Dopamine-Gated Plasticity in a Model of Drosophila Memory Formation

Marshall L. Mykietyshyn (marshall@bccn-berlin.de)
Computational Neuroscience, Technische Universität Berlin
Straße des 17. Juni 135, 10623 Berlin

Abstract

This paper reproduces the results of Jiang and Litwin-Kumar's model for dopamine-gated plastic learning in the mushroom body structures of Drosophila. After demonstrating that our model performs to the same qualitative degree of accuracy as the original, the complexity is reduced by removing redundant features and minimizing the number of mushroom body neurons, while still maintaining performance. It was found that the entire feedback neuron population, long-term potentiation, and half of the mushroom body neurons were not necessary to achieve similar performance. Following this, we simulated "optogenetically" imprinting a memory, consolidating that memory, and testing where the memory trace moved during consolidation. The results of this experiment were not conclusive and require further investigation.

Introduction

Dopaminergic neurons are thought to be responsible for gating learning plasticity in reward situations (Schultz, Dayan, & Montague, 1997). This report presents the replication of a dopamine-gated plasticity model for learning in the mushroom body circuit of Drosophila. Specifically, mushroom body neurons in Drosophila are responsible for associative olfactory learning (Jiang & Litwin-Kumar, 2020). Dopamine is often thought of as a global signal during reward learning. However, it may be that dopamine functions in a more targeted fashion. Jiang and Litwin-Kumar explore the role of dopamine in a more complex model of learning in the Drosophila, where dopamine neurons are paired with mushroom body neurons in a compartment structure. Specific mushroom body neurons are considered responsible for different behaviours of the fly. Therefore, closely spaced signals from olfactory inputs and dopamine neurons can modulate behavioural responses to stimuli (odours) through plasticity.

Jiang and Litwin-Kumar propose a biologically plausible model of the mechanisms underlying this type of dopamine-gated plastic learning, which is investigated through classical conditioning tasks. This paper replicates these experiments, then attempts to find a minimal version of the model by systematically reducing it and comparing the performance of each iteration. This minimal model is then used in a task where a false memory is imprinted and consolidated. This is meant to simulate an optogenetically imprinted memory, which is then consolidated. The goal of this analysis is to create a version of the Jiang and Litwin-Kumar model used to run simulations, and make predictions, which can subsequently be tested in real lab experiments.

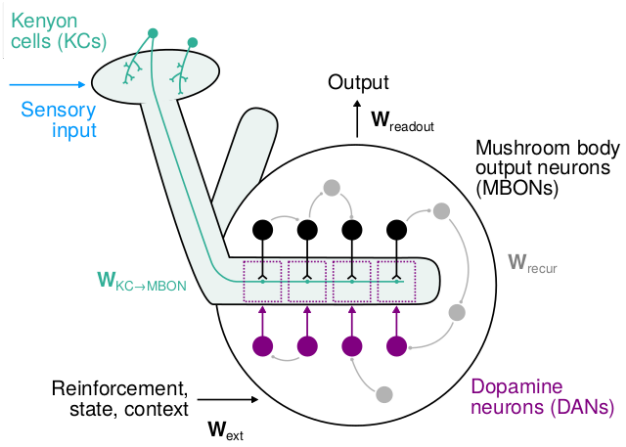

Figure 1: Drosophila memory circuit (Jiang & Litwin-Kumar, 2020)

Model

Jiang and Litwin-Kumar developed a model of Drosophila memory consisting of Kenyon cell (KC) and context (ext) inputs, a scalar output (readout), and computational circuitry which includes mushroom body output neurons (MBONs), dopaminergic neurons (DANs), and feedback neurons (FBNs). A schematic of this model can be seen in Figure 1. Odour inputs represented by the KC activities feed into the MBONs through a weight matrix $W_{KC\rightarrow MBON}$. Whether these odours are apetitive, aversive or neutral (i.e. their *valence*) is determined by an external context signal, which feeds into the FBNs through a weight matrix W_{ext}. The output circuitry consists of a network of the MBON, DAN and FBN neuron populations. These neurons are fully connected through a weight matrix W_{recur}, with the DAN to MBON connections set to zero. This ensures the DANs affect the MBON activity only indirectly, through plasticity-gating of the $KC \rightarrow MBON$ weights.

The model is a recurrent neural network (RNN) constructed in PyTorch, whose weights are trained in two ways.

1. W_{ext}, $W_{readout}$ and W_{recur} are trained over epochs using backpropagation of the readout loss.
2. $W_{KC\rightarrow MBON}$ are dynamically updated through a plasticity learning rule.

