Llenceu aquesta pàgina i substituïu-la per aquella que us faciliti la Unitat d'Informació i Projecció Institucionals (UIPI), disponible al formulari següent
http://www.upf.edu/uii/sgrafics/formulari_tesi.htm

The cerebellum and anticipatory actions

Computational modeling, robotics and neuroprosthetic studies

Ivan Herreros Alonso

Tesi Doctoral UPF / 2013

Supervisada pel

Dr. Paul Verschure

Synthetic, Perceptive, Emotive and Cognitive Systems (SPECS) and Institució Catalana de Recerca i Estudis Avançats (ICREA).

Dipòsit Legal: September 13th, 2013

Dedicated to the memory of my father.

Agraïments

My deepest gratitude to my supervisor, Prof. Paul Verschure for making this research a challenging but playful experience, illuminating and ambitious. I also thank my DEA supervisor, Prof. Leo Wanner, who long time ago allowed me to set foot on the UPF Technology Department, and Prof. Vito Pirrelli, from the Istituto di Linguistica Computazionale at the CNR at Pisa, who really started all this.

I acknowledge all my colleagues at Specs, specially Riccardo Zucca and Andrea Giovannucci, the team with which I started this research, and Martí Sanchez Fibla, Giovanni Maffei and Santiago Brandi, who have been fundamental in achieving the newest results that appear in this dissertation.

I am also thankful to the staff of the UPF and the Department of Technology. I have been very lucky to attend such an excellent institution, that, furthermore, is in my hometown.

My final and most heartfelt thank goes to my family. Specially, to Maja, this thesis is dedicated to her as well, and to the newest, handsomer and, without any doubt, most successful member of the Herreros-Denzer team: Emil. My little family made this last year of my PhD *das beste*.

Abstract

Animals excel in skilled motor behavior, especially when compared to artificial systems. Several adaptive and predictive motor control frameworks have been proposed to account for such performance, but the underlying biological implementation remains elusive. In vertebrates, the cerebellum is the most promising candidate locus where adaptive and predictive motor control primitives might be implemented. Indeed, several motor learning paradigms indicate that the cerebellum can drive anticipatory and well-timed coordinated motor actions. The most widely employed of such paradigms, classical conditioning of the eyelid reflex (eyeblink conditioning), has a well-defined circuitry blueprint as well as known input and output pathways. However, the underlying physiological mechanisms are not fully understood, and neither is it clear whether cerebellar function in eyeblink conditioning can be extrapolated to more general anticipatory actions. This dissertation addresses these two questions relying on computational, robotic and neuro-prosthetic approaches, following the hypothesis that the cerebellum, working as an adaptive feedforward filter, interacts with a reactive layer of feedback control. First, we show that GABAergic slow inhibitory currents provide a biologically grounded mechanism accounting for the representation of time in the cerebellar cortex. Based on such assumption, we build a computational model of the cerebellum that successfully controls robots in avoidance learning tasks and replaces the learning function of an inactivated rat cerebellum using a neuro-prosthetic implantation. Thus, we provide a biologically grounded explanation of the mechanisms underlying the acquisition of anticipatory responses by the cerebellum, that is fully functional in both robotic and neuro-prosthetic scenarios. Altogether, this work advances our understanding of the mechanisms at the basis of coordinated motor control in animals, and through this, it takes a step towards developing equivalent motor capabilities in artificial systems.

Resum

Els animals assoleixen unes habilitats motores que superen de llarg la dels sistemes artificials actuals. S'ha proposat que diferents marcs de control adaptatiu i predictiu estan a la base d'aquest acompliment, pero la seva implementació biològica roman desconeguda. En el cas dels vertebrats, és el cerèvel l'estructura cerebral on més probablement s'implementen aquestes primitives adaptatives i predictives. De fet, multiples paradigmes d'aprenentatge motor indiquen que el cerevel es capaç de controlar l'adquisició de accions motores anticipatòries i executar-les amb gran precisió temporal. El paradigma més estudiat, el condicionament clássic del parpelleig, té una anatomia ben definida així com també un circuits d'entrada i sortida coneguts. Tanmateix, els mecanismes a la base d'aquest tipus d'aprenentatge no es coneixen com tampoc es sap si la funció del cérevel al condicionament del parpelleig es pot generalitzar a altres tipus d'accions anticipatòries. Aquesta disertació adreça aquestes dues qüestion i ho fa mitjaçant estudis computacionals, robotics i neuro-prostètics, on apliquem l'hipòtesi que el cerevel actúa com a un filtre adaptatiu que complementa la funció d'una capa de control reactiva. En primer lloc, demostrem que les corrents inhibitories per versament proporcionen un sustrat fisiològic per a la represtació del temps en l'escorça del cerevel. Basant-nos en aquesta assumpció, desarrotlem un model computacional que es capaç de controlar un robot en un tasca de prevenció de colisions i que, implementat en un implant neuro-prostètic, també pot reemplaçar funcionalment el cerevel farmacologicament inactivat d'una rata. D'aquesta manera, donem un explicació biologicament plausible per als mecanismes que permeten l'aquisició de respostes anticipatories al cerevel, que a més a més es funcional tant en l'àmbit de la robòtica com en el de les neuro-pròtesis. Tot plegat, aquesta feina avança el coneixement relatiu als mecanismes que a la base del control motor en animals, i mitjançant això dòna un pas vers el desenvolupament de sistemes artificials amb capacitats equivalents.

Publications

Included in the thesis

Peer-reviewed

Herreros Alonso, I. & Verschure, P. F. (2013a). Nucleo-olivary inhibition balances the interaction between the reactive and adaptive layers in motor control. *Neural Networks*

Herreros Alonso, I., Maffei, G., Brandi, S., Sanchez Fibla, M., & Verschure, P. F. (2013b). Speed generalization capabilities of a cerebellar model on a rapid navigation task. Dins *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on*, ps. 8108–8113. IEEE

In preparation

Herreros Alonso, I., Giovannucci, A., & Verschure, P. F. (2013a). Replacing a cerebellar microcircuit by a synthetic system. *submitted*

Herreros Alonso, I. & Verschure, P. F. (2013b). Slow inhibitory currents as a substrate for the representation of time in the cerebellar cortex. *in preparation*

Other publications and abstracts

Peer reviewed

Brandi, S., Herreros, I., Sánchez-Fibla, M., & Verschure, P. F. (2013). Learning of motor sequences based on a computational model of the cerebellum. Dins *Biomimetic and Biohybrid Systems*, ps. 356–358. Springer

Maffei, G., Herreros, I., Sánchez-Fibla, M., & Verschure, P. F. (2013). Acquisition of anticipatory postural adjustment through cerebellar learning in a mobile robot. Dins *Biomimetic and Biohybrid Systems*, ps. 399–401. Springer

Bamford, S., Hogri, R., Giovannucci, A., Taub, A., Herreros, I., Verschure, P., Mintz, M., & Del Giudice, P. (2012). A vlsi field-programmable mixed-signal array to perform neural signal processing and neural modeling in a prosthetic system. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 20(4):455–467

Bobo, L., Herreros Alonso, I., & Verschure, P. F. (2012). A digital neuromorphic implementation of cerebellar associative learning. Dins *Biomimetic and Biohybrid Systems*, ps. 13–25. Springer

Prueckl, R., Taub, A., Herreros, I., Hogri, R., Magal, A., Bamford, S., Giovannucci, A., Almog, R., Shacham-Diamand, Y., Verschure, P., Mintz, M., Scharinger, J., Silmon, A., & Guger, C. (2011b). Behavioral rehabilitation of the eye closure reflex in senescent rats using a real-time biosignal acquisition system. Dins *Engineering in Medicine and Biology Society,EMBC, 2011 Annual International Conference of the IEEE*, ps. 4211 –4214. ISSN 1557-170X. doi: 10.1109/IEMBS.2011.6091045

Prueckl, R., Grunbacher, E., Ortner, R., Taub, A., Hogri, R., Magal, A., Segalis, E., Zreik, M., Nossenson, N., Herreros, I., et al. (2011a). The application of a real-time rapid-prototyping environment for the behavioral rehabilitation of a lost brain function in rats. Dins *Computational Intelligence, Cognitive Algorithms, Mind, and Brain (CCMB), 2011 IEEE Symposium on*, ps. 1–8. IEEE

Inderbitzin, M., Herreros Alonso, I., & Verschure, P. F. M. J. (2010). An integrated computational model of the two phase theory of conditioning. *World Congress of Computational Intelligence*

Herreros, I., Zucca, R., Giovannucci, A., & Verschure, P. F. M. J. (2009). The interaction of purkinje cell and inhibitory interneuron plasticity during classical conditioning. doi: 10.3389/conf.neuro.06.2009.03.190

Contents

List of Figures

CHAPTER 1

Introduction

Originally developed as a means for studying a simple form of associative learning in animals, the classical conditioning of the eye-blink response, has become a very useful tool for studying cerebellar function (Freeman & Steinmetz, 2011; Thompson & Steinmetz, 2009; Christian & Thompson, 2003; Hesslow & Yeo, 2002). In the eye-blink conditioning paradigm a neutral Conditioning Stimulus (CS), e.g., a tone, precedes by a fixed Inter Stimulus Interval (ISI) an innately aversive Unconditioned Stimulus (US), e.g., an airpuff. Prior to any training, the latter triggers a protective reflex-like eye-blink, the Unconditioned Response (UR). After a number of paired CS-US stimulation trials, an animal develops also a motor response to the CS, the Conditioned Response (CR), namely, an eye-blink that coincides with the expected arrival of the US. It is now well established that this form of learning depends on the cerebellum. Moreover, in the simplest eye-blink conditioning paradigm, when both the CS and US overlap in time and the ISI is below 1 second, learning within the cerebellum is sufficient for the acquisition of CRs.

The behavioral hallmark of eye-blink conditioning is that the animal acquires an anticipatory response that is specific both in space –which

muscles are activated– and in time –when are they activated–. Here we focus on the *when* aspect of the task, i.e., the cerebellum's ability to acquire adaptively-timed responses, as the mechanisms that achieve such a temporal specificity are not yet well understood. This is the first question addressed in this dissertation. Our response contributes to the state of the art by providing a parsimonious solution that refers to an idiosyncratic trait of the physiology of the cerebellar cortex, namely, the prevalence of the spillover inhibitory currents.

Once we have provided such a mechanism for generating a representation of the passage of time in the cerebellar cortex, we asses how our solution affects the learning dynamics of a computational model of the cerebellum. However, since we frame these tests in the context of avoidance learning, we address at the same time a more general issue. That is, how does our knowledge of the role of the cerebellum in classical conditioning generalize to avoidance learning? This issue, from our point of view, has not yet been sufficiently addressed in the field and it is pivotal for allowing to elaborate, departing from classical conditioning, a general theory about the cerebellar control of anticipatory actions. In relation to this, our contribution proofs that the adaptive and reactive layers of control can cooperate during avoidance behavior in a graded manner. Our proposal introduces the novelty that, instead of totally replacing the reactive by the anticipatory responses, we argue that the goal of cerebellar learning in anticipatory actions is to merge predictive feed-forward with reactive feedback control.

Next, we study whether an effect reported in the eye-blink conditioning can be functional in avoidance learning. For this, employing a robotic setup, we suggest that an experimental finding from the classical conditioning paradigm, that links the intensity of a predictive stimulus with the latency of an anticipatory response, can be operationally interpreted in behavioral terms. More precisely, we propose that a robot that learns to navigate a track at a slow velocity, can afterwards reproduce the same

skill at higher ones, and that this generalization, via a shaping-like process allows the robot to traverse the track at speeds that initially were beyond its reach. In short, with this interpretation, we aim to turn a single stimulus-response association into a sensorimotor contingency linking a family of stimuli with their appropriately modulated responses.

The final contribution of this dissertation uses a different approach to the validation of our computational theory of cerebellar learning. Indeed, despite having demonstrated that our computational models can replicate the behavior of a cerebellar microcircuit, the achievement was reached using synthetic signals acquired by a robot instead of noisy brain activity. Thus, in order to provide evidence that our model is developed along principles that are still valid in the biological context, we aim to deploy the model in a biological context. For this, we interface a synthetic cerebellum with the brain of a rat whose cerebellum has been inactivated by anesthesia with the goal of reproducing the normal acquisition of conditioned responses in a healthy animal. The result is a pioneering study that shows for the first time that a neuro-prosthesis in closed-loop with the central nervous system allows for the recovery of a lost learning function.

We have briefly summarized the main four contributions of this thesis. Now, before addressing each of them in detail in the following chapters, we will provide the necessary background in the remaining of the current one. We first quickly introduce remarkable aspects of the cerebellar anatomy. Afterwards we review the theories of cerebellar function. It follows a gross introduction of the cerebellar anatomy and physiology and, finally, a short review on the computational theories of cerebellar function.

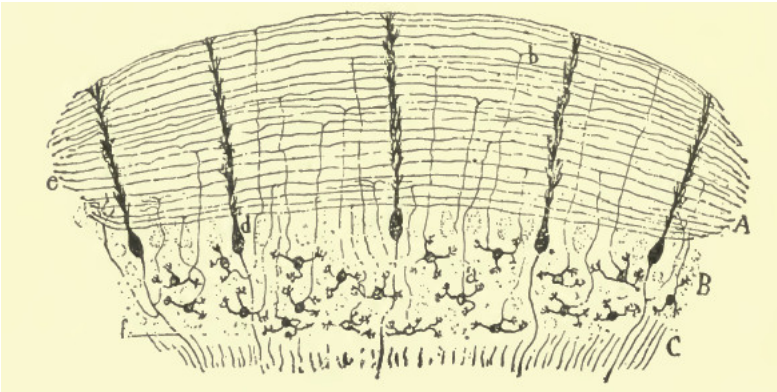

Figure 1.1: Longitudinal section of the cerebellar cortex. Illustration from Thomas (1911) after a drawing by Ramón y Cajal.

1.1 A first glance at the cerebellum

The cerebellum is a brain structure that sits in the ventral-posterior part of the head, below the occipital lobe and above the brain stem. Its placement is optimal for mediating the communication between the *mind* (embodied in the higher brain areas) and the body (that both receives downstream commands and carries upstream perceptions via the spinal cord). In numbers, the cerebellum occupies one-fifth of the volume of the central nervous systems and hosts half of its total number of neurons. From an anatomical standpoint, the cerebellum is a very remarkable structure both quantitatively and qualitatively. It contains from the tiny and structurally simple granule cells, to the big and complex Purkinje cells. In the beautiful planar dendritic trees of the Purkinje cells we find, at the same time, the highest convergence of input fibers to a single cell observed in the brain, $>10^5$ parallel fiber impinging each Purkinje cell, and the lowest possible, the *one*-to-*one* strong connection between a climbing fiber and a Purkinje cell.

The most prominent anatomical feature of the cerebellum is found in the cerebellar cortex: its lattice-like regular arrangement, which as de-

scribed in Braitenberg & Atwood (1958), can be obtained through the bidimensional translation of a cubic building-block. This cytoarchitecture, that is repeated all through the cerebellar cortex, clearly defines preferred directions for the propagation of information and induces orthogonal arrangements at multiple stages of this structure. For instance, the dendritic trees of Purkinje cells are contained in planes perpendicular to the parallel fibers, planes that also contain the axons of the inferior olivary cells and their end ramifications, the climbing fibers. Macroscopically, the structure itself, the cerebellar cortex, is folded only along one direction, like an accordion, with medio-lateral grooves, the folia. An unfolded hemisphere of a human cerebellum would cover an ellipsoidal surface with an antero-posterior extent of two meters and a medio-lateral width of 0.15 m (Braitenberg et al., 1997). For a comparison, a flattened hemisphere of human cerebral cortex is roughly circular with a 0.3 m diameter. The macroscopic and microscopical arrangements relate as follows: if we slice the cerebellum perpendicular to the surface and along the folia (longitudinal section, Fig. 1.1), we will section both Purkinje cell's dendritic trees and the axons the inferior olive cells capturing entire parallel fibers, whereas if we slice perpendicular the folia (transverse section, Fig. 1.2), we will preserve entire Purkinje cells and section the parallel fibers.

This impressionistic anatomical preamble is worth because from the perspective of a computational scientist, the main the motivation for studying of the cerebellum stems from its fascinating architecture. Its crystal-like arrangement of cells eloquently delineates the flow of information, and to a great extent, seems to indicate the underlying computational operations, such that it appears feasible to elucidate the cerebellar algorithm. Moreover, since the same structure is maintained through all the cerebellar tissue, arguably it is only necessary to *solve* the algorithm for a tiny piece of tissue to be able to describe the algorithm of the whole cerebellum (Apps & Garwicz, 2005). These hopes might be naive, but

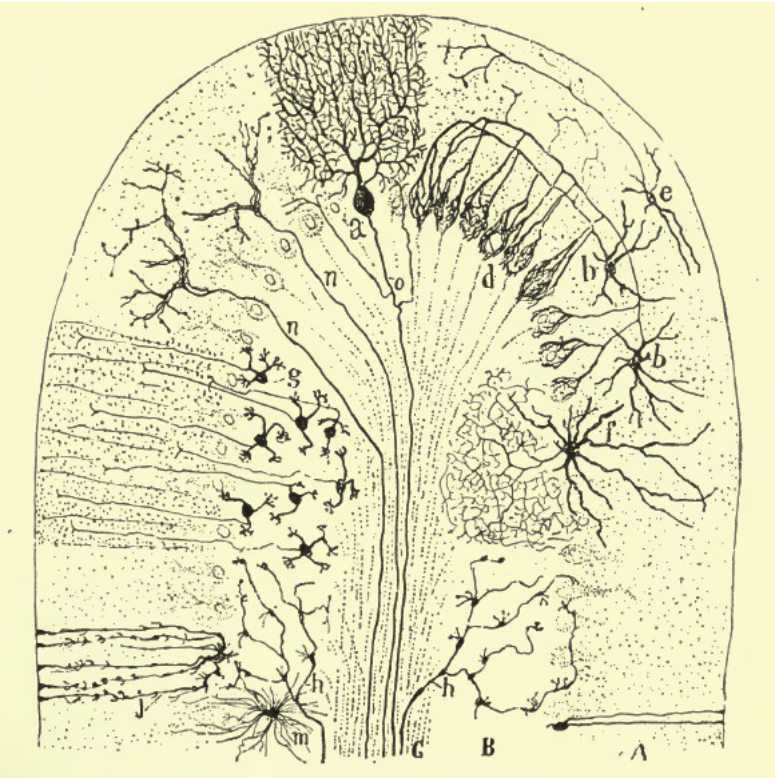

Figure 1.2: Transverse section of the cerebellar cortex. Illustration from Thomas (1911) after a drawing by Ramón y Cajal.

they have nonetheless been shared by the researchers that more prominently contributed to our understanding of the cerebellum: Eccles, Ito, Marr, Albus, Braitenberg, etc.

Indeed, the works that conform this dissertation constitute complementary efforts in a same endeavor: the deciphering of the cerebellar algorithm that supports the acquisition of adaptive reflexes, and through this the understanding of the basis of cerebellar motor control. But this goal, besides being an end in itself, is also a tractable means to solve the more general problem of unveiling the *general* cerebellar algorithm that makes the cerebellum such a successful product of evolution.

1.2 Cerebellar function

1.2.1 Cerebellar role in motor control

The concept of "cerebellar function" is subject to an endless debate (Thach, 2013; Glickstein, 2007; Thomas, 1912). The prevailing view states that the function of the cerebellum is the (fine) coordination of motor commands. It was first formulated in the XIX century by Flourens based on lesion studies in animals (Flourens, 1842). Quoting (page XII): "[...] *dans le cervelet réside une propriété dont rien ne donnait encore l'idée en physiologie, et qui consiste à coordoner les mouvements voulus par certaines parties du système nerveux* [...]" [1]. It is still customary to introduce the cerebellum as a brain structure dedicated to motor control. Evidence for this role has been accumulating through the decades in the clinical domain (Glickstein, 1994; Thomas, 1912) and in experimental psychology, with electrophysiological recording in animals and functional magnetic resonance studies in humans (Yamamoto et al., 2007; Shidara, 1993; Ojakangas & Ebner, 1992; Gilbert & Thach, 1977).

At this point, even though it is evident that the cerebellum is necessary for fine motor control, the debate is whether the cerebellar function *is* motor control or a more general function *necessary* for motor control. For instance, it has been suggested that the main function of the cerebellum is the processing of sensory information (Bower, 1997a,b; Paulin, 1993). According to this view, the cerebellum extracts relevant features from the sensory signal (Angelaki et al., 2004), or subtracts noise (Gao et al., 1996). Then, being able to extract information from the sensory domain is necessary for accurate motor control, but it is also necessary for cognitive tasks that do not require the performance of any motor action.

[1] translation: "within the cerebellum we find a property that was not yet predicted by the physiology, and that consists in the coordination of the movements commanded by other regions of the nervous system"

Another more specialized view links cerebellar function to motor control through the implementation of forward models (Bastian, 2006; Paulin, 1993; Miall et al., 1993). In this case, the cerebellum *anticipates* sensory signals to feedback controllers. This framework enables inverse feedback controllers to not depend on delayed feedback, what minimizes the effort and increases the stability of the control.

Two additional theories of cerebellar function make it an organ necessary for, but not restricted to, motor control. First, the presence of the long parallel fibers in the cerebellar cortex suggested to Braitenberg that the cerebellar function may consist in the detection and generation of sequences (Braitenberg et al., 1997).

On the other hand, and based on clinical and experimental evidence, it is suggested that the cerebellum might act as an internal clock. Cerebellar patients are specifically impaired in tasks requiring to perform explicit-timing (Spencer et al., 2003). If a subject has to tap with a finger on a table every second while remain motionless in the intervals, it is necessary to internally generate a code for the passage of time, i.e., to generate an explicit representation of time. This can be avoided if the subject continuously moves the finger between taps, e.g., drawing a circle in the air. In the latter case, the passage of time can be implicitly coded in the dynamics of the ongoing movement. Cerebellar patients are selectively impaired only in explicit-timing tasks. Moreover, since cerebellar patients are also impaired in the perception of time, it has been proposed that the cerebellum can act as a general internal time-keeping mechanism that is employed both by motor or perceptual systems (Ivry & Keele, 1989). We emphasize this distinction between explicit- and implicit-timing tasks because eye-blink conditioning, the task we focus in, is an explicit-timing task. That is, the CS stimulus has no dynamics that could convey time-related information for which the timing information has to be internally generated in the cerebellum.

So far, we have presented theories of cerebellar function that related intimately to motor control and that are elaborated, with the exception of Braitenberg's view, based on correlational evidence: lesion, clinical, electro-physiological or imaging studies that link observed behavior with neural activity or brain damage. In what follows we will introduce two lines of evidence that relate the (human) cerebellum with functions that, again, go beyond motor control but do so from evolutionary and phylogenetical perspectives.

1.2.2 Evolutionary perspective

Citing Sir Eccles, "With each further evolutionary development of the brain, the cerebellum seemed to be a necessary adjunct, presumably because it possessed some unique mode of processing information" (Eccles et al., 1967). We find a concrete example of this during the increase in encephalization that preceded the advent of the Mammalia order. Then, the increase on brain mass relative to body mass resulted from the expansion of the somatosensory and motor cortices, together with the cerebellum (Rowe et al., 2011). Congruently with the motor-centered view of cerebellar function, this evolutionary development was putatively associated with improved tactile and motor coordination in pre-mammals. The brain of the Hominoids (namely, geni Orangutan, Gorilla, Chimpanzee and Homo) provides another example. In that case, a remarkable increase in encephalization quotient caused an enlargement of all brain areas, but the size of the cerebellum increased even more than expected for the sheer enlargement of the cerebral cortices (Rilling, 2006).

The ratio between the cerebellum and the rest of the brain is far from being constant across species, and in the case of the evolution of modern humans, the history is specially telling. The evolutionary pulse that lead to bigger brains in our Homo ancestors (Cro-Magnon and Neandertal) affected more the cerebral cortices than the cerebellum, but the last evolutionary leap towards the Homo Sapiens saw a reversal of the trend and,

proportionally, our cerebellum occupies a bigger portion of our brain volume compared to our ancestors (Weaver, 2005). Two conclusions can be drawn from this picture. First, recent evidence confirms Eccles remark, and indeed the cerebellum co-developed with other brain areas throughout evolution (Barton & Harvey, 2000). And secondly, in the light of the evolution of the cerebellum in the Homo genus, the vision of cerebellar function as a purely motor one seems unlikely.

Indeed, compared to our human ancestors, modern humans stand out by the complexity of their social environment, the linguistic communication and abstract reasoning skills, and by the ability to use tools (technology). In general, these capabilities, except, arguably, the use of technology, go beyond motor control, so, why would recent evolution have resulted in such an enlarged cerebellum in humans? It is been suggested that the mechanisms that allow for the control of sequences of complex motor actions (as in tool making) are the basis over which syntax, and therefore the language ability, evolved (Sterelny, 2012; Fitch, 2011), and that these mechanisms are grounded in the cortico-cerebellar system (Barton, 2012).

Then, to summarize, from an evolutionary perspective, we can interpret that even though the main role of the cerebellum in primates might be the control of complex motor sequences, allowing for fine motor control, in humans, the same computational facilities provided by the cortico-cerebellar system sub-serve, *in addition* to motor control tasks, also cognitive operations.

1.2.3 Phylogenesis and comparative anatomy of the cerebellum

All vertebrates have a cerebellum and the majority of them have additional cerebellum-like structures (Bell, 2002). Alike the cerebellum, cerebellum-like structures have an histological division in their tissue, the molecular layer, that presents a lattice-like arrangement of parallel

fibers, originating from granule cells, which contact the spiny dendritic trees of the principal cells (Bell et al., 2008). The main structural divergence between cerebellum-like structures and the cerebellum is the absence of the climbing fibers in the former ones (Devor, 2000). The principal cells of the cerebellum-like structures, that are analogous to the cerebellar Purkinje cells, lack the hyper-specific 1-to-1 innervation provided the climbing fibers. However, both structures share the same pattern of high convergence of non-specific information through an input pathway (namely, through the parallel fibers in both cases) in contrast with a much more specific information coming from a separated pathway (through the climbing fibers in the case of the cerebellum and through afferent fibers in the case of the cerebellum-like structures).

The only exception to the presence of the cerebellum in vertebrates is raised by their most ancient living lineage, lampreys, where such a presence is debated (Bell et al., 2008; Northcutt, 2002). However lampreys and their evolutionary ancestors do have cerebellum-like structures, what suggests that cerebellum-like structures might be the evolutionary ancestors of the cerebellum (Montgomery et al., 2012; Bell et al., 2008). Interestingly, since cerebellum and cerebellum-like structures coexist in most vertebrates (including all mammals), it has been proposed that the cerebellum appeared by duplication of the cerebellum-like structures, that then was *subsumed* in the already existing brain architecture (Montgomery et al., 2012). Considering that the general plan of the vertebrate central nervous system, including spinal cord, basal ganglia, thalamus and a primitive cerebral cortex, namely the pallium, is already present in lampreys (Grillner et al., 2005), we may come to the surprising conclusion that the cerebellum is a late add-on to such a general plan.

In any case, independently of the evolutionary origin of the cerebellum, the already-mentioned structural similarity between cerebellum and cerebellar-like structures by itself already suggests that, given the analogous cyto-architectural design, both structures could perform similar

computations or have an analogous function (Bell et al., 2008; Devor, 2000). The previous is a relevant argument because whereas the function of the cerebellum is subject to debate, it is agreed that "cerebellum-like structures remove predictable features from the sensory flow" (quoted from (Bell, 2002)). This is, they subtract from the signal coming through the peripheral afferents the components that can be predicted based on the information that reaches through the parallel fibers. Sources of such information can be efference copy of motor commands, propioception and signals from other sensory modalities.

In a similar vein, Paulin (1993) has shown that across species the size of the different cerebellar lobes is related to the sensory complexity, and not to the motor complexity, of the limbs or organs that they are putatively connected to. For instance, the small echo-locating bats are among the most agile flying animals, what requires very fine motor coordination capabilities, but they posses a relatively small cerebellum for a mammalian. Indeed, the only expanded regions of the cerebellum in small bats are related to the echo-location sense, the same ones enlarged in marine echo-locating mammals, and not to the control of the fore-limbs, i.e., the wings. Another example is given by the platypus. Platypus are very idiosyncratic beaked mammals that live in Australia. Their beak is a very a sophisticated sensory organ, carrying an array of electrosensory and mechanosensory receptors, but from a motor control perspective, it is a simple structure. It results that Platypus have an unusually enlarged cerebellum, where the cerebellar lobe that receives sensory efferents from its beak greatly expanded.

In conclusion, two lines of evidence, phylogenetic and comparative anatomy, favor a sensory-centered view of cerebellar function instead of the more motor-centered one.

1.2.4 Other functions and summary

To complete the review on the theories of cerebellar function, we introduce two additional views, that inhabit opposite ends on the cognitive scale: its involvement in affective processing and the control of anticipatory reflexes. Clinical evidence summarized in Schmahmann & Sherman (1998) links cerebellar lesions with cognitive-affective deficits, the so-called cerebellar cognitive-affective syndrome.

Moreover, clinical (Clark & Squire, 1998), lesion (McCormick & Thompson, 1984), electrophysiological (Jirenhed et al., 2007) and imaging (Cheng et al., 2008) studies have shown that the cerebellum is crucial for the acquisition of anticipatory reflexes in the context of the eyeblink response. We will return to the latter function since it is the focus of this thesis.

In summary, we have provided a comprehensive review on the current views on cerebellar function. These theories have in common that they were developed without considering the internal structure of the cerebellum. At most, considerations regarding the cerebellar anatomy were limited to the appearance of its most external layer. At this point, we may ask why is there such a controversy and diversity on the interpretation regarding cerebellar function. Indeed, the problem may reside directly on the question and not in the answers. Why should there be a *unified* cerebellar function? There is no dispute regarding what the function of the cerebral cortex is because it is obvious that different parts of the cortex serve different functions: processing of visual stimuli, generation of motor commands, etc. It is beyond the scope of this dissertation to address what are the differences between both structures that disqualify such a question for the cerebral cortex but make it still possible to defend that the cerebellum has a single function. However, we can nonetheless mention the fact that whereas the cortical tissue of the cerebellum is remarkably uniform, there is a clear histological differentiation in the cerebral cortical structure. However, even if we assume that, based

on the homogeneity of its tissue, all cerebellar regions perform the same operations it still is a mistake to conclude that the cerebellum can only perform one function. Indeed, homogeneity of structure and function are statements that refer to different levels of analysis: algorithmic and functional.

1.3 Anatomy and physiology of the cerebellum

This section introduces a brief summary of the cerebellar anatomy and physiology. More than 45 years after its publication, Eccles et al. (1967) is still the best source where to find an exhaustive description of the anatomy of the cerebellum. In terms of the microcircuit design implemented in this dissertation, almost of all its elements can be traced back to Eccles' book. However, in what regards to the physiology we will base our models on findings that are posterior to that publication, such as the discovery of cerebellar long term depression (Ito et al., 1982), the recently demonstrated cumulative effect of glomerular spillover inhibition (Crowley et al., 2009), etc. The rest of the relevant literature for the specifics of each computational model is introduced in each chapter.

1.3.1 The cerebellar cortex

Histologically, the cerebellar cortex has three layers. From more external to more internal these are the molecular, Purkinje cell and granular layers. From a computational perspective one can simplify this division by adding the Purkinje cell layer to the molecular one.

Granular layer

The granular layer is exclusively an input layer where the information coming through the mossy fiber pathway reaches the cerebellar cortex.

Mossy fibers originate in the pontine nuclei where they can collect information coming from the rest of the brain. Their terminations, boutons, have excitatory synapses with the granule and Golgi cells. These two are the only cell types found in the granular layer.

Granule cells Granule cells are tiny excitatory cells with a very simple dendritic structure, having only from three to six branches, each branch contacting a single mossy fiber bouton. They receive excitatory input only from the mossy fibers and are inhibited by Golgi cells. Each granule cell has a unique ascending axon that reaches the molecular layer where it bifurcates into two opposite parallel branches, forming a so-called parallel fiber. Cerebellar granule cells are extremely numerous. Their estimated count in the humans ($>10^{10}$) makes up for more than one half of the total number of neurons in the whole brain.

Golgi cells The other cell bodies found in the granular layer correspond to Golgi cells. These are gabaergic inhibitory interneurons that contact granule cells. The dendritic tree of Golgi cells has two separate apical and basal ramifications. The basal dendrites are contacted by mossy fibers whereas the apical ones enter the molecular layer where they are synapsed by granule cells axons, be it in the parallel fiber or in the ascending section of the axon. Golgi cell activation results mainly in the inhibition of granule cells but also in the reciprocal inhibition of other Golgi cells. From the perspective of granule cells, Golgi cells implement two different inhibitory circuits. First, through the basal dendrites, mossy fiber activation that results in direct excitation of granule cells, recruits di-synaptic inhibition as well, thereby establishing a feed-forward inhibitory circuit. Second, granule cell activity activates Golgi cells through the apical dendrites, recruiting feedback inhibition of the granule cells. These are the two hypothesized mechanisms by which Golgi cells control the excitability of granule cells, thereby regulating the amount of input reaching the molecular layer.

Glomeruli In the granular layer we find an important non-cellular structure, the glomeruli. The glomeruli are myelinated aggregations of synapses. They are formed around single mossy fiber boutons, and include the dendrites of granule and Golgi cells and the axonal terminations of Golgi cells. In other words, they delimit the space for synaptic transmission in the granular layer. Such a tight glial encapsulation limits the neurotransmitter diffusion at these synapses, introducing an array of *exotic* dynamics to the synaptic transmission between mossy fibers, granule and Golgi cells, such as presynaptic cross-talk (Mitchell & Silver, 2000b,a) and spillover transmission (Crowley et al., 2009; Hamann et al., 2002).

Molecular layer

The molecular is the only output layer of the cerebellar cortex. It receives inputs from the mossy fiber pathway –via the parallel fibers and ascending axons of granule cells– and from the climbing fiber pathway. The molecular layer contains only inhibitory cells: the so-called Molecular Layer Interneurons (MLIs), namely the basket and stellate cells, and the Purkinje cells. The latter ones are the only output cells of the cerebellar cortex sending inhibitory axons down to the cerebellar nuclei.

Parallel fibers Parallel fibers excite MLIs and Purkinje cells, and MLIs inhibit Purkinje cells. In other words, relaying information from the mossy fibers that has already been pre-processed in the granular layer, parallel fibers establish a feed-forward circuit with direct mono-synaptic excitation and an indirect di-synpatic inhibition to Purkinje cells. As a result of this, the output of a granule cell can either (direcly) excite or (indirectly) inhibit a Purkinje cell. From a computational perspective, this will allow us to model the granule cell's effect on a Purkinje cell by a single gain factor that can take either positive or negative values.

Figure 1.3: Purkinje cell. Drawing by Ramón y Cajal.

Climbing fibers Cells from the molecular layer also receive input from the climbing fiber pathway. Since the times of Ramon y Cajal it is known that each Purkinje cell is contacted by a single climbing fiber. Indeed, the synapse between climbing fibers and Purkinje cell has been described as the more powerful synaptic contact of the central nervous system. Each climbing fiber action potential induces a spike into the post-synpatic Purkinje cell, that, due to its distinctive features, is referred to as a *complex spike* (Eccles et al., 1966a; Thach, 1967). In contrast with the regular *simple spikes* produced by the Purkinje cells, that reflect the integration of activity form the parallel fiber pathway (Häusser & Clark, 1997; Eccles et al., 1966b), complex spikes contain multiple spikelets (little spikes) usually followed by a pause response, this is, a short cessation of the simple spike firing. Crucially, climbing fibers also innervate MLIs (Sugihara et al., 1999), even though in a more modest manner.

Purkinje cells The Purkinje cells are the Eiffel tower of the cerebellum: its more distinct and recognizable feature, with an elaborate and beautiful architecture (Fig. 1.3). A Purkinje cell has a massive and planar dendritic tree that is traversed, and potentially contacted, by over a hundred thousand parallel fibers, and a long axon that reaching the cerebellar nuclei provides the only output of the cerebellum. From a computational perspective, this arrangement endows Purkinje cells with the capability of storing vast amounts of information. Importantly, it is now long established that the synapses from parallel fibers to Purkinje cells are plastic, and that their gain can be modulated by the coincidence of parallel fiber and climbing fiber activity (Linden & Connor, 1991; Ito et al., 1982). In other words, patterns of parallel fiber activity modify the synaptic configuration of Purkinje cells according to their coincidence with climbing fiber spikes. This particular feature of the cerebellar physiology was, initially as an hypothesis, and still is, as an experimental finding, the central tenet of the classical theory of cerebellar learning.

1.3.2 Cerebellar nuclei and inferior olive

Clearly, the most distinctive trait of the cerebellum is the anatomy and physiology of the cerebellar cortex, however, it is agreed that a functional unit of cerebellar computation is distributed among the cerebellar cortex, cerebellar nuclei and inferior olive (Apps & Garwicz, 2005). Such computational units, often referred to as cerebellar microcircuits or cerebellar microcomplexes, constitute parallel negative-feedback loops (Chaumont et al., 2013; Hofstotter et al., 2004).

Cerebellar nuclei All output leaves the cerebellum through the cerebellar nuclei. The cerebellar nuclei, or deep cerebellar nuclei, contain excitatory and inhibitory cells. Excitatory projections from the cerebellar nuclei reach brain stem structures and thalamus, and from there their information can propagate to the rest of the brain. In contrast, inhibitory

projections target the inferior olive, thus having an effect that remains local to the cerebello-olivary system. The last projection is known as nucleo-olivary inhibition. Knowledge regarding the anatomy and physiology of these nuclei is rather poor (Uusisaari & De Schutter, 2011). For this dissertation, it suffices to know that through the cerebellar nuclei Purkinje cell activity can excite extra-cerebellar structures and inhibit the inferior olive.

Inferior olive The inferior olive is a brainstem structure populated by cells that project only to the cerebellum. The physiology of the inferior olive is very idiosyncratic, i.e., olivary cells are electrically coupled through gap junctions and have a very low level of background activity, around 1 Hz. They project to the cerebellar nuclei and to the molecular layer of the cerebellar cortex, being the only source of the above-mentioned climbing fibers. In the classical cerebellar learning theory, the inferior olive has the important role of providing the teaching signal to the cerebellar cortex, and it is still an open question how their particular physiology interacts with such a role.

1.4 Computational theories of cerebellar information processing

The first complete description of the cerebellar anatomy and physiology provided by Eccles et al. (1967) coincided in time with the early peak of popularity of the Perceptron (Rosenblatt, 1958). The structural analogy between the Purkinje cell anatomy and the Perceptron design was so strikingly obvious that very rapidly cerebellar computation was interpreted on terms of learning in artificial neural networks (Brindley, 1964; Marr, 1969; Albus, 1971; Ito, 1972). This gave rise to what it is referred to as the classical theory of cerebellar learning or Marr-Albus-Ito theory. In a nutshell, the theory's main posit is that the climbing fibers provide

each Purkinje cell with a teaching signal that determines the information stored at the synapses between parallel fibers and Purkinje cells. This theory found strong experimental support in Ito et al. (1982), when it was shown that conjunctive parallel fiber and climbing fiber stimulation introduced a specific depression in the Purkinje cell response to the parallel fiber input.

