

# A network model of the barrel cortex combined with a differentiator detector reproduces features of the behavioral response to single-neuron stimulation

Davide Bernardi<sup>1,2\*</sup>, Guy Doron<sup>1</sup>, Michael Brecht<sup>1</sup>, Benjamin Lindner<sup>1,2</sup>,

**1** Bernstein Center for Computational Neuroscience Berlin, Philippstr. 13, Haus 2, 10115 Berlin, Germany

**2** Physics Department of Humboldt University Berlin, Newtonstr. 15, 12489 Berlin, Germany

\* [davide.bernardi@bccn-berlin.de](mailto:davide.bernardi@bccn-berlin.de)

## Abstract

The stimulation of a single neuron in the rat somatosensory cortex can elicit a behavioral response. The probability of a behavioral response does not depend appreciably on the duration or intensity of a constant stimulation, whereas the response probability increases significantly upon injection of an irregular current. Biological mechanisms that can potentially suppress a constant input signal are present in the dynamics of both neurons and synapses and seem ideal candidates to explain these experimental findings. Here, we study a large network of integrate-and-fire neurons with several salient features of neuronal populations in the rat barrel cortex. The model includes cellular spike-frequency adaptation, experimentally constrained numbers and types of chemical synapses endowed with short-term plasticity, and gap junctions. Numerical simulations of this model indicate that cellular and synaptic adaptation mechanisms alone may not be sufficient to account for the experimental results if the local network activity is read out by an integrator. However, a differentiator circuit can detect the single-cell stimulation with a reliability that barely depends on the length or intensity of the stimulus, but that increases when an irregular signal is used. This finding is in accordance with the experimental results obtained for the stimulation of a regularly-spiking excitatory cell.

## Author summary

It is widely assumed that only a large group of neurons can encode a stimulus or control behavior. This tenet of neuroscience has been challenged by experiments in which stimulating a single cortical neuron has had a measurable effect on an animal's behavior. Recently, theoretical studies have explored how a single-neuron stimulation could be detected in a large recurrent network. However, these studies missed essential biological mechanisms of cortical networks and are unable to explain more recent experiments in the barrel cortex. Here, to describe the stimulated brain area, we propose and study a network model endowed with many important biological features of the barrel cortex. Importantly, we also investigate different readout mechanisms, i.e. ways in which the stimulation effects can propagate to other brain areas. We show that a readout network which tracks rapid variations in the local network activity is in agreement with the experiments. Our model demonstrates a possible mechanism for how the stimulation of

a single neuron translates into a signal at the population level, which is taken as a proxy of the animal's response. Our results illustrate the power of spiking neural networks to properly describe the effects of a single neuron's activity.

## Introduction

A classical method used in neuroscience to understand cortical circuits is to determine how single neurons respond to a controlled sensory stimulus. If, for instance, one of a rat's whiskers is moved by the experimenter, a change in firing of specific neurons can be measured in the barrel cortex, one of the most well studied parts of the primary sensory cortex [1]. The concept of "reverse physiology" turns this approach around by studying the inverse situation in which neurons in higher brain areas are stimulated and a behavioral or motor response can be elicited (see [2] for early references on such experiments). The two kinds of experiments in combination then allow the linking of sensation and perception, one of the notoriously difficult problems in neuroscience.

In the case that the stimulation affects only a single neuron, the outcome of the unconventional and technically challenging reverse physiology experiments are particularly striking: stimulating a single neuron in the motor cortex can evoke a whisker movement [3], and single-cell stimulation in the barrel cortex (but not in the thalamus) leads to a weak but statistically significant behavioral response [4–6]. This contradicts prevailing hypotheses that relevant signals can only be encoded in the activity of large neural populations.

Both the enormity of cortical networks - tens of thousands of neurons in the case of the somatosensory cortex [7] - and the apparent randomness of single-neuron spiking [8,9] have classically been evoked as arguments for population coding. If single spikes are unpredictable and noisy how can a few externally induced spikes lead to changes in behavior?

On the theoretical side, cortical populations have been modeled as (locally) random networks of synaptically coupled excitatory and inhibitory cells [10–12] (many studies just take into account two distinct cell types). Even without the inclusion of explicit noise sources, these models can show asynchronous irregular activity [13–15] that is similar to that observed in the cortex of alert animals [16,17]; this kind of network noise can also be described analytically by stochastic mean-field methods (see for instance [10,13,18–20]). Besides the autonomous activity of such networks, their linear and nonlinear response to global stimuli (applied to *all* neurons in the network) has been in focus [20–24]. However, injecting a current into a *single* neuron in such a generic network model can lead to sizable changes in a subpopulation's activity as well [25,26]; this subpopulation can be regarded as a readout of the stimulus. If the readout population is somewhat oriented towards the direct postsynaptic partners of the stimulated cell, then the stimulus can be detected in the activity of this population [25]. If, more realistically, the readout is accomplished by a second recurrent network with feed-forward inhibition (as is most likely the case in the cortex), already a very small bias will lead to a detection performance comparable to that in the behavioral experiment [26]. Especially challenging for theoreticians are the results of the nanostimulation experiments in the barrel cortex of behaving rats [6] which demonstrated striking dependencies of the behavioral response on the properties of the stimulating current. The response does not depend on the duration of the stimulus, it depends weakly on its intensity, and strongly on its irregularity: the response is greatly enhanced if the current varies irregularly within the stimulation window instead of being held constant.

Here we tackle these challenges by studying a more detailed computational model that is much closer to the barrel cortex network than the generic setups previously

investigated in the context of single-cell stimulation [25, 26]. We also put forward a novel implementation of an alternative readout mechanism, based on the differentiation of the population activity. Our model includes three distinct populations of cells (one excitatory neuron type and two distinct inhibitory interneuron types) that are all modeled as integrate-and-fire neurons endowed with adaptation currents, mediating the omnipresent spike-frequency adaptation. These cells are coupled by chemical synapses, which undergo synaptic depression or facilitation, and by electrical synapses (gap junctions). Combined with the so-called *differentiator readout*, we achieve a response at the population level that is in several key aspects similar to the behavioral response in the experiments by [6].

## Results

The model consists of two parts: a recurrent network, in which a randomly selected excitatory regular spiking cell is stimulated to mimic the experiments, and a readout, which receives input from excitatory neurons of the recurrent network and can detect the stimulation.

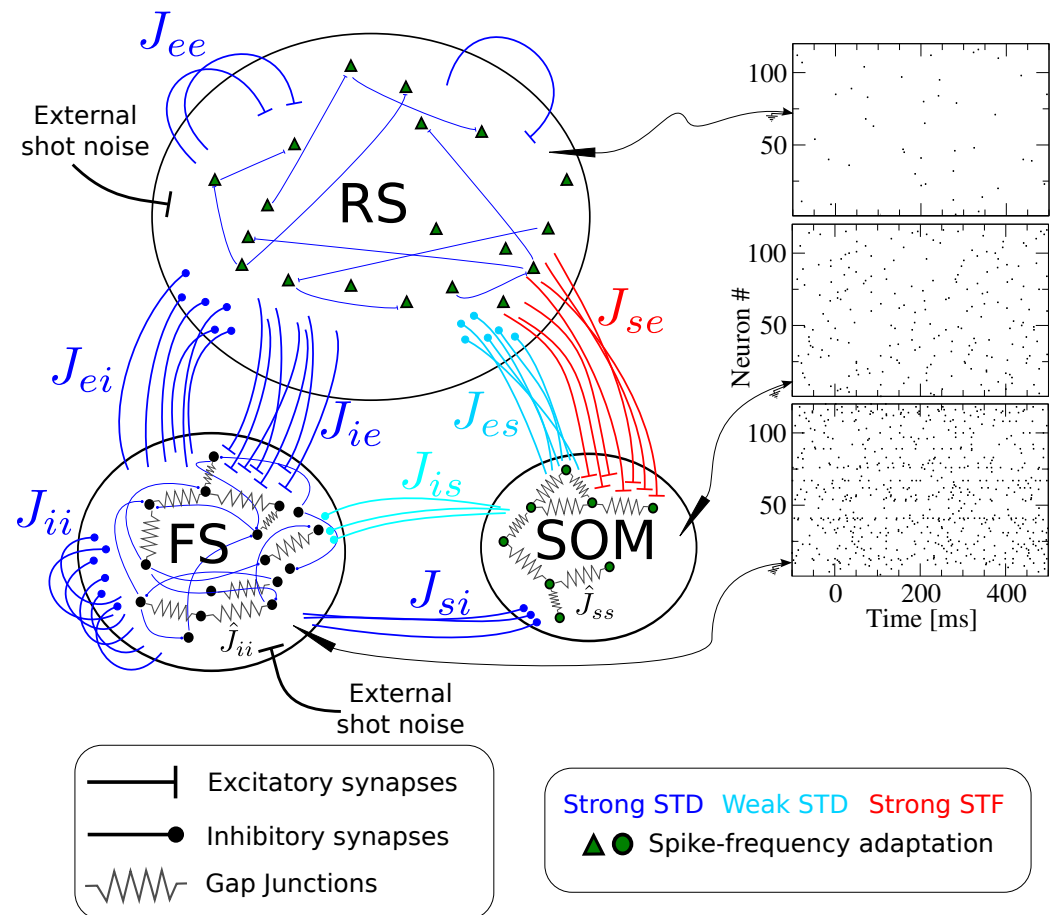
### Recurrent network model

Figure 1 shows a scheme containing all essential features of the BCN, briefly described in the figure caption. The recurrent network model represents the surroundings of the stimulated cell in a radius of about 200  $\mu\text{m}$ , in which connection probabilities can be considered as constant [27, 28]. A total network size of  $N = 2600$  is in line with estimates of neuron density in the barrel cortex [7]. Although this network size corresponds to a fraction of one barrel, this part of the model will be indicated as “barrel cortex network” (BCN). The BCN consists of three neuronal populations: excitatory regular spiking cells (RS), inhibitory fast-spiking (FS) cells, and somatostatin-expressing low-threshold spiking (SOM-LTS) inhibitory neurons. These three cell types account for a large fraction of the neurons found in the barrel cortex (about 99% of layer IV [29]). All neurons are modeled as one-compartment leaky integrate-and-fire (LIF) neurons. Besides leak conductance and spike generation, several other biological mechanisms are modeled, according to the cell type.

The largest population consists of 2000 excitatory RS neurons, which are sparsely connected to each other but densely connected to FS interneurons and SOM-LTS cells, as experimental studies report [29]. The membrane time constant of RS cells is lognormally distributed with mean  $\tau_{m,e} = 20$  ms and a standard deviation of 20% of the mean [29, 30]. The 400 inhibitory FS interneurons are characterized by faster membrane time constants (lognormal distribution with mean  $\tau_{m,i} = 10$  ms [29]) and are densely connected both to other FS neurons and to RS cells. The 200 SOM-LTS inhibitory neurons possess longer time scales (lognormal distribution with mean  $\tau_{m,s} = 20$  ms [29]) and a firing threshold which is 6 mV lower than in RS and FS neurons [29]. SOM-LTS neurons do not inhibit each other via chemical synapses, but form dense connections to and from RS neurons and sparser connections to and from FS interneurons [29, 31, 32].

Both FS and SOM-LTS neurons are densely coupled to cells of the same type via electrical synapses (gap junctions) [29, 31, 33], which are represented here by an effective global excitatory spiking coupling (the sub-threshold contribution of gap junctions has a much smaller effects on the network dynamics [34]).

In the barrel cortex, both RS and SOM-LTS neurons display spike-frequency adaptation, whereas FS neurons do not [29, 35]. Therefore, RS and SOM-LTS are endowed with a spike-triggered hyperpolarizing current [36–38]. Consistent with



**Fig 1. Recurrent network model representing the surroundings of the stimulated cell.** The network is formed by  $N_e = 2000$  excitatory regular spiking (RS) neurons,  $N_i = 400$  inhibitory fast spiking (FS) neurons, and  $N_s = 200$  inhibitory somatostatin-positive low-threshold spiking (SOM-LTS) neurons. Recurrent connections between RS neurons are sparse (15%), all connections involving FS neurons as well as those between RS and SOM-LTS neurons are dense (40%-50%). FS and SOM-LTS neurons are electrically coupled (only neurons of the same type). Gap junctions are represented by an effective all-to-all spiking coupling (see main text). Connections in blue are strongly depressing, connections in light blue are weakly depressing, and connections in red are strongly facilitating. RS and SOM-LTS neurons are endowed with a spike-frequency adaptation current. Input from the thalamus and from neighboring cortical regions is represented by Poissonian shot noise. SOM-LTS neurons do not receive external shot noise. The three raster plots show the spontaneous activity of 120 (from top to bottom) RS, SOM, and FS neurons. The spontaneous activity of all three populations is asynchronous and irregular. The spontaneous mean firing rate of excitatory RS neurons is  $r_{sp,e} \approx 0.8$  Hz, of SOM-LTS neurons is  $r_{sp,s} \approx 3$  Hz, and of FS neurons is  $r_{sp,i} \approx 10$  Hz.

experimental observations, the strength of the adaptation current is larger for RS neurons than for SOM-LTS neurons.

We drive RS and FS cells with Poissonian spike trains mimicking input from the thalamus and neighboring cortical areas. SOM-LTS cells are mostly subject to local input [29, 39] and therefore, in our model, they do not receive external input.

Experimental studies suggest that most synapses in the barrel cortex show short term depression, with the notable exception of connections from RS neurons to SOM-LTS cells, which have been found to be strongly facilitating [29, 40–42]. The short term plasticity of chemical synapses is simulated here by means of a standard model [43, 44]. Parameters were chosen such that all synapses except those from and to SOM-LTS neurons are strongly depressing. These synapses are depicted in blue in fig. 1. Parameters for synapses branching from SOM-LTS neurons (represented in light blue in fig. 1) are set such that they display a weak depression. Finally, parameters of synapses connecting RS cells to SOM-LTS neurons are chosen to generate a strong facilitation. A further property that distinguishes these synapses is that the transmission failure rate is high ( $\sim 50\%$ ) at low presynaptic firing rate, but the reliability increases upon repeated activation of the synapse [29]. This property is modeled by a variable that mimics the activity-dependent failure rate.

More details on model equations and parameters are given in the Methods.

## Readout models

Because outgoing long-range connections in the cortex mostly originate from pyramidal neurons, we assume that the readout can access a subset of the RS cells and consider three possible schemes, as illustrated in fig. 2. The first detection scheme (fig. 2A) receives input from a subset of the excitatory neurons of the BCN and reacts when the filtered activity of these neurons reaches a lower barrier. This readout scheme will be called integrator readout (IR). The second readout scheme (fig. 2B) filters the activity in the same way as the IR, but it subtracts a time-shifted copy of the same activity. In other words, it considers the difference between the filtered activity at different time points, thus acting as a sort of differentiator. For this reason, it will be referred to as differentiator readout (DR). The third readout scheme is the implementation of the DR by means of a simple network of LIF neurons and will be called differentiator network readout (DNR).

