Critical Scaling of Novelty in the Cortex

Tiago L. Ribeiro, Ali Vakili, Bridgette Gifford, Raiyyan Siddiqui, Vincent Sinfuego, Sinisa Pajevic, and Dietmar Plenz[†]

Section on Critical Brain Dynamics, National Institute of Mental Health, Bethesda, MD, USA

[†] Correspondence: plenzd@mail.nih.gov

Running Title: Critical scaling of novel cortical action potentials *in vivo*

Correspondence: Dietmar Plenz, Ph.D., Section on Critical Brain Dynamics, National Institute of Mental Health, Porter Neuroscience Research Center, Rm 3A-1000, 35 Convent Drive, Bethesda, MD 20892. office: (301) 675-5311

Keywords: Perturbation, population coding, criticality, scaling, neuronal avalanches, visual cortex, *in vivo*, 2-photon imaging, holography, all-optical

Abstract

The ability to detect unanticipated, novel events and rapidly relay this information across neural networks is fundamental to brain function, enabling the selection of appropriate behavioral responses. Here, we examine the transmission of holographically triggered action potentials in primary visual cortex of quietly resting mice, focusing on the dynamics of communication from pyramidal neurons. We demonstrate that these novel action potentials, which are uncorrelated with preceding activity, exert a disproportionally large influence on neighboring neurons. Their influence scales robustly to an exponent between 0.2 and 0.3 relative to their number. Remarkably, even a small number of novel action potentials can engage a majority of the local network, achieving high decoding accuracy of the perturbation origin in the face of high trial-by-trial variability and ongoing activity characterized by scale-invariant, parabolic neuronal avalanches. This heightened susceptibility to small, local perturbations aligns with the behavior of complex systems exhibiting critical dynamics. Our findings reveal that scaling underpins the efficient communication of unanticipated action potentials, suggesting it is a fundamental mechanism for detecting and processing novel events in the brain. These results provide new insights into the neural basis of novelty detection and highlight the importance of critical dynamics in cortical network function.

Introduction

Thriving in uncertainty requires detecting and responding to the unexpected. In mammals, unexpected events must be effectively communicated within and across brain areas to ensure accurate evaluation and appropriate behavioral responses[1]. A key step in this process is how novel action potentials in the cortex – those in a single pyramidal neuron that exceed near-future expectation based on ongoing activity – are transmitted to the surrounding cortical network.

Neuronal spiking has been found to shape cortical activity and behavior. A spike in a pyramidal neuron from superficial layers (L2/3) of cortex of mice has been estimated to influence fewer than 1% of nearby neurons, to evoke strong local inhibition[2-5], and in reptilian cortex to reliably recruit activity sequences[6]. Coordinated activity in small groups of L2/3 pyramidal neurons elicits spatially extended variable responses[7] and can drive trained behavior via pattern completion[8,9]. Similarly, spike bursts in deep layers (L5/6) can elicit sensitized behavioral responses[10,11] and alter global brain states[12,13]. These optogenetical studies in mice are in line with previous findings from cortical microstimulation in primates highlighting the capacity of few spikes to significantly alter brain activity under specific conditions (for review[14]). However, a comprehensive framework to quantify the impact of action potentials on the spatiotemporal dynamics of cortical networks remains elusive.

In principle, novel spikes representing unexpected events should induce propagation across the
cortex both broadly (to reach multiple functional domains for evaluation given their uncertain nature)
and rapidly (to be able to act quickly). However, this dissemination of information is constrained by
several factors. First, even in the resting state, the cortex exhibits substantial activity, with individual
neurons receiving thousands of action potentials every second. Accordingly, individual neurons operate
in a fluctuation-dominated[15], high-dimensional[16] regime that shapes their ability to respond to and
encode information from novel spikes[17] or sensory inputs[18]. Second, the sparse synaptic connectivity
between L2/3 pyramidal neurons — estimated at only 5–8% and diminishing sharply within 200 μm[19-23]
— further restricts the transmission of novel spikes[24]. Additionally, the weak and transient synaptic
currents of these failure-prone connections in vivo[2,22], coupled with strong local inhibition[25,26], further
reduce the potential influence of novel spikes introduced randomly into a pyramidal neuron in the
awake state. Given these limitations, understanding how novel spikes are encoded requires a focus on
the network's broader dynamics rather than the activity of individual neurons or microcircuits. However,
adopting this broader perspective presents a significant challenge due to the inherent complexity of
cortical activity.

Research suggests that brain networks operating near a critical state exhibit heightened
sensitivity to small perturbations[27,28], which might enable the cortex to detect and process new
information despite its complex internal activity and sparse topology. In this state, the system
undergoes internal fluctuations while remaining responsive to small, local events that can propagate
broadly, facilitating the recognition and dissemination of novel information. In the superficial cortical
layers, these fluctuations manifest as "neuronal avalanches" — scale-invariant patterns of synchronized
neuron activity characteristic of the awake state[29-31]. These avalanches align with theoretical predictions
of critical dynamics. However, direct experimental evidence demonstrating that novel, localized
information propagates within the neuronal avalanche regime and scales in line with predictions from
critically is lacking, leaving a gap in understanding whether and how criticality supports the spread of
novel information in the brain.

Using an all-optical approach, we holographically triggered spikes in individual L2/3 pyramidal
neurons in the primary visual cortex of awake mice during spontaneous, ongoing avalanche activity.
Simultaneously, we recorded the activity of hundreds of nearby neurons. Our results demonstrate
robust scaling of network responses, with novel spikes rapidly generating widespread, information-rich
activity across large cortical distances. These findings align with predictions for responses to small
perturbations in critical systems, highlighting the role of critical dynamics in efficient cortical information
processing.

Results

Holographically triggered spikes in a pyramidal neuron induce rapid, widespread excitation in
superficial cortical layers

In awake, quietly resting adult mice, we holographically targeted single L2/3 pyramidal neurons (Target
Cells, TC) in primary visual cortex (V1) that were chronically co-expressing the calcium indicator
jGCaMP7s and the opsin ChrimsonR (5 to 10 mW, 100 ms; ~100 – 150 trials per TC) while
simultaneously monitoring spiking activity from more than hundred nearby neurons using 2-photon
imaging (field of view: 450 µm x 450 µm; 50 – 140 µm cortical depth; temporal resolution: 22 ms; n = 6
mice; 16 experiments). Sparse co-expression, stimulation beam profiling, and anatomical
reconstruction confirmed select single neuron stimulation (Fig. 1a; Supplementary Figs. 1–3).
94   Holographic stimulation of a single pyramidal neuron caused rapid and wide-spread effects
across the surrounding network. Within 132 ms, approximately 15% of non-targeted neurons increased
spiking (Positive Responders, PosR), while fewer than 1% showed decreased spiking (Negative
Responders, NegR), despite high trial-to-trial variability (Fig. 1b, c; 2.5% significance threshold, one-
tailed Welch's *t*-test). Baseline activity was consistent across all groups (TC, PosR, NegR, and non-
responders, NonR), with neurons exhibiting spike count variability well above Poisson expectations
(Fig. 1d; baseline; n = 54 TC; n = ~9,000 PosR, NegR, NonR). Upon stimulation, PosR firing rates and
spike count variance increased markedly remaining well above expectations for a Poisson process as
reported previously[32] (Fig. 1d; response).

Novel spikes scale network responses without change in overall network variability

Holographically triggered spikes in TC neurons were classified as "novel" because they were not
predicted by ongoing activity, as demonstrated by autocorrelation analysis. During the baseline period,
ongoing spiking exhibited significant temporal order, with high autocorrelation across all groups (r ≈ 0.2,
Fig. 2a, b; base' vs. base). However, during the response, the autocorrelation for TC neurons (base vs.
resp) dropped to nearly zero (r ≈ 0; p < 10^{-3}, Wilcoxon rank sum), indicating that stimulation introduced
spikes randomly, disrupting the temporal organization of ongoing TC spiking. In contrast, temporal
correlations remained stable for PosR and NonR, while the observed decrease in NegR autocorrelation
reflected overall reduction in sparse spiking (Fig. 2b).

Fig. 1 | Holographically triggered spikes in a single pyramidal neuron rapidly elicit significant excitation in many visual cortex neurons during ongoing activity at rest. a Holographic stimulation (100 ms, *flash;*) targeting a single pyramidal neuron (*TC*; averaged over n = 150 trials) alters the relative fluorescence in the TC *soma*, its *dendrites*, and neighboring neurons (indicated by *arrows*), reflecting changes in spontaneous spiking activity. *Inset*: Mean from n = 54 TC somata. **b** Example of variable spike count (z-scored) in a holographically targeted TC (100 ms, 5mW stimulation) and neurons within a 450×450 µm² field of view (n = 166, ranked by mean spike count within 132 ms of stimulation; n = 50 consecutive trials; single experiment). *Bottom*: Summed z-score of population activity per trial, highlighting overall network response. *Arrows* indicate trials with a high number of co-active neurons. **c** Mean transient response to short-lasting holographic stimulation. Average time course (n = 6 mice; 16 experiments) of z-scored spike count at frame rate resolution (22 ms) for n = 54 TC, n = 1,221 Positive Responders (*PosR*), n = 69 Negative Responders (*NegR*) and n = 7,687 Non-significant Responders (*NonR*). Protracted and prolonged reconstructed responses are observed due to deep interpolation (see Supplementary Fig. 7). *Base', Base* and *Resp* indicate respective time windows of analysis. *Yellow bar*: Stimulation window. **d** The response to single TC stimulation is largely excitatory, while maintaining variability above Poisson-level expectation (*dashed line*). Summary scatterplot of variance vs. mean spike count/132 ms for baseline and response (n = 6 mice, n = 16 experiments). Approximately 15% of neurons exhibit a significant increase in spike count from baseline (*PosR*), while fewer than 1% show a significant decrease (*NegR*). For *NonR*, only marginal distributions are shown. *CPF*: Cumulative Probability Function.

