COMENIUS UNIVERSITY IN BRATISLAVA
FACULTY OF MATHEMATICS, PHYSICS AND INFORMATICS

BIOLOGICALLY INSPIRED MODELING OF ARTIFICIAL NEURO-GLIAL NETWORKS

Dissertation thesis

2019   Mgr. Peter Gergeľ

C#### COMENIUS UNIVERSITY IN BRATISLAVA
FACULTY OF MATHEMATICS, PHYSICS AND INFORMATICS

BIOLOGICALLY INSPIRED MODELING OF ARTIFICIAL NEURO-GLIAL NETWORKS

Dissertation thesis

Study programme:	Computer Science
Field of Study:	2508 Informatics
Department:	Department of Applied Informatics
Tutor:	prof. Ing. Igor Farkaš, Dr.

Bratislava, 2019   Mgr. Peter Gergeľ

Comenius University in Bratislava
Faculty of Mathematics, Physics and Informatics

THESIS ASSIGNMENT

Name and Surname: Mgr. Peter Gergeľ
Study programme: Computer Science (Single degree study, Ph.D. III. deg., full time form)
Field of Study: Computer Science, Informatics
Type of Thesis: Dissertation thesis
Language of Thesis: English
Secondary language: Slovak

Title: Biologically inspired modeling of artificial neuro-glial networks

Annotation: Until recently, it was commonly agreed that neurons are only functional units of the brain involved in information storing and processing and that more numerous astroglial cells (ACs) only served passive, supportive role. Research over the last decades has revealed that ACs are also involved in information processing tasks in the brain. Currently there exists evidence illustrating that neurons and AC bidirectionally communicate, ACs control synaptogenesis, provide maintenance of synapses, modulate neural excitation and so on. Models of such neuro-glial networks are still missing.

Aim: 1. Review existing models of neuro-glial networks in classification tasks, pattern recognition, abstraction, or similar problems.
2. Design novel model of neuro-glial network with learning algorithms and evaluate its performance.

Literature: Ikuta, C., Uwate, Y., and Nishio, Y. (2011). Performance and features of multi-layer perceptron with impulse glial network. International Joint Conference on Neural Networks, pages 2536–2541.
Verkhratsky, A. and Butt, A. M. (2007). Glial Neurobiology: A Textbook. John Wiley & Sons.
Nedergaard, M., Ransom, B., and Goldman, S. A. (2003). New roles for astrocytes: redefining the functional architecture of the brain. Trends in Neurosciences, 26(10):523–530.

Tutor: prof. Ing. Igor Farkaš, Dr.
Department: FMFI.KAI - Department of Applied Informatics
Head of department: prof. Ing. Igor Farkaš, Dr.

Assigned: 10.02.2015

Approved: 18.02.2015   prof. RNDr. Rastislav Kráľovič, PhD.
Guarantor of Study Programme

....
Student   Tutor

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Mgr. Peter Gergeľ
Študijný program: informatika (Jednoodborové štúdium, doktorandské III. st., denná forma)
Študijný odbor: informatika
Typ záverečnej práce: dizertačná
Jazyk záverečnej práce: anglický
Sekundárny jazyk: slovenský

Názov: Biologically inspired modeling of artificial neuro-glial networks
Biologicky inšpirované modelovanie umelých neuro-gliových sietí

Anotácia: Až donedávna sa predpokladalo, že neuróny sú jediné funkčné jednotky v mozgu, ktoré sa podieľajú pri uchovávaní a šírení informácie, a že omnoho početnejšie astrogliové bunky (AB) v mozgu zohrávajú iba pasívnu, podpornú úlohu. Výskumy z posledných desiatok rokov ukazujú, že aj AB sa podieľajú na informačných procesoch v mozgu. Existuje viacero dôkazov, ktoré poukazujú na to, že medzi neurónmi a AB existuje obojsmerná komunikácia a že AB regulujú synaptogenézu, starajú sa o údržbu synapsií, modulujú neurálnu excitáciu apod. Modely takýchto neuro-gliových sietí stále chýbajú.

Cieľ: 1. Preskúmajte existujúce modely neuro-gliových sietí pri klasifikácií, rozpoznávaní vzorov, abstrakcii, prip. iných podobných problémov.
2. Navrhnite nový model neuro-gliovej siete spolu s algoritmami pre ich učenie a vyhodnoťte jej správanie.

Literatúra: Ikuta, C., Uwate, Y., and Nishio, Y. (2011). Performance and features of multi-layer perceptron with impulse glial network. International Joint Conference on Neural Networks, pages 2536–2541.
Verkhratsky, A. and Butt, A. M. (2007). Glial Neurobiology: A Textbook. John Wiley & Sons.
Nedergaard, M., Ransom, B., and Goldman, S. A. (2003). New roles for astrocytes: redefining the functional architecture of the brain. Trends in Neurosciences, 26(10):523–530.

Školiteľ: prof. Ing. Igor Farkaš, Dr.
Katedra: FMFI.KAI - Katedra aplikovanej informatiky
Vedúci katedry: prof. Ing. Igor Farkaš, Dr.

Dátum zadania: 10.02.2015

Dátum schválenia: 18.02.2015

prof. RNDr. Rastislav Kráľovič, PhD.
garant študijného programu

....
študent   školiteľ

Acknowledgements

Abstract

Research in neuroscience over the past few decades has shed new light on glial cells which were originally considered as purely passive supportive cells. New data provide convincing evidence that astrocytes, a group of glial cells, possess important physiological functions distinguishing them from passive cells. Currently, it is known that astrocytes are involved in the neuronal activity regulation and the synaptic transmission. Similar to neurons, astrocytes form large networks (glial syncytium) allowing them to communicate with one another over the long distances using Ca^{2+} signals, however on a longer temporal scale compared to neurons. Since this is a relatively new area of research in neuroscience, little attention has been paid to the computational modeling using connectionist approaches. In this thesis, we investigate the potential role of artificial astrocytes in both feedforward and recurrent neural networks. The role of the astrocytes is to listen to and to modulate neurons based on their past activity. By systematic analysis we evaluate the performance of all proposed models on various tasks including data classification and memory capacity. In almost all cases we found hyperparameters allowing the models with astrocyte units to outperform traditional neural networks. Our results suggest that the field is worthy of investigation and deserves further research.

KEYWORDS: glial cells, astrocytes, computational models, artificial neural networks

Abstrakt

Neurovedný výskum za posledné dekády vniesol nové svetlo v nazeraní na gliové bunky, ktoré boli odjakživa považované výlučne za pasívne, podporné bunky. Nové štúdie prinášajú presvedčivé dôkazy, že astrocyty, skupina gliových buniek, disponujú dôležitými fyziologickými funkciami, ktoré ich odlišujú od pasívnych buniek. Aktuálne údaje naznačujú, že astrocyty sú aktívne zapojené v neurálnej regulácii a synaptickej transmisii. Podobne ako neuróny, astrocyty formujú siete (gliové syncytium) pomocou ktorých navzájom komunikujú na veľké vzdialenosti použitím Ca^{2+} signálov, avšak na dlhšej časovej škále v porovnaní s neurónmi. Pretože ide o relatívne novú oblasť výskumu v neurovede, málo pozornosti sa doteraz venovalo výpočtovému modelovaniu použitím konekcionistických prístupov. V tejto dizertačnej práci skúmame potenciálnu úlohu umelých astrocytov ako v dopredných, tak v rekurentných neurónových sieťach. Úlohou astrocytov je počúvať a modulovať neuróny na základe ich predchádzajúcej aktivity. Systematickou analýzou vyhodnocujeme úspešnosť všetkých navrhovaných modelov na rôznych úlohách ako sú klasifikácia dát, či pamäťová kapacita. Takmer vo všetkých prípadoch sme našli hyperparametre, ktoré umožňujú modelom s astrocytmi prekonať tradičné neurónové siete. Naše výsledky naznačujú, že oblasť je hodná pozornosti a zaslúži si ďalší výskum.

KĽÚČOVÉ SLOVÁ: gliové bunky, astrocyty, výpočtové modely, umelé neurónové siete

Table of contents

List of abbreviations

BP	backpropagation
CNS	central nervous system
ESN	echo state network
FFNN	feedforward neural network
MC	memory capacity
MLP	multi-layer perceptron
MSE	mean squared error
PNS	peripheral nervous system
RNN	recurrent neural network
SLP	single-layer perceptron
SOM	self-organizing map
SRN	simple recurrent network

Introduction

According to current knowledge of neuroscience, brain tissue consists of two cell populations: neurons and glia. The population of neurons is characterized by the ability to generate action potentials, whereas glia have been regarded as non-functional and supporting cells for several decades. Neurophysiological findings in the 1990s began to shift this view dramatically in a new direction, conveying evidence that glial cells are actively involved in modulation of neuronal excitability and synaptic plasticity, making them no longer merely passive cells.

Firstly identified in the 19th century, glial cells significantly contribute to the total brain mass with around 50% and the glia:neuron ratio in mammalian brains about 1:1 (Azevedo et al., 2009). The population of glia is commonly subdivided into four major groups: **oligodendrocytes**, **microglia**, **ependymal cells** and **astrocytes**. According to recent evidence, the first three types are closely specialized and account for myelination, immunity and cerebrospinal fluid production, respectively. Astrocytes, the most abundant and probably the most complex group, play a significant role in cognitive functions, traditionally attributed solely to neurons, such as learning and memory, information transfer and processing. Although not being able to be excited electrically and to generate action potentials as neurons do, they are incorporated in network **glial syncytium** where upon being excited *chemically* they propagate Ca^{2+} signals through the gap junctions.

In order to better understand these low-level mechanisms, computational modelling is often employed which recently has become an essential part of neuroscience. Such models may provide testable predictions for processes that are built upon these mechanisms such as neuronal regulation, or synaptic plasticity. Better knowledge about astrocyte–neuron cooperation may also provide building blocks for studying the regulatory capability of glial syncytium on a larger scale. Computational models of ANNs extended with astrocytes may not only be used as an interesting novel concept, but can mainly provide space for hypotheses to explain the potential roles of glia in biological neuronal circuits and networks.

The primary goal of our dissertation thesis is to study neuron–astrocyte coupling in connectionist systems. Inspired by contemporary evidence from neuroscientific research of astrocyte physiology and their interactions with surrounding neurons, and work by Ikuta et al. (2010), we propose several models of feedforward and recurrent artificial neural networks and evaluate their performance in classification tasks and memory capacity.

The dissertation thesis is organized as follows. In Chapter 1 we present a brief introduction to the history of glial cell discovery and recognition. In Chapter 2 we provide a morphological description of various types of glia. In Chapter 3 we summarize related work related to connectionist systems and artificial astrocytes. In Chapter 4 and Chapter 5 we propose feedforward and recurrent neural network extended by astrocyte units, respectively and analyze their performance on various tasks. We conclude our thesis and present the future work in Conclusion.

Chapter 1

Introduction to Glia

1.1 Origin of the glial cell research

Glial cells, also known as neuroglia, or simply glia[1], were first named by Rudolf Virchow (Virchow, 1856) and in later work (Virchow, 1858) defined and described as a substance filling the space between neuronal cells that holds them together and gives the brain structure its own shape. Virchow was not the first to study glia because only a few years ago Müller (1851) had recognized and described the so-called **Müller glia** in the retina where they play a supportive role for local neurons.

Deiters (1865) later showed that glia are separate cells that are not part of the neurons by claiming that cells that do not have axons cannot be neurons. Next significant advancement was made by Golgi (1871) who identified a separate class of star-shaped glia; however the name **astrocytes** was given to them by Lenhossék (1895). In the meantime, Golgi (1885) also developed a morphological description of glial cells and showed that there was a great diversity and differentiation of glia. He identified glia that make contact with blood vessels and glia that fill the space between the neuronal fibers. Later, Golgi together with Ramón y Cajal developed imaging techniques to create detailed images of neurons and glia. Their illustration of astrocytes can be seen in Fig. 1.1.

Since the role of glial cells was not yet known, there were several speculations and hypotheses about their function in the nervous system. Golgi (1885) assumed that glia played a single role of nourishing the neurons by forming intermediate links between vessels and neurons. Ramón y Cajal (1897) disagreed with this theory and, according to him, the role of glia was merely to provide electrical insulation, which was originally a hypothesis

---

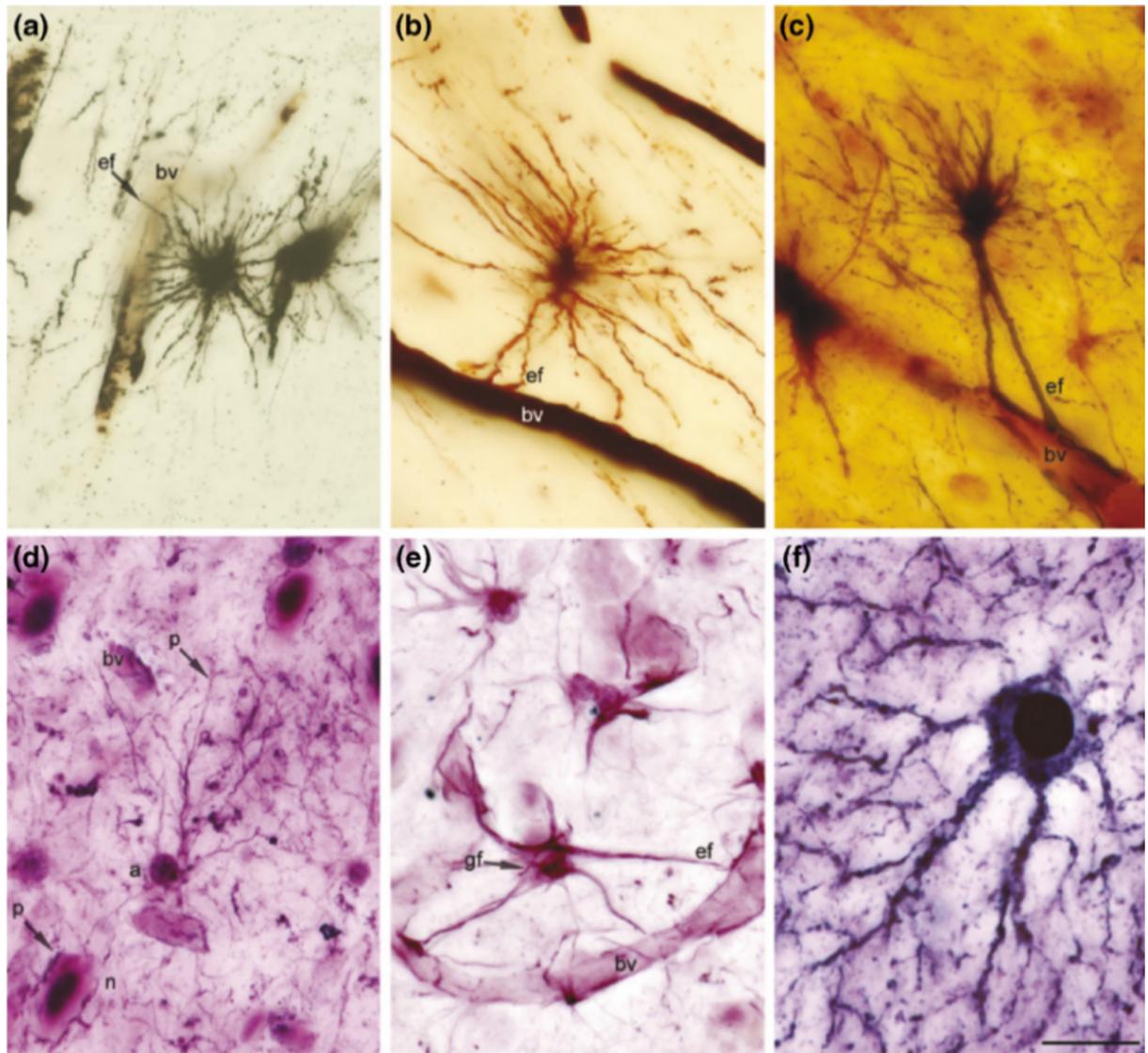[1]the name comes from the Greek language, in which *glia = glue*

Figure 1.1: Original Cajal histological images of protoplasmic astrocytes using **a)** Golgi-Cox method, **b)** Golgi-Kenyon method, **c)** formol-uranium nitrate method, **d)** gold chloride sublimated method, **e)** ammoniacal silver oxide method and **f)** silver carbonated method. Taken from García-Marín et al. (2007).

developed by his brother (Ramón y Cajal, 1891). The German physician Weigert (1895) had a different view of the matter and stated that neuroglia play a structural role in filling and occupying the space between the neurons. Cajal's student, Pio del Río Hortega, later identified and described other important glia types: **microglia** that play a immunity role in the central nervous system (Del Río Hortega, 1920) and **oligodendrocytes** that provide insulation of neuronal axons (Del Río Hortega, 1921).

At the beginning of the 20th century, scientific interest in glia became silent and re-oriented to neurons due the acceptance of the neuron doctrine (Waldeyer, 1891), which reports that only neurons are the basic anatomical and physiological units of nervous systems worth studying. The idea that there was an active interaction between neurons and glia came from Schleich (1894), who insisted that glia and neurons are equally important cells in terms of the information transfer and processing in the brain and demanded their integration into the neuron doctrine. However, this endeavour ended unsuccessfully.

Consequently, glia research stagnated for more than half a century and only in the 1960s the second wave of research started. The first evidence that glia respond to neuronal activity came from Hydén and Egyhazi (1963), where the authors discovered biochemical changes in glia that occur during learning. The work had several shortcomings and faced extensive criticism, but it was the first to propose the existence of the concept of the **tripartite synapse** (more in Section 1.5), which later served as an inspiration for others. Next evidence came from Orkand et al. (1966), who discovered a negative resting membrane potential in glia, permeability to K^+ ions, the existence of **gap junctions** between glia, and the fact that neurons can depolarize nearby glia membranes (shown in Fig. 1.2). In addition, the authors fiddled with an idea that K^+ accumulation is a form of cell signaling, which was later indeed confirmed. Another significant advance was made by Brightman and Reese (1969), who observed and described that glia form networks and are integrated in one large structure called **glial syncytium**. It should be noted, however, that the general view of glia was not changed at that time and they were still perceived exclusively as passive supportive and nutritional cells.

During the following two decades neuroglial research declined and was revived at the turn of the 1980s as a result of appearance of modern physiological techniques such as the *patch-clamp* method and the use of fluorescent dyes. Before that, glia were still considered passive, since they were assumed that they lack receptors for neurotransmitters and that depolarization occurs because of the increased extracellular level of K^+. Kettenmann et al. (1984) found, using these techniques, that this notion is not accurate and showed that
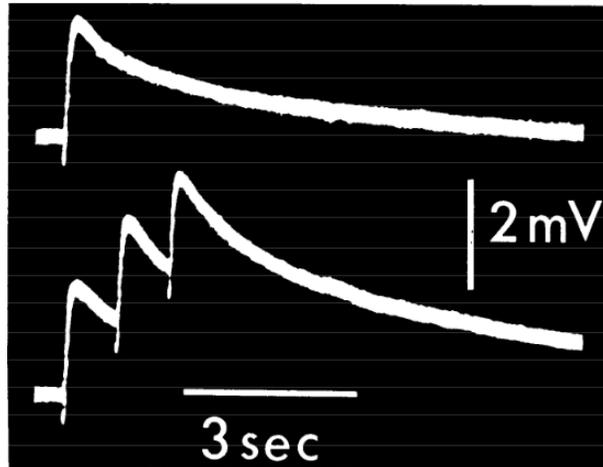
Figure 1.2: Glial cell membrane depolarization by neuronal impulses. The upper signal represents a change in a membrane potential over a single neuronal stimulus, whereas the lower signal represents a change by three neuronal stimuli. Taken from Orkand et al. (1966).

cultured astrocytes can undoubtedly respond to the presence of glutamate, GABA and aspartate (more in Fig. 1.3). The idea that astrocytes are not merely supportive cells, but "listen" to and respond to synapses by inducing biochemical changes was slowly coming to the surface. Last but not least, next important advance was made by Cornell-Bell et al. (1990), who demonstrated that activation of glutamate receptors on the astrocyte membrane induces an increase in the level of free intracellular Ca^{2+} that is propagated to adjacent astrocytes (more in Section 1.3).

1.2 Biophysical properties

Mature macroglia (astrocytes and oligodendrocytes) have a resting membrane potential of -80 to -90mV due to the high K^+ permeability that holds membrane potential near potassium equilibrium. Electrical depolarization of cell membranes induces electrotonic changes in membrane potential and ion distribution (at the membrane level) is similar to other cells: $K^+ \sim 120$mM[2], $Na^+ < 10$mM, $Ca^{2+} < 0.1\mu$M. The difference is in Cl^- ions whose concentration in astrocytes and oligodendrocytes is significantly higher (~ 30–40mM).

Thus, it is clear that glia express all basic ion channels including K^+ (most common), Na^+, Ca^{2+}, Cl^- and as well, the same receptors (ionotropic and metabotropic) as neurons do. Astrocytes mainly express glutamate receptors (GluRs), purine receptors, GABA re-
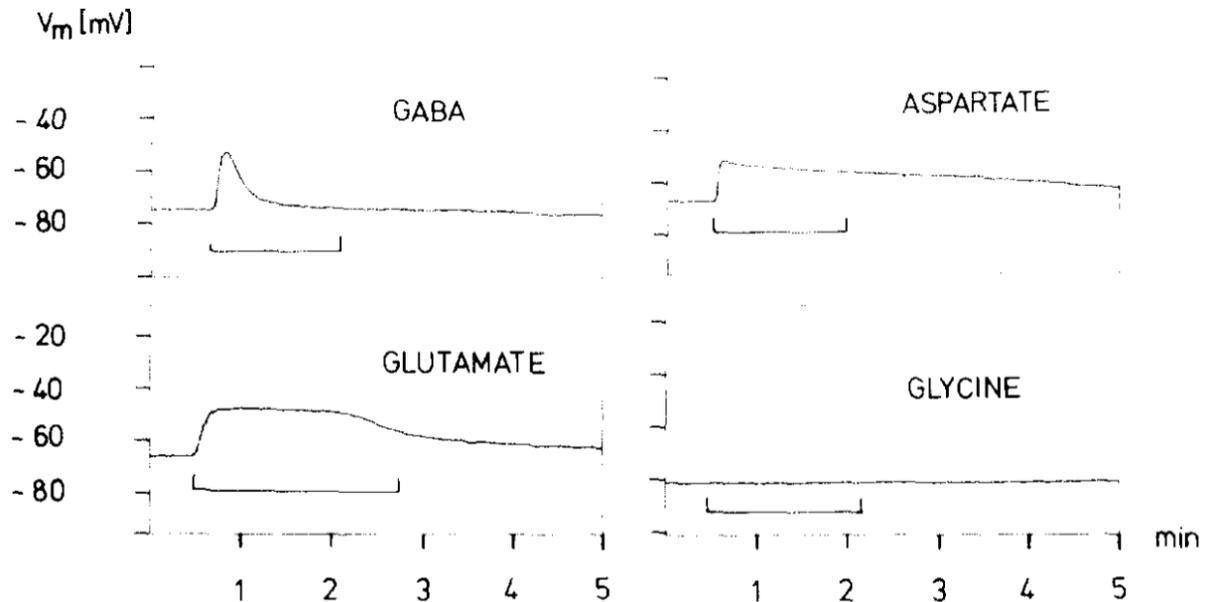
---

[2]mM = millimolar

Figure 1.3: Depolarization of cultured astrocyte membrane by application of various neurotransmitters in the absence of neurons. The membrane potential changes with the presence of GABA, aspartate, glutamate, but not glycine. Taken from Kettenmann et al. (1984).

ceptors, and other various neuropeptides, cytokines and chemokines. Oligodendrocytes express fewer types and include mostly metabotropic P2Y purine receptors, adenosine receptors (A1), ionotropic GluRs (iGluRs). In microglia the most common are purine receptors, iGluRs, metabotropic GluRs (mGluRs), GABA receptors, choline receptors. Finally, Schwann cells express purine receptors and, sporadically, endothelin and tachykinin receptors (Teichberg, 1991; Gallo and Russell, 1995).

