Learning Contrast-Invariant Cancellation of Redundant Signals in Neural Systems

Jorge F. Mejias[1][9]*, Gary Marsat[2,3][9], Kieran Bol[1], Leonard Maler[2,4], André Longtin[1,4]

1 Department of Physics, University of Ottawa, Ottawa, Ontario, Canada, 2 Department of Cellular and Molecular Medicine, University of Ottawa, Ottawa, Ontario, Canada, 3 Department of Biology, University of West Virginia, Morgantown, West Virginia, United States of America, 4 Centre for Neural Dynamics, University of Ottawa, Ottawa, Ontario, Canada

Abstract

Cancellation of redundant information is a highly desirable feature of sensory systems, since it would potentially lead to a more efficient detection of novel information. However, biologically plausible mechanisms responsible for such selective cancellation, and especially those robust to realistic variations in the intensity of the redundant signals, are mostly unknown. In this work, we study, via *in vivo* experimental recordings and computational models, the behavior of a cerebellar-like circuit in the weakly electric fish which is known to perform cancellation of redundant stimuli. We experimentally observe contrast invariance in the cancellation of spatially and temporally redundant stimuli in such a system. Our model, which incorporates heterogeneously-delayed feedback, bursting dynamics and burst-induced STDP, is in agreement with our *in vivo* observations. In addition, the model gives insight on the activity of granule cells and parallel fibers involved in the feedback pathway, and provides a strong prediction on the parallel fiber potentiation time scale. Finally, our model predicts the existence of an optimal learning contrast around 15% contrast levels, which are commonly experienced by interacting fish.

Introduction

For many neural systems, prediction and cancellation of redundant signals constitutes one of the most convenient features for efficiently processing behaviorally meaningful information. When processing sensory input, for instance, neural circuits must be able to discriminate a novel stimulus from the background of redundant or non-relevant signals. A well-known situation in which such a discrimination may be highly advantageous is the so called "cocktail party problem", in which a particularly relevant signal is extracted from a mixture containing other unimportant signals [1,2]. This is known to be useful, for instance, to identify particular voices or sounds for both human and nonhuman animals [1,3], or find and identify mates among conspecifics and heterospecifics [4]. However, the concrete mechanisms that the brain may employ to discriminate and cancel redundant information are presently unknown. It would be, therefore, convenient to identify and closely study natural systems displaying such a cancellation phenomenon, in order to isolate its fundamental principles. Of special interest might be the mechanisms able to conduct the cancellation process over a wide range of realistic conditions, such as canceling redundant signals of different intensities (or with time varying intensities due, e.g., to the relative movement of the receiver and the signal sources) while keeping novel stimuli intact.

The understanding of such a contrast-invariant cancellation mechanism would be beneficial not only for the "cocktail-party problem" in auditory systems, but also for visual neuroscience. Indeed, contrast invariance is a well known and well studied feature of the visual cortex, and particularly of the V1 area [5,6]. A number of ingredients are thought to play a role in contrast invariance in V1, such as inhibition [7,8], gain control [9,10] or membrane fluctuations [9,11], to name a few. However, many of the strategies giving rise to contrast invariance in V1 are still highly debated [12,13] or simply starting to be uncovered [9,14]. Consequently, new findings about how contrast invariance is achieved in other sensory modalities such as the simpler electrosensory system might contribute to understand contrast invariance in V1 and possibly to identify common principles for the corresponding biophysical mechanisms. The contrast-invariant cancellation sketched above stands as an interesting potential example.

A system able to perform cancellation of redundant information is known to exist in the electrosensory lateral-line lobe (ELL) of the weakly electric fish *Apteronotus leptorhynchus* [15–17]. This fish continuously emits a wave-type, high frequency ($600 \sim 1000\ Hz$) sinusoidal electric organ discharge (EOD) to sense its surroundings and communicate with conspecifics. Small objects such as prey produce spatially localized amplitude modulations (AMs) in the EOD. On the other hand, the presence of conspecifics or own-

Author Summary

The ability to cancel redundant information is an important feature of many sensory systems. Cancellation mechanisms in neural systems, however, are not well understood, especially when considering realistic conditions such as signals with different intensities. In this work, we study, employing experimental recordings and computational models, a cerebellar-like circuit in the brain of the weakly electric fish which is able to perform such a cancellation. We observe that *in vivo* recorded neurons in this circuit display a contrast-invariant cancellation of redundant stimuli. We employ a mathematical model to explain this phenomenon, and also to gain insight into several dynamics of the circuit which have not been experimentally measured to date. Interestingly, our model predicts that time-averaged contrast levels of around 15%, which are commonly experienced by interacting fish, would shape the circuit to behave as observed experimentally.

body movements such as tail bending induce spatially global AMs in the EOD. For example, since each fish has a fixed EOD frequency, the proximity of two fish produces an AM in the form of a beat of fixed frequency but time-varying amplitude due to the relative motion of the animals. The depth of these AMs, referred commonly as *contrast*, may depend on physical quantities such as the distance to conspecifics or the amplitude of the tail movement, in the case of global signals, or the size of (or distance to) the prey, in the case of local signals. Both spatially local and global AMs are encoded by electroreceptors (mainly, P-units) that densely cover the body of the fish [18]. In particular, AMs in the EOD are reliably encoded with a modulation in the firing rate of the P-units, which provide feedforward input to pyramidal neurons in the ELL. Interestingly, it has been found that a subpopulation of these pyramidal neurons, called superficial pyramidal (SP) cells, are able to respond selectively to local stimuli (i.e. prey) by removing low-frequency global redundant signals (i.e. tail bending), and thus maximizing the response to novel local stimuli [15,19]. In the following, we will denote this pathway from the P-units to the SP cells as the *feedforward* pathway.

This removal of global signals is also present in another family of electric fish, namely the mormyrid weakly electric fish, although the mechanism differs significantly [20,21]. These fish emit a pulse-type electric field instead of a wave-type field. The pacemaker generating the EOD also conveys spike discharges internally to ganglion neurons, to which the electroreceptors project. Through the so called anti-Hebbian spike-time-dependent plasticity, these ganglion neurons use this internal timing information (corollary discharge) to cancel out the redundant responses from the electroreceptors caused by the fish's own pulses, thus allowing an efficient detection of novel stimuli [22]. For both pulse-type and wave-type fish, the cancellation of global signals is achieved via the activation of a neural circuit denoted, by convention, as the indirect *feedback* pathway (it should be noted, however, that it is actually a longer feedforward circuit from the P-units to the SP cells via DP cells, as we will see below). Such a circuit, which we will denote here simply as feedback pathway, involves a granule cell population, the eminential granularis posterior (EGp), which projects a massive number of parallel fibers (PFs) onto SP cells.

In spite of this common architecture, the cancellation mechanism for the wave-type fish *A. leptorhynchus* is significantly different from the one used by pulse-type fish, not only because of the

nature of its EOD (wave-type), but also because the corollary discharge is not present in wave-type fish. For the particular case of wave-type fish, the presence of burst-induced long-term plasticity in the PF-SP cell synapses [23], together with the segregation of the PFs into frequency-specific channels [16,24], shapes the feedback input to the SP cell into a *negative image* of the redundant sensory stimulus, causing destructive interference and effectively canceling the global stimulus in the SP cells [16].

Little is known, however, about how different stimulus contrasts are processed in such a circuit. The AM contrast level of a signal is strongly correlated with the spatial proximity of the source, either for local input (i.e. distance to the prey) or global input (i.e. distance to conspecifics) [25], and thus constitutes a highly variable feature of the stimulus. Ideally, the fish would be expected to detect the presence of prey (and properly estimate the corresponding distance) while in the presence of other conspecifics at different distances from them. This would imply that SP cells display some form of cancellation for global stimuli of different contrasts. Neither the existence of such a cancellation nor its concrete dependence with the stimulus contrast have been experimentally quantified to date. Furthermore, the mechanisms that might lead to this phenomenon are not known. Arguably, a linear system would be expected to maintain the output as a given fraction of the input, regardless of the input strength. The neural circuits of interest here, however, are known to involve nonlinearities, including not only the input-output nonlinearity arising from the spiking threshold, but also the ones due to the presence of bursting and spike-timing dependent plasticity rules. Due to these nonlinearities, the particular configuration of PFs needed to properly cancel signals with a given contrast might be unable to provide consistent cancellation for another contrast. Also, since PFs are already segregated into different frequency-specific channels [16], it is unlikely that this strategy could be followed again to form contrast-specific channels, due to the limited number of PFs available. Therefore, both novel experimental observations and models are needed to address the question of cancellation of global stimuli with different contrasts.

In this work, we tackled this problem by employing a combination of experimental and modeling methods. We performed *in vivo* extracellular measurements of SP neuron activity for global and local stimuli of different AM frequencies and contrasts. Our measurements clearly show the existence of contrast-invariant cancellation of global stimulus for a wide, behaviorally relevant range of stimulus contrasts. Although a slight decay in cancellation for increasing contrasts is observed, the cancellation level decays only about 10% across all the range of contrasts considered, and thus cancellation remains at all times at values over 80% (100% being a perfect cancellation of the global signal). In addition, a computational model is fitted to our *in vivo* data and the *in vitro* results presented in [23], and it is employed to explore the origins of this contrast invariance. Our model is based on those from previous studies [16,24], which considered a feedback pathway composed of multiple delayed PFs projecting onto SP cells, with the strength of these PFs determined by a burst-induced long-term plasticity rule. In the model presented here, we also consider several novel mechanisms needed to understand contrast-invariant cancellation, which are: (*a*) the explicit modeling of the P-unit input/output characteristics, which affect both the feedforward and feedback pathways, and (*b*) the presence of saturating effects in the feedback pathway. We have also considered that plasticity does not shape PFs quickly, so that the model will have to deal with contrast levels that it was not explicitly trained to cancel. This last point is extremely important, as a system in which PFs are allowed, via long-term plasticity, to

relearn how to cancel every new stimulus would be highly unrealistic.

Employing this highly detailed model fitted to our *in vivo* data, we find that (*i*) in spite of nonlinearities associated with PF plasticity, the level of AM contrast is successfully transmitted through the feedback pathway, matching the contrast arriving at SP cells through the feedforward pathway and explaining the contrast-invariant cancellation found *in vivo*, (*ii*) the PF weights associated with a given contrast are able to provide good cancellation for other contrast levels, and (*iii*) the minor decay of cancellation with contrast is due to the saturation of activity in the feedback pathway. In addition, our model predicts that, in order to properly cancel global signals at the experimentally observed levels, the average contrast level that must drive the PF learning lies around 15% contrast levels. Interestingly, contrast levels around this one are commonly found within the natural environment for communication signals in the weakly electric fish [25]. This hypothetical link between social interaction and redundancy reduction in neural circuits might be used to uncover neural or synaptic mechanisms which are elusive to standard *in vivo* or *in vitro* techniques.

Results

Experimental observations

The goal of this study is to understand the mechanism that neurons in the ELL of the weakly electric fish employ to cancel spatially redundant signals with different contrasts. To do that, we first analyze experimental data from *in vivo* recordings. Fig. 1 shows the extracellularly recorded response of superficial pyramidal (SP) neurons in the ELL under different stimuli. As one can see, SP cells respond strongly and in a phase-locked fashion to local stimuli (Fig. 1A, black), whereas the response is much broader in phase, and smaller in amplitude, when the stimulus is global (Fig. 1A, gray). By considering peri-stimulus time histograms (PSTH), we confirm that, within the range of frequencies of the AM considered, the response to global signals is effectively cancelled (Fig. 1B). As previous studies have addressed, the cancellation is most pronounced within this AM frequency range [16], and it is achieved by the emergence of a *negative image* of the original signal, generated by the feedback pathway [15,16]. More importantly, whereas previous studies [16,24] characterized cancellation for a single stimulus strength, in this study we present the stimulus at different contrast levels (i.e. different strengths). We observe that the stimulus is cancelled efficiently for a wide range of input contrast levels, with cancellation values over 80% in all cases. Cancellation was measured by the ratio in gain between the local and global responses (see Methods for details).

