In *Proceedings of the Adaptive Learning Agents Workshop* (as part of AAMAS),
Virtual, May 2021

Is the Cerebellum a Model-Based Reinforcement Learning Agent?

Bharath Masetty
The University of Texas at Austin
Austin, Texas
bmasetty@utexas.edu

Reuth Mirsky
The University of Texas at Austin
Austin, Texas
reuth@cs.utexas.edu

Ashish Deshpande
The University of Texas at Austin
Austin, Texas
ashish@austin.utexas.edu

Michael Mauk
The University of Texas at Austin
Austin, Texas
mike@mail.clm.utexas.edu

Peter Stone
The University of Texas at Austin
Sony AI
Austin, Texas
pstone@cs.utexas.edu

ABSTRACT

The cerebellum is an integral part of the human brain and understanding its role in learning might present an opportunity for reciprocal enrichment of the fields of artificial intelligence and neuroscience. In this paper, we present a novel idea that the cerebellum's role goes beyond functioning as a supervised learning machine to performing model-based reinforcement learning. We revisit the current theories about the cerebellum's role in human learning processes and propose a novel way of evaluating it in the context of the simulated cerebellum. Based on the recent experimental findings, we propose that the cerebellum performs model-based reinforcement learning and we propose a way to evaluate the hypothesis using a simulated cerebellum. Finally, we discuss the necessary conditions to evaluate the hypothesis and the potential implications for future research of the hypothesis holds.

KEYWORDS

Cerebellum, Neuroscience, Reinforcement learning, Forward models, Reward function

ACM Reference Format:
Bharath Masetty, Reuth Mirsky, Ashish Deshpande, Michael Mauk, and Peter Stone. 2021. Is the Cerebellum a Model-Based Reinforcement Learning Agent?. In *Proc. of the 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2021), London, UK, May 3–7, 2021*, IFAAMAS, 5 pages.

1 INTRODUCTION

Over the years, many breakthroughs in artificial intelligence were inspired by neuro-scientific revelations about human learning processes. In a complementary manner, techniques in reinforcement learning have also been used to model neurological phenomena. The introduction of temporal difference learning [31] and the reward prediction error theory of dopamine [29] marked the first point of a remarkable convergence between the two domains. This event sparked a research line targeted at finding more correlations between ideas from reinforcement learning and neuroscience. Some notable examples are using actor-critic models to develop computational models of basal-ganglia [12], and distributional reinforcement learning was used to explain the probabilistic nature of dopaminergic neurons [5]. Most of this interdisciplinary research revolves around reward-related neural correlates, such as the basal ganglia and the ventral tegmental area. A commonly followed methodology is to use reinforcement learning models to explain animal behavioral data or neural recordings. In the paper, we call the reinforcement learning community's attention to another highly explored neural substrate in neuroscience, the cerebellum.

The cerebellum occupies only 10% of the human brain volume but contains more neurons than the rest of the brain. Its cortex is shown to consist of functionally distinct subdivisions called *microzones* which control specific muscle groups [16]. The cerebellum is unequivocally perceived as a brain component that plays a prominent role in movement coordination and motor control [6, 30, 33]. However, a topic of controversy in the current consensus about the cerebellum is its role in cognitive and motor learning [32]. A majority of the previous research identifies the nature of cerebellar learning as a supervised learning process [14, 19]. While this theory holds, we conjecture that the cerebellum's functionality goes beyond supervised learning. We review evidence that highlights the involvement of the cerebellum in learning the forward dynamics models, reward functions, and control policies. Based on the evidence showing these functionalities, we then present our hypothesis that the cerebellum performs model-based reinforcement learning.

In the subsequent sections, we first discuss the current neuroscientific understanding of the cerebellum's role in movement coordination and motor control. Next, we review some instances where simulated representations of the cerebellum were successfully used in robotics and machine learning applications that validate the cerebellum's role in the movement. Later we analyze some recent revelations about the cerebellum's role in learning reward functions of the environment. By combining current views and newly discovered evidence, we then present our hypothesis about the cerebellum's potential to perform reinforcement learning and propose a way to evaluate it using a simulated cerebellum. Finally, we conclude by discussing the potential implications on future research if the theory holds and some of the challenges involved in using cerebellum simulations.