The weights trained through backprop represent learning by evolution and over the lifetime of the animal, whereas the dynamically updated weights represent learning on a behavioural time scale. The dynamic weights

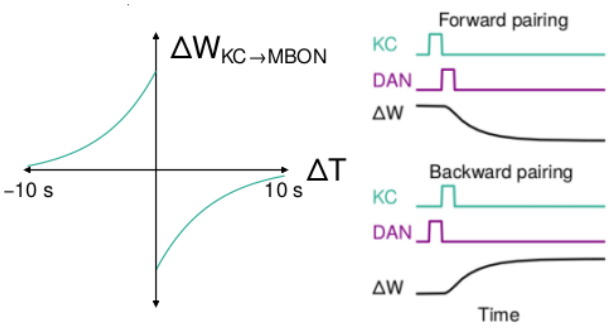

Figure 2: LTD/LTP order and timing (Jiang & Litwin-Kumar, 2020)

are updated through long-term depression/potentiation (LTD/LTP) rules. As seen in Figure 2, the DANs gate learning in the presence of KC inputs. This means when an odour is presented to the Drosophila (i.e. KCs are activated), if there is also DAN activity before (potentiation) or after (depression), a memory of that odour will be formed in the $KC \rightarrow MBON$ weights.

Paper Tasks and Results

Jiang and Litwin-Kumar train this model to perform several tasks. This paper considers three of them:

1. Classical conditioning with plasticity, including first-order conditioning, extinction conditioning, and second-order conditioning.
2. Classical conditioning without plasticity.
3. Continual learning, in which the weights are not reset between trials and multiple stimulus presentations occur in a single trial.

Classical Conditioning with Plasticity

There are three different classical conditioning tasks investigated by Jiang and Litwin-Kumar, first-order conditioning, extinction and second-order conditioning. Each task trial involves a set of intervals, in which stimuli are presented. The neuron activities are reset between each interval, to prevent memory through sustained activity, and the weights are reset between trials during training.

First-order conditioning is the most well-known version of classical conditioning, where an unconditioned stimulus (US) is used to condition a neutral stimulus, which is then said to be *conditioned* (CS). One trial consists of a conditioning and a test interval. Extinction adds a third interval to the first-order conditioning task, where the response should be reduced by presentation of the CS without the presence of the US. Second-order conditioning involves conditioning one odour first (CS1), then using that, in place of an US, to condition a second odour (CS2). Examples of these tasks can be seen in Figure 3.

These tasks were replicated using a full version of the Jiang and Litwin-Kumar model. The results can be seen in Figure 4. It is clear that the model qualitatively reproduces the results of (Jiang & Litwin-Kumar, 2020). Statistical comparisons were excluded due to time constraints, and because the ultimate goal was to reduce the model.

Classical Conditioning without Plasticity

Jiang and Litwin-Kumar compared the results of their dynamically updated model with those of an equivalent model, with no plasticity mechanism. In this case, the $KC \rightarrow MBON$ weights are initialized and set throughout the duration of the training. The goal is to determine whether plasticity is necessary to perform a conditioning task. Since this greatly reduces the models ability to form multiple associations, a limited set of odours is used throughout training and testing. The effects of using one vs. ten odours (two vs. twenty associations) is illustrated in Figure 5. As can be seen, the ability of the network to store associations decreases with the number of associations.

Jiang and Litwin-Kumar found that the model was not able to discriminate between conditioned and novel odours for more than a few associations. However, the reproduced model performed fairly well up to ten odours. It is not clear why this occurred. However, it is clear that the model's performance drops with the number of associations stored. Furthermore, the number of odours used to train and test the model without plasticity is approximately three orders of magnitude below that of the model trained with plasticity, since a new set of odours are generated for each training epoch in that case (5000 training epochs). Therefore, it can be concluded that plasticity is necessary to perform the tasks outlined in the previous section, which is what we wish to show.

Continual Learning with Plasticity

The final task reproduced is that of continual learning. In this case, a task trial consisted of a single interval, with multiple presentations of four different odours. Two of the odours were conditioned, one was given a positive valence (CS1+), while the other a negative valence (CS2+). The other two odours were neutral (CS-). The presentation times for these stimuli were drawn from a Poisson distribution. Neuron activities were again reset between trials. However, during training of the continual networks weights were initialized with a function of the values from the end of the previous trial. The weights were annealed to prevent early saturation. In the initial trial all weights had the same value. The influence of the previous trial's weights increased linearly until the midpoint of training, from which point the weights were passed directly for the remainder of epochs (see Eq. 12 from (Jiang & Litwin-Kumar, 2020)).

