Cell type specific information transfer for sparse coding

Fleur Zeldenrust*, Niccolò Calcini*[#], Xuan Yan, Ate Bijlsma[@], Tansu Celikel

Department of Neurophysiology, Donders Institute for Brain, Cognition, and Behaviour,
Radboud University, Nijmegen - the Netherlands
[#] Current address: Department for BioMedical Research, Universität Bern, Switzerland
[@]Current address: Departement Population Health Sciences / Department of Biology, Universiteit Utrecht, the Netherlands
* denotes equal contribution.
Correspondence should be addressed to f.zeldenrust@neurophysiology.nl

Abstract

Sensory neurons reconstruct the world from action potentials (spikes) impinging on them. Recent work argues that the formation of sensory representations are cell-type specific, as excitatory and inhibitory neurons use complementary information available in spike trains to represent sensory stimuli. Here, by measuring the mutual information between synaptic input and spike trains, we show that inhibitory and excitatory neurons in the barrel cortex transfer information differently: excitatory neurons show strong threshold adaptation and a reduction of intracellular information transfer with increasing firing rates. Inhibitory neurons, on the other hand, show threshold behaviour that facilitates broadband information transfer. We propose that cell-type specific intracellular information transfer is the rate-limiting step for neuronal communication across synaptically coupled networks. Ultimately, at high firing rates, the reduction of information transfer by excitatory neurons and its facilitation by inhibitory neurons together provides a mechanism for sparse coding and information compression in cortical networks.

Introduction

The intracellular computation from synaptic input to action potential shows a strong compression (Huang et al., 2020). Even though the spike train of a single neuron contains only limited stimulus information, the full stimulus information can be recovered using the spike trains of a small population of (tens) of neurons. How many neurons are needed for full information recovery depends on a combination of the type of code the presynaptic neurons use (rate or timing) and on whether spike trains from interneurons or pyramidal cells are used. So the *decoding* of neural spike trains depends critically on how they were *encoded*.

Cortical excitatory pyramidal cells and inhibitory interneurons encode different features of input stimuli. For instance, excitatory neurons respond to sensory evoked stimuli by sparse, unreliable and selective spike trains (Murray and Keller, 2011; Reyes-Puerta et al., 2015) and have narrower receptive fields(Bruno and Simons, 2002; Wu et al., 2008) compared to interneurons. This is a result of differences in both connectivity (for recent reviews, see (Griffen, 2014; Hofer et al., 2011; Isaacson and Scanziani, 2011; Markram et al., 2004; Wu et al., 2011) and intrinsic biophysical properties (Cardin et al., 2007; Nowak et al., 2008). Indeed, most cortical pyramidal cells have broader subthreshold receptive fields than spiking receptive fields(Tan et al., 2011), which is the result of a combination of both local inhibition and the spike threshold ('iceberg effect' (Priebe and Ferster, 2008; Rose and Blakemore, 1974)), suggesting an important role for the spike threshold in shaping receptive fields. This raises the question: what are the differences in the spike-generating process between cortical interneurons and pyramidal cells, and how does this result in differences in information encoding?

We measured the information encoding properties of pyramidal cells and interneurons in L2/3 of the mouse barrel cortex. We chose a combination of *ex-vivo* experiments (da Silva Lantyer et al., 2018) and computational modelling to unravel both the threshold behaviour and the information encoding properties of excitatory and inhibitory neurons, using a recently developed method to estimate the mutual information between input and output in an *ex-vivo* setup (Zeldenrust et al., 2017). In vivo, excitatory neurons are shown to have a high action potential threshold (Crochet et al., 2011) and high selectivity (Ranjbar-Slamloo and Arabzadeh, 2019). Here, we investigated the underlying neural mechanisms. We found that excitatory neurons show strong threshold adaptation, making them fire sparsely and resulting in a strong compression of information between input and output. Inhibitory neurons on the other hand have a threshold behaviour that favours fast-spiking, resulting in a higher information rate transferred through higher spike frequencies, possibly having a gating role in information transfer.

Results

Information transfer in inhibitory and excitatory neurons

Excitatory neurons show strong adaptation

Whole-cell recordings were made from pyramidal cells and interneurons in layer 2/3 (L2/3) of mouse barrel cortical slices (da Silva Lantyer et al., 2018). Cells were classified as either 'excitatory' or 'inhibitory' based on their electrophysiological responses to a standard current-step protocol (Fig. 1, see Materials & Methods). In response to depolarizing steps, excitatory neurons show strong spike-frequency adaptation,

limiting their maximum firing rate (Fig 1F and Supplementary Table S1), whereas inhibitory neurons can fire at much higher rates. To measure the information transfer in the spike-generating process, we used a recently developed method (Zeldenrust et al., 2017) that uses the output of an artificial neural network (ANN) to generate the frozen noise current input used in our ex-vivo experiments. The ANN responds to a randomly appearing and disappearing preferred stimulus or 'hidden state' (Markov process). The mutual information between the input current and the hidden state depends on three properties of the ANN: the number of neurons (N), the average firing rate of the neurons (r) and the time constant of the hidden state ($\boldsymbol{\tau}$). Because of the differences in adaptation and maximum firing rate between the excitatory and inhibitory cells, it was not possible to use the exact same frozen noise input current for the two cell types: $\boldsymbol{\tau}$ had to be long for the excitatory neurons (neurons firing at a low rate cannot transfer information about a fast-switching stimulus), but this is not the case for the inhibitory neurons. However, the *information* in the input could be kept constant (Fig 2). We used the parameters in table 1 in the frozen noise experiments to generate the input currents shown in Fig. 2.

Parameter	Excitatory cells	Inhibitory cells
Number of artificial neurons N	1000	1000
Hidden state time constant $\boldsymbol{\tau}$	250 ms	50 ms
Average firing rate artificial neurons μ_q	0.1 Hz	0.5 Hz
Baseline input current	(set so the cell was at -70 mV, see Fig 2)	(set so the cell was at -70 mV, see Fig 2)
Amplitude input current	2100 pA	700 pA
Analysis window size	100 s	20 s

Table 1: Parameters of the input in the frozen noise experiments.

<u>Inhibitory neurons show broadband information transfer; Pyramidal cells transfer less information and at low frequencies</u>

The information transfer of the spike-generating process of a single neuron can be estimated by calculating the mutual information between the spike train and the hidden state used for generating the input current (see Materials & Methods). More specifically, we define the fraction of transferred information (FI) as the mutual information between the spike train and the hidden state divided by the mutual information (MI) between the input current and the hidden state:

$$\mathrm{FI} = \frac{\mathrm{MI}_{\textbf{spike train}}}{\mathrm{MI}_{\textbf{input current}}}. \qquad (1)$$

FI quantifies how much information about the hidden state is transferred from the input current to the spike train, and thus quantifies which fraction of the information is kept during the spike-generating process. In Fig 3A we show FI as a function of the firing rate r, for inhibitory (blue) and excitatory (red) neurons, and compare it to the FI obtained from the optimised 'Bayesian neuron' (BN) model (Deneve, 2008) with the same parameters (see Materials & Methods) as for the input generated for the excitatory neurons (pink) or inhibitory neurons (turquoise). Pyramidal cells transfer more information at low firing rates (<~8 Hz) compared to interneurons. This is due to our choice of slower switching speed of the hidden state for excitatory neurons: a fast-switching hidden state cannot be properly tracked by neurons firing at a low firing rate (see also (Zeldenrust et al., 2017)). To compare inhibitory and excitatory neurons, we plotted FI as a function of the normalized firing rate $r_n = r*\boldsymbol{\tau}$ (unitless, Fig 3B) and fitted the measured values up to r_n = 1.5 to a saturating function:

$$\mathrm{FI} = a\left(\frac{2}{1+e^{-br_n}} - 1\right), \qquad (2)$$

where a is the saturation value and b is the rate with which this the saturation value a is reached (both unitless). In Fig. 3E and F the fit values and their 95% confidence intervals are shown. Inhibitory experimental and BN values saturate around similar values (a = 0.65 (0.64 - 0.66) and a = 0.64 (0.63 - 0.65) respectively), with experiments having a slightly lower rate (b = 5.8 (5.6 - 6.0) and 7.7 (7.3 - 8.0). Excitatory neurons saturate at lower experimental values (a = 0.51 (0.48 - 0.54)) and slightly lower BN values (a = 0.58 (0.54 - 0.63), and the saturation rates are also lower (b = 4.5 (4.0 - 4.9) and b = 6.1 (5.0 - 7.2) respectively). This shows that in the case of the excitatory neurons, the experimental spike trains transmit much less information than the spike trains of the BN, whereas in the inhibitory case the model and experimental spike trains perform similarly. As a control, we presented the input for the excitatory neurons also to inhibitory neurons (Fig. 3, green, a = 0.63 (0.52 - 0.75), b = 2.6 (1.6 - 3.7)); these inhibitory neurons fired at a higher normalized rate (Fig. 3C) and performed better than the excitatory neurons.

Inhibitory neurons transfer information more efficiently

To compare the efficiency of the pyramidal cells, the interneurons and the BN discussed above, we defined the (unitless) efficiency *E* as the fraction of information divided by the normalized firing rate (i.e. the firing rate relative to the switching speed of the hidden state):

$$\mathrm{E} = \frac{\mathrm{FI}}{r \cdot \tau}. \qquad (3)$$

This efficiency is shown as a function of the firing rate in Fig 4A and of the normalized firing rate in Fig 4B. If the fraction of information FI depends exponentially on the normalized firing rate, such as in Fig. 3B, the efficiency decreases as a function of the (normalized) firing rate (note that for very low firing rates the mutual information and hence the efficiency cannot be calculated reliably, due to the lack of spikes). In Fig 4B, the theoretical values for the fitted curves from Fig 3B and their inflection points (squares) are also shown (lines with circular black-lined markers). These curves have a limit value for a vanishing normalized firing rate at E=ab/2. With increasing (normalized) firing rate, the efficiency decreases. The inflection points of the theoretical curves are indicated by black-lined squares. Note that the model BN curves (turquoise for inhibitory neurons and pink for excitatory neurons) have higher efficiency values than their experimental counterparts, indicating that the BN model is indeed optimal in the sense that it sets an upper limit to the efficiency. The inhibitory experimental values are much closer to their BN counterparts than the excitatory experimental values to their BN counterparts, indicating that the inhibitory neurons are more efficient. However, when the interneurons respond to the slow-switching input made for the excitatory neurons (green markers), they perform similarly or even less efficiently than the excitatory neurons (red markers). Therefore, interneurons are more efficient than excitatory neurons at transferring information from the fast-switching stimulus, but not from the slow-switching stimulus.