Proc. of the 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2021), U. Endriss, A. Nowé, F. Dignum, A. Lomuscio (eds.), May 3–7, 2021, London, UK.

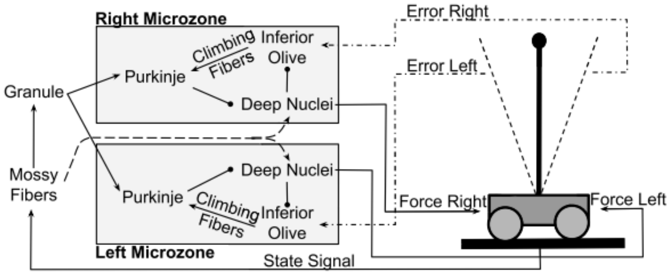

Figure 1: Cerebellum - cart pole interface adopted from [8]. The state information was sent via the mossy fibers and two microzones were used to control the left and right forces to balance the cart pole.

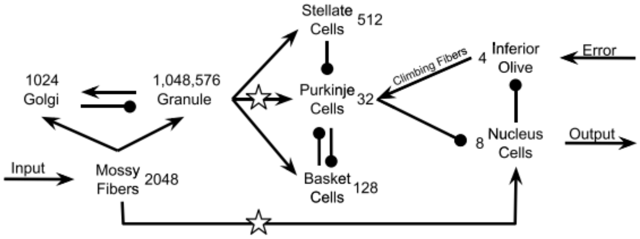

Figure 2: Connectivity and scale of the simulated cerebellum adopted from [8]. Arrows and circles represent excitatory and inhibitory connections respectively. The number of simulated cells are annotated for each region. Stars represent sites of synaptic plasticity.

2 BACKGROUND

Reinforcement learning is a paradigm which consists of an *agent* learning to take sequential decisions to perform a *task* in an *environment*. A task is usually modeled as a Markov Decision Process (MDP) [31]. An MDP M is defined as a 4-tuple $(\mathcal{S}, \mathcal{A}, p, r)$, where $\mathcal{S}$ defines the set of the possible agent's states in the environment, $\mathcal{A}$ defines the set of actions the agent can take, $p(s'|s,a)$ is the *transition function* which defines probability of being in state s' upon taking an action a in a state s and $r(s,a)$ is the *reward function* which determines the reward received by the agent upon taking an action a in a state s. The transition function can also be called a *forward dynamics model*. The forward dynamics model and the reward function together define the *model of an environment*. At every time-step, the agent in a state s takes action a, and the environment responds by giving the information about the future state s' and the reward r. We use the term *reward signal* to represent the single-step reward r in the environment's response. A *policy* is a mapping from $\mathcal{S}$ to $\mathcal{A}$ and the agent's goal is to learn a policy that maximizes the cumulative sum of reward signals received until the task terminates. Reinforcement learning algorithms can be broadly classified into two types: 1) model-based algorithms in which the learning agent has access to or learns the environment's model and 2) model-free algorithms that do not have access to the environment's model. In this paper, we discuss the cerebellum's potential to act as a model-based reinforcement learning agent.

The cerebellum is well known to play an important role in the execution of precisely-timed movements. This ability is shown to stem from its capability to modulate the strength of motor commands using a prediction of future sensory states [13, 15, 21]. The phenomenon of predicting future sensory states was explained via internal forward models of the body dynamics encoded in the cerebellum. In terms of reinforcement learning, an internal model is analogous to the environment's forward dynamics model. In addition to this analogy, the observation that the cerebellum can modulate motor commands implies that it can implement control policies as well.

We use the cart pole balancing task (Figure 1, on the right) as a running example throughout the paper. The cart pole task consists of a cart that can move along one dimension, and an inverted pendulum is set perpendicularly to the cart's moving axis. Four variables define the state space of this problem: pole angle (θ), pole angular velocity (ω), cart position (x), cart velocity (v). The cart is free to move either to the left or right depending on the net amount of force applied to it, which defines the action space. The task is to learn to balance the inverted pendulum by applying the right amount of force at every time step. For the cerebellum to learn the forward dynamics model in the cart pole balancing, it would mean learning to predict the state variables at the next time step, when the current state and applied force are provided as inputs.