Additionally, Jiang and Litwin-Kumar included a homeostasis mechanism, which they call *non-specific potentiation*, intended to prevent mass synaptic depression after many associations were formed (see Figure 6). This acts

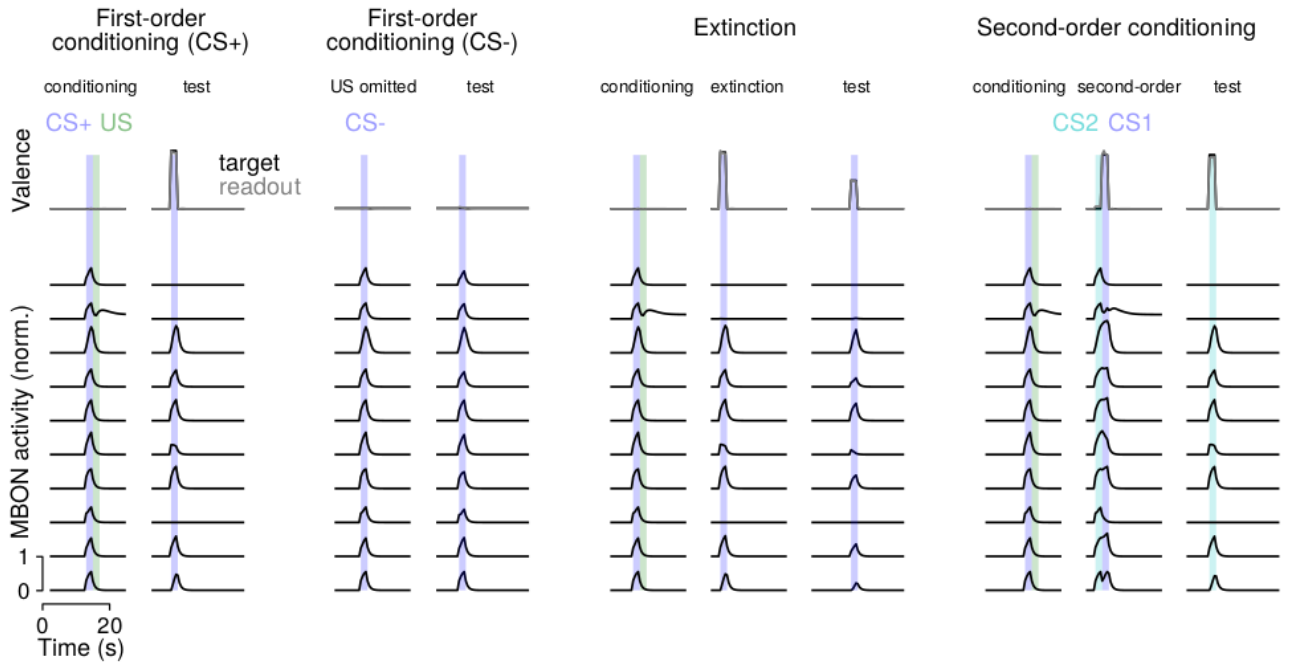

Figure 3: Classical conditioning tasks from (Jiang & Litwin-Kumar, 2020)

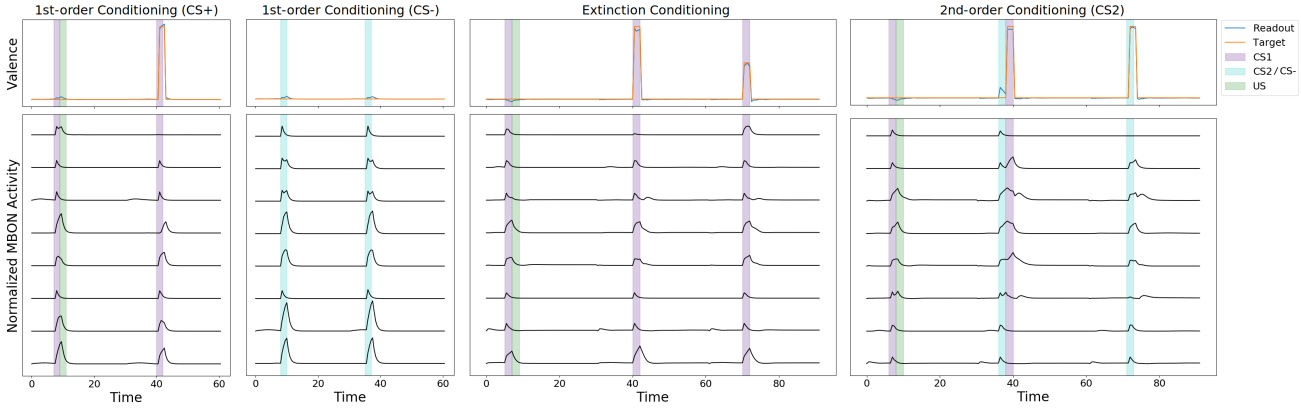

Figure 4: Reproduced classical conditioning tasks

as a form of forgetting, to prevent wide-spread synaptic depression (Jiang & Litwin-Kumar, 2020). The effects of this feature are compared in Figure 7.

It can be seen in the results from (Jiang & Litwin-Kumar, 2020) that non-specific potentiation appears to have a significant qualitative impact on the ability of the network to perform the task. However, this was not observed in the reproduced network. In fact, the trial shown in Figure 7 (bottom right) is one of the worse trials. On average, most trials looked similar to those with non-specific potentiation, and it was difficult to tell them apart in most cases. Again, it is not clear why there is this difference.

