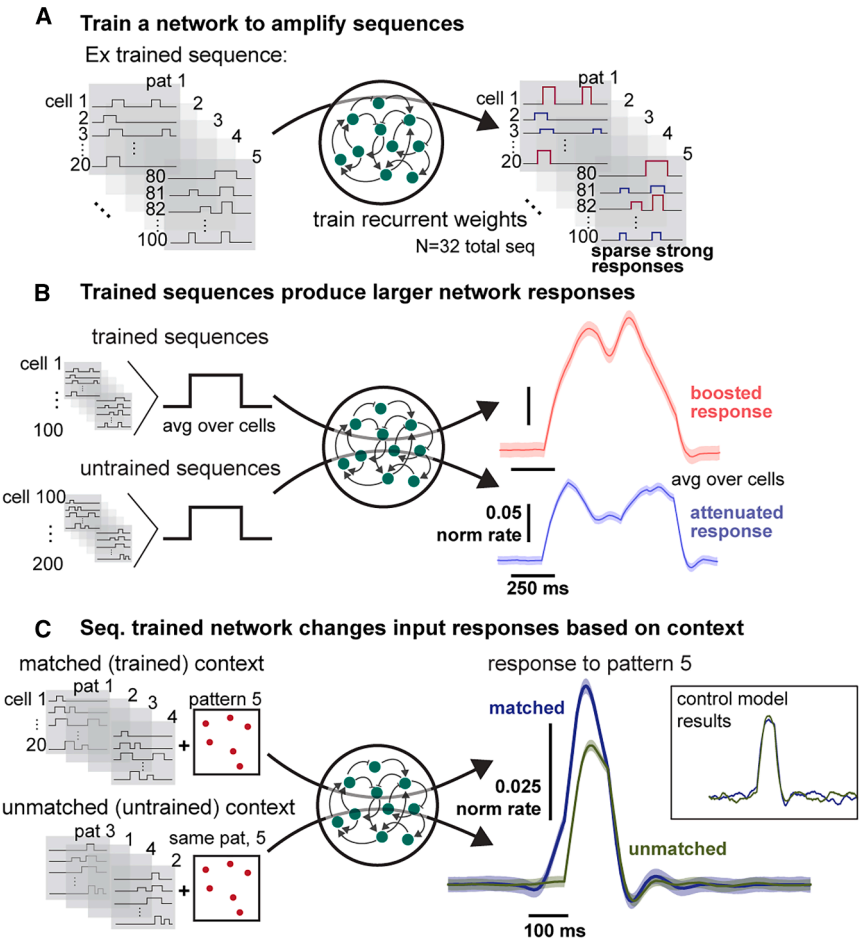

Figure 4. RNN trained to amplify patterns produces context-dependent modulation

(A and B) (A) Network is trained to amplify sequences of input (pattern duration: 120 ms), resulting in (B) boosted output time courses for trained input sequences.

(C) Simulating the experiment of (Figures 3A–3K) recapitulates the data: larger pattern responses in the correct (blue) context. Inset: control model: no context-dependent responses.

Error bars: SEM.

sustained after visual stimulus offset, with suppressed responses lasting for several hundred milliseconds (Figure 3L). This is sufficient for one frame of a natural visual movie to interact with the next. For comparison, a movie at 25 frames per second, a rate comfortably within the range where humans see smooth motion, has a 40 ms frame time. The mouse visual system likely supports even faster processing than the human, and yet here, we found that natural images flashed for short times produced responses 70 ms long for activated neurons and 210 ms for suppressed cells (Figure 3L). These data suggest that the dynamics of input during natural vision are sufficient to allow overlap between patterns, allowing recurrent influences from earlier input patterns to affect responses to later inputs.

An RNN model that boosts specific sequences produces the context-dependent responses we see in our data

The observation that responses to single input patterns are modulated by context—by earlier patterns or frames (Figures 3A–3K)—implies that longer sequences of input are also selectively filtered, as each movie frame response interacts with later movie frames. That is, the pairwise sequential interactions we found, where single frames of input are treated differently based on prior input, can also enable processing of longer sequences of dynamic input where longer natural sequences of input are selectively filtered. Biological motion, like observing a predator moving to attack, can be ongoing for some time and generate changing sequences of input patterns to the cortex through that time.

To determine whether boosting of even longer sequences can result from the temporal interactions we studied, we examined a recurrent neural network (RNN) trained to preferentially generate stronger responses for some seconds-long input sequences and not others (Figures 4A–4C). The target output was a stronger, sparser scaled version of the input (in the target, for 35% of units, output was 2× input, and for the remaining, output was 0.9× the input), to mimic natural movie response characteristics (Figure 4A). The model preferentially boosted the trained sequences and attenuated scrambled versions of those sequences (Figure 4B). We asked if the context-dependent effects seen in our Figure 3 experiment could explain this selective modulation. To test this, we chose a single input pattern as in the experiment, reflecting responses to a single frame of a sequence. We measured the response amplitude of that pattern when in the matched context versus the unmatched context. The two states are driven by presenting either the first 4 patterns of the trained sequence (matched) or of the untrained sequence (unmatched). These prior inputs in the model may be seen as simulating the L4 or lateral geniculate nucleus (LGN) inputs to L2/3 V1, with the final input pattern simulating the optogenetic stimulus. Indeed, with an RNN trained to boost a sequence of inputs, a single pattern extracted from the sequence and played back in the correct context produced a larger response compared with when it was presented in the incorrect context (Figure 4C). By contrast, a control network without amplification does not exhibit this context-dependent effect. The input pattern produces the same response magnitude in both matched and unmatched contexts (Figure 4C inset; Figure S4).

Thus, an RNN trained to boost a sequence of input mirrors the context-dependent stimulation effect seen in our experiments (Figures 3J and 4C). Together with the finding that neurons' responses to input are dependent on earlier inputs (Figure 2), these observations support the idea that the local recurrent

network is filtering extended sequences of input, specifically boosting input sequences corresponding to natural vision.

Recurrent network mechanism: One input pattern creates responses in other, non-targeted local neurons, which interact with later input patterns

If sequence modulation was created by local, recurrent interactions, we should be able to see signs of these interactions in the network activity. To look for specific influences—to see how particular input patterns filter later patterns—we separated stimulation patterns in time, inserting delays of several seconds (8 s, STAR Methods) between one or more patterns in a sequence (Figure 5A). We first inserted a delay between patterns B and C in an ABC sequence (A, B, and C patterns randomly chosen; STAR Methods).

The reduced sequence produces striking non-target responses (Figures 5A–5C) in cells not receiving stimulation, with some non-stimulated cells activated and others suppressed. These non-target neural responses are likely to be primarily due to local recurrent interactions, not due to axonal or dendrite activation, as here we used a somatically targeted opsin (stChrimsonR).[49]

These non-target responses induced by the AB sequence can occur in neurons that will later be stimulated in the C pattern. This could be a mechanism for sequential modulation: that is, a single input pattern produces non-target responses in other local neurons, and these responses interact with later patterns of input to modify responses to those later patterns.

To see how earlier patterns impacted later patterns, we examined how C cells changed their responses when preceded by the AB pattern and whether this change was predictable from the AB pattern responses. In prior work, we have found attenuation-by-suppression,[57] where V1 neurons, when suppressed, produce smaller (sublinear) responses to input due to the shape of neurons' activation functions. Such a nonlinearity could be a mechanism that creates preferential suppression of some patterns and sequences. To understand if sequence selectivity could depend on such nonlinearities, we examined non-target responses in C cells generated by the AB sequence. We compared the ABC response to a prediction computed by summing the C response and the AB response. We found that excited non-target cells did produce linear responses when they were stimulated next in sequence, with ABC responses well-predicted by the sum of non-target responses to AB and C (Figures 5D and 5F; Figure S5A). And consistent with prior findings on activation-by-suppression, suppressed non-target cells showed nonlinear responses (Figures 5E and 5F; Figure S5A).

These responses suggest that the sequential modulation we observed does involve some network mechanisms. Non-target responses, induced by earlier patterns of input through the local recurrent network, generate a non-target response pattern that interacts with later patterns.

Designed input sequences produce predicted effects, confirming the network mechanism

To provide an explicit test of this network mechanism, we constructed sequences that should be treated differently according to the mechanism described above.

To construct such sequences, we first measured responses to a single input pattern (Figure 5G). We then created later stimulation patterns based on the non-target responses to those single patterns. We constructed sequences that should be "boosted" by adding a pattern of stimulated cells that were excited by the first pattern and constructed "attenuated" sequences by adding a pattern of stimulated cells suppressed by the first pattern (Figures S5G–S5J). Any cell within a 10 μm radius from a previously stimulated cell was excluded from selection to avoid pattern overlap. We repeated this step a second time to create boosted and attenuated sequences that were three patterns long.

If our constructed sequences, when delivered to the network, produce the expected excitation or suppression, this would be strong evidence that this recurrent mechanism underlies sequence modulation.

We studied this by playing back these sequences onto the cortical population. Confirming our hypothesis, we found that the sequence designed to be boosted produced a larger response than the sequence designed to be attenuated (Figures 5H–5J). This interaction between non-target responses induced by one pattern affecting later patterns is therefore a mechanism where some sequences can be enhanced and others suppressed.

