Homeostatic plasticity - computational and clinical implications

Dissertation
zur Erlangung des Grades
Doktor der Naturwissenschaften

vorgelegt beim Fachbereich Informatik
der Goethe Universität
Frankfurt am Main

von

Cristina Savin
aus Sighisoara (Rumänien)

Frankfurt am Main 2010
(FB12)

Vom Fachbereich Informatik der Goethe Universität als Dissertation angenommen.

Dekan: Prof. Dr.-Ing. Detlef Krömker
1. Gutachter: Prof. Dr. Jochen Triesch
2. Gutachter: Prof. Dr. Christoph von der Malsburg

Zusammenfassung

Das menschliche Gehirn ist in der Lage sich an dramatische Veränderungen der Umgebung anzupassen. Wenn man eines Morgens aufwachen würde und die Welt um einen herum stünde auf dem Kopf, wäre man zunächst verwirrt und nicht fähig, sich in dieser neuen Umgebung zu bewegen. Nach ein paar Stunden allerdings würde das Gehirn gelernt haben, mit diesen neuen Sinneseindrücken umzugehen, und man wäre in der Lage, sich wieder annähernd normal zu verhalten. Mit etwas mehr Übung wäre es sogar möglich, wieder komplizierte Bewegungsabläufe wie Fahrrad oder Ski Fahren zu meistern.

Ein solches Beispiel mag sich anhören wie aus einem Science-Fiction-Roman, es stammt jedoch aus den persönlichen Beschreibungen eines Teilnehmers eines klassischen Experiments zum Wahrnehmungslernen. Freiwillige wurden gebeten, für einige Wochen ununterbrochen Prismen-Brillen zu tragen, welche die Bilder, die das Auge erreichen, in verschiedener Weise transformieren. Zum Beispiel werden die Bilder um 180 rotiert, so dass die Welt auf dem Kopf zu stehen scheint. Genau wie oben beschrieben passten sich die Teilnehmer nach einer Weile an die geänderten Verhältnisse an und konnten fortan wieder ihren alltäglichen Beschäftigungen nachgehen. Gegen Ende des Experiments, wenn die Brillen wieder abgenommen wurden, erschien die Welt wieder 'falsch'. Dieses Gefühl hielt jedoch nur für kurze Zeit an, bis sich alles wieder vollständig normalisiert hatte.

Hinter der Anpassungsfähigkeit des Gehirns stecken verschiedenste Lernmechanismen. Wenn man sich ansieht, wie sich die Nervenzellen im Gehirn untereinander verständigen, entdeckt man, dass fast jeder Aspekt der Kommunikation form- und veränderbar ist; der Transfer eines Signals zwischen zwei Neuronen über Synapsen, die Art und Weise wie einkommende Information im Zellkörper verarbeitet wird, oder der Transfer eines solchen Resultats in eine Ausgabe, überall wirken regulierende Prozesse.

Einige dieser Mechanismen sind bereits relativ gut erforscht, während bei anderen noch kaum bekannt ist, welche Rolle sie innerhalb der Informationsverarbeitungsprozesse im Gehirn spielen. Oft ist es schwierig die Funktion von einzelnen Prozessen experimentell zu untersuchen, da sie in einem Zellverband mit verschiedensten anderen Faktoren in einem komplexen Zusammenspiel wirken. In Computer-Simulationen kann man diese Komplexität dagegen bewusst einschränken oder kontrollieren, allerdings ist dieser Ansatz erst wenig erforscht worden.

Die vorliegende Arbeit beschäftigt sich mit der Untersuchung des Zusammenwirkens verschiedener Lernmechanismen und deren Beitrag zur Fähigkeit Neuronaler Netze komplexe Berechnungen ausführen zu können. Insbesondere soll gezeigt werden, dass das Zusammenspiel von 'Spike-Timing dependent Plasticity' (STDP) mit zwei weiteren Prozessen, 'Synaptic Scaling' und 'Intrinsic Plasticity' (IP), es ermöglicht, dass Nervenzellen einkommende Information effizient kodieren können. Die gleichen Mechanismen führen dazu, dass ein Netzwerk aus Neuronen in der Lage ist, ein 'Arbeitsgedächtnis' für vergangene Stimuli zu entwickeln. Abschliessend werden klinische Konsequenzen einiger Lern-Prozesse untersucht. Es wird gezeigt, dass Mechanismen, welche normalerweise für die Stabilität eines neuronalen Systems sorgen, können zu epileptischen Anfällen führen, unter speziellen Voraussetzungen.

Neuronale Lern-Mechanismen

Als erstes soll hier ein kurzer überblick über die wichtigsten, in dieser Arbeit verwendeten Lernmechanismen gegeben werden. Der Informations-Transfer zwischen Nervenzellen basiert auf mehreren Prozessen, sowohl in der präsynaptischen als auch in der postsynaptischen Zelle, welche wiederum durch diverse aktivitätsabhängige Lernmechanismen reguliert werden können. Von einigen, zusammengefasst unter dem Begriff Hebb'sches Lernen, glaubt man zu wissen, dass sie eine wichtige Rolle für Lernen und Gedächtnis spielen. Einigen der weniger gut verstandenen spricht man häufig stabilisierende Funktion zu. Sie reagieren welche auf globale Veränderungen der einkommenden Signale, um so das Netzwerk in einem für die Informationsverarbeitung geeigneten Regime zu halten.

Unter den Hebb'schen Mechanismen ist STDP besonders interressant, da es auf dem zeitlichen Verhältnis zischen ein- und ausgehenden Signalen basiert. Wenn das präsynaptische kurz vor dem postsynaptischen Neuron aktiv ist, verstärkt sich die Synapse, und erhöht damit die Wahrscheinlichkeit, dass das präsynaptische Neuron das postsynaptische Neuron zum feuern bringt.

Die homöostatischen Prozesse, welche im folgenden untersucht werden, sind 'Synaptic Scaling' und 'Intrinsic Plasticity' (IP). Ersteres kontrolliert die Gesamtstärke aller eingehenden Synapsen eines Neurons, IP dagegen verändert die Transferfunktion der Zelle basierend auf deren früheren Aktivitäts-Profil.

ICA mit spikenden Neuronen

Wie das Gehirn lernt, die eingehenden Informationen von den Sinnesorganen zu verarbeiten und zu repräsentieren, zählt sicher zu den wichtigsten Fragestellungen im Bereich der Neurowissenschaften. Viele computerbasierte Theorien weisen darauf hin, dass Nervenzellen die Redundanzen zwischen ihren Antwortprofilen für einen bestimmten Stimulus so weit wie möglich verringern sollten, um die Menge der kodierbaren Informationen zu maximieren. Ein spezieller Algorithmus, Independent Component Analysis (ICA), has sich zum Standard für das Lernen effizienter Kodierungen entwickelt. ICA erklärt u.a. einige Aspekte sensoricher Repräsentationen im Gehirn, wie etwa das Anwortverhalten von Neuronen in V1. Bisher blieb es allerdings ein Rätsel, wie ein realistisches Netzwerk aus spikenden Neuronen einen solchen Algorithmus konkret implementieren könnte und wie das Lernen mit Hilfe bekannter Plastizitätsmechanismen umgesetzt werden kann.

Um eine Antwort zu finden, haben wir ein Modell eines Netzwerk aus spikenden Neuronen untersucht, in welchem Anpassungen durch das Zusammenspiel von STDP, Synaptic Scaling und IP vorgenommen werden. Wir konnten zeigen, dass ein solches Netzwerk in der Lage ist, viele Standart-ICA-Probleme zu lösen. Wichtig dabei ist zu erwähnen, dass dieses Modell im Gegensatz zu allen bisherigen Vorschlägen für ICA in spikenden Neuronen den Lernprozess mit Hilfe von biologisch plausiblen Mechanismen erklären kann.

Eine nützliche Eigenschaft unseres Modells ist die Stabilität der gelernten effizienten Repräsentationen über einen weiten Bereich der Parameter. Damit zeigt es eine vor

allem für biologische Systeme wichtige Robustheit. Ferner, die hier verwendeten Methoden können mit einer Reihe klassischer mathematischer Formulierungen von ICA in Verbindung gebracht werden.

Um ICA-ähnliche Berechnungen in spikenden Neuronen beobachten zu können, ist es wichtig zu entscheiden, wie genau ein externer Stimulus in neuronalen Input transformiert wird. Da Nervenzellen ihre Informationen in einer Sequenz von Spikes kodieren, wäre eine Möglichkeit, einen Stimulus durch die Anzahl an Spikes innerhalb eines Zeitfensters zu repräsentieren ('Rate Coding'). Diese Möglichkeit wird von den meisten existiernenden Modellen gewählt. Die Alternative ist ein Code der auf den genauen Zeitpunkten von Spikes verschiedener Zellen basiert. Unser Modell hebt sich von anderen dadurch ab, dass es für beide Arten von Input-Codierungen effiziente neuronale Repräsentationen lernt.

Enstehung eines Arbeitsgedächtnisses durch Reward-abhängiges STDP

Auf die Frage wie Nervenzellen Informationen repräsentieren und kommunizieren gibt es bisher keine allgemein akzeptierte Antwort. Wahrscheinlich handelt es sich dabei aber auch um die falsche Frage, da jüngst experimentelle Hinweise gefunden werden konnten, die darauf hindeuten, dass mehrere Codes zur selben Zeit unterschiedliche Informationen übermitteln. Aus mit den Schnurhaaren verbundenen Nervenzellen in Ratten kann man zum Beispiel zur selben Zeit sowohl die Entfernung eines Objects aus der Spike-Rate, als auch die vertikale Position desselben aus den Spike-Zeiten dekodieren.

Bisher ist allerdings ungeklärt, wie diese verschiedenen Repräsentationen entstehen können. Eine interessante Hypothese geht davon aus, dass diese Codes das Resultat von Lernprozessen sind. Um diese Hypothese genauer zu untersuchen, haben wir ein weiteres Netzwerk spikender Neurone konstruiert, welches Informationen für eine kurze Zeit zu speichern hat. Wie im vorherigen Modell verwenden wir die selben Lern-Mechanismen. Diesmal soll das Netzwerk aber nicht nur passiv die Input-Statistiken lernen, sondern aktiv eine konkrete Aufgabe lösen. Aus diesem Grund wird das synaptische Lernen nun zusätzlich durch die erhaltenen Belohnung moduliert, und dadurch auf problemrelevante Input-Eigenschaften beschränkt. Diese Art von Belohnungsabhängigem Lernen ist experimentell gut fundiert, jedoch bisher in Modellen nicht benutzt worden, um interessante Aufgaben zu lösen.

In der hier verwendeten Aufgabe bekommt das Netzwerk einen zufällig aus einer begrenzten Menge ausgewählten Stimulus als Eingabe. Diese aktiviert eine Subpopulation von Nervenzellen für einen Zeitschritt. Nach einer kurzen Pause erscheint ein Signal, welches für den nächsten Zeitschritt das Ausführen der Aktion ankündigt. Abhängig davon, ob die Aktion zum aktuellen Input passt, wird ein Belohnungssignal gegeben. Diese Sequenz wird vielfach wiederholt, wobei die Pause eine konstante oder variable Länge haben kann. Diese Aufgabe ähnelt denjenigen, die häufig in Experimenten über das Arbeitsgedächtnis in Tieren verwendet werden.

Durch die Kombination von belohnungsabhängigem STDP und homöostatischen Mech-

anismen lernt das Netzwerk, die Stimulus-Repräsentationen für mehrere Zeitschritte verfügbar zu halten. Obwohl in unserem Modell-Design keinerlei Informationen über die bevorzugte Art der Kodierung enthalten sind, finden wir nach Ende des Trainings neuronale Repräsentationen, die denjenigen aus vielen Arbeitsgedächtnis-Experimenten gleichen. Unser Modell zeigt, dass solche Repräsentationen durch Lernen enstehen können und dass Reward-abhängige Prozesse eine zentrale Kraft bei der Entwicklung des Arbeitsgedächtnisses spielen können.

Die Tatsache, dass sich die Art, wie der gleiche Stimulus im Netzwerk kodiert wird, zwischen den Aufgaben mit konstanter bzw. variabler Pause unterscheidet, öffnet Raum für zukünftige Untersuchungen zur Enstehung von unterschiedlichen Kodierungen durch Lernen in Interaktion mit der Umgebung.

Neuroglia-vermitteltes Synaptic Scaling und die Entstehung von Epilepsie

In den vorangehenden Kapiteln haben wir uns mit den positiven Aspekten der Interaktion zwischen verschiedenen Lernmechanismen befasst. Dieser letzte Teil beschäftigt sich nun mit einigen möglichen klinischen Auswirkungen von homöostatischen Lern-Prozessen. Wir konnten zeigen, dass der selbe Mechanismus, der normalerweise die Aktivität im Gehirn in Balance hält, in speziellen Situationen auch zu Destabilisierung führen und epileptische Anfälle auslösen kann.

Es ist seit langer Zeit bekannt, dass Entzündungen des Gehirns, wie zum Beispiel bei Meningitis, das Risiko von epileptischen Anfällen signifikant erhöhen. Entzündungshemmende Medikamente zeigen oft zusätzlich anti-epileptische Wirkung. Aufgrund dieser Entdeckungen wurde die Hypothese aufgestellt, dass Entzündungen in neuronalen Netzwerken das Auftreten von Anfällen begünstigen können.

Als Erklärung für diesen Einfluss wurden verschiedene Mechanismen vorgeschlagen. In der vorliegenden Arbeit soll eine Form von ‘Synaptic Scaling’ untersucht werden, welche als erstes im Hippocampus entdeckt wurde. Unser Modell verwendet eine Form von synaptischer Regulierung, welche die Stärke von exzitatorischen Synapsen anpasst, um den Gesamt-Input eines Neurons konstant zu halten. Das entscheidende an dieser Regulierung ist jedoch, dass sie nicht durch die Nervenzelle selbst geregelt wird, sondern über Astrocyten, einer speziellen Form von Helferzellen im Gehirn. Die Astrocyten überwachen die Aktivitätsstärke eines Neurons und steuern Veränderungen in dessen Synapsen über die Freisetzung des Proteins TNF-α. Interessanterweise wird das gleiche Protein vom Immunsystem zur Signal-übertragung verwendet, was zu der hier untersuchten Annahme führt, dass diese Doppelfunktion die Quelle von entzündungsabhängigen Anfällen sein könnte.

Das hier vorgestellte Modell liefert eine neuartige Erklärung zur Entstehung von epileptischen Anfällen bei chronischen Entzündungen: Ein erhöhter Level des Immun-Botenstoffes führt zu erhöhten Synapsen-Stärken, welche wiederum epileptische Anfälle innerhalb des Netzwerkes begünstigen. Eine naheliegende Frage ist, warum beide, Immunsystem und Synaptische Stabilisierung, auf das gleiche Protein vertrauen. Grundsätzlich sind Immun-

und Nervensystem gut voneinander getrennt und die wenigen Interaktionen sind normalerweise zeitlich und räumlich begrenzt. Die negativen Effekte von TNF-α enstehen nur bei chronischen Entzündungen, welche sich weit ausbreiten und nicht nach kurzer Zeit wieder abheilen.

Eine weitere Besonderheit des hier untersuchten 'Synaptic Scaling' Mechanismus ist, dass regulatorische Prozesse nicht lokal auf eine Nervenzelle begrenzt sind. Ein einzelner Astrocyt reagiert auf Signale von vielen benachbarten Neuronen und durch Diffusion erreichen seine Botenstoffe auch weiter entfernte Zellen. Unser Modell weist darauf hin, dass das System durch diese Art der flächendeckenden Kommunikation anfälliger für lokale Störungen wird. Wenn zum Beispiel eine Störung in einem Bereich des Gehirns auftritt, welche den Input eine Reihe weiterer Neurone liefert, werden diese Neuronen seltener erregt und die entsprechenden Astrocyten produzieren TNF-α . Nervenzellen in benachbarten Arealen erhalten ebenfalls diffundiertes TNF-α und erhöhen daraufhin die Stärke ihrer Synapsen. Ein normaler Input führt in diesem Areal nun zu deutlich erhöhter Aktivität - der Start eines epileptischen Anfalls. Dieses Verhalten liefert eine Erklärung für den Zusammenhang von Läsionen im Gehirn und epileptischen Anfällen, welche ihren Ursprung häufig nahe der beschädigten Zellen haben.

Diskussion

Die Tatsache, dass einzelne Lernprozesse leicht durch spezifische Veränderungen der Netzwerkstruktur oder -aktivität beeinflusst werden können, könnte ein Grund für die Vielfalt der experimentell gefundenen, häufig in ihrer Funktion überlappenden, Mechanismen sein. Unabhängig von der spezifischen Aufgabenstellung soll diese Arbeit die Rolle der homöostatischen Lern-Mechanismen in neuronalen Simulationen hervorheben. Sowohl die Entstehung von Rezeptiven Feldern, als auch die Stabilisierung der Dynamiken eines rekurrenten Neuronalen Netzwerks während Belohnungsabhängigem Lernens stützen sich auf das Zusammenspiel von STDP und verschiedenen homöostatischen Mechanismen.

Die vorliegende Arbeit könnte auf unterschiedliche weise fortgesetzt und ausgeweitet werden. Zum Beispiel könnte erforscht werden, in wie weit die hier diskutierten Mechanismen für ICA sich für weitere Formen der Input-Codierung generalisieren lassen könnten. Zudem wäre es interessant zu sehen, wie sich das Modell verhält, wenn man es innerhalb von hierarchischen neuronalen Netzen einsetzt. Unser Modell des Arbeitsgedächtnisses würde wahrscheinlicherweise von zusatzlichen Lernmechanismen, welche z.B. die Inhibition des Netzwerkes steuern, profitieren können. Wir stehen erst am Anfang der Untersuchungen über die Rolle von regulierbarer Inhibition für Lernprozesse, allerdings konnten wir bereits zeigen, dass IP in inhibitorischen Neuronen die Repräsentationen verschiedener Stimuli verbessert, indem es sicherstellt, dass Rechen-Ressourcen gleichmässig über alle Inputs verteilt werden.

Contents

1 Introduction **1**

2 Mechanisms for cortical plasticity **4**
2.1 Hebbian forms of synaptic plasticity . . . 4
2.1.1 Plasticity at excitatory synapses . . . 4
2.1.2 Plasticity of inhibition . . . 8
2.2 Homeostasis . . . 9
2.2.1 Synaptic scaling . . . 9
2.2.2 Intrinsic plasticity . . . 10
2.3 Short-term plasticity . . . 14
2.4 Structural plasticity . . . 15
2.5 Plasticity and neuromodulation . . . 16
2.6 Interactions between different forms of plasticity . . . 17
2.7 Modeling plasticity . . . 18

3 ICA learning in spiking neurons **21**
3.1 Background . . . 22
3.1.1 Independent component analysis . . . 22
3.1.2 Models of receptive field development . . . 25
3.2 Computational model . . . 26
3.2.1 Neuron model . . . 27
3.2.2 Intrinsic plasticity . . . 28
3.2.3 Spike-timing dependent plasticity . . . 29
3.2.4 Synaptic scaling . . . 30
3.3 Results . . . 31
3.3.1 Linear demixing of two supergaussian sources . . . 31
3.3.2 One neuron learns an independent component . . . 31
3.3.3 ICA in a neuron population . . . 36
3.3.4 Is IP necessary for learning? . . . 36
3.3.5 Interaction between IP and STDP . . . 38
3.4 Discussion . . . 40

4 Working memory development by reward-dependent STDP **42**
4.1 Background . . . 43
4.1.1 Working memory development in children . . . 43
4.1.2 Working memory: experimental evidence . . . 44
4.1.3 Models of working memory . . . 45
4.2 Computational model . . . 47
4.2.1 Network model . . . 47
4.2.2 Reward-dependent learning . . . 48
4.2.3 Homeostatic and structural plasticity . . . 49
4.2.4 The task . . . 50
4.3 Results . . . 50
4.3.1 Network learns to solve the task . . . 50
4.3.2 Task-specific performance improvement . . . 51
4.3.3 Neurons develop stimulus specificity, in location and time . . . 51
4.3.4 Delayed categorization . . . 53
4.3.5 Network optimization: the role of structural plasticity during learning . . . 54

4.4 Discussion . . . 54

5 Glia-mediated synaptic scaling and epileptogenesis **57**
5.1 Background . . . 57
5.1.1 TNF-α mediated synaptic scaling . . . 57
5.1.2 Glia cells, more than brain glue . . . 59
5.1.3 Expression of immune proteins in the CNS . . . 62
5.1.4 Inflammation and epilepsy . . . 64
5.2 Computational model . . . 66
5.2.1 Neural network . . . 67
5.2.2 Glia Model . . . 68
5.3 Results . . . 69
5.3.1 Increases in TNF-α levels can induce seizures . . . 69
5.3.2 A sustained reduction in input can trigger seizures . . . 71
5.3.3 Local lesions cause seizures . . . 71
5.4 Discussion . . . 75

6 Conclusions and future work **77**
6.1 The role of homeostasis in learning . . . 77
6.2 Stability through diversity . . . 78
6.3 Future work . . . 78

A ICA with spiking neurons **80**
A.1 Mathematical derivation: IP rule . . . 80
A.2 Model robustness . . . 81
A.3 Learning with a two-dimensional input . . . 81
A.4 A link between IP and BCM . . . 84
A.5 RF estimation for the spike-based demixing . . . 87

B Working memory development **88**

C Synaptic scaling and epileptogenesis **90**
C.1 Network parameters . . . 90
C.2 Model calibration . . . 93
C.3 Stability analysis . . . 94

Glossary **119**

Acronyms **124**

Acknowledgements **125**

1 Introduction

Plasticity supports the remarkable adaptability and robustness of cortical processing. It allows the brain to learn and remember patterns in the sensory world, to refine motor control, to predict and obtain reward, or to recover function after injury.

Even basic sensory perception can be influenced by experience. After altering of the statistics of sensory inputs (e.g. by selective deprivation or over-presentation of certain stimuli), a variety of plasticity mechanisms act to compensate for these changes, dynamically relocating cortical processing to interesting areas of the input space (i.e. the spared or over-represented inputs). For example, V1 ocular dominance plasticity was shown to occur in response to monocular deprivation [Hubel & Wiesel 1998, Fox & Wong 2005, Hofer *et al.* 2006]. Similarly, tactile experience influences somatotopic maps, in rodent S1 [Feldman 2009], and exposing young rats to repeated auditory stimuli leads to an enhanced representation of the presented frequencies, altering auditory tuning curves and tonotopic maps [Keuroghlian & Knudsen 2007].

Interestingly, it is not only the mere presence or absence of inputs but also their temporal dynamics that can influence plasticity, through a mechanims termed spike timing dependent plasticity (STDP). In rat adult V1, temporally correlated inputs can lead to systematic shifts in visual tuning, related to stimulus order and timing [Dan & Poo 2006]. In humans, similar experimental procedures can be used to shift the perceived location of visual stimuli [Fu *et al.* 2002], or the orientation tuning and alter perception of orientation [Yao & Dan 2001]. Moreover, appropriately timed stimuli have also been shown to influence auditory [Dahmen *et al.* 2008] and motor [Wolters *et al.* 2003] cortex.

If at a young age map plasticity can be easily induced in response to passive stimulation, later into adulthood plasticity is harder to induce and slower. However, it remains effective for attended, task relevant stimuli, e.g. when explicitly paired with positive or negative reinforcement or neuromodulation [Feldman 2009]. For example, classic conditioning [Weinberger 2007, Kossut & Siuciska 1996, Siucinska & Kossut 2004] or perceptual learning [Dan & Poo 2006] can improve the representation of relevant stimuli in primary sensory areas. Depending on the task, the effects of learning can vary from an increase in response for trained stimuli, to the sharpening of some tuning curves or to their shift towards or away from the conditioned input, see [Feldman 2009].

Independent of brain area or the specific manipulation, we can identify several key principles shaping neural computation. First, as shown by experiments on sensory maps plasticity, neurons need to adapt to the statistics of their incoming signals in order to process information efficiently. Hence, the statistics of the natural environment greatly influence sensory processing [Simoncelli & Olshausen 2001]. Second, neurons must operate under certain biological constraints (e.g. metabolic or wiring limitations), which critically influence the way neurons can encode information [Laughlin 1981, Stemmler & Koch 1999, Laughlin 2001]. Lastly, the ultimate goal of the organism is to survive. Neuronal processing is shaped by the tasks the organism has to solve (reward-dependent learning) [Doya 2002]. Our work here investigates the implementation of these principles in neural circuits.

Within cortical networks, the three aforementioned principles translate into several plasticity mecha-

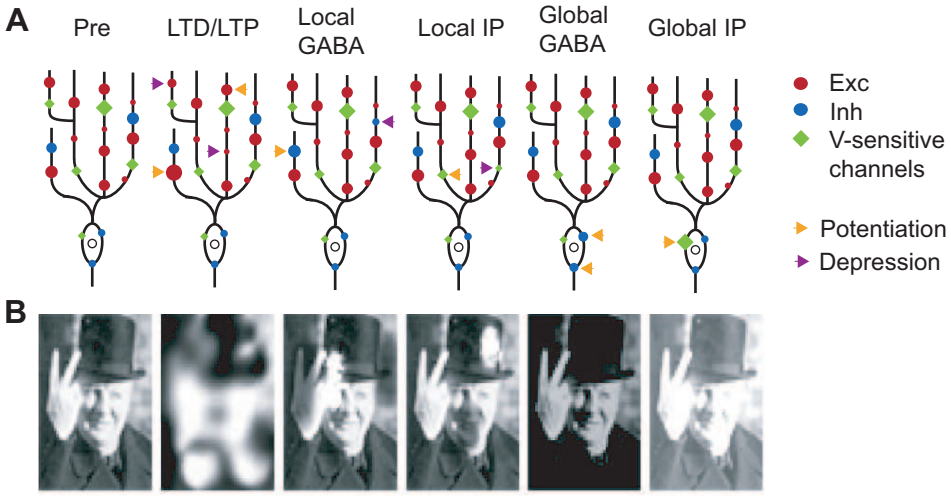

Figure 1.1. Example illustrating how memory storage is affected by different forms of plasticity (A) At the biological level, starting with a certain assignment of incoming weights to a pyramidal neuron, LTP/LTD alone changes the efficacy of different synapses as function of the incoming inputs and the neuronal output. Local inhibition affects the throughput of neighboring excitatory synapses, while global inhibition reduces the overall activity of the neuron. Similarly, local IP and global IP can increase the excitability of single dendrites or of neurons, causing an overall increase of the response to the same stimuli. (B) As a visualization of the computational implications of different mechanisms, consider a neuron whose synaptic weights correspond to the first image in the lower panel (in fact, what is depicted is the synaptic throughput, i.e. the probability of a synapse to elicit a spike in the postsynaptic neuron, white marking high values). LTP and LTD alone at a random subset of synapses can significantly change the receptive field of the neuron. Local GABA plasticity and local IP also change synaptic throughput, but their effects are restricted to a small subset of neighboring synapses. Lastly, global changes in inhibition and global IP change the activation of the neuron, while maintaining the relative weights of synapses. Adapted from [Kim & Linden 2007].

nisms. By far the best known form of cortical plasticity is Hebbian learning, which involves associative activity-dependent changes in the impact of a presynaptic spike on the postsynaptic firing probability. However, although the long-term potentiation and depression of excitatory synapses (LTP/LTD) have been at the heart of research into the cellular mechanism of learning and memory, these forms of plasticity are not the only ones relevant for cortical processing. Even the paper reporting LTP for the first time notes that the Hebbian changes in synaptic strength were accompanied by a change in the probability of a postsynaptic spike for the same input, a phenomenon later termed E-S potentiation [Bliss & Lomo 1973]. Since then, experiments have revealed a large array of cortical plasticity mechanisms affecting the synaptic efficacy of excitatory and inhibitory synapses and the intrinsic excitability of dendritic branches or neurons. Some of these processes, e.g. synaptic scaling or intrinsic plasticity (IP), have a homeostatic nature, bringing the cortical activity towards a certain fixed point, thus implementing some of the metabolic constraints of the system [Zhang & Linden 2003, Turrigiano 2008].

Ultimately, cortical plasticity arises through the complex interplay between these various forms of excitatory, inhibitory and intrinsic plasticity (see Fig. 1.1). Initial phases involve rapid activity-based potentiation or depression, stabilized by slower homeostatic processes and later consolidated by slower structural remodeling of the circuit [Feldman 2009]. Some of these plasticity mechanisms may be directly

involved in information storage, while others may be permissive, facilitating learning without taking part in the encoding, or meta-plastic, supporting higher order aspects of memory.

Unfortunately, little is known about the precise computational roles of different mechanisms, mostly due to the fact that they are difficult to investigate experimentally. The most important limitation is the fact that different plasticity mechanisms often share receptors and expression pathways, making it virtually impossible to constrain pharmacological manipulations to a single form of plasticity. This is why computational models hold great potential for advancing our knowledge on how different plasticity mechanisms contribute to information processing, and ultimately to behavior. They provide the best way to dissociate the effects of each form of plasticity and to investigate the effects of their interactions. Unfortunately, computational modeling has, so far, largely been restricted to Hebbian forms of plasticity, occasionally combined with synaptic scaling, with a few notable exceptions [Foldiák 1990, Triesch 2007], which consider also IP. This thesis extends this work in several directions, as it investigates the role of the interaction between Hebbian and homeostatic plasticity in the context of both unsupervised and reinforcement learning .

Thesis overview

Before describing these results, we first give an overview of the experimental evidence for various cortical plasticity mechanisms and briefly discuss some of the phenomenological models used for implementing them. In the following chapters, we present three computational models investigating different computational and clinical aspects of cortical plasticity, with a focus on homeostatic mechanisms.

First, we describe a model for receptive field development with spiking neurons, which shows how independent component analysis (ICA) can be performed in a network of stochastically spiking neurons through the combination of several synaptic and intrinsic plasticity mechanisms. Importantly, although alternative implementations have been proposed in the literature for solving the ICA problem in spiking neurons [Parra *et al.* 2009], ours is the first to offer a mechanistic explanation of how ICA-like computation could arise by biologically plausible learning mechanisms.

Second, in the context of reward-dependent learning, we show that similar mechanisms are needed for shaping the connectivity of recurrent neural networks to best perform a delayed response task. Although we make no *a priori* assumptions on how the stimulus should be encoded by the network, our learning procedure leads to representations similar to those observed in various working memory experiments. Hence, our model demonstrates that working memory dynamics may emerge naturally in a recurrent network with reward-modulated STDP, suggesting that reward-dependent learning may be a central driving force for the development of working memory.

Regardless of the specific problem, our work emphasizes the role of homeostatic plasticity in neural computation. Both the development of receptive fields in sensory neurons and stabilizing the dynamics of recurrent neural networks during reward-dependent learning critically depend on the interaction between Hebbian learning (STDP) and several homeostatic mechanisms.

The last part of our work focuses on another aspect of homeostatic plasticity, in relation with epileptic seizures. Paradoxically, there are situations in which the very mechanisms that normally are required to preserve the balance of the system, may act as a destabilizing factor. Specifically, we focus on a form of synaptic scaling which involves the interaction between neurons and glia cell, support cells in the brain, through a diffusible protein (TNF-α). Interestingly, this messenger is also a signaling molecule of the immune system, which may give rise to interactions between immune responses and neural activity regulation. Our model offers a novel explanation for the occurrence of seizures during chronic inflammation: rising levels of the immune messenger lead to increased synaptic connection strengths, which makes neural circuits prone to develop seizures.

We conclude by a discussion on the different roles of homeostatic plasticity in cortical processing and on interesting directions of investigation which were not yet covered in our work.

2

Mechanisms for cortical plasticity

Information transfer across a synapse is a complex process, involving the release of presynaptic neurotransmitter, followed by the transduction of synaptic signals by postsynaptic receptor, and their integration across different dendritic branches and then at the soma via voltage-gated ionic channels, finally resulting in a sequence of action potentials (see Fig. 2.1). Each of these phases is regulated by a variety of activity-dependent plasticity mechanisms, operating at multiple time scales, varying in spatial resolution (from the level of single synapses, to dendritic branches, and neurons) and direction of change. Some of them, such as Hebbian learning, are self-amplifying (i.e. the correlated activation of pre- and post-synaptic partners leads to weight changes that increase the probability of subsequent coactivation, thus implementing a positive feedback loop). Other serve a homeostatic purpose, bringing the neuronal activity back to the desired set point after changes in input (negative feedback loop).

At one end of the spectrum, we have classical Hebbian forms of plasticity, which act on a time scale of seconds to minutes, in a positive feedback loop, and are specific to single synapses. At the other end, different forms of homeostatic plasticity provide a global regulation meant to maintain the system in a regime that is favorable for information processing. Typically, the timescales of such negative feedback are slower than those for Hebbian plasticity, from hours to days. In between the two, both positive and negative feedback processes have been reported at different levels, from single synapses, to dendritic branches and neurons. Additionally, on faster time-scales, the dynamics of synapses can be subject to facilitation and depression, while on very slow time scales synaptic plasticity is supported by structural mechanisms, that cause the rewiring of neural circuits by synapse formation, elimination, or morphological changes.

This chapter gives an overview of the experimental evidence available on various cortical plasticity mechanisms and briefly discusses some of the phenomenological descriptions used for implementing them in computational models. For the specific plasticity mechanisms implemented in our work, detailed model descriptions are provided separately, in the Methods sections of the following chapters.

2.1 Hebbian forms of synaptic plasticity

2.1.1 Plasticity at excitatory synapses

Long term potentiation and depression

Long term potentiation and depression (LTP/LTD) are two forms of synaptic plasticity which are believed to play an important role in experience-dependent plasticity, memory and learning. Classically, LTP is induced by high-frequency tetanic stimulation (100-500 Hz), while low frequency stimulation (0.5-5 Hz) induces LTD. Both were first reported in the hippocampus, LTP at the perforant pathway [Bliss & Lomo 1973] and LTD at the Schaffer collaterals [Lynch *et al.* 1977]. Additionally, both involve an 'early' phase, lasting up to 60 minutes, followed by a 'late' phase, which requires protein synthesis [Malenka &

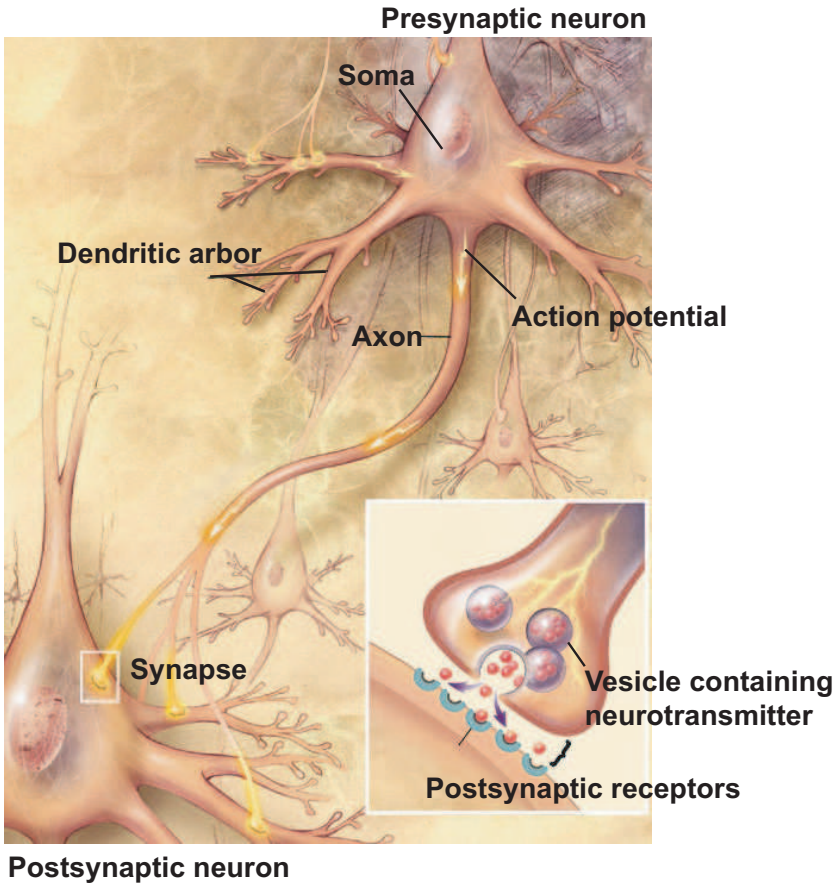

Figure 2.1. Information transmission in the brain. Neurons emit action potentials (spikes) which are transmitted to other neurons via synapses. When reaching the synapse, a presynaptic spike causes the release of a neurotransmitter (e.g. glutamate or GABA) into the synaptic cleft. This activates the corresponding receptors in the postsynaptic neurons, causing the opening of specific channels in the cell membrane, further leading to a change in membrane potential. Signals for different synapses are integrated along dendritic branches and the soma and cause action potentials in the neuron. Adapted from [ADEAR 2008].

Bear 2004].

Depending on brain region, LTP and LTD can vary in terms of induction mechanisms (NMDA or mGlu receptor dependent) and effects (pre- or post-synaptic) [Malenka & Bear 2004,Derkach *et al.* 2007]. In particular, NMDAR-dependent LTP and LTD at hippocampal synapses have been extensively studied, becoming a prototypical model for this form of activity-dependent synaptic regulation. Hippocampal LTP requires the synaptic activation of NMDARs during postsynaptic depolarization, which results in a Ca^{2+} influx at the postsynaptic site that activates a complex molecular cascade, involving, among other, the calcium/calmodulin-dependent protein kinase II (CaMKII), causing an increase of postsynaptic α-amino-3-hydroxyl-5-methyl-4-isoxazole-propionate (AMPA) receptors, see [Malenka & Bear 2004]. Similarly, hippocampal LTD requires activation of NMDARs and an increase in the concentration of postsynaptic Ca^{2+} , with the difference that it involves different NMDARs subunits, and may require Ca^{2+} release from

intracellular stores. In terms of mechanisms, LTD is correlated with dephosphorylation of postsynaptic PKA and PKC substrates and does not involve CaMKII [Malenka & Bear 2004]. Depending on the precise experimental paradigm, this NMDAR-dependent LTD can be accompanied by mGluR-dependent LTD, which is characteristic for parallel fibers in Purkinje cells [Kano *et al.* 2008]. Additionally, other forms of LTD can involve cannabinoid receptors, see [Malenka & Bear 2004, Feldman 2009].

Spike-timing dependent plasticity

Although it had been long recognized that LTP/LTD induction can exhibit a temporal dependence [Levy & Steward 1983], it was more than a decade later that experiments using precisely timed pre- and postsynaptic spikes at millisecond temporal resolution were performed [Markram *et al.* 1995, Markram *et al.* 1997b, Bi & Poo 1998a, Debanne *et al.* 1998, Zhang *et al.* 1998], demonstrating what became known as spike-timing dependent plasticity (STDP) [Abbott & Nelson 2000]. During such temporal-dependent learning, the arrival of a presynaptic spike a few milliseconds before a postsynaptic action potentials leads to LTP, whereas the reverse order causes LTD, see Fig. 2.2. In contrast to the forms of Hebbian learning discussed before, STDP emphasizes the importance of causality in determining the direction of synaptic modifications, in line with the original postulate of Hebb [Hebb 1949].

If STDP windows can vary across synapse types and brain regions [Abbott & Nelson 2000], the temporal asymmetry of classic STDP is preserved in many species (locust, frog, electric fish, zebra finch, rat, cat, and probably also humans) and brain areas (e.g. hippocampus, prefrontal, entorhinal, somatosensory and visual cortices), see review in [Caporale & Dan 2008], suggesting that STDP could be a general principle underlying experience dependent plasticity in the cortex.

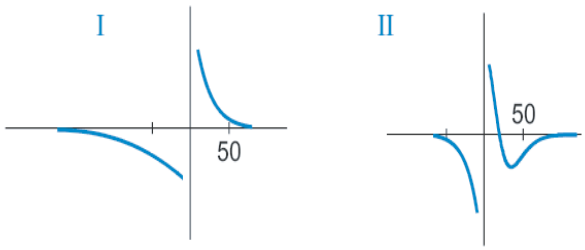

Figure 2.2. STDP windows at excitatory synapses. The temporal axis is in ms. Adapted from [Caporale & Dan 2008].

Mechanisms

As standard LTP/LTD, STDP modifications rely on NMDAR activation, which causes a rise in postsynaptic Ca^{2+} levels. Potentiation occurs if a presynaptic input is followed shortly by a backpropagating action potential (BAP), which releases the NMDAR Mg^{2+} block, causing a Ca^{2+} influx. Additionally, the incoming EPSP can temporally affect the conductance of the dendrite (depending on region, by either A-type K^+ or by voltage-gated Na^+ channels), amplifying the BAP signal, and thus controlling the magnitude of the weight potentiation. The specific time window for potentiation is determined by several factors, including BAP effects on the Mg^{2+} kinetics and the nonlinear interaction between BAPs and EPSPs [Caporale & Dan 2008].

Similarly, in the case of time-dependent synaptic depression a BAP triggers an afterdepolarization lasting for tens of milliseconds. An EPSP arriving at this time causes a much smaller Ca^{2+} influx, possibly

due to the fact the the BAP activates voltage-dependent calcium channels (VDCCs) and thus inactivates NMDARs, resulting in LTD. In an alternative interpretation, time-dependent depression depends on mGluRs and Ca^{2+} influx through VDCCs, a mechanism which does not require NMDARs activation [Karmarkar & Buonomano 2002].

Signal propagation through dendritic branches depends on both passive cable properties and on the presence of various active channels within dendrites. The fact that STDP relies on the backpropagation of the action potential through the dendritic arbor suggests that the window for LTP/LTD should vary as function of dendritic location. Indeed, a series of recent experiments have demonstrated several location specific effects. Overall, the learning rule will depend on dendritic morphology and the expression of different active channels, which can itself be modulated by activity, see below. Functionally, such spatial-dependent STDP could lead to different input selection in proximal and apical dendrites. Direct induction of potentiation in apical dendrites is much more difficult such that significant changes can only be achieved by triggering dendritic Ca^{2+} spikes. As this requires the synchronous activation of several synapses, plasticity at distal synapses rewards synaptic cooperation, for a more detailed discussion see [Caporale & Dan 2008].

STDP and learning

If most of the early STDP experiments were done in slice or neuron culture, recent studies have started to shed more light on the role of STDP in vivo. Sensory stimulation paired with electrical stimulation of the corresponding primary sensory area induces weight changes matching those predicted by STDP (visual stimulation in Xenopus [Mu & Poo 2006], cat [Schuett *et al.* 2001], or rat [Meliza & Dan 2006]; somatosensory stimulation in rat [Jacob *et al.* 2007]; olfactory stimulation in locust [Cassenaer & Laurent 2007]). Moreover, STDP has been also demonstrated in humans. Specifically, it was shown that paring an electric stimulation of a somatosensory area with the activation of motor cortex by Transcranial magnetic stimulation (TMS) induces changes in TMS-evoked motor potentials, which depend on the relative timing of the pairs on the scale of tens of milliseconds, as for in vitro STDP [Wolters *et al.* 2003].

In a more physiologically relevant setting, STDP was investigated using pure sensory stimulation. For example, the orientation selectivity of individual V1 neurons could be modified by the presentation of a pair of gratings, with the direction of the change depending on the temporal order between the two [Yao & Dan 2001, Yao *et al.* 2004]. Similar results could also be obtained for the spatial domain [Fu *et al.* 2002]. Moreover, the time scales observed in these experiments were again similar to those reported for STDP *in vitro*. Lastly, the equivalent psychophysical experiments in humans revealed perceptual changes consistent with the receptive field shifts observed in electrophysiology [Yao & Dan 2001, Fu *et al.* 2002, Yao *et al.* 2004]. Similar results have been obtained for motion stimuli and STDP was used to explain several known visual illusions involving motion, see [Caporale & Dan 2008].

Lastly, STDP can play a role in more general forms of experience-dependent plasticity, which do not explicitly involve timing. For example, it was shown that STDP can explain the reorganization of the sensory map that follows whisker trimming in awake behaving rats [Celikel *et al.* 2004].

Beyond the classic STDP model

As spike patterns used in experimental paradigms are very restricted, they provide limited information on the effects of STDP for naturalistic inputs. The typical STDP experiment uses pairs of pre- post- spikes, separated by long periods of inactivity. More complex experiments have been recently developed using different spike patterns (triplets or quadruplets). They reveal nonlinear interactions between different pairs, suggesting that the the overall outcome of STDP depends on activity patterns over several minutes [Caporale & Dan 2008]. As a special case, burst-dependent plasticity (BDP) has been reported in several systems. Depending on the system, BDP can behave in a manner reminiscent of classic correlation base learning, or can exhibit a dependence on burst timing. The cellular mechanisms behind BDP are not well

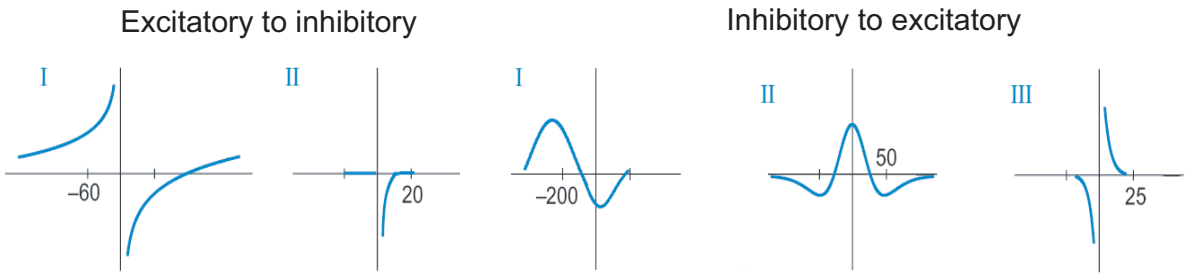

Figure 2.3. Diversity of STDP learning windows at inhibitory synapses. The temporal axis is expressed in ms. Adapted from [Caporale & Dan 2008].

understood. They may be related to STDP, but at least in some cases BDP seems to activate different mechanisms altogether [Caporale & Dan 2008].

Several models have been proposed to account for the nonlinear interactions between spike pairs, e.g. considering nearest neighbor interactions, with LTP winning over LTD [Sjöstrm *et al.* 2001], the short-term depression of presynaptic inputs and the frequency dependent attenuation of postsynaptic spikes [Froemke & Dan 2002], multiple timescales of integration [Pfister & Gerstner 2006], or voltage-dependent effects [Clopath *et al.* 2010]. A general framework for modeling spike-timing dependent plasticity follows in section 2.7.

2.1.2 Plasticity of inhibition

If the plasticity of ionotropic excitatory synapses has been extensively studied both experimentally and in theoretical work, comparatively little is known about the regulation of inhibitory synapses and circuits. There are several difficulties which have slowed down progress in this area: inhibitory neurons are much more diverse than pyramidal neurons; additionally, they are typically small in size and more difficult to access electrophysiologically. Despite these technical difficulties, it has been established that inhibitory circuits are also highly plastic: both inhibitory synapses and excitatory synapses to inhibitory neurons are capable of activity dependent long-term plasticity [Kim & Linden 2007, Feldman 2009].

As an example, monocular deprivation increases the strength of synapses from fast spiking interneurons onto excitatory synapses in primary visual cortex, potentially accompanied by the potentiation of excitatory synapses to the same interneurons. This adaptation has been suggested as a possible explanation for the depression of closed-eye response observed in these cases [Feldman 2009]. Furthermore, in adult primary auditory cortex, pairing auditory stimuli with cholinergic stimulation increases the response to the paired stimulus in two steps: a rapidly induced disinhibition to the relevant stimulus, followed by a gradual increase in tone-evoked excitation. Functionally, this fast change in inhibition may implement a 'tag', marking the sites which are to exhibit subsequent potentiation at the excitatory synapses, see [Feldman 2009].

As for synapses between pyramidal neurons, STDP has been reported for excitatory synapses onto inhibitory neurons and inhibitory synapses to pyramidal neurons. However, the windows for plasticity of GABAergic synapses are more complex than those at glutamatergic synapses (see Fig. 2.3). Despite the differences the induction mechanisms seem to be very similar in this case [Caporale & Dan 2008]. Additionally, inhibitory plasticity can include a homeostatic component. As example, sensory deprivation reduces, while classic conditioning and whisker overstimulation increases inhibition in primary sensory areas [Feldman 2009]. The balance between excitation and inhibition is readjusted by changing the strength of synapses and the number of functional inhibitory synapses. In general, the mechanisms

behind the plasticity of inhibition are more diverse and much less well understood than those for excitatory synapses, see [McBain & Kauer 2009].

Computationally, the plasticity of inhibition is critical for several aspects of cortical function. The stabilization of recurrent excitation by feedforward and feedback inhibition is a tightly regulated process [Zheng & Knudsen 1999, Wehr & Zador 2003, Gabernet *et al.* 2005, Wilent & Contreras 2005]. Disrupting the balance between the two can lead to runaway excitation, affecting sensory responses and altering experience dependent plasticity, see [Feldman 2009].

2.2 Homeostasis

Mounting experimental evidence suggests that many properties of the central nervous system can be regulated by homeostatic mechanisms. Homeostatic plasticity acts as a negative feedback loop compensating for chronic changes in neuronal activity. Several such processes operate within cortical circuits, on a wide range of temporal and spatial scales. The exact feature of activity that is regulated is unclear. It could be the average firing rate, average Ca^{2+} concentration or some more complex statistical measure of the activity, e.g. [Stemmler & Koch 1999]. However, in order to maintain this set-point, various aspects of activity are regulated, including the strength of excitatory and inhibitory connections and the intrinsic excitability of individual neurons. In the following, we focus on synaptic scaling and homeostatic intrinsic plasticity, while the regulation of the balance between excitation and inhibition was briefly discussed in section 2.1.2.

Computationally, it has long been suggested that the destabilizing influence of positive-feedback processes, such as Hebbian plasticity, needs to be compensated by homeostatic mechanisms [Abbott & Nelson 2000]. Interestingly, the Hebbian and homeostatic mechanisms often share the same molecular substrate, but have opposing effects on synaptic properties or neuronal activity [Turrigiano & Nelson 2000]. Importantly, the two operate on different time scales, as homeostatic processes must be slow enough to not interfere with information transmission, but fast enough to keep up with the changes caused by synaptic plasticity.

2.2.1 Synaptic scaling

Synaptic scaling was first demonstrated in neuron cultures, when it was shown that blocking inhibition triggers a compensatory reduction in the mEPSC amplitude of glutamatergic synapses, which brings the population firing rate back to the original value [Turrigiano *et al.* 1998]. Since then, many studies have shown that pharmacological manipulations of neuronal activity (with bidirectional effects) lead to similar homeostatic plasticity [Turrigiano 2007]. Moreover, it is now clear that this type of synaptic scaling is not restricted to culture and slices, but also occurs in vivo, e.g. following sensory deprivation [Kaneko *et al.* 2008].

As noted before, it is not clear what aspect of activity is controlled by synaptic scaling. As postsynaptic hyperpolarization can already trigger synaptic rescaling in some cases and given that synaptic scaling does not seem to require NMDA activation and can occur even when AMPA and NMDA receptors are blocked [Turrigiano & Nelson 2000], it is probably some function of postsynaptic activity [Turrigiano & Nelson 2004], which could be estimated by various Ca^{2+} receptors [Turrigiano 2008].

Depending on brain region, age and other factors, synaptic scaling can be implemented by either pre- or postsynaptic changes [Turrigiano 2007]. In many cases synaptic scaling changes the number of postsynaptic AMPA receptors, see [Turrigiano 2008]. In other situations, chronic activity deprivation can lead to an increase in the release probability and number of functional release sites, typically also accompanied by an increase in mEPSC amplitude (due to a presynaptic increase in quantal amplitude), see [Turrigiano 2008]. In that case, retrograde signaling could explain the presynaptic changes observed in some experiments. Alternatively, they could be owed to a slow growth of postsynaptic synapse, which

triggers a coordinated growth of the corresponding presynaptic terminal [Turrigiano & Nelson 2000]. Importantly, if postsynaptic changes affect neuron responsiveness, but not the dynamics of synaptic transmission, presynaptic scaling affects the probability of release, but also short-term synaptic plasticity (see below), thus changing the information transfer across the synapse.

Although several neurotransmitters have been implicated, such as brain-derived neurotrophic factor (BDNF), tumour necrosis factor-α (TNF-α) or Arc, a unified view of the molecular pathways involved in synaptic scaling is still missing. For example, Arc was shown to be necessary and sufficient for inducing synaptic scaling in cultured neurons, as activity regulates Arc levels bidirectionally, and homeostatic plasticity can be blocked by overexpressing or knocking down Arc. However, some caveats remain and Arc-mediated synaptic scaling cannot explain some experimental observations, see [Turrigiano 2007]. The data on TNF-α mediated synaptic scaling will be reviewed in detail in Chapter 3. To complicate things further, different modulators are related (TNF-α and BDNF share signaling pathways, while BDNF can also induce Arc expression). All in all, it remains to be established if any of these signalling molecules are part of the core pathway or just modulators of synaptic scaling, for an in-depth discussion on mechanisms see [Turrigiano 2008].

The spatial resolution of homeostatic synaptic scaling is generally considered to be that of a neuron, such that all incoming synapses change proportionally in response to changes in total drive or postsynaptic firing [Ibata *et al.* 2008]. However, some recent experiments seem to suggest that synaptic scaling can have, in some cases, either a finer (almost synapse-specific) or a coarser (e.g. in TNF-α -mediated scaling where a glial source suggests a network-level change in activity) resolution. The prevailing experimental paradigm for studying synaptic scaling has been a chronic exposure (by bath-applied blockers) of the entire neuron to a change in global activity, which makes it difficult to differentiate between global and local effects. Still, several experiments report non-uniform changes in synaptic weights, inconsistent with a global signal [Rabinowitch & Segev 2008]. For example, a weak depolarization of the postsynaptic neuron by increasing extracellular K^+ causes a subset of synapses to be deactivated, while the others remain unchanged, see [Rabinowitch & Segev 2008]. Moreover, some scaling requires local protein synthesis and can still be observed even if the dendrite is physically detached from the soma [Kang & Schuman 1996]. Lastly, when experimentally suppressing a fraction of the presynaptic partners of a neuron, localized scaling is observed after 2 days [Hou *et al.* 2008]. However, as this type of local regulation has only been demonstrated for cultured hippocampal pyramidal neurons, it remains unclear to which extent these effects are generalizable to other brain regions.