Model Reduction

The next step involved systematically reducing the Jiang and Litwin-Kumar model, to find a minimum version that could be analyzed more easily. Several different versions of the network were generated, and these were compared by their ability to perform the three classical conditioning tasks (first-order, second-order and extinction). Two types of modifications were considered in this reduction:

1. The complexity of the model.

2. The number of MBONs (and therefore DANs) in the system.

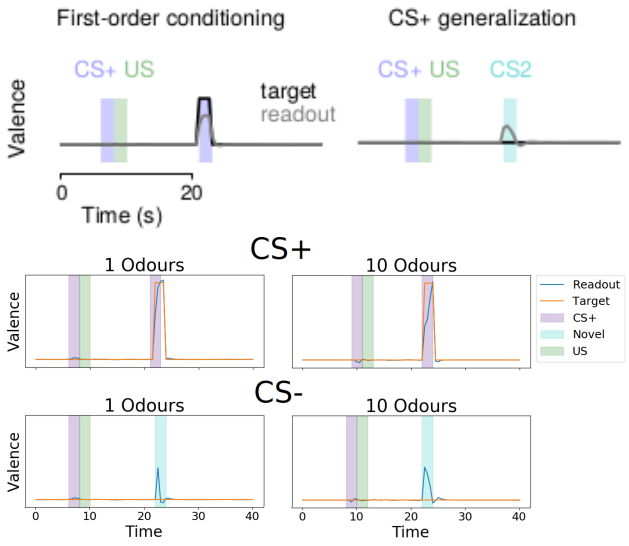

Figure 5: Classical conditioning tasks without plasticity, Top: from (Jiang & Litwin-Kumar, 2020). Bottom: reproduction.

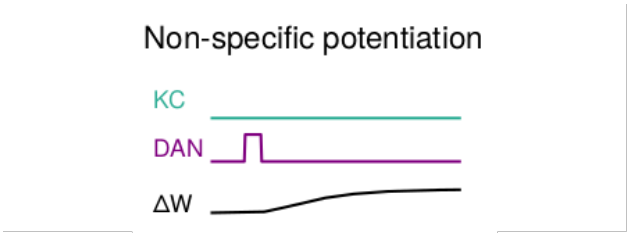

Figure 6: Non-specific Potentiation from (Jiang & Litwin-Kumar, 2020)

Output Circuitry Feedback Effects

The output circuitry consists of the MBONs, DANs and FBNs. To see which of these connections were necessary to perform the task, the output circuitry was reduced in stages. This was done by altering the connectivity matrix among the MBON, FBN and DAN populations. In the model of Jiang and Litwin-Kumar, only the DAN to MBON connections are set to zero, with the others being learned through training. The first stage of reduction involved removing all recurrent connections through the FBNs. This left two paths between the MBONs and DANs, a direct MBON to DAN connection (1-hop), and an indirect (2-hop) connection passing through the FBNs. This configuration will henceforth be referred to as *two-hop connectivity*.

The next reduction involved removing the effects of the FBNs entirely. In this case, only the direct (1-hop) connection from MBONs to DANs remained. This is referred to as *one-hop connectivity*. The final stage was completely separating the MBONs and DANs, so that no pathways existed between them. This is called *disconnected connectivity*. In all cases, the DAN to MBON connections were set to zero. A visualization of these stages can be seen in Figure 8.

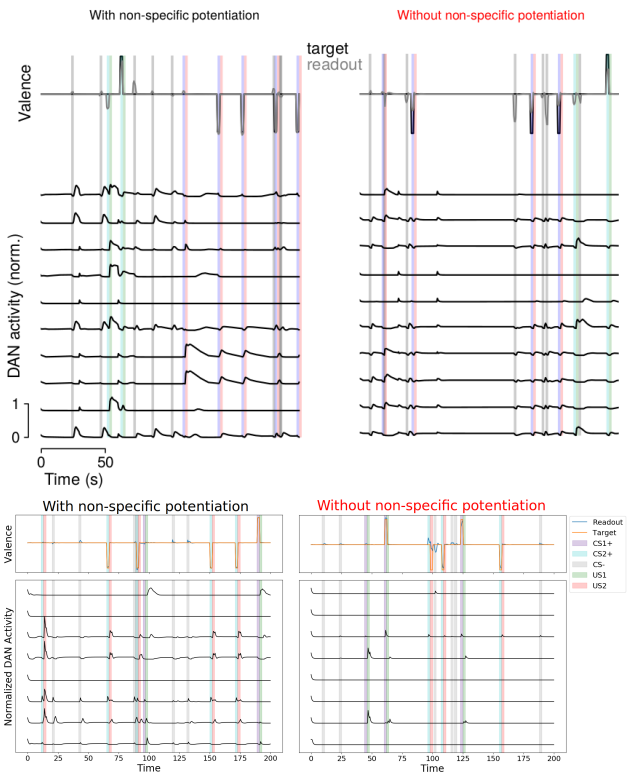

Figure 7: Continual learning task. Top: from (Jiang & Litwin-Kumar, 2020). Bottom: reproduction.