However, one additional nonlinear phenomenon was present (Figures 5A–5E). Patterns designed to be boosted produced responses that were sparser than suppressed sequences—a number of cells produced large responses that elevated the mean (Figures 5I and 5K). This is a kind of nonlinear amplification effect produced by later patterns falling on non-target cells that were excited by earlier patterns. This sparsity is a hallmark of responses to natural movies[32,33,58] and supports the idea that the cortical network can selectively boost sequences of responses associated with natural vision. Thus, we observe at least two kinds of nonlinearities that can contribute to sequence filtering, boosting some sequences and suppressing others: first, the attenuation-by-suppression effect that suppresses patterns falling on earlier non-target suppressed cells, and second, the increased strength and sparsity effect that can boost patterns that fall on earlier non-target excited cells.

In sum, experiments separating sequences in time (Figures 5A–5F) and constructing sequences designed to be excited or attenuated (Figures 5G–5K) together support a recurrent network mechanism for sequence tuning. Responses to an earlier pattern in a sequence create non-target responses in other local neurons that influence responses to later patterns. This modulation results in particular patterns, such as natural patterns (Figure 3), producing larger responses compared with other responses that are attenuated (Figure 6A).

DISCUSSION

Here, we find that individual V1 neurons' responses are modulated by prior inputs arriving to other V1 neurons (Figure 2). When we vary those inputs by changing prior visual stimuli, we find that earlier inputs boost responses to later inputs when the input sequence is derived from sequences found in natural vision (Figure 3). We show that a flashed natural scene produces a

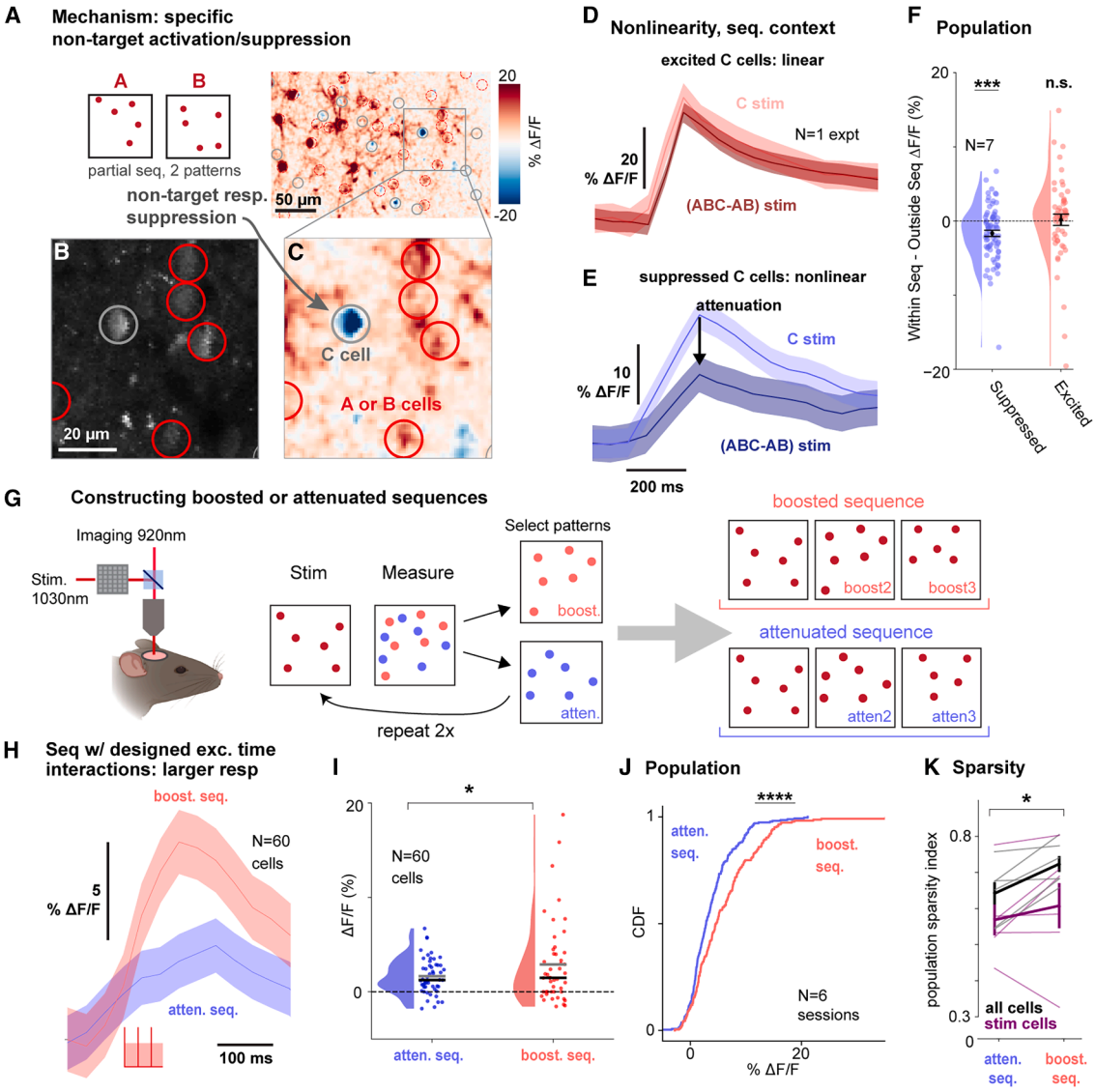

Figure 5. Sequences designed to be boosted produce strong and sparse responses, validating a recurrent network mechanism

(A–C) Non-target responses induced by earlier patterns. Red: stimulated cells (from A or B pattern). Gray circle: C cells, here unstimulated. 60 ms duration. (A) Responses to AB sequence. Left, schematic; right, GCaMP responses. (B) Anatomical FOV. (C) Zoomed view. Gray circle: unstimulated neuron.

(D) Excited cells from one example experiment ($N = 5$ cells). Light red: response to C stimulus alone. Dark red: response to ABC sequence minus AB response.

(E) Suppressed cells ($N = 5$ cells; same expt). Same conventions as (D). The C response within the ABC sequence (dark blue) is smaller than C alone (light blue; t test, 300 ms period after stim onset, $p = 0.013$).

(F) Population data ($N = 7$ expts; $N = 130$ cells, C cells only; compares ABC stim to AB + C stim as described in D and E). Blue: suppressed cells (response to AB stim < 0; t test versus linear prediction, $N = 76$ cells, $p < 0.001$). Red: excited cells (response to AB stim > 0; t test, $N = 54$ cells, $p = 0.84$). Distributions are significantly different (K-S test, $p < 0.01$).

(G) Experimental schematic: designing boosted or attenuated sequences based on non-target responses.

(H) Mean responses of boosted (pink) and attenuated (blue) sequences ($N = 20$ cells per pattern, 3 patterns in the sequence; 30 ms pattern duration).

(I) Stimulated cells' response distribution. Mean, gray horizontal line; median, black; means are significantly different (t test, $p = 0.033$; $N = 60$ cells from $N = 3$ patterns).

(J) Population data (K-S test, $p < 0.00001$, $N = 277$ exc cells, $N = 270$ Supp cells, $N = 6$ expts).

(K) Sparsity is greater for boosted patterns. Each point: average across neurons (t test, $p < 0.05$; $N = 6$ expts; all cells, gray).

Error bars: SEM.

response that lasts more than 100 ms (Figure 3L), meaning the integration window for the interaction of one pattern with a future pattern is at least this long. We then confirm that patterns interact across time via recurrent influences. To do this, we design sequences of input patterns that should be boosted or attenuated if non-target responses to prior patterns interact with later patterns. We find that the resulting sequences are indeed selectively boosted or attenuated. One intriguing result is that the boosted

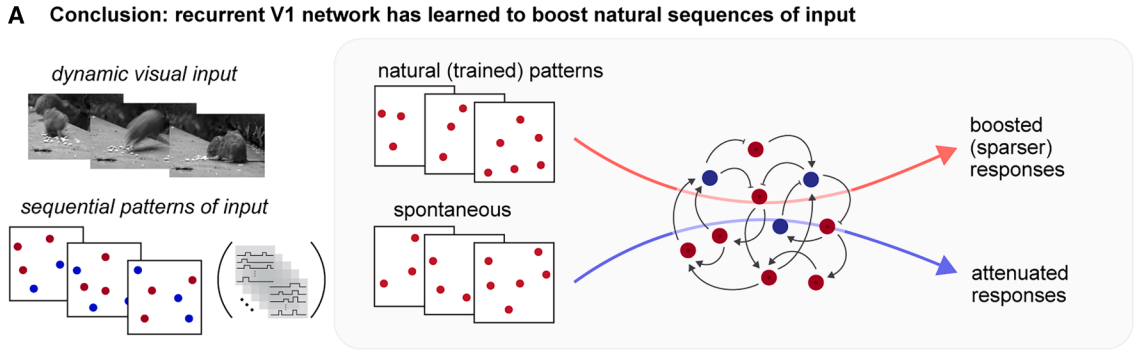

Figure 6. Active filtering of sequences by the cortical recurrent network

(A) Mechanism and conceptual model from this work: the recurrent network filters sequential inputs, boosting some and attenuating others.

sequences produce sparser responses than the attenuated sequences. This is consistent with the sparse responses to natural visual input observed by Gallant et al.,[32] supporting the idea that recurrent interactions do shape responses to natural visual input.