Functionally, synaptic scaling is important for the activity-dependent refinement of cortical circuits, particularly during development, when the number and strength of synapses change markedly as result of experience [Turrigiano & Nelson 2004]. Interestingly, as for Hebbian plasticity, synaptic scaling is developmentally regulated in sensory areas. Furthermore, the sites of homeostatic cortical plasticity migrate to different cortical layers in an age dependent manner, following Hebbian plasticity [Turrigiano & Nelson 2004].

2.2.2 Intrinsic plasticity

Another way to stabilize postsynaptic firing rates in response to changes in input is to change neuronal excitability, i.e. the way changes in membrane potential translate into a particular pattern of firing. This type of intrinsic plasticity (IP) homeostatically regulates a rich array of voltage-dependent sodium, potassium and calcium channels, and, through them, the integrative properties of the neuron [Desai *et al.* 1999, Daoudal & Debanne 2003, Zhang & Linden 2003].

Interestingly, the first report of LTP already mentions additional forms non-synaptic plasticity. Namely, the increase in population activity in that experiment was greater than expected from a LTP-induced change in EPSP alone [Bliss & Lomo 1973]. Moreover, this increased activity could sometimes be observed even in absence of a measurable change in synaptic efficacy, a phenomenon later termed EPSP-spike, or E-S potentiation [Zhang & Linden 2003, Kim & Linden 2007]. Since then, multiple forms of IP have been

reported in the literature which vary in the time-scale of adaptation or the type of ionic channels involved depending on species, brain region, or induction protocol, see Table 2.2.2 and [Zhang & Linden 2003].

Table 2.1. Experimental evidence for IP

Species	Area	Manipulation	Phenotype	References
Invertebrates				
Aplysia		Sensitization	+	[Cleary *et al.* 1998] [Brembs *et al.* 2002]
		Operant conditioning	+	[Cleary *et al.* 1998] [Brembs *et al.* 2002]
Xenopus	Optic tectum	Visual stimulation	+	[Aizenman *et al.* 2003]
Vertebrates				
Rabbit	Hippocampus	HFS	+	[Bliss & Lomo 1973]
Guinea pig	Hippocampus	HFS	+ NMDAR-d	[Andersen *et al.* 1980] [Hess & Gustafsson 1990]
Rat	Hippocampus	HFS	+ NMDAR-d pre	[McNaughton *et al.* 1994]
	Hippocampus	pairing	+ NMDAR-d	[Jester *et al.* 1995]
	Hippocampus	paring	+ NMDAR-d pre	[Ganguly *et al.* 2000]
	Hippocampus	HFS/LFS	± NMDAR-d	[Daoudal *et al.* 2002]
	Hippocampus	STDP	+/- NMDAR-d	[Wang *et al.* 2003]
	Hippocampus	ACPD	+ mGluR-d	[Cohen *et al.* 1999]
	Hippocampus	DHPG	+ mGluR-d	[Ireland & Abraham 2002]
	Hippocampus	KA	+	[Melyan *et al.* 2002]
Mouse	Hippocampus	Postsyn. Ca^{2+} depol.	+ GluR-indep.	[Tsubokawa *et al.* 2000]
	EC	Postsyn. depol.	+	[Egorov *et al.* 2002]
	EC	Postsyn. hyperpol.	-	[Egorov *et al.* 2002]
Rat	Cerebellum	HFS	+ NMDAR-d	[Aizenman & Linden 2000] [Armano *et al.* 2000]
	Cerebellum	Postsyn. Ca^{2+} depol.	+ GluR-indep.	[Aizenman & Linden 2000]
	Cortex	10Hz ACPD/CHPG	+ mGluR-d	[Sourdet *et al.* 2003]
	Cortex	TTX (48h)	+ homeostatic	[Desai *et al.* 1999]

Abreviations: HFS/LFS – high-/low- frequency stimulation, NMDAR-d/mGluR-d – NMDA/mGlu receptor-dependent plasticity, GluR-indep – glutamate receptor independent IP, KA – kainic acid, pre – presynaptic, ACPD – aminocyclopentanedicarboxylic acid (mGluR agonist) , DHPG – dihydroxyphenylglycine (mGluR agonist), CHPG – 2-chloro-5-hydroxyphenylglycine, TTX – tetrodotoxin. Adapted from [Daoudal & Debanne 2003].

Depending on the triggering activity patterns, IP may act at a various spatial scales, from changes affecting small clusters of neighboring synapses (see discussion above on spatial extent of synaptic scaling), to larger dendritic segments, which can act as individual units of plasticity [Husser & Mel 2003], and whole neurons [Zhang & Linden 2003]. For example, IP was reported to occur in a local subdomain of the apical dendrite of a CA1 pyramidal neuron [Frick *et al.* 2004], see Fig. 2.4. This form of localized IP could play a role in the propagation of EPSCs along dendrites, see [Marder *et al.* 1996] for a more detailed discussion.

Additionally, intrinsic plasticity mechanisms do not always serve a homeostatic function. It has been shown that in the cerebellum tetanic inputs produce a rapid and long-lasting increase in neuronal excitability, possibly implementing some form of heterosynaptic plasticity, such that strong activity from one input increases the sensitivity of the neuron to all of its inputs [Aizenman & Linden 2000].

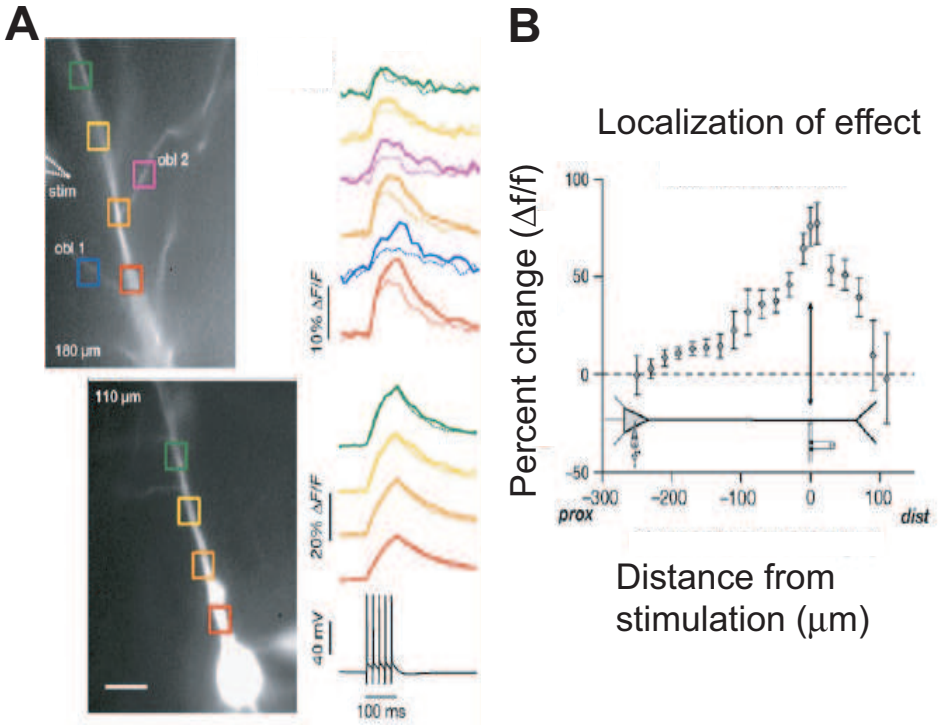

Figure 2.4. IP in a local subdomain of the apical dendrite of a CA1 pyramidal neuron. (A) A fluorescent dye is used for imaging the main apical (up) and an oblique (down) dendrite of a CA1 neuron. Ca^{2+} transients evoked by somatic current injection are measured before (dotted lines) and after (solid lines) LTP induction in several regions of interest, marked by different colors. (B) Changes in peak Ca^{2+} transient, normalized by distance to the site of synaptic stimulation (stim). Adapted from [Frick *et al.* 2004].

Molecular substrate

Mechanistically, IP and synaptic plasticity can often share common induction pathways, see Table 2.2.2. It is believed that changes in intrinsic excitability critically depend on postsynaptic depolarization, probably measured through Ca^{2+} elevation [Daoudal & Debanne 2003]. However, if LTP/LTD are induced by brief Ca^{2+} transients via NMDARs, IP may rely on a slower integration of Ca influx, possibly via L-type Ca^{2+} channels. In any case, this rise in Ca^{2+} activates several protein kinases and phosphatases, such as CaMKII, protein kinase C (PKC), protein kinase A (PKA), which also play a role in synaptic plasticity [Ganguly *et al.* 2000, Tsubokawa *et al.* 2000, Wang *et al.* 2003]. All in all, the signaling pathways leading to changes in the concentration of different ionic channels remain less well understood compared to synaptic plasticity.

The role of nonsynaptic plasticity in learning and memory

Persistent changes in intrinsic excitability can be induced by training in behaving animals [Zhang & Linden 2003, Mozzachiodi & Byrne 2009]. For a variety of organisms and tasks learning leads to increases in excitability, due to a combination of effects: the reduction of the spike threshold, spike accommodation and burst-evoked afterhyperpolarization (possibly involving several K^+ channels), see Table 2.2.2 and

Table 2.2. Potential molecular signals and substrates for IP

Trigger	Effectors	Enzymes	Functional targets
VDCCs	Ca^{2+}	PKC	$I_{K,A}$,I_K,$I_{K,AHP}$,$I_{K(Ca)}$
NMDARs	G-proteins	CaMKII	Na^+ channels
mGluRs		Adenylyl cyclase	Ca^{2+} channels, R- and T-type
mAChRs		Guanylyl cyclase	I_h,I_{CAN}
5HTRs		nNOS	
		PKA	
		PKG	
		MAPK	
		Protein phosphatases	

Abreviations: 5HTR – serotonin receptor, mAChR – muscarinic acetylcholine receptor, MAPK – mitogen-activated protein kinase, nNOS – neuronal nitric oxid synthase, PKA/PKC/PKG – protein kinase A/C/G, VDCCs – voltage-dependent calcium channels. Adapted from [Zhang & Linden 2003].

[Zhang & Linden 2003]. These changes are often shown to be intrinsic to the recorded neurons and can be accompanied by synaptic changes in the same neurons or in related circuits —IP can also occur in absence of synaptic changes, see review by [Mozzachiodi & Byrne 2009]— and, most importantly, they correlate to some measure of learning performance, pointing to a possible causal link between IP and behavior.

Table 2.3. Changes in neuronal excitability induced by learning

Species	Area	Manipulation	Phenotype	References
Invertebrates				
Hermissenda	Type B photoreceptor		+	[Alkon 1984] [Alkon *et al.* 1985]
Helix		Associative learning	+	[Gainutdinov *et al.* 1998] [Gainutdinov *et al.* 2000]
Aplysia		Sensitization	+	[Cleary *et al.* 1998]
Aplysia		Associative learning	+	[Antonov *et al.* 2001] [Antonov *et al.* 2003]
Hirudo		Habituation	-	[Sahley 1995]
		Sensitization	+	[Burrell *et al.* 2001]
Vertebrates				
Cat	Pericruciate cortex	Associative conditioning	+	[Brons & Woody 1980]
Rabbit	Hippocampus	Trace eyelid conditioning	+	[Disterhoft *et al.* 1986]
			+	[Coulter *et al.* 1989]
			+	[de Jonge *et al.* 1990]
	Cerebellum	Delay eyelid conditioning	+	[Schreurs *et al.* 1997] [Schreurs *et al.* 1998]
Rat	Pyriform cortex	Operant conditioning	+	[Saar *et al.* 1998] [Saar *et al.* 2001] [Saar *et al.* 2002]

Based on review by [Zhang & Linden 2003].

The experimental evidence suggests several distinct roles of IP in learning. On the one hand, in cat associative conditioning the changes in excitability persist even after extinction and might explain the faster acquisition during re-training [Brons & Woody 1980]. On the other hand, in rabbit eyetrace conditioning IP affects a large portion of hippocampal neurons, with increased excitability lasting only a few days, although the memory is maintained for much longer and is believed to facilitate synaptic learning by making the network more excitable [Disterhoft *et al.* 1986, Coulter *et al.* 1989, de Jonge *et al.* 1990]. Similarly, operant conditioning induces temporary changes in neuronal excitability which are believed to underly rule learning, see [Zhang & Linden 2003].

Computational implications

So far little is known about the computational implications of IP. Beyond maintaining system homeostasis, IP could fine-tune the output properties of the neuron to the statistics of its inputs [Stemmler & Koch 1999], regulate synaptic plasticity [Triesch 2007] —see also discussion on the link between IP and the Bienenstock-Cooper-Munro rule (BCM) rule in appendix A.4— and may represent a parallel substrate for learning and memory [Daoudal & Debanne 2003]. Moreover, the experiments mentioned before suggest IP could implement some form of adaptive generalization of memory [Zhang & Linden 2003]. However, computational models of how this could be achieved are still missing.

2.3 Short-term plasticity

When a presynaptic spike occurs this information is transmitted at the postsynaptic side by the membrane fusion of presynaptic vesicles containing a certain neurotransmitter, causing it to be released in the synaptic cleft. The neurotransmitter then binds to postsynaptic receptors, causing a voltage change in the receiving neuron. The release of neurotransmitter is influenced by the recent firing history of the presynaptic neuron, on timescales of milliseconds to minutes [Abbott & Regehr 2004], a process named short-term synaptic plasticity. In this way, synaptic transmission at dynamic synapses encodes not only the presynaptic spike, but also the previous history of spiking on that channel.

More specifically, periods of elevated presynaptic activity can cause either an increase or a decrease in neurotransmitter release, a process known as short-term synaptic facilitation and short-term synaptic depression, respectively. The two seem to coexist at the same synapse, with their relative weight influenced by factors such as brain area (e.g. facilitation dominates in prefrontal areas, while depression dominates in sensory areas [Mongillo *et al.* 2008]) or the probability of release prior to the presynaptic stimulation (high favors depression, low – facilitation) [Abbott & Regehr 2004]. Additionally, the properties of these short term processes are themselves regulated by activity, on a slower timescale. The repetitive pairing or pre- and post- synaptic spikes alters depression parameters in a way that emphasizes the temporal aspects of the presynaptic signal [Tsodyks & Markram 1997, Li *et al.* 2004]. Interestingly, the degree to which coactivated synapses share short-term plasticity parameters can dramatically influence their ability to stimulate the postsynaptic neuron. Specifically, synapses exhibiting the same type of synaptic plasticity are more effective in influencing their postsynaptic target [Abbott & Regehr 2004].

Mechanistically, depression can arise due to the temporary depletion of presynaptic vesicles or due to postsynaptic receptor desensitization [Tsodyks & Markram 1997, Abbott *et al.* 1997], while facilitation is mediated by the increase in residual calcium at the presynaptic terminal [Fortune & Rose 2001, Mongillo *et al.* 2008]. The two can be described by a simple phenomenological model, widely used in computational models [Tsodyks & Markram 1997].

From a computational perspective, short-term synaptic dynamics have been shown to implement a temporal filter of the presynaptic signal [Fortune & Rose 2001]. Predominantly facilitatory or depressing synapses (i.e. having low or high baseline probability of neurotransmitter release) function as high- and low-pass filters, respectively. Again, these properties are not fixed. The type of filtering at a synapse

can be altered by neuromodulators and other factors. For example, presynaptic inhibition can convert a synapse from a low-pass to a band-pass filter, or from a band-pass to a high-pass filter [Abbott & Regehr 2004]. Additionally, short-term depression implements a dynamic gain control of different neuronal afferents, which is critical for extracting meaning information from many complex inputs, spanning a wide range of firing rates [Abbott *et al.* 1997]. Short-term plasticity also influences the properties of the transmitted signal. It has been suggested that synaptic depression removes redundant correlations so that postsynaptic signals convey information in a more efficient manner [Goldman *et al.* 2002]. Conversely, facilitating synapses favor burst-like clusters of transmission, resulting in transmission sequences that are more irregular and more positively correlated than the presynaptic spike trains that evoked them. Hence, facilitation favors information encoded by bursts of action potentials [Lisman 1997]. Lastly, synaptic facilitation can provide a robust and metabolically efficient implementation of working memory, with information stored in slow decaying Ca^{2+} transients, which can be read by spiking activity [Mongillo *et al.* 2008].

2.4 Structural plasticity

In addition to the various physiological changes mentioned above, structural plasticity can alter the connectivity matrix in a network. As we learn new skills or adapt to changes in the environment, brain structure changes as well [Hihara *et al.* 2006, Yoshida *et al.* 2003]. Although the overall neuronal morphology is remarkably stable over the animal's lifetime, rapid structural changes occur continuously at the level of spines and synapses, effectively rewiring cortical microcircuits [Turrigiano & Nelson 2000, Alvarez & Sabatini 2007, Holtmaat & Svoboda 2009, Bhatt *et al.* 2009].

As for LTP and LTD, the degree of spine plasticity depends on age and region; spines are more dynamic in juveniles than in adults, and more dynamic in S1 relative to V1. This is maybe not surprising, given that spine plasticity often accompanies experimentally induced synaptic plasticity [Alvarez & Sabatini 2007]. For example, LTP protocols cause the sprouting of new spines near the site of potentiation, possibly as a way to compensate for saturation effects at the potentiated site, providing additional synaptic substrate for learning [Turrigiano & Nelson 2000]. However, as for the regulation of synaptic strength, the regulation of synapse number is likely to be a complex process, involving several mechanisms [Turrigiano & Nelson 2000]. Local increases in synaptic numbers are globally compensated as to maintain a total number of synapses [Ziv & Ahissar 2009]. This interaction between positive and negative feedback loops could implement a process of synaptic redistribution, such that increasing the number of connections to one site competitively decreases the number of connections elsewhere. This type of competitive process could explain e.g. the synaptic retraction happening during the formation of ocular dominance columns. Exactly how this type of regulation could be done remains unclear, however.

The precise role of structural plasticity in learning remains controversial. Rapid changes in synaptic spines have been implicated in the fast components of the reorganization of visual maps following monocular deprivation, and experimental evidence supports a model in which new excitatory synapses could mediate the potentiation of open-eye responses [Hofer *et al.* 2009]. More evidence comes recently from motor learning experiments demonstrating that the acquisition of new motor tasks is associated with the formation of new spines [Xu *et al.* 2009, Yang *et al.* 2009]. This can occur by the selective stabilization of a dendritic filopodium that had contacted a nearby axon before or by the emergence of a new spine, later followed by the elimination of other spines [Xu *et al.* 2009, Yang *et al.* 2009]. Furthermore, behavioral performance correlates with the number of new spines formed shortly after training and the extent of spine elimination. Exposure to an enriched environment induces similar spine remodeling, this time restricted to sensory areas, whereas motor learning affects preferentially synapses onto motor neurons [Yang *et al.* 2009].

Interestingly, the structural changes induced by the subsequent acquisition of another motor task has minimal interference with the spines formed during the acquisition of the first task [Xu *et al.* 2009].

This may be not too surprising given the small proportion of synapses affected by learning, but raises the question how baseline changes in spine number, known to be about two orders of magnitude greater [Xu *et al.* 2009, Yang *et al.* 2009], can occur without destabilizing the memory engram. This apparent contradiction could be explained by the fact that baseline spine changes are restricted to functionally unimportant synapses [Ziv & Ahissar 2009]. From a relatively large pool of weak synapses, activity dependent processes can then select a few relevant synapses to be stabilized. These few synapses would then survive and eventually underlie behavior.

Whether experience-driven spine remodeling represents an essential component of long-term learning or a side-effect of other synaptic processes, these experimental findings open interesting avenues for computational modeling. So far, this area has received little theoretical attention [Fares & Stepanyants 2009, Kalantzis & Shouval 2009]. It is likely that this will change in the future, as computational models are needed to elucidate the functional implications of structural plasticity, specifically in the context of reward-dependent learning in recurrent neural networks. Some initial results in this direction are included in Chapter 4.

2.5 Plasticity and neuromodulation

All the different plasticity mechanisms discussed above can be regulated by several neuromodulators, the most prominent being dopamine (DA). Both the striatum and several cortical ares, including prefrontal cortex (PFC) receive dopaminergic projections and DA receptors are often located in close proximity to AMPARs and NMDARs [Reynolds & Wickens 2002, Seamans & Yang 2004]. Computationally, it is generally believed that DA encodes a prediction error signal in reinforcement learning, see [Daw & Doya 2006]. In the cortex, DA might facilitate the acquisition of task dependent internal representations of states and action, needed by the basal ganglia for action selection and reward prediction [Doya 2002].

The effects of dopamine on neural processing are manifold. Importantly, dopamine can modulate the sign and amplitude of synaptic plasticity [Calabresi *et al.* 2007, Shen *et al.* 2008]. dopamine affects LTD/LTP induction and is essential for protein synthesis and the long term maintenance of synaptic changes (e.g. shown in the hippocampus for NMDAR-dependent LTP after D1 receptor activation) [Seamans & Yang 2004]. This effect is not restricted to excitatory synapses, but has also been reported in GABAergic neurons using LTP/LTD induction protocols [Seamans *et al.* 2001b]. Furthermore, neuromodulation can also affect the rules governing STDP [Seol *et al.* 2007]. Overall, the effects of DA on plasticity depend on various factors, such as neuron and synapse type, induction protocol, DA concentration, or activated receptor (D1 or D2). To complicate things further, prior exposure to DA affects subsequent synaptic plasticity and some of the DA effects are long-lasting, persisting after the neuromodulator has been washed out from the tissue [Seamans & Yang 2004]. The precise mechanisms through which DA regulates various aspects of plasticity remain poorly understood, however. From a computational perspective, several models have been proposed for dopamine-modulated synaptic plasticity [Izhikevich 2007, Florian 2007], which will be discussed in detail in Chapter 4, in the context of reward-dependent learning in spiking neural networks.

Additionally, DA influences synaptic transmission by changing the kinetics of NMDA channels, a process hypothesized to play an important role in the persistent activity in PFC associated to working memory [Durstewitz *et al.* 1999], and can influence short-term plasticity, as demonstrated for both excitatory and inhibitory synapses in several cortical areas [Seamans & Yang 2004, González-Burgos *et al.* 2004]. Moreover, DA can also induce bidirectional changes neuronal excitability, see [Seamans & Yang 2004]. The ionic mechanism underlying this DA modulation of intrinsic excitability include several slowly inactivating or persistent Na^+ and K^+ channels, and various Ca^{2+} currents, regulating, among other, the membrane time constant and the spike threshold [Seamans & Yang 2004].

In the context of structural plasticity, it remains unclear whether the mechanisms required for the stabilization of specific synapses involve DA. Given that training-associated spine remodeling was not

observed in mice that failed to learn the task, it has been suggested that this stabilization criterion may be reward dependent [Ziv & Ahissar 2009]. Interestingly, the blockage of dopaminergic receptors within the same motor cortex regions was shown to block the acquisition, but not the execution of a previously learned motor behavior [Molina-Luna *et al.* 2009]. However, conclusive experimental evidence supporting a dopamine-modulated model of spine plasticity is still missing.

2.6 Interactions between different forms of plasticity

Although technical limitations usually force experimentalists to study each plasticity mechanisms in isolation, recent data suggests that different forms of plasticity interact in nontrivial ways. For example, in organotopic hippocampal slices, different forms of homeostatic plasticity were shown to act together, on different temporal scales to maintain the stability of the system. Changes in intrinsic excitability occur first, followed by slower changes in inhibitory connections, which readjust the balance of excitation and inhibition [Karmarkar & Buonomano 2006].

In freely swimming Xenopus tadpoles, exposure to 4-5 h of visual stimulation increases neuronal excitability of neurons in the optic tectum, by increasing voltage-gated Na^+ currents, which compensates for an initial homeostatic AMPAR downregulation. Together, these coordinated synaptic and neuron-specific changes render the neurons more responsive to synaptic bursts, thus improving stimulus detection [Aizenman *et al.* 2003].

In region CA1 of the hippocampus, a classic STDP-induction protocol causes not only the expected change in synaptic efficacy, but also a regulation of the integration properties of the dendritic branch where the affected synapses are located [Campanac & Debanne 2008]. This form of local E-S potentiation is NMDAR-dependent and has been hypothesized to implement a form of functional redundancy, which enforces synaptic plasticity, see also [Hanse 2008].

Other interesting interactions occur between Hebbian and homeostatic synaptic plasticity. In the classic view, synaptic scaling uniformly scales all synapses of a neuron to maintain a desired level of activity in the postsynaptic neuron. This is computationally attractive, as this weight normalization preserves memory traces, presumably stored as relative differences in synaptic strength. In this light, the reports on dendrite-specific forms of synaptic scaling raise important questions about how the cross-talk between homeostatic and Hebbian synaptic plasticity can occur in this case. The answer may be that this type of regulation introduces competition within a group of adjacent synapses, which may act as a functional subunit within the dendritic tree. Such local normalization is beneficial, as it supports dendritic compartmentalization. More unexpectedly, localized synaptic scaling would select which spatial patterns of potentiation/depression can be stably maintained. Specifically, it penalizes clustered inputs, i.e. if neighbors are not potentiated at the same time, changes in synaptic strength are more likely to be stabilized. Moreover, such homeostasis is consistent with branch plasticity, see [Rabinowitch & Segev 2008].

Metaplasticity

Beyond classic synaptic plasticity, experience can also change the plasticity rules themselves, a process termed metaplasticity [Abraham 2008]. The typical experiment for metaplasticity involves an episode of priming activity (electrical stimulation), followed by a classic synaptic plasticity induction protocol (HFS, LFS, pairing, etc.).

Although it has been introduced for a long time, the concept of metaplasticity still generates a great deal of confusion in the literature. In some interpretations, metaplasticity is described as a means to compensate the weight saturation due to Hebbian plasticity [Abraham 2008], making it indistinguishable from homeostatic plasticity. In other situations, metaplasticity may refer to a distinct set of plasticity mechanisms [Feldman 2009], with unclear computational implications.

This confusion is amplified by the fact that the mechanisms behind different types of metaplastic changes observed experimentally are poorly understood. It is believed that they require NMDAR activation, although several forms of mGluR-dependent metaplasticity have also been reported, see [Abraham 2008]. Moreover, if metaplasticity is often homosynaptic, there are also cases in which inducing plasticity in some dendritic region can result in metaplastic changes elsewhere, e.g. NMDAR-independent heterosynaptic metaplasticity in the hippocampus [Abraham 2008] or sliding threshold BCM (as noted before, sliding threshold BCM can be explained by interaction between fast Hebbian plasticity and a slower IP process, as shown in Appendix A.4). In the heterosynaptic case, stimulation of protein synthesis by activity in one set of synapses can facilitate LTP persistence through a synaptic tag-and-capture process operating at a second set of weakly activated synapses. Additionally, heterosynaptic metaplasticity can also be mediated by altered postsynaptic ion-channel function and retrograde endocannabinoid signaling that reduces transmission at nearby inhibitory GABAergic terminals [Feldman 2009].

Regulating the balance of excitation and inhibition can also be metaplastic, altering the threshold for LTD/LTP, see [Feldman 2009]. Interestingly, the maturation of a specific class of inhibitory neurons has been shown to regulate the onset of the V1 critical period, presumably by changing the balance of excitation-inhibition, leading to a state more permissive for learning at the excitatory synapses.

All-in-all, significant work is needed both experimentally and in computational models to be able to clarify the mechanisms and computational roles of metaplasticity.

2.7 Modeling plasticity

Rate-based models of synaptic plasticity

In a rate-based formulation the activity of the pre- and postsynaptic neurons is described in term of firing rates (v_{pre} and v_{post}). Under the assumption of locality, synaptic changes at the synapse can be described in the most general form as a function of these two quantities and of the current weight of the synapse [Gerstner & Kistler 2002a]: $\frac{\delta w}{\delta t} = f(v_{\text{pre}}, v_{\text{pre}}, w)$. Note, however, that this formulation assumes that the history of synaptic activation does not affect learning and additional rules may be needed to describe slow changes in learning metaparameters.

If expanding the above function around the point $v_{\text{pre}} = 0, v_{\text{post}} = 0$, one gets:

$$\frac{\delta w}{\delta t} = c_0(w) + c_1^{\text{pre}}(w)v_{\text{pre}} + c_1^{\text{post}}(w)v_{\text{post}} + c_2^{\text{corr}}v_{\text{pre}}v_{\text{post}} + c_2^{\text{pre}}(w)v_{\text{pre}}^2 + c_2^{\text{post}}(w)v_{\text{post}}^2 + O(w),$$

with parameters c which are all functions of w. Specific values for these parameters yield different Hebbian learning rules (see Table 2.7).

This framework provides a straightforward way for modeling several aspects of synaptic plasticity. It covers the classic formulation of Hebbian learning, LTD, saturation effects of LTP (in biology, weights cannot grow arbitrarily large, and some constraints on the maximal size of a synapse need to be enforced in the model, using either soft or hard bounds). synaptic scaling can also be implemented, for example by the Oja rule [Oja 1982]. This implements competition between different synapses and is mathematically equivalent to a constraint on the incoming synapses to a neuron of the form $\sum_i w_{ij} = \text{const.}$.

As a generalization of the above model, we can include the history of observed firing rates for the pre- and post-synaptic site (such that the coefficients will be $c(w, \langle v_{\text{pre}}\rangle, \langle v_{\text{post}}\rangle)$). In this way, we can derive other classic synaptic plasticity rules, such as sliding threshold BCM [Bienenstock *et al.* 1982] or the covariance rule [Stanton & Sejnowski 1989]: $\frac{dw}{dt} = \alpha(v_{\text{pre}} - \langle v_{\text{pre}}\rangle,)(v_{\text{post}}\langle v_{\text{post}}\rangle$.

Spike-based models of synaptic plasticity

A similar approach can be taken for the case of a spike-based encoding. As before weight changes are assumed to depend only on information that is local to the synapse [Gerstner & Kistler 2002a]: $\frac{\delta w(t)}{\delta t} =$

Table 2.4. Effects of different rate-based synaptic plasticity rules

c_2^{corr}	c_2^{pre}	c_2^{post}	c_1^{pre}	c_1^{post}	c_0	Description
1						Classic Hebb
1					$-\gamma w$	LTD
1			$-v_\theta$			presynaptically gated LTD
1				$-v_\theta$		postsynaptically gated LTD
$\eta(w_{\max} - w)$						Weight saturation $w \leq w_{\max}$, soft bounds
η		$-\eta w$				Weights normalization (Oja)
1					$\gamma w(1-w)(w-w_0)$	Weight consolidation (bistability)

Parameters: η, γ - learning rates, v_θ - threshold for potentiation, $w_{\max}$ - upper bound for w, w_0 - weight threshold for swiching a bistable synapse between 0 and 1. Based on [Gerstner & Kistler 2002a].

$F[u_{\text{pre}}(t_1), u_{\text{post}}(t_2), w(t))$. However, here the membrane potential values $u_{\text{pre/post}}$ are not instantaneous, but functions of the previous history of the activity at the pre- and postsynaptic terminal, and f is a functional of these two and the current synaptic strength [Gerstner & Kistler 2002a].

In this case, we can use a Volterra expansion of the weight change, which yields:

$$\begin{aligned}\frac{dw}{dt} &= c_0(w) + \int_0^\infty \alpha_1^{\text{pre}}(w,s)u_{\text{pre}}(t-s)ds \\ &+ \int_0^\infty \alpha_1^{\text{post}}(w,s)u_{\text{post}}(t-s)ds \\ &+ \int_0^\infty \int_0^\infty \alpha_2^{\text{corr}}(w,s,s')u_{\text{pre}}(t-s)u_{\text{post}}(t-s')dsds' \\ &+ O(w).\end{aligned}$$

This expression may seem counterintuitive, but, under certain assumptions, it can be traced back to our previous rate formulation and to classic spike-based formulations of synaptic learning. To do this, we assume the presynaptic signal is defined as a spike train $u_{\text{pre}} = \sum_f \delta(t - t_{\text{pre}}^f)$. Additionally, we use the spike-response model introduced in [Gerstner & Kistler 2002b] for modeling the postsynaptic signal, which results in an expression of the form: $u_{\text{post}} = \eta(t - t_{\text{post}}^f) + h(t)$, where η defines the time course of a BAP,and h the local membrane potential at the synapse. Then, we can further simply the expression of $\frac{dw}{dt}$ by making additional assumptions on u_{post}.

Here, we consider two particular approximations. First, if the postsynaptic membrane potential is strongly influenced by BAPs, it can be modeled as a spike train, as done for the presynaptic neuron: $u_{\text{post}} = \sum_f \delta(t - t_{\text{post}}^f)$. This results in an update function of the form:

$$\begin{aligned}\frac{dw}{dt} &= c_0(w) + \sum_{t_{\text{pre}}^f} \alpha_1^{\text{pre}}(w, t - t_{\text{pre}}^f) \\ &+ \sum_{t_{\text{post}}^f} \alpha_1^{\text{post}}(w, t - t_{\text{post}}^f) \\ &+ \sum_{t_{\text{pre}}^f} \sum_{t_{\text{post}}^f} \alpha_2^{\text{corr}}(w, t - t_{\text{pre}}^f, t - t_{\text{post}}^f) \\ &+ O(w).\end{aligned}$$

Note that here weight changes are continuous in time, i.e. the effects a single spike event unfold over a certain time interval. Under a renotation: $c_1^{\text{pre/post}} = \int_0^\infty \alpha_1^{\text{pre/post}}(w,s)ds$, and assuming that this time interval is small, the above equation is equivalent with the rate formulation:

$$\begin{aligned}\frac{dw}{dt} &= c_0(w) + c_1^{\text{pre}} v_{\text{pre}} \\ &\quad + c_1^{\text{post}} v_{\text{post}} \\ &\quad + \cdots .\end{aligned}$$

The fact that for weak or uncorrelated pre/post- synaptic spikes, several spike-based models can be reduced to a rate-based formulation has been demonstrated before [Abbott & Nelson 2000, Izhikevich & Desai 2003, Toyoizumi *et al.* 2005]. Since the same holds for the generic formulation of spike-based learning, the considerations made for the rate model on the implication of different terms hold in this case as well.

At the other extreme, we could assume that the BAP has a negligible effect on u_{post}, i.e. $\eta(t - t^f_{\text{post}}) \approx 0$, resulting in a presynaptically gated voltage-dependent learning rule, similar to that introduced by [Fusi *et al.* 2000], in which presynaptic spikes trigger weight changes, with an amplitude and sign that depend on the postsynaptic membrane potential.

Lastly, the correlation term in expression 2.1 automatically introduces a learning window within which pre- and postsynaptic spikes interact to cause a change in synaptic strength. As a particular case, this could be the standard antisymmetric biphasic window of classic STDP, but it can also incorporate other learning windows (e.g. those reported for GABAergic neurons).

In practice, several STDP models have been proposed for the case of excitatory synapses, from standard all-to-all STDP [Abbott & Nelson 2000], nearest-neighbor STDP [Izhikevich & Desai 2003] or weight-dependent STDP [van Rossum *et al.* 2000], to more complex formulations, which take into account the spiking history on several time-scales, such as those proposed by [Froemke & Dan 2002] and [Toyoizumi *et al.* 2005, Pfister & Gerstner 2006, Clopath *et al.* 2010]. As several STDP versions are used in Chapter 3, these will described in detail in the Methods section of our ICA model.

Nonsynaptic plasticity

Unfortunately, so far little work has been done on modeling IP. Many models use some form of threshold adaptation for computational reasons, as it helps stabilize the dynamics, e.g. [Foldiák 1990], but without explicitly referring to IP. The first work to define a principled way for implementing IP was [Stemmler & Koch 1999]. This particular model used Hodgkin-Huxley neurons and derived an analytical formulation for the adaptation of the neuron conductances such that the distribution of the output firing rates is exponential. There, it was argued that the exponential ensures maximum output entropy, given the biological constraint of a fixed mean firing rate. The same idea was later used in [Triesch 2007] for a rate neuron with a sigmoidal transfer function, in combination with synaptic learning. In the work presented here, we have used the same IP derivation for a different transfer function, in the context of a stochastically spiking neuron [Toyoizumi *et al.* 2005]. The complete mathematical derivation of our IP implementation can be found in Appendix A.

3

ICA learning in spiking neurons

Independent component analysis (ICA) is a well-known signal processing technique for extracting statistically independent components from high-dimensional data. For the brain, ICA-like processing is essential for building efficient representations of sensory data [Barlow 2001, Simoncelli & Olshausen 2001, Simoncelli 2003]. However, although many algorithms have been proposed for solving the ICA problem, the mechanisms by which this type of processing can be realized in networks of spiking neurons remain unclear.

Classic ICA algorithms often exploit the nongaussianity principle, which allows the ICA model to be estimated by maximizing some nongaussianity measure, such as kurtosis or negentropy [Hyvärinen *et al.* 2001]. A related representational principle is sparse coding, which has been used to explain various properties of V1 receptive fields [Olshausen & Field 1997]. Sparse coding states that only a small number of neurons are activated at the same time, or alternatively, that each individual unit is activated only rarely [Olshausen & Field 1996a]. In the context of neural circuits, it offers a different interpretation of the goal of the ICA transform, from the perspective of metabolic efficiency. As spikes are energetically expensive, neurons have to operate under tight metabolic constraints [Lennie 2003], which affect the way information is encoded. Moreover, experimental evidence supports the idea that the activity of neurons in V1 is sparse [Baddeley *et al.* 1997].

As we have seen in the previous chapter, some forms of IP are thought to regulate the distribution of firing rates of a neuron [Stemmler & Koch 1999]. Computationally, such forms of homeostatic plasticity may maximize information transmission of a neuron, under certain metabolic constraints [Stemmler & Koch 1999]. Additionally, when interacting with Hebbian synaptic plasticity, IP allows the discovery of heavy-tailed directions in the input [Triesch 2007]. Here, we combine STDP [Gerstner *et al.* 1996,Markram *et al.* 1997a, Bi & Poo 1998b], synaptic scaling, and an IP rule similar to [Triesch 2007], which tries to make the distribution of instantaneous neuronal firing rates close to exponential. This model of IP is supported by experimental evidence on close to exponential distributions of firing rates in various visual areas in response to natural scenes [Baddeley *et al.* 1997] and by information theoretic arguments, as the exponential distribution is known to have maximum entropy for a fixed energy budget [Stemmler & Koch 1999, Triesch 2007].

We show that IP and synaptic scaling complement STDP learning, allowing single spiking neurons to learn useful representations of their inputs for several ICA problems. Firstly, output sparsification by IP together with synaptic learning is sufficient for demixing two zero mean supergaussian sources, a classic formulation of ICA. When using biologically plausible inputs and STDP, complex tasks, such as Foldiák's bars problem [Foldiák 1990], and learning oriented receptive fields for natural visual stimuli, can be tackled. Moreover, a neuron population learns to extract several independent components if neuron activities are decorrelated by adaptive lateral inhibition. We show that our IP rule is necessary for learning, as it enforces a sparse output, guiding learning towards heavy-tailed directions in the input. Moreover, for specific STDP implementations, it implements a sliding threshold for BCM learning [Bienenstock *et al.* 1982].

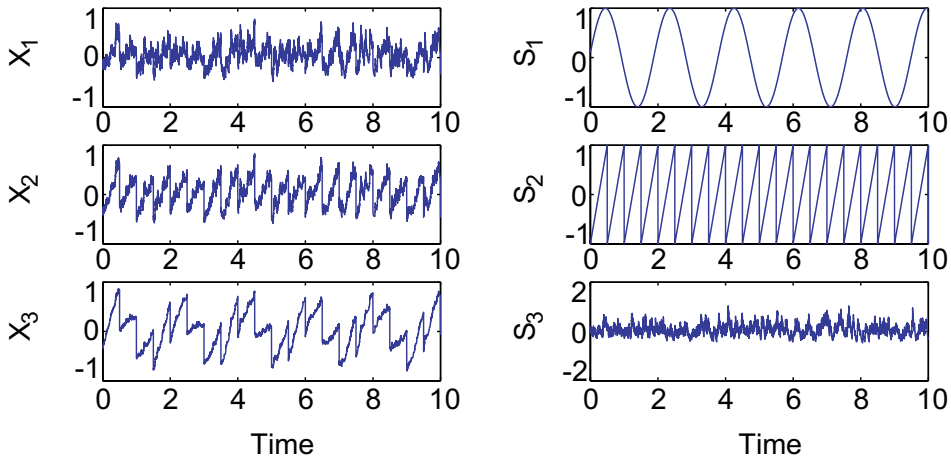

Figure 3.1. A blind source separation example. Given a set of mixtures X_i, BSS estimates the independent components S_i.

The underlying assumption behind our approach, implicit in all standard models of V1 receptive field development, is that both input and output information are encoded in rates. In this light, one may think of our current work as a translation of the model in [Triesch 2007] to a spike-based version. However, the principles behind this approach are more general than suggested by the previous work with rate neurons. Specifically, we show that the same rule can be applied when inputs are encoded as spike-spike correlation patterns, where a rate-based model would fail. The results presented in this chapter have been published as [Savin *et al.* 2010].

3.1 Background

3.1.1 Independent component analysis

ICA as blind source separation

At some point, we have all experienced the situation of finding ourselves in a crowded room, with many people speaking at the same time, and may have wondered how it is that, from this sea of noise, we still manage to follow a conversation with our neighbor. Known as 'the cocktail party problem', this example is a classical illustration of what in signal processing is called blind source separation (BSS), a well known ICA application.

More formally, BSS considers a set of signals $s(t)$ emitted by different sources, like several people speaking in a room, or electrical signals emitted by different brain areas. Multiple receivers, e.g. microphones in different locations in the room, our ears, electrodes from different positions on the scalp in EEG, capture different combinations of these signals $x(t)$. Assuming that signal superposition is linear, i.e. $x(t) = A \cdot s(t)$, BSS is the problem of recovering the source signals and the matrix A after observing only the signal mixtures x. Recovering the sources in this case is 'blind', as little information is available about the original sources: both the sources and the mixing matrix are unknown.

An example of some linear combinations of three independent signals is shown in Fig. 3.1. Surprisingly, retrieving the original signals and the mapping to the observed signals has a simple solution under one critical assumption: that the sources are statistically independent. This constraint is at the heart of ICA, differentiating it from other linear decompositions of multivariate data, e.g. PCA or factor analysis.

ICA definition

Formally, ICA can be defined as a generative model with independent latent variables. In this framework, a set of observations of the mixtures x_i is seen as being generated by the linear superposition of some unknown sources s_{i}: $x = A \cdot s$, where s is a column vector of the sources s_i. As for the BSS problem before, ICA means estimating A and s_{i}, given x_{i}. In its general form, s_{i} are sampled from a fixed distribution and need not be time dependent, as in the BSS example before. Also, the matrix A is often assumed to be square, for simplicity.

From a neural perspective, the relation is usually rewritten as: $s = W \cdot x$, with s being the response of neurons to the presentation of a sensory input x (e.g. the V1 response to an image) and $W = A^{-1}$ containing the weights of incoming synapses to each neuron.

For convenience, we typically assume x to be white (i.e. to have zero mean and unit variance), which implies that the matrix A is orthogonal ($A^{-1} = A^T$). In the visual system, this type of input preprocessing is believed to be done in the lateral geniculate nucleus (LGN) [Dan *et al.* 1996].

The ICA problem is well-defined provided that at most one of the sources has a gaussian distribution. This can be easily explained by noting that gaussian white distributions are rotation invariant, meaning that any orthogonal transform A would yield a distribution that is indistinguishable from the original. In real-world data, interesting distributions will often be supergaussian (peaked at zero and heavy-tailed).

ICA implementations

A variety of solutions have been proposed in the literature for solving the ICA problem, see [Hyvärinen *et al.* 2001]. Most implementations start by defining a certain objective function which measures the independence of different sources. Numerical methods are then used to optimize this function, typically by gradient descent. Combining different objective functions and optimization procedures yields a zoo of closely related ICA implementations. We will present the most prominent approaches in what follows.

A fundamental principle of ICA is the principle of *nongaussianity* — the independent components (ICs) are directions for which the projections of the data are the least gaussian. It is easy to see why this is the case: due to central limit theorem, the linear combination of several independent random variables is always more gaussian than each of the original sources. Hence, many classic ICA algorithms are trying to maximize the nongaussianity of the distribution of s.

In practice, different measures can be used for estimating nongaussianity. A simple and computationally efficient measure of nongaussianity is kurtosis, a generalization of variance using higher order cumulants, defined as: $kurt(y) = E(y^4) - 3 \cdot (E(y^2))^2$. Assuming that the input is whitened, this can be easily measured as a forth order cumulant of the distribution. As an alternative, entropy could be used as nongaussianity measure, by taking into account the fact that a gaussian has the largest entropy from all distributions of unit variance. In practice, the objective function is typically defined in terms of the negentropy: $J(x) = H(x_{\mathrm{gauss}}) - H(x)$, where x_{gauss} is a gaussian random variable with same covariance matrix C as x and the entropy of the gaussian distribution can be computed as: $H(x_{\mathrm{gauss}}) = \frac{1}{2} \log |\det(\mathrm{C})| + \frac{n}{2}(1 + \log(2\pi))$. This formulation is particularly nice because the negentropy is scale invariant, i.e. it does not change after multiplication by a constant. For both kurtosis and negentropy, practical algorithms have been derived based on either gradient or fixed point methods, see [Hyvärinen *et al.* 2001].

A closely related statistical method is *projection pursuit* [Huber 1985], initially introduced for visualizing high-dimensional data. Projection pursuit tries to find good representations of high-dimensional spaces by projecting data on an 'interesting' lower dimensional space. As for the ICA algorithms mentioned above, a typical measure of interest is the nongaussianity of the distribution in the low dimensional space. The initial motivation for this approach was that interesting clusters in the data have a multimodal distribution.

Another popular approach for ICA is *maximum likelihood* (ML), which tries to optimize the model

parameters (A) such that they best explain the data. The probability of the data, given the model can be easily computed as: $p(x) = |det(A^{-1})|\Pi_{\mathrm{i}} p_{\mathrm{i}}(s_{\mathrm{i}})$, where p_{i} denotes the probability density function of the individual components, giving an expression for the log-likelyhood:

$$\log(W) = \Sigma\Sigma \log(p_{\mathrm{i}}(w_{\mathrm{i}}^{\mathrm{T}})x) + N \log|det(W)|,$$

with N, the number of samples and $W = A^{-1}$, as before. Again, the above expression can be optimized using either gradient or fixed-point methods. A special case of this method, which uses the simple gradient, is known in the literature as the Bell-Sejnowski, or infomax algorithm [Bell & Sejnowski 1997]. As the name suggests, it was initially derived based on the *infomax principle*. More specifically, given a neural network with nonlinear units, the algorithm attempts to maximize the output entropy, or equivalently the input-output information flow. A caveat of this implementation is that the neuron transfer function has to be selected to be the cumulative probability distribution of p_{i}, assumed to be known *a priori*.

In some cases, the independence assumption may not be strictly true. Still, it could be interesting to find a linear representation of the data such that the components are as independent as possible. A good measure of the degree of dependence between different sources is their *mutual information* (MI), defined as $\mathrm{MI}(s_1, \ldots, s_{\mathrm{m}}) = \Sigma H(s_{\mathrm{i}}) - H(\mathrm{s})$, where $H(\cdot)$ denotes the entropy. That is, MI measures the deviation of the distribution from the null hypothesis that the variables s_{i} are independent. Estimating the ICs using MI is a very general method, as it makes no assumptions about the data. Moreover, it provides a rigorous theoretical explanation why the nongaussianity principle and the ML algorithms can be used for independent component estimation. More precisely, it can be easily shown that MI can be estimated using either the negentropy, or alternatively the likelihood. Hence, MI minimization yields essentially the same cost functions and algorithms as those described above.

What makes ICA more powerful than other linear component extraction methods, such as PCA, is the independence principle. More specifically, ICA goes beyond the constraint of decorrelation, requiring statistical independence. This means that ICA takes into account higher-order moments of the output distribution. More formally, ICA requires nonlinear uncorrelatedness, i.e. if s_1 and s_2 are two independent sources, they remain uncorrelated after any nonlinear transformations $h(s_1)$ and $g(s_2)$ [Hyvärinen *et al.* 2001]. This observation leads to an alternative ICA implementation, using a nonlinear version of a decorrelation method, such as PCA. The caveat in this case is how to choose suitable nonlinearities, a problem which has been investigated analytically in [Hyvärinen 1997]. Historically, nonlinear PCA is one of the early implementations of ICA [Oja 1997].

The discussion so far has been restricted to the case when A is square. In practice, it is often the case that there are fewer mixtures than the actual number of sources (overcomplete basis), such that information is lost in the mixing process. This makes the problem more difficult, because the matrix A is no longer invertible, and both the mixing matrix and the sources need to be estimated separately. First, assuming that A is know, the sources can be recovered using the pseudo-inverse $s = A^T(AA^T)^{-1}x$, or by ML (which yields an interesting linear programming solution when using a Laplacian prior) [Hyvärinen *et al.* 2001]. Second, the mixing matrix A can be estimated by maximizing the joint likelihood over A and s. An interesting by-product of an overcomplete basis is a sparse representation, which assumes that only a few sources are active in each mixture (more on this below).

Although all the ICA methods we have presented are closely related, they may differ in terms of statistical (e.g. robustness to outliers) and numerical (e.g. convergence speed, numerical stability) properties. Additionally, some methods require the input to be preprocessed by whitening or make use of prior knowledge about the ICs distribution. An important difference is between cost functions which allow the ICs to be estimated one at the time, termed single-unit contrast functions (our learning rule falls in this category), and those which simultaneously estimate the full basis. The first are more interesting from a biological perspective, as they tend to make use of only local information. Moreover, several units can be obtained by an additional constraint that the activity of different units has to be uncorrelated (the orthogonalization of the weight matrix can be done by standard methods, such as Gram-Schmidt or by symmetric orthogonalization, as done in FastICA, see [Hyvärinen *et al.* 2001]).

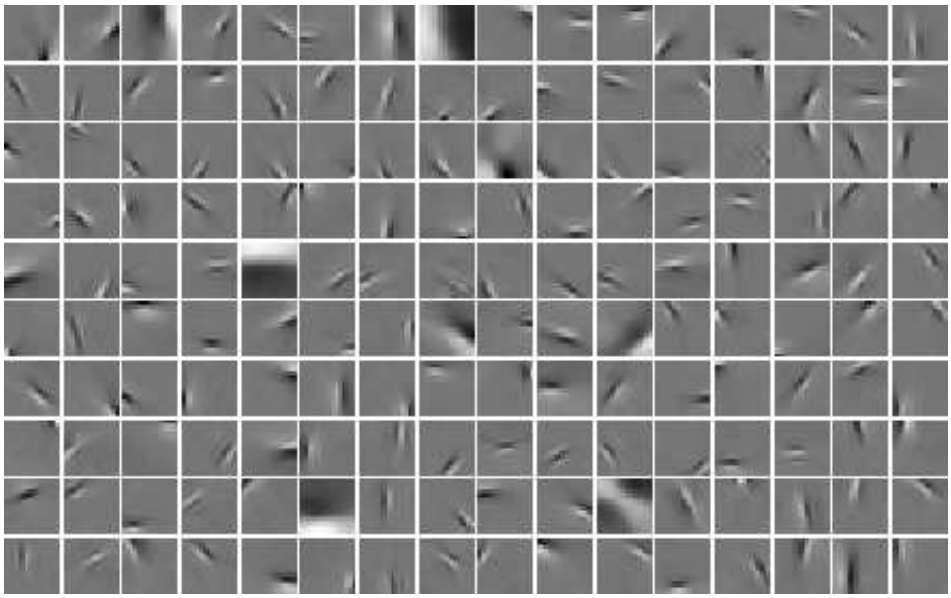

Figure 3.2. Example of a feature basis obtained when ICA is applied to natural images.

3.1.2 Models of receptive field development

An important application of ICA is modeling the properties of neurons in primary sensory areas, such as simple cells in primary visual cortex V1. When applied to natural images, ICA yields a decomposition very similar to that given by Gabor or wavelet analysis [Hyvärinen *et al.* 2001], with receptive fields localized in space and spatial frequency, and having a certain orientation (see Fig. 3.2). When compared to single-unit recordings from macaque V1, ICA gives a relatively good match, although this may vary for different algorithms [Bell & Sejnowski 1995, Bell & Sejnowski 1997, Olshausen & Field 1997, Weber & Triesch 2008, Lücke 2009].

Simple cells responses describe only a subset of neurons in V1. For the other neurons, termed complex cell, more elaborate models, such as independent subspace analysis [Hyvärinen *et al.* 2001], need to be used. All-in-all, our current understanding of V1 function remains limited [Olshausen & Field 2004a, Olshausen & Field 2005, Carandini *et al.* 2005].

Rate-based formulations of ICA

Several neural-based implementations of ICA have been proposed in the literature. Some of them, such as nonlinear PCA and infomax, we have already mentioned. Another prominent representational principle closely related to ICA is sparse coding, which has been used to explain various properties of V1 receptive fields [Olshausen & Field 1997]. Sparse coding is based on the assumption that only a small number of sources are active at the same time (population sparseness), or alternatively, that each individual source is activated only rarely (single-unit sparseness) [Olshausen & Field 1996b]. In neural terms, this means that individual stimuli activate only a few neurons within a neural population.

As sparseness is closely related to supergaussianity, linear sparse coding can be estimated using similar contrast functions as those used for ICA. It was originally formulated as a linear generative model with an additional white gaussian noise term. The sources could then be estimated by maximum likelihood, under the constraint of a sparse factorial prior for s, with some additional weight normalization. From

a neural perspective, this translates into a process similar to an analysis-synthesis loop [Mumford 1994], with weights updated by Hebbian learning determined by the outputs s and the residual image estimated through negative feedback connections [Olshausen & Field 1997].

