

Modelling dopamine modulated plasticity in the olfactory circuit of the fruit fly



Bayar Menzat

Green Templeton College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Michaelmas 2019

Abstract

In this thesis I investigate the role of dopamine modulated plasticity during olfactory learning in the fruit fly *Drosophila Melanogaster*. Although previous plasticity models explore many qualitative aspects of how dopamine orchestrates plasticity that also apply to the fruit fly, no computational model has reproduced recent learning experiments in the fruit fly to understand the link between dopamine modulated plasticity and behavioural changes. I address this problem by building a model that is based on the known anatomy of the fruit fly's olfactory circuit as well as recent behavioural and imaging data in three results chapters. In Chapter 2 I introduce a three layer spiking model based on the first layers of the olfactory circuit to study the transformations of dense spatio-temporal odour patterns into sparse representations in the fly. In Chapter 3 I introduce a model of the olfactory associative memory centre to study dopamine modulated learning rules.

I show that a learning rule that doesn't require postsynaptic spikes like in traditional models of plasticity is sufficient to reproduce learning experiments in the fruit fly. In Chapter 4 I investigate the role of potentiation with the goal of reproducing two experiments: relief learning and aversive re-learning. Relief learning is a type of experiment in which electrical shocks precedes the odour and the fly learns to approach that odour. The aversive re-learning experiment shows that a fly can learn to avoid an odour, reduce its avoidance and re-learns to avoid the odour in a flexible manner depending on the absence or presence of shock in consecutive trials. I show that both experiments can be explained by learning rules that allow for potentiation of synapses. Finally, I show that my model predicts that suppressing the activity of a select group of dopaminergic neurons during odour exposure and I propose experiments to validate my theoretical predictions. My results provide a new perspective on dopamine modulated learning in the fruit fly and suggest that dopamine orchestrates both potentiation and depression during plasticity.

Contents

1	Introduction	1
1.1	Neuroscience of reinforcement learning	1
1.1.1	Plasticity between neurons	2
1.1.2	Role of neuromodulators in plasticity	3
1.1.3	Reinforcement learning - the theoretical perspective	5
1.1.4	Izhikevich - Solving the distal reward problem	6
1.1.5	Drosophila Melanogaster, an ideal candidate to study dopamine mediated plasticity?	7
1.2	Introduction to learning in the fruit fly	7
1.2.1	Classical conditioning in the fruit fly	8
1.3	Learning paradigms in the fruit fly: learning, extinction, re-learning and relief learning	10
1.4	Anatomy of the fruit fly olfactory system	11
1.4.1	Early stages of odour processing in the fruit fly brain	12
1.4.2	The Mushroom Body is the centre for associative learning in the fruit fly	14
1.4.3	Anatomical subdivisions of the MB have unique role in odour processing	14
1.4.4	Mushroom body APL mediated recurrent activity stabilizes odour memories	15
1.5	Mushroom Body Output Neurons guide approach and retreat behaviours	16
1.6	Dopaminergic neurons convey reward and punishment signal	18
1.6.1	DAs can control memory expression based on animal's state .	19
1.6.2	DA modulated learning changes KC-MBON synapses	20
1.6.3	DANs encoding opposite values are required for extinction .	20
1.7	Spike Timing Dependent Plasticity (STDP) a candidate for olfactory plasticity	21

1.8	Abstract model of learning in the Mushroom Body	22
1.9	A phenomenological model of the neuron	22
1.10	Previous computation models of olfactory processing and learning	23
1.11	Questions addressed in my thesis	25
2	Sparse decorrelated odour representation in the mushroom body	26
2.1	Ch 2: Abstract	26
2.2	Introduction	27
2.3	Methods	28
2.3.1	Modelling ORN responses	29
2.3.2	Simulation conditions	30
2.3.3	Neuron model	31
2.3.4	Calculating separation between two odours	32
2.4	Results	33
2.4.1	Decorrelation of odour representation in the olfactory circuit	37
2.4.2	Studying APL mediated decorrelation using the Caron dataset	41
2.5	Discussion	43
3	Dopamine modulated aversive learning and extinction in the fruit fly	45
3.1	Ch3: Abstract	45
3.2	Introduction	46
3.3	Methods	48
3.3.1	MB Kenyon Cells model	50
3.3.2	Connectivity	50
3.3.3	MBONs and DANs	51
3.3.4	Learning rules	52
3.3.5	Rule 1: high-dopa-LTD rule	52
3.3.6	Choosing the threshold for the learning rule	54
3.3.7	Rule 2:R-STDP Rule	55
3.3.8	Learning rates	56
3.3.9	Valence and the performance index	57
3.3.10	Simulating reward and punishment	58
3.4	Results	59
3.4.1	Tuning a spiking olfactory model of the mushroom body	59
3.4.2	Dopa-pre-LTD rule can reproduce aversive learning and extinction experiments	60

3.4.3	Aversive extinction requires connections between opposing valence coding compartments	65
3.4.4	R-STDP rule can reproduce aversive extinction learning experiment	69
3.4.5	high-dopa-LTD rule can reproduce aversive learning experiment with no post spikes	70
3.4.6	Aversive re-learning decreases only MBON V2 firing rate	71
3.4.7	Second order conditioning forms two opposing memories	73
3.5	Discussion	74
4	Appetitive learning and extinction in the mushroom body	77
4.1	Ch4: Abstract	77
4.2	Introduction	77
4.3	Methods	79
4.4	Results	79
4.4.1	Appetitive learning and extinction	80
4.4.2	Extinction is blocked if reinforcement is present during re-exposure	85
4.4.3	Prediction: Blocking PAM- γ 5 dopaminergic neurons leads to an aversive memory	86
4.5	Discussion	87
5	Investigating the role of potentiation in the mushroom body	88
5.1	Ch 5: Abstract	88
5.2	Introduction	89
5.3	Methods	90
5.3.1	Rule 4:low-dopa-pre-LTP	90
5.3.2	Relief learning can be explained by potentiation of KC-MVP2 synapses	92
5.3.3	Re-learning with LTP enables re-extinction	94
5.3.4	Re-learning in the four layer olfactory circuit	98
5.4	Discussion	100
6	Conclusions and future work	102

List of Figures

1.1	Appetitive and aversive training protocols in the fruit fly	9
1.2	Experimental Performance Index for four different experiments.	11
1.3	Diagram of the complete olfactory circuit from odour detection to motor biasing neurons	13
1.4	Schematic of the circuits within the MB lobes.	15
2.1	Fruit fly olfactory circuit model diagram	29
2.2	Response curves for 16 ORN types to 110 odours (Hallem and Carlson, 2006)	30
2.3	Average number of spikes across all odours and receptors	34
2.4	Network simulation of a single odour	35
2.5	Network simulation of a single odour	35
2.6	Network activity during presentation of four different odours.	36
2.7	ORN, PN and KC responses to odours in the Hallem and Carlson, 2006 data set	37
2.8	Lateral inhibition facilitates pattern decorrelation in the antennal lobe.	39
2.9	APL mediated feedback inhibition facilitates pattern decorrelation in the MB.	40
2.10	Average angular distance in PNs and KCs as a function of lateral inhibition.	41
2.11	APL mediated feedback inhibition facilitates pattern decorrelation in the MB.	42
2.12	KC-APL mediated inhibition increases angular separation between odours and decreases number of responding KCs per odour	43
2.13	Network responses to odours in each layer	43
3.1	The circuit diagram of the mushroom body illustrating Kenyon Cells (KCs) , Mushroom Body output neurons (MBONs) and Dopaminergic neurons (DANs).	49

3.2	Learning rules: Rule 1 (dopa-pre-LTD)	53
3.3	Learning Rule 2 (R-STDP)	55
3.4	Calculating odour valence based on MBON firing rates elicited by the odour.	58
3.5	Tuning KC-MBON connections to yield equal firing rates for retreat and approach MBONs	59
3.6	Background dopamine and presynaptic levels do not change KC-MBON synapses.	60
3.7	Network activity during aversive training using rule 1	61
3.8	MBON Firing rates after learning protocol.	62
3.9	Changes in the MBON firing rates as number of shocks during training are increased	63
3.10	Changes in MBON MVP2 firing rates depending on timing of shock relative to odour onset.	64
3.11	Rule 1: Network activity during extinction learning protocol.	66
3.12	MBON Firing rates after aversive extinction protocol.	67
3.13	M6 and PAM- γ 5 neurons after aversive learning and aversive extinction	68
3.14	Changes in MBON M6 firing rate as number of extinction trials is increased.	68
3.15	Network activity during appetitive extinction simulation with R-STDP learning rule	70
3.16	Network activity during aversive training with no postsynaptic spikes	71
3.17	Network activity during aversive re-learning protocol	72
3.18	MBON Firing rates after appetitive extinction protocol.	73
3.19	Network activity during aversive re-learning protocol using rule 4: low-dopa LTP	74
4.1	Network activity during appetitive extinction protocol	80
4.2	MBON Firing rates after appetitive learning protocol.	81
4.3	Network activity during appetitive extinction protocol	82
4.4	MBON Firing rates after appetitive extinction protocol.	83
4.5	Neurons after aversive learning and aversive extinction	84
4.6	Appetitive extinction with different learning rates.	84
4.7	Rule 1 (high-dopa-LTD): Network activity during failed extinction learning protocol.	85
4.8	Suppressing PAM- γ 5 DANs leads to the formation of an aversive memory.	86

5.1	Learning rule low-dopa-LTP	92
5.2	Shock learning result depends on relative timing of shock and odour onset when using rule 3.	93
5.3	Rule low-dopa-LTP reproduces relief learning.	94
5.4	Network activity during aversive re-learning protocol using rule 4: low-dopa LTP	95
5.5	MBON Firing rates after aversive re-learning protocol.	96
5.6	Rule 4 potentiates MBON-MVP2 synapses when dopamine precedes odour	97
5.7	Network activity during aversive re-learning protocol using rule 4: low-dopa LTP	98
5.8	Learning, extinction, re-learning and re-extinction in the full network.	99

Chapter 1

Introduction

1.1 Neuroscience of reinforcement learning

The aim of the wise is not to secure pleasure, but to avoid pain.

Aristotle

Most animals, ranging from insects such as the fruit fly to human beings have brains that enable them to process sensory information and make decisions based on information they gather from the environment. The desire to avoid pain and search for pleasure is a defining characteristic of all animals. In the history of philosophy pain has been viewed with suspicion, Aristotle declaring that the aim of the wise is to avoid pain. But in the 18th century there were two great philosophers who had opposite views on the matter. Marquis de Sade believed that it is by way of pain that we arrive at pleasure, while Jeremy Bentham declared that the best course of action is one that maximizes pleasure. Although the feelings of pain is difficult to measure scientifically, especially in animals that are incapable of communicating, modern science can focus on the animal's behaviour: how will an animal respond to pain, or to the predictors of pain such as co-active stimuli? In fact this paradigm is so prevalent today that experiments on fruit flies and mice measure successful learning of a painful stimulus by the animal's ability to avoid it. Pleasure and pain are seen by science as reinforcers which drive learning, motivate and alter behaviour. The hypothesis stating the existence of internal reinforcer signals date back to [Olds and Milner, 1954](#). In these early experiments rats had areas of their hippocampus implanted with electrodes which they could activate using a lever in their cage. Researchers found that rats would develop an addiction to self-stimulation. Importantly, not every placement

of an electrode in the rat's brain led to self-stimulation: the most successful areas were situated in the basal ganglia: the ventral tegmental area (VTA) and substantia nigra pars compacta (SNc). These areas are known to contain populations of dopaminergic neurons. Dopamine is an example of a neuromodulator, a type of cell that can effect the ongoing spiking activity of one or many neurons. The activity of dopaminergic neurons is believed to affect behaviour and it has been linked to behavioural experiments such as Pavlovian conditioning ([Pavlov, 1927](#)). Pavlovian conditioning is a type of experiment in which a conditioned response such as salivation that normally occurs when food is seen by an animal can also be triggered by the stimulus that predicts the food. Dopaminergic neurons from VTA that were studied by [Schultz et al., 1997](#) were found to have similar temporal dynamics to Pavlovian conditioning by increasing their activity at the onset of the predictive stimulus. In the absence of reward dopaminergic neurons have been found to fire below their baseline level. This has led to the idea that dopamine signals the prediction error of the animal, an idea which was formalized in the following way:

$$\text{Dopamine Response} = \text{Reward occurred} - \text{Reward predicted} \quad (1.1)$$

While the experiments described above have found a link between dopaminergic activity and reward motivated behaviour, one of the fundamental question in neuroscience is understanding what is the role of dopamine during learning at the levels of neurons. At cellular level, the nervous system receives external input, processes it and generates behaviour. All neurons are capable of receiving input, processing and generating a response. The computation that a single neuron can perform is highly limited. This limitation can be overcome by combining many neurons in a network with different connectivity patterns, which can be changed through plasticity.

1.1.1 Plasticity between neurons

The most famous postulate that describes the relation between plasticity and learning was advanced by Hebb in 1949: "when an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in on or both cells such that A's efficiency, as on of the cells firing B, is increased" ([Hebb, 1949](#)). In the first experiment that confirmed Hebbian plasticity it was shown that persistent changes can be introduced in the hippocampus by high frequency stimulation protocols, in a phenomenon called long term potentiation (LTP). This is a mechanism for strengthening synaptic connections.

Later, a mechanism for weakening the synapse was discovered which is called long-term depotentiation (LTD). Today, different forms of Hebbian plasticity are believed to play a central role in learning as it has been found to be ubiquitous in the brain (Bliss et al., 1973, Markram et al., 1997, Bi and Poo, 1998). The next question posed by researchers was what is the role of spikes in the induction of LTP and LTD. In a theoretical study, it was suggested that presynaptic and postsynaptic spikes should determine whether synapses will undergo LTP or LTD. Experiments showed that this was indeed the case, first in vitro (Bi and Poo, 1998, Markram et al., 1997) and then in-vivo in the locust brain (Cassenaer and Laurent, 2012). The mechanism that modifies synapses based on the timing of spikes is called spike timing dependent plasticity (STDP). While the discovery of STDP opened many new avenues for studying plasticity in a paradigm called unsupervised learning, it became clear that pre and post spikes are not sufficient to explain learning, and neuromodulation played an important role in STDP (Pawlak et al., 2010).

1.1.2 Role of neuromodulators in plasticity

Neuromodulation is the process by which a neurotransmitter such as dopamine is released into extracellular space and affects the spiking activity of one or more neurons as opposed to a classical neurotransmitter which when released only affects a single postsynaptic partner. Dopamine is one such neuromodulator but there are many others such as noradrenaline or serotonin, each believed to play a role that affects animal behaviour in different ways. One possible explanation behind how neuromodulators affect behaviour is through the ion channel-specific effects. Small changes to a specific type of ion channels have been shown to be able to dramatically alter circuit behaviour (Goldman et al., 2001). In this situation a single neuromodulator that diffuses across the brain through volume transmission could have a coherent reliable effect each time. These discoveries inspired researchers to study the link between plasticity and dopamine. Corticostriatal neurons have been a popular target for studying the role of dopamine. In a landmark experiment, researchers studied slices of corticostriatal neurons in vitro from an area that contains nigrostriatal dopaminergic neurons. High frequency extracellular stimulation of corticostriatal neurons coupled with intracellular current injection led to LTD of the postsynaptic potential measured. Intriguingly, LTD was blocked when dopaminergic receptors were pharmacologically blocked or when DA neurons were lesioned (Calabresi et al., 1992). In a different experiment researchers managed to induce LTP of corticostriatal synapses and block

LTP again by dopamine receptor blocking. The results are difficult to interpret, however it was proposed that high dopamine levels could cause LTP, while intermediate cause LTD and absence of dopamine cause no changes at all (Reynolds and Wickens, 2002). A different study found a more direct mechanism of the interplay of STDP and dopamine: spike-timing-dependent LTP required D2 dopamine receptors and dopamine modulation led to suppression of GABAergic inhibition which led to more postsynaptic spikes and LTP (Bissière et al., 2003). However, when GABAergic activity was entirely suppressed pharmacologically, STDP was blocked even in the presence of dopamine suggesting the mechanism of interaction is complex.

In another *in vitro* study in a different brain area cell cultures with no dopaminergic cells were found to exhibit STDP (pre- before post LTP, post before pre- LTD) like plasticity (Zhang et al., 2009). Adding dopamine extended the time window for LTP and converted LTD into LTP.

The first *in-vivo* study of the interaction between STDP and dopamine was performed in the locust. Researchers injected current into so called odour coding Kenyon Cells (KCs) and their postsynaptic partners and found that they could reproduce the STDP time window. Adding octopamine, which in the locust is believed to signal reward, led to the change of LTP into LTD.

For theoreticians the interplay between dopamine and STDP has been hard to explain based on the experimental evidence due to the fact that the effect of dopamine on plasticity appears to be different depending on brain area. For example in the fruit fly in an experiment that optogenetically activated dopaminergic neurons while exposing the fly to an odour, plasticity was observed while postsynaptic neurons were blocked Hige et al. (2015a). This suggests that synaptic depression does not require STDP for plasticity in one brain area of the fruit fly. While multiple studies have shown that dopamine is required for plasticity to take effect at the synapse, the mechanism is not well understood. It's unclear whether dopamine concentration gates STDP or if dopamine concentration minus the baseline modulates plasticity bi-directionally depending whether the value is positive or negative. Furthermore, it remains unanswered whether there is a linear relationship between dopamine concentration and change in synaptic weight. These questions inspired researchers to look into the theoretical field of reinforcement learning which studies strategies for an agent to learn through trial and error in an environment by using feedback from its own actions and experiences.

1.1.3 Reinforcement learning - the theoretical perspective

Reinforcement learning is concerned with studying what are the actions an agent should take in order to maximize his reward. In a theoretical framework framed by [Sutton and Barto, 1998](#) the agent and the environment interact at discrete time steps $t = 1, 2, 3, \dots$. At each time step t the agent is able to sense the state of the environment $s_t \in S$ where S is the set of all states. In state s_t the agent selects an action $a_t \in A(s_t)$, from the set of actions possible in a state. In the next time step the agent will receive a reinforcement (positive or negative) based on the action taken represented by the numerical value $r_{t+1} \in R$ where R is the set of all possible reinforcements. The agent's action is based on a policy which is a function that maps a set of states to a set of actions

The policy π_t is the policy selected at time step t and $\pi(s, a)$ is the probability that $a_t = a$ if $s_t = s$. In the reinforcement learning framework algorithms are proposed to solve the problem of the search for the policy π which maximizes the expected accumulated utility. Solving the problem requires knowledge of the values associated with each state which in the real world is done by exploring the world. The trial-and-error process brings this algorithm closer to how animals learn. One such class of algorithms is the temporal difference (TD) algorithm:

$$V(s) \leftarrow V(s) + \alpha[r + \gamma V(s') - V(s)] \quad (1.2)$$

where V is a value function, $V(s)$ represents the agent's estimate for the value of being in state s . The estimate from the value of being in state s is updated based on the current value. α determines what the impact of new information will be. r is the reward for the action taken in the previous time step, and $V(s')$ is the estimate of that reward, so $r_t - V(s)$ is the difference between the actual and expected reward. γ is a discount factor used to discount expected future value to present value: $\gamma V(s)$ is the discounted expected value of the next state.

$$\text{TD error} = \text{reward at } t - \text{reward expected at } t + \text{discounted reward expected at } t+1 \quad (1.3)$$

A positive TD error tells us that the reward received was larger than expected and the action should be encouraged, while a negative TD error should lead to the decrease in the probability of taking that action. The formula is similar to the behaviour of dopaminergic neurons observed by [Schultz et al., 1997](#) with the exception of the future

expected reward term. The author observes that when an unexpected reward was received the dopaminergic firing rate was high for dopaminergic neurons. When an expected reward was received the firing rate remained unchanged while the absence of an expected reward led to the firing rate falling below normal. This similarity between neuroscience and theoretical neuroscience has prompted neuroscientist to search for other neural correlates of the TD algorithm. For example [Doya \(2002\)](#) suggests that "dopamine signals reward prediction error, serotonin determines the timescale of reward prediction, noradrenaline controls the randomness of action selection and acetylcholine the speed of memory update".

While theoreticians have been searching for neural correlates of the TD algorithm they have also struggled implementing the algorithm in a biologically realistic computational model. For example it's not clear how the brain implements a learning rule that is local which is to say uses information available at the synapse or how the brain knows which connections have to be modified when the reinforcement is delayed. The second problem is called the credit assignment problem and a solution has been proposed in a theoretical study that we will describe in the following section ([Izhikevich, 2007](#)).

1.1.4 Izhikevich - Solving the distal reward problem

In the Pavlovian experimental paradigm reward is delayed from the stimulus. In this section we discuss a model that seeks to solve the distal reward problem i.e how does the brain know which synapses are responsible for bringing the reinforcement? This problem is made difficult by the possibility that the firing patterns elicited by the stimulus are potentially gone by the time the reinforcement arrives. In [Izhikevich \(2007\)](#) the distal reward problem is approached by a framework that combines STDP and dopamine together. In this framework STDP provides both the sign and magnitude of the changes to the active synapses and is gated by extracellular concentration of dopamine. Extracellular concentration of dopamine is called d and changes over time as follows:

$$\frac{d}{dt}d = -\frac{d}{\tau_d} + DA(t) \quad (1.4)$$

where τ_d is a time constant of dopamine update which determines the rate of the reduction of the total amount of available dopamine. $DA(t)$ models the production of dopamine by dopaminergic neurons. $DA(t)$ is positive due to spontaneous firing of dopaminergic neurons. The strength of the synapse changes only when both dopamine

and an eligibility trace that is based on the STDP function are above zero:

$$\frac{d}{dt}w_{ij} = cd \quad (1.5)$$

Where w is the strength of the connection between neuron j and neuron i . d is the dopamine concentration. Higher firing rates of the dopaminergic neuron lead to higher concentration of dopamine and finally to more significant changes in synaptic strengths. Here, c represents the eligibility trace:

$$\frac{d}{dt}c = -c/\tau_c + STDP(t_{pre}, t_{post}) \quad (1.6)$$

and its value decays to $c = 0$ in an exponential manner, and τ_c is the variable that determines the rate of decay.

1.1.5 Drosophila Melanogaster, an ideal candidate to study dopamine mediated plasticity?

Thus far, we have presented an introduction into plasticity in the brain and recent theoretical and experimental discoveries that investigate the interaction between dopamine and plasticity. In this thesis I focused on modelling olfactory learning in the fruit fly, which I believe is an ideal candidate for studying dopamine mediated plasticity. Importantly, just like in human beings, dopamine signals reinforcement in the fruit fly in both aversive and appetitive conditioning paradigms ([Waddell, 2013](#)). The fly has been at the centre of recent discoveries that have unlocked neural circuit level understanding behind dopamine mediated learning ([Felsenberg et al., 2018](#)). Experimental tools allow scientists to study the role of individual neurons by silencing and activating them which can help us understand the link between plasticity and behaviour ([Aso et al., 2014](#)).

In the next section I present a review of olfactory associative learning and dopamine modulation in the fruit fly based on experiments and theoretical studies.

1.2 Introduction to learning in the fruit fly

The fruit fly *Drosophila Melanogaster* is a promising candidate animal for studying dopamine neuromodulation. In recent years the role of dopaminergic neurons has been uncovered in behavioural experiments that test the fly's ability to approach

or avoid an odour associated with punishment or reward. Classical conditioning is a widely studied form of associative learning in which a neutral stimulus (CS) becomes associated with positive or negative stimulus (US) through repeated pairing trials (Pavlov, 1927), thus rendering a conditioned response (CR). If the CS is presented multiple times without the US, the CR becomes attenuated in a process called extinction.

Since the 1970s it has been known that fruit flies are able to form associative memories (Quinn et al., 1974). Classical conditioning is tested in the flies by testing conditioned odour approach or avoidance in a T-maze after learning. At the level of the brain neural plasticity allows the fly to alter its behaviour as observed in behavioural experiments. Until now experimental studies have been unable to determine the precise form of plasticity that governs learning. Associative learning takes place in the area of the fly's brain called the mushroom body (MB), which is a large bilateral structure that receives predominantly olfactory information, but has also been found to receive visual, gustatory and mechanosensory information (de Belle and Heisenberg, 1994, Aso et al., 2014).

1.2.1 Classical conditioning in the fruit fly

Associative learning experiments in the fruit fly modify the Pavlovian conditioning paradigm. Rather than having a single stimulus (CS) the flies have to choose between two odours: the CS+ odour (paired with positive or negative reinforcement) and the CS- which is the odour unpaired with reinforcement and thus considered neutral (Fig. 1.1).

In the aversive training paradigm, flies are trained and tested in a T-maze apparatus which includes a training tube with an electric shock delivery system. Once the odour is paired with a shock (CS+) flies are transferred to the other side of the tube where they are exposed to the CS-. Finally, flies are moved to the test area where two symmetrical tubes can be approached from a choice area.

Flies can also form positive associations through reward training. In the reward training paradigm flies are introduced into two tubes each filled with either 1-octanol or MCH odours, and one of the odour exposure tubes is also coated in sucrose to create a positive association. Experiments calculate a so called performance index which is a behavioural metric used as a correlate of memory strength and represents the normalized preference of the trained odour compared to the untrained one. Aversive and appetitive training have similar scores and memory retention. Olfactory memories

and the valence of the odours are believed to be stored in the fly's brain through plasticity and these memories can be retrieved during behavioural testing.

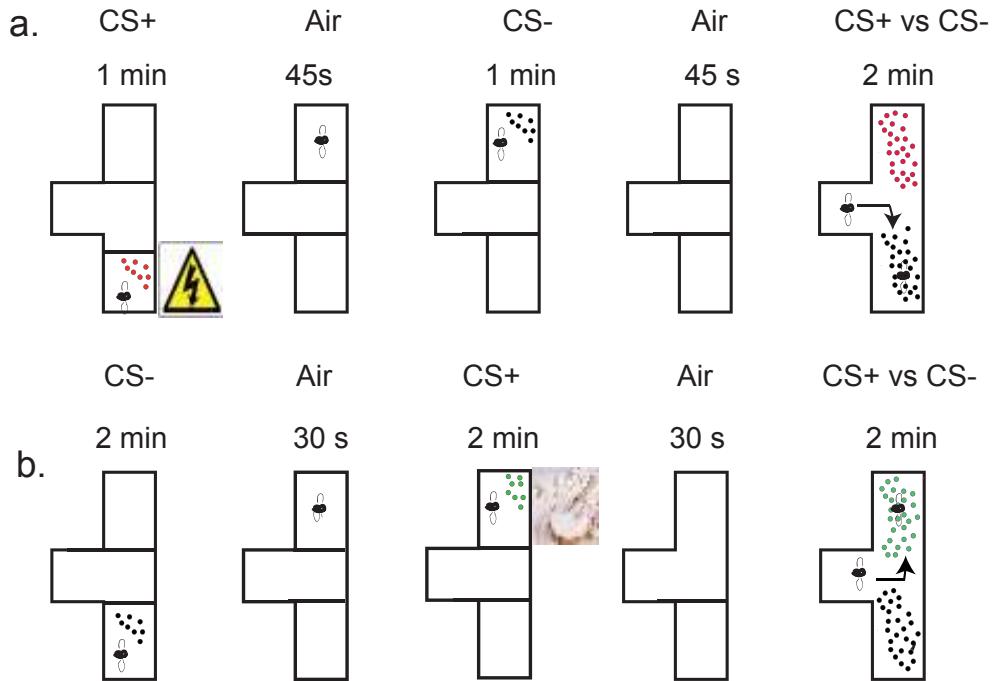


Figure 1.1: *Appetitive and aversive training protocols in the fruit fly* **a.** Flies are trained in a T-Maze to test aversive memory. Initially 100 flies are exposed to an odour while being shocked (CS+) for 60 seconds. After 30 seconds of air they are exposed to a second odour (CS-) which is not paired with reinforcement. Next, the same group of flies are allowed to make a choice between the CS+ and the CS- which results in flies retreating from the CS+ odour which was paired with shock and an approach behaviour towards the neutral CS-. **b.** Appetitive learning. In contrast to the aversive training protocol flies are initially exposed to the neutral odour (CS-) for 2 minutes. After 30 seconds of air flies are exposed to the CS+ odour in a chamber filled with sucrose (reward). After exposure to the second odour flies are tasked to choosing between the CS+ and CS- which results in an approach behaviour towards the CS+.

1.3 Learning paradigms in the fruit fly: learning, extinction, re-learning and relief learning

In the previous section we have introduced the appetitive and aversive learning experiments in the fruit fly. Other experiments have shown that the fly is capable to extinguish aversive and appetitive memories when the odour is re-exposed to the fly without the expected reinforcement. Aversive extinction is a protocol in which a previously aversive odour is re-exposed multiple times after aversive learning without the expected shock which results in the attenuation of the aversive (Felsenberg et al., 2018) Additionally, after aversive extinction flies have been shown to be able to re-learn to avoid the extinguished odour if it is re-exposed in the presence of a shock which results in a higher avoidance level in a T-maze behaviour experiment.

Appetitive extinction is the protocol in which an odour that was previously associated with a sugar reward is re-exposed without the expected reward. This results in the complete extinction of the approach behaviour exhibited after appetitive learning (Felsenberg et al., 2017).

Intriguingly flies, can also learn to associate an odour with relief from punishment. In such an experiment flies are exposed to a shock before the presentation of an odour and when tested will prefer the relief signalling odour compared to the neutral CS- (Yarali et al., 2008), (Fig. 1.2).

Another example that showcases the ability of the fly to learn higher order learning is second-order conditioning (Tabone and de Belle, 2011). Second-order conditioning is a learning process which requires two steps: first an odour (CS1) is paired with a shock (US). The a second, untrained odour (CS2) is paired with the CS1 in the absence of shock, which results in the CS2 being associated with an avoidance behaviour. Second-order conditioning provides a mechanism by which animals can use information from other stimuli that can provide the teaching signal instead of a US such as punishment or reward.

Experiments that study learning in the fruit fly measure associative memories at the behavioural level by using a metric called the Performance Index. In each experiment 100 flies are trained and are subsequently presented a choice between the reinforced odour (CS+) and a neutral odour (CS-). The Performance Index measures the proportion of flies that choose CS+ vs. the CS. Performance Index can range from -1 (all flies avoid the CS+) to 1 (all flies approach CS+). Most experiments have measure values that range from -0.5 to 0.5 (Fig 1.2).

In Chapter 3 and 4 we reproduce performance index results calculated for learning and extinction extinction by introducing a transformation from the firing rates of neurons that guide behaviour to a model performance index. In the past couple of years progress have been made in understanding the anatomy of the olfactory circuit that allows the fly to learn to avoid and approach odours. In the next section we discuss the experimental evidence for which neurons are responsible for olfactory associative learning.

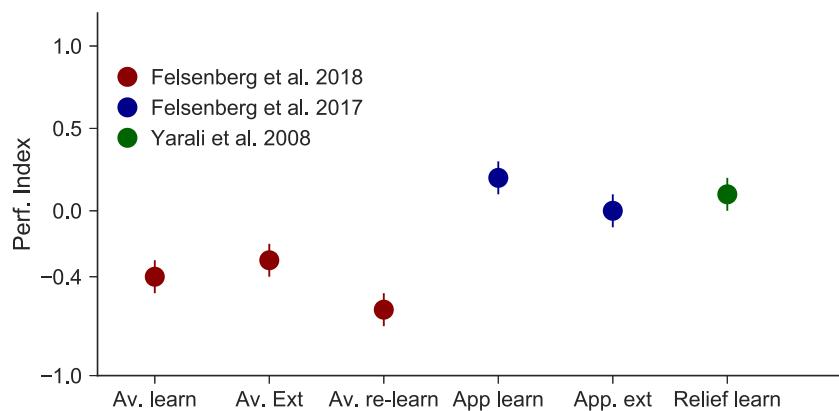


Figure 1.2: Experimental Performance Index for four different experiments. Performance Index calculates the proportion of flies that approach the reinforced CS+ odour compared to the CS-. Performance index from left to right is shown for the following experiments: aversive learning, aversive extinction protocol ([Felsenberg et al., 2018](#)) , appetitive learning, appetitive extinction ([Felsenberg et al., 2017](#)) and relief learning ([Yarali et al., 2008](#)) .

1.4 Anatomy of the fruit fly olfactory system

The mushroom body (MB) is the centre for associative learning in the fruit fly and receives olfactory information from the antennal lobe (AL) projection neurons (PNs) and valence information from both aversive and appetitive responding dopaminergic neurons (DANs) which converge at the output of the MB where mushroom body output neurons (MBONs) guide behaviour ([Aso et al., 2014](#)). We shall now discuss the staging or processing an odour from detection to behavioural output.

1.4.1 Early stages of odour processing in the fruit fly brain

Flies detect odours using approximately 1,200 olfactory receptor neurons (ORNs) housed in their antennae (Fig. 1.3). The tuning of each ORN is determined by a single odorant receptor gene, which controls whether the ORN is broadly or narrowly tuned to odours (Hallem and Carlson, 2006), which is to say whether it responds to a wide range of odours types or is selective. Axons from the ORN expressing the same receptor always converge on the same glomerulus in the antennal lobe (AL), which is made up of 51 glomeruli. The activation of multiple glomeruli in the AL allows for the discrimination of a broad range of odours. Each glomerulus receives an average input of 50 ORNs each connecting to approximately 3 PNs (Vosshall et al., 2000). The three projection neurons take input from each glomerulus deliver information to the calyces in the mushroom body while both inhibitory projection neurons and excitatory PNs send information to the Lateral Horn. The activity of ORN is picked up by inhibitory local neurons (LNs) and excitatory projection neurons (PNs) which are believed to perform normalisation of the responses of odours of different concentrations (Luo et al., 2010). The Lateral Horn is associated with innate behavioural responses to odours, such as sex-specific responses to pheromones during courtship (Datta et al., 2008), whereas the mushroom body is required for learned responses to odours (de Belle and Heisenberg, 1994). In contrast to PNs which respond to a wide range of odor through combinatorial activation of glomeruli, KCs display odour selectivity and their response to odours vary across individuals. Electron microscopy has shown that PNs converge randomly onto KCs (Eichler et al., 2017a) and each KC receives input from a range of 5-10 KCs.