Furthermore, the level of cancellation appears to be approximately the same for all contrasts, from very low (3.25%) to very high (30%) values (Fig. 1C), with a minor decay of cancellation levels observed for very high contrasts. Contrasts higher than 30% were not considered in this study, since P-unit electroreceptors encode AMs in a nearly linear manner up to $25 \sim 30\%$. After that, the activity of P-units saturates and biologically relevant information can not be processed in the same quasi-linear regime [26]. When averaging over all the frequencies considered, we can observe that the degradation of the cancellation process (defined as the complementary of the frequency-averaged cancellation, i.e. $Degrad = 100\% - Cancel$) is restricted to a range between 5% and 15%, and therefore the cancellation of global signals only varies in about 10% for all the range of biologically relevant contrast levels (Fig. 1D).

Modeling local stimulation

We first consider the response of the SP neuron to local stimuli. The dynamics of the neuron membrane potential is modeled following a leaky integrate-and-fire (LIF) formalism [27,28] with an extra burst-inducing mechanism. The subthreshold dynamics of the membrane potential is given by

$$\tau_m \frac{dV}{dt} = -V + [I + \sigma\xi_L(t) + S(t)]_+ + DAP(t), \qquad (1)$$

where $[\ ]^+$ denotes rectification of the input, $S(t)$ is the input from the P-units encoding the sensory stimulus, $DAP(t)$ is the burst-inducing mechanism needed to reproduce the behavior of *in vivo* SP cells [29,30], and $I + \sigma\xi_L(t)$ is a Gaussian low-pass filtered noise of mean I and standard deviation σ to fit the model to baseline (also referred here as *spontaneous*) activity conditions (i.e. no AM). As in the standard LIF formalism, when V reaches a certain fixed threshold, a spike is recorded, and after that V remains at a certain resting value during the absolute refractory period of the neuron.

When stimulated by a sinusoidal input, the model SP neuron responds with a modulation of its firing rate. Fig. 2A shows the maximum firing rate (solid black line) as a function of the amplitude of the sinusoidal-like signal entering the SP cell from the P-units. In color lines, we see the maximum firing rates observed experimentally for different contrasts entering the P-units. By looking at the intersections of the black curve with the color lines, we can determine the relationship between input contrast to the P-unit and input modulation to the SP cell (i.e. the P-unit output). The resulting input-output relationship for the P-units is shown in Fig. 2B. As we can see, the P-units display some degree of saturation for high input contrasts (of about 30%). This agrees with previously known results [26] which show that P-units behave as linear encoders for relatively low contrasts (up to 20%), and beyond that point they start to saturate. We can now easily include such a saturation effect into our model (see section Methods). Once this nonlinearity has been considered, the response of our model agrees very well with the experimental observations for local stimuli, as one can see in Fig. 2C, for different AM frequencies and contrasts.

Modeling global stimulation

We consider now the situation in which we have a spatially global stimulus in the system. The presence of the global stimulus activates the feedback pathway which projects onto the SP cells via the PFs. This implies considering an extra term in the dynamics of the membrane potential of the SP neuron, which is now

$$\tau_m \frac{dV}{dt} = -V + [I + \sigma\xi_L(t) + S(t)]_+ + DAP(t) + \Gamma(w_s - gV). \quad (2)$$

In the last term, Γ is the strength of the feedback (which will depend nonlinearly on the contrast since the feedback pathway is also driven by P-unit activity), and the term $-gV$ mimics the effect of disynaptic inhibition driven by the PFs. More precisely, since the reversal potential of the inhibitory synapse (GABA-A receptors) is close to the resting potential of the SP cell [31], inhibition was modelled as an extra shunting conductance [32]. The quantity w_s is the strength of the particular PF which is active (i.e. which is transmitting a burst arriving from a granule cell) at the phase segment s of the signal cycle. For simplicity, we assume that only one PF is active at a given time (see section Methods for details).
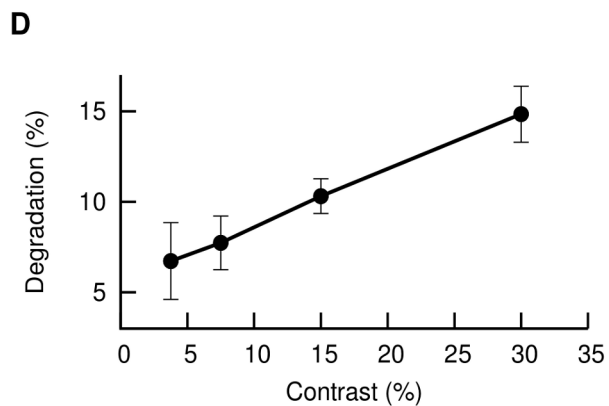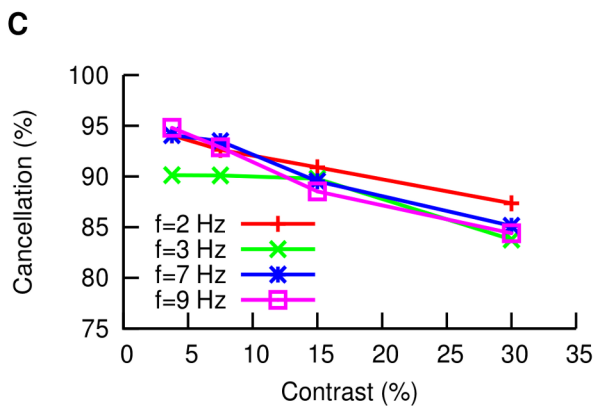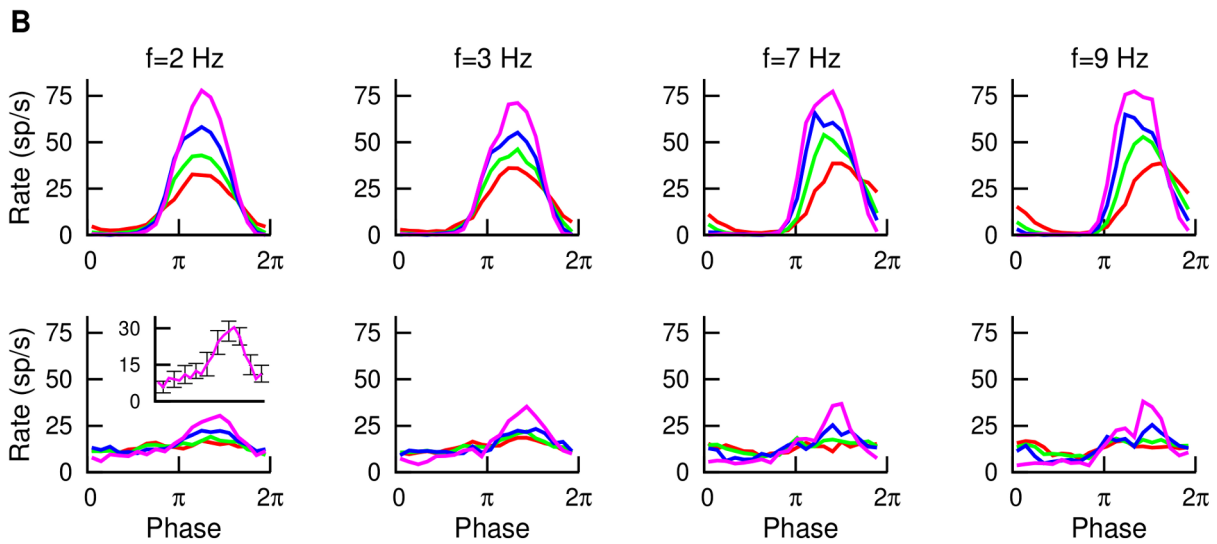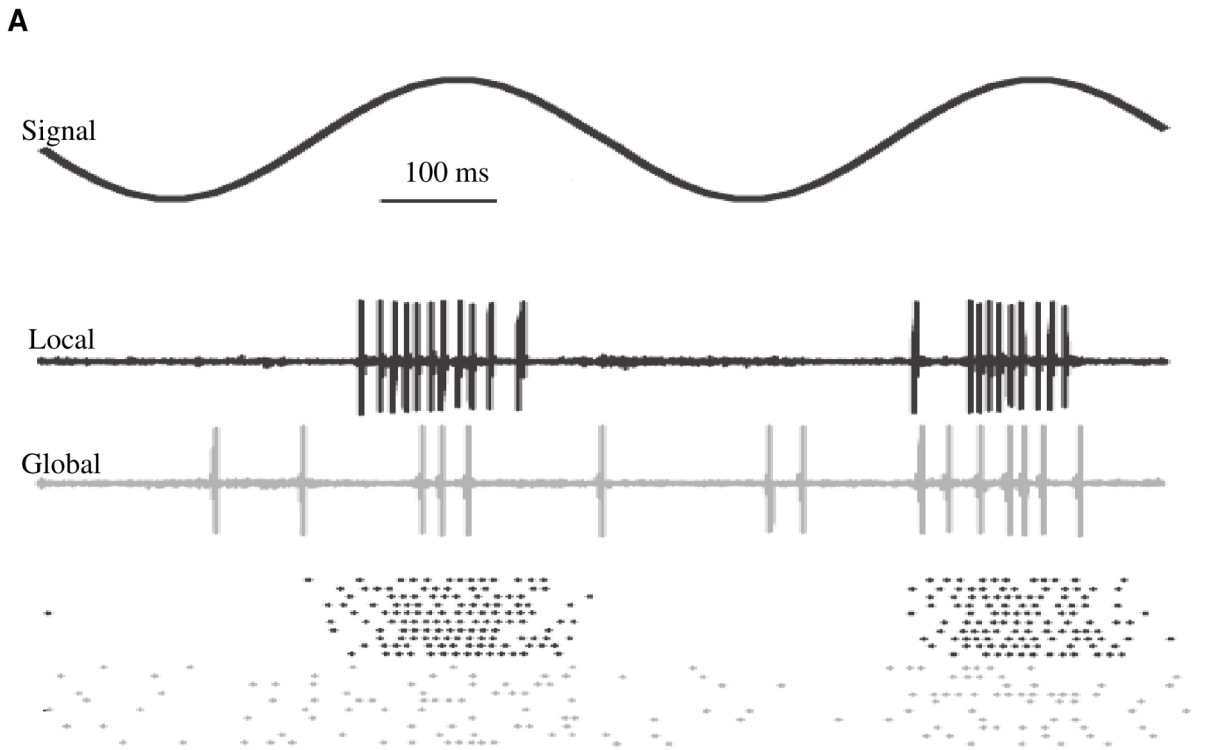
At the SP cell, we distinguish between small and large bursts, since the characteristics of the learning rule will be different depending on the burst size [23]. Adopting the burst definitions which were explicitly characterized in [23], we will consider the 2-spike burst as the typical small burst, and the 4-spike burst as the typical large burst (see Fig. 3A, and section Methods for further details on burst definition). Small and large bursts have different roles in the cancellation process, as it was found in [16]: large bursts cover long periods associated with low-frequency input,

while small bursts have a similar role for high frequencies and are also important in the timing of the plasticity for these input frequencies.

For burst timing purposes, the temporal location of a given burst is identified as the temporal location of its first spike. The burst-STDP learning rule employed, based on *in vitro* experimental recordings [23], is shown in Fig. 3B.

