THE UNIVERSITY
of EDINBURGH

Self-Organised Criticality via Retro-Synaptic Signals in Complex Neural Networks

Victor Hernandez-Urbina

Doctor of Philosophy

Institute of Perception, Action and Behaviour

School of Informatics

University of Edinburgh

2016

Abstract

The brain is a complex system *par excellence*. Its intricate structure has become clearer recently, and it has been reported that it shares some properties common to complex networks, such as the small-world property, the presence of hubs, and assortative mixing, among others. These properties provide the brain with a robust architecture appropriate for efficient information transmission across different brain regions. Nevertheless, how these topological properties emerge in neural networks is still an open question.

Moreover, in the last decade the observation of neuronal avalanches in neocortical circuits suggested the presence of self-organised criticality in neural systems. The occurrence of this kind of dynamics implies several benefits to neural computation. However, the mechanisms that give rise to critical behaviour in these systems, and how they interact with other neuronal processes such as synaptic plasticity are not fully understood.

In this thesis, we study self-organised criticality and neural systems in the context of complex networks. Our work differs from other similar approaches by stressing the importance of analysing the influence of hubs, high clustering coefficients, and synaptic plasticity into the collective dynamics of the system. Additionally, we introduce a metric that we call *node success* to assess the effectiveness of a spike in terms of its capacity to trigger cascading behaviour. We present a synaptic plasticity rule based on this metric, which enables the system to reach the critical state of its collective dynamics without the need to fine-tune any control parameter. Our results suggest that retro-synaptic signals could be responsible for the emergence of self-organised criticality in brain networks. Furthermore, based on the measure of node success, we find what kind of topology allows nodes to be more successful at triggering cascades of activity. Our study comprises four different scenarios: *i)* static synapses, *ii)* dynamic synapses under spike-timing-dependent plasticity (STDP), *iii)* dynamic synapses under node-success-driven plasticity (NSDP), and *iv)* dynamic synapses under both NSDP and STDP mechanisms. We observe that small-world structures emerge when critical dynamics are combined with STDP mechanisms in a particular type of topology. Moreover, we go beyond simple spike pairs of STDP, and implement spike triplets to assess their influence on the dynamics of the system. To the best of our knowledge this is the first study that implements this version of STDP in the context of critical dynamics in complex networks.

Acknowledgements

This thesis is the culmination of four years of PhD studies at the University of Edinburgh. As such, it marks the end of a personal stage, and the beginning of another. However, I didn't get to this point all by myself, and here I would briefly like to express my gratitude to some people that directly or indirectly helped me through my studies.

I would like to thank my family for all their material, immaterial and unconditional support through all these years. I'm in eternal debt with you! I would also like to thank Marina Papagianni for her love, her care and support over the last years of my PhD, but most importantly for her endless patience. Also, I would like to thank Tom Underwood for the insightful and motivating conversations that eventually became a part of this thesis, and also for his contagious optimism.

I am very grateful to the members of my supervision panel: Dr. Michael Herrmann, Dr. Matthias Hennig, and Dr. Vincent Danos. Many thanks for their advice, their feedback and recommendations to my work. In particular, I would like to thank Dr. Herrmann for his guidance, his motivation and his patience, but also for introducing me to the exciting world of *self-organised criticality*.

I would like to thank the examiners of my thesis, Dr. Peggy Series and Dr. Stefan Bornholdt, for their kind support and feedback.

Last but not least, I would like to thank the Mexican National Council on Science and Technology (CONACYT) fellowship no. 214055 for funding my research.

(JVHU. September, 2015)

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

(*Victor Hernandez-Urbina*)

To Adam Smith.
To the pioneers.

Table of Contents

Chapter 1

Introduction

"Essentially, all modeling of brain function from studying models of neural networks has ignored the self-organized aspects of the process, but has concentrated on designing a working brain by engineering all the connections of inputs and outputs."

—Per Bak

The thesis that the reader has in her hands lies at the intersection of complex networks, neural systems, and self-organised criticality. It is motivated by the idea that we can understand more about the brain machinery by studying it under the perspective of complex systems and network theory.

The following two sections of this Introduction have the purpose of providing the reader with a loose exposition of the ideas and concepts that motivated our research, as well as situating her in the context in which this thesis resides. Afterwards, in Sect. 1.3 we will provide an outline of the contents of this work.

1.1 Networks of the Brain

A healthy brain is a well-connected brain. Intelligence requires the capacity to process information in an efficient and robust manner. Thus, the brain should exhibit an architecture that enables optimal information transmission across its many regions, as well as stability and robustness in its communication. Additionally, such an architecture would have to provide the necessary mechanisms to allow for structural modifications that are essential for memory, learning, and recovery when its tissue is damaged. Neuronal plasticity refers to these structural changes that brain tissue undergo throughout its existence, and it might imply modulatory mechanisms at the level of individual

synapses, or the establishment or removal of existent connections among nerve cells. Brain structures are tightly intertwined with the dynamics occurring within them in such a way that structure affects collective dynamics, and vice versa; thus creating a loop that is responsible for the rich repertoire of functions that neural systems perform.

In recent years, the importance of studying the brain and neural systems under the perspective of graph and network theory has been stated, and research on this direction has been carried out with considerable success. The discovery of structural and functional connections among brain areas is an important, prolific and promising field of research in neuroscience. Charts of brain connectivity, from their structural or functional perspective, can be constructed nowadays thanks to the capacity to collect, store and analyse large brain databases. Neuroscience has thus entered the age of *big data*.

Brain networks can be described as graphs that are composed of nodes representing neural elements such as single neurons or brain regions which are connected by links denoting either physical or functional connectivity (Bullmore and Sporns, 2009). Here, the direction of the edges denotes the flow of information tranmission across the network. The *Human Connectome Project* sponsored by the National Institutes of Health in the United States is a project whose goal is to build a map of the whole brain connectivity in the same spirit as the *genome* project had the objective of mapping the entire human genome. Similar efforts are being carried out in Europe with the *Human Brain Project*, and in China with the *Brainnetome Project*. These ventures reflect the current restlessness in trying to understand the dynamics of the brain through understanding its structural properties.

Recent results in neuroimaging seem to imply that the brain's structural and functional systems exhibit features of complex networks, such as a small-world topology, modularity and highly connected nodes suggesting power-law degree distributions (Sporns et al., 2004; Sporns, 2010). The study of brain networks under the perspective of network theory is devoted to analyse particular network statistics that might reveal the mechanisms of particular brain functions. However, which network measures are most appropriate for the analysis of brain networks is still an open question (Rubinov and Sporns, 2010).

The brain exhibits features commonly associated to complex networks. Recently, it has been reported the existence of a *rich-club* organisation of the human brain, which is characterised by the presence of neocortical hubs that have a tendency to connect to other brain hubs rather than to poorly connected regions (van den Heuvel and Sporns,

2011)[1]. These brain hubs play a key role in the integration and propagation of global information giving rise to a backbone for global communication (van den Heuvel et al., 2012). Moreover, these observations have been applied to the study of brain pathologies such as Alzheimer's disease and schizophrenia, suggesting that these conditions imply abnormalities in brain topology (Bullmore and Sporns, 2009). Researchers have found that these diseases are associated with alterations in brain structure. For instance, they have found that such conditions are associated with a reduction of the small-world property in specific brain regions along with loss in hierarchical organization (Bullmore and Sporns, 2009).

Experimental evidence in this direction seem to lead to the same conclusion, that the architecture of the brain is not random (Song et al., 2005b). And although when we are born the brain is densely connected, its architecture evolves towards a sparse topology through the programmed death of cells that do not imply any benefit to brain function. Moreover, the brain is expensive in terms of the costs of neural tissue and metabolic consumption. Therefore, its dynamics and architecture must be the result of an economic trade-off between the network's physical cost and the adaptive value of its structure which translates into functional connectivity (Bullmore and Sporns, 2012). Can genetics account for the structural changes that the brain undergoes throughout its existence or it might be that a combination of genetic and activity-dependent mechanisms are responsible of shaping the brain's complex architecture? In this thesis, we implement two of the main properties commonly associated with complex networks, namely, the small-world property and the scale-invariance of degree distributions, to our models; and describe under what circumstances a densely connected network gives rise to a small-world structure under activity-dependent plasticity.

1.2 Phase Transitions in Brain Dynamics

Structure leads to function, for instance, the small-world property provides the structural requirements to allow synchronisation in brain networks (Yu et al., 2008). The brain is a very powerful computing machine which is performing complex calculations all the time, from the most abstract ones required for mathematical thinking, to the most habitual ones such as walking down a staircase without looking at every step. For these complex computations to take place, the brain must implement dynamics that take out the best from its efficient structure.

[1] In network theory this is known as assortative mixing.

Recently, it has been put forward the idea that brain computations are carried out at the *edge of chaos* (Beggs, 2008). This means that the dynamics of brain networks are undergoing a phase transition similar to those that occur in inorganic matter when a parameter is finely-tuned (e.g. liquid water turning into vapor). A typical signature of a phase transition is the power-law distribution of event sizes, which implies that at the critical point the dynamics of a system do not exhibit a particular scale. Power-law distributions in the activity of brain tissue have been detected in a modality of neuronal activity previously unknown whose *modus operandi* is avalanches of bursts of activity. These *neuronal avalanches* (Beggs, 2007) were first detected in cortical and acute slices of rat cortex (Beggs and Plenz, 2003, 2004); and since their first observation they have spawned subsequent research in different experimental settings which seem to suggest that these neuronal avalanches are the main modality of the brain at rest (Petermann et al., 2009; Klaus et al., 2011; Barbieri and Shimono, 2012; Friedman et al., 2012; Bellay et al., 2015; Shew et al., 2015). Model implementations of neuronal avalanches have elucidated the potential benefits of neuronal networks teetering at the edge of chaos. It is in this boundary between order and randomness, known as the critical point, that brain networks reach maximum computational power (Bertschinger and Natschläger, 2004; Kinouchi and Copelli, 2006; Haldeman and Beggs, 2005).

However, the idea of critical brain dynamics has not been received without some skepticism (Touboul and Destexhe, 2010), which include the unjustified enthusiasm regarding the ubiquity of power-law distributions in nature (Clauset et al., 2009; Stumpf and Porter, 2012). Moreover, it has been reported that brain dynamics are not precisely in the critical regime but instead they are wandering about it (Priesemann et al., 2014). Nevertheless, experimental evidence of the scale-invariant nature of neuronal avalanches in brain networks either *in vivo* or *in vitro* keeps piling up, implying that this is an active and promising field of research.

Another question regarding this modality of brain function refers to its mechanism of emergence. It is known from the theory of critical phenomena that the control parameter of systems that exhibit critical behaviour needs to be fine-tuned in order to give rise to this particular type of dynamics. However, under certain circumstances some systems are able to reach this state without the need of external tuning, which implies that the system implements some sort of feedback mechanism that gives it control over its parameters. Self-organised criticality (SOC) was proposed as a theory to explain the underlying mechanisms of this kind of systems, and at the same time provide explanations to the apparent ubiquity of $1/f$ fluctuations and power-law distributions of

events that are observed in phenomena as diverse as the dynamics of plate tectonics, solar flares, stock market crashes, mass extinctions, and forest fires, among others (Bak, 1997). Therefore, the concept of SOC has been put forward as a mechanism by which complexity arises in nature.

Self-organised critical models of neuronal networks have been developed to explain the emergence of power-law distribution of events and other observables that would imply that these networks are at the edge of a phase transition. These models can be divided in two categories, those that explain the emergence of critical behaviour through plasticity mechanisms such as spike-timing-dependent plasticity (Shin and Kim, 2006; Meisel and Gross, 2009) or short-term plasticity (Levina et al., 2007) or another Hebbian-like plasticity process (de Arcangelis et al., 2006); and those that explain it through non-plastic mechanisms like axonal re-wiring (Bornholdt and Rohlf, 2000; Bornholdt and Röhl, 2003) or the balance between neuronal excitation and inhibition (Beggs and Plenz, 2003). In this thesis we put forward yet another model of self-organised critical neuronal avalanches in neuronal networks. Our approach considers the introduction of retro-synaptic signals that inform the neuron about the behaviour at the level of its local surroundings.

As mentioned before, a healthy brain is a well connected brain, and this results in optimal neuronal information processing. Recently, it has been pointed out that critical dynamics are a signature of healthy brain systems (Massobrio et al., 2015). This hypothesis rests on the idea that a critical brain can show the fastest and most flexible adaptation to environments where unpredictability abounds. In such environment, we not only require fast, robust and reliable cognitive functions, but also sensory systems that can adapt quickly to changes in the environment, and provide responses that maximise the diversity of stimuli found in the world (Chialvo, 2006; Tagliazucchi and Chialvo, 2011). Moreover, it has been proposed that one of the purposes of sleep is to tune the brain for criticality (Pearlmutter and Houghton, 2009). The suggestion that criticality is a signature of a healthy brain implies a novel perspective to study brain disease whose effects result in a deviation from critical (or close-to critical) dynamics. Still, the relationship between criticality and cognition is an open question.

1.3 This Thesis

Simply put, this thesis is about self-organised criticality in complex networks, and how neuronal plasticity affects or benefits such a collective state. As stated in this chapter,

the motivation behind our research comes from complexity science, self-organisation, critical dynamics and the belief that network theory can help us to understand the most complex piece of machinery in the universe (or at least in this world), namely, the human brain.

Below we present an outline of what we will present in the following chapters.

- **Chapter 2** In this chapter we present the basic concepts that will be needed to understand the contents of this thesis. We divide these in two main categories, namely, topological notions and dynamical notions. The first refers to concepts from network science that will be recurrent in the thesis, whereas the latter refers to those concepts that capture the dynamical nature of the nodes in the networks.

- **Chapter 3** For the purposes of this thesis, we have generated networks that fit in very particular classes for the sake of carrying out a comparative study regarding the evolution of their collective dynamics under our model. This chapter is devoted to the statistical properties of these networks.

- **Chapter 4** In this chapter the question posed is *How do complex topological properties such as the small-world property and the presence of hubs affect the onset of criticality?* We observe that the presence of hubs in a sparse network gives rise to an extended region of parameter space in which the system reaches the critical state. Also, the small-world property boosts neural activity in complex networks. Moreover, we introduce a metric to assess the success of a spike in terms of its capacity to trigger cascading behaviour within the network. In this chapter, we present the most "successful network structure". By such an expression, we mean the network structure in which the spikes of nodes are more successful. Some of the contents of this chapter are also presented in a paper entitled *The success of complex networks at criticality*[2] and it is the result of a collaboration with Dr. Tom Underwood.

- **Chapter 5** In this chapter the question posed is *What happens to the network if we introduce spike-timing-dependent plasticity (STDP) mechanisms once the system is at criticality?* The answer is: we observe a small-world structure emerging from fully-connected networks, and at the same time the critical regime vanishes. The next question is *Do we observe the same behaviour when considering triplets of spikes instead of simple pairs of spikes in the plasticity rule?*

[2]Currently in pre-print: http://arxiv.org/abs/1507.07884

The answer is no, and still the critical regime vanishes. Some of the contents of this chapter are part of a paper entitled *Small-world structure induced by spike-timing-dependent plasticity in networks with critical dynamics*[3].

- **Chapter 6** This is the chapter that gives the name to this thesis. Here, we present our main contribution, which is that retro-synaptic feedback could in principle serve as a mechanism to achieve self-organised criticality in neural systems. This is shown in a model that implements a synaptic plasticity rule whose main ingredient is the metric introduced in Chapter 2 in order to assess the success of a spike. This plasticity rule goes under the name of *node-success-driven plasticity* (NSDP). Next, we test the model under the same STDP mechanisms used in the previous chapter, and observe that the critical state does not vanish this time. This implies that NSDP serves as a compensatory mechanism for the synaptic modulation induced by STDP. Such a compensatory mechanism makes the system stay in the critical state.

- **Chapter 7** In this chapter we briefly review the main results of this thesis, wrap everything up, and present the conclusions of this work. Some directions of future research are also presented.

- **Appendix A** This appendix contains the description of the algorithm used to generate the scale-free networks which were used extensively along this thesis.

[3]Currently in pre-print: http://arxiv.org/abs/1507.07879

Chapter 2

Preliminaries

In this chapter, we present the basic notions that are required to understand the concepts shown in the following chapters. As will be recurrent in this thesis, we divide this chapter in two sections, one devoted to topological notions, that is, those concerned to network structure, and a second one devoted to dynamical notions, that is, those concepts in the context of the collective behaviour that emerges from nodes accommodated in a particular network structure.

2.1 Topological Notions

2.1.1 Complex Networks

"What do metabolic pathways and ecosystems, the Internet, and propagation of HIV infection have in common? Until a few years ago, the answer would have been very little."

—Amaral and Ottino (2004)

Networks are ubiquitous: almost every process in nature or human activity can be modelled by a network and the interaction among its elements; they represent an extension from the mathematical notion of graphs, which were introduced in the 18th century by Leonhard Euler to solve the Königsberg bridge problem (Boccaletti et al., 2006).

Networks can occur physically as in electric power grids, highways or airline routes, the Internet; or they can occur as abstract entities when modelling a particular phenomenon in nature, like the network of acquaintances among individuals.

Networks gain the property of *complex* when considered in the context of *complex systems*. A complex system is loosely described as a system composed of rela-

tively simple components with no central control, in which an emergent behaviour is observed. By "simple components" it is meant that the individual components comprising the system are simple with respect to the behaviour achieved collectively. For instance, a single neuron is a complex entity by itself, however its functional role in the context of the whole brain is relatively simple as compared with the behaviour of the entire system (Mitchell, 2006). Emergence is the other main property of a complex system, as such it entails the description of a behaviour that emerges from the interaction of the components of the system, and whose properties quite frequently cannot be anticipated by the sole observation of the individual components.

A complex network is a structure that captures the interaction among dynamical units, whose number is so large that a graphical representation of such interaction becomes cumbersome. Unlike a graph, a complex network is a structure in which a process takes place and most of the time, this process affects the connectivity of the elements responsible for the process; creating, thus, a feedback loop. Thus, we should not only consider the dynamics *on* the network, but also the dynamics *of* the network (Gross and Blasius, 2008; Gross and Sayama, 2009). The notion of emergence in a complex network is implied by its structure. Quite often the structure of a complex network emerges as a result from a self-organizing local process of the elements comprising the network rather than from a design planned beforehand. For instance, the particular shape of the Internet emerges as a result of the process of optimizing the distribution of packages of information through routers providing service to a local group of computers.

Complex networks can be found everywhere and in diverse contexts. Examples of them include the chemical interactions inside the living cell, the Internet at the domain level or at the autonomous systems level, the spread of epidemics, the World Wide Web (WWW), air traffic and other transport systems, the brain, and even the interactions among transnational corporations that exert an important influence in the world economy, among many others (Albert and Barabási, 2002; Sporns, 2010; Vitali et al., 2011). Moreover, in recent years we have witnessed the rise in popularity of online platforms for social networking such as Facebook, LinkedIn and Twitter, among others.

The growing interest for studying complex networks can be explained by the fact that many real-world networks, either natural or manufactured, share many features in common such as a long-tailed degree distribution, correlations of node degrees, and the so-called small-world property, among others (Newman, 2003). Likewise, this

interest might be explained by the relative ease nowadays to collect and analyze large amounts of data given by a drop in the costs for computing and storage, paired up with an increment in computing power.

In general, a complex network is a network that possesses the following properties:

- *Feedback interaction.* The dynamics of the nodes have an effect on the structure of the network, which in turn affect the dynamics of nodes once again, and the process is repeated.

- *Structural emergence.* The network emerges as a result from a self-organising process occuring at the (local) level of nodes. There is no global control making decisions regarding the architecture of the network.

- *Behavioural emergence.* The dynamics of the nodes give rise to a collective behaviour that cannot be explained solely by the action of individual nodes. A property commonly summarized by the expression: *the whole is more than the sum of the parts*.

Some of the questions that today are trying to be answered by network scientists when thinking about real-world networks are (Mitchell, 2006):

- What network statistics can be used to characterise networks? That is, can we rely in a small set of network measures (e.g. degree distributions, clustering coefficients, characteristic path lengths, motifs, etc.) to capture the essence of network structure? Or are these statistics a mere shadow of the real object?

- What properties do real-world networks share despite the fact that they arise in very dissimilar contexts, and how? As mentioned above, some real-world networks share many properties in common despite the fact that they originate in so dissimilar contexts. Is this fact a result from a universal process or physical law yet to be discovered that shapes the connectivity among agents in a given context?

- How do we design efficient algorithms to determine such properties? That is, now that we have answered the first question satisfactorily, how can we design optimal algorithms to estimate such metrics in networks whose size increase constantly?

- How do such properties affect the dynamics *on* and *of* the network? That is, how do the metrics discovered by answering the first question affect the collective dynamics and in turn how this behaviour reshapes the network structure?

In this thesis we are concerned in how a set of network statistics, namely the average clustering coefficient, the average shortest path length among others, affect a type of collective behaviour namely the critical regime, and in turn how this behaviour affects network topology. For this purpose we are concerned as well in developing efficient algorithms to estimate such network properties.

2.1.2 Basic Properties of Networks

As mentioned earlier, networks are an extension of the mathematical notion of graphs. Like graphs, they are entities composed of a set of nodes and edges. A *node* is the fundamental unit of a network, they are also known as *vertices*. An *edge* is a line connecting two nodes, it also goes by the name of link or bond in other contexts. The *degree* of a node is simply the number of edges connected to a node.

A network is *directed* if all of its edges are directed, which means that edges in the network run in a single direction. We denote the direction of the edge by an arrow. On the other hand, a network is *undirected* if all of their edges run in both directions. For directed networks, a node has an *in-degree* describing the number of edges incoming to the node, as well as an *out-degree* describing the number of edges outgoing from the node.

The *degree distribution* of a network describes the frequency of all the different node degrees found in that network. In other words the degree distribution $P(k)$ of a network specifies the probability that a node taken uniformly at random from the network will possess degree k. Thus, directed networks have two degree distributions, one describing the *out-degree distribution* and another describing the *in-degree distribution*, which are not necessarily equal. One can also consider the joint distribution $P(j,k)$ representing the fraction of nodes that simultaneously have in-degree j and out-degree k. It has been observed that some real-world networks exhibit correlations between in- and out-degrees (Newman, 2003).

The *shortest path length* between nodes i and j is defined as the shortest path through the network that have nodes i and j as extremes. Let D be the matrix containing all the shortest path lengths for network G, that is, the entry d_{ij} from matrix D denotes the shortest path from node i to node j. The maximum value in D denotes the *diameter*

of the network. The average path length, also known as *characteristic path length*, is defined as the mean of the shortest paths over all couple of nodes (Boccaletti et al., 2006):

$$L = \frac{1}{N(N-1)} \sum_{i,j \in N, i \neq j} d_{ij} \tag{2.1}$$

If there are disconnected components in the network, then $L = \infty$. In the case of directed networks the direction of edges matter, and for nodes i and j not necessarily occurs that $d_{ij} = d_{ji}$. Moreover, for directed networks it might be the case that the network is connected (i.e. there are no disconnected components) but still a node might not be reachable. In such case, Eq. (2.1) diverges as well. The networks considered in this thesis are networks for which the characteristic path length is well defined, that is, every node in the network is reachable and every node is capable of spike transmission. Formally, for every node i, both in-degree and out-degree are larger than zero.

Clustering in the context of network theory refers to the presence of clusters of interconnected nodes. This property is commonly observed in acquaintance networks and it reflects a transitivity behaviour occurring within the networks. Transitivity refers to the fact that if node x is connected to node y, and node y is connected to node z, then there is a high probability that node x and z are connected as well. As a consequence, the network will exhibit a large number of triangles. Watts and Strogatz (1998) introduced the local clustering coefficient c_i of a given node i to measure the probability that the neighbours of i are themselves connected. The value of the local clustering coefficient of node i is given by counting the actual number of edges (denoted by e_i) in G_i, the subgraph of neighbours of i. The local clustering coefficient is defined as:

$$c_i = \frac{2e_i}{k_i(k_i-1)} \tag{2.2}$$

where $k_i(k_i-1)/2$ is the maximum number of possible edges in G_i. The average clustering coefficient of a network G is simply the average of local clustering coefficients of the nodes in the network:

$$C = \frac{1}{N} \sum_{i \in N} c_i \tag{2.3}$$

For directed networks, e_i stands for the total amount of edges regardless of direction found in G_i, whereas the maximum number of possible edges is given by $k_i(k_i-1)$, i.e. without dividing by 2 as for undirected networks in Eq. (2.2).

2.1.3 Random Networks

The most common model of random networks is the one proposed and studied rigorously by mathematicians Paul Erdös and Alfréd Rényi. Although illuminating, this model is inadequate to describe important features found in real-world networks, and so it has been extended in various ways (Newman, 2003).

There are two closely related variants of the Erdös and Rényi (ER) random network model:

- $R(n,p)$, in which we have n nodes and connect each pair with probability p.
- $R(n,e)$, in which a network is chosen uniformly at random from the set of all networks having n nodes and e edges.

Most of the results in one model can be derived for the other model in a straightforward fashion. In this thesis, unless stated otherwise, when we refer to a random network we refer to the $R(n,e)$ version of the ER model.

In a random network the degree distribution follows a Poisson distribution (Newman, 2010). However, it has been observed that the degree distribution of real-world networks differ from that of the random network. The degrees of nodes in most real-world networks are highly right-skewed providing the degree distribution with a long right tail, whose values are far from the mean degree (Newman, 2003). Moreover, it has been observed that random networks fail to replicate the high mean clustering coefficient (see Sect. 2.1.4) observed commonly in social networks.

2.1.4 Small-World Property

It has been found that the essence of real-world networks cannot be captured by the random network model introduced by Erdös and Rényi nor by regular structures such as lattices. The model by Watts and Strogatz (1998) was proposed to describe a class of networks that lie halfway between randomness and regularity. This class of networks are characterized by a small average shortest path length (a feature observed in random networks) and an average clustering coefficient significantly larger than expected by chance (a feature observed in regular lattices). Taken together, these properties offer a structural benefit to the processes taking place within the network, such as optimal information transmission that results from speeding up the communication among otherwise distant nodes. A term that summarizes the presence of these properties is that of the *small-world property*.

Interestingly, this class of networks fits better the features observed in real-world networks where very often connectivity among nodes obey a transitivity rule (that is, if a is connected to b and b is to c, then it is very likely that a is connected to c as well), and where long-range connections between nodes help building short-cuts in the network. Networks that exhibit the small-world property are so diverse and can be found in social, technological and biological contexts, to name a few (Humphries and Gurney, 2008). For instance, in their seminal paper, Watts and Strogatz applied their analysis to the electrical power network of western United States, the co-stardom network comprising movie actors and their collaborations, and the neuronal network of the nematode worm *C. elegans* and in all three cases the authors observed a small average shortest path length and a large average clustering coefficient (Watts and Strogatz, 1998).

Based on the ideas proposed by Humphries and Gurney (2008), we estimate the degree of small-world-ness in a network through the following process. Let G be a network consisting of n nodes and e edges. To test whether G exhibits the small-world property we construct a random network R with same number of nodes and same number of edges. Then, we estimate the mean clustering coefficient of both networks C_G and C_R along with their mean path length L_G and L_R. Finally, we compute S as the ratio of those values:

$$S = \frac{C_G/C_R}{L_G/L_R} \tag{2.4}$$

If $S > 1$, then G possesses the small-world property (Humphries and Gurney, 2008), which implies a more abundant presence of cliques and long-range connections among nodes than expected by chance.

2.1.5 Scale-Free Networks

When dealing with a network whose design was not conceived by anyone, but on the contrary, emerged as a result of a process of self-organisation (think of the Internet, or the biochemical interactions of the cell, for instance) a standard procedure for the analysis of such a structure starts by inspecting the distribution of the degrees of the nodes in the network.

Scale-invariance in the distribution of the node degrees of a network is a phenomenon observed in real-world networks. This scale-invariance is summarised by a power-law approximation of the degree distribution of the network under study, in which the probability $P(k)$ that a node connects to other k nodes decays as a power-

law following $P(k) \sim k^{-\beta}$. Networks that exhibit this particular feature are known as *scale-free networks*, given that their degree distribution does not exhibit a particular scale.

The observation that the degree distribution of a sample from the WWW can be approximated by a power law motivated Barabási and Albert to study the mechanisms that underlie the emergence of power law behaviour in real-world networks (Barabasi, 2002). These authors proposed a model to explain such emergence, the so-called *BA model* (Barabási and Albert, 1999). This model incorporates two features observed often in real-world networks, namely, network growth and preferential attachment, which refers to the preference of new nodes to connect to existing nodes that possess a large number of connections; a phenomenon appropriately summarised by the expression: *rich get richer*. Moreover, a long-tail in the degree distribution of a network implies the existence of many poorly connected nodes coexisting with very few but not negligible massively connected nodes known as *hubs* (see below). This observation is also related to the empirical $80/20$ rule proposed by the economist Vilfrido Pareto and know nowadays as the *Pareto principle*. In the context of network theory, the $80/20$ rule says that 20% of nodes possess 80% of the edges, and this is roughly what occurs in scale-free networks.

The BA model fails to reproduce the large mean clustering coefficient usually observed in real-world scale-free networks, and thus extensions to the model have been suggested in which an extra step is performed to improve clustering among nodes (Holme and Kim, 2002). Moreover, the BA model is based in global information of the network statistics, that is, it requires that at every time step the incoming nodes possess information regarding the degree of every node already in the network in order to choose the best-connected node to connect to. Obviously, this is unrealistic. If the BA model was actually a model to explain the power-law nature of the degree distribution of the WWW, then this implies that new websites have the information regarding the connectivity of every other website existing in the WWW. This is impossible. Power-law behaviour arises as a result of the local interactions of nodes comprising the network rather than from global information of the system. Extensions to the BA model based on local (rather than global) information have been already proposed (Caravelli et al., 2013).

Lastly, scale-invariant degree distribution and the small-world property are by no means exclusive and in fact many scale-invariant networks are also small-world networks.

2.1.6 Hubs

Hubs are loosely defined as *well-connected* nodes, that is, nodes that possess many connections compared to other nodes in the network. Thus, according to this definition the property of being a hub requires knowledge of the degree distribution of the network in order to know what is the degree of every node comprising the network. This would make of the notion of "hubness" of a node a *global* property in the sense that we require global information to assess this property for a given node.

We use the above definition intuitively when observing users in online social networks. For instance, if we observe a user in Twitter whose number of followers is large, we could not tell whether the user is a hub or not unless we compare his number of followers to that of other users in the network. It is then, after comparing the number of followers of different users, when we get an idea of the degree of "hubness" of an user.

Intuitively, hubs exist only in networks where there are *poorly-connected* nodes. That is, in networks in which hubs coexist with poorly-connected nodes. (Here, we will refer to this poorly-connected nodes as *dwarves*.) In this sense, fully-connected networks do not possess hubs although every node in them is *massively* connected to the rest of the nodes. But, this would imply that random networks following the Erdös-Renyi model could in principle feature the presence of hubs. This type of network exhibits a Poisson distribution with a well defined mean and variance. In contrast, scale-free networks, which motivated the study of hubs, exhibit a power-law in the degree distribution. In this type of distributions, the mean and variance are not always defined, as for different values of the exponent β larger moments of the distribution diverge (Newman, 2010). However, as we are dealing with a finite system, mean and variance exist; the problem is that they convey no appropriate information regarding the statistical properties of the nodes in the network.

A working definition for the notion of hub that we propose involves the mean degree of a given network, and provides a formal framework to categorize the hubness of a node. In this framework, it will be clear that fully-connected networks do not possess hubs, and as a network becomes less and less homogeneous, hubs start to emerge. We define a hub as a node whose degree (either in- or out-degree, see below) is greater or equal than the mean plus *two* times the standard deviation of the degree distribution of the network. The choice of two times the standard deviation is somehow arbitrary, but helps formalize the fact that fully-connected networks or homogeneous topologies in

general (e.g. lattices) do not possess hubs. If a network exhibited a Gaussian degree distribution, then 95% of the degrees of the nodes would fall below two standard deviations from the mean. Everything greater than that can be safely categorized as a hub. A more strict definition would require that the degree of a hub is greater or equal than the mean degree plus *three* standard deviations. But we do not use such constraint in our definition.

2.1.7 Broadcasting Hubs and Absorbing Hubs

Scale-free networks are characterized by a power-law approximation of their degree distribution (Barabási and Albert, 1999). In the case of directed networks, there are two degree distributions, one corresponding to the out-degrees of nodes and another one for the in-degree of nodes.

In general, real-world networks are directed, and very often both their degree distributions can be approximated by a power-law, e.g. the WWW, or at least one of them, e.g. citation networks (Newman, 2003). In either case, the presence of a long-tail in the *out-degree* distribution of a network implies the existence of *broadcasting hubs*, that is, nodes that have massive outgoing connections compared to other nodes in the system. On the contrary, the presence of a long-tail in the *in-degree distribution* implies the existence of *absorbing hubs*, that is, nodes that have massive incoming connections compared to any other nodes.

Here, we are interested in analyzing how collective dynamics develop for the case of networks with broadcasting hubs and for networks with absorbing hubs. In the following, scale-free networks with absorbing hubs will be labeled as *in-degree scale-free networks*, whereas those that contain broadcasting hubs will be termed *out-degree scale-free networks*.

2.1.8 Network Motifs

A more detailed picture of the network structure can be obtained by considering the relative frequency of particular configurations of small blocks in comparison to random networks (Milo et al., 2002). In particular, 3-node subgraphs, known as *motifs*, describe the relationship among nodes taken in threes. In Fig. 2.1 we show the thirteen different 3-node configurations for directed networks.

The identification of the motif configuration of a given network helps us construct the *motif profile* which is the distribution of the thirteen different 3-node configurations

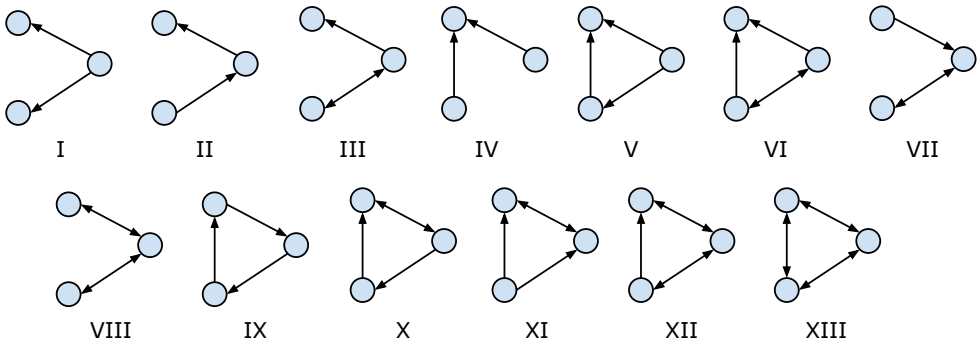


Figure 2.1: All configurations for 3-node connected subgraphs, representing the possible relations among nodes taken in threes from a directed network. Motifs with bidirectional connections are: *III*, *VI*, *VII*, *VIII*, *X*, *XI*, *XII* and *XIII*. The most basic representation of "hubness" is captured by motifs *I* (broadcasting hubs) and *IV* (absorbing hubs); whereas the most basic representation of clustering, in terms of triads of nodes, is captured by motif *V*.

possible for a directed network. Song et al. (2005b) observed that the motif distribution of acute slices taken from the visual cortex of rats exhibits an overrepresentation of some 3-node patterns when compared to random networks. This would imply that connectivity in brain networks is far from random.

Moreover, networks from a similar context share similarities in their motif profiles, so that these 3-node subgraphs might even define broad classes, or super-families, of networks (Milo et al., 2002, 2004). Similar motif distributions could point to similar dynamical processes in the network, e.g. food networks represent the flow of energy from bottom to top of the food web resulting in a particular motif configuration, whereas brain networks represent a flow of information without necessarily implying a particular direction of the flow, which results in a different motif profile (Milo et al., 2002).

2.2 Dynamical Notions

Having described the topological concepts that will become essential in the rest of this thesis, let us turn our attention now to notions regarding the behaviour that emerges inside the network due to the interaction among the elements comprising it.

2.2.1 Self-Organised Criticality: the influence of the neighbours

"How do we know that the creations of worlds
are not determined by falling grains of sand?"

—*Victor Hugo*, Les Miserables

The concept of *self-organized criticality* (SOC) was proposed first as a mechanism to explain the apparent ubiquity of $1/f$ noise in phenomena of very diverse nature, and thus as a mechanism to explain the rise of complexity in nature (Bak et al., 1988).

The concept of SOC, has been suggested to explain the dynamics of phenomena as diverse as plate tectonics (Gutenberg and Richter, 1956), piles of granular matter (Frette et al., 1996), forest fires (Bak et al., 1990), neuronal avalanches (Eurich et al., 2002) (see below), and mass extinctions (Bak, 1997), among several others including crashes in the stock market (Johansen and Sornette, 1998). This concept implies the existence of a critical point that becomes an attractor in the collective dynamics of a system. Such a critical point (or regime) denotes a state of the system in which the collective dynamics are undergoing a phase transition. As such, it represents the boundary between two different states of the system (e.g. order and randomness) and it is identified by the presence of power laws in the distribution of events, the divergence of the correlation length, among others (Bak et al., 1988). Thus, the expression *critical dynamics* refers to a state of a system in which its collective dynamics exhibit the features of SOC including the existence of clusters of activation, whose distribution can be approximated by a power-law. Given the appropriate conditions, a system comprising numerous interconnected elements interacting non-linearly can give rise to power-law behaviour, which relates the collective dynamics of such a system to the notion of phase transitions, divergence of correlation lengths, and $1/f^{\alpha}$ noise.

The study of phase transitions is generally situated within the realm of statistical physics and involves the analysis of a system under the influence of a control parameter, which when appropriately tuned gives rise to a phase transition. As mentioned earlier, a phase transition marks the boundary between two different regimes of the a system, for instance, a phase transition occurs when water goes from liquid (order) to gas (disorder), in which temperature is the control parameter of the system. The boundary between the qualitatively different regimes of the system is reached at a *critical* point of the control parameter (Kadanoff, 2009). Under certain conditions, a system might be able to self-organize in order to reach and maintain the critical point without the need of external supervision. These systems stay at the border of stability and chaos. Such systems and their properties are the study of SOC.

The paradigmatic model of SOC, the sand-pile model, was proposed by Bak and his collaborators to show how a system that tunes itself towards the critical point exhibits as well $1/f^{\alpha}$ noise (Bak et al., 1988). The model consists of grains of sands entering the system (a sandbox) at locations chosen at random. Eventually, piles of sand grow until they reach a (critical) point in which an addition of a single grain of sand will make the pile topple and spread the sand to neighbouring sites. This in turn has the possibility of trigger a domino effect in the neighbouring sites and start an avalanche of sand grains until the system reaches a stable configuration, that allows it to continue the accumulation of sand that enters the system externally.