Inhibitory neurons perform well as classifiers

The setup with the hidden state made it possible to show 'receiver-operator curves' (ROCs): we defined a 'hit' as a period during which the hidden state was 1 (up-state), in which at least 1 action potential was fired, and a 'miss' as an up-state in which no action potentials were fired. Similarly, we defined a 'false alarm' as a period during which the hidden state was 0 (down-state), in which at least 1 action potential was fired, and a 'correct reject' as a down state in which no action potentials were fired. We then defined the 'hit fraction' as the number of hits divided by the total number of up-states, and similarly the false alarm fraction for the number of false alarms divided by the total number of down-states. In Fig. 5A the results are shown, for the same five conditions as discussed above. For each experiment, a control experiment was simulated by generating a Poisson spike train with the same number of spikes as the original experiment. Note that

this 'control' is below the line hit fraction = false alarm fraction, because the hidden state is more often 0 than 1 ($P_1 = ⅓$). Since the hidden state is longer in the '0' state, the probability that a random spike occurs when the hidden state equals 0 is higher, hence the probability for a false alarm is higher than the probability for a hit.

Interneurons perform very similarly to the BN, as shown in Fig. 5, whereas the pyramidal cells perform less optimal than their model counterparts. We performed control experiments where input currents generated for excitatory neurons were injected into inhibitory neurons, (green triangles in Fig. 4 and 5). The results suggest that interneurons perform comparably to (on the same curve as) pyramidal cells, but with a lower discrimination threshold (i.e. with a higher firing rate), which is in agreement with our previous observation that interneurons responded with a higher firing rate than pyramidal cells. Note that inhibitory neurons fired slightly less spikes during the up-states (Fig. 5B) and the normalized firing rate in the up-state somewhat lower for the inhibitory neurons (Fig. 5F). Since the excitatory neurons fire more spikes during the down states (Fig. 5C and G), this corresponds to a lower efficiency for excitatory neurons (Fig. 4) and a worse performance on the binary classification task (Fig. 5A). Indeed, the number of spikes per down state (Fig. 5C) and normalized firing rate in the down state (Fig. 5G) differed between inhibitory and excitatory neurons (Supplementary Tables S2 and S3). Note that most 'incorrect' spikes are actually fired shortly after a down switch (Fig 5 H-K), so they might be 'correct' spikes that were a few milliseconds too late.

Interestingly, for large firing rates the experiments appear to outperform the model. This appears paradoxical, since the BN is an optimal model for the stimulus. However, the BN is an optimal model given a certain set of assumptions, one of which is the form of adaptation it uses: the BN increases its firing threshold each time a spike is fired, so that it only fires when this transfers 'new information' (Deneve, 2008). The neurons we recorded might use other forms of adaptation, that are less efficient (i.e. they use more spikes to transmit less information, see Figs 3 and 4) , but apparently also less prone to false alarms.

Pyramidal cells show diverse response properties

The information transfer, efficiency and ROC curves (Figs 3-5) show that inhibitory neurons transfer information more efficiently. They are close to optimal in transferring information about the hidden state (i.e. it shows *how much* information these neurons transfer, but not *what* features they respond to). In Fig. 6 we show the normalized spike-triggered averages (STAs) for spikes of inhibitory neurons (A and E) and excitatory neurons (C). The filter was whitened and regularized (see Materials & Methods). Next, the

projection values of spike-triggering and random currents were calculated (Figs 6B for an example for 1 cell), and the distance between the means of the distributions for random and spike-triggering currents was calculated for each cell (Figs 6D).

The average STAs for all inhibitory (Fig 6A, blue) and excitatory (Fig 6C, red) neurons were quite similar, but the traces for individual neurons (grey lines) were much more variable for excitatory neurons than inhibitory neurons. This could indicate that the excitatory neurons have a higher variance in their feature selectivity of incoming current stimuli than inhibitory neurons, but it is also possible that this is an effect of the lower number of spikes available for excitatory neurons. To control for this possibility, we calculated the STAs for spike trains of inhibitory neurons, where the number of spikes was reduced to match an excitatory trial (Fig 6E, brown). For all three groups (inhibitory, excitatory and inhibitory control spike trains) we calculated the inner product between all calculated STAs. Fig 6F shows the distributions of these inner products, and it is clear that both inhibitory full and control spike trains are much less *variable* (inner product closer to 1) than the excitatory spike trains (two-sample Kolmogorov-Smirnov test E-I $p=0$, E-C $p<1e-223$, I-C $p<1e-228$). The distribution of all distances between the means is shown in Figs 6D. The distances between the distributions, measured in standard deviations of the prior (random triggered currents) distribution, are much higher for excitatory neurons than for inhibitory neurons, indicating that excitatory neurons are more *selective* (p-values two-sample t-test: E-I $p<1e-28$, E-C $p<1e-24$, I-C $p=0.14$).

In conclusion, pyramidal cells fire less and are therefore more selective, but at the same time there is more variability between excitatory neurons in what input features they respond to. Excitatory neurons transfer information mainly at low frequencies, whereas interneurons transfer more information in a broadband spectrum and are more uniform with respect to the input features they respond to. Recently, we have shown (Huang et al., 2016) that the threshold behaviour of neurons is a crucial factor in their information processing. Therefore, in the next section the spike threshold behaviour of the excitatory and inhibitory neurons was investigated.

Threshold dynamics in inhibitory and excitatory neurons

Dynamic threshold of both neuron types

In Fig 7, we show the threshold behaviour of the inhibitory and excitatory neurons. The membrane potential threshold of each spike was determined based on the method of Fontaine et al. (2014) (see Materials & Methods). We show the distribution of the membrane potential as a function of the inter-spike interval (ISI, Fig. 7A and E). For both inhibitory and excitatory neurons, the membrane potential threshold goes up with

short ISIs, as expected, and for long ISIs the threshold is low. This effect has a long time scale (at least several tens of milliseconds), longer than expected based on the relative refractory period alone (typically less than ten milliseconds). The threshold for excitatory neurons is almost 10 mV higher than for inhibitory neurons (Fig 7A,E,G,H). Next to the ISI, the threshold also depended on the history of the membrane potential (Fig 7C,F): we calculated the regression between the action potential threshold and the average membrane potential in different windows preceding the spike. There is a strong correlation between the threshold and the membrane potential immediately preceding the spike for both neuron types, which reduces gradually with time before the spike. However, for both neuron types some relation between membrane potential and threshold is still visible several tens of milliseconds before the spike. The current clamp step protocol (Fig 7 G-J) confirms the overall higher threshold for excitatory neurons (Fig 7G, H) and strong spike-frequency adaptation (Fig 7 J) for excitatory neurons. The threshold adaptation rate however, shows significant differences between fast spiking and regular spiking neurons at current injection intensities ranging from +240 to +320pA, while they do not show significant changes at lower or higher intensities, possibly due to low firing rates or to reaching a steady state firing rate. (Fig 7J, table S1).

So in conclusion, both inhibitory and excitatory neurons show dynamic threshold behaviour, with interneurons having much lower thresholds, so they can fire at high rates, whereas the dynamic threshold of excitatory neurons promotes low-frequency firing and adaptation.

Voltage clamp experiments show a narrow sodium channel activation in inhibitory cells

To investigate the activation mechanism of excitatory and inhibitory cells, we performed voltage clamp experiments on a separate group of cells from the same area. We performed a sawtooth-protocol (Fig 8, A-C), in which the clamped membrane potential of the cells was linearly switched from -70 mV to 70 mV in 100 ms (Fig 8, left column), 50 ms (Fig 8, middle column) or 10 ms (Fig 8, right column). In response to such a rapidly changing membrane potential, the neuron generated current peaks (Fig. 8, A-C), which were abolished in control experiments with TTX application (Supplementary Fig S3), suggesting these are caused by sodium channel activation. We quantified the amplitude (not displayed), half-width (Fig S3, D-F), peak onset and offset duration (Fig 8, G-I and J-L respectively), adaptation of the amplitude of the 1st peak over consecutive voltage ramps (Fig 8, M-O) of the first of these peaks. The only significant, consistent and relevant difference we observed with this protocol between the inhibitory and excitatory neurons was in the half width of the first peak, which was strongly reduced in inhibitory neurons (Fig 8, D-F). By separating the half-width in onset and offset the difference remains in both measures, but interestingly the regular spiking neurons seem to be split into two different groups. Moreover, in this cell type the amplitude of the first peak was reduced in amplitude over consecutive sawtooths (Fig 8, M-O).

Discussion

Postsynaptic information processing requires decoding information from presynaptic action potentials. In an accompanying (Huang et al., 2020) paper we have shown that excitatory and inhibitory neurons *decode* spike-timing and firing rate information differentially. In particular, we found that when a 'Rate+Poisson' *encoding* paradigm was used, the spike trains of inhibitory neurons contain more stimulus information than the spike trains of excitatory neurons. However, whether the two neuronal classes *encode* information differently is not known. In this paper we have addressed this question by quantifying the internal information loss between input current and output spike train in *ex-vivo* current clamp experiments. We used a stimulus that is comparable to the 'Rate+Poisson' coding scheme: the input current was generated by an Artificial Neural Net (ANN), where each cell fires Poisson spike trains of which the firing rate is modulated by the absence or presence of the stimulus (Zeldenrust et al., 2017). As predicted in (Huang et al., 2020), we found that excitatory neurons transfer information at limited rates whereas inhibitory neurons transfer more information and at much higher rates. So excitatory cells show a stronger compression and are therefore more selective. Moreover, they show a strong variability in their response properties over the population. Interneurons on the other hand, show a near-optimal response, transferring much information about the input at relatively high rates. They are more uniform in the features they respond to and are more efficient at transferring information. These differences in information transfer between inhibitory and excitatory neurons are accompanied by a different threshold behaviour: excitatory cells show strong threshold adaptation, keeping their firing rates low, whereas interneurons have a lower threshold and can fire and transfer information at much higher rates.

