Model-Driven Analysis of Eyeblink Classical Conditioning Reveals the Underlying Structure of Cerebellar Plasticity and Neuronal Activity

Alberto Antonietti, Claudia Casellato, Egidio D'Angelo, and Alessandra Pedrocchi

***Abstract*— The cerebellum plays a critical role in sensorimotor control. However, how the specific circuits and plastic mechanisms of the cerebellum are engaged in closed-loop processing is still unclear. We developed an artificial sensorimotor control system embedding a detailed spiking cerebellar microcircuit with three bidirectional plasticity sites. This proved able to reproduce a cerebellar-driven associative paradigm, the eyeblink classical conditioning (EBCC), in which a precise time relationship between an unconditioned stimulus (US) and a conditioned stimulus (CS) is established. We challenged the spiking model to fit an experimental data set from human subjects. Two subsequent sessions of EBCC acquisition and extinction were recorded and transcranial magnetic stimulation (TMS) was applied on the cerebellum to alter circuit function and plasticity. Evolutionary algorithms were used to find the near-optimal model parameters to reproduce the behaviors of subjects in the different sessions of the protocol. The main finding is that the optimized cerebellar model was able to learn to anticipate (predict) conditioned responses with accurate timing and success rate, demonstrating fast acquisition, memory stabilization, rapid extinction, and faster reacquisition as in EBCC in humans. The firing of Purkinje cells (PCs) and deep cerebellar nuclei (DCN) changed during learning under the control of synaptic plasticity, which evolved at different rates, with a faster acquisition in the cerebellar cortex than in DCN synapses. Eventually, a reduced PC activity released DCN discharge just after the CS, precisely anticipating the US and causing the eyeblink. Moreover, a specific alteration in cortical plasticity explained the EBCC changes induced by cerebellar TMS in humans. In this paper, for the first time, it is shown how closed-loop simulations, using detailed cerebellar microcircuit models, can be successfully used to fit real experimental data sets. Thus, the changes of the model parameters in the different sessions of the protocol unveil how implicit microcircuit mechanisms can generate normal and altered associative behaviors.**

***Index Terms*— Cerebellum, distributed plasticity, eyeblink classical conditioning (EBCC), genetic algorithm (GA), long-term plasticity, spiking network model (SNN), transcranial magnetic stimulation (TMS).**

NOMENCLATURE

CFs	Climbing fibers.
CRs	Conditioned responses.
CS	Conditioned stimulus.
cTBS	continuous theta burst stimulation.
DCNs	Deep cerebellar nuclei cells.
EBCC	Eye blinking classical conditioning.
GA	Genetic algorithm.
GRs	Granular cells.
IOs	Inferior olive cells.
ISI	Interstimuli interval.
LTD	Long-term depression.
LTP	Long-term potentiation.
MFs	Mossy fibers.
PCs	Purkinje cells.
PFs	Parallel fibers.
PSTH	Peri-stimulus time histogram.
SNN	Spiking neural network.
US	Unconditioned stimulus.
TMS	Transcranial magnetic stimulation.

I. INTRODUCTION

SYNAPTIC plasticity modifies neurotransmission strength, and thus the probability that signals are transmitted through neural circuits. In this way, synaptic plasticity regulates information processing required to drive adaptive behaviors. The link between long-term synaptic plasticity and adaptive control has been suggested by a wealth of physiological and pathological data and by theoretical motor control models [1], [2]. However, how plasticity is engaged in dynamic processing during behavior is still unclear.

The cerebellum plays a critical role in adaptive motor control by implementing three fundamental operations: prediction, timing, and learning [3], [4]. These properties emerge in associative sensorimotor paradigms, such as the EBCC. This Pavlovian associative task is learned along with repeated presentations of paired stimuli, a CS (like a tone) followed by an US (like an air-puff or an electrical stimulation), eliciting the eyeblink reflex. The cerebellum learns to produce

Manuscript received January 9, 2016; revised April 13, 2016; accepted August 1, 2016. Date of publication September 1, 2016; date of current version October 16, 2017. This work was supported in part by the European Union: REALNET under Grant FP7-270434, and in part by the Human Brain Project, HBP-Regione Lombardia, under Grant HBP 604102.

A. Antonietti, C. Casellato, and A. Pedrocchi are with the Neuroengineering and Medical Robotics Laboratory, Department of Electronics, Information and Bioengineering, Politecnico di Milano, 20133 Milan, Italy (e-mail: alberto.antonietti@polimi.it).

E. D'Angelo is with the Brain Connectivity Center, Department of Brain and Behavioral Sciences, Istituto di Ricovero e Cura a Carattere Scientifico and the Istituto Neurologico Nazionale C. Mondino, University of Pavia, 27100 Pavia, Italy.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TNNLS.2016.2598190

a CR (an eyeblink) precisely timed to anticipate (or "predict") the US onset [5].

In a recent work [6], we have collected experimental data allowing to accurately determine the phases of EBCC in humans. Two subsequent sessions of EBCC acquisition and extinction were recorded, and TMS was applied on the cerebellum to alter circuit function and plasticity. These data suggest that TMS can dissociate EBCC extinction (related to the fast learning process) from consolidation (related to the slow learning process), probably by acting through a selective alteration of cerebellar plasticity. An extended multirate phenomenological model [7], [8] supported the multisite distribution of the learning process. However, the question on how the specific implementation of the cerebellar circuit was able to carry out the implicit computations eventually leading to EBCC learning [9]–[13] remained unresolved.

Since the formulation of the Marr's theory [10], it became clear that a gap still exists between the computation that the cerebellar network carries out and the implementation of the computational circuit, which requires that accurate microcircuit models are allowed to operate into the external large-scale circuitry of the brain [14]. In this paper, we have exploited a detailed computational model of the cerebellum operating in a sensorimotor circuit to match and interpret the EBCC experimental data. The model is a realistic SNN, equipped with distributed plasticity mechanisms, which is connected to an external circuit in order to generate the EBCC. SNN models [15]–[22] showed promising capabilities in reproducing behaviors similar to that of living brains due to their more faithful similarity to biological neural networks [23]. We used an event-driven SNN simulator based on look-up tables [24]–[26], which operates by compiling the dynamic responses of predefined cell models into lookup tables, thus significantly decreasing the simulation time with respect to other simulators, which has to solve complex systems of differential equations (e.g., NEURON [27] or Brian [28]).

In previous works [19], [29], we have demonstrated the learning properties and the capability of the cerebellar SNN to reproduce a general EBCC protocol. To build up the cerebellar model, we used SNNs instead of other modeling techniques, because this approach has proved to be very useful for neuroscience research, since it is capable to increase the understanding of the diverse features of the information processing that occurs in animals and human beings [30]. Our method exploited simulations of a neural architecture of acceptable size (thousands of spiking neurons), based on cells whose characteristics cannot be described by simple analytical expressions. It has been proved that an artificial NN whose circuitry is based on the synaptic organization of the cerebellum is capable to learn temporal associations [31].

In this contribution, we aimed at studying the way the general SNN cerebellar model defined in [19] has to be modified to adapt to real human data performing EBCC experiments under unaltered and perturbed conditions. In particular, we have developed a fitting procedure based on a cerebellar SNN endowed with multiple forms of plasticity and we have applied it to the real EBCC data set obtained by Monaco *et al.* [6], in which experiments were designed by the neuroscientists to study the interaction between the cerebellar associative learning and the TMS delivery.

While, on the one hand, the inclusion of realistic plasticity equations, spiking neural dynamics, and recurrent topologies enhanced the descriptive power of SNNs, on the other hand, this increased the number of free parameters requiring an efficient automated parameter tuning framework [32]. We exploited metaheuristic techniques, specifically evolutionary algorithms, to find out the near-optimal plasticity mechanism parameters identifying the cerebellar models that better reproduced the experimental results.

In summary, this approach is novel in several respects: 1) it implements and operates a detailed cerebellar microcircuit in closed-loop within a complete sensorimotor circuit; 2) it incorporates spiking (rather than analog) neurons and plasticity rules; 3) it uses plasticity at multiple sites rather than just at the PF-PC synapses of the cerebellar cortical layer; 4) it allows to simulate real data sets rather than formulating pure theoretical predictions; and 5) it allows to test hypotheses on circuit alterations (the TMS effect on plasticity) through modifications of specific neural mechanisms. Eventually, by reconnecting circuit implementation and computation to sensorimotor behavior, this approach provides a first direct test to the foundations of Marr's motor learning theory.

II. Materials and Methods

A. EBCC Protocol

The computational protocol used here was tailored on the experimental protocol by Monaco *et al.* [6], which is briefly explained for clarity. Human subjects underwent two sessions of EBCC, with a washout period interleaved [Fig. 1 (left)]. At the end of the first session, half of the subjects received a sham stimulation, while the other half received an effective cTBS. Therefore, the first session (session_1) included the data recorded from both groups; the second sessions ($\text{session}_{2\,\text{sham}}$ and $\text{session}_{2\,\text{tbs}}$) included the data recorded from the sham group and from the tbs group, respectively. Each session included 60 trials of acquisition and 10 trials of extinction. During the acquisition phase, the CS (a tone) was followed, after an ISI of 600 ms, by the US (a supraorbital nerve electric stimulation). During the extinction phase, the subjects were provided with the CS only. From these data, for each subject, the percent success rate (the number of CRs occurring within a moving window of ten trials) was computed. The experimental data showed two main features. First, the acquisition phase was faster in session_2 than in sessions_1, revealing a consolidation process during washout. Second, the extinction phases of $\text{session}_{2\,\text{tbs}}$ was smaller and slower than that of $\text{session}_{2\,\text{sham}}$.

The data used for the model fitting were the CR percentages (CR%), which corresponded to the success rates in the three different sessions (session_1, $\text{session}_{2\,\text{sham}}$, and $\text{session}_{2\,\text{tbs}}$).

B. Cerebellar Model and Protocol

The cerebellar model used for the computational simulations was based on a well-established cerebellar architecture [18], [33], which was built on physiological features of cerebellar microcomplex. The simulations were

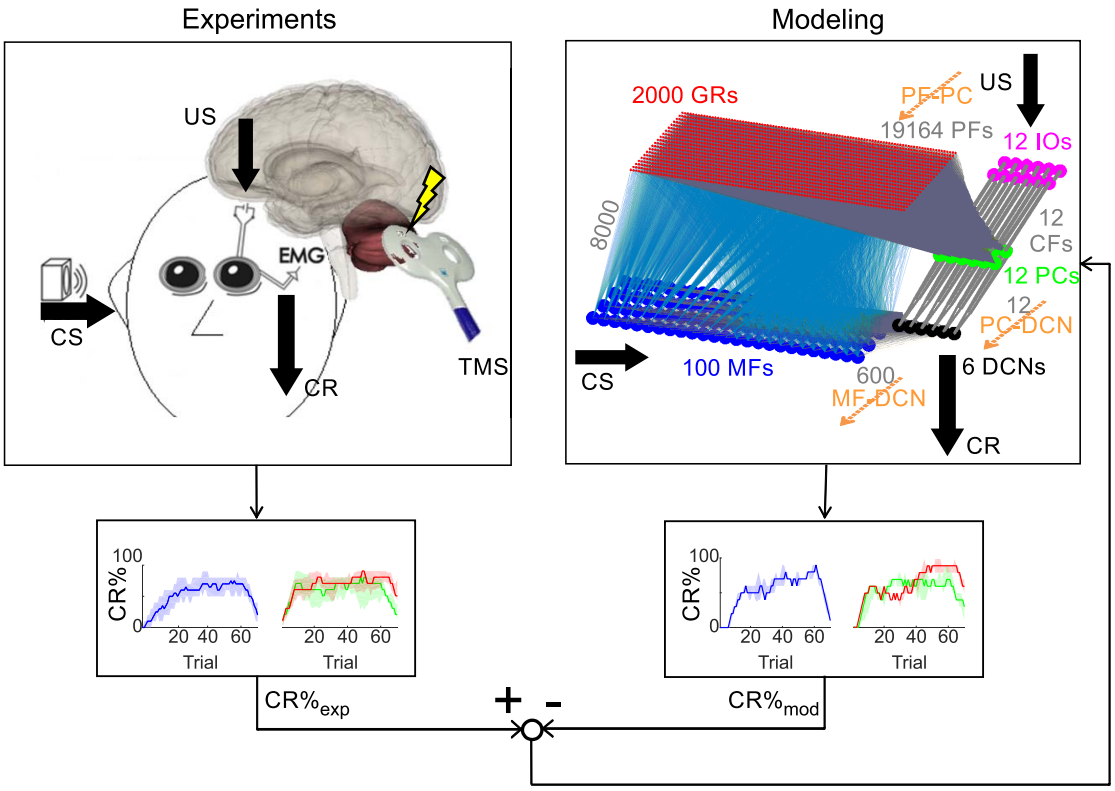

Fig. 1. Experimental and modeling approaches. Left: schematic of the EBCC experimental setup. The subject is stimulated with appropriate combinations of CS (tone) and US (electrical stimulation) organized in two sessions, each comprising an acquisition phase and an extinction phase; the CR (eyeblink) is detected by EMG on the orbicularis oculi muscle. Between the two sessions, one group of subjects receives cTBS on the posterior lobules of the lateral cerebellum [6]. Right: the model is endowed with three plasticity sites (PF-PC, MF-DCN, and PC-DCN), each bidirectional (LTP and LTD). Bottom: CR values obtained from the model are compared with those obtained from human subjects and used for optimal tuning of the model parameters through GAs. The optimal models are able to reproduce the real EBCC behavior.