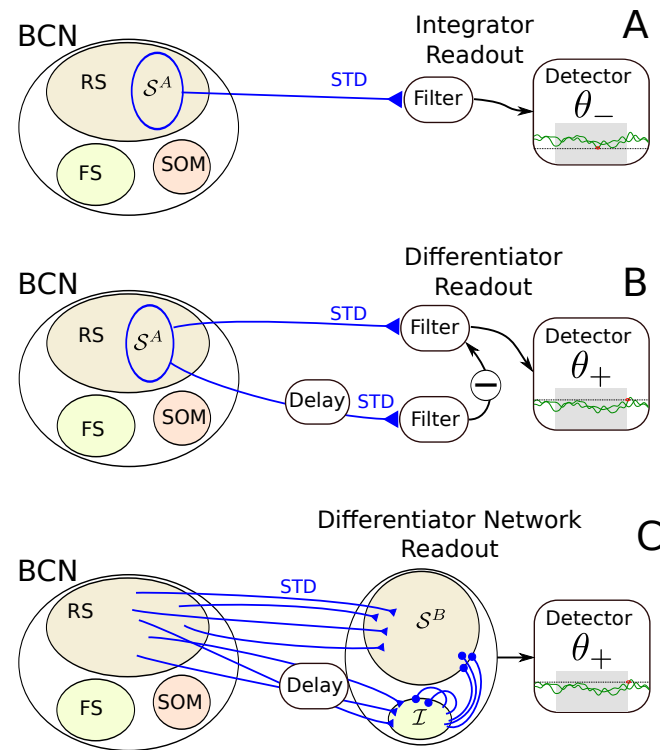
### Integrator readout

The first readout scheme, the integrator readout (IR), is based on a random selection of  $\hat{C} = 1000$  RS neurons that constitute the readout set  $\mathcal{S}^A$  (see fig. 2A). The spike trains emitted by all neurons within the readout population  $\mathcal{S}^A$  are linearly filtered by using the following dynamical equation:

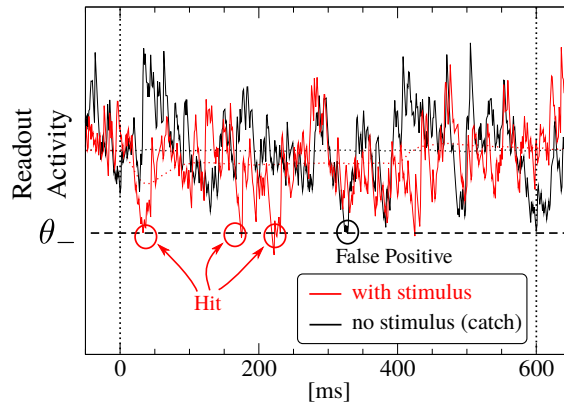
$$\tau_{m,e} \frac{dA_{ir}(t)}{dt} = -A_{ir} + R_{m,read} \left[ \sum_{i \in \mathcal{S}^A} J_{read,i}(t) x_i(t) \right], \quad (1)$$

where  $x_i$  is the spike train of the  $i$ th neuron within the readout set  $\mathcal{S}^A$ , the integration time constant is  $\tau_{m,e} = 20$  ms, and  $R_{m,read} = \tau_{m,e}/C_m$ . Consistent with the idea that synapses projecting to other brain areas will also undergo depression, the dynamic weights  $J_{read,i}(t)$  obey the same equation as all excitatory weights within the BCN.

To compute false positive and correct detection rates, a single lower decision boundary  $\theta_-$  is used by the IR, as depicted in fig. 3. A detection event is registered if  $A_{ir}$  falls at least once below the boundary  $\theta_-$  within the detection time window



**Fig 2. Readout models considered in this paper.** A cell selected at random from the barrel cortex network (BCN) is selected at random and stimulated. The BCN consists of three populations: excitatory regular-spiking neurons (RS), inhibitory fast-spiking neurons (FS), and somatostatin-positive low-threshold spiking neurons (SOM-LTS), all modeled by leaky integrate-and-fire neurons. The BCN includes several biological details of the barrel cortex (see fig. 1 and methods). Three readout schemes are considered. **A:** the integrator readout (IR) integrates the activity of a subset of the RS neurons within the BC and reacts to deviations in the negative direction. **B:** the differentiator readout (DR) evaluates the difference between the IR activity at two time points separated by a delay. This filtered running difference at fixed lag is processed by the detector, which reacts when an upper threshold is reached. **C:** the differentiator network readout (DNR) implements the operation of the DR with two populations of LIF neurons. The FS readout population provides delayed recurrent inhibition to itself and feed-forward inhibition to the RS readout population. All connections depicted in blue are dynamic and show short-term depression (STD).



**Fig 3. Working principle of the detector.** Here, the case of a lower boundary is shown. The black trace is one realization of  $A_{ir}(t)$  for a catch trial (no stimulus). Because the trajectory hits the detection boundary once, this trial generates a false positive event. The red trace represents one realization of the IR activity  $A_{ir}(t)$  in the presence of a stimulus. Because the trajectory hits the boundary at least once (in this case even multiple times), this trial triggers a hit, or correct detection event.

$(0, T_w = 600 \text{ ms})$ . In catch trials, no stimulus is present. These catch trials are used to determine the *false positive rate*:

$$\mathcal{FP}_{ir}(\theta_-) = \left\langle \max_{t \in (0, T_w)} \left\{ H(\theta_- - A_{ir}(t)) \right\} \right\rangle_{\text{catch}}, \quad (2)$$

where  $H(x)$  is the Heaviside step function, and angular brackets indicate trial average. The reasons why a single lower detection boundary is used are explained below (see Firing-rate response of the network). The *hit* or *correct detection* rate is computed exactly in the same way, but in the presence of a stimulus

$$\mathcal{CD}_{ir}(\theta_-) = \left\langle \max_{t \in (0, T_w)} \left\{ H(\theta_- - A_{ir}(t)) \right\} \right\rangle_{\text{stim.}}. \quad (3)$$

### Differentiator readout

The differentiator readout (DR) first reads in the input from the network in the same way as the IR and it takes the difference between  $A_{ir}$  evaluated at two times separated by a lag  $\Delta T$ . The result is then convolved with an exponential filter

$\mathcal{F}_{\tau_f}(t) = \exp(-t/\tau_f)/\tau_f$  to reduce the noise

$$A_{dr}(t) = (A_{ir}(t) - A_{ir}(t - \Delta T)) \star \mathcal{F}_{\tau_f}(t). \quad (4)$$

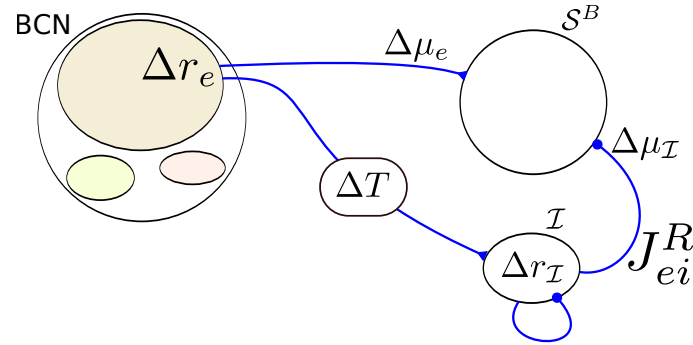
Trajectories computed from eq. (4) are used in combination with an upper detection threshold  $\theta_+$  to obtain the false positive and hit rates:

$$\mathcal{FP}_{dr}(\theta_+) = \left\langle \max_{t \in (0, T_w)} \left\{ H(A_{dr}(t) - \theta_+) \right\} \right\rangle_{\text{catch}} \quad (5)$$

$$\mathcal{CD}_{dr}(\theta_+) = \left\langle \max_{t \in (0, T_w)} \left\{ H(A_{dr}(t) - \theta_+) \right\} \right\rangle_{\text{stim.}}. \quad (6)$$

### Differentiator network readout

The operation performed by the DR, i.e. the subtraction of a delayed copy of the readout activity, can be approximately implemented by the simple network shown in



**Fig 4. Tuning of the differentiator readout network to implement the operation of the differentiator readout scheme.** A perturbation in the firing rate of the RS neurons in the BCN ( $\Delta r_e$ ) causes a perturbation in the mean input to the RS readout neurons,  $\Delta \mu_e$ , and a perturbation in the firing rate of the inhibitory readout population  $\mathcal{I}$ . This change in firing rate causes a shift in the input from  $\mathcal{I}$  to  $\mathcal{S}^B$  ( $\Delta \mu_{\mathcal{I}}$ ). The strength of the connection from  $\mathcal{I}$  to  $\mathcal{S}^B$  is adjusted such that  $\Delta \mu_e + \Delta \mu_{\mathcal{I}} = 0$ . This cancellation reaches  $\mathcal{S}^B$  with a time lag  $\Delta T$  so that the readout network roughly implements the operation of the DR eq. (4).

fig. 2C. The differentiator network readout (DNR) consists of two populations: one readout population of 10 000 RS neurons ( $\mathcal{S}^B$ ) and one population of 2000 FS inhibitory neurons ( $\mathcal{I}$ ). Both populations receive the same number of excitatory feed-forward connections from the RS population of the BCN. More precisely, each neuron in the DNR receives input from  $\hat{C} = 1000$  randomly chosen RS neurons. Neurons in the readout population  $\mathcal{S}^B$  evolve according to the same dynamical equation as RS neurons of the BCN, while neurons in  $\mathcal{I}$  follow the same dynamical equations as the FS neurons of the BCN.

If the purpose of the DNR is to implement the operation performed by the DR, it is crucial that the feed-forward inhibitory pathway from the BCN via  $\mathcal{I}$  to  $\mathcal{S}^B$  balances the direct feed-forward excitatory pathway input at a later time. To this end, the average weight of connections from  $\mathcal{I}$  to  $\mathcal{S}^B$ ,  $J_{ei}^R$ , is chosen such that a static change in the input from the direct pathway  $\Delta \mu_e$  is compensated by the static change in the input from the inhibitory pathway  $\Delta \mu_{\mathcal{I}}$  (see fig. 4). The value of  $J_{ei}^R$  that approximately satisfies the condition  $\Delta \mu_e + \Delta \mu_{\mathcal{I}} = 0$  is computed from a linear-response calculation, reported in the Methods. Importantly, the inhibitory pathway is given an additional transmission delay  $\Delta T = 10$  ms, so that changes in the input from the BCN to the DNR are counterbalanced at a later time. More details on the DNR can be found in the Methods along with all parameter values.

The DNR activity is obtained by filtering the average firing rate of the readout neurons  $\mathcal{S}^B$  with the same exponential filter used for the DR:

$$A_{\text{dnr}}(t) = \frac{1}{N_B} \sum_{x_k \in \mathcal{S}^B} x_k(t) \star \mathcal{F}_{\tau_f}(t). \quad (7)$$

False positive and hit rates are obtained in exactly the same way as done for the DR:

$$\mathcal{FP}_{\text{dnr}}(\theta_+) = \left\langle \max_{t \in (0, T_w)} \left\{ H(A_{\text{dnr}}(t) - \theta_+) \right\} \right\rangle_{\text{catch}}, \quad (8)$$

$$\mathcal{CD}_{\text{dnr}}(\theta_+) = \left\langle \max_{t \in (0, T_w)} \left\{ H(A_{\text{dnr}}(t) - \theta_+) \right\} \right\rangle_{\text{stim.}}. \quad (9)$$

## Effect size

The *effect size* is defined as the difference between the hit and the false positive rate [4]. It is a function of the detection threshold  $\theta$

$$\mathcal{Y}_X(\theta) = \mathcal{CD}_X(\theta) - \mathcal{FP}_X(\theta), \quad (10)$$

where  $X \in \{\text{ir}, \text{dr}, \text{dnr}\}$  indicates the detector type and  $\theta$  can be either an upper or lower boundary. The false positive rate of 0.25 corresponds approximately to the average false positive rate measured experimentally. For this reason, this value was chosen to compare the simulation results to the experimental data. More precisely, the threshold  $\bar{\theta}$  is chosen such that

$$\mathcal{FP}_X(\bar{\theta}) = 0.25, \quad (11)$$

which is then used to compute

$$\bar{\mathcal{Y}}_X = \mathcal{Y}_X(\bar{\theta}), \quad (12)$$

which is the final output of the detection procedure and will be compared to the experimental data.

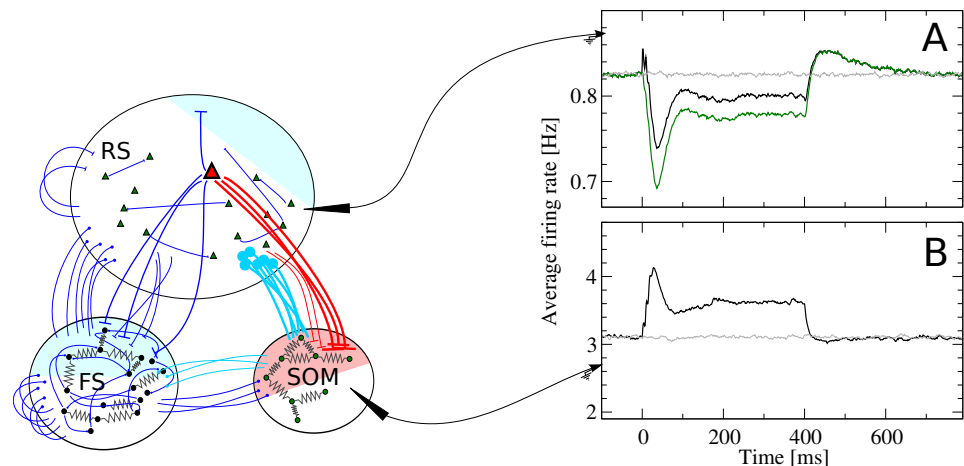
## Single-cell stimulation

In every trial, the network is initialized with random initial conditions and simulated for  $T_{\text{idle}} = 1200$  ms, to let the system forget the initial state. The spontaneous firing pattern of the network is asynchronous and irregular (fig. 1). The mean spontaneous firing rate of RS, FS, and SOM-LTS neurons is  $r_{\text{sp},e} \approx 0.8$  Hz,  $r_{\text{sp},i} \approx 10$  Hz, and  $r_{\text{sp},s} \approx 3$  Hz, respectively. These properties of the spontaneous firing activity are consistent with experimental observations [45, 46].

A neuron is randomly selected as site of the nanostimulation, which is switched on at  $t = 0$  and modeled as an additional current injected into the cell. The maximum stimulation current used here is  $\Delta I_{\text{max},e} = 5$  nA. This value is chosen to elicit a similar number of spikes as in the experiment and is in the range used experimentally [47]. After the stimulus is switched off, the network is further simulated until the time reaches  $t = T_{\text{end}} = 1200$  ms.

Following [6, 47], we use step currents of different lengths and intensities to investigate how the response probability depend on the properties of the stimulus. Furthermore, random permutations of steps of different length and amplitude will be used to generate irregular spike trains. Two equally-sized sets of catch trials, i.e. trials in which no stimulus was present, were simulated to estimate the size of random fluctuations in the detection rates.

The shot noise mimicking external input and the initial conditions were drawn anew in every trial. The same realization of the network (randomized cellular parameters and the connectivity matrix including weights and delays) was used once for each stimulus type. The total number of trials per stimulus type was  $N_{\text{trials}} = 10000$ .



