Olfactory learning through dopamine modulated plasticity in the fruit fly



Bayar Menzat

Green Templeton College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Hillary 2019

To my mother, Sevinci Menzat

Acknowledgements

I would like to thank the Doctoral Training Centre and its leader David Gavaghan for opening the door for me to study a DPhil at Oxford University and the Engineering and Physical Sciences Research council for funding my research. I also thank the wonderful people at Green Templeton College for providing their support. Over the course of my DPhil an extraordinary amount of people have contributed to my scientific development. First of all, I express my gratitude to my supervisor Tim Vogels. I've never before been challenged by such high standards until I worked with Tim. I also never believed I could reach those standards until I was supervised by him. Throughout the years, Tim has taught me how to conduct, communicate and tell a scientific story both visually and in written form, while also being a mentor that would encourage me to trust myself and think independently. I would also thank my co-supervisor Scott Waddell, who opened his lab to me and provided all of his lab's papers, preliminary data and the opportunity to test my model in experiments. I am indebted for being allowed to work on a theoretical Phd and at the same time enjoy being so close to the people conducting experiments. This Phd in its form would never have been possible without the help of Johannes Felsenberg and his collaborators. Three chapters in this PhD are focused on reproducing their work in a model. I've learned so much about being a passionate scientist from discussions with Johannes. He helped me understand why people dedicate their entire lives to studying the brain of small insects (it's because they are one of our best hope to understand the brain). I would also like to thank other great members of the CNCB who've helped me during my Phd: Diogo Pimentel, David Owald, Emmanuelle Perisse, Paola Cognigni, Pedro Jacob, and many others. A special thanks goes to Fiona Woods and Ruth Brain who do all the hard work behind the scenes to make this centre such a stellar pace.

I thank the Vogels lab with deep gratitude for making it a joy to come to work, especially in the final year. First of all I will thank William Podlaski, who was my first scientific collaborator and best friend. I would also like to thank the other students in Vogels lab who have been equally great friends: Yayoi Teramoto Kimura, Homero, David Jia, and Jake Stroud. Yayoi has been very supportive of my work both scientifically and as a friend who would lend an ear whenever I had a problem. A special thanks goes to the post-docs: Georgia Christodoulou who always encouraged me, Everton Agnes, Chaitanya Chintaluri Friedemann Zenke, Victor Hernandez-Urbina and Rui P. Costa. Friedemann gave me a supercar and taught me how to drive it: the Auryn simulator. Everton taught me how to develop my own learning rules, how to analyze them and honesty in science. He helped me climb the mountain. Chaitanya listened to every problem that I've encountered no matter how small or big, and always gave me invaluable insight in how to solve it. I would also like to thank my friends and family. Coming here to Oxford, I never knew I would meet such an amazing Romanian community. From the first moment they took me into their house they made me feel like home. Daniel Balanescu and Mariela Neagu have been like a second family for me. Mihai Vidrighin, Stefan Saftescu, Andre Renchin and Cristobal Garibay Peterson, who are now my brothers. Pumpernick, Mira Schneiders, Monica, Ioana and Roana who are my sisters. There are many others that I would like to thank but I hope that they know how much I care about them and will forgive me for missing them here. I would like to finish by thanking my family. My dad, Orhan, my sister Tijen, and my mother Sevinci, have supported me in ways that I can only justify by the fact that it's ok for family to offer unconditional love.

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 model explains recent relief learning and extinction experiments. In relief learning if a shock precedes an odour, the fly will learn to approach the relief signalling odour. Extinction is the disappearance of a previously learned behaviour when the behaviour is not reinforced. 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 four 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 olfactory aversive learning and extinction. 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 explore appetitive learning and extinction. I hypothesize that dopaminergic neurons can inhibit each other and that suppressing the activity of a select group of dopaminergic neurons during odour exposure forms an aversive memory. In chapter 5 I investigate the role of potentiation in two experiments: relief learning and aversive re-learning. The aversive re-learning experiment shows that a fly can learn to avoid an odour, reduce its avoidance and re-learn to avoid the odour in a flexible manner depending on the absence or presence of shock in consecutive trials. I show that with potentiation re-extinction is also possible in the fly. 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 Synaptic plasticity 3

1.1.2 Role of neuromodulators in synaptic plasticity 3

1.1.3 Reinforcement learning - the theoretical perspective 5

1.1.4 Solving the distal reward problem 7

1.1.5 Drosophila Melanogaster, an ideal candidate to study dopamine-mediated plasticity? 8

1.2 Introduction to learning in the fruit fly 9

1.2.1 Classical conditioning in the fruit fly 9

1.3 Learning paradigms in the fruit fly: learning, extinction, re-learning and relief learning 12

1.4 Anatomy of the fruit fly olfactory system 13

1.4.1 Early stages of odour processing in the fruit fly brain 14

1.4.2 The Mushroom Body is the centre for associative learning in the fruit fly 16

1.4.3 Anatomical subdivisions of the mushroom body have unique role in odour processing 16

1.4.4 Anterior paired lateral (APL) neuron mediates recurrent inhibition in the mushroom body 17

1.5 Mushroom Body Output Neurons guide approach and retreat behaviours 18

1.6 Dopaminergic neurons convey reward and punishment signal 20

1.6.1 Dopaminergic neurons can control memory expression based on animal's state 22

1.6.2 Dopamine modulated learning mediates plasticity in the Mushroom Body 22

1.6.3 Dopaminergic neurons encoding opposite values are required for extinction . . . 23
1.7 Spike Timing Dependent Plasticity (STDP) - a candidate for olfactory plasticity . . . 23
1.8 Previous computation models of olfactory processing and learning . . . 24
1.9 Abstract model of learning in the Mushroom Body . . . 26
1.10 A phenomenological model of a neuron . . . 26
1.11 Questions addressed in this thesis . . . 27

2 Sparse decorrelated odour representation in the mushroom body 29
2.1 Abstract . . . 29
2.2 Introduction . . . 30
2.3 Methods . . . 32
2.3.1 Modelling olfactory receptor neurons responses . . . 33
2.3.2 Simulation conditions . . . 35
2.3.3 Neuron model . . . 36
2.3.4 Calculating separation between two odours . . . 37
2.3.5 Measuring spatial and temporal sparseness . . . 38
2.4 Results . . . 38
2.4.1 Decorrelation of odour representation in the olfactory circuit . . . 42
2.4.2 Role of mushroom body feedback inhibition in decorrelation and sparse representation . . . 47
2.5 Discussion . . . 50

3 Olfactory aversive learning and extinction in the fruit fly 52
3.1 Abstract . . . 52
3.2 Introduction . . . 53
3.2.1 Experimental data . . . 55
3.3 Methods . . . 57
3.3.1 Kenyon Cells model . . . 58
3.3.2 Connectivity . . . 59
3.3.3 Mushroom body output neurons and dopaminergic neurons . . . 60
3.3.4 Learning rules . . . 60
3.3.5 Rule 1: pre-dopa-LTD rule . . . 61
3.3.6 Choosing the threshold for the learning rule . . . 62
3.3.7 Reward-modulated spike-timing-dependent plasticity (R-STDP) learning rule . . . 63

3.3.8 Learning rates . . . 65
3.3.9 Valence and the performance index . . . 66
3.3.10 Simulating reward and punishment . . . 67
3.4 Results . . . 68
3.4.1 Tuning a spiking olfactory model of the mushroom body . . . 68
3.4.2 Dopa-pre-LTD rule can reproduce aversive learning and extinction experiments . . . 69
3.4.3 Aversive extinction requires connections between opposing valence coding compartments . . . 72
3.4.4 R-STDP rule can reproduce aversive extinction learning experiment . . . 77
3.4.5 Pre-dopa-LTD rule can reproduce aversive learning experiment with no post spikes . . . 78
3.4.6 Aversive re-learning decreases only MBON V2 firing rate . . . 79
3.4.7 Second order conditioning forms two opposing memories . . . 81
3.5 Discussion . . . 82

4 Olfactory appetitive learning and extinction in the mushroom body 85
4.1 Abstract . . . 85
4.2 Introduction . . . 85
4.3 Methods . . . 86
4.4 Results . . . 87
4.4.1 Appetitive learning and extinction . . . 87
4.4.2 Prediction: Blocking PAM-γ5Ans during odour exposure leads to an aversive memory . . . 90
4.4.3 Extinction is blocked if reinforcement is present during re-exposure 92
4.5 Discussion . . . 93

5 Investigating the role of potentiation in the mushroom body 95
5.1 Abstract . . . 95
5.2 Introduction . . . 96
5.2.1 Experimental data . . . 98
5.3 Methods . . . 99
5.3.1 Neuron model . . . 99
5.3.2 Rule pre-dopa-LTP-LTD . . . 100
5.4 Results . . . 102

5.4.1 Relief learning can be explained by potentiation of KC-MVP2 synapses . . . 102
5.4.2 Re-learning with LTP enables re-extinction . . . 104
5.4.3 Aversive re-learning and re-extinction . . . 105
5.4.4 Appetitive re-learning and re-extinction . . . 107
5.4.5 Re-learning in the four-layer olfactory circuit . . . 109
5.5 Discussion . . . 113

6 Conclusions and future work **115**
6.1 The final learning rule . . . 115
6.2 Conclusions . . . 116
6.3 Future work . . . 117

List of Figures

1.1 Appetitive and aversive training protocols in the fruit fly 11
1.2 Experimental Performance Index calculated for learning, extinction and relief experiments. 13
1.3 Diagram of the complete olfactory circuit from odour detection to motor biasing neurons 15
1.4 Schematic of the circuits within the mushroom body lobes. 17

2.1 Fruit fly olfactory circuit model diagram 33
2.2 Response curves for 16 ORN types to 110 odours, adapted from Hallem and Carlson, 2006) 35
2.3 Average number spikes each odour elicits 38
2.4 Network simulation of a single odour 39
2.5 Network simulation of a single odour with spike-frequency adaptation 40
2.6 Voltage trace of the membrane potential for PNs and KCs. 41
2.7 Measure population and temporal sparseness 42
2.8 Network activity during presentation of four different odours. 43
2.9 Olfactory Receptor Neurons, Projection Neurons and Kenyon Cells responses to odours in the Hallem and Carlson, 2006 data set 44
2.10 Lateral inhibition facilitates pattern decorrelation in the antennal lobe. 45
2.11 APL mediated feedback inhibition facilitates pattern decorrelation in the MB. 46
2.12 Average angular distance in PNs and KCs as a function of lateral inhibition strength. 47
2.13 Feedback inhibition facilitates odour pattern decorrelation in the mushroom body. 48
2.14 KC-APL mediated inhibition increases angular separation between odours and decreases number of responding KCs per odour 49
2.15 Network responses to odours in each layer 49

3.2 The circuit diagram of the mushroom body illustrating Kenyon Cells (KCs), Mushroom Body output neurons (MBONs) and Dopaminergic neurons (DAns). . . . 57
3.3 Learning rules: Rule 1 (dopa-pre-LTD) . . . 61
3.4 Learning Rule 2 (R-STDP) . . . 64
3.5 Calculating odour valence based on MBON firing rates elicited by the odour. . . . 67
3.6 Tuning KC-MBON connections to yield equal firing rates for retreat and approach MBONs . . . 68
3.7 Background dopamine and presynaptic levels do not change KC-MBON synapses. . . . 69
3.8 Network activity during aversive training using rule 1 . . . 70
3.9 MBON Firing rates after learning protocol. . . . 71
3.10 Changes in the MBON firing rates as number of shocks during training are increased . . . 72
3.11 Rule 1: Network activity during extinction learning protocol. . . . 73
3.12 MBON Firing rates after aversive extinction protocol. . . . 74
3.13 Rule 1: Network activity during extinction learning protocol with MVP2-M6 inhibitory connection lesioned. . . . 75
3.14 M6 and PAM-γ5 neurons after aversive learning and aversive extinction 76
3.15 Changes in MBON M6 firing rate as number of extinction trials is increased. . . . 77
3.16 Network activity during appetitive extinction simulation with R-STDP learning rule . . . 78
3.17 Network activity during aversive training with no postsynaptic spikes 79
3.18 Network activity during aversive re-learning protocol . . . 80
3.19 MBON Firing rates after aversive re-learnprotocol. . . . 81
3.20 Network activity during second order conditioning . . . 82

4.1 Network activity during appetitive learning protocol . . . 87
4.2 MBON Firing rates after appetitive learning protocol. . . . 88
4.3 Network activity during appetitive extinction protocol . . . 89
4.4 MBON Firing rates after appetitive extinction protocol. . . . 90
4.5 Suppressing PAM-γ5opaminergic neurons leads to the formation of an aversive memory. . . . 91
4.6 Appetitive extinction with different learning rates. . . . 92

4.7 Rule 1 (pre-dopa-LTD): Network activity during failed extinction learning protocol. . . . 93

5.1 Pairing odour with a shock with different inter-stimulus intervals . . . 97
5.3 Learning rule pre-dopa-LTD-LTP . . . 100
5.4 Shock learning result depends on relative timing of shock and odour onset when using rule 3. . . . 103
5.5 Rule pre-dopa-LTD-LTP reproduces relief learning. . . . 104
5.6 Network activity during aversive re-learning protocol using rule pre-dopa-LTD-LTP . . . 105
5.7 MBON Firing rates after aversive re-learning protocol. . . . 106
5.8 Rule pre-dopa-LTD-LTP enbles aversive re-extinction. We show the average weight for CS+ KC-MVP2 (blue) an KC-M6 (purple) connections during aversive re-extinction simulation. . . . 107
5.9 Network activity during appetitive re-learning protocol using rule pre-dopa-LTD-LTP . . . 108
5.10 MBON Firing rates after appetitive re-learning protocol. . . . 108
5.11 Network activity during appetitive re-extinction protocol using rule pre-dopa-LTD-LTP . . . 109
5.12 Fruit fly olfactory circuit model diagram . . . 110
5.13 Learning, extinction, re-learning and re-extinction in the full network. 112
5.14 Stages of learning. . . . 114

Chapter 1

Introduction

Abbreviations

AL - antennal lobe
APL neuron - Anterior paired lateral neuron
DAns - Dopaminergic neurons
KC - Kenyon cell
LN - Local Neuron
LTD - Long-term depression
LTP - Long-term potentiation
MB - mushroom body
MBON - Mushroom body output neurons
ORN - Olfactory Receptor Neuron
PN - Projection Neuron
SFA - Spike-frequency adaptation
STDP - Spike-timing dependent plasticity

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 motivated by their desires. 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 maximises pleasure. Although the feeling 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 that can affect 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:

$$\text{Dopamine Response} = \text{Reward occurred} - \text{Reward predicted} \tag{1.1}$$

Experiments described above have proposed a link between dopaminergic activity and reward motivated behaviour. Today, one of the fundamental goals of neuroscience is understanding 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 Synaptic plasticity

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 one 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. Plasticity 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 the role of spikes in the induction of LTP and LTD is. 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 synaptic plasticity

Neuromodulators cause alteration of the response properties and synaptic efficiency of individual neuron of one or more neurons (Marder and Thirumalai, 2002).

Multiple classes of neuromodulators have been shown to affect the gain of input-output responses in the neurons of mammals and flies (Thurley et al., 2008, (Waddell, 2013). They have also been implicated in facilitating synaptic plasticity in visual, auditory and olfactory learning (Gu, 2002).

This is 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.

It's been proposed that neuromodulators can alter behaviour 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). As such, 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). Another team also managed to induce LTP of corticostriatal synapses and block LTP again by dopamine receptor blocking (Reynolds and Wickens, 2002). 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 GABergic inhibition, which led to more postsynaptic spikes and LTP (Bissière et al., 2003). However, when GABAergic activity was entirely suppressed pharmacologically, Bisierre observed that STDP was blocked even in the presence of dopamine suggesting the mechanism of interaction is complex.

In another in vitro study, 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 (Zhang et al., 2009).

A separate study has shown that high DA receptor activation promotes potentiation of corticostriatal synapses onto the direct pathway and learning from positive experiences, while low DA receptor activation promotes potentiation of corticostriatal synapses onto the indirect pathway and learning from negative experiences (Shen et al., 2008).

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 perhaps 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 (RL), which studies strategies for an agent to learn through trial and error in an environment. RL uses 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 $t = 1, 2, 3 \ldots$. 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:

$$V(s) \leftarrow V(s) + \alpha \underbrace{[r + \gamma V(s') - V(s)]}_{\text{TD error}}, \tag{1.2}$$

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 in update 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 estimated value of the new state, 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. The TD error can be explained as:

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

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. Schultz observed 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 neuroscientists 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. Equally

unclear is 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 Solving the distal reward problem

In the Pavlovian experimental paradigm reward is delayed from the stimulus. The outcomes of action that can be manifested as a reward or punishment often occur seconds after the actions that caused them.

The distal reward problem represents the problem of associating a reward with an earlier action when the neural activity that caused that action is no longer present is called the distal reward problem.

One key component of the learning paradigm are eligibility traces employed to represent a stimulus. Eligibility traces are used to represent the degree to which a stimulus is eligible for consideration when attempting to attribute rewards to actions or stimuli. When a stimulus is detected, an eligibility trace for that action is initiated and immediately begins to decay with time.

In the Izhikevich, 2007 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:

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

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:

$$\frac{d}{dt}w_{ij}(t) = c(t)d(t) \tag{1.5}$$

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:

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

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

Other studies have proposed modified versions of the STDP learning rule that use a third factor to represent reward. For exampe, Legenstein et al., 2008 have shown that they can learn a biofeedback task and showed the ability to memorize temporal spike patterns.

In a separate study Clopath et al., 2008 proposed a learning framework that can explain pharmacological experiments performed on neurons (Frey et al., 1997). The learning mechanism has three stages: tagging (when the synapse is in either high or low state which depends on Hebbian mechanisms it gets tagged), trigger (dopamine level above a certain threshold) and consolidation (if plasticity related proteins are available and if the synapse is tagged, the weight can change to a stable state allowing for consolidation).

Three factor learning rules using eligibility traces have also been used in insect models of learning. In Helgadottir et al., 2013, the authors have implemented a spiking model of the olfactory circuit using a learning rule inspired by experimental findings in the locust (Cassenaer and Laurent, 2012) to implement a learning rule in which octopamine gates learning. Importantly, this learning rule requires only presynaptic activity and a reward signal for learning to occur. The authors show that using this simple learning rule allows them to implement a SNN platform for biologically inspired closed-loop robotic control using sensors to learn object association with colours (Helgadottir et al., 2013).

In chapter 3 we modify the three-factor learning rule and apply it to fruit fly olfactory learning.

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

So 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 and extinction (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. It represents the normalized preference of the trained odour compared to the untrained one. Olfactory memories, and the valence of the odours are believed to be stored in the fly's brain through dopamine mediated plasticity forming memories that can be retrieved during behavioural testing (Fig 1.1).

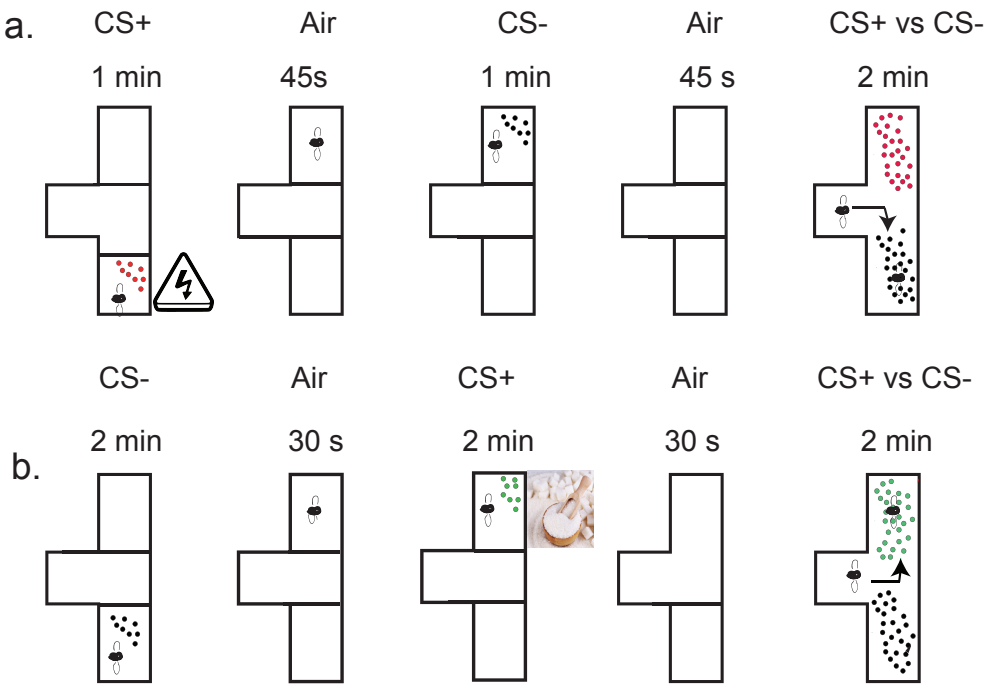


Figure 1.1: Aversive and appetitive training protocols in the fruit fly. **a.** Flies are trained in a T-Maze to test aversive memory. Initially 100 flies are exposed to an odour while being shocked (CS+) for 60 seconds. After 30 seconds of air they are exposed to a second odour (CS-) which is not paired 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 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.

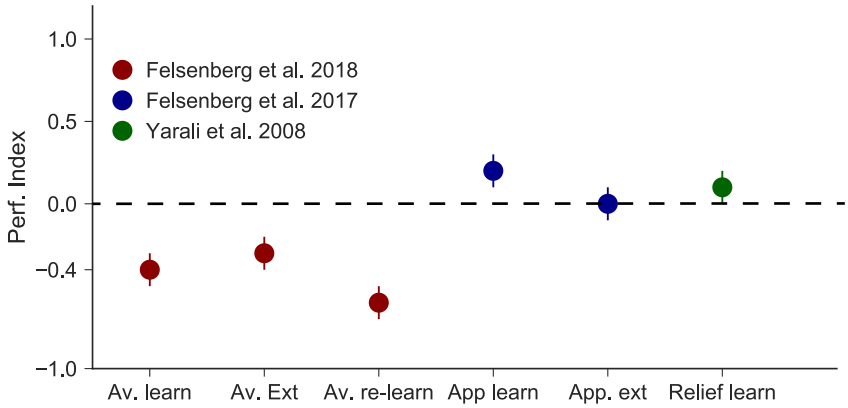


Figure 1.2: Experimental Performance Index calculated for learning, extinction and relief experiments. Performance Index calculates the proportion of flies that approach the reinforced CS+ odour compared to the CS-. Performance index from left to right is shown for the following experiments: aversive learning, aversive extinction and aversive re-learning (Felsenberg et al., 2018), appetitive learning, appetitive extinction (Felsenberg et al., 2017) and relief learning (Yarali et al., 2008). Vertical bars indicate standard error of the mean.

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. There, mushroom body output neurons (MBONs) guide behaviour (Aso et al., 2014). We shall now discuss the staging or processing of 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 three PNs (Vosshall et al., 2000).

In flies PNs can be divided in two major categories. Multigomerular PNs are GAGAergic and inhibitory and the majority of them project to the Lateral Horn and other region within the lateral protocerebrum (Galizia, 2014) . Uniglomerular PNs are excitatory are project to both the lateral horn and the mushroom body calyx (Galizia, 2014). We only model uniglomerular PNs that project to the mushroom body.

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 that each KC receives input from a range of 5-10 PNs. In Figure 3.2 we show a simplified version of the olfactory circuit and we note that DANs that represent the reward signal target different MBONs compared to DANs represent punishment. In fact there are more than 30 compartments that divide the mushroom body, and each compartments is innervated by mostly unique dopaminergic neurons (Aso et al., 2014).

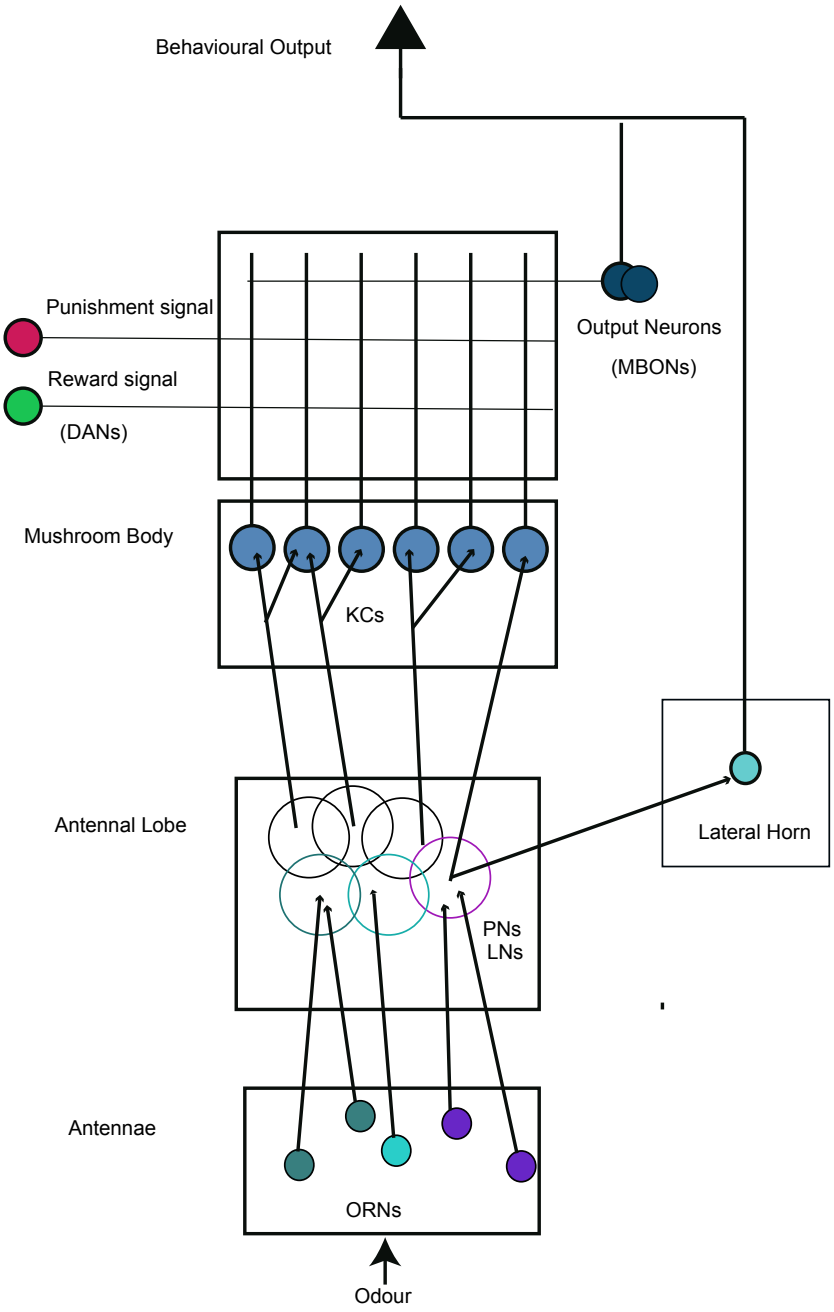


Figure 1.3: Diagram of the olfactory circuit from odour detection to motor biasing neurons (Adapted from Dylla et al., 2013, Fig. 2). Odours activate a subset of 1200 ORNs. ORNs expressing the same receptor converge to the same glomerulus in the antennal lobe. In the antennal lobe, activity is determined 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 of 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 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 mushroom body 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: $\alpha 1-3$, $\beta 1-2$, $\alpha' 1-3$, β' and γ 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 (Aso et al., 2014). 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 α'β' 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 into 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).

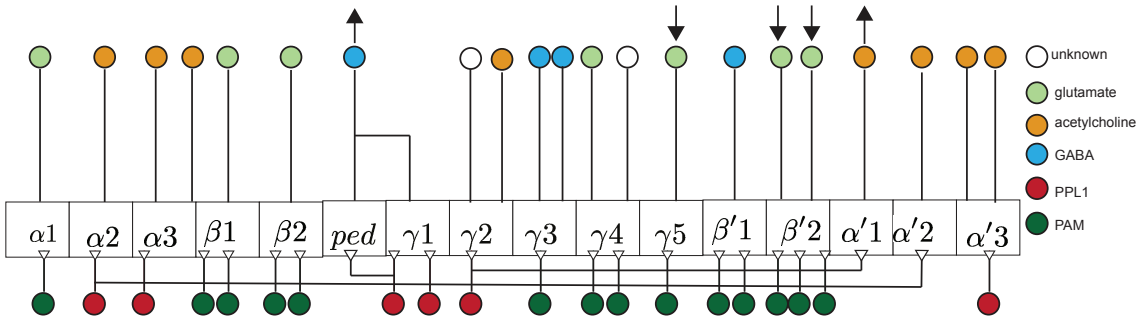


Figure 1.4: Schematic of the circuits within the mushroom body lobes. Adapted from Aso et al., 2014 DAN inputs (bottom) and MBON outputs (top) from 16 MB lobe compartments are shown. Colours indicate neurotransmitter types for MBONs. Three DAN cell types innervate multiple lobe compartments. Eight types 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. 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 Anterior paired lateral (APL) neuron mediates recurrent inhibition in the mushroom body

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

innervated 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 in normalizing KCs 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 project 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 crepine, 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 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 honeybees MBONs have been shown to encode valence allowing rewarded olfactory stimuli to be distinguished from non-rewarded stimuli and MBONs responded to a broad set of odours (Strube-Bloss et al., 2016), suggesting that the role of MBONs is similar across insects.

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. 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 and DAns

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 M4β'/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 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. A recent study has shown that MVP2 inhibits the dopaminergic neuron in its compartment and signalling from MVP2 is important for long term memory formation (Pavlowsky et al., 2018).

