Learning Representations with Neuromodulators

vorgelegt von
M. Sc.
Raphaël Holca-Lamarre
geb. in Montréal

von der Fakultät IV – Elektrotechnik und Informatik
der Technischen Universiät Berlin
zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften
- Dr. rer. nat. -

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender:	Prof. Dr. Hennig Sprekeler
Gutachter:	Prof. Dr. Klaus Obermayer
Gutachter:	Prof. Dr. Jörg Lücke
Gutachterin:	Prof. Dr. Shu-Chen Li

Tag der wissenschaftlichen Aussprache: 30. Juni 2017

Berlin 2017

To those from whom I learned,
alongside whom I went on adventures,
or with whom I shared a good laugh.

Acknowledgements

First of all, I would like to thank my supervisors Prof. Klaus Obermayer and Prof. Jörg Lücke. Klaus and Jörg provided valuable support throughout my PhD, from the initial application for funding to the execution of the work and up until the results publication. Then come all the members of the Neural Information Processing group who I thank, among many others, for their feedback during group talks and their company over lunch breaks–especially enjoyable when sitting outside in the sun. I would also like to express my gratitude to everyone at the Bernstein Centre, both staff and students, for providing a supportive environment, in particular for the many events and courses, the financial support, and the good times, among others at the annual retreat at Schloß Tornow. And last but not least, I acknowledge the funding bodies that made this research possible, the Studientstiftung des deutschen Volkes, the Fonds de Recherche du Québec – Nature et Technologies, and the Deutsche Forschungsgemeinschaft.

Abstract

Neurons in the cortex and in multi-layer perceptrons (MLPs) represent inputs as patterns of activity. In both systems, the nature of a neural representation carries important consequences and, accordingly, the mechanisms guiding representation learning are of great relevance. In MLPs, the most widespread learning method is the error-backpropagation algorithm. Although functionally effective, this method bears its limitations and is also unlikely to be implemented in the cortex. On the other hand, the mechanisms orchestrating plasticity in cortical representations remain unclear; neuromodulators appear to contribute to this process but their exact roles are largely unknown.

This thesis examines the plastic effects of two neuromodulators, acetylcholine (ACh) and dopamine (DA), with the following aims: first, to gain a functional understanding of ACh and DA transmission in shaping biological representations and, second, to explore neuromodulator-inspired learning rules for MLPs. I address these questions in a Hebbian-learning neural network model. I extend the model to simulate the plastic effects and physiological release properties of ACh and DA. I then study the impact of neuromodulatory transmission on the network's representation and its performance on a classification task.

In the model, ACh activation approximates the concept of attentional effort. I demonstrate that this signal redistributes neural preferences such that more neurons encode challenging or relevant stimulus classes, thereby boosting performance on these classes. DA activation in the model accompanies reward prediction errors. I show that this signal adjusts neural weights to the reward contingencies of a task, in turn enhancing class selectivity in neurons and yielding large gains in classification accuracy. These results suggest functional roles for ACh and DA in guiding biological representation learning. Additionally, in comparison with MLPs of the same architecture, neuromodulator-inspired learning produces lower error rates than those originally reported on the MNIST dataset and measures up with modern state-of-the-art optimisation methods. A single modulatory signal also proves to successfully guide learning in multiple layers concurrently in a deep convolutional network. Neuromodulator learning requires weak supervision signals and interacts with synaptically-local weight updates, thus offering potential applications in learning with weakly labelled data or in neuromorphic processors.

Zusammenfassung

Neuronen im Kortex und in künstlichen neuronalen Netzwerken (z.B. “Multi-layer perceptrons”, MPLs) repräsentieren Inputs als Aktivitätsmuster. In beiden Systemen ist die Natur des neuronalen Codes entscheidend und folglich sind die Mechanismen, die das Lernen von Repräsentationen beeinflussen, von großer Bedeutung. In künstlichen Netzwerken ist der ‘error-backpropagation’ Algorithmus die am weitesten verbreitete Lernmethode. Es ist unwahrscheinlich, dass diese Methode im Kortex implementiert ist, jedoch ist derzeit nicht bekannt, wie das Gehirn Repräsentationen lernt.

Die beiden Neuromodulatoren, Acetylcholin (ACh) und Dopamin (DA), lösen plastische Veränderungen in kortikalen Repräsentationen aus. In dieser Arbeit untersuche ich die funktionalen Rollen, die diese beiden Modulatoren, in Bezug auf die Verbesserung sensorischer Repräsentationen, haben. Dabei stehen zwei Ziele im Vordergrund: Zum einen, soll ein funktionales Verständnis der von ACh und DA ausgelösten Plastizität in der Biologie geschaffen werden. Zum anderen, sollen Lernregeln für MPLs erforscht werden, welche von Neuromodulatoren inspiriert sind. Ich behandle diese Fragen mithilfe eines neuronalen Netzwerkmodells, welches auf Hebb’schen Lernregeln basiert. Dazu simuliere ich ACh und DA als Modulatoren der Lernrate des Netzwerkes und bestätige dadurch, dass dieses Modell bekannte Plastizitätseffekte aus der Biologie reproduziert. Danach simuliere ich die physiologischen Eigenschaften der Ausschüttung von ACh und DA und quantifiziere deren Auswirkung mithilfe einer Klassifikationsaufgabe.

Im Modell folgt die Auschüttung von ACh den Anforderungen der Aufgabe und der Motivation, als Näherung des Konzepts des Aufmerksamkeitsaufwandes (‘attentional effort’). Ich zeige, dass dieses Signal neuronale Präferenz derart umverteilt, dass mehr Neuronen herausfordernde oder relevante Stimulusklassen kodieren, was die Klassifikation dieser Klassen deutlich verbessert. Ich gehe daher davon aus, dass ACh diese Rolle in der Biologie ausfüllt. DA Auschüttung wird im Modell von Belohnungsvorhersagefehlern (‘reward prediction errors’) hervorgerufen. Dieses Signal verbessert neuronale Selektivität, was wiederum zu einer Steigerung der Klassifikationsgenauigkeit führt. Ich vermute daher, dass die DA-induzierte Plastizität in senorischen Kortexarialen diese Funktion übernimmt. Angewandt auf den MNIST Datensatz zeigt DA-basiertes Lernen eine sehr gute Leistung, die vergleichbar mit modernen Optimisierungsmethoden in MLPs ist. Diese aussichtsreichen Ergebnisse motivieren weitere Forschungen auf dem Gebiet des neuromodulatorisch-inspirierten Lernens als Modell in der Biologie sowie für praktische Anwendungen.

Contents

List of Figures

List of Abbreviations

Abbreviation	Description
5-HT	Serotonin
A1	Primary Auditory Cortex
ANN	Artificial Neural Network
ACh	Acetylcholine
BvSB	Best versus Second-Best
ConvNet	Convolutional Neural Network
DA	Dopamine
EM	Expectation Maximisation algorithm
GPU	Graphical Processing Unit
IT	Inferotemporal area
KL	Kullback-Leibler divergence
LGN	Lateral Geniculate Nucleus
MAP	Maximum A Posteriori
MLP	Multi-Layer Perceptron
MNIST	Mixed National Institute of Standards and Technology dataset
NB	Nucleus Basalis of Meynert
NE	Noradrenaline
OFC	Orbitofrontal Cortex
PCA	Principal Component Analysis
RW	Rescorla-Wagner rule
SVM	Support Vector Machine
t-SNE	t-Distributed Stochastic Neighbour Embedding
VTA	Ventral Tegmental Area
V1	Primary Visual Cortex

List of Variables

Variable	Description
A	normalisation constant
$\tilde{y}$	activation of input neurons before normalisation
y	activation of input neurons after normalisation
d	index of input neurons
D	number of input neurons
W	weight from input to representation neuron
h	arbitrary transfer function
Q	weight from input to convolutional neuron
$\tilde{u}$	activation of convolutional neuron before lateral inhibition
u	activation of convolutional neuron
f	index of feature map
F	number of feature maps
$\tilde{v}$	activation of subsampling neuron before lateral inhibition
v	activation of subsampling neuron
$\tilde{s}$	activation of representation neuron before lateral inhibition
s	activation of representation neuron
c	index of the representation neurons
C	number of representation neurons
B	weight from representation to classification neuron
t	posterior of the classifier
k	index of the posterior of the classifier
K	number of posteriors or classes
m	label of image
$\hat{m}$	classification output of the network
M	number of unique labels
N	number of training images
p	rate of correct classification
κ	classification confidence
R	delivered reward
P	predicted reward
$S(\cdot)$	linearised logarithmic function
ACh	Acetylcholine activation
DA	Dopamine activation
$\mathcal{U}$	uniform distribution
$\mathcal{N}$	normal distribution

List of Parameters

Parameter	Description	Value
ϵ	learning rate of the shallow network	1e-4
ϵ^{conv}	learning rate of the convolutional layer	1e-6
ϵ^{feed}	learning rate of the fully connected feedforward layer	1e-2
A	normalisation constant	900
$\delta^{g}_{+/+}$	DA activation for (r_+, p_+) in the greedy network	0.10
$\delta^{g}_{+/0}$	DA activation for (r_+, p_0) in the greedy network	-1.00
$\delta^{\eta}_{+/+}$	DA activation for (r_+, p_+) in the explorative network	0.01
$\delta^{\eta}_{+/0}$	DA activation for (r_+, p_0) in the explorative network	-1.00
$\delta^{\eta}_{0/+}$	DA activation for (r_0, p_+) in the explorative network	4.00
$\delta^{\eta}_{0/0}$	DA activation for (r_0, p_0) in the explorative network	-0.25
δ^{lin}_{1}	parameter of the linear DA activation function	6.0
δ^{lin}_{2}	parameter of the linear DA activation function	-0.4
δ^{exp}_{1}	parameter of the exponential DA activation function	2.0
δ^{exp}_{2}	parameter of the exponential DA activation function	-1.1
τ	temperature of the softmax function	1.0
γ	spread of the normal distribution for noise injection (shallow net)	0.3
γ^{conv}	spread of the normal distribution for noise injection (conv. net)	0.5
γ^{feed}	spread of the normal distribution for noise injection (conv. net)	0.2
α	parameter of the ACh activation function (curve steepness)	15
β	parameter of the ACh activation function (curve height)	30
$\delta^{g,\text{conv}}_{+/+}$	DA activation for (r_+, p_+), greedy net, convolutional layer	0.0
$\delta^{g,\text{conv}}_{+/0}$	DA activation for (r_+, p_0), greedy net, convolutional layer	-2.0
$\delta^{g,\text{feed}}_{+/+}$	DA activation for (r_+, p_+), greedy net, feedforward layer	0.1
$\delta^{g,\text{feed}}_{+/0}$	DA activation for (r_+, p_0), greedy net, feedforward layer	-1.0
$\delta^{\eta,\text{conv}}_{+/+}$	DA activation for (r_+, p_+), explorative net, convolutional layer	0.1
$\delta^{\eta,\text{conv}}_{+/0}$	DA activation for (r_+, p_0), explorative net, convolutional layer	0.0
$\delta^{\eta,\text{conv}}_{0/+}$	DA activation for (r_0, p_+), explorative net, convolutional layer	0.0
$\delta^{\eta,\text{conv}}_{0/0}$	DA activation for (r_0, p_0), explorative net, convolutional layer	-0.5
$\delta^{\eta,\text{feed}}_{+/+}$	DA activation for (r_+, p_+), explorative net, feedforward layer	0.01
$\delta^{\eta,\text{feed}}_{+/0}$	DA activation for (r_+, p_0), explorative net, feedforward layer	-1.00
$\delta^{\eta,\text{feed}}_{0/+}$	DA activation for (r_0, p_+), explorative net, feedforward layer	6.00
$\delta^{\eta,\text{feed}}_{0/0}$	DA activation for (r_0, p_0), explorative net, feedforward layer	-1.00

Chapter 1

Introduction

1.1 The Brain's Shadow Theatre

A time-honoured form of storytelling, shadow theatre narrates stories of war, heroes and love. Puppets passing in front of a light cast shadows onto a canvas, provoking fear or laughter in the audience. The characters' shadows can be ambiguous, changing shape depending on their orientation (Fig. 1.1); a particular silhouette may be unclear and difficult to identify. Alternately, two puppets projecting similar or overlapping shadows are challenging to distinguish.

Alike puppets in shadow theatre, sensory events cast projections in our brains. Receptors sensing a stimulus in the periphery of our nervous system send swift signals to the brain where they elicit patterns of activity in neurons, the stimulus' neural shadow. The transformation of physical stimuli into neural codes form the brain's *representation* of its sensory environment. The particular form a representation takes matters: the silhouette a stimulus projects in the brain may be clearly recognisable–or not.

As an example, consider a student learning Mandarin Chinese. As a tonal language, Mandarin assigns meaning to the pitch of vowels: the syllables *mā*, *má*, *mǎ*, and *mà* each bear different interpretations ascribed by the tone on the vowel '*a*' (mother, hemp, horse, and scold, respectively). In a neural representation poorly suited to Mandarin, these syllables elicit similar activity patterns. The novice student then finds it difficult to tell the words apart.

Fortunately, evolution provided the mammalian brain with mechanisms to adapt sensory representations and facilitate perception. As the student practices at discerning the Mandarin tones, the representations of these tones adjust in his brain. From initially evoking identical responses, each vowel eventually gains a distinct representation: the learner can distinguish the tones at once.

This thesis is concerned with the mechanisms by which the brain refines its sensory representations, or its shadow theatre. In this chapter, I introduce the two main research objectives of this work. The first objective is to study the functions of two neuromodulators, acetylcholine and dopamine, in guiding plastic modifications in sensory representations. The second objective relates to artificial neural networks. Much like their biological counterpart, these networks transform data into neural representations. These algorithms learn representations through calculus-based optimisation methods which, although successful in practice, are unlike biological learning and bear

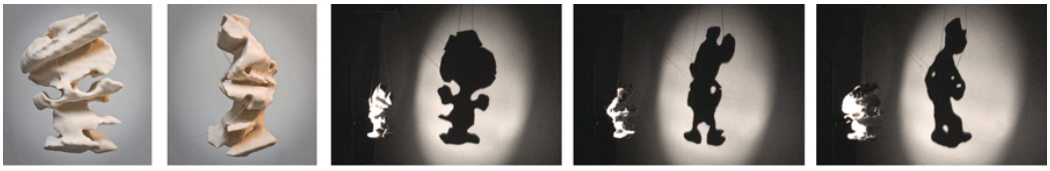

Figure 1.1: Shadow play. In shadow theatre, the same puppet may cast different silhouettes, here the characters Snoopy, Mickey Mouse, and Popeye (reproduced from [MP09] with permission). Alternately, two objects may project overlapping or similar shadows, making it difficult to tell them apart. Alike the puppets' silhouette, sensory events evoke patterns of neural activity in the brain. The specific neural representation of a stimulus makes it more or less evident to identify it.

their own limitations. Accordingly, the second aim of this work is to explore learning rules for artificial neural networks inspired from neuromodulation in the mammalian brain.

1.2 Biological Neural Representations

The turn of the 20th century saw the emergence of quantitative methods to study cognitive processes. The behaviourist movement then in place introduced stringent procedures to quantify the relationship between stimulus and behaviour. This line of research brought important contributions, notably the characterisation of classical and operant conditioning. Behaviourism, however, adopted an excessively reductionist view: it treated animal behaviour largely as a collection of reflexes and ignored transformations taking place within the brain.

In the 1960s, a new perspective emerged which regarded perception as a constructive process deriving not only from the stimulus but also from the mental structure of the perceiver. The discipline of cognitive psychology arose, investigating information processing within the brain. This new paradigm inherently assumed that percepts have internal representations in the brain.

Around those years, Evarts and Mountcastle developed novel tools to measure the responses of individual neurons in behaving animals. Such neural recordings allowed to relate the responses of nerve cells to specific sensory events. Soon after, Hubel and Wiesel famously discovered neurons in the primary visual cortex responding preferentially to edges of a particular orientation. These findings inspired a multitude of fruitful investigations into the neural representations of sensory events.

Researchers since observed neural correlates of all senses, from neurons responsive to specific odourants in the olfactory bulb to frequency-tuned cells in the auditory cortex. Cortical representations were also found for concepts disconnected from particular senses, like abstract rules or quantities and numbers. These neural representations often bear a topographic organisation such that nearby cells encode related information. Primary sensory areas additionally maintain ordered projections of sensory surfaces, for instance the roughly human-shaped tactile map of the body in the somatosensory cortex, or again the layout of the primary visual cortex following that of the retina in most mammals.

Sensory pathways are organised such that signals progress from one neural representation to another. This organisation is referred to as hierarchical due to the specific

ordering of processing stages along a sensory pathway. Each processing area extracts specific features from sensory signals; early stages typically encode simple attributes of limited spatial and temporal extend like edges in the visual system or pitch in the auditory system. Neurons in downstream stages respond to more intricate percepts such as the face or the name of an individual person. The transformation of sensory signals as they progress from one representation to another is thought to form the basis of perception in mammals.

Although at the time of their discoveries cortical representations were believed to be hard-wired and invariable throughout life, it is now clear that this is not the case. In particular, in the early 1990s, several experiments demonstrated that sensory representations adapt following extensive behavioural training. For instance, as monkeys repeatedly perform a fine tactile task with their finger tips, the area of their cortex devoted to these finger tips expands and neural responses sharpen. Improvements on such tasks were further shown to correlate with the extent of the cortical reorganisation, suggesting that perceptual sensitisation derives from refinements in neural representations.

It remains unclear how such refinements take place. Stimulation of the senses as well as active attentional involvement appear important ingredients. In addition, numerous experiments reveal an influential contribution of neuromodulators, organic molecules which modulate rather than cause neural responses. In particular, two compounds, acetylcholine (ACh) and dopamine (DA), promote modifications in neural representations. Released coincidentally with a sensory stimulus, both ACh and DA elicit enlargements in the cortical representation of the stimulus. Conversely, lesion of the cholinergic or dopaminergic system hinders forms of learning that require alterations in sensory representations. These findings implicate ACh and DA in the plasticity of cortical representations; the specific functions they carry, however, is uncertain. In this thesis, I examine the roles of ACh and DA transmission in a computational model of neural representation learning.

1.3 Artificial Neural Representations

Artificial intelligence has captivated researchers for generations. The history of the field is marked by a succession of bouts of enthusiasm and droughts of uncertainty in the abilities of synthetic intelligence. In the recent years, machine learning emerged as one of the most promising branch of artificial intelligence. Unlike some earlier methods which informed decisions on sets of predetermined logical rules, machine learning algorithms characteristically learn these rules from data. This approach proved far superior, requiring less human efforts and achieving greater performance, and it now enjoys most of the attention and funding in the field.

The accuracy of machine learning methods critically depends on the data representation an algorithm is to learn from. A representation in which data bear no structure makes learning altogether impossible; a representation partitioning the data well makes learning trivial. Early machine learning methods made use of hand-crafted features to transform data into a representation well suited for a learning task. However, such feature engineering is time consuming and inefficient.

Research efforts over the last decades produced multiple techniques that automatically extract useful features from data. Some operate exclusively on the basis of the

statistical structure of the input, for instance learning decorrelated or linearly independent features; others rely on supervision signals to learn which data attributes are discriminative for a specific task.

The perceptron, a computational unit inspired from biological neurons, was probably the first machine learning method invented. As other learning algorithms, the perceptron critically depends on the representation of its input. Incidentally, the dissemination of this observation by Minsky and Papert in the late 1960s prompted the first research drought in artificial intelligence. However, in the mid-1980s, several research groups independently discovered an algorithm to train perceptrons arranged in a layered structure, or multi-layer perceptrons (MLPs). The technique, known as error-backpropagation, allowed MLPs to find discriminative features for a task, learning their own data representation.

This approach soon gave rise to 'deep learning', artificial neural networks in which multiple layers of neurons learn hierarchical, compositional representations for data. Deep learning developed to be hugely successful, claiming the best performance on a wide variety of benchmarks and tackling problems as diverse as brain tumour detection, real-time speech translation, and bankruptcy prediction.

Although very effective, the error-backpropagation method suffers from limitations. First, in order to compute the error function, a target output must be specified for each training example, making training data expensive to acquire. Second, weight updates require information that is not available locally at the weights, such as gradient information from downstream neurons. This requirement limits the implementation of learning for deep networks in physical devices like neuromorphic chips. Additionally, error back-propagation is susceptible to vanishing error gradients whereby neurons far from the output layer learn prohibitively slowly. On this account, a second aim of this work is to examine alternate learning methods for deep neural networks based on the actions of neuromodulators in the brain.

1.4 Contributions of this Work

I address the topic of neuromodulator-inspired representation learning in a neural network model. This network contains rate-based neurons arranged in three feedforward processing layers. Learning follows Hebb's rule supplemented by synaptic scaling, feedforward inhibition and lateral competition, a combination of mechanisms implementing approximately optimal learning [KSL12]. I train the network on a classification task of images of handwritten digits from the MNIST dataset [LCB98]. The middle layer of the network then learns a representation of the data which I consider as a model of sensory representations in the cortex.

I extend this model of representation learning to simulate the effects of the neuromodulators ACh and DA on synaptic plasticity. Here, both ACh and DA modulate the learning rate of the neural network. This approach reproduces the observation that stimuli paired with activation of either the cholinergic or DAergic system occupy a greater area of the cortex. I then simulate the distinct physiological release properties of the neuromodulators. In the case of ACh, activation approximates the cognitive construct of attentional effort. For DA, activation follows reward prediction errors. I examine the influence of these two signals on the network's neural representation and on its classification performance.

This framework allows to formulate hypotheses regarding the functions of ACh and DA in cortical representation learning. Additionally, the model permits the evaluation of neuromodulator-like signals in artificial neural networks for practical purposes. Accordingly, four main contributions are made within this framework:

1. I demonstrate that a permissive plasticity signal whose activation properties approximate the construct of attentional effort improves the quality of a neural representation. The improvements result from altering the distribution of class preferences among neurons, in particular from enhancing the representations of challenging or relevant stimuli. I postulate that ACh serves this function in refining representations in the cortex.

2. I show that modulating synaptic plasticity as a function of reward prediction errors refines neural weights with respect to a classification task. Specifically, after training, weights reflect the boundaries of the conditions for reward delivery of the task, thereby raising classification accuracy. The results suggest that DA carries this role in modifying sensory representations in mammals. Although similar DA-based learning rules were already known, in the present work I derive the update rule not for its function, as done previously, but to emulate the plastic actions of DA on cortical representations. This work thus provides additional biological interpretations for DA-inspired learning rules.

3. I show that noisy neural responses give rise to a softmax exploration rule which, co-jointly with DA, benefits learning. In this context of exploratory learning, the optimal activation profile for DA closely agrees with the one observed in primates. These results suggest that, in mammals, variability in the responses of sensory neurons and DA release together serve a synergetic function in improving cortical representations.

4. In comparison with MLPs of the same architecture, the DA model yields performances greater than those originally reported on the MNIST dataset and rivalling those of state-of-the-art optimisation methods. A unique DA signal additionally proves to be capable of refining multiple representations concurrently in a deep convolutional network. Learning with neuromodulators requires weak supervision signals and synaptically-local information. Together, these characteristics make the learning model attractive in scenarios poorly suited to the error-backpropagation algorithm such as learning with sparse labelling or in physical devices.

1.5 Thesis Outline

The six remaining chapters of this thesis are organised as follow:

- **Chapter 2, Sensory Representations in the Cortex**, reviews background knowledge on the development and plasticity of cortical representations. In particular, this chapter covers the details of the activation properties of ACh and DA in mammals as well as their contributions to cortical plasticity.

- **Chapter 3, Representation Learning in ANNs**, describes various learning rules in artificial neural networks, setting the context for the model used in this work. This neural network model is then introduced along with a justification of its mathematical formulation. The chapter also outlines the basic models of the neuromodulators.
- **Chapter 4, Model of Acetylcholine**, introduces a model of the physiological release properties of ACh and presents the results obtained with this model. This includes various methods for approximating task demand and the assessment of other parameters on learning, like network size, label availability and data distribution.
- **Chapter 5, Model of Dopamine**, presents the results for different models of DA activation, examining the roles of exploration and probabilistic reward predictions in learning. In addition, the chapter reports on a comparison of the learning outcomes for DA and the error-backpropagation algorithms.
- **Chapter 6, Deep Convolutional Model**, presents an extension of the neural network model to add a convolutional layer. This chapter studies the ability of DA signals to orchestrate learning in multiple layers of a deep network concurrently.
- **Chapter 7, Conclusion**.

Chapter 2

Sensory Representations in the Cortex

2.1 Overview

An animal's body contains a myriad of sensory receptors. From cells detecting photons of specific frequencies in the retina to spindles responding to stretches in muscles and hair cells sensing acceleration in the inner ear, sensory neurons pick up on a host of physical stimuli and transduce them in electrical activity. These electrical signals travel down nerves and reach the cortex where they elicit patterns of neural activity. Those cortical activity patterns form the brain's representation of its sensory environment. The particular form a sensory representation takes is critical to many aspects of cognition, most importantly perception. For instance, refinements in representations explain our ability to learn to distinguish similar stimuli, such as when an enologist tells apart wines from their flavours. On the other hand, disruptions of representations underlie various perceptual disorders, for example not recognising a part of one's body as its own.

In this chapter, I present an overview of the current state of knowledge on cortical sensory representations. I first briefly describe the perinatal development of representations with regard to the complementary roles of genes and sensory experience in determining the tuning properties of neurons. I then cover modifications of representations taking place in adult animals as a result of behavioural training, alterations of sensory input, and pharmacological manipulations. Finally, I address the specific changes elicited by the neuromodulators acetylcholine and dopamine.

2.2 Development of Representations

Two opposing hypotheses were initially put forward to explain the development of cortical sensory representations. The *protomap* hypothesis [Rak88] held that a predetermined genetic program specifies the response tunings of neurons. The *protocortex* hypothesis [O'L89], on the other hand, stated that sensory experience shapes an initially clean slate, or *tabula rasa*. These two views, however, intrinsically entangle: no genes act independently of the environment and sensory experience needs a structure to act on. The modern view holds that, rather than following one or the other,

the response properties of neurons emerge through a close interaction between genetic material and sensory experience [SL01; SR05].

2.2.1 Genetic Program

The graded expression of transcription factors in the neural plate determines the regional patterning of the developing cortex [SR05]. These gradients control the proliferation, neurogenesis, migration, and connectivity of neurons. Alterations in the expression of molecular gradients can modify the size [Ham+04] or even duplicate [FSG01] cortical areas and shift the topography of thalamocortical projections [Gar+02]. Genes are also instrumental in guiding projections from the thalamus to their cortical target. In animals lacking the Ebf1 gene, for instance, the somatosensory thalamus projects to the immature visual cortex [GR04]. These findings underline the importance of genetic factors in determining the sensory identity of the cortical mantle.

The extent by which genes determine stimulus preferences is apparent in studies performed in young animals before experience impinges on representations. In the primary visual cortex (V1), fibres from the visual thalamus to layer IV target eye-specific regions before eye opening and independently of spontaneous retinal activity [CK00]. Similarly, orientation selectivity is present at birth in primates [WH74] and to a certain extent in rodents and cats [IB75; CS93]. In line with these findings, the early emergence of orientation selectivity as well as ocular dominance and orientation maps occur even during binocular suture [CGS98; FI84; SS82]. In the auditory system, tonotopic gradients [Har+97] and binaural selectivity [Til+06] are present in adult congenitally deaf cats. These results indicate that sensory experience is not required for the initial development of neural representations and, accordingly, that a genetic program is sufficient to explain basic properties of stimulus preferences.

2.2.2 Experience-dependent Development

Although experience is not necessary for stimulus selectivity to emerge, neural activity does play an central role in maintaining and refining neural tuning. For instance, blockade of spontaneous activity before birth reduces orientation tuning in single cells [CS93] and the clustering of horizontal connection in V1 [RS96]. Additionally, disrupting the normal waves of spontaneous activity in the retina results in imprecise retinotopic maps in the visual cortex [Can+05]. These observations argue for a role of endogenous activity in complementing genetic programs.

Soon after birth, sensory representations adjust to the statistical structure of an animal's sensory environment during a brief time window termed the critical period [Hen04; Hen05]. In infant animals, artificial strabismus–a misalignment of the eyes which decorrelates activity patterns in the two retinas–results in almost exclusively monocular V1 neurons [VSL80; CVN79] and in more sharply delineated ocular dominance columns [SLW+77]. Rearing cats in an environment containing a single orientation leads to a twice larger cortical area devoted to this orientation [SSB99]. Similarly, passively exposing rats to pure tones shortly after birth leads an over-representation of sites in the primary auditory cortex (A1) tuned to the presented tone [BPH11; VS+07] which affects the animals' perceptual sensitivity [Han+07]. These findings indicate that the statistical structure of sensory signals is instructive in shaping stimulus selectivity,

overriding the organisation initially laid out by genetic factors.

Studies in which sensory systems are surgically rewired are particularly telling on the relative roles of nature and nurture in establishing neural representations. For instance, in animals in which the optical nerve is rewired to project to the primary auditory cortex, axons in A1 segregate into eye-specific subregions like in V1 [Ang+97] and an orderly, 2-dimensional retinotopic map develops [Roe+90] with pin-wheel patterns [SAS00]. Additionally, rewired A1 neurons have orientation tuning properties that are quantitatively indistinguishable from V1 cells [Roe+92]. However, the rewired A1 retains connectivity patterns from the normal auditory cortex and orientation maps remain somewhat deficient compared to the normal visual cortex. Together, these results indicate that, if genes dictate the initial organisation of the cortex, patterned sensory information is also sufficient to give rise to various aspects of adult neural representations.

2.3 Modifications in Sensory Representations

After the end of the critical period, sensory representations lose much of their capacity to adapt to environmental changes. This reduction in plastic potential is thought to ensure the stability of perceptual systems and protect against detrimental changes. Nonetheless, representations can undergo modifications in specific circumstances in adulthood, such as following extensive behavioural training [Jen+90], as a result of altered sensory experience [PKF04], or through electric or pharmacologic stimulation of the cortex [Din+03; Gu02].

Neural representations are critical to perception and their modification carry both beneficial and harmful consequences. As an example of favourable effects, in humans, activation of a finger tip paired with transcranial neural stimulation expands the cortical representation of the finger tip and improves tactile discriminatory performance [Din+03]. Similarly, stimulation of the nucleus basalis concurrently with presentation of an auditory tone enlarges the tone's representation, in turn resulting in enhanced perceptual abilities [Ree+11]. Following discrimination training, improvements in perceptual sensitivity correlate with the degree of reorganisation in cortical representations [Rec+92; RSM93; PSM06; WFS04; Wei03; Din+03], suggesting that plasticity of neural representations explains the perceptual benefits. On the other hand, perceptual disorders like phantom limb [RSRR92; Hal+93; FNJ06], an often painful sensation in an amputated limb, or tinnitus [ER04], a ringing in the ear, appear to be correlates of degenerate sensory representations. These observations indicate that alterations in cortical representations bear major influences on perception, both meliorative and detrimental.

Here, I review experimental observations describing scenarios in which cortical representations undergo changes in adult animals.

2.3.1 Sensory Deprivation

After deafferation or lesion of sensory organs, initially silent cortical cells modify their stimulus selectivity so as to recover an input. For example, inflicting a lesion on the retina leads to an immediate loss of responsiveness in the cortical area innervated by the lesion. However, neurons just inside this area rapidly regain responses by shifting

their receptive fields to positions just outside the lesion [GE06]. Over the following weeks, visual responsiveness progresses inwards towards the centre of the affected zone. This reorganisation appears to be cortically-based: even after the denervated cortical area regained an input, a blind region remains in the thalamus [DSG95]. Similar adjustments take place throughout the cortex, such as in the somatosensory [CT88; Hal+93] or auditory cortices [KE07]. Here, plastic reorganisations of sensory maps ensure maximal usage of the cortex's processing potential.

2.3.2 Altered Sensory Experience

If sensory deprivation triggers cortical reorganisations, more modest manipulations of sensory experience can also do so. For instance, transferring adult rats from a normal laboratory cage to an enriched environment leads to alterations in the animals' somatosensory maps [PKF04]. In particular, whisker representations contract by about one half and receptive fields of barrel cortex neurons sharpen, yielding refined stimulus selectivity. Similarly, housing adult rats in enriched cages reactivates ocular dominance plasticity [Bar+10]. These findings indicate that enhanced sensory experience can boost cortical plasticity and refine perceptual systems. Conversely, however, over-stimulation of a whisker causes the representation of this whisker to shrink and weaken [Wel+92], indicating that enhanced stimulation may also be detrimental.

2.3.3 Perceptual Learning

Animal perception sharpens through training, a process termed perceptual learning [SD07; SW05; SNW12]. Natural examples of perceptual learning include greater tactile sensitivity in pianists [Rag+04] and improved ability to detect beer flavours in subjects experienced in the matter [PA88]. In laboratory settings, this form of learning is observed in all senses such as in the discrimination of auditory frequencies [RSM93; PSM06], visual orientations [YM04; Sch+01], tactile pressures [HHD01], odourants [FW02; WFS04], or flavours [OM79].

Except for some particular cases [WNS01; SW03; CSW09; TSW08; BS12], perceptual improvements are normally limited to attended, task-related stimulus features as well as their specific spatial location [AH93; AH93; PSM06; Sch+01]. For instance, practice at discriminating the orientation of gratings leaves sensitivity unchanged for orientations or retinal locations other than the trained one [YM04] and, likewise, improvements in tactile perception for a trained finger only partly transfer to other fingers [HHD01]. Such location and stimulus specificity suggests that perceptual learning takes place in sensory areas where distinct representations of trained stimuli are present [Wat+02].

Perceptual learning almost invariably alters cortical representations. For example, neurons in the primary auditory cortex of rats trained to discriminate the temporal repetition rate of auditory stimuli respond more strongly to task-relevant stimuli than neurons in control animals [Bao+04]. Practice on an orientation discrimination task results in neurons with narrower tuning bandwidth and greater amplitude [YM04]. In humans, learning Mandarin tones leads to an expansion of the cortical area activated by these tones [Wan+03]. Likewise, stimulation of the fingertips used for of Braille reading in blind patients evokes potentials over a greater area of the somatosensory

cortex then for the fingertips not used for reading [PLT93]. Importantly, improvements in perceptual sensitivity correlate with the degree of cortical reorganisation [Rec+92; RSM93; PSM06; WFS04; Wei03; Din+03]. In combination, these results convey that modifications in neural representations improve the processing of sensory information following extensive perceptual training.

2.3.4 Classical Conditioning

In classical conditioning, an animal learns to respond to a neutral or conditioned stimulus (e.g., a bell ring) that repeatedly accompanies a biologically potent or unconditioned stimulus (e.g., food). Such training alters the cortical representation of the conditioned stimulus. For instance, pairing an auditory tone with electric foot stimulation leads to rapid increases in the responses of neuron to the paired tone [Fri+03] and to shifts in the best frequency of neurons towards this tone [MS05]. The enhanced representation of a conditioned stimulus is thought to form a physiological memory trace that raises the likelihood of perceiving the learned stimulus.

2.3.5 Cortical Stimulation

In humans, repetitive transcranial magnetic stimulation over the cortical area receiving afferents from the index finger enlarges this area and lowers the tactile discrimination threshold for this index [Teg+05; Din+03]. In a similar fashion, direct electric stimulation of the bat's auditory cortex triggers a reorganisation of the animal's frequency map. In this case, cells spatially close to the stimulated neuron shift their best frequencies towards that of the target neuron while the opposite takes place in the farther surrounding [MS04]. In the cat visual cortex, intracortical micro-stimulation shifts receptive fields towards the preferred orientation of the stimulated cell, resulting in a magnification of the representation of the orientation [God+02]. Although artificial, these manipulations demonstrate that neural activity in the adult cortex may suffice to reshape how the brain represents sensory information.

2.3.6 Neuromodulation

Neuromodulators, like neurotransmitters, are organic molecules that trigger electrochemical responses in neurons. In contrast with neurotransmitters, however, modulators may be let out in large volumes of the brain at once, known as volume transmission, and remain in the cerebrospinal fluid after their release. Neuromodulators are thus typically effective on a broader temporal and spatial scale. Furthermore, while transmitters bind directly to ion channels and trigger rapid intracellular ion influx, modulators usually contact metabotropic receptors which conduct their actions through second messengers. This process is then slower and longer-lasting, "modulating" neuronal spiking rather than causing it.

Several neuromodulators influence cortical plasticity and, as a consequence, alter sensory representations. Here, I briefly review findings relative to two modulatory agents, noradrenaline (NE) and serotonin (5-HT). In the following section, I cover in more details two other neuromodulators, acetylcholine and dopamine, central to this thesis. This list is not exhaustive and other agents, for example histamine [Bek93] and norepinephrine [He+15], may also take part in shaping sensory representations.

Noradrenaline

NE exerts both excitatory and inhibitory effects on cortical neurons. Through this dual influence, NE reduces spontaneous activity more than evoked activity and is thus portrayed as amplifying the brain's signal-to-noise ratio [WW80; KR89].

In addition to sensory signal processing, NE bears a potent action on cortical plasticity. For example, NE depletion prevents plasticity in the kitten visual cortex [KP76] and adult somatosensory cortex [LCH88] following sensory deprivation. Conversely, NE infusion enhances ocular dominance plasticity [KK84] even in the adult cortex [HIK87] and promotes experience-dependent modification of receptive fields in V1 [GLS88]. These observations are evidence for a role of NE in permitting cortical reorganisations in both young and adult animals.

Serotonin

The serotonergic system is best known for contributing to feelings of well-being and, through its disorders, for its involvement in clinical depression. 5-HT additionally modulates synaptic plasticity and affects cortical representations. In young animals, 5-HT depletion diminishes the size of whisker-related patches in the somatosensory cortex [BC+94] while 5-HT excess results in a lack of patterned barrels in the somatosensory cortex [Cas+96]. Additionally, 5-HT receptor blockade in visual areas prevents ocular dominance plasticity following monocular deprivation [GS95; WGC97]. In adult animals, 5-HT promotes excitatory responses [RBC88; NEF87] as well as synaptic modifications [He+15] in neocortical cells. Together, these results suggest that 5-HT acts as a permissive plasticity agents which, at least in young animals, elicits alterations in cortical maps.

2.4 Acetylcholine

ACh is an organic molecule acting as a neurotransmitter at synaptic contacts throughout the mammalian nervous system. In the central nervous system, ACh also escapes synaptic contacts and stimulates extra-synaptic receptors [Fux+05]. ACh then carries its action through both synaptic and volume transmission [SPH09]. In line with this dual transmission mode, ACh operates on multiple timescales, triggering both second-scale and longer-lasting transients [Par+07].

In mammals, a few groups of neurons in the basal forebrain secrete ACh and send projections to the amygdala, the thalamus, the hippocampus, and the neocortex [Kit+94]. One of these nuclei, the nucleus basalis of Meynert (NB), provides the single major source of ACh in the cortex [Mes+83; Sem00], with its projections innervating all cortical areas [Mes+92; Kit+94; Záb92]. It remains debated whether these projections are diffuse or contact specific cortical targets [SPH09] but, in any case, ACh release takes place in a modality- and region-specific manner [FSR04].

Although we lack a definite theory describing ACh transmission in the cortex, attentional processes appear to be important contributors. In particular, the attentional demand of a stimulus or of a task exerts clear influences on the levels of ACh in the cortex.

2.4.1 ACh and Attentional Demand

A series of experiments demonstrate the influence of attentional processes on cortical ACh efflux. In these experiments, animals have to press either of two levers to signal the presence or absence of a short-duration visual cue, a task known to generate measures of sustained attention [MS95]. As the animals initiate the task, prefrontal ACh concentration rises to more than double its pre-task levels. The presence of ACh then remains high throughout the experiment and undergoes a protracted decay after termination of the task [Arn+02; HSB00; KBS06]. This effect is absent in control procedures with negligible attentional demand but otherwise similar motor requirements and reinforcer delivery [Arn+02]. Additionally, visual distractors [HSB00] or pharmacological manipulations [KBS06] that impair performance and increase attentional demand lead to further increases in ACh concentration. In the case of visual distractors, the additional rise in ACh parallels a recovery in performance. Together, these results indicate that cortical cholinergic transmission reflects effortful attentional control required to maintain performance under challenging conditions.

In addition to long-lasting modulations in ACh concentration, short duration ACh spikes also occur. In rats, the detection of a sensory cue predictive of a reward elicits a spur in prefrontal ACh. This event is accompanied by an attentional shift towards the reward ports [Par+07]. The timing of the ACh spike coincides with the onset of the animal's behavioural shift while the signal's amplitude correlates with the latency between cue presentation and reward retrieval. Missed cues or cues at the beginning of the training as well as reward delivery and retrieval do not elicit ACh transients. However, catch trials in which a cue is present but no reward is delivered still evoke ACh efflux. Together, these findings indicate that sensory events that bear a behavioural relevance and require an animal's attention elicit ACh efflux, with the amplitude of the signal denoting the motivational value of the event. On the other hand, rewards alone or reward prediction errors appear not to influence ACh release.

Activation of the cholinergic system thus appear to reflect attentional demand. This demand can derive either from the challenges an event poses or from its behavioural relevance. This account of ACh transmission relates to the cognitive construct of 'attentional effort' [SGK06]. Here, attentional effort describes a subject's motivated attempts to maintain performance under challenging circumstances, reflecting both subject motivation and task difficulty. For instance, the attentional performances of unpaid subjects drops during an hour-long experiment whereas that of paid subjects remains constant [TT96], indicating a role of motivational factors in maintaining attention. This dual aspect of attentional effort closely corresponds to the available evidence regarding the release properties of ACh. Attentional effort may thus be an accurate characterisation of the circumstances leading to cholinergic transmission.

2.4.2 Plastic Effects on Sensory Representations

In the cortex, ACh modulates the responses of neurons [Zin+06], facilitates the induction of long-term potentiation [BAS92; Chu+13], and is necessary for various forms of learning [Win+95; Eas+02]. Additionally, ACh contributes to modifications of cortical representations in a broad range of circumstances, including when paired with a sensory stimulus [KM98a; RD88; Web+91b; KM98b; GLS88; FMS07], during conditioning [FMS07; JGS01], following sensory deprivation [Sac+98; BSH97; BS86; ZW98;

JME91; Web+91a; Lea+13], and during task learning [Con+03; CKT10].

Pairing Experiments

Experiments in which a stimulus repeatedly coincides with cortical ACh release are perhaps the most telling in terms of the effects of ACh on representations. As an example, episodic pairing of an auditory tone with electrical activation of NB leads to a nearly doubled cortical representation of the tone [KM98a] (see Fig. 2.1). Similarly, coincident tactile and NB stimulation results in up to three times larger stimulus-evoked potentials in the somatosensory cortex [RD88; Web+91b]. Pairing high-repetition frequency tones with NB activation increases the ability of cortical neurons to fire in synchrony with high frequency stimuli [KM98b]. In V1 neurons, pairing ACh and NE with stimulation of the eye less effective in driving a neuron or with presentation of a neuron's suboptimal orientation results in shifts in the tuning of the neuron towards the paired stimulus [GLS88]. In all these cases, stimulation of NB or of sensory receptors separately or decorrelated stimulation of the two do no alter cortical representations.

These shifts in responses appear to take place through a transient disinhibition of principal excitatory cells mediated by ACh [FMS07]. In physiological conditions, excitation and inhibition are matched across a neuron's receptive field. Pairing a stimulus with NB activation results in a rapid reduction in synaptic inhibition for the paired stimulus soon followed by a large increase in excitation for this stimulus. Inhibition then slowly recovers to balance excitation. Jointly, these modifications yield a receptive field with a new preferred tuning at the target stimulus.