**Fig 5. Disynaptic inhibition mediated by somatostatin-positive low-threshold spiking (SOM-LTS) cells causes inhibitory response to the stimulation of a regular spiking (RS) cell.** **A:** when a RS cell is stimulated (the stimulus is switched here on at  $t = 0$  and off at  $t = 400$  ms), synapses from the stimulated cell (red triangle) to the SOM-LTS population strongly facilitate and cause a large increase in the firing rate of the SOM population, which then relaxes back to a plateau because of the spike-frequency adaptation (**b**). The inhibitory input from the SOM to the RS population produces a response in the RS cells which is almost a mirror image (**a**, black line). The initial positive peak in the RS response is due to the spikes fired by the stimulated cell itself, as it can be seen by excluding it from the population (**a**, green line). The curves in panels **a** and **b** are averages over 10000 trials and spikes are filtered with an exponential filter with decay constant  $\tau_f = 15$  ms. Gray lines indicate catch trials (no stimulation).

## Firing-rate response of the network

Before investigating to what extent the three readout procedures introduced above are capable to detect the single-cell stimulation, it is instructive to examine the trial-averaged firing rate response of the network to the stimulation of a single RS cell. In the following, the case of a constant step current with intensity at 25% of the maximum and a duration of 400 ms is considered.

When a single RS cell is stimulated (red triangle in fig. 5), its output spikes affect a relatively small set of RS neurons (blue shaded area in fig. 5) because RS-to-RS connections are sparse (15% connection probability). Furthermore, their average amplitude is smaller if compared to other connections and they are strongly depressing, so that the direct effect on the overall firing rate of the RS population is small.

Connections from RS cells to the FS population are dense (40%), so that the spikes of the stimulated RS neuron reach a large fraction of the FS population (blue shaded area in fig. 5). However, FS cells also strongly inhibit one another and thus counteract the input from the stimulated cell. Consequently, the average change in the firing rate of the entire FS population is small (not shown).

Finally and most importantly, the output of the stimulated cell reaches a large fraction of the SOM-LTS population (50%, red shaded area in fig. 5) via strongly facilitating synapses. As a result, they induce an appreciable increase in the average firing rate of the SOM-LTS population, shown in fig. 5b. Importantly, SOM-LTS do not inhibit each other. However, the spike-frequency adaptation causes a strongly damped oscillation which, after an initial peak around  $t = 30$  ms and a dip around  $t = 100$  ms, relaxes to a plateau lying about 20% above the spontaneous firing rate level.

The increased inhibition from the SOM-LTS population ultimately causes the average firing rate of the RS population (fig. 5a, black line) to drop below the spontaneous level (fig. 5a, gray line). Note that the time course of the response of the RS population is closely related to that of the SOM-LTS population, except for the small peak shortly after the stimulus onset and for the overshoot after the stimulus is switched off. The small peak is due to the spikes fired by the stimulated cell itself. This can be seen by omitting the spikes fired by the stimulated cell (fig. 5a, green line) and observing that the peak disappears. The overshoot after  $t = 400$  ms is due to the slow relaxation of the adaptation variable to its baseline value.

These observations are consistent with *in vitro* experiments showing that the strong activation of a single pyramidal cell in the barrel cortex has mostly an inhibitory effect on nearby pyramidal cells, and that this effect is due to disynaptic inhibition mediated by SOM-LTS neurons [40, 41]. More recent *in vivo* experiments pairing the stimulation of a single cell with calcium imaging suggest that bursts induced in a pyramidal cell have a very weak excitatory effect on other pyramidal cells and on FS neurons, but have a significant effect on neighboring SOM-LTS cells [48], in line with the behavior of our model.

## Relation between effect size and statistics of readout activity

In the previous subsection, the effects of the single-cell stimulation on the *trial-averaged* response of the RS population have been examined. The readout, however, must decide on the presence of the stimulus based on the RS population activity *in each single trial*, a much more difficult task (compare the smooth lines of fig. 5A,B with the noisy curves in fig. 3). The readout performance is quantified by the effect size, defined above. Before investigating how the effect size depends on the properties of the stimulus, we will examine how changes in the statistical properties of the readout activity  $A_X(t)$  can influence the effect size.

The simplest statistics that can be considered are the time-dependent mean and standard deviation of the readout activity (the averaging ensemble consists of the different trials). Statistics of higher order (skewness and kurtosis) were measured and did not display appreciable deviations from the spontaneous state and will be therefore omitted for brevity. Because we are interested in deviations from the spontaneous state, it makes sense to consider mean and standard deviation of the readout activity as standardized deviations from the spontaneous value. More precisely, we will consider first the time-dependent mean of the readout activity  $A_X(t)$  (here  $X = \text{ir, dr, dnr}$ , as defined in the subsection Readout models):

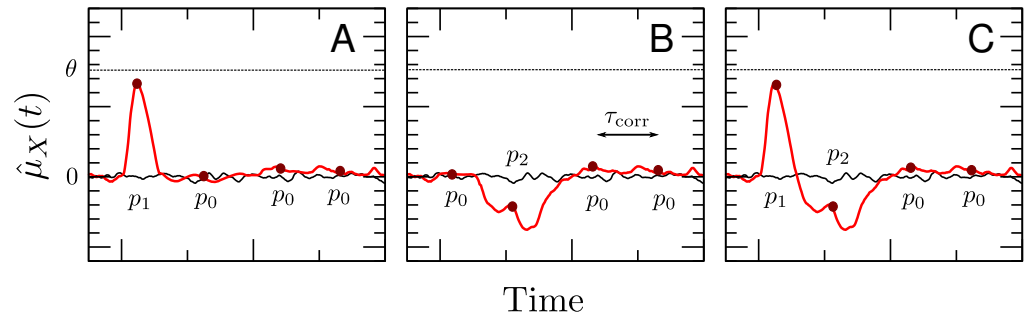
$$\hat{\mu}_X(t) = \frac{\langle A_X(t) \rangle - \mu_{X, \text{catch}}}{\sigma_{X, \text{catch}}}, \quad (13)$$

where  $\mu_{X, \text{catch}}$  and  $\sigma_{X, \text{catch}}$  are the average mean and standard deviation in the spontaneous state, respectively:

$$\mu_{X, \text{catch}} = \langle A_X(t) \rangle_{\text{catch}}, \quad (14)$$

$$\sigma_{X, \text{catch}} = \sqrt{\langle \Delta A_X^2(t) \rangle_{\text{catch}}}, \quad (15)$$

where  $\Delta A_X^2(t) = \left( A_X(t) - \langle A_X(t) \rangle \right)^2$ , and the time dependence in both last equations is self-averaging due to the stationary conditions. The time-dependent standard



**Fig 6. Illustration of simplified detection model used to interpret the simulation results.** The continuous-time problem is approximated by a discretized process, obtained by “sampling” trajectories at times separated by the correlation time. Each deviation from the spontaneous state changes the probability of *not* reaching the decision barrier from the spontaneous value  $p_0$  to  $p_1$  or to  $p_2$ .

deviation of the readout activity is defined in a similar way:

$$\hat{\sigma}_X(t) = \frac{\sqrt{\langle \Delta A_X^2(t) \rangle} - \sigma_{X, \text{catch}}}{\sigma_{X, \text{catch}}}. \quad (16)$$

Non-zero values of  $\hat{\mu}_X(t)$  and of  $\hat{\sigma}_X(t)$  at any time point within the detection time window can impact the effect size in different ways. Suppose, for instance, that the considered detector employs an upper boundary. Then, a positive deflection of  $\hat{\mu}_X(t)$  locally increases the probability of reaching the threshold, whereas a negative deflection reduces it. If a lower detection boundary is used, the opposite holds. Regardless of the type of threshold, a local increase of  $\hat{\sigma}_X(t)$  enhances the probability of reaching the threshold, whereas a local decrease in  $\hat{\sigma}_X(t)$  reduces the probability of crossing the decision barrier. This line of reasoning is qualitative only and holds under the assumption that  $A_X(t)$  is approximately normally distributed at all times.

To understand how multiple deviations from the spontaneous state within the decision time window jointly influence the effect size, it is useful to consider a simplified description of the decision model introduced in [25, 26, 49]. In this simplified theory, hit and false positive rates are approximated as the result of  $n = T_w / \tau_{\text{corr}}$  draws of a random variable, where  $T_w$  is the detection time window and  $\tau_{\text{corr}}$  is the autocorrelation time of the readout activity (in the example of fig. 6,  $n = 4$ ). If these draws are treated as independent, the false positive rate reads

$$\mathcal{FP}(\theta) = 1 - p_0^n(\theta), \quad (17)$$

where  $p_0(\theta)$  is the probability of *not* crossing the barrier  $\theta$  at a given time point and in the absence of the stimulus [ $p_0(\theta)$  does not depend on time and is therefore the same for each draw]. For concreteness, let us use an upper barrier at the value  $\bar{\theta}$ , which yields the false positive rate of 0.25. In this way, the dependence on  $\theta$  can be dropped, but the following considerations do not depend on the particular position or type of the boundary. Suppose now that  $\hat{\mu}_X(t)$  displays one peak at a certain position within the detection time window, as depicted in fig. 6A. Therefore, the probability that one trajectory of the readout activity triggers the detector is locally increased. Thus, in the vicinity of the peak, the probability of *not* triggering the detector will be

$p_1 = p_0 + \Delta p_1 < p_0$ . The correct detection rate for this situation is then

$$\bar{\mathcal{C}}\mathcal{D}_1 = 1 - p_1 p_0^{n-1}. \quad (18)$$

Consequently, the effect size reads

$$\bar{\mathcal{Y}}_1 = 1 - p_1 p_0^{n-1} - (1 - p_0^n) = p_0^n \left(1 - \frac{p_1}{p_0}\right). \quad (19)$$

Consider now the situation of a negative deflection in  $\hat{\mu}_X(t)$  occurring at a different time (as in fig. 6B). Locally, the probability of not triggering the detector is  $p_2 = p_0 + \Delta p_2 > p_0$ . In this case, the effect size is

$$\bar{\mathcal{Y}}_2 = 1 - p_2 p_0^{n-1} - (1 - p_0^n) = p_0^n \left(1 - \frac{p_2}{p_0}\right). \quad (20)$$

Suppose now that both features are present at sufficiently separated times within the same detection time window, as in fig. 6C. In this case, the effect size is

$$\bar{\mathcal{Y}}_{12} = p_0^n - p_1 p_2 p_0^{n-2} = p_0^n \left(1 - \frac{p_1 p_2}{p_0^2}\right). \quad (21)$$

Substituting  $p_1 = p_0 + \Delta p_1$  and  $p_2 = p_0 + \Delta p_2$  into eq. (21) and supposing  $\Delta p_1, \Delta p_2 \ll 1$  yields

$$\begin{aligned} \bar{\mathcal{Y}}_{12} &= p_0^n \left(1 - \frac{p_0^2 + p_0 \Delta p_1 + p_0 \Delta p_2 + \Delta p_1 \Delta p_2 + p_0^2 - p_0^2}{p_0^2}\right) \\ &\approx p_0^n \left(1 - \frac{p_0 + \Delta p_1}{p_0} + 1 - \frac{p_0 + \Delta p_2}{p_0}\right) \\ &= \bar{\mathcal{Y}}_1 + \bar{\mathcal{Y}}_2. \end{aligned} \quad (22)$$

This approximation generalizes to the case of more than two deviations from the spontaneous state [49], given that all deviations are small compared to  $p_0$ . For instance, when three features are present the effect size is

$$\bar{\mathcal{Y}}_{123} = p_0^n \left(1 - \frac{p_1 p_2 p_3}{p_0^3}\right) \approx \bar{\mathcal{Y}}_1 + \bar{\mathcal{Y}}_2 + \bar{\mathcal{Y}}_3. \quad (23)$$

The main insight here is that weak deviations from the spontaneous state appearing in the same detection window can (approximately) add or cancel each other. This will be useful in the following to interpret how the behavior of  $\hat{\mu}_X(t)$  and  $\hat{\sigma}_X(t)$  influence the response of the detector.

## Detection of the stimulation of a single neuron

In this subsection, we will investigate how the properties of the stimulus influence the response probability of the detector for each of the three readout procedures.

### Effect of stimulus duration

The effect of changing the stimulus duration will be considered first. To this end, stimuli of length 100, 200, and 400 ms are used (fig. 7, top). The stimulus intensity is kept constant at 25% of the maximum current. In the experiment, when the stimulated cell was a RS neuron, the three stimuli evoked  $6 \pm 3$ ,  $11 \pm 5$ , and  $23 \pm 10$  spikes, respectively. In the model, the number of evoked spikes was  $7 \pm 1$ ,  $12 \pm 2$ , and  $20 \pm 5$

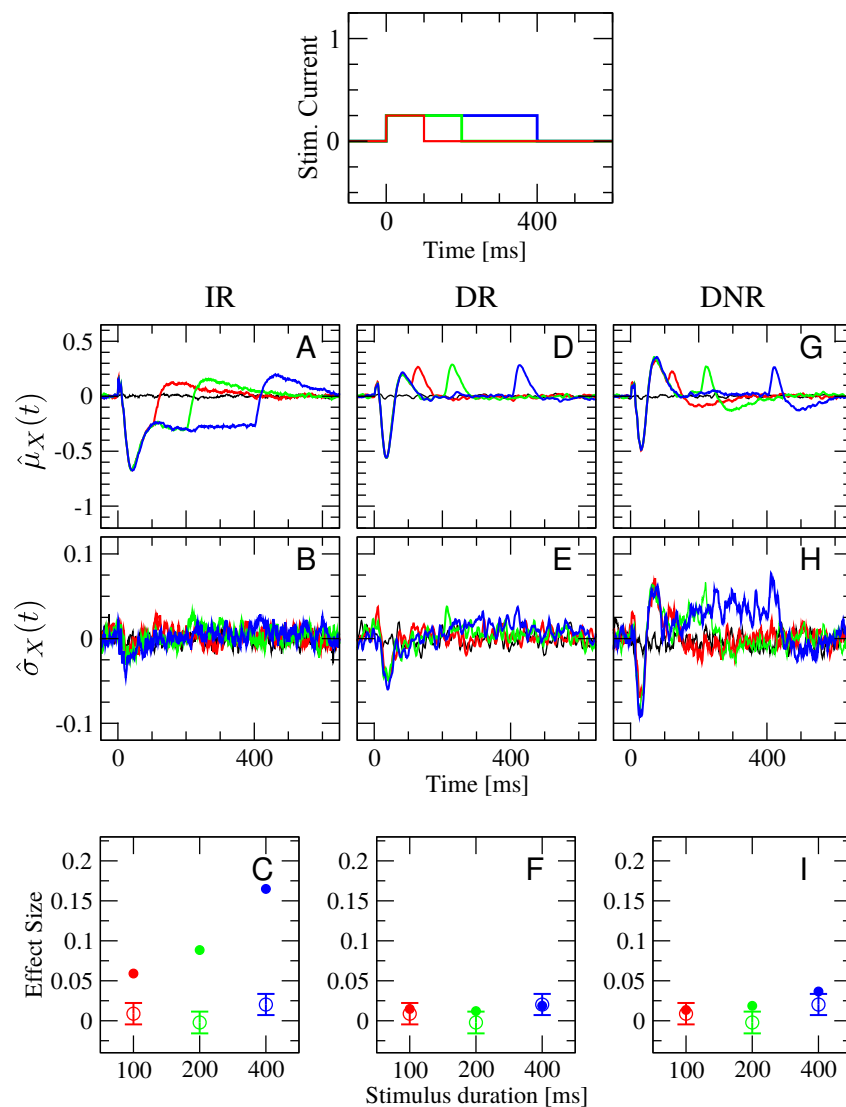
spikes, respectively. The average number of evoked spikes generated by the model is within one standard deviation of the experimental data. However, the spread of the spike count distribution is smaller in the model, which is not surprising, considering the multiple possible noise sources that are not modeled, and that only some of the cellular parameters are randomly distributed in the model.

