UNIVERSITY OF GRONINGEN

COMPUTATIONAL COGNITIVE SCIENCE

MASTER'S THESIS

The Influence of Connectivity Sparseness and Alzheimer's Disease on Pattern Separation in a Spiking Neuron Model of the EC-DG-CA3 Hippocampal Circuit

WRITTEN

BY

Chiel Wijs

Supervisor
Dr. Jelmer P. Borst

Abstract

The mechanism of pattern separation allows the hippocampal formation, a brain structure in the medial temporal lobe, to store similar memories using dissimilar patterns of neural activity, preventing collusion between them. A circuit consisting of the entorhinal cortex (EC), dentate gyrus (DG), and region CA3 of the hippocampus is seen as crucial for the process of pattern separation. The DG in particular is thought to contribute through a number of intrinsic neural as well as connectivity properties, one of which is the sparse connectivity of the mossy fibers connecting the DG principal neurons to region CA3. This research used a spiking neuron model of the EC-DG-CA3 circuit to determine how sparse connectivity of the mossy fibers affects pattern separation. Analysis of the model indicated that, within the used modelling framework, sparse connectivity as an isolated feature does not promote pattern separation. Rather, the general decrease in neural activity most heavily influences pattern overlap values. Additionally, the circuit was used to explore how changes that are observed in the DG of a transgenic mouse model of Alzheimer's disease (AD), a type of dementia, can be used to model the deficits in long-term memory characteristic of this neurological disorder. Modelling of an increase of the areas of the EC-DG synaptic junction surfaces, which correlate with synaptic efficacy, proved to be a consistent way to increase pattern overlap. This indicated it as a potential method for inducing AD-related cognitive deficits within the model. As this research focused on how neural and connectivity properties, as well as neurological symptoms of AD, affect a model of the hippocampus without any learning, implementation of a model that allows for increased pattern separation through learning across sparse connections presents itself as an interesting direction for future research into the progression of AD.

Contents

Chapter 1

Introduction

The term *hippocampus* refers to a group of brain structures in the medial temporal lobe as well as a component of that formation (i.e. the *hippocampus proper*; Lavenex, 2012). The group of structures, that is referred to as the*hippocampal formation* (HF), is known to play a role in the formation and recall of long-term declarative memory (Eichenbaum & Cohen, 2004). The role of the hippocampus as a structure for memory systems has been extensively studied through lesion studies. A well-known example is the case of patient H.M. (Eichenbaum, 2012), who underwent surgery to remove part of the HF to mitigate seizures. After his surgery, H.M. showed no true deficits in cognitive functioning other than for, specifically, long-term memory. While H.M.'s long-term memories from well before the surgery were intact, the formation of new long-term memories showed clear deficits. This deficit, however, did not show for *non-declarative*, or *procedural*, memories. Classical conditioning, priming, and skill learning were also preserved following the operation. The following works aims to investigate how certain aspects of a computational model of the hippocampus affects memory formation. More specifically, how these aspects affect the overlap in stimulus induced neural activity in the hippocampus. Both the effect of sparse connectivity between components of the HF, as well as the degradation of the brain that occurs due to Alzheimer's disease (AD), a form of dementia, are explored.

1.1 Pattern Separation and the Dentate Gyrus

One of processes posed as an important part of the functionality of the hippocampus is that of *pattern separation* (Figure 1.1; Bakker, Kirwan, Miller, & Stark, 2008). Pattern separation is the process of creating sparse non-overlapping neural re-representations from similar representations. This allows clear separation between memories of potentially similar stimuli, such that the encoding of new information does not overwrite information that is already stored (Yassa & Stark, 2011). This process is observed to happen in the dentate gyrus (DG) and region CA3.

The DG is a structure that lies between the entorhinal cortex (EC) and the rest of the HF (Figures 1.2 and 1.3; Witter, 2019). The EC itself serves as the relay station between the hippocampal formation (HF) and the neocortex. Excitatory neurons referred to as *granule cells* (GC) are the principal cell type of the DG, comprising 80-90% of its neurons (Vida,

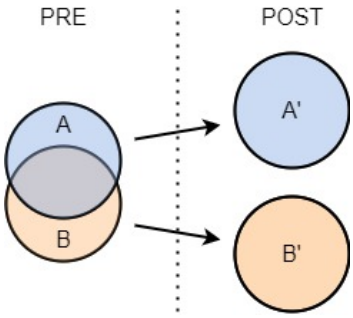

Figure 1.1: **Pattern separation.** To avoid overlap in the storage of similar information, the brain exhibits a phenomenon called pattern separation. Stimuli that induced similar patterns of activity in a presynaptic region produce dissimilar patterns of activity in a postsynaptic region.

Degro, & Booker, 2019). Their axons, called *mossy fibers*, are the main source of input from the DG to region CA3. The mossy fibers additionally innervate a very large number of inhibitory interneurons in both the DG and CA3, outnumbering the number of CA3 pyramidal cells contacted as much 10-fold in the rat hippocampus (Acsady, Kamondi, Sık, Freund, & Buzsáki, 1998).

CA3 is a subregion of the hippocampus proper[1] (Figures 1.2 and 1.3; Witter, 2019). It receives input from the EC both directly and through the DG. The direct connection is sometimes discarded in standard connectivity models of the hippocampus, as the one from Witter shown in Figure 1.2, or specified as a weak source of excitation of CA3 during memory formation as compared to the input from the DG (Lavenex, 2012). CA3 is suggested to play a crucial roll in the formation and retrieval of memories. Recurrent connections of the CA3 pyramidal cells form a so called autoassociative network. This network allows the binding and consolidation of (components of) episodic and long-term memories. Through these recurrent connections, activation of a small part of a memory's associated activity pattern results in the activation of the entire pattern, allowing for the retrieval of a memory based on partial cues (Senzai, 2019). This process is referred to as *pattern completion*.

Multiple aspects of the DG are thought to contribute to its role in the process of pattern separation (Bakker et al., 2008; Senzai, 2019; Lavenex, 2012): its large number of neurons as compared to the EC, its sparse representation of stimuli due to strong hyperpolarization of granule cells, and its sparse but very strong connections to the CA3 pyramidal cells. The lack of direct recurrent connections between granule cells, as opposed to the CA3 autoassociative network, might also help with keeping stimulus representation sparse as well as separate in the DG.

1.2 Modelling the Hippocampus

One method of investigating memory, the hippocampus and the effects of neurological disorders, is through computational modelling. Neural networks (NN) are one method of

[1] CA comes from *Cornu Ammonis*, an earlier name for the hippocampus

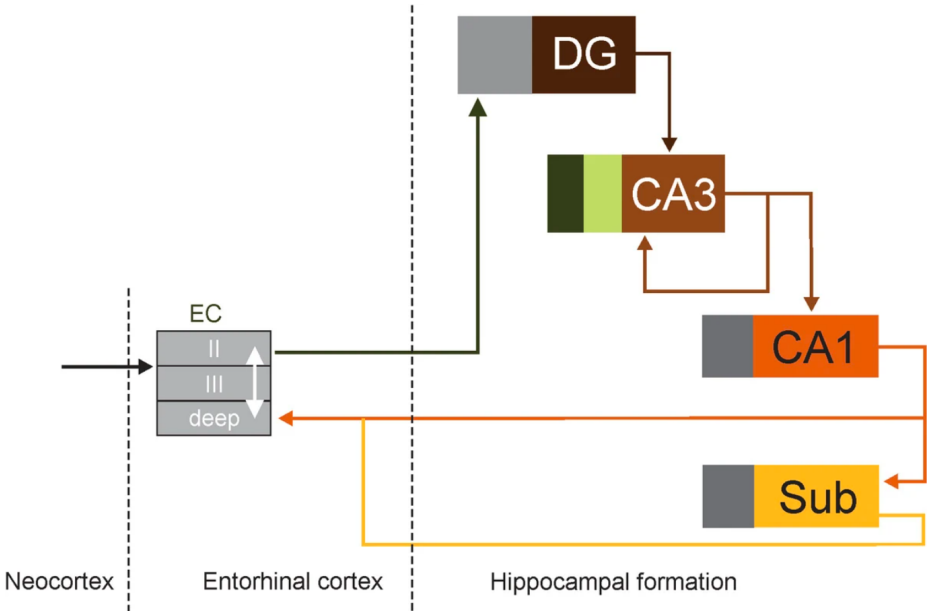

Figure 1.2: **Entorhinal-hippocampal network.** The entorhinal cortex (EC) serves as a relay station between the hippocampal formation and the neocortex, which harbors higher cognitive-functions. The entorhinal cortex projects to the dentate gyrus (DG) which in turn projects to region CA3 of the hippocampus proper. CA3 projects a recurrent connection onto itself, as well as it connects to region CA1. CA1 projects back to the EC both directly and through a structure called the subiculum (Sub). Figure from Witter (2019). Figure C.1 displays a more extensive model of the entorhinal-hippocampal formation from the same source.

computational modelling that is especially interesting, as well as suited, for the investigation of the workings of our brain; the principle of an NN is based on what neuroscientific research has discovered about the composition, structure, and connectivity of the brain. Spiking neural networks are particularly well suited as they mimic the spiking behaviour of cortical neurons. That biological plausibility allows for better evaluation of the computational model (Stewart, 2012); not only the behaviour that emerges can be compared to scientific measurements, but it can also be investigated whether the underlying cause of this behaviour is consistent with what is seen from the brain.

Chavlis and Poirazi (2017) provided a review of computational models of pattern separation in the hippocampus. The review explores the effect of connectivity features, adult neurogenesis in the DG, and morphological and intrinsic properties of the DG cells on pattern separation. A series of models coming from the "Scharfman lab" consisting of not only the principal granule cells of the DG, but also excitatory mossy cells (MCs) and inhibitory Hilar perforant path-associated (HIPP) cells that serve to regulate the activity in the DG, as well as a type of interneuron, provided insight into their contribution to pattern separation (Figure 1.3; Scharfman & Myers, 2016). Analysis of this model highlighted the role of the connectivity *between* different neurons of the DG for pattern separation. The review additionally highlights the role of a feedback connection from CA3 to the DG

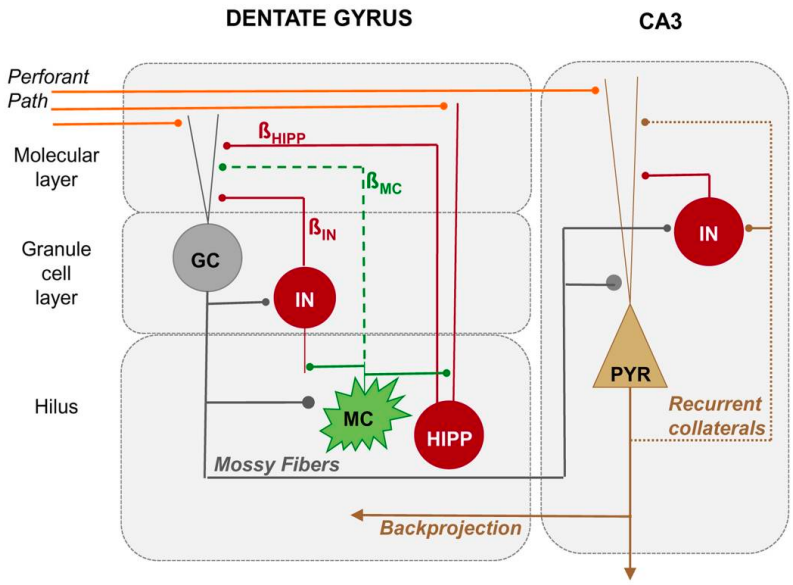

Figure 1.3: **A model of the DG-CA3 circuit.** The entorhinal cortex (EC) excites the dentate gyrus (DG) and CA3 through the perforant path. Both the DG and CA3 contain inhibitory interneurons (IN). The excitatory granule cells (GC) are the primary neuron type of the DG, and the main output of the DG to region CA3. The excitatory mossy cells (MC) and inhibitory Hilar perforant path-associated (HIPP) cells serve to regulate the activity in the DG. Recurrent collaterals of the CA3 pyramidal (PYR) form the autoassociative network that plays an important role for memory consolidation and retrieval. Figure adapted from Scharfman and Myers (2016).

for pattern separation and concludes the section on network connectivity with the notion that in all the reviewed works, deficits in patterns separation correlated with changes in network sparsity. This review however does not touch upon the role of the sparse connectivity of the mossy fibers, which was proposed as one of the aspects through which the DG promotes separated representations in region CA3.

Guzman et al. (2021) used a biologically realistic, full-scale EC–DG–CA3 circuit model to investigate the role of connectivity and synaptic properties on the efficacy of pattern separation in those regions. Contrary to the proposed role of the sparse connectivity of mossy fibers (as discussed so far), they found for their model that a decrease in the number of mossy fiber axon terminals actually decreases pattern separation, instead of contributing to it. Rather, they pose that presynaptic plasticity properties of the DG-CA3 synapses might regulate pattern separation.