Conditioning

During fear conditioning of an auditory stimulus, cells in the primary auditory cortex increase their responses to the conditioned stimulus. ACh applied on the auditory cortex during this procedure augments the shift in tuning while the ACh antagonist atropine abolishes it [JGS01]. In mammals, fear conditioning appears to require a disinhibition of pyramidal neurons triggered by ACh [Let+11]. These results point to a role of the cholinergic system in permitting alterations in neural tuning following conditioning.

Sensory Deprivation

In normal animals, sensory deprivation elicits synaptic reorganisations such that cortical neurons progressively regain responses to sensory stimulation. ACh appears to be required for this process to take place. For instance, destruction of both cholinergic and noradrenergic nuclei [BS86] or chronic blockade of ACh receptors [GS93] prevents plasticity of ocular dominance maps following monocular deprivation. After whisker trimming [Sac+98; BSH97], deafferation a single whisker field [ZW98] or finger amputation [JME91; Web+91a], ACh depletion precludes the normal expansion of the representations of the intact whiskers or fingers into the deprived cortex. Removal of NB impairs functional recovery and cortical reorganisation usually resulting from lesions of the motor cortex [CCT05]. Selective removal of cortex-projecting NB neurons prevents ferrets from re-weighting auditory localisation cues after occlusion of one

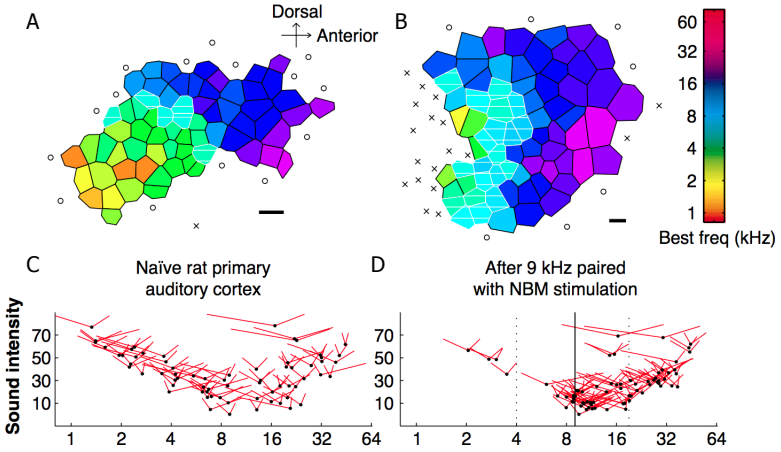

Figure 2.1: Pairing a tone with activation of the cholinergic system enlarges its representation in the rat auditory cortex. (A and B) Representative tonotopic maps in the rat primary auditory cortex (A) before and (B) after pairing a 9 kHz tone with NB stimulation. Each polygon represents an electrode penetration with the colour indicating the best frequencies of the recordings. Hatched areas have their best frequencies within 0.3 octave of 9 kHz. Scale bar is 200 µm. (C and D) Distribution of tuning curves for each penetration. The plot indicates the best frequency at the threshold (black dots) and the width of the tuning curve 10 dB above the threshold (red traces). The solid vertical line indicates the paired frequency; the dotted lines frequencies presented as often but not paired with NB stimulation. Reproduced with permission from [KM98a].

ear [Lea+13]. In combination, these observations indicate that cortical reorganisations resulting from sensory deprivation require cholinergic signalling.

Motor Learning

Homologously to sensory areas, the motor cortex contains representations of an animal's motor efferents. These representations are modified during acquisition of motor tasks such that more cortical resources are allocated to the motor commands involved in the task. Disrupting basal forebrain cholinergic function impairs both this reorganisation and motor learning [Con+03; CKT10], suggesting a causal link between ACh signalling, map plasticity, and learning.

Taken together, the findings presented in this section argue for a permissive role of ACh in various learning processes involving modifications in sensory representations.

2.5 Dopamine

Dopamine is an organic compound functioning as a neuromodulator in the mammalian brain. It is synthesised in nine cell groups which form distinct pathways [BD07] involved in functions as diverse as locomotion, motivation, and executive control [Sch07b]. Notably, DA-producing neurons in the ventral tegmental area (VTA) play a major role in reward processing and reinforcement learning [Sch07a; Sch10]. In this work, I limit the study of DA to its involvement in the reward pathway.

2.5.1 DA and Reward Processing

Two sets of experiments summarise our current understanding of the contribution of DA to reward processing. A first set comes from treating animals with neuroleptics drugs. These drugs act as DA antagonist and are known to blunt the effects of rewards. For instance, DA receptor blocker pimozide prevents hungry rats from learning to press a lever for food [WS81] or to perform lever presses if the task was already learned [WSG+78]. In this latter case, the action of DA agonists are similar to omitting food rewards altogether. Neuroleptics similarly eliminate reward-approaching behaviours regardless of the reward type, including hypothalamic electric stimulation [Gal+82], cocaine injection [DWW77], and water [GSW81]. This collection of results formed the first indication that DA signals rewarding events and triggers approaches towards rewards.

A second series of experiments was informed by observations and theories of reinforcement learning. At the time of the design of the experiments, it was known that a stimulus consistently predictive of a reward progressively gains appetitive characteristics itself, as in classical conditioning. Furthermore, such conditioning was known to depend on an unpredicted reward as learning does not take place if a cue fully predicts a reward. On account of these observations, it was postulated that a signal indicating the difference between a received and anticipated reward, or reward prediction error (RPE), could drive learning. Here, an unpredicted reinforcer would give rise to a positive RPE which, in turn, would allow learning. On the other hand, a fully anticipated reinforcer would not prompt a RPE, thereby blocking learning.

Schultz and colleagues [SDM97] famously showed that DA neurons in the VTA encode such error signals. In this study, monkeys were presented a visual and auditory cue that consistently announced a reward. Recordings in the VTA of the monkeys showed that neurons initially respond to the reward but progressively shift their responses to the predictive cue. Once the cue fully predicts the reward, firing of DA neurons to reward delivery ceases altogether. At this point, coupling a stimulus with the anticipated reward does not lead to learning of the stimulus. Furthermore, learning of a stimulus takes place primarily when DA neurons fire [WDS01], suggesting that the DA signal is responsible for conditioning. The firing of DA neurons thus obeys basic requirements of reinforcement learning theory.

Further results soon complemented the description of the features of rewards encoded by VTA neurons. For instance, in addition to RPEs, DAergic activation also signals the expected value of the reward distribution as well as uncertainty in the reward probability [FTS03; TFS05]. Conversely, aversive stimuli depress the responses of DAergic neurons, indicating that these neurons may encode 'negative' reward magnitudes [UMB04]. DA firing also encodes the expected timing of rewards [Nak+04]. VTA neurons additionally respond to salient or novel but not explicitly rewarding stimuli [Hor00]. In this case, other attention-inducing stimuli do not trigger such responses suggesting that DA activation relates not to attention but rather to the rewarding and approach-generating characteristics of novel and salient stimuli. Combined, these observations provide a description of the various characteristics of rewards represented in the firing of VTA neurons, of which RPEs appear the most prominent.

2.5.2 Plastic Effects on Sensory Representations

The VTA sends diffuse projections throughout the cortex [GSK92; HK10] and DA receptors are expressed in all cortical areas [Lid+91]. Correspondingly, reward delivery modulates neural responses in nearly all subdivisions of the brain [VCL11], also specifically in sensory areas like the visual [Ars+13], somatosensory [Ple+09], and auditory cortices [BSS11].

DA directly modulates synaptic plasticity in the prefrontal cortex [BCO02; MMO06; SZW05; MMO06] and, in sensory cortices, DA participates in plastic modifications of neural tuning [BCM01; Bao+03; FSV10; Poo+15]. The involvement of DA in cortical plasticity is evident in its contribution to various forms of learning, for instance rule [OJN14], sequence [Gra05], motor [Hos+11; LS09; ML+09], error-based [Kle+07], and sensory discrimination learning [KS06; Sch+12].

Of particular relevance for this work is the profound impact DA exerts on sensory representations. Similarly to the cholinergic system, repetitive activation of the DAergic system shortly after presentation of an auditory tone expands the area of A1 responsive to the tone (see Fig. 2.2). In contrast with ACh, however, the temporal order of sensory and DAergic activation is influential: here, reversing the presentation sequence such that DA precedes the auditory stimulus reduces the representation of the tone in the cortex [Bao+03].

Coupling stimuli with rewards bears a similar effect to pairing with direct VTA activation. In this case, the manipulation strengthens stimulus-evoked potentials [FSV10] and enhances stimulus representations [Poo+15; Sch+01; YM04; Rai+06]. The effects of reinforcers on perceptual systems are also evident behaviourally. In humans, learning on a visual discrimination task saturates in the absence of performance feedback. After this point, providing reinforcer following correct decisions leads to further improvements in sensitivity [Sei+06], putatively through a fine-tuning of the relevant sensory representations. Combined, these results describe a role of DAergic transmission in cortical plasticity and, specifically, in strengthening the representation of stimuli with which its release co-occurs.

2.6 Summary

Through molecular gradients, genes guide the growth of cortical connections in embryo, outlining crude sensory representations. Shortly after birth, the response properties of neurons are particularly malleable and sensory experience then refines the synaptic wiring established by genetic factors. After the end of this critical period, representations are less prone change. Nevertheless, several paradigms permit alterations of tuning properties in the adult cortex, including sensory deprivation, behavioural training, and pharmacological manipulations. In particular, the neuromodulators ACh and DA promote cortical plasticity and evoke modifications in the stimulus preferences of neurons. Release of either neuromodulator coincident with a stimulus induces neurons to shift their responses towards the stimulus. At a population level, this manipulation results in an enlargement in the cortical area devoted to the paired stimulus. In physiological conditions, ACh transmission appears to signal attentional effort, a construct reflecting both behavioural relevance and task difficulty. DA carries information relative to errors in reward predictions. Although their effects on cortical plasticity

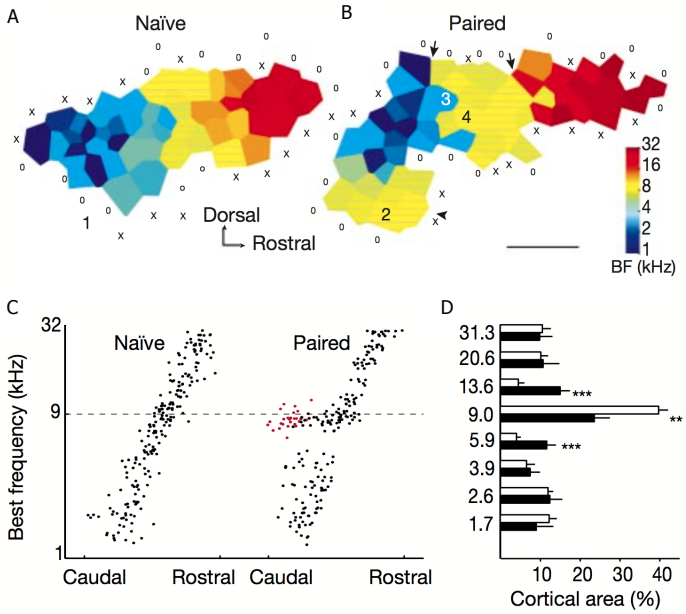

Figure 2.2: Pairing a tone with activation of the DAergic system enlarges its representation in the rat auditory cortex. (A) Representative tonotopic best frequency map of a naive animal. (B) Tonotopic map in an animal in which a 9 kHz tone repeatedly coincided with electric stimulation of the VTA. Hatched areas have their best frequencies within 0.3 octave of 9 kHz. Scale bar is 500 μm. In the experimental animal, the representation of 9 kHz tone expands while those of neighbouring frequencies shrink. (C) Distribution of preferred frequencies along the caudal-rostral axis of A1. (D) Percent of the auditory cortex tunes to each frequency. Black bar: naïve, white bat: paired. The pairing protocol significantly increases the cortical area devoted to 9 kHz. Double asterisk: $p < 0.005$; triple asterisk: $p < 0.001$. Reproduced with permission from [BCM01].

are relatively well defined, the functional roles these agents serve in altering neural representations remain unclear.

Chapter 3

Representation Learning in ANNs

3.1 Overview

In mammals, stimuli evoke responses in cells of the neocortex, forming a neural representation of sensory signals. From its primary target site in the cortex, activity propagates along a sensory pathway. At each discrete processing stage, neurons respond to distinct attributes of sensory signals. This progressive transformation of sensory signals into more elaborate neural representations is understood to form the basis of mammalian perception.

Artificial neural systems make use of similar computational principles. In particular, connection weights between two layers of artificial neurons transform activity in the input layer to a 'better' representation in the subsequent layer. This second representation should be better suited to the particular task the network faces, such as a classification or regression task. Deep learning extends this concept to a hierarchy of layers that progressively transform data to achieve a high-level, abstract representation in the ultimate layer.

In artificial neural networks (ANNs), the connection matrix between layers fully defines the data transformation and their representation in the following layer. The rules governing learning of the weight matrix are thus central to a network's representations. These rules can be thought of as the computational equivalent of experience-dependent learning in biology. As such, representation learning in ANNs may serve as a model for its biological counterpart. In this thesis, I make use of this correspondence to investigate the functions of ACh and DA in shaping representations. This modelling effort serves both to better understand the roles of these modulators in representation learning and to research into novel learning rules for ANNs.

This chapter is devoted to models of representation learning in ANNs. The text is organised in three sections. First, I conduct a short review of existing learning rules in ANNs. I then introduce the basic network model used in this thesis, including its theoretical foundation. In the third section, I present an extension of the model to simulate the influences of ACh and DA on plasticity.

3.2 Learning Rules in Neural Networks

In ANNs, at least three factors jointly shape a neural representation: the data, the network's architecture, and the learning rule. This section covers the learning aspect,

reviewing the most common weight update rules in ANNs. This review is not exhaustive but aims at describing the context surrounding the model employed in this work. For brevity, I limit the review to two families of learning rules, Hebbian learning and error-correcting rules, the latter of which includes reward-based learning. These rules each belong to a distinct class of learning methods, namely unsupervised, supervised, and reinforcement learning, respectively. Notable biology-related topics omitted from this review include time-dependent learning rules and aspects of sparse coding. I am indebted to the following resources for writing this section: [DA01; Sch15; GK02; Kur15; LBH15].

3.2.1 Hebbian Learning

In 1949, the psychologist Donald Hebb conceived a simple rule describing changes in the synaptic strength between two abstract neurons. Hebb's rule stated that if neuron A repeatedly contributes to the activation of neuron B, the synaptic connection from neuron A to neuron B should strengthen [Heb49]. Hebb's rule then recounts how patterned neural activity, and in particular causal relationships in the activation of neurons, shapes synaptic connections. This simple formulation is one of the most widely recognised approaches to learning in neural networks and constitutes the basis for a family of learning rules.

Basic Hebb's rule and Correlation-based Learning

The simplest mathematical formulation of Hebb's rule, for a single post-synaptic neuron, is given as:

$$\Delta W_d = \epsilon(s y_d)\,, \tag{3.1}$$

where y_d is the activation of pre-synaptic neuron d, s is the activation of the post-synaptic neuron, W_d is an element of the weight vector $\vec{W}$ connecting the pre- and post-synaptic neurons, and ϵ is the learning rate, a constant. The most straightforward approach to compute the post-synaptic activation s is to take the weighted sum of the neuron's input:

$$s = \sum_{d=1}^{D} W_d y_d\,, \tag{3.2}$$

where D is the number of pre-synaptic neurons, or of input dimensions. Hebbian learning is local in that it requires solely the activations of pre- and post-synaptic neurons and unsupervised as it involves only data, without additional environmental feedback.

The outcome of Hebbian learning can be understood as follows. First, consider the results of learning averaged over a whole dataset:

$$\Delta W_d = \epsilon\langle s y_d\rangle_N\,, \tag{3.3}$$

where $\langle\cdot\rangle_N$ indicates the average over all N data points of a training set. Substituting s by Eq. 3.2, we obtain:

$$\Delta W_d = \epsilon \sum_{d'=1}^{D} \langle y_d y_{d'} \rangle_N W_d \,. \tag{3.4}$$

Here, $\langle y_d y_{d'} \rangle_N = Q_{dd'}$, where $\mathbf{Q}$ is the input correlation matrix. Because of this correlation matrix, Eq. 3.4 is called a correlation-based plasticity rule.

It can be shown that Hebbian learning results in a weight vector that is parallel to the principal eigenvector of the input correlation matrix $\mathbf{Q}$. In other words, the weight vector $\vec{W}$ points in the direction of the input space in which the training data have the greatest variance. This weight vector is optimal if the data is to be represented by the activation of a single neuron s. From an information theoretic viewpoint, this projection maximises the amount of information the weight carries about the data. Hebb's straightforward rule thus produces a theoretically-sound neural representation from correlation patterns in data.

In the basic Hebb's rule, the variables y_d and s represent neural activations and are therefore always non-negative. This leads to ΔW_c also being always non-negative, producing a continuous growth of the weights. Variants of Hebbian learning make use of synaptic normalisation to prevent such unbounded growth, for instance in Oja's rule.

Oja's Rule

Erkki Oja [Oja82] introduced a synaptic normalisation term to Hebb's rule:

$$\Delta W_d = \epsilon(s y_d - s^2 W_d) \,. \tag{3.5}$$

This normalisation is said to be multiplicative because it is proportional to the weight vector $\vec{W}$. We can prove the stability of Oja's rule by considering the square length of the weight vector, given by: $|\vec{W}|^2 = 2s^2(1 - |\vec{W}|^2)$. From this equation we note that, as training progresses, the norm of the weight vector converges to a value of 1, thus preventing unbounded growth.

As with Hebb's rule, the final weight vector here is parallel to the principal eigenvector of the input correlation matrix, only normalised to 1. Oja's rule is also said to be local as it solely requires information available at the synapse, namely the activation of the pre- and post-synaptic neurons and the strength of the synapse.

Multiple Post-synaptic Neurons and Lateral Interactions

Hebbian learning can be extended to multiple post-synaptic neurons s_c. However, by itself, Hebb's rule yields identical weight vectors in all post-synaptic neurons, each parallel to the principal eigenvector of $\mathbf{Q}$. The information carried by the neurons is then perfectly redundant. A common solution to this problem is to include lateral interactions between neurons. For instance, unspecific inhibition (e.g., the softmax function) or fixed or plastic recurrent inhibitory connections decorrelate neural activations and ensure distinct weight vectors are learned.

In the case of Oja's rule, it can be shown that anti-Hebbian synaptic plasticity in lateral connections results in post-synaptic weight vectors that are the different

eigenvectors of the input correlation matrix $\mathbf{Q}$. Here, Hebbian learning can thus be interpreted as performing principal component analysis (PCA) of the input data.

Self-organising maps are a further extension of Hebbian learning to multiple post-synaptic neurons where lateral interactions include both cooperation and competition. These interactions are instantiated by short-range excitatory and long-range inhibitory connections, respectively. The connections result in close-by neurons having similar stimulus preferences, with preferences varying smoothly across the map, much like topographic representations in the cortex. In particular, when such networks receive features of visual receptive fields as input, they develop several key characteristics of the maps found in the primate visual cortex, including singularities ('pinwheels') located in monocular regions and iso-orientation contours running perpendicular to ocular dominance bands [OBS92; OSB92].

Although theoretically-sound and in line with several biological observations, Hebb's rule bears a weakness in that it does not learn from errors in its output. Indeed, learning is solely guided by correlations in neural activity. Error signals carry valuable information that, in animals, is beneficial for learning [Sei+06]. In the following section, I describe learning rules that are capable of making use of such signals through error correction.

3.2.2 Error-correcting Rules

Error-correcting rules aim at minimising an error function, often defined as the difference between a target output and the actual output. The perceptron, introduced by Frank Rosenblatt in 1957 [Ros57; Ros58], was the first artificial neuron to make use of such errors.

The Perceptron

As a psychologist, Rosenblatt initially conceived the perceptron as a mathematical model of biological neurons. Like Hebb's neurons, the perceptron computes its activation through a weighted sum of its input. This sum then feeds through a step activation function:

$$s = g(\sum_{d=1}^{D} W_d y_d)\,, \tag{3.6}$$

$$\text{where} \quad g(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{otherwise}\,. \end{cases}$$

Unlike Hebb's rule, however, the perceptron learns not from correlations in neural activity but from errors in its output. Specifically, the update rule is given by:

$$\Delta W_d = \epsilon(t - s) y_d\,. \tag{3.7}$$

where t is the network's target output.

The step activation function models the binary spikes of nerve cells and allows the perceptron to perform binary classification. For instance, built as specialised hardware, the perceptron could classify small images of simple shapes, which was probably the first form of machine learning. Notably, the perceptron could also carry out logical

operations like OR/AND/NOT. At the time, this generated significant enthusiasm as it was then believed that the ability to perform logical operations would essentially solve artificial intelligence. The perceptron was envisioned to soon be able to "walk, talk, see, write, reproduce itself and be conscious of its existence" [Nyt; Ola96].

This initial enthusiasm proved unjustified as several limitations plagued the perceptron, famously its inability to solve non-linearly separable problems like the exclusive OR (XOR). Another shortcoming of the perceptron is its non-differentiable activation function. A simple modification to make the activation function differentiable allows to minimise the error function using calculus, through error gradient descent, as was first done in the Adaptive Linear Neuron (ADALINE).

ADALINE and the Rescorla-Wagner Rule

Soon after the development of the perceptron, Widrow and Hoff [WH+60] realised that making the neuron's activation function linear, i.e., $g(x) = x$, permits finding the optimal weights through calculus. Here, the gradient of the error function with respect to the weights is used to adjust the weights so as to minimise the error:

$$\Delta W_d = -\epsilon \frac{\partial E}{\partial W_d}\,, \tag{3.8}$$

where $E = (t - s)$. In this simple form, the ADALINE is hindered by similar flaws as is the perceptron. This rule, however, has a few other interesting extensions.

First, Rescorla and Wagner [RW+72] showed that the ADALINE rule could be interpreted as a model of reinforcement learning. Here, the Rescorla-Wagner (RW) rule takes the same form as the ADALINE but with different meanings attached of the variables. In this case, t represents the magnitude of a delivered reward and s the network's prediction of the reward. This rule learns an association between a stimulus and the reward it predicts, emulating classical conditioning.

This simple formulation explained several characteristics of conditioning, for instance blocking and over-shadowing [MBG95], and predicted others like over-expectation [KM96]. Additionally, the RW rule highlighted the role of reward prediction errors (the difference $t - s$) in learning, an error signal later shown to be expressed by DAergic neurons of the VTA.

A further important extension of the ADALINE is its generalisation to other differentiable activation functions, a learning approach then known as the delta rule.

The Delta Rule and Error-backpropagation

The delta rule generalises ADALINE to compute the derivative of the error for any differentiable activation function $g(\cdot)$. The update rule is then given by:

$$\Delta W_d = \epsilon(t - s)g'(\sum_{d=1}^{D} W_d y_d)y_d\,. \tag{3.9}$$

The activation function can be chosen to saturate to binary values (e.g., the logistic function) to perform classification or as the identity function to carry out regression. In addition, non-linear activation functions are required if multiple layers of neurons are to be stacked in a network–without a non-linear activation function, a multi-layer

network always equally reformulates to a single-layered linear network. As hinted above, networks with a single layer cannot solve non-linearly separable problems like the exclusive OR. These problems require at least two layers to first transform the input in a linearly-separable representation before performing classification.

If the delta rule allows the use of non-linear activation functions required in multi-layer networks, it does not in itself specify learning for the intermediate layers. The back-propagation algorithm, probably first suggested by Werbos in 1974 [Wer74; Sch15], uses the derivative chain rule to iteratively compute (i.e., 'back-propagate') the error gradient of the delta rule from the output layer to intermediate layers. The back-propagation algorithm thus allows to compute the partial derivatives of the error function $\frac{\partial E}{\partial W_{ij}}$ for any weight W_{ij} in the network, a significant breakthrough in neural network research. Seminal work by Rumelhart, Hinton, and Williams [RHW85; RHW86] later contributed to the popularisation of the back-propagation algorithm. 'Deep learning', the use of neural networks with multiple intermediate layers, grew from these ideas. More recently, the discovery that GPUs dramatically speed up the optimisation of the weights [RMN09] allowed to train much bigger networks on a lot more data, significantly contributing in making deep learning one of the most successful machine learning techniques to date.

The hierarchical data representations learned through error back-propagation yield the best classification performances on several benchmarks, like image classification [KSH12] and speech recognition [Hin+12], and find a wide variety of applications, from predicting the activity of pharmaceutical agents [Ma+15] to identifying sub-atomic particles [Cio+12] or segmenting neurons in volumetric brain data [Hel+13]. Neural networks may also learn representations for sequential data. In particular, recurrent neural networks have shown promising results in natural language processing [BDV01]. For instance, learning to predict adjacent words in texts produces a representation that captures both semantic and syntactic regularities of a language [Mik+13b]. A striking illustration of the relevance of such representations is the ability to perform meaningful vectorial operations on words in a representation space: in such a case, the word vector 'queen' is for example shown be the closest to the vector ('king' - 'man' + 'woman') [MYZ13; Mik+13a].

Deep neural networks also bear some similarities to the mammalian brain. In the context of reinforcement learning, training a network to predict the value of state-action pairs produces a representation that is discriminative of the reward contingencies of the task [Mni+13; Mni+15], reminiscent of the results of reward-based learning in monkeys [SL02]. A modified architecture of deep networks, the convolutional neural network (ConvNet) [FM82; LeC+98], efficiently learns from data organised in multiple arrays, like images. The organisation of representations in a ConvNet resembles that of the ventral visual pathway in mammals and, when the same images are shown to a ConvNet and a monkey, activations of high-level units in the ConvNet predict those of cells in the monkey's inferotemporal cortex [Cad+14]. These findings suggest that the cost function minimised by the back-propagation algorithm may be analogous to the learning objectives of the cortex.

Despite some likeness of deep networks to the cortex and their otherwise outstanding practical successes, deep learning is difficult to motivate from a biological standpoint. In particular, the back-propagation algorithm requires information that is not available locally at the weights, such as gradient information from downstream neurons. Among

others, this requirement limits the implementation of back-propagation-based learning in physical artificial networks such as neuromorphic circuits. One of the aims of this thesis is to explore alternate learning rules for deep networks inspired by biological systems. In particular, I investigate the use of signals modelled on the neuromodulators ACh and DA to augment Hebbian representation learning. In the following section, I describe the basic neural network model used in this investigation.

3.3 Neural Network Model

This work aims at delineating the functions of ACh and DA in shaping cortical representations. To address this question, I take advantage of the correspondence between Hebbian and biological learning and use the former to model the development of representations in the cortex.

In this section, I describe in more details the Hebbian learning network employed in the present work. I first give an account of the learning mechanisms. Next, I present the relationship between learning in the model and in maximum likelihood estimation, justifying the network's particular formulation. Based on this relationship, I introduce a hierarchical version of the model to be used on a classification task. The material in this section was originally presented by Keck and colleagues [KSL12]; the reader should refer to this publication for more detailed derivations and proofs.

3.3.1 Learning Mechanisms

I introduce the learning mechanisms first in a network consisting of two layers, an input and a representation layer. These layers are made of D neurons y_d and C neurons s_c, respectively. Weights W_{cd} connect the two layers (Fig. 3.1). Input vectors activate the input layer where they are normalised as a model of feedforward inhibition.

Feedforward Inhibition

In animals, the responses of cortical neurons are largely invariant to contrast in sensory stimuli [SML90; Man+05; SJL03; Ass+07; OW08], in part due to rapid feedforward inhibition [Swa03; PS01; MKH05; WZ05; Pou+09; IS11]. To emulate this process, the activation of input neurons is normalised:

$$y_d = (A - D)\frac{\tilde{y}_d}{\sum_{d'} \tilde{y}_{d'}} + 1\,, \tag{3.10}$$

where A is a normalisation constant and $\tilde{y}_d$ is the activation of input neurons before the normalisation. The activation of input neurons then satisfies the constraint $\sum_{d=1}^{D} y_d = A$, with $A > 0$, making the responses of representation neurons invariant to the total incoming drive, or contrast-invariant. Feedforward inhibition also ensures that activity in the input layer is never inferior to 1, $y_d \geq 1$. This particular implementation of feedforward inhibition is discussed in Sec. 3.3.2.

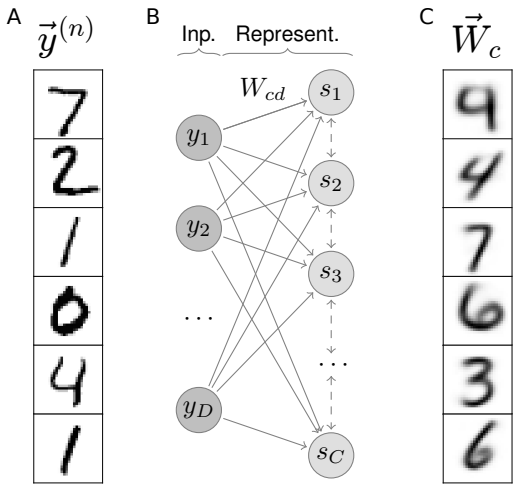

Figure 3.1: Network model of representation learning. (A) Example input vectors $\vec{y}^{(n)}$, illustrated as 2-D images. These input data depict images of handwritten digits from the MNIST dataset [LCB98]. (B) Illustration of the network with the input neurons y_d, weights W_{cd} and representation neurons s_c. Dashed arrows in the representation layer indicate lateral inhibition between neurons. (C) Example weights $\vec{W}_c$, also illustrated as 2-D images.

Input Integration

Representation neurons integrate incoming activation vectors $\vec{y}^{(n)}$ from the input layer through a weighted sum:

$$\tilde{s}_c = \sum_{d=1}^{D} S(W_{cd}) y_d \,, \tag{3.11}$$

where $S(\cdot)$ is a linearised logarithmic function given by:

$$S(W_{cd}) = \begin{cases} W_{cd} & \text{if } W_{cd} < 1 \\ \log(W_{cd}) + 1 & \text{if } W_{cd} \geq 1 \,. \end{cases} \tag{3.12}$$

Taking the logarithm of W_{cd} guarantees approximate optimal learning of the weights (see Sec. 3.3.2), with the linearisation ensuring that the weights are never negative.

Lateral Inhibition

As mentioned in the previous section, Hebb's rule with multiple post-synaptic neurons requires a mechanism to decorrelate neural responses and ensure that neurons learn distinct weights. For this model, a softmax function is chosen as a model of global, unspecific lateral inhibition:

$$s_c = \frac{\exp(\tilde{s}_c)}{\sum_{c'=1}^{C} \exp(\tilde{s}_{c'})} \,. \tag{3.13}$$

The softmax function gives rise to a soft winner-take-all where most neurons are silent and one or just a few neurons have non-zero activation (Fig. 3.2). The softmax also normalises the activation of neurons such that $\sum_{c=1}^{C} s_c = 1$.

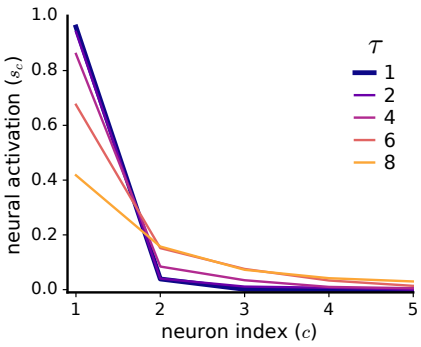

Figure 3.2: Effect of the temperature parameter of the softmax function on the sparseness of the neural code. Values on the x-axis are the indices of representations neurons, sorted from most to least active. The 5 most active neurons out of a network with 49 units are shown. Greater τ parameter values lead to a more distributed code; with $\tau \to \infty$ the activations of neurons tend to a uniform distribution. In this thesis, I use $\tau = 1$ (bold trace). Data are the mean responses for 50 randomly chosen stimuli.

The softmax function can be supplemented by a temperature parameter controlling the strength of the competition:

$$s_c = \frac{\exp(\tilde{s}_c/\tau)}{\sum_{c'=1}^{C} \exp(\tilde{s}_{c'}/\tau)}\,, \tag{3.14}$$

where τ is the temperature parameter. With $\tau \to 0$, the competition is a hard winner-take-all where a single neuron is active and all others are silent; with $\tau \to \infty$, the activations of neurons become uniform (Fig. 3.2). Except if stated otherwise, in this thesis I use $\tau = 1$, equivalent to Eq. 3.13. I examine the effects of code sparseness on learning in Ch. 5 (Sec. 5.5).

Hebbian Learning

Hebbian learning takes place between the input and representation neurons:

$$\Delta W_{cd} = \epsilon(s_c y_d - s_c W_{cd})\,, \tag{3.15}$$

where ϵ is the learning rate. The term $s_c y_d$ represents the classical form of Hebbian learning (post-synaptic activity $\times$ pre-synaptic activity). The term $-s_c W_{cd}$ is a multiplicative synaptic scaling term that prevents unbounded growth of the weights, similar to the additional term in Oja's rule (Eq. 3.5). This learning rule only requires synaptically-local values, namely pre- and post-synaptic activities and the strength of the weight.

Analysis of the Learning Outcome

To study the outcome of the learning mechanisms, it is informative to examine the sum of the incoming weights of a representation neuron c, $\overline{W}_c = \sum_{d=1}^{D} W_{cd}$. As a result of the constraint $\sum_{d=1}^{D} y_d = A$ imposed by feedforward inhibition, we obtain:

$$\Delta \overline{W}_c = \epsilon s_c (A - \overline{W}_c)\,. \tag{3.16}$$

This formulation indicates that the dynamics of $\overline{W}_c$ have a stationary point at $\overline{W}_c = A$. Furthermore, because neural activity s_c and the learning rate ϵ are positive, this stationary point is stable: $\overline{W}_c$ increases when smaller than A and decreases when larger. Thus, the combination of feedforward inhibition (Eq. 3.10) and synaptic scaling (Eq. 3.16) results in the sum of incoming weights $\overline{W}_c$ converging to A over the course of training. In this sense, the responses of representation neurons are invariant to contrast in the network's input.

The stationary point of the summed weights $\overline{W}_c$ is determined exclusively by the constant A and is therefore independent of the statistics of the input stimuli. On the other hand, the convergence values of individual weights W_{cd} depend on the data. If we abstract away the particular transfer function (Eqs. 3.11 and 3.13) and consider an arbitrary function $f_c(\vec{y}^{(n)}, W)$, we can show that the learning dynamics converge to the weights:

$$W_{cd} \approx A \frac{\sum_{n=1}^{N} f_c(\vec{y}^{(n)}, W) y_d^{(n)}}{\sum_{d'=1}^{D} \sum_{n=1}^{N} f_c(\vec{y}^{(n)}, W) y_{d'}^{(n)}} \,. \tag{3.17}$$

This approximation is very accurate for a small learning rate ϵ and a large number of input data N. This result will be useful to support the approximation of optimal learning in relation to maximum likelihood estimation presented in the following section.

Training Data

In this work, I make use of the MNIST (Mixed National Institute of Standards and Technology) dataset (available online at `http://yann.lecun.com/exdb/mnist/`). This dataset contains 70 000 images of digits hand-written by employees of the American census bureau and by American high-school students (see examples in Fig. 3.1 A). Each image consists of $28 \times 28 = 784$ grey-scale pixels and carries a label $m = \{0, 1, 2, \cdots, 9\}$ corresponding to the integer written in the image. These images provide inputs of intermediate complexity and high-dimensionality akin to natural sensory stimuli, making them a popular dataset to study neural information processing ([Nes+13; SPN14]). The dataset is split into a train and test set of 60 000 and 10 000 images, respectively.

The pixel values of the images are fed through the network as the $\tilde{y}$ variable in Eq. 3.10. These neural activation propagate through the network following Eqs. 3.10-3.16. The network then learns synaptic weights based on correlations in the activation of input neurons, with weights that resemble the different digit classes (see Fig. 3.1 C).

I also experimented with input images consisting of grey-scale sinusoidal gratings. These stimuli are interesting as they produce orientation-tuned neurons with properties reminiscent of that of simple cells in the primary visual cortex [HW62]. This correspondence allows a direct comparison of the changes in the weights of the neurons in the network to the modifications in receptive fields observed in animals (e.g., [Sch+01]). However, for this thesis, I concentrate on the MNIST data and only present a short overview of the results with sinusoidal gratings in Appendix A.

3.3.2 Interpretation as Maximum Likelihood Estimation

The neural network presented above can be understood as an approximation to a generative model performing maximum likelihood learning estimation. In particular, the weights of the neurons represent the parameters of centroids in a mixture model. The learning procedure approximates the expectation maximisation (EM) algorithm computing the maximum a posteriori (MAP) parameters. This section presents the relationship between the neural and mixture models, providing a justification for the network's mathematical formulation and giving a proof of the optimality of the learning dynamics.

Mixture of Poisson Distributions

A mixture model assumes that each data point $\vec{y}$ belongs to one of C classes, where each class c is a probability distribution described by the parameters $\vec{\mathcal{W}}_c$. The mixture distribution $\Pr(\vec{y}|c, \mathcal{W})$ defines the variation within a class c, where $\mathcal{W}$ is the matrix of all parameters. The probability $\Pr(c)$ defines how many inputs are generated by each class. Here, the distribution $\Pr(c)$ is assumed to be uniform, $\Pr(c) = 1/C$. Furthermore, since the activations of input neurons are strictly positive, a Poisson distribution is well suited to describe the variation within a class:

$$\Pr(\vec{y}|c, \mathcal{W}) = \prod_{d=1}^{D} \text{Poisson}(y_d; \mathcal{W}_{cd}) \,, \tag{3.18}$$

where D is the number of input dimensions of the data and the parameter $\mathcal{W}_{cd}$ represents the mean and variance of a Poisson distribution. To remain in line with the effects of feedforward inhibition presented in the previous section, the parameters $\mathcal{W}$ are assumed to satisfy the constraint $\sum_{d=1}^{D} \mathcal{W}_{cd} = A$, where $A > 0$.

With this statistical model, we can now derive an optimal solution for learning. In particular, the EM algorithm [DLR77] enables us to learn the MAP solutions for the parameters $\mathcal{W}$.

Expectation Maximisation

The EM algorithm consists of two alternating steps, an expectation (E-step) and maximisation (M-step) step. During the E-step, the posterior probabilities for the data are computed for each classes, thereby estimating the data likelihood under the current model parameters. In the following M-step, the posterior probabilities are used to compute the parameters of the statistical model that maximise the data likelihood.

For the Poisson mixture model, the computation of the posterior probabilities in the E-step is given by:

$$\Pr(c|\vec{y}^{(n)}, \mathcal{W}) = \frac{\exp(\tilde{s}_c)}{\sum_{c'=1}^{C} \exp(\tilde{s}_{c'})} \,, \tag{3.19a}$$

$$\text{where} \qquad \tilde{s}_c = \sum_{d=1}^{D} \log(\mathcal{W}_{cd}) y_d^{(n)} \,. \tag{3.19b}$$

The update of the parameters in the M-step is computed as:

$$\mathcal{W}_{cd}^{new} = A \frac{\sum_{n=1}^{N} \Pr(c|\vec{y}^{(n)}, \mathcal{W}) y_d^{(n)}}{\sum_{d'=1}^{D} \sum_{n=1}^{N} \Pr(c|\vec{y}^{(n)}, \mathcal{W}) y_{d'}^{(n)}} \,. \tag{3.20}$$

The alternate application of these rules guarantees that the data likelihood under the model parameters never decreases. In practice, the data likelihood always increases until it reaches a (possibly local) optimum. If the update rule attains a global maximum, the parameters $\mathcal{W}$ represent the best possible learning result for the data.

Relation to Hebbian Learning

As described above, learning in the neural network leads to weights W_{cd} converging to Eq. 3.17. The parameters $\mathcal{W}_{cd}$ of the mixture model converge to Eq. 3.20. We note the very similar structure of these two equations:

$$W_{cd} \approx A \frac{\sum_{n=1}^{N} f_c(\vec{y}^{(n)}, W) y_d^{(n)}}{\sum_{d'=1}^{D} \sum_{n=1}^{N} f_c(\vec{y}^{(n)}, W) y_{d'}^{(n)}}, \quad \mathcal{W}_{cd} = A \frac{\sum_{n=1}^{N} \Pr(c|\vec{y}^{(n)}, \mathcal{W}) y_d^{(n)}}{\sum_{d'=1}^{D} \sum_{n=1}^{N} \Pr(c|\vec{y}^{(n)}, \mathcal{W}) y_{d'}^{(n)}} \,.$$

In particular, if we choose the transfer function $f_c(\vec{y}^{(n)}, W)$ to be the same as the posterior $\Pr(c|\vec{y}^{(n)}, \mathcal{W})$, the weights of the neural network are equal to the parameters of the mixture model. This result justifies the use of Eqs. 3.19b and 3.19a for the input integration (Eq. 3.11) and the lateral inhibition (Eq. 3.13) in the neural network, respectively.

The close correspondence between the neural model and the EM algorithm guarantees that learning in the network is approximately optimal in terms of maximum likelihood estimation. In this case, the data likelihood under the network's parameters monotonously increases to a maximum, albeit possibly local (see Fig. 3.4 A).

Taking the logarithm of the weights and parameters in Eqs. 3.19b and 3.11 is more difficult to justify in biological terms as it might result in the weights changing sign during learning. The specific form of feedforward inhibition chosen for the model (Eq. 3.10) offers a solution to this concern. Indeed, this formulation ensures that all weights converge to a value superior or equal to 1. With this in mind, we can use a slightly modified version of the logarithmic function, the function $S(\cdot)$ (Eq. 3.12). This linearised logarithmic function, in combination with feedforward inhibition, prevents the weights from becoming negative while preserving the close correspondence between the network's fixed point and the maximum likelihood solution.

The interpretation of the neural network as a probabilistic model also allows to regard neural activations s_c as posterior probabilities for the classes c. We can then extend the model into a hierarchical probabilistic model in order to decode activity in the representation layer and perform classification.

3.3.3 Hierarchical Model for Classification

From the perspective of hierarchical generative models, we can assume that data is generated in the following way: first, a pattern type k is chosen (e.g., a digit class, with $k = \{0, \cdots, 9\}$), second, given k a pattern class c is chosen (e.g., a writing style), and, third, given c the actual pattern $\vec{y}$ is generated (see Fig. 3.3, left panel).

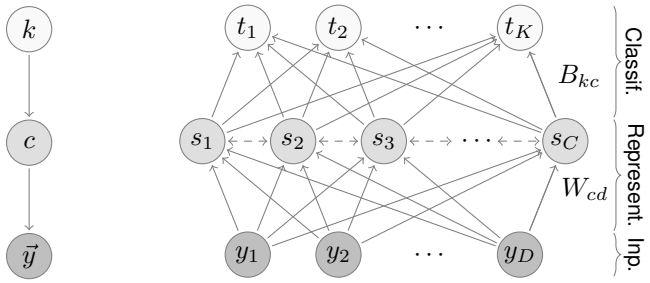

Figure 3.3: Hierarchical model. Left: illustration of the generative hierarchical model. Right: illustration of the corresponding neural network. The network consists of three layers: an input, a representation, and a classification layer. The input layer contains $D = 28 \times 28 = 784$ neurons, one for each pixel of the input images. The number of representation neurons is variable; for most results in this thesis I use $C = 7 \times 7 = 49$ neurons. For the MNIST dataset, there are $K = 10$ classification neurons, one for each of the integer class.