Figure 7 gives an overview of  $\hat{\mu}_X(t)$ ,  $\hat{\sigma}_X(t)$ , and the effect size measured by all three detectors with stimuli of different duration. In all plots, the color coding is as in the top panel and the black line represents catch trials. The various panels are organized as follows: the first, second, and third column represent results obtained from the IR, DR, and DNR, respectively; the first row shows the time-dependent trial average (expressed as standard score)  $\hat{\mu}_X(t)$ , the second row shows  $\hat{\sigma}_X(t)$ , and the third row displays the effect size defined as in eq. (12) (filled dots) superimposed with the experimental results (open circles with error bars).

The IR activity  $A_{ir}(t)$  is a low-pass filtered sum of a random subset of the spike trains fired by the RS population, as made clear by eq. (1). Hence, the time course of  $\hat{\mu}_{ir}(t)$  upon 400 ms stimulation (fig. 7A, blue line) closely resembles that of the RS population (see fig. 5 and the related discussion on p. 10). Except for the small initial peak due to the spikes fired by the stimulated cell (note that there is a 50% chance in each trial that the stimulated cell is included in the readout activity), the average response of the IR activity is a reduction that shows its deepest dip around  $t = 40$  ms and then relaxes back owing to the spike-frequency adaptation of RS and SOM-LTS populations. In the case of the 100 ms stimulus (red line), the mean deviation goes back to zero shortly after the stimulus is turned off, while  $\hat{\mu}_{ir}(t)$  for the other stimulus durations (green and blue line) settles around -0.3 for the remainder of the respective stimulation time window. In each case,  $\hat{\mu}_{ir}(t)$  shows an overshoot caused by the slower relaxation time scales of the adaptation variables; the overshoot is stronger the longer the signal's duration. The time-dependent standard deviation  $\hat{\sigma}_{ir}(t)$ , shown in fig. 7B, does not display appreciable deviations from zero except for a small dip that corresponds to that observed in the mean. The time courses of  $\hat{\mu}_{ir}(t)$  and  $\hat{\sigma}_{ir}(t)$  suggest that the readout activity has more chances of reaching a lower detection threshold for longer lasting stimuli. Indeed, the effect size measured by the IR strongly depends on the stimulus duration, as it can be seen in fig. 7C (filled circles), which is in contrast with the experimental data (open circles with error bars).

When the DR is used, the picture changes rather drastically. The time-dependent mean  $\hat{\mu}_{dr}(t)$  displays three peaks and one trough in response to all three signals (fig. 7D), even though the last two peaks partly overlap in the case of the 100 ms stimulus (red line). Because the DR considers differences in the IR readout activity, each peak corresponds to an upswing of  $\hat{\mu}_{ir}(t)$  and the trough to the initial sharp drop. The most prominent feature in  $\hat{\sigma}_{dr}(t)$  is the dip right after the stimulus onset, which is the same for all stimuli (fig. 7E). The main difference in the response to the three stimuli is the position of the last peak. Hence, it stands to reason that the DR activity  $A_{dr}(t)$  has similar chances to reach the upper barrier regardless of the signal length. Indeed, the effect size measured by the DR is very similar for the three signals and of the same magnitude as that measured in the experiments (fig. 7F).

The time course of  $\hat{\mu}_{dnr}(t)$  qualitatively resembles that of  $\hat{\mu}_{dr}(t)$  (see fig. 7G), thus suggesting that the DNR does approximately operate as a differentiator. The most evident difference between  $\hat{\mu}_{dnr}(t)$  and  $\hat{\mu}_{dr}(t)$  is the more pronounced undershoot after the last peak, which is likely due to the adaptation current in the readout population. The time-dependent standard deviation  $\hat{\sigma}_{dnr}(t)$  behaves similarly to the mean, and shows peaks and one dip at the same time as the time-dependent mean (fig. 7H). However, a persistent positive shift of  $\hat{\sigma}_{dnr}(t)$  with respect to the zero level can be observed during the entire duration of the signal, which slightly enhances the



**Fig 7. Only simulation results from differentiator readout (DR) and differentiator network readout (DNR) are compatible with the experimental data for the stimulation of a RS cell with three stimuli of equal intensity and different duration (top panel).** First row: standardized deviation from the spontaneous value of the time-dependent mean readout activity eq. (13). Second row: standardized deviation from the spontaneous value of the time-dependent standard deviation of the readout activity eq. (16). Third row: average effect size for the three stimuli. Open circles with error bars are experimental results, which are the same in each panel, and represent the average effect size computed from 119 RS cells (2407 trials in total). First column: integrator readout (IR). Second column: differentiator readout (DR). Third column: differentiator network readout (DNR). The color of each line corresponds to a stimulus as in the top panel (red: 100 ms, green: 200 ms, blue: 400 ms). Black line is catch trial condition (no stimulus).

detectability of the longer signal, as seen in fig. 7I. Still, the increase is moderate and the effect sizes are still close to the experimental values.

To summarize the results of fig. 7, the effect size measured by the IR is larger but strongly depends on the duration of the stimulus. The DR and the DNR detect the stimulus with a reliability that is essentially independent of the signal duration and that is of the same magnitude as the experimental data.

### Effect of stimulus intensity

In the second experiment, we vary the firing rate of the stimulated cell by changing the current intensity while keeping the total area under the step stimulus, i.e. the injected charge, constant. As depicted in the top panel of fig. 8, the stimulus lasting 100 ms (red), 200 ms (green), and 400 ms have an intensity corresponding to 100%, 50%, and 25% of the maximum current, respectively. In this way, the total number of elicited spikes is approximately the same for each stimulus. In the experiment, the three stimuli evoked a firing rate of  $(109 \pm 52)$  Hz,  $(54 \pm 23)$  Hz, and  $(30 \pm 10)$  Hz, respectively. In the model, the average evoked rates are  $(150 \pm 25)$  Hz,  $(103 \pm 20)$  Hz, and  $(50 \pm 12)$  Hz. Note that the maximum current in the model is chosen such that the number of elicited spikes roughly matches the data of the previous experiment (dependence on stimulus duration), which were based on a *different* set of cells. Still, the evoked firing rates in the model are in a similar range as the experimental values.

Figure 8 reports all detection statistics arranged in the same way as before. The shape of  $\hat{\mu}_{ir}(t)$ , the mean response of the IR to the three stimuli, is similar to the previous case, although the initial drop is much stronger here when the stronger stimuli are used (fig. 8A). A clear dip in the time-dependent standard deviation  $\hat{\sigma}_{ir}(t)$  is observed right after the signal onset for the two stronger stimuli (fig. 8B). The pronounced initial response to the stronger signals roughly compensates the shorter duration of the signal in terms of chances of reaching the detection barrier. As a result, the effect size measured by the IR depends weakly on the current intensity and is much stronger than in the experiments, as shown in fig. 8C.

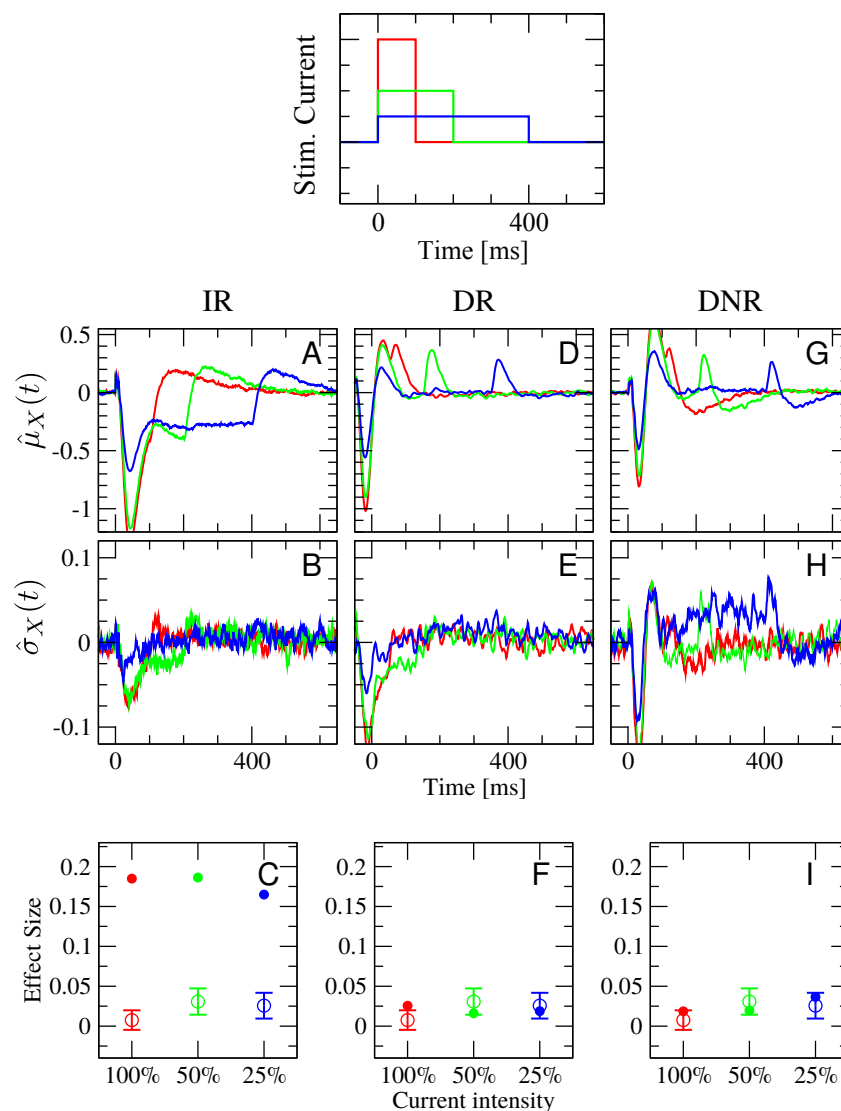
The time-dependent mean of the DR activity,  $\hat{\mu}_{dr}(t)$ , shown in fig. 8D, displays an initial drop followed by two peaks for each of the three stimuli. Here, however, both the first trough and the two subsequent peaks are more pronounced for signals of larger intensity, and the same holds for the initial dip in the time-dependent standard deviation  $\hat{\sigma}_{dr}(t)$  (fig. 8E). As these features have opposing effects and become stronger simultaneously upon growing stimulus intensity, the net effect on the response probability is barely noticeable, as shown in fig. 8F. Furthermore, the effect size is of magnitude comparable with the experimental observations.

Results obtained from the DNR are mostly similar to those from the DR. The principal differences are the mild trough after the last peak in the mean response  $\hat{\mu}_{dnr}(t)$  (fig. 8G) and the increased  $\hat{\sigma}_{dnr}(t)$  in the later part of the response (fig. 8H). The effect size measured from the DNR in the model shows a slight increase upon decreasing current intensity, which, considering the measurement uncertainties, is still consistent with the experimental observations.

Hence, the effect measured by all three readouts shows a weak dependence on the intensity of the stimulus, as it is observed in the experimental data. However, the effect size measured by the DR and the DNR is close to the experimentally observed average effect size of  $\approx 2\%$ , while the effect size obtained from the IR is considerably larger.

### Effect of stimulus regularity

In the third and last *in silico* experiment, random stimuli will be used to evoke irregular spike trains. These stimuli, in accordance with the experimental procedure [6, 47], is a



**Fig 8. Only results from differentiator readout (DR) and differentiator network readout (DNR) are compatible with the experimental data for the stimulation of a RS cell with three stimuli of intensity inversely proportional to duration (top panel).** As in the previous case, results from differentiator readout (DR) and differentiator network readout (DNR) are compatible with experimental data, whereas integrator readout (IR) gives qualitatively different results. Meaning of panels and color coding are as in fig. 7. Open circles with error bars are experimental results, which are the same in each panel, and represent the average effect size computed from 55 RS cells (1469 total trials).

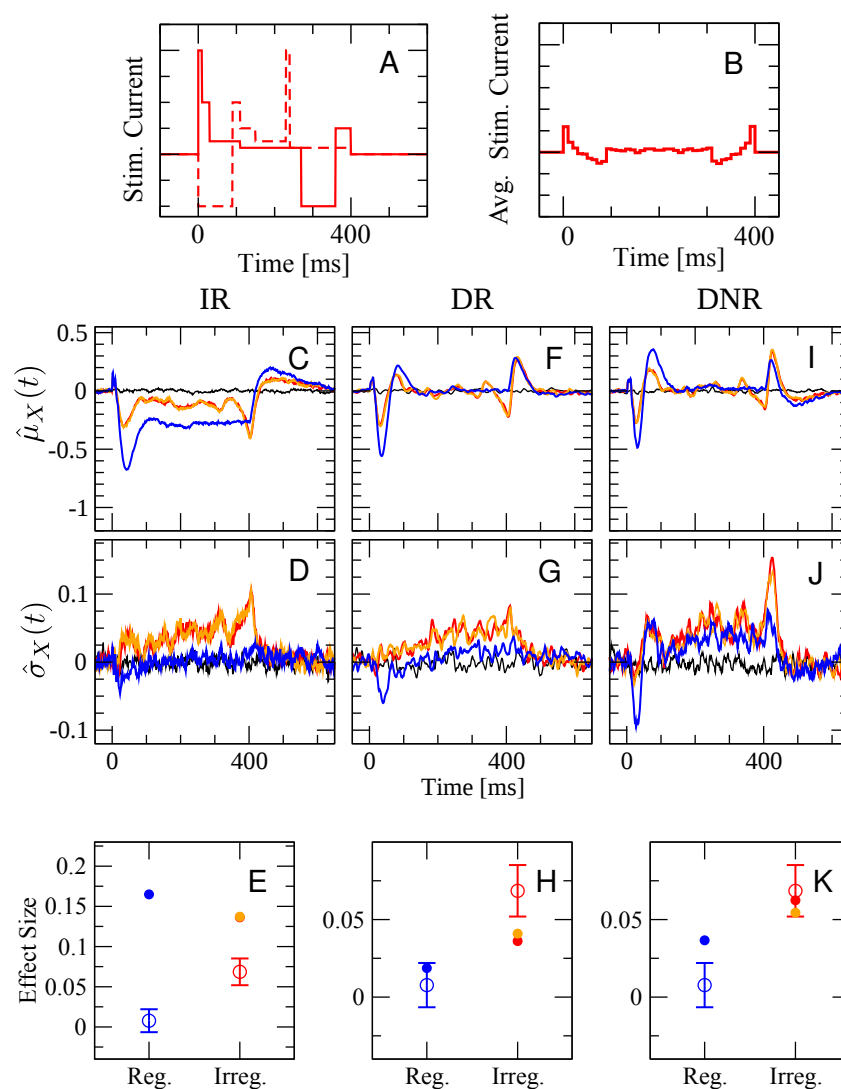
random shuffling of six current steps of length 10, 20, 40, 80, 160, and 90 ms, and with current intensity 100%, 50%, 25%, 12.5%, 6.25%, and -50% of the maximum current, respectively. In other words, each sequence consists of a random permutation of five positive (depolarizing) current steps with intensity inversely proportional to the duration and of one hyperpolarizing step, which inhibits the cell from firing. Two example signals are shown in fig. 9A. Note that stimuli are varied in each trial and not frozen. The irregular stimuli are constructed such that their total duration is 400 ms. The response probability to these stimuli will be compared to that of regular steps of 400 ms at 25% of the maximum current, which was used in both previous cases (plotted in blue). In the experiments, irregular current injections generated spike trains with an average firing rate of  $(24 \pm 11)$  Hz and average CV of  $(1.1 \pm 0.3)$ . In the model, the average rate is  $(27 \pm 5)$  Hz and the average CV  $(1.3 \pm 0.3)$ .