performed on a desktop PC (Intel Core i7-2600 CPU @3.40 GHz with 8 GB of RAM with Windows 7 64 bit).

The SNN [Fig. 1 (right)] was composed of 100 MFs, 2000 GRs, 12 IOs, 12 PCs, and 6 DCNs. All the neurons were modeled as leaky integrate-and-fire neurons, because they required only a few state variables to be implemented [34]–[36]. The MFs received the CS and were randomly connected with the granular layer; each GR received four random excitatory synapses from the MFs. The GRs activity was a sparse representation of the input signal, so each simulation time sample corresponded to a different state of the granular layer [37]. The IOs received the US and were connected one by one to PCs through the CFs. Each PC received synapses from the 80% of the GRs, through 19 164 PFs. Each DCN received excitatory connections from all the MFs and two inhibitory connections from two PCs. Within our model, the DCN-IO inhibitory loop [38] did not correspond to a physical connection, but it was implemented as a mechanism that decreased the IOs firing rate of the incoming US, if a CR was detected before the US onset. In this way, such DCN-IO inhibitory loop translated the motor command into a sensory modulation, meaning that a single cerebellar area simultaneously tackled both motor execution and sensory prediction [39], [40]. Successful CRs were identified when the output variable, related to the DCN population firing rate, crossed a predefined threshold before the US onset.

The protocol tested in computational simulations reproduced the experimental one. Each session consisted in 60 trials of acquisition and 10 trials of extinction. During the acquisition phase, the CS (MFs stimulation with a firing rate of ~40 Hz lasting 700 ms) was followed, after an ISI of 600 ms, by the US (100-ms IOs stimulation with a low firing rate around 1 Hz). During the extinction phase, the CS only was fed to the cerebellar model.

1) Learning Rules: The SNN model was equipped with three plasticity sites, cortical and nuclear. The synaptic connections in each site followed three different learning rules, which strengthen or weaken these connections. LTD or LTP mechanisms were modeled as modifications on the synaptic conductances [19], [29].

The first learning rule (1) models the well-known LTP-LTD mechanism at the cerebellum cortical level (PF-PC) [41]

$$\Delta W_{\mathrm{PF}_i \to \mathrm{PC}_j}(t) = \begin{cases} \mathrm{LTD}_1 \int_{-\infty}^{t_{\mathrm{IOspike}_j}} K_1(t-x)\delta_{\mathrm{PF}_i}(t-x)dx & \text{if } \mathrm{PC}_j \text{ active} \\ & t = t_{\mathrm{IOspike}_j} \\ \mathrm{LTP}_1 & \text{if } \mathrm{PC}_j \text{ active} \\ & t \neq t_{\mathrm{IOspike}_j} \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

where

$$\delta_{\mathrm{PF}_i}(s) = \begin{cases} 1 & \text{if } \mathrm{PF}_i \text{ is active at time } s \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

and the Kernel function is

$$K_1(z) = Ae^{-\frac{z-t_0}{\tau}}\left(\sin\left(2\pi\frac{z-t_0}{\tau}\right)\right)^{20} \tag{3}$$

where LTD_1 and LTP_1 are the first learning rule constants, t_{IOspike_j} is the time when the corresponding CF_j emits a spike, $K_1(z)$ is the integral kernel function, which has its peak at t_0 (100 ms) before t_{IOspike_j} and τ and A are normalization constants. LTD_1 and LTP_1 values were defined by the optimization process made by an *ad hoc* GA as described in Section II-C. The rationale of the kernel function is presented in detail in [26].

The second learning rule (4) regards the MF-DCN nuclear connections [13]. It was preliminary tested in computational simulation of the EBCC protocol in order to investigate the effect of multiple plasticity sites on cerebellar learning [19]

$$\Delta W_{\mathrm{MF}_i\to\mathrm{DCN}_j}(t) = \begin{cases} \mathrm{LTD}_2\int_{-\infty}^{+\infty} K_2(t-x)\delta_{\mathrm{MF}_i}(t-x)dx & \text{if } \mathrm{MF}_i \text{ active} \\ & t = t_{\mathrm{PCspike}_j} \\ \mathrm{LTP}_2 & \text{if } \mathrm{MF}_i \text{ active} \\ & t \neq t_{\mathrm{PCspike}_j} \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

where

$$\delta_{\mathrm{MF}_i}(s) = \begin{cases} 1 & \text{if } \mathrm{MF}_i \text{ is active at time } s \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

and the Kernel function is

$$K_2(z) = e^{-\frac{|z|}{\tau}}\left(\cos\left(\frac{z}{\tau}\right)\right)^2 \tag{6}$$

where LTD_2 and LTP_2 are the second learning rule constants, t_{PCspike_j} is the time when the corresponding PC_j emits a spike, $K_2(z)$ is the integral kernel function, and τ is used in order to normalize the arguments in the learning rule. LTD_2 and LTP_2 values were defined by the optimization process made by the GA as described in Section II-C.

The third learning rule regards the PC-DCN nuclear connections and it was implemented as a standard spike-timing-dependent plasticity [19], [29]. Considering the ith DCN (DCN_i) and the two PCs connected with this DCN, the following holds.

1) When one of the two PCs fires and then, within an LTP-time window equal to 20 ms, also the DCN_i fires, the two inhibitory synapses from PCs to DCN_i are increased. The amount of conductance increase depends on the delay between the PC and DCN spikes, with a maximum LTP change equal to LTP_3.
2) When the DCN_i emits a spike and then, within an LTD-time window equal to 50 ms, also one of the two PCs fires, the two PC-DCN connections are decreased. The amount of conductance decrease depends on the delay between the DCN and PC spikes, with a maximum LTD change equal to LTD_3.

The two time windows were chosen from neurophysiological ranges [42], whereas LTP_3 and LTD_3 values were defined by the optimization process made by the GA as described in Section II-C.

C. Genetic Algorithm

As done in previous works [32], [43]–[46], evolutionary algorithms were used to tune SNN parameters. Tuning the free parameters of models has always been a challenge, especially in computational neuroscience, where the complexity of network models makes the hand tuning impracticable. Automated parameter search methods have become increasingly important and various methods were used, such as brute force search, local random search, gradient descent, evolutionary algorithm, and so on [47]. In this paper, we preferred to use an evolutionary method for the parameter tuning, because it allowed us to tune both the synaptic weight initialization and the learning rates of the synaptic rules, without the requirement to specify the desired computations of the network, but just to specify a proper fitness function. We optimized both initial conditions and learning rates at each plasticity site. The GA was used for the initialization of the weights of the model (before the beginning of $\mathrm{session}_1$), but not for the training of the SNN. Indeed, the evolution of the weights was driven by the three plasticity rules and by the interaction of the network with the input/output signals (CS, US, and CR) in a closed-loop fashion. The final target behavior is derived from experimental data and it has an intrinsic variability, both intrasubject and intersubject. Thus, an analytic error-based approach (e.g., gradient descend) would have failed in tracking the general behavior. We defined a fitness function, where the overall features representing experimental data can be put together. To optimize the parameters in order to achieve maximal fitness, an evolutionary approach was selected to explore the whole parameters space. Since this method is inherently parallel [48] (i.e., a single generation is formed by multiple individuals that can be simulated in parallel), we exploited this property to significantly decrease the computational time required during the optimization process (e.g., from 11 hours and 10 min using a single CPU core to 3 h and 30 min with a parallelization on four CPU cores). We used a GA to find the values of the three pairs of LTP and LTD constants. For $\mathrm{session}_1$ only, also the initialization weights (w_0) for PF-PC, MF-DCN, and PC-DCN synapses were found out by means of the GA. It was programmed in MATLAB, which automatically triggered each simulation in C++, carrying out a complete EBCC session driven by the model equipped with the updated genes. Each tested generation was made up of 12 individuals. A single individual represented a simulation performed with a parameter set composed of its own genes. We built individuals as described in Table I, with nine genes for $\mathrm{session}_1$ (LTP_1, LTD_1, LTP_2, LTD_2, LTP_3, LTD_3, $w_{0\mathrm{PF\text{-}PC}}$, $w_{0\mathrm{MF\text{-}DCN}}$, and $w_{0\mathrm{PC\text{-}DCN}}$) and six genes for $\mathrm{session}_{2\,\mathrm{sham}}$ and $\mathrm{session}_{2\,\mathrm{tbs}}$ (LTP_1, LTD_1, LTP_2, LTD_2, LTP_3, and LTD_3). Each gene could vary within a predefined range during the GA optimization.

TABLE I
GENES FOR THE DIFFERENT SESSIONS AND THEIR MAXIMUM AND MINIMUM VALUES

	session_1	session_2 sham and tbs	Max Value	Min Value
Gene 1	LTP_1	LTP_1	0.01	10^{-10}
Gene 2	LTD_1	LTD_1	-10^{-10}	-1
Gene 3	LTP_2	LTP_2	10^{-5}	10^{-10}
Gene 4	LTD_2	LTD_2	-10^{-10}	-10^{-5}
Gene 5	LTP_3	LTP_3	10^{-5}	10^{-10}
Gene 6	LTD_3	LTD_3	10^{-5}	10^{-10}
Gene 7	w_{0PF-PC}		2 nS	10^{-10} nS
Gene 8	$w_{0MF-DCN}$		0.1 nS	10^{-10} nS
Gene 9	$w_{0PC-DCN}$		0.2 nS	10^{-10} nS

For the LTP and LTD constants, we took as reference previous works based on similar architectures [18], [19], [25], [26], [33] and neurophysiological constraints (e.g., LTD_1 greater than LTP_1 and nuclear plasticities with LTP and LTD constants lower than the cortical ones). To establish bounds for the three initial weights (genes 7–9) we set, for each gene, the minimum value near zero (10^{-10} nS) and the maximum value in order to limit the firing rate ranges of the different cell populations within neurophysiological values [13].