The observed differences between inhibitory and excitatory neurons in threshold behaviour and information transfer can be partially explained by differences in the sodium activation profile. We observed a clear difference in activation peak width between inhibitory and excitatory neurons, but not in the activation voltage. Moreover, the amplitude of these peaks adapted strongly for excitatory, but not for inhibitory neurons. This suggests that the main differences in threshold behaviour between inhibitory and excitatory neurons might not so much be explained by the initial sodium activation, but by what happens right after it: sodium inactivation and the activation of repolarizing currents. The high threshold we observed in interneurons could be the result of an interplay between the sodium current and a fast potassium current that promotes the de-inactivation of the sodium channels. For instance, Kv3 potassium channels only occur in fast-spiking cells (Erisir et al., 1999; Grissmer et al., 1994; Rowan et al., 2016, 2014; Rudy and McBain, 2001), suggesting a link between threshold facilitation and these potassium channels). Other candidates include sodium- and/or calcium-activated potassium channels (Sanchez-Vives et al., 2000), that have been shown to cause spike-frequency adaptation and after-hyperpolarizations. However, the exact interplay

between the spike history, the synaptic input and the spike threshold is complex, depending next to the intrinsic channel properties on for instance the axon initial segment location (for a review, see (Kole and Stuart, 2012)).

Neurons in the sensory cortices process information coming from the periphery by giving complex, non-linear spike responses to incoming stimuli (Koch and Segev, 2000). These responses are shaped by each neuron's place in the network (i.e. the connectivity of the network, (Harris and Mrsic-Flogel, 2013; Hofer et al., 2011; Ko et al., 2013, 2011; Okun et al., 2015)) and by its biophysical properties (i.e. which ion channels are expressed in its membrane). Cortical networks consist of excitatory pyramidal cells and inhibitory interneurons, two groups that differ from one another both in their connectivity (for recent reviews, see (Hofer et al., 2011; Markram et al., 2004)) and their intrinsic biophysical properties (for a recent review, see (Griffen, 2014)). The differences between the two groups of cells result in different responses to sensory stimuli (Murray and Keller, 2011), but to what extent these differences can be attributed to the differences in network structure or to the differences in single neuron properties remains an open question. For instance, the recent results of (Reyes-Puerta et al., 2015), who found using simultaneous recordings in rat barrel cortex that the spike trains of interneurons contained more stimulus-related information than those of pyramidal cells, could be the result of intrinsic neuronal properties, network connectivity differences between inhibitory and excitatory neurons, differences in adaptation, and so on. We show here that at least part of the observed differences between the response properties of inhibitory and excitatory neurons in layer 2/3 of the mouse barrel cortex can be explained by intrinsic neuronal properties, next to network properties.

Neurons in the cortex react to prolonged or repeated stimuli with a spike response that decreases in frequency over time, an effect called spike-frequency adaptation (Gutkin and Zeldenrust, 2014). This adaptation can be the result of a hyperpolarization of the membrane potential (for instance due to the activation of an outward current (Sanchez-Vives et al., 2000), an increase in threshold (Azouz and Gray, 2000), or both. It has been repeatedly shown that such adaptation is an essential part of the functioning of cortical neurons, and any model of cortical firing should include such adaptation (Brette and Gerstner, 2005; Fontaine et al., 2014; Gerstner and Naud, 2009; Jolivet et al., 2004; Rauch et al., 2003; Rossant et al., 2011). However, the functional implications of different forms of adaptation are still largely unknown. Here, we associated the functional implications of threshold adaptation with the information processing capability of inhibitory interneurons and excitatory pyramidal cells, which show very different forms of threshold behaviour.

Measuring the information transfer of neurons in an *ex-vivo* setup traditionally uses variations on one of two methods: 1) the 'direct method' (de Ruyter van Steveninck et al., 1997; Strong et al., 1998), in which a frozen-noise stimulus is repeated many times to estimate the signal-to-noise ratio, or 2) using long frozen noise stimuli to fit a reverse correlation model and estimate the mutual information between input and output from how well this model can predict the input (Bialek et al., 1991; de Ruyter van Steveninck and Bialek, 1988; Rieke et al., 1997). However, both these methods need long recordings (~1 hour) for a single estimate of the mutual information between the input current and the output spike train. Such long recording times are often not feasible for *ex-vivo* experiments, especially if different conditions need to be compared. Therefore, we used a recently developed method to measure the information transfer between input current and output spike train in *ex-vivo* experiments that can estimate the mutual information in recordings of tens to hundreds of seconds (Zeldenrust et al., 2017). In this method, the input current used in a current-clamp setup is the simulated output of an artificial neural network (ANN) that responds to a randomly appearing and disappearing preferred stimulus or 'hidden state' (Markov process). The information in the input can be explicitly controlled by varying the number of spikes of the ANN (the number of neurons and their firing rates) and the switching speed of the hidden state. In this method, all data can be used for the estimation of the mutual information, and trial repetition or model fitting is not necessary, making the required length of the recording much shorter. One assumption of the information method is that the output spike trains generated by the clamped neurons are (approximately) Poissonian. In Supplementary Fig. S5 it can be seen that there are slightly more spike doublets than expected from a Poisson process, but that there is otherwise no clear structure in the spike trains.

At the local network level, full stimulus information is preserved at the presynaptic level and in small groups of neurons (Huang et al., 2020). However, that does not mean that the full information is transferred to the next processing level. The fact that excitatory neurons, the neurons that connect not only locally but also across layers and areas, are more selective, as well as more variable in what stimulus features they represent, suggest that the network performs both a form of information compression and transformation. Moreover, (Huang et al., 2016) showed that an adaptive threshold model neuron is more informative for high temporal precision and low noise than a fixed threshold model neuron, suggesting that especially noisy stimuli with low temporal precision are suppressed. So the relation between threshold adaptation and information processing is very strong. The functional reason for this, remains an open topic. One theory suggests that such fast inhibition is needed for predictive coding (Boerlin et al., 2013; Denève and Machens, 2016; Hawkins and Ahmad, 2017; Zeldenrust et al., 2019). In addition to this, we suggest that the observed compression of excitatory neurons and the broadband information transfer of inhibitory neurons serve a common goal: sparse coding (Földiák, 1990; Foldiak and Endres, 2008; Olshausen and Field, 1996). In

vivo, the activity of neurons in barrel cortex is shown to be extremely sparse, with very narrow receptive fields (Ranjbar-Slamloo and Arabzadeh, 2019), possibly due to a large distance between the resting membrane potential and neural threshold in excitatory neurons (Crochet et al., 2011). Here, we show two underlying neural mechanisms that collaborate towards this sparse coding: strong threshold adaptation that compresses information in excitatory neurons and high activity levels in inhibitory neurons with little or no compression. Together, these mechanisms effectively 'gate' the message to the next processing layer, resulting in a specific and sparse code.

Materials & Methods

Experiments

Ethics statement.

Animals used were Pval-cre and SSt-cre mice from 9 to 45 weeks kept with unlimited access to water and food, housed in a 12 hour light/dark cycle. All experimental procedures were performed according to Dutch law and approved by the Ethical Committee for Animal Experimentation of Radboud University (RU DEC). Each mouse was perfused with iced and oxygenated (95%O_2/5%CO_2) Slicing Medium (composition in mM: 108 ChCl, 3 KCl, 26 $NaHCO_3$, 1.25 $NaH_2PO_4.H_2O$, 25 Glucose.H_2O, 1 $CaCl_2.2H_2O$, 6 $MgSO_4.7H_2O$, 3 Na-Pyruvaat) under anaesthesia with 1,5ml Isoflurane inhaled for 2 minutes.

Slice electrophysiology

The brain was covered in 2% agarose and submerged in Slicing Medium after which it was sliced in 300 μM thickness using a VF-300 compresstome (Precisionary Instruments LLC) and then incubated for 30 min in 37°C artificial cerebrospinal fluid (ACSF, composition in mM: 1200 NaCL, 35 KCL, 13 $MgSO_4$.7H2O, 25 $CaCl_2.2H_2O$, 100 Glucose.H_2O, 12.5 $NaH_2PO_4.H_2O$, 250 $NaHCO_3$), oxygenated (95%O_2/5%CO_2). The heating of the ACSF solution was then stopped and 30m were allowed for the ACSF bath to reach room temperature. Slices were then kept in this bath until use.

Slices were placed into the recording chamber under the microscope (Eclipse FN1, Nikon) and perfused continuously at a rate of 1 ml/min with the oxygenated ACSF at room temperature. Patch pipettes for whole-cell recordings were pulled from borosilicate glass capillaries, 1.0 mm outer diameter, 0.5mm inner diameter, on a pipette-puller (Sutter Instrument Co. Model P-2000), until an impedance of 8±2 MΩ for the tip was obtained. Pipettes were filled with a solution containing (in mM) *115 $CsMeSO_3$, 20 CsCl, 10 HEPES, 2.5 $MgCl_2$, 4 Na_2ATP, 0.4 NaGTP, 10 Na-phosphocreatine, 0.6 EGTA, 5 QX-314* (Sigma).

The whole cell access was obtained after reaching the GOhm seal and breaking the membrane. Upon entering the cell and the whole-cell mode, the membrane potential was kept fixed at -70mV, outside stimulation.

Input current generation

Data acquisition was performed with HEKA EPC9 amplifier controlled via HEKA's PatchMaster software (version 2.90x.2), subsequent analysis with MatLab (Mathworks, v.2016b).

The current clamp (CC) steps protocol was performed in every cell, and used to distinguish between cell type, according to the Firing Rate and spike shape. The protocol consisted in clamping the neuron at a baseline current $I_{baseline}$, corresponding to the one required to keep its membrane at -70mV, and providing a 500ms long stimulus of fixed current value I = $I_{baseline}$ + (40pA * step number), for a total of 10 steps, reaching a maximum current injected of $I_{baseline}$ + 400pA. Between each current injection step, a 5.5s recovery window was allowed.

To be able to use the method to measure the information transfer (Zeldenrust et al., 2017), the input current used in the current-clamp experiments is generated as the output of an artificial neural network (ANN) that responds to a randomly appearing and disappearing preferred stimulus or 'hidden state' (Markov process) x: a binary variable that can take the values of 1 (preferred stimulus present, 'on-state') and 0 (preferred stimulus absent, 'off-state'). The Markov process has rates r_{on} and r_{off}. This corresponds to a switching time constant $\tau = \frac{1}{r_{on}+r_{off}}$. A time constant $\boldsymbol{\tau}$ of 50 or 250 ms was used for inhibitory/fast spiking or excitatory/regular spiking neurons respectively, as excitatory neurons did not show a high enough firing rate to allow for faster time constants. The ANN consists of N=1000 neurons that fire Poisson spike trains, whose firing rates are modulated by the stimulus so that each neuron fires with rate q_{on}^{i} when x = 1, and $q_{off}^{i}$when x = 0. These rates are drawn from a Gaussian distribution with mean μ_q (see Table 1) and standard deviation $\sigma_q = \sqrt{\frac{1}{8}\mu_q}$. Each spike was convolved with an exponential kernel with unitary surface and a decay time of 5ms and the spike trains from different presynaptic neurons contribute to the output with weight $w_i = log\ \frac{q_{on}^{i}}{q_{off}^{i}}$.