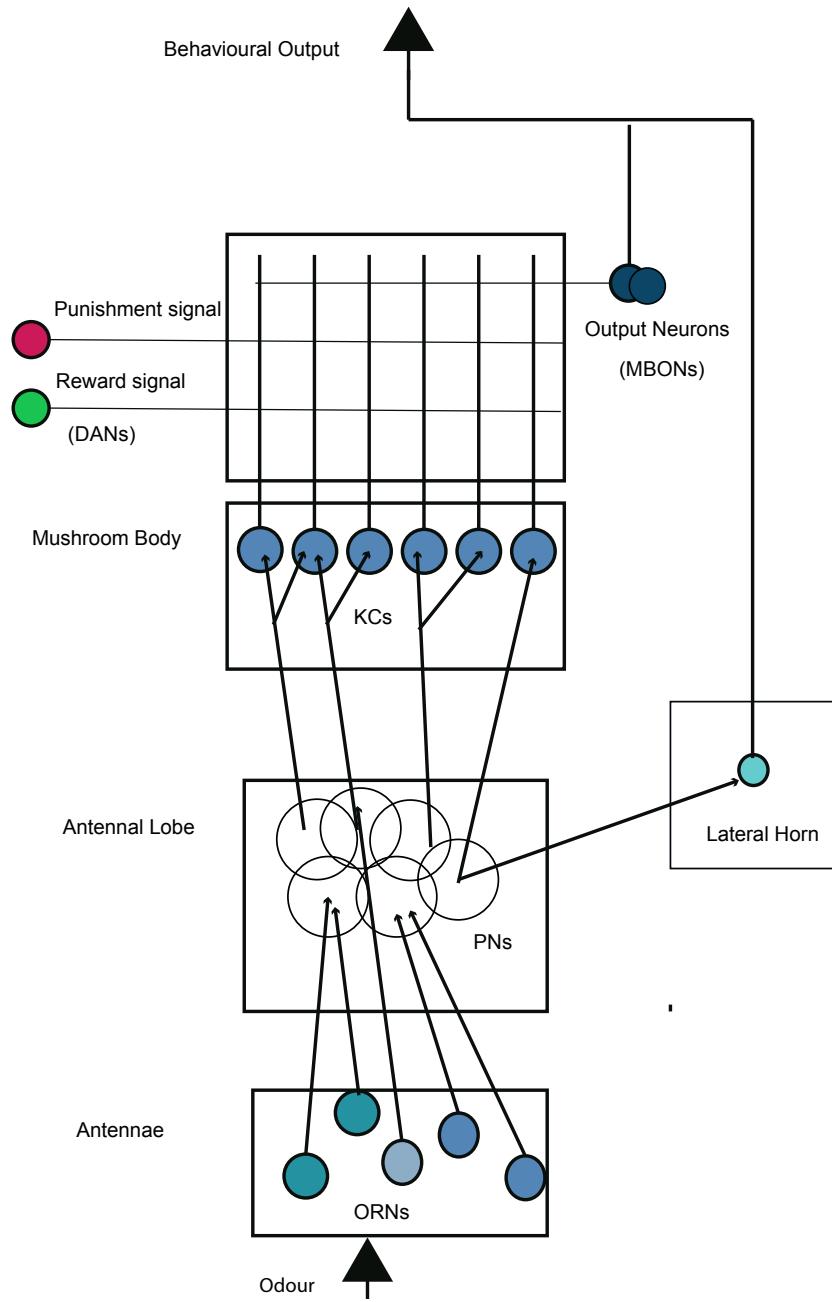


Figure 1.3: *Diagram of the olfactory circuit from odour detection to motor biasing neurons (Adapted from Dylla et al., 2017)*. Odours activate a subset of 1200 ORNs. ORNs expressing the same receptor converge to the same glomerulus in the antennal lobe. In the antennal lobe activity is determined by the interactions between ORNs, excitatory and inhibitory Local Neurons and projection neurons (PNs). 150 PNs converge randomly onto approximately 2000 Kenyon Cells (KCs) in the mushroom body. At the site of Kenyon Cells synapses modulatory dopaminergic neurons (DANs) signal punishment or reward. Co-activation of KCS and of DANs modify the activity of mushroom body output neurons (MBONs) which are KCs postsynaptic partners. Plasticity at KC-MBON site is required for learning odour specific approach and retreat behaviours.

1.4.2 The Mushroom Body is the centre for associative learning in the fruit fly

The mushroom bodies are symmetrical structures, with each mushroom body being composed of approximately 2,000 Kenyon Cells (Turner et al., 2008). All 2000 KCs extend their axons in two structures, the medial and vertical lobes of the mushroom body. Lesion experiments and synaptic blocking of mushroom bodies output have shown that the mushroom body plays a critical role in memory formation and retrieval (de Belle and Heisenberg, 1994). While odour coding in the antennae is dense, the MB is believed to maintain a sparse and high-dimensional odour representation with calcium imaging experiments showing that each odour activates approximately 10% of KCs (approximately 150 - 200 neurons in the mushroom body (Honegger et al., 2011)).

1.4.3 Anatomical subdivisions of the MB have unique role in odour processing

Each hemisphere of the MB is composed of approximately 2000 KC cell bodies (Aso et al., 2014). The Kenyon Cells project to three lobes: $\alpha\beta, \alpha'\beta'$ and the γ lobe. The lobes can be divided into 15 compartments: $\alpha 1 - 3, \beta 1 - 2, \alpha' 1 - 3, \beta'$ and $\gamma 1 - 5$ (Fig 1.4) The dendrites of the KCs form the calyx of the MB, and their axons project anteriorly to form the peduncle before those axons terminate in one or more lobes, termed $\alpha\beta, \alpha'\beta'$, and γ lobe (Krashes et al., 2007). KC axons in the α/α' lobes are clustered into anterior, middle, and posterior groups, in the β/β' lobes into posterior, surface, and core groups, and in the γ lobe into main and dorsal groups. The horizontal β lobe is ventrally adjacent to the γ lobe while the vertical lobe α extends to the dorsal end of the brain. Based on anatomical differences $\alpha\beta$ neurons can be further divided in $\alpha\beta$ core, surface and posterior neurons. Experimental evidence has shown these neurons fire on average at 2.2 Hz during odour presentation (Turner et al., 2008) The $\alpha'\beta'$ contains 600 KCs and again each KC has an axon in the horizontal and the vertical lobe. They have the highest baseline firing rate response to odours (4.9 Hz) (Turner et al., 2008). The γ lobe is formed by the anterior side of the horizontal lobe and consists of approximately 675 KCs which can be further subdivided in γ main and γ dorsal neurons (Aso et al., 2014) which also receives visual information. In a previous study, it has been shown by blocking different anatomical subdivisions of KCs, that the $\alpha\beta$ neurons are needed during both the retrieval of

aversive and appetitive memory, whereas $\alpha'\beta'$ neurons only affect appetitive memory (Perisse et al., 2013). Furthermore, it appears that within the $\alpha\beta$ lobe aversive and appetitive memories are represented by distinct non-overlapping KCs (Perisse et al., 2013).

The mushroom body can be divided in 32 anatomically distinct compartments. Each compartment is innervated by a distinct population of DANs and MBONs and are involved in different forms of associative learning that can be categorized by duration of memory (STM, MTM, LTM), type of behaviour they promote (approach or retreat) or type of reinforcement that drives plasticity (sugar, water, shock) (Fig. 1.4).

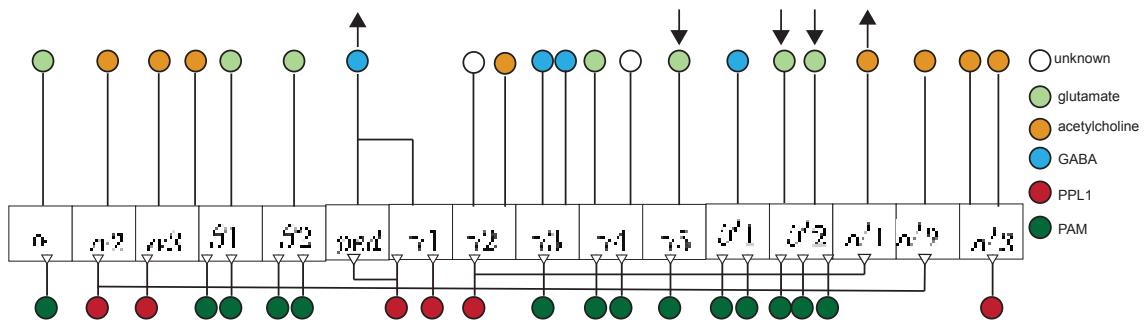


Figure 1.4: *Schematic of the circuits within the MB lobes.* Adapted from Aso et al., 2014 DAN inputs (bottom) and MBON outputs (top) from 16 MB lobe compartments are shown. Colours indicate neurotransmitter types for MBONs. Three DAN cell types innervate multiple lobe compartments. Eight types of MBONs receive input from more than one compartment. Aversive dopaminergic neurons (PPL1 red) and appetitive dopaminergic neurons (PAM, green) innervate unique compartments which are not overlapping. Different MB lobe has been attributed to different functions such as short term and long term memory and MBONs are believed to coordinate behaviour in relation to different phases of memory. Arrows are shown for four MBON types that we model in Chapter 3 denoting the behaviour they have been shown to promote when they are artificially activated during odour exposure.

1.4.4 Mushroom body APL mediated recurrent activity stabilizes odour memories

Experimental studies have shown that the mushroom body is a recurrent circuit. Olfactory representations in the mushroom bodies are modulated by the inhibitory anterior paired lateral (APL) neuron (Lin et al., 2014a). The mushroom bodies are

innervates by two symmetrical APL neurons which sends inhibition both to the calyx as well as the vertical and medial lobes of the MB. The KCs and APL neurons appear to form a negative feedback inhibitory circuit, with APL inhibiting the activity of the KCs while also receiving excitatory input from the KCs . The APL releases GABA at a rate proportional to its level of depolarization and appears to play a role normalizing KC responses to odours and maintaining sparseness ([Lin et al., 2014a](#)). In Chapter 2 we model APL mediated feedback inhibition and investigate the role it plays in facilitating sparse odour representation.

1.5 Mushroom Body Output Neurons guide approach and retreat behaviours

Outputs from approximately 2,000 Kenyon cells of the mushroom body converge onto a population of 32 mushroom body output neurons (MBONs). The MBON dendrites divide the two lobes and the base of the peduncle into a small number non-overlapping compartments, within which the dendrite densely contacts either all KCs or an anatomically distinct sub-population. The MBONs innervating each compartment are genetically distinct and each projects characteristically to other regions of the brain. MBONs project from the mushroom body to several regions. A small number of excitatory neurons innervating the α or α' lobes project to lateral horn, a region implicated in the control of innate odour-related behaviours. Every MBON projects to four neuropils just outside the mushroom body: the crepines, the superior medial protocerebrum (SMP) the superior intermediate protocerebrum and the superior lateral protocerebrum where input from the approach and retreat MBONs is believed to be transformed into motor control. The majority of MBONs respond to a large number of odours ([Hige et al., 2015b](#)) and in contrast to KCs which have a sparse activation pattern across odours, MBON odour tuning is highly correlated across odours ([Hige et al., 2015b](#)). Odour responses have been shown to be consistent between hemispheres but are very different across individual flies. This is consistent with the model that MBONs code for experience dependent learning. MBONs can be classified according to the behaviour they promote Two recent studies have shown that silencing and activating each of the 32 MBONs they can change the behaviour of the animal from approach to retreat or from retreat to approach depending on the MBON type ([Aso et al., 2014; Owald et al., 2015](#)). In this thesis we will refer to MBONs as either retreat or approach MBONs based on the behaviour

they promote. In the next sections we will summarize experimental evidence that highlights the behavioural role of individual MBONs which we model in Chapter 3 and 4.

MBON-M4/6 show bi-directional changes in response to training

The M4 and M6 MBONs are located in the γ lobe of the mushroom body in two different compartments. Calcium imaging experiments of the M4 and M6 MBONs before and after appetitive learning have shown bi-directional changes in the firing rate: their response to the trained odour decreases after reward learning and increased after aversive learning. We define bi-directional as the ability of an MBON to either increase or decrease its activity after learning, depending on the reinforcement used (appetitive or aversive). Furthermore, optogenetic activation of the M4/6 MBONs changed a neutral behaviour into avoidance behaviour, while blocking the neuron converted a naive odour approach into odour retreat in untrained flies i.e flies that had an innate repulsion to the odour without learning ([Owald et al., 2015](#)). Interestingly bi-directional change appears limited to retreat MBONs, as no experiment have shown approach MBONs can change in the same way. Activation of the M6 neuron has also been shown to excite the dopaminergic neuron innervating its own compartment. This is not the case for the M4 MBON. Furthermore, a recent aversive extinction experiment has shown that while aversive learning increases the activity of the M4/6 MBONs, aversive extinction decreases the activity of the M6 MBON while the M4 MBON remains unchanged ([Felsenberg et al., 2018](#)).

MVP2 MBON is GABAergic and inhibits other MBONs

The MVP2 MBON is also located in the γ lobe and is an inhibitory MBON which releases the GABA neurotransmitter ([Aso et al. 2014](#)). When the MVP2 MBON is activated during odour exposure it promotes approach behaviour. Recent experimental evidence has shown that the MVP2 MBON inhibits the M4 β' /6 retreat MBONs ([Owald et al., 2015](#)). Following aversive training, MVP2 exhibits a reduced response believed to be the result of plasticity. Thus, M4/6 MBON increase in activity after aversive learning is believed to be due to an indirect effect: MVP2 decreases in firing rate through dopamine modulated plasticity which results in the reduction of the inhibition onto M4/6 MBONs. An electrophysiological study has shown that the firing rate of the MVP2 decreases from approximately 100 Hz to 20 Hz after an odour is paired with the activation of the dopaminergic neuron in its compartment ([Hige](#)

[et al., 2015a](#)). We use this data to tune our plasticity model for aversive learning in Chapter 3.

MBON-V2 is required for both STM and LTM retrieval

MBON-V2 innervates the $\alpha 2$ and $\alpha' 3$ compartments and were found to be required for the retrieval of aversive STM and LTM after two different training protocols ([Séjourné et al., 2011](#)). Like in the case of other MBONs, odour activity was found to be depressed after aversive learning. Furthermore, evidence has shown that activation of the MBON-V2 MBONs triggers spikes in the dopaminergic neurons that modulate plasticity in its compartment suggesting an excitatory feedback loop similar to the case of the M4/6 neuron.

The four MBONs (M4, M6, MVP2 and V2) we have presented in this section are modelled in Chapter 3 and 4 based on experimental evidence of their connectivity to investigate how their activities pool together to determine the behaviour of the fruit fly and how plasticity changes their activities and ultimately the behaviour.

1.6 Dopaminergic neurons convey reward and punishment signal

Valence in the MB is signalled by dopaminergic neurons (DANs). Dopaminergic neurons (DAs) are the most prevalent modulatory neurons that innervate the mushroom bodies. There are distinct dopaminergic neurons that provide positive and negative value signals. The two major clusters where most DAs reside are the PPL and the PAM clusters. Positive reinforcement is provided by subsets of approximately one hundred DAs in the PAM cluster ([Burke et al., 2012](#)). During appetitive training sucrose in normally used which is both sweet and nutritious but water and pain relief has also been shown to activate positive reinforcement responding DANs. Until recently it has been believed that in marked difference to mammalian brains, in flies octopamine signals reward such as sugar taste. However, recent evidence has shown that dopaminergic receptors such as dDA1 is required for appetitive memory formation and that octopamine relay the presence of sugar through DANs that are in the PAM cluster. It is now believed that OA mediated reward signals activate PAM DANs to trigger reward based plasticity. When OA neurons are activated during training, memories have been shown to last no more than 3 hours ([Burke et al., 2012](#)) which suggests OA signals only sweetness. When flies are trained with arabinose, which

is a sweet but non nutritious sugar memories formed also last 3 hours. It has also been shown that OA-mediated reward reaches only compartments in the γ 4 and β' lobes. Conversely, long lasting nutritious reinforcement memories require activity in the α and β lobes. They predominantly innervate nearby zones in the β , β' , γ lobes. Odours can be also reinforced with waters when flies are thirsty (Lin et al., 2014b) Water learning requires PAM- γ 4 neurons for STM (short term memory) while LTM (long term memory) water memories depending on the activity of PAM- β' 1 neurons. Negative reinforcement such as from electric shock or bitter substances appears to be conveyed by DAs housed in the PPL cluster. Interestingly the PAM- γ 3 neurons have also been found to be involved in relaying aversive reinforcement.

The PPL DANs appear to innervate α or α' lobe where only 12 PPL1 neurons have been identified (Waddell, 2013). Different PPL1 neurons are involved in STM and LTM memory like in the case of appetitive learning. A recent paper (Cohn et al., 2015) has shown that during electric shock PPL1- γ 2 and PAM- γ 3 show increased level of calcium activity, while PAM- γ 4 and PAM- γ 5 were inhibited. Aversive reinforcement is not only limited to electric shock. Heat can also be used an US to train flies to avoid an odour CS (Galili et al., 2014). Bitter taste from substances such as quinine is signalled to the MB by the PPL1 DAN cluster. Thus flies learn to avoid dangerous stimuli associated with odours through the activity of PPL1 DANs inner-vating the MB vertical lobe as well as γ 1 and γ 2 and PAM- β 2 and PAM- γ 3. The dendrites of MBONs are restricted to few DA zones, which correspond to one type of reinforcement delivery (either aversive or appetitive). For example, in the case of the retreat MBON (M4 β'), experimental data exists to support that the axons from sugar rewarding dopaminergic neurons overlap with the dendrites of the MBON (Owald et al., 2015). Conversely, shock responding dopaminergic neurons overlap with the dendrites of approach MBONs (Yamagata et al., 2015). Thus, anatomical evidence suggests that reward DAs modulate the synapses onto retreat MBONs while punishment DAs modulate synapses onto approach MBONs.

1.6.1 DAs can control memory expression based on animal's state

Sugar memory is most robustly expressed in hungry flies and similarly, thirst promotes the expression of water memory. It appears that to form these memories, different zones of the MB provide the anatomical requirements in which state control can be implemented. For example, a group of DA neurons that provide the inhibitory

constraint to the expression of sugar memory (Lin et al., 2014a) occupy the same zones in the MB as the MVP2 MBON, whose activity drives approach behaviour. Thus it seems likely that the internal state of hunger skews the balance of MBON pathways to favour approach. In this vein, a current study has shown that the MVP2 MBON neuron increases its response to an odour after it has been starved which biases the behaviour towards approach. This evidence adds state modulation of behaviour as another layer of complexity to the valence learning model. There are also reinforcements that can activate both aversive and appetitive pathways.

Pairing an odour with ethanol induces short-term avoidance (30 minutes) while 24 hours later triggers an approach response (Kaun et al., 2011). A recent study has shown that a bitter nutritious food (Das et al., 2014) can activate separate PAM and PPL1 compartment and create either a weak aversive memory or a neutral memory depending on how bitter the food is.

1.6.2 DA modulated learning changes KC-MBON synapses

A hypothesis of how valence learning is achieved in the olfactory circuit proposes that learning modifies subsets of KC-MBON pathways which in naive flies are balanced (Owald et al., 2015). Appetitive responding DAs promote odour approach by depressing odour drive to avoidance promoting MBONs and possibly by strengthening approach pathways. Conversely, aversive responding DAs promote depressing odour drive to approach promoting MBONs (Aso et al., 2014). It is likely that a feedback loop from MBON to DAs influences learning as well. Blocking the M4 β' neuron induced a depression of the KC-M4 β' MBON connection (Owald et al., 2015), which could be explained by a scenario in which M4 β' inhibits some appetitive DAs.

1.6.3 DANs encoding opposite values are required for extinction

While PPL1 and PAM DANs have been attributed roles in aversive and appetitive learning respectively, recent evidence (Felsenberg et al., 2018) has shown that they each play a role in extinction. PPL1 shock responding DANs are required for appetitive memory extinction, while PAM sugar responding DANs are required for aversive memory extinction. This experimental evidence suggest that absence of expected reward is interpreted as punishment while absence of expected punishment is interpreted as reward.

1.7 Spike Timing Dependent Plasticity (STDP) a candidate for olfactory plasticity

A candidate mechanism for valence learning has been proposed in a recent study in the locust, a similar system to Drosophila ([Cassenaer and Laurent, 2012](#)). They have shown that KC-MBON synapses usually obey the spike-timing-dependent plasticity (STDP) learning rule. STDP can be seen as a spike-based variant of the Hebbian rule, where synaptic change depends on the relative timing of pre- and postsynaptic action potential ([Bi and Poo, 1998](#)). During STDP, a synapse potentiates when the postsynaptic spike happens after the presynaptic spike within a time window of the order of tens of milliseconds and depresses when the order is reversed ([Bi and Poo, 1998](#)). Plotting the change in synaptic efficacy against the relative timing of pre- and postsynaptic spike results in a bimodal efficacy curve. Spike timing, however, is not sufficient to explain how plasticity strengthens or weakens synapses. Neuromodulators such as dopamine can gate plasticity, determining when learning can take place ([Pawlak et al., 2010](#)). Indeed, in the locust STDP experiment when octopamine was injected 1 s after spike pairing, synapses were only observed to decrease their efficacy. In flies, octopamine is known to act on the mushroom body indirectly, by activating PAM dopaminergic neurons ([Waddell, 2013](#)). Therefore, evidence from studying the locust suggests the hypothesis of dopamine modulated STDP as the candidate mechanism for learning in the fruit fly. There are however problems with the STDP mechanism when applied to fruit flies. First of all the time-scale of STDP and olfactory conditioning is different by three orders of magnitude. Furthermore, experiments have shown that while Kenyon Cells (pre-) are necessary for flies to form olfactory memories the same can not be said for postsynaptic (MBON) spikes which are required in STDP. Experiments with the temperature sensitive dynamin transgene shibire has shown that blocking output synapses during training has no effect on memory acquisition and a recent experiment has shown that even by completely blocking MBON spikes during training, LTD still occurs, suggesting the learning mechanism has a component which is independent of postsynaptic spikes.

1.8 Abstract model of learning in the Mushroom Body

An increasing body of evidence suggests the activity of MBONs directly influence the Performance Index which measures the behavioural outcome of learning. Evidence has been accumulating from previous experiments that the Mushroom Body learns odour identity as sparse KC activity in the MB ([Honegger et al., 2011](#)). Odour identity information is subsequently relayed to different MB compartments, where neutral stimuli drive MB output pathways that bias for approach or retreat behaviours equally. Each of the 15 compartments codes for a specific quality of the US such as sweetness, nutritional value, water or is specialized in forming memories of different duration - Short Term Memory, Medium Term Memory, Long Term Memory ([Aso et al., 2014](#)). The role of MBONs is believed to read-out all the changes that have occurred during learning ([Hige et al., 2015b](#), [de Belle and Heisenberg, 1994](#)) through dopamine mediated plasticity. All though a great deal is now known from recent experiments, so far no mechanistic model has been put forward that can reproduce recent experiments and unifies our understanding of dopamine modulated plasticity in the fly. In Chapter 3 I present a simple model of dopamine modulated learning that can explain how changes in firing rates of MBONs can affect behaviour. We assume that an odour is initially neutral due to a balance of activity of approach and retreat MBONs. Aversive learning disrupts this balance and decreases the activity of approach MBONs and increases the activity of retreat MBONs which subsequently determine behavioural parameters such as the performance index. In contrast, appetitive learning decreases the activity of Retreat MBONs and increases the desire of the fly to approach an odour. The model I introduce is a spiking neural circuit with an architecture that remains unchanged during all the experiments we reproduce. To model individual neurons in the mushroom body I use a simplified phenomenological model of a neuron called the adaptive exponential leaky integrate and fire neuron ([Gerstner and Brette, 2009](#)) which I introduce in the next section.

1.9 A phenomenological model of the neuron

Spiking neuron models seek to reproduce the action potential. The interior of the neuron has a negative charge with respect to their surroundings. When the neuron receives excitatory stimuli the neuron is depolarized until it reaches a threshold value. When the threshold value is reached the neuron fires an action potential. After the

neuron fires it enters a refractory period, where the neuron is harder to excite. If we order spiking models by complexity, the most complete model is the Hodgkin-Huxley model which was based on pioneering studies in the neuron of the giant squid. The model consists of four equations, describing the membrane potential and currents of Na^+ and K^+ ions and uses 10 parameters to configure. However, it is notoriously difficult to tune so theoreticians have looked at more simple phenomenological models such as the leaky integrate-and-fire model. The integrate-and-fire type elicits a spike when the membrane potential exceeds a threshold. While this is a crude approximation it is computationally far less expensive to simulate in a network of spiking neurons compared to the Hodgkin-Huxley model. In this thesis we use the adaptive integrate and fire model (AdEx) ([Gerstner and Brette, 2009](#)) which has the advantage that its parameters can be easily related to physiological quantities. The AdEx model is able to reproduce many electrophysiological features seen in real neurons, with only two variables and four free parameters.

1.10 Previous computation models of olfactory processing and learning

Previous computational models of olfaction in the fruit fly studied the role of the MB as a classifier ([Wessnitzer et al., 2011](#)), sparse coding in the mushroom body ([Luo et al., 2010](#)) and learning in both the antennal lobe and the mushroom body ([MaBouDi et al., 2017](#)).

The [Luo et al., 2010](#) model has used an available dataset of ORN responses to 110 different odours ([Hallem and Carlson, 2006](#)) to study decorrelation of odour representation in the mushroom body and highlight the role of normalization and feedback inhibition to generate sparse odour representation. Similar to this model we also study the role of feedback inhibition and obtain similar results in a spiking neural network which is more biologically plausible compared to a rate network. The authors observe that odours exhibit a highly variable response in the recorded ORN population and they propose an equation that transforms ORN to PN firing rates using a feedforward nonlinearity:

$$r_{PN} = \frac{R_{max} r_{ORN}^{1.5}}{\tau^{1.5} + r_{ORN}^{1.5}} \quad (1.7)$$

where r_{PN} and r_{ORN} are the firing rates of a corresponding ORN and PN pair and R_{max} and τ are constants. This transformation helps obtain similar average PN firing

rates for all odours.

The [Wessnitzer et al., 2011](#) model is able to learn to distinguish odour mixtures. In this model a single odour activates activates only glomeruli that respond to the same receptor type while multiple odours combined activate glomeruli that respond to different receptors. The model also implements a spike timing-dependent plasticity (STDP) learning rule at the KC-MBON synapse with neurotransmitter-based signalling representing presence of the reinforcement. Co-active pre and post spike will only modify the strength of the connection if reinforcement is activated within a time window. Each odour has a unique ORN receptor activation profile and activates an equal number of PNs. For PN dynamics the authors used a self organizing map algorithm. The KC which acts as coincidence detectors connect randomly to 10 PNs out of 150. KCs are tuned to fire only when 3 to 6 PN fire together. The authors model a single mushroom body output neuron and show that the fly can learn to separate trained and untrained odours. Significantly, the authors shows that the computational model can perform separation of two trained odours A and B and an untrained odour mixture A+B. Odour mixtures are modelled as linear combinations of the single odour activation patterns.

In Chapter 3 we build a mushroom body network in which we test learning rules similar to the model of [Wessnitzer et al., 2011](#). However, we extend this model by adding multiple MBONs and we implement learning rules that do not require postsynaptic spikes.

- Famaghi second order conditioning
- Betkowicz
- Bogacz paper

No recent olfactory circuit model has incorporated the most recent experimental discoveries in the field of fruit fly olfactory learning such as the role of individual mushroom body output neurons (MBONs) which are believed to code either retreat and approach behaviours. Experimental evidence such as electron microscopy connectivity data shows that the mushroom body is a recurrent circuit, however no model has incorporated recent connectivity data. We have based our model on the most recent experimental data which results in a novel model that makes predictions that can be tested in the future.

1.11 Questions addressed in my thesis

While there are many unanswered questions in the field of associative learning it is beyond the scope of this thesis to build a complete model on the fruit fly's olfactory brain. In the following three results chapter I investigate the following questions:

1. How is an odour transformed from a dense representation into a sparse one in the mushroom body?
2. What is the mechanistic implementation of olfactory valence learning at the KC-MBON synapses?
3. Can a depression only learning rule explain enhanced firing rate after aversive learning of some MBONs? Is it sufficient to explain aversive and appetitive extinction? Can it also explain relief learning and re-learning?
4. What are the functional requirements at the circuit level for aversive and appetitive extinction in the olfactory learning circuit (i.e is feedback from MBONs onto DANs necessary) ?
5. How are changes in MBON firing rate translated into behavioural output and experimental observable variables such as the Performance Index?

Chapter 2

Sparse decorrelated odour representation in the mushroom body

2.1 Ch 2: Abstract

The fruit fly, *Drosophila Melanogaster* can detect an odour using its antennae and distinguish odours with very similar chemical structures. In the first layers of processing, the odour is transformed from a dense spatio-temporal odour pattern of activity in the antennae into a sparse representation in the mushroom body where only a subset of odour coding cells are activated.

Odour detection in the fruit fly begins with 1200 olfactory receptor neurons (ORNs), each of which expresses one of 51 odourant receptors. Information is then transferred to projection neurons (PNs) in the antennal lobe, which then send axons to multiple areas of the brain including the mushroom body (MB), where Kenyon Cells (KCs) form a sparse odour representation. Finally, the sparse odour evoked activity in the mushroom body is pooled by so called Mushroom Body Output Neurons (MBONs) which determine approach or retreat behaviour in response to an odour.

Here, we investigate how dense olfactory receptor neurons activity is transformed into spatio-temporally sparse KC activity in the mushroom body. Our model shows that a diverse number of mechanisms facilitate sparse representation in the mushroom body. We tune our model to elicit realistic firing activity and show that multiple mechanisms enable decorrelation and sparse activation of neurons in the mushroom body. In our results section we highlight a few of them: PN-KC connectivity, spike frequency adaptation, lateral inhibition in the antennal lobe and APL feedback inhibition in the mushroom body. We also find that adaptation in PNs and KCs contribute increasing

temporal sparseness in KC odour responses.

We test our model by using a data-set of ORN responses to 110 different odours as input to our network and show that our network encodes the odour from the dataset in a sparse and decorrelated manner. We show that feedback inhibition reduces the average number of KCs that respond to an odour. Finally we visualize MBON responses to odurs in the dataset, which lays the groundwork to test learning rules the modify KC-MBON synapse, which we investigate in Chapter 3,4 and 5.

2.2 Introduction

Flies, like other animals are capable to identify relevant stimulus features and neural computation helps it achieve this goal. A large body of evidence has identified sparse stimulus encoding as an essential feature of neural coding in both invertebrates and vertebrates ([Olshausen and Field, 1997](#), [Lin et al., 2014a](#)). Sparse coding is a coding strategy by which a stimulus activates only a small subset of the total population of neurons. In the fruit fly sparse coding has been shown to help reduce overlap between odour representations and enhance the ability to discriminate between similar odours.

Flies detect odours using approximately 1200 ORNs in their antennae. The tuning of each ORN is determined by a single odorant receptor gene, which controls whether the ORN is broadly or narrowly tuned to odours ([Hallem and Carlson, 2006](#)). ORNs expressing the same receptor converges to the same glomerulus where their activity is picked up by excitatory PNs and inhibitory LNs. PNs in the antennal lobe have been shown to normalize odour responses by gain modulation depending on the concentration of the odour and have also shown adaptation over time ([Bhandawat et al., 2007](#)). PNs project their axons to KCs and it is believed that the connectivity between PNs and KCs is random with each PN connecting to approximately 6 KCs. Experiments have shown that even when exposed to an odour for 1 second KCs show adaptation immediately after odour onset ([Murthy et al., 2008](#)). Thus KCs have been shown to respond to odours in both a spatially and temporally sparse manner. It is unknown how KCs spike during a 60 second odour presentation which is the duration of exposure in learning protocols. KCs connect to MBONs which are the neurons that determine approach and retreat behaviour in the fruit fly ([Aso et al., 2014](#)). It is unknown how MBONs fire to odours over long odour exposure time.

Here we show that while KCs respond selectively to odours, MBONs respond to all odours in our simulations. This model forms the basis of testing plasticity the site of KC-MBONs synapses. To test transformations of a dense odour representation we

use both synthetic generated ORN responses based on a function and a dataset that recorded 24 ORN type responses to 110 different odours as input into our network. We tune our network to elicit realistic responses at each layer in our network (ORN, PNs, KCs and MBONs). We find that lateral inhibition in the antennal lobe and feedback inhibition in the mushroom body is sufficient to decorrelate odour representations. In Chapter 4 we show that by decorrelating overlapping representation, we can learn to associate an odour with an aversive stimulus without affecting the MBONs response to other untrained odours.

Our model is based on previous theoretical work that has highlighted the role of feedback and lateral inhibition to support sparse coding in the fruit fly ([Wessnitzer et al., 2011](#), [Betkiewicz et al., 2018](#)).

We model a four layer olfactory circuit (ORNs, PNs, KCs and MBONs) based on experimental evidence that studied the anatomy of the olfactory circuit.

2.3 Methods

We modelled a four layer (ORN, PNs, KCs and MBONs) using the spiking neural network simulator Auryn ([Zenke and Gerstner, 2014](#)). We base our model on previous work ([Wessnitzer et al., 2011](#), [Betkiewicz et al. \(2018\)](#)) that investigated sparse decorrelated olfactory representations in a three layer model of the olfactory circuit (ORN-PNs-KCs) and studied the role of lateral inhibition and spike frequency adaptation in facilitating temporal and spatial sparseness. We use the same method to generate synthetic odours and we modify the previous model by adding feedback inhibition in the mushroom body and by also adding a fourth layer (MBONs) that receives inputs from KCs. HACK Our model includes 24 ORN types, 24 PNs, 24 LNs, and 1000 KCs. The connections between the four network layers are feed-forward and excitatory. LN provide lateral inhibition to PNs in the second layer while the APL neuron provides feedback inhibition in the third layer. In our model each of the 24 LN-PN pairs constitute one of the 24 glomeruli that exist in the fly ([Hallem and Carlson, 2006](#)). KCs are connected to the APL neuron in an all to one manner, while the APL also inhibits all of the KCs in our model. ORNs provide input to PNs and LNs. All ORNs of the same receptor type target the same glomerulus (PN-LN pair). Every LN has inhibitory connections with all PNs that is not part of the same glomerulus, mediating lateral inhibition within the layer. Connections between PNs and KCs were randomly drawn. For a KC layer size of 1000 KCs, a PN connects to approximately 200 KCs). Each KC receives input from an average of 10 PNs based on

experimental data([Turner et al., 2008](#)). Responses to a set of 4 stimuli over 50 trials, each odour lasting 5 seconds were simulated either with synthetic odours or with odours from the Hallem and Carlson dataset. At the Hallem and Carlson dataset records the activity of only 24 ORN types, we reduced our network size from 50 PNs to 24 PNs and doubled the connection strengths from PNs to KCs.