1.3 Alzheimer's disease

One illness that affects hippocampal connectivity and synaptic properties is Alzheimer's disease (Setti, Hunsberger, & Reed, 2017). Alzheimer's disease affects more than 36 million

people worldwide, accounting for 60-70% of all dementia cases (World Health Organization, 2021). It is a progressive neurodegenerative disease. As the illness progresses, more and more brain tissue is affected and/or lost (Pihlajamäki & Soininen, 2012). Impairment of the formation of episodic memories is one of the early symptoms of AD. Neuropathologically, AD is characterized by so called *plaques* and *tangles*, as well as brain atrophy (i.e. loss of neurons). Plaques, or more specifically β-amyloid (Aβ) plaques, are extracellular clusters of β-amyloid peptides[2]. Intracellular neurofibrillary tangles (NFTs) consist of aggregates made up of hyperphosphorylated tau protein. Two major alterations to the HF as part of the progression of AD are a loss of the number of synapses between the EC and the DG and an increase in the area of the *synaptic apposition surface* of the remaining synapses, resulting in a stronger synaptic connection.

An often employed method for the research of Alzheimer's disease are transgenic mice (Kitazawa, Medeiros, & M LaFerla, 2012). Transgenic mice are genetically engineered to overexpress a certain mutated gene, most commonly APP, presenilin-1 (PS1) and/or presenilin-2 when used for AD research (PS2; Kitazawa et al., 2012). Mutations of these genes are associated with abnormal metabolism of the compounds that underlie the neuropathological hallmarks of AD.

Alonso-Nanclares, Merino-Serrais, Gonzalez, and DeFelipe (2013) showed that APP/PS1 transgenic mice exhibiting plaques have 37% less synapses per volume in the DG than non-transgenic mice, drawing a similar conclusion as Smith, Adams, Gallagher, Morrison, and Rapp (2000), although the difference is larger. Additionally, the authors show that the size of the DG synapses, or more specifically the *synaptic apposition surface* (SAS) area, was on average 41% larger for the transgenic mice.

The SAS refers to two separate surfaces, with similar areas, that are part of the synaptic junction (Morales, Rodríguez, Rodríguez, DeFelipe, & Merchán-Pérez, 2013; Holtmaat & Svoboda, 2009). The area of the presynaptic active zone (AZ) of the axon terminal positively correlates with the probability of neurotransmitter release. The postsynaptic density (PSD) is the area on which the postsynaptic receptors are situated, and its area is proportional to the number of receptors.

Interestingly, these results come from measurements of plaque-free zones of the DG. As only 4% of DG volume is affected by these plaques, Alonso-Nanclares et al. (2013) pose that the cognitive deficits observed in these mice might be most heavily influenced by the changes to the plaque-free regions.

1.4 Thesis Layout

This thesis aims to investigate two topics concerning pattern separation in the hippocampus. The goals are to explore i) how sparse connectivity affects pattern separation in a simple model of the hippocampus and ii) how we can model the debilitating effects of Alzheimer's disease on pattern separation. For this sake, a model of part of the hippocampus will be created, covering the EC-DG-CA3 circuit, as these regions are deemed relevant for pattern separation.

[2]Peptides are short chains of amino acids and form the building blocks for larger proteins.

The first component of this thesis will be an exploration of the sparse connectivity of the mossy fibers (i.e. the axons of the principal dentate gyrus cells) to region CA3 of the hippocampus and its contribution to pattern separation. As mentioned in Section 1.1, this might be one of the attributes of the DG that contributes to pattern separated representation in CA3, where memory consolidation and retrieval occurs. More specifically, Senzai (2019) states that "The sparseness of the mossy fiber input to CA3 pyramidal cells are assumed to have a randomizing effect on the representations in CA3, making them as different as possible from each other (Treves and Rolls, 1992; Rolls and Treves, 1998; Rolls, 2013)" (p. 46). The network of Guzman et al. (2021) indicated the opposite: an increase in sparseness, by itself, decreases pattern separation. To investigate the effect of connectivity sparseness on pattern separation, different rates of, and methods for, creating sparse connections will be explored.

The second component will be to investigate how the loss of synapses between the entorhinal cortex (EC) and the dentate gyrus (DG), as well as the strengthening of the remaining synapses, affects the activity and pattern separation in the same partial model of the hippocampus. Because AD patients exhibit issues with the formation of long term memories, the expected outcome would be an increase in pattern overlap in region CA3 following the "decay" of the connection between the EC and the DG. Examining the effects separately, as well as combined, allows for the identification of the synaptic strengthening as either a compensatory mechanism against the debilitating consequences of synaptic loss or as a separate synaptic malfunction (Senzai, 2019).

This work continues firstly with a chapter on methodology that covers the modelling framework that was used, a description of stimulus creation, experimental methodology, the partial model of the hippocampus, and analysis methodology, among which a method for quantifying pattern overlap. Then, results relevant to both research topics are presented. The last chapter contains a discussion of these results as well as suggestions for future research.

Chapter 2

Methods

This chapter will first cover the relevant aspects of the neural engineering framework (NEF; Stewart, 2012), the framework used for creating the neural network model. Those are the *encoding* of a vector signal into neural activity, the *decoding* of neural activity into a vector signal, and the calculation of weight matrices used to connect neuron populations. Second, a description of the stimuli and experiment are presented. Then, a description of the model and arguments for relevant design choices are provided as well as description of the implementation of the changes that occur in the DG due to AD. Lastly, a method for measuring overlap in neural activity patterns, called the *Jaccard score* is discussed.

All code relevant to this project can be found on:
https://github.com/ChielWijs/Hippocampus-model-of-sparse-connectivity-and-Alzheimers-disease.

2.1 The Neural Engineering Framework

The smallest computational unit of interest in the NEF is a neuron. Groups of neurons, or neural populations, can represent any signal, which is a vector of arbitrary dimension, through their spiking behaviour. When an individual neuron spikes is determined by its tuning curve. Taking the example of 1-dimensional (1D) neurons, the tuning curves (can) look like the ones displayed in Figure 2.1 (top right). This figure shows that neurons produce varying firing rates depending on the incoming signal. There are multiple attributes used to determine this response. The first attribute is an encoding vector $\mathbf{e}_i$, which indicates the preferred direction of a neuron. A neuron responds most strongly to a signal that has the same direction as its encoding vector. Secondly, an intercept value that indicates at which signal value the neuron starts firing. And third, a maximum firing rate.

As a signal is presented to a population, these properties are used to determine each neuron's input current. How the resulting potential is converted to spiking behaviour is dependent on the type of neuron model that is used. The standard neuron model for the NEF is the leaky integrate-and-fire (LIF) neuron. With this model a spike is generated when the postsynaptic potential reaches a certain threshold, after which the potential is set back to a baseline value that would represent a neuron's resting potential. In addition, the model also continually "leaks" charge, such that the model returns to that resting

potential when not presented with an input current. The bottom row of figure 2.1 shows the subthreshold voltages and spiking behaviour of four neurons in response to a linearly increasing signal (depicted in the upper left corner of the same figure). The spike train is in turn passed through a low-pass filter to finally determine a continuous neural activity value a_i. This filter acts as a model for how the current across a biological synapse is translated to the current induced in a postsynaptic neuron.

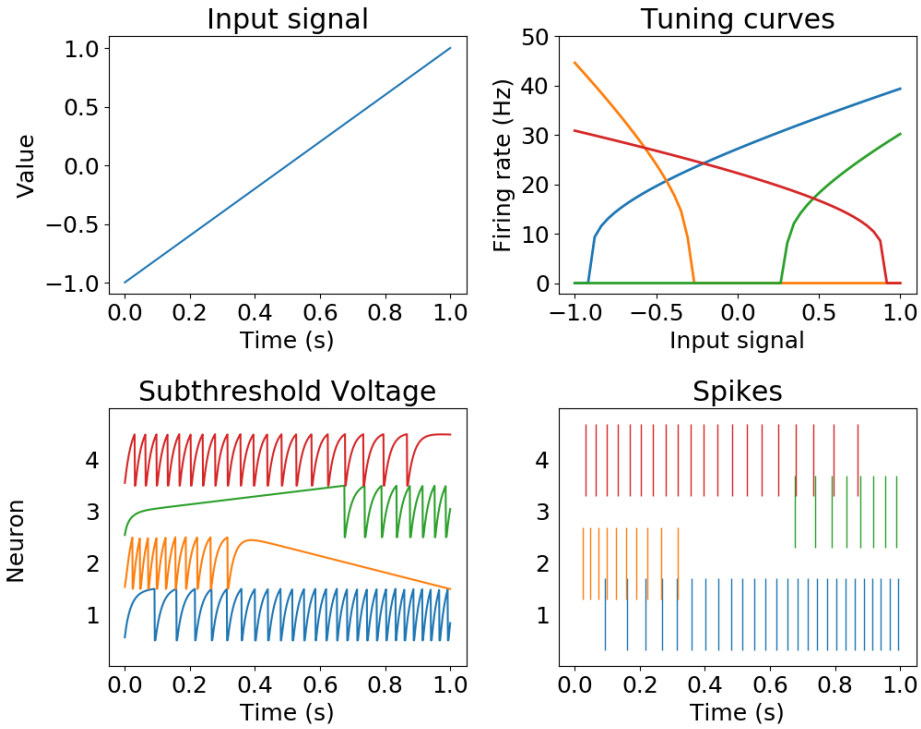

Figure 2.1: **NEF: encoding spikes from a time-variate signal.** (**Top left**) A 1-dimensional time variate signal that linearly increases with time. (**Top right**) Neurons' tuning curves determine to which signal values the neurons respond. (**Bottom left**) Dependent on a neuron's intrinsic properties, the incoming signal is converted to a subthreshold voltage. How these voltages are converted to a spiking behaviour (**bottom right**) depends on the neuron model that is used.

Naturally, to analyze a network, one would want to be able to determine what signal a neuron population represents, that is, *decoding* a signal from the filtered spiking activity. That is, to calculate an estimate $\hat{\mathbf{x}}$ of the original vector $\mathbf{x}$ based on the activity of a neural population. This is possible using a linear *decoder* such that

$$\hat{\mathbf{x}} = \sum_i a_i \mathbf{d}_i$$

As the difference between $\hat{\mathbf{x}}$ and $\mathbf{x}$ should be as small as possible, finding $\mathbf{d}_i$ comes down to a least-squares optimization problem, which can be solved for algebraically. It should be noted that decoders can be solved (or approximated) for any arbitrary function (e.g. a

square root function). The standard connection in the NEF solves for the function $\mathbf{f}(\mathbf{x}) = \mathbf{x}$, which means that the information is just passed along the connection.

With this information, calculating the weight matrix between neural populations is straightforward. Suppose that information is to be passed from a neural population A with decoding vectors $\mathbf{d}_i$ to a population B with encoding vectors $\mathbf{e}_j$. Instead of first decoding the neural activity from A to a signal estimate only to encode that signal to neural activity in B, a dot product multiplication can be used to calculate a set of weights

$$w_{ij} = \mathbf{d}_i \cdot \mathbf{e}_j$$

that suits the same purpose, directly connecting A to B.

For computer-based simulation, the Nengo Python package provides a way to construct large-scale (spiking) neural networks based on the theoretical methods of the NEF (Bekolay et al., 2014). It is an object-oriented implementation. The objects important for this study are the *ensemble*, *node*, *probe*, *connection*, and *network*. A Nengo ensemble is a group of neurons that collectively represent a vector. There are a variety of neuron models to choose from, all with different non-linear functions used to convert the incoming signal to an outgoing one. A Nengo node is an object that can be used to provide non-neural input to other objects or as process output. A Nengo probe can be used to retrieve data from the simulation. A Nengo connection can be used to connect various Nengo objects. Depending on the object types that are connected, a connection can work with decoders and encoders, or with connection weights. A Nengo network is an object that can contain all of the other objects just listed, including other networks. Once specified, a network is passed to a *simulator*. This object first *builds* the network. During this process all the settings and parameters are used to create the values for the objects that are contained within the network, the neuron tuning curves and connection weights among other things. Once the model is build, it can be used for simulation of experiments.

2.2 Semantic Pointer Architecture

Up until now, the concept of transfer of information has been described as the transfer of vectors. However, people do not see vectors floating around, nor do they hear a vector when someone talks to them. As such, a method is needed to assign vector values to certain concepts such as words that allows for the systematic representation and manipulation of these concepts. For this purpose, Eliasmith (2013) created a framework called the Semantic Pointer Architecture (SPA), which uses the computational theories of the NEF to model higher-level cognitive functioning. The SPA is based on the semantic pointer hypothesis. The hypothesis states that higher-level cognitive function is made possible by semantic pointers. Such a semantic pointer (SP) is a neural representation that contains both semantic content as well as (syntactic) information that support the processing of this content.

