Temporal whitening by power-law adaptation in neocortical neurons

Christian Pozzorini, Richard Naud, Skander Mensi and Wulfram Gerstner

School of Computer and Communication Sciences and School of Life Science,
Brain Mind Institute,
Ecole Polytechnique Federale de Lausanne
1015 Lausanne EPFL, Switzerland

E-mail: christian.pozzorini@epfl.ch

1 Abstract

Spike-frequency adaptation is widespread in the central nervous system, but its functional role remains unclear. In neocortical pyramidal neurons, adaptation manifests itself by an increase in neuronal firing threshold and by adaptation currents triggered after each spike. Combining electrophysiological recordings with modeling, we found that these adaptation processes last for more than 20 seconds and decay over multiple time scales according to a power-law. The power-law decay associated with adaptation mirrors and cancels the temporal correlations of input current received *in-vivo* at the soma of L2/3 somatosensory pyramidal neurons. These findings suggest that, in the cortex, spike-frequency adaptation causes temporal decorrelation of output spikes (*temporal whitening*), an energy efficient coding procedure that, at high signal-to-noise ratio, improves the information transfer.

2   Introduction

Neural signaling requires a large amount of metabolic energy [1]. Consequently, neurons are thought to communicate using efficient codes in which redundant information is discarded [2]. Theories of efficient coding [3] successfully predict several features of sensory systems. At early stages of visual processing, inputs coming from the external world are *decorrelated* both in space and time [4–7]; through sensory adaptation [8], codes are dynamically modified so as to maximize information transmission [9–12]; and sensory adaptation on multiple timescales [11, 13, 14] could possibly reflect the statistics of the external world [15].

Sensory adaptation is at least partially due to intrinsic properties of individual neurons and, in particular, to spike-frequency adaptation (SFA). SFA is not only observed at the early stages of sensory processing, but is also widespread in cortical neurons embedded in highly recurrent networks. Often modeled by a single process with one specific timescale [16, 17], SFA also occurs on multiple timescales [18–20]. In pyramidal neurons of the rat somatosensory cortex, three or more processing steps away from the sensory receptors, SFA is scale-free [21], meaning that the effective speed at which individual neurons adapt is not fixed but depends on the input. Scale-free adaptation can be captured by simple threshold models with a power-law decaying spike-triggered process [22] that possibly describes the combined action of Na-channel inactivation [23–25] and ionic channels mediating adaptation currents [26–28].

Three questions therefore arise: First, can the temporal features of spike-triggered currents and spike-triggered changes in firing threshold, possibly spanning multiple timescales, directly be extracted from experimental data? Second, can SFA be explained by these spike-triggered effects? And finally, do the timescales of SFA match the temporal statistics of the inputs received by individual neurons? If temporal characteristics of inputs and SFA were matched, SFA could lead to a perfect decorrelation of the information contained in one spike with that of the previous one of the same neuron, a phenomenon called *temporal whitening* [29]. Temporal whitening in turn implies that, at high signal-to-noise ratio, information transmission is enhanced [30].

3 Results

The question of whether SFA is optimally designed for efficient coding can only be addressed if both the dynamics of SFA and the statistical properties of the inputs generated in biologically relevant situations are known. Therefore, the *Results* section is organized as follow. We start with a combined theoretical and experimental approach so as to extract the dynamics of spike-triggered processes and SFA directly from *in-vitro* recordings of cortical neurons. Then, we analyze the synaptically driven membrane potential dynamics recorded *in-vivo* from somatosensory neurons during active whisker sensation (data from [31]). Our overall goal is to study whether adaptation optimally removes the temporal correlations in the input to single neurons embedded in the highly recurrent network of the cortex.

3.1 Spike-frequency adaptation is mediated by power-law spike-triggered processes

To reveal adaptation on multiple timescales, L5 somatosensory pyramidal neurons were stimulated with sinusoidal noisy currents of period T (see Methods) chosen between 500 ms and 16 s (Fig. 1a). For each cell, the baseline I_0 and the amplitude ΔI_{mean} of the sinusoidal waves were calibrated to produce an output firing rate oscillating between 2 and 6 Hz and a time-averaged mean of $r_0 = 4$ Hz. Single neurons responded with a firing rate $r(t)$ characterized by fast fluctuations around a sinusoidal mean $r_{\mathrm{mean}}(t)$ given by the first-order approximation

$$r_{\mathrm{mean}}(t) = r_0 + \hat{H}_{\mathrm{A}}(w) \cdot \Delta I_{\mathrm{mean}} \sin(wt + \hat{H}_{\Phi}(w)) \tag{1}$$

where $w = 2\pi/T$ is the angular frequency of the input modulation, r_0 is the average firing rate, $\hat{H}_{\mathrm{A}}(w)$ is the amplitude response and $\hat{H}_{\Phi}(w)$ is the phase response. In the Fourier domain, the transfer function $\hat{H}(w) = \hat{H}_{\mathrm{A}}(w)e^{i\hat{H}_{\Phi}(w)}$ constitutes a linear model for the modulation of the output firing rate (Fig. 1).
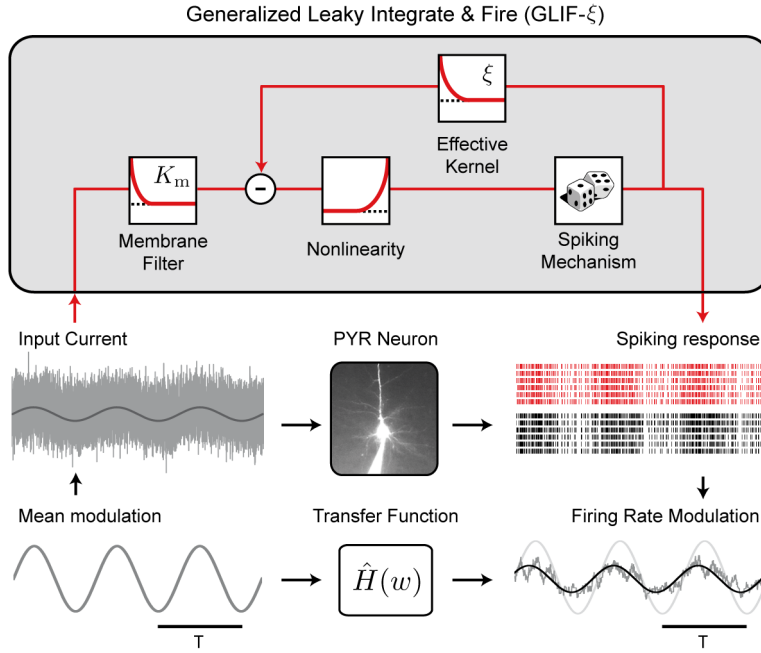
Figure 1: Experimental protocol and spiking neuron model GLIF-ξ. To reveal SFA on multiple timescales, synaptically isolated L5 pyramidal neurons (*PYR Neuron*) were repeatedly stimulated with fluctuating currents (*Input Current*) generated by adding filtered Gaussian noise to sinusoidal waves with different angular frequencies $w = 2\pi/T$ (*Mean Modulation*). The horizontal bars (bottom left and right) indicates the period T of modulation. The single neuron response (*Spiking response*, black) was recorded intracellularly and the firing rate $r(t)$ was estimated by counting the number of spikes in every time bin (*Firing Rate Modulation*, gray). The periodic oscillations of the firing rate $r_{\mathrm{mean}}(t)$ (*Firing Rate Modulation*, black) is related to the mean input (*Firing Rate Modulation*, light gray) by a linear rate model defined in the Fourier domain by the the transfer function $\hat{H}(w)$. The intracellular recordings were used to fit the Generalized Leaky Integrate & Fire model GLIF-ξ (black-lined box, top). In this model, the input current is first low-pass filtered by the membrane filter $K_m(t)$ and then transformed into a firing intensity by an exponential nonlinearity. Spikes are emitted stochastically (*Spiking response*, red) and trigger an adaptation process described by the effective adaptation kernel $\xi(s)$.

Since SFA is at least partly due to spike-triggered effects, the simple firing rate picture of Equation 1 must be complemented by a spike-based description. Intracellular recordings were therefore used to fit a generalized leaky integrate-and-fire model (GLIF-ξ) with escape-rate noise [32] for stochastic spike emission (Fig. 1a). To capture spike-triggered adaptation, the model features an *effective* dynamic threshold, described by the function $\xi(s)$. This function (also called effective adaptation *filter* or *kernel*) summarizes the stereotypical sequence of biophysical events triggered by the emission of an action potential and accounts for both adaptation currents and *physiological* changes of the firing threshold. Since the effects induced by consecutive spikes accumulate, the effective dynamic threshold produces SFA. Importantly, the functional shape of

$\xi(s)$, like all the other parameters of the model, were extracted from the data (see Methods and ref. [33]).
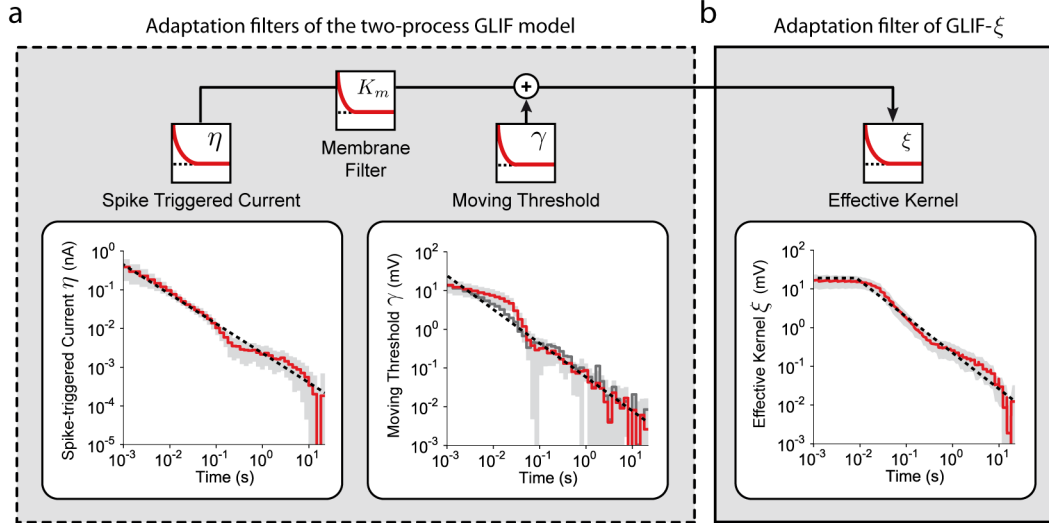

