LETTER **Communicated by Angus Silver**

How Synaptic Release Probability Shapes Neuronal Transmission: Information-Theoretic Analysis in a Cerebellar Granule Cell

Angelo Arleo
angelo.arleo@upmc.fr
CNRS, UPMC, UMR 7102 Neurobiology of Adaptive Processes, F-75005, Paris, France, and Neuroscience Group, SONY Computer Science Laboratory, 75005, Paris, France

Thierry Nieus
thierry.nieus@iit.it
Neuroscience and Brain Technologies, Italian Institute of Technology, 16163 Genova, Italy, and Department of Physiology, University of Pavia, and IRCCS C. Mondino, Pavia 27100, Italy

Michele Bezzi
michele.bezzi@gmail.com
Neuroscience Group, SONY Computer Science Laboratory, 75005, Paris, France

Anna D'Errico
derrico@sissa.it
Department of Physiology, University of Pavia, and IRCCS C. Mondino, 27100, Pavia, Italy

Egidio D'Angelo
dangelo@unipv.it
Department of Physiology, University of Pavia, and IRCCS C. Mondino, 27100, Pavia, Italy

Olivier J.-M. D. Coenen
olivier@oliviercoenen.com
Neuroscience Group, SONY Computer Science Laboratory, 75005, Paris, France

Note: A. A., T. N., and M. B. contributed equally to this work. M Bezzi is now at SAP Research Center, Aophia-Antipolis, France. A. D'Errico is now at Sector of Neurobiology, International School for Advanced Studies, Trieste, Italy. O. Coenon is now at Computational Neuroscience Research Team, Intelligent Systems Research Center, School of Computing and Intelligent System, University of Ulster, Londonderry, BT48 7JL, Northern Ireland.

A nerve cell receives multiple inputs from upstream neurons by way of its synapses. Neuron processing functions are thus influenced by changes in the biophysical properties of the synapse, such as long-term potentiation (LTP) or depression (LTD). This observation has opened new perspectives on the biophysical basis of learning and memory, but its quantitative impact on the information transmission of a neuron remains partially elucidated. One major obstacle is the high dimensionality of the neuronal input-output space, which makes it unfeasible to perform a thorough computational analysis of a neuron with multiple synaptic inputs. In this work, information theory was employed to characterize the information transmission of a cerebellar granule cell over a region of its excitatory input space following synaptic changes. Granule cells have a small dendritic tree (on average, they receive only four mossy fiber afferents), which greatly bounds the input combinatorial space, reducing the complexity of information-theoretic calculations. Numerical simulations and LTP experiments quantified how changes in neurotransmitter release probability (p)-modulated information transmission of a cerebellar granule cell. Numerical simulations showed that p shaped the neurotransmission landscape in unexpected ways. As p increased, the optimality of the information transmission of most stimuli did not increase strictly monotonically; instead it reached a plateau at intermediate p levels. Furthermore, our results showed that the spatiotemporal characteristics of the inputs determine the effect of p on neurotransmission, thus permitting the selection of distinctive preferred stimuli for different p values. These selective mechanisms may have important consequences on the encoding of cerebellar mossy fiber inputs and the plasticity and computation at the next circuit stage, including the parallel fiber–Purkinje cell synapses.

1 Introduction

Theoretically, neurons can be considered as transmitting devices encoding information in terms of digital spike trains. Spikes are transmitted between neurons at the synapses, where they are converted into analog signals by elaborate nonlinear transformations based on the time-dependent properties of neurotransmitter release and diffusion, postsynaptic receptor activation, and intrinsic electroresponsiveness. Synapses undergo plasticity via activity-dependent modifications, such as long-term potentiation (LTP) and depression (LTD) (Bliss & Collingridge, 1993; Bliss & Lomo, 1973; Malenka & Bear, 2004). These modifications affect the synaptic dynamics (e.g., by regulating the number of neurotransmitter quanta released or the phosphorylation of postsynaptic receptors), and ultimately they alter the computational and transmitting properties of the whole postsynaptic neuron (Abbott & Regehr, 2004; Tsodyks & Markram, 1997).

Neuronal computation can be analyzed in terms of information content by quantifying how much information the neural responses convey about the input stimuli (Bialek, Rieke, de Ruyter van Steveninck, & Warland, 1991; Borst & Theunissen, 1999; Fuhrmann, Segev, Markram, & Tsodyks, 2002; Quian Quiroga & Panzeri, 2009). In this framework, neurons are treated as stochastic communication channels, and information theory (Cover & Thomas, 1991; Shannon, 1948) provides the mathematical tools to measure their transmitting properties. Information theory has been used to analyze neuronal computation and quantify the information transmitted by a neuron following sensory stimulation (e.g., in the cat visual cortex: Sharpee et al., 2006; in the fly visual system: Brenner, Bialek, & de Ruyter van Steveninck, 2000; Brenner, Strong, Koberle, Bialek, & de Ruyter van Steveninck, 2000; de Ruyter van Steveninck, Lewen, Strong, Koberle, & Bialek, 1997; in the mammalian auditory system: Lu & Wang, 2004; Smith & Lewicki, 2006; in the cricket cercal sensory system: Dimitrov, Miller, Gedeon, Aldworth, & Parker, 2003; Roddey & Jacobs, 1996; Theunissen, Roddey, Stufflebeam, Clague, & Miller, 1996; Theunissen & Miller, 1991; in the rat somatosensory system: Wan et al., 2004). It has also been used to characterize the relevant regions of a neuron tuning curve in relation to the variability in its sensory encoding properties (Butts & Goldman, 2006) or to characterize the adaptation in receptive fields to visual stimuli (Sharpee et al., 2006). Information theory has been found useful to measure the efficiency of information transmission at a single synapse (de la Rocha, Nevado, & Parga, 2002; Fuhrmann et al., 2002; Goldman, 2004; London, Schreibman, Hausser, Larkum, & Segev, 2002; Manwani, Steinmetz, & Koch, 2002; Tiesinga, 2001), as well as to characterize the information transmission in simplified integrate-and-fire neurons (Manwani et al., 2002; Zador, 1998), or in more complex models, but under specific constraints, for example, in the absence of interaction among presynaptic inputs (Manwani & Koch, 2001) or between inhibitory inputs and phase lags (Tiesinga, Fellous, Jose, & Sejnowski, 2002). In the current work, an information-theoretic approach was used to characterize the processing of a neuron over a region of its excitatory input space and investigate how neuronal processing changes following synaptic plasticity.

The mean information transmitted between the neuron responses r and its inputs s can be quantified by using Shannon's mutual information (MI) (Cover & Thomas, 1991; Rolls & Deco, 2002; Shannon, 1948):

$$MI(R, S) = \sum_{s \in S} \sum_{r \in R} p(s)p(r \mid s) \log_2 \frac{p(rs)}{p(r)}, \tag{1.1}$$

where S and R are the input and output spaces, respectively; $p(s)$ and $p(r)$ are the a priori probability distributions; and $p(r \mid s)$ is the conditional probability distribution.

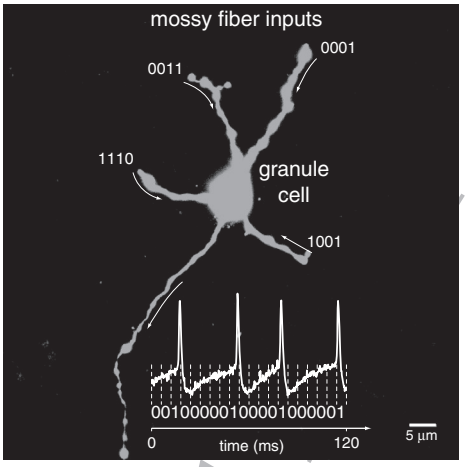


Figure 1: Cerebellar granule cell morphology and spike signal digitalization. Confocal image of a biocytin-stained granule cell (GC) in the rat cerebellum (slice preparation). Granule cells are characterized by a compact electrotonic structure (D'Angelo et al., 1995; Saviane & Silver, 2006) and have an exceptionally low number of synapses (four on average) (Eccles et al., 1967; Jakab & Hamori, 1988) receiving spike trains from the mossy fibers (MFs) (Chadderton et al., 2004). GCs are numerous ($\sim 10^{11}$ in humans) and constitute more than half of the neurons in the brain. White trace: The membrane potential of a GC recorded over 120 ms. Spike trains were digitized as strings of 0s and 1s, within discrete time windows of 6 ms (time bins).

The MI measure can be used to estimate how much the neuron response reflects the input stimuli versus the intrinsic variability of the neuron (e.g., if the response is independent from the inputs, MI will be zero). Thus, MI essentially measures the difference between the signal and noise entropy (Borst & Theunissen, 1999) and provides a sound statistical tool to dissect the relative contributions of different factors to neural information transmission (e.g., spike count versus spike timing) under different conditions (e.g., before and after LTP induction). One major factor that makes it difficult to estimate the MI see equation 1.1, and assess how synaptic plasticity affects neuronal processing is the high dimensionality of the input-output space (Borst & Theunissen, 1999). Indeed, assessing MI requires determining the probability distribution of the output spike trains given any input spike train. In general, this is impracticable even for a single neuron due to the multiple mechanisms of nonlinear integration at individual synapses; the large number of synapses, typically 10^3 to 10^4; and their location on wide dendritic trees with complex electrotonic and active properties (Borst & Theunissen, 1999; Koch & Segev, 2000).

To overcome these issues, the cerebellar granule cell, GC (see Figure 1), was considered. GCs are tiny neurons (6 μm diameter) (Eccles, Ito, &

Szentagothai, 1967; Jakab & Hamori, 1988) located at the major input stage of the cerebellum, the granular layer. GCs play a major role in the early stages of cerebellar computation, and their synapses have been proposed to regulate the input-output relationship through gain modulation (Albus, 1971; Mitchell & Silver, 2003; Rothman, Cathala, Steuber, & Silver, 2009) and long-term adaptation (Hansel, Linden, & D'Angelo, 2001; Philipona & Coenen, 2004; Schweighofer, Doya, & Lay, 2001).