Sparse coding provides a different interpretation of the goal of this transform, in terms of information compression or denoising, see [Hyvärinen *et al.* 2001]. From a neural perspective, it models the constraint of a limited energy budget [Attwell & Laughlin 2001]. Moreover, it has been argued that, when combined with overcompleteness, sparse coding produces a simpler representation of the curved manifold structure of data for natural scenes, easing the decoding by higher visual areas [Olshausen & Simoncelli 2001]. Interestingly, primary sensory cortical areas often over-represent their sensory inputs (from the thalamus), e.g. in cat V1 the ratio between the number of axons projecting from L2/3 to higher areas relative to the number of inputs from the LGN is approximately 25:1 [Olshausen & Field 2004b]. In this context, it has been suggested that this overcomplete basis could explain some of the nonlinearities in the response of simple cells, see [Olshausen & Field 1997].

ICA in spiking neurons

If several neural implementations of ICA-like learning exist for rate neurons, see [Hyvärinen 1999, Hyvärinen *et al.* 2001], relatively little has been done using spiking neurons [Klampfl *et al.* 2009, Parra *et al.* 2009]. First, a generalization of the information bottleneck contrast function has been used for minimizing the mutual information between two neurons [Klampfl *et al.* 2009]. This approach is however, very limited in scope as the cost function cannot straightforwardly generalize to a larger population. Moreover, the resulting synaptic learning rule is nonlocal.

A more interesting approach was introduced in [Parra *et al.* 2009]. Considering a deterministic transfer function, with the output forced to have maximum entropy (i.e. a set of independent homogeneous Poisson processes of given mean) this algorithm maximizes the likelihood of the input, under the inverse transfer function (similar to infomax). This optimization yields a population based formulation for weight changes, with a non-local update rule. Unlike infomax, where this expression could be reduced to a local form by using the natural gradient, in the spike-based version this is not feasible, due to the complexity of the expressions. Nonetheless, the advantage of this formulation is that it is very general, allowing information to be encoded either in precise spike times, or using a population code.

Lastly, given that BCM-learning can discover heavy-tailed directions in the input [Intrator & Cooper 1992, Intrator 1998] and learn Gabor receptive fields [Blais *et al.* 1998] in linear neurons, an interesting candidate for a spike-based ICA implementation is a variant of STDP which implements a generalization of BCM [Toyoizumi *et al.* 2005] (although relatively straightforward, this has not been shown before, see appendix 1). This STDP rule optimizes the information transmission between input and output, under the constraint of a fixed mean. We will see in the following that this approach can be viewed as a special case of our ICA implementation.

3.2 Computational model

A schematic view of the learning rules used is shown in Fig. 3.3. The stochastically spiking neuron [Toyoizumi *et al.* 2005] generates spikes as an inhomogeneous Poisson process, with mean expressed as a function of the total incoming current to the neuron $g(u)$, parametrized by variables r_0, u_0 and u_α. This transfer function of the neuron is optimized by adapting the three parameters to make the distribution of instantaneous firing rates of the neuron exponential. Additionally, incoming weights change by nearest-neighbor STDP [Izhikevich & Desai 2003] and a synaptic scaling mechanism keeps the sum of all incoming weights constant over time. The complete mathematical description of the model follows below.

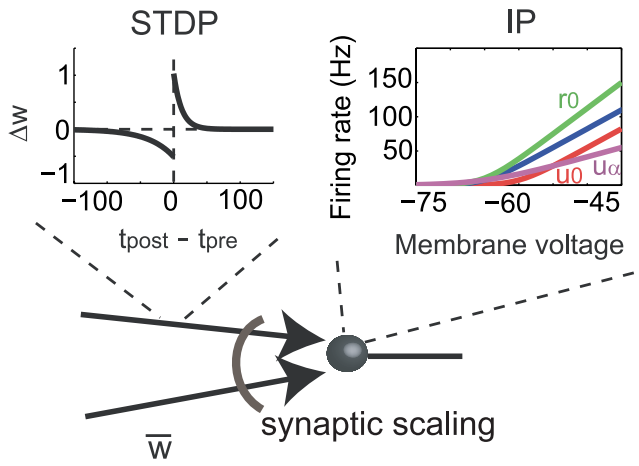

Figure 3.3. Overview of plasticity rules used for ICA-like learning. Synapses are modified by nearest-neighbor STDP and synaptic scaling. Additionally, intrinsic plasticity changes the neuron's transfer function by adjusting three parameters r_0, u_0, and u_α. Different transfer functions show the effects of changing each of the three parameters individually relative to the default case depicted in blue. Namely, r_0 gives the slope of the curve, u_0 shifts the entire curve left or right, while u_α can be used for rescaling the membrane potential axis. Here, r_0 is increased by a factor of 1.5, u_0 by 5 mV, u_α by a factor of 1.2.

3.2.1 Neuron model

We consider a stochastically spiking neuron with refractoriness [Toyoizumi *et al.* 2005]. The model defines the neuron's instantaneous probability of firing as a function of the current membrane potential and the refractory period of the neuron, given as a function of the time since its last spike. More specifically, the membrane potential is computed as:

$$u(t) = u_\mathrm{r} + \sum_{j,f} w_\mathrm{j} \epsilon(t - t_\mathrm{j}^f),$$

where $u_\mathrm{r} = -70\,\mathrm{mV}$ is the resting potential, while the second term represents the total incoming drive to the neuron, computed as the linear summation of post-synaptic potentials evoked by incoming spikes. Here, w_j gives the strength of synapse j, t_j^f is the time of a presynaptic spike, and $\epsilon(t - t_\mathrm{j}^f)$ the corresponding evoked post-synaptic potential, modeled as a decaying exponential, with time constant $\tau = 10\,\mathrm{ms}$ and amplitude 1 mV.

The refractory state of the neuron, with values in the interval $[0, 1]$, is defined as a function of the time of the last spike $\hat{t}$, namely:

$$R(t) = \frac{(t - \hat{t} - \tau_\mathrm{abs})^2}{\tau_\mathrm{refr}^2 + (t - \hat{t} - \tau_\mathrm{abs})^2} \Theta(t - \hat{t} - \tau_\mathrm{abs}),$$

where $\tau_\mathrm{abs} = 3\,\mathrm{ms}$ gives the absolute refractory period, $\tau_\mathrm{refr} = 10\,\mathrm{ms}$ is the relative refractory period and $\Theta(\cdot)$ is the Heaviside function.

The probability $\rho(t)$ of the stochastic neuron firing at time t is given as a function of its membrane potential and refractory state [Toyoizumi *et al.* 2005]:

$$\rho(t) = 1 - e^{-g(u(t))R(t)\Delta t} \approx g(u(t))R(t)\Delta t,$$

where $\Delta t = 10^{-3}$ s is the time step of integration, and $g(u(t))$ is a gain function, defined as:

$$g(u) = r_0 \log\left(1 + e^{\frac{u-u_0}{u_\alpha}}\right).$$

Here r_0 , u_0 and u_α are model parameters, whose values are adjusted by intrinsic plasticity.

3.2.2 Intrinsic plasticity

Our intrinsic plasticity model attempts to maximize the mutual information between input and output, for a fixed energy budget [Triesch 2007, Joshi & Triesch 2009]. More specifically, it induces changes in neuronal excitability that lead to an exponential distribution of the instantaneous firing rate of the neuron [Stemmler & Koch 1999]. The specific shape of the output distribution is justified from an information theoretic perspective by the fact that the exponential distribution has maximum entropy, for a fixed mean. This is true for distributions defined on the interval $[0, \infty)$, but under certain assumptions it can be a good approximation for the case where the interval is bounded, as it happens in our model, due to the neuron's refractory period (see below). Optimizing information transmission under the constraint of a fixed mean is equivalent to minimizing the Kullback-Leibler divergence between the neuron's firing rate distribution and that of an exponential with mean μ:

$$\begin{aligned} D &= d(p_{\text{neuron}} \| p_{\text{exp}}) \\ &= \int f_Y(y) \log\left(\frac{f_Y(y)}{\frac{1}{\mu}e^{-y/\mu}}\right) dy \\ &= -H(Y) + \frac{1}{\mu}E(Y) - \log(\mu), \end{aligned}$$

with $y = g(u)$, $H(\cdot)$ denoting the entropy and $E(\cdot)$ the expected value. Note that the above expression assumes that the instantaneous firing rate of the neuron is proportional to $g(u)$, that is that $R(t) \approx 1$. When taking into account the refractory period of the neuron, which imposes an upper-bound $R = \frac{1}{\tau_{\text{abs}}}$ on the output firing rate, the maximum entropy distribution for a specific mean $\mu \leq R$ is a truncated exponential [Kapur 1990]. The deviation between the optimal exponential for the infinite and the bounded case depends on the values of μ and R, but it is small in cases in which $\mu \ll R$. hence, our approximation is valid as long as the instantaneous firing rate μ is significantly lower than $1/\tau_{\text{abs}}$, that is when the mean firing rate of the neuron is small. In our case, we restrict $\mu \leq 10$ Hz. If not otherwise stated, all simulations have $\mu = 2$ Hz. Note also that the values considered here are in the range of firing rates reported for V1 neurons [Olshausen & Simoncelli 2001].

Computing the gradient of D for r_0, u_0 and u_α, and using stochastic gradient descent, the optimization process translates into the following update rules [Joshi & Triesch 2009]:

$$\begin{aligned} \Delta r_0 &= \frac{\eta_{\text{IP}}}{r_0}(1 - \frac{g}{\mu}), \\ \Delta u_0 &= \frac{\eta_{\text{IP}}}{u_\alpha}\left(\left(1 + \frac{r_0}{\mu}\right)\left(1 - e^{-\frac{g}{r_0}}\right) - 1\right), \\ \Delta u_\alpha &= \frac{\eta_{\text{IP}}}{u_\alpha}\left(\frac{u - u_0}{u_\alpha}\left(\left(1 + \frac{r_0}{\mu}\right)\left(1 - e^{-\frac{g}{r_0}}\right) - 1\right) - 1\right). \end{aligned}$$

Here, $\eta_{\text{IP}} = 10^{-5}$ is a small learning rate.

Additionally, as a control, we have considered a simplified rule, which adjusts a single transfer function parameter in order to maintain the mean firing rate of the neuron at a constant value μ. More specifically,

a low-pass-filtered version of the neuron firing rate is used to estimate the current mean firing rate of the neuron:

$$\frac{d\bar{g}}{dt} = -\frac{\bar{g}}{\tau_{\rm m1}} + \delta(t - t_{\rm f}^f),$$

where δ is the Dirac function and $t_{\rm f}^f$ is the time of firing of the post-synaptic neuron and $\tau_{\rm m1} = 100$ms. Based on this estimate, the value of the parameter r_0 is adjusted as:

$$r_0 = r_0 - \eta_{\rm m1}(\bar{g} - \mu).$$

Here, μ is the goal mean firing rate, as before and $\eta_{\rm m1}$ is a learning rate, set such that, for a fixed gaussian input distribution, convergence is reached as fast as for our IP rule described before ($\eta_{\rm m1} = 10^{-4}$).

3.2.3 Spike-timing dependent plasticity

The STDP rule implemented here considers only nearest-neighbor interactions between spikes [Izhikevich & Desai 2003]. The change in weights is determined by:

$$\Delta w_{\rm ij} = \begin{cases} A_+ e^{-\frac{\Delta t}{\tau_+}}, & \text{if} \quad \Delta t > 0 \\ A_- e^{\frac{\Delta t}{\tau_-}}, & \text{if} \quad \Delta t < 0 \end{cases},$$

where $A_{+/-}$ is the amplitude of the STDP change for potentiation and depression, respectively (default values $A_+ = 1.03 \cdot 10^{-4}$ and $A_- = -0.51 \cdot 10^{-4}$), $\tau_{+/-}$ are the time scales for potentiation and depression ($\tau_+ = 12$ ms, $\tau_- = 38$ ms; for learning spike-spike correlations $\tau_+ = 10$ ms) [Bi & Poo 1998b], and $\Delta t = t^{\rm post} - t^{\rm pre}$ is the time difference between the firing of the pre- and post-synaptic neuron. For the lateral inhibitory connections, the STDP learning is faster, namely $A_+ = 5 \cdot 1.03 \cdot 10^{-4}$ and $A_- = -5 \cdot 0.51 \cdot 10^{-4}$. In all cases, weights are always positive and clipped to zero if they become negative.

This STDP implementation is particularly interesting as it can be shown that, under the assumption of uncorrelated or weakly correlated pre- and post-synaptic Poisson spike trains, it induces weight changes similar to a BCM rule, namely [Izhikevich & Desai 2003]:

$$\Delta w \approx xy \left(\frac{A_+}{\tau_+^{-1} + y} + \frac{A_-}{\tau_-^{-1} + y} \right),$$

where x and y are the firing rates of the pre- and post-synaptic neuron, respectively.

For the above expression, the fixed BCM threshold can be computed as:

$$\nu = -\frac{A_+/\tau_- + A_-/\tau_+}{A_+ + A_-},$$

which is positive when potentiation dominates depression on the short time scale, while, overall, synaptic weakening is larger than potentiation:

$$\begin{aligned} A_+ &> |A_-|, \\ |A_-|\tau_- &> A_+\tau_+. \end{aligned}$$

In some experiments we also consider the classical case of additive all-to-all STDP [Song *et al.* 2000], which acts as simple Hebbian learning, the induced change in weight being proportional to the product of the pre- and post-synaptic firing rates, see [Izhikevich & Desai 2003] for comparison of different STDP implementations. The parameters used in this case are: $A_+ = 8.33 \cdot 10^{-6}$, $A_- = -2.63 \cdot 10^{-6}$ and the same time constants as for the nearest neighbor case.

Additionally, when comparing the effects of IP and BCM, we use the minimal triplets STDP model [Pfister & Gerstner 2006], with an additional sliding threshold. The model is a generalization of STDP learning, which takes into account spike triplets. It is mathematically equivalent to the model in [Toyoizumi *et al.* 2005], i.e. BCM-like, but computationally simpler. Specifically, the model computes one presynaptic (r_1) and two postsynaptic (o_1 and o_2) activity traces, with different time scales:

$$\begin{aligned}\frac{dr_1}{dt} &= -\frac{r_1}{\tau_+} + \delta(t - t^f_{\text{pre}}) \\ \frac{do_1}{dt} &= -\frac{o_1}{\tau_-} + \delta(t - t^f_{\text{post}}) \\ \frac{do_2}{dt} &= -\frac{o_2}{\tau_{\text{y}}} + \delta(t - t^f_{\text{post}})\end{aligned}$$

where δ is the Dirac function and $t^f_{\text{pre/post}}$ is the time of firing of the pre- and post-synaptic neuron, respectively. The timescales of integration τ_+ and τ_- are similar to those in classical STDP ($\tau_+ = 16.8$ ms, $\tau_- = 33.7$ ms), while τ_{y} is slower ($\tau_{\text{y}} = 114$ ms).

The change in weight can be computed as:

$$\frac{dw}{dt} = A_3^+ \cdot r_1 \cdot o_2 \cdot y(t) - A_2^- \cdot o_1 \cdot x(t) \cdot \frac{\Theta}{\Theta_{\text{goal}}},$$

with the parameters $A_3^+ = 6.5 \cdot 10^{-4}$, $A_2^- = 7.1 \cdot 10^{-4}$ and $\Theta_{\text{goal}} = 25$ is a scaling parameter which implicitly defines the goal mean firing rate of the neuron. For implementing sliding threshold BCM, the original triplet model model is enhanced with a slowly varying threshold Θ, which estimates a low-pass filtered version of the square of the postsynaptic firing rate, similar to the classical BCM sliding threshold [Pfister & Gerstner 2006]. Namely:

$$\frac{d\Theta}{dt} = -\frac{\Theta}{\tau_{\text{th}}} + o_2^2,$$

with $\tau_{\text{th}} = 100$ s.

3.2.4 Synaptic scaling

Input weights are all positive and, as in approaches which directly maximize kurtosis or similar measures [Hyvärinen 1999, Hyvärinen 1997, Triesch 2007], the weight vector is normalized: $\sum_{\text{i}} w_{\text{i}} = w_{\text{tot}}$, with $w_{\text{tot}} = 2.5$. This value is arbitrary, as it represents a scaling factor of the total current, which can be compensated for by IP. It was selected in order to keep the final parameters close to those in [Toyoizumi *et al.* 2005]. For experiments using natural image patches, the normalization was done independently for the on- and off- populations, using the same value for w_{tot} in each case.

In a neural population, the same normalization is applied for the lateral inhibitory connections. As before, weights do not change sign and are constraint by the L_1 norm: $\sum_{\text{i}} w_{\text{i}}^{\text{inh}} = w_{\text{tot}}^{\text{inh}}$, with $w_{\text{tot}}^{\text{inh}} = -12$.

Currently, the normalization is achieved by dividing each weight by $\frac{\sum_{\text{i}} w_{\text{i}}}{w_{\text{tot}}}$, after the presentation of each sample. Biologically, this operation is implemented by a synaptic scaling mechanism, which multiplicatively scales the synaptic weights to preserve the average input drive received by the neuron [Turrigiano & Nelson 2004].

Experimental setup

In all experiments, excitatory weights were initialized at random from the uniform distribution and normalized as described before. The transfer function was initialized to the parameters in [Toyoizumi

et al. 2005] ($r_0 = 11$ Hz, $u_0 = -65$ mV, $u_\alpha = 2$ mV). Unless otherwise specified, all model parameters had the default values defined in the corresponding sections above. The details of each specific experiments are described in the corresponding sections below.

3.3 Results

3.3.1 Linear demixing of two supergaussian sources

To illustrate the basic mechanism behind our approach, we first ask if enforcing a sparse prior by IP and Hebbian learning can yield a valid ICA implementation for the classic problem of demixing two supergaussian independent sources. In the standard form of this problem, zero mean, unit variance inputs ensure that the covariance matrix is the identity, such that simple Hebbian learning with a linear unit (equivalent to principal component analysis) would not be able to exploit the input statistics and would just perform a random walk in the input space. This is, however, a purely mathematical formulation, and does not make much sense in the context of biological neurons. Inputs to real neurons are bounded and —in a rate-based encoding— all positive. Nonetheless, we chose this standard formulation to illustrate the basic underlying principles behind our model. Bellow, we will consider different spike-based encodings of the input and learning with STDP.

As a special case of a demixing problem, we use two independent Laplacian distributed inputs, with unit variance: $p_{U_i}(u_i) = \frac{1}{\sqrt{2}} e^{-\sqrt{2}|u_i|}$, $p(u_1, u_2) = p_{U_1}(u_1) \cdot p_{U_2}(u_2)$. For the linear superposition, we use a rotation matrix A:

$$A = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{pmatrix},$$

where α is the angle of rotation, resulting in a set of inputs $u' = Au$. Samples are drawn at each time step from the input distribution and are mapped into a total input to the neuron as $u_\mathrm{t} = w_1 u'_1 + w_2 u'_2$, with the weight vector w normalized). The neuron's transfer functions $g(u_\mathrm{t})$, the same as for our spiking model (Fig.3.3), is adapted based on our IP rule, to make the distribution of firing rates exponential ($\eta_\mathrm{IP} = 10^{-4}$). For simplicity, here weights change by classic Hebbian learning: $\Delta w_i = \eta u'_i g(u_\mathrm{t})$, with η being the synaptic learning rate ($\eta = 10^{-7}$). Each sample from the input distribution is presented for one time step.

In Fig. 3.4A we show the evolution of synaptic weights for different starting conditions. As our IP rule adapts the neuron parameters to make the output distribution sparse (Fig. 3.4B, C), the weight vector aligns itself along the direction of one of the sources. In this case, the weight normalization procedure can influence the final solution. Namely, positive weights with constant L_1 norm always yield a weight vector in the first quadrant, but this limitation can be removed by a different normalization, which keeps the L_2 norm of the vector constant ($|w|_{\mathrm{L}_2} = 1$). With this simple model, we are able to demix a linear combination of two independent sources for different mixing matrices and different weight constraints (Fig. 3.4D), as any other single-unit implementation of ICA.

3.3.2 One neuron learns an independent component

After showing that combining IP and synaptic learning can solve a classical formulation of ICA, we now focus on spike-based, biologically plausible inputs. In the following, STDP is used for implementing synaptic learning, while the IP and the synaptic scaling implementations remain the same. Additionally, samples are presented for a time interval $T = 100$ms (on the order of a fixation duration [Martinez-Conde *et al.* 2004]), followed by the weight normalization.

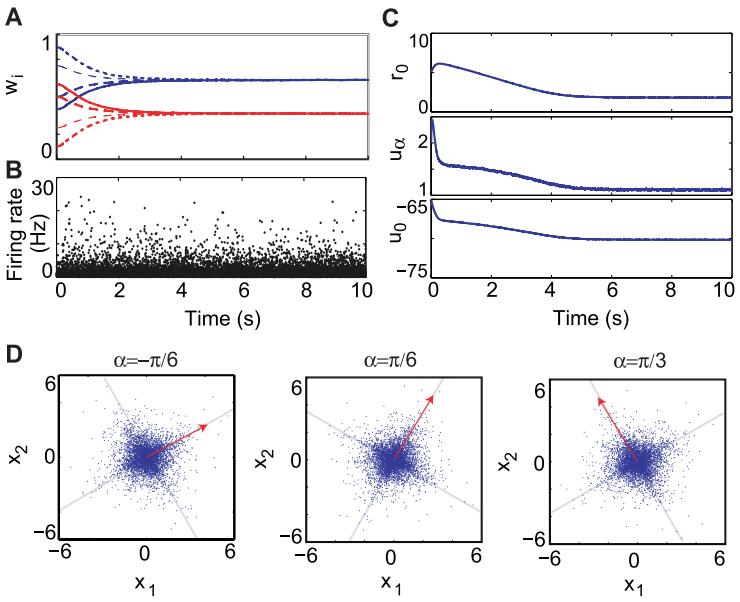

Figure 3.4. A demixing problem. (A) Evolution of the weights for different initial conditions, for the case where $\alpha = -\frac{\pi}{6}$, with L_1 weight normalization. (B) Corresponding evolution of the neuron's output and (C) transfer function parameters. (D) Final weight vector obtained when mixing two Laplace inputs by a rotation matrix with a rotation angle α. For the left example, normalization was done by $|w|_{\mathrm{L}_1} = 1$; for the others $|w|_{\mathrm{L}_2} = 1$ was used. In all cases the final vector was scaled by a factor of 5, to improve visibility.

Demixing with spikes

The demixing problem above can be solved also in a spike-based setting, after a few changes. First, the positive and negative inputs have to be separated into on- and off- channels ($u_i^{+/-}$) and converted into Poisson spike trains of duration T (see Methods), with the corresponding rates $u_i^{+/-}$ (note that the inputs are no longer white in this four-dimensional space). Secondly, to avoid the unbiological situation of having few very strong inputs, such that single presynaptic spikes always elicit a spike in the postsynaptic neuron, each channel consists of several (by default, 25) synapses, with independent inputs having the same firing rate. Synapses are all positive and adapt by STDP, under the normalization $\sum w_i = 1$. In order to speed up learning, we scale the firing rates of the input on the on- and off- channels by a factor of 20.

The evolution of the weights is shown in Fig. 3.5A. As in the original formulation of the problem, the receptive field of the neuron, obtained by projecting the final weights back onto the original two-dimensional space (see appendix A5), slowly aligns itself along one of the independent components (Fig. 3.5B). Specifically, after convergence the total weight of each channel can be estimated as the sum of individual weights corresponding to that input. The resulting four-dimensional weight vector is then projected back to the original two-dimensional input space using: $w = \max(\mathrm{w_{on}}, \mathrm{w_{off}})$.

Foldiák's bars: input encoded as firing rates

As a second test case for our model, we consider Foldiák's well-known bars problem [Foldiák 1990]. This is a classic non-linear ICA problem and is interesting from a biological perspective, as it mimics relevant

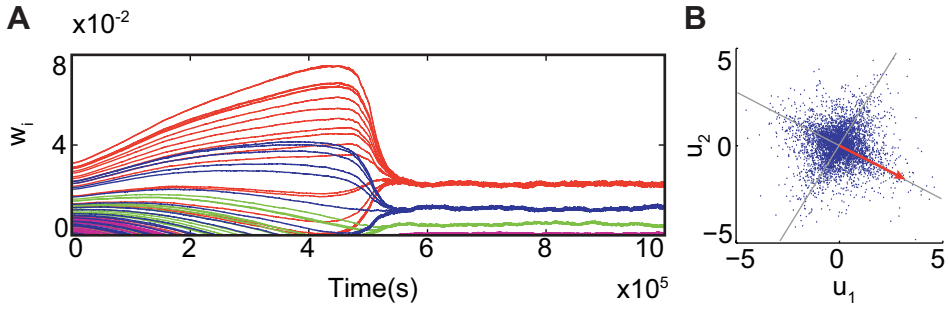

Figure 3.5. The demixing problem with inputs encoded as spike trains. (A) Evolution of the weights for a rotation angle $\alpha = \frac{\pi}{6}$; w_1^+ in red, w_1^- in magenta,w_2^+ in green, and w_1^- in blue. (B) Final corresponding weight vector in the original two-dimensional space. As before, the final weight vector is scaled by a factor of 5, to improve visibility.

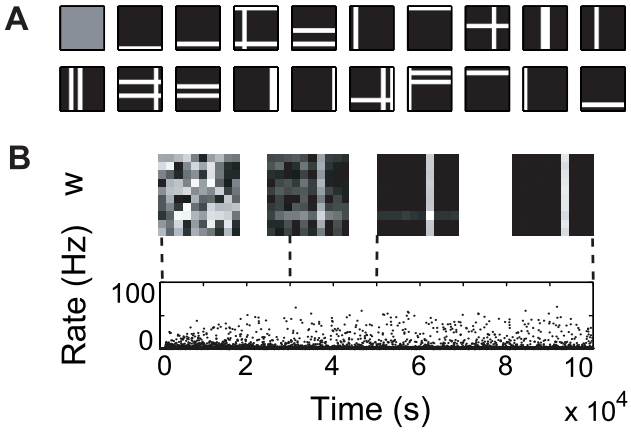

Figure 3.6. Learning a single independent component for the bars problem. (A) A set of randomly generated samples from the input distribution. (B) Evolution of the neuron's receptive field as the IP rule converges and instantaneous firing rate of the neuron, sampled every 10 s. Each dot corresponds to the firing rate $(g(u(t)) \cdot R(t))$ for a single sample, estimated in bins of 100 ms. (C) Time evolution of the parameters r_0, u_0, u_α of the neuron transfer function.

nonlinearities, e.g. occlusions. In the classic formulation, for a two-dimensional input x, of size $N \times N$, a single bar must be learned after observing samples consisting of the nonlinear superposition of $2N$ possible individual bars (see Fig. 3.6A), each appearing independently, with probability $\frac{1}{2N}$ $(N = 10)$. The superposition is non-linear, as the intersection of two bars has the same intensity as the other pixels in the bars (a binary OR).

In our implementation, the input vector is normalized to $|x| = N$, with $|\cdot|$ defining the L_1 norm, as in [Butko & Triesch 2007]. Additionally, the value of each pixel $x_{\mathrm{i,j}}$ is converted into a corresponding Poisson spike train with mean firing rates computed as $f_{\mathrm{bkgnd}} + x_{i,j} \cdot f_{\max}$, where $f_{\mathrm{bkgnd}} = 0.1$ Hz is the frequency of a background pixel and $f_{\max} = 100$ Hz gives the maximum input frequency, corresponding to a sample containing a single bar.

As the IP mechanism begins to take effect, making neuronal activity sparse, the receptive field of the

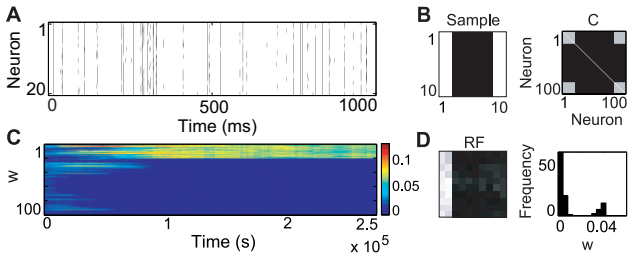

Figure 3.7. Bars in a correlation-based encoding. (A) A spike train with C=0.75. (B) Example of a sample containing two 2-pixel wide bars and the corresponding covariance matrix used for it's encoding. (B) Evolution of the weights in time. (C) Final receptive field and its histogram.

neuron slowly adapts to one of the bars (Fig. 3.6B). This effect is robust for a wide range of parameters (see appendix A2) and, as suggested by our previous results for rate neurons [Triesch 2007], does not critically depend on the particular implementation of the synaptic learning. We obtain similar results using additive [Song *et al.* 2000] and a simple triplet [Pfister & Gerstner 2006] STDP.

The input normalization makes a bar in an input sample containing a single IC stronger than a bar in a sample containing multiple ICs. This may suggest that a single component emerges by preferentially learning 'easy' examples, i.e. examples with a single bar. However, this is not the case and one bar is learned even when single bars never appear in the input, as in [Lücke & Sahani 2008]. Specifically, we use a variant of the bars problem, in which the input consists of always 4 distinct bars, selected at random. All model parameters are the same as before. In this case, the neuron correctly learns a single component. Moreover, similar results can be obtained for always 2-5 bars and for a wide range of parameters (see appendix A2).

Foldiák's bars: input encoded by spike-spike correlations

In the previous experiments, input information was encoded as firing rates. However, this stimulus encoding is not critical and the presence of spike-spike correlations is sufficient to learn a stable receptive field, even in absence of a presynaptic rate modulation. To demonstrate this, we considered a slight modification of the above bars problem, with input samples consisting always of two, 2 pixel wide, bars (with a full basis consisting of N bars). This is a slightly more difficult task, as wide bars emphasize the non-linearities due to bars overlap. However, our model can solve this problem when formulated using a rate encoding (see appendix A3).

More interestingly, a bar is learned when all inputs have the same mean firing rate ($f = 25$ Hz) and the information about whether a pixel belongs to a bar or not is encoded in correlation patterns between different inputs (Fig. 3.7). Specifically, background inputs were uncorrelated, while inputs belonging to bars were all correlated, with a correlation coefficient $C = 0.75$. Poisson processes with such correlation structure can be generated in a computationally efficient fashion by using dichotomous gaussian distributions [Macke *et al.* 2009]. As before, each sample was presented for a time interval T, followed by the weight normalization. Similar results were obtained for the version with always two, one pixel wide bars.

The fact that our approach works also for correlated inputs may be not too surprising, given the properties of STDP and IP. We know that, in the original rate formulation, strong inputs lead to higher firing in the postsynaptic neuron, causing the potentiation of the corresponding synapses. Similarly, if presynaptic inputs fire all with the same rate, the correlated inputs are more successful in driving the neuron and hence their weight is strengthened [Song *et al.* 2000]. In parallel, IP enforces sparse postsynaptic responses, guiding learning towards a heavy-tailed direction in the input (a more formal analysis

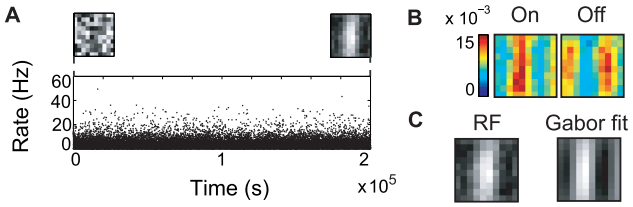

Figure 3.8. Learning a Gabor-like receptive field. (A) Evolution of the neuronal activity during learning. (B) Learned weights corresponding to the inputs from the on and off populations. (C) The receptive field learned by the neuron, and its l.m.s. Gabor fit.

of the interaction between IP and STDP follows below).

Stable receptive fields with a single 2 pixel wide bar can also be obtained for lower correlation coefficients ($C = 0.5$), but convergence slows down with lower C values. The stochastic nature of our neuron model makes it not particularly sensitive to input correlations, hence C cannot be too small. We would expect better results for a deterministic model, such as the leaky integrate-and-fire (LIF). Our previous experiments indicate that matching the first two moments of the output distribution to those of an exponential gives a reasonable IP approximation in this case.

As an additional limitation, STDP-induced competition [Song *et al.* 2000] enforces some constraints on the model parameters, to ensure the stability of a solution with multiple non-zero weights. This could be done by increasing the size of the input, restricting the mean overall firing of the neuron μ or by slightly changing the STDP parameters. These parameter changes do not affect learning with the rate-based encoding, however.

Natural scenes input

The last classical ICA problem we considered is the development of V1-like receptive fields for natural scenes [Simoncelli 2003]. Several computational studies have emphasized that simple cell receptive fields in V1 may be learned from the statistics of natural images by ICA or other similar component extraction algorithms [Bell & Sejnowski 1997, Olshausen & Field 1997, Hyvärinen 1999]. We hypothesized that the same type of computation could be achieved in our spiking neuron model, by combining different forms of plasticity. Only a rate encoding was used for this problem, partly for computational reasons and partly because it is not immediately obvious how a correlation-based encoding would look like in this case.

We used a set of images from the van Hateren database [van Hateren 1998], with standard preprocessing. The rectified values of the resulting image were mapped into a firing frequency for an on- and off-input population, as done for the bars. As before, each sample was presented for $T = 100$ ms. Weight normalization was done separately for the on- and off-population. The STDP, IP and other simulation parameters were the same as for the bars tests.

As preprocessing step, the images were convolved with a difference-of-gaussians filter with center and surround widths of 1.0 and 1.2 pixels, respectively. Random patches of size 10×10 were selected from various positions in the images. Patches having very low contrast were discarded. The individual input patches were normalized to zero mean and unit variance, similar to the processing in [Butko & Triesch 2007]. The rectified values of the resulting image were mapped into a firing frequency for an on- and off-input population ($f_{i,j}^{\mathrm{On/Off}} = f_{\mathrm{bkgnd}} + x_{i,j}^{\mathrm{On/Off}} \cdot f_{\mathrm{max}}$).

As shown in Fig. 3.8A, the receptive field of the neuron computed as the difference between the weights of the on- and the off- input populations (depicted in Fig. 3.8B) evolved to an oriented filter, similar to those obtained by other ICA learning procedures [Bell & Sejnowski 1997, Olshausen & Field 1997, Hyvärinen 1999, Falconbridge *et al.* 2006, Weber & Triesch 2008]. A similar receptive field can be obtained

by reverse correlation from white noise stimuli. The least-mean-square-error fit to a Gabor wavelet [Lücke 2009] is shown in Fig. 3.8C. As vertical edges are usually over-represented in the input, the neuron will typically learn a vertical edge filter, with a phase shift depending on the initial conditions. Although the receptive field obtained in this case has low spatial frequency, more localized receptive fields can be obtained in a neural population (see below).

3.3.3 ICA in a neuron population

So far, learning has been restricted to a single neuron. For learning multiple independent components, we implemented a neuron population in which the activities of different neurons were decorrelated by adaptive lateral inhibition. Our approach is similar to that of other feature extraction methods based on single-unit contrast functions [Hyvärinen 1999]. Namely, we considered a simple scheme for parallel (symmetrical) decorrelation, based on lateral inhibition (Fig. 3.9A). The all-to-all inhibitory weights changed based on the same STDP rule used for learning the forward connections. Also, similar to the excitatory connections, the inhibitory synapses are subject to synaptic scaling.

The input-related parameters used here are the same as for the experiments before, but $\mu = 5$Hz, to speed up learning. As done for the excitatory weights, the inhibitory weights are initialized at random, with no self-connections, and normalized.The initial parameters of the neuron transfer function are uniformly distributed around the default values mentioned above, with variance 0.1, 5, and 0.2 for r_0, u_0, and u_α, respectively.

We first investigated the bars problem. Unfortunately, computational overhead significantly limits the size of networks we can simulate. For the same reason, we only use rate-based encoding, which has faster convergence in single neurons. We show the response of a population of 10 neurons in Fig. 3.9B. The neurons learn distinct receptive fields, each consisting of one independent component. As lateral inhibition begins to take effect, the pairwise correlation between the responses of different neurons in the population decreases to zero (see the average over all pairs in Fig.3.9C), making the final inhibitory weights unspecific (Fig.3.9D). As decorrelation is not a sufficient condition for independence, we show that, simultaneously to the decrease in correlation, the normalized mutual information, computed as $\mathrm{MI}^*(X,Y) = \frac{\mathrm{MI}(X,Y)}{H(X)+H(Y)}$, with $\mathrm{MI}(\cdot)$ denoting the mutual information, and $H(\cdot)$ the entropy, see [Butko & Triesch 2007], between the units is reduced (Fig. 3.9C).

For the natural scenes problem, we can only consider the undercomplete case, due to computational limitations. In this case we obtain oriented, localized receptive fields (Fig. 3.9E).

Due to the adaptive nature of the IP rule, the balance between excitation and inhibition does not need to be tightly controlled, allowing for great robustness to changes in scaling parameters. However, the inhibition strength may influence the time required for IP convergence (the stronger the inhibition, the longer it takes for the system to reach a stable state). A more important constraint we find is that the adaptation of inhibitory connections needs to be faster than that of feedforward connections to allow for efficient decorrelation (see Methods for parameters).

3.3.4 Is IP necessary for learning?

Given the results presented, one must wonder what the role of IP is in this learning procedure. Does IP simply find an optimal nonlinearity for the neuron's transfer function, given the input, something that could be computed offline, or is the interaction between IP and STDP critical for learning? To answer this question, we go back to the first experiment using Foldiák's bars. We repeat the first experiment for a fixed gain function, given by the final parameters obtained after learning ($r_0 = 23.8$, $u_0 = -66.4$, $u_\alpha = 1.1$). In this case, the receptive field of the neuron does not evolve to an IC (Fig. 3.10), which suggests that ICA-like computation relies on the interplay between weight changes and the corresponding readjustment of neuronal excitability, which forces the output to be sparse. Note that this result holds for simulation times significantly larger than in the experiment before, where a bar emerged after $5 \cdot 10^4$ s,

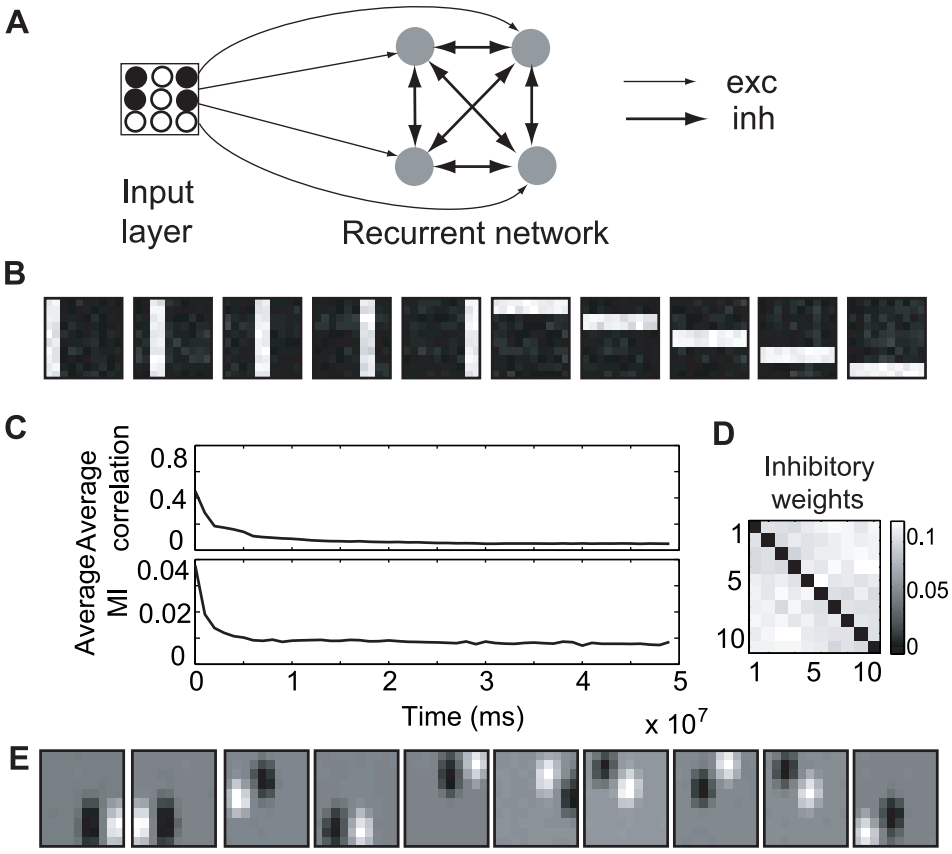

Figure 3.9. Learning multiple ICs. (A) Overview of the network structure. All neurons receive signals from the input layer and are recurrently connected by all-to-all inhibitory synapses. (B) A set of receptive fields learned by the neuron population for the bars problem. (C) Evolution of the mean correlation coefficient and mutual information in time, both computed by dividing the neuron output in bins of width 1000 s and estimating C and MI^* for each bin. (D) Learned inhibitory lateral connections. (E) A set of receptive fields learned for natural image patches.

suggesting that, even if the neuron would eventually learn a bar, it would take orders of magnitude longer to do so.

It may be assumed that the failure to learn a bar for the learned transfer function is simply owed to the low postsynaptic firing observed in this case. Hence, it may be the case that a simpler rule, which tries to regulate just the mean firing rate of the neuron would suffice to learn an IC. We tested this hypothesis by constructing an alternative IP rule, which regulates a single parameter (r_0) to preserve the average firing rate of the neuron. In this case, for the same experiment parameters as before, no bar is learned and the output distribution is gaussian, with a small standard deviation around the goal average value μ (Fig. 3.11A, upper row). However, it can happen that a bar is learned, provided that the input distribution has sufficient variance to start the learning process (variance was regulated by the parameter w_{tot}), as shown in Fig. 3.11B, lower row. When this occurs, the final output distribution is highly kurtotic, due to the receptive field. The dependence on model parameters in this case suggests that regulating the first moment is not sufficient for reliably learning a bar and higher order moments

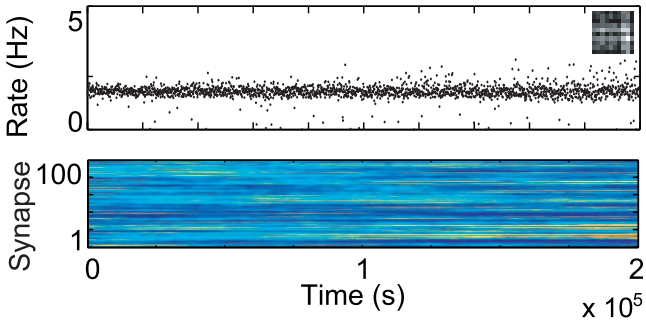

Figure 3.10. IP is critical for learning. Evolution of (A) the activity and (B) receptive field for a neuron with a fixed gain function, given by the final parameters obtained after learning in the previous bars experiment. A bar cannot be learned in this case.

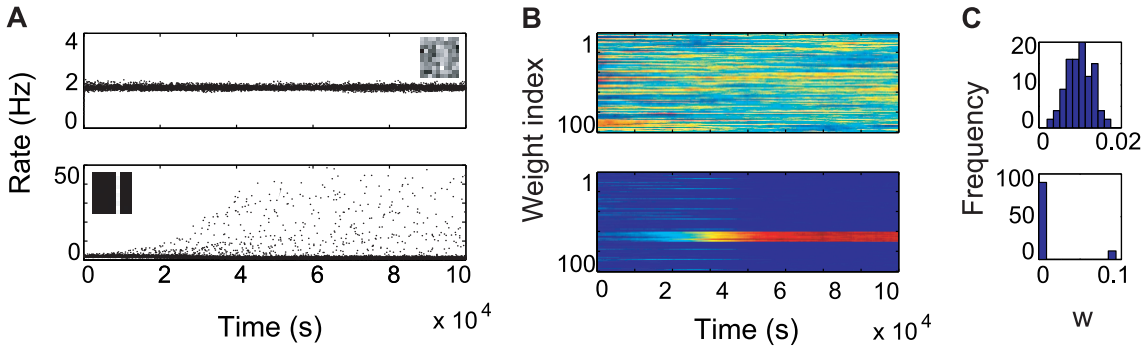

Figure 3.11. Mean firing constraint is not sufficient for reliable learning. (A) Evolution of neuron activation for a neuron with a gain function regulated by a simplified IP rule, which adjusts r_0 to maintain the same mean average firing μ. $w_{\text{tot}} = 2.5$ or $w_{\text{tot}} = 10$, in the first and second row, respectively. Inset illustrates final receptive field for each case. (B) Corresponding evolution of weights and (C) their final distribution.

need to be considered as well.

3.3.5 Interaction between IP and STDP

For elucidating the mechanism by which the interaction between STDP and IP may facilitate the discovery of an independent component, we go back to the simplified problem of a single unit receiving a two dimensional input. This analysis is restricted to the theoretical formulation of zero mean, unit variance inputs and does not consider the time of individual spikes. Moreover, for mathematical convenience, the rotational angle α is zero in this case. Previously, we have shown in simulations that for a bounded, whitened, two dimensional input the weight vector tends to rotate towards the heavy-tailed direction in the input [Triesch 2007]. Here, we show analytically that the same result holds for the zero mean, unit variance uniform and the standard Cauchy distribution. The two inputs are independent ($p(u_1, u_2) =$

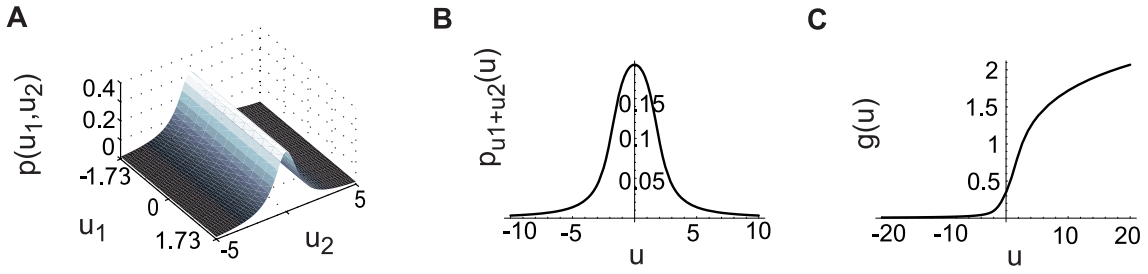

Figure 3.12. The simplified case of a two dimensional input space. (A) The joint probability distribution of the input. (B) Distribution of the membrane potential, given the total input for a weight vector along the first diagonal. (C) Corresponding optimal neuronal gain function.

$p_{U_1}(u_1) \cdot p_{U_2}(u_2)$), with the PDFs defined by:

$$\begin{aligned} p_{U_1}(u_1) &= \begin{cases} \frac{1}{2\sqrt{3}}, & \text{for} |u_1| \leq \sqrt{3} \\ 0, & \text{otherwise} \end{cases}, \\ p_{U_2}(u_2) &= \frac{1}{\pi(1+u_2^2)}. \end{aligned}$$

As before, a linear relation is assumed between the input and the induced changes of voltage in the postsynaptic neuron and the IP adaptation is taken to be much faster than synaptic plasticity [Triesch 2007].

Under the above assumptions, given the two inputs and the corresponding weight vector (w_1, w_2), the probability distribution for the total input u of the neuron ($u = u_1 + u_2$) can be computed as a convolution of the probability distributions of u_1 and u_2. For simplicity, we only consider here a weight vector along the first diagonal, but the analysis could be extended to any other weight vector. For this case, the result of the convolution is:

$$p_{U=U1+U2}(u) = p_{U_1}(u_1) * p_{U_2}(u_2) = \frac{1}{2\sqrt{3}\pi} \left(\arctan(u_+) - \arctan(u_-)\right),$$

where * marks the convolution operator and $u_{+/-} = u \pm \sqrt{3}$.

Under the assumption of a fast IP, the neuron's gain becomes close to the optimal function $\tilde{g}$, while the weights remain approximately unchanged. Knowing the input distribution, $\tilde{g}$ can be computed analytically. Specifically, we exploit the standard relation $p_G(g) = \left(\frac{\mathrm{d}g}{\mathrm{d}u}\right)^{-1} p_U(u)$ and the constraint that the output distribution should be exponential. By integrating the above equation, with p_U as computed before and $p_G(g) = \frac{1}{\mu} e^{-\frac{g}{\mu}}$, we obtain the optimal transfer function (Fig. 3.12):

$$\tilde{g} = -\mu \log\left(\frac{1}{2} + \frac{1}{2\sqrt{3}\pi}\left(u_- \arctan(u_-) - u_+ \arctan(u_+) + \frac{1}{2}\log\left(\frac{1+u_+^2}{1+u_-^2}\right)\right)\right).$$

Of course, given the family of functions considered for modeling g, the gradient based IP rule will approximate this solution more or less well.

Knowing $\tilde{g}$, the estimated change in weights by classic Hebbian learning, under the assumption that

$E(R(t)) \approx 1$, can be computed as:

$$\begin{aligned} \Delta w_1 &= E(u_1 g) = \int x\tilde{g}(y) p_{U_1}(x) p_{U_2}(y-x) \mathrm{d}x\mathrm{d}y, \\ \Delta w_2 &= E(u_2 g) = \int (y-x)\tilde{g}(y) p_{U_1}(x) p_{U_2}(y-x) \mathrm{d}x\mathrm{d}y. \end{aligned}$$

Integrating the above relations, we obtain that $\Delta w_1 < \Delta w_2$, meaning that the learning procedure rotates the weight vector towards the heavy-tailed direction in the input. Similar results can be obtained numerically for our STDP rule (see Methods). Intuitively, the largest changes in weight occur for large neuron responses, which corresponds to samples from the right tail of the input distribution. As the uniform distribution has only finite support, here this tail is determined by the other input. More generally, we can show that significant changes in weights occur only on the tail of the output distribution and that they are much larger for the heavy-tailed directions (see appendix A3).These results apply to any IP rule that enforces an exponential output distribution. In practice, implementations corresponding to different families of transfer functions may work more or less well, depending on the specific problem.

The experiment before can be repeated in simulations with different pairs of input distributions. In appendix A3, we demonstrate that the heavy-tailed direction is always selected, regardless of whether or not one of the distributions is bounded, that the effect is not specific to exponential directions in the input and, more importantly, that the weights change as expected from a rule that tries to maximize the kurtosis of the output. Most importantly, we show that for simple problems when a solution can be obtained by nonlinear PCA our IP rule significantly speeds up learning of an independent component.

One way to understand these results could be in terms of nonlinear PCA theory. Given that for a random initial weight vector, the total input distribution is close to Gaussian, in order to enforce a sparse output, the IP has to change the transfer function in a way that 'hides' most of the input distribution (for example by shifting u_0 somewhere above the mean of the Gaussian). As a result, the nonlinear part of the transfer function will cover the 'visible' part of the input distribution, facilitating the discovery of sparse inputs by a mechanism similar to nonlinear PCA. In this light, IP provides the means to adapt the transfer function in a way that makes the nonlinear PCA particularly efficient.

Lastly, from an information-theoretic perspective, our approach can be linked to previous work on maximizing information transmission between neuronal input and output by optimizing synaptic learning [Toyoizumi *et al.* 2005]. This synaptic optimization procedure was shown to yield a generalization of the classic BCM rule [Bienenstock *et al.* 1982]. We can show that, for a specific family of STDP implementations, which have a quadratic dependence on postsynaptic firing, IP effectively acts as a sliding threshold for BCM learning (see appendix A4).

3.4 Discussion

Although ICA and related sparse coding models have been very successful in describing sensory coding in the cortex, it has been unclear how such computations can be realized in networks of spiking neurons in a biologically plausible fashion. We have presented a network of stochastically spiking neurons that performs ICA-like learning by combining different forms of plasticity. Although this is not the only attempt at computing ICA with spiking neurons, in previous models synaptic changes were not local, depending on the activity of neighboring neurons within a population [Klampfl *et al.* 2009, Parra *et al.* 2009]. In this light, our model is, to our knowledge, the first to offer a mechanistic explanation of how ICA-like computation could arise by biologically plausible learning mechanisms.

In our model, IP, STDP and synaptic scaling interact to give rise to robust receptive field development. This effect does not depend on a particular implementation of STDP, but it does require an IP mechanism which enforces a sparse output distribution. Although there are very good theoretical arguments why this should be the case [Stemmler & Koch 1999, Lennie 2003, Triesch 2007], the experimental evidence

supporting this assumption is limited [Baddeley *et al.* 1997]. A likely explanation for this situation is the fact that it is difficult to map the experimentally observable output spikes into a probability of firing. Spike count estimates cannot be used directly, as they critically depend on the bin size. Additionally, the inter-spike interval (ISI) of an inhomogeneous Poisson process with exponentially distributed mean $\lambda(t)$ is indistinguishable from the ISI of a homogeneous Poisson distribution with mean $\lambda = E(\lambda(t))$. Hence, more complex statistical analyses are required for disentangling the two, see [Cox 1955]. Nonetheless, from a functional perspective, V1 receptive fields are known to adapt to the statistics of visual inputs as to increase the information carried by the neural response about the input stimulus [Sharpee *et al.* 2006]. Given that this adaptation occurs on a timescale of tens of seconds, it is interesting to hypothesize that IP could play a role in this type of reorganization.

From a computational perspective, our approach is reminiscent of several by-now classic ICA algorithms. As mentioned before, IP enforces the output distribution to be heavy-tailed, like in sparse coding [Olshausen & Field 1997]. Our model also shares conceptual similarities to InfoMax [Bell & Sejnowski 1997], which attempts to maximize output entropy (however, at the population level) by regulating the weights and a neuron threshold parameter. Maximizing information transmission between pre- and post-synaptic spike trains under the constraint of a fixed mean postsynaptic firing rate links our method to previous work on synaptic plasticity. A spike-based synaptic rule optimizing the above criterion [Toyoizumi *et al.* 2005] yields a generalization of the BCM rule [Bienenstock *et al.* 1982], a powerful form of learning, which is able to discover heavy-tailed directions in the input [Intrator & Cooper 1992,Intrator 1998] and to learn Gabor receptive fields [Blais *et al.* 1998] in linear neurons. We have shown that, for some specific types of synaptic plasticity, our IP mechanism effectively acts as a sliding threshold in BCM. However, our rule is more general, as it does not require a specific shape for the synaptic learning.