Every time a pair of presynaptic-postsynaptic bursts occur, each PF weight w_s is updated according to the following rule
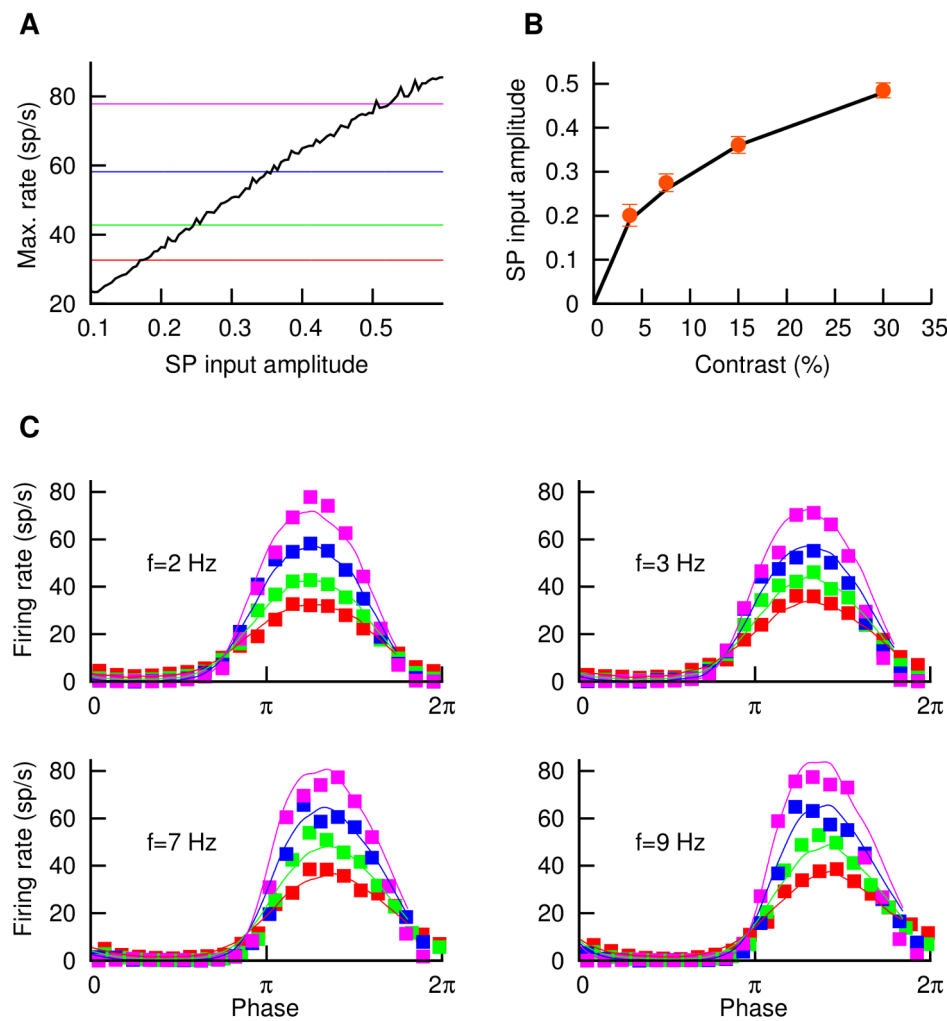
Figure 2. Fitting of the model to local response. (A) Maximum firing rate (solid black line) reached by the model SP neuron as a function of the amplitude of the sine wave $S(t)$ resembling the input from P-units. The AM frequency is $2\ Hz$, and similar responses were found for other frequencies. Colored lines indicate the maximum firing rate observed in the experiments for four different AM contrasts. (B) By considering the crossing points between the black line and the colored lines in panel A, we establish a dependence between AM contrasts (i.e. input to P-units) and amplitude of the signal arriving at the SP cell, $S(t)$, and thus obtaining an AM input-output function for the P-units. Points denote frequency-averaged quantities, while bars denote the standard deviation of each average. (C) Once this P-unit nonlinear feature is considered, the model (solid lines) is able to properly fit the experimental observations (points, shown previously in Fig. 1B) for different input frequencies and contrasts of the local stimulus. For panels A and C, the color code for contrast is red (3.25% contrast), green (7.5%), blue (15%) and violet (30%). doi:10.1371/journal.pcbi.1003180.g002

Figure 3. Parallel fiber weights and learning contrast. (A) Example of small and large bursts as defined in the text. Two spikes occurring within a 15 *ms* window constitute a small burst, whereas four spikes within 45 *ms* constitute a large burst. (B) Burst-driven learning rules for small and large bursts, as a function of the timing Δt of the presynaptic and postsynaptic bursts. Points indicate data from *in vitro* recordings (taken from [23], each point being averaged over 100 trials in the experimental pl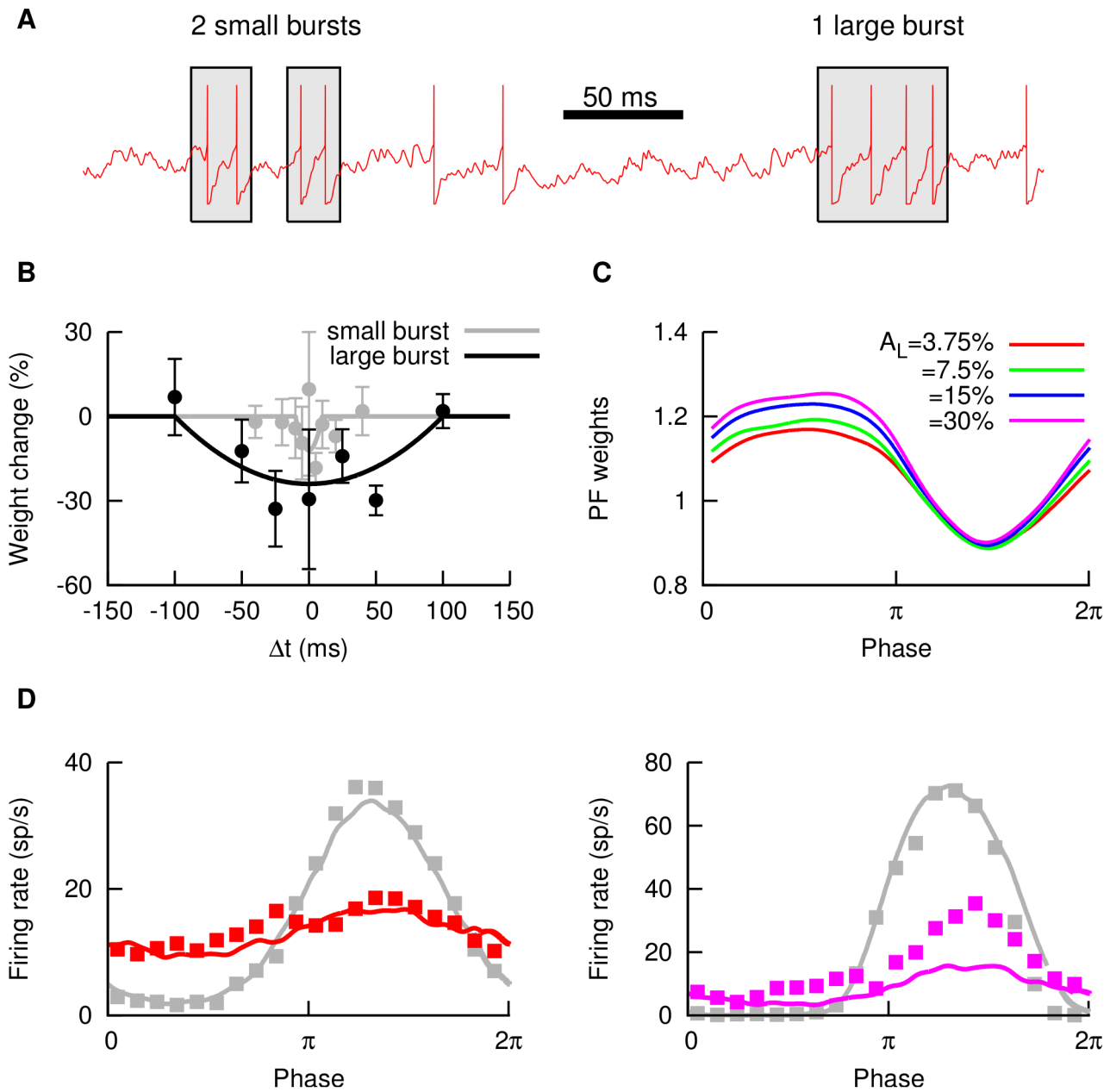asticity protocol), and lines are the fit employed in the model. (C) PF weights as a function of the stimulus phase, for different learning contrasts A_L. (D) Firing rate of the SP cell as a function of the stimulus phase, for local (gray) and global (colored) stimulation. The learning contrast was set at $A_L = 15\%$ and the test contrast was 3.75% (left panel) or 30% (right panel). For both panels, lines correspond to model predictions and points to experimental data shown previously in Fig. 1B.
doi:10.1371/journal.pcbi.1003180.g003

$$w_s \rightarrow w_s - w_s \eta_{2,4} \left[1 - \left(\frac{t_s - t_B}{L_{2,4}} \right)^2 \right]_+, \quad (3)$$

where η_2 and L_2 are used if the burst of the SP cell is a small burst, and η_4 and L_4 are used if it is a large burst. The presynaptic burst is assumed to match the burst type occurring at the SP cell. Once

again, $[\,]^+$ symbolizes rectification, which means this rule is applied to all weights whose phase segment began at a time t_s as long as $|t_s - t_B| < L$. Beyond this range, the weights are unchanged.

Note that the burst-induced depression found *in vitro* is purely depressing and would eventually decrease all PF weights to zero. To avoid that, we include a non-associative potentiating rule so that the weights slowly relax back to w_{max} with a time constant of

τ_w according to

$$\tau_w \frac{dw_s}{dt} = w_{max} - w_s \qquad (4)$$

This rule maintains the independence of synaptic weights and is biologically plausible, since Lewis and Maler [33] demonstrated a presynaptic form of synaptic enhancement in PFs. This enhancement was elicited when PF discharge occurred, without the need of a concomitant pyramidal cell burst response. Such a form of potentiation lasted for many minutes, and a weak potentiation with this or larger τ_w would have been difficult to detect experimentally. Furthermore, similar potentiation rules have been experimentally observed in mormyrid fish, and it has been shown to play an important role in cancellation in these fish [34], in which there is a corollary discharge.

Below we will see that the homeostatic time constant τ_w may play an interesting role in the learning dynamics. The response of the SP neuron model was always recorded after a certain *learning period*, during which all PF weights reached their equilibrium state.

One of the main points to take into account is that the strength of the PF synapses will depend on the stimulus contrast employed during the learning period (we denote such contrast level as *learning contrast*, A_L). This is due to the fact that the occurrence of SP bursts, which are the driving events of PF plasticity, depends strongly on the stimulus contrast. The effect of the contrast on PF weights is shown in Fig. 3C, where one can see that the distribution of PF strengths is similar for different learning contrasts A_L, although not exactly the same. In all cases, the sinusoidal stimulus shapes the PF weights to form a negative image of the signal, which constitutes the basis of the cancellation phenomenon. The weights for different learning contrasts are almost identical around the peak of the stimulus (corresponding here to a phase of $3\pi/2$), where the SP neuron is mainly driven by the stimulus (and noise plays a relatively minor role), and bursts are more likely to occur. The variability of PF weights with the learning contrast is higher for the signal trough (around phase of $\pi/2$), where bursts occur scarcely and are not able to efficiently shape the weight distribution. As we can see, PF weights around the signal trough are higher for high learning contrasts. This is due to the fact that a high-contrast stimulus induces a strong hyperpolarization in the SP membrane potential at the stimulus trough, lowering the chances of bursting for that stimulus phase and preventing the depressing LTP rule to decrease the weights.

Interestingly, even though both the feedforward and the feedback inputs are sinusoidally driven, the distribution of PF weights significantly deviates from a sinusoidal function, as one may clearly observe from Fig. 3C. Such a deviation has its origin in the highly nonlinear nature of burst dynamics in SP cells. Indeed, it is known that bursting rate displays a highly nonlinear, exponential-like relationship with input (see Fig. 5B in [24]). Since bursts are the main events driving the PF learning, the nonlinear input-burst relationship is translated into a nonlinearity in the PF weight distribution emerging from learning. This, successively, turns the feedback input to the SP cell into a highly nonlinear contribution that prevents treatment of the cancellation phenomenon as a trivial linear summation of sine waves that are out of phase with one another.

In Fig. 3D, two examples of cancellation of a global signal of a frequency of $f = 3\ Hz$, a learning contrast of $A_L = 15\%$, and different stimulus contrasts are shown (in each panel, the corresponding SP neuron response to same-frequency, same-contrast local signals is displayed in gray for comparison purposes). As we can see, the cancellation is very good in both cases, although

the model overestimates the degree of cancellation for the 30% contrast case (violet line in right panel). In both panels, experimental data are plotted with points and model results are displayed with lines.