For classification, given an input pattern $\vec{y}$ and the model parameters Θ, we want to infer the class k of the input pattern. In other words, we want to compute the posterior probability $\Pr(k \,|\, \vec{y}, \Theta)$ for all classes k. Here, I assume a uniform prior over the K pattern types, $\Pr(k) = 1/K$, and approximate the posterior using the labels of the input patterns. I first compute a value B_{kc}:

$$B_{kc} := \frac{1}{N_m} \sum_{n=1}^{N_m} \Pr(c|\vec{y}^{(n)}, W) = \frac{1}{N_m} \sum_{n=1}^{N_m} s_c^{(n)} \,, \tag{3.21}$$

with N_m input patterns $\vec{y}^{(n)}$ bearing a label $m = k$. B_{kc} can be interpreted as the weight between the hidden neuron s_c and output neuron t_k. The posterior is then approximated as:

$$\Pr(k \,|\, \vec{y}, \Theta) \approx t_k = \sum_{c=1}^{C} \frac{B_{kc} s_c}{\sum_{k'=1}^{K} B_{k'c}} \,. \tag{3.22}$$

As a classification result $\hat{m}$, I take the unit with the largest value of approximation to the posteriors:

$$\hat{m} = \underset{k=1}{\overset{K}{\operatorname{argmax}}}(t_k) \,. \tag{3.23}$$

This hierarchical formulation allows to decode activity in the representation layer, providing a probabilistic classification of the input images.

Classification Performance

I assess the performance of the network as its rate of correct classification, or its accuracy, on the test set of the MNIST dataset. I use this measure as a quantification of the quality of the neural representation learned by the network. I report the percentage of correct classification and, at times, also the error rate (simply the complement of correct classification). I use the error rate to compare the different models as small differences in performances are more noticeable with this measure.

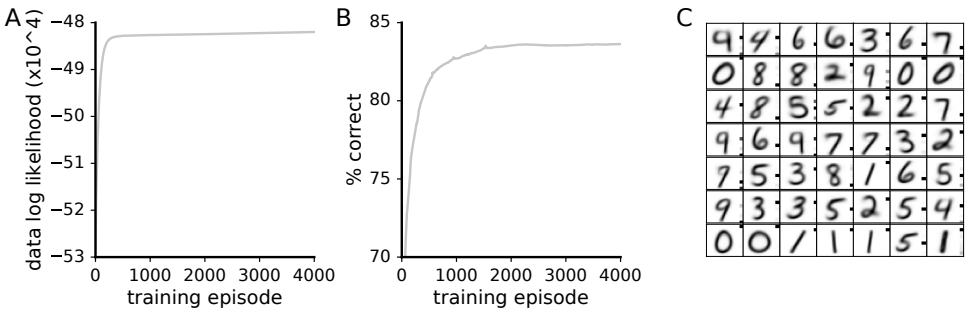

Figure 3.4: Classification performance of the Hebbian network. (A) Progression of the log likelihood of the data under the model parameters. (B) Progression of the classification accuracy on the MNIST dataset. (C) Representation weights learned by the network. The images represent weight vectors $\vec{W}_c$; immediately to the left of each image are the weight vectors $\vec{B}_c$ of the classification layer, with classes '0' to '9' from top to bottom. Data likelihood, and to a close approximation classification accuracy, increase monotonously during training.

3.3.4 Supplementary Methods

Code

The model was written in the Python programming language and run on a personal computer or on a computer cluster. For parameters exploration, I use the PyPet package developed in our group [MO16]. To benchmark the network, I compare its accuracy with that of MLPs and ConvNets. These networks where run with the Python modules Scikit-learn [Ped+11] (version 18.dev0, downloaded on 04/29/16) and TensorFlow [Aba+15] (version 1.0, downloaded on 04/22/17), respectively.

Training

I pre-compute the activation of input neurons $\vec{y}^{(n)}$ through Eq. 3.10 for all N training images. Learning proceeds as full iterations over the dataset during which input neuron activations $\vec{y}^{(n)}$ are presented in a random order to the representation layer.

Weight Initialisation

Weights of representation neurons are initialised using the statistics of the input images. Specifically, I initialise the weights with the mean activation of input neurons taken over the whole dataset, with the addition of noise to break symmetry:

$$W_{cd} = \mu(y_d) - \sigma^2(y_d)\eta_{\text{init}}\,, \tag{3.24}$$

where $\mu(\cdot)$ and $\sigma^2(\cdot)$ are the mean and variance taken over all training images N, respectively, and η_{init} is noise drawn from a uniform distribution in the interval $[0, 2)$.

Mini-batch Learning

To speed up computation, I train the network using mini-batches. The weight updates ΔW_{cd} are then computed over batches of 50 training examples. Using mini-batches only negligibly affects the representation learned by the network and the performance of the algorithm.

Update of the Classification Weights

The value B_{kc} is updated after presentation of every 100 input images, or roughly 600 times during one iteration over the dataset.

Cross-validation

Parameter search is performed through five-fold cross validation on the training dataset. The network performance is reported as the mean classification accuracy over the five validation sets. To compare the networks' performances across models (e.g., DA versus ACh signalling), networks with the optimal parameters are re-trained on the whole training data (including the validation set) and their performances are reported for a separate test set. This approach was used for all parameter searches except for comparison between models with identical parameters spaces (Fig. 4.8 and 6.14). In these cases, measuring accuracy on the testing set (without cross validation) yields unbiased results.

3.4 Models of the Neuromodulators

In the previous section, I introduced a neural network that learns representations from input data which I consider as a model of cortical learning. The representations are obtained in a fully unsupervised manner and only reflect correlations in the data. The work presented in this thesis is concerned with how factors additional to regularities of unlabelled data can shape representations. In particular, I investigate the effects of two neuromodulators, ACh and DA, in guiding representation learning. I address this question by modelling the impact of cholinergic and DAergic transmission on synaptic plasticity and examining the results on the network's neural representation and classification performance.

In this section, I present the models of ACh and DA used for the remainder of the thesis. I derive the models based on the observation that, when pairing a stimulus with either ACh [KM98a; RD88; Web+91b; KM98b; GLS88; FMS07] or DA [BCM01; FSV10], the representation of this stimulus enlarges in the cortex. That is, the modulators alter the tuning of cortical neurons to raise neural responses to the paired stimulus. At the population level, this translates in an expansion of the cortical area responsive to the stimulus. This phenomenon can be emulated in the model as follows.

3.4.1 Modifications in Representations

The neural network introduced above can be interpreted as a generative model. As such, the network performs density estimation: the distribution of its weights is determined by the density of data points in the input space. For instance, the number of data points in a cluster (relative to the total number of data points) dictates the relative number of weights encoding this cluster.

To illustrate this observation, let us consider the following manipulation: an input $\vec{y}^{(*)}$ is duplicated and presented N^* times to the network. From Eq. 3.17, we note that this manipulation brings the weight vector of neuron c, $\vec{W}_c$, closer to the input vector $\vec{y}^{(*)}$ by a factor proportional to $\frac{N^*}{N}$, where N is the total number of stimuli (see Fig. 3.5 B).

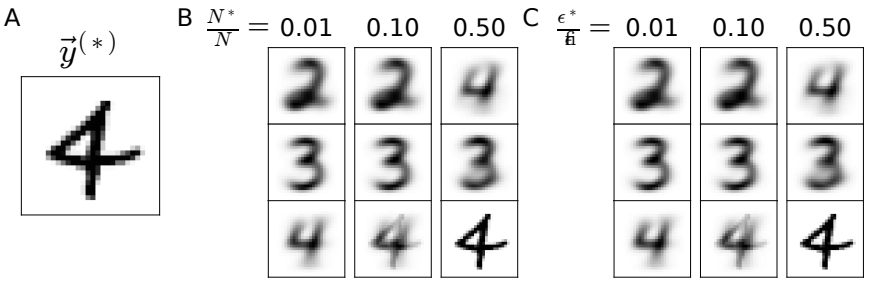

Figure 3.5: Effects of increased stimulus frequency and learning rate on the weights. (A) Illustration of a target stimulus $\vec{y}^{(*)}$. (B - C) Depiction of the representation weights of a neural network. The network contains three neurons and is trained on a dataset containing only '2', '3', and '4'. (B) The network is trained with $\vec{y}^{(*)}$ presented with a greater frequency ($\frac{N^*}{N}$). (C) The target stimulus $\vec{y}^{(*)}$ is paired with a higher learning rate ($\frac{\epsilon^*}{\bar{\epsilon}}$). The two manipulations result in almost identical alterations, namely an over-representation of the target stimulus.

Pairing a stimulus with a greater learning rate bears a similar effect as raising the frequency of this stimulus. For instance, presenting a training image twice is similar to presenting this image once but with a twice larger learning rate (these are almost identical for small learning rates). In more details, coupling an input vector $\vec{y}^{(*)}$ with a learning rate ϵ^* will bring the weight vector $\vec{W}_c$ closer to $\vec{y}^{(*)}$ by a factor proportional to $\frac{\epsilon^*}{\bar{\epsilon}}$, where $\bar{\epsilon}$ is the summed learning rate for all stimuli, $\bar{\epsilon} = \sum_{n=1}^{N} \epsilon^{(n)}$. In practice, this manipulation is equivalent to increasing the frequency of an input vector (see Fig. 3.5).

I make use of this characteristic of the learning dynamics to emulate the effects of ACh and DA on cortical representations.

3.4.2 Models of ACh and DA

Based on the observations presented above, I include ACh and DA in the network as a modulation of the learning rate:

$$\text{acetylcholine} \qquad \Delta W_{cd} = ACh \cdot \epsilon(s_c y_d - s_c W_{cd})\,, \tag{3.26a}$$

$$\text{dopamine} \qquad \Delta W_{cd} = DA \cdot \epsilon(s_c y_d - s_c W_{cd})\,. \tag{3.26b}$$

This model abstracts away the biological details of the neuromodulator's impact on synaptic plasticity. Following such an approach is nonetheless in general agreement with biological observations in that both ACh [BAS92; Chu+13] and DA [BCO02; MMO06; SZW05; MMO06] are reported to promote synaptic plasticity, as they do in the present mathematical formulation.

The models of ACh and DA allow to simulate the basic results of the pairing experiments presented in the previous chapter. For instance, if a target stimulus class is coupled with $DA = \rho$, for $\rho > 1$, while all other classes have $DA = 1$, the target class will be over-represented in the network's weights. On the other hand, for $\rho < 1$, this class will be under-represented (Fig. 3.6).

Here, the models of ACh and DA are mathematically identical and the pairing protocol gives rise to the same changes in weights in both cases. It should be noted, however, that in animals there are some distinctions in the effects of pairing with

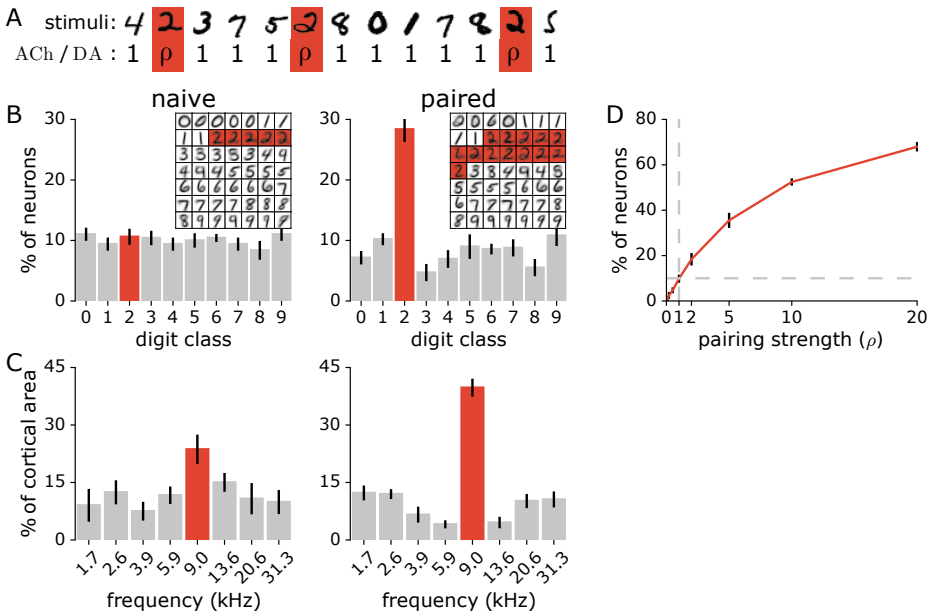

Figure 3.6: The models of ACh and DA emulate the effects of neuromodulator pairing protocols in animals. (A) Stimuli of a target class (in this case '2') are paired with an activation ρ of either the ACh or DA variable in Eq. 3.25. For all other classes ACh and $DA = 1$. (B) Percentage of neurons responsive to the target class, before (naïve) and after pairing with $\rho = 5$. Inset is an example the network's representation, with highlighted neurons responding preferentially to the target class. (C) Similar pairing protocol in rats. Electrical stimulation of the VTA triggers DA release shortly after the presentation of an auditory tone of a 9 kHz. Modified with permission from [Bao+03]. (D) Percent of neurons responding preferentially to the target class as a function of the pairing strength, ρ. Data are the mean of 10 runs, each with a different target class, error bars are the variance. In rats as in the model, the pairing manipulation result in an increase in the proportion of a representation dedicated to the target stimulus.

ACh or DA. For instance, Kilgard and colleagues [KM98a] report that coupling a tone with ACh release increases the representation of the paired frequency as well as those of adjacent frequencies in A1. A similar manipulation with DA decreases the representation of adjacent frequencies [BCM01]. DA also bears a temporally asymmetric influence such that reversing the presentation order of the stimulus and DA suppresses the representation of the stimulus [Bao+03], an effect absent with ACh. For this work, however, these differences are not taken into account.

For the experiments with ACh and DA signalling, I model the critical period observed in animals as a pre-training of the network with pure Hebbian learning. This training takes place until the representation stabilises, i.e., until performance saturates. After this point, ACh and DA transmission begins. The pre-training is not required for any of the results obtained in this work except for the optimal DA activation values found through parameter search; without the pre-training, the optimal DA activation profile does not match that reported in animals (see Ch. 5, Sec. 5.3).

3.4.3 Relation to Previous Models

Previous work addressed the impacts of ACh and DA in neural circuits. I briefly summarise some of the approaches taken below.

Acetylcholine

Weinberger and Jonathan [WB98] develop a model of ACh signalling to study its role in classical conditioning. The authors make use of a modified version of Hebb's rule and simulate the action of ACh as an amplification in the post-synaptic activation of target neurons. An *in vivo* micro-stimulation study validates this model. For the Hebbian rule used in this thesis, the two models of ACh are mathematically equivalent; this work thus offers support to the formulation described above.

Li and Cleland present a detailed biophysical model of ACh neuromodulation in the olfactory bulb [LC13]. Their simulation reproduces the effects of ACh in transiently sharpening receptive fields and in increasing spike synchrony but does not address its role in synaptic plasticity.

Dopamine

Numerous work addressed the question of reward-based learning in neural networks. Roelfsema and colleagues show that a RPE signal affecting the magnitude of Hebbian updates allows a neural network to learn various classification tasks [RO05; ROW10; RRB12]. Other models similarly make use of a reinforcement signal affecting the learning rate of Hebbian plasticity to learn stimulus-response associations [LLD10; LG09]. In these latter cases, however, the tuning properties of representation neurons are set manually and only the weights of the read-out units are learned.

The DA learning mechanism is closely related to algorithms such as REINFORCE [Wil92] in which a reinforcement signal acts on the learning rate of a neural network's weight update rule. Of note, despite this close correspondence, the decision to model DA as a modulation of the network's learning rate was made not to resemble those rules but rather to mirror biology. Here, the model of DA (and ACh) reproduces the results of pairing experiments (Fig. 3.6). The close similarity between this model's and previous' update rules can thus be taken as further support for the biological realism of the latter.

3.5 Summary

Various learning rules enable ANNs to extract useful features from data. These rules include correlation-based and error-correcting methods, each pursuing different objectives. Specifically, correlation-based rules aim at finding statistically sound representations, for instance linearly uncorrelated features, while error-correcting rules aim at finding representations that minimise errors in a network's output. The features learned in these networks make up neural representations of data. These representations are of great practical relevance and can also serve as valuable models of sensory representations in the cortex.

For this research, I make use of such an ANN to investigate the beneficial roles of neuromodulators ACh and DA in shaping neural representations. In this chap-

ter, I described a Hebbian learning mechanism originally developed by Keck and colleagues [KSL12] that approximates optimal learning in term maximum likelihood estimation. I derive a simple mathematical formulation for the effects of ACh and DA in which both signals modulate the learning rate of the network. This approach is in line with a few experimental observations. In particular, in animals as in the model, stimulus-modulator coupling enlarges the representation of the paired stimulus. The model presented here permits the study of this plasticity from a functional perspective. In the following chapters, I present the results of this investigation.

Chapter 4

Model of Acetylcholine

4.1 Overview

Animals can selectively attend to relevant aspects of their environment while ignoring others. For instance, a student focused on writing his thesis in a busy café won't notice the conversations going on at neighbouring tables. This ability derives from the suppression of neural responses to distracting stimuli [MCW08; MK13] and allows animals to efficiently allocate their processing resources to relevant sensory information.

ACh release in the mammalian neocortex is tightly linked with such attentional processes. For instance, as a rat detects a behaviourally meaningful sensory cue, a spike in cortical ACh accompanies the reorientation of its attention towards the cue [Par+07]. Additionally, when a rat performs a task requiring sustained attention, the concentration of ACh in its prefrontal cortex more than doubles compared to control [Arn+02; KBS06]. In the course of such tasks, distractors that further tax an animal's attention evoke additional ACh release [HSB00; KBS06]. These observations indicate that the cholinergic system responds to events demanding an animal's attention such as relevant stimuli or challenging tasks. In this sense, ACh transmission reflects the cognitive construct of attentional effort [SGK06].

ACh transmission is permissive of cortical reorganisation in various circumstances. In particular, stimuli that repeatedly coincide with activation of the cholinergic system excite a larger area in the cortex [RD88; Web+91b; KM98a; KM98b; SK83]. It remains unknown why a signal associated with attentional efforts elicits such changes in cortical representations.

In this chapter, I address this question using the model of ACh introduced above. I extend the model to simulate the physiological release properties of ACh. Here, ACh activation approximates attentional efforts, reflecting both the requirements of a task and the motivation to perform on this task. The following sections are organised such that I deal separately with these aspects of task demand and motivation to perform. In the first section, I present a basic model for approximating task demand from the network's classification performance. I then look at alternate models of task demand, such as classification confidence. Finally, I examine the motivational aspect of attention by assigning varying levels of relevance to the different stimulus classes.

4.2 Classification Performance as Proxy of Task Demand

Attentional demand is an intuitive concept experienced in everyday life. For instance, consider driving on a winding countryside road in different conditions: on quiet day (control), while talking on the phone (distractor), during a snowstorm (task difficulty), after a long day of work (fatigue), or after a glass of wine (pharmacological manipulation). Compared to control, the other conditions pose further attentional demands and require additional efforts to achieve similar performance levels.

Although we routinely encounter examples of attentional demand, it is unclear how the mammalian brain encodes it. Sarter and colleagues [SGK06] review evidence suggesting that indications of performance decrease, such as rises in error rates or decays in reward rates, trigger effortful cognitive control in animals. In particular, electrophysiological and neuroimaging studies suggest that errors and reward losses activate the anterior cingulate cortex (ACC) [Haj+05; UVC04]. The authors suggest that the ACC acts as a performance monitoring system that activates the cholinergic system, in turn initiating attentional efforts to prevent further erroneous behaviour [SGK06]. Based on these observations, I simulate the ACC by measuring the network's performance on the classification task. This measure of performance determines the strength of ACh activation. I present the details of this approach below.

4.2.1 Model Description

The network's classification accuracy is a straightforward indication of task demand. The rate of correct classification is readily computed by comparing the labels of the training images with the network's classification output. I compute the mean rate of correct classification for the last training episode as:

$$\bar{p} = \frac{1}{N} \sum_{n=1}^{N} p^{(n)} \,, \tag{4.1}$$

$$\text{where} \quad p^{(n)} = \begin{cases} 1 & \text{if } \hat{m}^{(n)} = m^{(n)} \\ 0 & \text{otherwise} \,, \end{cases}$$

where $m^{(n)}$ and $\hat{m}^{(n)}$ are the label and the classification output during the last training episode for image $\vec{y}^{(n)}$, respectively, and N is the number of training images. I similarly compute the mean rate of correct classification separately for each label class m, denoted as $\bar{p}_m$. Fig. 4.1 A depicts the rates of correct classification for the individual classes and averaged over all classes.

In the model, the variable ACh regulates the network's learning rate (Eq. 3.26a), emulating the effects of the cholinergic system on synaptic plasticity. For each stimulus, the value of the ACh variable is given by the ratio of the performance averaged over the class of the current stimulus to the performance averaged over all classes, fed through a sigmoid function:

$$ACh = \frac{\beta}{1.0 + \exp(\alpha \cdot (\bar{p}_m / \bar{p} - 1.0))} \,, \tag{4.2}$$

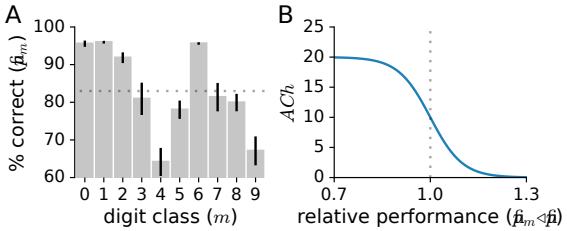

Figure 4.1: ACh activation is inversely related to classification performance. (A) Classification performance $\bar{p}_m$ for the 10 digit classes m. A dashed line indicates the average performance over all classes, $\bar{p}$. The data are the mean of 10 runs; error bars indicate the variance. (B) ACh activation is given by the performance ratio $\bar{p}_m/\bar{p}$ fed through a sigmoid function (Eq. 4.2; here $\alpha = 15$ and $\beta = 20$). This measure quantifies the difficulty of the current stimulus. The lower the performance ratio (i.e., the more difficult a stimulus class), the larger the ACh signal.

where α and β are parameters of the sigmoid function (see Fig. 4.1 B). The ratio $\bar{p}_m/\bar{p}$ provides a measure of how challenging the current stimulus is, or of the task demand. Stimuli from classes that are easy for the network (those from classes with correct classification higher than the average) elicit weak ACh activation. Stimuli that are difficult for the network (those from classes with correct classification less than average) elicit strong ACh activation. The sigmoid function ensures that ACh release saturates for particularly easy or difficult stimuli (note that I considered other functions including an exponential and a polynomial; the sigmoid gave the best results, probably because of its saturating properties).

4.2.2 Results

As a model of the critical period in animals (see Ch. 3), I pre-train the network through Hebbian learning. After performance reaches a plateau, I allow the activation of ACh. At this point, I perform parameter search for the α and β parameters of Eq. 4.2.

Parameter Search

The α and β parameters control the height and sharpness of the ACh activation function, respectively. I determine the values of these parameters through grid search, selecting the values that maximise classification performance. Fig. 4.2 shows the results of this parameter exploration, with $\alpha = 15$ and $\beta = 30$ yielding the greatest classification rates. I use these parameters to train the network model.

Changes in the Network

Visual inspection of the weights of the neural network (see Fig. 4.3 A) indicates that ACh activation alters the number of neurons dedicated to the different digit classes. For instance, there are more neurons resembling a '4' and fewer neurons resembling a '1' after training with ACh release. In order to quantify this redistribution, I determine the preferred digit class of a representation neuron, m_c, by taking $m_c = \text{argmax}_{k=1}^{K}(B_{kc})$. B_{kc} can be understood as the weight from representation neuron c to output neuron k and is computed as the mean activation of neuron c over all input images having a

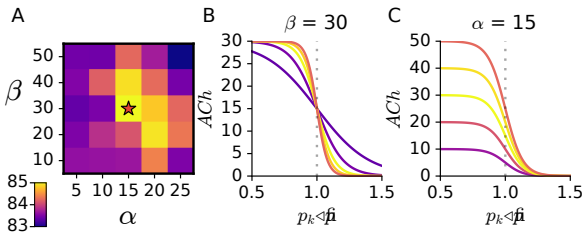

Figure 4.2: Exploration of the parameters of the ACh activation curve. (A) 2-dimensional grid search for the α and β parameters. A star indicates the best parameter set. (B-C) A selection of explored ACh activation functions, in (B) with the β parameter fixed to its best value ($\beta = 30$) and in (C) with the α parameter fixed to its best value ($\alpha = 15$). The colour axis indicates the rate of correct classification on the test dataset. Data are the mean performances over five validation sets.

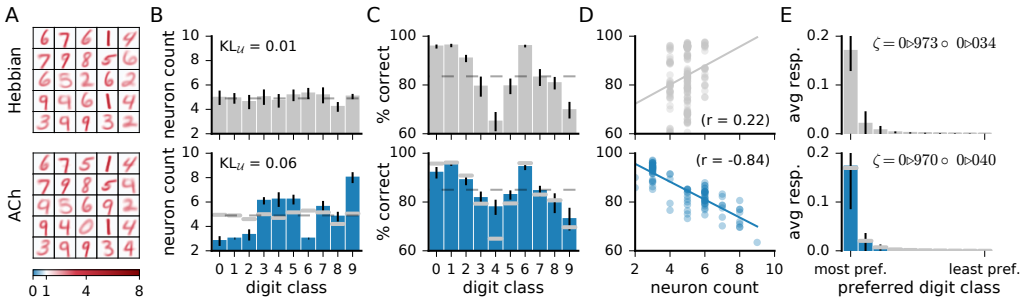

Figure 4.3: ACh augments the neural representations of challenging classes. (A) Example weights of representation neurons (25 out of 49). (B) Distribution of the neurons preferred digit classes. Dashed line is a uniform distribution. $\mathrm{KL}_\mathcal{U}$ indicates the mean KL divergence from a uniform distribution $\mathcal{U}$ to the observed distribution. (C) Performance of the network on the different classes. Dashed line is the average over all classes. (D) Performance on a class as a function of the number of neurons dedicated to this class. (E) Average response of neurons to the digit classes, ordered from most to least preferred class. Data are the mean of 20 runs, error bars indicate the variance. Grey overlaid bars in (B), (C) and (E) are values for Hebbian learning for comparison.

label $m = k$ (see Eq. 3.21). In this sense, $\mathrm{argmax}_{k=1}^{K}(B_{kc})$ gives the digit class to which neuron c maximally responds, or its preferred class.

We can then look at the distribution of digit class preferences (Q) in networks trained with Hebbian and ACh-based learning (Fig. 4.3 B). In the case of Hebbian learning, the distribution Q_{Hebb} is close to uniform. We can quantify the distance from a uniform distribution ($\mathcal{U}$) to the observed distribution using the Kullback-Leibler (KL) divergence; here, $\mathrm{KL}(Q_{\mathrm{Hebb}} \,||\, \mathcal{U}) = 0.01$. As this measure indicates, the distribution Q_{Hebb} is not perfectly uniform and some classes are represented with more neurons than others. There is a positive correlation between the number of neurons dedicated to a class and the network's performance on this class ($r = 0.22$, Fig. 4.3 D), indicating that representing a class with more neurons is beneficial to performance.

Training with ACh redistributes class preferences in the network, leading to a less uniform distribution ($\mathrm{KL}(Q_{\mathrm{ACh}} \,||\, \mathcal{U}) = 0.06$, Fig. 4.3 B). Specifically, ACh increases the number of neurons dedicated to challenging classes while easier classes are represented

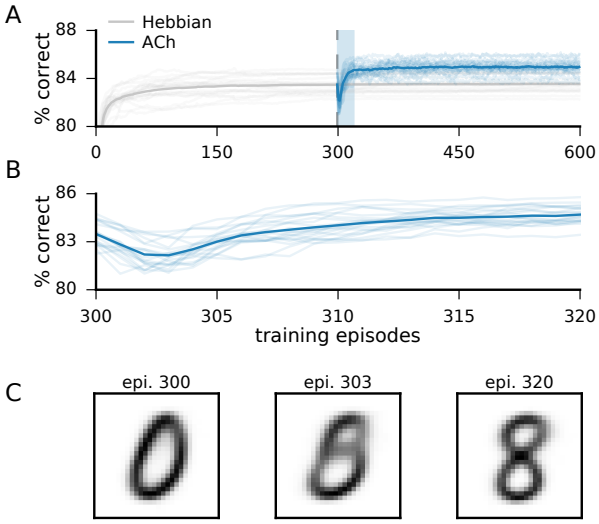

Figure 4.4: Performance during training with ACh initially drops before recovering. (A) Rate of correct classification on the test dataset over the training episodes. A vertical dashed line indicates the start of ACh transmission. (B) Magnified area highlighted in (A) (episodes 300 to 320, immediately after beginning the activation of ACh). Performance with ACh initially drops before recovering. (C) Example weights of a representation neuron, changing from a '0' to an '8'. The point of lowest accuracy corresponds to an episode at which the neuron is poorly tuned to either class.

with fewer units. Consider for example the rate of correct classification for classes '1' and '4', the classes on which the network performs best and worst, respectively. ACh release leads to a respective decrease and increase in the number of neurons preferring these classes.

Performance Measure

The redistribution of neurons elicited by ACh raises the network's accuracy on the difficult classes (e.g., '4') and lowers performance on the easy classes (e.g., '1', Fig. 4.3 C). ACh thus reverses the correlation between neuron count and performance ($r = -0.84$, Fig. 4.3 D). On average over all classes, performance grows from $83.5 \pm 0.7\%$ with Hebb's rule alone to $85.0 \pm 0.6\%$ when supplemented with ACh, corresponding to a relative decrease of 12% in the error rate.

At the onset of training with ACh, the rate of correct classification shortly drops before recovering and surpassing that of Hebbian learning (Fig. 4.4). This event is due to neurons weights being poorly representative of any digit class as they "move" from one class (or datapoint cluster) to another.

Interpretation of the Parameter Search

In light of these results, we can interpret the roles of the parameters α and β of the sigmoid function. β gives the scale of the ACh activation function. As the ACh variable multiplies the learning rate of the network, β is equivalent to uniformly scaling this learning rate. α determines the sharpness of the ACh activation function. In other words, α sets the relative activation of ACh for an easy versus a difficult stimulus. A

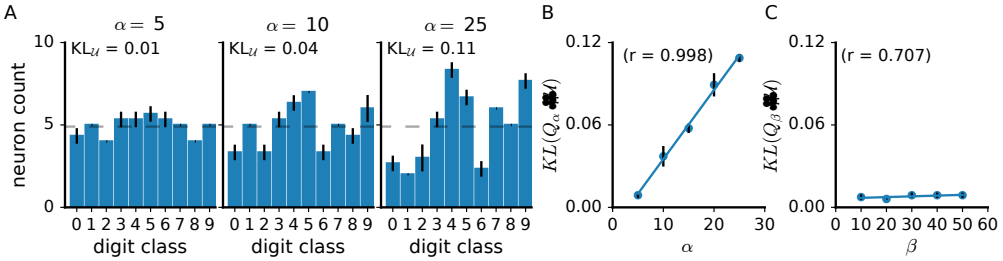

Figure 4.5: Interpretation of the results of the parameter exploration. (A) Distribution of the neurons' preferred digit class for different values of the α parameter. The mean KL divergence from a uniform distribution $\mathcal{U}$ to the observed distributions Q is indicated on the plots ($\mathrm{KL}_{\mathcal{U}}$). (B-C) This KL divergence for networks trained with different values for (B) α and (C) β. The α parameter strongly influences the divergence from a uniform distribution whereas β does not. The data are the means of 3 runs; the error bars indicate the variance.

large α leads to a greater redistribution of the neurons' preferred classes and therefore to a less uniform distribution of class preferences. The value of α strongly predicts the divergence from a uniform distribution to the observed distribution of neurons per digit class ($r = 0.999$, Fig. 4.5).

4.3 Alternate Models of Task Demand

Classification performance is a natural measure to quantify task demand. However, computing correct classification requires environmental feedback to determine whether each classification decision is correct. Introspection suggests that such feedback is not necessary to assess attentional demand. For example, imagine listening to someone speak in your mother tongue or in a foreign language. Understanding the speech in the foreign language experienced as more challenging and as requiring greater attention, this even in the absence of environmental feedback. This example indicates that perception by itself can be experienced as demanding without requiring additional information from the environment. In this section, I consider alternate methods to quantify task demand in order to reduce the need for environmental feedback. Specifically, I examine the use of classification confidence as a proxy of task demand. This measure is computed by the network itself and does not require immediate environmental feedback for its evaluation.

Additionally, thus far I quantified task demand as an average for the different digit classes (i.e., $\bar{p}_m$). This approach requires labels (m) both to compute the average demand of a class and to retrieve this value for the current stimulus. To remove this requirement for labels, I examine two alternatives: first, using the stimulus classes inferred by the classifier instead of the true labels and, second, discarding the average over the classes altogether by computing task demand in a stimulus-wise manner.

4.3.1 Classification Confidence

In animals, decision confidence is a good indicator of task difficulty. For instance, in rats trained on an odourant discrimination task, discrimination accuracy depends on the distance of an odourant to the decision boundary; rats make fewer mistakes

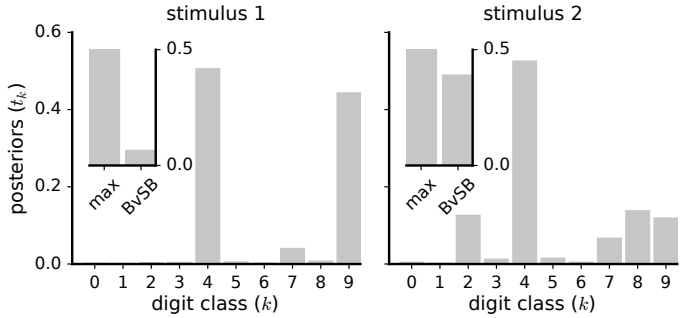

Figure 4.6: Illustration of two measures of classification confidence. The plots represent the classifier's posterior probabilities t_k over the digit classes for two stimuli. The inset plots depict classification confidence quantified through maximal posterior and BvSB posteriors. In both cases, the network classifies the input stimuli as a '4' with a posterior probability of around 50%. The two cases differ, however, in that the remaining probability mass is either concentrated mostly in one class (stimulus 1) or spread among multiple classes (stimulus 2). In the first case, the nearly equal posterior probabilities for classes '4' and '9' may indicate lower classification confidence. The BvSB approach captures this lower confidence while the maximal posterior does not.

on odourants far from the decision boundary. Neurons in the orbitofrontal cortex (OFC) appear to encode the animals' decision confidence on this task [Kep+08]. Here, the firing of some OFC neurons correlates with stimulus difficulty (measured as the stimulus' distance from the classification boundary) as well as with response accuracy. For instance, a neuron fires more for a difficult than easy stimulus and for an incorrectly classified than correctly classified stimulus. This neuron's firing pattern thus reflects the animal's perceived response confidence. Furthermore, OFC firing predicts the rats' willingness to wait for a reward following a decision, indicating that this firing represents a measure of the animals' trust in their decisions that also guides behaviour. This measure of confidence strongly correlates with the animals' response accuracies.

Animal brains thus encode measures of classification confidence which relate to the task's difficulty. In the model, classification confidence can be quantified using the classifier's posterior probabilities over the digit classes, with a large posterior indicating high confidence. I consider two measures to quantify classification confidence: either taking the maximal posterior probability over the classes or taking the difference between the maximal and second maximal posterior probabilities (known as Best versus Second-Best, BvSB [JPP09]). The BvSB posteriors may provide a more accurate estimation of classification confidence as it considers the distribution of the probability mass among competing classification options (see Fig. 4.6). Formally, the two measures of classification confidence κ are given as:

$$\text{max posterior} \qquad \kappa = \max_{k=1}^{K}(t_k) \tag{4.3a}$$

$$\text{BvSB posteriors} \qquad \kappa = \max_{k=1}^{K}(t_k) - \max_{k=1}^{K}{}_2(t_k) \tag{4.3b}$$

where the $\max_2$ operator gives the value second to the maximum.

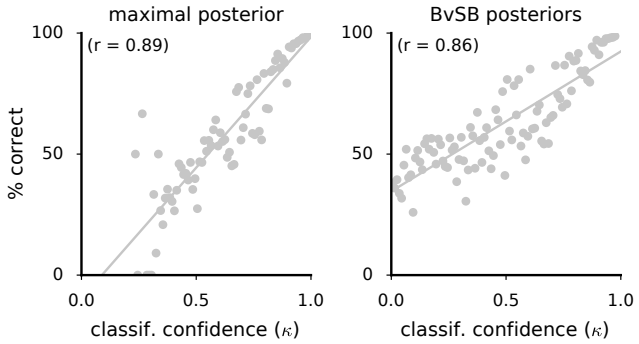

Figure 4.7: Classification confidence strongly predicts rates correct of classification. I train a network through Hebbian learning, have it classify the test images of the MNIST dataset, measure the network's classification confidence for each of those stimuli, and bin the classification confidence (bin size of 0.02%). For each bin, I measure the average correct classification. This rate of correct classification for each bin is plotted against the classification confidence for this bin. Classification confidence is measured either as the maximal posterior probability or as BvSB posterior probabilities. The rate of correct classification exhibits a strong positive correlation with both measures of classification confidence.

Both the maximal posterior and the BvSB posteriors strongly correlate with correct classification ($r = 0.89$ and $r = 0.86$, respectively, see Fig. 4.7). These strong correlations indicate that classification confidence indeed expresses how difficult a stimulus is for the network.

Here, I use the same approach as for classification performance to compute the value of the ACh variable. I measure the classification confidence averaged over each class separately as well as averaged over all classes. ACh activation is given by the ratio of the classification confidence for the class of the current stimulus to the average classification confidence for all stimuli, fed through a sigmoid function:

$$ACh = \frac{\beta}{1.0 + \exp(\alpha \cdot (\bar{\kappa}_m / \bar{\kappa} - 1.0))} \tag{4.4}$$

where $\bar{\kappa}_m$ is the network's average classification confidence on the class m of the current stimulus, $\bar{\kappa}$ is the average classification confidence over all stimuli, and α and β are parameters of the sigmoid function.

4.3.2 Inferred Stimulus Class and Stimulus-Wise Confidence

So far, tracking averages over classes required a knowledge of the label of each stimulus. To remove this need for labels, I first consider making use of the network's inferred stimulus class $\hat{m}$ instead of the true label m. In a second approach, I remove the averages over the classes altogether. Specifically, instead of retrieving the average classification confidence for the class of the current stimulus, I simply take the classifier's output confidence for the current stimulus. In this case, ACh activation is given by the ratio of the classification confidence for the current stimulus to the average over all

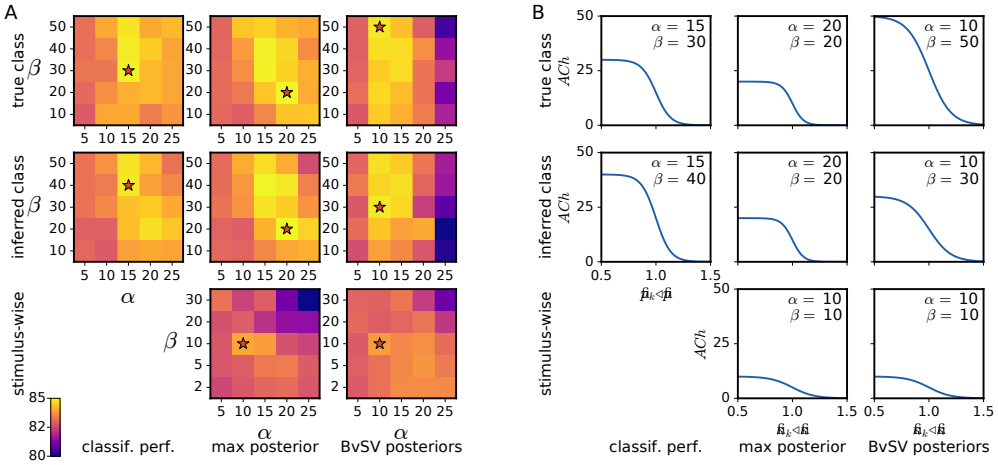

Figure 4.8: Grid search for optimal values of the α and β parameters. (A) Results of the grid search, with the colour axis indicating classification performance. A star indicates the best parameter set. Performance values are averages of 3 runs. (B) Optimal ACh activation function.

stimuli:

$$ACh = \frac{\beta}{1.0 + \exp(\alpha \cdot (\kappa_n/\bar{\kappa} - 1.0))}, \tag{4.5}$$

where κ_n is the confidence for the current stimulus n.

4.3.3 Results

I introduced three measures of task demand (classification performance, maximal posterior, and BvSB posteriors) and three ways of computing these values (average over true classes, average over inferred classes, and computed in a stimulus-wise measure). Each measure of task demand can be combined with each computing approach, except for classification performance and stimulus-wise computation. This results in $3 \times 3 - 1 = 8$ different models. In the following text, I present the results of the exploration for the optimal α and β parameters. I then compare the performance measures for these models.

Parameter Exploration

Fig. 4.8 depicts the results of the exploration for the α and β parameters of the sigmoid function for the 8 different models. The approaches using averages over classes (two top rows) result in similar patterns of parameter performance values and optimal ACh activation functions (and equivalently for the stimulus-wise approaches).

Performance Measures

All models of ACh activation perform statistically significantly better than Hebbian learning alone ($p < 0.0001$, see Fig 4.9). The best performing approach uses the maximum posterior to quantify task demand and computes this measure as an average

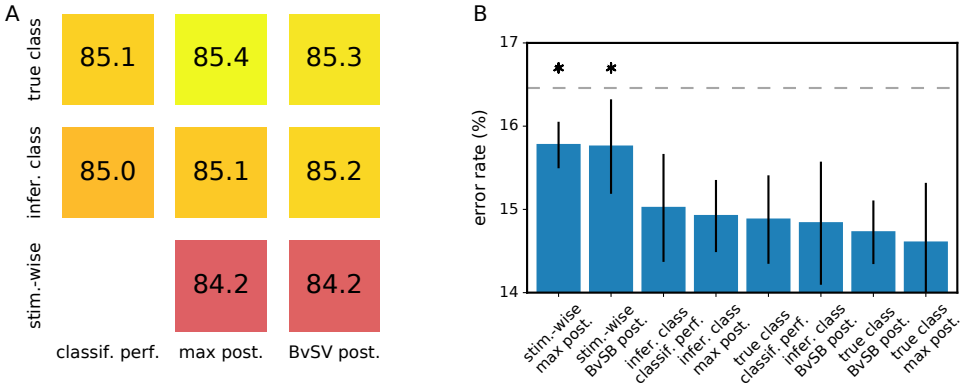

Figure 4.9: Performance comparison for the different models. (A) Colours and numbers indicate the average rate of correct classification on the test MNIST dataset for a model. (B) Error rates for the models, sorted in decreasing order. A dashed line indicates mean error rate for Hebbian learning alone. Values are the average of 10 runs, error bars are the variance. A star indicates error rates that are statistically significantly greater than that of the best model ($p < 0.01$). All models perform statistically significantly better than Hebbian learning alone ($p < 0.0001$)

over the true class label. This method gives a rate of correct classification of 85.4±0.7%, a relative decrease of 12% in error rates compared to Hebbian learning.