GCs have several remarkable properties. First, they have a compact electrotonic structure (D'Angelo, De Filippi, Rossi, & Taglietti, 1995; Saviane & Silver, 2006; Silver, Traynelis, & Cull-Candy, 1992), which maintains their whole cell membrane equipotential, eliminating spatial effects on computation (Koch & Segev, 2000). Second, they have a low number of mossy fiber (MF) afferents (4.17 on average) (Eccles et al., 1967; Jakab & Hamori, 1988), which generates a tractable number of presynaptic input combinations and greatly reduces the complexity of information-theoretic calculations. Third, they have a stereotyped synaptic and excitable behavior, which simplifies the implementation of models (D'Angelo et al., 1995). Fourth, MFs have been shown to respond with high-frequency bursts to punctuate stimulation (Arenz, Silver, Schaefer, & Margrie, 2008; Chadderton, Margrie, & Hausser, 2004; Jorntell & Ekerot, 2006; Rancz et al., 2007), which have been carefully characterized. Fifth, both GC output bursts and plasticity at MF-GC synapses are controlled by input patterns and Golgi cell inhibition (Mapelli & D'Angelo, 2007). The dynamics of repetitive stimulation have been clarified to a considerable extent (Nieus et al., 2006; Saviane & Silver, 2006; Sola, Prestori, Rossi, Taglietti, & D'Angelo, 2004). Finally, MF-GC synaptic transmission is based on nonlinear transformations determined by presynaptic short-term facilitation and depression, glutamate spillover, postsynaptic AMPA and NMDA receptor gating, and multiple voltage-dependent channel interactions regulating intrinsic electroresponsiveness (D'Angelo et al., 1995; Nielsen, DiGregorio, & Silver, 2004; Nieus, DiGregorio, & Silver, 2006; Sargent, Saviane, Nielsen, DiGregorio, & Silver, 2005; Sola et al., 2004). The mechanisms of synaptic transmission and plasticity at the MF-GC synapses have been intensely investigated, revealing that LTP is largely determined by a raise in the presynaptic neurotransmitter release probability (p) (D'Angelo et al., 2001; Nieus et al., 2006; Saviane & Silver, 2006; Sola et al., 2004).

The study presented here focused on the effects of release probability (p) changes at MF-GC synapses on the overall GC information transmission properties. The same information-theoretic quantification was applied to analyze data from numerical simulations of biophysical synaptic and GC models, as well as data issued from in vitro intracellular GC recordings. At a first level, the mutual information MI between MF inputs and GC responses was measured as a function of release probability, p. Expectedly, MI increased significantly with p, as MI is a measure of the information transmission averaged over the entire input set considered for examination.

A second level of analysis assessed the contribution of specific stimuli to information transmission. For this purpose, the stimulus-specific surprise measure (Butts & Goldman, 2006; DeWeese & Meister, 1999; Theunissen & Miller, 1991) was used to quantify the optimality of single stimulus transmission (see section 4). This analysis investigated which stimulus patterns were "preferred" by the neuron under different release probability conditions, and it demonstrated that maximum synaptic release probability did not constitute a necessary condition in order to achieve optimal transmission. Rather, for a significant set of stimuli, the surprise values saturated at intermediate p values (consistent with those found in brain recordings at MF-GC synapses: Sola et al., 2004; at CA3-CA1 hippocampal synapses: Dobrunz & Stevens, 1997; and at neocortical pyramid-to-pyramid connections: Koester & Johnston, 2005) with optimal information transmission occurring over a large range of release probabilities (from about 0.4 to the maximum value tested, i.e., 0.8). Finally, a third level of analysis concentrated on the spatiotemporal characteristics of the stimuli and quantified, for different p values, the spike timing contribution to information transmission by means of the surprise-per-spike measure (see section 4). The surprise per spike was typically higher for either long-correlated stimuli at low p or short-correlated stimuli at high p.

The work presented in this letter provides a set of tools to investigate neuronal coding and information transfer in the cerebellar granular layer network, which enables a quantitative exploration of the relative importance of the coding strategies for different input patterns and synaptic parameters.

2 Results

A first fundamental question is how the transmitting properties of a GC vary under different release probability conditions (e.g., LTP). To this aim, the average amount of information transmitted by the cell (i.e., the MI) was measured (see section 2.1) computationally with a GC model and experimentally. Experimental results were obtained with a limited set of stimuli; the numerical simulations were used for corroborating the experimental findings and extending them to a larger region of the input space. The question of how changes in release probability shape the neurotransmission of specific MF inputs is addressed in section 2.2 using the stimulus-specific surprise. Finally, section 2.3 investigates how correlations across input spike trains affect single-stimulus transmission and assesses the informative contribution of single pulses in the presence of distinct spatiotemporal stimulus structures and multiple presynaptic release probabilities.

2.1 Impact of Neurotransmitter Release Probability Changes on Information Transmission. The average information transmitted by a single GC was quantified before and after induction of long-term synaptic

plasticity at MF-GC synapses, a condition shown to modify release probability (p) at the MF synaptic terminals (Sola et al., 2004). MI analysis was performed using experimental data obtained by whole-cell patch recordings of GCs during in vitro electrophysiological recordings (see section 4). To measure MI, one to four MFs were stimulated by a set of spike trains, according to a protocol inspired by the bursting discharge of GCs following punctuate tactile stimulation in vivo (Chadderton et al., 2004). Because our experimental techniques did not allow us to stimulate the four MFs independently, this analysis could be done only over a limited input set (16 distinct stimuli, each made of identical spike trains on the four MF inputs). The GCs responded with noisy spike trains that had higher average frequency and occurred earlier after LTP (see Figure 2A). The neurotransmitter release probability p was estimated before and after LTP induction (see Supplementary Material).[1] Because p was the average value over the different synapses, it will henceforth be indicated with $\bar{p}$. LTP caused an average MI increase of $32 \pm 4\%$ for $\bar{p}$ changing by $48 \pm 5\%$ ($n = 9$; paired student's t-test, $p < 0.05$). The average amount of information carried by the firing frequency was 51%, meaning that half of the information transfer was due to interspike temporal relationships.

The same stimulation protocol was employed to run the numerical simulations with the detailed GC model, and MI was measured for different p values at the model MF-GC synapses. As shown in Figure 2B, MI increased as a function of $\bar{p}$ for both experimental and simulated data. The numerical results predicted an increase in MI from 0.2 to 3.5 bits for $\bar{p}$ varying within the range [0.1, 1]. Due to the restricted input space explored (16 stimuli), MI tended to saturate when the number of simultaneously active MFs was greater than 2 and $\bar{p} >= 0.5$ (see below). The vectors representing the experimental MI changes during LTP fell within the limits of the model predictions for all the GCs examined. The similarity between experimental and simulation results indicated that the model could predict the information transfer dynamics following long-term synaptic plasticity. It also supported the hypothesis that the major noise source of GCs is stochastic neurotransmitter release, as anticipated by quantal analysis (Sola et al., 2004). A control analysis focusing on the role of presynaptic depression and postsynaptic receptor desensitization revealed that setting either one or the other to zero affected spike timing only to a small extent (see Supplementary Material, section S2.2), implying minor effects on information transmission.

The model permitted extending numerical simulations to the experimentally impracticable case of independent activation of the four MF afferents. Independent input spike trains (with a maximal frequency of 100 Hz) were generated at each of the four MFs, and all the possible

[1]Supplementary material referred to throughout the letter is available online at http://www.mitpressjournals.org/doi/suppl/10.1162/NECO_a_00006-Arleo.

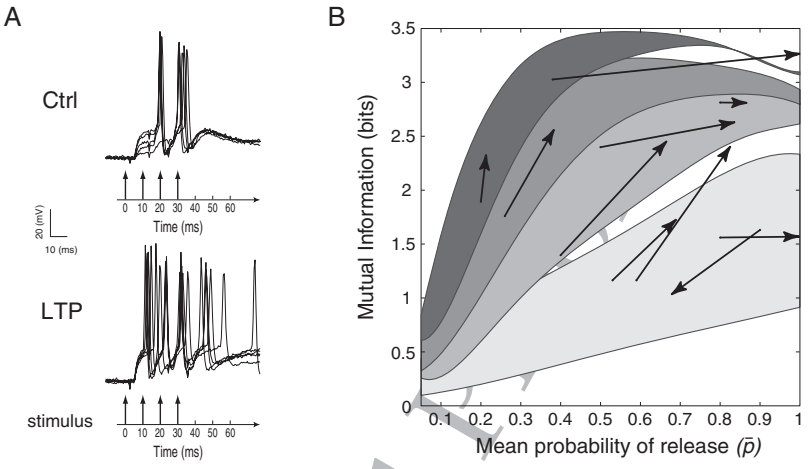


Figure 2: Increases in neurotransmitter release probability enhance mutual information in cerebellar granule cells— experiments and simulations with a restricted mossy fiber input set. (A) A GC was maintained at a membrane potential between −60 mV and −70 mV, and it was activated through MF stimulations at the times indicated by the arrows. This protocol was repeated 25 times. The responses were recorded before and after delivering a theta-burst stimulation (TBS) (Nieus et al., 2006) (five selected voltage traces are shown superposed). In this cell, release probability was $\bar{p} = 0.4$ before and $\bar{p} = 0.65$ after LTP induction (see Supplementary Material for the $\bar{p}$ estimation process). Effective long-term potentiation (LTP) induction is demonstrated by earlier and more intense spike activation (Nieus et al., 2006). (B) Mutual information (MI) changed significantly as a function of neurotransmitter release probability, $\bar{p}$. The shaded regions show the MI values obtained with the GC model, whereas the vectors indicate the experimental changes in MI before and after induction of long-term synaptic plasticity in 10 GCs. The four regions (from bottom to top) correspond to different numbers of active MFs (one to four). The lower and upper borders of each computed region were obtained by setting the resting potential of the GC model at −70 and −60 mV, the two experimental extremes, respectively. LTP, associated with a $\bar{p}$ increase, was observed in nine cases, while in one case with very high initial $\bar{p}$, LTD (long-term depression) was observed associated with a $\bar{p}$ decrease. Note that a 60 ms time window was used to sample both experimental and simulated GC responses.