Training Model Only on Second-order Conditioning

The future work to be done with this model is focused on memory consolidation. We hypothesize that this is a form of second-order conditioning. Therefore, in order to simplify the model, to make it easier to analyze, we consider the effects of training the network only on first- and second-order conditioning tasks. In this case, there were no alterations to the network architecture itself; however, during training no extinction trials were presented to the model. These networks are referred to as *CS2 only*.

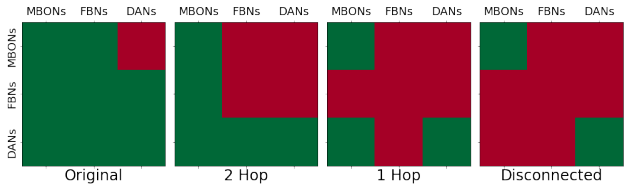

Figure 8: Representations of output circuitry connections during model reduction. Left: (Jiang & Litwin-Kumar, 2020) connectivity. Center left: '2-hop' connectivity. Center right: '1-hop' connectivity. Right: 'Disconnected' connectivity.

Removal of Long-term Potentiation

Based on the way the LTP/LTD rules are defined, the dynamics of the network, and the timing of stimulus presentations in the trials, it was thought that LTP may not significantly affect the network's ability to perform the task. Therefore, a version of the (Jiang & Litwin-Kumar, 2020) model, with only the LTD rule was created.

MBON Sensitivity

The MBONs and DANs are paired through the plasticity rule in a 1-to-1 configuration. Therefore, any changes to the number of MBONs also affected the number of DANs. A standard sensitivity study was performed on the number of MBONs in the model. First an exponential reduction in the number of MBONs was plotted, then a linear assessment was performed to refine the results. The minimum network was found when an increase in the number of MBONs had no clear effect on network performance.

This step of reduction was done following the reduction in model complexity. Since it was found that removing the FBNs from the system had little impact on the effectiveness of the network to perform the tasks (see Figure 10), the MBON sensitivity used the reduced one-hop connectivity network. This was done so that the combined effects of reduced complexity and number of MBONs would be seen directly, since the goal is to produce a minimal model with roughly the same performance as the original. The results of this study for the network trained only on 2nd-order conditioning tasks is seen in Figure 9. The results of the studies for the first-order and all classical conditioning trained networks are similar. It was found that the minimum number of MBONs required to perform all tasks at an MBON-independent level is ten. This is a qualitative assessment of the threshold, rather than a statistical one.

Reduction Results

Ultimately, it was found that the model did not require the FBN population at all to function, that LTP had minimal impact on performance, and that the model's performance was not seriously affected by reducing the number of MBONs to ten. The minimal model used in the next section includes LTP, but has only ten MBONs and one-hop connectivity. The comparison results are seen in Figure 10.

Imprint Memory Task

Once a minimal network was determined, it was possible to start performing new tasks using it. The task presented here is meant to mimic optogenetically imprinting an aversive memory into the Drosophila, then allowing that memory to consolidate over time. This experiment was performed as two consecutive trials, one to imprint the memory and a second to consolidate it. At the end of each trial, the MBONs were deactivated in sequence (from most negative to most positive readout weight) to determine the effect on the DAN activities. During deactivation, $W_{KC \to MBON}$ plasticity is turned off.