Analysis

Cell classification

Cells were classified using the following procedure. Before the frozen noise injection, for each cell, the response to a current-clamp step (CC-step) protocol was recorded. From these recordings, the maximum firing rate, the average spike-halfwidth and the average after hyperpolarization (AHP) amplitude were extracted (Fig 1). On-site, the cells were classified by the experimenter based on the firing rate and the spike width. Based on this initial classification the cell received the frozen noise with either $\boldsymbol{\tau}$ = 250 ms (excitatory neurons) or $\boldsymbol{\tau}$ = 50 ms (inhibitory neurons).

Offline, the initial classification was verified using an agglomerative clustering protocol (MATLAB 'clusterdata') to cluster the data into 2 groups (separated following Ward's method(Ward, 1963)), according to the maximum firing rate and the average spike-half width (normalized to zero mean and unit standard deviation) reached during the CC-Step protocol (Fig 1E). Only for a single cell, the initial classification and the post-hoc classification were in disagreement (Fig 1E, pink star). We decided to keep this cell in the original (inhibitory) group due to its position between the two clusters.

Calculation of mutual information

The mutual information between the hidden state and the input (MI_I) or a spike train ($MI_{spike\ train}$) can be estimated with the help of this hidden state x. The method is explained in detail in (Zeldenrust et al., 2017) and follows derivations from (Denève, 2008; Lochmann and Denève, 2008).

The log-odds ratio L that the hidden state is 1, given the history of the input until now $I(t)$can be estimated using

$$\frac{\hat{L}}{dt} = r_{\mathrm{on}}(1-e^{-\hat{L}}) - r_{\mathrm{off}}(1+e^{\hat{L}}) + I(t) - \theta, \tag{4}$$

Where $\theta = \sum_{i=1}^{N} \quad q_{on}^{i} - q_{off}^{i}$is the constant offset of the input, which is chosen to equal 0 here. Using the log-odds ratio, we can estimate the conditional entropy

$$\hat{H}_{xy} = -\langle x \log_2\left(\frac{1}{1+e^{-L}}\right) + (1-x)\log_2\left(1-\frac{1}{1+e^{-L}}\right)\rangle_{\mathrm{time}}. \tag{5}$$

Because the hidden state follows a memoryless Markov process, its entropy at every moment in time can be calculated by

$$H_{xx} = -p_1 \log_2(p_1) - (1-p_1) \log_2(1-p_1). \quad (6)$$

Where $p_1 = \frac{r_{on}}{r_{on}+r_{off}}$ is the prior probability that the hidden state equals 1. With the canonical $MI = H_{xx} - H_{xy}$ the mutual information between the input and the hidden state can now be estimated. Similarly, the mutual information between a spike train and the hidden state can be estimated by integrating equation (4) where the input I is now replaced by

$$I_{\text{spike train}} = w \cdot \rho(t), \quad (7)$$

where $\rho(t)$ is the spike train of the neuron, and its weight w is given by

$$w = \frac{\hat{q}_{\text{on}}}{\hat{q}_{\text{off}}} = \frac{\#\text{ spikes while } x = 1}{\#\text{ spikes while } x = 0} \quad (8)$$

and θ is calculated similarly based on the observed q_{on} and q_{off}.

<u>Bayesian neuron</u>

The `Bayesian neuron' (Denève, 2008) is a spiking neuron model that optimally integrates evidence about the hidden state from the ANN described above. It is optimal given an efficient coding or redundancy reduction assumption: it only generates new spikes if that spikes transfers *new* information about the hidden state, that cannot be inferred from the past spikes in the spike train. In practice, the neuron performs a leaky integration of the input, in order to calculate the log-odds ratio L for the hidden state being 1:

$$\frac{L}{dt} = r_{\text{on}}(1-e^{-L}) - r_{\text{off}}(1+e^{L}) + I(t) - \theta, \quad (9)$$

where r_{on} and r_{off} are the switching speeds of the hidden state, and $\theta = \sum_{i=1}^{N} q_{on}^{i} - q_{off}^{i}$ is the constant offset of the input, which is chosen to be equal to 0 in this paper. The neuron compares this log-odds ratio from the input with the log-odds ratio of its own spiketrain, G:

$$\frac{G}{dt} = r_{\text{on}}(1-e^{-G}) - r_{\text{off}}(1+e^{G}), \quad (10)$$

which is updated each time a spike is fired:

$$\text{if } L > G + \frac{\eta}{2} \begin{cases} \text{a spike is fired} \\ G \to G + \eta \end{cases} , \tag{11}$$

where $\eta$is the only free parameter that describes the distance between the threshold and the reset of the Bayesian neuron. Note that this neuron model has a form of threshold adaptation: if it did not spike for a long time, G decays to its prior value $G_{prior} = ln\frac{r_{on}}{r_{off}}$. With each spike, G is increased by η, and more input (larger L) is needed to fire a spike.

Threshold detection

The membrane potential threshold of each spike was determined based on the method of (Fontaine et al., 2014): in a window of 1 to 0.25 before each spike maximum, the earliest time in the window at which either the first derivative exceeded 18 mV/ms (i.e. exceeded a value of 0.9 mV for a sampling rate of 20000 Hz) or the second derivative exceeded 140 mV/ms^2 (i.e. exceeded a value of 0.35 mV for a sampling rate of 20000 Hz) was designated as the threshold-time, and the threshold value was determined as the corresponding membrane potential.

ROC curves

We defined a 'hit' as a period during which the hidden state was 1, in which at least 1 action potential was fired, and a 'miss' as a period during which the hidden state was 1, in which no action potentials were fired. Similarly, we defined a 'false alarm' as a period during which the hidden state was 0, in which at least 1 action potential was fired, and a 'correct reject' as a period during which the hidden state was 0, in which no action potentials were fired. So each period in which the hidden state was 1, was either defined as a 'hit' or a 'miss', and each period in which the hidden state was 0, was either defined as a 'false alarm' or a 'correct reject'. The total number of hits was divided by the total number of periods during which the hidden state was 1, which resulted in the fraction of hits $0 \leq f_h \leq 1$ and similarly for the misses, false alarms and correct rejects. We calculated the fractions of hits, misses, false alarms and correct rejects for each spike train, as well as for a corresponding Poisson spike train of the same length and with the same number of spikes. Note that for these Poisson spike trains, the hit fraction is actually below the hit fraction = false alarm fraction line, due to the nature of the hidden state: because the hidden state is more often 0 than 1, a random spike will have a higher chance of being timed during a period where the hidden state equals 0. Therefore, the false alarm fraction will be higher than the hit fraction for Poisson spike trains.

Spike-triggered average

The whitened and regularized spike-triggered average (STA) was calculated as

$$\mathrm{STA} = (X^T X \lambda \mu_{X^T X} I) \backslash (X^T s), \ (12)$$

where $X$is a stimulus-lag matrix, where each row is the stimulus vector with a different lag (see (Chichilnisky, 2001; Paninski, 2003; Sharpee et al., 2004; Simoncelli et al., 2004)), $X^T X$is the correlation matrix and $I$is the identity matrix. Parameter $\lambda$is the regularization parameter set to 10 and $\mu_{X^T X}$is the mean of the diagonal of the correlation matrix. Finally, s denotes the spike train. The resulting STA was normalized with the L2 norm. For each neuron, the inner product of all spike-triggering stimuli was calculated, as well as the same number of random-triggered stimuli. With the random-triggered stimuli, the prior distribution of the input could be calculated, and compared to the distribution of spike-triggering stimuli (the posterior distribution). The difference in mean between the prior and posterior was calculated for each neuron, and the distribution of means over all neurons is shown in Fig 6.

Acknowledgements

This work was supported by grants from the European Commission (Horizon2020, nr. 660328), European Regional Development Fund (MIND, nr. 122035) and the Netherlands Organisation for Scientific Research (NWO-ALW Open Competition, nr. 824.14.022) to TC and by the Netherlands Organisation for Scientific Research (NWO Veni Research Grant, nr. 863.150.25) to FZ.

References

Azouz R, Gray CM. 2000. Dynamic spike threshold reveals a mechanism for synaptic coincidence detection in cortical neurons in vivo. *Proceedings of the National Academy of Sciences* **97**:8110–8115. doi:10.1073/pnas.130200797

Bialek W, Rieke F, de Ruyter van Steveninck RR, Warland D. 1991. Reading a Neural Code. *Science* **252**:1854–1857.

Boerlin M, Machens CK, Denève S. 2013. Predictive Coding of Dynamical Variables in Balanced Spiking Networks. *PLoS Computational Biology* **9**. doi:10.1371/journal.pcbi.1003258

Brette R, Gerstner W. 2005. Adaptive exponential integrate-and-fire model as an effective description of neuronal activity. *Journal of neurophysiology* **94**:3637–42. doi:10.1152/jn.00686.2005

Bruno RM, Simons DJ. 2002. Feedforward mechanisms of excitatory and inhibitory cortical receptive fields. *The Journal of Neuroscience* **22**:10966–75.

Cardin JA, Palmer LA, Contreras D. 2007. Stimulus Feature Selectivity in Excitatory and Inhibitory Neurons in Primary Visual Cortex. *The Journal of Neuroscience* **27**:10333–10344. doi:10.1523/JNEUROSCI.1692-07.2007

Chichilnisky EJ. 2001. A simple white noise analysis of neuronal light. *Network: Computation in Neural Systems* **12**:199–213.

Crochet S, Poulet JFA, Kremer Y, Petersen CCH. 2011. Synaptic Mechanisms Underlying Sparse Coding

of Active Touch. *Neuron* **69**:1160–1175. doi:10.1016/j.neuron.2011.02.022

da Silva Lantyer A, Calcini N, Bijlsma A, Kole K, Emmelkamp M, Peeters M, Scheenen WJJ, Zeldenrust F, Celikel T. 2018. A databank for intracellular electrophysiological mapping of the adult somatosensory cortex. *GigaScience* **7**:1–9. doi:10.1093/gigascience/giy147

de Ruyter van Steveninck RR, Bialek W. 1988. Real-time performance of a movement-sensitive neuron in the blowfly visual system: coding and information transfer in short spike sequences. *Proceedings of the Royal Society of London Series B* **234**:379–414.

de Ruyter van Steveninck RR, Lewen GD, Strong SP, Koberle R, Bialek W. 1997. Reproducibility and Variability in Neural Spike Trains. *Science* **275**:1805–1808. doi:10.1126/science.275.5307.1805

Denève S. 2008. Bayesian spiking neurons I: inference. *Neural Computation* **20**:91–117. doi:10.1162/neco.2008.20.1.91

Denève S, Machens CK. 2016. Efficient codes and balanced networks. *Nature neuroscience* **19**:375–82. doi:10.1038/nn.4243

Erisir A, Lau D, Rudy B, Leonard CS. 1999. Function of Specific K+ Channels in Sustained High-Frequency Firing of Fast-Spiking Neocortical Interneurons. *Journal of Neurophysiology* **82**:2476–2489. doi:10.1152/jn.1999.82.5.2476

Földiák P. 1990. Forming sparse representations by local anti-Hebbian learning. *Biol Cybern* **64**:165–170.