The cerebellum is known to possess the required components to be able to learn the forward dynamics models. Figure 2 depicts the basic components of the cerebellum: The mossy fibers are the primary input to the cerebellum. They directly or indirectly receive input from almost all areas of the cortex. The mossy fibers then map onto a huge granular cell layer, thus creating an architecture capable of encoding large state-action spaces and making the cerebellum suitable for learning forward dynamics models [35]. The climbing fibers in the cerebellum provide the necessary error signals known as *sensory prediction errors* (SPE), which is the difference between the predicted future state by the cerebellum and the actual future, and they help in encoding the forward dynamics models. Additionally, tasks that involve simple movements like eyelid conditioning and vestibulo-ocular reflex are known to have direct engagement with the cerebellum, and they are often used to validate and tune the simulated models [8, 26]. It is important to note that the exact inputs provided to the cerebellum and the semantics of the output signals from the cerebellar nucleus are not entirely known, especially in the context of cognitive and motor learning.

The cerebellum's well-understood topology and synaptic organization make it a ripe candidate for simulation neuroscience. Simulation neuroscience is a branch of research that studies computational simulations of neural substrates to exploit their underlying mechanisms [7]. Simulating the cerebellum has been an active area of research over more than a decade, with a continual evolution of increasingly detailed and accurate models [25–27]. Several researchers have used cerebellar simulations to demonstrate its capability to encode internal models by using them in applications like robotic motion control [10]. The adaptive cerebellar spiking model was used to control a robotic arm [20], a cerebellum-inspired neural network was used for dynamic state estimation and model predictive control [2], and the concept of synaptic eligibility traces was used to handle the variable temporal delays between predictive signals and subsequent motor responses in the context of predictive control [17, 23, 24]. The ideas discussed in this paper are based on work done using a biologically constrained bottom-up model of a cerebellum [3, 26] that is built based on the Marr-Albus-Ito theory of cerebellar function [1, 11, 22].

In previous work by Hausknecht et al. [8], machine learning capabilities of a cerebellum simulator were investigated. Their simulation had the convergence/divergence ratios of connectivity between each component layer close to the biological cerebellum. Figure 2 shows the connectivity layout of the simulated cerebellum used in their work. The simulator was used to perform six tasks - eyelid conditioning, cart pole balancing, PID control, robot balancing, pattern recognition, and MNIST handwritten digit recognition. Figure 1 shows the interface between the cart pole environment and the simulated cerebellum. Two microzones were used to represent the left pushing and the right pushing forces. The task is to learn to balance the pole successfully, and it was achieved by using a handcrafted task error function for both microzones, thus reducing the task to a supervised learning problem. This study's main conclusion was that the simulated cerebellum could learn to perform supervised learning and control tasks that have a clear definition of the task error. However, the simulator was unsuccessful in reinforcement learning and temporal pattern recognition tasks. This result was attributed to the simulated cerebellum's inability to handle the *credit assignment problem*, which determines how each of the agent's actions contributes to a specific future outcome.

All the experimental findings and simulated cerebellum studies discussed so far align with an established notion that the cerebellum performs supervised learning [14, 28]. However, some of the recent studies [4, 34] aimed to understand the cerebellum's role in addiction and social behavior have uncovered a previously unappreciated functionality. Experiments on rodents have revealed that the cerebellum has direct excitatory projections onto the ventral tegmental area (VTA), otherwise known as the brain's *rewarding center.* VTA is part of multiple dopaminergic pathways in the brain which have projections onto the cortex and ventral striatum in basal ganglia. External stimulation of the ventral tegmental area has revealed that the cerebellum's excitatory projections are rewarding in nature. Two more studies [9, 18] revealed that the climbing fibers located in fore-limb controlling areas of the cerebellar-cortex in mice were found to modulate their activity in response to the external rewards. The climbing fibers' activity was shown to closely resemble the *reward prediction error* (RPE), which is the difference between the predicted reward and the actual reward received in the next state is thought to be important for reinforcement learning in the brain. These new revelations indicate the cerebellum possibly learns the external reward functions along with the forward dynamics models.