1.3 Glial syncytium

Similarly to neurons in the nervous system interconnecting and forming neuronal circuits and networks, macroglia are also linked together and integrated in network called **glial syncytium**. The main difference is in the modality of the cellular connections. Neurons communicate with each other by **synapses** in which the electrical signal in the presynaptic neuron is converted into a chemical signal mediated by neurotransmitters that are bound to postsynaptic neurons, which in turn trigger an electrical response or activation of secondary messengers. Glia, on the other hand, are not linked by synapses, but by **gap**

junctions, which are specialized intercellular connections, where the cellular membranes are significantly close to one another (2–2.5 nm) and the molecular transfer is mediated by intercellular channels, the so-called **connexons**, which are essentially large pores formed by six **connexin** proteins, as shown below in Fig. 1.4.
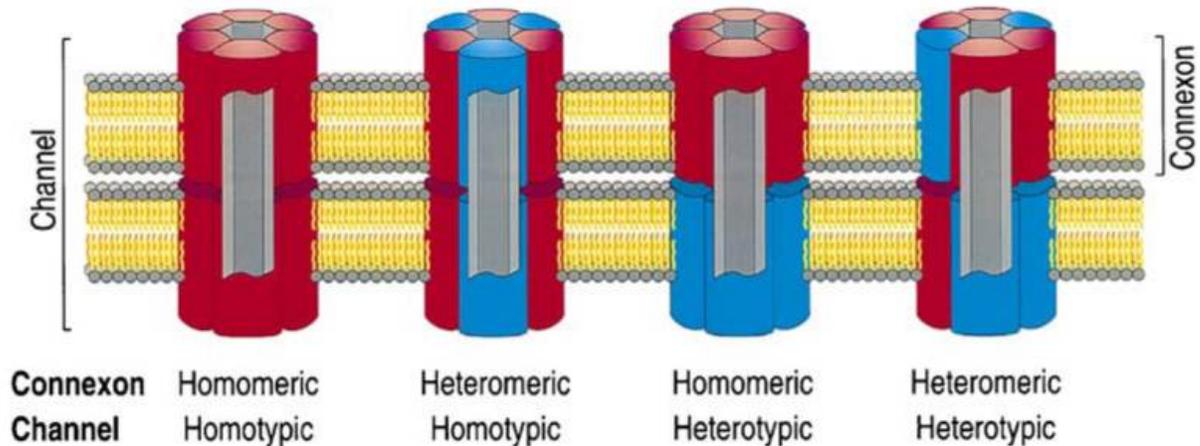

Figure 1.4: Illustration of gap junction between two glia membranes which consists of various connexons. Each connexon can be of either **homotypic** (between the same cell type, e.g. astrocyte–astrocyte), or **heterotypic** (between different cell types, e.g. astrocyte–oligodendrocyte). In addition, each connexon may be made from the same type of protein (**homomeric**), or different types (**heteromeric**). Taken from Kumar and Gilula (1996).

The biophysical behavior of connexons is very similar to the standard ion channels, which operate in two states (open and closed) and can switch rapidly between them. Ion permeability is modulated by several factors such as cytoplasmic Ca^{2+} concentration, acidification, or secondary messenger activation (cAMP, PKC).

Concentration of gap junctions is highest in astrocytes and in the grey matter. Two astrocytes are linked by approximately 230 junctions and one astrocyte connects with 50–100 adjacent astrocytes. Charles et al. (1991), in their work using fluorescent techniques, illustrated the activation of astrocyte syncytium in time as shown in Fig. 1.5. Glial syncytium is not ubiquitous in the brain and the degree of interconnection varies across regions. For example, all cortical astrocytes are integrated in syncytium, whereas only 80% are integrated from the visual center and 50% from the hippocampus. Oligodendrocytes form connections with other oligodendrocytes and astrocytes, but the degree of connectiveness is substantially lower and one oligodendrocyte has only 2 to 4 neighbors. Microglia in syncytium are not integrated and do not create connections with one another (Verkhratsky and Butt, 2007).
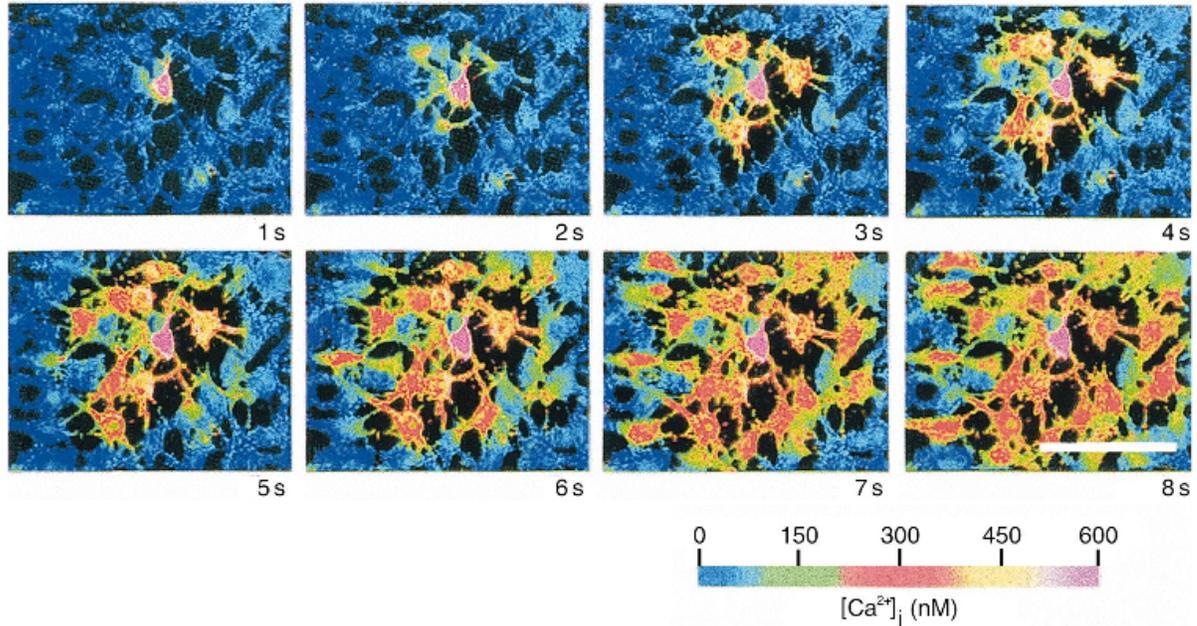
Figure 1.5: Imaging of astrocyte activation within syncytium in time using fluorescence microscopy. Mechanical stress of single astrocyte induces an increase in the Ca^{2+} concentration by influx from extra- and intra-cellular storage, which is subsequently spread into the entire syncytium. Taken from Charles et al. (1991).

1.4 Calcium signaling

The role of calcium ions is ubiquitous in cellular systems and well described. They operate as intercellular messengers important in regulating nearly all known cellular processes. There are, however, exceptions such as the action potential propagation, which depend exclusively on Na^+ and K^+ ions and does not require Ca^{2+}. Physiological responses in the detection of these ions are mediated by multiple cellular sensors, such as various enzymes with different affinities for the Ca^{2+} ions. These enzymes are distributed differently amongst different regions of the cell, thus forming local Ca^{2+} gradients that regulate Ca^{2+} dependent processes. The actual molecular systems responsible for controlling intercellular Ca^{2+} homeostasis and generating calcium signaling are limited to different protein families represented by Ca^{2+} channels and transporters.

Free intracellular calcium is only a small fraction of total calcium in the cell. Because the Ca^{2+} concentrations are unevenly distributed, a potential electrochemical force is generated that imposes Ca^{2+} from high-concentrated to low-concentrated regions, which are separated by biological membranes. The primary ways in which Ca^{2+} moves from one

region to another are mediated by using voltage-dependent or ligand-dependent channels that have different activation mechanisms and different calcium permeability. Glia respond to various electrical, mechanical and chemical stimulation (such as neurotransmitters, neuromodulators and hormones) by increasing the intracellular Ca^{2+} level. These Ca^{2+} signals can then pass through gap junctions to neighboring glia without any decrease in the concentration. In addition, Ca^{2+} responses are consistent with changes in the extracellular environment, indicating Ca^{2+} mediated form of glial excitability. The Ca^{2+} elevations propagate throught syncytium as a wave with the velocity of 10-20 $\mu m s^{-1}$, which travels approximately 100 μm (Finkbeiner, 1995). Blocking gap junctions completely stops Ca^{2+} wave propagation (Enkvist and McCarthy, 1992).

In the grey matter of the central nervous system, astrocytes are able to sense synaptic activity. By expressing the same set of receptors as postsynaptic neurons do and because their membranes are closely associated with the synapses, astrocytes can be co-excited by synaptic transmission. Evidence suggests that the electrical stimulation of neurons in the hippocampus can initiate action potentials in postsynaptic neurons and also Ca^{2+} concentration increase in neighboring astrocytes (Dani et al., 1992). Alvarez-Maubecin et al. (2000) showed using glial cell measurements of *Locus Coeruleus* in mice, that changes in membrane potential are synchronized with the frequency of neuronal oscillations (Fig. 1.6).

1.5   Tripartite synapse

In the grey matter, astrocytes are intimately associated with neuronal membranes and synapses in a way that the astrocytes encircle completely or partially presynaptic and postsynaptic terminals. It is known that about 60% of all synapses in the hippocampus are surrounded by astrocyte membranes (Bushong et al., 2002). The percentage is even higher in the cerebellum where almost all synapses of **Purkinje cells** are matched with membranes of **Bergmann's glia**, a type of astroglial cell in the cerebellum, where each individual Bergmann glia encircles 2000 to 6000 synaptic connections (Reichenbach et al., 1994). These glia–neuron connections are relatively intimate as the average spacing of two neighboring membranes is as close as $1\mu m$ suggesting that the glial cells are able to capture the neurotransmitters secreted at the synapse and provide a corresponding response. This process is initiated by receptors located on the glial membrane, which are expressed according to the dominant neurotransmitters at the synapse. In this sense, the glial cell resembles a postsynaptic neuron to some extent.
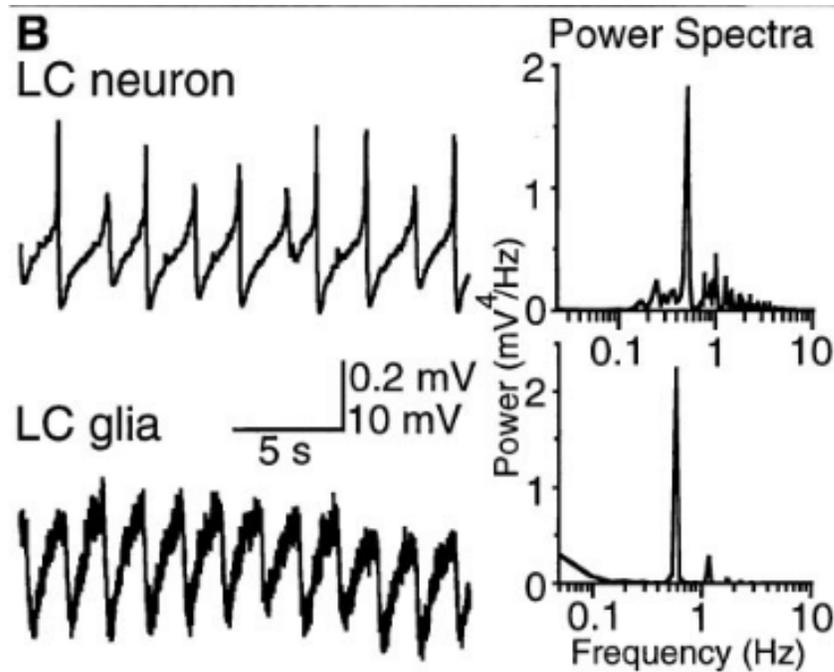
Figure 1.6: Rhythmic oscillations of membrane potential of neurons and glia in Locus Coeruleus. The neuronal activity frequency was 0.51 Hz, whereas the glial response frequency was 0.58 Hz matching the former very well. Taken from Alvarez-Maubecin et al. (2000).

Realizing such a close morphological connection between glia and neurons gave birth to a concept of the **tripartite synapse** (Fig. 1.7). According to the concept, a synapse consists of three equally important components: i) a presynaptic neuron, ii) a postsynaptic neuron, and iii) a glial cell. The neurotransmitter released from the presynaptic neuron activates receptors in the postsynaptic neuron and simultaneously in the glial cell. While in the former the well-known postsynaptic potential is initiated, in the latter, on the other hand, the Ca^{2+} signal arises and begins to propagate to adjacent glial cells through the glial syncytium. This concept is extensively supported by experimental evidence that demonstrates the existence of neuron-to-glia and glia-to-neuron connections.

Neuronal stimulation initiates Ca^{2+} astrocyte signaling *in vitro* and *in situ*. Remarkably, astrocytes can differentiate between intensities of neuronal activity and the frequency of Ca^{2+} oscillation is induced by synaptic activity caused by neuronal stimulation. Low-frequency stimulation of *Schaffer collateral* in the CA1 region of the hippocampus reveals no changes in astrocytes, but the high frequency forces astrocytes to resonate with neurons (Honsek et al., 2012). Similarly to neurons, astrocytes possess cellular memory in which prolonged potentiation is stored after intense synaptic stimulation, which persists even
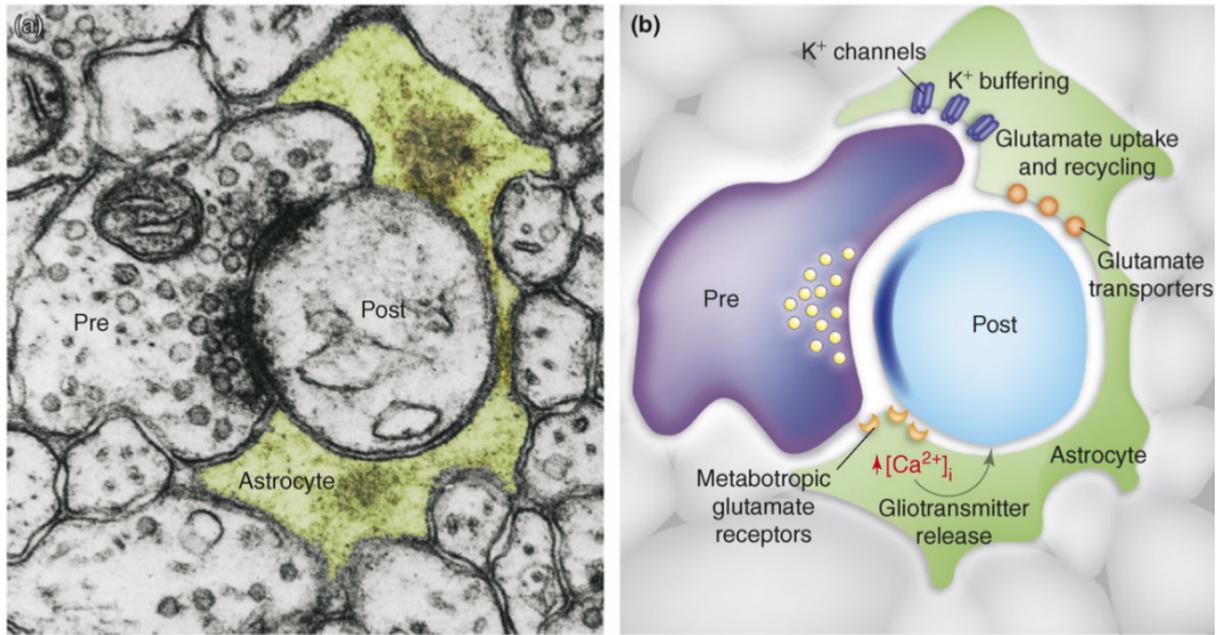
Figure 1.7: The tripartite synapse made of i) a presynaptic neuron, ii) a postsynaptic neuron, and iii) a glial cell. **Left:** electron micrograph illustrating presynaptic (*Pre*) and postsynaptic (*Post*) terminal encircled by an astrocyte. **Right:** illustration showing the role of astrocyte on a synapse. Astrocyte buffers K^+ ions released by the presynaptic neuron using K^+ channels on the membrane, cleans the extracellular space, captures and recycles excess glutamate with glutamate transporters, initiates Ca^{2+} signaling by activating mGluRs and modulates postsynaptic neuron by glutamate release. Taken from Halassa et al. (2007).

after the stimulus ends. This process indeed resembles *long-term potentiation* in neurons where intense synaptic stimulation induces an amplitude increase in postsynaptic potential. The difference in astrocytes is that the response amplitude does not increase, but the frequency of Ca^{2+} signaling does (Verkhratsky and Butt, 2013).

Ca^{2+} waves are commonly initiated in two ways: 1) as a response to neuronal activity, or 2) spontaneously. Glial activity has the potential to release neurotransmitters from intracellular stores (called **gliotransmitters**), which directly affect close neurons (Malarkey and Parpura, 2008). Such glia–neuron signaling is mediated by ionotropic and metabotropic receptors present on the neuronal membrane. The release of glutamate by an astrocyte can therefore directly depolarize and thereby modulate neuronal excitability (Köles et al., 2016). Jourdain et al. (2007) clarified the way in which synaptic modulation is accompanied by the release of glutamate into the presynaptic terminal space where NMDA receptors are activated to enhance the synaptic transmission.

As already stated, release of glutamate by astrocytes is regulated by the frequency of Ca^{2+} oscillations, where each increase in Ca^{2+} induces pulse release of glutamate. Since the frequency of Ca^{2+} oscillations is controlled by the intensity of synaptic activity, its increase causes significant release of glutamate by astrocytes, which ultimately enhances and increases the original synaptic signal. In addition, release of glutamate is also present outside synaptic connections, therefore activating GluRs and regulating neuronal membrane potential. It is important to note that Ca^{2+} oscillations are present throughout the full glial network, suggesting that glutamate is released intermittently by individual astrocytes, as Ca^{2+} waves progress deeper within syncytium. Therefore, distant neurons that were not originally part of the synaptic activity controlled by Ca^{2+} signaling, are also modulated. Moreover, glutamate release by single astrocyte is not limited to modulation of one, but multiple neurons.

Chapter 2

Morphology of glial cells

2.1 Astrocytes

Astrocytes (or generally astroglial cells) represent heterogeneous cell populations in the nervous system and are considered to be the most abundant and functionally richest cells. According to Privat et al. (1995), they are morphologically divided into several basic classes due to their star shape (Fig. 1.1): **protoplasmic astrocytes** found in grey matter and spinal cord, **radial astrocytes**, which are pluripotent neuronal precursors during development, **fibrous astrocytes** in white matter, radial **Müller glia** found in the vertebrate retina and **Bergmann glia** cooperating with the *Purkinje neurons* in the cerebellum.

Since astrocytes in the nervous system participate in range of processes, the functions themselves can be conceptually divided into the following categories: developmental function, structural function, vascular function, metabolic function and homeostatic function.

2.1.1 Developmental function

Neurogenesis is present in vertebrates throughout their lives, but it only occurs in certain areas of the brain. Fish (Zupanc, 2006) and reptiles (Font et al., 2002) have multiple proliferative zones believed to be capable of providing neurons to any region of the brain, whereas in mammals, neurogenesis is, according to current knowledge, limited to hippocampus and olfactory cortex (Eriksson et al., 1998). Immature astrocytes are precursor cells that, during their existence, differentiate into neurons or mature astrocytes. The neurons formed in the subventricular zone travel to the olfactory cortex, but the hippocampus-born neurons remain in the place and integrate into neuronal networks.

Gliogenesis, as opposed to neurogenesis, occurs in all brain areas. The type of glia cell is predominantly governed by the region where the glia is born – oligodendrocytes dominate in the subcortical white matter, but in the spinal cord the astrocytes and oligodendrocytes are equivalent.

While being alive, the brain constantly creates new synapses, eliminates old ones and strengthens or weakens some of them. These processes take place when interacting with the outside world and are essential for learning and memory. Neurons themselves produce a large number of synaptic proteins important for synaptogenesis (Scheiffele, 2003). Interestingly, before triggering synaptogenesis in the CNS, neurons supply the target region at least a week in advance, while it takes the very same time to produce new astrocytes in the region (Pfrieger and Barres, 1997). This interval, common to neurons and astrocytes, suggests the possibility that astrocytes are necessary to control the formation of a new synapse.

Pfrieger and Barres (1997) showed, in purified neuronal cultures, that although synaptogenesis occurs in the absence of astrocytes, it does so at a slow pace. However, the addition of astrocytes can accelerate this process up to 100 times. There are several possible explanations and it is assumed that synaptogenesis is thought to be almost impossible without the presence of the cholesterol, which is important in the formation of membranes and is produced and supplied by astrocytes. In addition, astrocytes also produce specific proteins essential for synaptogenesis.

2.1.2 Structural function

Each astrocyte in the grey matter controls its own anatomical domains with the minimal overlap with neighboring astrocytes (Fig. 2.1). Within their territory, astrocytes wrap around adjacent blood vessels, neurons, synapses, and other nearby glia, which are integrated into astrocytic processes (Verkhratsky and Butt, 2007).