Foldiak P, Endres D. 2008. Sparse Coding. *Scholarpedia* **3**:2984.

Fontaine B, Peña JL, Brette R. 2014. Spike-Threshold Adaptation Predicted by Membrane Potential Dynamics In Vivo. *PLoS Computational Biology* **10**:1–11. doi:10.1371/journal.pcbi.1003560

Gerstner W, Naud R. 2009. How good are neuron models? *Science* **326**:379–80. doi:10.1126/science.1181936

Griffen TC. 2014. GABAergic synapses: their plasticity and role in sensory cortex - See more at: http://journal.frontiersin.org/Journal/10.3389/fncel.2014.00091/abstract#sthash.l5jGe6MC.dpuf. *Front Cell Neurosci* **8**.

Grissmer S, Nguyen AN, Aiyar J, Hanson DC, Mather RJ, Gutman GA, Karmilowicz MJ, Auperin DD, George Chandy K. 1994. Pharmacological Characterization of Five Cloned Voltage-Gated Expressed in Mammalian Cell Lines. *Molecular Pharmacology* **45**:1227–1234.

Gutkin BS, Zeldenrust F. 2014. Spike frequency adaptation. *Scholarpedia* **9**:30643. doi:10.4249/scholarpedia.30643

Harris KD, Mrsic-Flogel TD. 2013. Cortical connectivity and sensory coding. *Nature* **503**:51–8. doi:10.1038/nature12654

Hawkins J, Ahmad S. 2017. Why Neurons Have Thousands of Synapses, a Theory of Sequence Memory in Neocortex. *Frontiers in neural circuits* **10**:1–13. doi:10.3389/fncir.2016.00023

Hofer SB, Ko H, Pichler B, Vogelstein JT, Ros H, Zeng H, Lein E, Lesica N a, Mrsic-Flogel TD. 2011. Differential connectivity and response dynamics of excitatory and inhibitory neurons in visual cortex. *Nature neuroscience* **14**:1045–1052. doi:10.1038/nn.2876

Huang C, Englitz B, Reznik A, Zeldenrust F, Tansu Celikel. 2020. Information transfer and recovery for the sense of touch. *In preparation*.

Huang C, Resnik A, Celikel T, Englitz B. 2016. Adaptive Spike Threshold Enables Robust and Temporally Precise Neuronal Encoding. *PLoS Computational Biology* **12**. doi:10.1371/journal.pcbi.1004984

Isaacson JS, Scanziani M. 2011. How inhibition shapes cortical activity. *Neuron* **72**:231–43. doi:10.1016/j.neuron.2011.09.027

Jolivet R, Lewis TJ, Gerstner W. 2004. Generalized integrate-and-fire models of neuronal activity approximate spike trains of a detailed model to a high degree of accuracy. *Journal of neurophysiology* **92**:959–76. doi:10.1152/jn.00190.2004

Ko H, Cossell L, Baragli C, Antolik J, Clopath C, Hofer SB, Mrsic-Flogel TD. 2013. The emergence of functional microcircuits in visual cortex. *Nature* **496**:96–100. doi:10.1038/nature12015

Ko H, Hofer SB, Pichler B, Buchanan K a, Sjöström PJ, Mrsic-Flogel TD. 2011. Functional specificity of local synaptic connections in neocortical networks. *Nature* **473**:87–91. doi:10.1038/nature09880

Koch C, Segev I. 2000. The role of single neurons in information processing. *Nature Neuroscience* **3**:1171–1177. doi:10.1038/81444

Kole MHP, Stuart GJ. 2012. Signal Processing in the Axon Initial Segment. *Neuron* **73**:235–247. doi:10.1016/j.neuron.2012.01.007

Lochmann T, Denève S. 2008. Information transmission with spiking Bayesian neurons. *New Journal of Physics* **10**:055019. doi:10.1088/1367-2630/10/5/055019

Markram H, Toledo-rodriguez M, Wang Y, Gupta A, Silberberg G, Wu C. 2004. Interneurons of the Neocortical Inhibitory System. *Nature reviews Neuroscience* **5**:793–807. doi:10.1038/nrn1519

Murray PD, Keller A. 2011. Somatosensory response properties of excitatory and inhibitory neurons in rat motor cortex. *J Neurophysiol* **106**:1355–1362.

Nowak LG, Sanchez-Vives MV, McCormick DA. 2008. Lack of orientation and direction selectivity in a subgroup of fast-spiking inhibitory interneurons: Cellular and synaptic mechanisms and comparison with other electrophysiological cell types. *Cerebral Cortex* **18**:1058–1078. doi:10.1093/cercor/bhm137

Okun M, Steinmetz N a., Cossell L, Iacaruso MF, Ko H, Barthó P, Moore T, Hofer SB, Mrsic-Flogel TD, Carandini M, Harris KD. 2015. Diverse coupling of neurons to populations in sensory cortex. *Nature* **521**:511–515. doi:10.1038/nature14273

Olshausen BA, Field DJ. 1996. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature* **381**:607–609.

Paninski L. 2003. Convergence properties of three spike-triggered analysis techniques. *Network: Computation in Neural Systems* **14**:437–64.

Priebe NJ, Ferster D. 2008. Inhibition, Spike Threshold, and Stimulus Selectivity in Primary Visual Cortex. *Neuron* **57**:482–497. doi:10.1016/j.neuron.2008.02.005

Ranjbar-Slamloo Y, Arabzadeh E. 2019. Diverse tuning underlies sparse activity in layer 2/3 vibrissal cortex of awake mice. *Journal of Physiology* **597**:2803–2817. doi:10.1113/JP277506

Rauch A, La Camera G, Luscher H-R, Senn W, Fusi S. 2003. Neocortical pyramidal cells respond as integrate-and-fire neurons to in vivo-like input currents. *Journal of neurophysiology* **90**:1598–612. doi:10.1152/jn.00293.2003

Reyes-Puerta V, Kim S, Sun J, Imbrosci B, Kilb W. 2015. High Stimulus-Related Information in Barrel Cortex Inhibitory Interneurons. *PLoS Computational Biology* **11**:e1004121. doi:10.1371/journal.pcbi.1004121

Rieke F, Warland D, de Ruyter van Steveninck RR, Bialek W. 1997. Spikes: exploring the neural code. Cambridge, Massachusets: MIT Press.

Rose D, Blakemore C. 1974. Effects of bicuculline on functions of inhibition in visual cortex. *Nature* **249**:375–377.

Rossant C, Goodman DFM, Fontaine B, Platkiewicz J, Magnusson AK, Brette R. 2011. Fitting neuron models to spike trains. *Frontiers in neuroscience* **5**:9. doi:10.3389/fnins.2011.00009

Rowan MJM, DelCanto G, Yu JJ, Kamasawa N, Christie JM. 2016. Synapse-Level Determination of Action Potential Duration by K(+) Channel Clustering in Axons. *Neuron* **91**:370–383.

Rowan MJM, Tranquil E, Christie JM. 2014. Distinct Kv channel subtypes contribute to differences in spike signaling properties in the axon initial segment and presynaptic boutons of cerebellar interneurons. *J Neurosci* **34**:6611–6623.

Rudy B, McBain CJ. 2001. Kv3 channels: voltage-gated K+ channels designed for high-frequency repetitive firing. *Trends Neurosci* **24**:517–526.

Sanchez-Vives MV, Nowak LG, McCormick D a. 2000. Cellular mechanisms of long-lasting adaptation in visual cortical neurons in vitro. *The Journal of Neuroscience* **20**:4286–4299. doi:20/11/4286 [pii]

Sharpee TO, Rust NC, Bialek W. 2004. Analyzing neural responses to natural signals: maximally informative dimensions. *Neural Computation* **16**:223–250. doi:10.1162/089976604322742010

Simoncelli EP, Paninski L, Pillow JW, Schwartz O. 2004. Characterization of Neural Responses with Stochastic Stimuli In: Gazzaniga M, editor. The Cognitive Neurosciences. MIT Press. p. 1385.

Strong SP, Koberle R, de Ruyter van Steveninck RR, Bialek W. 1998. Entropy and Information in Neural Spike Trains. *Physical Review Letters* **80**:197–200.

Tan AYY, Brown BD, Scholl B, Mohanty D, Priebe NJ. 2011. Orientation Selectivity of Synaptic Input to Neurons in Mouse and Cat Primary Visual Cortex. *The Journal of Neuroscience* **31**:12339–12350. doi:10.1523/JNEUROSCI.2039-11.2011

Wu GK, Arbuckle R, Liu B hua, Tao HW, Zhang LI. 2008. Lateral Sharpening of Cortical Frequency Tuning by Approximately Balanced Inhibition. *Neuron* **58**:132–143. doi:10.1016/j.neuron.2008.01.035

Wu GK, Tao HW, Zhang LI. 2011. From elementary synaptic circuits to information processing in primary auditory cortex. *Neuroscience and Biobehavioral Reviews* **35**:2094–2104. doi:10.1016/j.neubiorev.2011.05.004

Zeldenrust F, de Knecht S, Wadman WJ, Denève S, Gutkin BS. 2017. Estimating the Information Extracted by a Single Spiking Neuron from a Continuous Input Time Series. *Frontiers in Computational Neuroscience* **11**:49. doi:0.3389/fncom.2017.00049

Zeldenrust F, Gutkin B, Denève S. 2019. Efficient and robust coding in heterogeneous recurrent networks. *bioRxiv*. doi:https://doi.org/10.1101/804864