115   Spike count changes in TC in relation to other neuron categories exhibited distinct differences

116   for baseline and response which we quantified using cross-correlations and scaling functions. During

117   baseline, TC spike counts (TC_{sc}) were significantly correlated with all groups (<cc> ≈ 0.05; Fig. 2c) and

118   followed a consistent scaling relationship:

119                              $PosR_{sc}, NegR_{sc}, NonR_{sc} \sim TC_{sc}^{h}$,   (1)

120   with h ≅ 0.5 (linear regression; Fig. 2d). This scaling applied both to per-neuron spike counts and totals

121   across all neurons and was observed for the majority of baseline fluctuations from 1 to ~ 7 spikes/132

122   ms (transient firing rate increases of 8 – 53 Hz; see also Supplementary Fig. 4).

123   Upon stimulation, the excitatory response in PosR showed similar scaling:

124                              $PosR_{sc} \sim TC_{sc}^{h}$,   (2)

with a significantly lower exponent of h ≅ 0.3 (linear regression; Fig. 2e, f). In contrast, NonR exhibited

no scaling, while NegR showed negative scaling. This demonstrates that network responses to single

spikes scale similar to responses to multiple novel spikes from a single neuron. Despite our partial

control of holographical spike generation due to variability from external (e.g., head movement) and

internal (e.g., network inhibition) factors, these scaling relationships, nevertheless, derived a posteriori

by averaging trials, confirm that driving a single pyramidal neuron distinctly affects neuronal populations

compared to baseline. Our 1–7 added spikes per 132 ms closely match the 1–7 spikes/132 ms seen in

spontaneous fluctuations, showing our holographic manipulation stays within normal cortical firing

ranges (<20% of trials exceeded 7 spikes/132 ms).

The distinct scaling relationships between TC stimulation and the excitatory network response

was mirrored in corresponding positive, zero, and negative cross-correlations between TC and the

other neuron categories (Fig. 2c). The positive cross-correlation found for TC and PosR exhibited only

weak overlap in evoked and functional connectivity (R^2 = 0.018; linear regression), demonstrating that

novel spikes scale with neighboring sub-networks largely unexplored during ongoing activity. These

scaling relationships were robust to the significance level alpha employed. At ~10 times lower alpha,

the scaling exponent *h* for PosR was still close to 0.2 (Fig. 2g).

The sparse synaptic connections between local L2/3 pyramidal neurons — estimated at only 5–

8% and largely absent beyond 200 μm[19-23,33] suggest that among 200 neurons observed, less than 16

neurons are synaptically connected to a given TC. PosR were consistently found near TCs within a

~200 μm radius as reported previously[4], in line with expectations for locally connected neighbors

providing the strongest spike count increases. However, a substantial portion of PosR was distributed

throughout the network (Fig. 2g–h) and this proportion increased as the significance threshold (α) was

lowered. At the lowest significance level examined, nearly half of all recorded neurons showed

significant positive responses, with an average scaling exponent of *h* ≥ 0.25.

Fig. 2 | Scaling of cortical network responses to novel, holographically triggered spikes in individual pyramidal neurons in awake mice. a Schematic view of activity windows to calculated auto- and cross-correlations. **b** Holographically triggered TC spikes are novel, as evidenced by their zero autocorrelation during ongoing activity. In contrast, autocorrelations remain high for other neuron categories, except NegR, due to spike failure during the response. **c** Ongoing TC spike fluctuations positively correlate with all neuron groups. However, evoked TC spikes selectively correlate positively with PosR, show no correlation with NonR, and correlate negatively with NegR. **d** Scaling for ongoing fluctuations in TC with respect to all other groups. Note limited range in spontaneous TC count fluctuations of <10 spikes/132 ms. **e** Robust scaling for novel spikes in TC with respect to PosR but not NonR or NegR. **f** Scaling of population spike count for each neuron category separately and combined (*black*). **g** Strong positive scaling for PosR remains even when significance level alpha is greatly reduced. Change in scaling exponent as a function of alpha per neuron (*avg; solid lines*) and population (*sum; dashed lines*) for each group. **h** PosR are predominantly located within a 100 μm radius of the TC but also distribute throughout the network as alpha decreases. **i** Relative number of neurons per category changes with alpha. Note that PosR and NegR make up the majority of the observed network at the lowest significance level at which robust scaling is still observed. **j** The Fano Factor (FF) remains unchanged for PosR and NonR despite significant scaling in PosR (α = 2.5%). *Opaque bars*: Mean spike count corrected FF (see Methods). **k** Population FF also shows no change across categories. *Shaded areas* (d − f) and error bars (b, c, j, k): SD. When comparing correlations (b, c) and FF (j), 1, 2 & 3 stars indicate p < 0.05, 10^{-3} & 10^{-6}, respectively, using Wilcoxon rank sum test.

The corresponding recruitment of both PosR and NegR from the NonR pool (Fig. 2g–i) suggests

that novel spikes are transmitted across the network via multi-synaptic pathways and concomitant

increases as well as decreases in spiking as recently described when stimulating groups of pyramidal

neurons simultaneously[7]. The engagement of a large number of neurons distributed over broad spatial

regions when stimulating a single pyramidal neuron contrasts with expectations from anatomy and

single cell physiology regarding the efficacy of transmitting single action potentials.

A reduction in spiking variability, measured by the Fano factor (FF; variance-to-mean spike

count), is a well-documented response to sensory stimulation[34] and may underlie the robust scaling

observed with TC stimulation. On the other hand, simultaneous holographical stimulation of groups of

neurons exhibits variable spike counts in individual TCs through recurrent circuit dynamics[7], which

should counter-balance FF decreases in the network. We found that repeated holographic stimulation

reduced the FF in TC neurons as expected from the increase in regular driving, however, the FF for

PosR and NonR remained unchanged, even after accounting for differences in mean spike count[35] (Fig.

2j). Similarly, no change in FF was observed for PosR and NonR population responses (Fig. 2k). To

investigate whether variability in TC spike triggering masked changes in FF for PosR and NonR, we

analyzed subgroups of TC trials with differing FF, population activity, and post-response time windows.

However, the FF remained consistently high and unaltered for both groups across all conditions

(Supplementary Figs. 5, 6).

Our findings demonstrate that novel spikes in a single pyramidal neuron elicit significant scaling

responses in a large group of neighboring neurons, while preserving a fluctuation-dominated regime.

This indicates that cortical network in vivo operate in a fluctuation-dominated regime in which local

fluctuations exhibit responses in line with critical scaling[28,36].

Temporally accurate information on the origin of novel spikes distributes throughout the

network

We next demonstrate that widespread spike count changes in response to novel spikes encode crucial

information about their origin, which is rapidly and transiently distributed across the cortical network. In

n = 7 experiments, up to 10 TCs were stimulated in a semi-random order (100 ms each, with 2–10 s

intertrial intervals; Fig. 3a). The origin of these novel spikes was successfully decoded from the activity

of the non-driven neurons using two machine learning approaches: XGBoost[37] (XGB) and Random

Forest[38,39] (RF; see Materials and Methods).

Decoding accuracy and F1-scores exceeded 90% for both algorithms when utilizing the full

network excluding TC neurons (Fig. 3b; Supplementary Fig. 7a), with a positive correlation between F1-

scores and the number of trials that elicited spikes in TCs (Supplementary Fig. 7d). Importantly,

between 10–70% of best decoding neurons needed removal for accuracy to drop to chance level (Fig.

3c). In a representative experiment (Fig. 3d), F1-scores of TCs dropped to chance levels after removing

>70% of the best-encoding neurons. On average, 20–30% of the best-encoding neurons were sufficient

to achieve high decoding accuracy, regardless of the algorithm used (Fig. 3e; Supplementary Fig. 7c).

Fig. 3 | Information on the origin of novel spikes propagates widely and transiently in superficial layers 2/3 of the awake and resting mouse. a Example field of view with targeted TCs (*circled*) and network neighborhood (mean baseline activity). **b** High decoding accuracy (*violin plot*) and F1 scores (*box plot*) for identifying TC origin from network responses using XGBoost (XGB) or Random Forest (RF; n = 7 experiments; chance level 10 – 14%). **c** Removing 10–70% of best decoding neurons (ranked by Shapley values) drops accuracy to chance (*dashed lines*; A – G experiments; *brackets*: # TCs). **d** Distributed TC spike information: up to 100 top decoding neurons must be removed to reach chance (single experiment). F1-score of individual TC and probability of removing a PosR (*purple area*). **e** Mean drop in F1-score across experiments (XGB). *Inset*: Average percentage of best decoding neurons removal to reach chance. **f** Radius-dependent exclusion confirms spatially restricted TC stimulation, excluding neighbors. *Top*: F1 drop vs. radial exclusion from TCs (single experiment). *Bottom*: ΔF1 (*black*) summarizes probability of F1 changes >3σ (SD: *grey*) and probability of neuron drop (*red*; all experiments). **g** Summary of likelihood, L of dropped neuron being PosR for all experiments. **h** Individual neurons can qualify as PosR to multiple TCs. Average probability distribution for the number of TCs a neuron is categorized as PosR across experiments (*lines*). **i** Novel spike Information is transient: decoding success exceeds chance (*dashed lines*) only when the response window (*pink areas*) overlaps with the stimulation period (0– 100 ms).