2.1.3 Vascular function

Astrocytes play several vascular functions, including the maintenance of the blood–brain barrier, and the regulation of vascular tone. In the formation and maintenance of the blood–brain barrier they express the necessary endothelial proteins, transporters and enzymes (Abbott et al., 2006). With an increased neuronal activity, astrocytes accelerate blood flow in a given area so that neurons can get the energy they need to function – the process called **hyperaemia**. The Ca^{2+} signaling within the syncytium is triggered
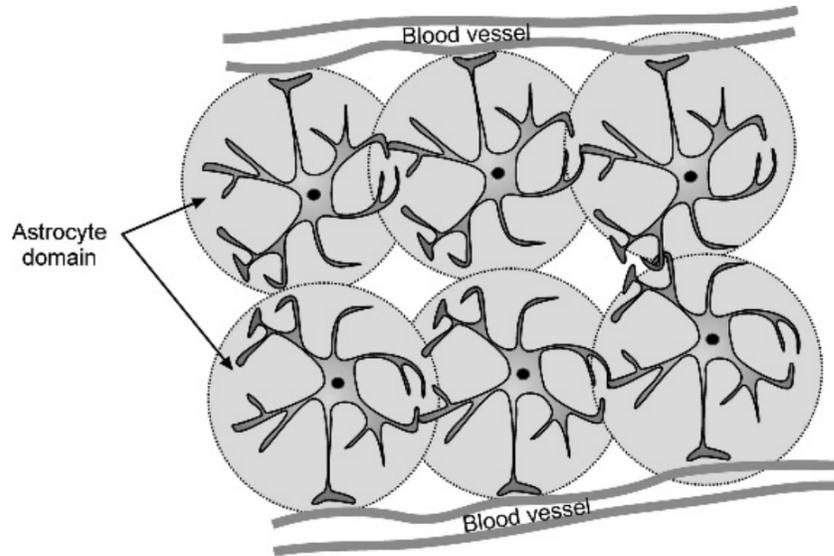
Figure 2.1: Anatomical domains of astrocytes. Taken from Verkhratsky and Butt (2007).

which leads to the release of several important vasoactive substances such as arachidonic acid, prostaglandin E2, or K^+ ions that ultimately regulates the tone of vascular smooth muscles (Attwell et al., 2010).

2.1.4 Metabolic function

The idea that the role of astroglial cells is to provide nutrients to the neurons is more than a century old with Golgi (1885) being the first to propose it. A few decades later, several studies confirmed this concept and showed by imaging techniques that astroglia produce several important nutrients for neurons. The main ones are glucose and oxygen, without whom the life of neurons would not be possible. Rouach et al. (2008) showed that in hippocampus samples, the fluorescent glucose is spread to the entire astrocyte network 20 minutes after injection into a single astrocyte. They further demonstrated that in a metabolic demand, where neurons were glucose deprived, synaptic transmission was inhibited and glucose delivery to a single distant astrocyte revived synaptic transmission. Similar research was done by MacVicar and Newman (2015), who showed in samples of rat hippocampus that neuron stimulation at high/low oxygen levels causes arterial constriction/dilation, respectively. Since astrocytes have limited capabilities for a long-term storage of glucose in the form of glycogen, they instead convert it into a more easily metabolized lactate and pass it to the neuronal mitochondria (Hertz, 2004).

2.1.5 Homeostatic function

Proper levels of extracellular ions are critical to the overall functioning of the brain, since each fluctuation affects the membrane potential that determines neuronal excitability. The extracellular space is occupied predominantly by Na^+, Cl^- ions and in a smaller concentration by K^+ and Ca^{2+} ions. During neuronal activity, ion levels fluctuate and it is the astrocytes themselves that govern the **homeostasis** and maintain the proper concentrations. Throughout action potential propagation, they primarily regulate K^+ ions that are released by neurons into the extracellular space. If K^+ ions will accumulate in the extracellular space, it would cause long lasting membrane depolarization and refractory period, since high levels of K^+ inactivate Na^+ channels. By absorbing K^+ and preventing accumulation in the extracellular space, astrocytes allow neurons to generate next action potentials. Moreover, astrocytes are not limited to regulation of only K^+ ions, but regulate all types present in the CNS (Gee and Keller, 2005).

Not only do astrocytes regulate extracellular ion concentrations, but they also maintain pH level using bicarbonate and proton transporters, and by withdrawing HCO_3^-. Correct pH concentration is physiologically important because even small fluctuations can significantly affect synaptic transmission and neuronal excitability. For example, when consuming glucose by neurons during neuronal activity or synaptic transmission, the CO_2 and H^+ are produced as a metabolic waste leading to pH fluctuation (Chesler, 2003). Chen et al. (1998) demonstrated that lowering pH below 7.0 can completely inhibit NMDA receptors.

Glutamate, an anion of glutamic acid, is the major excitatory neurotransmitter exocytotically released by presynaptic terminals. In glutamatergic neurotransmission, the glutamate is secreted to the synaptic cleft where approximately only 20% activate the receptors of the postsynaptic neuron and the remaining 80% are collected by astrocytes. Since redundant glutamate is neurotoxic, it is necessary to remove it from extracellular space. Collected glutamate is metabolized by astrocytes to glutamine and forwarded back to neurons whom recycle it back to the glutamate. Neurons per se are dependent on astrocytes for glutamine, since they are unable to recycle it and the necessary enzyme, *glutamine synthetase*, is astrocyte specific (Rosenberg, 1991).

2.2 Oligodendrocytes

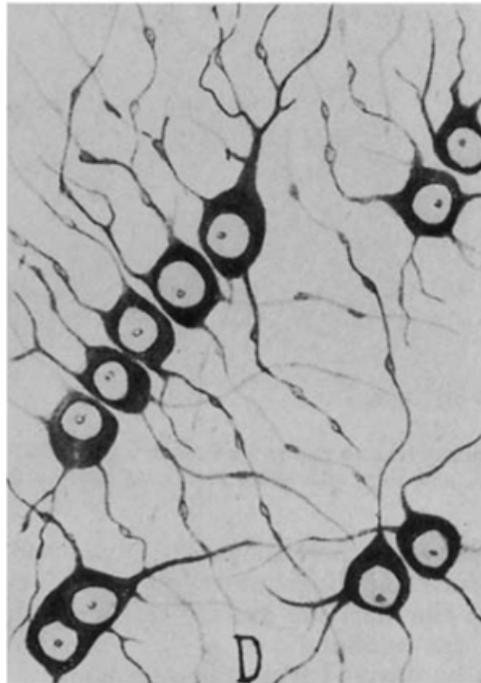Del Río Hortega (1921) identified two new cell types in the CNS: **oligodendrocytes**

Figure 2.2: Del Río Hortega's illustration of oligodendrocytes in white matter. Taken from Somjen (1988).

and **microglia**. A characteristic feature of oligodendrocytes (Fig. 2.2) is the small soma containing large amounts of chromatin and the relatively low number of dendrites.

Del Río Hortega distinguished four types of oligodendrocytes: **type I** producing multiple different myelin segments for small, variable axons, **type II** similar to type I, but for parallel axons, **type III** myelinating axons of larger dimensions and **type IV** that has a soma shape similar to Schwann's cells.

The primary role of oligodendrocytes is to produce myelin for neuronal axons, which serves as an insulator of axonal segments and which is necessary for neuronal transmission at high speeds, up to 200 ms^{-1}. Since the axon may be **unmyelinated** or **myelinated**, the action potential may be propagated in two ways: by *continuous* or by *saltatory*[1] conduction, respectively. Each oligodendrocyte produces several myelin segments that are interrupted on the axon by **nodes of Ranvier**, which densely populate Na$^+$ channels on the membrane allowing saltatory conduction. Action potentials arise only in these nodes (not within a myelin), where the depolarization occurs, which rapidly hops to subsequent nodes and restores the action potential there. This is a quicker way compared to the unmyelinated axon and, moreover, it saves energy because Na$^+$ ions only accumulate at

---

[1]from Latin language, *saltare = to hop*

these nodes. After the action potential has been propagated, Na^+ ions must be released into the extracellular environment, governed by Na^+/K^+-ATPase. In unmyelinated axons, there are more Na^+ channels which are randomly distributed along the axon.

Myelination is a complex process driven by axons consisting of (1) oligodendrocyte progenitor cell (OPC) proliferation and migration in white matter, (2) target axon recognition and localization, (3) OPC differentiation to myelinating oligodendrocytes, (4) surrounding the target axon, (5) transferring important biochemical molecules to axons, (6) production of the myelin and (7) formation of a myelin node (Barres and Raff, 1999).

Oligodendrocytes are predominantly located in white matter, but are also present in grey matter. Like astrocytes, they are integrated in glial syncytium using gap junctions. In addition to providing myelin, oligodendrocytes play a supportive trophic role for neurons by producing brain-derived neurotrophic factor (BDNF) and insulin-like growth factor-1 (IGF-1) (Bradl and Lassmann, 2010).

2.3 Schwann cells

The myelin sheath was first observed in the peripheral nervous system (PNS) by Remak (1838), who described and illustrated the cells bound to the axon membrane, but it was actually Schwann (1839) who studied them at roughly the same time and gave them a name. Schwann cells play a similar role as olidendrocytes in the formation of myelin, except that they work in the PNS (oligodendrocytes in the CNS) and they form only one full segment of axon isolation. Three types of Schwann cells are distinguished: (1) **unmyelinating Schwann cells** that play similar roles as astrocytes in the CNS – structure formation, metabolic support, regulatory function, (2) **myelinating Schwann cells**, which are about as much as non-myelinated, but are responsible for myelin production and (3) **perisynaptic Schwann cells** maintaining axon terminals which they encapsulate and hence maintain proper synaptic functionality. More than that, they regulate the development of axon terminals and efficacy of synaptic transmission by controlling perisynaptic ion levels (Mirsky and Jessen, 1999).

In addition, Schwann cells are involved in the growth of neurons and various neuronal structures in the PNS by providing trophic support (Riethmacher et al., 1997). Lin et al. (2000) showed that in mice, deficient in the gene erbB2, development disorders of Schwann cells occur, which are accompanied by misformation of neuromuscular synaptic connections and motor axons. Other authors, e.g. Reddy et al. (2003), confirmed that by removing

perisynaptic Schwann cells the formation of new synapses is dramatically reduced and the existing synapses retract.

2.4 Microglia

Last but not least, important group of glial cells are microglia (in the CNS), which represent about 10% of all glial cells. Microglia are a class of macrophages and play an important role in the development of CNS, homeostasis and almost all pathological conditions of the CNS. They protect immature neurons and glia from glutamate-induced apoptosis, but also contribute to the opposite: initiating programmed cell death in defective neurons. Microglia modulate synaptic activity by regulating density of synapses and glutamate receptors (Ji et al., 2013). They are distributed throughout all brain structures and are most concentrated in the hippocampus, olfactory telencephalon, basal ganglia and substantia nigra. The morphological difference of the microglia in grey and white matters is well described. While in grey matter they are distributed in all directions, in white matter they actually concentrate only within the axon bundles.

In the **resting state**, microglia remain stationary, but their internal processes constantly monitor the extracellular environment and directly interact with neurons, astro-
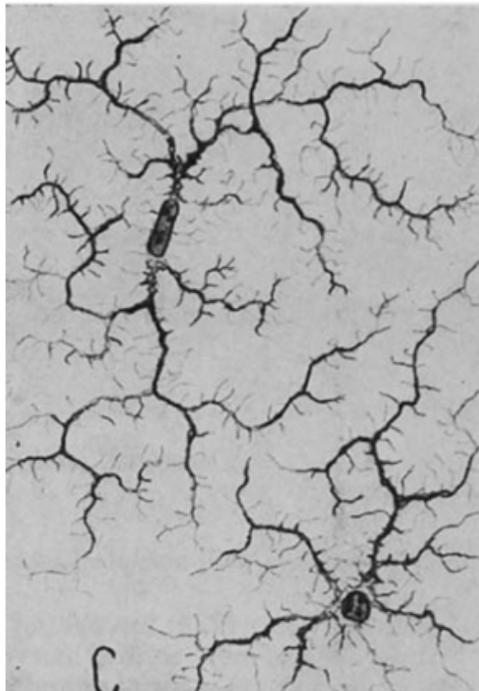

Figure 2.3: Del Río Hortega's illustration of microglia. Taken from Somjen (1988).

cytes and blood vessels. This continuous state of motion allows them to respond promptly to any CNS threat, such as damage or infection. If that happens, they immediately transform into the **active form**, which is accompanied by morphological changes that allow them to quickly deal with the current problem. This conversion is observed in almost all neuropathological conditions, such as degenerative disorders (Alzheimer's disease, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis), infections (bacterial, viral and parasitic), stroke, tumors, or injuries. Remarkably, microglia have a pattern detection algorithm that allows them to detect foreign and potentially harmful substances. Upon detecting something potentially toxic, they respond by producing pro- and anti-inflammatory cytokines, chemokines, free radicals, various tropical factors, and other substances. In addition, they proliferate and migrate to areas where inflammation occurs (Nayak et al., 2014).

Chapter 3

Related work

Computational neuroscience distinguishes two modeling paradigms: **biophysical** and **connectionist**. While the former focuses on physical and chemical properties of a biological system using various mathematical methods, the latter makes a significant reduction in the complexity of low-level mechanisms which in turn may lead to better comprehension of the system at a higher level. Despite the plethora of biophysical models of astrocytes per se and neuron–astrocyte coupling, connectionist modeling has so far been out of scientific interest. For an overview of biophysical models, we recommend Oschmann et al. (2018), Volman et al. (2012) and Wade et al. (2014).

It should be noted that there is also a another way of modeling the neuronal systems using digital circuits. We refer to articles Joshi et al. (2011), Irizarry-Valle et al. (2013) and Irizarry-Valle and Parker (2015), where instead of computer modeling, the authors chose the neuromorphic engineering and designed CMOS circuits with a small neural network extended with "digital" astrocytes to modulate neuronal excitatory postsynaptic potentials. However, since this is not related to our thesis in terms of modeling modality, we do not delve into this topic.

3.1 Modeling of neuronal regulation

Ikuta et al. (2009) first came up with the concept of artificial glia in ANNs, and in the course of a few years they proposed several models taking into consideration cooperation between neurons and glia. The common architectural layout present in their published papers is the traditional MLP with the extension of the hidden layer by artificial glia, as illustrated in Fig. 3.1. The net function for hidden neurons is calculated as the linear
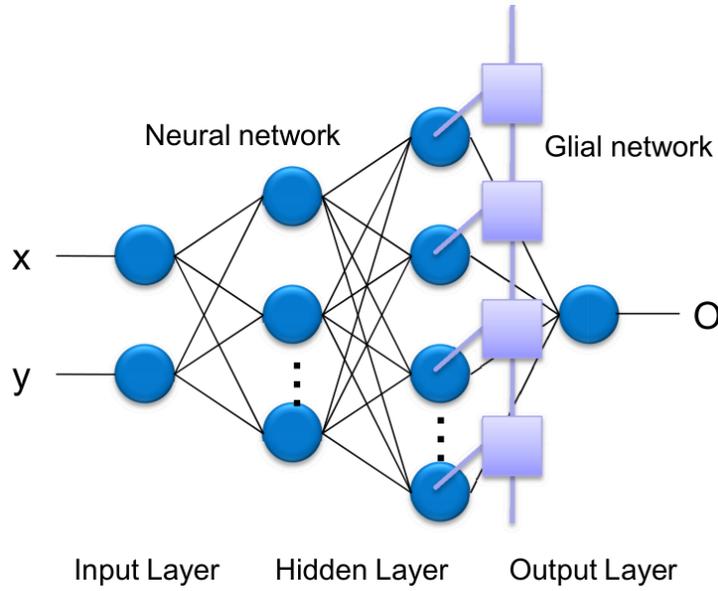
Figure 3.1: An MLP with the second hidden layer extended with a chain of artificial glial units. Taken from Ikuta et al. (2009).

combination of weights and neurons from the previous layer together with the activation of glia weighted by α (eq. 3.1). The glial activation $\Psi_i(t)$ is the sum of m transient glial outputs, $\psi_{i+k}(t)$, each weighted by the factor $\beta^{|k|}$ where $k \in \{-m, (-m+1), \ldots, m\}$. The exact formula is stated in eq. 3.2.

$$h_i^{(2)}(t+1) = f(\sum_{j=0}^{N} w_{ij}(t)h_j^{(1)}(t) + \alpha\Psi_i(t)) \tag{3.1}$$

$$\Psi_i(t) = \sum_{k=-m}^{m} \beta^{|k|}\psi_{i+k}(t) \tag{3.2}$$

The use of specific function $\psi_i(t)$ is the subject of the particular article, however an overall architecture remains the same. In Table 3.1 we summarize the meaning of each parameter and hyperparameter common for all networks that we further describe in the following sections. Although the authors published multiple papers (up to 20), we review only the papers we found to have the greatest contribution.

3.1.1 Chaos glial network

In their very first model, Ikuta et al. (2009) attempted to solve the problem of two spirals interleaved together (highly nonlinear problem), as shown in Fig. 3.2. The architecture of the network had 2 input neurons (x and y coordinate), 20 neurons on the first hidden

$\Psi_i(t)$	final activation of i-th glia
$\psi_i(t)$	transient activation of i-th glia
α	shared glial weight
β	attenuation factor
m	propagating range of transient glia activation

Table 3.1: Parameters and hyperparameters used across various papers by Ikuta et al.

layer, 40 in the second hidden layer (extended with artificial glia, one for each neuron) and a single output neuron.
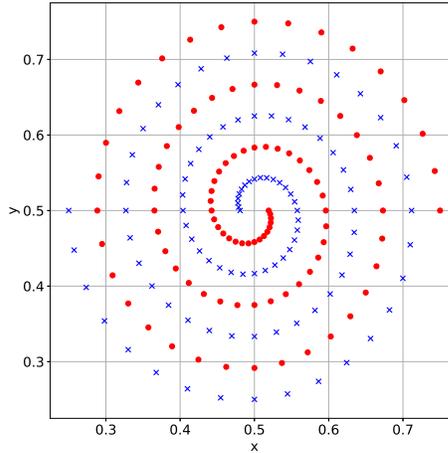

Figure 3.2: Two spirals problem. The input is a two-dimensional point (x, y) and the task is to classify it to either red (o) or blue (x) spiral.

In ANNs it is well known, that adding white noise to either input data or to error gradients improves performance and increases generalization (Murray and Edwards, 1994). Taking that into account and the fact that astrocytes in the CNS spontaneously generate Ca^{2+} waves that propagate to neighboring units (as stated in previous chapters), the authors combined both ideas and proposed artificial glia connected in a single chain generating chaotic oscillations defined as

$$\psi_i(t+1) = \begin{cases} \frac{2\psi(t)+1-A}{1+A}, & -1 \le \psi(t) \le A \\ \frac{-2\psi(t)+1+A}{1-A}, & A < \psi(t) \le 1 \end{cases} \quad (3.3)$$

The achieved performance was better compared to traditional MLP and authors claim that chaotic noise generated by glia allows the network to avoid getting stuck in local minima. At the same time, the model worked better than the model with simple white noise instead of glia (Fig. 3.3).
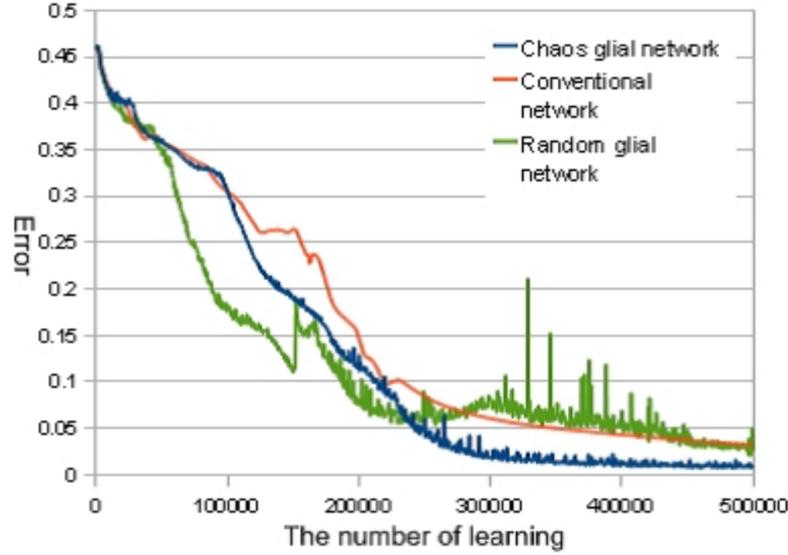
Figure 3.3: Error rate for three models on two spirals problem. The traditional MLP with no glial units (orange line) achieved comparable results as the MLP with random noise (green line), although with faster rate of convergence. The proposed model yielded the best performance (blue line). Taken from Ikuta et al. (2009).

3.1.2 Impulse glial network

The model proposed in the following work (Ikuta et al., 2010) is largely based on the previous one, but the authors included bidirectional interaction between glia and neurons instead of unidirectional chaotic noise. The computation of the transient glial output is given as

$$\psi_i(t) = \begin{cases} 1, & \theta_n < h_i(t-1) \quad \text{and} \quad \theta_g > \psi_i(t-1) \\ \gamma\psi_i(t-1), & \text{otherwise} \end{cases} \tag{3.4}$$

and the final glial output is computed as

$$\Psi_i(t) = \sum_{k=-m}^{m} \beta^{|k|}\psi_{i+k}(t-|k|) \tag{3.5}$$

Three new hyperparameters are introduced and the activation of the glial unit itself is determined directly by the neural activity. The glia is excited when the neuron exceeds the threshold θ_n and at the same time the glia is not within the interval of the refractory period given by θ_g. When excited, its activity spreads to neighboring glia and at the same time exponentially decays by γ.

The problem the authors used for the assessment of the performance was the classi-

fication of two chaotic time series given by four adjacent points (Fig. 3.4). Hence, the network had 4 input neurons, a single hidden layer of 10 neurons and glia and one output neuron to distinguish between the two time series.
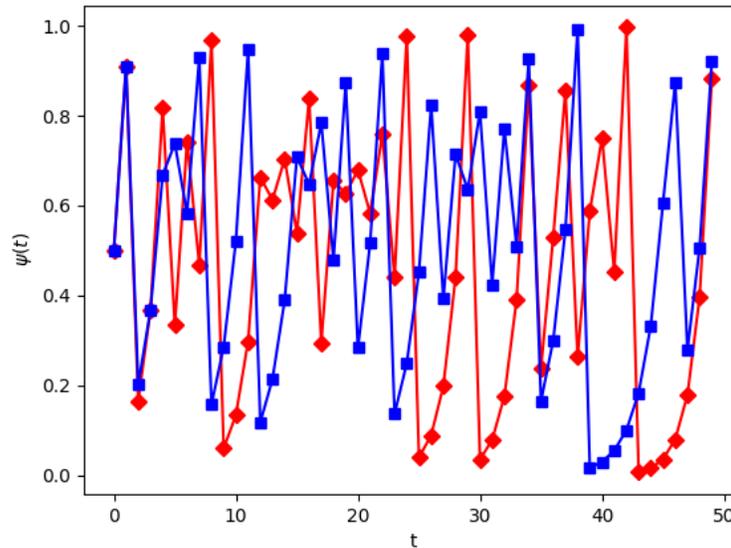

Figure 3.4: Two chaotic time series dataset used by the authors for the evaluation of their models. By presenting four neighboring time points, the task is to identify the first or the second time series.