The model exhibits sand avalanches of all sizes whose distribution can be approximated by a power-law. Moreover, in the model, the correlation lengths diverge implying that a toppling grain of sand at one side of the system can influence the behaviour of a grain of sand at the other side of the system. Also, a power-law distribution of avalanche durations (lifetimes) is observed, which is equivalent to a power-law frequency spectrum of the form $1/f^{\alpha}$ (Bak et al., 1988). Therefore, the sand-pile model with its simple local rules giving rise to collective scale-invariant behaviour serves as an appropriate example of complexity.

Some of the suggested requirements for a system to exhibit SOC are (Jensen, 1998):

- Threshold elements. The system consists of numerous components. These components can be thought of accumulators who gather energy (or any other physical quantity) and release it when going beyond a given threshold. One effect of the threshold is the emergence of metastable states in the system, which in other contexts is known as *punctuated equilibria*.

- Network structure. The elements of the system are arranged in such a way that their interactions are specified formally by a network topology. This also corresponds to the identification of SOC systems as *interaction-dominated threshold systems*. This means that the many degrees of freedom in the system are interacting; and the dynamics of the system are dominated by such interactions

- External driving. Energy enters the system externally and is accumulated by the components of the system. As well, the external driving can take the form of noise in the elements of the system.

- Separation between time scales of external driving and relaxation dynamics. The

process responsible for the driving of the system must be much slower than that of the internal relaxation. A clear example of this is the dynamics of earthquakes: the stress in the earth's crust builds up on the scale of years due to the movement of the tectonic plates, however the accumulated stress is released in a matter of seconds as an earthquake.

2.2.2 Neuronal Avalanches

"It now seems that you have dug yourself into a hole from which you cannot escape. If power laws are so unexceptional, then why should I be so excited about seeing them in neural data."

—*Mnemo* from (Beggs and Timme, 2012)

Already Bak pointed out a link between SOC and the dynamics of the human brain. He suggested that the brain teeters on the edge of a phase transition, hovering between order and disorder. As well, a mechanism by which SOC could be incorporated in the context of brain networks was suggested before it was observed in real brain tissue (Eurich et al., 2002).

However, it is not that the brain is undergoing a material or structural phase transition as when water turns into ice or as when a giant component emerges in a sparse random network (Strogatz, 2001; Dorogovtsev et al., 2008). Rather, the phase transition taking place is dynamical in the sense of the collective dynamics of the system. The regimes reached by such a dynamical phase transition are related to information processing, and how neuronal activity propagates through a neuronal network.

Recall that the sand-pile model consists of grains of sand being propagated via avalanches when the piles reach a certain threshold given by the slope of the pile. In a neuronal network, an action potential is propagated to neighbouring cells everytime the neuron surpasses its threshold. Thus, in one regime of the system the propagation of an action potential ceases almost as soon as it is emitted; this regime is called *subcritical*. Here, avalanches that span through the whole system are practically non-existent and only small avalanches abound. Another regime corresponds to neuronal activity that is amplified pathologically at every time step; this corresponds to the *supercritical* regime. Here, large avalanches are very frequent and can be triggered at any time. In between these two regimes lies the *critical* regime in which an action potential is propagated in a sustained manner. Here, small avalanches are very frequent but large avalanches that span to the entire system exist too. It is in this regime that the

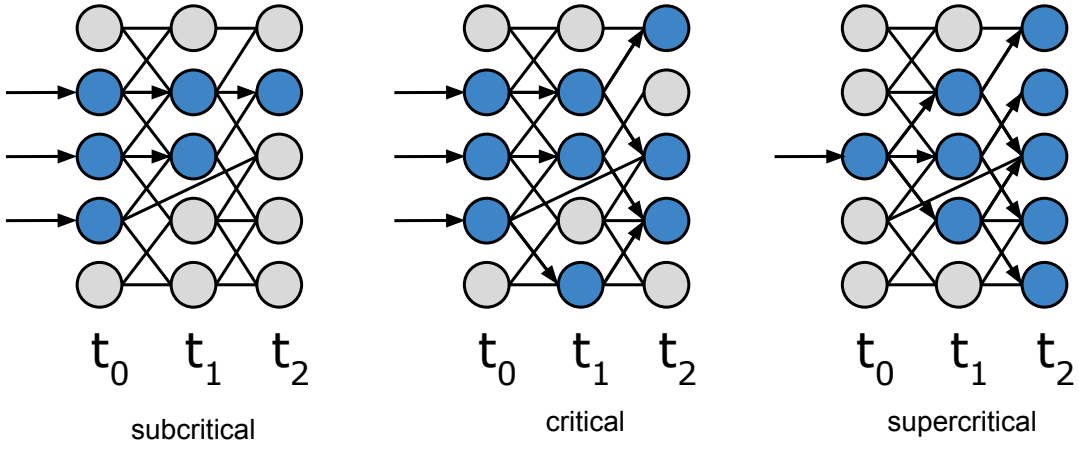


Figure 2.2: Dynamical regimes of a system exhibiting neuronal avalanches. Adapted from (Chialvo, 2006). The subcritical regime features activity that dies out quickly and large avalanches are non-existent. The supercritical regime, on the other hand, exhibits activity that becomes amplified without control. In between these two regimes there is the critical regime in which activity is sustained in a healthy manner and action potentials are reproduced in almost the same numbers at each time step until finite size effects take place.

distribution of avalanche sizes (among other observables) can be approximated by a power-law. Fig. 2.2 shows an example of these three regimes.

The experimental evidence for critical avalanches in brain networks came a few years after SOC was suggested as another operational mode of neural systems along with oscillations and synchronisation. Avalanches of action potentials were observed as the result of spontaneous activity in local field potentials of cultured slices of rat cortex (Beggs and Plenz, 2003), and later in the superficial cortical layers of awake, resting primates (Petermann et al., 2009). As the name suggests, neuronal avalanches are bursts of activity that resemble a domino effect triggered by spiking in groups of neurons. The observed avalanches are stable and repeatable spatiotemporal patterns of activity (Beggs and Plenz, 2004), which might relate them to memory mechanisms inside the brain.

It has been observed that the distribution of avalanche sizes as well as the distribution of their durations can be approximated by a power law with respective exponents of $\gamma = -3/2$ and $\delta = -2$, both in experiments (Beggs and Plenz, 2003) and models; as well as scaling relationships among system sizes and exponents (Eurich et al., 2002; Levina et al., 2007).

Beggs and Plenz (2003) mapped the observed neuronal avalanches to a *branching process* in order to derive an expression for the critical exponent of the distribution of avalanches sizes and their duration. Branching processes were introduced initially in the 19th century to explain the disappearance of British family names. Since then, they have been used extensively in mathematics, biology and physics. Particularly, they have been used as a mean-field approximation to model the dynamics of the system during avalanches (Levina, 2008). A branching process describes the development of a system in which the dynamics are based on the offspring produced by the components. Such an offspring takes the form of subsequent activation of units once an avalanche has been triggered. As such, it requires the definition of the branching ratio σ which is defined as the ratio of descendants that become active at time $t+1$ to ancestors that were active at time t. The quantity σ has been used to characterize the critical state of a system (Beggs and Plenz, 2003), and to identify the regimes surrounding such a state. When $\sigma < 1$ the system is subcritical, whereas when this value is above unity the system is supercritical. In between these two states lies the critical state which corresponds to $\sigma = 1$. In Sect. 2.2.6, we will present a statistic similar to the branching ratio that will be used to develop a plasticity learning rule capable of reaching SOC.

Power-law behaviour in the dynamics of neuronal avalanches relate this biological process to the notion of SOC described above, in which the critical regime is reached without external tuning, correlation lengths diverge, and the system is at the boundary of a (dynamical, rather than structural) phase transition. Critical dynamics of brain networks have been studied thoroughly in artificial models, and it has been found that the critical regime, that is, when the system is in the border of stability and disorder, implies several computational benefits for the system, to name a few:

- Optimal information transmission and maximum dynamic range (Kinouchi and Copelli, 2006).
- Maximum information storage (Haldeman and Beggs, 2005; Uhlig et al., 2013).
- Stability of information transmission (Bertschinger and Natschläger, 2004).

Hence the *criticality hypothesis* for brain dynamics, which states that neural networks operate at the *edge of chaos*, that is, at the critical point in a phase transition between total randomness and boring order (Beggs, 2008). Recently, it has been suggested that criticality might play a role in neuronal networks *in vivo* during focused attentional states in particular it has been observed that the *default mode network* exhibits

critical behaviour (Linkenkaer-Hansen et al., 2001). As well, neuronal avalanches have been observed as a result of irregular spiking of piramidal neurons of layer $2/3$ of rat neocortex during the transition from the anaesthetised to the awake state, suggesting that the convergence to the critical state is a signature of brains becoming awake (Bellay et al., 2015). Neuronal avalanches have also been reported in the visual cortex during adaptation to strong visual stimuli (Shew et al., 2015). These results seem to confirm the hypothesis that our senses are tuned to criticality (Chialvo, 2006).

2.2.3 Eurich model

The Eurich model was suggested as a mathematical model of spiking neurons in a fully-connected topology that gives rise to critical behaviour when the synaptic weights among nodes are appropriately tuned (Eurich et al., 2002). Therefore it is not an example of SOC, but of critical dynamics identified by the presence of power-law distributed avalanches and their durations.

Synapses in the model are static (that is, non-plastic), connectivity is all-to-all, and neurons are all excitatory; these attributes make the model simple enough to be studied analytically. The model consists of N non-leaky integrate-and-fire nodes. Each node j is characterized by a continuous variable known as the membrane potential h_j, which is updated in discrete time according to the equation:

$$h_j(t+1) = h_j(t) + \sum_{i=1}^{N} A_{ij} w_{ij} s_i(t) + I_{ext} \tag{2.5}$$

where A denotes the adjacency matrix with entries $A_{ij} = 1$ if node i sends and edge to node j, and $A_{ij} = 0$ otherwise; w_{ij} denotes the synaptic weight from node i to node j; $s_i(t) \in \{1,0\}$ represents the state of node i (active or quiescent, respectively) at time t; and I_{ext} denotes external input which is supplied to a node depending on the state of the system at time t. This mechanism of external driving works as follows: if there is no activity at time t, then a node is chosen uniformly at random and its membrane potential is increased by a fixed amount through the variable I_{ext}. If $h_i(t)$ exceeds the threshold θ, then node i emits a spike, which changes the state of this node to active ($s_i(t) = 1$) and propagates its activity through its synaptic output. Afterwards, the node is reset, ie. $h_i(t+1) = 0$.

The coupling strength w_{ij} for every node i sending an edge to node j is set according to the equation:

$$w_{ij} = \frac{\alpha}{\langle e \rangle} \tag{2.6}$$

where $\alpha \in (0,1]$ is the control parameter of the model, and $\langle e \rangle$ denotes the mean degree of the network.

The model resembles the sand-pile model described in Sect. 2.2.1 in the way that activity is supplied externally to the system and the propagation of activity once a site of the system goes beyond threshold. Here, avalanches take the form of neuronal activity being propagated as a domino effect. While the system is relaxing, external drive stops. This guarantees that relaxation time occurs at a different time scale as the external driving. This corresponds to an infinite separation of time scales between external driving and relaxation dynamics that has been suggested as a necessary condition for SOC (Jensen, 1998; Levina, 2008).

For a particular interval of the control parameter α the sizes of the neuronal avalanches, (identified as the number of nodes that become active in between two quiescent periods) and their durations (identified as the number of time steps in between two quiescent periods) can be approximated by a power-law. In analytical examinations, the exponents derived for such distributions are $\gamma = -3/2$ and $\delta = -2$ for avalanches sizes and durations respectively, in the thermodynamic limit of fully-connected networks (Eurich et al., 2002; Levina et al., 2007).

The Eurich model is interesting because it predicted accurately the critical exponent found by Beggs and Plenz in real brain tissue (Beggs and Plenz, 2003). In order to become analytically solvable, the Eurich model had to make strong assumptions which are not suitable for realistic neural modelling. For instance, the model is based on fully-connected networks, which do not occur naturally in real-world networks (see Sect. 2.1.1) and it only considers static synapses. Therefore, in this thesis we will extend the model by considering not only fully-connected networks but also heterogeneous topologies, such as random and scale-free networks with varying degree of the small-world property, which makes them fit perfectly into the context of complex networks.

2.2.4 Spike-Timing-Dependent Plasticity

An important area of research in network science is focused on studying the mechanisms by which nodes connect and disconnect during the development of a network. As mentioned before, complex networks often possess feedback mechanisms by which

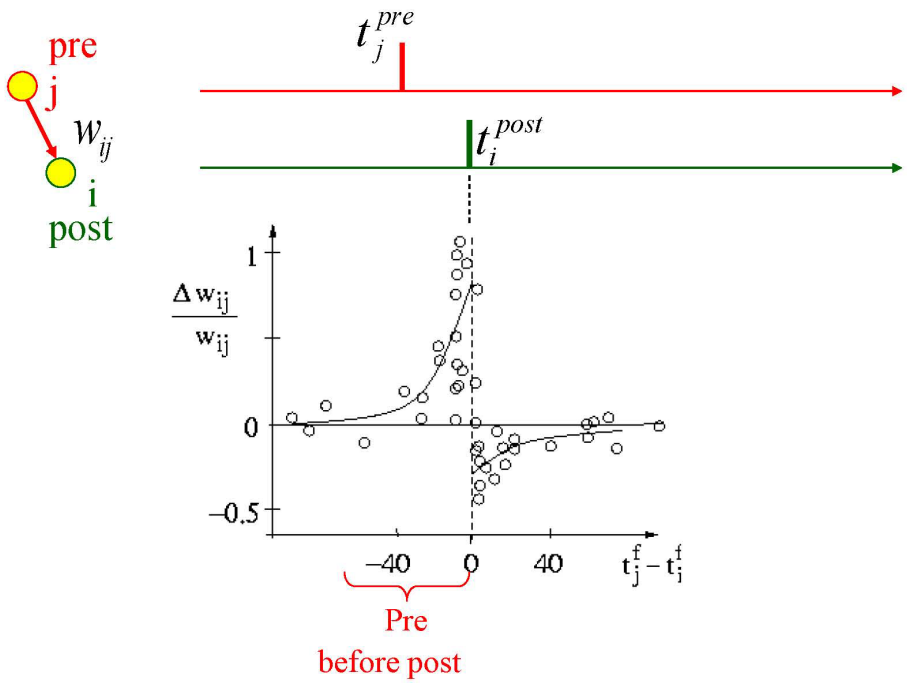


Figure 2.3: Spike-timing-dependent plasticity (STDP) is an asymmetric learning rule in brain networks that captures the temporal correlations between pre- and post-synaptic neurons. In this figure, the x axis represents the time difference between spikes of pre- and post-synaptic units, whereas the y axis corresponds to the synaptic weight update. Note the asymmetry between weight update for potentiation and depression. Image taken from (Sjostrom and Gerstner, 2010).

the node dynamics and their interactions affect the structure of the network, which in turn alters the behaviour of the nodes, etc. It is thus useful to consider both the dynamics *on* networks as well as the dynamics *of* networks.

In brain networks, synaptic plasticity refers to the structural changes that neuronal networks undergo through the strengthening or weakening of synaptic connections in response to the in-going activity. In particular, *spike-timing dependent plasticity* (STDP) captures the existing temporal correlations between the spikes of pre- and post-synaptic units resulting in a temporally asymmetric learning rule that has been proposed as a mechanism for learning and memory in the brain (Bi and Poo, 1998). STDP is a *local* rule responsible of long-term synaptic modulation that emphasises the precise timing of each individual spike, it also incorporates and extends the essential mechanism of Hebbian learning by including the notion of long-term potentiation (LTP) and long-term depression (LTD) of synapses based on the differences in activa-

tion times for pre- and post-synaptic neurons (Bi and Poo, 2001). Figure 2.3 shows the experimental observations of the STDP mechanism at work. Synaptic modulation is in function of the temporal difference between spike emission of pre- and post-synaptic neurons. Potentiation is achieved when the pre-synaptic node fires shortly before the post-synaptic one. Depression is achieved when the opposite occurs, namely, the post-synaptic neuron fires shortly before the pre-synaptic unit.

We have just described the simplest and perhaps the most popular version of STDP which involves only a couple of spikes: one from the presynaptic and another one from the post-synaptic neuron. As mentioned above, the time difference between these spikes determines the amount of synaptic modulation between the two neurons. This basic STDP mechanism is known as *pair-wise* STDP (pSTDP). However, it has been reported that pSTDP fails to replicate observations in experimental data. Intuitively, a pre-post pairing followed by a post-pre pairing of the same magnitude would cancel out any synaptic modulation triggered, however this is not what it has been observed in experiments. Pair-wise models of STDP fail to account for the dependence on the repetition frequency of the spike pairs. In experiments, pSTDP is not able to explain the synaptic modulation triggered by higher-order plasticity rules such as triplets and quadruplets of spikes (Pfister and Gerstner, 2006; Clopath and Gerstner, 2010). For this reason, pSTDP has been extended in order to consider *triplets* of spikes rather than just pairs of them. Here, we will refer to this triplet-spike version of STDP as tSDTP in order to differenciate it from pSTDP.

A triplet rule for tSTDP involves sets of three spikes: two pre- and one post-synaptic, or one pre- and two post-synaptic spikes. With a triplet rule of this form it is possible to fit experimental data from visual cortical slices as well as from hippocampal cultures. Interestingly, when this rule is based on Poisson spike trains, the learning rule can be mapped to a Bienestock-Cooper-Munro (BCM) learning rule (Bienenstock et al., 1982). BCM has been suggested as a model of learning in the visual cortex and the formation of receptive fields. Moreover, it has been proposed that the triplet rule for STDP is a mechanism used by neuronal networks to perform ICA-like computations (Gjorgjieva et al., 2011).

As well, tSTDP involves not only the temporal correlations between pre- and post-synaptic units but also the spike rate of neurons, which makes it a more complete learning rule than pSTDP. It has been pointed out that timing dependence and frequency dependence of synaptic plasticity interact (Clopath and Gerstner, 2010). Early experiments showed that the amount of plasticity as a result from pre-post pairing does

not depend exclusively on the delay between pre- and post-synaptic spikes but also on the frequency at which such pairings are repeated.

When we consider triplets of spikes for STDP, we might feel inclined to extend the model in order to consider higher-order STDP rules. However, experiments show that spike triplets are able to reproduce data generated by higher-order terms (Pfister and Gerstner, 2006).

2.2.5 Computational Implementations of STDP

As mentioned in Sect. 2.2.4, STDP is a temporally asymmetric form of Hebbian learning induced by temporal correlations between pre- and postsynaptic neurons. Synaptic weight between pre- and postsynaptic nodes is potentiated (increased), if the postsynaptic neuron fires shortly *after* the presynaptic neuron. It is depressed (decreased) if the opposite happens, namely the post-synaptic neuron fires shortly *before* the presynaptic neuron.

We implemented pair-wise STDP (pSTDP) mechanisms in our simulations through the following set of equations:

$$w_{ij}(t+1) = w_{ij}(t) + \Delta w_{ij}(\Delta t) \tag{2.7}$$

where $\Delta t = t_{post} - t_{pre}$ denotes the difference between the spiking times of pre- and post-synaptic neurons, and:

$$\Delta w_{ij}(\Delta t) = \begin{cases} a_p \exp\left\{\frac{-\Delta t}{T_p}\right\} & \text{if } \Delta t \geq 0 \\ -a_d \exp\left\{\frac{\Delta t}{T_d}\right\} & \text{if } \Delta t < 0 \end{cases}$$

where parameters a_p and T_p set the amount and duration of LTP, whereas a_d and T_d set the amount and duration of LTD. In our experiments we set $a_p = a_d = 0.1$. Observations of STDP in brain tissue suggest that the time window for potentiation is typically shorter than the depression time window (Bi and Poo, 1998), for that reason we let $T_p = 10$ time steps and $T_d = 20$ time steps. However, it also has been observed that time-windows and amount of potentiation/depression vary across species and brain structures (Bi and Poo, 2001). We cannot expect that a single theoretical model of synaptic plasticity can account for all experimental facts (Morrison et al., 2008).

We impose hard bounds on synaptic weights, that is, $0 < w_{min} < w_{ij} < w_{max}\ \forall i, j$, which prevents unbounded weight growth, gives rise to strong competition between inputs to a neuron and results in a bimodal distribution of the synaptic weights at the

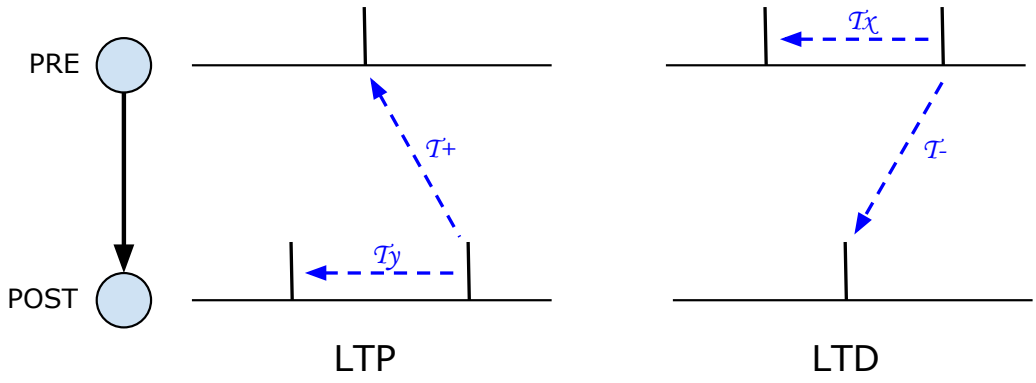


Figure 2.4: Mechanism of tSTDP. Potentiation is achieved when the pre-synaptic unit fires in between two post-synaptic spikes, whereas depression occurs when the post-synaptic unit fires in between two pre-synaptic spikes (see text).

end of simulation time (Billings and van Rossum, 2009). In order to allow for activity-dependent pruning of synapses, we set $w_{ij} = 0$, if $w_{ij} \leq w_{min}$ following application of Eq. (2.7). If the connection is to be terminally deleted, we set also $A_{ij} = 0$ once $w_{ij} = 0$.

A model for tSTDP (see above) is implemented in the following way. The mechanism is similar to that of pSTDP in the sense that the difference between spike times in pre- and post-synaptic neurons determines the amount of synaptic modulation, however one extra term is added to the weight update function. This extra term considers the temporal difference between the two most recent spikes of one of the spiking nodes. Hence, the spike triplet. Potentiation occurs in a similar way as in pSTDP, however the amount of synaptic modulation is also in function of the difference between the two most recent spikes of the post-synaptic unit. Therefore, this rule is identified as *post-pre-post*. Similarly, depression occurs analogously to pSTDP, but the synaptic update is in function of the temporal difference between the two most recent spikes of the pre-synaptic neuron. So, the rule is summarized as *pre-post-pre*. Figure 2.4 shows a schematic representation of these two rules.

Thus, in order to extend pSTDP into tSTDP we add some extra terms to the equation for $\Delta w_{ij}(\Delta t)$ shown above which yields:

$$\Delta w_{ij}(\Delta T_1) = \begin{cases} a_p \exp\left\{\frac{-\Delta T_1}{T_p}\right\} \exp\left\{\frac{-\Delta T_2}{T_y}\right\} & \text{if } \Delta T_1 \geq 0 \\ -a_d \exp\left\{\frac{\Delta T_1}{T_d}\right\} \exp\left\{\frac{-\Delta T_2}{T_x}\right\} & \text{if } \Delta T_1 < 0 \end{cases}$$

where parameters a_p and T_p set the amount and duration of LTP, whereas a_d and T_d set

the amount and duration of LTD, as with pSTDP (see above); T_1 represents the difference between pre- and post-synaptic spike times, T_2 denotes the temporal difference between the two most immediate post-synaptic spikes (if $\Delta T_1 \geq 0$) or the temporal difference between the most immediate pre-synaptic spikes (if $\Delta T_1 < 0$); and T_x and T_y are two parameters that in a similar fashion as with parameters T_d and T_p set the amount of influence of immediate spikes for depression and potentiation, respectively.

Lastly, the computational implementations that have been just described take into account only *nearest spikes*, that is, only the immediate spikes that have been emitted, unlike other models based in *all-to-all* interactions, in which the weight update takes into account the complete spike history of a neuron in order to estimate the synaptic modulation (Pfister and Gerstner, 2005).

2.2.6 Node success

We introduce a local measure of the performance of a node during simulation time. The node success of node i at time t is the fraction of *out-neighbours* of this node that become active at time $t+1$ when node i spikes at time t, in other words:

$$\varphi_i(t) = \frac{\sum_{j=1}^{N} A_{ij} s_j(t+1)}{\sum_{j=1}^{N} A_{ij}} \tag{2.8}$$

where A is the adjacency matrix, and $s_j(t+1)$ the state of node j at time $t+1$. The sum in the numerator runs for all out-neighbours j of node i.

Thus, node success measures the performance of an individual spike in terms of the subsequent spikes triggered by such initial activation, which occur within the out-neighbourhood of a given node. In contrast to other network statistics (e.g. degree distribution, mean shortest path length, mean clustering coefficient, etc.) node success is a local measure of performance. Unlike the branching ratio (see Sect. 2.2.3), node success captures the local performance of a spike. This makes it a more detailed description of the development of an avalanche than the branching ratio. Through this statistic we will be able to track the type of nodes that contribute more to the critical state of the system (e.g. hubs, dwarves, high locally clustered nodes or low locally clustered ones). Moreover, we will be able to know what type of topology yields the most successful nodes during the critical state (e.g. fully-connected, random or scale-free networks). This type of analysis is only possible when considering a local measure (node success) instead of a global one (branching ratio).

We consider two different averages of node success. First, the *mean node success per node* $\langle\varphi\rangle_i$ which results from considering only the times in which a node spikes and then averaging its node success at each of these times. In other words,

$$\langle\varphi\rangle_i = \frac{1}{|\tau_i|}\sum_{t\in\tau_i}\varphi_i(t) \tag{2.9}$$

where $\tau_i = \{t : s_i(t) = 1\}$. The other mean success that we consider is the *mean node success per time step* $\langle\varphi(t)\rangle$ which results from averaging the node successes of all nodes in the system at a particular time step. In other words,

$$\langle\varphi(t)\rangle = \frac{1}{N}\sum_{i=1}^{N}\varphi_i(t) \tag{2.10}$$

Node success is the basic ingredient for a long-term plasticity rule that we propose as a mechanism to achieve SOC in neural systems, and that we will refer to in subsequent chapters as *node success driven plasticity* (NSDP).

Node success and the branching ratio: The branching ratio has been used previously to study the critical state of neuronal avalanches in slices of neurocortex (Beggs and Plenz, 2003). It is defined as the ratio σ of active nodes at time t to the number of active nodes at time $t-1$. In other words: $\sigma = descendants/ancestors$. It might seem that the measure of node success is similar to the branching ratio. However, the latter is a global statistic that provides details of the system as a whole, as such it disregards all the details of avalanche evolution in terms of what type of nodes are contributing more or less to the avalanche, that is, what nodes are more *successful*. The branching ratio can be thought of the "system's (node) success". It can be used to identify the regime of the collective dynamics of the system, whereas the node success cannot. In other words, a high or low node success for a given node does not provide enough information in order to characterise the critical state. Lastly, a *local* interpretation of the branching ratio has been defined as $\sigma_i = \sum_{j\in O(i)} P_{ij}$, where $O(i)$ denotes the out-neighbourhood of node i, and P_{ij} represents the probability that j will become active as a consequence of i becoming active (Beggs, 2007). This local measure is not analogous to node success. The term P_{ij} in the definition of the local branching ratio implies the presence of some stochasticity in the way that nodes become active, whereas in our working model (the Eurich model, Sect. 2.2.3) the fact that nodes accumulate activity in their membrane potential, and when this variable goes beyond threshold the node will become active, implies that their spiking is deterministic.

Finally, in Fig. 2.5 we show the development of the dynamics of a 6-node graph. There we also present the value of the node success φ for nodes that become active

throughout the simulation. As mentioned earlier, external driving occurs when there is no node active in the system. A node is chosen uniformly at random and its membrane potential in increased by a fixed amount of energy I_{ext}. As soon as a node's membrane potential goes beyond threshold as a result of the external driving, relaxation dynamics take place in the form of an avalanche of node activations (that is, spike propagation). During this process the driving is stopped. This results in a separation of timescales between driving and relaxation dynamics which is required for criticality to occur in a system (see Sect. 2.2.1). When there is no node being triggered by spike propagation, the system reaches a metastable state and external driving resumes. The size of the avalanche is the number of nodes that became active during the avalanche, whereas its duration is the number of time steps until the avalanche stops.

2.3 Related Work

In this section we present the work done by other researchers which is closest to ours. We provide minimal details, and offer a briefly description about how their work is related to our research, and why it is not. In general terms, our work differs from theirs by introducing complex networks along with other dynamic notions such as tSTDP or the node success metric, in the study of self-organised criticality in neural systems. We stress the importance of using this class of networks because of their apparent ubiquity in natural and man-made systems. In particular, because it has been reported that brain networks exhibit some features of complex networks.

2.3.1 Bornholdt model

The model by Bornholdt and Röhl (2003) is an example of a system that exhibits a structural phase transition rather than a dynamical one. Their model self-organises to a particular topological configuration through a re-wiring activity-dependent mechanism that is based on the amount of correlation (or anti-correlation) among nodes.

The model consists of N neurons arranged in a 2D lattice with random asymmetric weights. Node connections are based on Moore's neighbourhood, which is a nearest neighbour connectivity in a lattice which results in 8 neighbours for each node. Each node is in state $\sigma = +1$ or $\sigma = -1$, and weights w_{ij} for nodes i and j are taken in the interval $[-1, 1]$. States are updated stochastically in parallel based on the internal input received from a node's neighbourhood following these equations:

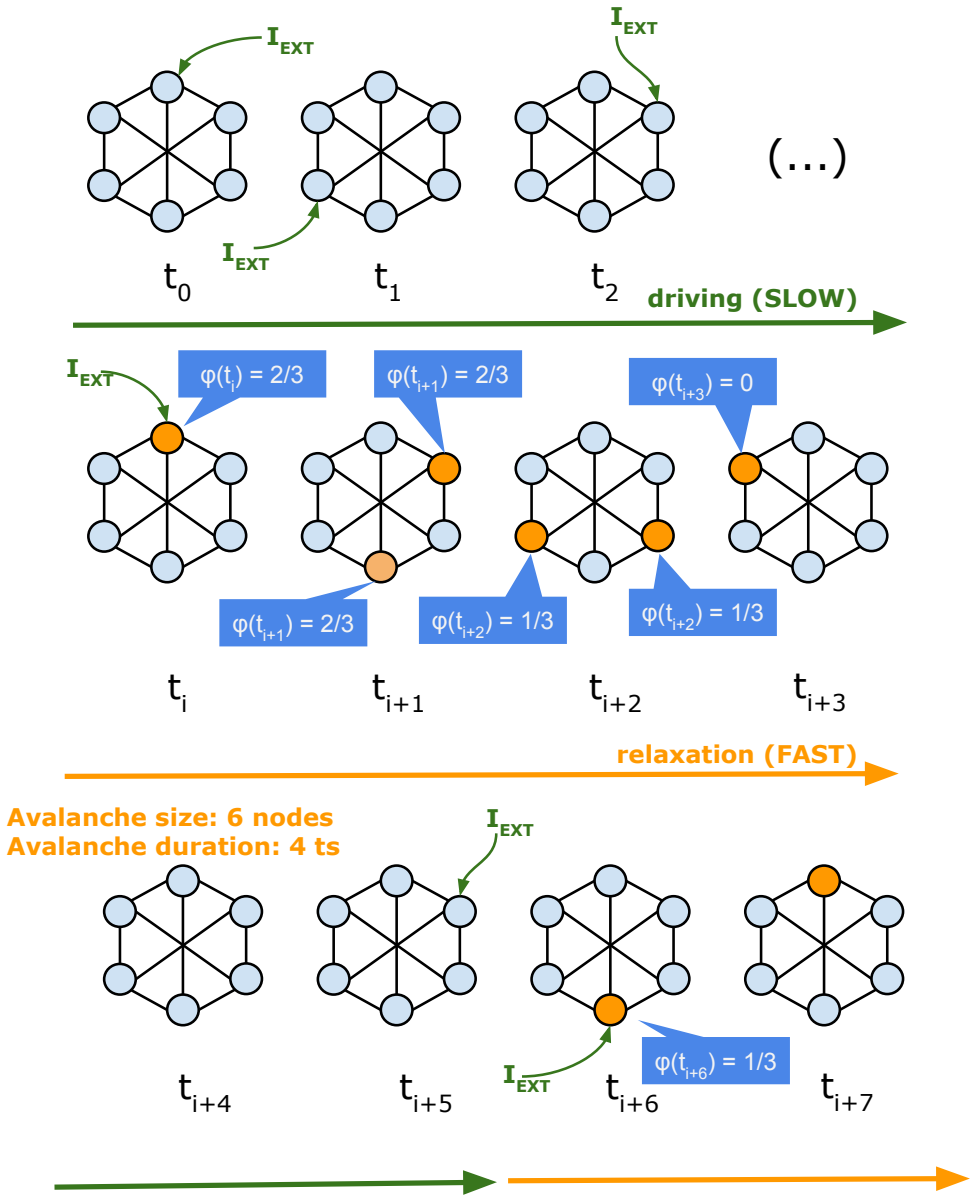


Figure 2.5: Avalanche dynamics (see text). External driving takes place on the system when there are no active nodes, and stops as soon an avalanche starts. At time t_i a node becomes active and propagates a spike to its out-neighbours. The number of out-neighbours that become active as a result of spike propagation is captured by the node success statistic φ. The avalanche ends when there is no node susceptible of going beyond threshold. The second avalanche continues after t_{i+7}; and $\varphi(t_{i+7})$ for the only active node at this time step is estimated based on the number of nodes active at t_{i+8}.

$P(\sigma_i(t+1)=+1)=g_\beta(f_i(t))$, and $P(\sigma_i(t+1)=-1)=1-P(\sigma_i(t+1)=+1)$. Moreover, $f_i(t)=\sum_{j=1}^{N} w_{ij}\sigma_j(t)+\theta_i$, where θ_i is a threshold for node i. Lastly, $g_\beta(f_i(t))=1/(1+\exp\{-2\beta f_i(t)\})$.

The above set of equations specify the dynamics of each node, whereas the local re-wiring rule follows this mechanism. Average correlation for nodes i and j over a time interval τ is estimated by $C_{ij}(\tau)=1/(\tau+1)\sum_{t=t_0}^{t_0+\tau}\sigma_i(t)\sigma_j(t)$. If activity of these nodes is highly correlated (above a threshold α) the nodes will obtain a common link, if the opposite happens, they lose a link.

These equations allow the system to self-organise to a critical topological configuration that is robust against perturbations and independent of initial parameters. Such configuration reflects patterns of correlated activity that occur within the network, thus resembling mechanisms of Hebbian plasticity. As there is no fine-tuning of a control parameter, the system exhibits self-organised criticality in the collective dynamics.

How is this model related to our work? This is a candidate model to explain the emergence of criticality in neural systems. It implements local activity-driven plasticity mechanisms to re-wire the network which ultimately lead to the critical state. How is it not related to our work? Unlike this work, our research involve spike-timing-dependent plasticity mechanisms to modulate synaptic efficacies. Moreover, our work go beyond simple lattices and incorporates complex topologies such as scale-free small-world networks.

2.3.2 Jost model

This model is aimed to show how a simple version of the spike-timing-dependent mechanisms (STDP, Sect. 2.2.4) is able to prune an initially globally connected network in such a way that the resulting network is scale-free (Jost and Kolwankar, 2009). This work might be inspired by the fact that the human brain is densely connected at very early stages of life, and later, behaviour and learning alter this initial structure into a more robust and efficient topology.

The model is comprised by a network of N neurons, whose state is specified by an N-dimensional column vector X_n. These states are values in the interval $[0,1]$. Initially, all neurons are connected to each other, and their synaptic strength is specified by an $N\times N$ coupling matrix G, which can be considered as a weighted adjacency matrix. G_{ij} $(i\neq j)$ is the coupling strength of the edge from j to i, and authors impose the balancing condition: $G_{ii}=1-\sum_{i\neq j}G_{ij}$. Initially, coupling strengths are taken randomly

from the interval $[0, g_{in}^{max}]$, where g_{in}^{max} is a boundary parameter that prevents initial synchronisation of any pair of nodes. $G_{ij} = 0$ means that there is no link from j to i. The values in the matrix G are non-negative, which implies that inhibitory synapses are not considered.

At every time step, the state vector is updated by equation $X_{n+1} = G\, f(X_n)$, where G is the adjacency matrix and f is a map from $\Omega = [0,1]^N$ onto itself modeling the dynamics of the node. This function is the logistic map $f(x) = \mu x(1-x)$. At each timestep the synaptic strengths of each node is updated based on their activity according to the discrete version of the STDP rule give by the equation $G_{ij}(n+1) = G_{ij}(n) + \varepsilon(X_j(n-1)X_i(n) - X_j(n)X_i(n-1))$; for $i \neq j$, and ε is the learning rate.

The authors report that the final structure of the network follows a power law, which implies that their STDP rule prunes a globally coupled network so that the final structure is scale-free, and hence, it contains highly connected nodes.

One observation to the model is that it does not involve very realistic features that could be related to neural systems, particularly their choice for the node state updating rule does not have a biological interpretation. Something similar could be said about their STDP rule. Nevertheless, it is remarkable how this simple rule shapes the network in such a way that a scale-free network could result from its implementation. Moreover, although this could have not be the intention of the authors, there is no analysis regarding the critical state of the dynamics of the system. How is it related to this thesis? The authors present a activity-dependent plasticity rule that turns a fully-connected network into a scale-free network. It shows how a simple local rule is able to give rise to a complex topology. How is it not related to this thesis? This model does not exhibit critical dynamics, and thus, do not imply all the benefits on neural computation described in Sect. 2.2.2.

2.3.3 Meisel model

The model of Meisel and Gross (2009) implements the ideas behind adaptive networks by considering a network of threshold elements whose activity influences the coupling strengths in each synapse. Thus, the topology of the network is altered by the inner dynamics of the network, and this in turn has an impact on the overall dynamics.

The model consists of a network of N leaky integrate-and-fire neurons, where 80% of them correspond to excitatory neurons, whereas the remaining 20% are inhibitory. The model also implements a transmission delay constant τ.

The topology of the network changes due to STDP but in a slower timescale than the spiking dynamics. That is, the update of the topology is carried out after a certain amount of time and not immediately at each time step as in the case of the Jost model.

The simulations in this model proceed as follows. The simulation runs according to the standard leaky integrate-and-fire model for a fixed time t_{sim}. At each time step in this stage, the activity among neurons is recorded in a couple of variables for each node. Once t_{sim} is over, the topology of the network is updated considering the mechanism of STDP and based on the values of the recorded variables for each node. After the update rule has been applied, the system is re-started assigning random membrane potentials to the neurons. The procedure of simulating the dynamics and then applying the topology update rule is iterated to allow the network topology to relax to a self-organized state.

The authors reported that the connectivity in the network evolved towards a characteristic connectivity independent of initial conditions through their own implementation of STDP mechanisms. Moreover, they found that this characteristic connectivity scales with system size following a very particular scaling law.

