Simulation of Neurofeedback-Induced Striatal Learning using a simple EEG signal generator

Nathan Godey
nathan.godey@eleves.enpc.fr

1. Critical study of the original article

1.1. Motivations

Neurofeedback is a type of neural training that consists in presenting to a subject a feedback based on a measurement taken on its own brain, and to ask it to decrease or increase this measurement. In recent years, this approach seems to have been validated on a number of occasions in the clinical domain, where it appears to yield interesting results in some applications, as one can read in [1, 2, 3]. Usually, it is thought that neurofeedback enhances attention and focus, which is a key factor in curing diseases such as attention-deficit hyperactivity disorder (ADHD) [4] or severe depression [5].

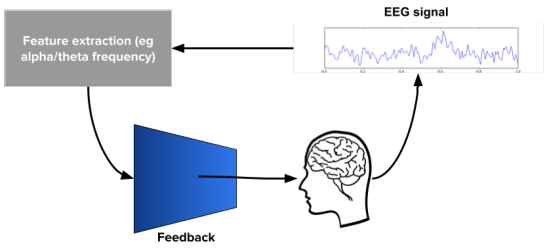


Figure 1: Principle of Neurofeedback

At the same time, this clinical appeal is often lacking strong scientific foundations, and it might be that neurofeedback's success thrives on both a financial and ideological enthusiasm as pointed out in [6], . Some work also expose the impact of psychological phenomenas related to the experiments themselves, such as placebo (or neuroplacebo) effect [7].

The article [8] we are going to study is an attempt to introduce a computational framework for neurofeedback, and more specifically EEG-based neurofeedback. An accurate and rigourous modelling of neurofeedback learning would actually help in two different ways:

- Provide researchers with a fast way to simulate neurofeedback learning using various hypothesis, and to tune the range of useful parameters for their experiments
- Allow clinical research to take advantage of a computational support for more robustness in its presented results, which would also help gather a better understanding of the underlying processes involved in neurofeedback learning and avoid purely empirical conclusions that may be flawed (see [7, 6])

The authors of [8] advocate that even a basic model allow to model neurofeedback efficiently and to yield conclusions regarding the threshold setting for instance (at what level of the EEG-based measurement to set positive or negative feedback).

1.2. Mechanisms of neurofeedback

A key hypothesis in Davelaar [8] is that neurofeedback learning mainly occurs in the striatum.

This hypothesis is based on empirical observations (see Levesque et al., Hinterberg et al., and Johnston et al.). Another empirical finding is that the thalamus plays a regulating role for the Slow Cortical Potential (or SCP), which is highly related to the EEG signal involved in Neurofeedback.

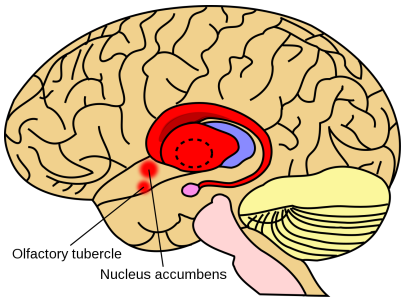


Figure 2: Striatum (in red) and Thalamus (in blue)

Davelaar et al. propose a theory for the underlying mechanism behind neurofeedback learning which is based on three distinct phases:

- **Exploration stage**: The subject (or trainee) tries different strategies, either coming from the exterior (the trainer, the internet, ...), or idiosyncratic attempts. Each strategy corresponds to a feedback, which allows the trainee to evaluate their performance (consciously or subconsciously). The learning process is expected to happen in the frontal-striatal connectivity.

- **Consolidation stage**: The trainee has found a working strategy (again, consciously or subconsciously). A reinforcement learning process occurs between the striatal representation obtained in the first stage, and thalamic pathways between the striatum and the cortex. This phenomenon has actually been observed empirically (Ghaziri et al. 2013 [9]). The trainee also becomes able to give a subjective description of how it can enhance its own neurofeedback performance (e.g. relaxation, tension or meditative thoughts).