It is interesting to think of the mechanism presented here in relation to projection pursuit [Huber 1985], which tries to find good representations of high-dimensional spaces by projecting data on a lower dimensional space. The algorithm searches for interesting projection directions, a typical measure of interest being the nongaussianity of the distribution of data in the lower dimensional space. The difference here is that, although we do not explicitly define a contrast function maximizing kurtosis or other similar measure, our IP rule implicitly yields highly kurtotic output distributions. By sparsifying the neuron output, IP guides the synaptic learning towards the interesting (i.e. heavy-tailed) directions in the input.

From a different perspective, we can relate our method to nonlinear PCA. It is known that, for zero mean whitened data, nonlinear Hebbian learning in a rate neuron can successfully capture higher order correlations in the input [Hyvärinen 1997, Oja 1997]. Moreover, it has been suggested that the precise shape of the Hebbian nonlinearity can be used for optimization purposes, for example for incorporating prior knowledge about the sources' distribution [Hyvärinen 1997]. IP goes one step further in this direction, by adapting the transfer function online, during learning. From a biological perspective, there are some advantages in adapting the neuron's excitability during computation. In this way, the system gains greater robustness to changes in input levels. Additionally, IP regulation plays a homeostatic role, making constraints on the input mean or second order statistics unnecessary. In the end, all the methods we have mentioned are closely related and, though conceptually similar, our approach is another distinct solution.

Previous work on IP-based receptive field development has been restricted to a rate model neuron [Triesch 2007,Butko & Triesch 2007,Weber & Triesch 2008]. Beyond translating these results to a spiking neuron model, we have shown here that similar principles can be applied when information is encoded as spike-spike correlations, where a model relying just on firing rates would fail. It is an interesting challenge for future work to investigate to which extent the mechanism presented here generalize to other types of input encoding. As a first step, the following chapter investigates how different encodings could emerge as a result of reward dependent learning, for the task of short-term memory.

4

Working memory development by reward-dependent STDP

Working memory (WM), defined as a temporary storage of stimulus-specific information, is critical for various cognitive processes. It is believed that WM has as neural correlate a selective persistent activation of neurons in several cortical areas [Miyashita 1988]. Additionally, this activity can be time-varying [Brody *et al.* 2003b]. Neuronal firing rates can change systematically during the delay period [Brody *et al.* 2003a, Pastalkova *et al.* 2008] and this time-varying neural response is thought to convey as much information about the initial stimulus as persistent firing [Baeg *et al.* 2003].

Inspired by abstract models such as the Hopfield network [Hopfield 1982], WM has been traditionally modeled by recurrent networks with attractor dynamics, in which a non-selective background activity attractor coexists with multiple stimulus-specific attractors [Amit & Brunel 1997, Brunel & Wang 2001, Mongillo *et al.* 2008]. These, however cannot account for the reported time-varying representations of stimuli. Alternatively, it has been suggested that short-term memory may rely on 'feed-forward' structures within recurrent networks, which propagate activity through a chain of transiently activated network states. In this case, a persistent stimulus-specific response can be obtained by an appropriate linear combination of a set of temporal filters with different time constants [Goldman 2009]. Regardless of the model, the question how such representations could emerge as result of a learning process remains largely unanswered. This is particularly problematic, as prefrontal neural responses can be considerably shaped by experience [Baeg *et al.* 2003, Rainer & Miller 2000, Kennerley & Wallis 2009]. In a few sessions of training, the encoding of novel stimuli in prefrontal neurons can change significantly, in parallel to behavioral learning, allowing for better prediction of both encoded stimuli and actions [Baeg *et al.* 2003].

Moreover, from a developmental perspective, WM is known to improve considerably, in both capacity and maintenance time during childhood and adolescence [Luciana & Nelson 1998]. This improvement can be attributed to a variety of factors, from the maturation of cortical tissue, to the development of better strategies to encode and store information, e.g. chunking [Gathercole 1999]. However, part of this improvement is also be owed to learning.

Here, we ask if WM can develop in generic recurrent neural networks by reward-dependent plasticity while learning to perform tasks that require some temporary storage of information. Specifically, we ask if an initially unstructured neural network can acquire WM properties while learning to perform a delayed response task. Unlike classical WM models, we make no *a priori* assumptions on stimulus encoding, leaving the network to discover an appropriate representation by reward-dependent learning.

A critical problem when trying to optimize recurrent networks is how to stabilize the dynamics during learning. This is particularly difficult to achieve in recurrent networks, as typically such networks require a fine tuning of gain parameters, which cannot be achieved while learning. As for the problem of unsupervised learning in the previous chapter, our solution involves several synaptic and neuronal homeostatic mechanisms. Specifically, we combine the reward-modulated STDP at excitatory synapses with synaptic scaling and IP controlling the excitability of both excitatory and inhibitory neurons (see below). The

homeostatic mechanisms at excitatory neurons stabilize the network during the task-dependent learning, while the adaptive inhibition makes sure that the representation space is split evenly between different stimuli. Additionally, activity-dependent structural plasticity is used for optimizing the sparse cortical connectivity to best encode information.

In the following, we give a brief introduction of the most important WM experiments and the neural models used to explain them, highlighting some of the limitations of the existing models in terms of WM dynamics and learning. We continue by a description of our recurrent network and the different plasticity mechanisms shaping its connectivity. We show that, after learning, the network is able to perform our WM task. Moreover, the representations that emerge are similar to those observed experimentally, with neurons exhibiting stimulus and time specific activation. Lastly, the performance of our model is significantly better than that of a randomly initialized recurrent network, suggesting that reward-modulated STDP can be used for optimizing network performance in a task-dependent manner. Some preliminary results of this work have been published as [Savin & Triesch 2009]

4.1 Background

4.1.1 Working memory development in children

One of the most important factors influencing working memory performance is age [Luciana & Nelson 1998]. The WM function improves significantly until adolescence, with an approximately 4-fold increase between the ages of 4 and 14, see Fig. 4.1. Typically, WM performance increases steeply for infants that are 8 years old or smaller, with a more gradual increase to adult performance, reached —depending on the task and measure— between the ages of 11 and 16 [Gathercole 1999].

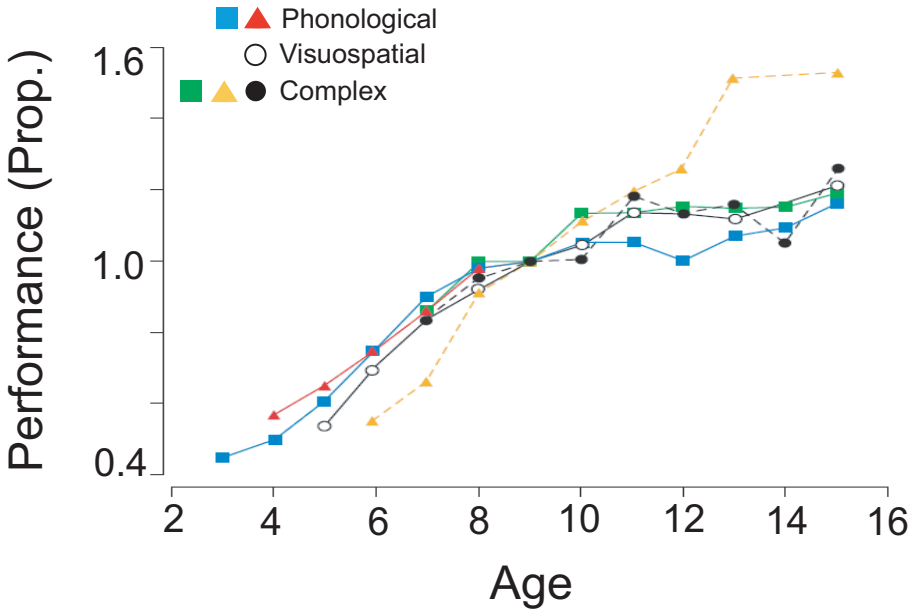

Figure 4.1. Working memory performance as function of age. Mean performance for each age group is plotted as proportion relative to the 9-years old group. WM tasks: blue squares – digit span, red triangles – non-word repetition, open circles – forward digit span, green squares – Corsi blocks, yellow triangles – listening span, filled circles – backward digit span. Adapted from [Gathercole 1999].

It is not completely understood what are the sources of these age-related changes in WM performance. They could reflect the maturation of the underlying cortical architecture, an increase in processing efficiency of a range of processes involved in encoding, storage and retrieval of memory items, the development of attention, or the usage of memorization strategies such as rehearsal and chunking [Gathercole 1999]. In the context of our work, we can assume that the cortical substrate underlying WM can be optimized by reward-dependent learning to improve both WM capacity and maintenance time, similar to the plastic changes observed in animal models [Rainer & Miller 2000, Kennerley & Wallis 2009] described below.

4.1.2 Working memory: experimental evidence

Working memory experiments have demonstrated that the firing rate of prefrontal cortex (PFC) neurons is selectively enhanced during the delay period [Miyashita 1988, Goldman-Rakic *et al.* 1990, Goldman-Rakic 1990]. This persistent neural activation is correlated with the previously presented cue, the forthcoming action or a more complex function of the two, see [Durstewitz *et al.* 2000b], and its disruption (e. g. by electrical stimulation, or due to distractors) leads to a decay in performance [Funahashi *et al.* 1989]. Hence, it is generally believed that this selective, persistent activation of prefrontal neurons represents the neural correlate of WM [Goldman-Rakic 1995, Durstewitz *et al.* 2000b].

Although most WM experiments focus on prefrontal cortex, the same type of stimulus specific activation can also be observed in various other brain areas, including parietal cortex, IT, hippocampus and motor areas [Watanabe & Niki 1985, McFarland & Fuchs 1992, Miller *et al.* 1993, Constantinidis & Steinmetz 1996, Quintana & Fuster 1999]. Nonetheless, delay activity in PFC is the most prominent and most robust to distractor stimuli [Durstewitz *et al.* 2000b].

Time-dependent encoding of stimuli

An increasing body of experimental evidence suggests that the firing rates of persistently active neurons in several areas can change systematically during the delay period [Brody *et al.* 2003b,Pastalkova *et al.* 2008]. For example, most PFC neurons engaged in a delayed match-to-sample task (frequency of mechanical vibration applied to fingertip, in monkey) showed a response profile that was time-varying. Only about 25% of the recorded neurons displayed persistent activation (approximally constant high firing rate), whereas the rest displayed ramping effects ('late' neurons), or fired predominantly at the beginning of the delay period ('early' neurons) [Brody *et al.* 2003a]. In this way, the neuron population could store information not only about the stimulus, but also about time passed since the stimulus presentation [Jin *et al.* 2009].

Several researchers have reported similar time-dependent responses, such that now such patterns seem ubiquitous for delay period responses in various cortical areas [Miller *et al.* 1996, Chafee & Goldman-Rakic 1998, Pesaran *et al.* 2002, Rainer & Miller 2002]. Importantly, these time-varying responses are as predictive of the animal's response as the population of persistent neurons [Baeg *et al.* 2003].

Plasticity of working memory

Classical WM experiments are performed in animals having extensive training with the task. Recent experiments in naïve animals reveal that response of PFC neurons to stimuli changes significantly during training. In parallel to the increase in performance, after learning, the prefrontal neurons activation become sparser, individual neurons develop a narrow tuning, and the representation becomes more robust to input noise [Rainer & Miller 2000]. Moreover, the functional connectivity between different prefrontal neurons (defined by pairwise correlations) also changes during learning, being maximal in the early phases of learning, the time when the most learning progress is observed [Baeg *et al.* 2007]. Although the mechanisms behind these changes are not well understood, it is likely that the observed effects are at least partially owed to learning at prefrontal neurons.

The optimization of the ensemble representation of stimuli during training has been reported in several other experiments [Baeg *et al.* 2003] and it is likely to involve reward-dependent learning. Experiments show that, during a spatial WM task, the magnitude of the expected reward determines the spatial selectivity of prefrontal neurons and, correspondingly, the behavioral performance of the animals [Kennerley & Wallis 2009]. This suggests that the expected reward modulates the representation of information about stimuli to ensure an efficient allocation of resources.

4.1.3 Models of working memory

Since the first experiments on the neural correlates of WM [Miyashita 1988], a variety of computational models of WM have been developed, at different levels of abstraction. The most important classes include recurrent excitatory networks with attractor dynamics, feedforward-connected subgroups which propagate activity as synfire chains, or cell-specific bistable mechanisms, see [Durstewitz *et al.* 2000b]. They were used to model different aspects of WM including the neuronal activity profile during the delay period or the role of neuromodulation in WM tasks [Durstewitz *et al.* 1999].

Attractor dynamics

The idea that persistent activity can be obtained through recurrent excitation has been ellaborated by Hopfield for binary threshold units [Hopfield 1982]. In this framework, several discrete symbols can be stored in the synaptic weight matrix of the excitatory connections and retrieved as point attractors of the network dynamics. As this type of network has the ability to retrieve a stored pattern given partial or noise-corrupted inputs, it reproduces the type of similarity-based retrieval characteristic of human memory. Moreover, due to its simplicity, it is possible to analytically investigate the robustness and the memory capacity of the model. However, due to its high level of abstraction, Hopfield networks cannot be mapped directly to biology.

A series of subsequent models have tried to bring the Hopfield model closer to a biologically-plausible implementation. First one can separate the network into excitatory and inhibitory subpopulations, with positive (rate) activation [Amit & Brunel 1995,Amit & Brunel 1997]. This model can further be extended to spiking neurons [Compte *et al.* 2000]. In all cases, stimuli are encoded by distinct cell assemblies of strongly connected neurons which are weakly connected with the rest of the network, possibly with an additional cell assembly corresponding to the baseline activation of the network [Amit & Brunel 1997]. This type of connectivity could be learned by Hebbian synaptic plasticity [Amit & Brunel 1997].

Under the aforementioned constraints on the weight matrix and assuming all-to-all connectivity, a mean-field analysis can be derived and the stability of the network can be investigated as function a of the input to the network and the average synaptic gain of excitatory synapses, see Fig. 4.2. In the absence of input, the network resides in the baseline fixed-point, with small non-zero firing rate of the population. When a stimulus is presented to the network, the activity increases sufficiently for the recurrent excitation to switch the network to the corresponding attractor, where it will remain even after the input is removed. Activating the inhibitory subpopulation can destabilize this attractor, bringing the network back to baseline, see [Compte *et al.* 2000].

As an important caveat, AMPAR dynamics are too fast to account for the robust persistent activity at 15-30 Hz observed experimentally and slower NMDAR currents are required for robust persistent dynamics in the physiological range for neuronal firing rates. Moreover, the ratio between AMPAR and NMDAR is important for determining the robustness of the network to noise and distractors [Compte *et al.* 2000]. Interestingly, the NMDAR/AMPAR ratio in prefrontal neurons is the largest in the cortex [Scherzer *et al.* 1998], which could explain why PFC is particularly effective in sustaining neural activity during the delay.

An elegant extension of the models above has been recently introduced in [Mongillo *et al.* 2008]. Here, complementary to classic attractor dynamics, information can be stored in the short-term dynamics of

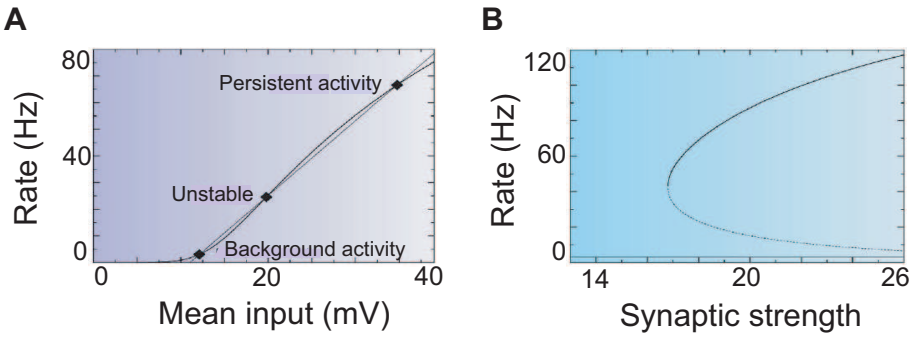

Figure 4.2. Attractor dynamics in a network of spiking neurons. (A) I-f curve showing the mean firing rate for a population of leaky IF neurons, with low reset, connected by excitatory synapses. (B) Bifurcation diagram showing the population firing rate as function of synaptic strength. Adapted from [Barbieri & Brunel 2008].

excitatory synapses. We have mentioned before that in PFC, unlike other brain areas, excitatory synapses have predominantly facilitatory dynamics. Given that the time scale of this process is of the order of hundreds of milliseconds, a trace of the stimulus can persist in the form of a Ca^{2+} transient even after the activity of the network has seemingly returned to baseline. The information can be then retrieved by a brief unspecific input to the network (this signal could also be intrinsically generated). The metabolic cost of such an implementation is significantly lower than for persistent neural activation. Moreover, it allows for a gating of information, such that a stored stimulus can be recalled on request, and it makes it significantly easier to store several patterns in memory at the same time, see [Mongillo *et al.* 2008].

Neuromodulators, in particular dopamine (DA), are known to affect the functioning of PFC during WM tasks. The DA production increases during WM tasks [Schultz *et al.* 1993, Watanabe 1996] and is essential for WM performance (via the activation of D1 receptors) [Sawaguchi & Goldman-Rakic 1994, Durstewitz *et al.* 1999, Durstewitz *et al.* 2000a, Durstewitz & Seamans 2002]. These effects are believed to reflect the DA-dependent modulation of Na^+ and NMDAR and $GABA_A$ conductances reported experimentally [Seamans *et al.* 2001a, Gorelova & Yang 2000, Yang & Seamans 1996, Seamans & Yang 2004]. These conductances affect the network activity through nonlinear interactions, such that D1 activation strengthens the currently activated cell assembly, while suppressing the activity of other attractors (in the phase plane, this can be viewed as a deepening of the basin of attraction corresponding to the active stimulus) [Durstewitz *et al.* 1999, Brunel & Wang 2001, Durstewitz & Seamans 2002]. In this way, the WM activation is focused towards behaviorally relevant inputs.

So far, the discussion has been restricted to the problem of storing a discrete set of stimuli, what is known in the literature as object working memory. The same principle can be generalized for storing continuous values, by using line attractors. This is a particular class of systems, in which the dynamics are stable to perturbations that push the system away from this line (of surface), but not to those along this line (or surface). Such models have been considsered as an implementation of spatial working memory [Seung *et al.* 2000, Compte *et al.* 2000, Miller *et al.* 2003]. However, the existing models tend to require extremely careful tuning of synapses and robust models implementing a continuum of persistent states are still missing.

Alternatives

The classic framework of attractor dynamics cannot account for time-varying representations of stimuli found experimentally. An alternative way to sustain activity locally is through the feedforward propa-

gation of activity through a chain of transiently activated network states. Such synfire chains [Aertsen *et al.* 1991, Aertsen *et al.* 2001] could explain the temporal dynamics observed during the delay period in motor cortex, see [Diesmann *et al.* 1999]. Although it is possible to implement synfire chains that are robust and noise-tolerant [Diesmann *et al.* 1999, Gewaltig *et al.* 2001, Aertsen *et al.* 2001], due to the fact that this solution relies on precise spike timings, it is believed to be more sensitive to noise and GABAergic inputs compared to attractor dynamics [Durstewitz *et al.* 1999].

Recently, the idea of a short-term memory which relies on feedforward structures within recurrent networks has been reevaluated in [Goldman 2009]. Specifically, the model assumes neurons to act as a set of linear low-pass filters, with different time-constants. In this scheme, a persistent stimulus-specific response can be obtained by an appropriate linear combination of this set of temporal filters. Moreover, the same scheme can capture the rich dynamic repertoire observed *in vivo*. Of course, real networks are likely to exhibit a hybrid structure, combining functionally feedforward and feedback components.

A problem that still needs to be addressed by theoretical work is how such feedforward and feedback architectures can emerge as result of learning. Unfortunately, so far, synaptic learning has been limited to very contrived situations, e.g. nonoverlapping inputs, that cause the reliable coactivation of a specific subpopulation with no effect on the rest of the network, see [Amit & Brunel 1997]. This approach only considers small changes in the gain of the system, ignoring the complexities of learning in recurrent networks, where strong positive feedback loops can quickly dominate the dynamics. A general model of learning for WM would, ideally, be able to stabilize the dynamics in a relatively large parameter range. Moreover, in the context of goal-directed behavior it could adapt the network architecture as a function of the tasks to be performed. It is likely that such optimization would result in a solution in which several of the mechanisms discussed before may coexist, to allow for information to be flexibly encoded and manipulated.

4.2 Computational model

4.2.1 Network model

An overview of the network is shown in Fig.4.3 A. Input units encoding different stimuli activate small non-overlapping subsets of excitatory neurons within the recurrent layer (e.g. the green sub-population corresponding to the second stimulus). The output layer receives inputs from all excitatory units and, through a winner-take-all (WTA) mechanism, selects the action to be taken. Depending on this behavioral response, a reward is given which determines synaptic changes through reward-modulated STDP [Izhikevich 2007] (R-STDP). Learning affects both synapses within the recurrent network and synapses connecting to the motor layer.

The recurrent network consists of N units (80% excitatory and 20% inhibitory), with sparse connectivity (with the default value $N = 150$, and $N = 250$ in the structural plasticity experiments, and connection probabilities $p_{ee} = 0.1$, $p_{ei} = 0.7$, $p_{ie} = 0.5$, $p_{ii} = 0$, with indices marking the excitatory and inhibitory populations). The input and motor layers have $K + 1$ (one additional input unit is used for the cue, see task description below) and K units, respectively. Additionaly, the cue signaling the time when the response has to be provided by the motor layer is implemented as the $(K + 1)^{\text{th}}$ input, which activates a distinct subpopulation of neurons within the recurrent layer (alternatively, this input could only affect the motor layer). The connectivity matrix is initialized randomly, with weights drawn from the uniform distribution ($w_{ij} \in [0, 1]$), followed by a sum-to-one weight normalization.

For modeling the neurons, we consider a simplified linear threshold unit [van Vreeswijk & Sompolinsky 1996], with a binary output y described by:

$$y = 1 \quad \text{if} \quad Wx \geq \Theta \quad \text{or} \quad y = 0, \quad \text{otherwise.} \tag{4.1}$$

Here W is the weight matrix corresponding to the neuron inputs x and Θ is the spike threshold. We

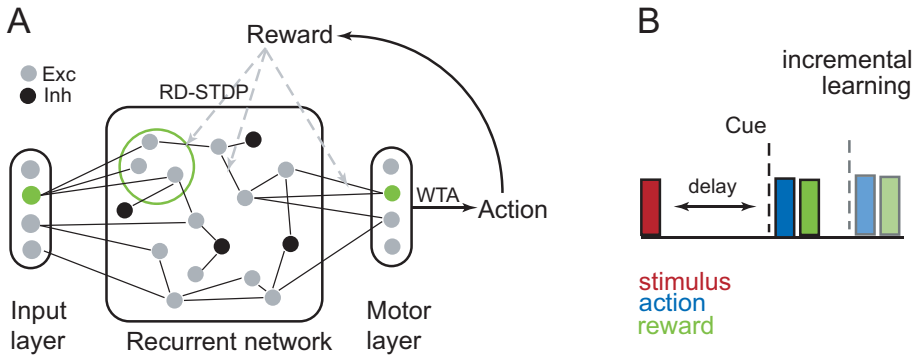

Figure 4.3. Overview of the model and task (A) The model: the network receives localized, stimulus specific inputs. Based on the activation of the neurons in the motor layer, a WTA mechanism selects the action to be performed. Synapses adapt by reward-modulated STDP. (B) A delayed response task: one of K stimuli presented at the beginning of the trial must be held in memory during the delay period. The cue signals the time of response, after which the corresponding reward is given. We use an incremental learning paradigm, where the delay increases in subsequent blocks of trials.

will show that, despite its extreme simplicity, this model is able to capture some key aspects of neuronal dynamics during working memory tasks. Moreover, as the model neuron has no memory itself, a memory trace can only develop as a network effect, requiring an appropriate connectivity matrix, which forms through reward-modulated STDP (R-STDP).

Additionally, a background input is delivered causing the random activation of N_{bk} units at each time step. The input subpopulations are non-overlapping and each include the same number of neurons N_{in}. As mentioned before, the cue activates another subpopulation of N_{c} neurons. Default values are $N_{\mathrm{in}} = 12$, $N_{\mathrm{bk}} = 7$, $N_{\mathrm{c}} = 12$.

4.2.2 Reward-dependent learning

Several experiments have demostrated that STDP depends on the presence or absence of neuromodulators such as dopamine [Pawlak & Kerr 2008, Zhang *et al.* 2009]. They can affect the threshold for potentiation and depression, turn plasticity on and off, and, in some cases, even change the temporal profile of STDP.

In our model as well, network connectivity is modified through a reward-modulated STDP rule (R-STDP), based on that introduced in [Izhikevich 2007]. Specifically, an eligibility trace e_{ij} is associated to each synapse w_{ij}, storing a history of potential weight changes, with an exponential decay to zero:

$$\frac{de_{ij}}{dt} = -\frac{e_{ij}}{\tau_{\mathrm{e}}} + y_i(t)y_j(t-1) - y_i(t-1)y_j(t) + \alpha y_i(t)y_j(t). \tag{4.2}$$

The weight changes include a causal component and a Hebbian term, the relative contribution of each being determined by an additional parameter α (by default, $\alpha = 0.1$). Consistent with the classical STDP framework, potentiation occurs when the presynaptic neuron j fires before the postsynaptic neuron i, while the reverse pattern induces depression. Due to the specifics of our model, causal interactions can only occur between two successive time steps. An additional Hebbian term encourages neurons that are co-activated to continue to fire together, allowing for a potential development of attractor-like dynamics, which would not be possible otherwise. This choice is also justified by a variety of studies indicating that STDP can lead to Hebbian learning, see [Gerstner & Kistler 2002b].

At the time of reward delivery, the strength of synapses is modified proportionally to the eligibility trace e_{ij} and the received reward r:

$$\frac{dw_{ij}}{dt} = \beta r(t) e_{ij}(t), \tag{4.3}$$

where $\beta = 10^{-4}$ is the synaptic learning rate. The weights are rectified such that $w_{ij} \geq 0$.

The reward r can be either positive or negative, as biological evidence from cortico-striatal synapses suggests that dopamine can induce both potentiation or depression in response to tetanic stimulation, depending on its concentration relative to a baseline value [Reynolds & Wickens 2002]. Learning affects only excitatory synapses, both within the recurrent network, and those connecting to neurons in the motor layer.

4.2.3 Homeostatic and structural plasticity

To keep the activity of the recurrent network stable during learning, additional homeostatic mechanisms are required [Lazar *et al.* 2007]. Specifically, the spike threshold for excitatory neurons adapts such that a certain mean average firing rate is preserved:

$$\Delta\Theta = \lambda_{\text{exc}}(y(t) - y_0). \tag{4.4}$$

Here, y_0 is the desired average firing rate, $y_0 \in (0, 1)$, and $\lambda_{\text{exc}} = 10^{-3}$ is the learning rate for the threshold adaptation. In a similar manner, the threshold of inihibitory neurons is regulated by:

$$\Delta\Theta_{\text{inh}} = \lambda_{\text{inh}}(< y_{\text{exc}}(t) > - y_0). \tag{4.5}$$

As before, y_0 is the desired average firing rate of the excitatory neurons, and $\lambda_{\text{inh}} = 10^{-3}$ is the learning rate. Importantly, as demonstrated experimentally, the excitability of inhibitory neurons is determined by the population firing rate for the excitatory population $< y_{\text{exc}} >$, estimated via the release of difussible messengers, such as BDNF [Rutherford *et al.* 1998, Turrigiano & Nelson 2000]. This type of non-local homeostasis may play an important role in maintaining network stability, see discussion in chapter 6.

When modeling synaptic scaling, the total drive received by neurons is constrained through the normalization of incoming weights: $\Sigma_j w_{ij} = 1$, which affects all synapses. Here, this is done explicity, although, as we have seen in chapter 2, the same could be implemented by a local, weight dependent rule.

In addition, structural plasticity is implemented in two steps. First, very weak synapses ($w_ij < \Theta_{\text{syn}}$, whit Θ_{sym} set to a small value) are pruned, as dendritic spines are known to retract in the absence of synaptic activity [Lamprecht & LeDoux 2004]. Second, new synapses can grow between neurons which exhibit correlated activity, but are not yet synaptically connected. The probability of forming a new connection depends on the degree of correlated activity, measured using eligibility traces similar to those used for implementing R-STDP:

$$\frac{ds_{ij}}{dt} = -\frac{s_{ij}}{\tau_{\text{s}}} + y_i(t)y_j(t-1) - y_i(t-1)y_j(t) + \alpha y_i(t)y_j(t), \tag{4.6}$$

with parameter α as defined before and $\tau_{\text{s}} = 2 \cdot K \cdot T_{\text{trial}}$, where T_{trial} is the duration of a single trial. Unlike the eligibility traces used for reward-dependent learning, these correlated activity estimates are computed for all neuron pairs which are not connected by synapses.

Lastly, the retraction and growth processes are balanced, such that the overall connectivity of the network is preserved, as suggested by experimental data [Turrigiano & Nelson 2000]. Due to computational limitations, structural plasticity is investigated only in section 4.3.5 and is not used in the other experiments.

4.2.4 The task

We consider a delayed response task requiring a stimulus to be stored in memory for a certain time such that a stimulus-specific response can be delivered after an external cue. A schematic view of this experiment is shown in Fig.4.3 B. During a trial, one object, selected at random from a set of K, is presented to the network, inducing the activation of the corresponding subpopulation within the recurrent network for one time step. After a delay —either fixed for all trials, or selected independently from a uniform distribution between 1 and a maximum $D_{\max}$— a cue is presented indicating that at the next time step the action will be selected and the corresponding reward will be delivered. The cue is encoded as the $K+1$-th object in the stimulus set, while the reward is ± 1 for correct and incorrect responses, respectively. Action neurons are allowed to fire during the delay, without any effect on the reward.

We use an incremental learning paradigm, in which the difficulty of the task is slowly increased during the experiment. This type of paradigm, in which the agent is allowed to learn from 'easy' examples first has often been used in classical reinforcement learning to speed up convergence. In our model, we implement this by incrementing the delay for the fixed, or maximum delay for the variable delay task, after each T_{inc} trials. This strategy is biologically justified, as during development an infant is confronted with more and more difficult tasks. Moreover, incremental learning of this kind is commonly used for training animals used in WM experiments.

4.3 Results

4.3.1 Network learns to solve the task

Our network learns to correctly perform this task (Fig.4.4). Performance is influenced by task difficulty; it decreases with increasing delay, with better results for the fixed compared to the variable delay task. The incremental learning paradigm improves network performance in both cases.

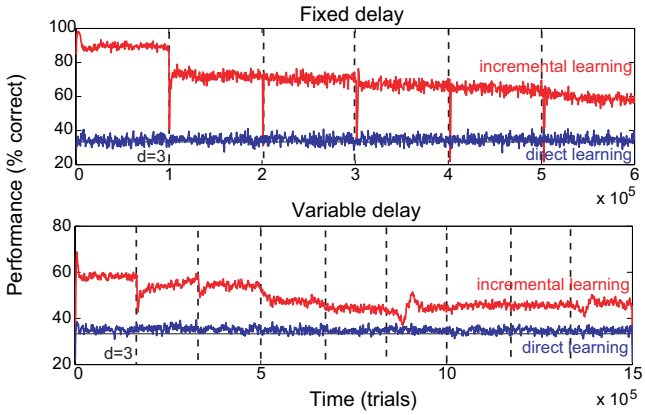

Figure 4.4. Evolution of network performance during learning. The experiment considers $K = 3$ stimuli, a fixed or variable delay task, with or without incremental learning. The initial delay is $d = 3$; vertical lines mark the end of a trial block, when d is incremented.

The less-than-perfect performance can be attributed to several factors. It is partially owed to accumulated noise, as the network receives random background input during the delay. Additionally, because of the simple units and the sparse connectivity, the network capacity is rather small, and often a better representation for one stimulus comes at the cost of a poorer representation for the others. Lastly, it has

been suggested that simple reward-dependent trace learning in spiking neuron networks, as done here, may suffer from some intrinsic limitations due to the limited variability in the response of different neurons, see [Urbanczik & Senn 2009]. This could be overcome by a more complex learning rule, in which the firing rate of each neuron is compared to that of the entire population and the difference between the two is used for modulating learning. Unfortunately, the solution proposed in [Urbanczik & Senn 2009] cannot be straightforwardly generalized to our model, due to the fact that it implicitly assumes a rate-based encoding for the neurons in the network, which is not the case here. How population-level information could be used without any assumptions on the encoding, as requiered by our model, remains unclear.

4.3.2 Task-specific performance improvement

As various plasticity mechanisms act on the network, it is interesting to ask to which extent these contribute to the observed effects. Specifically, we compare our reward-dependent learning to a similarly constructed recurrent network, in which weights within the recurrent network remain fixed, or alternatively are modified by STDP, independently of the obtained reward. In both cases, the learning of the readout to the motor layer continues to be done by R-STDP and the homeostatic mechanisms act as before on all excitatory neurons in the system. The results of this experiment are presented in Fig.4.5.

We see that the performance of the network with R-STDP remains consistently higher than that of the other two networks. Note that, for the fixed connectivity task, the network with static synapses is unable to maintain a stable memory trace for more than 3 time steps, and performance quickly decreases back to chance levels. STDP on its own is able to improve the representation to some extent, as it also leads to a sparse weight distribution. However, without information about the outcome of actions, this improvement is limited.

The two homeostatic mechanisms considered influence the outcome of learning as well. The absence of any of the two leads to a significant drop in performance, often to chance level in the case of the threshold regulation. Synaptic scaling is useful, as it enforces competition between different synapses, contributing to the specificity of the representation and the development of a exponential weight distribution. The spike threshold regulation prevents the network from falling into epileptic-like regimes, which are likely to occur during learning, due to positive feedback from Hebbian learning.

4.3.3 Neurons develop stimulus specificity, in location and time

To better understand the nature of the encoding that emerges after learning, we analyze the properties of the neurons for the final weight matrix. We start by computing the average neural activation during the delay period for each stimulus. Based on this measure, we can conclude that a large percentage of the neurons in the network ($> 40\%$) exhibit stimulus specificity. Examples of such neurons are shown in Fig.4.6 A.

To better quantify the degree of specificity of the neurons, we additionally use the depth of selectivity, often used for describing experimental data [Rainer *et al.* 1998]. Namely, the depth of selectivity is defined as:

$$S = \frac{K - \frac{\Sigma R_i}{R_{\max}}}{K - 1}, \tag{4.7}$$

where K is the number of stimuli as defined before, R_{i} is the firing rate corresponding to stimulus i and $R_{\max} = max\{R_{\mathrm{i}}\}$. This measure takes the value zero when the neural response is identical for all objects and can reach the maximum of one when the neuron responds exclusively to one of the stimuli. The depth of specificity of all excitatory neurons within one network is shown in Fig.4.6 B. With this measure, we can confirm that many neurons in the network exhibit stimulus-specific activation during the delay period, as reported experimentally [Brody *et al.* 2003a].

Additionally, neurons exhibit specificity in time. This can be seen by computing a post-stimulus time histogram (PSTH), for each of the possible stimuli. As seen in Fig.4.7 A, given a certain stimulus at

time $t = 1$, certain neurons will respond at specific times after stimulus onset. This property is a result of learning, as it is not seen in trials at the beginning of the experiment.

The fact that neurons respond to inputs in a time-dependent way suggests that stimulus identity is encoded in a neuronal trajectory. For a better visualization of this encoding, we perform a dimensionality reduction, by principal component analysis (PCA). In Fig.4.7 B, we show the mean trajectories corresponding to each stimulus, in the space of the first three principal components (PCs). The types of

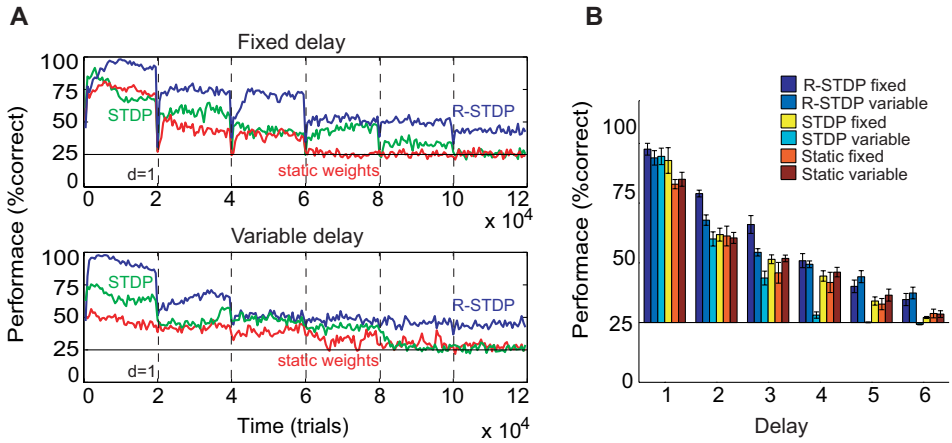

Figure 4.5. R-STDP can be used for reservoir optimization. (A) The evolution of task performance for both fixed and variable delays, in the case in which the connectivity is shaped by reward dependent STDP (blue), STDP (green) or remains fixed (red), $K = 4$, initial delay is 1. (B) Statistical analysis of the performance for the experiment described before. Results averaged over 10 experiments, with error bars measuring standard error.

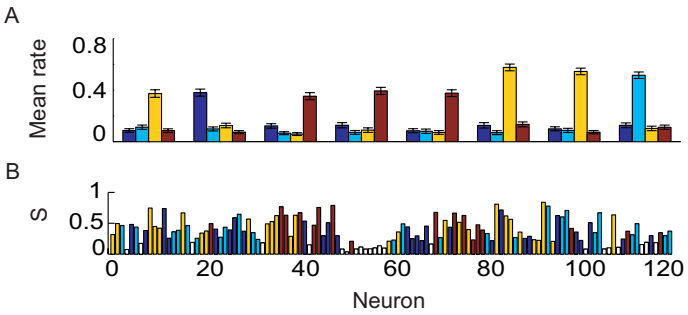

Figure 4.6. Neurons develop stimulus specificity. (A) Examples of network neurons, with their mean firing rate during the delay period for each of the input stimuli ($K = 4$, different colors mark different stimuli). (B) Depth of specificity and neuron selectivity for the entire excitatory population. Different colors mark different stimuli, white corresponding to unspecific neurons, defined as having $S < 0.4$.

trajectories observed tend to differ between the fixed and the variable delay task (synfire-chains for the variable and limit-cycles for the fixed delay, respectively). Small limit cycles can persist in the network for a relatively long time, which makes them a good solution for stimulus representation. The difference may arise due to the fact that it is much easier for the readout to detect a single time-point in a limit cycle (for fixed delay), compared to the set of all states (for variable delay).

The degree of variability of different stimulus-specific trajectories is analyzed in Fig.4.7 C. The cluster size for each time-point increases over time, as suggested before by the stimulus-specific PSTH. Importantly, this type of stimulus-induced reproducible trajectory emerges as a result of reward-modulated STDP, and is not observed before learning (Fig.4.7 D).

Due to the Hebbian component in the synaptic update, one could have assumed classic attractor-like solutions would be preferred for stimulus encoding. However, the emerging representation is usually time-varying. We have compared the behavior of the network with and without the Hebbian component in Fig.4.8 A. Performance is initially higher for the network with Hebbian learning. However, it decays below that of the STDP-only network for larger delays. A closer look at the network dynamics hints to one possible explanation why this may be the case. The Hebbian component encourages a more compact, attractor-like representation. However, this encoding requires stimulus-specific neurons to fire consistently for long delays, activating the homeostatic threshold regulation (Eq.4.4) and destabilizing the representation. In both cases, the weight distribution is sparse and approximately exponential (Fig.4.8 B).

4.3.4 Delayed categorization

Areas typically associated with persistent activity do more than passively storing information. Maintenance is often intertwined with decision making and motor planning [Durstewitz *et al.* 2000b]. In our framework, one can think that neuronal activations can encode either the initial stimulus, or the action

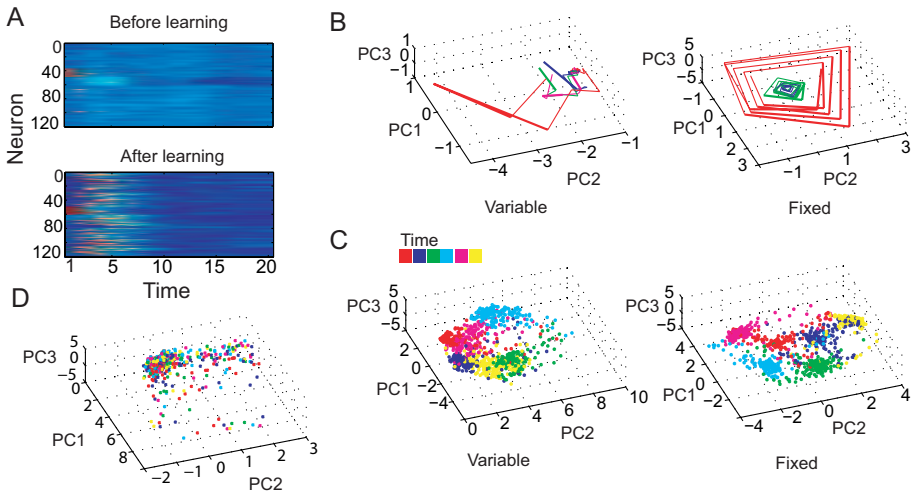

Figure 4.7. Stimulus encoding is time-dependent. (A) PSTH of excitatory neurons in the recurrent network: given a certain stimulus, neurons in the network respond in a time-dependent manner. (B) Emerging neuronal trajectories for different stimuli (in color), in a fixed and variable delay task. (C) Degree of variability for stimulus-specific trajectories. Different colors mark different time-points after stimulus onset. (D) Before learning, the network activity following the presentation of one stimulus is highly variable and not organized in time.

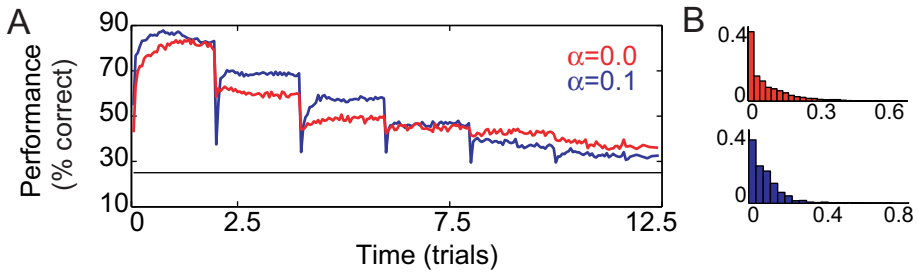

Figure 4.8. Comparison of different synaptic plasticity rules. (A) Performance in a fixed delay task when synaptic plasticity includes ($\alpha = 0.1$) or does not include ($\alpha = 0$) a Hebbian term. (B) Corresponding weight distribution for $\alpha = 0$ (up, red) and $\alpha = 0.1$ (down, blue).

to be taken in response to it. However, for the simple delayed response task we have considered so far, there would be no difference between the two, as actions simply signal stimulus identity.

Due to this fact, we consider a modified version of the task, which can distinguish between stimuli and actions. Specifically, stimuli are divided in two disjoint subsets and the task requires the motor layer to signal the class to which the stimulus belongs. For simplicity, we take stimuli with odd indices to belong to class A, and even ones to class B. Of course, due to the random initialization of the connectivity, any other subdivision would be equivalent.

The network is able to solve this task (Fig.4.9 A) as well. Interestingly, an analysis of the neuronal activation reveals both stimulus- and action-specific neurons emerge after learning (Fig.4.9B-D).

4.3.5 Network optimization: the role of structural plasticity during learning

When investigating the performance of individual networks, one notes a great degree of variability between different network instances. Moreover, classic reservoir theory shows that networks of binary inputs are rather sensitive to topology, which significantly influences both network dynamics and computational performance [Büsing *et al.* 2009]. In our particular case as well, due to the sparse connectivity required for ensuring appropriate network dynamics (see Appendix B), performance critically depends on the particular instance of the weight matrix.

As a solution for this type of problem, structural plasticity can further optimize network connectivity, on a slower time-scale. After learning, most of the synapses are very weak (exponential weight distribution) such that some of them can be removed without affecting the network response. Then, they can be replaced by other weak synapses, from which reward-dependent learning can select a few for stabilization. These synapses would then survive and eventually contribute to the network behavior.

Indeed, when comparing networks implementing structural plasticity to networks with fixed random connectivity, we see that the performance can be significantly improved by network reorganization. In contrast, the activity-dependent synaptic reorganization corrects a 'bad' initial choice, such that the network can encode the input stimuli more reliably (Fig. 4.3.5). Interestingly, the Fano factor for the distribution of incoming synapses is small, resembling values reported for cortical networks, as seen in [Fares & Stepanyants 2009].

4.4 Discussion

We have shown that a recurrent network of binary neurons can solve a delayed response task by reward dependent learning, without any constraints imposed on how stimulus identity should be represented.

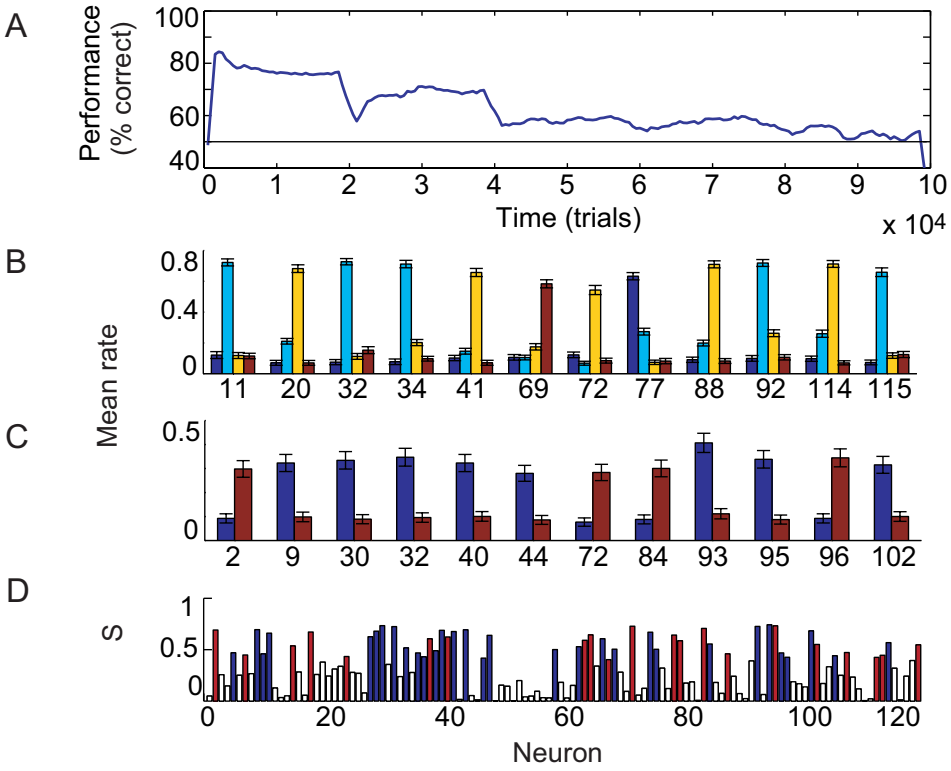

Figure 4.9. A categorization task. (A) Performance for a variable delay categorization task with 4 stimuli divided in 2 categories. (B) Example of stimulus-specific neurons. (C) Activity of neurons within the recurrent network exhibiting action-specificity. (D) Depth of specificity for the two categories, with color coding as described for Fig.4.6.

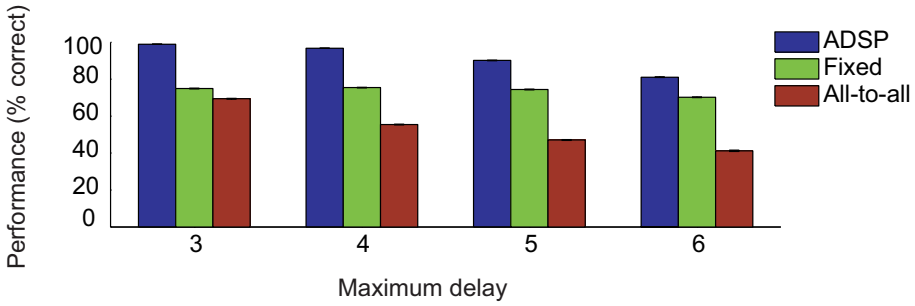

Figure 4.10. Activity-dependent structural plasticity improves performance. Comparison of the performance for an incremental variable delay task for networks with fixed sparse random connectivity, all-to-all connections and networks implementing activity-dependent structural plasticity. In all cases, synapses adapt by reward-dependent STDP, as before.

Our simple model has been able to capture several essential properties reported in WM experiments. Specifically, through R-STDP neurons acquire stimulus specificity and develop a temporally varying representation of their input. Additionally, our model's performance is significantly higher than that of networks with fixed connectivity or networks trained with STDP only.

Temporally varying representations are preferred to more classical attractors in our model. This may be attributed to homeostatic regulation, as the sliding threshold discourages neurons to remain activate for a long time. Analytically, it has been shown that a Hopfield network with dynamic thresholds which depend on a unit's past activity can lead to temporal activity sequences similar to those observed here [Horn & Usher 1989]. Importantly, the emerging trajectory encoding is similar to that observed experimentally.

In learning WM experiments, novel stimuli elicit a higher activation within prefrontal areas. However, the specificity of the representation is significantly better for familiar stimuli, and more robust to stimulus degradation [Rainer & Miller 2000, Rainer & Miller 2002]. In light of these results, it would be interesting to analyze the response of the optimized network to noisy stimuli. Additionally, future work should consider a more detailed analysis of the properties of the weight matrix obtained after learning.

As possible experimental predictions, our model suggests that pharmacological manipulations of R-STDP could be detrimental for learning new WM tasks. However, one must be careful to disentangle effects owed to a disruption in learning to other dopamine effects, given that dopamine is known to modulate synaptic and voltage gated currents in PFC neurons *in vitro* and —through the activation of D1 receptors— can change the properties of NMDA currents.

From a more general computational perspective, it is interesting to relate our work to classic frameworks for cortical computation, such as the liquid state machine (LSM) [Maass *et al.* 2002]. How to optimize reservoirs for specific problems or classes of problems remains an unanswered question within the LSM community [Schmidhuber *et al.* 2007, Haeusler *et al.* 2009]. In the framework of our simple model, we have been able to show that the connectivity of recurrent networks can be optimized in a biologically-plausible, task-specific way to solve a problem that would have been difficult to solve with a similar randomly connected network. It would be interesting to try to generalize these findings with a network more similar to traditional LSM reservoirs.

Moreover, given that transient dynamics seem to be ubiquitous in various cortical areas [Buonomano & Maass 2009], it is tempting to think that the results presented here are not restricted to structures traditionally related to working memory. It has long been known that in the olfactory system, for example, odors are represented as a succession of transient states, which eventually converge to an attractor [Laurent *et al.* 2001]. In this case, it is this transient trajectory that encodes the most information about the stimulus and allows for maximum discriminability. Mathematically, such input-specific transients have been modeled by stable heteroclinic channels, consisting of a sequence of saddle points [Rabinovich *et al.* 2000]. So far, it has remained unclear what are the mechanisms through which such representations could emerge in a biological neural network. It is tempting to think that reward dependent learning could play a role in this case as well.

5

Glia-mediated synaptic scaling and epileptogenesis

So far, we have investigated the computational benefits of the interaction between homeostatic and Hebbian plasticity. In the last part of this work, we will see that homeostatic plasticity is not always beneficial, as under certain pathological conditions, attempts to compensate disruptions in neuronal activity become detrimental and lead to hyperactivity and seizures.

The work presented here considers a certain form of synaptic scaling, which has been reported to occur in several cortical regions in the mouse [Stellwagen & Malenka 2006,Kaneko *et al.* 2008]. Specifically, it has been demonstrated that glia —support cells in the brain— modulate synaptic strength through the soluble form of the proinflammatory cytokine tumour necrosis factor α (TNF-α) [Beattie *et al.* 2002, Stellwagen & Malenka 2006].

Interestingly, TNF-α also plays an important role in the immune system. It is a pro-inflammatory cytokine whose levels can rise dramatically during local acute immune responses. A 10-fold increase in blood serum is frequently found, and even 100-fold increases are seen during sepsis [Haagmans *et al.* 1994, Damas *et al.* 1992, Galic *et al.* 2008]. Produced by immune cells such as monocytes, T-lymphocytes, and phagocytes, TNF-α can activate neutrophils and macrophages, control the recruitment of immune agents from the blood and regulate the permeability of the blood-brain-barrier (blood-brain barrier) to soluble molecules [Deli *et al.* 1995, Rosenberg *et al.* 1995, Bechmann *et al.* 2007].

This dual role of TNF-α as both immune system messenger and neuromodulator raises the interesting question whether TNF-α produced during an immune response in the brain could interfere with the homeostatic regulation of synapses, potentially increasing synaptic strength sufficiently to trigger paroxysmal activity. If so, this interference between different signaling pathways could explain some of the recent evidence suggesting an immune system influence in seizure initiation in certain pathological conditions [Vezzani 2005, Lucas *et al.* 2006].

To investigate this hypothesis, we have developed a computational model of TNF-α-mediated synaptic scaling. Our model shows that an overall increase in TNF-α levels following chronic inflammation or TNF-α overexpression by glia cell can push the network activity into a paroxysmal regime. In addition, it shows that neuronal hyperexcitability can arise after localized disruptions in network structure, resulting from simulated local lesions. In particular, following partial deafferentation, TNF-α produced by glial cells within the lesion area diffuses to the neighboring tissue and triggers network bursts.

The results presented here have been published as [Savin *et al.* 2009].