Positive spike count changes (PosR; α = 2.5%) contributed significantly, accounting for 30–60%

of decoding success (Fig. 3d, g). As a further control, we performed a radius exclusion analysis which

demonstrated direct stimulation of non-TC cells did not influence decoding statistics. Excluding neurons

within <10 μm of a TC had no effect on performance, but stepwise removal of neighboring non-TC

neurons progressively reduced decoding accuracy (Fig. 3f).

Individual neurons responded significantly to the stimulation of up to five different TCs,

suggesting that they are capable of contributing to the decoding of multiple TCs, with significant overlap

in the functional networks that encode where novel spikes originate (Fig. 3h). Information about these

novel spikes faded quickly once the stimulus ended (Fig. 3i), aligning with the short-lived response

profile in all neurons (cf. Fig. 1c). The increased decoding accuracy before stimulation can be attributed

to temporal averaging (132 ms) and the use of Deep-Interpolation[40], which reduces noise to improve

accuracy but slightly broadens the time of responses without impacting the overall results

(Supplementary Fig. 7b).

These findings demonstrate that the cortical superficial layers can reliably and briefly transmit

precise information about the timing and source of novel spikes, highlighting a robust, distributed

mechanism for processing unexpected inputs across the neural network.

Novel spikes scale network responses during the fluctuation dominated regime of ongoing

neuronal avalanches

For many neurons, the spike count in encoding of novel spikes showed higher variability than expected

from a Poisson process (Fig. 1), with Fano factors similar to those observed during ongoing neuronal

activity (Fig. 2). This variability extended to the population of subclasses, such as PosR. To better

understand this fluctuation-dominated activity regime, we studied its statistical properties in relation to

novel spike encoding.

Holographic stimulation occurred during large fluctuations in ongoing activity (Fig. 4a),

measured as continuous population activity above a small threshold. These suprathreshold periods

exhibited heavy-tail distributions in size (summed spikes), and duration as described previously for

neuronal avalanches[29,41] (Fig. 4b). Notably, when correcting for subsampling biases using temporal

coarse graining, mean avalanche size scaled quadratically with avalanche duration – a hallmark of

parabolic avalanches – which includes avalanches lasting up to ~0.7 s, accounting for 87% of all

avalanches and ~20% of all spikes during ongoing activity (Fig. 4c, d). Avalanches exceeding >1 s

exhibited linear scaling close to 1, representing activity too large to resolve within our field of view[29,41,42].

Notably, novel spikes influenced the timing of ongoing avalanche initiation and termination, due to

added population activity during stimulation (Fig. 4e).

Fig. 4 | High decoding accuracy of novel spikes occur during scale-free avalanche dynamics despite highly fluctuating levels of baseline activity. a *Top*: Population activity in visual cortex of the awake mouse exhibits large fluctuations during quiet resting. *Bottom*: Zoomed in view of population activity in response to 3 (i – iii) holographic stimulations of a single TC. *Vertical lines*: TC stimulation. *Red line*: TC spike count. *Grey stripe*: Population threshold for avalanche detection. Shown is the population spike count (excluding TCs) per 22-ms frame. **b** Suprathreshold contiguous fluctuations in population activity exhibit the characteristic hallmarks of neuronal avalanches. Avalanche size (*left*) and duration (*right*) distributions as a function of the temporal coarse-graining factor k. **c** Scaling of avalanche dynamics as a function of temporal coarse-graining k. **d** Approximately 87% of avalanches are parabolic in nature, with durations <0.7 s, containing about 18% of thresholded spikes. Summary scaling of parabolic avalanches is shown for durations up to ~0.7 s ($k = 4$). **e** Holographic stimulation influences avalanche initiation, with a decreased likelihood for avalanche start and end during stimulation. **f** Mean response spike count scaling as a function of mean baseline spike count for non-TC neurons. This analysis reveals a low baseline activity regime (mean baseline spike count < ~0.3), in which responses are mostly independent from base activity, distinct from a high baseline activity regime in which response activity can be predicted from base. **g** Mean avalanche duration is more than double for avalanches that start during LOW base trials compared to HIGH. 87% of all LOW avalanches are parabolic (scaling of $\chi = 2$), whereas only ~52% of HIGH avalanches are parabolic. **h** Decoding accuracy remains largely unchanged when obtained using only LOW or HIGH baseline activity regimes, regardless of the threshold used to define the two regimes. Overall accuracy for Low base trials (*solid lines*) and High base trials (*dashed lines*) for each dataset (same color code as in Fig. 3), as function of the threshold used to define Low and High base trials.

Decoding of novel spikes remained remarkably robust across the avalanche scaling regime
despite large fluctuations in baseline activities. For baseline activity below a threshold (Ө < 0.3
spikes/132 ms), response spike count scaled only moderately, while for Ө > 0.3 spikes/132 ms, the
scaling approached 1, the latter indicating strong correlation between pre-stimulus and stimulus spike
counts (Fig. 4f). This enabled classification of baseline activity into a low- and high-activity regime
respectively which prominently featured parabolic avalanches (Fig. 4g). Importantly, novel spike
encoding differed by less than 4% in accuracy on average between these two regimes independent of
Ө (Fig. 4h, Supplementary Fig. 7e).

Discussion

Efficient communication of novel neural signals is vital for processing surprises and guiding
future actions. Using an all-optical approach in awake, resting mice, we demonstrated that
holographically introducing action potentials in a single L2/3 pyramidal neuron rapidly activates
widespread responses in the surrounding cortical network, encoding the novel stimulus' origin.
Remarkably, these responses scaled robustly and were efficiently decoded amidst ongoing, large
fluctuations in network activity organized as scale-invariant, parabolic neuronal avalanches. This high
susceptibility to local perturbations[27,28], with a scaling exponent of ~0.2–0.3, highlights critical dynamics
as a cornerstone of cortical processing, offering key insights into brain function.

Our findings challenge a longstanding discussion on cortical coding fueled by the intricate
statistical anatomy of the cortex and its synaptic connections. Pyramidal neurons in associative layers
rarely establish synaptic connections with their neighbors[19-23]. These excitatory connections are
relatively weak, prone to failure, and decay within milliseconds[22,43]. This combination of sparse and
weak connectivity makes it unlikely that a spike in a single neuron could propagate reliably in the
surround network. To address this, most cortical codes rely on activity patterns that amplify synaptic
impact on postsynaptic pyramidal neurons.

For instance, coincident activation of numerous synapses from diverse neurons converging onto
a postsynaptic target enhances the likelihood of activity propagation, favoring coding models based on
local synchrony, such as synfire chains[44,45], waves[46,47] or phase-locked oscillations[48,49]. Alternatively,
repeated or burst firing of individual neurons[12,50], strong synaptic connections[51], subthreshold priming
by background activity[15,52] supported by subcellular amplification mechanisms such as short-term
synaptic facilitation[53,54] and non-linear dendritic processing[55] can dynamical increase individual
connection strength supporting rate-based coding.

Our findings do not support a dichotomic separation into synchrony vs. firing rate separation in
the propagation of neuronal activity (see also [24,56]). We show highly efficient scaling of novel spikes
within the cortical neighborhood from single to multiple action potentials delivered within 100 ms. This
suggests that neither tight synchrony nor repeat firing is strictly necessary for effective encoding of
novel spikes. This conclusion is in line with findings that spike counts over 100 ms predict behavior
independent of considerations of synchrony at the millisecond level[57]. By emphasizing the sufficiency of
a few action potentials in cortical codes, our results thus challenge existing paradigms and highlight a
minimalist yet efficient mechanism for cortical information processing supporting the prominence of
sparse, single action potential firing found in vivo for associative layers[58].

Our findings demonstrate the remarkable impact of just a few action potentials amid the
thousands of spikes generated every second in the awake brain's cortical neighborhood. In a cortical
volume with a neuronal density of ~100,000 neurons per cubic mm and a connectivity radius of 0.2
mm[19], the local spherical neighborhood of ~3,300 neurons exhibiting a mean firing rate of ~1 Hz at rest
(*cf*. Fig. 1d) produces approximately 3,300 spikes per second that could influence a single neuron's
state. This apparent spike abundance might suggest that a few added action potentials have negligible
influence. However, our results challenge this assumption, showing that introducing 1–7 novel spikes in
a randomly chosen, single pyramidal neuron, not synchronized with neighboring activity, can have
extensive spatial reach and convey significant information, even within the highly fluctuating activity of
the ongoing network. Given our 0.45×0.45 mm field of view and ~170 imaged neurons (*cf*. Fig. 2i), we
sample ~5% of this neighborhood (<170/3300). At a conservative significance level of $\alpha = 2.5\%$, we
observed ~23 PosR per TC (1221/54; Fig. 1d), which translates to more than 450 neighboring neurons
significantly *increasing* their firing within tens of milliseconds in response to our holographic stimulation.
This estimate, however, is likely highly conservative as lowering alpha further increases the number of
neighboring neurons with detectable changes in spiking, as confirmed by our decoding results. Notably,
reducing α, also reveals a role for local inhibition in intracortical circuits[25,26]. This adjustment
increasingly recruits both positive and negative responses from the pool of previously non-significant
responders. These recruitments are most likely composed of a mixture of direct (excitation, inhibition)
as well as indirect (disfacilitation, disinhibition) local circuit interactions[4,7], while preserving the general
scaling relationship (cf. Fig. 2). Simulations have found that in such a fluctuation-dominated but
relatively balanced regime, a reduction in membrane conductance due to increased synaptic
bombardment still enables neurons to generate few spikes in response to rapid changes in synaptic
input[17].