Our results show a computational role for the dense recurrent network in layer 2/3 of V1. In several cortical regions, this dense recurrent connectivity is thought to generate ongoing or slow dynamics: for example, in the motor cortex, recurrent dynamics are thought to be associated with muscle movement, and in the prefrontal cortex, delay dynamics are associated with short-term memory.[59,60] By contrast, V1 does not generate complex sustained or ongoing dynamics. Here, we find the V1 network does support complex temporal processing, not by generating sequences but by filtering sequences of input, boosting natural sequences of patterns while attenuating others (Figure 6A).

One feature of our results is that matched responses and opto-only responses are similar in average magnitude, while unmatched responses are smaller than both (Figure 3). One might have expected the cortex to explicitly amplify the response in the correct natural or sequence context, and instead, it appears that the cortical network may be selectively suppressing non-matched input patterns. But this similarity to the opto-only responses could still be a signature of selective amplification, combined with a change in cortical state during natural vision. The cortical network may produce responses that are on average smaller during complex visual input than during simple or static vision, so that both the matched and non-matched responses are reduced relative to the opto-only case.

This would mean some sequences are indeed selectively boosted or amplified by the network based on preceding patterns, while natural vision also results in an overall, average, or blanket suppression of incremental neural responses (gain reduction) that combines with that selective amplification. Supporting the idea that natural vision may lead to a blanket gain reduction compared with spontaneous activity states with simple or static vision, some data and models suggest that incremental responses decrease as cortical networks receive more input[50,61,62] (see also prior work section below). This proposal of amplification paired with blanket response reduction would also be consistent with the evidence we found (Figure 5) that sublinearity can be produced by suppressing neurons (see also LaFosse et al.[57])—a blanket suppression of many neurons would suppress their responses to input while leaving responses to natural input unchanged. A possible alternative to this proposal could be a mechanism of selective suppression of some patterns. This would mean the cortical network is organized to suppress a large number of non-natural patterns specifically. Given the larger space of non-natural patterns of activity compared with natural patterns of activity, selective suppression seems less likely than selective amplification combined with blanket suppression, but this is an interesting direction for future work.

Past work may not have identified this sequence filtering phenomenon because physiological recording methods are limited by the curse of dimensionality. It is impractical, in finite experimental time, to show sufficient stimuli to fully estimate models of receptive fields that encode complex features of natural vision. The other relevant constraint imposed by recording methods is the inability to easily determine which parts of neural responses to a stimulus come from feedforward input versus local, recurrent interactions. Prior recording studies did observe that natural responses were stronger and sparser on average than expected from simple receptive field combinations.[32,33] But they could not identify that response gain varies depending on prior visual stimuli. This would have required either a complex receptive field model impractical to estimate or another way to identify how a group of neurons was affected by different patterns of input to other local neurons. Two-photon stimulation allows us to match one specific excitatory pattern of input to the visual context in which it occurs. We can thus test whether a single pattern influences another, and the direction of its effect, without needing to identify interacting patterns via sampling large numbers of patterns.

Relation to prior work: Structure in spontaneous activity, straightening, cortical suppression, and hippocampal sequences

Prior observations that an echo of natural input can be seen in spontaneous activity[63–67] are consistent with our findings. During spontaneous activity, neurons in the cortex fluctuate[68–70] with measurable, but weak, correlation to natural visual responses. Our data suggest that while sequences of neural

responses during spontaneous activity states are largely suppressed by the cortical network, some spontaneous patterns that partially match natural patterns can be weakly amplified by the recurrent network. However, because responses during spontaneous activity lack the structured feedforward driving inputs that occur during vision, spontaneous patterns are not amplified to the same extent as visual inputs, resulting in the relatively weak spontaneous correlations previously reported.

Martin and Douglas[3,4] noted the large number of local recurrent synapses in the cortex and hypothesized that they could be used for amplification of input. Our work confirms this and extends their ideas by showing the kinds of inputs that can be amplified, including complex temporal statistics.

Another related observation is the straightening of perceptual and neural responses seen in natural vision.[71,72] Straightening refers to the relationship between response patterns at different moments of a natural movie. Prior work has found that cortical responses at one moment are more geometrically similar to later responses than is seen in the cortical inputs. The sequence filtering by the recurrent network we find could well be the mechanism for the straightening effect.

There has also been some evidence for nonlinear spatiotemporal processing in V1 at the level of single neuron receptive fields. David, Reid, et al.,[34,44,73] (reviewed by Ringach[45]) have found in primate or carnivore V1 some non-separable or nonlinear interactions in single neuron receptive field centers. This nonlinearity, like the complex "non-classical surround" in single neuron responses, could arise from sequence filtering at the population level. One pattern of input due to natural vision affects responses to future patterns of input, boosting responses in the way we have described here. Single neuron responses might be well-predicted if their interactions with many other local neurons could be measured, but in recordings of small numbers of neurons, this effect appears as receptive field nonlinearities.

Driving one cell in V1 creates an average suppressive effect that falls off with distance from the stimulated cell (also seen for small input patterns,[16,74], confirmed for balanced networks in simulation[50]). This might be summarized by saying that stimulating one cell creates suppression. But this is only an average effect and therefore not at odds with our findings. The broad suppressive effect seen in these studies is observed by averaging across many neurons, and the average effect can be suppressive even as individual neurons respond in heterogeneous ways. In fact, we observe this heterogeneity in cells' responses, as patterned optogenetic stimulation excites some non-target cells and suppresses others (Figure 5A; Figure S5E). Moreover, the statement that stimulating one cell in the cortex creates suppression is not always even true on average. Simulation of balanced, strongly coupled networks shows average suppression in some states and not others. In particular, single-cell stimulation combined with visual input produces average suppression,[74] but single-cell stimulation in the absence of visual stimulation has a nearly zero average effect.[50,75]

Another related prior observation is the pervasive suppressive response at the end of external cortical stimulation, widely observed in electrical microstimulation studies.[76,77] Stimulation is followed by an average suppression that lasts for approximately 150 ms. This duration is consistent with the timescale of suppression we find following a flashed natural scene (Figure 3L). Our results put forth a possible explanation for the previously reported suppression: stimulation is triggering the recurrent non-target effects that we observe (Figure 5). As above with single-cell stimulation, the non-target responses to patterned input are heterogeneous at the level of individual cells (Figure 2). The suppression in other studies reflects the average of many cells whose target (electrically stimulated) or non-target status is not known. Cortical post-stimulation suppression thus could well be a signature of the sequence filtering we observe and a phenomenon that arises from heterogeneous single neuron responses.

Last, our sequence effects are fundamentally different than the sequential replay that is observed in the hippocampus.[78] Our effect depends on firing rates over a few tens or hundreds of milliseconds, a distinct effect from the ordered spike sequences with millisecond-level precise timing that are generated during sharp-wave replay.[79,80]

The dense E-E cortical network in the sensory cortex may encode the structure of natural sensation

Our data suggest the large number of excitatory-excitatory recurrent connections in the cortex is used for learning the structure of the natural world. Depending on which cells are previously activated, the recurrent network creates effects in other neurons that allow relative boosting (or suppression) of responses to particular input sequences.

The numerous excitatory recurrent connections in the cortex can give the V1 network great capacity to selectively process the space of natural visual inputs.[36] The responses to visual input are high-dimensional: any given frame from a natural movie can affect the firing of many thousands of neurons in the cortex. So there is a large space of population responses, forming the set of patterns of activity that result from many different frames or different visual stimuli. While some sorts of dendritic nonlinearities (discussion in Histed[81]) and nonlinearities due to short-term synaptic plasticity[82] may also contribute, our data suggest that the high-dimensional space of visual input can be met with high-dimensional processing in the recurrent network. The recurrent network has numerous synapses, which can set the non-target pattern produced by any input pattern. There are hundreds of millions of excitatory-excitatory recurrent synapses in a cubic millimeter of cortex,[1] so the dense recurrent network of the cortex seems well-placed to support the space of transformations required for sequence filtering.

While specific inhibitory connections[83,84] may play a role in producing these non-target recurrent responses, no specific inhibitory population is needed to drive our effects. The mechanisms we observe may be controlled primarily by excitatory-excitatory synapses, without being mapped to a set of inhibitory cells. Here, we express opsin in excitatory cells only. This excitatory activation can cause suppression in other excitatory cells in steady state. For example, exciting excitatory cells may drive a few spikes in a number of inhibitory cells, which then increase their input to many excitatory cells. In turn, the excitatory cells' firing rates decrease, and that withdraws excitation from other connected excitatory cells. In this way, the steady-state firing rates of the excitatory network can be primarily controlled by the pattern of connections amongst the excitatory cells, not by

specific inhibitory connections, and without subsets of inhibitory cells that are substantially activated (reviewed in Histed[81]). Computations that are not mapped in a 1:1 way to inhibitory subclasses are attractive for high-dimensional coding, since the space of excitatory-excitatory connections is larger than the space of inhibitory-excitatory connections. This type of mechanism, with excitatory activity in one population driving either excitation or suppression in other excitatory cells in the presence of balancing inhibition, is consistent with the balanced state or inhibition-stabilized state of cortical operation.[50,69,70,75,85,86]

The sequence filtering computation is also a form of predictive processing.[27,87] Because responses to a given visual input are influenced by previous inputs, the natural, expected sequence of inputs is boosted. This predictive processing—changing the response to the next input based on the previous—is achieved by altering the network's input-output transformation, or network gain, based on prior inputs. This occurs via short-timescale non-target recurrent effects (Figures 2 and 3), without large error or mismatch signals that might be expected from traditional predictive coding theories.[27,88–91]

Sequence filtering seems likely to occur in other sensory regions beyond V1

While we have made these measurements in V1, we speculate that recurrent connections in other sensory regions—auditory, somatosensory, etc.—also are used for active filtering.[81] Those sensory modalities also process natural inputs that vary in time, though often with faster changes than in vision, predicting different timescales of sequence filtering in other sensory regions. Additionally, associative brain areas downstream of the primary sensory cortex can have tuning for complex sequences[92] and may filter at longer timescales.[93] The temporal precision of this filtering for different sensory modalities could be explored in future studies.