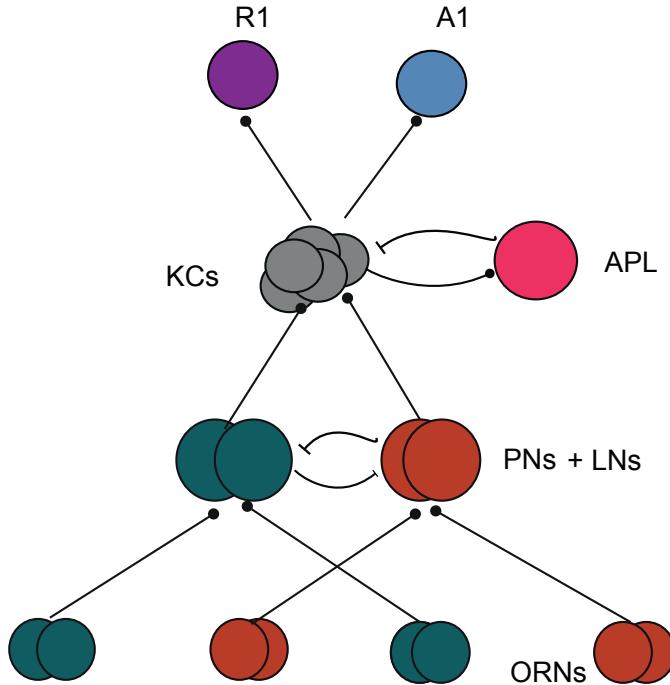


Figure 2.1: Olfactory circuit model diagram. Circuit diagram shows the main components of our model starting from bottom: 24 ORN types, 4 of each kind. ORNs of the same type excite 1 unique PN and 1 LN which form a glomerulus. LNs inhibit PNs which are not in the same glomerulus. PNs excite 1000 KCs with each KC receiving input from 10 random PNs. KCs are randomly connected with excitatory connections to two MBONs.

2.3.1 Modelling ORN responses

We used published data-set ([Hallem and Carlson, 2006](#)) that records firing rates of 24 different classes of ORNs in response to 110 different odours as input to our model. In the experiment that recorded the ORN responses, each odour was presented for six 500 ms trials, and ORN responses were recorded as spikes per second.

Most odour responses were found to be excitatory, however 20% recorded ORN decreases in firing rate in responses to the odour. We modelled ORN responses based on the available dataset by transforming the firing rates into Poisson spike trains.

The firing rates from the data-set were transformed into homogeneous Poisson spike trains for each ORN.

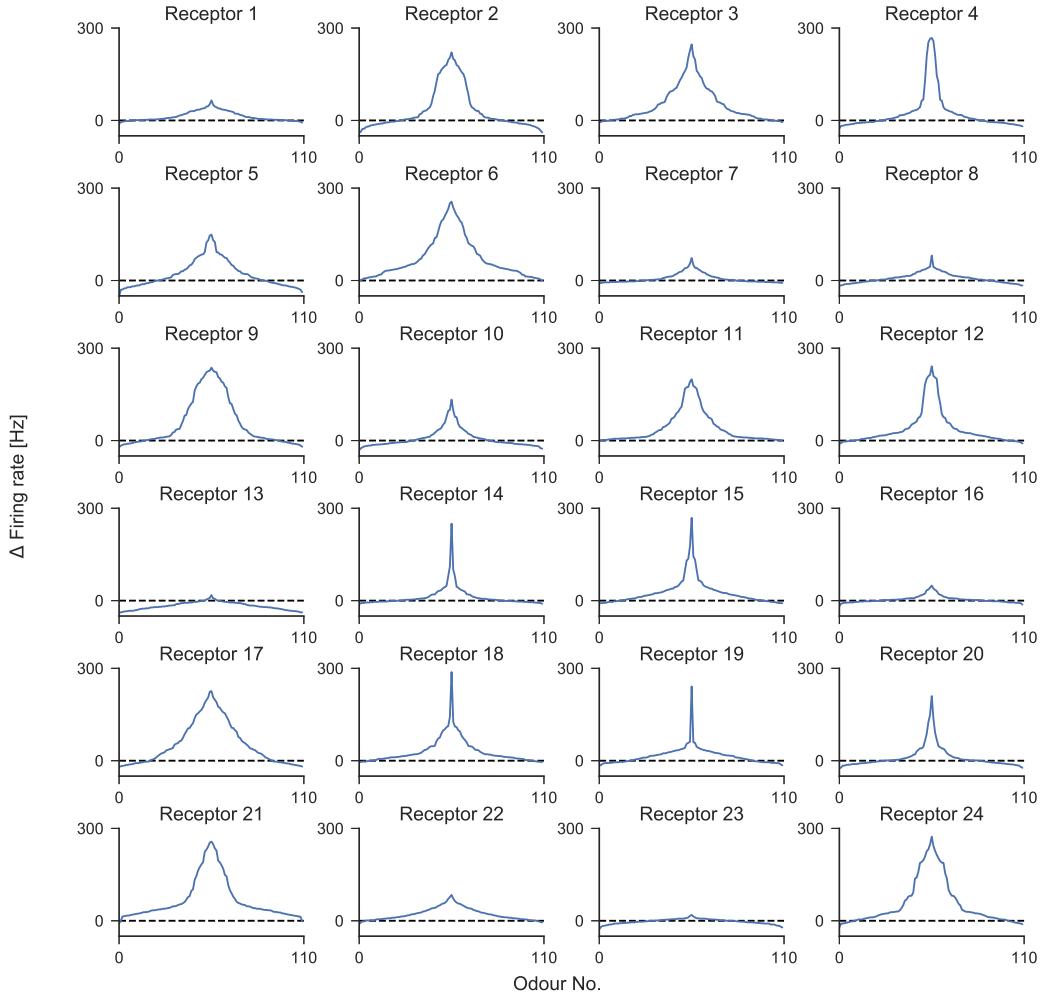


Figure 2.2: Response curves for 24 ORN types to 110 odours (Hallem and Carlson, 2006). Odourants are sorted along the x axis according to the response strengths they elicit from a receptor. Odourants that elicit strongest responses to the receptor shown are displayed near the centre of the curve while the weakest responding are at the edges.

2.3.2 Simulation conditions

Four different scenarios were tested: i) without lateral inhibition and feedback inhibition ii) with lateral inhibition iii) with feedback inhibition and iv) with lateral inhibition and feedback inhibition. During simulations PNs are tuned to fire at 4 Hz during spontaneous activity. KCs are tuned to fire at 0.1 Hz during spontaneous activity. Each KC receives input from 8 to 10 PNs assigned randomly. We also model

one approach MBON and one retreat MBONs that pool random connections from all of the 1000 KCs.

2.3.3 Neuron model

PN, LN, KC and MBONs were modelled using The AdEx model with conductance based synapses. The model is based on the integrate and the membrane potential is described by the following equation:

$$\tau_m \frac{dV}{dt} = (V_{rest} - V(t)) + g_{ex}(t)(E_{ex} - V(t)) + g_{inh}(t)(E_{inh} - V(t) + w). \quad (2.1)$$

Here, V is the membrane potential of the neuron as a function of time, τ_m is the membrane time constant, V_{rest} is the resting membrane potential, E_{ex} is the excitatory reversal potential and E_{inh} is the inhibitory reversal potential. w represents adaptation and follows the following dynamics:

$$\tau_w \frac{dw}{dt}(t) = a(V(t) - v_r) - w(t) + b\tau_w \sum_k \delta(t - t_{k,i}^*), \quad (2.2)$$

where τ_w is the adaptation time constant while a is the subthreshold adaptation conductance. When the membrane voltage reaches a predefined threshold a spike is fired which resets the membrane voltage to a reset value and the adaptation current is increased by b which is a parameter that controls the effect of spike triggered adaptation. Finally $t_{i,k}^*$ is the timing of the spike, where i is the index of the neuron and k is the index of the spike.

The synaptic conductances are expressed by g_{ex} (excitatory) and g_{inh} (inhibitory). They are modelled according to the following equations:

$$\tau_{ex} \frac{dg_{ex}(t)}{dt} = -g_{ex}(t) \quad (2.3)$$

and

$$\tau_{inh} \frac{dg_{inh}(t)}{dt} = -g_{inh}(t), \quad (2.4)$$

where τ_{ex} and τ_{inh} are the synaptic time constants for the excitatory and the inhibitory conductance, respectively.

When the neuron receives an action potential from a presynaptic cell, the postsynaptic conductance increases by the following formulas: $g_{ex} \rightarrow g_{ex} + \Delta g_{ex}$ and

$g_{inh} \rightarrow g_{inh} + \Delta g_{inh}$ for excitatory and inhibitory synapses, respectively. Parameters for our model PNs and KCs can be seen in Table 2.1

Parameter name	Symbol	Value
Membrane time constant	τ_m	20 ms
Leak conductance	E_L	30 nS
Reset potential	V_r	-70.6 mV
Threshold potential	V_{th}	-50.4 mV
Refractory Time	τ_{ref}	20 ms
Inhibitory synaptic potential	E_i	-80 mV
Excitatory synaptic potential	E_e	0 mV
Excitatory time constant	τ_e	5 ms
Inhibitory time constant	τ_{inh}	10 ms
Synaptic parameters no-adaptation		
ORN-PN Synaptic weight	w_{orn-pn}	1 nS
PN-KC Synaptic weight	w_{pn_kc}	2 nS
KC-MBON Synaptic weight	w_{pn_kc}	2 nS
Synaptic parameters with adaptation		
ORN-PN Synaptic weight	w_{orn-pn}	1 nS
PN-KC Synaptic weight	w_{pn_kc}	2 nS
KC-MBON Synaptic weight	w_{pn_kc}	2 nS
Adaptation parameters		
PN subthreshold adaptation	pn_a	1 nS
PN spike frequency adaptation	pn_b	2 nS
PN adaptation time constant	pn_{τ_w}	2 nS
KC subthreshold adaptation	kc_a	0 mV
KC spike frequency adaptation	kc_b	2 ms
KC adaptation time constant	kc_{τ_w}	5 ms
MBON subthreshold adaption	E_e	0 mV
MBON spike frequency adaptation	τ_e	2 ms
MBON adaptation time constant	τ_{inh}	5 ms

Table 2.1: Neuron parameters

2.3.4 Calculating separation between two odours

Angular separation of odour was calculated for PN and KC responses. Population vectors representing firing rates in response to each odour in the data set were used to calculate angular separation as between each pair of odours:

$$\text{Angular distance } i,j = 1 - \cos(\theta) \quad (2.5)$$

where $\cos(\theta)$ is

$$\cos(\theta) = \frac{\mathbf{n}_i \cdot \mathbf{n}_j}{\|\mathbf{n}_i\| \cdot \|\mathbf{n}_j\|} \quad (2.6)$$

where the cosine of the angle between these vectors represents the similarity measure between the vectors. \mathbf{n}_i and \mathbf{n}_j represents the spike train of each neuron in the population in response to odour i and odour j respectively. The value of similarity range from 0 (for orthogonal vectors of $\theta = 90^\circ$) to 1 (for parallel vectors of $\theta = 0$).

2.4 Results

To study odour representation in the fruit fly's mushroom body we built a four layer spiking neural network (Fig. 2.1) composed of the following features: an input layer made of 24 ORN types, the antennal lobe (AL) and the mushroom body (MB).

After studying olfactory processing using a synthetic odour dataset we tested our model with a published dataset that recorded responses for 24 ORN types to 110 different odours. The receptor response profile is diverse, with some receptors responding broadly to many odours in the dataset while some only responded to very few (Fig 2.2). To avoid introducing bias into the dataset we reduced the number of PNs to 24 which receive input from 24 types of ORNs as they have been recorded in the dataset. We have otherwise kept the network from the previous section unchanged. We also plot 16 randomly selected odour responses profiles to the 24 ORNs (Fig 11). Here, we also observe that odour profiles are diverse as some odours activate ORNs strongly (Odour 4, 5, 6 Fig ??) while others activate them weakly (Odour 2,11 Fig ??). We also calculate the mean number of spikes elicited by ORNs and odours (Fig 2.3) and sort them in increasing order of mean number of spikes.

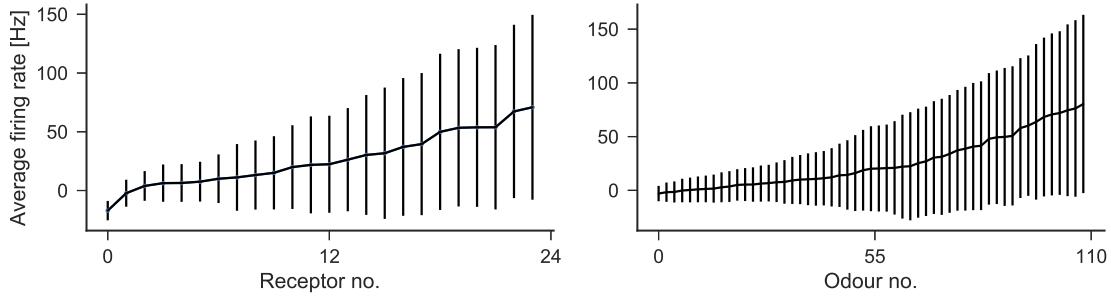


Figure 2.3: Average number of spikes elicited by each odour and each receptor from [Hallem and Carlson, 2006](#) data-set. **a.** The average number of spikes for each of the 24 ORN types to all of the odours in the data-set. **b.** The average number of spikes for each of the 110 odours in response to all of the ORNs. Vertical bars indicate one standard deviation.

To test our network response we selected four odours in our dataset based on the response profile we calculated in Fig ?? (two odours that activate ORN weakly and two odours that activate ORNs strongly). We find that PNs and KCs responded to all four odours (Fig ??) but each odours elicits a different number of average spikes in the PN and KC population (Fig ?? c,d).

Here we present results by simulating response to a single odour. Without spike frequency adaptation Fig 2.5 b,c illustrates PN and KC responses to four different stimuli. PNs excited by the stimulus showed a strong response at stimulus onset, followed by an adaptation during the stimulus period. PNs with no direct input from odour stimulated ORNs were suppressed during stimulus onset. In the MB, KCs showed close to zero spikes during spontaneous activity and responded to odours and a smaller number of spikes compared to PN (Fig 2.1 b,d).

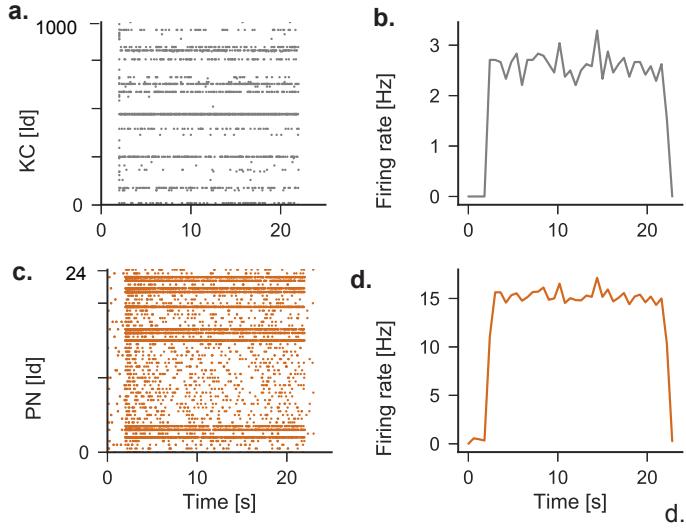


Figure 2.4: Network activity during simulation of one odour with no spike frequency adaptation **a.** Raster plot of KC responses to one odour simulated for 20 seconds. Each dot (grey) represents a spike. **b.** Population firing rate of KCs during simulation of one odour. **c** Like in **a.** we show a raster plot of PNs. **d.** Like in **b.**, we show the population firing rate of PN.

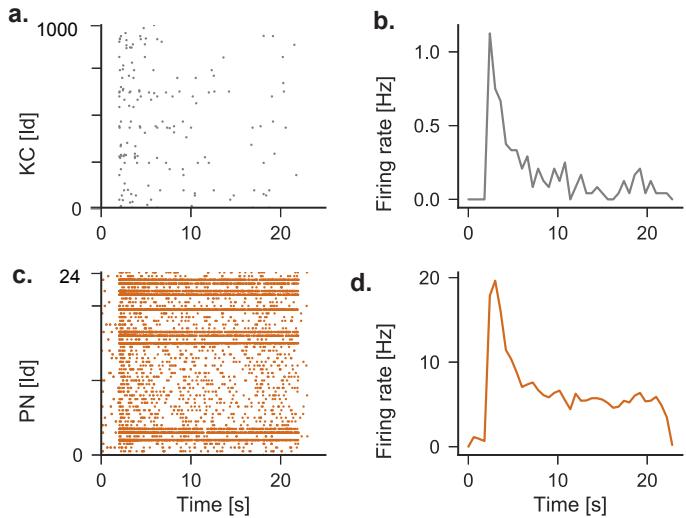


Figure 2.5: Network activity during simulation of one odour with spike frequency adaptation enabled **a.** Raster plot of KC responses to one odour simulated for 20 seconds. Each dot (grey) represents a spike. **b.** Population firing rate of KCs during simulation of one odour. **c** Like in **a.** we show a raster plot of PNs. **d.** Like in **b.**, we show the population firing rate of PN.

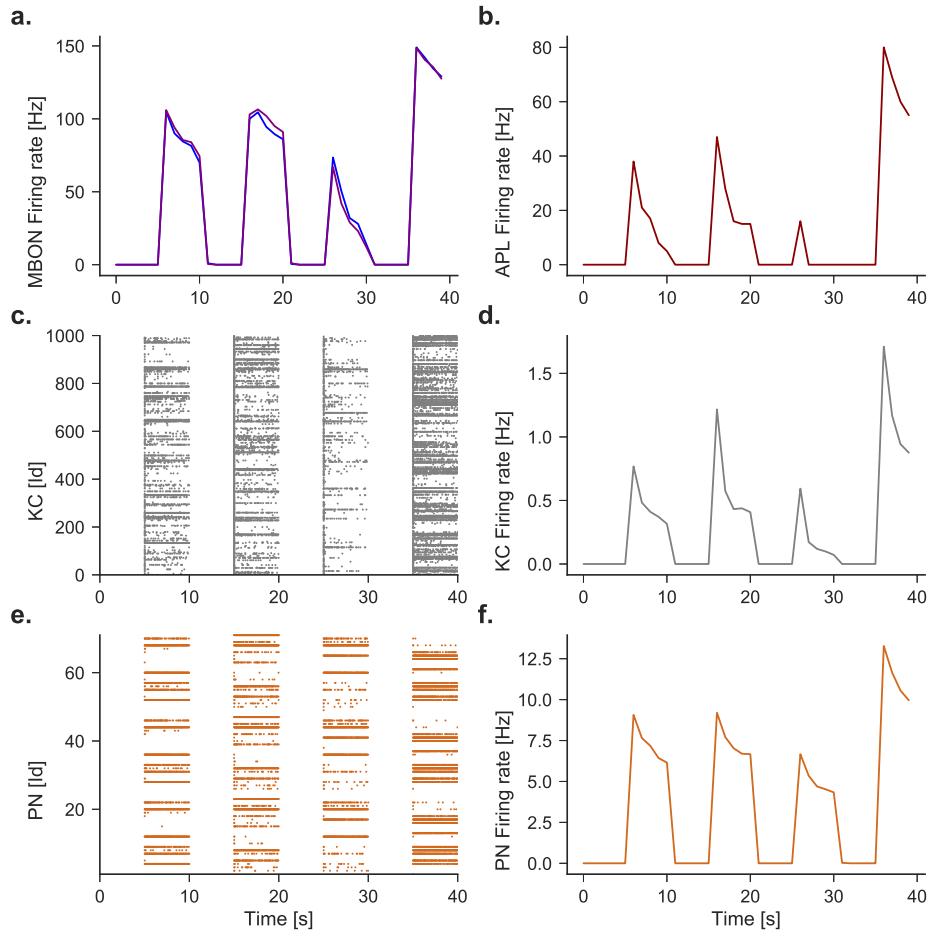


Figure 2.6: Network activity during presentation of four different odours. **a.** Shown here are the firing rates of two MBON: M6 (purple) and MVP2 (blue) in response to four different odours. Each odour is exposed for 5 seconds followed by five seconds of spontaneous activity. **b.** Firing rate of the GABA-ergic APL neuron during the simulation of four odours. **c.** Raster plot of KCs during the simulation. Each grey dot represents a single spike. **d.** Population firing rate of all the KCs in response to the four odours. **e.** Raster plot of PNs during the simulation. **f.** Population firing rate of the PN population (orange).

Average KC activity was different across the four different odours in single trials. This is because of random PN-KC synaptic weights and the wiring of PNs to KCs can lead to odours activating some KCs stronger than others.

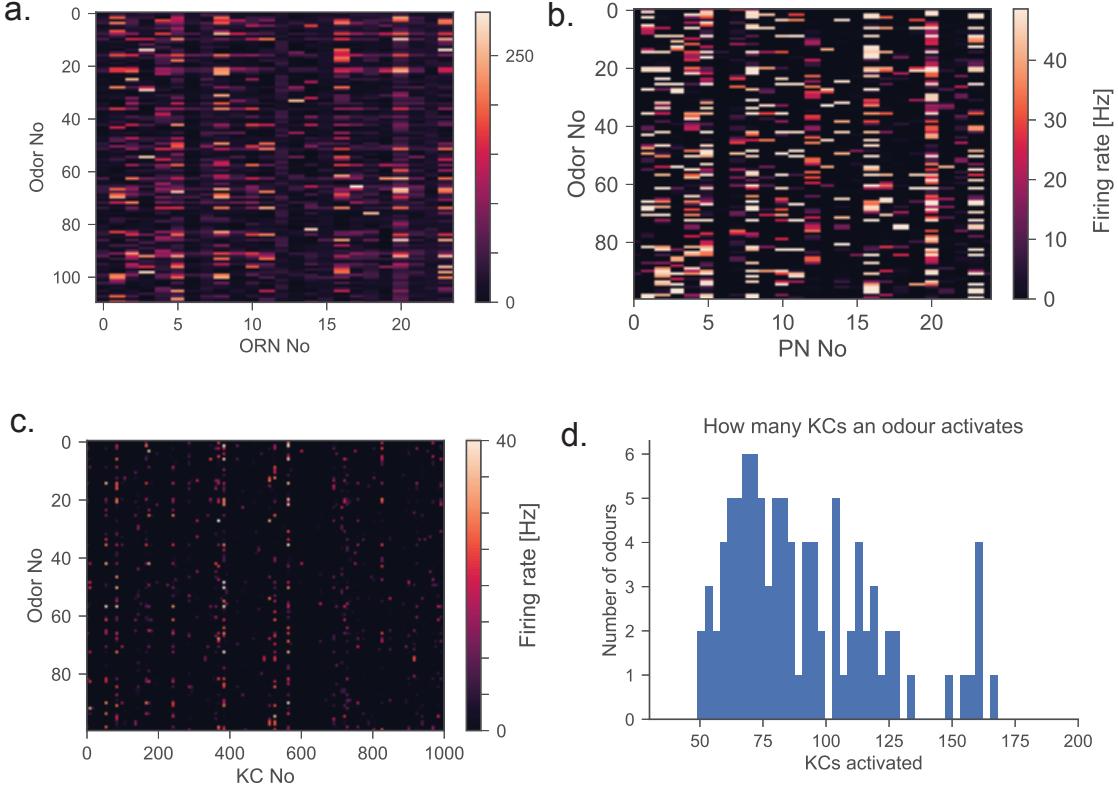


Figure 2.7: ORN, PN and KC responses to odours in the [Hallem and Carlson, 2006](#) data set. **a.** We show a colormap of ORN responses to 110 different odours in the data set. **b.** We show the entire population of 24 PN responses to the same data as in **a.** and in the same format. Each group of three PNs receive input from one out of 24 ORN types. **c.** 1000 KC responses to the odours in the data set with the same format as in **a.** and **b.**. **d.** Histogram showing how many KC an odour activates.

2.4.1 Decorrelation of odour representation in the olfactory circuit

In order to investigate the effect of lateral inhibition in the antennal lobe and APL mediated feedback inhibition in the MB we simulated odours responses by activating or deactivating each mechanism separately. We deactivated lateral inhibition and adjusted the ORN-PN weights to keep 4 Hz PN firing rate during spontaneous activity. When lateral inhibition was activated PNs not driven by the stimulus were suppressed (Fig ?? d). The average number of PN and KCs activated by a stimulus decreased when we activated lateral inhibition (Fig 2.8 c,f). Lateral inhibition suppresses PNs which are not strongly activated by the stimulus increasing population sparseness in the PN population (Fig 2.8 c). Reducing the number of activated PN results in the

reduction of the number of activated KCs as well (from 25% to 10% trial average, Fig 2.8 c,f).

As we have shown that lateral inhibition reduces number of activated PNs, we were also interested to see whether it can reduce similarity of responses in the PN layer for odours which have overlapping representations in ORN responses. To study similarity of responses in different layers we calculated the angular distance (where 1 is maximum separation and 0 represents patterns which are identical). We plot these results in (Fig 2.8) and we find that mean angular distance increases when lateral inhibition is activated (Fig 2.8 d).

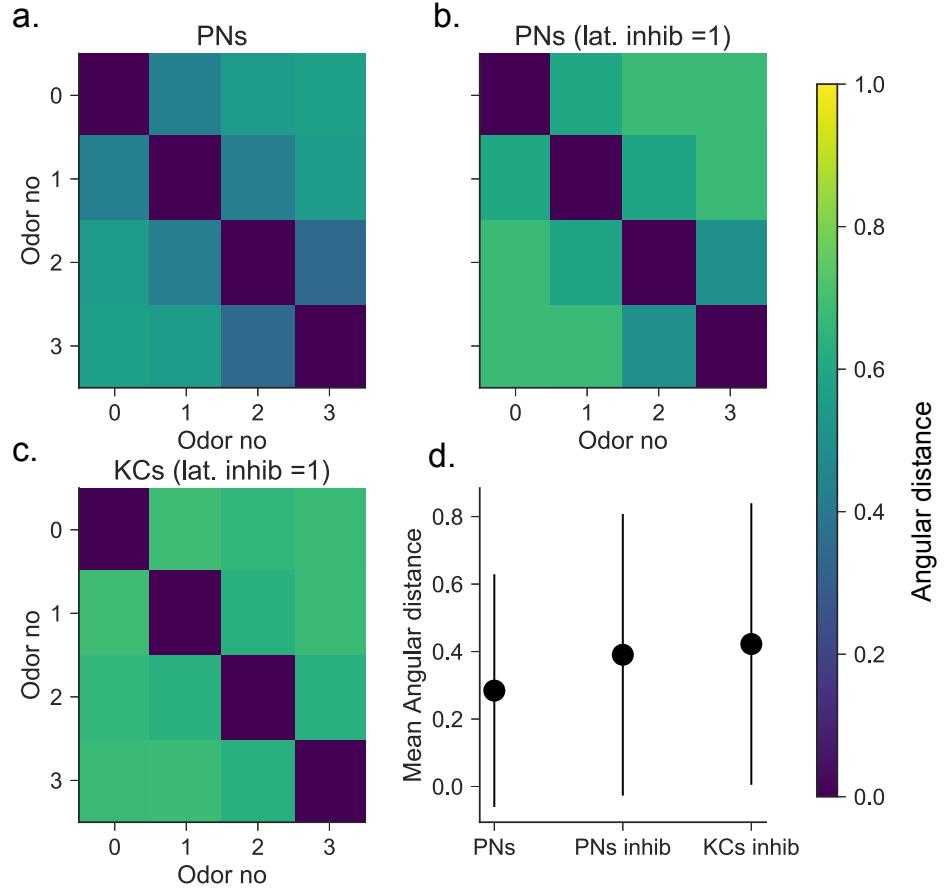


Figure 2.8: Lateral inhibition facilitates pattern decorrelation in the MB. We simulated four patterns with approximately 20% overlap. **a.**, We show odour pattern correlation (using angular distance) in the PN population with no lateral inhibition. **b.**, We show pattern correlation in PNs with lateral inhibition active. **c.**, Similar to **b.** we calculated the odour pattern correlation in the KC layer. **d.**, We illustrate the the mean angular distance for **a.**, **b.** and **c.**. Vertical bars indicated 1 standard deviation.

We showed that lateral inhibition increases dissimilarity of odour patterns in the PN layer. We also study the effect of APL mediated feedback inhibition without lateral inhibition on pattern similarity. We find that feedback inhibition increase the mean angular distance between odour patterns represented by KC activity (Fig 2.9 d.).

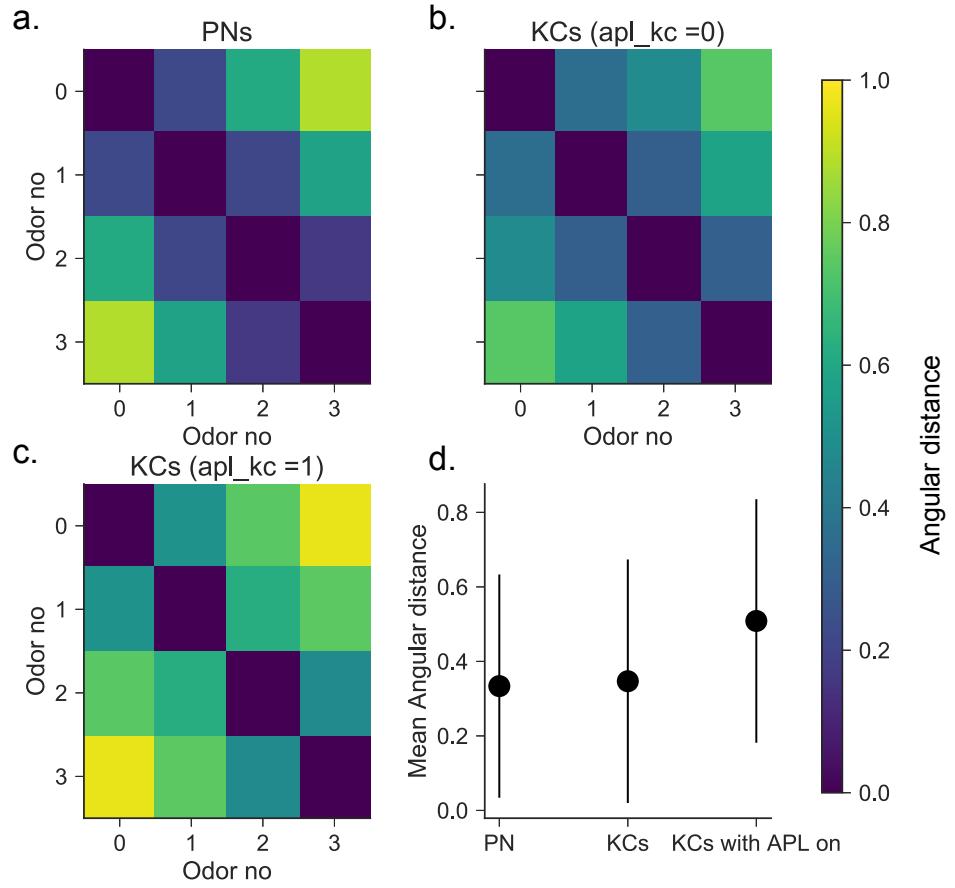


Figure 2.9: APL mediated feedback inhibition facilitates pattern decorrelation in the MB. We simulated four patterns with approximately 20% overlap. **a.**, We show odour pattern correlation (using Pearson correlation) in the PN population. **b.** shows pattern correlation in KCs with no APL feedback inhibition. **c.**, We show the same as **b.** but with APL-mediated inhibition activated. **d.** We illustrate the the mean angular distance for PNs, KC and KCs with APL feedback inhibition active as shown in **a.**, **b.** and **c.**

Our results show that both lateral inhibition in the PN layer and feedback inhibition in the KC layer can increase pattern separation in the activities of PNs and KCs respectively. We now investigate how the two mechanism interact when they are both activate like in the biological implementation in the fruit fly's olfactory circuit where both mechanisms are active. To study this we fixed the strength of APL feedback inhibition and changed the strength of lateral inhibition in the interval [0, 25] nS. We calculated mean angular distance both in the PN layer (Fig 10 blue) and in the KC layer (Fig 10, purple) and found that separation of odour patterns increased mono-

tonically with the strength of lateral inhibition. We find that the standard deviation of the mean angular distance is smaller in the KC layer compared to the PN layer (Fig 10 shaded blue, shaded purple).

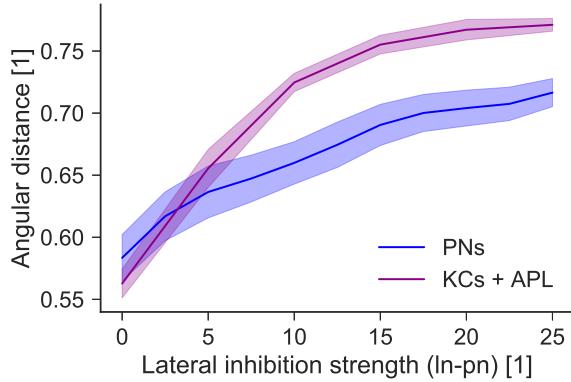


Figure 2.10: Average angular distance in PNs and KCs as a function of lateral inhibition. We show changes in average pattern similarity (angular distance) calculated across four stimuli in the PN (blue) and KC (purple) layers as we increase lateral inhibition. Shaded areas indicate 1 standard deviation.

The reduction of pattern correlation from PN to KC population has been observed in experiments ([Wolfe et al., 2010](#)). APL mediated feedback inhibition has been shown to be required to both maintain a sparse representation and decorrelated patterns([Lin et al., 2014a](#)). In the next section we study a dataset that recorded ORN responses to 110 different odours to verify whether our results hold when we use a more realistic input into our network. We also study the effect of APL mediated inhibition on sparse representation.

2.4.2 Studying APL mediated decorrelation using the Caron dataset

We now investigate the angular separation of the odour response in PNs and KCs by simulating all 110 odours. Our results show that APL mediated feedback inhibition increasing angular separation across the entire population of odours (Fig ?? d).