In fig. 9 we compare the simulation results obtained when irregular stimuli are used to those obtained from the 400 ms regular current injection. The latter is shown once more in fig. 9 because it serves here as reference case but it will not be discussed in depth (see discussion above). Results for irregular stimulation are based on two sets of irregular stimuli, constructed by choosing from all possible permutations with equal probability. Despite the large total number of trials used here ( $N_{\text{trials}} = 10000$ ), it is advisable to compare results for two independent sets of stimuli because the large number of possible permutations ( $6! = 720$ ) implies that the number of trials per signal is limited and that finite-size fluctuations due to the particular choice of signals may be non-negligible. Results for the two independent sets of irregular stimuli are plotted in red and orange in fig. 9. We recall that  $\hat{\mu}_X(t)$  and  $\hat{\sigma}_X(t)$  are obtained by averaging over different realizations of the irregular stimuli and are not related to the two particular signals shown in fig. 9A.

The *average* signal is plotted in fig. 9B and displays a peak just after the stimulus onset, followed by a mild trough and then by a plateau barely above the zero level. Just before the end of the stimulation time window, another trough followed by a peak can be seen. Accordingly,  $\hat{\mu}_{\text{ir}}$  shows two dips at the same time where the two peaks in the average signal are seen (fig. 9C). Note that although the first and second half of the average signal are perfectly symmetrical, the second dip in  $\hat{\mu}_{\text{ir}}(t)$  is more pronounced than the first.

Considering how different the average signal is from each particular realization of the irregular sequence, it may seem questionable to average over signals that provoke rather heterogeneous responses. However, the detector as well as the animals in the actual experiments do not know which realization of the irregular sequence is used in each trial. Therefore, this averaging ensemble, in which the stimulus is drawn in each trial, correctly represents the experimental situation and it makes sense to consider its time-dependent mean, as done above. The variability due to the particular realization of the input signal, which mostly averages out in the mean, reveals itself in an increased time-dependent standard deviation,  $\hat{\sigma}_{\text{ir}}(t)$  (fig. 9D), which is above the zero level in the entire stimulation time window and grows further towards the end of the stimulus. This increase of  $\hat{\sigma}_{\text{ir}}(t)$  above the zero level enhances the chances of reaching the detection threshold. As a result, the effect size upon irregular stimulation is large, but not as large as that observed for the regular stimulus (fig. 9E filled dots), which is not consistent with the experimental observations, in which it is the other way around (fig. 9E, open circles with error bars).

The average DR activity in response to the irregular stimuli (fig. 9F, red and orange line) and to the 400 ms regular stimulus (blue line) are rather similar to each other. The main difference is that both the initial trough and peak are somewhat smaller for irregular stimulation. Furthermore, a small dip is observed for irregular stimuli just before the last peak. On the contrary, the standard deviation  $\hat{\sigma}_{\text{dr}}(t)$  is markedly



**Fig 9. Summary of detection statistics upon stimulation of a RS cell with irregular stimuli (random permutation of multiple current steps), compared to regular 400 ms stimulation.** Blue line refers to regular stimulus, as in the previous cases. Red and orange lines represent two different random samples of 10000 stimuli (the stimulus is changed in each trial) from the 720 possible permutations of the six steps (two specific realizations are shown in the top panel, while the average stimulus is shown in the inset). Black line is catch trial condition (no stimulus). First row: standardized deviation from the spontaneous value of the time-dependent mean readout activity eq. (13). Second row: standardized deviation from the spontaneous value of the time-dependent standard deviation of the readout activity eq. (16). Third row: mean effect size for regular and irregular stimulation. Open circles with error bars are experimental results, which are the same in each panel, and represent the average effect size computed from 62 RS cells (1780 trials in total).

different for the two stimulus types (fig. 9G). When the regular stimulus is used,  $\hat{\sigma}_{\text{dr}}(t)$  displays a small dip at the beginning and a slight increase in the later part of the stimulation time window (blue). In contrast, upon irregular stimulation it steadily grows over the entire stimulus range (red and orange lines), similarly to  $\hat{\sigma}_{\text{ir}}(t)$ . Consequently, the average effect size upon irregular stimulation measured by the DR is larger than that upon regular stimulation (fig. 9H, red and orange vs. blue full circles), which is qualitatively consistent with the experimental observations.

The time-dependent mean DNR activity  $\hat{\mu}_{\text{dnr}}(t)$  (shown in fig. 9I) is similar to  $\hat{\mu}_{\text{dr}}(t)$ , whereas  $\hat{\sigma}_{\text{dnr}}(t)$  is generally larger than  $\hat{\sigma}_{\text{dr}}(t)$ , as in the previous cases. Nevertheless, fig. 9J shows that  $\hat{\sigma}_{\text{dnr}}(t)$  is larger in response to irregular stimuli (red and orange line) than its counterpart measured upon regular stimulation (blue line) and displays a strong peak at the end. The larger  $\hat{\sigma}_{\text{dnr}}(t)$  leads to a larger average effect size when irregular stimuli are presented (fig. 9K).

In summary, the simulation results in fig. 9 show that once again the DR and the DNR are consistent with the experimental observation that irregular stimuli are easier to detect than a regular current step, as opposed to the IR, which yields a smaller effect upon irregular stimulation than upon regular stimulation.

## Discussion

The development of juxtacellular stimulation has brought remarkable experimental opportunities, ranging from reliably evoking prescribed spike trains [50–52] to probing the role of a single neuron in the perception of sensory inputs and motor responses [2–6]. Recently, first attempts were made to model the behavioral responses to the stimulation of a single cortical neuron in rodents [25, 26]. However, how the behavioral response probability is influenced by the properties of the injected current [6, 47] is still unaccounted for theoretically. The aim of the present study was to construct a model that can reproduce some of these findings, i.e. that the probability of the behavioral response is not substantially influenced by the duration or by the intensity of a constant stimulus, but it is strongly dependent on whether irregular or regular stimuli are used.

The spiking network model we constructed in this study incorporates several features of the barrel cortex, and its parameters were consistent with the experimental literature. Among the biological processes included in the model, short-term depression and spike frequency adaptation could be expected to oppose slow changes in the input. However, our results indicate that these mechanisms may not be sufficient to explain the data if an integrator readout (IR) is employed. If a differentiator readout (DR or DNR) is used instead, simulation results are in agreement with the data.

How plausible is the scenario of a DNR and the implied transmission delays? The finite time difference used by the differentiator was chosen to be  $\Delta T = 10$  ms because it roughly matches the timescales of the changes induced in the readout population by the current jumps and thus ensures a good signal-to-noise ratio. The intersomatic distance required to achieve such a time difference in the barrel cortex would be approximately 2 mm [53], which is a large but not unphysiological value [54]. Moreover, although in the model the additional delay was entirely assigned to the connections from the BCN to the inhibitory readout population  $\mathcal{I}$  (fig. 2C), it would be possible to distribute the total latency among these connections and those from  $\mathcal{I}$  to the readout population  $\mathcal{S}^B$ . In this way, the disynaptic inhibitory pathway would need, for instance, to travel back and forth with a distance of only one millimeter.

It has been shown that a simple classifier can discriminate highly correlated inputs to a spiking network which exhibits chaotic spontaneous firing activity [55]. It is likely that some explicit training of the readout weights would drastically increase the effect size; in fact, the experimental subjects did undergo a training phase before the

single-neuron stimulation sessions [4, 6]. However, we chose not to explicitly train the readout to detect specific cells, because the training in experiments was performed by employing microstimulation pulses, which intricately affect a large area rather than a specific cell directly [56–59]. Instead, here we assumed that the training had already occurred and that it resulted in the formation of the differentiator circuit. It is possible that training the readout to detect the microstimulation with a suitable learning rule would produce a detector that is also more efficient in detecting the single-cell stimulation than the simple differentiator considered here. It would be substantially more complicated, however, to construct a proper model for the complex and only partially understood cascade of events triggered by cortical microstimulation.

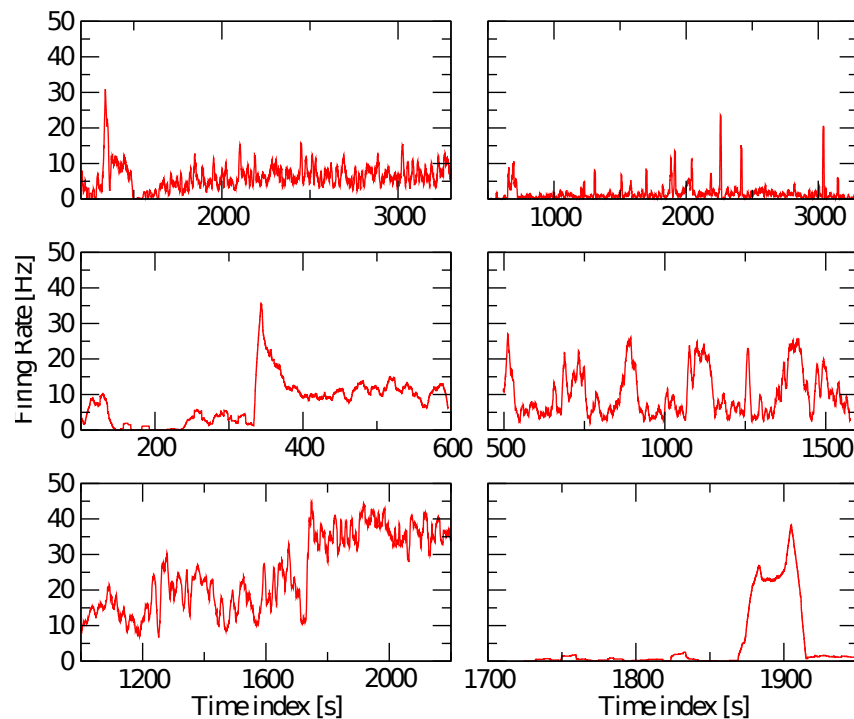
In our model, the main effect of inducing sustained firing in a single excitatory cell was to recruit SOM-LTS cells, which, in turn, inhibited the surrounding excitatory neurons. These results are consistent with the disynaptic inhibition observed *in vitro* [40, 41] and with recordings under anesthesia showing that bursts in pyramidal cells mostly activate surrounding SOM cells, hardly affecting other pyramidal cells or neighboring fast-spiking neurons [48].

Many different classes of inhibitory interneurons have been identified in the neocortex [60]. In this study we decided to include only two types of interneurons to try to limit the already high degree of complexity of the model. Modeling PV neurons is necessary as they are the most common interneuron type and form the backbone of the inhibitory system. SOM-LTS cells were included both because they are the second most common type of interneuron in the barrel cortex [61] and because other experimental studies hinted at their possible functional role when a single pyramidal cell is firing at high rates [40, 41, 48]. Another important class of cortical interneurons is formed by vasoactive intestinal peptide (VIP) neurons. These neurons do not directly provide inhibitory input to pyramidal cells and receive comparatively weak input from pyramidal cells. However, they have been found to make connections to and receive connections from SOM-LTS cells [32]. A recent computational study shows that the mutual inhibition between VIP and SOM-LTS cells can modulate the response of pyramidal cells to external input [62]. On this basis, it may be speculated that VIP neurons also amplify the response to the single-cell stimulation through disinhibition. In other words, SOM-LTS cell activation would inhibit VIP neurons which, in turn, would disinhibit SOM-LTS and amplify the effects of single-cell stimulation. Because VIP neurons are believed to receive top-down input, this conjecture would explain how the attention level of the experimental subjects positively influences the ability to detect the single-cell stimulation [4, 6].

The network model considered here represents the surroundings of the stimulated cells, which justifies the choice of a random unstructured connection topology within each neuronal population. Expanding the model beyond the local scale requires a structured or distance-dependent connectivity profile. Spatial connection profiles and non-random topologies have strong repercussions on the cross-correlations between the spike trains in a network [63–66]. Cross-correlations, in turn, largely contribute to the fluctuations in the pooled activity of a large readout population, as was considered here [25, 49, 67–70], and in general has consequences for the propagation of information about a stimulus to subsequent processing stages [71]. Hence, it is important that future studies investigate how different network topologies influence both the signal (the single-cell stimulation) and the noise (the fluctuations in the network’s activity).

Our results indicate that a readout circuit operating as a differentiator is in better agreement with the experimental data, although the effect size measured by the integrator readout was larger in all considered cases. Recalling that experimental subjects are rewarded for each correct detection, why should they opt for a sub-optimal readout procedure? One possible answer is that the integrator readout only works

better for the stationary situation considered here. In reality, the detector would have to deal with strong and slow variations in the network's firing state. It is possible to have a sense of this kind of variability by looking at fig. 10, which shows the spontaneous firing rate of some neurons during single-cell stimulation recording sessions. The strong changes in firing rate are mostly unrelated to any stimulation event. A readout that integrates spiking in a sliding time window would experience difficulty in distinguishing the possibly small changes induced by the single-cell stimulation from the strong background variations, whereas a differentiator readout might still separate the timescales of stimulus and background noise.



**Fig 10. Nonstationarity of the firing rate of cells measured *in vivo* in the barrel cortex.** Spikes are filtered with a 1 s sliding window.

## Methods

### Detailed description of the recurrent network model

#### Single-neuron properties and total input to neurons

We modeled all neurons as leaky integrate-and-fire point neurons [72]. The  $k$ th neuron follows the differential equation

$$\tau_{m,k} \frac{dv_k}{dt}(t) = -v_k(t) + R_{m,k} I_{\text{total},k}(t), \quad (24)$$

where the membrane time constant  $\tau_{m,k}$  was drawn from a lognormal distribution with mean  $\tau_{m,e} = \tau_{m,s} = 20$  ms if  $k$  is a RS neuron or a SOM-LTS neuron, or with mean  $\tau_{m,i} = 10$  ms if  $k$  is a FS neuron. The standard deviation of all three distributions was set to 20% of the mean. These values are compatible with experimental measurements

for the rat barrel cortex [29,30]. The membrane resistance is  $R_{m,k} = \tau_{m,k}/C_m$ , where the capacitance  $C_m = 150$  pF is assumed equal for all neurons. Equation (24) is complemented with the rule that whenever  $v_k(t)$  reaches the threshold value  $v_{T,k}$ , the neuron emits a spike and  $v_k(t)$  is reset and clamped at  $v_R = 10$  mV for the duration of the refractory period  $\tau_{\text{ref},k}$ . The value of the firing threshold was drawn independently for each neuron from a Gaussian distribution [30] with mean  $v_{T,E} = v_{T,I} = 20$  mV if  $k$  is an RS or FS neuron [29,30] and with mean  $v_{T,S} = 14$  mV if the  $k$ th neuron belongs to the SOM-LTS population, because the distance from resting potential to threshold is 5 mV to 7 mV lower in SOM-LTS neurons than in RS and FS neurons [29]. The standard deviation was set to 10% of the mean for all three neuron types [29]. The refractory time is  $\tau_{\text{ref},k} = \tau_{\text{ref},0} + \hat{\tau}_{\text{ref},k}$ , where  $\tau_{\text{ref},0} = 4$  ms and  $\hat{\tau}_{\text{ref},k}$  was drawn from a lognormal distribution with mean 2 ms and standard deviation 1 ms. The variability in the refractory time serves the purpose of mimicking the variability in the maximum firing rate of neurons [29].