However, Marr and Albus described the cerebellum as a pattern storing device that acquired a mapping between discrete inputs and discrete outputs. Fujita (1982) proposed instead that the cerebellum worked as an adaptive filter, acquiring a mapping of continuous input signals into a continuous output signal. Essentially, Fujita's contribution was to frame the function of the granular layer in dynamic, instead of static, terms. Whereas Marr and Albus used the concept of pattern separation as the goal of granular layer computation, Fujita suggested that the granular layer decomposed the input signals into different components, very much as a Fourier or a Laplace transform does in an engineering application.

More recently, this theory has been refined into the adaptive filter theory of cerebellar computation (Dean et al., 2013). According to this theory the cerebellum acts as an analysis-synthesis adaptive filter that uses a decorrelation learning rule to acquire a given input-output mapping. The adaptive filter theory can be seen as an evolution of the classical cerebellar learning theory. Thus, by implementing an adaptive filter model of the cerebellum in this thesis we are, in substance, adhering to the updated classical theory or cerebellar learning.

CHAPTER 2

A model for the representation of time in the cerebellar cortex based on slow inhibitory currents

We begin by addressing a problem that is implicit in the control of anticipatory actions, that is, how are anticipatory actions timed. We study this in the context of the classical conditioning of the eye-blink response, since the responses acquired in such a paradigm are well-timed. At the physiological level, behavioral responses are triggered an adaptively-timed depression of the firing of Purkinje cells. Thus, the question of how the timing of this anticipatory action is acquired is reduced to where do Purkinje cells find the information to time their responses. Here our view is that such information is coded in the granular layer of the cerebellar cortex.

More precisely, we depart from previous work from our group (Verschure & Mintz, 2001), where it was assumed that the timing information was coded, at the level of the cerebellar cortex, as a decaying activity trace.

Now we refine such solution from a perspective that borrows from the adaptive filter theory (Dean et al., 2010). This is, we assume that a cerebellar Purkinje cell has access to multiple decaying traces with different temporal dynamics, not just one, and that by combining them, a Purkinje cell is able to produce an arbitrary response. Note that such an output curve incorporates not only the timing information but also the amplitude one. In other words, from this perspective, we assume that timing information and amplitude information are not dissociated.

We complete our modeling hypothesis by suggesting that the slow spillover inhibitory currents that dominate the integration of information in the cerebellar glomeruli provide the mechanism that governs the temporal dynamics of the granule cells.

This chapter reproduces a manuscript in preparation entitled "Slow inhibitory currents as a substrate for the representation of time in the cerebellar cortex". The abstract reads:

> Experimental research using the eye-blink conditioning paradigm has demonstrated that the cerebellar cortex is able to build associative memories between a conditioned (CS) and an unconditioned stimulus (US), and to generate well-timed conditioned responses (CR). Several computational models assume that the parallel fibers —the axons of granule cells— encode the elapsed Time Since Onset (TSO) of the CS. The plausibility and nature of this encoding is under debate. Here we show that the TSO can be encoded by the granule cells firing rate modulated by spillover inhibition. Namely, by the build-up of slow inhibitory currents produced by the interaction between Golgi and granule cells. Adding the assumption that the molecular layer acts as a linear adaptive filter, we demonstrate that our integrated computational model of the cerebellar cortex can trigger adaptively timed CRs, repro-

ducing the over-expectation result and CS-intensity effect. Moreover, the model confirms the crucial computational role of the intrinsically-driven spontaneous activity of Purkinje cells. Our results demonstrate that the granule cell dynamics provides a precise representation of time that can be exploited by Purkinje cells for the adaptive timing of deep nuclear responses.

2.1 Introduction

Classical conditioning of the eye-blink is widely used to study timing in the cerebellum because of the simplicity of the experimental paradigm and because of its dependence on the intact cerebellum (Yeo & Hesslow, 1998) (see Figure 2.1). It has been shown that the cerebellar cortex is necessary for the acquisition and retention of well-timed Conditioning Responses (CRs) (Hesslow & Yeo, 2002; Christian & Thompson, 2003) but the mechanisms that underlie this capacity are still not well understood.

The elucidation of the neural pathways of the US and CS (Figure 2.2) corroborated that in essence, the involvement of the cerebellar cortex in eye blink conditioning is in agreement with the Marr-Albus theory of learning in the cerebellum (Albus, 1971; Marr, 1969): The US, acting as the teaching signal, reaches the cerebellar cortex through the climbing fiber (CF) pathway. CF signals converge with the CS information conveyed by the parallel fibers (PFs). At the level of single Purkinje cells (PCs), the repeated coincidence of CS-evoked PF activity patterns and US-induced firing of the CF induces a depression of the PC response to the CS. This response causes the closure of the eyelid (Hesslow, 1994). Furthermore, experiments with decerebrated ferrets have shown that for Inter Stimulus Intervals (ISI) up to 1 second the cerebellar cortex can solve the problem of generating well-timed CRs (Jirenhed et al., 2007).

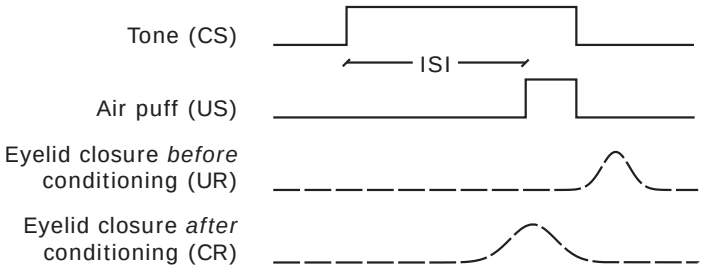

Figure 2.1: **Eye-blink conditioning**. In eye-blink conditioning an Unconditioned Stimulus (US), e.g. an airpuff to the eye, triggers an innate defensive eye-blink reflex or Unconditioned Response (UR). Through repetitive pairings of the US with an initially neutral Conditioned Stimulus (CS), like a tone, the subject develops a Conditioned Response (CR) similar to the UR (Gormezano et al., 1987; Kehoe & Macrae, 2002a). During conditioning the CS precedes the US by a fixed Inter Stimulus Interval (ISI). Crucially, the acquired CR peaks at the expected time of the US, thus matching the ISI.

Within the cerebelar cortex, the granular layer has been previously proposed as the generator of a Time Since Onset (TSO) code. Recent models rely on random connectivity in the granular layer (Medina & Mauk, 2000; Yamazaki & Tanaka, 2007a), where the apparently chaotic response of the granule cell population to the CS provides the substrate of the TSO code. However, the complex firing dynamics of the GrCs given sustained stimulation that these models predict have so far not been observed *in vivo*.

Here we describe a novel solution to the problem of the representation of time in the granular layer of cerebellar cortex, proposing that slow inhibitory currents are the biological mechanism underlying TSO encoding. The solution builds on our previous work in cerebellar modeling (Verschure & Mintz, 2001; Hofstotter et al., 2002, 2004), where we assumed that TSO was coded by a decaying trace at the GrC to PC synapse. With this study we provide a concrete implementation of such trace by

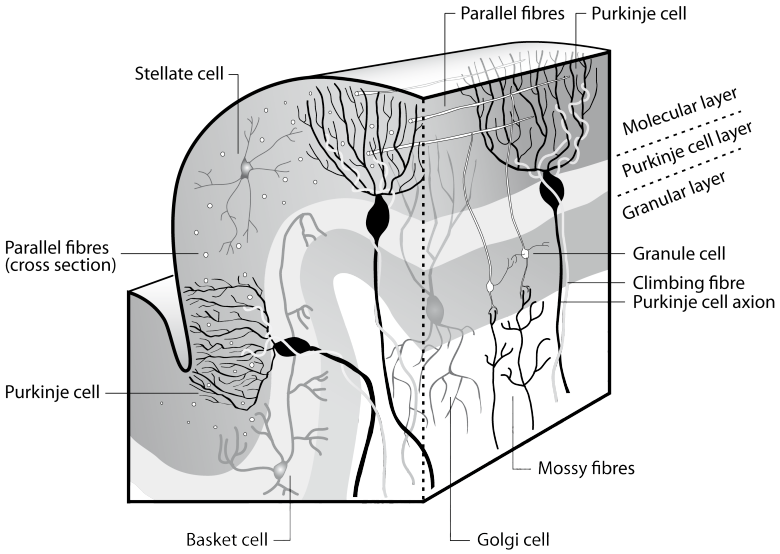

Figure 2.2: **Scheme of cerebellar cortex anatomy**. Histologically the cerebellar cortex is divided in three layers ranging from the deep granular layer to the Purkinje cell (PC) layer and the superficial molecular layer. It receives two input pathways [MFs and climbing fibers (cfs)] and projects to its target structures (the deep cerebellar nuclei) only through the Purkinje Cell (PC) axons. MFs originating in the Pons terminate in the granular layer, where they form glutamatergic synaptic contacts with the GrCs and the GABAergic GoCs. These contacts are formed in complex anatomical structures, the so-called glomeruli. A glomerulus is a structure encapsulated by a glial shell, that besides the mossy fiber boutons and GrC and GoC dendritic terminations also contains the axon terminals of the GoCs.

means of the decaying activity of GrCs, produced by the accumulation of slow inhibitory currents of the Golgi cells.

To test this hypothesis, we model the granular layer paying special attention to the time dynamics of the synaptic currents. Next, we use an abstract model of the molecular layer to validate the granular layer model in the context of eye-blink conditioning and to identify the im-

plications of the TSO code, particularly its generalization to different stimulus properties.

2.2 Methods

2.2.1 Computational model of the granular layer

Architecture Our model contains two neuronal populations, Golgi and granule cells, of Integrate-and-Fire cells (modelled as in (Shadlen & Newsome, 1998)), and a third population of cells, Pontine nuclei neurons, whose output is simulated as poissonian spike trains (Figure 2.3). The neurons of all the populations are distributed in rectangular lattices. The population sizes are listed in Table 2.1. We include excitatory connections from MF to GrCs and GoCs, and from GrCs to GoCs. Inhibitory contacts from GoCs to GrCs are modelled by two different connection types, fast and slow. We implemented the model in IQR, a large scale neural-network simulator (Bernardet et al., 2002).

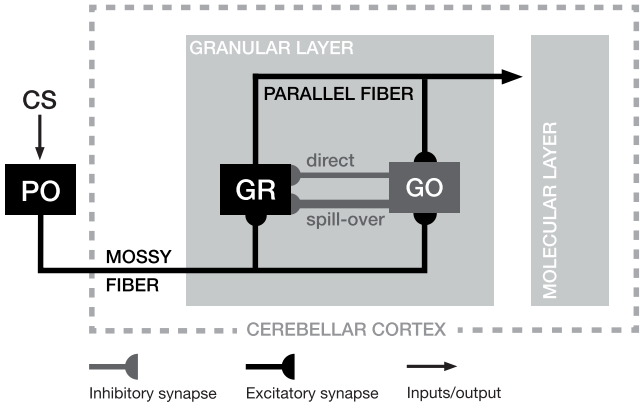

Figure 2.3: **Granular Layer model**. Excitatory connections are depicted in black, inhibitory in gray.

	τ (ms)	ρ	Firing rate σ (s^{-1})	Population Size
Golgi	20	2	10	4×5
Granule	6	1	15	40×25
Pontine (MF)	-	-	12 or 60	20×5

Table 2.1: **Neuron properties and population sizes**. τ Time constant of the membrane decay. ρ Refractory period. **Firing rate** mean firing rate, pontine cells have two values, the first for the spontaneous activity; the second for CS evoked activity. **Population size**: Size of the rectangular lattice containing the cells.

	$[C]$	$[D]$	R	τ_r (ms)	τ_d (ms)	max(A)
MF - GrC	5	50	2	0	5	0.145
MF - GoC	20	4	3	0	5	0.008
GrC - GoC	30	0.3	10	0	5	0.003
GoC - GrC (fast)	4	200	2	0	10	0.04
GoC - GrC (slow)	8	400	3	20	750	0.007

Table 2.2: **Connectivity parameters**. [**C**] Mean convergence: average number of pre-synaptic cells contacting a post-synaptic cell. [**D**] Mean divergence: average number of post-synaptic cells contacted by a pre-synaptic cell. **R** Connection radius: radius of the gaussian distribution generating the connections measured in the pre-synaptic population. $\tau_{\mathbf{r}}$: time constant for the exponential rise. $\tau_{\mathbf{d}}$: time constant for the exponential decay. max **A**: maximum amplitude of the evoked potential in the post-synaptic cell (arbitrary units).

Connections Connections are generated randomly according to a bi-dimensional Gaussian probability distribution. The parameters of the model connectivity are summarized in Table 2.2.

Neuron models To match the different neuron populations we used a neuron model (Shadlen & Newsome, 1998) with two parameters: the time constant for the decay of the membrane potential, τ, and the refractory time, ρ (see table 2.1). The subthreshold state of each cell is described by the time of the last spike $s(t)$ and the membrane potential $v(t)$. The

membrane potential is defined in arbitrary units ranging from 0 to 1, where 0 is the resting potential and 1 the spiking threshold. Such abstract membrane potential $v(t)$ can be remapped to a physiological value ($V(t)$) using $V(t) = v(t)(V_{th} - V_{res}) + V_{res}$, where V_{res} is the resting potential (approx. $-70mV$) and V_{th} the spiking threshold (approx $-50mV$).

After each time step, the new membrane potential is updated according to the decay factor δ (δ<1), plus the excitatory input $e(t)$ minus the inhibitory input $i(t)$:

$$v(t) = [v(t-1)\delta + e(t) - i(t)]^+$$

where,

$$[x]^+ = max(0, x) \tag{2.1}$$

For each neuron population the value of δ depends on time constant parameter τ and the simulation time-step size. With τ expressed in milliseconds and a time step of 1 ms, we have $\delta = e^{\tau^{-1}}$.

As indicated in equation 2.1 the minimum value of $v(t)$ was bound to 0, setting a limit to the effect of inhibitory currents. This reflects the fact that, generally, synaptic inhibition do not hyperpolarize cells below their resting potential as the reversal potential for inhibitory synaptic currents is close to the resting potential (Coombs et al., 1955).

A neuron's output, $o(t)$, is defined such that a spike is generated when $v(t)$ reaches 1:

$$o(t) = \begin{cases} 1 & \text{if } v(t) >= 1 \\ 0 & \text{otherwise} \end{cases} \tag{2.2}$$

After each spike we update the time of the last spike, $s(t)$, and maintain $v(t)$ at 0 for the duration of the refractory period:

$$v(t) = 0 \quad \text{if } t < s(t) + \rho$$

where $s(t)$ contains the time of the last spike at time t and ρ is the neuron refractory period.

Synaptic currents We describe the post-synaptic current $y(t)$ produced by a pre-synaptic action potential at time $t = 0$ as the difference of two exponentials.

$$y(t) = \frac{1}{Z}(e^{-t/\tau_r} - e^{-t/\tau_d})$$

where τ_r and τ_d are the time constants for the rise and the decay of the current, respectively, and Z is a normalizing factor. We can compute the difference of exponentials using two coupled linear ordinary differential equations:

$$\tau_d \dot{y}(t) = -y(t) + g(t)$$
$$\tau_r \dot{g}(t) = -g(t) + o(t)$$

where $o(t)$ is the output of a pre-synaptic neuron, defined as in equation 2.2.

For our simulation, we approximate the two differential equations using the forward Euler method with time step of 1 ms.

Slow currents We employ a phenomenological model to simulate the effect of inhibitory spill-over currents onto GrCs. For this we assume that for the GoC spikes to evoke a slow Inhibitory Post-Synaptic Current (IPSC) in a GrC, the extra-synaptic concentration of GABA neurotransmitter, $y_e(t)$, has to reach a certain threshold value Θ_s (Rossi et al., 2003). Once $y_e(t)$ exceeds Θ_s, the magnitude of the slow IPSC, $y_i(t)$, is defined by:

$$y_i(t) = [y_e(t) - \Theta_s]^+$$

We set Θ_s such that during baseline $y_i(t)$ is 0, implying that spill-over inhibition only affects the cell firing during the CS presentation. Moreover, we simulate the evolution of the GABA concentration again as a difference of exponentials, with a raise constant of 10 ms and a decay constant of 750 ms (Rossi & Hamann, 1998; Rossi et al., 2003).

Intrinsic currents Inhibitory neurons of the cerebellar cortex display intrinsic activity (Häusser & Clark, 1997). To model such currents phenomenologically we added spontaneous currents generated by the addition of poisson processes with a constant frequency such that we obtained a constant input with bound variability.

Simulation stability For each cell population type, we constrained our model to display anactivity similar to the one reported in literature; namely a firing rate of $15s^{-1}$ for GrC (Jörntell & Ekerot, 2006) and of $10s^{-1}$ for GoCs (van Kan et al., 1993). Since the complex architecture of the system does not allow us to find an analytical solution for setting the synaptic weights, we use optimization to achieve this goal. Concretely, we implemented a homeostatic mechanism to automatically adjust the baseline firing rate of the single neurons. For every neuron j at time t we control an internal excitability value $\epsilon_j(t)$, that acts as a multiplicative factor of the excitatory input. During the optimization we update $\epsilon_j(t)$ to minimize the difference between the estimated firing rate $\hat{\sigma}_j(t)$ of each neuron and a target firing rate σ. We estimate $\hat{\sigma}_j(t)$ over large temporal window according to:

$$\hat{\sigma}(t) = \int_0^t \phi(t-x) o_j(x) dx \tag{2.3}$$

where the filter ϕ is a normalized alpha function with a time constant of 5 seconds and $o_j(t)$ is the output of neuron j (equation 2.2). Informally, the result of the integration in equation 2.3 is an estimation of the mean firing rate in a 5 seconds interval before time t.

Therefore, after the estimation of $\hat{\sigma}(t)$ at every time step, the excitability factor $\epsilon_j(t)$ is updated according to the difference between the estimated and the desired firing rates:

$$\dot{\epsilon}_j = \alpha_\epsilon(t)(\sigma - \hat{\sigma}_j(t))$$

where α_ϵ controls the rate of change of the excitability, $epsilon_j(t)$. Therefore, to get the final $epsilon_j$ that would approximate the desired firing rates during baseline activity we apply this optimization process for 60 seconds of simulated time. During period the only input to the system, the PN cells, fire with baseline activity ($12s^{-1}$). Applying this method by the end of the optimization process he mean firing rates of the cell populations where less than a 5% away of the desired value, σ.

Simulated eye-blink conditioning In the simulation a CS is signaled as an increase of activity in a subset of MFs. Namely, a bundle of 3×4 mossy fibers increments their firing rate from 12 spikes/s to 60 spikes/s citepfreeman1999neuronal.

2.3 Results

In this series of simulations, we have addressed the physiological validity of the firing dynamics of the modeled cells, the generation of an encoding of the TSO in the granule cell population and, using a setup mimicking delay eye-blink conditioning with the goal of triggering well-timed CRs, we have tested the performance of a linear readout applied to this TSO code. Moreover, we have analysed the contribution of the CS-excited cells to the TSO code and the generalization of the whole cerebellar model to the CS-effect and over-expectation paradigm in classical conditioning.

Granule and Golgi cells response to a CS We first assess the physiological validity of the activity dynamics of the neurons in our model. After the optimization process (see Methods), cells in our simulation display baseline activity and coefficients of variation (CV) resembling data from *in vivo* experiments (Jörntell & Ekerot, 2006; Bengtsson & Jörntell, 2009a; van Kan et al., 1993). Namely, the distribution of mean firing rates for the GrCs is 10.0 ± 1.7 spikes/s with a CV of 1.02. For the GoCs the results are 16.5 ± 2.7 spikes/s with a CV of 0.67.

Next, we examine the TSO encoding at the level of single cells during a CS presentation. We first analyze a CS-driven GrC (Figure 2.4A). The cell receives increased excitation and inhibition during the CS period. However, while the excitation remains constant during the CS, the inhibition gradually builds due to the effect of the slow spill-over currents. In consequence, the cell fires strongly at stimulus onset and falls through quiescence later in the presentation (Figure 2.4A, raster plot and instant firing rate). Thus, the gradual increase in slow inhibition in the GrCs constitutes an analog trace of the TSO that modulates their firing rate.

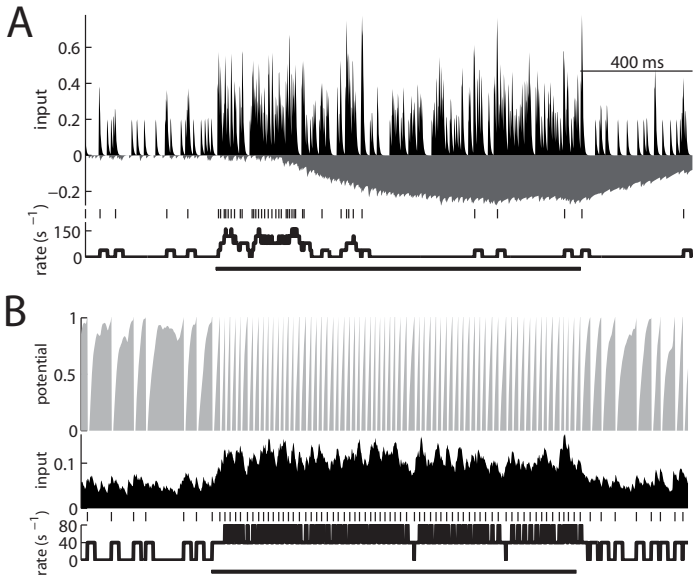

Figure 2.4: **Single trial response of modelled cells**. **A** Granule Cell. From top to bottom: EPSCs (black) and IPSCs (dark gray). Spike trace. Instantaneous firing rate using a 25 ms sliding window. The bar indicates the duration of the CS signal. Time scale indicated in the top inset. **B** Golgi Cell. From top to bottom: Membrane potential (light gray). EPSCs (black). Spike trace. Instantaneous firing rate. The bar. CS signal. Time scale as in B

We next examine a GoC from the simulation. In the absence of stimu-

lation the cell receives excitatory input from MFs and PFs, and a constant drive of excitation that simulates the effects of intrinsic currents (see (Häusser & Clark, 1997) and Methods). During the CS a constant amount of excitation adds to such baseline level, leading to an elevated firing rate that is maintained for the entire stimulus duration (Figure 2.4B).

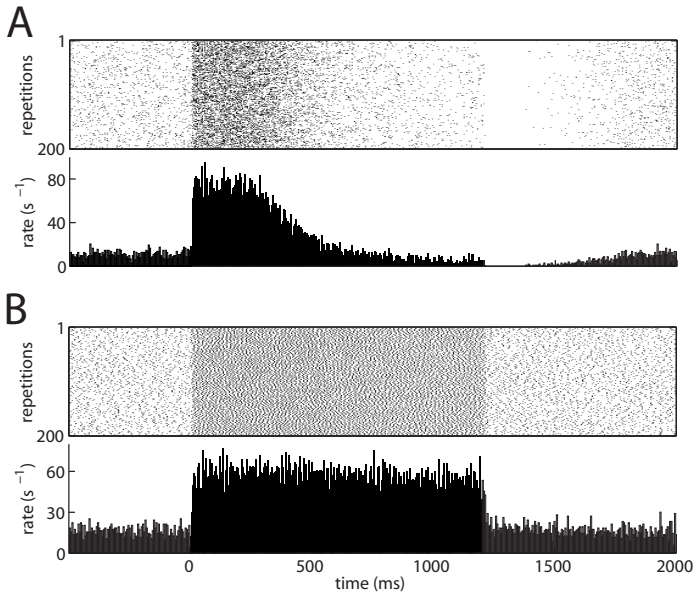

Figure 2.5: **Modelled cells response over repeated CS presentations**. **A** Raster plot of the same cell as in 2.4A during 200 presentations of a 1200 ms CS (above). PSTH obtained from the raster plot (below). **B** Raster plot of the same Golgi cell as in 2.4B during 200 presentations of the same stimulus (above). PSTH obtained from the raster plot (below); response to the CS (solid bars) and during baseline (outlined bars).

Examining the response of the same GrC and GoC to multiple repetitions of the CS we see that the GrC displays a decrease in its firing activity during the stimulus presentation (Figure 2.5A), whereas the GoC main-

tains the same level of activity during the entire presentation (Figure 2.5B).

Summing up the results so far, the large integration time provided by the slow inhibitory currents turns the sustained change in GoC activity into a representation of time at the level of the GrCs (Figure 2.4A, inhibition trace). Such inhibition causes a gradual variation in the firing rate of the GrCs that can code the TSO. From now on, since the firing rate of the GrC depends on the TSO of the CS, we refer to the characterization of a GrC response to a sustained stimulus as its *TSO tuning curve*.

Dynamics of the population of Granule cells After analyzing a single GrC, we examine the mean activity of the GrCs during a stimulus presentation (Figure 2.6). We observe that only a subset of GrCs increase their firing rate during the stimulus presentation (only 130 GrCs at least doubled their firing rate at the stimulus onset; approximately 1/4 of the population) whereas after the CS offset most of the GrCs fire below their baseline level, what indicates that CS recruited inhibition affects more GrCs than excitation. Therefore, as expected from the connectivity defined in table 2.2, a stimulus produces local excitation and global inhibition.

Importantly, due to the random generation of synaptic contacts (excitatory and inhibitory), and to the different excitability values yielded by the optimization process (see methods), different GrCs display uneven response profiles. Closely inspecting the cells excited by the CS, we observe that they present different TSO tuning curves (Figure 2.6C). From this we can interpret that individual GrCs jointly code stimulus identity and time, in agreement with the hypothesis that the function of the granular layer is to increase the representational capacity of the cerebellar cortex by producing combinatorial codes (Marr, 1969; Albus, 1971; Mapelli et al., 2010).

In addition, some of the GrCs increase their mean activity during the

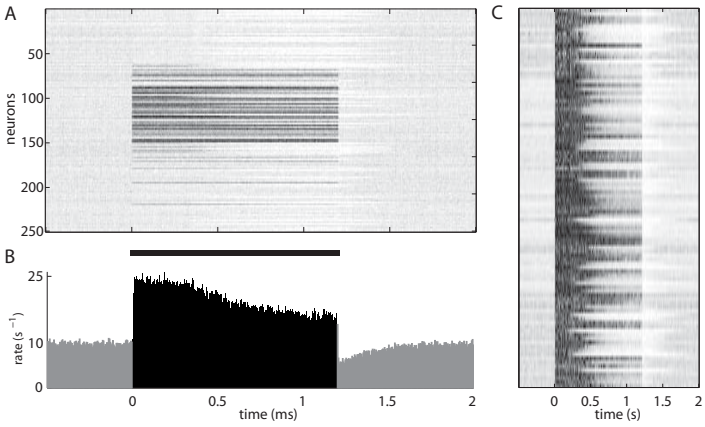

Figure 2.6: **Population response of the modelled GrCs to the CS**. **A**. Raster plot displaying the average response of 120 GrCs during 200 repetitions of a 1200 ms CS (middle, black horizontal bar). Darker colours represent higher firing rates (maximum 110 spikes/s). **B**: PSTH of 500 GrCs, including the 120 displayed above. Both graphs use 5 ms time bins. **C**. Relative raster plot of 100 GrCs increasing their firing rate to the CS presentation. The gray scale of each GrC is scaled to its maximum firing rate.

CS to 10 times their average baseline level. Considering that in our simulation CS-related mossy fibers augment their activity by a factor 5 to baseline (from 12 to 60 spikes/s) then, informally, the granular layer model increases the signal to noise level of the CS signal. This result is consistent with data regarding granular layer transduction obtained *in vivo* (Bengtsson & Jörntell, 2009a; Chadderton et al., 2004).

Thus, our GrC model yields an expanded representation of the MF signal, as predicted by the Marr-Albus theory (Marr, 1969; Albus, 1971), while increasing the signal to noise ratio as proposed by some experimentalists (Bengtsson & Jörntell, 2009a; Rossi et al., 2003)

Read out of the TSO We want to assess whether the GrC population provides a TSO that can be read out by a linear filter. Several

authors have likened the molecular layer computation (or the PC computation) to a filter combining linearly the PF input (Walter & Khodakhah, 2006, 2009; Dean et al., 2010; Porrill & Dean, 2008; Brunel et al., 2004). Moreover, the adaptive-filter theory proposes that given a decorrelation learning rule, the molecular layer can obtain a linear readout of the GrC population such that it minimizes the difference between the molecular layer output and a desired output function (Fujita, 1982). Based on this our computational model for the molecular layer is a weighted summation obtained by least-squares optimization. Likewise, we use the set of TSO curves as the basis for the generation of output functions alike in shape to the CR response. Precisley, CRs are modelled as rectangular pulses lasting 100 ms. The position of the response is determined by an ISI parameter, starting 100 ms before the expected US. For this, we first define a function, $g(t, t_{CS})$, that produces a single CR-like response:

$$g(t, t_{CS}) = \begin{cases} 1 & \text{if } t_{CS} + ISI - 100 \leq t < t_{CS} + ISI \\ 0 & \text{otherwise} \end{cases}$$

where t_{CS} is the onset time of a CS. If we let T_{CS} be a list with all the CS onset times, then we can define the desired output as superposition of all CRs:

$$b(t) = \sum_{t_{CS} \in T_{CS}} g(t, t_{CS}) \tag{2.4}$$

It is important to note that $b(t)$ is also defined outside of the CS period, where the desired output is 0.

Next, we define the following linear system:

$$Gw = b + \gamma \tag{2.5}$$

where G is a matrix with the output of 500 granule cells during 200 trials using 25 ms time bins. Importantly, the matrix contains a continuous recording activity where the CSs were delivered at the times stored in

T_{CS}. b is a column vector with $b(t)$ read out each 25 ms and γ is the error term. w is a column vector with the linear weights of the 500 GrCs. Our goal is to find a the value of w that minimizes the norm of the error term, γ. We denote this optimal solution by $\hat{w}$.

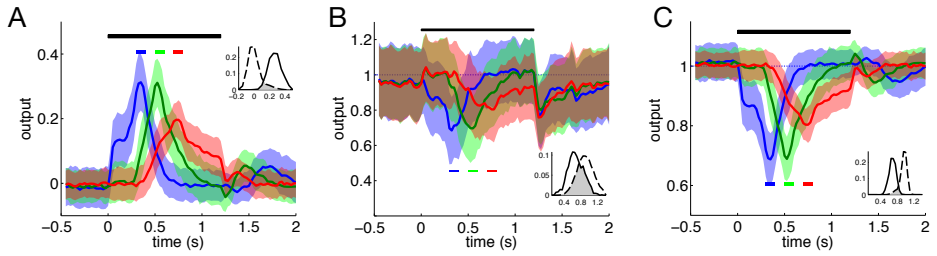

Figure 2.7: **Read-out of the time representation by linear combination**. **A**. Result of mapping the output of 500 cells during 200 trials to three different CRs: 300 ms (blue), 500 ms (red) and 700 ms (green). Mean and standard deviations. **Insets**. Output probability for within CR period (solid) and outside it (dashed) **B**. Same as in **A** but setting as target a pause response. **C**. Same as in **B** but adding a constant input bias.

Once we obtain $\hat{w}$ we evaluate the resulting mapping $G\hat{w}$. Examining the results for three different ISIs, we see that for each ISI the mapping produces a mean response that peaks at the middle of the desired CR, thus correctly preceding the onset of the expected US (Figure 2.7A). Moreover, since the variability during the baseline is small compared to the mean change between baseline and CR, ensuring a good decodability of the signal. We assess this with the Uncertainty Coefficient (UC) (Theil, 1972). For this, we consider the whole system as a noisy channel, where the information to be transmitted is b and the output signal is $G\hat{w}$. If we evaluate this measure for the 500 ms ISI we obtain an UC of 0.61 (see the output probability distribution for the 500 ms ISI in Figure 2.7 inset). This implies that our cerebellar model, containing only 500 GrCs, provides 60% of the information needed in order to encode CRs with 100% precision. Indeed if we evaluate the model using a simple

maximum likelihood classifier we see that, respectively, 91% of the CRs and 95% of the non-CRs are correctly read-out. Therefore, we conclude that the quality of the TSO code provided by the GrC population can be read-out by linear combination, allowing for the generation of adaptively-timed CR-like responses.

At this point, we aim for a behaviorally valid output from our model of the molecular layer. For this we mimic the PC response after eye-blink conditioning (the PC-CR) that consists of a diminution or total cessation of the spiking output. To obtain an output b with such a shape we redefine $b(t)$:

$$b(t) = 1 - \sum_{t_{CS} \in T_{CS}} g(t, t_{CS}) \tag{2.6}$$

where individual CR responses provided by $g(t)$ have a subtractive effect on a constant baseline output of 1. With this definition of b we get the new $\hat{w}$. As desired, $G\hat{w}$ produces responses that reach their minimum in the middle of the target interval, but now the trial to trial variability is greatly increased (Figure 2.7B). Consequently, if we consider the same 500 ms ISI, the UC drops to 0.1, meaning a loss of 80% of the information if we compare to the previous result (Figure 2.7A). How can we explain such a drop in performance? The system described in equation 2.5 requires non-null values in G to produce a non-null output. Hence, having 1 as the default output requires to constantly weight the values in G to produce it. This explains the worst performance in Figure 2.7B compared to Figure 2.7A; informally, whereas in the first case the system solves just one task, namely generation of well-timed CRs, in the second it is solving two: CR generation plus maintaining a constant non-null output. Moreover, from a physiological perspective, if we consider Gw as a model of the molecular layer, it makes a wrong modeling assumption, implying that PCs need GrCs activity to produce any output. Instead it is known that PCs in a slice fire action potentials even with excitatory transmission blocked (Häusser & Clark, 1997). Put it differently, with

equation 2.6 we have constrained the output to be more ecologically valid, without having constrained the model accordingly. We correct this by expanding G with an additional column of 1s and adding an extra weight w_β.

$$[G\,\mathbf{1}]\begin{bmatrix} w \\ w_\beta \end{bmatrix} = b + \gamma \tag{2.7}$$

where the added vector of 1s acts as an independent source of activity. This system can be rewritten as $Gw + w_\beta = b + \gamma$, where we see explicitly that w_β acts as an output bias setting the default output (or a spontaneous level of activity). With equation 2.7 we can solve together for the *synaptic weights* (the original vector w) and the output bias (w_β). From a mathematical perspective, the latter system is a multi-linear regressor with a constant term whereas the system in equation 2.5 has no constant term. The $G\hat{w}$ mapping obtained (Figure 2.7C) is a mirror version of the first result in Figure 2.7A. Quantitatively the results are almost identical. Namely, for the 500 ms ISI we get again a UC of 0.5. Furthermore, the value w_β is 0.98 and the mean output during baseline is 0.99, entailing that in our model the contribution of spontaneous GrC firing to the spontaneous output of the molecular layer is only of 1%. Concluding, if the molecular layer acts as a linear filter of the activity of our model GrC population, a good approximation of the desired output requires an independent source of input, allowing to decouple spontaneous PC output from spontaneous GrC activity.

From now on we will analyze the properties of the cerebellar cortex model comprising the detailed granular layer implementation, and the molecular layer with the input bias (equation 2.7) where the output is a PC-CR (equation 2.6).

Labeled line code A fundamental aspect of the TSO representation is whether the timing information for a given stimulus-response association is local (restricted to GrCs activated by the stimulus) or global

(involving the whole population). The latter possibility exists since CS-recruited inhibition also reaches GrCs that do not receive CS-recruited excitation (Figure 2.6). To answer this question we obtain a new mapping using only the one third of cells with the greater increase in their firing rate at the CS onset (namely, the cells that respond to the CS with at least a 2.5-fold increased discharge). The resulting map produces an output comparable to the one obtained with the whole population (Figure 2.8). Using the 500 ms ISI as our benchmark, we obtain a UC of 0.54 (vs. a UC of 0.61 with the whole population), meaning that with only a third of the cells we retain 89% of the information. In consequence, GrCs that are not strongly excited by the CS contribute only marginally to the CR generation. This implies that, first, if different stimuli excite non-overlapping sets of PF their read-outs can be be acquired independently, which favors the stability of learned associations. Secondly, if only strongly excited cells participate in the read-out the learning mechanism can be gated to synapses with high frequency activity. This last condition is congruent with the requisites of PF→PC LTD and LTP (Wang et al., 2000), and likely with long term plasticity at the PF→MLI synapse (Jörntell & Hansel, 2006) as well.

Generalization to other conditioning paradigms Drawing an analogy with the adaptive filter theory of the cerebellar cortex, in our model the granular layer implements an analysis step decomposing the signal in multiple components (the TSO tuning curves), and the molecular layer implements the synthesis step recombining the components to obtain the desired output. Notably, in our model, the components generated in the analysis step have particular dynamics. For instance, the TSO curve of a GrC depends on the MF intensity, and via Golgi-released inhibition, it can be additionally modulated by the activity in distant GrCs (see Figure 2.13A & B). Here we address the question whether these dynamics are sufficient to reproduce two benchmarks of classical conditioning: (1) the CS intensity effect (Pavlov & Anrep, 1927; Svens-

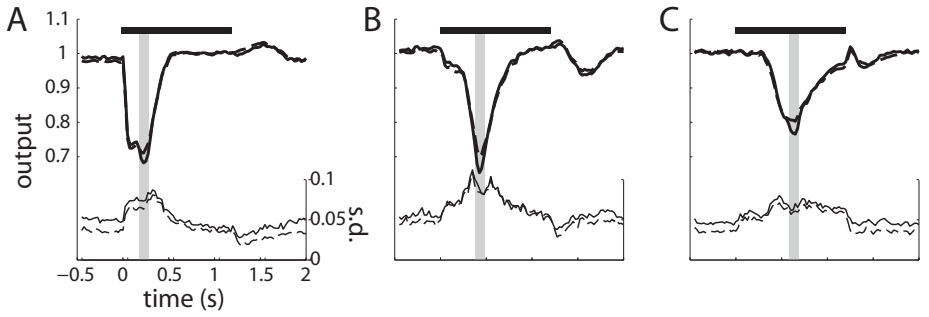

Figure 2.8: **Labeled-line code**. Result of mapping the output of all 500 cells to a timed pause response (thick solid line) compared to the response obtained only with the 33% of most active cells (thick dashed line). The thin lines indicate the standard deviation (sd). **A**. 300 ms ISI. **B**. 500 ms ISI. **C**. 700 ms ISI.

son et al., 2000) and (2) the over-expectation effect (Pavlov & Anrep, 1927; Kehoe & White, 2004).