MBON-V2 is required for both STM and LTM retrieval

MBON-V2s innervate 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 (DAns) 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 DAns reside are the PPL and the PAM clusters. Positive reinforcement is provided by subsets of approximately one hundred DAns 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 (Waddell, 2013). 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 $\gamma 4$ 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-$\gamma 4$ neurons for STM (short term memory) while LTM (long term memory) water memories depending on the activity of PAM-$\beta' 1$ neurons. Negative reinforcement such as from electric shock or bitter substances appears to be conveyed by DANs housed in the PPL cluster. Interestingly the PAM-$\gamma 3$ 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-$\gamma 2$ and PAM-$\gamma 3$ show increased level of calcium activity, while PAM-$\gamma 4$ and PAM-$\gamma 5$ 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 innervating the MB vertical lobe as well as $\gamma 1$ and $\gamma 2$ and PAM-$\beta 2$ and PAM-$\gamma 3$. The dendrites of MBONs are restricted to few DA zones, which correspond to one type of reinforcement delivery (either aversive of appetitive). For example, in the case of the retreat MBON (M4β'), 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 DANs modulate the synapses onto retreat MBONs while punishment DANs modulate synapses onto approach MBONs.

1.6.1 Dopaminergic neurons 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 Dopamine modulated learning mediates plasticity in the Mushroom Body

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 DANs promote odour approach by depressing odour drive to avoidance promoting MBONs and possibly by strengthening approach pathways. Conversely, aversive responding DANs promote depressing odour drive to approach promoting MBONs (Aso et al., 2014). It is likely that a feedback loop from MBON to DANs influences learning as well. Blocking the M4β' neuron induced a depression of the KC-M4β' MBON connection (Owald et al., 2015), which could be explained by a scenario in which M4β' inhibits some appetitive DANs. Neurons in the mushroom body express at least two types of dopamine receptors: DAMB and dDA1. dDA1 has been shown to be required for both appetitive and aversive learning (Kim et al., 2007). DAMB has been shown to be richly expressed in the

mushroom bodies and has been speculated to mediate dopamine triggered increases in cAMP (Han et al., 1996).

1.6.3 Dopaminergic neurons 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. In the next section I discuss how dopamine can orchestrate plasticity in the fruit fly.

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 cannot 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 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:

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

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 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 show 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.

In Faghihi et al., 2017, the authors have proposed a model for second-order conditioning. In second order conditioning after CS+ trained with shock a second odour, CS2 is paired with the CS+ odour. After pairing, odour 2 will acquire the valence of odour 1, which is to say it will become aversive. The authors have proposed a mechanism of retrograde signalling via the MBON. After shock training KC-MBON synapses are strengthened which increases the amount of retrograde signal produced by the MBON which diffuses across neighbours KCs. In the model, retrograde signal triggers potentiation of KC-MBON synapses proportional to the activity of KCs. Thus, in second order conditioning when odour 2 is alternated with the trained odour, retrograde signalling will potentiate KC-MBON synapses for the untrained odour.

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.

1.9 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. Although 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 will introduce in the next section.

1.10 A phenomenological model of a 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. We've used the leaky integrate-and-fire model to simulate fruit fly neurons in the following chapter with conductance-based synapses like in the COBA model (Vogels and Abbott, 2005).

In many neurons, spike frequency adaptation has been observed, which reduces the firing rate of a neuron when constant current is injected. A modified model of the LIF is 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. We've used this model to represent firing patterns of neurons in the fruit fly that decay rapidly after responding to odours for less than one second, such as Kenyon Cells (Murthy et al., 2008).

1.11 Questions addressed in this 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 four results chapters 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 Abstract

The fruit fly, Drosophila Melanogaster can reliably detect odours using its antennae and can distinguish odours with very similar chemical structures. Transformations in the olfactory circuit give the fly the ability to distinguish many different classes of odours. 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 is 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 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 olfactory receptor neurons activity is transformed into spatio-temporally sparse KC activity in the mushroom body. We tune our model to elicit realistic firing activity and show that multiple mechanisms enable decorrelation and sparse odour representation 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 test our model using a data-set of ORN responses to 110 different odours

as input to our network and show that our network encodes the majority of odours from the data-set in a sparse and decorrelated manner. We also show that feedback inhibition reduces the average number of KCs that respond to an odour. Finally, we visualize MBON responses to odours in the data-set, which lays the groundwork to test learning rules that modify KC-MBON synapse, which we develop in the following chapters.

2.2 Introduction

Flies, like other animals are capable of identifying relevant stimulus features using neural computation that helps it achieve this goal. A large body of evidence has identified sparse stimulus encoding as an essential feature of neural coding in both vertebrates and invertebrates (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.

Previous modelling studies have shown that the sparse coding caused by lateral inhibition within the dentate gyrus of the hippocampus results in improved pattern separation (O'Reilly and McClelland, 1994) which is useful for distinguishing highly similar patterns.

Inhibition in the antennal lobe has been proposed by Cleland and Linster, 2005 to explain the experimentally observed impairment of sensory discrimination when inhibition is blocked. It has also proposed by a modelling study that lateral inhibition contributes to gain control of odour responses that explain experimental results in the honeybee mushroom body (Schmuker et al., 2011).

In a locust model of the antennal lobe, oscillations observed in the PN population at high odour concentrations have been reproduced by strong activation and periodic modulation of LNs which provide lateral inhibition (Bazhenov et al., 2001).

Another model of the locust antennal lobe has shown that excitatory connections between KCs can facilitate the preservation of the temporal order of which PNs get activated in response to odours which increases the number of spatio-temporal patterns that KCs can recognize (Nowotny et al., 2003) .

Pattern separation, or orthogonalization, is characterized by creating distinct, non-redundant representations in an attempt to reduce the likelihood of interference during retrieval (O'Reilly and McClelland, 1994).

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 normalise odour responses by gain modulation depending on the concentration of the odour, and it is believed that the connectivity between PNs and KCs is random with each KC receiving input from approximately 10-12 PNs (Turner et al., 2008).

Experiments have shown that when exposed to an odour for just one second, KCs show adaptation immediately after odour onset (Murthy et al., 2008). Thus, KCs are believed to respond to odours in both a spatially and temporally sparse manner.

One mechanism that has been shown to facilitate temporal sparseness in insects and vertebrates (Farkhooi et al., 2013) is spike-frequency adaptation (SFA), which can be explained by spike dependent activation of potassium currents such as the calcium-activated K+ current that is activated by the increase in Ca2+ at the peak of each action potential, which triggers a form of slow negative feedback to the excitability of the cell (Benda and Herz, 2003). K+ currents are only activated after spiking activity. These currents cause an after-hyperpolarizing potential (AHP) after a spike or burst of spikes (Benda and Herz, 2003).

We model SFA using leaky integrate-and-fire neurons with an adaptation variable that increases each time the neuron spikes and falls back to zero when the neuron is silent (see Methods, Eq 2.1).

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 build a network in which KCs respond in a sparse manner, while MBONs respond to every odour. This model forms the basis of testing plasticity the site of KC-MBONs synapses. To test transformations of a dense odour representation we use a data-set that recorded 24 ORN type responses to 110 different odours as input into our network (Hallem and Carlson, 2006). We tune our network to elicit realistic responses in 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 are sufficient to decorrelate odour representations. In chapter 4 we show that by decreasing overlap between odours, 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 improve previous models (Wessnitzer et al., 2011, Betkiewicz et al., 2018) based on recent experimental data (Aso et al., 2014, Lin et al., 2014a) and based on these results we added feedback inhibition in the mushroom body mediated by a single GABAergic neuron and separated our MBONs into two categories: approach and retreat MBONs to calculate the valence of an odour in the naive state before learning. This allows us to study how odours get represented in terms of MBON firing rates after transformations in the antennal lobe and the mushroom body.

2.3 Methods

We modelled a four-layer olfactory circuit (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 of odour representation in the mushroom body. Our model includes feedback inhibition in the mushroom body and a fourth layer (MBONs) that receives inputs from KCs.

Our model is composed of 24 ORN types, 24 PNs, 24 LNs, and 1000 KCs. The connections between the four network layers are feed-forward and excitatory. LNs 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 all-to-one, while the APL also inhibits all KCs in our model. ORNs provide excitatory 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 KC). Each KC receives input from an average of 10 PNs based on experimental data (Turner et al., 2008). Finally, KCs excite two MBONs (M6 and MVP2) (Fig 2.1).

In the data-set (Hallem and Carlson, 2006) there are a number of odours which mostly elicit inhibitory responses in ORNs, five of which we could not get KCs and MBONs to respond to. We removed the following odours from the simulation: ammonium hydroxide, putrescine, cadaverine, trans-caryophyllene and eugenol.

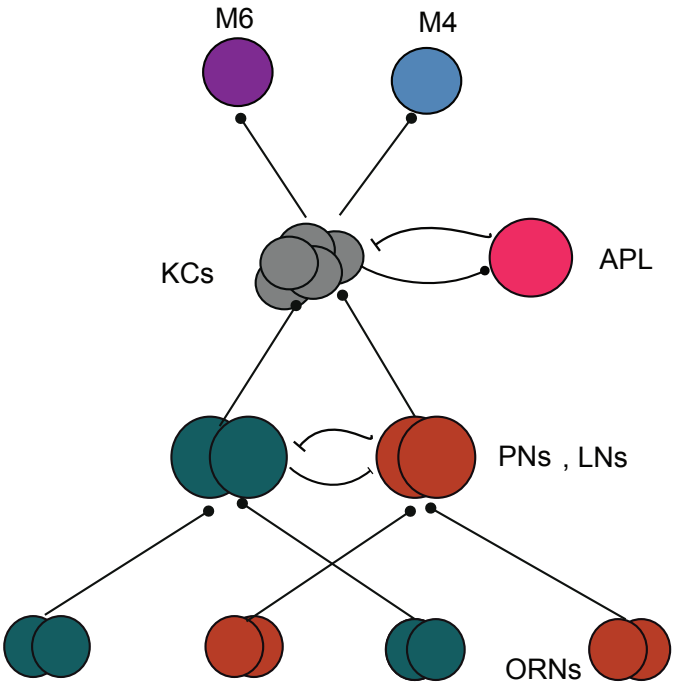


Figure 2.1: Olfactory circuit model diagram. Circuit diagram shows the main components of our model starting from bottom: 24 ORN types, 4 of each kind. ORNs of the same type excite 1 unique PN and 1 LN which form a glomerulus. LNs inhibit PNs which are not in the same glomerulus. PNs excite 1000 KCs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

2.3.1 Modelling olfactory receptor neurons responses

We used a 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 reproduced a figure from the original paper which plots the tuning curves of the 24 ORNs in response to 110 odours (Fig 2.2). ORNs have diverse tuning curves, some ORNs respond to a broad range of odours

while others are narrowly tuned. We modelled ORN responses based on the available dataset by transforming the firing rates into Poisson spike trains. We transformed the firing rates from the data-set into homogeneous Poisson spike trains for each ORN.

To efficiently generate Poisson spike trains from such a configuration we consider a grid spanned by N rows corresponding to the Poisson neurons (y-axis) and discrete-time with bins of size Δ t=0.1 seconds (x-axis). To create Poisson spikes, we fill each column at each time step during the simulation by drawing exponentially distributed inter-spike-intervals. The distribution of inter-spike-intervals is the same both on the x-axis (time) and y-axis (neuron number).

When a jump leads beyond the number of neurons in the Poisson population it is simply continued in the next time step. This way every random number yields a spike.

Using the ORN data-set, we simulated responses to a set of four odours over 50 trials, each odour lasting 5 seconds. We selected four odours to study how diverse PN and KC responses are to stimuli in the data-set. We used the entire data-set of odours to study the role of lateral inhibition in the PNs and feedback inhibition in the KCs in decorrelating odour representations.

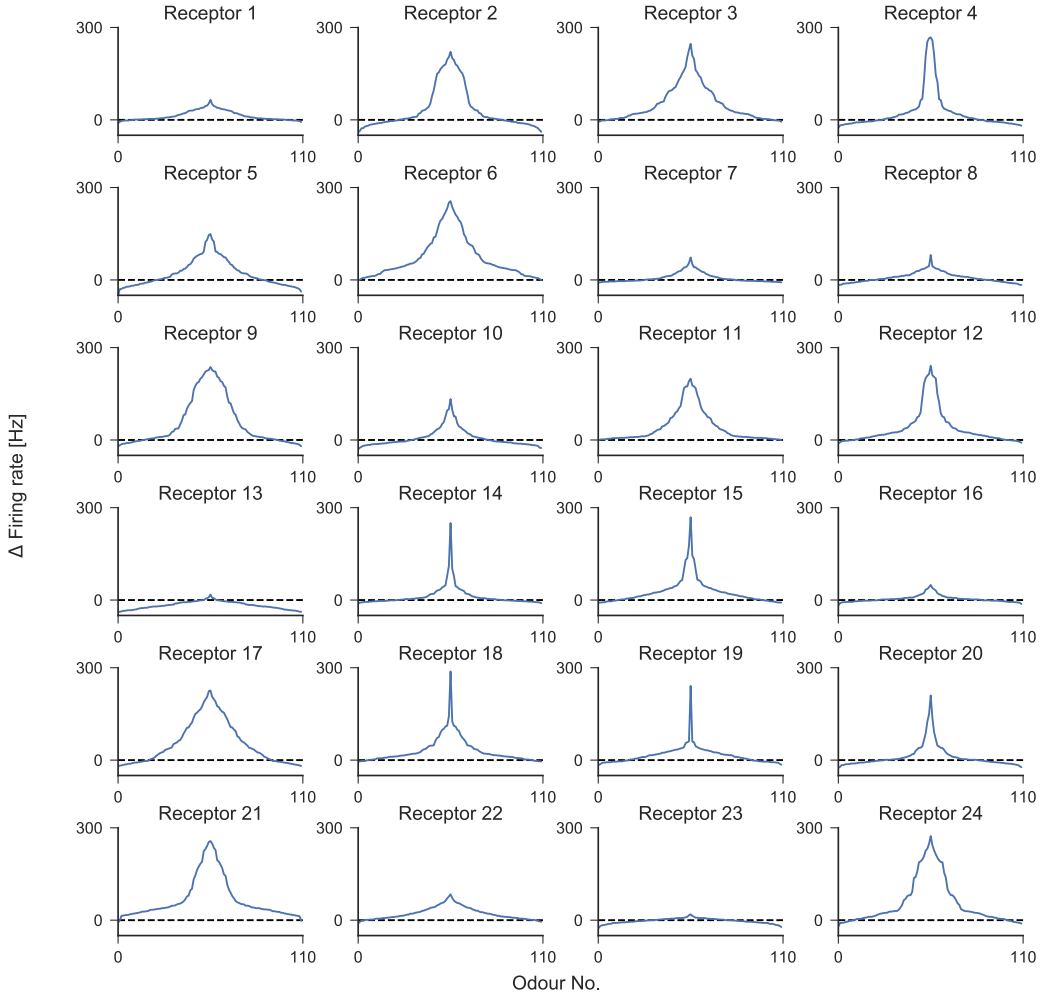


Figure 2.2: Response curves for 24 ORN types to 110 odours, figure adapted from Hallem and Carlson, 2006. Odours are sorted along the x axis according to the response strengths they elicit from a receptor. Odours that elicit strongest responses to the receptor shown are displayed near the centre of the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

2.3.2 Simulation conditions

To study the effect of different cellular mechanism on olfactory processing we deactivated and activated spike-frequency adaptation in PNs and KCs, lateral inhibition in the antennal lobe and feedback inhibition in the mushroom body. 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 their input from all of the 1000 KCs.

2.3.3 Neuron model

PN, LN, KC and MBONs were modelled using the leaky integrate-and-fire neuron with spike-frequency adaptationusing conductance based synapses (Gerstner and Brette, 2009). The membrane potential is described by the following equation:

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

where 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. $g_a(t)$ represents adaptation and evolves according to

$$\tau_w \frac{dg_a}{dt}(t) = -g_a(t) + b \sum_k \delta(t - t_k^*), \quad (2.2)$$

where τ_w is the adaptation time constant. When the membrane voltage reaches a predefined threshold a spike is fired which resets the membrane voltage to a reset value and $g_a(t)$ is the adaptation of the neuron, modelled by a change in conductance which increased by b at each spike. Finally, δ denotes the Dirac δ function and t_k^* is the timing of the k-th spike.

The synaptic conductances are expressed by $g_{ex}(t)$ (excitatory) and $g_{inh}(t)$ (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: $g_{ex} \rightarrow g_{ex} + \Delta g_{ex}$ 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}	5 ms
Inhibitory synaptic potential	E_{inh}	-80 mV
Excitatory synaptic potential	E_{ex}	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}	3 nS
ORN-LN Synaptic weight	w_{orn-ln}	1 nS
LN-PN Synaptic weight	w_{orn-ln}	2 nS
Synaptic parameters with adaptation		
ORN-PN Synaptic weight	w_{orn-pn}	1.2 nS
PN-KC Synaptic weight	w_{pn_kc}	2.5 nS
Adaptation parameters		
Spike frequency adaptation	b	0.1 nS
Spike adaptation time constant	τ_w	100 ms

Table 2.1: All neurons parameters (ORN, LNs, PNs, KCs, MBONs).

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 between each pair of odours:

$$\text{Angular distance }_{i,j} = 1 - cos(\theta_{ij}), \tag{2.5}$$

where $cos(\theta)$ is:

$$cos_{ij}(\theta) = \frac{n_i \cdot n_j}{\|n_i\| \cdot \|n_j\|}. \tag{2.6}$$

The cosine of the angle between these vectors represents the similarity measure between the vectors. n_i and n_j represent the spike train of each neuron in the population in response to odour i and odour j respectively. The value of similarity ranges from 0 (for orthogonal vectors of $\theta = 90°$) to 1 (for parallel vectors of $\theta = 0$).

2.3.5 Measuring spatial and temporal sparseness

We used the Treves-Rolls measure for calculating both temporal and population sparseness (Willmore and Tolhurst, 2001):

$$S = 1 - \frac{(\sum_j^N r_j/N)^2}{\sum_j^N r_j^2/N} \tag{2.7}$$

where r_j denotes one neuron's firing counts in response to the j-th bin (bin size =20 ms) during the two seconds of odour exposure (in this case, lifetime sparseness can be obtained) or the firing counts of the j-th neuron out of the group during a given period of the stimulation (in this case, population sparseness can be calculated).

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). In the first layer ORN Poisson spike trains were modelled with spike rates based on experimental recording of different odours responses.

We plotted the average number of ORN spikes elicited by each odour in the data-set (Hallem and Carlson, 2006) after subtracting spontaneous ORN activity. Some odours do not elicit many ORN spikes (less than 50 spikes on average) while others activate ORNs strongly (more than 100 spikes on average) (Fig 2.3).

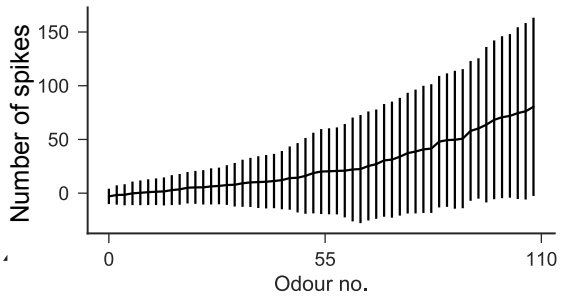


Figure 2.3: We show the average number of spikes elicited by each odour in the ORN population based on the Hallem and Carlson, 2006 data-set. The average number of spikes for each of the 110 odours in response to all of the ORNs. Vertical bars indicate one standard deviation.

As we discussed in the methods section, we removed five of the odours that elicited the fewest spikes. For those five odours, we could not elicit KC and MBON responses.

In our data-set of 105 odours, we initially presented a single odour to our network to study its response. Without spike frequency adaptation (Fig 2.4 b, d) PNs, and KCs fired at a constant firing rate while the odour was present. The odour we presented only activated a subset of KCs (Fig 2.4 a), while the majority of PNs were active (Fig 2.4 c).

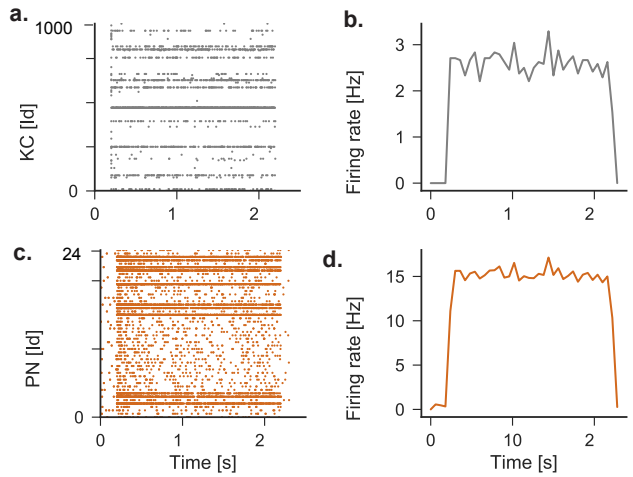


Figure 2.4: Network activity during simulation of one odour with no spike frequency adaptation **a.** Raster plot of KCs responses to one odour simulated for 2 seconds. Each dot (grey) represents a spike. **b.** Population firing rate of KCs during simulation of one odour. **c** Like in **a.** we show a raster plot of PNs. **d.** Like in **b.**, we show the population firing rate of PN.

After adding spike frequency adaptation, PNs excited by the stimulus showed a strong response at stimulus onset, followed by an adaptation during the stimulus period (Fig 2.5 d). KCs followed a similar temporal dynamic (Fig 2.5 b). The majority of PNs responded to the odour, while only a subset of KCs was activated (Fig 2.5 a, c).

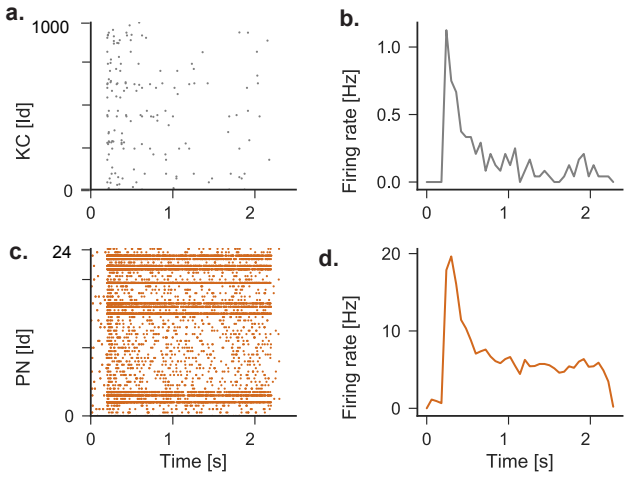


Figure 2.5: Network activity during simulation of one odour with spike frequency adaptation enabled **a.** Raster plot of KCs responses to one odour simulated for 2 seconds. Each dot (grey) represents a spike. **b.** Population firing rate of KCs during simulation of one odour. **c** Like in **a.** we show a raster plot of PNs. **d.** Like in **b.**, we show the population firing rate of PN.

We looked at odour activity of single PNs and KCs from a trial of odour exposure. The odour activated PN responded throughout the duration of the odour (Fig 2.6 a), while the KC fired the majority of spikes in the first second (Fig 2.6 b). We also show a silent KC that was not activated by the odour (Fig 2.6 c). Silent KCs are important for maintaining population sparseness.

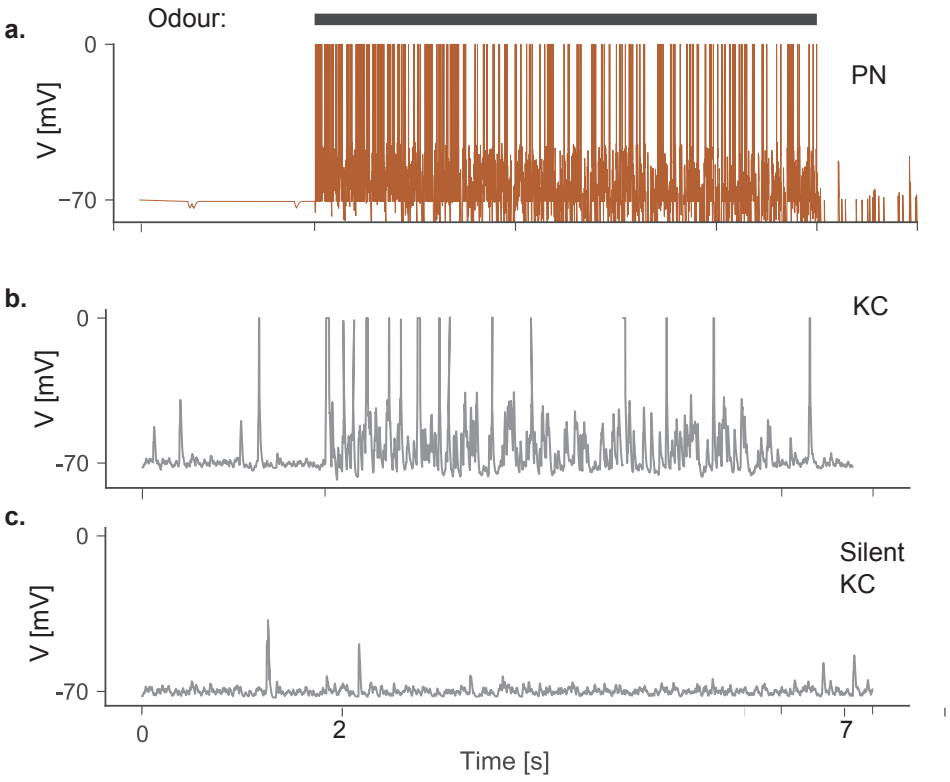


Figure 2.6: Voltage trace of membrane potential for PN and KC. An odour was simulated for five seconds from t=2 s (black bar). **a.** An odour activated PN is randomly selected. Orange trace, single trial. **b.** An odour activated KC is randomly selected from the same trial as in **a.** to show mem. pot. (grey) **c.** Another KC was randomly selected to view the membrane potential from a KC that was silent during odour presentation.

We calculated temporal and spatial sparseness of KC responses to odours in our simulations using the Treves-Rolls sparseness measure (Treves and Rolls, 1991, see Methods). Our results show that spike-frequency adaptation increases temporal sparseness (Fig 2.7 a) while APL mediated feedback inhibition facilitates population

sparseness (Fig 2.7 b).

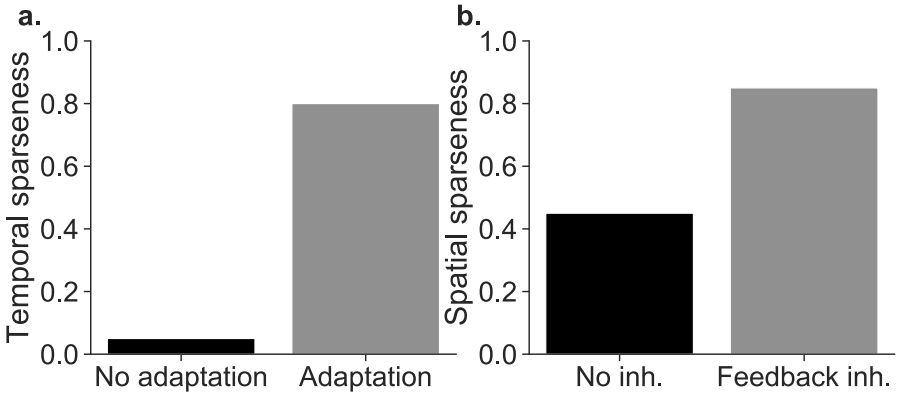


Figure 2.7: Measuring temporal and spatial sparsness. **a.** We measured temporal sparseness in the absence of spike-frequency adaptation (black) and with spike-frequency adaptation activated (grey). **b.** We also measured population sparseness without APL mediated feedback inhibition in the mushroom body (black) and with feedback inhibition activated (grey).

2.4.1 Decorrelation of odour representation in the olfactory circuit

We selected four odours from the data-set (1-butanol, 2-pentanone, isobutyl acetate and 1-proponal) and visualized our models' response at each layer. PNs, KCs and MBONs responded to all four odours with different average firing rates (Fig 2.8 a,b,c). The APL neuron which provides feedback inhibition to Kcs, also responded to all four odours (Fig 2.8 d).

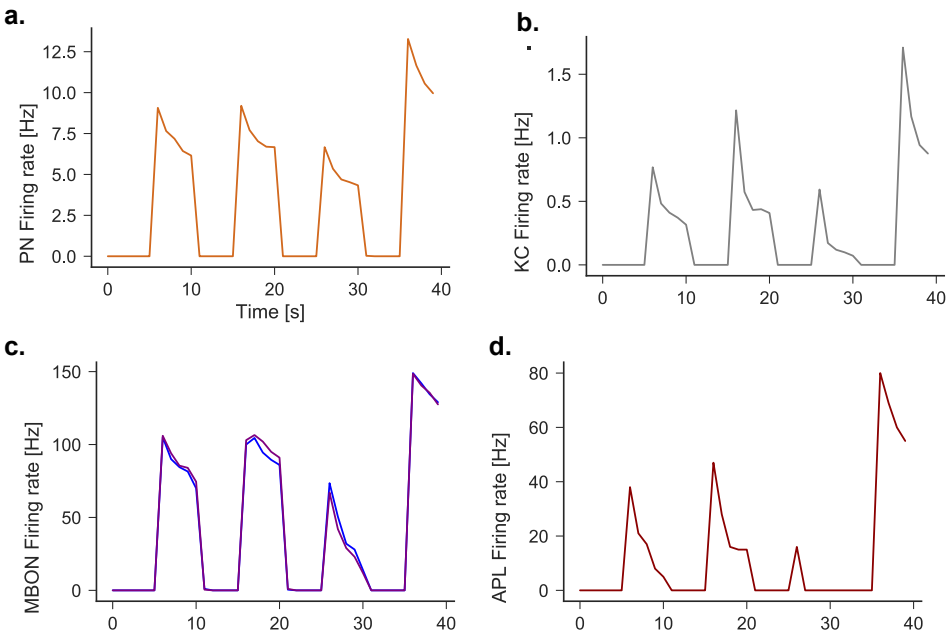


Figure 2.8: Network activity during presentation of four different odours. **a.** Shown here are the population firing rates of PNs (orange) in response to four odours, each exposed for 5 seconds. **b.** Similarly to **b.** we show the population firing rate of KCs in response to the four odours. **c.** 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. **d.** Firing rate of the GABA-ergic APL neuron during the simulation of four odours.