The Nengo implementation of the SPA is used for creating the vectors that represent concepts. Before an experiment is simulated, a collection of vectors ($\vec{u}_i$) is created such that each is of unit length ($||\vec{u}_i|| = 1$) and, to a certain threshold, dissimilar to one another.

The SPA defines three vector operations that allow for the manipulation of information: *superposition*, *binding* and *unbinding*[1]. Superposition is the act of creating a new vector by the addition of two existing ones such that $\vec{w} = \vec{u} + \vec{v}$. For this research, the vector $\vec{w}$ is additionally scaled to unit-length, such that

$$\vec{w} = \frac{\vec{u} + \vec{v}}{||\vec{u} + \vec{v}||}$$

The resulting SP vector is equally similar to, and has the same magnitude as, either of the SP vectors from which it was created (Figure 2.2, top). As a result there will be some overlap between the neural activity representing the new SP and the neural activity representing either component SP, but it is not the sum of the two activity patterns (Figure 2.2, bottom). This is the case as presenting $\vec{w}$ to a model is not the same as presenting $\vec{u}$ and $\vec{v}$ simultaneously.

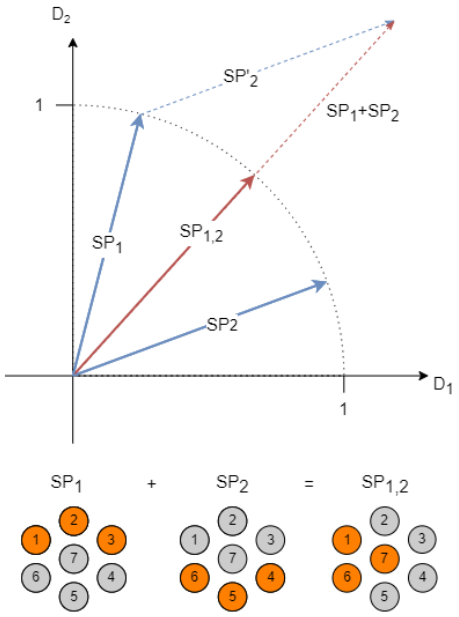

Figure 2.2: **SPA: vector superposition.** (**Top**) The addition of two semantic pointer (SP) vectors results in a vector that is equally similar to both. The newly created vector is scaled to unit-length such that all SP vectors are of the same magnitude. (**Bottom**) Because the new $SP_{1,2}$ vector is similar to either of its components SP_1 and SP_2, there will be some overlap between the neural activity representing $SP_{1,2}$ and the activity patterns representing SP_1 and SP_2. The oranges circles indicate active neurons, the grey circle indicate non-active neurons.

[1]The latter two are not relevant for the current research, and will not be further discussed. Please see Eliasmith (2013), Section *Symbol Processing* for a discussion of these operations.

2.3 Experiment

The model of the ED-DG-CA3 hippocampal circuit performed an experiment to test the effects of connectivity sparseness and AD-related changes to the model on pattern separation. The stimuli presented to the model were those used for an *associative recognition task* (Borst, Schneider, Walsh, & Anderson, 2013; Borst, Ghuman, & Anderson, 2016). Subjects that perform an associative recognition task assess if a combination of two presented items, usually words, is a novelty to them. For that purpose, during a training stage, they are presented with a number of word pairs. During a subsequent testing phase, they are presented with the original pairs, re-pairings of the original pairs, as well as novel pairs.

A behavioural phenomenon called the *fan-effect* is observed for the performance (i.e. response time and error rate) on this type of associative memory task (Borst et al., 2013, 2016). Subjects perform better (i.e. shorter response time and lower error rate) for pairs of which each word was used for only that unique pair during the training phase, so-called fan-1 pairs. Performance is worse for pairs for which each word was used for two unique pairs during the training phase, fan-2 pairs. Because of this fan-effect, I hypothesize that there should be more overlap between the representations (i.e. weaker pattern separation) of fan-2 than fan-1 pairs (but see Anderson (2007) for a different explanation).

The stimuli presented to the model were a set of sixteen word pair stimuli, of which eight fan-1 pairs and eight fan-2 pairs. A 16-dimensional SP was first created for each of the individual words. The SP vector for each word pair was then created by means of superposition and scaling as previously explained in Section 2.2. As the research is not concerned with the effects of learning on pattern separation, there were no separate training and testing phase. Each stimulus was presented to the model once for a period of one second. The simulation was performed using ten models, all with a different random seed to ensure that none are identical (e.g. different encoding vectors for the neuron populations and different semantic pointers for the word pairs).

2.4 Model

The architecture of the model is depicted in Figure 2.3. The model consist of four parts: an input node that is used to present the stimuli to the EC, and three ensembles that represent the EC, the DG, and region CA3 of the hippocampus proper.

The number of neurons in the EC, DG, and CA3 are 500, 5000, and 1000, respectively, all of the leaky integrate-and-fire neuron type. West and Gundersen (1990) estimated that the human DG and CA3-2 [2] are comprised of around $17 * 10^6$ and $2.7 * 10^6$ neurons, respectively (DG:CA3 ratio of 10:1.6). Estimates were done for 5 subjects with a mean age of 66.8 (SD=15.9) that had shown no neurological disorders prior to their death. Šimić, Kostović, Winblad, and Bogdanović (1997) reported similar numbers with the DG containing $19 * 10^6$ neurons and CA3-2 containing $2.6 * 10^6$ neurons (DG:CA3 ratio of 10:1.4), for both young (mean=29.8, SD=13.5; n=8) and older (mean=80.2, SD=7.9; n=10) subjects. The DG-to-CA3 neuron ratio of 10:2 used for the model equals that used by Norman and O'Reilly (2003) for their model of the hippocampus, and is close to that of the human subjects. The

[2] Regions CA3 and CA2 are often treated as a single region due to their similarity

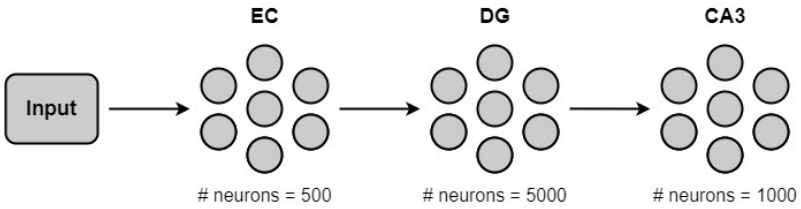

Figure 2.3: **Model architecture.** The model consist of one input node and three ensembles (i.e. groups of neurons) representing the entorhinal cortex (EC), the dentate gyrus (DG) and region CA3 of the hippocampus. The connection from the input node to the EC converts the signal representing a stimulus to neural activity. The EC forwards this representation through the DG to region CA3. The connection from the DG to region CA3 is the connection of interest with regard to the modelling of sparse connectivity and its effect on pattern separation. The earlier connection from the EC to the DG is the connection of interest with regard to the effects of changes in the plague-free zones of the brain as it is affected by Alzheimer's Disease.

ensemble size for the EC is similarly based on the model from Norman and O'Reilly. The choice for neuron counts from subject without a background of neurological disorders was preferred to isolate the effects of the specific AD related changes of synaptic loss and increased synaptic strength in the plague free zones, as loss of neurons is a major hallmark of later stages of AD.

One notable aspect of the neurons in the DG and CA3 ensembles is that all their intercepts, the signal value at which a neuron becomes active, were set to a zero value. With the SPA, during an ISI, the vector representing the word "0" (read: *zero*) is usually presented. This results in an n-dimensional vector with the value zero at each position being presented/shown to the model. Setting the intercepts of these neurons to zero thus causes these ensembles to be silent (i.e. no spiking activity), unless a non-zero n-dimensional vector is presented. The reason that this point is important is that we see pattern separation increase in the hippocampus during learning (Bakker et al., 2008; Yassa & Stark, 2011). If the default intercepts would be used, this learning would also occur when the ISI is presented, which is unwanted. Even though learning is not a part of this research, taking this into account helps generalize the findings to models that do incorporate learning. Borst, Aubin, and Stewart (2021) similarly made use of this intercept distribution.

All objects were connected in the order of input node to EC to DG to CA3 with connections implementing the default identity function, as discussed in Section 2.1.

2.5 Modelling connectivity sparseness and the effects of AD

Connectivity sparseness was implemented by dropping (i.e. converting to zero) a percentage of values of the weight matrix that connects two groups of neurons. Various methods were used to determine which weights were to be dropped: The strongest connections, those with the highest absolute values, could be dropped; The weakest connections, those with the lowest absolute value, could be dropped; Or connections could be dropped at

random. To investigate the effect of connectivity sparseness of the mossy fibers on pattern separation in CA3, 10 to 90% of connections between the DG and CA3 were dropped in 10% increments, using all three methods.

To compensate for the loss in overall activity, which by itself most probably results in stronger pattern separation in a postsynaptic ensemble (see Appendix A), each weight was linearly scaled such that the postsynaptic activity from a sparse network was withing a 10% range of the activity of that same ensemble when fully connected. This process will henceforth be referred to as *activity matching*. Linearly scaling of the weights can be easily accomplished within the NEF by linearly transforming the vector communicated between the two populations, for which a parameter is available for a Nengo connection object. The function of that connection will stay the same, the identity function in the context of this research, but a linear scaling of the mapping is performed after the function is applied. Given that the default transformation value is 1, a transformation value of 1.5 increases each weight to 150% of its initial value, as does a transformation value of 0.8 decrease each weight to 80% of its initial value. As the intercepts for the neurons in the DG and CA3 are all zero, a higher transform value results in more spiking activity (Figure A.1).

The effects of synaptic loss over the EC-DG connections as seen from the AD-transgenic mice was modelled following the same methodology, without the additional activity matching. The effect was tested using both the random and lowest weight dropping methods. This according to my hypothesis that AD would either affect synaptic connections at random, or that it would affect the smallest, and therefore the weakest, synaptic connections. Synaptic loss of the EC-DG synapses was modelled as progressing from no loss to a loss of 37% (Alonso-Nanclares et al., 2013).

The effect of the increase of the synaptic apposition surface area of the EC-DG synapses was modelled by upscaling the weights of a connection, similarly to the activity matching method used on the DG-CA3 connection. SAS area was modelled as progressing from no increase to an increase of of 41% (Alonso-Nanclares et al., 2013).

Lastly, the effect of both progressing in unison was simulated.

2.6 Quantifying Pattern Separation

To quantify the effect of sparse connectivity and the consequences of AD on pattern separation in the DG and CA3, a metric called the Jaccard score was used, following the methodology of Poli, Wheeler, DeMarse, and Brewer (2018). The Jaccard score[3] is a measure of similarity between two sets and is calculated by dividing the union of two sets by the intersection of the two sets such that

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|}$$

where A and B are sets of neurons that are active during presentation of stimulus a and b, respectively. A Jaccard score of one would indicate that two sets would perfectly overlap. A Jaccard score of zero would indicate that there is no overlap between the two sets.

[3] Also referred to as the Jaccard *index* or Jaccard *(similarity) coefficient*

Stronger pattern separation as described earlier would thus be seen as a decrease in the Jaccard score.

As per Poli et al. (2018), a neuron was deemed active when it fired a minimum number of times within a certain time frame. As there is no standard for a spiking frequency threshold (SFT) that one would use to declare that a neuron is part of an activity pattern, Jaccard scores were determined for various SFTs. The number of spikes that each neuron produced were summed for non-overlapping periods of 100 ms, and compared to a threshold of 10, 20, and 30 spikes (corresponding to a SFT of 100 HZ, 200 Hz, and 300 Hz, respectively). The Jaccard score was calculated for each time frame and averaged to reach a final quantification of pattern overlap.

Chapter 3

Results

In this chapter, the results from the experiments are presented. First, a baseline analysis, not pertaining to the effects of changes made to the network, concerning the fan effect and spiking activity throughout the network are presented. Second, the results from the activity-matched sparse connectivity implementation are presented. Lastly, this Section contains the results from the experiment concerning the AD related changes to the hippocampal network.

3.1 Baseline analysis

Analysis of the baseline model[1] model shows that the DG and CA3 exhibit a larger overlap between fan-2 pairs than fan-1 pairs (Figure 3.1, left). The difference in overlap for the different stimuli seems a lot smaller for the EC than for the other two ROIs. Overlap is higher in the EC for all stimuli when compared to the DG and EC, which exhibit similar values in pattern overlap overall.

Similarly, for all SFTs, the EC exhibits a larger portion of active neurons when compared to both the DG and CA3 (Figure 3.1, right). The DG and EC, as was the case with the pattern overlap values, exhibit very similar values for the portion of neurons that exceed the various SFTs. While the different stimuli induced different overlap values, this effect of stimulus fan seems not present for the portion of active neurons. Notably, for the EC, DG, and CA, only a small portion of neurons (<5%) meet the SFT of 300 Hz.