- **Interoceptive stage**: The trainee is now in an *interoceptive* state, i.e. it consciously tries to remain in a physiological state that results in successful feedback, based on the two first phases.

1.3. Problem

Davelaar et al. [8] study a neurofeedback experiment based on alpha frequency that is described and tested in the article from Zoefel et al. [1]. In the actual experiment, the goal is to increase the Peak Alpha Frequency (PAF) in the EEG signal, i.e. the maximum amplitude of the signal in the [8Hz;12Hz] frequency range. In order to do so, the Upper Alpha Frequency (UAF), defined by the average amplitude of the signal in the [PAF;PAF+2Hz] frequency range, is the target metric. The EEG signal is captured in several areas of the brain (P3, Pz, O1, O2 and P4). After measuring a baseline UAF ($UAF_{baseline}$) when the trainee is not stimulated or trained, Zoefel et al. provide it with a visual feedback regarding whether its current UAF is above $UAF_{baseline}$ (positive feedback) or below $UAF_{baseline}$ (negative feedback).

The Peak Alpha Frequency is clinically associated with better *cognitive preparedness* [10]. One should also notice that in [10], the PAF is said to be affected by cognitive tasks, which may indicate that neurofeedback learning's efficiency in that situation could be in part due to the fact it is a cognitive task *per se*.

1.4. Simulations

1.4.1 Experiment 1

In their first experiment, Davelaar et al. [8] propose a basic model to simulate the neurofeedback learning presented in [1]. This model is made of two components:

- The **striatum**, represented by 1000 disconnected Medium Spiny Neurons (MSN) units, that each have a probability p_i of being activated. We want about 1% of the MSN units

to be activated at the same time so we make sure that $\sum p_i = 10$ at any time step via normalization

- The **EEG generator** based on the *Simple Model of Spiking Neurons* made by Izhikevich in [11], made of 800 excitatory and 200 inhibitory spiking neurons.

In the striatum model, the MSN units have a dynamic activation probability p_i, which acts as a Bernoulli parameter at each time step (1ms). At each time step, we normalize the probabilities so that $\sum p_i = 10$ to have 10 active units by time step in average.

The MSN units receive neurofeedback in $\{-1; 1\}$ each 100ms, and each p_i are decreased (negative feedback) or increased (positive feedback) by the *count* of each unit's activation in the last 1024 iterations.

For every time step, each probability p_i is decreased by each unit's activation *rate* in the last 1024 iterations, before normalization. When the target unit is active, the thalamic input is increased by one.

In the EEG Generator, the dynamical behaviour of the membrane potential of a neuron v is based on the level of the simulated thalamic input, both on excitatory and inhibitory neurons. The excitatory thalamic input is drawn from $\mathcal{N}(0, \sqrt{5})$ and the inhibitory thalamic input is drawn from $\mathcal{N}(0, \sqrt{2})$. On the other hand, a membrane recovery variable u tend to reduce spiking rates, based on previous values of v (making strictly consecutive spikes less likely). The exact system definition is :

$$\begin{cases} v' = 0.04v^2 + 5v + 140 - u + I \\ u' = a(bv - u) \\ \text{if } v \geq 30, v = c \text{ and } u = u + d \end{cases} \quad (1)$$

where a, b, c and d are parameters defined independently for each neuron. Each excitatory neuron (resp. inhibitory neuron) has a corresponding r_e parameter (resp. r_i) drawn from $\mathcal{U}([0, 1])$. Then, we define :

$$\begin{cases} a = 0.02 + 0.08 r_i \\ b = 0.2 + 0.05(1 - r_i) \\ c = -65 + 15 r_e^2 \\ d = 2 + 6(1 - r_e^2) \end{cases} \quad (2)$$