Figure 2: Adaptation filter of the GLIF-ξ model extracted from *in-vitro* recordings. a: To obtain the effective adaptation filter $\xi(s)$ of the GLIF-ξ model, intracellular recordings were first fitted with a two process GLIF model which accounts for SFA with both a spike-triggered current $\eta(s)$ and spike-triggered movement of the firing threshold $\gamma(s)$. **Left**: The mean spike-triggered current $\eta(s)$ obtained by averaging the results of different cells ($n = 14$) is shown in red on a log-log scale. The dashed black line shows the fit of a power-law function $\eta_{\mathrm{PL}}(s) = \alpha_\eta s^{-\beta_\eta}$ with parameters $\alpha_\eta = 0.44$ nA, $\beta_\eta = 0.76$ and s in milliseconds. **Right**: The mean moving threshold $\gamma(s)$ obtained by averaging the results of different cells ($n = 14$) is shown in red on a log-log scale. The dashed black line shows the fit of a power-law function $\gamma_{\mathrm{PL}}(s) = \alpha_\gamma s^{-\beta_\gamma}$ with parameters $\alpha_\gamma = 24.4$ mV, $\beta_\gamma = 0.87$ and s in milliseconds. The dark gray line is a control showing an independent estimation of the average moving threshold $\gamma(t)$ obtained with an alternative fitting procedure (see Supplementary Material). **b:** As illustrated by the block diagram, the spike-triggered current $\eta(s)$ and the moving threshold $\xi(s)$ are combined to obtain the effective adaptation filter $\xi(s)$ of the GLIF-ξ model (see Methods). The mean adaptation filter $\xi_{\mathrm{L}}(s)$ (red, GLIF-ξ_{L}) obtained by averaging the effective spike-triggered adaptation measured in individual cells ($n = 14$, see Supplementary Fig. S7) is shown on a log-log scale. The optimal fit of a truncated power-law $\xi_{\mathrm{PL}}(s)$ (dashed black, GLIF-ξ_{PL}) yields an exponent $\beta_\xi = 0.93$ (c.f. Eq. 2). In all panels, the gray area indicates one standard deviation for the distribution of filters across different cells (asymmetric errors are due to log-scales).

As previously reported [33], neocortical pyramidal neurons adapt their firing rates by means of two distinct biophysical mechanisms that respectively increase the firing threshold and lower the membrane potential after each spike. To get an accurate estimation of the effective adaptation filer $\xi(s)$, intracellular recordings were first fitted with a two-process model that explicitly features both a dynamic threshold and an adaptation current, described by the filters $\gamma(s)$ and

$\eta(s)$, respectively [33] (see Methods and Supplementary Fig. S4). Since in the model the emission of action potentials only depends on the difference between the membrane potential and the firing threshold, spike-triggered currents $\eta(s)$ and movements of the firing threshold $\gamma(s)$ can then be combined to obtain the effective adaptation filter $\xi(s)$ of the more parsimonious model GLIF-ξ (see Methods and Fig. 2).

We found that 22 seconds after the emission of an action potential a small but significant deflection remains in both the spike-triggered current $\eta(t)$ and the moving threshold $\gamma(t)$. Moreover, when displayed on log-log scales, the decay of both adaptation kernels is approximately linear over four orders of magnitude, meaning that both the adaptation current and the moving threshold are characterized by scale-free spike-triggered dynamics (Fig. 2a). Fitting $\eta(t)$ and $\gamma(t)$ with a power-law function (i.e. $f_{\mathrm{PL}}(t) = \alpha_{\mathrm{f}} t^{-\beta_{\mathrm{f}}}$), revealed that both spike-triggered processes have similar scaling exponents ($\beta_{\eta} = 0.76$, $\beta_{\gamma} = 0.87$). Consequently, the effective adaptation filter $\xi(t)$ is well described by a truncated power-law

$$
\xi_{\mathrm{PL}}(t) = \begin{cases} \alpha_{\xi} \cdot \left(\frac{t}{T_{\xi}}\right)^{-\beta_{\xi}} & \text{if } t > T_{\xi} \\ \alpha_{\xi} & \text{if } 0 < t \le T_{\xi} \end{cases} \tag{2}
$$

with parameters $\alpha_{\xi} = 19.2$ mV, $\beta_{\xi} = 0.93$ and $T_{\xi} = 8.3$ ms for the average kernel (Fig. 2b) and slightly different values for each individual cell (Supplementary Fig. S7), indicating that scale-free SFA is an intrinsic property of individual neurons and not only of the average over several cells.

In the following, we will refer to a model with a single spike-triggered adaptation filter as GLIF-ξ_L, where GLIF stands for Generalized Leaky Integrate and Fire and ξ_L indicates that SFA is implemented by a 22-second *Long* filter obtained by combining the moving threshold and the spike-triggered current extracted from the experimental data. With the same logic, we denote GLIF-ξ_{PL} a model in which the effective filter $\xi(s)$ is described by the truncated power law ξ_{PL} defined by Equation 2. A list with all the GLIF-ξ_{PL} model parameters is given in Table 1.

3.2 Power-law spike-triggered adaptation explains the single neuron activity on short timescales

Valid single neuron models should predict the occurrence of individual spikes with a millisecond precision [34]. In response to a single injection of a fluctuating current (Fig. 3a) the neuron emits spikes that the GLIF-ξ_L model is able to predict with a high degree of accuracy (Fig. 3b, red). When the same current is injected repetitively, the spiking responses reveal the stochastic nature of single neurons: certain action potentials are emitted reliably with a high temporal precision, while others do not occur at each repetition or are characterized by larger temporal jitters. The GLIF-ξ_L model also captures this aspect (Fig. 3c). To validate our model, its predictive power was quantified using a similarity measure denoted M_d^* (see Methods and ref. [35]). On average, GLIF-ξ_L was able to predict more than 80% of the spikes ($M_d^* = 0.807$, s.d. $= 0.04$) with a precision of ± 4 ms (see Supplementary Fig. S6g). Very similar performances, statistically not different ($n=12$, Student t-test, p=0.89), were achieved by GLIF-ξ_{PL} ($M_d^* = 0.804$, s.d. $= 0.05$), indicating that spike-triggered processes are well described by a truncated power law (see Supplementary Fig. S5d).

As expected, the subthreshold response observed *in-vitro* is systematically overestimated by GLIF-ξ_L (Fig. 3b, red). This is explained by the fact that GLIF-ξ_L artificially translates spike-triggered currents into effective threshold movements. In a two-process model, where adaptation currents and threshold movements are described as two distinct features (i.e. when $\eta(t)$ and $\gamma(t)$ are not combined in a single effective kernel), model prediction of the membrane voltage and experimental data are indeed in good agreement (Fig. 3b, gray), confirming the validity of our fitting procedure. In terms of mere spike-timing prediction, the two-process GLIF model and the more parsimonious GLIF-ξ model are equivalent (Fig. 3c). For this reason, we work in the following with single-process model GLIF-ξ.

Overall, the spike time prediction paradigm demonstrates the ability of both GLIF-ξ_L and GLIF-ξ_{PL} to capture the spiking activity of single neurons on the timescale of milliseconds.
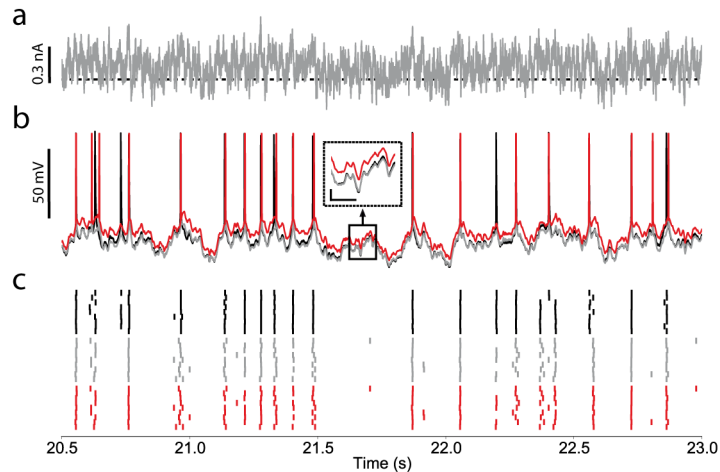
Figure 3: The GLIF-ξ model predicts the occurrence of single spikes with a millisecond precision. a: Typical 2.5-second segment of injected current. The same fluctuating current is presented several times (*frozen-noise*). The dashed black line represents 0 nA. **b:** The spiking response, but not the subthreshold membrane potential, predicted by the GLIF-ξ_L model (red) is in close agreement with the experimental data (black). In the two process-model (gray), where spike-triggered currents and threshold movements are modeled by two distinct processes (i.e. $\gamma(s)$ and $\eta(s)$), the dynamics of the subthreshold membrane potential predicted by the model is in excellent agreement with the experimental data. Inset: comparison of subthreshold membrane potential (scale bars: 40 ms, 5 mV). **c:** The raster plots show the spiking response of both the neuron (black) and the GLIF-ξ_L model (red) to repetitive presentation of the same current. By construction, the spiking response of the GLIF-ξ_L model is identical to that of the two-process model (gray).

3.3 Power-law spike-triggered adaptation explains the single neuron activity on slow timescales

We wondered whether the 22-second long adaptation filter $\xi(s)$ can also predict the firing rate modulation on the much slower timescale of seconds. To this end, the GLIF-ξ_L model fitted on responses to different frequencies of modulation ($0.5 \leq T \leq 16$ s) was used to predict the firing rates recorded in the second part of the experiment, where one of the two slowest modulations ($T = 8$ or 16 s) was chosen and repetitively presented to the cell. Comparison of the raster plots obtained by injecting the same current in both the neuron and the GLIF-ξ_L model shows that the spiking activity of the real neuron closely resembles the one predicted by the model (Fig. 4a-c). Furthermore, the match between the running-mean PSTHs constructed for the model and the experimental data reveals that both responses share a similar phase advance (Fig. 4d), indicating that our GLIF-ξ_L model is sufficient to capture the characteristic signature of SFA

under slow sinusoidal stimulation [21].