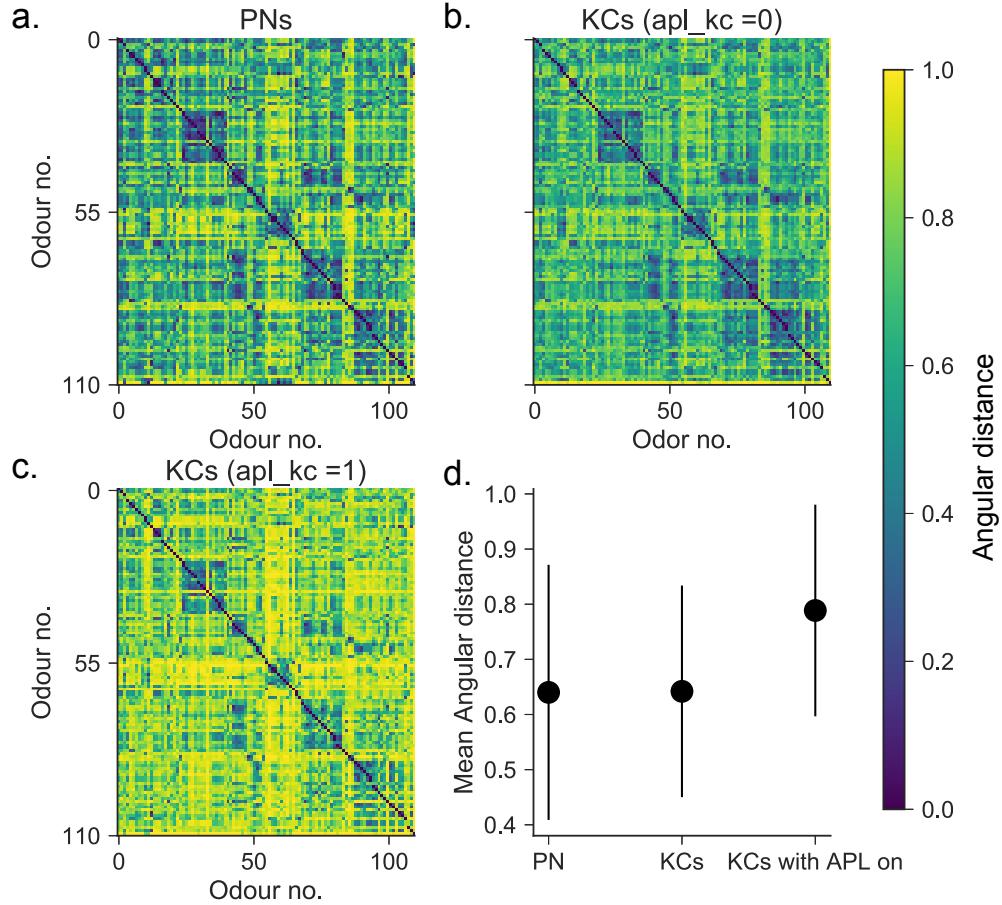


Figure 2.11: APL mediated feedback inhibition facilitates pattern decorrelation in the MB. We simulated 110 odour patterns from the Caron et al. 2006 data set and calculated odour odour similarity for all pairs using angular distance. **a.** shows odour similarity based on angular distance in the PN population. **b.** shows angular distance in KC population with no APL feedback inhibition. **c.** is the same as **b.** with APL-mediated inhibition activated. **d.** illustrates the the mean angular distance for PNs, KC and KCs with APL feedback inhibition active as shown in **a.**, **b.** and **c.**

Finally we studied how APL mediated feedback inhibition affects population sparseness in our dataset. We found that increasing the feedback inhibition in the mushroom body leads to a decrease in average number of KCs that respond to odours (Fig 2.12 a) and the number of odours that an individual KC responds to (Fig 2.12 b.).

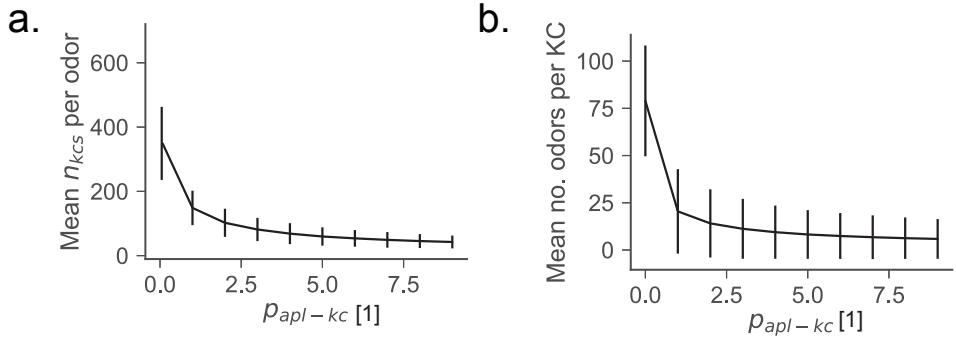


Figure 2.12: KC-APL mediated inhibition increases angular separation between odours and decreases number of responding KCs per odour. Angular separation for 110 different odours were calculated for all pairs of odours in the KCs with (A) and without (B) APL mediated inhibition. c. shows number of KCs responding to an odour on average as a function of strength of APL mediated feedback inhibition. d. shows how many odours does 1 KC respond on average as a function of APL mediated inhibition.

Here we show the firing rates elicited by MBONs in response to 110 odours we simulated based on the Carlson data set of recorded ORN responses. **a.** We show KC responses to 20 of the 110 odours we simulated. Each odour was exposed for 5 seconds followed by 5 seconds of spontaneous activity. The diagonal represents odours that elicit equal approach and retreat MBON firing artes.

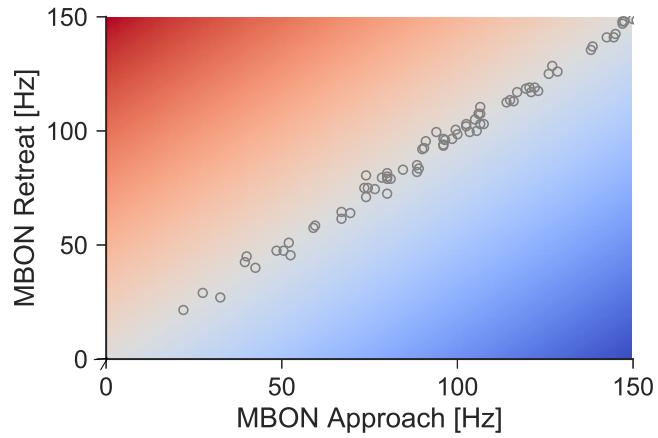


Figure 2.13: We show MBON responses to 110 odors from the

2.5 Discussion

In this chapter we laid the foundations in understanding how dense odour information is transmitted to the MBONs which guide the behaviour of the fruit fly in an

odour specific manner. A spiking network of 4 layers (ORNs, PNs and LNs, KCs and MBONs) was simulated to investigate sparse representation in the PN and KC layers of the fruit fly's olfactory circuit. We reduced the complexity of the olfactory circuit by making a number of simplifications. We modelled 4 ORNs for each of the 50 types rather than the full number of 1200 ORNs. We also modelled only 50 PNs corresponding to the 50 glomeruli rather than the total number of 150 PNs. Connectivity in the olfactory circuit is more complex than we have shown in our model. A number of studies have shown that there are recurrent connections even in the early stages of processing from PNs to ORNs for example (Robinson et al. 2018). However, these simplifications helped us highlight the role of inhibitory processes in decorrelating overlapping representations. Our spiking model suggests lateral inhibition and feedback inhibition are sufficient mechanisms to explain how dense responses in the antennal lobe are transformed in sparse representations in KCs. Sparse coding reduces overlap between representations, which has been shown to require APL feedback inhibition in the mushroom body ([Lin et al., 2014a](#)). To study responses to different odors we created synthetic ORN odours which allowed us to control the degree of overlap between odours. This approach allowed us to study how overlap in ORN representation gets transformed in the antennal lobe and the mushroom body. Lateral inhibition and feedback inhibition increases population sparseness in KC odour responses and reduces overlap between representation in our model as has also been shown in past theoretical studies ([Olshausen and Field, 1997](#)). We have obtained similar results for both the synthetic odour dataset and the experimental dataset ([Hallem and Carlson, 2006](#)). We tuned APL feedback inhibition to match population sparseness in KCs in our model matches calcium imaging experimental results in Drosophila ([Honegger et al., 2011](#)). In conclusion we have shown that a four layer olfactory circuit model can create sparse decorrelated odour representations that propagate from ORNs to MBONs.

In Chapter 3 and 4 we explore plasticity at KC-MBON synapses. However, plasticity has also been observed in the antennal lobe with a role in habituation. Feedback inhibition (LN-PN) activity is strengthened after repeated exposures to the same odour [Sudhakaran et al., 2012](#).

Chapter 3

Dopamine modulated aversive learning and extinction in the fruit fly

3.1 Ch3: Abstract

The fruit fly can learn to avoid odours that were paired with shock during past experiences and approach odours paired with sugar by forming associative memories in an area of the brain called the mushroom body. Changes in the behaviour of the fly are reversible: if the odour is re-exposed without the expected reinforcement approach and retreat behaviours will be attenuated. Learning modifies synaptic connections between odour coding Kenyon Cells (KCs) and behaviour influencing mushroom body output neurons (MBONs) and requires dopamine release from dopaminergic neurons (DANs). Recent experiments have shown that aversive and appetitive memories form in different compartments of the mushroom body where the synapses of retreat or approach promoting MBONs are modified. The mechanism behind the modification of KC-MBON synapses appears to be predominantly depression dominated – i.e aversive learning reduces the firing rate of approach MBONs, and appetitive learning reduces the firing rate of retreat MBONs. To study olfactory learning in a computational model of the fruit fly’s brain, we modelled three classes of neurons in the mushroom body: KCs, MBONs and DANs. We tuned our model to elicit biologically plausible firing rates and we connected neurons in our model based on available connectivity data. We introduced a simple transformation from MBON firing rates into a behavioural measurement called the performance index which measures how many flies approach a reinforced odour and we show that our model can obtain very similar results to those measured in experiments. After tuning our network, we

tested dopamine-modulated synaptic learning rules for spiking neurons to reproduce fruit fly olfactory learning experiments. Specifically, we used two learning rules: rule high-dopa-LTD which requires only presynaptic activity and dopamine to modify KC-MBON synapses and rule R-STDP which is based on reward modulated spike timing dependent plasticity.

Our results show that both learning rules can reproduce aversive learning and extinction experiments. Furthermore, experimental evidence also shows that the fly can learn aversive associations when postsynaptic spikes are suppressed which only Rule 1 reproduces suggesting that olfactory learning depends only on presynaptic and dopaminergic activity. In conclusion our model shows that learning rules that only allow for weakening of KC-MBON synapses are sufficient to explain how associative memories form in the fruit fly. Furthermore, our model suggests that connections between compartments that store appetitive and aversive memories required to reproduce extinction experiments.

3.2 Introduction

The fruit fly forms olfactory associations in the mushroom body and modifies its memories to incorporate new information about an environment in which dangers and opportunities change over time. This mechanism is important for animals that use olfaction since the same odour can lead to reward (food) or punishment, when there is a nearby predator to the food source. Flies are not only able to learn to associate an odour with a punishment but an aversive odour can also be used to train a neutral, in a paradigm called second-order conditioning ([Tabone and de Belle, 2011](#)).

At the level of the brain circuit, the site where odour-reinforcement associations are learned is the mushroom body. Appetitive and aversive dopaminergic neurons (DANs) project to distinct compartments in the MB which gives the fruit fly the ability to learn reinforcement related information about a stimulus in anatomically distinct regions of the brain. Different compartments learn different qualities of an odour (reward or punishment) but they are also governed by different learning rules ([Aso and Rubin, 2016](#)).

The main neurons involved in associative learning are believed to be odour coding Kenyon Cells (KCs) which excite MBONs and DANs which target the synapses of KCs and MBONs ([Aso et al., 2014](#)). KCs have been shown to fire predominantly at odour onset, although it is unknown if this is true across the entire population of KCs.

Optogenetic experiments have shown that learning in some compartments requires shorter duration DAN activation than in others, suggesting that some compartments learn with a higher learning rate than others (Aso and Rubin, 2016).

Memory performance is quantified at the level of behaviour by calculating a so-called "performance index" which measures how many flies choose the safe odour (CS-) versus how many flies select the shock predicting odour (CS+). The magnitude of the performance index measures how strong the memory is. Behavioural choice of the fly has been shown to be directly influenced by the activity of MBONs which are classified by the behaviour they promote: approach or retreat (Aso et al., 2014) while calcium imaging and electrophysiological experiments have shown how learning modifies the activity of MBONs.

During aversive learning, electrophysiological experiments have shown that the firing rate of approach promoting MBONs decreases (Honegger et al., 2011), while calcium imaging experiments have shown the activity of retreat promoting MBONs increases (Owald et al., 2015). Conversely, during appetitive learning retreat promoting MBONs have been shown to decrease their activity. In behavioural experiments an odour is exposed for 60 seconds while flies are being shocked once every 5 seconds. Experiments have also shown that shocks at the end of the odour exposure protocol increase the strength of the memory as measured by the performance index metric (Beck et al., 2000).

Flies are believed to keep a trace of the odour as they have been shown to learn an aversive association even when there is a delay between stimulus and shock (Dylla et al., 2017). After learning, the memory enters a phase of consolidation. However, this process can be disrupted if the odour is re-exposed in the absence the reinforcement that was used during initial training.

Two recent experiments have shown that re-exposing the odour without the expected reinforcement can partially extinguish aversive memories and fully extinguish appetitive memories (Felsenberg et al., 2017, Felsenberg et al., 2018).

The mechanism behind extinction has been shown to require the activity of opposite signalling DANs compared to those required for initial learning: appetitive extinction requires the activity of shock signalling DANs while aversive extinction requires reward signalling DANs to be active (Felsenberg et al., 2017, Felsenberg et al., 2018). The mechanism behind aversive extinction has been described in detail, compared to the mechanism of appetitive extinction. Aversive learning reduces the activity of the MVP2 MBON that inhibits M4/6 retreat MBONs increase their activity in response to the trained odour through disinhibition. The M6 retreat MBON has

been shown to excite the *PAM* – γ 5 dopaminergic neurons in its own compartment, which could allow an odour to trigger plasticity during extinction (Felsenberg et al., 2018). After aversive extinction, a subsequent training session of the odour paired with the shock leads to re-learning of the aversive association.

A recent study that has reconstructed the wiring diagram of the entire connectome of the larva fly’s mushroom body has proposed a canonical microcircuit that enables associative learning (Eichler et al., 2017b). Interestingly, they discovered that there are as many connections from KCs to DANs as there are connections from KCs to MBONs. This may explain why most DANs respond to untrained odours (Dylla et al., 2017).

Here, we investigate a plasticity model that can reproduce behavioural and electrophysiological experiments to understand the mechanism behind olfactory plasticity in the fruit fly. We introduce a three layer model (KCs, MBONs and DANs) to study dopamine modulated plasticity (Fig 3.1). We show that after aversive learning both appetitive PAM- γ 5 DANs and aversive *PPL1* DANs increase their response to the aversive odour. We show that increased PAM- γ 5 response can signal extinction during odour re-exposure while increased *PPL1* DAN response can be used to perform second order conditioning.

3.3 Methods

To study olfactory associative learning in the fruit fly we modelled a two layer spiking neural network (Fig 3.1) using conductance-based synapses. Our network models KCs, MBONs and DANs to study dopamine modulated plasticity at the site of KC-MBON synapses. Each of the four compartments we modelled is made of a unique DAN and MBON (Table 3.1). For the sake of brevity we simplified the names of MBONs and DANs in our model based on the naming convention used in the two extinction papers we reproduced (Felsenberg et al., 2017, Felsenberg et al., 2018) (Table 3.1).

Name	Literature name	Role
Compartment 1		
MVP2	<i>MBONγ1ped</i>	Approach
PPL1	<i>PPL1γ1ped</i>	Shock reinforcement
Compartment 2		
V2	<i>MBONαV2sc</i>	Approach
PPL1-2	<i>PPL1α'3</i>	Shock reinforcement
Compartment 3		
M6	<i>MBONγ5β'2a</i>	Retreat
PAM- γ 5	PAM- γ 5	Reward reinforcement
Compartment 4		
M4	<i>MBONβ'2mp</i>	Retreat
PAM β'	<i>PAMβ'2mp</i>	Reward reinforcement

Table 3.1: Neurons modelled in our simulations

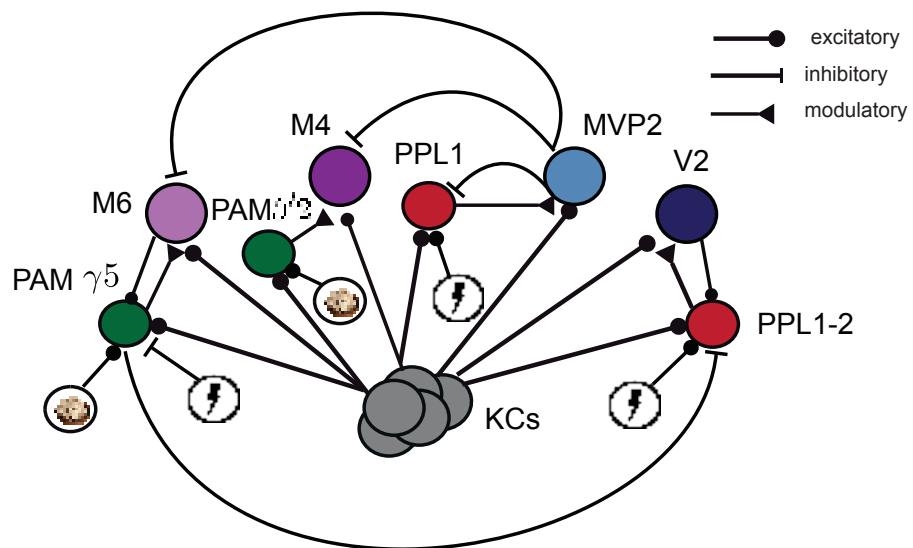


Figure 3.1: Circuit diagram of our model of the mushroom body illustrating KCs, MBONs and DANs. Schematic of our circuit which is composed of 1000 KCs (grey) exciting four MBONs (MVP2 - light blue, V2 - dark blue, M6 - light purple, M4 - dark purple). M6 MBON and M4 MBON are inhibited by the MVP2 MBON. All four DANs in our model (PPL1, PPL1-2, PAM- γ 5 and PAM- β' 2 receive excitatory input from KCs; PAM- γ 5 excited by external reward and by MBON M6 and inhibited by shocks; PPL1-2 excited external shocks and by MBON-V2 and also inhibited by PAM- γ 5; PPL1 excited by shock and PAM- β' 2 is excited by reward.

3.3.1 MB Kenyon Cells model

KCs responding to an odour were modelled as inhomogeneous Poisson spike generators: at every time step the neuron has a probability of emitting a spike, according to instantaneous firing rate:

$$r(t) = 0.1 \text{ Hz} + p_0 f(t) \quad (3.1)$$

where $f(t)$ is:

$$f(t) = \begin{cases} \exp\left(\frac{-t}{\tau_{kc}}\right), & \text{if } t > od_{start} \text{ and } t < od_{end} \\ 0, & \text{otherwise} \end{cases} \quad (3.2)$$

In our simulations p_0 is 10 and τ_{kc} is 1. We modelled KCs based on firing properties observed experimentally (Turner et al., 2008), (Murthy et al., 2008) which show that KCs fire tonically at odour onset (10-30 Hz) and decay rapidly to spontaneous activity rates where they fire at 0.1 Hz. In our simulations odours are presented for 60 seconds in all of the protocols. CS+ and CS- odour activate 100 non-overlapping KCs out of 1000. KCs that do not respond to the odour fire at spontaneous activity rates during the simulation.

3.3.2 Connectivity

Here we describe how we connected the neurons in our model based on experimental evidence. MBONs in general receive input from one or two of the lobes of the KCs (Aso et al., 2014). We simplified the complexity of KC population and modelled a single group of 1000 KCs that send excitatory input to all 4 MBONs (M6, M4, MVP2 and V2) in our model. KC-MBON weights are tuned so that MBONs fire at approximately 100 Hz when exposed to an odour which is based on experimental electrophysiological data that recorded odour responses of the MVP2 MBON (Hige et al., 2015a). Weights for MBON-MVP2 and MBON-V2 are tuned to an average of 3 nS while MBONs-M6 and MBON-M4 have an average of 6 nS (Table 3.2). Retreat MBONs (M6, M4) are inhibited by MBON-MVP2 based on experimental data showing such an inhibitory connection exists (Owald et al., 2015).

All DANs receive excitatory input from KCs based on experimental evidence (Eichler et al., 2017b). PAM- γ 5 is excited by the M6 MBON and inhibited by the presence of shock based which is based on experimental data showing PAM- γ 5 decrease their activity during shock exposure and that activation of M6 MBON increase

the activity of PAM- γ 5 (Cohn et al., 2015, Felsenberg et al., 2018). PPL1 DAN is excited by shock and inhibited by the MVP2 MBON based on a recent experimental study suggesting that such a connection exists (Pavlovsky et al., 2018). In our model PPL1-2 is inhibited by PAM- γ 5, a connection that we propose to exist, and we require to reproduce the appetitive extinction experiment. In chapter 4 we propose an experiment that can validate such a connection exists.

Synaptic parameters		
KC-MVP2 Synaptic weight	$w_{kc-MVP2}$	3 nS
KC-V2 synaptic weight	w_{kc-M6}	3 nS
Ex KC-M6 synaptic weight	w_{kc-M6}	6 nS
Ex KC-M4 synaptic weight	w_{kc-M4}	6 nS
Ex KC-DAN synaptic weight	w_{kc-DAN}	1 nS
Inh MVP2-M6 synaptic weight	$w_{MVP2-M6}$	4 nS
Inh MVP2-M4 synaptic weight	$w_{MVP2-M62}$	4 nS
Inh MVP2-PPL1 synaptic weight	$w_{MVP2-PPL1}$	6 nS
Inh PAM γ 5-PPL1-2 synaptic weight	$w_{PAM5-PPL1-2}$	4 nS

Table 3.2: KC parameters

3.3.3 MBONs and DANs

MBONs and DANs were modelled as conductance based leaky integrate-and-fire neurons like in Chapter 2. For convenience the equations are repeated here. The conductance based LIF equations are as follows:

$$\tau_m \frac{dV(t)}{dt} = (V_{rest} - V(t)) + g_{ex}(t)(E_{ex} - V(t)) + g_{inh}(t)(E_{inh} - V(t)). \quad (3.3)$$

Here, $V(t)$ is the membrane potential of the neuron as a function of t (time), τ_m is the membrane time constant, V_{rest} is the resting membrane potential, E_{ex} is the excitatory reversal potential, and E_{inh} is the inhibitory reversal potential. The synaptic conductances are expressed by g_{ex} (excitatory) and g_{inh} (inhibitory). They are modelled according to the following equations:

$$\tau_{ex} \frac{dg_{ex}(t)}{dt} = -g_{ex}(t) \quad (3.4)$$

and

$$\tau_{inh} \frac{dg_{inh}(t)}{dt} = -g_{inh}(t), \quad (3.5)$$

When the neuron receives an action potential from a presynaptic cell, the postsynaptic conductance increases by the following formulas: $g_{ex} \rightarrow g_{ex} + \Delta g_{ex}$ and $g_{inh} \rightarrow g_{inh} + \Delta g_{inh}$ for excitatory and inhibitory synapses, respectively. Finally, τ_{ex} and τ_{inh} are the synaptic time constants for the excitatory and the inhibitory conductance, respectively. For parameters used to model MBON see Table 3.3. For parameters used to model DANs see Table 3.4.

Parameter	Default Value	Symbol
MBON Membrane Time Constant	90 ms	τ_m
MBON Spiking threshold	-40 mV	V_{th}
MBON Resting membrane potential	-60 mV	V_{inh}
MBON Excitatory reversal potential	0 mV	E_{ex}
MBON Inhibitory reversal potential	-80 mV	E_{inh}

Table 3.3: MBON neuron parameters.

Parameter	Default Value	Symbol
DAN Membrane Time Constant	90 ms	τ_m
DAN Spiking threshold	-40 mV	V_{th}
DAN Resting membrane potential	-60 mV	V_{inh}
DAN Excitatory reversal potential	0 mV	E_{ex}
DAN Inhibitory reversal potential	-80 mV	E_{inh}

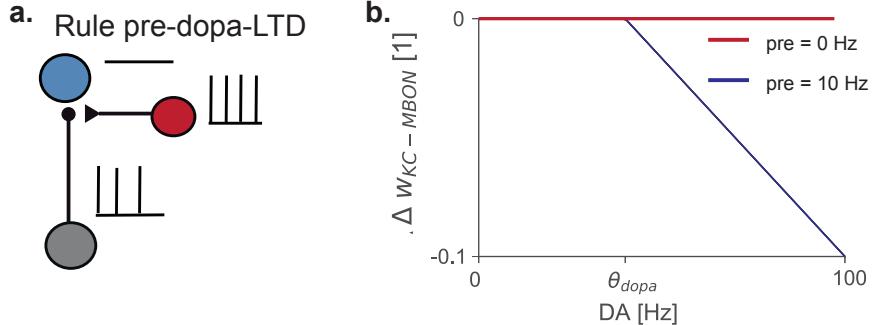
Table 3.4: DAN neuron parameters.

3.3.4 Learning rules

We implemented two types of learning rules. Rule high-dopa-LTD requires only presynaptic activity and dopamine to change the weight between KC and MBON synapses (Fig 5.1). In contrast, Rule R-STDP is part of the three factor learning rules family (Frémaux et al., 2010) which have a Hebbian term that depends on pre and post activity and a third factor to represent the effect of dopamine (Fig 3.3).

3.3.5 Rule 1: high-dopa-LTD rule

Rule 1 is constructed based on experimental evidence showing that postsynaptic spikes appear to be dispensable for learning (Hige et al., 2015b):



$$\Delta w \propto \text{pre} \times \text{dopa} \times \mathcal{H}(\text{dopa} - \theta_{\text{dopa}})$$

Figure 3.2: **a.** Learning rule 1 (dopa-pre-ltd). Rule high-dopa-LTD requires only presynaptic activity and dopamine activity above a threshold. Coincident presynaptic and dopaminergic activity trigger LTD of the the synapse. **b.** Show is change in KC-MBON weight ($w_{KC-MBON}$) when there is either no presynaptic activity (orange) or high presynaptic activity (10 Hz) and variable dopamine activity (weight change only occurs when there is both presynaptic activity and dopaminergic activity above a set threshold (θ_{dopa})).

$$\frac{d}{dt}w_{ij}(t) = -A_- \underbrace{e_j(t)}_{\text{presynaptic}} \underbrace{S_i(t)\mathcal{H}(z_i(t) - \theta_{\text{dopa}})}_{\text{dopamine}} \quad (3.6)$$

where A_- is a positive parameter that determines the maximum weight change a connection can undergo. Synaptic weight change w_{ij} requires the synaptic trace $e_j(t)$ to be above zero and dopamine trace z_i to be above θ_{dopa} which is a threshold set to be equal to background odour activity. The eligibility trace e_j is updated each time the presynaptic neuron j fires as spike as follows:

$$\tau_e \frac{d}{dt}e_j(t) = -e_j + S_j(t), \quad (3.7)$$

where τ_e is the characteristic time, and $S_j(t)$ is the spike train of the j -th presynaptic neuron, described by the following equation:

$$S_j(t) = \sum_k \delta(t - t_{j,k}^*), \quad (3.8)$$

where $t_{j,k}^*$ is the timing of the k -th spike of the j -th neuron. Similarly, $S_i(t)$ is the spike train of the i -th dopaminergic neuron:

$$S_i(t) = \sum_k \delta(t - t_{i,k}^*), \quad (3.9)$$

where $t_{i,k}^*$ is the timing of the k -th spike of the i -th dopaminergic neuron and k is the index of the spike. The dopamine concentration, $z_i(t)$ represents a global eligibility trace that increases each time the dopamine neuron fires a spike:

$$\frac{dz_i(t)}{dt} = -\frac{z_i(t)}{\tau_d} + S_i(t), \quad (3.10)$$

where τ_e is the characteristic time of τ_d and $S_i(t)$ is the spike train of the i -th dopaminergic neuron. We tuned the learning rates parameters (Table 3.5) to reproduce learning experiments we present in the results section.

Parameter name	Symbol	Value
Pre eligibility time constant	τ_e	10 s
MBON MVP2, M4 threshold	θ_{dopa}	10
MBON M6, V2 threshold	θ_{dopa}	200

Table 3.5: Rule 1: high-dopa-LTD parameters

3.3.6 Choosing the threshold for the learning rule

We implemented background level dopamine activity, based on experimental evidence which shows that DANs are active even in the absence of reinforcement (Dylla et al 2014). To subtract baseline activity from the learning rule we add a threshold that is equal to the firing rate of the dopaminergic neuron during odour exposure multiplied by the time constant (τ_z) of the dopaminergic eligibility trace (z_i). This ensures that learning occurs only when DANs fire above baseline activity.

3.3.7 Rule 2:R-STDP Rule

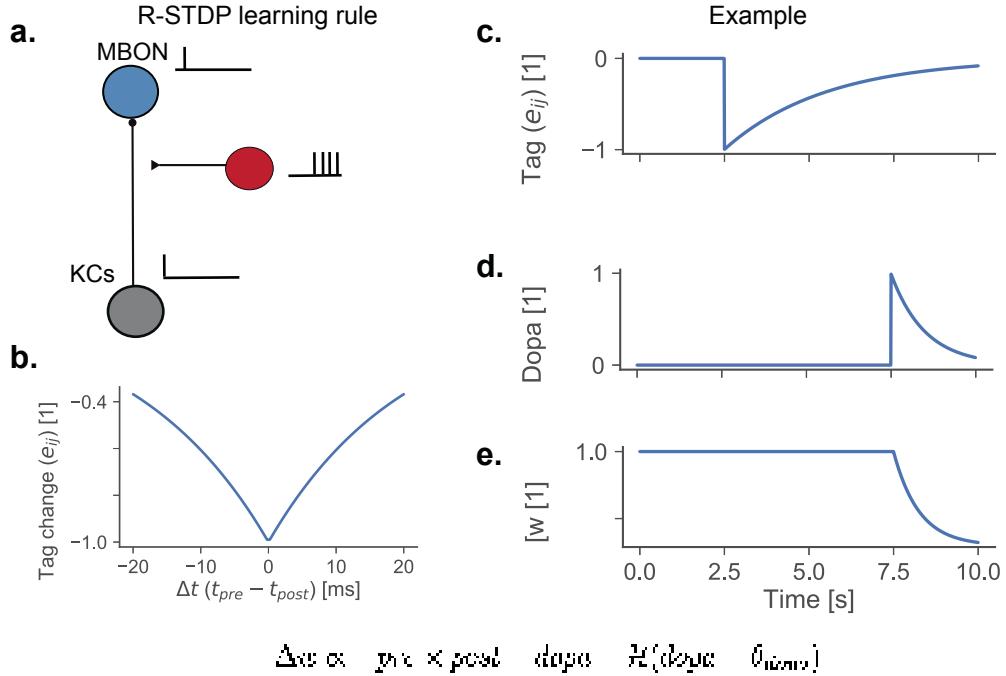


Figure 3.3: **a.** Learning rule 2 (R-STDP) requires presynaptic, postsynaptic and dopaminergic activity to trigger change in KC-MBON weights. **b.** Pre-post spike pairs or post pre-spike pairs within a time window create an eligibility trace (e_{ij}) that decays over time. On the x-axis we show the timing difference between pre-post spike pairs while on the y-axis we show the value of the eligibility trace (e_{ij}). **c.** We show the evolution of the eligibility trace (e_{ij}) after a single pre-post spike pairing event. **d.** We show the evolution of dopamine eligibility trace ($Dopa$) when dopamine spikes are triggered after 7.5 seconds of simulation. **e.** Shown here is the weight between KC and MBONs (w) which changes when both e_{ij} and $dopa$ traces are non-zero.

In traditional Hebbian learning, weight change depends on spikes of both the presynaptic neuron and the postsynaptic neuron. This can be described by a Hebbian function $H(pre, post)$ that can take the form of STDP. Here, We assume that synapses keep track of pre-post activity by updating a synaptic eligibility trace and we base our three factor learning on a previous model by [Izhikevich, 2007](#), with weight dynamics given by:

$$\frac{d}{dt}w_{ij}(t) = -A_- e_{ij}(t) S_z(t) \mathcal{H}(z_k(t) - \theta_{dopa}), \quad (3.11)$$

where like in our Rule dopa-pre-LTD, A_- is a positive parameter that determines maximum weight change. In contrast to our first rule, the synaptic tag $e_{ij}(t)$ depends

on both pre and postsynaptic activity:

$$\tau_e \frac{d}{dt} e_{ij}(t) = -e_{ij}(t) + \underbrace{x_j(t)S_i(t) + y_i(t)S_j(t)}_{H(\text{pre}_j, \text{post}_i)}, \quad (3.12)$$

and the tag (e_{ij}) will increase whenever there is a post spike and the presynaptic trace is above zero or whenever there is presynaptic spike and the postsynaptic trace is above zero. $S_i(t)$ is the spike train of the i -th neuron, given by

$$S_i(t) = \sum_k \delta(t - t_{i,k}^*), \quad (3.13)$$

where $t_{i,k}^*$ is the timing of the spike, where i is the i -th neuron and k is the index of the spike. e_{ij} increases when the postsynaptic neuron fires an action potential and depends on the timing of the last spike timings of the presynaptic neuron through the variable $x_j(t)$ which evolves according to the following equation:

$$\frac{dx_j(t)}{dt} = -\frac{x_j(t)}{\tau_-} + S_j(t) \quad (3.14)$$

The tag (e_{ij}) also increases when the presynaptic neuron fires an action potential if the postsynaptic neuron activity trace is above given by:

$$\frac{dy_i(t)}{dt} = -\frac{y_i(t)}{\tau_-} + S_i(t), \quad (3.15)$$

where τ_- is the characteristic time of the y_i and x_j eligibility traces. When the dopamine eligibility trace (z_k) is above threshold θ_{dopa} and the synaptic eligibility trace (e_{ij}) is above zero the learning rule will depress the weight of the synapse.

The STDP parameters used in the simulations were measured experimentally in vitro in mammalian brains (Abbott et al., 2000) and are described in Table 3.6.

Parameter name	Symbol	Value
STDP depression time constant	τ_-	20 ms
STDP potentiation time constant	τ_+	20 ms
STDP eligibility time constant	τ_e	15 s
Dopamine (z) time constant	τ_z	1 s

Table 3.6: Rule 2: R-STDP parameters

3.3.8 Learning rates

In our model KC-MBONs from different compartment have different learning rate. This is based on experimental evidence suggesting that the KC-MVP2-PPL1 com-

partment has a higher learning rate than other compartments (Aso and Rubin, 2016), but we also tuned the learning rate to reproduce aversive learning and extinction experiments. T We tuned our learning rates (Table 3.5) to reproduce aversive and appetitive learning experiments and we selected a higher learning rate for KC-MVP2 synapses which is based on experimental data that suggests synapses that target the MVP2 MBON have a higher learning rate

Parameter name	Symbol	Value
Learning rate KC-MVP2	η_{MVP2}	0.004
Learning rate KC-V2	η_{V2}	0.002
Learning rate KC-M4	η_{M4}	0.002
Learning rate KC-M6	η_{M6}	0.002

Table 3.7: Rule 1: high-dopa-LTD parameters

3.3.9 Valence and the performance index

Behavioural experiments in the fruit fly are based on a metric called Performance Index. The performance index is defined as the number of flies chose the reinforced odour (CS+) compared to how many flies chose the neutral odour (CS-) normalized by the total number of flies. Our theoretical model is based on the assumption that the fly compares the valences of CS+ and CS- and we use this idea to propose a theoretical performance index based on the two valences. Valence in our model is given by:

$$\text{Valence} = \frac{2r_{V2} - r_{M6} - r_{M4}}{2r_{V2} + r_{M6} + r_{M4}} \quad (3.16)$$

where r_{v2} is the maximum firing rate of MBON-V2 in response to an odour while $r_{M4,M6}$ are the the firing rates of MBON-M6 and MBON-M4 in response to the odour. Notice that we do not use the firing rate of MBON-MVP2 to calculate the performance index. Experimental evidence has shown that MBON-MVP2 does not project to the motor output (Aso et al. (2014)). We calculate the model Performance Index by subtracting the valences of of the reinforced $CS+$ odour and the $CS-$ as follows:

$$\text{Performance Index} = \text{Valence}_{CS+} - \text{Valence}_{CS-} \quad (3.17)$$

In our simulations the valence of the neutral $CS-$ odour does not change, thus the performance index is determined by the the valence of the $CS+$ (Fig 3.4).