The two models that compute classification confidence in a stimulus-wise manner yield a significantly poorer classification accuracy than the best model ($p < 0.01$). All other models (those relying on averages over classes) are not statistically significantly worse than the best model ($p > 0.01$).

Although the models using true classes perform on average better than those using inferred classes, this trend is not statistically significant; the classifier thus seems to provide an estimate of class labels that is adequate for ACh learning. Similarly, the accuracies of the different methods to quantify task difficulty are not significantly different from each others; here again, the classifier's approximation suffices. Additionally, the two estimates of classification confidence, maximal or BvSB posteriors, appear equivalent. Therefore, of the approaches examined in this section, only that of computing ACh as an average over classes bears a statistically meaningful advantage.

Comparison of Stimulus-wise versus Class-wise ACh Release

The models with stimulus-wise ACh activation perform significantly worse than those with an activation averaged over the classes. To determine the cause of this effect, I examine the statistics of the ACh signal in two networks trained with each approach. I use the maximal posterior as a measure of task demand in both cases and the true class label to compute the average over classes.

The mean ACh activation for corresponding classes is not statistically different from one network to the other (Fig. 4.10 A). However, the distributions of ACh activation within classes largely differ. For example, Fig. 4.10 B depicts the release of ACh for all stimuli labelled as '7', averaged over the 50 training episodes. In the network with class-wise ACh activation, all stimuli trigger identical ACh release; variance derives from fluctuations between episodes (see Fig. 4.10 C). In contrast, in the network with stimulus-wise ACh transmission, the distribution of ACh values within a class is

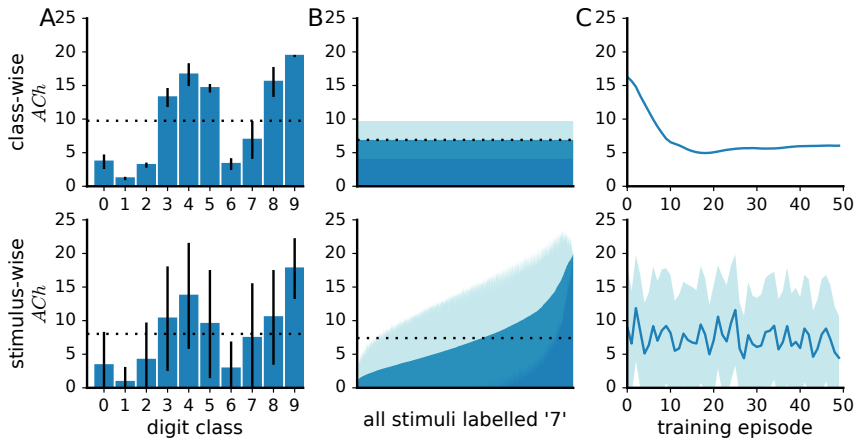

Figure 4.10: Comparison of the distributions of ACh activation. (A) Mean ACh activation over the classes. The mean ACh activations for corresponding classes are not statistically different for the two networks. Average of all classes is indicated as a dashed line, error bars show the variance. (B) ACh activation for stimuli labelled '7', sorted in increasing order. Values are the average over 50 training episodes; the shaded area is the variance. Dotted line is the mean over all stimuli. Although the mean ACh activations are not statistically significantly different from each others, the distributions of activations within a class largely differ. (C) Mean ACh activation for stimuli labelled '7' over the training episodes.

markedly non-uniform, in addition to varying across episodes.

Visualising the data in two dimensions reveals that the strengths of ACh activation are clustered in regions of the input space (Fig. 4.11, dimension reduction performed using t-Distributed Stochastic Neighbour Embedding (t-SNE) [MH08]). This clustering of ACh activation appears to disrupt learning by pulling weights away from the centre of mass of a datapoint cluster.

4.4 Signal of Stimulus Relevance

The cognitive construct of attentional effort is thought to comprise both the demand of a task and an animal's motivation to perform this task [SGK06]. In this section, I simulate the latter aspect of motivation in the model of ACh. In particular, I assign varying degrees of "relevance", or weights, to the integer categories. This relative relevance here is understood as reflecting the network's motivation to perform on the different classes.

I consider two perspectives on asymmetrical class relevance: either classes have a balanced distribution in the dataset but dissimilar relevance or, alternatively, classes have an unbalanced distribution but equal relevance. In both cases, relevance is asymmetric relative to the data distribution.

4.4.1 Non-uniform Stimulus Relevance

In a first approach, I assign relevances r_m to the various classes in decreasing arithmetic progression such that stimuli labelled '0' have a relevance value of '10' while stimuli

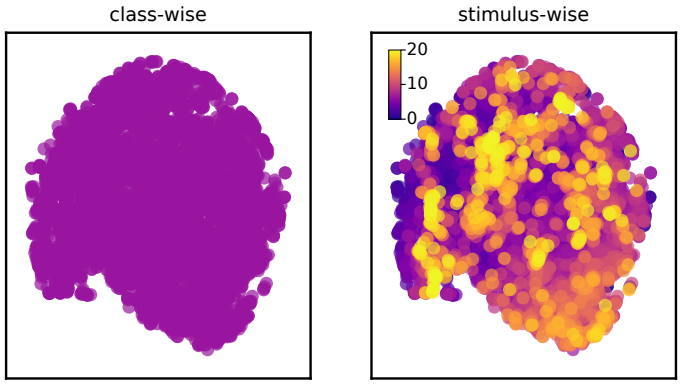

Figure 4.11: Spatial distribution of ACh activation. 2-dimensional representations of all images labelled as '7' (dimension reduction performed using t-SNE [MH08]). The colour axis shows ACh activation triggered by individual stimuli (averaged over 50 episodes). ACh activation in uniformly distributed within a class in the case of class-wise ACh computation but not in the case of stimulus-wise computation. The non-uniform distribution in the latter case appears to disrupt learning, probably by pulling weights away from the centre of mass of the datapoint cluster.

carrying a label '9' have a relevance of '1' (see Fig. 4.12 A). Formally, we can write:

$$r_m = M - m\,, \tag{4.6}$$

where M is the number of class labels (10 in this case).

I model ACh activation simply by setting the value of the ACh variable in Eq. 3.26a to the relevance of the class of the current stimulus (the class is determined by the true label of the stimulus):

$$ACh = r_m\,. \tag{4.7}$$

Results

The release of ACh reorganises weights such that more neurons represent digits with greater relevance (Fig. 4.12 B). This effect translates to a drop in the error rates for more relevant classes, and vice-versa (Fig. 4.12 C), resulting in a strong negative correlation between performance and relevance ($r = -0.87$).

In a similar manipulation, I pair a single stimulus class with varying levels of relevance (other classes have $ACh = 1$). This protocol augments the representation of the paired stimulus and improves accuracy for this class, in line with biological observation [Ree+11; Wei03; PSM06] (Fig. 4.13). ACh thus boosts classification accuracy for relevant classes.

4.4.2 Non-uniform Stimulus Distribution

As an alternative take on stimulus relevance, I investigate the impact of ACh on a dataset with uniform class relevance but unequal data distribution. In more details, I

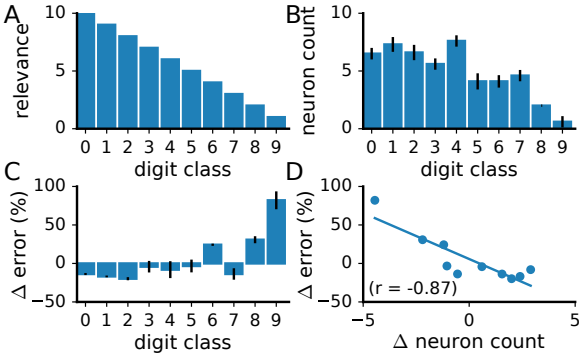

Figure 4.12: ACh enhances classification accuracy for relevant stimulus classes. (A) I assign different 'relevance' or 'motivational' weights to the digits. ACh activation is set to these values at stimulus presentation. (B) Distribution of stimulus preferences after training with ACh. More units prefer the more relevant digits. (C) Changes in error rates on the different classes. (D) Changes in error rates as a function of changes in neuron count. The redistribution of class preferences orchestrated by ACh yields gains in performance on the more relevant classes and decay on the less relevant classes. Data are the average of 10 runs, error bars indicate variance.

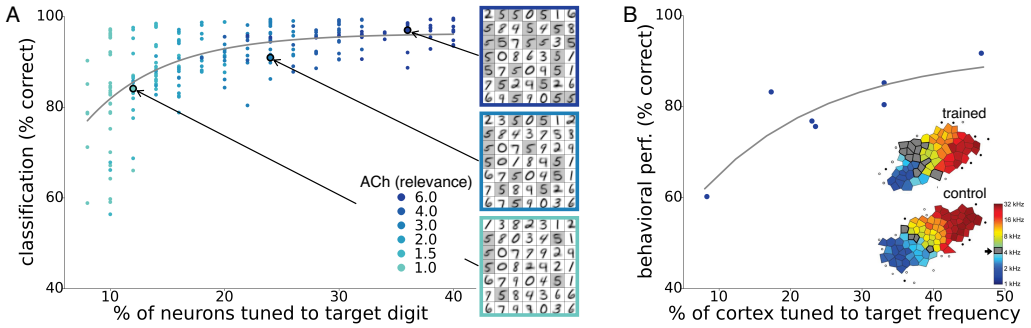

Figure 4.13: Accuracy correlates with the extent of a stimulus representation in the model and in biology. (A) Pairing a stimulus class with greater class relevance (i.e., ACh value) expands its representation in the model, in turn yielding enhanced performance on this class. (B) In rats, extensive frequency discrimination training improves behavioural performance and augments the representation of the target frequency in the cortex (from [Wei03]). Inset: A1 tonotopic maps of control and trained rats. The colour of each polygon depicts the best frequency associated with neurons at that position in the map. Gray polygons indicate recording sites with best frequencies within 0.375 octaves of the trained frequency (from [PSM06]). The cortical expansion of the target stimuli may derive from a greater behavioural relevance of those stimuli learned during training and signalled by ACh. In both model and biology, performance correlates with the extent of the representation occupied by relevant stimuli.

make use of a modified version of the MNIST dataset in which some of the classes are over-represented: the training data contain classes '0', '2', '3', '5', and '8', and there are 60 times more '0's and '2's than examples of the other classes. To model equal class relevance, I take the test dataset to be uniformly distributed over the classes.

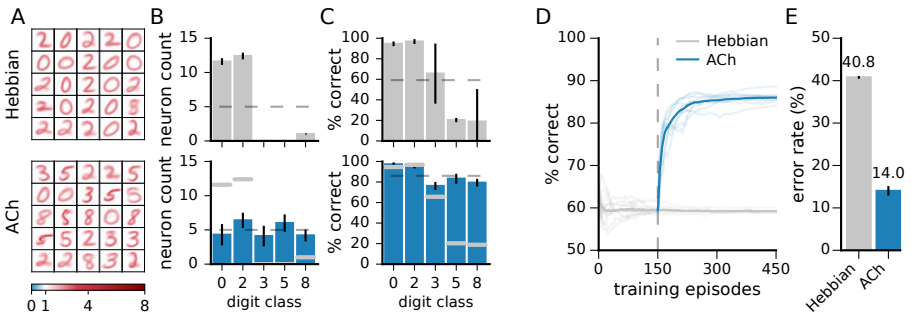

Figure 4.14: ACh activation greatly boosts performance on the non-uniform MNIST dataset. (A) Weights of the networks. The colour bar indicates weight strength. (B) Distribution of the neurons' preferred digit classes. Dashed line is a uniform distribution. (C) Correct classification on the test dataset. Dashed line is the mean over the classes. (D) Progression of performance on the test dataset. Light traces are individual runs, dark traces are averages over 10 runs. (E) Final test error rate. For all bar plots, the data are the mean of 10 runs, error bars indicate the variance, and grey traces are data for the Hebbian network for comparison.

Results

Without neuromodulation, the network performs poorly on the under-represented integer classes as it dedicates only few or no neurons to those digits (Fig. 4.14). Including ACh neuromodulation yields large gains in performance. As with the standard MNIST dataset, ACh carries its effect by attributing more neurons to classes on which performance is low (here, those that are under-represented). In this case, however, because of the markedly non-uniform data distribution, the impact of ACh is highly consequential.

4.5 Additional Effects on Learning

In the previous sections, I showed that a model of ACh signalling significantly improves the quality of a neural representation in two conditions, either when task difficulty is uneven among classes or again if stimuli bear unequal relevance. In this section, I examine how other variables may influence this learning mechanism, namely the number of neurons in the network and the availability of labels.

4.5.1 Network Size

ACh carries its beneficial effects by modifying the number of units encoding each stimulus class. The total number of units of a network may thus alter the influence of ACh. For large network sizes, we may expect ACh to bear minimal benefits as plenty of units represent each class. To examine this hypothesis, I train networks containing a different quantity of representation neurons with ACh-based learning. I perform this experiment for both the full MNIST dataset and its non-uniform version.

Results

Fig. 4.15 illustrates the influence of network size on performance. Larger networks fare better, both for Hebbian and ACh learning. However, as anticipated, ACh carries

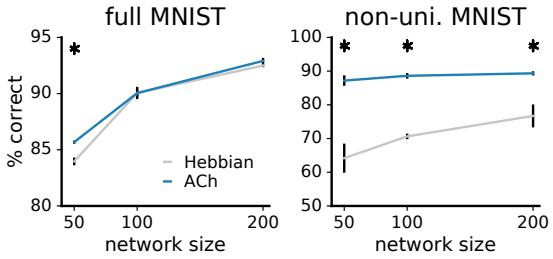

Figure 4.15: The benefits of ACh decrease with network size. Network performance as a function of the number of neurons in the representation layer. Data are the average of 5 runs, error bars indicate the variance. Asterisks indicate statistically significantly greater accuracy for ACh than for Hebbian learning ($p < 0.01$).

a weaker effect on accuracy in large networks. For the original dataset, the mean performances for ACh and Hebbian learning are not statistically significantly different for network sizes other than the smallest ($p > 0.01$). For the non-uniform dataset, the gains in accuracy deriving from ACh drop by half as network size increases from 50 to 200 units. These results highlight the role of ACh in reassigning the *limited* computational resources of a network. If a network holds a plenty of representation units, the worth of ACh diminishes. Note, however, that whether of network is adequate in size is relative to the number of target classes as well as to the variability of stimuli within these classes. As a result, the benefits of ACh should be restored in a large network facing substantial stimulus variability.

4.5.2 Label Availability

In section 4.2, I tested two methods to reduce dependancy on external labels, namely using inferred instead of true stimulus classes and classification confidence instead of performance. Even in these cases, however, the methods requires labels to train the classifier in order to infer classes and measure confidence. Here, I examine the impact of label availability on learning by training networks with a varying fraction of the total labels, from 100% down to 0.1% of labels.

Results

The accuracy of the network decreases with label scarcity, both when trained with Hebb's rule and ACh (Fig. 4.16). For the Hebbian network, labels only ever affect the classification layer; the decay in performance therefore derives exclusively from lower classifier accuracy. Although rendition also decreases with ACh, the signal yields improvements that are of similar magnitude for all label fractions. The constant improvements over declining label availability suggest that although the classifier's precision degrades with rare labelling this lower precision does not impact ACh learning. Therefore, in effect, ACh learning only minimally relies on labels. Together, these results indicate that ACh signalling is beneficial even for scarcely labelled data.

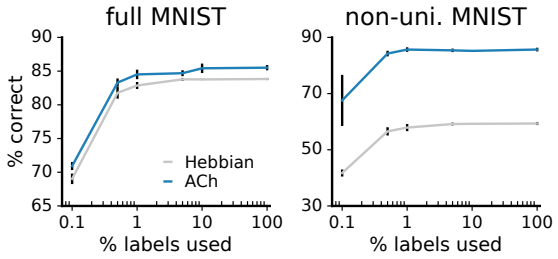

Figure 4.16: ACh benefits performance even with very few available labels. Network performance as a function of the percentage of training labels used. ACh activation relies on inferred stimulus class and classification confidence. Data are the average of 3 runs, error bars indicate the variance.

4.6 Discussion

As an animal detects a decline in its performance, it initiates effortful attentional processes to prevent further deteriorations [SGK06]. This surge in attentional effort is paralleled by heightened activation of cholinergic neurons in the basal forebrain [Arn+02; HSB00; Dal+01; Pas+00; McG+02; KBS06] which send a modulatory signal to the cortical mantle [HS11]. For instance, engaging in a demanding motor [CKT10] or tactile task [BTD97] enhance ACh release in the motor and somatosensory cortices, respectively.

There is broad evidence that ACh acts as a permissive plasticity agent at its projection sites [Buc+10; GS10]. In sensory cortices, for instance, ACh transmission promotes alterations of neural representations [BAS92; GLS88; Chu+13; JGS01; MS05; Sug12; KM98a; KM98b]. The scientific literature contains several hypotheses regarding the functional role of the modifications elicited by ACh. Froemke and colleagues [FMS07] suggest that shifts in neural tunings toward a stimulus paired with ACh activation serves as a long-term enhancement of attention to this stimulus. Others propose that this modification stores the behavioural relevance of the stimulus [KM98a; Wei03] or generally improves signal processing [Fro+13; Gu03]. In this chapter, I show that a signal modulating synaptic plasticity as a function of task difficulty or stimulus relevance improves the quality of a neural representation with respect to a classification task. The gains in performance result from assigning more neurons to challenging or relevant stimulus classes. The results presented above suggests that ACh serves this role in mammalian cortices.

Experimental evidence offer support to this hypothesis. For instance, motor skill acquisition and the accompanying enlargement of relevant representations in the motor cortex require ACh transmission [Con+03; CKT10]. Conversely, discrimination abilities improve for a tone whose representation is expanded as a result of repeated pairing with ACh activation [Ree+11]. More generally, ACh antagonists or lesion of the cholinergic system impairs perceptual [WFS04; FW02; BH95; Lea+13] and motor skill learning [Con+03]. These results indicate that the cholinergic system is crucial for forms of learning involving modifications in sensory maps, especially those affecting the relative extent of cortical representations.

In the model, the gains in performance brought fourth by ACh are relatively modest on the full MNIST dataset, amounting to a relative decrease of 12% in the error rate. However, these improvements take place even for negligible data labelling, indicating

that ACh signalling is helpful even with limited environmental feedback. For the non-uniform MNIST data, ACh brings substantial benefits, cutting down the error rate by two-thirds. This result in turn argues for a particularly valuable role of ACh when data follow an largely uneven distribution.

The worth of ACh diminishes with the size of a neural representation, underlining its role in redistributing a limited computational allowance. The brain relies on a large but finite neural population to represent sensory signals. Those signals contain considerable variance and potentially fall in infinitely many perceptual classes. In such a case where stimulus variability exceeds representational capacity, ACh is expected to bear significant benefits.

Functionally, the ACh learning mechanism is reminiscent of some boosting methods in machine learning, for instance AdaBoost [FS95]. Here, the learning algorithm attributes greater weights to misclassified training examples such that the learner is forced to focus on difficult examples. In contrast, however, other machine learning algorithms proceed in the opposite manner, for instance by identifying noisy and detrimental data and under-sampling them [SM14]. These two objectives appear contradictory–difficult data are either made more relevant or ignored–but both raise accuracy. This conflict may explain the poorer performance of the ACh model with stimulus-wise activation. In this case, ACh may attribute a too great importance to difficult but detrimental data points that should rather be under-sampled, thereby corrupting learning.

In this chapter, I examined two approximations of task difficulty, maximal posterior and best versus second-best posteriors (see Fig. 4.6). These two estimates provide classification accuracies that are not statistically significantly different from each other. Further examination of the data indicates that most training examples elicit a large maximal posterior, a case where the maximal and BvSB posteriors are similar, explaining their comparable renditions.

4.7 Summary

The mammalian brain discharges cholinergic signals that reflect attentional efforts; these discharges elicit alterations in sensory representations. In this chapter, I presented a model for ACh to examine its functional role in representation learning. I showed that a signal related to the difficulty or relevance of stimulus classes benefits representations by raising the computational resources dedicated to those classes. More units better represents variations in classes resulting in greater classification accuracy. This mechanism proved particularly helpful when data are largely non-uniformly distributed and performs well even when data labelling is scarce. Future work should address the poorer performances resulting from stimulus-wise ACh activation and examine the possibility that some difficult stimuli should rather be under-sampled.

Chapter 5

Model of Dopamine

5.1 Overview

In their "struggle for life", organisms universally seek nutrients and reproduction. As humans, for instance, we experience a desire for food and water as well as for sexual contact and parental investment. These stimuli are intrinsically pleasurable and form so-called primary rewards. The value of other rewards, like money, is learned. The subjective value an individual attaches to these percepts is a crucial factor in guiding decisions and behaviour. It determines, for instance, the amount of effort one is willing to expend or the risks one is willing to take to obtain a good. Economists represent subjective value with the mathematical construct of utility; the mammalian brain represent it in the firing of neurons in the midbrain. In particular, DAergic transmission originating in the VTA encodes errors in the prediction of the utility, or reward value, of a stimulus.

Through the diffuse and widespread projections of the VTA, errors in reward predictions attain most brain regions where they modulate neural firing and orchestrate learning processes. DAergic signalling contributes to multiple forms of error-based learning, including reinforcement learning, rule learning, and sensory discrimination learning. In this chapter, I investigate if such signals may also aid sensory representation learning.

As a mean to assess the potential benefits of RPE signals on representation learning, I design a model for the physiological release of DA in the neural network. In this model, the network receives rewards for correct classification decisions and makes reward predictions based on its exploratory behaviour. The difference between the predicted and received rewards, the RPE, activates DAergic transmission. This signal modulates the learning rate of the neural network.

I investigate three approaches to exploration and reward predictions. In a first approach, the network always makes the classification decisions it regards as optimal, known as a 'greedy' behaviour. Because of this greedy behaviour, the network is said to predict a reward on every trial. In a second approach, the network takes explorative classification decisions. In this case, the network makes binary reward predictions based on its exploratory behaviour: it predicts the presence or the absence of a reward on exploitative or explorative trials, respectively. Finally, I consider an approach in which the network's reward predictions are taken as the expected values of the rewards, quantified using the classifier's posterior probability.

5.2 Greedy Model

Decision making problems require a method to select one option amongst those available. In the absence of information about the options, an agent's only choice is to act randomly. If, on the other hand, an agent can estimate the values of the different alternatives, it may use this knowledge to guide its decision. The agent could then 'exploit' the environment, choosing the option presumed optimal, or 'explore' the other alternatives. In general, decision methods aim at balancing exploitation and exploration to maximise returns in unknown or changing environments.

One of the simplest approaches to decision making is that of a greedy strategy. A greedy algorithm takes decisions that are locally optimal at each step with the expectation that these decisions will lead to a global optima. Because of their lack of exploratory behaviour, however, greedy algorithms usually converge to local optima. Nonetheless, they may find a good approximation to the global optimum. An intuitive example of a greedy algorithm is to always travel in the direction of a geographic target although this strategy might lead to an impasse like a river. Interestingly, most currencies have unit subdivisions (e.g., the Euro with 1, 2, 5, 10, 20 and 50 cents) such that, when handing back change, a greedy approach is optimal. In a first model of DA transmission, I make use of such a greedy algorithm to guide the network's classification decisions.

5.2.1 Model Description

Classification Decision

As in the previous chapter, the network's classification decisions are determined using the classifier's posterior probabilities. On each trial, the network computes the posteriors t_k for each output class k (Eqs. 3.22 and 3.23). The class with the highest posterior is the presumed best option; the network following a greedy strategy always chooses this option as its output:

$$\hat{m} = \text{argmax}_{k=1}^{K}(t_k)\,. \tag{5.1}$$

Reward Delivery

In this work, only two reward sizes are delivered, $R = \{r_+, r_0\}$. Positive rewards r_+ are delivered deterministically on every trial on which the network makes a correct classification. The rewards are of arbitrary but identical sizes; for simplicity, I assume $r_+ = 1$. Null rewards $r_0 = 0$ are delivered when the network makes an incorrect classification. We can write:

$$R = \begin{cases} r_+ = 1 & \text{if } \hat{m} = m \\ r_0 = 0 & \text{otherwise}\,. \end{cases} \tag{5.2}$$

Reward Prediction

In mammals, activation of the midbrain DAergic system signals the difference between a predicted and a received reward. In particular, the magnitude of the predicted

reward represents its expected value [TFS05], reflecting both the size and probability of a reward. To remain in line with this observation, I assume for the model that reward predictions P are given by the expected value of the reward R:

$$P := \mathrm{E}[R] = \sum_{i=1}^{\infty} r_i \cdot \Pr(r_i)\,, \tag{5.3}$$

where r_i is a realisation of the reward R. As a result of the binary reward values, Eq. 5.3 simplifies to:

$$\begin{aligned} P := \mathrm{E}[R] &= r_+ \cdot \Pr(r_+) + r_0 \cdot \Pr(r_0) \\ &= \Pr(r_+)\,. \end{aligned} \tag{5.4}$$

In the last section of this chapter, I make use of the classifier's posteriors to approximate the probability $\Pr(r_+)$. In this section and the following one, I consider a simplification by which $\Pr(r_+)$ is determined by the network's exploratory behaviour. Specifically, the reward probability is approximated to be one on exploitative trials and zero on explorative trials:

$$\Pr(r_+) \approx \begin{cases} 1 & \text{if } \hat{m} = \mathrm{argmax}_{k=1}^{K}(t_k) \\ 0 & \text{otherwise}\,. \end{cases} \tag{5.5}$$

This discretisation of reward probabilities results in only two possible values for reward predictions, $P = \{p_+ = 1,\, p_0 = 0\}$. For the greedy model, because the network only takes exploitative decisions, its reward predictions are always of one: $P = p_+ = 1$.

Reward Prediction Error and Dopamine Activation

The difference between a predicted and delivered reward gives rise to a RPE. In the greedy approach, there are only two discrete possible RPE scenarios: either the network predicts a reward and is rewarded (p_+, r_+), or the network predicts a reward but is not rewarded (p_+, r_0). RPEs affect the learning rate of the network through the DA variable in Eq. 3.26b. In each of the two RPE cases, the value of the DA variable takes a constant value $\delta^g_{./.}$:

$$DA = \begin{cases} \delta^g_{+/+} & \text{if } p_+ \text{ and } r_+ \\ \delta^g_{+/0} & \text{if } p_+ \text{ and } r_0\,, \end{cases} \tag{5.6}$$

where the superscript g denotes parameters for the greedy model. The values of the $\delta^g_{./.}$ constants are determined through grid search to maximise the rate of correct classification.

5.2.2 Results

Parameter Search

Through 2-dimensional parameter sweep, I find that DA activation should be close to zero when the network makes a correct reward prediction ($\delta^g_{+/+} = 0.1$) and negative when the network expects a reward but does not get rewarded ($\delta^g_{+/0} = -1.0$, see

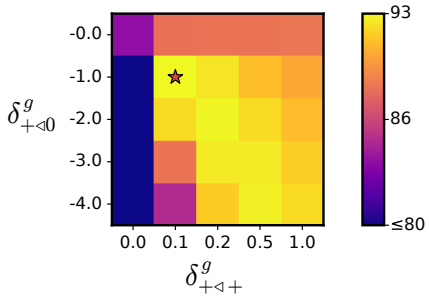

Figure 5.1: Parameter search for the DA values in the greedy model. The optimal parameter set has $\delta^g_{+/+}$ small and positive and $\delta^g_{+/0}$ larger and negative. Elements on the diagonal of the grid search matrix give comparable classification performances as they have similar ratios. Data are the mean performances over five validation sets. The colour axis indicates rates of correct classification on the test MNIST data.

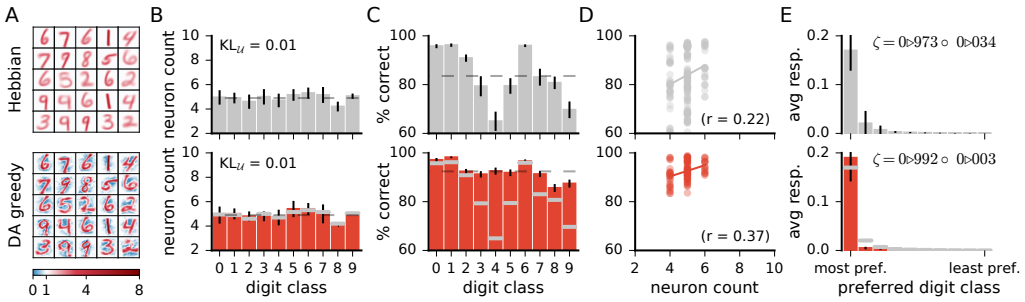

Figure 5.2: Network changes following training with DA in the greedy model. (A) Example weights of representation neurons (25 out of 49). (B) Number of neurons responsive to the different classes. Dashed line is uniform distribution. (C) Performance of the network on the different classes. Dashed line is average over all classes. (D) Performance on the different classes as a function of the number of neurons dedicated to this class. (E) Average response of neurons to the digit classes, ordered by class preference. (B), (C) and (E) are the mean of 10 runs, error bars indicate the variance. Grey overlaid bars in (B), (C) and (E) are values for Hebbian learning for comparison.

Fig. 5.1). The parameter search indicates that constant ratios of $\delta^g_{+/+}$ to $\delta^g_{+/0}$ give similarly good classification performances (more or less on the diagonal of the search matrix). This is so because scaling both $\delta^g_{+/+}$ and $\delta^g_{+/0}$ by the same constant is equivalent to scaling the learning rate of the network, small changes of which have little effect on the network's performance.

Changes in the Network

In contrast with ACh signalling, DA bears little effect on the number of neurons responsive to the different classes (Fig. 5.2 B). For both Hebbian learning and DA-based learning, the distribution of the neurons' preferred digit class is close to uniform ($\mathrm{KL}(Q_{Hebb}\,||\,\mathcal{U}) = 0.01$, $\mathrm{KL}(Q_{DA}\,||\,\mathcal{U}) = 0.01$). The positive correlation between neuron count and classification performance also remains after training with DA (Hebbian: $r = 0.22$, DA: $r = 0.37$).

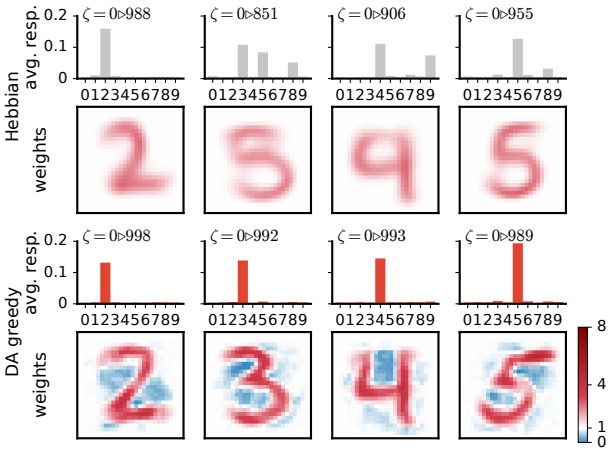

Figure 5.3: DA enhances class selectivity in neurons. Top rows: average responses of example neurons to the different integer classes. Bottom rows: depiction of the weights. The colour axis represents the values of the weights. While neurons in the Hebbian network respond to stimuli of multiple classes, those trained with DA respond almost exclusively to a single class.

Visual inspection of the weights suggests that DA makes neurons' weights more selective to specific digit classes. Consider the example weights shown in Fig. 5.3. Weights in one column are for corresponding neurons in a Hebbian and DA network (the networks were initialised with the same random seed). Weights in the Hebbian model are rather poorly tuned (e.g., the neuron resembling a '3', a '5', and an '8'). On the other hand, DA-based learning leads to well-defined digit representations. This observation can be quantified by measuring the average responses of neurons to the different classes (top rows in Fig. 5.3). This measure indicates that Hebbian learning yields neurons exhibiting strong responses to multiple stimulus classes, i.e., with a broad tuning. Training with DA yields more sharply tuned weights as units respond almost exclusively to a single digit category.

On average over all neurons, DA generates a 12% increase in the activations of neurons to their preferred classes, accompanied by a 58% reduction to non-preferred classes (Fig. 5.2 E). These modifications amount to neuron weights being more selective to specific integers, or having a sharper tuning. We can quantify such neural selectivity as the difference between a neuron's mean response to stimuli of its preferred class and its mean response to stimuli of all other classes:

$$\zeta_c = \frac{\bar{s}_c^{\bullet} - \bar{s}_c^{\circ}}{\bar{s}_c^{\bullet}}, \tag{5.7}$$

where $\bar{s}_c^{\bullet}$ and $\bar{s}_c^{\circ}$ are the average responses of neuron c to stimuli of its preferred and non-preferred classes, respectively. Here, $\zeta_c = [0, 1]$, where $\zeta_c = 0$ is a neuron that responds equally strongly to all stimuli (non selective) and $\zeta_c = 1$ is a neuron that responds exclusively to one digit category (perfect selectivity). The selectivity of individual neurons are indicated on Fig. 5.3; the selectivity averaged over all neurons, ζ, is indicated on Fig. 5.2 E. This measure indicates that DA statistically significantly increases neural selectivity ($p < 0.001$).

We can also quantify a neural network's selectivity for a specific digit class m as

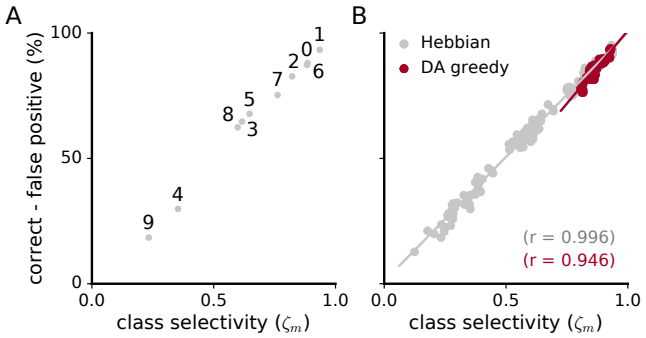

Figure 5.4: Neural selectivity strongly correlates with classification performance. (A) Difference between rates of correct and false positives for each digit class as a function of a network's selectivity for this class (Eq. 5.8). Data are for a single Hebbian neural network. (B) Same as (A) but for 20 networks, for both Hebbian- and DA-based learning. DA enhances neural selectivity, which translates in greater classification accuracy

the sum of the selectivity of the neurons whose preferred stimulus class is m:

$$\zeta_m = \frac{1}{C_m} \sum_{c_m=1}^{C_m} \zeta_{c_m} \,, \tag{5.8}$$

where the index c_m indicates a neuron whose preferred stimulus class is m and C_m is the total number of such neurons. Fig. 5.4 depicts the relationship between a network's selectivity for a class and its performance on this class.

In the weights plotted in Fig. 5.3, we note that DA transmission leads to some weights collapsing to zero (indicated in blue in the figure, white background weights have a strength of 1). These weights correspond to input neurons whose activation was repeatedly paired with a misclassification. Such zero-valued weights are absent in the network trained solely with Hebbian learning.

We can further visualise the outcome of DA learning by reducing the the dimensionality of input images to 2 features (using t-SNE [MH08]) and train the network on these data (see Fig. 5.5). In Hebbian learning, the neural network acts as a clustering algorithm. As the learning mechanism is agnostic to the labels of the stimuli, the classification boundaries miss some aspects of the data classes. In particular, boundaries are poorly defined between close-by clusters such '3', '5' and '8'. Following DA signalling, weights adjust to match the boundaries for the conditions for reward delivery of the task.

Performance

DA induces large and statistically significant improvements in classification accuracy ($92.51 \pm 0.07\%$ for DA compared to $83.62 \pm 0.13\%$ for Hebbian learning, $p < 0.0001$), corresponding to halving the error rate, down by of 54%. Performance for a class strongly correlates with neural selectivity for this class, for both the Hebbian and DA networks ($r = 0.996$ and $r = 0.946$, respectively, see Fig. 5.4). This strong correlation suggests that enhanced neural selectivity explains the rise in correct responses following DA learning. The variance not explained by neural selectivity may derive from a

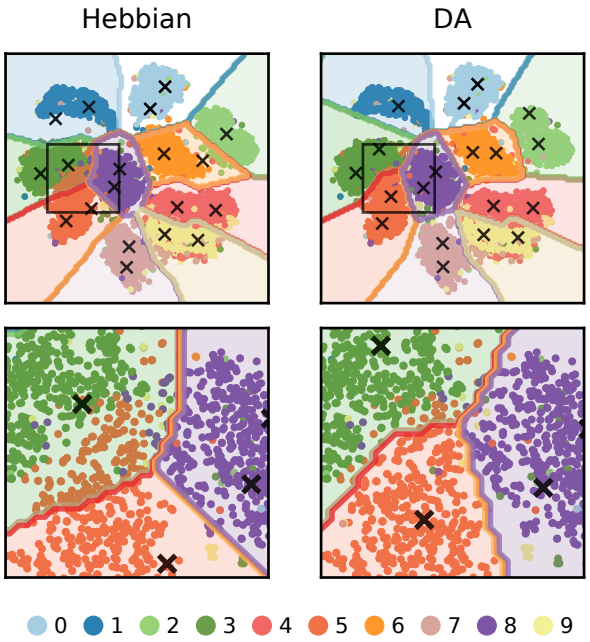

Figure 5.5: 2-dimensional visualisation of the outcome of DA learning. I reduce the dimensionality of input images from 784 to 2 features using t-SNE and a train network on these data. The input stimuli are represented as coloured dots and the weights of hidden neurons as black crosses. The classification boundaries are overlaid on top. In Hebbian learning, the network finds clusters in the data. For classes that are well separated from others, the network retrieves close to perfect boundaries (e.g., '1' or '0'). However, for classes with close-by perimeters (e.g., '3', '5', and '8', magnified below), the boundaries poorly match the true labels. In these cases, DA transmission matches the network's classification boundaries to the reward contingencies of the task.

population code: as neural selectivity is measured for each neuron separately, it is possible that neural selectivity increases also at a population level without affecting selectivity at the single neuron level.

5.3 Explorative Model

Most living creatures exhibit explorative behaviours: fungi send filament to explore distant nutrient sources [Wat+05], ants cooperatively survey a terrain to find a suitable location for their nest [PS06] while primates learn to forage for food from older animals [RB08]. Distinct from vain curiosity, exploration is critical to a creature's ability to survive and reproduce, for instance allowing to maximise the harvest of resources. In the present neural network simulation, exploration translates to sampling alternate classification decisions for a given input image. Learning from such exploration offers the prospect of escaping local optima in the weight space.

There are no optimal exploration policies, even when the objectives are well defined [CMA07]. Rather, various meta-heuristics exist, each with their own characteristics. These algorithms can generally be divided into two families: undirected and directed strategies. Undirected exploration selects suboptimal options randomly, with-

out any preference. Conversely, directed exploration makes use of selection criteria to prefer a suboptimal option over the others. A criterion may be the expected value of the available choices or uncertainty in the options' worth. Below I present three algorithms for exploration, one undirected the two others directed. I discuss possible realisations of these algorithms in a neural network and use this discussion to inform my design decisions in implementing exploration in the model.

5.3.1 Methods of Exploration

ϵ-greedy

The simplest techniques for exploration are those of undirected strategies, of which the ϵ-greedy algorithm is the best known example. Following this method, an agent selects the option considered the richest with probability $1 - \epsilon$ and an explorative decision with probability ϵ. This strategy is undirected in that, when making an explorative decision, the choice probability is uniformly distributed over all suboptimal options. The probability of selecting an option $\hat{m}_i$ is written as:

$$\Pr(\hat{m}_i) = \begin{cases} 1 - (K-1) \cdot \epsilon & \text{if } i = \operatorname{argmax}_{k=1}^{K} \mathrm{E}[\hat{m}_k] \\ \epsilon & \text{otherwise}\,, \end{cases} \tag{5.9}$$

where $\epsilon = [0, 1]$.

The ϵ-greedy method may be implemented in a neural network by modifying the activity of neurons either in the representation or in the output layer. In both cases, exploration would be achieved by setting the activation of a single neuron to one and that of all other neurons to zero. The probability of selecting a given neuron would be given by Eq. 5.9 where the expected value would be approximated by the posterior probabilities of the classifier: $\mathrm{E}[\hat{m}_i] \approx t_i$.

Performing such exploration in the output layer would have the most direct effect on the classification decision. However, this would not impact activation in the representation layer where the DA signal is available. If exploration and DA transmission occur in different layers, they cannot interact to direct learning. A solution would be to have the result of the exploration (i.e., the activation in the output layer) feed back to the representation layer, affecting activation and learning there. Alternatively, exploration could be implemented by modulating neural activity in the representation layer, affecting both the classification decision and learning in this layer.

Softmax

A second algorithm for exploration is that of the softmax rule [SB98]. In this rule, the choice probability is proportional to the option's expected value. In contrast with the ϵ-greedy algorithm, the softmax rule is directed as the probability mass is not uniformly distributed over the available options. Formally, the softmax rule is given as:

$$\Pr(\hat{m}_i) = \frac{\exp(\mathrm{E}[\hat{m}_i]/\tau)}{\sum_{k=1}^{K} \exp(\mathrm{E}[\hat{m}_k]/\tau)}\,, \tag{5.10}$$

where τ determines the breadth of exploration.

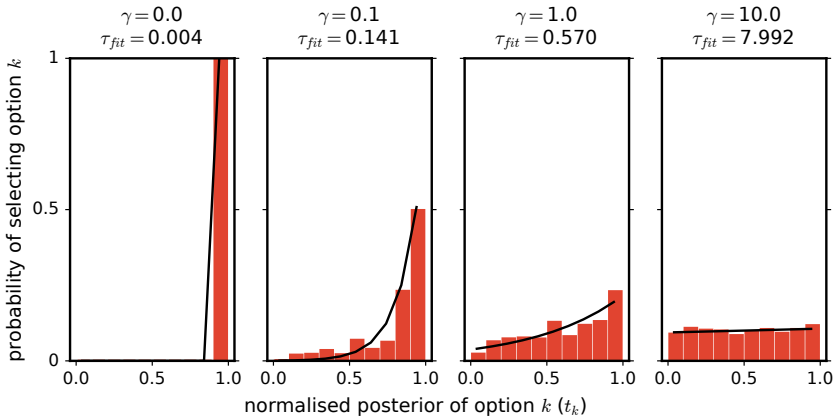

Figure 5.6: Noise injection in neural activity approximates the softmax rule. For this figure, I train a network through Hebbian learning and add noise in the activation of representation neurons following Eq. 5.11. The classification decision of the network is taken as $\text{argmax}_{k=1}^{K}(t_k)$ after noise injection (i.e., the greedy decision after noise injection). I measure the probability of selecting an option k and plot this as a function of the posterior probability t_k of this option prior to noise injection (red histograms). I repeat this measure for networks with different values of γ. The black trace is a softmax function (Eq. 5.10) with the parameter τ fitted to the empirically measured choice probability. Noise addition is a good approximation to the softmax function. The variance of the injected noise, γ, is comparable to the temperature parameter τ of the softmax function: as the variance increases, the distribution of decision probability tends towards a uniform distribution.

In a neural network, the softmax algorithm may be realised explicitly by setting the activation of one neuron to one and all others to zero. Here, the probability of a neuron being activated would follow the expected reward value. Alternatively, this rule could be implemented implicitly by adding noise in the activation of representation neurons:

$$\begin{aligned} \tilde{s}_c &= \sum_{d=1}^{D} S(W_{cd}) y_d + \eta_c \,, \\ \eta_c &\sim \mathcal{N}(0, \gamma) \,, \end{aligned} \tag{5.11}$$

where $\mathcal{N}$ is a normal distribution and γ is a hyper-parameter determining the spread of this distribution. Adding independently and identically distributed noise to the activation of neurons results in a decision strategy approximating the softmax rule (see Fig. 5.6). In this case, the parameter γ corresponds to the temperature parameter τ of the softmax: for $\gamma \to \infty$, all classification decision have equal probabilities; for $\gamma \to 0^+$, the network is purely exploitative. For intermediate values, the network favours options with high expected values.