Figures

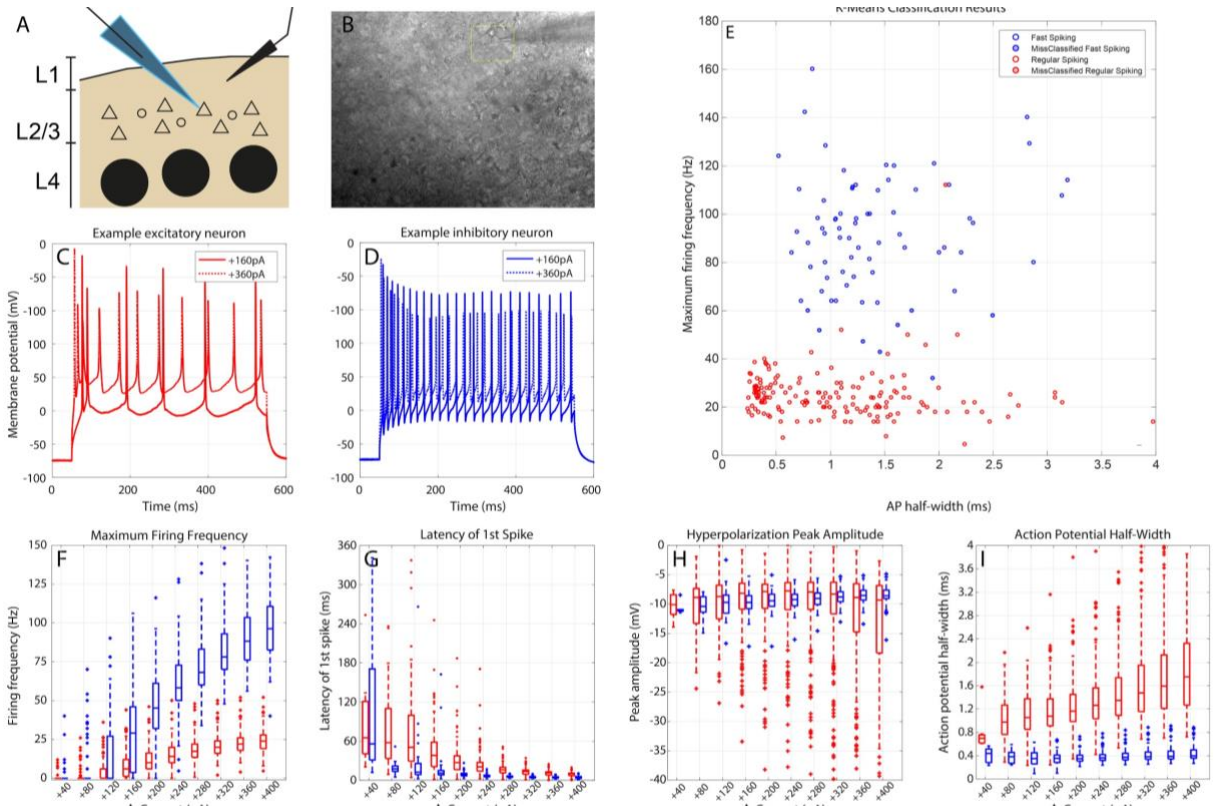

Fig 1. Cell classification. A, B) We selected neurons to record from the mouse somatosensory cortex (barrel cortex), in L2/3. Visually, the shape and size of soma were a good indicator of the cell type: smaller and roundish shapes would point towards fast spiking neurons, while slightly larger and triangular shapes would point to regular spiking (pyramidal) neurons. C) Example responses of an excitatory cell to a constant injected current. D) Example responses of an inhibitory cell to a constant injected current. E) Cell classification using agglomerative clustering based on the maximum firing frequency and spike width. Cells were classified as inhibitory (blue) when they had a small spike half-width combined with a high maximum firing rate, and as excitatory (red) with a large spike half-width and low maximum firing rate. There was one cell (pink star) in between the clusters, where the agglomerative clustering and the initial classification disagreed. This cell was classified as an inhibitory neuron (see Materials & Methods). F) Maximum firing frequency distribution for incremental current injection amplitudes for inhibitory (blue) and excitatory (red) neurons. G) Same as F), but for the latency of the first spike. H) After-hyperpolarization distribution. I) Spike half-width distribution. For threshold behaviour in the current-clamp step-and-hold protocol, see Supplementary Fig. S1.

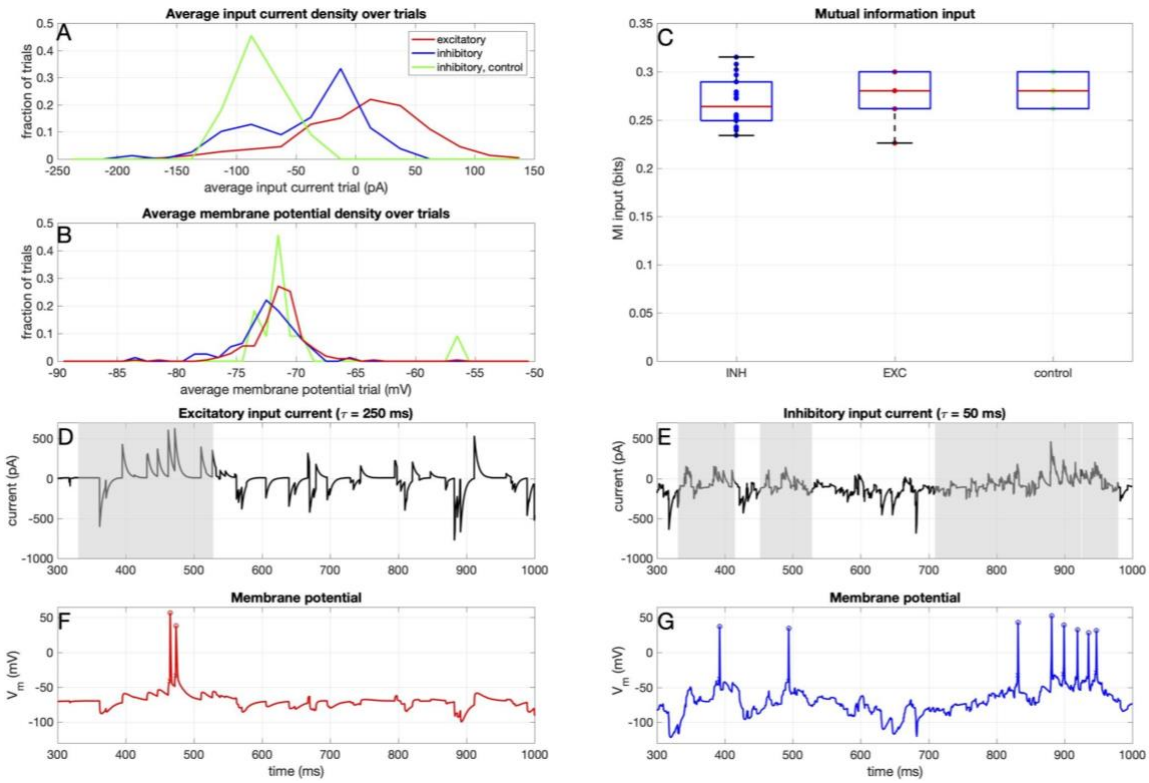

Fig 2. Input. A) Average (over the trial) input current and B) membrane potential for all trials. Green data points/lines denote the control experiments where the inhibitory neurons received the input current that was otherwise given to the excitatory neurons. C) Mutual information between the hidden state and the input current, for all trials. Note that because frozen noise was used, every frozen noise trial was actually the same. Therefore, there are not many different realizations and hence not many different MI values. D) Example injected frozen noise current for an excitatory neuron. The grey shaded area corresponds to times when the hidden state was 1. E) Example injected frozen noise current for an inhibitory neuron. F) Example resulting membrane potential of an excitatory neuron. G) Example resulting membrane potential of an inhibitory neuron.

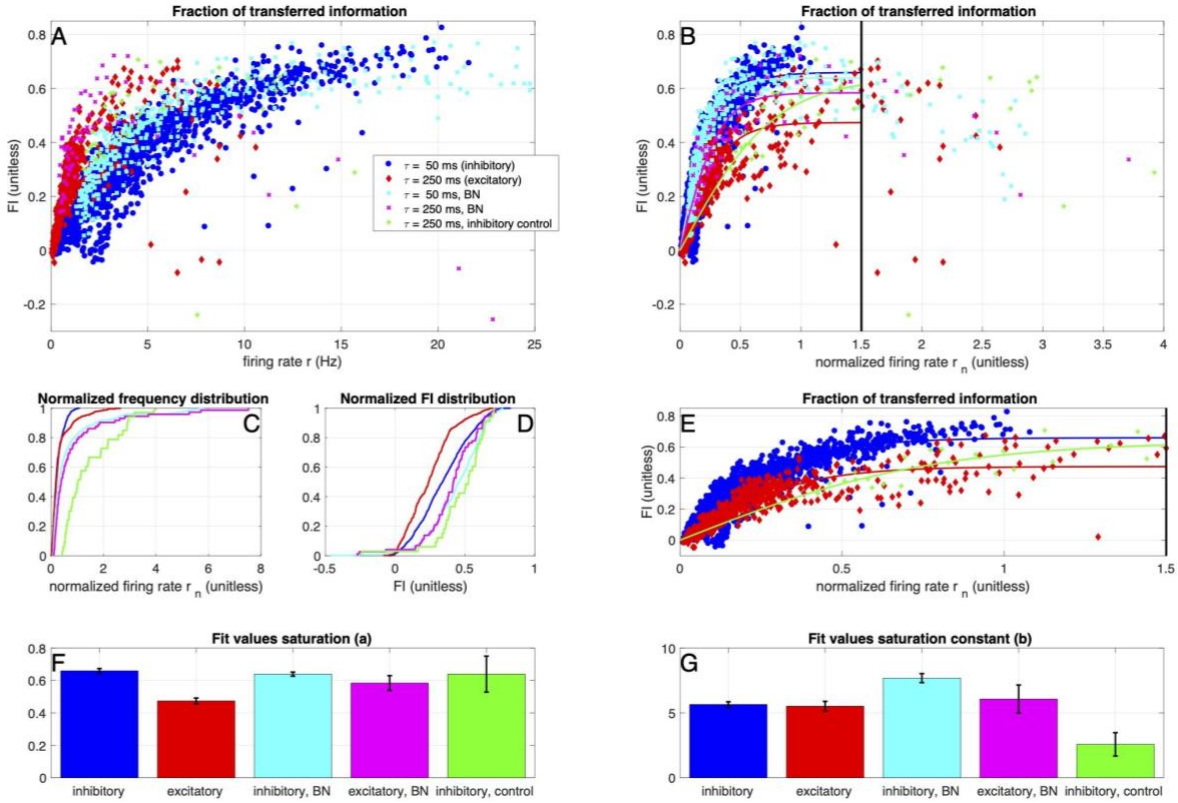

Fig 3: Inhibitory neurons transfer more information. A) Fraction of information kept during the spike generating process (FI, see eq. (1)) as a function of the firing rate, for inhibitory neurons (blue) and excitatory neurons (red). In green, the control experiments where the inhibitory neurons received the input current that was normally given to the excitatory neurons. In turquoise and pink, the simulations with the Bayesian Neuron Materials and Methods, see table 1 for parameter values and Supplementary Fig. S2 for simulation values). B) Fraction of information kept during the spike generating process (FI), as a function of the firing rate normalized by the switching speed of the hidden state (see table 1). The solid lines denote fits of the data up to a normalized firing frequency of r_n =1.5 (eq. (2)). Colors/markers the same as in A. C) and D) Normalized firing frequency and FI distribution of the spike trains in all conditions. E) Zoom of B). F) and G) Fit values and their 95% confidence intervals (error bars) for parameter a (F) and b (G). Data from 144 excitatory neurons (220 trials), 72 inhibitory neurons (78 trials) and 9 control inhibitory neurons (11 trials).