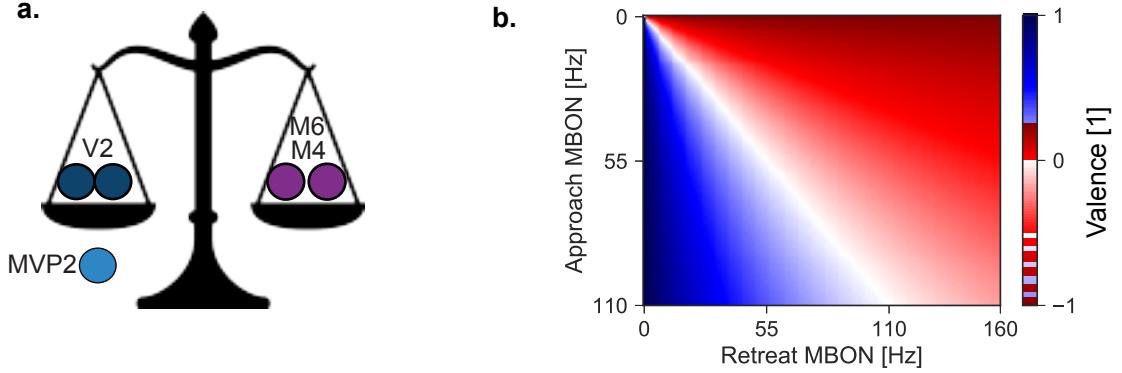


Figure 3.4: Calculating odour valence based on MBON firing rates elicited by the odour. **a.** Valence in our model is calculated by comparing the firing rates elicited by approach biasing V2 MBON to the retreat biasing M4 and M6 MBONs according to Eq 3.16. MVP2 MBON is excluded from calculated the valence but it affects the result by inhibiting M4 and M6 MBONs. **b.** We show the resulting valence calculated for different Approach (V2) and Retreat MBON (M4, M6) firing rates. When Approach MBONs and Retreat MBONs have equal firing rates the resulting valence is zero, corresponding to a neutral state (white diagonal). When Retreat MBON firing rate is higher than Approach MBON firing rate the valence will be negative (red shaded area). Conversely, when Approach MBON firing rate is higher than retreat MBON firing rate the valence will be positive (blue shaded area).

3.3.10 Simulating reward and punishment

To simulate a single shock we activate PPL1 and PPL1-2 neurons by increasing their activity to elicit a firing rate of approximately 60 Hz each time we simulate a shock for 1 seconds. Similary, for reward we increase the firing rate of PAM- γ 5 and PAM2 dopaminergic neurons to approximately 50-100 Hz depending on the magnitude of the reward.

Parameter	Default Value	Symbol
Synaptic parameters		
Input PPL background	r_{ppl_input}	2 Hz
Input PAM background	r_{ppl_input}	2 Hz
Input PPL shock	r_{ppl_input}	10 Hz
Input PAM reward	r_{ppl_input}	10 Hz
Ex PAM- γ 5 background	w_{pam_input}	2 nS
Ex PPL background	w_{ppl_input}	2 nS

Table 3.8: DAN neuron parameters.

3.4 Results

In this section we discuss our results simulating a two layer spiking network of the mushroom body with two different learning rules. We compare our proposed learning rule that depends only on presynaptic and dopaminergic activity to a dopamine modulated Hebbian learning rule (Rule 2 R-STDP) and show that they both can reproduce the aversive learning experiment. We transform MBON firing rate into a theoretical Performance Index and we tune our learning rules to reproduce Performance Index that were measured in aversive learning experiments.

3.4.1 Tuning a spiking olfactory model of the mushroom body

We first tuned our network to elicit approximately equal MBON firing rates to an untrained odour. As MBON-MVP2 inhibits both MBON-M6 and MBON-M4, we increased the connections strength to both MBON-M6 and MBON-M4 (Fig 3.5). We vary the connection strength to both MBON-MVP2 and MBON-M6 to elicit equal firing rates for the two neurons (Fig 3.5). We choose the parameters that yield firing rates of approximately 100 Hz for all of our MBONs (Fig 3.5.). When the connection strength of KC-MVP2 weights decrease the firing rate of MBON-MVP2 decreases while the firing rate of MBON-M6 increases (Fig 3.5 a).

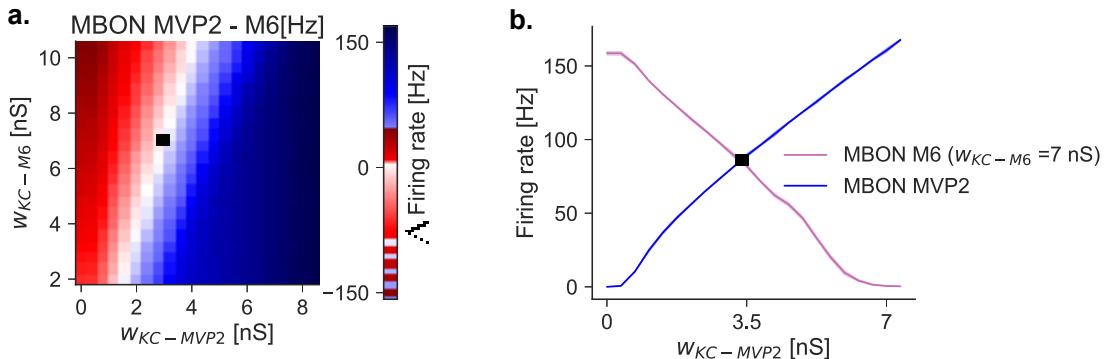


Figure 3.5: *Tuning KC-MBON connections to yield equal firing rates for retreat and approach MBONs.* **b.** As in a. we show the relationship between MVP2 and M6 firing rate for different tuning of initial connection strengths. On the x-axis we show the KC-MVP2 connection strength while on the y-axis we show the KC-M6 connection strength. Red circle denotes the tuning value we chose in our simulations, which yields equal MVP2 and M6 firing rates in response to a stimulus. Black square denotes the parameter values we chose in our simulations.

During spontaneous activity KCs fire at approximately 0.1 Hz and DA neurons fire at 5 Hz. Our learning rules implement a threshold for dopamine activity as described

in Eq 3.6. In all of our simulations background dopamine activity rates do not modify KC-MBON weights (Fig 3.6).

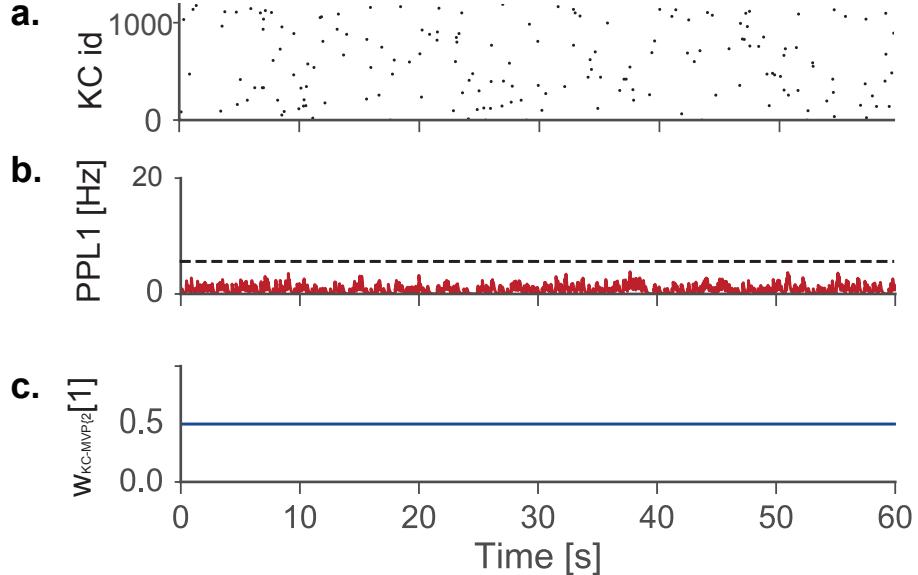


Figure 3.6: No changes in KC-MBON weights occurs during spontaneous activity. Showing spontaneous network activity with no odour present. **a.** We show a raster plot of 1000 KCs during 60 seconds of spontaneous activity. Each point represents an individual spike. **b.** We show background dopamine levels of PPL1 DAN. simulating background KC (top) activity and background dopamine (middle). **c.** We show KC-MVP2 mean weight during simulation of spontaneous activity.

3.4.2 Dopa-pre-LTD rule can reproduce aversive learning and extinction experiments

In this section we present results we obtained from testing Rule high-dopa-LTD. We simulated 60 seconds of an odour paired with 12 shocks (CS+) followed by 60 seconds of unpaired second odour exposure (CS-), which is based on the aversive learning experiment (Felsenberg et al. (2018)). In our simulation odour presentation led to a transient increase of KC population firing rate (Fig 3.7 a) lasting for 1 second and decaying afterwards, while simulation of shocks increased the firing rate of PPL DANs (Fig 3.7 d) each time a shock was presented. KC activity at odour onset created an eligibility trace (Fig 3.7 f) with $\tau_j = 10s$ while shock presentation increased the dopamine eligibility trace above the threshold for plasticity (Fig 3.7 d). The dopamine and presynaptic traces decrease CS+ KC-MVP2 average weight (Fig 3.7 d). After odour exposure, MBON responses to CS- were unchanged (Fig 3.7 d).

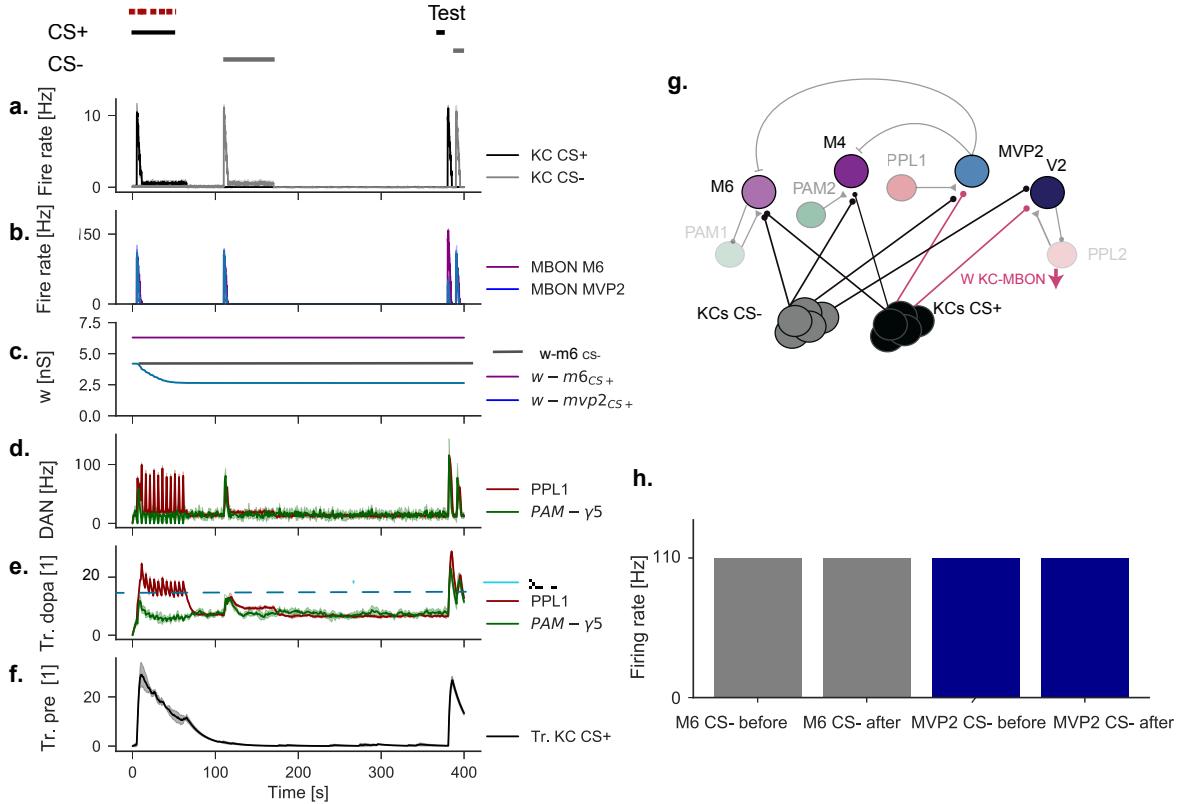


Figure 3.7: Network activity during aversive training. Black bar on the top denotes CS+ exposure duration, while grey bar represents CS- exposure. Red dots represent shock exposure. After training and CS- exposure, CS+ and CS- odours are re-exposed to test MBON responses with plasticity de-activated. **a.** Population activity of KCs responding to CS+ (black) and CS- (grey) odours during learning and testing. **b.** Firing rates of MBONs MVP2 and M6 during the simulation. **c.** Aversive training using high-dopa-LTD depresses CS+ KC-MVP2 (blue) weights while CS+ KC-M6 (purple) and CS- KC-MVP2 weights remain unchanged. **d.** We show the firing rates of PPL1 and PAM- γ 5 DAN during the simulation. **e.** We show PPL1 DAN and PAM- γ 5 DAN eligibility traces during the simulation. Dashed line shows the threshold (θ_{dopa}) for enabling plasticity at KC-MVP2 synapses **f.** Shown here is the eligibility trace of one CS+ responding KC with $\tau_j = 10s$. Shading indicates 1 standard deviation. **g.** A circuit diagram of the network is shown with two populations of KCs responding to CS- (grey) and CS+ (black). CS+ and CS- responding KCs connections to MBONs are shown as circled arrows. Black arrows show connections that are unchanged after aversive learning while red arrows denote connections that have experienced depression. **h.** Firing rates of MVP2 and M6 MBONs to CS- odour before and after odour exposure.

After training, odour response to the CS+ changed bi-directionally: MVP2 MBON decreased, while M6 MBON response increased compared to CS- which remained unchanged (Fig 3.7 a.). This is qualitatively similar to experimental results (Owald

et al., 2015, Felsenberg et al., 2018). After exposure MBON odour responses to the CS- remained unchanged (Fig 3.7 b) We tuned out network so that the changes in MBON MVP2 firing rate match experimental data (Fig 3.8 a). We also tuned our learning rule to obtain a theoretical performance index that matches aversive learning experiments (Felsenberg et al. (2018)), (Fig 3.8 c).

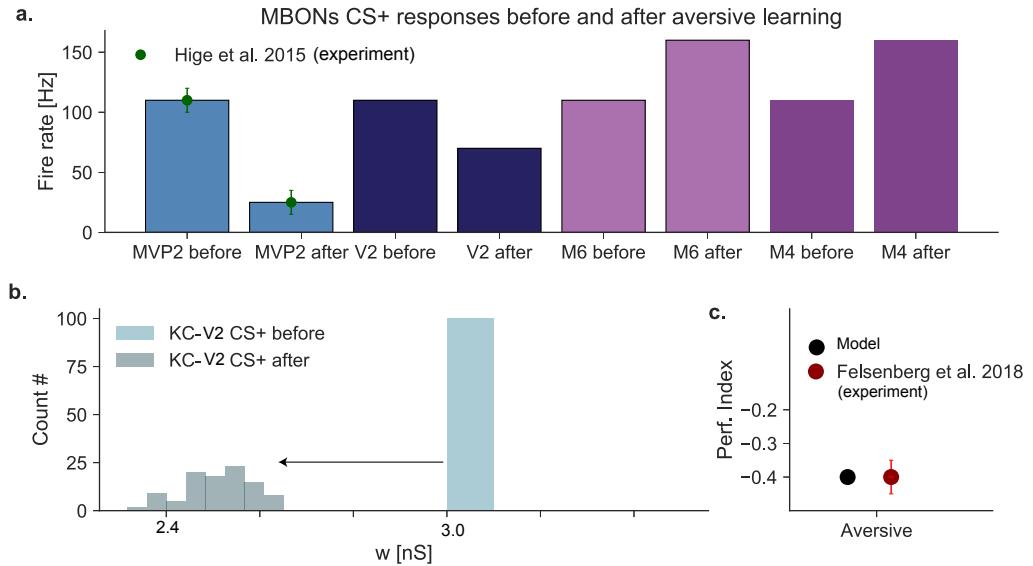


Figure 3.8: MBON firing rates, Performance Index and weights after aversive learning. **a.** We show firing rates for MBON MVP2 (light blue), V2 (dark blue), M6 (light purple), M4 (dark purple) before and after aversive learning simulation. **b.** KC-V2 connection weights for KCs that respond to the CS+ are shown before (dark blue) and after (dark blue) aversive learning. **c.** Model performance index and experimental data from Felsenberg et al., 2018 (dark red) is shown after aversive training. Red bar indicates 1 standard deviation.

We tested different time constants for the presynaptic eligibility trace $\tau_j = 0.1, 2, 10s$ and we found that increasing the time constant increased the the change in the firing rate of MBON MVP2 (Fig 3.9 a). Our results show that increasing the number of shocks leads to a stronger memory performance, our theoretical Performance Index decreases monotonically with number of shocks, which has been shown to occur by experiments (Beck et al., 2000), (Fig 3.9 b).

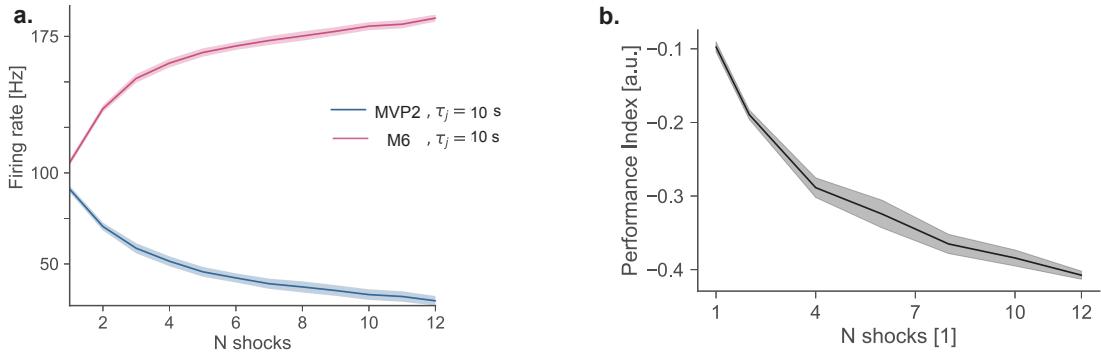


Figure 3.9: Changes in the MBON firing rates as number of shocks during training are increase using learning rule pre-dopa LTD. **a.** MBON MVP2 and MBON M6 firing rates after aversive training with increasing number of shocks, $\tau_j = 10$ s. **b.** Model performance index for aversive training with increasing number of shocks. Shaded areas illustrate 1 standard deviation.

By increasing τ_e we also increased the change in the firing rates of MBON MVP2 at the time of the arrival of the last shocks (Fig 3.9 a.). We now investigate the effect of changing the duration of KC onset response during odour simulation to 1, 5 , 10 , 15 , 20 and 30 seconds. Increasing the KC odour onset response increase the overlap with shocks that arrive with different delays compared to odour onset (5s to 60 s delay). We also tested different time constant for the presynaptic eligibility trace and ($\tau_j = 1$ s to 25 s). Increasing the τ_e enhanced the change in the firing rate of MBON-MVP2 for shocks the final shocks which arrives 60 seconds after odour onset (Fig 3.10).

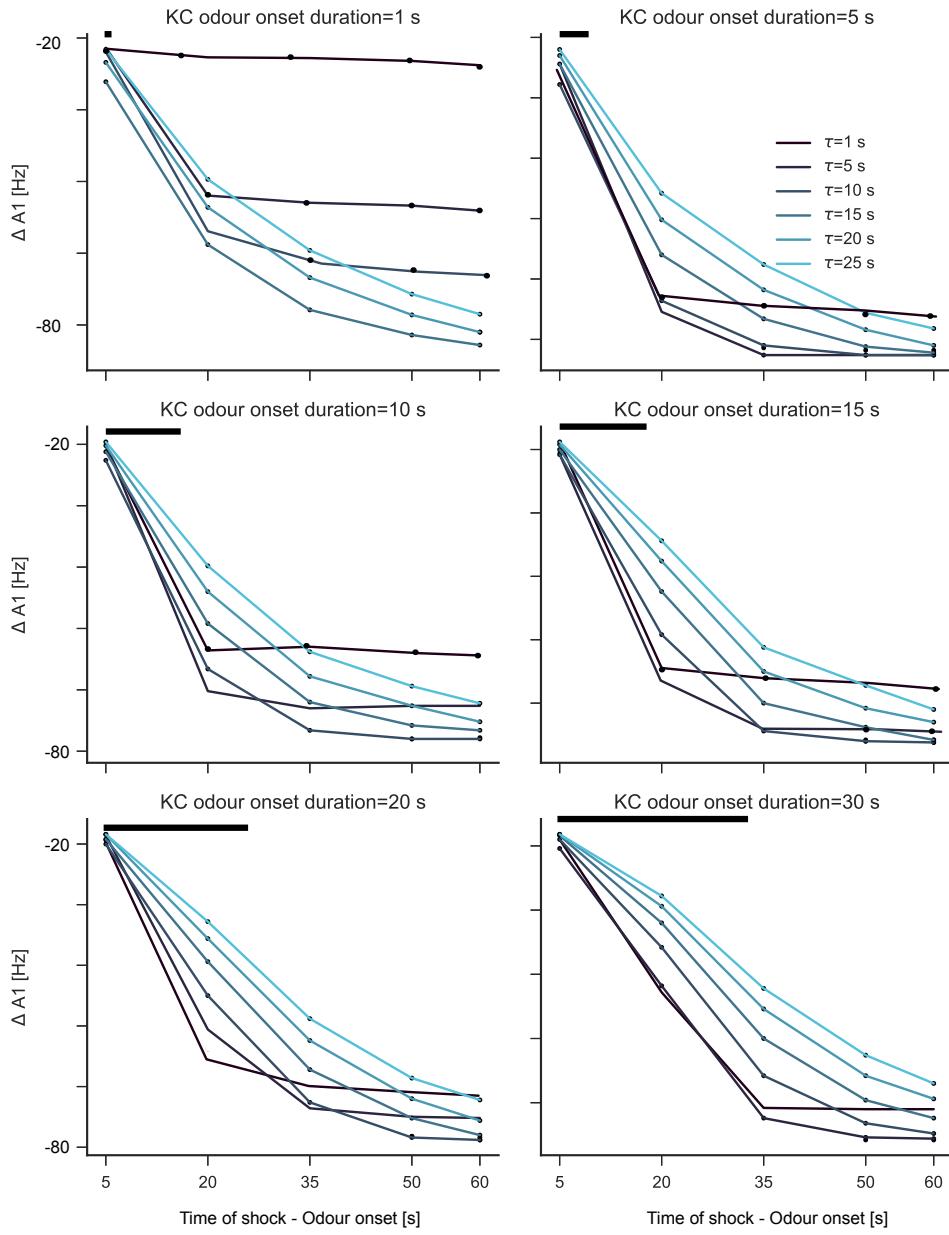


Figure 3.10: MBON MVP2 firing rates changes during 60 second odour exposure training with 12 shocks. X-axis shows the time of shock relative to the start of odour exposure while y-axis denotes the effect the shock on the MBON-MVP2 firing rate. Curves show results for simulation with different presynaptic activity trace time constants ($\tau_j=1$ to 25 s). Black bar denotes KC odour onset response: 1 to 30 s.

3.4.3 Aversive extinction requires connections between opposing valence coding compartments

We now investigate the aversive extinction protocol which involves simulating aversive learning followed by re-exposing the same odour 5 times without the punishment. As we showed in the previous section, after aversive learning the firing rates of MVP2,V2 approach MBON decreases while the firing rate of M6 and M4 retreat MBONs increased. Similarly, the firing rate of $PAM - \gamma 5$ increased. After aversive learning, when we re-exposed the odours the average KC-M6 CS+ weight decreased each time we presented the CS+ odour without shock (Fig 3.11 c).

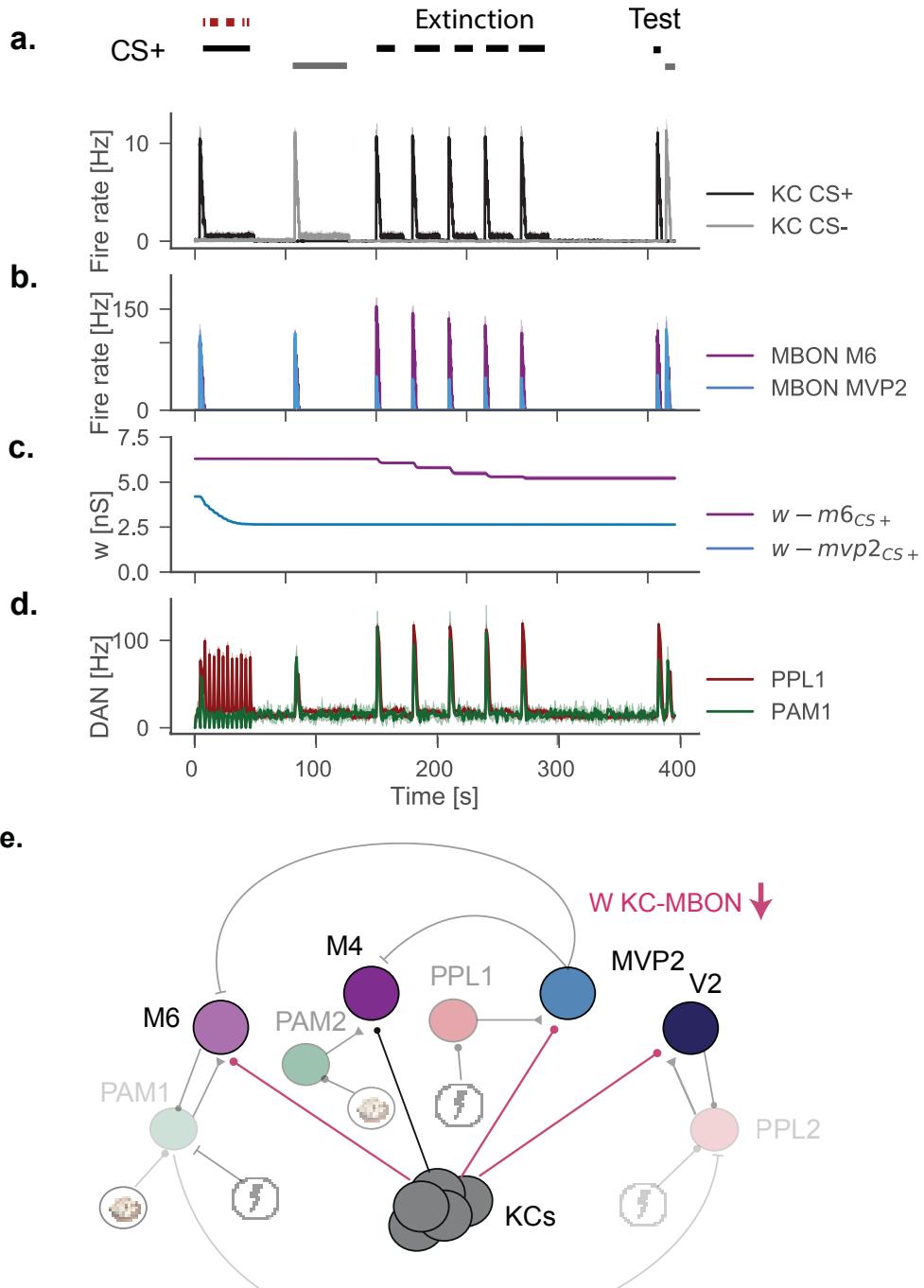


Figure 3.11: Network activity during aversive extinction learning protocol. CS+ odour (black) is exposed for 60s, followed by 45 pause, 60 s of CS- exposure and 5 CS+ re-exposures without shock. **a.** Population activity firing rate of KCs responding to CS+ and CS- stimuli during learning and testing. **b.** Firing rates of MBONs MVP2 and M6 during the simulation. **c.** Aversive training using M6 depresses CS+ KC-MVP2 (blue) weights while KC-M6 (purple) remain unchanged. **d.** Firing rates of PPL1 and PAM- γ 5 DAN during the simulation. **e.** We show the circuit representation and connections in response to CS+ after aversive learning. Black lines from KC to MBONs are unchanged connections while red lines are connections that underwent LTD.

After aversive extinction the firing rate MBON M6 decreased from 160 Hz to 110 Hz while the firing rate of MBON M4 remained unchanged (Fig 3.11 b.). We tuned plasticity to change the theoretical performance index to match closely the experimental data (Fig 3.12 c. black and red circles). We observe that aversive extinction is partial, decreasing from -0.4 to -0.3 and that only one retreat coding MBON changes its firing rate (MBON M6 but not MBON M4). Aversive learning leads to the decrease of KC-MVP2 weights (Fig 3.11 b.) while aversive extinction decreases KC-M6 weights (Fig 3.11 b.).

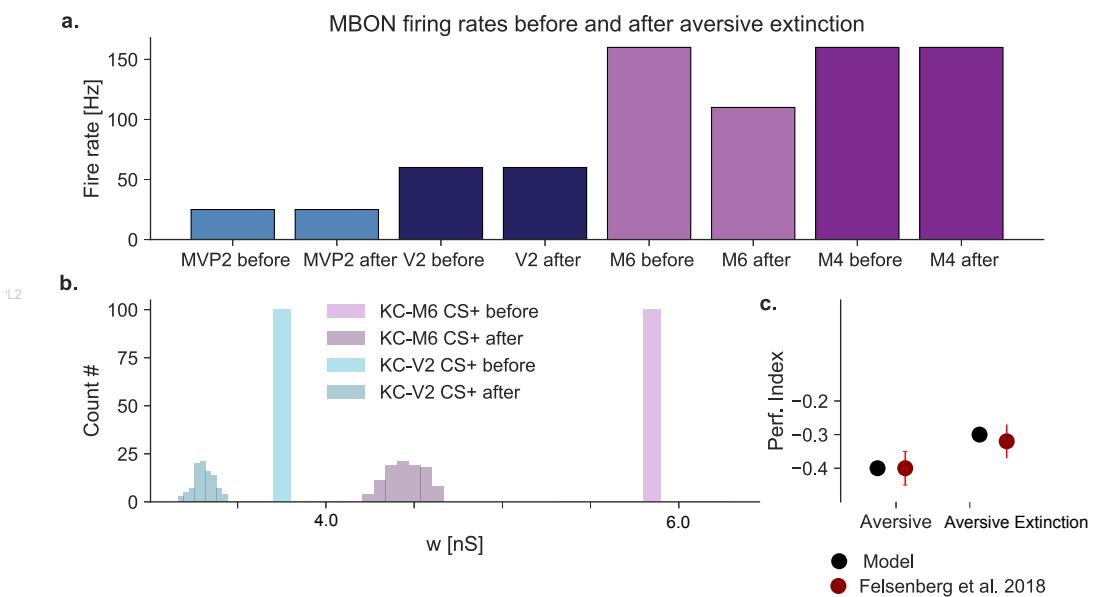


Figure 3.12: MBON firing rates, Performance Index and weights after aversive extinction. **a.** MBON MVP2 (light blue), V2 (dark blue), M6 (light purple), M4 (dark purple) are shown before and after aversive extinction simulation. **b.** KC-M6 connection weights for KCs that respond to CS+ are shown before (light purple) and after (dark purple) aversive extinction. **c.** Model (black) performance index and experimental performance index based on data from Felsenberg et al. (2018) (dark red) is shown after aversive training and after aversive extinction.

We now test the effect of aversive learning and extinction learning on the activity of PAM- γ 5 which is excited by MBON M6 in response to the trained odour. Our results show that aversive learning leads to the increase of both MBON M6 and PAM- γ 5 activity in response to CS+ while aversive extinction reduces their responses (Fig 3.13 a,b).

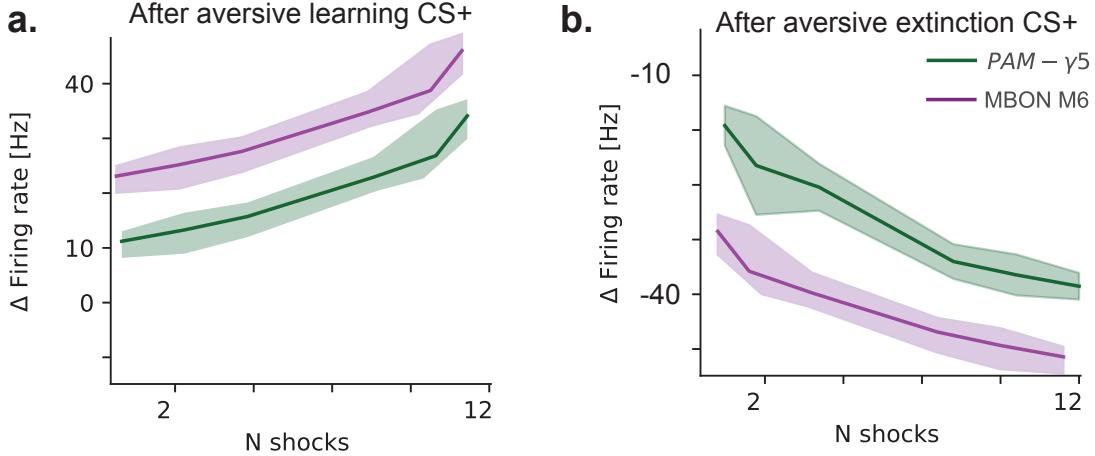


Figure 3.13: **a.** We show the changes in the responses of MBON M6 and PAM- γ 5 neurons during odours exposure after aversive learning and extinction as we increase the number of shocks. **b.** We show the responses of MBON M6 and PAM- γ 5 after aversive extinction of aversive memories of different number of shocks.

In our model MBON-M6 excites PAM- γ 5. The firing rate of the MBON M6 decreased monotonically with the number of extinction trials when $\eta_{m6} = 5e - 7$. When we increase the learning rate for KC-M6 connections (η_{m6}) the effect of subsequent extinction trials was reduced (Fig 3.14). In the case of a high learning rate $\eta_{m6} \geq 1e - 6$ only the first extinction trial modified the firing rate of MBON M6.