We plotted the response of each neuron to all 110 odours in the ORN, PN and KC layers (Fig 2.9 a, b, c). Most odours activate a large number of ORNs and PNs, while the majority of KCs were silent to each odour (Fig 2.9 c). Our results show that KCs activate between 50 and 160 KCs on average (Fig 2.9 d).

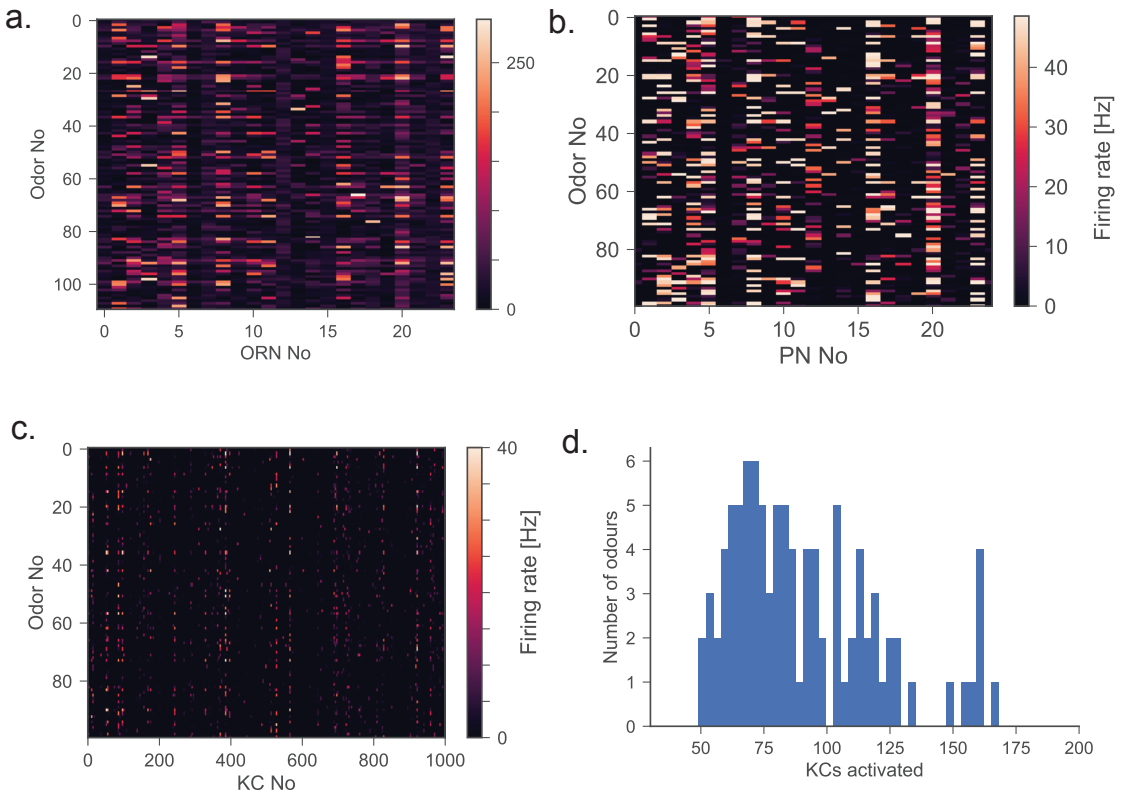


Figure 2.9: ORN, PN and KC responses to odours based on ORN responses in the Hallem and Carlson, 2006 data set. **a.** We show a color map of ORN responses to 110 different odours in the data set (Hallem and Carlson, 2006). **b.** Shown here are 24 PN responses after simulating PN responses to the same odours show in A **a.** and in the same format. Each group of three PNs receive input from one out of 24 ORN types. **c.** Shown here are simulated responses of 1000 KCs responses to odours in the data set with the same format as in **a.** and **b.**. **d.** Histogram showing how many KC an odour activates.

In order to investigate the role of lateral inhibition in the antennal lobe and APL mediated feedback inhibition in the MB, we simulated the four previous odours responses (Fig 2.8) by activating or deactivating each mechanism separately. We deactivated lateral inhibition and adjusted the ORN-PN weights to maintain the same population firing rates in PNs and KCs.

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 plotted angular distance between the odours in the PN layer with and without lateral inhibition and in the KC layer (Fig 2.10 a, b, c) and although mean of angular trend seem to be increasing when lateral inhibition is activated, this difference in not

significant (Fig 2.10 d, one-way ANOVA p=0.15).

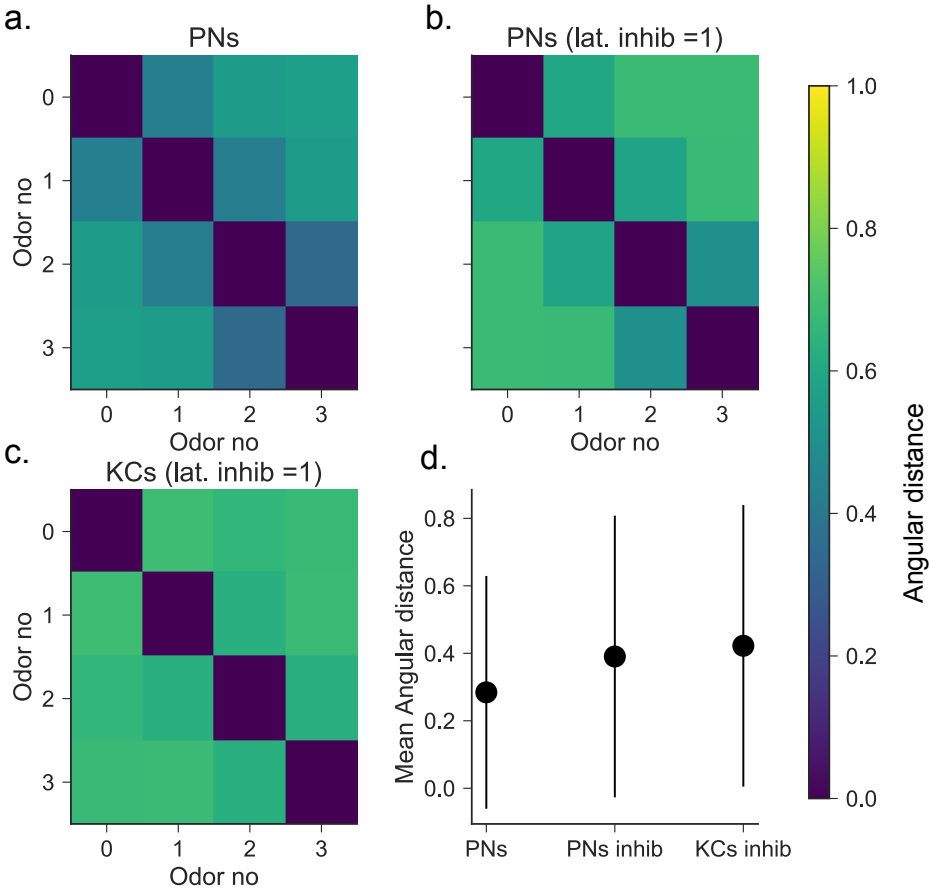


Figure 2.10: Lateral inhibition facilitates pattern decorrelation in the MB. We simulated four odours from the Hallem and Carlson, 2006 data-set. **a.** We show odour pattern correlation (using angular distance) in the PN population with no lateral inhibition. **b.** We show pattern correlation in PNs 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 indicate 1 standard deviation.

We showed that lateral inhibition increases dissimilarity of odour patterns in the PN layer. By switching off and on APL feedback inhibition, we found a trend that shows feedback inhibition leads to the increase of the mean angular separation between odour patterns represented by KC activity (Fig 2.11, one-way ANOVA p=0.083).

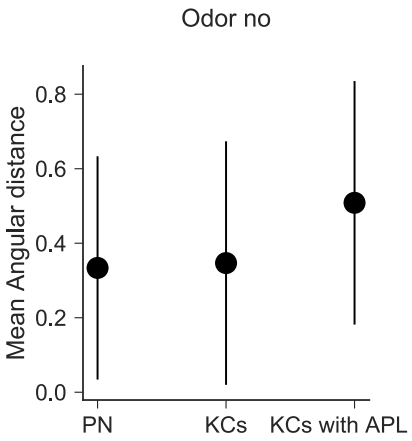


Figure 2.11: APL mediated feedback inhibition facilitates pattern decorrelation in the MB. We simulated four odours from the Hallem and Carlson, 2006 data-set. We illustrate the the mean angular distance for PNs, KC and KCs with APL feedback inhibition active. Vertical bars illustrate one standard deviation.

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 2.12 blue) and in the KC layer (Fig 2.12, purple) and found that separation of odour patterns increased monotonically 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 2.12) shaded blue, shaded purple).

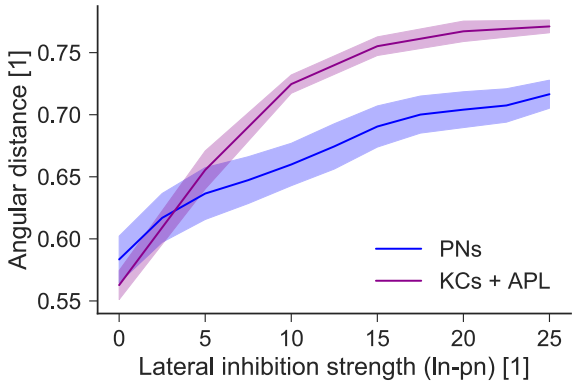


Figure 2.12: Average angular distance in PNs and KCs as a function of lateral inhibition. We show changes in average pattern similarity (angular distance) calculated across four stimuli in the PN (blue) and KC (purple) layers as we increase lateral inhibition. Shaded areas indicate 1 standard deviation.

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 Role of mushroom body feedback inhibition in decorrelation and sparse representation

We also investigated the angular separation of the odour response in PNs and KCs for all the remaining 105 odours in the data-set. We plot pairwise angular distance for all odours in the Hallem and Carlson, 2006 data-set. We plot angular distance in PN population, KCs with no feedback inhibition and KCs with APL mediated feedback inhibition (Fig 2.13 a,b,c). Our results show that APL mediated feedback inhibition increases mean angular separation across the entire population of odours (Fig 2.13 d, one-way ANOVA p=0.065).

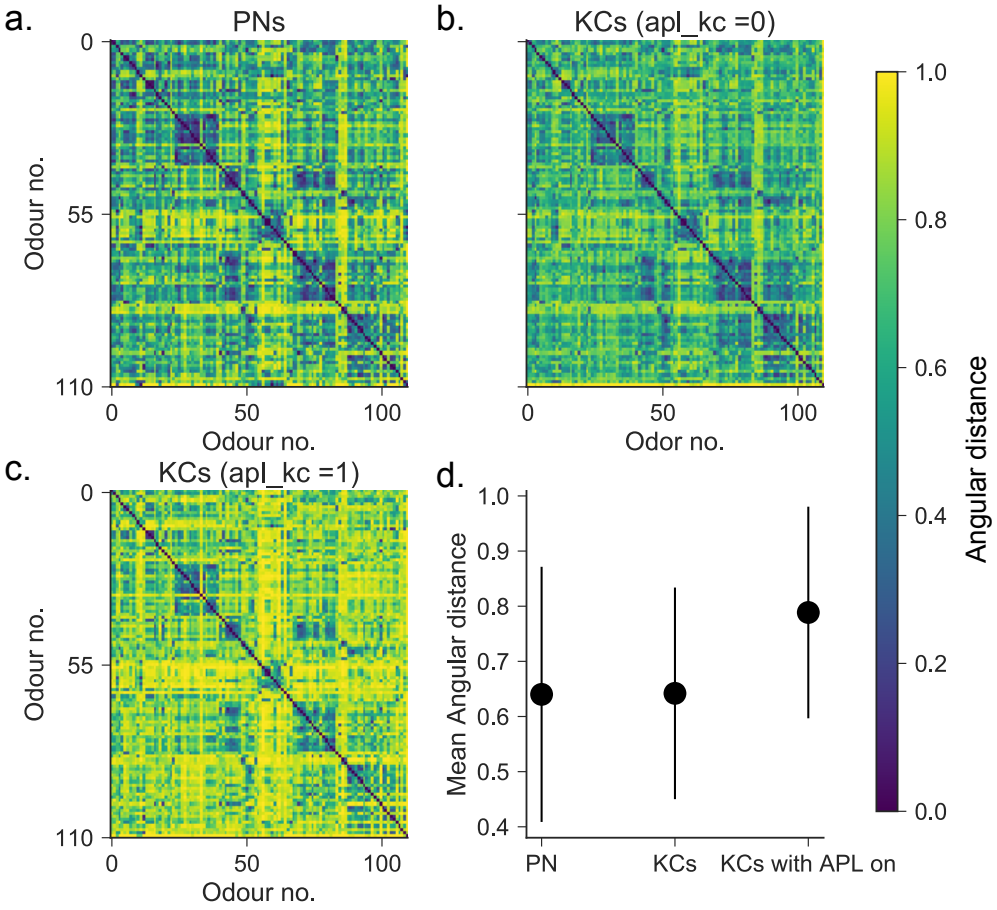


Figure 2.13: APL mediated feedback inhibition facilitates pattern decorrelation in the MB. We simulated 110 odour patterns from the Hallem and Carlson, 2006 data-set and calculated odour odour similarity for all pairs using angular distance. **a.** shows odour similarity based on angular distance in 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.** Vertical bars indicate one standard deviation.

APL mediated feedback inhibition has been shown to be important to maintain sparse KC population responses to odours. 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 2.14 a) and the number of odours that an individual KC responds to (Fig 2.14 b).

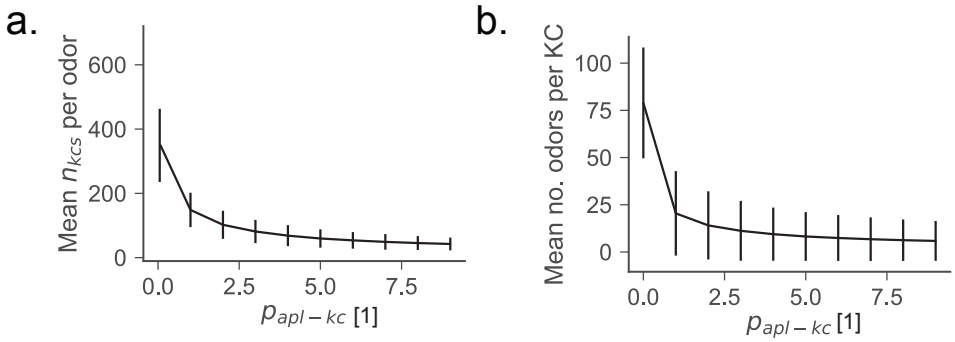


Figure 2.14: KC-APL mediated inhibition decreases number of responding. **a.** We show the number of KCs responding to an odour on average as a function of strength of APL mediated feedback inhibition. **b.** Shown here how many odours does a single KC respond to on average as a function of APL mediated inhibition. Vertical bars indicate 1 standard deviation.

We plotted the firing rate of each of the 105 odours from the Hallem and Carlson, 2006 data-set. All of the odours elicited a response in the MBONs ranging between 20 and 100 Hz (Fig 2.15). The odours elicited approximately equal firing rates in the approach MVP2 MBON and retreat M6 MBON. In the next chapter we define this as an odour being in a naive state. In chapter 3 we introduce learning rules that can change the firing rates elicited by an odour.

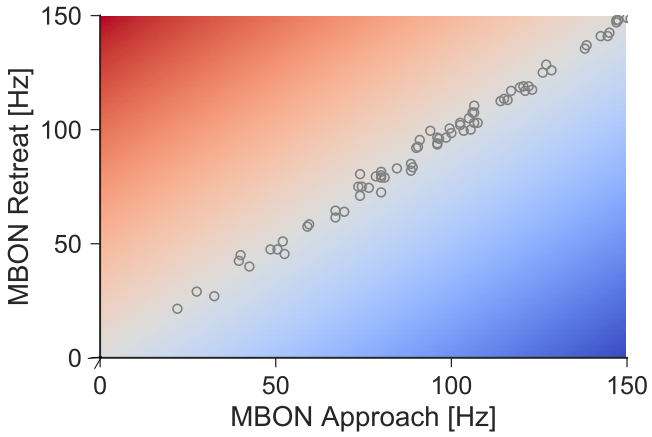


Figure 2.15: Here we show the firing rates elicited by MBONs in response to 110 odours we simulated based on the Carlson data set of recorded ORN responses. The diagonal represents odours that elicit equal approach and retreat MBON firing artes.

2.5 Discussion

In this chapter we laid the foundations for 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 24 PNs corresponding to the 24 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, as ORN have been found to inhibit each other (Su et al., 2012). 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 odours 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 increase 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 using a data-set of recorded ORN responses to odours (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). The olfactory circuit of the fruit fly is characterized by recurrent connections with both excitatory and inhibitory synapses. Does the network maintain a balanced state between excitation and inhibition? A number of theoretical studies have suggested that a balanced state can improve functionality of the network (Vogels and Abbott, 2009). In our results we had difficulties obtaining responses to all of the odours because some odours activate ORNs very weakly. A balanced network with a gain modulation system could potentially solve that problem (Stroud et al., 2018).

In in the next chapters we explore plasticity at KC-MBON synapses. Plasticity has also been observed in the antennal lobe with a role in habituation. Lateral inhibition (LN-PN) activity is strengthened after repeated exposures to the same odour (Sudhakaran et al., 2012). A theoretical study has proposed that plasticity in the antennal lobe could improve decorrelation performance of ORN odour responses that have significant overlap (Kasap and Schmuker, 2013). In this study, they used the inhibitory synaptic plasticity learning rule, which adjusts the strength of inhibitory connections (Vogels et al., 2011). The authors showed that through inhibitory plasticity, lateral inhibition can reduce input correlation and overlap of odour representation in the antennal lobe and improve classification performance (Kasap and Schmuker, 2013). In this chapter we have shown that a four-layer olfactory circuit model can create sparse decorrelated odour representations that propagate from ORNs to MBONs, which provides the basis for exploring plasticity at in the mushroom body.

Chapter 3

Olfactory aversive learning and extinction in the fruit fly

3.1 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 pre-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. Experiments have shown that the fly can form a memory when postsynaptic spikes are blocked. Our results show that both learning rules can reproduce aversive learning and extinction experiments. However, only rule pre-dopa-LTD can learn when postsynaptic spikes are supressed, while rule R-STDP fails. 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 are 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-γ5 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 larvae 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 two-layer model (KCs, MBONs and DANs) to study dopamine modulated plasticity (Fig 3.2). We show that after aversive learning both appetitive PAM-γ5 DANs and aversive $PPL1$ DANs increase their response to the aversive odour. We show that increased PAM-γ5 response can signal extinction during odour re-exposure, while increased PPL1 DAn response can be used to perform second order conditioning.

3.2.1 Experimental data

We developed learning rules to reproduce aversive learning and extinction experiments as seen in Felsenberg et al., 2018. Our learning rules were tuned to reproduce changes observed in CS+ responses in MVP2, M6 and M4 MBONs after aversive learning and after aversive extinction. In a previous study, it has been shown using electrophysiology the firing rate of MVP2 MBON in response to the CS+ odour before and after aversive learning (Hige et al., 2015a). Furthermore, it was also shown that the firing rate of MVP2 in response to the CS- odour did not change. Our first assumption in developing our learning rule is that response to the CS- odour does not change.

In the study we focus on reproducing, the authors imaged CS+ and CS- responses in M6 and M4 MBONs using two-photon fluorescence microscopy. The baseline fluorescence (F_0) represents mean fluorescence from 2 s before odour presentation. $F/F0$ represents fluorescence relative to baseline. The area under the curve was measured as the integral of F/F_0. To account for variance between flies, the responses of the CS+ and CS- were normalized to responses to a novel odour (isoamyl acetate). Results showed that after aversive learning response to CS+ in M6 MBON was enhanced compared to CS- (Fig 3.1 a). However, after extinction, M6 CS+ response

was approximately equal to CS- (Fig 3.1 b). Intriguingly, the M4 MBON increases its response to the CS+ odour after aversive learning and maintains this after extinction (Fig 3.1 c,d). The authors have also shown that the M6 MBON excites the $PAM-\gamma5$ DAn in its compartment while the M4 MBON does not seem to excite the DAn in its compartment (Felsenberg et al., 2018). Helmchen et al., 1996 have shown that calcium imaging of neurons is related to spike train. In our results, we respect the trends observed in changes in the mean calcium imaging of M6, M4 and MVP2 imaging but we do not convert the results directly into firing rates. We also do not reproduce the variance observed, which the authors believe is due to individual differences between flies (Felsenberg et al., 2018).

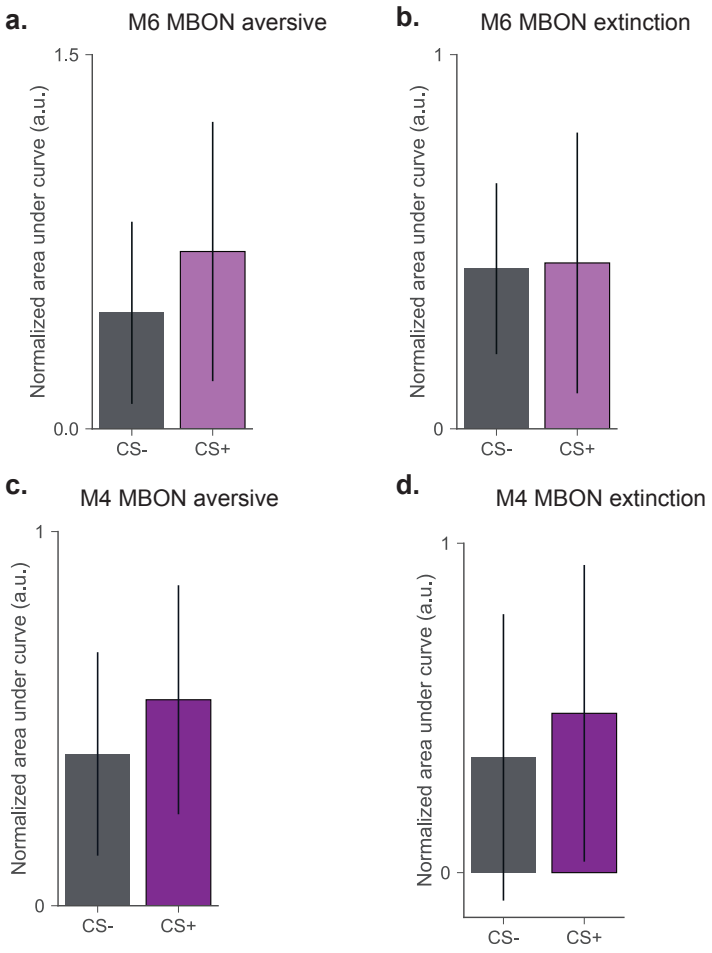


Figure 3.1: GCamp imaging of MBON M6 and M4 responding to 5 seconds of odour exposure. **a.** M6 MBON CS+ and CS- responses after aversive learning **b.** M6 MBON CS+ and CS- responses after aversive extinction **c.** M4 MBON CS+ and CS- odour responses after aversive learning **d.** M4 MBON CS+ and CS- odour responses after aversive extinction. Vertical bars represent one standard deviation.

3.3 Methods

To study olfactory associative learning in the fruit fly we modelled a two-layer spiking neural network of the mushroom body. We have modelled three types of neurons: KCs, MBONs and DANs (Fig 3.2).

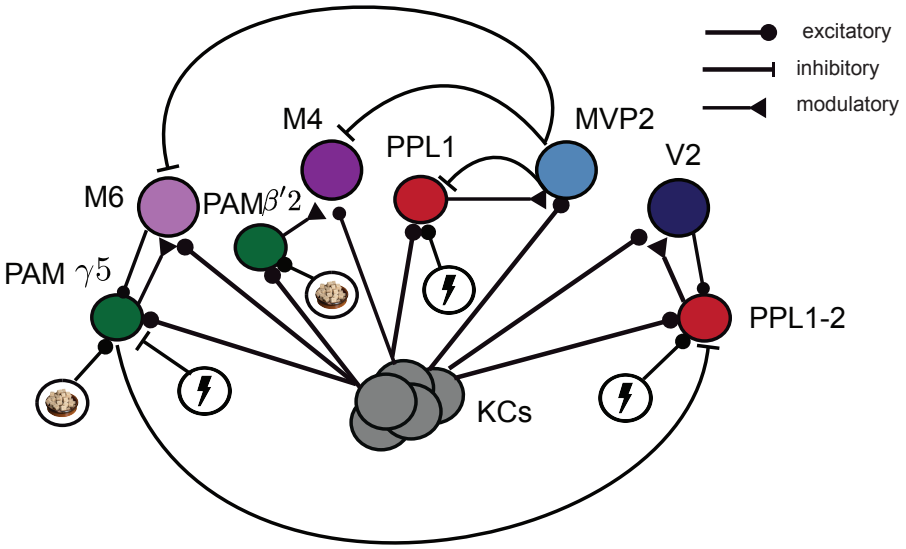


Figure 3.2: Circuit diagram of our model of the mushroom body illustrating KCs, MBONs and DANs. Schematic of our circuit which is composed of 1000 KCs (grey) exciting four MBONs (MVP2 - light blue, V2 -dark blue, M6 - light purple, M4 -dark purple). M6 MBON and M4 MBON are inhibited by the MVP2 MBON. All four DANs in our model (PPL1, PPL1-2, PAM-$\gamma5$ and PAM-$\beta'2$ receive excitatory input from KCs; PAM-$\gamma5$ excited by external reward and by MBON M6, and inhibited by shocks; PPL1-2 is excited by external shocks and by MBON-V2 and also inhibited by PAM-$\gamma5$; PPL1 excited by shock and PAM-$\beta'2$ is is excited by reward.

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	$MBON\gamma 1ped$	Approach
PPL1	$PPL1\gamma 1ped$	Shock reinforcement
Compartment 2		
V2	$MBON\alpha V2sc$	Approach
PPL1-2	$PPL1\alpha'3$	Shock reinforcement
Compartment 3		
M6	$MBON\gamma 5\beta'2a$	Retreat
PAM-$\gamma 5$	PAM-$\gamma 5$	Reward reinforcement
Compartment 4		
M4	$MBON\beta'2mp$	Retreat
PAMβ'	$PAM\beta'2mp$	Reward reinforcement

Table 3.1: Neurons modelled in our simulations.

3.3.1 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:

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

where $f(t)$ is:

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

In our simulations p_0 is 10 and τ_{kc} is 1 *second*. 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 of 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.

We chose to model KCs as Poisson processes to focus on the learning rule and to be able to make sure there is no overlap between CS+ and CS- odour representation. In chapter 2, we created a more complete model, however responses to each odour were variable and with different degrees of overlap. We initially want to study our learning

rules under conditions that we can control (both odours elicit similar responses in MBONs before learning and there is no overlap). In chapter 5 we revisit our four-layer model and show that we can use our learning rule in more realistic scenarios.

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 like in electrophysiological experiments 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 which has been shown to happen in the fruit fly (Owald et al., 2015).

All DANs receive excitatory input from KCs, which a recent connectome of the fly's larvae has shown to occur (Eichler et al., 2017b). PAM-γ5 is excited by the M6 MBON and inhibited by the presence of shock which is based on experimental data showing PAM-γ5 decrease their activity during shock exposure and that activation of M6 MBON increase the activity of PAM-γ5 (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 (Pavlowsky et al., 2018). In our model PPL1-2 is inhibited by PAM-γ5, 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	Symbol	Value
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 γ5-PPL1-2 synaptic weight	$w_{PAM5-PPL1-2}$	4 nS

Table 3.2: Network connectivity parameters.

3.3.3 Mushroom body output neurons and dopaminergic neurons

To model MBONs and DANs we used the AdEx model we introduced in chapter 2 (Eq 2.1). 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.

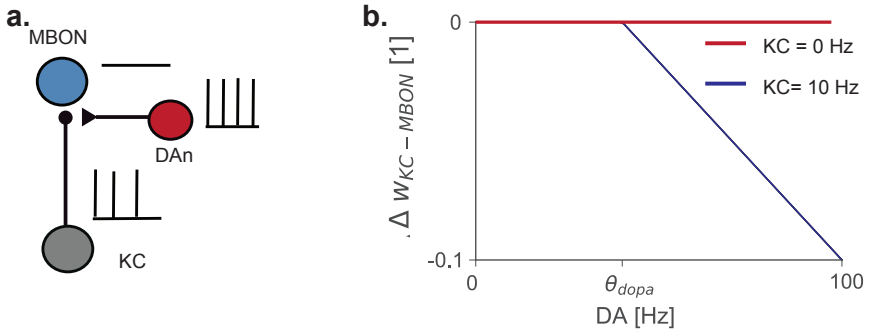
3.3.4 Learning rules

We implemented two types of learning rules. Rule pre-dopa-LTD requires only presynaptic activity and dopamine to change the weight between KC and MBON

synapses (Fig 3.3). 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 3.4).