Emergence of contrast invariance

The PF weights are modified via long-term plasticity mechanisms, which operate in the order of minutes to hours. Since changes in stimulus contrast associated with behavior (i.e. tail bending becoming narrower) may occur on the scale of milliseconds to seconds, one can not expect that the weights will be able to adapt fast enough to new presented contrasts in realistic situations. More likely, PF weights will reach some stationary level (as a result of some time-averaged contrast level provided by day-to-day natural stimuli), and then such an equilibrium level will be used to cancel any particular contrast level that the fish receives. In such a situation, we could expect certain differences in the quality of the cancellation depending on the learning contrast assumed in the simulations.

The model prediction of the cancellation level for different frequencies and contrasts, and assuming different learning contrasts, is shown in Fig. 4. In all cases, the cancellation levels are maintained on values over 80% for different contrast levels, as in the experimental observations (see Fig. 1C), and thus indicate the emergence of contrast invariance in the cancellation of global stimuli in the model. It is particularly interesting to note, from the model results, that the specific learning contrast chosen has little impact on the results, contrary to what was expected, and that the levels of cancellation are broadly the same for all AM frequencies (with high frequencies having slightly lower values, as seen also in the experimental data).

Note that, in addition to this counterintuitive lack of impact of the learning contrast in the cancellation, some qualitative differences appear with respect to the experimental data. Concretely, for all frequencies and learning contrasts, the level of cancellation slightly increases with the stimulus contrast according to the model predictions, while in the experiments it slightly decreases. The origins of such model/experiment discrepancy may be diverse, but it is reasonable to assume that they could be mainly due to the lack of a key ingredient in the model's feedback pathway, since the response of the model and experiment for the local signal were in good agreement both qualitatively and quantitatively (see Fig. 2).

Since the discrepancy is mostly evident for large contrast values, one might want to consider, as a first approach, possible features of the real system that may be particularly relevant at those conditions. A relevant factor to consider here is the existence of saturation effects along the feedback pathway. Saturation is inherent to all spiking neurons, and high contrast input to granule cells might cause saturation in two, not mutually exclusive ways: higher contrasts might evoke discharge in a greater number of granule cells and/or it might evoke a higher frequency discharge in granule cells. The first situation was examined by Berman and Maler who used stronger PF stimulation to activate greater numbers of PFs; clear saturation of the PF response was observed (Fig. 5B in [32]). The second scenario was studied by Lewis and Maler [35]; this study demonstrated a saturating SP response to increasing PF stimulation frequency (Fig. 5 in [35]). The presence of saturation in the PF-SP synapses imposes a limit in the feedback strength for increasing stimulus contrasts, which would naturally lead to a devaluation of the cancellation quality for high contrasts as observed *in vivo*. Furthermore, it is reasonable to think that, in addition to this PF saturation, the bursting activity of granule cells might as well saturate for high enough contrast values (since only a
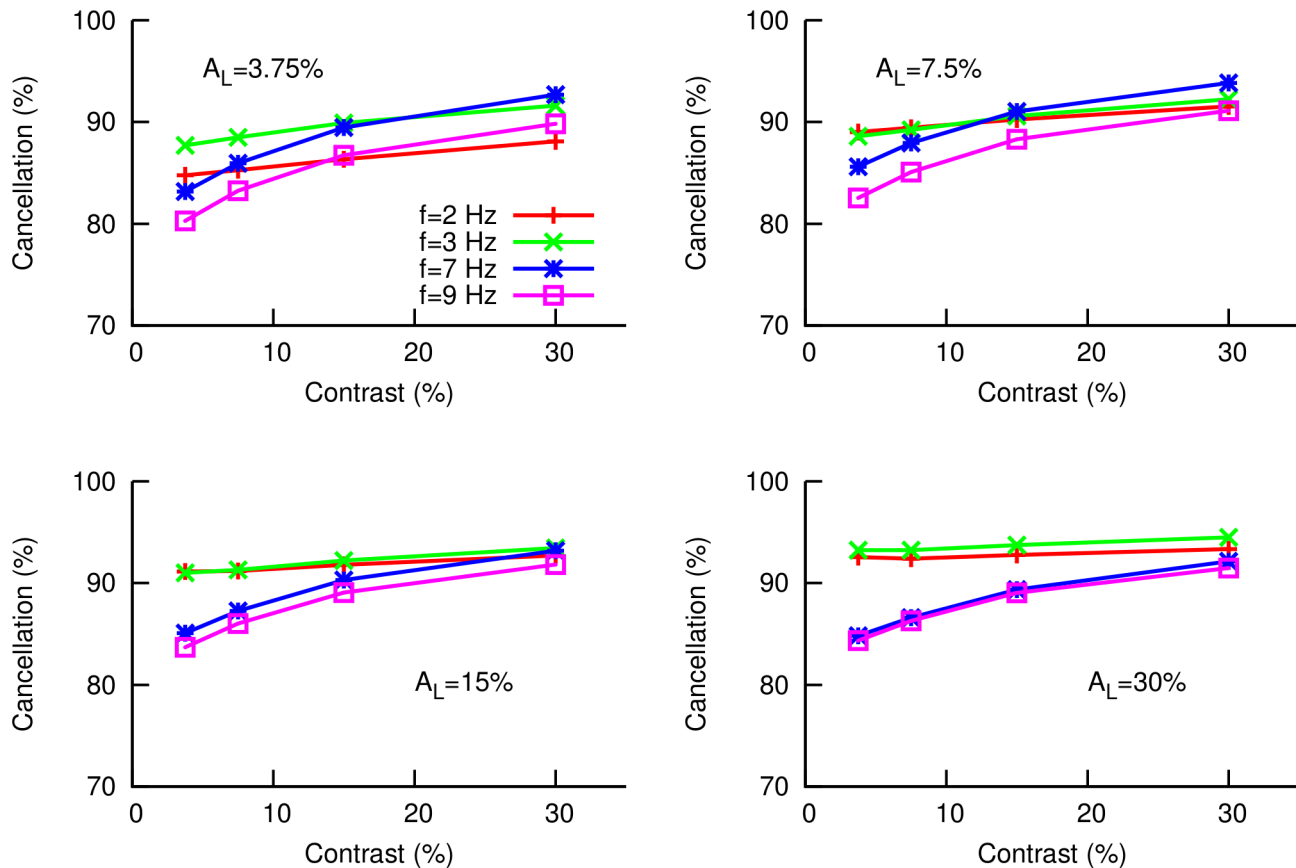
Figure 4. Cancellation of global stimulus without feedback saturation. Model predictions of the level of cancellation of global signals as a function of contrast, for different signal AM frequencies (colored lines) and different learning contrasts considered (different panels). doi:10.1371/journal.pcbi.1003180.g004

limited number of granule cell bursts can be generated within one stimulus cycle), adding an extra layer of saturation to the feedback pathway.

To take into account this saturating behavior in the model, we consider a small correction in the feedback gain for high contrast values by introducing a factor G_s in the last term of Eq. 2. This extra factor will be one for low contrasts (i.e., 3.25% and 7.5%) and less than one of higher values (we chose $G_s = 0.85$ for 15% contrast and $G_s = 0.65$ for 30% contrast, although other values are possible without qualitatively varying our conclusions). The cancellation levels with this new assumption are shown in Fig. 5A, for different frequencies and learning contrasts. As we can see now, the slight decrease in cancellation levels with increasing contrasts is present for all learning contrasts, in agreement with experimental data (shown as a gray line in panels of Fig. 5A for a direct visual comparison). Therefore, we have identified, via a computational model, the saturation of the feedback pathway as a plausible origin of the slight cancellation decrease with contrast. Our model also provides some insight into the plausible saturating dynamics of these granule cells, which have not been recorded *in vivo* to date.

Optimal learning contrast

A second conclusion that we can make from our modeling results concerns the level of cancellation for different learning contrasts. As we can see in Fig. 5A, considering different learning contrasts has now a clear effect on the cancellation properties, as opposed to the case in which saturation of the feedback pathway

was not considered. Indeed, high learning contrasts shift the cancellation curves to higher values, leading to higher cancellation levels for all contrasts and frequencies considered. The closest agreement with the experimental data is obtained with a learning contrast of $A_L = 15\%$, as we can see in Fig. 5A and more clearly in Fig. 5B. Lower learning contrasts lead to low values of overall cancellation, mainly because PF weights are tuned to cancel only weak modulations and are not able to overcome a large-amplitude signal completely. On the other hand, higher learning contrasts produce an over-cancellation at low input contrasts (that is, the SP cells have a peak of firing rate where the stimulus displays a trough, and vice versa) which is not observed experimentally. Due to the saturation of the feedback pathway with contrast, the PFs have to span a wider range of weight values in order to obtain a proper cancellation, and this has a negative effect when trying to cancel low contrast signals. Therefore, the optimal learning contrasts are those situated just below the appearance of a strong saturation in the feedback pathway, but strong enough to allow cancellation for the whole regime of linear encoding of P-units, around 15%. For such an optimal learning contrast, individual firing rate responses are also shown in Fig. 6 as a function of the global stimulus phase for different contrasts and frequencies. The figure also shows the corresponding experimental SP response to local stimulation for comparison purposes.

It might be argued that $A_L = 15\%$ is only optimal when compared to the few other values of the learning contrast considered here. To better characterize the optimal learning contrast and its robustness, we have employed our model to extend

A

B
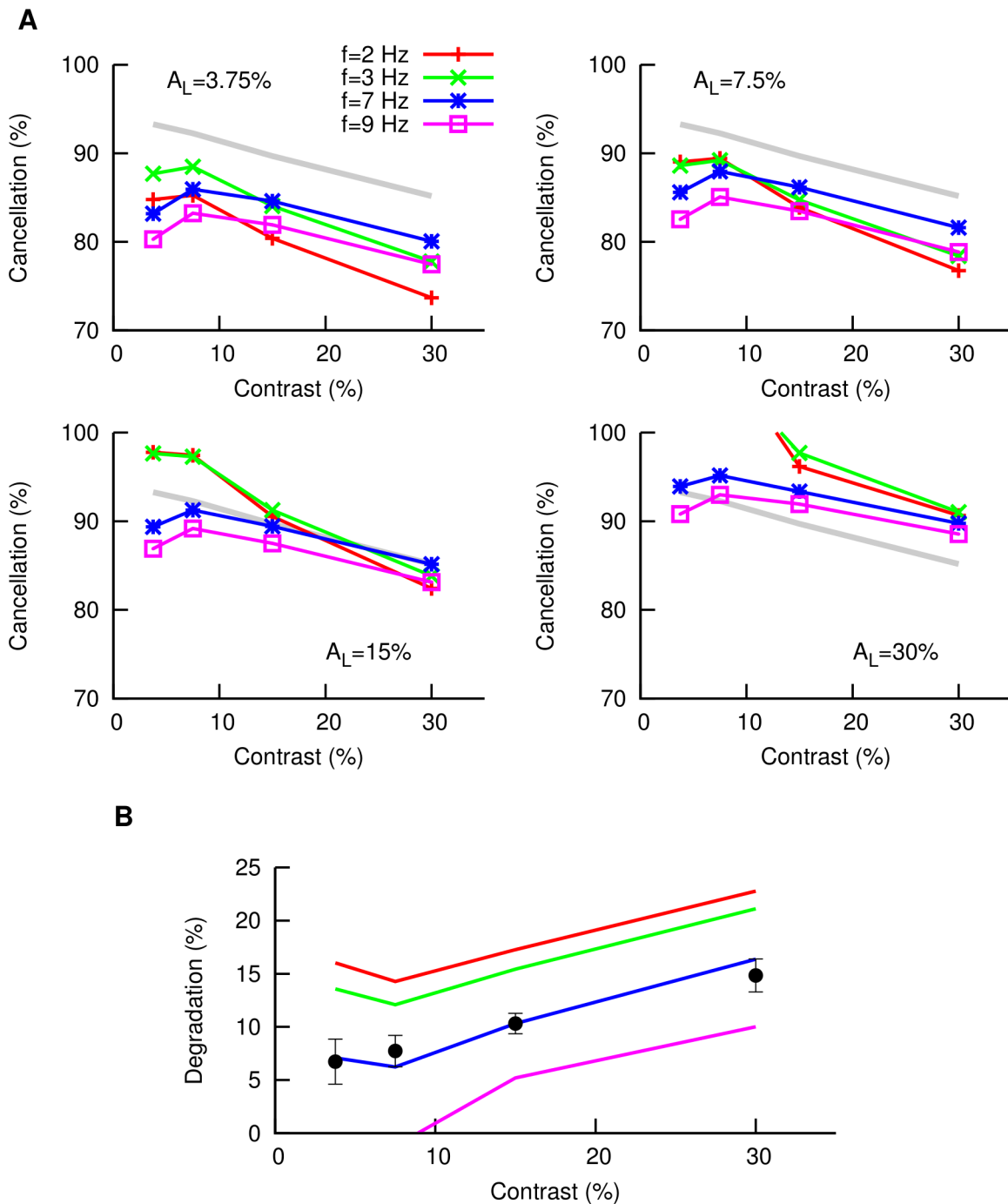
Figure 5. Cancellation of global stimulus with feedback saturation. (A) Model predictions of the level of cancellation of global signals as a function of contrast, for different signal AM frequencies (colored lines) and different learning contrasts considered (different panels). We have introduced here a factor G_s which takes into account the saturation of the feedback pathway for high stimulus contrasts. The gray line indicates the frequency-average cancellation levels measured experimentally (from data in Fig. 1C). (B) Degradation level (defined as the quantity $Degrad = 100\% - Cancel$, once frequency is averaged) as a function of the contrast for different learning contrasts (red $A_L = 3.75\%$, green $A_L = 7.5\%$, blue $A_L = 15\%$, violet $A_L = 30\%$). Symbols denote experimental data. As we can see, the optimal learning contrast is around 15%.
doi:10.1371/journal.pcbi.1003180.g005