If the  $k$ th neuron belongs to the FS population, its total input current  $I_{\text{total},k}$  is just the sum of the external input and of the recurrent input, i.e. it reads:

$$R_{m,k} I_{\text{total},k}(t) = R_{m,k} [I_{\text{ext},k}(t) + I_{\text{rec},k}(t)], \quad (25)$$

where the first term on the right side of eq. (25) represents the external input from other brain areas and the second term models the recurrent local input from other neurons within the network. If the considered  $k$ th neuron belongs either to the RS or to the SOM-LTS population, the total input current includes an additional adaptation term  $a_k(t)$ :

$$R_{m,k} I_{\text{total},k}(t) = R_{m,k} [I_{\text{ext},k}(t) + I_{\text{rec},k}(t) - a_k(t)]. \quad (26)$$

The adaptation current in the last equation obeys [36,37]:

$$\tau_{a,k} \frac{da_k}{dt}(t) = -a_k(t) + \tau_{a,k} \Delta a_k x_k(t), \quad (27)$$

where  $x_k(t) = \sum_j \delta(t - t_{k,j})$  is the spike train emitted by neuron  $k$ . In other words, every time the neuron fires, the adaptation current jumps by  $\Delta a_k$ . Otherwise, it decays to zero with the time constant  $\tau_{a,k}$ .

Both  $\Delta a_k$  and  $\tau_{a,k}$  are randomly drawn from a lognormal distribution with standard deviation equal to 20% of the mean. For RS neurons, the mean of the two distributions are  $\tau_{a,e} = 100$  ms and  $\Delta a_e = 0.3$  nA, respectively; for SOM-LTS neurons they are  $\tau_{a,s} = 50$  ms and  $\Delta a_s = 0.2$  nA, respectively.

With this choice of parameters, the strength of the spike-frequency adaptation roughly agrees with *in vitro* measurements from the layer IV of the rat barrel cortex [29,35].

## External input to the network

The external input encompasses one constant term and two excitatory Poissonian shot-noise processes:

$$R_{m,k} I_{\text{ext},k}(t) = R_{m,k} I_0 + \tau_{m,k} \left[ \sum_{j=1}^{C_{\text{ext},\text{th},k}} \sum_l J_{k,j,l}^{\text{th}} \delta(t - t_{k,j,l}^{\text{th}}) + \sum_{p=1}^{C_{\text{ext},\text{bc},k}} \sum_q J_{k,p,q}^{\text{bc}} \delta(t - t_{k,p,q}^{\text{bc}}) \right]. \quad (28)$$

The constant term is set to  $R_{m,k}I_0 = 10$  mV for all neurons. The second term in eq. (28) represents the input from the thalamus, and the third mimics incoming spikes from surrounding cortical areas. Because the thalamus has a higher average firing rate, “thalamic” input spikes at times  $t_{k,j,l}^{th}$  occur at an average rate of  $r_{ext,th} = 10$  Hz, while “cortical” input spikes at times  $t_{k,p,q}^{bc}$  have a lower rate of  $r_{ext,bc} = 2$  Hz. The number of external input spike trains depends on the cell type. Experimental studies suggest that SOM cells, in contrast to RS and FS cells, receive only weak input from the thalamus and from distant brain areas [29, 39]. Therefore, if the  $k$ th neuron belongs to the SOM-LTS population, then the number of external inputs is set to zero  $C_{ext,th,k} = C_{ext,bc,k} = 0$ , whereas when  $k$  is a RS or a FS neuron, then  $C_{ext,th,k} = 500$ . Furthermore, dendrites of FS neurons tend to be more localized than those of pyramidal cells, i.e. to receive more input from local RS neurons and less from distant ones. Hence, the number of inputs mimicking the cortical surroundings is  $C_{ext,bc,e} = 2000$  when  $k$  is a RS neuron, and  $C_{ext,bc,i} = 1000$  when  $k$  is a FS neuron. Each input spike causes a PSP drawn independently from an exponential distribution with mean  $J_{ext,e} = 0.1$  mV when  $k$  is a RS neuron, and from an exponential distribution with mean  $J_{ext,i} = 0.2$  mV when  $k$  is a FS neuron, because both thalamic and cortical excitatory postsynaptic potential (EPSP) amplitudes are larger in FS cells than in RS cells [29].

### Recurrent input to RS neurons

The recurrent input term  $I_{rec,k}(t)$  depends on the neuron type. If  $k$  is a RS cell, it is

$$R_{m,k}I_{rec,k}(t) = \tau_{m,k} \left[ \sum_{i \in \mathcal{P}_e(k)} J_{ki}(t)x_i(t - D_{ki}) - \sum_{j \in \mathcal{P}_i(k)} J_{kj}(t)x_j(t - D_{kj}) - \sum_{\ell \in \mathcal{P}_s(k)} J_{k\ell}(t)x_\ell(t - D_{k\ell}) \right], \quad (29)$$

where  $x_i(t - D_{ki})$  indicates the spike train emitted by neuron  $i$ ,  $D_{ki}$  represents the total transmission delay resulting from the axonal propagation, the neurotransmitter diffusion, and the dendritic propagation from neuron  $i$  to neuron  $k$ ,  $J_{ki}$  stands for the synaptic strength from neuron  $i$  to neuron  $k$ , which depends on the spiking history (see below).

Connections to neuron  $k$  originate from three sets of neurons:  $\mathcal{P}_e(k)$ , formed by  $C_{ee} = 300$  randomly selected RS neurons,  $\mathcal{P}_i(k)$ , consisting of  $C_{ei} = 200$  randomly selected FS neurons, and  $\mathcal{P}_s(k)$ , composed of  $C_{es} = 100$  randomly selected SOM-LTS neurons. Hence, the connection probability of RS-to-RS synapses is  $C_{ee}/N_e = 15\%$ , of FS-to-RS and of SOM-LTS-to-RS is  $C_{ei}/N_i = C_{es}/N_s = 50\%$ , consistent with the experimental observations that the connections between RS cells are sparse whereas those between RS and inhibitory cells are dense [27, 29, 40, 73–75]. Transmission delays are drawn uniformly in the range 0.5 ms to 1.0 ms [75]. All synaptic weights in eq. (29) undergo short-term depression (STD)

$$J_{ki}(t) = J_{ki}R_{ki}(t^-), \quad (30)$$

where the maximum coupling amplitudes  $J_{ki}$  (corresponding to the first spike after neuron  $i$  has not been firing for a long time) are drawn independently from an exponential distribution with mean  $J_{ee} = 0.1$  mV for RS-to-RS connections,  $J_{ei} = 0.5$  mV for FS-to-RS coupling, and  $J_{es} = 0.25$  mV for SOM-to-RS connections. The variables  $R_{ki}(t)$  represent the fraction of available synaptic resources, and  $t^-$  indicates that the function is evaluated immediately before a spike. Model and parameters of STD, i.e. the time evolution of the  $R_{ki}(t)$ , are described below.

## Recurrent input to FS neurons

The recurrent input to a FS neuron reads:

$$R_{m,k}I_{\text{rec},k}(t) = \tau_{m,k} \left[ \sum_{i \in \mathcal{Q}_e(k)} J_{ki}(t)x_i(t - D_{ki}) - \sum_{j \in \mathcal{Q}_i(k)} J_{kj}(t)x_j(t - D_{kj}) - \sum_{p \in \mathcal{Q}_s(k)} J_{kp}(t)x_p(t - D_{kp}) + \sum_{\ell \in FS} \hat{J}_{k\ell}x_{\ell}(t - D_{k\ell}) \right], \quad (31)$$

where the first term represents the synaptic input from  $\mathcal{Q}_e(k)$ , a set of  $C_{ie} = 800$  randomly selected RS cells (connection probability  $C_{ie}/N_e = 40\%$ ), the second term is the input from the inhibitory FS presynaptic population  $\mathcal{Q}_i(k)$  with size  $C_{ii} = 200$  (connection probability  $C_{ii}/N_i = 50\%$ ), and the third term represents the inputs from  $\mathcal{Q}_s(k)$ ,  $C_{is} = 50$  randomly selected SOM-LTS neurons (connection probability 25%). All weights appearing in these three terms follow eq. (30) and their peak value is drawn from an exponential distribution of mean  $J_{ie} = 0.2$  mV,  $J_{ii} = 1.0$  mV, and  $J_{is} = 0.1$  mV, respectively. Transmission delays are the same as for RS-to-RS connections. These values reflect the fact that FS neurons receive strong and dense connections both from RS and from FS neurons, and that synapses from SOM to FS neurons are comparatively weaker [29,32]. The fourth and last term in eq. (31) is an effective model for the electrical coupling among FS cells (gap junctions), see next subsection.

## Effective model for gap junctions

Both FS and SOM neurons in the rat somatosensory cortex are coupled by gap junctions [31,33,39,76]. In a simplified picture, gap junctions act as a passive conductance coupling between the membrane voltage of two neurons. The standard way of mimicking the effect of a gap junction between neuron  $k$  and neuron  $\ell$  would be the following additional current for neuron  $k$  [77,78]:

$$R_{m,k}I_{GJ,k\ell} = \gamma_{k\ell}(v_{\ell} - v_k) + \tau_{m,k}\hat{J}_{k\ell}x_{\ell}(t - D_{k\ell}), \quad (32)$$

where  $\gamma_{k\ell}$  is proportional to the Ohmic conductance between the two neurons and modulates the strength of the subthreshold coupling, and  $\hat{J}_{k\ell}$  models the effect of spikes fired by neuron  $\ell$ , which has to be added *ad hoc*, because LIF neurons do not explicitly generate action potentials. Gap junctions typically form between dendrites of different neurons. Therefore, the effect of spikes must travel from the soma along the dendrite of the first neuron to the gap junction and then from it along the dendrite into the soma of the second neuron. For this reason, the time necessary for this propagation can be as large as 0.5 ms [79]. Hence the delay term  $D_{k\ell}$  is drawn from a uniform distribution in the range 0.1 ms to 0.5 ms. As reported in the main text, the subthreshold coupling is completely neglected here, i.e.  $\gamma_{j\ell} = 0$  is set for all neuron pairs. The subthreshold coupling was shown to have a very weak influence on the firing rate, synchrony, and oscillation frequency of networks of LIF neurons, as opposed to the spike-related coupling [34]. The amplitude of gap-junction-related post-synaptic potentials measured in FS neurons of the rat somatosensory cortex is rather variable and, on average, about half as large as excitatory post-synaptic potentials induced by RS neurons [79,80]. Hence,  $\hat{J}_{k\ell}$  was drawn from an exponential distribution of mean  $\hat{J}_{ii} = J_{ie}/2 = 0.05$  mV.

The probability of a gap junction connecting two neighboring inhibitory neurons of the same type (FS with FS and SOM with SOM) is high (60% to 80% [39, 76]). For simplicity, the gap-junction coupling was approximated here as all-to-all (without self-coupling).

### Recurrent input to SOM-LTS neurons

Finally, the recurrent input to a SOM-LTS neuron is

$$R_{m,k}I_{\text{rec},k}(t) = \tau_{m,k} \left[ \sum_{i \in \mathcal{L}_e(k)} J_{ki}(t)x_i(t - D_{ki}) - \sum_{j \in \mathcal{L}_i(k)} J_{kj}(t)x_j(t - D_{kj}) + \sum_{\ell \in \text{SOM}} \hat{J}_{k\ell}x_{\ell}(t - D_{k\ell}) \right]. \quad (33)$$

The three terms in eq. (33) represent the input from excitatory RS neurons, from inhibitory FS neurons, and from gap junctions, respectively. Gap-junctions are modeled in the same way as for FS neurons: their amplitudes and delays are drawn from the same distributions. The first term in eq. (33) is the input from  $C_{se} = 1000$  randomly chosen RS neurons (connection probability  $C_{se}/N_e = 50\%$ ). These are the only connections that undergo short-term *facilitation* instead of depression, and for which random *transmission failures* were modeled (details on the model below). The static baseline amplitudes  $J_{ki}$  of each synapse are drawn independently from an exponential distribution and have mean  $J_{se} = 0.1$  mV. The second term in eq. (33) represents the input from  $C_{si} = 100$  randomly selected FS neurons (connection probability  $C_{si}/N_i = 25\%$ ). These connections have the average maximum strength  $J_{si} = 0.25$  mV, undergo short-term depression and obey eq. (30). Chemical synapses between SOM neurons are infrequent and weak [29, 39] and were omitted for simplicity.

### Model of short-term depression

Except for those connecting RS to SOM neurons, all chemical synapses in the model undergo short-term depression. Each weight  $J_{kj}(t)$  has a time dependence described by

$$J_{ki}(t) = J_{ki}R_{ki}(t^-), \quad (34)$$

where the variable  $R_{ki}(t)$  stands for the fraction of available synaptic resources and is described by the standard model by Tsodyks and Markram [43],

$$\dot{R}_{ki}(t) = \frac{1 - R_{ki}(t)}{\tau_D} - U_{se}R_{ki}(t^-) \sum_j \delta(t - \hat{t}_{i,j}), \quad (35)$$

where  $\hat{t}_{i,j}$  are the times at which the spikes of neuron  $i$  arrive at the synapse, and  $t^-$  indicates that the function is evaluated at  $t - \varepsilon$  ( $\varepsilon > 0$  is a small positive number), i.e. just before a spike. The parameter  $U_{se}$  represents the release probability and  $\tau_D$  is the recovery time scale. Note that the time evolution of  $R_{ki}(t)$  depends on the spike times of the presynaptic neuron  $i$  only. Hence, if  $\tau_D$  and  $U_{se}$  do not depend on  $k$ , the time course of each variable  $R_{ki}(t)$  is a time-shifted copy of a single master variable  $R_i(t)$

$$R_{ki}(t) = R_i(t - D_{ki}), \quad (36)$$

where  $R_i(t)$  obeys the same equation as  $R_{ki}(t)$ , except that the arrival times  $\hat{t}_{i,j}$  in eq. (35) are replaced by  $t_{i,j}$ , the spike times of neuron  $i$ . Here, it is assumed that  $\tau_D$  and  $U_{se}$  only depend on the type of the source and target neuron, but not on the identity of the particular neuron within a population so that eq. (36) holds. In this way, the actual number of dynamic variables required to simulate the network is reduced from one variable per synapse to one variable per neuron, which is an enormous computational advantage.