In session_1, the cerebellar model started in a neutral configuration, where all the connections of the three synaptic sites were initialized at the same values ($w_{0\text{PF-PC}}$, $w_{0\text{MF-DCN}}$, and $w_{0\text{PC-DCN}}$) as they were naïve bundles. During the learning process along session_1, the synapses differentiated themselves. Thus, the weight state of the median model of session_1 at the end of the simulation was saved as the initial state of the synaptic weights for both $\text{session}_{2\,\text{sham}}$ and $\text{session}_{2\,\text{tbs}}$.

Since we aimed at extracting which parameter-sets generated models with a behavioral outcome (i.e., the success rate) as much similar as possible to the experimental data, we defined the GA fitness function as described in (7)

$$\begin{aligned}\text{fitness} &= \left(\left(1-\frac{\sum_{i=1}^{60}|\text{CR}\%_{\text{exp}}(i)-\text{CR}\%_{\text{mod}}(i)|}{60}\right)\cdot 0.4\right. \\ &\quad \left.+\left(1-\frac{\sum_{i=61}^{70}|\text{CR}\%_{\text{exp}}(i)-\text{CR}\%_{\text{mod}}(i)|}{10}\right)\cdot 0.6\right) \\ &\quad \cdot\left(1-\frac{\sum_{i=1}^{70}\text{OUT}(i)}{70}\right)\cdot\left(1-\frac{\sum_{i=1}^{10}\text{OUT}(i)}{10}\right) \qquad (7)\end{aligned}$$

where $\text{CR}\%_{\text{exp}}(i)$ is the CRs percentage of the median of experimental data at the ith trial and $\text{CR}\%_{\text{mod}}(i)$ is the CRs percentage of the model at the ith trial. The extinction was more weighted (0.6) than the acquisition (0.4), since it was the most critical phase, as emerged from the experimental data analyses [6]. The last two factors, containing $\text{OUT}(i)$, represent penalty parameters, which decreased the fitness values of that model if the CR percentage was outside the allowed ranges defined by the lower and upper quartiles of the experimental data, in particular, in the first ten trials where the intersubject variability was lower. $\text{OUT}(i)$ was a Boolean variable equal to 0 if the $\text{CR}\%_{\text{mod}}(i)$ was between the quartiles of $\text{CR}\%_{\text{exp}}(i)$; otherwise, it was equal to 1.

The fitness function was designed in order to be equal to 1 when the $\text{CR}\%_{\text{mod}}(i)$ coincided with the median of experimental data; then, it decreased if the difference between the experimental success rate and the model success rate increased (Fig. 1).

The GA process for the definition of the 12 individuals of the following generation consisted in three parts: selection, crossover, and mutation. The four best individuals with the highest fitness among all the individuals of the generation were saved without any crossover or mutation; the other eight individuals were generated by means of the following steps. In the selection process (roulette wheel), the 12 individuals of the current generation were sorted in the descending order of their fitness, the probability of being selected as one of the parents of the following generation was proportional to the fitness [49]. At the end of the selection, eight individuals were chosen as parents. The probability of a crossover between ordered couples of two parents was 80%; if the crossover happened, four randomly selected genes were swapped between the two parents. After the crossover process, each individual had a probability of 90% to go through a mutation. Individuals 5–8 underwent mutation by an uniform random reextraction from the genes exploration space, whereas, individuals 9–12 by a Gaussian mutation starting from their current values (i.e., the mutated gene was extracted using a Gaussian distribution with the mean equal to the current gene value and the standard deviation equal to the 10% of the variation range of the gene). After these three processes, the final 12 individuals of the following generation were defined and the new 12 EBCC simulations could start. The GA stopped when, for 100 consecutive generations, the fitness improvement between two generations was lower than 0.1%.

D. Data Analysis

1) Experimental Versus Model Outcomes: For the whole analysis, we tested the data normality with the Anderson–Darling test, in order to choose the proper statistical test. For variables that were nonnormally distributed, we indicated median [25th percentile 75th percentile].

We evaluated the results of the GA optimization considering, for each session, the best (i.e., with the highest fitness values) 25% models of the total number of individuals. We did not limit the analysis only to the single best individual in order to guarantee the robustness of our results. Indeed, we considered multiple good solutions that well-fitted the experimental data, so as to have a deeper insight into the parameters space. In this way, we obtained for each session a group of models. As expected by its selection criteria, the GA led to a monotonic increase of the maximum fitness values for each generation [Fig. 2(A)]; the stopping criterion was satisfied after 139 generations for session_1, after 196 generations for $\text{session}_{2\,\text{sham}}$, and after 251 generations for $\text{session}_{2\,\text{tbs}}$. The total number of tested individuals was 1668 for session_1, 2352 for $\text{session}_{2\,\text{sham}}$, and 3012 for $\text{session}_{2\,\text{tbs}}$, for an overall number of 7032 tested combinations of genes. The high number of tested individuals and the stability of the fitness function supported the hypothesis that the parameters found by the GA reliably correspond to those

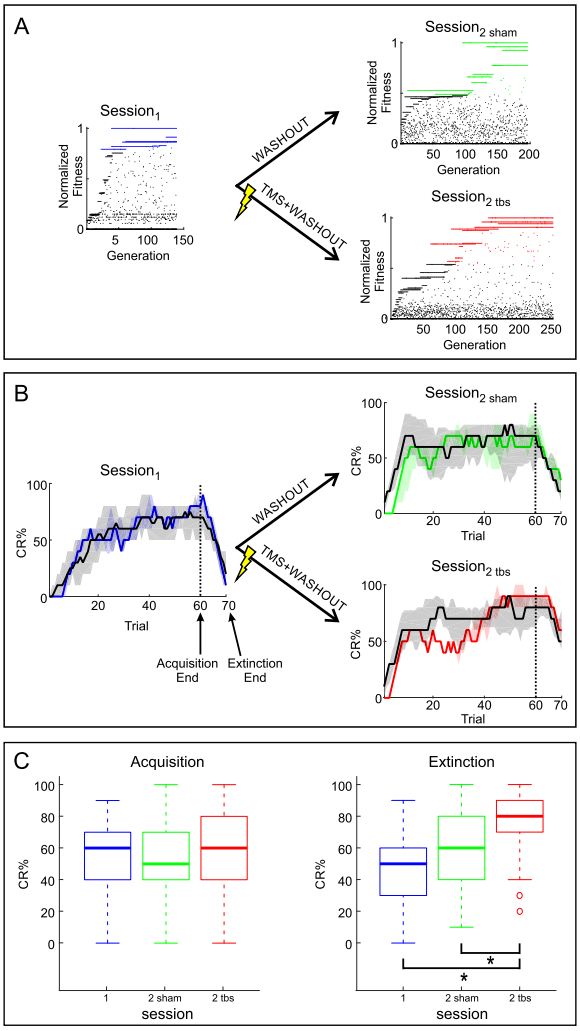

Fig. 2. Model fittings on experimental data. (A) Normalized fitness values along generations of GA for session$_1$, session$_{2\,\text{sham}}$, and session$_{2\,\text{tbs}}$. The best models (i.e., the best 25% generated by the GA) are depicted in color. (B) CR% in session$_1$, session$_{2\,\text{sham}}$, and session$_{2\,\text{tbs}}$ along the 60 acquisition and 10 extinction trials. For each panel, the median of experimental data among subjects is reported in black (the gray area represents the quartile range). The median of model fittings is reported in color (the surrounding colored areas represent the quartile range). (C) CR% averaged along the acquisition and extinction trials is reported for the session$_1$, session$_{2\,\text{sham}}$, and session$_{2\,\text{tbs}}$. Both for acquisition and extinction, a Mixed-Effect ANOVA was applied, to consider the session factor (three groups) and the trial sequence factor (1st–60th or 61st–70th, in chronological order). * indicates statistical difference ($p < 0.05$).

that would generate a behavioral outcome mostly similar to the experimental data. The computational time that was required to simulate the overall number of individuals was 3 h and 30 min for session$_1$, 4 h and 25 min for session$_{2\,\text{sham}}$, and 4 h and 15 min for session$_{2\,\text{tbs}}$. Taking into account the best 25% individuals, we selected 417 individuals with a normalized fitness equal to 0.86 [0.82 0.91] for session$_1$, 588 individuals with a normalized fitness equal to 0.77 [0.53 0.96] for session$_{2\,\text{sham}}$, and 753 individuals with a normalized fitness equal to 0.90 [0.88 0.96] for session$_{2\,\text{tbs}}$. The mean computational time to simulate a single individual (56 s of simulation time) was 48.1 ± 4.6 s for session$_1$, 43.4 ± 2.0 s for session$_{2\,\text{sham}}$, and 42.9 ± 1.9 s for session$_{2\,\text{tbs}}$. Since the network size is the same for each session, variations on the computational time for different individuals were due to different amounts of the spiking activity generated by the SNN.

In order to evaluate the fitting goodness, we computed the Pearson correlation coefficient between the median of experimental CR% and the median of model CR% along sessions.

Then, we focused separately on the acquisition and extinction phases carried out by the selected models in each session, comparing the CR% of the three groups of models (i.e., session$_1$, session$_{2\,\text{sham}}$, and session$_{2\,\text{tbs}}$). We exploited a linear mixed-effect ANOVA test, where the fixed effect factor was the group, and the random effect factor was the trial number. Post-hoc comparison used Bonferroni correction and Tukey contrasts, in order to highlight significant pairwise differences ($p < 0.05$) among the three sessions. To validate the model considering another variable, different from the CR% that was used for the optimization training, we compared the presence/absence of CRs trial by trial of the selected models in the three sessions and of the three median experimental individuals. For each of the three sessions, we computed the sensitivity (8) and the specificity (9) of the model on the experimental behavior (e.g., model 1 of session$_1$ versus median of experimental subjects 1–22 of session$_1$, model 2 of session$_1$ versus median of experimental subjects 1–22 of session$_1$, and so on).

$$\text{sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \cdot 100 \tag{8}$$

where TP (true positive) is the number of trials where both the model and the experimental data showed a CR, and FN (false negative) is the number of trials where the model predicted an absence of CR, whereas the experimental data showed a CR.

$$\text{specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \cdot 100 \tag{9}$$

where TN (true negative) is the number of trials where both the model and the experimental data showed an absence of CR, and FP (false positive) is the number of trials where the model predicted the presence of a CR, whereas the experimental data showed an absence of CR.

Due to the high intersubject variability of the experimental data set, the sensitivity and specificity values of the model had to be compared with the sensitivity and specificity values of each subject in the three sessions against the three median behaviors (e.g., subject 1 of session$_1$ versus median of subjects 1–22 of session$_1$, subject 2 of session$_1$ versus median of subjects 1–22 of session$_1$, and so on).

2) Session-Specific Models Parameters: We analyzed the resulting parameters (LTP$_1$, LTD$_1$, LTP$_2$, LTD$_2$, LTP$_3$, and LTD$_3$ for all the three sessions and $w_{0\text{PF-PC}}$, $w_{0\text{MF-DCN}}$, and $w_{0\text{PC-DCN}}$ for session$_1$ only) and their distribution within the exploration space.

For both the groups (session$_{2\,\text{sham}}$ and session$_{2\,\text{tbs}}$), we quantified the difference (Δ) of each plasticity parameter between sessions$_2$ values and the median value of session$_1$ [e.g., LTP$_1$(session$_{2\,\text{sham}}$) $-$ LTP$_1$(session$_1$)] as median and quartiles among the best selected models were normalized as the percentages of the range of motion of each parameter. We used Wilcoxon rank-sum test (MannWhitney test) to check if there were significant differences ($p < 0.05$) between Δ in session$_{2\,\text{sham}}$ and in session$_{2\,\text{tbs}}$ for the three plasticity sites.