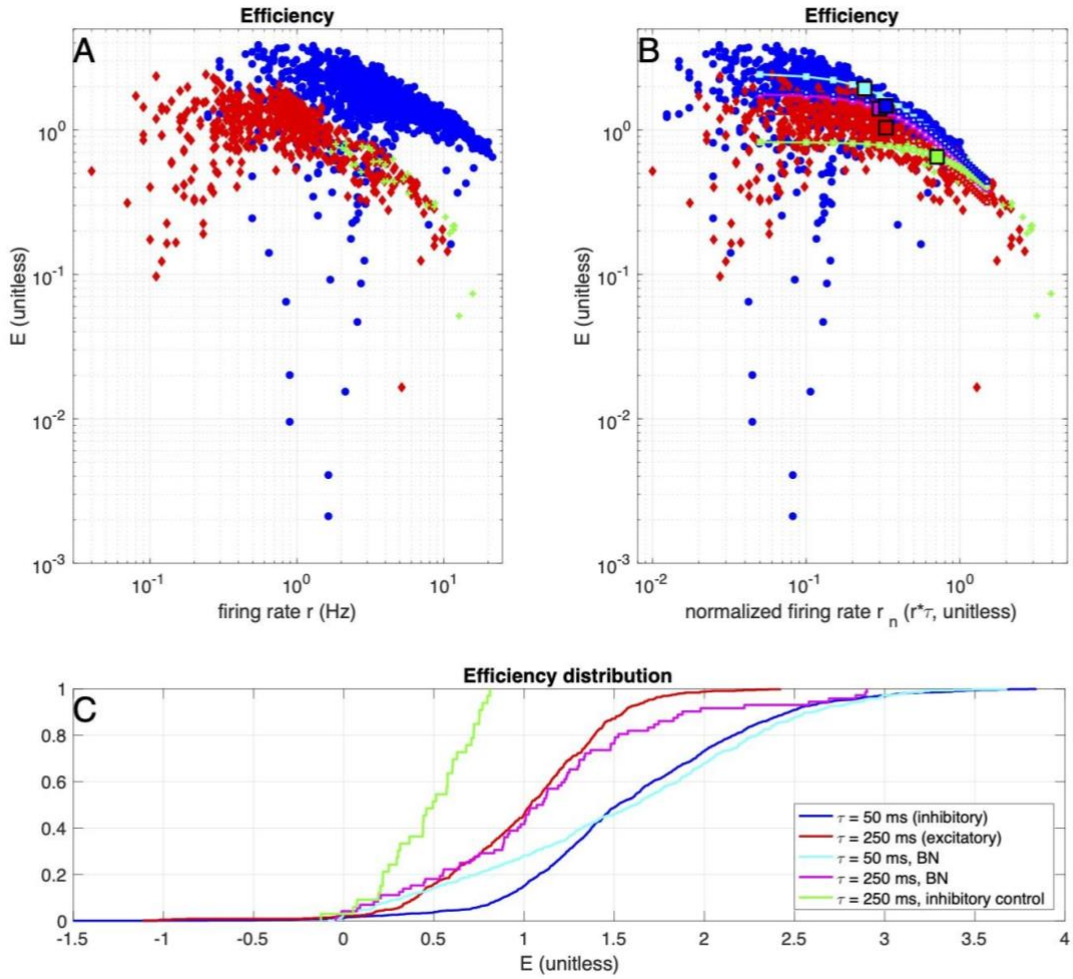

Fig. 4 Inhibitory neurons transfer information more efficiently. A) Efficiency E (eq. (3)) as a function of the firing rate, for inhibitory neurons (blue) and excitatory neurons (red). In green, the control experiments where the inhibitory neurons received the input current that was normally given to the excitatory neurons. B) Same as in A), but now as a function of the normalized firing rate. The lines with white squares denote the fitted curves from Fig 3B and their inflection points (large squares). The fitted values for the theoretical `Bayesian neuron' (see Materials & Methods) are shown in pink and turquoise (parameter values: see table 1, simulation values: see Fig S1). C) Cumulative distribution of the efficiency E. Data from 144 excitatory neurons (220 trials), 72 inhibitory neurons (78 trials) and 9 control inhibitory neurons (11 trials). See also Supplementary Fig. S2 for data of simulations of the Bayesian Neuron.

641

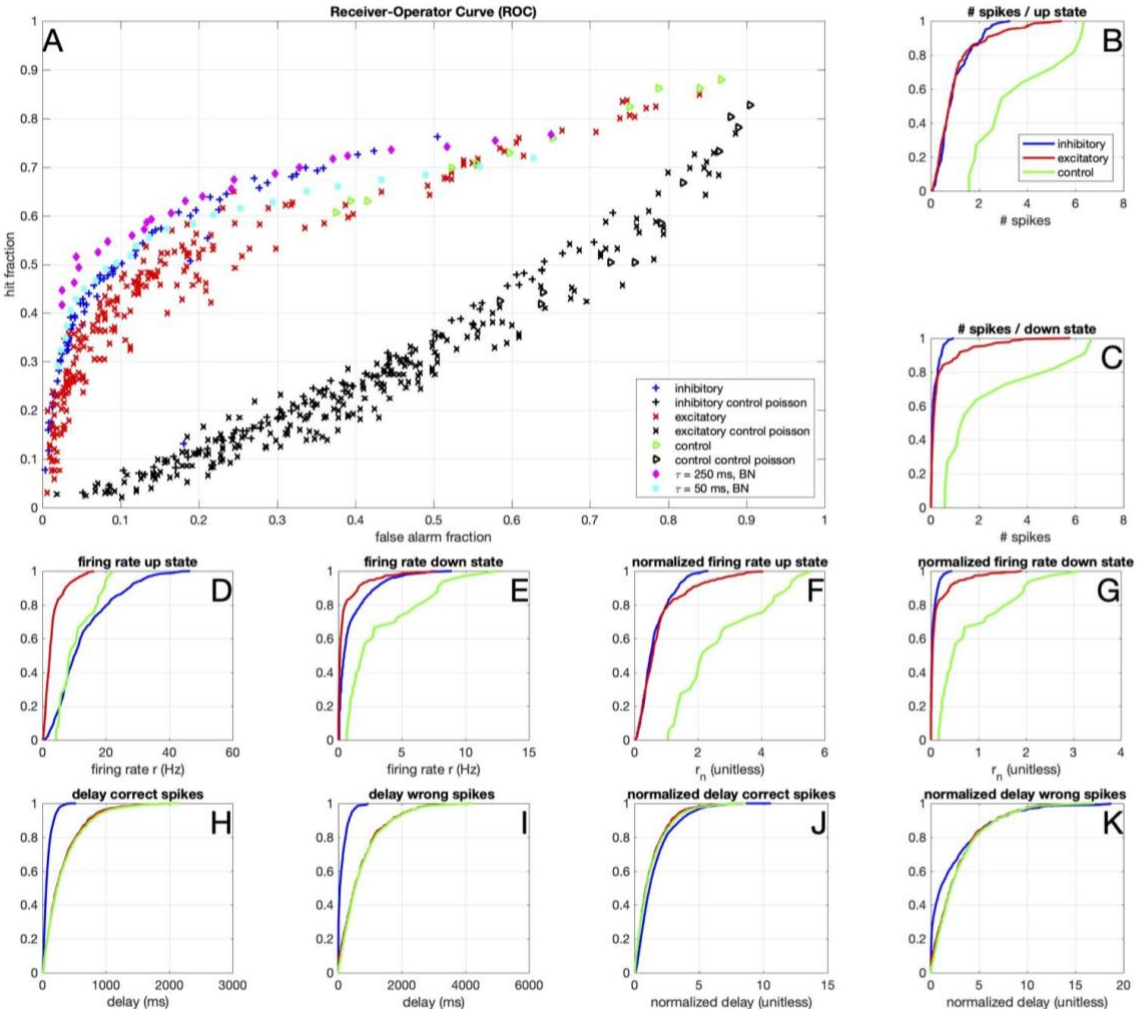

Fig. 5 Binary classification. A) Receiver Operator Curve (ROC), where the hit rate was defined as the fraction of up-states, in which at least 1 action potential was fired. Similarly, the false alarm rate was defined as the fraction of down-states, in which at least 1 action potential was fired. In black the results for Poisson spike trains with firing rates matched to those of the experimental/simulation conditions are shown. B) Distribution of the number of spikes per period where the hidden state was 1 (up state), for inhibitory neurons (blue) and excitatory neurons (red). C) Same as B), but for periods where the hidden state was 0 (down state). D) Firing rate r distribution in the up-state. E) Firing rate r distribution in the down-state.. F) Normalized firing rate r_n distribution in the up-state G) Normalized firing rate r_n distribution in the down-state. H) Delay (in ms) of each correct spike since the state switches from down to up. I) Delay (in ms) of each incorrect spike since the state switches from up to down. J) Normalized delay (delay/τ, unitless) of each correct spike since the state switch from down to up. K) Normalized delay of each incorrect spike since the state switch from up to down. Results of hypotheses test for A-F are in Supplementary Table S2 and S3. Data from 144 excitatory neurons (220 trials), 72 inhibitory neurons (78 trials) and 9 control inhibitory neurons (11 trials).