In the experiments, four different models were compared: 1) traditional MLP, 2) MLP with random noise, 3) MLP with chaotic glial noise (explained in Section 3.1.1) and 4) MLP with impulse glial network. Results in Fig. 3.5 reveal that the basic MLP apparently got stuck in the local minima and yielded the highest MSE, while the proposed network performed the best.

3.1.3 Network with local glial connections

In the next work, Ikuta et al. (2013a) included the concept of neural inhibition controlled by glial units and local glial connections. Instead of connecting one neuron with one glia, the glia now sums neural activity over the range of neurons amongst both hidden layers, although the specific formula for the computation is not present in the paper. Architecture of the network is illustrated in Fig. 3.6 and the activation function for transient glia activation is defined in eq. 3.6[1].

---

[1]Authors actually used different symbols for θ_g and θ_n, but we have used the symbols as stated to stay consistent with the previous notation.
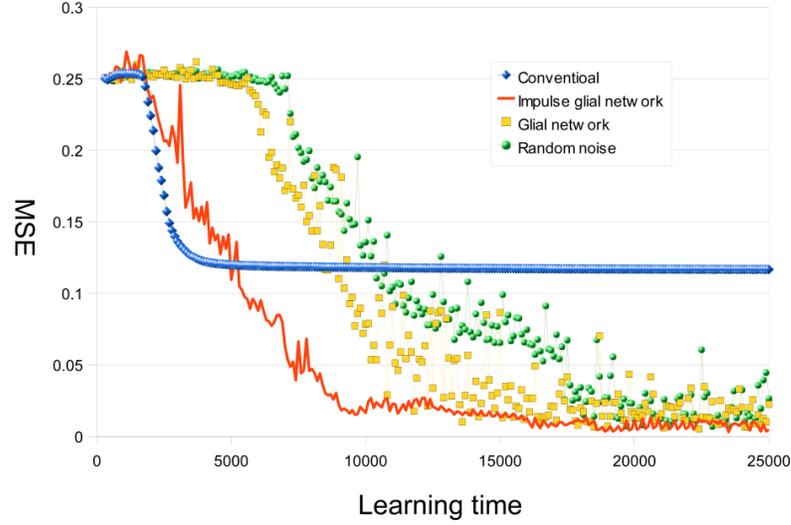
Figure 3.5: Comparison of four different models on the task of two series classification. The worst was the basic MLP, whereas the best was the proposed model with the impulse glial network. Taken from Ikuta et al. (2010).
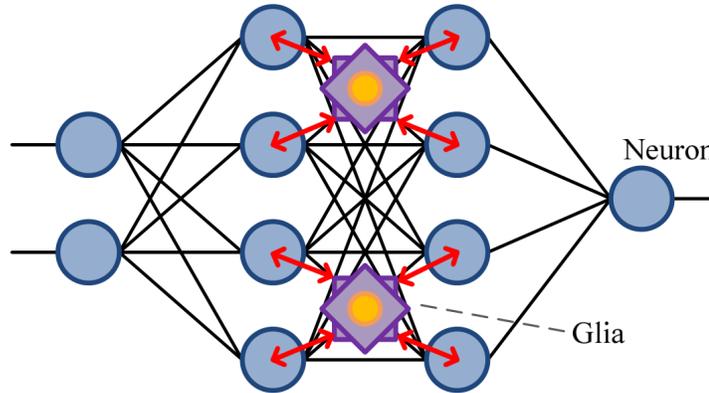
Figure 3.6: MLP with two hidden layers where each glia sums activity from specific range of neurons amongst both layers. Taken from Ikuta et al. (2013a).

$$
\psi_i(t+1) = \begin{cases} 1, & I_i = 0, & \text{if} \quad h_i(t) > \theta_n \quad \text{and} \quad I_i > \theta_g \\ -1, & I_i = 0, & \text{else if} \quad h_i(t) < -\theta_n \quad \text{and} \quad I_i > \theta_g \\ \gamma\psi_i(t), I_i = I_i + 1, & \text{otherwise} \end{cases} \tag{3.6}
$$

In case of high neural activity, when θ_n is exceeded, the glia is excited by outputting value 1. On the contrary, when a neuron produces a value lower than $-\theta_n$, the glia returns -1. In both cases, the refractory period interval must be satisfied, which is defined by a

discrete number of steps determined by θ_g.

The authors tested performance of the proposed network (with 10 neurons in both hidden layers) again on two spirals problem and obtained better results as compared to traditional MLP. However, they did not compare the model with previous networks with artificial glial units (chaotic or impulse).

3.1.4 Pulse glial network with a dynamic period of inactivity

Ikuta et al. (2013b) proposed the MLP with a pulse glial network having a dynamic period of inactivity. The architecture is the same as in previous models, with a hidden layer extended by artificial glia that is excited by hidden neurons. Upon being excited, the glial activation (pulse) propagates back to neurons and to neighboring glia. Whereas in the previous networks, the period of inactivity was fixed and given by θ_g, in the proposed model the period varies dynamically according to the activity of neighboring glia. This behaviour is formalized as

$$
\psi_i(t+1) = \begin{cases} 1, & [\theta_n < h_i(t) \quad \text{or} \quad \psi_{i+1}(t-iD) \equiv 1 \quad \text{or} \quad \psi_{i-1}(t-iD) \equiv 1] \\ & \text{and} \quad \tau_i \geq \theta_{gi} \\ \gamma\psi_i(t), & \text{otherwise} \end{cases} \tag{3.7}
$$

The new hyperparameter D is a delay time of a glial effect and parameter τ_i is local time present in each glia during a period of inactivity. The dynamic period of inactivity, θ_{gi} is described as

$$
\theta_{gi}(t+1) = \begin{cases} \theta_{gi}(t) - 1, & \psi_i(t) = 1 \quad \text{and} \quad \psi_i(t - \theta_{gi}(t)) = 1 \\ \theta_{gi}(0), & \text{otherwise} \end{cases} \tag{3.8}
$$

The dataset for evaluating the performance was once again two spirals problem and the architecture consisted of a single hidden layer with 40 neurons. In the experiments, the proposed model yielded better results than all the models described so far (Fig. 3.7).

3.1.5 MLP with a pulse glial network and neurogenesis

The slightly different approach is presented in Ikuta et al. (2014) where the authors came up with a concept of neurogenesis driven by glial units (illustrated in Fig. 3.8). The
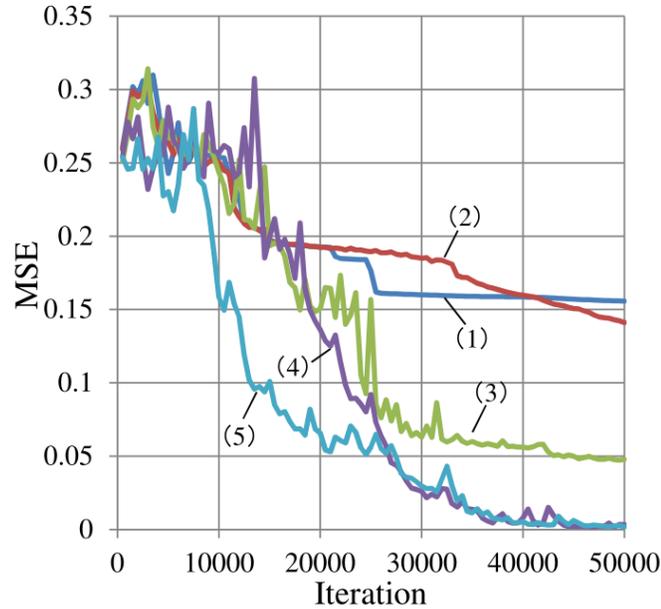
Figure 3.7: MSE for five models on the problem of two spirals: (1) traditional MLP, (2) MLP with random noise (3) MLP with a static period of inactivity, (4) MLP with a random period of inactivity, (5) proposed MLP with a variable period of inactivity. Taken from Ikuta et al. (2013b).

computation for activation output of neural and glial units stays the same as in previous model from Section 3.1.4. If, at a given time, the amount of glial activation does not exceed a preselected threshold (meaning that the corresponding neuron is not sufficiently active), the neuron is canceled and replaced with a new one with random pre- and post-synaptic weights. Thus, the total number of neurons is always constant and less important neurons are given a new chance.

Figure 3.8: Glia-controlled neurogenesis in an MLP. If activation of glial units is not sufficient, meaning that the associated neuron give only little contribution, the pre- and post-synaptic weights of the neuron are randomized to give the neuron a new chance. Taken from Ikuta et al. (2014).

Ikuta et al. showed that a proposed network performed better on a problem of two spirals compared to the model with pulse glial network without neurogenesis as described in previous sections.

3.1.6 Learning controlled by glial units

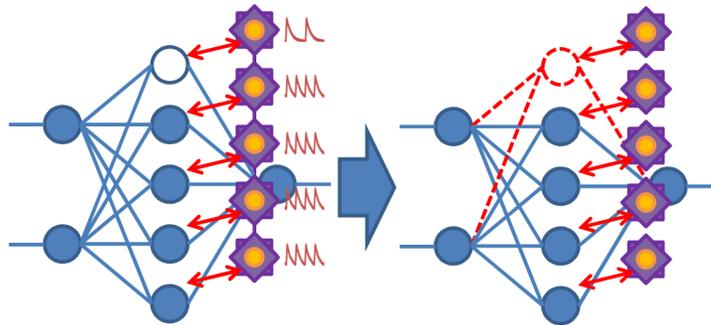A very similar idea was presented in Ikuta et al. (2013c), where a hidden layer was divided into multiple groups of neurons and each group was assigned with a single glial unit to control whether the associated group will be allowed to learn or not. The decision for learning is based on the glial activity and a higher activity switches neurons to the learning mode, whereas the lower activity stops it (illustration of the concept is in Fig. 3.9). The authors did not provide exact equations for glial activation, but it may be assumed that they used some of them mentioned in previous sections. Weights of the neurons could only be updated if glial activation are within a certain range, but the authors did not provide a specific interval in the article. Simulations on two spirals problem (shown in Fig. 3.10) confirmed that the proposed model is more successful than traditional MLP and MLP with random glial noise.
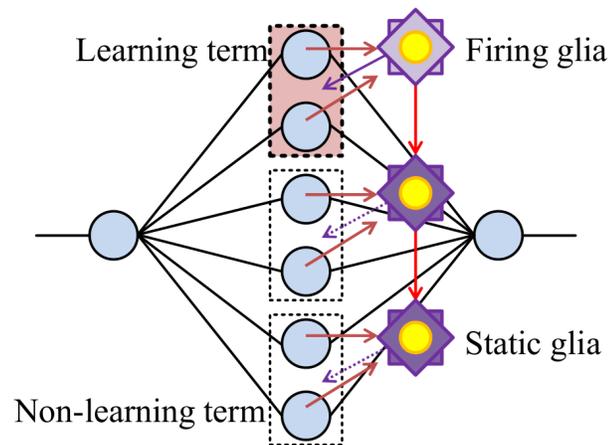
Figure 3.9: The concept of learning controlled by glial units. The hidden layer is divided into groups of learning and non-learning neurons. The learning is decided by the assigned glia based on their activity. Taken from Ikuta et al. (2013c).

3.1.7 Hopfield neural network with glial units

The last paper we present from the same authors is Ikuta et al. (2012) where the Hopfield network is proposed with the addition of artificial glial units. The neurons are placed in
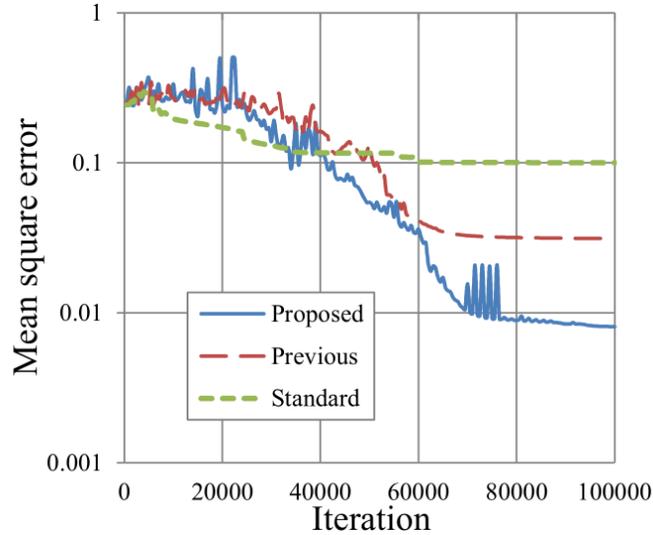
Figure 3.10: Learning curves on the problem of two spirals. Comparison of basic MLP (green line), MLP with random glial noise (red line) and the proposed model (blue line). Taken from Ikuta et al. (2013c).

a two-dimensional grid and each neuron is paired with a single glia that listens to and regulates its activity (Fig. 3.11).
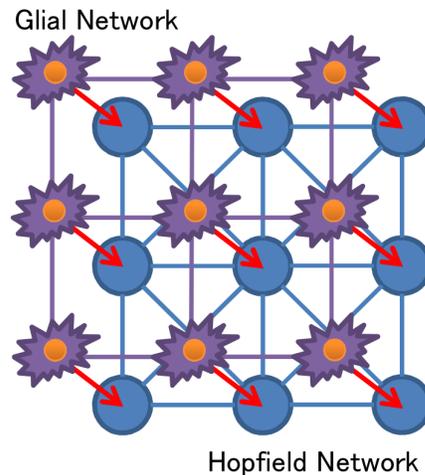
Figure 3.11: Architecture of the Hopfield neuro-glial network. Neurons are arranged in a grid, and each neuron is associated with exactly one glia. Taken from Ikuta et al. (2012).

Neural activity update is determined as

$$\boldsymbol{u}(t+1) = f(\boldsymbol{W}\boldsymbol{x}(t) - \boldsymbol{h} + \boldsymbol{\phi}(t)) \tag{3.9}$$

where $\boldsymbol{x}$ is the vector determining neural state, $\boldsymbol{W}$ is the weight matrix of neural connec-

tions, $\boldsymbol{h}$ is the vector of neural thresholds and $\boldsymbol{\phi}(t)$ is the vector of glial activations.

The exact formula for glial activation is not stated in the paper, however the authors mention that glia is excited when neurons exceed a preselected threshold, then it stays inactive (not excitable) for some given period, while the output exponentially decays, suggesting the use of eq. 3.4.

During the simulation, the authors evaluated the performance on the traveling salesman problem. In each step they selected 8 neurons with the highest output that activated glia (Fig. 3.12), while the rest of glia decays and activates their neighbors. The proposed model provided better results than the basic Hopfield model, but actually worse than the model with random noise.

(a) t=1. (b) t=2. (c) t=3.

(d) t=4. (e) t=5. (f) t=6.
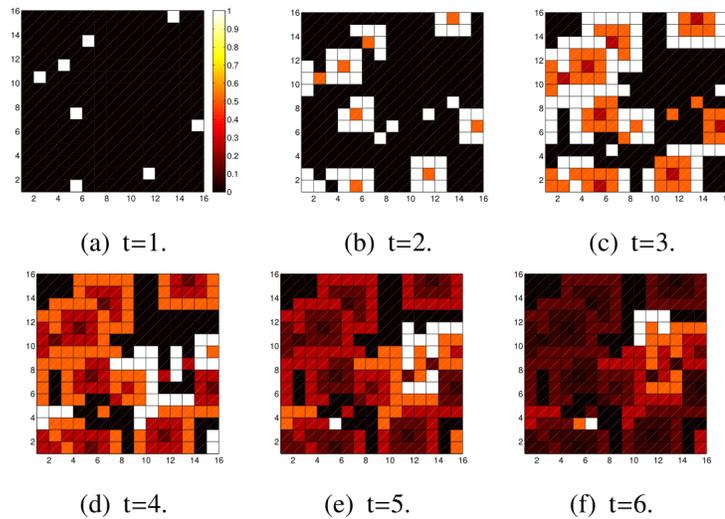
Figure 3.12: Glial activation in each time step in the Hopfield network with glial units. Taken from Ikuta et al. (2012).

3.1.8 Model SONG-Net

Finally, the last model we found in the literature dealing with neuronal regulation by glial cells comes from Marzouki (2015) where the author combined MLP with the Kohonen's self-organizing map (SOM) (Kohonen, 1982). Instead of using traditional fully-connected hidden layer, the author decided for Kohonen's layer which computes the distance between an input pattern and the weights of hidden neurons using the Euclidean metric, which is finally used for the computation of output layer activation. The architecture layout is illustrated in Fig. 3.13.

By evaluating the performance on four tasks, the author showed that the proposed
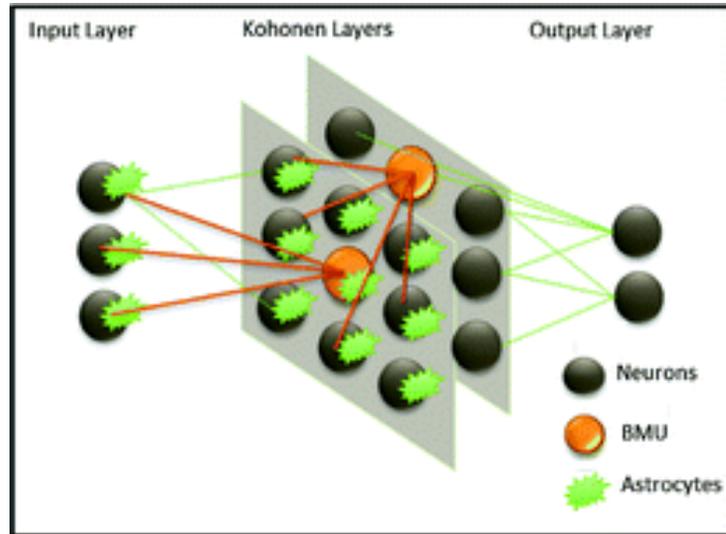
Figure 3.13: Architectural layout of proposed SONG-Net network. Instead of having fully-connected hidden layer, the SOM is used instead. Taken from Marzouki (2015).

model achieved faster convergence rate up to twelve times, while having lower MSE. However, the author did not present glia as individual functional units, but instead he used them only as an inspiration for the concept of neuronal regulation. No glial units (in the traditional sense of the term) were actually used in the model.

3.2 Modeling of synaptic plasticity modulation

The second role of artificial glia in ANNs is a synaptic plasticity modulation, firstly introduced by Porto-Pazos et al. (2011) and further extended in Alvarellos-González et al. (2012). The authors use more precise term *artificial astrocytes* instead of *artificial glia*, since glia represent the vast majority of non-neuronal cells in the nervous system with multiple functions, whereas only astrocytes are currently considered to play a vital role in information processing tasks. Their model was also based on MLP in which every single neuron is accompanied by a single astrocyte that monitors its activity and updates postsynaptic weights accordingly. The scheme of the network is illustrated in Fig. 3.14.

Every astrocyte registers the activity of the neuron, by applying function $u(x) : \mathbb{R} \to \mathbb{Z}$ over the output of the neuron as

$$u(x_j(t)) = \begin{cases} -1, & x_j(t) \leq 0 \\ 1, & x_j(t) > 0 \end{cases} \tag{3.10}$$
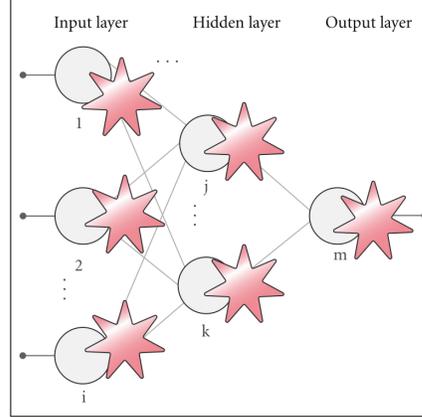
Figure 3.14: Architecture of neuro–glial network with artificial astrocytes. Astrocytes are paired with each neuron within all layers, controlling the synaptic plasticity of output synapses depending on the neural activity. Taken from Alvarellos-González et al. (2012).

while the history of past k activities is summed and stored in r_j as

$$r_j(t) = \sum_{i=0}^{k-1} u(x_j(t-i)) \tag{3.11}$$

The change in synaptic weights is then defined by equations eq. 3.12 to eq. 3.14:

$$z(t) = \begin{cases} a, & r_j(t) \equiv \mu \\ b, & r_j(t) \equiv -\mu \end{cases} \tag{3.12}$$

$$\Delta w_i(t) = |w_i(t)| z(t) \tag{3.13}$$

$$w_i(t + \Delta t) = w_i(t) + \Delta w_i(t) \tag{3.14}$$

where the hyperparameters a, b, and μ are chosen manually for each dataset. The authors proposed several rules of synaptic plasticity that we do not mention, but all of them are based upon this simple idea.

The authors evaluated the performance of such a trained model on four different problems (Fig. 3.15) and confirmed that the model was able to learn the problems better than the equivalent model without astrocytes in most cases, but its efficacy depended on the complexity of the given problem. It should be noted that an equivalent model without astrocytes (traditional MLP) was not trained by the gradient-descent method, but by the genetic algorithm.
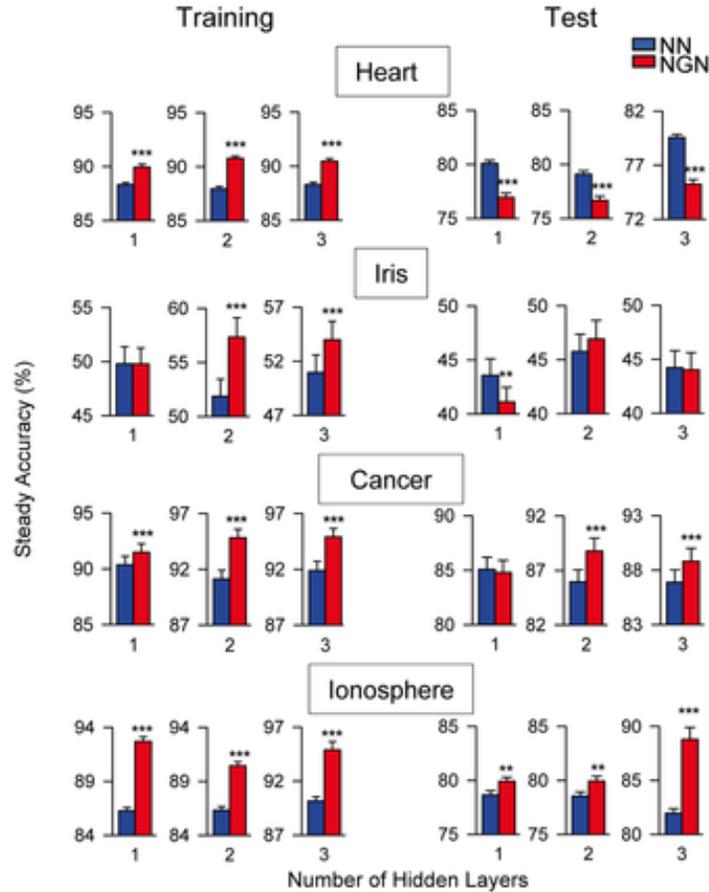
Figure 3.15: Performance comparison of traditional neural network (NN) with the proposed neuro-glial network (NGN). As can be seen, the NGN model worked better for most problems. The first problem, Heart, was to determine the presence of heart disease from 13 input parameters. The second problem, Iris, was the classification of plants into 3 different classes according to 4 input parameters. In the third problem, Cancer, the presence of the tumor from 9 parameters was determined. Finally, in the last problem, Ionosphere, the authors tried to determine the state of the ionosphere (good / bad) from the 34 input features. Taken from Porto-Pazos et al. (2011).