To study the role of the 22-second long adaptation filter of GLIF-ξ_L, we then fitted the same single-process model under the assumption that the adaptation filter $\xi(t)$ has a duration of only 1 second (GLIF-ξ_S, where S stands for *short* adaptation filter). Compared to GLIF-ξ_L, the firing rate predicted by GLIF-ξ_S (Fig. 4e, green) appears to be in phase with the input (Fig. 4e, dark gray) and not with the spike output of the cells, indicating that GLIF-ξ_S is unable to capture the slow components of SFA (i.e. the model with a short adaptation filter predicts a wrong phase advance). To provide even stronger evidence, the ability to predict both the mean firing rate r_0 (Fig. 4f) and the phase lead $\hat{H}_\Phi$ (Fig. 4g) was systematically quantified. Whereas the GLIF-ξ_L model is capable of very good predictions which are in statistical agreement with the experimental data (errors $\Delta r_0 = -0.01$ Hz, s.d. $= 0.67$; $n=12$, Student t-test, $t = -0.04$, $p = 0.97$ and $\Delta \hat{H}_\Phi = -0.17$ deg, s.d. $= 5.7$; $n=12$, Student t-test, $t = -0.10$, $p = 0.92$), GLIF-ξ_S has the tendency to both overestimate the average firing rate and underestimate the phase advance (errors $\Delta r_0 = 0.47$ Hz, s.d. $= 0.72$; $n=12$, Student t-test, $t = 2.16$, $p = 0.05$ and $\Delta \hat{H}_\Phi = -17.9$ deg, s.d. $= 6.5$ deg; $n=12$, Student t-test, $t = -9.16$, $p < 10^{-6}$), demonstrating that an adaptation filter of 1 second is not sufficient.
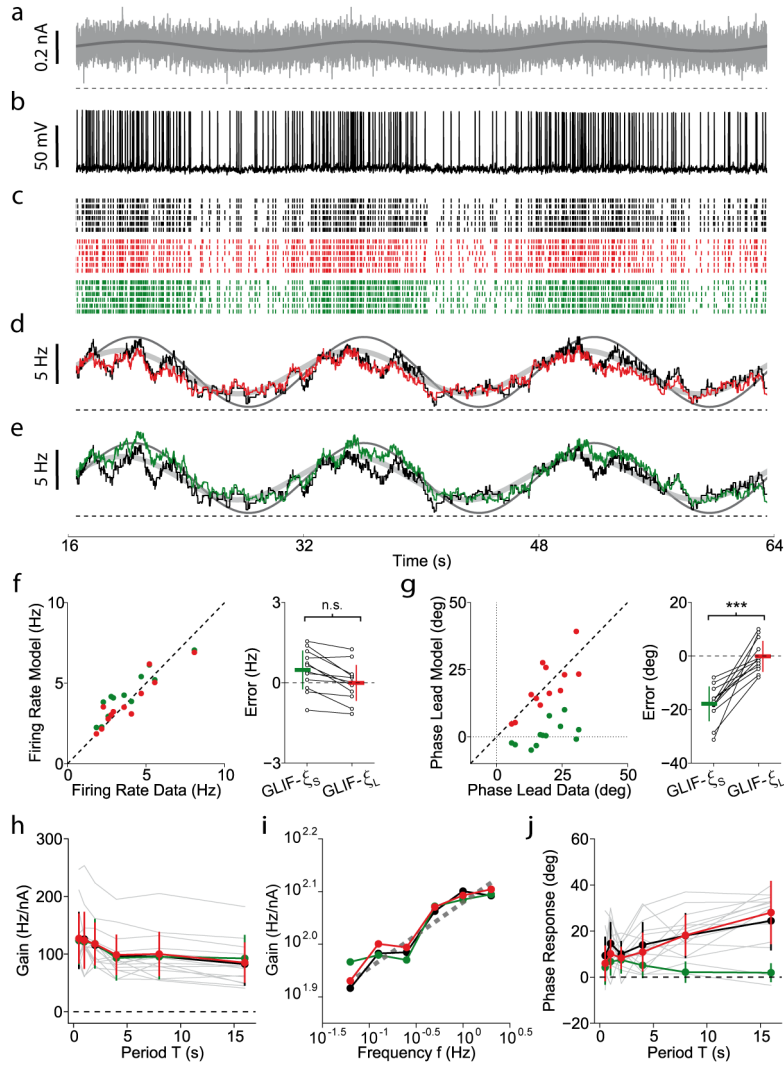
Figure 4: The GLIF-ξ model accurately predicts the firing rate response on multiple timescales.
a: An input current consisting of a *frozen-noise* signal on top of a small sinusoidal wave of period $T = 16$ s
(dark gray) is presented several times to the same L5 pyramidal neuron. **b:** Typical membrane potential response
in a single trial. **c:** The firing activity obtained by repetitive presentation of the same input signal (black) is
compared with the predictions made by GLIF-ξ_L (red) and GLIF-ξ_S (green). **d:** Data shown in panel *c* were used
to build two continuous PSTHs (spikes were filtered with a moving average over 1 second). Experimental data
(black) and GLIF-ξ_L model prediction (red) are in good agreement. The PSTHs were fitted with two sinusoidal
functions to quantify the phase shift $\hat{H}_\Phi$ (Eq. 1) with respect to the modulation of the mean input (dark gray
sinusoidal wave). For clarity, only the fit to experimental data is shown (light gray sinusoidal wave, $\hat{H}_\Phi = 17.6$
deg). The phase shift is measured with respect to the periodic input (dark gray line). **e:** The prediction made
by the GLIF-ξ_S model is shown in green. Compared to panel *d*, the prediction is less accurate. In particular, the
PSTH produced by GLIF-ξ_S is approximatively in phase with the input modulation (dark gray) but not with the
PSTH of the experimental data, meaning that the model is unable to correctly capture the correct phase advance
of $\hat{H}_\Phi = 17.6$ deg. Experimental PSTH (black), sinusoidal fit of the experimental data (light gray) and input
modulation (dark gray) are copied from panel *d*. **f:** The performances of GLIF-ξ_L and GLIF-ξ_S in predicting
the average firing rate r_0 of a new stimuli (not used for data fitting) are compared. Left: for the two models,
the predicted r_0 is plotted against the observed one. Each dot represents a different cell. Right: the error in
predicting r_0 is plotted for the two models. Each couple of open circles shows the performances of the two models
on one cell. On average, the errors produced by GLIF-ξ_L (red) and GLIF-ξ_S (green) are not significantly different
($n = 12$, Student t-test, $t = -1.62$, $p = 0.12$; n.s. $p > 0.05$). Error bars indicate one standard deviation. **g:**
The two models are significantly different if compared by measuring the ability to predict the phase shift $\hat{H}_\Phi$
($n = 12$, Student t-test, $t = 6.77$, $p = 0.8 \cdot 10^{-6}$; triple star: $p < 0.001$). Conventions are as in panel *f*. **h-j:** The
linear response of L5 pyramidal neurons to a slow modulation of the mean input was characterized by the transfer
function $\hat{H}(w)$. The gain $\hat{H}_A(w)$ (**h**) and the phase response $\hat{H}_\Phi(w)$ (**j**) are plotted as a function of the period
of modulation $T = 2\pi/w$. **i:** log-log plot of the gain $\hat{H}_A$ as a function of the input frequency $f = T^{-1}$. The
amplitude response observed *in-vitro* was fitted by power law with scaling exponent $\beta_H = 0.12$ (dashed gray). In
panels *h-j*, data from individual cells ($n = 14$, gray lines) are averaged (black) and compared with the predictions
of GLIF-ξ_L (red) and GLIF-ξ_S (green). Error bars indicate one standard deviation. The horizontal dashed lines
appearing in panels *a, d, e, h* and *j* indicate zero.

Finally, the transfer function $\hat{H}(w)$ was measured for both real neurons and spiking models by fitting Equation 1 on the firing rates observed in response to six frequencies of modulation (Fig. 4h-j). For both real neurons and GLIF-ξ_{L}, the amplitude response $\hat{H}_{\mathrm{A}}(w)$ was stronger at higher frequencies compared to lower ones revealing high-pass filtering, a characteristic feature of SFA (Fig. 4h). Consistent with observations in L2/3 pyramidal neurons [21], plotting on log-log scales the amplitude response $\hat{H}_{\mathrm{A}}$ as a function of the input frequency $f = T^{-1}$, revealed that the gain of L5 pyramidal neurons is also approximatively power law (Fig. 4i). Moreover, the phase response $\hat{H}_{\Phi}(w)$ was always positive meaning that, for all the frequencies tested in this study, the output firing rate led the input modulation (Fig. 4j). Overall, GLIF-ξ_{L} was able to capture the features of the transfer function observed in L5 pyramidal neurons. Similar results were obtained with GLIF-ξ_{PL} (Supplementary Fig. S5), confirming that the spike-triggered processes observed *in-vitro* are correctly modeled by a truncated power-law lasting 22 seconds.

The experimental results reported in Figure 4h-j are very similar to those obtained in L2/3 pyramidal neurons [21] and provide an independent evidence for multiple timescales of adaptation. However, in contrast to [21], we could reject the hypothesis that the phase response is frequency-independent (n=14, one-way ANOVA, $F = 4.38$, $p = 0.0015$) so that a model where the firing rate encodes a fractional derivative of the input is not applicable to L5 pyramidal neurons of the mouse somatosensory cortex.

Overall, these results show that accounting for long-lasting spike-triggered effects with an appropriate adaptation filter is crucial to capture the response of L5 pyramidal neurons on multiple timescales.

3.4 Power-law adaptation is optimally tuned to perform *temporal whitening*

Our model describes how the net current resulting from dendritic integration is encoded into a spike train at the soma of neocortical pyramidal neurons. To investigate the implications of power-law adaptation, we considered a situation in which a population of N uncoupled GLIF-

ξ_{PL} neurons encodes a common input $I(t) = I_0 + \Delta I(t)$ in their instantaneous firing rate $A(t)$, also called population activity. Note that, since the neurons in our population are all identical and receive the same input, the population activity $A(t)$ is identical to the PSTH measured by repetitively injecting the same current into one single cell. For relatively small fluctuations around a mean activity A_0, we can assume that the population operates in a linear regime and responds to an external input fluctuation $\Delta I(t)$ according to the first-order approximation

$$A(t) = A_0 + \int_0^t \Delta I(t-s)H(s)ds + n(t) \tag{3}$$

where the impulse response $H(t)$ is the inverse Fourier transform of $\hat{H}(w)$, the noise $n(t)$ is due to stochastic firing in a finite population and both terms depend on the intrinsic properties of the individual neurons and in particular on the precise shape of the adaptation filter $\xi(t)$.