our study and to consider other learning contrasts. We consider now a range of possible learning contrasts (around eleven values from 3.75% to 30%) and we compute, for each one of them, the degradation of cancellation as a function of the stimulus contrast, as in Fig. 5B. For learning contrasts not considered

in the experiments, such as $A_L = 10\%$, values for the P-unit adaptation and feedback saturation in the model were obtained by linear interpolation between known values. We also define a function which quantifies the discrepancy between the model prediction for a given A_L and the experimental data. The error

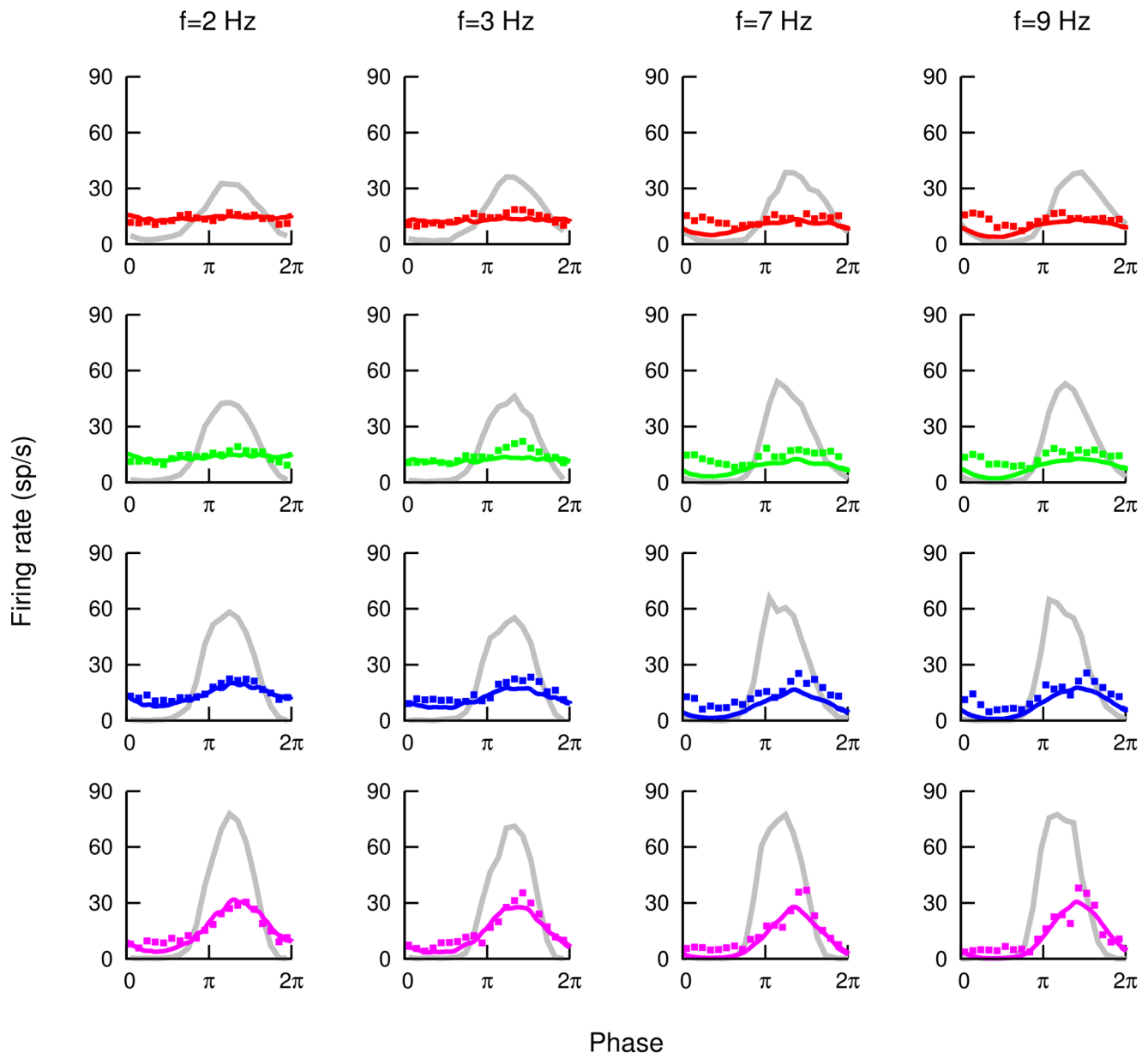**Figure 6. Cancellation at different frequencies and contrasts.** Comparison between experimental data (points) and model predictions (lines) for the cancellation of a global signal for different AM frequencies (columns) and contrasts (red 3.75% contrast, green 7.5% contrast, blue 15% contrast, violet 30% contrast). In each panel, the corresponding experimentally measured SP response to local stimulus is also shown in gray. The learning contrast chosen for the model was $A_L = 15\%$, which optimizes the agreement between model predictions and experimental data.
doi:10.1371/journal.pcbi.1003180.g006

function is given by

$$\text{Error} = \frac{1}{n} \sum_{i=1}^{n} (x_i - y_i)^2, \qquad (5)$$

where i runs over all stimulus contrasts considered (up to $n = 4$ for an easier comparison with experimental data), and x_i, y_i are, respectively, the model and experimental degradation for the stimulus contrast i. As we can see in Fig. 7A, the error function is minimal for $A_L = 15\%$. Considering surrounding contrast levels with similar error function values would give us a range of

optimal learning contrast of $12\% \sim 20\%$ contrast levels, which correspond to reasonably low error levels in the figure. Interestingly, contrast levels around this range are commonly found within the natural environment of the weakly electric fish (Yu *et al.*, personal communication). In particular, it has been experimentally shown that the presence of free-swimming conspecifics induces a certain range of contrast levels in the electric fish, being these levels centered and more common around 15% (see Fig. 3B in [25]), in strong agreement with the optimal contrast level predicted by our model.

In addition to identifying an optimal learning contrast around 15%, our model gives us insight into the dynamics of weak PF potentiation. As it has been argued, the potentiation rule is hard to
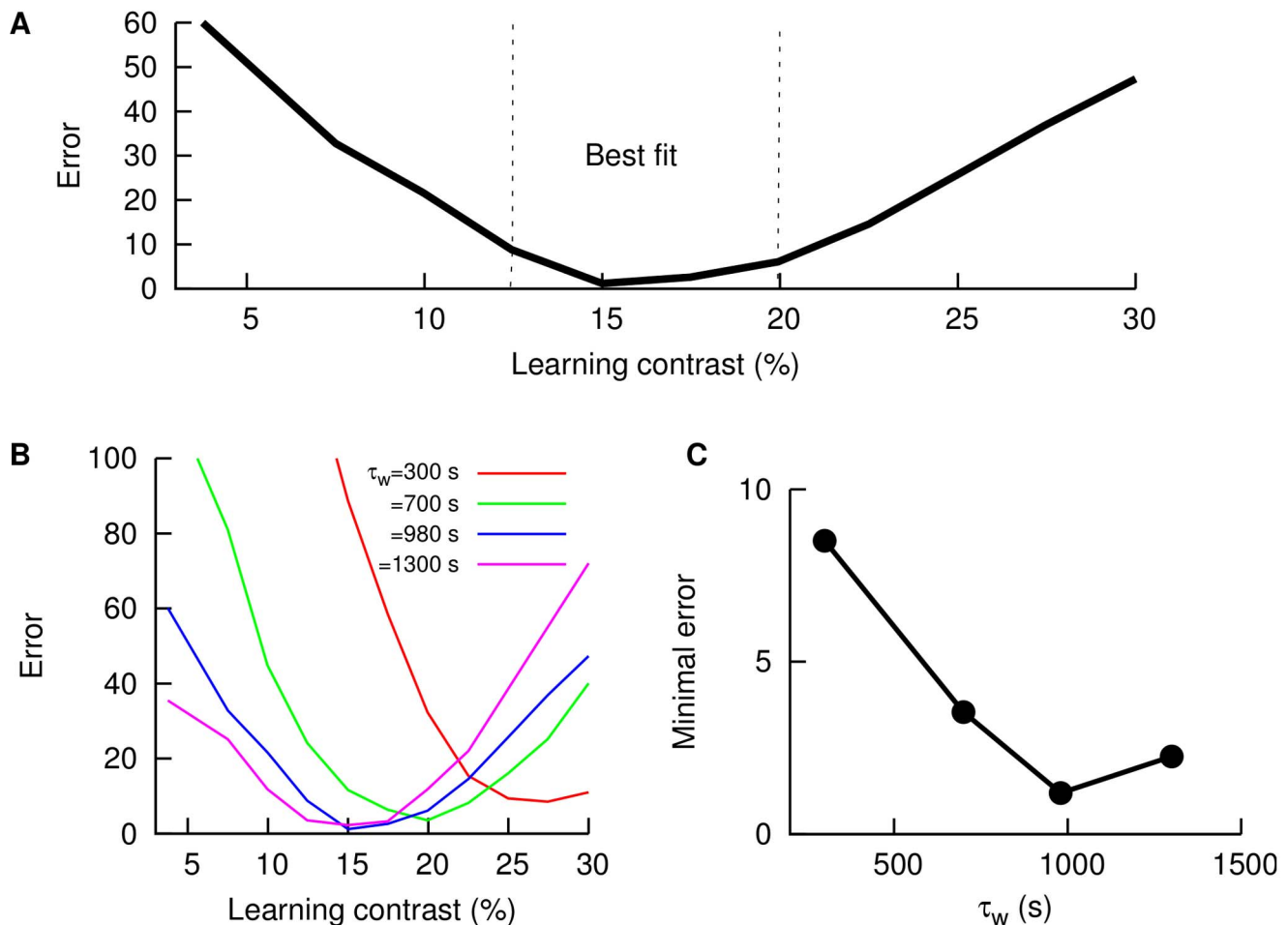
Figure 7. Optimal learning contrast and potentiation time constant. (A) Error function, defined as the sum of the squared distances between experimental and model points, as a function of the learning contrast. Dashed lines, located at 12% and 20%, enclose the region of learning contrast levels which give a reasonably good fit (Error < 10) between data and simulations. (B) Same as panel A, but for different time constant values of the potentiation learning rule for comparison. (C) Minimum of each error curved displayed in panel B, as a function of the potentiation time constant. The lowest error is obtained for potentiation time constants of about 1000 seconds.
doi:10.1371/journal.pcbi.1003180.g007

find experimentally, since it is expected to work at very long time scales and would therefore have a hardly appreciable effect during *in vitro* recordings [23] (associative potentiation rules have been, however, found in mormyrid fish [34]). To study the potentiation time scale in detail, we evaluate the error measurement defined above as a function of the learning contrast, for different values of the potentiation time constant τ_w. As Fig. 7B shows, different τ_w values lead to different error curves. Time constants of $\sim 1000\,s$ or above yield mainly the same results, i.e. the optimal learning contrast lies around 15%. Smaller time constants significantly deviate from this value, which is to be expected since a small time constant would lead to more rapid forgetting of the phase-specific synaptic strength and thus significantly modify the learning rule and the PF weight distributions. The model learning rule would thus not fit the experimental data anymore. For time constants of $\sim 700\,s$ or lower, the optimal learning contrast is found to be higher, which makes intuitive sense because the system forgets faster, but the minimal error reached in these cases is substantially larger than for larger time constants. This can be clearly seen in Fig. 7C: the minimal error decreases as τ_w increases, until a global minimum is reached for $\tau_w \simeq 1000\,s$ (corresponding to an optimal learning contrast of about 15%). This value for the potentiation

time constant is therefore a strong prediction of our model. *In vitro* experiments that pharmacologically eliminate the confounding effects of postsynaptic depression [23] and disynaptic inhibition [35] should be able to precisely estimate τ_w and therefore test our prediction.