Beyond biology, recurrent networks in artificial systems also often are used to create temporally structured computations. Our model findings (Figure 4) are consistent with the idea that recurrent artificial networks can learn temporal statistics.[94] They support the idea that densely connected recurrent networks can perform sequence processing and learn temporal structure, not just in artificial systems but also in biological brains. Current state-of-the-art generative language models also learn temporal statistics, but using non-recurrent architectures.[95] Understanding the factors that favor different architectures for computation is an important continuing question at the intersection of neuroscience and AI.

These data show a new and powerful purpose for recurrent connectivity in the sensory cortex: to confer sensitivity to sequential input. The visual cortical L2/3 network is sensitive to dynamic visual context, boosting responses to sequences of input corresponding to dynamic natural sensation and suppressing others.

RESOURCE AVAILABILITY

Lead contact

Requests for further information and resources should be directed to and will be fulfilled by the lead contact, Mark Histed (mark.histed@nih.gov).

Materials availability

No new materials were generated in this study.

Data and code availability

All data and code are available on Zenodo (https://doi.org/10.5281/zenodo.18726362).

ACKNOWLEDGMENTS

We thank B. Averbeck, M. Stopfer, A. Levine, S. David, and members of the Histed Lab for comments and discussion. This work is supported by the following funding sources: National Institutes of Health grant U19NS107464 (M.H.), NIH intramural project ZIAMH002956, Pew Biomedical Scholars Program (N.S.), and Klingenstein-Simons Fellowship in Neuroscience (N.S.).

AUTHOR CONTRIBUTIONS

Conceptualization, C.E.D., Z.Z., M.H.H., and P.K.L.; methodology, C.E.D., Z.Z., P.K.L., Y.D., S.M., N.S., and M.H.; investigation, C.E.D., Z.Z., S.M., and P.K.L.; formal analysis, C.E.D. and Z.Z.; visualization, C.E.D., P.K.L., and Z.Z.; funding acquisition, M.H.H. and N.S.; project administration, M.H.H. and N.S.; supervision, M.H.H. and N.S.; writing – original draft, C.E.D. and M.H.H.; writing – review & editing, C.E.D., M.H.H., S.M., and N.S.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Animals
- METHOD DETAILS
  - Viral injection and cranial window implants
  - Retinotopic mapping
  - Two-photon holographic imaging and stimulation
  - Electrophysiology
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Two-photon data analysis
  - Cell pattern selection
  - RNN Model
  - Statistics

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.neuron.2025.12.024.

Received: May 13, 2025
Revised: August 27, 2025
Accepted: December 17, 2025
Published: March 3, 2026

REFERENCES

1. Braitenberg, V., and Schüz, A. (2013). Anatomy of the Cortex: Statistics and Geometry (Springer Science & Business Media).
2. Liu, Y.-J., Ehrengruber, M.U., Negwer, M., Shao, H.-J., Cetin, A.H., and Lyon, D.C. (2013). Tracing inputs to inhibitory or excitatory neurons of mouse and cat visual cortex with a targeted rabies virus. Curr. Biol. *23*, 1746–1755. https://doi.org/10.1016/j.cub.2013.07.033.

3. Douglas, R.J., Koch, C., Mahowald, M., Martin, K.A., and Suarez, H.H. (1995). Recurrent excitation in neocortical circuits. Science *269*, 981–985. https://doi.org/10.1126/science.7638624.
4. Douglas, R.J., and Martin, K.A.C. (2007). Recurrent neuronal circuits in the neocortex. Curr. Biol. *17*, R496–R500. https://doi.org/10.1016/j.cub.2007.04.024.
5. Hubel, D.H., and Wiesel, T.N. (1959). Receptive fields of single neurones in the cat's striate cortex. J. Physiol. *148*, 574–591. https://doi.org/10.1113/jphysiol.1959.sp006308.
6. Hubel, D.H., and Wiesel, T.N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. J. Physiol. *160*, 106–154. https://doi.org/10.1113/jphysiol.1962.sp006837.
7. Lien, A.D., and Scanziani, M. (2013). Tuned thalamic excitation is amplified by visual cortical circuits. Nat. Neurosci. *16*, 1315–1323. https://doi.org/10.1038/nn.3488.
8. Chung, S., and Ferster, D. (1998). Strength and orientation tuning of the thalamic input to simple cells revealed by electrically evoked cortical suppression. Neuron *20*, 1177–1189. https://doi.org/10.1016/S0896-6273(00)80498-5.
9. Li, Y.T., Ibrahim, L.A., Liu, B.H., Zhang, L.I., and Tao, H.W. (2013). Linear transformation of thalamocortical input by intracortical excitation. Nat. Neurosci. *16*, 1324–1330. https://doi.org/10.1038/nn.3494.
10. Priebe, N.J., and Ferster, D. (2012). Mechanisms of neuronal computation in mammalian visual cortex. Neuron *75*, 194–208. https://doi.org/10.1016/j.neuron.2012.06.011.
11. Ko, H., Hofer, S.B., Pichler, B., Buchanan, K.A., Sjöström, P.J., and Mrsic-Flogel, T.D. (2011). Functional specificity of local synaptic connections in neocortical networks. Nature *473*, 87–91. https://doi.org/10.1038/nature09880.
12. Lee, W.-C.A., Bonin, V., Reed, M., Graham, B.J., Hood, G., Glattfelder, K., and Reid, R.C. (2016). Anatomy and function of an excitatory network in the visual cortex. Nature *532*, 370–374. https://doi.org/10.1038/nature17192.
13. Rossi, L.F., Harris, K.D., and Carandini, M. (2020). Spatial connectivity matches direction selectivity in visual cortex. Nature *588*, 648–652. https://doi.org/10.1038/s41586-020-2894-4.
14. Douglas, R.J., Martin, K.A.C., and Whitteridge, D. (1989). A canonical microcircuit for neocortex. Neural Comput. *1*, 480–488. https://doi.org/10.1162/neco.1989.1.4.480.
15. Marshel, J.H., Kim, Y.S., Machado, T.A., Quirin, S., Benson, B., Kadmon, J., Raja, C., Chibukhchyan, A., Ramakrishnan, C., Inoue, M., et al. (2019). Cortical layer–specific critical dynamics triggering perception. Science *365*, eaaw5202. https://doi.org/10.1126/science.aaw5202.
16. Oldenburg, I.A., Hendricks, W.D., Handy, G., Shamardani, K., Bounds, H.A., Doiron, B., and Adesnik, H. (2024). The logic of recurrent circuits in the primary visual cortex. Nat. Neurosci. *27*, 137–147. https://doi.org/10.1038/s41593-023-01510-5.
17. Marr, D., and Hildreth, E. (1980). Theory of edge detection. Proc. R. Soc. Lond., B *207*, 187–217. https://doi.org/10.1098/rspb.1980.0020.
18. Olshausen, B.A., and Field, D.J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. Nature *381*, 607–609. https://doi.org/10.1038/381607a0.
19. Yoshida, T., and Ohki, K. (2020). Natural images are reliably represented by sparse and variable populations of neurons in visual cortex. Nat. Commun. *11*, 872. https://doi.org/10.1038/s41467-020-14645-x.
20. Ringach, D.L. (2002). Spatial structure and symmetry of simple-cell receptive fields in macaque primary visual cortex. J. Neurophysiol. *88*, 455–463. https://doi.org/10.1152/jn.2002.88.1.455.
21. Olshausen, B.A., and Field, D.J. (2006). What Is the Other 85 Percent of V1 Doing? In 23 Problems in Systems Neuroscience, L. van Hemmen and T. Sejnowski, eds. (Oxford: Oxford University Press), pp. 182–211.
22. Kang, I., Talluri, B.C., Yates, J.L., Niell, C.M., and Nienborg, H. (2025). Is the impact of spontaneous movements on early visual cortex species specific? Trends Neurosci. *48*, 7–21. https://doi.org/10.1016/j.tins.2024.11.006.
23. Niell, C.M., and Stryker, M.P. (2010). Modulation of visual responses by behavioral state in mouse visual cortex. Neuron *65*, 472–479. https://doi.org/10.1016/j.neuron.2010.01.033.
24. Vinck, M., Batista-Brito, R., Knoblich, U., and Cardin, J.A. (2015). Arousal and locomotion make distinct contributions to cortical activity patterns and visual encoding. Neuron *86*, 740–754. https://doi.org/10.1016/j.neuron.2015.03.028.
25. Nienborg, H., and Cumming, B.G. (2009). Decision-related activity in sensory neurons reflects more than a neuron's causal effect. Nature *459*, 89–92. https://doi.org/10.1038/nature07821.
26. Krug, K. (2004). A common neuronal code for perceptual processes in visual cortex? Comparing choice and attentional correlates in V5/MT. Philos. Trans. R. Soc. Lond. B Biol. Sci. *359*, 929–941. https://doi.org/10.1098/rstb.2003.1415.
27. Keller, G.B., and Mrsic-Flogel, T.D. (2018). Predictive Processing: A Canonical Cortical Computation. Neuron *100*, 424–435. https://doi.org/10.1016/j.neuron.2018.10.003.
28. Vetter, P., Smith, F.W., and Muckli, L. (2014). Decoding sound and imagery content in early visual cortex. Curr. Biol. *24*, 1256–1262. https://doi.org/10.1016/j.cub.2014.04.020.
29. Iurilli, G., Ghezzi, D., Olcese, U., Lassi, G., Nazzaro, C., Tonini, R., Tucci, V., Benfenati, F., and Medini, P. (2012). Sound-driven synaptic inhibition in primary visual cortex. Neuron *73*, 814–828. https://doi.org/10.1016/j.neuron.2011.12.026.
30. Ghazanfar, A.A., and Schroeder, C.E. (2006). Is neocortex essentially multisensory? Trends Cogn. Sci. *10*, 278–285. https://doi.org/10.1016/j.tics.2006.04.008.
31. Bimbard, C., Sit, T.P.H., Lebedeva, A., Reddy, C.B., Harris, K.D., and Carandini, M. (2023). Behavioral origin of sound-evoked activity in mouse visual cortex. Nat. Neurosci. *26*, 251–258. https://doi.org/10.1038/s41593-022-01227-x.
32. Vinje, W.E., and Gallant, J.L. (2000). Sparse coding and decorrelation in primary visual cortex during natural vision. Science *287*, 1273–1276. https://doi.org/10.1126/science.287.5456.1273.
33. Froudarakis, E., Berens, P., Ecker, A.S., Cotton, R.J., Sinz, F.H., Yatsenko, D., Saggau, P., Bethge, M., and Tolias, A.S. (2014). Population code in mouse V1 facilitates readout of natural scenes through increased sparseness. Nat. Neurosci. *17*, 851–857. https://doi.org/10.1038/nn.3707.
34. David, S.V., Vinje, W.E., and Gallant, J.L. (2004). Natural stimulus statistics alter the receptive field structure of v1 neurons. J. Neurosci. *24*, 6991–7006. https://doi.org/10.1523/JNEUROSCI.1422-04.2004.
35. Cossell, L., Iacaruso, M.F., Muir, D.R., Houlton, R., Sader, E.N., Ko, H., Hofer, S.B., and Mrsic-Flogel, T.D. (2015). Functional organization of excitatory synaptic strength in primary visual cortex. Nature *518*, 399–403. https://doi.org/10.1038/nature14182.
36. Stringer, C., Pachitariu, M., Steinmetz, N., Carandini, M., and Harris, K.D. (2019). High-dimensional geometry of population responses in visual cortex. Nature *571*, 361–365. https://doi.org/10.1038/s41586-019-1346-5.
37. Li, W., and Gilbert, C.D. (2002). Global contour saliency and local colinear interactions. J. Neurophysiol. *88*, 2846–2856. https://doi.org/10.1152/jn.00289.2002.
38. Olshausen, B.A., and Field, D.J. (2005). How close are we to understanding v1? Neural Comput. *17*, 1665–1699. https://doi.org/10.1162/0899766054026639.
39. Benucci, A., Ringach, D.L., and Carandini, M. (2009). Coding of stimulus sequences by population responses in visual cortex. Nat. Neurosci. *12*, 1317–1324. https://doi.org/10.1038/nn.2398.