So far, we have discussed evidence that shows that the cerebellum learns forward dynamics models and reward functions. These two functions together constitute the definition of a model of the environment in a reinforcement learning system. Therefore, we can summarize the following points about learning processes in the cerebellum: 1) The cerebellum learns models of the environment, and 2) It is involved in modulation of the action commands that control the muscle movements in the body, which can also be explained as implementing a *control policy.* Based on these two points, we hypothesize that *model-based reinforcement learning could be one of the cerebellum's functionalities.* In the next section, we discuss a possible way to test this hypothesis using a cerebellum simulation.

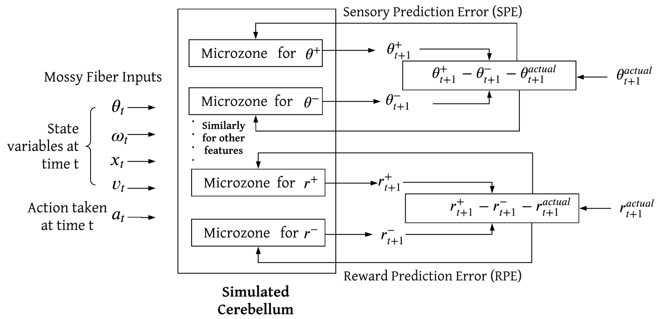

Figure 3: An example supervised learning framework for learning the forward dynamics and reward functions using the simulated cerebellum for the cart pole balancing task.

3 METHODOLOGY

We propose to use a simulated cerebellum to test if it can spontaneously optimize a policy given that it is provided the necessary error signals to learn a forward dynamics function and reward function. We do not assume that the simulated cerebellum acts as a reinforcement learning agent and evaluate its performance. We instead wish to check *if* the cerebellum *can* act as a reinforcement learning agent by creating an appropriate environment guided by our hypothesis. We first explain how learning forward dynamics models and reward functions can be modeled in the cerebellum simulation. We then present a possible way to implement the policy learning process.

3.1 Model Learning

The learning of the forward dynamics models and reward functions can still be modeled as a supervised learning process and do not require an RL agent's decision-making component, namely, a policy. The mossy fiber signals can be used to represent the input variables, and the climbing fiber signals can be used to deliver the error signal, which will act as the SPE and RPE for the forward dynamics model and the reward function, respectively. Following the configuration in [8], Figure 3 shows an example for implementing this approach using the simulated cerebellum for the cart pole balancing task. The four state variables and two action outputs are provided as inputs to the simulation. Since the simulated cerebellum output is typically inferred from the modeled nucleus neurons' firing rates, which can only be positive values, we use two microzones to represent the positive and negative ranges of every output that the simulation needs to learn. For example, a^+ for the amount of left pushing force to be applied on the cart and a^- for the amount of right pushing force. The output would then be the difference in firing rates of

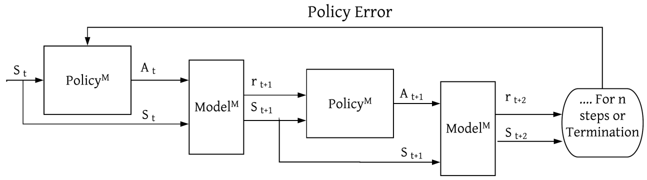

Figure 4: Policy learning framework showing the interaction between the model microzone and policy learning microzones in the proposed n-step update mechanism.

the two microzones of the corresponding output feature. Following this model, for the cart pole task, we would need eight microzones to predict each of the four state variables and two microzones to represent the reward as shown in Figure 3.

Let s be a signal to be learned, and s^+ and s^- be the outputs of microzones associated with positive and negative ranges of signal s. Let $p(s^+)$ and $p(s^-)$ represent the probability of delivering an error signal to positive and negative microzones of s. If s_{t+1} is the signal value at time step t+1, the error probabilities to learn the signal s can be defined as follows:

If $s^+ - s^- < s_{t+1}^{actual}$

$$p(s^+) = 1 \tag{1}$$

$$p(s^-) = 0 \tag{2}$$

If $s^+ - s^- > s_{t+1}^{actual}$

$$p(s^-) = 1 \tag{3}$$

$$p(s^+) = 0 \tag{4}$$

The error probabilities defined above represent the probability for delivering a climbing fiber signal which increases the output of the corresponding microzone. This error function ensures that whenever a predicted state variable is less than its actual value, a deterministic error signal is delivered to increase the value of the corresponding state variable and vice-versa. This error definition applies for both state feature microzones and reward microzones. So far, we have discussed how the cerebellum can be used as a model learning architecture. This approach is in line with the previous section's neuroscientific evidence and does not require any special assumptions.