Discussion

Removal of redundant information is a key task to accomplish for an optimal detection of novel stimuli in sensory systems. Unfortunately, not many neural mechanisms are known to provide such a filtering under realistic conditions. In this work, we have analyzed one of the few neural circuits clearly identified as a system able to cancel redundant information, which involves the indirect feedback pathway to the ELL of the weakly electric fish *Apteronotus leptorhynchus*. Our results, obtained from a combination of *in vivo* extracellular recordings and detailed computational modeling, reveal a plausible framework which explains the cancellation of redundant information observed in the electric fish [15,16]. They further reveal that this cancellation displays contrast invariance over the entire range of behaviorally relevant contrast levels [26]. The key ingredients for this contrast-invariant

cancellation can be summarized in (*i*) the efficient transmission of the contrast level through the (nonlinear) feedback pathway, resulting in a match of the feedback input to the feedforward signal in the SP, and (*ii*) the fact that, when the PFs adjust their synaptic weights to cancel a given contrast, they provide a good cancellation input for other contrasts as well. Due to these two features, the model is able to explain the high levels of cancellation (over 80% at all times) observed in experiments for a broad range of AM frequencies and contrasts, thus providing a theoretical framework for the contrast invariance in cancellation. This theoretical framework may be helpful to understand other neural systems where cancellation of redundant signals occurs (such as auditory systems [36] or other neural circuits confronting "cocktail party" problems [1,2]) and may also provide novel useful points of view to understand contrast invariance in visual systems [6,9]. In addition, our finding might be seen as a very simple form of context-specific adaptation [37], since the adaptation mechanism (the feedback input into SP cells) would make the SP response different depending on the behavioral context (e.g. prey vs conspecific).

In order to achieve the good agreement of this model with the experimental data, an additional ingredient has been necessary to explain the minor decay of cancellation levels with contrast found *in vivo*. According to our modeling results, a weak saturation in the feedback pathway is a sufficient condition to explain the decay in cancellation for high stimulus contrasts, and this saturation may have different sources. It is known, for instance, that a strong PF stimulation activating a large number of PFs induces a prominent saturation in the PF transmission [32]. Such a saturation phenomenon in PFs would be enough to provide the weak level of saturation needed to explain our experimental findings. Furthermore, PFs are also known to saturate for increasing frequency [35], and other factors such as a possible saturation in granule cell firing or the presence of short-term plasticity mechanisms found in PF synapses [38] may also contribute to the saturation of the feedback pathway. In particular, short-term synaptic depression could be a plausible candidate to induce feedback saturation, as it can induce synaptic fatigue causing nonlinear gain control [39] and, when interacting with short-term facilitation, can produce important effects in the dynamics of recurrent neural circuits [35,40,41].

While we have assumed in this work that granule cells in the EGp fire in a bursty fashion and phase-lock to the periodic stimulus, there is an ongoing debate concerning the propensity of the granule cells to fire in bursts [42–46]. Importantly, we are dealing here with a specialized group of cerebellar cells, the Zebrin-2 negative cells [47], whose firing patterns are not known to date. However, the presence of granule cells with a strong bursting behavior is not required for our conclusions to hold. Since the global stimulus is of a periodic nature, it will likely induce the clustering of granule cell spikes around a certain range of stimulus phases, even if the granule cells do not have a tendency to burst. A simple scaling of the granule cell response with the stimulus, as it occurs for SP cells, is therefore the only essential requirement of our model. Similarly, the concrete input/output characteristics of granule cells do not have a strong impact on our results, as saturation for large contrasts has already been found experimentally in parallel fibers [32], and therefore it is not necessary to impose this condition to granule cells.

The possible role of different types of inhibition in the cancellation of global signals has been experimentally addressed previously. For instance, Maler *et al.* [48] presented morphological evidence suggesting the presence of lateral inhibition (but not recurrent inhibition), as the one we are considering in our model. Bastian *et al.* [49] demonstrated the existence of inhibitory surrounds for superficial pyramidal cells, but it was later observed [15] that the cancellation could be completely prevented by blocking the EGp feedback, and thus suggesting that these inhibitory surrounds do not have an important role for cancellation. Therefore, the role of other types of inhibition can not be completely ruled out, but their effects on cancellation have been found to be much less important than the PF-triggered feedback inhibition that we are considering in our study. On the other hand, local input activates feedforward inhibition, but this type of input does not trigger cancellation as we also illustrate in Fig. 1B.

It is also important to mention that, due to the fact that the same stimulus is driving both the SP cells (via the feedforward pathway) and the EGp (via the feedback pathway), parallel fibers will be naturally time-locked to the stimulus. As a consequence, any initial phase displacement in the stimulus (with respect to previous stimuli) will affect both the SP cells and the EGp in the same way and will not affect the cancellation. Sudden and fast phase shifts like the ones associated with communication signals (i.e. small chirps [50]), however, will not be predicted and cancelled by the present mechanism, and they will be treated as novel stimuli since they may carry useful behavioral information.

In addition to different stimulus contrasts, we have considered signals with different AM frequencies in our study. Local signals (such as prey), which should not be cancelled, usually fall into the range of frequencies considered here (from 2 to 9 *Hz*). Such a range of frequencies lies within the band of good cancellation observed experimentally and, consequently, we do not observe major differences among AM frequencies in the cancellation level [16]. A slight decrement in cancellation is however observed for 9 *Hz*, which is indeed to be expected since cancellation starts to decay around 8 *Hz* and is practically inexistent at 20 *Hz* [16,24]. On the other hand, while frequency-specific channels have been identified in this system [16], the width of a given frequency channel is not known to date. In our model, we have assumed that frequencies as close as 2 and 3 *Hz* are canceled via different frequency-specific channels, but it might be possible that both frequencies fall into the cancellation domain of a single channel. However, the detailed mechanisms that such a broadband channel could employ to cancel close (but different) frequencies are unknown and they fall out of the scope of our study. Therefore we assumed here that each frequency was processed by a specific channel. The good agreement between our experimental observations and model predictions suggests that our approach may be indeed adequate.

It is also worth mentioning that the specific definition of burst does not have a fundamental importance in our model. In particular, the one used here (i.e. the occurrence of a number of spikes within a fixed time window) has been chosen for being computationally adequate, but also for being consistent from a biophysical point of view. This is explained by the following two factors: (*i*) the ISI distribution of SP cells is bimodal [16], with a clear frequency-independent peak at small ISI values which highlights the existence of bursting [51], and (*ii*) in our system, plasticity is not triggered by single-pulse paired stimulation [23]. The combination of these two factors suggests that SP bursts are structured and well located in time, and highlights the importance of bursts as functionally meaningful events which clearly differ from single spikes.

The PF plasticity rule, as presented in [23], would constantly weaken PF strength until all synapses would reach zero strength. To avoid that, a weak, phase-independent potentiation mechanism was considered here following previous studies [16]. This potentiation rule might have to be extremely slow such that its effect was not detected in standard *in vitro* protocols. Such a plasticity mechanism would therefore be hard to measure for most direct methods. Aided by our model and experimental findings, we

were able to predict a reasonable value for the potentiation time constant, of about $\tau_w \simeq 1000\ s$. This estimation constitutes a strong prediction of our model, and further modeling and experimental studies will be necessary to corroborate and extend this prediction.

Our model indicates the existence of a certain *optimal learning contrast*, which presumably resembles the time-averaged contrast that should drive the PF plasticity to obtain the cancellation values observed experimentally. We have found that this optimal learning contrast lies around 15% contrast levels (or, considering a small range of contrast levels, around $12\%\sim20\%$). Interestingly, contrast levels around this value are commonly found within the natural environment for communication signals in the weakly electric fish, such as, for instance, when surrounded by free-swimming conspecifics [25]. Indeed, assuming that a global AM of 15% contrast level is due to the presence of a conspecific, this would correspond roughly to a distance of $12\sim16$ cm between both fish [52]. In addition, this indicates that a close experimental measurement of the common contrast levels found in the fish's natural environment (which has been the goal of recent studies [25]) might provide a good indirect confirmation of the existence of weak potentiation rules which are hard to find in *in vitro* conditions. Further experimental and modeling work is needed, however, to clarify these possible links, as well as the impact of other realistic assumptions in our circuit, such as considering heterogeneous populations of superficial neurons [53,54].

Finally, the study of cancellation of global signals needs to be extended to situations in which more realistic stimuli are considered. Although of remarkable usefulness, the assumption of global sinusoidal signals would correspond only to the case of a perfectly periodic tail movement, or to the presence of a static conspecific at a certain fixed distance. However, tail bending is commonly an aperiodic movement in real conditions and, in addition, the distance between two electric fish would constantly vary as they swim. This implies that the stimulus contrast will vary in time, as for instance following an Ornstein-Uhlenbeck process as shown by Yu *et al.* [25], and such variations are likely to be relevant in the cancellation process. This constitutes a much more complex situation than the one studied here (in which each of the contrast levels we have investigated is constant in time), although preliminary work suggests that the present framework may be extended to explain the cancellation for those complex situations as well. For instance, in a complex global signal constituted by several coexisting frequency components, each one of them could be cancelled independently by frequency-specific channels present in the feedback circuit [16]. Furthermore, slow variations in the contrast level for any given frequency might activate other adaptation mechanisms that could aid in the cancellation, as we are currently investigating.

Methods

Ethics statement

All experimental procedures were approved by the University of Ottawa Animal Care Committee.

In vivo electrophysiology

Experimental recordings were performed as in [16]. Briefly, craniotomy is performed under general anesthesia. During the experiment the fish is awake but paralyzed with curare and locally anesthetized. Single-unit extracellular recordings from superficial pyramidal E cells of the centro-lateral segment of the electro-sensory lateral line lobe were performed. These cells can be easily identified based to their location (depth and centro-medial position), their receptive field, their baseline firing rate and response properties. Stimuli consisted of amplitude modulations of

the fish's own electric field. The stimulus was delivered through two large global electrodes placed on each side of the fish thereby achieving a global stimulation. For local stimulation, a small dipole was placed in the center of the cell's receptive field; the distance between the dipole and the skin was adjusted to maximally stimulate the whole receptive field of the cell while avoiding stimulation of receptors outside the classical receptive field.

Network architecture

To understand the biological mechanisms responsible for contrast-invariant cancellation of global stimuli, we consider a simplified model of the ELL and the indirect feedback pathway (see Fig. 8). AMs of the sensory input are encoded in firing rate modulations of the P-units. If the stimulus is spatially local, P-units transmit these modulations to the SP neuron, which projects to other higher brain regions. If the stimulus is spatially global, the feedback pathway is also activated (in addition to the feedforward pathway sketched above) and a population of EEL neurons called deep pyramidal (DP) cells transmit the signal from the P-units to a granule cell population, the eminential granularis posterior (EGp), via the Nucleus praeminentialis (nP). Each granule cell projects through parallel fibers onto the SP neuron, closing the feedback loop.