3.3.5 Rule 1: pre-dopa-LTD rule

Our first learning rule (pre-dopa-LTD) is based on experimental evidence showing that postsynaptic spikes appear to be dispensable for learning (Fig 3.3 a, Hige et al., 2015b). In our implementation, when both KCs and DANs are active, KC-MBON synapses will be modified. Importantly, weights are bounded and cannot take values below zero. We also require MBON activity to be above a threshold, to exclude baseline activity levels (Fig 3.3 b).



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

Figure 3.3: **a.** Learning rule 1 (dopa-pre-ltd). Rule pre-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})).

Weight dynamics ($w_{ij}(t)$) are given by:

$$\frac{dw_{ij}(t)}{dt} = -A_{-} \underbrace{e_j(t)}_{presynaptic} \underbrace{S_{dopa}(t)\mathcal{H}(z_{dopa}(t) - \theta_{dopa})}_{dopamine}, \tag{3.3}$$

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_{dopa} to be above θ_{dopa}, which is a threshold set

to be equal to background odour activity. $\mathcal{H}$ is the Heaviside function. The eligibility trace e_j is updated each time the presynaptic neuron j fires as spike as follows:

$$\tau_e \frac{de_j(t)}{dt} = -e_j(t) + S_j(t), \tag{3.4}$$

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:

$$S_j(t) = \sum_k \delta(t - t^*_{j,k}), \tag{3.5}$$

where $t^*_{j,k}$ is the timing of the k-th spike of the j-th neuron.

Similarly, $S_{dopa}(t)$ is the spike train of the dopaminergic neuron:

$$S_{dopa}(t) = \sum_k \delta(t - t^*_k), \tag{3.6}$$

where t^*_k is the timing of the k-th spike of the dopaminergic neuron. The dopamine trace, $z_{dopa}(t)$ represents a global eligibility trace that increases each time the dopamine neuron fires a spike:

$$\frac{dz_{dopa}(t)}{dt} = -\frac{z_{dopa}(t)}{\tau_d} + S_{dopa}(t), \tag{3.7}$$

where τ_d is the characteristic time of z_{dopa} and $S_{dopa}(t)$ is the spike train of the dopaminergic neuron.

3.3.6 Choosing the threshold for the learning rule

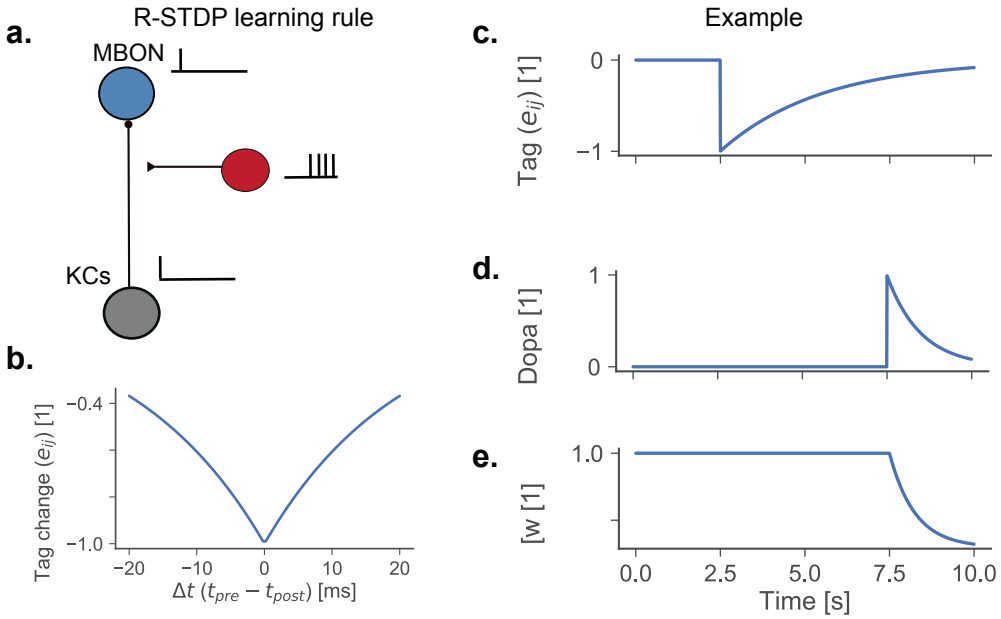
We implemented background level dopamine activity, based on experimental evidence which shows that DANs are active even in the absence of reinforcement (Dylla et al., 2017). To subtract baseline activity from the learning rule, we added a threshold that is equal to the firing rate of the dopaminergic neuron during odour exposure multiplied by the time constant (τ_z) of the dopaminergic eligibility trace (z_{dopa}). This ensures that learning occurs only when DANs fire above baseline activity. Parameters can be seen in Table 3.5.

Parameter name	Symbol	Value
Dopamine trace time constant	τ_z	2 s
MBON KC-MVP2 threshold	$\theta_{dopa-MVP2}$	50
MBON KC-M4 threshold	$\theta_{dopa-M4}$	50
MBON KC-M6 threshold	$\theta_{dopa-M6}$	75
MBON KC-V2 threshold	$\theta_{dopa-V2}$	75

Table 3.5: Rule 1: pre-dopa-LTD parameters.

3.3.7 Reward-modulated spike-timing-dependent plasticity (R-STDP) learning rule

We also explored a second lecond learning rule, in which synapses keep track of pre-post activity by updating a synaptic eligibility trace. We base our three factor learning on a previous model by Izhikevich, 2007, that requires presynaptic, postsynaptic and dopaminergic activity to modify the strength of the synapse (Fig 3.4 a). 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. A pre-post spike pair creates an eligibility trace that decays over time (Fig 3.4 c). The initial value of the eligibility trace $e_{ij}(t)$ depends on the relative timing difference between pre and post spikes (Fig 3.4 b) If there is dopamine activity before the pre-post spike decays to zero (Fig 3.4 d), the KC-MBON connections corresponding to the activity KCs will depress (Fig 3.4 e).



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

Figure 3.4: **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 ($e_{ij}(t)$) . **c.** We show the evolution of the eligibility trace ($e_{ij}(t)$) 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.

Here, weight dynamics of weight (w_{ij}) are given by:

$$\frac{dw_{ij}(t)}{dt} = -A_{-} e_{ij}(t) S_z(t) \mathcal{H}(z_{dopa}(t) - \theta_{dopa}), \tag{3.8}$$

where similar to our previous 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:

$$\tau_e \frac{de_{ij}(t)}{dt} = -e_{ij}(t) + \underbrace{x_j(t) S_i(t) + y_i(t) S_j(t)}_{H(pre_j, post_i))}, \tag{3.9}$$

and the tag ($e_{ij}(t)$) 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

$$S_i(t) = \sum_k \delta(t - t^*_{i,k}), \tag{3.10}$$

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. $e_{ij}(t)$ 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:

$$\frac{dx_j(t)}{dt} = -\frac{x_j(t)}{\tau_-} + S_j(t) \tag{3.11}$$

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

$$\frac{dy_i(t)}{dt} = -\frac{y_i(t)}{\tau_-} + S_i(t), \tag{3.12}$$

where τ_- is the characteristic time of the y_i and x_j eligibility traces. When the dopamine eligibility trace ($z_{dopa}(t)$) is above threshold θ_{dopa}, and the synaptic eligibility trace ($e_{ij}(t)$) 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	10 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 compartment 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. We tuned our learning rates (Table 3.7) 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	$A{-}_{MVP2}$	7×10^{-7}
Learning rate KC-V2	$A{-}_{V2}$	4×10^{-7}
Learning rate KC-M4	$A{-}_{M4}$	5×10^{-7}
Learning rate KC-M6	$A{-}_{M6}$	5×10^{-7}

Table 3.7: Rule 1: pre-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:

$$\text{Valence} = \frac{2r_{V2} - r_{M6} - r_{M4}}{2r_{V2} + r_{M6} + r_{M4}}, \tag{3.13}$$

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, Fig 3.5 a). We calculate the model Performance Index by subtracting the valences of the reinforced $CS+$ odour and the $CS-$ as follows:

$$\text{Performance Index} = Valence_{CS+} - Valence_{CS-}. \tag{3.14}$$

In our simulations the valence of the neutral CS- odour does not change, thus the performance index is determined by the valence of the CS+ (Fig 3.5 b).

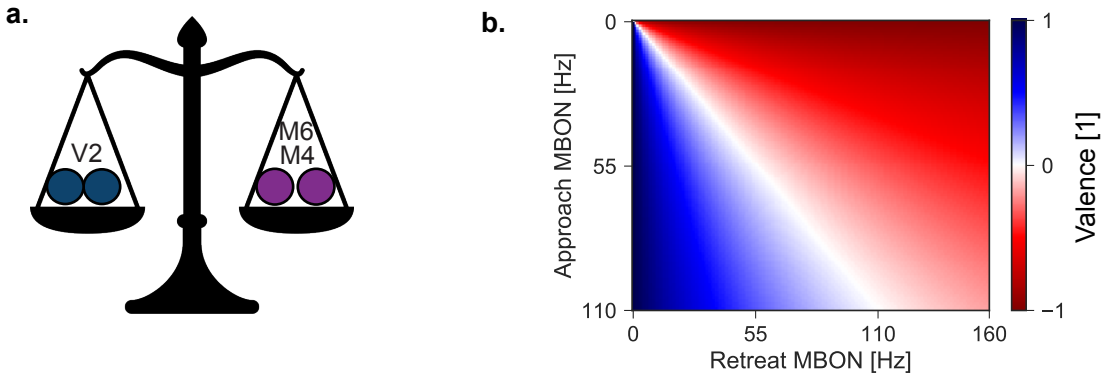


Figure 3.5: Calculating odour valence based on MBON firing rates elicited by the odour. **a.** Valence in our model is calculated by comparing the firing rates elicited by approach biasing V2 MBON to the retreat biasing M4 and M6 MBONs according to Eq 3.13. MVP2 MBON is excluded from calculated the valence 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 100-200 Hz each time we simulate a shock for 1 seconds. Similarly, for reward we increase the firing rate of PAM-γ5 and PAM2 dopaminergic neurons to approximately 50-100 Hz depending on the magnitude of the reward. Parameters for simulating reward and punishments can be viewed in Table 3.8.

Parameter name	Symbol	Value
PPL1 background	$r_{ppl-bkg}$	10 Hz
PAM background	$r_{pam-bkg}$	10 Hz
PPL1 shock	$r_{ppl-shock}$	50-100 Hz
PAM reward	$r_{pam-reward}$	50-100 Hz

Table 3.8: DAN neuron firing rates.

3.4 Results

In this section we discuss our results obtained 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 the 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 from KCs to both MBON-M6 and MBON-M4 (Fig 3.6). We vary the connection strength to both MBON-MVP2 and MBON-M6 to elicit equal firing rates for the two neurons (Fig 3.6). We choose the parameters that yield firing rates of approximately 100 Hz for all of our MBONs (Fig 3.6 a). 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 3.6 b).

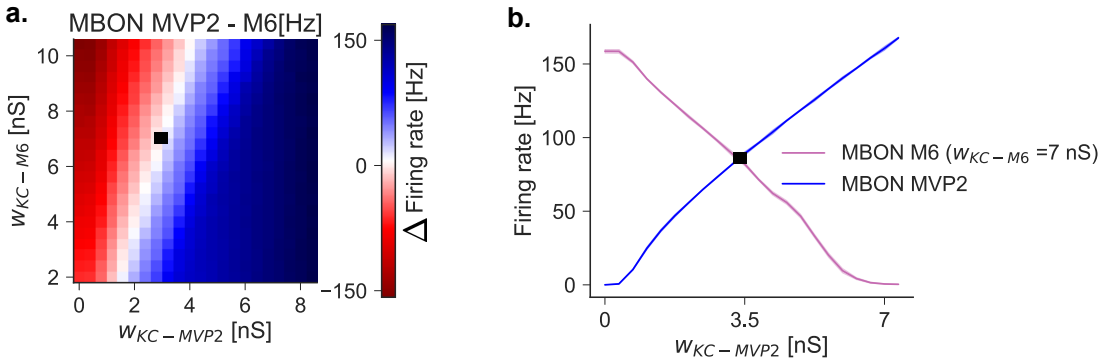


Figure 3.6: Tuning KC-MBON connections to yield equal firing rates for retreat and approach MBONs. **b.** As in a. we show the relationship between MVP2 and M6 firing rate for different tuning of initial connection strengths. On the x-axis we show the KC-MVP2 connection strength while on the y-axis we 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 (Fig 3.3 a) and DA neurons fire at 5 Hz (Fig 3.3 b). Our learning rule modifies KC-MBON connections

only if DAn activity is above a threshold as described in Eq 3.3. Background dopamine activity rates do not modify KC-MBON weights (Fig 3.7 c).

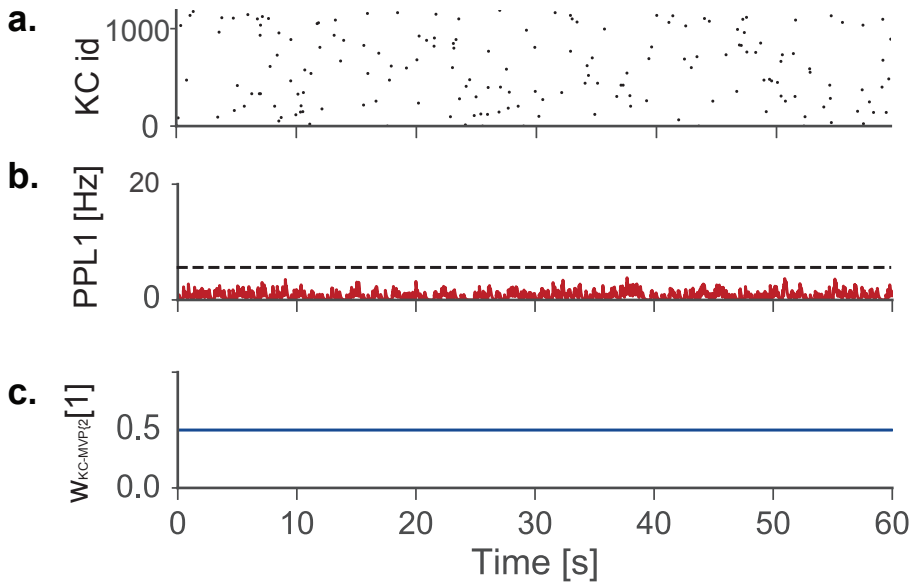


Figure 3.7: No changes in KC-MBON weights occurs during spontaneous activity. Showing spontaneous network activity with no odour present. **a.** We show a raster plot of 1000 KCs during 60 seconds of spontaneous activity. Each point represents an individual spike. **b.** We show background dopamine levels 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 obtained from testing Rule pre-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). Odour presentation led to a transient increase of KC population firing rate (Fig 3.8 a) lasting for 1 second and decaying afterwards, while simulation of shocks increased the firing rate of PPL1 DANs (Fig 3.8 d) each time a shock was presented. KC activity at odour onset created an eligibility trace for each active KC (Fig 3.8 f) with $\tau_j = 10s$, while shock presentation increased the dopamine eligibility trace above the threshold for plasticity (Fig 3.8 d). The dopamine and presynaptic traces multiplied together lead to the decrease of CS+ KC-MVP2 average weight (Fig 3.8 d). After odour exposure, MBON responses to CS- were unchanged (Fig 3.8 a, h), while M6 CS+ response increased (Fig 3.8 a).

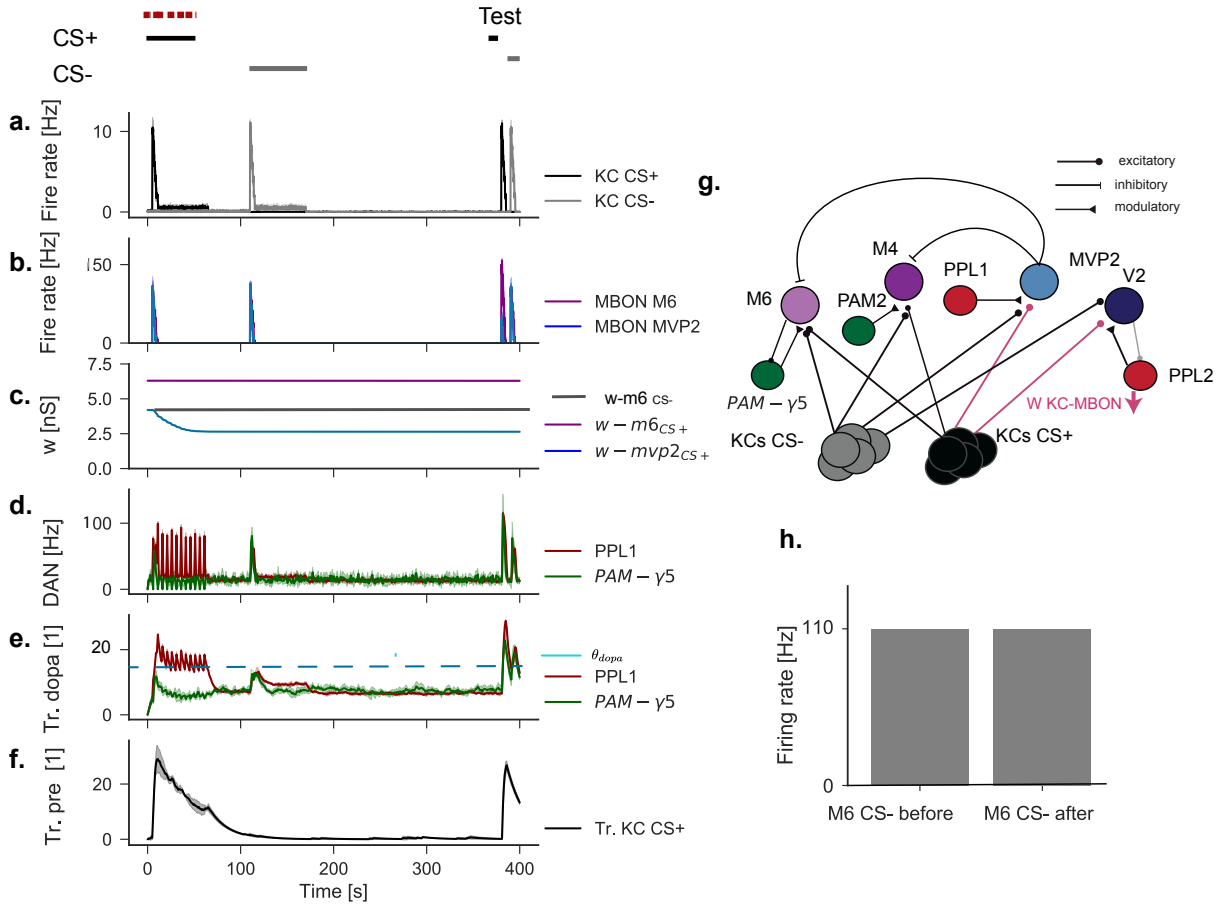


Figure 3.8: Network activity during aversive training. Black bar on the top denotes CS+ exposure duration, while grey bar represents CS- exposure. Red dots represent shock exposure. After training and CS- exposure, CS+ and CS- odours are re-exposed to test MBON responses with plasticity de-activated. **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 pre-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-γ5 DAN during the simulation. **e.** We show PPL1 DAN and PAM-γ5 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. Lines with circle arrow ends represent excitatory connections (KCs to MVP2 or M6 or M4). Lines with vertical line end are inhibitory connections (MVP2 to M6 or M4), while lines with triangle ends are modulatory connections (PPL1 to KC-MVP2). Black arrows show connections that are unchanged after aversive learning while red arrows denote connections that have experienced depression. The colours of the circuit are consistent with the colours in the simulation on the left. **h.** Firing rates of M6 MBONs in response to CS- odour before and after odour exposure (no change).

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 3.8 a). This is qualitatively similar to experimental results (Owald et al., 2015, Felsenberg et al., 2018). After CS- exposure, MBON responses to the CS- remained unchanged (Fig 3.8 b).

We tuned out network so that the changes in MBON MVP2 firing rate match experimental data (Fig 3.9 a). We also tuned our learning rule to obtain a theoretical performance index that matches aversive learning experiments (Felsenberg et al., 2018, Fig 3.9 c). After aversive learning, all CS+ KC-V2 connections decreased in strength (Fig 3.9 b).

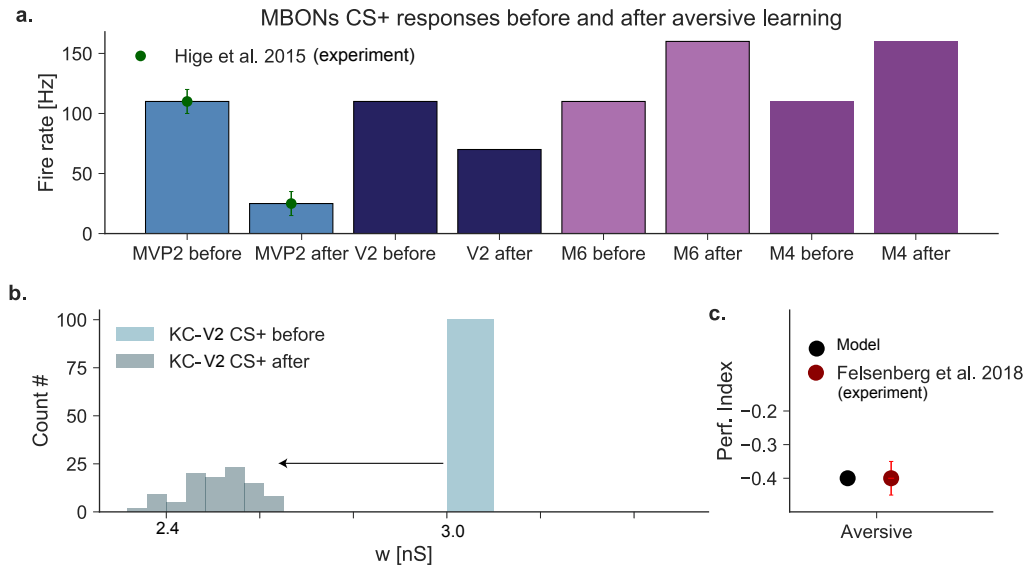


Figure 3.9: MBON firing rates, Performance Index and weights after aversive learning. **a.** We show firing rates for MBON MVP2 (light blue), V2 (dark blue), M6 (light purple), M4 (dark purple) before and after aversive learning simulation. **b.** KC-V2 connection weights for KCs that respond to the CS+ are shown before (light blue) and after (dark blue) aversive learning. **c.** Model performance index (black) and experimental data from Felsenberg et al., 2018 (dark red) is shown after aversive training. Red bar indicates standard error of the mean.

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 is is accordance with experimental results (Beck et al., 2000, Fig 3.10 b).

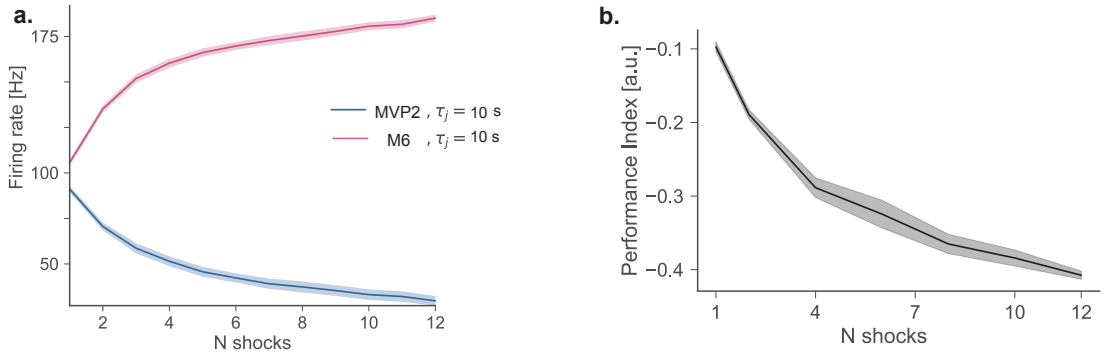


Figure 3.10: Changes in the MBON firing rates as number of shocks during training are increased using learning rule pre-dopa LTD. **a.** MBON MVP2 and MBON M6 firing rates after aversive training with increasing number of shocks, τ_j =10 s. **b.** Model performance index for aversive training with increasing number of shocks. Shaded areas illustrate 1 standard deviation.

3.4.3 Aversive extinction requires connections between opposing valence coding compartments

We now investigate the aversive extinction protocol, which in the real experiment requires re-exposing the same odour multiple times after aversive learning without the expected shock to partially extinct the memory (Felsenberg et al., 2018). We activated CS+ KCs both during aversive learning and extinction (Fig 3.11 a). M6 and MVP2 MBON initially responded with equal firing rates, however after aversive learning MBON M6 firing rate increased and MVP2 After aversive learning the firing rates of MVP2 decreased, while the firing rate of M6 increased (Fig 3.11 b). Similarly, the firing rate of PAM-γ5 increased (Fig 3.11 d). 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).

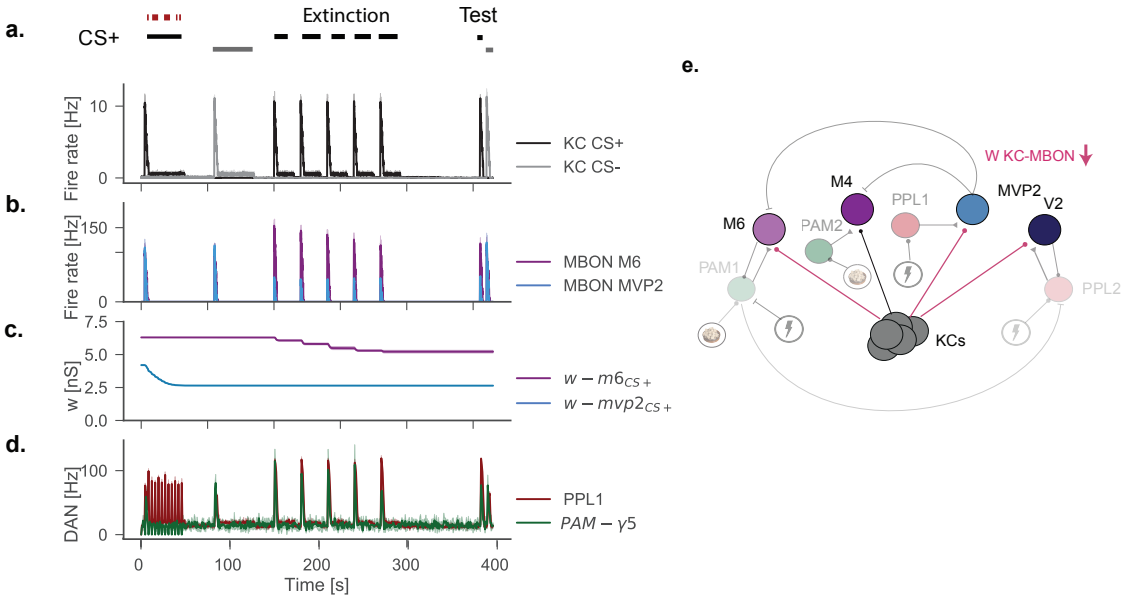


Figure 3.11: Network activity during aversive extinction learning protocol. CS+ odour (black) is exposed for 60s, followed by 45 pause, 60 s of CS- exposure (grey) and 5 CS+ re-exposures without shock (extinction). **a.** Population activity firing rate of KCs responding to CS+ and CS- stimuli during learning and testing. **b.** Firing rates of MBONs MVP2 and M6 during the simulation. **c.** Aversive training using M6 depresses CS+ KC-MVP2 (blue) weights while KC-M6 (purple) remain unchanged. **d.** Firing rates of PPL1 and PAM-γ5 DAN during the simulation. Shaded area indicates 1 standard deviation. **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 a). We tuned plasticity to change the theoretical performance index to match closely the experimental data (Fig 3.12 c). 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).

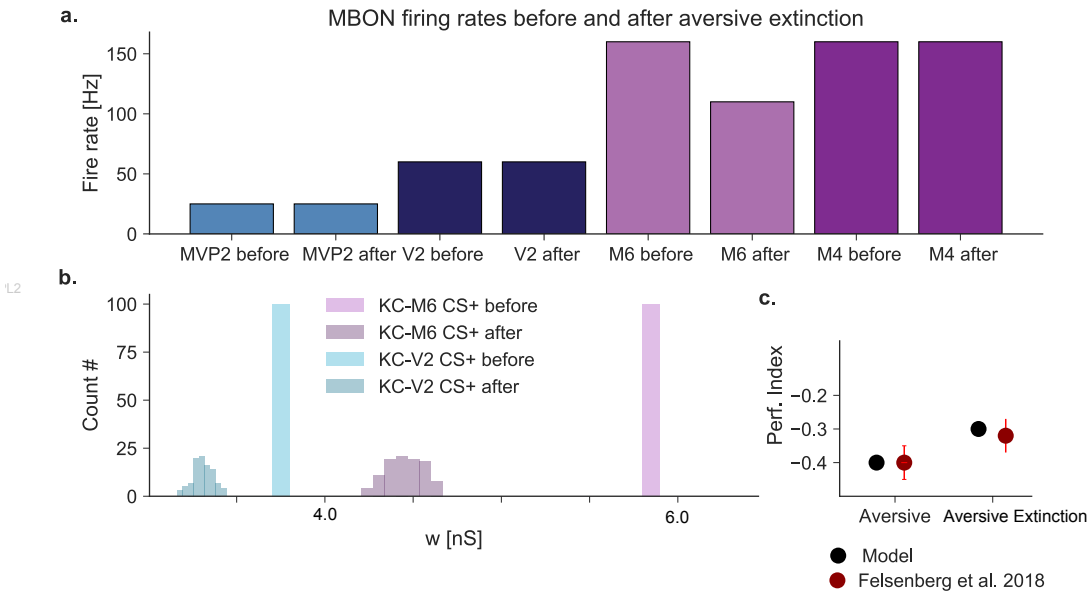


Figure 3.12: MBON firing rates, Performance Index and weights after aversive extinction. **a.** MBON MVP2 (light blue), V2 (dark blue), M6 (light purple), M4 (dark purple) are shown before and after aversive extinction simulation. **b.** KC-M6 connection weights for KCs that respond to CS+ are shown 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. This is in continuation of Fig 3.9 c. Vertical bars indicate standard error of the mean.