To retrieve the EEG signal, we compute the sum of membrane potentials of the excitatory neurons at a given time $\sum v_e(t)$. In order to increase the significance of the alpha frequency band, the raw EEG signal is processed. We first need to compute the low-pass filtered signal :

$$EEG(t) = 0.9 EEG(t-1) + 0.1 \sum v_e(t)$$

Finally, we apply a Hamming filter over the last 1024 time iterations and we compute the power spectrum of the signal to get the UAF. The feedback is then defined by :

$$feedback = \mathbb{1}_{UAF_t > UAF_{base}} - \mathbb{1}_{UAF_t \leq UAF_{base}}$$

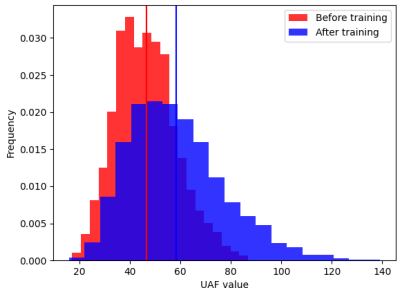


Figure 3: UAF distributions for the baseline experiment and after neurofeedback learning. One can see how the UAF distribution was transported to higher UAF values.

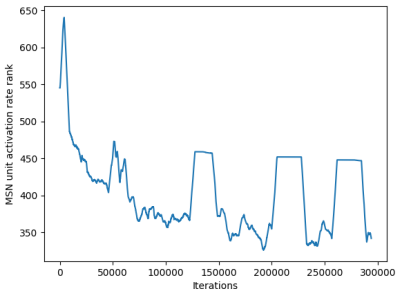
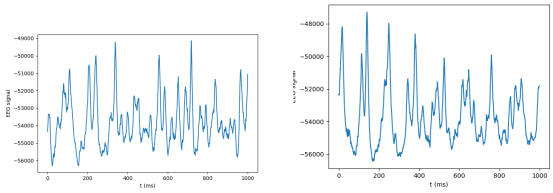


Figure 4: Evolution of the rank of the target MSN unit in terms of activation rate (over the 1000 units). We can see that the activation rate happens to decrease when reaching convergence, which might be caused by the EEG generator maintaining a high UAF value by itself for some time after the target unit has been activated more frequently.



(a) Baseline (b) During learning

Figure 5: Generated EEG time signal

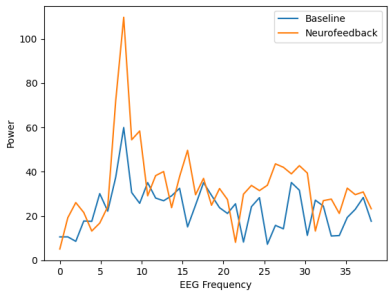


Figure 6: Power spectrum of the EEG signal. We clearly see the increase in the PAF value induced by neurofeedback training.

1.4.2 Experiment 2

In that second experiment, we are going to consider that the distributions of the Upper Alpha Frequency we obtained earlier are fixed throughout time.

Namely, we have two distributions $\mathcal{D}_{UAF|MSN_{target}=1}$ and $\mathcal{D}_{UAF|MSN_{target}=0}$ corresponding to the UAF values that were obtained when the target MSN unit was either activated or deactivated (see Figure 6). Thus, in this experiment, we will directly draw the UAF values from these distributions based on the state of the target unit. This will allow us to focus on the dynamics of the striatal model without simulating the costly EEG signal generation.