5.1 Background

5.1.1 TNF-α mediated synaptic scaling

The experimental findings which have guided the design of our model are briefly summarized in Fig. 5.1, while a complete list of reported effects of TNF-α on plasticity is shown in Table 5.1.1. First, the acute application of TNF-α was shown to increase the AMPA receptor surface expression in hippocampal

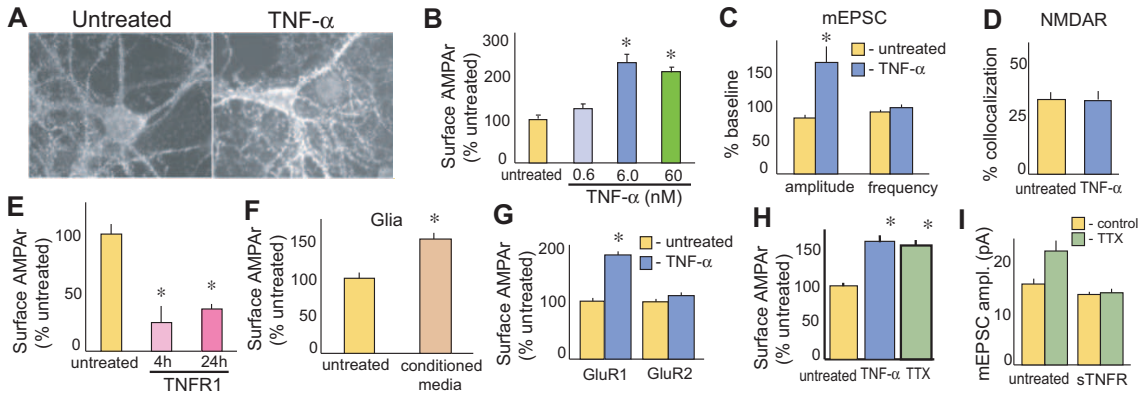

Figure 5.1. Overview of the experimental results on TNF-α mediated synaptic scaling. A) Example of surface AMPAR staining. B) Effects of TNF-α on surface AMPAR staining ($^*P < 0.01$). C) Change in mEPSC frequency and amplitude for TNF-α treated and untreated neurons. D) TNF-α does not affect the localization of NMDA receptors. E) TNFR1 application changes AMPAR surface expression. F) Effects of glial conditioned media on surface AMPAR expression are similar as for TNF-α application. G) TNF-α preferentially increases synaptic expression of GluR2-lacking AMPARs. H) Two days of TTX treatment increase AMPAR surface expression, similar to direct TNF-α application. I) Addition of sTNFR during the TTX treatment blocks the effect, suggesting that the synaptic scaling observed after chronic activity blockade is TNF-α-mediated. Adapted from [Beattie *et al.* 2002, Stellwagen *et al.* 2005, Stellwagen & Malenka 2006].

neuron cultures, but had no effect on N-methyl-D-aspartic acid (NMDA) receptors [Beattie *et al.* 2002], see Fig. 5.1 A, B, D. The increase in AMPAR surface expression was accompanied by a significant increase in synaptic efficacy, measured by mEPSP amplitude, after 10 min of TNF-α exposure, as shown in Fig. 5.1 C. Moreover, experiments showed that a certain amount of TNF-α is necessary for maintaining the synaptic strength of neurons and blocking TNF-α receptors by the application of TNF-α receptor 1 (TNFR1, which functions as a TNF-α antagonist) for 4 h causes a clear decrease in AMPAR surface expression (see Fig. 5.1 E).

In a subsequent paper, the same group demonstrated that, during this process, TNF-α interacts with TNFR1 expressed by hippocampal neurons, causing the exocytosis of glutamate receptor 2-lacking AMPA receptors, through a phosphatidylinositol kinase 3 process and the endocytosis of $GABA_A$ receptors [Stellwagen *et al.* 2005], see Fig. 5.1 G.

Interestingly, the paper by [Beattie *et al.* 2002] also showed that endogenous TNF-α produced by astrocytes has similar effects on synaptic excitability, suggesting that glia could play a role in controlling the strength of excitatory synapses by TNF-α secretion (Fig. 5.1 F). In a subsequent paper [Stellwagen & Malenka 2006], the relevance of these findings for normal brain function has been investigated. It was demonstrated that after chronic activity blockade (2 days) by tetrodotoxin the strength of synapses is multiplicatively scaled up, as in classic synaptic scaling experiments [Turrigiano *et al.* 1998] and that this effect is owed to an increase in TNF-α. Moreover, this surplus TNF-α is owed to local glial cells' production (see Fig. 5.1H,I). Moreover, as for the direct application of TNF-α, soluble TNFR which scavenges endogenous TNF-α blocks the ability of the conditioned medium or tetrodotoxin to induce the scaling of mEPSC and AMPARs. This effect of TNF-α is not restricted to excitatory synapses. After chronic activity blockade, TNF-α decreases the amplitude of mIPSCs as well [Stellwagen *et al.* 2005, Stellwagen & Malenka 2006].

The functional relevance of TNF-α-mediated synaptic scaling was recently confirmed in vivo, in a monocular deprivation model using TNF-α knockout mice [Kaneko *et al.* 2008]. It is known that, in wild type mice monocular deprivation within the critical period leads to a decrease in the response to inputs from the deprived eye, followed by a later strengthening of the response to input from the open eye. In the TNF-α knockouts, cortical neurons receiving binocular inputs still show a weakening of the inputs from the deprived eye, but no increase of responses to the open eye. The same effect can be mimicked by the temporary inhibition of TNF-α signaling (using sTNFR1), suggesting that it not owed to defects in early circuit formation, but to an active process mediated by TNF-α. These results are consistent with a model in which the increased response to the open eye is owed to synaptic scaling, which attempts to compensate for the decrease in total input to the neuron, as demonstrated in [Stellwagen & Malenka 2006].

These findings, combined, suggest that glial cells estimate the synaptic drive to neurons via neurotransmitter spillover at the synapses [Volterra & Meldolesi 2005]. Activated glial cells, in turn, stimulate the post-synaptic neurons via TNF-α, inducing an increase in excitatory synaptic strength. The process is accompanied by the endocytosis of $GABA_A$ receptors [Stellwagen & Malenka 2006], which results in a decrease in inhibitory synaptic strength. In sum, the neuromodulator TNF-α [Pan *et al.* 1997, Vitkovic *et al.* 2000] seems to be of vital importance for balancing neuronal activity in various cortical regions.

Current experimental evidence shows that TNF-α is involved only in the up-scaling of synapses following a decrease in activity, suggesting that homeostatic synaptic scaling is mediated by at least two distinct neuromodulators. While TNF-α shifts synaptic weights to more excitation and less inhibition, other signals push synapses towards less excitation and more inhibition. A potential candidate is the brain-derived neurotrophic factor (BDNF), as BDNF has a reverse effect of synaptic scaling and can reverse the synaptic potentiation induced by tetrodotoxin activity blockade [Turrigiano 2006].

5.1.2 Glia cells, more than brain glue

For every neuron in the brain there are roughly ten glia cells supporting its function. While the role of oligodendrocytes and microglia is well understood (axon myelinization and immune defense, respectively), that of astrocytes remains elusive. For long, astrocytes have been viewed as 'brain glue', passive elements needed only to provide a scaffold for neurons, with no role in brain function. As new experimental evidence became available, they started to be considered support cells, needed to ensure the optimal functioning of neurons [Volterra & Meldolesi 2005]. But more interestingly, astrocytes have been recently shown to mediate a variety of neuron processes. It is now known that glia cells promote the formation of excitatory synapses, and may be important for hippocampal morphogenesis [Barry *et al.* 2008]. Moreover, astrocytes were shown to integrate and process synaptic information and control synaptic transmission and plasticity, leading to the hypothesis that astrocytes are active partners in the synaptic function [Haydon 2001, Newman 2003, Perea *et al.* 2009, Henneberger *et al.* 2010].

Ca^{2+}-mediated astrocytic excitability

One of the reasons why astrocytes have long been considered nonexcitable is their inability to generate action potentials. However, although they cannot communicate through electrical signals, glia cells can still be activated through an increase in intracellular Ca^{2+}, owed mainly to the mobilization of Ca^{2+} stored in the endoplasmatic reticulum [Perea *et al.* 2009]. This activation can occur spontaneously (intrinsic oscillations) or due to synaptic activity [Perea & Araque 2005a], suggesting that glia can respond to the activity of neighboring neurons. This has been demonstrated in culture, brain slices, but also in vivo, in response to neural activation, see [Perea *et al.* 2009].

As their processes wrap around the synaptic cleft of synapses (a single astrocyte can contact 100,000 synapses) and they express a range of neurotransmitter receptors (mostly metabotropic), astrocytes are well placed to sense the overall activity of the neighboring neurons. During astrocytic activation, either spontaneous or activity dependent, the rise in Ca^{2+} is initially restricted to a small area within

Table 5.1. Effects of TNF-α on synaptic transmission

Manipulation	Phenotype	References
Acute application		
Exogenous TNF-α	mEPSC frequency increase, EPSP enhanced, decreased inhibition, AMPA/NMDA ratio increased	[Beattie et al. 2002] [Stellwagen *et al* 2005, 2006]
Soluble TNFR	mEPSC decreased in vitro and slices; mEPSC amplitude decreased in vitro; AMPA/NMDA ratio increased	[Beattie *et al* 2002]
Hebbian plasticity		
Exogeneous TNF-α	LTP inhibited	[Tancredi *et al* 1992] but see [Stellwagen 2006]
TNFR$^{-/-}$	normal LTP; impaired LTP	[Albensi & Manson 2000] but see [Stellwagen 2006]
TNF-$\alpha^{-/-}$	LTP and LTD normal	[Stellwagen 2006]
Homeostatic plasticity (under TTX)		
TNF-$\alpha^{-/-}$	no increase in mEPSC amplitude	[Stellwagen 2006]
soluble TNFR	no increase in mEPSC amplitude, no decrease in mIPSC amplitude	[Stellwagen 2006]
Visual cortical plasticity (Deprivation induced)		
TNF-$\alpha^{-/-}$	normal decrease in deprived eye; no increase in strength of response to open eye	[Kaneko *et al* 2008]
soluble TNFR	normal decrease in deprived eye; no increase in strength of response to open eye	[Kaneko *et al* 2008]

Adapted from [Boulanger 2009].

the astrocytic processs (a 'microdomain') and can then extend to other regions of the cell. Moreover, astrocytic Ca^{2+} signals can propagate to neighboring glia cells as intercellular Ca^{2+} waves [Charles *et al.* 1991, Cornell-Bell *et al.* 1990, Scemes & Giaume 2006].

Interestingly, astrocytes do not respond linearly to activity neighboring synapses. They discriminate the activity from different synaptic pathways, e.g. in area CA1, astrocytes respond to glutamatergic activity coming from the Schaffer collaterals, but not the alveus [Perea & Araque 2005b]. Combining information from multiple sources is a nonlinear process (except when superimposing GABA and glutamate, probably because they activate different pathways). Moreover, the response is either supralinear or sublinear, depending on the firing rate of the stimulated neurons [Perea & Araque 2005b], leading to the view that astrocytes integrate synaptic information in nontrivial ways, potentially relevant for information processing in the brain [Perea *et al.* 2009].

Gliomodulators and synaptic transmission

The communication between neurons and glia is bidirectional. In response to Ca^{2+} activation, astrocytes release various molecules, such as glutamate, adenosine triphosphate (ATP), adenosine, D-serine, GABA, TNF-α, which can, in their turn, regulate neuronal and synaptic activity (see Fig. 5.3). Among other things, they were shown to increase neuronal excitability facilitating neuronal synchronization, to modulate or even induce long-term synaptic potentiation (LTP) [Perea & Araque 2007] and, as already discussed, to mediate synaptic scaling [Stellwagen & Malenka 2006]. A more detailed overview of the effects of different gliotransmitters is shown in Table 5.1.2. So far, there is little experimental evidence

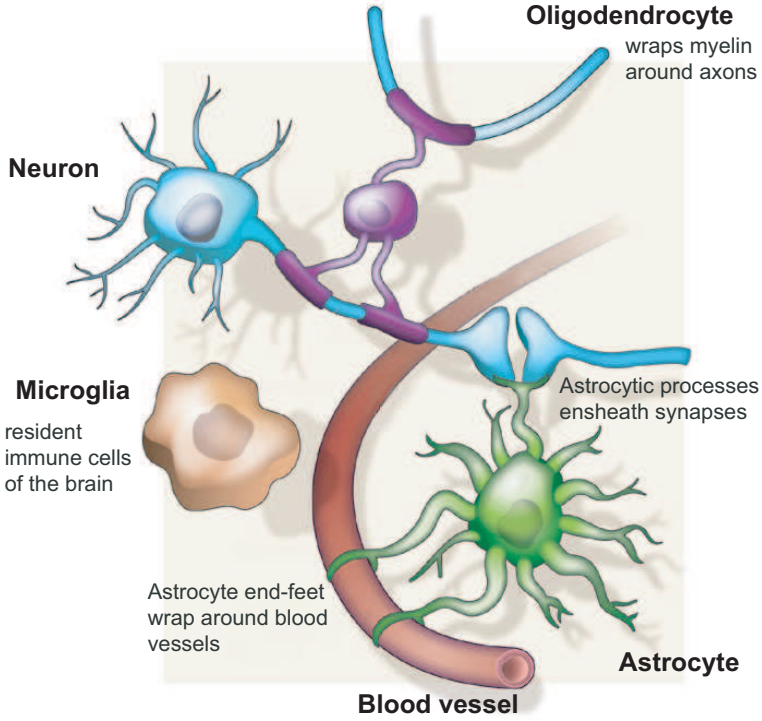

Figure 5.2. Overview of glia-neuron interactions. Several types of glia cells exist in the cortex, having different functions. Microglia are the resident immune cells in the brain, protecting it from damage or infection. Oligodendrocytes wrap myelin around axons to speed up neuronal transmission. Astrocytes extend processes that ensheath blood vessels and synapses. Adapted from [Allen & Barres 2009].

on the role of gliotransmitter release on behavior. As a first step in this direction, it was recently demonstrated, using genetic manipulations, that astrocyte-secreted ATP contributed to the deficits associated with loss of sleep in mice [Halassa *et al.* 2009].

Experimental evidence suggests that at least some gliotransmitters are released in a Ca^{2+}-dependent manner through vesicle or lysosome exocytosis. Moreover, gliotransmitter vesicles are sometimes located within the astrocytic processes, in close proximity to synapses, see [Perea *et al.* 2009]. Despite these findings, it is not clear what really is the degree of locality of neuron-astrocytic processes. In some cases (ATP), the interactions seem to be rather specific, while in others astrocytes seem to act as a bridge mediating nonsynaptic interactions between neighboring neurons [Navarrete & Araque 2008]. It remains to be established in which cases this neuron-glia communication is point-to-point and in which cases it is done through unspecific gliotransmitter spillover. The degree of specificity of these interactions may be relevant for the stability of the network, under certain pathological conditions, as we will show below for our model.

A single gliotransmitter can have different effects depending on the site of action, the receptor subtype and potentially network state (it is still unclear how these processes are affected by the presence of neuromodulators such as dopamine), which makes the study of these interactions sometimes complicated. Moreover, we may imagine that several gliotransmitters act on neurons at the same time. Currently, the effects of these interactions remain poorly understood.

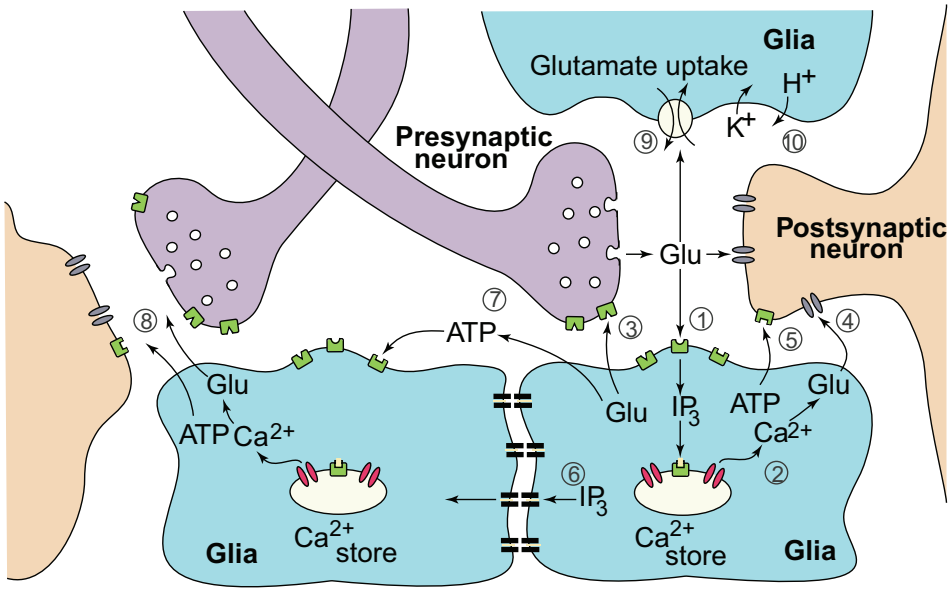

Figure 5.3. Schematic view of glutamate-mediated neuron-glia interactions. Adapted from [Newman 2003]. Glutamate release (Glu) from the presynaptic terminal activates glial receptors (1), evoking an increase in Ca^{2+} levels (2) and the release of glutamate from glia. Presynaptically, glial Glu regulates transmitter release (3), while postsynaptically it directly depolarizes neurons (4). Activated glia also release ATP, which inhibits postsynaptic neurons by activating A_1 receptors (5). Additionally, intercellular Ca^{2+} waves can propagates between astrocytes by diffusion of inositol (1,4,5)-trisphosphate (IP3) through gap junctions (6) and by release of ATP (7), leading to the modulation of distant synapses (8). Glia can also modulate synaptic transmission by uptake of glutamate (9) and by regulating extracellular K^+ and H^+ levels (10).

5.1.3 Expression of immune proteins in the CNS

The presence of the bbb and the fact that immune responses are attenuated in the CNS originally lead to the view of the brain as an immunopriviledged site, largely unaffected by systemic infections, which interacts with the immune system only during disease or trauma [Perry *et al.* 1995]. Recent experiments paint a more complex picture, however. Immune agents are frequently observed in the brain [Bechmann *et al.* 2007] and glial cells can also acquire immune functions [Sebire *et al.* 1993]. Moreover, neurons themselves express immune proteins and can respond to cytokines secreted by infiltrating leukocytes, microglia or astrocytes.

Immune proteins expression in the CNS during inflammation

It was demonstrated that the CNS shows a robust immune response to brain infections and a variety of injuries (trauma, ischemia). Additionally, CNS inflammation can follow a systemic immune response, by blood transmission or local production [Ransohoff 2009]. Inflammation in the brain involves both the innate and adaptive immune system and shares molecules and pathways of systemic infection [Vezzani & Granata 2005], with inflammatory mediators being secreted by microglia, astrocytes, neurons and oligodendrocytes.

Although it is true that the brain responds differently to immune challenges, in that leukocyte recruit-

Table 5.2. Gliotransmitters and synaptic transmission

Region	Phenotype	References
Glutamate		
Hippocampus	Depression of evoked EPSCs and ICSCs	[Araque 1998] [Liu 2004]
	Frequency increase of mEPSCs, mIPSCs	[Kang 1998] [Araque 1998]
	Frequency increase of spontaneous EPSCs	[Fiacco 2004] [Jourdain 2007]
	Frequency increase of spontaneous IPSCs	[Liu 2004]
	Postsynaptic SIC	see [Perea 2009]
	Increase of neuronal excitability	[Bezzi 1998]
	Heterosynaptic depression	[Andersson 2007]
Cortex	Postsynaptic SIC	[Ding 2007]
Ventrobasal thalamus	Postsynaptic SIC	[Parri 2001]
Nucleus accumbens	Postsynaptic SIC	[D'Ascenzo 2007]
Olfactory bulb	Postsynaptic SIC	[Kozlov 2006]
Retina	Light-evoked neural activity	[Newman 1998]
ATP/Adenosine		
Hippocampus	Heterosynaptic depression	[Serrano 2006] [Zhang 2003]
	Modulation of LTP	[Pascual 2005]
	Synaptic depression	[Pascual 2005]
Cerebellum	Depression of spontaneous EPSCs	[Brockhaus 2002]
HPV	Insertion of AMPARs	[Gordon 2005]
Retina	Depression of light-evoked	[Newman 2003]
D-Serine		
Hippocampus	Modulation of LTP	[Yang 2003]
HSN	Modulation of LTP	[Panatier 2006]
Retina	Potentiate NMDAR transmission	[Stevens 2003]
TNF-α		
Hippocampus	AMPAR insertion, synaptic scaling	[Stellwagen *et al.* 2006]
Cortex	Synaptic scaling	[Kaneko *et al.* 2009]
GABA		
Olfactory bulb	Postsynaptic SOC	[Kozlov 2006]

Abreviations: SIC/SOC, slow inward/outward current; HPN, hypothalamic paraventricular nucleus; HSN, hypothalamic supraoptic nucleus. Adapted from [Perea *et al.* 2009].

ment is limited and delayed, and macrophages acquire a microglia phenotype after crossing the bbb [Perry *et al.* 1995], the activation of microglia is rapid [Lucas *et al.* 2006]. A CNS insult (e.g. trauma, or infection) causes microglia, the resident immune cells, to produce the receptors and soluble proteins which are normally produced by leukocytes (major histocompatibility complex II (MHCII), toll-like receptorss, chemokines, cytokines). This leads not only to an immune response within the cortical tissue, but it was shown that microglia can also change the permeability of the blood-brain barrier [Bechmann *et al.* 2007], allowing proteins from the blood and leukocytes to infiltrate the brain [McAllister & van de Water 2009].

Immune proteins and normal brain function

Apart for their roles during inflammation, several proteins of the immune system are constitutively expressed in the normal brain, being involved in a range of functions related to brain development and adult plasticity. Some of these roles are generally accepted, others still debated. Inconsistencies in the data on the presence of cytokines in normal brains have to do with low concentrations (fmols) or highly specific activity and high turnover, making concentrations difficult to detect and quantify [Rostne *et al.* 2007].

We have already seen that TNF-α mediates synaptic plasticity in hippocampus and visual cortex, but this is not a singular example. Cytokines were shown to play important roles development of neural circuitry, e.g. neural growth and the activity dependent synaptic refinement [Boulanger 2009], in stem cell renewal and neuronal differentiation [McAllister & van de Water 2009] and the regulation of ionic channels permeability [Houzen *et al.* 1997]. Also, they are needed for normal synaptogenesis and synaptic stability [Boulanger 2009]. During embryogenesis, immune molecules act as chemo-attractants, critical for neural migration, neurite outhgrowth and neurogenesis [Rostne *et al.* 2007].

Beyond their roles in in brain development, immune proteins act as neuromodulators and are involved in some forms of functional plasticity in mature synapses [Boulanger 2009]. For example, it was shown that major histocompatibility complex I (MHCI) expression is particularly high in regions undergoing activity-dependent plasticity, such as the developing visual system and the adult hippocampus [Boulanger 2009]. Interestingly, MHCI is bidirectionally regulated by changes in activity and, together with TNF-α, is believed to be necessary for synaptic scaling [Boulanger 2009]. As TNF-α is known to alter MHCI cell surface expression on neurons [Collins *et al.* 1986, Shatz 2009], it seems likely that TNF-α and MHCI induced synaptic scaling share a common pathway.

Immune proteins mediate not only the homeostatic regulation of neural activity, but also Hebbian forms of synaptic plasticity. For example, IL-1β and TNF-α (acute application or endogenous production induced by lipopolysaccharide (LPS)) were shown to inhibit LTP, potentially due to the fact that IL-1β inhibits Ca influx into neurons [Cunningham *et al.* 1996]. Additionally, long-term synaptic depression (LTD) is reduced or abolished in mice lacking TNF-α receptors, an effect which can be mimicked by the acute application of a variety of other related proinflammatory cytokines.

Moreover, cytokines are involved in the range of behavioral changes that accompany systemic inflammation. Specifically, blocking TNF-α or IL-1β was shown to alter the regulation of sleep [Vitkovic *et al.* 2000] and affects the release of dopamine and vasopressin [Rostne *et al.* 2007]. It was suggested that the role of proinflammatory cytokines as neuromodulators which modulate LTP/LTD may explain the poor learning performance observed during CNS inflammation.

5.1.4 Inflammation and epilepsy

It is interesting to note that CNS infection is often followed by epilepsy, with a probability 7 times higher than that in normal population [Lancman & Morris 1996]. This lead to the suggestion that immune mechanism may be involved in the pathogenesis of some forms of epilepsy [Aarli 2000]. The hypothesis that CNS inflammation can enhance the predisposition of brain tissue to develop seizures and brain damage is supported by the finding that anti-inflammatory drugs can provide effective anti-epileptic treatment in some cases [Vezzani 2005].

Although the transformation of a healthy neural tissue into a seizure generating one is a complex process, which cannot be reduced to a single mechanism, given the reported effects of inflammatory molecules in the brain, it seems likely that they would contribute to the changes in excitability observed during seizures. They are known to affect excitotoxicity by interfering with the mechanisms of glutamate release and reuptake and enhance production of nitric oxide (NO) [Vezzani & Granata 2005]. In this light, a preexisting chronic inflammatory state can predispose to seizures and promote neurological dysfunctions.

Moreover, additional mechanisms through which inflammation may contribute or predispose to seizures and cell death have been suggested by recent experimental evidence, which suggests that, in some cases, seizures may be owed to the break of the bbb [Kim *et al.* 2009, Fabene *et al.* 2008]. Specifically, it is know that cytokines can damage the bbb, leading to increased permeability to external substances which can cause seizures [Oby & Janigro 2006, Marchi *et al.* 2007]. There is still an ongoing debate if the compromised integrity of bbb causes or is a result of seizures. Still, it is certain that changes in bbb properties often accompany neurological conditions which are associated with the late onset of epilepsy [Vezzani & Granata 2005].

Interestingly, a seizure itself triggers a strong immune response, which is rapid (1h) and long lasting (IL upregulated 60 days after) [Vezzani 2005]. This has lead to the suggestion that cytokines synthesized and released during or after seizures, as result of cell injury, may contribute to epileptogenesis. In this view, an initial epileptic event (which can occur due to structural abnormalities, e.g. cerebral malformations, functional abnormalities, such as mutations in ionic channel properties, or due to trauma) can itself lead to recurring seizures [Vezzani 2005, Vezzani & Granata 2005].

The evidence on the role of inflammation in triggering seizures comes from both clinical practice and animal models. We review the most important findings below.

Evidence from clinical cases

Conditions associated with an increase in pro-inflammatory markers (e.g. Rassmunsen encephalitis, West syndrome and tuberous sclerosis) are often accompanied by seizures [Vezzani & Granata 2005, Aarli 2000]. For example, the prevalence of epilepsy in patients with systemic lupus erythematosus (SLE) is 8 times higher than in the normal population [Aarli 2000]. Additionally, a certain variant of temporal lobe epilepsy (TLE) has been reported to have genetic causes associated with enhanced ability to produce cytokines after a immune stimulus. Unexpectedly, increased expression of proinflammatory molecules has been reported in resected tissue from patients with drug-resistant epilepsy, also for disorders without a known inflammatory pathophysiology [Vezzani & Granata 2005].

Inflammation in animal models of epilepsy

Additional evidence of a role of inflammation in seizures comes from animal models. Chronic inflammation was shown to trigger spontaneous seizures in mouse models of pneumococcal meningitis, cerebral malaria, or cysticercosis. Additionally, inflammation induced by LPS, for example, is known to enhance the effect of proconvulsive drugs (such as kaininc acid), an effect blocked by anti-inflammatory drugs [Vezzani & Granata 2005]. In a different study, systemic infection caused by *Shigella dysenteriae* was reported to enhance the seizure inducing effect of pentylenetetrazol, an effect mediated by TNF-α and interleukin-1β which a subsequent study revealed to be non-monotonic as a function of the TNF-α concentration [Yuhas *et al.* 2003].

TNF-α and epilepsy

Within the CNS, TNF-α may originate from microglia, infiltrating leukocytes, astrocytes and neurons. In its soluble form, TNF-α signaling is mediated the receptors p55 and p75, both present on neurons (p55 is said to be predominant in CNS cells) [Nadeau & Rivest 1999]. TNF-α expression is region specific, but in absence of inflammation the baseline concentration is low [Pan *et al.* 1997]. Brain TNF-α levels increase in a range of CNS disorders, such as bacterial meningitis, cerebral malaria, trauma, ischemia and multiple sclerosis (MS), many of which are associated with increased risk of seizures [Vezzani 2005].

At present, the TNF-α influence on seizures remains controversial, which is not too surprising, given that the cytokine is part of a complicated molecular network, with multiple regulatory mechanisms running in parallel. After brain injury (or seizures), TNF-α can have either a protective or a damaging role, depending on the timing and extent of TNF-α activation and the presence of other mediators [Shohami *et al.* 1999]. Specifically, TNF-α can induce cell death when the p55 pathway is activated and can be neuroprotective and reduce the incidence of seizures, when activating the p75 receptor pathway [Balosso *et al.* 2005].

In animal models, systemic infection by *Shigella* causes seizures, an effect blocked by cytokine inactivation using TNF-α and IL-1β antibodies [Yuhas *et al.* 2003]. This effect was concentration dependent, being reduced for extremely high TNF-α concentrations, presumably due to a saturation in the p55 pathway, combined with an increased activation of p75. Additionally, the incidence of seizures induced by

pneumoccocal meningitis was decreased by a treatment using an inhibitor of a TNF-α-converting enzyme, which reduces the levels of soluble TNF-α [Meli *et al.* 2004]. Lastly, it was shown that mice overexpressing TNF-α and IL-1β in glia, but not neurons, are susceptible to seizures and neurodegeneration [Campbell *et al.* 1993, Akassoglou *et al.* 1997], suggesting that astrocyte-mediated synaptic scaling may play a role in the effect.

All in all, the outcome of pharmacological manipulations of TNF-α levels depends on a variety of factors including concentration, time-scale, activated pathway, and receptors involved. Typically, the response to infection is rapid, localized and reversible, meant to eliminate the pathogen; it is only in conditions of chronic inflammation that this response becomes detrimental and can lead to seizures.

5.2 Computational model

In order to investigate the relationship between TNF-α and epileptiform activity, we developed a computational model of the interaction of neurons and glial cells during homeostatic synaptic scaling. To the best of our knowledge, it is the first model to simulate a network of spiking neurons interacting with a population of glial cells. As a first attempt, our model does not aspire to capture the full complexity of neuron-glia interactions. Namely, we focused on TNF-α-mediated homeostatic regulation, and did not model other ways through which glial cells can influence neuronal excitability, such as the regulation of extracellular potassium [Kofuji & Newman 2004], or the release of glutamate, or ATP [Newman 2003].

For modeling the neuron-glia interactions during synaptic scaling, we considered a two-dimensional representation of a population of neurons and glial cells. An overview of the model is shown in Fig. 5.4. Within a lattice topology, we constructed a recurrent network of spiking neurons co-localized with a similarly organized glia population. As suggested by the results from [Beattie *et al.* 2002, Stellwagen & Malenka 2006], in our model glial cells estimate the total synaptic drive to neurons by glutamate spillover and respond by producing TNF-α. This TNF-α diffuses to neighboring neurons and its concentration controls the strength of synaptic connections by scaling all excitatory weights of a neuron in a multiplicative fashion. For simplicity, the translation of local glutamate levels into TNF-α and the adjustment of synaptic weights as function of TNF-α levels was modeled by two sigmoids, with parameters established by a calibration step (see Appendix C2).

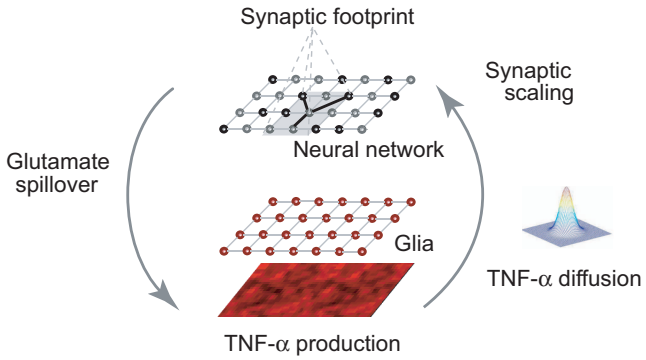

Figure 5.4. Schematic view of the homeostatic glia-mediated synaptic scaling. Glial cells estimate the synaptic drive received by neurons via glutamate spillover and adjust their TNF-α production as a function of the estimated glutamate level. Diffusing TNF-α reaches the neighboring neurons, triggering the multiplicative scaling of excitatory synapses.

5.2.1 Neural network

We used a two-dimensional spiking neuron model [Izhikevich 2003] that can produce a rich set of dynamical behaviors, while remaining computationally feasible:

$$\frac{dv(t)}{dt} = 0.04 \cdot v(t)^2 + 5 \cdot v(t) + 140 - u(t) + I(t) \quad (5.1)$$

$$\frac{du(t)}{dt} = a(b \cdot v(t) - u(t)), \quad (5.2)$$

where $v(t)$ is the membrane potential, $u(t)$ is the recovery variable, $I(t)$ the total post-synaptic current for the neuron, and a, b are model parameters, determining the dynamical behavior of the neuron.When the membrane potential reaches the threshold value $v_{\mathrm{th}} = 30$, a spike occurs, the membrane potential is reset to its rest value, and the recovery variable is updated: $v(t) \leftarrow c; \quad u(t) \leftarrow u(t) + d$, where c is the membrane rest potential, and d a parameter of the recovery variable. Different dynamic behaviors can be obtained by varying the parameters a, b, c, and d. Here, we consider only pyramidal neurons (regularly spiking RS, with $a = 0.02$, $b = 0.1$, $c = -65$, $d = 8$) and inhibitory interneurons (fast spiking FS, $a = 0.1$, $b = 0.2$, $c = -65$, $d = 2$).

AMPA and $\mathrm{GABA_A}$ synapses were modeled as exponentially decaying conductances g, which are instantaneously increased upon the arrival of an afferent spike:

$$\frac{dg(t)}{dt} = -\frac{g(t)}{\tau_{\mathrm{syn}}} + \bar{g} \sum_{\mathrm{i}} D(t_{\mathrm{i}}) \delta(t - t_{\mathrm{i}}), \quad (5.3)$$

where $\bar{g}$ represents the maximum synaptic increase per spike event (here, $\bar{g} = 1$), t_i is the time of the i^{th} spike, and $D(t_{\mathrm{i}})$ is a synaptic depression variable, described below. The time constant τ_{syn} gives the synaptic conductance decay, with default values 10 ms and 20 ms for AMPA and GABA_A synapses, respectively [Moreno-Bote & Parga 2005] (a more detailed description of the dependence of network dynamics on these parameters can be found in Appendix C1).

Additionally, excitatory synapses in our model exhibit short-term synaptic depression, which arises due to the temporary depletion of presynaptic vesicles, with an exponential recovery to baseline level [Tsodyks & Markram 1997, Abbott *et al.* 1997]:

$$\frac{dD(t)}{dt} = \frac{1 - D(t)}{\tau_{\mathrm{D}}} - U \sum_{\mathrm{i}} D(t_{\mathrm{i}}) \delta(t - t_{\mathrm{i}}), \quad (5.4)$$

where U is the fraction of synaptic resources that is consumed by a single event ($U = 0.05$) and τ_{D} is the time constant of recovery for the synaptic resources, in the range $450 - 700$ ms [Abbott *et al.* 1997]. For some of the experiments investigating network dynamics, we included also synaptic delays drawn from a uniform distribution with the range $[0, d_{\mathrm{max}}]$ ($d_{max} = 32$ ms).

The total input current I received by a neuron is obtained by summing up the post-synaptic currents for all incoming synapses:

$$I(t) = \sum_{\mathrm{i}} A_{\mathrm{i}} w_{\mathrm{i}} g_{\mathrm{i}}(t)(E_{\mathrm{i}} - v(t)), \quad (5.5)$$

where A_{i} is a scaling factor, with different values for a synapse connecting an excitatory/inhibitory neuron to an excitatory/inhibitory neuron (default values are $A_{\mathrm{EE}} = 0.02$ for excitatory-excitatory connections and $A_{\mathrm{EI}} = A_{\mathrm{IE}} = A_{\mathrm{II}} = 0.03$ for the rest), w_{i} is the strength of the i^{th} synapse ($w_{\mathrm{i}} \in [0, 1]$), $g_{\mathrm{i}}(t)$ measures the instantaneous synaptic conductance for synapse i, E_{i} is the reversal potential of the synapse (0 mV for excitatory synapses and -70 mV for inhibitory synapses), and $v(t)$ the membrane potential of the post-synaptic neuron.

The network used in our simulation consists of $N \times N$ neurons ($N = 25$), 80% RS excitatory and 20% FS inhibitory neurons, organized in a two-dimensional lattice. Neurons connect locally, within a square synaptic footprint of size $r \times r$ ($r = 7$), with probability $p_{\text{conn}} = 0.2$ and periodic boundary conditions. A weak excitatory input ($A_{\text{in}} = 0.03$) was provided from $N_{\text{in}} = 25$ input neurons, that connect to each neuron in the network with probability $p_{\text{in}} = 0.2$ and spike as independent Poisson processes with frequency $f_{\text{in}} = 10$ Hz.

5.2.2 Glia Model

The glial tissue was organized in a lattice similar to the neuron network. For simplicity, the model considers one glial cell per neuron. As biological evidence suggests a ration 1:4 between neurons and astrocytes, each of our glia cells actually models a small glia population neighboring a neuron. Consistent with the assumptions in [Stellwagen & Malenka 2006], glial cells monitor excitatory drive to neighboring neurons via glutamate spillover, considered to be proportional to the total drive received by the neuron at each position (x, y), averaged over a time window τ_{glut}:

$$\bar{I}_{\text{x,y}}(t) = \frac{1}{\tau_{\text{glut}}} \int_{t-\tau_{\text{glut}}}^{t} I_{\text{x,y}}(t)dt. \tag{5.6}$$

The neuronal activity is regulated by changes in the excitatory synaptic strength, corresponding to the AMPAR increases described experimentally. As the reported reduction in GABA_A receptor is smaller in relative magnitude [Stellwagen & Malenka 2006] and given that the analysis of the network properties reveals a weak dependence of the activity on A_{IE} (see Appendix A3), our model does not consider GABA_A receptor regulation. However, such regulation is expected to exacerbate the results described below (simulations showed no quantitative difference in the case when the same gain factors are considered for both the excitatory and the inhibitory regulation). Although experimental reports on synaptic scaling involve pyramidal neurons [Turrigiano *et al.* 1998, Ogoshi *et al.* 2005], the modeled scaling affects excitatory synapses on both pyramidal and inhibitory neurons, for simplicity.

To account for the astrocytic arborisation, the local glutamate estimates are convolved with a normalized Gaussian kernel: $C_{\text{glut}}(t) = \bar{I}(t) * G_{\sigma_1}$, where: $G_\sigma = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$, and $\bar{I}(t) = \left(\bar{I}_{\text{x,y}}(t)\right)_{\text{x,y}}$ is the matrix of glutamate estimates for all neurons at time t ($\sigma_1 = 1.22$, corresponding to a cell radius 3).

The TNF-α concentration is determined by the glial production, computed as function of the local glutamate concentration, with an exponential decay:

$$\frac{dc_{\text{tnf}}(t)}{dt} = -\frac{c_{\text{tnf}}(t) - c_{\text{tnf}\infty}(t)}{\tau_{\text{tnf}}} \tag{5.7}$$

$$c_{\text{tnf}\infty}(t) = 1 - \frac{1}{1 + e^{-\frac{c_{\text{glut}}(t) - c_{\text{glut0}}}{K_{\text{glut}}}}}, \tag{5.8}$$

where $c_{\text{tnf}}(t)$ is the local TNF-α concentration at time step t, $c_{\text{tnf}\infty}(t)$ represents target TNF-α concentration for the current activity level, and τ_{tnf} is a time constant for TNF-α concentration decay. The asymptotic value of the TNF-α concentration is a function of the glutamate concentration at a glial cell $c_{\text{glut}}(t)$, while the parameter c_{glut0} specifies the target glutamate concentration value, and K_{glut} is a scaling parameter (unless specified otherwise, $c_{\text{glut0}} = 0.52$, $K_{\text{glut}} = 2.5$).

The diffusion of TNF-α was modeled by the convolution with another Gaussian kernel G_{σ_2}, with the default value $\sigma_2 = 1.58$ (cell radius 4): $C' = C * G_{\sigma_2}$, where C and C' denote the matrices corresponding to the concentration at each glial cell c and neuron c'. The TNF-α triggers a change in the average synaptic conductance of the neuron $w(t)$, $w(t) = \frac{1}{N}\sum_{\text{i}} w_{\text{i}}(t)$, described by the equation:

$$\frac{dw(t)}{dt} = -\frac{w(t) - w_\infty(t)}{\tau_{\text{w}}}, \tag{5.9}$$

$$w_\infty(t) = \frac{1}{1 + e^{-\frac{c'(t) - c_{\mathrm{tnf0}}}{K_\mathrm{c}}}}, \tag{5.10}$$

where $c'(t)$ is the local TNF-α concentration at the neuron, τ_w gives the time scale of the synaptic strength change, c_0 and K_c are model parameters with default values $c_0 = 0.5$ and $K_\mathrm{c} = 0.03$.

The total change in conductance is distributed to the synapses w_i in a way that preserves their relative strength: $\frac{dw_\mathrm{i}(t)}{dt} = \frac{dw(t)}{dt} \frac{w_\mathrm{i}(t)}{\sum_\mathrm{i} w_\mathrm{i}(t)}$.

While the estimation of synaptic drive and the production of TNF-α by glial cells are very slow processes, such that observable changes occur on time scales of minutes to days [Stellwagen & Malenka 2006], the model uses much faster time constants in order to reduce simulation time. Specifically, for all results presented, the homeostatic regulation of synaptic strength occurs on the order of seconds rather than hours or days (time constants for glutamate estimation τ_glut, the glial TNF-α production τ_tnf and the synaptic modification τ_w were reduced to 1s, 10s and 1s, respectively). However, since the time constants of the homeostatic plasticity are still much slower than the activity dynamics of the spiking network, the qualitative behavior of the overall model should be preserved.

Currently, little is known —in quantitative terms— about the processes underlying TNF-α production and the corresponding synaptic scaling. Further experiments are needed to constrain the model parameters relating neuronal activity to TNF-α production and increase in synaptic strength. In order to set these 'free' parameters, we used a calibration procedure that ensured that the synaptic scaling robustly maintains homeostasis of the neuronal activity after changes in input (see Appendix C2 for details).

5.3 Results

5.3.1 Increases in TNF-α levels can induce seizures

As mentioned before, CNS inflammation is known to increase the risk of seizures in various diseases [Vezzani & Granata 2005], suggesting that TNF-α or other pro-inflammatory cytokines might be part of the mechanisms inducing an increased susceptibility to seizures during brain infections (e.g. bacterial meningitis). To investigate this hypothesis, we modeled chronic inflammation in the brain.

It is established that during chronic inflammation the TNF-α concentration inside the brain can increase [Gutierrez *et al.* 1993, Hanisch & Kettenmann 2007]. Additional TNF-α can either be produced locally, or can originate in serum and penetrate the bbb. Local production is owed mostly to activated microglia, but additional TNF-α sources are monocytes or lymphocytes, which can enter the brain due to changes in bbb permeability [Prat *et al.* 2005]. Also, systemic infections can trigger the activation of immune system agents inside the brain, in response to endotoxemia [Rivest *et al.* 2000, Galic *et al.* 2008], for example.

To keep things simple, the increase in TNF-α levels caused by immune system activation was modeled as an additional spatially homogeneous TNF-α source, by adding a small constant amount of TNF-α at each time step to the local glial production. Due to the fact that synaptic weights were initialized at random, this disruption was induced only after the network has reached the homeostatic regime ($t = 150$ s). The consequences of this manipulation can be seen in Fig. 5.5 A-C.

If the contribution from external TNF-α is big enough ($> 10^{-6}$, see Fig. 5.5 B), the homeostatic mechanism is no longer able to compensate for it by reducing local production. Consequently, the TNF-α concentration remains elevated, causing an increase in average synaptic conductance, and a corresponding rise in average neuronal firing rates (Fig. 5.5 C). Additionally, as excitatory synapses are strengthened, the activity becomes increasingly synchronized (a detailed analysis of how neuronal synchronization depends on the scaling of excitatory synapses can be found in Appendix C1). Transient increases in the input, due to normal activity fluctuations, are amplified by the recurrent excitatory connections to full seizure-like bursts (similar to those in Fig. 5.6 A). As suggested by experimental findings [Nita *et al.* 2006], regular spiking neurons exhibit spike bursts (Fig.5.6 B, lower), which synchronize over the population.

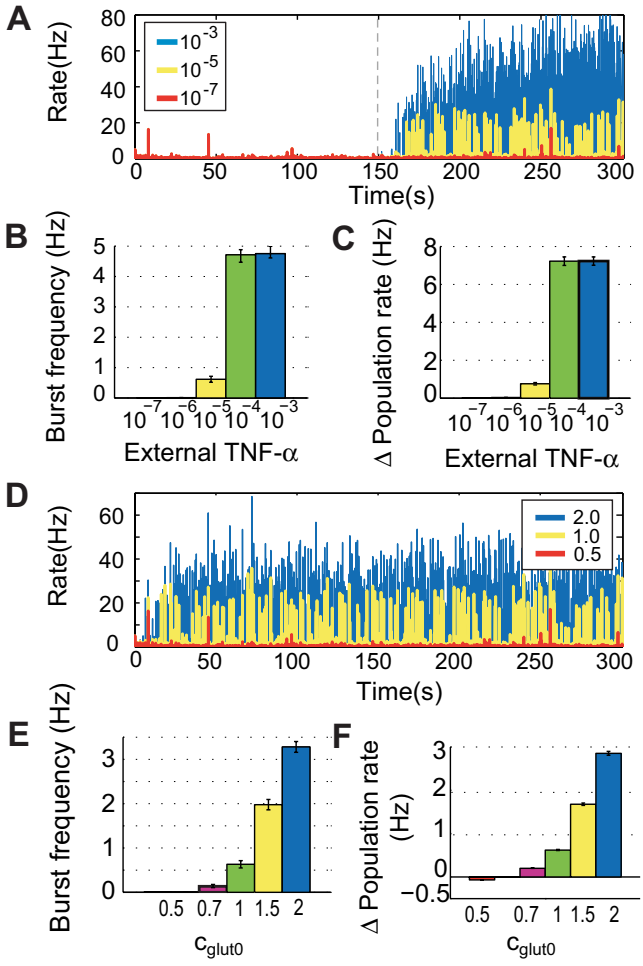

Figure 5.5. Effects on network activity due to a TNF-α increase. A) Changes in average neuronal activity after chronic inflammation (different external TNF-α values displayed in different colors). The severity of the immune response affects B) burst frequency and C) induces an increase in the average firing rate. D) Changes in average neuronal activity in the case of TNF-α overexpression by glia (for different c_{glut0} values). The degree of overexpression affects E) burst frequency and F) the average firing rate. Results are averaged over 10 trials, with error bars indicating standard error.

For high values of external TNF-α, the effect saturates (Fig. 5.5 B), presumably due to neuronal spike frequency adaptation and to synaptic depression (see Sec. Model-Neural network). The relation between average excitatory synaptic strength and the development of network bursts was investigated in detail in Appendix C1. Importantly, when analyzing a reduced population-level model of the system (Appendix C3), we showed that system robustness to additional TNF-α sources depends on the parameters K_{glut} and K_{C}, a result confirmed in simulation. Specifically, the stable state of the system is determined by the total gain of the feedback loop (i.e. by $k_{\text{glut}} * k_{\text{C}}$). The stability of the system to a certain destabilizing scenario, such as a chronic immune response, depends critically on the individual parameters, however. For a mild inflammatory state, the network can either remain stable, or develop strong seizure-like activity patterns, as function of K_{C} (see Fig. C.8).

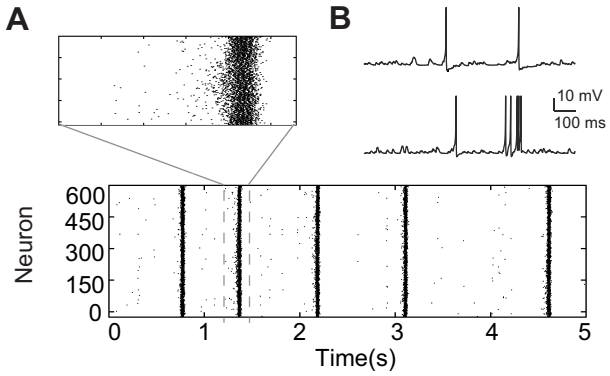

Figure 5.6. Network seizures. B) Typical example of network bursts occurring after a 10% increase in excitatory synaptic strength. B) Voltage traces of a neuron in the control case (upper) and during a network burst (lower).

An additional experimental finding captured by the model is the fact that transgenic mice mildly overexpressing TNF-α can develop spontaneous seizures [Akassoglou *et al.* 1997] (strong overexpression is usually fatal). To model various levels of TNF-α overexpression, we simulated an increase in the target glutamate level of the system (determined by parameter c_{glut0}). This manipulation forces glia to produce more TNF-α than in the control case. Similar to the effects observed experimentally, the network reaches a hyperexcitable state, with the probability of developing seizures being related to the degree of TNF-α overexpression. As shown in Fig. 5.5 D-E, if the change in synaptic strength is big enough, the network activity becomes paroxysmal.

5.3.2 A sustained reduction in input can trigger seizures

Simulations used for the calibration of system parameters, which involve changes in input firing rate (see Appendix A3), suggest that the dynamics of the network can change qualitatively due to chronic changes in the input rates. Therefore, we also investigated the degree of network synchronization for different input frequencies. As before, the system was first allowed to converge to the homeostatic state (150s), after which the mean input frequency was changed in a step function-like fashion to different levels and the frequency of the network bursts was evaluated (see Fig. 5.7).

After a sudden but sustained reduction in input, the network initially fell almost silent, but later recovered its normal activity level by a strengthening of excitatory synapses due to synaptic scaling. However, as a result of the adaptation, the network generated bursts at irregular intervals, provided that the remaining input is sufficient to drive the network (the remaining input was insufficient to trigger any activity in the 2Hz case, see Fig. 5.7 B). Correspondingly, after a sudden increase in input, the network responded with an initial burst of activity, after which it slowly returned to a low activity regime without any seizures. In both cases, the homeostatic mechanism brought the average firing rate of the neuronal population back to baseline (Fig. 5.7 C). These results extend those from a recently published work [Fröhlich *et al.* 2008], in which only a simplified one-dimensional network structure is considered, with local connections that facilitate burst propagation.

5.3.3 Local lesions cause seizures

The experiments involving variations in input rates described above suggest that the dynamics of the network can change qualitatively for a long-term reduction in input. Thus, it is plausible to assume that,

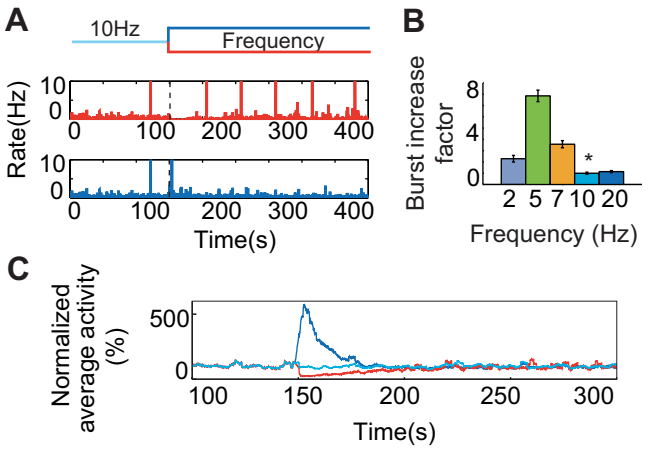

Figure 5.7. Network adaptation following changes in input frequency. A) Time evolution of average population rate after a +/-50% change in mean firing of the input population. B) Increase in burst probability relative to baseline (marked by *). C) A low-pass filtered version of the average firing rate. Results are averaged over 10 trials, with error bars representing the standard error.

as previously suggested by both experimental and theoretical evidence [Houweling *et al.* 2005, Timofeev *et al.* 2000], brain lesions can enhance seizure predisposition. In contrast to previous approaches using a global homeostatic regulation, here we focus on effects due to our specific mechanisms of synaptic scaling, specifically on spatial effects owed to TNF-α diffusion.

We considered localized lesions in the shape of a square with varying size. As before, the network was initialized randomly. After synaptic scaling stabilized ($t = 150\ s$), deafferentation was induced, causing a certain percentage of the synapses from the input neurons to neurons within the lesion area to be removed (Fig. 5.8 A). We studied network dynamics following the lesion for various lesion sizes and different degrees of synaptic damage.

When varying the proportion of deafferentation for a fixed lesion size (15×15), the strength of remaining synapses reached a value large enough to facilitate network bursts in several cases (Fig. 5.8 B). The burst probability increased with the severity of the lesion (Fig. 5.8 C), consistent with results in [Houweling *et al.* 2005]. Furthermore, the effect was observable for a wide range of partial deafferentation levels, as observed experimentally [Timofeev *et al.* 2000]. Similarly, when the lesion affected a fixed percentage of the input synapses (100%), the network dynamics depended on the size of the affected area. For small lesions the effect was negligible, but increased with the size of the lesion, as shown in Fig. 5.8 E. For almost complete deafferentation, however, seizure-like behavior disappeared. As neurons in our model have no intrinsic spontaneous activity, the network is driven by the external input. Consequently, when the lesion damages a large number of input connections, the remaining drive is insufficient to generate any activity (also, the overall firing rate decreases, as shown in Fig. 5.8 F). It may trigger large bursts if the input configuration is just right, but most times the network is quiescent.

Investigating the evolution of average excitatory strength within and outside the lesion area (Fig. 5.8 G) revealed a potential explanation for the effect. Due to the decrease in glutamatergic input, glia cells within the lesion area start producing TNF-α and the average excitatory strength increases (to a value higher than before the lesion, as the rate of the input is higher than the population firing rate). However, TNF-α diffusion causes a similar increase of weights outside the lesion, which would not have needed this regulation.