By defining a set of variables to measure the amount of synchronization among nodes in the network, the authors were able to perform a test for criticality throughout the simulation. They found power law behavior in a variable that measures the spike-time correlation in the evolving network. Moreover, the authors report that the synaptic strengths evolve following a power law with exponent -1.5. However, although they implement an update rule on the network topology they did not carry out an analysis on network statistics that could describe the structure of the network.

How is this model related to this thesis? This model suggests that self-organised criticality emerges in neural systems through STDP mechanisms. Unlike this model, we suggest that SOC is achieved by a different plasticity mechanism. One that is based in the metric of node success that we described in Sect. 2.2.6. Moreover, as we will describe in Chap. 5, in our model STDP has a negative effect over the critical behaviour of the system. How is this model not related to ours? Unlike this model, we implement synaptic modulation throughout the entire simulation time which is different from their mechanism in which simulation time is interrupted, synaptic weights are modulated, and afterwards the system is re-started. Additionally, we go beyond simple spike-pairs to implement the STDP rule, and consider triplets of spikes, which fit better experimental data.

2.3.4 Basalyga model

The model presented by Basalyga et al. (2011) consists of 100 leaky integrate-and-fire excitatory and inhibitory neurons connected randomly. The authors introduce STDP mechanisms to implement activity-dependent plasticity and synaptic pruning. They observe that a small-world structure emerges from the initial network due to the plasticity rule. Each node in the network possesses a neuronal membrane potential which is updated according to equation $C\dot{V} = -g_L(V - E_L) + S(t) + G(t)$, where C is the specific capacitance, g_L is the leak conductance, E_L the leak reversal potential, $S(t)$ represents the spiking mechanism (which in their experiments is Poissonian), and $G(t)$ denotes conductance based synaptic interactions, which specify the interaction between neighbouring nodes.

In the model each node has three types of synapses: fixed excitatory synapses, plastic (STDP) excitatory synapses and fixed inhibitory synapses. This means that not all connections are plastic and a single node can behave both as an excitatory and an inhibitory neuron, which seem to be in opposition with Dale's principle that states that a neuron releases the same type of neurotransmitter in all of its synapses (Strata and Harvey, 1999). In the model STDP is implemented through equations $w_{ij} = w_{ij} + \Delta w_{ij}(\Delta t)$, where Δt is the STDP function which depends on the time difference between pre- and post-synaptic spikes (Sect. 2.2.5).

$$\Delta w_{ij}(\Delta t) = \begin{cases} a_p \exp\left\{\frac{-\Delta t}{T_p}\right\} & \text{if } \Delta t \geq 0 \\ -a_d \exp\left\{\frac{\Delta t}{T_d}\right\} & \text{if } \Delta t < 0 \end{cases}$$

The authors impose hard bounds on the synaptic update in order to avoid unbounded growth in the weights. Moreover, they assess the amount of *small-world-ness* based on ideas presented by Humphries and Gurney (2008) (Sect. 2.1.4). In their experiments they observe the emergence of a small-world structure emerging from random topologies.

How is this model related to this thesis? The model implements activity-dependent plasticity in random networks which "carves" a small-world structure out of the initial topology. Motivated by their methods, we use the metric S described by Eq. (2.4) to assess the presence of the small-world property in our systems. How is it not related to this thesis? The model is not based on critical dynamics, which also leaves out the computational benefits of this dynamical regime for the system. Moreover, the fact that individual neurons exhibit three types of plasticity and the fact that only some of them are plastic is somehow unrealistic.

2.3.5 Levina model

The model proposed by Levina et al. (2007) was developed as an extension to the Eurich model (Sect. 2.2.3) to include activity dependent dynamics in the synaptic weights, which ultimately would drive the system to the critical regime without external supervision. Thus achieving *self-organised criticality* (SOC).

The model consists of N threshold elements each of which is characterised by a state variable $h \in (0, \theta)$ representing the neuronal membrane potential. This variable evolves according to equation $\dot{h} = \delta_{i,\xi} I^{ext} + \frac{1}{N} \sum_{j=1}^{N} u\, J_{ij} \delta(t - t_{sp}^{j} - \tau_d)$. The first term at the right-hand side of the equation refers to the external input that is supplied to the system in case there is no activity within it. This corresponds to the external slow driving that has been proposed as a requirement for systems exhibiting SOC. The second term refers to the internal input integration from nodes in the *out-neighbourhood* of the current node. Each neuron integrates inputs in this manner until they reach the threshold θ. Once this occurs, the node emits a spike that is delivered to all post-synaptic nodes at a fixed delay τ_d. Afterwards, the membrane potential is reset. The variables J_{ij} are the synaptic weights between nodes i and j. These variables are dynamic and subject to the update rule $\dot{J}_{ij} = \frac{1}{\tau_J}\left(\frac{\alpha}{u} - J_{ij}\right) - u\, J_{ij} \delta(t - t_{sp}^{j})$, which describes the amount of available neurotransmitter in the synapse. If a spike arrives at the synapse, the available transmitter is diminished by a fraction u. If the pre-synaptic node is inactive, then the synapse recovers and its strength J_{ij} approaches the value α/u at a slow time scale specified by τ_J. In this way, the maximal strength of a connection is determined by the parameter α, and can be observed only when the synapse is fully recovered. It is this interplay between available resources and recovery that provides the system with a feedback mechanism essential for criticality to occur without external supervision.

How is it related to this thesis? The Levina model is a model of neuronal avalanche dynamics that succeeds in exhibiting SOC suggesting at the same time how this might be achieved by brain networks. Its dynamical synapses make it a strong candidate along with the excitatory/inhibitory balance suggested by Beggs and Plenz (Beggs and Plenz, 2003) (Sect. 2.2.2) to explain how criticality occurs in neural systems. How is it not related to this thesis? Although the authors experiment with different topologies and sizes, the notion of complex networks is still not in the model. As well, although they experiment with STDP, their model does not consider the spike-triplet version of this plasticity mechanism, which has been suggested as a more accurate model of neuronal plasticity (Sect. 2.2.4).

Chapter 3

Network Topology

In this chapter we present the statistical properties of the networks created for the purpose of experimenting with the Eurich model and its variations. Additionally to fully-connected networks (generally considered a homogeneous toy network), we consider complex networks in our experiments, that is, heterogeneous structures that are more akin to real-world networks than fully-connected ones (see Sect. 2.1.1).

In general terms, the *directed* networks that we generated can be classified in the following way:

1. Fully-connected networks.
2. Random networks.
3. Scale-free networks:
    - *In-degree* scale-free networks with *low* mean clustering coefficient.
    - *In-degree* scale-free networks with *high* mean clustering coefficient.
    - *Out-degree* scale-free networks with *low* mean clustering coefficient.
    - *Out-degree* scale-free networks with *high* mean clustering coefficient.

As mentioned in Sect. 2.1.5, scale-free networks feature a degree distribution that follows a power-law of the form $P(k) \sim k^{-\beta}$. In our work we set $\beta = 1$. As well, directed networks possess both an in-degree and an out-degree distribution. A power-law distribution could occur in either of them. Because dynamics could in principle differ by this fact we consider scale-free networks in which the out-degree distribution follows a power-law, as well as scale-free networks in which the in-degree distribution follows a power-law. As mentioned in Sect. 2.1.7, an *out-degree* scale-free network

features the presence of *broadcasting hubs*, whereas an *in-degree* scale-free network features the presence of *absorbing hubs*. Details on how we build this type of networks can be found in Appendix A.

Moreover, we consider two levels of mean clustering coefficient for scale-free networks (*low* and *high*) by tuning a parameter in an algorithm based on ideas by Holme and Kim (2002). Real-world networks often feature a power-law degree distribution along with a high amount of clustering which relates them to the *small-world property* (see Sect. 2.1.4). The small-world property is not a binary one, and as such, there exists a degree of what we would call *small-world-ness*. The process of varying the mean clustering coefficient in our scale-free networks has an immediate effect in the degree of small-world-ness of such structures (see Sect. 3.5). Scale-free networks with low mean clustering coefficient possess a lower degree of small-world-ness when compared against scale-free networks with high mean clustering coefficient. By considering structures with different amounts of clustering, we can study the critical behaviour of complex networks with different degrees of *small-world-ness*. Once again, details on how we construct such networks can be found in Appendix A.

Lastly, we consider four different system sizes for the sake of studying the scaling properties of the system dynamics. Sizes considered are $N = 128, 256, 512$ and $1,024$. For the sake of statistical robustness we created 30 networks for each size and class.

The statistical features that we are interested to analyse in the networks that we generated are:

1. Degree distributions: in-degree, out-degree.

2. Local clustering coefficients and mean clustering coefficient of network.

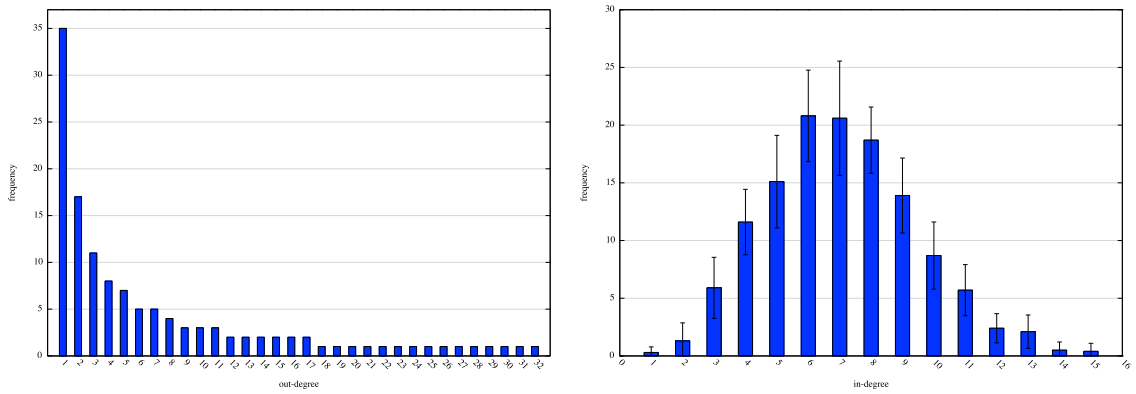
3. Characteristic path length.

4. Mean connectivity and density of network.

5. Degree of small-world-ness as estimated by measure S (see Sect.2.1.4).

6. Local clustering coefficient vs. in- and out-degree.

7. Motif profile of network.



(a) Out-degree distribution

(b) In-degree distribution

Figure 3.1: Degree distributions for scale-free networks of size $N = 128$ with *low* mean clustering coefficient and out-degree power-law distribution. Error bars in (b) represent standard deviations.

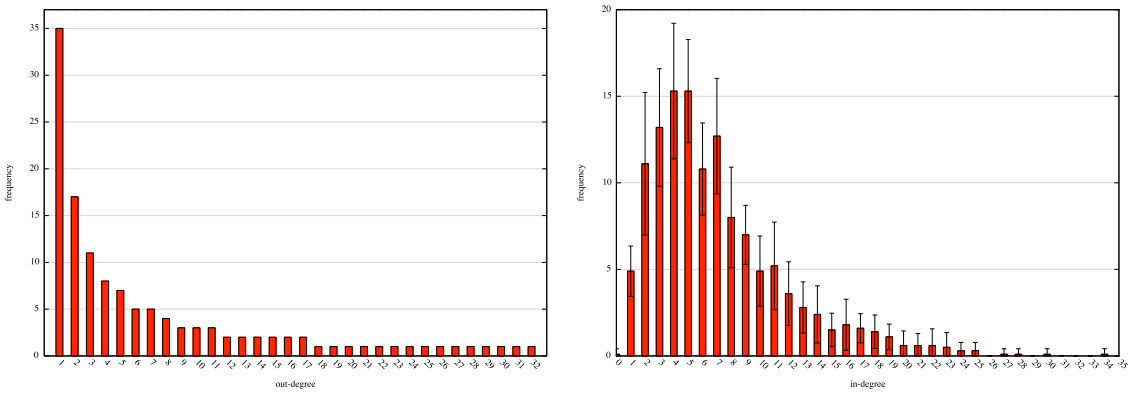
3.1 Degree Distributions

The degree distribution of a fully-connected network is just a spike at the degree shared by all nodes in the network. Therefore, in this section we will not refer to this type of network, and instead we will focus only in random and scale-free networks.

We created sets of directed networks of different structures and sizes, in particular, our scale-free networks differ in two different ways: *i)* the in-degree or the out-degree distribution follows a power-law, and *ii)* the mean clustering coefficient is high or low. This yields four different sets of scale-free networks, which for the different sizes (128, 256, 512, and $1,024$) results in a total of 16 different sets of scale-free networks. Our process of building a *directed* scale-free network results in a network whose out-degree distribution follows a power-law, whereas its in-degree distribution can *in principle* take any shape. The process of transposing the adjacency matrix associated with a network is similar to the action of reversing the direction of the edges in it. Thus, by transposing the adjacency matrix of a network whose out-degree distribution follows a power-law we obtain a network with reversed edges, and therefore whose in-degree distribution follows a power-law.

As expected the in- and out-degree distributions of our random networks follow Poisson distributions (see Sect. 2.1.3).

Figure 3.1 shows the in- and out-degree distributions for the networks of size $N = 128$ with *low* mean clustering coefficient and whose out-degree distribution fol-



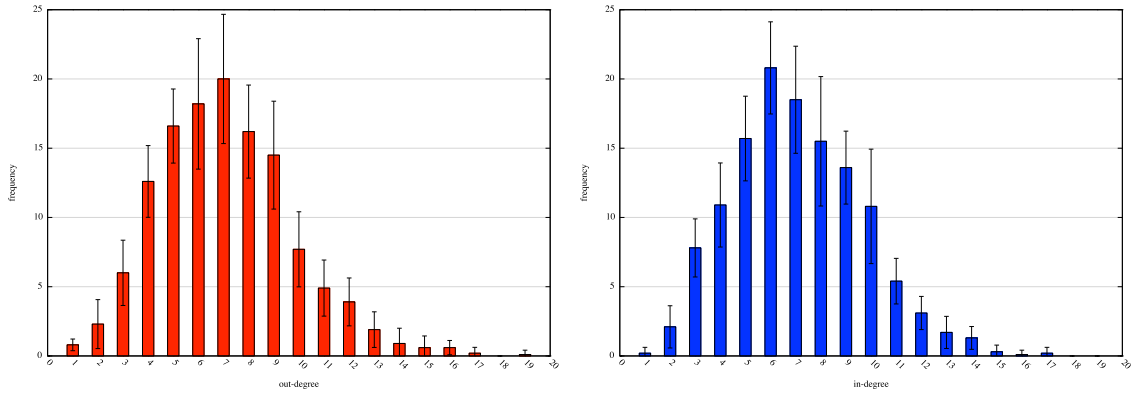
(a) Out-degree distribution

(b) In-degree distribution

Figure 3.2: Degree distributions for scale-free networks of size $N = 128$ with *high* mean clustering coefficient and out-degree power-law distribution. The right-skewed distribution in (b) suggests the presence of absorbing hubs as a result of maximising the mean clustering coefficient in these networks. Error bars represent standard deviations.

lows a power-law. A low mean clustering coefficient implies that the small-world property in this network is also low. We show error bars in Fig. 3.1b but not in Fig. 3.1a as the process of producing scale-free networks yields always the same power-law distribution, that is, its standard deviation is null. We introduce randomness in such a process by choosing uniformly at random the nodes to connect to, which at the end results in variations in the shape of the in-degree distribution but not in the out-degree distribution (see Appendix A).

Similarly, Fig. 3.2 shows the degree distributions for a directed scale-free network of similar properties as in Fig. 3.1 except that now the networks possess a *high* mean clustering coefficient, which results in a higher amount of small-world-ness than in the previous case. Moreover, note that the out-degree distribution, the distribution that follows a power-law, has exactly the same shape as in Fig. 3.1a. This is due to our algorithm to generate scale-free networks, in which the shape of the out-degree distribution is fixed but the nodes chosen to connect to are taken at random, which results in variations of the in-degree distributions across networks. Finally, the *clustering step* (see Appendix A) introduces a bias in order to choose nodes that will increase the overall clustering coefficient of the network. This might explain why the in-degree distribution in Fig. 3.2b looks more right-skewed than the distribution shown in Fig. 3.1b. Furthermore, the shape of the distribution in Fig. 3.2b suggests the emergence of in-degree hubs (that is, absorbing hubs) coexisting with the out-degree hubs (ie. broadcasting



(a) Out-degree distribution (b) In-degree distribution

Figure 3.3: Degree distributions for random networks of size $N = 128$. Error bars represent standard deviations.

hubs) which were intentionally produced and represented in Fig. 3.2a. Does the process of maximising the mean clustering coefficient in a scale-free network have the consequence of producing both absorbing and broadcasting hubs? Answering this question is beyond the scope of this thesis but by posing such a question we would like to point directions of future research.

Finally, Fig. 3.3 shows in- and out-degrees for our random networks of size $N = 128$. Here, degree distributions resemble a Poisson distribution as mentioned in Sect. 2.1.3. In-degree distributions of scale-free and random networks have a similar shape, Fig. 3.4 shows a comparison among the in-degree distributions of our scale-free networks with low and high mean clustering coefficient along with the in-degree distributions of our random networks of size $N = 128$. The scale-free network with high mean clustering coefficient (light blue in figure) exhibits a long right tail which suggests the presence of hubs in the network, in this case, absorbing hubs.

For larger system sizes (ie. 256, 512 and $1,024$) the observations described above are very similar. In fact, degree distributions exhibit some interesting scaling properties. For instance, Fig. 3.5a shows the out-degree distribution for scale-free networks of all system sizes. As we mentioned before, out-degree distributions for scale-free networks share the same shape for networks with high and low mean clustering coefficient. The long tail in Fig. 3.5a for system sizes larger than 128 is not shown for the purposes of presentation. However, Fig. 3.5b shows the out-degree distributions of scale-free networks of $1,024$ nodes. Figure 3.5b shows the distribution in linear coordinates, whereas Fig. 3.5c shows it with log-log plots. In this latter case the power-law

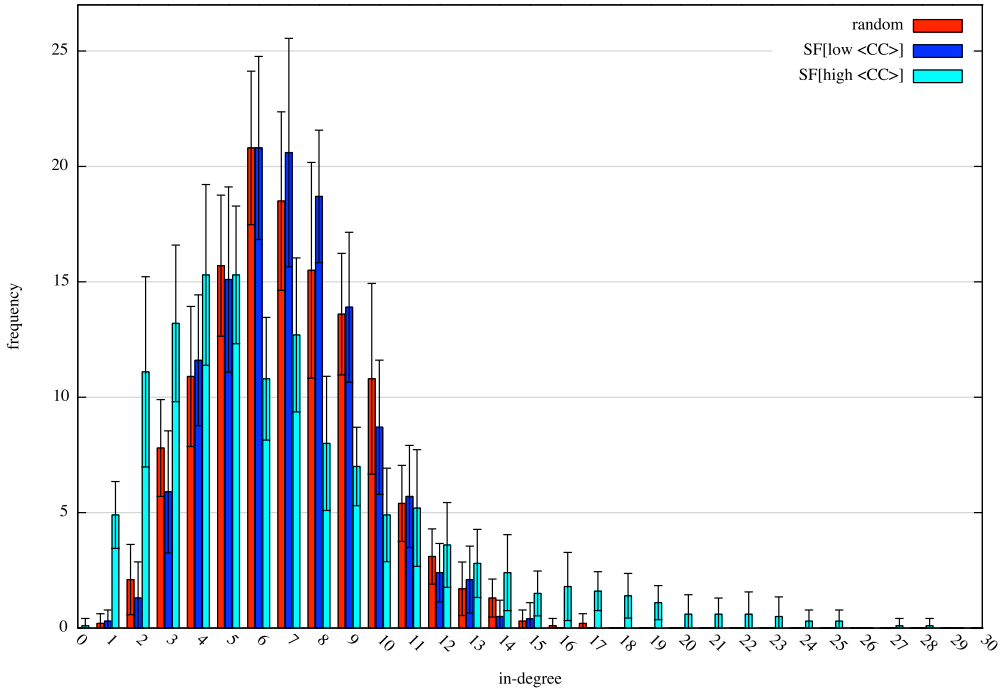
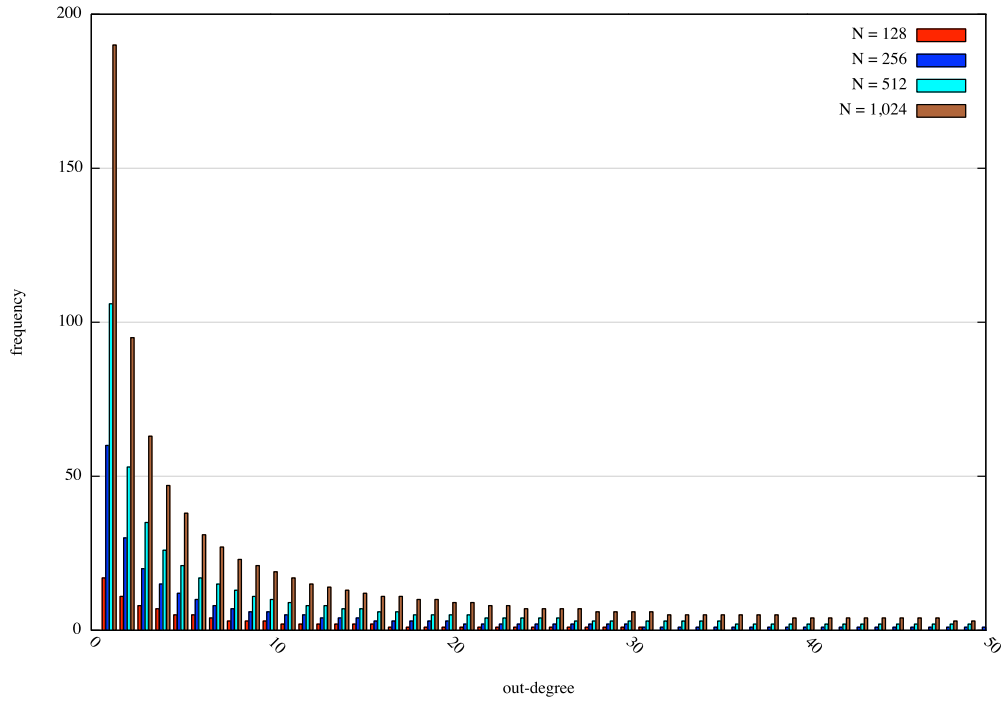


Figure 3.4: In-degree distributions for random and scale-free networks of size $N = 128$. The out-degree scale-free network with high mean clustering coefficient (light blue) exhibits the presence of absorbing hubs as a result from the process of maximising the overall clustering coefficient.

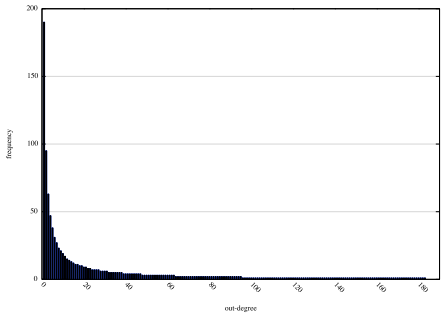
distribution is evident.

Figure 3.6 shows the in-degree distributions for out-degree scale-free networks of all sizes and levels of the mean clustering coefficient. Here we do not show error bars for the purpose of clarity. Figure 3.6a shows the in-degree distributions for a low average clustering coefficient. The distributions are bell-shaped exhibiting fast decaying tails. On the other hand, Fig. 3.6b shows the same type of distributions for high mean clustering coefficient. In this case the distributions exhibit a heavy-tailed shape with a slow decaying right tail suggesting the presence of absorbing hubs. Does the presence of absorbing hubs along with broadcasting hubs have an impact on the collective dynamics of the system? We will answer this and other questions in the upcoming chapters.

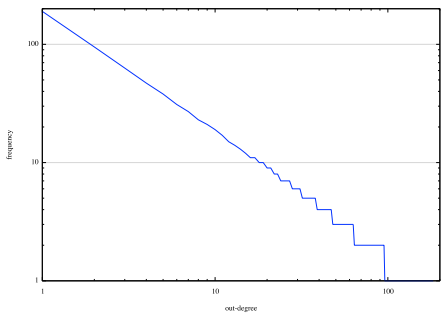
Finally, Fig. 3.7 shows both in- and out-degree distributions for our random networks in all sizes considered. As expected, the shape of those distributions follows a Poisson distribution with fast decaying tails. We do not present error bars in Fig. 3.7 for the sake of clarity. In summary, our scale-free networks exhibit a power-law in the



(a) Out-degree distributions

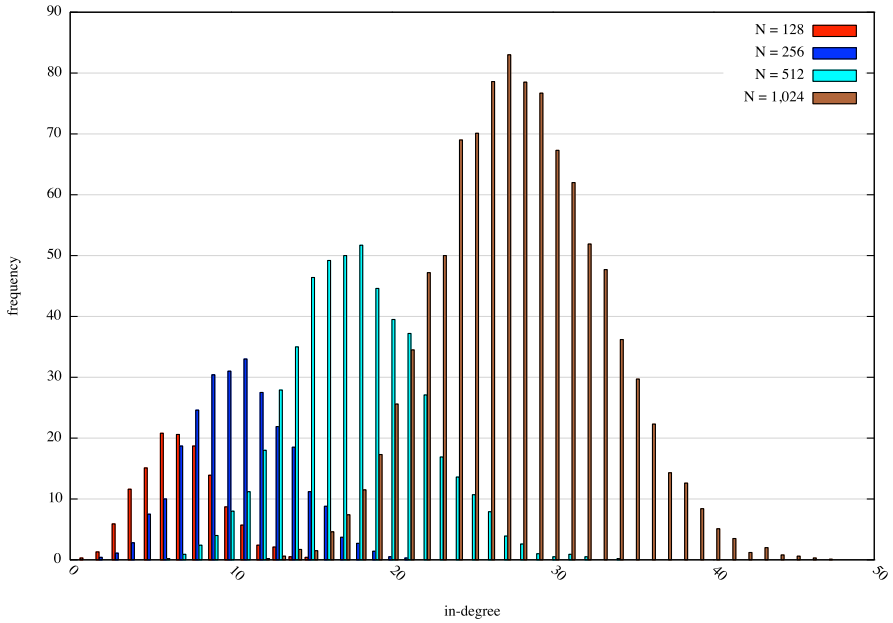


(b) Linear coordinates

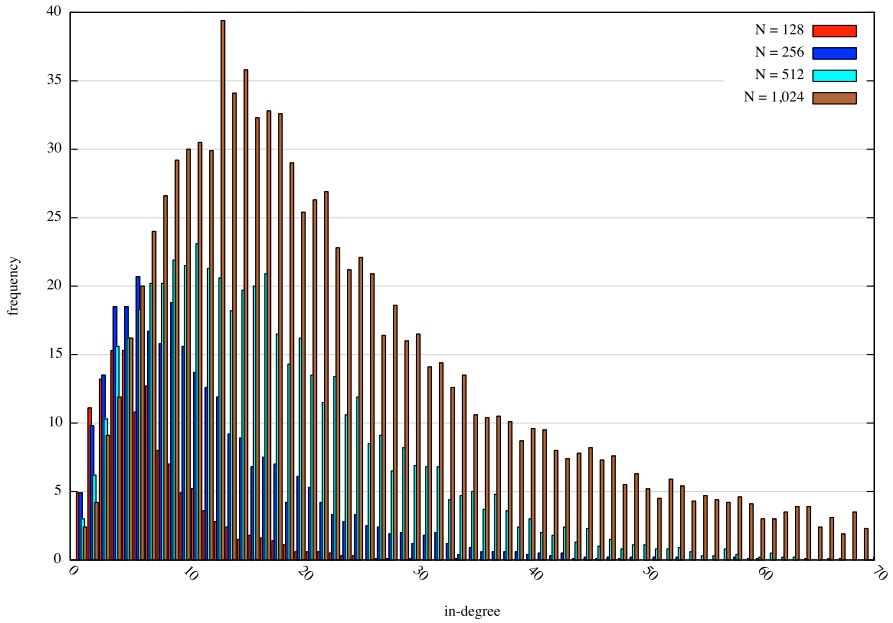


(c) Log-log coordinates

Figure 3.5: Out-degree distributions for scale-free networks. (a) shows a comparison of out-degree distributions for scale-free networks of sizes $N = 128, 256, 512$, and $1,024$. For the purpose of presentation, we do not show the whole tail of larger system sizes. (b) shows the distribution of networks with $1,024$ nodes in linear coordinates revealing a long tail. (c) shows the same distribution in log-log coordinates exhibiting a straight line which is the hallmark of power-law distributions.



(a) Low average clustering coefficient



(b) High average clustering coefficient

Figure 3.6: In-degree distributions of scale-free networks of sizes $N = 128$, 256, 512, and $1,024$. (a) shows in-degree distributions for networks with low mean clustering coefficient, whereas (b) shows them for networks with high mean clustering coefficient. Unlike scale-free networks with low mean clustering coefficient, high clustered scale-free networks exhibit the presence of a long tail not only in the out-degree distribution but also in the in-degree distribution.

out-degree distribution; this yields broadcasting hubs. On the other hand, the in-degree distributions can take the form of a symmetric Poisson distribution when the average clustering coefficient is low, and a right-skewed Poisson distribution with slow decaying tail when the average clustering coefficient is high. This is observed across all system sizes.

The same is true when reversing the direction of the edges, that is, when transposing the adjacency matrix associated to each network. In this situation, the out-degree distribution becomes the in-degree distribution and vice versa. Therefore, where before we had broadcasting hubs now we have absorbing hubs. Lastly, random networks behave as expected in all system sizes considered.

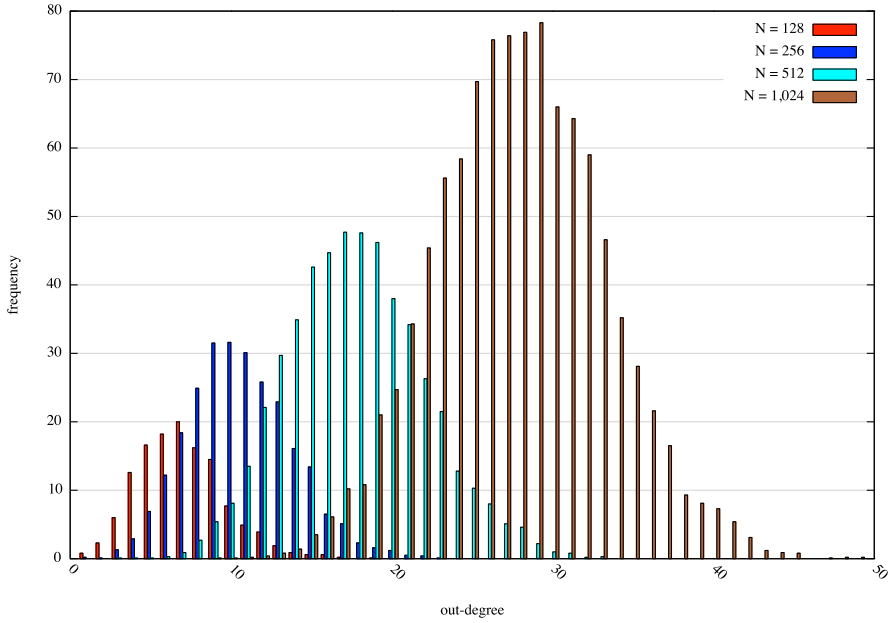
3.2 Local Clustering Coefficients

In Sect. 2.1.2 we described how the local clustering coefficient is computed for any given node. The mean clustering coefficient of a network is simply the mean of local clustering coefficients of all nodes in the network. Here, we will not consider the local clustering coefficients of nodes in a fully-connected network as it is unity for all nodes. That leaves us only with the local clustering coefficients of nodes in our scale-free and random networks for all different system sizes considered.

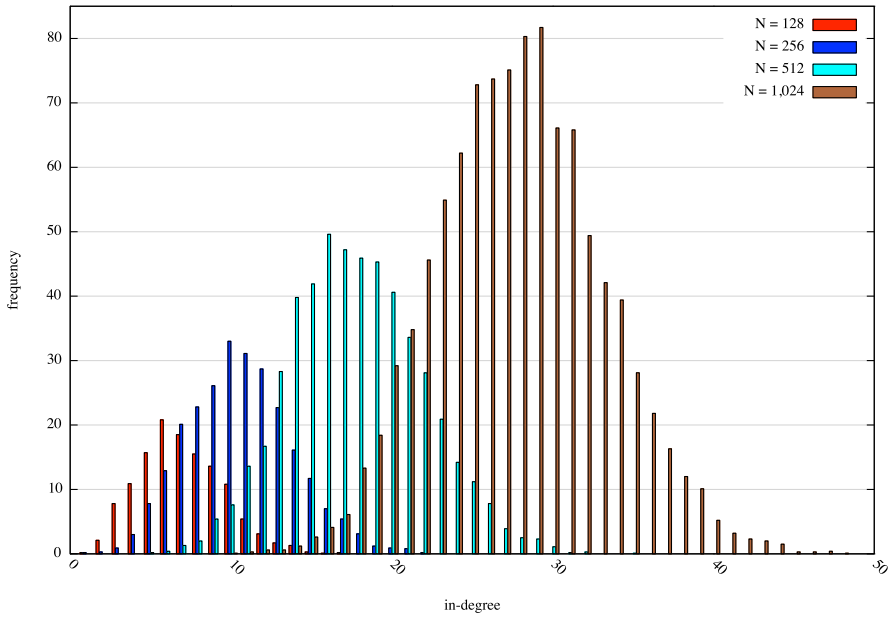
The purpose of this section is to present the distribution of the local clustering coefficients for scale-free and random networks. Are they normally distributed? Do they have a long tail yielding the value of the mean clustering coefficient rather uninformative? Are they power-law distributed for scale-free networks? In Sect. 2.1.4 we said that the small-world property is present when a network exhibits a *low* mean shortest path length and a *high* mean clustering coefficient. However, if the distribution of local clustering coefficients for a given network exhibits a power-law, then its mean might not reveal any information about the distribution.

Let us start with scale-free networks. Here, again it is enough to study out-degree scale-free networks (that is, networks whose out-degree distribution follows a power-law) because in-degree scale-free networks result from transposing the adjacency matrix associated to them, which in turn yields the same local clustering coefficient per node.

Figure 3.8 shows an example of the distribution of local clustering coefficients for scale-free and random networks of size $N = 1,024$. Random networks show a bell-shaped distribution of local clustering coefficients (see Fig. 3.8c). For out-degree scale-

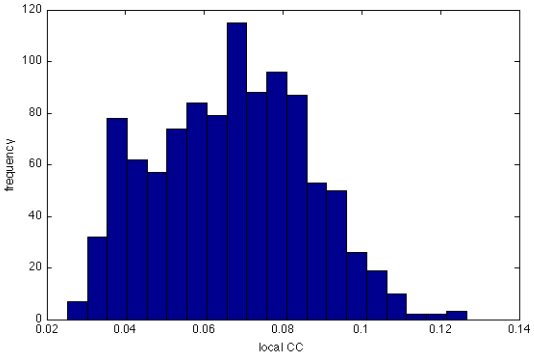


(a) Out-degree distribution

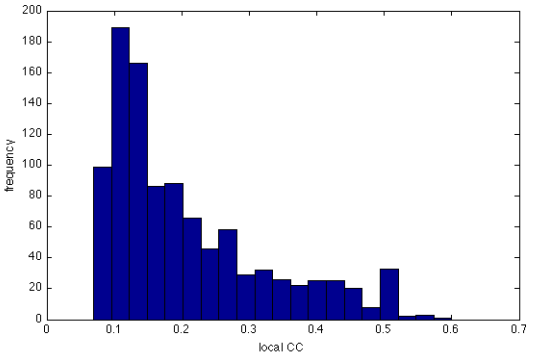


(b) In-degree distribution

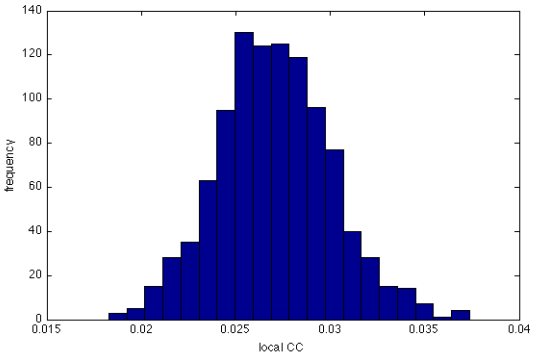
Figure 3.7: Degree distributions for random networks. Both types of distributions are bell-shaped which implies the absence of hubs in these networks. Error bars not shown for clarity.



(a) Scale-free net with *low* mean CC



(b) Scale-free net with *high* mean CC



(c) Random network

Figure 3.8: Distribution of local clustering coefficients (CC) for scale-free and random networks of size $N = 1,024$. The distribution of local CCs of random networks exhibit a clear bell shape, whereas those of scale-free networks do not. In particular, the distribution of local CCs in scale-free nets with high mean CC is right skewed, and reaches larger values of the CC than its scale-free nets with low mean CC and random structures. (This is an example of a single trial.)

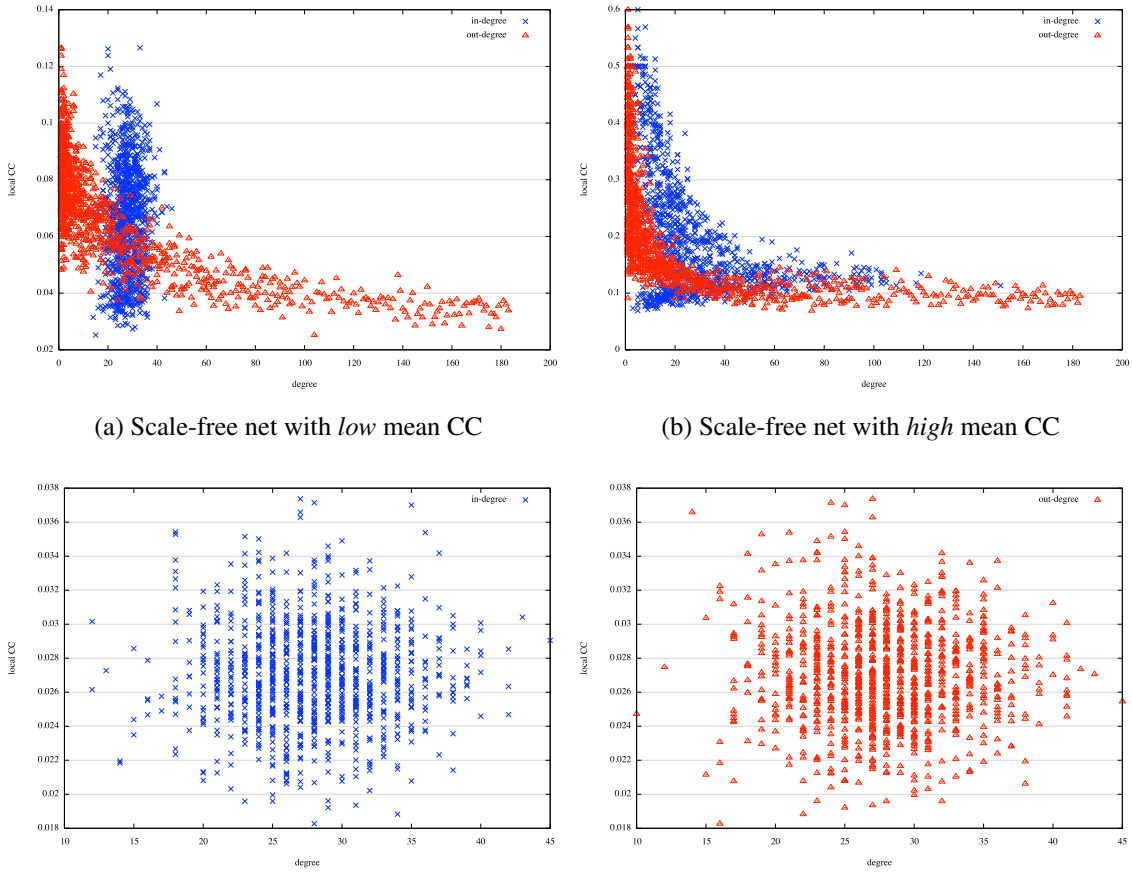
free networks with *low* mean clustering coefficient the distributions of local clustering coefficients are somehow bell-shaped with a well defined mean and a fast decaying right tail (see Fig. 3.8a). But, then for out-degree scale-free networks with *high* mean clustering coefficient the distribution of local clustering coefficients looks different (see Fig. 3.8b). In this case, the distribution exhibits a slow decaying tail as the value of the local clustering coefficient increases. Nevertheless, this right-skewed distribution is not a power-law.