3.2 The effect of connectivity sparseness on pattern separation

Figure 3.2 (top) shows the effect of connectivity sparseness for the DG-CA3 connection on pattern overlap in region CA3, with the implementation of activity matching. Using the method of dropping the lowest weights, pattern overlap, across SFTs, increases after connectivity decreases to 40-50%. Using the method of dropping the highest weights, overlap is generally increasing for a SFT of 100 Hz. For the higher SFTs, the overlap value fluctuates as the connection becomes sparser: it initially decreases, then increases, and then

[1] i.e. all ensembles are fully connected (no sparsity), and the connection weights are not scaled.

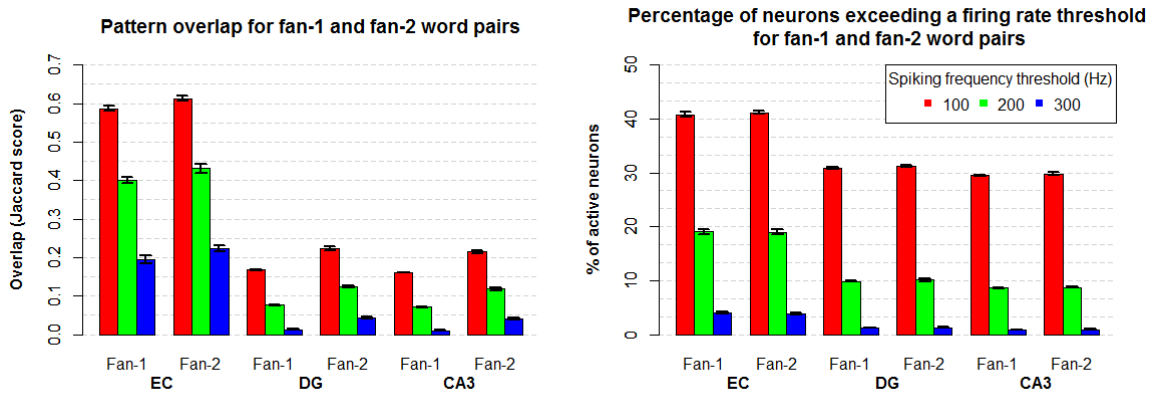

Figure 3.1: **Pattern overlap and neural activity in the baseline model.** (**Left**) The effect of fan on pattern overlap is apparent for the DG and CA3, and to a lesser extend for the EC. Pattern overlap, in general, is higher in the EC than the DG and CA3, which have a different neural intercept distribution. (**Right**) In the EC, About 40% of neurons have a spiking frequency (SF) $\geq$ 100 Hz during stimulus presentation, and about half of those exceed a SF of $\geq$ 200. In both the DG and CA3, about 30% of neurons exhibit a SF of $\geq$ 100 Hz, about a third of those have a SF of $\geq$ 200 Hz. For any of the ROI, only a small portion of neurons (<5%) have a FR $\geq$ 300 Hz. The fan-effect is not apparent for the neural firing rates. Error bars indicate one standard error of the mean.

decreases again, the shape of it is reminiscent of a negated sine wave. The overlap value for the sparsest connection is slightly higher than the value at full connectivity for the 200 Hz SFT and slightly lower for the 300 Hz SFT. Dropping weights at random, in combination with the activity matching process, seems to keep the overlap values, for all SFTs, rather consistent.

Figure 3.2 (bottom) show how the fan effect is affected by the increased connectivity sparseness. For the method of dropping the lowest weights, the effect on relative overlap difference for fan-1 and fan-2 stimuli exhibits a pattern opposite to that seen for the pattern overlap values itself. As the connectivity decreases to 40-50% the relative difference between stimuli decreases for both SFTs, but especially for the neurons with a higher firing rate. When dropping the highest weights, the relative difference perhaps slightly decreases for the low SFT, but even then the effect seems minimal. For the higher SFT an interesting pattern emerges. It exhibits a peek from 100% to 60% connectivity, after which it remains fairly stable. When dropping weights at random, the relative difference in pattern overlap between the fan-1 and fan-2 pairs seems consistent across connectivity sparseness, as was the case with the overlap values.

3.3 The effect of AD related changes in the hippocampus

The figures presented here are those where synaptic loss was modelled as affecting the weakest connections in the hippocampus, as this was deemed the most successful method

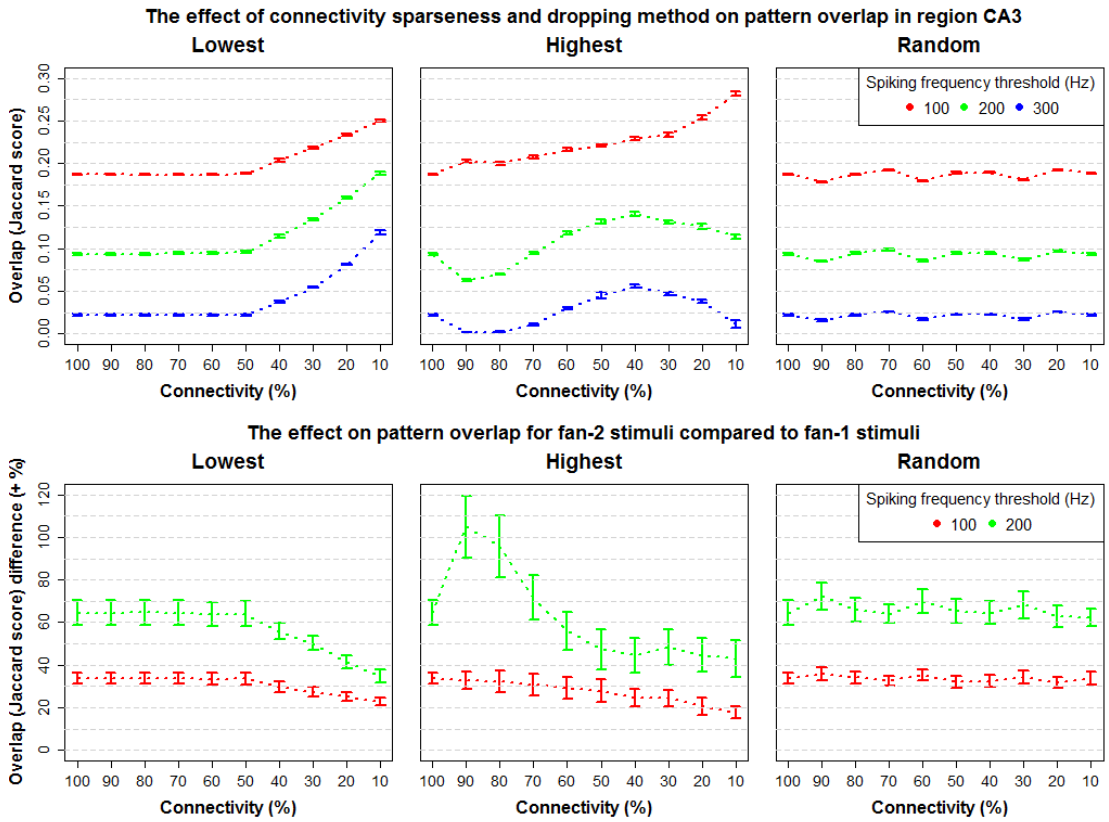

Figure 3.2: **The effect of activity-matched sparse connectivity at the DG-CA3 connection on pattern overlap in CA3.** (**Top**) Dependent on the method used for dropping weights, a different relation between connectivity sparseness and pattern overlap in CA3 emerges. Most notable is that for none of the methods there seems to be a relation where pattern overlap generally decreases as sparseness increases. Rather, the opposite seems to be true when dropping the weakest connections (left), for any spiking frequency threshold (SFT). Similarly, an upward trend is seen for pattern overlap for a SFT of 100 *Hz* when dropping the strongest connections (middle). Dropping weights at random seems to have little effect on pattern overlap (right). (**Bottom**) The relative difference in pattern overlap between stimuli of different fan seems to decrease as the weakest connections are removed (left). When the strongest connections are removed (middle), the relative difference in overlap for a SFT of 200 *Hz* first peaks and then stabilizes. Error bars indicate one standard error of the mean.

of modelling the effect of AD on pattern separation in CA3. This was deemed the most successful method as it was the one that resulted in an increase in pattern overlap in CA3. As discussed in Sections 1.1 and 1.4, increased pattern overlap could be a viable reason for worse performance on memory tasks, possibly resulting in collusion between memories or the overwriting of existing memories during memory formation. With this idea, increased pattern overlap could provide an explanation for the long-term memory deficits exhibited by AD patients.

To keep the current Section readable, figures regarding the model of AD where synaptic loss occurred at random connections, as well as the intermediate effect on the DG for both methods, can be found in Appendix B. A description of these results are presented in the current section however.

Figure 3.3 (top) shows the effect of the progression of the AD related neurological changes to the synapses of the EC-DG connection on pattern overlap in CA3, where memory consolidation and retrieval occurs. Synaptic loss of the weakest EC-DG synapses (without the activity-matching compensation) seems to have no effect on pattern overlap in CA3. The increase of the SAS area of the EC-DG synapses, which results in a stronger connection, seems to correlate with an increase in pattern overlap, rather consistently across participant models. The in unison progression of both synaptic loss and the increase of the SAS area of the EC-DG synapses displays a similar, nearly identical, effect as the SAS area increase by itself.

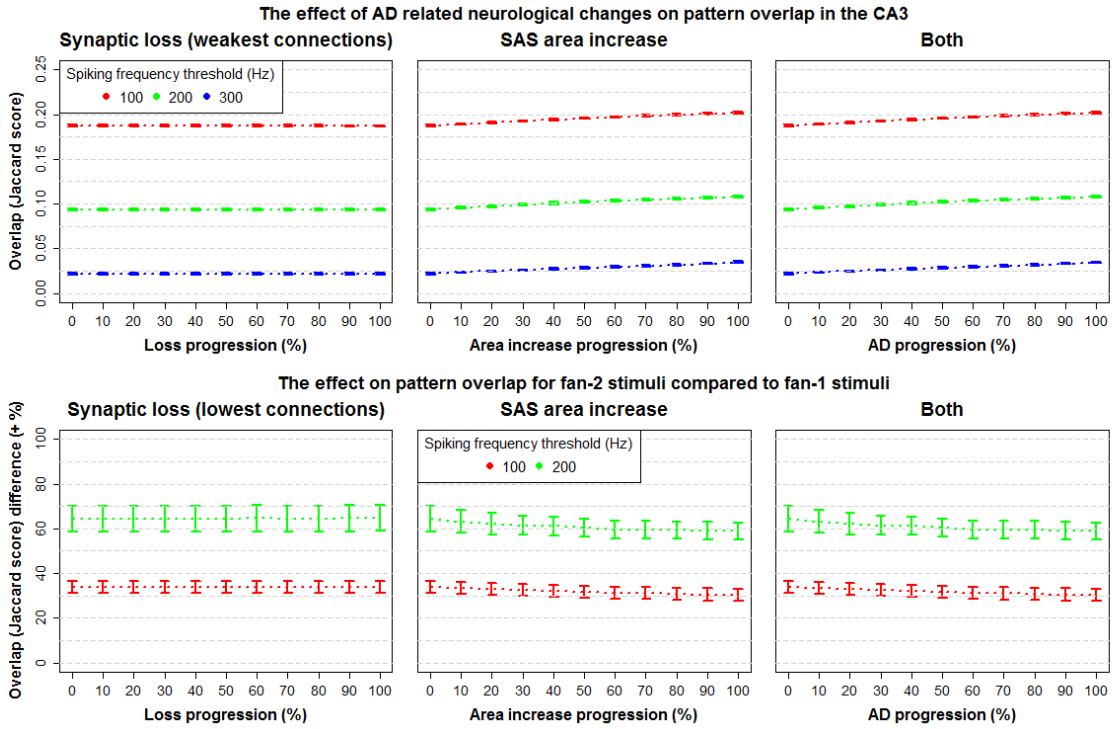

Figure 3.3: **The effect of AD related changes to the EC-DG connection on pattern overlap in CA3 in a fully connected network.** (**Top**) Synaptic loss of the weakest EC-DG synapses seems to have no effect on pattern separation in CA3 (left). An increase of the synaptic apposition surface (SAS) area of the EC-DG synapses seems to increase pattern overlap in CA3 (middle). The progression of both in unison shows an increasing trend for the pattern overlap in CA3. **Bottom** None of the AD related changes, nor the combination of both, seems to have an effect on the relative difference in pattern overlap between fan-1 and fan-2 stimuli in CA3. Error bars indicate one standard error of the mean.