input combinations were explored (yielding an extensive set of 65536 stimuli; see section 4). Neurotransmitter release probability p was also regulated independently at each MF-GC synapse (from $p = 0.1$ to 0.8, in steps of 0.1). Consequently the information transmission analysis was performed for many different p combinations across the four MFs (e.g., $p_{MF_1} = 0.2$, $p_{MF_2} = 0.8$, $p_{MF_3} = 0.3$, $p_{MF_4} = 0.4$). Figure 3A displays a

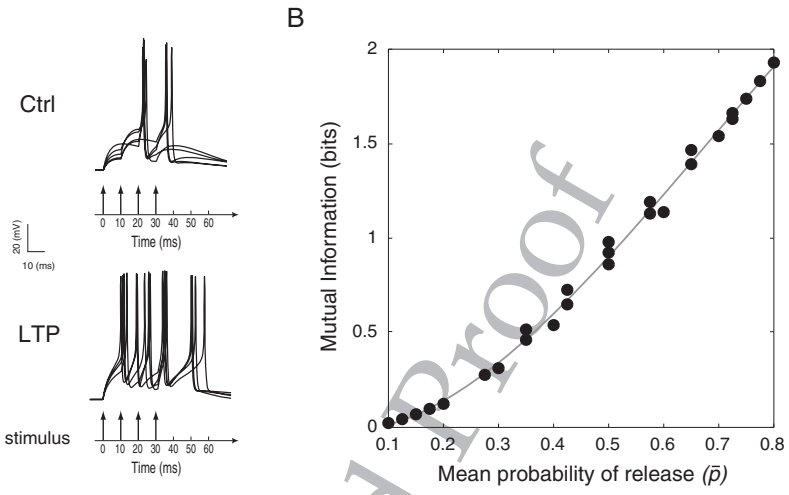


Figure 3: Increases in neurotransmitter release probability enhance mutual information in cerebellar granule cells: simulation results with an extended mossy fiber input set. (A) Responses of the model GC increased their time locking to the start of the input stimulation after a large increase in release probability $\bar{p}$ (five sample traces are shown superposed). (B) MI calculated by stimulating the MF-GC synapses with the extended input set ($k = 2^{16} = 65536$ stimuli) and by varying the release probability p on the four MFs independently, within the range [0.1, 0.8] (Sola et al., 2004). Each data point indicates the MI value corresponding to a different combination of p across the four MFs, and the x-axis provides the $\bar{p}$ averaged over the four MFs (therefore different MI values can coexist for any value of $\bar{p}$). The larger $\bar{p}$ is, the larger is MI. Note that MI shows a smooth increase rather independent from the specific p combination used over the different synapses.

sample response of the model GC to the same stimulus before and after simulated LTP (i.e., for two different p combinations, with $\bar{p} = \langle p_{MF_i} \rangle_{i=1,4}$). Figure 3B shows that the MI computed over the extended stimulus set increased as a function of $\bar{p}$ without saturating (in contrast to the limited set case of Figure 2B), suggesting that the information transmission averaged over a large set of MF stimuli may benefit linearly from increasing $\bar{p}$ values.

2.2 Impact of Release Probability Changes on Single Stimulus Transmission. The surprise measure (see section 4) was used to study the influence of release probability changes on the transmission of specific MF inputs. For each $\bar{p}$ value, all stimuli were ranked according to their surprise value (e.g., in Figure 4A for $\bar{p} = 0.42$), and then different subsets of stimuli were considered. In the subset of stimuli with surprise larger than 90% of the maximum (see Figure 4B), after an initial rapid growth for $0.1 \leq \bar{p} \leq 0.5$,

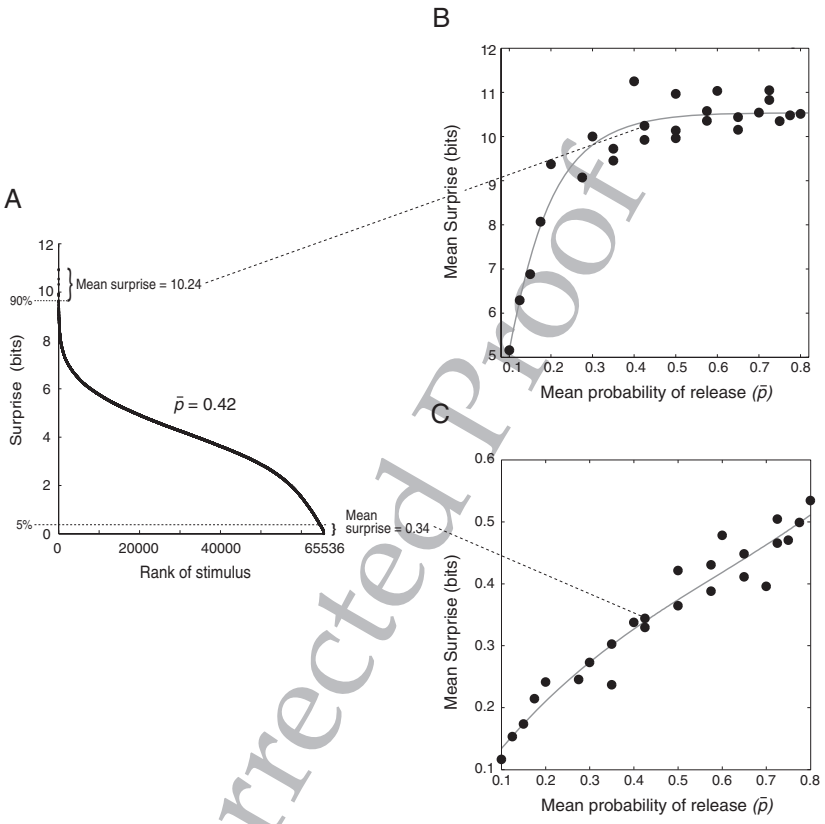


Figure 4: Surprise as a function of neurotransmitter release probability: Simulation results. (A) The surprise measure was used to rank the extensive stimulus set ($k = 2^{16} = 65{,}536$ stimuli), which had p values set independently on the four MFs within the range [0.1,0.8]. The example given here had an average (over the four MF-GC synapses) release probability of $\bar{p} = 0.42$. (B) The mean surprise for the (top 10%) most informative stimuli tended to plateau as $\bar{p}$ increased, even though MI increased with larger $\bar{p}$ values (see Figure 3B). For a particular set of p values, the mean surprise was obtained by averaging over the subset of stimuli with a surprise larger than 90% of the maximum. Note that the subset of stimuli contributing to the mean surprise may change with different set of p values. (C) The mean surprise for the least informative stimuli—stimuli with a surprise less than 5% of the maximum—was the only subset for which the surprise was monotonously increasing with $\bar{p}$. See also Supplementary Material, figure S15 and section S2.4.

the average surprise tended to plateau around 10.5 bits. This saturation effect was observed for most of the stimuli (see Supplementary Material, Figure S15 and section S2.4). Only the surprise averaged over the subset of stimuli having a surprise smaller than 5% of the maximum showed a quasi-linear increase with $\bar{p}$ (see Figure 4C), suggesting that only these stimuli continued to benefit from increasing $\bar{p}$ values. Although on average the information transmitted by the GC increased monotonically with $\bar{p}$ (see Figure 3B), optimal single-stimulus transmission had already occurred at intermediate release probability values for most MF inputs, and further increases in $\bar{p}$ did not result in larger surprise values.

This finding suggests that transmission of specific inputs was differentially affected by changes in presynaptic release probability. To begin characterizing the properties of stimuli with optimal transmission at different $\bar{p}$, simple features such as low versus high firing rate were first considered. Low-rate stimuli tended to benefit the most from $\bar{p}$ increases (see Figure S16). The hypothesis that high-frequency stimuli (as opposed to low-rate stimuli) had less to gain from an increasing release probability $\bar{p}$ was confirmed by quantitative analysis (see Figure S17). But how does the spatiotemporal structure of MF inputs affect neurotransmission in GCs?

2.3 Influence of MF Input Correlations and Spike Timing on Information Transmission. In order to reveal the effect of input correlations on information transmission, the relationship between the surprise of a stimulus and the correlation C across the four MF spike trains forming the stimulus was analyzed for each release probability, $\bar{p}$ (similar to Butts & Goldman, 2006). The coefficient C measured the average number of coincident input spikes across the four MFs (see section 4). The surprise tended to increase as a function of C for all $\bar{p}$ values, showing that correlated activity across MFs improves GC neurotransmission (see Figure 5). Moreover, these findings supplement the observations that GCs operate as coincidence detectors (Moreno-Bote & Parga, 2004) requiring the coactivation of two or more MFs (D'Angelo et al., 1995; Jorntell & Eckerot, 2006; Eccles et al., 1967).

In order to discard the influence of firing rate on information transmission and focus on the contribution of single spikes, the surprise per spike was employed (see section 4). The entire MF stimulus set was ranked according to the surprise-per-spike value for all release probabilities, $\bar{p}$ (see Figure 6A). As shown in Figure 6B.B1, the stimuli with the largest surprise per spike at low $\bar{p}$ were characterized by long-correlated MF spike trains (the blue-labeled patterns), whereas as $\bar{p}$ increased, the stimuli with greatest surprise per spike were short correlated trains (red-labeled stimuli). As shown in Figure 6B.B2, the stimuli with lowest surprise per spike at small $\bar{p}$ did not have a stereotyped structure, whereas as $\bar{p}$ increased, they rapidly became stimuli with at most one spike per bin and therefore no coactivated spikes.