The parameter values chosen to model strong depression (all connections depicted in blue in fig. 1) are  $\tau_D = \tau_{D,s} = 150$  ms and  $U_{se} = U_{se,s} = 0.2$ . With this choice, the eighth PSP of a 40 Hz pre-synaptic regular spike train is about one half of the maximal amplitude [29]. Most chemical synapses in the barrel cortex are depressing ([29, 42, 53]). However, inhibitory synapses originating from SOM-LTS neurons and terminating onto RS neurons show only weak depression or slight facilitation. Here, these connections are modeled as mildly depressing (fig. 1, light blue) by setting  $\tau_D = \tau_{D,w} = 50$  ms and  $U_{se,w} = 0.05$ . For simplicity, also SOM-to-FS connections were given the same STD parameters.

### Short-term facilitation and transmission failures

Excitatory synapses from RS neurons to SOM-LTS neurons (depicted in red in fig. 1) are strongly facilitating ([29, 40, 41]). If the parameter  $U_{se}$  in eq. (36) is turned into a dynamical variable,  $u(t)$ , facilitating synapses can be described [44, 81]. The amplitude of the PSPs is proportional to the product  $R(t)u(t)$ . Considering the connection from the RS neuron  $i$  to the SOM-LTS neuron  $k$ , the time evolution of the synaptic amplitude is described by (note that the conventions have been slightly changed with respect to ref. [44]):

$$J_{ki}(t) = J_{ki}R_i(t^- - D_{ki})\frac{u_i(t^+ - D_{ki})}{U_b}, \quad (37)$$

where  $t^+$  means that the function is evaluated at  $t + \varepsilon$ , i.e. the value of  $u_i(t)$  just after the occurrence of a spike. The variables  $R_i(t)$  and  $u_i(t)$  obey

$$\dot{u}_i(t) = \frac{U_b - u_i(t)}{\tau_F} + (1 - u_i(t^-))U \sum_j \delta(t - t_{i,j}) \quad (38)$$

$$\dot{R}_i(t) = \frac{1 - R_i(t)}{\tau_D} - u_i(t^-)R_i(t^-) \sum_j \delta(t - t_{i,j}), \quad (39)$$

where  $t_{i,j}$  indicate the spike times of neuron  $i$ . The first term in eq. (38) governs the relaxation of the facilitation variable to the baseline level  $U_b$  and the second term determines a positive jump upon each pre-synaptic spike. The time evolution of the depression variable  $R_i(t)$  has the same form of eq. (35), i.e. a purely depressing synapse, except that the release probability is the time-dependent function  $u_i(t)$ . The choice of the parameters  $U, \tau_F, \tau_D$  dictates whether, for a given firing rate, the synapse facilitates, depresses, or both [44]. Here, the four parameters appearing in eqs. (38) and (39) were set as follows:  $\tau_F = 300$  ms,  $\tau_D = \tau_{D,f} = 100$  ms,  $U_b = 0.01$ , and  $U = 0.03$ . With this choice and for a pre-synaptic stimulation of 40 Hz, the synapse is purely facilitating [29].

RS-to-SOM synapses stand out from all other synapses considered here because of a much higher occurrence of transmission failures at low presynaptic firing rates (the average failure rate is  $\approx 10\%$  for RS-to-RS synapses,  $\approx 5\%$  for synapses to and from FS neurons, and  $\gtrsim 50\%$  for RS to SOM-LTS synapses [29]). However, the failure rate of RS-to-SOM-LTS synapses decreases to  $\approx 10\%$  upon repeated stimulation at 40 Hz (failure rates for other synapses weakly depend on the presynaptic firing rate [29]). Here,

transmission failures are modeled only for RS-to-SOM synapses via a stochastic binary variable  $S(p_f)$ :

$$S(p_f) = \begin{cases} 1 & \text{with probability } 1 - p_f \\ 0 & \text{with probability } p_f \end{cases}, \quad (40)$$

where  $p_f$  describes the failure rate, which obeys the following dynamical equation

$$\dot{p}_f(t) = \frac{p_{f,\text{rest}} - p_f(t)}{\tau_f} - G(p_f, \Delta p_f, p_{\min}) \sum_j \delta(t - t_{i,j}). \quad (41)$$

In the last equation,  $p_{f,\text{rest}} = 0.5$  is the baseline failure rate. Upon each presynaptic spike, the failure rate decreases by  $G(p_f, \Delta p_f, p_{\min})$  and relaxes back to the baseline value with the time constant  $\tau_f = 250$  ms. The size of each downward jump is  $\Delta p_f = 0.1$  but is constrained to values above  $\Delta p_f = p_{\min} = 0.1$ , a condition which is imposed by the piecewise linear function

$$G(p_f, \Delta p_f, p_{\min}) = \begin{cases} 0 & \text{if } p_f \leq p_{\min} \\ p_f - p_{\min} & \text{if } p_{\min} < p_f < p_{\min} + \Delta p_f \\ \Delta p_f & \text{if } p_{\min} + \Delta p_f \leq p_f \end{cases}. \quad (42)$$

In the end, the synaptic weight from the RS neuron  $i$  to the SOM-LTS neuron  $k$  obeys the following equation:

$$J_{ki}(t) = J_{ki}R_i(t^- - D_{ki}) \frac{u_i(t^+ - D_{ki})}{U_b} S(p_{f,i}(t^- - D_{ki})). \quad (43)$$

If the average effect of synaptic failures is taken into account, a 40 Hz presynaptic stimulation causes the eighth PSP to be about eight times larger than the first, which is in a reasonable qualitative agreement with the strong amplification measured *in vitro* [29, 41].

## Detailed description of the readout network model

The differentiator network readout (DNR) consists of one population of  $N_B = 10\,000$  RS neurons ( $\mathcal{S}^B$ ) and of one population of 2000 FS neurons ( $\mathcal{I}$ ). Each neuron in the DNR follows the same dynamical equation as its counterpart within the BCN and receives feedforward input from  $\hat{C} = 1000$  randomly selected RS neurons of the BCN. The number of independent Poisson processes mimicking thalamic input is  $C_{\text{ext,th}}^R = 250$  and the constant external input is  $R_m I_0 = 15$  mV. The number of independent Poisson processes representing cortical input is the same as for the BCN. Each neuron in the DNR receives 200 random connections from the local inhibitory population ( $\mathcal{I}$ ). Hence, the only recurrent connections within the DNR are inhibitory.

All connections from the BCN to the DNR and within the DNR are randomly drawn from the same distributions as for the corresponding class of neurons within the BCN, except for the connections from the inhibitory readout population  $\mathcal{I}$  to the excitatory readout population  $\mathcal{S}^B$ , the average strength of which,  $J_{ei}^R$ , is tuned to a value that enables the DNR to approximate the function of a differentiator circuit, as explained in the following.

Referring to fig. 4, we have to calculate the value of  $J_{ei}^R$  such that the input  $\Delta\mu_{\mathcal{I}}$  via the indirect path to the readout population  $\mathcal{S}^B$  equals a negative and temporally delayed image of the direct input  $\Delta\mu_e$ . This value can approximately be determined by the following linear-response calculation.

Consider a perturbation of the firing rate of the RS neurons within the BCN and indicate it with  $\Delta r_e$ . We assume that the perturbation is slow compared to the most

important system time constants so that time-dependencies can be neglected. As a consequence of the firing rate perturbation within the BCN, the mean input from the BCN to  $\mathcal{S}^B$  changes by

$$\Delta\mu_e = \tau_{m,e} J_{ee}^{FF} \bar{R}(r_e) \hat{C} \Delta r_e, \quad (44)$$

where the term

$$\bar{R}(r) = \frac{1}{1 + \tau_{D,s} U_{se,s} r} \quad (45)$$

represents the average effect of the short-term depression (STD), given a presynaptic firing rate  $r$ . In eq. (44),  $J_{ee}^{FF}$  represents the average synaptic strength of the connections from the BCN to the excitatory readout population  $\mathcal{S}^B$ . Likewise, the mean input from  $\mathcal{I}$  to  $\mathcal{S}^B$  changes by

$$\Delta\mu_{\mathcal{I}} = -\tau_{m,e} J_{ei}^R \bar{R}(r_{\mathcal{I}}) C_{ei}^R \Delta r_{\mathcal{I}}, \quad (46)$$

where  $C_{ei}^R = 200$  is the number of input connections from  $\mathcal{I}$  to  $\mathcal{S}^B$  per postsynaptic neuron, and  $\Delta r_{\mathcal{I}}$  is the change in the firing rate of the population  $\mathcal{I}$  from the spontaneous value  $r_{\mathcal{I}}$ .

The linear-response approximation of  $\Delta r_{\mathcal{I}}$  is

$$\Delta r_{\mathcal{I}} = \frac{d\phi_{sn}}{d\mu} \tau_{m,i} \left( J_{ie}^{FF} \bar{R}(r_e) \hat{C} \Delta r_e - J_{ii}^R \bar{R}(r_{\mathcal{I}}) C_{ii}^R \Delta r_{\mathcal{I}} \right), \quad (47)$$

where  $d\phi_{sn}/d\mu$  is the so-called DC susceptibility, i.e. the linear response of the firing rate of a LIF neuron to a slow change in its total mean input  $\mu$ . The value of the DC susceptibility can be approximated by taking the derivative of the firing rate of a white-shot-noise-driven LIF neuron [82] with respect to its mean input. The explicit expression for  $d\phi_{sn}/d\mu$  with a non-zero refractory period can be found in the first appendix of [49].

First, eq. (47) can be solved for  $\Delta r_{\mathcal{I}}$  and substituted into eq. (46). Then, we require that the perturbation in the mean input from direct and indirect pathways cancel each other (see fig. 4). In other words, we impose  $\Delta\mu_e + \Delta\mu_{\mathcal{I}} = 0$  and finally solve for  $J_{ei}^R$ , which yields

$$J_{ei}^R = \frac{J_{ee}^{FF} \left( 1 + \tau_{m,i} \frac{d\phi_{sn}}{d\mu} J_{ii}^R \bar{R}(r_{\mathcal{I}}) C_{ii}^R \right)}{\tau_{m,i} \frac{d\phi_{sn}}{d\mu} J_{ie}^{FF} C_{ei}^R \bar{R}(r_{\mathcal{I}})}. \quad (48)$$

The only unknown quantity on the right hand side of eq. (48) is  $r_{\mathcal{I}}$ , the spontaneous firing rate of  $\mathcal{I}$ . This firing rate can be estimated from the numerical solution of the following self-consistency condition:

$$r_{\mathcal{I}} = \phi_{sn}(J_{ee}^{FF}, J_{ii}^R, r_{tot}^{in}, C_{ii}^R, r_{\mathcal{I}}, I_{ext}), \quad (49)$$

where  $r_{tot}^{in} = \hat{C} r_e + C_{ext,bc,e}^R r_{ext,bc} + C_{ext,th,e}^R r_{ext,th}$  is the total excitatory input rate to  $\mathcal{I}$  and  $\phi_{sn}(a_e, a_i, R_e, R_i, I_0)$  is the firing rate of a LIF neuron driven by white shot-noise with exponentially distributed weights [82]. The first two arguments,  $a_e, a_i$  are the excitatory and inhibitory mean input weights, respectively. The third and fourth argument  $R_e, R_i$  are the input rates of the excitatory and inhibitory input, respectively. The last argument  $I_0$  is the constant input. The explicit expression with non-zero input current and non-zero refractory period is

$$\begin{aligned} \phi_{sn}(a_e, a_i, R_e, R_i, I_0) = & \left( \tau_{ref} \right. \\ & \left. + \tau_m \int_0^{1/a_e} \frac{ds}{s} Z_0^{-1}(s) \left[ \frac{e^{s\hat{v}_T}}{1 - a_e s} - e^{s\hat{v}_R} \right] \right)^{-1} \end{aligned} \quad (50)$$

where  $\hat{v}_R = v_R - R_m I_0$ ,  $\hat{v}_T = v_T - R_m I_0$ , and  $Z_0^{-1}(s) = (1 - a_e s)^{\tau_m R_e} (1 + a_i s)^{\tau_m R_i}$ .

Substituting numerical values in eq. (48) reveals that  $J_{ei}^R = 0.65$  mV approximately satisfies the imposed condition.

## Experimental data

The experimental data appearing in figs. 7 to 9 are a part of the dataset of references [6, 47]. In particular, the data shown in fig. 7 are the average effect size (for each stimulus duration) of most cells shown in Fig. 18A of reference [47] (800 ms stimuli were not used in the present study); the experimental effect size in fig. 8 is the average for each stimulus intensity (and duration) of the cells appearing in Fig. 14A of reference [47]; the average effect size for regular and irregular stimulation of fig. 9 is based on the same dataset used for Fig. 21C of reference [47]. For experimental procedures, we refer to [6, 47].

## Acknowledgments

This work was supported by the German Research Foundation (DFG), GRK 1589/2 and by the German Federal Ministry of Education and Research (Bernstein Center II, grant no. 01GQ1001A).

## References

1. Feldmeyer D, Brecht M, Helmchen F, Petersen CCH, Poulet JFA, Staiger JF, et al. Barrel cortex function. *Prog Neurobiol.* 2013;103:3–27. doi:10.1016/j.pneurobio.2012.11.002.
2. Doron G, Brecht M. What single-cell stimulation has told us about neural coding. *Philos Trans R Soc Lond B Biol Sci.* 2015;370:20140204. doi:10.1098/rstb.2014.0204.
3. Brecht M, Schneider M, Sakmann B, Margrie TW. Whisker movements evoked by stimulation of single pyramidal cells in rat motor cortex. *Nature.* 2004;427(6976):704–710. doi:10.1038/nature02266.
4. Houweling AR, Brecht M. Behavioural report of single neuron stimulation in somatosensory cortex. *Nature.* 2008;451(7174):65–68. doi:10.1038/nature06447.
5. Voigt BC, Brecht M, Houweling AR. Behavioral detectability of single-cell stimulation in the ventral posterior medial nucleus of the thalamus. *J Neurosci.* 2008;28(47):12362–12367. doi:10.1523/JNEUROSCI.3046-08.2008.
6. Doron G, von Heimendahl M, Schlattmann P, Houweling AR, Brecht M. Spiking irregularity and frequency modulate the behavioral report of single-neuron stimulation. *Neuron.* 2014;81(3):653–663. doi:10.1016/j.neuron.2013.11.032.
7. Meyer HS, Wimmer VC, Oberlaender M, de Kock CPJ, Sakmann B, Helmstaedter M. Number and laminar distribution of neurons in a thalamocortical projection column of rat vibrissa cortex. *Cereb Cortex.* 2010;20:2277–2286. doi:10.1093/cercor/bhq067.
8. Softky WR, Koch C. The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. *J Neurosci.* 1993;13(1):334–350.