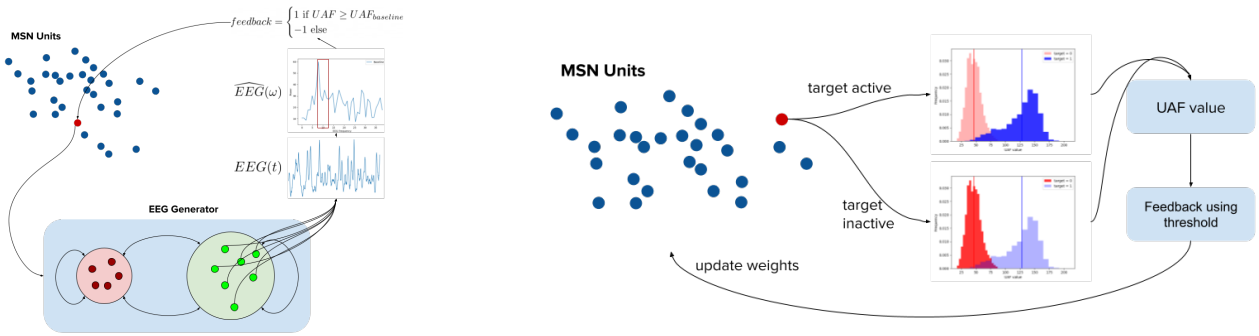
The main motivation of this model is to reduce computational complexity to make it easier to compare simulations for different parameters (here, the feedback threshold).

The striatum model is also modified. The MSN units now have weights of initial value 1. At each iteration, *exactly* 10 units are activated according to their normalized weights (acting as probabilities). If the target unit is activated, the UAF value is drawn from $\mathcal{D}_{UAF|MSN_{target}=1}$, and else from $\mathcal{D}_{UAF|MSN_{target}=0}$. The sampled UAF value is compared to the chosen threshold T :

- if the sampled UAF is smaller than the threshold, the (unnormalized) weights of the 10 activated units are decreased by η
- if the sampled UAF is larger than the threshold, the (unnormalized) weights of the 10 activated units are increased by η

The learning rate η is set at 0.1. The range of possible thresholds is limited by $\mathcal{T} = Conv(supp(\mathcal{D}_{UAF|MSN_{target}=0} + \mathcal{D}_{UAF|MSN_{target}=1}))$.

In the following figures, 50 simulated learning sessions of 10000 time iterations (i.e. 10s) were done for each of 30 threshold levels linearly distributed in $\mathcal{T}$.



(a) First experiment

(b) Second experiment

Figure 7: Schemas of the designed experiments

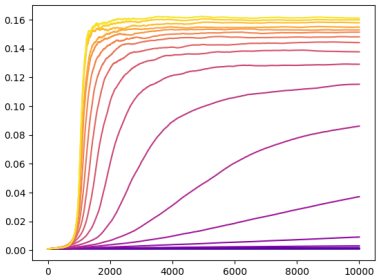
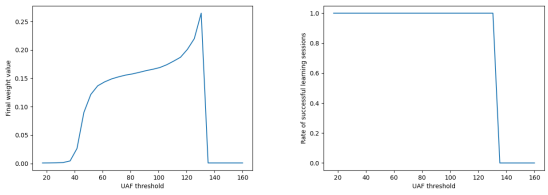


Figure 8: Evolution of the average synaptic weight of the target unit with for 10000 time iterations for several threshold values. The curves with hotter colors represent simulations with higher thresholds.



(a) Final weight

(b) Rate of learners

Figure 9: Final results with regards to the threshold value.

As mentioned in Davelaar et al. [8], we observe an optimal threshold value for which the final synaptic weight is maximized. This threshold is critical since increasing it further seems to make learning fail (see Figure 12b). This is described as an attractor system, where either of two phenomena can occur:

- Once the target unit is activated once, the probability that the drawn UAF value is greater than the threshold $P(UAF > T|MSN_{target} = 1)$ is greater than $\frac{1}{2}$ so that the weight of the target unit is increased more frequently than it is decreased, which results in a positive learning feedback loop. This phenomenon happens at smaller T values.
- $P(UAF > T|MSN_{target} = 1) \leq \frac{1}{2}$ and the weight of the target unit is decreased more frequently than it is increased, which means no positive learning feedback loop can occur. This phenomenon happens for larger T values.

Another constraint on the threshold T for the model to learn is that it needs not mistake the other units for the target unit. This happens when $P(UAF > T|MSN_{target} = 0)$ is too high, i.e. when T is too big with respect to $\mathcal{D}_{UAF|MSN_{target}=0}$. This explains why we find an optimal threshold T_{opt} that belongs approximately at the "intersection" of $\mathcal{D}_{UAF|MSN_{target}=0}$ and $\mathcal{D}_{UAF|MSN_{target}=1}$, both maximizing $P(UAF > T|MSN_{target} = 1)$ and minimizing $P(UAF > T|MSN_{target} = 0)$.