These findings were corroborated by analyzing how the surprise per spike changed as a function of the correlation C for different release

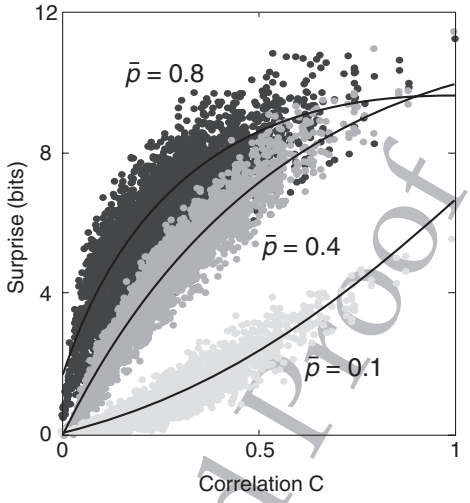


Figure 5: Surprise as a function of the correlation across the four mossy fiber inputs: simulation results. The larger the correlation coefficient C of the stimulus became, the larger was the information transmitted by the stimulus. C measures the average number of coincident spikes across the four MF afferents and over the four time bins (see section 4). This result holds for all release probability values $\bar{p} \in [0.1,0.8]$. Nevertheless, the larger the $\bar{p}$ value was, the more the surprise saturated with higher C values.

probability values (see Figure 6C). At low $\bar{p}$, the surprise per spike increased quasi-linearly with the MF spike correlation C, with the peak of the distribution occurring for long-correlated inputs (blue-labeled stimuli in Figure 6C, left). By contrast, as $\bar{p}$ increased, the distribution of the surprise per spike as a function of C became non linear, and the peak of the distribution increased and occurred for shorter correlated stimuli (red dots in Figure 6C, center and right). Thus, an increase in $\bar{p}$ enhanced the surprise per spike of short correlated stimuli, whereas it favored only partially that of long highly correlated stimuli, whose surprise per spike saturated and even decreased at high $\bar{p}$.

2.4 Influence of Specific Patterns on Information Transmission. Why did long correlated stimuli benefit less, in terms of information per spike, from $\bar{p}$ increases than short correlated ones? A possible explanation emerged by examining at different $\bar{p}$ values the GC responses to distinct MF stimuli. Figure 7A shows both the spikegram and the poststimulus time histogram (PSTH) of the GC response to a long correlated stimulus (with four bins with coactivated spikes). For $\bar{p} > 0.5$, spike doublets were elicited by the second set of input spikes (second arrow, $\bar{p} = 0.6$ and $\bar{p} = 0.8$). The

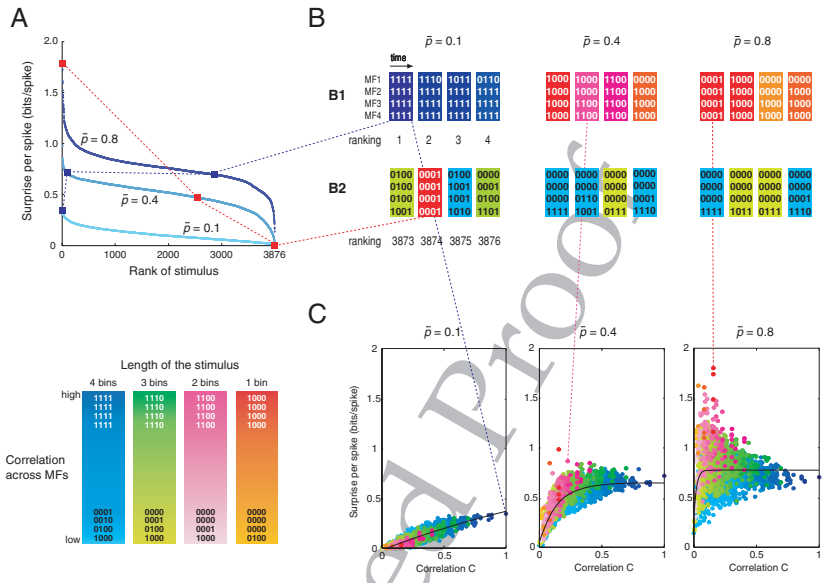


Figure 6: Surprise per spike as a function of release probability: simulation results. (A) The stimuli were ranked as a function of their surprise per spike for every $\bar{p}$ value [0.1, 0.8, step of 0.1]. Here, the release probability was equal for all four MF synapses at the GC; that is, any permutation of the four input spike trains was equivalent. This reduced the number of different stimuli to 3876 from the initial $2^{16} = 65{,}536$. (B) The sets providing the largest (B1) and the smallest (B2) contribution to the surprise per spike were selected for three different $\bar{p}$ values. A different color map was chosen to identify the number of bins occupied (blue, green, pink, red for 4, 3, 2, 1 bins, respectively). The number of spikes per bin modulated the color map (e.g., red to orange for four to one spikes per bin, within the red color map). At low $\bar{p}$, the stimuli with the largest surprise per spike are long-correlated spike trains: four bins with coactivated spikes at $\bar{p} = 0.1$ (e.g., the blue-labeled stimuli in B1). As $\bar{p}$ increases, the stimuli with largest surprise per spike are short correlated trains with two or three bins of coactivated spikes at $\bar{p} = 0.2$–0.3, one to two bins at $\bar{p} = 0.4$–0.5 (red and pink-labeled stimuli), and one bin for $\bar{p} > 0.5$ (red and orange-labeled stimuli). The stimuli with the smallest surprise per spike have no stereotyped structure at low $\bar{p}$, whereas they are long noncorrelated patterns at high $\bar{p}$ (e.g., the cyan-labeled stimuli in B2). (C) Whereas the surprise per spike increased almost linearly with correlation at the lowest $\bar{p}$ value ($\bar{p} = 0.1$), the peak of the distribution increased and moved back as $\bar{p}$ increases: starting with the highest correlated stimuli at $\bar{p} = 0.1$ with four coactivated spikes in all four bins (blue dots), the peak at $\bar{p} = 0.8$ is found for a stimulus with a smaller correlation value, with four coactivated spikes in only one bin (red dot and red-labeled stimuli). The distribution of points at $\bar{p} = 0.4$ already shows saturation for the long correlated stimuli while for the shorter correlated stimuli, the distribution starts to peak.

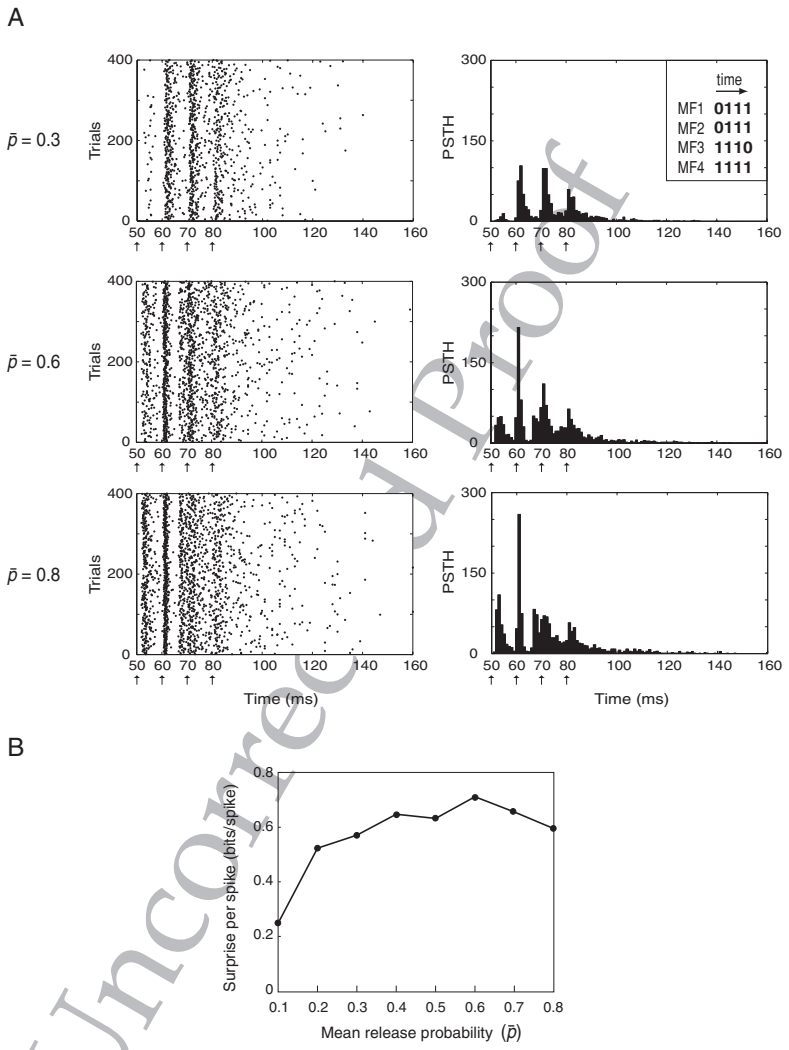


Figure 7: Spike doublet appearance in GC responses in the presence of high release probability: Simulation results. (A) GC response (left: spikegram; right: poststimulus time histogram, PSTH) to a long correlated stimulus at three different $\bar{p}$ values. The stimulus structure is shown in the top-right inset, and arrows in the time line of the diagrams indicate the input spike timing. For $\bar{p} = 0.6$ and 0.8, the second set of spikes (second arrow) elicited a spike doublet, made of one spike occurring at ∼61 ms, followed by another spike at ∼67 ms (before the occurrence of the next set of input spikes at 70 ms). Note the double-peaked PSTH around 70 ms. (B) Surprise per spike of the long correlated stimulus used in *A* as a function of release probability.

doublets disrupted the precise spike timing of the response (evident in the PSTH after the second response spike), limiting the surprise per spike of the stimulus despite the improved timing of the first response spike with respect to the $\bar{p} < 0.5$ case (see Figure 7B). The longer the stimulus (i.e., the larger the number of bins with spikes), the higher was the probability of eliciting spike doublet responses at high $\bar{p}$ values (see Figure 8C for other spike doublet examples). This behavior was probably determined by the engagement of smoothly varying currents, such as the NMDA current and the persistent and resurgent Na^+ currents (D'Angelo et al., 2001; Magistretti, Castelli, Forti, & D'Angelo, 2006; Nieus et al., 2006), with increasing $\bar{p}$. These currents drive repetitive firing and decrease the probability of having precise stimulus-locked response spikes.