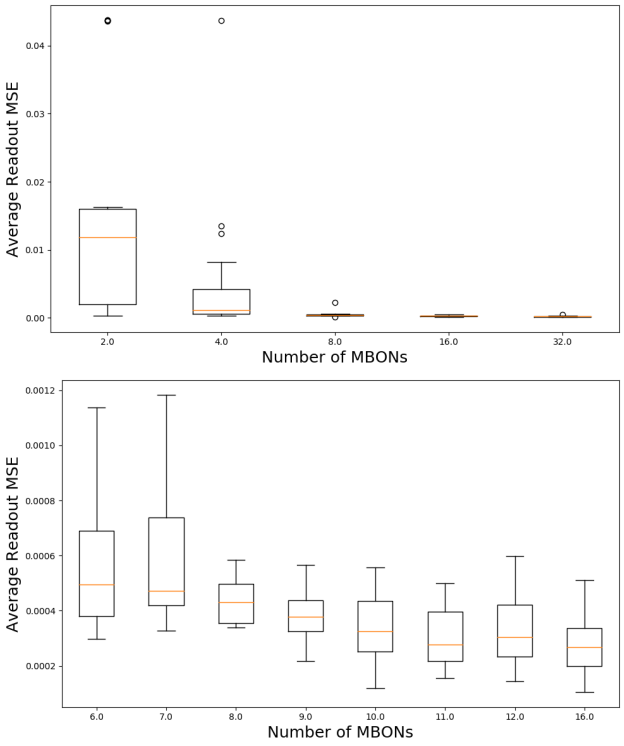

Figure 9: MBON sensitivity, CS2 training only

Since the behaviour of the model is determined by the readout weights $W_{readout}$, the sign and magnitude of these weights indicate how the corresponding MBON codes for aversive or appetitive behaviour and how strong that reaction is. Therefore, imprinting an aversive memory in simulation should require artificially activated the DAN corresponding to the most strongly negative readout weight, during presentation of a neutral odour. No US (context input) is presented with the odour, because the goal is to artificially control where the memory is created. See Figure 11 for an example of the trial and DAN activities.

After imprinting the memory, a novel stimulus was presented to activate the DANs, and the steady state activity of the DANs was recorded as each MBON was deactivated in sequence. The relative impact on steady state DAN activity of knocking out each MBON was then calculated. This effect was calculated as the difference of activity compared to when no MBONs were deactivated (control).

The consolidation trial consisted of a series of second-order conditioning intervals, where both CS1 and CS2 were the "optogenetically" conditioned odour (Figure 11. These trials activated the KCs and DANs such that the memory of the odour moved to a new location (i.e. the $KC \to MBON$ weights updated). Following this, the MBONs were again deactivated in sequence, a novel stimulus presented, and the relative steady state DAN activities were recorded.

The effects of imprinting and consolidating the memory on the $KC \to MBON$ weights is shown in Figure 12. We