2. Remarks and variations

2.1. Theoretical bound on the threshold T

A necessary condition for a learning process to happen in the second model (i.e. for the target unit weight to increase relatively to the other units) is that the target (unnormalized) weight w_t^T tend to increase at iterations when the MSN target is active, for a given threshold T. Formally, we want to have:

$$E((w_{t+1}^T - w_t^T)|MSN_t^{target} = 1) > 0$$

The dynamic of w_t^T can be introduced in the above equation, with η as the learning rate used for the MSN units :

$$\begin{aligned} &E((w_{t+1}^T - w_t^T)|MSN_t^{target} = 1) \\ &= \eta(P(UAF \geq T|MSN_t^{target} = 1) \\ &- P(UAF < T|MSN_t^{target} = 1)) \\ &= \eta(1 - 2P(UAF < T|MSN_t^{target} = 1)) \end{aligned} \tag{3}$$

By combining our two results, we get :

$$P(UAF < T|MSN_t^{target} = 1) < \frac{1}{2}$$

Finally, if $F_1(T)$ is the CDF of the continuous random variable $(UAF < T|MSN_t^{target} = 1)$, we get that the model can successfully learn only if:

$$T < F_1^{-1}(\frac{1}{2})$$

By repeating the second experiment on a range of thresholds surrounding the value of the observed bound, we see that the bound is very close to the optimal threshold.

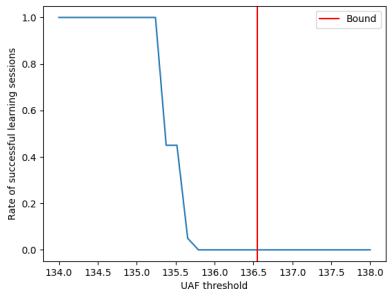


Figure 10: Rate of successful learning sessions in the neighbourhood of the bound

2.2. Relaxing constraints

In order to make the second experiment simulation faster, we can try to relax the constraint on the number of active units at each iteration. Namely, the current way to draw the active units is to pick randomly *exactly* 10 units using the probabilities $p_{i,t}^T$ at iteration t for every unit i as weights. The algorithmic way to pick 10 units in such manner at a given iteration is to stack $10 - k$ units (where k is the current number of drawn units) according to $(p_{i,t}^T)_i$, making sure we don't pick a unit twice, until $k = 10$. The relaxation we implemented is to pick units according to a simple Bernoulli law with $(p_{i,t}^T)_i$, and to normalize the activation probabilities so that:

$$E(N) = \sum_{i=1}^{1000} p_{i,t}^T = 10$$

where N is the r.v. representing the number of active units. This allows us to get 10 active units in average, with:

$$V(N) = \sum_{i=1}^{1000} (1 - p_{i,t}^T) p_{i,t}^T$$

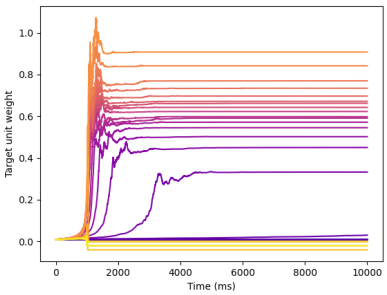
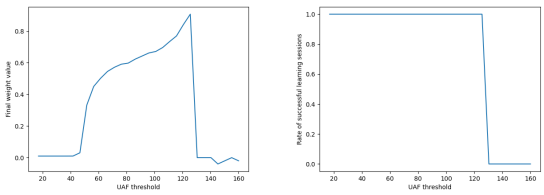


Figure 11: Evolution of the average synaptic weight (relaxed dynamic)



(a) Final weight (b) Rate of learners

Figure 12: Final results with respect to the threshold value (relaxed dynamic)

As we can see in Figures 11 and 12, the shape of the dynamic does not change much and the final results for each threshold yield the same conclusions about the existence of an optimal threshold. On the other hand, the time complexity for the whole simulation seems to be divided by 4, mostly because the draw of random values can be done once and for all, which allows to take advantage of NumPy's efficiency.