642

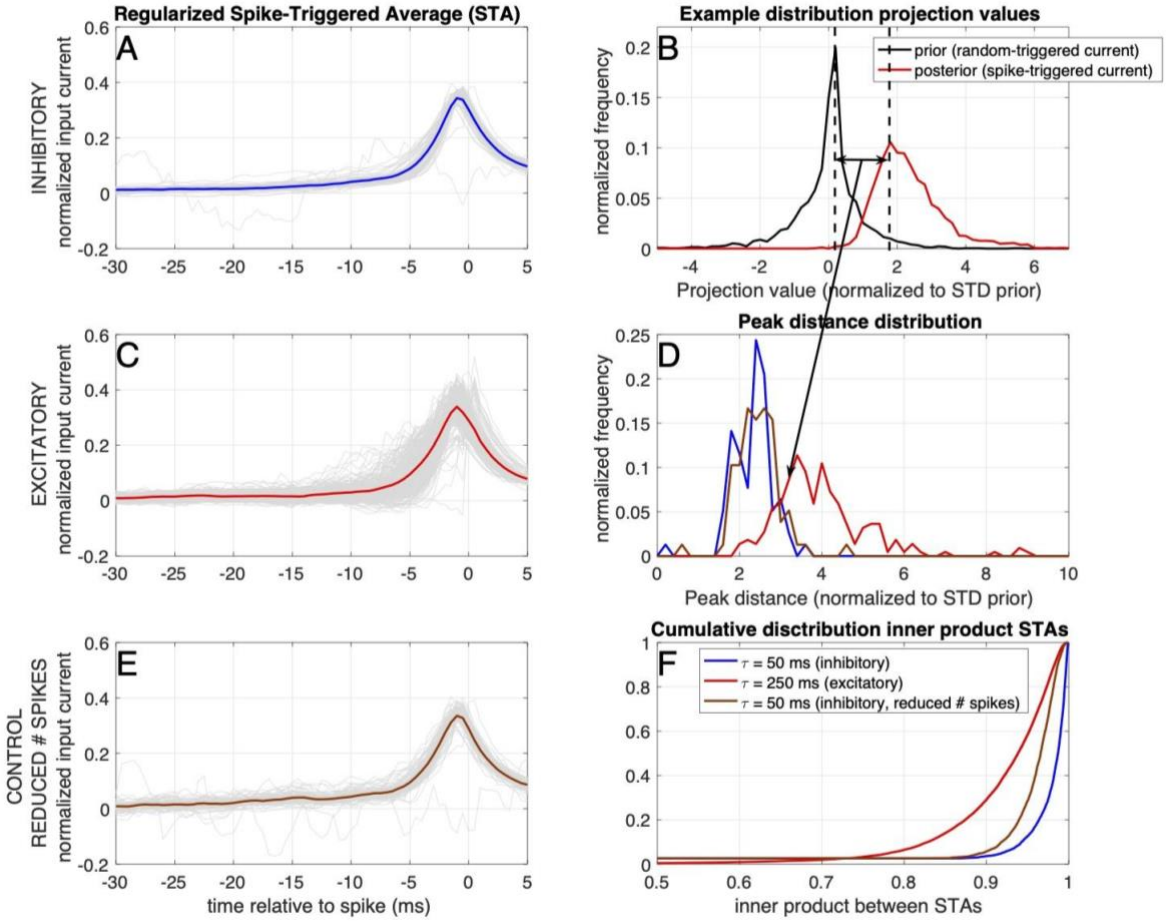

Fig 6. Linear filtering properties. A) Whitened and regularized (see Materials & Methods) spike-triggered average (STA) for inhibitory neurons. The STAs for individual neurons are shown as thin grey lines, and the average over neurons is shown as a thick coloured line. B) Example of a prior (random triggered, black line) and posterior (spike-triggered, blue line) distribution of stimulus projection values for a single inhibitory neuron. C) Same as A), but for excitatory neurons. D) Distribution of the differences between the means (see arrow in B) between the prior and posterior distribution over all neurons.-E) Same as A), but for the reduced spike trains of the inhibitory neurons. F) Distribution of the inner products between the STAs for the three groups (note that because the STAs are normalized by the L2-norm, the maximal value of the inner product is limited to 1). Data from 144 excitatory neurons (220 trials) and 72 inhibitory neurons (78 trials).

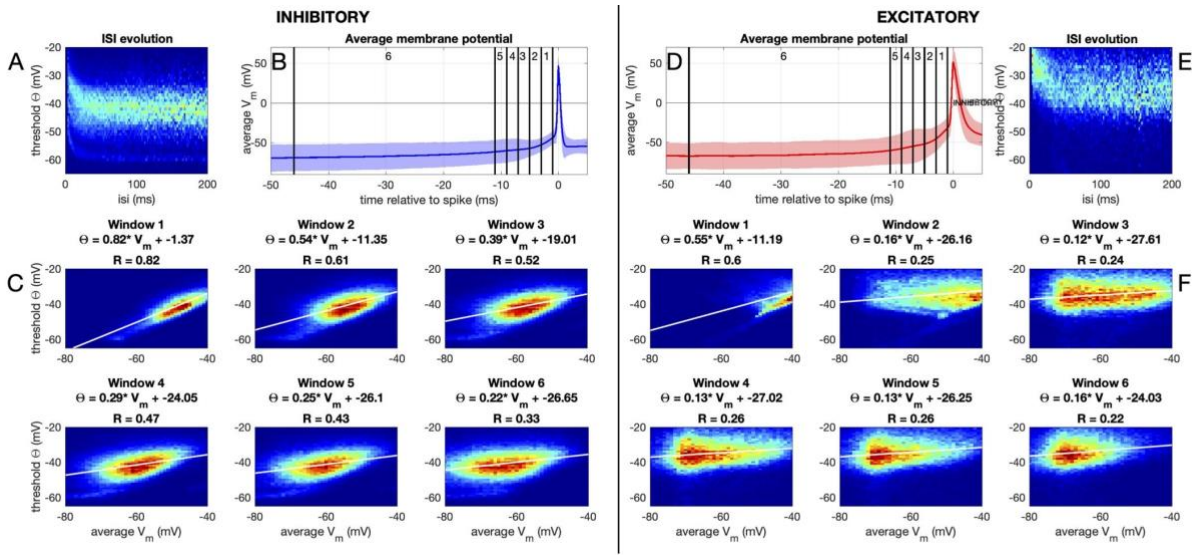

Fig. 7 Dynamic threshold A-C) Inhibitory neurons. D-F) Excitatory Neurons. A) Distribution of membrane potential threshold values (see Materials & Methods) for each inter-spike interval (ISI); normalized per ISI. B) Average spike shape (shaded region denotes standard deviation). Vertical lines denote the windows in C. C) Heatmap and regression for the relation between the threshold and the average membrane potential in the given window. D-F) Same as in A-C, but for excitatory neurons. This is all in the Frozen Noize protocol, for threshold behaviour in the current-clamp step-and-hold protocol, see Supplementary Fig. S1.

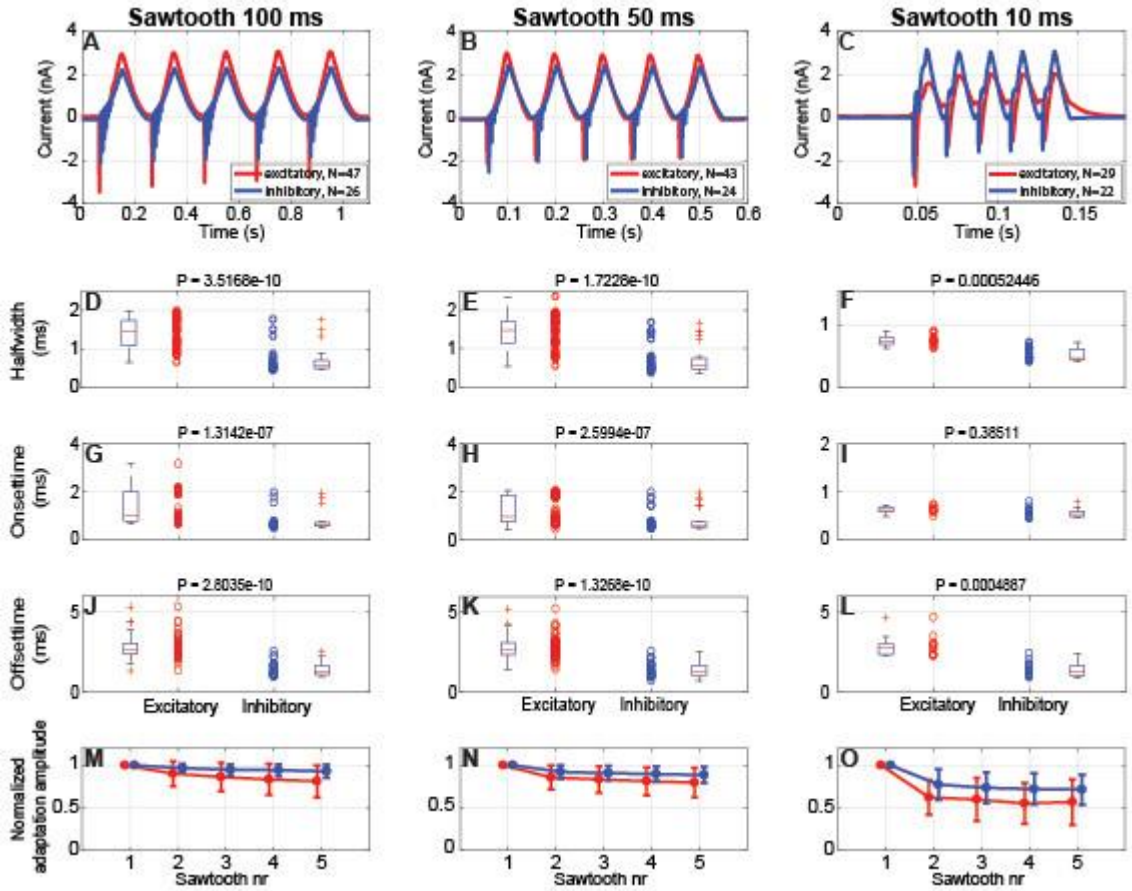

Fig. 8 Voltage clamp sawtooth Sawtooth voltage clamp experiments, where the clamped membrane potential was shifted linearly from -70 mV to +70 mV in 100 ms (left column), 50 ms (middle column) and 10 ms (right column). A-C: Example traces. D-F: Halfwidth of the first peak. G-I: Onset Time calculated by measuring the time between initiation and the maximum amplitude of the first peak. For G and H we reported the p-value for completeness, but a Kolomogorov-Smirnov test results in a relevant non-gaussianity of the distributions. J-L: Offset Time calculated by measuring the time between maximum amplitude and conclusion of the peak. M-O:Adaptation of the amplitude of the first peak for each sawtooth, compared between consecutive sawtooths (for statistical comparison, see Supplementary Table S4). All statistics displayed are computed with 2-sample t-test. Supplementary Fig. S4 shows the amplitude, initiation voltage and latency of the first peak.

645