This is interesting as both types of network (low- and high mean-clustered) are scale-free networks, and they possess exactly the same out-degree distribution. The process of creating triads in order to maximise the mean clustering coefficient in a scale-free network results in the particular shape of the distribution in Fig. 3.8b. Are the nodes in the tail of Fig. 3.8b the (broadcasting) hubs of the network? In the following sections we will see that in fact that is not the case. Hubs are the hardest nodes to cluster because their numerous edges decrease the probability that their neighbours will be connected among themselves. The limit-case in which massively connected nodes are also high clustered is -obviously- the fully-connected network. Thus, the *triad-formation* step in our algorithm (see Appendix A) has an effect mainly over dwarf nodes, that is, in non-hubs. These dwarf nodes will be better clustered than hubs, because the fact of having less connections increase the probability that their neighbours are connected among themselves. Lastly, although we only present the distribution of a single network of each type in Fig.3.8, we observe the same behaviour in all networks considered by class and size.

3.3 Local Clustering Coefficients vs. Degree

In this section we will explore the correlations between local clustering coefficients and node degrees. How well clustered are hubs? How well clustered are dwarves? How do they behave in random topologies? Once again, the discussion in this section comprises only random and scale-free networks as in fully-connected structures nodes are of one kind only and they possess the maximum value of local clustering coefficient, that is, unity.

In Fig. 3.9 we present how the local clustering coefficients are distributed against the two degrees (in- and out-degree) in a network. This figure provides more information than Fig. 3.8 as it describes how the local clustering coefficients are distributed per degree for the networks considered there. We present an example of the distribution



(a) Scale-free net with *low* mean CC

(b) Scale-free net with *high* mean CC

(c) Random network: in-degree vs. local CC

(d) Random network: out-degree vs. local CC

Figure 3.9: Scatter plots of in- and out-degrees versus local clustering coefficient (CC) for scale-free and random networks of size $N = 1,024$. (a), (b) A negative correlation between out-degree and local CC (red triangles) in scale-free networks is observed. The larger the out-degree the smaller the local CC. (c), (d) This correlation is not observed in random networks. (This is an example of a single trial.)

of degrees (in- and out-) against local clustering coefficients for three networks: two scale-free networks (low and high mean clustering coefficient) and a random network. They consist of $1,024$ nodes; and once again, although we present results of a single network for the sake of clarity, the behaviour that we describe in the following lines is observed across all networks and sizes. As well, it is sufficient to consider out-degree scale-free networks as the in-degree scale-free networks that result from transposing their adjacency matrices (inverting the direction of edges) gives a similar situation as in Fig. 3.9 but now the out-degree becomes the in-degree and vice versa.

Figures 3.9a and 3.9b show an inverse correlation between out-degree and local clustering coefficient (shown in red) that is not observed when comparing in-degree

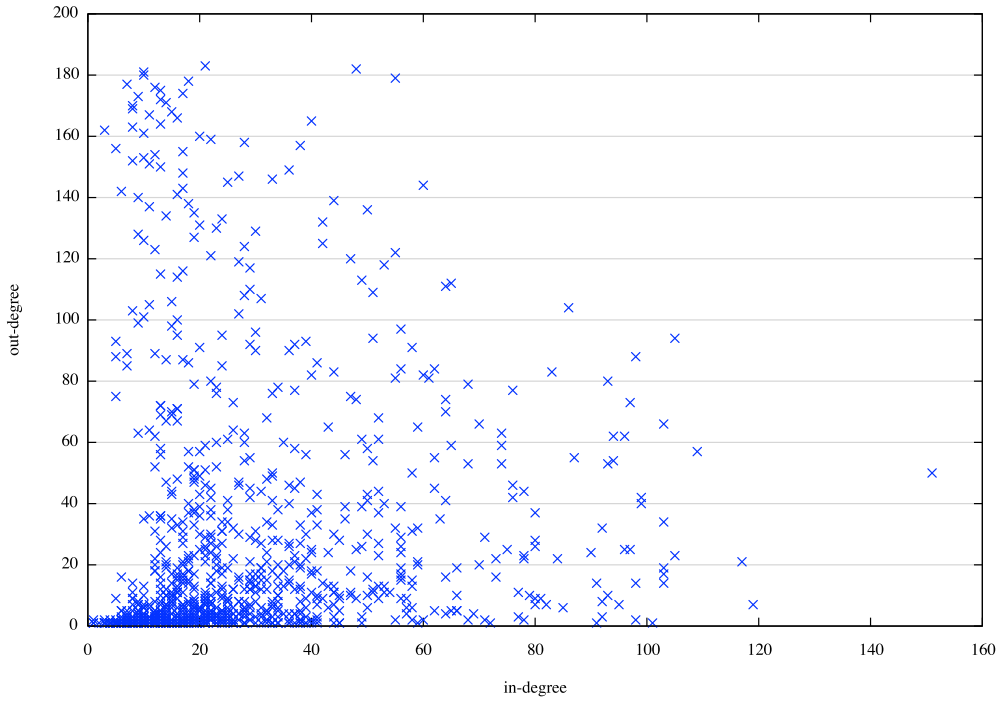


Figure 3.10: Scatter plot of nodes' in-degree vs. out-degree for out-degree scale-free network with high mean clustering coefficient and $1,024$ nodes. There is no correlation between out- and in-degree which implies that absorbing hubs and broadcasting hubs are two different classes of nodes. The Pearson correlation coefficient between in- and out-degree yields a value around 0.03 (This is an example of a single trial.)

against local clustering coefficient (shown in blue). The higher the out-degree of a node the lower its local clustering coefficient. In other words, hubs are low clustered nodes. It is important to mention that this behaviour is observed only in scale-free networks. This correlation between out-degree and local clustering coefficient is not observed in random networks (see Fig.3.9d).

In our scale-free networks the higher the out-degree of a node the lower its local clustering coefficient, and this can be explained by the fact that the high connectivity of hubs decrease the probability that their (out-)neighbours will be connected among themselves. The only situation in which this is possible is in a dense network converging to a fully-connected structure.

As mentioned in the previous section, the distribution of local clustering coefficients for scale-free networks with high mean clustering coefficient exhibit a slow-decaying right tail (see Fig. 3.8b). This tail is related not to well-connected nodes (ie. hubs) but to dwarf nodes, that is, poorly-connected ones, whose probability of neigh-

bours becoming connected among themselves is higher than for hubs (see Fig. 3.9b).

In Fig. 3.6b we showed how the in-degree distribution of out-degree scale-free networks with high mean clustering coefficient exhibits a right-tail which suggests the presence of absorbing hubs coexisting with broadcasting hubs in this type of network. How well clustered are these absorbing hubs compared to other nodes? Figure 3.9b shows that these nodes with massive incoming connections possess a low local clustering coefficient.

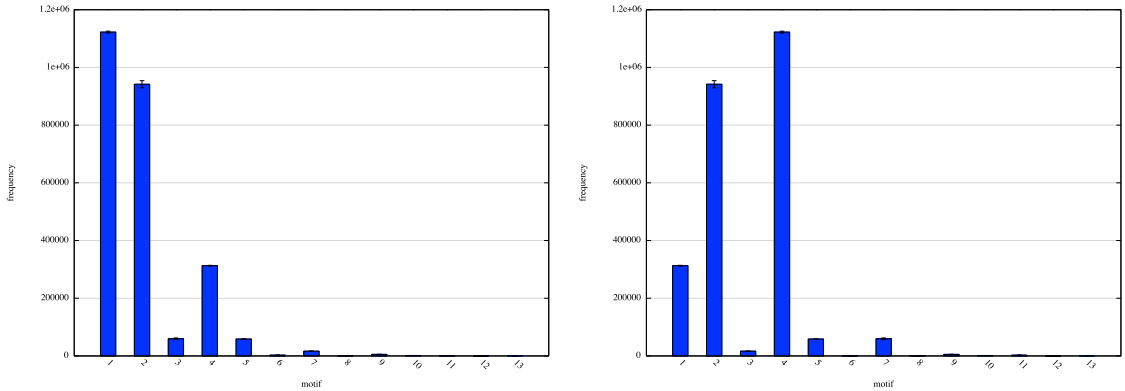
At this point, we might ask whether the absorbing hubs emerging in a scale-free network as it becomes more and more clustered are also the broadcasting hubs. This is not the case, for instance, Fig. 3.10 shows how the in-degree is distributed per out-degree in an out-degree scale-free network with high mean clustering coefficient of size $N = 1,024$. In general, we do not observe any correlation between in- and out-degree, which implies that the emerging absorbing hubs are not as well broadcasting hubs. Lastly, random networks do not exhibit any kind of correlation between in- or out-degree and local clustering coefficient (see Figs. 3.9c and 3.9d)

3.4 Motifs

In this section we will describe the motif profile of our networks. As presented in Sect. 2.1.8, a motif profile is a description of the distribution of the 13 different configurations of 3-node subgraphs in a directed network. We do not consider the motif profile of fully-connected structures as they imply the presence of only one type of motif, namely, motif *XIII*.

Figure 3.11 shows the motif profiles of two types of networks: one is an out-degree scale-free network of size $N = 1,024$ and low mean clustering coefficient, and the other is its transpose, that is, the same type of network but with reversed edges, which yields an in-degree scale-free network. This results in a permutation of motif labels. Reversing the direction of edges changes some motifs, for instance, motif *I* becomes motif *IV* and vice versa, while others might remain the same, for instance, motif *II* and motif *XIII* (see Fig. 2.1 in page 19). Thus, to simplify our discussion, in this section we will consider only motifs of out-degree scale-free networks and random networks. Our observations are valid for in-degree scale-free networks up to a permutation of motif labels.

Figure 3.12 shows the motif profiles of all our scale-free and random networks. Insets show the case for in-degree scale-free networks that result from reversing the



(a) Out-degree scale-free network

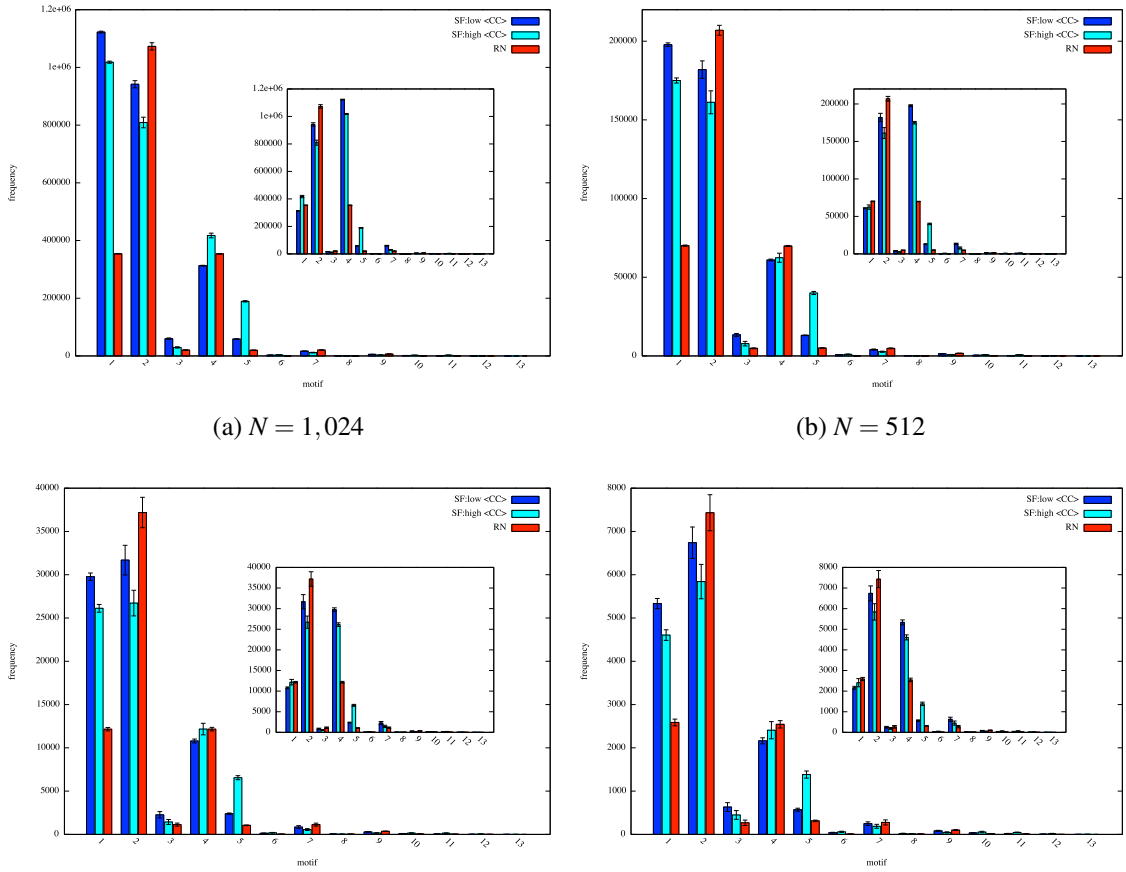
(b) Its transpose, ie. in-degree scale-free network

Figure 3.11: Motif profiles for scale-free networks of size $N = 1,024$ and low mean clustering coefficient. The motif profiles of a directed network and its transpose (same network with reversed edges) are similar up to a change of motif labels (see text). (Error bars represent standard deviations.)

direction of edges of our out-degree scale-free networks.

Scale-free networks abound in motifs *I* and *II*, which are the simplest relationships among nodes taken in threes in a directed network. Motif *I* is also the simplest representation of the "out-hubness" of a node when considering triplets of nodes. It is then natural that it is over-represented in out-degree scale-free networks, which are characterised by the presence of broadcasting hubs. It is also natural that when reversing the direction of edges and when broadcasting hubs become absorbing hubs, the over-representation in motif *I* becomes now an over-representation in motif *IV* (see insets in Fig. 3.12). Thus, motif *IV* is the simplest representation of the "in-hubness" of a node. This motif is also present in out-degree scale-free networks, however not as numerous as motifs *I* or *II*.

For a given directed network, a greater number of motifs *I* and *IV* compared to other motifs of the same network might be a sign of scale-free-ness. As mentioned above, motif *I* might be representing the presence of broadcasting hubs, whereas motif *IV* does it for absorbing hubs. Comparing these motifs to those present in a random network with same size and same number of edges, we observe that motif *I* is more numerous in out-degree scale-free than in random networks due to the presence of broadcasting hubs, whereas motif *IV* is more abundant in in-degree scale-free than in random networks due to the presence of absorbing hubs. In this way, motif-counting might provide a local method to discover the scale-free nature of a directed network,



(a) $N = 1,024$ (b) $N = 512$

(c) $N = 256$ (d) $N = 128$

Figure 3.12: Motif profiles for out-degree scale-free and random networks. (Inset: transpose of out-degree scale-free nets, ie. in-degree scale-free nets.) Motif I is more abundant in out-degree scale-free than in random nets due to the presence of broadcasting hubs, whereas motif IV is more numerous in in-degree scale-free than in random nets due to the absorbing hubs.

which does not require global information of the network as it would be to estimate the degree distribution. However, this hypothesis is beyond the scope of this thesis.

Maximising the clustering coefficient in our scale-free networks yields and over-representation of motif V. This motif captures the most basic triad closure in a directed network. Its over-representation is shown in Fig. 3.12 in the way that motif V is more numerous in out-degree scale-free networks with high mean clustering coefficient compared to their low mean clustering coefficient counterparts and random networks.

In summary, the presence of absorbing and broadcasting hubs is being represented by the abundance in motifs I and IV in out- and in-degree scale-free networks, respectively; whereas the high clustering induced by our algorithm is represented by an

abundance in motif V compared to random networks with same number of nodes and edges.

3.5 Other Network Statistics

In this last section we present other network statistics that can be summarized easier than the properties so far described. These statistics include the mean clustering coefficient, the mean shortest path length, the measure S of small-world-ness, among others.

Table 3.1 in page 61 shows the averaged value (and standard deviations) of some basic network statistics estimated from all networks generated for different types and sizes. These statistics are valid as well for out- and in-degree scale-free networks, that is, they are invariant under inversion of the direction of edges. The network statistics shown in Table 3.1 are:

- Mean clustering coefficient (mean CC). The mean clustering coefficient is estimated for a single network as described in Sect. 2.1.2 and following Eqs. (2.2) and (2.3) in page 13. In Table 3.1 we present the mean of the mean clustering coefficient over all trials (ie. different networks generated) per network type and size along with its standard deviation. Random structures are the type of network with the lowest mean clustering coefficient, whereas the maximum mean clustering coefficient is reached -naturally- by the fully-connected network. In-between these two are scale-free topologies with low and high mean clustering coefficient. As expected, the triad-formation step of our algorithm increases the overall clustering coefficient in scale-free networks, allowing them to reach a mean clustering coefficient larger than random networks and than scale-free networks with no triad-closure. As discussed in Sect. 2.1.5, scale-free networks occurring naturally in the world possess a high level of clustering.

- Mean shortest path length (mean L). As defined in Sect. 2.1.2 the shortest path between nodes i and j in a directed network is the minimum number of (directed) edges needed to traverse to go from node i to node j, and given that the network is directed, it is not necessarily the same as going from j to i. Moreover, this value might diverge even if the network has no disconnected components. This situation occurs when a node is not reachable (or cannot be left once reached) given the direction of edges connecting to it. As shown in Table 3.1, all of our

networks possess a well-defined mean shortest path length. As in the case above, we present the mean of the mean shortest path length along with its standard deviation. Fully-connected topologies possess the lowest value for the mean shortest path length, in this type of network every node is one edge away from each other. However, the shortest path length is not that large for our scale-free and random networks compared to the size of it. For instance, a node in a random network of size $N = 1,024$ is -on average- 2.42 steps away from any other node. It is as if a low mean shortest path length is a property that comes for free in random and scale-free networks.

- Number of edges. As the name suggests, these are the total number of edges in a network. Note that for a given network size our scale-free and random structures possess exactly the same number of edges. Thus, the network structure emerges from a particular permutation of edges. One permutation yields a random network, whereas another one yields a scale-free network; and yet another one yields a scale-free network with high mean clustering coefficient. This fact will be important to keep in mind when we present the collective dynamics that take place in these different topologies. As it is our intention to keep this number equal for scale-free and random networks (for the sake of valid comparisons) the standard deviation is null.

- Mean connectivity/degree (mean Conn). The mean connectivity refers to the mean degree in the network. Our procedure to generate networks keep the mean degree practically the same for every network in a particular type and size. One fact that it is important to keep in mind is that the mean degree for scale-free networks, although existent because it is a finite size structure, is uninformative given the nature of the power-law degree distribution in them. Nevertheless, we consider this value in this type of network in order to use the Eurich model as defined in Sect. 2.2.3.

- Density. This value refers to the number of edges present in a network from all possible ones, and thus it describes if a network is dense or sparse. Its value is obtained by computing $e/[N(N-1)]$ where e is the number of edges in the network. As can be seen in Table 3.1 our scale-free and random networks are sparse. For instance, a node in a random network of size $N = 1,024$ is only connected to 2% of the network. Why do we consider sparse networks? The number of neurons in the human brain has been estimated to be around 10^{11}, whereas the number

of synapses (ie. connections) to lie around 10^{14} (Herculano-Houzel, 2009). The brain is a directed network, thus its density can be computed using the above equation. This yields $10^{14}/[10^{11}(10^{11}-1)] \approx 10^{14}/10^{22} = 10^{-8}$, which implies that the brain is sparse.

- Degree of small-world-ness (S). The small-world property is not a binary one, rather it is a property that occurs in different amounts. In a network, this property is identified by two conditions: a small mean shortest path length combined with a large mean clustering coefficient. In Sect. 2.1.4 we presented a simple equation to know whether a network possesses the small-world property and to what extent. If a network has a value of S greater than unity, then the network exhibits the small-world property. As mentioned above, a low mean shortest path length comes almost for free in our networks, then it is the mean clustering coefficient who is setting the amount of small-world-ness in a network. Varying this parameter has an immediate effect on the small-world property. As expected, our scale-free networks with high mean clustering coefficient reach the largest value of S for all different types considered.

This concludes our presentation regarding the networks that we generated for the purposes of experimenting on them following the dynamics of the Eurich model and its variations. In the upcoming chapters we will present the results of our experiments, and when required we will refer to the network statistics that are relevant for a particular observation.

In the upcoming chapters we will consider two scenarios when considering the dynamics of the system, namely: *i) static* and *ii) dynamic* synapses. In our static synapses scenario we do not implement any plasticity mechanism which would ultimately alter the structure of the network, creating thus a feedback loop between topology and collective dynamics. In our dynamic synapses scenario we do allow such plasticity to take place. In such scenario, we consider three different types of plasticity mechanisms: pair-wise STDP, triplet-wise STDP, and lastly, we introduce a new type of plasticity based on our node-success statistic (see Sect. 2.2.6) that we call *node-success-driven plasticity*.

Type	Subtype	Size	Mean CC	Mean L	Edges	Mean Conn	Density	S
Scale-free	Low Mean CC	128	0.094 ± 0.0044	3.11 ± 0.10	905	7.07	0.055	1.47 ± 0.11
		256	0.077 ± 0.0019	3.04 ± 0.034	2,724	10.64	0.041	1.62 ± 0.098
		512	0.075 ± 0.0011	2.91 ± 0.029	9,074	17.72	0.034	1.83 ± 0.088
		1,024	0.066 ± 0.00057	2.86 ± 0.021	28,481	27.81	0.027	2.07 ± 0.038
	High Mean CC	128	0.24 ± 0.0068	3.40 ± 0.17	905	7.07	0.055	3.53 ± 0.30
		256	0.23 ± 0.0054	3.34 ± 0.09	2,724	10.64	0.041	4.45 ± 0.24
		512	0.24119 ± 0.003	3.19 ± 0.053	9,074	17.72	0.034	5.39 ± 0.21
		1,024	0.20696 ± 0.003	3.11 ± 0.029	28,481	27.81	0.027	5.93 ± 0.20
Random		128	0.056 ± 0.0029	2.68 ± 0.019	905	7.07	0.055	
		256	0.041 ± 0.00084	2.60 ± 0.0037	2,724	10.64	0.041	
		512	0.034 ± 0.00042	2.48 ± 0.00075	9,074	17.72	0.034	
		1,024	0.027 ± 0.00018	2.42 ± 0.00075	28,481	27.81	0.027	
Fully-connected		128	1	1	16,256	127	1	1
		256	1	1	65,280	255	1	1
		512	1	1	261,632	511	1	1
		1,024	1	1	1,047,552	1,023	1	1

Table 3.1: Other network statistics.

Chapter 4

Static Synapses

" If we're interested in network topology, we're interested in how the nodes of a network are connected to each other [⋯] If we're interested in criticality, we're interested in how the network behaves. The two topics are certainly related, but it is possible for non-scale-free networks to exhibit critical behavior and it is possible for scale-free networks to not exhibit critical behavior. The network connectivity affects the critical behavior of that network."

—*Critio* from (Beggs and Timme, 2012)

This chapter is devoted to the *static synapses* scenario, and will serve as a base to develop the *dynamic synapses* scenario, and the three different cases derived from it. Our main motivation to conduct the experiments presented in this chapter was to enquire about the relationship between network structure and collective dynamics. In particular, we were interested in studying how the complexity of a network, identified by properties such as the small-world property and the presence of hubs, influences the onset of critical dynamics when the elements of the system follow the dynamics of the Eurich model.

In other words, our motivation was to study the Eurich model in the context of complex networks. As mentioned in Sect. 2.2.3, this model was suggested as a model of neuronal avalanches, and it allowed to derive some interesting properties of the nature of avalanche durations and their lifetimes. Here, we decided to extend the model to consider not only fully-connected but also scale-free structures, along with random networks to work as a baseline. Scale-free networks are known to be ubiquitous in manufactured and natural systems, therefore it is an interesting underlying structure in which to study model dynamics.

The findings presented in this chapter can be summarised as follows:

- Collective dynamics of complex networks exhibit critical behaviour (Sect. 4.3).
- Scale-free and random networks exhibit extended critical intervals for the control parameter (Sect. 4.2).
- Small-world property boosts network activity (Sect. 4.3.5).
- Scale-free topologies comprise more successful nodes (Sect. 4.3.6).
- Existence of an upper bound for the value of the average node success in fully-connected networks (Sect. 4.3.7).

4.1 Methods

Our results are derived mainly from computer simulations over the network types and sizes described in Chapter 3. As mentioned in Sect. 2.2.3, under the Eurich model every node j possesses a membrane potential, which is updated according to the following equation:

$$h_j(t+1) = h_j(t) + \sum_{i=1}^{N} A_{ij} w_{ij} s_i(t) + I_{ext} \tag{4.1}$$

where A denotes the adjacency matrix; w_{ij} denotes the synaptic weight from node i to node j; $s_i(t)$ represents the state of node i at time t; and I_{ext} denotes external input which is supplied to a node according to the current state of the system. If there is no active node at time t, then a node is chosen uniformly at random and its membrane potential is increased by I_{ext}. If $h_i(t)$ exceeds the threshold θ, then node i becomes active and propagates its activity through its synaptic output. Afterwards, the node is reset. The coupling strength w_{ij} for every node i sending an edge to node j is set according to the equation:

$$w_{ij} = \frac{\alpha}{\langle e \rangle} \tag{4.2}$$

where α is the control parameter of the model, and $\langle e \rangle$ denotes the mean degree of the network.

When starting simulations, all membrane potentials are initialised at random taking values in the interval $(0,1)$, whereas all states are set to inactive. By means of external driving, activity inside the system in the form of neuronal avalanches is guaranteed to occur. However, avalanche sizes and their durations will not always be the same nor can they be predicted.

4.1.1 Assessment of critical behaviour

Our first concern lies on how to determine that the collective dynamics of the system is in the critical state. In Sect. 2.2.1 we mentioned that critical dynamics feature the presence of power-law distribution of events (e.g. size and duration of avalanches). Thus, a simple and straightforward way to look for criticality is to inspect the distribution of avalanche sizes and durations of a system for different values of the control parameter α. If the distribution resembles a power-law distribution then we have reasons to suspect that the system is undergoing a phase transition.

We assess the quality of such a power law through the mean-squared deviation $\Delta\gamma$ from the best-matching power law with exponent γ obtained through regression in log-log scales. Our choice of using this method is due to its simplicity and justified by the asymptotic unbiasedness of the estimation. When this error function is at its minimum, that is, when the data is best approximated by a power-law distribution of avalanche sizes with exponent γ, is when the system is at the critical state.

A power-law fit of the distribution of avalanche sizes is by no means a sufficient condition for a system to be critical, but it is a necessary one. Other tests for criticality that we will carry out in our experiments include:

- Observation of the *trichotomy* of critical behaviour: It has been pointed out the relative ease to produce power-laws through natural or artificial means (Touboul and Destexhe, 2010). In other words, there are power-law distributions without criticality. However, real critical systems exhibit a trichotomy of dynamical regimes that is non existent in other contexts where power-law distributions can be found. These three dynamical regimes are the sub-, super- and critical regimes that we described in Sect. 2.2.2. For a system that features the presence of power-law behaviour to be critical it should exhibit as well this dynamical trichotomy (Beggs and Timme, 2012).

- Analysis of the largest eigenvalue Λ associated to the matrix of synaptic weights: Larremore et al. (2011) studied the spectral decomposition of the weight matrix of a system at criticality, and concluded through analytical inspections along with numerical simulations that the largest eigenvalue of the weight matrix governs the network's dynamic range. When this value is unity, the system is in a critical state and its dynamic range is maximized.

- Analysis of the relationships between critical exponents of avalanche sizes and

their lifetimes: From the theory of critical phenomena, we know that at criticality the distribution of several observables follow power-laws with mathematical expressions linking each other (Sornette, 2004). In particular, there is a power-law positive correlation between avalanche sizes and their lifetimes which only occurs at criticality (Beggs and Timme, 2012; Shew et al., 2015; Bellay et al., 2015).

- Analysis of data collapse: At criticality the dynamics of a system show no particular scale, thus resulting in a fractal geometry of its observables (Sornette, 2004; Kadanoff, 2010). A standard procedure to analyse the critical regime in a model of neuronal avalanches is to observe if avalanches exhibit a fractal structure. To this purpose, the average "shape" of avalanches is estimated by keeping track of the the lifetime of an avalanche and the number of nodes involved at each avalanche step. If the system is at criticality, then avalanches of different sizes would exhibit the same shape up to a scaling function, in such a way that data could be collapsed in order to observe how avalanche shapes resemble each other (Friedman et al., 2012).

Our fist approach to analyse the critical behaviour of our systems will involve estimating the $\Delta\gamma$ function due to its ease to be be implemented in running time. When more formal analyses are required we will use the tests described above.

4.1.2 Numerical Implementation

We divide our experiments in two main stages, namely:

1. *Finding the critical interval.* In this stage, we find the values for which the control parameter α in Eq. (4.2) makes the system reach the critical regime. For this purpose, we examine the values of the parameter α for which the deviation $\Delta\gamma$ reaches its minimum. This results in a function that depends on the control parameter. The shape of the $\Delta\gamma$ function differs across topologies and system sizes. In Sect. 4.2 we describe how this function behaves for different structures. In this stage we considered networks of sizes 128, 256, and 512. We perform 10 trials per experiment, that is, 10 different networks per type and size.

2. *Analysis of dynamics during criticality.* Once we find the interval in which the parameter α leads the system to the critical regime, we re-start the system, but

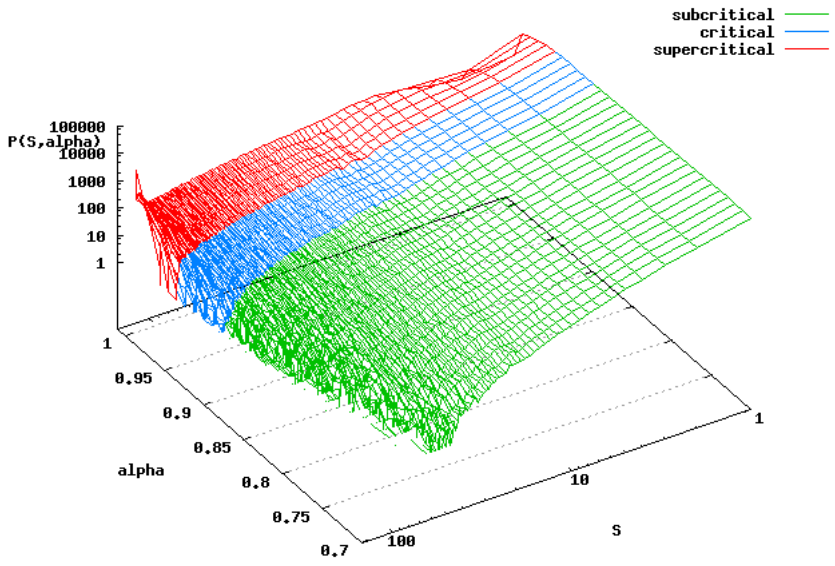
this time the values for α in Eq. 4.2 will be taken uniformly at random from the interval found in the previous step. We call this interval the *critical interval*. In such interval the system is at criticality, and it is here that we perform all further analyses. In this stage we considered 30 trials per experiment for system sizes 128, 256 and 512, whereas for networks of $1,024$ nodes we considered only 10 trials.

Both the relaxation time towards the critical state as well as the sampling time needed to assess criticality depend on the system size. For networks consisting of 128 nodes we allow critical dynamics to set in for one million time steps according to the Eurich model; for networks of 256 nodes we allow critical dynamics to set in for two million time steps, for networks comprising 512 we allow for three million time steps, and finally for networks of $1,024$ elements the dynamics run for four million time steps. This selection of times is appropriate for large events (that is avalanches that extend to the whole network) to take place during simulation time. With this in mind, we expect to have small events (i.e. small avalanches) coexisting with large events (i.e. avalanches that span the whole system). An inspection of the distribution of avalanche sizes after this driving stage shows a distribution that can be approximated by a power-law with a cut-off due to the finite nature of the system. The power-law approximation of such a distribution implies that the system is in the critical regime with very frequent small events coexisting with rare but not negligible large events.

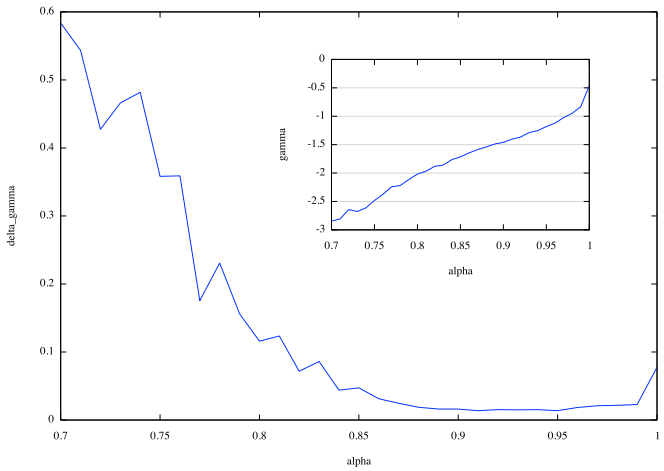
4.2 Critical Intervals

Our first target in this section is to present how the onset of criticality varies through different topologies and system sizes. For that purpose we consider Eq. (4.2), and devote our efforts in finding the critical interval for the six different network classes enumerated at the beginning of Chapter 3. That is, we want to find the values of the control parameter α for which the $\Delta\gamma$ curve reaches its minimum.

Figure 4.1a shows the distribution of avalanche sizes and control parameter α for a fully-connected network with 128 nodes. For some values of this parameter, the distribution does not exhibit avalanches of the size of the system (subcritical regime). As we increase the value of the parameter, the large avalanches become more frequent, and for some values, the distribution becomes bimodal (supercritical regime). However, for an interval of the parameter α (between 0.9 and 0.95) large avalanches are few but not negligible, and coexist with small avalanches which are very frequent. This



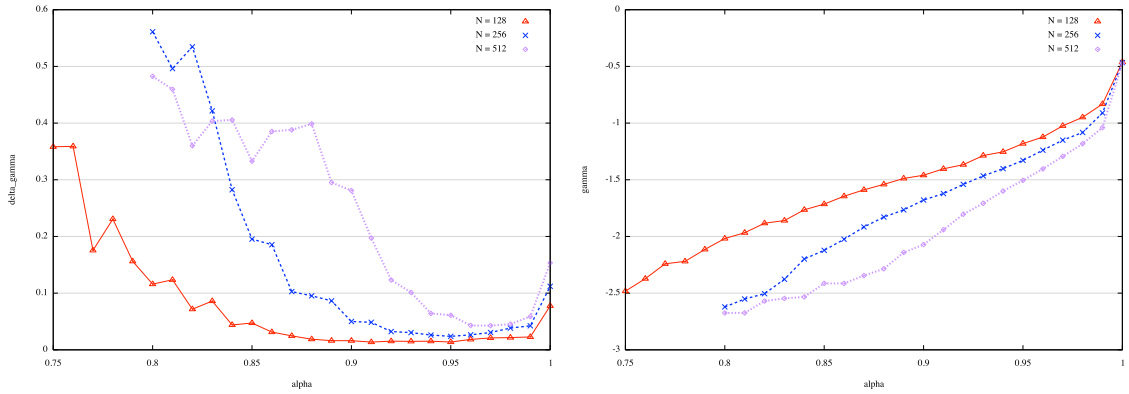
(a) Distribution $P(S,\alpha)$ of avalanche sizes and parameter α



(b) $\Delta\gamma$ curve and exponent γ (inset).

Figure 4.1: Distribution $P(S,\alpha)$ and $\Delta\gamma$ curve and exponent γ for fully-connected network of size $N = 128$ in double logarithmic plots. Error is minimised for values of $\alpha \in (0.9, 0.95)$. The inset shows the exponent γ of the best matching power-law distribution for different values of α. When the error is minimum ($\alpha \in (0.9, 0.95)$) the exponent takes values between -1.5 and -1.

corresponds to the critical regime, which can also be identified with a low error in the power-law fitting shown in Fig. 4.1b. This is the $\Delta\gamma$ curve and shows the values of α for which the distribution can be well approximated by a power-law. Moreover, the exponent γ of the best-fit power-law (associated with a small deviation) lies in the



(a) $\Delta\gamma$ curve. (b) Exponent γ of best-fit power-law.

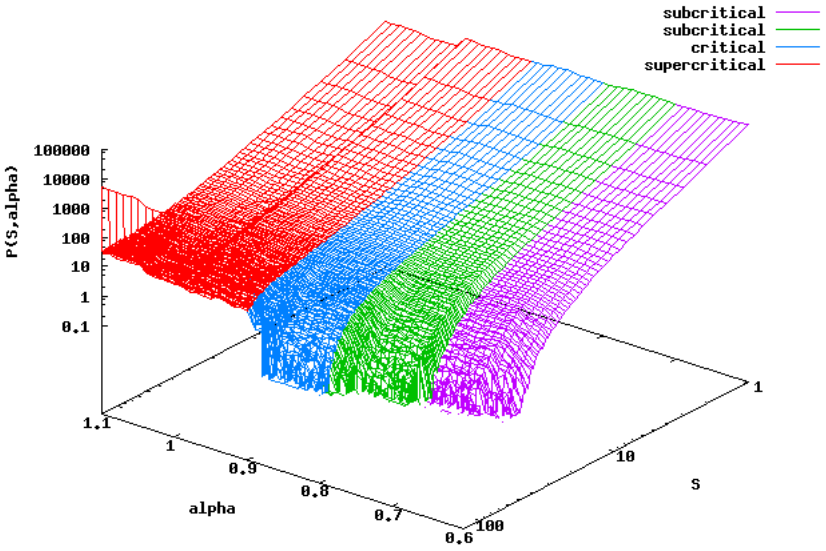
Figure 4.2: $\Delta\gamma$ curve and exponent γ of best-fit power-law for fully-connected networks and different system sizes. The basin in the curve narrows down as system size increases. This suggests that in the thermodynamic limit, $\alpha = 1$ is the value that minimises the deviation from the best-matching power-law.

interval $[-1.5, -1]$ (Fig.4.1b, inset). Recall that *in vivo* experiments as well as analytical derivations have shown an exponent of -1.5 for the power-law approximation of critical dynamics in neural systems (see Sect. 2.2.2).