As a matter of fact, in the original implementation, the random process of drawing active MSN units needed to be done at each time step because each unit's activity was not independent, which is the case in our implementation

2.3. Importance of the EEG Generator

A very important factor in the result of the modelling is the design of the EEG Generator. While the a, b, c and d variables (or more precisely their ranges) seem to be inherent to the structure of the individual MSN units, we can think that different cortical areas need to be modelled by different proportions of excitatory units and inhibitory units.

What is worth-studying is the impact of the choice of the number of units of each category in the baseline UAF distribution.

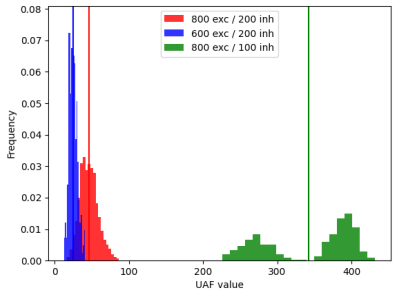


Figure 13: UAF distributions for different EEG settings

As one can see in Figure 13, using different EEG settings can yield distributions that may vary much both in shape and value range. Interestingly, as the proportion of excitatory units relatively increases, the distribution becomes a "mixture" and does not resemble a Gaussian / χ^2 law.

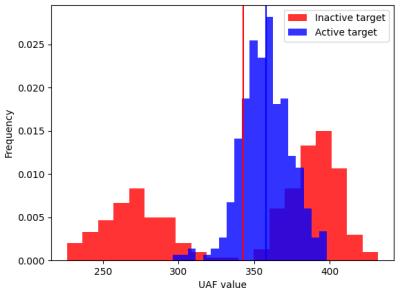


Figure 14: UAF distributions when the target MSN unit is either active or inactive for 800 excitatory cells and 100 inhibitory cells in the Izhikevich model

If we focus on this specific setting (800 excitatory cells against 100 inhibitory cells), and we try

to reproduce the results obtained in the second experiment, we first see that the UAF distributions do not allow for easy threshold discrimination as pictured in Figure 14. The mean UAF value is still higher when the target is active than when it is not, but if we set a threshold T between these two means, we will still obtain a high probability $P(UAF > T|MSN_{target} = 0)$, implying a high probability to get a positive feedback when the target unit is inactive.

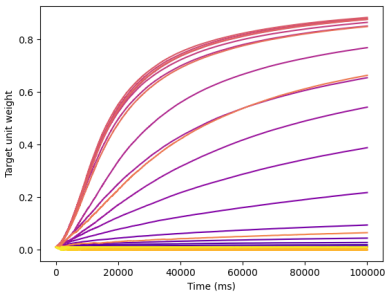
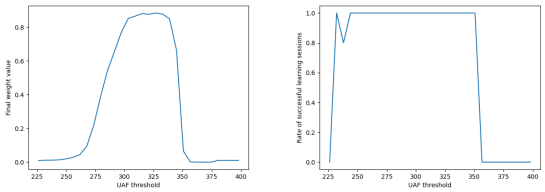


Figure 15: Evolution of the average synaptic weight (800 excitatory units, 100 inhibitory units)



(a) Final weight (b) Rate of learners

Figure 16: Final results with respect to the threshold value (800 excitatory units, 100 inhibitory units)

As we can see in Figure 16, the target unit weight takes much longer to converge in average, as it is still increasing for all threshold levels after 100000 iterations. Moreover, Figure 16a shows that the final activation weight of the target unit is similar for a significant range of thresholds, which questions the relevance of choosing an "optimal" threshold in that situation.