For large populations, the noise term in Equation 3 becomes negligible and optimal coding is achieved by the removal of temporal correlations potentially present in the input [30, 36]. This encoding strategy is known as *temporal whitening* and requires the population activity to have a flat power spectrum $A(f) = \text{Const.}$ (see Supplementary Material).

SFA is known to implement high-pass filtering of the input current [37, 38]. In the particular case of power law adaptation, the population response is characterized by a power law gain (Fig. 4h,i and ref. [21]) suggesting that, in neocortical pyramidal neurons, spike-triggered processes might be optimally tuned to efficiently encode scale-free signals (i.e. signals that are temporally correlated across multiple timescales). However, the issue of whether the functional role of power law adaptation is to implement *temporal whitening* can only be solved if the statistical properties of the input received *in-vivo* by neocortical pyramidal neurons are known.

To this end, we analyzed the synaptically driven membrane potential dynamics recorded from somatosensory pyramidal neurons during active whisker sensation (see Methods). A spectral analysis performed on the data of Crochet et al. 2011 [31] revealed that, at low frequencies, the power spectrum of the subthreshold membrane potential fluctuations is characterized by a

power-law decay (Fig. 5a, red), indicating that *natural stimuli* received by somatosensory pyramidal neurons are indeed scale-free.
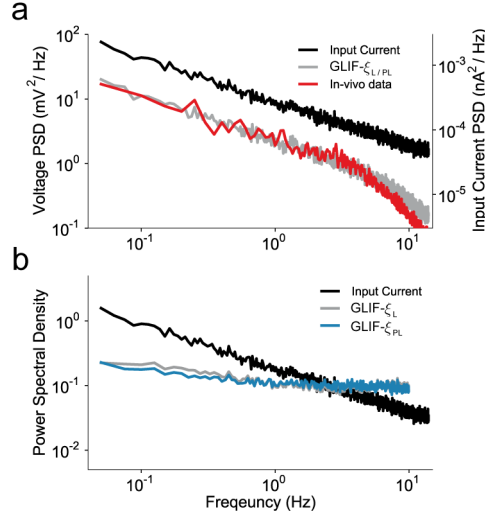

Figure 5: Power-law adaptation is near-optimally tuned to perform *temporal whitening*.
a: Power spectral density of the intracellular membrane potential fluctuations recorded *in-vivo* from L2/3 pyramidal neurons (Voltage PSD, red). The power spectrum was computed using 20-second long recordings ($n = 57$) obtained from 7 different cells (data from [31]). Fitting a power law (not shown) on the frequency band $0.05 < f < 2$ Hz yields a scaling exponent $\beta_I = 0.67$. A population of GLIF-ξ_{PL} neurons ($N = 100$) was simulated in response to a scale-free input characterized by a power-law frequency spectrum $\Delta I(f) \propto f^{-\beta_I}$, with $\beta_I = 0.67$ (Input Current PSD, black). The subthreshold response of individual GLIF-ξ_{PL} neurons is characterized by a frequency spectrum (Voltage PSD, gray) which is in good agreement with the *in-vivo* recordings. **b:** The population activity of a group of GLIF-ξ_{PL} neurons (blue) in response to an *in-vivo* like input (black, copied from panel *a*) has a nearly flat spectrum $A(f)$. Very similar results were obtained by simulating a population of GLIF-ξ_L ($N = 100$) neurons (gray). To allow a direct comparison between input and output powers, all the spectra shown in panel *b* were normalized to have the same total power. All the simulations were performed with the average parameters extracted from *in-vitro* recordings (Table 1). The mean activity was tuned to $A_0 = 4$ Hz by adjusting the baseline current to I_0. However, the precise value of A_0 does not have a big impact on the result.

To provide further evidence, the activity of a population of GLIF-ξ_{PL} neurons was simulated in response to an *in-vivo* like input characterized by a scale-free spectrum (Fig. 5a, black). The statistics of the subthreshold responses obtained in individual GLIF-ξ_{PL} neurons were consistent with the ones observed *in-vivo* (Fig. 5b, gray). Moreover, we found that the power spectrum of the population activity $A(f)$ (Fig. 5b, blue) was much closer to a horizontal line than that of the input, indicating that a population of GLIF-ξ_{PL} neurons efficiently encodes *in-vivo* like signals

by removing temporal correlations present in the input. Similar results were obtained with a population of GLIF-ξ_L neurons, where the adaptation filter $\xi(t)$ was not an idealized power law, but the average kernel extracted from intracellular recordings (Figure 5b, gray).

Overall, our results suggest that in neocortical pyramidal neurons, power-law spike-triggered adaptation mirrors and approximately cancels the temporal correlations of signals generated in a biologically relevant situation. This result provides evidence for efficient coding at the level of single neurons embedded in the highly recurrent network of the cortex.

4   Discussion

Neocortical pyramidal neurons are known to adapt their firing rate on multiple timescales [20,21]. Here we found that SFA is due to two separable spike-triggered mechanisms: each time an action potential is fired, both an adaptation current and a movement of the firing threshold are induced. Our results show that these spike-triggered effects are surprisingly long (more than 20 s) and decay with a power-law (Fig. 2), highlighting the fact that SFA does not have a specific timescale. A generalized leaky integrate-and-fire model with an effective power-law spike-triggered process simultaneously captures the fast dynamics critical for the prediction of individual spikes (Fig. 3) and the slow processes that modulate the firing rate (Fig. 4 and Supplementary Fig. S5). Most importantly, we found that, in behaving mice, the currents resulting form dendritic integration and received as input at the soma of pyramidal neurons are characterized by long-range temporal correlations that are partially removed by power-law spike-triggered adaptation (Fig. 5). This final observation indicates that, in cortex, power-law SFA is near-optimally tuned to efficient coding.

4.1   Extent of spike-triggered effects

Our results show that, 20 seconds after its emission, an individual spike can still affect the firing activity of a neuron. Possibly, spike-triggered effects have an even longer duration. However after 22 seconds the magnitudes of both the moving threshold and the spike-triggered current become too small to be measured by our method (for $t > 20$ s, $\eta(t) < 0.1$ pA and $\gamma(t) < 0.01$ mV, see Fig. 2a). Since the effects of consecutive spikes accumulate, these small contributions shape the single neuron response in a significant way (see Fig. 4).

Whereas power-law adaptation is necessary to capture the firing rate fluctuations, a model with spike-triggered processes that only last for 1 second (GLIF-ξ_S) achieves very high performances ($M_d^* = 0.80$, s.d. $= 0.03$) in predicting the occurrence of individual spikes (Supplementary Fig. S6g). This fact probably explains why power-law adaptation has not been observed in previous studies where model validation was only based on spike time prediction.

4.2 Biophysical implementation of power-law spike-triggered adaptation

Our fitting procedure enabled us to discriminate between adaptation processes implemented by spike-triggered currents and physiological changes of the firing threshold. However, the biophysical details concerning the implementation of power-law dynamics are not part of our model. In principle, power-law relaxations can be approximated by a sums of exponentials covering a wide range of timescales [20, 22]. It is therefore likely that the spike-triggered current $\eta(s)$ we found results from the combined action of multiple ion-channels operating on different timescales like for example Ca-dependent, Na-depedent and high-voltage activated potassium channels. Note, however, that a match of the relative strength of different currents implies a fine-tuned regulation of gene expression levels. In line with this hypothesis, multiple timescales of SFA have been previously modeled by biophysical models with several channels mediating adaptation currents [20, 21, 29]. Alternatively, scale-free dynamics could also be an intrinsic property of single channels. In particular, the power-law decay we found in the moving threshold $\gamma(s)$ might reflect the scale-free dynamics observed during Na-channel de-inactivation [39]. In this alternative view, scale-free dynamics is likely to emerge from the presence of multiple inactivated states of ionic-channels [19, 40].

4.3 How general is power-law spike-triggered adaptation?

All the *in-vitro* results presented in this paper are from mouse layer 5. We also investigated SFA in L2/3 and obtained very similar results (data not shown). In particular, we found that L2/3 pyramidal neurons adapt by means of power-law filters that closely resemble the ones observed in L5 and cause positive phase lead of the firing rate response to slow sinusoidal currents. These preliminary results suggest that L2/3 and L5 somatosensory pyramidal neurons share similar adaptation mechanisms[1]. We also fitted GLIF models to the data of Lundstrom et al. [21] and found that both L2/3 and L5 pyramidal neurons of the rat somatosensory cortex adapt by means of spike-triggered power-law processes (data not shown) indicating that this mechanism is conserved across species and could be a common feature of cortical pyramidal neurons.

---

[1]Long recordings from L2/3 were often unstable and we could not collect a complete dataset.

4.4 Functional implications

Both the moving threshold and the spike-triggered current are characterized by power-law decays with very similar scaling exponents. This suggests that the particular shape of the adaptation filters is not due to chance, but plays an important role. Neural signaling consumes a large amount of metabolic energy [1, 2]. The brain should therefore represent information using codes in which redundant information is discarded. According to efficient coding theory, optimality is achieved by adapting to the stimulus statistics and, at high signal-to-noise ratio (SNR), by completely removing correlations that are potentially present in the signals to be encoded [3]. Efficient coding theory has been used to explain neural processing at early stages of the visual system. In the retina, center-surround receptive fields coupled with nonlinear processing strongly attenuate spatial correlations of natural images [4, 7]. Similarly, in primary visual cortex (V1), spatial decorrelation of features has been found [15]. In the temporal domain, neural firing was found to be decorrelated in the lateral geniculate nucleus of the cat [6]. Interestingly, pyramidal neurons of V1 adapt on multiple timescales, providing further temporal decorrelation [29]. However, it remained unclear whether SFA serves for temporal redundancy reduction in the cortex.

To solve this issue, we estimated the statistics of input currents received *in-vivo* at the soma of L2/3 somatosensory pyramidal neurons during active whisker sensation (data from [31]). This current, which reflects spatio-temporal statistics of spike arrivals at the synapses as well as subsequent filtering in the dendritic tree, can be seen as the driving current for spike generation. We found that input currents of pyramidal neurons do not have a preferred timescale but are characterized by scale-free dynamics. Moreover, our numerical simulations indicate that power-law spike-triggered processes are near-optimally tuned to completely remove the temporal correlations revealed by the power-law decay of the input power spectrum (Fig. 5). Overall, these results provide evidence for efficient coding in single neocortical neurons stimulated with behaviorally relevant signals.

The GLIF-ξ model implements a form of predictive coding. Indeed, the sum of adaptation processes $\xi(s)$ triggered by past spikes can be interpreted as a linear predictor of the future input.