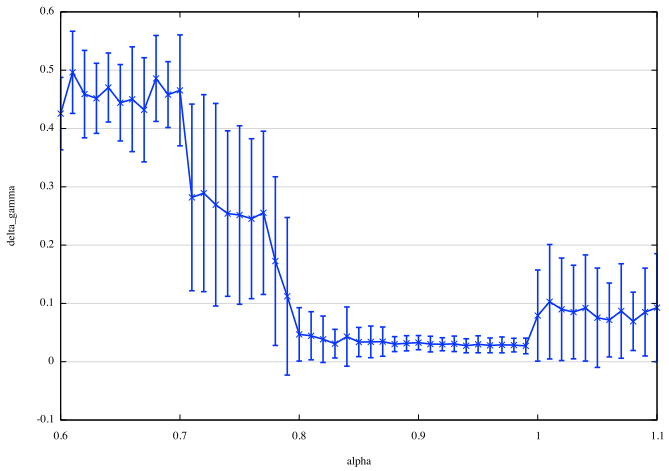
A similar shape of the curve in Fig. 4.2a can be observed as system size increases in fully-connected networks (see Fig. 4.2a). As size increases the system needs more time to reach a critical state identified by a power-law distribution of avalanche sizes. This might explain why in Fig. 4.2a as size grows the minimum value of the power-law deviation is larger.

Moreover, the size of the *plateau* in the $\Delta\gamma$ curve becomes narrower as system size grows. This suggests that in the thermodynamic limit $\alpha = 1$ is the only value that minimises the deviation from the best-matching power-law distribution. Already this result has been derived analytically by Eurich et al. (2002). Lastly, Fig. 4.2b shows the exponent γ of the best-matching power-law approximation varies as the parameter α grows. We know that this class of systems feature an exponent of $\gamma = -1.5$ when the system is at criticality. From Fig. 4.2b we can see how this value is reached for larger values of α as system size increases; for instance, for a network of size $N = 128$ an exponent of $\gamma = -1.5$ is reached for $\alpha \approx 0.9$, whereas for $N = 256$ such a value for γ is reached when $\alpha \approx 0.95$. Again, this suggests that in the thermodynamic limit, such value of the exponent is reached when α equals unity.

Let us consider now our other types of network. The situation for heterogeneous



(a) Distribution $P(S,\alpha)$ of avalanche sizes and parameter α

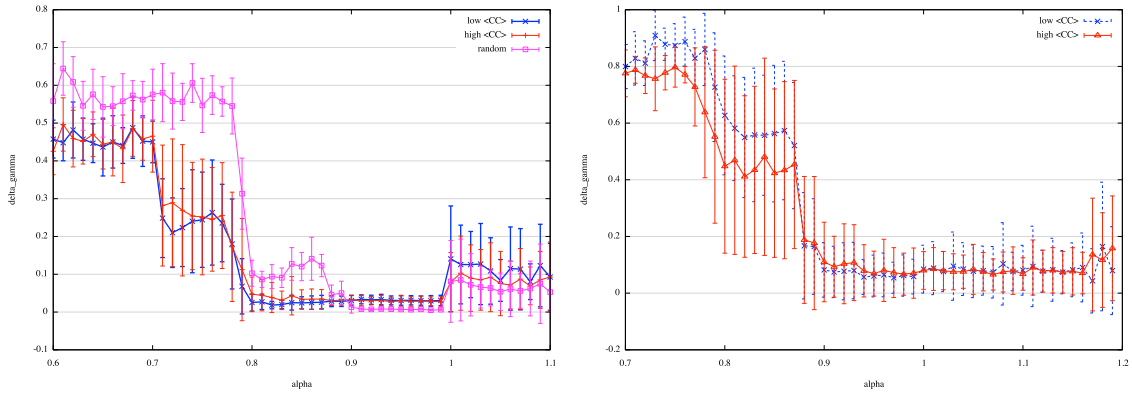


(b) $\Delta\gamma$ curve.

Figure 4.3: Distribution $P(S,\alpha)$ and $\Delta\gamma$ curve for out-degree scale-free networks of size $N = 128$ and high mean clustering coefficient. The $\Delta\gamma$ curve exhibits abrupt transitions and a plateau where the function reaches its minimum. Such plateau covers the interval $[0.8, 1.0]$ indicating an interval of values for which the avalanche size distribution is well approximated by a power-law. Error bars denote standard deviations.

networks looks quite different than for fully connected structures. Here, the $\Delta\gamma$ curve exhibits abrupt changes in its shape suggesting a discontinuous phase transition as the value of the parameter α increases.

Figure 4.3b presents the $\Delta\gamma$ curve for all instances of the out-degree scale-free



(a) Random and out-degree SF nets with high and low mean CC.

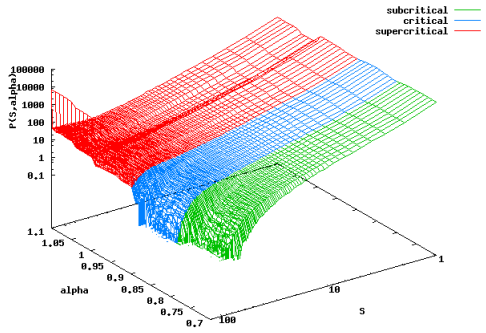
(b) In-degree SF nets with high and low mean CC.

Figure 4.4: $\Delta\gamma$ curve for random and in- and out-degree scale-free (SF) networks with high and low mean clustering coefficient (CC) of size $N = 128$. The function exhibits abrupt transitions as the value of α increases. Moreover, there is a clear plateau where the function is at its minimum. The existence of an interval rather than a single value indicates that the system is at criticality for more than a single value of the control parameter. Error bars denote standard deviations.

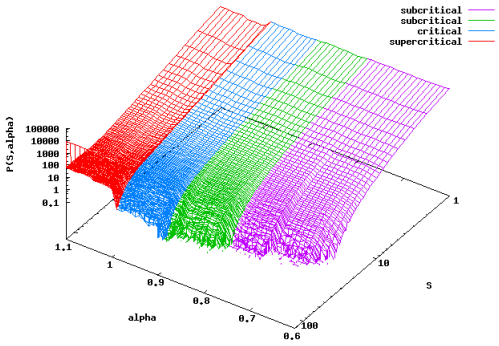
networks with high mean clustering coefficient. This curve is no longer smooth as in Fig. 4.1a but shows pronounced changes in the collective behaviour of the system. Moreover, it exhibits a clear plateau where the deviation from the best-matching power-law reaches its minimum pointing out the interval in which the control parameter α drives the system to the critical state. These abrupt transitions can be observed also in the distribution of avalanche sizes. Unlike the distribution of avalanche sizes for fully-connected networks (Fig. 4.1a) the distribution of avalanche sizes for this type of scale-free networks exhibits abrupt changes as the parameter α grows (Fig. 4.3a).

A plateau is also observed in the other scale-free networks used, namely out-degree scale-free networks with low mean clustering coefficient, and in-degree scale-free networks with high and low mean clustering coefficient. At principle, one might think that this behaviour is particular of scale-free networks given the particular nature of their degree distribution, in which hubs coexist with dwarves. However, we observe a plateau also in random structures bringing down such possibility (see below).

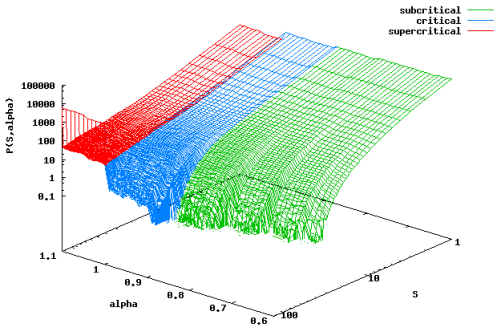
Figure 4.4a shows the $\Delta\gamma$ curve for random and out-degree scale-free topologies comprising 128 nodes. Out-degree scale-free networks with low mean clustering coefficient exhibit a $\Delta\gamma$ curve whose shape is similar to that of its high mean clustered



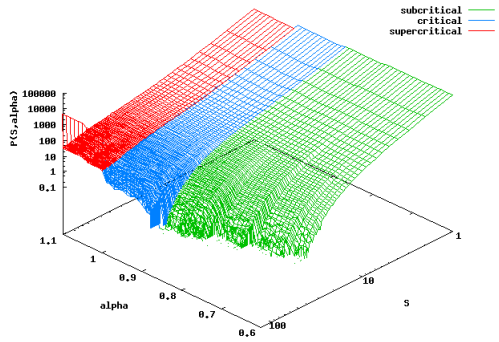
(a) Out-degree scale-free networks with low mean CC.



(b) Random networks.



(c) In-degree scale-free networks with low mean CC.

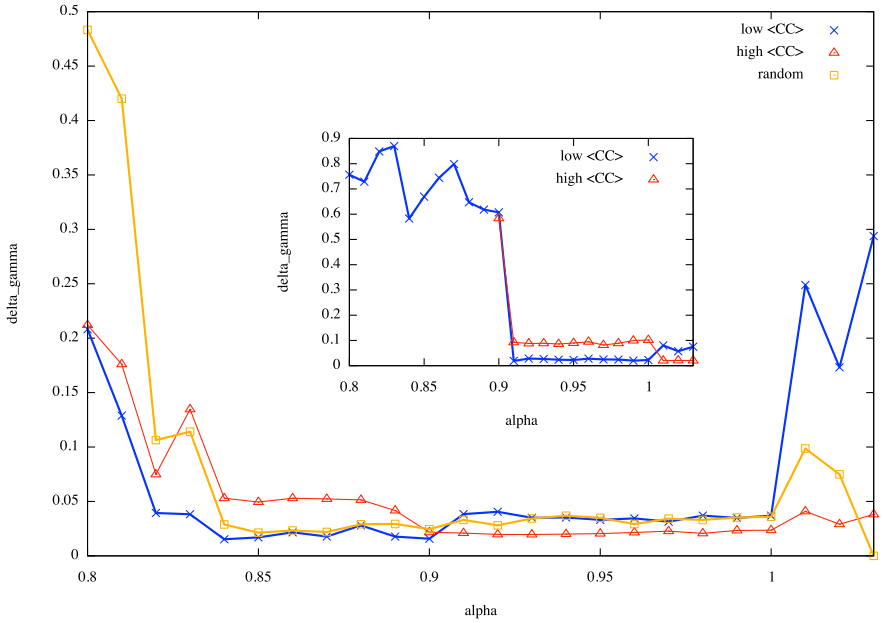


(d) In-degree scale-free networks with high mean CC.

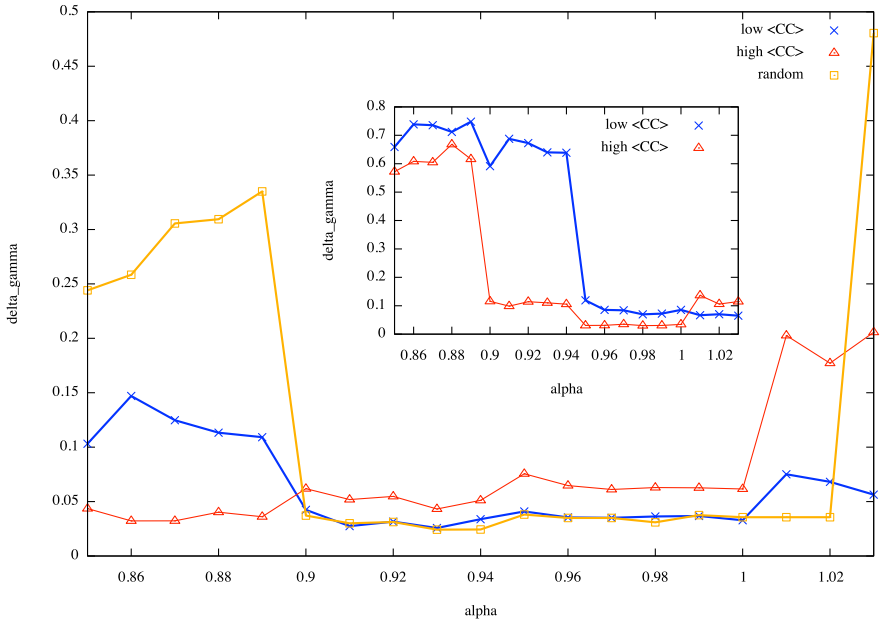
Figure 4.5: Distribution $P(S,\alpha)$ of avalanche sizes and parameter α for networks of size $N = 128$ in double logarithmic plots. Abrupt transitions occur when the system goes from one regime to the other. In Fig. 4.5b (as in Fig. 4.3a) we observe an abrupt transition occurring within the subcritical regime (plotted in purple). Less abrupt transitions also occur within other regimes, however they are not so evident to be plotted differently.

counterpart, and as expected, their distributions of avalanche sizes look similar as well (Fig. 4.5). In such distributions we observe abrupt transitions from one regime to the other. Random networks also present a clear plateau in their $\Delta\gamma$ curve (Fig. 4.4a). However such a plateau is shorter than in scale-free networks. Distribution of avalanche sizes of random networks also reflect these abrupt transitions of the $\Delta\gamma$ curve (Fig. 4.5).

In-degree scale-free networks exhibit a $\Delta\gamma$ curve different from that of random and out-degree scale-free networks (see Fig. 4.4b). Here, as in previous cases we observe

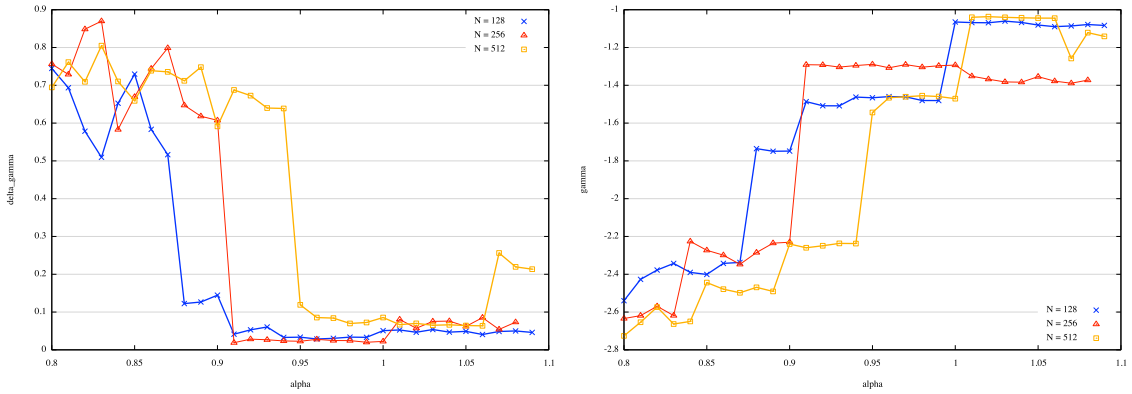


(a) $N = 256$



(b) $N = 512$

Figure 4.6: $\Delta\gamma$ curve for random and scale-free networks for networks of sizes 256 and 512. Abrupt changes in the shape of the curve occur when the system goes from one regime to the other. A plateau appears and within it the function reaches its minimum. Outsets show this behaviour for random and out-degree scale-free networks with high and low mean clustering coefficient, whereas insets show the same for in-degree scale-free networks.



(a) $\Delta\gamma$ curve.

(b) Exponent γ of best-fit power-law.

Figure 4.7: $\Delta\gamma$ curve and exponent γ of best-fit power-law for out-degree scale-free networks with low mean clustering coefficient and system sizes 128, 256, and 512. Similar to fully-connected networks (Fig. 4.2) the plateau in the curve narrows down as system size increases.

a plateau, but this time its length is larger than in other type of topology. We can interprete that this results from the presence of absorbing hubs which are nodes that contribute more than any other type of node to the dynamics of the system by possessing a higher firing rate (see Sect. 4.3.5). As with out-degree scale-free networks, in-degree scale-free structures with high mean clustering coefficient have a $\Delta\gamma$ curve with a shape similar to that of in-degree scale-free networks with low mean clustering coefficient (Fig. 4.4b). Lastly, the distribution of avalanche sizes of in-degree scale-free networks also reflect the abrupt changes found in their $\Delta\gamma$ curve (see Fig. 4.5).

For larger system sizes the situation is not different. We observe as well abrupt regime changes and plateaus in the $\Delta\gamma$ curve. Figure 4.6 shows an example of this behaviour for random and scale-free networks for sizes $N = 256$ and $N = 512$.

Similar to fully-connected networks we observe a scaling behaviour when considering the $\Delta\gamma$ curve and exponent for a particular network type and different system sizes. Here as well the plateau narrows down as the system size increases (Fig. 4.7a), and the value for which the parameter α reaches an exponent of $\gamma = -1.5$ increases with system size (Fig. 4.7b).

We observe an interesting phenomenon in scale-free and random networks when considering the $\Delta\gamma$ curve and the exponent γ associated with the best-fit power-law distribution. A plateau in the $\Delta\gamma$ curve implies that for different values of the parameter α the distribution of avalanche sizes are well approximated by a power-law. However,

the exponent γ does not settle down asymptotically to a specific value, rather is exhibits an increasing behaviour as α grows. For values of α within the plateau of $\Delta\gamma$ we have a power-law distribution of avalanche sizes. The exponent of such power-law is not the same for all the values of α within the plateau, which implies that the distribution of avalanche sizes is well approximated by a power-law with different values of the exponent in which the already known exponent $\gamma = -1.5$ is also included.

Lastly, we might wonder about the reason of the particular shape of the $\Delta\gamma$ curve for our heterogeneous topology, which contrast remarkably from that of fully-connected structures. As mentioned earlier, our first hypothesis was related to the nature of the degree distribution of our scale-free networks. However, after observing the same behaviour in random networks we had to revise our idea. Our next approach was to enquire about what our scale-free and random networks have in common. As mentioned in Sect. 3.5, our scale-free and random networks have in common the number of edges, which results in the same density (see Table 3.1). Thus, we claim that the particular shape of the $\Delta\gamma$ curve in such structures is a result from the sparseness of such networks.

Moreover, the existence of such plateaus might be related to the presence of *Griffith phases* as a result of the collective dynamics of the system. The existence of Griffith phases in models of neuronal avalanches has been suggested as a mechanism by which neural systems extend their "range of criticality", which implies that the critical point is replaced by an extended critical region (Moretti and Muñoz, 2013). This hypothesis has been proven in hierarchical-modular networks, which resemble the structure of brain networks. Here, we suggest that this phenomenon might also be observed in random and scale-free networks due to sparseness. Therefore, neural systems might prefer sparse complex topologies such as scale-free networks in order to extend the critical region, and at the same time obtain the benefits of a scale-free structure such as robustness (Albert et al., 2000). However, testing this hypothesis is beyond the scope of our work.

At the start of this section, we defined the critical interval for a network as the interval of values of the control parameter α for which the coupling strength in Eq. (4.2) leads the system to the critical state in which avalanche sizes are distributed as a power-law. The results presented in this section helped us to find the critical interval per network in order to re-start the system with values for the coupling strengths taken uniformly at random from such an interval. In the next section we show how our networks behave at the critical regime.

4.3 Complex Networks at Criticality

Most of the contents of this and subsequent sections are presented in a paper entitled *The success of complex networks at criticality*[1] co-authored with Dr. Tom Underwood. Therefore, some of the text is taken *verbatim* from our paper.

As mentioned previously, the system has to be fine-tuned in order to reach the critical regime (see Sect. 2.2.3). This is achieved by setting the coupling strengths with values taken from the critical interval as described in the previous section.

In this stage of our experiments we are interested in analysing not only the distribution of avalanche sizes, the $\Delta\gamma$ curve and the exponent γ, but also other observables which include the distribution of avalanche lifetimes, the average node success of the system, the data collapse in avalanche shapes, among others, which also serve as a test for criticality in these systems; as well as, comparing the system's behaviour in different topologies.

4.3.1 Power-law distributions of avalanche sizes

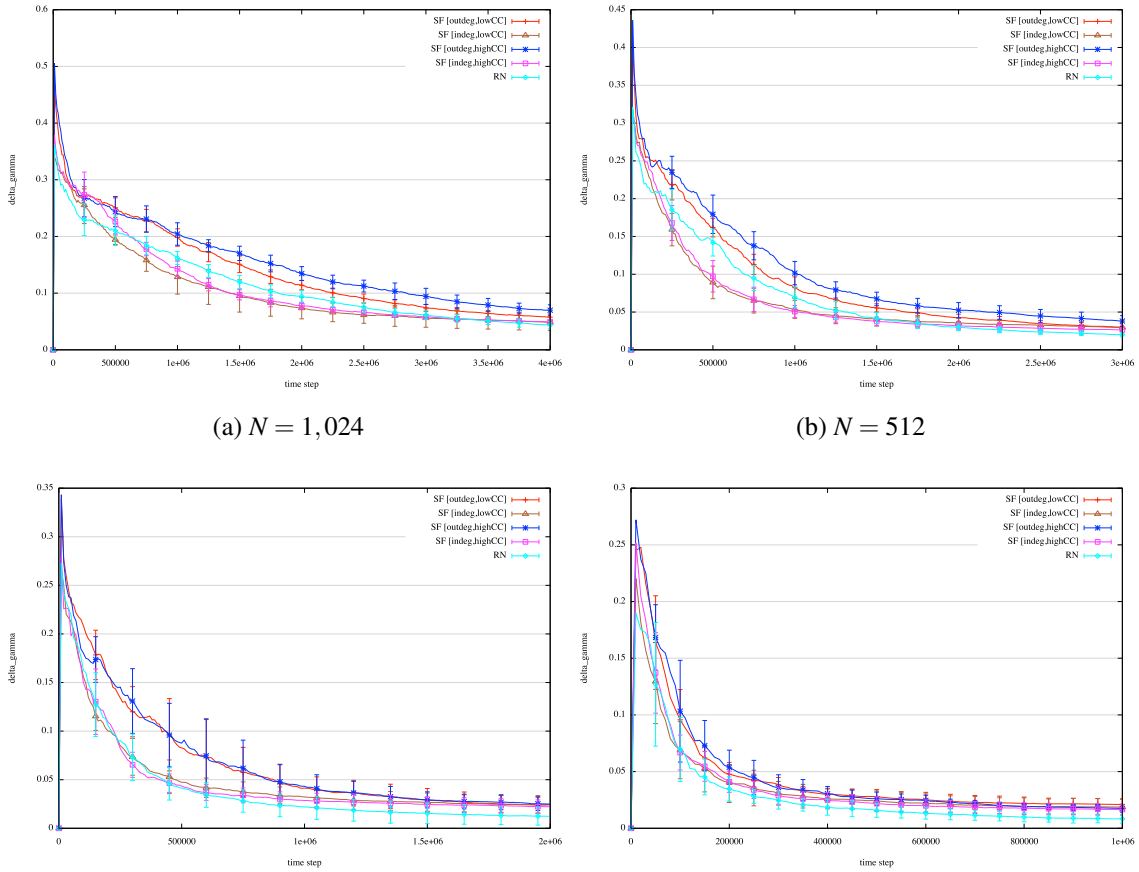
As mentioned in Sect. 4.1.1, we assess the quality of the power-law approximation to the distribution of avalanche sizes by estimating the deviation from the best power-law fit. When such an error function reaches a minimum value of less than or equal to 0.05, we consider the event-size distribution as well approximated by a power-law.

Fig. 4.8 shows the power-law fitting error as a function of simulation time for the distribution of avalanche sizes for all scale-free and random networks. This figure shows the deviation ($\Delta\gamma$ curve) of our data from the best matching power law with exponent γ (see Sect. 4.3.2). In this figure, we present mean values and standard deviations of $\Delta\gamma$ obtained from the different realisations of our experiments.

The evolution of the $\Delta\gamma$ curve for all systems considered exhibit a descending trend which reveals the power-law nature of the distribution of avalanche sizes, a phenomenon that relates the collective dynamics of such systems to criticality. Moreover, some scaling behaviour can be observed as well in Fig. 4.8. In the four system sizes considered we observe a pattern regarding the behaviour of the curve. For instance, in all cases considered out-degree scale-free networks take longer than in-degree scale-free networks to reach the critical state.

Figure 4.9 shows the distribution of avalanches sizes for all scale-free and random networks at criticality (ie. when $\Delta\gamma \leq 0.05$) in log-log plots. This visualisation allows

[1] http://arxiv.org/abs/1507.07884



(a) $N = 1,024$ (b) $N = 512$

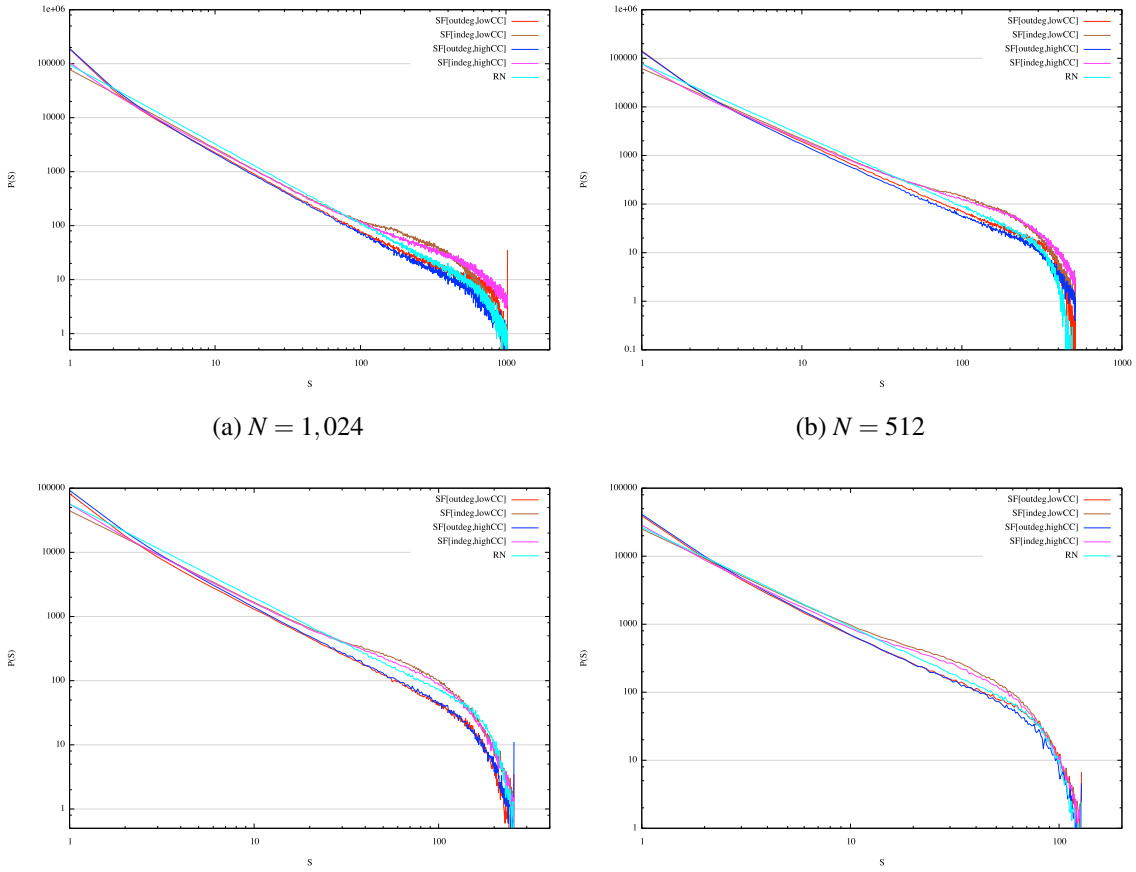
(c) $N = 256$ (d) $N = 128$

Figure 4.8: $\Delta\gamma$ curves showing the deviation from power-law matching per time step for scale-free and random networks of all sizes. The curve exhibits a descending trend as time evolves which implies that the system reaches a critical state after some time. Out-degree scale-free nets converge faster to the critical state, whereas random nets exhibit the slowest convergence. Error bars denote standard deviations.

us to identify the straight line that is the hallmark of power-law distributions. However such a straight line can only be identified for a certain interval of the distribution and then finite-size effects take place resulting in a cut-off in the distribution. We show the averaged value of all realisations, but we do not present error bars in Fig. 4.9 in order to make its presentation more accessible.

Behaviour that can be observed in the deviation function ($\Delta\gamma$ curves in Fig. 4.8) can also be observed in the distribution of avalanche sizes. For instance, the shape of the distribution of avalanche sizes is similar for out-degree scale-free networks. Likewise, in-degree scale-free networks possess a similar avalanche size distribution shape.

In the next section we present how this similarity also extends to the value of the



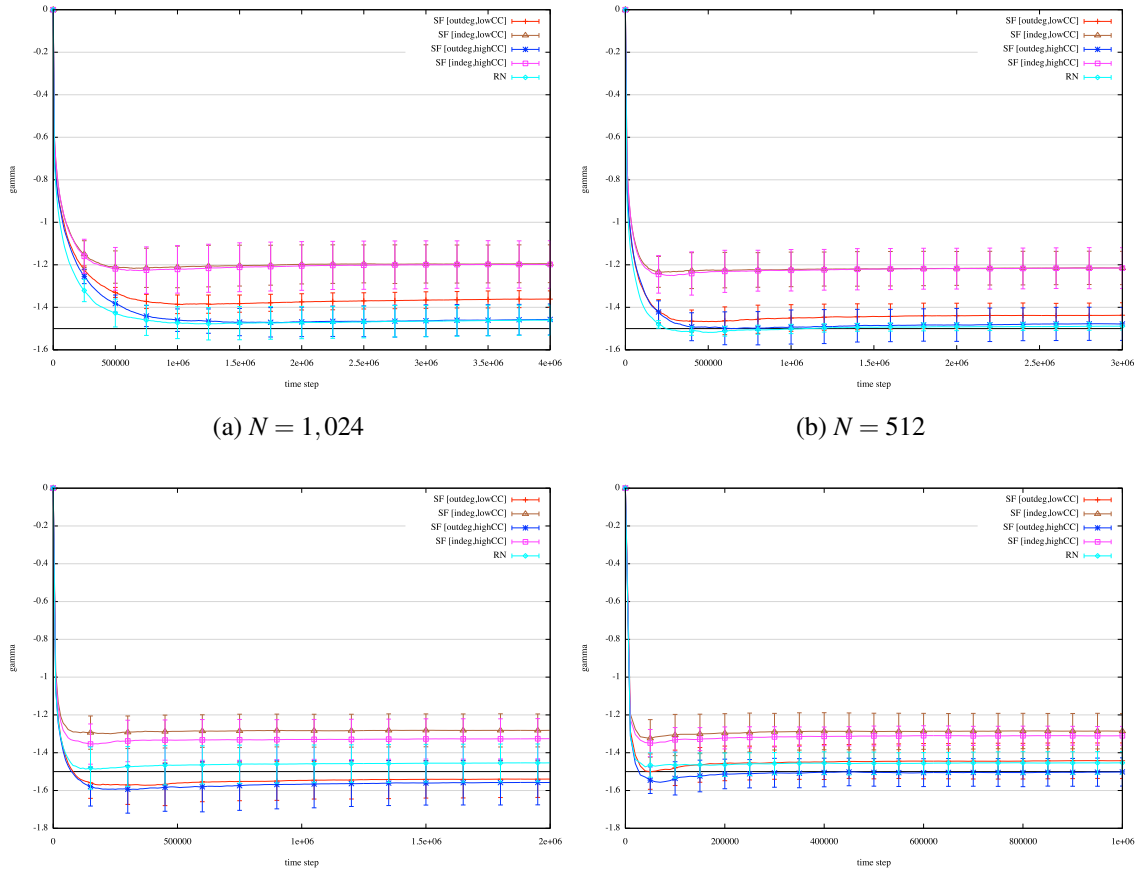
(a) $N = 1,024$ (b) $N = 512$

(c) $N = 256$ (d) $N = 128$

Figure 4.9: Double logarithmic plot of the distribution of avalanche sizes for scale-free and random networks of all sizes considered. The straight-line shape extending for more than two decades in these plots imply an appropriate power-law approximation for the distribution. This is a hallmark of criticality. We present only mean values for the sake of clarity.

exponent γ. As far as avalanche size distributions can tell us, out-degree scale-free networks with high and low mean clustering coefficient possess a similar collective behaviour. This behaviour differs from their in-degree counterparts as well as from random networks.

Along with the observation of power-law statistics of avalanche sizes, we test for criticality in our systems by inspecting the value of the largest eigenvalue Λ associated to the matrix W of synaptic weights w_{ij}. Larremore et al. (2011) showed that a system at criticality exhibits a largest eigenvalue Λ of unity. In Table 6.1 in page 158 we present the value of the largest eigenvalue Λ for the weight matrices with static synapses used here (column Λ_{static}). As it can be seen in this table, the value of Λ is



(a) $N = 1,024$ (b) $N = 512$

(c) $N = 256$ (d) $N = 128$

Figure 4.10: Evolution of the exponent γ of the best-fit power-law distribution of avalanche sizes. The value of the exponent does not exhibit fluctuations for most of the simulation time. For all networks shown, the exponent lies in the interval $[-1.60, -1.05]$. In-degree scale-free nets with low and high mean CC reach a similar exponent. The same occurs with out-degree scale-free nets, which suggests that the exponent is sensitive to the presence of absorbing or broadcasting hubs.

very close to unity (due to finite-size effects) in this type of systems.

4.3.2 Critical exponents for distributions of avalanche sizes

The exponent of the power-law approximation of the distribution of avalanche sizes at criticality varies across network structures with some similarities shared for certain network classes. Figure 4.10 shows the evolution of the value of the exponent γ for scale-free and random networks of all sizes considered. It can be seen that the value of the exponent does not exhibit any fluctuations for most of the simulation time sug-

gesting that although the $\Delta\gamma$ curve has not reached its minimum, the distribution of avalanche sizes possess a steady exponent from the early stages of the simulation.

As with the $\Delta\gamma$ curve and the distribution of avalanche sizes, out-degree scale-free networks with high and low mean clustering coefficient achieve similar exponents. Likewise, in-degree scale-free networks obtain an exponent that is similar to each other (high and low mean clustering coefficient). In all cases considered the value of the exponent γ lies in the interval $[-1.60, -1.05]$, which seem to be at odds with results regarding fully-connected networks (Eurich et al., 2002) and experimental observations (Beggs and Plenz, 2003) in which the exponent takes a value of -1.5. However, this contrast might be due to finite-size effects.

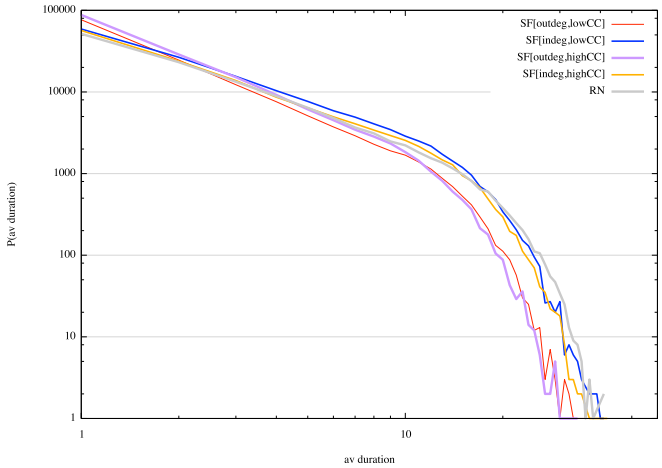
For any given system size, the exponents of in-degree scale-free networks with high mean clustering coefficient differ from the exponents of out-degree scale-free networks with high mean clustering coefficient. The same behaviour is observed for networks with low mean clustering coefficient, that is, the exponent of in-degree scale-free networks differ from their out-degree counterparts. This implies that the action of reversing the direction of edges (which yields a transposed adjacency matrix) has an effect on the value of the exponent γ.

Moreover, the nature of the network's structure seems to be captured by the values of the exponent γ. In-degree scale-free networks with low mean clustering coefficient take values from the same interval as the in-degree scale-free networks with higher mean clustering coefficient. In the same spirit, out-degree networks with low mean clustering coefficient take values from the same interval as their high mean clustering coefficient counterparts. This behaviour implies that the presence of absorbing (similarly, broadcasting) hubs seem to determine the range of values that the exponent γ can take independently of the level of clustering (and consequently of small-world-ness) that the network possess.

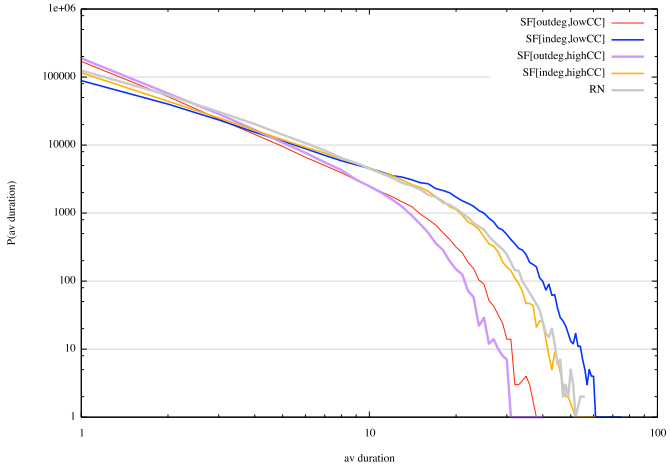
The aforementioned behaviour of the exponent γ for scale-free networks is observed in all the system sizes considered, thus implying an underlying scaling nature of the dynamics of the model.

4.3.3 Power-law distributions of avalanche lifetimes

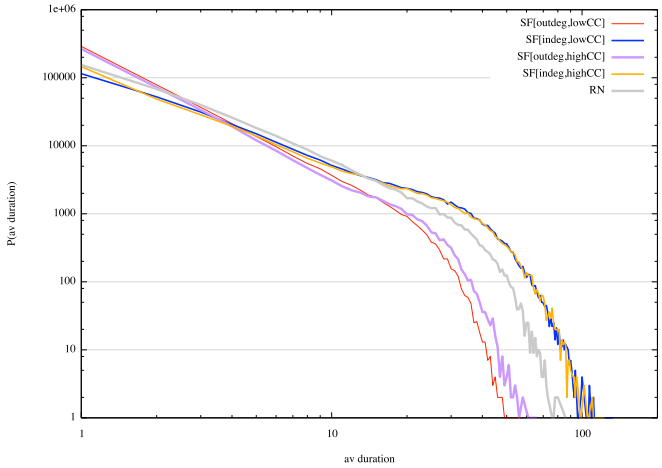
It is known from criticality theory that a system at criticality exhibits power-law distributions in more than one observable (e.g. avalanche sizes and durations, average dynamic correlation versus distance, etc.) (Stanley, 1999; Beggs and Timme, 2012).