3.2 Policy Learning

This section proposes a mechanism for policy learning in a simulated cerebellum, which is the second step in testing our hypothesis that the cerebellum can act as a reinforcement learning agent. In the implementation by Hausknecht et al. [8], the control policy was directly learned for the cart pole balancing task using a hand-crafted error function. The error function was carefully defined for the two microzones whose outputs represent the left-ward and right-ward forces. However, to evaluate if the simulated cerebellum performs model-based reinforcement learning, it is essential to show that it can learn an optimal control policy *spontaneously*, using the learned reward function and forward model. As a start, we present a version of such a potential policy learning mechanism, as outlined in Figure 4. The current state features and reward signal are provided as inputs to the policy microzones to get an output representing the action. These microzones were not part of previous simulated architectures. However, the simulated cerebellum structure can very well support them, just as it can support other microzones. The input-output pairs of the policy microzones are then provided as inputs to the model microzones to get the predicted next state and reward, which are fed back to the policy microzones, and the process is repeated. An error is delivered to the policy microzones if the task terminates or fails after n steps. The entire process can be thought of as a finite step look-ahead using the forward dynamics and reward microzones. This way, no particular error function for learning the optimal policy is defined in advance. The success of the approach relies on how learning the forward model and reward function influence the policy microzones in the simulated cerebellum internally.

4 DISCUSSION

To summarize, we combine the cerebellum's established functionalities with some recent discoveries to hypothesize that the cerebellum possibly performs model-based reinforcement learning. We then propose a two-stage method to test this hypothesis using a simulated cerebellum. The first step is to learn the model of the environment consisting of forward dynamics and reward functions. The next step is to perform the n-step look-ahead on policy microzones using the model microzones. This way, we test if the simulated cerebellum can optimize the policy by itself without a perfectly defined task error. This proposed architecture is only one possible approach to test the cerebellum functionality as a reinforcement learning agent. Still, it has the advantage of leveraging a realistic simulation instead of tackling the challenges involved in experiments on natural cerebella and noisy environments. On the other hand, the outcome of the testing approach dramatically depends on the biological accuracy of the simulated cerebellum used to test and the underlying assumptions used in its construction. For example, in the proposed policy learning mechanism, the only possible way for the policy microzones to optimize their output is through their internal interaction with the model microzones. If the simulator's architecture constrains this phenomenon, then the proposed approach will always fail. Another potential issue comes from parameter tuning. In artificial agents, the learning outcome often depends on the choice of the hyperparameters such as discount factors, learning rates, and exploration/exploitation tradeoffs. Simulated cerebellum models can also have similar sensitivities to hyperparameter tuning. In practice, simulated cerebellum models are tuned to match their performance to fit the observations from biological experiments like eyelid conditioning and vestibulo-ocular reflex. However, a single set of hyperparameters might not generalize to all applications, especially when used in non-biological tasks like cartpole balancing. There is also a risk of divergence because the model is iteratively fed with its own outputs, especially when used simulations are modeled imperfectly. These risks can be minimized by using a validated cerebellum simulator, such as the one introduced by Medina and Mauk [26].

If the proposed theory holds, and the cerebellum is shown to perform as a reinforcement learning agent, it could offer many insights to both the AI and neuroscience communities. The AI community's main benefit would stem from disentangling the policy learning from the model and reward learning in the cerebellum at an algorithmic level. This understanding will help explain how the cerebellum handles the credit assignment problem and reward maximization. It could inspire building more sample-efficient and human-like learning agents inspired by the cerebellum's architecture. For neuroscientists, the confirmation of this theory would help create more accurate simulated representations of the cerebellum and pave a path for benchmarking and tuning the simulated representations on reinforcement learning tasks with longer horizons.