As the increased TNF-α production within the lesion affects the neighboring 'healthy' neurons, we

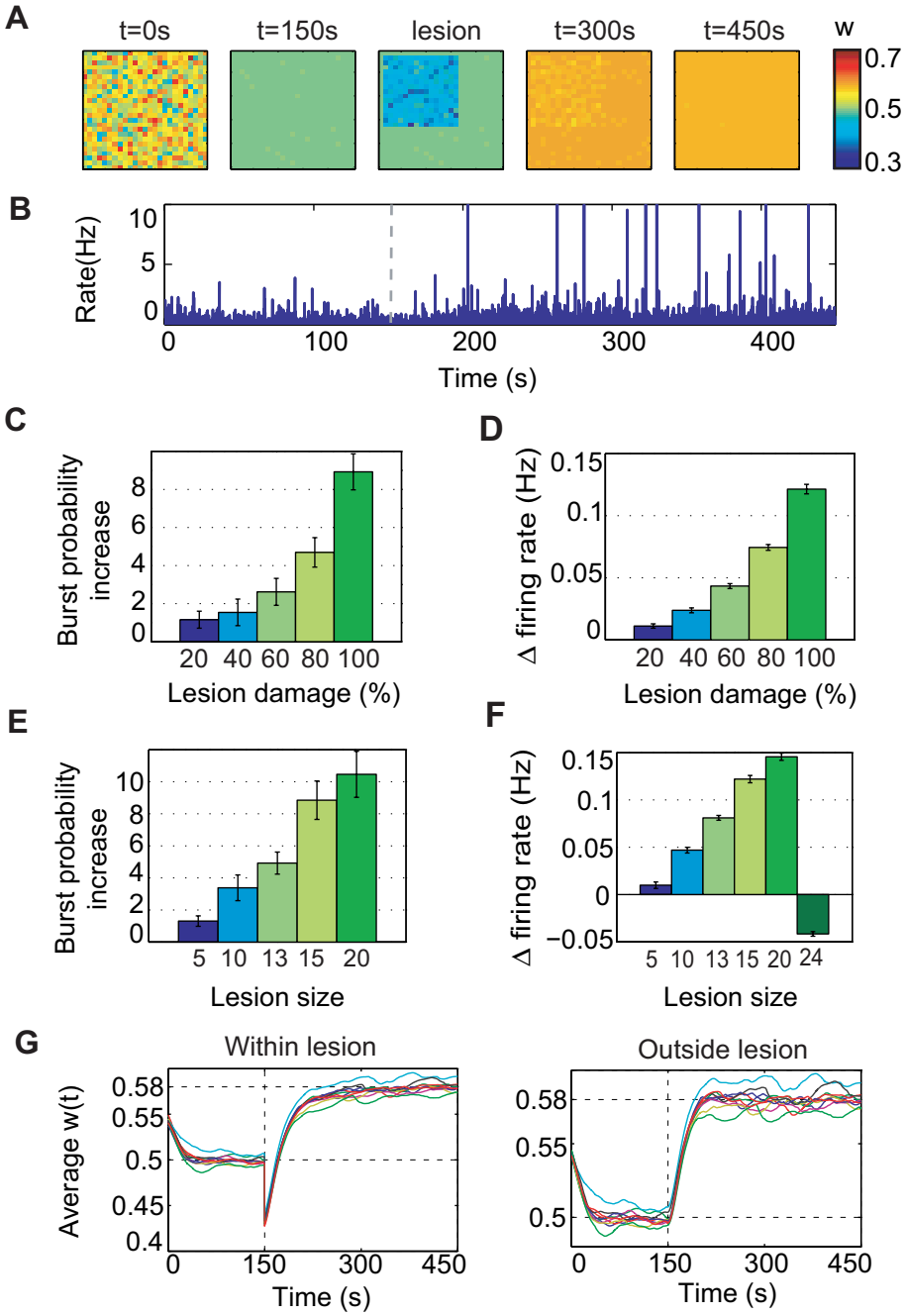

Figure 5.8. Effects of local lesions. A) Evolution in time of average synaptic strength for individual neurons after a local lesion. B) Corresponding average firing rate for the neuronal population. C) Increase in burst frequency and D) firing rate relative to baseline for different lesion severity levels and E,F) lesion sizes. G) Time evolution of average excitatory synaptic strength for the neurons within and outside lesion, for a lesion size 15×15 and 100% damage of input synapses. Different colors mark different trials. Results C-F) estimated over 10 trials, with error bars measuring the standard error.

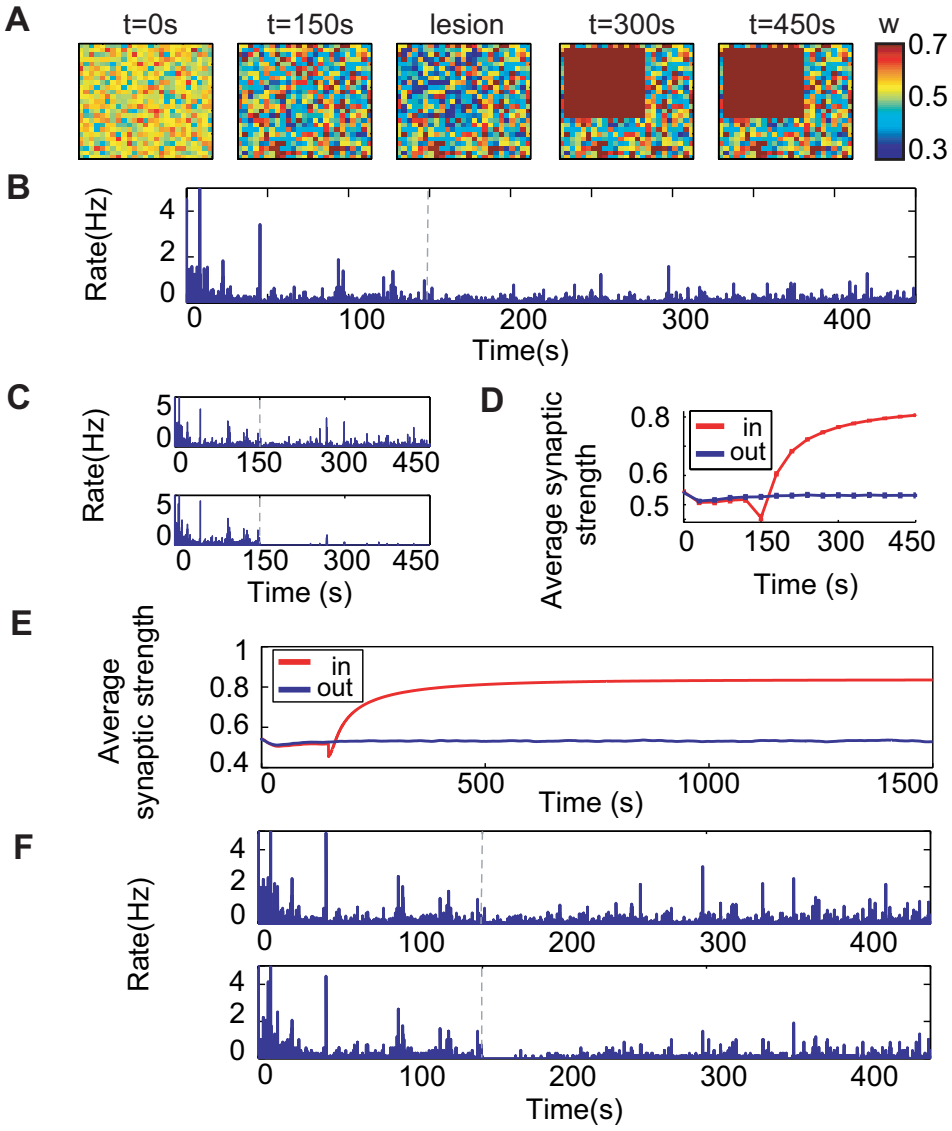

Figure 5.9. Effects of local lesions for local synaptic scaling. A) Evolution in time of average synaptic strength for individual neurons during a 100%, 15×15 local lesion. B) Corresponding average activity for the entire neuronal population, and C) for separate affected and unaffected neuron populations. D) Evolution of average weights within and outside the lesion. E) Single trial converge of average excitatory weights for a 100% lesion. F) Average activity within and outside a 80% 15×15 lesion.

assumed that TNF-α diffusion could play a role in the increased network synchrony. In order to test this hypothesis, we compared the dynamics of the network after 100% deafferentation, within a 15×15 area, for various diffusion coefficients. As predicted, when the scaling was restricted to the lesion area (Fig. 5.9 A), bursts disappeared completely, as shown in Fig. 5.9 B. Also, note that local synaptic scaling

leads to an increased variability in average excitatory strength for neurons (Fig. 5.9 A). For the case of 100% lesion, the damage was too severe for the homeostatic mechanism to be able to recover the original activity level (see Fig. 5.9 C). For a 80% lesion, activity within the lesion went back to baseline. (Fig. 5.9 F). Interestingly, convergence became significantly slower in the case of local regulation (see Fig. 5.9 D, E), suggesting that the spatial averaging of signals may be more advantageous, as it allows for increased system responsiveness.

Our results demonstrate that strengthening of synapses in neurons around the lesion due to TNF-α diffusion is responsible for the network hyperexcitability in the case of localized lesions. When diffusion does occur, the actual parameters controlling the astrocytic arborisation range and TNF-α diffusion were not very important and no systematic differences were observed in the system dynamics. Importantly, this result suggests that homeostatic regulation mechanisms that rely on the diffusion of neuromodulators are prone to become maladaptive in cases of localized disruptions in the system.

5.4 Discussion

The work presented here represents a first model of glial cells interacting with a population of neurons by homeostatic synaptic plasticity. Our model suggests that the dual role of TNF-α as both a pro-inflammatory messenger of the immune system and as a mediator of synaptic scaling can lead to interesting interactions. Specifically, it offers a novel mechanism through which immune activity in the brain influences the dynamics of cortical circuits increasing seizure susceptibility. In addition, our model builds on previous data linking homeostatic mechanisms and seizures [Houweling *et al.* 2005, Timofeev *et al.* 2000, Fröhlich *et al.* 2008] in the context of localized lesions. It implements a biologically-plausible mechanism for the synaptic up-regulation following deafferentation. Interestingly, in our model diffusion of TNF-α through cortical circuits was shown to be critical for the development of paroxysmal activity. This represents a clear distinction to previous models [Houweling *et al.* 2005, Fröhlich *et al.* 2008], which explicitly implement the homeostatic regulation of synapses as function of the global population activity (corresponding to a large TNF-α diffusion in our model). This is particularly relevant because the degree of locality of the process was shown to affect the final outcome after deafferentation. This effect could be enhanced by other means of spreading TNF-α from the lesion area, such as the range of the astrocytic arborisation (a result confirmed in simulations) or by glial communication, e.g. via gap junctions [Giaume & McCarthy 1996]. Currently, little is known about how TNF-α and other proteins diffuse through cortical tissue and more data is needed to constrain future models.

Recent experimental evidence [Galic *et al.* 2008] has shown that a LPS-induced infection in rats, occurring during a critical period in development, induces long-lasting increase in neuronal excitability and seizure susceptibility. The effect is mediated by TNF-α and can be mimicked by an intracerebral administration of rat recombinant TNF-α. The nature of the changes induced by a transient inflammatory response during development is still unclear. However, is interesting to note that, although the baseline TNF-α levels seem to not be altered in the adult, the cytokine levels following an induced seizure are increased in these animals (as observed in TLE patients with a specific genetic mutation), potentially also due to an increase in the number of astrocytes. In the context of our model, one could imagine that an initial inflammation within a critical period may alter the 'gain function' for the astrocytic TNF-α production, making the system more unstable and thus more prone to seizures. However, further experiments are needed to clarify whether TNF-α-mediated synaptic scaling plays a role in this case.

An interesting question in this context is why the immune system and the synaptic scaling mechanism rely on the same messenger protein, if this can lead to such unwanted cross-talk. One possible answer is that the immune and nervous systems are usually well isolated from one another through the bbb. In this case, it may be that the 'immune privilege' of the brain is an adaptation meant to minimize the cross-talk between the two systems [Boulanger 2009]. According to an alternative view, TNF-α has an immune-related role in the brain under normal circumstances. From this perspective, TNF-α is part

of a well balanced network of molecular mechanisms in the homeostatic state. It is only under chronic conditions that the excess of TNF-α turns harmful and increases seizure susceptibility, while short and local brain inflammation does not affect the stability of the system.

Taken together, our results illustrate that the reliance of immune signaling and synaptic scaling on the same messenger molecule, TNF-α, may be responsible for infection related seizures in a number of conditions. A great challenge for future experiments would be to carefully analyze the interference between the signaling pathways regulating an inflammatory response and homeostatic synaptic regulation. This is particularly important as it was shown that, when activating a different signaling pathway, the cytokine can have beneficial effects, improving neuronal survival by the release of neurotrophic factors [Akassoglou *et al.* 1997, Balosso *et al.* 2005]. In the case of a brain-immune system interference, beneficial pharmacological manipulations would ideally block the synaptic scaling mechanism, without disturbing the protective effects of TNF-α. A potential target for selectively disabling homeostatic regulation is the TNF-α receptor p55, as protective effects are mediated by a different receptor (p75) [Balosso *et al.* 2005], but is not yet clear how this can be achieved in practice.

6 Conclusions and future work

We have shown that the interaction between different plasticity mechanisms underlies several important aspects of cortical processing. One one side, it allows a network of stochastically spiking neurons to adapt to the statistics of its inputs such that it can transmit information more efficiently and this type of unsupervised learning explains receptive field development in sensory areas. On the other side, in the context of goal-directed behavior, similar mechanisms are needed for adapting an initially unstructured recurrent neural network to the constraints imposed by a particular task. Moreover, for the delayed response problem considered here, our learning procedure leads to representations similar to those observed experimentally, suggesting that reward-dependent learning may be a central driving force for the development of WM. Lastly, homeostatic plasticity can have important clinical implication. We have shown that, in certain pathological conditions, attempts to preserve the balance of the system by synaptic scaling may actually have detrimental effects and lead to seizures.

6.1 The role of homeostasis in learning

Regardless of the specific problem, our work emphasizes the role of homeostatic plasticity in neural computation. Both the development of receptive fields in sensory neurons and stabilizing the dynamics of recurrent neural networks during reward-dependent learning critically depend on the interaction between STDP, intrinsic plasticity and synaptic scaling. One one hand, IP enforces a sparse prior on the neural responses, enabling STDP to orient the receptive field of the neurons towards heavy-tailed directions in the input. On the other side, in the context of recurrent networks, a simpler implementation of IP is critical for stabilizing the dynamics of the system during learning. Additionally, in both cases, synaptic scaling constrains the weight vector to a fixed norm and provides a convenient means for bounding individual weights, compensating for Hebbian effects due to STDP.

Given that work in associative memory networks has long suggested that sparse coding can increase the memory capacity in Hopfield networks, by minimizing the overlap between the encoding of different patterns, e. g. [Kanerva 1993], it is interesting to think about the roles a more complex IP rule, similar to that used for implementing ICA, could have for the problem of reward-dependent learning introduced in chapter 4. This type of model adaptation has not yet been used in this context. However, an IP rule that changes the transfer function of rate neurons to make the output distribution Gaussian has been used, with some success, for reservoir adaptation in the context of echo-state-networks [Schrauwen *et al.* 2008]. As a different approach, it was recently shown that, for a specific class of neural networks, the more complex IP adaptation destabilizes fixed-point attractors (e. g. Hopfield networks), leading to complex dynamics, such as intermittent bursting or self-organized chaos [Markovich & Gros 2010].

6.2 Stability through diversity

Cortical processing involves a variety of Hebbian and homeostatic mechanisms, acting on different time and spatial scales. One can say that if experimentalists examine any aspect of cortical function long enough, they will find it undergoes some form of plasticity [Kim & Linden 2007].

From a computational perspective, the experimental data is rather puzzling, as often different forms of plasticity seem to overlap in function. It is likely that, although they may not be all strictly required computationally, this diversity of mechanisms is needed to ensure system robustness. The various plasticity mechanisms offer flexibility, such that the system is maintained in a functional range after disruptions affecting different cortical targets.

Cortical plasticity spans a wide range of temporal scales, from fast short term synaptic plasticity, acting on scales of hundreds of milliseconds, or LTP/LTD induced after a few minutes, to slow homeostatic synaptic scaling and intrinsic plasticity, with a time scale of hours to days, or even slower structural plasticity. It seems likely that this variety of temporal scales is needed to allow the system to respond accurately to a wide range of sensory perturbations and to permit a range of dynamic timescales needed for action. Additionally, from an engineering perspective, combining several integral feedback systems (different homeostatic mechanisms) with different temporal scales could be an efficient way to regulate cortical activation without significant oscillations, or overshoots.

Along the same lines, the degree of spatial selectivity of different Hebbian and homeostatic mechanisms is unclear. We have seen in our model for TNF-α mediated synaptic scaling that homeostatic regulation mechanisms which rely on the diffusion of neuromodulators are prone to become maladaptive in cases of localized disruptions in the network. In such cases, more localized homeostatic plasticity could provide a mechanism to compensate for the system overexcitability and maintain stable function. However, the question still remains why such global regulation would be needed in the first place.

One aspect only briefly investigated in our work is the plasticity of inhibitory networks. We know from experimental data that, unlike excitatory neurons, interneurons do not attempt to preserve their firing rates at a certain set-point [Rutherford *et al.* 1998]. However, it is likely that this asymmetry is need to coordinate for maintaining higher order network function [Maffei & Fontanini 2009]. Though seemingly different, the behavior of both excitatory and inhibitory neurons still works towards maintaining system homeostasis. If homeostatic mechanisms would try to push both populations towards a certain fixed point in a cell-autonomous way, this would negatively affect network dynamics. If interneurons would scale up their response after periods of inactivation, this would reduce the activation of excitatory neurons, triggering a new wave of compensatory processes. Hence, the homeostasis of inhibitory neurons has to be sacrificed for the stability of the overall network response.

6.3 Future work

There are several directions in which the work presented here could be extended further. First, since our ICA implementation was shown to work even when input information was encoded as spike-spike correlations, it would be an interesting challenge to investigate to which extent the mechanism presented here generalize to other types of input encoding.

From another perspective, several extensions can be imagined for a neuron population. Preliminary work using rate neurons [Gerhard *et al.* 2009] suggests that topographic ICA could be implemented by limiting the range of the adaptive lateral inhibition to a certain neighborhood of the neuron. Moreover, given that the output of several neurons remains somewhat correlated after learning, we hypothesize that a subsequent network layer, receiving this activity as input, could discover some higher-order structure in the input. So far, attempts to construct such iterative or hierarchical models for ICA have had little success [Shan *et al.* 2007], as linear ICA models require some additional nonlinear processing for converting the output of one layer into a suitable input for the subsequent one. The nonlinear model

proposed in chapter 3 naturally works with non-negative, heavy-tailed distributions for both input and output, making our network a feasible building block for more complex hierarchical structures.

Many questions remain also in relation to our working memory model. First, as currently the memory capacity of the network is rather limited, the most important issue is to mathematically explore the theoretical limitations of the model. Additional improvements might be obtained through additional plasticity mechanisms. Specifically, we have only begun to explore the role of adaptive inhibition. Inhibitory IP could already improve the representation of different inputs, by making sure that encoding resources are distributed evenly between different stimuli. However, the role of STDP at inhibitory synapses remains still unclear.

In the framework of our simple model, we have been able to show that the connectivity of recurrent networks can be optimized in a task-specific way to solve a problem that would have been difficult to solve with a similar randomly connected network. This seems to suggest that our reward-dependent learning scheme could be used for problem-specific reservoir optimization. Along these lines, it would be interesting to generalize our findings to a network more similar to traditional LSM networks.

Lastly, our results seem to suggest that the representation of stimuli at the neuronal level can depend on the specific task a network needs to perform. It is tempting to think that reward-dependent learning could provide an unifying explanation for different coding schemes reported in the literature [Knutsen & Ahissar 2009].

A

ICA with spiking neurons

A.1 Mathematical derivation: IP rule

The intrinsic plasticity model used for spike-based ICA changes neuronal excitability such that the distribution of the instantaneous firing rate of the neuron becomes exponential [Joshi & Triesch 2009]. This type of IP has been previously implemented for a Hodgkin-Huxley [Stemmler & Koch 1999] and a rate model using a sigmoid transfer function [Triesch 2007]. The mathematical derivation here follows closely that in [Triesch 2007], for a different transfer function.

From an information theoretic perspective, this type of output distribution is interesting because, among all distributions defined on the interval $[0, \infty)$, the exponential has maximum entropy, for a fixed mean. In our case, due to the refractory period of the neuron, the output firing rate is bounded. However, the exponential can be still a good approximation in this case, under the assumption that the mean firing rate of the neuron is small.

Optimizing information transmission under the constraint of a fixed mean is equivalent to minimizing the Kullback-Leibler divergence between the neuron's firing rate distribution and that of an exponential with a given mean μ [Triesch 2007]:

$$\begin{aligned} D &= d(p_{\text{neuron}} \| p_{\text{exp}}) && \text{(A.1)} \\ &= \int f_Y(y) \log\left(\frac{f_Y(y)}{\frac{1}{\mu} e^{-y/\mu}}\right) dy && \text{(A.2)} \\ &= -H(Y) + \frac{1}{\mu} E(Y) - \log(\mu), && \text{(A.3)} \end{aligned}$$

with $y = g(u)$, $H(\cdot)$ denoting the entropy and $E(\cdot)$ the expected value. Using $p_Y(y) = \frac{dg}{du} \cdot p_U(u)$, the entropy term above can be rewritten as:

$$H(Y) = -E\left(\log(p_Y(y))\right) = E\left(\log\left(\frac{dg}{du}\right) + E(p_U(u))\right), \tag{A.4}$$

with the last term being constant for a given input distribution.

Considering φ as one arbitrary model parameter, the gradient of the KL-divergence becomes:

$$\frac{\partial D}{\partial \varphi} = \frac{1}{\frac{dg}{du}} \frac{\partial}{\partial \varphi}\left(\frac{\partial g}{\partial u}\right) + \frac{1}{\mu} E\left(\frac{\partial g}{\partial \varphi}\right). \tag{A.5}$$

For our neuron model, we have:

$$\frac{d\partial g}{\partial u} = \frac{r_0}{u_\alpha}\left(1 - e^{-\frac{g}{r_0}}\right). \tag{A.6}$$

Replacing the derivatives in the equation before, for each of the three model parameters, yields:

$$\frac{\partial D}{\partial r_0} = E\left(\frac{1}{r_0}(1-\frac{g}{\mu})\right), \tag{A.7}$$

$$\frac{\partial D}{\partial u_0} = E\left(\frac{1}{u_\alpha}\left(\left(1+\frac{r_0}{\mu}\right)\left(1-e^{-\frac{g}{r_0}}\right)-1\right)\right), \tag{A.8}$$

$$\frac{\partial D}{\partial u_\alpha} = E\left(\frac{1}{u_\alpha}\left(\frac{u-u_0}{u_\alpha}\left(\left(1+\frac{r_0}{\mu}\right)\left(1-e^{-\frac{g}{r_0}}\right)-1\right)-1\right)\right). \tag{A.9}$$

Using stochastic gradient descent, the above gradients yield set of update rules for the neuron parameters:

$$\Delta r_0 = \frac{\eta_{\mathrm{IP}}}{r_0}\left(1-\frac{g}{\mu}\right), \tag{A.10}$$

$$\Delta u_0 = \frac{\eta_{\mathrm{IP}}}{u_\alpha}\left(\left(1+\frac{r_0}{\mu}\right)\left(1-e^{-\frac{g}{r_0}}\right)-1\right), \tag{A.11}$$

$$\Delta u_\alpha = \frac{\eta_{\mathrm{IP}}}{u_\alpha}\left(\frac{u-u_0}{u_\alpha}\left(\left(1+\frac{r_0}{\mu}\right)\left(1-e^{-\frac{g}{r_0}}\right)-1\right)-1\right). \tag{A.12}$$

with η_{IP} being a small learning rate.

A.2 Model robustness

The ICA learning model proposed in chapter 3 is extremely robust to changes in a variety of parameters. In the case of the bars problem, for example, the neuron will always learn a bar, independent of initial conditions (w_{tot}, initial transfer function parameters r_0, u_0 and u_α), the IP-enforced mean firing rate ($\mu \in [1, 10]$ Hz), variations in the learning rate for IP ($\eta \in 10^{-4}$, 10^{-7}), the STDP parameters (the threshold between potentiation and depression ν, varied by changing the ratio A_+/A_- or one of the time constants $\tau_\pm$, see Methods), or variations in the average probability of a bar ($[1/2N, 1/N]$). Moreover, similar receptive fields can be obtained with different synaptic plasticity rules, such as additive [Song *et al.* 2000] or simple triplet [Pfister & Gerstner 2006] STDP.

As seen in Fig. A.1 A and B, a single bar is learned for different input sizes (5 to 25) for the original bars problem. Moreover, the rule handles equally well more difficult variants of the bars problem [Lücke & Sahani 2008], in which samples consist always of the same number of bars (e.g. 4 in Fig. A.1C), or in which bars that are two-pixel wide, non-overlapping (Fig. A.1D), emphasizing the nonlinearity of the superposition.

A.3 Learning with a two-dimensional input

To build more intuition, and to show that the analytical results hold also for distributions with infinite support, we repeat the two-dimensional input experiment for a Gaussian and a Laplacian distribution:

$$p_{U_1}(u_1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u_1^2}{2}}, \tag{A.13}$$

$$p_{U1}(u_2) = \frac{1}{\sqrt{2}} e^{-\sqrt{2}|u_2|}. \tag{A.14}$$

As before, the PDF for the total input is $p_U(u) = p_{U_1}(u_1) * p_{U_2}(u_2)$, which, unfortunately, does not have a closed-form solution. However, as in [Triesch 2007], we can generate samples from the distribution

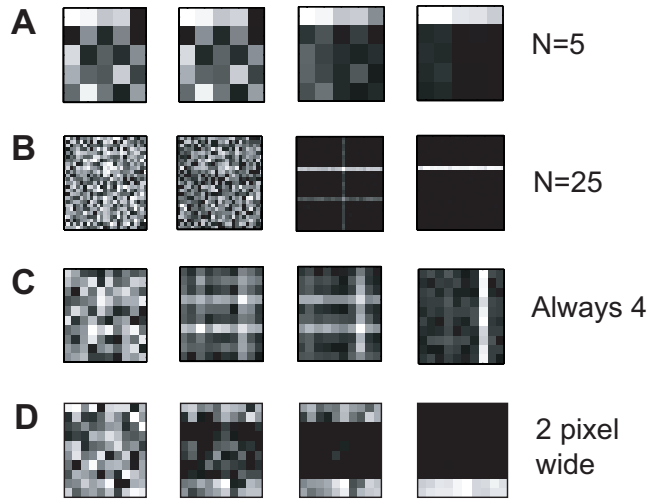

Figure A.1. Learning an IC. (A) and (B) The original bars problem, with different input sizes $N = 5$ and $N = 25$, respectively, (C) Modified bars problem, in which each sample consists of exactly 4 superimposed bars, (D) Modified bars problem, in which bars are two pixel wide and each sample consists of exactly 2 superimposed bars. Except for the varied variable (N), all model parameters are those used also in the main text.

and obtain an estimate of the optimal transfer function by our IP rule. For the approximated optimal parameters $\tilde{r}_0$ $\tilde{u}_0$ $\tilde{u}_\alpha$, we can again estimate the changes in weights with nearest-neighbor STDP (see [Izhikevich & Desai 2003] and the Methods section in Chapter 3) as:

$$\Delta w_{\mathrm{i}} \propto u_{\mathrm{i}} \tilde{g}(u) \left(\frac{A_+}{\tau_+^{-1} + \tilde{R} g(u)} + \frac{A_-}{\tau_-^{-1} + \tilde{r} g(u)} \right), \tag{A.15}$$

with $u = w_1 u_1 + w_2 u_2$, $i = 1, 2$. The proportionality factor is given by the learning rate of the STDP rule and by the average value of the refractory state of the neuron $R(t)$ (as before, we consider $\tilde{R} \approx 1$).

In Fig. A.2, we plot the change in weights for different input samples, ordered by the corresponding neuron activation. One can easily note that significant changes in weights occur only on the tail of the output distribution and that they are much larger for the heavy-tailed directions in the input. Importantly, this property does not depend on the initial state of the weight vector (except for the extreme case where $w_1 \gg w_2$, where the gradient becomes very flat, making it difficult for the IP rule to converge). It holds also when weights change by classic Hebbian learning, as in the analytical case considered before. Consequently, when the STDP and IP rules are acting together, the neuron will learn to respond preferentially to the Laplace direction (u_2) in the input, independent of the initial value of w (Fig. A.2F).

Maximizing kurtosis

In the case of multiple heavy-tails in the input, it is expected that ICA would select one single such direction, as we have demonstrated for the demixing problem with a two-dimensional input. This can be explained mathematically by the fact that the kurtosis of a linear mixture of several nongaussian distributions is always smaller than that of each of its component distributions [Hyvärinen *et al.* 2001]. Hence, maximizing kurtosis—a very common objective for ICA—means reducing the weighted sum of several directions to a single direction, by making all but one of the weights equal to zero. It is easy

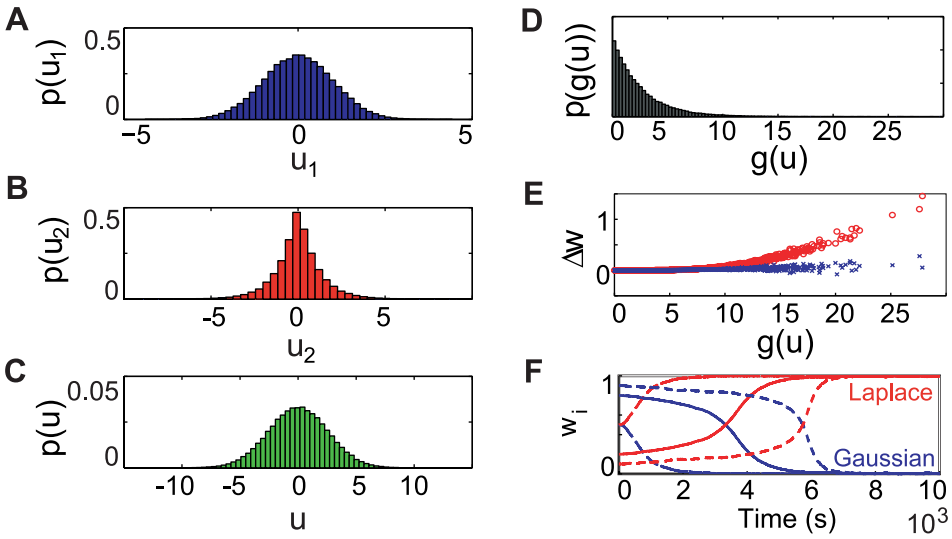

Figure A.2. A two dimensional input space, both directions with infinite support. (A) and (B) Probability distributions for the two independent directions, (C) The distribution of total input for $w = (0.6, 0.4)$, (D) Output distribution of the neuron, after the IP rule has converged, (E) Change in weights for the Gaussian (u_1, blue crosses) and Laplacian (u_1 red circles) directions, ordered by the instantaneous firing rate of the output neuron, (F) The actual evolution of the weight vector when the IP and STDP rule are acting in parallel (w_1 in blue, w_2 in red, different dashes show variations in initial conditions). The weight vector aligns with the heavy-tailed direction u_2, independent of initial conditions.

to show that the same principle applies to our rule (Fig. A.3A), for two independent Laplace inputs. Depending on the initial conditions, the neuron will develop sensitivity to one of the two directions. When the initial weights are equal (the solid lines), the winning input is selected at random.

Although the shape of the output distribution enforced by IP may suggest a bias for finding exponential tails in the input, we found that our rule generally selects the heaviest tail distribution in the input. Specifically, when the input consists of a Laplace and a Cauchy distribution, the latter is preferred, as required in ICA (see Fig. A.3B).

IP speeds up learning

We have seen before that a fixed transfer function does not facilitate learning for the bars problem. However, theoretical studies suggest that, with some preprocessing and an appropriate fixed nonlinearity, a receptive field could develop by nonlinear PCA [Oja 1997] in a rate neuron model for a linear superposition of independent sources. Even for the linear superposition, adapting the transfer function online may be advantageous, e.g. in terms of convergence time. To analyze this, we consider again the problem of learning the heavy tail direction in a two-dimensional input, with a Gaussian and Laplacian component, normalized to zero mean and unit variance. Although in this simplified case a fixed nonlinearity is sufficient for discovering the heavy-tailed direction, as predicted, learning is faster in the case when IP interacts with synaptic learning, compared to the case when the final transfer function obtained after learning is used (Fig. A.4).

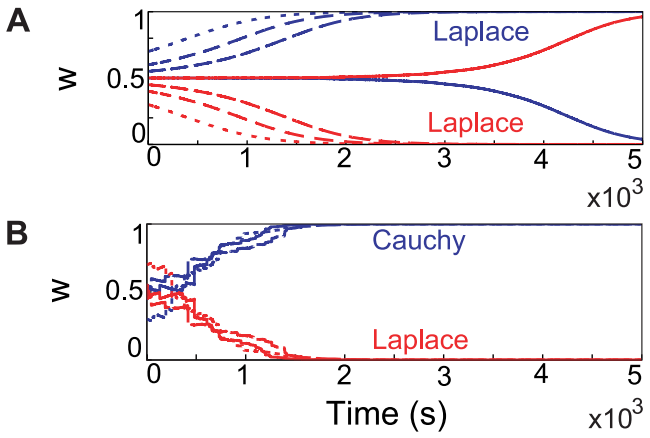

Figure A.3. Maximizing the kurtosis of the output by IP. (A) Evolution of the weight vector when the IP and STDP rule are acting in parallel for the case of two Laplacian inputs. As function of initial conditions, the learned receptive field prefers one of the directions, maximizing kurtosis. (B) In the case of a Cauchy (red) and Laplace (blue) distribution, the more heavy-tailed direction is preferred.

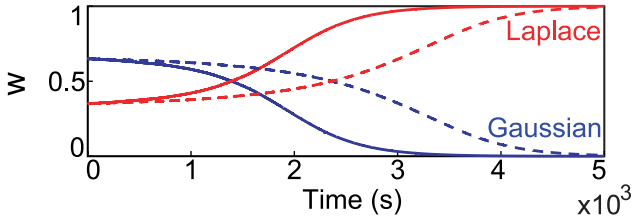

Figure A.4. IP speeds learning. Evolution of weights for a Gaussian and a Laplacian input, with (solid line) and without IP (dotted line). In the second case, the neuron transfer function is fixed to the parameters obtained by IP at the time of convergence.

A.4 A link between IP and BCM

From an information-theoretic perspective, it is interesting to relate our approach to previous work on maximizing information transmission between neuronal input and output by optimizing synaptic learning [Toyoizumi *et al.* 2005]. This optimization procedure results in a spike-based rule that implements a generalized version of the classic BCM rule. A link between BCM and IP-based learning has been noted previously [Triesch 2007], when it has been argued that the slow sliding threshold of BCM could play a homeostatic role similar to that of IP. Here we address this question again, from the perspective of learning.

For this comparison, we use the minimal triplets STDP model [Pfister & Gerstner 2006], with an additional sliding threshold. The model is a generalization of STDP learning, which takes into account spike triplets. It is mathematically equivalent to the model in [Toyoizumi *et al.* 2005], i.e. BCM-like, but computationally simpler. Specifically, the model computes one presynaptic (r_1) and two postsynaptic (o_1

and o_2) activity traces, with different time scales:

$$\frac{dr_1}{dt} = -\frac{r_1}{\tau_+} + \delta(t - t^f_{\text{pre}}) \tag{A.16}$$

$$\frac{do_1}{dt} = -\frac{o_1}{\tau_-} + \delta(t - t^f_{\text{post}}) \tag{A.17}$$

$$\frac{do_2}{dt} = -\frac{o_2}{\tau_y} + \delta(t - t^f_{\text{post}}) \tag{A.18}$$

where δ is the Dirac function and $t^f_{\text{pre/post}}$ is the time of firing of the pre- and post-synaptic neuron, respectively. The timescales of integration τ_+ and τ_- are similar to those in classical STDP ($\tau_+ = 16.8$ ms, $\tau_- = 33.7$ ms), while τ_y is slower ($\tau_y = 114$ ms).

The change in weight can be computed as:

$$\frac{dw}{dt} = A_3^+ \cdot r_1 \cdot o_2 \cdot y(t) - A_2^- \cdot o_1 \cdot x(t) \cdot \frac{\Theta}{\Theta_{\text{goal}}}, \tag{A.19}$$

with the parameters $A_3^+ = 6.5 \cdot 10^{-4}$, $A_2^- = 7.1 \cdot 10^{-4}$ and $\Theta_{\text{goal}} = 25$ is a scaling parameter which implicitly defines the goal mean firing rate of the neuron. For implementing sliding threshold BCM, the original triplet model model is enhanced with a slowly varying threshold Θ, which estimates a low-pass filtered version of the square of the postsynaptic firing rate, similar to the classical BCM sliding threshold [Pfister & Gerstner 2006]. Namely:

$$\frac{d\Theta}{dt} = -\frac{\Theta}{\tau_{\text{th}}} + o_2^2, \tag{A.20}$$

with $\tau_{\text{th}} = 100$ s.

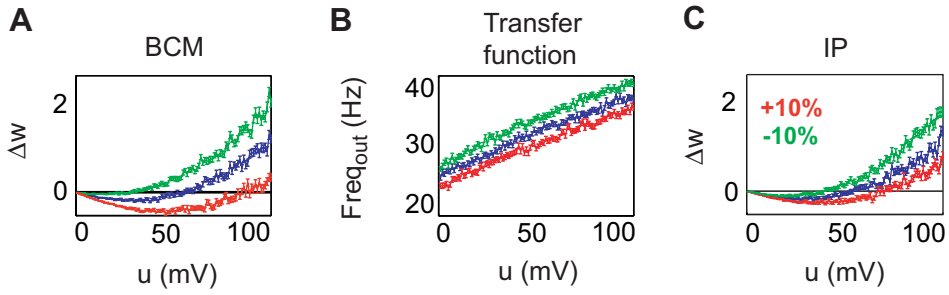

Figure A.5. Comparing sliding threshold BCM with IP (A) Weight changes for the minimal triplets model. Depending on the neuronal history (by $k = 1 \pm 0.1$), the threshold between potentiation and depression varies. (B) For the same variations in input (as before, the input is assumed to be Gaussian, with mean varying such that $x_1 = k \cdot x_0$) IP changes the neuronal transfer function to preserve the same neuronal output distribution ($u^1_\alpha = k \cdot u^0_\alpha$ and $u^1_0 = k \cdot u^0_0 - (k-1) \cdot u_r$, see Methods). (C) Corresponding to the shift of the transfer function, the weight changes for the same input vary. All estimates were averaged over 100 trials, each lasting 1 second.

When trying to compare IP-guided learning and sliding-threshold BCM one should note there is a fundamental difference between the two. While the sliding threshold of BCM has a direct effect on learning (by changing the threshold between potentiation and depression), IP affects plasticity only indirectly, through a change in neuronal output. As an illustration, consider the classic BCM weight change curve

in Fig. A.5A (blue curve), obtained with the minimal triplets STDP model described above. Depending on the history of the neuron, the threshold between potentiation an depression shifts, but the output of the neuron to the current sample remains the same. For the same alteration of the input distribution, IP causes a change in the transfer function (Fig. A.5B), which also results in a shift of the synaptic learning curve (Fig. A.5C), while the actual STDP threshold remains the same. The amplitude of this shift may depend on the IP parameters, but essentially it achieves the same type of alteration as sliding threshold BCM.

Given the similarity between our rule and BCM, and previous work showing the triplets STDP exhibit input selectivity [Toyoizumi *et al.* 2005, Pfister & Gerstner 2006], we hypothesized that a BCM rule with slowly-adapting threshold could also solve the bars problem. Indeed, as seen in Fig. A.6, spiking neurons by BCM-like spike-based synaptic learning can extract an independent component (IC). Moreover, as expected, after learning neuron responses become sparse and the output distributions highly kurtotic (see Fig. A.6C).

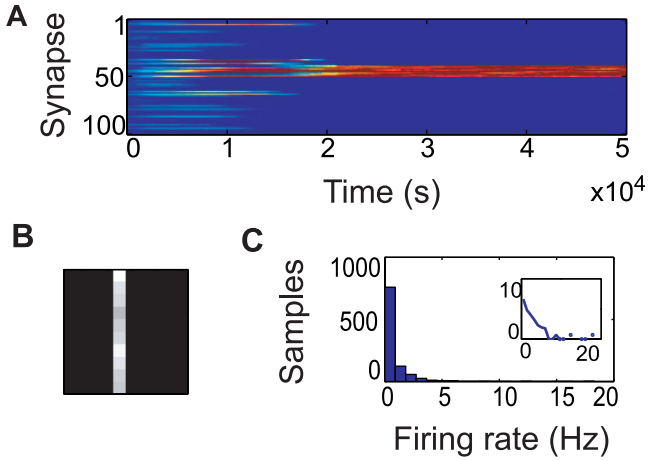

Figure A.6. Learning an IC with sliding threshold BCM. (A) Evolution of input weights for the bars problem. (B) Final receptive field learned. (C) Distribution of firing rates of the neuron; inset with the same measure in logarithmic scale.

Several constraints are necessary for the development of a stable receptive field with the sliding-threshold BCM rule. Firstly, as in the case of IP, synaptic scaling is required for stabilizing the receptive field. This is somewhat surprising, as BCM theory would predict that no constraint on the weight vector should be needed. Secondly, the parameter Θ_{goal} must be adjusted to the input characteristics for a bar to be a stable solution. The obvious advantage of IP is that it is robust to parameters changes and does not require tight tuning of any parameter. Moreover, directly optimizing the neuron excitability allows for various synaptic learning implementations, which may make a combined approach more suitable given different biological constraints.

A.5 RF estimation for the spike-based demixing

In order to recover the receptive field of the neuron after learning, we need to project the final weight vector back to the subspace defined by the input. Given that inputs can either be positive or negative but not both, this subspace can be easily defined by the constraint $w_i^{\text{on}} \cdot w_i^{\text{off}} = 0$, where the index i refers to the different input dimensions, and takes here the values 1 or 2. Hence, the problem of recovacrering the receptive field of the neuron can be cast as a constraint optimization problem: we want to find the vector $\tilde{w} = (\tilde{w}_1^{\text{on}} \tilde{w}_1^{\text{off}} \tilde{w}_2^{\text{on}} \tilde{w}_2^{\text{off}})$ from the subspace defined by the constraint above, such that the distance between $\tilde{w}$ and the original vector w is minimized. For simplicity, we consider here the Euclidean distance as metric.

Our cost function is then:

$$D(\tilde{w}) = \sum \left[(\tilde{w}_i^{\text{on}} - w_i^{\text{on}})^2 + (\tilde{w}_i^{\text{off}} - w_i^{\text{off}})^2 \right],$$

which needs to be minimized under the constraints $\tilde{w}_i^{\text{on}} \cdot \tilde{w}_i^{\text{off}} = 0$, with $i = \{1, 2\}$.

This problem can be solved using Lagrange multipliers. Namely, we define an alternative cost function:

$$L(\tilde{w}) = \sum_i \left[(\tilde{w}_i^{\text{on}} - w_i^{\text{on}})^2 + (\tilde{w}_i^{\text{off}} - w_i^{\text{off}})^2 \right] + \sum_i \lambda_i \tilde{w}_i^{\text{on}} \cdot \tilde{w}_i^{\text{off}}.$$

When computing the derivative of L, we can easily see that each pair of on- and off- weights give an independent system of equations, with identical form:

$$\begin{cases} 2 \cdot (\tilde{w}_i^{\text{on}} - w_i^{\text{on}}) + \lambda_i \cdot \tilde{w}_i^{\text{off}} &= 0 \\ 2 \cdot \left(\tilde{w}_i^{\text{off}} - w_i^{\text{off}}\right) + \lambda_i \cdot \tilde{w}_i^{\text{on}} &= 0 \\ \tilde{w}_i^{\text{on}} \cdot \tilde{w}_i^{\text{off}} &= 0 \end{cases}$$

By solving the above system, we obtain for each dimension solutions of the form: $\tilde{w}_i^{\text{on}} = w_i^{\text{on}}$ and $\tilde{w}_i^{\text{off}} = 0$, or alternatively $\tilde{w}_i^{\text{off}} = w_i^{\text{off}}$ and $\tilde{w}_i^{\text{on}} = 0$. Both solutions are extrema of the cost function L, but only one is our searched minimum. A quick check of the nature of each extremum shows that the final solution corresponds to the larger of the two weights. This means that the receptive field of the neuron can be computed by $|w_i| = \max\{w_i^{\text{on}}, w_i^{\text{off}}\}$, with the sign given by $\text{argmax}\{\text{w}_\text{i}^{\text{on}}, \text{w}_\text{i}^{\text{off}}\}$, i.e. positive if $w_i^{\text{on}} > w_i^{\text{off}}$, negative otherwise.

B

Working memory development

The stimulus representation which emerges as result of our learning procedure is very similar to a synfire chain [Abeles *et al.* 1993]. This assumes a feedforward underlying structure, consisting of multiple neuron pools which enables a volley of spikes to propagate synchronously from pool to pool.

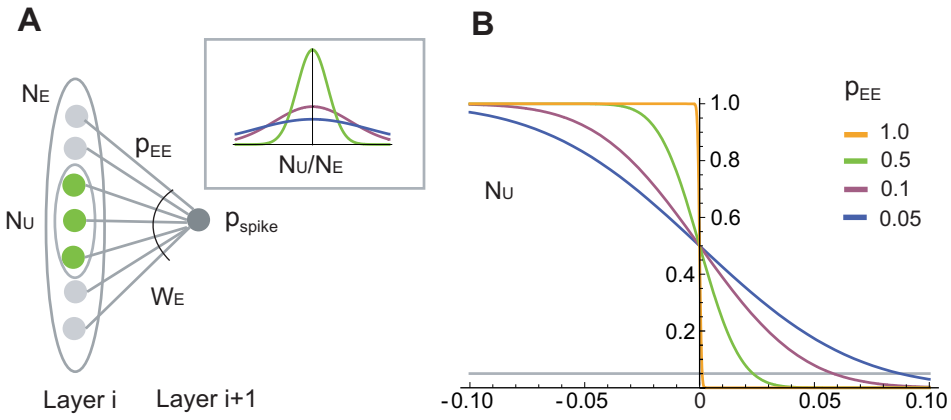

Figure B.1. Stable feedforward activity propagation. (A) We consider the propagation of activity between two pools of neurons in a synfire chain. If N_{U} out of N_{E} neurons are active in layer i, the total input distribution is a Gaussian with fixed mean and standard deviation dependent on p_{syn}.

It is possible to study the propagation probability of such a chain using the transmission function of the neurons [Diesmann *et al.* 1999]. As a special case, we consider only excitatory inputs from the previous layer (see Fig. B.1A). If N_{U} out of N_{E} neurons are active in layer i, and under our weight constraint $\sum w_j = 1$, the input distribution X of a neuron in layer $i+1$ is approximatively Gaussian, with mean $\mu = \frac{N_{\mathrm{U}}}{N_{\mathrm{E}}}$. The standard deviation of the distribution can be estimated in simulations, for randomly generated input weights. Not too surprinsingly, this value is very small for $p_{\mathrm{syn}} = 1$ and increases with sparser connectivity. Given the input distribution, the probability of firing of the neuron can be computed as:

$$\begin{aligned} p(\text{spike}) &= P(X > \theta) \\ &= \int_{\theta}^{\infty} \frac{1}{\sqrt{2\Pi}\sigma} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\ &= 0.5 \cdot \left(1 - Erf\left(\frac{x-\mu}{\sigma}\right)\right), \end{aligned}$$

where θ is the neuron threshold, σ is a function of p_{syn} and Erf is the error function.

Fig. B.1B shows the neuron's probability of firing, for different levels of connectivity sparseness. For a group of identical neurons, the firing rate of layer $i+1$ is equivalent to the firing probability of a single neuron. A stable synfire chain propagation can be achieved for the value of θ corresponding to the intersection between p(spike) and the horizontal line $y = \frac{N_{\text{U}}}{N_{\text{E}}}$. Assuming that $H = \frac{N_{\text{U}}}{N_{\text{E}}}$, our IP mechanism will always push θ towards this point.

In practice, due to sampling θ will oscillate around the 'true' value. In the case of all-to-all inter-layer connectivity, small perturbations are sufficient to dramatically change the value of p(spike), which makes this configuration unstable. Networks having a small p_{syn} the slope of the function describing the neuron's spike probability is much less steep, making it more robust to the unavoidable fluctuations of θ during learning. This suggests that sparse connectivity may be a desirable property for stabilizing the dynamics during learning.

In the case when inhibition is present, the inhibitory drive to the neuron can be similarly modeled as a Gaussian, with corresponding mean and variance. The total input distribution in this case will be another Gaussian distribution, with parameters given by the convolution of the excitatory and inhibitory components. In this case, one could imagine that the needed variability for the input distribution could alternatively be obtained by an appropriately setup of inhibitory connections. Further investigation is needed to determine exactly what is the optimal way to connect the inhibitory and excitatory populations for this purpose.

C

Synaptic scaling and epileptogenesis

C.1 Network parameters

We investigated the effects of changes in network parameters on the dynamics of the neural circuit without homeostatic regulation, focusing on the generation of seizure-like events. Seizure susceptibility was quantified in terms of probability or frequency of network bursts. Such a burst event was considered to occur whenever the population firing rate (estimated in 10ms time bins here) was higher than a given burst threshold (default, 10Hz). Here, we analyzed the effects of varying the connectivity, synaptic properties and input frequency in order to determine the conditions under which the network becomes hyperexcitable. These results were also used for optimizing the time bin width and the burst threshold for the synaptic scaling simulations (specifically, 30ms and 10Hz).

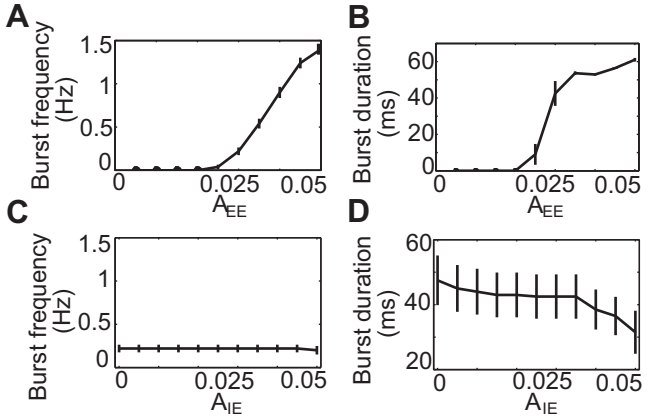

Figure C.1. The dependence of burst frequency and average duration on A, B) the strength of excitatory-to-excitatory, and C, D) inhibitory-to-excitatory synapses, respectively. Values estimated over 10 trials, with error bars indicating standard error.

We start by analyzing the importance of varying the synaptic strength of excitatory-excitatory connections. As the strength of excitatory synapses is increased, the network dynamics changes from a low firing to large synchronized population bursts. As shown in Fig. C.1 A and B, both the number and duration of burst events can be significantly altered by varying the strength between excitatory synapses. As high TNF-α concentrations scale up excitatory synapses, these results support the idea that for inflammatory responses which raises the TNF-α levels inside the brain sufficiently the network activity can become paroxysmal.

In order to determine the effects of inhibition, we study the impact of varying the amount of inhibition within the neuron population, i.e. the scaling factor for the inhibitory to excitatory synapses (A_{IE}). The results in Fig. C.1 C, D show a decrease in burst duration, but no significant effect on their frequency, in the parameter range considered.

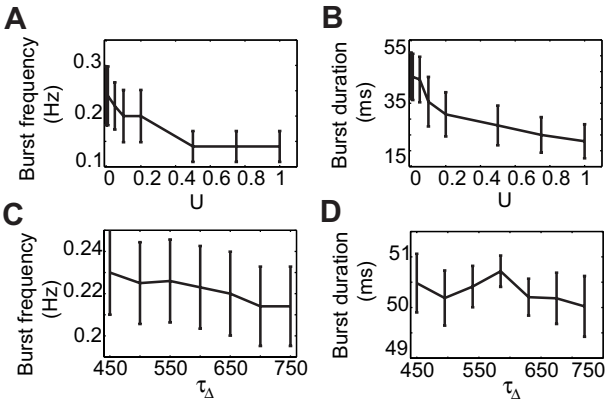

Figure C.2. The effects of synaptic depression on network bursts. Burst frequency and duration as a function of parameters A) and (A) U, and B) and D) τ_{D}, respectively.

It was shown that some amount of inhibition is required for preventing the network activity from 'exploding', but burst termination can also be achieved by other mechanisms. As synaptic depression reduces the efficacy of synapses in response to high activity, it is probable that the mechanism is also important for burst termination. To test this hypothesis, we study the parameters influencing synaptic resources consumption (U) and recovery (τ_{D}) and their effects on burst properties. The results illustrated in Fig. C.2 demonstrate that short-term synaptic depression plays a role in burst termination, consistent with the findings reported in [Houweling *et al.* 2005]. The fraction of resources consumed per burst event influences significantly the burst duration, as slower exhaustion of the synaptic resources makes bursts last longer. It also increases to some extent the burst frequency. The time constant of the recovery τ_D influences the burst frequency, since the longer it takes to recover the initial amount of synaptic resources, the longer it takes to trigger another network burst event.