The importance of critical dynamics for sensory cortical networks has been argued from the
point of view of optimal decoding of stimulus intensities through maximization of the dynamic range[59]. In
that context, a critical network is expected to produce response-input curves that scale as a power law,
with exponents that are typically much smaller than one[28]. The study of psychophysics has
demonstrated that psychological value scales with stimulus intensity as a power law with <1 exponents,
a result known as Stevens' law of psychophysics[60]. Therefore, it has been argued that the origins of
Steven's low exponents for response curves might originate precisely at this critical networks' ability to
amplify low-intensity stimuli and thus respond to a wide range of stimulus intensities[61]. We were able to
show that *in vivo* sensory cortical networks do display low-exponent, power-law behavior for their
response to varying stimuli intensity (as proxied by the number of spikes evoked in a target cell during
stimulation). This is in line with predictions from critical dynamics, with the observed exponents closely
matching the expected value (0.285) for a two-dimensional directed percolation system[28].

Our decoding analysis provided important insights about the fine-grained nature of this
communication of novelty. First, strongly decoding neurons were found in the immediate neighborhood
of TCs supporting a spatial weighing in the response, in line with expectations from the local, synaptic
connectivity statistics of pyramidal neurons. On the other hand, a large fraction of neurons with high
decoding contribution did not fall into the PosR category or were located at long distances from the TC,
supporting the contribution of indirect pathways. This also de-emphasizes reliance on distinct, non-
overlapping network connectivity.

The ability of PosR neurons to contribute significantly to the decoding of multiple TCs is in line
with the spatial overlap of neuronal populations encoding different stimuli and their temporal separation.
Our findings align with the notion that cortical networks are organized to maximize coding efficiency and
redundancy, where overlapping populations can fluctuate in their response patterns to encode different
stimuli[62,63]. The rapid decay found for the encoding of novel spikes further underscores this dynamic,
ensuring that novel signals do not disrupt ongoing network activity but still convey novel information
effectively within the constraints of the temporal and spatial coding architecture described for e.g. visual
stimuli of the primary visual cortex in the awake mouse[64].

Our findings of maintained high Fano factor in evoked responses is in contrast to quenched
variability demonstrated in response to sensory stimuli[34]. This difference may arise because single-
neuron stimulation introduces a smaller perturbation than sensory stimulation, which in the awake state
is dominated by inhibition[25], affecting response variability. Early studies in anesthetized animals[18] (e.g.
in the visual cortex of cats) demonstrated that trial-by-trial variability in stimulus responses can be
approximated as average response overlaid on a correlated, variable background. In the awake state,

our observation of consistently high FF values for responses in the presence of scale-invariant
parabolic avalanches indicates the need for a conceptual approach that bypasses the reliance on
averaging. Avalanches emerge in superficial layers of cortex when animals recover from
anesthesia[65,66], exhibit high neuronal selectivity[67,68] and spatiotemporal order[69-73] characterized by
power law statistics with exponents ≤2 that is, averages are not defined and variability is high. We
demonstrate that decoding of novel spikes is similar across the fluctuation dominated avalanche regime
(Fig. 4).

Our results reveal the efficient communication of single spikes within the scale-invariant, critical-
state dynamics of cortical networks, where synchronized fluctuations are neither overly constrained nor
entirely random. We propose this remarkable sensitivity of the mammalian cortex, to efficiently
disseminate novel information from individual neurons, underscores the fundamental role of critical
dynamics in cortical communication and computation.

Materials and Methods

Animal surgery

All procedures were approved by the NIH Animal Care and Use Committee (ACUC) and experiments
followed the NIH *Guide for the Care and Use of Laboratory Animals*. Mice were obtained from Jackson
labs, bred inhouse with C57BL/6 backgrounds (Jackson Laboratory) under a reversed 12:12 hr
light/dark cycle. Chronic 2PI imaging in adult (>6 weeks) mice was enabled by using a head bar in
combination with a cranial window placed centered at ~2.5 mm from the midline (right hemisphere) and
~1 mm rostral to the lambdoid suture and consisting of a stack of circular glass cover slips using
established protocols[74]. Mice were injected with a 7:1 mixture of a viral construct to express
jGCaMP7s[75] and ChrimsonR[76] in pyramidal neurons using the CaMKII promotor. A total of 2 − 3
injections of virus (100 − 400 nL; <1 μL in total; 10^{13} vg/mL; pGP-AAV-syn-jGCaMP7s-WPRE AAV9 &
pAAV-CamKIIa-ChrimsonR-mScarlet-KV2.1, Addgene) were administered into the right hemisphere.
Chronic 2PI started after >2 weeks in identified V1 at a depth of ~100–200 μm.

Identification of V1 maps

Retinotopic maps of V1 and higher visual areas (HVAs) were generated for all mice prior to recording
using published protocols[77,78]. Briefly, awake, head-fixed mice faced with their left eye a 19" LCD
monitor placed at 10 cm distance and tilted 30º towards the mouse's midline. Using Psychophysics
toolbox[79], contrast-reversing, spherically corrected checkerboard bars were drifted across the screen
vertically (altitude) and horizontally (azimuth) for each of the four directions (30 repeats per direction).

Simultaneous wide-field imaging (Quantalux, Thorlabs) captured jGCaMP7s fluorescence, which was

averaged for each direction. Altitude and azimuth phase maps were calculated by phase-wrapping the

first harmonics of the 1D Fourier transform for each of the four averages and subsequently subtracting

the maps of the opposite directions[78]. Sign maps were generated by taking the sine of the angle

between the gradients in the altitude and azimuth maps and processed[77]. Borders were drawn around

visual area patches and overlaid onto anatomical reference images to identify V1.

Holographic stimulation of single neurons

Holographic stimulation[80] was done by manually selecting individual pyramidal neurons as target from

the two-photon image of the field of view in both red and green channels. Neurons that express the

opsin (ChrimsonR, with mScarlet tag) were selected for stimulation. Based on the selected targets, a

map of top-hat patterns with diameter of 10 μm was used to generate the hologram on the Spatial Light

Modulator (SLM, Meadowlark, LCoS high resolution 1920×1152 phase modulator). Light from the

laser/OPA system (Carbide/Orpheus-F – Light Conversion, λ = 1064 nm, repetition rate = 600 kHz,

pulse width <100 fs) was expanded (4x) to fill the area of the SLM (8 mm × 8 mm). The hologram on

the pupil plane was shrunk to fit the size of the Galvo-Galvo scanners (~3 mm diameter) and form the

desired pattern of top hats after the objective lens (Nikon 16X, working distance = 3mm and effective

focal length = 12.5 mm) on the selected targets. Each target neuron was continuously stimulated by the

top-hat beam shape for 100 ms for a total of ~100—150 trials per experiment, with an interval of 2 s (10

s in some recordings) between trials. Power at target neurons ranged from ~2.5 – 10 mW.

2PI imaging, pre-processing pipeline and meta data collection

For standard 2PI, images were acquired by a scanning microscope (Bergamo II series, B248, Thorlabs

Inc.) coupled to a pulsed femtosecond Ti:Sapphire 2-photon laser with dispersion compensation

(Chameleon Discovery NX, Coherent Inc.). The microscope was controlled by ThorImageLS and

ThorSync software (Thorlabs Inc.). The wavelength was tuned to 940 nm to excite jGCaMP7s. Signals

were collected through a 16× 0.8 NA microscope objective (Nikon). Emitted photons were collected

through a 525/50 nm band filter using GaAsP photomultiplier tubes. The field of view was ~450×450

μm. Imaging frames of 512×512 pixels were acquired at 45.527 Hz by bidirectional scanning of a 12

kHz Galvo-resonant scanner. Beam turnarounds at the edges of the image were blanked with a

Pockels cell. The average power for imaging was <70 mW, measured at the sample.

The obtained tif-movies in uint16 format were rigid motion-corrected via the python-based

software package '*suite2p*'[81]. Registered images were further denoised using machine-learning based,

deep interpolation[40] (see below) and then semi-automatically processed by suite2p for ROI selection

and fluorescence signal extraction. For each labeled neuron, raw soma and neuropil fluorescence

signals over time were extracted for each ROI. Spiking probabilities were obtained from neuropil-

corrected fluorescence traces ($F_{corrected} = F_{ROI} - 0.7*F_{neuropil}$) via MLspike (https://github.com/MLspike),

utilizing its autocalibration feature to obtain unitary spike event amplitude, decay time, and channel

noise for individual ROIs.