ACKNOWLEDGMENTS

This work has taken place in the Learning Agents Research Group (LARG) at UT Austin. LARG research is supported in part by NSF (CPS-1739964, IIS-1724157, NRI-1925082), ONR (N00014-18-2243), FLI (RFP2-000), ARO (W911NF-19-2-0333), DARPA, Lockheed Martin, GM, and Bosch. Peter Stone serves as the Executive Director of Sony AI America and receives financial compensation for this work. The terms of this arrangement have been reviewed and approved by the University of Texas at Austin in accordance with its policy on objectivity in research.

REFERENCES

[1] James S Albus. 1971. A theory of cerebellar function. *Mathematical biosciences* 10, 1-2 (1971), 25–61.

[2] Christopher Assad, Sanjay Dastoor, Salomon Trujillo, and Ling Xu. 2005. Cerebellar dynamic state estimation for a biomorphic robot arm. In *2005 IEEE International Conference on Systems, Man and Cybernetics*, Vol. 1. IEEE, 877–882.

[3] Dean V Buonomano and Michael D Mauk. 1994. Neural network model of the cerebellum: temporal discrimination and the timing of motor responses. *Neural computation* 6, 1 (1994), 38–55.

[4] Ilaria Carta, Christopher H Chen, Amanda L Schott, Schnaude Dorizan, and Kamran Khodakhah. 2019. Cerebellar modulation of the reward circuitry and social behavior. *Science* 363, 6424 (2019).

[5] Will Dabney, Zeb Kurth-Nelson, Naoshige Uchida, Clara Kwon Starkweather, Demis Hassabis, Rémi Munos, and Matthew Botvinick. 2020. A distributional code for value in dopamine-based reinforcement learning. *Nature* 577, 7792 (2020), 671–675.

[6] John C Eccles, Masao Ito, and János Szentágothai. 1967. The mossy fiber input into the cerebellar cortex and its inhibitory control by Golgi cells. In *The cerebellum as a neuronal machine*. Springer, 116–155.

[7] Xue Fan and Henry Markram. 2019. A brief history of simulation neuroscience. *Frontiers in neuroinformatics* 13 (2019), 32.

[8] Matthew Hausknecht, Wen-Ke Li, Michael Mauk, and Peter Stone. 2016. Machine learning capabilities of a simulated cerebellum. *IEEE transactions on neural networks and learning systems* 28, 3 (2016), 510–522.

[9] William Heffley, Eun Young Song, Ziye Xu, Benjamin N. Taylor, Mary Anne Hughes, Andrew McKinney, Mati Joshua, and Court Hull. 2018. Coordinated cerebellar climbing fiber activity signals learned sensorimotor predictions. *Nature Neuroscience* 21, 10 (01 Oct 2018), 1431–1441. https://doi.org/10.1038/s41593-018-0228-8

[10] Constanze HofstoÈtter, Matti Mintz, and Paul FMJ Verschure. 2002. The cerebellum in action: a simulation and robotics study. *European Journal of Neuroscience* 16, 7 (2002), 1361–1376.

[11] Masao Ito. 1989. Long-term depression. *Annual review of neuroscience* 12, 1 (1989), 85–102.

[12] Daphna Joel, Yael Niv, and Eytan Ruppin. 2002. Actor–critic models of the basal ganglia: New anatomical and computational perspectives. *Neural networks* 15, 4-6 (2002), 535–547.

[13] Mitsuo Kawato. 1999. Internal models for motor control and trajectory planning. *Current opinion in neurobiology* 9, 6 (1999), 718–727.

[14] Mitsuo Kawato, Shinya Kuroda, and Nicolas Schweighofer. 2011. Cerebellar supervised learning revisited: biophysical modeling and degrees-of-freedom control. *Current opinion in neurobiology* 21, 5 (2011), 791–800.

[15] Mitsuo Kawato, Shogo Ohmae, Huu Hoang, and Terry Sanger. 2020. 50 years since the Marr, Ito, and Albus models of the cerebellum. *arXiv preprint arXiv:2003.05647* (2020).

[16] Garrett T Kenyon. 1997. A model of long-term memory storage in the cerebellar cortex: a possible role for plasticity at parallel fiber synapses onto stellate/basket interneurons. *Proceedings of the National Academy of Sciences* 94, 25 (1997), 14200–14205.