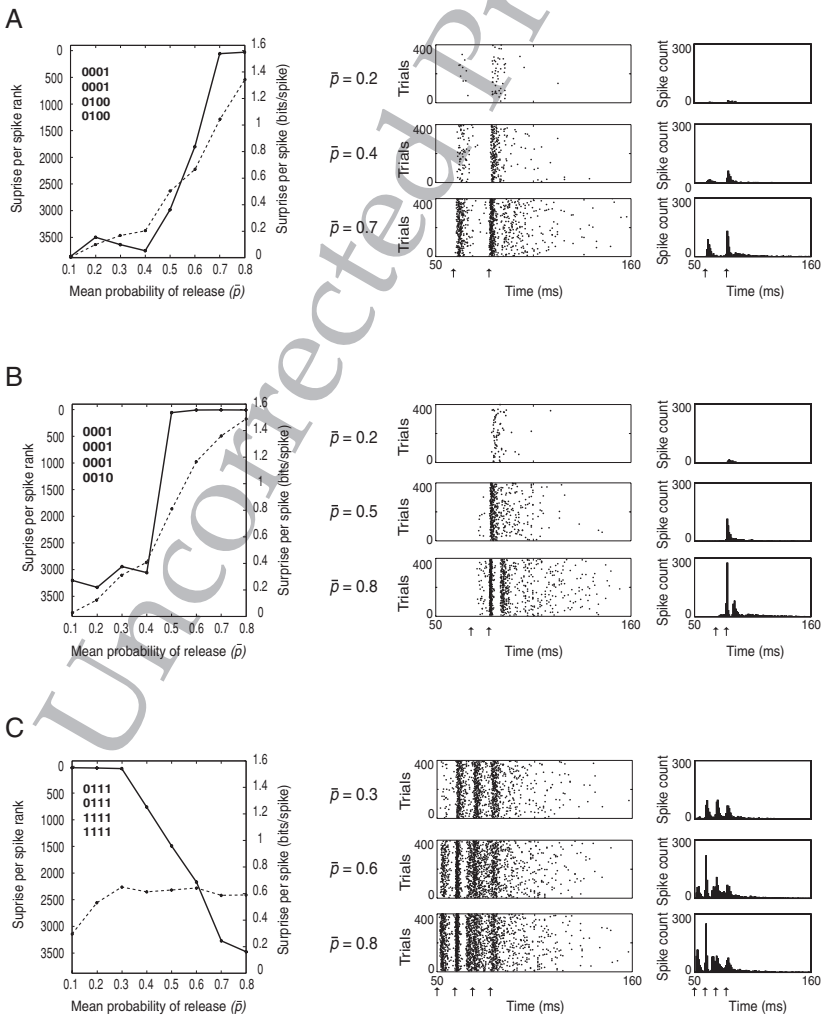
Which stimuli were the most sensitive to release probability changes? An answer was obtained by considering the stimuli showing the largest and the steepest ranking changes in their surprise-per-spike values following release probability modifications. The largest surprise-per-spike increases were observed for short stimuli with at least two coactivated spikes in one bin (see Figure 8A and Supplementary Material, Figure S6). A similar spatiotemporal structure was typically observed for the steepest, and therefore most rapid, ranking transitions in surprise per spike (see Figure 8B and Supplementary Material, Figure S7). For these short stimuli, their surprise per spike increased significantly following small plasticity changes ($\Delta\bar{p} = 0.1$). As shown by the spikegrams and the PSTHs of Figures 8A and 8B, the stimuli with the largest and steepest ranking increases of surprise per spike produced little activity at low $\bar{p}$, whereas they elicited better time-locked responses at high $\bar{p}$. Finally, the largest and the steepest surprise per spike ranking decreases were observed for long and highly correlated stimuli (for which most of the bins were filled with three or four coactivated spikes, Figure 8C and Supplementary Material, Figure S8).

3 Discussion

This letter provides an analysis of information transmission over a region of the excitatory input space of a neuron. The direct method to compute mutual information, MI (Zador, 1998), adopted in this work is usually computationally impracticable, and it was made possible by the peculiar structure of the studied system, the cerebellar granule cell (GC) (see Figure 1). Indeed, the fact that GCs are electronically compact and receive on average only four mossy fiber (MF) excitatory inputs reduces the state-space dimensionality drastically. Still, the MF-GC system is representative because MF-GC synapses call on the same complex mechanisms mediating information processing at most brain synapses (D'Angelo et al., 1995; Nielsen et al., 2004; Nieus et al., 2006; Sola et al., 2004).

The information-theoretic study presented here explored the relative contribution of various factors (e.g., spike timing, specific input patterns)

to neurotransmission and investigated how their interrelations changed following long-term plasticity. The analysis showed that the MF-GC relay transmitted information via firing rate, spike timing, and spike correlation of MF discharge. Along with its ability to regulate spike frequency and timing (Nieus et al., 2006), the increase of release probability (p) accompanying LTP regulated the amount of transmitted information (Zador, 1998). The main observation is that, on average, information was carried almost equally by spike frequency and spike timing, and it increased quasi-linearly with p. On the other hand, it was shown that maximizing release



probability is not a necessary condition to optimize stimulus-specific information for most stimuli. Indeed, single-stimulus transmission reached a plateau at intermediate p values for most MF inputs, with optimal transmission taking place over a large p range (from intermediate values to the maximum). Interestingly, in this range, the preferred stimulus patterns were very sensitive to p variations; in other words, small variations of p changed the subset of stimuli that are best transmitted by the cell. Such intermediate release probability values reflect the values observed experimentally in the cerebellar granular layer (Saviane & Silver, 2006; Sola et al., 2004), as well as in other brain areas (e.g., neocortical pyramid-to-pyramid connections: Koester & Johnston, 2005; CA3-CA1 hippocampal synapses: Dobrunz & Stevens, 1997).

Figure 8: Surprise-per-spike changes for exemplar stimuli with highest information transmission modifications following release probability increases: Simulation results. (A–C) Information transmission for stimuli with the fourth largest increase (A), the fourth steepest increase (B), and the fourth largest decrease (C) in the rank of the surprise per spike as a function of $\bar{p}$ (left panels: the continuous lines indicate the changes in the rank, whereas the dashed curves denote the absolute values of the surprise per spike). The middle and right panels show the GC responses to the stimulus as spikegrams and PSTHs, respectively, for three different $\bar{p}$. The steepest increases and decreases typically occurred between 0.4–0.6 and 0.3–0.4, respectively. In *C*, at p = 0.6 and 0.8, note the spike doublets in response to the second set of input spikes at ∼62 and ∼68 ms before the third set of input spikes at 70 ms. This doublet perturbed the timing of the following spikes as seen by the wider spread in the PSTH. This spread is largely responsible for the decline in rank for this stimulus at high $\bar{p}$. In *B*, note the persistent high rank for this stimulus after the rank increase at $\bar{p} = 0.5$ (left panel). A doublet to the three-spike bin (second arrow) starts appearing at $\bar{p} = 0.6$ (not shown) and is clear at $\bar{p} = 0.8$ with spikes at ∼82 and ∼88 ms in response to the second set of input spikes (middle and right panels). Hence, the doublet appearance at $\bar{p} = 0.6$ is not the cause of the rank increase at $\bar{p} = 0.5$ or of a transmission decrease as for the stimuli in *C* since its rank remains high beyond $\bar{p} = 0.6$. Thus, when *B* is compared with *A* and *C*, the effect of the appearance of doublets on surprise per spike is highly stimulus dependent. For the left panels, the 3876 different stimuli obtained with the same p over the four synapses ($\bar{p} = p$) were ranked according to the surprise per spike obtained at a specific p between 0.1 and 0.8. The 10 largest and steepest increases and decreases in rank were analyzed. In general, it was found that in short notation (number of spikes per bin with implicit permutation), the largest increases were for stimuli of the form x2xx or x202 with x = 0, 1; whereas the steepest increases were for x2xx or x3xx, x = 0, 1, 2; and the largest decreases were for xyy4, x = 1,2,3,4; y = 3,4.

In this study, the influence of the Golgi cells (GoCs) was omitted to focus on the GC encoding of its excitatory inputs only. In fact, the series of analyses presented here did not aim at providing a comprehensive characterization of the information transmission process in the granular layer; rather, it was meant to set forth an information-theoretic framework suitable for studying the efficacy of GC information processing quantitatively. At this stage, GoC inhibition was not included because a thorough understanding of the synaptic transmission at the MF-GoC relay is still missing (Forti, Cesana, Isope, Dieudonné, & D'Angelo, 2008). In addition, the study of quantal release properties of GoC-GC synapses has been addressed only very recently (Mapelli, Rossi, Nieus, & D'Angelo, 2009). Further analyses will be carried out to quantify the effects of Golgi inhibition onto GC neurotransmission by means of a cerebellar GABA synaptic model accounting for recent experimental evidence (Mapelli et al., 2009) at this synapse (see also Bezzi, Nieus, Arleo, D'Angelo, & Coenon, 2004, for preliminary work in this direction). GoCs fire autorhythmically at 2 to 4 Hz in vitro (Forti, Cesana, Mapelli, & D'Angelo, 2006) and at around 10 Hz in vivo (Vos, Wijnants, Taeymans, & de Schutter, 1999), and their discharge frequency can be raised up to 300 Hz on stimulation. It is worth mentioning that even under such strong inhibitory conditions, GCs are still able to display short-burst activity (D'Angelo & De Zeeuw, 2009). Following stimulation of the whisker pad, GoCs in anesthetized rats follow a stereotyped firing pattern characterized by bursts of two or three spikes, followed by pauses of around 100 ms. Moreover, the cerebellum displays oscillatory activity at theta frequency (Hartmann & Bower, 1998), and a cerebellar network model suggested that these oscillations are characterized by a highly synchronous activity of GoCs and GCs (Maex & Schutter, 1998). In all these studies, the GoCs appear to regulate the narrow time window during which GCs may fire (D'Angelo & De Zeeuw, 2009; Kistler & De Zeeuw, 2003; Solinas et al., 2007a, 2007b).