Deep-interpolation

Deep-interpolation[40] (Deep-IP; https://github.com/AllenInstitute/deepinterpolation) removes independent

noise by using local spatiotemporal data across a noisy image stack of N_{pre} + N_{post} frames to predict, or

interpolate, pixel intensity values throughout a single withheld central frame. The deep neural network

is a nonlinear interpolation model based on a UNet inspired encoder-decoder architecture with 2D

convolutional layers where training and validation are performed on noisy images without the need for

ground truth data. As described previously in detail[68], after rigid motion correction, individual denoised

frames were obtained by streaming one 60-frame (N_{pre} = N_{post} = 30 frames) registered, image stack

through the provided Ai-93 pretrained model[40] for each frame to be interpolated. At an imaging rate of

~45 Hz, these 60 frames correspond to a combined ~1.3 seconds of data surrounding the frame to be

interpolated.

Postprocessing pipeline

Our postprocessing pipelines were custom-written in Matlab (Mathworks) and Python

(www.python.org). Some routines utilized NumPy (https://numpy.org/) and Matplotlib

(https://matplotlib.org/).

Stimulus response measurement

Unless stated otherwise, spike count over a window of 6 imaging frames (~132 ms) starting at the

stimulation onset was used as a measure of the response to holographic stimulus of each target cell for

all cells. This was compared to the baseline count, defined as a 6-frame window starting ~264 ms

before stimulus onset. Due to difficulties and limitations of spike deconvolution of calcium traces, some

parts of the time series resulted in artificially high spike rates. To avoid artificial increase of variability

measures due to these artifacts, we introduced an outlier removal procedure in which trials with spike

count above 10 standard deviations over the mean are removed from the analysis. This procedure

resulted in <0.1% trials removed, across all neurons for all experiments.

Target selection and exclusion

Up to 25 target cells were visually selected based on expression and location in the field of view (cells

close to the edges were avoided) for each experiment. Not all targets were responsive to stimulation.

Targets in which fewer than 20% of the trials evoked a spike count above the 91% percentile of its

baseline count were excluded from the analysis. This procedure resulted in 54 target cells across 16

experiments from 6 mice being considered out of 193 targets attempted across 33 experiments from 11

mice. We estimate an average of 46.3 ± 17.9 ChrimsonR-labeled cells per field of view, which

translates to an average density of labeled cells of ~$2.3×10^{-4}$ cells per μm^2. This sparse labeling

facilitates the holographic stimulation of individual targets (see Supplementary Fig. 3).

Definition of significant responders

Cells with significant response to each target's stimulation were defined as those whose spike count

distribution was significantly different from its baseline count distribution. Specifically, we used a z-test

to compare the stimulus response spike count to the average and standard deviation of the baseline

count for each non-target cell. A cell was defined as a positive responder if the z statistic is above the

highest 2.5% of the distribution and as a negative responder if the z statistic is below the lowest 2.5%.

Otherwise, the cell is classified as a non-significant responder. This procedure is done for each target

stimulated in an experiment separately. Therefore, in experiments where multiple targets were

selected, each non-target cell is tested for significance of its response to each target's stimulation,

possibly being a significant responder for some targets and non-significant responder to others (see

Fig. 3h).

Fano Factor calculation and significance

Fano Factor was calculated as usual, for each cell separately, by dividing the variance of their spike

count by the mean of their spike count across trials. To correct for changes in firing rate across

conditions[35], we introduced a procedure in which the window used to measure the spike count during

stimulation varied from 1 to 9 frames (~22 ms to 198 ms), starting at the onset. The window length was

determined by minimizing the normalized difference between the mean baseline and stimulus counts.

Cells for which this difference was above 20% were excluded from the analysis. This procedure

resulted in less than 30% of cells removed, on average, across the non-target cells and an average of

65% of targets removed (minimum of 13% or 7 targets considered). Values were reported as mean ±

error over the mean across cells, unless stated otherwise. Significance of Fano Factor changes across

conditions was obtained by employing the Wilcoxon rank sum test at a significance level of 0.05 (p-

values reported in Results section and figure legends).

Trial separation and controls

To control for different conditions affecting the variability, the Fano Factor was recalculated and

compared using a subset of trials. To control for target cell response variability affecting non-target

cells, we separated trials into low target count trials and high target count trials. We defined low count

trials as those for which the target cell produced a spike count during stimulation that was above the

91st percentile of the baseline count distribution and up to twice that number of spikes, whereas high

count trials were any trial with count above that limit. To control for the influence of ongoing fluctuation

of the network activity we separated trials into low and high population baseline count, defined as trials

for which the population count during baseline was lower than the 25th percentile of its distribution or

higher than 75th percentile, respectively. To control for temporal influence of the responses to the

stimulus, we defined two consecutive 6-frame windows, centered at the peak of the target's average

response to the stimulus (2—3 frames after onset), as early and late response windows.

Correlation analysis

Spike counts across trials were compared by computing correlations across time (autocorrelation,

calculated using the spike counts over pairs of windows separated by a temporal lag of 132 ms) and

across cells (comparing spike counts of targets *vs* non-targets). Values are reported as mean ± error

over the mean. Significance of correlation changes across conditions was obtained by employing the

Wilcoxon rank sum test at a significance level of 0.05 (p-values reported in Results section and figure

legends).

Decoding Analysis

Classification model selection, performance metrics, and data preprocessing

In our prediction and classification scheme, we used target neuron ensembles as "classes" and the

activity of the responding/non-target neurons as the "features". As described in Stimulus response

measurement section, we use spike count over a window of 6 imaging frames as a feature value for

each neuron. For the temporal analysis, this window is offset relative to stimulus onset, as described

below. In a preliminary analysis, we compared several machine learning strategies and found that the

ensemble methods of boosting (XGBoost[37]) and bagging (scikit-learn[82] Random Forest classifier)

performed best to several other common classification schemes such as Logistic Regression, Naive

Bayes, Neural Networks, and Support Vector Machines. These two approaches yielded the highest

accuracy and were then applied to all of our analyses presented in this work. In all cases we split the

data samples randomly so that 80% was used for training and 20% for testing. We repeat this split 100

times and from these report the mean and standard errors of our metrics. As metrics we use overall

accuracy, per-class recall and precision, as well as F1-scores (harmonic mean of recall and precision).

For the results in Figs. 3c, h we use all data and all trials, but for others we performed additional pre-

processing, before running classification.

Since not all trials were equally successful in eliciting spikes from target cells, for all other

decoding analysis we selectively filtered trials and targets, with a criterion that the target neuron had to

produce at least one spike across any of the six frames after the onset of stimulation. We then selected

valid targets as those that spiked on at least half of their trials. Thus, non-significantly driven TC (based

on conservative response spike count) excluded some TC in the spike count analysis which

nevertheless had significantly F1-score in the decoding task (gray TC in Supplementary Fig. 7a). To

characterize feature importances and interpret our classification schemes we used Shapley analysis[83],

as described below.

Radius Analysis

For each target neuron, we implemented an exclusion zone based on the distance from the target

neuron's soma, *r*. We progressively increased *r*, and the neurons/features falling within this distance

were systematically eliminated from the feature matrix, enabling quantification of the relationship

between F1-score degradation and exclusion zone radius for individual target neurons. We computed

the probability distribution of any feature neuron being excluded, as a function of distance, $P_{ex}(r)$.

Specifically, we aggregated the frequency of neuron eliminations at each radius across all datasets and

normalized by the total number of exclusions.

The spatial dependency of predictive performance was characterized by analyzing the

differential changes in F1-score across the range of exclusion radii. For each target neuron across all

datasets, we identified the radius at which the decrease in F1-score first exceeded three standard

deviations below the mean rate of change. These radii were averaged across datasets and smoothed

using a Gaussian kernel ($\sigma = 1$) to obtain an estimate of the spatial scale of predictive influence.

Neuron Drop Analysis and Positive Responder Probability

To evaluate the distributed nature of neural information and assess the robustness of our decoding

framework, we conducted an iterative feature elimination analysis for each target neuron across all

datasets. For each iteration, including initial model fitting, we use Shapley values (employed using

SHAP package: SHapley Additive exPlanations[84] to quantitatively rank feature neurons based on their

predictive contribution for a given target neuron. Through sequential elimination of the highest-ranked

feature neurons and subsequent model retraining, we systematically assessed how the removal of

informative neurons impacted decoding performance. This iterative process of SHAP-guided feature

elimination and model reconstruction was repeated until a single neuron remained. At each iteration we

recorded F1-score, precision, and overall accuracy and the entire procedure was replicated one

hundred times, utilizing different out-of-bag random selection of training and testing sets.

To characterize the relationship between feature importance and significant responders, we

quantified the probability that eliminated neurons exhibited significant responses. For each target

neuron, we constructed a binary vector corresponding to the sequential elimination order, where entries

were assigned values of 1 or 0 based on whether the eliminated neuron was a significant responder.

These binary sequences were averaged across all iterations of the elimination procedure to obtain

target-specific probabilities and then averaged across all target neurons.

Temporal profiles of the prediction accuracy

To further characterize our holographic stimulation data we explored the accuracy of our prediction

scheme as a function of time. This procedure allowed us to study how long the information about the

origin of stimulation persists in network. Note that the prediction scheme is a more stringent

requirement than studying only the causal effects of stimulation, since even though the activity after the

stimulation can reverberate in a network for a long time, identifying the node that initiated it is more

challenging.

We assessed the temporal course of the prediction accuracy across datasets, each with

different levels of responsiveness of the targets. We varied the size of our response window (2 through

8 frames) and then moved it in steps of 2. The main results are present in Figure 3h. The fading "pink"

rectangles indicate the amount of data used from the stimulation region (100 ms) window, so that the

fully white regions use data strictly before or after stimulation is over. Note that the bins and the

stimulation region are not perfectly aligned, as the trigger for the stimulation was independent from the

image scanning raster. We define the reference bin 0 as the first bin which fully overlaps with the

stimulation.