Both experimental paradigms have in common that the subjects are tested with CSs similar but not equal than the one(s) used for training. Thereby, these paradigms study generalization of the learned response to unseen stimuli. To map this contingency to our model we introduce a simplifying notation. We denote with $G(CS_1)$ the response of the granular layer to the stimulus CS_1, and $\hat{w}_1$ is the least-squares solution of the system $G(CS_1)w = b + \gamma$, where b is the desired output defined as in equation 2.6. Then, once the cerebellar model has been *trained* with CS_1, we check its response to an unseen CS_2 by producing $G(CS_2)\hat{w}_1$. Notice that since the readout stays constant, any variation in the response relies on the dynamics of the Golgi-granule system.

CS intensity effect The CS-intensity effect is observed when a subject is conditioned with a stimulus of a given intensity (e.g.: a tone of 65 dB) and then tested with the same stimulus with an increased intensity (e.g.: an increased loudness). Under such circumstances the latency of

the CR relative to the CS onset shortens and the response likelihood and amplitude increase. Notably, this result has been reproduced at the level of PC-CRs, using MF-stimulation as a CS and mapping the increased stimulus intensity to a greater charge injection (Svensson et al., 2000).

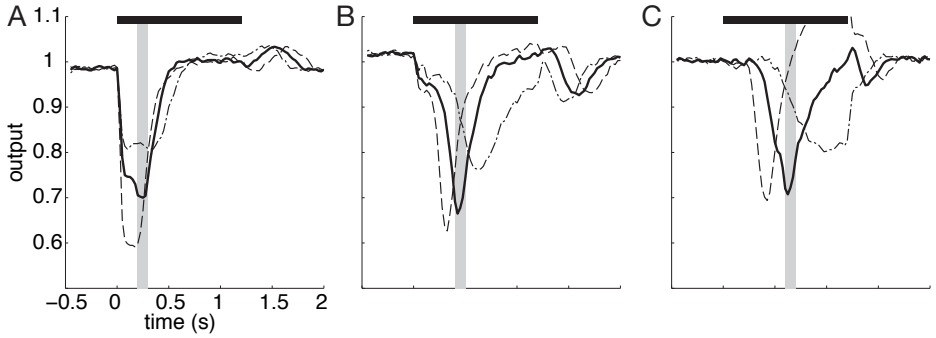

Figure 2.9: **CS-intensity**. Mean response for the normal intensity CS (thick line), higher intensity CS (dashed line) and lower intensity CS (dash-point line). The horizontal black line indicates the CS presentation and the grey vertical region indicates the PC-CR used for training. **A**. 300 ms ISI. **B**. 500 ms ISI. **C**. 700 ms ISI.

To reproduce the CS-intensity effect in our set-up we map different intensities as different firing rates of the set of MFs coding the CS. The CS sub-indexes indicate the rate of these MFs in spikes per second: namely CS_{45}, CS_{60} and CS_{90}. After getting the response of the granular layer to 200 repetitions of each separate stimulus, we find the optimal mapping for the stimulus of medium intensity, $\hat{w}_{60}$. Next, we check the output this mapping produces when applied to the lower and higher intensity stimuli; namely, $G(CS_{45})\hat{w}_{60}$ and $G(CS_{90})\hat{w}_{60}$. The result agrees with the physiological data: the response to the CS of higher intensity displays a shift in the peak timing and a modulation of the amplitude, with an earlier and deeper drop compared to the original response (Figure 2.9), whereas the response to the lower intensity CS produces a delayed and shallower decrease. In summary, the model reproduces the results of the

CS-intensity effect both in amplitude and latency modulation (Figure 2.11A).

Over-expectation The over-expectation effect is obtained when the subject is conditioned using two different CSs that have an association with the same US (Pavlov & Anrep, 1927; Kehoe & White, 2004). During training, trials with CS_a and CS_b are interleaved such that the subject produces CRs to each individual stimulus. Once performance is stabilized the subject is presented a compound stimulus comprising both CSs (CS_c). In a first phase, the result is similar to the CS intensity; the likelihood of the response to the compound CS increases. However, if the presentation of the compound stimulus is sustained for a sufficient number of trials (or sessions) the earlier responses are (partially) extinguished whereas the response to the compound stimulus is stabilized (Kehoe & White, 2004). Since literature regarding the over-expectation in eye-blink conditioning has only reported the results in terms of response likelihood, it is unclear the effect in terms of latency modulation. However the measures of responding in eye-blink conditioning covariate (Garcia et al., 2003), and there is an inverse relation between amplitude and latency during acquisition, extinction and generalization. For this, it is likely that also in over-expectation an increase in the response amplitude may be accompanied by a shortening of its latency.

To reproduce the over-expectation effect we perform two experiments. With the first we check the model's response to the compound stimulus (CS_c) after separate acquisition of both CS-US associations (CS_a and CS_b)[1]. We refer to this experiment as the S→C stage of the over-expectation. Now, the response to the compound stimulus ($G(CS_c)\hat{w}_{a,b}$)

[1]Namely, to simulate the separate acquisition we solve the system

$$\begin{bmatrix} G(CS_a) \\ G(CS_b) \end{bmatrix} w = \begin{bmatrix} b \\ b \end{bmatrix} + \gamma$$

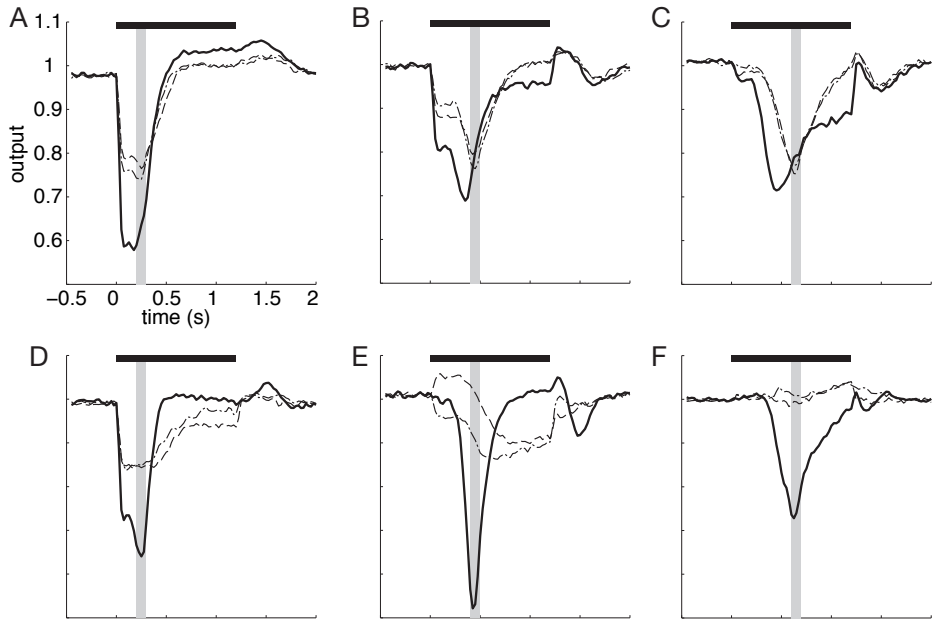

Figure 2.10: **Over-expectation**. Mean response for single CSs (thin dashed lines) and the compound CS (thick line) after acquisition with the single CSs. **A**. 300 ms ISI. **B**. 500 ms ISI. **C**. 700 ms ISI. Mean response for single CSs (thin dashed lines) and the compound CS (thick line) after acquisition with the compound CS. **D**. 300 ms ISI. **E**. 500 ms ISI. **F**. 500 ms ISI. The horizontal black line indicates the CS presentation and the grey vertical region indicates the PC-CR used for training.

displays an increased amplitude and shortened latency (Figure 2.10A,B and C).

For the second stage of the over-expectation we should have employed an incremental learning algorithm to properly reproduce the gradual extinction of the response of the single CS due to training with the compound stimulus. However, for the sake of simplicity, we approximate this result by checking the responses to CS_a and CS_b after having performed acquisition with the compound CS. Consquently, we make the simplify-

ing assumption that given extensive training with a compound stimulus, the result is independent of the earlier single CS associations. We refer to this experiment as the C→S stage of the over-expectation. The results show that $G(CS_a)\hat{w}_c$ and $G(CS_b)\hat{w}_c$ display shallower responses. In both cases, the modulation is greater than for CS intensity result, and, notably, for the longer ISI, the responses to the individual CSs are completely gone (Figure 2.10D,E and F). Indeed, the suppression of the single CS responses is so pronounced that it is not possible to quantify their latency (see Figure 2.10D,E and F and Figure 2.11C). Summing up, the results regarding the two phases of the over-expectation paradigm correctly reproduce the response amplitude modulation and predict a latency effect (Figure 2.11B & C).

Different substrates of the amplitude and latency effects The CS-intensity effect and the over-expectation paradigm allow us to find a qualitative difference among the molecular layer models with and without input bias. With standard classical conditioning both models yielded an equal response pattern, with a well-timed drop in the mean response, and only the performance decreased, albeit by a 80% (see Figure 2.6B and C). However, in the CS-intensity and over-expectation setups the results of the model without input bias are correct only as long as the latency modulation is concerned, but not in terms of the amplitude modulation. Briefly, stronger CSs do not yield a CR of increased amplitude, despite shortening the latency of the response (Figure 2.12).

On the one hand, in the light of these results, we can explain the modulation of the response latency by the dynamics of the spillover currents. In short, manipulations of the stimuli yielding higher discharge of the MFs cause an increased accumulation of inhibition released by GoCs, which anticipates the decay in activity of the GrCs (Figure 2.13A & B). The individual dynamics of the GrCs TSO tuning curves are transferred to any response built by their combination. In Figure 2.13C we display a

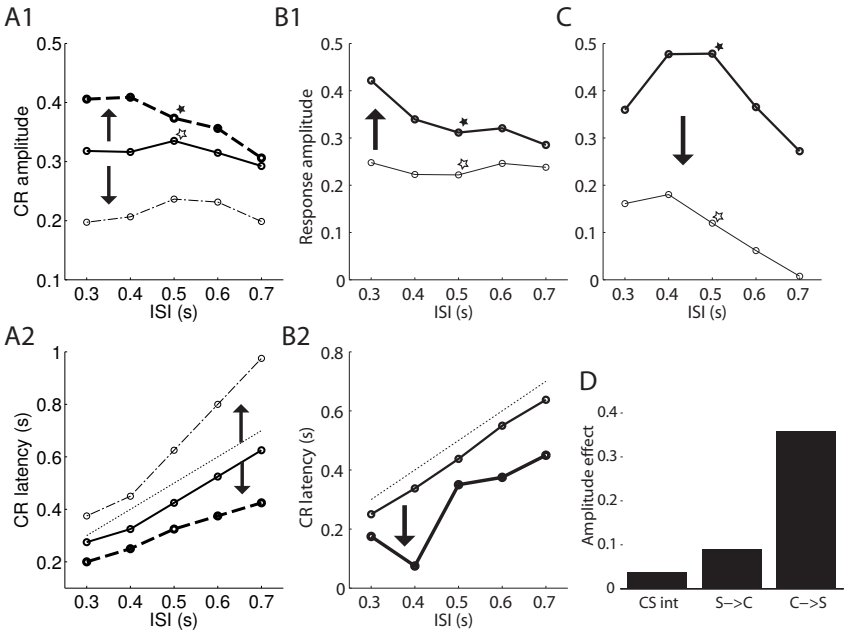

Figure 2.11: **Summary of amplitude and latency effects**. *CS-intensity.* **A1** Peak amplitude of the CR (mean). Each line corresponds to low (dashed-point), medium (solid line) and high (dash) stimulus intensity. Model tested with 5 different ISIs. **A1** Mean latency to the peak amplitude of the CR and US latency (dotted line). Same stimuli as in panel A1. *Over-expectation* **B1** Stage S→C. Peak amplitude of the CR (mean) for the single CS (mean of both CRs) and for the compound CS (thick line). **B2** Latency effect for the same stimuli in B1. **C** Stage S→C. Same stimuli as in B1. In panels A1 to C the arrows indicate the sense of the generalization (from trained to unseen stimuli). **D** Range of the modulation of the CR, computed for the 500 ms ISI using the responses marked with a star in the upper panels.

response produced by the combination of two cell outputs where we observe an inverse relation between the stimulus intensity and the latency of the downwards deflection.

On the other hand, these results suggest a necessary role for the input bias in the reproduction of the amplitude modulation. Note that in Figure

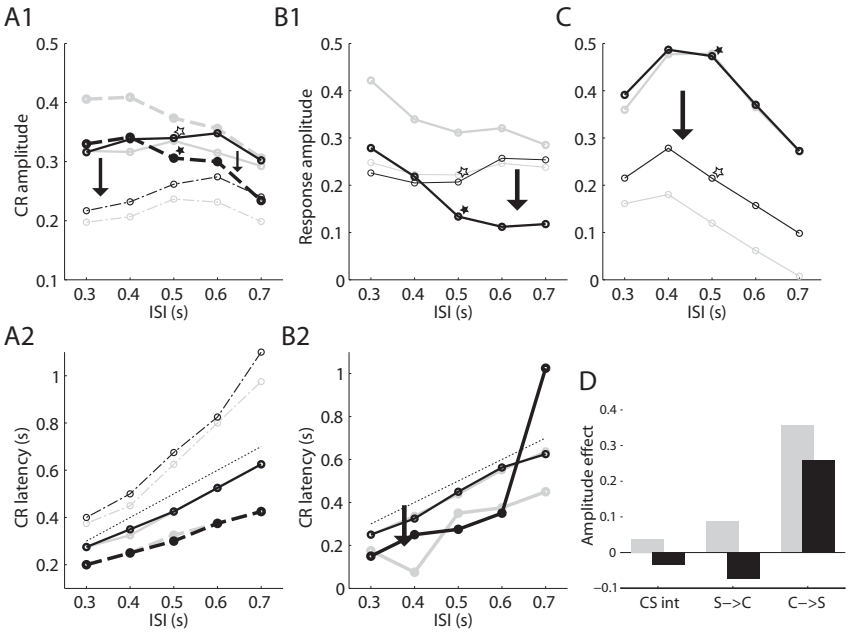

Figure 2.12: **Summary of amplitude and latency effects in absence without the input-bias term**. Results from the same experiments as in Figure 2.11 plotted for the model without bias term. For comparison, the outputs of the model with bias appear in light gray.

2.13C the amplitude increases when we increase the stimulus intensity. If we consider that such response is added to a default output level, then it is clear how the amplitude effect is obtained in the model with input bias (see Figure 2.13D). Namely, given that the discharge of the GrCs is proportional to the stimulation intensity and given the modulatory role of their output, then the more the stimulus intensity the more the modulation, what yields a higher response amplitude. On the contrary, if GrCs instead of modulating *build* the whole PC output, then increasing the stimulus intensity will increase the entire PC response curve (see Figure 2.13E for an explanation). In this case, if we compute the amplitude of the response relative to the baseline output level, a higher stimulation results in a lower PC response amplitude. We will return to this argument

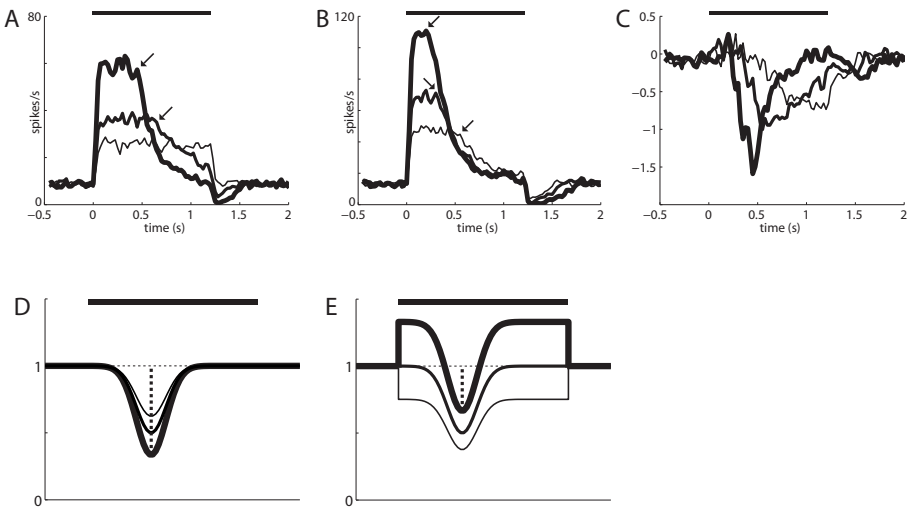

Figure 2.13: **Explanation of the modulation and latency effects**. **A & B**. Response of single GrCs to stimuli of different intensities: low (thin line), medium and high (thick line). The onset of the downwards deflection in the discharge is marked with an arrow. **C**. Response built by a linear combination output of the two previous cells. The coefficients are computed such that the minimum of the response to the stimulus of medium intensity is at -1. **D & E**. Schema of the output of the molecular layer to stimuli of different intensities in a model where the output of the GrCs has a modulatory effect on the discharge rate of the PCs (**D**) *vs* a model where they sustain the whole discharge (**E**). The vertical doted line indicates the amplitude of the response to the stimulus of higher intensity. In all panels the stimulus intensity is represented as in **A**. The horizontal bar on the top of each panel signals the presentation of the CS.

in the discussion.

2.4 Discussion

We have addressed the question how the TSO can be represented in the cerebellum in order to support its learning capabilities. We have demonstrated that the slow inhibitory currents arising from the spill-over of the neurotransmitter GABA in the glomeruli provide a substrate for the representation of time in the cerebellum, thereby explaining the distinct capability of the cerebellar cortex to generate well-timed responses in the classical conditioning of the eye-blink. We have shown that such a representation can be decoded with a linear readout and that the GrCs that are excited by the CS carry the timing information. Finally, we have reported that the combined model of the granular and molecular layers reproduces the CS-intensity effect and the over-expectation result. With these last results we have disentangled how the slow inhibitory currents and the intrinsic activity of the PCs mediate different aspects of CR generalization.

The key feature of the model presented here is the role of slow inhibitory currents in the representation of time. Slow inhibitory currents in the cerebellar GrCs are well-studied. They arise because of the spill-over of the neurotransmitter GABA in the synapse from GoCs to GrCs. Due to the tight glial encapsulation (Eccles et al., 1967), neurotransmitter diffusion within the glomeruli is restrained, such that spill-over molecules have an increased probability of activating non-local synapses (Rossi & Hamann, 1998). Slow inhibitory currents have been reported both *in-vitro* (Rossi & Hamann, 1998; Hamann et al., 2002; Rossi et al., 2003) and *in-vivo* (Jörntell & Ekerot, 2006). Crucially for our model, the cumulative effect of trains of slow inhibitory currents has been recently demonstrated in-vitro by Crowley and colleagues (Crowley et al., 2009). In summary, our model provides a functional role for what appears to be key aspects of the physiology of the cerebellar glomeruli.

More generally, any slow current endows a neuron with the capacity of

integrating information over long time intervals. Previous models of the representation of time in the granular layer included slow currents in the form of NMDA or mGluR currents (Medina et al., 2000; Medina & Mauk, 2000; Yamazaki & Tanaka, 2007a). Although these excitatory currents have been reported *in vitro* (Mitchell & Silver, 2000b), their significance *in vivo* is still unclear. Recently, another kind of extra-synaptic current generated by the spill-over of neurotransmitter has been shown to promote prolonged activity in GrCs over hundreds of milliseconds. Namely, inhibitory mGluR2 expressed in GoCs activated by GrC-released glutamate (Holtzman et al., 2011). Therefore it seems plausible that spill-over currents play a major role in cerebellar function supporting sub-second timing.

Our computational model relies on random connectivity to obtain a diversity of GrC responses. *In vitro* studies report that GrCs receive a varying number of MF and GoC contacts, which results in a variety of response dynamics to short lasting stimulations (Mapelli et al., 2010). Intracellular GrC recordings identified other putative sources of response diversity such as the variability in synaptic gains (Bengtsson & Jörntell, 2009a) and the variability of the temporal dynamics of inhibitory synaptic currents (Crowley et al., 2009). We expect that this diversity of responses exists in the GrCs of the cerebellar cortex in order for the ensemble of GrC TSO curves to contain a large set of independent basis, allowing the generation of complex output responses such as the PC–generated CR by linear combination .

Random connectivity is a common ingredient of a number of models of the representation of time in the granular layer (Buonomano & Mauk, 1994; Medina & Mauk, 2000; Yamazaki & Tanaka, 2007a,b). However such models are inspired on the, so called, liquid–state machine (Maass et al., 2002). Essentially, a liquid–state machine requires a network with positive feedback, that in this case would be the cerebellar cortex. This is at odds with cerebellar anatomy since the cerebellar cortex is mainly a feed-

forward structure whose only feedback connections are inhibitory. Moreover, from a physiological perspective, these models assume that GrCs have irregular, bursty and/or delayed responses to stimuli which, to our knowledge, still lack experimental support (even though detailed computational models have predicted similar dynamics (D'Angelo & De Zeeuw, 2009) and see (Holtzman et al., 2011)). In addition, we have shown previously that liquid state machines are very sensitive to noise (Knusel et al., 2004).

In contrast to these models, we have made more conservative physiological assumptions. First, we assumed a labelled line code for the GrCs where the cells coding the CS are only the ones that increase their activity at CS onset (fig 2.8). We assumed that given tonic excitation from MFs, GrCs sustain an increased firing rate with a progressive attenuation. To our knowledge there is only one *in-vivo* study that reports the response of the GrCs to prolonged stimulation. In this case somatosensory stimulation results in a sustained but decreasing response (Jörntell & Ekerot, 2006). Of course, in such a study the mechanism of the depression might be adaptation occurring upstream to the cerebellum, but nonetheless these GrC dynamics are fully consistent with our model. Furthermore recent work has shown *in vitro* that the build up of slow inhibition in a GrC depresses its spike output after but not at stimulus onset (Crowley et al., 2009).

We assumed that GoCs increase their activity during the CS. It has been shown that GoCs can sustain high firing rates for hundreds of milliseconds in response to the initiation of voluntary movements (van Kan et al., 1993) but so far their most typically reported response to sensory stimulation consists in a short excitation followed by a long lasting depression (Holtzman et al., 2006, 2011). Therefore, the answer to the plausibility of sustained GoC responses *in vivo* in the case of auditory stimuli (or electrical MF stimulation) remains unclear. However, we can make two inferences. First, even though our simulations contain just 20

GoCs we are able to produce well timed responses, thus it might not be necessary to have a large population of GoCs with a sustained response to obtain a stable and robust TSO representation. Second, even in the case of short phasic GoC responses, they also elicit slow inhibitory currents *in vivo* (Jörntell & Ekerot, 2006). In this case, sustained summation of slow currents would not occur, as our hypothesis requires. However, this finding does advocate a role of slow inhibitory currents in any hypothetical representation of time in the granular layer.

We have shown that in our model most of the timing information can be decoded from the GrCs receiving strong sustained CS-excitation. Thus, as long as learning is concerned, the cerebellar cortex acts as a high-pass filter of the information coming through the MF pathway. In classical conditioning, this implies that before recruiting the cerebellar cortex, a CS has to produce a strong response in the brain areas upstream to the cerebellum, i.e.: the auditory cortex for a tone CS. This gives a functional interpretation to the two-stage theory of classical conditioning (Konorski, 1967). We predict that the GoCs act as a gate of these extra-cerebellar representations of the CS. Only responses that are sufficiently boosted will effectively drive the granular layer and thus the molecular learning machinery (Boele et al., 2009; Inderbitzin et al., 2010).

As already noted by Eccles, a cell with a spontaneous level of activity can code information bi-directionally by increasing or decreasing its default firing rate (Eccles et al., 1967). Additionally, if such a cell has the capability of fixing its spontaneous rate independently of the activity of its efferent neurons, then it is reasonable to infer that its coding properties will be enhanced. We had exploited this feature in our previous model of cerebellar learning (Verschure & Mintz, 2001). The physiology of the PCs justifies the introduction of a source of activity other than direct GrC output (Häusser & Clark, 1997). Our results highlight the significance of an independent input bias, while drawing an exciting analogy between intrinsic activity and the constant term of a multi-regression

system (fig 2.6). Moreover, in our model the generalization inherent to the over-expectation and CS-intensity effects requires such an independent input bias, specifically for the amplitude effect. Interestingly such a generalization appears to entail prior assumptions underlying adaptive avoidance behaviour; e.g.: in the case of the CS-intensity effect, once a particular stimulus is identified as noxious, in this case a given CS, its intensity can be interpreted as a predictor of perilousness, or time to contact, therefore the more the intensity, the faster and bolder the protective action. Our results suggest that the intrinsic activity of PCs may be fundamental to implement this latter generalization. This prediction could be tested by measuring the CS-intensity effect in transgenic mice having disrupted PC intrinsic activity. Precisely, we hypothesize that despite showing correct acquisition, these genetically-modified mice will not respond to a stimulus of an increased intensity with an increased amplitude.

2.5 Conclusion

Several lines of evidence indicate that the cerebellar cortex has the distinct capacity of solving the temporal extent of the CS-US association in delay eye-blink conditioning for ISIs in the sub-second range. Here we have provided a novel explanation for this capacity based on a unique feature of the cerebellar cortex: the pervasive presence of slow inhibitory-currents originated by the spill-over of neurotransmitter within the glomeruli.

CHAPTER 3

Nucleo-Olivary Inhibition balances the interaction between the reactive and adaptive layers in motor control

In the previous chapter we addressed the problem of the coding of the passage of time in the cerebellar cortex and we showed that it is possible to generate such a code taking into account the slow dynamics of the spillover inhibition. Here we exploit such an encoding framework while testing a cerebellar-based controller in a simulation of an avoidance learning task.

Instead of modeling just the granular layer of the cerebellum, now we use a computational model that includes cerebellum and inferior olive. However, all structures are modeled at a higher level of abstraction. In what follows we summarize the key features of this chapter and how it relates to Chapter 2:

1. We implement a rate-based model of the granular layer of the cerebellar cortex that reproduces the profile of the output signals from the spiking model previously shown. We refer to these signals, that simulate the output of granule cells, as cortical bases. Despite having simplified the generation of each cortical basis, we replicate the basic dynamics of the responses in Chapter 2, i.e., at stimulus presentation each basis has a fast rise in activity followed by a slow decay (see 3.2).

2. The computational controls a simulated robot. Therefore, the teaching signal is derived from the interaction of the robot with the world, i.e., it is computed as a function of the proximity of the robot to the walls, and as such, it varies during the simulation contingently on the robot's learned behavior.

3. To mimic the signaling properties of the inferior olivary-cells, a Poisson process maps the real valued proximity signal into a binary code. This binary code is superimposed over a further Poisson process that generates spontaneous activity with a physiological rate of 1 Hz.

4. Crucially, we include in the controller design the deep nuclear inhibitory projections to the inferior olivary cells, i.e., the Nucleo Olivary Inhibition (NOI). It is known that such connection is instrumental in allowing the cerebellum to erase acquired responses that are no longer adaptive. Here we show that the behavioral consequence of such a connection in avoidance learning is that in the trained agent, the reactive and adaptive actions are blended, and that the exact balance of both actions is governed by the gain of the NOI.

In short, our approach here is to study how the anatomy of the cerebellar microcircuit studied in classical conditioning constrains the behavior of a robot in an avoidance learning task. This chapter reproduces a paper

with its same title, which was published at the journal Neural Networks (Herreros Alonso & Verschure, 2013a). The abstract reads:

> In the acquisition of adaptive motor reflexes to aversive stimuli, the cerebellar output fulfills a double purpose: it controls a motor response and it relays a sensory prediction. However, the question of how these two apparently incompatible goals might be achieved by the same cerebellar area remains open. Here we propose a solution where the inhibition of the Inferior Olive (IO) by the cerebellar Deep Nuclei (DN) translates the motor command signal into a sensory prediction allowing a single cerebellar area to simultaneously tackle both aspects of the problem: execution and prediction. We demonstrate that having a graded error signal, the gain of the Nucleo-Olivary Inhibition (NOI) balances the generation of the response between the cerebellar and the reflexive controllers or, in other words, between the adaptive and the reactive layers of behavior. Moreover, we show that the resulting system is fully autonomous and can either acquire or erase adaptive responses according to their utility.

3.1 Introduction

The execution of an avoidance action seems to involve both sensory prediction and motor control: the prediction of a noxious stimulus triggers an anticipatory motor command. A similar division between sensory prediction and actuation is also found in control theory when a forward model provides predicted feedback to a feedback controller (Miall et al., 1993). In this latter case, the reactive commands of the feedback controller are caused by the factual feedback anticipated by the forward model. On the contrary, in the case of an avoidance action, a common sense interpretation suggests that the predicted sensory event is *counter-*

factual, i.e., not the factual sensory event is predicted but the one that would be perceived if the avoidance action is not carried out. Here we will show that to understand the role of the cerebellum in Avoidance Learning (AL) one might have to drop this assumption.

Acquisition of anticipatory responses has been extensively studied with the paradigm of pavlovian classical conditioning (Pavlov & Anrep, 1927), e.g., classical conditioning of the eyeblink reflex (Gormezano et al., 1987) (henceforth, eyeblink conditioning). In eyeblink conditioning a neutral cue such as a sound or a light, the Conditioning Stimulus (CS), precedes by a fixed time-interval the delivery of a noxious Unconditioned Stimulus (US) to the eye, e.g., a periorbital electric shock. The US occurrence triggers a reflexive protective action (the closure of the eyelid) that constitutes the Unconditioned Response (UR). After a number of paired CS-US repetitions, the subject reacts to the delivery of the CS by closing the eyelids in anticipation of the expected US, i.e., producing a Conditioned Response (CR) (Pavlov & Anrep, 1927; Mackintosh, 1974; Gormezano et al., 1987). Once acquired, CRs can be deleted by extinction training, i.e., presenting CSs not followed by the US.

There is broad agreement that the substrate of learning in eyeblink conditioning is located in the cerebellum (Christian & Thompson, 2003; Yeo & Hesslow, 1998). The well known cerebellar circuitry (Eccles et al., 1967) helped to accurately delineate the neural pathways of CS, US and CR (Steinmetz et al., 1985; Mauk et al., 1986). The roles of the different stimuli accord with Marr-Albus-Ito cerebellar learning theory (Marr, 1969; Albus, 1971; Ito et al., 1982): the US signal relayed by the IO reaches the cerebellar cortex through the climbing fibers where it induces plasticity at the synapses of the parallel fibers that transmit the CS information. After repeated coincidence of these two signals, the Purkinje cells –the sole output of the cerebellar cortex– acquire a response to the CS, namely, a drop in their firing activity, that drives the behavioral CR (Jirenhed et al., 2007). Moreover, as with the overt CR, extinction

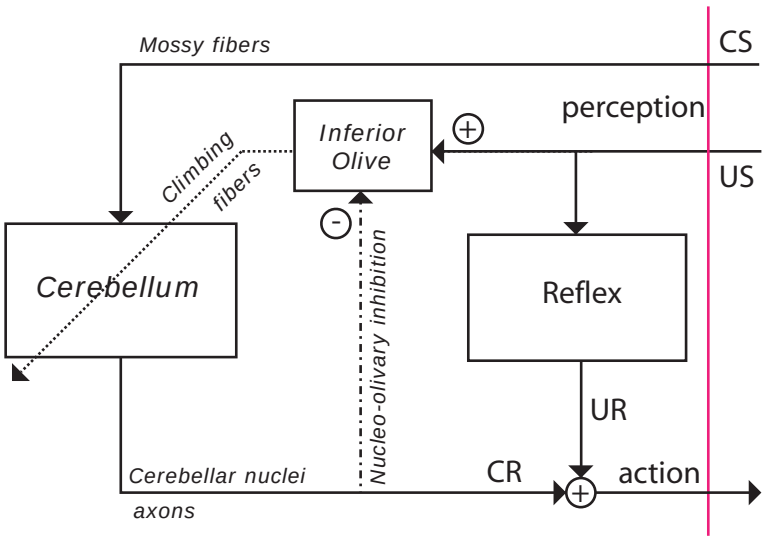

Figure 3.1: Layout of the cerebellar controller and the reflex arc. Information pathways are tagged according to the inputs outputs that they relay in classical conditioning. The labels in italic identify anatomically information pathways and processes. For a detailed description of the cerebellar architecture see (Eccles et al., 1967).

training suppresses the Purkinje cell response.

Learning in classical conditioning regards sensory prediction. As the Rescorla-Wagner model formalized, animals learn in classical conditioning only when events violate their expectations (Rescorla & Wagner, 1972). Therefore, to support this kind of learning the cerebellum must acquire and generate sensory predictions. In general, according to the adaptive filter theory, cerebellar learning is explained in terms of decorrelation (Fujita, 1982). A corollary of this theory is that the cerebellum only learns when the IO activity is perturbed from baseline. In this context, the inhibitory connections from the cerebellar deep nuclear cells to the Inferior Olive, the Nucleo Olivary Inhibition (NOI) (Andersson et al., 1988), are key to interpret cerebellar learning as the acquisition

of sensory predictions. The NOI subtracts the cerebellar output relayed from the US signal reaching the IO, such that if both signals match, they cancel each other leaving IO activity at baseline. Therefore, in eyeblink conditioning, if after the CS either the excitation produced by the US or the inhibition produced by the CR (via the NOI) outweighs the other, the perturbation of the IO activity would recruit cerebellar plasticity such that in following trials IO activity will remain closer to baseline. Remarkably, the NOI has an unusual long latency for a monosynaptic transmission in the order of the tens of milliseconds (Hesslow, 1986).

Regarding motor control, it is well-established that the output of the cerebellum drives the CR (Hesslow, 1994). In itself, this does not contradict the sensory prediction interpretation if the predicted US stimulus and the amplitude of the CR are correlated (although it is not obvious why such a correlation should exist). In other words, since correlation between neural activity and stimulus intensity –or action amplitude– is interpreted as evidence for the neural activity *coding* the stimulus –or the action– then, in classical conditioning, the cerebellar output may code both the perception *and* the response if perception and response are themselves correlated. However, the question remains whether the NOI, fundamental for sensory prediction, is functional from a motor learning perspective. AL, which, as a paradigm, is closely related to classical conditioning, serves us to address this issue.

In a classical conditioning preparation the CR is required to not ameliorate or reduce the noxiousness of the US. For instance, with a peri-orbital shock US, the CR has no effect in reducing the intensity of the shock. In AL, the CR modifies the effect of the US. For instance, if we use an airpuff without restraining the eyelids, then the effective or sensed intensity of the US will decrease as the CR increases the degree of eyelid closure, i.e., the noxiousness of the US will diminish as it reaches a more protected cornea. Therefore, whereas in classical conditioning the CR and the US can only be compared internally (and by means of the NOI),

in AL an implicit *comparison* between action and stimulus takes place in the external world. This difference between classical conditioning and AL is not always explicitly made in the literature, since some eyeblink conditioning studies, specially with humans subjects, are made with an airpuff and an unconstrained CR (Clark & Squire, 1998).

However, attempting to apply the cerebellar microcircuit studied classical conditioning to a task of AL raises a series of questions.

First, if the cerebellum outputs a motor command and the IO receives a peripheral sensory signal, then the NOI performs a non consistent comparison between information from different domains. In such a case, why should the temporal profile of the signal masking a phasic US be similar to the motor command controlling the eyelid muscles? (Lepora et al., 2010). Note that the same inconsistency of the temporal dynamics appears when we consider the avoidance of a noxious stimulus as a comparison performed in the external world. E.g., the temporal profile of the eyelid closure and the physical US stimulus might be different.

And second, AL introduces a contingency between the motor action and the sensory prediction: the CR diminishes the effective intensity of the US. We refer to this link as the behavioral negative feedback loop in contrast to the internal negative feedback provided by the NOI. But if the behavioral learning can avoid the US, what is the use of the internal negative feedback? Remark that in cases where avoidance can be complete (to hit against a wall or to completely avoid it) the role of the NOI is not evident, i.e., since both negative feedback loops are superposed, the NOI might halt learning before it leads to the total avoidance of the US.

However, it has been shown both with modeling and animal preparations that inactivation of the NOI prevents extinction in classical conditioning (Medina et al., 2002; Bengtsson & Hesslow, 2006). Extrapolating this result to AL, then the NOI has the functional role of suppressing acquired responses that are no longer adaptive. Therefore, even though

it could be possible for a cerebellar microcircuit lacking the NOI to optimally acquire an avoidance response in AL, such circuit would require an extra-cerebellar brain structure to generate the signal driving extinction. In other words, in the absence of an external signal reflecting the cost of an unnecessary avoidance action, this signal, playing the role of a hypothetical 'negative US', has to be computed internally, and the NOI provide a means for its generation.

We propose that the key to reconcile sensory prediction and motor control lies in the nature of the US signal. Considering a graded rather than an all-or-none US signal, the NOI can halt learning once the US intensity drops below a certain *safety* level, that is, once the US is as mild as to lose its noxiousness. Moreover, this residual signal can play an important functional role, i.e., in a trial-by-trial basis, it can validate the suitability of the anticipatory action. For instance, in the case of AL of the eyeblink response, once the eyelids are closed, perceiving the airpuff confirms the suitability of keep triggering the anticipatory action the next time the CS is presented.

To summarize, we propose that the NOI allows balancing the level of control between a reactive and an adaptive layer. We validate this proposal in a series of simulations where a robot has to perform a collision avoidance task in a track with a single turn. For the adaptive layer we use a controller based on the anatomy of the cerebellum (Fig. 3.1) (Eccles et al., 1967). Using the principles behind adaptive filter modeling of the cerebellum, we implement an analysis-synthesis filter with a decorrelation learning rule (Dean et al., 2010). With this setup, we study the effects of different parametrizations of the NOI gain, showing that it fixes the balance between reactive and adaptive actions, and that besides being required for extinction, the NOI is fundamental for a correct timing of the adaptive responses.

3.2 Methods

3.2.1 Cerebellar model

The model of the cerebellum consists of a set of parallel cerebellar microcircuits, each one connected to its IO component. Each cerebellar microcircuit encapsulates information processing from cerebellar cortex and cerebellar nuclei together. The inputs displayed in Fig. 3.2 correspond to the mossy fiber and the climbing fiber pathways, and relay the cue and the error signals, respectively. In nature, the output of a microcircuit module will be carried by the axons of the deep-nuclear cells (for a review of the cerebellar cytoarchitecture see (Eccles et al., 1967)).