(a) $N = 128$



(b) $N = 256$



(c) $N = 512$

Figure 4.11: Double logarithmic plot of the distribution of avalanche lifetimes for scale-free and random networks of all sizes. A straight line can be observed in the distribution before finite-size effects take place in the form of a cut-off. Power-law distribution of avalanche lifetimes is also a hallmark of criticality.

Therefore, we would expect a power-law approximation of the distribution of avalanche durations as a proof of critical behaviour. This was reported previously in experiments and models based principally in fully-connected structures (see Sect. 2.2.2). Here, we report that complex networks also exhibit this behaviour at criticality. Figure 4.11 shows the distribution of avalanche lifetimes.

Moreover, it is known that at criticality the distribution of avalanche lifetimes and their sizes will obey a power-law correlation that associate their exponents in a mathematical relationship (Beggs and Timme, 2012; Friedman et al., 2012; Bellay et al., 2015; Shew et al., 2015). Thus, another test for criticality would require us to look for such a power-law correlation between avalanche sizes and their lifetimes. In Fig. 4.12 we show in double logarithmic scales the distributions of avalanche sizes and lifetimes along with the best-fit approximation (red line).

The plots in Fig. 4.12 show a linear trend in log-log scales between avalanche lifetimes and sizes implying, thus, a power-law relationship between these two observables, which can be expressed as $S \sim D^{\beta}$, where β is the exponent of the power-law correlation.

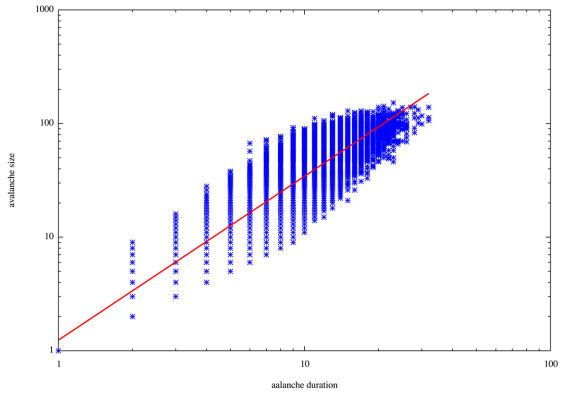
4.3.4 Avalanche shapes and data collapse

"If these cascades, or avalanches, are truly critical then there should be some way to capture a relationship between the avalanches in a fractal way."

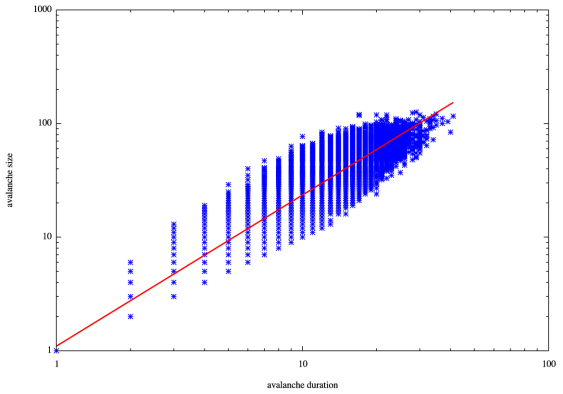
—*Critio* from (Beggs and Timme, 2012)

In this section we will have a look at the shape of the avalanches in all our different network types and sizes. The shape of an avalanche is defined as the number of nodes involved in an avalanche per time step, or better said, per avalanche time (or step). For example, in the current model for any network type and size, all avalanches start with one node, this corresponds to the first avalanche step. In subsequent time steps we record the number of nodes that become active and average them until the avalanche stops. There are avalanches that involve a single node and therefore occur in a single time step, but also there are avalanches that span to the whole network and that take more than one time step to develop. By the end of each simulation we end up with the number of nodes that on average become active in the second avalanche time, the third avalanche time, and so on.

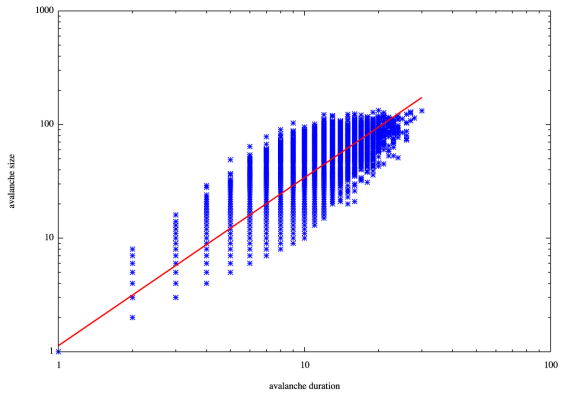
The plot of avalanche step and the average numbers of nodes involved per avalanche step defines the shape of the avalanche for a given network. The maximum avalanche



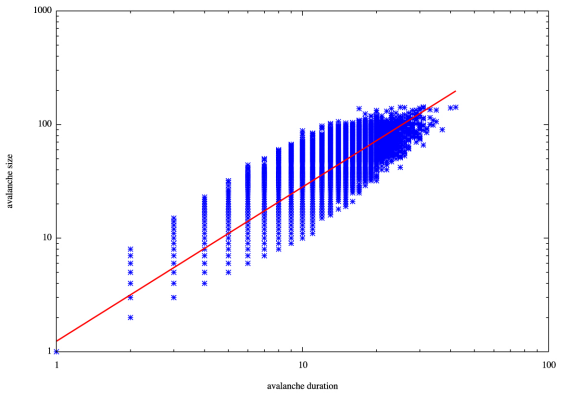
(a) Out-degree SF net with low mean CC. ($\beta = 1.44$)



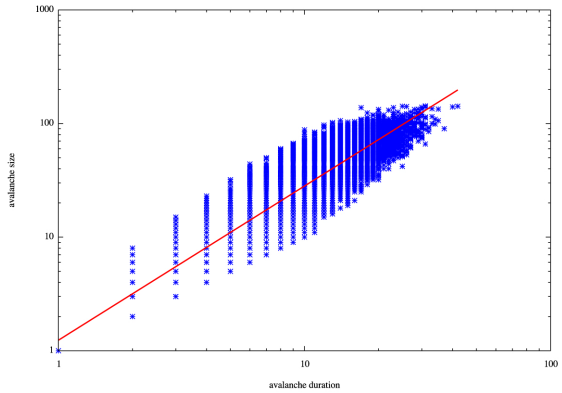
(b) In-degree SF net with low mean CC. ($\beta = 1.32$)



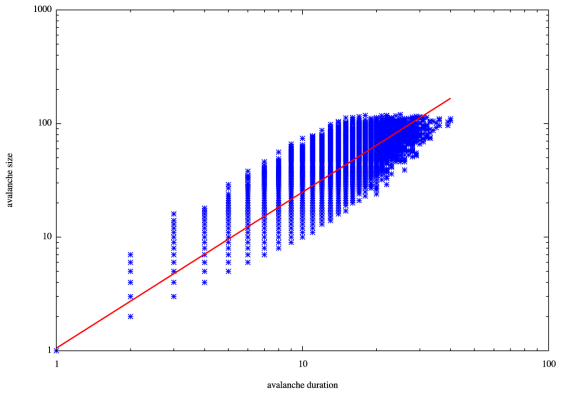
(c) Out-degree SF net with high mean CC. ($\beta = 1.47$)



(d) In-degree SF net with high mean CC. ($\beta = 1.35$)

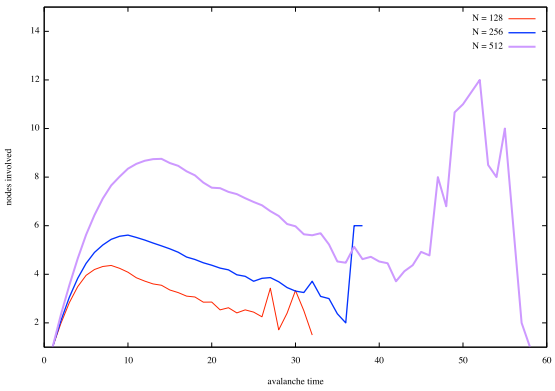


(e) Random net. ($\beta = 1.35$)

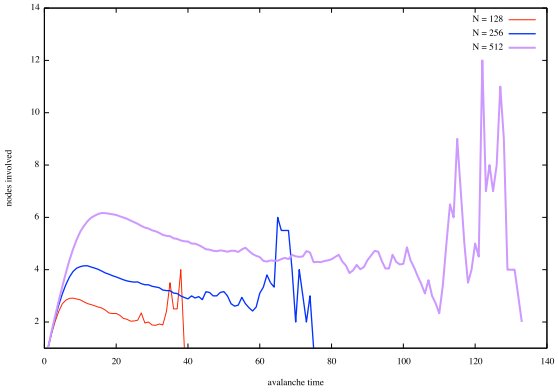


(f) Fully-connected net. ($\beta = 1.37$)

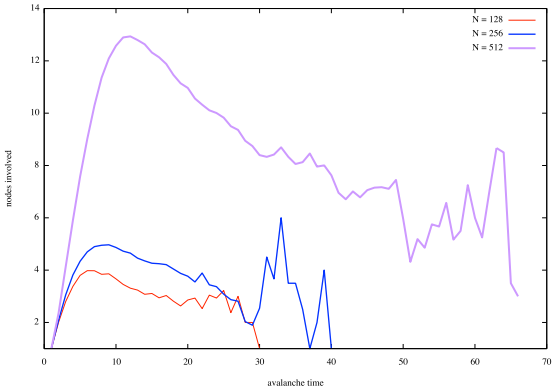
Figure 4.12: Linear relationship on logarithmic axes between avalanche sizes and their lifetimes for networks of 128 nodes. This behaviour implies a power-law correlation of the form $S \sim D^{\beta}$ which is a signature of criticality. In blue dots we present our data whereas in red we present the best-fit approximation.



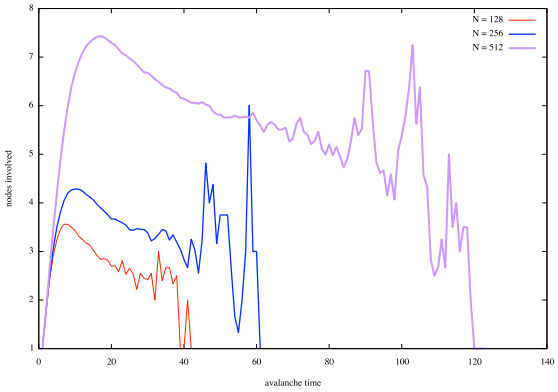
(a) Out-degree SF nets with low mean CC.



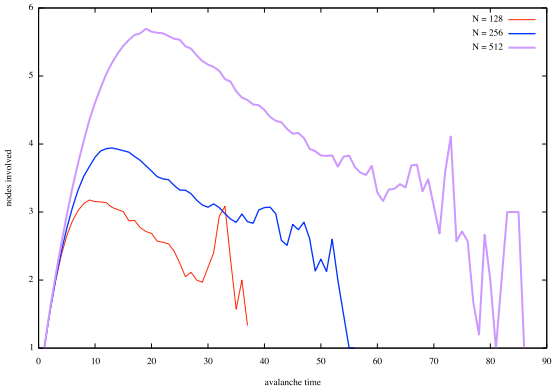
(b) In-degree SF nets with low mean CC.



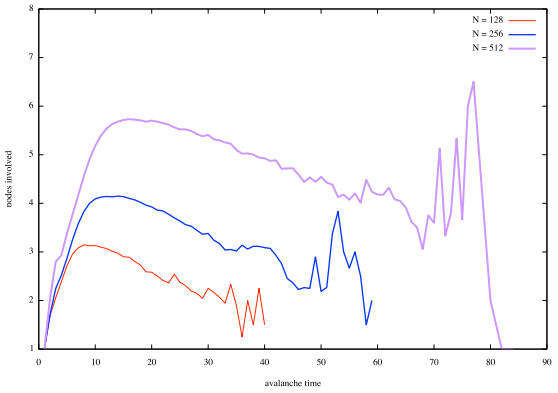
(c) Out-degree SF nets with high mean CC.



(d) In-degree SF nets with high mean CC.



(e) Random nets.



(f) Fully-connected nets.

Figure 4.13: Average avalanche shapes for all networks and sizes considered. The similarity of shapes among different sizes for a given network type suggests the possibility of data collapsing in these systems. This fact implies that avalanches possess a fractal structure which is also a signature of criticality. (We present only mean values for the sake of clarity.)

duration corresponds to the largest number of time steps in which an avalanche takes place for a given network and size.

Figure 4.13 shows the average shape of avalanches for all types of networks and sizes considered in our experiments. For a given network type, we can observe a scaling behaviour occuring among different system sizes in the shape of avalanches. Regarding the nature of these avalanche shapes, in-degree scale-free networks exhibit longer-lasting avalanches than other system types. All avalanches start with a single node becoming active, then the shape of the avalanche shows a growing trend until it reaches a maximum denoting a large number of nodes involved during avalanches. Afterwards, the shape shows a decreasing trend, and towards the end it exhibits large fluctuations in which the avalanche shape might reach a high peak implying that a large number of nodes become active towards the end of the avalanche.

As mentioned in the last paragraph, a scaling behaviour can be appreciated when observing the shapes of avalanche resulting from different system sizes in Fig. 4.13. For instance, consider the shapes shown in Fig. 4.13a for out-degree scale-free networks with low mean clustering coefficient. Here, shapes look very similar and this fact suggests a fractal structure occurring in the avalanches of this particular type of network. This implies that avalanches in system sizes larger than 128 possess a structure similar to that of smaller sizes.

It is known from the theory of critical phenomena that systems at criticality exhibit self-similar cascades of activity which is also related to the scale-invariance of the distribution of events (Sornette, 2004). Data collapse refers to the possibility of re-scaling observables such as avalanche shape in a system at criticality in order to reveal the fractal nature of its cascading activity (Beggs and Timme, 2012; Friedman et al., 2012). This also predicts the behaviour of larger system sizes than the ones considered. As expected, our systems exhibit this feature which accounts as another proof of their state of criticality. We present an example of data collapse taken from our out-degree scale-free networks with low mean clustering coefficient in Fig. 4.14. Here, the collapse of curves into each other suggest a fractal nature on the avalanche dynamics of the system.

The results presented in the previous sections serve as a proof that our systems are indeed at criticality, and therefore their power-law behaviour does not result from an statistical artifact. In the following sections we describe how the network properties of these systems give rise to particular behaviour at criticality.

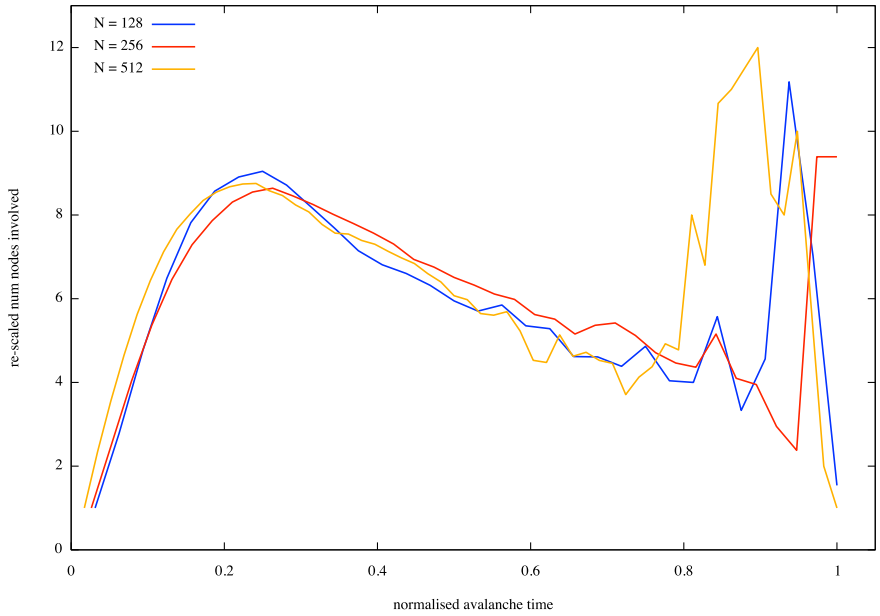


Figure 4.14: Data collapse for avalanche shapes from out-degree scale-free nets with low mean CC (Fig. 4.13a). Data collapsing reveals the fractal nature of avalanches in a system at criticality. (We present only mean values for the sake of clarity.)

4.3.5 Small-world property boosts network activity

The *firing rate* of a neuron is defined in several ways. The simplest of which is also known as *spike-count rate* and is obtained by counting the number of spikes that occur during a trial and dividing by the duration of such trial (Dayan and Abbott, 2001). In this thesis, we use this definition of firing rate and the duration of the trial is taken to be the total simulation time (see Sect. 4.1).

The small-world property affects the rate of firing of nodes comprising a network. Fully-connected networks are structures in which all nodes exhibit a similar firing rate, giving rise to a well defined mean and variance (see Fig. 4.15) unlike scale-free networks in which the variance of the firing rate seemingly diverges, and thus its mean cannot characterise the network activity. In fact, as we will describe in this section, in this latter type of structure, absorbing hubs possess a total spike count that can far exceed the mean spike count of nodes in fully-connected networks. We pose several questions regarding the relationship between network structure and dynamics. The first of these questions is: are the nodes with higher local clustering coefficient those that spike more often, that is, do better-clustered nodes fire more? Surprisingly, nodes with low local clustering coefficient exhibit a larger spiking rate than more clustered nodes.

Figure 4.16 shows this behaviour for scale-free and random networks of all sizes.

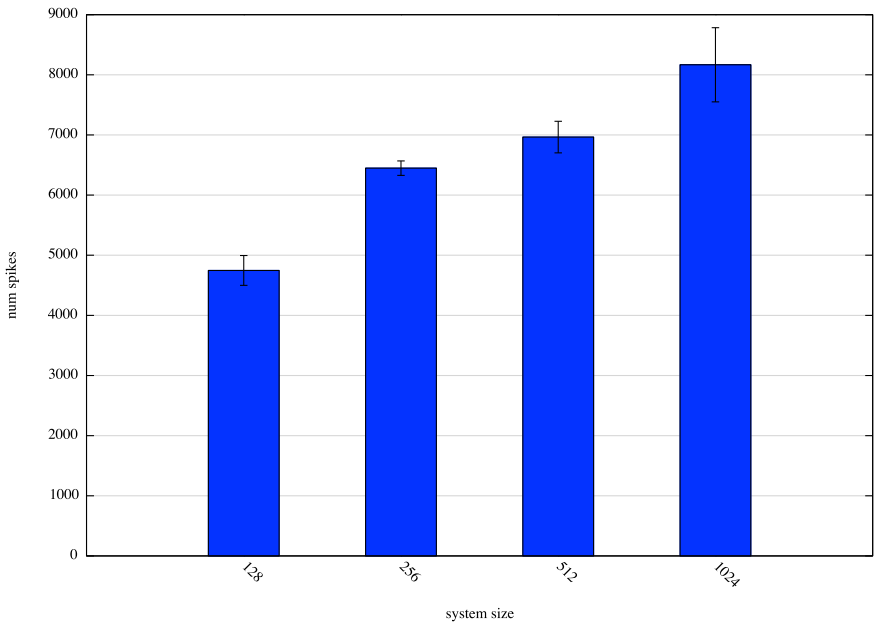


Figure 4.15: Average spike number per node in fully-connected nets for all sizes considered. In fully-connected nets nodes spike less than in random and scale-free nets (see text). We believe this is due to a phenomenon called *spike jamming* that we describe in Sect. 4.3.7. Error bars denote standard deviations.

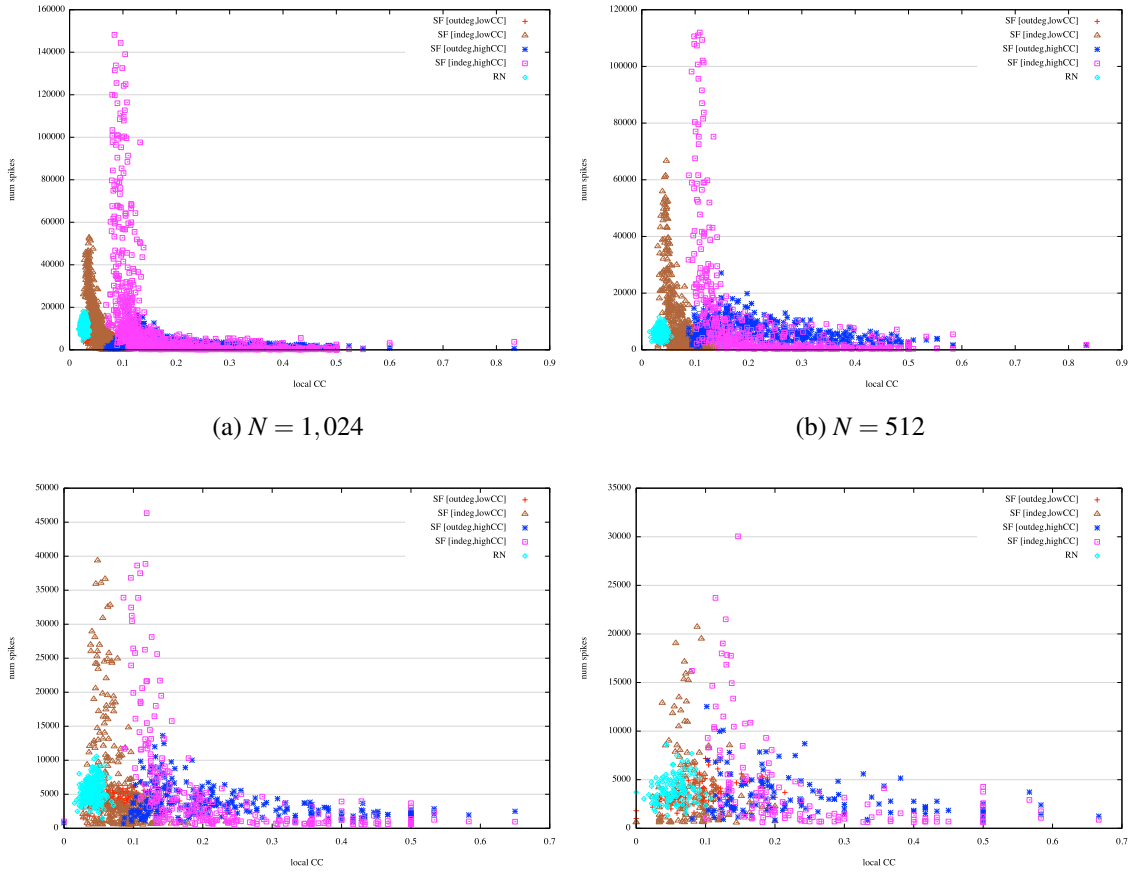
Here we show not only that low locally clustered nodes fire more but also that in in-degree scale-free networks nodes can fire more than in any other type of structure.

A question that arises at this point is the following: Are those low locally clustered nodes who spike so frequently in in-degree scale-free networks the absorbing hubs? Or are they dwarf nodes?

Recall that in Sect. 3.3 we pointed out that hubs (either absorbers or broadcasters) are in general low locally clustered. Thus, we have reasons to suspect that well-connected nodes spike more than any other type of node. In Fig. 4.17 we present how the two different degree distributions (in and out) are related to spiking activity in scale-free networks. Figure 4.17a shows this behaviour for low mean clustered scale-free networks, whereas Fig. 4.17b shows it for high mean clustered networks.

For the case of out-degree scale-free networks, that is, networks that possess broadcasting hubs we do not observe any correlation between node out-degree and firing activity (red triangles in Fig. 4.17). This is valid for out-degree scale-free networks with low and high mean clustering coefficient, and all system sizes although in Fig. 4.17 we only present the case for $N = 1,024$.

In in-degree scale-free networks, that is networks that include absorbing hubs, we

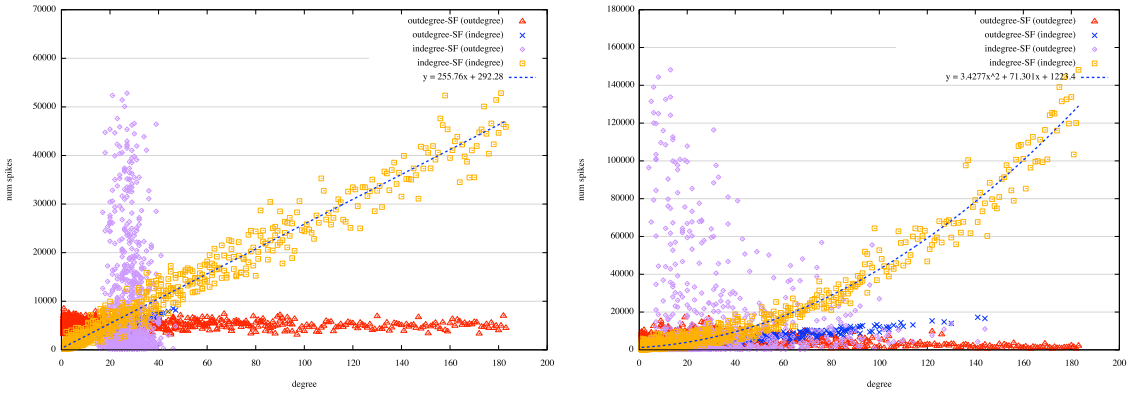


(a) $N = 1,024$ (b) $N = 512$

(c) $N = 256$ (d) $N = 128$

Figure 4.16: Scatter plot of total spike count per local clustering coefficient for scale-free and random networks of all sizes. Low clustered nodes are responsible for spiking more frequent in scale-free networks. An in-degree scale-free structure with high mean clustering coefficient is a configuration in which nodes spike more often (purple squares in all plots). This is an example of a single trial.

observe a positive correlation between node in-degree and spiking (Fig. 4.17, orange squares). This behaviour occurs in all system sizes. Interestingly, the behaviour of in-degree scale-free networks with low mean clustering coefficient differ from the behaviour of the same type of network with high mean clustering coefficient. Both exhibit a positive correlation between in-degree and total number of spikes, nevertheless low mean clustered networks exhibit a linear trend, whereas high mean clustered ones exhibit a non-linear trend. For in-degree scale-free networks with low mean clustering coefficient, the Pearson correlation coefficient (PCC) yields a value around 0.98, whereas for the same type of network with high mean clustering coefficient the PCC is around 0.94 due to the non-linearity of the trend.



(a) Low mean CC (b) High mean CC

Figure 4.17: Scatter plot of degrees vs. spike count. In in-degree scale-free nets the absorbing hubs spike more than any other type of node, and their in-degree is correlated with the amount of spikes fired. For networks with higher mean clustering coefficient (CC) this correlation is quadratic (b, orange squares). We show this behaviour for a network of size $N = 1,024$.

We inspect how in-degree goes from a linear correlation to a quadratic one as the mean clustering coefficient grows. Figure 4.18 shows the relationship between in-degree and total number of spikes for in-degree scale-free networks and different values for the mean clustering coefficient. As the clustering coefficient grows the positive correlation between in-degree and spiking exhibits a clearer quadratic trend.

This suggests that as a network becomes more clustered (and in consequence, more *small-worldly*) the activity of their nodes exhibit a more quadratic dependency on the node's in-degree until it saturates. We do not explore this hypothesis in this thesis, but it is the direction of future research.

At first, the observation of a positive correlation between spiking and node in-degree would seem very obvious, in the sense that nodes with high in-degree are driven beyond threshold frequently by the action of their in-neighbours. However a large in-degree cannot explain completely the high firing rate of this type of nodes, because this would also predict that nodes in fully-connected networks would possess a large spike count, which is not the case (see Fig. 4.15).

What is happening in fully-connected networks that prevents nodes from firing as much as the other heterogeneous structures considered even when these nodes are massively connected? We suggest that an obstructing behaviour is occurring in this globally-coupled structure. We name this phenomenon *spike jamming*, and we will

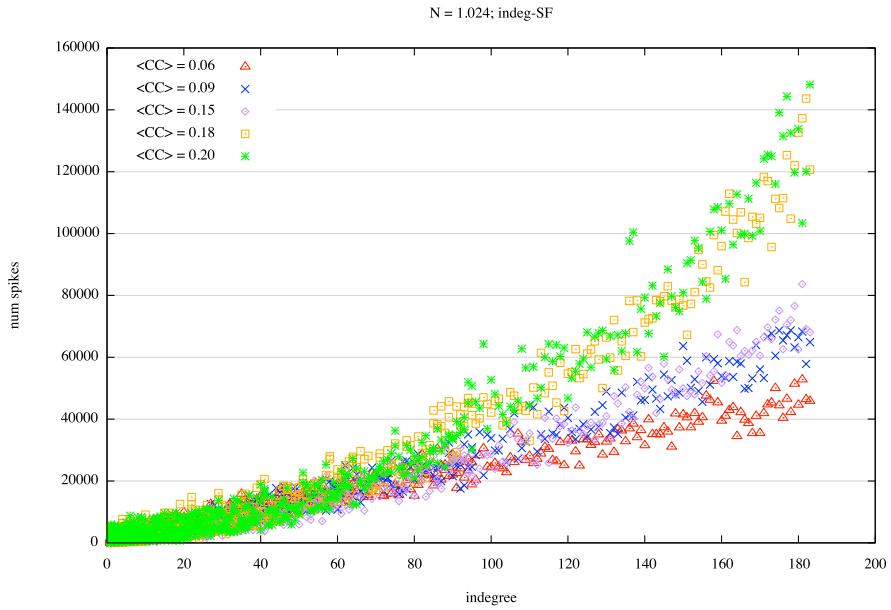
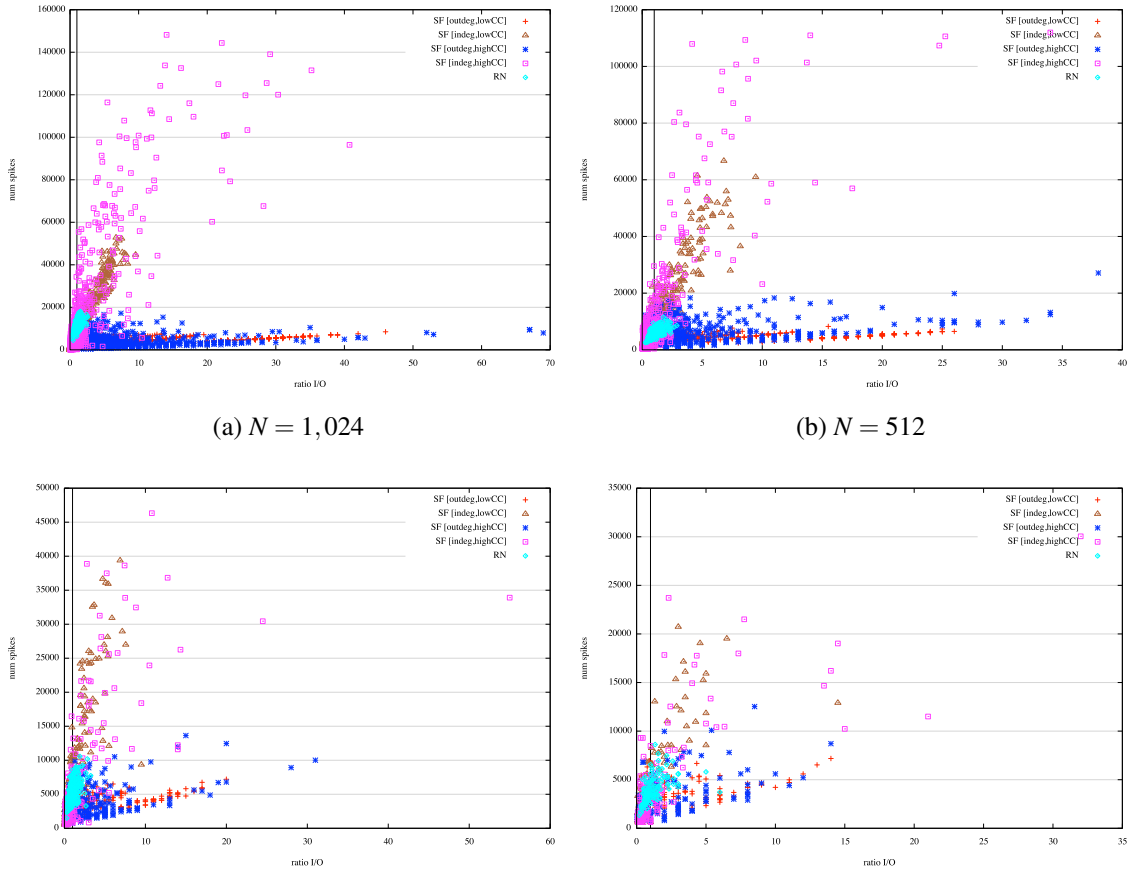


Figure 4.18: Scatter plot of in-degree vs. spike count in in-degree scale-free networks of $1,024$ nodes with varying mean clustering coefficient. A quadratic correlation between in-degree and number of spikes emerges as the mean clustering coefficient grows. This implies that the more amount of small-world-ness an in-degree scale-free network has, the more its nodes will spike.

discuss it in detail in Sect. 4.3.7.

Coming back to heterogeneous structures, random networks behave somehow similarly to scale-free networks in the sense that there is a positive correlation between node in-degree and spiking, but not between out-degree and spiking. However, this type of structure does not reach the same amount of spiking per node as scale-free networks due to the random nature of its connectivity. Thus, the presence of hubs account for the high firing rate in heterogeneous structures. For instance, in Fig. 4.17 we show how in-degree correlates with spike count. The maximum number of spikes emitted by a single hub in Fig. 4.17b is around $120,000$ spikes. This far exceeds the *circa* $8,000$ spikes emitted on average by nodes in fully-connected structures of the same size (Fig. 4.15). In scale-free networks, a single node could potentially fire much more than all the nodes in a fully-connected structure. Actually, we observed that hubs (as defined in Sect. 2.1.6) account for something between 30% and 50% of all spikes occurring in an in-degree scale-free network with high mean clustering coefficient.

Does a large in-degree in scale-free networks account by itself for high spiking activity? Or is it the joint action of in- and out-degree that explain this particular



(a) $N = 1,024$ (b) $N = 512$

(c) $N = 256$ (d) $N = 128$

Figure 4.19: Scatter plot of total number of spikes per ratio ρ of in-degree to out-degree for scale-free and random nets of all sizes considered. The solid line at the left lies at $\rho = 1$ where in-degree equals out-degree. An equal number of incoming and outgoing connections cannot account for higher spike rate. A node must have a larger in-degree than out-degree in order to spike more in these type of structures.

behaviour? In other words, could this high firing rate be explained by a specific configuration of in- and out-degree? To explore this question we considered the *ratio of in-degree to out-degree* per node, which for a node i is given by:

$$\rho_i = \frac{\sum_{k=1}^{N} A_{ki}}{\sum_{j=1}^{N} A_{ij}} \tag{4.3}$$

where the numerator is the in-degree of node i, and the denominator is its out-degree. The quantity ρ is equal to unity when a node has the same number of incoming and outgoing connections. In fully-connected networks all nodes have $\rho = 1$. If $\rho_i > 1$, then the in-degree of node i is larger than its out-degree, and $\rho_i < 1$ when the opposite

occurs.

From fully-connected networks we know that homogeneity in node degree, that is, $\rho = 1$ for every node, is not a property suitable for large spiking. Moreover, we verify this fact in heterogeneous structures where nodes with $\rho \gg 1$ fire more than any other nodes. Figure 4.19 shows this behaviour for scale-free and random networks. There we present how spiking is improved as ρ grows larger than unity. The solid black line at left of Fig. 4.19 marks the point where $\rho = 1$, that is, where in-degree equals out-degree. For all heterogeneous structures considered a larger in-degree than out-degree correlates with higher firing rate. For instance, for the networks considered in Fig. 4.19a, the PCC between ρ and spike count for in-degree scale-free networks with low mean clustering coefficient is approximately 0.97, whereas for its out-degree counterpart the PCC is approx. 0.13.

In summary, in fully-connected networks nodes fire less than in any other topology. This is explained by the homogeneity of the nodes comprising the network, which gives rise to the phenomenon of spike jamming. In heterogeneous topologies, scale-free networks fire more than random networks, and the firing activity is improved by the presence of absorbing hubs and high mean clustering coeffcient, which implies a larger degree of small-world-ness. However, for this to happen absorbing hubs must possess the right amount of outgoing connections. An equal number of outgoing connections (i.e. $\rho = 1$) does not correlate with high spike rate. The only situation where spiking is improved is when nodes possess more incoming connections than outgoing ones (i.e. $\rho \gg 1$).

4.3.6 Scale-free topologies comprise more successful nodes

Spiking is not all that matters, since we should also consider the fate of a spike that has just been emitted. Here, we consider a *successful spike* one that triggers subsequent spikes from the nodes in the out-neighbourhood of the node where the initial spike originated. As we are interested in the propagation of activity within the system, we would like to observe the sustained activation of nodes in subsequent time steps.

In Sect. 2.2.6 we defined the success of a node i at time t as the fraction of out-neighbours of i that become active at time $t+1$ when node i spiked at time t. Here, we consider the *mean node success of the system* as the mean of the mean node success

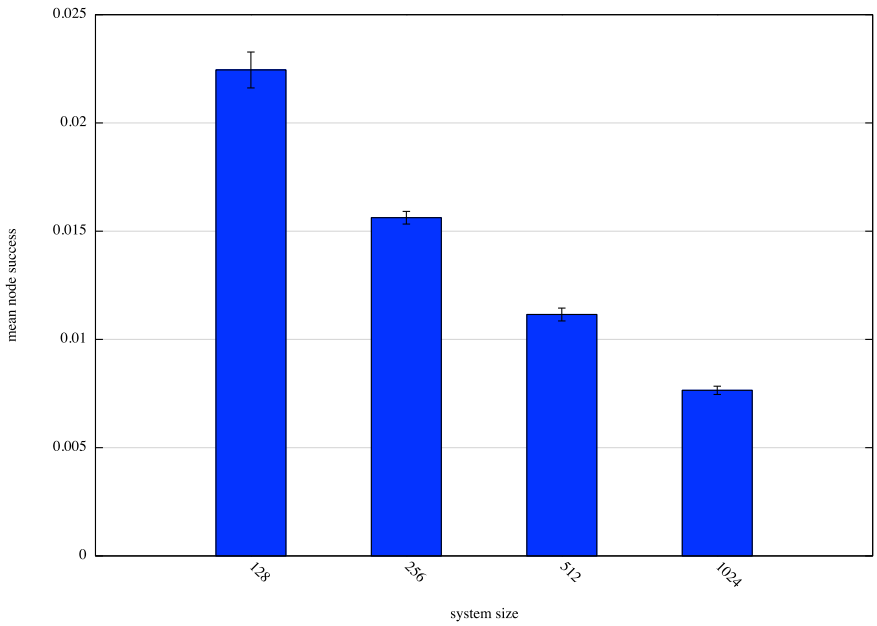


Figure 4.20: Average node success for fully-connected networks as computed using Eq. (4.4) for all system sizes considered. Larger networks exhibit a larger average spike count (Fig.4.15). However, they achieve lower mean node success due to the *spike jamming* effect that we describe in Sect. 4.3.7. (Error bars denote standard deviations.)