40. Tolhurst, D.J., Walker, N.S., Thompson, I.D., and Dean, A.F. (1980). Nonlinearities of temporal summation in neurones in area 17 of the cat. Exp. Brain Res. *38*, 431–435. https://doi.org/10.1007/BF00237523.

41. Shamma, S. (2001). On the role of space and time in auditory processing. Trends Cogn. Sci. *5*, 340–348. https://doi.org/10.1016/S1364-6613(00)01704-6.

42. Theunissen, F.E., David, S.V., Singh, N.C., Hsu, A., Vinje, W.E., and Gallant, J.L. (2001). Estimating spatio-temporal receptive fields of auditory and visual neurons from their responses to natural stimuli. Network *12*, 289–316. https://doi.org/10.1080/net.12.3.289.316.

43. Machens, C.K., Wehr, M.S., and Zador, A.M. (2004). Linearity of cortical receptive fields measured with natural sounds. J. Neurosci. *24*, 1089–1100. https://doi.org/10.1523/JNEUROSCI.4445-03.2004.

44. DeAngelis, G.C., Ohzawa, I., and Freeman, R.D. (1993). Spatiotemporal organization of simple-cell receptive fields in the cat's striate cortex. II. Linearity of temporal and spatial summation. J. Neurophysiol. *69*, 1118–1135. https://doi.org/10.1152/jn.1993.69.4.1118.

45. Ringach, D.L. (2004). Mapping receptive fields in primary visual cortex. J. Physiol. *558*, 717–728. https://doi.org/10.1113/jphysiol.2004.065771.

46. Ringach, D.L., Hawken, M.J., and Shapley, R. (1997). Dynamics of orientation tuning in macaque primary visual cortex. Nature *387*, 281–284. https://doi.org/10.1038/387281a0.

47. Emiliani, V., Cohen, A.E., Deisseroth, K., and Häusser, M. (2015). All-Optical Interrogation of Neural Circuits. J. Neurosci. *35*, 13917–13926. https://doi.org/10.1523/JNEUROSCI.2916-15.2015.

48. Packer, A.M., Peterka, D.S., Hirtz, J.J., Prakash, R., Deisseroth, K., and Yuste, R. (2012). Two-photon optogenetics of dendritic spines and neural circuits. Nat. Methods *9*, 1202–1205. https://doi.org/10.1038/nmeth.2249.

49. LaFosse, P.K., Zhou, Z., Friedman, N.G., Deng, Y., Li, A.J., Akitake, B., and Histed, M.H. (2023). Bicistronic Expression of a High-Performance Calcium Indicator and Opsin for All-Optical Stimulation and Imaging at Cellular Resolution. eNeuro *10*. ENEURO.0378-22.2023. https://doi.org/10.1523/ENEURO.0378-22.2023.

50. O'Rawe, J.F., Zhou, Z., Li, A.J., LaFosse, P.K., Goldbach, H.C., and Histed, M.H. (2023). Excitation creates a distributed pattern of cortical suppression due to varied recurrent input. Neuron *111*, 4086–4101.e5. https://doi.org/10.1016/j.neuron.2023.09.010.

51. Bauer, J., Margrie, T.W., and Clopath, C. (2025). Movie reconstruction from mouse visual cortex activity. eLife. https://doi.org/10.7554/eLife.105081.2.

52. Hecht, S., and Verrijp, C.D. (1933). The Influence of Intensity, Color and Retinal Location on the Fusion Frequency of Intermittent Illumination. Proc. Natl. Acad. Sci. USA. *19*, 522–535. https://doi.org/10.1073/pnas.19.5.522.

53. Nomura, Y., Ikuta, S., Yokota, S., Mita, J., Oikawa, M., Matsushima, H., Amano, A., Shimonomura, K., Seya, Y., and Koike, C. (2019). Evaluation of critical flicker-fusion frequency measurement methods using a touchscreen-based visual temporal discrimination task in the behaving mouse. Neurosci. Res. *148*, 28–33. https://doi.org/10.1016/j.neures.2018.12.001.

54. Umino, Y., Solessio, E., and Barlow, R.B. (2008). Speed, Spatial, and Temporal Tuning of Rod and Cone Vision in Mouse. J. Neurosci. *28*, 189–198. https://doi.org/10.1523/JNEUROSCI.3551-07.2008.

55. Bonin, V., Histed, M.H., Yurgenson, S., and Reid, R.C. (2011). Local diversity and fine-scale organization of receptive fields in mouse visual cortex. J. Neurosci. *31*, 18506–18521. https://doi.org/10.1523/JNEUROSCI.2974-11.2011.

56. Smith, S.L., and Häusser, M. (2010). Parallel processing of visual space by neighboring neurons in mouse visual cortex. Nat. Neurosci. *13*, 1144–1149. https://doi.org/10.1038/nn.2620.

57. LaFosse, P.K., Zhou, Z., O'Rawe, J.F., Friedman, N.G., Scott, V.M., Deng, Y., and Histed, M.H. (2024). Cellular-resolution optogenetics reveals attenuation-by-suppression in visual cortical neurons. Proc. Natl. Acad. Sci. USA. *121*, e2318837121. https://doi.org/10.1073/pnas.2318837121.

58. Olshausen, B.A., and Field, D.J. (2004). Sparse coding of sensory inputs. Curr. Opin. Neurobiol. *14*, 481–487. https://doi.org/10.1016/j.conb.2004.07.007.

59. Sussillo, D., Churchland, M.M., Kaufman, M.T., and Shenoy, K.V. (2015). A neural network that finds a naturalistic solution for the production of muscle activity. Nat. Neurosci. *18*, 1025–1033. https://doi.org/10.1038/nn.4042.