Consistent with predictive coding, further spiking only occurs when the real input exceeds the prediction. In line with our results, it has been shown that predictive coding of scale-free inputs by means of power-law spike-triggered kernels largely reduces the number of action potentials required to achieve a certain signal-to-noise ratio [41].

4.5 *Temporal Whitening* vs. *Noise-Shaping*

For deterministic signals encoded in absence of noise, efficient coding theory states that redundancy reduction is the optimal solution. However, in presence of noise, complete decorrelation can be detrimental. Redundancy can indeed improve the robustness of a code [30]. To assess optimal coding in small populations of neurons, the noise term $n(t)$ associated with stochastic firing (c.f. Eq. 3) has therefore to be considered.

Previous studies have shown that non-renewal firing activity with negatively correlated interspike intervals can achieve higher information rates by *noise-shaping* [42–44]. In this coding strategy, the SNR is increased in the frequency band of the input signal by transferring the effective noise power to other frequencies (see Supplementary Material). As already hypothesized in [45], our results show that at low-frequencies, spike-triggered adaptation results in a reduction of noise which is completely counterbalanced by a similar modification of the transfer function that controls the amplitude of the signal, so that the SNR remains unchanged. Consequently, modifying the adaptation filter $\xi(s)$ does not affect the power spectrum of the effective noise (see Supplementary Fig. S8), indicating that *noise-shaping* is probably not the functional role of power-law adaptation. The question of how this result generalizes to different stimulation paradigms is out of the scope of this study and will be addressed in a future paper.

Finally, since our model produces *temporal whitening* but not *noise-shaping*, it is tempting to speculate that to reduce noise, a relatively large number of cortical neurons simultaneously encode the same signal, enabling fast processing [46]. Indeed, while temporal averaging requires time, spatial averaging is instantaneous.

4.6 Conclusion

In computational studies of memory and learning in neural networks, SFA is often neglected and, when considered, it is usually assumed to operate on short timescales. From our perspective, the power-law of spike-triggered adaptation could be helpful in bridging the gap between the millisecond timescale of spike timing and behavioral timescales. Moreover, our results suggest that power-law adaptation causes temporal decorrelation of output spikes, a procedure that, at high signal-to-noise ratio, improves information transfer.

5   Acknowledgments

We thank C. C. H. Petersen, B. N. Lundstrom, G. Hennequin and A. Seeholzer for helpful discussions. We are also grateful to S. Crochet for sharing his *in-vivo* recordings and to B. N. Lundstrom for sharing the data that inspired this work. Finally, a particular thanks goes to S. Naskar for his precious help with the *in-vitro* recordings.

6 Author contributions

C.P. and R.N. conceived the study. C.P. designed the experiments, analyzed the data, performed the modeling and wrote the initial draft. S.M. contributed to data analysis and modeling. W.G. supervised the project. All the authors worked on the manuscript.

Figure 1: Experimental protocol and spiking neuron model GLIF-ξ.

To reveal SFA on multiple timescales, synaptically isolated L5 pyramidal neurons (*PYR Neuron*) were repeatedly stimulated with fluctuating currents (*Input Current*) generated by adding filtered Gaussian noise to sinusoidal waves with different angular frequencies $w = 2\pi/T$ (*Mean Modulation*). The horizontal bars (bottom left and right) indicates the period T of modulation. The single neuron response (*Spiking response*, black) was recorded intracellularly and the firing rate $r(t)$ was estimated by counting the number of spikes in every time bin (*Firing Rate Modulation*, gray). The periodic oscillations of the firing rate $r_{\mathrm{mean}}(t)$ (*Firing Rate Modulation*, black) is related to the mean input (*Firing Rate Modulation*, light gray) by a linear rate model defined in the Fourier domain by the the transfer function $\hat{H}(w)$. The intracellular recordings were used to fit the Generalized Leaky Integrate & Fire model GLIF-ξ (black-lined box, top). In this model, the input current is first low-pass filtered by the membrane filter $K_m(t)$ and then transformed into a firing intensity by an exponential nonlinearity. Spikes are emitted stochastically (*Spiking response*, red) and trigger an adaptation process described by the effective adaptation kernel $\xi(s)$.

Figure 2: Adaptation filter of the GLIF-ξ model extracted from *in-vitro* recordings.

a: To obtain the effective adaptation filter $\xi(s)$ of the GLIF-ξ model, intracellular recordings were first fitted with a two process GLIF model which accounts for SFA with both a spike-triggered current $\eta(s)$ and spike-triggered movement of the firing threshold $\gamma(s)$. **Left**: The mean spike-triggered current $\eta(s)$ obtained by averaging the results of different cells ($n = 14$) is shown in red on a log-log scale. The dashed black line shows the fit of a power-law function $\eta_{\mathrm{PL}}(s) = \alpha_\eta s^{-\beta_\eta}$ with parameters $\alpha_\eta = 0.44$ nA, $\beta_\eta = 0.76$ and s in milliseconds. **Right**: The mean moving threshold $\gamma(s)$ obtained by averaging the results of different cells ($n = 14$) is shown in red on a log-log scale. The dashed black line shows the fit of a power-law function $\gamma_{\mathrm{PL}}(s) = \alpha_\gamma s^{-\beta_\gamma}$ with parameters $\alpha_\gamma = 24.4$ mV, $\beta_\gamma = 0.87$ and s in milliseconds. The dark gray line is a control showing an independent estimation of the average moving threshold $\gamma(t)$ obtained with an alternative fitting procedure (see Supplementary Material). **b:** As illustrated by the block diagram, the spike-triggered current $\eta(s)$ and the moving threshold $\xi(s)$ are combined to obtain the effective adaptation filter $\xi(s)$ of the GLIF-ξ model (see Methods). The mean adaptation filter $\xi_{\mathrm{L}}(s)$ (red, GLIF-ξ_{L}) obtained by averaging the effective spike-triggered adaptation measured in individual cells ($n = 14$, see Supplementary Fig. S7) is shown on a log-log scale. The optimal fit of a truncated power-law $\xi_{\mathrm{PL}}(s)$ (dashed black, GLIF-ξ_{PL}) yields an exponent $\beta_\xi = 0.93$ (c.f. Eq. 2). In all panels, the gray area indicates one standard deviation for the distribution of filters across different cells (asymmetric errors are due to log-scales).

Figure 3: The GLIF-ξ model predicts the occurrence of single spikes with a millisecond precision.

a: Typical 2.5-second segment of injected current. The same fluctuating current is presented several times (*frozen-noise*). The dashed black line represents 0 nA. **b:** The spiking response, but not the subthreshold membrane potential, predicted by the GLIF-ξ_L model (red) is in close agreement with the experimental data (black). In the two process-model (gray), where spike-triggered currents and threshold movements are modeled by two distinct processes (i.e. $\gamma(s)$ and $\eta(s)$), the dynamics of the subthreshold membrane potential predicted by the model is in excellent agreement with the experimental data. Inset: comparison of subthreshold membrane potential (scale bars: 40 ms, 5 mV). **c:** The raster plots show the spiking response of both the neuron (black) and the GLIF-ξ_L model (red) to repetitive presentation of the same current. By construction, the spiking response of the GLIF-ξ_L model is identical to that of the two-process model (gray).

Figure 4: The GLIF-ξ model accurately predicts the firing rate response on multiple timescales.

a: An input current consisting of a *frozen-noise* signal on top of a small sinusoidal wave of period $T = 16$ s (dark gray) is presented several times to the same L5 pyramidal neuron. **b:** Typical membrane potential response in a single trial. **c:** The firing activity obtained by repetitive presentation of the same input signal (black) is compared with the predictions made by GLIF-ξ_L (red) and GLIF-ξ_S (green). **d:** Data shown in panel c were used to build two continuous PSTHs (spikes were filtered with a moving average over 1 second). Experimental data (black) and GLIF-ξ_L model prediction (red) are in good agreement. The PSTHs were fitted with two sinusoidal functions to quantify the phase shift $\hat{H}_\Phi$ (Eq. 1) with respect to the modulation of the mean input (dark gray sinusoidal wave). For clarity, only the fit to experimental data is shown (light gray sinusoidal wave, $\hat{H}_\Phi = 17.6$ deg). The phase shift is measured with respect to the periodic input (dark gray line). **e:** The prediction made by the GLIF-ξ_S model is shown in green. Compared to panel d, the prediction is less accurate. In particular, the PSTH produced by GLIF-ξ_S is approximatively in phase with the input modulation (dark gray) but not with the PSTH of the experimental data, meaning that the model is unable to correctly capture the correct phase advance of $\hat{H}_\Phi = 17.6$ deg. Experimental PSTH (black), sinusoidal fit of the experimental data (light gray) and input modulation (dark gray) are copied from panel d. **f:** The performances of GLIF-ξ_L and GLIF-ξ_S in predicting the average firing rate r_0 of a new stimuli (not used for data fitting) are compared. Left: for the two models, the predicted r_0 is plotted against the observed one. Each dot represents a different cell. Right: the error in predicting r_0 is plotted for the two models. Each couple of open circles shows the performances of the two models on one cell. On average, the errors produced by GLIF-ξ_L (red) and GLIF-ξ_S (green) are not significantly different ($n = 12$, Student t-test, $t = -1.62$, $p = 0.12$; n.s. $p > 0.05$). Error bars indicate one standard deviation. **g:** The two models are significantly different if compared by measuring the ability to predict the phase shift $\hat{H}_\Phi$ ($n = 12$, Student t-test, $t = 6.77$, $p = 0.8 \cdot 10^{-6}$; triple star: $p < 0.001$). Conventions are as in panel f. **h-j:** The linear response of L5 pyramidal neurons to a slow modulation of the mean input was characterized by the transfer function $\hat{H}(w)$. The gain $\hat{H}_A(w)$ (**h**) and the phase response $\hat{H}_\Phi(w)$ (**j**) are plotted as a function of the period of modulation

$T = 2\pi/w$. **i:** log-log plot of the gain $\hat{H}_{\mathrm{A}}$ as a function of the input frequency $f = T^{-1}$. The amplitude response observed *in-vitro* was fitted by power law with scaling exponent $\beta_{\mathrm{H}} = 0.12$ (dashed gray). In panels *h-j*, data from individual cells ($n = 14$, gray lines) are averaged (black) and compared with the predictions of GLIF-ξ_{L} (red) and GLIF-ξ_{S} (green). Error bars indicate one standard deviation. The horizontal dashed lines appearing in panels *a*, *d*, *e*, *h* and *j* indicate zero.