[17] RE Kettner, S Mahamud, H-C Leung, N Sitkoff, JC Houk, BW Peterson, and AG Barto. 1997. Prediction of complex two-dimensional trajectories by a cerebellar model of smooth pursuit eye movement. *Journal of neurophysiology* 77, 4 (1997), 2115–2130.

[18] Dimitar Kostadinov, Maxime Beau, Marta Blanco-Pozo, and Michael Häusser. 2019. Predictive and reactive reward signals conveyed by climbing fiber inputs to cerebellar Purkinje cells. *Nature Neuroscience* 22, 6 (01 Jun 2019), 950–962. https://doi.org/10.1038/s41593-019-0381-8

[19] John W Krakauer, Alkis M Hadjiosif, Jing Xu, Aaron L Wong, and Adrian M Haith. 2019. Motor learning. *Comprehensive Physiology* 9 (2019), 613–663.

[20] Niceto R Luque, Jesús Alberto Garrido, Richard R Carrillo, Silvia Tolu, and Eduardo Ros. 2011. Adaptive cerebellar spiking model embedded in the control loop: Context switching and robustness against noise. *International Journal of Neural Systems* 21, 05 (2011), 385–401.

[21] Mario Manto, James M Bower, Adriana Bastos Conforto, José M Delgado-García, Suzete Nascimento Farias Da Guarda, Marcus Gerwig, Christophe Habas, Nobuhiro Hagura, Richard B Ivry, Peter Mariën, et al. 2012. Consensus paper: roles of the cerebellum in motor control—the diversity of ideas on cerebellar involvement in movement. *The Cerebellum* 11, 2 (2012), 457–487.

[22] David Marr and W Thomas Thach. 1991. A theory of cerebellar cortex. In *From the Retina to the Neocortex*. Springer, 11–50.

[23] Jeffrey L McKinstry, Gerald M Edelman, and Jeffrey L Krichmar. 2006. A cerebellar model for predictive motor control tested in a brain-based device. *Proceedings of the National Academy of Sciences* 103, 9 (2006), 3387–3392.

[24] Jeffrey L McKinstry, Anil K Seth, Gerald M Edelman, and Jeffrey L Krichmar. 2008. Embodied models of delayed neural responses: Spatiotemporal categorization and predictive motor control in brain based devices. *Neural networks* 21, 4 (2008), 553–561.

[25] Javier F Medina, Keith S Garcia, William L Nores, Nichole M Taylor, and Michael D Mauk. 2000. Timing mechanisms in the cerebellum: testing predictions of a large-scale computer simulation. *Journal of Neuroscience* 20, 14 (2000), 5516–5525.

[26] Javier F Medina and Michael D Mauk. 2000. Computer simulation of cerebellar information processing. *nature neuroscience* 3, 11 (2000), 1205–1211.

[27] A Pellionisz, R Llinas, and DH Perkel. 1977. A computer model of the cerebellar cortex of the frog. *Neuroscience* 2, 1 (1977), 19–35.

[28] Jennifer L Raymond and Javier F Medina. 2018. Computational principles of supervised learning in the cerebellum. *Annual review of neuroscience* 41 (2018), 233–253.

[29] Wolfram Schultz, Peter Dayan, and P Read Montague. 1997. A neural substrate of prediction and reward. *Science* 275, 5306 (1997), 1593–1599.

[30] Gordon M Shepherd. 2004. *The synaptic organization of the brain.* Oxford university press.

[31] Richard S Sutton. 1988. Learning to predict by the methods of temporal differences. *Machine learning* 3, 1 (1988), 9–44.

[32] William T Thach. 1998. What is the role of the cerebellum in motor learning and cognition? *Trends in cognitive sciences* 2, 9 (1998), 331–337.

[33] W Thomas Thach, HP Goodkin, and JG Keating. 1992. The cerebellum and the adaptive coordination of movement. *Annual review of neuroscience* 15, 1 (1992), 403–442.

[34] Mark J Wagner and Liqun Luo. 2020. Neocortex–Cerebellum Circuits for Cognitive Processing. *Trends in Neurosciences* 43, 1 (2020), 42–54.

[35] Daniel M Wolpert, R Chris Miall, and Mitsuo Kawato. 1998. Internal models in the cerebellum. *Trends in cognitive sciences* 2, 9 (1998), 338–347.