As mentioned above, the algorithms require the glial parameters a, b, and μ to be set manually for each individual problem and cannot be transferred between each other. As a consequence, before evaluating the performance of the model itself, an adequate hyperparameter subset must be found for each problem using some exhaustive search method such as grid search, which tends to be time-consuming and error-prone. Therefore Mesejo et al. (2015) designed and implemented a method based on evolutionary algorithms that automatically search for the desired optimal parameters.

Last but not least, Sajedinia (2014) extended the model by introducing random net-

works of artificial astrocytes. The model consists of several astrocyte networks, where each network involves a random subset of all astrocytes (Fig. 3.16). When being activated, astrocyte results in the activation of all other astrocytes within the same network where each of them will increase its value $r_j(t)$. Although the authors did not compare their model with model by Porto-Pazos et al. (2011), they claim that the network with random astrocyte networks outperforms traditional MLP without astrocyte units.
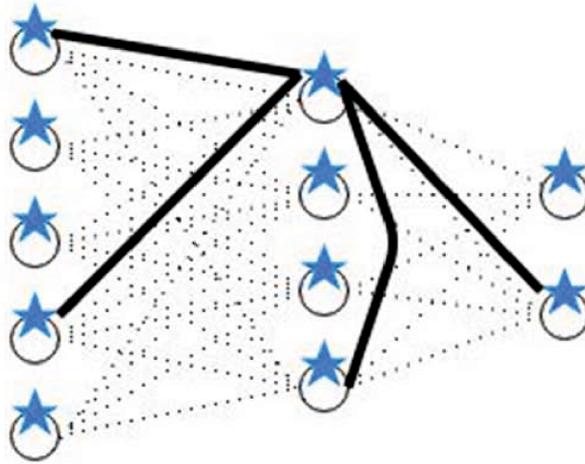
Figure 3.16: Architecture of the model with random astrocyte networks. Taken from Sajedinia (2014).

Chapter 4

Astrocytes in feedforward neural networks

Before we delve into the study of artificial astrocytes in feedforward neural networks, we start by providing a brief introduction into feedforward neural networks.

4.1 Feedforward neural networks

4.1.1 Single-layer perceptron

The very first model of an artificial neuron was introduced by McCulloch and Pitts (1943) mimicking the functionality of a biological neuron. It is the simplest model possible with boolean inputs (fire or no fire) and boolean weights (excitatory or inhibitory). Inputs are simply aggregated resulting into the final binary output. All weights are manually selected and remain hard-coded, therefore the model does not change with the time.

Major advance was the introduction of the **perceptron** by Rosenblatt (1958). The model is extended with real number inputs and weights, and training procedure to adapt the weights. Architecture of the model, as depicted in Fig. 4.1, consists of a single neuron whose parameters (input weights and threshold) are learned in a supervised manner with the teacher providing desired outputs. If the patterns belong to two linearly separable classes, Rosenblatt proved that his model converges and finds the decision boundary in finite number of steps. Computation of the final class is calculated as a linear combination of the input vector $\boldsymbol{x}$ with the corresponding weight vector $\boldsymbol{w}$ subtracted the threshold θ and passed to the signum activation function as shown in eq. 4.1.
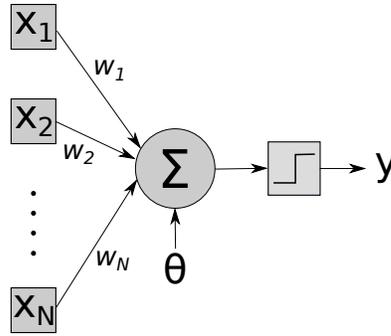
Figure 4.1: Architecture of Rosenblatt's perceptron. Input vector $\boldsymbol{x}$ is linearly combined with weights $\boldsymbol{w}$ and subtracted the threshold θ. The resulting sum is then passed to the signum activation function which returns 0 or 1 deciding the corresponding class y. Having other activation functions (i.e. continuous) than (discrete) signum is also possible.

$$y = \text{sgn}(\sum_i x_i w_i - \theta) \tag{4.1}$$

where the signum function is defined as

$$\text{sgn}(net) = \begin{cases} 1, & net \geq 0 \\ 0, & net < 0 \end{cases} \tag{4.2}$$

Training of the model consists of a few steps starting with (1) the initialization of parameters (the weights and the threshold) with random values, (2) selecting an input $\boldsymbol{x}$, (3) computing the output class y and finally (4) the adaptation of the weights $\boldsymbol{w}$ using error-correction learning rule given the desired class d

$$\Delta \boldsymbol{w} = \eta(d - y)\boldsymbol{x} \tag{4.3}$$

Hyperparameter η is the learning rate which is manually selected from a range $0 < \eta \leq 1$. Since the threshold θ can be considered as a special case of the synaptic weight with the input fixed to -1, the same learning rule can be applied to its adaption.

By combining several perceptrons into single model with shared inputs, but unique parameters, it is possible to train the model to distinguish more than two classes. In order to perform relatively well, all classes have to be linearly separable. Instead of having one output neuron, the model has several neurons in a single layer and is therefore called **Single-layer perceptron** (SLP).
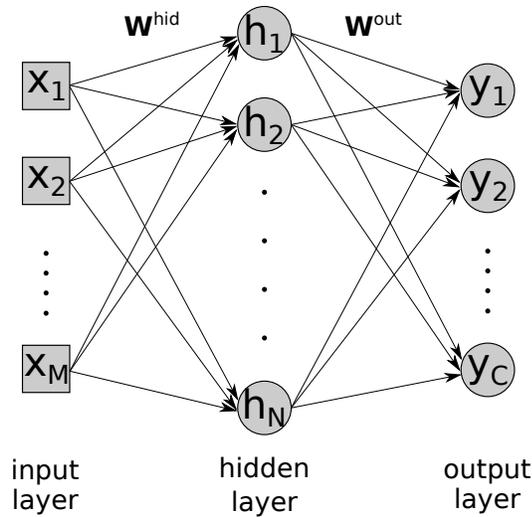
Figure 4.2: Architecture of MLP which distinguishes from SLP by addition of hidden layers and weight matrices. While this figure illustrates only one hidden layer, the model can have multiple hidden layers.

4.1.2 Multi-layer perceptron

Since SLP is suited for the classification of linearly separable classes only, extending the model with the **nonlinear activation function** and **hidden layers** allows to overcome this limitation and significantly broadens the domain applicability. Instead of having an input and output layers only, the new layer of intermediate neurons, called hidden layer, is provided which computes nonlinear transformation of input space into **feature space** which is hence used in the output layer computation. Architecture of the model is presented in Fig. 4.2.

Regarding the activation functions, several variants are considered such as logistic sigmoid, hyperbolic tangent, rectified linear unit. In our work we consider logistic sigmoid with the following definition:

$$f(net) = \frac{1}{1 + \exp{(-net)}} \tag{4.4}$$

It has been proven that MLP with a single hidden layer is capable of approximating any continuous function to any desired degree of accuracy (Hornik et al., 1989). The tradeoff for having superior representational power and efficiency is more complex and a time consuming training of the model, which includes, but is not limited to finding suitable input data encoding, selection of model parameters such as number of neurons, or type of activation functions. The most common approach for the training of MLPs are first-order

iterative gradient based methods that require the definition of loss function to measure model's error in approximating the desired output. Usually the sum of squared errors is considered with the following definition

$$E(t) = \frac{1}{2}\sum_{p=1}^{P}(\boldsymbol{d}^{(p)} - \boldsymbol{y}^{(p)}(t))^2 \tag{4.5}$$

Training per se consists of two phases. In **forward pass** the signal computation flows from the input layer through the hidden layers to the output layer and the decision of the output layer y_i is measured against the desired label d_i. Upon computing the error for all P patterns, the **backward pass** is performed in which the error signal is propagated backwards from the output layer towards the input layer and successive adjustments are made to the synaptic weights of the network. Regarding the learning equations, both weight matrices (input–hidden and hidden–output) are updated by moving in the direction of the negative gradient (hence the name gradient descent):

$$\boldsymbol{W}(t+1) = \boldsymbol{W}(t) - \eta\frac{\partial E(t)}{\partial \boldsymbol{W}(t)} \tag{4.6}$$

4.2 Artificial astrocytes in FFNN

4.2.1 Fixed-weights astrocytes in A-MLP

As stated in chapter 3, two major roles of astrocytes in ANNs are considered in general, either as neuronal regulators or as synaptic plasticity modulators. In this work we study solely the former function: the neuronal regulation. We start with a simplest model of a feedforward neural network and astrocyte per se, and we gradually move toward adding more complex mechanisms.

Inspired by Ikuta et al. (2010), we work with MLPs as described in the previous part with one hidden layer that is augmented with artificial astrocytes. Each neuron is paired with a single astrocyte and each astrocyte regulates only one neuron. The architecture of the proposed model is depicted in Fig. 4.3.

The output of i-th hidden neuron is given by the following formula

$$h_i(t+1) = f(\sum_{j=0}^{M} w_{ij}x_j(t) + \alpha\psi_i(t)) \tag{4.7}$$

with the logistic sigmoid as the activation function $f(net)$. The astrocyte activity is
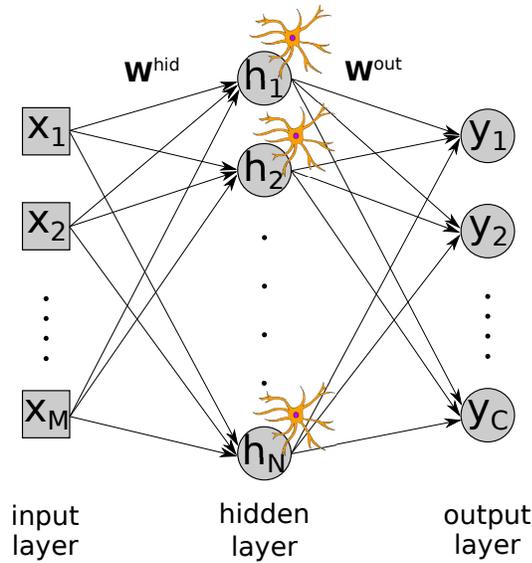
Figure 4.3: Architecture of A-MLP. The hidden layer of NN is augmented with astrocytes that listen to and regulate neurons.

modified according to

$$\psi_i(t) = \begin{cases} 1, & \text{if } \theta < h_i(t-1) \\ \gamma\psi_i(t-1), & \text{otherwise} \end{cases} \tag{4.8}$$

Each astrocyte contributes, with a weight α, to the activity of the hidden neuron (eq. 4.7). When the neuron output exceeds the given threshold θ, the astrocyte activation is set to 1 and then it starts to decay by a factor γ, where $0 < \gamma < 1$. We call this model in our work **A-MLP**.

Note that the model consists of three free hyperparameters (α, γ, θ) whose optimal values have to be found experimentally. Since each problem requires a different set of optimal parameters, finding them requires time-intensive computations. As we explain later, we try to solve these issues by replacing constant parameters with modifiable versions.

4.2.2 Dynamic weights in A-MLP(α)

The astrocytes regulate neurons in the hidden layer by the factor α being shared for all astrocytes, which is, nevertheless, not biologically plausible. Plethora of regulatory mechanisms are well described from the current research in biochemistry that includes neuronal excitation or inhibition by astrocytes (Fellin et al., 2006). For that reason we consider an individual weight α_i for each astrocyte which may be either excitatory or

inhibitory. The activation rule for the hidden unit then becomes

$$h_i(t+1) = f(\sum_{j=0}^{M} w_{ij}x_j(t) + \alpha_i\psi_i(t)) \tag{4.9}$$

As explained in Section 4.1.2, the neuronal synaptic weights are updated using the gradient descent method. Since the astrocytic weight in eq. 4.7 can be treated as any other weight, we can apply the same optimization method for its update. The goal is to minimize the loss function stated in eq. 4.5, by moving the astrocytic weights along the negative gradient, i.e. $\Delta\alpha_i = -\partial E(w)/\partial\alpha_i$. Since E is differentiable with respect to α_i, we can write using the chain rule:

$$\Delta\alpha_i = -\frac{\partial E}{\partial y}\frac{\partial y}{\partial net_y}\frac{\partial net_y}{\partial h_i}\frac{\partial h_i}{\partial net_{hi}}\frac{\partial net_{hi}}{\partial\alpha_i} \tag{4.10}$$

$$\Delta\alpha_i = -\overbrace{(d-y(x))y(x)(1-y(x))}^{\delta_y}w_{yh_i}h_i(1-h_i)\psi_i \tag{4.11}$$

$$\Delta\alpha_i = -\overbrace{\delta_y w_{yh_i}h_i(1-h_i)}^{\delta_i}\psi_i \tag{4.12}$$

which yields the final formula:

$$\Delta\alpha_i = -\delta_i\psi_i \tag{4.13}$$

Note that this derivation is valid for the model with a single output neuron which is the case for our experiments. However it is trivial to derive the rule for the model with multiple output neurons.

4.2.3 Dynamic threshold in A-MLP(θ)

In the case of other two free hyperparameters, γ and θ, since it is not straightforward to compute the derivation for the activation function written in eq. 4.8 with respect to γ and θ, we consider an alternative unsupervised rule.

Generally during training of NNs it happens quite often that some neurons get trapped in one of the two extremes, by becoming either silent or permanently active. The gradient update of such neuronal weights is then problematic, because either the gradient is close to zero and therefore no errors would propagate through a silent neuron leading to no

update of neuronal weight. On the other hand in permanently active neuron the weights might grow into large values, even infinite, leading to numerical problems, thus making the model unstable.

The same issue may happen in our model with artificial astrocytes when the threshold θ is set too low, making the astrocytes fire all the time. On the contrary, too high a value may prevent the neurons from exceeding the required threshold. This would however not advance into numerical problems, but the regulatory function of astrocytes would be lacking. Moreover, since each neuron in the neural network develops its own role in the classification task and for the same reason we explained in the previous section, we employ each astrocyte with a custom weight θ_i.

To incorporate unsupervised dynamic change of θs during training to accommodate for the change in neuronal behaviour, we propose an update rule with two variations. In order to stabilize the astrocytic regime, we can set the threshold θ either directly to the mean value $\langle . \rangle_t$ of an astrocyte unit (eq. 4.14) or only shift the threshold slightly closer to the mean value (eq. 4.15) using the learning speed η_θ. This forces the astrocyte to move only within its mean values avoiding the critical values of 0 and 1. With a higher θ it becomes harder for the neuron to overpass, thus the activity decays and vice versa. Hence, the update rules are

$$\theta_i(t+1) = \langle \psi_i(t) \rangle_t \tag{4.14}$$

and

$$\theta_i(t+1) = \theta_i(t) + \eta_\theta(\langle \psi_i(t) \rangle_t - \theta_i(t)) \tag{4.15}$$

where we introduced another hyperparameter, namely an averaging window of present length.

4.2.4 Dynamic activity decay in A-MLP(γ)

Hyperparameter γ can be updated based on the same principle as explained before. Now instead we update γ to achieve inverse correlation with the mean value of the astrocytic activity as (also two variations)

$$\gamma_i(t+1) = 1 - \langle \psi_i(t) \rangle_t \tag{4.16}$$

$$\gamma_i(t+1) = \gamma_i(t) + \eta_\gamma(1 - \langle \psi_i(t) \rangle_t - \gamma_i(t)) \tag{4.17}$$

Higher values of γ are achieved during a lower activity, thus a hypo-excited astrocyte

holds its activation value for a longer period. Contrarily, the lower γ triggers faster activity decay forcing the astrocyte to avoid excessive simulation.

4.2.5   Combination of previous models

The last two models are simple combinations of key ideas described in previous parts. Model A-MLP (γ, θ) combines dynamic thresholds, θs, and activity decays, γs. Model A-MLP(α, γ, θ) includes dynamic regulatory weights, αs, as well.

4.3   Experiments

To assess the performance of proposed networks with artificial astrocytes, we have chosen four classification tasks: 1) two spirals, 2) nested circles, 3) chessboard, and 4) N-parity. First three datasets (illustrated in Fig. 4.4) consist of two sets of two-dimensional points interleaved together with a high level of inseparability. This is considered difficult for a standard ANN due to a high number of potential local minima, which are generally rather problematic for gradient-based models.
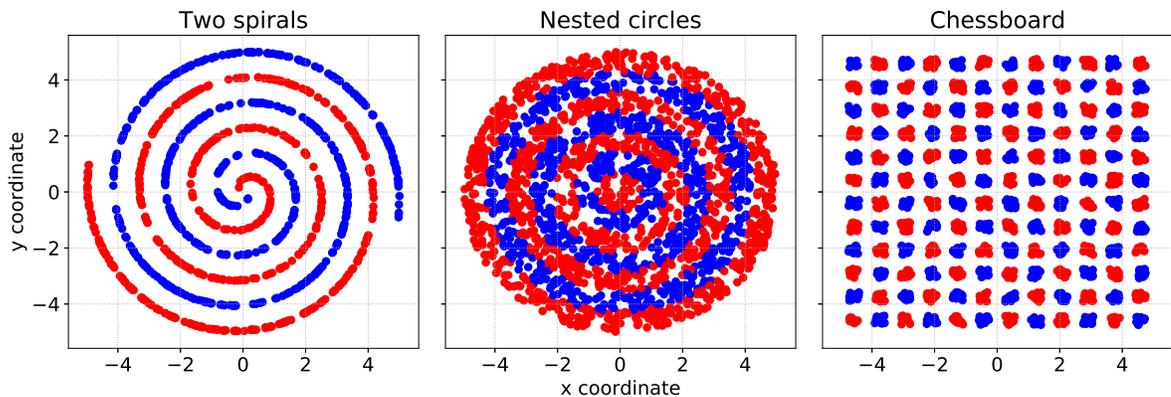
Figure 4.4: Three datasets from the same domain of the two dimensional plane that we use in evaluating the performance of our models. The two spirals are rotated three times, the circles are nested together seven times and the chessboard has size 12×12. The problem is, given a point (x, y), to decide whether it belongs to the first or the second class.

We compare all results with the traditional MLP without astrocyte units which is used as a baseline. To eliminate the possibility that astrocyte units act as a random noise generator which is well described as a mean of regularization to avoid overfitting

([Holmstrom and Koistinen](#), 1992; [Bishop](#), 1995), we include in our comparison the model **N-MLP**, which simply extends the hidden layer activation with noise from the uniform distribution $Uni(-1, 1)$ – the very same interval that astrocytes produce (after weight multiplication).

In the first place for each dataset we found optimal hyperparameters using a grid search for MLP (number of hidden neurons, learning rate, weight initialization and number of epochs) that yielded the lowest MSE. Then we transferred the very same set of hyperparameters to all A-MLP models and searched for remaining hyperparameter values (individually for each model). Each dataset was randomly split to train/test set in the ratio 80:20.

4.3.1 Two spirals

The optimal number of hidden neurons was selected as $N = 30$, since more units did not produce better results. The learning rate was $\eta = 0.1$ and all weight matrices were initialized from the uniform distribution $Uni(-0.1, 0.1)$. Regarding the models with astrocyte units, the outcome of the grid search is presented in [Fig. 4.5](#) where it can be seen that the best values tend to cluster around $\alpha = -0.1$, $\gamma = 0.5$, $\theta = 0.1$, which are the very same values we also used. The length of an averaging window in models with dynamic $\boldsymbol{\theta}$ and $\boldsymbol{\gamma}$ was set to 50.
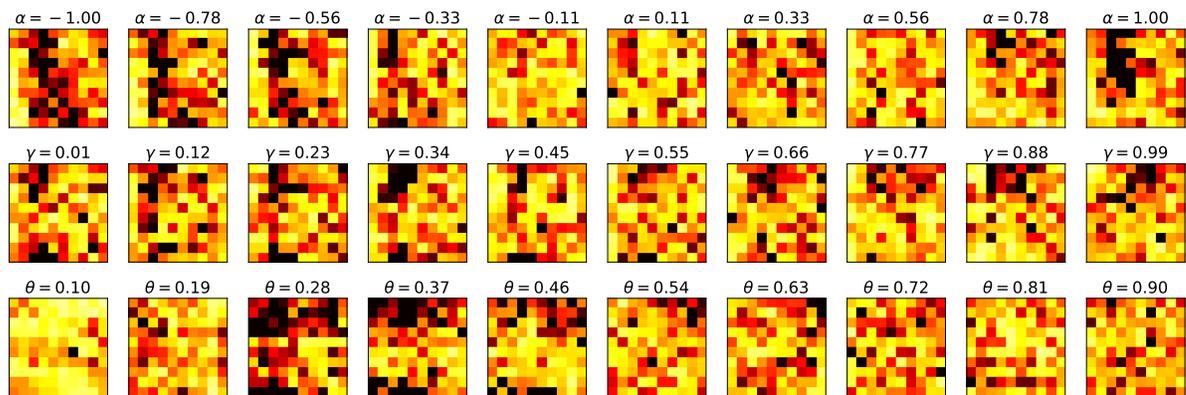
Figure 4.5: Grid search for optimal values of hyperparameters. Each heatmap uses a fixed single parameter (shown in the title) and displays all combinations for the other two parameters. Each cell in every heatmap is averaged over 5 simulations. Lighter color denotes better performance.

[Table 4.1](#) shows results of all models (training lasted for 5000 epochs), averaged over

100 runs. Learning curves can be seen in Fig. 4.6. The best model, A-MLP(θ), yields 50% lower error rate compared to the standard MLP.