We implemented each cerebellar microcircuit as an analysis-synthesis adaptive filter (see (Dean et al., 2010) for a review), where the information coming through the mossy fiber pathway is mapped into deep-nuclear activity and adjusted according to the teaching signal provided by the climbing-fiber signal. Even though the assumptions inherent to this model are described in a series of publications (Fujita, 1982; Porrill & Dean, 2008; Dean et al., 2010) we briefly describe them here. The cerebellar cortex acts as a filter that maps mossy-fiber activity into the Purkinje cell output. Upon entering the cerebellar cortex, the mossy fiber information is expanded into multiple components or *basis* into the granular layer. Such basis arise by means of the interaction between the mossy fibers, the excitatory granule cells and the inhibitory Golgi cells (Yamazaki & Tanaka, 2007b). Different factors such as the diversity in the connectivity between these cells (Mapelli et al., 2010) and in the synaptic gain distribution (Crowley et al., 2009) generate a repertoire of responses in the output of the granule cells. We will refer to these outputs as the *cortical basis*. Such cortical basis, relayed by the parallel fibers, serve to modulate the output of Purkinje cells. Since parallel fibers can directly excite a Purkinje cell or inhibit it by disynaptic inhibition through the molecular layer interneurons, then a same parallel fiber can

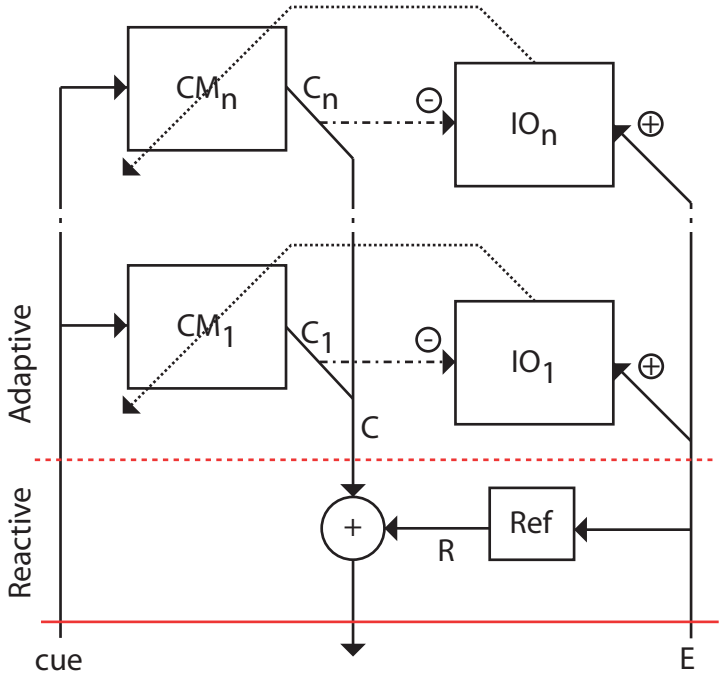

Figure 3.2: Computational architecture with the reactive and the adaptive controllers including N cerebellar microcircuits (CM), N IO components and a reflex arc (Ref). Each microcircuit generates an output command C_i. The individual commands are averaged into C and then added to R to generate the final action.

have either a positive or a negative effect in the projecting Purkinje cell. This allows the weights applied to the cortical basis in the model to have positive or negative values (Porrill & Dean, 2008). Such weights model the gain of synaptic contacts made by the parallel fibers with inhibitory interneurons and Purkinje cells. Their values are adjusted according to the teaching signal provided IO activity, that reaches the cerebellar cortex through the climbing fibers.

This implementation of the adaptive filter model includes the following novelties:

1. we use a random spike generator model for the IO with a rate of 1 Hz, consistent with the range of low firing rates observed in vivo (Eccles et al., 1967);
2. we set a series of parallel microcircuits each one with its own IO component;
3. we collapse the cerebellar cortex and cerebellar nuclei together.

The total output of the cerebellar controller is produced averaging the output of all the microcircuits.

Cortical basis To generate the cortical basis we convolve the signal coming through the mossy fiber pathway with two exponentials. In such a way, the response of each basis to a unitary pulse resembles an alpha function. The time constants governing the exponentials are randomly drawn from two flat distributions (a fast time constant, τ_r, ranging from 2 to 50 ms and a slow one, τ_d, ranging from 50 to 750 ms). The first set of fast time constants control the rise of the basis and the second, the decay. These values are within the physiological range of the time constants of the slow currents in the granular layer, e.g., slow spillover inhibitory currents (Crowley et al., 2009; Hamann et al., 2002; Rossi & Hamann, 1998).

Given an input $m(t)$, the output of a cortical basis $p_j(t)$ is generated according to the next difference equations:

$$\begin{aligned} p_j^r(t) &= \gamma_j^r p_j^r(t-1) + m(t-1) \\ p_j^d(t) &= \gamma_j^d p_j^d(t-1) + p_j^r(t-1) \\ p_j(t) &= \sigma_j [p_j^d(t-1) - \theta_j]^+ \end{aligned}$$

where j indexes a particular basis. p_j^r and p_j^d compute a convolution and, informally, each one governs the rise and decay of the p_j basis, respectively. They are controlled by the persistence factors γ_j^r and γ_j^d,

which are computed from τ_r and τ_d (see above). The third equation adds a non-linearity (a threshold θ_j) allowing to produce steeper responses (note that $[x]^+ = \max(x, 0)$). Such threshold was set to 0.7 times the maximum value attained by p_j^d when $m(t)$ carried a spike during a single time-step. Finally, all basis are scaled by σ_j such that their maximum amplitude is fixed to 1.

The output of each microcircuit is obtained by a linear combination of the its current components:

$$C_i(t) = [\mathbf{p}_i(t)\mathbf{w}_i(t)^T]^+ \tag{3.1}$$

where i indexes a particular microcircuit and $\mathbf{w}_i(t)$ is the vector of weights and $\mathbf{p}_i(t)$, the vector of basis. As in (Lepora et al., 2010) we clipped the output of the adaptive filter to remove negative values.

The weights are updated using the de-correlation learning rule:

$$\Delta w_{ij}(t) = \beta \; e_i(t) \; p_{ij}(t - \delta) \tag{3.2}$$

where β controls the learning rate and $e_i(t)$ is the error signal of the microcircuit i, computed from the inferior olive output (see below). δ corresponds to the latency of the error feedback (Miall et al., 2007). This parameter establishes how far ahead happened (or should have happened) the action that caused (or could have prevented) the current error. For this, δ resolves the temporal credit assignment problem (Sutton & Barto, 1998). In practice, this parameter sets up the relative timing of the CS relative to the US. For this, it has to account not only for the delays in transmission of the signals, but also for the latency associated to the action execution. In our case, given the temporal dynamics of the temporal basis, turns last on the order of 100 ms, making such a value a sensible choice for δ.

3.2.2 Inferior Olive

Since this work focus on the NOI, we found relevant to reproduce in our model a distinctive trait of the IO physiology: its low spontaneous firing rate. To this end, each IO component generates randomly and independently spikes with a *poisson* firing rate of 1 Hz. Having a simulation time step orders of magnitude below this rate, we generate the poissonian firing pattern converting that rate into a baseline firing probability, b. The baseline probability is then modulated at each time step both by the error signal, $E(t)$, and by the cerebellar command of the i-th microcircuit, $C_i(t)$, yielding the probability of spike in the IO at each time step, $P_i^{IO}(t)$, according to:

$$P_i^{IO}(t) = b + E(t) - k_c C_i(t - \delta)$$

where k_c is the gain of the NOI. The last term of this equation implements the NOI, subtracting the cerebellar output to the probability of firing in the IO. Note, however, that δ, the same parameter that controlled the temporal eligibility trace in Eq. 3.2, is used here to delay the effect cerebellar output into the IO. In other words, as in nature, we set up a NOI with a long latency.

3.2.3 Error signal

To compute the teaching signal term in Eq. 3.2 each microcircuit compares the instant IO activity $r_i^{IO}(t)$ and the baseline $\tilde{r}_i^{IO}(t)$.

$$e_i(t) = r_i^{IO}(t) - \tilde{r}_i^{IO}(t)$$

To compute $\tilde{r}_i^{IO}(t)$ we used exponential averaging with a time constant of 10 seconds.

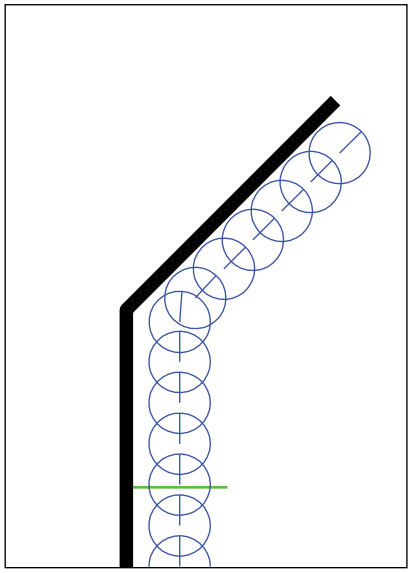

Figure 3.3: A sample trial. Trajectory of the robot during a trial, replotted every 50 time steps. The green mark in the ground is the cue signal and the thick black line is the wall.

3.2.4 Scenario

We tested our model in a simulation where a robot performs a collision avoidance task. The simulated robot has a cylindrical shape and incorporates two sensors: a proximity sensor that detects the presence of a wall and a rudimentary visual sense that allows the robot to detect a color mark in the ground. At every trial, as the robot traverses the track, it passes over a green mark in the ground (cue) and after a fixed timed interval, determined by the distance between the cue and the turn, detects the proximity to the wall and turns /(Fig. 3.3). Note that with this scenario the robot avoidance task is analogous to an eyeblink AL experiment. We already exploited a similar analogy in (Hofstotter et al., 2002).

To remark that this is a scenario of AL and not of classical conditioning,

we use a different nomenclature. Instead of a US, an error signal E is relayed by the proximity sensors. E depends both in the physical closeness, p, between the robot and the wall, and we defined it relative to the radius of the robot r as

$$p = \left[1 - \frac{|\mathbf{c} - \mathbf{x}| - r}{r}\right]^+$$

where the vector $\mathbf{x}$ vector contains robot position, $\mathbf{c}$ the closest point in the wall and the operator $||$ computes the vector norm. Note that the value of p ranges between 0 and 1. Then E is generated multiplying p by the approaching angle:

$$E = [\cos(\dot{\mathbf{x}}, \mathbf{c} - \mathbf{x})\ p]^+$$

where $\dot{\mathbf{x}}$ is the robot velocity. E can only have positive values, since negative ones imply that the robot moves away from the wall.

Instead of a CS we have a *cue* signal, which is a binary event fired when the robot crosses over the green line. The reactive response, similar to the unconditioned response is denoted as R, and the adaptive response, which is generated by the cerebellum and is similar to the CR, is denoted as C. Both responses induce a turn in the robot trajectory by controlling the angular velocity and are added to generate the total turn at each time step. The command C is generated by the already introduced cerebellar controller. The command R is generated by multiplying E by the gain of the reflexive controller (k_r).

3.3 Results

3.3.1 Acquisition of a response

Before discussing the experimental results, we highlight a key difference between the obstacle avoidance task we use and AL of the eyeblink reflex. From the sensory prediction view, in the eye-blink paradigm the target

function for the cerebellar controller to match –the sensory response to the US– varies in amplitude as the adaptive response evolves trial by trial; the noxiousness of the airpuff diminishes as it reaches a more closed eyelid. In the collision avoidance setup both the amplitude of the US – the proximity to the wall– and its timing vary. Any turn, insufficient as it might be to avoid hitting the wall, delays the collision, increasing the interval between cue and collision. Given this contingency, it is difficult to predict for a track configuration the optimal timing of the turn. However, we can predict that once the responses become stable, the relative timing of C and R will be determined by the delay of the NOI (δ, see Methods).

We display the results of a representative simulation where we use a track with a turn of 30 ° and the delay between cue and the onset of the proximity signal is of 300 ms. To maintain the analogy with conditioning we refer to this time interval as the Inter Stimulus Interval (ISI). After a few trials the controller produces adaptive responses, that during the remaining of the simulation are slowly adjusted both in timing and amplitude (Fig. 3.4A). The final responses are narrower and peak later than the initial ones (Fig. 3.4C). All adaptive trajectories are smoother than the purely reactive one, but within the adaptive trajectories we observe that the robot turns faster and closer to the wall in the last trials (Fig. 3.4B) and that by the end of training the adaptive response provides a higher proportion of the total response (Fig. 3.4D).

3.3.2 Adaptability of the responses in time and amplitude.

We show that the controller adapts to different turn angles, R_0, (Fig. 3.5A and B) and ISIs (Fig. 3.5C and D). For a same ISI, the ratio of the adaptive response to R_0 (C/R_0) remains constant regardless of R_0 (Fig. 3.5B). This is not the case for different ISIs: the larger the distance from cue to collision, the minor the ratio of the adaptive response (C/R_0) (Fig. 3.5D). In short, as in nature, performance decreases as the ISI increases.

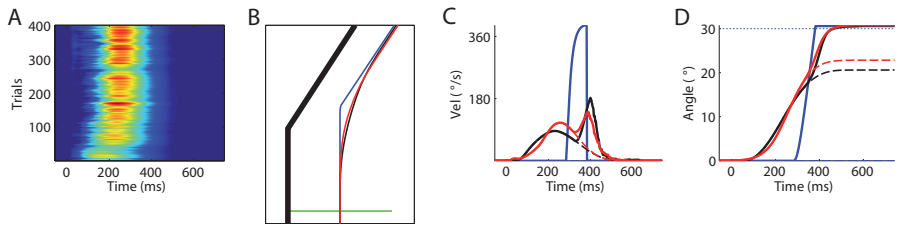

Figure 3.4: **Single simulation.** **A**. Raster plot of the adaptive responses (C) trial by trial. **B**. Three sample trajectories plotted on the track: trial 1 (blue), mean of trials 51 to 100 (black) and trials 351 to 400 (red). The cue signal (green line) and the walls of the track are shown (thick black line). **C**. Profile of the total response, $C + R$, (solid lines) for the same data in B. The dashed lines separate the adaptive component, C, (below) from the reactive, R, (above). **D**. Cumulative responses. The dotted line marks the target response (R_0)

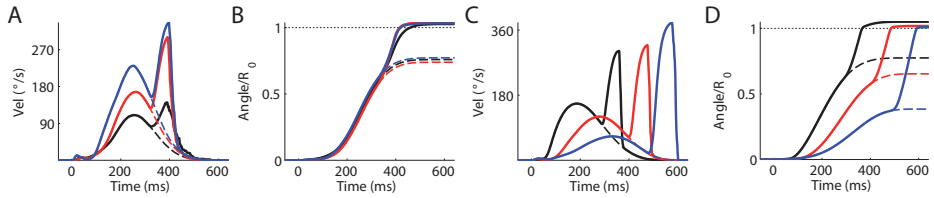

Figure 3.5: Adaptability to different amplitudes and timings. **A**. Mean responses in the last 50 trials for three different target amplitudes: 30° (black), 45° (red) and 60° (blue). The dashed lines separate the adaptive component, C, (below) from the reactive, R, (above) **B**. Cumulative responses scaled to the target amplitude (R_0). Same data as in A. **C**. Adaptive and reactive responses for three different timings: 230 ms (black), 330 ms (red) and 430 ms (blue). Data displayed and trial selected as in A. **D**. Cumulative responses scaled to the target amplitude (R_0). Same data as in C.

3.3.3 Control of the IO over the dynamics of the learning.

We validate that the controller acquires and extinguishes actions according to changes in the stimuli contingencies. For this, we run a protocol with acquisition, extinction and reacquisition stages. For the acquisition and reacquisition we use the same track, with a turn of 45 ° and an ISI of 300 ms, and for extinction we use a straight track.

In the results we observe that after the regular acquisition, the response is erased with extinction training (Fig 3.6A and B). Remarkably, during reacquisition both the amplitude of the response and the timing are more rapidly adjusted than in acquisition (Fig 3.6C). We analyze this result in detail since it goes beyond the expected properties of this adaptive filter implementation. In short, fast reacquisition might occur because after extinction the weight configuration does not return to the initial one, but to a closer configuration in the weight space (in this case a 25% closer). This relative proximity is not totally explained by the trace of the adaptive response that still remains after extinction, since the magnitude of this response, measured as area under response, is only a 17% of magnitude prior to extinction. In other words, extinction affects more the overt behavior than the underlying memory. If we only consider the first two principal components describing both the trajectory of the weights and the evolution of C during the experiment, the same results hold (Fig 3.6D). However, by visual inspection we also appreciate that after just 100 trials of reacquisition both the weights and C are very close to their configurations of the previous acquisition, suggesting a faster (more direct) gradient descent during reacquisition.

3.3.4 Behavioral role of the NOI.

We check how the different values of k_c affect the behavior. We use a track with a 45 ° for acquisition, and afterwards a straight track for extinction.

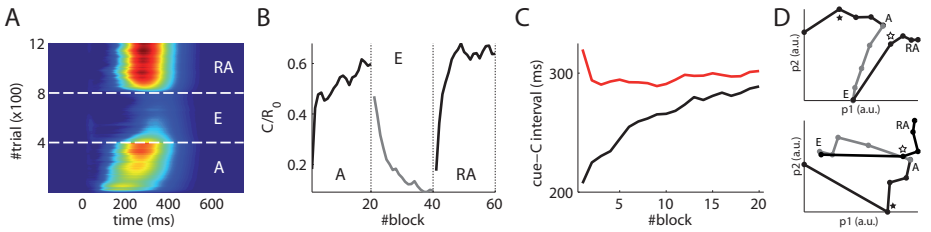

Figure 3.6: Acquisition, extinction and reacquisition. **A**. Raster plot of the adaptive response C during acquisition (A), extinction (E)and reacquisition (RA). Warmer colors for higher amplitude. **B**. Evolution of the ratio of adaptive to total response (C/R_0). Blocks of 20 trials. **C**. Evolution of the timing of the adaptive response C during acquisition (black) and reacquisition (red). The timing of C is computed as $\frac{\sum C(t)t}{\sum C(t)}$. **D**. (Up) Trajectories of the weights during learning (First two principal components). The labels identify the end of each stage. The solid star marks the position after 100 trials of acquisition and the empty star, after 100 trials of reacquisition. (Down) Trajectory of the evolution of the adaptive responses (C).

We see that the higher the value of k_c the smaller the adaptive response but the faster and more effective the extinction (Fig. 3.7A). As expected, after removing the NOI no extinction is observed. More interestingly, the NOI is required for the precise timing the responses ($Fig.3.7B$). Since the basis generated in the cerebellar cortex are not optimal to adjust the timing, part of the initial response has to be deleted or unlearned (Fig 3.4C). This latter process relies on the NOI. Thus, in absence of the NOI, the initial response, poorly timed, is maintained as long as it avoids the collision (Fig. 3.7B).

3.3.5 Transference of responses between layers controlled by the gain of the NOI.

Finally test how different values of k_c balance the adaptive and reactive responses. The values of k_c are chosen in proportion to the inverse of the

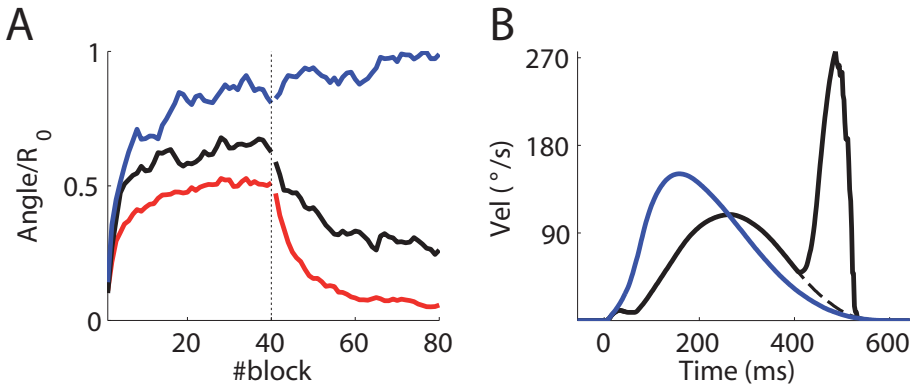

Figure 3.7: Effect of the nucleo-olivary inhibition. **A**. Accumulated adaptive response per trial proportional to the target response (R_0) in the two stages of the experiment (acquisition and extinction) for three gains of the NOI (k_c); values k_b^{-1} (red), $k_b^{-1}/4$ (black) and 0 (blue). **B**. Mean final responses C (solid) and $C + R$ (dashed) for two of the models.

gain of the reflexive controller k_r. First, we see that for a same ISI, the ratio between the adaptive and reactive actions is set by k_c independently of the angle of the turn (Fig. 3.8A). This is, at the end of acquisition simulations with a same k_c have a same proportion of the initial response transferred to the adaptive layer. Next, in a same track with a turn of 45 ° and an ISI of 300 ms, we run simulations varying the value of k_c between $2k_r^{-1}$ and $2^{-6}k_r^{-1}$. Higher values of k_c result on a minor transfer of the response to the adaptive layer (Fig. 3.8A). When k_c is equal to the inverse of k_r so that the feedforward connection of the reflexive controller and the feedback connection of the NOI apply inverse transformations, then the final response is roughly equally distributed between the adaptive and the reactive layers. As k_c becomes smaller, the maximum proportion of transfer achieved is of 0.8, i.e., the anticipatory action executes an 80% of the total turn. However, minimizing the reactive action by transferring more control to the adaptive layer comes at the expense of the timing

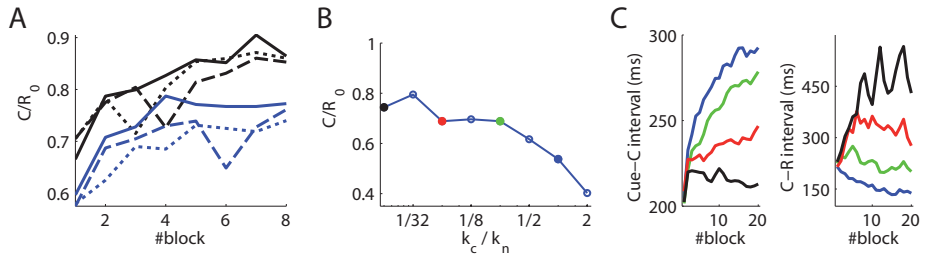

Figure 3.8: **Control of the adaptive-reactive balance by the NOI.** Evolution of the proportion of the adaptive response (C/R_0) for different gains of the NOI (k_c) and total angle of the turn (R_0). **A**. Proportion of the adaptive over the total turn (C/R_0). Angles are 30° (dotted lines), 45° (dashed lines) and 60° (solid lines). We use gains of the NOI (k_c) of value $k_r^{-1}/2$ (blue) and $k_r^{-1}/4$ (black). **B**. Final values of C/R_0 as a function of k_c in a track with a turn of 45° and an ISI of 300 ms. **C**. Timing of C. Time between cue and C. Left. Time between C and R. The values of k_c correspond to the points with the same color in B. The timings are computed as in Fig. 3.6.

accuracy of the anticipatory action. Only for the smaller values of k_c C is closely timed to the ISI, with C anticipating R by 100 ms. This is, for small values of the k_c the temporal arrangement of C and R reflects accurately the latency of the NOI δ. Summing up, k_c trades off the amount of anticipatory control against the timing accuracy.

3.4 Discussion

It is suggested that one of the functions of the cerebellum is to replace reflexes by anticipatory avoidance actions (Wolpert et al., 1998). In this paper, we proposed that the NOI might allow such a replacement to be partial, resulting in a cooperation between the reactive and adaptive layers of control. To test this hypothesis we built a computational model including a reflex controller –the reactive layer– and a cerebellar controller –the adaptive layer– implementing the latter along the lines of the

cerebellar adaptive filter theory (Dean et al., 2010; Fujita, 1982). This architecture was then applied to a simulation of a robot collision avoidance task. The results confirmed that the strength of the NOI inhibition determined the degree of replacement of the original reactive turn by an anticipatory adaptive one. Moreover, since the question of whether the cerebellum performs motor control or sensory prediction in classical conditioning still remains open (Lepora et al., 2010), such results advocate for the possibility of both roles coexisting in cerebellar function.

The originality of our approach lies in the inclusion of the NOI in a motor control scenario. On the one hand, previous computational simulations of the cerebellum that included the NOI have only been applied to classical conditioning paradigms, where the CR affects the US signal reaching the cerebellum through the NOI but not behaviorally (Medina & Mauk, 2000). This left open the question of how the NOI might interact with a behavioral avoidance response. On the other hand, simulations that used a cerebellar model in motor tasks, such as navigating a robot through a curved path (McKinstry et al., 2006), lacked the NOI feedback projection.

Multiple roles are attributed to the inhibitory projections from deep nuclear cells to the IO (Bengtsson & Hesslow, 2006): controlling the tonic activity of the whole olivo-cortico-nuclear circuit (Demer et al., 1985; Andersson & Hesslow, 1987), regulating the teaching signal provided by the climbing fibers (Medina et al., 2002; Jirenhed et al., 2007) and also adjusting dynamically the coupling between of olivary cells through gap junctions (De Zeeuw et al., 1998; Llinas & Welsh, 1993). In this paper we have only dealt with the control of the teaching signal.

3.4.1 NOI in timing and optimality of the cortical basis

Comparing avoidance responses at initial and later stages of learning, we see that the peak of the responses is delayed with learning until a

more optimal timing is achieved. This operation requires suppressing the early part of the initial responses. But to correct these initial poorly timed responses, that are already effective from a behavioral point of view, the cerebellum has to detect the mismatch between its positive output and the shape of the error signal.

However, if we had used a set of basis that was optimal for the response generation (e.g., a set of Gaussian functions of constant amplitude but increasing delays as in (Lepora et al., 2010)) then it would have not been necessary to suppress early components of the response. Interestingly, in classical conditioning the time of CRs is adjusted in a similar manner during learning in mice: training delays the peak of the CR until it matches the US onset (Koekkoek et al., 2003). Thus, using the adaptive filter stance for speculation, we can infer that the basis produced in the cerebellar cortex might not be optimal for the timing of delayed eyeblink responses.

3.4.2 NOI and extinction

The necessity of the NOI for extinction has already been shown in studies of classical conditioning both computationally and in physiology (Verschure & Mintz, 2001; Medina & Mauk, 2000). However we highlight that in the case of the learning of avoidance behavior, a controller both able to acquire and to erase adaptive responses can function in a completely autonomous manner.

3.4.3 Plausibility of a graded error signal

Our solution requires a graded error signal. We assume that the residual error either does not require correction or that the reactive response prevents any noxious consequences. In our simulations, even though we have not referred to the concepts of harm or noxiousness in the error signal, we have modeled the latter case; the residual error still induces

a turn, a reactive response. It is our view that there are AL paradigms that provide a graded error signal. For instance, in eyeblink AL, the activation of the cornea receptors by the airpuff will be more harmful that the activation of the eyelid receptors. And obviously, even if the noxious US is completely avoided, i.e., even if the eyelids are totally closed protecting the cornea, this does not precludes the eyelid receptor to sense the airpuff.

3.4.4 Cerebellum as a forward model

We have shown that in the context of adaptive reflexes, interpreting the cerebellum as a forward model helps to define the goal of learning, as it to be the exact prediction of the sensory input. However, since in this architecture the same signal supports the sensory prediction and the motor control, this solution implies that the CR amplitude and the residual US signal are proportional. This proportionality might be non-intuitive. For instance, in eyeblink AL, we might expect the successful CR to be defined in terms of degree of eyelid closure independently of the intensity of the airpuff. In contrast, our model predicts that the CR amplitude at the asymptotic level depends on the US intensity. This prediction could be easily validated experimentally, or by a further analysis of existing data.

However a question stands regarding the suitability of the controller architecture here presented. Since there is evidence of cerebellar regions acting as forward models (Miall et al., 2007), why do we not adhere to the standard forward model-forward controller architecture (Miall et al., 1993)? Quite simply, AL is not a feedback control task and for this, a forward model-feedback controller architecture cannot be applied. For instance, in collision avoidance, we are not trying to act faster after the collision by predicting it, but to avoid the collision altogether. In our case, only the reflex controller performs feedback control, but the cerebellar controller uses the feedback information to improve performance

in future trials.

3.4.5 Role of the Inferior Olive

Our computational model can acquire correct responses with the IO sustaining a physiologically correct rate of activity of 1 Hz. However, such a low firing rate slowed learning and caused performance of the individual microcircuits to fluctuate (data not shown). A problem that we practically by-passed by a adding independent IO cells to the model.

However, within each microcircuit, the low firing rate of the IO limited the gain of the NOI. In our implementation the error information resides the spiking probability of the IO (P^{IO}, see Methods). Therefore, whenever the NOI drives this probability below 0, information regarding the magnitude of the error is lost. For this, it is hard to see how can be beneficial to have a teaching signal with such a low firing rate. One view is that the output of the IO, besides acting as a teaching signal, might play other roles wherein low firing rate is beneficial. This role should be supported by the specific response of Purkinje cells to the the climbing fiber signal, namely, the complex spikes. A second explanation is that the low rate of complex spikes might prevent them to interfere with the information carried at a much higher firing rate by the simple spikes.

3.4.6 Fast reacquisition

Although not directly related to the claim of this paper, we also reported that our model reproduces the fast reacquisition observed after extinction (Jirenhed et al., 2007). Previous computational studies reproduced fast acquisition by including different learning mechanisms with different learning rates (Porrill & Dean, 2008; Medina et al., 2001). In our model fast reacquisition stems for a non-linearity in the cerebellar adaptive filter and not from the interaction of different learning mechanisms. Indeed, by the rectification applied to the cerebellar output (Eq. 3.1) ex-

tinction can be achieved without returning to the original configuration of null weights, i.e., any configuration yielding a negative output before the rectification is also effective. Such a configuration might be closer in *weight space* to the weight configuration that triggers the CRs, for which at reacquisition returning to this configuration might be achieved faster.

3.4.7 Means to adjust the gain of the NOI

As we manually set the gain of the NOI (k_c) to different values, a next step would be to devise a heuristic or learning rule to set its gain adaptively. On the one hand, our controller has to produce an adaptive avoidance response managing to decrease the error signal under a safety level. In addition, we know that the greater the gain of the NOI the smaller the adaptive response produced. For this, an heuristic should decrease such gain to allow a sufficient transfer of control from the reactive to the adaptive layer. On the other hand, the smaller the NOI the slower the extinction of no longer adaptive responses. Considering that the drive for rapid extinction might be related to the cost of unnecessary avoidance actions (for instance, closing the eyelids to protect the eyeballs results in a momentary loss of potentially relevant visual input), this cost function could be used to learn an optimal k_c. In any case, a model solving this issue would have to hypothesize which extra-cerebellar structures evaluate this cost, and how does their information reach nucleo-olivary circuit.

3.4.8 Conclusion

We have a cerebellar controller including the NOI negative feedback-loop to the acquisition of an adaptive reflex in a collision avoidance task. Within that domain, we have shown a) that we can simultaneously interpret the output of the cerebellar controller as a sensory prediction and as a motor control signal and b) that the degree of transference of the initial reactive response into an adaptive response after learning is deter-

mined by the gain of the NOI. These results require the error signal, the US, to be graded. Moreover, under these conditions the NOI controls the temporal accuracy of the responses and allows the extinction of no longer adaptive responses. With this work we have shown that from the cerebellar microcircuit studied in classical conditioning we can obtain a functional controller for AL. A next question now is whether once more realistic plants are taken into consideration, such as the muscles controlling the eyelids, this architecture is complete or additional components are necessary to simultaneously meet the requirements of sensory prediction and motor control.

CHAPTER 4

The CS-intensity effect as a built-in cerebellar sensorimotor contingency

In the previous chapter we demonstrated a cerebellar-based controller applied to an avoidance learning task. The results confirmed that the model was functional and that in the trained robot, the commands of the reactive and adaptive components are blended. We concluded that the collision avoidance learning scenario developed for the previous study offered us a test-bed where to reproduce findings of the experimental eyeblink conditioning paradigm in the context of avoidance learning.

In this chapter we advance this work in two ways. First, we move from the simulated to the physical world, to check whether the control architecture is still functional in the noisy conditions inherent to a real-world control problem. Secondly, we investigate whether an effect observed in the eye-blink classical conditioning paradigm, i.e., the CS-intensity effect, can be considered as a built-in sensorimotor contingency enhancing the computational flexibility of the cerebellum.

The CS-intensity effect is a contingency observed in classical conditioning that links the intensity of the CS with the latency of the acquired CR. Thus, the rationale behind this work is that this modulation of the response latency may allow the cerebellar model to control the execution of the same skill with different rapidities, according to the context. The skill we use here is a turn, and the necessity to modulate the execution rapidity is given by the different velocities at which the robot navigates the track.

The key features relating this chapter with the previous ones are the following:

1. The computational layout employed in this model is the same as in Chapter 3 with the difference that instead of using multiple microcircuits that receive a low bandwidth binary teaching signal we use a single microcircuit receiving a scalar teaching signal.
2. The generation of the cortical bases now explicitly introduces the slow inhibitory component that was the core of the solution proposed in Chapter 2 for the representation of time in the cerebellar cortex.
3. We employ the same collision avoidance scenario developed for Chapter 3, even though we move from a simulated to a real-world setup. Crucially, we increment the speed of the robot as the training progresses.

This chapter reproduces a paper with the title "Speed Generalization Capabilities of a Cerebellar Model on a Rapid Navigation Task" that will appear in the proceedings of the International Conference on Intelligent Robots and Systems (IROS) (Herreros Alonso et al., 2013b) to be held in Tokyo, Japan, during November 3-8, 2013. The abstract reads:

> The cerebellum is a brain structure necessary for skilled motor behavior and has a well understood and repetitive ar-

chitecture. Such an architecture inspired the Marr-Albus-Ito theory of cerebellar learning, that provides an explanation for the acquisition of motor skills by the cerebellum. Numerous computational models inspired in such a theory have already been employed in robotic tasks. Here we look into one of the suggested roles of the cerebellum, the replacement of reflexes by anticipatory actions and we apply it to a robot navigation task. The acquisition of anticipatory actions has been thoroughly studied in the field of classical conditioning. Of particular interest is the so-called CS-intensity effect, an effect that links the rapidity of execution of an anticipatory protective action, the Conditioned Response (CR), to the intensity of a predictive signal, the Conditioning Stimulus (CS). We propose that the CS-intensity effect implements a built-in sensory-motor contingency that allows to carry over a skill learned in a safe and easy context, e.g., turning at slow velocity, to a more difficult one, e.g., a turning at a faster speed. We demonstrate this hypothesis in a series of experiments where a robot has to navigate a track that has a turn. We show that after being trained at a slow velocity, by means of the CS-intensity effect, the cerebellar controller modulates the turning such that its onset anticipates as the robot speed increases. Ultimately, through incremental learning, this generalization allows the robot to learn to navigate the track at its maximum speed.

4.1 Introduction

Since the first theories of cerebellar learning were proposed (Albus, 1971; Marr, 1969), computational models of the cerebellum have been applied to robotic control tasks (Albus et al., 1975). The aim of such implemen-

tations has been not only theoretical, to validate the functionality of cerebellar theories (Herreros Alonso & Verschure, 2013a; McKinstry et al., 2006; Luque et al., 2011; Hofstotter et al., 2002), but also practical, since computational models of the cerebellum perform competitively in some domains (e.g. bipedal walking (Sabourin & Bruneau, 2005), robotic arm control (Fag et al., 1997)). Competitive applications are often obtained with models that abstract away from the cerebellar physiology. Such models reproduce 1) the overall cerebellar architecture at the level of the information pathways and 2) the error-based learning rule. However poor attention is devoted to the fine-grained dynamics of the cerebellar computation. We believe that as much as the robotic community has gained from the implementation of control systems blueprinted from the cerebellar architecture, it will gain from reproducing the dynamics of the real cerebellum. Here we look into the dynamics of anticipatory reflexes controlled by the cerebellum as studied in classical conditioning. More concretely, we look into the CS-intensity effect, that links the speed and amplitude of execution of an anticipatory reflex to the intensity of the cue signal (Svensson et al., 2010, 1997). We show that the CS-intensity effect implements a sensory-motor contingency that allows generalizing a learned motor skill along different speeds of execution. Finally, we demonstrate that this generalization enables to safely learn the rapid execution of a motor action, namely, to perform a turn at high speed avoiding collisions.

In classical conditioning the experimenter sets up a contingency between a neutral Conditioning Stimulus (CS) and a noxious Unconditioned Stimulus (US) such that the CS becomes a predictor for an upcoming threat (Pavlov & Anrep, 1927). Initially, the US triggers an innate and reflex-like protective Unconditioned Response (UR) that after training will be preceded by a similar anticipatory action, i.e., the Conditioned Response (CR) . For instance, in eye-blink conditioning a usual setup has a tone preceding a mild electric shock to the peri-orbital area of an animal by a

time interval below a second (Gormezano et al., 1987) that elicits a protective eyeblink. After a number of repetitions of such paired $CS-US$ presentations, animals develop an anticipatory blink, the CR, aiming to mitigate the harm caused by the US.

Here we have translated the delay eyeblink paradigm to a collision avoidance task. In such a task, a robot has to traverse a track avoiding collisions. The track contains a single turn, that is preceded by a series of stripes on the ground. The robot is equipped with sensors through which it detects the marks on the ground (CS) and the proximity to the walls (US). The proximity signal triggers first a reactive turn (UR_t) and, over a certain threshold, a reactive braking (UR_b). Initially, the robot is tested at a velocity for which the reactive control safely avoids collisions even though the close proximity to the wall forces the robot to brake. Thus, firstly, cerebellar learning will be expressed as an anticipatory and smooth turn, the CR, preventing the robot from reducing the velocity.

Our rationale is to apply a sensory motor contingency like the CS-intensity effect to generalize the turn learned at the slower and *safer* velocity to a faster and *dangerous* one, i.e., a velocity for which the reactive controller would not prevent the robot collision. Ultimately, this would allow the robot to learn to navigate the track at higher velocities. In order to map the speed into the intensity of the CS, we use the derivative of the visual signal, i.e., an analogous to simple optic flow signal. Likewise a faster robot will experience a more intense (albeit shorter) CS.

We use an adaptive filter model of the cerebellum with a de-correlation learning rule (Dean et al., 2010). In accordance with this model, the signal generated by the visual input is decomposed on a series of components with different temporal profiles, we call them cortical bases. The cerebellar learning consists then in finding the weights of the linear combination that maps the response of cortical bases into a correct output.

In our case, a correct output is the one producing an anticipatory turn that either allows the robot to navigate the track without reducing the velocity or, if this is not possible, to simply stay within the track boundaries avoiding collision.

According to this scheme, when the CS-intensity is increased, the variation of the CR learned at a given intensity depends on the response of the cortical bases to the new CS. For instance, if doubling the intensity of the stimulus only doubles the response of the cortical bases, but does not alter their temporal profile, the cerebellar output will only be scaled in amplitude but not shifted in time. Thus, in practical terms, the response of cortical bases responses has to be non-linear. To achieve this, we implement in our model the signal transduction mechanism that is applied in the input stage of the cerebellar cortex, i.e., at the granular layer (Eccles et al., 1967). Namely, we assume that the granule cells, the cells that code the cortical bases, act as linear thresholded filters (Spanne & Jörntell, 2013) and that their output results from the interaction of a fast excitatory component minus a slow inhibitory one (Crowley et al., 2009).