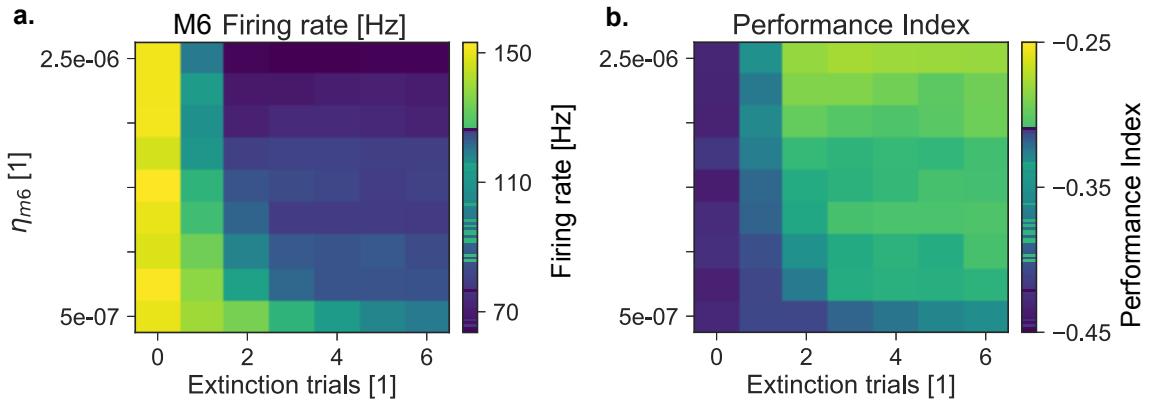


Figure 3.14: Changes in the MBON M6 firing rate as number of extinction trials is increased. **a.** Colormap shows change in the firing rate of MBON-M6 after aversive extinction with increasing number of extinction trials and increasing learning rate ($\eta_{PAM - \gamma 5}$) for the KC-MBON-M6 connections **b.** We show the change in the performance index for the same configuration in a.

3.4.4 R-STDP rule can reproduce aversive extinction learning experiment

We also tested Rule R-STDP based on Eq. 3.11 and we obtained very similar results to using Rule 1. KC and MBON activity in the first second of odour exposure create an eligibility trace (e_{ij}) (Fig 3.15 e.) which modifies the weight when shocks are present (Fig 3.15 c, d) After aversive learning MVP2 decreases its response to CS+ and M6 and PAM- γ 5 increase their CS+ responses (Fig 3.11 b,d) KC-M6 CS+ average weight decreases during the 5 odour presentations after aversive learning which reproduces the aversive extinction experiment (Fig 3.15 c)

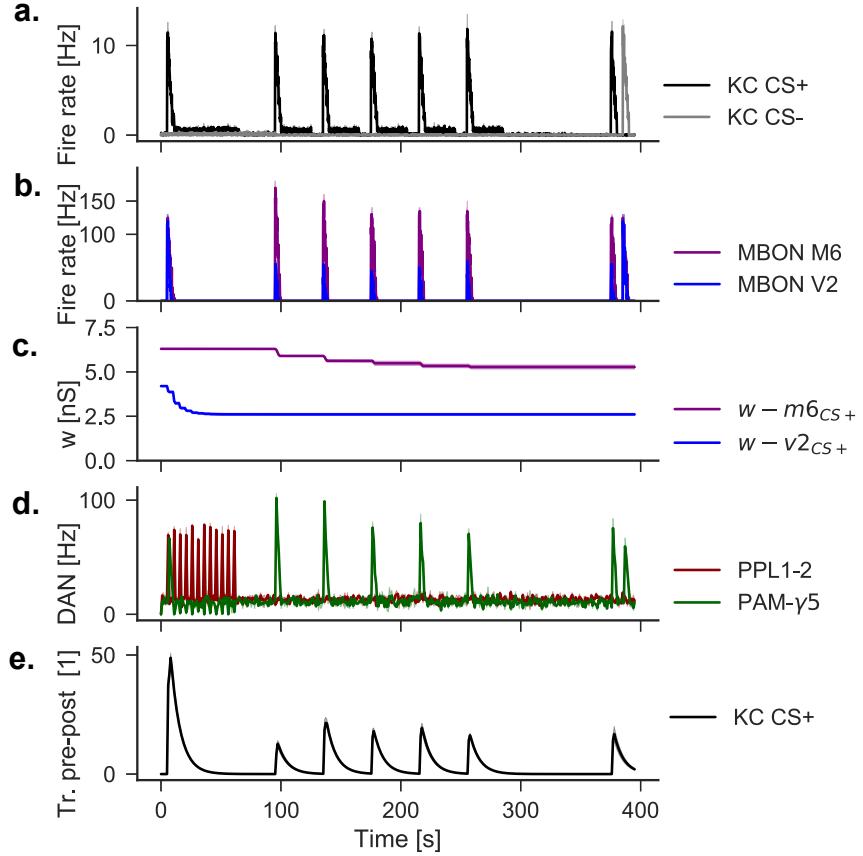


Figure 3.15: Network activity during appetitive extinction simulation with R-STDP learning rule **a.** Shown here are population firing for CS+ responding KCs (black) and CS- responding KCs (grey) **b.** MBON MVP2 (blue) and M6 (purple) firing rates during the simulation. **c.** Average CS+ KC-MVP2 weight (light blue) and average CS+ KC-M6 weight (purple) shown during simulation. KC-MVP2 weight change. **d.** Panel shows the firing rates of PAM- γ 5 (green) and PPL1 (red) DANs during the simulation. **e.** Shown here is a pre-post e_{ij} tag KC-MBON synapse that responds to the CS+ with $\tau_e = 10s$. As there are no MBON spikes e_{ij} is zero throughout the simulation

3.4.5 high-dopa-LTD rule can reproduce aversive learning experiment with no post spikes

We simulated aversive learning using both Rule 1 (high-dopa-LTD) and Rule 2 (R-STDP) while blocking MBON activity completely to reproduce experimental results (Hige et al., 2015a) that show plasticity occurs in the absence of postsynaptic spikes. Our results show that Rule 1 (high-dopa-LTD) can modify the firing rate of MBON MVP2 without postsynaptic spikes (Fig 3.16 a.) while Rule 2 (R-STDP) fails to do so (Fig 3.16 b.). This is because the dopa-pre-LTD learning rule has no postsynaptic

component and depends only on presynaptic and dopaminergic activity. Our results also show that while aversive learning can be reproduced , aversive extinction fails if we block the MBON during the extinction trials since in this case it requires feedback from MBON to DAN.

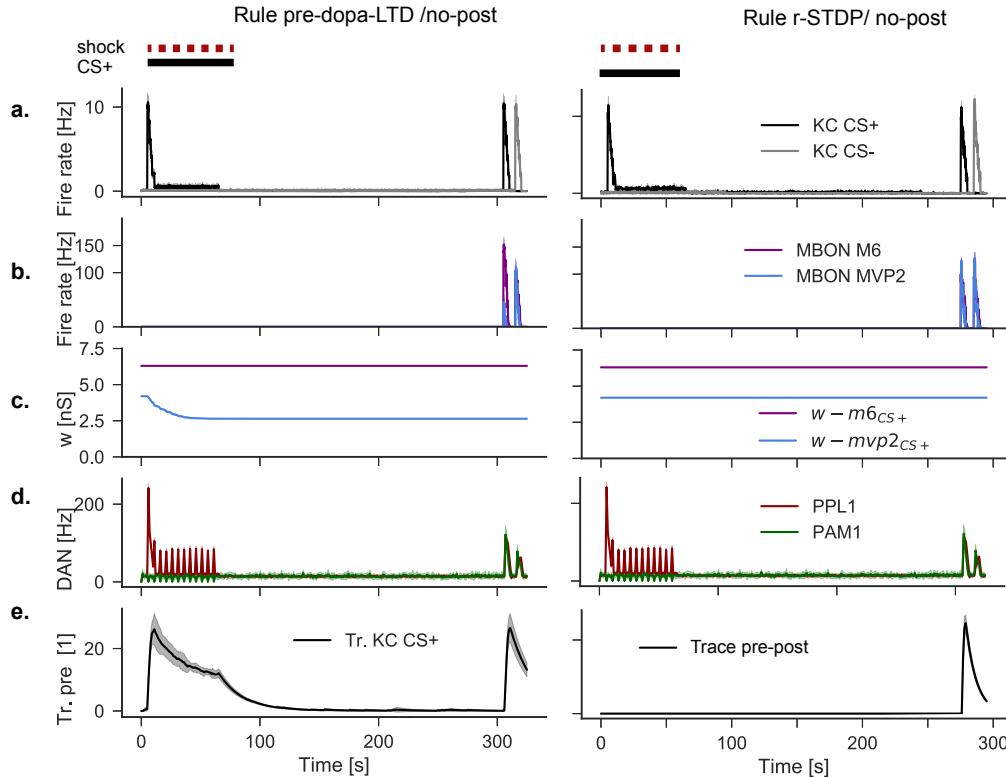


Figure 3.16: Network activity during aversive training with no postsynaptic spikes using Rule 1 (pre-dopa LTD, left) and Rule 2 (R-STDP, right). **a.** KC population firing rates in response to CS+ and CS-. **b.** MBON MVP2 and M6 firing rates during the simulation. **c.** KC-MVP2 and KC-M6 CS+ weight change during simulation. **d.** Panel shows the firing rates of PAM- γ 5 and PPL1 DANs during the simulation. **e.** Panel shows the KC eligibility trace of one KC that responds to the CS+ with $\tau_j = 10s$ (left) and an eligibility trace of one pre-post pair (e_{ij}) on the right.

3.4.6 Aversive re-learning decreases only MBON V2 firing rate

We also simulated aversive re-learning immediately after extinction. A series of 12 shocks were administered 30 seconds after aversive extinction in our simulation (Fig 3.17). During aversive re-learning KC-V2 decreased further than during aversive learning while KC-MVP2 weights remained unchanged (3.17 c.).

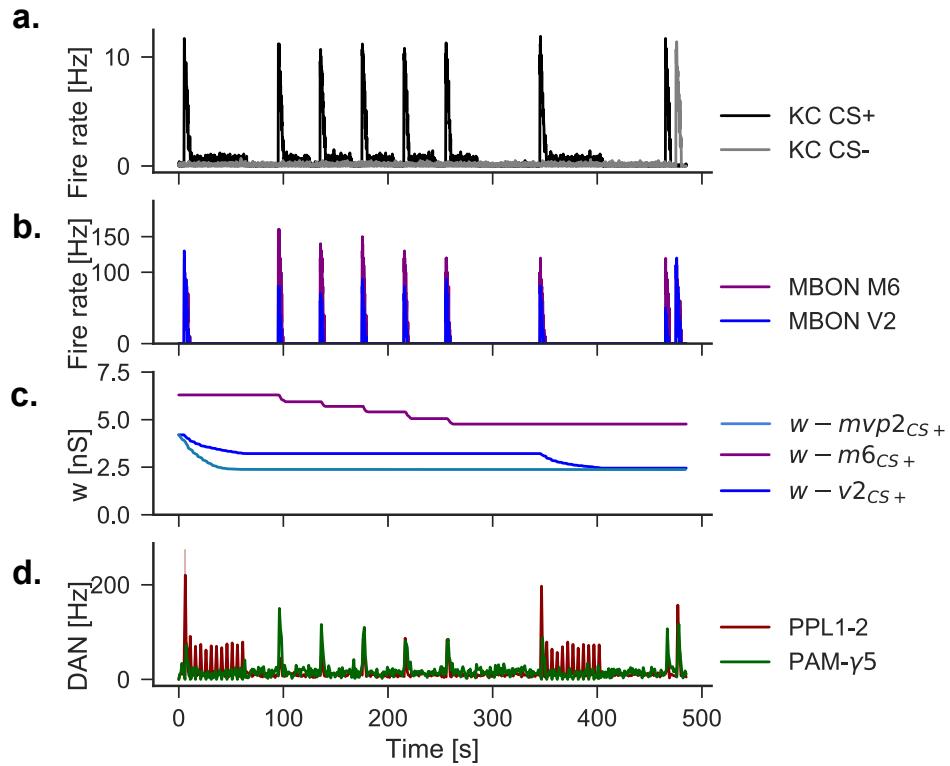


Figure 3.17: Network activity during aversive re-learning and extinction. **a.** Population activity firing rate of KCs responding to CS+ and CS- stimuli during learning and testing. **b.** Firing rates of MBONs V2 and M6 during the simulation. **c.** We show the average CS+ KC-V2 (blue), CS+ KC-MVP2 (light blue) and KC-M6 (purple) weights during the simulations. **d.** Firing rates of PPL1 (red) and PAM- γ 5 (green) DANs during the simulation.

After aversive re-learning only the V2 MBON decreased its firing rate compared to before re-learning (Fig 3.18 b.). We obtained a similar Performance Index to the result observed experimentally when we converted CS+ firing rates into a performance index using Eq. 3.16 (Fig 5.5 c.).

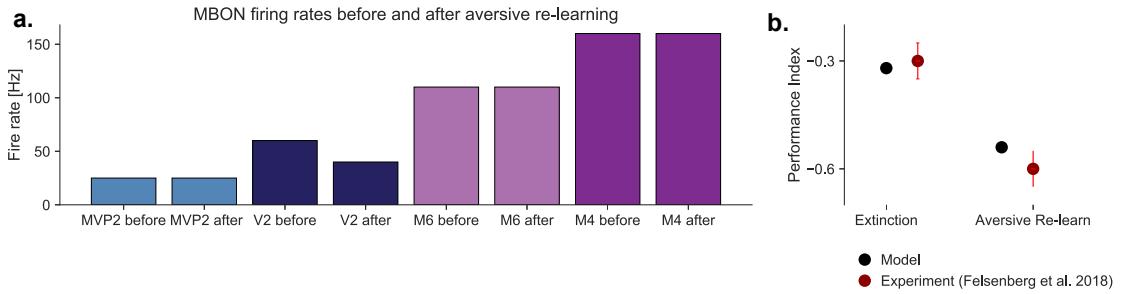


Figure 3.18: MBON firing rates, Performance Index and weights after appetitive extinction. **a.** MBON MVP2 (light blue), V2 (dark blue), M6 (light purple), M4 (dark purple) are shown before and after aversive extinction simulation. **b.** Model performance index and experimental data from Felsenberg et al. 2018 (dark red) is shown after aversive extinction and after aversive re-learning and our resulting performance index calculated from MBON firing rates is shown in black.

3.4.7 Second order conditioning forms two opposing memories

We return to our model that we used in our previous Chapter 3 and we discuss second order conditioning. We have shown that after aversive learning two DANs increase their responses to the aversive odour: PPL1 DAN (weaker inhibition from MVP2 MBON) and PAM- γ 5. We tested whether we could use an aversive CS+ odour to replace shock and we simulated 4 trials of CS2 followed by CS-. Each time we obCS+ odour was exposed after the CS2 odour we observed depression for both KC-MVP2 CS2 weights and KC-M6 CS2 weights (Fig second order c.). After the second order conditioning protocol we calculated a theoretical performance index ($PI = -0.02$).

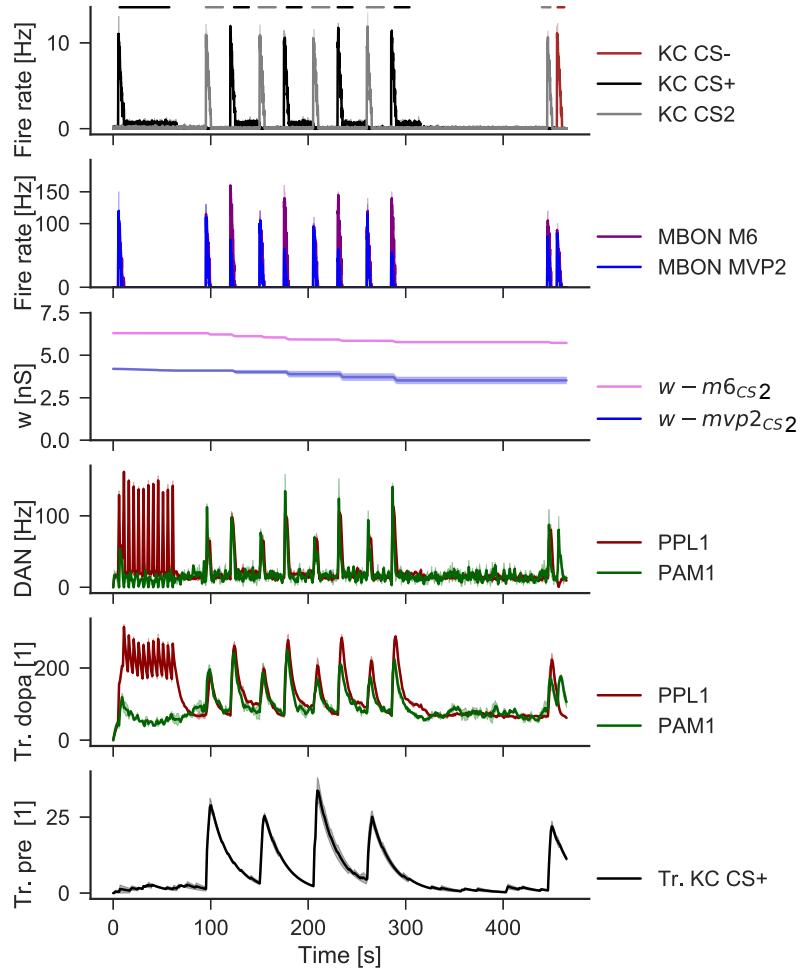


Figure 3.19: Network activity during aversive re-learning protocol using rule 4: low dopa LTP. **a.** KC population firing rates in response to CS+ and CS-. **b.** MBON MVP2 and M6 firing rates during the simulation. **c.** KC-MVP2 and KC-M6 CS+ weight change during simulation **d.** Panel shows the firing rates of PAM1 and PPL1 DANs during the simulation. **e.** Panel shows the eligibility traces of PAM1 and PPL1 DANs during the simulation. **f.** Panel shows the KC eligibility trace of one KC that responds to the CS+ with $\tau_j = 10s$.

3.5 Discussion

In this chapter, we studied the effects of dopamine modulated learning rules on the firing rate of MBONs in a two layer spiking neural network. To study the link between changes in MBON firing rates and changes in behaviour we first tuned out network to elicit approximately equal firing rates for all MBONs for an untrained odour (Fig 3.5). We also introduced a transformation function from MBON firing rates into a

predicted performance index and obtained similar results in our model to experimental data (Fig 3.4).

While the performance index is a behavioural measurement of number of flies that move towards the reinforced odour and depends on multiple factors such as the state of the fly (hungry satiated or thirsty) our model performance index is a simplification that assumes the difference between approach and retreat biasing MBON determines the behaviour of the fly.

Fly experiments have shown that shocks that arrives tens of seconds after stimulus onset increase the strength of memory (they affect outcome of plasticity at KC-MBON synapses). To reproduce this observation, we tuned our learning rule so that dopaminergic activity that occurs at the end of odour exposure can change the synaptic connections. We simulated different odour response profiles for KCs while also modifying the duration of the time constant of the presynaptic trace. Our odour response profile and duration of eligibility trace is similar to experiments that showed KCs respond transiently to odour (Murthy et al., 2008) and maintain a trace of the odour that allows the fly to learn even if the shock arrives after odour exposure has finished (Dylla et al., 2017).

We showed that an LTD-only learning rule that requires only presynaptic and dopaminergic activity can reproduce olfactory learning experiments in the fruit fly. We also tested a dopamine modulated STDP rules (r-STDP) and found that r-STDP fails at reproducing fruit fly experiments where MBON spikes were blocked during odour exposure paired with dopamine release (Fig 3.16) (Hige et al., 2015a).

We also reproduced experimental measurements of MBON-MVP2 ($MBON - \gamma 1pedc$) before and after an odour was paired with dopamine activation. By adding an inhibitory connections from MVP2 MBON onto M6 and M4 MBONs we can also explain the increase in the activity of retreat MBONs after aversive learning using an LTD only learning rule (Fig 3.8)(Owald et al., 2015). In our simulation aversive simulating the extinction experiment requires MBON activity together with excitatory feedback from M6 MBON on PAM- $\gamma 5$, which has been observed experimentally (Felsenberg et al., 2018). We have shown that aversive extinction only partially extinguishes memory by measuring the performance index. Similarly, experiments have shown that only M6-MBON decreases after aversive extinction while the M4-MBON remains unchanged (Fig 3.12 a.).

Our theoretical performance index required a decrease in M4 MBON to achieve complete aversive extinction. We propose that partial aversive extinction can be explained by the fact that only a subset of retreat MBON decrease their activity

during aversive extinction. In the case of the aversive extinction simulations we used only connections that were shown to exists by experiments (Felsenberg et al., 2018). We found that we could also reproduce performance index measurements obtained after aversive re-learning. However, this was possible by decreasing the activity of the approach V2 MBON. Re-extinction would be impossible in our proposed framework as it requires the increase in the firing of the M6 MBON. After aversive learning M6 MBON was dis-inhibited by the MVP2 MBON. This mechanism can not work multiple times as the firing rate of the MVP2 MBON will reach zero spikes. In Chapter 4 we investigate a mechanism for potentiation that increases the flexibility of our computational model to re-learn and re-extinct multiple times.

Dopamine modulated plasticity is a promising framework for explaining experimental results in the brain of the fruit fly and shows that plasticity models based on pre- and post is not sufficient to explain current experiments. While, post-synaptic spikes may not be explicitly required in the learning rule they appear to be playing an important role in signalling DAN changes in the internal representation of the valence of an odour. In Chapter 4 we will show that our model can also reproduce appetitive extinction experiments and we will investigate why potentiation is required in the fly system.

Chapter 4

Appetitive learning and extinction in the mushroom body

4.1 Ch4: Abstract

In this chapter, we investigate the mechanism behind appetitive learning and extinction. In chapter 3, we showed that a learning rule that depends only on presynaptic spikes and dopaminergic activity can reproduce aversive learning and extinction experiments. In this chapter we show that the same model can also reproduce appetitive learning and extinction experiments. After appetitive learning the activity of aversive PPL1-2 DAns is enhanced through dis-inhibition. Odour re-exposure without the expected reward triggers plasticity in the V2 compartment, which decreases the firing rate of approach biasing MBONs. Appetitive extinction in our model requires appetitive PAM- γ 5 DAns to inhibit aversive *PPL1 – 2* DAns. To test whether such a connection exists, we predict that blocking PAM- γ 5 neurons during odour exposure will create an aversive memory. Finally, we show that in both aversive and appetitive learning cases, extinction is blocked in the reinforcement that was used to train the odour is present during re-exposure to odour.

4.2 Introduction

The mechanism behind aversive extinction is better understood than for appetitive extinction. Aversive learning reduces the activity of the MVP2 MBON that inhibits M4/6 retreat MBONs increase their activity in response to the trained odour through disinhibition. The M6 retreat MBON has been shown to excite the *PAM – γ 5*

dopaminergic neurons in its own compartment, which could allow an odour to trigger plasticity during extinction([Felsenberg et al., 2018](#)). After aversive extinction, a subsequent training session of the odour paired with the shock leads to re-learning of the aversive association. Thus far in this thesis, we have shown that dopamine modulated learning rules that only depress KC-MBONs synapses are sufficient to explain learning aversive learning and extinction experiments.

Here, we show that our model can also reproduce appetitive learning and extinction experiments. Reward memories are stored in compartments that are targeted by dopaminergic neurons that are activated by reward([Aso et al. \(2014\)](#)). During reward learning appetitive dopaminergic neurons depress the odour evoked activity of retreat biasing MBONs. Re-exposure to the appetitive odour has been shown to lead to the complete extinction of the memory ([cite felsenberg 2017](#)). Furthermore, appetitive memory extinction appear to require the activity of aversive dopaminergic neurons. In contrast to aversive extinction, the exact mechanism for appetitive extinction is currently unknown. In this chapter we will show that our model can also reproduce appetitive extinction experiments.

As extinction experiments have shown, the fly is capable of assigning a valence to an odour and re-evaluate that valence when the odour is re-exposed in future trials. The mechanisms for extinction can be explained by reducing the activity of MBONs that signal opposite behaviours. It is still unclear when KC-MBON synapses potentiate. Studies have shown in the $\gamma 4$ compartment when high dopamine activity levels precede Kenyon Cell activation, the activity of the MBON is potentiated. This could potentially explain relief learning, an experimental paradigm that has shown that when shock precedes an odour an appetitive memory forms ([Yarali et al., 2008](#)). Furthermore, it is sufficient to activate PPL1- $\gamma 1$ ped DAN which innervates KC-MVP2 synapses to observe this effect. Here, we propose to explain this effect by proposing a learning rule that will trigger potentiation or depression depending on the timing of dopamine and odour onset.

In theoretical neuroscience papers, synaptic depression and potentiation have been proposed as mechanisms of achieving stability in a network. If synapses can only potentiate they could . Conversely, if synapses only experience depression then eventually neurons will become completely silent. Aversive extinction experiments have shown that the M6 MBON increases its activity after aversive learning through disinhibition. We have proposed that the M6 MBON drives extinction by activating the *PAM* – $\gamma 4$ DANs above their baseline activity level. However, after aversive extinction the activity of the M6 MBON appears to return to baseline level. If flies can

re-evaluate information and change the valence of an odour multiple times, how does the fly perform re-extinction? Or to ask this question at a mechanistic level, how does the M6 MBON signal re-extinction in subsequent trials if there is conflicting information. If the M6 MBON enhances its activity through MVP2 mediated dis-inhibition, eventually the MVP2 MBON will reach its minimum level which doesn't allow the flexibility for re-evaluating the valence of an odour multiple times. Here, we propose a learning rule that potentiates KC-M6 MBON synapses during re-learning and can perform re-extinction.

Learning experiments have also shown that an aversive odour can also be used to train a neutral odour to become aversive in a paradigm called second-order conditioning ([Tabone and de Belle, 2011](#)). In the study, the authors have shown that second-order conditioning

Here we introduce two learning rules that implement potentiation of KC-MBON synapses with different mechanisms. The first rule (dopa-pre-LTP) potentiates KC-MBON synapses when shock arrives before odour. This allows us to reproduce the relief experiment by ([Yarali et al., 2008](#)) that shows the fly will learn opposite behaviour depending on the order of dopamine and shock during learning. Relief learning explain

Our second learning rule (low-dopa-LTP) potentiates KC-MBON synapses at very low dopamine levels. Low dopamine levels in our model occur during odour exposure after extinction or if dopaminergic neurons are blocked.

Our learning rule allows the fly to learn, extinguish and re-learn multiple times which is impossible with a learning rule that allows only for depression of KC-MBON synapses to occur.

4.3 Methods

We use the two layer mode as in Chapter 3, with the same parameters to investigate the mechanism behind appetitive learning and extinction.

4.4 Results

In this section we discuss our results using our learning rules for relief learning (Rule 3 and aversive re-learning (Rule 4). We used the same network as in Chapter 3 and modified the learning rule at KC-MVP2 synapses and KC-M6 synapses. Using

Rule 3 (dopa-pre-ltp) (Equation 4.1) we simulated shock learning with different inter-stimulus intervals (ISIs) ranging from exposing shock 150 seconds before odour onset to a 60 seconds delay. When shock precedes odour onset (ISI = 130) (Fig 33 A) dopamine activity will create a trace of dopamine activity that will lead to LTP at the time of odour onset (Fig 33 c). When shock exposure is partially overlapping with odour exposure the outcome will be LTD as has been observed experimentally (Fig 33 B)

4.4.1 Appetitive learning and extinction

In this section we investigate appetitive learning and extinction in our model. After reward learning KC-M6 weights decreased (Fig 4.3) while extinction simulation led to the decrease KC-V2 weights (Fig 4.3). Reward learning led to the increase in PAM- γ 5 dopaminergic activity above baseline (Fig 4.3 d. green), while during extinction learning PPL1-2 dopaminergic fired above baseline (Fig 4.3 d. red).

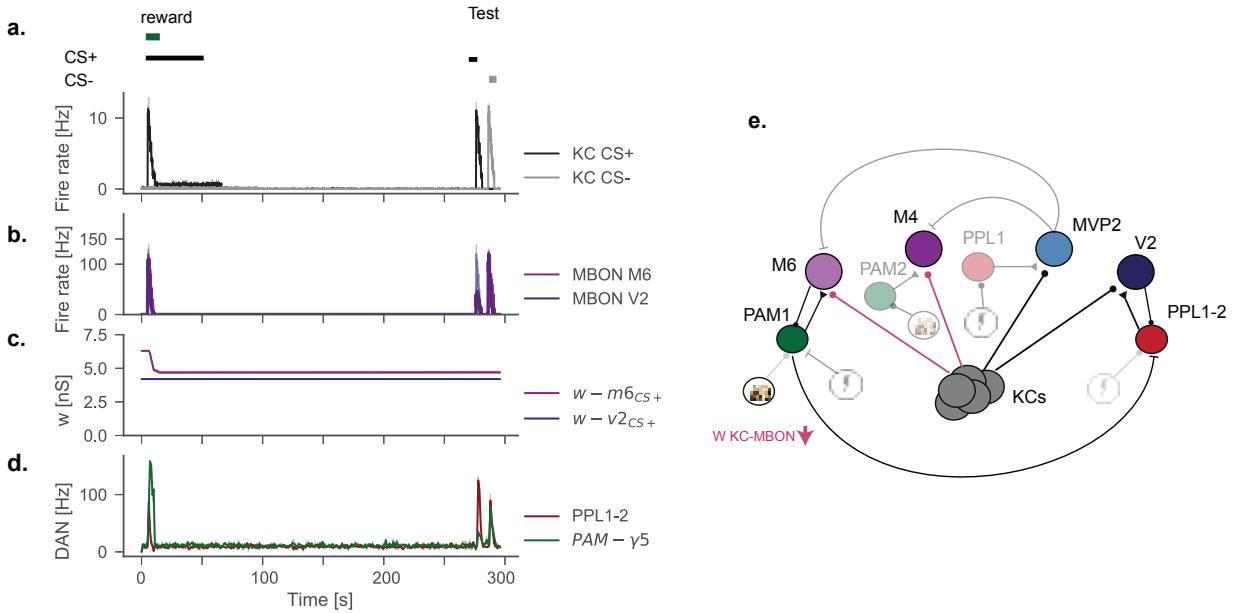


Figure 4.1: Network activity during appetitive learning and extinction. **a.** Population activity firing rate of KCs responding to CS+ and CS- stimuli during learning and testing. **b.** Firing rates of MBONs MVP2 and M6 during the simulation. **c.** We show average CS+ KC-V2 (blue) and KC-M6 (purple) during the simulations. **d.** Firing rates of PPL1-2 and PAM- γ 5 DANs during the simulation. **e.** We show the circuit representation and connections in response to CS+ after appetitive extinction. Black lines from KC to MBONs are unchanged connections while red lines are connections that underwent LTD.

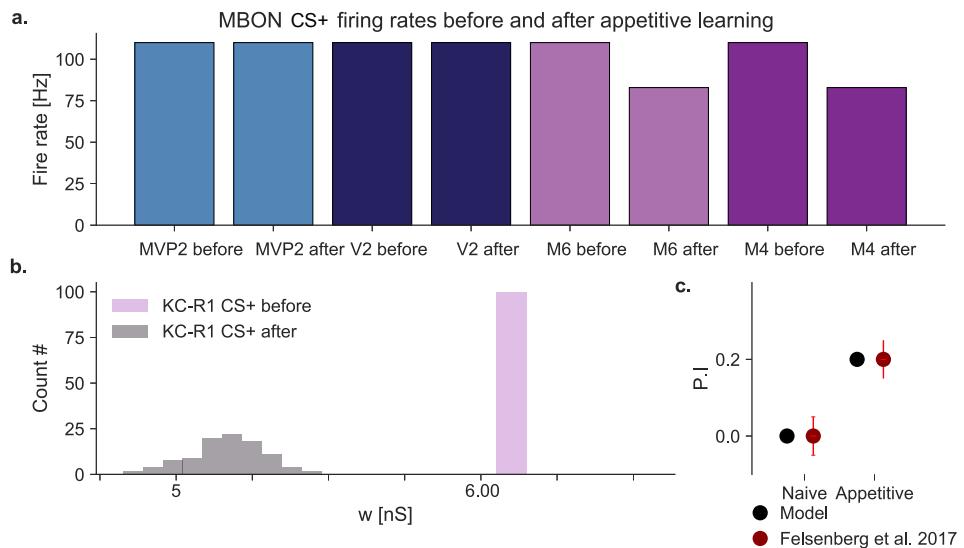


Figure 4.2: MBON firing rates, Perf. Index and weights after appetitive learning. **a.** MBON MVP2 (light blue), V2 (dark blue), M6 (light purple), M4 (dark purple) are shown before and after appetitive learning simulation. **b.** KC-M6 connection weights for KCs that respond to CS+ are shown before (light purple) and after (dark purple) appetitive learning. **c.** Model performance index and experimental performance index based on data from Felsenberg et al. 2018 (dark red) is shown after appetitive training.

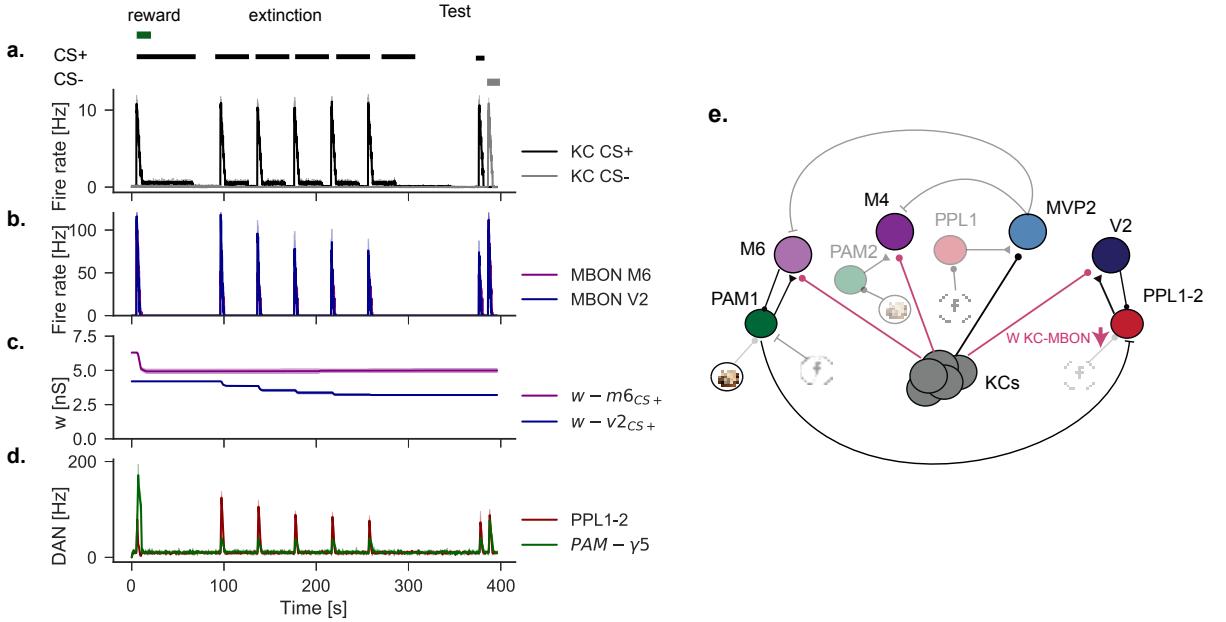


Figure 4.3: Network activity during appetitive learning and extinction. **a.** Population activity firing rate of KCs responding to CS+ and CS- stimuli during learning and testing. **b.** Firing rates of MBONs MVP2 and M6 during the simulation. **c.** We show average CS+ KC-V2 (blue) weights and KC-M6 (purple) during the simulations . **d.** Firing rates of PPL1-2 and PAM- γ 5 DANs during the simulation.**e.** We show the circuit representation and connections in response to CS+ after appetitive extinction. Black lines from KC to MBONs are unchanged connections while red lines are connections that underwent LTD.