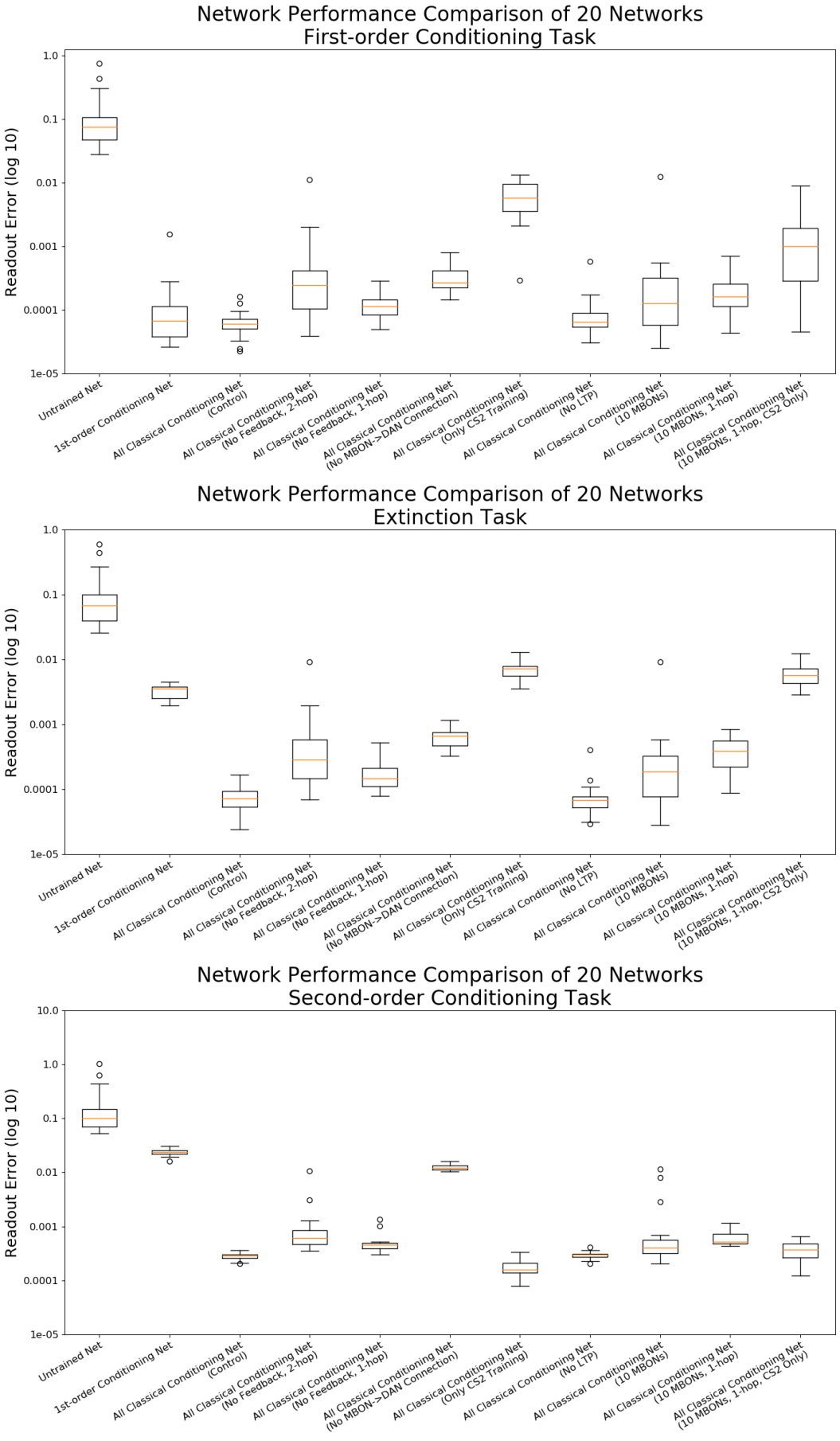

Figure 10: Model comparison. Top: first-order conditioning task. Middle: extinction conditioning task. Bottom: Second-order conditioning task.

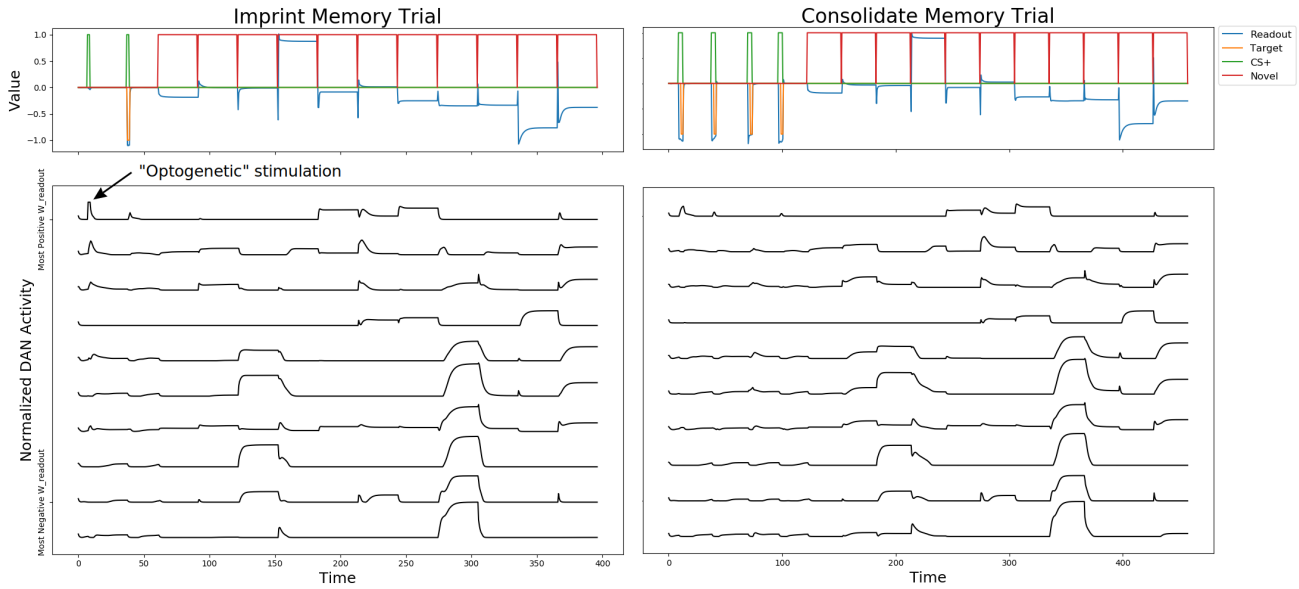

Figure 11: Memory imprint experiment, simulated trials

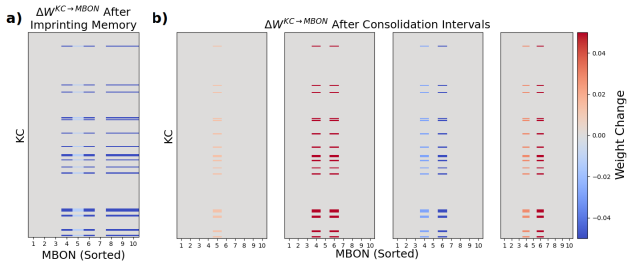

Figure 12: Memory imprint experiment, change in $KC \rightarrow MBON$ weights. a) Change in weights after imprint trial. b) Change in weights after each consolidation trial.

see here that the weights updated after both the memory imprinting trial, and the consolidation trials.

The difference of the relative steady state DAN activities is shown in Figure 13. It shows the difference in steady state DAN activity after the imprint task and after the consolidation task. The rows correspond to DANs sorted by the readout weights. The columns correspond to the effect on the activity when knocking out each MBON. The furthest left columns correspond to the MBONs with the most negative (aversive) readout weights, while the furthest right columns correspond to the MBONs with the most positive (appetitive) readout weights.