Model	train set	test set
MLP	0.075 ± 0.067	0.094 ± 0.066
N-MLP	0.056 ± 0.046	0.072 ± 0.049
A-MLP	0.073 ± 0.067	0.088 ± 0.068
A-MLP(α)	0.050 ± 0.049	0.078 ± 0.050
A-MLP(θ)	0.035 ± 0.038	0.053 ± 0.042
A-MLP(γ)	0.072 ± 0.063	0.088 ± 0.061
A-MLP(γ, θ)	0.039 ± 0.048	0.055 ± 0.049
A-MLP(α, γ, θ)	0.060 ± 0.043	0.098 ± 0.044

Table 4.1: MSE (mean $\pm$ standard deviation) over 100 instances on the two spirals task trained for 5000 epochs. The best model, A-MLP(θ), yields more than 50% lower error rate compared to the MLP with statistical significance ($p < 0.001$).
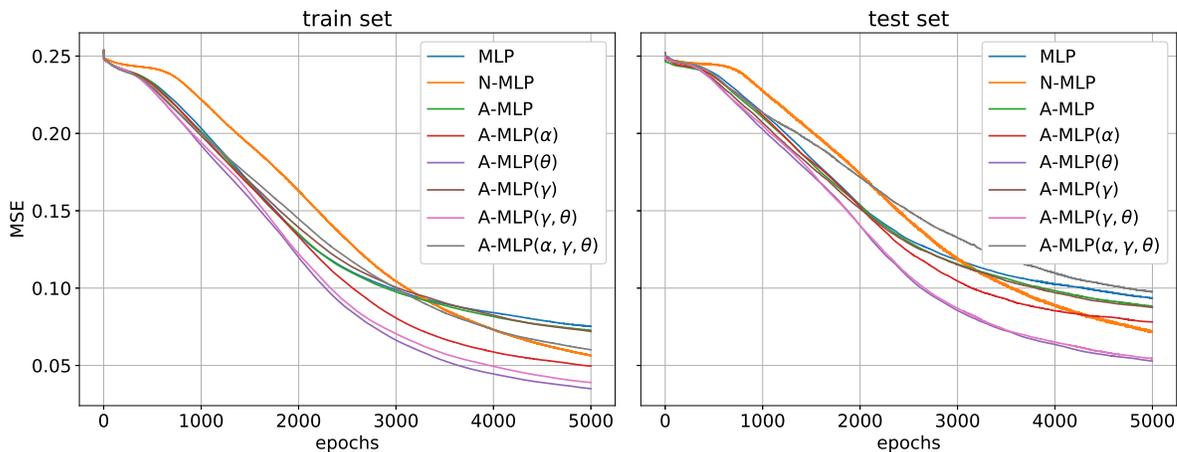
Figure 4.6: Learning curves for all models on the two spirals problem. Although the rate of convergence matches all other models, the final MSE is lower for each model with astrocyte units.

Next, in order to get insight into learned parameters, we displayed the distributions of final $\boldsymbol{\theta}$ at the end of training, accumulated over 100 simulations (see Fig. 4.7). It is clear that the values advance into positively skewed distribution with $\langle \boldsymbol{\theta} \rangle \sim 0.34$. Although the figure presents the best model, A-MLP(θ), similar distribution arises on all models with dynamic $\boldsymbol{\gamma}$ or $\boldsymbol{\theta}$.
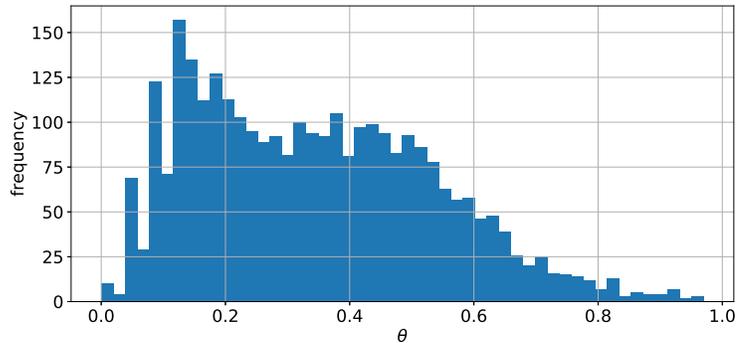
Figure 4.7: Positively skewed distribution of $\boldsymbol{\theta}$ values ($N = 30$) for the best performing model A-MLP(θ) at the end of training, accumulated over 100 simulations for the two spirals dataset. The length of average windows was set to 50, hence the slow update pace of the vector $\boldsymbol{\theta}$ that adapts to overall activity of astrocytes.

4.3.2 Nested circles

We found optimal number of hidden neurons $N = 50$, the learning rate $\eta = 0.1$ and the initialization of weight matrices from the uniform distribution $Uni(-1, 1)$. Using the same settings of grid search for the models with astrocyte units we obtained values $\alpha = -0.7$, $\gamma = 0.99$, $\theta = 0.6$ and the length of the averaging window $= 2$, which were shared among all models that required the specific hyperparameter. Table 4.2 shows results of all models, averaged over 100 instances. Learning curves can be seen in Fig. 4.8.

Model	train set	test set
MLP	0.119 ± 0.035	0.145 ± 0.033
N-MLP	0.123 ± 0.016	0.146 ± 0.018
A-MLP	0.112 ± 0.032	0.136 ± 0.030
A-MLP(α)	0.098 ± 0.020	0.136 ± 0.024
A-MLP(θ)	0.086 ± 0.020	0.116 ± 0.024
A-MLP(γ)	0.125 ± 0.033	0.151 ± 0.030
A-MLP(γ, θ)	0.091 ± 0.016	0.121 ± 0.018
A-MLP(α, γ, θ)	0.140 ± 0.015	0.239 ± 0.029

Table 4.2: MSE (mean $\pm$ standard deviation) over 100 instances on the nested circles problem trained for 5000 epochs. The best model, A-MLP(θ), is the same as in case of two spirals. However, the worst model, A-MLP(α, γ, θ), has significantly higher MSE (especially on the test set) compared to MLP which is quite a surprise given the fact that on the previous task it performed quite well.

Concerning the results, it can be seen that use of white noise in model N-MLP did not help at all and the performance was very similar to the standard MLP. All networks behave
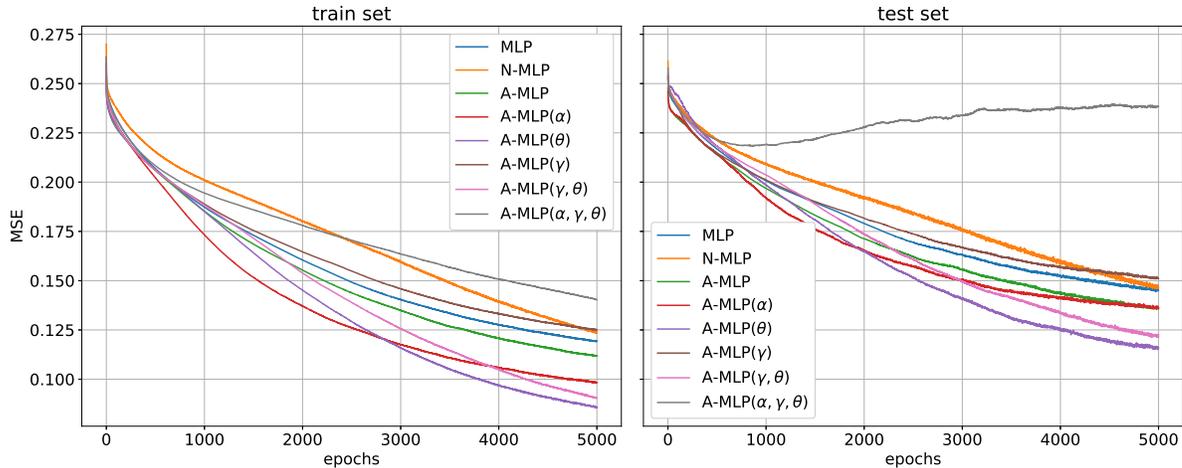
Figure 4.8: Learning curves for all models on the nested circles problem.

more or less the same as on previous task with the same winning model A-MLP(θ). For reasons unknown to us, the performance of A-MLP(α, γ, θ) appears to be overly atypical with the severe overfitting of the train set leading to poor generalization. Out of curiosity we plotted (shown in Fig. 4.9) output space for the models MLP, A-MLP(θ) and A-MLP(α, γ, θ) to grasp an idea on how the networks perform. At first glance it may appear convincing that the output space for the MLP matches original dataset probably the most, however the network was able to learn only six levels of nested circles. The best model, A-MLP(θ), managed to learn even the last level (the red circle in the middle) with the lower MSE and higher classification accuracy despite the fact that the final figure does not look that plausible. Regarding the last model, A-MLP(α, γ, θ), it can be clearly seen that the model did not generalize at all.

Last but not least, we look at the final distribution of $\boldsymbol{\theta}$ values as presented in Fig. 4.10. Since the optimal length of average window was found with the grid search to be only two, suggesting that the astrocytes "prefer" to chase immediate neural activity instead of focusing on the overall behaviour. It can be seen that $\boldsymbol{\theta}$ values tend to cluster around the maximum which is due to astrocytic activation of two consecutive times.

4.3.3 Chessboard

The hyperparameters values we used were $N = 400$ hidden neurons, weights initialization from the uniform distribution $Uni(-1, 1)$ and 5000 training epochs. Concerning the hyperparameters for astrocyte units, using the grid search we found: $\alpha = 0.42$, $\gamma = 0.29$,
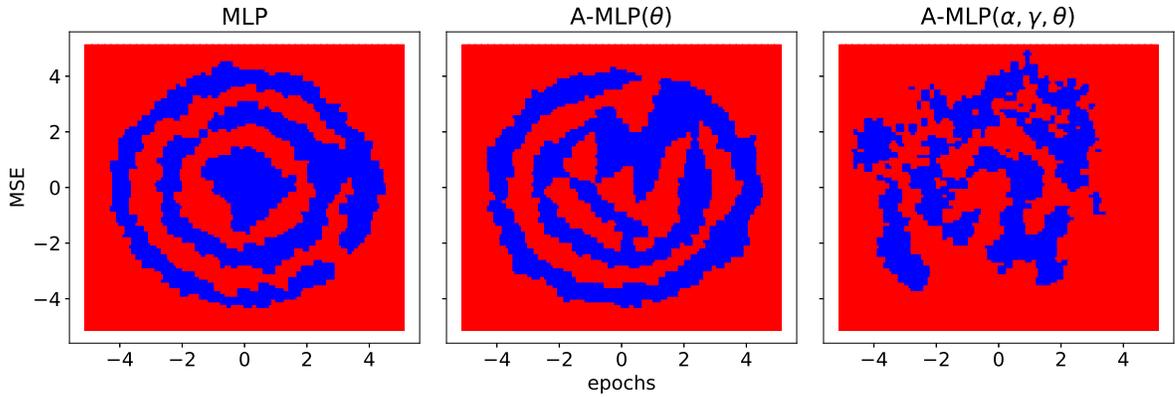
Figure 4.9: Comparison of models MLP, A-MLP(θ) (the best) and A-MLP(α, γ, θ) (the worst). The output space for MLP matches the original dataset very well, however it learned only six levels of nested circles. A-MLP(θ) managed to learn even the most inner circle and performed the best, although the separation of the classes appears imperfect. The last network A-MLP(α, γ, θ) did not generalize at all.
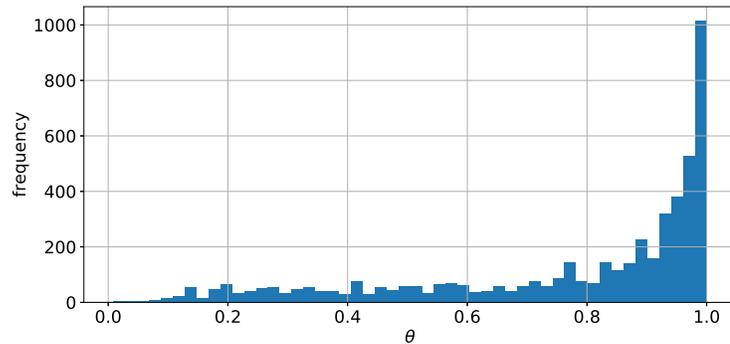
Figure 4.10: Distribution of $\boldsymbol{\theta}$ values for the best performing model A-MLP(θ) at the end of training, accumulated over 100 simulations for the nested circles dataset. The length of averaging windows was set to two implying the almost immediate change of the vector $\boldsymbol{\theta}$ based upon the mean of two last astrocytic activations.

$\theta = 0.44$ and average window length $= 10$. The results are presented in Table 4.3 and Fig. 4.11. This time, the network with the best performance was A-MLP(α, γ, θ) yielding almost perfect generalization on the testing set, however the error rate did not differ much compared with the other models with astrocyte units. On the other hand, MLP and N-MLP performed significantly worse.

Model	train set	test set
MLP	0.042 ± 0.070	0.049 ± 0.077
N-MLP	0.115 ± 0.094	0.124 ± 0.097
A-MLP	0.012 ± 0.066	0.015 ± 0.065
A-MLP(α)	0.005 ± 0.009	0.009 ± 0.015
A-MLP(θ)	0.007 ± 0.031	0.014 ± 0.041
A-MLP(γ)	0.005 ± 0.024	0.011 ± 0.026
A-MLP(γ, θ)	0.002 ± 0.009	0.006 ± 0.011
A-MLP(α, γ, θ)	0.001 ± 0.003	0.005 ± 0.007

Table 4.3: MSE (mean $\pm$ standard deviation) over 100 instances on the chessboard problem trained for 5000 epochs. To our surprise, the model A-MLP(α, γ, θ) achieved the lowest error rate on both training and testing sets. On the other hand, the worst performing model was N-MLP with rather low performance compared to other models.
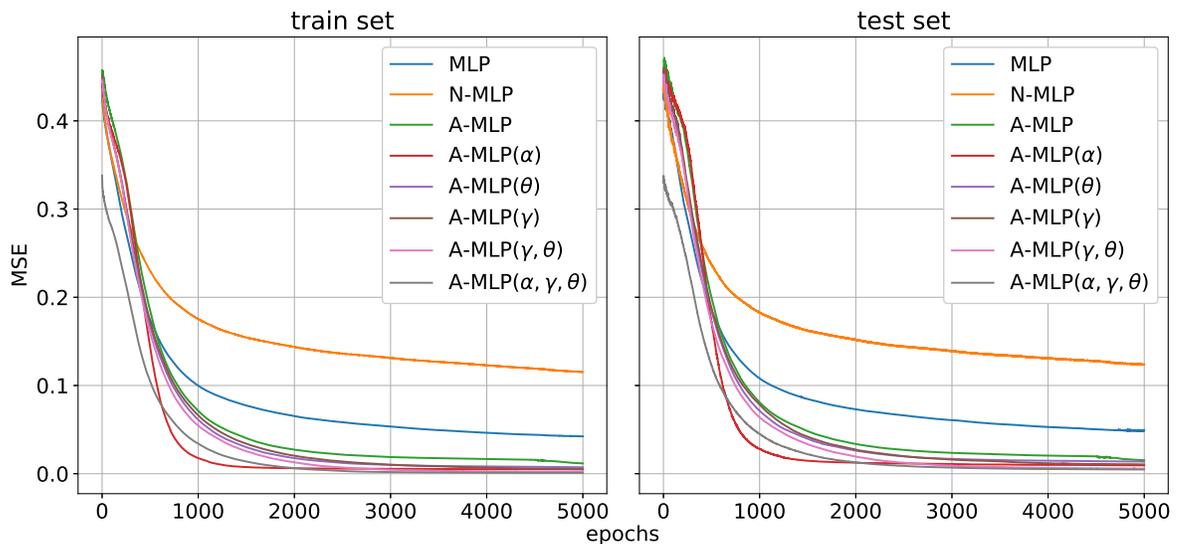
Figure 4.11: Learning curves for all models on the chessboard dataset.

4.3.4 N-parity problem

Our last dataset that we tested is quite different, yet belonging to a binary classification domain. The task is to determine whether a binary input vector has even or odd number of ones. More formally, an input vector has the form $\boldsymbol{x} = [x_1, \ldots, x_N], x_i = \{0, 1\}$ and the target $y = (1 + \sum_{i=1}^{N} x_i) \mod 2$. Since the problem is notoriously difficult to generalize to unseen patterns for machine learning algorithms, we train the models on full dataset (no train/test split) whose total size is 2^N.

Starting with an MLP, we chose the hidden layer with N neurons (a higher number did not yield better results), the output layer with a single neuron (0 = odd input vector,

1 = even input vector) and 10 000 training epochs. Proposed models with astrocyte units had the following values for fixed hyperparameters: $\alpha = -0.5, \gamma = 0.5, \theta = 0.5$ (previously found using the grid search). In Table 4.4 we present performance of all models and although we see models with astrocyte units lead on average to better performance, the differences are not statistically significant ($p > 0.1$).

Finally we looked at distribution of $\boldsymbol{\theta}$ at the end of the training (shown in Fig. 4.12). It can be seen that astrocytes develop various regimes depending on the problem complexity. With lower N it is possible to clearly detect N peaks, but with higher N the profiles gradually lose their multimodality, albeit remaining non uniformly distributed.

Model	4-parity	6-parity	8-parity
MLP	0.081 ± 0.060	0.065 ± 0.035	0.046 ± 0.070
N-MLP	0.104 ± 0.062	0.049 ± 0.031	0.021 ± 0.015
A-MLP	0.083 ± 0.086	0.059 ± 0.034	0.039 ± 0.023
A-MLP(α)	0.080 ± 0.065	0.072 ± 0.054	0.073 ± 0.069
A-MLP(γ)	0.087 ± 0.065	0.062 ± 0.034	0.042 ± 0.026
A-MLP(θ)	0.083 ± 0.075	0.065 ± 0.036	0.037 ± 0.021
A-MLP(γ, θ)	0.074 ± 0.051	0.063 ± 0.055	0.042 ± 0.027
A-MLP(α, γ, θ)	0.092 ± 0.072	0.078 ± 0.056	0.056 ± 0.028

Table 4.4: Mean squared error (MSE) $\pm$ standard deviation of 100 instances on three parity problems trained for 10 000 epochs. Models with astrocyte units yielded lower error rate although no statistical significance was found.
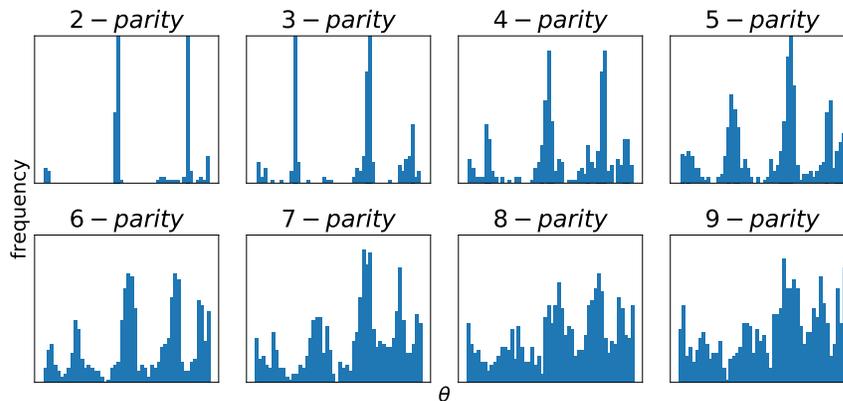
Figure 4.12: Distribution of $\boldsymbol{\theta}$ values (across 100 simulations) after being fully trained on a parity problem for the model A-MLP(θ). With lower N it is possible to detect N peaks assuming that each astrocyte handles a single bit from an input vector. On the other hand, with higher N, the peaks become less visible.

Chapter 5

Astrocytes in recurrent neural networks

Whereas in a feedforward neural network the information flows only in one direction, from the input layer directly to the output layer, in a recurrent neural network, the feedback connections are present allowing to pass and process the information typically in a loop. These feedback connections can take a variety of forms including feedback from the hidden layer to the input layer, feedback from the outer layer to the hidden layer, or alternatively feedback from the output layer directly to the input layer. A combination of multiple feedback types is also feasible.

RNNs are generally built upon (as an extension of) MLPs in terms of the architectural layout. However incorporation of feedback connections allows them to process inputs and outputs of a variable length which is a major limitation of MLPs. The nature of RNNs implies the existence of their intrinsic **state space** (hidden unit representations) emerging from the fact that the computation of neural activation for the hidden layer depends not only on an input layer, but also on the activation of the hidden layer in the previous time steps allowing the model to retain and change the state in the hidden units. This makes RNNs suitable for modeling temporal dynamic behavior, such as sequence recognition, generation or temporal association, prediction.

Although architectural layout of RNNs takes many different forms, here we present only the layout related to our study: the **simple recurrent network**.

5.1 Simple recurrent network

The adjective *simple* in the name of SRN connotes the fact that computation of the hidden layer activity and the error derivative "simply" depends on the activity and the error derivative attained from the unit-time delay. This simplification, however, does not imply that the model does not store information from the distant past. The model was originally conceived and used by Elman (1990) and the term **Elman's network** is often used interchangeably.

An SRN typically consists of three standard neural layers (input, hidden, output) where the hidden layer is copied and used as an input together with the input pattern to the computation of hidden layer activation in the next time step (Fig. 5.1). The copy of hidden layer outputs is retained in the **context layer** which is fully connected with the hidden layer using the feedback connections.
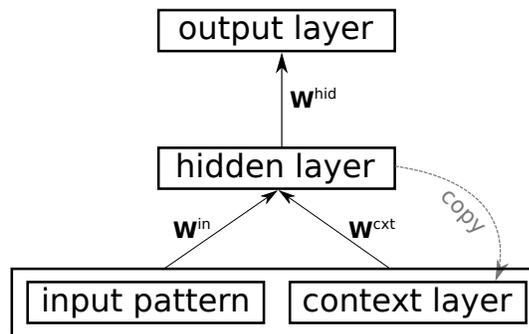
Figure 5.1: SRN architecture consisting of input, hidden and output layers. Input pattern is extended with the context layer where the output of hidden layer from previous time step is stored. The backward arrow, from the hidden layer to the context layer, denotes a copy operation.

Training of SRNs is similar to traditional training of FFNNs using gradient descent methods. There is, however, significant difference related to the error propagation. Whereas in FFNNs the errors are propagated layer by layer until they reach the input layer, in RNNs it is important to take into account the feedback connections. Use of the BP algorithm directly results in taking into account only the last input signal from the input pattern, leading to a rather short memory of the network. However, unfolding the hidden layer in time a number of steps back through time, therefore changing the RNN into FFNN, allows the BP algorithm to be applied for calculating and accumulating errors across each input signal resulting in greater memory of the network. Afterwards, the update of the weights is the same as in FFNN, using eq. 4.6. The algorithm described here (unfolding the model

and using traditional BP) is called **Back-propagation through time** (Werbos, 1988).

5.2 Echo state networks

Training traditional RNNs is considered to be difficult because of limitations of gradient descent methods which tend to be computationally expensive, to have slow convergence and to generally lead to poor local minima. Hence, the full adaption of all network weights is often omitted, yet still yielding excellent performance. This approach serves as a foundation for ESNs which were introduced by Jaeger (2001) for nonlinear system identification and time series modeling. ESNs are characterized by having randomly generated input weights and reservoirs with the training only the readout weights.