To summarize, we propose that the CS-intensity effect allows to generalize a sensory motor association learned at one speed of execution to different ones and that this generalization can help a robot to master the execution of skilled behavior at high speeds through incremental training. In order to reproduce the CS-intensity effect in the computational model of the cerebellum, we generate the cortical bases mimicking the fast excitation followed by slow inhibition signal transduction of the cerebellar cortex. Finally, we validate this proposal with a series of robotic experiments where a robot navigates a track at gradually increasing velocities.

4.2 Methods

4.2.1 Computational architecture

The computational architecture implemented two layers of control (Fig. 4.1). The reactive controller maps the proximity signals at both sides of the robot (US_l and US_r) into reflex-like avoidance turns (UR_{tl} and UR_{tr}) and braking actions (UR_{bl} and UR_{br}). In addition, an adaptive controller, implemented as a cerebellar analysis-synthesis adaptive filter controller ((Dean et al., 2010)), will eventually acquire an anticipatory turn after the CS and ahead of the UR. In these experiments, since the tracks included just a single right turn, we only implemented a controller to anticipate the turn to the right. However, we mounted reactive controllers in both sides, to ensure that the robot stayed on the track even when the anticipatory turn was overshooting.

4.2.2 Cerebellar model

The neurobiological assumptions underpinning this model have been described elsewhere (Herreros Alonso & Verschure, 2013a). We just highlight that this model implements the Nucleo-Olivary Inhibition (NOI), through which the cerebellum can compare its output signal with the sensory signal carrying the US information, and adjust its mapping until the mismatch is minimized. Additionally, here we add a slow inhibitory components to the computation of each basis computation, that allows for more precise responses.

To generate the signal of the cortical bases we produce two components for each basis: a fast excitatory and a slow inhibitory one. Each component consists of a double exponential convolution. The time constants of the convolutions for the excitatory and inhibitory component are randomly drawn from two flat probability distributions ranging from 0.05 to 0.1 seconds, and from 0.2 to 5.5 seconds, respectively. These values are

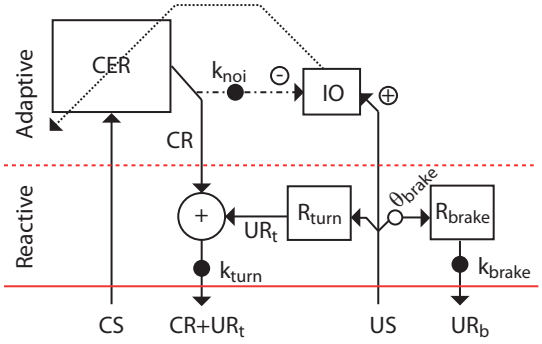

Figure 4.1: Computational architecture with the reactive and adaptive controllers. CS, conditioning stimulus (visual input); US, unconditioned stimulus (proximity signal); CR, conditioned response (anticipatory turn); $\mathrm{UR_t}$, unconditioned response (reactive turn); $\mathrm{UR_b}$, unconditioned response (reactive brake); CER, cerebellar controller; $\mathrm{R_{turn}}$, reactive turn controller; $\mathrm{R_{brake}}$, reactive brake controller; θ_{brake}, threshold for braking; $\mathrm{k_{noi}}$, gain of the Nucleo-olivary inhibition; $\mathrm{k_{turn}}$, gain of the turning command; $\mathrm{k_{brake}}$, gain of the braking command.

significantly over the physiological time constants of the synaptic currents found in the cerebellar cortex, but they are appropriated for the current experimental setup. The value obtained after the two convolutions is then thresholded and scaled for each basis. The whole computation is formalized as follows:

$$
\begin{aligned}
e_j^r(t) &= \gamma_j^r e_j^r(t-1) + CS(t) \\
e_j^d(t) &= \gamma_j^d e_j^d(t-1) + e_j^r(t) \\
e_j(t) &= \sigma_j [e_j^d(t) - \theta_j]^+
\end{aligned}
$$

where j indexes a particular basis. e_j^r and e_j^d compute a convolution and, informally, each one governs the rise and decay of the e_j basis, respectively. They are controlled by the persistence factors γ_j^r and γ_j^d, which generate the appropriated exponential decay. The third equation adds a

non-linearity (a threshold θ_j) that in the current implementation is critical for obtaining the response latency modulation by the CS intensity (note that $[x]^+ = \max(x, 0)$). I.e., without such a threshold the computational model would act as a linear filter. For each cortical basis an inhibitory component, i_j is generated with the exact same process, only using time constants 10 times larger.

The final value of the basis, p_j, is computed as follows:

$$p_j(t) = [e_j(t) - i_j(t)]^+$$

The output of the cerebellar controller is given by:

$$CR(t) = [\mathbf{p}(t)^T \mathbf{w}(t)]^+$$

where $\mathbf{w}(t)$ is the vector of weights and $\mathbf{p}(t)$, the vector of bases.

The weights are updated using the de-correlation learning rule:

$$\Delta w_j(t) = \beta \, E(t) \, p_j(t - \delta)$$

where β is the learning rate and $E(t)$ is the error signal, computed by the inferior olive output (see below). δ provides the latency of the NOI. The value of δ determines how much the adaptive action anticipates the reactive one and has to exceed the feedback delay (Miall et al., 1993). In our experiments we used a β of 0.01 and a δ of 1.0 s. We estimated that the feedback delay in our system was in the order of 0.2 seconds.

Finally, the error signal for the cerebellar system is computed as the difference between the scaled cerebellar output and the US signal as follows:

$$E(t) = US(t) - k_{noi} CR(t - \delta) \tag{4.1}$$

The factor scaling the CR is the gain of the NOI, k_{noi}. Note that through this computation, the error signal for the cerebellum is suppressed before

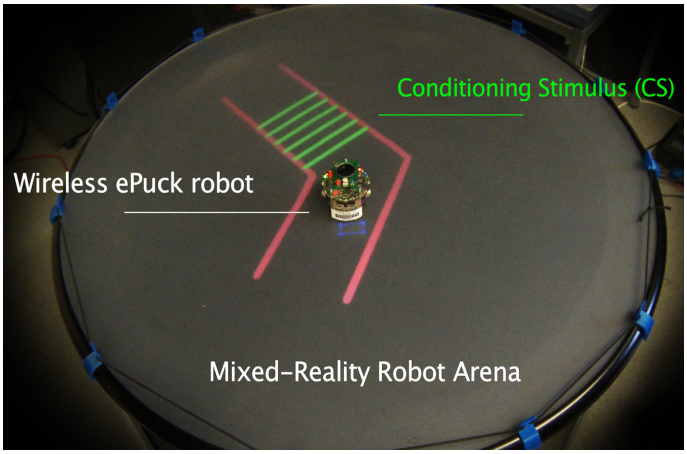

Figure 4.2: Experimental setup. Mixed reality environment with the back-projected arena and the physical robot (e-puck).

the sensory US is completely prevented, for which a fraction of the initial reactive response, UR, still prevails after training. The amplitude of this residual UR depends on the k_{noi} value and has functional implications (see (Herreros Alonso & Verschure, 2013a)). In short, with a k_{noi} of 1 the final amplitude of both actions would be similar whereas with a k_{noi} equal to 0 the adaptive response would completely replace the reactive one. In this setup we use a value of 0.4, obtaining a CR that is bigger than the UR at the end of training.

4.2.3 Experimental setup

The setup consists of an epuck robot (Mondada et al., 2009) and a Mixed-Reality Robot Arena (MRRA) (Fibla et al., 2010). The robot navigates a track back-projected onto the table displaying the CS signal as a series of green stripes. A tracking system captures the position and direction of the robot in order to compute the sensory data, proximity and visual signals. These virtual sensory signals are provided to the controller system, that then issues the appropriated motor commands to the epuck

robot via a bluetooth connection. In summary this setup mixes the constraints of controlling a real robot with the versatility of generating a virtual scenario.

Stimuli The robot is equipped with *virtual* sensors. Namely, a camera and proximity sensors. The camera allows to detect the green stripes displayed on the ground of the track. From this input, we compute the proportion of the visual field occupied by CS stimulus (the green stripes). The CS is then computed as the differential of the previous signal. With this procedure, the intensity of the CS signal is linked to the speed of the robot, since the instantaneous variation in the visual field is proportional to the robot velocity. The proximity sensors are mounted at each side, frontally and 15 degrees away from the forward direction. They have a range of 6 cm, and are normalized such that at maximum proximity their value is 1.

Motor commands The motor of each wheel is controlled independently, with a signal that blends the output of all controllers as follows:

$$\begin{aligned} M_l = \quad & M_{init} + k_{turn}(CR + UR_{tl} - UR_{tr}) \\ & -k_{brake}(UR_{bl} + UR_{br}) \\ M_r = \quad & M_{init} + k_{turn}(CR - UR_{tl} + UR_{tr}) \\ & -k_{brake}(UR_{bl} + UR_{br}) \end{aligned}$$

M_l and M_r denote the left and right commands, respectively. M_{init} sets the initial velocity, that is maintained constant for each trial. We use values ranging from 8 cm/s (corresponding to a M_{init} equal to 20 units) to 20 cm/s (corresponding to a M_{init} equal to 50 units). The reactive turns (UR_{tl} and UR_{tr}), that convey the same value as their contra-lateral US signals, are added to the anticipatory turn (CR) and multiplied by the motor gain for the turn k_{turn}. Each brake command is computed as

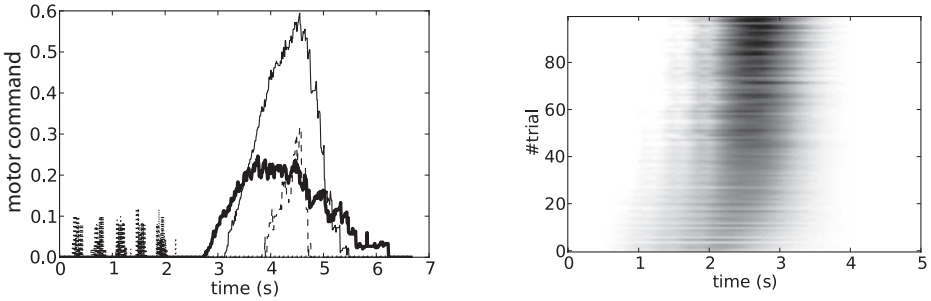

Figure 4.3: Evolution of the responses. (*Left*). Reactive commands at the first trial (turning [thin line] and braking [thin dashed line]) and at last trial (reactive turning [thick solid line], there is no braking at the last trial). The cue signals are displayed near the onset of the trial [grey dotted lines]. (*Right*). Evolution of the adaptive command. Darker color indicates higher amplitude of the response.

follows:

$$UR_{bl} = (1 - \theta_{brake})[(US_r - \theta_{brake})]^+$$

In short, each UR_b is computed as the proximity signal exceeding the θ_{brake} threshold, normalized. Both braking actions are then added and multiplied by the corresponding motor gain (k_{brake}). In our experiment we set k_{turn} and k_{brake} to values of 8 and 20, respectively, and set the braking threshold to 0.5, what corresponds approximately to a 3 cm distance from the wall.

4.3 Results

4.3.1 Experiment 1

The goal of this experiment is to train the robot to acquire a predictive turn allowing it to traverse the track without decreasing the initial velocity. The training session lasts 100 trials. We set up an initial velocity of 8 cm/s. This speed is sufficiently slow to prevent the robot from

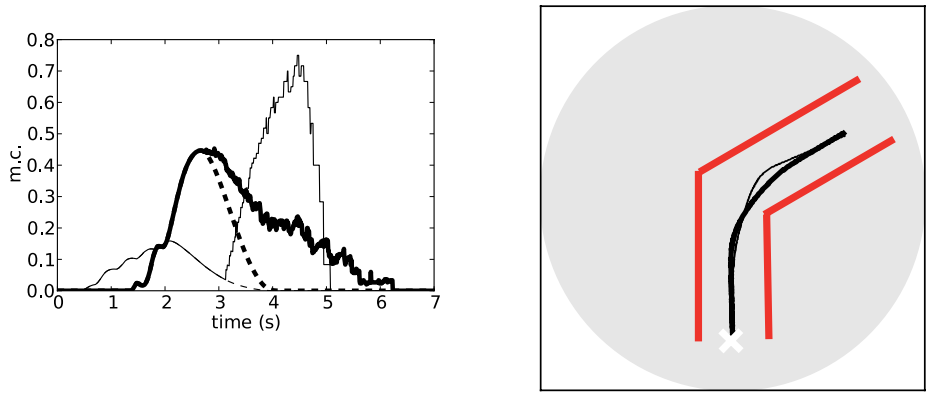

Figure 4.4: Integration of the reactive and adaptive responses. (*Left*). Total command in an early trial (thin lines) and at the end of learning (thick lines). The dashed lines separate the adaptive (below) and reactive components of the response (above). (*Right*). Trajectory of the robot for the same two trials

hitting the wall under sole reactive control. However, even at this slow velocity the proximity signal reaches the braking level (Fig. 4.3 *left*). As training progresses we observe the acquisition of a predictive turn that slowly grows in amplitude and becomes more accurately timed (Fig. 4.3 *right*). Note that at the end of the training the reactive response is not completely erased: the proximity signal still grows over the threshold causing a reactive turn, but stays below the braking threshold (Fig. 4.3 *left*). Thus, once successfully trained, the robot balances the predictive and the reactive actions such that it can traverse the track as fast as possible preserving the initial speed.

Concerning the relative timing of both actions, δ_{noi} determines how much the adaptive response will anticipate the reactive one (Fig. 4.5). We set up this parameter to 1 s, obtaining an optimal merge of both actions (Fig. 4.4). Namely, the adaptive response peaks at the onset of the reactive one, which is the textbook definition of adaptive timing in the classical

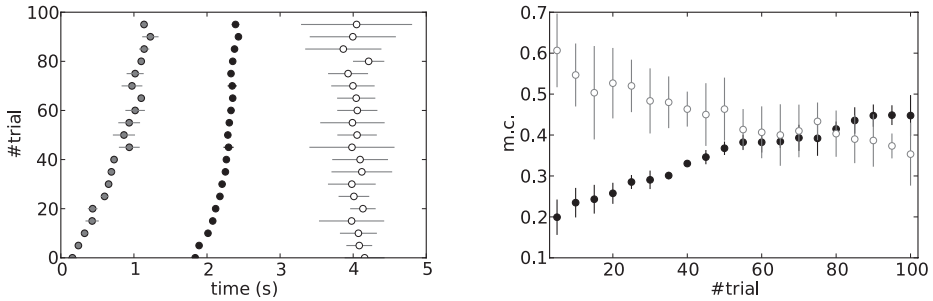

Figure 4.5: Quantification of the responses during training. (*Left*). Timings relative to the CS onset. Onset (gray markers) and peak (black) of the adaptive response and peak of the reactive one (empty marker). Mean and standard deviations computed every five trials. (*Right*). Maximum amplitudes of both responses.

eye-blink conditioning paradigm. This merge results in a final trajectory where the robot displays a single turn different from the trajectory during the early trials, where both turns can be singled out (Fig. 4.4).

4.3.2 Experiment 2

With the second experiment we assess how well the response acquired at the initial safe velocity generalizes to faster velocities. For this, we applied the cerebellar controller that we trained in the previous experiment and we froze its memory by setting the learning rate to 0. Afterwards, we increased the initial velocity every five trials by a step of 0.8 cm/s. In this way, at the trained velocity the cerebellar controller outputs the same response acquired during the training, but the response at higher velocities depends on how the increased intensity of the CS is translated into the response.

We first evaluate the experiment in behavioral terms. For this we test which is the maximum speed that the robot can achieve before colliding

with the walls. We observe that after having only been trained at a velocity of 8 cm/s, the robot can safely navigate the track at speeds up to 17.6 cm/s. Note that the highest safe speed with sole reactive control was of 14.4 cm/s (see supplementary video). Therefore, even though the robot cannot learn from scratch to traverse the track at a speed of 17.6 cm/s, because it will crash with the wall in the first trial, it can traverse the track at such a velocity if it is initially trained at the *safe* velocity of 8 cm/s. This suggests that, in principle, we can use this controller to train the robot optimally navigate this track at the speed of 17.6 if we first train it at 8 cm/s.

Now, we examine the output of the cerebellar controller to assess whether the navigation of the track at higher velocities is achieved, 1) because the exact response that was learned at the initial training speed is triggered at higher velocities but *still* facilitates the turning, or 2) because the learned response is generalized in congruence with the requirements of the increased velocity. Note that as the velocity increases, the interval between the CS and the collision shortens and that, given the dynamics of the motor plant, the amplitude of the turning command also increases (i.e., reproducing the same curvature at a higher speed requires a higher control signal). Thus, to generalize from the previous learned motor command, the cerebellum has, 1) to anticipate the adaptive response such that the robot turns earlier and 2) increase the amplitude of the motor command. We recall here that the speed of the robot is implicitly coded in the intensity of the CS.

We obtain that as the velocity increases (as the intensity of the CS stimulus increases) the timing of the adaptive response is anticipated. The anticipation is such that at all the velocities tested the adaptive response still peaks ahead of the reactive one (Fig. 4.6 *left*). The adaptive turn remains anticipatory even when the increase of speed triggers an earlier reactive turn. However, we do not obtain an increase in the amplitude of the response together with the increased velocity (Fig. 4.6 *right*).

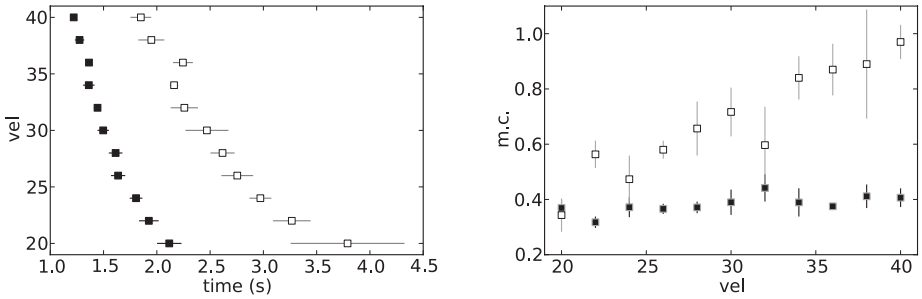

Figure 4.6: Quantification of the responses during generalization. (*Left*). Timings relative to the CS onset. Peak of the adaptive response (black) and peak of the reactive one (empty marker). Mean and standard deviations computed every five trials. (*Right*). Maximum amplitudes of both responses.

This lack of generalization of the amplitude may be the reason why, even though the learned response remains anticipatory, at some point it becomes insufficient to keep the robot on the track.

4.3.3 Experiment 3

Finally, we want to find out whether the controller allows the robot to traverse the track at its maximum velocity (20 cm/s). For this we run an experiment in which the robot is incrementally trained at higher velocities. As in the previous experiment we depart from a cerebellar controller already trained during 100 trials at the initial velocity of 8 cm/s. Whenever the robot navigates the track without braking for 5 consecutive trials or after performing 10 trials at the same velocity, we increase the velocity by a step of 0.8 cm/s. Our hypothesis is that, since the controller generalizes the response learned at a slower velocity to a higher one, learning to perform a correct turn at an increased speed would only require a fine-tuning of the initial acquired CR.

We obtain that the robot is able to navigate the track at the maximum

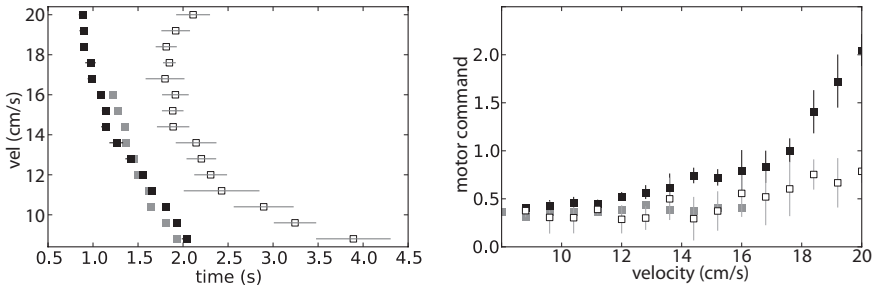

Figure 4.7: Quantification of the responses during continuous learning. (*Left*). Timings relative to the CS onset. Peak of the adaptive response (black) and of the reactive one (empty marker). For comparison, the peak of the CRs obtained in the previous experiment are also displayed (gray). Mean and standard deviations computed every five trials. (*Right*). Maximum amplitudes of both responses.

velocity, even though it cannot avoid braking for speeds above 17.6 cm/s. To assess how much the generalization from slower to higher velocities facilitates the learning, we compare the results from experiment 2, that were generalized from a single initially trained velocity, with the commands learned in this experiment, where the learning proceeds at each velocity step, assuming that a major similarity implies a better generalization. We observe that having learned to perform the turn at the velocity of 8 cm/s, the controller correctly generalizes the timing at the velocities up to 17.6 cm/s, i.e., there is no systematic difference between the timing of the generalized and the learned commands (Fig. 4.7 *left*). However, both commands are very different in amplitude (Fig. 4.7 *right*). As our controller did not reproduce the amplitude component of the generalization, the learned commands have a bigger amplitude than the generalized ones. Finally, besides delaying the collision with the wall, that in the current experiment did not occur even at the maximum velocity of the robot, learning a more precise response allows the robot to navigate the track faster (Fig 4.8).

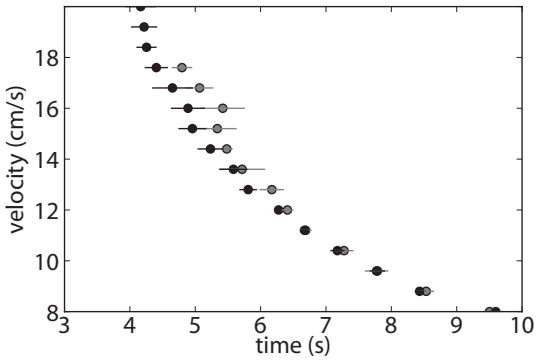

Figure 4.8: Time to navigate the track at different velocities with continuous learning (dark) or only training at the slowest speed (gray). Mean and standard deviation of the last five trials in each condition.

In conclusion, we observe that the correct generalization of the timing of the CR to higher velocities allows the robot to learn to navigate the track at its maximum velocity.

4.4 Conclusions

We have presented a control architecture inspired by the cerebellum that allows a robot to navigate a track avoiding collisions. Such controller learns to transform a purely reactive avoidance response into a more complex response that includes both anticipatory and reactive components, i.e., the CR and the UR, respectively. We have also shown that even if the robot is only trained at the slowest velocity, the generalized response is still adaptive for higher velocities, with the CR correctly anticipating the US onset. In addition, we show that this generalization facilitates the process of learning to navigate the track at velocities much higher that the ones that could be safely handled by the reactive control alone.

To achieve this result we have extended previously existing computational

models of the cerebellum (Herreros Alonso & Verschure, 2013a; Dean et al., 2010) with a method for generating the cortical bases inspired by the cerebellar physiology. More concretely, here we added two features that mimic the computation of the cerebellar granule cells: the interplay between fast excitation and slow inhibition (Crowley et al., 2009) and the non-linearity of their responses (Spanne & Jörntell, 2013). The first modification allows the system to acquire precise responses and the second, the non-linearity of the bases, namely, the addition of a threshold to the currents, is necessary for achieving the modulation of the CR latency by the intensity of the CS. Since in our set-up the intensity of the CS is linked to the velocity of the robot, this means that as the velocity of the robot increased the latency of the response was correctly advanced.

This addition allowed us to explore the generalization of an avoidance action over different speeds of execution, a feature not studied in other previous studies that used cerebellar controllers for collision avoidance (Hofstotter et al., 2002; McKinstry et al., 2006).

With this controller the reactive turn, the UR, is not completely suppressed by the CR. The amount of the final response transferred to the adaptive controller is determined by the gain of the NOI (k_{noi}) (Herreros Alonso & Verschure, 2013a). The residual UR is the only means for the controller architecture to know that the CR is necessary. For instance, if we remove the US of the current set-up (removing the turn while keeping the CS) this controller would gradually extinguish the CR. Thus thanks to the comparison between cerebellar output and sensory input performed via the NOI this controller can manage the acquisition and extinction of an anticipatory adaptive reflex in a totally autonomous way.

The proposed controller architecture is not task specific and, in general, it could be applied to scenarios having a feedback signal that has to be kept under a certain safety level. An interesting suggestion is that this

type of controller could be involved safe limb control in soft robots (Dean et al., 2013). In that case, the cerebellar controller would take care of avoiding the limbs to too strongly collide with the robot's own limbs or external objects. The major difference then in that case is that the signal playing the role of the CS will be internally generated, reflecting the robot's intent to move the limb.

Regarding the delay of the NOI, some questions remain open. First, how can this parameter be learned? And secondly, can it also be modulated? In particular, the second question has a functional relevance in our scenario, because even though the CS-CR interval is learned by the system, the CR-UR is fixed. In other words, the CR always anticipates the UR by the same fixed time interval, even if it it would be more effective for this time interval to be adjusted according to the velocity.

Thus, to conclude we have presented a controller for the acquisition of anticipatory reflexes that besides being completely autonomous, it includes a built-in sensory motor contingency that modulates the timing of the protective action according to the intensity of the predictive cue. For the first time we have implemented this controller with a real robot. The next step will be to apply this controller with more complex motor plants to validate whether this built-in sensory motor contingency facilitates learning the rapid execution of motor actions when the dynamics are more complex.

CHAPTER 5

Replacing a cerebellar microcircuit by a synthetic system

In the previous chapters we have deployed computational models onto progressively more constrained scenarios. This is, starting from a purely computational study, continuing with a computational model applied to a simulated task, ending with the application of a computational model to a real-world robot task. In all cases, we dealt with the issuing of well-timed actions. The last experimental work that we present addresses the same issue but in the most constrained scenario for a neural model: controlling a living animal via the (functional) replacement of one of its brain structures.

So far, the positive results obtained are existence proofs for a given articulated computational hypothesis of the function of the cerebellar cortex. The components of such a hypothesis are the following: 1) the granular layer produces a representation of the passage of time presynaptically to the Purkinje cells; 2) Purkinje cells exploit such a code by means of a synaptic plasticity mechanism, therefore being able to trigger adaptively-

timed responses; 3) the Purkinje cells receive a teaching signal via the climbing fibers; 4) the inhibition of the IO by the deep nuclear cells is a necessary component of the circuitry, that modulates the acquisition and extinction dynamics, and that is fundamental for the adaptive timing of the CRs. Now, we go beyond the existence proof to validate whether a synthetic system built along this computational hypotheses is compatible with the brain of a living animal. In other words, if instead of artificially generated signals we connect the model to biological signals retrieved in real-time, is our computational model still functional?

Since this test is done in the context of the development of an implantable neuro-prosthesis, further implementation constraints apply. In short, it was necessary to reduce the computational model to its functional core. For this reason, the computational model we present in this chapter is qualitatively different from the models presented so far. Namely, we have been assuming that the cerebellar output is produced as a linear combination of multiple bases, each one providing a different coding of time. Now, to allow implementation of the model in a low power aVLSI platform, we use a minimal representation of time, a single basis that via a non-linear mapping controls the issuing of a stereotyped conditioned response. Indeed, the overall design of the computational model follows the line of previous aVLSI implementations by our own group (Hofstotter et al., 2004) but adapts it to the requirement of dealing with a teaching signal that is encoded as a binary signal with a low level of spontaneous activity (this same requirement was considered in Chapter 3).

This chapter reproduces a manuscript Herreros Alonso et al. (2013a) that has been submitted to a journal from the field of neural circuits and neuroprosthetics. The abstract of the manuscript reads as follows:

> The reproduction of the computations performed by a brain structure opens the possibility of developing a neuro-prosthetic system that could replace the original structure. Here,

we demonstrate the feasibility of this approach with respect to the motor learning capabilities of the cerebellum. Specifically we show that we can sustain the acquisition, retention and extinction of the eye-blink reflex in a rat when the biological circuit that performs this task, the cerebellum, has been inactivated by anesthesia. To achieve this, we first develop a computational model of the cerebellar microcircuit involved in the acquisition of conditioned reflexes and test it with synthetic data that has the characteristics of the signals directly obtained from the brain. Later, in a closed-loop bio-hybrid preparation, we interface the model with the brain of an anaesthetized rat, feeding into the synthetic system the biosignals recorded from the cerebellar input structures and injecting back the result of the computation into an area targeted by cerebellar output. In consequence, we show that an anaesthetized rat can be classically conditioned to the acquisition of an eye-blink response with the aid of our neuro-prosthetic system. Finally, to overcome stability problems in the dynamics of the recorded biosignals, we propose an updated computational model and validate it with the experimental data recorded. The neuro-prosthetic system presented here is unique in that it replaces a function of the central nervous system receiving inputs from the brain and returning its outputs back into the brain. These results represent an early example of science based medicine where on one hand the neuro-prosthetic system directly validates a theory of cerebellar learning that informed the design of the system, and on the other, it takes a step towards the development of neuro-prostheses that could recover lost learning functions in animals and, in the longer term, humans.

5.1 Introduction

Neural prostheses between the central nervous system and peripheral systems exist from some time. Some examples are retinal and cochlear implants (Cohen, 2007; Zrenner, 2002; Wilson et al., 1991; Eddington et al., 1978), and brain computer interface systems controlling artificial limbs (Hochberg et al., 2012; Moritz et al., 2008; Schwartz et al., 2006; Chapin et al., 1999). However, the bi-directional coupling of a prosthetic system with the Central Nervous System (CNS) has only very recently been addressed (Berger et al., 2011; Giovannucci et al., 2010). Here we demonstrate the functional bi-directional coupling of an artificial system and the CNS in the context of classical conditioning.

Classical conditioning is one of the most essentials forms of associative learning (Pavlov & Anrep, 1927). In classical conditioning, an initially neutral Conditioned Stimulus (CS) precedes an aversive or appetitive Unconditioned Stimulus (US), leading to the acquisition of a Conditioned Response (CR). A widely employed paradigm in classical conditioning is eye-blink reflex conditioning where an animal is exposed to a CS, e.g., a tone, followed after a certain Inter-Stimulus Interval (ISI) by an aversive US to the eye or periorbital area, e.g. an air-puff (Kehoe & Macrae, 2002b). After repeated trials of this paired stimulus presentation, the animal closes the eyelids in anticipation of the expected air-puff; this anticipatory action is known as the conditioned response (CR). If a so trained animal is subsequently exposed to tones followed by the absence of the air-puff US (CS-alone stimulation or extinction training), the previously acquired associative CR will disappear and the tone will return to its neutral nature. Remarkably, if we repeat the initial training after extinction, the CRs are more rapidly acquired, a phenomenon known as *savings* (Napier et al., 1992).

The cerebellum is critical for the acquisition of CRs in eyeblink conditioning (Christian & Thompson, 2003; Hesslow & Yeo, 2002). The CS

signal reaches the cerebellum through the mossy fibers originating in the Pontine Nuclei (PN), while the US signal is projected through the climbing fibers originated in the Inferior Olive (IO). These two projections converge on the cerebellar Purkinje Cells (PCs), the sole output cells of the cerebellar cortex, that acquire a reduction in their activity to the CS (Jirenhed et al., 2007) that through dis-inhibition of their target, the deep nuclear cells, leads to the production of the CR.

Following these assumptions, we have previously developed a computational model of learning in the cerebellum (Hofstotter et al., 2002; Verschure & Mintz, 2001) that was implemented in aVLSI hardware and tested in a robot learning task (Hofstotter et al., 2004). Here we show how this computational model can be deployed as a prototype of a neuroprosthetic device by interfacing an implementation of the model with the brain of a living animal, replacing the function of the animal's inactivated cerebellum.

The validation proposed here entails feeding the computational model not with the artificial signals generated in the robot experiments, but with biosignals acquired through electrodes. Indeed, a very relevant difference among this so-called bio-hybrid set up and the robot one previously tested (Hofstotter et al., 2004) resides in the coding of the CS and the US: whereas in previous work, the computational model received an unambiguous CS/US signal, the signal we now extract from the animal's pons and IO is reflecting the complexity of the biological system. In particular the IO derived signals contain US-evoked events interspersed with many events caused by the spontaneous activity of the IO. Indeed, the characteristics of the signal we retrieve in our experiment is a good match with the IO physiology showing a baseline firing rate in the range of 0.5-2 Hz (De Zeeuw et al., 1998). In other words, in our experiment, the encoding of the US in the IO is very noisy. Therefore, in order to adapt the computational model to the characteristics of the experimental biosignals, our first goal is to develop a set of methods allowing the

computational model to work even when there is a low level of signal to noise in the IO channel.

In addition to providing a proof of concept of our approach towards neuro-prosthetics, our second goal is to validate the algorithms designed for the interfacing and operation of a neuro-prosthetic system in an in-vivo bio-hybrid preparation. The system under evaluation implements a real-time model of cerebellar learning that is driven by signals recorded directly from the PN and IO, detecting CS and US events respectively from these recording channels, and transduce CRs acquired by the neuro-prosthetic system by stimulation in a physiologically appropriate location in the Facial Nucleus (FN) in order to elicit a well-timed CR, with latencies matched to the biological circuit being replaced (see Prueckl et al. (2011a) for specifics on the physiology of this experiment). Since in our preparation the acquisition of natural CRs is precluded by anesthesia, any overt CRs observed in the experiment are the result from associative learning occurring solely within the synthetic system.

Finally, we address a drawback encountered in the experimental setup. Namely, during the recording the level of spontaneous activity in the IO, that has a crucial effect on the stability of the learning, slowly varies during the experiment $\pm$ 1 Hz above the initial value. Therefore, and foreseeing that such fluctuations are even more likely to occur in a future chronic implant, our final goal is to implement a variation of the cerebellar algorithm that is robust under similar conditions, i.e. slow fluctuations in the activity of IO.

We believe that our approach defines a specific paradigm for the generation of neuro-prosthetic systems that evolves following 4 steps: 1: identify the input and output structures and their encoding, 2: identify the anatomical and physiological principles underlying the computations performed by the target system, 3: integrate steps 1 and 2 with the appropriate signal processing in a single device and 4: miniaturise the

neuro-prosthetic system while optimising its power consumption. Here we emphasize steps 1-3 since we already previously have demonstrated step 4 (Hofstotter et al., 2004).

In summary, with this work we sought to provide further evidence for the fundamental principle underlying our model; namely, that the activity of the IO constitutes a teaching signal that controls the acquisition or extinction of CRs, and that by regulating the IO activity, the Nucleo Olivary Inhibition (NOI) stabilizes the CRs during paired CS-US training and drives extinction during CS alone stimulation. Our results provide such evidence, and additionally, they demonstrate at the design level the possibility of realizing an implantable low power neuro-prosthesis that would support the controlled acquisition, retention and extinction of novel behavior even if the biological substrate has lost its learning capability due to trauma or ageing.

5.2 Methods

5.2.1 Cerebellar model

Latencies

It is well known in the domain of control theory that the latencies and delays inherent in a system to be controlled play an important role in the design of the controller. Here, our controller is based on the cerebellar microcircuit involved in eye-blink conditioning. In nature, such a microcircuit must have internalized the latencies to the eyeblink system in several ways, one of them arguably being through the unusually long latency of the NOI (Hesslow, 1986) that we had previously interpreted as allowing for the matching of the system delays (Hofstotter et al., 2002). Informally, once an error signal reaches the IO, such a delay indicates how far ahead of it, the cerebellum should have taken a protective action for it to be effective. Consistently with this view, in the computational

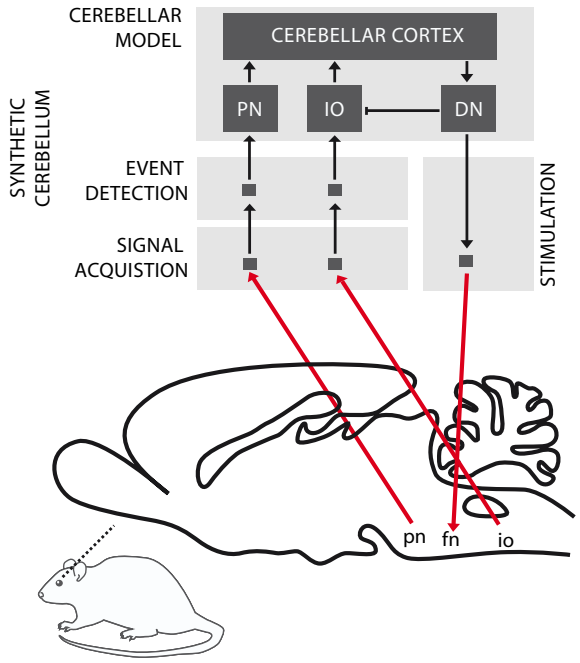

Figure 5.1: **Biological microcircuit and synthetic counterpart.** Recording (pn and io) and stimulation sites. After amplification and filtering of the signals recorded in the afferent structures, discrete events retrieved from Multi Unit Activity (MUA) are isolated by the event detection stages of the system, such that they are fed to their counterparts in the synthetic cerebellum (PN and IO). In nature, the repeated coincidence of CS and US signals within the cerebellar cortex induces plasticity causing the cerebellum to respond to the CS with a CR. In our model, once such a CR is acquired, it is relayed via the synthetic DN to the fn of the rat as an electrical stimulation that causes the animal to trigger the behavioral CR, i.e., the eyeblink. In addition, within the model, the CR triggered by the DN inhibits the IO, preventing a US derived signal from reaching the cerebellum once a protective action has already been issued. Since anesthesia prevents acquisition in the rodent's cerebellum, behavioral CRs expressed in the set up studied here are controlled by the synthetic circuit.

model that we employ (see below), the latency between the activation of the Deep Nucleus (DN) and the onset of the inhibition of the IO (the NOI delay, Λ_{noi}) sets the anticipation of the CR execution relative to the expected US arrival (Hofstotter et al., 2002) (see Fig. 5.2). Therefore, we will first discuss the latencies associated with the task of classical conditioning, since their specific properties underlie the cerebellar computational model.

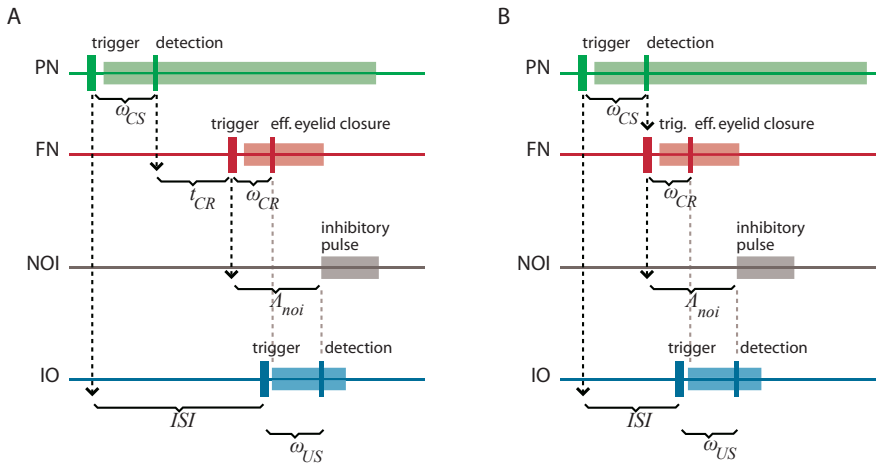

Figure 5.2: **Intrinsic latencies of the eye-blink conditioning preparation**. **A**: ISI, Inter-stimulus interval; ω_{CS}, latency between the peripheral CS stimulation and the detection of its associated neuronal response in the PN; t_{CR}, internal response timing learned by the model between the CS detection and the CR triggering; ω_{CR}, latency between the neuronal triggering of the CR and the effective eyelid closure, Λ_{noi}, delay between the CR trigger and the onset of the negative feedback loop inhibition; ω_{US}, latency between the US trigger and the detection of its associated neural response in the IO. **B**: Same latencies as in **A** for the minimum learnable ISI.