To study which connections are required for aversive extinction we simulated the learning protocol with a virtual lesion to the MVP2-M6 connection. Exactly like in Figure 3.11, we activated CS+ KCs both during aversive learning and extinction (Fig 3.13 a). M6 and MVP2 MBON initially responded with equal firing rates. After aversive learning the firing rates of MVP2 decreased (Fig 3.13 b). The firing rate of PAM-$\gamma 5$ remained the same (Fig 3.13 d). After aversive learning, when we re-exposed the odours, the average KC-M6 CS+ remained unchanged as the MVP2-M6 connection was lesioned (Fig 3.13 c).

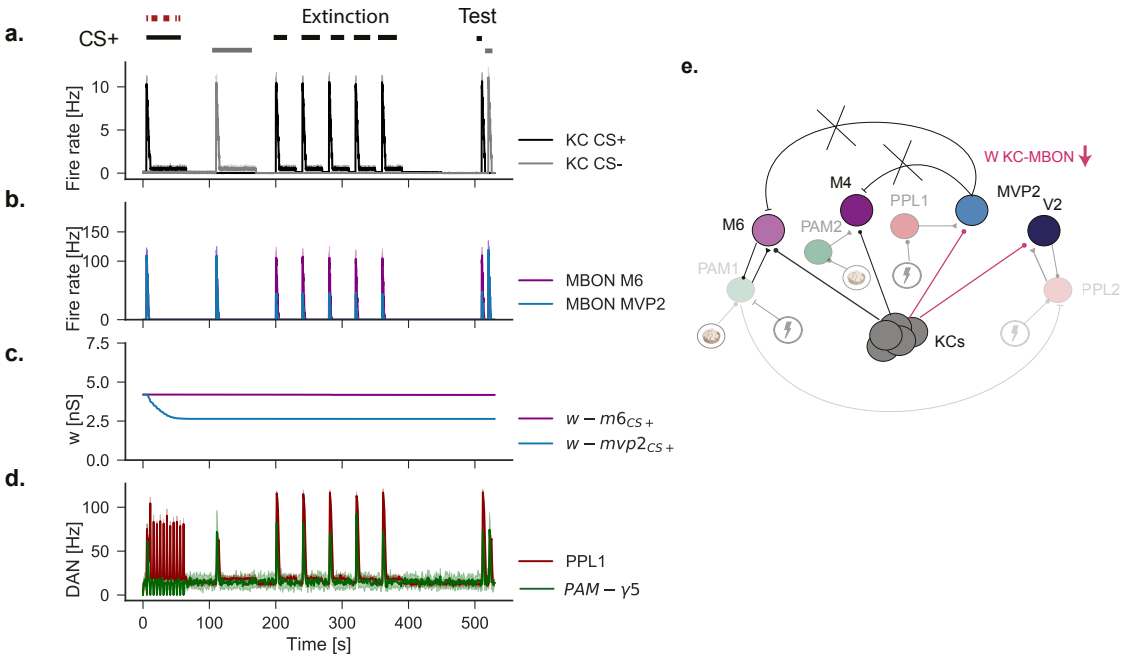


Figure 3.13: Network activity during aversive extinction learning protocol with MVP2-M6 inhibitory connection lesioned. CS+ odour (black) is exposed for 60s, followed by 45 pause, 60 s of CS- exposure (grey) and 5 CS+ re-exposures without shock (extinction). **a.** Population activity firing rate of KCs responding to CS+ and CS- stimuli during learning and testing. **b.** Firing rates of MBONs MVP2 and M6 during the simulation. **c.** Aversive training using M6 depresses CS+ KC-MVP2 (blue) weights while KC-M6 (purple) remain unchanged. **d.** Firing rates of PPL1 and PAM-γ5 DAN during the simulation. Shaded area indicates 1 standard deviation **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. **e.** Circuit diagram of our model like in Fig 3.2 showing which connections are lesioned (MVP2-M6 and MVP2-M4).

We now test the effect of aversive learning and extinction learning on the activity of PAM-γ5 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-γ5 activity in response to CS+, while aversive extinction reduces their responses (Fig 3.14 a, b).

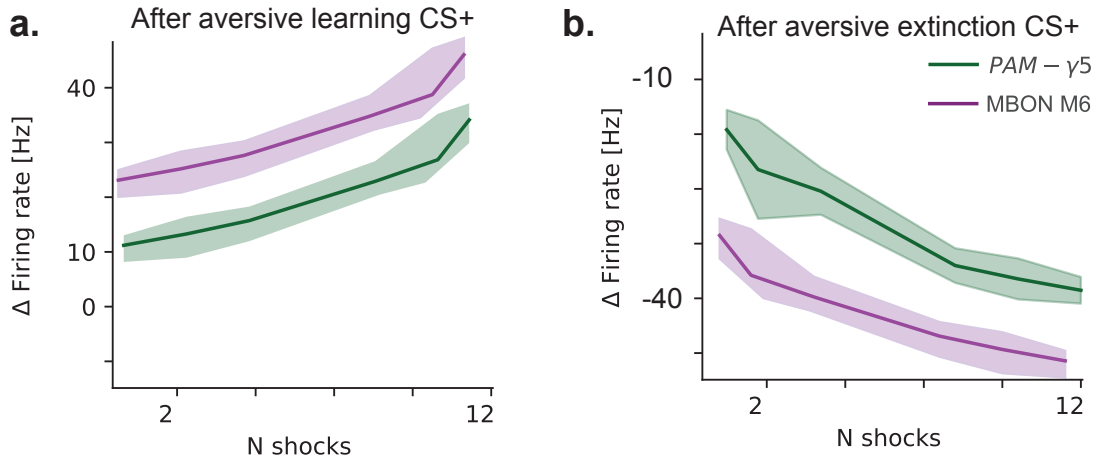


Figure 3.14: Changes in the firing rates of M6 MBON and PAM-γ after aversive learning and extinction. **a.** We show the changes in the responses of MBON M6 and PAM-γ5 neurons during odours exposure after aversive learning and extinction as we increase the number of shocks we used during aversive learning. **b.** We show the responses of MBON M6 and PAM-γ5 after aversive extinction of aversive memories of different number of shocks. Shaded area indicates 1 standard deviation.

In our model MBON-M6 excites PAM-γ5, which can extinction when its firing rate increases. We tested the effect of changing the learning rate in the M6 MBON compartment. The firing rate of the MBON M6 decreased monotonically with the number of extinction trials when $\eta_{m6} = 5 \times 10^{-7}$ (Fig 3.15 a). When we increase the learning rate for KC-M6 connections (η_{m6}) the effect of subsequent extinction trials was reduced (Fig 3.15 a). In the case of a high learning rate $\eta_{m6} \geq 10^{-6}$ only the first extinction trial modified the firing rate of MBON M6 (Fig 3.15 a). Learning rate also determines change in performance index after extinction (Fig 3.15 b).

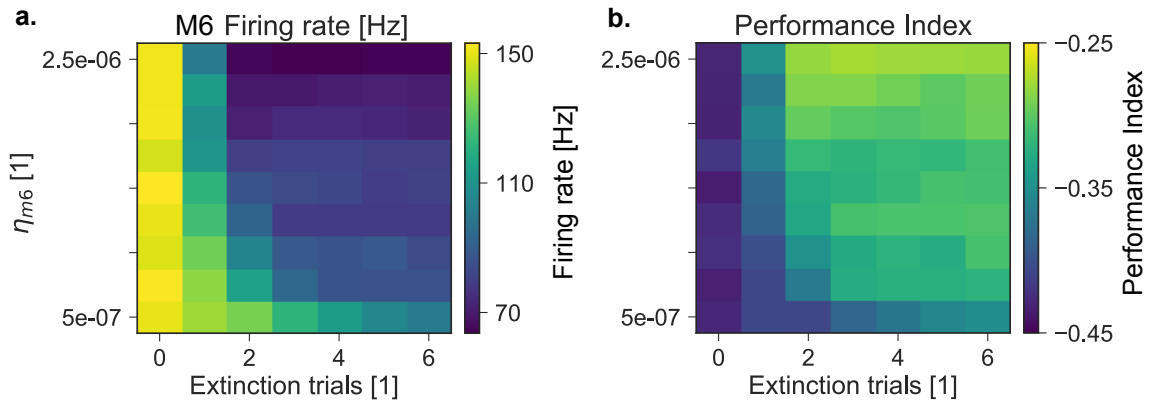


Figure 3.15: Changes in the MBON M6 firing rate as number of extinction trials is increased. **a.** Colormap shows change in the firing rate of MBON-M6 after aversive extinction with increasing number of extinction trials and increasing learning rate ($\eta_{PAM-\gamma5}$) for the KC-MBON-M6 connections **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. 3.8 and we obtained very similar results to using Rule pre-dopa-LTD. KC and MBON activity in the first second of odour exposure create an eligibility trace ($e_{ij}(t)$) (Fig 3.16 e) which modifies the weight when shocks are present (Fig 3.16 c) After aversive learning, MVP2 decreases its response to CS+ and M6 and PAM-γ5 increase their CS+ responses (Fig 3.8 b, d). KC-M6 CS+ average weight decreases during the 5 odour presentations after aversive learning which reproduces the aversive extinction experiment (Fig 3.16 c).

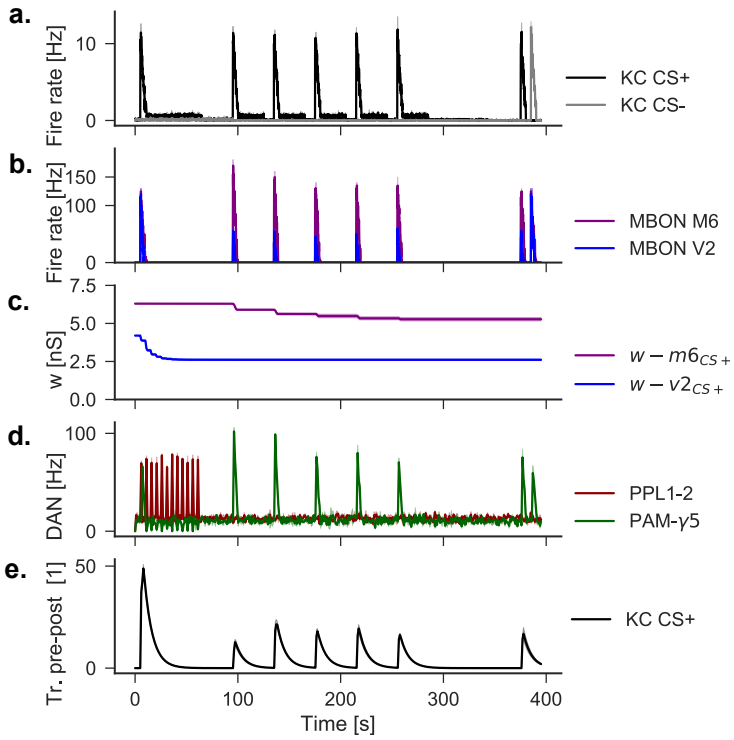


Figure 3.16: Network activity during appetitive extinction simulation with R-STDP learning rule **a.** Shown here are population firing for CS+ responding KCs (black) and CS- responding KCs (grey) **b.** MBON MVP2 (blue) and M6 (purple) firing rates during the simulation. **c.** Average CS+ KC-MVP2 weight (light blue) and average CS+ KC-M6 weight (purple) shown during simulation. KC-MVP2 weight change. **d.** Panel shows the firing rates of PAM-γ5 (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 Pre-dopa-LTD rule can reproduce aversive learning experiment with no post spikes

We simulated aversive learning using both rule pre-dopa-LTD and rule 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. We activated CS+ KCs and PPL1 neurons to simulate pairing an odour with shock (Fig 3.17 a, d).

Our results show that Rule pre-dopa-LTD can modify the firing rate of MBON MVP2 without postsynaptic spikes (Fig 3.17 a) while Rule R-STDP fails to do so (Fig 3.17 b). When MBON activity is blocked, absence of pre-post spike pair results in the

eligibility trace never increasing (Fig 3.17 e, right), and KC-MVP2 weight remains unchanged (Fig 3.17 c). After aversive learning using rule pre-dopa-LTD the CS+ response of M6 increased (Fig 3.17 b).

Blocking MBON will also block excitatory input from MBON-M6 to PAM-γ5. Thus, while aversive learning does not require MBON activity using our learning rule dopa-pre-LTD, aversive extinction requires it.

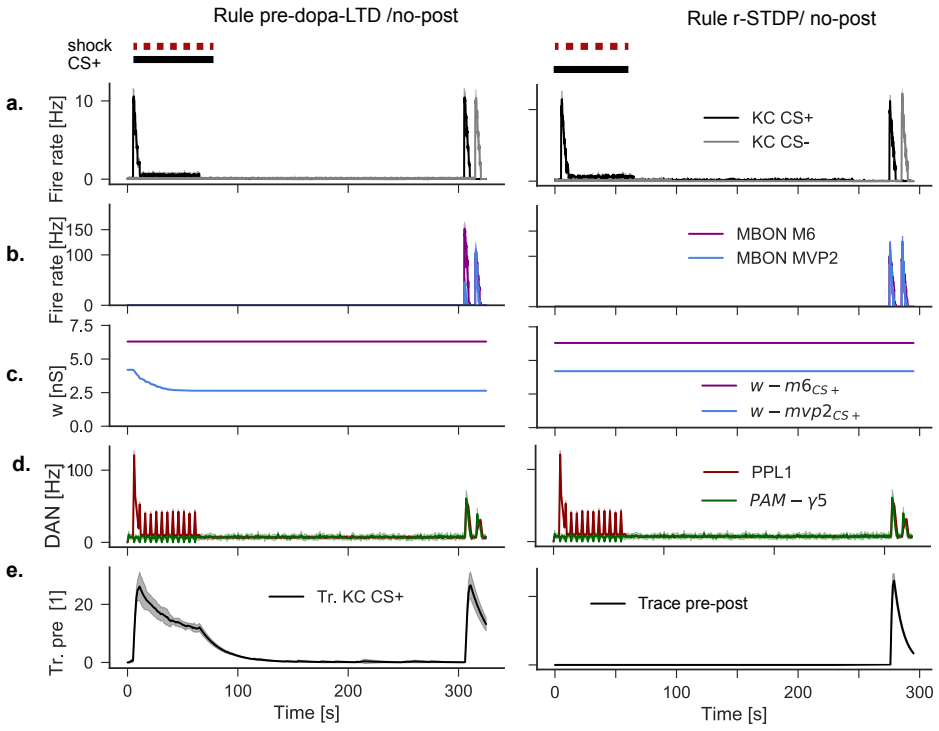


Figure 3.17: Network activity during aversive training with no postsynaptic spikes using Rule 1 (pre-dopa LTD, left) and Rule 2 (R-STDP, right). **a.** KC population firing rates in response to CS+ and CS- **b.** MBON MVP2 and M6 firing rates during the simulation. **c.** KC-MVP2 and KC-M6 CS+ weight change during simulation **d.** Panel shows the firing rates of PAM-γ5 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. Shaded areas indicate one standard deviation.

3.4.6 Aversive re-learning decreases only MBON V2 firing rate

We also simulated aversive re-learning immediately after extinction. We reactivatingivated CS+ KCs and PPL1 and PPL1-2 DAns after aversive extinction to simulate re-learning (Fig 3.18 a, d).

During aversive re-learning CS+ KC-V2 weights decreased further than during aversive learning while KC-MVP2 weights remained unchanged (Fig 3.18 c). During testing we observed that the firing rate of V2 MBON decreased in response to CS+ (Fig 3.18 b).

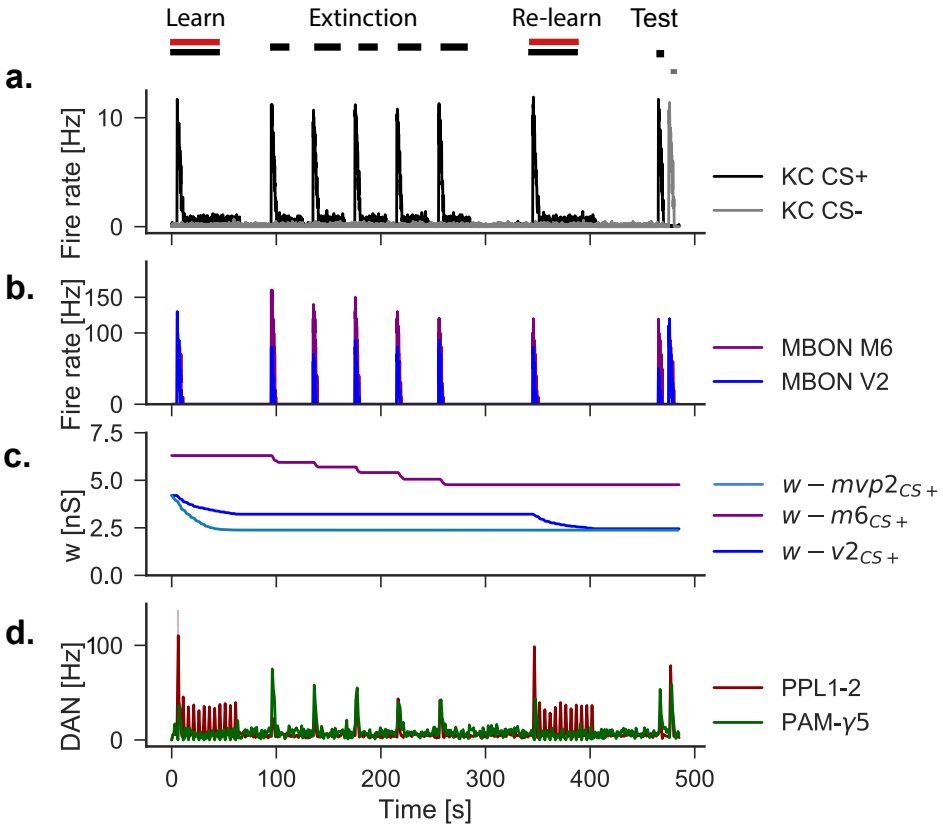


Figure 3.18: Network activity during aversive re-learning. **a.** Population activity firing rate of KCs responding to CS+ and CS- stimuli during learning and testing. **b.** Firing rates of MBONs V2 and M6 during the simulation. **c.** We show the average CS+ KC-V2 (blue), CS+ KC-MVP2 (light blue) and KC-M6 (purple) weights during the simulations. **d.** Firing rates of PPL1 (red) and PAM-γ5 (green) DAns during the simulation. Shaded areas indicate 1 standard deviation.

After aversive re-learning only the V2 MBON decreased its firing rate compared to before re-learning (Fig 3.19 a). We obtained a similar Performance Index to the result observed experimentally when we converted CS+ firing rates into a performance index using Eq. 3.13 (Fig 5.7 b).

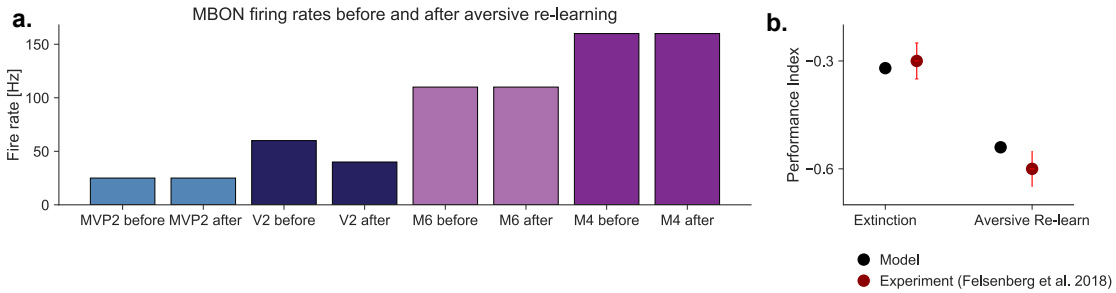


Figure 3.19: MBON firing rates, Performance Index and weights after aversive re-learn. **a.** MBON MVP2 (light blue), V2 (dark blue), M6 (light purple), M4 (dark purple) are shown before and after aversive extinction simulation.**b.** Model performance index and experimental data from Felsenberg et al., 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. Vertical bars indicate standard error of the mean.

3.4.7 Second order conditioning forms two opposing memories

In our model MVP2 inhibits PPL1 DAn, while M6 excites PAM-γ5 DAn. After aversive learning both DANs increase their responses to the aversive odour (CS+): PPL1 DAN (decreased inhibition from MVP2 MBON) and PAM-γ5 (increased excitation from M6).

To test the possibility of second-order conditioning we simulated exposures to CS+ and CS2 after either CS+ aversive training or aversive extinction (Fig 3.20 a). MBONs were activated by both CS+ and CS2 (Fig 3.20 b). Second order conditioning after aversive learning led to the decrease of CS2 KC-MVP2 and KC-M6 weights (FIg 3.20 c). When we simulated second order conditioning after aversive extinction only CS2 KC-MVP2 weights were modified.

After aversive learning, re-exposure to CS+ triggers an increased response in both PAM-γ5 and PPL1 DANs (Fig 3.20 d left). Second order conditioning after CS+ aversive learning leads to the creation of two opposing memories in the M6 and MVP2 compartments. In contrast, second-order conditioning after CS+ aversive extinction only modifies CS2 KC-MVP2 weights (Fig 3.20 c, right). This is because CS+ aversive extinction reduces PAM-γ5 CS+ response.

Unfortunately, neither result was close to experimental values, when we converted CS2 firing rates into the performance index. This is because second order conditioning only modifies the firing rates of MVP2 and M6, without affecting V2 and M4 MBONs.

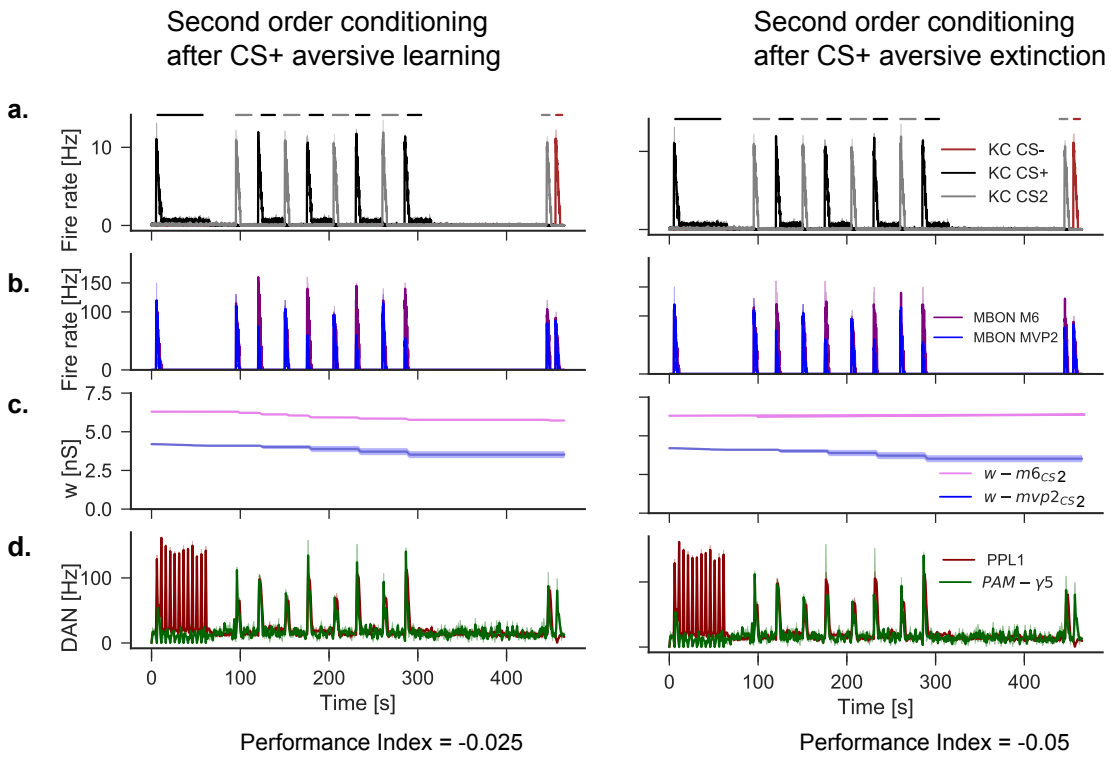


Figure 3.20: Network activity during second order conditioning after CS+ aversive learning (left) and after CS+ aversive extinction (right). KC population firing rates in response to CS+ and CS- **b.** MBON MVP2 and M6 firing rates during the simulation. **c.** KC-MVP2 and KC-M6 CS+ weight change during simulation **d.** Panel shows the firing rates of PAM-γ5 and PPL1 DAns during the simulation. Shaded areas indicate 1 standard deviation.

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 3.6). 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 3.5).

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 in firing rate between approach and retreat biasing MBON

determines the behaviour of the 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) (Yarali et al., 2008). 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 is similar to experiments that showed KCs respond transiently to odours (Murthy et al., 2008). Our presynaptic trace, based on experiments, shows that flies maintain a trace of the odour which allows the fly to learn even if the shock arrives after odour exposure has ended (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 3.17) (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 3.9)(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). 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 disinhibited by the MVP2 MBON. This mechanism cannot work multiple times as the firing rate

of the MVP2 MBON in response to CS+ will eventually reach zero spikes. In chapter 5 we investigate a mechanism for potentiation that increases the flexibility of our computational model to re-learn and re-extinct more than once. Interestingly, after aversive learning both PAM-γ5 and PPL1 enhance their respond to the CS+ odour.

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, postsynaptic spikes may not be explicitly required in the learning rule they appear to be playing an important rule 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

Olfactory appetitive learning and extinction in the mushroom body

4.1 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 disinhibition. 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-γ5 DAns to inhibit aversive PPL1-2 DAns. To test whether such a connection exists, we predict that blocking PAM-γ5 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-γ5

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. So 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 (Felsenberg et al., 2017). Furthermore, appetitive memory extinction appears 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. We hypothesize a link between PAM-γ5 DANs and PPL1-2 DAn. Aversive DANs have been shown to be inhibited by sugar (Cohn et al., 2015), but it is unknown if this contributes to extinction. Extinction requires the activity of DANs that code for the opposite reinforcer to the US after learning. In the previous chapter we have shown that aversive learning can increase the activity of PAM-γ5 through MVP2 mediated disinhibition. In this chapter we explore whether PAM-γ5 DANs could also help orchestrate appetitive extinction by decreasing their response to the appetitive odour after learning and disinhibiting aversive DANs.

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. Rewards are simulated as transient increases in PAM DAns activity for two seconds, corresponding to calciumg imaging responses of $PAM-\gamma5$ DAns during sucrose ingestion (Waddell lab, personal correspondence).

4.4 Results

4.4.1 Appetitive learning and extinction

In this section we investigate appetitive learning and extinction in our model. During reward learning we paired CS+ KC activation with 1 second of PAM-γ5 and PAM-β'2 activation (Fig 4.1 a, d). After reward learning CS+ KC-M6 and KC-M4 weights decreased (Fig 4.3 c, e) and the firing rates of MBON M4 and MBON M6 decreased in response to the CS+ odour decreased as well (Fig 4.3 b).

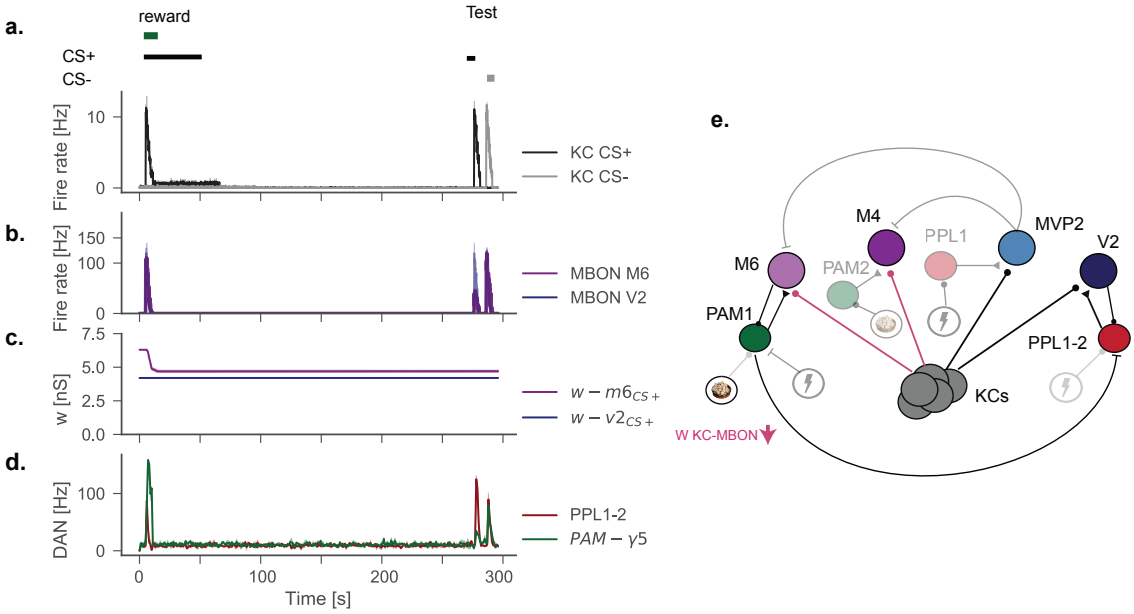


Figure 4.1: Network activity during appetitive learning. **a.** Population activity firing rate of KCs responding to CS+ and CS- stimuli during learning and testing. **b.** Firing rates of MBONs MVP2 and M6 during the simulation. **c.** We show average CS+ KC-V2 (blue) weights and KC-M6 (purple) during the simulations. **d.** Firing rates of PPL1-2 and PAM-γ5 DAns during the simulation. **e.** We show the circuit representation and connections in response to CS+ after appetitive learning. Black lines from KC to MBONs are unchanged connections while red lines are connections that underwent LTD. Shaded areas indicate 1 standard deviation.