Appetitive learning decreased the firing rates of MBON M6 and M4 while MBON MVP2 and MBON V2 remained unchanged (Fig 4.2 a.). Our learning rule decreases KC-M6 and KC-M4 weights (Fig 4.2 b.) after appetitive learning. We tuned our learning rule to reproduce the experimental performance index after appetitive learning and our results show we obtained similar values (Fig 4.2 c. black and red).

After appetitive extinction the firing rate of MBON V2 decreases (Fig 4.2 a.) and both the weights of KC-V2 and KC-M6 and M4 are reduced (Fig 4.4 b.). Here we tuned our learning rule to match the experimental performance index and we obtained similar results (Fig 4.4 c.).

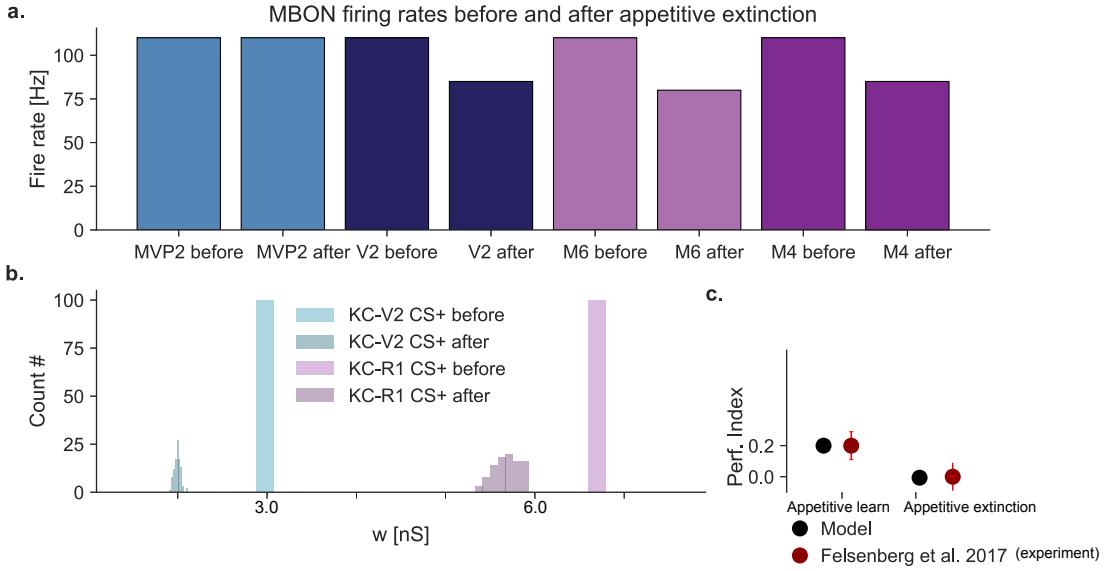


Figure 4.4: MBON firing rates, Perf. Index and weights after appetitive extinction. **a.** MBON MVP2i (light blue), V2 (dark blue), M6 (light purple), M4 (dark purple) are shown before and after aversive extinction simulation. **b.** KC-M6 connection weights for KCs that respond to CS+ are shown before (light purple) and after (dark purple) aversive extinction. **c.** Model performance index and experimental data from Felsenberg et al. 2018 (dark red) is shown after aversive training and after aversive extinction. Model 1 (black) represents simulation where only the firing rate of M6 decreases after extinction simulation. Model 2 (blue) represents the case where both M6 and M4 MBON decrease in their firing rate by the same amount.

We now investigate the mechanism behind appetitive extinction in our model. Our results show that reward learning leads to the decrease in the firing rate of PAM- γ 5 in response to CS+ which inhibits PPL1-2 (Fig 4.4 a green). PPL1-2 increased its activity in response to CS+ after reward learning (Fig 4.4 a red). After appetitive extinction PPL1-2 decreased its CS+ response while PAM- γ 5 remained unchanged (Fig 4.4 b.)

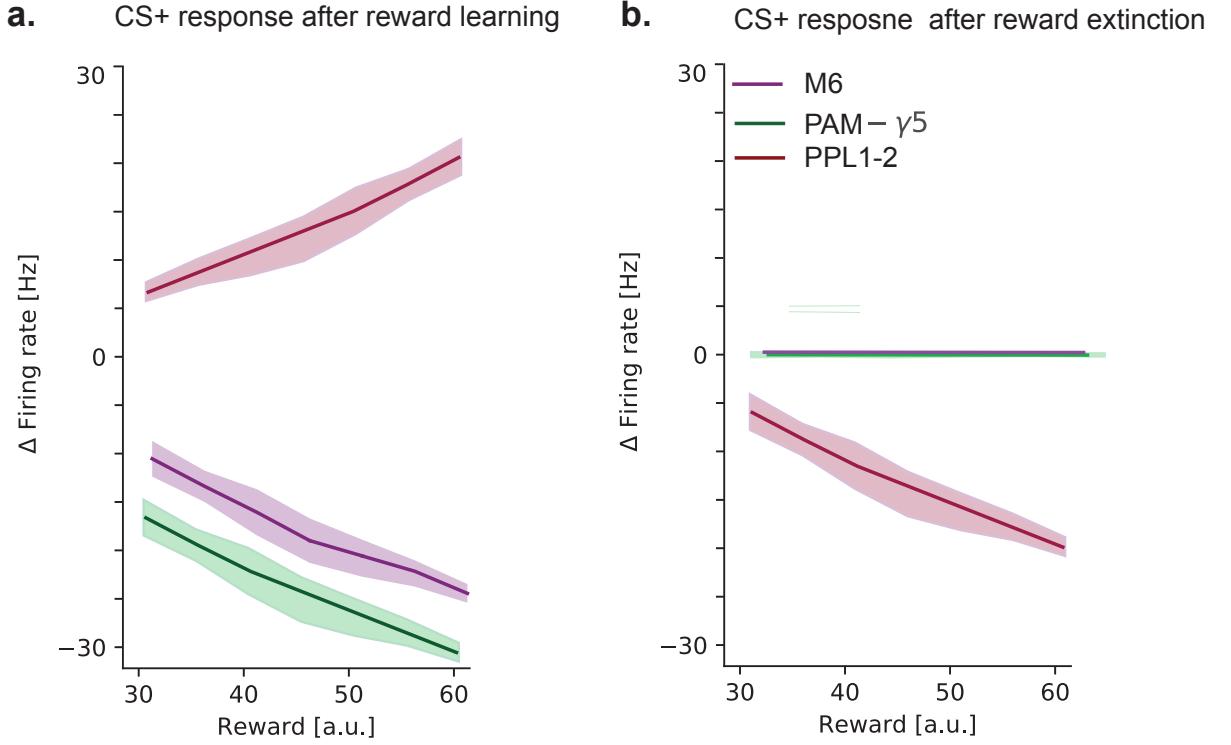


Figure 4.5: PAM- γ 5, PPL1 and M6 after reward learning and reward extinction protocol. **a.** We show the changes in the firing rate of PAM- γ 5., PPL1, and M6 as we increase the reward during appetitive learning. **b.** Similar to a. we show the result of appetitive extinction for appetitive learning with increasing reward.

We also investigated the effect of changing the learning rates for KC-V2 and KC-M6 compartments (Fig 4.6). Our results show that by tuning the learning rates we can extinguish an appetitive memory completely (Performance Index = 0, Fig 4.6 c).

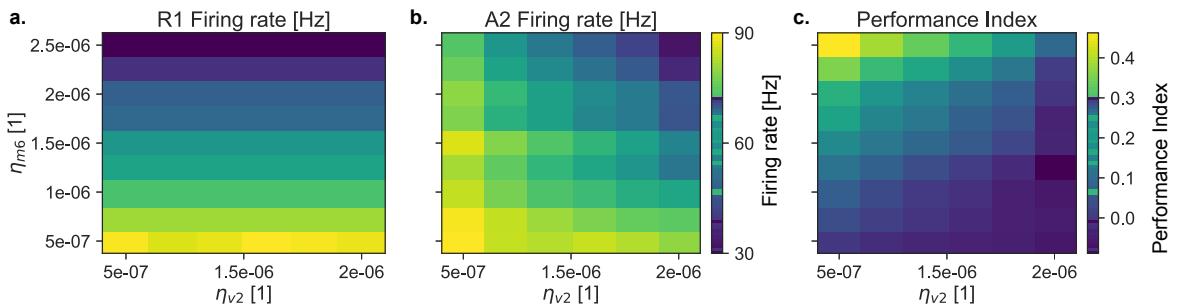


Figure 4.6: Appetitive extinction with different learning rates. **a** We show a colormap of MBON M6 firing rate after appetitive extinction as we change learning rates for KC-MVP2 connection on the x-axis and KC-M6 learning rate on the y-axis. **b.** We show a colormap with the same axes as in a. but we illustrate changes in V2 MBON firing rate. **c.** We show a colormap using the same configuration as in a and b. but we show the changes in the performance index after appetitive extinction.

4.4.2 Extinction is blocked if reinforcement is present during re-exposure

We also tested whether we can block extinction of a memory if the reinforcement that was used to train the CS+ is present during re-exposure trials. Our results show that during the modified aversive extinction protocol KC-M6 do not undergo any changes, while KC-MVP2 weights are modified (Fig 4.7 c left). When we simulated the appetitive extinction protocol, presence of reward did not modify KC-V2 weight, thus extinction was blocked in both appetitive and aversive extinction protocols if reinforcement is present. We calculated the model performance index and our results show that after re-exposing the odour with punishment we obtained a performance index higher than initial aversive learning (Fig ?? b). Similarly, re-exposing the odour after appetitive training with reward led to a performance index higher than initial appetitive learning (Fig 4.7 b).

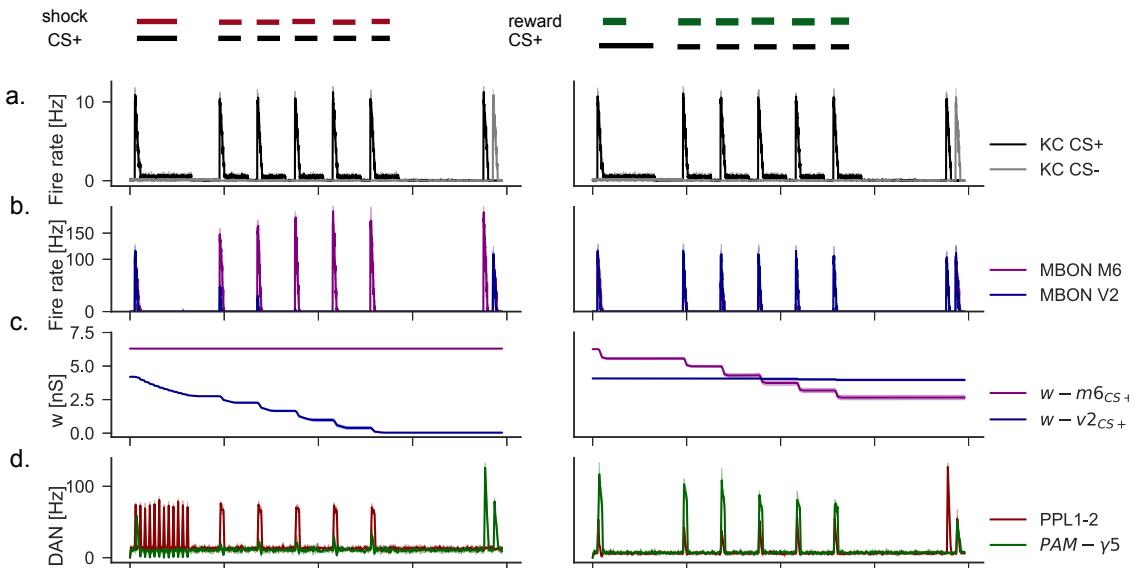


Figure 4.7: Network activity during extinction protocol in the presence of reinforcement. Left: CS+ odour (black) is exposed for 60s, followed by 45 pause and 5 re-exposures **with** shock. Right: CS+ odour (black) is exposed for 60, s followed by 44 pause and 2 re-exposures **with** reward. The dynamics of neurons and traces during the aversive extinction protocol are shown: KC activity for both CS+ (black) and CS-(yellow) odours; MVP2 (blue) and M6 (purple) activity; PAM- γ 5 (green) and PPL1 (red) as well as KC-MVP2 and KC-M6 (purple) weights and PAM- γ 5 PPl1, and KC traces are shown for completeness.

4.4.3 Prediction: Blocking PAM- γ 5 dopaminergic neurons leads to an aversive memory

To reproduce appetitive extinction we added an inhibitory connection between the PAM- γ 5 DAN and the PPL1 DAN. We simulated a scenario that can be tested in experiments and we blocked the activity of PAM- γ 5 during odour exposure. Blocking PAM- γ 5 led to the decrease in the firing rate of MBON V2 after odour exposure which in an experimental scenario would increase the aversion towards the odour (Fig 39).

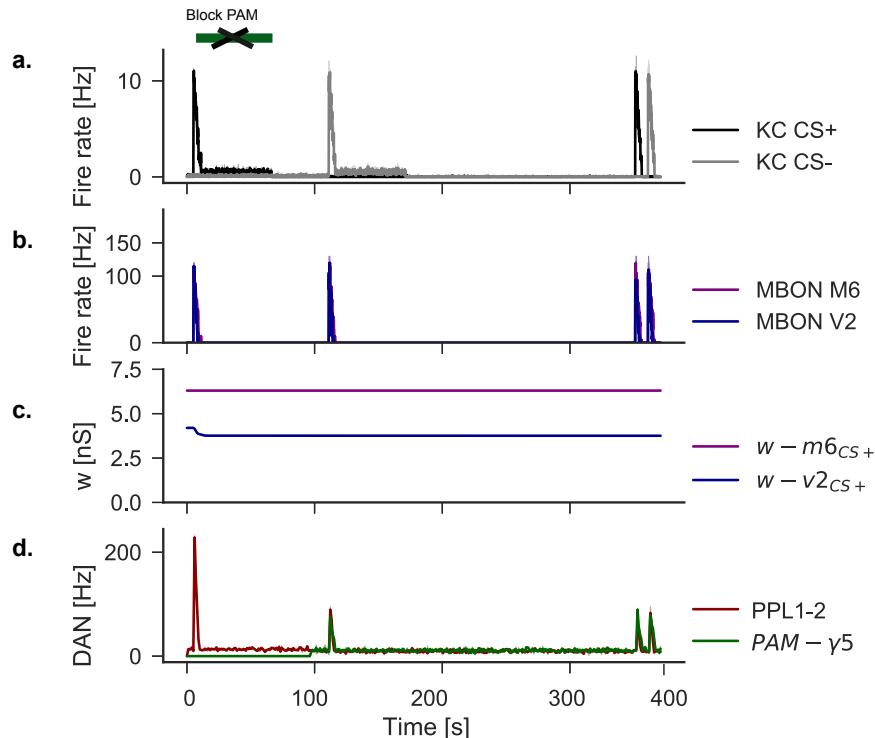


Figure 4.8: Suppressing PAM- γ 5 during odour exposure. **a.** Population activity firing rate of KCs responding to CS+ and CS- stimuli during learning and testing. **b.** Firing rates of MBONs M6 and V2 during the simulation. **c.** We show average CS+ KC-V2 (blue) weights and KC-M6 (purple) during the simulations. **d.** Firing rates of PPL1-2 and PAM- γ 5 DANs during the simulation. **e.** We show PPL1-2 DAN and PAM- γ 5 DAN eligibility traces during the simulation. **f.** We show one KC CS+ eligibility trace with $\tau_j = 10s$. *Performance index score in this model was -0.1 for the odour exposed while blocking PAM- γ 5 DAN.*

4.5 Discussion

We have reproduced appetitive and extinction learning in the mushroom body using the same model we introduced in Chapter 3. We propose a mechanism that is different from aversive extinction. In the case of aversive learning we found M6 MBON increases its CS+ response which also enhances PAM- γ 5 DAN CS+ response due to excitatory feedback. This can explain aversive extinction but is not a suitable mechanism to explain appetitive extinction. Here, we found that by adding an inhibitory connection from *PAM* - γ 5 to a PPL1 DAN, after appetitive learning the activities of the two DANs will change in opposite direction. *PAM* - γ 5 will decrease its CS+ response while *PPL1* - 2 increases. Using our high-dopa-LTD learning rule we found that we could trigger plasticity in the V2 compartment during odour re-exposure which facilitates extinction of the appetitive memory. To test our model we proposed a simple experiment that involves blocking appetitive DANs. Exposure to an odour while DANs are blocked should create an aversive memory.

Chapter 5

Investigating the role of potentiation in the mushroom body

5.1 Ch 5: Abstract

In this chapter we investigate the role of potentiation in the mushroom body between KC-MBON synapses.

In chapter 3 and 4 we introduced two learning rules that can only weaken KC-MBOn synapses and showed that they are sufficient to model learning and extinction experiments in the fruit fly. We also showed that while we can perform aversive learning, extinction and re-learning, re-extinction is not possible in a model that can only decrease the strength of KC-MBON weights. In our previous learning rules LTD was triggered when dopamine activity levels were above baseline levels of activity during odour presentation. Here, we introduce a new learning rule, low-dopa-LTP that triggers potentiation when an odour is present and dopamine firing rate is below baseline activity. An experiment suggesting potentiation at KC-MBON synapses has shown that shock encoding PPL1 activation before odour activation leads to the formation of a positive memory. Here, we modify the PPL1 DAN we modelled in Chapter 3 by adding spike frequency adaptation that lasts 20 seconds which reduces the activity of PPL1 below baseline level after shock presentation which leads to potentiation of KC-MVP2 synapses using learning low-dopa-LTP and a positive performance index. While we have shown that depression is sufficient to reproduce learning, extinction and re-learning experiments we show that by using our learning rule we can enable re-extinction as well. We propose that during both aversive and appetitive re-learning DANs in compartments that are required for extinction are inhibited by the reinforce-

ment which using our learning rule triggers potentiation of KC-MBON synapses and enables re-extinction. Finally, we show that our new learning rule can be tested in the full model of the olfactory circuit we presented in Chapter 2, and a single odour can learn, extinct and re-learning an aversive association without affecting other odours.

5.2 Introduction

Experimental studies have shown that synapses can increase and decrease their strength depending on the activity of presynaptic and postsynaptic neurons, as well as neuromodulators (cite Bliss 1973, Bi and Poo 1998, Cassenaer). In the fruit fly high dopamine activity pair with an odour has been shown to trigger a decrease in the firing rate of MBONs (cite Hige).

However, there has been experimental evidence that potentiation occurs in the mushroom body. A recent experiment has shown that when DAn activation in the $\gamma 4$ compartment was activated more than 45 seconds apart from KC activation the $\gamma 4$ MBON activated was enhanced. The effect was fully reversible once KC and DAn activation were paired together. This experiment suggests that unpaired DAn activation potentiates all KC-MBON synapses (lack of odour specificity) in a single compartment or that, DAn activation that precedes KC activation leaves a trace for tens of seconds, which leads to the potentiation once KC are activated, which would be odour specific potentiation.

Evidence for odour specific potentiation has been observed in the relief learning paradigm. Relief memories refers to the ability of animals to learn the cues that follow a traumatic experience. A behavioural experiment has shown that when a fly is exposed to an odour after getting shocked, it will tend to approach that odour (cite Yarali 2008)

In a more recent experiment, specific DANs were activated and it was found that photostimulation of the PPl-1 DAN was sufficient to create a relief memory (cite Yarali 2018).

A previous model by Abbot and Drew was proposed to explain the relief learning experiment. The authors showed that by extending the time window of STDP to tens of seconds they could explain the relief learning experiment.

While the model fails to explain why postsynaptic spikes appear to be dispensable in the fruit fly, the model does highlight the need to have a mechanism for both depression and potentiation to reproduce relief learning.

The fruit fly navigates an environment that is dynamic and odours that predict rewards and punishment can change their predictions multiple times. In Chapters 3 and 4 we showed that initial learning in the fruit fly can be explained by decreasing the activity of the MBONs which bias for the undesired behaviour (i.e during appetitive learning the activity of retreat MBONs decreases). However, our model can not perform re-learning and re-extinction multiple times without potentiation. Here we show that the same learning rule we use to reproduce relief learning can also be used to enable re-learning and re-extinction multiple times.

5.3 Methods

In this chapter we continue to use the model we developed in Chapter 3 to investigate the role of potentiation in re-learning and re-extinction. We introduce a new learning rule (low-dopa-pre-LTP) that potentiates KC-MBON weights when the activity of DANs is below baseline activity.

We also revisit the model we developed in Chapter 2 to create a unified 4 layer model that can create a sparse representation of an odour and learn to associate an odour with an aversive stimulus, extinct and re-learn without changing the valence of other odours.

5.3.1 Rule 4:low-dopa-pre-LTP

Here we introduce Rule low-dopa-LTP that combines our rule high-dopa-LTD with a second component that allows KC-MBON synapses to potentiate when dopamine activity is below a defined threshold.

$$\frac{d}{dt}w_{ij}(t) = \underbrace{A_+e_j(t)S_i\mathcal{H}(\theta_{dopa}^+ - z_i(t))}_{LTP} - \underbrace{A_-e_j(t)S_i\mathcal{H}(z_i(t) - \theta_{dopa})}_{LTD} \quad (5.1)$$

where A_- and A_+ are positive parameter that determines the maximum depression and potentiation a connection can undergo respectively. Synaptic weight change w_{ij} requires the synaptic tag e_j to be above zero and dopamine tag z_i to be above θ_{dopa} which is a threshold set to be equal to background odour activity. The eligibility trace e_j is updated each time the presynaptic neuron j fires as spike as follows:

$$\tau_e \frac{d}{dt}e_j(t) = -e_j(t) + S_j(t), \quad (5.2)$$

where τ_e is the characteristic time of e_j and $S_j(t)$ is the spike train of the j-th presynaptic neuron, described by the following equation:

$$S_j(t) = \sum_k \delta(t - t_{j,k}^*), \quad (5.3)$$

where $t_{j,k}^*$ is the timing of the spike, where j is the -th neuron and k is the index of the spike.

Similarly $S_i(t)$ is the spike train of the i-th dopaminergic neuron:

$$S_i(t) = \sum_k \delta(t - t_{i,k}^*), \quad (5.4)$$

where $t_{i,k}^*$ is the timing of the spike, where i is the i-th dopaminergic neuron and k is the index of the spike.

The dopamine concentration, $z(t)$ which gates synaptic change represents a global eligibility trace that increases each time the dopamine neuron fires a spike,

$$\frac{dz(t)}{dt} = -\frac{z(t)}{\tau_d} + S_i(t), \quad (5.5)$$

where τ_e is the characteristic time of τ_d and $S_i(t)$ is the spike train of the j-th presynaptic neuron,

Parameter name	Symbol	Value
Pre eligibility time constant	τ_e	10 s
Dopamine (z) time constant	τ_z	1 s
Learning rate	η	0.001
MBON MVP2, M4 threshold	θ_{dopa}	10
MBON M6, V2 threshold	θ_{dopa}	20

Table 5.1: Rule 1: high-dopa-LTD parameters

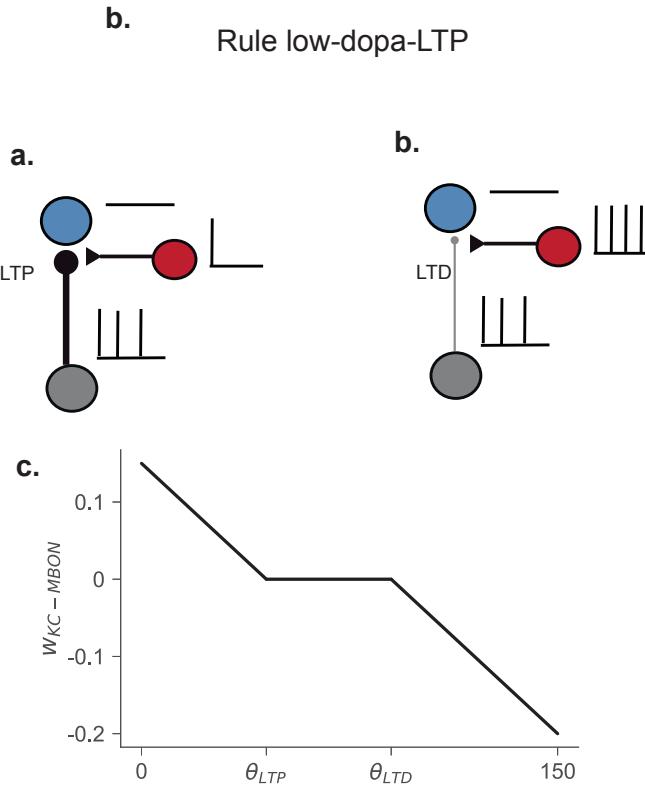


Figure 5.1: Learning rule low-dopa LTP **a.** Learning rule 1 (dopa-pre-ltd). Rule high-dopa-LTD requires only presynaptic activity and dopamine activity above a threshold. Coincident presynaptic and dopaminergic activity trigger LTD of the the synapse. **b.** Shown is change in KC-MBON weight ($w_{KC-MBON}$) when there is either no presynaptic activity (orange) or high presynaptic activity (10 Hz) and variable dopamine activity (weight change only occurs when there is both presynaptic activity and dopaminergic activity above a set threshold (θ_{dopa})).

5.3.2 Relief learning can be explained by potentiation of KC-MVP2 synapses

Here we show that our learning rule pre-dopa-LTP can reproduce relief learning experiments. We studied relief learning by reversing the order between the shock and odour exposures. The relative timing difference between shock and odour onset is called the ISI (inter stimulus interval). Positive ISIs reflect shock precedes odour exposure. ISIs higher than 60 seconds correspond with the situation of no overlap between shock and odour exposure protocol. We simulated ISI intervals between 140 seconds and -10 seconds using our learning rule pre-dopa-LTP. For ISIs higher than 60 seconds KC-MVP2 weights potentiation (Fig 24 c) which led to the increase of the firing rate

of the MVP2 MBON. Fo ISIs lower than 40 seconds KC-MVP2 synapses depressed. This is in line with experimental evidence (Yarali et al 2018).

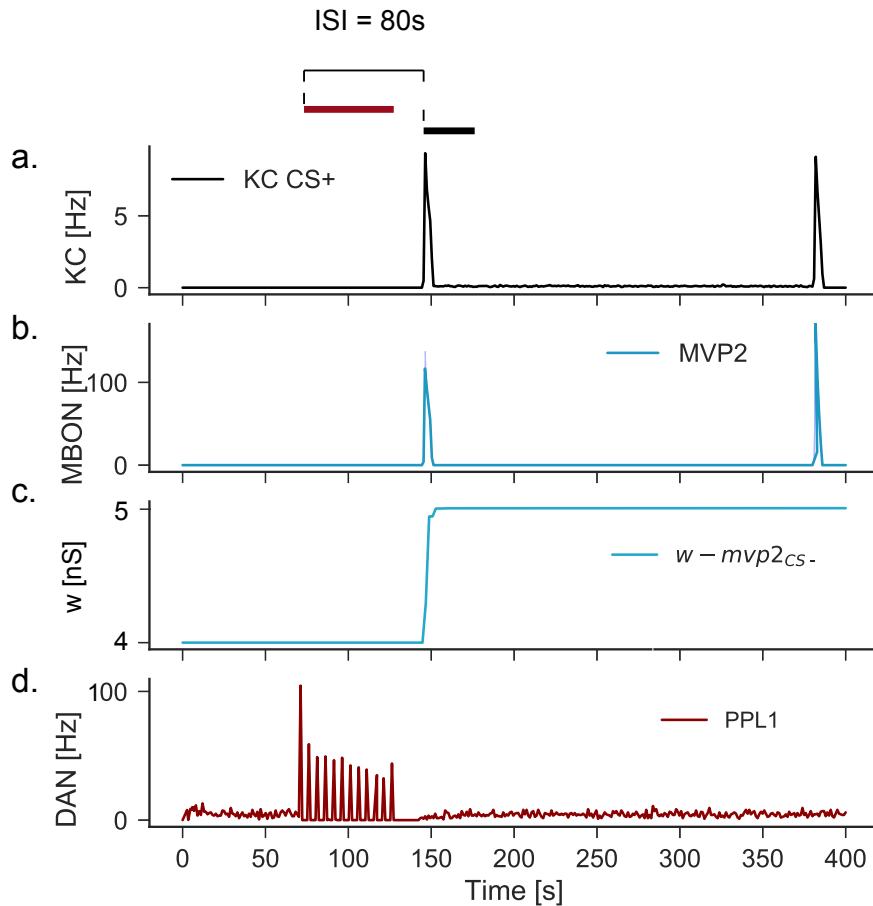


Figure 5.2: Shock learning depends on the relative timing of shock and odour onset. Here we show simulation result using two different ISI ($t_{odor}^{start} - t_{shoc}^{start}$). In a. shock onset precedes odour onset by 150 s (ISI = 150) which results in the potentiation of KC-MVP2 weights (middle left) while in b. shock onset is 20 seconds before odour presentation which results in the depression of KC-MVP2 weights (b iii).

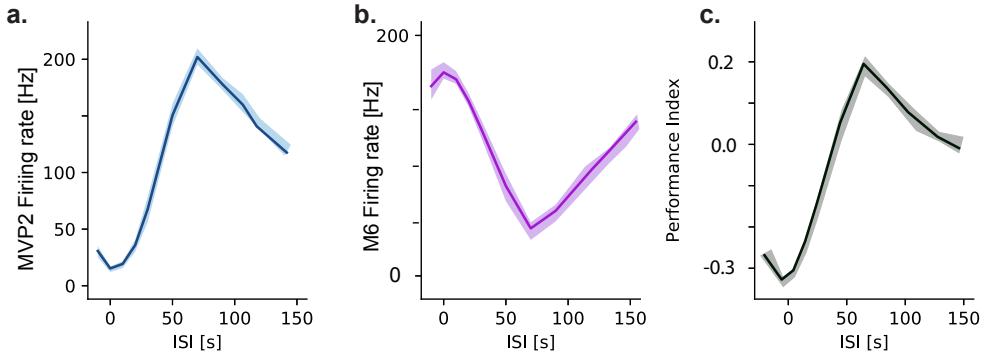


Figure 5.3: Rule low-dopa-LTP reproduces relief learning experiment. Shown here are changes in the MBON firing rates and performance index for ISI ranging from -10 to 150 seconds. ISI is the difference between the timing of the first shock and odour exposure. **a.** MVP2 firing rate for simulations of shock-odour pairs with ISI ranging from -10 to 150 seconds. **b.** Same as in **a.** We show the M6 MBON firing rate. **c.** Shown here is the model performance index for different ISI.

5.3.3 Re-learning with LTP enables re-extinction

In Chapter 3 we modelled re-learning using a learning rule that allows only for synaptic depression. Here we test a learning rule that potentiates KC-M6 synapses when PAM1 dopaminergic neurons fire below baseline levels. Experimental results show that the performance index score after aversive re-learning is higher than after initial learning (-0.6 vs -0.4). When we simulated aversive re-learning after aversive extinction using an LTD only learning rule with the parameters we used to reproduce the aversive extinction score we found that we could not get the same result. When we added potentiation to KC-M6 synapses we obtained a very similar score to that in experimental results. We added this rule for KC-MBON M6 connections to reproduce the re-learning experiment. We found that by using our new learning rule 4 we can re-learn at the same level as the initial aversive learning. After aversive learning presenting the odour with shock again depression KC-MBON MVP2 synapses. With our normal rule initial learning already reduces MBON MVP2 firing rate by 80%. Further learning would not be able to bring back the performance index to the initial level (refre-learning-fail).

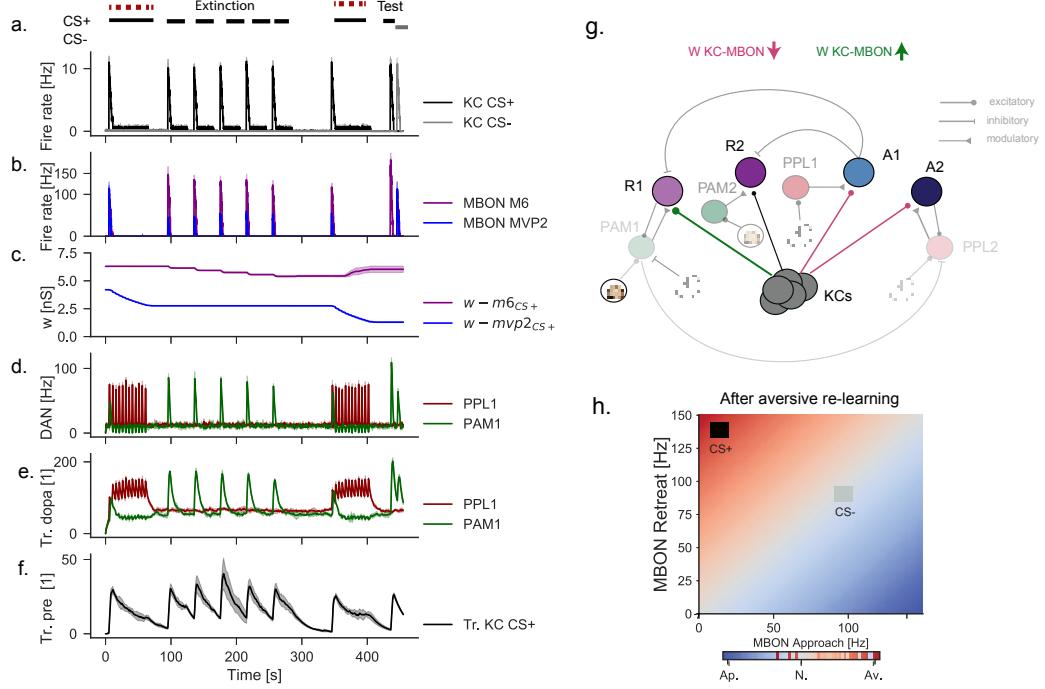


Figure 5.4: Network activity during aversive re-learning protocol using rule 4: low dopa LTP. **a.** KC population firing rates in response to CS+ and CS-. **b.** MBON MVP2 and M6 firing rates during the simulation. **c.** KC-MVP2 and KC-M6 CS+ weight change during simulation **d.** Panel shows the firing rates of PAM1 and PPL1 DANs during the simulation. **e.** Panel shows the eligibility traces of PAM1 and PPL1 DANs during the simulation. **f.** Panel shows the KC eligibility trace of one KC that responds to the CS+ with $\tau_j = 10s$.