Burst duration values suggested by experimental findings are typically larger than those produced in our model, in the order 200-400ms [Nita *et al.* 2006], as compared to 60-100ms in simulations. We hypothesize that the network size is too small for sustaining the synchronous activity for longer times. In order to test this assumption, we have systematically varied the network size, while maintaining the connectivity parameters constant. This variation is still insufficient to match experimental values, but burst duration is increased for larger networks, as seen in Fig. C.3. Synaptic delays can be a potential mechanism for prolonging bursts further. By introducing synaptic delays ($d_{max} = 32$ms) we have been able to prolong burst events by as much as 25-30%. Due to the computational overhead, most experiments do not consider synaptic delays, however.

The range of lateral connectivity also plays a role in the generation of bursts. When varying the range of connectivity and the connection probability such that the average incoming drive to neurons is kept constant, bursts occur only if the synaptic footprint is large enough, suggesting that excitatory loops are essential for the emergence of seizure-like behavior.

Due to the variability in times between burst events, it is likely that bursts are triggered by fluctuations in network input. In order to test this hypothesis, we compute a cross-correlogram of the instantaneous (bin size of 10ms) rates of input and the population rate of the network, on one hand, and of the inputs

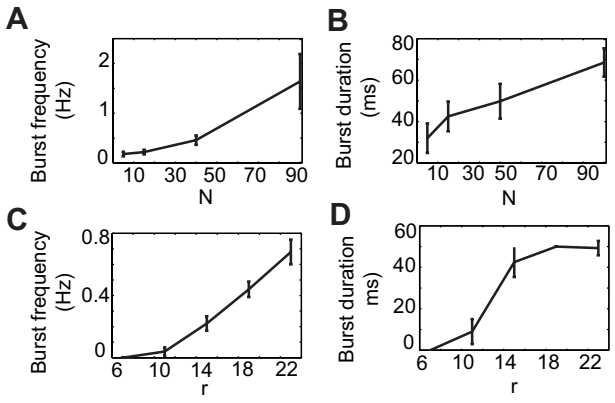

Figure C.3. Dependence of burst properties on A), B) network size and C), D) connectivity range.

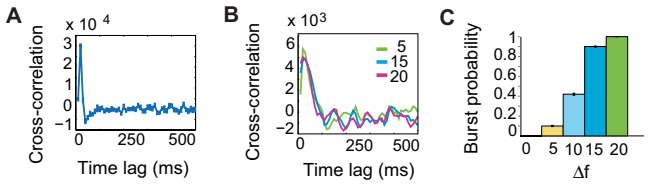

Figure C.4. Input-triggered synchronization A) Cross-correlation between input and population rate for a 5s trial, in bursting conditions, ($A_{\mathrm{EE}} = 0.03$). B) Cross-correlation between input rate and burst events, for transient increases the input frequency by an amount Δf ($A_{\mathrm{EE}} = 0.02$). C) Increase in burst probability, following a transient increase in input frequency.

and burst events, on the other. Both show a strong peak for the network at a time lag of about 20 ms, as shown in Fig. C.4 A. To further confirm that transient increases in input can, on their own, trigger bursts, we perform an additional experiment. We change the average frequency of the Poisson process in a short time window (of length 10ms) and monitor whether or not a burst occurs 100ms after this input increase. The cross-correlation of the input-output rates shows similar results to the initial experiment (Fig. C.4 B). Additionally, the burst probability increases with the increase in frequency during the window, further supporting the idea of bursts as input triggered events (Fig. C.4 C).

It is reasonable to assume that network bursts also lead to an increase in overall neuronal synchronization. More specifically, we analyze how network synchronization depends on average excitatory synaptic strength, the quantity affected by synaptic scaling (specifically, A_{EE} and A_{EI}). For different scaling factors, the network activity is analyzed (10 trials, each lasting 10s). A total of 1000 pairs of neurons are selected at random and the cross-correlation coefficient is computed for each pair. The average cross-correlation coefficients (see Fig. C.5, blue) show that strengthening of lateral excitatory synapses increases the synchronization of individual neurons in the network.

The synaptic time constants used in our model are larger than those used in similar experiments [Houweling *et al.* 2005]. However, these values do not play a critical role in the network dynamics. Theoretical work predicts that for an exponentially coupled network of excitatory and inhibitory neurons synchronized firing is facilitated by shorter synaptic decay times [Kanamaru & Sekine 2005]. We observe the same result in simulation for $\tau_{syn} = 5$ms, after changing the synapse scaling parameters such that

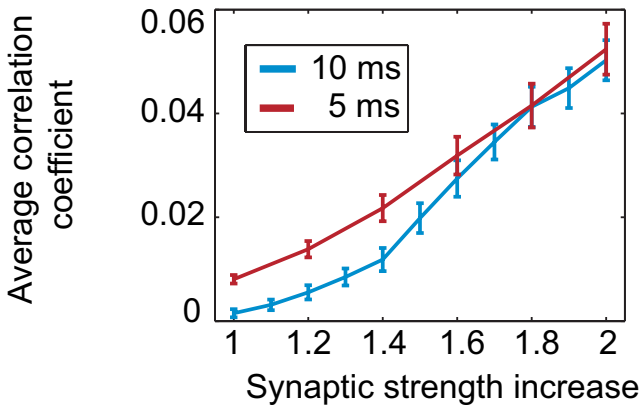

Figure C.5. Synchrony dependence on synaptic time constants. Average correlation coefficient of neurons for an increase of average excitatory strength relative to baseline, for different values of τ_{syn} (see text for details).

the total current of an ESP is preserved (namely, $A_{\text{EE}} = 0.04$, $A_{\text{EI}} = 0.06$, $A_{\text{IE}} = 0.12$, $A_{\text{II}} = 0.12$). Our experiments show an increase in synchronization at lower values of the excitatory coupling, for faster synaptic dynamics (see Fig. C.5, red).

To summarize the results above, in the neuronal network model considered bursts are triggered by transient increases in the input, which are amplified by the recurrent excitatory architecture to a full network burst. The activation of the local inhibitory population, together with the depletion of synaptic resources terminate a burst event.

C.2 Model calibration

Current experimental data is insufficient for fully determining the parameters of the processes involved in TNF-α mediated synaptic scaling. A reasonable constraint is that, under normal conditions, the synaptic scaling mechanism should be able to maintain homeostasis of neuron activity. In particular, the synaptic regulation should be able to preserve the average firing of the neurons in response to changes in the input firing rate.

The specific experiment performed involves slow ramp changes in input (see Fig. C.6 A). As previously, the system is allowed 150s to reach a homeostatic state before the input rates are modulated. The analysis focuses on two of the model parameters—the gain values for the TNF-α production as function of glutamate (K_{glut}), and the adjustment of excitatory synaptic strength as function of TNF-α concentration (K_{c}). The reciprocal of their product is an important parameter of the model and can be viewed as a total gain of the feedback loop (see next section).

As a calibration step, we select the gain that maintains the average population rate constant in time, independent of the change in input. For smaller gains, the homeostatic mechanism is not able to fully compensate for the changes in input, while for large gains the system overcompensates for these variations. As the gain defines a set of possible values for the two model parameters mentioned above, we select one such pair for the subsequent experiments. The particular choice can however be important for some pathological disruptions affecting the feedback loop (during an immune response, for example), and the robustness of the system to such events can be enhanced by larger values for K_{c} (see the analysis of the population behaviors below for a more detailed discussion).

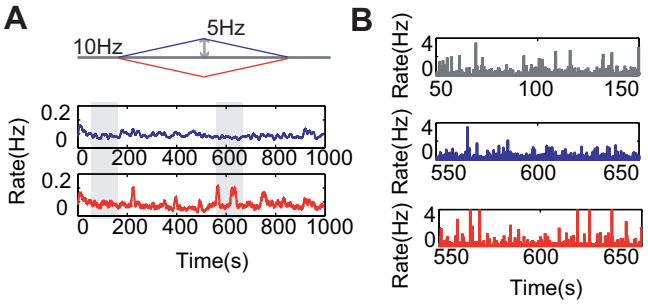

Figure C.6. Model calibration experiments. A) Homeostatic regulation of activity in response to changes in input. Left panel shows the low-pass filtered population firing rate after input increases (red) and decreases (blue). A closer view of the activity corresponding to the gray boxes is shown in the right panel. B) After chronic input reduction (red) the strengthening of excitatory synapses makes neuron firing more synchronized, compared to the homeostatic regime (gray) or the reverse input change (blue). Population rates are estimated in 30ms time bins.

For the selected parameters, the firing rates of the neuron population are maintained, as seen in Fig. C.6 A (the time average is computed by convolving the output firing rate with a Gaussian kernel, with $\sigma = 6s$). A more detailed study of the structure of the population firing (Fig. C.6 B) reveals similar firing patterns for the increase in input (blue) and constant firing (gray) regime. In the case of a chronic decrease in input (red), although the mean output rate is maintained the activity of individual neurons becomes more correlated, as the excitatory weights are increased to compensate for the input change.

C.3 Stability analysis

Although the homeostatic loop is not linear, looking at a large scale linear approximation of the system may enhance our understanding of its behavior. For a homogeneous connectivity structure, it is possible to consider a spatial average of the variables and construct a reduced model of the system, whose stability can be analyzed in the framework of linear control systems.

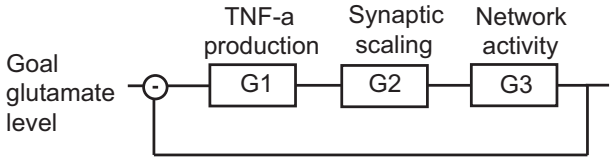

Figure C.7. Diagram of the large scale linear approximation of the system.

The block diagram of the reduced system is shown in Fig. C.7. Several assumptions are required for this approximation to be valid. Firstly, we consider the case of small variations around the fixed point. Hence, the two sigmoid functions that describe the processes translating glutamate in TNF-α and corresponding synaptic strength can be linearized around this fixed point. Secondly, we assume a linear mapping between the average value of excitatory synaptic strength and the glutamate level averaged for the neuronal population.

More specifically, when considering the asymptotic value of the TNF-α concentration, as function of

glutamate, linearized around the fixed point c_{tnf0}, together with the exponential decay, we obtain the transfer function for the TNF-α production as:

$$G_1(s) = \frac{\alpha_1}{\tau_{\text{tnf}} s + 1},$$

where $\alpha_1 = -\frac{1}{4K_{glut}}$ is the gain of the process.

Similarly, the transfer function for the synaptic strengths regulation is:

$$G_2(s) = \frac{\alpha_2}{\tau_{\text{w}} s + 1},$$

with $\alpha_2 = \frac{1}{4K_c}$.

Based on simulated data, we develop a phenomenological model of the network activity as function of average synaptic strength and input frequency. For this, we consider average synaptic weights in the range 0.4 to 0.6 and input frequencies between 8 and 12Hz (10 trials are considered for each parameter pair). For each trial, the strength of excitatory synapses of each neuron are multiplicatively scaled such that w is the same for all neurons. The glutamate levels are estimated for 30ms large bins and a low-pass filtered version of the signal is computed, similar to the averaging assumed for the glial cells. The resulting data points are fitted with a linear function by using the least-mean-squares method, resulting in a transfer function: $G_3(s) = \alpha_3$, with $\alpha_3 = 0.44$.

The transfer function of the system can be then computed as:

$$G(s) = \frac{G_3(s)}{1 + G_1(s)G_2(s)}$$

After replacing the expressions for the transfer functions above, the stability analysis is reduced to studying the poles of $G(s)$, i.e. the solutions s_1 and s_2 of the equation:

$$(1 + \tau_{\text{tnf}} s)(1 + \tau_{\text{c}} s) + \alpha_1 \alpha_2 \alpha_3 = 0$$

The roots have negative real parts in all cases, and the output is oscillatory for complex solutions, corresponding to:

$$(\tau_{\text{tnf}} + \tau_{\text{c}})^2 - 4\tau_{\text{tnf}}\tau_{\text{c}}(1 + \alpha_1\alpha_2\alpha_3) < 0$$

which reduces to a bound for the product $K_{\text{tnf}}K_{\text{c}} < 0.0643$. The analytical results are confirmed by simulations in the neuron network model. Firstly, experiments confirm that network dynamics do not change when the total gain of the feedback loop is maintained constant. Secondly, for the lesion experiment (80%, $r = 10$) lower gains ($K_{\text{tnf}}K_{\text{c}} = 0.125$) result in low burst probability (less than 10%), while for the underdamped regime ($K_{\text{tnf}}K_{\text{c}} = 0.0075$) bursts occur in 80% of the cases (results averaged over 10 trials), as predicted by the constraint on the product $K_{\text{tnf}}K_{\text{c}}$.

The actual values for the parameters K_{tnf} and K_{C} matter in the cases when disruptions are induced within the control loop, for example, during inflammation. In this case, although the total gain is preserved, the network dynamics change, depending on how sensitive the synaptic strength is to TNF-α fluctuations. For a mild inflammatory state, induced by an homogeneous TNF-α source with $c_{ext} = 10^{-5}$, the network can either remain stable, or develop strong seizure-like patterns of activity, as function of α_2 (see Fig. C.8). Note that network dynamics prior to the disruption are virtually indistinguishable for all pairs, as predicted by the population level analysis.

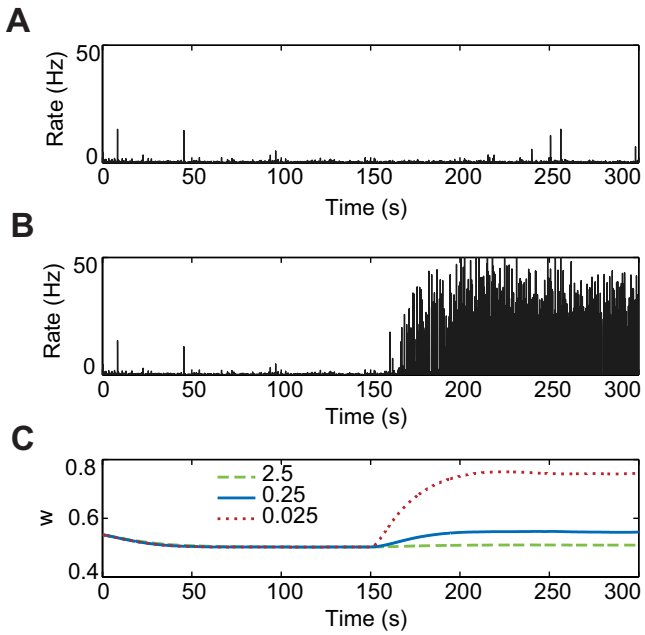

Figure C.8. Dependence of excitability increase following inflammation on K_c, for constant total gain. A) Population activity for $K_{\mathrm{C}} = 2.5$, single trial. B) Population activity for $K_{\mathrm{C}} = 0.025$, single trial. C) Time evolution of average excitatory strength for different values of K_{C}, averaged over 10 trials.

Bibliography

[Aarli 2000] J. Aarli. *Epilepsy and the immune system.* Arch. Neurology, vol. 57, pages 1689–1692, 2000.

[Abbott & Nelson 2000] L.F. Abbott and S.B. Nelson. *Synaptic plasticity: taming the beast.* Nature Neuroscience, vol. 3, pages 1178–1183, 2000.

[Abbott & Regehr 2004] L.F. Abbott and W.G. Regehr. *Synaptic computation.* Nature, vol. 431, no. 7010, pages 796–803, Oct 2004.

[Abbott *et al.* 1997] L.F. Abbott, J.A. Varela, K. Sen and S.B. Nelson. *Synaptic depression and cortical gain control.* Science, vol. 275, no. 5297, pages 220–224, Jan 1997.

[Abeles *et al.* 1993] M. Abeles, H. Bergman, E. Margalit and E. Vaadia. *Spatiotemporal firing patterns in the frontal cortex of behaving monkeys.* J Neurophysiol, vol. 70, no. 4, pages 1629–1638, Oct 1993.

[Abraham 2008] W.C. Abraham. *Metaplasticity: tuning synapses and networks for plasticity.* Nat Rev Neurosci, vol. 9, no. 5, page 387, May 2008.

[ADEAR 2008] ADEAR, editeur. Alzheimer's disease fact sheet. NIH, 2008.

[Aertsen *et al.* 1991] A. Aertsen, E. Vaadia, M. Abeles, E. Ahissar, H. Bergman, B. Karmon, Y. Lavner, E. Margalit, I. Nelken and S. Rotter. *Neural interactions in the frontal cortex of a behaving monkey: signs of dependence on stimulus context and behavioral state.* J Hirnforsch, vol. 32, no. 6, pages 735–743, 1991.

[Aertsen *et al.* 2001] A. Aertsen, M. Diesmann, M.O. Gewaltig, S. Grün and S. Rotter. *Neural dynamics in cortical networks–precision of joint-spiking events.* Novartis Found Symp, vol. 239, pages 193–204; discussion 204–7, 234–40, 2001.

[Aizenman & Linden 2000] C.D. Aizenman and D.J. Linden. *Rapid, synaptically driven increases in the intrinsic excitability of cerebellar deep nuclear neurons.* Nature Neuroscience, vol. 3, pages 109–112, 2000.

[Aizenman *et al.* 2003] C. Aizenman, C. Akerman, K. Jensen and H. Cline. *Visually Driven Regulation of Intrinsic Neuronal Excitability Improves Stimulus Detection In Vivo.* Neuron, vol. 39, no. 5, pages 831–842, August 2003.

[Akassoglou *et al.* 1997] K. Akassoglou, L. Probert, G. Kontogeorgos and G. Kollias. *Astrocyte-specific but not neuron-specific transmembrane TNF triggers inflammation and degeneration in the central nervous system of transgenic mice.* J Immunol, vol. 158, no. 1, pages 438–445, Jan 1997.

[Alkon *et al.* 1985] D.L. Alkon, M. Sakakibara, R. Forman, J. Harrigan, I. Lederhendler and J. Farley. *Reduction of two voltage-dependent K+ currents mediates retention of a learned association.* Behav Neural Biol, vol. 44, no. 2, pages 278–300, Sep 1985.

[Alkon 1984] D.L. Alkon. *Calcium-mediated reduction of ionic currents: a biophysical memory trace.* Science, vol. 226, no. 4678, pages 1037–1045, Nov 1984.

[Allen & Barres 2009] Nicola J Allen and Ben A Barres. *Neuroscience: Glia - more than just brain glue.* Nature, vol. 457, no. 7230, pages 675–677, Feb 2009.

[Alvarez & Sabatini 2007] V.A. Alvarez and B.L. Sabatini. *Anatomical and physiological plasticity of dendritic spines.* Annu Rev Neurosci, vol. 30, pages 79–97, 2007.

[Amit & Brunel 1995] D. Amit and N. Brunel. *Learning internal representations in an attractor neuron network with analogue neurons.* Network, vol. 6, 1995.

[Amit & Brunel 1997] D. Amit and N. Brunel. *Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex.* Cereb Cortex, vol. 7, no. 3, pages 237–252, 1997.

[Andersen *et al.* 1980] P. Andersen, S.H. Sundberg, O. Sveen, J.W. Swann and H. Wigström. *Possible mechanisms for long-lasting potentiation of synaptic transmission in hippocampal slices from guinea-pigs.* J Physiol, vol. 302, pages 463–482, May 1980.

[Antonov *et al.* 2001] I. Antonov, I. Antonova, E.R. Kandel and R.D. Hawkins. *The contribution of activity-dependent synaptic plasticity to classical conditioning in Aplysia.* J Neurosci, vol. 21, no. 16, pages 6413–6422, Aug 2001.

[Antonov *et al.* 2003] I. Antonov, I. Antonova, E. Kandel and R.D. Hawkins. *Activity-dependent presynaptic facilitation and hebbian LTP are both required and interact during classical conditioning in Aplysia.* Neuron, vol. 37, no. 1, pages 135–147, Jan 2003.

[Armano *et al.* 2000] S. Armano, P. Rossi, V. Taglietti and E. D'Angelo. *Long-term potentiation of intrinsic excitability at the mossy fiber-granule cell synapse of rat cerebellum.* J Neurosci, vol. 20, no. 14, pages 5208–5216, Jul 2000.

[Attwell & Laughlin 2001] D. Attwell and S.B. Laughlin. *An energy budget for signaling in the grey matter of the brain.* J Cereb Blood Flow Metab, vol. 21, no. 10, pages 1133–1145, Oct 2001.

[Baddeley *et al.* 1997] R. Baddeley, L. Abbott, M. Booth, F. Sengpiel, T. Freeman, E. Wakeman and E. Rolls. *Responses of neurons in primary and inferior temporal visual cortices to natural scenes.* Proceedings Biological Sciences, vol. 264, no. 1389, pages 1775–1783, 1997.

[Baeg *et al.* 2003] E.H. Baeg, Y.B. Kim, K. Huh, I. Mook-Jung, H.T. Kim and M.W. Jung. *Dynamics of population code for working memory in the prefrontal cortex.* Neuron, vol. 40, no. 1, pages 177–188, Sep 2003.

[Baeg *et al.* 2007] E.H. Baeg, Y.B. Kim, J. Kim, J.W. Ghim, J.J. Kim and M.W. Jung. *Learning-induced enduring changes in functional connectivity among prefrontal cortical neurons.* J Neurosci, vol. 27, no. 4, pages 909–918, Jan 2007.

[Balosso *et al.* 2005] S. Balosso, T. Ravizza, C. Perego, J. Peschon, I.L. Campbell, M.G. De Simoni and A. Vezzani. *Tumor Necrosis Factor-alpha Inhibits Seizures in Mice via p75 Receptors.* Ann Neurology, vol. 57, pages 804–812, 2005.

[Barbieri & Brunel 2008] F. Barbieri and N. Brunel. *Can attractor network models account for the statistics of firing during persistent activity in prefrontal cortex?* Front Neurosci, vol. 2, no. 1, pages 114–122, Jul 2008.

[Barlow 2001] H.B. Barlow. *Redundancy reduction revisited.* Network, vol. 12, pages 241–253, 2001.

[Barry *et al.* 2008] G. Barry, M. Piper, C. Lindwall, R. Moldrich, S. Mason, E. Little, A. Sarkar, S. Tole, R.M. Gronostajski and L.J. Richards. *Specific glial populations regulate hippocampal morphogenesis.* J Neurosci, vol. 28, no. 47, pages 12328–12340, Nov 2008.

[Beattie *et al.* 2002] E.C. Beattie, D. Stellwagen, W. Morishita, J.C. Bresnahan, B.K. Ha, M. Von Zastrow, M.S. Beattie and R.C. Malenka. *Control of Synaptic Strength by Glial TNF-α.* Science, vol. 295, pages 2282–2285, 2002.

[Bechmann *et al.* 2007] I. Bechmann, I. Galea and V.H. Perry. *What is the bloodbrain barrier (not)?* Trends in Neuroimmunology, vol. 28, pages 5–11, 2007.

[Bell & Sejnowski 1995] A.J. Bell and T.J. Sejnowski. *An information-maximization approach to blind separation and blind deconvolution.* Neural Comput, vol. 7, no. 6, pages 1129–1159, Nov 1995.

[Bell & Sejnowski 1997] A.J. Bell and T.J. Sejnowski. *The "independent components" of natural scenes are edge filters.* Vision Res, vol. 37, no. 23, pages 3327–3338, Dec 1997.

[Bhatt *et al.* 2009] D.H. Bhatt, S. Zhang and W.B. Gan. *Dendritic spine dynamics.* Annu Rev Physiol, vol. 71, pages 261–282, 2009.

[Bi & Poo 1998a] G.Q. Bi and M.M. Poo. *Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type.* J Neurosci, vol. 18, no. 24, pages 10464–10472, Dec 1998.

[Bi & Poo 1998b] G.Q. Bi and M.M. Poo. *Synaptic modifications in cultured hippocampal neurons: Dependence on spike timing, synaptic strength, and postsynaptic cell type.* J. Neurosci., vol. 18, page 10464 10472, 1998.

[Bienenstock *et al.* 1982] E. Bienenstock, L. Cooper and P. Munro. *Theory for the development of neuron selectivity: Orientation specificity and binocular interactions in visual cortex.* Journal of Neuroscience, vol. 2, pages 32–48, 1982.

[Blais *et al.* 1998] B. Blais, N. Intrator, H. Shouval and L. Cooper. *Receptive field formation in natural scene environments: comparison of single cell learning rules.* Neural Computation, vol. 10, pages 1797–1813, 1998.

[Bliss & Lomo 1973] T.V. Bliss and T. Lomo. *Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path.* J Physiol, vol. 232, no. 2, pages 331–356, Jul 1973.

[Boulanger 2009] L.M. Boulanger. *Immune proteins in brain development and synaptic plasticity.* Neuron, vol. 64, no. 1, pages 93–109, Oct 2009.

[Brembs *et al.* 2002] B. Brembs, F.D. Lorenzetti, F.D. Reyes, D.A. Baxter and J.H. Byrne. *Operant reward learning in Aplysia: neuronal correlates and mechanisms.* Science, vol. 296, no. 5573, pages 1706–1709, May 2002.

[Brody *et al.* 2003a] C. Brody, A. Hernández, Zainos A. and R. Romo. *Timing and neural encoding of somatosensory parametric working memory in macaque prefrontal cortex.* Cereb Cortex, vol. 13, no. 11, pages 1196–1207, Nov 2003.

[Brody *et al.* 2003b] C. Brody, R. Romo and K. Kepecs. *Basic mechanisms for graded persistent activity: discrete attractors, continuous attractors, and dynamic representations.* Curr Opin Neurobiol, vol. 13, no. 2, pages 204–211, Apr 2003.

[Brons & Woody 1980] J.F. Brons and C.D. Woody. *Long-term changes in excitability of cortical neurons after Pavlovian conditioning and extinction.* J Neurophysiol, vol. 44, no. 3, pages 605–615, Sep 1980.

[Brunel & Wang 2001] N. Brunel and X.J. Wang. *Effects of neuromodulation in a cortical model of object working memory dominated by recurrent inhibition.* Journal of Computational Neuroscience, vol. 11, pages 63–85, 2001.

[Buonomano & Maass 2009] D. Buonomano and W. Maass. *State-dependent computations: spatiotemporal processing in cortical networks.* Nat Rev Neurosci, vol. 10, no. 2, pages 113–125, Feb 2009.

[Burrell *et al.* 2001] B.D. Burrell, C.L. Sahley and K.J. Muller. *Non-associative learning and serotonin induce similar bi-directional changes in excitability of a neuron critical for learning in the medicinal leech.* J Neurosci, vol. 21, no. 4, pages 1401–1412, Feb 2001.

[Büsing *et al.* 2009] L. Büsing, B. Schrauwen and R. Legenstein. *Connectivity, Dynamics, and Memory in Reservoir Computing with Binary and Analog Neurons.* Neural Comput, Dec 2009.

[Butko & Triesch 2007] N. Butko and J. Triesch. *Learning sensory representations with intrinsic plasticity.* Neurocomputing, vol. 70, 2007.

[Calabresi *et al.* 2007] P. Calabresi, B. Picconi, A. Tozzi and M. Di Filippo. *Dopamine-mediated regulation of corticostriatal synaptic plasticity.* Trends in Neurosciences, vol. 30, no. 5, pages 211–219, 2007.

[Campanac & Debanne 2008] E. Campanac and D. Debanne. *Spike timing-dependent plasticity: a learning rule for dendritic integration in rat CA1 pyramidal neurons.* J Physiol, vol. 586, no. 3, pages 779–793, Feb 2008.

[Campbell *et al.* 1993] I. L. Campbell, C. R. Abraham, E. Masliah, P. Kemper, J. D. Inglis, M. B. Oldstone and L. Mucke. *Neurologic disease induced in transgenic mice by cerebral overexpression of interleukin 6.* Proc Natl Acad Sci U S A, vol. 90, no. 21, pages 10061–10065, Nov 1993.

[Caporale & Dan 2008] N. Caporale and Y. Dan. *Spike timing-dependent plasticity: a Hebbian learning rule.* Annu Rev Neurosci, vol. 31, pages 25–46, 2008.

[Carandini *et al.* 2005] M. Carandini, J.B. Demb, V. Mante, D.J. Tolhurst, Y. Dan, B.A. Olshausen, J. Gallant and N.C. Rust. *Do we know what the early visual system does?* J Neurosci, vol. 25, no. 46, pages 10577–10597, Nov 2005.

[Cassenaer & Laurent 2007] S. Cassenaer and G. Laurent. *Hebbian STDP in mushroom bodies facilitates the synchronous flow of olfactory information in locusts.* Nature, vol. 448, no. 7154, pages 709–713, Aug 2007.

[Celikel *et al.* 2004] T. Celikel, V.A. Szostak and D.E. Feldman. *Modulation of spike timing by sensory deprivation during induction of cortical map plasticity.* Nat Neurosci, vol. 7, no. 5, pages 534–541, May 2004.

[Chafee & Goldman-Rakic 1998] M.V. Chafee and P.S. Goldman-Rakic. *Matching patterns of activity in primate prefrontal area 8a and parietal area 7ip neurons during a spatial working memory task.* J Neurophysiol, vol. 79, no. 6, pages 2919–2940, Jun 1998.

[Charles *et al.* 1991] A. C. Charles, J. E. Merrill, E. R. Dirksen and M. J. Sanderson. *Intercellular signaling in glial cells: calcium waves and oscillations in response to mechanical stimulation and glutamate.* Neuron, vol. 6, no. 6, pages 983–992, Jun 1991.

[Cleary *et al.* 1998] L.J. Cleary, W.L. Lee and J.H. Byrne. *Cellular correlates of long-term sensitization in Aplysia.* J Neurosci, vol. 18, no. 15, pages 5988–5998, Aug 1998.

[Clopath *et al.* 2010] C. Clopath, L. Büsing, E. Vasilaki and W. Gerstner. *Connectivity reflects coding: A model of voltage-based spike-timing-dependent-plasticity with homeostasis.* Nature Neuroscience, vol. In press, 2010.

[Cohen *et al.* 1999] A.S. Cohen, C.M. Coussens, C.R. Raymond and W.C. Abraham. *Long-lasting increase in cellular excitability associated with the priming of LTP induction in rat hippocampus.* J Neurophysiol, vol. 82, no. 6, pages 3139–3148, Dec 1999.

[Collins *et al.* 1986] T. Collins, L. A. Lapierre, W. Fiers, J. L. Strominger and J. S. Pober. *Recombinant human tumor necrosis factor increases mRNA levels and surface expression of HLA-A,B antigens in vascular endothelial cells and dermal fibroblasts in vitro.* Proc Natl Acad Sci U S A, vol. 83, no. 2, pages 446–450, Jan 1986.

[Compte *et al.* 2000] A. Compte, N. Brunel, P. S. Goldman-Rakic and X. J. Wang. *Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model.* Cereb Cortex, vol. 10, no. 9, pages 910–923, Sep 2000.

[Constantinidis & Steinmetz 1996] C. Constantinidis and M. A. Steinmetz. *Neuronal activity in posterior parietal area 7a during the delay periods of a spatial memory task.* J Neurophysiol, vol. 76, no. 2, pages 1352–1355, Aug 1996.

[Cornell-Bell *et al.* 1990] A. H. Cornell-Bell, S. M. Finkbeiner, M. S. Cooper and S. J. Smith. *Glutamate induces calcium waves in cultured astrocytes: long-range glial signaling.* Science, vol. 247, no. 4941, pages 470–473, Jan 1990.

[Coulter *et al.* 1989] D.A. Coulter, J.J. Lo Turco, M. Kubota, J.F. Disterhoft, J.W. Moore and D.L. Alkon. *Classical conditioning reduces amplitude and duration of calcium-dependent afterhyperpolarization in rabbit hippocampal pyramidal cells.* J Neurophysiol, vol. 61, no. 5, pages 971–981, May 1989.

[Cox 1955] D.R. Cox. *Some statistical methods connected with series of events.* Journal of the Royal Statistical Society, Series B, vol. 17, no. 2, pages 129–164, 1955.

[Cunningham *et al.* 1996] A.J. Cunningham, C.A. Murray, L. O'Neill, M.A. Lynch and J.J. O'Connor. *Interleukin-1β (IL-1β) and tumour necrosis factor (TNF) inhibit long-term potentiation in the rat dentate gyrus in vitro.* Neuroscience Letters, vol. 203, pages 17–20, 1996.

[Dahmen *et al.* 2008] J.C. Dahmen, D. Hartley and A.J. King. *Stimulus-timing-dependent plasticity of cortical frequency representation.* J Neurosci, vol. 28, no. 50, pages 13629–13639, Dec 2008.

[Damas *et al.* 1992] P. Damas, D. Ledoux, M. Nys, Y. Vrindts, D. De Groote, P. Franchimont and M. Lamy. *Cytokine serum level during severe sepsis in human IL-6 as a marker of severity.* Ann. Surg., vol. 215, no. 4, pages 356–362, 1992.

[Dan & Poo 2006] Y. Dan and M. Poo. *Spike timing-dependent plasticity: from synapse to perception.* Physiol Rev, vol. 86, no. 3, pages 1033–1048, Jul 2006.

[Dan *et al.* 1996] Y. Dan, J.J. Atick and R.C. Reid. *Efficient coding of natural scenes in the lateral geniculate nucleus: experimental test of a computational theory.* J Neurosci, vol. 16, no. 10, pages 3351–3362, May 1996.

[Daoudal & Debanne 2003] G. Daoudal and D. Debanne. *Long-term plasticity of intrinsic excitability: learning rules and mechanisms.* Learn Mem, vol. 10, no. 6, pages 456–465, 2003.

[Daoudal *et al.* 2002] G. Daoudal, Y. Hanada and D. Debanne. *Bidirectional plasticity of excitatory postsynaptic potential (EPSP)-spike coupling in CA1 hippocampal pyramidal neurons.* Proc Natl Acad Sci U S A, vol. 99, no. 22, pages 14512–14517, Oct 2002.

[Daw & Doya 2006] N. Daw and K. Doya. *The computational neurobiology of learning and reward.* Curr Opin Neurobiol, vol. 16, no. 2, pages 199–204, Apr 2006.

[de Jonge *et al.* 1990] M.C. de Jonge, J. Black, R.A. Deyo and J.F. Disterhoft. *Learning-induced after-hyperpolarization reductions in hippocampus are specific for cell type and potassium conductance.* Exp Brain Res, vol. 80, no. 3, pages 456–462, 1990.

[Debanne *et al.* 1998] D. Debanne, B.H. Gähwiler and S.M. Thompson. *Long-term synaptic plasticity between pairs of individual CA3 pyramidal cells in rat hippocampal slice cultures.* J Physiol, vol. 507 (Pt 1), pages 237–247, Feb 1998.

[Deli *et al.* 1995] M. A. Deli, L. Descamps, M. P. Dehouck, R. Cecchelli, F. Joo, C. S. Abraham and G. Torpier. *Exposure of tumor necrosis factor-alpha to luminal membrane of bovine brain capillary endothelial cells cocultured with astrocytes induces a delayed increase of permeability and cytoplasmic stress fiber formation of actin.* Journal of Neuroscience Research, vol. 41, pages 717–726, 1995.

[Derkach *et al.* 2007] V. Derkach, M. Oh, E. Guire and T. Soderling. *Regulatory mechanisms of AMPA receptors in synaptic plasticity.* Nat Rev Neurosci, vol. 8, no. 2, pages 101–113, Feb 2007.

[Desai *et al.* 1999] N.S. Desai, L.C. Rutherford and G. Turrigiano. *Plasticity in the intrinsic excitability of cortical pyramidal neurons.* Nat Neurosci, vol. 2, no. 6, pages 515–520, Jun 1999.

[Diesmann *et al.* 1999] M. Diesmann, M. O. Gewaltig and A. Aertsen. *Stable propagation of synchronous spiking in cortical neural networks.* Nature, vol. 402, no. 6761, pages 529–533, Dec 1999.

[Disterhoft *et al.* 1986] J.F. Disterhoft, D.A. Coulter and D.L. Alkon. *Conditioning-specific membrane changes of rabbit hippocampal neurons measured in vitro.* Proc Natl Acad Sci USA, vol. 83, no. 8, pages 2733–2737, Apr 1986.

[Doya 2002] K. Doya. *Metalearning and neuromodulation.* Neural Netw, vol. 15, no. 4-6, pages 495–506, 2002.

[Durstewitz & Seamans 2002] D. Durstewitz and J.K. Seamans. *The computational role of dopamine D1 receptors in working memory.* Neural Netw, vol. 15, no. 4-6, pages 561–572, 2002.

[Durstewitz *et al.* 1999] D. Durstewitz, M. Kelc and O. Gntrkn. *A neurocomputational theory of the dopaminergic modulation of working memory functions.* J Neurosci, vol. 19, no. 7, pages 2807–2822, Apr 1999.

[Durstewitz *et al.* 2000a] D. Durstewitz, J.K. Seamans and T.J. Sejnowski. *Dopamine-mediated stabilization of delay-period activity in a network model of prefrontal cortex.* J Neurophysiol, vol. 83, no. 3, pages 1733–1750, Mar 2000.

[Durstewitz *et al.* 2000b] D. Durstewitz, J.K. Seamans and T.J. Sejnowski. *Neurocomputational models of working memory.* Nat Neurosci, vol. 3 Suppl, pages 1184–1191, Nov 2000.

[Egorov *et al.* 2002] A.V. Egorov, B.N. Hamam, E. Fransén, M.E. Hasselmo and A.A. Alonso. *Graded persistent activity in entorhinal cortex neurons.* Nature, vol. 420, no. 6912, pages 173–178, Nov 2002.

[Fabene *et al.* 2008] P.F. Fabene, G. Navarro Mora, M. Martinello, B. Rossi, F. Merigo, L. Ottoboni, S. Bach, S. Angiari, D. Benati, A. Chakir, L. Zanetti, F. Schio, A. Osculati, P. Marzola, E. Nicolato, J.W. Homeister, L. Xia, J.B. Lowe, R. McEver, F. Osculati, A. Sbarbati, E.C. Butcher and G. Constantin. *A role for leukocyte-endothelial adhesion mechanisms in epilepsy.* Nat Med, vol. 14, no. 12, pages 1377–1383, Dec 2008.

[Falconbridge *et al.* 2006] M.S. Falconbridge, R.L. Stamps and D.R. Badcock. *A simple Hebbian/anti-Hebbian network learns the sparse, independent components of natural images.* Neural Computation, vol. 18, no. 2, pages 415–429, 2006.

[Fares & Stepanyants 2009] T. Fares and A. Stepanyants. *Cooperative synapse formation in the neocortex.* Proc Natl Acad Sci U S A, vol. 106, no. 38, pages 16463–16468, Sep 2009.

[Feldman 2009] D.E. Feldman. *Synaptic mechanisms for plasticity in neocortex.* Annu Rev Neurosci, vol. 32, pages 33–55, 2009.

[Florian 2007] R.V. Florian. *Reinforcement learning through modulation of spike-timing-dependent synaptic plasticity.* Neural Comput, vol. 19, no. 6, pages 1468–1502, Jun 2007.

[Foldiák 1990] P. Foldiák. *Forming sparse representations by local anti-Hebbian learning.* Biological Cybernetics, vol. 64, pages 165–170, 1990.

[Fortune & Rose 2001] E.S. Fortune and G.J. Rose. *Short-term synaptic plasticity as a temporal filter.* Trends Neurosci, vol. 24, no. 7, pages 381–385, Jul 2001.

[Fox & Wong 2005] K. Fox and R. Wong. *A comparison of experience-dependent plasticity in the visual and somatosensory systems.* Neuron, vol. 48, no. 3, pages 465–477, Nov 2005.

[Frick *et al.* 2004] A. Frick, J. Magee and D. Johnston. *LTP is accompanied by an enhanced local excitability of pyramidal neuron dendrites.* Nature Neuroscience, vol. 7, no. 2, pages 126–135, Feb 2004.

[Froemke & Dan 2002] R.C. Froemke and Y. Dan. *Spike-timing-dependent synaptic modification induced by natural spike trains.* Nature, vol. 416, no. 6879, pages 433–438, Mar 2002.

[Fröhlich *et al.* 2008] F. Fröhlich, M Bazhenov and T.J. Sejnowski. *Pathological effects of homeostatic synapting scaling on network dynamics in diseases of the cortex.* The Journal of Neuroscience, vol. 28, no. 7, pages 1709–1720, 2008.

[Fu *et al.* 2002] Y.X. Fu, K. Djupsund, H. Gao, B. Hayden, K. Shen and Y. Dan. *Temporal specificity in the cortical plasticity of visual space representation.* Science, vol. 296, no. 5575, pages 1999–2003, Jun 2002.

[Funahashi *et al.* 1989] S. Funahashi, C. J. Bruce and P. S. Goldman-Rakic. *Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex.* J Neurophysiol, vol. 61, no. 2, pages 331–349, Feb 1989.

[Fusi *et al.* 2000] S. Fusi, M. Annunziato, D. Badoni, A. Salamon and D. J. Amit. *Spike-driven synaptic plasticity: theory, simulation, VLSI implementation.* Neural Comput, vol. 12, no. 10, pages 2227–2258, Oct 2000.

[Gabernet *et al.* 2005] L. Gabernet, S.P. Jadhav, D.E. Feldman, M. Carandini and M. Scanziani. *Somatosensory integration controlled by dynamic thalamocortical feed-forward inhibition.* Neuron, vol. 48, no. 2, pages 315–327, Oct 2005.

[Gainutdinov *et al.* 1998] K. L. Gainutdinov, L. J. Chekmarev and T. H. Gainutdinova. *Excitability increase in withdrawal interneurons after conditioning in snail.* Neuroreport, vol. 9, no. 3, pages 517–520, Feb 1998.

[Gainutdinov *et al.* 2000] K. L. Gainutdinov, V. V. Andrianov, T. K. Gainutdinova and E. A. Tarasova. *The electrical characteristics of command and motor neurons during acquisition of a conditioned defensive reflex and formation of long-term sensitization in snails.* Neurosci Behav Physiol, vol. 30, no. 1, pages 81–88, 2000.

[Galic *et al.* 2008] M. Galic, K. Riazi, J. Heida, A. Mouihate, N. Fournier, S. Spencer, L. Kalynchuk, G. Teskey and Q. Pittman. *Postnatal inflammation increases seizure susceptibility in adult rats.* The Journal of Neuroscience, vol. 28, no. 27, pages 6904–6913, 2008.

[Ganguly *et al.* 2000] K. Ganguly, L. Kiss and M. Poo. *Enhancement of presynaptic neuronal excitability by correlated presynaptic and postsynaptic spiking.* Nat Neurosci, vol. 3, no. 10, pages 1018–1026, Oct 2000.

[Gathercole 1999] S. Gathercole. *Cognitive approaches to the development of short-term memory.* Trends in Cognitive Sciences, vol. 3, no. 11, pages 410–419, 1999.

[Gerhard *et al.* 2009] F. Gerhard, C. Savin and J. Triesch. *A robust biologically plausible implementation of ICA-like learning.* In Proc. ESANN, page In press, 2009.

[Gerstner & Kistler 2002a] W. Gerstner and W. Kistler. *Mathematical formulations of Hebbian learning.* Biol Cybern, vol. 87, no. 5-6, pages 404–415, Dec 2002.

[Gerstner & Kistler 2002b] W. Gerstner and W.M. Kistler. Spiking neuron models: Single neurons, populations, plasticity. Cambridge University Press, New York, 2002.

[Gerstner *et al.* 1996] W. Gerstner, R. Kempter, J.L. van Hemmen and H. Wagner. *A neuronal learning rule for sub-millisecond temporal coding.* Nature, vol. 383, pages 76–78, 1996.

[Gewaltig *et al.* 2001] M.O. Gewaltig, M. Diesmann and A. Aertsen. *Propagation of cortical synfire activity: survival probability in single trials and stability in the mean.* Neural Netw, vol. 14, no. 6-7, pages 657–673, 2001.

[Giaume & McCarthy 1996] C. Giaume and K.D. McCarthy. *Control of gap-junctional communication in astrocytic networks.* Trends in Neurosciences, vol. 19, no. 8, pages 319–325, 1996.

[Goldman-Rakic *et al.* 1990] P.S. Goldman-Rakic, S. Funahashi and C.J. Bruce. *Neocortical memory circuits.* Cold Spring Harb Symp Quant Biol, vol. 55, pages 1025–1038, 1990.

[Goldman-Rakic 1990] P. S. Goldman-Rakic. *Cellular and circuit basis of working memory in prefrontal cortex of nonhuman primates.* Prog Brain Res, vol. 85, pages 325–35; discussion 335–6, 1990.

[Goldman-Rakic 1995] P.S. Goldman-Rakic. *Cellular basis of working memory.* Neuron, vol. 14, no. 3, pages 477–485, Mar 1995.

[Goldman *et al.* 2002] M.S. Goldman, P. Maldonado and L.F. Abbott. *Redundancy reduction and sustained firing with stochastic depressing synapses.* J Neurosci, vol. 22, no. 2, pages 584–591, Jan 2002.

[Goldman 2009] M.S. Goldman. *Memory without feedback in a neural network.* Neuron, vol. 61, no. 4, pages 621–634, Feb 2009.

[González-Burgos *et al.* 2004] G. González-Burgos, L.S. Krimer, N. Urban, G. Barrionuevo and D.A. Lewis. *Synaptic efficacy during repetitive activation of excitatory inputs in primate dorsolateral prefrontal cortex.* Cereb Cortex, vol. 14, no. 5, pages 530–542, May 2004.

[Gorelova & Yang 2000] N.A. Gorelova and C.R. Yang. *Dopamine D1/D5 receptor activation modulates a persistent sodium current in rat prefrontal cortical neurons in vitro.* J Neurophysiol, vol. 84, no. 1, pages 75–87, Jul 2000.

[Gutierrez *et al.* 1993] E. G. Gutierrez, W. Banks and J. Kastin. *Murine tumor necrosis factor alpha is transported from blood to brain in the mouse.* Journal of Neuroimmunology, vol. 47, pages 169–176, 1993.

[Haagmans *et al.* 1994] B. Haagmans, A. von den Eertwegh, E. Claassen, M. Hoarzinek and V. Schijns. *Tumour necrosis factor-α production during cytomegalovirus infection in immunosuppressed rats.* J. Gen. Virol., vol. 75, pages 779–787, 1994.

[Haeusler *et al.* 2009] S. Haeusler, K. Schuch and W. Maass. *Motif distribution and computational performance of two data-based cortical microcircuit templates.* J. of Physiology, vol. In Press, 2009.

[Halassa *et al.* 2009] M. Halassa, C. Florian, T. Fellin, J.R. Munoz, S.Y. Lee, T. Abel, P.G. Haydon and M.G. Frank. *Astrocytic modulation of sleep homeostasis and cognitive consequences of sleep loss.* Neuron, vol. 61, no. 2, pages 213–219, Jan 2009.

[Hanisch & Kettenmann 2007] U.K. Hanisch and H. Kettenmann. *Microglia: active sensor and versatile effector cells in the normal and pathologic brain.* Nature Neuroscience, vol. 10, no. 11, pages 1387 – 1394, 2007.

[Hanse 2008] E. Hanse. *Associating synaptic and intrinsic plasticity.* J Physiol, vol. 586, no. 3, pages 691–692, Feb 2008.

[Haydon 2001] P. G. Haydon. *GLIA: listening and talking to the synapse.* Nat Rev Neurosci, vol. 2, no. 3, pages 185–193, Mar 2001.

[Hebb 1949] D. Hebb. The organization of behavior; a neuropsychological theory. New York: Wiley, 1949.

[Henneberger *et al.* 2010] C. Henneberger, T. Papouin, S. Oliet and D.A. Rusakov. *Long-term potentiation depends on release of D-serine from astrocytes.* Nature, vol. 463, pages 232–236, 2010.

[Hess & Gustafsson 1990] G. Hess and B. Gustafsson. *Changes in field excitatory postsynaptic potential shape induced by tetanization in the CA1 region of the guinea-pig hippocampal slice.* Neuroscience, vol. 37, no. 1, pages 61–69, 1990.

[Hihara *et al.* 2006] S. Hihara, T. Notoya, M. Tanaka, S. Ichinose, H. Ojima, S. Obayashi, N. Fujii and A. Iriki. *Extension of corticocortical afferents into the anterior bank of the intraparietal sulcus by tool-use training in adult monkeys.* Neuropsychologia, vol. 44, no. 13, pages 2636–2646, 2006.

[Hofer *et al.* 2006] S.B. Hofer, T.D. Mrsic-Flogel, T. Bonhoeffer and M. Hübener. *Lifelong learning: ocular dominance plasticity in mouse visual cortex.* Curr Opin Neurobiol, vol. 16, no. 4, pages 451–459, Aug 2006.

[Hofer *et al.* 2009] S. Hofer, T.D. Mrsic-Flogel, T. Bonhoeffer and M. Hübener. *Experience leaves a lasting structural trace in cortical circuits.* Nature, vol. 457, no. 7227, pages 313–317, Jan 2009.

[Holtmaat & Svoboda 2009] A. Holtmaat and K. Svoboda. *Experience-dependent structural synaptic plasticity in the mammalian brain.* Nat Rev Neurosci, vol. 10, no. 9, pages 647–658, Sep 2009.

[Hopfield 1982] J. J. Hopfield. *Neural networks and physical systems with emergent collective computational abilities.* Proc Natl Acad Sci U S A, vol. 79, no. 8, pages 2554–2558, Apr 1982.

[Horn & Usher 1989] D. Horn and M. Usher. *Neural networks with dynamical thresholds.* Phys Rev A, vol. 40, no. 2, pages 1036–1044, 1989.

[Hou *et al.* 2008] Q. Hou, D. Zhang, L. Jarzylo, R.L. Huganir and H.Y. Man. *Homeostatic regulation of AMPA receptor expression at single hippocampal synapses.* Proc Natl Acad Sci U S A, vol. 105, no. 2, pages 775–780, Jan 2008.

[Houweling *et al.* 2005] A.R. Houweling, M. Bazhenov, I. Timofeev, M. Steriade and T.J. Sejnowski. *Homeostatic Synaptic Plasticity Can Explain Post-traumatic Epileptogenesis in Chronically Isolated Neocortex.* Cerebral Cortex, vol. 15, pages 834–845, 2005.

[Houzen *et al.* 1997] H. Houzen, S. Kikuchi, M. Kanno, K. Shinpo and K. Tashiro. *Tumor necrosis factor enhancement of transient outward potassium currents in cultured rat cortical neurons.* J Neurosci Res, vol. 50, no. 6, pages 990–999, Dec 1997.

[Hubel & Wiesel 1998] D.H. Hubel and T.N. Wiesel. *Early exploration of the visual cortex.* Neuron, vol. 20, no. 3, pages 401–412, Mar 1998.

[Huber 1985] P. Huber. *Projection pursuit.* The Annals of Statistics, vol. 13, no. 2, pages 435–475, 1985.

[Hyvärinen *et al.* 2001] A. Hyvärinen, J. Karhunen and E. Oja. Independent Component Analysis. Wiley Series on Adaptive and Learning Systems for Signal Processing, Communication and Control, 2001.

[Hyvärinen 1997] A. Hyvärinen. *One-unit contrast functions for independent component analysis: a statistical analysis.* In Neural Networks for Signal Processing, pages 388–397, 1997.

[Hyvärinen 1999] A. Hyvärinen. *Survey on Independent Component Analysis.* Neural Computing Surveys, vol. 2, pages 94–128, 1999.

[Husser & Mel 2003] M. Husser and B. Mel. *Dendrites: bug or feature?* Curr Opin Neurobiol, vol. 13, no. 3, pages 372–383, Jun 2003.

[Ibata *et al.* 2008] K. Ibata, Q. Sun and G. Turrigiano. *Rapid synaptic scaling induced by changes in postsynaptic firing.* Neuron, vol. 57, no. 6, pages 819–826, Mar 2008.

[Intrator & Cooper 1992] N. Intrator and L. Cooper. *Objective function formulation of the BCM theory of visual cortical plasticity: Statistical connections, stability conditions.* Neural Networks, vol. 5, pages 3–17, 1992.

[Intrator 1998] N. Intrator. *Neuronal goals: efficient coding and coincidence detection.* In Springer-Verlag, editeur, ICONIP Hong Kong: Progress in Neural Information Processing, volume 1, pages 29–34, 1998.

[Ireland & Abraham 2002] D.R. Ireland and W.C. Abraham. *Group I mGluRs increase excitability of hippocampal CA1 pyramidal neurons by a PLC-independent mechanism.* J Neurophysiol, vol. 88, no. 1, pages 107–116, Jul 2002.

[Izhikevich & Desai 2003] E. Izhikevich and N Desai. *Relating STDP to BCM.* Neural Computation, vol. 15, pages 1511–1523, 2003.

[Izhikevich 2003] E.M. Izhikevich. *Simple Model of Spiking Neurons.* IEEE Transactions on Neural Networks, vol. 14, pages 1569– 1572, 2003.

[Izhikevich 2007] E. Izhikevich. *Solving the distal reward problem through linkage of STDP and dopamine signaling.* Cerebral Cortex, vol. 17, pages 2443–2452, 2007.

[Jacob *et al.* 2007] V. Jacob, D.J. Brasier, I. Erchova, D. Feldman and D.E. Schulz. *Spike timing-dependent synaptic depression in the in vivo barrel cortex of the rat.* J Neurosci, vol. 27, no. 6, pages 1271–1284, Feb 2007.

[Jester *et al.* 1995] J.M. Jester, L.W. Campbell and T.J. Sejnowski. *Associative EPSP–spike potentiation induced by pairing orthodromic and antidromic stimulation in rat hippocampal slices.* J Physiol, vol. 484 (Pt 3), pages 689–705, May 1995.

[Jin *et al.* 2009] D.Z. Jin, N. Fujii and A.M. Graybiel. *Neural representation of time in cortico-basal ganglia circuits.* Proc Natl Acad Sci U S A, vol. 106, no. 45, pages 19156–19161, Nov 2009.

[Joshi & Triesch 2009] P. Joshi and J. Triesch. *Rules for information-maximization in spiking neurons using intrinsic plasticity.* In Proc. IJCNN, 2009.

[Kalantzis & Shouval 2009] G. Kalantzis and H.Z. Shouval. *Structural plasticity can produce metaplasticity.* PLoS One, vol. 4, no. 11, page e8062, 2009.