60. Fujisawa, S., Amarasingham, A., Harrison, M.T., and Buzsáki, G. (2008). Behavior-dependent short-term assembly dynamics in the medial prefrontal cortex. Nat. Neurosci. *11*, 823–833. https://doi.org/10.1038/nn.2134.

61. Ahmadian, Y., Rubin, D.B., and Miller, K.D. (2013). Analysis of the stabilized supralinear network. Neural Comput. *25*, 1994–2037. https://doi.org/10.1162/NECO_a_00472.

62. Sanzeni, A., Histed, M.H., and Brunel, N. (2020). Response nonlinearities in networks of spiking neurons. PLoS Comput. Biol. *16*, e1008165. https://doi.org/10.1371/journal.pcbi.1008165.

63. Kenet, T., Bibitchkov, D., Tsodyks, M., Grinvald, A., and Arieli, A. (2003). Spontaneously emerging cortical representations of visual attributes. Nature *425*, 954–956. https://doi.org/10.1038/nature02078.

64. Luczak, A., Barthó, P., and Harris, K.D. (2009). Spontaneous events outline the realm of possible sensory responses in neocortical populations. Neuron *62*, 413–425. https://doi.org/10.1016/j.neuron.2009.03.014.

65. Tsodyks, M., Kenet, T., Grinvald, A., and Arieli, A. (1999). Linking spontaneous activity of single cortical neurons and the underlying functional architecture. Science *286*, 1943–1946. https://doi.org/10.1126/science.286.5446.1943.

66. Nguyen, N.D., Lutas, A., Amsalem, O., Fernando, J., Ahn, A.Y.-E., Hakim, R., Vergara, J., McMahon, J., Dimidschstein, J., Sabatini, B.L., et al. (2024). Cortical reactivations predict future sensory responses. Nature *625*, 110–118. https://doi.org/10.1038/s41586-023-06810-1.

67. Yao, H., Shi, L., Han, F., Gao, H., and Dan, Y. (2007). Rapid learning in cortical coding of visual scenes. Nat. Neurosci. *10*, 772–778. https://doi.org/10.1038/nn1895.

68. Shadlen, M.N., and Newsome, W.T. (1998). The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. J. Neurosci. *18*, 3870–3896. https://doi.org/10.1523/JNEUROSCI.18-10-03870.1998.

69. van Vreeswijk, C., and Sompolinsky, H. (1996). Chaos in neuronal networks with balanced excitatory and inhibitory activity. Science *274*, 1724–1726. https://doi.org/10.1126/science.274.5293.1724.

70. Brunel, N. (2000). Dynamics of Sparsely Connected Networks of Excitatory and Inhibitory Spiking Neurons. J. Comput. Neurosci. *8*, 183–208. https://doi.org/10.1023/A:1008925309027.

71. Hénaff, O.J., Goris, R.L.T., and Simoncelli, E.P. (2019). Perceptual straightening of natural videos. Nat. Neurosci. *22*, 984–991. https://doi.org/10.1038/s41593-019-0377-4.

72. Hénaff, O.J., Bai, Y., Charlton, J.A., Nauhaus, I., Simoncelli, E.P., and Goris, R.L.T. (2021). Primary visual cortex straightens natural video trajectories. Nat. Commun. *12*, 5982. https://doi.org/10.1038/s41467-021-25939-z.

73. Reid, R.C., Soodak, R.E., and Shapley, R.M. (1991). Directional selectivity and spatiotemporal structure of receptive fields of simple cells in cat striate cortex. J. Neurophysiol. *66*, 505–529. https://doi.org/10.1152/jn.1991.66.2.505.

74. Chettih, S.N., and Harvey, C.D. (2019). Single-neuron perturbations reveal feature-specific competition in V1. Nature *567*, 334–340. https://doi.org/10.1038/s41586-019-0997-6.

75. Sanzeni, A., Palmigiano, A., Nguyen, T.H., Luo, J., Nassi, J.J., Reynolds, J.H., Histed, M.H., Miller, K.D., and Brunel, N. (2023). Mechanisms underlying reshuffling of visual responses by optogenetic stimulation in mice and monkeys. Neuron *111*, 4102–4115.e9. https://doi.org/10.1016/j.neuron.2023.09.018.

76. Butovas, S., and Schwarz, C. (2003). Spatiotemporal effects of microstimulation in rat neocortex: a parametric study using multielectrode recordings. J. Neurophysiol. *90*, 3024–3039. https://doi.org/10.1152/jn.00245.2003.

77. Sombeck, J.T., Heye, J., Kumaravelu, K., Goetz, S.M., Peterchev, A.V., Grill, W.M., Bensmaia, S., and Miller, L.E. (2022). Characterizing the short-latency evoked response to intracortical microstimulation across a multi-electrode array. J. Neural Eng. *19*. https://doi.org/10.1088/1741-2552/ac63e8.

78. Davidson, T.J., Kloosterman, F., and Wilson, M.A. (2009). Hippocampal replay of extended experience. Neuron *63*, 497–507. https://doi.org/10.1016/j.neuron.2009.07.027.

79. Roumis, D.K., and Frank, L.M. (2015). Hippocampal sharp-wave ripples in waking and sleeping states. Curr. Opin. Neurobiol. *35*, 6–12. https://doi.org/10.1016/j.conb.2015.05.001.

80. Buzsáki, G. (1986). Hippocampal sharp waves: their origin and significance. Brain Res. *398*, 242–252. https://doi.org/10.1016/0006-8993(86)91483-6.

81. Histed, M.H. (2025). Active filtering: a predictive function of recurrent circuits of sensory cortex. Annu. Rev. Vis. Sci. *11*, 193–215.

82. Buonomano, D.V., and Maass, W. (2009). State-dependent computations: spatiotemporal processing in cortical networks. Nat. Rev. Neurosci. *10*, 113–125. https://doi.org/10.1038/nrn2558.

83. Sadeh, S., and Clopath, C. (2020). Patterned perturbation of inhibition can reveal the dynamical structure of neural processing. eLife *9*, e52757. https://doi.org/10.7554/eLife.52757.

84. Znamenskiy, P., Kim, M.-H., Muir, D.R., Iacaruso, M.F., Hofer, S.B., and Mrsic-Flogel, T.D. (2024). Functional specificity of recurrent inhibition in visual cortex. Neuron *112*, 991–1000.e8. https://doi.org/10.1016/j.neuron.2023.12.013.

85. Ahmadian, Y., and Miller, K.D. (2021). What is the dynamical regime of cerebral cortex? Neuron *109*, 3373–3391. https://doi.org/10.1016/j.neuron.2021.07.031.

86. Sanzeni, A., Akitake, B., Goldbach, H.C., Leedy, C.E., Brunel, N., and Histed, M.H. (2020). Inhibition stabilization is a widespread property of cortical networks. eLife *9*, e54875. https://doi.org/10.7554/eLife.54875.

87. Rao, R.P., and Ballard, D.H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nat. Neurosci. *2*, 79–87. https://doi.org/10.1038/4580.

88. den Ouden, H.E.M., Kok, P., and de Lange, F.P. (2012). How prediction errors shape perception, attention, and motivation. Front. Psychol. *3*, 548. https://doi.org/10.3389/fpsyg.2012.00548.

89. Schultz, W., and Dickinson, A. (2000). Neuronal coding of prediction errors. Annu. Rev. Neurosci. *23*, 473–500. https://doi.org/10.1146/annurev.neuro.23.1.473.

90. Lindsay, G. (2021). The Challenges of Proving Predictive Coding. Simons Foundation Global Brain. https://www.simonsfoundation.org/2021/06/03/the-challenges-of-proving-predictive-coding/.

91. Furutachi, S., Franklin, A.D., Aldea, A.M., Mrsic-Flogel, T.D., and Hofer, S.B. (2024). Cooperative thalamocortical circuit mechanism for sensory prediction errors. Nature *633*, 398–406. https://doi.org/10.1038/s41586-024-07851-w.

92. Purandare, C., and Mehta, M. (2023). Mega-scale movie-fields in the mouse visuo-hippocampal network. eLife *12*, RP85069. https://doi.org/10.7554/eLife.85069.

93. Murray, J.D., Bernacchia, A., Freedman, D.J., Romo, R., Wallis, J.D., Cai, X., Padoa-Schioppa, C., Pasternak, T., Seo, H., Lee, D., et al. (2014). A hierarchy of intrinsic timescales across primate cortex. Nat. Neurosci. *17*, 1661–1663. https://doi.org/10.1038/nn.3862.

94. Graves, A. (2013). Generating Sequences With Recurrent. Neural Networks Preprint at. arXiv. https://doi.org/10.48550/ARXIV.1308.0850.

95. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., and Polosukhin, I. (2017). Attention is all you need. Preprint at arXiv. https://doi.org/10.48550/arXiv.1706/03762.

96. Gorski, J.A., Talley, T., Qiu, M., Puelles, L., Rubenstein, J.L.R., and Jones, K.R. (2002). Cortical excitatory neurons and glia, but not GABAergic neurons, are produced in the Emx1-expressing lineage. J. Neurosci. *22*, 6309–6314. https://doi.org/10.1523/JNEUROSCI.22-15-06309.2002.

97. Ye, Z., Bull, M.S., Li, A., Birman, D., Daigle, T.L., Tasic, B., Zeng, H., and Steinmetz, N.A. (2023). Brain-wide topographic coordination of traveling spiral waves. Preprint at bioRxiv. https://doi.org/10.1101/2023.12.07.570517.