Concretely, setting a functional delay for the NOI requires knowledge of the transmission or mechanical latencies involved in the task, otherwise

the internal timing of the CR and US signals would result in non-adaptive CRs coming too late or too early with respect to the US air-puff. In other words, for the blink to anticipate the air-puff, Λ_{noi} has to exceed the sum of the sensory latency between the air-puff reaching the cornea and the US detection (ω_{US}) plus the effector latency between the CR triggering by the DN and the effective eyelid closure (ω_{CR}):

$$\Lambda_{noi} \geq \omega_{CR} + \omega_{US} \tag{5.1}$$

In the literature this sum of afferent plus efferent latencies is referred to as the delay of the error feedback (Miall et al., 1993). Setting Λ_{noi} to this minimal latency the CR and the US onsets will match. However, to achieve a better protection form the US, the best temporal arrangement of CR and US is that of the US onset coinciding with the middle of the CR. Given that we elicit the CR by an electrical stimulation lasting 150 ms ((Prueckl et al., 2011b)), such a temporal arrangement is achieved adding 75 ms to the minimal latency in Eq 5.1.

On the other hand, the sum of the latency between the onset of the CS delivery and its detection (ω_{CS}) plus the latency between the FN stimulation and the CR execution (ω_{CR}) affects the optimal internal timing (t_{CR}) that the model has to acquire for a given ISI:

$$t_{CR} = ISI - (\omega_{CS} + \omega_{CR}) \tag{5.2}$$

This time interval is shorter than the external ISI since it accounts for the detection and execution latencies. Note that diminishing t_{CR} towards 0 we get the minimum ISI that is learnable by the model, where CS detection immediately triggers a CR:

$$\min_{ISI} = \omega_{CS} + \omega_{CR}$$

For an ISI shorter than $\min_{ISI}$, a CR initiated by the cerebellum will always come after the US. For this reason, we will design a controller that only acquires CRs whenever the ISI exceeds this value.

Computational model

In what follows we summarize the biological model that is a modification of an earlier model (Hofstotter et al., 2002; Verschure & Mintz, 2001). Our model is based on the following assumptions: (1) the cerebellum is the brain area principally involved in the acquisition of a CR in the delay classical conditioning paradigm; (2) the only inputs received by the cerebellum are the Mossy Fibers (MF), carrying CS-related information, and the Climbing Fibers (CF), carrying US-related information; (3) the mechanism responsible for the acquisition of a conditioned response is plasticity at the Parallel Fiber (PF) to PC synapses; (4) such plasticity is induced by the stimulation of PF, alone (Long Term Potentiation – LTP) or jointly with CF (Long Term Depression - LTD); (5) IO, cerebellar cortex and DN are organized in distinct micro-complexes, which constitute negative feedback loops over IO; (6) the timing of the CR is adapted to the length of the ISI by these olivo-cortico-nuclear feedback loops that control the plasticity at PF-PC synapses by gating the cf error signal; (7) the training procedure leads to a pause in PC activity following CS presentation; (8) a CR is triggered by dis-inhibition of the deep nucleus by the cessation of PC firing; (9) PCs operate in two distinct modes: a spontaneous and a CS-driven mode. Informally, the PC is always maintained active during spontaneous activity of the input PF. However, during a CS presentation, PCs switch to a decaying activity. For a detailed explanation see Hofstotter et al. (2004).

Here, in order to deploy the cerebellar model on a low power aVLSI platform, we generated a computational model functionally equivalent to previous versions (Hofstotter et al., 2004, 2002; Verschure & Mintz, 2001) albeit more abstract from an anatomical standpoint to ensure com-

putational efficiency.

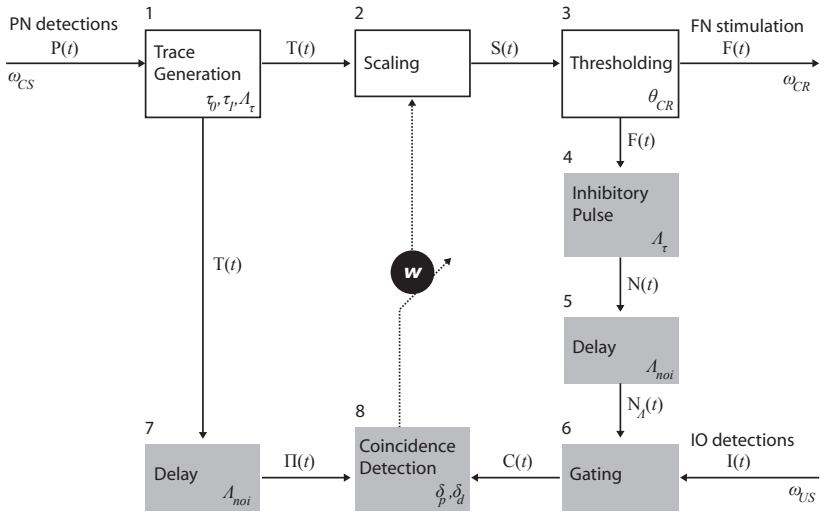

Figure 5.3: **Functional model of the cerebellum.** The processes in the top row (white boxes) map PN activity into action; in the case of eyeblink conditioning, tone detections into eyeblinks. Such mapping is controlled by the memory parameter w. The shaded processes adapt the mapping, namely, they are involved in the adjustment of w. The numbers identify specific processes. The latencies affecting each of the recording and stimulating channels as well as the parameters used in each process (wee main text for an explanation).

Process descriptions The Trace generation, Scaling and Thresholding processes (1, 2 and 3 in Fig. 5.3) model the processing of information that enters the cerebellum via the mossy fibers and leaves it through the excitatory axons of deep nuclear cells that project to the FN (Hesslow & Yeo, 2002; Christian & Thompson, 2003). The Trace generation (1) process codes the time since the CS onset with a decaying trace having a fixed initial value (τ_0), final value (τ_1) and duration (Λ_τ). This trace defines the memory span of the system; i.e. the maximum temporal gap

between CS detection and a CR execution learnable by the system. The Scaling (2) process multiplies the trace with the memory parameter w, which is the only parameter modified by learning. With w we mimic the changes in synaptic efficacy that occur in the molecular layer, due to LTD in the PF to PC synapse and/or other kinds of associative plasticity (Dean et al., 2010). Lastly the Thresholding (3) process triggers a CR whenever the value of the scaled trace falls below a decision threshold (θ_{CR}). Within this process we collapse all the transductions that occur post-synaptically from the PCs down to the efferents of the deep nuclei. In short, these three processes, map event detections in the PN into stimulation of the rat FN. The parameter w regulates the mapping and, by scaling the trace signal, controls whether a response is triggered or not, and if so, determines its timing in a way analogous to the biological system.

Three processes control the negative feedback loop that stabilizes learning: Inhibitory pulse (4), Delay (5) and Gating (6). The role of the negative feedback in classical conditioning is to prevent the error signal triggered by the US from reaching the cerebellar cortex if a CR has already been triggered (Medina et al., 2001; Verschure & Mintz, 2001). Processes 4 and 5 set the shape of the inhibitory square pulse. Its duration matches the duration of the CS trace, such that the IO can be inhibited for the whole duration of the CS trace. Process 4 delays the pulse by Λ_{noi} seconds. In practice, the value used was on the order of 100 ms. The Gating process (6) suppresses IO detections that co-occur with the inhibitory pulse. In summary, these components functionally reproduce the inhibitory control of the deep nuclear cells over the IO (Bengtsson & Hesslow, 2006).

The last two components, Delay (7) and Coincidence detection (8), update the associative weight w, thereby controlling the learning of the CR timing. Process Delay delays the CS stimulus trace by Λ_{noi} seconds (same value introduced above). The resulting trace defines the temporal

window wherein errors detected by IO can be prevented by the cerebellum. For instance, if a US is detected following a CS but ahead of this temporal window, then the CS-US interval is too short and any CR initiated by the cerebellum after the CS detection could not avoid the US (see eq. 5.1). Likewise, such a trace defines a so-called eligibility window wherein IO activity can be associated with a given PN detection. In short the system has a minimum ISI of Λ_{noi} s, and a maximum ISI of $\Lambda_{noi}+\Lambda_{\tau}$ s. Lastly, the Coincidence detection process (8) checks whether event detections in the CS and US pathways coincide and updates w accordingly. Namely, it decreases w every time an IO detection overlaps with the eligibility trace, and increases w if no IO detection occurs during that period. The function performed by these last two processes mimics the control of plasticity in the PF-PC synapse (Safo & Regehr, 2008; Wang et al., 2000; Sarkisov & Wang, 2008). The initial value of w is set to w_0.

Calibration of the cerebellar model

Definition of the optimization problem We mentioned that the Coincidence Detection process in Fig. 5.3 modulates the w parameter thereby controlling the acquisition and timing of CRs. In our implementation, the synaptic efficacy w is modified in linear steps, namely, δ_d for depression and δ_p for potentiation. The cerebellar model optimization consists of selecting the plasticity parameters δ_d and δ_p that result in a desired learning behavior. We solve this optimization problem in two different scenarios: with synthetic data or with data directly obtained from the brain activity of the animal. With the former we assess the properties of the model, whereas with the latter is applied in the biohybrid preparation. In both cases, the data consists of a set of detections in both recording sites of the system (PN and IO) and each set might contain evoked-detections (caused by the CS or the US, respectively) of spontaneous events. We refer to the former as True Detections (TDs)

and to the second as False Alarms (FAs).

Informally, we impose that the learning dynamics of the system mirror the behavior: when CSs and USs are paired the circuit should learn to produce CRs within tens of trials; when in a trained animal CSs are not paired with USs, the circuit should unlearn to produce CRs within tens of trials; all other conditions should not alter the circuit transfer function. More formally, the optimization problem is described by a linear system representing three types of constraints: acquisition, extinction and stability (see Table 5.1).

ID	Experimental condition	Description
1	Acquisition	Paired CS-US presentation leading to acquisition of CRs
2	Extinction	CS-alone trials with CR leading to extinction of CRs
3	Stability	CS-alone or unpaired CS-US trials with no CR, causing no modification of the memory parameter

Table 5.1: Stimulation conditions for the closed-loop experiment. See text for further explanation.

Estimation of plasticity events. PN (CS) and IO (US) detected or artificially generated events are coded in binary vectors $\mathbf{P}$ and $\mathbf{I}$, where each element is a time step and a value of 1 signals an event. The vector of eligibility traces ($\mathbf{\Pi}$, box 1 in Fig. 5.3) is obtained by convolving $\mathbf{P}$ with the eligibility trace waveform (ϵ):

$$\mathbf{\Pi} = \mathbf{P} \star \epsilon$$

where ϵ is a rectangular pulse lasting Λ_τ s and delayed by Λ_{noi} s. Here we fixed these values to 0.3 s and 0.1 s, respectively. The first value is in

good accordance with the maximum interval between CS and US bridged by the cerebellum in eyeblink conditioning (Kalmbach et al., 2009; Moyer et al., 1990), whereas the second matches the most effective interval between parallel fiber and climbing fiber stimulation for the induction of cerebellar LTD (Safo & Regehr, 2008; Wang et al., 2000; Sarkisov & Wang, 2008).

Eligibility Trace Vector	US Vector	Plasticity
$\mathbf{\Pi}(t) = 1$	$\mathbf{I}(t) = 1$	Depression
$\mathbf{\Pi}(t) = 1$	any $\mathbf{I}(t)$	Potentiation
$\mathbf{\Pi}(t) = 0$	any $\mathbf{I}(t)$	No plasticity event

Table 5.2: Plasticity Conditions. See text for further explanation.

Plasticity events occur under the conditions specified in table 5.2. Firstly, a necessary condition for a potentiation or depression event to occur at a given time step (t) is that the eligibility trace is non-zero. Secondly, the number of potentiation events is $P = \sum \mathbf{\Pi}$ for potentiation occurs for every time step with a non-zero CS eligibility trace. Thirdly, depression occurs when US detection overlaps with the eligibility trace. Hence, the number of depression events can be obtained with the scalar product of $\mathbf{\Pi}$ and $\mathbf{I}$:

$$D = \mathbf{\Pi}^T \mathbf{I} \tag{5.3}$$

Note that whenever a depression event occurs, it outweighs the default potentiation events triggered by the plasticity trace $\mathbf{\Pi}$, resulting in a net depression.

In the presence of CRs, D must be corrected to account for the IO events (spontaneous or US-evoked) suppressed by the NOI. Note that since the timing of inhibition depends on the triggering of the CR and the eligibility window is anchored to the CS, rapidly elicited CRs are more effective in gating plasticity than late CRs. In other words, the effectiveness of

the gating decreases as the CRs become more delayed. We can heuristically approximate the reduction in IO events reaching the coincidence detection by multiplying the number of IO detections by an estimated *mean* proportion of IO events *not* suppressed by the NOI ($\bar{\sigma}$),

$$D = \mathbf{\Pi}^T(\bar{\sigma}\mathbf{I}) = \bar{\sigma}\mathbf{\Pi}^T\mathbf{I} \tag{5.4}$$

where $\bar{\sigma} = 1 - \sigma$, with σ being the proportion of IO events suppressed by inhibition.

As the equation illustrates, this result can be computed simply by multiplying the result of Eq. 5.3 by the factor $\bar{\sigma}$.

Optimization of the plasticity parameters. At this point, having estimated the number of plasticity events produced by two sets of event detections in PN and IO, we obtain the optimal plasticity parameters (δ_p and δ_d) by solving with weighted least squares the following system:

$$\begin{bmatrix} \bar{P}_1 & \bar{D}_1 \\ \bar{P}_2 & \bar{D}_2 \\ \bar{P}_3 & \bar{D}_3 \end{bmatrix} \begin{bmatrix} \delta_p \\ -\delta_d \end{bmatrix} = \begin{bmatrix} -\Delta_a/T_a \\ \Delta_e/T_e \\ 0 \end{bmatrix} \tag{5.5}$$

$\bar{P}_i$ and $\bar{D}_i$ are the mean plasticity events per trial, potentiation and depression respectively, and the sub-indexes indicate the experimental condition (see Table 5.1). They are obtained by dividing $\mathbf{D}$ and $\mathbf{P}$ by the number of trials contained in the training set. Δ_a is an estimate of the change in w necessary for acquisition and T_a sets the desired number of trials for acquisition. These two values set a target mean change of w per trial. For instance, if the initial value of w is 0.5 and we estimate that well-timed CRs occur when w reaches a value of 0.3, then we set Δ_a to 0.2. Δ_e and T_e are the same values applied to extinction. As we declared in the assumptions of the biological model, and for consistency

with classical cerebellar learning theory, that links learning in the cerebellum with LTD in PF-PC synapses (Ito et al., 1982; Albus, 1971), we suppose that CR acquisition requires depression (decrease) of the value of w and extinction, a potentiation (increase). Regarding the optimization algorithm, we weighted more the stability constraint since it by itself guarantees the convergence of the learning dynamics, i.e., paired CS-US stimulation yields acquisition and CS-alone stimulation yields extinction. Informally, if under spontaneous IO activity w has an average of 0 drift, then an increase in IO activity will lead to acquisition and a decrease, to extinction. Once this constraint is satisfied, the acquisition and extinction constraints modulate the rate of either learning process.

Adaptive calibration of the model

In the previous section we have introduced a calibration method that assumes stationary bio-signals during the experiment. Crucially, this is a strong assumption that will hardly ever be met under in-vivo conditions. In our case, for instance, the rate of IO activity in the bio-hybrid experiment, fluctuated markedly producing non-associative modifications in the synaptic efficacy w. For this reason, here we introduce an adaptive version of the calibration method that considers non-stationary responses in IO activity. Since the recalibration has to occur without resorting to additional training data, we keep the same acquisition and extinction constraints used for the initial calibration, and we only update the stability constraint, introducing in this constraint the current estimation of the rate of spontaneous IO activity.

The recalibration is periodically performed, with a fixed time interval. In the experiment we used 150 seconds, that corresponds roughly to 10 trials. Such recalibration requires an estimate of the ongoing level of spontaneous activity in IO, the IO_{far}, where the sub-index far stands for the false alarm rate (see below). To compute this estimate we count the number of IO detections between recalibrations. Note that, since

the system is blind to whether the detections are spontaneous or evoked, i.e. it has no knowledge whether stimuli are presented or not, for the estimation of IO_{far} all detections are considered spontaneous. During acquisition, given that some of the IO detections will be US-evoked, this results in an over-estimation of the true IO_{far}. In that case, the estimate is more accurate for a higher proportion of spontaneous detections to evoked ones, what can be easily achieved using large Inter-Trial Intervals (ITI).

From here, since the IO rate only affects the number of depression events, we only have to update one value in the linear system in Eq. 5.5, namely, $\bar{D}_3$ that accounts for the number of depression events during spontaneous activity. This is so, because the other term of the stability constraint has no dependence on IO activity. Hence, having only to update a single term of the system allows us to fulfill one of the requirements of the system prototyped here: providing algorithms simple enough to be implemented in a low power VLSI solution. Indeed, algebraic manipulation shows that we can compute the solution to the optimization for each of the two plasticity parameters as a ratio of two polynomials with a maximum degree of 1 (for the derivation see appendix). For instance, in the case of δ_p we have:

$$\delta_p = \frac{\alpha_2 IO_{far}^2 + \alpha_1 IO_{far} + \alpha_0}{\beta_2 IO_{far}^2 + \beta_1 IO_{far} + \beta_0}$$

where the coefficients of the polynomials are determined only by the training data. For the detailed derivation of this formula see Appendix 2.

5.2.2 Work-flow of the bio-hybrid experiment

The methods introduced so far where common to the simulation and in-vivo experiments. In what follows we introduce the methodology specif-

ically developed for the bio-hybrid preparation.

Preparation The details of the animal preparation, surgery and electrophysiological setup with the exact placement of the electrodes have been fully described in a separate publication (Prueckl et al., 2011b). Briefly, we introduce a titanium-nitride micro-electrode array (Faculty of Engineering, Tel Aviv University) in the PN in order to detect the neural response to the CS (an auditory signal: white noise), and a tungsten needle electrode (A-M Systems, USA) in the IO to detect the response to the US (the air-puff to the eye of the rat). Signals from both recording sites are band-pass filtered (300 - 3000 kHz) to work in the MUA range. A stimulating electrode is placed in the facial nucleus and tested to induce reliable eye-blinks when 200 mA 0.1 ms constant-current pulses with a frequency of 80 Hz for 150 ms are delivered.

For the classical conditioning preparation we use as the CS a white-noise stimulus at 67-70 dB with a duration of 450 ms and a 150 ms duration air-puff as the US. The presence of CRs was verified recording the periorbital electromyography. The Inter-Stimulus Interval is set to 300 ms and the ITI is randomized between 10 and 15 seconds.

After validating the responsiveness of the MUA signal to the air-puff and the tone by visual inspection of both Peri-Stimulus Time Histograms (PSTHs) of MUA events, we record a training data-set that comprises 30 trials with paired CS-US presentation followed by 2 minutes of spontaneous activity. This data-set is then used to calibrate the synthetic cerebellum (see below). After calibration we proceed with the classical conditioning paradigm, providing the animal with CS-US stimulations until we observe stable CRs, followed by CS-alone stimulation until extinction of the CRs has been achieved.

Signal processing The goal of the signal processing stage is to detect in the MUA signal the onset of the responses to the CS and US, i.e., to the

tone and the aripuff stimuli, respectively. Given the intended implementation in a low power VLSI solution, we limit ourselves to low complexity algorithms. Briefly, we detect sustained increases in the variability of the MUA signal occurring after each stimulus presentation. This is achieved with the following steps: first we subtract from the signal a running estimation of the mean and rectify the resulting signal. Secondly, the signal is smoothed to obtain a short-term temporal average that serves as a measure monotonically related to the variability increase. Lastly, event detection occurs every time the resulting variability signal surpasses a certain detection threshold.

We defined a priori the windows of possible True Detections (TDs) for each channel (10-150 ms after the trigger in the PN and 5-205 ms after the trigger in the IO). Likewise, the performance of the signal detection can be summarized with the True Detections Ratio (TDR) and the False Alarm Rate (FAR), where the TDR indicates the proportion of stimuli raising *at least* one detection within the TD window, and the FAR the frequency of events detections during the periods of no stimulation, i.e., outside the TD window. Note that since in the IO we found a FAR between 0.5 and 2 Hz, we can compare the IO_{far} with the spontaneous levels of activity in the IO (Jirenhed et al., 2007).

Estimation of the event detections. With the calibration dataset we estimate for both channels the detection performance during early acquisition trials (before any CR is triggered). To estimate the number of detections during CS-alone trials we combine the PN data from the paired stimulation trials with IO data from the spontaneous activity period. This is done in order not to excessively extend length of the calibration phase.

Optimization of the signal detection regimes. To signal an event detection we have to first set a detection threshold. Such selection poses

a multi-objective optimization problem since we want to simultaneously maximize the TDs and minimize the false alarms, and we do not know *a priori* which is the best trade-off of both measures that maximizes the chance of success in our experiment. To overcome this problem we iterate the optimization process over a set of threshold configurations for both input channels, and then select *a posteriori* the one yielding the minimum error in the optimization of the plasticity parameters (eq. 5.5). Likewise, the optimization process, and not us, selects the optimum trade-off between TDs and false alarms. Note that the simplicity of the calibration method previously introduced allows us to iterate over a great number of possible threshold configurations in very little time.

5.2.3 Use of the model with simulated data

Before deploying the computational model in the bio-hybrid setup, we tested the performance of the model with artificially generated data. To this end we generate a set of detections for each channel according to a certain pair of TDR and FAR statistics. From these two statistics we produce the binary vectors of detections **P** and **I**, and for this we convert the TDR and FAR to event probabilities per time step. We obtain the probability of detection in the absence of stimulation by multiplying the FAR by the model time step, which is 2 ms. Regarding TDs, to convert the TDR to a probability we have to consider the size of the TD window. Operationally, since we interpret the TDR as the probability of getting *at least* one event within the detection window, we have to find the event detection probability yielding no events during the TD window with a probability of $1 - \mathrm{TDR}$. This event detection probability is equal to $1 - \sqrt[n]{1 - \mathrm{TDR}}$, where n is the size of the TD window in time steps.

5.3 Results

5.3.1 Simulation experiments

Performance of the model with spontaneous activity in the IO

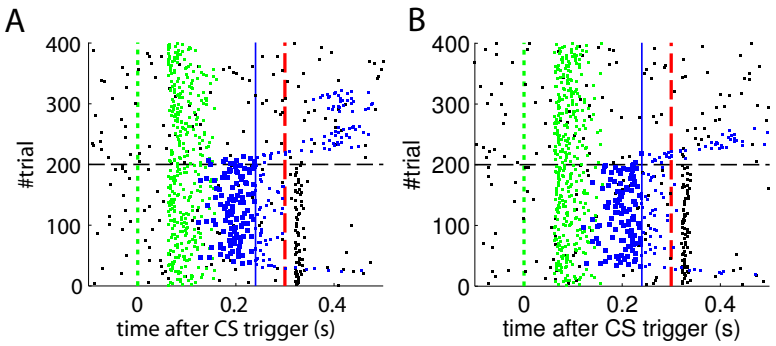

Figure 5.4: **Raster plots of the inputs and outputs of the model with and without stability constraint. A.** Model with stability constraint. PN detections (green), IO detections (black) and CR triggers (blue), well-timed (thick) and late (thin). CS (dotted green line) and US (dashed red line) onset times. The horizontal dashed black line separates acquisition and extinction phases. **B.** Model without stability constraint. Data plotted as in A.

In previous work we used this cerebellar model assuming no baseline activity in the IO (Verschure & Mintz, 2001; Hofstotter et al., 2004). As a first step, we test whether this model supports the acquisition and extinction of CRs when the IO displays spontaneous activity (see parameters in Table 5.3). The outcome of a representative simulation shows that indeed the model adapts well to the case of baseline IO activity (Fig. 5.4A). It acquires well-timed CRs in CS-US paired trials and extinguishes them in CS-alone unpaired trials (Fig. 5.4A and 5.5A), and importantly the parameter w reaches a stable plateau after complete extinction (Fig. 5.5B). We stress that the stabilization occurring at the end of extinction even in the presence of CS alone stimulations, stems from

the stability condition in equation 5.5. If we remove this constraint the overt behavioral results are similar (Fig. 5.4B and Fig. 5.5A) but the underlying memory dynamics differ (Fig. 5.5B). Indeed, behaviorally both models only differ in the extinction phase, which is slower for the model with stability. However, in regard to the model's state, without stability, the synaptic efficacy w grows also after extinction of the CRs has been accomplished. Note that, in consequence this can make reacquisition harder than acquisition if the extinction training is maintained, which goes against the experimental evidence (Kehoe & Macrae, 2002b). In conclusion, the computational model of the cerebellum is also functional when the IO has baseline activity, requiring only a proper calibration of the plasticity parameters.

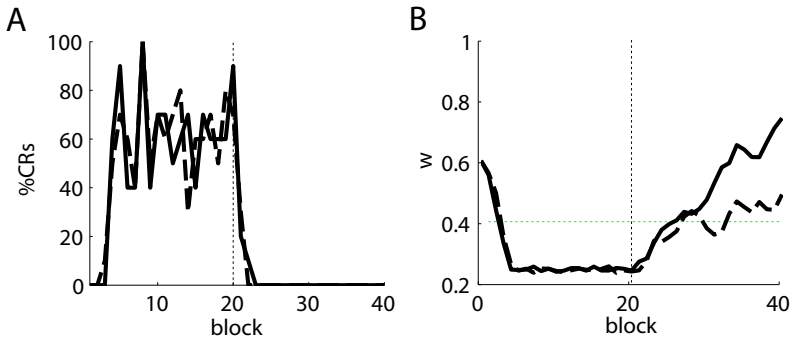

Figure 5.5: **Behavior of the model with simulated data.** Behavioral performance. Percentage of CRs per block of trials of the model fitted with stability constraint (solid line) and without (dashed line). The vertical dotted line separates acquisition and extinction training. **B.** Trajectory of w in the model fitted with plasticity constraint (solid line) and without (dashed line). The horizontal green dotted line marks the level above which the model does not trigger any CRs. Blocks of 10 trials.

Parameter	Value
Cerebellar model	
CS trace:	
τ_0	1
τ_1	0.5
Λ_τ	350 ms
Λ_{noi}	100 ms
θ_{CR}	0.2
w_0	0.5
Signal Detection	
PN_{td}	0.95
PN_{far}	0 Hz
IO_{td}	0.75
IO_{far}	1.0 Hz
Fitting constraints	
Δ_a, Δ_e	0.2
T_a, T_e	40

Table 5.3: Parameters of the simulation.

Effect of the latencies of the cerebellar model

We previously discussed the relevance of the latencies in the design of the controller. Here we will illustrate with two examples the functional implications of the two latencies implemented in the model, namely, the latency of the NOI and the latency of the plasticity trace. We recall that in our model, both latencies are set to the same value, namely, Λ_{noi} s.

The effects of the NOI latency on the timing of the CRs has already been discussed in the literature at the theoretical level (Hofstotter et al., 2002; Hesslow & Ivarsson, 1994), and demonstrated in experimental set-ups (Herreros Alonso & Verschure, 2013a). Here, and because of the noisy input conditions, we see that if we do not apply any delay to the NOI, the CR triggers eventually anticipate the US, but by too short a latency too be considered effective (Fig. 5.6A). Therefore, even though the model

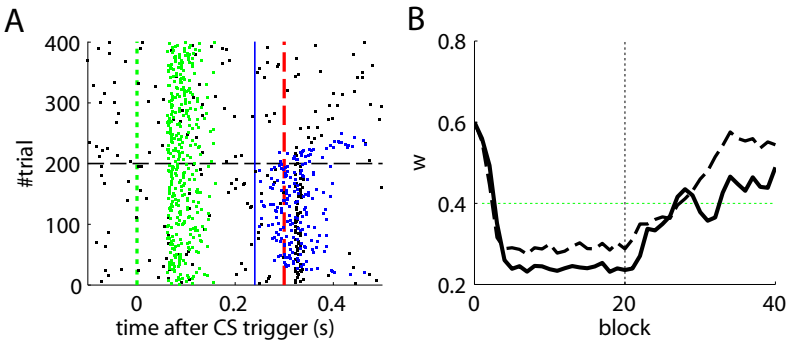

Figure 5.6: **Results with and without delayed NOI inhibition. A.** Raster plot with the output of the model with the delay of the NOI set to 0 s. **B.** Trajectory of w in the model with a delay of 100 ms in the NOI (solid line) and with no delay (dashed line). The horizontal green dotted line marks the level above which the model does not trigger any CRs. Blocks of 10 trials.

triggers CRs, they are maladaptive. Indeed, the synaptic efficacy w fails to reach a level sufficiently low to initiate well-timed CRs, as it does when the latency of the NOI is properly set (Fig. 5.6B). Note however that the jitter of the trace of the synaptic efficacy w occasionally brings the CR triggers close to the criterion of correct timing. Given that, if such a jitter will be incresead it would be possible for occasional CRs to anticipate the US sufficiently to be characterized as *well-timed*. This occurs if, for instance, the signal to noise ratio of the IO signal decreases (5.7B, with TDs in the IO lowered from 70% to 50%) or if we force the learning to be faster (5.7A). This by no means indicates that the model works better if the signal conditions are worse, it only indicates that as the dynamics of the model become more noisy (5.7C), some well-timed CRs may incidentally be triggered, even if the delay of the NOI is not correctly set.

Not delaying the plasticity trace leads to a different kind of maladaptive behavior. In this case, if we set an ISI below the minimum ISI described in Eq. 5.1, the computational model without delayed plasticity acquires

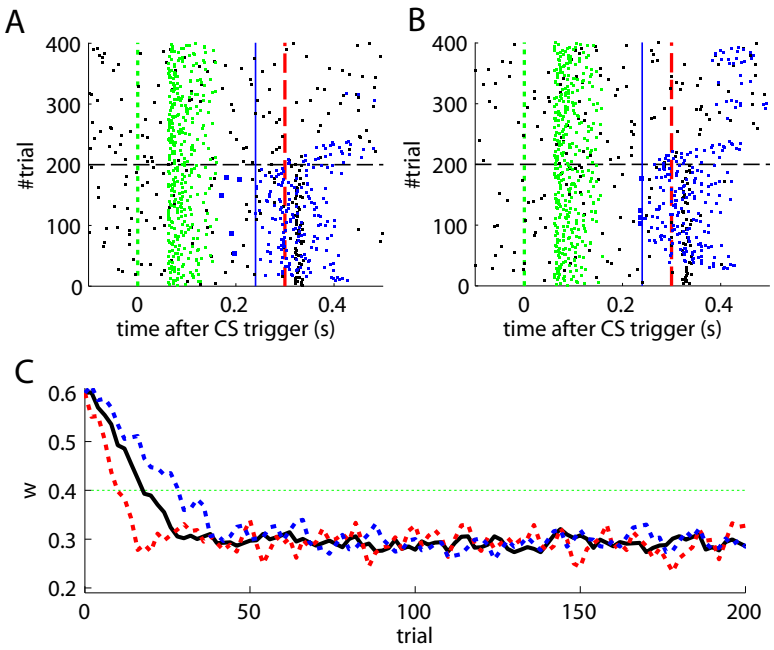

Figure 5.7: **Results with non-delayed NOI inhibition in different conditions.** Raster plots with the output of the model with the delay of the NOI set to 0 s. **B.** the model constraint to acquire CRs twice as fast or with **A.** a ratio of TDs in the IO lowered to 50%. **C.** Traces of the synaptic efficacy w for the simulation in Fig. 5.6 (black) compared to the simulations in panel A (dotted red) and B (dotted blue).

CRs that can only be late CRs by definition (Fig. 5.8B). In contrast, setting the appropriate delay to the plasticity trace avoids building any association between CS and US that are too close in time (Fig. 5.8A).

5.3.2 Bio-hybrid experiment

Evaluation of the training data We started the bio-hybrid experiment recording a training data set composed of 30 trials of paired CS-US stimulation, with an ISI of 300 ms and an ITI of 10 s. After applying

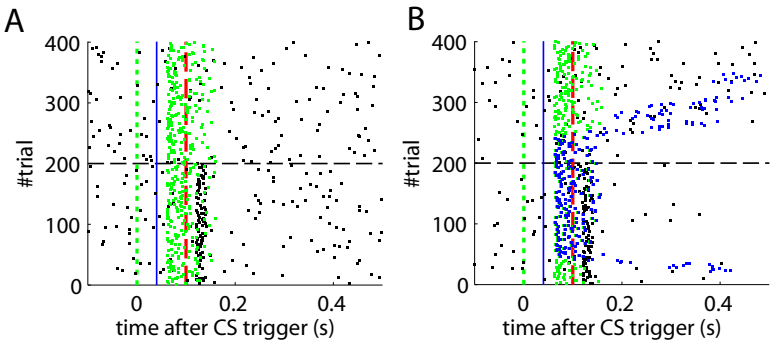

Figure 5.8: **Effect of the delayed plasticity trace on the behavior. A** Model with plasticity trace starting Λ_{noi} s after each PN detection. **B.** Model with plasticity trace starting right after each PN detection. Data plotted as in Fig. 5.4

.

the signal processing algorithms (see Methods) we built the Receiver Operating Characteristic (ROC) curve for each of the channels (Fig. 5.9). The PN channel TDs reached 100% with a False Alarm Rate close to 0.1 Hz while the IO displayed TDs near 50% for the range of optimal FARs ($\sim$ 1 Hz). Therefore the PN channel was reliable while the IO channel was poor from the detection standpoint.

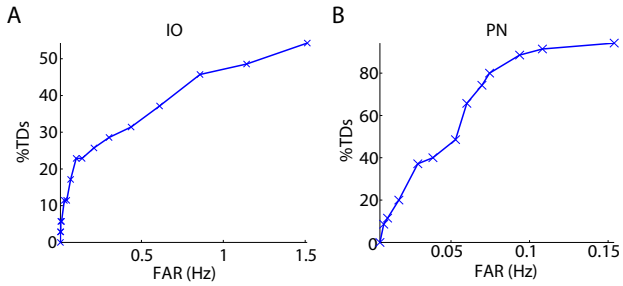

Figure 5.9: **Event detection performance for the recording sites.** ROC curves for the IO (**A**) and the PN (**B**) event detections.

Optimization of the plasticity parameters and signal detection regimes The following phase entailed tuning the plasticity parameters and selecting the optimal signal detection thresholds. The optimization process selected detection thresholds yielding a percentage of TDs of 48.6% and a FAR of 1.14 Hz for the IO channel, and a 91.4% of TDs with a FAR of 0.11 Hz for the PN.

The model calibration set the potentiation and depression steps (δ_p and δ_d) to 0.0161 and 3.36e-5 respectively.

The offline simulation parameterized with the previous values is shown in Fig. 5.10. First of all, acquisition occurs in 40 trials with an asymptotic performance of 40% well-timed CRs. Secondly, there is a 20% chance of obtaining total extinction after 120 trials of CS-alone stimulation. Thus, we observe *a priori* that the low detection quality of the US driven IO signals makes will pose a challenge for successful extinction.

Evaluation of the bio-hybrid experiment After the preliminary assessment of the quality of the biosignals, we proceeded with the online classical conditioning experiment (Fig. 5.11). The experiment lasted 1h 20min and comprised 190 CS-US stimulation trials (acquisition) followed by 180 CS-alone trials (extinction), with randomized ITIs between 10 and 15 s. In Fig. 5.11 we display events detected and stimulations triggered by the neuro-prosthetic system. For the whole experiment, in the PN there was a TDR in of 75.5% and a FAR of 0.1 Hz, that include a high number of late CS detections (Fig. 5.11). Notably, the number of PN detections during baseline was very low (only 5 false alarms in 80 minutes). In the IO we obtained a TDR of 38% and a FAR of 1.2 Hz.

Detections in both channels were delayed by tens of milliseconds with respect to the stimulus trigger. The mean latency of the TDs in the PN (ω_{CS}) was of 96.2 ms after the CS trigger (Fig. 5.12A) whereas the mean latency in the IO channel (ω_{US}) was of 68.5 ms (Fig. 5.12B).

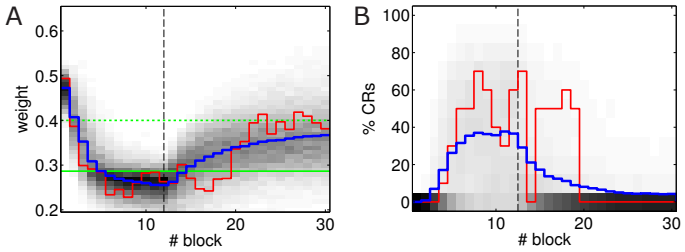

Figure 5.10: **Performance of the experiment predicted by the training data. A.** Trajectory of the memory parameter after 2500 simulations plotted in blocks of 10 trials. The simulated experiment contained 120 trials of acquisition and 180 trials of extinction. Distribution of the block-by-block values of w (gray-scale) with mean (blue) and and output of a sample simulation (red). We indicate the levels of the weight that result in late (upper green line) and well-timed CRs (lower green line). The transition from acquisition to extinction training is marked by a vertical line. **B.** Predicted behavioral performance after 2500 simulations. Percentage of well-timed CRs. Distribution of the block-by-block performance (gray-scale) with mean (blue) and result (red) of a sample simulation (same as in A).

The experiment was successful in terms of behavior: well-timed CRs were triggered with regularity towards the end of the acquisition phase, and no CR was triggered during the last 90 trials of the extinction training (Fig. 5.11 and 5.12C). The first response appeared at trial 29, but the first well-timed CR came only at trial 118. Notice that towards the end of acquisition the series of well-timed CRs appeared regular. After the onset of the extinction trials (trial 191) well-timed CRs were rapidly extinguished. A block-by-block analysis reveals that the performance fluctuated during acquisition (Fig. 5.12C) and that extinction of well-timed CRs was very rapid, in total there are only 4 well-timed CRs during extinction, the last one appearing at trial 220, i.e., after 30 trials of extinction. However, the extinction of late CRs was more gradual, encompassing blocks 21 to 29, i.e. total extinction required 100 trials.