However, in order to work well, ESNs require delicate tuning of several hyperparameters including the reservoir, the spectral radius ρ, and input weight scaling τ. ESN must have the **Echo State Property** which says that regardless of initial conditions, the hidden layer must converge to the same state given the same input signal. If this is met, only readout weights adaption is sufficient to obtain the ESN with high performance.

The model has indeed the same architectural layout as SRNs (Fig. 5.2) and is usually equipped with a large amount of sparse feedback connections that remain untrained. The hidden neurons are driven by an input signal and are widely referred as the **reservoir** (in the context of ESNs).
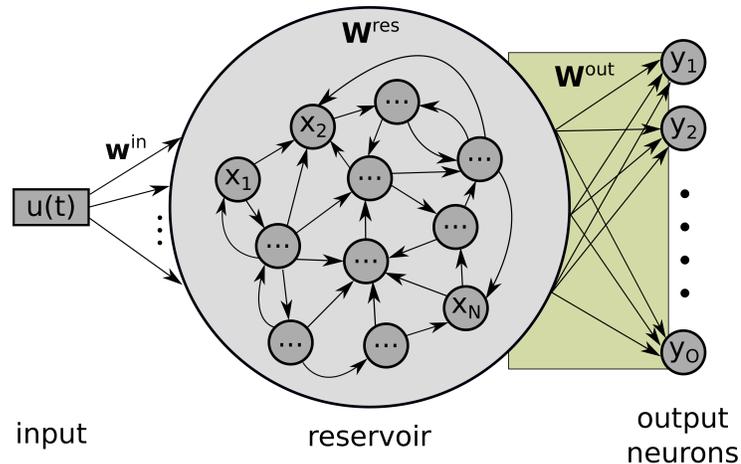
Figure 5.2: Architecture of ESN containing single input, N hidden units and O outputs units. Only the readout ($\boldsymbol{W}^{\text{out}}$) connections are adapted during training. Feedback connections from output layer to the reservoir are also possible, however they are not shown here.

Reservoir activation vectors $\boldsymbol{x}(t) = [x_1(t), ..., x_N(t)]$ and output activations $\boldsymbol{y} = [y_1, ..., y_C]$

for given input pattern $\boldsymbol{u} = [u(1), ..., u(T)]$ are updated according to ESN dynamics given by the formulas

$$\boldsymbol{x}(t) = \mathrm{f}^{\mathrm{res}}(\boldsymbol{w}^{\mathrm{in}}u(t) + \boldsymbol{W}^{\mathrm{res}}\boldsymbol{x}(t-1)) \tag{5.1}$$

$$\boldsymbol{y}(t) = \mathrm{f}^{\mathrm{out}}(\boldsymbol{W}^{\mathrm{out}}\boldsymbol{x}(T)) \tag{5.2}$$

where $\mathrm{f}^{\mathrm{res}}$, $\mathrm{f}^{\mathrm{out}}$ are suitable activation functions, $\boldsymbol{w}^{\mathrm{in}}$ is the input weight vector, $\boldsymbol{W}^{\mathrm{res}}$ and $\boldsymbol{W}^{\mathrm{out}}$ are recurrent and output weight matrices, respectively. In our study we use $\mathrm{f}^{\mathrm{res}}(\mathrm{net}) = 1/(1 + \exp(-net))$ and $\mathrm{f}^{\mathrm{out}} = \boldsymbol{id}$.

5.3 Fixed-weights astrocytes in A-ESN

Here we propose a model of ESN augmented with the same model of astrocytes as described in Section 4.2.1. Similarly, we omit the concept of glial syncytium in which astrocytes are connected using gap junctions and communicate sharing slow Ca^{2+} signals (as opposed to neuronal firing), but we start studying the simplest possible model instead. We consider merely the role of neuronal regulation by astrocytes themselves and equip each reservoir neuron with one astrocyte as shown in Fig. 5.3. We call this model **A-ESN**.
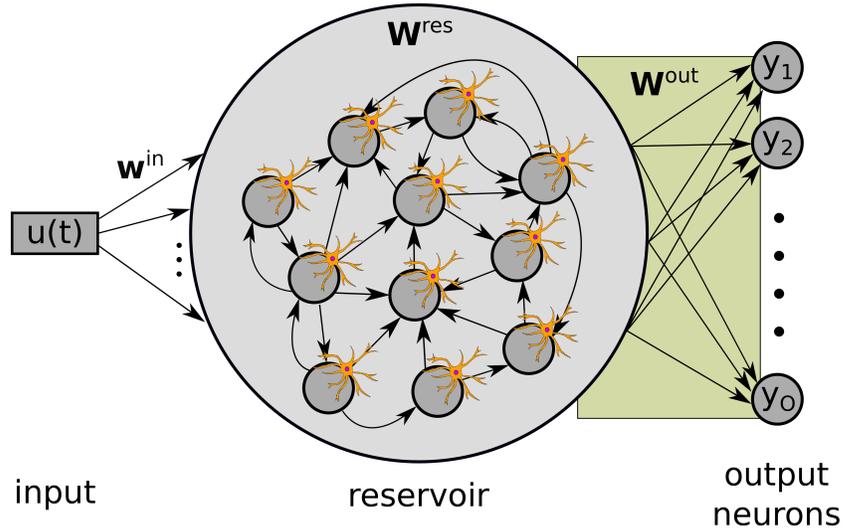
Figure 5.3: The architecture of the proposed model, A-ESN, with a reservoir of neurons and astrocytes. Each neuron is paired with an astrocyte that listens to it and regulates neuron's behaviour based on its past.

Reservoir activation $x_i'(t)$ takes into account input pattern $u(t)$, previous time step

activation vector $\boldsymbol{x}'(t-1)$ and astrocyte activation $\psi_i(t)$ weighted by a single shared weight w^α, which is expressed in the vector form as

$$\boldsymbol{x}'(t) = f(\boldsymbol{w}^{\text{in}}u(t) + \boldsymbol{W}^{\text{res}}\boldsymbol{x}'(t-1) + w^\alpha\boldsymbol{\psi}(t)) \qquad (5.3)$$

Astrocytes $\psi_i(t)$ listen to their associated neurons and when some of the neurons exceed the threshold θ, astrocytes produce the activation value of 1. The rest of them decay by factor γ as

$$\psi_i(t) = \begin{cases} 1, & \text{if } \theta < x'_i(t-1) \\ \gamma\psi_i(t-1), & \text{otherwise} \end{cases} \qquad (5.4)$$

This ESN dynamics is graphically depicted in Fig. 5.4.
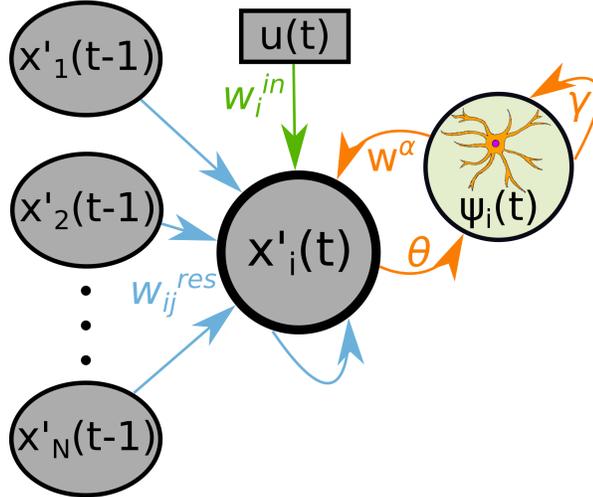
Figure 5.4: Neuron–astrocyte coupling. The astrocyte, weighted by w^α, regulates the associated neuron by contributing to its input. When the neuron surpasses the threshold θ, the astrocyte outputs 1 and slowly decays by factor γ in the next time steps. Blue arrows depict the reservoir weights, the green arrow an input weight and orange arrows the astrocyte parameters.

5.4 Hebbian-weights astrocytes in A-HL-ESN

Since using a single shared weight w^α for all astrocytes may be too constraining, we consider an individual weight for each astrocyte. Although astrocytes are not considered to be able to trigger neuronal action potential, they still modulate their membrane potential by the release of gliotransmitters including glutamate (exciting the neuron) or ATP (in-

hibiting the neuron) (Fellin et al., 2006). For that reason we consider randomly generated weights from a uniform distribution $Uni(-1, 1)$.

The exact relationship of neuronal regulation by astrocytes is still not well understood and we can only guess to which extent is this process plastic and what are the specific mechanisms of plasticity. For that matter we speculate using Hebbian learning which is in great detail described in Hebb (1949). The basic principle is that the change of a synaptic weight w_{ji} between neurons x_i and y_j, with the learning rate η, is expressed as

$$\Delta w_{ji}(t) = \eta x_i(t) y_j(t) \tag{5.5}$$

In our case we apply this rule for the change of the weight w^α between a neuron x_i' and an astrocyte ψ_i. Repeated application, however, may lead to an exponential change of the weight which is not biologically plausible and this is solved by incorporating some form of stabilization. This is in many cases the normalization of the final weights. We consider Oja (1982)'s rule which introduces a nonlinear, forgetting factor for the weight change

$$\Delta w_i^\alpha(t+1) = \eta x_i'(t)[\psi_i(t) - x_i'(t) w_i^\alpha(t)] \tag{5.6}$$

To take into account this new dynamics, we split our training algorithm into two phases: 1) once the unsupervised learning of the weights $\boldsymbol{w}^\alpha$ (eq. 5.6) in the reservoir is complete, 2) a supervised learning algorithm is applied to the readout weights. Instead of using eq. 5.3 for the reservoir update, we consider

$$\boldsymbol{x}'(t) = f(\boldsymbol{w}^{\text{in}} u(t) + \boldsymbol{W}^{\text{res}} \boldsymbol{x}'(t-1) + \boldsymbol{w}^\alpha * \boldsymbol{\psi}(t)) \tag{5.7}$$

with operator '$*$' denoting the element-wise product of vectors. We call the model with Hebbian learning as **A-HL-ESN**.

5.5 Experiments

For both classification tasks and memory capacity we consider the following training procedure:

1. Generate random input weights $\boldsymbol{w}^{\text{in}}$ and reservoir weights $\boldsymbol{W}^{\text{res}}$ scaled by $\rho/|\lambda_{\max}|$, where $\lambda_{\max}$ denotes the largest absolute eigenvalue of $\boldsymbol{W}^{\text{res}}$ and ρ is manually selected.

2. Run ESN using the training inputs $\boldsymbol{u}_{train}$ and collect the required reservoir activation state $\boldsymbol{x}(t)$ (more precisely explained in the specific section for each task).

3. Compute the linear readout weights using formula

$$\boldsymbol{W}^{\text{out}} = \boldsymbol{Y}^{\text{tgt}} \boldsymbol{X}^+ \tag{5.8}$$

where $\boldsymbol{Y}^{\text{tgt}}$ is a matrix of concatenated target vectors (in columns) and $\boldsymbol{X}^+$ is the pseudoinverse matrix of concatenated reservoir activation states from step 2.

4. Use the trained network on new input data $\boldsymbol{u}_{test}$ and evaluate the performance.

5.5.1 Classification experiments

For the classification tasks we have decided for the UCR Time Series Classification Archive (Chen et al., 2015) which consists of 85 real world problems and is often used for benchmarking of Machine Learning models. In our study we selected 8 random datasets upon which we assess the performance of the proposed methods. We use a standard ESN (without astrocytes) as a baseline and compare it with models A-ESN and A-HL-ESN.

For the training of readout weights we use only the last activation state vector $\boldsymbol{x}(T)$ obtained by processing each training pattern $\boldsymbol{u}_{train} = [u(1), ..., u(T)]$. Target vectors are represented using one-hot encoding and have shape $\boldsymbol{y}^{\text{tgt}} = [y(1), ..., y(C)]$ where C is the number of output classes. After training the model, the class of a new input $\boldsymbol{u}_{test}$ is decided by selecting output neuron with maximum activation

$$\text{class}(\boldsymbol{u}_{test}) = \arg \max_k y_k \tag{5.9}$$

Using the grid search we systematically investigated each hyperparameter (averaged over five instances) and selected the values with the lowest error rate on the testing dataset. Regarding the ranges for each hyperparameter we chose the values presented in Table 5.1.

The UCR archive already provides train/test split of the datasets, but we found this rather problematic because of the high risk of overfitting the hyperparameters to a particular test dataset. In order to avoid this, we merged both train and test datasets into a single set and used 5-fold cross-validation instead. To eliminate the random fluctuation in performance, we executed training procedures with random weights, random permutations of datasets and averaged error rates over 100 instances.

Parameter	Tested values
N	20 to 500 with step=20
τ	10, 5, 1, 0.5, 0.1, 0.05, 0.01, 0.001, 0.0001
ρ	0.8 to 1.4 with step=0.05
w^α	-1.0 to 1.0 with step=0.1
γ	0.0 to 1.0 with step=0.1
θ	0.0 to 1.0 with step=0.1

Table 5.1: Hyperparameter value ranges used in the grid search for each dataset.

Allowing for possibility of imbalanced datasets in which one class is over-represented with the respect to the others, we use *Matthews correlation coefficient* (MCC) (Matthews, 1975) as a metrics for performance evaluation score rather than the mean-squared error, accuracy or F1-score which does not work well on imbalanced datasets. The value MCC = 1 corresponds to a perfect match between model predictions and observations, whereas -1 indicates total disagreement between the two.

In all experiments, we used hyperparameters summarized in Table 5.2 resulting in the largest MCC on testing datasets.

Dataset	ESN			A-ESN			A-HL-ESN	
	N	ρ	τ	w^α	γ	θ	γ	θ
Earthquakes	20	1.40	0.01	-0.6	0.9	0.7	0.7	0.1
FaceFour	20	0.95	0.05	-0.4	0.1	0.2	0.2	0.2
MoteStrain	120	1.30	0.001	-0.3	0.9	0.5	0.3	0.3
OSULeaf	60	0.95	0.01	0.6	0.6	0.8	0.2	0.1
PhalOutlCorr	80	0.95	5.0	1.0	1.0	0.8	1.0	0.3
ProxPhalOutlCorr	40	0.90	1.0	-0.4	0.9	0.7	0.0	0.1
SwedishLeaf	160	1.40	0.001	-0.4	0.2	0.9	0.2	0.2
ToeSegmentation1	60	1.30	1.0	-0.5	1.0	0.9	0.3	0.1

Table 5.2: Optimal hyperparameters selected using the grid search for each dataset. Non-astrocytic hyperparameters (N, ρ, τ) were shared in all models on a given dataset.

Results in terms of MCC averaged over 100 simulations are presented in Table 5.3. It is clear that model with Hebbian connections, A-HL-ESN, significantly outperforms models ESN and A-ESN. Despite having more complex training procedure and thus higher time complexity, gain in terms of performance is clearly notable. Model with fixed connections, A-ESN, have yielded results equivalent to standard ESN (assuming correct settings of

hyperparameters), although it is speculative why on the last dataset (ToeSegmentation1), the error rate is significantly better (MCC of 0.5 ± 0.1 vs 0.32 ± 0.11).

Dataset	ESN	A-ESN	A-HL-ESN
Earthquakes	0.20 ± 0.12	0.21 ± 0.12	0.24 ± 0.11
FaceFour	0.44 ± 0.12	0.43 ± 0.13	0.56 ± 0.14
MoteStrain	0.65 ± 0.04	0.67 ± 0.06	0.85 ± 0.03
OSULeaf	0.41 ± 0.06	0.42 ± 0.06	0.57 ± 0.06
PhalOutlCorr	0.37 ± 0.04	0.38 ± 0.04	0.43 ± 0.03
ProxPhalOutlCorr	0.48 ± 0.06	0.52 ± 0.07	0.53 ± 0.06
SwedishLeaf	0.64 ± 0.03	0.63 ± 0.03	0.84 ± 0.03
ToeSegmentation1	0.32 ± 0.11	0.50 ± 0.10	0.59 ± 0.11

Table 5.3: MCC (mean±standard deviation) averaged over 100 simulations on each dataset. In each case, the model A-HL-ESN is superior regarding the performance.

In order to better understand the role of astrocytes with Hebbian connections, we were interested to know how the astrocyte weights develop during learning. For the fully trained models (all 100 instances), we plotted final distributions of the weights w_i^α as depicted in Fig. 5.5. We can observe that the weight distributions are skewed in the interval (1,2), roughly independent of the dataset, with an exception being MoteStrain, where some of the weights are also between 0 and 1. This can be explained by the use of Oja's rule (eq. 5.6) which shifts the weights towards the interval (0,1) when the neural activity is higher than astrocytic activity and towards the interval (1,2) when the opposite happens. Since we use unipolar activation functions for both neurons and astrocytes with an output (0,1), the weights will always be forced to stay within positive region. We may conclude this implies excitatory nature of the astrocytes in terms of neural regulation and that astrocytic activity remains higher the activity of neurons.

5.5.2 Memory capacity experiments

Memory capacity (MC) is defined by Jaeger (2002) as a measure of network's ability to reconstruct the past information from the reservoir on the network output by computing correlations. In our work (Farkaš et al., 2016; Farkaš and Gergel', 2017) we systematically investigated the effect of proper reservoir initialization on MC and proposed two gradient descent iterative methods that approach a maximum of theoretical limit of ESN's MC that drive the reservoir dynamics towards the critical regime (the transition zone between
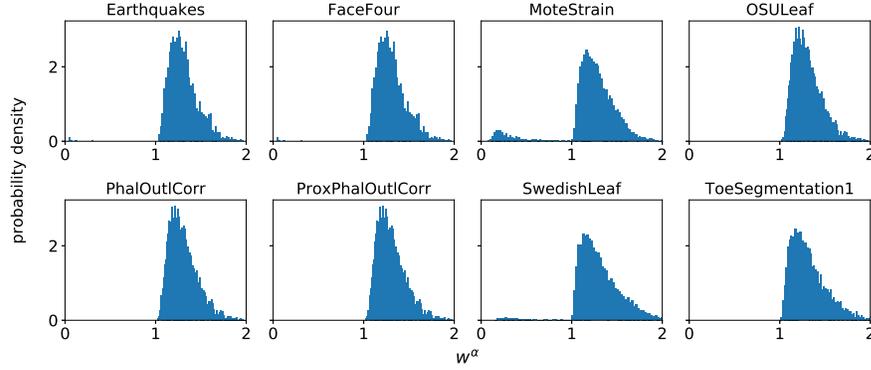
Figure 5.5: Distribution of the weights w_i^α in the trained models A-HL-ESN reveals excitatory role of the astrocytes.

a stable and an unstable dynamics regime).

Jaeger defined the MC as

$$\text{MC} = \sum_{k=1}^{k_{\max}} \text{MC}_k = \sum_{k=1}^{k_{\max}} \frac{cov^2(u(t-k), y_k(t))}{var(u(t)) \cdot var(y_k(t))} \tag{5.10}$$

where functions *cov* and *var* denote covariance and variance, respectively. $u(t-k)$ is the input presented k-steps before the current input, $k_{max} = \infty$, and $y_k(t)$ is the reconstruction at the network output. The computation of MC is approximated using $k_{max} = O$ output neurons. The computation of MC takes into account the network ability to retrieve the past input signal (for various delays k) from the reservoir using the linear combinations of reservoir unit activations observed at the output (quantified by MC_k). Jaeger proved that the MC for recalling an indepdendent, identically distributed (i.i.d) input by an ESN of N-units with identity activation function is bounded by N.

For assessing the performance of all models, we measure the total MC using randomly generated sequences $\boldsymbol{u} \in \mathbb{R}^T$ of length $T = 1100$, drawn from a uniform distribution, hence $u_i \in Uni(-1, 1)$. Such sequence has no underlying structure and is random: u_p and u_r are independent for $p \neq r$. We fed the first 100 inputs to the network to remove the initial transient which is normally not present once the network has "warmed up" to the task. Next 500 inputs to the network are utilized for the training, while collecting the reservoir activations $\boldsymbol{x}$ and target output vectors $\boldsymbol{y}^{\text{tgt}}$ and storing them into the matrices $\boldsymbol{X}$ and $\boldsymbol{Y}^{\text{tgt}}$, respectively. The optimal output weight matrix $\boldsymbol{W}^{out}$ is computed analytically according to eq. 5.8 and final 500 input patterns are used for assessing the MC.

Regarding the initialization, we use $N = 100$ neurons within a reservoir, initialize the

input weights $\boldsymbol{w}^{in}$ from uniform distribution $Uni(-\tau, \tau)$ ($\tau = 1e{-}6$), and the reservoir weight matrix $\boldsymbol{W}^{res}$ from normal distribution $\mathcal{N}(0, 1)$ scaled by $\rho/|\lambda_{\max}|$ ($\rho = 5.5$) as stated in step 1 of Section 5.5.

Similarly as in case of the classification experiments, we used a grid search for obtaining the best hyperparameters for models with astrocyte units. The values $\gamma = 1.0$ and $\theta = 0.6$ turned up to be optimal in both models. Regarding the A-ESN, we used $\alpha = -1.9$, implying extensive inhibitory role of *some* astrocytes as a result of relatively high threshold θ, which activates only few astrocytes. Upon being activated, their activity remains constant for the rest of input sequence.

A-HL-ESN, on the other hand, did not perform very well. By taking into account the fact that the weight vector $\boldsymbol{w}^\alpha$ remains within positive range, which contradicts weight value for the previous model, out of curiosity we simply swapped the term $+\boldsymbol{w}^\alpha * \boldsymbol{\psi}(t)$ with the $-\boldsymbol{w}^\alpha * \boldsymbol{\psi}(t)$ in eq. 5.7, which turned out to perform better, however, it did not exceed MC of the simpler model, A-ESN. We call these models with excitatory and inhibitory contribution as A-HL-ESN$^+$ and A-HL-ESN$^-$. The results are summarized in Table 5.4.

	ESN	A-ESN	A-HL-ESN$^+$	A-HL-ESN$^-$
Memory capacity	32.14 ± 3.13	40.03 ± 2.24	25.59 ± 2.32	39.94 ± 4.86

Table 5.4: Measure of MC on randomly generated sequences from uniform distribution on three models. The higher value signifies better performance, hence the best performing model is A-ESN with the static astrocytic weights.

Conclusion

The neuroscientific research for the last decades has highlighted the importance of glial cells in information processing context. Astrocytes regulate neuronal functionality in a variety of ways, particularly by maintaining the concentration of ions and neurotransmitters, by releasing gliotransmitters, and modulating both neuronal excitability and synaptic plasticity. However, limited amount of research has been done in the field of ANNs equipped with artificial astrocytes.

Inspired by Ikuta et al. (2010) and the subsequent work, as well as by recent findings from biological research of astrocyte physiology and their interactions with surrounding neurons, in our thesis we have proposed artificial astrocyte units to be integrated in feed-forward and recurrent neural networks. The role of astrocytes in both models is reduced to the regulation of neuronal excitability. The interaction is bidirectional and by listening to neural activity, astrocytes provide positive or negative feedback helping the neurons to stabilize.