Softmax with Uncertainty

A third approach to exploration, also directed, is to prefer options that bear a high uncertainty in their expected value. For instance, this method could build on the

softmax rule to add a bonus to options based on the variance in the rewards [Daw+06]:

$$\Pr(\hat{m}_i) = \frac{\exp[(\mathrm{E}[\hat{m}_i] + \sigma^2(\hat{m}_i))/\tau]}{\sum_{k=1}^{K} \exp[(\mathrm{E}[\hat{m}_k] + \sigma^2(\hat{m}_k))/\tau]}\,. \tag{5.12}$$

This algorithm could be executed as the explicit realisation of the softmax rule with uncertainty taken as the variance in the rewards received after selecting option i for an inferred stimulus class k: $\sigma^2(\hat{m}_i^{(k)})$.

5.3.2 Model Description

Exploration

Of the exploration strategies described above, the softmax algorithm achieved through noise addition in the activation of representation units is the most straight-forward to execute. In contrast with the other approaches, this one does not require the explicit computation of probabilities, making it simpler and more biologically realistic. Furthermore, implementing this rule in the representation layer avoids the need for feedback projections from the classification layer. In addition, humans appear to make use of such a softmax rule during exploratory decision making [Daw+06]. For these reasons, I select this method for exploration in the model. A comparison of the performances of the alternate meta-heuristics may be the subject of future work.

The injection of noise in Eq. 5.11 does not necessarily produce explorative behaviour. To determine whether a trial is explorative or not, I compare the classification output without and with noise addition. Specifically, I compute the posteriors over all classes k both without and with noise addition, denoted t_k and t_k^η, respectively. I then compare $\mathrm{argmax}_{k=1}^{K}(t_k)$ and $\mathrm{argmax}_{k=1}^{K}(t_k^\eta)$. If these are the same, the trial is said to be exploitative; if they differ, the trial is said to be explorative. The classification output is taken as the class with the greatest posterior after noise injection, $\hat{m} = \mathrm{argmax}_{k=1}^{K}(t_k^\eta)$.

Reward Prediction

As for the greedy model, I make use of the approximation by which the network's exploratory behaviour determines its reward prediction:

$$\Pr(r_+) := \begin{cases} 1 & \text{if } \hat{m} = \mathrm{argmax}_{k=1}^{K}(t_k) \\ 0 & \text{otherwise}\,. \end{cases} \tag{5.13}$$

In other words, the network predicts a reward with a probability of 1 for exploitative choices and with a probability of 0 for explorative choices.

For the explorative model, there are four discrete RPEs scenarios deriving from the combination of each of the two reward predictions and reward deliveries (Fig. 5.7). These scenarios are that the network may (1) predict a reward and get rewarded (p_+, r_+), (2) predict a reward but not get rewarded (p_+, r_0), (3) not predict a reward but get rewarded (p_0, r_+), and (4) not predict a reward and not get rewarded (p_0, r_0). The variable DA in Eq. 3.26b takes a constant value $\delta^\eta_{\cdot/\cdot}$ depending on the RPE. We

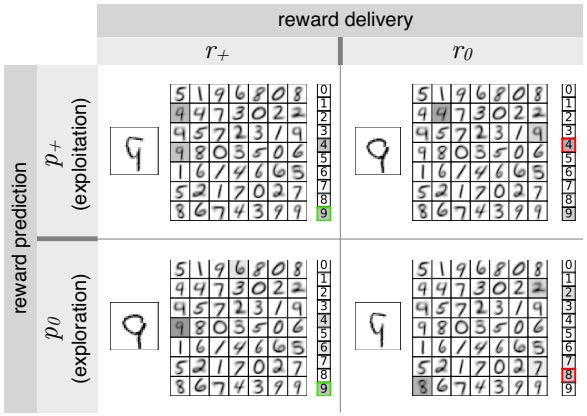

Figure 5.7: Four reward prediction error scenarios. The RPEs are the four possible combinations of reward predictions (p_+ and p_0) and reward deliveries (r_+ and r_0). The figure depicts a network in each of the four cases with the current input, the weights of 49 representation neurons and the 10 classification units. Grey highlights indicate neural activity, green and red edges indicate correct and incorrect classifications, respectively. Note that networks on the diagonals receive the same input but that noisy activations in the representation layer culminate in different outputs. In each of the four scenarios, DA in Eq. 3.26b takes a distinct value $\delta^{\eta}_{./.}$.

can write:

$$DA = \begin{cases} \delta^{\eta}_{+/+} & \text{if } p_+ \text{ and } r_+ \\ \delta^{\eta}_{+/0} & \text{if } p_+ \text{ and } r_0 \\ \delta^{\eta}_{0/+} & \text{if } p_0 \text{ and } r_+ \\ \delta^{\eta}_{0/0} & \text{if } p_0 \text{ and } r_0\,, \end{cases} \tag{5.14}$$

where $\delta^{\eta}_{./.}$ are constants whose values are determined through parameter search to maximise classification performance (the superscript η indicates parameters for the explorative model).

5.3.3 Results

Parameter Search

I perform parameter exploration to identify the optimal values of the four $\delta^{\eta}_{./.}$ constants and of the exploration parameter γ through 5-dimensional grid search. Fig. 5.8 shows the results of this search. The optimal γ is of 0.3 (compare with Fig. 5.6). In the case of the $\delta^{\eta}_{./.}$ constants, I find that for surprising rewards (p_0, r_+) the optimal $\delta^{\eta}_{0/+}$ is positive while in the absence of an expected reward (p_+, r_0) the optimal $\delta^{\eta}_{+/0}$ is negative. For correctly predicted rewards (either p_+, r_+ or p_0, r_0) the optimal $\delta^{\eta}_{+/+}$ and $\delta^{\eta}_{0/0}$ are close to zero. This parameter set is in close correspondence with the release schedule of DA in mammals (see Fig. 5.9).

Changes in the Network

Allowing the network to take explorative decisions alters the weights in a comparable manner as for the greedy model (Fig. 5.11). Namely, neurons see their activations rise

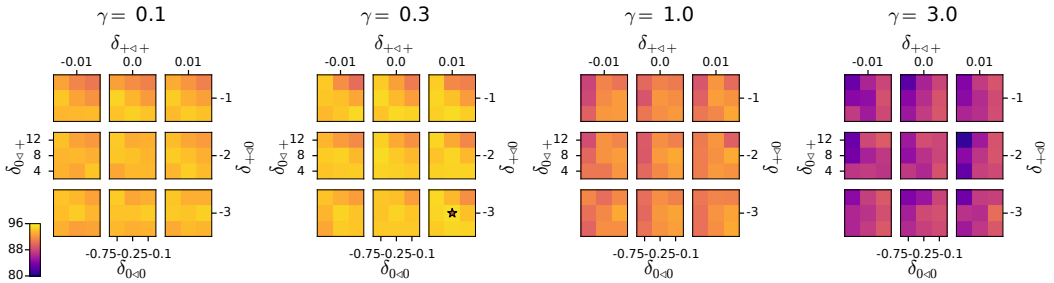

Figure 5.8: Parameter search for γ and the $\delta^{\eta}_{./.}$ constants in the explorative model. 5-dimensional grid search: each plot is for a different value of the exploration parameter γ, containing the results of the exploration for the $\delta^{\eta}_{./.}$ constants. The coloured axis indicates the rate of correct classification (averaged over 3 runs). A star indicates the the best parameter set with respect to classification performance.

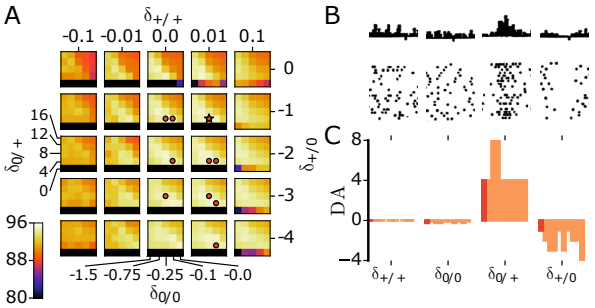

Figure 5.9: Detailed parameter search for the $\delta^{\eta}_{./.}$ constants. (A) Exploration of the four $\delta^{\eta}_{./.}$ constants through exhaustive search (coloured axis indicates performance, averaged over the 5-fold cross-validation runs). A star indicates the best parameter set, dots indicate parameter sets yielding accuracies not statistically significantly different from that of the best set ($p > 0.01$). (B) Firing of dopaminergic neurons in monkeys in reward prediction error scenarios homologous to those of the explored $\delta^{\eta}_{./.}$ constants (modified with permission from [SDM97]). (C) Bar plot of the best parameter set (dark red) and sets not significantly different from best (light red). The parameter sets are sorted in decreasing order of their classification accuracies from left to right.

for the preferred integer and drop for all others (by +17% and −84%, respectively). These modifications are more pronounced than in the greedy approach (+12% and +58%, respectively). Visually, the neurons' weights are noisier than for the greedy strategy but still clearly resemble specific digits (Fig. 5.11). I present a further comparison of the models in Sec. 5.5.

Performance Measure

The network's classification performance increases further, reaching a correct classification rate of $95.57 \pm 0.05\%$ in comparison with $92.71 \pm 0.07\%$ for the greedy approach.

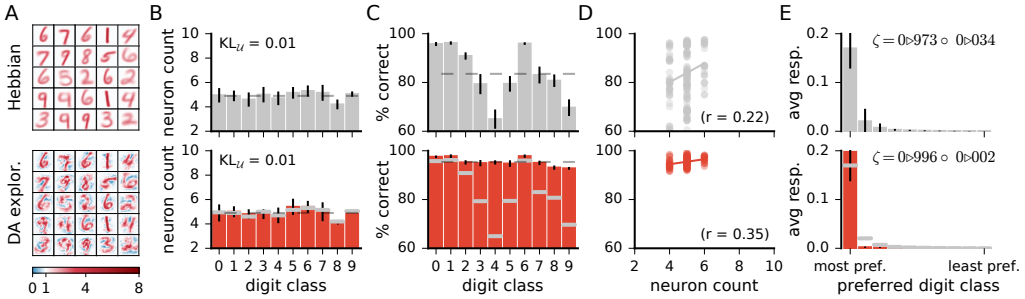

Figure 5.10: Network changes for DA with exploration. (A) Example weights of representation neurons. (B) Number of neurons responsive to the different classes. Dashed line is uniform distribution. (C) Performance of the network on the different classes. Dashed line is average over all classes. (D) Performance on the different classes as a function of the number of neurons dedicated to this class. (E) Average response of neurons to the digit classes, ordered by class preference. (B), (C) and (E) are the mean of 10 runs, error bars indicate the variance. Grey overlaid bars in (B), (C) and (E) are values for Hebbian learning for comparison.

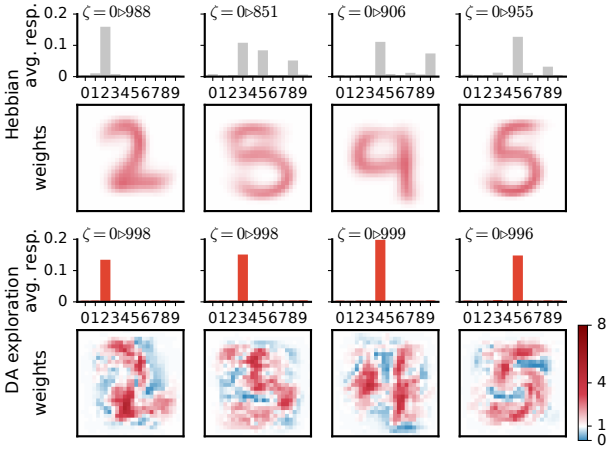

Figure 5.11: DA in the explorative model boosts class selectivity. As for DA in the greedy approach, DA with exploration leads to weights that are better tuned to specific integers.

Interpretation of the Parameter Search Results

To interpret the results of the parameter exploration, consider the input images $\vec{y}^{(n)}$ and the weights of neurons $\vec{W}_c$ as vectors in a high-dimensional space (here, a 784-dimensional space, for each input pixel). The activations of neurons s_c are computed as the dot product between an input and the weight vectors; the smaller the difference between these vectors, the stronger the activation. Lateral inhibition introduces a soft winner-take-all competition resulting in a few neurons having a strong response and other neurons being silent. Hebbian learning then induces weight modifications $\Delta\vec{W}_c = DA \cdot \epsilon \cdot s_c(\vec{y} - \vec{W}_c)$ (Eq. 3.26b). We note that, for each weight, $\Delta\vec{W}_c$ points from the weight towards the current input. The variable DA modulates the magnitude of $\Delta\vec{W}_c$, $\|\Delta\vec{W}_c\|$.

DA takes a value $\delta^{\eta}_{\cdot/\cdot}$ specified by the current RPE scenario. According to the parameter search, for correct reward predictions (p_+, r_+ or p_0, r_0), the optimal $\delta^{\eta}_{+/+}$

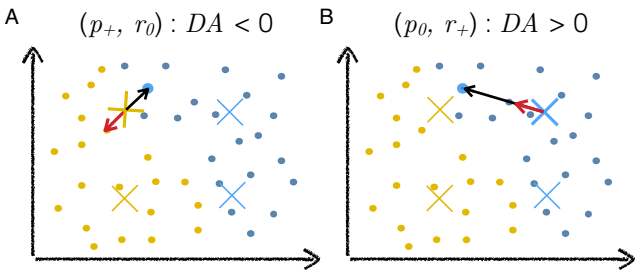

Figure 5.12: Cartoon explanation of DA-based learning. The plots represent a toy example of a two-dimensional input space with dots as input examples and crosses as neuron weights, the colours of which indicate classes. The highlighted blue training example is the current input to the network. Black arrows depict the vector $(\vec{y} - \vec{W}_c)$ which determines the direction of the weight change vector $\Delta\vec{W}_c$. Red arrows depict the weight change vector $\Delta\vec{W}_c$, including the multiplication by the DA variable. As a scalar multiplier, DA only affects the magnitude (and sign) of $\Delta\vec{W}_c$ and leaves its direction unchanged. (A) The network makes an incorrect greedy decision: the blue input activates a yellow cross, no reward is delivered. In this case, $DA < 0$. This negates the direction of $\Delta\vec{W}_c$, moving the active weights away from the training example (red arrow). (B) The networks makes an correct explorative decision: noise injection leads to another neuron (blue) being more active. In this case, the value of DA is positive, moving the weight towards the current input (red arrow).

and $\delta^{\eta}_{0/0}$ are both of approximatively zero. In both these cases, $\|\Delta\vec{W}_c\| \approx 0$; all the network's weights remain unchanged.

When the network takes an exploitative decision that turns out to be wrong (p_+, r_0), the optimal $\delta^{\eta}_{+/0}$ is negative. The vector $\Delta\vec{W}_c$ is negated so that it points away from the current input (Fig. 5.12 A). Active neurons will have their weights move away from the input; these neurons are less likely to win the softmax competition at future presentations of the same data point.

When the network takes an explorative decision that is surprisingly correct (p_0, r_+), the optimal $\delta^{\eta}_{0/+}$ is positive. The weights of active neurons move towards the input (Fig. 5.12 B). The explorative decision (expected incorrect) turned out to be right; this decision should be taken again on future presentation of the same stimulus. DA-based learning can be understood as reinforcement learning at the level of sensory representations.

5.4 Expectations as Reward Predictions

The mammalian brain represents a payoff's expected value, the product of a reward's size and probability, in the phasic responses of DAergic cells of the VTA. For instance, in monkeys, an image predicting a juice reward of 0.15 ml delivered with certain probability elicits significantly greater activation in DAergic midbrain neurons than an image predicting the same juice volume but delivered with half the probability [TFS05]. Both the magnitude and the probability of a payoff determine its effective worth and, accordingly, both these values guide animal behaviour [RMR08].

If a reward's expected value is relevant to behaviour, it may also benefit repre-

sentation learning. In particular, in the model described above, reward predictions were binary, reflecting solely whether a decision was explorative or not. Identical DA transmission–and modulation of learning–took place regardless of how daring the network had been in its exploration. However, the extent of the exploration may be relevant to plasticity. For instance, an explorative decision with a very low probability of being correct that, extraordinarily, turns out right might require greater weight changes than an only mildly improbable decision which turns out right. Here, the probability of a predicted reward for a classification decision may serve to quantify the scale of exploration: the lower the reward probability, the more explorative the network is being. In this section, I test the hypothesis that modulating plasticity as a function of reward probabilities further benefits learning compared to binary reward predictions.

5.4.1 Model Description

As in the previous models, reward delivery takes two discrete values, $R = \{r_+ = 1, r_0 = 0\}$. Reward prediction, on the other hand, differs from the prior approaches. Here, I use the classifier's posteriors to approximate the reward probability. Specifically, as before, I compute the posteriors over all classes for the current stimulus both without and with noise injection, t_k and t_k^η, respectively. The classification output is taken as the class with the greatest posterior when noise is added, $\hat{m} := \text{argmax}_{k=1}^{K}(t_k^\eta)$. The reward probability is then approximated as the posterior prior to noise injection for the class chosen as an output:

$$\Pr(r_+) \approx t_{\hat{m}}\,. \tag{5.15}$$

The posterior probability $t_{\hat{m}}$ strongly correlates with the empirically-measured probability of being rewarded for choosing class $\hat{m}$ as an output ($r = 0.98$), with a fitted slope of $a = 0.98$ and an intersect of $b = 0.01$, validating the approximation $\Pr(r_+) \approx t_{\hat{m}}$ (see Fig. 5.13).

In the preceding explorative network, RPEs fell in one of four discrete cases, with the DA variable taking a distinct value in each case. Here, the RPE is continuous-valued, given by the difference between the continuous-valued reward prediction $P = [0, 1]$ and the discrete-valued reward delivery $R = \{0, 1\}$:

$$RPE = R - P\,, \tag{5.16}$$

with $RPE = [-1, 1]$.

I explore different functions mapping $RPE \mapsto DA$. The search for the optimal $\delta_{./.}^\eta$ constants revealed that DA should grow with RPE. Based on this observation, I consider two monotonously increasing functions of $RPE \mapsto DA$, a linear and an exponential function:

$$\text{linear} \qquad DA = \delta_1^l \cdot RPE + \delta_2^l \tag{5.17a}$$

$$\text{exponential} \qquad DA = \exp(\delta_1^e \cdot RPE) + \delta_2^e\,, \tag{5.17b}$$

where δ_1^l and δ_2^l, and δ_1^e and δ_2^e are parameters of the linear and exponential functions, respectively. I explore the values of these parameters to maximise classification performance.

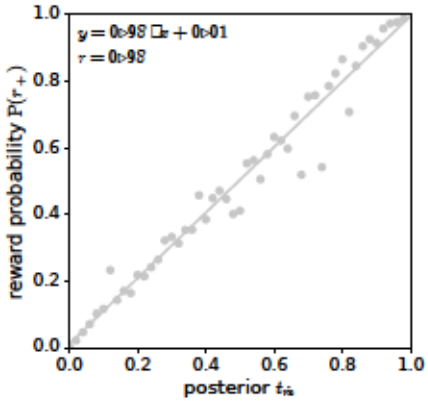

Figure 5.13: The posterior probability of the class selected as output strongly predicts the reward probability. For this figure, I train a network with exploratory behaviour implemented through noise injection in the activation of representation neurons. I keep track of the posteriors prior to noise injection of the class selected as output, $t_{\hat{m}}$. I bin these posteriors and compute the probability of reward delivery for each bin. This reward probability is plotted against the posteriors. The two variables exhibit a strong correlation, with a linear fit close to an identity function, validating the approximation of $\Pr(r_+) \approx t_{\hat{m}}$ (Eq. 5.15).

5.4.2 Results

Parameter Search

Through grid search, I find that a steeper slope (δ_1^l and $\delta_1^e > 0$) and a small negative y-offset ($\delta_1^l < 0$ and $\delta_1^e < -1$) yield the best performances for both the linear and exponential functions (Fig. 5.14). The exponential function produces better classification results ($93.74 \pm 0.09\%$ versus $92.78 \pm 0.80\%$) but this difference does not reach statistical significance ($p = 0.389$). Nonetheless, I make use of the exponential function for further analysis of the model.

Changes in the Network

Similar modifications take place here as for the network with binary reward predictions, albeit to a lesser extent. Specifically, neurons become more selective to a single digit class, some weights collapse to zero, and, visually, the weights appear somewhat noisy (Figs. 5.15 and 5.16). Quantitatively, the mean responses of neurons increases by 13% to their preferred classes and decrease by 64% for all other classes (compared to 17% and 84% for the discrete explorative model).

Performance Measure

DA with continuous reward expectations yields large improvements in performance compared to control, reaching correct classification rates of $93.74 \pm 0.09\%$. This is greater than the greedy DA model ($92.51 \pm 0.07\%$) but lower than the explorative DA model with discrete RPE scenarios ($95.53 \pm 0.05\%$), with all differences reaching statistical significance. I compare the different models in more details in the next section.

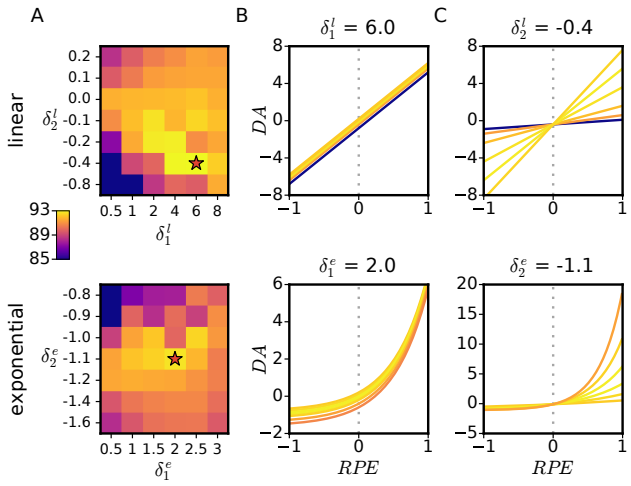

Figure 5.14: Parameter exploration for continuous reward expectations. (A) Results of the parameter sweep; coloured axis indicates rates of correct classification, a star indicates the best parameter set. (B) Example functions $RPE \mapsto DA$ for different δ_2 values, with δ_1 set to its best performing value. (C) Same as (B) but for different δ_1 values, with δ_2 set to its best performing value. The colour of the curve indicates the resulting performance. Data are for one network instance.

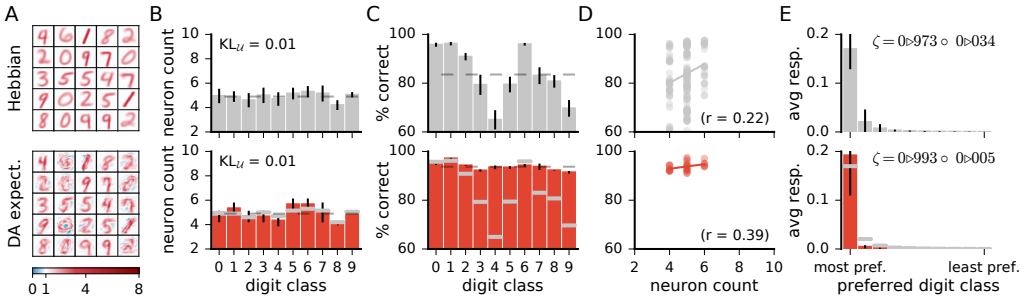

Figure 5.15: Network changes for DA with reward expectation. (A) Example weights of representation neurons. (B) Number of neurons responsive to the different classes. Dashed line is uniform distribution. (C) Performance of the network on the different classes. Dashed line is average over all classes. (D) Performance on the different classes as a function of the number of neurons dedicated to this class. (E) Average response of neurons to the digit classes, ordered by class preference. (B), (C) and (E) are the mean of 20 runs, error bars indicate the variance. Grey overlaid bars in (B), (C) and (E) are values for Hebbian learning for comparison.

5.5 Model Comparison

All models of DA lead to statistically significant enhancements in classification accuracy compared to Hebbian learning alone ($p < 0.0001$, see Fig. 5.17 A). The mean performance of the different models also all differ from one another ($p < 0.0001$). In this section, I attempt to explain these discrepancies.

5.5.1 Greedy versus Explorative

Across the different models, the mean selectivity of a neural population strongly correlates with the network's classification performance ($r = 0.998$, Fig. 5.17 C), indicating

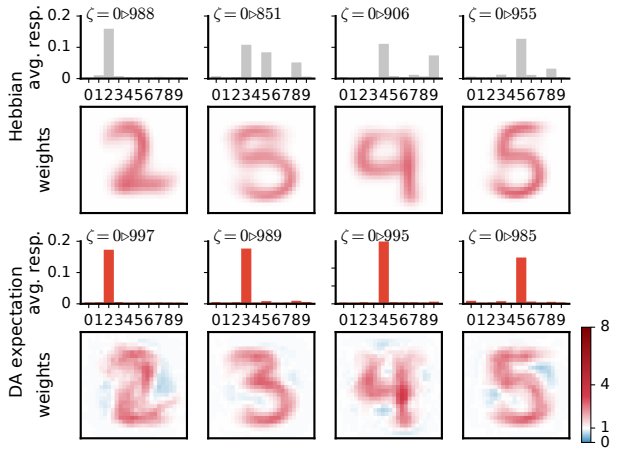

Figure 5.16: Weight comparison before and after DA with reward expectation. DA enhances neural selectivity.

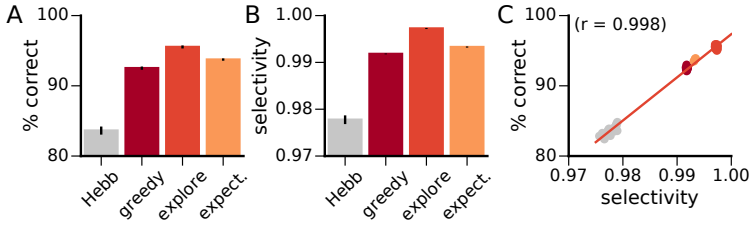

Figure 5.17: Performance comparison of the DA models. (A) Rate of correct classification. (B) Mean selectivity of representation neurons. (C) Correlation between mean neural selectivity for a network and rate of correct classification for this network. Each data point is one network; the means of 10 networks are represented in the bar plots. Error bars indicate the variance.

that discrepancies in neural selectivity may explain variations in performances. Of the different models of DA, the greedy approach yields the lowest performances. The greedy approach seeks to minimise miss-classifications and maximise correct classifications but does not explore alternate decisions. On the other hand, the explorative model seeks to minimise miss-classifications and to maximise explorative, unexpectedly correct categorisations. I attempt here to interpret quantitatively the role of exploration. To do so, I train a greedy and an explorative network with the additional constraint that DA activation takes place only for images classified as a target class m_t:

$$DA' = \begin{cases} DA & \text{if } \hat{m} = m_t \\ 0 & \text{otherwise}\,, \end{cases} \tag{5.18}$$

where DA' modulates learning in Eq. 3.26b.

Fig. 5.18 presents the results of this training. Panel C of the figure depicts the average responses of neurons whose preferred stimulus class is m_t. Those responses are for both the preferred stimulus class and the sum of all non-preferred classes. The bars of lighter colours are for a Hebbian (control) network for comparison. We note that, compared to control, the greedy approach decreases the neurons' responses to the non-preferred classes while leaving the response to the preferred class unchanged. This yields a decrease in miss-classifications (i.e., a decrease in false-positives for the class

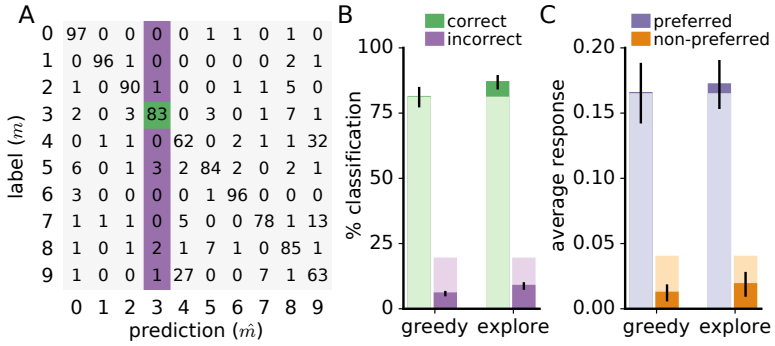

Figure 5.18: Effects of exploration on classification performance. (A) Confusion matrix for a greedy network. In this example, the target class $m_t = 3$; DA is then only released for stimuli classified as '3' (highlighted). (B) Correct and incorrect classification rates for the target class m_t. Lighter-coloured bars are for a Hebbian network (control). (C) Average responses of neurons whose preferred stimulus class is m_t. The responses are for the preferred (i.e., m_t) and non-preferred classes. Panels B and C of the figure are almost identical due to the strong correlations between a neuron selectivity and the rate of correct classification presented in Fig. 5.17 C. Data are the mean of 10 runs, each for a different target class m_t; error bars indicate the variance. The greedy approach minimises miss-classifications for the target class while the explorative approach additionally maximises explorative correct classifications.

m_t) but does not increase the rate of correct classification (correct positives for class m_t, Fig. 5.18 B). Of course, the decrease in false-positives for class m_t translates in an increase in correct positives for the other classes. When DA is released for all classes, this produces a rise in correct classifications for all classes. These results indicate that enhanced accuracy for the greedy network derives from a reduction in false-positives.

For the explorative network, DA-based learning results in a rise in neural response for the preferred class and to a drop for non-preferred classes. The greater response to the preferred class derives from stimuli of non-target classes being tentatively ("exploratively") classified as a target class (m_t), and this being correct. The learning mechanism maximises such surprisingly correct classifications, thereby enhancing the responses of neurons to their preferred class. These changes translate to improvements in accuracy which, in contrast with pure exploitation, both boosts true-positives and lessens false-positives.

5.5.2 Discrete versus Continuous Reward Predictions

Approximating predicted rewards through the classifier's posteriors offers poorer performance than binary reward predictions. This result comes as a surprise as the former approach brings additional information to the learning mechanism. Here, I attempt to explain this disparity in performance.

Fig. 5.19 shows empirically-measured DA values as a function of RPEs for the two approximation methods. Although the two activation functions closely correspond to one another for a significant range of RPEs, they also largely deviate in some cases, for instance for unexpected rewards ($\delta_{0/+}$, $RPE \gtrsim 0.5$). To determine which of these deviations are responsible for the difference in performance, I conduct an experiment

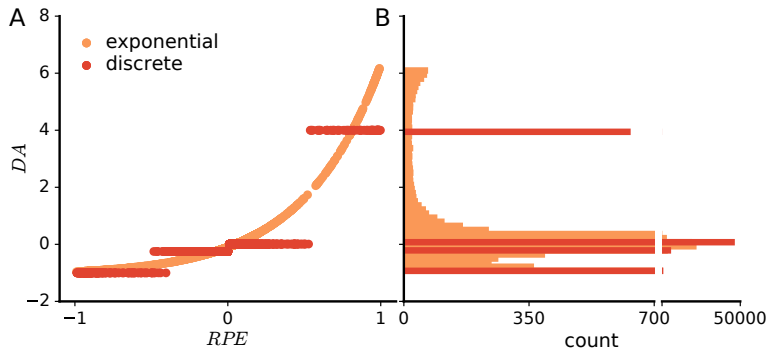

Figure 5.19: Comparison of DA activation functions. (A) Empirically measured DA values for the discrete and exponential $RPE \mapsto DA$ functions. (B) Histogram count for the DA activations for one iteration over the dataset (60 000 examples).

in which I linearise the exponential release function for each $\delta_{./.}$ value separately. For instance, for the linearisation for correctly predicted rewards (p_+, r_+), DA activation is given as:

$$DA = \begin{cases} \delta^{\eta}_{+/+} & \text{if } p_+ \text{ and } r_+ \\ \exp(\delta^{e}_{1} \cdot RPE) + \delta^{e}_{2} & \text{otherwise}\,. \end{cases} \tag{5.19}$$

The empirically-measured DA activations for each of theses scenarios is presented in Fig. 5.20 A. The performances of the models are presented in Fig. 5.20 B. Linearising the exponential function for correctly predicted rewards (+/+) recovers the performance of the fully discrete model (the two approaches are then not statistically significantly different, $p > 0.01$). In contrast, linearising any other part of the function leaves performance unchanged. These results indicate that the loss in performance for the exponential release function owes to the a poor mapping function for correctly predicted rewards.

In the discrete model, RPEs for correct reward predictions map to values close to zeros ($\delta_{+/+} = 0.01$). In the exponential curve, although most RPEs map to values around zeros, others lead to much greater DA activation, up to $DA = 2$. These greater values deteriorate learning. In more details, in the discrete model, having $\delta_{+/+}$ and $\delta_{0/0}$ close to zero for correct predictions ensures that incorrect predictions have maximal impact on the weights. In order words, error-based learning supplants correlation-based Hebbian learning. For the exponential activation function, large DA values for correct predictions restores correlation-based learning at the detriment of error-based learning.

This effect is visible in the distribution of the weight strengths of representation neurons (Fig. 5.21). Weight normalisation and feedforward inhibition ensures that the means of the distributions are the same for all models ($\mu = 1.28$). However, the spreads of these distributions differ. In particular, the variance in the explorative model is twice as large that of the Hebbian network ($\sigma^2 = 0.55$ and $\sigma^2 = 0.28$, respectively). On the other hand, for the exponential activation function, the variance is equal that of Hebbian learning. Linearising the exponential function for p_+, r_+ restores a distribution that is virtually identical to that of the discrete model ($\sigma^2 = 0.55$). These results suggest that high DA activation values for correctly predicted rewards (p_+, r_+) disrupt error-based learning by increasing the effects of statistical Hebbian learning on weight

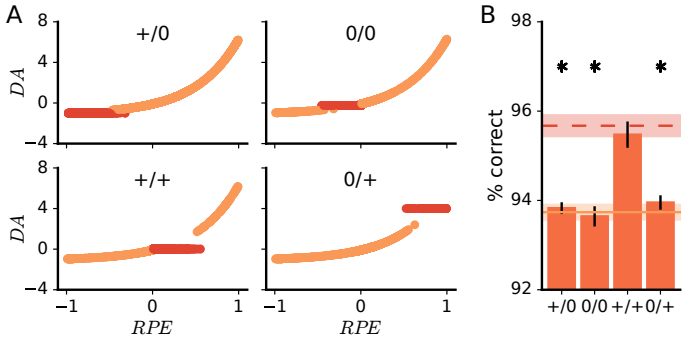

Figure 5.20: Linearisation of the exponential DA activation function. (A) Empirically-measured DA activations for the linearisation of the exponential function for each of four the discrete RPE scenarios. Red data points indicate the linearised portion of the function. (B) Performance of the models with different linearisation. Horizontal solid and dashed lines are the performances of the exponential and discrete models, respectively. Error bars and shaded areas are the variance. Stars indicate performances that are statistically significantly different from the discrete model ($p < 0.001$).

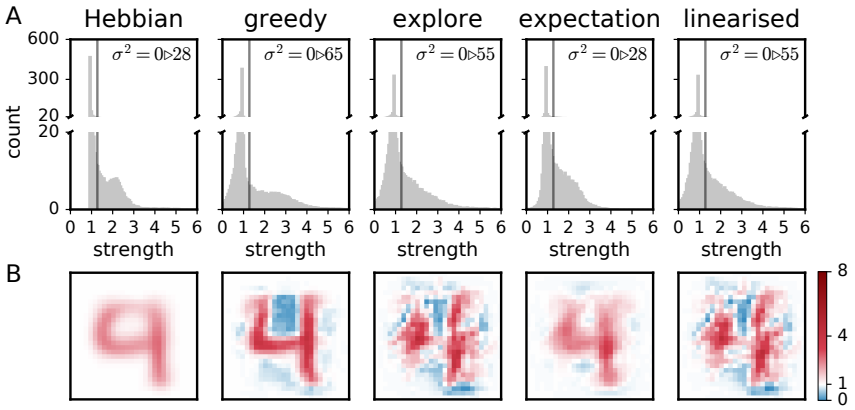

Figure 5.21: Weight distribution for the different models. (A) Histogram of weight strengths for networks trained through each of the different approaches. Data are for all the neurons of a single network. 'Explore' is the explorative model with discrete RPE scenarios, 'expectation' is the explorative model with an exponential $RPE \mapsto DA$ mapping function, and 'linearised' is the expectation model with a linearisation for correctly predicted rewards (+/+ in Fig. 5.20). The vertical line and σ^2 indicate the mean and variance of a distribution, respectively. (B) Depiction of the weights of the same representation neuron in different models. The variance of the weight distribution for the explorative model is twice that of Hebbian learning. On the other hand, the variance of the continuous reward prediction model is equal to that of Hebbian learning. The linearisation for correctly predicted rewards restores the distribution observed in the explorative network.

distributions.

I perform a last experiment to test whether reward expectations may be beneficial to refine sensory representations. Here, I make use of the discrete $RPE \mapsto DA$ function for all RPE scenarios except for the surprising reward case (p_0, r_+) for which DA

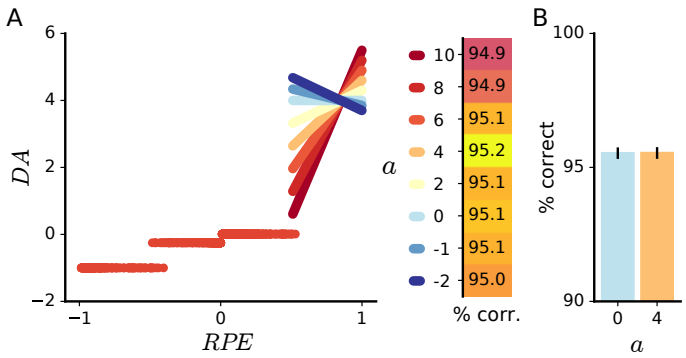

Figure 5.22: Discrete DA activation function with a linear component for surprising rewards. (A) Activation function for different slope parameters a in Eq. 5.20. The results of the parameter search for the optimal value of a are illustrated on the right. (B) Performance comparison for a control network ($a = 0$) and a network trained with the optimal value of $a = 4$. This comparison was made with a large sample size (20 runs) and after more training episodes (150). The small difference observed during parameter search (0.1%) disappears in this case.

activation is given by a linear relationship to the RPE:

$$DA = \begin{cases} \delta^{\eta}_{+/+} & \text{if } p_+ \text{ and } r_+ \\ \delta^{\eta}_{+/0} & \text{if } p_+ \text{ and } r_0 \\ a \cdot (RPE - \mathrm{E}[RPE_{0/+}]) + \delta^{\eta}_{0/+} & \text{if } p_0 \text{ and } r_+ \\ \delta^{\eta}_{0/0} & \text{if } p_0 \text{ and } r_0\,, \end{cases} \tag{5.20}$$

where a is a hyper-parameter determining the slope of the linear relationship and $\mathrm{E}[RPE_{0/+}]$ is the expected value of the RPE for the surprising reward scenario. The RPE here is computed using the posterior probability of the class chosen as output. This linear release function ensures that the mean DA activation for surprising rewards is the same for any value of a; only the distribution of the activation values changes. For $a = 0$, this relationship is identical to the discrete activation function. For $a > 0$, more daring explorative decisions elicit greater weight changes if correct. I perform parameter search for different values of a (Fig. 5.22 A). This search indicates that a positive relationship between RPE and DA activation, $a = 4$, gives slightly better performances than the control case, $a = 0$ (+0.1%). However, when tested with a larger sample size (20 runs) and with more training episodes (150 episodes), this difference disappears (Fig. 5.22 B). These results indicate that the additional information brought by the probability of a predicted reward is not useful to refine synaptic weights in the present model of DA-based learning.

5.5.3 Comparison with Acetylcholine

On the uniform MNIST dataset, DA yields a more consequential rise in performance than ACh (Fig. 5.23). Additionally, when training a network with both neuromodulators, first with ACh and then with DA, the performance attained is not statistically significantly different than from DA alone. On the other hand, on the non-uniform MNIST dataset (see Sec. 4.4.2) the performance gains deriving from ACh are homologous in size to that of DA and combining the two signals yields further improvements

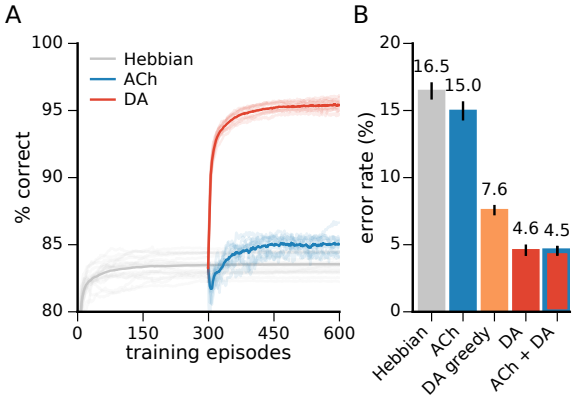

Figure 5.23: Performance comparison of the neuromodulator models. (A) Performance progression on the test dataset. Darker traces are averages over 20 runs, lighter traces are individual runs. (B) Final test performances for various models. All approaches lead to significant improvements over Hebbian learning alone ($p < 0.001$). DA alone and DA with ACh yield the best performance. Data are the mean of 20 runs, error bars are the variance.

compared to either separately (Fig. 5.24). This result indicates that the neuromodulators carry complimentary effects that can successfully combine to refine neural representations.

Fig. 5.25 depicts the impact of diminishing label availability on the accuracy of the models. While label scarcity affects both neuromodulators, the consequences are more substantial for DA. In error-based learning, labels are necessary to determine the correctness of an output. Reducing the ratio of labelled data consequently significantly hinders DA learning. In particular, when less than 1% of labels is used, the benefits of DA drop below those of ACh, this for both versions of the MNIST dataset. Thus, ACh might be particularly favourable when labels are scarce. It should be noted, however, that error-based learning does not require labels per se but only indications of whether outputs are right or wrong. DA might still perform well in a scenario where true labels are in short supply but reinforcement feedback is available.

5.5.4 Comparison with Error-backpropagation

It is informative to compare the learning outcome of DA with that of the error-backpropagation algorithm. Here, I train multi-layer perceptrons (MLPs) with two state-of-the-art gradient descent methods, the L-BFGS [Zhu+97] and Adam [KB14] algorithms. All the networks compared have identical architectures, namely 784 input, 300 hidden, and 10 output neurons. The DA model achieves an error rate comparable to those of the MLPs ($2.87 \pm 0.13\%$, $2.14 \pm 0.14\%$, and 1.88 ± 0.07 for DA, L-BFGS, and Adam, respectively, see Fig. 5.26). Additionally, the DA model outperforms the MLP reported in the original publication of the MNIST dataset by LeCun and colleagues [LeC+98], also for an identical network architecture.