Avalanche analysis

Continuous epochs of suprathreshold population activity

Continuous periods of population activity were identified by applying a threshold Θ on the population

activity $p(t)$, the sum of the spikes from all neurons at a given time t, such that:

553
$$p_\Theta(t) = \begin{cases} p(t) - \Theta & p(t) > \Theta \\ 0 & p(t) \leq \Theta \end{cases} \tag{1}$$

This procedure is known as soft-thresholding and was employed for all analysis unless otherwise

stated. For a given recording $p(t)$ and coarse-graining value k, the dependence of the number N of

epochs on the threshold Θ, N(Θ) was obtained for a range of thresholds $\Theta \in [\Theta_1, \Theta_2]$ such that Θ_1 was

low enough that it removed no population activity from the time course and Θ_2 was high enough that it

would remove all population activity from the time course. The function N(Θ) was typically well-

approximated by a log-normal distribution and a corresponding fit yielded shape parameters μ and σ.

The threshold used in the analysis for all recordings and coarse-graining factors was chosen such that

$\Theta = \mu$ and estimated for each k. See Capek & Ribeiro et al.[68] for more details.[30,69,85,86]

Temporal coarse-graining

A temporal coarse-graining operation was applied to the thresholded population activity $p_\Theta(t)$. For a

given temporal coarse-graining factor k an ensemble of K different coarse-grained time series

$p_k^0(\tau), p_k^1(\tau), ..., p_k^{K-1}(\tau)$ was arrived at through the following method:

$$p_k^j(\tau) = \sum_{i=k\tau+j}^{k(\tau+1)+j-1} p_\Theta(i) \quad for\ \tau \in [0, 1, ..., \left\lfloor \frac{T-j}{k} \right\rfloor] \tag{2}$$

For each time series $p_k^j(\tau)$, epochs were extracted by finding pairs (τ_1, τ_2) such that $p_k^j(\tau_1) = 0$,

$p_k^j(\tau_2) = 0$ and $p_k^j(\tau') > 0$ for all $\tau' \in [\tau_1 + 1, ..., \tau_2 - 1]$. The size of the epoch is given by $S = \sum_{i=\tau_1}^{\tau_2} p_k^j(i)$

and its corresponding duration given by $\tau_2 - \tau_1 - 1$. For a given ensemble of coarse-grained time

series, all epochs were combined.

Scaling curve fit

For more precise evaluation of χ_{sh} and χ_{lg}, we introduced the following fitting function:

$$S(d) = \frac{C\,d^{\chi_{sh}}}{(1+(d/\Phi)^\gamma)^{(\chi_{sh}+\chi_{lg})/\gamma}} \tag{3}$$

This function is a double power law with initial slope χ_{sh}, transitioning to a second slope χ_{lg} at around the

point $d = F$. The parameter g controls how abruptly that transition happens and has been fixed at 4 for

all the curves presented. The other parameters were free to adjust to the data, and all fits were

performed in log-space, i.e., the log(S(log(d)) was fit to the log of data (taking the log of both average

sizes as well as durations).

Data availability

The preprocessed imaging data used in this study are available in the general repository Zenodo using

the following access DOI:TBD (Ref. [87]). The source data for all figures and supplemental figures in this

study are provided for this paper.
586

Code availability

588 Computer code used in this study is available at the following github link: TBD.
589

Acknowledgements

591 We thank Craig C. Stewart and members of the Plenz lab for help with animal surgery, animal care. This
592 research was supported by the Division of the Intramural Research Program (DIRP) of the National
593 Institute of Mental Health (NIMH), USA, ZIAMH002797, ZIAMH002971, and the BRAIN initiative Grant
594 U19 NS107464-01. This research utilized the supercomputing resources of the National Institutes of
595 Health (NIH, USA; Biowulf, http://hpc.nih.gov).
596

Author Contributions

598 TLR and DP conceived and planned the study; TLR, BG, and VS performed experiments; AV took the
599 lead in holographic setup design. TLR took the lead in data analysis with support from BG and VS. SP
600 and RS took the lead on machine learning based decoding of experimental data. All authors contributed
601 to the analyses. TLR and DP wrote the manuscript.
602

Competing interests

604 The authors declare no competing interests.
605

606 **Supplementary Figures**

607 **Supplementary Figure 1**

Supplementary Fig. 1 | All optical interrogation set-up to image and perturb select neuronal populations with cellular specificity. **a** Schematics of all-optical holographic stimulation combined with 2-photon imaging path (DM: dichroic mirror, M: Mirror, F: filter, PMT: photomultiplier, L: lens). **b** Mean 2-photon image in L2/3 of V1 with neurons co-expressing jGCaMP7s and ChrimsonR allowing for holographic stimulation of a single pyramidal neurons while reconstructing spike activity in the surround network using 2-PI. **c** Widefield expression image of GCaMP7s and CrimsonR across the craniotomy and corresponding merged yellow combined picture indicating co-expression (*single mouse*). **d** Co-expression of GCaMP7s with ChrimsonR pyramidal neurons over changing the titer ratio of the two corresponding virus. N = 23 mice.

611 **Supplementary Figure 2**

Supplementary Fig. 2 | Schematic design and evaluation of the stimulation path using spatial light modulation. a Light from the laser (Light Conversion, λ = 1064 nm, repetition rate = 600 kHz, pulse width <100 fs) is expanded (4x) to fill the area of the spatial light modulator (SLM; Meadowlark, LCoS high resolution 1920×1152 phase modulator). The hologram on the pupil plane is shrunk to fit the size of the Galvo-Galvo scanners (~3 mm diameter) that guide the stimulation beam in the focal plane of image (Nikon 16X, working distance = 3 mm and effective focal length = 12.5 mm). Neurons are manually selected from the mean 2-PI image of the field of view in both red and green channels. Target cells (TC) that express both the opsin (ChrimsonR, with tdTomato tag) and GCaMP7s are selected for stimulation and simultaneous estimation of the evoked spike count. Based on the selected targets, a map of top-hat patterns with diameter of 10 μm were used to generate the hologram on the SLM. Optical power per TC was adjusted to 5 — 10 mW, depending on the depth and expression level. **b** Spiking above base line increases with stimulation power (N = 3 TC; 100 ms stimulation duration). **c** No difference was found in evoked spike count above baseline for 5 and 10 mW stimulation when averaging across all TC. **d** Schematics of reconstruction of the stimulation xy-profile using fluorescence microbeads. Galvo-Galvo (GG) controlled stimulation of gel-embedded microbeads (3 μm diameter) and simultaneous Galvo-Resonance (GR) 2-PI of elicited response. *Circles*: Individual microbeads before and during stimulation. **e** Experimentally reconstructed and smoothed xy-profile of a flat-hat stimulation (10 μm diameter; 5 mW) at 4 different positions within the field of view at center (0,0), off-x (+100 μm, 0), off-y (0, +100 μm), and off-xy (+100 μm, +100 μm). *Inset*: Non-smoothed reconstruction.

613 **Supplementary Figure 3**

Supplementary Fig. 3 | Summary in successfully isolating individual pyramidal neurons in layer 2/3 of primary visual cortex during rest using holographic stimulation combined with 2-PI. For each TC included in the analysis the mean luminance after motion correction and denoising (*left*) and the ΔF/F in response to stimulation for each pixel in a 50×50-pixel area centered on the TC (*right*) are shown.

614
615

616 **Supplementary Figure 4**

Supplementary Fig. 4 | Spike count in all different subpopulations scale with spike count of target cells during baseline activity, but only Positive Responders scale during stimulation. a Mean spike count on PosR (*blue*), NonR (*grey*) and NegR (*purple*) as function of spike count of TC, in log-log, during baseline. Power law fits are indicated by solid lines, with obtained exponents shown. *Left*: mean calculated per neuron. *Right*: mean calculated after summing over the population. **b** Same as in a, but during holographic stimulation.

619   **Supplementary Figure 5**

Supplementary Fig. 5 | High response variability in PosR and NonR does not originate from variability in evoked spikes of TC. a Example of separating trials from one TC into those with a high spike count and low spike count. *Top*: Spike count distribution during baseline allows for the identification of 91% threshold above chance as well as twice that threshold. *Bottom*: Using the two thresholds from the baseline distributions categorizes stimulation trials of the TC into low and high target counts, respectively. **b** Summary statistics of mean TC trials separated by low and high target counts. Note the remaining high variance of PosR despite the large drop in TC variance. **c** FF does not change if trials within a TC are separated into high and low spike counts which changes the FF for TC.

621   **Supplementary Figure 6**

Supplementary Fig. 6 | Maintained variability during response is not due to change in preceding baseline activity and is robust throughout the response. No change in FF when subdividing into TC trials with high or low population spike count (**a**) or early vs. late spiking responses (**b**).