No CR was triggered by the system in the last 60 trials of the experiment. Regarding the timing, well-timed CRs occurred on average 50 ms ahead of the US trigger (Fig. 5.11E).

The evolution of the synaptic efficacy w is displayed in Fig. 5.13. We estimate that given our setup CRs follow a PN event whenever the value of w goes below 0.4. However, for such CRs to be anticipatory, w should settle below or at 0.28. During the experiment, w decreased steadily during the first 60 trials, down to a value of 0.29. Afterwards the decrease decelerated. The dynamics of w suggest that learning has reached an asymptotic-level by the end of the acquisition stage (Fig. 5.13). The

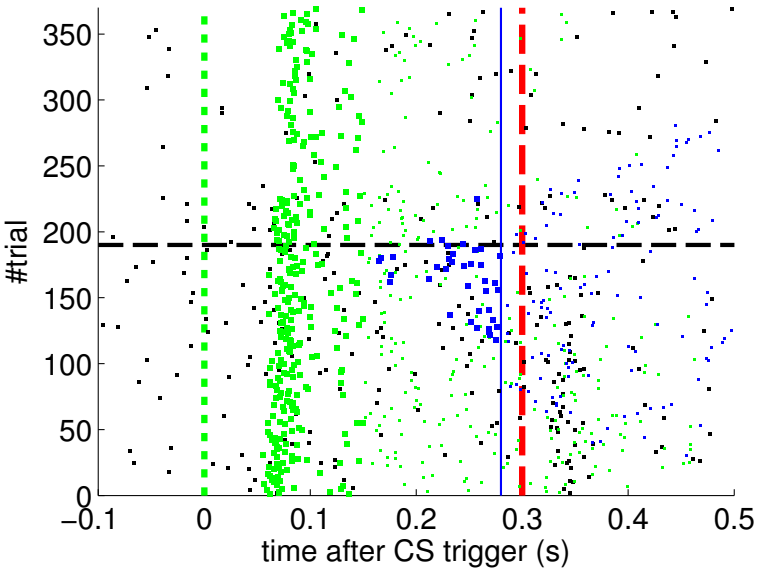

Figure 5.11: **Event detections and triggers during the online experiment. A.** Raster plot with the PN detections (blue dots; well-timed PNs are thicker), IO detections (black), and CR triggers (blue dots; well-timed triggers are thicker). CS trigger (dotted green line) and US trigger (dashed red line) are also indicated. Blue line separates well-timed from late CRs. Horizontal dashed-line defines transition from acquisition to extinction training.

mean value of w during well-timed CRs was 0.25, indicating that well-timed CRs were triggered on the average 140 ms after the PN detection. Thus, for an ISI of 300 ms the model acquired an internal timing (t_{CR}) of 140 ms.

Thus, in conclusion, at a first level of analysis, the results of the bio-

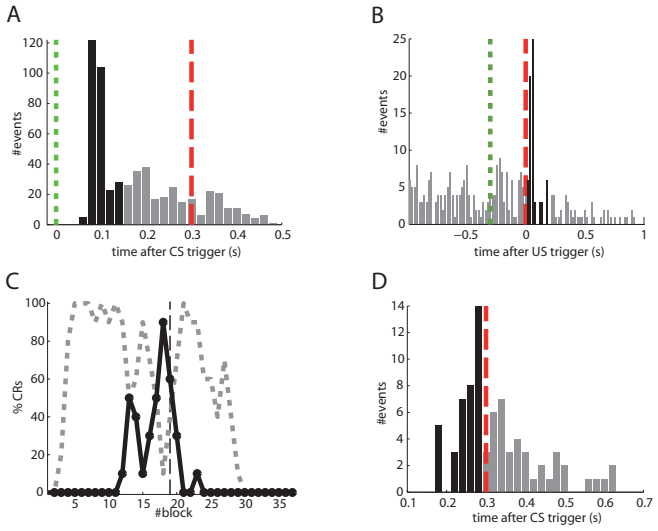

Figure 5.12: **Quantitative results. A. Events detected in the PN.** PSTH of PN detections relative to the CS-trigger: TDs (black bars) and FAs (gray bars). CS trigger (dotted green line) and US trigger (dashed red line). **B. Events detected in the IO.** Detections in the IO sorted relative to the US-trigger. Data plotted as in A). **C. Behavioural performance of the bio-hybrid.** Percentage of well-timed CRs during acquisition (solid red line) and extinction (dashed red line - starting at block 20). CRs that were not triggered at least 20 ms ahead of the US trigger appear as late-CRs (light red line). Each block contains 10 trials. **D. Timing of CRs.** PSTH of the Condition Responses. The information is extracted from trials 118 to 190. Data plotted as in A

hybrid experiment were correct both from the behavioral point of view and also regarding the dynamics of the underlying memory parameter stored in the synthetic cerebellum.

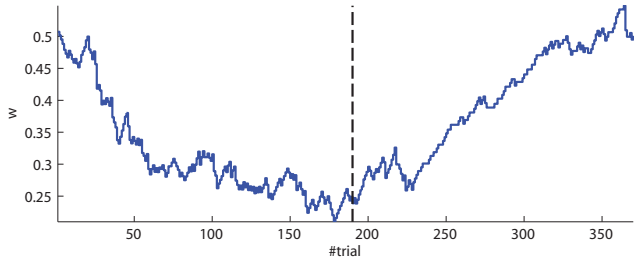

Figure 5.13: **Weight trajectory during the experiment.** The dashed vertical line separates the acquisition and extinction phases.

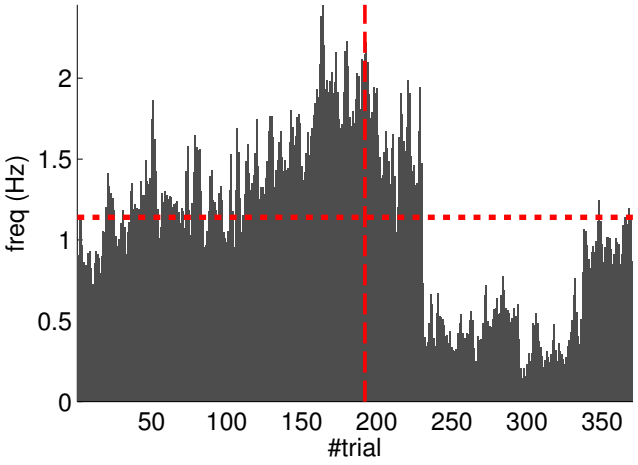

Figure 5.14: **Fluctuations in the spontaneous IO rate.** Mean IO rate in each trial of the experiment. The horizontal dotted line marks the 1.14 Hz level of activity. The vertical dashed line marks the transition from acquisition to extinction trials.

Instability of the activity during the recording Having said that, there were two major caveats in the experiment. First, due to an

artifact introduced by the electrical stimulation of the CR, the signal of the IO was masked for 2 seconds after each CR. This masking was performed at the signal acquisition stage (Prueckl et al., 2011b). For this reason, no events reached the computational model of the cerebellum for 2 seconds after each CR trigger. Under such circumstances the model's NOI became superfluous, because for all its extent there was no IO detection to be inhibited. Or, in other words, the mask at the signal acquisition stage acted as a NOI with 0 latency and longer duration. We have already argued that the latency of the NOI is necessary for consistently achieving a correct timing of the CRs (Fig. 5.6). Thus, in summary, on the one hand, it is reasonable to assume that the well-timed CRs were in part a consequence of the noisy conditions of the input setup (e.g., a IO_{td} of 38%), and on the other, it is also reasonable to expect that in the absence of the stimulation artifact the synthetic cerebellum would have achieved a higher proportion of well-timed CRs.

Second, our calibration of the cerebellar model assumes that the level of spontaneous activity in the IO remains constant. If the IO spontaneous activity significantly deviates from the one estimated during calibration, then w will drift, eventually leading to either acquisition or extinction. Since during the conditioning experiment we observed that the spontaneous activity fluctuated (Fig. 5.14), there is the possibility that the behavior observed did not result from associative learning but from changes in w due to oscillations in the IO spontaneous activity. In particular, given that the IO spontaneous activity increased during acquisition and decreased during extinction such fluctuations might have caused or favored the behavioral result.

To perform an *a posteriori* control for this, we check whether the observed oscillations in spontaneous activity may lead to acquisition by themselves even in the presence of temporally unrelated CSs and USs. We tested this simulating unpaired CS-US presentations. For this, we generate experiments with shuffled IO detections within each trial. After performing

20000 simulations, we observe that the increased spontaneous activity of the IO causes a decrement in w during the acquisition phase for *unpaired* stimulation (Fig. 5.15A). Considering the behavior (Fig. 5.15B) in the average simulation the decrease of w yielded to the triggering of a small number of CRs. These results both at the level of the memory parameter and the behavior were clearly below the performance observed on the bio-hybrid experiment but demonstrated nonetheless that under experimental conditions with big variability in the recorded signals, the model might acquire spurious associations.

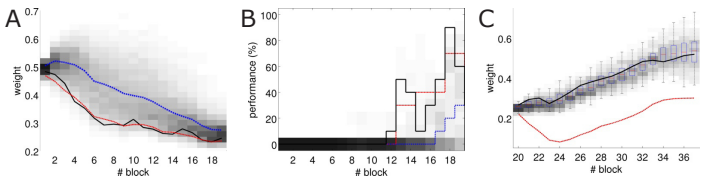

Figure 5.15: **Observed performance vs. performance during simulated unpaired acquisition. A.** Acquisition during paired CS-US training versus simulated unpaired CS-US. Trajectory of the weight during the acquisition phase of the experiment (black line) plotted against results of 20000 simulations of unpaired training. Distribution of the simulation results (grayscale), median (blue dotted line) and the 0.05 bottom of the distribution (red line). Blocks of 10 trials. **B.** Behavioral performance during acquisition against performance in the simulations. Percentage of CRs during acquisition in the experiment (black line) plotted against the percentage obtained in the simulations. Distribution of the simulation performances (grayscale), median (blue dotted) and the upper 0.1 of the distribution (red line). **C.** Acquisition during paired CS-US training versus simulated unpaired CS-US with and without negative feedback loop. Trajectory of the weight during the extinction phase of the experiment (black line) plotted against results of 20000 simulations of unpaired training. Distribution of the simulation results with the negative feedback-loop (grayscale and boxplots) and without it (red line, mean).

5.3.3 Calibration method adapted to fluctuating IO activity

Lastly, we tackle the problem of the instability of the IO activity evidenced in the previous section. For this we apply the adaptive calibration method (see Methods). We test this method with data from the bio-hybrid preparation, aiming to show that with the adaptive calibration, the cerebellar model becomes robust against slow fluctuations in the baseline IO activity. As a caveat, notice however, that in the bio-hybrid experiment, by definition of a *closed-loop* set-up, the data recorded during the session depended on the output of the model. In this case, the data recorded was affected by the electrical stimulation in the FN driving the CR. Thus, to cancel out this effect we replaced the 3 seconds of the IO signal occurring after each CR, by the signal extracted from random trials with no CR.

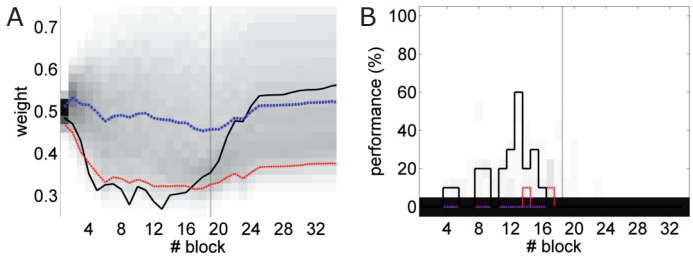

Figure 5.16: **Performance with the adaptive calibration method. A.** Trajectory of the memory parameter w plotted in blocks of 10 trials. Distribution of the performace of the simulations of unpaired CS-US stimulation (grayscale), with mean (dotted blue) and lower 10% (dotted red), for a total of 2500 simulations of 36 blocks. Trajectory of the simulated classical conditioning experiment (solid black), with 18 blocks of acquisition and 18 blocks of extinction. The transition from acquisition to extinction training is marked by a vertical line. **B.** Behavioral performance of the same simulations. Percentage of well-timed CRs per block. Results plotted with the same convention as in A.

The results now separate clearly the performance between the unpaired- and paired-stimulation experiments. Most importantly, there is no acquisition of CRs during unpaired stimulation. In this case, the fluctuations of the baseline IO rate do not push w further than +/- 0.1 from the starting value, both during acquisition and extinction. Regarding the performance, in the mean, there are no well-timed CRs with unpaired CS-US training. To the contrary, in the simulated acquisition and extinction experiment, the overall behavior of acquisition followed by extinction is preserved. In this experiment the CR performance decreases relative to the result with the bio-hybrid, especially by the end of acquisition, when the recorded IO baseline rate was higher. The reason is that now a high rate of spontaneous detections in the IO diminishes the relative saliency of the US-evoked events, for which such a high rate harms rather than helps acquisition. Thus, this simulation confirms, that if same conditions of signal instability of the bio-hybrid experiment are repeated, with the adaptive calibration method we can ensure that any CRs observed will specifically be due to the CS-US association.

5.4 Discussion

Here we have addressed the challenge of defining, interfacing and validating a neuro-prosthetic system for the cerebellum. In particular, we have first: defined a biologically grounded computational model of the circuit targeted for substitution, second: defined its input and output structures and their encoding of their respective events and third , we have implemented this cerebellar prosthetic and interfaced it to the brain of a rat. Our results show that our bio-hybrid preparation shows behaviorally and physiologically valid forms of acquisition and extinction of the conditioned eye-blink response. Our neuro-prosthetic system learned to associate a tone with an air-puff, and as a result to trigger an anticipatory closure of the eyelid. Since the anaesthesia used impairs the

cerebellar circuitry dedicated to learning such a task, the observed eyeblinks are produced by the synthetic system. The fact that acquired CRs can be abolished by extinction training also indicates that the CRs result from a learning process induced by the contingent association of the CS and the US, thus reproducing the hallmark result of Pavlovian classical conditioning.

Here we have presented a step towards the enhancement and/or recovery of the capabilities of CNS through neuro-prosthetic solutions. Recently, another closed-loop solution targeting a different structure of the brain, the hippocampus (Berger et al., 2011) has been presented. This system, however, follows a different approach where first, the subject had to first acquire a specific stimulus response association that was recorded by the neuro-prosthetic system and subsequently the recorded state was used to recover this association after lesion to the hippocampus. Compared to this model, instead of aiming at restoring an acquired memory, our neuro-prosthesic targets to fully replace its target structure and to realize the capability to form new memories.

In a parallel effort, the computational model of the cerebellum here presented, together with the signal detection algorithms and the signal acquisition components, have been implemented in a low power VLSI (Bamford et al., 2012). Hence, with the results presented here we provide the last step in showing that our silicon cerebellum has satisfied all requirements of a neuro-prosthetic system.

An earlier version of the computational model presented here was implemented in an aVLSI platform and interfaced with a robot that was conditioned to a visual stimulus predicting a collision (Verschure & Mintz, 2001; Hofstotter et al., 2002, 2004). Thus, after showing that our approach allowed miniaturization and autonomous performance, we have now demonstrated that the model can be applied to the processing of inputs coming from a living brain that are specific to the computation

performed, i.e. CS, US, and trigger its specific output: the CR.

In our experiment, the IO channel provided the only teaching signal to the system. This channel displayed a spontaneous level of activity in the 0.5-2 Hz range, i.e., the level of activity expected in a single IO cell. However, in healthy animals, acquisition of an eyeblink CR is controlled by a cerebellar micro-complex, encompassing not one, but a number of IO cells. This imposes different constraints on the learning system because the IO derived error signal for our neuro-prosthetic is in all likelihood much impoverished as compared to the its biological counterpart. Hence, we expect that the key feature to strengthen in our approach is the quality and precision of the data acquisition of the biological preparation. For this we are planning further experiments in a chronic implant together with higher bandwidth physiology.

We reported two major caveats in the experimental preparation: the instability of the IO spontaneous rate, and presence of a stimulation artefact that precluded reliable read-out of the IO signal. Regarding the first problem, in the bio-hybrid experiment we computed the plasticity parameters assuming that the spontaneous IO rate inferred from the calibration data remained stable throughout the experiment. However, we observed that fluctuations in the spontaneous IO firing rate induced a drift in the w synaptic weight. Comparing this performance with simulated unpaired CS-US experiments, we saw that the performance with unpaired stimulation tended to be below the one observed in our experiment, but we also saw that the system triggered non-associative CRs. Next, we showed in simulations, that with an adaptive calibration method it is possible to control for the fluctuations in the IO activity such that they do not induce the acquisition of the these non-associative CRs. Thus we can conclude from the bio-hybrid experiment that our silicon cerebellum neuro-prosthetic can tuned to deal with marked fluctuations of its input brain signals. This also demonstrates the robustness of the learning principles implemented in the cerebellum and in particular the negative

feedback implemented in the DN-IO system.

Regarding the problem of the stimulation artifact, addressing it falls outside the scope of the current analysis. We emphasize that this problem is an issue of engineering of the stimulation system that the biological system does not encounter. We are investigating two solutions. First, given the very short duration of the stimulation pulses (see (Prueckl et al., 2011b)) it is possible to apply a more precise masking to the IO signal, timed to these pulses, that would minimize the signal loss. A second possibility is to avoid electrical stimulation altogether using optogenetic methods. These aspects need to be taken into account in a next iteration of the neural-prosthesis development.

To conclude, from a bio-engineering perspective we demonstrate that our approach supports outsourcing to a linked neuro-prosthetic system the acquisition and extinction of an adaptive reflex in an acute preparation. Given the modularity of the cerebellum, and the common assumption that the cerebellar algorithm performs similar computations throughout its different microcircuits (Dean et al., 2010; Albus, 1971; Marr, 1969), this work could be applied to support other adaptive reflexes as well, as long as their afferent and efferent circuitry could be identified. Additional further work is required to reproduce this result with the aVLSI synthetic system, testing this approach with a chronic implant, where one could check the stability of the acquired memory along days and study the long-term bio-compatibility.

Appendix 1: Mathematical definition of the cerebellar model

In what follows we provide a compact definition of the cerebellar computational model, describing each of the processes separately. The only additional notation introduced here for convenience are the sample time

s, that it is used to access the state of a variable in the previous step of the computation, and the the Heaviside function, $H(x)$, defined as

$$H(x) = \begin{cases} 1 & \text{if } x > 1 \\ 0 & \text{elsewhere} \end{cases}$$

The signals and the processes correspond to the ones presented in Fig. 5.3. Note that except $\mathbf{T}(t)$, $\mathbf{S}(t)$ and $w(t)$, that are real-valued, the rest of the signals are binary.

Process 1: Trace generation

$$\mathbf{T}(t) = \tau_0^{\mathbf{P}(t)} \left\{ \mathrm{H}(\mathbf{T}(t-s) - \tau_1) \left(\mathbf{T}(t-s) - \frac{\tau_0 - \tau_1}{\Lambda_\tau} \right) \right\}^{(1-\mathbf{P}(t))}$$

Process 2: Scaling

$$\mathbf{S}(t) = w(t)\mathbf{T}(t)$$

Process 3: Thresholding

$$\mathbf{F}(t) = H(\theta_{CR} - \mathbf{S}(t))\, H(\mathbf{S}(t-s) - \theta_{CR})$$

Process 4: Inhibitory Pulse

$$\mathbf{N}(t) = \mathbf{N}(t-s) + \mathbf{F}(t) - \mathbf{F}(t-\Lambda_\tau)$$

Process 5: Delay of inhibitory pulse

$$\mathbf{N}_\Lambda(t) = \mathbf{N}(t - \Lambda_{noi})$$

Process 6: Gating

$$\mathbf{C}(t) = \mathbf{I}(t)(1 - \mathbf{N}_\Lambda(t))$$

Process 7: Delay of plasticity trace

$$\mathbf{\Pi}(t) = H(\mathbf{T}(t - \Lambda_{noi}))$$

Process 8: Coincidence detection

$$w(t) = w(t - s) + \mathbf{\Pi}(t)(\delta_p - \mathbf{C}(t)\delta_d)$$

Appendix 2: Derivation of the update function of the adaptive calibration model

We assume that during spontaneous activity, detections in the IO and the PN are independent. Consequently, the probability of a simultaneous detection in both channels as $PN_{far}IO_{far}$.

$$\bar{D}_3 = PN_{far}\Lambda_\tau IO_{far}T_3 \tag{5.6}$$

where T_3 is the duration of the original recording used for the calibration and Λ_τ is the duration of the plasticity trace.

The linear system to optimize is given by:

$$\begin{bmatrix} \bar{P}_1 & \bar{D}_1 \\ \bar{P}_2 & \bar{D}_2 \\ \bar{P}_3 & \bar{D}_3 \end{bmatrix} \begin{bmatrix} \delta_p \\ -\delta_d \end{bmatrix} = \begin{bmatrix} -\Delta_a/T_a \\ \Delta_e/T_e \\ 0 \end{bmatrix} \tag{5.7}$$

The expressed in matrix notation becomes:

$$\mathbf{A}\,\mathbf{x} = \mathbf{b} \tag{5.8}$$

Adding the cost matrices for the weighted least squares we get:

$$\mathbf{C}\,\mathbf{A}\,\mathbf{x} = \mathbf{C}\,\mathbf{b}$$

where

$$\mathbf{C} = \begin{bmatrix} c_1 & 0 & 0 \\ 0 & c_2 & 0 \\ 0 & 0 & c_3 \end{bmatrix} \tag{5.9}$$

With this, the least squares solution for $\mathbf{x}$ is given by

$$\mathbf{x} = (\mathbf{A}^T\mathbf{C}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{C}\,\mathbf{b}$$

First, $\mathbf{A}^T\mathbf{C}\mathbf{A}$ expressed in terms of plasticity events and costs is the following matrix

$$\begin{bmatrix} c_1\bar{P}_1^2 + c_2\bar{P}_2^2 + c_3\bar{P}_3^2 & c_1\bar{P}_1\bar{D}_1 + c_2\bar{P}_2\bar{D}_2 + c_3\bar{P}_3\bar{D}_3 \\ c_1\bar{P}_1\bar{D}_1 + c_2\bar{P}_2\bar{D}_2 + c_3\bar{P}_3\bar{D}_3 & c_1\bar{D}_1^2 + c_2\bar{D}_2^2 + c_3\bar{D}_3^2 \end{bmatrix} \quad (5.10)$$

Grouping and renaming all the terms that do not depend on $\bar{D}_3$ we can express this matrix as

$$\begin{bmatrix} \chi_1 & \chi_2 + \psi_2\bar{D}_3 \\ \chi_2 + \psi_2\bar{D}_3 & \chi_3 + \psi_3\bar{D}_3^2 \end{bmatrix} \quad (5.11)$$

The expressions for the χ and ψ terms are given at the end of this appendix.

The determinant of this matrix, $\mathfrak{D}$, expressed as polynomial of $\bar{D}_3$ is

$$(\chi_1\psi_3 + \psi_2^2)\bar{D}_3^2 + 2\chi_2\psi_2\bar{D}_3 + (\chi_1\chi_3 - \chi_2^2)$$

using Eq. 5.6, we can simplify the notation, and express this determinant as a polynomial of IO_{far}, which is the input variable that will be updated at each re-calibration.

$$\beta_2 IO_{far}^2 + \beta_1 IO_{far} + \beta_0$$

Now we can express $(\mathbf{A}^T\mathbf{C}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{C}$ as

$$\frac{1}{\mathfrak{D}}\begin{bmatrix} \chi_3 + \psi_3\bar{D}_3^2 & -\chi_2 - \psi_2\bar{D}_3 \\ -\chi_2 - \psi_2\bar{D}_3 & \chi_1 \end{bmatrix}\begin{bmatrix} c_1\bar{P}_1 & c_2\bar{P}_2 & c_3\bar{P}_3 \\ c_1\bar{D}_1 & c_2\bar{D}_2 & c_3\bar{D}_3 \end{bmatrix}$$

where the first two factor are the inverse of $\mathbf{A}^T\mathbf{C}\mathbf{A}$ and the third term in the result of $\mathbf{A}^T\mathbf{C}$.

Performing the algebra, we obtain the following matrix expression

$$\begin{bmatrix} (\chi_3 + \psi_3 \bar{D}_3^2)c_1\bar{P}_1 - (\chi_2 + \psi_2 \bar{D}_3)c_1\bar{D}_1 & (\chi_3 + \psi_3 \bar{D}_3^2)c_2\bar{P}_2 - (\chi_2 + \psi_2 \bar{D}_3)c_2\bar{D}_2 \\ \chi_1 c_1 \bar{D}_1 - (\chi_2 + \psi_2 \bar{D}_3)c_1\bar{P}_1 & \chi_1 c_2 \bar{D}_2 - (\chi_2 + \psi_2 \bar{D}_3)c_2\bar{P}_2 \end{bmatrix}$$

were we have omitted the third column of the matrix, since it will be canceled by the zero in the third of row $\mathbf{b}$.

The last step to solve the system is to multiply by b, after what we obtain the solution for the δ_p and δ_p:

$$\begin{aligned} \delta_p = \quad & -\tfrac{1}{\mathfrak{D}} \tfrac{\Delta_a}{T_a} \left((\chi_3 + \psi_3 \bar{D}_3^2)c_1\bar{P}_1 - (\chi_2 + \psi_2 \bar{D}_3)c_1\bar{D}_1 \right) \\ & +\tfrac{1}{\mathfrak{D}} \tfrac{\Delta_e}{T_e} \left((\chi_3 + \psi_3 \bar{D}_3^2)c_2\bar{P}_2 - (\chi_2 + \psi_2 \bar{D}_3)c_2\bar{D}_2 \right) \end{aligned}$$

and

$$\begin{aligned} \delta_d = \quad & \tfrac{1}{\mathfrak{D}} \tfrac{\Delta_a}{T_a} \left(\chi_1 c_1 \bar{D}_1 - (\chi_2 + \psi_2 \bar{D}_3)c_1\bar{P}_1 \right) \\ & -\tfrac{1}{\mathfrak{D}} \tfrac{\Delta_e}{T_e} \left(\chi_1 c_2 \bar{D}_2 - (\chi_2 + \psi_2 \bar{D}_3)c_2\bar{P}_2 \right) \end{aligned}$$

where we see that $\bar{D}_3$ appears in the numerator of δ_p with degree 2 and in the numerator of δ_d with degree 1. In both cases, the denominator is given by the determinant $\mathfrak{D}$, which is another polynomial of degree 2. Therefore, we can express each parameter as a ratio of two polynomials.

$$\delta_p = \frac{\alpha_2 IO_{far}^2 + \alpha_1 IO_{far} + \alpha_0}{\beta_2 IO_{far}^2 + \beta_1 IO_{far} + \beta_0}$$

$$\delta_p = \frac{\gamma_1 IO_{far} + \gamma_0}{\beta_2 IO_{far}^2 + \beta_1 IO_{far} + \beta_0}$$

In consequence, the adaptive calibration algorithm requires to pre-compute 8 coefficients. The expression for each coefficient is detailed next:

$$
\begin{aligned}
\beta_0 &= (\chi_1\chi_3 - \chi_2^2) \\
\beta_1 &= 2\chi_2\psi_2 \\
\beta_2 &= (\chi_1\psi_3 + \psi_2^2) \\
\alpha_0 &= \tfrac{\Delta_a}{T_a}(\chi_2 c_1\bar{D}_1 - \chi_3 c_1\bar{P}_1) + \tfrac{\Delta_e}{T_e}(\chi_3 c_1\bar{P}_1 - \chi_2 c_2\bar{D}_2) \\
\alpha_1 &= \tfrac{\Delta_a}{T_a}\psi_2 c_1\bar{D}_1 - \tfrac{\Delta_e}{T_e}\psi_3 c_2\bar{D}_2 \\
\alpha_2 &= \tfrac{\Delta_e}{T_e}\psi_3 c_2\bar{P}_2 - \tfrac{\Delta_a}{T_a}\psi_3 c_1\bar{P}_1 \\
\gamma_0 &= \tfrac{\Delta_a}{T_a}(\chi_1 c_1\bar{D}_1 - \chi_2 c_1\bar{P}_1) + \tfrac{\Delta_e}{T_e}(\chi_1 c_2\bar{D}_2 - \chi_2 c_2\bar{P}_2) \\
\gamma_1 &= \tfrac{\Delta_e}{T_e}\psi_2 c_2\bar{P}_2 - \tfrac{\Delta_a}{T_a}\psi_2 c_1\bar{P}_1
\end{aligned}
$$

where χ and ψ are given by

$$
\begin{aligned}
\chi_1 &= c_1\bar{P}_1^2 + c_2\bar{P}_2^2 + c_3\bar{P}_3^2 \\
\chi_2 &= c_1\bar{P}_1\bar{D}_1 + c_2\bar{P}_2\bar{D}_2 \\
\chi_3 &= c_1\bar{D}_1^2 + c_2\bar{D}_2^2 \\
\psi_2 &= c_3\bar{P}_3 \\
\psi_3 &= c_3
\end{aligned}
$$

Abbreviation	Meaning
CNS	Central Nervous System
IO	Inferior Olive
PN	Pontine Nuclei
DN	Deep Nuclei
FN	Facial Nucleus
NOI	Nucleo Olivary Inhibition
Λ_{noi}	Latency of the NOI
δ_p	potentiation step size
δ_d	depression step size
TDR	proportion of true detections
FAR	False Alarm Rate
PN_{td}	proportion of true detections in the PN
PN_{far}	rate of false alarms in the PN
IO_{td}	proportion of true detections in the IO
IO_{far}	rate of false alarms in the IO
w	synaptic efficacy
PSTH	Peri-Stimulus Time Histogram
ω_{CS}	latency between CS trigger and CS event detection
ω_{US}	latency between US trigger and US event detection
ω_{CR}	latency between CR trigger and its physical execution
σ	probability of a US signal suppression by the NOI
P_i	number of potentiation events in the condition i
$\bar{P}_i$	mean number of potentiation events per trial in the condition i
D_i	number of depression events in the condition i
$\bar{D}_i$	mean number of depression events per trial in the condition i
Δ_a	estimated amount of change in w required for acquisition
Δ_e	estimated amount of change in w required for extinction
T_a	desired number of trials required for acquistion
T_e	desired number of trials required for extinction

Table 5.4: List of the abbreviations used in this chapter.

CHAPTER 6

Conclusion

This dissertation constitutes an effort towards the understanding of the function carried out by the cerebellum in the context of the acquisition of anticipatory actions and of the mechanisms that underlie such a function. To this end we have addressed the following theoretical questions: what is the representation of time in the granular layer of the cerebellar cortex? (Chapter 2); how is the teaching signal provided by the IO affected by its spontaneous activity, and how is this teaching signal modulated by the Nucleo Olivary Inhibition (NOI)? (Chapters 3 and 5)? and finally, how does the microcircuit studied in classical conditioning of the eye-blink response apply to the more ecologically valid context of avoidance learning? (Chapters 3 and 4). All these issues have been studied in an integrated manner, and tested in scenarios that were increasingly more constrained first by the embedding of the computational models in simulated and physical robots and finally by the testing of the functional validity of the solution in a neuro-prosthetic implementation.

This dissertation's main contributions, reported in the same order as they appear in the main text, are the following:

- We have introduced a parsimonious model that accounts for the

generation of a code of the passage of time in the cerebellar cortex, that is able to code intervals up to 1 second, in good match with the described range of operation of the cerebellum in timing tasks (Buhusi & Meck, 2005). We have proposed that the substrate for such a code is the well documented presence of slow spillover inhibitory currents in the cerebellar granule cells.

- We have shown that the cerebellar microcircuit studied in classical conditioning of the eye-blink response can be applied to an avoidance learning scenario. We have shown that a crucial component of this microcircuit, the NOI, shapes the behavior of the controlled agent, in this case a robot, causing the adaptive and reactive actions to become blended after training.

- We have accounted for the so-called CS-intensity effect of classical conditioning as the fingerprint of the cerebellum's capacity to acquire sensorimotor contingencies. Again, to demonstrate this we employed an avoidance learning setup where a robot traversed a track at gradually increasing velocities. We demonstrated that the sensorimotor contingency linking the intensity of the predictive stimulus with the latency of the avoidance response allowed the robot to generalize a learned skill, in this case a turn, along different speeds of execution.

- We have demonstrated that a computational model of a structure of the central nervous system, in this case the cerebellum, can be used to implement a neuro-prosthetic device allowing the functional replacement of its inactivated biological counterpart. Regarding this contribution, we have provided a complete analysis of this pioneer result, and also provided an evolved version of the computational model used in the experiment to be tested in the next iteration of the development of the cerebellar neuro-prosthesis.

Now we will address each of these contributions in more detail, discussing the opportunities of further research that they open.

6.1 Coding of time in the cerebellar cortex

In Chapter 2 we have introduced which, to our knowledge, is the more parsimonious model of the encoding of the passage of time in the cerebellar cortex. Yamazaki & Tanaka (2007a) and Medina et al. (2000), among others, have likened the computation at the granular layer of the cerebellar cortex to that of a reservoir in echo-state networks (Jaeger, 2003) or liquid-state machines (Maass et al., 2002). However, such a hypothesis lacks experimental evidence since, so far, the response of granule cells to sustained stimuli seems to lack the complex dynamics required for such models (Bengtsson & Jörntell, 2009b; Jörntell & Ekerot, 2006).

In comparison, from the output perspective, our coding hypothesis simply assumes that granule cells progressively adapt their response to sustained stimuli, as it was shown for example by Jörntell & Ekerot (2006), and that the rate of the adaptation is in the same order of magnitude as the timing capacities of the cerebellum, i.e., up to 1 second. Based on the well-known presence of spillover inhibition in the granular layer (Hamann et al., 2002; Rossi & Hamann, 1998) we further hypothesized that the mechanism causing such adaptation is the build-up of such slow inhibitory currents in granule cells, and demonstrated such a hypothesis with a computational simulation. After having presented this work as a poster (Herreros et al., 2009), this idea found experimental support in a work by Crowley et al. (2009), where the accumulation of inhibitory spillover currents was measured intra-cellularly *in vitro*.

This by no means implies that we have provided the definite answer to the coding of time in the cerebellar cortex, since some of our assumptions still need experimental support, such as the dynamics of the Golgi cells dur-

ing eye-blink conditioning. Regarding the type of code, such a question will be answered once the response of granule cells to the conditioning stimulus is recorded *in vivo* during an eye-blink conditioning experiment either electrophysiologically or by imaging techniques. It is very likely that the response will come out soon, since this type of experiment is now in the agenda of the field (Rasmussen et al., 2013b; Halverson et al., 2010).

The question regarding the substrate of the code is more complex to answer since spillover inhibition and standard fast synaptic inhibition cannot be differentially blocked, as both are mediated by the same receptors (Hamann et al., 2002). For this, a blockade of spillover inhibition will also affect local synaptic transmission, or in other words, remove inhibition from the whole granular layer circuit, what will presumably have widespread consequences in cerebellar function. Therefore, finding the mechanism that supports such a capacity will require intra-cellular recordings.

However, we have indirectly validated our timing hypothesis by testing its functional implications. In Chapters 3 and 4 we employed such a representation of time in an adaptive controller that was tested in a robot collision avoidance task. The controller was able to learn that task and there were qualitative similarities between the behavior of the robot and the one observed in nature in similar associative learning tasks. For instance, in our model the learning of a well-timed response progressed by incrementally delaying the peak of the initial CR and the same has been reported for mice in classical conditioning (Koekkoek et al., 2003).

6.2 The cerebellar microcircuit in avoidance learning

An idiosyncrasy of the eye-blink conditioning paradigm is that, in principle, the CR should be devoid of any adaptive value, i.e., it should not

ameliorate the US by making it less noxious. This is not the case in the more ecologically valid setup of avoidance learning, where, in contrast, the elicited protective behavior reduces the noxiousness of the aversive stimulus (Kreider & Mauk, 2010). For this it was unclear how the functionality of the cerebellar microcircuit studied in classical conditioning generalized to the avoidance learning scenario, an issue of poignant interest since, arguably, in nature the function of such a microcircuit is homologous to avoidance learning, i.e., to support the acquisition of anticipatory reflexes. Specifically, a major caveat was to discern whether the internal negative feedback provided by the NOI was functional in avoidance learning or if, instead, it interfered with the negative behavioral feedback, that is non-existent in classical conditioning. This matter is the focus of Chapter 3.

There, we used the cerebellar microcircuit as the core of a controller for a robotic collision avoidance task. Such approach highlighted an aspect that until very recently was overlooked in the eye-blink conditioning literature: that the US does not necessarily have to raise an internal all-or-none signal, but that it can be coded in a graded manner (Rasmussen et al., 2013a; Najafi & Medina, 2013)). Indeed, it is well-known in the field that rather than stereotypical discrete responses both the CR and the UR have temporal dynamics and graded amplitudes that evolve through the training process. Our results claim the same status for the US signal.

This distinction is relevant since it allows a new interpretation of the goal of the cerebellum in avoidance learning. Such a goal being to attenuate the noxiousness of the aversive stimulus through an anticipatory protective action bringing the stimulus intensity below a certain safety level. That is, the aversive stimulus is not completely erased because 1) it is not necessary to do so once its intensity is sufficiently diminished and because 2) its presence validates the suitability of the avoidance action. For the last point one has to consider that protective actions may have

a *cost*, e.g., blinking has the cost of interrupting the visual input stream, therefore it may be desirable to *avoid avoiding too much*. In that case, the presence of a residual US signal allows to confirm the appropriateness of the action at every single execution while its absence drives extinction.

Besides highlighting these points, the fundamental contribution of our proposal in Chapter 3 is that the balance between the anticipatory and the reactive actions can be set by a single parameter, this being the gain of the NOI. The next logical step in future experiments is to find a principled way to adjust such a parameter, for instance, in terms of cost minimization. We are already addressing this issue form a computational perspective (Brandi, 2013) but, in addition, it will be of major interest to devise an eye-blink conditioning paradigm where the cost the anticipatory action is under experimental control, to see whether we can act upon the balance of adaptive and reactive actions. This last point would demonstrate a second degree of adaptivity in the acquisition of avoidance actions; a meta-learning of the adaptive reflex.

6.3 Acquisition of sensorimotor contingencies by the cerebellum

In an example of a transfer of knowledge from experimental psychology to robotics, we have shown that the cerebellum can use a sensorimotor contingency like the CS-intensity effect to enable the flexible execution of motor skills. Indeed, the reproduction of the CS-intensity effect, that in Chapter 2 we used as a means to further validate our model for the representation of time in the cerebellar cortex, now is upgraded from an experimental phenomenon to a critical feature of cerebellar operation, along the lines of what was put forth in Svensson et al. (1997, 2010). This is, it shows that by controlling the intensity of the cerebellar input it might be possible to flexibly adjust the timing of its output.