To get further insights into the causes of the poorer performances of DA learning relative to modern optimisation methods, I compare the weights learned by the Hebbian network and the MLPs. Error back-propagation produces weights that do not resemble examples of the training data (see Fig. 5.27). This observation is also apparent in the low class selectivity of MLP neurons (Fig. 5.28). Here, neurons exhibit broad tunings,

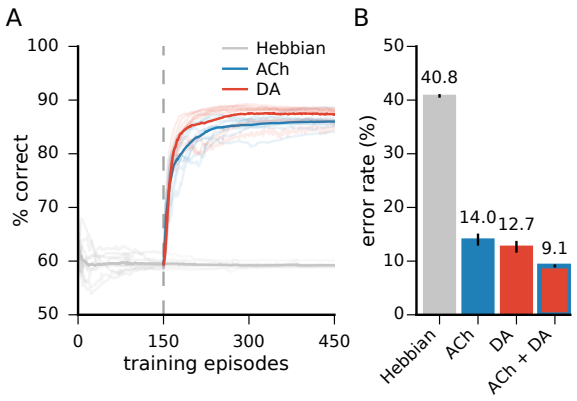

Figure 5.24: Performance of the neuromodulator models on the non-uniform MNIST dataset. (A) Progression of networks accuracy on the test dataset. Darker traces are averages over 10 runs, lighter traces are individual runs. (B) Performance comparison of the neuromodulators. The combination of DA and ACh yield the lowest error rate. Data are the mean of 10 runs, error bars indicate the variance.

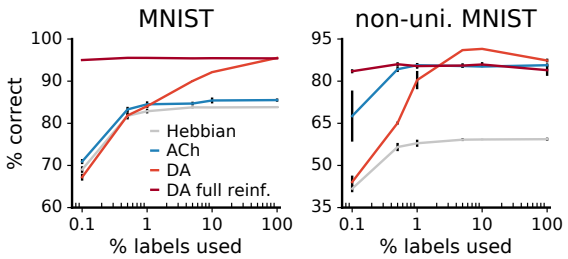

Figure 5.25: Comparison of label reliance for ACh and DA. While labels scarcity affects both neuromodulators, the impact is stronger for DA. Data is the mean of 3 runs, error bars indicate the variance.

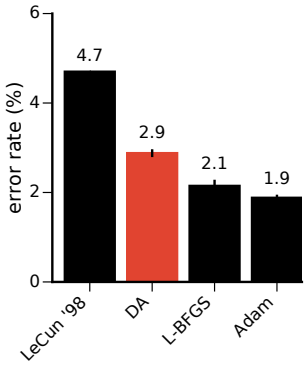

Figure 5.26: Performance comparison with MLPs. All results are for networks with an identical architecture, i.e., a single hidden layer containing 300 units. 'LeCun '98' are the original results from LeCun and colleagues [LeC+98] on the MNIST dataset. 'L-BFGS' and 'Adam' are state-of-the-art optimisation methods for MLPs. DA outperforms the MLP reported in the original publication of the MNIST dataset and achieves error rates comparable to those of state-of-the-art optimisation methods.

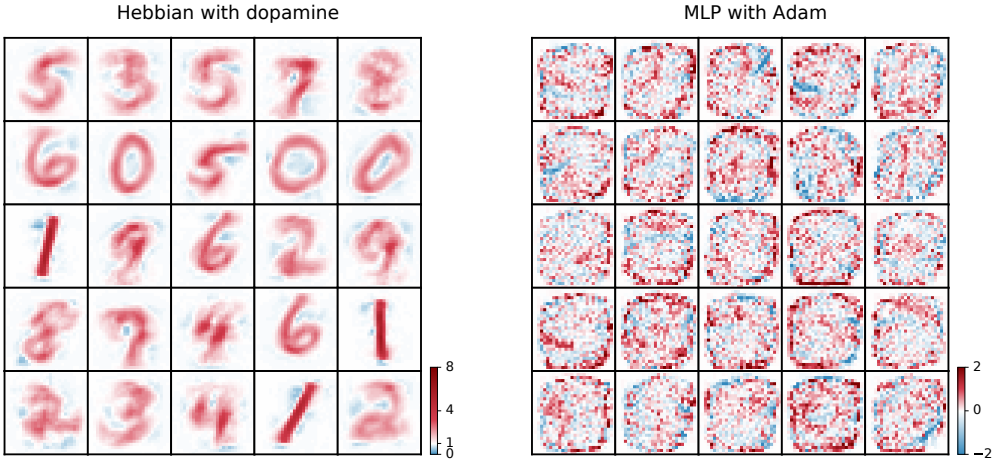

Figure 5.27: Comparison of representations in a Hebbian net and an MLP. Weights of two networks with 25 hidden neurons, one trained with Hebbian learning with DA activation, the other with the Adam algorithm. Weights in the Hebbian network clearly resemble the input data whereas those in the MLP are more difficult to interpret.

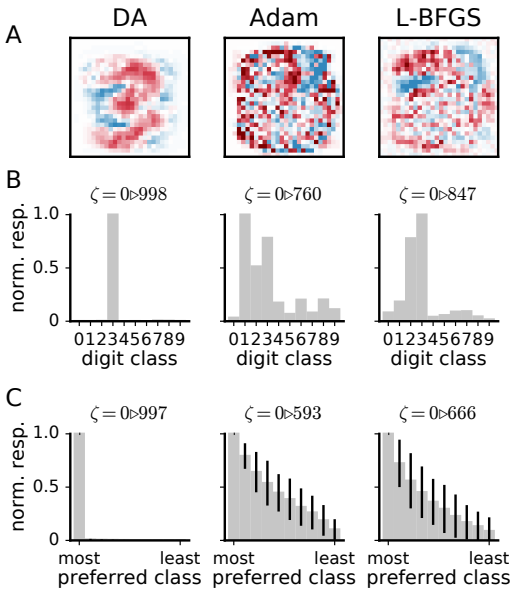

Figure 5.28: MLP neurons display low class selectivity. (A) Example weights for three networks. (B) Average responses of neurons in (A) to stimuli of the 10 digit classes. Data are the mean of 50 000 stimuli, values are normalised to the maximal mean response. (C) Average responses of all neurons in a network, ordered from most to least preferred classes. Data are the mean of all neurons of a single network, error bars are the variance.

responding strongly to multiple classes. Additionally, activations in the MLPs are less sparse than in the Hebbian network (Fig. 5.29). Together, the results indicate that MLPs make use of a distributed code which may explain their greater performance. Augmenting code distribution in the biology-inspired network, however, deteriorates performance (Fig. 5.29 C).

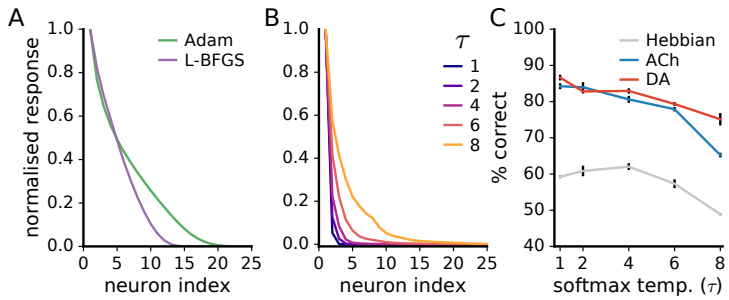

Figure 5.29: Code sparsity in MLPs and the Hebbian model. (A) Mean normalised responses of the 25 hidden neurons of MLP networks, ordered from most to least active. (B) Same responses in the Hebbian network for different inhibition strength (quantified as the temperature parameter τ of the softmax function in Eq. 3.14). For all results presented in this thesis, $\tau = 1$. (C) Performance of the biology-inspired models for increasingly distributed neural codes (i.e., greater τ). Data are mean of 5 runs, error bars indicate the variance. MLPs exhibit a more distributed code which may explain their higher performances. However, augmenting the breadth of the code in the Hebbian network degrades performance.

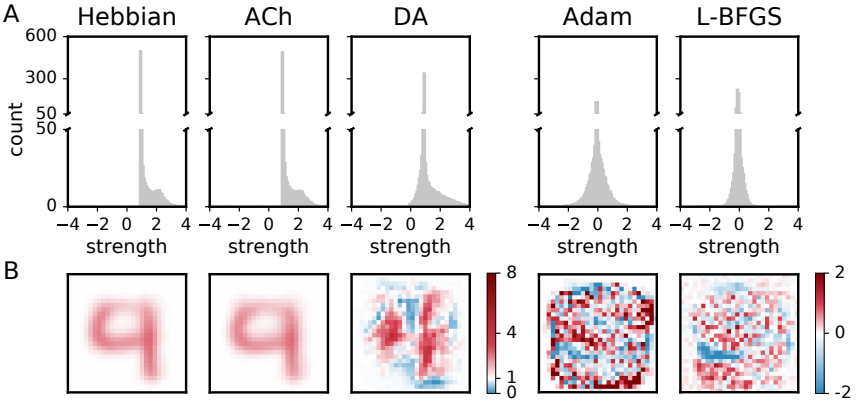

Figure 5.30: Hebbian nets and MLPs have dissimilar weight strength distributions. (A) Histogram of weight strengths for the Hebbian models and the MLPs. Data are the average of histograms for all neurons of a single network. (B) Example weights for each network. Note the different colour scales for the Hebbian and MLP nets. While the Hebbian nets contain exclusively non-negative weights, MLPs have about half their weights below zero.

Another distinguishing characteristic of MLPs is the distribution of their weight strengths: while connections in the Hebbian model are never negative, about half those in MLPs are (see Fig. 5.30). Connection strengths inferior to zero enable an MLP to negatively weight features that invalidate a classification output, conferring a valuable computational advantage.

5.6 Discussion

5.6.1 Relationship to Biology

DAergic neurons of the midbrain encode various features of rewards [TFS05; Sat+03] and, in particular, strongly respond to the difference between predicted and received rewards, or RPEs [SDM97; Sch10]. Midbrain neurons neurons project to the entire cortex [HK10] and the reward signals they carry modulate neural activity in most cortical areas [VCL11] including primary sensory cortices [Ars+13; Ple+09; BSS11]. DA affects plasticity at the sites where it is released, as measured both at the level of synapses [BHL03; BCO02; MMO06; Ota+98; SZW05; Cal+07; Cen+99; Li+03; NSF07] and behaviourally [Bre+02; Wis04; Gra05; Kle+07; OJN14; Hos+11; LS09; ML+09; KS06; Sch+12]. In sensory cortices, DA efflux, triggered either by electric stimulation of the midbrain [BCM01; Bao+03] or by reward delivery [FSV10; Poo+15; Bei+03] elicits plastic changes in the responses of primary sensory neurons.

The role of the plastic modifications induced by DA are usually understood in terms of reinforcement learning, for instance to learn the appetitive value of stimuli [FSV10; Wis04; Bre+02] or to acquire reward-directed behaviours [Sch+12; OJN14; Wis04]. In sensory representations, the changes elicited by DA were previously hypothesised to enhance the saliency of stimuli predictive of rewards [BCM01] and to adapt cortical representations to task requirements [BSS11]. Here, I show that a signal modulating plasticity as a function of RPEs adapts synaptic weights to the reward contingencies of a task. Specifically, in the model, the responses of neurons become matched to the boundaries in conditions for reward delivery. In the digit classification task, this corresponds to neurons being more selective to the distinct classes, in this way improving response accuracy. I suggest that, in mammals, DA carries this role of adapting sensory representations to the reward contingencies of a task.

After training monkeys on a visual discrimination task, neural responses become matched to the stimulus features that discriminate between the reward conditions of the task [SL02]. This process is homologous to the effect of DA in the model. The results presented above suggest that lesioning the dopaminergic system would prevent this form of learning. Animal experiments show that interfering with DA signalling impairs sensory discrimination learning [KS06; Sch+12], supporting this prediction.

Noisy neural activations in the representation layer gives rise to a softmax exploration rule. Coupled with DA release, such exploration provides further improvements compared to a greedy approach. Additionally, in this case, the optimal values of the $\delta_{./.}$ constants closely correspond to the release profile of DA in animals [SDM97; TFS05] (see Fig. 5.9). DAergic activation properties in animals may thus have been selected through evolutionary pressures to refine sensory representations in the presence of variable neural responses.

In animals, training on a discrimination task leads to a progressive population-wide increase in neural selectivity to task-relevant stimuli. This effect is explained by both the recruitment of more neurons to encode those stimuli and by a refinement of single cell responses [Poo+15]. Thus, there appears to be two phenomena at play concurrently: one inducing large-scale, population-wide modifications and the other refining tuning properties at the single neuron-level. The model presented here suggests that these alterations can derive from two separate signals both acting similarly on synaptic plasticity. A first signal following task difficulty or stimulus relevance results in

large-scale modifications whereby neurons may change altogether the stimulus class they respond to. A second signal accompanying RPEs induces smaller-scale alterations such that weights better match the classes of a task without alterations to their preferred classes.

On the full MNIST dataset, the combination of both DA and ACh does not further improves classification accuracy compared to DA alone (Fig. 5.23). This observation is in line with some animal experiments. Following training on a discrimination task, cortical representations in animals reorganise such that task-relevant stimuli become over-represented [Bao+04; FM11; SP14; GLP09; SNW12; Wan+03; WB98; RSM93; Poo+15]. Furthermore, a short time after training, the degree of population-level reorganisation correlates with improvements in performance [RSM93; PSM06; WFS04; Wei03; Din+03]. However, months after the initial training, this correlation disappears as performance on the task remains high but cortical maps become indistinguishable from control [BIP04; PSM06]. That is, recruiting more neurons to encode task-relevant stimuli does not provide further improvements over small-scale synaptic modifications. Similarly to this observation, the alterations orchestrated by ACh on the uniform MNIST dataset do not yield benefits over those brought by DA. On the other hand, the model predicts that changes in the extent of cortical representations are beneficial in circumstances where some stimuli are grossly under-represented in a training set. That is, infrequent but highly relevant stimuli should bear a greater cortical representation than would be expected from their frequency alone.

On the non-uniform data, ACh and DA both improve the learning outcome by a similar extent. Additionally, the combination of the two signals performs better than each separately. These results argue for complimentary effects of ACh and DA in representation learning, with each modulatory agent serving a distinct purpose.

The DA learning mechanism yields error rates that measure up to those of state-of-the-art optimisation methods in MLPs. Since evolutionary pressures must have favoured well performing learning mechanisms in the brain, any candidate model of cortical learning must offer strong functional performances. The DA mechanism meets this criteria, making it a suitable model for learning in biological neural structures.

5.6.2 Relationship to Machine learning

Comparing the representations in the Hebbian and MLP networks revealed that the latter make use of a more distributed code. In the Hebbian model, increasing the breadth of the code degrades performance. Future work should address the use of distributed neural codes as it may improve the learning outcome. Furthermore, allowing negative weight values may bring additional computational capacity to the Hebbian network.

It is interesting to note that, although its accuracy is less than that of modern learning methods, the DA mechanism outperforms MLPs of the same architecture presented in the original publication of the MNIST dataset [LeC+98]. Since this publication in the last 1990s, advances in gradient-based learning methods in MLPs resulted in error rates dropping to less than half the value originally reported. The biologically-inspired method presented in this work is at a relatively immature stage, perhaps comparable to the early days of error back-propagation, and we may expect similar improvements to derive from future research.

From a practical perspective, the algorithm presented in this work offers several interesting advantages over traditional gradient backpropagation-based learning. First, it requires a weaker supervision signal, making use only of binary rewards instead of explicit labels. Second, it can learn concurrently with and without environmental feedback. Finally, weight modifications are based on synaptically-local information and a signal broadcasted identically to all neurons. These features make the learning mechanism suitable for implementation in physical devices such as neuromorphic chips. Additionally, neuromodulation is likely to further improve performance of hierarchical networks with Hebbian learning [FSL16], which have a functional focus on learning from data with very few labels.

5.7 Summary

DAergic firing in the brain encodes RPEs and elicits plasticity in sensory representations. In this chapter, I demonstrate that a signal modulating synaptic plasticity as a function of RPEs enhances neural selectivity and boosts classification performance. In the presence of exploration, the activation profile of DA matches that observed in animals. The network then achieves error rates in the range of those of state-of-the-art optimisation methods for MLPs. To extend the functional capacity of the model, future work should investigate the use of more distributed neural codes as well as of negative weight values.

Chapter 6

Deep Convolutional Model

6.1 Overview

The mammalian perceptual system is made up of hierarchies of cortical areas that encode progressively more complex sensory features. This organisation is apparent in the pattern of inter-areal wiring [VEM83; ME83; FVE91; KH00] as well as in the tuning of neurons [Pat+02; GSW14], most evidently so in the primate ventral visual pathway. Here, neurons in the primary area prefer simple stimuli such as oriented edges, sinusoidal gratings, and blobs [HW62; JP87; Rin02] while cells in V2 respond most strongly to more complex features like arcs, angles, intersecting lines, and polar gratings [HVE00]. V4 neurons require yet more complicated shapes to reach maximal activation [KT94] and, further downstream, cells exhibit a preference for whole percepts like faces, animals, and objects [KKF00; SL02; Qui+05].

In artificial neural networks, research indicated early on that networks containing at least one hidden layer can solve non-linearly separable problems that shallow networks cannot [Ros58; MP69; RHW86]. More recent developments show that 'deep' networks with multiple hidden layers can learn intricate input-output relationships, achieving record-breaking results on a plethora of tasks [Dah+12; KSH12; Hin+12; Ma+15; Cio+12; Hel+13; Mni+15]. Such a layered structure captures the intrinsic compositional hierarchy of natural signals where low-level motifs (e.g., object parts or speech phoneme) assemble into higher-level entities (whole objects or words). These results argue for a fundamental role of hierarchical processing in interpreting natural signals.

Convolutional weight filters recently proved another valuable architectural characteristics in ANNs. In data spatially organised in multiple arrays, such as 2-dimensional images, local groups of data highly correlate and often form motifs invariant to location. ConvNets leverage this property by convolving spatially localised weight filters over data. Neurons at different locations share the same filter weights, thereby greatly reducing the number of trainable parameters. After the convolutional layer, a so-called pooling layer merges semantically similar features which creates an invariance to small shifts and distortions. Interestingly, ConvNets are directly inspired from the mammalian visual system where cells encode comparable features at different retinal locations [HW62; FM82]. ConvNets rival the primate visual area IT in terms of performance on object recognition tasks and predicts the response of neurons in this area [Cad+14]. Such networks thus make both excellent functional algorithms and

compelling models of visual sensory processing.

Additional processing layers in neural networks bring further challenges for learning algorithms. In particular, learning in deeper networks broadens the scope of the credit-assignment problem. In this case, a learning algorithm must identify the units along a processing hierarchy responsible for a network's incorrect behaviour in order to modify their weights.

In this chapter, I implement the DA learning mechanism in a two-layered convolutional network to examine its rendition in deeper neural structures. The rationale for using a convolutional architecture is the following. First, the convolution operation is line with the organisation of the mammalian visual system. Second, ConvNets currently claim the best performance on a wide range of problems, including the MNIST dataset used in this thesis [CMS12]. Third, practicalities of the network's functioning favour the use parallel layers, such as the feature maps in ConvNets. Specifically, in the Hebbian network, softmax inhibition results in mostly one neuron being active at stimulus presentation. Stacking layers with a single active neuron is futile as downstream layers have no correlations to learn from. Using a convolution in the first layer circumvents this problem as the softmax competition here takes place locally for each filter. Therefore, although a single unit fires in each local competition, many units respond in the whole layer, giving rise to correlation patterns.

The following text is divided in four sections: first, I describe the network's structure, then I present results for Hebbian learning alone, next I describe results for DA in a greedy network, and finally I investigate the use of DA in an explorative model.

6.2 Convolutional Network Model

The inner working of the ConvNet is similar to the one described in Ch. 3; the main difference here lies in the network's architecture. As in the shallow network, the ConvNet is made up of an input layer, a representation (here named fully connected feedforward) layer, and an output layer. In addition, the ConvNet contains a convolutional and a pooling layer between the input and representation layers (Fig. 6.1).

6.2.1 Model Description

Input Layer

As in the shallow network, the input layer normalises the activation of the input neurons, emulating the process of feedforward inhibition in biology:

$$y_d = (A - D)\frac{\tilde{y}_d}{\sum_{d'} \tilde{y}_{d'}} + 1\,, \tag{6.1}$$

where A is a normalisation constant, D is the number of input neurons and $\tilde{y}$ is the activation of individual input neurons before normalisation.

Convolutional Layer

The convolutional layer is made of F feature maps. All neurons of a map f share the same weight matrix $\mathbf{Q}^f$. This weight matrix corresponds to a kernel of size $G \times H$

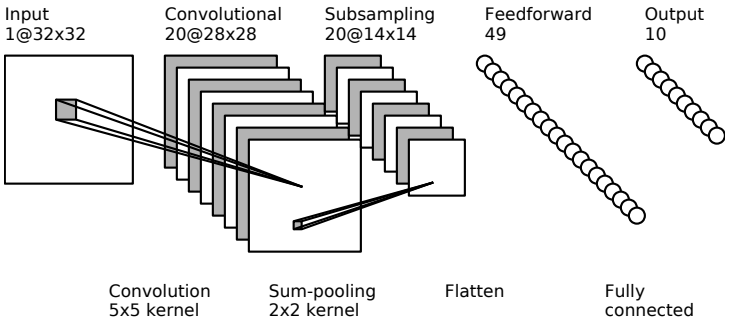

Figure 6.1: Architecture of the convolutional network. Five layers compose the ConvNet. The input, representation (here named fully connected feedforward), and output layers are identical to those of the shallow network. In addition, the network contains a convolutional and a subsampling layer between the input and feedforward layers. All neurons within a map of the convolutional layer share the same weight filter. This weight filter is a spatially limited, 5×5 kernel convolved over the input layer. There are no trainable parameters between the convolutional and subsampling layers, only a sum-pooling operation (see text). The subsampling layer is fully connected to the feedforward layer.

convolved over the input layer. In order to consider all pixels of the input images equally, zero-padding is added on all sides of the images (added before the normalisation above). The padding is of size $(G-1)/2$ and $(H-1)/2$ on the vertical and horizontal axes, respectively. Except if stated otherwise, for all results presented in this chapter I use $F = 20$ feature maps and a square convolutional kernel of size $G = H = 5$. These dimensions result in input images of 32×32 pixels and in $20 \times 5 \times 5 = 500$ weights in the convolutional layer. The padding and convolution yield feature maps containing the same number of neurons as the original images, or $28 \times 28 = 784$.

The activations of neurons in the convolutional layer are given by:

$$\tilde{u}^f_{i,j} = \sum_{g=i}^{i+G-1} \sum_{h=j}^{j+H-1} S(Q^f_{g,h}) \cdot y_{i,j} \,, \tag{6.2}$$

where f is the index of the feature map, g and h are the indices of weights along the vertical and horizontal axes of the convolutional kernel, respectively, i and j are the indices of the neurons in the input and convolutional layers along the vertical and horizontal axes, respectively, and $S(\cdot)$ is a linearised logarithmic function (see Eq. 3.12).

Lateral Inhibition in the Convolutional Layer

In the representation layer of the shallow network, all neurons compete for activation against one another as a result of the softmax inhibition (Eq. 3.13). This process is the same in the feedforward layer of the ConvNet. In the convolutional layer, however, the softmax is taken across all F convolutional maps but only over neurons (i, j). In other words, neurons compete with those at the same location in the different maps but not with other neurons in the same map. We can write:

$$u^f_{i,j} = \frac{\exp(\tilde{u}^f_{i,j})}{\sum_{f'} \exp(\tilde{u}^{f'}_{i,j})} \,. \tag{6.3}$$

Subsampling Layer

The subsampling layer performs a pooling operation that reduces the number of trainable parameters and introduces partial translation invariance in the responses of neurons. The pooling operation is a summation ("sum-pooling") over a kernel of size $E \times E$ convolved over each feature map individually with a stride E. (Max-pooling, i.e., taking the maximum value over a kernel, is more common in the machine learning literature than sum-pooling but the latter gives better performances in the present network.) Formally, the subsampling operation is defined as:

$$\tilde{v}^f_{a,b} = \sum_{i=a\cdot E}^{(a+1)\cdot E-1} \sum_{j=b\cdot E}^{(b+1)\cdot E-1} u^f_{i,j}\,, \tag{6.4}$$

where a and b are indices of the vertical and horizontal axes of the subsampling maps, respectively. The pooling is followed by a softmax operation identical to the one performed in the convolutional layer:

$$v^f_{a,b} = \frac{\exp(\tilde{v}^f_{a,b})}{\sum_{f'} \exp(\tilde{v}^{f'}_{a,b})}\,. \tag{6.5}$$

For this work, I use a subsampling kernel of size $E = 2$. The subsampling layer is fully connected to the representation layer. The subsampling operation leads to a 4 fold reduction in the number of trainable parameters between the convolutional and feedforward layers, from $20 \times 28 \times 28 \times 49 = 768\,320$ to $20 \times 14 \times 14 \times 49 = 192\,080$.

Fully Connected Feedforward Layer

Except is stated otherwise, the feedforward layer contains $C = 7 \times 7 = 49$ units, the same number as in the representation layer of the shallow network. The activation of the feedforward layer is computed like for the shallow network, only taking as input the activation of the subsampling layer:

$$\tilde{s}_c = \sum_{a=0}^{A} \sum_{b=0}^{B} S(W^f_{c(a,b)}) v^f_{a,b}\,. \tag{6.6}$$

Softmax inhibition follows:

$$s_c = \frac{\exp(\tilde{s}_c)}{\sum_{c'} \exp(\tilde{s}_{c'})}\,. \tag{6.7}$$

Learning

As for the shallow network, weights are trained through Hebbian learning augmented with a weighted normalisation term:

$$\Delta Q^f_{g,h} = \epsilon^{\text{conv}}(u^f_{i,j} y_{i+g,j+h} - u^f_{i,j} Q^f_{g,h})\,, \tag{6.8}$$

$$\Delta W^f_{c(a,b)} = \epsilon^{\text{feed}}(s_c u^f_{a,b} - s_c W^f_{c(a,b)})\,, \tag{6.9}$$

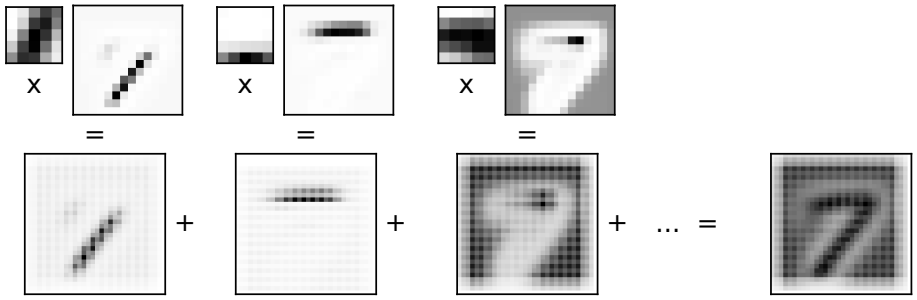

Figure 6.2: Reconstruction of the 'receptive field' of a feedforward unit. On the top row, the smaller weight matrices are the weights $Q^f_{g,h}$ of three convolutional filters f. The larger weight matrices are the weights $W^f_{c(a,b)}$ of the corresponding filter f for one feedforward unit c. On the bottom row are the partially reconstructed 'receptive fields', i.e., the convolutional filters tiled and multiplied by the corresponding feedforward weight. These partial reconstructions are for each filter (left) and summed for all filters (right). This figure shows a selection of 3 out of 20 filters; the final (summed) receptive field is for all 20 filters.

where ϵ^{conv} and ϵ^{feed} are learning rates for the convolutional and feedforward layers, respectively.

In total, the ConvNet contains $500 + 192\,080 + 490 = 193\,070$ trainable weights. This large number of parameters (5 times greater than in the shallow network) and the the convolution and subsampling operations significantly slow down the simulation. To compensate, I increase the learning rate of the network, allowing the performance to plateau within 10 training episodes, with only minimal impacts on the results (the lower learning rate in the shallow network was necessary for the model of ACh). On a personal computer, these 10 training episodes take about 2 hours to run on the full MNIST dataset. (I implemented the most computationally demanding functions using the Numba package which compiles the instructions into machine code, yielding performances similar to C++. The computational efficiency of the code could certainly be further improved but this is beyond the scope of this work.)

6.2.2 Receptive Field Reconstruction

In contrast with the shallow network, the representation weights in the ConvNet are not easily interpretable in themselves. To address this problem, I reconstruct the 'receptive fields' of feedforward neurons in the following manner. For each filter, I tile the filter over a white canvas and multiply it at each location by the corresponding feedforward weight. I add the filters where their tiling overlaps. The tiling is done with a stride of two to compensate for the subsampling. This operation results in 20 canvases, one for each feature map. I then scale the canvases by their maximal values and sum them to form the final 'receptive field' (Fig. 6.2). This reconstruction is only for visualisation purposes and does not serve for further quantitative analyses.

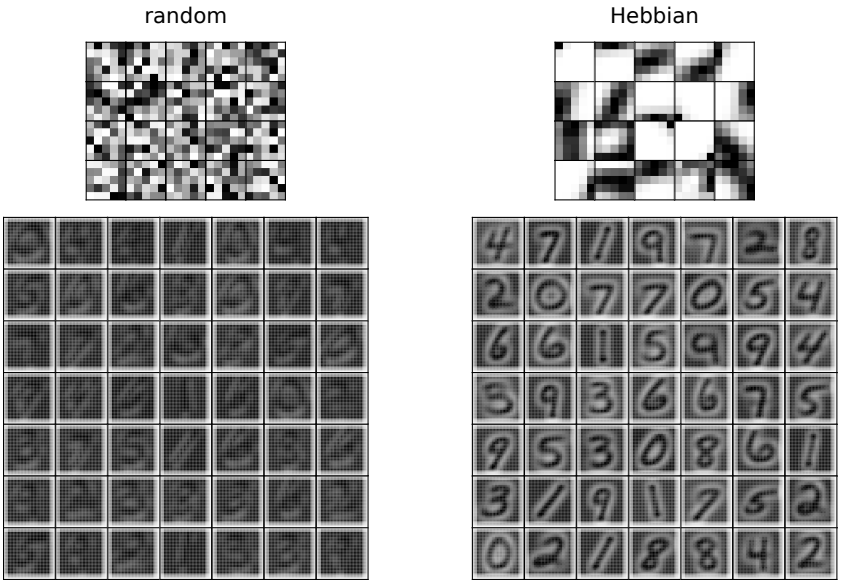

Figure 6.3: Weights in the convolutional network. Top: weights of the convolutional filters, either random or acquired through Hebbian learning. Bottom: reconstruction of the receptive fields of the feedforward neurons.

6.3 Hebbian Learning

As a first test of the ConvNet, I train the model only through Hebbian learning, without DA. To assess the effects of the convolutional layer, I train a first network without learning in this layer: the weights are initialised randomly and left as is. In a second network, I train both the convolutional and feedforward layers.

6.3.1 Results

Weights

Fig. 6.3 depicts the weights learned in two ConvNets. In the network with learning in both layers, the convolutional filters represent oriented bars and blobs reminiscent of the receptive fields of primary visual neurons in mammals. The reconstructed receptive fields of the feedforward layer resemble the different digit classes, much like the weights learned in the shallow network.

In the ConvNet with the convolutional filters left random, the weights in the feedforward layer are less clearly identifiable. Nonetheless, we can still make out distinct integers. This observation suggests that even random convolutional filters may have some representational power.

Performance

Both approaches yield classification performances that are significantly superior to chance level (Fig. 6.4). Not surprisingly, the network with learned convolutional filters outperforms the one with random filters. This former network also reaches higher rates of correct classification than its shallow equivalent.

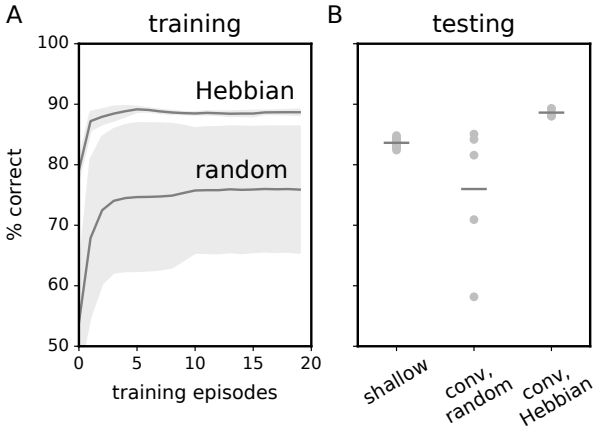

Figure 6.4: Performance of the Hebbian convolutional network. (A) Progression of the rate of correct classification on the training dataset for networks with random convolutional filters ('random') or with filters learned with Hebb's rule ('Hebbian'). The trace is the mean of 5 runs, shaded area is the variance. (B) Classification performance on separate test data for the ConvNet in (A) and for a shallow Hebbian network. The shallow network contains 49 representation units, the same as the number of feedforward units in the ConvNet. Data points are different network instances; horizontal bars indicate mean performances.

On average, the network with random filters produce poorer results than the shallow network. Surprisingly, however, some of the networks with random filters achieve accuracies close to that of the ConvNet with learned filters, superior to those of the shallow network. This again indicates that even random convolutional filters can yield relevant representations.

6.4 Dopamine in Greedy Model

In this section, I extend the model of DA transmission introduced in Ch. 5 to a ConvNet following a greedy classification strategy.

6.4.1 Model Description

In animals, DAergic signals reach most of the cortical mantle [LS09; Hos+11; KS06; GSK92; Lid+91; HK10] where they modulate neural activity [Ars+13; Ple+09; VCL11] and contribute to synaptic plasticity [BCO02; MMO06; Ota+98; BCM01; Bao+03; FSV10; Hos+11; KS06]. These signals appear to broadcast diffusely, with little topographical organisation, as evidence by the intermixing of cortical-projecting DAergic neurons of the VTA and the co-labelling of these neurons following retrograde tracing in multiple cortical areas [GSK92; HK10]. These findings suggest that DA neurons of the VTA send a unique RPE signal uniformly throughout the cortex.

To remain in line with these observations, I compute a single RPE value for the whole ConvNet. The calculation of the RPE follows Eqs. 5.2-5.5 for discrete reward predictions presented in Ch. 5. In the present section, I consider a network behaving exclusively in a greedy manner. The network is thus said to predict a reward on every trial, which yields two discrete RPE scenarios: (p_+, r_+) and (p_+, r_0).

Although the RPE signal is the same for both layers, the specific values of DA activation differ for the convolutional and feedforward layers. In biology, this may be equivalent to having varying receptor concentrations in different cortical areas, as is indeed observed [Lid+91]. In the model, the value of the DA variable follows separate $\delta_{./.}$ constants for the two layers. Specifically, in the convolutional layer, DA activation is given by:

$$DA^{\text{conv}} = \begin{cases} \delta^{g,\,\text{conv}}_{+/+} & \text{if } p_+ \text{ and } r_+ \\ \delta^{g,\,\text{conv}}_{+/0} & \text{if } p_+ \text{ and } r_0\,, \end{cases} \tag{6.10}$$

while in the feedforward layer, DA activation is determined as:

$$DA^{\text{feed}} = \begin{cases} \delta^{g,\,\text{feed}}_{+/+} & \text{if } p_+ \text{ and } r_+ \\ \delta^{g,\,\text{feed}}_{+/0} & \text{if } p_+ \text{ and } r_0\,. \end{cases} \tag{6.11}$$

The values of the $\delta_{./.}$ variables are found through two-dimensional grid-search for each layer separately.

The DA signal affects learning in Eqs. 6.8 and 6.9 for the convolutional and feedforward layers, respectively:

$$\Delta Q^f_{g,h} = DA^{\text{conv}} \cdot \epsilon^{\text{conv}}(u^f_{i,j} y_{i+g,j+h} - u^f_{i,j} Q^f_{g,h})\,, \tag{6.12}$$

$$\Delta W^f_{c(a,b)} = DA^{\text{feed}} \cdot \epsilon^{\text{feed}}(s_c u^f_{a,b} - s_c W^f_{c(a,b)})\,, \tag{6.13}$$

In order to assess the influence of DA on the two layers separately, I consider networks trained with four approaches: one with Hebbian learning, one with DA only in the convolutional layer, one with DA only in the feedforward layer, and one with DA in both layers.

6.4.2 Results

Parameter Exploration

To minimise computational requirements, parameters are explored for each layer individually (one layer undergoes DA-based learning while the other learns through Hebbian learning, and vice-versa). For each parameter set, a single network instance is tested, also to reduce computational overhead. Fig. 6.5 presents the results of the parameter search. For both layers, the optimal parameter set has $\delta_{+/+}$ close or at zero and $\delta_{+/0}$ larger and negative, in line with the values found in the shallow network.

Weight Changes

Fig. 6.6 depicts the weights of ConvNets trained solely with Hebb's rule or with DA in both layers. DA alters neural tuning in the convolutional and feedforward layers. To quantify the modifications induces by DA, it is informative to consider neural selectivity ζ (defined in Ch. 5, Eq. 5.7). The convolutional weights display much less stimulus selectivity than those in the feedforward layer (Fig. 6.7). This is explained by the small receptive field size of these neurons (5×5 pixels) which then encode little class-specific information. DA statistically significantly increases stimulus selectivity in the feedforward ($p < 0.0001$) but not in the convolutional layer ($p > 0.01$).

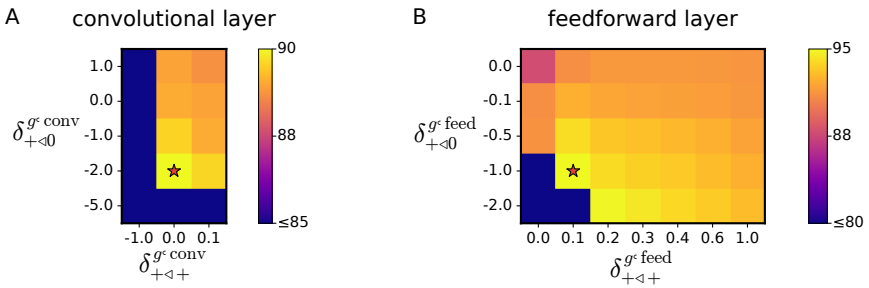

Figure 6.5: Exploration of DA parameters in the greedy ConvNet. (A) Parameters for the convolutional and (B) feedforward layers. The colour axis represents classification accuracy on test data, a star indicates the best performing parameter set. Data are for a single network.

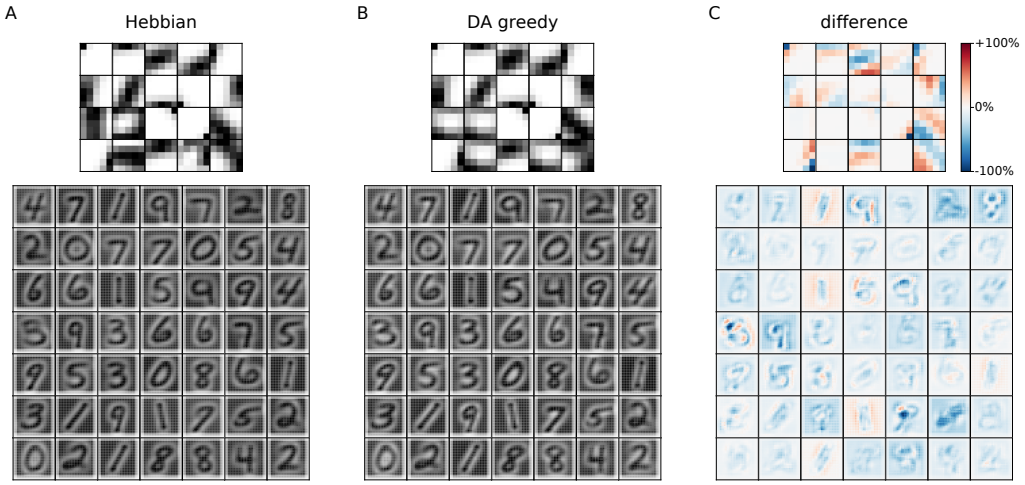

Figure 6.6: Weight changes. Top: weights of the convolutional filters; bottom: reconstruction of the receptive fields of the feedforward layer. (A) Weights following Hebbian and (B) DA-based learning. For DA learning, the weights of the convolutional and feedforward layers are from different networks in which only this layer underwent DA activation (otherwise, DA in the convolutional layer may indirectly affect affect downstream weights). (C) Difference in weights between Hebbian- and DA-based learning, relative to the maximal difference.

Performance Improvements

In the convolutional layer, DA bears minimal repercussions on the network's performance; rates of correct classification are not statistically greater than for Hebbian learning alone (Fig. 6.8, $p > 0.01$). On the other hand, DA in the feedforward layer greatly enhances performance, halving the error rate.

6.4.3 Additional Effects on Learning

Filter Size

It comes as a surprise that DA signalling in the convolutional layer yields no gain in classification accuracy. This may be a consequence of the small dimension of the convolutional filters which does not allow neurons to represent task-relevant information. To test this hypothesis, I examine the influence of filter size on the learning outcomes.

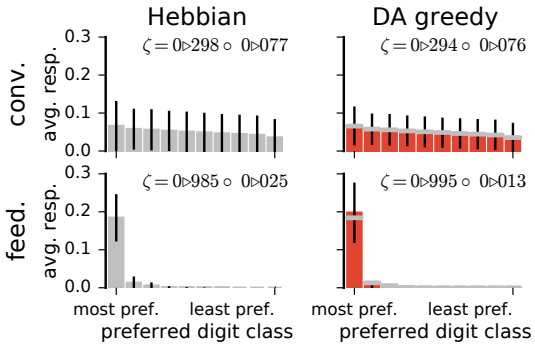

Figure 6.7: Neural selectivity in the ConvNet. Average responses are computed as the mean activation of a neuron to images of the different digit classes and sorted from most to least preferred classes. The data are the mean for all neurons of five networks; error bars are the variance. On each plot, ζ indicates the mean neural selectivity $\pm$ variance (see Eq. 5.7). DA statistically significantly increases selectivity in the feedforward ($p < 0.0001$) but not in the convolutional layer ($p > 0.01$).

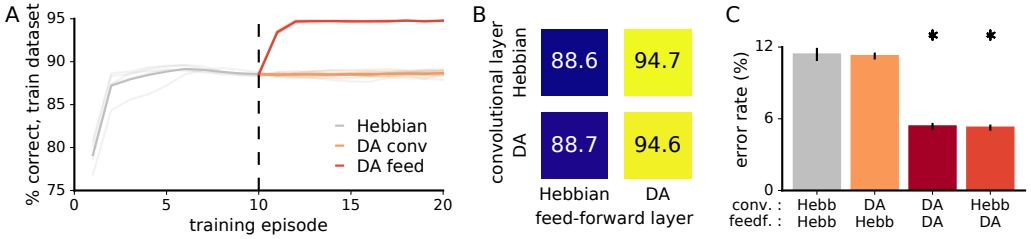

Figure 6.8: Performance in the greedy ConvNet. (A) Progression of the training performance. The dashed vertical line indicates the start of DA release. Heavy traces are the mean of 5 runs, light traces are individual runs. (B) Colour and numbers indicate the average rate of correct classification for a model on the test MNIST dataset. (C) Error rates for the models, sorted in decreasing order. Error bars are the variance. Asterisks indicate error rates that are statistically significantly lower than that of the Hebbian model ($p < 0.0001$).

In view of computational efficiency, I make use of a smaller network (16 feedforward neurons) trained on a subset of the data (containing only classes '4', '7' and '9'). The results of this manipulation are shown in Fig. 6.9. Three observations can be made from this figure. First, performance in the Hebbian model declines with larger filters, indicating that small filters encoding task-unspecific features produce better data representations. Second, the larger the filters are, the greater the impact of DA is in the convolutional layer. While DA does not statistically significantly enhance performance for small filters (e.g., 5×5), it has an important effect for larger filters (e.g., 27×27). This observation is also apparent in the class selectivity of filters of different dimensions: here, DA enhances stimulus selectivity for broad but not narrow filters (Fig. 6.10). Third, for intermediate filter sizes (13×13), simultaneous DA activation in both layers yields performances greater than DA applied separately in each layer. This finding indicates that a unique DA signal can refine data representations in multiple layers concurrently.