After training, M6 and M4 responses to CS+ odour were reduced (4.2 a). All of the KC-M6 synapses responding to CS+ decreased after appetitive learning (Fig 4.2 b). We also tuned our learning rule to obtain a theoretical performance index that matches appetitive learning experiments (Felsenberg et al., 2017, Fig 3.9 c).

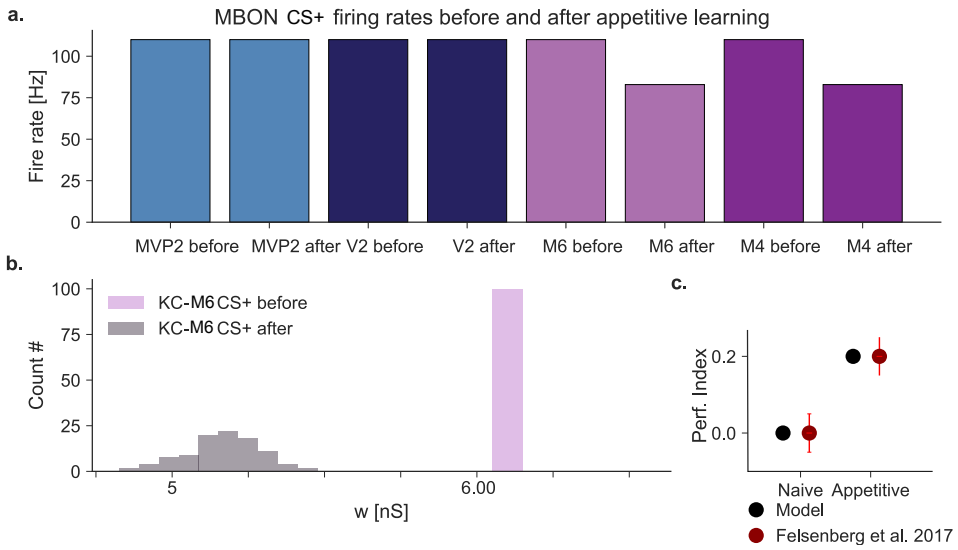


Figure 4.2: MBON firing rates, Performance Index and weights after appetitive learning. **a.** MBON MVP2 (light blue), V2 (dark blue), M6 (light purple), M4 (dark purple) are shown before and after appetitive learning simulation. **b.** KC-M6 connection weights for KCs that respond to CS+ are shown before (light purple) and after (dark purple) appetitive learning. **c.** Model performance index and experimental performance index based on data from Felsenberg et al., 2017 (dark red) is shown after appetitive training. Vertical bars indicate standard error of the mean.

After appetitive learning we simulated extinction and re-exposed the CS+ odour five times (Fig 4.3 a). Reward learning led to the decrease of PAM-γ5 firing rate in response to CS+ (Fig 4.3 d green). After reward learning we note that when the CS+ odour was re-exposed both the M6 and PAM-γ5 responses decreased, while PPL1-2 firing rate increased (Fig 4.3 b, d). Each time the CS+ odour was re-exposed CS+ KC-V2 weights decreased (Fig 4.3 c).

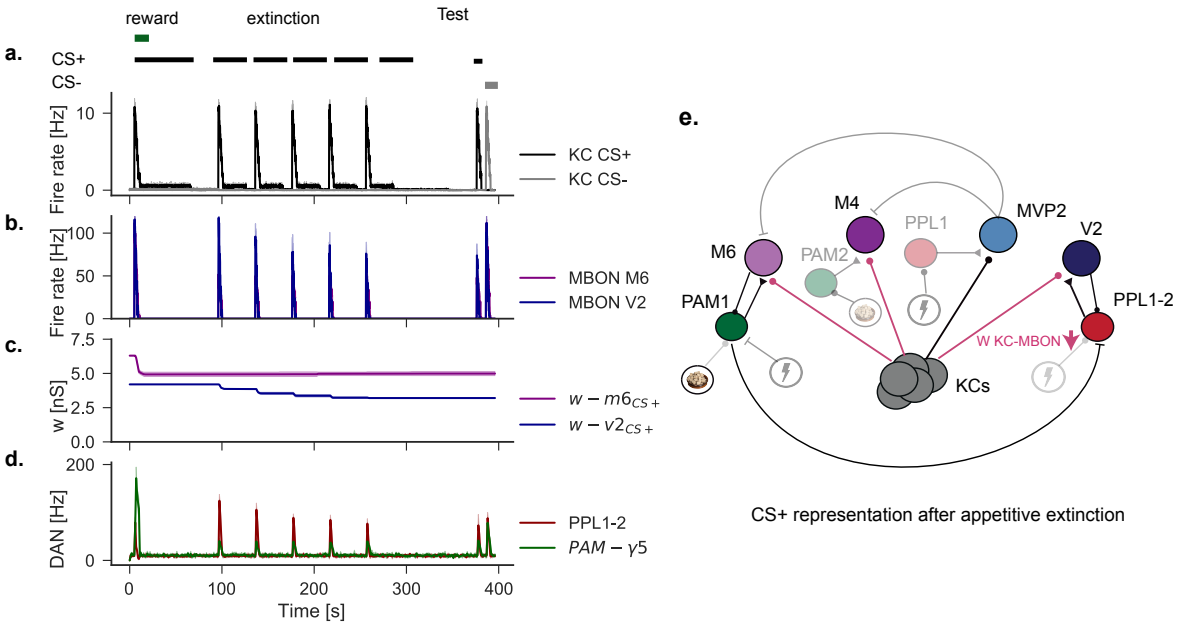


Figure 4.3: Network activity during appetitive learning and extinction. **a.** Population activity firing rate of KCs responding to CS+ and CS- stimuli during learning and testing. **b.** Firing rates of MBONs MVP2 and M6 during the simulation. **c.** We show average CS+ KC-V2 (blue) weights and KC-M6 (purple) during the simulations . **d.** Firing rates of PPL1-2 and PAM-γ5 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. Shaded areas indicate 1 standard deviation.

Appetitive extinction decreased the firing rates of V2 MBON to approximately the same level as M4 and M6 (Fig 4.4 a). Both the connection representing CS+ between KC-V2 and KC-M6 an M4 are reduced (Fig 4.4 b). Our results for appetitive extinction match experimental results (Fig 4.4 c).

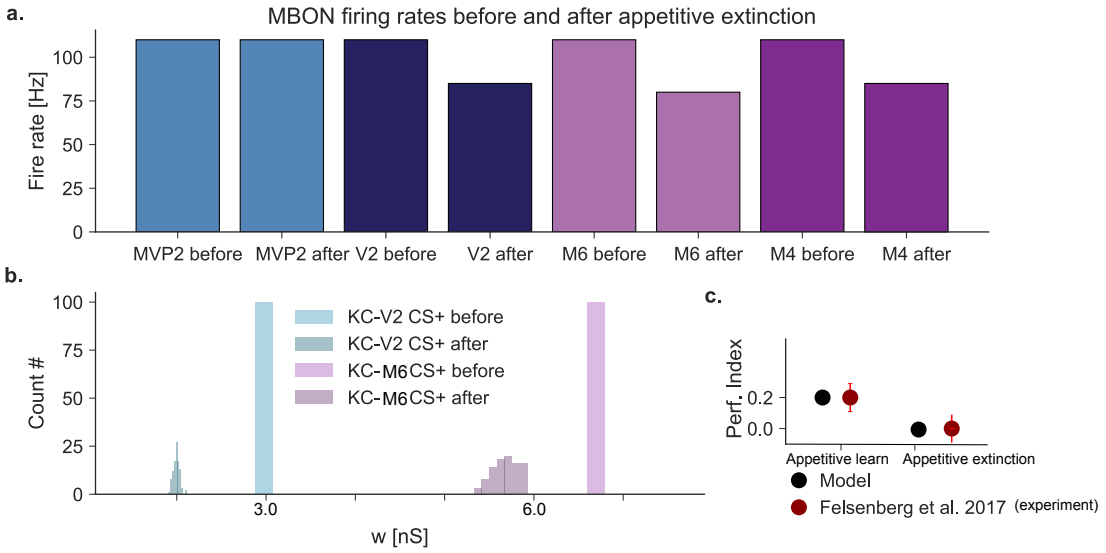


Figure 4.4: MBON firing rates, Performance Index and weights after appetitive extinction. **a.** MBON MVP2i (light blue), V2 (dark blue), M6 (light purple), M4 (dark purple) are shown before and after aversive extinction simulation. **b.** KC-M6 connection weights for KCs that respond to CS+ are shown before (light purple) and after (dark purple) aversive extinction. **c.** Model performance index and experimental data from Felsenberg et al., 2017 (dark red) is shown after appetitive training and after appetitive extinction. Model (black) represents simulation. Vertical bars indicate standard error of the mean.

4.4.2 Prediction: Blocking PAM-γ5Ans during odour exposure leads to an aversive memory

To test whether appetitive extinction in our model is caused by a reduction in odour response of PAM-γ5 DAN which inhibits PPL1-2, we exposed an odour while blocking PAM-γ5 (Fig 4.5 a, d). Blocking PAM-γ5 led to an increase of PPL1-2 DAn activity (Fig 4.5 d) and the decrease in the CS+ KC-V2 weights and the firing rate of MBON V2 in response to CS+, similar to the pattern observed during appetitive extinction (Fig 4.5 b, c). We propose that our simulation result can be tested in an experiment by blocking PAM-γ5 DAns.

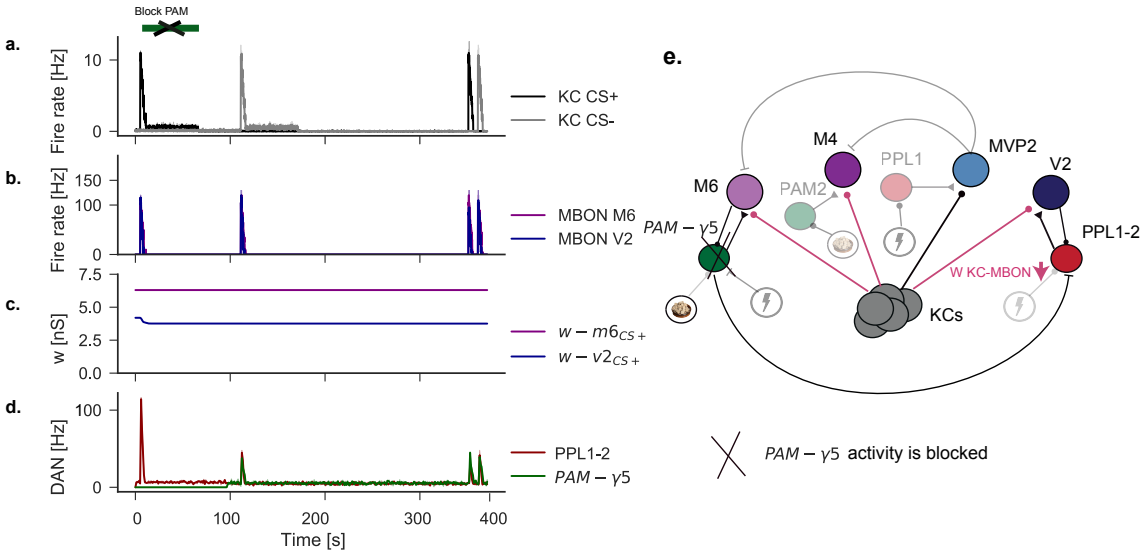


Figure 4.5: Suppressing PAM-γ5 during odour exposure. **a.** Population activity firing rate of KCs responding to CS+ and CS- stimuli during learning and testing. **b.** Firing rates of MBONs MVP2 and M6 during the simulation. **c.** We show average CS+ KC-V2 (blue) weights and KC-M6 (purple) during the simulations. **d.** Firing rates of PPL1-2 and PAM-γ5 DANs during the simulation. Performance index score in this model was -0.1 for the odour exposed while blocking PAM-γ5 DAN. **e.** Circuit diagram of the circuit we simulated (same as in Fig 3.2) with the activity of PAM-γ5 blocked.

We also changed 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 or partially (Performance Index = 0, Fig 4.6).

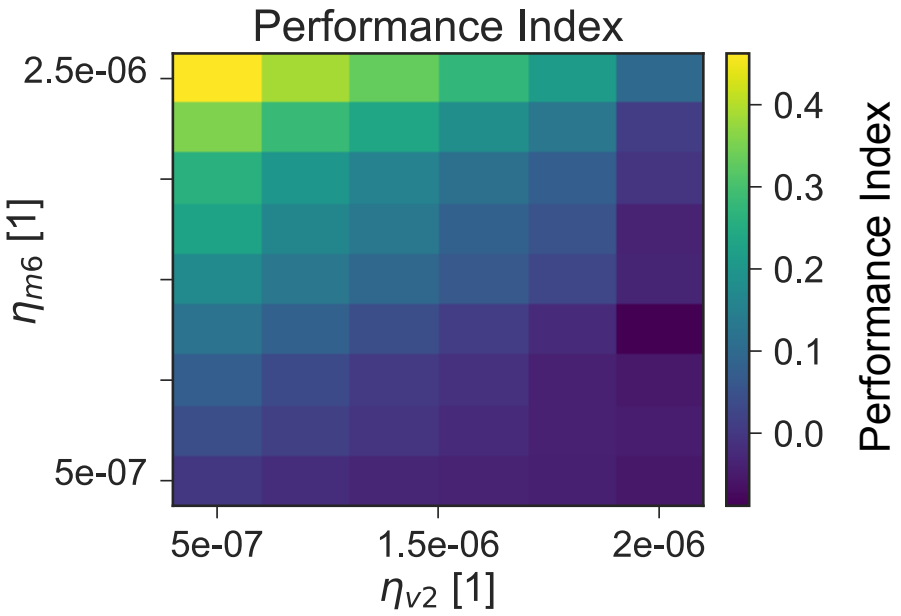


Figure 4.6: Appetitive extinction with different learning rates. We show a color map of the resulting performance index after appetitive extinction simulated with different learning rates (x-axis denotes η_{v2} learning rate while y-axis denotes η_{M6} learning rate.)

4.4.3 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 4.7 b). Similarly, re-exposing the odour after appetitive training with reward led toe a performance index higher than initial appetitive learning (Fig 4.7 b).

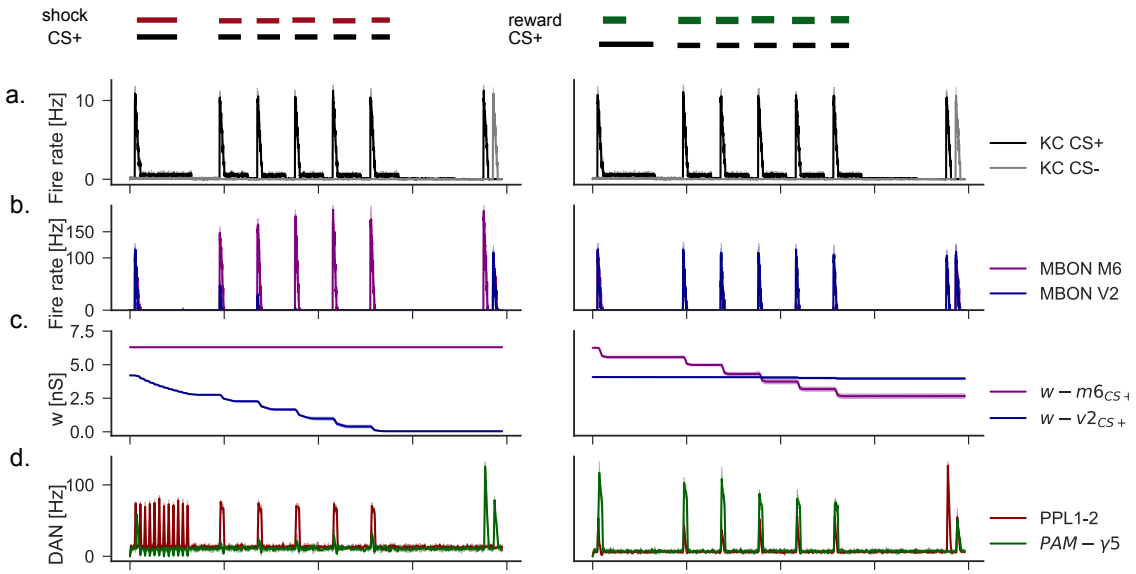


Figure 4.7: Network activity during extinction protocol in the presence of reinforcement. Left: CS+ odour (black) is exposed for 60s, followed by 45 pause and 5 re-exposures **with** shock. Right: CS+ odour (black) is exposed for 60, s followed by 44 pause and 2 re-exposures **with** reward. The dynamics of neurons and traces during the aversive extinction protocol are shown: **a.** KC activity for both CS+ (black) and CS-(yellow) odours; **b** MVP2 (blue) and M6 (purple) activity; **c** KC-MVP2 and KC-M6 (purple) weights and PAM-γ5 PPl1, and KC traces are shown for completeness. **d** PAM-γ5 (green) and PPL1 (red) KC-MVP2 and KC-M6 (purple) weights and PAM-γ5 PPl1, and KC traces are shown for completeness. Shaded areas indicate 1 standard deviation.

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-γ5 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 $PAM - \gamma 5$ to a PPL1 DAN, after appetitive learning the activities of the two DANs will change in opposite direction. $PAM - \gamma 5$ will decrease its CS+ response while PPL1-2 increases. Using our pre-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

propose that blocking appetitive DAns while exposing an odour will form an aversive memory.

Chapter 5

Investigating the role of potentiation in the mushroom body

5.1 Abstract

In this chapter we investigate the role of potentiation at the site of KC-MBON synapses in the mushroom body. In chapter 3 we examined two learning rules that can only weaken KC-MBON synapses. Synaptic depression was triggered when dopamine activity levels were above baseline levels of activity during odour presentation. We showed that our model could reproduce learning and extinction experiments but could not perform re-extinction. Here, we introduce a new learning rule, pre-dopa-LTD-LTP that extends the previous learning rule from chapter 3 that can modify KC-MBON synapses bi-directionally. Potentiation occurs when an odour is present and dopamine firing rate is below baseline activity. Weakening of the synapse is still triggered by dopamine levels above baseline activity.

We use our new learning rule to reproduce relief learning experiments that have shown that PPL1 activation in the $\gamma 1$ compartment before odour activation leads to the formation of a positive memory. We modify the PPL1 DAn we modelled in chapter 3 by adding spike-frequency adaptation, which reduces the activity of PPL1 below baseline level after shock activation of PPL1. This leads to potentiation of KC-MVP2 synapses using our new learning rule and a positive memory is formed when shock precedes odour.

Our results show that potentiation of KC-MBON synapses is required for both re-extinction and relief learning. We propose an experiment that requires imaging the M6 MBON after aversive re-learning and we predict that the response to the trained

odour will be enhanced.

We test our new learning rule in the four-layer model we studied in chapter 2, and show that we can reproduce all of the aversive learning paradigms in the full model that we examined so far in the simplified two-layer model. We have created the most complete model of the fruit fly's olfactory circuit that can reproduce the stages of transformation of odour information from ORN firing rates into a sparse representation in the mushroom body and learn the valence of the odour.

5.2 Introduction

The fruit fly navigates an environment that is dynamic with odours that could change value multiple times. In chapters 3 and 4 we showed that olfactory learning and extinction experiments can be reproduced by learning rules that can only depress synapses. However, after re-learning, re-extinction is no longer possible since the firing rates of the neurons required for extinction are too low. This suggests synaptic depression is not sufficient to explain plasticity at the site of KC-MBON synapses.

Multiple experimental studies have shown that synapses can both increase and decrease their efficiency depending on the activity of presynaptic and postsynaptic neurons, and of neuromodulators such as dopamine (Bliss et al., 1973, Bi and Poo, 1998, Cassenaer and Laurent, 2012, Gerstner et al., 2018). In the fruit, fly high dopamine activity paired with an odour has been shown to trigger a decrease in the firing rate of MBONs (Hige et al., 2015a).

Additionally, recent experimental evidence has suggested that potentiation occurs in the mushroom body too. One such experiment has shown that when DANs in the $\gamma 4$ compartment were activated more than 45 seconds before KC activation, the $\gamma 4$ MBON response to KC activation was increased (Cohn et al., 2015). The effect was fully reversible once KC and DAN activation were paired together. This experiment suggests that DAn activation before odour presentation either potentiates all KC-MBON synapses in a single compartment or that DAn activation leaves a trace lasting tens of seconds that triggers odour specific potentiation.

Evidence for odour specific potentiation has been suggested by the relief learning experiment. Relief learning represents the ability to form positive memories to stimuli that follow a traumatic experience. Experiments in the fly have shown that when a fly is exposed to an odour after getting shocked, it will tend to approach that odour (Yarali et al., 2008). In the relief learning experient, the outcome of changes the inter-stimulus interval on learning was studied (Fig 5.1). Inter-stimulus interval measures

the relative different between shock and odour onset. To form a relief memory shock must precede odour and there must be no overlap between odour and shock (ISI > 60 seconds). In a more recent experiment, photostimulation of only the PPL1 DAn in the $\gamma 1$ compartment was sufficient to create a positive, relief memory to the odour that followed PPL1 activation (König et al., 2018).

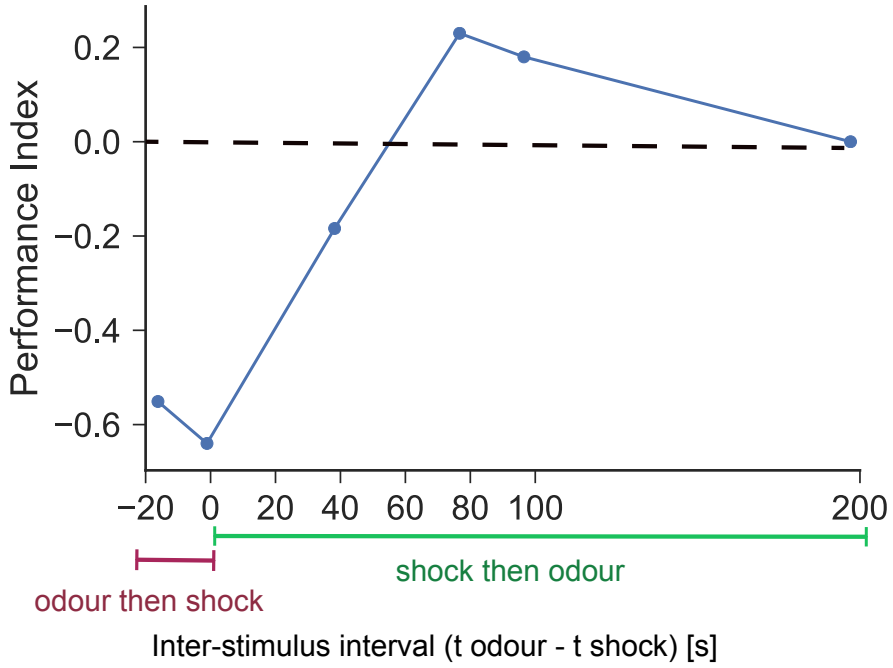


Figure 5.1: Result of pairing an odour with a shock with different inter-stimulus intervals. Curve shows performance index for different inter-stimulus intervals. Black bar denotes an odour exposed for 60 seconds, while dashed line denotes shock. The inter-stimulus interbal (ISI) calculates the relative timing difference between odour onset and the timing of the first shock (first second of exposure). Dashed line denotes neutral performance index.

One possible interpretation of this experiment is that PPL1 activation before odour presentation potentiates KC-MVP2 synapses that are activated by the odour.

In a theoretical study of relief learning, the authors showed that by extending the time window of STDP to tens of seconds they could explain the relief learning experiment (Drew and Abbott, 2006). While the model requires postsynaptic spikes, which appear to be dispensable in the fruit fly, it does highlight the need to have a mechanism for both depression and potentiation to reproduce relief learning.

In our model, we propose an alternative explanation to explain relief learning. After shock activation of PPL1, spike-frequency adaptation decreases the neuron's activity below baseline level until the odour is exposed. Finally, when the odour is

exposed, KC activation paired with DAn activity below baseline, triggers potentiation of KC-MBON synapses. A central requirement of our model is spike-frequency adaptation that can last tens of seconds. Adaptation is a prominent feature of neural dynamics. A previous theoretical study that modelled pyramidal neurons in the neocortex found that to reproduce firing patterns observed experimentally, the model needed adaptation that lasts 20 seconds and decays over multiple time scales according to a power law (Pozzorini et al., 2013). Furthermore, spike-frequency adaptation has also been discovered in dopaminergic neurons in the substantia nigra (Vandecasteele et al., 2011).

If high levels of dopamine trigger synaptic depression, it appears reasonable to assume that very low levels could trigger potentiation. When we implement this hypothesis into our novel learning rule (pre-dopa-LTD-LTP), we find that we can reproduce both relief learning and re-extinction experiments.

5.2.1 Experimental data

In the fruit fly it has been established that odour exposure paired with activation of dopaminergic neurons in the same compartment leads to depression of the KC-MBON synapse (Hige et al., 2015a). In this chapter we are exploring whether potentiation occurs when dopamine activity during odour exposure is below baseline level. A novel experiment has shown that shocks exposure inhibits reward coding PAM dopaminergic neurons, while sugar ingestion inhibits the activity of punishment coding PPL1 dopaminergic neurons (Cohn et al., 2015). In chapter 3 we have shown that a rule that depresses synapses can perform learning, extinction and re-learning as well but not re-extinction due to firing rates always decreasing. Here we test a learning rule in which potentiates KC-MBON connections when dopamine is below baseline level. In our model this occurs during each learning trial as shock or sugar activates some DANs and inhibits opposing coding dopaminergic neurons.

However, we don't allow MBONs to potentiate during initial learning. MBON M6 was shown to increase its CS+ response after aversive learning (Fig 5.2 c, Felsenberg et al., 2018). However, the M6 CS+ responses were measured both at the level of dendrites, where there is no overlap with inhibitory connections from MVP2 MBON and at the level of axons. An increase in response was only observed at the level of the axon (Fig 5.2 b, c), which suggests that increased M6 activity is due to disinhibition from diminished MVP2 activity in response to CS+ (Fig 5.2 b, c).

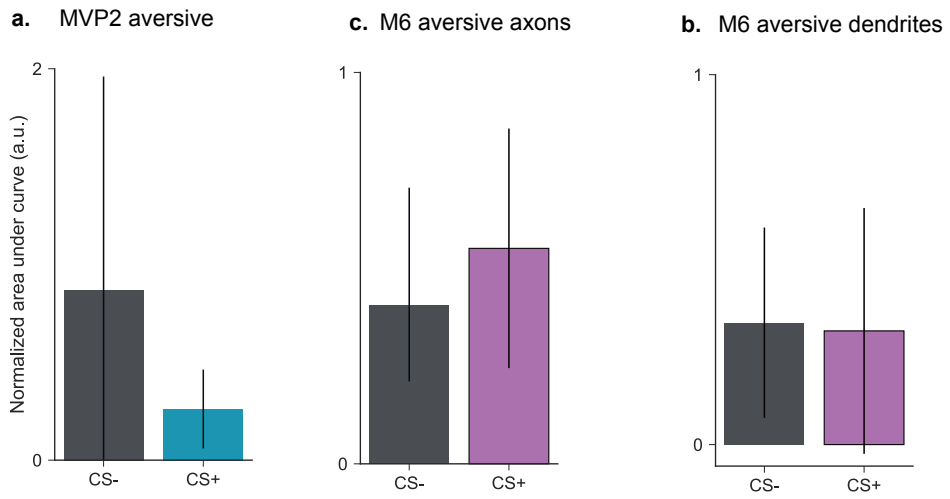


Figure 5.2: GCamp imaging of MBON MVP2 and M6 responding to 5 seconds of odour exposure. **a.** MVP2 MBON CS+ and CS- responses after aversive learning. **b.** M6 MBON CS+ and CS- responses after aversive learning, measured in the axons of M6 **c.** M6 MBONMVP2 CS+ and CS- responses after aversive learning. CS+ and CS- responses after aersive learning measured in the dendrites where there is no overlap with MVP2 inhibitory connections. Vertical bars indicate 1 standard deviation.

5.3 Methods

In this chapter we use the two-layer spiking neural network we introduced in chapter 3 to investigate the role of potentiation in re-learning and re-extinction. We introduce a new learning rule (pre-dopa-LTD-LTP) that potentiates KC-MBON weights when the activity of DAns is below baseline odour induced level.

We also use the model we developed in chapter 2 to show that our new learning rule works in a more complete model of the olfactory circuit.

5.3.1 Neuron model

We used the AdEx model as in the previous chapters. However, we changed the time constant for adaptation for the PPL1 DAn, which can be seen in Table 5.1.