Furthermore, for each parameter, we quantified the coefficient of variation among models (i.e., the interquartile range normalized for the parameter range). Indeed, it is relevant to focus on the dispersion of the optimal parameters values, to evaluate the robustness of the modifications, and thus to infer the role of each parameter in driving the model to reproduce the experimental behavior. For each of the three pairs of plasticity parameters, we performed a k-mean clustering with three clusters. In this way, we evaluated any systematic separation related to the three sessions by computing the number of misclassified elements and by computing the distances between the centroid of session_1-related cluster and the centroids of $\text{session}_{2\,\text{sham}}$ and $\text{session}_{2\,\text{tbs}}$ clusters.

3) Models Evolution (Synaptic Weights and Spiking Activities): We analyzed the synaptic weight modification along trials, reporting two significant trials for each session: the end of the acquisition phase (60th trial) and the end of the extinction phase (70th trial). The 19 164 PF-PC connections, the 600 MF-DCN connections, and the 12 PC-DCN connections were considered.

Synaptic modifications provoke changes in neural activity; thus, we computed the median firing rates of each cell population in order to compare them between different phases and between different sessions and to verify that they did not exceed the neurophysiological values [13].

We analyzed the spiking activity of the different populations of cells, by generating the PSTH of the spikes gave to the SNN as input (i.e., MFs and IOs) and generated by the SNN dynamics (i.e., GRs, PCs, and DCNs). For each model and for each cell population, we computed the PSTH with a time-bin of 10 ms. We considered the onset of CS as the starting point of each PSTH. Regarding PCs and DCNs, we analyzed the end of acquisition (60th trial) and the end of extinction (70th trial), since the strong modulation of their activity along sessions' phases.

PSTH considered the intratrial trends, albeit only at two specific simulation points. Thus, we also inspected the evolution of spiking activity of PCs and DCNs along each of the 70 trials of each session and along the intratrial time, computing the number of spikes of the cell population in each time-bin of 10 ms.

III. Results

In this paper, we developed a detailed computational model of the cerebellum operating inside a sensorimotor control system and capable of reproducing EBCC experimental data. In order to do so, a realistic spiking cerebellar network endowed with distributed plasticity mechanisms [19], [29] was connected to an external circuit accounting for the critical neural centers involved in EBCC. Evolutionary algorithms allowed to find out a family of near-optimal plasticity parameters determining the best models able to reproduce the experimental EBCC data.

A. Model Fitting to Experimental Results

GA simulations were run to fit the EBCC experimental data reported on human subjects by Monaco *et al.* [6] (Fig. 1). The simulations achieved a good fitting in terms of CR success rate, for all the three original EBCC sessions (session_1, $\text{session}_{2\,\text{sham}}$, and $\text{session}_{2\,\text{tbs}}$), as shown in Fig. 2(A). We assured the robustness of results evaluating the best 25% models for each session.

In session_1, the median of CR percent success rate of GA individuals [Fig. 2(B)] started from zero, and then, after six trials, some CRs occurred and the CR percentage increased progressively attaining a stable level around 60%–70%. The CR percentage showed a fast decrease toward zero within a few trials of extinction (starting at the 61st trial). During all the 70 trials, the CR% expressed by the model was very similar to the experimental data. The only remarkable difference was that, in the very first trials the human subjects were already able to produce a few temporal association between the CS and the US, while the simulations started to generate CRs after about six trials (see [50] for a potential explanation and remedy to the problem). The Pearson correlation coefficient computed between the median of models and the median of experimental data was 0.94, confirming the goodness of fit.

In $\text{session}_{2\,\text{sham}}$ and $\text{session}_{2\,\text{tbs}}$, the median CR percentage of individuals [Fig. 2(B)] started from zero, and then, it rapidly increased after three to four trials, reaching 60%–70% and remaining stable until the beginning of the extinction phase (60th trial). A peculiarity of $\text{session}_{2\,\text{sham}}$ and $\text{session}_{2\,\text{tbs}}$ was that reacquisition was more rapid than acquisition in session_1. Then, while extinction in $\text{session}_{2\,\text{sham}}$ showed a fast decrease toward zero, extinction in $\text{session}_{2\,\text{tbs}}$ showed a slower decrease reaching just 50%–60% at the 70th trial. In both $\text{session}_{2\,\text{sham}}$ and $\text{session}_{2\,\text{tbs}}$, the shape of the learning curves was very similar to the experimental curves. The Pearson correlation coefficients computed between the median of GA models and the median of experimental data were 0.80 for $\text{session}_{2\,\text{sham}}$ and 0.79 for $\text{session}_{2\,\text{tbs}}$, confirming the goodness of fit in both cases.

In order to determine whether and where the fitting results differed among session_1, $\text{session}_{2\,\text{sham}}$, and $\text{session}_{2\,\text{tbs}}$, we performed a mixed-effect ANOVA test and post-hoc analysis. These showed that, concerning the whole acquisition phase [Fig. 2(C)], there were no significant differences between models ($p = 1.00$). For the extinction phase, $\text{session}_{2\,\text{tbs}}$ was significantly different from both session_1 ($p = 0.035$) and $\text{session}_{2\,\text{sham}}$ ($p = 1.19 \cdot 10^{-5}$), whereas there was no significant difference between session_1 and $\text{session}_{2\,\text{sham}}$ ($p = 0.804$). These tests confirmed that the parameter modifications of $\text{session}_{2\,\text{tbs}}$ individuals effectively reflected the change of experimental conditions.

The sensitivity and specificity values of models versus experimental median, and experimental subjects versus experimental median, are reported in Table II. It is evident that the sensitivity and specificity values of the selected models are comparable with the values of the experimental subjects against their medians.

B. Session-Specific Model Parameters

Since simulations showed a behavior comparable with that observed experimentally, we analyzed model parameters (Table III) in order to shed light on how combinations of LTP

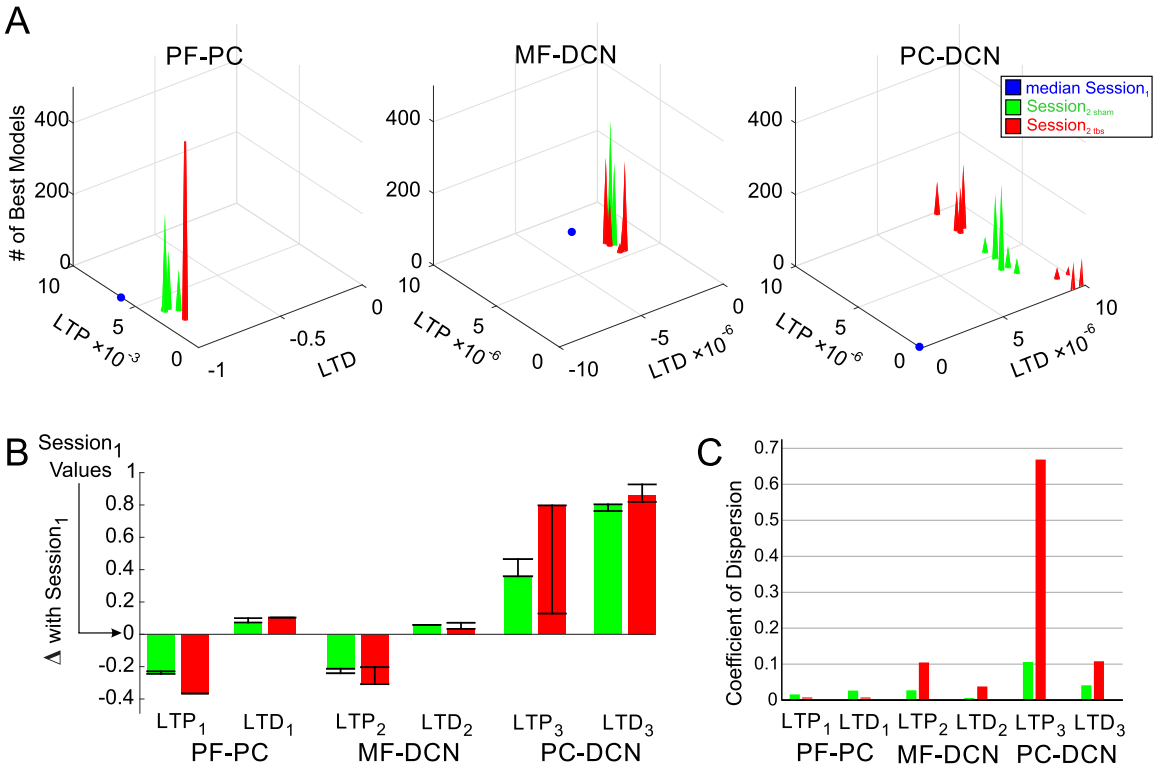

Fig. 3. (A) Plasticity parameters of the model LTP_1 and LTD_1 of the first plasticity site (PF-PC); LTP_2 and LTD_2 of the second plasticity site (MF-DCN); LTP_3 and LTD_3 of the third plasticity site (PC-DCN). For each panel, the parameters of the median best model for $session_1$ (blue dot) and of the best model groups, for $session_{2\,sham}$ (green) and $session_{2\,tbs}$ (red area), are reported. (B) Normalized differences (Δ) of LTP and LTD values (median represented by the bar edge and 25th–75th percentiles represented by black bars) in $session_{2\,sham}$ (green bars) and $session_{2\,tbs}$ (red bars), with respect to the corresponding median values in $session_1$ (baseline common reference). (C) Variability normalized within each range of the optimal parameters found by the GA for $session_{2\,sham}$ (green bars) and $session_{2\,tbs}$ (red bars), for each of the six plasticity parameters.

TABLE II
SENSITIVITY AND SPECIFICITY TESTS

	Models vs experimental median		
	$session_1$	$session_{2\,sham}$	$session_{2\,tbs}$
sensitivity	65%	55%	71%
specificity	71%	49%	38%
	Experimental subjects vs experimental median		
	$session_1$	$session_{2\,sham}$	$session_{2\,tbs}$
sensitivity	63%	64%	73%
specificity	59%	54%	45%

and LTD accounted for learning in the different EBCC phases.

In the three EBCC sessions, the LTP and LTD constants occupied a well-defined subspaces [Fig. 3(A)]. In particular, synaptic plasticity parameters settled on different optimal values for $session_{2\,sham}$ and $session_{2\,tbs}$ [Fig. 3(B)]. The difference with respect to the common baseline, represented by the median of $session_1$, was in the same direction for all the parameters in both $session_{2\,sham}$ and $session_{2\,tbs}$ (e.g., for LTP_1, both values decreased), but the amplitude variation was usually bigger for $session_{2\,tbs}$ than $session_{2\,sham}$ (except for LTD_2).

It should also be noted that the dispersion of these learning optimal parameters was different for the different plasticities [Fig. 3(C)]. The PF-PC weight showed the lowest variability ($\leq$ 3%), the MF-DCN weights had moderate variability ($\leq$ 11%), whereas the PC-DCN weights had a high variability, with a maximum value of 67% for LTP_3 in $session_{2\,tbs}$. The Wilcoxon rank-sum test for each of the three plasticity sites unveiled that the variation of plasticity parameters was significantly higher when the cTBS was administered, for the connections involving PCs ($\Delta session_{2\,sham}$ versus $\Delta session_{2\,tbs}$: PF-PC, $p = 5.7039 \cdot 10^{-5}$; MF-DCN, $p = 0.2839$; and PC-DCN $p = 2.9454 \cdot 10^{-7}$). But it is worth noting that the three plasticity parameters show different scatters. That suggests a systematic role of cTBS in modifying the plasticity parameters at the cortical layer (PF-PC) much more than at the other synaptic connections.

For each plasticity site, we quantified how much k-means clusters coincided with the three model groups ($session_1$, $session_{2\,sham}$, and $session_{2\,tbs}$) generated by the GA. We found a misclassification rate of 18.14% for PF-PC plasticity, 23.61% for MF-DCN plasticity, and 31.97% for PC-DCN plasticity. The relatively modest misclassification implied a systematic dependence of plasticity changes on the session.