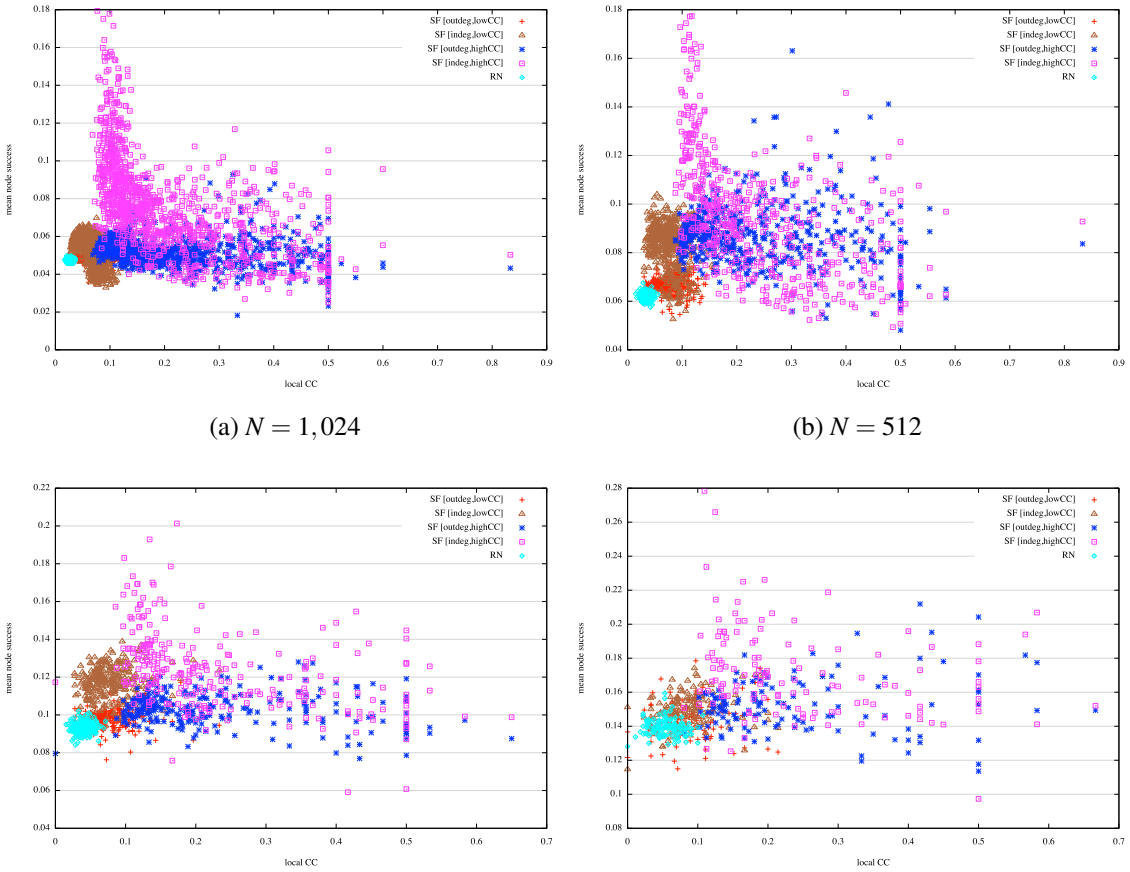
$\langle\varphi\rangle_i$ (as defined by Eq. (2.9) in Sect. 2.2.6) for all nodes in the system. In other words:

$$\langle\varphi\rangle = \frac{1}{N}\sum_{i=1}^{N}\langle\varphi\rangle_i \tag{4.4}$$

We observe that for all system sizes fully-connected networks comprise nodes with very low mean node success (see Fig. 4.20).

Next, we turn our attention to heterogeneous structures. Here, we repeat the same questions that we considered in the previous section, namely, how successful are nodes with high local clustering coefficient? How successful are hubs? And finally, how network structure affect node success?

In Fig. 4.21 we present a scatter plot of the mean success per node $\langle\varphi\rangle_i$ and local clustering coefficient for scale-free and random networks. In scale-free networks, nodes with low local clustering coefficient are the most successful nodes. This behaviour is more evident in in-degree scale-free networks with high mean clustering coefficient for larger system sizes (purple dots in Fig. 4.21). These low local clustered nodes are the hubs. However, unlike the firing activity, $\langle\varphi\rangle_i$ does not exhibit a very clear correlation between in- or out-degree and node success. For instance, the PCC between in-degree and $\langle\varphi\rangle_i$ for the in-degree scale-free network presented in Fig. 4.22a



(a) $N = 1,024$ (b) $N = 512$

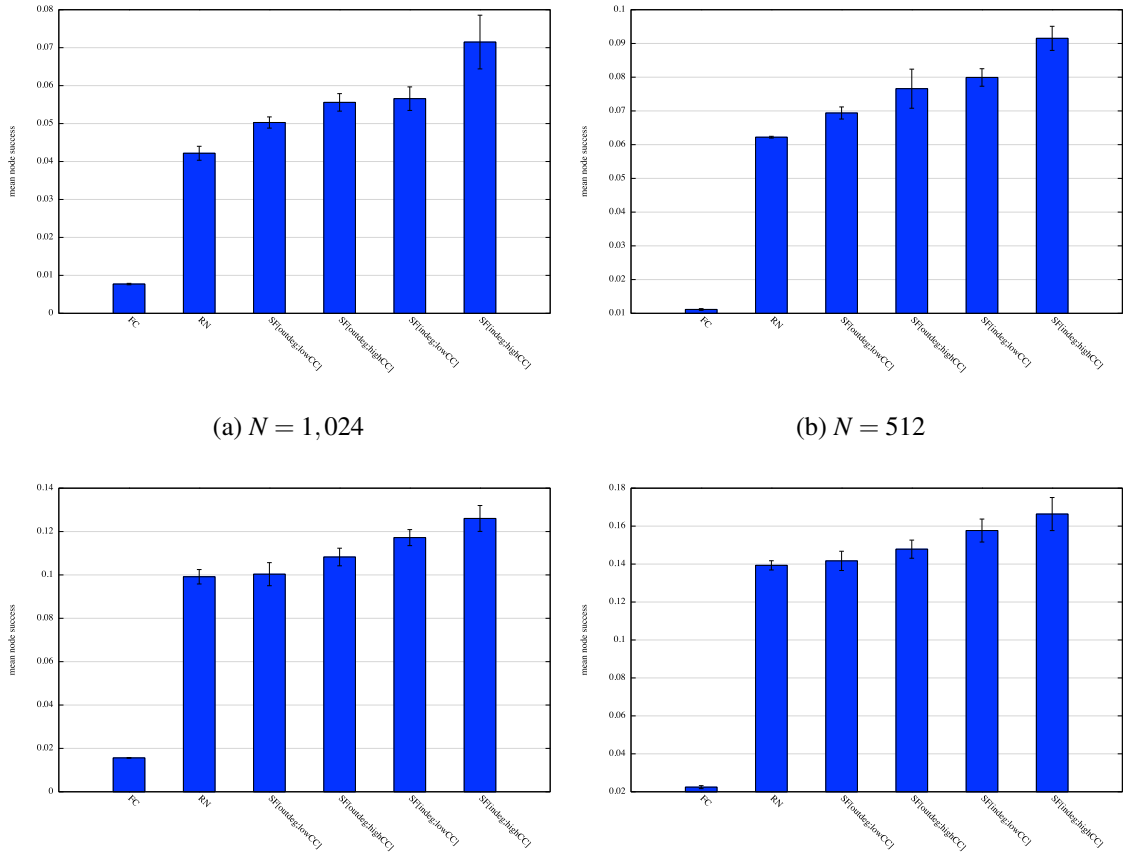
(c) $N = 256$ (d) $N = 128$

Figure 4.21: Scatter plot of local clustering coeff. (CC) vs. mean node success $\langle \varphi \rangle_i$ for random and scale-free nets of all sizes considered. Low locally clustered nodes possess a larger mean node success than better clustered nodes. This behaviour is more evident in in-degree scale-free nets with high mean CC with $1,024$ nodes (purple dots).

is around 0.75, which is the largest value from all networks considered in such figure. Lastly, random networks do not exhibit any particular pattern regarding the success of their nodes (with a PCC of approx. 0.03 for the network in Fig. 4.22a).

Which is the most successful topology? In other words, what is the structure that maximises the success per node? To answer this question we present in Fig. 4.22 the mean node success of the system $\langle \varphi \rangle$ as computed using Eq. (4.4) for all the topologies considered.

In all system sizes considered, in-degree scale-free networks with high mean clustering coefficient are the most successful topologies. Second in place are in-degree scale-free networks with low mean clustering coefficient, followed by out-degree scale-



(a) $N = 1,024$ (b) $N = 512$

(c) $N = 256$ (d) $N = 128$

Figure 4.22: Mean node success per network structure and all sizes considered: fully-connected (FC), random (RN), out-degree scale-free networks with low mean clustering coefficient (SF[outdeg;lowCC]), high mean clustering coefficient (SF[outdeg;highCC]), and in-degree scale-free networks with similar features (SF[indeg;lowCC] and SF[indeg;highCC]). In-degree scale-free network is the type of structure that maximises the success of spikes emitted by nodes, whereas fully-connected structures perform the worst.

free networks with high and low mean clustering coefficient, respectively. Thus, absorbing hubs in a scale-free structure allow for more node success per node if accompanied by a high degree of small-world-ness. Last in the group are random networks followed by fully-connected structures, which are the worst performing networks in our simulations.

As mentioned in Sect. 3.5, scale-free and random topologies have the same number of edges. Thus, in-degree scale-free-ness with high mean clustering coefficient is the permutation of edges that maximises node success under the dynamics of the Eurich

model.

Lastly, for a system to remain in the critical regime the mean node success must remain below a certain value. In other words, a large mean node success $\langle \varphi \rangle$ is related to the supercritical regime in which nodes fire constantly. Therefore, if we are interested in maximising the value of the mean node success we would have compromise the critical regime. In Fig. 4.23 we show an example of this behaviour for systems of 128 nodes.

4.3.7 Upper bound of mean node success for fully-connected nets

As mentioned in the previous section, fully-connected networks perform worst. In this section we present an analytical expression for the upper bound of this metric for globally-coupled structures.

Recall from Eq. (4.1) that $h_i(t)$ denotes the membrane potential of node i at time step t, and that $\theta > 0$ denotes the threshold in the membrane potential required to trigger a spike in a node. In other words, node i will spike at t if $h_i(t) \geq \theta$. At the start of the simulation all membrane potentials are set below threshold. These potentials are then *driven externally* at each time step until the membrane potential of one node is goes above threshold triggering an avalanche (see Fig. 2.5 in page 34). Thus, during an avalanche the membrane potentials of all nodes i in the network evolve as follows:

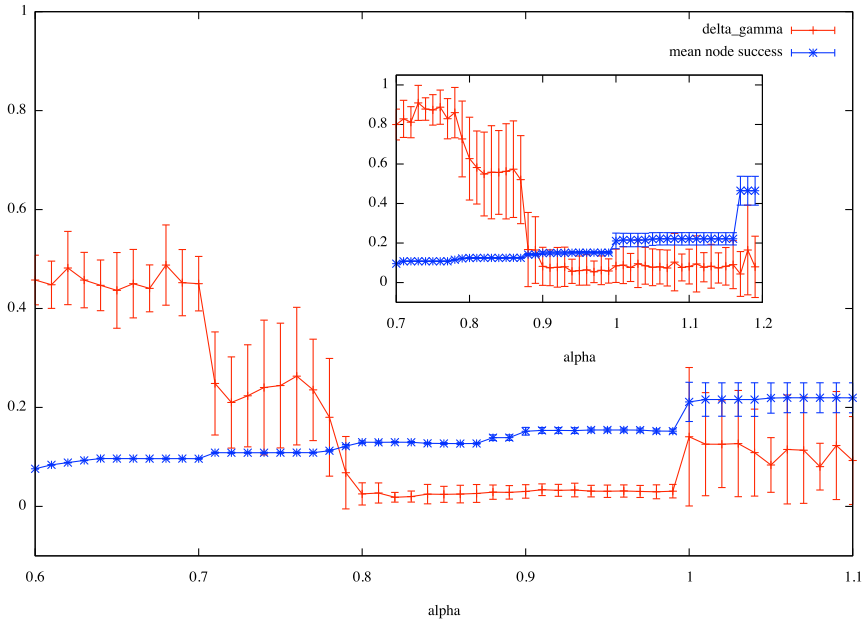
$$h_i(t+1) = \begin{cases} 0 & \text{if } h_i(t) \geq \theta \\ h_i(t) + \sum_{j \in I_i} w_{ij} s_j(t) & \text{otherwise,} \end{cases} \tag{4.5}$$

where I_i denotes the set of in-neighbours of i, w_{ij} describes the synaptic weight between nodes i and j, and $s_j(t) = 1$ denotes the state of node j at time t, and it is given by the following expression:

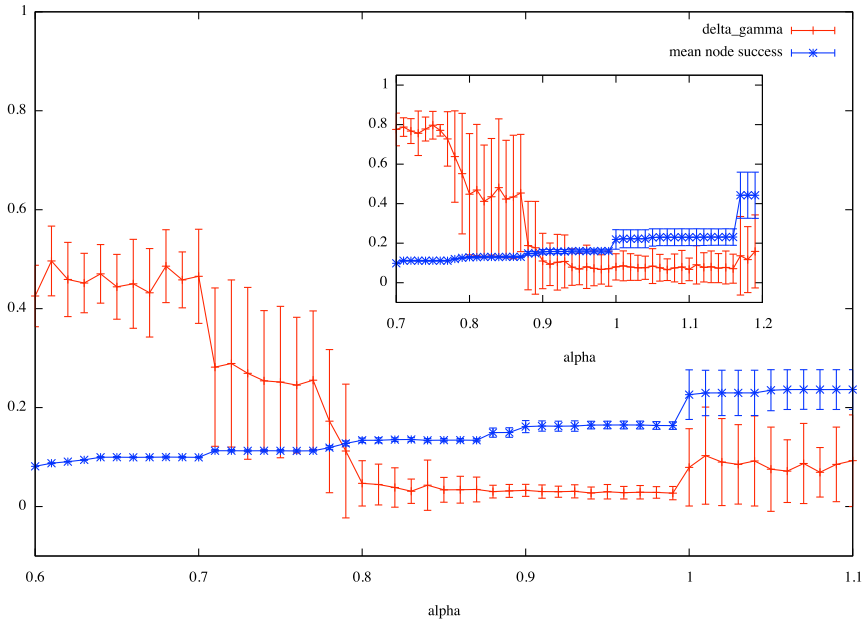
$$s_j(t) = \begin{cases} 1 & \text{if } h_j(t) \geq \theta \\ 0 & \text{otherwise.} \end{cases} \tag{4.6}$$

During an avalanche the external driving stops. An avalanche ends when there are no more active nodes, and then external driving resumes until a new avalanche is triggered. This process continues until the simulation is over.

In the following presentation we will consider only fully-connected networks. In this type of structure, each node has $(N-1)$ out-neighbours where N denotes the number of nodes in the network. In other words, for node i, all its out-neighbours are the



(a) Out-degree and low mean CC (inset: in-degree and low mean CC.)



(b) Out-degree and high mean CC (inset: in-degree and high mean CC.)

Figure 4.23: $\Delta\gamma$ curve and mean node success $\langle\varphi\rangle$ per value of control parameter α for scale-free networks of size $N = 128$. Outsets show out-degree scale free networks with two levels of the mean clustering coefficient (CC), whereas insets show their in-degree equivalent. Maximising the mean node success of the system takes the dynamics out of the critical regime towards the supercritical regime. The plateau in the $\Delta\gamma$ curve, representing the set of values of the control parameter where the system reaches the critical state, is accompanied with a mean node success that is not at its maximum. The $\langle\varphi\rangle$ function increases with α which implies that larger values of $\langle\varphi\rangle$ are related to supercriticality.

nodes j such that $j \neq i$. The node success $\varphi_i(t)$ of node i at time step t is defined as the fraction of out-neighbours of i which spike at time step $t+1$, given that i spikes at t (Eq. (2.8) in Sect. 2.2.6). Therefore, if i spikes at t, Eq. (2.8) becomes[2]:

$$\varphi_i(t) = \frac{S(t+1)}{N-1} \tag{4.7}$$

where $S(t)$ denotes the number of nodes that spike at time step t. Note that the right-hand side of the above expression is independent of i. This reflects the fact that $\varphi_i(t)$ is identical for all nodes i which spike at t in a fully-connected network. For this reason we will henceforth omit the subscript i, and deal only with the quantity $\varphi(t)$, the node success of *any* node which spikes at t.

Consider the mean node success for nodes which spike during any period of τ time steps. Since $S(t)$ nodes spike at t, this is given by

$$\langle\varphi\rangle = \frac{1}{\mathcal{S}} \sum_{t=1}^{\tau} S(t)\varphi(t), \tag{4.8}$$

where without loss of generality we have chosen the τ time steps to be $t = 1, 2, \ldots, \tau$, and

$$\mathcal{S} \equiv \sum_{t=1}^{\tau} S(t) \tag{4.9}$$

is the total number of spikes which occur during this period. A crucial aspect of the dynamics described above is that $h_i(t+1) = 0$ if node i spikes at time t. It is therefore impossible for node i to spike on two adjacent time steps. This constraint can be expressed mathematically as

$$s_i(t) + s_i(t+1) \leq 1 \quad \text{for all } i, t, \tag{4.10}$$

and leads to the following theorem.

Theorem 1. $\langle\varphi\rangle$ *has an upper bound of*

$$\langle\varphi\rangle_{max} = \frac{N}{2(N-1)} \tag{4.11}$$

in the limit $\tau \to \infty$. *This upper bound is reached when* $S(t) = N/2$ *for all* $t > 1$.

Proof: Substituting Eq. (4.7) into Eq. (4.8) gives

$$\langle\varphi\rangle = \frac{1}{\mathcal{S}(N-1)} \sum_{t=1}^{\tau} S(t)S(t+1). \tag{4.12}$$

[2] $\varphi_i(t)$ is undefined if node i does not spike at t.

Defining the quantity

$$\tilde{S}(t) \equiv S(t)/\sqrt{\mathcal{S}}, \tag{4.13}$$

Eq. (4.12) can be expressed as

$$\langle \varphi \rangle = \frac{1}{(N-1)} \sum_{t=1}^{\tau} \tilde{S}(t)\tilde{S}(t+1). \tag{4.14}$$

Now, taking the summation of Eq. (4.10) over all nodes i yields [3]

$$S(t) + S(t+1) \leq N \quad \text{for all } t \tag{4.15}$$

after noting that

$$S(t) = \sum_i s_i(t). \tag{4.16}$$

Subtracting $S(t+1)$ from both sides of Eq. (4.15) and then multiplying throughout by $S(t+1)$ yields

$$S(t)S(t+1) \leq S(t+1)\big[N - S(t+1)\big] \quad \text{for all } t, \tag{4.17}$$

which can be expressed as

$$\tilde{S}(t)\tilde{S}(t+1) \leq \tilde{S}(t+1)\left[\frac{N}{\sqrt{\mathcal{S}}} - \tilde{S}(t+1)\right] \quad \text{for all } t \tag{4.18}$$

after dividing both sides by $\mathcal{S}$ [4]. The right-hand side of the above inequality, which we denote as

$$\xi(t+1) \equiv \tilde{S}(t+1)\left[\frac{N}{\sqrt{\mathcal{S}}} - \tilde{S}(t+1)\right], \tag{4.19}$$

is hence an upper bound for term t on the right-hand side of Eq. (4.14). Therefore we can write

$$\langle \varphi \rangle \leq \frac{1}{(N-1)} \sum_{t=1}^{\tau} \xi(t+1). \tag{4.20}$$

Now, $\xi(t+1)$ is maximised when $\tilde{S}(t+1) = N/\big(2\sqrt{\mathcal{S}}\big)$. Therefore $\tilde{S}(t+1) = N/\big(2\sqrt{\mathcal{S}}\big)$ for $t = 1, 2, \ldots, \tau$ provides an upper bound for $\langle \varphi \rangle$, which, from substituting the aforementioned $\tilde{S}(t+1)$ into into Eqs. (4.19) and (4.20), can be shown to be

$$\langle \varphi \rangle_{\max} = \frac{1}{(N-1)} \left(\frac{N}{2}\right)^2 \frac{\tau}{\mathcal{S}}. \tag{4.21}$$

[3] In obtaining Eq. (4.15) we use the fact that if $x \leq a$ and $y \leq b$ then $x + y \leq a + b$.

[4] In obtaining these equations we have used the following rules for manipulating inequalities: if $x \leq a$ then $x + c \leq a + c$ for all c; if $x \leq a$ then $xc \leq ac$ if $c \geq 0$; if $x \leq a$ then $x/c \leq a/c$ if $c > 0$. In applying the second of these rules we used $c = s(t+1)$, which is ≥ 0. In applying the last of these rules we used $c = \sqrt{\mathcal{S}}$, which is > 0.

However, from Eq. (4.13), $\tilde{S}(t+1) = N/\left(2\sqrt{S}\right)$ for $t = 1,2,\ldots,\tau$ corresponds to $S(t+1) = N/2$ for $t = 1,2,\ldots,\tau$, and hence from Eq. (4.9) also corresponds to

$$\mathcal{S} = S(1) + \sum_{t=2}^{\tau} S(t) = S(1) + (\tau - 1)N/2. \tag{4.22}$$

Substituting this into Eq. (4.21) gives

$$\langle \varphi \rangle_{\max} = \frac{1}{(N-1)} \left(\frac{N}{2}\right)^2 \frac{\tau}{S(1) + (\tau - 1)N/2}. \tag{4.23}$$

We emphasise that this is realised when $S(t+1) = N/2$ for $t = 1,2,\ldots,\tau$, or equivalently, when $S(t) = N/2$ for $t = 2,3,\ldots,\tau+1$. Noting that $S(1)$ cannot exceed N, Theorem 1 results when the limit $\tau \to \infty$ is taken. ■

We should point out some remarks regarding Theorem 1. First, $\langle \varphi \rangle$ as $\tau \to \infty$, which we henceforth refer to simply as $\langle \varphi \rangle$, describes the mean node success over all nodes over all time.

Moreover, Theorem 1 applies in a very general way to fully-connected networks, in the sense that for the purpose of proving the theorem we have made no assumptions regarding how the system is driven between avalanches, or the initial values of the membrane potentials. Furthermore, we have made no assumptions regarding the specific values of θ or the synaptic weights w_{ij}.

Also, since Theorem 1 concerns only to fully-connected networks, it follows that any network with $\langle \varphi \rangle > \langle \varphi \rangle_{\max}$ *cannot* be a fully connected network. In a similar fashion, if one wishes to construct a network with $\langle \varphi \rangle > \langle \varphi \rangle_{\max}$ starting from a fully connected network, it is necessary that some connections between nodes are removed, that is, such a network should part from a massively connected to a less connected structure. In Sect. 5.2.5 we show how plasticity mechanisms can maximise the node success reachable by single nodes in a fully-connected network through activity-dependent edge pruning.

Furthermore, $\langle \varphi \rangle_{\max}$ decreases monotonically with N, and $\langle \varphi \rangle_{\max} \to 1/2$ in the thermodynamic limit, i.e., $N \to \infty$.

Lastly, and most importantly, the theorem is non-trivial in the sense that one can easily conceive networks of size N whose global node successes can potentially exceed $\langle \varphi \rangle_{\max}$. For instance, consider the network corresponding to a "directed ring", where $A_{12} = 1, A_{23} = 1, \ldots, A_{(N-1)N} = 1, A_{N1} = 1$, and $A_{ij} = 0$ for all other elements of the adjacency matrix. If $w_{ij} = \theta$, then assuming, without loss of generality, that first spike in the network occurs on time step $t = 1$ at node 1, then the spike propagates around the

ring indefinitely, that is, at $t = 2$ node 2 spikes, at $t = 3$ node 3 spikes, at $t = N$ node N spikes, at $t = N + 1$ node 1 spikes, and so on. In this case it is easy to see that $\langle \varphi \rangle = 1$, which is greater than $\langle \varphi \rangle_{\max}$ for $N > 2$. Therefore the existence of an upper bound for fully-connected networks stems from some particular property of their topology in combination with the dynamics of the Eurich model.

What is this property of fully-connected networks that results in such an upper bound? As mentioned previously, it is the fact that membrane potentials are reset to zero after they spike which ultimately results in the low performance (in terms of the mean node success) of this type of networks. This "quiescent" state gives rise to the phenomenon of *spike jamming* that we mentioned in Sect. 4.3.5, and which we describe in detail below.

Consider a single node i firing at time step t in a fully-connected network. For this node to be maximally successful, it must trigger all $N - 1$ of its out-neighbours, i.e., all other nodes in the network, to spike at $t + 1$. Suppose this happens, in which case $\varphi_i(t) = 1$. Consider now one of the nodes $j \neq i$ which spikes at $t + 1$. For j to be maximally successful, all other nodes in the network must spike at $t + 2$. However, on account of refractoriness, this is impossible. To elaborate, at $t + 1$, all nodes except for i, and including j, are spiking. Therefore all these nodes must be frozen at $t + 2$ (they cannot spike at $t + 2$). On the other hand i, which spiked at t, while frozen at $t + 1$, is free to spike at $t + 2$. Hence, at best, only one of the $N - 1$ out-neighbours of j, namely i, can spike at $t + 2$, and therefore at best $\varphi_j(t + 1) = 1/(N - 1)$. For large N, j is clearly very unsuccessful. The same applies for all other nodes which fire at $t + 1$.

Therefore the result is that, while i is maximally successful, the remaining $N - 1$ nodes are extremely unsuccessful, and thus on average the whole network is unsuccessful during this avalanche, which we assume ends at $t + 2$. This example illustrates the effect which supports the upper bound for fully-connected networks: if a node i spikes synchronously with one of its out-neighbours, then that out-neighbour is quiescent for the next time step, and hence cannot spike on the time step after i spikes, which curtails the potential node success of i, and correspondingly the propagation of spikes throughout the network. Therefore, we refer to this effect as spike jamming. Note that the aforementioned effect occurs in all networks, not just fully-connected networks. However, fully-connected networks are special in that all nodes are out-neighbours of each other, and thus this effect has more potential to curtail the node success in fully-connected networks than in any other network.

4.3.8 Complex networks exhibit non-Hamiltonian avalanches

In the context of graph theory a *path* in a graph (or network) refers to a sequence of edges in which two consecutive edges possess a vertex in common. A *Hamiltonian path* is defined as a path in a directed (or undirected) graph or network that visits each vertex exactly once. Thus, a *Hamiltonian avalanche* is an avalanche in which no nodes are active more than once for the entire duration of the avalanche.

It has been reported previously that fully-connected networks under the Eurich model exhibit only Hamiltonian avalanches, that is, avalanches in which no node is active more than once in a single avalanche. Furthermore, it has been proven that the only situation in which the opposite occurs, that is, when a node is active more than once in a single avalanche, occurs only when the system is in the supercritical regime related with avalanches larger than the size of the system (Eurich et al., 2002; Levina, 2008).

The situation in complex networks looks quite different, and in fact non-Hamiltonian avalanches are the norm in this type of networks at criticality. This kind of systems exhibit the properties associated with the critical state that we described in previous sections, namely, power-law distributions, correlations among critical exponents, data collapse, etc.; but these systems also exhibit avalanches in which nodes can be active more than once in a single avalanche, which is incompatible with criticality in fully-connected networks.

Table 4.1 in page 106 shows the average number of total avalanches, average number of Hamiltonian avalanches, and average number of non-Hamiltonian avalanches with standard deviations for all our networks. Avalanches are more numerous in out-degree scale-free networks than in any other type of network, whereas non-Hamiltonian avalanches are more abundant in in-degree scale-free networks than in any other system. One might be inclined to think that non-Hamiltonian avalanches result from the action of hubs, but non-Hamiltonian avalanches are not limited to scale-free networks, as random network also exhibit this type of avalanches. Thus, we believe that this particular type of behaviour is a result of the sparsity and heterogeneity of the nodes in the network. However, this hypothesis remains yet to be explored.

Non-Hamiltonian avalanches imply nodes that become active more than once in a single avalanche. However, we believe that a particular set of features make nodes more prone to become active more than once in a single avalanche. What particular node feature is responsible for the re-activation of nodes during a single avalanche?

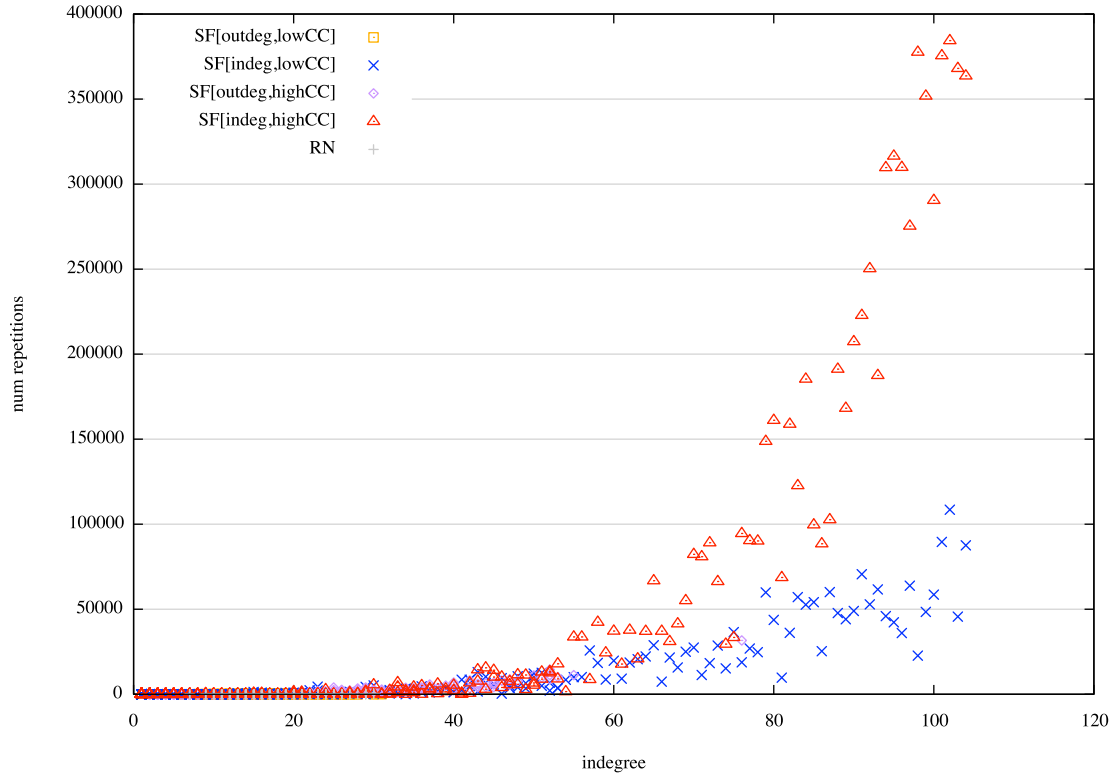


Figure 4.24: Scatter plot of in-degree vs. number of node reactivations for scale-free and random networks with $N = 512$. Nodes with largest in-degree have also the largest number of re-activations in a single avalanche. Therefore, in-degree scale-free networks possess nodes with the largest number of re-activations, and thus, the largest number of non-Hamiltonian avalanches. A quadratic trend is observed as in Fig. 4.18 for networks with high mean clustering coefficient.

Could it be their in-degree or out-degree, or their local clustering coefficient? We observe that indeed the in-degree is positively correlated with the number of re-activations of a node in a single avalanche. Thus, in-degree scale-free networks, which feature the presence of absorbing hubs, are the type of network that exhibit more re-active nodes and therefore more non-Hamiltonian avalanches.

Figure 4.24 shows an example of how in-degree is correlated with number or re-activations in scale-free and random networks of 512 nodes. The larger the in-degree a node has the more it is susceptible to re-activate in a single avalanche. Moreover, clustering has also an effect on node re-activations. A higher mean clustering coefficient results in a quadratic trend between in-degree and node re-activation. Thus, if a network possesses absorbing hubs and exhibits the small-world property, then avalanches will exhibit re-activation of nodes.

4.4 Discussion

We began this chapter with the hypothesis that network complexity exerts an effect over the dynamics of a critical system. The features of network complexity that we focused on were the small-world property, and the presence of hubs. We compared networks with these features against random and fully-connected structures, confirming that such properties exert a considerable effect particularly in terms of spiking activity and the success of spikes in terms of their capacity to trigger subsequent activations. We presented arguments to establish the importance of network heterogeneity over its internal dynamics. In particular, we described how the presence of hubs and high clustering improves the overall network performance than when nodes are all the same as in fully-connected networks, or when there is no particular pattern of connectivity as in random networks. With this in mind, we conclude that scale-free networks perform much better than any other topology when considered together with the small-world property.

Given that our scale-free and random networks possess exactly the same number of edges, we conclude that "scale-free-ness" with high degree of "small-world-ness" is a permutation of edges that allows nodes to be more successful and to be more active in terms of spike count. In particular, we observed that in-degree scale-free networks with high mean clustering coefficient is the structure that allows nodes to perform better due to the presence of absorbing hubs. However, real-world networks often comprise a more complex ecosystem in which absorbing hubs and broadcasting hubs coexist in the same network adding another layer of complexity to the dynamics within the system. Moreover, it is often the case that the structure of real-world networks is not static, but they possess mechanisms by which nodes become connected and disconnected over time as well as network growth or shrinkage; features that affect the collective activity in ways that cannot be predicted with the current model.

Nevertheless, we believe that the introduction of the analysis of the success of a spike can be applied seamlessly to these situations as well. Moreover, our analyses could potentially be applied to contexts outside neural systems. For instance, a spike can be thought of the transmission of an infection among contacts, the death of a species in models of ecosystems, the failure of a power generator in power networks, and even in online social networks such as Facebook or Twitter we might regard a spike as the action of writing a *post* or a *tweet*. In all these contexts, the fate of a spike is as relevant to the collective dynamics as is the network topology. Here we have

shown that the combination of individual dynamics of nodes and topology determine the success of the spikes that spread across the system.

Lastly, we observed how the critical regime is extended by the presence of a plateau in scale-free and random networks. In particular, the former network type exhibits larger plateaus than its random counterpart. We suggest that this might be related to the presence of *Griffith phases* in neural systems (Moretti and Muñoz, 2013). This hypothesis is interesting as it proposes that brain networks prefer sparse heterogeneous structures such as scale-free small-world networks not only for the benefits that these structures imply in terms of robustness, information processing and transmission, but also because it allows the system to extend the range in parameter space in which the critical state is reached.

Type	Subtype	Size	Total Avalanches	Non-Hamiltonian Avalanches	Hamiltonian Avalanches
Out-degree Scale-free	Low Mean CC	128	$73,412 \pm 3,040$	763 ± 303	$72,649 \pm 3,249$
		256	$14,6818 \pm 3,233$	573 ± 552	$14,6245 \pm 3,534$
		512	$243,411 \pm 2,117$	993 ± 513	$242,417 \pm 2,621$
		1,024	$310,908 \pm 5,172$	$1,580 \pm 1,053$	$309,327 \pm 5,918$
	High Mean CC	128	$76,081 \pm 2,808$	$2,166 \pm 750$	$73,915 \pm 2,722$
		256	$164,141 \pm 3,469$	$3,128 \pm 1,302$	$161,012 \pm 4,146$
		512	$230,994 \pm 3,351$	$3,399 \pm 1,157$	$227,594 \pm 4,025$
		1,024	$312,760 \pm 5,966$	$5,211 \pm 1,301$	$307,548 \pm 6,417$
In-degree Scale-free	Low Mean CC	128	$68,973 \pm 5,185$	$6,639 \pm 1093$	$62,333 \pm 5,834$
		256	$126,318 \pm 4,230$	$11,965 \pm 2,119$	$114,353 \pm 5,509$
		512	$177,127 \pm 8,393$	$17,092 \pm 1,790$	$160,034 \pm 10,082$
		1,024	$222,977 \pm 10,460$	$18,498 \pm 2,243$	$204,478 \pm 12,671$
	High Mean CC	128	$68,334 \pm 3,570$	$7,871 \pm 1,177$	$60,463 \pm 3,364$
		256	$136,228 \pm 6,516$	$14,345 \pm 2,632$	$121,882 \pm 7,961$
		512	$186,404 \pm 7,869$	$23,444 \pm 2,282$	$162,960 \pm 9,683$
		1,024	$243,972 \pm 14,006$	$26,098 \pm 3,178$	$217,874 \pm 17,061$
Random		128	$66,914 \pm 3,469$	790 ± 357	$66,124 \pm 3,546$
		256	$147,929 \pm 4,078$	$1,521 \pm 747$	$146,407 \pm 4,804$
		512	$201,719 \pm 1,185$	549 ± 76	$201,170 \pm 1,219$
		1,024	$254,533 \pm 19,034$	414 ± 134	$253,535 \pm 20,868$
Fully-connected		128	$74,465 \pm 563$		
		256	$141,138 \pm 575$		
		512	$202,050 \pm 966$		
		1,024	$214,844 \pm 18,236$		

Table 4.1: Total number of avalanches in all networks considered.

Chapter 5

Dynamic Synapses: Spike-Timing-Dependent Plasticity

Neuronal plasticity refers to the changes that brain networks undergo as a result of adaptation, learning and maintenance of brain function. In particular, spike-timing-dependent plasticity (STDP) is a mechanism for synaptic modulation that has been suggested as the underlying mechanism responsible for learning and memory in the brain (Bi and Poo, 2001). In Sect. 2.2.4 we introduced this concept as a temporally asymmetric form of Hebbian learning induced by temporal correlations between pre- and post-synaptic neurons, which accounts for the emergence of long-term potentiation (LTP), and long-term depression (LTD) in brain networks.

After exploring the Eurich model under the perspective of complex networks, we asked ourselves what would result if we added STDP mechanisms to the dynamics of the system at criticality. This chapter is our first approach in the scenario of dynamic synapses. Here, we describe the collective behaviour of systems under the Eurich model combined with activity-driven plasticity. Such plasticity mechanisms provide the network with adaptive capabilities which result from a feedback loop between the dynamics *in* the network and the dynamics *of* the network. In our study, we consider two different types of STDP mechanisms: pair-wise spike-timing-dependent plasticity (pSTDP) and triplet-wise spike-timing dependent plasticity (tSTDP).

We started by adding simple pSTDP mechanisms to our systems at criticality, and upon observing that a small-world structure emerged from fully-connected networks through activity-dependent synaptic pruning, we hypothesised that STDP carves this particular network structure in fully-connected networks at criticality. With the purpose of testing this idea we set up the protocol that we describe in this chapter.

The results presented in this chapter can be summarised as follows:

- A small-world structure emerges in fully-connected networks from pSTDP, and not from tSTDP (Sect. 5.2.1).
- Both types of plasticity (pSTDP and tSTDP) impair criticality (Sects. 5.2.2 and 5.3.1).
- Pair-wise STDP prunes direct-feedback connections establishing a one-way information flow based on the activity of nodes (Sect. 5.2.3).
- Moreover, this plasticity rule improves node success in fully-connected networks (Sect. 5.2.5).