Parameter name	Symbol	Value
Adaptation time constant	τ_w	25 s
Adaptation strength	b	0.1 nA

Table 5.1: PPL1 parameters

5.3.2 Rule pre-dopa-LTP-LTD

Here we introduce a new learning rule (pre-dopa-LTD-LTP). Based on the available experiments we discussed in the introduction, KC-MBON in M6 and V2 compartments cannot potentiate during learning as the weights are set to a maximum level. This is based on evidence that increased M6 activity after learning is due to disinhibition (Felsenberg et al., 2018). Only KC-MVP2 weights are allowed to potentiate, to reproduce relief learning experiment. When an odour is present and dopamine is below a threshold of activity KC-MBON synapses will potentiate (Fig 5.3 a). Conversely, when dopamine is above a threshold, KC-MBON synapses will depress (Fig 5.3 b). Our learning rule has two thresholds one for LTD and one for LTP (Fig 5.3 c).

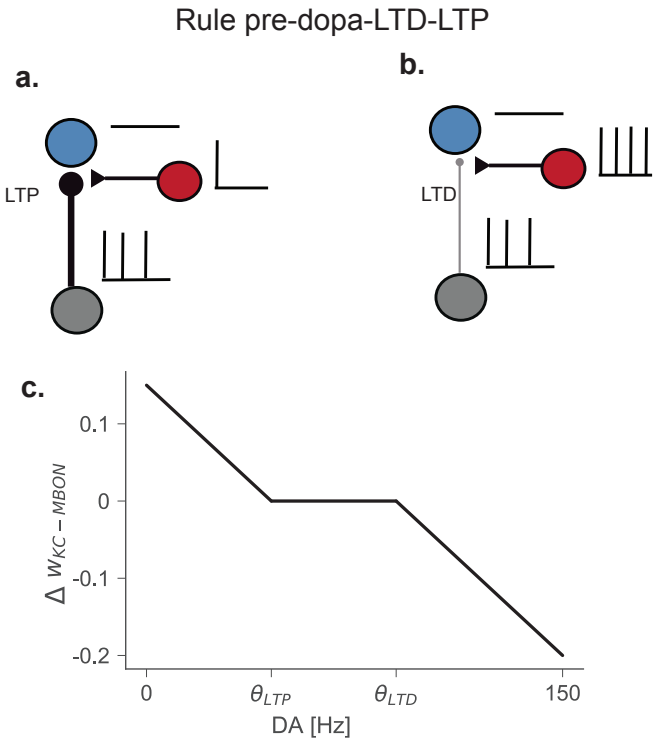


Figure 5.3: Learning rule pre-dopa-LTD-LTP **a.** Potentiation (LTP) requires presynaptic activity and dopamine activity below a threshold. Coincident presynaptic and low DAn activity rates trigger potentiation of the the synapse. **b.** Conversely, synaptic weakening (LTD) requires presynaptic activity and dopamine activity above a threshold. **c.** Shown is change in KC-MBON weight ($w_{KC-MBON}$) as a function of dopamine activity (weight change only occurs when there is both presynaptic activity and dopaminergic activity is either above θ_{LTD} or below θ_{LTP} threshold).

The temporal evolution of the synaptic strength w_{ij} is described by:

$$\frac{d}{dt}w_{ij}(t) = e_j(t)(A_+(\theta_{LTP} - z_{dopa}(t)\mathcal{H}(\theta_{LTP} - z_{dopa}(t)) - A_- S_{dopa}\mathcal{H}(z_{dopa}(t) - \theta_{LTD})), \tag{5.1}$$

where A_- and A_+ are positive parameter that determines the maximum depression and potentiation a connection can undergo respectively. In these simulations A_- is the same value as in our previous learning rule and A_+ is twice A_-. Synaptic weight change w_{ij} requires the presynaptic tag $e_j(t)$ to be above zero and the dopamine tag $z_{dopa}(t)$ to be either above θ_{LTD} or below θ_{LTP}. The presynaptic tag $e_j(t)$ is updated each time the presynaptic neuron j fires as spike as follows:

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

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

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

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, $S_{dopa}(t)$ is the spike train of the dopaminergic neuron: The dopamine concentration, $z_{dopa}(t)$ which gates synaptic change represents a global eligibility trace that increases each time the dopamine neuron fires a spike and is implemented in the same way as in our first learning rule pre-dopa-LTD (Eq 3.7).

Fig 5.3c illustrates the change in weight w_{ij} for different DAn firing rates when there is constant presynaptic activity.

So far, we have studied learning in a simplified model of the olfactory circuit that represented KCs as inhomogeneous Poisson process. In chapter 2 we presented a model that can use experimental recordings of ORN firing rates (Hallem and Carlson, 2006 to show that we can create a sparse representation of an odour in the mushroom body.

5.4 Results

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

As a first test to our novel learning rule, we show that we can reproduce the relief learning experiments by simulating odour and shock exposures with different inter stimulus intervals (ISI). ISI represents the relative timing difference between shock and odour onset. A positive ISI value means that shock onset precedes odour one. Both shock and odour are simulated for 60 seconds. ISIs higher than 60 seconds corresponds with the situation when there is no overlap between shock and odour induced activities.

We presented a single odour (Fig 5.4 a) to the network 20 seconds after PPL1 activation (Fig 5.4 d) and observed potentiation of KC-MVP2 weights in response to the odour (Fig 5.4 b). After simulating relief learning we found that the firing rate of the MVP2 MBON increased its response to the odour (Fig 5.4 b).

PPL1 DAn fired at baseline activity levels for the first 60 seconds. When we simulated twelve one second shocks, we observed a transient increase in the firing rate of PPL1, followed by a period of decreased activity due to spike-frequency adaptation after the final shock (Fig 5.4 d).

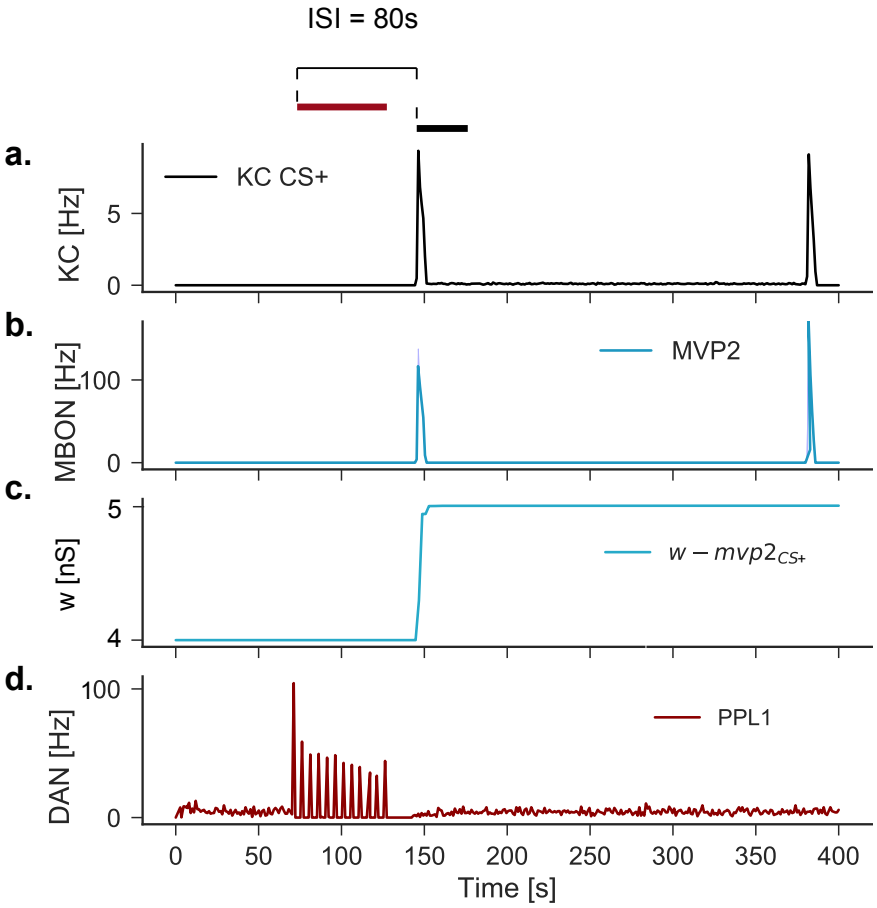


Figure 5.4: Relief learning simulation when inter stimulus interval ($t^{start}_{shock} - t^{start}_{odour}$) is 80 seconds. Black bar at the top denotes odour duration, while red bar denotes shock duration. **a.** Shown here is the population firing rate of the CS+ responding KCs. **b.** The firing rate of the MVP2 MBON during the simulation. **c.** We show the average weight between KCs that respond to the CS+ odour and the MVP2 MBON ($w_{KC-MVP2}$) **d.** We show the firing rate of the PPL1 DAn during relief learning.

In relief learning, shock that precedes an odour by up to 40 seconds (ISI = 100) creates a positive odour specific memory in the fly (Yarali et al., 2008). To test whether we can reproduce this result, we simulated ISIs ranging from -10 seconds to 150 seconds and visualized the firing rates of MVP2, M6 MBONs as well as the resulting model performance index (Fig 5.5). Our results show that the strongest aversive memory was formed when odour was delivered together with shock exposure (ISI =0, $PI = -0.35$). Shock delay (ISI=-10) created a weaker aversive memory (PI=-0.28).

MVP2 MBON changed its fire bi-drectionally depending on the ISI. For ISIs between 60s and 100s the firing rate of MVP2 was potentiated. ISIs higher than 100s

did not affect the firing rate of the MVP2 neuron, while ISIs lower than 60s decreased the firing rate of the MVP2 neuron (Fig 5.5 a). Conversely, the M6 MBON which in our model is inhibited by the MVP2 MBON increased its firing rate when there was overlap between the shock and the odour (ISI < 60) and decreased when ISI was between 60s and 100s (Fig 5.5 b).

When shock preceded the odour, the strongest appetitive memory was formed when the last shock came immediately before the odour (ISI $= 60$ seconds, PI $=$ 0.11). This is due to spike-frequency adaptation that begins decreasing after the last shock and PPL1 spontaneous activity returns to the level it had before the first shock. Appetitive memories were formed for ISI up to 100 seconds which is similar to experimental results (Fig 5.5 c, König et al., 2018).

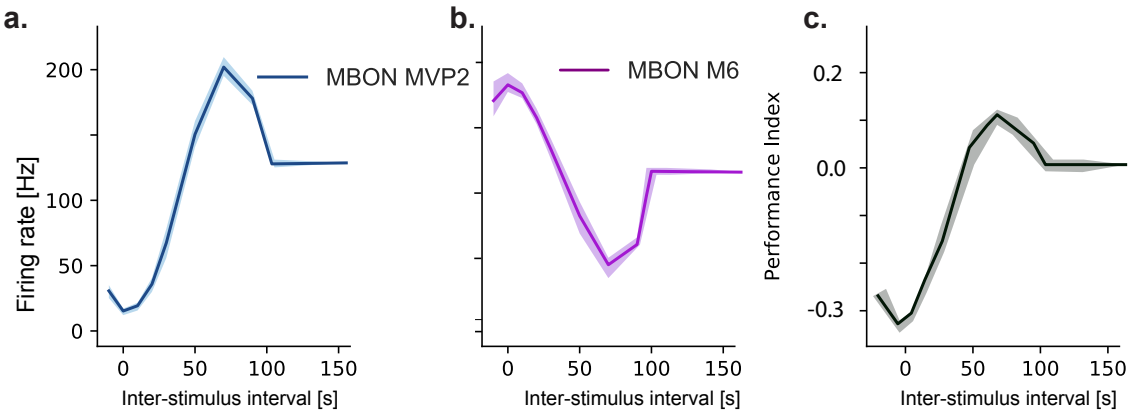


Figure 5.5: Rule pre-dopa-LTD-LTP reproduces relief learning experiment. Shown here are changes in the MBON firing rates and performance index for ISI ranging from -10 to 150 seconds. ISI is the difference between the timing of the first shock and odour exposure. **a.** MVP2 firing rate for simulations 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. Shaded areas indicate 1 standard deviation.

5.4.2 Re-learning with LTP enables re-extinction

In chapter 3 and 4 we modelled learning and re-learning using a learning rule that allows only for synaptic depression. We showed that we could not perform re-extinction because DAn firing rates in response to the trained odour were too low. Here, we test our new learning rule during re-learning and re-extinction and show that we can perform both aversive and appetitive re-extinction. In the next section we describe our results performing appetitive and aversive re-learning and re-extinction.

5.4.3 Aversive re-learning and re-extinction

We simulated the aversive re-learning protocol by reactivatingivating CS+ KCs after aversive extinction (Fig 5.6 a). During learning the odour is paired with 12 shocks that activate PPL1 (Fig 5.6 d) and PPL1-2 DAns. This leads to the decrease of CS+ KC-MVP2 average weight (Fig 5.6 c.). We have shown in chapter 3 that aversive learning leads to enhanced CS+ responses for M6 MBON and PAM-γ5DAn. After learning, the trained odour was re-exposed which led to the decrease of the average CS+ KC-M6 weight (Fig 5.6 c.). Finally, after extinction, we re-trained the odour with shock. During re-learning CS+ KC-M6 weight potentiated (Fig 5.6 c.). When we tested the two odours, we observed that the MBON-M6 firing rates and PAM-γ5 responses were higher for the CS+ odours than for the CS- (Fig 5.6 b, c).

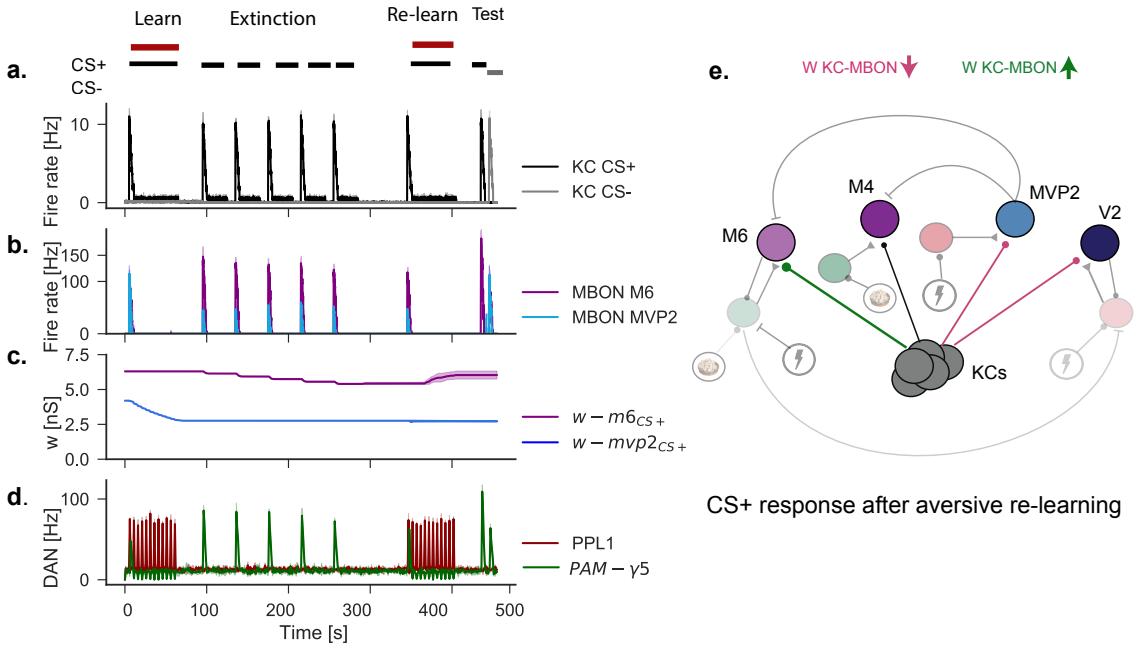


Figure 5.6: Network activity during aversive re-learning protocol using rule pre-dopa-LTD-LTP. **a.** We show KC population firing rates for cells that respond to CS+ (black) and CS- (grey) **b.** MBON MVP2 and M6 firing rates are shown during the simulation. **c.** KC-MVP2 (blue) and KC-M6 CS+ (purple) weight change during simulation **d.** Panel shows the firing rates of PAM1 and PPL1 DAns during the simulation. **e.** Shown here is the circuit and the state of the KC-MBON connections after re-learning. After re-learning CS+ KC-V2 and KC-MVP2 weights are still depressed, while KC-M6 synapses have undergone potentiation. Shaded areas indicate 1 standard deviation.

We subsequently looked at changes in CS+ responses in the MBON population

after re-learning. We tuned our learning rate for LTP to increase the CS+ firing rate of M6 MBON to the firing rate after initial learning (Fig 5.7 a). As M6 MBON excites PAM-γ5, an increase in the CS+ response of M6 will also result an increase in the CS+ response of PAM-γ5 which enables re-extinction. We also visualized the distribution of CS+ KC-M6 synapses and we observed that all of the weights were potentiated after re-learning (Fig 5.7 b). Our models' performance index after re-learning matches the experimental result (Fig 5.7 c.).

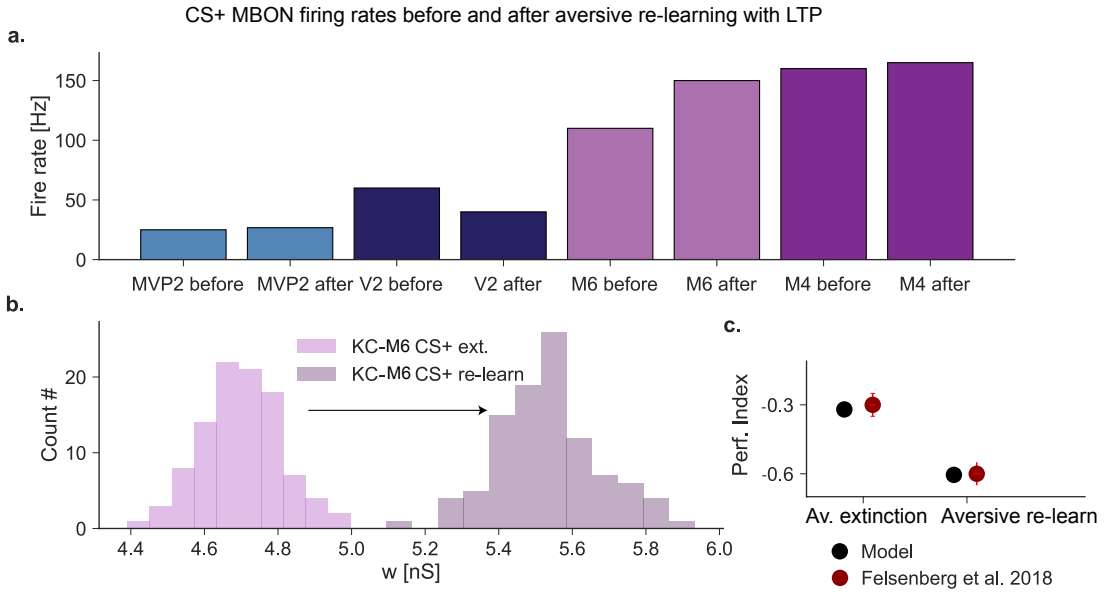


Figure 5.7: MBON firing rates, Performance Index and weights after aversive re-learning. **a.** MBON MVP2j (light blue), V2 (dark blue), M6 (light purple), M4 (dark purple) are shown before and after aversive re-learning. **b.** KC-M6 connection weights for KCs that respond to CS+ are shown before (light purple) and after (dark purple) aversive re-learning. **c.** Model performance index and experimental data from Felsenberg et al., 2018 (dark red) is shown after aversive training and after aversive extinction. Model (black) represents modelling result calculated from MBON firing rates using Equation 3.13. Vertical bars indicate standard error of the mean.

After investigating potentiation of CS+ KC-M6 synapses after aversive re-learning, we simulated aversive re-extinction. Aversive re-extinction triggered a decrease of CS+ KC-M6 which returned back to the level after initial learning (Fig 5.8). We conclude that using our new learning rule there is no limit into how many times an odour can change when there is LTP of KC-M6 synapses during re-learning.

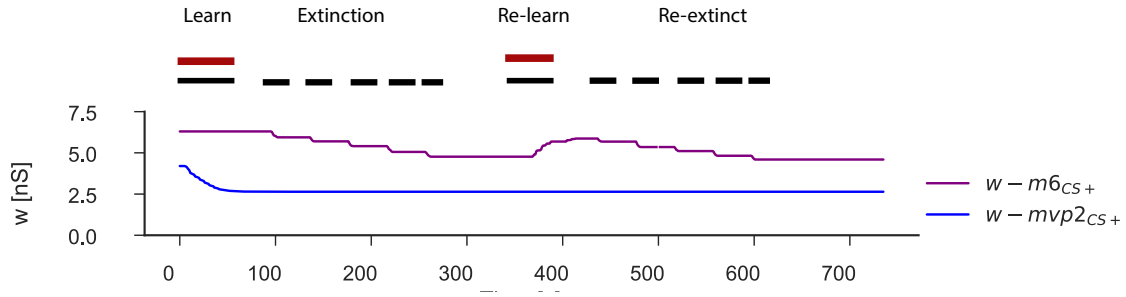


Figure 5.8: Rule pre-dopa-LTD-LTP enbles aversive re-extinction. We show the average weight for CS+ KC-MVP2 (blue) an KC-M6 (purple) connections during aversive re-extinction simulation.

5.4.4 Appetitive re-learning and re-extinction

We also tested appetitive re-learning using our pre-dopa-LTD-LTP learning rule by reactivating CS+ KCs (Fig 5.9 a) together with 1 second of PAN DAns activation (Fig 5.9 d). After appetitive extinction the firing rate of the V2 MBON which excites PPL1-2 DAn decreased. During appetitive re-learning when the CS+ odour is repaired with reward KC-V2 average weight increase (Fig 5.9 c). During testing our results show that the firing rate of V2 MBON increased (Fig 5.9 b). Strong activation of PAM-γ5 DAn, during re-learning triggers an inhibitory response in PPL1-2 neurons (Fig 5.9 d) as they are inhibited by PAM-γ5. This in turn is the reason why LTP of KC-V2 synapses occurs during re-learning.

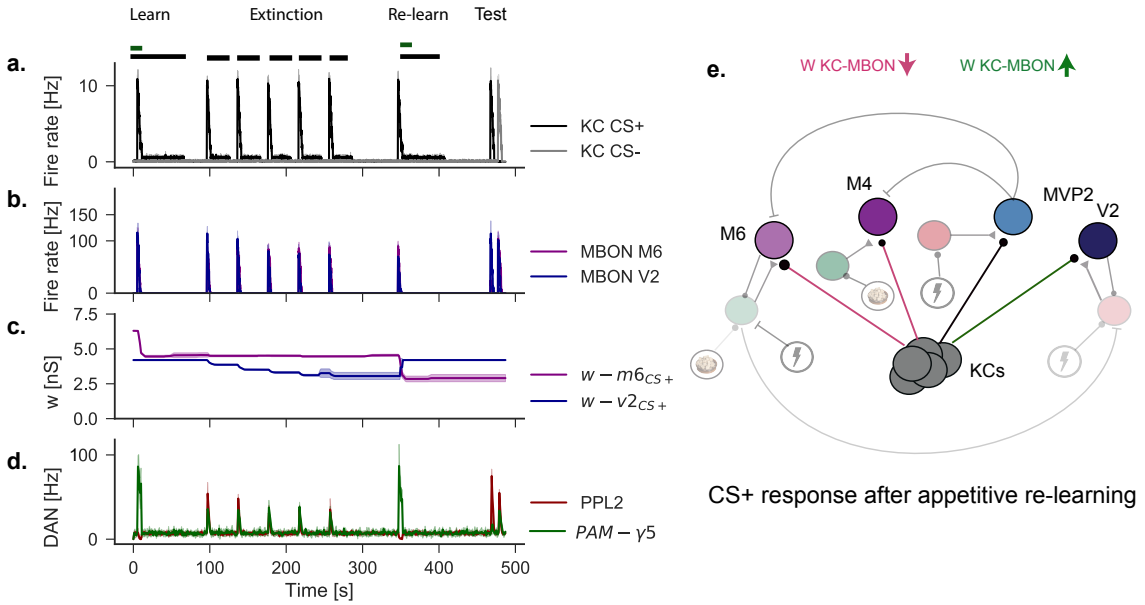


Figure 5.9: Network activity during appetitive re-learning protocol using rule pre-dopa-LTD-LTP. **a.** KC population firing rates in response to CS+ and CS- **b.** MBON MVP2 and M6 firing rates during the simulation. **c.** KC-MVP2 and KC-M6 CS+ weight change during simulation **d.** Panel shows the firing rates of PAM1 and PPL1-2 DAns during the simulation. **e.** We show the circuit and state of KC-MBON weights after appetitive re-learning. Shaded areas indicate one standard deviation.

After appetitive re-learning MBON CS+ odour responses increased in the V2 approach MBON and decreased in the M6 and M4 MBONs (Fig 5.10 a). We calculated the performance index which was higher after appetitive re-learning compared to initial learning (Fig 5.10 b).

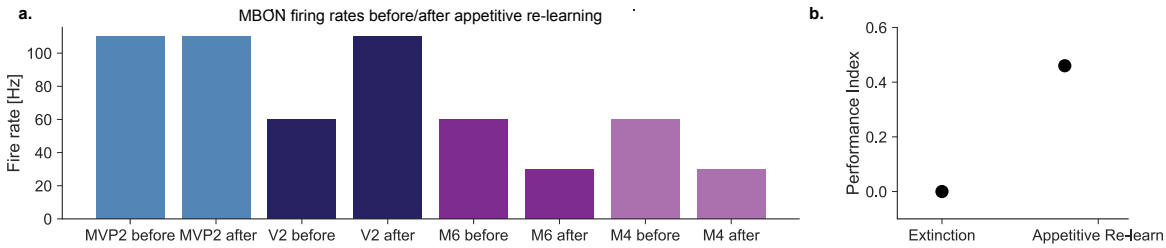


Figure 5.10: MBON firing rates, Performance Index and weights after appetitive re-learning. **a.** MBON MVP2 (light blue), V2 (dark blue), M6 (light purple), M4 (dark purple) are shown before and after aversive re-learning. **b.** Model performance index is shown after appetitive extinction and appetitive re-learning. Model (black) represents modelling result calculated from MBON firing rates using Equation 3.13.

Our learning rule also allows us to obtain appetitive re-extinction after re-learning,

that requires an increase in the CS+ V2 response. CS+ KC-V2 connections are strengthened during re-learning and depressed during re-extinction (Fig 5.11).

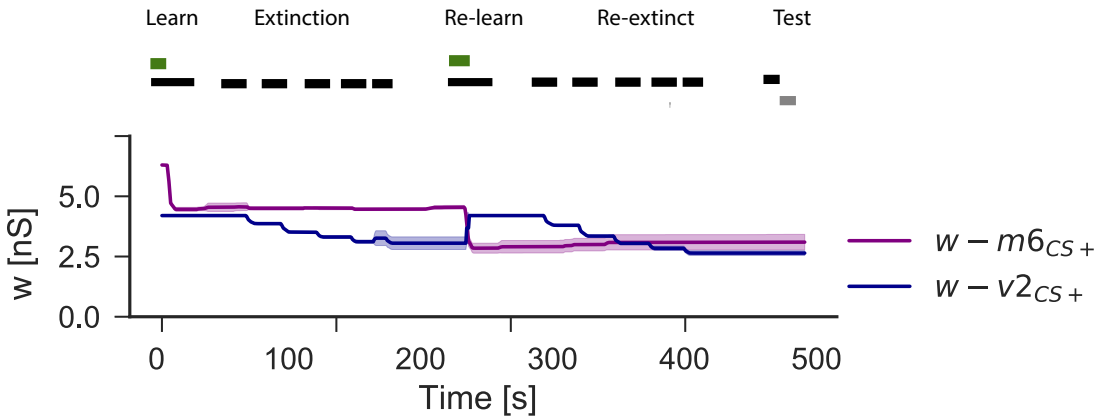


Figure 5.11: We show KC-V2 (blue) and KC-M6 (purple) CS+ weight change during appetitive re-extinction simulation. Shaded area illustrates one standard deviation.

5.4.5 Re-learning in the four-layer olfactory circuit

In chapter 2 we showed that a four-layer model of the olfactory circuit can decorrelate odours from a published (Hallem and Carlson, 2006) data-set, and can create a sparse representation in the mushroom body that only activates approximately 100 KCs per odour. To test learning in the full circuit, we have selected four odours that have very little overlap in KC representation (Odour 23 - CS1, 55 - CS2, 69 - CS3, 89 - CS+) from the data-set to test whether we could learn one odour without affecting the representation of the other three. We wanted to to test whether our learning rule could learn an odour without affecting the representation of the remaining three odours.

In the final section of this chapter we test the four-layer model from chapter 2 and we add DANs to the model to study whether our learning rule works in the complete model (Fig 5.12).