In the real system, the cerebellar feedback pathway to the ELL is bilateral [18,55] and can take several routes before returning to the ELL. Furthermore, DP cells, which constitute the origin of the feedback pathway, display a variety of phase relationships with the stimulus depending on its location (i.e., the side of the body) and the specific cell type (E-cells, which fire at the signal peak, or I-cells, firing at the trough) [56,57]. Finally, each granule cell will be located at a certain position in space and therefore it will be characterized by a particular distance to the target SP cell. All these elements together produce a wide range of feedback temporal delays, suggesting that the PFs can provide feedback to SP neurons at all possible phases of the global periodic stimulus. In addition to this, granule cells have been reported to phase-lock to periodic signals and to burst to sensory stimuli and be silent elsewhere [43,45,58,59].
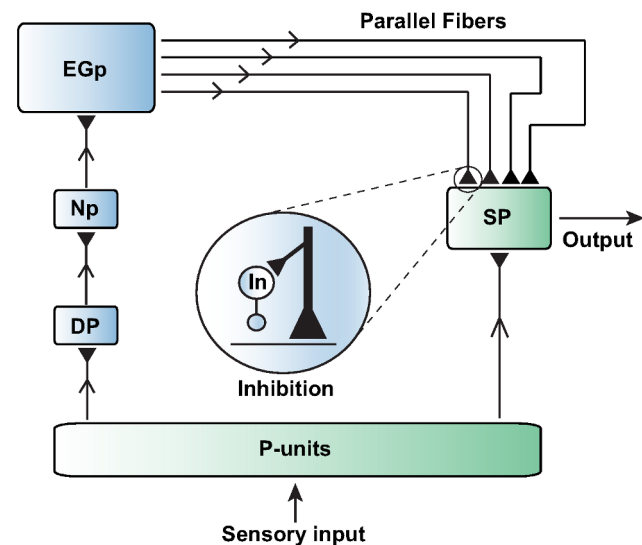

Figure 8. Schematic diagram of the model considered. The network architecture considered in the model involves the feedforward circuit (in green) and the indirect feedback pathway (in blue), which is active only for global stimulus.
doi:10.1371/journal.pcbi.1003180.g008

To include these features in our simplified model, we assume that (1) the array of PFs provide feedback to SP neurons at all possible phases of the sensory stimulus, and (2) granule cells in the feedback pathway respond in a bursty fashion, phase-locked to the AM frequency signal. We also take into account in the model that PFs also synapse onto inhibitory interneurons, which provide some level of inhibition to the SP neuron.

It has also been shown that certain long-term plasticity rules may adjust the weights of the parallel fibers. According to recent *in vitro* experiments, PFs projecting onto SP cells display a long-term depression (LTD) rule that depends on the timing of presynaptic and postsynaptic bursts [23]. Such a burst-driven learning rule, combined with the presence of PFs displaying a wide set of temporal delays, is thought to be responsible for the generation of a negative image of the input AMs, providing the substrate for signal cancellation [16].

LIF model

For local stimulation, the SP neuron is modelled following a leaky integrate-and-fire (LIF) formalism (Eq. 1) with an extra term accounting for the bursting dynamics (DAP). Noise was introduced in the system via a low-pass filtered (with cut-off frequency f_{cut}) Gaussian noise $\xi_L(t)$ of zero mean and variance one. The variance is later adjusted via the parameter σ, and a constant bias I is introduced to fit the experimental firing rate in spontaneous (i.e. no AMs) conditions. As in the standard LIF model, a spike is recorded when V reaches the threshold V_{th}, and after that V remains at a certain resting value V_r during the absolute refractory period τ_{ref} of the neuron.

The EOD signal arriving at electroreceptors can be described, as a first approach, as a sinewave of amplitude A_{EOD} and high frequency ($f_{EOD} \sim 1000\ Hz$). The presence of stimuli induces EOD amplitude modulations of frequency $f (\ll f_{EOD})$ and contrast A, so that the EOD amplitude is also a sinewave given by $A_{EOD}(t) = A \sin(2\pi f t)$. Since electroreceptors encode AMs by modulating their firing rate, the AM frequency f and the contrast A are enough to characterize their behavior. The output firing rate $S(t)$ of the P-unit population, which is driven by this input, is given by

$$S(t) = \kappa(A)\sin(2\pi f t). \qquad (6)$$

P-units are known to display some level of saturation for high contrasts [26] (this effect is denoted in the above equation as $\kappa(A)$), and we observe such saturation in our *in vivo* recordings via the nonlinear response of the SP cell to different stimulus contrasts. By fitting the model SP cell response to the *in vivo* measured SP cell response, we determine the input-output amplitude relationship of the P-units (see Fig. 2A and B). The corresponding values are $\kappa(3.75\%) = 0.201$, $\kappa(7.5\%) = 0.275$, $\kappa(15\%) = 0.361$ and $\kappa(30\%) = 0.485$. When using values other than these ones, linear interpolation was employed to estimate the new values of $\kappa(A)$. To incorporate the effect of P-unit adaptation at low frequencies (as in [16]), we multiply the signal $S(t)$ by a small constant factor of 1.15 for $f > 5\ Hz$.

In addition, as electroreceptor input is strictly excitatory, the input to the SP neuron is rectified and $[\ ]^+$ in Eq. 1 symbolizes rectification (that is, $[x]_+ = x$ if $x > 0$, and $[x]_+ = 0$ otherwise). This aided us in incorporating the rectified nature of the pyramidal cell activity into the model. Values for these parameters are displayed in Table 1.

Bursting mechanism

Superficial cell bursting drives the long-term plasticity rules operating in the PFs of the feedback pathway. The term $DAP(t)$

in Eq. 1 models the depolarizing after-potential (DAP), an injection of current into the soma of the neuron after an action potential is fired due to presence of active channels in the cell's dendrites. This effect has been modeled previously in superficial cells [29,30], and we adapt the parameter values used in these works to match the bursting rate of the model to our experimental observations.

The mechanism responsible for the generation of bursting is the following: after the cell fires ($V = V_{th}$) at time t_n, it will receive a DAP (i.e. a small current injection) a short time later, as long as the previous time the cell fired is not too recent. This additional stimulation is modeled as a difference in alpha functions $s(t,z)$ (Eq. 8), one generated by the soma voltage, and the other by some mean dendrite voltage. However, if the time interval between this spike time t_n and the previous spike time t_{n-1} is less than the refractory period of the dendrite, r_d, then the DAP is inactive for the current spike. The refractory period r_d is modeled as a dynamic variable $r_d(t)$ that changes according to a secondary variable, b, which is updated for each spike. The time just after the most recent spike was fired is referred as t_n^+. The equations governing the DAP [30] are

$$DAP(t) =$$
$$\begin{cases} 0 & \text{if } t - t_n < r_s \\ \alpha\{s(t - t_n, \beta b(t_n^+)) - s(t - t_n, \gamma)\} & \text{if } t - t_n > r_s \text{ and } t_n - t_{n-1} > r_d(t_n^+) \\ 0 & \text{if } t - t_n > r_s \text{ and } t_n - t_{n-1} < r_d(t_n^+) \end{cases} \qquad (7)$$

$$s(t,z) = \frac{t e^{\frac{-t}{z}}}{z} \qquad (8)$$

$$r_d(t) = \mu_3 + \mu_4\ b(t) \qquad (9)$$

Table 1. Parameter values for the equations of the model which approximate experimental data.

Parameter	Value
V_{th}	1
V_r	0
τ_m	7 ms
τ_{ref}	0.7 ms
I	0.59
σ	0.768
g	1.44
f_{cut}	500 Hz
τ_w	980 s
w_{max}	1.5
η_2	0.0018
η_4	0.0036
L_{w_2}	10 ms
L_{w_4}	100 ms

doi:10.1371/journal.pcbi.1003180.t001

$$\frac{db}{dt} = -b/\tau_b + \left(\mu_1 + \mu_2\, b^2\right) \sum_n \delta(t - t_n) \qquad (10)$$

The parameters used in the above equations are listed in Table 2.

Parallel fibers

The feedback pathway is initiated by DP cells, which do not exhibit global cancellation behavior since they do not receive feedback, and these cells project onto the nucleus praeminentialis (nP) which in turn projects onto EGp granule cells. Finally, granule cells project massive numbers of excitatory PFs to the ELL, where they provide input to SP cells as well as local interneurons (stellate cells). The stellate cells in turn inhibit the SP cells via shunting GABA-A receptor channels.

Due to difficulties in recording them, the firing activity of EGp granule cells in the electric fish is not known. Similar granule cells in mammals, however, have been shown to respond to sensory input [43,58] and to phase-lock their bursting to sinusoidal input [45]. We assume here, therefore, that the activity of each PF is one burst per stimulus period.

Considering the natural distribution of temporal PF delays (see details on network architecture above), the total bursting PF input was assumed to be continuous in time. In the model, the feedback cycle associated with the stimulus cycle was discretized into segments of 2.5 ms each. This implies that the number of segments changes with the AM frequency considered (for instance, we would have 200 feedback segments for a 2 Hz stimulus, and 100 segments for an 4 Hz stimulus). Each segment, labeled s, becomes active at time t_s, has a global strength Γ (common for all PFs) and a synaptic weight w_s (particular for each PF, see Eq. 2), and then becomes inactive at $t_s + 2.5$ ms. Each segment is associated with the activity of a given PF for simplicity, although it could be associated with the activity of a certain set of coincident PFs as well. The total excitatory feedback input is therefore a step-wise continuous and periodic signal given by Γw_s, for each segment s as time moves from segment to segment during a period. Disynaptic inhibition, which is also modulated by PF activity, is modeled as an extra shunting conductance $-gV$.

The global strength of the feedback is driven by DP cells, which in turn are driven by P-units. This implies a dependence of the feedback strength on the P-unit response which is modeled as a dependence of Γ with the contrast A:

$$\Gamma = \Gamma_0 G_s\, \kappa(A), \qquad (11)$$

Table 2. Parameters used in the DAP model.

Parameter	Value
α	20
β	2.45 ms
γ	1.4 ms
μ_1	0.6
μ_2	2
μ_3	0.7 ms
μ_4	24.5 ms
r_s	0.7 ms
τ_b	7 ms

with $\kappa(A)$ being the input/output relationship of the P-units. The parameter Γ_0 is set to 4.16 for the global stimulus (in order to fit the mean firing rate measured experimentally at global stimulation of 2 Hz and 7.5% contrast), and to zero for the local stimulus (since this type of stimulus does not activate the feedback pathway). The saturation of the feedback pathway is being taken into account in the parameter G_s, which will be one for low contrasts (i.e., 3.25% and 7.5%) and less than one of higher values (we chose $G_s = 0.85$ for 15% contrast and $G_s = 0.65$ for 30% contrast). When feedback saturation is not being considered in the model, we just set $G_s = 1$ for all contrasts. Again, since P-units drive feedback, the P-unit adaptation at low AM frequencies discussed above will affect as well the feedback input. Therefore, we follow the same criterion as with $S(t)$ and we multiply the whole feedback function by a small factor 1.15 if $f > 5$ Hz to account for P-unit adaptation.

A necessary condition for cancellation is to have a stable phase relationship for each segment and, hence, each weight. Such a requirement is fulfilled by considering that there is a particular set of PFs responsible for the cancellation of a given AM frequency. This has been corroborated with *in vivo* electrophysiological measurements [16], and therefore we assume here that our feedback pathway is frequency-specific.

As also observed experimentally in [16], cancellation starts to decay at high AM frequencies, hypothetically due to failures of granule cells for bursting at least once per cycle under global stimulation (and thus failing to drive learning properly). In agreement with these observations, we notice a slightly consistent decay of cancellation for $f = 9$ Hz, which can be easily taken into account by reducing Γ_0 to a value of 3.12 for this frequency, to improve the fitting of experimental data.

Burst definition and cancellation measurement

Following the definition of a burst that induces plasticity [23], the spike train of the model SP cell was constantly monitored for small (2 spikes within 15 ms) and large (4 spikes within 45 ms) bursts. These particular definitions of burst are only adopted here to simplify the online computations, as the quantitative behavior of our model does not depend sensitively on such assumptions, or even on the presence of strong intrinsic mechanisms for bursting generation (see Discussion for details).

It is worth mentioning here that spikes in each SP burst must be independent (i.e. there cannot be a small burst in a large burst, or a large and a small burst in 5 spikes). Since each PF segment produces a presynaptic burst arriving at the apical dendrite, there is one PF burst at every time t_s in the model, and thus PF bursts are spaced 2.5 ms apart. When the SP cell bursts under global stimulation at time t_B, the burst learning rule identified *in vitro* (Eq. 3) is immediately invoked for all PF segments.