2.4. Continuous Neurofeedback

An important setting that is used in the implementation from Davelaar et al. [8] is the binary behaviour of the feedback system. In their article, the feedback is computed as:

$$feedback = \begin{cases} 1 \text{ if } UAF > UAF_{baseline} \\ -1 \text{ else} \end{cases} \quad (4)$$

This is an computationally efficient way to represent neurofeedback, but one can wonder if it enables the model to capture the complexity of the actual feedback that is provided in the experiments. As a matter of fact, the designed protocols often integrate feedback through media (video, audio, game...), which questions the binary aspect of the threshold. For instance, if the protocol consists in asking the subject to increase the volume of an audio, a more realistic approach would be to use a *continuous* modelling of the feedback, which would in this case be related to the volume level.

A basic way to represent continuous feedback is to use a $tanh$-like function:

$$feedback = 1 - \frac{2}{1 + exp(\frac{UAF - UAF_{baseline}}{\tau})}$$

The τ parameter accounts for the softness of the feedback curve (see Figure 17).

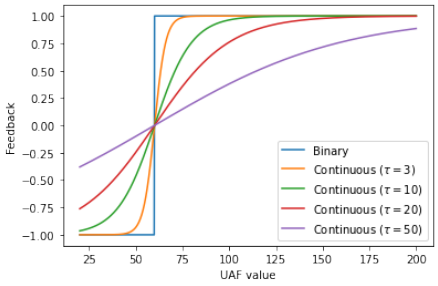


Figure 17: Feedback curve in function of the UAF value for different τ settings ($UAF_{baseline} = 60$)

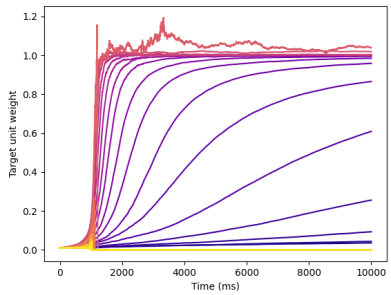
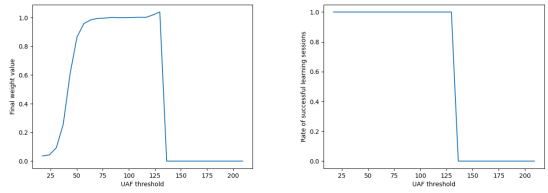


Figure 18: Evolution of the average synaptic weight (continuous neurofeedback, $\tau = 20$)



(a) Final weight (b) Rate of learners

Figure 19: Final results with respect to the threshold value (continuous neurofeedback, $\tau = 20$)

As one can see in the figures above, using continuous feedback does not change much the global behaviour of the model. An optimal threshold can still be found, around the same value as for the binary feedback ($T_{opt} \simeq 135$).

If we zoom in to compare the specific behaviour of the final weight values, we obtain Figure 20.

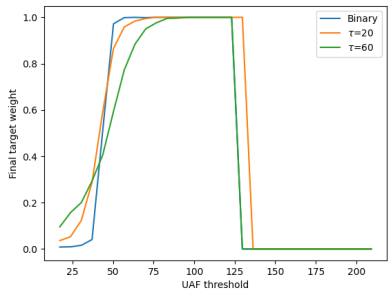


Figure 20: Evolution of the final weight of the target unit with respect to the set threshold for different types of feedback

In Figure 20, we observed several phenomena:

- For smaller threshold values, the continuous feedback allows faster convergence of the target unit weight. This can be explained by the ability to distinguish feedbacks even above the threshold value, which motivates for activating the target unit.
- The binary feedback yields a binary convergence behaviour : a larger band of threshold levels has converged than for the continuous feedback systems, while smaller thresholds did not converge at all
- For a reasonable τ setting ($\tau = 20$), the optimal threshold seems to be pushed slightly further

All in all, we see that the choice of the feedback system is not trivial, and that it has an undeniable impact on the way neurofeedback learning works in our model.