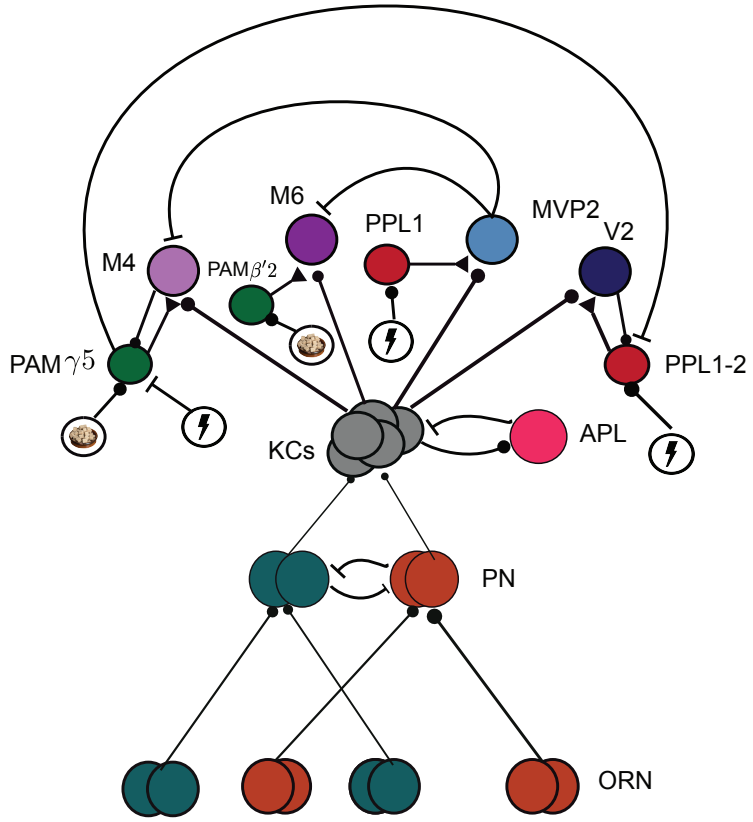


Figure 5.12: Olfactory circuit model diagram. Circuit diagram shows the main components of our model starting from bottom: 24 ORN types. ORNs of the same type excite 1 unique PN and 1 LN which form a glomerulus. LNs inhibit PNs which are not in the same glomerulus. PNs excite 1000 KCs, with each KC receiving input from 10 random PNs. KCs are randomly connected with excitatory connections to MBON V2, MVP2, M6 and M4. Connectivity to all four DAns (PAM-γ5, PAM-β'2, PPL1, PPL1-2 is identical to the two-layer model from chapter 2 (Fig 3.2)
.

We simulated aversive learning, extinction, re-learning and re-extinction in the full circuit. Each of the four odours elicited a different response in the population of PNs and KCs (Fig 5.13 a, b). After learning the response in the MBONs MVP2 and M6 decreased (Fig 5.13 c). During aversive learning average KC-MVP2 weight decreased (Fig 5.13 e) due to the PPL1 being activated to signal shock (Fig 5.13 d). 30 seconds later, during re-learning KC-M6 weights potentiated and finally, when

the odour was re-exposed the weights decreased, showing we can reproduce all of the aversive learning examples in the full circuit. We visualized the firing rates of the MVP2 (approach) MBON and M6 (retreat) in response to each of the four odours by plotting them in a 2D space. Our results show that all four odours elicited equal M6 and MVP2 responses before training the CS+ odours (Fig 5.13 g). After aversive learning, the response to CS+ changed while CS1, CS2, CS3 remained unchanged (Fig 5.13 h). Similarly, aversive extinction and re-learning only modified the M6 response to the CS2 odour (Fig 5.13 i,j).

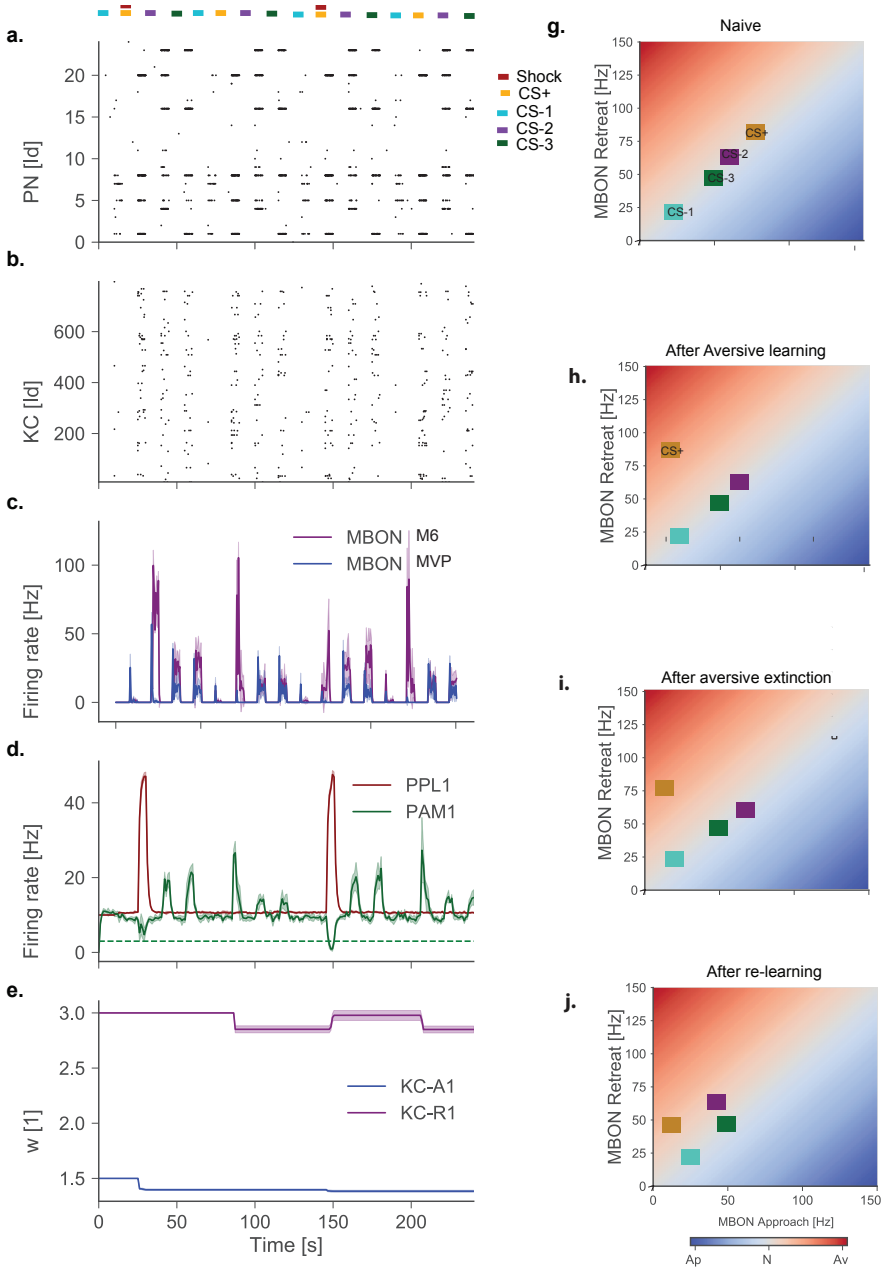


Figure 5.13: Learning, extinction, re-learning and re-extinction in the full network. We plot the responses of the neurons in our network to 4 stimuli: CS-1 (light blue), CS+ (yellow), CS-2 (purple), CS-3 (dark green). **a.** Spike raster of PN activity during the simulation. **b.** Spike raster of KC 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. **j.** Firing rates elicited by MBONs after aversive re-learning in response to the four odours.

5.5 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 the anatomy of the fruit fly: appetitive DAns target MBONs which promote retreat behaviour, while aversive DAns target MBON which code for approach (Aso et al., 2014). Optogenetic activation of DAns paired with an odour has been shown to lead to the decrease in the firing rates of MBONs (Hige et al., 2015a). This suggests that high concentration of dopamine during odour presentation leads to depression of KC-MBON synapses. This makes sense as 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 pre-dopa-LTD-LTP) to reproduce experiments that were not reproducible with a learning rule that only allows LTD to occur. Specifically, in relief learning $PPL1-\gamma1ped$ 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 re-extinction. During appetitive re-learning PPL1-2 DAN is inhibited by $PAM-\gamma5$ 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 $PAM-\gamma5$ is inhibited by shocks which leads to the potentiation of KC-M6 synapses. After re-learning the CS+ firing rates of M6 MBON and $PAM-\gamma5$ were enhanced which allowed to 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 pre-dopa-LTD-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 Hallem and Carlson, 2006 data-set and trained one odour with an 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. Training a single odour will affect all the odours that have overlapping KC representation. Here, we have only tested aversive learning mechanisms. In the future we would like to test appetitive learning mechanism as well and test the network's ability to classify odours.

We have proposed a new learning rule that can explain both how synaptic depression and potentiation is orchestrated by the activity of DAns. We proposed 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.In our previous chapters we have shown that precise circuit connectivity is required to reproduce learning experiments. Aversive and appetitive learning, extinction and re-learning change the activity of a subset of MBONs in our model (Fig 5.14) and it's currently unclear how many MBONs are modified by each stage of learning.

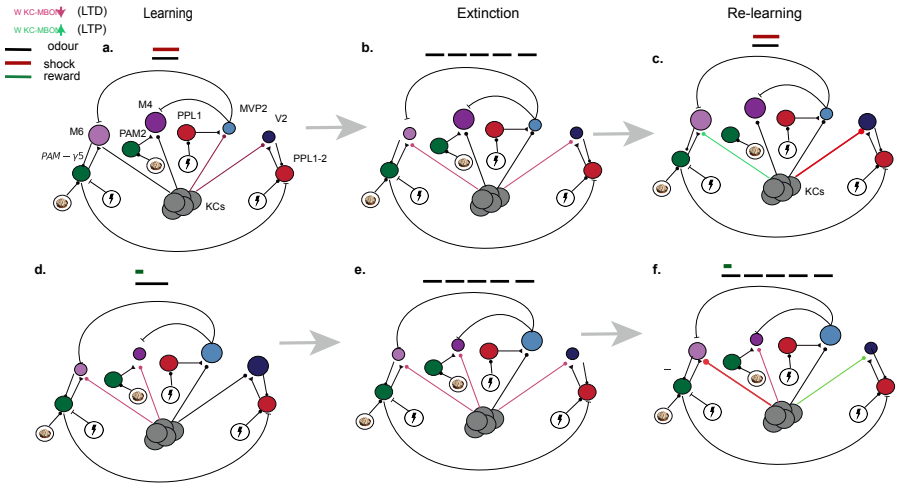


Figure 5.14: Stages of learning. We show the changes that occur in the trained odour responses of MBONs after learning, extinction and re-learning. **a.** After aversive learning the activity of approach promoting MBONs V2 and MVP2 are areduced. **b.** After aversive extinction retreat promoting M6 MBON decreases its response to the trained odour. **c.** After aversive re-learning the activity of M6 MBON is potentiated in response to the trained odour. Conversely, after appetitive learning. **d.** CS+ responses of retreat promoting M6 and M4 MBONs are reduced. **e.** Appetitive extinction leads to the reduction of the CS+ response of the approach promoting V2 MBON while re-learning. **f.** leads to the increase of V2 MBON response.

Chapter 6

Conclusions and future work

6.1 The final learning rule

In chapter 5 we showed that potentiation enables re-extinction and can explain relief learning. While high levels of dopamine paired with an odour triggers LTD, we added a second component to our learning rule that leads to the decrease of 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 and Carlson, 2006 data-set to train one odour without affecting the representation of the other odours. It's important to note that the final learning rule can reproduce all the experiments we've modelled: aversive and appetitive learning, extinction, re-learning, relief learning and second order conditioning.

Pre-dopa-LTD-LTP learning rule could potentially be explained by a two-receptor model. One receptor has a low affinity to dopamine and activated by high levels and orchestrates LTD, while a second receptor is activated at low dopamine concentration levels and orchestrates LTP. We speculate dDA1 is a candidate for the low affinity receptor, since it's required for both aversive and appetitive learning. 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 tuned to exhibit plausible dynamics of KCs such as transient responses to odours. 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 models' MBONs can respond to the majority of odours in the data-set.

6.2 Conclusions

In chapter 2 we explored sparse representation in the olfactory circuit of the fruit fly. We highlighted the role of lateral and feedback inhibition. However, some odours from the data-set elicited very few ORN spikes, which led to very weak PN and KC responses. Luo et al., 2010 have proposed that PNs normalize odour responses through a non-linear transformation. It would be very interesting to investigate whether we could reproduce their transformation proposed by using mechanisms such as gain modulation or inhibitory synaptic plasticity in the antennal lobe.

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 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. A model in the honeybee has also proposed a learning rule that is very similar to ours, which requires only presynaptic activity and dopamine to decrease KC-MBON weights (Peng and Chittka, 2017).

We propose that aversive extinction requires MBON-DAn excitatory. Appetitive extinction requires an appetitive DAN to aversive DAN inhibitory connection. In both cases MBON-M6 signalling to PAM-γ5 is required to either increase PAM-γ5 CS+ response after aversive learning or decrease it after appetitive learning. We conclude that while it appears learning does not require postsynaptic spikes, for 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-γ5 DAn to PPL1-2 DAn. Such a connection predicts that suppressing the activity of PAM-γ5 during odour exposure will create an aversive memory. We also explored a possible mechanism for second order conditioning. In our model, MVP2

MBON inhibits PPL1 DAn. After aversive learning, MVP2 decreases its CS+ response, while PPL1 increases. When we simulated second order conditioning by alternating CS+ exposure with a second untrained odour. If an order is exposed immediately before an aversive odour, PPL1 DAn activated by the CS+ odour can orchestrate the formation of an aversive memory in the MVP2 compartment of the previously untrained odour. However, in our model second order conditioning triggers extinction and is also affected by it. As the CS+ odour activates both the PPL1 DAn and the PAM-γ DAn. The authors of the second-order conditioning study have not observed extinction during the experiment but they observed that second order conditioning only works when the two odours are mixed together (Tabone and de Belle, 2011).

6.3 Future work

If we could normalize odour responses in PNs we might also be able to obtain very similar MBON responses for all of the odours, enabling us to use our learning rule to train the entire data-set. If each odour activates approximately 80 KCs, the overlap will affect learning using our rule. If our learning rules would modify only the most active KCs it would be interesting to see how many odours can be correctly classified. In summary, my thesis extends the understanding of the different stages of formation of memory in the olfactory system of the fruit fly and show that both depression and potentiation are required in the mushroom body. My results suggest that DAns in different compartments coordinate and communicate to learn and update the memory of an odour when the expected reward or punishment is absent and describes the mechanism for potentiation that can play a role both relief learning and re-learning. Such modelling efforts narrow the number of experiments necessary to test different hypotheses.

The fruit fly holds a lot of promise for understanding how memories are formed, the path ahead is best summarized by the following quote: "In the end there will be no substitute for the head-butting forward-genetic approach .... At heart this amounts to walking the beaches and peeking into bottles, with the hope that nature might have left a message inside... With luck, they will enlighten us about our capacity to remember" (Waddell and Quinn, 2001).

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, 2000. [Cited on page 65.]

Yoshinori Aso and Gerald M Rubin. Dopaminergic neurons write and update memories with cell-type-specific rules. *Elife*, 5:e16135, 2016. [Cited on pages 53 and 65.]

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, 2014. [Cited on pages 9, 13, 14, 16, 17, 18, 19, 22, 26, 31, 32, 53, 54, 59, 66, 86, and 113.]

Maxim Bazhenov, Mark Stopfer, Mikhail Rabinovich, Ramon Huerta, Henry DI Abarbanel, Terrence J Sejnowski, and Gilles Laurent. Model of transient oscillatory synchronization in the locust antennal lobe. *Neuron*, 30(2):553–567, 2001. [Cited on page 30.]

CDO Beck, Bradley Schroeder, and Ronald L Davis. Learning performance of normal and mutantdrosophila after repeated conditioning trials with discrete stimuli. *Journal of Neuroscience*, 20(8):2944–2953, 2000. [Cited on pages 54 and 71.]

Jan Benda and Andreas VM Herz. A universal model for spike-frequency adaptation. *Neural computation*, 15(11):2523–2564, 2003. [Cited on page 31.]

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. [Cited on page 32.]

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, 1998. [Cited on pages 3, 23, and 96.]

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

Tim VP Bliss, AR Gardner-Medwin, and T Lømo. Synaptic plasticity in the hippocampal formation. *Macromolecules and behaviour*, 193:203, 1973. [Cited on pages 3 and 96.]

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, 2012. [Cited on pages 20 and 21.]

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. [Cited on page 4.]

Stijn Cassenaer and Gilles Laurent. Conditional modulation of spike-timing-dependent plasticity for olfactory learning. *Nature*, 482(7383):47, 2012. [Cited on pages 3, 8, 23, and 96.]

Thomas A Cleland and Christiane Linster. Computation in the olfactory system. *Chemical senses*, 30(9):801–813, 2005. [Cited on page 30.]

Claudia Clopath, Lorric Ziegler, Eleni Vasilaki, Lars Büsing, and Wulfram Gerstner. Tag-trigger-consolidation: a model of early and late long-term-potentiation and depression. *PLoS computational biology*, 4(12):e1000248, 2008. [Cited on page 8.]

Raphael Cohn, Ianessa Morantte, and Vanessa Ruta. Coordinated and compartmentalized neuromodulation shapes sensory processing in drosophila. *Cell*, 163(7): 1742–1755, 2015. [Cited on pages 21, 59, 86, 96, and 98.]

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. [Cited on page 22.]

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. [Cited on page 14.]

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, 1994. [Cited on pages 9, 14, 16, and 26.]

Kenji Doya. Metalearning and neuromodulation. *Neural Networks*, 15(4-6):495–506, 2002. [Cited on page 6.]

Patrick J Drew and LF Abbott. Extending the effects of spike-timing-dependent plasticity to behavioral timescales. *Proceedings of the National Academy of Sciences*, 103(23):8876–8881, 2006. [Cited on page 97.]

Kristina V Dylla, Dana S Galili, Paul Szyszka, and Alja Lüdke. Trace conditioning in insects—keep the trace! *Frontiers in physiology*, 4:67, 2013. [Cited on page 15.]

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. [Cited on pages 54, 55, 62, and 83.]

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. [Cited on page 14.]

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. [Cited on pages 55 and 59.]

Faramarz Faghihi, Ahmed A Moustafa, Ralf Heinrich, and Florentin Wörgötter. A computational model of conditioning inspired by drosophila olfactory system. *Neural Networks*, 87:96–108, 2017. [Cited on page 25.]

Farzad Farkhooi, Anja Froese, Eilif Muller, Randolf Menzel, and Martin P Nawrot. Cellular adaptation facilitates sparse and reliable coding in sensory pathways. *PLoS computational biology*, 9(10):e1003251, 2013. [Cited on page 31.]

Johannes Felsenberg, Oliver Barnstedt, Paola Cognigni, Suewei Lin, and Scott Waddell. Re-evaluation of learned information in drosophila. *Nature*, 544(7649):240, 2017. [Cited on pages 12, 13, 54, 57, 86, 87, 88, and 90.]

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. [Cited on pages 9, 12, 13, 19, 23, 54, 55, 56, 57, 59, 69, 71, 72, 74, 81, 83, 86, 98, 100, and 106.]

Nicolas Frémaux, Henning Sprekeler, and Wulfram Gerstner. Functional requirements for reward-modulated spike-timing-dependent plasticity. *Journal of Neuroscience*, 30(40):13326–13337, 2010. [Cited on page 61.]

Uwe Frey, Richard GM Morris, et al. Synaptic tagging and long-term potentiation. *Nature*, 385(6616):533–536, 1997. [Cited on page 8.]

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. [Cited on page 21.]

C Giovanni Galizia. Olfactory coding in the insect brain: data and conjectures. *European Journal of Neuroscience*, 39(11):1784–1795, 2014. [Not cited.]

Wulfram Gerstner and Romain Brette. Adaptive exponential integrate-and-fire model. *Scholarpedia*, 4(6):8427, 2009. [Cited on pages 26, 27, and 36.]

Wulfram Gerstner, Marco Lehmann, Vasiliki Liakoni, Dane Corneil, and Johanni Brea. Eligibility traces and plasticity on behavioral time scales: experimental support of neohebbian three-factor learning rules. *Frontiers in neural circuits*, 12, 2018. [Cited on page 96.]

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. [Cited on page 4.]

Q Gu. Neuromodulatory transmitter systems in the cortex and their role in cortical plasticity. *Neuroscience*, 111(4):815–835, 2002. [Cited on page 4.]

Elissa A Hallem and John R Carlson. Coding of odors by a receptor repertoire. *Cell*, 125(1):143–60, 2006. [Cited on pages v, 14, 24, 31, 32, 33, 35, 38, 44, 45, 46, 47, 48, 49, 50, 101, 109, 113, and 115.]

Kyung-An Han, Neil S Millar, Michael S Grotewiel, and Ronald L Davis. Damb, a novel dopamine receptor expressed specifically in drosophila mushroom bodies. *Neuron*, 16(6):1127–1135, 1996. [Cited on page 23.]

DO Hebb. The organization of behavior. 1949. *New York Wiely*, 2(7):8, 1949. [Cited on page 3.]

Lovisa Irpa Helgadottir, Joachim Haenicke, Tim Landgraf, Raul Rojas, and Martin P Nawrot. Conditioned behavior in a robot controlled by a spiking neural network. In *2013 6th International IEEE/EMBS Conference on Neural Engineering (NER)*, pages 891–894. IEEE, 2013. [Cited on page 8.]

Fritjof Helmchen, Keiji Imoto, and Bert Sakmann. Ca2+ buffering and action potential-evoked ca2+ signaling in dendrites of pyramidal neurons. *Biophysical journal*, 70(2):1069–1081, 1996. [Cited on page 56.]

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. [Cited on pages 5, 20, 55, 59, 78, 83, 96, 98, and 113.]

Toshihide Hige, Yoshinori Aso, Gerald M Rubin, and Glenn C Turner. Plasticity-driven individualization of olfactory coding in mushroom body output neurons. *Nature*, 2015b. [Cited on pages 18, 26, and 61.]

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, 2011. doi: 10.1523/JNEUROSCI.1099-11.2011. [Cited on pages 16, 26, 50, and 54.]

Eugene M Izhikevich. Solving the distal reward problem through linkage of stdp and dopamine signaling. *Cerebral cortex*, 17(10):2443–2452, 2007. [Cited on pages 7 and 63.]

Bahadir Kasap and Michael Schmuker. Improving odor classification through self-organized lateral inhibition in a spiking olfaction-inspired network. In *2013 6th*

International IEEE/EMBS Conference on Neural Engineering (NER), pages 219–222. IEEE, 2013. [Cited on page 51.]

Karla R Kaun, Reza Azanchi, Zaw Maung, Jay Hirsh, and Ulrike Heberlein. A drosophila model for alcohol reward. *Nature neuroscience*, 14(5):612, 2011. [Cited on page 22.]

Young-Cho Kim, Hyun-Gwan Lee, and Kyung-An Han. D1 dopamine receptor dda1 is required in the mushroom body neurons for aversive and appetitive learning in drosophila. *Journal of Neuroscience*, 27(29):7640–7647, 2007. [Cited on page 22.]

Christian König, Afshin Khalili, Mathangi Ganesan, Amrita P Nishu, Alejandra P Garza, Thomas Niewalda, Bertram Gerber, Yoshinori Aso, and Ayse Yarali. Reinforcement signaling of punishment versus relief in fruit flies. *Learning & Memory*, 25(6):247–257, 2018. [Cited on pages 97 and 104.]

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, 2007. [Cited on page 16.]

Robert Legenstein, Dejan Pecevski, and Wolfgang Maass. A learning theory for reward-modulated spike-timing-dependent plasticity with application to biofeedback. *PLoS computational biology*, 4(10):e1000180, 2008. [Cited on page 8.]

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. [Cited on pages 17, 18, 22, 30, 32, 47, and 50.]

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. [Cited on page 21.]

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, 2010. doi: 10.1073/pnas.1005635107. [Cited on pages 14, 24, and 116.]

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. [Cited on page 24.]

Eve Marder and Vatsala Thirumalai. Cellular, synaptic and network effects of neuromodulation. *Neural Networks*, 15(4-6):479–493, 2002. [Cited on page 3.]

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. [Cited on page 3.]

Mala Murthy, Ila Fiete, and Gilles Laurent. Testing odor response stereotypy in the drosophila mushroom body. *Neuron*, 59(6):1009–1023, 2008. [Cited on pages 27, 31, 58, and 83.]

Thomas Nowotny, Mikhail I Rabinovich, Ramon Huerta, and Henry DI Abarbanel. Decoding temporal information through slow lateral excitation in the olfactory system of insects. *Journal of computational neuroscience*, 15(2):271–281, 2003. [Cited on page 30.]

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. [Cited on page 2.]

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. [Cited on pages 30 and 50.]

Randall C O'Reilly and James L McClelland. Hippocampal conjunctive encoding, storage, and recall: Avoiding a trade-off. *Hippocampus*, 4(6):661–682, 1994. [Cited on page 30.]

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, 2015. [Cited on pages 18, 19, 21, 22, 54, 59, 71, and 83.]

Ivan Petrovich Pavlov. Conditional reflexes: an investigation of the physiological activity of the cerebral cortex. *Oxford Press*, 1927. [Cited on pages 2 and 9.]

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. [Cited on pages 20 and 59.]

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. [Cited on pages 3 and 23.]

Fei Peng and Lars Chittka. A simple computational model of the bee mushroom body can explain seemingly complex forms of olfactory learning and memory. *Current Biology*, 27(2):224–230, 2017. [Cited on page 116.]

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, 2013. [Cited on page 17.]

Christian Pozzorini, Richard Naud, Skander Mensi, and Wulfram Gerstner. Temporal whitening by power-law adaptation in neocortical neurons. *Nature neuroscience*, 16(7):942, 2013. [Cited on page 98.]

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, 1974. [Cited on page 9.]

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

Michael Schmuker, Nobuhiro Yamagata, Martin Nawrot, and Randolf Menzel. Parallel representation of stimulus identity and intensity in a dual pathway model inspired by the olfactory system of the honeybee. *Frontiers in neuroengineering*, 4: 17, 2011. [Cited on page 30.]

Wolfram Schultz, Peter Dayan, and P Read Montague. A neural substrate of prediction and reward. *Science*, 275(5306):1593–1599, 1997. [Cited on pages 2 and 6.]

Julien Séjourné, Pierre-Yves Plaçais, Yoshinori Aso, Igor Siwanowicz, Séverine Trannoy, Vladimiros 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. [Cited on page 20.]

Weixing Shen, Marc Flajolet, Paul Greengard, and D James Surmeier. Dichotomous dopaminergic control of striatal synaptic plasticity. *Science*, 321(5890):848–851, 2008. [Cited on page 5.]

Jake P Stroud, Mason A Porter, Guillaume Hennequin, and Tim P Vogels. Motor primitives in space and time via targeted gain modulation in cortical networks. *Nature neuroscience*, 21(12):1774, 2018. [Cited on page 50.]

Martin F Strube-Bloss, Martin P Nawrot, and Randolf Menzel. Neural correlates of side-specific odour memory in mushroom body output neurons. *Proceedings of the Royal Society B: Biological Sciences*, 283(1844):20161270, 2016. [Cited on page 19.]

Chih-Ying Su, Karen Menuz, Johannes Reisert, and John R Carlson. Non-synaptic inhibition between grouped neurons in an olfactory circuit. *Nature*, 492(7427):66, 2012. [Cited on page 50.]

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. [Cited on page 51.]

Richard S Sutton and Andrew G Barto. *Introduction to reinforcement learning*, volume 135. MIT press Cambridge, 1998. [Cited on page 5.]

Christopher J Tabone and J Steven de Belle. Second-order conditioning in drosophila. *Learning & Memory*, 18(4):250–253, 2011. [Cited on pages 12, 53, and 117.]

Kay Thurley, Walter Senn, and Hans-Rudolf Luscher. Dopamine increases the gain of the input-output response of rat prefrontal pyramidal neurons. *Journal of neurophysiology*, 99(6):2985–2997, 2008. [Cited on page 4.]

Alessandro Treves and Edmund T Rolls. What determines the capacity of autoassociative memories in the brain? *Network: Computation in Neural Systems*, 2(4): 371–397, 1991. [Cited on page 41.]

Glenn C Turner, Maxim Bazhenov, and Gilles Laurent. Olfactory Representations by Drosophila Mushroom Body Neurons. *Journal of Neurophysiology*, 99(2):734–746, 2008. [Cited on pages 16, 31, 32, and 58.]

M Vandecasteele, J-M Deniau, and L Venance. Spike frequency adaptation is developmentally regulated in substantia nigra pars compacta dopaminergic neurons. *Neuroscience*, 192:1–10, 2011. [Cited on page 98.]

Tim P Vogels and Larry F Abbott. Signal propagation and logic gating in networks of integrate-and-fire neurons. *Journal of neuroscience*, 25(46):10786–10795, 2005. [Cited on page 27.]

Tim P Vogels and LF Abbott. Gating multiple signals through detailed balance of excitation and inhibition in spiking networks. *Nature neuroscience*, 12(4):483, 2009. [Cited on page 50.]

Tim P Vogels, Henning Sprekeler, Friedemann Zenke, Claudia Clopath, and Wulfram Gerstner. Inhibitory plasticity balances excitation and inhibition in sensory pathways and memory networks. *Science*, 334(6062):1569–1573, 2011. [Cited on page 51.]

Leslie B Vosshall, Allan M Wong, and Richard Axel. An olfactory sensory map in the fly brain. *Cell*, 102(2):147–159, 2000. [Cited on page 14.]

Scott Waddell. Reinforcement signalling in Drosophila; dopamine does it all after all. *Current opinion in neurobiology*, 23(3):324–329, 2013. [Cited on pages 4, 9, 21, and 23.]

Scott Waddell and William G Quinn. Flies, genes, and learning. *Annual review of neuroscience*, 24(1):1283–1309, 2001. [Cited on page 117.]

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, 2011. [Cited on pages 24, 25, and 32.]

Benjamin Willmore and David J Tolhurst. Characterizing the sparseness of neural codes. *Network: Computation in Neural Systems*, 12(3):255–270, 2001. [Cited on page 38.]

Jason Wolfe, Arthur R Houweling, and Michael Brecht. Sparse and powerful cortical spikes. *Current opinion in neurobiology*, 20(3):306–312, 2010. [Cited on page 47.]

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, 2015. [Cited on page 21.]

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. [Cited on pages 12, 13, 83, 96, and 103.]

Friedemann Zenke and Wulfram Gerstner. Limits to high-speed simulations of spiking neural networks using general-purpose computers. *Frontiers in neuroinformatics*, 8:76, 2014. [Cited on page 32.]

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. [Cited on page 4.]