9. London M, Roth A, Beeren L, Häusser M, Latham PE. Sensitivity to perturbations in vivo implies high noise and suggests rate coding in cortex. *Nature*. 2010;466(7302):123–127. doi:10.1038/nature09086.
10. Treves A. Mean-field analysis of neuronal spike dynamics. *Network*. 1993;4(3):259–284.
11. Amit DJ, Brunel N. Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. *Cereb Cortex*. 1997;7(3):237–252.
12. Kumar A, Schrader S, Aertsen A, Rotter S. The high-conductance state of cortical networks. *Neural Comput*. 2008;20:1–43. doi:10.1162/neco.2008.20.1.1.
13. Brunel N. Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *J Comput Neurosci*. 2000;8(3):183–208. doi:10.1023/A:1008925309027.
14. Ostojic S. Two types of asynchronous activity in networks of excitatory and inhibitory spiking neurons. *Nat Neurosci*. 2014;17(4):594–600. doi:10.1038/nn.3658.
15. Wieland S, Bernardi D, Schwalger T, Lindner B. Slow fluctuations in recurrent networks of spiking neurons. *Phys Rev E*. 2015;92(4):040901. doi:10.1103/PhysRevE.92.040901.
16. Renart A, de la Rocha J, Bartho P, Hollender L, Parga N, Reyes A, et al. The asynchronous state in cortical circuits. *Science*. 2010;327(5965):587–590. doi:10.1126/science.1179850.
17. Harris KD, Thiele A. Cortical state and attention. *Nat Rev Neurosci*. 2011;12(9):509–523. doi:10.1038/nrn3084.
18. van Meegen A, Lindner B. Self-Consistent Correlations of Randomly Coupled Rotators in the Asynchronous State. *Phys Rev Lett*. 2018;121:258302. doi:10.1103/PhysRevLett.121.258302.
19. Vellmer S, Lindner B. Theory of spike-train power spectra for multidimensional integrate-and-fire neurons. *Phys Rev Research*. 2019;1(2):023024.
20. Sanzeni A, Histed MH, Brunel N. Response nonlinearities in networks of spiking neurons. *bioRxiv*. 2019;doi:10.1101/856831.
21. Doiron B, Chacron MJ, Maler L, Longtin A, Bastian J. Inhibitory feedback required for network oscillatory responses to communication but not prey stimuli. *Nature*. 2003;421(6922):539–543.
22. Kumar A, Rotter S, Aertsen A. Conditions for propagating synchronous spiking and asynchronous firing rates in a cortical network model. *J Neurosci*. 2008;28(20):5268–5280.
23. Ledoux E, Brunel N. Dynamics of networks of excitatory and inhibitory neurons in response to time-dependent inputs. *Front Comput Neurosci*. 2011;5:25. doi:10.3389/fncom.2011.00025.
24. Alijani AK, Richardson MJE. Rate response of neurons subject to fast or frozen noise: from stochastic and homogeneous to deterministic and heterogeneous populations. *Phys Rev E*. 2011;84:011919. doi:10.1103/PhysRevE.84.011919.

25. Bernardi D, Lindner B. Optimal detection of a localized perturbation in random networks of integrate-and-fire neurons. *Phys Rev Lett.* 2017;118:268301. doi:10.1103/PhysRevLett.118.268301.
26. Bernardi D, Lindner B. Detecting single-cell stimulation in a large network of integrate-and-fire neurons. *Phys Rev E.* 2019;99(3):032304.
27. Avermann M, Tamm C, Mateo C, Gerstner W, Petersen CCH. Microcircuits of excitatory and inhibitory neurons in layer 2/3 of mouse barrel cortex. *J Neurophysiol.* 2012;107(11):3116–3134. doi:10.1152/jn.00917.2011.
28. Schoonover CE, Tapia JC, Schilling VC, Wimmer V, Blazeski R, Zhang W, et al. Comparative strength and dendritic organization of thalamocortical and corticocortical synapses onto excitatory layer 4 neurons. *J Neurosci.* 2014;34(20):6746–6758. doi:10.1523/JNEUROSCI.0305-14.2014.
29. Beierlein M, Gibson JR, Connors BW. Two dynamically distinct inhibitory networks in layer 4 of the neocortex. *J Neurophysiol.* 2003;90(5):2987–3000. doi:10.1152/jn.00283.2003.
30. Harrison PM, Badel L, Wall MJ, Richardson MJ. Experimentally verified parameter sets for modelling heterogeneous neocortical pyramidal-cell populations. *PLoS Comput Biol.* 2015;11(8):e1004165. doi:10.1371/journal.pcbi.1004165.
31. Beierlein M, Gibson JR, Connors BW. A network of electrically coupled interneurons drives synchronized inhibition in neocortex. *Nat Neurosci.* 2000;3:904–910. doi:10.1038/78809.
32. Pfeffer CK, Xue M, He M, Huang ZJ, Scanziani M. Inhibition of inhibition in visual cortex: the logic of connections between molecularly distinct interneurons. *Nat Neurosci.* 2013;16(8):1068–1076. doi:10.1038/nn.3446.
33. Galarreta M, Hestrin S. A network of fast-spiking cells in the neocortex connected by electrical synapses. *Nature.* 1999;402(6757):72. doi:10.1038/47029.
34. Holzbecher A, Kempter R. Interneuronal gap junctions increase synchrony and robustness of hippocampal ripple oscillations. *Eur J Neurosci.* 2018;48(12):3446–3465.
35. Gottlieb JP, Keller A. Intrinsic circuitry and physiological properties of pyramidal neurons in rat barrel cortex. *Exp Brain Res.* 1997;115(1):47–60. doi:10.1007/pl00005684.
36. Benda J, Herz AV. A universal model for spike-frequency adaptation. *Neural Comput.* 2003;15(11):2523–2564. doi:10.1162/089976603322385063.
37. Schwalger T, Lindner B. Patterns of interval correlations in neural oscillators with adaptation. *Front Comput Neurosci.* 2013;7:164.
38. Gutkin B, Zeldenrust F. Spike frequency adaptation. *Scholarpedia.* 2014;9(2):30643. doi:10.4249/scholarpedia.30643.
39. Gibson JR, Beierlein M, Connors BW. Two networks of electrically coupled inhibitory neurons in neocortex. *Nature.* 1999;402:75–79. doi:10.1038/47035.
40. Silberberg G, Markram H. Disynaptic inhibition between neocortical pyramidal cells mediated by Martinotti cells. *Neuron.* 2007;53(5):735–746. doi:10.1016/j.neuron.2007.02.012.

41. Kapfer C, Glickfeld LL, Atallah BV, Scanziani M. Supralinear increase of recurrent inhibition during sparse activity in the somatosensory cortex. *Nat Neurosci.* 2007;10(6):743–753. doi:10.1038/nn1909.
42. Lefort S, Petersen CCH. Layer-Dependent Short-Term Synaptic Plasticity Between Excitatory Neurons in the C2 Barrel Column of Mouse Primary Somatosensory Cortex. *Cereb Cortex.* 2017;27:3869–3878. doi:10.1093/cercor/bhx094.
43. Tsodyks MV, Markram H. The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proc Natl Acad Sci USA.* 1997;94:719–723.
44. Tsodyks M, Pawelzik K, Markram H. Neural networks with dynamic synapses. *Neural Comput.* 1998;10:821–835. doi:10.1162/089976698300017502.
45. Middleton JW, Omar C, Doiron B, Simons DJ. Neural correlation is stimulus modulated by feedforward inhibitory circuitry. *J Neurosci.* 2012;32:506–518. doi:10.1523/JNEUROSCI.3474-11.2012.
46. Gentet LJ, Kremer Y, Taniguchi H, Huang ZJ, Staiger JF, Petersen CCH. Unique functional properties of somatostatin-expressing GABAergic neurons in mouse barrel cortex. *Nat Neurosci.* 2012;15:607–612. doi:10.1038/nn.3051.
47. Doron G. Psychophysical characterization of single neuron stimulation effects in rat barrel cortex. Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät; 2012.
48. Kwan AC, Dan Y. Dissection of cortical microcircuits by single-neuron stimulation in vivo. *Curr Biol.* 2012;22(16):1459–1467. doi:10.1016/j.cub.2012.06.007.
49. Bernardi D. Detecting Single-Cell Stimulation in Recurrent Networks of Integrate-and-Fire Neurons. Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät; 2019.
50. Dose J, Doron G, Brecht M, Lindner B. Noisy Juxtacellular Stimulation In Vivo Leads to Reliable Spiking and Reveals High-Frequency Coding in Single Neurons. *J Neurosci.* 2016;36:11120–11132. doi:10.1523/JNEUROSCI.0787-16.2016.
51. Stüttgen MC, Nonkes LJP, Geis HRAP, Tiesinga PH, Houweling AR. Temporally precise control of single-neuron spiking by juxtacellular nanostimulation. *J Neurophysiol.* 2017;117(3):1363–1378.
52. Dose J, Lindner B. Evoking prescribed spike times in stochastic neurons. *Physical Review E.* 2017;96(3):032109.
53. Helmstaedter M, Staiger JF, Sakmann B, Feldmeyer D. Efficient recruitment of layer 2/3 interneurons by layer 4 input in single columns of rat somatosensory cortex. *J Neurosci.* 2008;28(33):8273–8284. doi:10.1523/JNEUROSCI.5701-07.2008.
54. Schnepel P, Kumar A, Zohar M, Aertsen A, Boucsein C. Physiology and Impact of Horizontal Connections in Rat Neocortex. *Cereb Cortex.* 2014;25:3818–3835. doi:10.1093/cercor/bhu265.

55. Lajoie G, Lin KK, Thivierge JP, Shea-Brown E. Encoding in Balanced Networks: Revisiting Spike Patterns and Chaos in Stimulus-Driven Systems. *PLoS Comput Biol*. 2016;12:e1005258. doi:10.1371/journal.pcbi.1005258.
56. Butovas S, Schwarz C. Spatiotemporal effects of microstimulation in rat neocortex: a parametric study using multielectrode recordings. *J Neurophysiol*. 2003;90(5):3024–3039. doi:10.1152/jn.00245.2003.
57. Butovas S, Hormuzdi SG, Monyer H, Schwarz C. Effects of electrically coupled inhibitory networks on local neuronal responses to intracortical microstimulation. *J Neurophysiol*. 2006;96(3):1227–1236. doi:10.1152/jn.01170.2005.
58. Histed MH, Bonin V, Reid RC. Direct activation of sparse, distributed populations of cortical neurons by electrical microstimulation. *Neuron*. 2009;63(4):508–522. doi:10.1016/j.neuron.2009.07.016.
59. Overstreet CK, Klein JD, Tillery SIH. Computational modeling of direct neuronal recruitment during intracortical microstimulation in somatosensory cortex. *J Neural Eng*. 2013;10(6):066016. doi:10.1088/1741-2560/10/6/066016.
60. Tremblay R, Lee S, Rudy B. GABAergic Interneurons in the Neocortex: From Cellular Properties to Circuits. *Neuron*. 2016;91:260–292. doi:10.1016/j.neuron.2016.06.033.
61. Meyer HS, Schwarz D, Wimmer VC, Schmitt AC, Kerr JND, Sakmann B, et al. Inhibitory interneurons in a cortical column form hot zones of inhibition in layers 2 and 5A. *Proc Natl Acad Sci USA*. 2011;108(40):16807–16812. doi:10.1073/pnas.1113648108.
62. Hertäg L, Sprekeler H. Amplifying the redistribution of somato-dendritic inhibition by the interplay of three interneuron types. *PLoS Comput Biol*. 2019;15(5):e1006999. doi:10.1371/journal.pcbi.1006999.
63. Kriener B, Helias M, Aertsen A, Rotter S. Correlations in spiking neuronal networks with distance dependent connections. *J Comput Neurosci*. 2009;27:177–200. doi:10.1007/s10827-008-0135-1.
64. Trousdale J, Hu Y, Shea-Brown E, Josić K. Impact of network structure and cellular response on spike time correlations. *PLoS Comput Biol*. 2012;8(3):e1002408. doi:10.1371/journal.pcbi.1002408.
65. Pernice V, Staude B, Cardanobile S, Rotter S. Recurrent interactions in spiking networks with arbitrary topology. *Phys Rev E*. 2012;85:031916. doi:10.1103/PhysRevE.85.031916.
66. Rosenbaum R, Smith MA, Kohn A, Rubin JE, Doiron B. The spatial structure of correlated neuronal variability. *Nat Neurosci*. 2017;20(1):107–114. doi:10.1038/nn.4433.
67. Helias M, Tetzlaff T, Diesmann M. Echoes in correlated neural systems. *New J Phys*. 2013;15:023002. doi:10.1088/1367-2630/15/2/023002.
68. Dipoppa M, Gutkin BS. Correlations in background activity control persistent state stability and allow execution of working memory tasks. *Front Comput Neurosci*. 2013;7:139. doi:10.3389/fncom.2013.00139.

69. Helias M, Tetzlaff T, Diesmann M. The Correlation Structure of Local Neuronal Networks Intrinsically Results from Recurrent Dynamics. *PLoS Comput Biol.* 2014;10(1):1–21. doi:10.1371/journal.pcbi.1003428.
70. Kühn T, Helias M. Locking of correlated neural activity to ongoing oscillations. *PLoS Comput Biol.* 2017;13:e1005534. doi:10.1371/journal.pcbi.1005534.
71. Zylberberg J, Pouget A, Latham PE, Shea-Brown E. Robust information propagation through noisy neural circuits. *PLoS Comput Biol.* 2017;13:e1005497. doi:10.1371/journal.pcbi.1005497.
72. Gerstner W, Kistler WM, Naud R, Paninski L. Neuronal dynamics: From single neurons to networks and models of cognition. Cambridge University Press; 2014.
73. Lefort S, Tómm C, Sarria JCF, Petersen CCH. The excitatory neuronal network of the C2 barrel column in mouse primary somatosensory cortex. *Neuron.* 2009;61(2):301–316. doi:10.1016/j.neuron.2008.12.020.
74. Packer AM, Yuste R. Dense, unspecific connectivity of neocortical parvalbumin-positive interneurons: a canonical microcircuit for inhibition? *J Neurosci.* 2011;31:13260–13271. doi:10.1523/JNEUROSCI.3131-11.2011.
75. Koelbl C, Helmstaedter M, Lübke J, Feldmeyer D. A barrel-related interneuron in layer 4 of rat somatosensory cortex with a high intrabarrel connectivity. *Cereb Cortex.* 2015;25(3):713–725. doi:10.1093/cercor/bht263.
76. Amitai Y, Gibson JR, Beierlein M, Patrick SL, Ho AM, Connors BW, et al. The spatial dimensions of electrically coupled networks of interneurons in the neocortex. *J Neurosci.* 2002;22(10):4142–4152. doi:20026371.
77. Lewis TJ, Rinzel J. Dynamics of spiking neurons connected by both inhibitory and electrical coupling. *J Comput Neurosci.* 2003;14(3):283–309.
78. Ostojic S, Brunel N, Hakim V. Synchronization properties of networks of electrically coupled neurons in the presence of noise and heterogeneities. *J Comput Neurosci.* 2009;26(3):369–392.
79. Tamás G, Buhl EH, Lörincz A, Somogyi P. Proximally targeted GABAergic synapses and gap junctions synchronize cortical interneurons. *Nat Neurosci.* 2000;3(4):366.
80. Sun QQ, Huguenard JR, Prince DA. Barrel cortex microcircuits: thalamocortical feedforward inhibition in spiny stellate cells is mediated by a small number of fast-spiking interneurons. *J Neurosci.* 2006;26(4):1219–1230. doi:10.1523/JNEUROSCI.4727-04.2006.
81. Markram H, Wang Y, Tsodyks M. Differential signaling via the same axon of neocortical pyramidal neurons. *Proc Natl Acad Sci USA.* 1998;95:5323–5328. doi:10.1073/pnas.95.9.5323.
82. Richardson MJE, Swarbrick R. Firing-rate response of a neuron receiving excitatory and inhibitory synaptic shot noise. *Phys Rev Lett.* 2010;105(17):178102. doi:10.1103/PhysRevLett.105.178102.