We demonstrated this result in a real robot setup, thus providing a more

constrained validation of the applicability of our cerebellar-based control model. Even though this point was not elaborated in the main text, in the experiments we observed that the anticipatory responses triggered in the real arena were wider than in the simulated environment and, for this, the real robot turned more smoothly than the simulated one, even though both were mounted with the exact same controller. A possible explanation for this is that the variability of the movement of the physical robot introduced variability in its perceptions, for instance, flattening the distribution of the error signal across trials, and this impacted on the learned response, that became flatter as well. In other words, this illustrates, as put forth in Verschure et al. (2003), that there are properties of an autonomous learning system that only manifest themselves when the synergies between behaviour and perception enter the picture.

It has been recently proposed that a control framework very similar to the one presented here can be used as a basis for safe limb control (Dean et al., 2013), which opens an array of applications in the emerging field of soft robotics. A crucial step in this direction can be taken by simply replacing the externally generated CS by an internal signal. In that case we would be able to apply this control scheme to support the skilled execution of volitional actions. Or, in other words, to provide a grounded robotic demonstration of the 170 years-old hypothesis advanced by Flourens (1842): that the cerebellum shapes and coordinates the motor commands triggered by other brain structures.

6.4 Recovery of a lost cerebellar learning function through neuro-prosthetics.

With the neuro-prosthetic application (Chapter 5) we showed that our understanding of the information processing of the cerebellar microcircuit allowed us to build a synthetic microcircuit that can functionally replace its biological counterpart in the context of eye-blink conditioning. Such

a result required not only the work presented in this thesis, but also the coordinated effort of a broad and multidisciplinary team (see Prueckl et al. (2011b) and Bamford et al. (2012)).

Regarding the computational model, the practical constraints of the experiment, i.e., the intended implementation in an aVLSI device, informed its design. For this we implemented a model along the line of Verschure & Mintz (2001) while incorporating aspects highlighted in the previous chapters of this dissertation, i.e., the effect of the low spontaneous activity of the IO on the coding of the teaching signal.

The results obtained showed that it is possible to design closed-loop neuro-prosthetic systems that replicate the function of a brain structure by reproducing, to some extent, the anatomical layout of the original structure. The novelty of our approach resides in that, besides performing the same input-output transformation of the biological system, by replicating its neuro-anatomy, our neuro-prosthesis implements a similarly structured algorithm.

In addition, the goal of our neuro-prosthesis was not just to rebuild a damaged link between different brain areas, or to reproduce previously observed patterns of activation thereby restoring already acquired memories, but to recover the capability to form new memories. In the long run, these are the first steps on a vast and uncharted territory: that of the direct closed-loop interaction of our brain with machines. Leaving ethical concerns aside, and the metaphysical considerations regarding how this will affect he concept of *self*, we have illustrated the possibility to develop neuroscience-based treatments for the diseases that affect specific structures of the central nervous system.

In the short run, we will see an incremental increase in the computational capacity of the neural-prosthesis, in the bio-compatibility of the implants, etc. but the real progress in this field, requires clearly delineated theories on the function of the different structures of the central nervous system.

6.5 Concluding remarks

The computational models introduced in this dissertation are currently being used to address other aspects of cerebellar function that were not dealt with here, such as the generation motor sequences (Brandi et al., 2013), the minimization of the overall cost in avoidance behavior (Brandi, 2013) and the acquisition of anticipatory postural adjustments (Maffei et al., 2013), what confirms that such models were not an end in themselves but that they constitute a means for performing further research and advance our general understanding of the cerebellar function.

How general is the cerebellar algorithm that he have gradually developed throughout this work is a question that will only be answered in the future by extending its range of application. Here, in the last robot experiments that we have presented, we applied the microcircuit design to an avoidance learning task that was performed at different speeds. This is already a long shot from the initial computational simulation of the eyeblink conditioning paradigm. But, as we mentioned in the introduction, the diversity of theories about the cerebellar function suggests, to us, that the cerebellum serves a diversity of functions.

The next step is to include this circuit in the loop of voluntary motor control, using as the triggers for the cerebellar output internal signals and not external stimuli. These signals will be now generated by higher brain structures, for which it will be necessary to embed the computational models of the cerebellum in complex cognitive structures. And as the volitional commands reaching the cerebellum will become more complex, composed of multiple concurrent or sequential elements, the cerebellum will have to coordinate them, probably by adjusting the relative timing of execution of each sub-command.

Finally, an exciting venue is to model the control of cognitive operations by the cerebellum (Ito, 2008). Then we could address issues such as how does the cerebellum might contribute to logical reasoning? (Balsters &

Ramnani, 2011) Or, how does it contribute to our conscious experience? (Itō, 2012; Schmahmann et al., 2002).

In general, these questions will require to embed the cerebellum in elaborated cognitive architectures, such as the one provided by the distributive adaptive control theory (Verschure, 2012), and explicitly delineate the interaction between its different components and the cerebellum. This has the potential of leading towards the development of robots or autonomous agents that will not only be more agile from a physical point of view, but also be more adaptive from a cognitive perspective.

Following this path we will begin to provide a general answer to the questions of: What is the key computational feature that the cerebellum adds to the central nervous system of (almost) all vertebrates? Why is it preserved all through evolution almost since the body plan of vertebrates was originated? And how does it enhance the fitness of living organisms? We believe that the integrative approach we took in this thesis is the key towards providing a joint answer to all of the above questions.

Bibliography

Albus, J. (1971). A theory of cerebellar function. *Mathematical Biosciences*, 10(1-2):25–61.

Albus, J. S. et al. (1975). A new approach to manipulator control: The cerebellar model articulation controller (cmac). *Journal of dynamic systems, measurement and control*, 97(3):220–227.

Andersson, G., Garwicz, M., & Hesslow, G. (1988). Evidence for a gaba-mediated cerebellar inhibition of the inferior olive in the cat. *Experimental Brain Research*, 72(3):450–456.

Andersson, G. & Hesslow, G. (1987). Activity of purkinje cells and interpositus neurones during and after periods of high frequency climbing fibre activation in the cat. *Experimental Brain Research*, 67(3):533–542.

Angelaki, D. E., Shaikh, A. G., Green, A. M., & Dickman, J. D. (2004). Neurons compute internal models of the physical laws of motion. *Nature*, 430(6999):560–564.

Apps, R. & Garwicz, M. (2005). Anatomical and physiological foundations of cerebellar information processing. *Nature Reviews Neuroscience*, 6(4):297–311.

Balsters, J. H. & Ramnani, N. (2011). Cerebellar plasticity and the automation of first-order rules. *The Journal of Neuroscience*, 31(6):2305–2312.

Bamford, S., Hogri, R., Giovannucci, A., Taub, A., Herreros, I., Verschure, P., Mintz, M., & Del Giudice, P. (2012). A vlsi field-programmable mixed-signal array to perform neural signal processing and neural modeling in a prosthetic system. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, 20(4):455–467.

Barton, R. A. (2012). Embodied cognitive evolution and the cerebellum. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1599):2097–2107.

Barton, R. A. & Harvey, P. H. (2000). Mosaic evolution of brain structure in mammals. *Nature*, 405(6790):1055–1058.

Bastian, A. J. (2006). Learning to predict the future: the cerebellum adapts feedforward movement control. *Current opinion in neurobiology*, 16(6):645–649.

Bell, C. (2002). Evolution of cerebellum-like structures. *Brain, Behavior and Evolution*, 59(5-6):312–326.

Bell, C. C., Han, V., & Sawtell, N. B. (2008). Cerebellum-like structures and their implications for cerebellar function. *Annu. Rev. Neurosci.*, 31:1–24.

Bengtsson, F. & Hesslow, G. (2006). Cerebellar control of the inferior olive. *The Cerebellum*, 5(1):7–14.

Bengtsson, F. & Jörntell, H. (2009a). Sensory transmission in cerebellar granule cells relies on similarly coded mossy fiber inputs. *Proceedings of the National Academy of Sciences*, 106(7):2389.

Bengtsson, F. & Jörntell, H. (2009b). Sensory transmission in cerebellar granule cells relies on similarly coded mossy fiber inputs. *Proceedings of the National Academy of Sciences*, 106(7):2389–2394.

Berger, T., Hampson, R., Song, D., Goonawardena, A., Marmarelis, V., & Deadwyler, S. (2011). A cortical neural prosthesis for restoring and enhancing memory. *Journal of neural engineering*, 8:046017.

Bernardet, U., Blanchard, M., & Verschure, P. (2002). IQR: a distributed system for real-time real-world neuronal simulation. *Neurocomputing*, 44:1043–1048.

Bobo, L., Herreros Alonso, I., & Verschure, P. F. (2012). A digital neuromorphic implementation of cerebellar associative learning. Dins *Biomimetic and Biohybrid Systems*, ps. 13–25. Springer.

Boele, H., Koekkoek, S., & De Zeeuw, C. (2009). Cerebellar and extracerebellar involvement in mouse eyeblink conditioning: the acdc model. *Frontiers in Cellular Neuroscience*, 3.

Bower, J. M. (1997a). Control of sensory data acquisition. *International review of neurobiology*, 41:489–513.

Bower, J. M. (1997b). Is the cerebellum sensory for motor's sake, or motor for sensory's sake: the view from the whiskers of a rat? *Progress in brain research*, 114:463–496.

Braitenberg, V. & Atwood, R. P. (1958). Morphological observations on the cerebellar cortex. *Journal of Comparative Neurology*, 109(1):1–33.

Braitenberg, V., Heck, D., & Sultan, F. (1997). The detection and generation of sequences as a key to cerebellar function: experiments and theory. *Behavioral and Brain Sciences*, 20(02):229–245.

Brandi, S. (2013). *Optimization of anticipatory reflexes on a computational model of the cerebellum.* Projecte F. de Carrera o Tesina de L., Univeristat Pompeu Fabra, Spain.

Brandi, S., Herreros, I., Sánchez-Fibla, M., & Verschure, P. F. (2013). Learning of motor sequences based on a computational model of the

cerebellum. Dins *Biomimetic and Biohybrid Systems*, ps. 356–358. Springer.

Brindley, G. (1964). The use made by the cerebellum of the information that it receives from sense organs. *International Brain Research Organization Bulletin*, 3(3):80.

Brunel, N., Hakim, V., Isope, P., Nadal, J., & Barbour, B. (2004). Optimal Information Storage and the Distribution of Synaptic Weights:: Perceptron versus Purkinje Cell. *Neuron*, 43(5):745–757. ISSN 0896-6273.

Buhusi, C. V. & Meck, W. H. (2005). What makes us tick? functional and neural mechanisms of interval timing. *Nature Reviews Neuroscience*, 6(10):755–765.

Buonomano, D. & Mauk, M. (1994). Neural network model of the cerebellum: temporal discrimination and the timing of motor responses. *Neural Computation*, 6(1):38–55.

Chadderton, P., Margrie, T., & Häusser, M. (2004). Integration of quanta in cerebellar granule cells during sensory processing. *Nature*, 428(6985):856–860. ISSN 0028-0836.

Chapin, J. K., Moxon, K. A., Markowitz, R. S., & Nicolelis, M. A. (1999). Real-time control of a robot arm using simultaneously recorded neurons in the motor cortex. *Nature neuroscience*, 2(7):664–670.

Chaumont, J., Guyon, N., Valera, A. M., Dugué, G. P., Popa, D., Marcaggi, P., Gautheron, V., Reibel-Foisset, S., Dieudonné, S., Stephan, A., et al. (2013). Clusters of cerebellar purkinje cells control their afferent climbing fiber discharge. *Proceedings of the National Academy of Sciences*, p. 201302310.

Cheng, D. T., Disterhoft, J. F., Power, J. M., Ellis, D. A., & Desmond, J. E. (2008). Neural substrates underlying human delay and trace eyeblink conditioning. *Proceedings of the National Academy of Sciences*, 105(23):8108–8113.

Christian, K. & Thompson, R. (2003). Neural substrates of eyeblink conditioning: acquisition and retention. *Learning & Memory*, 10(6):427.

Clark, R. & Squire, L. (1998). Classical conditioning and brain systems: the role of awareness. *Science*, 280(5360):77–81.

Cohen, E. D. (2007). Prosthetic interfaces with the visual system: biological issues. *Journal of neural engineering*, 4(2):R14.

Coombs, J., Eccles, J., & Fatt, P. (1955). The specific ionic conductances and the ionic movements across the motoneuronal membrane that produce the inhibitory post-synaptic potential. *The Journal of Physiology*, 130(2):326.

Crowley, J., Fioravante, D., & Regehr, W. (2009). Dynamics of fast and slow inhibition from cerebellar Golgi cells allow flexible control of synaptic integration. *Neuron*, 63(6):843–853.

D'Angelo, E. & De Zeeuw, C. (2009). Timing and plasticity in the cerebellum: focus on the granular layer. *Trends in neurosciences*, 32(1):30–40. ISSN 0166-2236.

De Zeeuw, C. I., Simpson, J. I., Hoogenraad, C. C., Galjart, N., Koekkoek, S., Ruigrok, T., et al. (1998). Microcircuitry and function of the inferior olive. *Trends in neurosciences*, 21(9):391.

Dean, P., Anderson, S., Porrill, J., & Jörntell, H. (2013). An adaptive filter model of cerebellar zone c3 as a basis for safe limb control? *The Journal of physiology*.

Dean, P., Porrill, J., Ekerot, C., & Jörntell, H. (2010). The cerebellar microcircuit as an adaptive filter: experimental and computational evidence. *Nature Reviews Neuroscience*, 11(1):30–43.

Demer, J., Echelman, D., & Robinson, D. (1985). Effects of electrical stimulation and reversible lesions of the olivocerebellar pathway on purkinje cell activity in the flocculus of the cat. *Brain research*, 346(1):22–31.

Devor, A. (2000). Is the cerebellum like cerebellar-like structures? *Brain research reviews*, 34(3):149–156.

Eccles, J., Ito, M., & Szentágothai, J. (1967). *The cerebellum as a neuronal machine*. Springer Berlin.

Eccles, J., Llinas, R., & Sasaki, K. (1966a). The excitatory synaptic action of climbing fibres on the purkinje cells of the cerebellum. *The Journal of Physiology*, 182(2):268–296.

Eccles, J., Llinas, R., & Sasaki, K. (1966b). Parallel fibre stimulation and the responses induced thereby in the purkinje cells of the cerebellum. *Experimental Brain Research*, 1(1):17–39.

Eddington, D. K., Dobelle, W. H., Brackmann, D., Mladejovsky, M., & Parkin, J. (1978). Auditory prostheses research with multiple channel intracochlear stimulation in man. *The Annals of otology, rhinology, and laryngology*, 87(6 Pt 2):1.

Fag, A. H., Sitkoff, N., Houk, J. C., & Barto, A. G. (1997). Cerebellar Learning for Control of a Two-Link Arm in Muscle Space. *Robotics and Automation, IEEE Intemational Conference*, (April):2638–2644.

Fibla, M. S., Bernardet, U., & Verschure, P. F. (2010). Allostatic control for robot behaviour regulation: An extension to path planning. Dins *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, ps. 1935–1942. IEEE.

Fitch, W. T. (2011). The evolution of syntax: an exaptationist perspective. *Frontiers in evolutionary neuroscience*, 3.

Flourens, P. (1842). *Recherches expérimentales sur les propriétés et les fonctions du système nerveux dans les animaux vertébrés.* Ballière.

Freeman, J. H. & Steinmetz, A. B. (2011). Neural circuitry and plasticity mechanisms underlying delay eyeblink conditioning. *Learning & Memory*, 18(10):666–677.

Fujita, M. (1982). Adaptive filter model of the cerebellum. *Biological Cybernetics*, 45(3):195–206. ISSN 0340-1200.

Gao, J.-H., Parsons, L. M., Bower, J. M., Xiong, J., Li, J., & Fox, P. T. (1996). Cerebellum implicated in sensory acquisition and discrimination rather than motor control. *Science*, 272(5261):545–547.

Garcia, K., Mauk, M., Weidemann, G., & Kehoe, E. (2003). Covariation of alternative measures of responding in rabbit (oryctolagus cuniculus) eyeblink conditioning during acquisition training and tone generalization. *Behavioral neuroscience*, 117(2):292.

Gilbert, P. & Thach, W. (1977). Purkinje cell activity during motor learning. *Brain research*, 128(2):309–328.

Giovannucci, A., Bamford, S., Herreros-Alonso, I., Taub, A., Hogri, R., Prueckl, R., Zucca, R., Guger, C., Mintz, M., Silmon, A., Giudice, P. D., & Verschure, P. F. M. J. (2010). Replacing a cerebellar microcircuit with an autonomous neuroprosthetic device. Dins *Program No. 786.18. 2010 Neuroscience Meeting Planner. San Diego, CA: Society for Neuroscience, 2010. Online.*

Glickstein, M. (1994). Cerebellar agenesis. *Brain*, 117(5):1209–1212.

Glickstein, M. (2007). What does the cerebellum really do? *Current Biology*, 17(19):R824–R827.

Gormezano, I., Prokasy, W., & Thompson, R. (1987). *Classical conditioning*. Lawrence Erlbaum. ISBN 089859507X.

Grillner, S., Hellgren, J., Menard, A., Saitoh, K., & Wikström, M. A. (2005). Mechanisms for selection of basic motor programs–roles for the striatum and pallidum. *Trends in neurosciences*, 28(7):364–370.

Halverson, H. E., Davis, T., & Mauk, M. D. (2010). Neuronal responses in cerebellar cortex during associative learning.

Hamann, M., Rossi, D., & Attwell, D. (2002). Tonic and spillover inhibition of granule cells control information flow through cerebellar cortex. *Neuron*, 33(4):625–633.

Häusser, M. & Clark, B. (1997). Tonic synaptic inhibition modulates neuronal output pattern and spatiotemporal synaptic integration. *Neuron*, 19(3):665–678.

Herreros, I., Zucca, R., Giovannucci, A., & Verschure, P. F. M. J. (2009). The interaction of purkinje cell and inhibitory interneuron plasticity during classical conditioning. doi: 10.3389/conf.neuro.06.2009.03.190.

Herreros Alonso, I., Giovannucci, A., & Verschure, P. F. (2013a). Replacing a cerebellar microcircuit by a synthetic system. *submitted*.

Herreros Alonso, I., Maffei, G., Brandi, S., Sanchez Fibla, M., & Verschure, P. F. (2013b). Speed generalization capabilities of a cerebellar model on a rapid navigation task. Dins *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on*, ps. 8108–8113. IEEE.

Herreros Alonso, I. & Verschure, P. F. (2013a). Nucleo-olivary inhibition balances the interaction between the reactive and adaptive layers in motor control. *Neural Networks*.

Herreros Alonso, I. & Verschure, P. F. (2013b). Slow inhibitory currents as a substrate for the representation of time in the cerebellar cortex. *in preparation.*

Hesslow, G. (1986). Inhibition of inferior olivary transmission by mesencephalic stimulation in the cat. *Neuroscience letters*, 63(1):76–80.

Hesslow, G. (1994). Inhibition of classically conditioned eyeblink responses by stimulation of the cerebellar cortex in the decerebrate cat. *The Journal of Physiology*, 476(2):245.

Hesslow, G. & Ivarsson, M. (1994). Suppression of cerebellar Purkinje cells during conditioned responses in ferrets. *Neuroreport*, 5:649–652.

Hesslow, G. & Yeo, C. (2002). The functional anatomy of skeletal conditioning. Dins *A neuroscientist s guide to classical conditioning. Springer, New York*, ps. 86–146. Springer.

Hochberg, L. R., Bacher, D., Jarosiewicz, B., Masse, N. Y., Simeral, J. D., Vogel, J., Haddadin, S., Liu, J., Cash, S. S., van der Smagt, P., et al. (2012). Reach and grasp by people with tetraplegia using a neurally controlled robotic arm. *Nature*, 485(7398):372–375.

Hofstotter, C., Gil, M., Eng, K., Indiveri, G., Mintz, M., Kramer, J., & Verschure, P. (2004). The cerebellum chip: an analog vlsi implementation of a cerebellar model of classical conditioning. *Advances in Neural Information Processing Systems*, 17:577–584.

Hofstotter, C., Mintz, M., & Verschure, P. (2002). The cerebellum in action: A simulation and robotics study. *European Journal of Neuroscience*, 16(7):1361–1376.

Holtzman, T., Mostofi, A., Phuah, C., & Edgley, S. (2006). Cerebellar Golgi cells in the rat receive multimodal convergent peripheral inputs via the lateral funiculus of the spinal cord. *The Journal of Physiology*, 577(1):69.

Holtzman, T., Sivam, V., Zhao, T., Frey, O., Wal, P., Koudelka-Hep, M., Dalley, J., & Edgley, S. (2011). Multiple extra-synaptic spillover mechanisms regulate prolonged activity in cerebellar golgi cell-granule cell loops. *The Journal of Physiology*.

Inderbitzin, M., Herreros Alonso, I., & Verschure, P. F. M. J. (2010). An integrated computational model of the two phase theory of conditioning. *World Congress of Computational Intelligence*.

Ito, M. (1972). Neural design of the cerebellar motor control system. *Brain research*, 40(1):81–84.

Ito, M. (2008). Control of mental activities by internal models in the cerebellum. *Nature Reviews Neuroscience*, 9(4):304–313.

Itō, M. (2012). *Cerebellum: The Brain for an Implicit Self*. FT Press.

Ito, M., Sakurai, M., & Tongroach, P. (1982). Climbing fibre induced depression of both mossy fibre responsiveness and glutamate sensitivity of cerebellar Purkinje cells. *The Journal of Physiology*, 324(1):113.

Ivry, R. B. & Keele, S. W. (1989). Timing functions of the cerebellum. *Journal of Cognitive Neuroscience*, 1(2):136–152.

Jaeger, H. (2003). Adaptive nonlinear system identification with echo state networks. Dins *Advances in Neural Information Processing Systems*, ps. 593–600.

Jirenhed, D., Bengtsson, F., & Hesslow, G. (2007). Acquisition, extinction, and reacquisition of a cerebellar cortical memory trace. *Journal of Neuroscience*, 27(10):2493.

Jörntell, H. & Ekerot, C. (2006). Properties of somatosensory synaptic integration in cerebellar granule cells in vivo. *Journal of Neuroscience*, 26(45):11786.

Jörntell, H. & Hansel, C. (2006). Synaptic memories upside down: bidirectional plasticity at cerebellar parallel fiber-Purkinje cell synapses. *Neuron*, 52(2):227–238.

Kalmbach, B., Ohyama, T., Kreider, J., Riusech, F., & Mauk, M. (2009). Interactions between prefrontal cortex and cerebellum revealed by trace eyelid conditioning. *Learning & Memory*, 16(1):86. ISSN 1072-0502.

Kehoe, E. & Macrae, M. (2002a). Fundamental behavioral methods and findings in classical conditioning. Dins *A neuroscientist s guide to classical conditioning. Springer, New York*, ps. 171–231.

Kehoe, E. & White, N. (2004). Overexpectation: Response loss during sustained stimulus compounding in the rabbit nictitating membrane preparation. *Learning & Memory*, 11(4):476. ISSN 1072-0502.

Kehoe, E. J. & Macrae, M. (2002b). Fundamental behavioral methods and findings in classical conditioning. Dins *A neuroscientist s guide to classical conditioning*, ps. 171–231. Springer.

Knusel, P., Wyss, R., Koenig, P., & Verschure, P. (2004). Decoding a temporal population code. *Neural computation*, 16(10):2079–2100.

Koekkoek, S., Hulscher, H., Dortland, B., Hensbroek, R., Elgersma, Y., Ruigrok, T., & De Zeeuw, C. (2003). Cerebellar LTD and learning-dependent timing of conditioned eyelid responses. *Science*, 301(5640):1736.

Konorski, J. (1967). Integrative activity of the brain.

Kreider, J. C. & Mauk, M. D. (2010). Eyelid conditioning to a target amplitude: adding how much to whether and when. *The Journal of Neuroscience*, 30(42):14145–14152.

Lepora, N., Porrill, J., Yeo, C., & Dean, P. (2010). Sensory prediction or motor control? application of marr-albus type models of cerebellar function to classical conditioning. *Frontiers in Computational Neuroscience*, (0):12. ISSN 1662-5188. doi: 10.3389/fncom.2010.00140.

Linden, D. J. & Connor, J. A. (1991). Participation of postsynaptic pkc in cerebellar long-term depression in culture. *Science*, 254(5038):1656–1659.

Llinas, R. & Welsh, J. (1993). On the cerebellum and motor learning. *Current opinion in neurobiology*, 3(6):958–965.

Luque, N. R., Garrido, J. A., Carrillo, R. R., Tolu, S., & Ros, E. (2011). Adaptive cerebellar spiking model embedded in the control loop: Context switching and robustness against noise. *International Journal of Neural Systems*, 21(05):385–401.

Maass, W., Natschläger, T., & Markram, H. (2002). Real-time computing without stable states: A new framework for neural computation based on perturbations. *Neural computation*, 14(11):2531–2560.

Mackintosh, N. (1974). *The psychology of animal learning*. Academic Press New York.

Maffei, G., Herreros, I., Sánchez-Fibla, M., & Verschure, P. F. (2013). Acquisition of anticipatory postural adjustment through cerebellar learning in a mobile robot. Dins *Biomimetic and Biohybrid Systems*, ps. 399–401. Springer.

Mapelli, J., Gandolfi, D., & D'Angelo, E. (2010). Combinatorial responses controlled by synaptic inhibition in the cerebellum granular layer. *Journal of neurophysiology*, 103(1):250–261. ISSN 0022-3077.

Marr, D. (1969). A theory of cerebellar cortex. *The journal of physiology*, 202(2):437.

Mauk, M., Steinmetz, J., & Thompson, R. (1986). Classical conditioning using stimulation of the inferior olive as the unconditioned stimulus. *Proceedings of the National Academy of Sciences*, 83(14):5349.

McCormick, D. A. & Thompson, R. F. (1984). Cerebellum: essential involvement in the classically conditioned eyelid response. *Science*, 223(4633):296–299.

McKinstry, J., Edelman, G., & Krichmar, J. (2006). A cerebellar model for predictive motor control tested in a brain-based device. *Proceedings of the National Academy of Sciences of the United States of America*, 103(9):3387–3392.

Medina, J., Garcia, K., & Mauk, M. (2001). A mechanism for savings in the cerebellum. *The Journal of Neuroscience*, 21(11):4081–4089.

Medina, J., Garcia, K., Nores, W., Taylor, N., & Mauk, M. (2000). Timing mechanisms in the cerebellum: testing predictions of a large-scale computer simulation. *Journal of Neuroscience*, 20(14):5516.

Medina, J. & Mauk, M. (2000). Computer simulation of cerebellar information processing. *Nature Neuroscience*, 3:1205–1211.

Medina, J., Nores, W., & Mauk, M. (2002). Inhibition of climbing fibres is a signal for the extinction of conditioned eyelid responses. *Nature*, 416(6878):330–333.

Miall, R., Christensen, L., Cain, O., & Stanley, J. (2007). Disruption of state estimation in the human lateral cerebellum. *PLoS Biol*, 5(11):e316.

Miall, R., Weir, D., Wolpert, D., & Stein, J. (1993). Is the cerebellum a smith predictor? *Journal of motor behavior*, 25:203–203.

Mitchell, S. & Silver, R. (2000a). GABA spillover from single inhibitory axons suppresses low-frequency excitatory transmission at the cerebellar glomerulus. *Journal of Neuroscience*, 20(23):8651.

Mitchell, S. & Silver, R. (2000b). Glutamate spillover suppresses inhibition by activating presynaptic mGluRs. *Nature*, 404(6777):498–502.

Mondada, F., Bonani, M., Raemy, X., Pugh, J., Cianci, C., Klaptocz, A., Magnenat, S., Zufferey, J.-C., Floreano, D., & Martinoli, A. (2009). The e-puck, a robot designed for education in engineering.

Montgomery, J. C., Bodznick, D., & Yopak, K. E. (2012). The cerebellum and cerebellum-like structures of cartilaginous fishes. *Brain, Behavior and Evolution*, 80(2):152–165.

Moritz, C. T., Perlmutter, S. I., & Fetz, E. E. (2008). Direct control of paralysed muscles by cortical neurons. *Nature*, 456(7222):639–642.

Moyer, J. R., Deyo, R. A., & Disterhoft, J. F. (1990). Hippocampectomy disrupts trace eye-blink conditioning in rabbits. *Behavioral neuroscience*, 104(2):243.

Najafi, F. & Medina, J. F. (2013). Beyond "all-or-nothing" climbing fibers: graded representation of teaching signals in purkinje cells. *Frontiers in Neural Circuits*, 7:115.

Napier, R. M., Macrae, M., & Kehoe, E. J. (1992). Rapid reacquisition in conditioning of the rabbit's nictitating membrane response. *Journal of Experimental Psychology: Animal Behavior Processes*, 18(2):182.

Northcutt, R. G. (2002). Understanding vertebrate brain evolution. *Integrative and comparative biology*, 42(4):743–756.

Ojakangas, C. L. & Ebner, T. J. (1992). Purkinje cell complex and simple spike changes during a voluntary arm movement learning task in the monkey. *Journal of neurophysiology*, 68(6):2222–2236.

Paulin, M. G. (1993). The role of the cerebellum in motor control and perception. *Brain, Behavior and Evolution*, 41(1):39–50.

Pavlov, I. & Anrep, G. (1927). *Conditioned reflexes*. Dover Pubns. ISBN 0486430936.

Porrill, J. & Dean, P. (2008). Silent synapses, LTP, and the indirect parallel-fibre pathway: computational consequences of optimal cerebellar noise-processing. *PLoS computational biology*, 4(5).

Prueckl, R., Grunbacher, E., Ortner, R., Taub, A., Hogri, R., Magal, A., Segalis, E., Zreik, M., Nossenson, N., Herreros, I., et al. (2011a). The application of a real-time rapid-prototyping environment for the behavioral rehabilitation of a lost brain function in rats. Dins *Computational Intelligence, Cognitive Algorithms, Mind, and Brain (CCMB), 2011 IEEE Symposium on*, ps. 1–8. IEEE.

Prueckl, R., Taub, A., Herreros, I., Hogri, R., Magal, A., Bamford, S., Giovannucci, A., Almog, R., Shacham-Diamand, Y., Verschure, P., Mintz, M., Scharinger, J., Silmon, A., & Guger, C. (2011b). Behavioral rehabilitation of the eye closure reflex in senescent rats using a real-time biosignal acquisition system. Dins *Engineering in Medicine and Biology Society,EMBC, 2011 Annual International Conference of the IEEE*, ps. 4211 –4214. ISSN 1557-170X. doi: 10.1109/IEMBS.2011.6091045.

Rasmussen, A., Jirenhed, D.-A., Zucca, R., Johansson, F., Svensson, P., & Hesslow, G. (2013a). Number of spikes in climbing fibers determines the direction of cerebellar learning. *The Journal of Neuroscience*, 33(33):13436–13440.

Rasmussen, A., Zucca, R., Jirenhed, D., Johansson, F., Ortenblad, C., Svensson, P., & Hesslow, G. (2013b). Golgi cell activity during eyeblink conditioning in decerebrate ferrets. *The Cerebellum*, ps. 1–4.

Rescorla, R. & Wagner, A. (1972). A theory of pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical conditioning II: Current research and theory*, ps. 64–99.

Rilling, J. K. (2006). Human and nonhuman primate brains: Are they allometrically scaled versions of the same design? *Evolutionary Anthropology: Issues, News, and Reviews*, 15(2):65–77.

Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6):386.

Rossi, D. & Hamann, M. (1998). Spillover-Mediated Transmission at Inhibitory Synapses Promoted by High Affinity [alpha] 6 Subunit GABAA Receptors and Glomerular Geometry. *Neuron*, 20(4):783–795.

Rossi, D., Hamann, M., & Attwell, D. (2003). Multiple modes of GABAergic inhibition of rat cerebellar granule cells. *The Journal of Physiology*, 548(1):97.

Rowe, T. B., Macrini, T. E., & Luo, Z.-X. (2011). Fossil evidence on origin of the mammalian brain. *Science*, 332(6032):955–957.

Sabourin, C. & Bruneau, O. (2005). Robustness of the dynamic walk of a biped robot subjected to disturbing external forces by using cmac neural networks. *Robotics and Autonomous Systems*, 51(2):81–99.

Safo, P. & Regehr, W. (2008). Timing dependence of the induction of cerebellar LTD. *Neuropharmacology*, 54(1):213–218.

Sarkisov, D. & Wang, S. (2008). Order-dependent coincidence detection in cerebellar Purkinje neurons at the inositol trisphosphate receptor. *Journal of Neuroscience*, 28(1):133.

Schmahmann, J. D., Anderson, C. M., Newton, N., & Ellis, R. (2002). The function of the cerebellum in cognition, affect and consciousness:

Empirical support for the embodied mind. *Consciousness & emotion*, 2(2):273–309.

Schmahmann, J. D. & Sherman, J. C. (1998). The cerebellar cognitive affective syndrome. *Brain*, 121(4):561–579.

Schwartz, A. B., Cui, X. T., Weber, D. J., & Moran, D. W. (2006). Brain-controlled interfaces: movement restoration with neural prosthetics. *Neuron*, 52(1):205–220.

Shadlen, M. & Newsome, W. (1998). The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *The Journal of Neuroscience*, 18(10):3870. ISSN 0270-6474.

Shidara, M. (1993). Inverse-dynamics encoding of eye movements of purkinje cells in the cerebellum. *Nature*, 365:50–52.

Spanne, A. & Jörntell, H. (2013). Processing of multi-dimensional sensorimotor information in the spinal and cerebellar neuronal circuitry: A new hypothesis. *PLoS computational biology*, 9(3):e1002979.

Spencer, R., Zelaznik, H., Diedrichsen, J., & Ivry, R. (2003). Disrupted timing of discontinuous but not continuous movements by cerebellar lesions. *Science*, 300(5624):1437.

Steinmetz, J., Lavond, D., & Thompson, R. (1985). Classical conditioning of the rabbit eyelid response with mossy fiber stimulation as the conditioned stimulus. *Bulletin of the Psychonomic Society*, 23(3):245–248.

Sterelny, K. (2012). Language, gesture, skill: the co-evolutionary foundations of language. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1599):2141–2151.

Sugihara, I., Wu, H., & Shinoda, Y. (1999). Morphology of single olivocerebellar axons labeled with biotinylated dextran amine in the rat. *The Journal of comparative neurology*, 414(2):131–148.

Sutton, R. & Barto, A. (1998). *Reinforcement learning: An introduction*, volum 28. Cambridge Univ Press.

Svensson, P., Ivarsson, M., & Hesslow, G. (1997). Effect of varying the intensity and train frequency of forelimb and cerebellar mossy fiber conditioned stimuli on the latency of conditioned eye-blink responses in decerebrate ferrets. *Learning & Memory*, 4(1):105–115.

Svensson, P., Ivarsson, M., & Hesslow, G. (2000). Involvement of the cerebellum in a new temporal property of the conditioned eyeblink response. *Progress in Brain Research*, 124:317–323. ISSN 0079-6123.

Svensson, P., Jirenhed, D.-A., Bengtsson, F., & Hesslow, G. (2010). Effect of conditioned stimulus parameters on timing of conditioned purkinje cell responses. *Journal of neurophysiology*, 103(3):1329–1336.

Thach, W. (1967). Somatosensory receptive fields of single units in cat cerebellar cortex. *Journal of neurophysiology*, 30(4):675–696.

Thach, W. (2013). Does the cerebellum initiate movement? *The Cerebellum*, ps. 1–12.

Theil, H. (1972). *Statistical decomposition analysis: with applications in the social and administrative sciences*. North-Holland Pub. Co.

Thomas, A. (1911). *La fonction cérébelleuse*, volum 16. O. Doin et fils.

Thomas, A. (1912). *Cerebellar functions*. 12. Journal of Nervous and Mental Diseases Publishing Company.

Thompson, R. & Steinmetz, J. (2009). The role of the cerebellum in classical conditioning of discrete behavioral responses. *Neuroscience*, 162(3):732–755.

Uusisaari, M. & De Schutter, E. (2011). The mysterious microcircuitry of the cerebellar nuclei. *The Journal of physiology*, 589(14):3441–3457.

van Kan, P., Gibson, A., & Houk, J. (1993). Movement-related inputs to intermediate cerebellum of the monkey. *Journal of neurophysiology*, 69(1):74.

Verschure, P. & Mintz, M. (2001). A real-time model of the cerebellar circuitry underlying classical conditioning: A combined simulation and robotics study. *Neurocomputing*, 38:1019–1024.

Verschure, P. F. (2012). Distributed adaptive control: A theory of the mind, brain, body nexus. *Biologically Inspired Cognitive Architectures*.

Verschure, P. F., Voegtlin, T., & Douglas, R. J. (2003). Environmentally mediated synergy between perception and behaviour in mobile robots. *Nature*, 425(6958):620–624.

Walter, J. & Khodakhah, K. (2006). The linear computational algorithm of cerebellar Purkinje cells. *Journal of Neuroscience*, 26(50):12861.

Walter, J. & Khodakhah, K. (2009). The advantages of linear information processing for cerebellar computation. *Proceedings of the National Academy of Sciences*, 106(11):4471.

Wang, S., Denk, W., & Häusser, M. (2000). Coincidence detection in single dendritic spines mediated by calcium release. *nature neuroscience*, 3(12):1266–1273.

Weaver, A. H. (2005). Reciprocal evolution of the cerebellum and neocortex in fossil humans. *Proceedings of the National Academy of Sciences of the United States of America*, 102(10):3576–3580.

Wilson, B. S., Finley, C. C., Lawson, D. T., Wolford, R. D., Eddington, D. K., & Rabinowitz, W. M. (1991). Better speech recognition with cochlear implants. *Nature*, 352(6332):236–238.

Wolpert, D., Miall, R., & Kawato, M. (1998). Internal models in the cerebellum. *Trends in Cognitive Sciences*, 2(9):338–347. ISSN 1364-6613.

Yamamoto, K., Kawato, M., Kotosaka, S., & Kitazawa, S. (2007). Encoding of movement dynamics by purkinje cell simple spike activity during fast arm movements under resistive and assistive force fields. *Journal of neurophysiology*, 97(2):1588–1599.

Yamazaki, T. & Tanaka, S. (2007a). A spiking network model for passage-of-time representation in the cerebellum. *European Journal of Neuroscience*, 26(8):2279–2292.

Yamazaki, T. & Tanaka, S. (2007b). The cerebellum as a liquid state machine. *Neural Networks*, 20(3):290–297.

Yeo, C. & Hesslow, G. (1998). Cerebellum and conditioned reflexes. *Trends in Cognitive Sciences*, 2(9):322–330.

Zrenner, E. (2002). Will retinal implants restore vision? *Science*, 295(5557):1022–1025.