623 **Supplementary Figure 7**

Supplementary Fig. 7 | Decoding accuracy, stimulation efficacy, impact of deep interpolation.
a Multi-target cell stimulation experiments analyzed and corresponding break-down of TC performances. Overview of F1-scores for TCs (*green*) and TC that did not reach 20% of spike success (*gray*) upon stimulation. F1-scores are compared to chance level (*red bars*) and overall accuracy (*red*). vs. late spiking responses. XGBoost (*left*) and Random Forest (*right*) results are very similar across all datasets. **b** Deep interpolation slightly advances stimulus information and significantly increases accuracy thereby slightly prolonging transient response time. Exemplary impact for weak (A) and strong (D) data set of similar chance level (10 TC each). The enhanced accuracy achieved with Deep Interpolation demonstrates the method's noise-reducing benefits, which improve the model's ability to capture relevant neural signals. However, it also shows that it undermines our ability to study temporal progression of activations, as its window of smoothing ranges across 30 frames. **c** Z-scored F1-scores calculated for each TC and then averaged, for each dataset (same color code as in Fig. 3), as function of number of best predictors dropped. *Inset*: Average number of best predictors dropped before F1-scores reach chance levels. Random Forest method. **d** For each data set, stimulation success (at least 1 spike) varied among TC. Target stimulation success positively correlated with decoding power, i.e. F1-score, of a TC in RF and XGB. **e** Difference in accuracy for Low and High trials is small (with Low trials having slightly higher accuracy), regardless of the threshold used to separate them (cp. Fig. 4h).

References

Allen, W. E. *et al.* Global Representations of Goal-Directed Behavior in Distinct Cell Types of Mouse Neocortex. *Neuron* **94**, 891-907.e896 (2017). https://doi.org:10.1016/j.neuron.2017.04.017

London, M., Roth, A., Beeren, L., Häusser, M. & Latham, P. E. Sensitivity to perturbations in vivo implies high noise and suggests rate coding in cortex. *Nature* **466**, 123-127 (2010).

Meyer, J. F., Golshani, P. & Smirnakis, S. M. The Effect of Single Pyramidal Neuron Firing Within Layer 2/3 and Layer 4 in Mouse V1. *Frontiers in Neural Circuits* **12** (2018). https://doi.org:10.3389/fncir.2018.00029

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

Jouhanneau, J.-S., Kremkow, J. & Poulet, J. F. A. Single synaptic inputs drive high-precision action potentials in parvalbumin expressing GABA-ergic cortical neurons in vivo. *Nature communications* **9**, 1540 (2018). https://doi.org:10.1038/s41467-018-03995-2

Hemberger, M., Shein-Idelson, M., Pammer, L. & Laurent, G. Reliable sequential activation of neural assemblies by single pyramidal cells in a three-layered cortex. *Neuron* **104**, 353-369. e355 (2019).

O'Rawe, J. F. *et al.* Excitation creates a distributed pattern of cortical suppression due to varied recurrent input. *Neuron* **111**, 4086-4101.e4085 (2023). https://doi.org:10.1016/j.neuron.2023.09.010

Marshel, J. H. *et al.* Cortical layer–specific critical dynamics triggering perception. *Science*, eaaw5202 (2019). https://doi.org:10.1126/science.aaw5202

Carrillo-Reid, L., Han, S., Yang, W., Akrouh, A. & Yuste, R. Controlling Visually Guided Behavior by Holographic Recalling of Cortical Ensembles. *Cell* **178**, 447-457.e445 (2019). https://doi.org:10.1016/j.cell.2019.05.045

Houweling, A. R. & Brecht, M. Behavioural report of single neuron stimulation in somatosensory cortex. *Nature* **451**, 65-68 (2008). https://doi.org:10.1038/nature06447

Wolfe, J., Houweling, A. R. & Brecht, M. Sparse and powerful cortical spikes. *Curr. Opin. Neurobiol.* **20**, 306-312 (2010). https://doi.org:https://doi.org/10.1016/j.conb.2010.03.006

Cheng-yu, T. L., Poo, M.-m. & Dan, Y. Burst spiking of a single cortical neuron modifies global brain state. *Science* **324**, 643-646 (2009).

Zucca, S. *et al.* An inhibitory gate for state transition in cortex. *Elife* **6**, e26177 (2017).

Histed, M. H., Ni, A. M. & Maunsell, J. H. Insights into cortical mechanisms of behavior from microstimulation experiments. *Prog. Neurobiol.* **103**, 115-130 (2013). https://doi.org:10.1016/j.pneurobio.2012.01.006

Vaadia, E. *et al.* Dynamics of neuronal interactions in monkey cortex in relation to behavioural events. *Nature* **373**, 515-518 (1995).

Stringer, C. *et al.* Spontaneous behaviors drive multidimensional, brainwide activity. *Science* **364**, eaav7893 (2019). https://doi.org:doi.1126/science.aav7893

Kuhn, A., Aertsen, A. & Rotter, S. Neuronal integration of synaptic input in the fluctuation-driven regime. *J. Neurosci.* **24**, 2345-2356 (2004). https://doi.org:10.1523/JNEUROSCI.3349-03.2004

Arieli, A., Sterkin, A., Grinvald, A. & Aertsen, A. Dynamics of ongoing activity: Explanation of the large variability in evoked cortical responses. *Science* **273**, 1868-1871 (1996).

Braitenberg, V. & Schüz, A. *Anatomy of the Cortex: Statistics and Geometry*. (Springer-Verlag, 1991).

Thomson, A. M., West, D. C., Wang, Y. & Bannister, A. P. Synaptic connections and small circuits involving excitatory and inhibitory neurons in layers 2-5 of adult rat and cat neocortex: triple intracellular recordings and biocytin labelling in vitro. *Cereb. Cortex* **12**, 936-953 (2002).

Levy, R. B. & Reyes, A. D. Spatial Profile of Excitatory and Inhibitory Synaptic Connectivity in Mouse Primary Auditory Cortex. *J. Neurosci.* **32**, 5609 (2012). https://doi.org:10.1523/JNEUROSCI.5158-11.2012

Jouhanneau, J.-S., Kremkow, J., Dorrn, A. L. & Poulet, J. F. In vivo monosynaptic excitatory transmission between layer 2 cortical pyramidal neurons. *Cell reports* **13**, 2098-2106 (2015).

Campagnola, L. *et al.* Local connectivity and synaptic dynamics in mouse and human neocortex. *Science* **375**, eabj5861 (2022). https://doi.org:doi:10.1126/science.abj5861

Kumar, A., Rotter, S. & Aertsen, A. Spiking activity propagation in neuronal networks: reconciling different perspectives on neural coding. *Nat. Rev. Neurosci.* **11**, 615-627 (2010). https://doi.org:10.1038/nrn2886

Haider, B., Hausser, M. & Carandini, M. Inhibition dominates sensory responses in the awake cortex. *Nature* **493**, 97-100 (2013). https://doi.org:http://www.nature.com/nature/journal/v493/n7430/abs/nature11665.html#supplementary-information

Tremblay, R., Lee, S. & Rudy, B. GABAergic interneurons in the neocortex: From cellular properties to circuits. *Neuron* **91**, 260-292 (2016). https://doi.org:10.1016/j.neuron.2016.06.033

Chialvo, D. R., Cannas, S. A., Grigera, T. S., Martin, D. A. & Plenz, D. Controlling a complex system near its critical point via temporal correlations. *Scientific Reports* **10**, 12145 (2020). https://doi.org:10.1038/s41598-020-69154-0

Muñoz, M. A., Dickman, R., Vespignani, A. & Zapperi, S. Avalanche and spreading exponents in systems with absorbing states. *Phys. Rev. E* **59**, 6175-6179 (1999).

Capek, E. *et al.* Parabolic avalanche scaling in the synchronization of cortical cell assemblies. *Nature communications* **14**, 2555 (2023). https://doi.org:10.1101/2022.11.02.514938

Plenz, D. *et al.* Self-Organized Criticality in the Brain. *FrPhy* **9** (2021). https://doi.org:10.3389/fphy.2021.639389

Beggs, J. M. & Plenz, D. Neuronal avalanches in neocortical circuits. *J. Neurosci.* **23**, 11167-11177 (2003). https://doi.org:10.1523/JNEUROSCI.23-35-11167.2003

Softky, W. R. & Koch, C. The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. *J. Neurosci.* **13**, 334-350 (1993).

Holmgren, C., Harkany, T., Svennenfors, B. & Zilberter, Y. Pyramidal cell communication within local networks in layer 2/3 of rat neocortex. *The Journal of physiology* **551**, 139-153 (2003).

Churchland, M. M. *et al.* Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nat. Neurosci.* **13**, 369-378 (2010).

Rajdl, K., Lansky, P. & Kostal, L. Fano factor: a potentially useful information. *Frontiers in computational neuroscience* **14**, 569049 (2020).

Marro, J. & Dickman, R. Nonequilibrium phase transitions in lattice models. *Nonequilibrium Phase Transitions in Lattice Models* (2005).

Chen, T. & Guestrin, C. in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 785–794 (Association for Computing Machinery, San Francisco, California, USA, 2016).

Ho, T. K. in *Proceedings of 3rd international conference on document analysis and recognition.* 278-282 (IEEE).