For PF-PC plasticity, the distance between the centroids of $session_1$ and $session_{2\,sham}$ clusters was 0.07 and between $session_1$ and $session_{2\,tbs}$ was 0.11. For MF-DCN plasticity, the distance between the centroids of $session_1$ and $session_{2\,sham}$ clusters was $2.31 \cdot 10^{-6}$ and between $session_1$ and $session_{2\,tbs}$ was $3.20 \cdot 10^{-6}$. For the third plasticity site (PC-DCN), the distance between the centroids of $session_1$ and $session_{2\,sham}$ clusters was $5.25 \cdot 10^{-6}$ and between $session_1$ and $session_{2\,tbs}$ was $8.21 \cdot 10^{-6}$. Thus, the distances between $session_1$ and $session_{2\,tbs}$ parameters were always greater than the distances between $session_1$ and $session_{2\,sham}$, which is consistent with

TABLE III
MEDIAN AND QUARTILE OF EACH OPTIMIZED PARAMETER IN THE THREE DIFFERENT SESSIONS

		$session_1$			$session_{2\,sham}$			$session_{2\,tbs}$		
PF-PC	LTP_1	0.0061	[0.0048	0.0079]	0.0037	[0.0037	0.0038]	0.0024	[0.0024	0.0024]
	LTD_1	-0.99	[-0.99	-0.99]	-0.92	[-0.92	-0.89]	-0.88	[-0.88	-0.88]
MF-DCN	LTP_2	$8.57{\cdot}10^{-6}$	[$8.57{\cdot}10^{-6}$	$8.78{\cdot}10^{-6}$]	$6.44{\cdot}10^{-6}$	[$6.17{\cdot}10^{-6}$	$6.44{\cdot}10^{-6}$]	$5.56{\cdot}10^{-6}$	[$5.49{\cdot}10^{-6}$	$6.54{\cdot}10^{-6}$]
	LTD_2	$-2.49{\cdot}10^{-6}$	[$-2.59{\cdot}10^{-6}$	$-1.96{\cdot}10^{-6}$]	$-1.91{\cdot}10^{-6}$	[$-1.91{\cdot}10^{-6}$	$-1.91{\cdot}10^{-6}$]	$-2.15{\cdot}10^{-6}$	[$-2.15{\cdot}10^{-6}$	$-1.78{\cdot}10^{-6}$]
PC-DCN	LTP_3	10^{-10}	[10^{-10}	$3.81{\cdot}10^{-6}$]	$3.59{\cdot}10^{-6}$	[$3.59{\cdot}10^{-6}$	$4.65{\cdot}10^{-6}$]	$7.97{\cdot}10^{-6}$	[$1.28{\cdot}10^{-6}$	$7.97{\cdot}10^{-6}$]
	LTD_3	$5.18{\cdot}10^{-8}$	[10^{-10}	$3.09{\cdot}10^{-6}$]	$8.01{\cdot}10^{-6}$	[$7.67{\cdot}10^{-6}$	$8.09{\cdot}10^{-6}$]	$8.67{\cdot}10^{-6}$	[$8.24{\cdot}10^{-6}$	$9.32{\cdot}10^{-6}$]
w_{0PF-PC}		1.56	[1.56	1.56]						
$w_{0MF-DCN}$		0.078	[0.078	0.078]						
$w_{0PC-DCN}$		0.12	[0.093	0.14]						

a more similar outcome behaviors generated by $session_{2\,sham}$ and $session_1$ than by $session_{2\,tbs}$ and $session_1$. These metrics support the higher changes induced when not only a washout goes by, but also when a cTBS perturbation interferes.

C. Synaptic Plasticity Changes

Different weight changes characterized cortical plasticity (PF-PC) with respect to the two nuclear plasticities (MF-DCN and PC-DCN): PF-PC plasticity underwent faster changes along session trials with respect to MF-DCN and PC-DCN plasticities.

At the beginning of $session_1$, all the connections were equal to the initialization values found by the GA (genes 7–9): 1.57 nS for PF-PC, 0.078 nS for MF-DCN, and 0.094 nS for PC-DCN. At the beginning of session2, the connection weights for sham and tbs were set to the synaptic state at the 70th trial of $session_1$.

By comparing $session_1$ with $sessions_2$, differences occurred at the different synapses. At PF-PC synapses, there were lower weights in $session_2$ than $session_1$, both at the end of acquisition and extinction. At nuclear synapses, there were more dispersed values for MF-DCN (with numerous connections that were decreased toward zero) and PC-DCN weights in $sessions_2$ than $session_1$ (Fig. 4).

Comparing the weight configuration at the end of acquisition, PF-PC synapses in $session_{2\,tbs}$ were more depressed with respect to $session_{2\,sham}$. The synaptic weights for the nuclear plasticities were just slightly higher for MF-DCN, and lower for PC-DCN, in the tbs session than in the sham session. Thus, DCN synapses were in a potentiated state compared with $session_1$, facilitating task recalling (savings).

Regarding the differences between $session_{2\,sham}$ and $session_{2\,tbs}$, there were not observable behavioral alterations at the end of acquisition, but modifications emerged during the rapid extinction process. Thus, PF-PC plasticity differences were responsible for this differentiation, since the cortical plasticity reacts to fast changes, as happens during the extinction, whereas the nuclear plasticity sites affect long-term learning processes.

D. Neuronal Firing Rates

The spiking activity of MFs and IOs remained the same for each trial and for each session [Fig. 5(A)]. MFs had a random activity with an average firing rate of 39 [range 32–44] Hz. IOs were active only during the US (i.e., between 600 and 700 ms after the CS onset) with a frequency around 1 Hz. Finally, the GRs had a continuous activity directly related to the MFs input, with a firing rate of 10.20 [range 6.81–13.72] Hz. The PC activity was modulated within trials and along trials, with a minimum firing rate around 50 Hz within the trials when a CR was generated [Fig. 5(B)] and a maximum firing rate at the beginning of the acquisition and back at the end of the extinction around 100 Hz, completely inhibiting the circuit output [Fig. 5(C)]. Consequently, DCN activity was modulated within trials and along trials, with frequency peaks just before US onset to generate CRs. The firing rates of all these cells (Table IV) matched the known neurophysiological ranges [13], [51].

TABLE IV
MEDIAN AND QUARTILE OF PCs AND DCNs FIRING RATES [Hz] IN THE DIFFERENT SESSIONS/PHASES

	$session_1$		$session_{2\,sham}$		$session_{2\,tbs}$	
Acquisition PC	63	[45 88]	51	[28 77]	46	[26 71]
Acquisition DCN	5	[1 11]	16	[2 17]	16	[4 17]
Extinction PC	91	[70 114]	58	[40 81]	48	[27 74]
Extinction DCN	5	[2 7]	14	[1 17]	16	[7 17]

The PCs firing rate was higher and the DCNs firing rate was lower in $session_1$ with respect to $sessions_2$, during both acquisition and extinction phases. This effect was due to the higher strength of PF-PC and PC-DCN connections in $session_1$ than $sessions_2$, yielding higher PC activity and stronger PC-DCN inhibition.

At the beginning of acquisition, the PCs were spontaneously active and supplied tonic inhibition to DCN (Fig. 6). Then, along trials, PC activity decreased within a specific time-window (400–600 ms), in which DCN activity correspondingly increased. Finally, during extinction, PC activity increased again, decreasing the DCN firing rate. Interestingly, the DCN activity peak responsible for CR generation tended to disappear. In particular, in $session_1$ extinction, PCs recovered toward the initial tonic activity nearly silencing DCN cells. In $session_{2\,sham}$ extinction, PC activity recovered more slowly but still fast enough to cancel the DCN activity peaks anticipating US. In $session_{2\,tbs}$ extinction, the PC activity recovery was negligible, so that the differences between PC firing rate at the end of acquisition and at the end of extinction were thin. This abnormal behavior did not lead, on average, to an

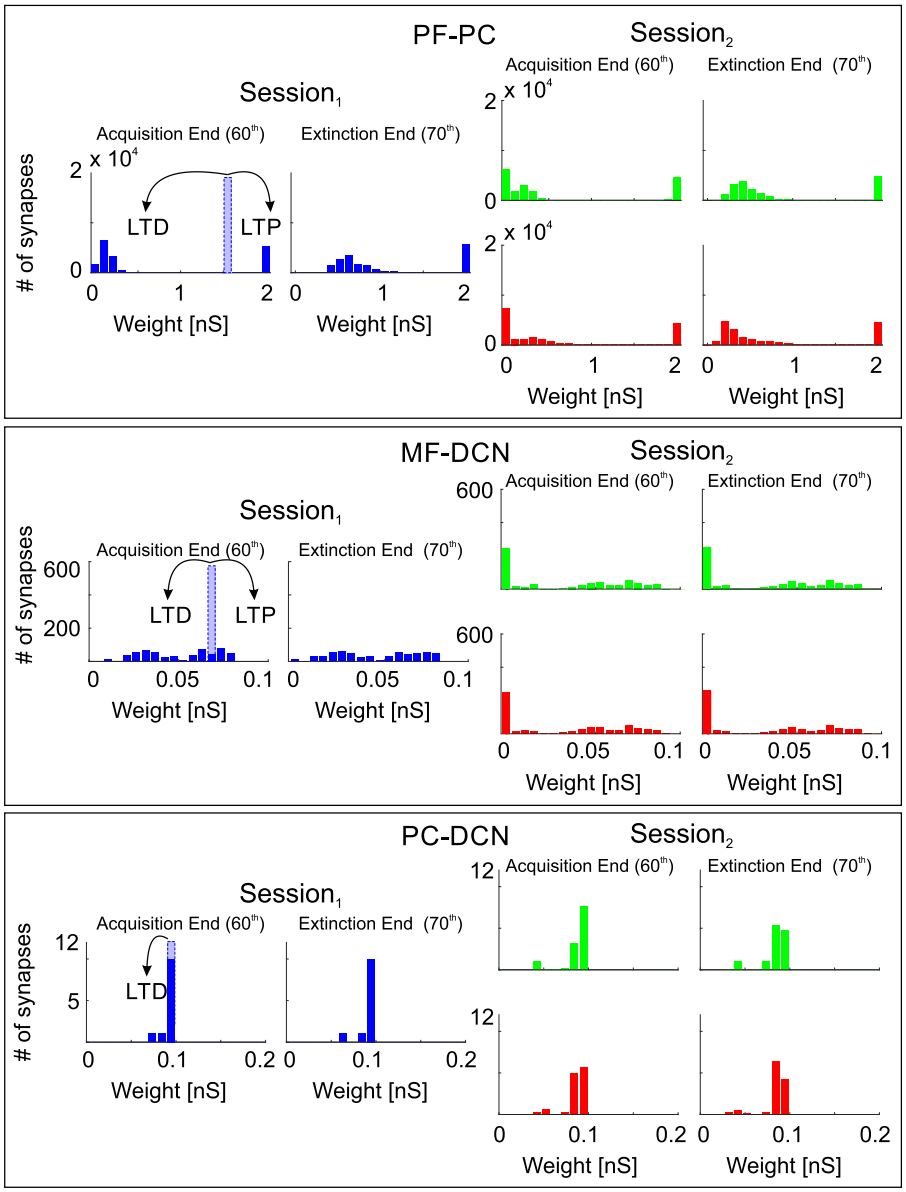

Fig. 4. Synaptic weights' distribution. For each of the three plasticity sites, the synaptic weights at the end of acquisition (60th trial) and at the end of extinction (70th trial) are reported in each session. Light blue blocks represent the initialization weights of session$_1$. The initialization weights found by GA were distant from their maximum and minimum allowed values, which is reasonable since the synaptic weights could vary in both directions (LTP and LTD) during the trial sequence. Top: PF-PC synaptic weights (19 164 connections). Middle: MF-DCN synaptic weights (600 connections). Bottom: PC-DCN synaptic weights (12 connections).

immediate reduction of the DCN peaks anticipating US (CRs were generated until the 66th trial) but made extinction slow and incomplete, as observed in the experiments.