The synaptic model used here was stochastic (see section 4; see also Sun, Lyons, & Dobrunz, 2005). It presented both facilitation and depression depending on the release probability p and spike activity (Nieus et al., 2006; Tsodyks & Markram, 1997). Hence, the calculations of information presented here extend those that have thus far been investigated using depressing synapse (de la Rocha & Parga, 2005; Goldman, 2004), and constant-transmission-probability models (Manwani & Koch, 2001; Zador, 1998). Recently an information-theoretic measure, namely, the synaptic information efficacy (SIE) (London et al., 2002), has been introduced to quantify the transmitting properties of a single synapse. SIE is actually the mutual information (MI) between the presynaptic input spike trains and the postsynaptic responses. In this study, MI was measured across only a single synapse, while considering the rest of the dendritic input as background noise (London et al., 2002). Similarly, other studies using different analytical approaches have focused on the transmitting properties of single synapses

(de la Rocha et al., 2002; Goldman, 2004; Manwani et al., 2002; Tiesinga, 2001). In addition, although input correlations have been shown to be relevant to neuronal computation (de la Rocha & Parga, 2005; Schreiber, Fellous, Tiesinga, & Sejnowski, 2004; Tiesinga & Toups, 2005), many studies have investigated conditions in which they are absent (London et al., 2002; Manwani & Koch, 2001; Zador, 1998). This study complements these previous works by presenting a quantitative characterization of how the correlation across multiple synaptic inputs of a neuron may contribute to information transmission.

Our results corroborate the hypothesis that the correlation among the inputs plays a crucial role in affecting neurotransmission at all levels of release probability p. Our analysis strengthened this observation by showing how the spatiotemporal structure of MF inputs affected GC neurotransmission as a function of p. The most sensitive stimuli to release probability changes were identified in terms of both surprise and surprise per spike. Short correlated MF bursts resulted to benefit the most from p increases. Long correlated stimuli caused changes in neurotransmission and excitation dynamics at high p, which bounded their transmission reliability. These results are consistent with previous findings suggesting that when the synapses are tuned toward specific input stimuli, one of the roles of LTP and LTD may be that of generating spike train–specific nonlinear detectors (Natschlager & Maass, 2001; Sharpee et al., 2006), which would regulate the transmission of specific spatiotemporal input patterns at the level of the neuron. The preferential recoding of the information contained in certain patterns may be of biological relevance for computation at subsequent stages in the cerebellar cortex. For instance, two spikes in close succession lead to the opening of presynaptic NMDA channels and start a cascade of events leading to long-term plasticity at the synapses between the Purkinje cells (PCs) (Casado, Isope, & Ascher, 2002) and the parallel fibers (PFs), which are the axons of the GCs. Recoding following granular layer plasticity could also have the effect of reducing the length of PF spike trains, thereby increasing the timing accuracy of the PC responses.

Naturally, to provide definite answers beyond the neuronal level, coding strategies need to be evaluated within a cerebellar network (Coenen, Arnold, Sejnowski, & Jabri, 2001; Philipona & Coenen, 2004) by means of large-scale simulations, where plasticity may also regulate the average GC population firing and the duration of multiple PF activation necessary to elicit a PC response. Therefore, the information-theoretic approach presented here constitutes a step forward in the investigation of neural information transfer in the granular layer network of the cerebellum. Scaling from elementary cellular mechanisms such as synaptic release probability to network computation is fundamental to understanding how MF-GC long-term plasticity, by being instrumental in the control of information transmission, may regulate the operations to which the cerebellum participates (Dum, Li, & Strick, 2002).

4 Materials and Methods

4.1 Biophysical Model of the MF-GC Synaptic Relay. A stochastic version of our previous models (D'Angelo et al., 2001; Nieus et al., 2006) was developed for this study (see also Sun et al., 2005). The GC model was provided with four independent MF-GC synaptic contacts endowed with stochastic neurotransmitter release mechanisms. Each synaptic contact consisted of three independent releasing sites (RS) (Saviane & Silver, 2006; Sola et al., 2004), each governed by a three-state presynaptic model (Tsodyks & Markram, 1997). Neurotransmitter release was modeled as a system of four first-order differential equations relating the probability of release (p) to the available (X), released (Y), and recovered (Z) neurotransmitter resources (Nieus et al., 2006; Tsodyks & Markram, 1997) (see Supplementary Material). The release was made probabilistic and modeled as an all-or-none process by comparing a random number (ε) drawn from a uniform [0, 1] probability distribution with the released resources Y. Hence, a wave of glutamate (Nieus et al., 2006) was released from the RS whenever ε was less than Y, so that averaging over numerous independent Ys led to recovery of the behavior of the deterministic release model (Nieus et al., 2006; Saviane & Silver, 2006; Sola et al., 2004). Each RS activated an independent postsynaptic site endowed with AMPA and NMDA receptors. Because three RS/contact (Saviane & Silver, 2006) were introduced, each postsynaptic site was calibrated to contribute one-third of the total conductance (Nieus et al., 2006). Since no evidence for p heterogeneity was reported (Sola et al., 2004), p was set at the same value for all releasing sites at the same synapse. (Further details on the biophysical model can be found in Supplementary Material, section S1.1.)

First, a series of simulations was run by adopting the same stimulation protocol used for the patch clamp experiments (see section 2 and Figure 2). Then the input space was extended by considering the four MF afferents as four independent spike trains and regulating the neurotransmitter release probability p at each MF-GC synapse independently (e.g., see Figure 3). All possible combinations of input spike trains lasting up to 40 ms, as binary words of 10 ms bins (i.e. with a maximum rate of 100 Hz), were then created. Thus, each MF spike train was encoded as a four-bit binary word, and the cell input was a 4 (number of MFs) $\times$ 4 (number of bits for each MF) $=$ 16 bit binary word. Accordingly, the entire input space consisted of $2^{16} = 65536$ stimuli. GC responses were digitalized using temporal bins of 6 ms over a period of 120 ms (see Supplementary Material, section S2.1, for a discussion on the dependence of the results upon the bin size used to sample GC responses). Each stimulus was presented 400 times, and the effect of limited sampling on MI computation was taken into account (see Supplementary Material, section S1.3).

To simulate natural firing activity in MFs, all 65,536 stimuli occurring with different probabilities, $p(s)$, were considered. The actual firing

distribution of MFs is not known, but following Treves, Panzeri, Rolls, Booth, and Wakeman (1999), a continuous unimodal distribution (with a single peak close to the spontaneous activity, i.e., close to zero in our case) with an exponential tail was assumed. This distribution has been observed in other brain areas, among them, the frontal cortex (Abeles, Vaadia, & Bergman, 1990), hippocampus and close structures (Barnes, McNaughton, Mizumori, Leonard, & Lin, 1990), visual cortex (Baddeley et al., 1997; Franco, Rolls, Aggelopoulos, & Jerez, 2007), as well as used in previous neural network models (Treves & Rolls, 1991). For this study, a decreasing exponential distribution of firing rate with a mean of 10 Hz was chosen. Another distribution with similar features (i.e., Poisson) was also tested, and compatible results were obtained (see Supplementary Material, section S2.3).

The model was implemented with NEURON 5.4. The most demanding numerical simulations were run on a cluster of 20 CPUs (1.7 GHz each).

4.2 Electrophysiological Recordings. Whole-cell patch clamp recordings of GCs were performed from acute cerebellar slices of P18-P23 Wistar rats according to published procedures (D'Angelo et al., 1995; Sola et al., 2004) (see Supplementary Material, section S1.2). Varying the intensity of MF stimulation permitted generating simultaneous activity in one to four MFs. The core experiment was carried out in current-clamp mode. MFs were stimulated by a set of spike trains lasting 40 ms and with a frequency up to 100 Hz. To investigate the effect of long-term synaptic plasticity on MI, the same set of stimuli was presented before and after LTP induction by theta-burst stimulation (TBS: 8 100 Hz bursts lasting 100 ms every 250 ms) (Nieus et al., 2006; Sola et al., 2004) (see Supplementary Material, Figure S2). The mean MF-GC release probability $\bar{p}$ was estimated in voltage clamp, at the beginning and at the end of each experiment, by analyzing the excitatory postsynaptic currents (EPSCs) elicited by four stimuli at 100 Hz (Nieus et al., 2006; see Supplementary Material, Figure S3).

GC responses were spike-sorted via a threshold crossing procedure (Igor; WaveMetrics Inc.). To compute MI, both input (MF) and output (GC) spike trains were transformed into binary words (see Figure 1). MF spike trains were digitized by using a temporal bin of 10 ms. Because each stimulus lasted 40 ms and all MFs received the same spike train, the input set consisted of $2^4 = 16$ stimuli. Each stimulus was presented 25 times. The GC responses were digitized by using a 6 ms bin width. A 60 ms time window was used for sampling the GC output; the spikes occurring beyond this time window (less than 1% of all spikes) were not taken into account to reduce the dimensionality of the sampling space. The significance of MI measurements was assessed by a bootstrapping procedure (Lu & Wang, 2004) of 1000 repetitions). MI proved to be statistically significant ($p < 0.05$) for MI > 0.4 bits.

All experiments were carried out according to the guidelines and regulations laid down by the institution's animal welfare committee.