The progression of either of the afflictions separately or in unison does not seem to

have an effect on the relative difference in pattern overlap between stimuli of different fan (Figure 3.3, bottom). While there might be a slight descending trend visible for the relative difference for the progression of the increase of the SAS area of the EC-DG synapses and the progression of AD as a whole, all data points are within a range of one standard error of the mean from one another.

The effect of random synaptic loss of the EC-DG synapses on pattern overlap in CA3 (Figure B.2) shows a different trend than the effect of the loss of the weakest synapses as just described. As synaptic loss of the EC-DG synapses increases, pattern overlap decreases. While the effect of the increase of the SAS area of the EC-DG synapses remains the same, the combination of both also results in a slight decrease in pattern overlap in region CA3 as both afflictions progress.

Interestingly, the effect of at random synaptic loss of the EC-DG synapses on pattern separation in the DG show an opposite trend to that seen in CA3 (Figure B.4): The synaptic loss results in an increase in pattern overlap in the DG. The trends are the same for the DG and CA3 when synaptic loss was modelled as affecting the weakest synapses between the EC and the DG (Figures 3.3 and B.3).

Given the non-existent to negligible effect of the synaptic loss of the weakest connection of the EC-DG synapses on pattern overlap in CA3, an additional simulation was performed. Instead of starting the connection at 100% connectivity, the connection started at 10% connectivity. As such, this simulation affected connections with much higher weights, and thus stronger influence, than the first simulation of synaptic loss. Inspecting Figure A.2 it seems that the 10% strongest connection have a stronger influence on pattern separation than the lowest 90%. Figure 3.4 Shows the results of this additional simulation of the effect of AD related changes to the EC-DG synapses.

Indeed, synaptic loss occurring at already sparse EC-DG synapses induces an effect on pattern overlap in CA3. An increase in sparseness results in weaker pattern overlap, similar to the result from the experiment with fully connected EC-DG synapses which were dropped at random (Figure B.2). The initial sparseness of the EC-DG synapses does not seem to change how the increase in SAS area affects pattern overlap.

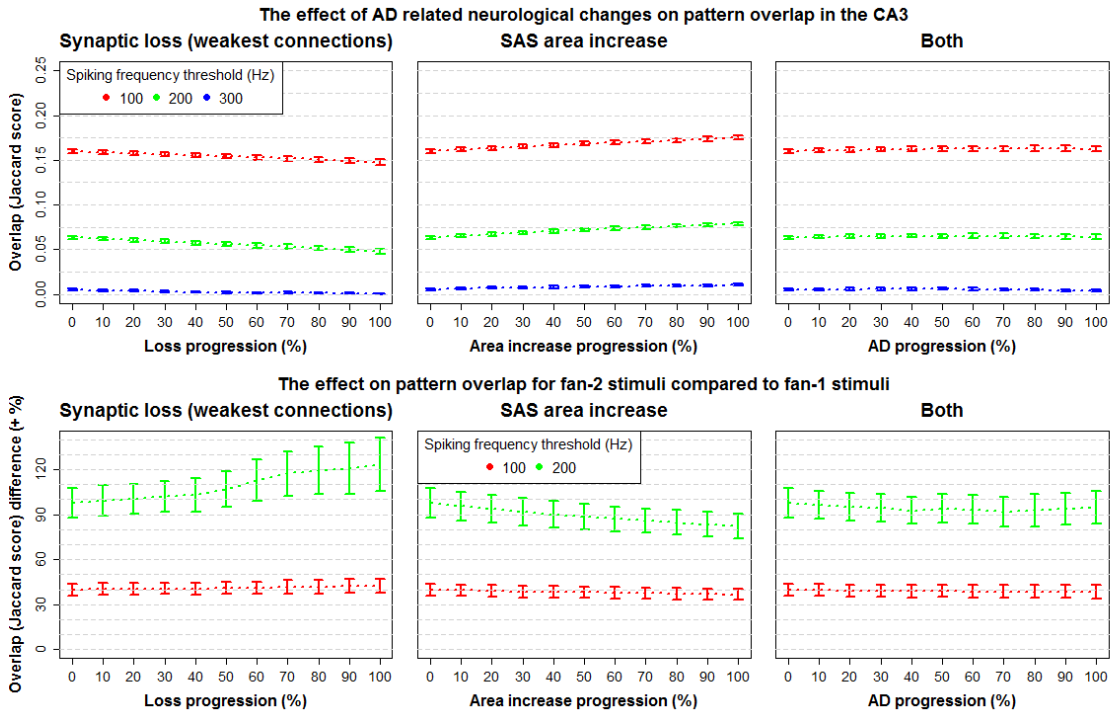

Figure 3.4: **The effect of AD related changes to the EC-DG connection on pattern overlap in CA3 in an already sparse network.** (**Top**) Synaptic loss of the weakest EC-DG synapses seems to decrease pattern overlap in CA3 (left). An increase of the synaptic apposition surface (SAS) area of the EC-DG synapses seems to increase pattern overlap in CA3 (middle). The progression of both in unison, when initial connectivity of the EC-DG synapses was set at 10%, seem to have a zero net effect on pattern overlap in CA3 (right). (**Bottom**) None of the AD related changes, nor the combination of both, seems to have an effect on the relative difference in pattern overlap between fan-1 and fan-2 stimuli in CA3. Error bars indicate one standard error of the mean.

Chapter 4

Discussion

For this study, two topics with regard to structural connectivity of the hippocampus and its effect on overlap between neural activity patterns have been investigated. Both these topics were researched using a spiking neuron model of the network consisting of populations representing the entorhinal cortex (EC), the dentate gyrus (DG), and region CA3 of the hippocampus proper. The first topic that was researched was how sparse connectivity of the mossy fibers that connect the DG and CA3 affects pattern overlap in region CA3. The sparse but strong connectivity of these synapses seen at the rodent and human hippocampus is thought to be one of the aspects through which the DG contributes to pattern separation in region CA3, where the consolidation during memory storage and pattern completion during memory retrieval is thought to occur (Senzai, 2019). The second topic was to experiment with a way to simulate the effects of Alzheimer's disease (AD) related neurological changes to plaque-free zones in the DG on, again, pattern overlap in region CA3. Alonso-Nanclares et al. (2013) showed that plaque bearing mice exhibited synaptic loss of the synapses connecting the EC and the DG. In addition, the area of the synaptic apposition surface (SAS), the surfaces at both ends of the synaptic cleft, whose areas correlate with synaptic efficacy, increases. This phenomenon is speculated either to be a compensatory mechanism for the loss of transmission due to synaptic loss or to be a synaptic malfunction by itself (Senzai, 2019).

4.1 The effect of increased network sparsity

Analysis of the effects of weight scaling, without any change to network connectivity, indicated that a decrease in synaptic activity in and of itself can be the cause of a decrease in pattern overlap (Figure A.1). The weight scaling directly fulfills a scaling of the information vector that is passed through a connection. As the intercepts of the CA3 neurons population were all set to zero, them representing a shorter vector results in less spiking activity. With the idea to prevent this effect from colluding with the influence of the connectivity sparseness a method referred to as activity matching was applied, where the loss in post synaptic neural activity due to synaptic loss was compensated through scaling of residual connection weights.

With this activity matching approach, it seems like increased connectivity sparseness

does not produce the effect as we might expect it. After all, the general train-of-thought is that the sparseness of the modeled connection might contribute to pattern separation (Bakker et al., 2008; Senzai, 2019; Lavenex, 2012). None of the methods of dropping weight to increase sparseness (i.e. discarding the lowest, discarding the highest weight, or discarding them at random) results in a consistent decrease in pattern overlap, as can be seen from Figure 3.2 (top).

A possible explanation for the effect observed when dropping the lowest weights comes from the fact the that strongest connections seem disproportionately influential on pattern overlap. Dropping the lowest weights without the activity matching process has seemingly no effect on pattern separation (Figure A.2, top). However, it does have an effect on neural activity, which decreases as the weakest connections are dropped (Figure A.2, bottom). As such, during the experiment, the activity matching procedure results in a scaling of the connection weights as the connection becomes sparser. This results in the strengthening of the remaining, already strongest, connections, increasing pattern overlap.

Not only do the strongest connections seems most influential on pattern overlap, they also seems disproportionately influential on the amount of neural activity (Figure A.2, bottom); dropping the 20% strongest connection results in less activity than dropping the 80% weakest. When dropping the highest weights, the weight scaling that is necessary to accomplish activity matching must thus be quite severe. Even though the remaining weakest connections have less influence on pattern overlap, their extreme excitation apparently still results in an increase in pattern overlap. Even though the excitation is assumed to be rather extreme due to the activity matching process, it seems that the effect on overlap between neurons at higher spiking frequency thresholds ($200 - 300$) does not necessarily follow the presented reasoning. It might be that this effect is less constant/predictable due the smaller portion of neurons that meet these spiking frequency thresholds (Figure 3.1, right and Figure D.1), but that is merely speculative. The effect of dropping the highest weights displays a somewhat similar wave pattern for pattern overlap and the portion of active neurons for a spiking frequency $\geq 300\ Hz$.

When inspecting Figures A.2 and A.1, it seems that dropping weights at random and down scaling the connection weights (from one down to zero) show a similar effect on both pattern overlap and neural activity. Given that dropping weights at random, with activity matching, results in fairly constant overlap values, it thus seems that dropping these weights, without activity matching, has a similar effect on the network as the weight scaling.

The effect of the activity matched sparseness does compare to that found by Guzman et al. (2021). The analysis of their biologically realistic model of the EC, DG, CA3 network also indicated that the sparsity of the mossy fibers decreased, rather than increased pattern separation. Interestingly, the authors applied the same concept of activity matching to preserve the activity level of their network. The varying sparsity was modelled as an increase/decrease of the number of synaptic boutons (i.e. axon terminals) of each dentate gyrus granule cell. They do not seem to specify to which target neurons, new or existing, these additional boutons connect.

4.2 The modeller's toolbox: How to regulate pattern separation

It should be noted that the way through which sparseness is created within the model does not (necessarily) reflect the natural process of brain development. While it might be helpful to try and isolate the effect of connectivity sparseness within the modelling approach, I see no reason why the brain would adhere to this principal of activity matching during development, nor does it make sense that the brain would start with a fully connected network. Considering this, it thus seems that the findings could be used to support the idea that increased connectivity sparseness aids pattern separation.

Figure A.2 shows that dropping either the strongest connections, or connections at random, without activity matching, results in stronger pattern separation in region CA3. Another method for reducing pattern overlap would be to apply a linear mapping to the function over the connection between two stimuli, not for activity matching, but to purposefully reduce the amount of neural activity. As said, there seems to be no argument to preserve neural activity at a certain level from a biological point-of-view when initiating the model. It should be clear that both these methods work with, and perhaps due to, the all zero intercepts of the post population. While further work should indicate how sparseness affects pattern overlap with the default random distribution of intercepts, based on the underlying mathematics, it seems not the case that the linear mapping would produce the same consistent effect on neural activity with the default random distribution of intercepts.

Also, further application for modelling with Nengo does not require a certain level of neural activity necessarily[1]. This research has so far focused on the process of creating sparse representations of memories in CA3 from the information that reaches the hippocampal formation from the neocortex. This however is only one part of hippocampal functionality. In order to "relive" one's memory, there has to be a way to have a memory that is stored in CA3 induce activity in the neocortex that would be similar to the information that was originally encoded. Some model of hippocampal functioning successfully implement this process through error-driven learning mechanism that connects CA3 to the next-in-line region CA1 (Ketz, Morkonda, & O'Reilly, 2013). In a nutshell, the error-driven learning signal is created through the activity in CA1, and a signal coming from the EC, whose activity was also used to drive activity in CA3 for initial memory encoding. Borst et al. (2021) used an error-driven learning rule in a similar fashion for their model of associative recognition. In Nengo, the learning rule that they used, called the prescribed error-sensitivity (PES; MacNeil & Eliasmith, 2011) rule, is used to adapt the connections weights. For the calculation of this weight change, a leaning rate constant and the variable presynaptic activity a_i (see Section 2.1 for a description of how this value is calculated) both act as a linear scalar of the error signal. As such, a decrease in overall spiking activity, with subsequently lower a_i values, can be compensated for by selecting a larger learning rate value if necessary.

Intercept selection has been presented a couple of times as a one of the reasons as to why sparseness and weight scaling have their particular influence on pattern overlap. The analysis of the baseline model shows that intercept selection by its own right could be con-

[1] Within reason that is, it might be necessary to at least have enough activity to represent a vector consistently to some degree for example.

sidered a method for regulating pattern overlap. Using the identity function to connect all the neural populations, it seems that population size is not influential on pattern overlap, as the DG and CA3 exhibit very similar pattern overlap values (Figure 3.1, left) even though the DG ensemble contains five times as much neuron's. What is apparent is the large difference in overlap values for the EC and both DG and CA3. What differentiates the EC from the other two regions is that its neural intercepts are distributed according to the default random distribution. As such, it seems that, due to the assumptions of the NEF, populations size by itself cannot be used to regulate pattern overlap. It should be considered that this conclusion is based on the use of the Jaccard score as the measure of overlap, and that the reported scores are all averages. Within the NEF, larger populations are capable of more accurately representing a value over time. Future experiments could indicate how population size might thus still influence the behaviour of a spiking neuron model of cognition.