It should be noted that the peak of DCN activity at ~650 ms from the beginning of a trial was not related to CR, since it happened after the US onset. This peak was a sort of unconditioned response triggered by the US itself and not by a predictive learned stimuli association.

IV. Discussion

The main result of this paper is that a 3-plasticity-site spiking cerebellar model embedded into a control system proved able to reproduce a biologically relevant associative task, the EBCC, which is largely accepted as a test-bench to investigate the sensorimotor learning capabilities of the cerebellum [52]. These were reliably emulated in terms of timing and success rate of the CR over a set of properties typical of human behavior: rapid acquisition, consolidation, extinction, and fast reacquisition following extinction [53]. The relevance of this finding is that, for the first time, a realistic modeling reconstruction of the cerebellum has been successfully used to fit an experimental data set, unveiling the implicit microcircuit computations of the network operating in closed loop. Other computational models using large-scale SNNs [13] were not matched against quantitative experimental data, so that the specific roles of each plasticity site in the multiple time-scale learning process could not be predicted. In this paper, the parameterization of network plasticity mechanisms was carried out by adapting the model response to human data through a metaheuristic process based on a GA. The emergence of multiple learning phases (incl uding acquisition, extinction, and reacquisition) was governed by the multiple learning sites

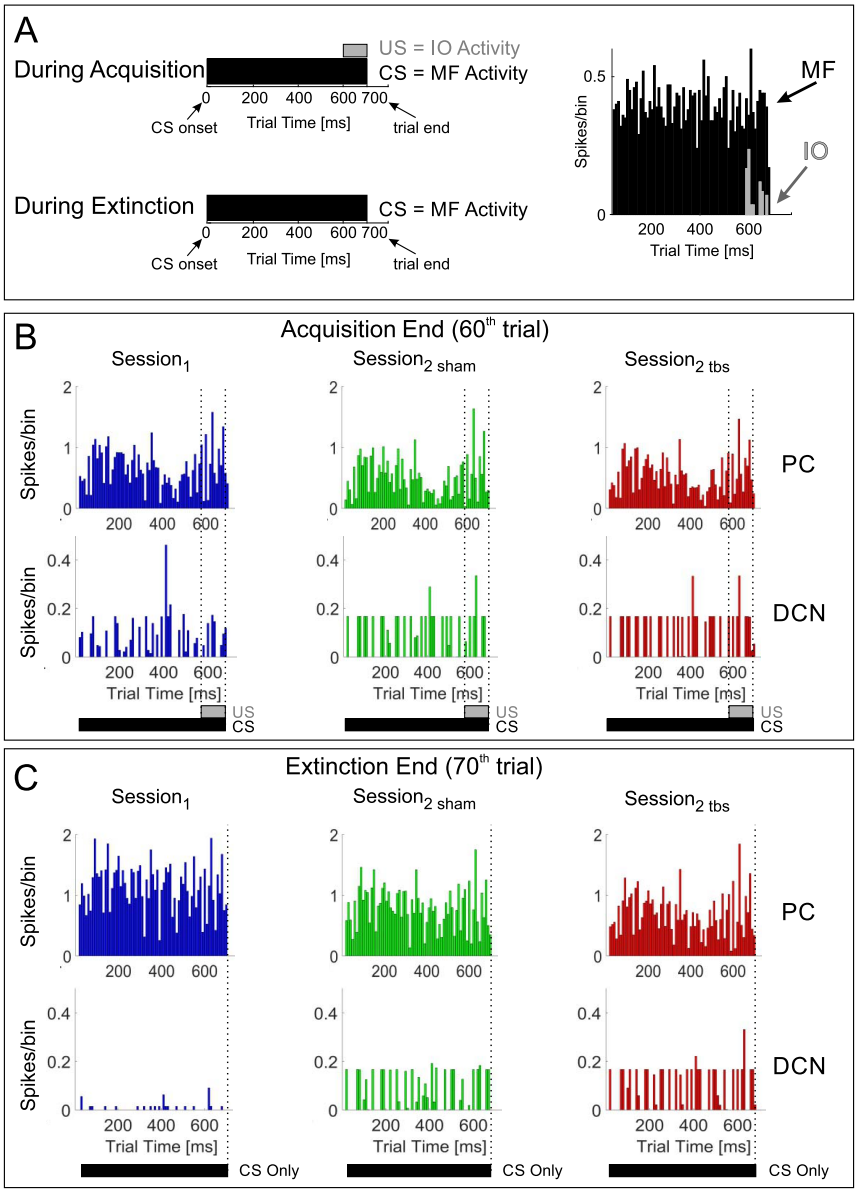

Fig. 5. Number of spikes generated within each time-bin (10 ms) is reported, along trial time (700 ms). The values are averaged across cell populations and across the models groups. (A) PSTH for cell populations of MFs and IOs. These populations activities were triggered by external stimuli, and thus, they are reported regardless of session. (B) PSTH for cell populations of PCs and DCNs during the last acquisition trial, in the three different sessions: session_1, $\text{session}_{2\,\text{sham}}$, and $\text{session}_{2\,\text{tbs}}$. (C) PSTH for cell populations of PCs and DCNs during the last extinction trial, in the three different sessions: session_1, $\text{session}_{2\,\text{sham}}$, and $\text{session}_{2\,\text{tbs}}$.

of the network [54] and finally caused a change in spike firing of the different neuronal types. The altered behavior induced in human subjects by TMS between the first and second sessions of training was reproduced by allowing the model to retune on the altered data set supporting the concept that TMS interfered especially with plasticity in the cerebellar cortical layer [6].

A. Dynamical Changes in Neuronal Firing and Synaptic Weights During Associative Learning

Neuronal activity in the cerebellar network showed a typical evolution during the learning process (session_1). Soon after a few trials, a strong inhibition of PC activity occurred just before the US. This in turn transiently released DCN neurons from inhibition increasing their firing and causing the behavioral response. The DCN activity increase was precisely timed and anticipated the US, as typically occurs in behavior. During extinction, PC and DCN firing recovered toward basal levels. Interestingly, during relearning (sessions_2), firing changes occurred much more rapidly due to the memory traces maintained in DCN.

It should be noted that, before training, no cues were given to the cerebellar network to evolve in the observed manner, except that it was equipped with structure, connectivity, and plasticity rules derived from biology. The evolution in neuronal firing was, thus, fully driven by long-term synaptic plasticity depending on the dynamical evolution of the inputs and of the local neuronal firing.

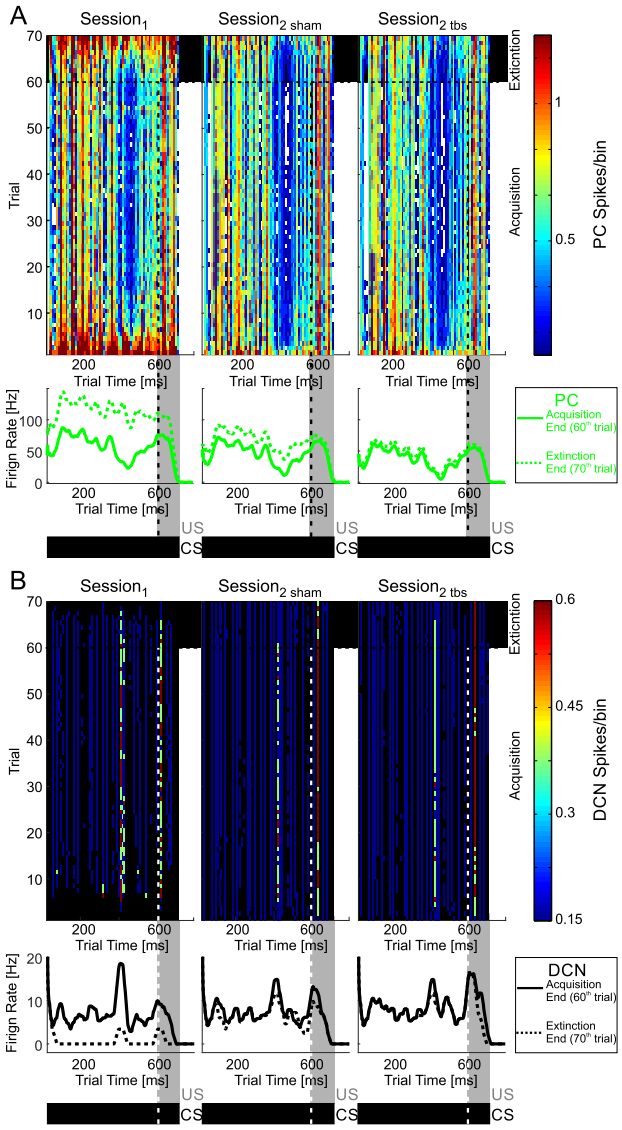

Fig. 6. Global spiking activity evolution. (A) PCs. (B) DCNs. The number of spikes generated within each time-bin (10 ms) is reported in color scale, along trial time (700 ms); the values are averaged across population cells and across the models groups, whereas the evolution along the session trials is depicted on the y-axis, for each session (session$_1$, session$_{2\ \mathrm{sham}}$, and session$_{2\ \mathrm{tbs}}$). Under each spike pattern, the instantaneous firing rate along trial time at the end of acquisition (60th trial) and at the end of extinction (70th trial) is reported.

The PF-PC synapses were the first to change, owing to its fast rate constants, followed by PC-DCN and MF-DCN synapses, which evolved at a much slower rate. This made learning biphasic, with a first rapid phase taking place in the cerebellar cortex and a slower phase taking place in the DCNs. These observations are consistent with the hypothesis that the cerebellar cortex is necessary to generate an adaptive well-timed CRs, but it is insufficient to determine the full set of learning properties by itself [51]. Indeed, multiple processes may contribute to motor skill acquisition, which usually proceeds through a rapid convergence toward a stable state before being consolidated into persistent memory [8]. As suggested by Medina *et al.* [55] through *ad hoc* lesions and computational simulations, a site of plasticity outside the cerebellar cortex (possibly in the cerebellar nuclei) can indeed protect from permanent EBCC extinction, so that residual plasticity can later contribute to savings seen during relearning. The cerebellar cortex, thus, operated as a fast learning module, while deeper structures operated as a slow learning module where the motor skill can be transferred and consolidated into more persistent memory [55]. Accordingly, the DCN activity can be modulated by PCs, and DCN spike timing is strongly correlated with memory acquisition [56].

B. Model Predictions on Neuronal Firing and Synaptic Plasticity in the Cerebellar Network

One of the major difficulties in the experimental analysis of neuronal firing during behaviors is that potential changes compared with baseline activity, even if significant, cannot be easily interpreted given to simultaneous changes in other neurons, in the synapses in between and in the afferent fiber activity. The model provided a series of testable predictions of neuronal properties during EBCC that can represent a guide to interpret experimental data.

As far as firing rates were concerned, PC and DCN cell firing rates were in the physiological ranges [13] and then changed in a characteristic manner along the acquisition and extinction phases. These changes proved consistent with those revealed in electrophysiological measures in decerebrate ferrets [51]. The characteristic PC and DCN firing rate changes during different phases of learning could be used to predict the ongoing changes in PC and DCN connections in *in vivo* experiments.

Finally, the synaptic weights change correlated with different behavioral phases. In relationship with the different kinetics of plasticity, the involvement of nuclear synapses was more evident during the long acquisition processes while that of cortical synapses during the fast extinction process [6]–[8]. Not unexpectedly, alteration in cortical synapses provided an explanation to TMS perturbations, which occur superficially affecting the faster learning process (see the following).