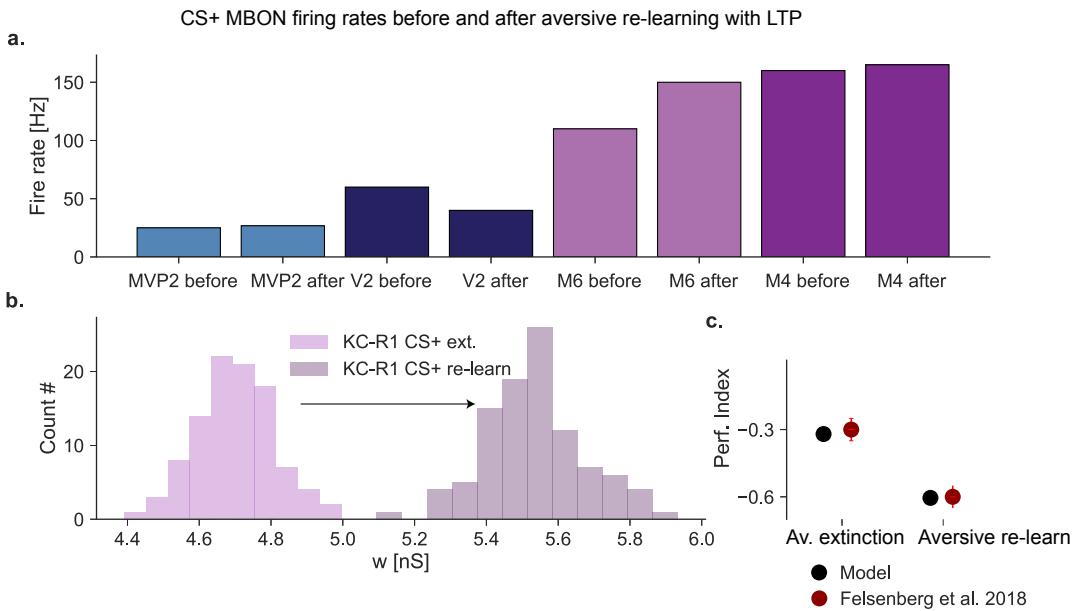


Figure 5.5: MBON firing rates, Perf. Index and weights after aversive re-learning. **a.** MBON MVP2j (light blue), V2 (dark blue), M6 (light purple), M4 (dark purple) are shown before (after aversive extinction) and after aversive re-learning. **b.** KC-M6 connection weights for KCs that respond to CS+ are shown before (light purple) and after (dark purple) aversive extinction. **c.** Model performance index and experimental data from Felsenberg et al. 2018 (dark red) is shown after aversive training and after aversive extinction. Model 1 (black) represents simulation where only the firing rate of M6 decreases after extinction simulation. Model 2 (green) represents the case where we use a learning rule that can not re-learn.

Our learning rule triggers LTP when an odour is present and dopamine is below a certain level. When dopamine is blocked, KC-M6 synapses potentiate. This would in turn create an aversive phenotype.

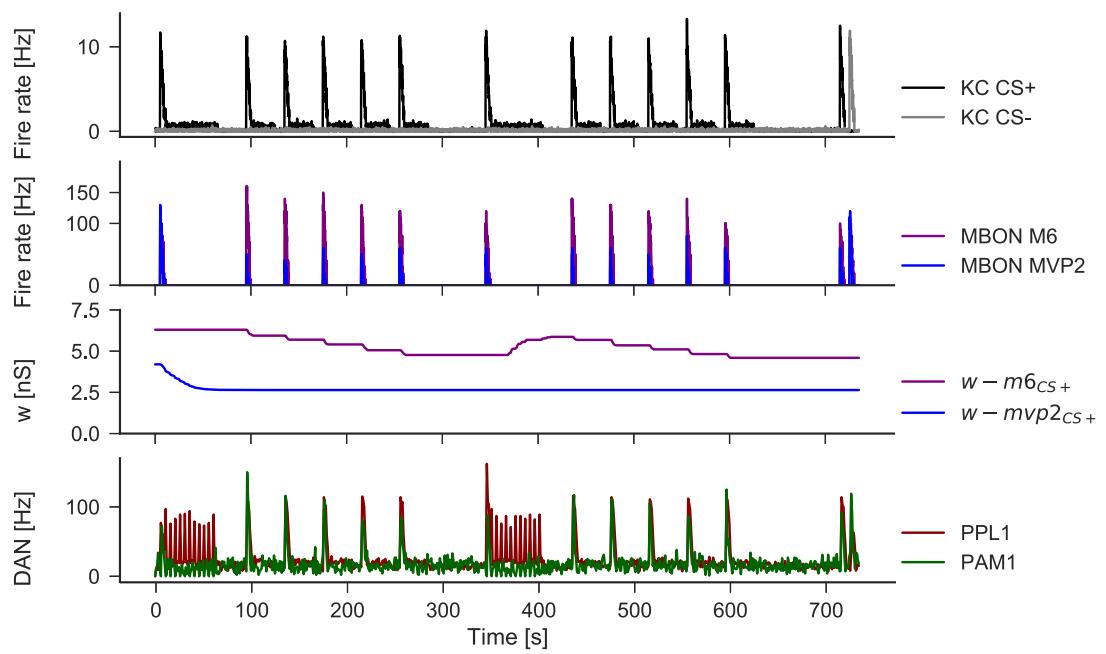


Figure 5.6: Rule 4 potentiates MBON-MVP2 synapses when dopamine precedes odour

a. Changes in the firing rate of the approach firing rate depending on the timing difference between odour and reinforcement (ISI). b. c.

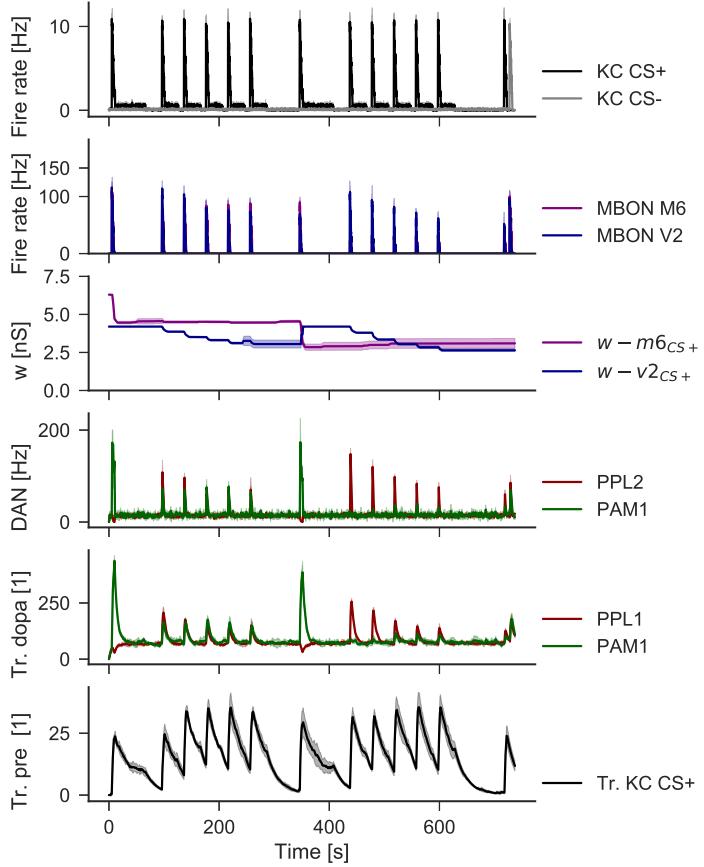


Figure 5.7: Network activity during aversive re-learning protocol using rule 4: low dopa LTP. **a.** KC population firing rates in response to CS+ and CS-. **b.** MBON MVP2 and M6 firing rates during the simulation. **c.** KC-MVP2 and KC-M6 CS+ weight change during simulation **d.** Panel shows the firing rates of PAM1 and PPL1 DANs during the simulation. **e.** Panel shows the eligibility traces of PAM1 and PPL1 DANs during the simulation. **f.** Panel shows the KC eligibility trace of one KC that responds to the CS+ with $\tau_j = 10s$.

5.3.4 Re-learning in the four layer olfactory circuit

We now show that we can simulate aversive learning, extinction, re-learning and re-extinction in the full circuit. Using the model we developed in Chapter 2, we selected four odours (CS1, CS2, CS3, CS4) and paired CS2 with shock. Using Rule 4 (LTP) we trained the full circuit to create an aversive memory for odour 2, then extinguish it, re-learn in and re-extinguish it (Fig 43).

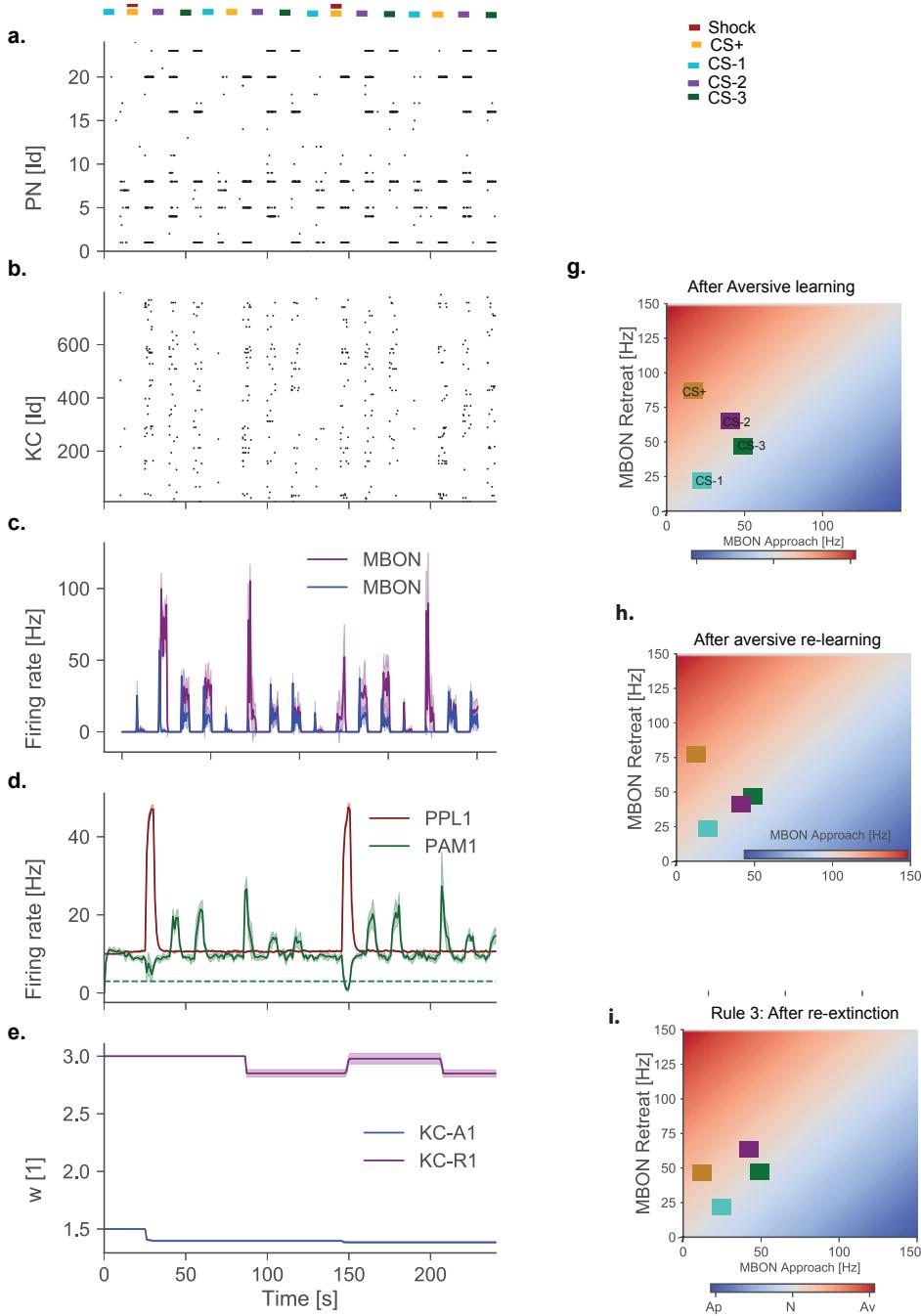


Figure 5.8: Learning, extinction, re-learning and re-extinction in the full network. We plot the responses of the neurons in our network to 4 stimuli: CS-1 (light blue), CS+ (yellow), CS-2 (purple), CS-3 (dark green). **a.** Spike raster of PN activity during the simulation. **b.** Spike raster of KC activity during simulation. **c.** Firing rates of MBON-MVP2 and MBON-M6 during the simulation. **d.** Firing rates of PAM1 and PPL1 during the simulation. **e.** Average weight of KC-MVP2 (blue) and KC-M6 (purple) during the simulation. **f.** Circuit illustrating 4-layer model used in simulation: ORNs, PNs, KCs and MBONs. This is the same model we used in chapter 2. **g.** Firing rate elicited by MBON MVP2 (Approach x-axis) and MBON-M6 (Retreat y-axis) for the four stimuli (CS+, CS-1, CS-2, CS-3, CS-4) after aversive training. **h.** and **i.** are the same as **g** but they show firing rates after aversive extinction and re-learning respectively.

Using Rule2 in the same scenario fails to re-learn (Fig 44).

5.4 Discussion

Experiments in the fruit fly have found mostly evidence for dopamine modulated depression of KC-MBON synapses after learning. This can be explained by evidence that appetitive DANs target MBONs which promote retreat behaviour, while aversive DANs target MBON which code for approach (cite Aso). Optogenetic activation of DANs paired with an odour has been shown to lead to the decrease in the firing rates of MBONs (cite Hige).

This suggests that high concentration of dopamine during odour presentation leads to depression of KC-MBON synapses, which is sensible since aversive learning will reduce the activity of approach MBONs, while appetitive learning will decrease the activity of retreat MBONs in response to the trained odour.

In this chapter we investigated the functional consequences of potentiation in the mushroom body. Under what circumstances is potentiation necessary?

We proposed a single learning rule (Rule low-dopa-LTP) to explain experimental observations. Specifically, in relief learning *PPL1* – $\gamma 1ped$ activation before odour exposure creates an appetitive memory. We added adaptation that decreases the firing rate of *PPL1* DAN for up to 20 seconds after shock presentation, which led to potentiation of KC-MVP2 synapses when we simulated the relief learning experiment.

We also showed that potentiation of KC-MBON synapses enables e-extinction. During appetitive re-learning PPL1-2 DAN is inhibited by *PAM* – $\gamma 5$ which responds to reward. After appetitive re-learning the firing rate of MBON-V2 is enhanced which is necessary for re-extinction if the trained odour is exposed again without the expected reward. Conversely, during aversive re-learning *PAM* – $\gamma 5$ is inhibited by shocks which leads to the potentiation of KC-M6 synapses. After re-learning the CS+ firing rates of M6 MBON and *PAM* – $\gamma 5$ were enhanced which allowed to trigger re-extinction when we presented the odour again.

Finally, we tested whether our learning rule could work in a multi-layer model of the olfactory circuit. We tested the model we developed in Chapter 4 by adding plasticity at KC-MBON synapses using our low-dopa-LTP learning rule. In Chapter 2 we have shown that lateral inhibition in the antennal lobe and feedback inhibition in the mushroom body increases decorrelation between odours. We presented four odours from the Carlson et al. 2009 data-set and trained one odour with an aversive stimulus. Our results show we can perform aversive learning, extinction,

re-learning and re-extinction on one odour without changing the representation of the other three odours as long as there is no significant overlap in KC representation. When we selected odours that had a higher degree of overlap, our protocol failed as an untrained odour that was similar to the trained odour triggered changes in KC-M6 weight. Our results show that we can train a full model of the olfactory circuit to create a sparse representation of the odour in KC populations and to reproduce learning experiments in the fruit fly. We have proposed a new learning rule that can explain both how depression and potentiation is orchestrated by the activity of DANs. We propose an experiment that images the activity of M6 MBON in response to CS+ after re-learning and we predict that the activity will be potentiated compared to its activity after extinction.

Chapter 6

Conclusions and future work

In this thesis we have shown that learning rules that depend on the activity of presynaptic and dopaminergic neurons are sufficient to explain learning experiments in the fruit fly. We used spiking neural networks that are tuned to exhibit biologically plausible firing rates. We focused our efforts on faithfully reproducing experiments in the fruit fly. In Chapter 2 we tuned a four layer model to study how odour information is transformed from the antennae to the mushroom body. We found that lateral inhibition in the antennal lobe contributes to decorrelation of odour representation. We also found that APL mediated feedback inhibition also contributes to decorrelation and increases population sparseness in response to an odour. We finished by showing that our model's MBONs can respond to all odours in a data-set.

In Chapter 3 and 4 we tested learning rules that trigger weakening of KC-MBON synapses. We showed that a learning rule that is independent of postsynaptic spikes can reproduce learning and extinction experiments. To reproduce extinction experiment we discovered that MBON-DA_n excitatory feedback is required. Thus, we conclude that while our learning rule does not require postsynaptic spikes, to perform both aversive and appetitive extinction MBON activity is required.

We also proposed a simple linear transformation from MBON firing rates to performance index. In our model the valence of the CS- remains unchanged before and after odour exposure, so the valence of the CS+ odour determines the model performance index we calculated. Based on our target performance index we tuned the learning rates in our model and we obtained very similar results to values measured in experiments.

To reproduce appetitive extinction we proposed an inhibitory connection from PAM- γ 5 DAn to PPL1-2 DAn. Such a connection predicts that suppressing the activity of PAM- γ 5 during odour exposure will create an aversive memory. Intriguingly,

we also discovered a possible mechanism for second order conditioning. In our model, MVP2 MBON inhibits PPL1 DAn. After aversiv learning MVP2 decreases its firing rate while PPL1 increases. When we simulated second order conditioning if an order is exposed immedidately before an aversive odour then PPL1 DAn can form an aversive memory in the MVP2 compartment of the previously untrained odour.

In Chapter 5 we proposed a solution to reproduce aversive re-extinction. While high levels of dopamine paired with an odour triggers LTD, we added a second component to our learning rule that depresses KC-MBON synapses when DANs fire below baseline level. This mechanism can both explain relief learning and allow our model to perform re-extinction. Finally, we showed that we can use our learning rule to train odours from the Hallem Carlson data-set to train one odour without affecting the representation of the other odours.

We have proposed a series of experiments that can validate our learning rule and model. In Chapter 2 some odour from the data-set elicited very few ORN spikes, which led to very weak PN and KC responses. Luo et al have proposed that PNs normalize odour responses through a non-linear transformation. It would be very interesting to investigate whether we could reproduce teh transformation proposed by Luo et al. using short term plasticity between ORN and PNs.

If we can normalize odour responses in PNs we might also be able to obtain very similar MBON responses from all of the odours. If all the odours activate the same number of KCs we could test whether we could train all the 110 odours in the data-set to be either aversive or appetitive.

It would also be interesting to investigate how odour mixtures activated ORNs and what would happen to the valence of an odour mixture if only one odour is trained.

In summary, my thesis extends the understanding of the different stages of formation of memory in the fruit fly and show that both depression and potentiation are required in the mushroom body. My results suggest that DANs in the fruit fly give the fly the ability to form associatiations with rewards and punishment and update those memories whenever the value of the stimulus changes.

Bibliography

L F Abbott, Sen Song, and Kenneth D Miller. Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. - PubMed - NCBI. *Nature Neuroscience*, 3(9):919–926, September 2000. pages 56

Yoshinori Aso and Gerald M Rubin. Dopaminergic neurons write and update memories with cell-type-specific rules. *eLife*, 5:e16135, 2016. pages 46, 47, 57

Yoshinori Aso, Divya Sitaraman, Toshiharu Ichinose, Karla R Kaun, Katrin Vogt, Ghislain Belliart-Guérin, Pierre-Yves Plaçais, Alice A Robie, Nobuhiro Yamagata, Christopher Schnaitmann, William J Rowell, Rebecca M Johnston, Teri-T B Ngo, Nan Chen, Wyatt Korff, Michael N Nitabach, Ulrike Heberlein, Thomas Preat, Kristin M Branson, Hiromu Tanimoto, Gerald M Rubin, and Liqun Luo. Mushroom body output neurons encode valence and guide memory-based action selection in *Drosophila*. *eLife*, 3:e04580, December 2014. pages 7, 8, 11, 14, 15, 16, 20, 22, 27, 46, 47, 50, 57, 78

CDO Beck, Bradley Schroeder, and Ronald L Davis. Learning performance of normal and mutant *Drosophila* after repeated conditioning trials with discrete stimuli. *Journal of Neuroscience*, 20(8):2944–2953, 2000. pages 47, 62

Rinaldo L Betkiewicz, Martin Paul Nawrot, and Benjamin Lindner. Circuit and cellular mechanisms facilitate the transformation from dense to sparse coding in the insect olfactory system. *bioRxiv*, page 240671, 2018. pages 28

Vikas Bhandawat, Shawn R Olsen, Nathan W Gouwens, Michelle L Schlief, and Rachel I Wilson. Sensory processing in the *Drosophila* antennal lobe increases reliability and separability of ensemble odor representations. *Nature neuroscience*, 10(11):1474, 2007. pages 27

G Q Bi and M M Poo. Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. *The*

Journal of neuroscience : the official journal of the Society for Neuroscience, 18(24):10464–10472, December 1998. pages 3, 21

Stephanie Bissière, Yann Humeau, and Andreas Lüthi. Dopamine gates ltp induction in lateral amygdala by suppressing feedforward inhibition. *Nature neuroscience*, 6(6):587, 2003. pages 4

Tim VP Bliss, AR Gardner-Medwin, and T Lømo. Synaptic plasticity in the hippocampal formation. *Macromolecules and behaviour*, 193:203, 1973. pages 3

Christopher J Burke, Wolf Huetteroth, David Owald, Emmanuel Perisse, Michael J Krashes, Gaurav Das, Daryl Gohl, Marion Silies, Sarah Certel, and Scott Waddell. Layered reward signalling through octopamine and dopamine in Drosophila. *Nature*, 492(7429):433–437, December 2012. pages 18

Paolo Calabresi, Roberto Maj, Nicola B Mercuri, and Giorgio Bernardi. Coactivation of d1 and d2 dopamine receptors is required for long-term synaptic depression in the striatum. *Neuroscience letters*, 142(1):95–99, 1992. pages 3

Stijn Cassenaer and Gilles Laurent. Conditional modulation of spike-timing-dependent plasticity for olfactory learning. *Nature*, 482(7383):47, 2012. pages 3, 21

Raphael Cohn, Ianessa Morantte, and Vanessa Ruta. Coordinated and compartmentalized neuromodulation shapes sensory processing in drosophila. *Cell*, 163(7):1742–1755, 2015. pages 19, 51

Gaurav Das, Martín Klappenbach, Eleftheria Vrontou, Emmanuel Perisse, Christopher M Clark, Christopher J Burke, and Scott Waddell. Drosophila learn opposing components of a compound food stimulus. *Current biology*, 24(15):1723–1730, 2014. pages 20

Sandeep Robert Datta, Maria Luisa Vasconcelos, Vanessa Ruta, Sean Luo, Allan Wong, Ebru Demir, Jorge Flores, Karen Balonze, Barry J Dickson, and Richard Axel. The drosophila pheromone cva activates a sexually dimorphic neural circuit. *Nature*, 452(7186):473, 2008. pages 12

J S de Belle and M Heisenberg. Associative odor learning in drosophila abolished by chemical ablation of mushroom bodies. *Science (New York, N.Y.)*, 263(5147):692–5, Feb 1994. pages 8, 12, 14, 22

Kenji Doya. Metalearning and neuromodulation. *Neural Networks*, 15(4-6):495–506, 2002. pages 6

Kristina V Dylla, Georg Raiser, C Giovanni Galizia, and Paul Szyszka. Trace conditioning in drosophila induces associative plasticity in mushroom body kenyon cells and dopaminergic neurons. *Frontiers in neural circuits*, 11:42, 2017. pages 13, 47, 48, 75

Katharina Eichler, Feng Li, Ashok Litwin-Kumar, Youngser Park, Ingrid Andrade, Casey M Schneider-Mizell, Timo Saumweber, Annina Huser, Claire Eschbach, Bertram Gerber, et al. The complete connectome of a learning and memory centre in an insect brain. *Nature*, 548(7666):175, 2017a. pages 12

Katharina Eichler, Feng Li, Ashok Litwin-Kumar, Youngser Park, Ingrid Andrade, Casey M Schneider-Mizell, Timo Saumweber, Annina Huser, Claire Eschbach, Bertram Gerber, et al. The complete connectome of a learning and memory centre in an insect brain. *Nature*, 548(7666):175, 2017b. pages 48, 50

Johannes Felsenberg, Oliver Barnstedt, Paola Cognigni, Suewei Lin, and Scott Waddell. Re-evaluation of learned information in drosophila. *Nature*, 544(7649):240, 2017. pages 10, 11, 47, 48

Johannes Felsenberg, Pedro F Jacob, Thomas Walker, Oliver Barnstedt, Amelia J Edmondson-Stait, Markus W Pleijzier, Nils Otto, Philipp Schlegel, Nadiya Sharifi, Emmanuel Perisse, et al. Integration of parallel opposing memories underlies memory extinction. *Cell*, 175(3):709–722, 2018. pages 7, 10, 11, 17, 20, 47, 48, 51, 60, 62, 67, 75, 76, 78

Nicolas Frémaux, Henning Sprekeler, and Wulfram Gerstner. Functional requirements for reward-modulated spike-timing-dependent plasticity. *Journal of Neuroscience*, 30(40):13326–13337, 2010. pages 52

Dana Shani Galili, Kristina V Dylla, Alja Lüdke, Anja B Friedrich, Nobuhiro Yamagata, Jin Yan Hilary Wong, Chien Hsien Ho, Paul Szyszka, and Hiromu Tanimoto. Converging circuits mediate temperature and shock aversive olfactory conditioning in drosophila. *Current biology*, 24(15):1712–1722, 2014. pages 19

Wulfram Gerstner and Romain Brette. Adaptive exponential integrate-and-fire model. *Scholarpedia*, 4(6):8427, 2009. pages 22, 23

Mark S Goldman, Jorge Golowasch, Eve Marder, and LF Abbott. Global structure, robustness, and modulation of neuronal models. *Journal of Neuroscience*, 21(14):5229–5238, 2001. pages 3

Elissa A Hallem and John R Carlson. Coding of odors by a receptor repertoire. *Cell*, 125(1):143–60, Apr 2006. pages iv, 12, 23, 27, 28, 29, 30, 34, 37, 44

DO Hebb. The organization of behavior. 1949. *New York Wiely*, 2(7):8, 1949. pages 2

Toshihide Hige, Yoshinori Aso, Mehrab N Modi, Gerald M Rubin, and Glenn C Turner. Heterosynaptic plasticity underlies aversive olfactory learning in drosophila. *Neuron*, 88(5):985–998, 2015a. pages 4, 17, 50, 70, 75

Toshihide Hige, Yoshinori Aso, Gerald M Rubin, and Glenn C Turner. Plasticity-driven individualization of olfactory coding in mushroom body output neurons. *Nature*, September 2015b. pages 16, 22, 52

Kyle S Honegger, Robert A A Campbell, and Glenn C Turner. Cellular-resolution population imaging reveals robust sparse coding in the drosophila mushroom body. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 31(33):11772–85, Aug 2011. doi: 10.1523/JNEUROSCI.1099-11.2011. pages 14, 22, 44, 47

Eugene M Izhikevich. Solving the distal reward problem through linkage of stdp and dopamine signaling. *Cerebral cortex*, 17(10):2443–2452, 2007. pages 6, 55

Karla R Kaun, Reza Azanchi, Zaw Maung, Jay Hirsh, and Ulrike Heberlein. A drosophila model for alcohol reward. *Nature neuroscience*, 14(5):612, 2011. pages 20

Michael J Krashes, Alex C Keene, Benjamin Leung, J Douglas Armstrong, and Scott Waddell. Sequential Use of Mushroom Body Neuron Subsets during Drosophila Odor Memory Processing. *Neuron*, 53(1):103–115, January 2007. pages 14

Andrew C Lin, Alexei M Bygrave, Alix De Calignon, Tzumin Lee, and Gero Miesenböck. Sparse, decorrelated odor coding in the mushroom body enhances learned odor discrimination. *Nature neuroscience*, 17(4):559, 2014a. pages 15, 16, 20, 27, 41, 44

Suewei Lin, David Owald, Vikram Chandra, Clifford Talbot, Wolf Huetteroth, and Scott Waddell. Neural correlates of water reward in thirsty drosophila. *Nature neuroscience*, 17(11):1536, 2014b. pages 19

Sean X Luo, Richard Axel, and L F Abbott. Generating sparse and selective third-order responses in the olfactory system of the fly. *Proceedings of the National Academy of Sciences of the United States of America*, 107(23):10713–8, Jun 2010. doi: 10.1073/pnas.1005635107. pages 12, 23

HaDi MaBouDi, Hideaki Shimazaki, Martin Giurfa, and Lars Chittka. Olfactory learning without the mushroom bodies: Spiking neural network models of the honeybee lateral antennal lobe tract reveal its capacities in odour memory tasks of varied complexities. *PLoS computational biology*, 13(6):e1005551, 2017. pages 23

Henry Markram, Joachim Lübke, Michael Frotscher, and Bert Sakmann. Regulation of synaptic efficacy by coincidence of postsynaptic aps and epsps. *Science*, 275 (5297):213–215, 1997. pages 3

Mala Murthy, Ilia Fiete, and Gilles Laurent. Testing odor response stereotypy in the drosophila mushroom body. *Neuron*, 59(6):1009–1023, 2008. pages 27, 50, 75

James Olds and Peter Milner. Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *Journal of comparative and physiological psychology*, 47(6):419, 1954. pages 1

Bruno A Olshausen and David J Field. Sparse coding with an overcomplete basis set: A strategy employed by v1? *Vision research*, 37(23):3311–3325, 1997. pages 27, 44

David Owald, Johannes Felsenberg, Clifford B Talbot, Gaurav Das, Emmanuel Perisse, Wolf Huetteroth, and Scott Waddell. Activity of Defined Mushroom Body Output Neurons Underlies Learned Olfactory Behavior in Drosophila. *Neuron*, 86 (2):417–427, April 2015. pages 16, 17, 19, 20, 47, 50, 61, 75

Ivan Petrovich Pavlov. Conditional reflexes: an investigation of the physiological activity of the cerebral cortex. *Oxford Press*, 1927. pages 2, 8

Alice Pavlowsky, Johann Schor, Pierre-Yves Plaçais, and Thomas Preat. A gabaergic feedback shapes dopaminergic input on the drosophila mushroom body to promote appetitive long-term memory. *Current Biology*, 28(11):1783–1793, 2018. pages 51

Verena Pawlak, Jeffery R Wickens, Alfredo Kirkwood, and Jason ND Kerr. Timing is not everything: neuromodulation opens the stdp gate. *Frontiers in synaptic neuroscience*, 2:146, 2010. pages 3, 21

Emmanuel Perisse, Yan Yin, Andrew C Lin, Suewei Lin, Wolf Huetteroth, and Scott Waddell. Different Kenyon Cell Populations Drive Learned Approach and Avoidance in Drosophila. *Neuron*, 79(5):945–956, April 2013. pages 15

W G Quinn, W A Harris, and S Benzer. Conditioned behavior in drosophila melanogaster. *Proceedings of the National Academy of Sciences of the United States of America*, 71(3):708–12, Mar 1974. pages 8

John NJ Reynolds and Jeffery R Wickens. Dopamine-dependent plasticity of corticostriatal synapses. *Neural Networks*, 15(4-6):507–521, 2002. pages 4

Wolfram Schultz, Peter Dayan, and P Read Montague. A neural substrate of prediction and reward. *Science*, 275(5306):1593–1599, 1997. pages 2, 5

Julien Séjourné, Pierre-Yves Plaçais, Yoshinori Aso, Igor Siwanowicz, Séverine Tranoy, Vladimirov Thoma, Stevanus R Tedjakumala, Gerald M Rubin, Paul Tchénio, Kei Ito, et al. Mushroom body efferent neurons responsible for aversive olfactory memory retrieval in drosophila. *Nature neuroscience*, 14(7):903, 2011. pages 18

Indulekha P Sudhakaran, Eimear E Holohan, Sahar Osman, Veronica Rodrigues, K Vijayraghavan, and Mani Ramaswami. Plasticity of recurrent inhibition in the drosophila antennal lobe. *Journal of Neuroscience*, 32(21):7225–7231, 2012. pages 44

Richard S Sutton and Andrew G Barto. *Introduction to reinforcement learning*, volume 135. MIT press Cambridge, 1998. pages 5

Christopher J Tabone and J Steven de Belle. Second-order conditioning in drosophila. *Learning & Memory*, 18(4):250–253, 2011. pages 10, 46, 79

Glenn C Turner, Maxim Bazhenov, and Gilles Laurent. Olfactory Representations by Drosophila Mushroom Body Neurons. *Journal of Neurophysiology*, 99(2):734–746, February 2008. pages 14, 29, 50

Leslie B Vosshall, Allan M Wong, and Richard Axel. An olfactory sensory map in the fly brain. *Cell*, 102(2):147–159, 2000. pages 12

- Scott Waddell. Reinforcement signalling in Drosophila; dopamine does it all after all. *Current opinion in neurobiology*, 23(3):324–329, 2013. pages 7, 19, 21
- Jan Wessnitzer, Joanna M Young, J Douglas Armstrong, and Barbara Webb. A model of non-elemental olfactory learning in Drosophila. - PubMed - NCBI. *Journal of Computational Neuroscience*, 32(2):197–212, June 2011. pages 23, 24, 28
- Jason Wolfe, Arthur R Houweling, and Michael Brecht. Sparse and powerful cortical spikes. *Current opinion in neurobiology*, 20(3):306–312, 2010. pages 41
- Nobuhiro Yamagata, Toshiharu Ichinose, Yoshinori Aso, Pierre-Yves Plaçais, Anja B Friedrich, Richard J Sima, Thomas Preat, Gerald M Rubin, and Hiromu Tanimoto. Distinct dopamine neurons mediate reward signals for short- and long-term memories. *Proceedings of the National Academy of Sciences of the United States of America*, 112(2):578–583, January 2015. pages 19
- Ayse Yarali, Thomas Niewalda, Yi-chun Chen, Hiromu Tanimoto, Stefan Duerrnagel, and Bertram Gerber. Pain relief learning in fruit flies. *Animal Behaviour*, 76(4): 1173–1185, 2008. pages 10, 11, 78, 79
- Friedemann Zenke and Wulfram Gerstner. Limits to high-speed simulations of spiking neural networks using general-purpose computers. *Frontiers in neuroinformatics*, 8:76, 2014. pages 28
- Ji-Chuan Zhang, Pak-Ming Lau, and Guo-Qiang Bi. Gain in sensitivity and loss in temporal contrast of stdp by dopaminergic modulation at hippocampal synapses. *Proceedings of the National Academy of Sciences*, 106(31):13028–13033, 2009. pages 4