Discussion and Conclusion

The first part of this paper focuses on reproducing the results of (Jiang & Litwin-Kumar, 2020). Overall, the rerproduced results compare well. The no-plasticity version, and the continual learning model without non-specific potentiation both performed significantly better than the model of Jiang and Litwin-Kumar. It is not clear why this occurred.

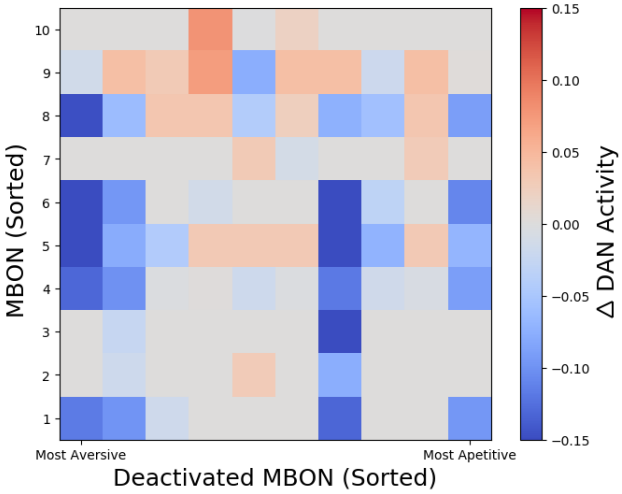

Figure 13: Memory imprint experiment, change in DAN activity after consolidation.

However, it is a little suspicious to us that the parts of their model they claim are necessary for it to perform seem to be less important than they claim. Since these are also the novel aspects of the Jiang and Litwin-Kumar model, there is an incentive for them to oversell their importance. Therefore, we believe this is something to be aware of when considering their results and claims.

Upon reducing the model, it was shown that the entire FBN population is redundant. All output circuitry connections through FBNs were removed, with minimal effect on the model's performance. Additionally, LTP rarely seems to be invoked, as removing this also had little effect on performance. However, this was left in, as removing it did

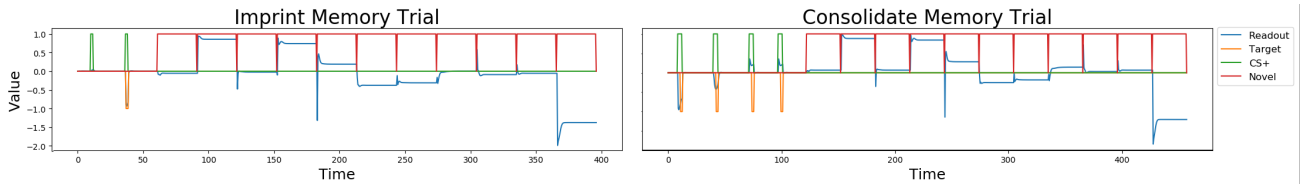

Figure 14: Memory imprint experiment, simulated trials

not simplify analysis of the model. The MBON sensitivity study illustrated that the model could perform nearly as well with half the original number of MBON-DAN pairs. All of these alterations combined to make the model much simpler to analyze. The final minimal model had no feedback neurons (FBNs), a one-hop connection from MBONs to DANs, and ten MBON-DAN pairs. It performs comparatively with the original (Jiang & Litwin-Kumar, 2020) model.

The imprint task is meant to mimic an optogenetic experiment that could be replicated by real lab conditions. Ideally, predictions made using this model could be verified experimentally. Imprinting the memory worked well, with one oddity. Intuitively, the DAN associated with the most negative readout weight should be activated to imprint an aversive memory. However, in practice, it was found that activating the most positive readout weight imprinted an aversive memory. It is not clear why this switch occurs. It may have to do with how the weights train, or it could be an error in the code that wasn't found.

Unfortunately, the consolidation step did not have the expected impact on DAN activity. The memory did usually strengthen during the consolidation intervals. However, the memory trace was not most strongly affected by deactivating the MBON used to imprint the original memory, as was expected. Additionally, the steady state DAN activity patterns did not significantly change after consolidation. On the other hand, the changes in the $KC \rightarrow MBON$ weights show that the memory did move during consolidation. Therefore, the lack of clear response to deactivating MBONs in the DAN activity could be due to the recurrent connections within the MBONs, or another mechanism that is not fully understood. More work would be necessary to properly assess the system dynamics.

Finally, there was a large variation in behaviour across networks. For example, in some networks, if the memory was imprinted with a relatively weak artificial DAN activity magnitude, the memory would flip from an aversive to an appetitive memory during consolidation (see Figure 14). Response variability would be expected in real animals, so it was interesting to see it in the simulated network as well. It presented challenges to the analysis, but perhaps also opportunity to develop methods which can be used to enhance the analysis of future experimental data.

References

Jiang, L., & Litwin-Kumar, A. (2020). Models of heterogeneous dopamine signaling in an insect learning and memory center. *bioRxiv*. doi: https://doi.org/10.1101/737064

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *14*, 1593-1599. doi: 10.1126/science.275.5306.1593