39   Breiman, L. Random forests. *MLear* **45**, 5-32 (2001).
40   Lecoq, J. *et al.* Removing independent noise in systems neuroscience data using
DeepInterpolation. *Nat. Methods* (2021). https://doi.org/10.1038/s41592-021-01285-2
41   Srinivasan, K., Ribeiro, T. L., Kells, P. & Plenz, D. The recovery of parabolic avalanches in
spatially subsampled neuronal networks at criticality. *Scientific Reports* **14**, 19329 (2024).
https://doi.org/10.1038/s41598-024-70014-4
42   Yu, S., Klaus, A., Yang, H. & Plenz, D. Scale-invariant neuronal avalanche dynamics and the
cut-off in size distributions. *PLoS One* **9**, e99761 (2014).
https://doi.org/10.1371/journal.pone.0099761
43   Seeman, S. C. *et al.* Sparse recurrent excitatory connectivity in the microcircuit of the adult
mouse and human cortex. *eLife* **7**, e37349 (2018). https://doi.org/10.7554/eLife.37349
44   Abeles, M. *Corticonics: Neural Circuits of the Cerebral Cortex*. Vol. 1 (Cambridge University
Press, 1991).
45   Diesmann, M., Gewaltig, M. O. & Aertsen, A. Stable propagation of synchronous spiking in
cortical neural networks. *Nature* **402**, 529-533 (1999). https://doi.org/10.1038/990101
46   Prechtl, J. C., Cohen, L. B., Pesaran, B., Mitra, P. P. & Kleinfeld, D. Visual stimuli induce waves
of electrical activity in turtle cortex. *Proc.Natl.Acad.Sci.U.S.A* **94**, 7621-7626 (1997).
47   Rubino, D., Robbins, K. A. & Hatsopoulos, N. G. Propagating waves mediate information
transfer in the motor cortex. *Nat. Neurosci.* **9**, 1549-1557 (2006).
48   Singer, W. Cortical dynamics revisited. *Trends Cogn Sci* (2013).
49   Gray, C. M., Konig, P., Engel, A. K. & Singer, W. Oscillatory responses in cat visual cortex
exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature* **338**,
334-337 (1989).
50   Lisman, J. E. Bursts as a unit of neural information: making unreliable synapses reliable. *Trends*
Neurosci. **20**, 38-43 (1997).
51   Riquelme, J. L., Hemberger, M., Laurent, G. & Gjorgjieva, J. Single spikes drive sequential
propagation and routing of activity in a cortical network. *eLife* **12**, e79928 (2023).
https://doi.org/10.7554/eLife.79928
52   Shu, Y., Hasenstaub, A., Badoual, M., Bal, T. & McCormick, D. A. Barrages of synaptic activity
control the gain and sensitivity of cortical neurons. *J. Neurosci.* **23**, 10388-10401 (2003).
53   Tsodyks, M., Uziel, A. & Markram, H. Synchrony generation in recurrent networks with
frequency-dependent synapses. *J. Neurosci.* **20**, RC50 (2000).
54   Molnár, G. *et al.* Complex Events Initiated by Individual Spikes in the Human Cerebral Cortex.
PLoS Biol. **6**, e222 (2008). https://doi.org/10.1371/journal.pbio.0060222
55   Ferrarese, L. *et al.* Dendrite-Specific Amplification of Weak Synaptic Input during Network
Activity In Vivo. *Cell Reports* **24**, 3455-3465.e3455 (2018).
https://doi.org:https://doi.org/10.1016/j.celrep.2018.08.088
56   Harris, K. D. & Mrsic-Flogel, T. D. Cortical connectivity and sensory coding. *Nature* **503**, 51-58
(2013). https://doi.org/10.1038/nature12654
57   Histed, M. H. & Maunsell, J. H. Cortical neural populations can guide behavior by integrating
inputs linearly, independent of synchrony. *Proc. Natl. Acad. Sci. U. S. A.* **111**, E178-187 (2014).
https://doi.org/10.1073/pnas.1318750111
58   Barth, A. L. & Poulet, J. F. A. Experimental evidence for sparse firing in the neocortex. *Trends*
Neurosci. **35**, 345-355 (2012).
59   Kinouchi, O. & Copelli, M. Optimal dynamical range of excitable networks at criticality. *Nat.*
Phys. **2** 348-351 (2006).

60   Stevens, S. S. *Psychophysics: Introduction to its perceptual, neural and social prospects*.
(Routledge, 2017).

61   Assis, V. R. & Copelli, M. Dynamic range of hypercubic stochastic excitable media. *Phys Rev E*
Stat Nonlin Soft Matter Phys **77**, 011923 (2008). https://doi.org:10.1103/PhysRevE.77.011923

62   Jun, N. Y. *et al.* Coordinated multiplexing of information about separate objects in visual cortex.
eLife **11**, e76452 (2022). https://doi.org:10.7554/eLife.76452

63   Caruso, V. C. *et al.* Single neurons may encode simultaneous stimuli by switching between
activity patterns. *Nature communications* **9**, 2715 (2018). https://doi.org:10.1038/s41467-018-
05121-8

64   Levy, M., Sporns, O. & MacLean, J. N. Network analysis of murine cortical dynamics implicates
untuned neurons in visual stimulus coding. *Cell Reports* **31** (2020).
https://doi.org:10.1016/j.celrep.2020.03.047

65   Scott, G. *et al.* Voltage imaging of waking mouse cortex reveals emergence of critical neuronal
dynamics. *J. Neurosci.* **34**, 16611 - 16620 (2014). https://doi.org:10.1523/JNEUROSCI.3474-
14.2014

66   Bellay, T., Klaus, A., Seshadri, S. & Plenz, D. Irregular spiking of pyramidal neurons organizes
as scale-invariant neuronal avalanches in the awake state. *eLife* **4**, e07224 (2015).
https://doi.org:10.7554/eLife.07224

67   Bellay, T., Shew, W. L., Yu, S., Falco-Walter, J. J. & Plenz, D. Selective participation of single
cortical neurons in neuronal avalanches. *Frontiers in Neural Circuits* **14** (2021).
https://doi.org:10.3389/fncir.2020.620052

68   Capek, E. *et al.* Parabolic avalanche scaling in the synchronization of cortical cell assemblies.
Nature communications **14**, 2555 (2023).

69   Petermann, T. *et al.* Spontaneous cortical activity in awake monkeys composed of neuronal
avalanches. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 15921-15926 (2009).
https://doi.org:10.1073/pnas.0904089106

70   Beggs, J. M. & Plenz, D. Neuronal avalanches are diverse and precise activity patterns that are
stable for many hours in cortical slice cultures. *J. Neurosci.* **24**, 5216-5229 (2004).

71   Stewart, C. V. & Plenz, D. Inverted-U profile of dopamine-NMDA-mediated spontaneous
avalanche recurrence in superficial layers of rat prefrontal cortex. *J. Neurosci.* **26**, 8148-8159
(2006). https://doi.org:10.1523/JNEUROSCI.0723-06.2006

72   Lombardi, F., Herrmann, H. J., Plenz, D. & de Arcangelis, L. Temporal correlations in neuronal
avalanche occurrence. *Scientific Reports* **6**, 24690 (2016). https://doi.org:10.1038/srep24690

73   Lombardi, F., Herrmann, H., Plenz, D. & de Arcangelis, L. On the temporal organization of
neuronal avalanches. *Front. Syst. Neurosci.* **8**, 204 (2014).
https://doi.org:10.3389/fnsys.2014.00204

74   Goldey, G. J. *et al.* Removable cranial windows for long-term imaging in awake mice. *Nature*
Protocols **9**, 2515 (2014). https://doi.org:10.1038/nprot.2014.165

75   Dana, H. *et al.* High-performance GFP-based calcium indicators for imaging activity in neuronal
populations and microcompartments. *bioRxiv* (2018).

76   Klapoetke, N. C. *et al.* Independent optical excitation of distinct neural populations. *Nat.*
Methods **11**, 338 (2014). https://doi.org:10.1038/nmeth.2836
https://www.nature.com/articles/nmeth.2836#supplementary-information

77   Juavinett, A. L., Nauhaus, I., Garrett, M. E., Zhuang, J. & Callaway, E. M. Automated
identification of mouse visual areas with intrinsic signal imaging. *Nature Protocols* **12**, 32-43
(2017). https://doi.org:10.1038/nprot.2016.158

78   Sit, K. K. & Goard, M. J. Distributed and retinotopically asymmetric processing of coherent
motion in mouse visual cortex. *Nature communications* **11**, 1-14 (2020).
79   Kleiner, M., Brainard, D. & Pelli, D. What's new in Psychtoolbox-3? (2007).
80   Adesnik, H. & Abdeladim, L. Probing neural codes with two-photon holographic optogenetics.
Nat. Neurosci. **24**, 1356-1366 (2021). https://doi.org:10.1038/s41593-021-00902-9
81   Pachitariu, M. *et al.* Suite2p: beyond 10,000 neurons with standard two-photon microscopy.
bioRxiv, 061507 (2017). https://doi.org:10.1101/061507
82   Pedregosa, F. Scikit-learn: Machine learning in python Fabian. *Journal of machine learning*
research **12**, 2825 (2011).
83   Shapley, L. S. Notes on the n-person game—ii: The value of an n-person game. (RAND
Coorperation, Santa Monica, CA, USA, 1951).
84   Scott, M. & Su-In, L. A unified approach to interpreting model predictions. *Advances in neural*
information processing systems **30**, 4765-4774 (2017).
85   Poil, S. S., Hardstone, R., Mansvelder, H. D. & Linkenkaer-Hansen, K. Critical-state dynamics
of avalanches and oscillations jointly emerge from balanced excitation/inhibition in neuronal
networks. *J. Neurosci.* **32**, 9817-9823 (2012).
86   Villegas, P., Di Santo, S., Burioni, R. & Muñoz, M. A. Time-series thresholding and the
definition of avalanche size. *Phys. Rev. E* **100** (2019).
https://doi.org:10.1103/PhysRevE.100.012133
87   Capek, E. *et al.*   (Zenodo, 2023).

829