4.3 Theoretical Analysis

4.3.1 Mutual Information and "Surprise." The mean information transmitted between the GC responses r and the MF inputs s was calculated by using Shannon's mutual information (MI), equation 1.1, where both the stimuli s and the responses r were represented as either binary words (see Figure 1) or spike counts. The binary word coding preserves the information about spike timing, up to a certain precision, whereas the spike count coding simply assesses the information transmitted by the neuron average firing. The data processing inequality (Cover & Thomas, 1991) ensures that MI obtained using spike counts is always not greater than the MI obtained using binary words.

To isolate the specific contribution of a single stimulus s to MI, the stimulus specific surprise (called simply *surprise* here) was used (Butts & Goldman, 2006; DeWeese & Meister, 1999):

$$I_1(s) = \sum_{r \in R} p(r \mid s) \log_2 \frac{p(r \mid s)}{p(r)}. \tag{4.1}$$

The surprise measures how much the conditional distribution $p(r \mid s)$ differs from the prior probability distribution $p(r)$, corresponding to the so-called Kullback-Leibler distance. The surprise per spike was then computed by dividing the surprise by the spike count of the input stimulus. Normalizing by the number of input spikes essentially removes the linear part of the contribution of MF firing rates to information transmission.

Both MI and the surprise measure have a systematic bias due to the limited data samples available (Nemenman, Bialek, & de Ruyter van Steveninck, 2004; Paninski, 2003; Strong, Koberle, de Ruyter van Steveninck, & Bialek, 1998; Treves & Panzeri, 1995). This issue is treated in the Supplementary Material, section S1.3.

4.3.2 Correlation Measure. To estimate the number of coincident spikes across the four MF inputs, the normalized average of the pair-wise correlation over all the different pairs of MFs was computed. Let $s \in S$ be a stimulus constituted by four MF patterns, and let x_i and x_j denote the input spike trains representing the activity on the ith and the jth MF, respectively. The correlation coefficient $C(s)$ was taken as

$$C(s) = \frac{1}{N} \cdot \langle (K(x_i) - \bar{x}) \cdot (K(x_j) - \bar{x})' \rangle_{i>j},$$

where $\bar{x}$ represents the average firing rate, N is the normalization factor equal to the number of input bins, and $K(x)$ denotes a smoothing function (i.e., an exponential kernel with τ equal to the input bin width) that takes into account the effects of the short-time integration over the membrane

time constant. The $C(s)$ vector, for all $s \in S$, was then rescaled into the range [0, 1]. For example, $C(s) = 0.4039$ for $s = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}$, whereas $C(s) = 0.1561$ for $s = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$.

Acknowledgments

We kindly acknowledge the contribution of David Gall (Université Libre de Bruxelles) for the confocal image in Figure 1. This work was supported by European projects SENSOPAC IST-2005–028056 (www.sensopac.org) and SpikeFORCE IST-2001-35271 (www.spikeforce.org), both created by O.J.-M.D.C., by Sony CSL-Paris, and by the Italian MIUR project PRIN-2004053317 and CNR-INFM project FIRB-RBNE01AAS7 to E.D'A.

References

Abbott, L. F., & Regehr, W. G. (2004). Synaptic computation. *Nature*, *431*, 796–803.

Abeles, M., Vaadia, E., & Bergman, H. (1990). Firing patterns of single units in the prefrontal cortex and neural network models. *Network*, *1*, 13–25.

Albus, J. S. (1971). A theory of cerebellar function. *Math. Biosc.*, *10*, 25–61.

Arenz, A., Silver, R. A., Schaefer, A. T., & Margrie, T. W. (2008). The contribution of single synapses to sensory representation in vivo. *Science*, *321*, 977–980.

Baddeley, R., Abbott, L. F., Booth, M. C., Sengpiel, F., Freeman, T., Wakeman, E. A., et al. (1997). Responses of neurons in primary and inferior temporal visual cortices to natural scenes. *Proc. Biol. Sci.*, *264*, 1775–1783.

Barnes, C. A., McNaughton, B. L., Mizumori, S. J., Leonard, B. W., & Lin, L. H. (1990). Comparison of spatial and temporal characteristics of neuronal activity in sequential stages of hippocampal processing. *Prog. Brain Res.*, *83*, 287–300.

Bezzi, M., Nieus, T., Arleo, A., D'Angelo, E., & Coenon, O. J. M. (2004). *Information transfer at the mossy fiber-granule cell synapse of the cerebellum.* Society for Neuroscience Abstracts No. 827.5, San Diego.

Bialek, W., Rieke, F., de Ruyter van Steveninck, R. R., & Warland, D. (1991). Reading a neural code. *Science*, *252*, 1854–1857.

Bliss, T. V., & Collingridge, G. L. (1993). A synaptic model of memory: Long-term potentiation in the hippocampus. *Nature*, *361*, 31–39.

Bliss, T. V., & Lomo, T. (1973). Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *J. Physiol.*, *232*, 331–356.

Borst, A., & Theunissen, F. E. (1999). Information theory and neural coding. *Nat. Neurosci.*, *2*, 947–957.

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

Brenner, N., Strong, S. P., Koberle, R., Bialek, W., & de Ruyter van Steveninck, R. R. (2000). Synergy in a neural code. *Neural. Comput.*, *12*, 1531–1552.

Butts, D. A., & Goldman, M. S. (2006). Tuning curves, neuronal variability, and sensory coding. *PLoS Biol.*, *4*, e92.

Casado, M., Isope, P., & Ascher, P. (2002). Involvement of presynaptic N-methyl-D-aspartate receptors in cerebellar long-term depression. *Neuron*, *33*, 123–130.

Chadderton, P., Margrie, T. W., & Hausser, M. (2004). Integration of quanta in cerebellar granule cells during sensory processing. *Nature*, *428*, 856–860.

Coenen, O. J.-M. D., Arnold, M. P., Sejnowski, T. J., & Jabri, M. A. (2001). Parallel fiber coding in the cerebellum for life-long learning. *Autonomous Robots*, *11*, 291–297.

Cover, T. M., & Thomas, J. A. (1991). *Elements of information theory.* New York: Wiley.

D'Angelo, E., De Filippi, G., Rossi, P., & Taglietti, V. (1995). Synaptic excitation of individual rat cerebellar granule cells in situ: Evidence for the role of NMDA receptors. *J. Physiol.*, *484*, 397–413.

D'Angelo, E., & De Zeeuw, C. I. (2009). Timing and plasticity in the cerebellum: Focus on the granular layer. *Trends Neurosci.*, *32*, 30–40.

D'Angelo, E., Nieus, T., Maffei, A., Armano, S., Rossi, P., Taglietti, V., 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.

de la Rocha, J., Nevado, A., & Parga, N. (2002). Information transmission by stochastic synapses with short-term depression: Neural coding and optimization. *Neurocomputing*, *44–46*, 85–90.

de la Rocha, J., & Parga, N. (2005). Short-term synaptic depression causes a non-monotonic response to correlated stimuli. *J. Neurosci.*, *25*, 8416–8431.

de Ruyter van Steveninck, R. R., Lewen, G. D., Strong, S. P., Koberle, R., & Bialek, W. (1997). Reproducibility and variability in neural spike trains. *Science*, *275*, 1805–1808.

DeWeese, M. R., & Meister, M. (1999). How to measure the information gained from one symbol. *Network*, *10*, 325–340.

Dimitrov, A. G., Miller, J. P., Gedeon, T., Aldworth, Z., & Parker, A. E. (2003). Analysis of neural coding through quantization with an information-based distortion measure. *Network*, *14*, 151–176.

Dobrunz, L. E., & Stevens, C. F. (1997). Heterogeneity of release probability, facilitation, and depletion at central synapses. *Neuron*, *18*, 995–1008.

Dum, R. P., Li, C., & Strick, P. L. (2002). Motor and nonmotor domains in the monkey dentate. *Ann. N. Y. Acad. Sci.*, *978*, 289–301.

Eccles, J. C., Ito, M., & Szentagothai, J. (1967). *The cerebellum as a neuronal machine.* Berlin: Springer Verlag.

Forti, L., Cesana, E., Isope, P., Dieudonné, & D'Angelo, E. (2008). The mossy fiber input to Golgi cells in the cerebellum. FENS Abst., vol. 4, 221.5.

Forti, L., Cesana, E., Mapelli, J., & D'Angelo, E. (2006). Ionic mechanisms of autorhythmic firing in rat cerebellar Golgi cells. *J. Physiol.*, *574*, 711–729.

Franco, L., Rolls, E. T., Aggelopoulos, N. C., & Jerez, J. M. (2007). Neuronal selectivity, population sparseness, and ergodicity in the inferior temporal visual cortex. *Biol. Cybern.*, *96*, 547–560.

Fuhrmann, G., Segev, I., Markram, H., & Tsodyks, M. (2002). Coding of temporal information by activity-dependent synapses. *J. Neurophysiol.*, *87*, 140–148.

Goldman, M. S. (2004). Enhancement of information transmission efficiency by synaptic failures. *Neural Comput.*, *16*, 1137–1162.

Hansel, C., Linden, D. J., & D'Angelo, E. (2001). Beyond parallel fiber LTD: The diversity of synaptic and non-synaptic plasticity in the cerebellum. *Nat. Neurosci.*, *4*, 467–475.

Hartmann, M. J., & Bower, J. M. (1998). Oscillatory activity in the cerebellar hemispheres of unrestrained rats. *J. Neurophysiol.*, *80*, 1598–1604.

Jakab, R. L., & Hamori, J. (1988). Quantitative morphology and synaptology of cerebellar glomeruli in the rat. *Anat. Embryol. (Berl.)*, *179*, 81–88.

Jorntell, H., & Ekerot, C. F. (2006). Properties of somatosensory synaptic integration in cerebellar granule cells in vivo. *J. Neurosci.*, *26*, 11786–11797.

Kistler, W. M., & De Zeeuw, C. I. (2003). Time windows and reverberating loops: A reverse-engineering approach to cerebellar function. *Cerebellum*, *2*, 44–54.