3. Conclusion

As mentioned in the *General Discussion* section of the article from Davelaar et al. [8], these models pave the way for a more complex simulation, taking into account multiple areas of the brain for EEG signal generation for instance. The initial goal of this project was to replace the Izhikevich model with an adaptive linear-nonlinear cascade model as a reduction of a biophysically grounded model of Ad-Ex neuron populations [12] for EEG signal generation.

However, one can see that the models defined in this article have a strong complexity, both mathematically and in terms of time computation. There seems to be room for many variations, which could explain the more robust articles that emerge in the field of Neurofeedback. It should also be noted that there is a lack of open access biophysical data for Neurofeedback experiments, which makes the prospect of building a robust model harder, since modelled results cannot be compared to measurements.

Hence, future work could help understand better the way the two models presented here work, both from a modelling and a mathematical point of view, by investigating the way feedback mechanics affect the dynamics of the learning process for instance. Another way to validate this model would be to compare its results to physical measurements on another neurofeedback training task, e.g. for Theta-Beta Ratios (TBR) enhancements in ADHD treatment. Finally, as mentioned above, a real validation of this framework could be achieved by replacing the EEG signal generator with a more complex model with multiple brain areas and an actual modelling of the thalamic behaviour.

3.1. Acknowledgement

This project benefited from the precious help of Pr. Eddy Davelaar from the University of Birbeck, whom I hereby thank.

References

[1] Benedikt Zoefel, René J. Huster, and Christoph S. Herrmann. Neurofeedback training of the upper alpha frequency band in eeg improves cognitive performance. *NeuroImage*, 54(2):1427–1431, 2011.

[2] Pietraszewski P et al Maszczyk A, Gołaś A. Neurofeedback for the enhancement of dynamic balance of judokas. *Biol Sport*, 35(1):99–102, 2018.

[3] Mariela Rance, Christopher Walsh, Denis G. Sukhodolsky, Brian Pittman, Maolin Qiu, Stephen A. Kichuk, Suzanne Wasylink, William N. Koller, Michael Bloch, Patricia Gruner, Dustin Scheinost, Christopher Pittenger, and Michelle Hampson. Time course of clinical change following neurofeedback. *NeuroImage*, 181:807–813, 2018.

[4] Stefanie Enriquez-Geppert, Diede Smit, Miguel Garcia Pimenta, and Martijn Arns. Neurofeedback as a treatment intervention in adhd: Current evidence and practice. *Current psychiatry reports*, 21(6):46–46, 2019.

[5]

[6] Robert T Thibault, Michael Lifshitz, and Amir Raz. The climate of neurofeedback: scientific rigour and the perils of ideology. *Brain*, 141(2):e11–e11, 12 2017.

[7] Robert T. Thibault, Michael Lifshitz, and Amir Raz. Neurofeedback or neuroplacebo? *Brain*, 140(4):862–864, 03 2017.

[8] Eddy J. Davelaar. Mechanisms of neurofeedback: A computation-theoretic approach. *Neuroscience*, 378:175–188, May 2018.

[9] Jimmy Ghaziri, Alan Tucholka, Vanessa Larue, Myriam Blanchette-Sylvestre, Gabrielle Reyburn, Guillaume Gilbert, Johanne Lévesque, and Mario Beauregard. Neurofeedback training induces changes in white and gray matter. *Clinical EEG and Neuroscience*, 44(4):265–272, March 2013.

[10] Efthymios Angelakis, Joel F Lubar, Stamatina Stathopoulou, and John Kounios. Peak alpha frequency: an electroencephalographic measure of cognitive preparedness. *Clinical Neurophysiology*, 115(4):887–897, April 2004.

[11] E.M. Izhikevich. Simple model of spiking neurons. *IEEE Transactions on Neural Networks*, 14(6):1569–1572, November 2003.

[12] Cakan Caglar and Obermayer Klaus. Biophysically grounded mean-field models of neural populations under electrical stimulation. *PLOS Computational Biology*, 16:1–30, 04 2020.