The seemingly non-existent effect of stimulus fan on neural activity (Figure 3.1, right) most probably arises from the fact that the vectors used to represent the word pairs are all of equal length. As discussed, the amount of activity of the DG and CA3 neurons correlates with the length of the represented vector. As such, stimulus fan has no effect on the amount of neural activity in these regions. The relation between vector length and neural activity for the randomly distributed intercepts of the EC population was not investigated for this work. However, also for the EC we see that stimulus fan has no effect on the amount of neural activity.

4.3 Pattern separation through learning

While this research so far has focused on a static model of the hippocampus, behavioural data shows that pattern separation also occurs as an active process: Repetition aids our memory performance, for example during associative recognition tasks (Borst et al., 2013, 2016). It is even so that the term *pattern separation* as used in scientific literature most commonly refers to this active process.

Borst et al. (2021) presents a new learning rule for the NEF/Nengo that accomplishes this active pattern separation in such a way that a model of associative recognition (AR) produces neurological activity, in addition to behaviour, similar to that exhibited by human subjects. The learning rule relies on reorientation of the neural encoders, in effect changing the signals for which the neurons are most active. During presentation of a stimulus, post-synaptic neurons that are active above a certain threshold orient their encoders towards the incoming signal, while neurons that are active below a certain threshold orient their encoders away from the incoming signal. This results in a fraction of the neurons becoming more responsive the next time that the stimulus is presented, while the rest becomes less active. The results is a decrease in pattern overlap given selection of a threshold.

This model of associative recognition shows neural dynamics quite comparable to that seen from human subjects, both in terms of timing of activity of certain regions of the brain as well as the amount of activity's dependence of stimulus fan. On the downside, the difference in error for fan-1 and fan-2 stimuli does not match the human data as well: the model exhibits a larger difference in error rate. Here the opportunity presents itself

to incorporate the findings of the current research with this existing literature. Figure 3.2 (bottom) shows that increasing sparsity by dropping the weakest connection results in a smaller relative difference in overlap value between fan-1 and fan-2 stimuli. The major issue is that the learning rule used by Borst et al. (2021) does not allow the use of sparse connections, as it was developed for use with a connection that makes use of decoders and encoders, not connection weights.

A solution would be to create a learning rule that allows for adaptation of synaptic weights, instead of encoders. Consider the commonly used Hebbian theory of long-term potentiation which postulates that synchronous pre- and postsynaptic activity strengthens, while asynchronous activity weakens, the connection between neurons (Hebb, 1949). Add to this a threshold that postsynapatic activity should exceed for strengthening of the connection to take place during synchronous pre- and postsynaptic activity. Given a sufficiently high threshold, after a time, only a number of the neurons that were initially responsible for representing a stimuli will still be active during stimulus presentation. While considerable efforts have been made to implement such a learning rule, they have not yet proven successful.

One way by which the AR model of Borst et al. (2021) might still benefit from the current findings, without the need for the additional learning rule, is to introduce an intermediary ensemble to the memory system. When placed between an input ensemble and the memory ensemble (which fulfills the purpose of region CA3), sparse connectivity over the connection from the input to the intermediary ensemble might serve the purpose of reducing the relative difference in pattern overlap between fan-1 and fan-2 stimuli even before the signal reaches the memory ensemble. This proposed solution however is speculative, future implementation should indicate if this makes a feasible solution to the "faulty" error rates produced by the AR model.

4.4 Simulating Alzheimer's disease

The simulations indicate that loss of EC-DG synapses in and of itself is not detrimental to pattern separation in region CA3. The opposite even seems to emerge for this particular model, assuming that the EC-DG synapses are affected at random, the pattern overlap values in CA3 decreases. While the increased SAS area increases overlap, the progression of both in unison still results in an overall decrease. In that case, we would actually expect increased memory performance. Overall, it seems that any increase in overlap in CA3, for this model, is solely dependent on the increased SAS area, implement as a linear mapping of the communicated vector, and therefore a scaling of the connection weights.

While specified as a separate topic of research withing this paper, there is overlap in the findings of the research into the effect of sparse connectivity and that of Alzheimer's disease. As a fact, the main difference between the simulations is which connection was affected, and to what degree. As such, it seems like results from the first part of this research provides us with an interesting proposition with regards to the "role" of the increased synaptic apposition surface area. As presented earlier, there is still speculation as to the cause of the increase in the SAS area. Is it a way by which the brain tries to compensate for the loss of synapses, or is it a separate, simultaneously occurring malfunction? Let us

hypothesize that it is compensatory. In that case we should wonder what it exactly tries to compensate for: maybe the loss on overall neural activity, or perhaps even an observed decline in cognitive functioning. The results from the activity matched experiment into the effect of increased sparsity (i.e. synaptic loss) shows us that the former could potentially induce increased pattern overlap in CA3, with a deficit in memory performance as a consequence. Were it a separate synaptic malfunction, then this model also shows that. The separate effect produces a consistent increase in pattern overlap. However, as previously described, this is most likely the results of the intercept selection for this model.

For future works, the effect of synaptic loss and increases SAS area might prove quite informative. As was described in the previous Section, an NEF learning rule that implements pattern separation awhile allowing for the sparsification of the connection will have to be developed. It would be interesting to see how the AD related affliction affect the brain during stimulus presentation; After all, our brain is a dynamic system, not a static one as is the model used for this research.

4.5 Limitations and considerations

One rather obvious limitation, but no less major because of it, is the architecture of the model. It is quite simple, relying only on feed-forward connections and a single neuron type. While the neurons themselves might be a fair model of biological neurons, the model as a whole can hardly be considered biologically accurate. While Figure 1.2 was presented as a representative of the connectivity of the hippocampal formation in the Introduction, more extensive schematics show the intricacies and extend of the connectivity of the various components (See Figure C.1). While the weight matrices of the model contain both positive and negative weights, resembling excitatory and inhibitory synapses, it does not encompass the connectivity of the different neurons types, even within the DG itself (Figure C.2), and the influence they have on activity in the DG. The simplification of the observable is of course a necessity when it comes to modelling the brain. Even the very extensive model of the hippocampus presented by Trujillo (2014) and the models discussed by (Chavlis & Poirazi, 2017) cannot encompass every single connection that we find in the HF. In addition, models that produce the expected behaviour, even though the underlying cause of it might not be true to biology, are not necessarily less useful because of it.

One consideration – for the study of Alzheimer's disease in general – is that much of our knowledge is based on research into non-human mammals, especially rodents. For example, the mouse model of AD investigated by Alonso-Nanclares et al. (2013) is based only on the presence of plaques, the mice exhibited no neurofibrillary tangles, even though these are one of the neuropathological hallmarks of AD for human patients. As they themselves state, this model allows them to isolate the effects of plaque formation. However, due to the differences in brain organization between mice and humans (Felipe, 2010; as cited by Alonso-Nanclares et al., 2013), even if the isolated effects would be clear, extrapolation of their findings to human subject might prove difficult. It should be noted however that this does not mean that findings as presented by them or the findings from the current research are not valuable for the overall research into AD. Even if the underlying mechanisms do not work as we expect, finding a consistent way to regulate something such as

pattern overlap can help us with modelling the behavioural consequences of AD.

One of the most valuable additions for the research into Alzheimer's disease would be data from human participants similar to the type we see from rodents from, for example, Alonso-Nanclares et al. (2013). While Alzheimer's disease is ultimately fatal, being accounted the seventh-leading cause of death in the United States of America in 2020 and 2021, it is mostly older subjects that die from it (Alzheimer's Association, 2022). As such, there is little opportunity for the use of invasive methods for the collection of human data regarding the development of AD. While neuroscientists can currently select from a more than impressive arsenal of non-invasive imaging techniques, the invasive techniques used for the study of animals that have met their end prematurely provide much more detail. With regard to data from behavioural experiments, it would be interesting for this research to see how Alzheimer's disease affect the fan effect on performance on associative recognition tasks.

As discussed earlier, I see the implementation of a learning rule that results in pattern separation while allowing for the use of sparse connection as a most interesting addition to the Nengo toolbox. With it, it would be possible to see how the sparse connectivity of the mossy fibers affects learning, and how the progression of AD influences the formation of long-term memory.

4.6 Conclusion

In conclusion, analysis of the model has shown that both intrinsic neuron properties as well as network connectivity can be selected to regulate the amount of pattern overlap in neural populations. These findings pertain to the overlap values as present when the model is first initiated. As such, this knowledge could be used to tune models where an expected output cannot be accomplished through the use of learning or the setting of network parameters not discussed in this work. Future implementation of an algorithm that allows for the modelling of the effect of learning on pattern separation in a sparse network would present an interesting direction for the modelling of the progression of Alzheimer's disease (AD). Concerning this neurological illness, this research has provided a systematic approach for the regulation of increased pattern separation in the neural network model of the hippocampus, inspired by the changes that occur in the brain during the progression of AD. A consistent method for modelling the consequences of AD on behaviour might allow us to make predictions about the performance of AD patients on particular tasks, which could be a great benefit as data from this type of patients can be difficult to acquire.

References

Acsady, L., Kamondi, A., Sık, A., Freund, T., & Buzsáki, G. (1998). Gabaergic cells are the major postsynaptic targets of mossy fibers in the rat hippocampus. *Journal of neuroscience*, *18*(9), 3386–3403.

Alonso-Nanclares, L., Merino-Serrais, P., Gonzalez, S., & DeFelipe, J. (2013). Synaptic changes in the dentate gyrus of app/ps1 transgenic mice revealed by electron microscopy. *Journal of Neuropathology & Experimental Neurology*, *72*(5), 386–395.

Alzheimer's Association. (2022). *2022 alzheimer's disease facts and figures.* Retrieved 2022-08-31, from `https://www.alz.org/media/documents/alzheimers-facts-and-figures.pdf`

Anderson, J. R. (2007). *How can the human mind occur in the physical universe?* (Vol. 3). Oxford University Press.

Bakker, A., Kirwan, C. B., Miller, M., & Stark, C. E. (2008). Pattern separation in the human hippocampal ca3 and dentate gyrus. *science*, *319*(5870), 1640–1642.

Bekolay, T., Bergstra, J., Hunsberger, E., DeWolf, T., Stewart, T., Rasmussen, D., ... Eliasmith, C. (2014). Nengo: a Python tool for building large-scale functional brain models. *Frontiers in Neuroinformatics*, *7*(48), 1–13. doi: 10.3389/fninf.2013.00048

Borst, J. P., Aubin, S., & Stewart, T. C. (2021). *A functional spiking-neuron model of memory formation and recall.* (Manuscript submitted for publication)

Borst, J. P., Ghuman, A. S., & Anderson, J. R. (2016). Tracking cognitive processing stages with meg: A spatio-temporal model of associative recognition in the brain. *NeuroImage*, *141*, 416–430.

Borst, J. P., Schneider, D. W., Walsh, M. M., & Anderson, J. R. (2013). Stages of processing in associative recognition: Evidence from behavior, eeg, and classification. *Journal of cognitive neuroscience*, *25*(12), 2151–2166.

Chavlis, S., & Poirazi, P. (2017). Pattern separation in the hippocampus through the eyes of computational modeling. *Synapse*, *71*(6), e21972.

Eichenbaum, H. (2012). Amnesia and the hippocampal memory system. In (Vol. 151). Oxford University Press.

Eichenbaum, H., & Cohen, N. J. (2004). Hippocampal function in humans. In H. Eichenbaum & N. J. Cohen (Eds.), . Oxford University Press on Demand.

Eliasmith, C. (2013). *How to build a brain: A neural architecture for biological cognition*. Oxford University Press.

Guzman, S. J., Schlögl, A., Espinoza, C., Zhang, X., Suter, B. A., & Jonas, P. (2021). How connectivity rules and synaptic properties shape the efficacy of pattern separation in the entorhinal cortex–dentate gyrus–ca3 network. *Nature Computational Science*,

1(12), 830–842.

Hebb, D. O. (1949). The first stage of perception: growth of the assembly. *The Organization of Behavior*, *4*, 60–78.

Holtmaat, A., & Svoboda, K. (2009). Experience-dependent structural synaptic plasticity in the mammalian brain. *Nature Reviews Neuroscience*, *10*(9), 647–658.