Figure 5: Power-law adaptation is near-optimally tuned to perform *temporal whitening*.

a: Power spectral density of the intracellular membrane potential fluctuations recorded *in-vivo* from L2/3 pyramidal neurons (Voltage PSD, red). The power spectrum was computed using 20-second long recordings ($n = 57$) obtained from 7 different cells (data from [31]). Fitting a power law (not shown) on the frequency band $0.05 < f < 2$ Hz yields a scaling exponent $\beta_{\mathrm{I}} = 0.67$. A population of GLIF-$\xi_{\mathrm{PL}}$ neurons ($N = 100$) was simulated in response to a scale-free input characterized by a power-law frequency spectrum $\Delta I(f) \propto f^{-\beta_{\mathrm{I}}}$, with $\beta_{\mathrm{I}} = 0.67$ (Input Current PSD, black). The subthreshold response of individual GLIF-ξ_{PL} neurons is characterized by a frequency spectrum (Voltage PSD, gray) which is in good agreement with the *in-vivo* recordings. **b:** The population activity of a group of GLIF-ξ_{PL} neurons (blue) in response to an *in-vivo* like input (black, copied from panel *a*) has a nearly flat spectrum $A(f)$. Very similar results were obtained by simulating a population of GLIF-ξ_{L} ($N = 100$) neurons (gray). To allow a direct comparison between input and output powers, all the spectra shown in panel *b* were normalized to have the same total power. All the simulations were performed with the average parameters extracted from *in-vitro* recordings (Table 1). The mean activity was tuned to $A_0 = 4$ Hz by adjusting the baseline current to I_0. However, the precise value of A_0 does not have a big impact on the result.

7 Methods

7.1 *In-vitro* electrophysiological recordings

All animal experiments were done using published procedures [47, 48] in accordance with the rules of the Swiss Federal Veterinary Office. Briefly, somatosensory brain slices were obtained from P14-18 Wild Type mice (C57BL6/J) and whole-cell patch-clamp recordings were performed at 35°C from L5 pyramidal neurons. The pipette solution was comprised of (in mM): 135 K-gluconate, 4 KCl, 4 Mg-ATP, 10 Na2-phosphocreatine, 0.3 Na3-GTP and 10 HEPES (pH 7.3, 290 mOsm). During the experiments, transmission at all excitatory synapses was blocked by adding CNQX (20 μM) and D-AP5 (50 μM) to the bath solution. All electrophysiological data were low-pass Bessel filtered at 10 kHz and digitized at 20 kHz. Membrane potential measurements were not corrected for the liquid junction potential. Recordings characterized by instabilities in the action potential shape and/or large drifts in the baseline firing rate r_0 were excluded from the dataset upon visual inspection.

7.2 Current Injections

To characterize single neurons with the standard tools of linear system analysis, 64-s long experiments were performed in which noisy currents modulated by sinusoidal means were delivered in current-clamp mode. Similar protocols have already been used in previous studies (see, for example, refs. [38] and [21]). The injected current, denoted I_{ext}, was generated according to the following equation

$$I_{\text{ext}}(t) = I_0 + \Delta I_{\text{mean}} \cdot \sin\left(\frac{2\pi}{T}t\right) + \Delta I_{\text{noise}} \cdot \mathcal{N}(t) \tag{4}$$

were I_0 is a constant offset, ΔI_{mean} controls the amplitude of the sinusoidal mean and ΔI_{noise} defines the standard deviation of the noise. The noise $\mathcal{N}(t)$ was generated with an Ornstein-Uhlenbeck process with zero mean, unitary variance and a temporal correlation of 3 ms.

Each experiment consisted of many injections of currents generated according to Equation 4. In the first half of the experiment (*training set*), the currents $I_{\text{ext}}(t)$ were presented six times with different periods of modulation $T \in \{0.5, 1, 2, 4, 8, 16\}$ in seconds. Stimuli were delivered

in random order and, for each of the six injections, a new realization of the noise $\mathcal{N}(t)$ was used. The results of these experiments were used to both fit the GLIF-ξ models and estimate the transfer function $\hat{H}(w)$ (Fig. 4h-j). In the second part of the experiment (*test set*), one of the two slowest modulations ($T = 8$ or 16 s) was chosen and more injections were performed. To assess the reliability of single neurons, the same realization of noise $\mathcal{N}(t)$ was used (*frozen-noise*), meaning that neurons were repetitively stimulated with exactly the same current. Thus, in the test set, $\mathcal{N}(t)$ plays the role of a known and rapidly changing temporal signal (on the timescale of 3 ms), rather than that of an unknown noise; at the same time the periodic modulation plays the role of a slow temporal signal (T on the timescale of 8-16 s). The results of these experiments were used to assess the predictive power of GLIF-ξ models (see Fig. 3 and Fig. 4a-g). All injections were performed with interstimulus intervals of 1 minute.

To monitor the stability of the recordings, before and after each injection, the neuron was stimulated with two additional inputs. The first was a 2.5-s long current composed of a hyperpolarizing step followed, after 500 ms, by a suprathreshold step. The response to this stimulus was used to identify the neuronal type (L5 burst-generating cells were not included in the dataset). The second was a 4-s long noisy current generated with an Ornstein-Uhlenbeck process with zero mean and temporal correlation of 3 ms. The amplitude of this current was scaled to avoid action potentials. We used this second injection to estimate the electrode filter $K_e(s)$ and perform Active Electrode Compensation (see Supplementary Material).

At the beginning of each experiment, the parameters I_0, ΔI_{mean} and $\Delta I_{\mathrm{noise}}$ were tuned to obtain a firing rate r_{mean} oscillating periodically between 2 and 6 Hz. Typical values obtained after calibration were comprised in the range 100-450 pA for I_0, 15-30 pA for ΔI_{mean} and 50-150 pA for $\Delta I_{\mathrm{noise}}$.

7.3   Linear analysis

For each neuron, the transfer function $\hat{H}(w)$ shown in Figure 4h-j was estimated with standard methods already used in previous studies [21, 38]. Briefly, the experimental spike train $\{\hat{t}_j\}$ was

built by selecting the times at which the membrane potential $V(t)$ crossed 0 mV from below. The firing rate $r(t)$ was then obtained by building a histogram of the spike times. The bin size was chosen such that each period of modulation T was divided in 30 bins. For each input frequency $w = 2\pi/T$, the transfer function $\hat{H}(w)$ was estimated by fitting the sinusoidal function $C_0 + C_1 \cdot \sin(wt + \phi)$ on the firing rate $r(t)$. Each fit was performed by minimizing the sum of squared errors (SSE), with C_0, C_1 and ϕ being the only free-parameters. The amplitude response $\hat{H}_{\mathrm{A}}(w)$ was obtained by dividing C_1 by ΔI_{mean}; c.f. Equation 1. The optimal ϕ extracted at the stimulation frequency w gave the phase response $\hat{H}_\Phi(w)$. The transfer functions of the GLIF-ξ models (Fig. 4h-j) were obtained with the same method, except that spikes were generated by the model, rather than extracted from experiments.

7.4 Generalized Leaky Integrate-and-Fire model (GLIF-ξ)

The spiking neuron models discussed in this study are generalized leaky integrate-and-fire models equipped with a spike-triggered mechanism for SFA and with escape rate noise for stochastic spike emission (Fig. 1). Spikes are produced according to a point process with conditional firing intensity $\lambda(t)$ which exponentially depends on the momentary distance between the membrane potential $V(t)$ and the *effective* firing threshold $V_T(t)$ [49]:

$$\lambda(t) = \lambda_0 \exp\left(\frac{V(t) - V_T(t)}{\Delta V}\right) \tag{5}$$

where λ_0 has units of s^{-1} so that $\lambda(t)$ is in Hz and ΔV defines the sharpness of the threshold. Consequently, the probability of a spike to occur at a time $\hat{t} \in [t; t + \Delta t]$ is

$$P(\hat{t} \in [t; t + \Delta t]) = 1 - \exp\left(-\int_t^{t+\Delta t} \lambda(s)\, ds\right) \approx \lambda(t)\Delta t. \tag{6}$$

In the limit of $\Delta V \to 0$, the model becomes deterministic and action potentials are emitted at the moment when the membrane potential crosses the firing threshold. For finite ΔV and a membrane potential at threshold (i.e. when $V = V_T$), λ_0^{-1} defines the mean latency until a spike is emitted.

The subthreshold dynamics is modeled as a standard leaky integrator defined by the following ordinary differential equation for the membrane potential V

$$C\dot{V} = -g_L\left(V - E_L\right) + I_{\text{ext}} \tag{7}$$

where the three parameters C, g_L and E_L determine the passive properties of the membrane, the *dot* denotes the temporal derivative and I_{ext} is the injected current.

The dynamics of the *effective* firing threshold $V_T(t)$ in Equation 5 is given by

$$V_T(t) = V_T^* + \sum_{\hat{t}_j < t} \xi(t - \hat{t}_j - T_{\text{ref}}) \tag{8}$$

where V_T^* is a constant, $\{\hat{t}_1, \hat{t}_2, \hat{t}_3, \dots\}$ are the times at which action potentials have been fired and $\xi(s)$ is an effective adaptation filter that accounts for all the biophysical events triggered by the emission of an action potential. According to Eq. 8, each time a spike is emitted, a threshold movement with stereotypical shape $\xi(s)$ is triggered, after a delay of absolute refractoriness T_{ref}. Threshold movements induced by different spikes accumulate and therefore produce SFA, if $\xi > 0$. For $s < 0$, we fixed $\xi(s) = 0$ so that only spikes in the past can affect the momentary value of the firing threshold (i.e. $\xi(s)$ is *causal*). Importantly, the adaptation filter $\xi(s)$ also accounts for adaptation processes mediated by spike-triggered currents. Consequently, $V_T(t)$ does not describe the *physiological* threshold (i.e. the membrane potential at which action potentials are initiated *in-vitro*) but has to be interpreted as a *phenomenological* model of spike-triggered adaptation. Finally, the functional shape of $\xi(s)$ is not defined *a priori* but is obtained by combining the effects of both spike-triggered currents and spike-triggered movement of the *physiological* threshold which, in turn, are extracted from the experimental data (see Fitting GLIF-ξ on *in-vitro* recordings).

In principle, an absolute refractory period can be included in the adaptation kernel $\xi(s)$. However, here we prefer to work with an explicit reset after a dead time. Each time a spike is emitted the numerical integration is stopped and the membrane potential is reset to V_r. Then,

after a short period of absolute refractoriness T_{ref}, the numerical integration is restarted. The GLIF-ξ model only differs from a Generalized Linear Model [50] due to this explicit reset.