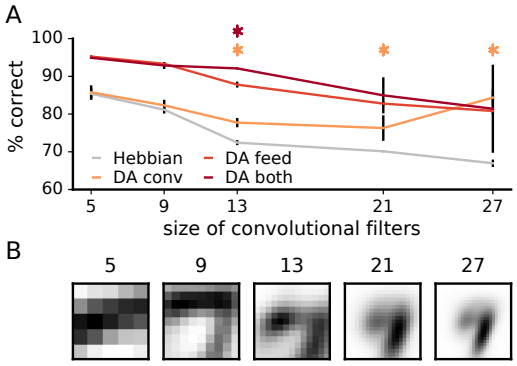

Figure 6.9: The benefits of DA grow with the size of the convolutional filters. (A) Classification performance on test data as a function of filter size. Data are the mean of 3 runs, error bars are variance. Red and orange asterisks indicate statistical significance for 'DA both' > 'DA feed' and 'DA conv' > 'Hebbian', respectively ($p < 0.01$). (B) Example convolutional weights in Hebbian networks for each filter size. Compared to small filters, larger filters encode more task-related information; DA has a greater impact on performance for these filters.

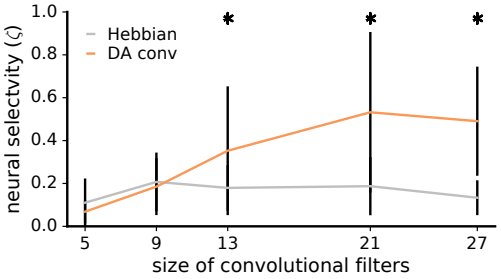

Figure 6.10: DA's impact on neural selectivity increases with filter extent. Class selectivity of convolutional weights as a function of filter size. DA significantly increases the selectivity of convolutional weights for larger but not smaller filter sizes ($p < 0.01$). Data is the mean of all neurons of 20 maps for 3 networks (i.e., 60 data points per filter size), error bars indicate variance.

Effects of Subsampling

The subsampling layer of convolutional networks provides tolerance to distortions and translations in the local features. To assess the extent of this role in the present network, I vary the ratio of the subsampling operation. This ratio is equivalent to $1/E^2$, where E is the size and stride of the subsampling kernel in Eq. 6.4. Fig. 6.11 presents the results of this experiment. The ratio 1:4 is the one used in the network models presented above. Greater subsampling slightly deteriorates the benefits of DA but does not affect Hebbian learning. The minor losses in performance probably derive from too coarse image representations. On the other hand, after removing subsampling altogether (ratio of 1:1), performance of the ConvNet drops below that of the shallow network, both for DA and Hebbian learning. This result suggests that the representational advantage of the ConvNet derives from the subsampling operation; without subsampling, the representation learned by the convolutional layer is worse than simply using raw input pixels.

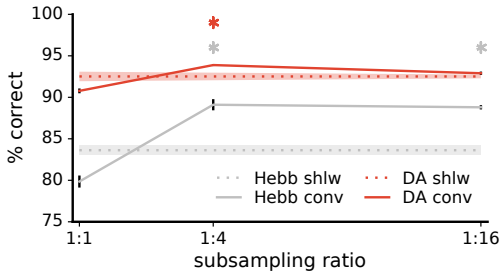

Figure 6.11: Subsampling improves the quality of a ConvNet's representation. Classification accuracy for different subsampling ratios. This ratio is given by $1/E^2$, where E is the size and stride of the subsampling kernel. This ratio also gives the reduction in the number of units from the convolutional to subsampling layers. Grey and red asterisks indicate statistically greater accuracy in the ConvNet than in the shallow network for Hebbian and DA learning, respectively ($p < 0.01$). Data is the mean of 3 runs, error bars are the variance. Subsampling enhances the network's representational power by providing tolerance to distortions and translations in the local features.

6.5 Dopamine in Explorative Model

Explorative decision making proved to further improve the outcome of DA-based learning in the shallow neural network. In this section, I test whether such exploration also benefits the ConvNet.

6.5.1 Model Description

Exploration follows the approach presented in Sec. 5.3.2 with discrete reward predictions. Specifically, normally distributed noise is added in the activation of the convolutional and feedforward layers (Eqs. 6.2 and 6.6):

$$\tilde{u}^f_{i,j} = \sum_{g=0}^{G-1} \sum_{h=0}^{H-1} S(Q^f_{g,h}) \cdot y_{i,j} + \eta^f_{i,j} \tag{6.14}$$

$$\tilde{s}_c = \sum_{c=0}^{C-1} S(W_{cd}) \cdot y_d + \eta_c \,, \tag{6.15}$$

where $\eta^f_{i,j} \sim \mathcal{N}(0, \gamma^{\text{conv}})$ and $\eta_c \sim \mathcal{N}(0, \gamma^{\text{feed}})$, with γ^{conv} and γ^{feed} being hyperparameters controlling the how explorative the network is. This approach emulates a softmax decision rule (see Sec. 5.3.1). For the results presented in this section, I use $\gamma^{\text{conv}} = 0.5$ and $\gamma^{\text{feed}} = 0.2$.

The are four RPE scenarios, the same as for the shallow network: (r_+, p_+), (r_0, p_+), (r_0, p_0), and (r_+, p_0). The outcome of a classification dictates reward delivery; whether a trial is explorative or not determines the network's reward prediction. A trial is said to be explorative if the output of the network with noise addition differs from the network's output without noise addition. Which of the four distinct RPE cases the current trial falls into determines the magnitude of DA transmission, with the two layers having a separate set of $\delta_{./.}$ variables, $\delta^{\eta,\,\text{conv}}_{./.}$ and $\delta^{\eta,\,\text{feed}}_{./.}$.

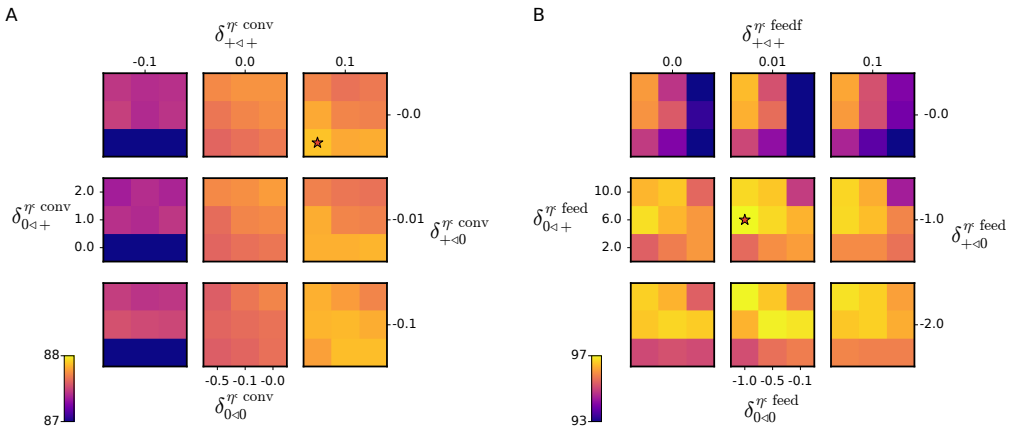

Figure 6.12: Parameters search in the explorative ConvNet. (A) Convolutional layer, (B) feedforward layer. Colour indicates test performance, a star indicates the best parameter sets. Data are for a single network instance.

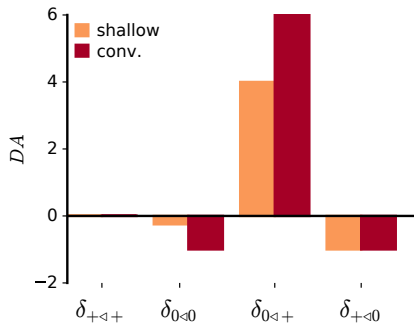

Figure 6.13: Comparison of optimal DA parameters for the shallow and convolutional networks. The similarity in theses values suggest that this activation profile is generally beneficial to refine representations during exploration- and error-based learning.

6.5.2 Results

Parameter Exploration

I explore the $\delta_{./.}$ parameter space for each layer independently (see Fig. 6.12). To limit computational costs, I train a single network for each parameter set. Fig. 6.13 shows a comparison of the optimal parameters for the feedforward layer and for the shallow network (the parameters for the convolutional layer may not be relevant as exploration in this layer does not significantly improves performance, see below). Although the optimal parameter sets differ in some respects (in particular for $\delta_{0/0}$), the overall agreement between the parameters suggest that this activation profile is generally beneficial to error-based representation learning.

Performance Improvements

Fig. 6.14 shows a comparison of the performances of some of the networks presented in this chapter. These networks have convolutional filters of 5×5 pixels. For these networks, DA release in the convolutional layer does not significantly improve network

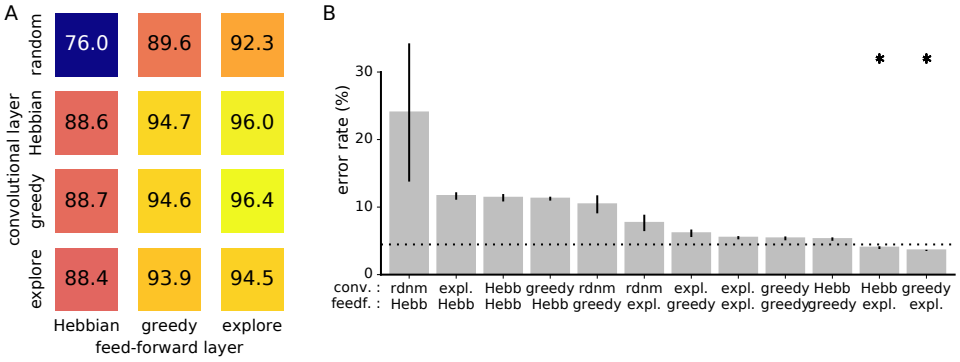

Figure 6.14: Performance comparison for different ConvNet models. (A) Colour and numbers indicate the average rate of correct classification for a model (mean of 10 runs). (B) Error rate for the models, sorted in decreasing order. Dashed line indicates mean performance for an explorative shallow network with 49 hidden neurons, the same as the number of feedforward units in the ConvNets shown here. Asterisks indicate error rates that are statistically significantly lower than for the shallow network ($p < 0.0005$). Data are the mean of 5 runs, error bars are the variance.

accuracy, this for both greedy or explorative approaches. On the other hand, DA transmission in the feedforward layer does boost performance. Exploration in this layer additionally provides small but statistically significant gains in accuracy (from 94.74 ± 0.25 to 96.00 ± 0.20 for the greedy and explorative strategies, respectively; $p < 0.001$).

The best performing network has greedy DA-based learning in the convolutional layer and explorative DA-based learning in the feedforward layer, reaching a mean error rate of $3.59 \pm 0.12\%$. This is statistically significantly lower than in the explorative shallow network (the best performing shallow model with the same number of representation units), which achieved a mean error rate of $4.47 \pm 0.05\%$ ($p < 0.0005$). The additional data transformation provided by the convolutional layer thus produces a relative reduction of 20% in error rates.

6.5.3 Additional Effects on Learning

Filter Size

For small convolutional kernel sizes, DA transmission in the convolutional layer of the greedy network did not improve classification results. It is therefore not a surprise that DA in the explorative network also does not enhance accuracy. To test if exploration at the level of the convolutional layer may be beneficial for larger filters, I vary their dimension and examine the impact on performance. I find that, with increasing filter size, DA in the convolutional layer raises classification scores compared to control (Fig. 6.15). However, the explorative model is not statistically significantly superior to the greedy one for DA in the convolutional filters ($p > 0.01$). Therefore, we must reject the hypothesis that exploration in the convolutional layer benefits learning.

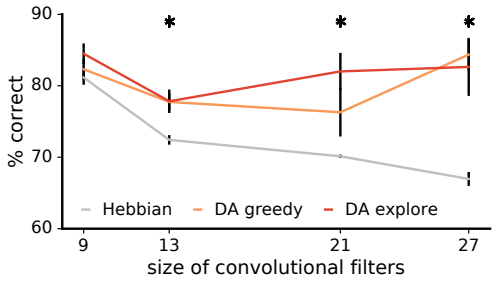

Figure 6.15: Larger filters do not allow beneficial exploration in the convolutional layer. With increasing filter dimension, DA in the convolutional layer of the explorative and greedy networks significantly enhance classification accuracy compared to Hebb's rule alone (asterisks, $p < 0.001$). However, the difference between the greedy and explorative approaches does not reach statistical significance for any filter size ($p > 0.01$).

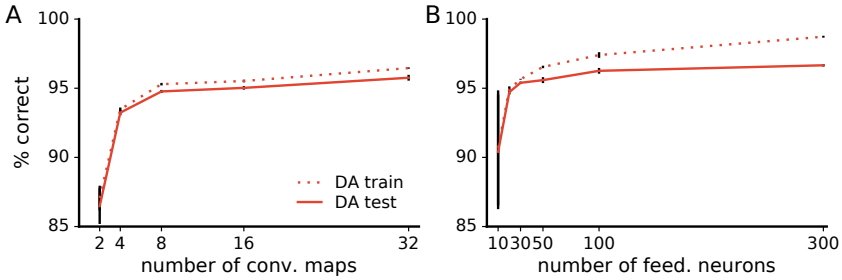

Figure 6.16: More neurons in either layer improves classification results. For the feedforward layer, however, performance saturates rather early (around 100 neurons) as gains on the training set fails to generalise to the test set.

Neuron Number

I study the influence of the number of neurons in the convolutional and feedforward layers on learning (for the convolutional layer, this corresponds to the number of feature maps). More neurons in either layer translates to enhanced performances (Fig. 6.16). In the feedforward layer, although learning keeps improving until 300 neurons on the training set, this does not generalise to the test set.

6.6 Discussion

For Hebbian learning alone, the convolutional layer cuts down the error rate by about a third compared to the shallow network (from $16.4 \pm 0.6\%$ to $11.4 \pm 0.6\%$). With DA, the decrease is of about 20% (from $4.5 \pm 0.2\%$ to $3.6 \pm 0.1\%$, for 49 representation units, see Fig. 6.17). These gains are relatively minor compared to the 5 fold increase in trainable parameters. In comparison, adding a convolutional layer to an MLP reduces the error rate by 74% (from $3.4 \pm 0.1\%$ to $0.95 \pm 0.05\%$). These results indicates that the error-backpropagation algorithm makes better use of multiple hidden layers than DA learning does.

The benefits brought by DA depends on the size of the convolutional filters. This result is in line with the function of DA in the model in enhancing neural selectivity. Restricted filter dimensions do not allow representing task-relevant features on

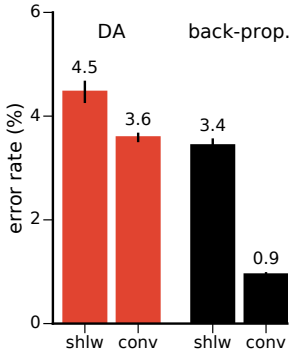

Figure 6.17: Performance comparison of DA and error-backpropagation in ConvNets. Error rates for shallow and convolutional networks of similar architectures, namely 784, 20×784 (in ConvNets only), 49, and 10 neurons. Adding a convolutional layer to the shallow network yields statistically significant benefits in performance for both DA- and error-backpropagation-based learning ($p < 0.001$). However, the magnitude of the change is more than 4 times greater in the latter than in the former.

the MNIST data and, therefore, DA may only carry a limited impact in this layer. With appropriate filter sizes, however, DA elicits advantageous modifications in both the convolutional and feedforward layers, indicating that a single RPE signal can be used concurrently in many layers. In this case, the network may serve as a model of learning in hierarchical biological networks, predicting at which processing stages weight alterations might take place.

It is uncertain why the convolutional layer did not bring more important benefits to the network. We may speculate that the DA learning mechanisms do not suitably solve the credit assignment problem. Whereas the backpropagation algorithm computes a weight update $\frac{\partial E}{\partial W_{ij}}$ for any weight W_{ij} independently, DA only carries a global signal uniquely affecting all weights. This characteristic may limit its performance. Additional mechanism such as attention [ROW10] may be required to address the credit assignment problem in deep networks.

6.7 Summary

Convolutional neural networks leverage local correlations in data to learn location invariant filters. In combination with subsampling, these filters provide tolerance to distortions, thereby increasing the generalisation power of the classifier. In this chapter, I considered the influence of DA in a Hebbian ConvNet. I showed that ConvNets perform better than their shallow equivalent and that a unique DA signal significantly improves the quality of representations in both layers of the network concurrently. However, the advantages deriving from the extra convolutional layer are marginal compared to those observed in ConvNets trained with error-backpropagation. These lower benefits might derive from poorly resolving the credit assignment problem across layers. Future should investigate mechanisms to address this deficit.

Chapter 7

Conclusion

7.1 Findings Summary

Biological and artificial neural structures represent inputs in the activation patterns of 'neurons', either living nerve cells or arrays of numbers. Although distinct in their implementations, these representations serve comparable functions: extracting relevant features from signals. Appropriately represented information enables high-level processing like perception and cognition in animals or intricate data classification in machines.

Both organic and synthetic systems have the ability to refine neural representations. In this work, I investigated the contributions of neuromodulators to such refinements. I demonstrated that two signals regulating synaptic plastic in a Hebbian learning network bring distinct and complimentary benefits to neural representations. One signal follows task demand and task relevance, approximating the cognitive construct of attentional effort. I showed that this signal raises performance by rearranging the distribution of stimulus preferences in a network. The second signal follows reward prediction errors. This learning process adapts neural weights to the reward contingencies of the task, in this way significantly improving the quality of a representation with respect to this task. I postulate that ACh and DA respectively carry these roles in mammalian sensory cortices.

The model of ACh benefits learning when the capacity of a representation is limited relative to the variability of input data. In particular, the ACh model is valuable if the difficulty or the relevance of stimuli is unevenly distributed among the data; the greater the non-uniformity, the more useful ACh is. Additionally, ACh produces comparable gains in accuracy regardless of the amount of available labelled data, indicating that ACh bears particular advantages when limited environmental feedback exists.

In the model, stimulus-wise ACh activations exhibit poorer standings than activations computed as a mean over classes. I speculated that the non-uniform distribution of ACh release disrupts learning by attributing a too great importance to detrimental data. Future work should investigate if the additional information brought by stimulus-wise ACh transmission can in some cases benefit learning.

For the model of DA, I showed that noisy neural activation gives rise a softmax exploration rule beneficial to learning. In this case, the DA activation profile matches the one reported in animals. These results suggest that DA transmission and noisy neural responses serve a synergistic function in refining neural tuning in animal sensory

cortices.

In the model, DA signalling yields error rates lower than those originally reported for the MNIST data and comparable to those of state-of-the-art optimisation methods for MLPs of the same architecture. Since evolutionary pressures must have resulted in performant learning mechanisms, these results support the notion that the model of DA is implemented in the brain. Furthermore, in addition to strong performances, neuromodulator-based learning requires weak supervision signals and interacts with information available at the synapse. These characteristics make modulator-inspired learning appealing in conditions poorly suited to the error-backpropagation algorithm, for instance with weakly labelled data or in neuromorphic processors.

Using a reward's expected value as a measure of task demand performed worse than simple binary reward predictions. It remains surprising that additional information carried by reward probabilities does not benefit the learning outcomes. In addition, the comparison of DA with error-backpropagation indicated that the greater performances of the latter might result from a more distributed code as well as from allowing negative weight values. Future work should investigate these characteristics as the may enhance the quality of a neural representation.

I conducted additional experiments to test the worth of DA in a deep convolutional neural network, with ambivalent results. On the one hand, the additional convolutional layer statistically significantly enhanced the network's accuracy, both without and with DA signalling. These experiments further showed that a unique DA signal can improve representations in both of the network's layers concurrently. The extent of the benefits depended on the amount of task-relevant information encoded in a layer. On the other hand, the improvements in performance deriving from the convolutional layer were of modest magnitude, marginal in comparison to those observed with error-backpropagation learning. These limited gains may derive from the network's poor resolution of the credit assignment problem. Together, these results support the notion that global neuromodulatory signals can direct learning in multi-layer networks but that, at the moment, they are significantly less apt at doing so than the error-backpropagation algorithm.

7.2 Outlook

This work both brings encouraging results as well as it opens questions for future research. Perhaps the most important unresolved challenge remains that of the credit assignment problem. In the model presented in this thesis, strong lateral inhibition ensured that inputs activate only a few neurons at a time, thereby effectively removing the credit assignment problem in shallow networks. However, more distributed codes appear beneficial in MLPs and, consequently, imposing too stringent sparseness seems an inadequate solution. In addition, learning in multi-layered networks raises another version of the credit assignment problem, here across layers. In this case, appropriately solving this problem may enable neuromodulator-inspired learning methods to attain the prowess of error-backpropagation in deep MLPs.

Representations play overarching functions in learning for both animals and machines. As we gain a deeper understanding of the processes contributing to the plasticity of representations in humans, we may further our ability to acquire languages later in life, for instance, or cure disabilities like phantom limb pain and perceptual agnosia.

On the other hand, novel biology-inspired learning rules in MLPs may open the way to unconventional network architectures, allow efficient learning with little supervision, or permit to train neural systems in physical devices. The inner workings of the brain's shadow theatre certainly offer exciting perspectives as subjects of future research.

Appendix A

Training with Sinusoidal Gratings

In addition to the MNIST dataset, I experimented with training networks with sinusoidal gratings, or Gabor patches (see Fig. A.1 A). On these data, the network conducts a binary classification task in which the orientation of Gabor patches must be categorised as clockwise or counter-clockwise of a target orientation. This task is homologous to training protocols in animal experiments (e.g., [Sch+01; YM04; Rai+06]). This approach produces weights that resemble sinusoidal gratings (Fig. A.1 B) and neural responses that can be characterised by tuning curves and preferred orientations (Fig. A.1 C), alike cortical cells. This correspondence is interesting if one is to compare the modifications in tuning in the model with those reported in biology. However, by itself, this task turns out too be to simple for the network as it achieves close to perfect classification even with purely Hebbian learning (Fig A.2).

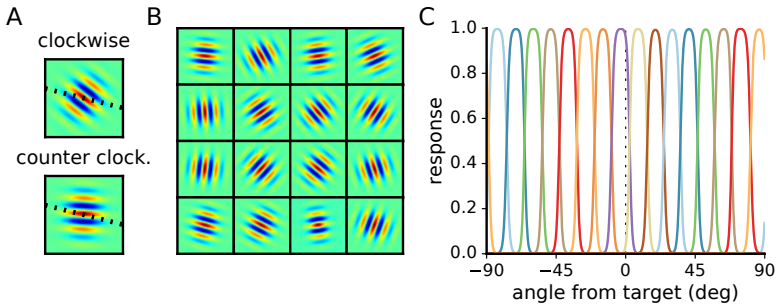

Figure A.1: Training networks with sinusoidal gradings produces well-defined tuning curves. (A) Orientation discrimination task. The network must decide whether stimuli are oriented clockwise (top) or counter-clockwise (bottom) of a target orientation (dotted line). (B) Weights learned by a Hebbian network with 16 representation neurons. (C) Tuning curves of neurons in (B). This dataset allows to compare the response properties of model neurons with those of cells in the visual cortex of animals.

I experimented with different solutions to increase task difficulty, including adding noise in the input images or in the activation of neurons. In addition, I examined increasing the strength of lateral inhibition (τ parameter of the softmax inhibition, Eq. 3.14). This method prevents the Hebbian network from performing the task optimally as it enforces too sharp tuning curves (see Fig. A.3 C). Specifically, the imperfect learning outcome is due to the target orientation at times falling at a flat portion of the neurons' tuning curves where they poorly encode orientation information (Fig. A.3 A).

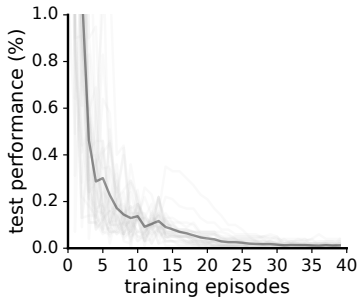

Figure A.2: The network achieves close to perfect performance on the orientation discrimination task. Test performance of Hebbian networks on the binary orientation classification task. Light traces are individual runs, dark trace is the mean of 30 runs.

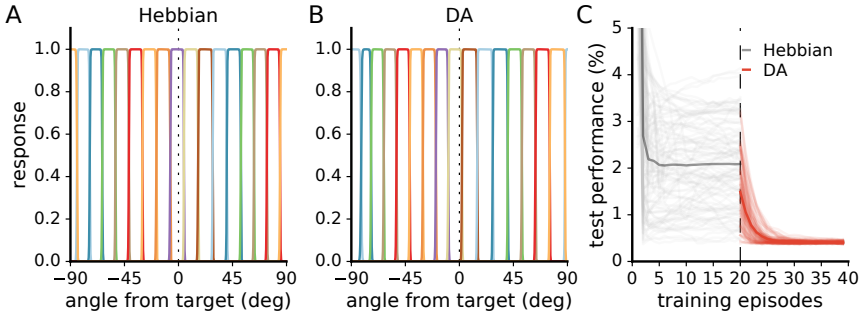

Figure A.3: Increasing lateral inhibition prevents perfect performance. (A) Tuning curves for Hebbian and (B) DA learning. (C) Test performance of networks trained with Hebb's rule (grey) and supplemented with DA (red traces). Light traces are individual runs, dark trace is the mean of 100 runs. DA significantly enhances performance by matching the slope of the tuning curves with the discrimination boundary.

Training with DA matches the tuning curves slopes with the target orientation, enhancing discrimination abilities (Fig. A.3 B-C). The slopes of tuning curves carry maximal information about the stimuli and performance is thus greater where their density is the highest (Fig. A.4), in agreement with animal experiments [Sch+01]. For future work, using homologous datasets to train animals and models may allow to test the validity of a model and to make predictions about learning outcomes in animals.

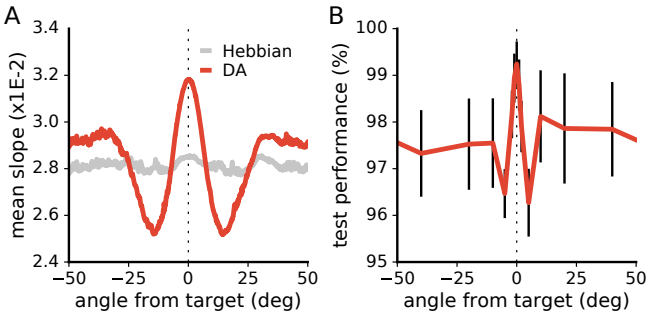

Figure A.4: Performance is highest where slope density is maximal. (A) Mean slope of neurons, averaged for 100 networks. (B) Discrimination performance for different orientations. Data are mean of 100 networks, error bars are variance. Across networks, classification accuracy is maximal for regions of the input space where slope density is at its highest, and vice-versa.

Bibliography

[Aba+15] Martín Abadi et al. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. Software available from tensorflow.org. 2015. URL: `http://tensorflow.org/`.

[AH93] Merav Ahissar and Shaul Hochstein. "Attentional control of early perceptual learning." In: *Proceedings of the National Academy of Sciences of the United States of America* 90.12 (1993), p. 5718.

[Ang+97] Alessandra Angelucci et al. "Experimentally induced retinal projections to the ferret auditory thalamus: development of clustered eye-specific patterns in a novel target". In: *Journal of Neuroscience* 17.6 (1997), pp. 2040–2055.

[Arn+02] HM Arnold et al. "Differential cortical acetylcholine release in rats performing a sustained attention task versus behavioral control tasks that do not explicitly tax attention". In: *Neuroscience* 114.2 (2002), pp. 451–460.

[Ars+13] John T Arsenault et al. "Dopaminergic reward signals selectively decrease fMRI activity in primate visual cortex". In: *Neuron* 77.6 (2013), pp. 1174–1186.

[Ass+07] Collins Assisi et al. "Adaptive regulation of sparseness by feedforward inhibition". In: *Nature Neuroscience* 10.9 (2007), pp. 1176–1184.

[Bao+03] Shaowen Bao et al. "Suppression of cortical representation through backward conditioning". In: *Proceedings of the National Academy of Sciences* 100.3 (2003), pp. 1405–1408.

[Bao+04] Shaowen Bao et al. "Temporal plasticity in the primary auditory cortex induced by operant perceptual learning". In: *Nature Neuroscience* 7.9 (2004), pp. 974–981.

[Bar+10] Laura Baroncelli et al. "Experience-dependent reactivation of ocular dominance plasticity in the adult visual cortex". In: *Experimental neurology* 226.1 (2010), pp. 100–109.

[BAS92] Susanne Bröcher, Alain Artola, and Wolf Singer. "Agonists of cholinergic and noradrenergic receptors facilitate synergistically the induction of long-term potentiation in slices of rat visual cortex". In: *Brain Research* 573.1 (1992), pp. 27–36.

[BC+94] Carol A Bennett-Clarke et al. "Effect of serotonin depletion on vibrissa-related patterns of thalamic afferents in the rat's somatosensory cortex". In: *Journal of Neuroscience* 14.12 (1994), pp. 7594–7607.

[BCM01] Shaowen Bao, Vincent T Chan, and Michael M Merzenich. "Cortical remodelling induced by activity of ventral tegmental dopamine neurons". In: *Nature* 412.6842 (2001), pp. 79–83.

[BCO02] Olivier Blond, Francis Crépel, and Satoru Otani. "Long-term potentiation in rat prefrontal slices facilitated by phased application of dopamine". In: *European Journal of Pharmacology* 438.1 (2002), pp. 115–116.

[BD07] Anders Björklund and Stephen B Dunnett. "Dopamine neuron systems in the brain: an update". In: *Trends in neurosciences* 30.5 (2007), pp. 194–202.

[BDV01] Yoshua Bengio, Réjean Ducharme, and Pascal Vincent. "A Neural Probabilistic Language Model". In: *Advances in neural information processing systems*. Citeseer. 2001.

[Bei+03] Ralph E Beitel et al. "Reward-dependent plasticity in the primary auditory cortex of adult monkeys trained to discriminate temporally modulated signals". In: *Proceedings of the National Academy of Sciences* 100.19 (2003), pp. 11070–11075.

[Bek93] JM Bekkers. "Enhancement by histamine of NMDA-mediated synaptic transmission in the hippocampus". In: *Science* 261.5117 (1993), pp. 104–106.

[BH95] Allen E Butt and Gordon K Hodge. "Acquisition, retention, and extinction of operant discriminations in rats with nucleus basalis magnocellularis lesions." In: *Behavioral Neuroscience* 109.4 (1995), p. 699.

[BHL03] Stephanie Bissière, Yann Humeau, and Andreas Lüthi. "Dopamine gates LTP induction in lateral amygdala by suppressing feedforward inhibition". In: *Nature Neuroscience* 6.6 (2003), pp. 587–592.

[BIP04] Mel Brown, Dexter RF Irvine, and Valerie N Park. "Perceptual learning on an auditory frequency discrimination task by cats: association with changes in primary auditory cortex". In: *Cerebral Cortex* 14.9 (2004), pp. 952–965.

[BPH11] Tania Rinaldi Barkat, Daniel B Polley, and Takao K Hensch. "A critical period for auditory thalamocortical connectivity". In: *Nature Neuroscience* 14.9 (2011), pp. 1189–1194.

[Bre+02] Björn Brembs et al. "Operant reward learning in Aplysia: neuronal correlates and mechanisms". In: *Science* 296.5573 (2002), pp. 1706–1709.

[BS12] Anna Byers and John T Serences. "Exploring the relationship between perceptual learning and top-down attentional control". In: *Vision Research* 74 (2012), pp. 30–39.

[BS86] Mark F Bear and Wolf Singer. "Modulation of visual cortical plasticity by acetylcholine and noradrenaline". In: *Letters to Nature* 320 (1986), pp. 172–176.

[BSH97] KA Baskerville, JB Schweitzer, and P Herron. "Effects of cholinergic depletion on experience-dependent plasticity in the cortex of the rat". In: *Neuroscience* 80.4 (1997), pp. 1159–1169.

[BSS11] Michael Brosch, Elena Selezneva, and Henning Scheich. "Representation of reward feedback in primate auditory cortex". In: *Frontiers in systems neuroscience* 5 (2011), p. 5.

[BTD97] Allen E Butt, Guy Testylier, and Robert W Dykes. "Acetylcholine release in rat frontal and somatosensory cortex is enhanced during tactile discrimination learning". In: *Psychobiology* 25.1 (1997), pp. 18–33.

[Buc+10] K. A. Buchanan et al. "Facilitation of Long-Term Potentiation by Muscarinic M1 Receptors Is Mediated by Inhibition of SK Channels". In: *Neuron* 68.5 (2010), pp. 948–963.

[Cad+14] Charles F Cadieu et al. "Deep neural networks rival the representation of primate IT cortex for core visual object recognition". In: *PLoS Computational Biology* 10.12 (2014), e1003963.

[Cal+07] Paolo Calabresi et al. "Dopamine-mediated regulation of corticostriatal synaptic plasticity". In: *Trends in Neurosciences* 30.5 (2007), pp. 211–219.

[Can+05] J. Cang et al. "Development of precise maps in visual cortex requires patterned spontaneous activity in the retina". In: *Neuron* 48.5 (2005), pp. 797–809.

[Cas+96] Olivier Cases et al. "Lack of barrels in the somatosensory cortex of monoamine oxidase a–deficient mice: role of a serotonin excess during the critical period". In: *Neuron* 16.2 (1996), pp. 297–307.

[CCT05] James M Conner, Andrea A Chiba, and Mark H Tuszynski. "The basal forebrain cholinergic system is essential for cortical plasticity and functional recovery following brain injury". In: *Neuron* 46.2 (2005), pp. 173–179.

[Cen+99] Diego Centonze et al. "Unilateral dopamine denervation blocks corticostriatal LTP". In: *Journal of Neurophysiology* 82.6 (1999), pp. 3575–3579.

[CGS98] Michael C. Crair, Deda C. Gillespie, and Michael P. Stryker. "The role of visual experience in the development of columns in cat visual cortex". In: *Science* 279.5350 (1998), pp. 566–570.

[Chu+13] Sungkun Chun et al. "Thalamocortical long-term potentiation becomes gated after the early critical period in the auditory cortex". In: *The Journal of neuroscience: the official journal of the Society for Neuroscience* 33.17 (2013), pp. 7345–7357.

[Cio+12] T Ciodaro et al. "Online particle detection with neural networks based on topological calorimetry information". In: *Journal of physics: conference series*. Vol. 368. 1. IOP Publishing. 2012, p. 012030.

[CK00] Justin C Crowley and Lawrence C Katz. "Early development of ocular dominance columns". In: *Science* 290.5495 (2000), pp. 1321–1324.

[CKT10] JM Conner, M Kulczycki, and MH Tuszynski. "Unique contributions of distinct cholinergic projections to motor cortical plasticity and learning". In: *Cerebral Cortex* 20.11 (2010), pp. 2739–2748.

[CMA07] Jonathan D Cohen, Samuel M McClure, and J Yu Angela. "Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration". In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 362.1481 (2007), pp. 933–942.

[CMS12] Dan Ciresan, Ueli Meier, and Jürgen Schmidhuber. "Multi-column deep neural networks for image classification". In: *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*. IEEE. 2012, pp. 3642–3649.

[Con+03] James M Conner et al. "Lesions of the basal forebrain cholinergic system impair task acquisition and abolish cortical plasticity associated with motor skill learning". In: *Neuron* 38.5 (2003), pp. 819–829.

[CS93] Barbara Chapman and Michael P Stryker. "Development of orientation selectivity in ferret visual cortex and effects of deprivation". In: *Journal of Neuroscience* 13.12 (1993), pp. 5251–5262.

[CSW09] Hoon Choi, Aaron R Seitz, and Takeo Watanabe. "When attention interrupts learning: inhibitory effects of attention on TIPL". In: *Vision Research* 49.21 (2009), p. 2586.

[CT88] Michael B Calford and Rowan Tweedale. "Immediate and chronic changes in responses of somatosensory cortex in adult flying-fox after digit amputation". In: *Nature* 332.6163 (1988), pp. 446–448.

[CVN79] ML Crawford and GK Von Noorden. "The effects of short-term experimental strabismus on the visual system in Macaca mulatta." In: *Investigative Ophthalmology & Visual Science* 18.5 (1979), pp. 496–505.

[DA01] Peter Dayan and Laurence F Abbott. *Theoretical neuroscience*. Vol. 806. Cambridge, MA: MIT Press, 2001.

[Dah+12] George E Dahl et al. "Context-dependent pre-trained deep neural networks for large-vocabulary speech recognition". In: *Audio, Speech, and Language Processing, IEEE Transactions on* 20.1 (2012), pp. 30–42.

[Dal+01] Jeffrey W Dalley et al. "Distinct changes in cortical acetylcholine and noradrenaline efflux during contingent and noncontingent performance of a visual attentional task". In: *The Journal of neuroscience: the official journal of the Society for Neuroscience* 21.13 (2001), pp. 4908–4914.

[Daw+06] Nathaniel D Daw et al. "Cortical substrates for exploratory decisions in humans". In: *Nature* 441.7095 (2006), pp. 876–879.

[Din+03] Hubert R Dinse et al. "Pharmacological modulation of perceptual learning and associated cortical reorganization". In: *Science Signaling* 301.5629 (2003), p. 91.

[DLR77] Arthur P Dempster, Nan M Laird, and Donald B Rubin. "Maximum likelihood from incomplete data via the EM algorithm". In: *Journal of the royal statistical society. Series B (methodological)* (1977), pp. 1–38.

[DSG95] C Darian-Smith and CD Gilbert. "Topographic reorganization in the striate cortex of the adult cat and monkey is cortically mediated". In: *The Journal of neuroscience: the official journal of the Society for Neuroscience* 15.3 (1995), pp. 1631–1647.

[DWW77] Harriet De Wit and Roy A Wise. "Blockade of cocaine reinforcement in rats with the dopamine receptor blocker pimozide, but not with the noradrenergic blockers phentolamine or phenoxybenzamine." In: *Canadian Journal of Psychology/Revue canadienne de psychologie* 31.4 (1977), p. 195.

[Eas+02] A Easton et al. "Unilateral lesions of the cholinergic basal forebrain and fornix in one hemisphere and inferior temporal cortex in the opposite hemisphere produce severe learning impairments in rhesus monkeys". In: *Cerebral Cortex* 12.7 (2002), pp. 729–736.

[ER04] Jos J Eggermont and Larry E Roberts. "The neuroscience of tinnitus". In: *Trends in neurosciences* 27.11 (2004), pp. 676–682.

[FI84] Yves Frégnac and Michel Imbert. "Development of neuronal selectivity in primary visual cortex of cat." In: *Physiological Reviews* 64.1 (1984), pp. 325–434.

[FM11] Robert C Froemke and Ana Raquel O Martins. "Spectrotemporal dynamics of auditory cortical synaptic receptive field plasticity". In: *Hearing Research* 279.1 (2011), pp. 149–161.

[FM82] Kunihiko Fukushima and Sei Miyake. "Neocognitron: A new algorithm for pattern recognition tolerant of deformations and shifts in position". In: *Pattern recognition* 15.6 (1982), pp. 455–469.

[FMS07] R. C. Froemke, M. M. Merzenich, and C. E. Schreiner. "A synaptic memory trace for cortical receptive field plasticity". In: *Nature* 450.7168 (2007), pp. 425–429.

[FNJ06] Herta Flor, Lone Nikolajsen, and Troels Staehelin Jensen. "Phantom limb pain: a case of maladaptive CNS plasticity?" In: *Nature Reviews Neuroscience* 7.11 (2006), pp. 873–881.

[Fri+03] Jonathan Fritz et al. "Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex". In: *Nature Neuroscience* 6.11 (2003), pp. 1216–1223.

[Fro+13] Robert C Froemke et al. "Long-term modification of cortical synapses improves sensory perception". In: *Nature Neuroscience* 16.1 (2013), pp. 79–88.

[FS95] Yoav Freund and Robert E Schapire. "A desicion-theoretic generalization of on-line learning and an application to boosting". In: *European conference on computational learning theory*. Springer. 1995, pp. 23–37.

[FSG01] Tomomi Fukuchi-Shimogori and Elizabeth A Grove. "Neocortex patterning by the secreted signaling molecule FGF8". In: *Science* 294.5544 (2001), pp. 1071–1074.

[FSL16] Dennis Forster, Abdul-Saboor Sheikh, and Jörg Lücke. "Neural Simpletrons–Minimalistic Directed Generative Networks for Learning with Few Labels". In: *Stat; Bulletin of the Wisconsin Nurses Association* 1050 (2016), p. 23.

[FSR04] GN Fournier, K Semba, and DD Rasmusson. "Modality-and region-specific acetylcholine release in the rat neocortex". In: *Neuroscience* 126.2 (2004), pp. 257–262.

[FSV10] Edit Frankó, Aaron R Seitz, and Rufin Vogels. "Dissociable neural effects of long-term stimulus–reward pairing in Macaque Visual Cortex". In: *Journal of Cognitive Neuroscience* 22.7 (2010), pp. 1425–1439.

[FTS03] Christopher D Fiorillo, Philippe N Tobler, and Wolfram Schultz. "Discrete coding of reward probability and uncertainty by dopamine neurons". In: *Science* 299.5614 (2003), pp. 1898–1902.

[Fux+05] K Fuxe et al. "Dynamics of volume transmission in the brain. Focus on catecholamine and opioid peptide communication and the role of uncoupling protein 2". In: *Journal of neural transmission* 112.1 (2005), pp. 65–76.

[FVE91] Daniel J Felleman and David C Van Essen. "Distributed hierarchical processing in the primate cerebral cortex". In: *Cerebral Cortex* 1.1 (1991), pp. 1–47.

[FW02] Max L Fletcher and Donald A Wilson. "Experience modifies olfactory acuity: acetylcholine-dependent learning decreases behavioral generalization between similar odorants". In: *The Journal of neuroscience: the official journal of the Society for Neuroscience* 22.2 (2002), RC201.

[Gal+82] CR Gallistel et al. "Does pimozide block the reinforcing effect of brain stimulation?" In: *Pharmacology Biochemistry and Behavior* 17.4 (1982), pp. 769–781.

[Gar+02] Sonia Garel et al. "The early topography of thalamocortical projections is shifted in Ebf1 and Dlx1/2 mutant mice". In: *Development* 129.24 (2002), pp. 5621–5634.

[GE06] Dimitrios V Giannikopoulos and Ulf T Eysel. "Dynamics and specificity of cortical map reorganization after retinal lesions". In: *Proceedings of the National Academy of Sciences* 103.28 (2006), pp. 10805–10810.

[GK02] Wulfram Gerstner and Werner M Kistler. "Mathematical formulations of Hebbian learning". In: *Biological cybernetics* 87.5-6 (2002), pp. 404–415.

[GLP09] Charles D Gilbert, Wu Li, and Valentin Piech. "Perceptual learning and adult cortical plasticity". In: *The Journal of physiology* 587.12 (2009), pp. 2743–2751.

[GLS88] Joachim M Greuel, Heiko J Luhmann, and Wolf Singer. "Pharmacological induction of use-dependent receptive field modifications in the visual cortex". In: *Science* 242.4875 (1988), pp. 74–77.

[God+02] Ben Godde et al. "Plasticity of orientation preference maps in the visual cortex of adult cats". In: *Proceedings of the National Academy of Sciences* 99.9 (2002), pp. 6352–6357.