Koch, C., & Segev, I. (2000). The role of single neurons in information processing. *Nat. Neurosci.*, *3* (Suppl.), 1171–1177.

Koester, H. J., & Johnston, D. (2005). Target cell–dependent normalization of transmitter release at neocortical synapses. *Science*, *308*, 863–866.

London, M., Schreibman, A., Hausser, M., Larkum, M. E., & Segev, I. (2002). The information efficacy of a synapse. *Nat. Neurosci.*, *5*, 332–340.

Lu, T., & Wang, X. (2004). Information content of auditory cortical responses to time-varying acoustic stimuli. *J. Neurophysiol.*, *91*, 301–313.

Maex, R., & Schutter, E. D. (1998). Synchronization of Golgi and granule cell firing in a detailed network model of the cerebellar granule cell layer. *J. Neurophysiol.*, *80*, 2521–2537.

Magistretti, J., Castelli, L., Forti, L., & D'Angelo, E. (2006). Kinetic and functional analysis of transient, persistent, and resurgent sodium currents in rat cerebellar granule cells in situ: An electrophysiological and modelling study. *J. Physiol.*, *573*, 83–106.

Malenka, R. C., & Bear, M. F. (2004). LTP and LTD: An embarrassment of riches. *Neuron*, *44*, 5–21.

Manwani, A., & Koch, C. (2001). Detecting and estimating signals over noisy and unreliable synapses: Information-theoretic analysis. *Neural Comput.*, *13*, 1–33.

Manwani, A., Steinmetz, P. N., & Koch, C. (2002). The impact of spike timing variability on the signal-encoding performance of neural spiking models. *Neural Comput.*, *14*, 347–367.

Mapelli, J., & D'Angelo, E. (2007). The spatial organization of long-term synaptic plasticity at the input stage of cerebellum. *J. Neurosci.*, *27*, 1285–1296.

Mapelli, L., Rossi, P., Nieus, T., & D'Angelo, E. (2009). Tonic activation of GABA-B receptors reduces release probability at inhibitory connections in the cerebellar glomerulus. *J. Neurophysiol.*, *101*(6), 3089–3099.

Mitchell, S. J., & Silver, R. A. (2003). Shunting inhibition modulates neuronal gain during synaptic excitation. *Neuron*, *38*(3), 433–445.

Moreno-Bote, R., & Parga, N. (2004). Role of synaptic filtering on the firing response of simple model neurons. *Phys. Rev. Lett.*, *92*, 028102.

Natschlager, T., & Maass, W. (2001). Computing the optimally fitted spike train for a synapse. *Neural Comput.*, *13*, 2477–2494.

Nemenman, I., Bialek, W., & de Ruyter van Steveninck, R. R. (2004). Entropy and information in neural spike trains: Progress on the sampling problem. *Phys. Rev. E*, *69*, 056111.

Nielsen, T. A., DiGregorio, D. A., & Silver, R. A. (2004). Modulation of glutamate mobility reveals the mechanism underlying slow-rising AMPAR EPSCs and the diffusion coefficient in the synaptic cleft. *Neuron*, *42*, 757–771.

Nieus, T., Sola, E., Mapelli, J., Saftenku, E., Rossi, P., & D'Angelo, E. (2006). LTP regulates burst initiation and frequency at mossy fiber–granule cell synapses of rat cerebellum: Experimental observations and theoretical predictions. *J. Neurophysiol.*, *95*, 686–699.

Paninski, L. (2003). Estimation of entropy and mutual information. *Neural Comput.*, *15*, 1191–1254.

Philipona, D., & Coenen, O. J.-M. D. (2004). Model of granular layer encoding of the cerebellum. *Neurocomputing*, *58–60*, 575–580.

Quian Quiroga, R., & Panzeri, S. (2009). Extracting information from neuronal populations: Information theory and decoding approaches. *Nat. Rev Neurosci.*, *10*, 173–185.

Rancz, E. A., Ishikawa, T., Duguid, I., Chadderton, P., Mahon, S., & Hausser, M. (2007). High-fidelity transmission of sensory information by single cerebellar mossy fibre boutons. *Nature*, *450*, 1245–1248.

Roddey, J. C., & Jacobs, G. A. (1996). Information theoretic analysis of dynamical encoding by filiform mechanoreceptors in the cricket cercal system. *J. Neurophysiol.*, *75*, 1365–1376.

Rolls, E. T., & Deco, G. (2002). *Computational neuroscience of vision.* Oxford: Oxford University Press.

Rothman, J. S., Cathala, L., Steuber, V., & Silver, R. A. (2009). Synaptic depression enables neuronal gain control. *Nature*, *457*, 1015–1018.

Sargent, P. B., Saviane, C., Nielsen, T. A., DiGregorio, D. A., & Silver, R. A. (2005). Rapid vesicular release, quantal variability, and spillover contribute to the precision and reliability of transmission at a glomerular synapse. *J. Neurosci.*, *25*, 8173–8187.

Saviane, C., & Silver, R. A. (2006). Fast vesicle reloading and a large pool sustain high bandwidth transmission at a central synapse. *Nature*, *439*, 983–987.

Schreiber, S., Fellous, J. M., Tiesinga, P., & Sejnowski, T. J. (2004). Influence of ionic conductances on spike timing reliability of cortical neurons for suprathreshold rhythmic inputs. *J. Neurophysiol.*, *91*, 194–205.

Schweighofer, N., Doya, K., & Lay, F. (2001). Unsupervised learning of granule cell sparse codes enhances cerebellar adaptive control. *Neuroscience*, *103*, 35–50.

Shannon, C. E. (1948). The mathematical theory of communication. *Bell Syst. Tech. J.*, *27*, 379–423.

Sharpee, T. O., Sugihara, H., Kurgansky, A. V., Rebrik, S. P., Stryker, M. P., & Miller, K. D. (2006). Adaptive filtering enhances information transmission in visual cortex. *Nature*, *439*, 936–942.

Silver, R. A., Traynelis, S. F., & Cull-Candy, S. G. (1992). Rapid-time-course miniature and evoked excitatory currents at cerebellar synapses in situ. *Nature*, *355*, 163–166.

Smith, E. C., & Lewicki, M. S. (2006). Efficient auditory coding. *Nature*, *439*, 978–982.

Sola, E., Prestori, F., Rossi, P., Taglietti, V., & D'Angelo, E. (2004). Increased neurotransmitter release during long-term potentiation at mossy fibre-granule cell synapses in rat cerebellum. *J. Physiol.*, *557*, 843–861.

Solinas, S., Forti, L., Cesana, E., Mapelli, J., De Schutter, E., & D'Angelo, E. (2007a). Computational reconstruction of pacemaking and intrinsic electroresponsiveness in cerebellar Golgi cells. *Front Cell Neurosci.*, *1*, 2.

Solinas, S., Forti, L., Cesana, E., Mapelli, J., De Schutter, E., & D'Angelo, E. (2007b). Fast-reset of pacemaking and theta-frequency resonance patterns in cerebellar Golgi cells: Simulations of their impact in vivo. *Front Cell Neurosci.*, *1*, 4.

Strong, S. P., Koberle, R., de Ruyter van Steveninck, R. R., & Bialek, W. (1998). Entropy and information in neural spike trains. *Phys. Rev. Lett.*, *80*(1), 197–200.

Sun, H. Y., Lyons, S. A., & Dobrunz, L. E. (2005). Mechanisms of target-cell specific short-term plasticity at Schaffer collateral synapses onto interneurones versus pyramidal cells in juvenile rats. *J. Physiol.*, *568*, 815–840.

Theunissen, F. E., & Miller, J. P. (1991). Representation of sensory information in the cricket cercal sensory system. II. Information theoretic calculation of system accuracy and optimal tuning-curve widths of four primary interneurons. *J. Neurophysiol.*, *66*, 1690–1703.

Theunissen, F., Roddey, J. C., Stufflebeam, S., Clague, H., & Miller, J. P. (1996). Information theoretic analysis of dynamical encoding by four identified primary sensory interneurons in the cricket cercal system. *J. Neurophysiol.*, *75*, 1345–1364.

Tiesinga, P. H. (2001). Information transmission and recovery in neural communication channels revisited. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.*, *64*, 012901.

Tiesinga, P. H., Fellous, J. M., Jose, J. V., & Sejnowski, T. J. (2002). Information transfer in entrained cortical neurons. *Network*, *13*, 41–66.

Tiesinga, P. H., & Toups, J. V. (2005). The possible role of spike patterns in cortical information processing. *J. Comput. Neurosci.*, *18*, 275–286.

Treves, A., & Panzeri, S. (1995). The upward bias in measures of information derived from limited data samples. *Neural Comput.*, *7*, 399–407.

Treves, A., Panzeri, S., Rolls, E. T., Booth, M., & Wakeman, E. A. (1999). Firing rate distributions and efficiency of information transmission of inferior temporal cortex neurons to natural visual stimuli. *Neural Comput.*, *11*, 601–632.

Treves, A., & Rolls, E. T. (1991). What determines the capacity of autoassociative memories in the brain? *Network*, *2*, 371–398

Tsodyks, M. V., & Markram, H. (1997). The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proc. Natl. Acad. Sci. USA*, *94*, 719–723.

Vos, B. P., Wijnants, M., Taeymans, S., & de Schutter, E. (1999). Miniature carrier with six independently moveable electrodes for recording of multiple single-units in the cerebellar cortex of awake rats. *J. Neurosci. Methods*, *94*, 19–26.

Wan, Y. H., Jian, Z., Wen, Z. H., Wang, Y. Y., Han, S., Duan, Y. B., et al. (2004). Synaptic transmission of chaotic spike trains between primary afferent fiber and spinal dorsal horn neuron in the rat. *Neuroscience*, *125*, 1051–1060.

Zador, A. (1998). Impact of synaptic unreliability on the information transmitted by spiking neurons. *J. Neurophysiol.*, *79*, 1219–1229.

Received September 18, 2007; accepted February 16, 2010.