The three GLIF-ξ models discussed in the paper differ in the duration and shape of the adaptation filter $\xi(s)$. In GLIF-ξ_{L} and GLIF-ξ_{S}, the functional shape of $\xi(s)$ is the one directly extracted from intracellular recordings (see below). In these two models the duration of the adaptation filter is of 22 s and 1 s, respectively. In the third model, GLIF-ξ_{PL}, the adaptation filer $\xi(s)$ is modeled as a truncated power law and lasts for 22 s.

7.5 Fitting the GLIF-ξ model on *in-vitro* recordings

The recorded potential was preprocessed to remove the bias due to the voltage drop across the recording electrode. For that, Active Electrode Compensation (AEC) [51, 52] was performed in such a way as to remove experimental drifts due to slow changes in the electrode properties (see Supplementary Material and Supplementary Fig. S1-S3). To fit GLIF-ξ models, the method introduced in [33] was extended to get a more accurate estimate of $\xi(s)$. This was done with an additional hidden variable I_{drift} able to absorb small drifts that are likely to occur in long recordings (see Supplementary Material).

In neocortical pyramidal neurons, SFA is mediated by both adaptation currents and changes in the firing threshold [33]. To get an accurate estimation of the effective adaptation filter $\xi(s)$, *in-vitro* recordings were first fitted with a two-process GLIF model that explicitly features both a spike-triggered current $\eta(s)$ and a spike-triggered movement of the firing threshold $\gamma(s)$ (see Fig. 2 and Supplementary Fig. S4). The effective adaptation filer $\xi(s)$ of the GLIF-ξ_{L} model was then obtained by combining $\eta(s)$ and $\gamma(s)$ according to the following formula

$$\xi(t) = \int_0^\infty K_m(t-s)\eta(s)ds + \gamma(t), \tag{9}$$

where $K_m(s) = \Theta(s)\frac{R}{\tau_m}e^{-\frac{s}{\tau_m}}$ is the membrane filter and $\Theta(s)$ is the Heaviside step function (see Fig. 2). Importantly, the functional shapes of $\eta(s)$ and $\gamma(s)$ were not assumed *a priori* but were

directly extracted from the experimental data by the following two-step procedure.

In the first step, the functional shape of $\eta(s)$, together with all the parameters that determines the subthreshold dynamics, were extracted by fitting $\dot{V}$ on the experimental voltage derivative $\dot{V}^{(\mathrm{data})}$. Since adaptation currents directly affect the membrane potential dynamics, $\dot{V}^{(\mathrm{data})}$ was fitted with the following model

$$C\dot{V} = -g_L\left(V - E_L\right) + I_{\mathrm{ext}} - \sum_{\hat{t}_j < t} \eta(t - \hat{t}_j - T_{\mathrm{ref}}) + I_{\mathrm{drift}}(t) \tag{10}$$

where Eq. 7 is extended with a spike-triggered current $\eta(s)$ (Fig. 2a) and the additional term I_{drift} captures experimental drifts. To avoid any *a priori* assumption on the functional shape of the spike-triggered current, $\eta(s)$ was defined as linear combination of rectangular basis functions. Consequently, optimal parameters (minimizing the SSE between $\dot{V}$ and $\dot{V}^{(\mathrm{data})}$) could be obtained by solving a multilinear regression problem [53].

Given the subthreshold dynamics, the second step consists in estimating the parameters of the effective threshold. Since spike-triggered currents are properly modeled by the filter $\eta(s)$, the effective threshold of Equation 8 is replaced by

$$V_T^{(\mathrm{bio})}(t) = V_T^* + \sum_{\hat{t}_j < t} \gamma(t - \hat{t}_j - T_{\mathrm{ref}}), \tag{11}$$

where $V_T^{(\mathrm{bio})}(t)$ describes the *physiological* threshold (Fig. 2a) at which action potentials are initiated *in-vitro*. In contrast to $\xi(s)$, $\gamma(s)$ is not a phenomenological model but describes physiological changes of the firing threshold triggered by the emission of previous spikes. Similar to $\eta(s)$, the moving threshold $\gamma(s)$ was defined as a linear combination of rectangular basis function and its functional shape was extracted from experimental data by maximizing the *log-likelihood* of the observed spike-train. With the exponential function in Equation 5, the *log-likelihood* to maximize is a concave function of the parameters [54]. Consequently, the fit could be performed using standard gradient ascent methods.

A detailed description of the fitting procedure can be found in the Supplementary Material.

7.5.1 Power-law fit of the effective adaptation filter $\xi(s)$

For GLIF-ξ_{PL}, the effective adaptation filter $\xi_{\mathrm{L}}(s)$ extracted from the intracellular recordings was fitted with a truncated power law $\xi_{\mathrm{PL}}(s)$ (Eq. 2). The fit was performed in two steps. In the first step, the magnitude α_ξ and the scaling exponent β_ξ were estimated using a least-square linear regression performed in log-log space. For that, data points were logarithmically resampled and excluded from the fit if $\xi_{\mathrm{L}}(s) < 5 \cdot 10^{-3}$ mV or $s < 5$ ms. In a second step, the cutoff T_ξ was obtained by calculating the intercept between the power law fitted in the first step and the average value of the extracted kernel $\xi_{\mathrm{L}}(s)$ computed on the first 5 ms.

A similar procedure (i.e. least-square linear regression in log-log space with logarithmically resampled points) was used for the power law fit of the spike-triggered current $\eta(s)$ and the spike-triggered movement of the firing threshold $\gamma(s)$ shown in Fig. 2.

7.6 Performance evaluation

7.6.1 Cross-validation

All the performances reported in this study were evaluated on datasets that have not been used for parameter extraction. This strategy avoids problems due to *overfitting* and enables a direct comparison of models that do not have the same number of parameters (which is true for GLIF-ξ_{S}, GLIF-ξ_{L} and GLIF-ξ_{PL}). For the predictions reported in Figure 3 and Figure 4a-g, the model fitted on the first half of the experiment (*training set*) was used to predict the responses observed in the second half (*test set*). Since in certain experiments the average firing rates r_0 observed in the *test set* were slightly different than the ones of the *training set*, the parameter V_T^* was readjusted using the firsts 16 s of all the *test set* injections and models were validated on the responses recorded in the remaining 48 s. According to this procedure, models that do not capture SFA on slow timescales are expected to overestimate the average firing rate r_0. Finally, for the predictions reported in Figure 4h-j, a *leave-one-out strategy* was used: the models fitted on the responses to 5 different periods of modulation were used to predict the sixth one.

7.6.2 Spike-train metrics

To assess the ability to predict spike times, we used the similarity measure M_d^* [35]. M_d^* quantifies the similarity between two groups of spike trains generated by two stochastic processes and corrects the bias caused by the small number of available repetitions. M_d^* takes values between 0 and 1, where $M_d^* = 0$ indicates that the model is unable to predict any of the observed spikes and $M_d^* = 1$ means that the two groups of spike trains have the same instantaneous firing rate and are statistically indistinguishable. M_d^* can also be interpreted as the number of spikes correctly predicted divided by an estimate of the number of reliable spikes. To compute M_d^* we proceeded as follow. 500 spike trains were generated using our optimal model. Then, the average number of coincident spikes n_{NM} between the model M and the real neuron N was computed by counting, for each possible pair of spike trains, the number of action potentials predicted by the model that fell within ± 4 ms of a recorded spike. M_d^* was finally obtained by applying the formula:

$$M_d^* = \frac{2 n_{\mathrm{NM}}}{n_{\mathrm{MM}} + n_{\mathrm{NN}}} \qquad (12)$$

where n_{MM} and n_{NN} are the average number of coincident spikes (with a precision of ± 4 ms) between distinct repetitions generated by the model (n_{MM}); and repetitions recorded from a real neuron (n_{NN}).

7.7 Estimating the statistical properties of the input current received *in-vivo* by neocortical pyramidal neurons

To test the hypothesis that power-law adaptation contributes to efficient coding by *whitening* the single neuron output, the power spectrum of the currents $\Delta I(f)$ received as input at the soma of a neocortical pyramidal neuron *in-vivo* was characterized. According to Equation 10, in absence of spikes, the membrane potential $\Delta V(t)$ is a low-pass filtered version of the input current, where the cutoff frequency $f_c = \tau_m^{-1}$ is defined by the membrane timescale. Consequently, at all frequencies $f \ll f_c$, we have that $\Delta I(f) = \Delta V(f)/R^2$, with $\Delta V(f)$ being the power spectrum of the subthreshold membrane potential fluctuations and R the input resistance.

We estimated $\Delta V(f)$ using 20-second long whole-cell recordings ($n = 57$) of the synaptically driven membrane potential dynamics obtained from 7 different L2/3 pyramidal neurons of behaving mice (data from [31]). All the *in-vivo* recordings are from the primary somatosensory barrel cortex and were performed during active whisker sensation. Further details on the experimental protocol can be found in ref. [31]. Spike-triggered currents last for more than 20 seconds and can in principle affect $\Delta V(f)$ even at very low frequencies. For this reason, only trials with low firing rates $r_0 < 0.5$ Hz were used. However, including recordings with $r_0 > 0.5$ Hz did not affect the results.

7.8 Simulating the population response to *in-vivo* like inputs

The results reported in Figure 5 were obtained by simulating a population of $N = 100$ unconnected GLIF-ξ_{PL} neurons in response to 4000-second long currents $I(t)$ characterized by a power spectrum $\Delta I(f) \propto f^{-\beta_{\mathrm{I}}}$, with $\beta_{\mathrm{I}} = 0.67$. Model parameters are given in Table 1 and input currents were generated as in ref. [29] by numerically solving the following inverse Fourier transform

$$I(t) = I_0 + \mathcal{C} \cdot \int_{-\infty}^{+\infty} \sqrt{\Delta I}(f)\mathcal{N}(f)e^{i(2\pi ft + \phi(f))}df, \tag{13}$$

where $\mathcal{N}(f)$ is a Gaussian white-noise process, the phases $\phi(f)$ are independently sampled from a uniform distribution and $\mathcal{C}$ is a scaling factor adjusted to fit the power spectrum of the sub-threshold membrane potential fluctuations observed *in-vivo* (see Fig. 5a). To avoid unrealistic large power at slow frequencies, we introduced a cutoff $\Delta I(f) = 0$, for $f < 0.025$ Hz. The highest frequency in the signal was determined by the time step $\Delta T = 0.5$ ms used for numerical simulations. The mean input I_0 was adjusted to obtain a plausible average activity of $A_0 = 4$ Hz, which was consistent with the firing rates obtained *in-vitro*.