Ketz, N., Morkonda, S. G., & O'Reilly, R. C. (2013). Theta coordinated error-driven learning in the hippocampus. *PLoS computational biology*, *9*(6), e1003067.

Kitazawa, M., Medeiros, R., & M LaFerla, F. (2012). Transgenic mouse models of alzheimer disease: developing a better model as a tool for therapeutic interventions. *Current pharmaceutical design*, *18*(8), 1131–1147.

Lavenex, P. (2012). Functional anatomy, development, and pathology of the hippocampus. In T. Bartsch (Ed.), (Vol. 151). Oxford University Press.

MacNeil, D., & Eliasmith, C. (2011). Fine-tuning and the stability of recurrent neural networks. *PloS one*, *6*(9), e22885.

Morales, J., Rodríguez, A., Rodríguez, J.-R., DeFelipe, J., & Merchán-Pérez, A. (2013). Characterization and extraction of the synaptic apposition surface for synaptic geometry analysis. *Frontiers in neuroanatomy*, *7*, 20.

Norman, K. A., & O'Reilly, R. C. (2003). Modeling hippocampal and neocortical contributions to recognition memory: a complementary-learning-systems approach. *Psychological review*, *110*(4), 611.

Pihlajamäki, M., & Soininen, H. (2012). Alzheimer's disease. In T. Bartsch (Ed.), (Vol. 151). Oxford University Press.

Poli, D., Wheeler, B. C., DeMarse, T. B., & Brewer, G. J. (2018). Pattern separation and completion of distinct axonal inputs transmitted via micro-tunnels between co-cultured hippocampal dentate, ca3, ca1 and entorhinal cortex networks. *Journal of neural engineering*, *15*(4), 046009.

Scharfman, H. E., & Myers, C. E. (2016). Corruption of the dentate gyrus by "dominant" granule cells: Implications for dentate gyrus function in health and disease. *Neurobiology of learning and memory*, *129*, 69–82.

Senzai, Y. (2019). Function of local circuits in the hippocampal dentate gyrus-ca3 system. *Neuroscience research*, *140*, 43–52.

Setti, S. E., Hunsberger, H. C., & Reed, M. N. (2017). Alterations in hippocampal activity and alzheimer's disease. *Translational issues in psychological science*, *3*(4), 348.

Šimić, G., Kostović, I., Winblad, B., & Bogdanović, N. (1997). Volume and number of neurons of the human hippocampal formation in normal aging and alzheimer's disease. *Journal of Comparative Neurology*, *379*(4), 482–494.

Smith, T. D., Adams, M. M., Gallagher, M., Morrison, J. H., & Rapp, P. R. (2000). Circuit-specific alterations in hippocampal synaptophysin immunoreactivity predict spatial learning impairment in aged rats. *Journal of Neuroscience*, *20*(17), 6587–6593.

Stewart, T. C. (2012). A technical overview of the neural engineering framework. *University of Waterloo*, *110*.

Thomas, E. A., Reid, C. A., Berkovic, S. F., & Petrou, S. (2009). Prediction by modeling that epilepsy may be caused by very small functional changes in ion channels. *Archives of neurology*, *66*(10), 1225–1232.

Trujillo, O. (2014). *A spiking neural model of episodic memory encoding and replay in hippocampus* (Unpublished master's thesis). University of Waterloo.

Vida, I., Degro, C. E., & Booker, S. A. (2019). Connectivity of the hippocampus. In V. Cutsuridis, B. P. Graham, S. Cobb, & I. Vida (Eds.), *Hippocampal microcircuits: a computational modeler's resource book* (pp. 29–90). Springer.

West, M. J., & Gundersen, H. (1990). Unbiased stereological estimation of the number of neurons in the human hippocampus. *Journal of Comparative Neurology*, *296*(1), 1–22.

Witter, M. P. (2019). Connectivity of the hippoocampus. In V. Cutsuridis, B. P. Graham, S. Cobb, & I. Vida (Eds.), *Hippocampal microcircuits: a computational modeler's resource book* (pp. 5–28). Springer.

World Health Organization. (2021). *WHO dementia fact sheet.* Retrieved 2021-11-18, from `https://www.who.int/news-room/fact-sheets/detail/dementia`

Yassa, M. A., & Stark, C. E. (2011). Pattern separation in the hippocampus. *Trends in neurosciences*, *34*(10), 515–525.

Appendix A

The effect of weight scaling on pattern separation and the need for activity matching.

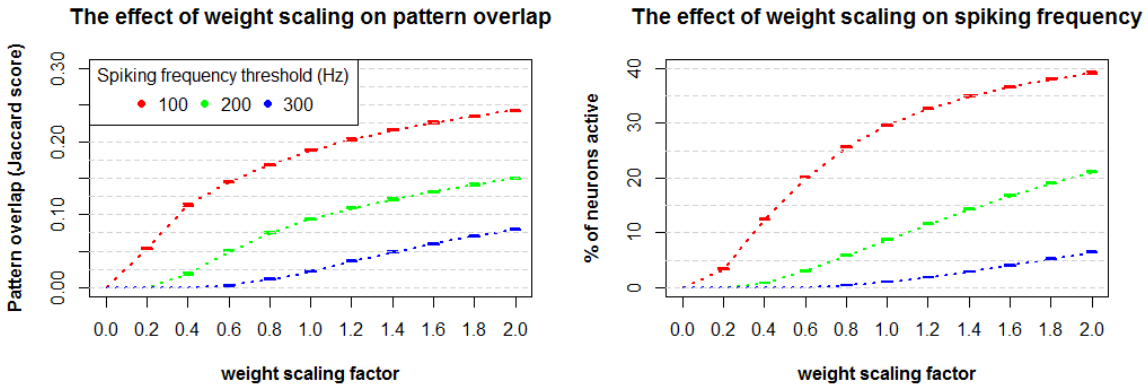

Figure A.1: **The effect of weight scaling on pattern overlap in CA3.** Application of a linear mapping over the connection between the dentate gyrus (DG) and regions CA3 results in linear scaling of the connection weights. For a given spiking frequency threshold it seems that pattern overlap (left) and the fraction of active neurons (right) strongly correlate.

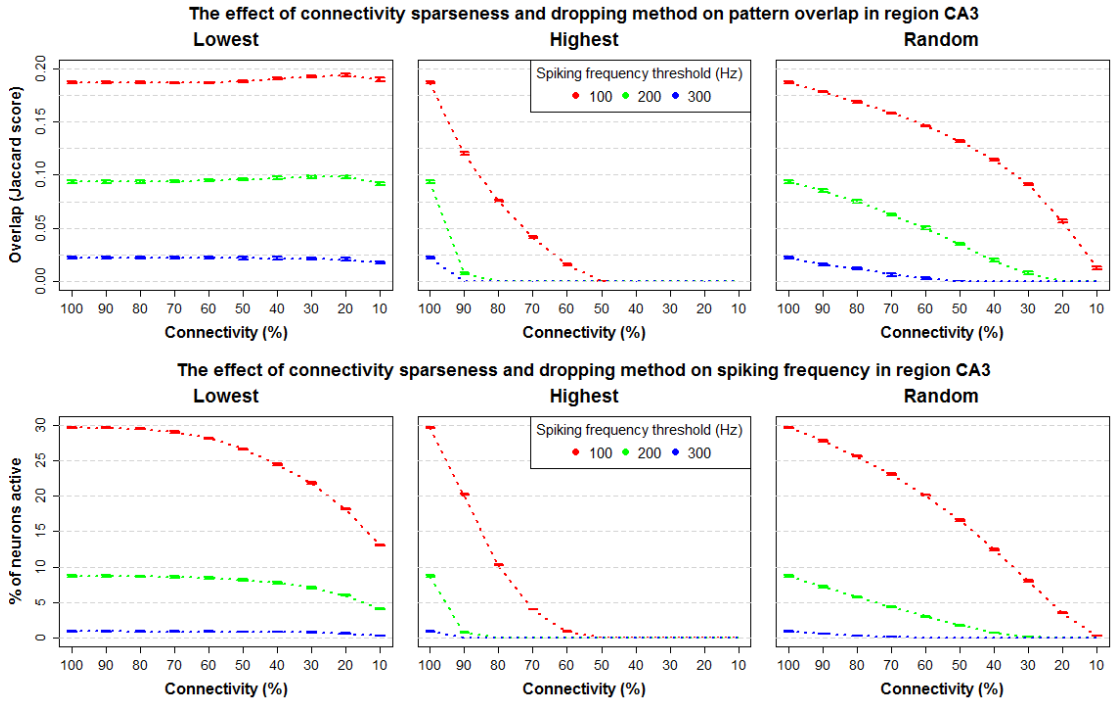

Figure A.2: **The effect of sparse connectivity at the DG-CA3 connection on pattern overlap in CA3. (Top)** As connections become sparse, pattern overlap decreases when the highest weights are dropped, or when they are dropped at random. From these figures, it becomes apparent, that the 10% strongest connections exert a disproportional amount of influence on the pattern separation value. **(bottom)** For any of the dropping methods, neural activity decreases as the connections become sparser. Because of this it is hard to say whether an observed reduction in pattern overlap is due to the effect of sparseness itself, or the general decrease in neural activity.

Appendix B

Complete results for the AD experiment

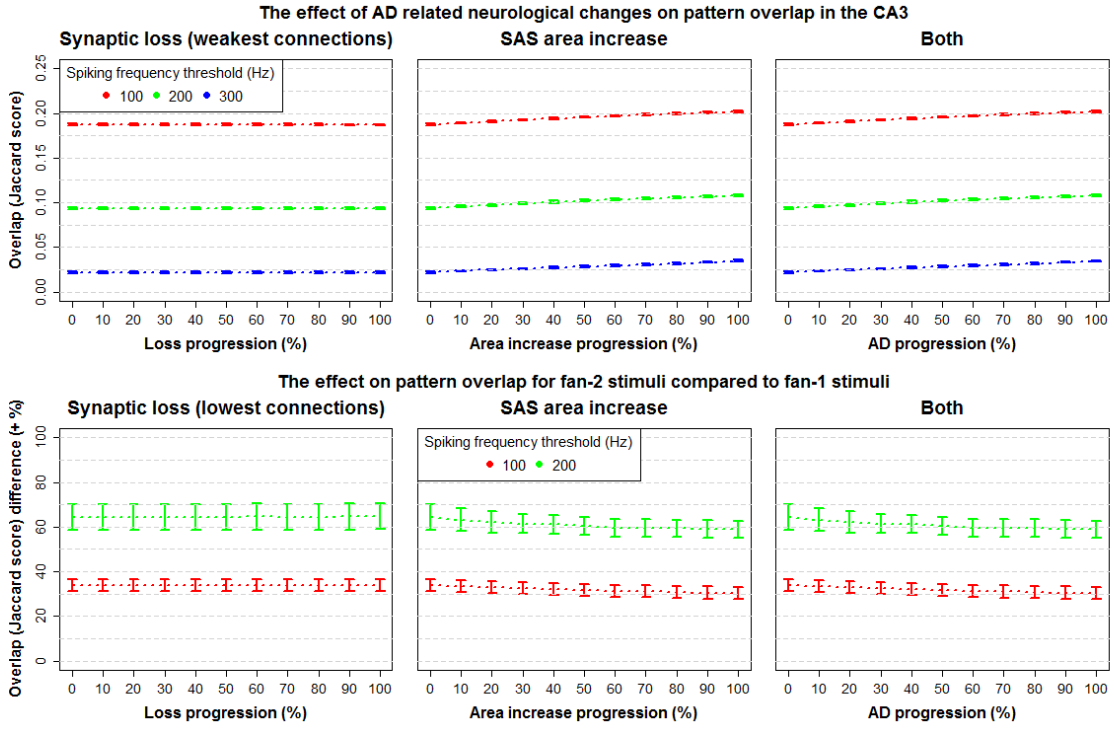

Figure B.1: **The effect of AD related changes to the EC-DG connection on pattern overlap in CA3.** Synaptic loss affects the weakest connections between the EC and the DG. At full progression, 41% of synapses are lost. The SAS area increase is implemented as a scaling of the connection weights between the EC and the DG. At full progression, weights were scaled to 137% of their original value