In case of the FFNN, the original model with astrocyte units consists of several hyperparameters including glial weight, threshold, attenuation factor, propagating range of astrocyte activation, refractory period and an activation decay that needs to be selected manually for each problem, which is a time-consuming and error-prone process. Since we found the model to be too complex to start with and it turned out to be challenging to obtain the very same results as Ikuta et al. did in their paper, we have simplified the model by omitting the concept of syncytium and kept astrocytes as individual units, not connected with each other within a single network, which in the end turned out to perform significantly better.

Instead of using a single constant glial weight for all astrocytes, we have proposed a gradient-descent method that updates the parameters along the negative gradient of the loss function for each astrocyte individually. For the threshold and the activation decay, we have introduced two unsupervised rules (eq. 4.14 and 4.15 for the threshold and eq. 4.16 and 4.17 for the decay) which sets the specific value according to the history of astrocytic

activity. Since both rules turned out to perform practically the same, we used the first variation that updates the value directly to the averaging window of the astrocytic activity.

We have evaluated the performance of the proposed modifications on four classification problems: 1) two-spirals, 2) nested circles, 3) chessboard and 4) N-parity. For all problems we first selected an MLP with optimal hyperparameters found using an exhaustive grid search (the number of hidden neurons, the learning rate, initial weight distribution) and then used them in models with astrocyte units. The results obtained for N-parity did not outperform MLP, however all models already converged to the global minimum with zero classification error. In case of the first three tasks, all our models performed better in terms of the lower errors with statistical significance ($p < 0.001$). In nested circles problem, the simulations turned up to be relatively time-consuming, thus we trained all models only for 5000 epochs which was not enough (Fig. 4.8), albeit the models with astrocyte units had better convergence rate. Also, it is a mystery to us why the model A-MLP(α, γ, θ) was not able to generalize at all, since in other tasks it performed relatively well and in case of the chessboard problem it was even the best.

Inspired by positive results from the FFNNs, we transferred the same model of astrocytes to RNNs and explored their influence. Such models are closer to biological realism than FFNNs, because recurrent connections are critical and ubiquitous in the CNS. Since training recurrent neural networks is difficult for various problems, we considered ESNs instead. In addition, we incorporated Hebbian learning for weights between astrocytes and their associated neurons. By systematic analysis of this new dynamics on eight classification tasks we found very little contribution of astrocytes with fixed weights, but in case of Hebbian learning the performance yielded significantly positive outcome. We also examined the dynamic change of the threshold and the activation decay as in FFNNs, bipolar activation functions for neurons and astrocytes (with their output activation within the interval (-1, 1)), various options for training the readout weights. However, these modifications did not perform that well. By swapping astrocytes with neurons in eq. 5.6 the network performed more or less the same.

Future research in this area may follow several directions. The activation function for the astrocyte, as formulated in eq. 4.8, is definitely not the only one and there are several varieties to be considered. Since Ca^{2+} signalling within glial syncytium operates on a much slower pace as opposed to neuronal firing, it may be beneficiary to incorporate this slow, temporal dynamics into astrocytic behaviour. Although our model of an artificial astrocyte includes slow decay, "firing", however, remained still instant. Despite focusing on

the astrocytes as single separate units, it is possible to model glial syncytium and design an astrocytic network of astrocytes connected together, hence fulfilling the biologically plausible spatiotemporal dynamics. Last but not least, instead of modeling the regulation of neuronal excitability, it is possible to design models that also incorporate the rules for synaptic plasticity.

Bibliography

Abbott, N. J., Rönnbäck, L., and Hansson, E. (2006). Astrocyte–endothelial interactions at the blood–brain barrier. *Nature Reviews Neuroscience*, 7(1):41–53.

Alvarellos-González, A., Pazos, A., and Porto-Pazos, A. B. (2012). Computational models of neuron-astrocyte interactions lead to improved efficacy in the performance of neural networks. *Computational and Mathematical Methods in Medicine*, 2012(20):476324.

Alvarez-Maubecin, V., García-Hernández, F., Williams, J. T., and Van Bockstaele, E. J. (2000). Functional coupling between neurons and glia. *Journal of Neuroscience*, 20(11):4091–4098.

Attwell, D., Buchan, A. M., Charpak, S., Lauritzen, M., MacVicar, B. A., and Newman, E. A. (2010). Glial and neuronal control of brain blood flow. *Nature*, 468(7321):232–243.

Azevedo, F. A., Carvalho, L. R., Grinberg, L. T., Farfel, J. M., Ferretti, R. E., Leite, R. E., Filho, W. J., Lent, R., and Herculano-Houzel, S. (2009). Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. *Journal of Comparative Neurology*, 513(5):532–541.

Barres, B. A. and Raff, M. C. (1999). Axonal control of oligodendrocyte development. *Journal of Cell Biology*, 147(6):1123–1128.

Bienenstock, E. L., Cooper, L. N., and Munro, P. W. (1982). Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. *Journal of Neuroscience*, 2:32–48.

Bishop, C. M. (1995). Training with noise is equivalent to Tikhonov regularization. *Neural Computation*, 7(1):108–116.

Bradl, M. and Lassmann, H. (2010). Oligodendrocytes: biology and pathology. *Acta Neuropathologica*, 119(1):37–53.

Brightman, M. and Reese, T. (1969). Junctions between intimately apposed cell membranes in the vertebrate brain. *Journal of Cell Biology*, 40(3):648–677.

Bushong, E. A., Martone, M. E., Jones, Y. Z., and Ellisman, M. H. (2002). Protoplasmic astrocytes in ca1 stratum radiatum occupy separate anatomical domains. *Journal of Neuroscience*, 22(1):183–192.

Charles, A. C., Merrill, J. E., Dirksen, E. R., and Sandersont, M. J. (1991). Intercellular signaling in glial cells: calcium waves and oscillations in response to mechanical stimulation and glutamate. *Neuron*, 6(6):983–992.

Chen, Y., Keogh, E., Hu, B., Begum, N., Bagnall, A., Mueen, A., and Batista, G. (2015). The UCR time series classification archive. www.cs.ucr.edu/~eamonn/time_series_data/.

Chen, Y.-H., Wu, M.-L., and Fu, W.-M. (1998). Regulation of presynaptic NMDA responses by external and intracellular pH changes at developing neuromuscular synapses. *Journal of Neuroscience*, 18(8):2982–2990.

Chesler, M. (2003). Regulation and modulation of pH in the brain. *Physiological Reviews*, 83(4):1183–1221.

Cornell-Bell, A. H., Finkbeiner, S. M., Cooper, M. S., and Smith, S. J. (1990). Glutamate induces calcium waves in cultured astrocytes: long-range glial signaling. *Science*, 247(4941):470.

Dani, J. W., Chernjavsky, A., and Smith, S. J. (1992). Neuronal activity triggers calcium waves in hippocampal astrocyte networks. *Neuron*, 8(3):429–440.

Deiters, O. (1865). *Untersuchungen über Gehirn und Rückenmark des Menschen und der Säugethiere.* Friedrich Veiweg & Sohn.

del Río Hortega, P. (1920). *Estudios sobre la neuroglía: la microglía y su transformación en células en bastoncito y cuerpos gránulo-adiposos.* Imprenta y Librería de Nicolás Moya.

del Río Hortega, P. (1921). *Estudios sobre la neuroglia: La glia de escasas radiaciones (oligodendroglia).* Junta para Ampliación de Estudios e Investigaciones Científicas.

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, 14(2):179–211.

Enkvist, M. and McCarthy, K. D. (1992). Activation of protein kinase C blocks astroglial gap junction communication and inhibits the spread of calcium waves. *Journal of Neurochemistry*, 59(2):519–526.

Eriksson, P. S., Perfilieva, E., Björk-Eriksson, T., Alborn, A.-M., Nordborg, C., Peterson, D. A., and Gage, F. H. (1998). Neurogenesis in the adult human hippocampus. *Nature Medicine*, 4(11):1313–1317.

Farkaš, I., Bosák, R., and Gergel', P. (2016). Computational analysis of memory capacity in echo state networks. *Neural Networks*, 83:109–120.

Farkaš, I. and Gergel', P. (2017). Maximizing memory capacity of echo state networks with orthogonalized reservoirs. In *International Joint Conference on Neural Networks*, pages 2437–2442. IEEE.

Fellin, T., Pascual, O., and Haydon, P. G. (2006). Astrocytes coordinate synaptic networks: balanced excitation and inhibition. *Physiology*, 21(3):208–215.

Finkbeiner, S. M. (1995). Modulation and control of intracellular calcium. In *Neuroglia*, pages 273–288. Oxford University Press.

Font, E., Desfilis, E., Pérez-Cañellas, M. M., and Garcia-Verdugo, J. M. (2002). Neurogenesis and neuronal regeneration in the adult reptilian brain. *Brain, Behavior and Evolution*, 58(5):276–295.

Gallo, V. and Russell, J. (1995). Excitatory amino acid receptors in glia: different subtypes for distinct functions? *Journal of Neuroscience Research*, 42(1):1–8.

García-Marín, V., García-López, P., and Freire, M. (2007). Cajal's contributions to glia research. *Trends in Neurosciences*, 30(9):479–487.

Gee, J. R. and Keller, J. N. (2005). Astrocytes: regulation of brain homeostasis via apolipoprotein E. *International Journal of Biochemistry & Cell Biology*, 37(6):1145–1150.

Golgi, C. (1871). *Contribuzione alla fina anatomia degli organi centrali del sistema nervoso*. Tipi Fava e Garagnani.

Golgi, C. (1885). *Sulla fina anatomia degli organi centrali del sistema nervoso*. S. Calderini.

Halassa, M. M., Fellin, T., and Haydon, P. G. (2007). The tripartite synapse: roles for gliotransmission in health and disease. *Trends in Molecular Medicine*, 13(2):54–63.

Hebb, D. O. (1949). *The Organization of Behavior: A Neuropsychological Theory.* Wiley, New York.

Hertz, L. (2004). The astrocyte-neuron lactate shuttle: a challenge of a challenge. *Journal of Cerebral Blood Flow & Metabolism*, 24(11):1241–1248.

Holmstrom, L. and Koistinen, P. (1992). Using additive noise in back-propagation training. *IEEE Transactions on Neural Networks*, 3(1):24–38.

Honsek, S. D., Walz, C., Kafitz, K. W., and Rose, C. R. (2012). Astrocyte calcium signals at Schaffer collateral to CA1 pyramidal cell synapses correlate with the number of activated synapses but not with synaptic strength. *Hippocampus*, 22(1):29–42.

Hornik, K., Stinchcombe, M., and White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5):359–366.

Hydén, H. and Egyhazi, E. (1963). Glial RNA changes during a learning experiment in rats. *Proceedings of the National Academy of Sciences*, 49(5):618–624.

Ikuta, C., Uwate, Y., and Nishio, Y. (2009). Multi-layer perceptron with chaos glial network. *IEEE Workshop on Nonlinear Circuit Networks*, pages 11–13.

Ikuta, C., Uwate, Y., and Nishio, Y. (2010). Multi-layer perceptron with impulse glial network. *IEEE Workshop on Nonlinear Circuit Networks*, pages 9–11.

Ikuta, C., Uwate, Y., and Nishio, Y. (2011). Performance and features of multi-layer perceptron with impulse glial network. In *International Joint Conference on Neural Networks*, pages 2536–2541.

Ikuta, C., Uwate, Y., and Nishio, Y. (2013a). Four-layer feed-forward neural network with local glia connections. In *Shikoku-Section Joint Convention Record of The Institutes of Electrical and Related Engineers*, page 22.

Ikuta, C., Uwate, Y., and Nishio, Y. (2013b). Improvement of multi-layer perceptron by pulse glial network with dynamic period of inactivity. In *International Joint Conference on Neural Networks*, pages 1638–1643.

Ikuta, C., Uwate, Y., and Nishio, Y. (2014). Investigation of multi-layer perceptron with pulse glial chain including neurogenesis. *IEEE Workshop on Nonlinear Circuits*, pages 70–72.

Ikuta, C., Uwate, Y., Nishio, Y., and Yang, G. (2012). Hopfield neural network with glial network. *IEEE Workshop on Nonlinear Circuits*, pages 369–372.

Ikuta, C., Uwate, Y., Nisho, Y., and Yang, G. (2013c). Investigation of glia-neuron network with group learning of hidden-layer neurons. *IEEE Workshop on Nonlinear Circuits*, pages 440–443.

Irizarry-Valle, Y. and Parker, A. C. (2015). An astrocyte neuromorphic circuit that influences neuronal phase synchrony. *IEEE Transactions on Biomedical Circuits and Systems*, 9(2):175–187.

Irizarry-Valle, Y., Parker, A. C., and Joshi, J. (2013). A CMOS neuromorphic approach to emulate neuro-astrocyte interactions. In *International Joint Conference on Neural Networks*, pages 1–7. IEEE.

Jaeger, H. (2001). The "echo state" approach to analysing and training recurrent neural networks-with an erratum note. *Bonn, Germany: German National Research Center for Information Technology GMD Technical Report*, 148(34):13.

Jaeger, H. (2002). Short term memory in echo state networks. In *GMD-German National Research Institute for Computer Science)*. Citeseer. http://www.faculty.jacobs-university.de/hjaeger/pubs/STMEchoStatesTechRep.pdf.

Ji, K., Akgul, G., Wollmuth, L. P., and Tsirka, S. E. (2013). Microglia actively regulate the number of functional synapses. *PLOS ONE*, 8(2):e56293.

Joshi, J., Parker, A. C., and Tseng, K.-C. (2011). An in-silico glial microdomain to invoke excitability in cortical neural networks. In *IEEE International Symposium of Circuits and Systems*, pages 681–684.

Jourdain, P., Bergersen, L. H., Bhaukaurally, K., Bezzi, P., Santello, M., Domercq, M., Matute, C., Tonello, F., Gundersen, V., and Volterra, A. (2007). Glutamate exocytosis from astrocytes controls synaptic strength. *Nature Neuroscience*, 10(3):331–339.

Kettenmann, H., Backus, K., and Schachner, M. (1984). Aspartate, glutamate and γ-aminobutyric acid depolarize cultured astrocytes. *Neuroscience Letters*, 52(1):25–29.

Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43(1):59–69.

Köles, L., Kató, E., Hanuska, A., Zádori, Z. S., Al-Khrasani, M., Zelles, T., Rubini, P., and Illes, P. (2016). Modulation of excitatory neurotransmission by neuronal/glial signalling molecules: interplay between purinergic and glutamatergic systems. *Purinergic Signalling*, 12(1):1–24.

Kumar, N. M. and Gilula, N. B. (1996). The gap junction communication channel. *Cell*, 84(3):381–388.

Lenhossék, M. (1895). *Der feinere Bau des Nervensystems im Lichte neuester Forschungen: eine allgemeine Betrachtung der Strukturprinzipien des Nervensystems, nebst einer Darstellung des feineren Baues des Rückenmarkes.* Fischer.

Lin, W., Sanchez, H. B., Deerinck, T., Morris, J. K., Ellisman, M., and Lee, K.-F. (2000). Aberrant development of motor axons and neuromuscular synapses in erbB2-deficient mice. *Proceedings of the National Academy of Sciences*, 97(3):1299–1304.

MacVicar, B. A. and Newman, E. A. (2015). Astrocyte regulation of blood flow in the brain. *Cold Spring Harbor Perspectives in Biology*, 7(5):a020388.

Malarkey, E. B. and Parpura, V. (2008). Mechanisms of glutamate release from astrocytes. *Neurochemistry International*, 52(1):142–154.

Marzouki, K. (2015). Neuro-Glial Interaction: SONG-Net. In *International Conference on Neural Information Processing*, pages 619–626. Springer.

Matthews, B. W. (1975). Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochimica et Biophysica Acta (BBA)-Protein Structure*, 405(2):442–451.

McCulloch, W. S. and Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, 5(4):115–133.

Mesejo, P., Ibánez, O., Fernández-Blanco, E., Cedrón, F., Pazos, A., and Porto-Pazos, A. B. (2015). Artificial neuron–glia networks learning approach based on cooperative coevolution. *International Journal of Neural Systems*, 25(04):1550012.

Mirsky, R. and Jessen, K. R. (1999). The neurobiology of Schwann cells. *Brain Pathology*, 9(2):293–311.

Müller, H. (1851). Zur Histologie der Netzhaut. *Zeitschrift für wissenschaftliche Zoologie*, 3:234–237.

Murray, A. F. and Edwards, P. J. (1994). Enhanced MLP performance and fault tolerance resulting from synaptic weight noise during training. *IEEE Transactions on Neural Networks*, 5(5):792–802.

Nayak, D., Roth, T. L., and McGavern, D. B. (2014). Microglia development and function. *Annual Review of Immunology*, 32:367–402.

Nedergaard, M., Ransom, B., and Goldman, S. A. (2003). New roles for astrocytes: redefining the functional architecture of the brain. *Trends in Neurosciences*, 26(10):523–530.

Oja, E. (1982). Simplified neuron model as a principal component analyzer. *Journal of Mathematical Biology*, 15(3):267–273.

Orkand, R., Nicholls, J., and Kuffler, S. (1966). Effect of nerve impulses on the membrane potential of glial cells in the central nervous system of amphibia. *Journal of Neurophysiology*, 29(4):788–806.

Oschmann, F., Berry, H., Obermayer, K., and Lenk, K. (2018). From in silico astrocyte cell models to neuron-astrocyte network models: A review. *Brain Research Bulletin*, 136:76–84.

Parpura, V. (2005). Exocytotic release of glutamate from astrocytes. *Journal of Neurochemistry*, 94:72.

Perea, G. and Araque, A. (2007). Astrocytes potentiate transmitter release at single hippocampal synapses. *Science*, 317(5841):1083–1086.

Pfrieger, F. W. and Barres, B. A. (1997). Synaptic efficacy enhanced by glial cells in vitro. *Science*, 277(5332):1684–1687.

Porto-Pazos, A. B., Veiguela, N., Mesejo, P., Navarrete, M., Alvarellos, A., Ibáñez, O., Pazos, A., and Araque, A. (2011). Artificial astrocytes improve neural network performance. *PLOS ONE*, 6(4):e19109.

Privat, A., Gimenez-Ribotta, M., and Ridet, J. (1995). Morphology of astrocytes. In *Neuroglia*, pages 3–22. Oxford University Press New York.

Ramón y Cajal, P. (1891). El encéfalo de los reptiles. *Trab. Lab. Histol. Fac. Med.(Zaragoza)*, 1891:1–33.

Ramón y Cajal, S. (1897). Algo sobre la significación fisiológica de la neuroglia. *Revista Trimestral Micrografía*, 1:3–47.

Reddy, L. V., Koirala, S., Sugiura, Y., Herrera, A. A., and Ko, C.-P. (2003). Glial cells maintain synaptic structure and function and promote development of the neuromuscular junction in vivo. *Neuron*, 40(3):563–580.

Reichenbach, A., Siegel, A., Rickmann, M., Wolff, J., Noone, D., and Robinson, S. (1994). Distribution of Bergmann glial somata and processes: implications for function. *Journal für Hirnforschung*, 36(4):509–517.

Remak, R. (1838). *Observationes anatomicae et microscopicae de systematis nervosi structura.* Reimer.

Riethmacher, D., Sonnenberg-Riethmacher, E., Brinkmann, V., Yamaai, T., Lewin, G. R., and Birchmeier, C. (1997). Severe neuropathies in mice with targeted mutations in the ErbB3 receptor. *Nature*, 389(6652):725–730.

Rosenberg, P. A. (1991). Accumulation of extracellular glutamate and neuronal death in astrocyte-poor cortical cultures exposed to glutamine. *Glia*, 4(1):91–100.

Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6):386.

Rouach, N., Koulakoff, A., Abudara, V., Willecke, K., and Giaume, C. (2008). Astroglial metabolic networks sustain hippocampal synaptic transmission. *Science*, 322(5907):1551–1555.

Sajedinia, Z. (2014). Artificial astrocyte networks, as components in artificial neural networks. In *International Conference on Unconventional Computation and Natural Computation*, pages 316–326. Springer.

Scheiffele, P. (2003). Cell-cell signaling during synapse formation in the CNS. *Annual Review of Neuroscience*, 26(1):485–508.

Schleich, C. L. (1894). *Schmerzlose Operationen: Örtliche Betäubung mit indifferenten Flüssigkeiten. Psychophysik des natürlichen und künstlichen Schlafes.* Springer-Verlag.

Schwann, T. (1839). *Mikroskopische Untersuchungen über die Uebereinstimmung in der Struktur und dem Wachsthum der Thiere und Pflanzen: mit 4 Kupfertafeln.* Reimer.

Somjen, G. G. (1988). Nervenkitt: notes on the history of the concept of neuroglia. *Glia*, 1(1):2–9.

Teichberg, V. I. (1991). Glial glutamate receptors: likely actors in brain signaling. *The FASEB Journal*, 5(15):3086–3091.

Verkhratsky, A. and Butt, A. M. (2007). *Glial Neurobiology: A Textbook.* John Wiley & Sons.

Verkhratsky, A. and Butt, A. M. (2013). *Glial Physiology and Pathophysiology.* John Wiley & Sons.

Verkhratsky, A., Orkand, R. K., and Kettenmann, H. (1998). Glial calcium: homeostasis and signaling function. *Physiological Reviews*, 78(1):99–141.

Virchow, R. (1856). *Gesammelte Abhandlungen zur Wissenschaftlichen Medicin.* Meidinger.

Virchow, R. (1858). *Die Cellularpathologie in ihrer Begründung auf physiologische und pathologische Gewebelehre: zwanzig Vorlesungen, gehalten während der Monate Februar, März und April 1858 im pathologischen Institute zu Berlin.* Hirschwald.

Volman, V., Bazhenov, M., and Sejnowski, T. J. (2012). Computational models of neuron-astrocyte interaction in epilepsy. *Frontiers in Computational Neuroscience*, 6:58.

Wade, J., Kelso, S., Crunelli, V., McDaid, L., and Harkin, J. (2014). *Biophysically based Computational Models of Astrocyte $\sim$ Neuron Coupling and their Functional Significance.* Frontiers in Computational Neuroscience.

Waldeyer, W. (1891). Ueber einige neuere forschungen im gebiete der anatomie des centralnervensystems1. *DMW-Deutsche Medizinische Wochenschrift*, 17(44):1213–1218.

Weigert, C. (1895). *Beiträge zur Kenntnis der normalen menschlichen Neuroglia*, volume 12. Diesterweg.

Werbos, P. J. (1988). Generalization of backpropagation with application to a recurrent gas market model. *Neural Networks*, 1(4):339–356.

Zupanc, G. (2006). Neurogenesis and neuronal regeneration in the adult fish brain. *Journal of Comparative Physiology A*, 192(6):649.