[Kanamaru & Sekine 2005] T. Kanamaru and M. Sekine. *Synchronized firings in the networks of class 1 excitable neurons with excitatory and inhibitory connections and their dependences on the forms of interactions.* Neural Computation, vol. 17, no. 6, pages 1315–1338, 2005.

[Kaneko *et al.* 2008] M. Kaneko, D. Stellwagen, R. Malenka and M. Stryker. *Tumor necrosis factor-alpha mediates one component of competitive, experience-dependent plasticity in developing visual cortex.* Neuron, vol. 58, no. 5, pages 673–680, Jun 2008.

[Kanerva 1993] P. Kanerva. *Sparse distributed memory and related models.* In Proceedings of ICPR, 1993.

[Kang & Schuman 1996] H. Kang and E. M. Schuman. *A requirement for local protein synthesis in neurotrophin-induced hippocampal synaptic plasticity.* Science, vol. 273, no. 5280, pages 1402–1406, Sep 1996.

[Kano *et al.* 2008] M. Kano, K. Hashimoto and T. Tabata. *Type-1 metabotropic glutamate receptor in cerebellar Purkinje cells: a key molecule responsible for long-term depression, endocannabinoid signalling and synapse elimination.* Philos Trans R Soc Lond B Biol Sci, vol. 363, no. 1500, pages 2173–2186, Jun 2008.

[Kapur 1990] N.J. Kapur. Maximum entropy models in science and engineering. Wiley-Interscience, 1990.

[Karmarkar & Buonomano 2002] U.R. Karmarkar and D.V. Buonomano. *A model of spike-timing dependent plasticity: one or two coincidence detectors?* J Neurophysiol, vol. 88, no. 1, pages 507–513, Jul 2002.

[Karmarkar & Buonomano 2006] U. Karmarkar and D. Buonomano. *Different forms of homeostatic plasticity are engaged with distinct temporal profiles.* European Journal of Neuroscience, vol. 23, pages 1575–1584, 2006.

[Kennerley & Wallis 2009] S. Kennerley and J. Wallis. *Reward-dependent modulation of working memory in lateral prefrontal cortex..* J Neurosci, vol. 29, no. 10, pages 3259–3270, Mar 2009.

[Keuroghlian & Knudsen 2007] A.S. Keuroghlian and E.I. Knudsen. *Adaptive auditory plasticity in developing and adult animals.* Prog Neurobiol, vol. 82, no. 3, pages 109–121, Jun 2007.

[Kim & Linden 2007] S.J. Kim and D.J. Linden. *Ubiquitous plasticity and memory storage.* Neuron, vol. 56, no. 4, pages 582–592, Nov 2007.

[Kim *et al.* 2009] J.V. Kim, S.S. Kang, M.L. Dustin and D.B. McGavern. *Myelomonocytic cell recruitment causes fatal CNS vascular injury during acute viral meningitis.* Nature, vol. 457, no. 7226, pages 191–195, Jan 2009.

[Klampfl *et al.* 2009] S. Klampfl, R. Legenstein and W. Maass. *Spiking Neurons Can Learn to Solve Information Bottleneck Problems and Extract Independent Components.* Neural Computation, vol. 21, no. 4, pages 911–959, 2009.

[Knutsen & Ahissar 2009] P.M. Knutsen and E. Ahissar. *Orthogonal coding of object location.* Trends Neurosci, vol. 32, no. 2, pages 101–109, Feb 2009.

[Kofuji & Newman 2004] P. Kofuji and E.A. Newman. *Potassium buffering in the central nervous system.* Neuroscience, vol. 129, pages 1045–1056, 2004.

[Kossut & Siuciska 1996] M. Kossut and E. Siuciska. *Overlap of sensory representations in rat barrel cortex after neonatal vibrissectomy.* Acta Neurobiol Exp (Wars), vol. 56, no. 2, pages 499–505, 1996.

[Lamprecht & LeDoux 2004] R. Lamprecht and J. LeDoux. *Structural plasticity and memory.* Nat Rev Neurosci, vol. 5, no. 1, pages 45–54, Jan 2004.

[Lancman & Morris 1996] M. Lancman and H. Morris. *Epilepsy after central nervous system infection: clinical characteristics and outcome after epilepsy surgery.* Epilepsy Research, vol. 25, pages 285–290, 1996.

[Laughlin 1981] S.B. Laughlin. *A simple coding procedure enhances a neuron's information capacity.* Z Naturforsch C, vol. 36, no. 9-10, pages 910–912, 1981.

[Laughlin 2001] S.B. Laughlin. *Energy as a constraint on the coding and processing of sensory information.* Curr Opin Neurobiol, vol. 11, no. 4, pages 475–480, Aug 2001.

[Laurent *et al.* 2001] G. Laurent, M. Stopfer, R. W. Friedrich, M. I. Rabinovich, A. Volkovskii and H. D. Abarbanel. *Odor encoding as an active, dynamical process: experiments, computation, and theory.* Annu Rev Neurosci, vol. 24, pages 263–297, 2001.

[Lazar *et al.* 2007] A. Lazar, G. Pipa and J. Triesch. *Fading memory and time series prediction in recurrent networks with different forms of plasticity.* Neural Netw, vol. 20, no. 3, pages 312–322, Apr 2007.

[Lennie 2003] P. Lennie. *The cost of neural computation.* Current Biology, vol. 13, pages 493–497, 2003.

[Levy & Steward 1983] W.B. Levy and O. Steward. *temporal contiguity requirements for long-term associative potentiation/depression in the hippocampus.* Neuroscience, vol. 8, pages 791–797, 1983.

[Li *et al.* 2004] C. Li, J. Lu, C. Wu, S. Duan and M. Poo. *Bidirectional modification of presynaptic neuronal excitability accompanying spike timing-dependent synaptic plasticity.* Neuron, vol. 41, no. 2, pages 257–268, Jan 2004.

[Lisman 1997] J.E. Lisman. *Bursts as a unit of neural information: making unreliable synapses reliable.* Trends Neurosci, vol. 20, no. 1, pages 38–43, Jan 1997.

[Lucas *et al.* 2006] S.M. Lucas, N. Rothwell and R. Gibson. *The role of inflammation in CNS injury and disease.* British Journal of Pharmacology, vol. 147, pages 232–240, 2006.

[Luciana & Nelson 1998] M. Luciana and C.A. Nelson. *The functional emergence of prefrontally-guided working memory systems in four- to eight-year-old children.* Neuropsychologia, vol. 36, no. 3, pages 273–293, Mar 1998.

[Lücke & Sahani 2008] J. Lücke and M. Sahani. *Maximal causes for non-linear component extraction.* Journal of Machine Learning Research, vol. 9, pages 1227–1267, 2008.

[Lücke 2009] J. Lücke. *Receptive Field Self-Organization in a Model of the Fine-Structure in V1 Cortical Columns.* Neural Computation, page In press, 2009.

[Lynch *et al.* 1977] G.S. Lynch, T. Dunwiddie and V. Gribkoff. *Heterosynaptic depression: a postsynaptic correlate of long-term potentiation.* Nature, vol. 266, no. 5604, pages 737–739, Apr 1977.

[Maass *et al.* 2002] W. Maass, T. Natschläger and H. Markram. *Real-time computing without stable states: a new framework for neural computation based on perturbations.* Neural Comput, vol. 14, no. 11, pages 2531–2560, Nov 2002.

[Macke *et al.* 2009] J. Macke, P. Berens, A. Ecker, A. Tolias and M. Bethge. *Generating spike trains with specified correlation coefficients.* Neural Comp., vol. 21, pages 397–423, 2009.

[Maffei & Fontanini 2009] A. Maffei and A. Fontanini. *Network homeostasis: a matter of coordination.* Curr Opin Neurobiol, vol. 19, no. 2, pages 168–173, Apr 2009.

[Malenka & Bear 2004] R. Malenka and M. Bear. *LTP and LTD: an embarrassment of riches.* Neuron, vol. 44, no. 1, pages 5–21, Sep 2004.

[Marchi *et al.* 2007] N. Marchi, L. Angelov, T. Masaryk, V. Fazio, T. Granata, N. Hernandez, K. Hallene, T. Diglaw, L. Franic, I. Najm and D. Janigro. *Seizure-promoting effect of blood-brain barrier disruption.* Epilepsia, vol. 48, no. 4, pages 732–742, Apr 2007.

[Marder *et al.* 1996] E. Marder, L.F. Abbott, G.G. Turrigiano, Z. Liu and J. Golowasch. *Memory from the dynamics of intrinsic membrane currents.* Proc Natl Acad Sci USA, vol. 93, no. 24, pages 13481–13486, Nov 1996.

[Markovich & Gros 2010] D. Markovich and C. Gros. *Self-organized chaos through heterostatic optimization.* arXiv:1001.0663v1, 2010.

[Markram *et al.* 1995] H. Markram, P.J. Helm and B. Sakmann. *Dendritic calcium transients evoked by single back-propagating action potentials in rat neocortical pyramidal neurons.* J Physiol, vol. 485 (Pt 1), pages 1–20, May 1995.

[Markram *et al.* 1997a] H. Markram, J. Lübke, M. Frotscher and B. Sackmann. *Regulation of synaptic efficacy by coincidence of postsynaptic APS and EPSPS.* Science, vol. 275, no. 5297, pages 213–215, 1997.

[Markram *et al.* 1997b] H. Markram, J. Lbke, M. Frotscher and B. Sakmann. *Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs.* Science, vol. 275, no. 5297, pages 213–215, Jan 1997.

[Martinez-Conde *et al.* 2004] S. Martinez-Conde, S. Macknik and D. Hubel. *The role of fixational eye movements in visual perception.* Nature Reviews Neuroscience, vol. 5, pages 229–240, 2004.

[McAllister & van de Water 2009] K. McAllister and J. van de Water. *Breaking boundaries in neural-immune interactions.* Neuron, vol. 64, no. 1, pages 9–12, Oct 2009.

[McBain & Kauer 2009] C.J. McBain and J.A. Kauer. *Presynaptic plasticity: targeted control of inhibitory networks.* Curr Opin Neurobiol, vol. 19, no. 3, pages 254–262, Jun 2009.

[McFarland & Fuchs 1992] J.L. McFarland and A.F. Fuchs. *Discharge patterns in nucleus prepositus hypoglossi and adjacent medial vestibular nucleus during horizontal eye movement in behaving macaques.* J Neurophysiol, vol. 68, no. 1, pages 319–332, Jul 1992.

[McNaughton *et al.* 1994] B.L. McNaughton, J. Shen, G. Rao, T.C. Foster and C.A. Barnes. *Persistent increase of hippocampal presynaptic axon excitability after repetitive electrical stimulation: dependence on N-methyl-D-aspartate receptor activity, nitric-oxide synthase, and temperature.* Proc Natl Acad Sci U S A, vol. 91, no. 11, pages 4830–4834, May 1994.

[Meli *et al.* 2004] D. Meli, J. Loefflera, P. Baumann, U. Neumann, T. Buhlb, D. Leppert and S. Leib. *In pneumococcal meningitis a novel water-soluble inhibitor of matrix metalloproteinases and TNF-α converting enzyme attenuates seizures and injury of the cerebral cortex.* Journal of Neuroinflammation, vol. 151, pages 6–11, 2004.

[Meliza & Dan 2006] D.C. Meliza and Y. Dan. *Receptive-field modification in rat visual cortex induced by paired visual stimulation and single-cell spiking.* Neuron, vol. 49, no. 2, pages 183–189, Jan 2006.

[Melyan *et al.* 2002] Z. Melyan, H.V. Wheal and B. Lancaster. *Metabotropic-mediated kainate receptor regulation of IsAHP and excitability in pyramidal cells.* Neuron, vol. 34, no. 1, pages 107–114, Mar 2002.

[Miller *et al.* 1993] E.K. Miller, L. Li and R. Desimone. *Activity of neurons in anterior inferior temporal cortex during a short-term memory task.* J Neurosci, vol. 13, no. 4, pages 1460–1478, Apr 1993.

[Miller *et al.* 1996] E.K. Miller, C.A. Erickson and R. Desimone. *Neural mechanisms of visual working memory in prefrontal cortex of the macaque.* J Neurosci, vol. 16, no. 16, pages 5154–5167, Aug 1996.

[Miller *et al.* 2003] P. Miller, C. Brody, R. Romo and X.J. Wang. *A recurrent network model of somatosensory parametric working memory in the prefrontal cortex.* Cereb Cortex, vol. 13, no. 11, pages 1208–1218, Nov 2003.

[Miyashita 1988] Y. Miyashita. *Neuronal correlate of visual associative long-term memory in the primate temporal cortex.* Nature, vol. 335, no. 6193, pages 817–820, Oct 1988.

[Molina-Luna *et al.* 2009] K. Molina-Luna, A. Pekanovic, S. Röhrich, B. Hertler, M. Schubring-Giese, M. Rioult-Pedotti and A.R. Luft. *Dopamine in motor cortex is necessary for skill learning and synaptic plasticity.* PLoS One, vol. 4, no. 9, page e7082, 2009.

[Mongillo *et al.* 2008] G. Mongillo, O. Barak and M. Tsodyks. *Synaptic Theory of Working Memory.* Science, vol. 319, no. 5869, pages 1543–1546, 2008.

[Moreno-Bote & Parga 2005] R. Moreno-Bote and N. Parga. *Simple model neurons with AMPA and NMDA filters: role of synaptic time scales* . Neurocomputing, vol. 65, pages 441–448, 2005.

[Mozzachiodi & Byrne 2009] R. Mozzachiodi and J.H. Byrne. *More than synaptic plasticity: role of nonsynaptic plasticity in learning and memory.* Trends Neurosci, Nov 2009.

[Mu & Poo 2006] Y. Mu and M. Poo. *Spike timing-dependent LTP/LTD mediates visual experience-dependent plasticity in a developing retinotectal system.* Neuron, vol. 50, no. 1, pages 115–125, Apr 2006.

[Mumford 1994] D. Mumford. Large scale neuronal theories of the brain, chapitre Neuronal architectures for pattern-theoretic problems, page 125152. MIT Press, 1994.

[Nadeau & Rivest 1999] S. Nadeau and S. Rivest. *Effects of circulating tumor necrosis factor on the neuronal activity and expression of the genes encoding the tumor necrosis factor receptors (p55 and p75) in the rat brain: a view from the blood-brain barrier.* Neuroscience, vol. 93, no. 4, pages 1449–1464, 1999.

[Navarrete & Araque 2008] M. Navarrete and A. Araque. *Endocannabinoids mediate neuron-astrocyte communication.* Neuron, vol. 57, no. 6, pages 883–893, Mar 2008.

[Newman 2003] E. Newman. *New roles for astrocytes: regulation of synaptic transmission.* TRENDS in Neuroscience, vol. 26, no. 10, pages 536–542, 2003.

[Nita *et al.* 2006] D.A. Nita, Y. Cissé, I. Timofeev and M. Steriade. *Increased propensity to seizures after chronic cortical deafferentation in vivo.* Journal of Neurophysiology, vol. 95, no. 2, pages 902–913, 2006.

[Oby & Janigro 2006] E. Oby and D. Janigro. *The Blood-Brain Barrier and Epilepsy.* Epilepsia, vol. 47, no. 11, pages 1761–1774, 2006.

[Ogoshi *et al.* 2005] F. Ogoshi, H.Z. Yin, Y. Kuppumbatti, B. Song, S. Amindari and J.H. Weiss. *Tumor necrosis-factor-alpha (TNF-α) induces rapid insertion of Ca2+-permeable α-amino-3-hydroxyl-5-methyl-4-isoxazole-propionate (AMPA)/kainate (Ca-A/K) channels in a subset of hippocampal pyramidal neurons.* Experimental Neurology, vol. 193, pages 384–393, 2005.

[Oja 1982] E. Oja. *A simplified neuron model as a principal component analyzer.* J Math Biol, vol. 15, no. 3, pages 267–273, 1982.

[Oja 1997] E. Oja. *The nonlinear PCA learning rule in independent component analysis.* Neurocomputing, vol. 17, no. 1, pages 25–46, 1997.

[Olshausen & Field 1996a] B.A. Olshausen and D.J. Field. *Emergence of simple-cell receptive field properties by learning a sparse code for natural images.* Nature, vol. 381, no. 6583, pages 607–609, Jun 1996.

[Olshausen & Field 1996b] B.A. Olshausen and D.J. Field. *Natural image statistics and efficient coding.* Network, vol. 7, no. 2, pages 333–339, May 1996.

[Olshausen & Field 1997] B.A. Olshausen and D.J. Field. *Sparse coding with an overcomplete basis set: A strategy employed by V1?* Vision Research, vol. 37, no. 23, pages 3311–3325, 1997.

[Olshausen & Field 2004a] B.A. Olshausen and D.J. Field. Problems in systems neuroscience, chapitre What is the other 85% of V1 doing. Oxford University Press, 2004.

[Olshausen & Field 2004b] B.A. Olshausen and D.J. Field. *Sparse coding of sensory inputs.* Curr Opin Neurobiol, vol. 14, no. 4, pages 481–487, Aug 2004.

[Olshausen & Field 2005] B.A. Olshausen and D.J. Field. *How close are we to understanding V1?* Neural Comput, vol. 17, no. 8, pages 1665–1699, Aug 2005.

[Olshausen & Simoncelli 2001] B.A. Olshausen and E. Simoncelli. *Natural image statistics and neural representation.* Annu. Rev. Neuroscience, vol. 24, pages 1193–1216, 2001.

[Pan *et al.* 1997] W. Pan, J.E. Zadina, R.E. Harlan, J.T. Weber, W.A. Banks and A.J. Kastin. *Tumor Necrosis Factor-α: a Neuromodulator in the CNS.* Neuroscience and Biobehavioural Reviews, vol. 21, no. 5, pages 603–613, 1997.

[Parra *et al.* 2009] L.C. Parra, J.M. Beck and A.J. Bell. *On the maximization of information flow between spiking neurons.* Neural Comput, vol. 21, no. 11, pages 2991–3009, Nov 2009.

[Pastalkova *et al.* 2008] E. Pastalkova, V. Itskov, A. Amarasingham and G. Buzsáki. *Internally generated cell assembly sequences in the rat hippocampus.* Science, vol. 321, no. 5894, pages 1322–1327, Sep 2008.

[Pawlak & Kerr 2008] V. Pawlak and J. Kerr. *Dopamine receptor activation is required for corticostriatal spike-timing-dependent plasticity.* J Neurosci, vol. 28, no. 10, pages 2435–2446, Mar 2008.

[Perea & Araque 2005a] G. Perea and A. Araque. *Glial calcium signaling and neuron-glia communication.* Cell Calcium, vol. 38, no. 3-4, pages 375–382, 2005.

[Perea & Araque 2005b] G. Perea and A. Araque. *Properties of synaptically evoked astrocyte calcium signal reveal synaptic information processing by astrocytes.* J Neurosci, vol. 25, no. 9, pages 2192–2203, Mar 2005.

[Perea & Araque 2007] G. Perea and A. Araque. *Astrocytes potentiate transmitter release at single hippocampal synapses.* Science, vol. 317, no. 5841, pages 1083–1086, Aug 2007.

[Perea *et al.* 2009] G. Perea, M. Navarrete and A. Araque. *Tripartite synapses: astrocytes process and control synaptic information.* Trends Neurosci, vol. 32, no. 8, pages 421–431, Aug 2009.

[Perry *et al.* 1995] H.V. Perry, M.D. Bell, H.C. Brown and M.K. Matyszak. *Inflammation in the nervous system.* Current Opinion in Neurobiology, vol. 5, pages 636–641, 1995.

[Pesaran *et al.* 2002] B. Pesaran, J.S. Pezaris, M. Sahani, P. Mitra and R.A. Andersen. *Temporal structure in neuronal activity during working memory in macaque parietal cortex.* Nat Neurosci, vol. 5, no. 8, pages 805–811, Aug 2002.

[Pfister & Gerstner 2006] J.P. Pfister and W. Gerstner. *Triplets of Spikes in a Model of Spike Timing-Dependent Plasticity.* Journal of Neuroscience, vol. 26, no. 38, pages 9673–9682, 2006.

[Prat *et al.* 2005] A. Prat, K. Biernacki and J. P. Antel. *Th1 and Th2 lymphocyte migration across the human BBB is specifically regulated by interferon beta and copolymer-1.* Journal of Autoimmunology, vol. 24, pages 119–124, 2005.

[Quintana & Fuster 1999] J. Quintana and J.M. Fuster. *From perception to action: temporal integrative functions of prefrontal and parietal neurons.* Cereb Cortex, vol. 9, no. 3, pages 213–221, 1999.

[Rabinovich *et al.* 2000] M. Rabinovich, R. Huerta, A. Volkovskii, H. D. Abarbanel, M. Stopfer and G. Laurent. *Dynamical coding of sensory information with competitive networks.* J Physiol Paris, vol. 94, no. 5-6, pages 465–471, 2000.

[Rabinowitch & Segev 2008] I. Rabinowitch and I. Segev. *Two opposing plasticity mechanisms pulling a single synapse.* Trends Neurosci, vol. 31, no. 8, pages 377–383, Aug 2008.

[Rainer & Miller 2000] G. Rainer and E.K. Miller. *Effects of visual experience on the representation of objects in the prefrontal cortex.* Neuron, vol. 27, no. 1, pages 179–189, Jul 2000.

[Rainer & Miller 2002] G. Rainer and E.K. Miller. *Timecourse of object-related neural activity in the primate prefrontal cortex during a short-term memory task.* Eur J Neurosci, vol. 15, no. 7, pages 1244–1254, Apr 2002.

[Rainer *et al.* 1998] G. Rainer, W. F. Asaad and E. K. Miller. *Selective representation of relevant information by neurons in the primate prefrontal cortex.* Nature, vol. 393, no. 6685, pages 577–579, Jun 1998.

[Ransohoff 2009] R. Ransohoff. *Chemokines and Chemokine Receptors: Standing at the Crossroads of Immunobiology and Neurobiology.* Immunity, Oct 2009.

[Reynolds & Wickens 2002] J. Reynolds and J. Wickens. *Dopamine-dependent plasticity of corticostriatal synapses.* Neural Networks, vol. 15, no. 4-6, pages 507–521, 2002.

[Rivest *et al.* 2000] S. Rivest, S. Lacroix, L. Valliéres, S. Nadeau, L. Zhang and N. Laflamme. *How the Blood Talks to the Brain Parenchyma and the Paraventricular Nucleus of the Hypothalamus During Systemic Inflammatory and Infectious Stimuli.* Proc. Soc. Exp. Biol., vol. 223, no. 1, pages 22–38, 2000.

[Rosenberg *et al.* 1995] G. A. Rosenberg, E. Y. Estrada, J. E. Dencoff and W. G. Stetler-Stevenson. *Tumor necrosis factor-alpha-induced gelatinase B causes delayed opening of the blood-brain barrier: an expanded therapeutic window.* Brain Research, vol. 703, pages 151–155, 1995.

[Rostne *et al.* 2007] W. Rostne, P. Kitabgi and S.M. Parsadaniantz. *Chemokines: a new class of neuromodulator?* Nat Rev Neurosci, vol. 8, no. 11, pages 895–903, Nov 2007.

[Rutherford *et al.* 1998] L. C. Rutherford, S. B. Nelson and G. G. Turrigiano. *BDNF has opposite effects on the quantal amplitude of pyramidal neuron and interneuron excitatory synapses.* Neuron, vol. 21, no. 3, pages 521–530, Sep 1998.

[Saar *et al.* 1998] D. Saar, Y. Grossman and E. Barkai. *Reduced after-hyperpolarization in rat piriform cortex pyramidal neurons is associated with increased learning capability during operant conditioning.* Eur J Neurosci, vol. 10, no. 4, pages 1518–1523, Apr 1998.

[Saar *et al.* 2001] D. Saar, Y. Grossman and E. Barkai. *Long-lasting cholinergic modulation underlies rule learning in rats.* J Neurosci, vol. 21, no. 4, pages 1385–1392, Feb 2001.

[Saar *et al.* 2002] D. Saar, Y. Grossman and E. Barkai. *Learning-induced enhancement of postsynaptic potentials in pyramidal neurons.* J Neurophysiol, vol. 87, no. 5, pages 2358–2363, May 2002.

[Sahley 1995] C.L. Sahley. *What we have learned from the study of learning in the leech.* J Neurobiol, vol. 27, no. 3, pages 434–445, Jul 1995.

[Savin & Triesch 2009] C. Savin and J. Triesch. *Developing a working memory with reward-modulated STDP.* In Computational and systems neuroscience, 2009.

[Savin *et al.* 2009] C. Savin, J. Triesch and M. Meyer-Hermann. *Epileptogenesis due to glia-mediated synaptic scaling.* J R Soc Interface, vol. 6, no. 37, pages 655–668, Aug 2009.

[Savin *et al.* 2010] C. Savin, P. Joshi and J. Triesch. *Independent component analysis with spiking neurons.* Plos Computational Biology, vol. In press, 2010.

[Sawaguchi & Goldman-Rakic 1994] T. Sawaguchi and P. S. Goldman-Rakic. *The role of D1-dopamine receptor in working memory: local injections of dopamine antagonists into the prefrontal cortex of rhesus monkeys performing an oculomotor delayed-response task.* J Neurophysiol, vol. 71, no. 2, pages 515–528, Feb 1994.

[Scemes & Giaume 2006] E. Scemes and C. Giaume. *Astrocyte calcium waves: what they are and what they do.* Glia, vol. 54, no. 7, pages 716–725, Nov 2006.

[Scherzer *et al.* 1998] C.R. Scherzer, G.B. Landwehrmeyer, J.A. Kerner, T.J. Counihan, C.M. Kosinski, D.G. Standaert, L.P. Daggett, celebi G. Veli J.B. Penney and A.B. Young. *Expression of N-methyl-D-aspartate receptor subunit mRNAs in the human brain: hippocampus and cortex.* J Comp Neurol, vol. 390, no. 1, pages 75–90, Jan 1998.

[Schmidhuber *et al.* 2007] J. Schmidhuber, D. Wierstra, M. Gagliolo and F. Gomez. *Training recurrent networks by Evolino.* Neural Comput, vol. 19, no. 3, pages 757–779, Mar 2007.

[Schrauwen *et al.* 2008] B. Schrauwen, M. Wardermann, D. Verstraeten and J. Steil. *Improving reservoirs using intrinsic plasticity.* Neurocomputing, vol. 71, pages 1159–1171, 2008.

[Schreurs *et al.* 1997] B G. Schreurs, D. Tomsic, P.A. Gusev and D.L. Alkon. *Dendritic excitability microzones and occluded long-term depression after classical conditioning of the rabbit's nictitating membrane response.* J Neurophysiol, vol. 77, no. 1, pages 86–92, Jan 1997.

[Schreurs *et al.* 1998] B.G. Schreurs, P.A. Gusev, D. Tomsic, D.L. Alkon and T. Shi. *Intracellular correlates of acquisition and long-term memory of classical conditioning in Purkinje cell dendrites in slices of rabbit cerebellar lobule HVI.* J Neurosci, vol. 18, no. 14, pages 5498–5507, Jul 1998.

[Schuett *et al.* 2001] S. Schuett, T. Bonhoeffer and M. Hübener. *Pairing-induced changes of orientation maps in cat visual cortex.* Neuron, vol. 32, no. 2, pages 325–337, Oct 2001.

[Schultz *et al.* 1993] W. Schultz, P. Apicella and T. Ljungberg. *Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task.* J Neurosci, vol. 13, no. 3, pages 900–913, Mar 1993.

[Seamans & Yang 2004] J.K. Seamans and C.R. Yang. *The principal features and mechanisms of dopamine modulation in the prefrontal cortex.* Prog Neurobiol, vol. 74, no. 1, pages 1–58, Sep 2004.

[Seamans *et al.* 2001a] J.K. Seamans, D. Durstewitz, B.R. Christie, C.F. Stevens and T.J. Sejnowski. *Dopamine D1/D5 receptor modulation of excitatory synaptic inputs to layer V prefrontal cortex neurons.* Proc Natl Acad Sci U S A, vol. 98, no. 1, pages 301–306, Jan 2001.

[Seamans *et al.* 2001b] J.K. Seamans, N. Gorelova, D. Durstewitz and C.R. Yang. *Bidirectional dopamine modulation of GABAergic inhibition in prefrontal cortical pyramidal neurons.* J Neurosci, vol. 21, no. 10, pages 3628–3638, May 2001.

[Sebire *et al.* 1993] G. Sebire, D. Emilie, C. Wallon, C. Hery, O. Devergne, JF. Delfraissy, P. Galanaud and M. Tardieu. *In vitro production of IL-6, IL-1 beta, and tumor necrosis factor-alpha by human embryonic microglial and neural cells.* Journal of Immunology, vol. 150, pages 1517–1523, 1993.

[Seol *et al.* 2007] G.H. Seol, J. Ziburkus, S. Huang, L. Song, I.T. Kim, K. Takamiya, R.L. Huganir, H.K. Lee and A. Kirkwood. *Neuromodulators control the polarity of spike-timing-dependent synaptic plasticity.* Neuron, vol. 55, no. 6, pages 919–929, Sep 2007.

[Seung *et al.* 2000] H. S. Seung, D. D. Lee, B. Y. Reis and D. W. Tank. *Stability of the memory of eye position in a recurrent network of conductance-based model neurons.* Neuron, vol. 26, no. 1, pages 259–271, Apr 2000.

[Shan *et al.* 2007] H. Shan, L. Zhang and G.W. Cottrell. *Recursive ICA.* Advances in Neural Information Processing Systems, vol. 19, 2007.

[Sharpee *et al.* 2006] T. Sharpee, H. Sugihara, A. Kurgansky, S. Rebrik, A. Stryker and K.D. Miller. *Adaptive filtering enhances information transmission in visual cortex.* Nature, vol. 439, no. 7079, pages 936–942, Feb 2006.

[Shatz 2009] C.J. Shatz. *MHC class I: an unexpected role in neuronal plasticity.* Neuron, vol. 64, no. 1, pages 40–45, Oct 2009.

[Shen *et al.* 2008] W. Shen, M. Flajolet, P. Greengard and D.J. Surmeier. *Dichotomous dopaminergic control of striatal synaptic plasticity.* Science, vol. 321, no. 5890, pages 848–851, Aug 2008.

[Shohami *et al.* 1999] E. Shohami, I. Ginis and J. M. Hallenbeck. *Dual role of tumor necrosis factor alpha in brain injury.* Cytokine Growth Factor Rev, vol. 10, no. 2, pages 119–130, Jun 1999.

[Simoncelli & Olshausen 2001] E. Simoncelli and B. Olshausen. *Natural Image statistics and neural representations.* Annu. Rev. Neuroscience, vol. 24, pages 1193–1216, 2001.

[Simoncelli 2003] E. Simoncelli. *Vision and the statistics of the visual environment.* Current Opinion in Neurobiology, vol. 13, pages 144–149, 2003.

[Siucinska & Kossut 2004] E. Siucinska and M. Kossut. *Experience-dependent changes in cortical whisker representation in the adult mouse: a 2-deoxyglucose study.* Neuroscience, vol. 127, no. 4, pages 961–971, 2004.

[Sjöstrm *et al.* 2001] P.J. Sjöstrm, G.G. Turrigiano and S.B. Nelson. *Rate, timing, and cooperativity jointly determine cortical synaptic plasticity.* Neuron, vol. 32, no. 6, pages 1149–1164, Dec 2001.

[Song *et al.* 2000] S. Song, K.D. Miller and L.F. Abbott. *Competitive Hebbian learning through spike-timing-dependent synaptic plasticity.* Nature Neurosci, vol. 3, pages 919–926, 2000.

[Sourdet *et al.* 2003] V. Sourdet, M. Russier, G. Daoudal, N. Ankri and D. Debanne. *Long-term enhancement of neuronal excitability and temporal fidelity mediated by metabotropic glutamate receptor subtype 5.* J Neurosci, vol. 23, no. 32, pages 10238–10248, Nov 2003.

[Stanton & Sejnowski 1989] P.K. Stanton and T.J. Sejnowski. *Associative long-term depression in the hippocampus induced by hebbian covariance.* Nature, vol. 339, no. 6221, pages 215–218, May 1989.

[Stellwagen & Malenka 2006] D. Stellwagen and R.C. Malenka. *Synaptic scaling mediated by glial TNF-α.* Nature, vol. 440, pages 1054–1059, 2006.

[Stellwagen *et al.* 2005] D. Stellwagen, E. Beattie, Jae Y. Seo and R.C. Malenka. *Differential regulation of AMPA receptor and GABA receptor trafficking by Tumor Necrosis Factor-α.* Journal of Neuroscience, vol. 25, no. 12, pages 3219–3228, 2005.

[Stemmler & Koch 1999] M. Stemmler and C. Koch. *How voltage-dependent conductances can adapt to maximize the information encoded by neuronal firing rate.* Nature Neuroscience, vol. 2, no. 6, pages 521–527, 1999.

[Timofeev *et al.* 2000] I. Timofeev, F. Grenier, M. Bazhenov, T. Sejnowski and M. Steriade. *Origin of Slow Cortical Oscillations in Deafferented Cortical Slabs.* Cerebral Cortex, vol. 10, pages 1185–1199, 2000.

[Toyoizumi *et al.* 2005] T. Toyoizumi, J.P. Pfister, K. Aihara and W. Gerstner. *Generalized Bienenstock-Cooper-Munro rule for spiking neurons that maximizes information transmission.* PNAS, vol. 102, no. 14, pages 5239–5244, 2005.

[Triesch 2007] J. Triesch. *Synergies between intrinsic and synaptic plasticity mechanisms.* Neural Computation, vol. 19, pages 885–909, 2007.

[Tsodyks & Markram 1997] M. Tsodyks and H. Markram. *The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability.* PNAS, vol. 94, pages 719–723, 1997.

[Tsubokawa *et al.* 2000] H. Tsubokawa, S. Offermanns, M. Simon and M. Kano. *Calcium-dependent persistent facilitation of spike backpropagation in the CA1 pyramidal neurons.* J Neurosci, vol. 20, no. 13, pages 4878–4884, Jul 2000.

[Turrigiano & Nelson 2000] G.G. Turrigiano and S.B. Nelson. *Hebb and homeostasis in neuronal plasticity.* Current Opinion in Neurobiology, vol. 10, no. 3, pages 358–364, 2000.

[Turrigiano & Nelson 2004] G. Turrigiano and S. Nelson. *Homeostatic plasticity in the developing nervous system.* Nature Reviews Neuroscience, vol. 5, pages 97–107, 2004.

[Turrigiano *et al.* 1998] G. Turrigiano, K. Leslie, N. Desai, L. Rutherford and S. Nelson. *Activity-dependent scaling of quantal amplitude in neocortical neurons.* Nature, vol. 391, pages 892–896, 1998.

[Turrigiano 2006] G. Turrigiano. *More than a sidekick: glia and homeostatic synaptic plasticity.* Trends Mol Med, vol. 12, no. 10, pages 458–460, Oct 2006.

[Turrigiano 2007] G. Turrigiano. *Homeostatic signaling:the positive side of negative feedback.* Current Opinion in Neurobiology, vol. 17, pages 318–324, 2007.

[Turrigiano 2008] G. Turrigiano. *The self-tuning neuron: synaptic scaling of excitatory synapses.* Cell, vol. 135, no. 3, pages 422–435, Oct 2008.

[Urbanczik & Senn 2009] R. Urbanczik and W. Senn. *Reinforcement learning in populations of spiking neurons.* Nat Neurosci, vol. 12, no. 3, pages 250–252, Mar 2009.

[van Hateren 1998] J.H. van Hateren. *Independent component filters of natural images compared with simple cells in primary visual cortex.* Proceedings of the Royal Society B: Biological Sciences, vol. 265, no. 1394, pages 359–366, 1998.

[van Rossum *et al.* 2000] M.C. van Rossum, G.Q. Bi and G.G. Turrigiano. *Stable Hebbian learning from spike timing-dependent plasticity.* J Neurosci, vol. 20, no. 23, pages 8812–8821, Dec 2000.

[van Vreeswijk & Sompolinsky 1996] C. van Vreeswijk and H. Sompolinsky. *Chaos in Neuronal Networks with Balanced Excitatory and Inhibitory Activity.* Science, vol. 274, no. 5293, pages 1724 – 1726, 1996.

[Vezzani & Granata 2005] A. Vezzani and T. Granata. *Brain inflammation in epilepsy: experimental and clinical evidence.* Epilepsia, vol. 46, no. 11, pages 1724–1743, 2005.

[Vezzani 2005] A. Vezzani. *Inflammationn and Epilepsy.* Epilepsy Currents, vol. 5, no. 1, pages 1–6, 2005.

[Vitkovic *et al.* 2000] L. Vitkovic, J. Bockaert and C. Jacque. *Inflammatory cytokines: Neuromodulators in normal brain?* Journal of Neurochemistry, vol. 74, pages 457–471, 2000.

[Volterra & Meldolesi 2005] A. Volterra and J. Meldolesi. *Astrocytes, from brain glue to communication elements: the revolution continues.* Nature Reviews Neuroscience, vol. 6, pages 626–640, 2005.

[Wang *et al.* 2003] Z. Wang, N.I. Xu, C.P. Wu, S. Duan and M.M. Poo. *Bidirectional changes in spatial dendritic integration accompanying long-term synaptic modifications.* Neuron, vol. 37, no. 3, pages 463–472, Feb 2003.

[Watanabe & Niki 1985] T. Watanabe and H. Niki. *Hippocampal unit activity and delayed response in the monkey.* Brain Res, vol. 325, no. 1-2, pages 241–254, Jan 1985.

[Watanabe 1996] M. Watanabe. *Reward expectancy in primate prefrontal neurons.* Nature, vol. 382, no. 6592, pages 629–632, Aug 1996.

[Weber & Triesch 2008] C. Weber and J. Triesch. *A Sparse Generative Model of V1 Simple Cells with Intrinsic Plasticity.* Neural Computation, vol. 20, no. 5, pages 1261–1284, 2008.

[Wehr & Zador 2003] M. Wehr and A.M. Zador. *Balanced inhibition underlies tuning and sharpens spike timing in auditory cortex.* Nature, vol. 426, no. 6965, pages 442–446, Nov 2003.

[Weinberger 2007] N.M. Weinberger. *Auditory associative memory and representational plasticity in the primary auditory cortex.* Hear Res, vol. 229, no. 1-2, pages 54–68, Jul 2007.

[Wilent & Contreras 2005] W.B. Wilent and D. Contreras. *Dynamics of excitation and inhibition underlying stimulus selectivity in rat somatosensory cortex.* Nat Neurosci, vol. 8, no. 10, pages 1364–1370, Oct 2005.

[Wolters *et al.* 2003] A. Wolters, F. Sandbrink, A. Schlottmann, E. Kunesch, K. Stefan, L.G. Cohen, R. Benecke and J. Classen. *A temporally asymmetric Hebbian rule governing plasticity in the human motor cortex.* J Neurophysiol, vol. 89, no. 5, pages 2339–2345, May 2003.

[Xu *et al.* 2009] T. Xu, X. Yu, A.J. Perlik, W.F. Tobin, J.A. Zweig, K. Tennant, T. Jones and Y. Zuo. *Rapid formation and selective stabilization of synapses for enduring motor memories.* Nature, vol. 462, no. 7275, pages 915–919, Dec 2009.

[Yang & Seamans 1996] C.R. Yang and J.K. Seamans. *Dopamine D1 receptor actions in layers V-VI rat prefrontal cortex neurons in vitro: modulation of dendritic-somatic signal integration.* J Neurosci, vol. 16, no. 5, pages 1922–1935, Mar 1996.

[Yang *et al.* 2009] G. Yang, F. Pan and W. Gan. *Stably maintained dendritic spines are associated with lifelong memories.* Nature, vol. 462, no. 7275, pages 920–924, Dec 2009.

[Yao & Dan 2001] H. Yao and Y. Dan. *Stimulus timing-dependent plasticity in cortical processing of orientation.* Neuron, vol. 32, no. 2, pages 315–323, Oct 2001.

[Yao *et al.* 2004] H. Yao, Y. Shen and Y. Dan. *Intracortical mechanism of stimulus-timing-dependent plasticity in visual cortical orientation tuning.* Proc Natl Acad Sci U S A, vol. 101, no. 14, pages 5081–5086, Apr 2004.

[Yoshida *et al.* 2003] M. Yoshida, Y. Naya and Y. Miyashita. *Anatomical organization of forward fiber projections from area TE to perirhinal neurons representing visual long-term memory in monkeys.* Proc Natl Acad Sci USA, vol. 100, no. 7, pages 4257–4262, Apr 2003.

[Yuhas *et al.* 2003] Y. Yuhas, A. Weizman and S. Ashkenazi. *Bidirectional Concentration-Dependent Effects of Tumor Necrosis Factor Alpha in Shigella dysenteriae-Related Seizures.* Infection and Immunity, pages 2288–2291, 2003.

[Zhang & Linden 2003] W. Zhang and D.J. Linden. *The other side of the engram: experience-dependent changes in neuronal intrinsic excitability.* Nature Reviews Neuroscience, vol. 4, pages 885–900, 2003.

[Zhang *et al.* 1998] L.I. Zhang, H.W. Tao, C.E. Holt, W.A. Harris and M. Poo. *A critical window for cooperation and competition among developing retinotectal synapses.* Nature, vol. 395, no. 6697, pages 37–44, Sep 1998.

[Zhang *et al.* 2009] J.C. Zhang, P.M. Lau and G.Q. Bi. *Gain in sensitivity and loss in temporal contrast of STDP by dopaminergic modulation at hippocampal synapses.* Proc Natl Acad Sci U S A, vol. 106, no. 31, pages 13028–13033, Aug 2009.

[Zheng & Knudsen 1999] W. Zheng and E.I. Knudsen. *Functional selection of adaptive auditory space map by GABAA-mediated inhibition.* Science, vol. 284, no. 5416, pages 962–965, May 1999.

[Ziv & Ahissar 2009] N. Ziv and E. Ahissar. *New tricks and old spines.* Nature, vol. 462, pages 859–861, 2009.

Glossary

action potential all-or-none electric signal used for transmitting information between neurons; also called spike. 4

afterhyperpolarization membrane hyperpolariation following a spike. 12

AMPAR subtype of ligand-gated glutamate receptor, responsible for the majority of excitatory current in cortical synapses. 16, 58

apical dendrite subtype of dendrite, characteristic for pyramidal cells. 11

Arc activity-regulated cytoskeleton-associated protein; can be used as a marker for plastic changes in the brain. 10

astrocyte are characteristic star-shaped glial cells in the brain, with multiple functions (e.g. regulating ionic concentrations in the extracellular space, synaptic reuptake, modulating neuronal transmission). 58, 59, 65

astrocytic process extension of an astrocyte, which can dynamically sample the neighboring intercellular space. 60

axon the long extension from a neuron that transmits outgoing signals to other cells. 15

BAP backpropagating action potential; an action potential initiated at the soma propagates back into the dendrites, influencing the integration of synaptic signals and synaptic plasticity. 6

blood-brain barrier structure separating the blood and cerebrospinal fluid, which prevents microscopic objects, such as bacteria, and large molecules to enter the brain. 57, 63

CA1 subregion of the hippocampus. 11, 17

cannabinoid receptor class of membrane receptors activated by cannabinoids. 6

chemokine specific protein which can induce directed chemotaxis, i.e. movement along a chemical gradient, in nearby responsive cells. 63

chunking refers to a strategy for making more efficient use of short-term memory by recoding information in larger units. 42

classic conditioning a form of associative learning which involves presentations of a neutral stimulus (conditioned stimulus) along with a stimulus which evokes an innate, often reflexive, response (unconditioned stimulus). If the two are repeatedly paired, eventually the organism begins to produce the behavioral response for the neutral stimulus alone. 8

complex cell V1 cell responding to appropriately oriented stimuli, independent of its location in the receptive field. Many complex cells are also direction-selective. 25

critical period developmental time window within which experience can induce changes in a cortical region. 18

cytokine category of signaling molecules that are secreted by specific cells of the immune system, involved in inter-cell communication. 63, 65

delay period in working memory experiments, the time between stimulus presentation and recall, when the stimulus is no longer visible and needs to be actively maintained in memory. 44

dendrite a branch-like extension through which a neuron receives messages from other neurons. 6

dopamine neuromodulator produced by dopaminergic neurons in several areas of the brain, including the substantia nigra and the ventral tegmental area. Known to affect channel dynamics and modulate plasticity at cortico-striatal synapses. 16, 64

E-S potentiation EPSP-spike potentiation; an activity-dependent increase in synaptic throughput, or the probability of an input to elict a postsynaptic spike.. 10

excitatory synapse synapse using glutamate as neurotransmitter, which, when activated, induces a depolarization of the postsynaptic membrane. 2, 8, 15

excitotoxicity excessive activation of glutamate receptors in cortical neurons, which can lead to cell death. 64

exocytosis the process by which a cell sends the contents of vesicles out of the cell membrane. The vesicles bind to the membrane, releasing soluble proteins into the extracellular environment. 58

extinction the termination of the reinforcement contingency that maintains the response. The overall outcome is a reduction or elimination of the conditioned response. 14

eyetrace conditioning a form of classical conditioning, which involves the pairing of an auditory or visual stimulus with an eye blink-eliciting unconditioned stimulus (e.g. a mild puff of air to the eye). 14

filopodium long, thin protrusion extending from an axon, at the end of which a presynaptic terminal can form. 15

glia cell support cell in the brain. 3, 57

gliotransmitter chemical released by glia cells, usually astrocytes, that enable communication with neurons or other glial cells. 60

glutamate most abundant excitatory neurotransmitter in the vertebrate nervous system.. 66

habituation the cessation of a response after repeated presentation of the stimulus. 13

heterosynaptic plasticity activity at one synapses induces changes in the property of other synapses. 11

hippocampus important brain region which belonging to the limbic system of mammals, known to play important roles in memory and spatial navigation. In humans and other primates, it is located inside the medial temporal lobe. 4, 17, 18

inhibitory synapse synapse using different types of GABA as neurotransmitter, which, when activated, induces a hyperpolarization of the postsynaptic membrane. 2

kaininc acid agonist for the kainate receptor which mimics the effect of glutamate. Used in animal experiments for inducing seizures.. 65

knockout genetic technique in which an organism is engineered to carry genes that have been made inoperative. 59

leukocyte are cells of the immune system defending the body against both infectious disease and foreign materials. Also termed white blood cells. 63, 65

limit cycle closed trajectory in phase space. 53

lymphocyte is a type of white blood cell in the vertebrate immune system. 69

mEPSC see quantal amplitude. 9, 58

metaplasticity refers to higher-order regulation of synaptic plasticity, otherwise put how does the activity history affect synaptic plasticity itself. 17

microglia specific class of glia cells, involved in immune responses in the brain. 59, 65, 69

mIPSC see quantal amplitude. 58

monocyte a part of leukocytes, used for replenishing resident macrophages and dendritic cells. They mature into different type of macrophages, and are responsible for the phagocytosis (ingestion) of foreign substances in the body. 69

nearest-neighbor STDP model of STDP in which the synaptic changes induced by a spike depend only on the temporally closest spikes at the other synaptic terminal. 20, 26

neuromodulator a chemical agent that is released by a neurosecretory cell and acts on neurons by modulating their response to neurotransmitters. 16, 64

neuron culture experimental preparation in which neurons are grown in a dish, forming a network with a simpler structure than that in situ. 7

neurotransmitter endogenous chemical which can relay, amplify or modulate signals between neurons. Packaged into synaptic vesicles, they are released into the synaptic cleft after a presynaptic spike, where they bind to postsynaptic receptors. 4, 10, 14, 59

NMDAR subtype of ligand and voltage gated glutamate receptors which are calcium permeable. 5, 16

ocular dominance in the mature primary visual cortex of mammals neurons respond predominantly to inputs from one of the eyes. When looked at the cortical surface, such neurons are organized in alternating stripes corresponding to the left and the right eye, termed ocular dominance columns. 1, 15

oligodendrocyte a special type of glia cell, whose main function is to produce myelin, which is wrapped around the axons to increase the speed of neuronal transmission. 59

operant conditioning a form of conditioning in which the subject learns the consequences of this actions, thereby altering his behavior. 14

optic tectum symmetric structure which is part of the vertebrate midbrain. 17

orientation tuning selective response of neurons in primary visual cortex to images containing edges of a specific orientation. 1

p55 TNF-α receptor, expressed predominantly in neurons, part of the pathway triggering neurodegeneration and cell death. 65

p75 TNF-α receptor implicated in some of the neuroprotective effects of TNF-α after seizures. 65

perceptual learning involves long lasting changes in human perception following repeated exposed to specific stimuli, e.g. an increased ability to discriminate between very similar stimuli after training. 1

perforant pathway hippocampal pathway, including entorhinal cortex projections to the subiculum. 4

phagocyte white blood cell that protects the body by ingesting (by phagocytosis) harmful foreign particles, bacteria, and dead or dying cells. 57

quantal amplitude the amplitude of the postsynaptic response to a single vesicle; also termed mEPSC or mIPSC, for excitatory and inhibitory synapses, respectively. 9

receptor specialized membrane proteins which mediate the communication between the the cell and the outside world. Specific signaling molecules (e.g. neurotransmitters) attach to the receptor initiating changes on the intracellular side of the membrane. 3, 4

refractory period time interval following a spike within which a neuron cannot fire another action potential. 27, 80

reinforcement learning involves learning from interaction with an environment, from the consequences of action, rather than from explicit teaching. RL theory formalizes how an agent ought to take actions in an environment in order to maximize reward. 50

S1 primary sensory cortex. 1, 15

Schaffer collaterals axons connecting pyramidal neurons from region CA3 to apical dendrites of CA1 neurons, in the hippocampus. 4

sensitization specific augmentation of a stimulus response after the presentation of a injurious stimulus. 13

short-term memory conscious, brief retention of information that is currently being processed in a person's mind. 42

short-term synaptic depression form of short-term synaptic plasticity in which a first spike causes a decrease in the postsynaptic potential induced by a subsequent spike. 14

short-term synaptic facilitation form of short-term synaptic plasticity in which the postsynaptic potential induced by a spike increases if another spike was recently transmitted by that synapse. 14

short-term synaptic plasticity dependence of synaptic transmission on the history of activity at the synapse, on a timescale of hundreds of milliseconds. 10, 14

simple cell V1 cell which respond best to elongated stimuli at a specific location. 25

soma neuron body, which includes the cell nucleus; center of protein synthesis. 4

somatotopic map organization of neurons' responses based on location of tactile sensors, typically investigated for rodent whiskers. 1

spike threshold the minimal voltage required for the initiation of an action potential. 16

spine small membranous protrusion from a neuron's dendrite that typically receives input from a single presynaptic terminal. 15

structural plasticity activity dependent changes of neuronal connectivity, through which new synapses are formed and old synapses disappear due to spine retraction. 15, 43

synaptic cleft tiny gap between the pre- and post-synaptic terminal where neurotransmitters are released during synaptic transmission. 14

synaptic scaling scaling up or down of the quantal amplitude of all synapses to a neuron, in response to long lasting changes in neuronal activity. 2, 18, 21, 42, 90

synfire chain feedforward network consisting of multiple neuron pools which enables a volley of spikes to propagate synchronously from pool to pool. 45

T-lymphocyte subtype of lymphocytes which plays a central role in cell-mediated immunity. Also called T-cells. 57

tetrodotoxin potent neurotoxin which blocks action potentials by binding to the pores of the voltage-gated, fast sodium channels, which are needed for spike generation. 58, 59

toll-like receptors a class of proteins which play a key role in the innate immune system. 63

tonotopic map spatial organization of neurons responses in primary auditory cortex as function of sound frequency. 1

V1 primary visual cortex. 1, 7, 15, 18, 21–23, 25, 26, 28, 35, 41

vasopressin is a hormone found in most mammals, including humans. Secreted by the posterior pituitary gland it is released in the bloodstream, but also directly in the brain by neurons of the Supraoptic nucleus. It has been implicated in a variety of functions, including aggression, blood pressure and temperature regulation and possibly memory. 64

voltage-gated ionic channel membrane channel which enables the flow of certain ions in a voltage dependent way. 4

Acronyms

TNF-α tumour necrosis factor-α. 10

AMPA α-amino-3-hydroxyl-5-methyl-4-isoxazole-propionate. 5, 57

ATP adenosine triphosphate. 60, 66

BCM Bienenstock-Cooper-Munro rule. 14, 18, 86

BDNF brain-derived neurotrophic factor. 10

BDP burst-dependent plasticity. 7

IC independent component. 86

ICA independent component analysis. 3, 21, 80

IP intrinsic plasticity. 2, 10, 21, 42, 82, 89

LGN lateral geniculate nucleus. 23

LPS lipopolysaccharide. 64, 65

LTD long-term synaptic depression. 2, 4, 15, 64

LTP long-term synaptic potentiation. 2, 4, 15, 60

MHCI major histocompatibility complex I. 64

MHCII major histocompatibility complex II. 63

NMDA N-methyl-D-aspartic acid. 5, 58

NO nitric oxide. 64

PKA protein kinase A. 12

PKC protein kinase C. 12

R-STDP reward-modulated STDP. 48, 49

RL reinforcement learning. 3

STDP spike timing dependent plasticity. 1, 3, 7, 8, 16, 20, 21, 29, 42, 47, 82, 86

TLE temporal lobe epilepsy. 65, 75

TMS Transcranial magnetic stimulation. 7

VDCC voltage-dependent calcium channel. 6

WM working memory. 42

Acknowledgments

I am very thankful to my family and friends for their constant support. I would like to thank Prof. Dr. Jochen Triesch, Prof. Dr. Michael Meyer-Hermann and Prof. Dr. Christoph von der Malsburg for their supervision, to Jörg Lücke for interesting collaborations and for providing the code for the Gabor l.m.s. fit, Felipe Gerhard for his help with the natural images processing and being a great student, to Claudia Clopath for many comments on the ICA project, to Dr. Kai Willadsen for giving always good advice, to Dr. Cornelius Webber, Thomas Weisswange, Daniel Krieg, Marc Henniges and Christian Keck for proofreading parts of this manuscript. Last but not least, I thank my officemates Hasnaa, Jenia, Maneesh and Adilah for many good memories.