To quantify the level of cancellation of the global stimuli, we follow [16] and employ the following criterion:

$$\text{Cancellation}(\%) = 100\left(1 - \frac{Z_G}{Z_L}\right), \qquad (12)$$

where Z_G, Z_L are, respectively, the amplitude of SP response (measured in spikes per second) for global and local stimulation. More precisely, the PSTH (i.e. firing rate as a function of stimulus phase) of the SP cell was fitted to a sinusoidal function (plus baseline level) for the global stimulation case, and the amplitude of such sinusoidal was taken as Z_G. Because of the rectification, the same fit could not be applied to the local stimulation case, since the response clearly deviates from a sine wave. The SP response to local stimulation was then fitted to a Gaussian distribution (plus baseline level) and the height of such a Gaussian was taken as Z_L.

This criterion was followed for both experimental data and model predictions.

The degradation measurement is the complementary of the frequency-averaged cancellation, and it was employed to show, in a clear manner, how much the cancellation level is degraded when increasing the stimulus contrast (Fig. 1D). It is defined as

$$\text{Degradation}(\%) = \left(100 - \frac{1}{N_f} \sum_{i}^{N_f} \text{Cancellation}_i(\%) \right) \quad (13)$$

with $N_f = 4$ is the number of AM frequencies considered in the

study (which are 2, 3, 7 and 9 Hz) and the sum runs over all these frequencies.

Author Contributions

Conceived and designed the experiments: JFM GM KB LM AL. Performed the experiments: JFM GM. Analyzed the data: JFM GM LM AL. Contributed reagents/materials/analysis tools: JFM GM. Wrote the paper: JFM GM LM AL.

References

1. Haykin S, Chen Z (2005) The cocktail party problem. Neural Comput 17: 1875–1902.
2. Clopath C, Longtin A, Gerstner W (2008) An online hebbian learning rule that performs independent component analysis. Advances in Neural Information Processing Systems 20: 312–328.
3. Bee MA, Micheyl C (2008) The cocktail party problem: What is it? How can it be solved? And why should animal behaviorists study it? J Comparat Psychol 122: 235–251.
4. Bee MA (2008) Finding a mate at a cocktail party: spatial release from masking improves acoustic mate recognition in grey treefrogs. Animal Behavior 75: 1781–1791.
5. Skottun B, Bradley A, Sclar G, Ohzawa I, Freeman R (1987) The effects of contrast on visual orientation and spatial-frequency discrimination - a comparison of single cells and behavior. J Neurophysiol 57: 773–786.
6. Ferster D, Miller K (2000) Neural mechanisms of orientation selectivity in the visual cortex. Ann Rev Neurosci 23: 441–471.
7. Troyer T, Krukowski A, Priebe N, Miller K (1998) Contrast-invariant orientation tuning in cat visual cortex: thalamocortical input tuning and correlation-based intracortical connectivity. J Neurosci 18: 5908–5927.
8. Anderson J, Carandini M, Ferster D (2000) Orientation tuning of input conductance, excitation, and inhibition in cat primary visual cortex. J Neurophysiol 84: 909–926.
9. Finn IM, Priebe NJ, Ferster D (2007) The emergence of contrast-invariant orientation tuning in simple cells of cat visual cortex. Neuron 54: 137–152.
10. Nowak LG, Barone P (2009) Contrast adaptation contributes to contrast-invariance of orientation tuning of primate V1 cells. PLoS One 4: e4781.
11. Hansel D, van Vreeswijk C (2002) How noise contributes to contrast invariance of orientation tuning in cat visual cortex. J Neurosci 22: 5118–5128.
12. Holt G, Koch C (1997) Shunting inhibition does not have a divisive effect on firing rates. Neural Comput 9: 1001–1013.
13. Sutherland C, Doiron B, Longtin A (2009) Feedback-induced gain control in stochastic spiking networks. Biol Cybern 100: 475–489.
14. Carandini M (2007) Melting the iceberg: contrast invariance in visual cortex. Neuron 54: 11–13.
15. Bastian J, Chacron M, Maler L (2004) Plastic and nonplastic pyramidal cells perform unique roles in a network capable of adaptive redundancy reduction. Neuron 41: 767–779.
16. Bol K, Marsat G, Harvey-Girard E, Longtin A, Maler L (2011) Frequency-tuned cerebellar channels and burst-induced LTD lead to the cancellation of redundant sensory inputs. J Neurosci 31: 11028–11038.
17. Marsat G, Longtin A, Maler L (2012) Cellular and circuit properties supporting different sensory coding strategies in electric fish and other systems. Curr Op Neurobiol 22: 686–692.
18. Maler L (2007) Neural strategies for optimal processing of sensory signals. Prog. Brain Res 165: 135–154.
19. Sawtell NB, Williams A (2008) Transformations of electrosensory encoding associated with an adaptive filter. J Neurosci 28: 1598–1612.
20. Bell C, Bodznick D, Montgomery J, J B (1997) The generation and subtraction of sensory expectations within cerebellum-like structures. Brain, Behav and Evol 50: 17–31.
21. Requarth T, Sawtell NB (2011) Neural mechanisms for filtering self-generated sensory signals in cerebellum-like circuits. Curr Op Neurobiol 21: 602–608.
22. Roberts PD, Bell CC (2000) Computational consequences of temporally asymmetric learning rules: II. Sensory image cancellation. J Comput Neurosci 9: 67–83.
23. Harvey-Girard E, Lewis J, Maler L (2010) Burst-induced anti-Hebbian depression acts through short-term synaptic dynamics to cancel redundant sensory signals. J Neurosci 30: 6152–6169.
24. Bol K, Marsat G, Mejias JF, Maler L, Longtin A (2013) Modeling cancelation of periodic inputs with burst-STDP and feedback. Neural Networks. DOI: 10.1016/j.neunet.2012.12.011.
25. Yu N, Hupe G, Garfinkle C, Lewis JE, Longtin A (2012) Coding conspecific identity and motion in the electric sense. PLoS Comput Biol 8: e1002564.
26. Gussin D, Benda J, Maler L (2007) Limits of linear rate coding of dynamic stimuli by electroreceptor afferents. J Neurophysiol 97: 2917–2929.
27. Lapicque L (1097) Recherches quantitatives sur l'excitation électrique des nerfs traitée comme une polarization. J Physiol Pathol Gen 9: 620–635.
28. Gerstner W, Kistler W (2002) Spiking neuron models- Single neurons, population, plasticity. Cambridge University Press.
29. Doiron B, Longtin A, Turner RW, Maler L (2001) Model of gamma frequency burst discharge generated by conditional backpropagation. J Neurophysiol 86: 1523–1545.
30. Noonan L, Doiron B, Laing C, Longtin A, Turner RW (2003) A dynamic dendritic refractory period regulates burst discharge in the electrosensory lobe of weakly electric fish. Neurosci Res 23: 1524–1534.
31. Berman NJ, Maler L (1998) Inhibition evoked from primary afferents in the electrosensory lateral line lobe of the weakly electric fish (Apteronotus leptorhynchus). J Neurophysiol 80: 3173–96.
32. Berman NJ, Maler L (1998) Distal versus proximal inhibitory shaping of feedback excitation in the electrosensory lateral line lobe: implications for sensory filtering. J Neurophysiol 80: 3214–32.
33. Lewis JE, Maler L (2004) Synaptic dynamics on different time scales in a parallel fiber feedback pathway of the weakly electric fish. J Neurophysiol 91: 1064–70.
34. Han V, Grant K, Bell C (2000) Reversible associative depression and nonassociative potentiation at a parallel fiber synapse. Neuron 27: 611–622.
35. Lewis JE, Maler L (2002) Dynamics of electrosensory feedback: short-term plasticity and inhibition in a parallel fiber pathway. J Neurophysiol 88: 1695–706.
36. Chechik G, Anderson MJ, Bar-Yosef O, Young ED, Tishby N, et al. (2006) Reduction of information redundancy in the ascending auditory pathway. Neuron 51: 359–368.
37. Kramer P, Shelhamer M, Peng G, Zee D (1998) Context-specific short-term adaptation of the phase of the vestibulo-ocular reflex. Exp Brain Reseach 120: 184–192.
38. Dittman J, Kreitzer A, Regehr W (2000) Interplay between facilitation, depression, and residual calcium at three presynaptic terminals. J Neurosci 20: 1374–1385.
39. Abbott LF, Valera JA, Sen K, Nelson SB (1997) Synaptic depression and cortical gain control. Science 275: 220–224.
40. Mejias JF, Torres JJ (2009) Maximum memory capacity on neural networks with short-term synaptic depression and facilitation. Neural Comput 21: 851–871.
41. Mejias JF, Hernandez-Gomez B, Torres JJ (2012) Short-term synaptic facilitation improves information retrieval in noisy neural networks. EPL 97: 48008.
42. Chadderton P, Margrie T, Hausser M (2004) Inhibitory feedback required for network burst responses to communication but not to prey stimuli. Nature 428: 856–860.
43. Rancz E, Ishikawa T, Duguid I, Chadderton P, Mahon S, et al. (2007) High-fidelity transmission of sensory information by single cerebellar mossy fibre boutons. Nature 450: 1245–1248.
44. D'Angelo E, De Filippi G, Rossi P, Taglietti V (1998) Ionic mechanism of electroresponsiveness in cerebellar granule cells implicates the action of a persistent sodium current. J Neurophysiol 80: 493–503.
45. D'Angelo E, Nieus T, Maffei A, Armano S, Rossi P, et al. (2001) Theta-frequency bursting and resonance in cerebellar granule cells: experimental evidence and modeling of a slow K- dependent mechanism. J Neurosci 21: 759–770.
46. Jorntell H, Ekerot C (2006) Properties of somatosensory synaptic integration in cerebellar granule cells in vivo. J Neurosci 26: 11786–11797.
47. Brochu G, Maler L, Hawkes R (1990) Zebrin-II - a polypeptide antigen expressed selectively by Purkinje-cells reveals compartments in rat and fish cerebellum. J Comp Neurol 291: 538–552.
48. Maler L, Sas EK, Rogers J (1981) The cytology of the posterior lateral line lobe of high-frequency weakly electric fish (Gymnotidae): dendritic differentiation and synaptic specificity in a simple cortex. J Comp Neurol 195: 87–139.

49. Bastian J, Chacron M, Maler L (2002) Receptive field organization determines pyramidal cell stimulus-encoding capability and spatial stimulus selectivity. J Neurosci 22: 4577–4590.

50. Walz H, Hupe GJ, Benda J, Lewis JE (2013) The neuroethology of electrocommunication: how signal background influences sensory encoding and behaviour in Apteronotus leptorhynchus. J Physiol Paris 107: 13–25.

51. Longtin A, Chialvo D (1998) Stochastic and deterministic resonances in excitable systems. Phys Rev Lett 81: 4012.

52. Hupe G (2012) Electrocommunication in a species of weakly electric fish, Apteronotus leptorhynchus: signal patterning and behaviour. University of Ottawa Press.

53. Marsat G, Maler L (2010) Neural heterogeneity and efficient population codes for communication signals. J Neurophysiol 104: 2543–2555.

54. Mejias JF, Longtin A (2012) Optimal heterogeneity for coding in spiking neural networks. Phys Rev Lett 108: 228102.

55. Sas E, Maler L (1987) The organization of afferent input to the caudal lobe of the cerebellum of the gymnotid fish Apteronotus leptorhynchus. Anat and Embryol 177: 55–79.

56. Bastian J, Nguyenkim J (2001) Dendritic modulation of burst-like firing in sensory neurons. J Neurophysiol 85: 10–22.

57. Marsat G, Proville RD, Maler L (2009) Transient signals trigger synchronous bursts in an identified population of neurons. J Neurophysiol 102: 714–723.

58. Chadderton P, Margrie TW, Hausser M (2004) Integration of quanta in cerebellar granule cells during sensory processing. Nature 428: 856–860.

59. Sawtell NB (2010) Multimodal integration in granule cells as a basis for associative plasticity and sensory prediction in a cerebellum-like circuit. Neuron 66: 573–84.