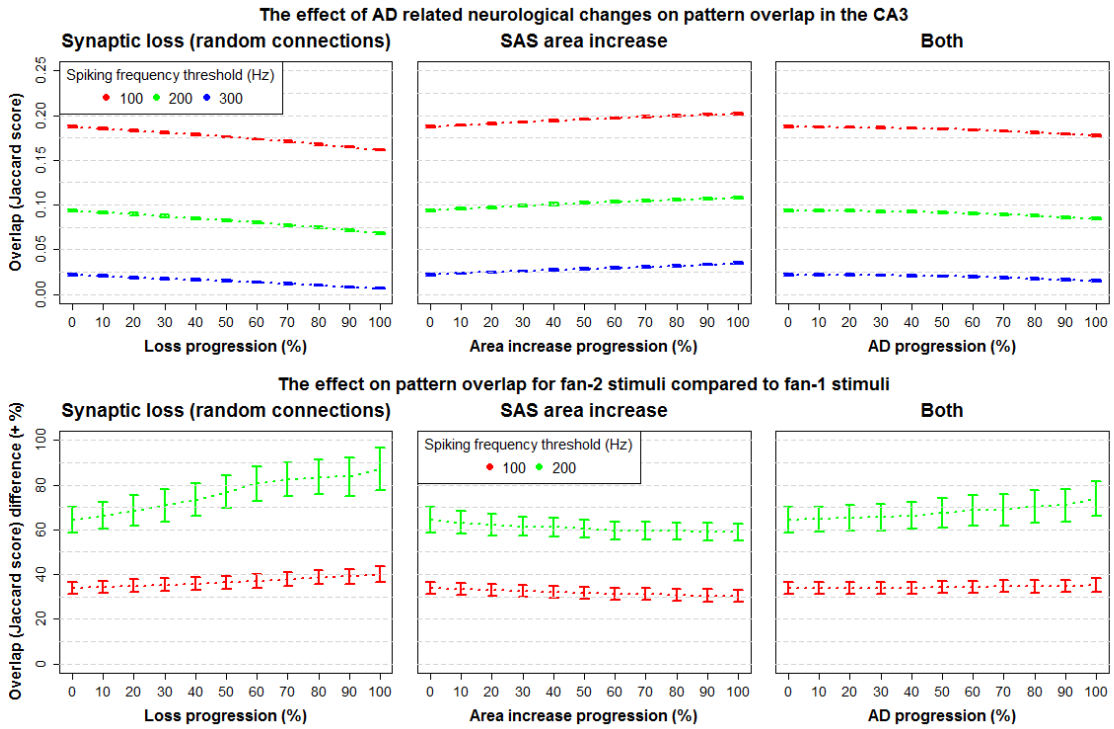

Figure B.2: **The effect of AD related changes to the EC-DG connection on pattern overlap in CA3.** Synaptic loss affects the connections between the EC and the DG at random. At full progression, 41% of synapses are lost. The SAS area increase is implemented as a scaling of the connection weights between the EC and the DG. At full progression, weights were scaled to 137% of their original value

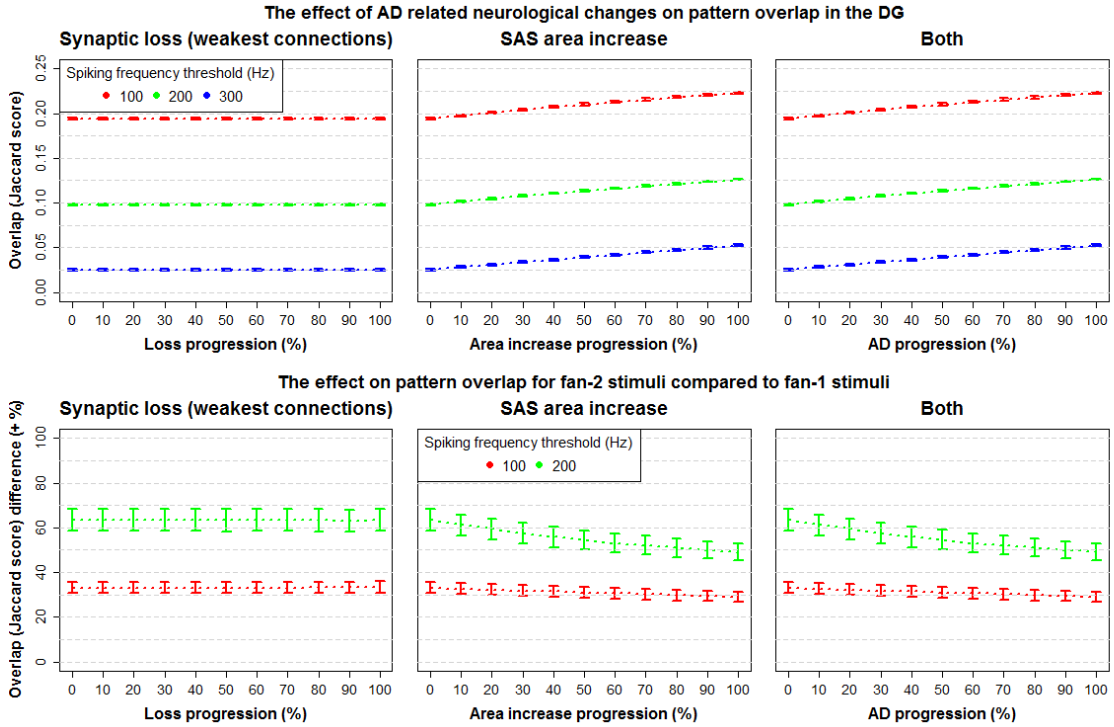

Figure B.3: **The effect of AD related changes to the EC-DG connection on pattern overlap in the DG.** Synaptic loss affects the weakest connections between the EC and the DG. At full progression, 41% of synapses are lost. The SAS area increase is implemented as a scaling of the connection weights between the EC and the DG. At full progression, weights were scaled to 137% of their original value

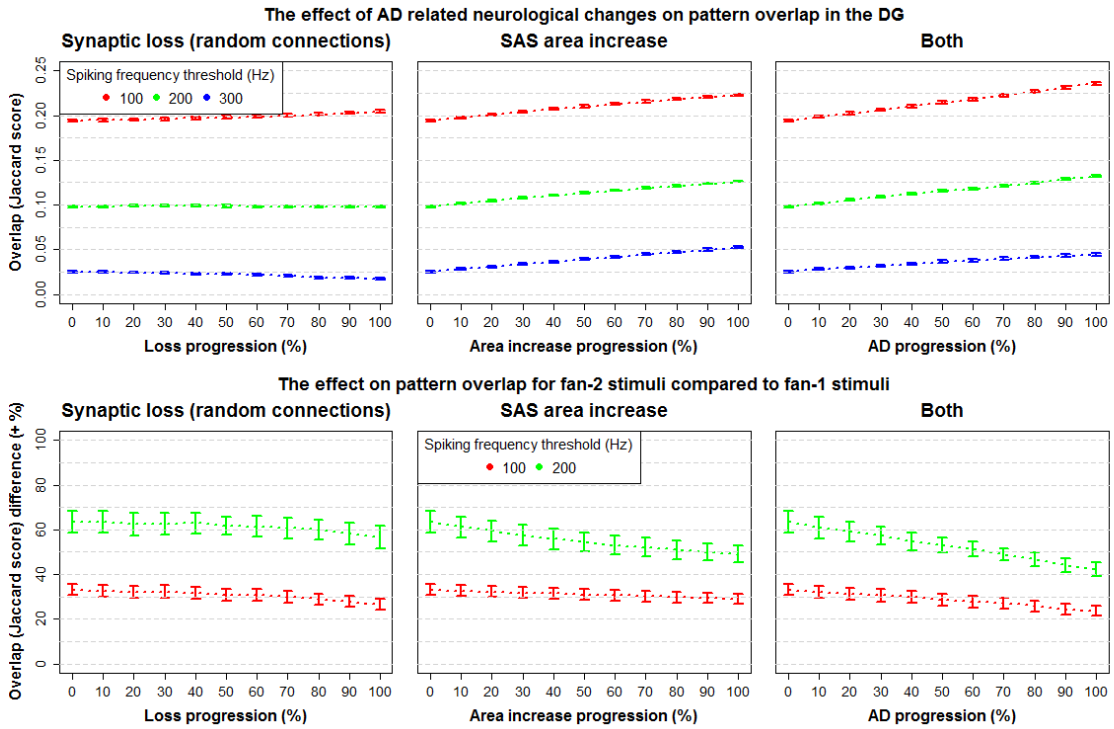

Figure B.4: **The effect of AD related changes to the EC-DG connection on pattern overlap in the DG.** Synaptic loss affects the connections between the EC and the DG at random. At full progression, 41% of synapses are lost. The SAS area increase is implemented as a scaling of the connection weights between the EC and the DG. At full progression, weights were scaled to 137% of their original value

Appendix C

Hippocampal and DG connectivity

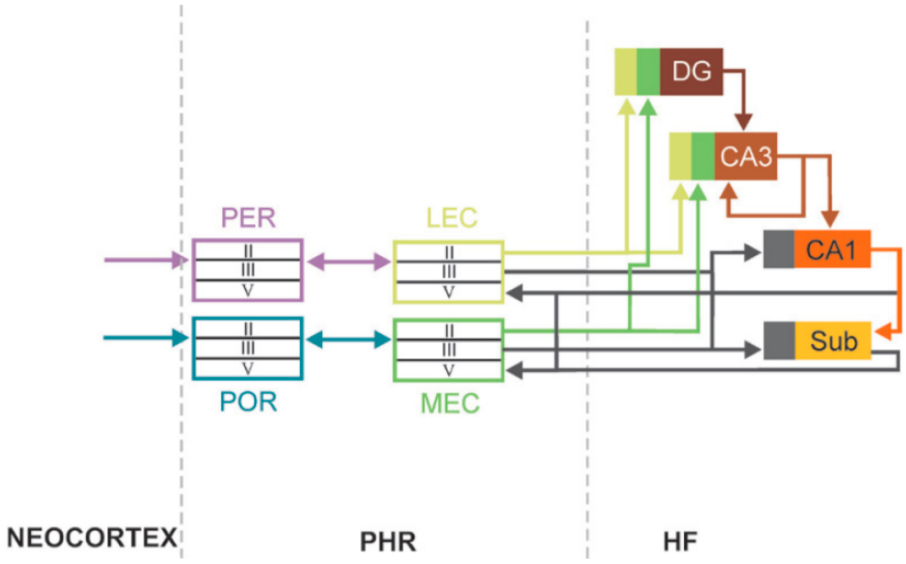

Figure C.1: **Extensive schematic of the connectivity between the various components of the entorhinal-hippocampal network.** Figure from Witter (2019)

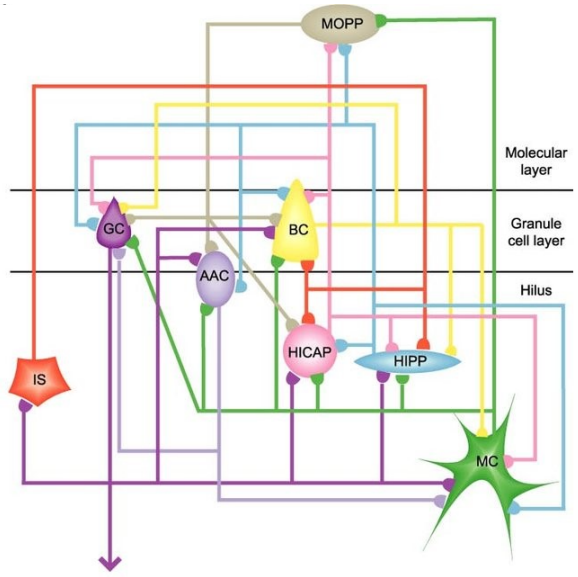

Figure C.2: **Schematic of the inter connectivity of the different neuron types in the dentate gyrus**. Figure from Thomas et al. (2009)

Appendix D

Neuron firing rates for the activity matching experiment

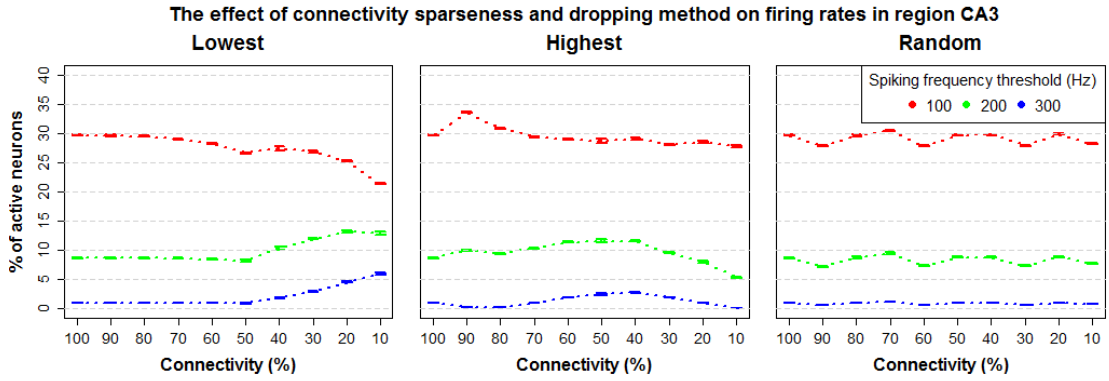

Figure D.1: **The effect of activity-matched sparse connectivity at the DG-CA3 connection on firing rates in CA3.**