98. Goldbach, H.C., Akitake, B., Leedy, C.E., and Histed, M.H. (2021). Performance in even a simple perceptual task depends on mouse secondary visual areas. eLife *10*, e62156. https://doi.org/10.7554/eLife.62156.

99. Burgess, C.P., Lak, A., Steinmetz, N.A., Zatka-Haas, P., Bai Reddy, C., Jacobs, E.A.K., Linden, J.F., Paton, J.J., Ranson, A., Schröder, S., et al. (2017). High-Yield Methods for Accurate Two-Alternative Visual Psychophysics in Head-Fixed Mice. Cell Rep. *20*, 2513–2524. https://doi.org/10.1016/j.celrep.2017.08.047.

100. Steinmetz, N.A., Aydin, C., Lebedeva, A., Okun, M., Pachitariu, M., Bauza, M., Beau, M., Bhagat, J., Böhm, C., Broux, M., et al. (2021). Neuropixels 2.0: A miniaturized high-density probe for stable, long-term brain recordings. Science *372*, eabf4588. https://doi.org/10.1126/science.abf4588.

101. Giovannucci, A., Friedrich, J., Gunn, P., Kalfon, J., Brown, B.L., Koay, S.A., Taxidis, J., Najafi, F., Gauthier, J.L., Zhou, P., et al. (2019). CaImAn an open source tool for scalable calcium imaging data analysis. eLife *8*, e38173. https://doi.org/10.7554/eLife.38173.

102. Pachitariu, M., Stringer, C., Dipoppa, M., Schröder, S., Rossi, L.F., Dalgleish, H., Carandini, M., and Harris, K.D. (2017). Suite2p: beyond 10,000 neurons with standard two-photon microscopy. Preprint at bioRxiv *30*. https://doi.org/10.1101/061507.

103. Friedrich, J., Zhou, P., and Paninski, L. (2017). Fast online deconvolution of calcium imaging data. PLoS Comput. Biol. *13*, e1005423. https://doi.org/10.1371/journal.pcbi.1005423.

104. Allen Brain Atlas. Brain Observatory. Natural Movies. https://observatory.brain-map.org/visualcoding/stimulus/natural_movies.

105. Kingma, D.P., and Ba, J. (2014). Adam: A method for stochastic optimization. Preprint at arXiv. https://doi.org/10.48550/ARXIV.1412.6980.

106. Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., et al. (2019). PyTorch: An imperative style, high-performance deep learning library. Preprint at arXiv. https://doi.org/10.48550/arXiv.1912.01703.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
pAAV-hSyn-DIO-ChrimsonR-mRuby2-ST	Addgene	Addgene Plasmid #105448
pGP-AAV-syn-jGCaMP8s-WPRE	Addgene	Addgene Plasmid #162374
pAAV-hSyn-DIO-jGCaMP8s-P2A-stChrimsonR	Addgene	Addgene Plasmid #174007
Experimental models: Organisms/strains		
Emx1-Cre mice	the Jackson Laboratory	Strain #:005628 RRID:IMSR_JAX:005628

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Animals

All experiments were conducted in adherence to NIH and Department of Health and Human Services (HHS) guidelines for animal research and were approved by the Institutional Animal Care and Use Committee (IACUC) at the relevant institutions (optogenetic experiments: NIMH; Neuropixels experiments: University of Washington.)

Optogenetic experiments

Emx1-Cre animals (N=16, https://www.jax.org/strain/005628)[96] of both sexes were used. No systematic differences were observed between males and females. Viral injection and window implants were done at ages 2-7 months. After procedures, animals were singly housed on a reverse 12-hour dark/light cycle.

Neuropixels experiments

Male C57BL/6J mice (N=5), aged two to three months, were used. After procedures, animals were singly housed on a standard 12-hour dark/light cycle.

All animals were put on water schedule five or more days after head-plate surgery, and their weights were carefully monitored to ensure they remained above 85% of their baseline body weight.

METHOD DETAILS

Viral injection and cranial window implants

We performed injections and implants as described in [49] (optogenetic experiments; NIMH) or[97] (2023; Neuropixels; University of Washington). Minor differences in the two sets of approaches are not expected to impact results.

Optogenetic experiments

Briefly, mice were given dexamethasone (3.2 mg/kg) 30 minutes before surgery and anesthetized with 1%–3% isoflurane (in 100% O_2). A titanium headplate was affixed using C&B Metabond (Parkell) and a 3 mm craniotomy was made over V1 (-3.1 mm ML, +1.5 mm AP from lambda) and a glass optical window chronically inserted. Mice were injected with either a mixture of two AAV viruses to induce GCaMP and stChrimsonR into the cells (mixed virus injections), or a bicistronic virus expressing both GCaMP and stChrimsonR in each cell.[49] For the mixed injection, the fluorescent calcium indicator, (AAV9-syn-jGCaMP8s; titers: 5.0–10x10^{12} genome copies (GC)/ml) and the soma-targeted opsin (AAV9-hsyn-DIO-stChrimsonR-mRuby2; titer: 3.0x10^{12} GC/ml) were mixed in phosphate-buffered saline. In the case of the bicistronic virus injections, we used AAV9-hsyn-DIO-jGCaMP8s-p2a-stChrimsonR (titers: 2.9-5.9x10^{12} GC/mL). For each mouse, 3 to 5 injections (100 nL/min, 200 μm depth) were made to cover a wider area of the cortex. A custom made light-blocking cap was fixed onto the implant to limit ambient light exposure and prevent debris from contacting the window. Experiments began three weeks after these procedures.

Neuropixels experiments

Briefly, mice were induced into anesthesia with 5% isoflurane, and maintained at 2-3%. Carprofen (5 mg/kg) and Lidocaine (2 mg/kg) were administered for analgesia. A titanium head plate and a 3D-printed chamber were affixed using Metabond. Carprofen was administered at 0.05 mg/ml in the water post-surgery for three days. Five days following recovery, mice were put on water restriction. After two days, they were habituated to head fixation for two days, during which they received up to thirty random 5μL sucrose water 10% rewards. Subsequently, animals were exposed to a set of 20 natural images presented over 1000 trials, followed by a 5μL sucrose water reward for five days. This step was performed to provide a control training condition for another experiment. None of the 20 images in this step were used as part of the experiment mentioned in this work. One or two days before recordings, a 2 x 2 mm craniotomy was performed over V1 using a dental drill and stereotaxic techniques. The area was then covered with silicone gel. This

process was conducted under anesthesia procedures used for the head plate implant, and Carprofen (5 mg/kg) was given for analgesia. A cap was placed on the chamber to prevent debris from getting into the area around the craniotomy.

Retinotopic mapping

Before optogenetic experiments, we determined the location of V1 in the cranial window using a hemodynamic intrinsic imaging protocol previously described in [98]. Briefly, small visual stimuli were presented while 530 nm light was delivered using a fiber-coupled LED (M350F2; Thorlabs, Newton NJ). Hemodynamic response was calculated as the change in reflectance of the cortical surface between the baseline period and a response window starting 3 ms after stimulus onset. Imaging was done on a Zeiss Discovery stereo microscope with a 1x widefield objective through a green long-pass emission filter, acquired at 2Hz. An average retinotopic map was fit to the cortical responses based on the centroids of the hemodynamic response for each stimulus location.

Two-photon holographic imaging and stimulation

Two-photon imaging and stimulation procedures are described in detail in.[49] Briefly, animals were awake and alert under a 16x water-immersion objective (Nikon; Tokyo, Japan; NA=0.8) manually positioned over the implanted optical window. Imaging was done with a custom-built microscope and controlled by ScanImage software in MATLAB (The Mathworks, Natick, MA). Calcium responses were measured ~100-200 μm below the surface of the pia (L2/3 of V1) with an imaging field of view of 414 x 414 μm. Imaging was performed using 920 nm wavelength light at 15-20 mW and frames acquired at 30 Hz.

Holographic stimulation was performed using a femtosecond pulsed laser (1030 nm, Satsuma, Amplitude Laser, or 1040nm, Monaco, Coherent Inc.) A spatial light modulator (SLM) shaped the laser wavefront to create stimulation patterns (10 μm diameter disks; 2.5-5.3mW/target, 500 kHz pulse rate). Timing of sequential pattern presentation was selected to mimic the temporal structure of natural visual inputs: pattern duration was 30, 60, 91, or 121 ms (1-4 imaging frames; frame duration approx 30.1 ms), and overlap between patterns (0-1 frames, 0-30 ms) was sometimes used to simulate the effect of a dynamic visual stimulus, which can change more frequently than pattern period. Overlap in patterns: Fig. 1IJ, 30 ms; Figure 2, 0 ms; Figure 3, 0 ms, Figure 4, 0 ms, Figures 5A–5F, 30 ms. The response to a 50 ms flashed single visual input lasts around 70 ms (Figure 3L), overlapping with the next input from a natural visual stimulus. The radial point-spread function (PSF) of diffraction limited spots generated by the SLM was 9.4 μm and the axial PSF was 54 μm. The laser was gated on during horizontal flyback periods and off during the imaging pixel acquisition to allow for approximately simultaneous stimulation and imaging. Reported stimulation power is the average power over on and off periods (i.e. reduced by a factor of 0.3 from the laser power measured without this gating.) To minimize crosstalk we chose stChrimsonR opsin. Previous work with the bicistronic virus used here has not shown evidence of crosstalk activation.[49] This work additionally characterized off-target stimulation and found the majority of off-target stimulation occurs in cells within this 10 μm radius when stimulating ~20 or fewer cells. Thus, we generally avoid selecting cells whose centroid is within a 10 μm of another selected cell.