The population activity $A(t)$ was constructed by counting the number of spikes falling in bins of 50 ms and its power spectrum $A(f)$ was finally computed using time series of 40 s.

References

[1] Attwell, D. & Laughlin, S. B. An energy budget for signaling in the grey matter of the brain. *J. Cereb. Blood Flow Metab.* **21**, 1133–1145 (2001).

[2] Laughlin, S. B. Energy as a constraint on the coding and processing of sensory information. *Curr. Opin. Neurobiol.* **11**, 475–480 (2001).

[3] Barlow, H. Possible principles underlying the transformation of sensory messages. *Sensory communication* 217–234 (1961).

[4] Srinivasan, M. V., Laughlin, S. B. & Dubs, A. Predictive coding: a fresh view of inhibition in the retina. *Proc. R. Soc. Lond., B, Biol. Sci.* **216**, 427–459 (1982).

[5] Dong, D. & Atick, J. Temporal decorrelation: a theory of lagged and nonlagged responses in the lateral geniculate nucleus. *Network* **6**, 159–178 (1995).

[6] Dan, Y., Atick, J. & Reid, R. C. Efficient coding of natural scenes in the lateral geniculate nucleus: experimental test of a computational theory. *J. Neurosci.* **16**, 3351–3362 (1996).

[7] Pitkow, X. & Meister, M. Decorrelation and efficient coding by retinal ganglion cells. *Nat. Neurosci.* **15**, 628–635 (2012).

[8] Wark, B., Lundstrom, B. N. & Fairhall, A. Sensory adaptation. *Curr. Opin. Neurobiol.* **17**, 423–429 (2007).

[9] Wainwright, M. J. Visual adaptation as optimal information transmission. *Vision Res.* **39**, 3960–3974 (1999).

[10] Brenner, N., Bialek, W. & de Ruyter van Steveninck, R. Adaptive rescaling maximizes information transmission. *Neuron* **26**, 695–702 (2000).

[11] Fairhall, A. L., Lewen, G. D., Bialek, W. & de Ruyter Van Steveninck, R. R. Efficiency and ambiguity in an adaptive neural code. *Nature* **412**, 787–792 (2001).

[12] Maravall, M., Petersen, R. S., Fairhall, A. L., Arabzadeh, E. & Diamond, M. E. Shifts in Coding Properties and Maintenance of Information Transmission during Adaptation in Barrel Cortex. *PLoS Biol.* **5**, e19 (2007).

[13] Baccus, S. A. & Meister, M. Fast and slow contrast adaptation in retinal circuitry. *Neuron* **36**, 909–919 (2002).

[14] Ulanovsky, N., Las, L., Farkas, D. & Nelken, I. Multiple time scales of adaptation in auditory cortex neurons. *J. Neurosci.* **24**, 10440–10453 (2004).

[15] Simoncelli, E. & Olshausen, B. Natural image statistics and neural representation. *Annu. Rev. Neurosci.* **24**, 1193–1216 (2001).

[16] Izhikevich, E. M. Simple model of spiking neurons. *IEEE Trans. Neural Netw.* **14**, 1569–1572 (2003).

[17] Brette, R. & Gerstner, W. Adaptive exponential integrate-and-fire model as an effective description of neuronal activity. *J. Neurophysiol.* **94**, 3637–3642 (2005).

[18] Spain, W. & Schwindt, P. Two transient potassium currents in layer V pyramidal neurones from cat sensorimotor cortex. *J. Physiol.* (1991).

[19] Gilboa, G., Chen, R. & Brenner, N. History-dependent multiple-time-scale dynamics in a single-neuron model. *J. Neurosci.* **25**, 6479–6489 (2005).

[20] La Camera, G. *et al.* Multiple time scales of temporal response in pyramidal and fast spiking cortical neurons. *J. Neurophysiol.* **96**, 3448–3464 (2006).

[21] Lundstrom, B. N., Higgs, M. H., Spain, W. J. & Fairhall, A. L. Fractional differentiation by neocortical pyramidal neurons. *Nat. Neurosci.* **11**, 1335–1342 (2008).

[22] Drew, P. J. & Abbott, L. F. Models and properties of power-law adaptation in neural systems. *J. Neurophysiol.* **96**, 826–833 (2006).

[23] Fleidervish, I. A., Friedman, A. & Gutnick, M. J. Slow inactivation of Na+ current and slow cumulative spike adaptation in mouse and guinea-pig neocortical neurones in slices. *J. Physiol.* **493 (Pt 1)**, 83–97 (1996).

[24] Mickus, T., Jung, H. y. & Spruston, N. Properties of slow, cumulative sodium channel inactivation in rat hippocampal CA1 pyramidal neurons. *Biophys. J.* **76**, 846–860 (1999).

[25] Melnick, I. V., Santos, S. F. A. & Safronov, B. V. Mechanism of spike frequency adaptation in substantia gelatinosa neurones of rat. *J. Physiol.* **559**, 383–395 (2004).

[26] Madison, D. V. & Nicoll, R. A. Control of the repetitive discharge of rat CA 1 pyramidal neurones in vitro. *J. Physiol.* **354**, 319–331 (1984).

[27] Schwindt, P. C., Spain, W. J. & Crill, W. E. Long-lasting reduction of excitability by a sodium-dependent potassium current in cat neocortical neurons. *J. Neurophysiol.* **61**, 233–244 (1989).

[28] Sanchez-Vives, M. V., Nowak, L. G. & McCormick, D. A. Cellular mechanisms of long-lasting adaptation in visual cortical neurons in vitro. *J. Neurosci.* **20**, 4286–4299 (2000).

[29] Wang, X.-J., Liu, Y., Sanchez-Vives, M. V. & McCormick, D. A. Adaptation and temporal decorrelation by single neurons in the primary visual cortex. *J. Neurophysiol.* **89**, 3279–3293 (2003).

[30] Rieke, F., Warland, D., de Ruyter van Steveninck, R. & Bialek, W. Spikes: Exploring the neural code. *Cambridge: MIT Press* (1999).

[31] Crochet, S., Poulet, J. F. A., Kremer, Y. & Petersen, C. C. H. Synaptic mechanisms underlying sparse coding of active touch. *Neuron* **69**, 1160–1175 (2011).

[32] Gerstner, W. & Kistler, W. Spiking Neuron Models: Single Neurons, Populations, Plasticity. *Cambridge University Press New York, NY, USA* (2002).

[33] Mensi, S. *et al.* Parameter Extraction and Classification of Three Cortical Neuron Types Reveals Two Distinct Adaptation Mechanisms. *J. Neurophysiol.* (2011).

[34] Jolivet, R. *et al.* The quantitative single-neuron modeling competition. *Biol. Cybern.* **99**, 417–426 (2008).

[35] Naud, R., Gerhard, F., Mensi, S. & Gerstner, W. Improved similarity measures for small sets of spike trains. *Neural. Comput.* **23**, 3016–3069 (2011).

[36] Atick, J. Could information theory provide an ecological theory of sensory processing? *Network* **3**, 213–251 (1992).

[37] Benda, J. & Herz, A. V. M. A universal model for spike-frequency adaptation. *Neural. Comput.* **15**, 2523–2564 (2003).

[38] Köndgen, H. *et al.* The dynamical response properties of neocortical neurons to temporally modulated noisy inputs in vitro. *Cereb. Cortex* **18**, 2086–2097 (2008).

[39] Toib, A., Lyakhov, V. & Marom, S. Interaction between duration of activity and time course of recovery from slow inactivation in mammalian brain Na+ channels. *J. Neurosci.* **18**, 1893–1903 (1998).

[40] Lowen, S. B., Liebovitch, L. S. & White, J. A. Fractal ion-channel behavior generates fractal firing patterns in neuronal models. *Phys. Rev. E Stat. Nonlin. Soft.* **59**, 5970–5980 (1999).

[41] Bohte, S. M. & Rombouts, J. O. Fractionally Predictive Spiking Neurons. In *NIPS*, 253–261 (2010).

[42] Mar, D. J., Chow, C. C., Gerstner, W., Adams, R. W. & Collins, J. J. Noise shaping in populations of coupled model neurons. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 10450–10455 (1999).

[43] Shin, J. Adaptation in spiking neurons based on the noise shaping neural coding hypothesis. *Neural Netw.* **14**, 907–919 (2001).

[44] Chacron, M., Lindner, B. & Longtin, A. Noise Shaping by Interval Correlations Increases Information Transfer. *Phys. Rev. Lett.* **92** (2004).

[45] Avila-Akerberg, O. & Chacron, M. J. Nonrenewal spike train statistics: causes and functional consequences on neural coding. *Exp. Brain Res.* **210**, 353–371 (2011).

[46] Gerstner, W. Population dynamics of spiking neurons: fast transients, asynchronous states, and locking. *Neural. Comput.* **12**, 43–89 (2000).

[47] Lefort, S., Tomm, C., Sarria, J. C. F. & Petersen, C. C. H. The Excitatory Neuronal Network of the C2 Barrel Column in Mouse Primary Somatosensory Cortex. *Neuron* **61**, 301–316 (2009).

[48] Avermann, M., Tomm, C., Mateo, C., Gerstner, W. & Petersen, C. Microcircuits of excitatory and inhibitory neurons in layer 2/3 of mouse barrel cortex. *J. Neurophysiol.* **107**, 3116–3134 (2012).

[49] Jolivet, R., Rauch, A., Lüscher, H. & Gerstner, W. Predicting spike timing of neocortical pyramidal neurons by simple threshold models. *J Comput Neurosci* **21**, 35–49 (2006).

[50] Pillow, J. W. *et al.* Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature* **454**, 995–999 (2008).

[51] Brette, R. *et al.* High-resolution intracellular recordings using a real-time computational model of the electrode. *Neuron* **59**, 379–391 (2008).

[52] Badel, L. *et al.* Dynamic I-V curves are reliable predictors of naturalistic pyramidal-neuron voltage traces. *J. Neurophysiol.* **99**, 656–666 (2008).

[53] Paninski, L., Pillow, J. & Simoncelli, E. Comparing integrate-and-fire models estimated using intracellular and extracellular data. *Neurocomputing* **65**, 379–385 (2005).

[54] Paninski, L. Maximum likelihood estimation of cascade point-process neural encoding models. *Network* **15**, 243–262 (2004).