In sessions$_2$, DCN synapses were potentiated compared with session$_1$, facilitating task recalling (savings). Since reacquisition was almost the same for both the sham and tbs groups, the saving mechanisms were preserved regardless of TMS perturbation. Therefore, savings derived from a constructive interaction between cortical and nuclear synapses modulation. It was modeled as a process occurring during the washout between sessions, when the SNN parameters governing the plastic changes at the multiple sites underwent a remodulation, especially selective and controlled at level of PF-PC.

C. Model Predictions on TMS Data

It was previously shown that TMS stimulation applied on the cerebellum influenced its learning processes, but the underlying mechanisms were unclear. Miall and King [57], by applying TMS to human subjects during dynamic actions, measured an increase of the positional error, hypothesizing that the state estimation provided by the cerebellum was somehow compromised. However, they could not address the localization of such an effect within the multilayer cerebellar structure. The present model fittings to TMS data suggested

that TMS should mostly alter plasticity in the cerebellar cortex, i.e., in most superficial layers directly affected by TMS. Also other synapses tended to change (though nonsignificantly), reflecting redistribution of weights over the whole network.

Conceptually, this bears about important implications for understanding the place and nature of changes induced by TMS. First, fast and slow processes were updated simultaneously from motor learning errors, supporting a parallel architecture of motor memory [58]. Second, cerebellar TMS affected memories based on large-magnitude errors, i.e., it altered the fast process operations, in line with recent studies showing that cerebellar degeneration impaired the ability to learn from large-magnitude errors, but had a modest impact on learning from small errors [59]. The model, thus, proved a useful complement to TMS, which is commonly used to reversibly disrupt normal brain functions, thereby allowing to dissociate and to study the underlying plasticity mechanisms [60]–[62].

D. Properties and Limits of the Model

A major advantage of our model is that, while plasticity at the PF-PC synapses was originally the only one to be considered, we have embedded three reversible plasticity forms. These plasticities, based on recent observations at the cellular level, have different trigger signals and timescales, improving neurophysiological realism and expanding the computational and learning capabilities of the circuit. Actually, the differential engagement of these multiple plasticity sites allowed to better emulate the complex properties of learning than with a single plasticity alone [26].

A second advantage was that, rather than using the model to formulate a pure theoretical hypothesis, we have proved that the model can be tuned against a real data set. Clearly, the effectiveness of this approach depends on the richness of neuronal mechanisms and synaptic plasticities embedded into the model itself. In fact, our model successfully associated neuronal-scale to behavioral-scale features and was able to reveal potential mechanisms of alteration following a perturbation imposed to the system, i.e., cerebellar TMS. It could be envisaged that imposed modifications to neuronal functioning or plasticity mechanisms in the model would allow it to predict the consequences of cerebellar alterations in human pathologies, a promising aspect that deserves future investigation.

In spite of these advantages, the model has also limits. A first limit is due to the elementary representation of neurons and synapses. It will be interesting to see how the system will respond with more advanced neuron and plasticity mechanisms, as the center surrounds organization of the granular layer or coherent oscillations [63]. A second limit is in the plasticity rules, which could be more complex then represented here [64]. Actually, without embedding the whole model with intrinsic temporal dynamics (oscillations and resonance), we expect that adding further molecular complexity would not be very useful, so that an increasing biological realism should require to redesign the system as a whole. Finally, we did not include cerebro-cerebellar recurrent loops in the control system. These would be most useful in motor task involving closed-loop planning and execution, but are probably not very important in the present context, in which only part of the functionalities of the sensorimotor control system are exploited.

The choice of the fitness function to be used during the optimization process is another limitation of the GA approach. As a matter of fact, the fitness function (7) designed for the parameter set search could represent solely some features of the experimental behaviors. In a single value, it had to summarize the similarity between the experimental CRs profile and the models one. However, we designed an accurate fitness function, which considered all the important features of the desired (optimal) behavior, to diminish the information loss.

From a computational point of view, the proposed model was simulated exploiting an event-based SNN simulator that used, for now, the computational power of a single processor. In this way, it was possible to simulate different parameter sets in parallel on the four CPU cores, but this limited the simulation performances. To further reduce the simulation time, maintaining the network dimension, it could be possible to run a single simulation in parallel on multiple CPU cores or using a GPU-approach, which speeds up the simulation [15], [34].

Finally, the similarity of the proposed model with the cytological structure of the cerebellum represented a crucial point of this paper. As a side effect of the biological realism of the SNN, it was impossible to directly compare the performances of the model with other classical neural network topologies (e.g., single or multilayers perceptrons, fuzzy neural networks, and so on), which do not consider the main working principles of the biological network.

In an extensive theoretical work, Maass [65] demonstrated that SNNs are powerful from a computational point of view, even more powerful than conventional NNs, reducing the overall need of units. This consideration is based on the intrinsic time-dependent dynamics of spiking neurons, which allow capturing and exploiting more efficiently the temporal patterns of sensory-motor events.

In our scenario, the use of SNN has been preferred to other statistical tools for manifold reasons. First of all, SNNs are algorithms patterned after the brain structures and contain a series of mathematical equations that are used to simulate biological processes. Indeed, they represent a technique that has emerged as a potential alternative to logistic regression analysis and other statistical methods [66]. Furthermore, SNNs are not constrained by a predefined mathematical relationship between dependent and independent variables, and have the possibility to reproduce complex nonlinear relationships. In addition to this, while a generic NN is an implicit (black-box) approach, the SNN exploited by us is no more a black-box. In fact, each neuron and each synaptic connection have a precise meaning and a biological alias that we can understand by means of the neurophysiological knowledge. If the aim was to merely reproduce the exact output behavior showed by human subjects, other classical or advanced methods (e.g., logistic regression) would have obtained higher sensibility and sensitivity, possibly reaching values near 100%. However, it would be impossible to link the high-level parameters in the models to low-level features in the biological learning mechanisms. On the contrary, this is possible using

realistic, biologically inspired models, such as the SNN, that we have exploited.

E. From Microcircuit Implementation and Computation to Cerebellar Algorithms

The Marr's theory about brain functional principles envisaged that a circuit algorithm could be resolved on the basis of microcircuit computation and implementation [9], [10]. In this paper, we have implemented a detailed neuronal microcircuit generating implicit spiking computations able to produce associative sensorimotor behaviors, that is, we have reversed the original procedure: rather than anticipating an algorithm and looking for possible computations and implementations capable of generating it (inverse problem), we have followed a bottom-up approach yielding a behavioral response (an adaptive sensorimotor association) built on network constructive principles and plasticity rules. We have therefore moved a first step toward a direct demonstration of Marr's predictions on the cerebellar operating mechanisms in a human-like behavior. A further critical challenge will then be investigating the responses of an advanced cerebellar circuit model engaged into the feedback and feedforward loops representing an entire sensorimotor system operating in a closed loop.

V. Conclusion

As a remarkable advance with respect to the state of the art, the model approach to data interpolation can provide a new key to understand the physiological mechanisms of associative motor learning in the cerebellar circuit and to predict the potential changes in dysfunctional conditions. In silico manipulations of a realistic model-based cerebellar platform can be a key approach to understand cerebellar functioning and intervene on cerebellar diseases. Several cerebellar impairments can be found in the literature; with an approach similar to what we proposed in this paper, it would be possible to translate these physiological damages into models' modifications (e.g., a decrease of the volume of the cerebellar cortex, due to a cortical degeneration [67], could be translated by decreasing the number of PCs in the model).

Acknowledgment

The authors would like to thank Dr. G. Koch, Research Director of the Department of Clinical and Behavioral Neurology of Foundation Santa Lucia, Rome, for having provided the human subjects data.

Replication Data

The results showed in this paper can be replicated. We have published online the instructions and the data set needed. The data can be retrieved at Harvard Dataverse: doi:10.7910/DVN/XUYXKC.

References

[1] J. E. Steinmetz and D. R. Sengelaub, "Possible conditioned stimulus pathway for classical eyelid conditioning in rabbits. I. Anatomical evidence for direct projections from the pontine nuclei to the cerebellar interpositus nucleus," *Behavioral Neural Biol.*, vol. 57, no. 2, pp. 103–115, Mar. 1992.

[2] G. W. Crabtree and J. A. Gogos, "Synaptic plasticity, neural circuits, and the emerging role of altered short-term information processing in schizophrenia," *Frontiers Synaptic Neurosci.*, vol. 6, p. 28, Nov. 2014.

[3] E. D'Angelo and S. Casali, "Seeking a unified framework for cerebellar function and dysfunction: From circuit operations to cognition," *Frontiers Neural Circuits*, vol. 6, p. 116, Jan. 2012.

[4] E. D'Angelo *et al.*, "Realistic modeling of neurons and networks: Towards brain simulation," *Funct. Neurol.*, vol. 28, no. 3, pp. 153–166, 2013.

[5] J. F. Medina, W. L. Nores, T. Ohyama, and M. D. Mauk, "Mechanisms of cerebellar learning suggested by eyelid conditioning," *Current Opinion Neurobiol.*, vol. 10, no. 6, pp. 717–724, Dec. 2000.

[6] J. Monaco, C. Casellato, G. Koch, and E. D'Angelo, "Cerebellar theta burst stimulation dissociates memory components in eyeblink classical conditioning," *Eur. J. Neurosci.*, vol. 40, no. 9, pp. 3363–3370, Nov. 2014.

[7] M. A. Smith, A. Ghazizadeh, and R. Shadmehr, "Interacting adaptive processes with different timescales underlie short-term motor learning," *PLoS Biol.*, vol. 4, no. 6, p. e179, Jun. 2006.

[8] R. Shadmehr, M. A. Smith, and J. W. Krakauer, "Error correction, sensory prediction, and adaptation in motor control," *Annu. Rev. Neurosci.*, vol. 33, pp. 89–108, Jan. 2010.

[9] M. Ito, "Adaptive control of reflexes by the cerebellum," *Prog. Brain Res.*, vol. 44, pp. 435–444, Dec. 1976.

[10] D. Marr, "A theory of cerebellar cortex," *J. Physiol.*, vol. 202, no. 2, pp. 437–470, Jun. 1969.

[11] J. Porrill and P. Dean, "Cerebellar motor learning: When is cortical plasticity not enough?" *PLoS Comput. Biol.*, vol. 3, no. 10, p. e197, Oct. 2007.

[12] N. F. Lepora, J. Porrill, C. H. Yeo, and P. Dean, "Sensory prediction or motor control? Application of Marr–Albus type models of cerebellar function to classical conditioning," *Frontiers Comput. Neurosci.*, vol. 4, p. 140, Jan. 2010.

[13] J. F. Medina, K. S. Garcia, W. L. Nores, N. M. Taylor, and M. D. Mauk, "Timing mechanisms in the cerebellum: Testing predictions of a large-scale computer simulation," *J. Neurosci.*, vol. 20, no. 14, pp. 5516–5525, 2000.

[14] P. Dean, J. Porrill, C.-F. Ekerot, and H. Jörntell, "The cerebellar microcircuit as an adaptive filter: Experimental and computational evidence," *Nature Rev. Neurosci.*, vol. 11, no. 1, pp. 30–43, Jan. 2010.

[15] T. Yamazaki and J. Igarashi, "Realtime cerebellum: A large-scale spiking network model of the cerebellum that runs in realtime using a graphics processing unit," *Neural Netw.*, vol. 47, pp. 103–111, Nov. 2013.

[16] C. Hofstötter, M. Mintz, and P. F. M. J. Verschure, "The cerebellum in action: A simulation and robotics study," *Eur. J. Neurosci.*, vol. 16, no. 7, pp. 1361–1376, Oct. 2002.

[17] W.-K. Li, M. J. Hausknecht, P. Stone, and M. D. Mauk, "Using a million cell simulation of the cerebellum: Network scaling and task generality," *Neural Netw.*, vol. 47, pp. 95–102, Nov. 2013.

[18] C. Casellato *et al.*, "Adaptive robotic control driven by a versatile spiking cerebellar network," *PLoS ONE*, vol. 9, no. 11, p. e112265, Nov. 2014.