Visual responses were measured in awake, head-fixed mice viewing flashed single natural image frames (40° circular mask with neutral gray background; 120 ms) or full natural movie stimuli (full-field, 2 seconds, 30Hz frame rate) presented on an LCD monitor. Mice were given small volumes of water on 20% of trials (3 μL).

Electrophysiology

For Neuropixels electrophysiology the mice were head-fixed for the recordings while seated in plastic support with forepaws on a wheel,[99] with the ability to move body parts other than the head but no ability to locomote. The behavioral state of the animals was monitored using two Basler acA2440-75um cameras at 560x560 resolution. One camera focused on the eye area was used to monitor the pupil diameter, and the other focused on the face was used to detect licks and ensure the animal was not in distress. Animals were awake during the duration of the experiment. Recordings were conducted from all layers of V1 using Neuropixel 2.4 probes[100] with a sampling rate of 30 kHz. The recording sites were targeted based on the stereotaxic coordinates in the CCF Allen atlas using the Pinpoint system.[97] The reference for the recordings was set at the tip of the electrode. Three recordings were performed: one from the first animal and two (on subsequent days) from the second animal. Spike sorting was executed using Kilosort 2.5 (https://github.com/MouseLand/Kilosort; RRID: SCR_016422)[100] on each of the four shanks separately.

Visual stimuli were presented on three 60 Hz screens (LG LP097QX1), surrounding the mice and covering 270 x 70 degrees (azimuth x elevation) visual angle. Each trial featured a natural image from a set of two, sourced from the Allen Institute Brain Observatory, and repeated across all screens. These images were equiluminant and underwent histogram equalization for contrast consistency. Images were warped to appear rectilinear from the animals' viewpoint. Each image was displayed 30 times for 50 ms. A reward of 5 μL of 10% sucrose water was administered at 1150 ms post-stimulus offset to maintain engagement. The trials were randomly interleaved with each other and with additional trials from another experiment, which also included natural image presentations and sucrose rewards. The inter-trial interval followed an exponential distribution with a minimum and mean of 2 and 2.6 seconds, respectively.

All electrophysiology analysis was done in Python. Baseline value computed in a 40 ms window ending 5 ms prior to stimulus presentation. Response value was calculated across a 50 ms interval beginning 5 ms after stimulus presentation. Excited cells are those with average response greater than baseline; suppressed cells those with response less than baseline.

QUANTIFICATION AND STATISTICAL ANALYSIS

Two-photon data analysis

We performed motion correction using the CaImAn toolbox[101] and cell segmentation using Suite2p.[102] All data analysis was done in Python (https://www.python.org). For pixel based analyses, we computed $\Delta F/F_0$ ($\Delta F/F_0$; F: raw fluorescence intensity; F_0: average fluorescence across the 45 imaging timepoints, 1.5 s, prior to stimulus presentation) at every pixel of the image stacks for display across the FOV. The $\Delta F/F_0$ for all time courses was calculated from fluorescent traces output from Suite2p. Neuropil is subtracted during the Suite2p preprocessing to account for extraneous signals. However by design, because the stimulations occur at different time points, changes in pixel response due to cell body or dendrites would be modulated by temporal context thus the conclusion would be unchanged. Off-target suppression measured in Figure 5 would only be diluted by dendritic activation.

For full sequence trials (ABC, CBA), averaging windows for each pattern (A, B, C) started at stimulation onset of that pattern and lasted 300 ms; shifted for pattern order so the response window always began on the first frame of stimulation of that pattern (raw time course data without shift is shown in Figure 2B). To isolate the optogenetic response following visual input, we subtracted off the response to the visual stimulus alone, [(vis+opto)-vis] (Figure 3). Averaged response interval for individual cells began at optogenetic stimulus onset and ended at stimulus offset (120 ms, Figure 3K). Activity of the C pattern in ABC sequence was calculated as [ABC – AB] (Figures 5D and 5E). In Figures 5I–5K, the averaging interval for individual cell responses started at stimulus onset and lasted 200 ms. Response intervals for all pixel based analyses started at stimulus onset and lasted 300 ms (Figures 3B–3H and 5A). Deconvolution (Figure 1J) was done using OASIS with an autoregressive constant of 1.[103] Population sparsity[32] (Figure 5K) was calculated as:

$$S = \left\{1 - \left[\left(\sum r/n\right)^2 \Big/ \sum\left(r^2/n\right)\right]\right\} \Big/ \left[1 - (1/n)\right]$$

Where r is the cell responses in the frame following the last stimulation and n is the number of cells.

Cell pattern selection

For analysis of visual response to flashed natural images, individual cell responses were averaged over a 300 ms interval starting from the stimulus onset. Cells activated by the visual stimulus were defined by responses exceeding a 5% $\Delta F/F_0$ threshold. All cells that met this threshold were included in the optogenetic pattern (11–15 cells per pattern, Figures 3D and 3E). Data includes three randomly selected time points across two natural movies (Mouse movie: Video S1; Touch of Evil Movie[104]).

To construct the boosted and attenuated patterns (Figure 5), we first selected at random (not based on visual responses) the first 15–20 cell pattern in the sequence. Average $\Delta F/F_0$ response to stimulation of that pattern was calculated across a 300 ms time interval starting at stimulation onset. To select the latter patterns, cells within a 10 μm radius of a previously stimulated cell were excluded to control for off-target effects. Excited non-target cells meeting a minimum threshold of 5% $\Delta F/F_0$ were included in the next pattern for the boosted sequence (15–20 cells per pattern). Conversely, suppressed non-target cell responses below -5% $\Delta F/F_0$ (15–20 cells per pattern) were selected for the attenuated sequence pattern. If more than 20 cells met these criteria, the 20 cells with the largest magnitude response were used. This process was repeated for the two sets of first and second patterns to generate the third pattern for each sequence. Each pattern in both sequences had an equal number of cells. Any cells included in previous patterns were excluded from selection for the subsequent patterns (Figures 5G–5J).

RNN Model

We trained a recurrent neural network (RNN) consisting of N=500 units, whose input dynamics for the ith neuron are given by:

$$\tau\frac{dx_i}{dt} = -x_i + \sum_{j=1}^{N} W_{ij}^{rec}\phi(x_i) + I_{im}(t) + \eta_i$$

The input structure is defined by:

$$I_{i,m}(t) = w_i^{in} \cdot \sum_{k=1}^{N_k} P_{k,m}^{wf}(t) \cdot P_{i,k}^{pat}$$

Where P_{km}^{wf} is a matrix specifying the input waveform for each pattern, k, and sequence,m. $P_{ik}^{pat} = \{0, 1\}$ is an indicator matrix denoting which cell i is in each pattern k. The weights of the input pattern w_{in} are positive and exponentially distributed for a fraction $p = 0.3$ of units, and zero otherwise.

The readout of the network is defined as:

$$Z_i(t) = w_i^{out}\phi(x_i(t))$$

The transfer function of single units is $\phi(x) = tanh(x)$. The readout weights, w_{out}, are an Identity matrix. The initial recurrent weights W_{ij}^{rec}, before any training, are independently sampled from a random Gaussian distribution with a mean zero and standard deviation $g_0/\sqrt{N}$. The noise term η_i is randomly sampled from a zero mean distribution with standard deviation 0.0005 at every time step.

We trained the recurrent weights W_{ij}^{rec} of the RNN using backpropagation-through-time (ADAM optimizer in PyTorch)[105,106] such that the network readout Z for designated "trained sequences" matches a sparse scaled version of the time-varying input $I_{im}(t)$. To mimic the strong sparse activity seen in response to natural movies (and in Figures 5G–5K), we set the target output so 35% of the cells in the pattern (7 of 20) have a target level that is 2x the input and the other 65% (13/20) have a target that is 0.9x their input. The network readout Z for "control sequences" was set to match the time-varying input without any scaling. Sequences were generated by selecting 5 patterns from a set of 18 distinct input patterns. Each pattern had independently sampled weights from the exponential distribution. Untrained sequences were scrambled versions of the numerically ordered trained sequences. The input and output weights remained fixed. We trained the network on 500 epochs to produce these strong sparse responses. Control training had a target readout that exactly matched the time-varying input for all sequences. Parameters: $\tau = 60\,ms$, $g_0 = 0.8$, Euler integration timestep $\Delta t = 1\,ms$, learning rate 0.01.

Statistics

All statistical analyses were performed using Python. Statistical tests used include t-tests (Figures 2D, 5D–5F, 5I, and 5K), Kolmogorov-Smirnov and Anderson-Darling tests (Figures 2E, 2N, 2O, 5F, and 5J), Mann-Whitney U tests (Figures 2G–2L and 3K) and Mood's median test (Figure 2Q). Significance threshold was held at α=0.05; n.s., not significant ($p>0.05$); *$p\leq0.05$, **$p\leq0.01$, ***$p\leq0.001$. All experiments were replicated in multiple animals. All errorbars shown are SEM unless otherwise noted.