[GR04] Sonia Garel and John LR Rubenstein. "Intermediate targets in formation of topographic projections: inputs from the thalamocortical system". In: *Trends in neurosciences* 27.9 (2004), pp. 533–539.

[Gra05] Ann M Graybiel. “The basal ganglia: learning new tricks and loving it”. In: *Current Opinion in Neurobiology* 15.6 (2005), pp. 638–644.

[GS10] Andrew J Giessel and Bernardo L Sabatini. “M1 muscarinic receptors boost synaptic potentials and calcium influx in dendritic spines by inhibiting postsynaptic SK channels”. In: *Neuron* 68.5 (2010), pp. 936–947.

[GS93] Qiang Gu and Wolf Singer. “Effects of intracortical infusion of anticholinergic drugs on neuronal plasticity in kitten striate cortex”. In: *European Journal of Neuroscience* 5.5 (1993), pp. 475–485.

[GS95] Qiang Gu and Wolf Singer. “Involvement of serotonin in developmental plasticity of kitten visual cortex”. In: *European Journal of Neuroscience* 7.6 (1995), pp. 1146–1153.

[GSK92] P Gaspar, I Stepniewska, and JH Kaas. “Topography and collateralization of the dopaminergic projections to motor and lateral prefrontal cortex in owl monkeys”. In: *Journal of Comparative Neurology* 325.1 (1992), pp. 1–21.

[GSW14] Kalanit Grill-Spector and Kevin S Weiner. “The functional architecture of the ventral temporal cortex and its role in categorization”. In: *Nature Reviews Neuroscience* 15.8 (2014), pp. 536–548.

[GSW81] Gary J Gerber, John Sing, and Roy A Wise. “Pimozide attenuates lever pressing for water reinforcement in rats”. In: *Pharmacology Biochemistry and Behavior* 14.2 (1981), pp. 201–205.

[Gu02] Q Gu. “Neuromodulatory transmitter systems in the cortex and their role in cortical plasticity”. In: *Neuroscience* 111.4 (2002), pp. 815–835.

[Gu03] Qiang Gu. “Contribution of acetylcholine to visual cortex plasticity”. In: *Neurobiology of Learning and Memory* 80.3 (2003), pp. 291–301.

[Haj+05] Greg Hajcak et al. “On the ERN and the significance of errors”. In: *Psychophysiology* 42.2 (2005), pp. 151–160.

[Hal+93] Peter W Halligan et al. “Thumb in cheek? Sensory reorganization and perceptual plasticity after limb amputation.” In: *Neuroreport* 4.3 (1993), pp. 233–236.

[Ham+04] Tadashi Hamasaki et al. “EMX2 regulates sizes and positioning of the primary sensory and motor areas in neocortex by direct specification of cortical progenitors”. In: *Neuron* 43.3 (2004), pp. 359–372.

[Han+07] Yoon K Han et al. “Early experience impairs perceptual discrimination”. In: *Nature Neuroscience* 10.9 (2007), pp. 1191–1197.

[Har+97] R Hartmann et al. “Response of the primary auditory cortex to electrical stimulation of the auditory nerve in the congenitally deaf white cat”. In: *Hearing research* 112.1 (1997), pp. 115–133.

[He+15] Kaiwen He et al. “Distinct Eligibility Traces for LTP and LTD in Cortical Synapses”. In: *Neuron* 88.3 (2015), pp. 528–538.

[Heb49] Donald Olding Hebb. *The organization of behavior*. New York: Wiley and Sons, 1949.

[Hel+13] Moritz Helmstaedter et al. “Connectomic reconstruction of the inner plexiform layer in the mouse retina”. In: *Nature* 500.7461 (2013), pp. 168–174.

[Hen04] Takao K Hensch. “Critical period regulation”. In: *Annual Review of Neuroscience* 27 (2004), pp. 549–579.

[Hen05] T. K. Hensch. “Critical period plasticity in local cortical circuits”. In: *Nature Reviews. Neuroscience* 6.11 (2005), pp. 877–888.

[HHD01] Justin A Harris, Irina M Harris, and Mathew E Diamond. “The topography of tactile learning in humans”. In: *The Journal of neuroscience: the official journal of the Society for Neuroscience* 21.3 (2001), pp. 1056–1061.

[HIK87] P Heggelund, K Imamura, and T Kasamatsu. “Reduced binocularity in the noradrenaline-infused striate cortex of acutely anesthetized and paralyzed, otherwise normal cats”. In: *Experimental brain research* 68.3 (1987), pp. 593–605.

[Hin+12] Geoffrey Hinton et al. “Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups”. In: *Signal Processing Magazine, IEEE* 29.6 (2012), pp. 82–97.

[HK10] Suzanne N Haber and Brian Knutson. “The reward circuit: linking primate anatomy and human imaging”. In: *Neuropsychopharmacology* 35.1 (2010), pp. 4–26.

[Hor00] Jon C Horvitz. “Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events”. In: *Neuroscience* 96.4 (2000), pp. 651–656.

[Hos+11] Jonas A Hosp et al. “Dopaminergic projections from midbrain to primary motor cortex mediate motor skill learning”. In: *The Journal of neuroscience: the official journal of the Society for Neuroscience* 31.7 (2011), pp. 2481–2487.

[HS11] Michael E Hasselmo and Martin Sarter. “Modes and models of forebrain cholinergic neuromodulation of cognition”. In: *Neuropsychopharmacology* 36.1 (2011), pp. 52–73.

[HSB00] Anne Marie Himmelheber, Martin Sarter, and John P Bruno. “Increases in cortical acetylcholine release during sustained attention performance in rats”. In: *Cognitive Brain Research* 9.3 (2000), pp. 313–325.

[HVE00] Jay Hegdé and David C Van Essen. “Selectivity for complex shapes in primate visual area V2”. In: *The Journal of neuroscience: the official journal of the Society for Neuroscience* 20.5 (2000), pp. 61–61.

[HW62] David H Hubel and Torsten N Wiesel. “Receptive fields, binocular interaction and functional architecture in the cat’s visual cortex”. In: *The Journal of physiology* 160.1 (1962), pp. 106–154.

[IB75] Michel Imbert and Pierre Buisseret. “Receptive field characteristics and plastic properties of visual cortical cells in kittens reared with or without visual experience”. In: *Experimental Brain Research* 22.1 (1975), pp. 25–36.

[IS11] J. S. Isaacson and M. Scanziani. “How inhibition shapes cortical activity”. In: *Neuron* 72.2 (2011), pp. 231–243.

[Jen+90] William M Jenkins et al. “Functional reorganization of primary somatosensory cortex in adult owl monkeys after behaviorally controlled tactile stimulation”. In: *Journal of neurophysiology* 63.1 (1990), pp. 82–104.

[JGS01] Weiqing Ji, Enquan Gao, and Nobuo Suga. “Effects of acetylcholine and atropine on plasticity of central auditory neurons caused by conditioning in bats”. In: *Journal of Neurophysiology* 86.1 (2001), pp. 211–225.

[JME91] Sharon L Juliano, Wu Ma, and Don Eslin. “Cholinergic depletion prevents expansion of topographic maps in somatosensory cortex”. In: *Proceedings of the National Academy of Sciences* 88.3 (1991), pp. 780–784.

[JP87] Judson P Jones and Larry A Palmer. “An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex”. In: *Journal of neurophysiology* 58.6 (1987), pp. 1233–1258.

[JPP09] Ajay J Joshi, Fatih Porikli, and Nikolaos Papanikolopoulos. “Multi-class active learning for image classification”. In: *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*. IEEE. 2009, pp. 2372–2379.

[KB14] Diederik Kingma and Jimmy Ba. “Adam: A method for stochastic optimization”. In: *arXiv preprint arXiv:1412.6980* abs/1412.6980 (2014).

[KBS06] Rouba Kozak, John P Bruno, and Martin Sarter. “Augmented prefrontal acetylcholine release during challenged attentional performance”. In: *Cerebral Cortex* 16.1 (2006), pp. 9–17.

[KE07] Andrej Kral and Jos J Eggermont. “What’s to lose and what’s to learn: development under auditory deprivation, cochlear implants and limits of cortical plasticity”. In: *Brain Research Reviews* 56.1 (2007), pp. 259–269.

[Kep+08] Adam Kepecs et al. “Neural correlates, computation and behavioural impact of decision confidence”. In: *Nature* 455.7210 (2008), pp. 227–231.

[KH00] Jon H Kaas and Troy A Hackett. “Subdivisions of auditory cortex and processing streams in primates”. In: *Proceedings of the National Academy of Sciences* 97.22 (2000), pp. 11793–11799.

[Kit+94] Cheryl A Kitt et al. “Cholinergic innervation of mouse forebrain structures”. In: *Journal of Comparative Neurology* 341.1 (1994), pp. 117–129.

[KK84] BD Kuppermann and T Kasamatsu. “Enhanced binocular interaction in the visual cortex of normal kittens subjected to intracortical norepinephrine perfusion”. In: *Brain research* 302.1 (1984), pp. 91–99.

[KKF00] Gabriel Kreiman, Christof Koch, and Itzhak Fried. “Category-specific visual responses of single neurons in the human medial temporal lobe”. In: *Nature neuroscience* 3.9 (2000), pp. 946–953.

[Kle+07] Tilmann A Klein et al. “Genetically determined differences in learning from errors”. In: *Science* 318.5856 (2007), pp. 1642–1645.

[KM96] Yacoub Khallad and Jay Moore. "Blocking, unblocking, and overexpectation in autoshaping with pigeons". In: *Journal of the experimental analysis of behavior* 65.3 (1996), pp. 575–591.

[KM98a] Michael P Kilgard and Michael M Merzenich. "Cortical map reorganization enabled by nucleus basalis activity". In: *Science* 279.5357 (1998), pp. 1714–1718.

[KM98b] Michael P Kilgard and Michael M Merzenich. "Plasticity of temporal information processing in the primary auditory cortex". In: *Nature Neuroscience* 1.8 (1998), pp. 727–731.

[KP76] Takuji Kasamatsu and John D Pettigrew. "Depletion of brain catecholamines: failure of ocular dominance shift after monocular occlusion in kittens". In: *Science* 194.4261 (1976), pp. 206–209.

[KR89] Arlette Kolta and Tomás A Reader. "Modulatory effects of catecholamines on neurons of the rat visual cortex: single-cell iontophoretic studies". In: *Canadian journal of physiology and pharmacology* 67.6 (1989), pp. 615–623.

[KS06] Masaharu Kudoh and Katsuei Shibuki. "Sound sequence discrimination learning motivated by reward requires dopaminergic D2 receptor activation in the rat auditory cortex". In: *Learning & Memory* 13.6 (2006), pp. 690–698.

[KSH12] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. "Imagenet classification with deep convolutional neural networks". In: *Advances in neural information processing systems*. 2012, pp. 1097–1105.

[KSL12] Christian Keck, Cristina Savin, and Jörg Lücke. "Feedforward Inhibition and Synaptic Scaling–Two Sides of the Same Coin?" In: *PLoS Computational Biology* 8.3 (2012), e1002432.

[KT94] Eucaly Kobatake and Keiji Tanaka. "Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex". In: *Journal of Neurophysiology* 71 (1994), pp. 856–856.

[Kur15] Andrey Kurenkov. *A 'Brief' History of Neural Nets and Deep Learning*. Dec. 2015. URL: `http://www.andreykurenkov.com/writing/a-brief-history-of-neural-nets-and-deep-learning/`.

[LBH15] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. "Deep learning". In: *Nature* 521.7553 (2015), pp. 436–444.

[LC13] Guoshi Li and Thomas A Cleland. "A two-layer biophysical model of cholinergic neuromodulation in olfactory bulb". In: *The Journal of neuroscience: the official journal of the Society for Neuroscience* 33.7 (2013), pp. 3037–3058.

[LCB98] Yann LeCun, Corinna Cortes, and Christopher JC Burges. *The MNIST database of handwritten digits*. 1998.

[LCH88] Barry E Levin, Rebecca L Craik, and Peter J Hand. "The role of norepinephrine in adult rat somatosensory (SmI) cortical metabolism and plasticity". In: *Brain research* 443.1 (1988), pp. 261–271.

[Lea+13] Nicholas D Leach et al. "Cortical Cholinergic Input Is Required for Normal Auditory Perception and Experience-Dependent Plasticity in Adult Ferrets". In: *The Journal of neuroscience: the official journal of the Society for Neuroscience* 33.15 (2013), pp. 6659–6671.

[LeC+98] Yann LeCun et al. "Gradient-based learning applied to document recognition". In: *Proceedings of the IEEE* 86.11 (1998), pp. 2278–2324.

[Let+11] Johannes J Letzkus et al. "A disinhibitory microcircuit for associative fear learning in the auditory cortex". In: *Nature* 480.7377 (2011), pp. 331–335.

[LG09] Chi-Tat Law and Joshua I Gold. "Reinforcement learning can account for associative and perceptual learning on a visual decision task". In: *Nature Neuroscience* 12.5 (2009), p. 655.

[Li+03] Shaomin Li et al. "Dopamine-dependent facilitation of LTP induction in hippocampal CA1 by exposure to spatial novelty". In: *Nature Neuroscience* 6.5 (2003), pp. 526–531.

[Lid+91] MS Lidow et al. "Distribution of dopaminergic receptors in the primate cerebral cortex: quantitative autoradiographic analysis using [3 H] raclopride,[3 H] spiperone and [3 H] SCH23390". In: *Neuroscience* 40.3 (1991), pp. 657–671.

[LLD10] Jiajuan Liu, Zhong-Lin Lu, and Barbara A Dosher. "Augmented Hebbian reweighting: Interactions between feedback and training accuracy in perceptual learning". In: *Journal of Vision* 10.10 (2010), p. 29.

[LS09] Andreas R Luft and Stefanie Schwarz. "Dopaminergic signals in primary motor cortex". In: *International Journal of Developmental Neuroscience* 27.5 (2009), pp. 415–421.

[Ma+15] Junshui Ma et al. "Deep neural nets as a method for quantitative structure–activity relationships". In: *Journal of chemical information and modeling* 55.2 (2015), pp. 263–274.

[Man+05] Valerio Mante et al. "Independence of luminance and contrast in natural scenes and in the early visual system". In: *Nature Neuroscience* 8.12 (2005), pp. 1690–1697.

[MBG95] Ralph R Miller, Robert C Barnet, and Nicholas J Grahame. "Assessment of the Rescorla-Wagner model." In: *Psychological bulletin* 117.3 (1995), p. 363.

[McG+02] Jill McGaughy et al. "Selective behavioral and neurochemical effects of cholinergic lesions produced by intrabasalis infusions of 192 IgG-saporin on attentional performance in a five-choice serial reaction time task". In: *The Journal of neuroscience: the official journal of the Society for Neuroscience* 22.5 (2002), pp. 1905–1913.

[MCW08] Kerry McAlonan, James Cavanaugh, and Robert H Wurtz. "Guarding the gateway to cortex with attention in visual thalamus". In: *Nature* 456.7220 (2008), pp. 391–394.

[ME83] JH Maunsell and DAVID C van Essen. "The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey". In: *The Journal of neuroscience: the official journal of the Society for Neuroscience* 3.12 (1983), pp. 2563–2586.

[Mes+83] M Mesulam et al. "Cholinergic innervation of cortex by the basal forebrain: cytochemistry and cortical connections of the septal area, diagonal band nuclei, nucleus basalis (substantia innominata), and hypothalamus in the rhesus monkey". In: *Journal of Comparative Neurology* 214.2 (1983), pp. 170–197.

[Mes+92] M Mesulam et al. "Differential cholinergic innervation within functional subdivisions of the human cerebral cortex: a choline acetyltransferase study". In: *Journal of Comparative Neurology* 318.3 (1992), pp. 316–328.

[MH08] Laurens van der Maaten and Geoffrey Hinton. "Visualizing data using t-SNE". In: *Journal of Machine Learning Research* 9.Nov (2008), pp. 2579–2605.

[Mik+13a] Tomas Mikolov et al. "Distributed representations of words and phrases and their compositionality". In: *Advances in neural information processing systems*. 2013, pp. 3111–3119.

[Mik+13b] Tomas Mikolov et al. "Efficient estimation of word representations in vector space". In: *arXiv preprint arXiv:1301.3781* (2013).

[MK13] Shreesh P Mysore and Eric I Knudsen. "A shared inhibitory circuit for both exogenous and endogenous control of stimulus selection". In: *Nature neuroscience* 16.4 (2013), pp. 473–478.

[MKH05] Wolfgang Mittmann, Ursula Koch, and Michael Häusser. "Feed-forward inhibition shapes the spike output of cerebellar Purkinje cells". In: *The Journal of physiology* 563.2 (2005), pp. 369–378.

[ML+09] Katiuska Molina-Luna et al. "Dopamine in motor cortex is necessary for skill learning and synaptic plasticity". In: *PLoS One* 4.9 (2009), e7082.

[MMO06] Yoshiki Matsuda, Aude Marzo, and Satoru Otani. "The presence of background dopamine signal converts long-term synaptic depression to potentiation in rat prefrontal cortex". In: *The Journal of neuroscience: the official journal of the Society for Neuroscience* 26.18 (2006), pp. 4803–4810.

[Mni+13] Volodymyr Mnih et al. "Playing atari with deep reinforcement learning". In: *arXiv preprint arXiv:1312.5602* (2013).

[Mni+15] Volodymyr Mnih et al. "Human-level control through deep reinforcement learning". In: *Nature* 518.7540 (2015), pp. 529–533.

[MO16] Robert Meyer and Klaus Obermayer. "pypet: A Python Toolkit for Data Management of Parameter Explorations". In: *Frontiers in neuroinformatics* 10 (2016).

[MP09] N. J. Mitra and M. Pauly. "Shadow Art". In: *ACM Transactions on Graphics* 28.5 (2009), 156:1–156:7.

[MP69] Marvin Minsky and Seymour Papert. "Perceptrons." In: (1969).

[MS04] Xiaofeng Ma and Nobuo Suga. “Lateral inhibition for center-surround reorganization of the frequency map of bat auditory cortex”. In: *Journal of Neurophysiology* 92.6 (2004), pp. 3192–3199.

[MS05] Xiaofeng Ma and Nobuo Suga. “Long-term cortical plasticity evoked by electric stimulation and acetylcholine applied to the auditory cortex”. In: *Proceedings of the National Academy of Sciences of the United States of America* 102.26 (2005), pp. 9335–9340.

[MS95] Jill McGaughy and Martin Sarter. “Behavioral vigilance in rats: task validation and effects of age, amphetamine, and benzodiazepine receptor ligands”. In: *Psychopharmacology* 117.3 (1995), pp. 340–357.

[MYZ13] Tomas Mikolov, Wen-tau Yih, and Geoffrey Zweig. “Linguistic Regularities in Continuous Space Word Representations.” In: *Hlt-naacl*. Vol. 13. 2013, pp. 746–751.

[Nak+04] Hiroyuki Nakahara et al. “Dopamine neurons can represent context-dependent prediction error”. In: *Neuron* 41.2 (2004), pp. 269–280.

[NEF87] Steen Nedergaard, Ingemar Engberg, and John A Flatman. “The modulation of excitatory amino acid responses by serotonin in the cat neocortexin vitro”. In: *Cellular and molecular neurobiology* 7.4 (1987), pp. 367–379.

[Nes+13] Bernhard Nessler et al. “Bayesian computation emerges in generic cortical microcircuits through spike-timing-dependent plasticity”. In: *PLoS Computational Biology* 9.4 (2013), e1003037.

[NSF07] Sheeja Navakkode, Sreedharan Sajikumar, and Julietta Uta Frey. “Synergistic requirements for the induction of dopaminergic D1/D5-receptor-mediated LTP in hippocampal slices of rat CA1 in vitro”. In: *Neuropharmacology* 52.7 (2007), pp. 1547–1554.

[Nyt] “NEW NAVY DEVICE LEARNS BY DOING; Psychologist Shows Embryo of Computer Designed to Read and Grow Wiser”. In: *New York Times* (1958), p. 25.

[O’L89] Dennis DM O’Leary. “Do cortical areas emerge from a protocortex?” In: *Trends in neurosciences* 12.10 (1989), pp. 400–406.

[OBS92] K. Obermayer, G. G. Blasdel, and K. Schulten. “Statistical-mechanical analysis of self-organization and pattern formation during the development of visual maps”. In: *Phys. Rev. A* 45 (10 1992), pp. 7568–7589.

[Oja82] Erkki Oja. “Simplified neuron model as a principal component analyzer”. In: *Journal of mathematical biology* 15.3 (1982), pp. 267–273.

[OJN14] Torben Ott, Simon Nikolas Jacob, and Andreas Nieder. “Dopamine receptors differentially enhance rule coding in primate prefrontal cortex neurons”. In: *Neuron* 84.6 (2014), pp. 1317–1328.

[Ola96] Mikel Olazaran. “A sociological study of the official history of the perceptrons controversy”. In: *Social Studies of Science* 26.3 (1996), pp. 611–659.

[OM79] Dean H Owen and Peter K Machamer. “Bias-free improvement in wine discrimination”. In: *Perception* 8.2 (1979), pp. 199–209.

[OSB92] Klaus Obermayer, K. Schulten, and Gary G. Blasdel. "A comparison between a neural network model for the formation of brain maps and experimental data". In: *Advances in neural information processing systems*. Citeseer. 1992.

[Ota+98] S Otani et al. "Dopamine facilitates long-term depression of glutamatergic transmission in rat prefrontal cortex". In: *Neuroscience* 85.3 (1998), pp. 669–676.

[OW08] Shawn R Olsen and Rachel I Wilson. "Lateral presynaptic inhibition mediates gain control in an olfactory circuit". In: *Nature* 452.7190 (2008), pp. 956–960.

[PA88] Rosalynn M Peron and Gary L Allen. "Attempts to train novices for beer flavor discrimination: a matter of taste". In: *The Journal of general psychology* 115.4 (1988), pp. 403–418.

[Par+07] Vinay Parikh et al. "Prefrontal acetylcholine release controls cue detection on multiple timescales". In: *Neuron* 56.1 (2007), pp. 141–154.

[Pas+00] F Passetti et al. "Increased acetylcholine release in the rat medial prefrontal cortex during performance of a visual attentional task". In: *European Journal of Neuroscience* 12.8 (2000), pp. 3051–3058.

[Pat+02] Roy D Patterson et al. "The processing of temporal pitch and melody information in auditory cortex". In: *Neuron* 36.4 (2002), pp. 767–776.

[Ped+11] F. Pedregosa et al. "Scikit-learn: Machine Learning in Python". In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830.

[PKF04] Daniel B Polley, Eugen Kvašňák, and Ron D Frostig. "Naturalistic experience transforms sensory maps in the adult cortex of caged animals". In: *Nature* 429.6987 (2004), pp. 67–71.

[Ple+09] Burkhard Pleger et al. "Influence of dopaminergically mediated reward on somatosensory decision-making". In: *PLoS Biology* 7.7 (2009), e1000164.

[PLT93] Alvaro Pascual-Leone and Fernando Torres. "Plasticity of the sensorimotor cortex representation of the reading finger in Braille readers". In: *Brain* 116.1 (1993), pp. 39–52.

[Poo+15] Jasper Poort et al. "Learning enhances sensory and multiple non-sensory representations in primary visual cortex". In: *Neuron* 86.6 (2015), pp. 1478–1490.

[Pou+09] Frédéric Pouille et al. "Input normalization by global feedforward inhibition expands cortical dynamic range". In: *Nature Neuroscience* 12.12 (2009), pp. 1577–1585.

[PS01] Frédéric Pouille and Massimo Scanziani. "Enforcement of temporal fidelity in pyramidal cells by somatic feed-forward inhibition". In: *Science* 293.5532 (2001), pp. 1159–1163.

[PS06] Stephen C Pratt and David JT Sumpter. "A tunable algorithm for collective decision-making". In: *Proceedings of the National Academy of Sciences* 103.43 (2006), pp. 15906–15910.

[PSM06] Daniel B Polley, Elizabeth E Steinberg, and Michael M Merzenich. "Perceptual learning directs auditory cortical map reorganization through top-down influences". In: *The Journal of neuroscience: the official journal of the Society for Neuroscience* 26.18 (2006), pp. 4970–4982.

[Qui+05] R Quian Quiroga et al. "Invariant visual representation by single neurons in the human brain". In: *Nature* 435.7045 (2005), pp. 1102–1107.

[Rag+04] Patrick Ragert et al. "Superior tactile performance and learning in professional pianists: evidence for meta-plasticity in musicians". In: *European Journal of Neuroscience* 19.2 (2004), pp. 473–478.

[Rai+06] Steven Raiguel et al. "Learning to see the difference specifically alters the most informative V4 neurons". In: *The Journal of neuroscience: the official journal of the Society for Neuroscience* 26.24 (2006), pp. 6589–6602.

[Rak88] Pasko Rakic. "Specification of cerebral cortical areas". In: *Science* 241.4862 (1988), p. 170.

[RB08] Lisa G Rapaport and Gillian R Brown. "Social influences on foraging behavior in young nonhuman primates: learning what, where, and how to eat". In: *Evolutionary Anthropology: Issues, News, and Reviews* 17.4 (2008), pp. 189–201.

[RBC88] JN Reynolds, A Baskys, and PL Carlen. "The effects of serotonin on N-methyl-D-aspartate and synaptically evoked depolarizations in rat neocortical neurons". In: *Brain research* 456.2 (1988), pp. 286–292.

[RD88] DD Rasmusson and RW Dykes. "Long-term enhancement of evoked potentials in cat somatosensory cortex produced by co-activation of the basal forebrain and cutaneous receptors". In: *Experimental Brain Research* 70.2 (1988), pp. 276–286.

[Rec+92] Gregg H Recanzone et al. "Topographic reorganization of the hand representation in cortical area 3b owl monkeys trained in a frequency-discrimination task". In: *Journal of neurophysiology* 67.5 (1992), pp. 1031–1056.

[Ree+11] Amanda Reed et al. "Cortical map plasticity improves learning but is not necessary for improved performance". In: *Neuron* 70.1 (2011), pp. 121–131.

[RHW85] David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. *Learning internal representations by error propagation*. Tech. rep. DTIC Document, 1985.

[RHW86] David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. "Learning representations by back-propagating errors". In: *Letters to Nature* 323.9 (1986), p. 533.

[Rin02] D. L. Ringach. "Spatial structure and symmetry of simple-cell receptive fields in macaque primary visual cortex". In: *Journal of Neurophysiology* 88.1 (2002), pp. 455–463.

[RMN09] Rajat Raina, Anand Madhavan, and Andrew Y Ng. "Large-scale deep unsupervised learning using graphics processors". In: *Proceedings of the 26th annual international conference on machine learning*. ACM. 2009, pp. 873–880.

[RMR08] Edmund T Rolls, Ciara McCabe, and Jerome Redoute. "Expected value, reward outcome, and temporal difference error representations in a probabilistic decision task". In: *Cerebral Cortex* 18.3 (2008), pp. 652–663.

[RO05] Pieter R Roelfsema and Arjen van Ooyen. "Attention-gated reinforcement learning of internal representations for classification". In: *Neural Computation* 17.10 (2005), pp. 2176–2214.

[Roe+90] Anna W Roe et al. "A map of visual space induced in primary auditory cortex". In: *Science* 250.4982 (1990), p. 818.

[Roe+92] Anna W Roe et al. "Visual projections routed to the auditory pathway in ferrets: receptive fields of visual neurons in primary auditory cortex". In: *Journal of Neuroscience* 12.9 (1992), pp. 3651–3664.

[Ros57] Frank Rosenblatt. *The perceptron, a perceiving and recognizing automaton Project Para*. Cornell Aeronautical Laboratory, 1957.

[Ros58] Frank Rosenblatt. "The perceptron: A probabilistic model for information storage and organization in the brain." In: *Psychological review* 65.6 (1958), p. 386.

[ROW10] Pieter R Roelfsema, Arjen van Ooyen, and Takeo Watanabe. "Perceptual learning rules based on reinforcers and attention". In: *Trends in cognitive sciences* 14.2 (2010), p. 64.

[RRB12] Jaldert Rombouts, Pieter Roelfsema, and Sander M Bohte. "Neurally Plausible reinforcement learning of working memory tasks". In: *Advances in Neural Information Processing Systems*. 2012, pp. 1871–1879.

[RS96] Edward S Ruthazer and Michael P Stryker. "The Role of Activity in the Development of Long-Range Horizontal Connections in Area 17 of the Ferret". In: *Journal of Neuroscience* 16.22 (1996), pp. 7253–7269.

[RSM93] GH al Recanzone, CE Schreiner, and Michael M Merzenich. "Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys". In: *The Journal of neuroscience: the official journal of the Society for Neuroscience* 13.1 (1993), pp. 87–103.

[RSRR92] Vilayanur S Ramachandran, M Stewart, and DC Rogers-Ramachandran. "Perceptual correlates of massive cortical reorganization." In: *Neuroreport* 3.7 (1992), pp. 583–586.

[RW+72] Robert A Rescorla, Allan R Wagner, et al. "A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement". In: *Classical conditioning II: Current research and theory* 2 (1972), pp. 64–99.

[Sac+98] Robert NS Sachdev et al. "Role of the basal forebrain cholinergic projection in somatosensory cortical plasticity". In: *Journal of Neurophysiology* 79.6 (1998), pp. 3216–3228.

[SAS00] Jitendra Sharma, Alessandra Angelucci, and Mriganka Sur. "Induction of visual orientation modules in auditory cortex". In: *Nature* 404.6780 (2000), pp. 841–847.

[Sat+03] Takemasa Satoh et al. "Correlated coding of motivation and outcome of decision by dopamine neurons". In: *The Journal of neuroscience: the official journal of the Society for Neuroscience* 23.30 (2003), pp. 9913–9923.

[SB98] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*. Vol. 135. MIT press Cambridge, 1998.

[Sch+01] Aniek Schoups et al. "Practising orientation identification improves orientation coding in V1 neurons". In: *Nature* 412.6846 (2001), pp. 549–553.

[Sch+12] Horst Schicknick et al. "Dopamine modulates memory consolidation of discrimination learning in the auditory cortex". In: *European Journal of Neuroscience* 35.5 (2012), pp. 763–774.

[Sch07a] Wolfram Schultz. "Behavioral dopamine signals". In: *Trends in neurosciences* 30.5 (2007), pp. 203–210.

[Sch07b] Wolfram Schultz. "Multiple dopamine functions at different time courses". In: *Annual Review of Neuroscience* 30 (2007), pp. 259–288.

[Sch10] Wolfram Schultz. "Review Dopamine signals for reward value and risk: basic and recent data". In: *Behav. Brain Funct* 6 (2010), p. 24.

[Sch15] Jürgen Schmidhuber. "Deep learning in neural networks: An overview". In: *Neural Networks* 61 (2015), pp. 85–117.

[SD07] Aaron R Seitz and Hubert R Dinse. "A common framework for perceptual learning". In: *Current Opinion in Neurobiology* 17.2 (2007), pp. 148–153.

[SDM97] Wolfram Schultz, Peter Dayan, and P Read Montague. "A neural substrate of prediction and reward". In: *Science* 275.5306 (1997), pp. 1593–1599.

[Sei+06] Aaron R Seitz et al. "Two cases requiring external reinforcement in perceptual learning". In: *Journal of Vision* 6.9 (2006), p. 9.

[Sem00] Kazue Semba. "Multiple output pathways of the basal forebrain: organization, chemical heterogeneity, and roles in vigilance". In: *Behavioural Brain Research* 115.2 (2000), pp. 117–141.

[SGK06] Martin Sarter, William J Gehring, and Rouba Kozak. "More attention must be paid: the neurobiology of attentional effort". In: *Brain research reviews* 51.2 (2006), pp. 145–160.

[SJL03] Mark Stopfer, Vivek Jayaraman, and Gilles Laurent. "Intensity versus identity coding in an olfactory system". In: *Neuron* 39.6 (2003), pp. 991–1004.

[SK83] Adam M Sillito and John A Kemp. "Cholinergic modulation of the functional organization of the cat visual cortex". In: *Brain research* 289.1 (1983), pp. 143–155.

[SL01] Mriganka Sur and Catherine A Leamey. “Development and plasticity of cortical areas and networks”. In: *Nature Reviews Neuroscience* 2.4 (2001), pp. 251–262.

[SL02] Natasha Sigala and Nikos K Logothetis. “Visual categorization shapes feature selectivity in the primate temporal cortex”. In: *Nature* 415.6869 (2002), pp. 318–320.

[SLW+77] Carla J Shatz, S Lindstro, Thorsten N Wiesel, et al. “The distribution of afferents representing the right and left eyes in the cat’s visual cortex”. In: *Brain research* 131.1 (1977), pp. 103–116.

[SM14] Michael R Smith and Tony Martinez. “Reducing the effects of detrimental instances”. In: *Machine Learning and Applications (ICMLA), 2014 13th International Conference on*. IEEE. 2014, pp. 183–188.

[SML90] Gary Sclar, John HR Maunsell, and Peter Lennie. “Coding of image contrast in central visual pathways of the macaque monkey”. In: *Vision Research* 30.1 (1990), pp. 1–10.

[SNW12] Yuka Sasaki, José E Náñez, and Takeo Watanabe. “Recent progress in perceptual learning research”. In: *Wiley Interdisciplinary Reviews: Cognitive Science* 3.3 (2012), pp. 293–299.

[SP14] Christoph E Schreiner and Daniel B Polley. “Auditory map plasticity: diversity in causes and consequences”. In: *Current Opinion in Neurobiology* 24 (2014), pp. 143–156.

[SPH09] Martin Sarter, Vinay Parikh, and W Matthew Howe. “Phasic acetylcholine release and the volume transmission hypothesis: time to move on”. In: *Nature Reviews Neuroscience* 10.5 (2009), pp. 383–390.

[SPN14] Michael Schmuker, Thomas Pfeil, and Martin Paul Nawrot. “A neuromorphic network for generic multivariate data classification”. In: *Proceedings of the National Academy of Sciences* 111 (2014), pp. 2081–2086.

[SR05] Mriganka Sur and John LR Rubenstein. “Patterning and plasticity of the cerebral cortex”. In: *Science* 310.5749 (2005), pp. 805–810.

[SS82] S MURRAY Sherman and PETER D Spear. “Organization of visual pathways in normal and visually deprived cats.” In: *Physiological reviews* 62.2 (1982), pp. 738–855.

[SSB99] Frank Sengpiel, Petra Stawinski, and Tobias Bonhoeffer. “Influence of experience on orientation maps in cat visual cortex”. In: *Nature Neuroscience* 2.8 (1999), pp. 727–732.

[Sug12] Nobuo Suga. “Tuning shifts of the auditory system by corticocortical and corticofugal projections and conditioning”. In: *Neuroscience and Biobehavioral Reviews* 36.2 (2012), pp. 969–988.

[SW03] Aaron R Seitz and Takeo Watanabe. “Psychophysics: Is subliminal learning really passive?” In: *Nature* 422.6927 (2003), pp. 36–36.

[SW05] Aaron Seitz and Takeo Watanabe. “A unified model for perceptual learning”. In: *Trends in cognitive sciences* 9.7 (2005), pp. 329–334.

[Swa03] Harvey A Swadlow. "Fast-spike interneurons and feedforward inhibition in awake sensory neocortex". In: *Cerebral Cortex* 13.1 (2003), pp. 25–32.

[SZW05] Xiu Sun, Yun Zhao, and Marina E Wolf. "Dopamine receptor stimulation modulates AMPA receptor synaptic insertion in prefrontal cortex neurons". In: *The Journal of neuroscience: the official journal of the Society for Neuroscience* 25.32 (2005), pp. 7342–7351.

[Teg+05] Martin Tegenthoff et al. "Improvement of tactile discrimination performance and enlargement of cortical somatosensory maps after 5 Hz rTMS". In: *PLoS Biology* 3.11 (2005), e362.

[TFS05] Philippe N Tobler, Christopher D Fiorillo, and Wolfram Schultz. "Adaptive coding of reward value by dopamine neurons". In: *Science* 307.5715 (2005), pp. 1642–1645.

[Til+06] J Tillein et al. "Sensitivity of primary auditory cortex to binaural cues in congenitally deaf cats". In: *Assoc. Res. Otolaryngol* 29 (2006), p. 47.

[TSW08] Yoshiaki Tsushima, Aaron R Seitz, and Takeo Watanabe. "Task-irrelevant learning occurs only when the irrelevant feature is weak". In: *Current biology: CB* 18.12 (2008), R516.

[TT96] Phillip D Tomporowski and Veronica F Tinsley. "Effects of memory demand and motivation on sustained attention in young and older adults". In: *The American journal of psychology* (1996), pp. 187–204.

[UMB04] Mark A Ungless, Peter J Magill, and J Paul Bolam. "Uniform inhibition of dopamine neurons in the ventral tegmental area by aversive stimuli". In: *Science* 303.5666 (2004), pp. 2040–2042.

[UVC04] Markus Ullsperger and D Yves Von Cramon. "Neuroimaging of performance monitoring: error detection and beyond". In: *Cortex* 40.4 (2004), pp. 593–604.

[VCL11] Timothy J Vickery, Marvin M Chun, and Daeyeol Lee. "Ubiquity and specificity of reinforcement signals throughout the human brain". In: *Neuron* 72.1 (2011), pp. 166–177.

[VEM83] David C Van Essen and John HR Maunsell. "Hierarchical organization and functional streams in the visual cortex". In: *Trends in neurosciences* 6 (1983), pp. 370–375.

[VS+07] Etienne de Villers-Sidani et al. "Critical period window for spectral tuning defined in the primary auditory cortex (A1) in the rat". In: *The Journal of neuroscience: the official journal of the Society for Neuroscience* 27.1 (2007), pp. 180–189.

[VSL80] RICHARD C Van Sluyters and FRANCES B Levitt. "Experimental strabismus in the kitten". In: *Journal of Neurophysiology* 43.3 (1980), pp. 686–699.

[Wan+03] Yue Wang et al. "fMRI evidence for cortical modification during learning of Mandarin lexical tone". In: *Journal of Cognitive Neuroscience* 15.7 (2003), pp. 1019–1027.

[Wat+02] Takeo Watanabe et al. “Greater plasticity in lower-level than higher-level visual motion processing in a passive perceptual learning task”. In: *Nature Neuroscience* 5.10 (2002), pp. 1003–1009.

[Wat+05] Sarah C Watkinson et al. “New approaches to investigating the function of mycelial networks”. In: *Mycologist* 19.1 (2005), pp. 11–17.

[WB98] Norman M Weinberger and Jonathan S Bakin. “Learning-induced physiological memory in adult primary auditory cortex: receptive field plasticity, model, and mechanisms”. In: *Audiology and Neurotology* 3.2-3 (1998), pp. 145–167.

[WDS01] Pascale Waelti, Anthony Dickinson, and Wolfram Schultz. “Dopamine responses comply with basic assumptions of formal learning theory”. In: *Nature* 412.6842 (2001), pp. 43–48.

[Web+91a] Harry H Webster et al. “Basal forebrain lesions with or without reserpine injection inhibit cortical reorganization in rat hindpaw primary somatosensory cortex following sciatic nerve section”. In: *Somatosensory & motor research* 8.4 (1991), pp. 327–346.

[Web+91b] HH Webster et al. “Long-term enhancement of evoked potentials in raccoon somatosensory cortex following co-activation of the nucleus basalis of Meynert complex and cutaneous receptors”. In: *Brain research* 545.1 (1991), pp. 292–296.

[Wei03] Norman M Weinberger. “The nucleus basalis and memory codes: auditory cortical plasticity and the induction of specific, associative behavioral memory”. In: *Neurobiology of Learning and Memory* 80.3 (2003), pp. 268–284.

[Wel+92] E Welker et al. “Plasticity in the barrel cortex of the adult mouse: effects of chronic stimulation upon deoxyglucose uptake in the behaving animal”. In: *Journal of Neuroscience* 12.1 (1992), pp. 153–170.

[Wer74] Paul Werbos. “Beyond regression: New tools for prediction and analysis in the behavioral sciences”. PhD thesis. Harvard University, 1974.

[WFS04] Donald A Wilson, Max L Fletcher, and Regina M Sullivan. “Acetylcholine and olfactory perceptual learning”. In: *Learning and Memory* 11.1 (2004), pp. 28–34.

[WGC97] Yongchang Wang, Q Gu, and Max S Cynader. “Blockade of serotonin-2C receptors by mesulergine reduces ocular dominance plasticity in kitten visual cortex”. In: *Experimental brain research* 114.2 (1997), pp. 321–328.

[WH+60] Bernard Widrow, Marcian E Hoff, et al. “Adaptive switching circuits”. In: *IRE WESCON convention record*. Vol. 4. 1. New York. 1960, pp. 96–104.

[WH74] Torsten N Wiesel and David H Hubel. “Ordered arrangement of orientation columns in monkeys lacking visual experience”. In: *Journal of comparative neurology* 158.3 (1974), pp. 307–318.

[Wil92] Ronald J Williams. “Simple statistical gradient-following algorithms for connectionist reinforcement learning”. In: *Machine learning* 8.3-4 (1992), pp. 229–256.

[Win+95] J Winkler et al. "Essential role of neocortical acetylcholine in spatial memory". In: *Letters to Nature* 375 (1995), pp. 484–487.

[Wis04] Roy A Wise. "Dopamine, learning and motivation". In: *Nature Reviews. Neuroscience* 5.6 (2004), pp. 483–494.

[WNS01] Takeo Watanabe, José E Náñez, and Yuka Sasaki. "Perceptual learning without perception". In: *Nature* 413.6858 (2001), pp. 844–848.

[WS81] Roy A Wise and Howard V Schwartz. "Pimozide attenuates acquisition of lever-pressing for food in rats". In: *Pharmacology Biochemistry and Behavior* 15.4 (1981), pp. 655–656.

[WSG+78] Roy A Wise, Joan Spindler, GJ Gerberg, et al. "Neuroleptic-induced" anhedonia" in rats: pimozide blocks reward quality of food". In: *Science* 201.4352 (1978), pp. 262–264.

[WW80] Barry D Waterhouse and Donald J Woodward. "Interaction of norepinephrine with cerebrocortical activity evoked by stimulation of somatosensory afferent pathways in the rat". In: *Experimental neurology* 67.1 (1980), pp. 11–34.

[WZ05] Michael Wehr and Anthony M Zador. "Synaptic mechanisms of forward suppression in rat auditory cortex". In: *Neuron* 47.3 (2005), pp. 437–445.

[YM04] Tianming Yang and John HR Maunsell. "The effect of perceptual learning on neuronal responses in monkey visual area V4". In: *The Journal of neuroscience: the official journal of the Society for Neuroscience* 24.7 (2004), pp. 1617–1626.

[Zhu+97] Ciyou Zhu et al. "Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization". In: *ACM Transactions on Mathematical Software (TOMS)* 23.4 (1997), pp. 550–560.

[Zin+06] W Zinke et al. "Cholinergic modulation of response properties and orientation tuning of neurons in primary visual cortex of anaesthetized Marmoset monkeys". In: *European Journal of Neuroscience* 24.1 (2006), pp. 314–328.

[ZW98] XO Zhu and PM Waite. "Cholinergic depletion reduces plasticity of barrel field cortex." In: *Cerebral Cortex* 8.1 (1998), pp. 63–72.

[Záb92] László Záborszky. "Synaptic organization of basal forebrain cholinergic projection neurons". In: *Neurotransmitter interactions and cognitive function*. Springer, 1992, pp. 27–65.