[19] A. Antonietti, C. Casellato, J. A. Garrido, E. D'Angelo, and A. Pedrocchi, "Spiking cerebellar model with multiple plasticity sites reproduces eye blinking classical conditioning," in *Proc. 7th Int. IEEE/EMBS Conf. Neural Eng. (NER)*, Apr. 2015, pp. 296–299.

[20] Q. Yu, H. Tang, K. C. Tan, and H. Li, "Rapid feedforward computation by temporal encoding and learning with spiking neurons," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 24, no. 10, pp. 1539–1552, Oct. 2013.

[21] E. Ros, E. M. Ortigosa, R. Agís, R. Carrillo, and M. Arnold, "Real-time computing platform for spiking neurons (RT-spike)," *IEEE Trans. Neural Netw.*, vol. 17, no. 4, pp. 1050–1063, Jul. 2006.

[22] N. R. Luque, J. A. Garrido, R. R. Carrillo, O. J.-M. D. Coenen, and E. Ros, "Cerebellar input configuration toward object model abstraction in manipulation tasks," *IEEE Trans. Neural Netw.*, vol. 22, no. 8, pp. 1321–1328, Aug. 2011.

[23] J. Ranhel, "Neural assembly computing," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 23, no. 6, pp. 916–927, Jun. 2012.

[24] E. Ros, R. Carrillo, E. M. Ortigosa, B. Barbour, and R. Agís, "Event-driven simulation scheme for spiking neural networks using lookup tables to characterize neuronal dynamics," *Neural Comput.*, vol. 18, no. 12, pp. 2959–2993, Dec. 2006.

[25] R. R. Carrillo, E. Ros, B. Barbour, C. Boucheny, and O. Coenen, "Event-driven simulation of neural population synchronization facilitated by electrical coupling," *Biosystems*, vol. 87, nos. 2–3, pp. 275–280, Feb. 2007.

[26] R. R. Carrillo, E. Ros, C. Boucheny, and O. J.-M. D. Coenen, "A real-time spiking cerebellum model for learning robot control," *Biosystems*, vol. 94, nos. 1–2, pp. 18–27, 2008.

[27] M. Hines and N. Carnevale, "The NEURON simulation environment," *Neural Comput.*, vol. 9, no. 6, pp. 1179–1209, Aug. 1997.

[28] D. Goodman and R. Brette, "Brian: A simulator for spiking neural networks in python," *Frontiers Neuroinform.*, vol. 2, no. 11, p. 5, 2008.

[29] A. Antonietti *et al.*, "Spiking neural network with distributed plasticity reproduces cerebellar learning in eye blink conditioning paradigms," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 1, pp. 210–219, Jan. 2016.

[30] S. Ghosh-Dastidar, "Spiking neural networks," *Int. J. Neural Syst.*, vol. 19, no. 4, pp. 295–308, Aug. 2009.

[31] D. V. Buonomano and M. D. Mauk, "Neural network model of the cerebellum: Temporal discrimination and the timing of motor responses," *Neural Comput.*, vol. 6, no. 1, pp. 38–55, Jan. 1994.

[32] K. D. Carlson, J. M. Nageswaran, N. Dutt, and J. L. Krichmar, "An efficient automated parameter tuning framework for spiking neural networks," *Frontiers Neurosci.*, vol. 8, no. 2, p. 10, Feb. 2014.

[33] C. Casellato, A. Antonietti, J. A. Garrido, G. Ferrigno, E. D'Angelo, and A. Pedrocchi, "Distributed cerebellar plasticity implements generalized multiple-scale memory components in real-robot sensorimotor tasks," *Frontiers Comput. Neurosci.*, vol. 9, p. 24, Feb. 2015.

[34] F. Naveros, N. R. Luque, J. A. Garrido, R. R. Carrillo, M. Anguita, and E. Ros, "A spiking neural simulator integrating event-driven and time-driven computation schemes using parallel CPU-GPU co-processing: A case study," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 7, pp. 1567–1574, Jul. 2015.

[35] E. M. Izhikevich, "Simple model of spiking neurons," *IEEE Trans. Neural Netw.*, vol. 14, no. 6, pp. 1569–1572, Nov. 2003.

[36] E. M. Izhikevich, "Which model to use for cortical spiking neurons?" *IEEE Trans. Neural Netw.*, vol. 15, no. 5, pp. 1063–1070, Sep. 2004.

[37] T. Yamazaki and S. Tanaka, "The cerebellum as a liquid state machine," *Neural Netw.*, vol. 20, no. 3, pp. 290–297, Apr. 2007.

[38] S. Brandi, I. Herreros, and P. F. M. J. Verschure, "Optimization of the anticipatory reflexes of a computational model of the cerebellum," in *Biomimetic and Biohybrid Systems* (Lecture Notes in Computer Science), vol. 8608, A. Duff, N. F. Lepora, A. Mura, T. J. Prescott, and P. F. M. J. Verschure, Eds. Switzerland: Springer, 2014.

[39] I. Herreros and P. F. M. J. Verschure, "Nucleo-olivary inhibition balances the interaction between the reactive and adaptive layers in motor control," *Neural Netw.*, vol. 47, pp. 64–71, Nov. 2013.

[40] I. Herreros, G. Maffei, S. Brandi, M. Sánchez-Fibla, and P. F. M. J. Verschure, "Speed generalization capabilities of a cerebellar model on a rapid navigation task," in *Proc. IEEE Int. Conf. Intell. Robots Syst.*, Nov. 2013, pp. 363–368.

[41] N. R. Luque, J. A. Garrido, R. R. Carrillo, S. Tolu, and E. Ros, "Adaptive cerebellar spiking model embedded in the control loop: Context switching and robustness against noise," *Int. J. Neural Syst.*, vol. 21, no. 5, pp. 385–401, Oct. 2011.

[42] N. Caporale and Y. Dan, "Spike timing-dependent plasticity: A Hebbian learning rule," *Annu. Rev. Neurosci.*, vol. 31, pp. 25–46, Jul. 2008.

[43] R. Batllori, C. B. Laramee, W. H. Land, and J. D. Schaffer, "Evolving spiking neural networks for robot control," *Procedia Comput. Sci.*, vol. 6, pp. 329–334, Jan. 2011.

[44] H. Hagras, A. Pounds-Cornish, M. Colley, V. Callaghan, and G. Clarke, "Evolving spiking neural network controllers for autonomous robots," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, vol. 5. Apr./May 2004, pp. 4620–4626.

[45] P. Trhan, "The application of spiking neural networks in autonomous robot control," *Comput. Inform.*, vol. 29, no. 5, pp. 823–847, 2012.

[46] J. Thangavelautham and G. M. T. D'Eleuterio, "Tackling learning intractability through topological organization and regulation of cortical networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 23, no. 4, pp. 552–564, Apr. 2012.

[47] W. Van Geit, E. De Schutter, and P. Achard, "Automated neuron model optimization techniques: A review," *Biol. Cybern.*, vol. 99, nos. 4–5, pp. 241–251, Nov. 2008.

[48] D. E. Goldberg and J. H. Holland, "Genetic algorithms and machine learning," *Mach. Learn.*, vol. 3, pp. 95–99, Oct. 1988.

[49] D. E. Goldberg and K. Deb, "A comparative analysis of selection schemes used in genetic algorithms," *Found. Genetic Algorithms*, vol. 1, no. 1, pp. 69–93, 1991.

[50] N. R. Luque, J. A. Garrido, R. R. Carrillo, E. D'Angelo, and E. Ros, "Fast convergence of learning requires plasticity between inferior olive and deep cerebellar nuclei in a manipulation task: A closed-loop robotic simulation," *Frontiers Comput. Neurosci.*, vol. 8, p. 97, Aug. 2014.

[51] A. Rasmussen, D.-A. Jirenhed, D. Z. Wetmore, and G. Hesslow, "Changes in complex spike activity during classical conditioning," *Frontiers Neural Circuits*, vol. 8, p. 90, Aug. 2014.

[52] P. J. E. Attwell, S. Rahman, and C. H. Yeo, "Acquisition of eyeblink conditioning is critically dependent on normal function in cerebellar cortical lobule HVI," *J. Neurosci.*, vol. 21, no. 15, pp. 5715–5722, Aug. 2001.

[53] B. S. Hoffland *et al.*, "Cerebellar theta burst stimulation impairs eyeblink classical conditioning," *J. Physiol.*, vol. 590, no. 4, pp. 887–897, Feb. 2012.

[54] J. A. Garrido, N. R. Luque, E. D'Angelo, and E. Ros, "Distributed cerebellar plasticity implements adaptable gain control in a manipulation task: A closed-loop robotic simulation," *Frontiers Neural Circuits*, vol. 7, no. 10, p. 159, Oct. 2013.

[55] J. F. Medina, K. S. Garcia, and M. D. Mauk, "A mechanism for savings in the cerebellum," *J. Neurosci.*, vol. 21, no. 11, pp. 4081–4089, 2001.

[56] W. Zhang and D. J. Linden, "Long-term depression at the mossy fiber–deep cerebellar nucleus synapse," *J. Neurosci.*, vol. 26, no. 26, pp. 6935–6944, Jun. 2006.

[57] R. C. Miall and D. King, "State estimation in the cerebellum," *Cerebellum*, vol. 7, no. 4, pp. 572–576, Jan. 2008.

[58] J.-Y. Lee and N. Schweighofer, "Dual adaptation supports a parallel architecture of motor memory," *J. Neurosci.*, vol. 29, no. 33, pp. 10396–10404, Aug. 2009.

[59] S. E. Criscimagna-Hemminger, A. J. Bastian, and R. Shadmehr, "Size of error affects cerebellar contributions to motor learning," *J. Neurophysiol.*, vol. 103, no. 4, pp. 2275–2284, 2010.

[60] A. G. Richardson *et al.*, "Disruption of primary motor cortex before learning impairs memory of movement dynamics," *J. Neurosci.*, vol. 26, no. 48, pp. 12466–12470, 2006.

[61] A. Hadipour-Niktarash, C. K. Lee, J. E. Desmond, and R. Shadmehr, "Impairment of retention but not acquisition of a visuomotor skill through time-dependent disruption of primary motor cortex," *J. Neurosci.*, vol. 27, no. 49, pp. 13413–13419, 2007.

[62] J. M. Galea, N. B. Albert, T. Ditye, and R. C. Miall, "Disruption of the dorsolateral prefrontal cortex facilitates the consolidation of procedural skills," *J. Cognit. Neurosci.*, vol. 22, pp. 1158–1164, Jun. 2010.

[63] S. Solinas, T. Nieus, and E. D'Angelo, "A realistic large-scale model of the cerebellum granular layer predicts circuit spatio-temporal filtering properties," *Frontiers Cellular Neurosci.*, vol. 4, p. 12, Jan. 2010.

[64] E. D'Angelo, "The organization of plasticity in the cerebellar cortex: From synapses to control," *Prog. Brain Res.*, vol. 210, pp. 31–58, Jan. 2014.

[65] W. Maass, "Networks of spiking neurons: The third generation of neural network models," *Neural Netw.*, vol. 10, no. 9, pp. 1659–1671, Dec. 1997.

[66] J. V. Tu, "Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes," *J. Clin. Epidemiol.*, vol. 49, no. 11, pp. 1225–1231, Nov. 1996.

[67] A. Dimitrova *et al.*, "Correlation of cerebellar volume with eyeblink conditioning in healthy subjects and in patients with cerebellar cortical degeneration," *Brain Res.*, vol. 1198, pp. 73–84, Mar. 2008.

Alberto Antonietti, photograph and biography not available at the time of publication.

Claudia Casellato, photograph and biography not available at the time of publication.

Egidio D'Angelo, photograph and biography not available at the time of publication.

Alessandra Pedrocchi, photograph and biography not available at the time of publication.