Most of the contents presented in the following sections are part of a paper entitled *Small-world structure induced by spike-timing-dependent plasticity in networks with critical dynamics*[1]. Therefore, some of the text is taken *verbatim* from this paper.

5.1 Methods

For the purposes of this chapter we build upon the model and methods presented in Sect. 4.1, and consider the addition of the computational implementation of STDP described in Sect. 2.2.5. So, we run our experiments in the same fashion as in Sect. 4.3 for static synapses, which means that the control parameter α takes values from the critical interval. Depending on system size, the simulation will run for a number of time steps enough to allow for power-law distribution of avalanche sizes to emerge in the system. Afterwards, we introduce STDP mechanisms for another 4 million time steps, whereupon we analyse the resulting network structure and its dynamics.

We consider three system sizes: 128, 256, and 512. For each network type and size, we perform 30 trials for the sake of statistical robustness. In the case of fully-connected networks, as there exists only one fully-connected network of size N, randomness is introduced in our code for each realisation of the experiment rather than in the structure as for the other network classes considered. After each realisation of the experiment we generate 100 random networks to compute the metric S (see Sect. 2.1.4) for each of the networks that result from our simulations.

In the following section we present the results of our experiments when considering the standard and most popular version of STDP, namely, pair-wise STDP (Bi and Poo,

[1] http://arxiv.org/abs/1507.07879

1998). Afterwards, in Sect. 5.3, we present the results for the spike-triplet version of STDP (Pfister and Gerstner, 2005).

5.2 pSTDP

As mentioned in Sect. 2.2.4, pSTDP considers only two spikes, one from the pre- and another from the post-synaptic unit, to modulate the synaptic weight between the two neurons; the order of which will ultimately determine if the weight will be potentiated or depressed.

These dynamics are summarised with the following equations:

$$w_{ij}(t+1) = w_{ij}(t) + \Delta w_{ij}(\Delta t) \tag{5.1}$$

where $\Delta t = t_{post} - t_{pre}$ denotes the difference between spikes of pre- and post-synaptic neurons, and:

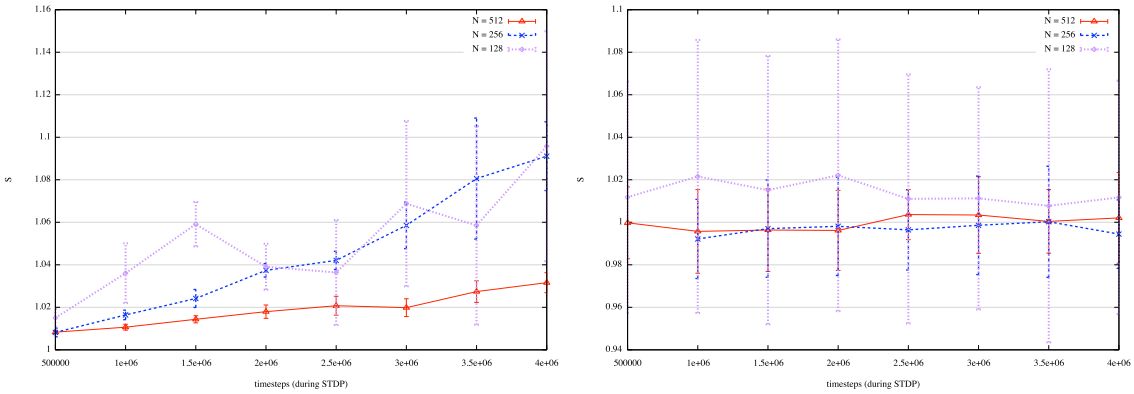
$$\Delta w_{ij}(\Delta t) = \begin{cases} a_p \exp\left\{\frac{-\Delta t}{T_p}\right\} & \text{if } \Delta t \geq 0 \\ -a_d \exp\left\{\frac{\Delta t}{T_d}\right\} & \text{if } \Delta t < 0 \end{cases}$$

where parameters a_p and T_p set the amount and duration of LTP, whereas a_d and T_d set the amount and duration of LTD. In our experiments we set $a_p = a_d = 0.1$, $T_p = 10$ time-steps and $T_d = 20$ time-steps. Also, we impose hard bounds on synaptic weights, that is, $0 < w_{min} < w_{ij} < w_{max}\ \forall i, j$, which prevents unbounded weight growth. In order to allow for activity-dependent pruning of synapses, we set $w_{ij} = 0$, if $w_{ij} \leq w_{min}$ following application of Eq. (5.1). If the connection is to be terminally deleted, we set also $A_{ij} = 0$ once $w_{ij} = 0$.

In order to assess the small-world-ness of the structures emerging from activity-dependent pruning we use the metric S introduced in Sect. 2.1.4, which is estimated in the following way. Let G be a network consisting of n nodes and e edges. To test whether G exhibits the small-world property we construct a random network R with same number of nodes and same number of edges. Then, we estimate the mean clustering coefficient of both networks C_G and C_R along with their mean characteristic path length L_G and L_R. Finally, we compute S as the ratio of those values:

$$S = \frac{C_G/C_R}{L_G/L_R} \tag{5.2}$$

If $S > 1$, then G possesses the small-world property (Humphries and Gurney, 2008). In the upcoming sections we present our results when using this type of synaptic plasticity.



(a) Fully-connected (b) Random

Figure 5.1: Evolution of small-world-ness measured by Eq. (5.2) for (a) fully-connected and (b) random networks. Small-world-ness (as measured by the metric S) increases due to pSTDP for fully-connected networks, which is not observed for random networks. This more evident for smaller system sizes. We show mean values with error bars representing standard deviations.

5.2.1 Small-world structure emerges in fully-connected networks

We observe that after simulation time a small-world structure emerges from fully-connected networks as measured by the metric S when applying pSTDP mechanisms. In Fig. 5.1 we show the evolution of S per time step for the fully-connected and random topologies considered. We present mean values and standard deviations estimated from the different realisations of our experiment.

Unlike the fully-connected case, the random structure does not show any particular trend regarding the evolution of S. Therefore, pSTDP does not imply any improvement in the structure of random networks. Moreover, as we can see in Fig. 5.1, pSTDP acts faster in smaller fully-connected networks than in larger ones, thus its effect is clearer in the networks comprising 128 nodes than in those of 512 nodes. In any case, we observe a positive trend in the evolution of S.

Our next question was to find out how essential is the critical regime for a small-world structure to emerge in fully-connected networks. Would a small-world structure emerge independently of the dynamical regime of the system? For this purpose, we considered two additional settings in our experiments. In the first one, the system would be initialised with synaptic weights taken below the critical interval. This would result in subcritical dynamics in the system. For the second setting, we initialised the system with synaptic weights taken above the critical interval. This would

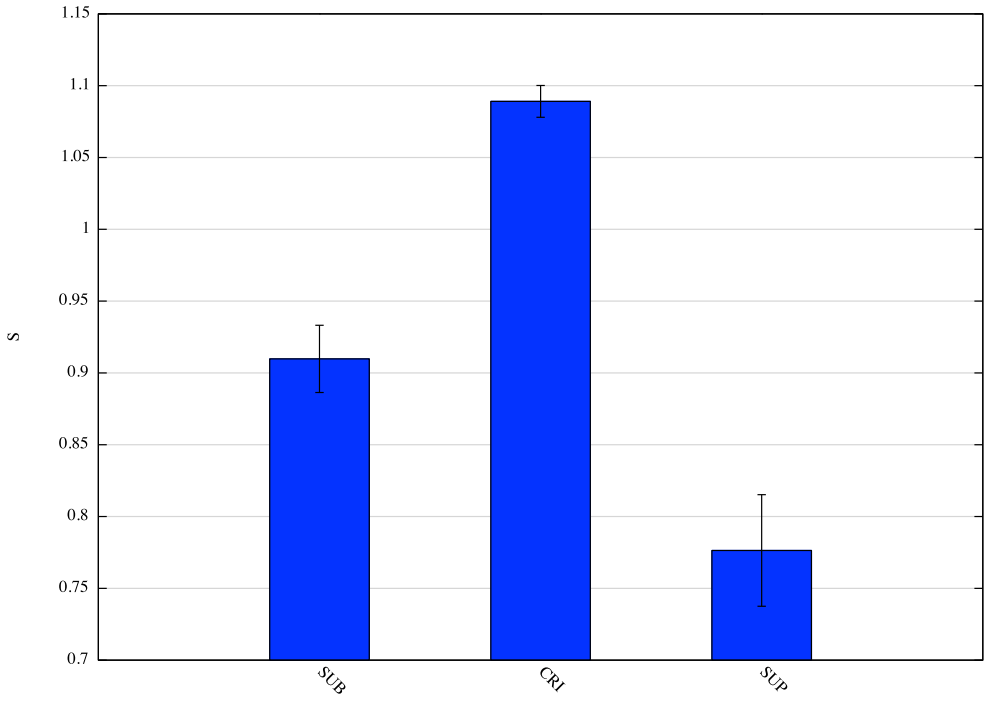
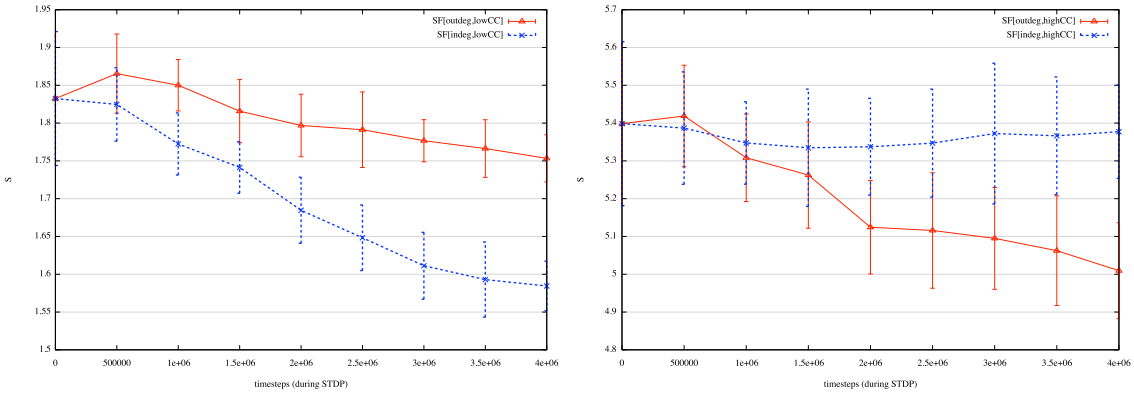


Figure 5.2: Value of small-world-ness S post-pSTDP regime in fully-connected networks of 128 nodes, whose synaptic weights yield subcritical dynamics (*SUB*, left), critical dynamics (*CRI*, middle), and supercritical dynamics (*SUP*, right). Only when the collective behaviour of the system is in the critical state, the small-world structure emerges as pointed out by $S > 1$.

result in supercritical dynamics in the collective behaviour. We observe that a small-world structure emerges *only* when the system is initialised with weights taken from the critical interval, that is, only with parameters that would allow the system to reach the critical state. In Fig. 5.2 we show the value of the metric S (Eq. (5.2)) post-pSTDP for fully-connected networks of 128 nodes, whose synaptic weights were taken from the three different regimes of the system. As can be seen, $S > 1$ only for the critical regime, which implies that criticality is essential for the small-world structure to emerge.

We believe that criticality provides the appropriate balance in the dynamics of the system to allow for a small-world structure to emerge from this type of networks. A subcritical regime would result in low activity in the system which would not allow for the appropriate amount of pruning to occur in the network, whereas a supercritical regime would result in amplified activity that would prune too many connections yielding disconnected components in the network.



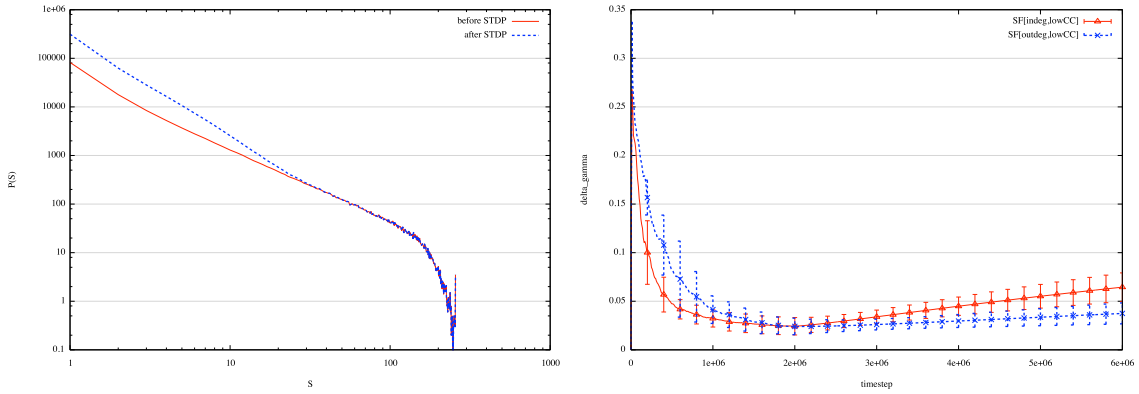
(a) Low CC (b) High CC

Figure 5.3: Effects of pSTDP on scale-free networks of size $N = 512$. pSTDP reduces the degree of small-world-ness in scale-free networks. This is observed across all system sizes and when the network has either (a) low, or (b) high mean clustering coefficient. Solid lines represent out-degree scale-free nets, whereas dashed lines represent in-degree scale-free nets.

What is the effect of pSTDP in the scale-free networks considered? As mentioned previously, these structures already possess the small-world property, that is $S > 1$ for all of them, however the degree of *small-world-ness* in them vary as we detailed in Table 3.1 in page 61. We considered *high* and *low* values of the mean clustering coefficient for this type of networks, this results respectively in high and low degrees of small-world-ness in the networks considered. The reason we considered such classes is because we are interested in assessing how pSTDP mechanisms affect scale-free structures with varying clustering coefficients. Additionally, as stated previously, we consider scale-free networks with *broadcasting hubs* and networks with *absorbing hubs*.

For these topologies, pSTDP has a negative effect regarding their small-world-ness, identified by a decrease in the value of S once pSTDP sets in. For instance, in Fig. 5.3a we show a comparison of the evolution of S for two scale-free networks with 512 nodes and low mean clustering coefficient (which implies low degree of small-world-ness). The network identified by the continuous line is an out-degree scale-free network, whereas the dashed line represents its transpose, that is, an in-degree scale-free network. As it can be seen in such figure, pSTDP affects negatively the small-world property in these networks.

Similarly, Fig. 5.3b traces the changes of S for two scale-free networks of same size as above but with high degree of small-world-ness. The continuous line represents a



(a) Log-log distribution of avalanche sizes pre- and post-pSTDP

(b) $\Delta\gamma$ curve post-pSTDP

Figure 5.4: Effects of pSTDP on collective dynamics for scale-free nets with $N = 256$. Before pSTDP, small avalanches coexist with large ones, and their distribution can be approximated accurately by a power-law (a, continuous line), which is also identified by a small error around time-step 2×10^6 (b, solid and dashed lines). However, after pSTDP, large avalanches cease to occur (a, dashed line) and the error exhibits an incremental trend for the rest of the simulation (b, both lines after time step 2×10^6).

scale-free network with power-law out-degree distribution, whereas the dashed line represents its transpose, a network with a power-law in-degree distribution. In the best of cases, S does not exhibit an noticeable decrease. These situations are verified in all other systems sizes considered for scale-free networks.

It is clear that pSTDP does not have the same effects when the power-law of the degree distribution is present in the out-degree distribution or in the in-degree distribution; nor when the clustering coefficient of network is low or high. The trend, however, is not clear; that is, we cannot tell if a particular network feature results in faster or slower decrease of the small-world property in this kind of structures. Nevertheless, we find that pSTDP affects negatively the degree of small-world-ness in scale-free networks.

5.2.2 pSTDP impairs criticality

As mentioned in Sect. 4.1.1, we assess the quality of the power-law approximation to the distribution of avalanches by estimating the deviation from the best power-law fit. When such an error function reaches a minimum value of less than 0.05, we consider the event-size distribution as well approximated by a power-law and conclude

that the system is in a critical state. After the system has approached a power-law behaviour, we allow pSTDP mechanisms to set in. We observe that criticality is generally lessened while pSTDP is modulating the synaptic strengths. When pSTDP sets in the difference between the distribution of avalanche sizes and the best-fit power-law increases again, such that eventually a power-law ceases to appropriately model the system, see Fig. 5.4b. This is observed in all networks and sizes considered.

The deterioration of the power-law approximation is explained by the fact that once pSTDP sets in large avalanches cease to occur and only small avalanches take place. In this sense, it can be said that pSTDP favours local, clustered events. Fig. 5.4a shows an example of this behaviour. Here, the continuous line shows the system during criticality before pSTDP mechanisms take place, which can also be identified in the minimum value shown in Fig. 5.4b around time step 2×10^6. Once pSTDP takes place there are no more large avalanches to be added up to the distribution, but only small ones, which results in the particular shape of the dashed line in Fig. 5.4a. As well, in Fig. 5.4b after pSTDP has set in right after the 2×10^6 time step the deviation from a power-law distribution increases and we observe a larger error as time goes by.

We conclude that the critical state vanishes as pSTDP mechanisms take place in our model. We verify this by observing the value of Λ, the largest eigenvalue associated to the matrix with entries w_{ij}. In Table 6.1 in page 158 we present the value of this eigenvalue for our systems post-pSTDP regime. As can be seen in column Λ_{pSTDP} the value of $\Lambda < 1$ pointing out the deviation from criticality in the systems.

5.2.3 pSTDP prunes direct-feedback connections

We analyse the motif profile of the networks after the pSTDP regime. As mentioned in Sect. 2.1.8 there are 13 different motifs representing the possible relations among nodes taken in threes from a directed network. The different configurations are shown in Fig. 2.1 in page 19.

A motif profile shows the distribution of these 13 different 3-node configurations for a single network. For example, before pSTDP, a fully-connected network possess a single motif distribution: all 3-node relationships are of the *XIII* type. However, after pSTDP the motif *XIII* breaks apart and instead motifs of the types *I*, *II*, *IV*, *V* and *IX* abound. Interestingly, none of these contain direct feedback connections. Fig. 5.5 shows the motif profiles of the fully-connected networks considered after pSTDP (sizes 256 and 512 in the insets). For the smallest system size it is more evident how motifs

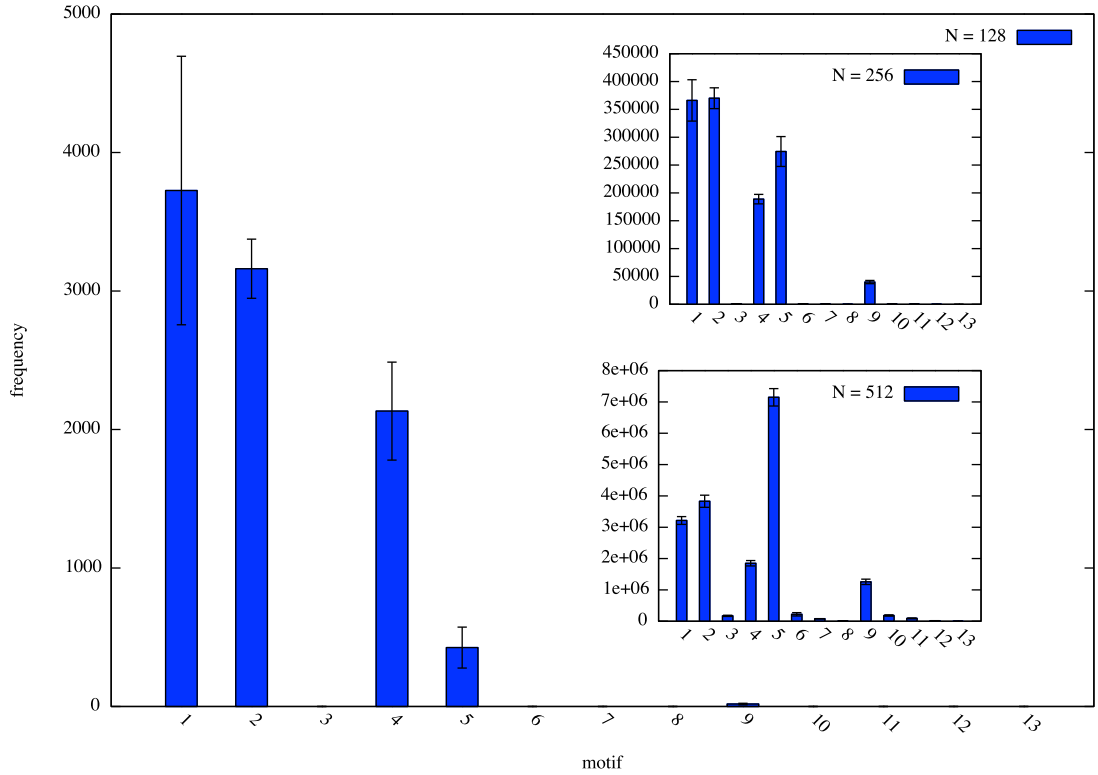
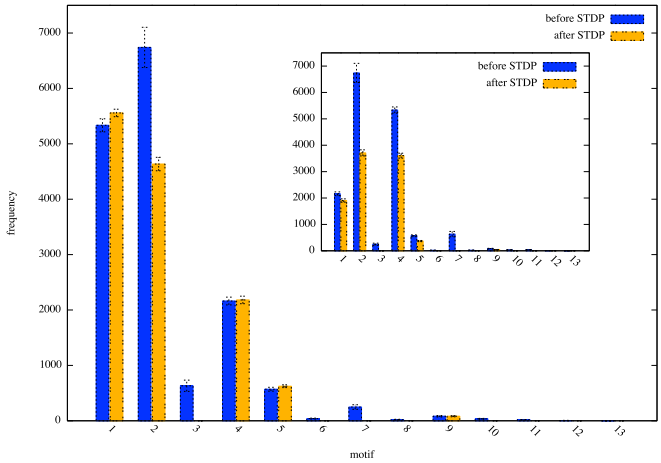


Figure 5.5: Post-pSTDP motif profile for fully connected networks of size $N = 128$. (Inset: sizes $N = 256$ and $N = 512$.) After pSTDP, motif *XIII* breaks apart and motifs *I*, *II*, *IV*, *V* and *IX* emerge. None of these new motifs implement direct feed-back connectivity among the nodes comprising the motif. This is more evident for smaller system sizes, where pSTDP requires less time for pruning.

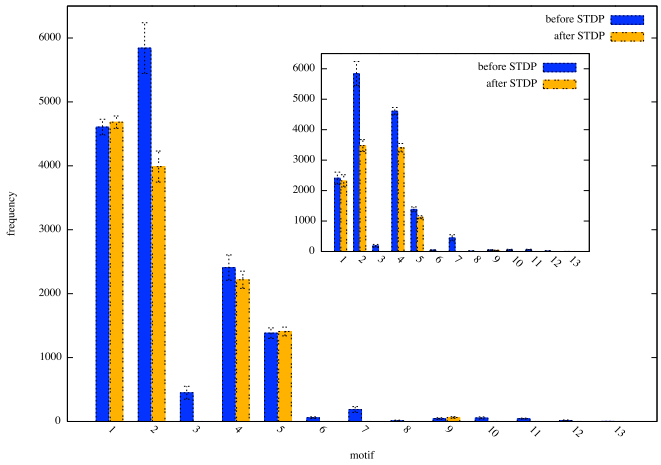
I, *II*, *IV*, *V* and *IX* grow where previously there was only motif *XIII*. As the system grows, pSTDP requires more time to prune the network and profiles differ from one another.

In the case of heterogeneous topologies, pSTDP *attacks* motifs with bidirectional connections affecting in this way the local clustering of the global structure, which in turn affects the degree of small-world-ness of the network. Motifs with bidirectional connections are: *III*, *VI*, *VII*, *VIII*, *X*, *XI*, *XII* and *XIII*; all of them impaired by the action of pSTDP. Fig. 5.6 shows this behaviour for our networks of 128 nodes. However, we observe this particular behaviour in all the different system sizes considered.

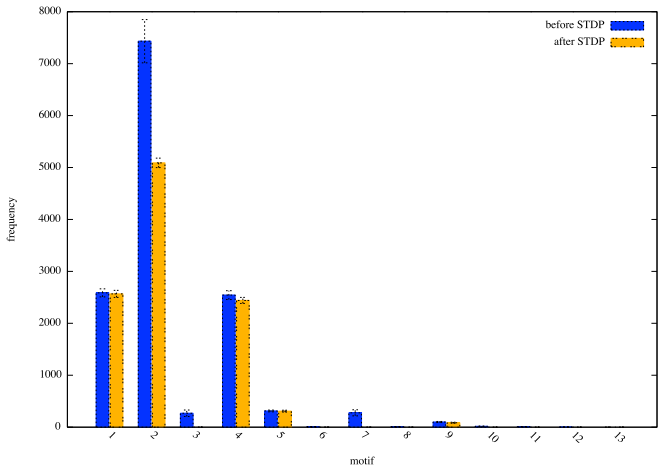
In particular, motif *II* is severely impaired in all cases considered. This motif represents the most basic feed-forward flow of information comprising 3 nodes, and as such the most elementary pre- and post-synaptic relationship among three nodes, and as such it is expected to be a target for pSTDP. Interestingly, some motifs vanish from



(a) outdegree-SF net with low CC (inset: its transpose).



(b) outdegree-SF net with high CC (inset: its transpose).



(c) Random net.

Figure 5.6: Motif profiles for networks of size $N = 128$ pre- (blue) and post-pSTDP regime (orange). pSTDP particularly affects motifs that contain direct feed-back connectivity. This is observed across all system sizes and when the network possess (a) high or (b) low mean clustering coefficient (CC), as well as when the network has an in-degree power-law distribution (insets). A similar behaviour is observed in (c) random networks.

the profile. These are motifs *III*, *VI*, *VII*, *VIII*, *X*, *XI*, *XII* and *XIII*; all of them involving direct feedback connectivity.

Therefore, in a sense it can be said that pSTDP prevents direct feedback connectivity, preferring indirect feedback flow in which a third node serves as intermediary. For example, motif *X* includes a direct feedback connection between two nodes; however, after pSTDP this motif might transform into motif *IX*, in which the feedback flow is now mediated by a third node.

pSTDP favours *one-way* connectivity rather than *two-way*, thus establishing a direction for the flow of stimuli given by the current activity in the system. Simply put, if node i sends an edge to node j, but this latter node fires shortly before the former most of the time, then pSTDP acts by severing this somewhat *useless* connection in terms of the causal relationship between the activation of nodes i and j.

5.2.4 Other effects on topology

Previously, we mentioned how pSTDP affects the degree of small-world-ness in the networks considered. The process of activity-dependent pruning has other consequences in topology, namely, the decrease in edge density and eventually, the disconnection of the network and the emergence of different connected components across the system.

The density of a network is the ratio of the number of edges to the number of possible edges (see Sect. 3.5). As such, a fully-connected network possess all possible connections in the network resulting in the maximum possible value of the density, which is 1. As mentioned in Sect. 3.5, for all system sizes, our heterogeneous structures have the same number of edges, which yields the same density for all of them. However, how pSTDP affects this ratio varies depending on the particular type of network considered.

Network density is an indicator of sparseness in a structure. As described in Sect. 3.5 our heterogeneous networks are sparse, a feature that has an effect on the way pSTDP acts upon the structure. Fig. 5.7 shows the effects of pSTDP mechanisms on network density for all the fully-connected cases considered, whereas Fig. 5.8 shows a similar case for scale-free and random networks of 128 nodes.

The effects of pSTDP are soon visible in a dense network topology. In a fully-connected network, pSTDP acts fast removing a considerable amount of edges as soon as it sets in. In contrast, for sparse networks, either scale-free or random, pSTDP acts

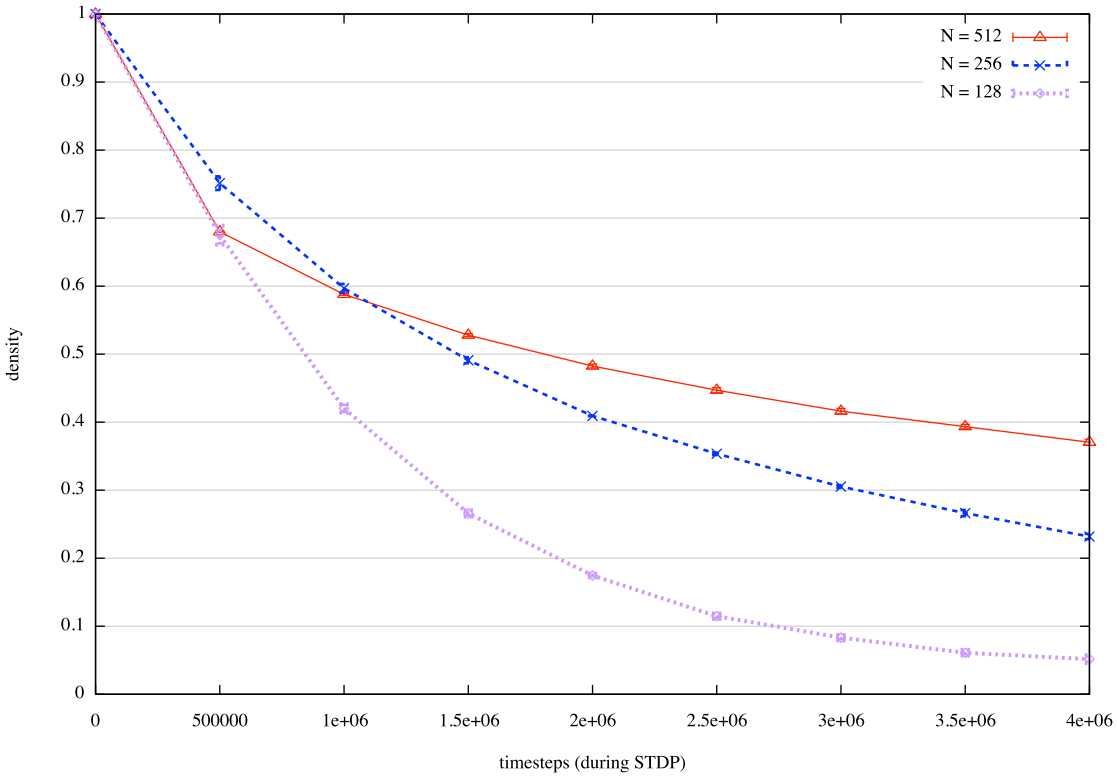
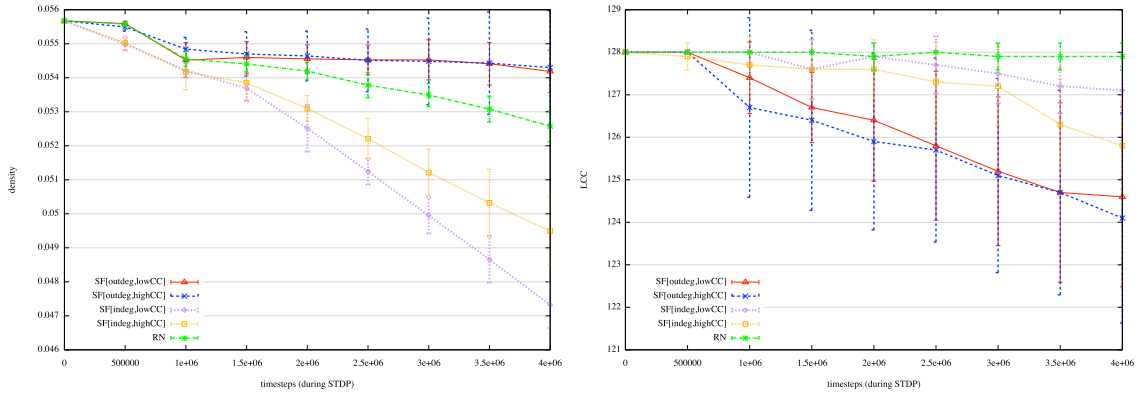


Figure 5.7: Density of fully-connected nets. pSTDP prunes edges here faster than in the other structures (see Fig. 5.8 for comparison). By the end of the simulation, fully-connected networks have lost more than 50% of their connections. In fact, networks of 128 nodes lose more than 90% by the end of the simulation. This is a big contrast from heterogeneous structures, which lose ca. 1% of their connections by the end of the experiments (Fig. 5.8).

more slowly not even removing 1% of the initial configuration under otherwise similar conditions.

However, pSTDP does not act the same in every heterogeneous topology. In-degree scale-free networks lose density faster than any other network, specially when the mean clustering coefficient is low. This can be explained by the fact that in in-degree scale-free networks absorbing hubs possess a higher firing rate, which in turn has a strong influence on pSTDP mechanisms. However, clustering seems to serve as a protective mechanism to prevent edge pruning. This can be observed in the fact that our scale-free networks with low mean clustering coefficients lose edges faster than their high clustered counterparts.

During simulation time, the effects of pSTDP pruning result on a disconnected network where components of different sizes emerge. All our observations regarding



(a) Network density (b) LCC size

Figure 5.8: pSTDP effects on (a) network density and (b) the size of the largest connected component (LCC) for heterogeneous structures of size $N = 128$. Unlike fully-connected nets, the pruning process of pSTDP is milder in heterogeneous structures. In this particular figure, pSTDP removes around 1% of edges by the end of simulation time. This can be explained by the sparseness of these networks. In-degree scale-free networks with low mean clustering coefficient (CC) lose edges faster than any other network (a, purple dashed line), whereas the LCC of out-degree scale-free networks with high CC shrinks faster than any other structure (b, blue dashed line).

topology refer to the largest connected component (LCC) of the network. Because pSTDP behaves differently across the different topologies considered, we expect the LCC to be different as well. The LCC of out-degree scale-free networks suffer a faster shrinkage than random networks and in-degree scale-free networks, even if it is this latter type of networks the one that loses density faster than any other topology. This behaviour is shown in Fig. 5.8b and it is observed across all system sizes considered.

In summary, scale-free networks with absorbing hubs lose edges faster than any other structure. However, these networks show more resilience to become disconnected (Fig. 5.8a). On the contrary, scale-free networks with broadcasting hubs show more resilience to lose edges, but they are more prone to become disconnected when losing edges (Fig. 5.8b).

5.2.5 pSTDP improves node success

In Sect. 2.2.6, we introduced the concept of node success as a local measure of node performance described by Eq. (2.8). We inquire about the effects of pSTDP on node

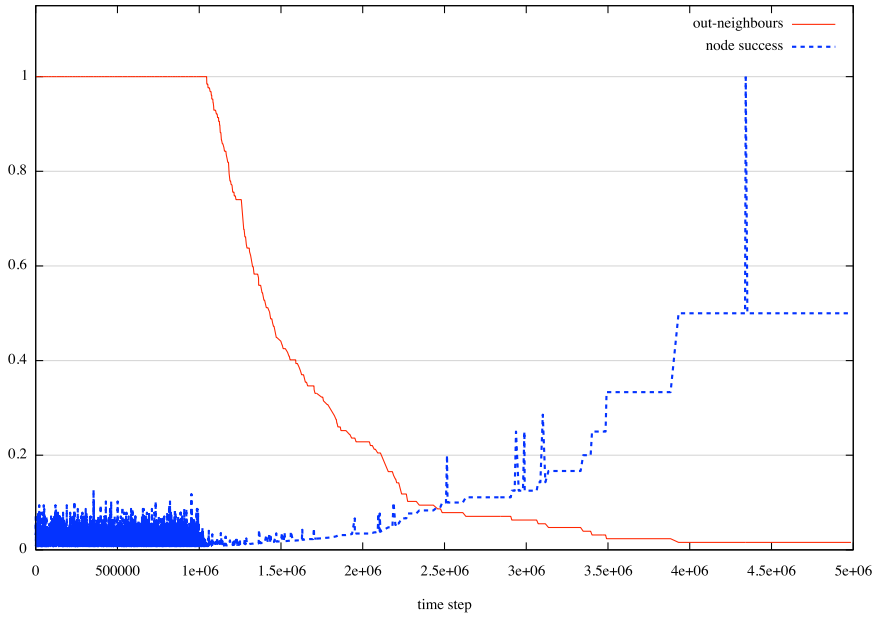


Figure 5.9: Evolution of synaptic pruning and node success for an individual node in a fully-connected net of 128 nodes. The red solid line represents the normalised numbers of out-neighbours of the node. At the start of the simulation the node has 127 neighbours but after pSTDP mechanisms take place the node starts losing connections among them as a result of activity-dependent pruning. Nevertheless, at the same time the node success of this node (represented by the blue dashed line) increases as the pruning process deprives it from "indifferent" neighbours. Around time step 4.5×10^6 this node reaches a node success of 1, which is represented by the spike at this time step. We only present values of the node success where $\varphi > 0$ for the sake of presentation.

performance per time step. Does pSTDP increase the probability that out-neighbours of a given node will become active following a spike?

Indifferent out-neighbours are those nodes in the out-neighbourhood of a given node, which do not become active following a spike from the node connecting to them. A quick inspection on the mechanisms of pSTDP suggest that this type of plasticity has all the ingredients required for a node to get rid of its indifferent out-neighbours: a *successful* spike potentiate connections with non-indifferent out-neighbours, whereas an *unsuccessful* spike causes depression between the spiking node and the indifferent neighbours, which ultimately might lead to synaptic pruning.

We observe that pSTDP through synaptic pruning increases the chances of nodes to become more successful in fully-connected networks. Fig. 5.9 shows the evolu-

tion of synaptic pruning and node success for a node in a fully-connected network of 128 nodes. The node starts with 127 neighbours represented by the red solid line in the figure (the number of out-neighbours has been normalised to simplify the presentation). Once pSTDP mechanisms take place around time step 10^6 the node starts losing connections among out-neighbours, but at the same time the value of its node success increases. This behaviour was observed in nodes of fully-connected structures, which would imply that pSTDP not only carves a small-world structure out of this type of topologies, but also maximises the node success achievable by individual nodes in these networks. Therefore, pSTDP is a mechanism by which starting from a fully-connected network we obtain a small-world network in which the node success of individual nodes is increased.

5.2.6 Discussion

We started this chapter enquiring what would be the effects of simple STDP mechanisms (pSTDP) in a system at criticality. Our initial observations were that pSTDP carves a small-world structure out of a fully-connected topology. Therefore we hypothesised that critical dynamics combined with this plasticity mechanism could explain the emergence of this particular type of topology in neural systems. We confirmed our hypothesis with the results presented in the previous sections.

One important observation regarding small-world-ness in networks is that of the case of fully-connected topologies. It can be argued that this particular type of structure possesses the small-world property *ab initio* as it has both the maximum value of mean clustering coefficient and the lowest value of mean path length possibles, therefore any alteration to its structure can only impair its degree of small-world-ness. This is a very valid observation. However, we should point out that a fully-connected network is a blank slate, in which no dynamic process has yet taken place. pSTDP carves a structure out of this topology, that changes the ratio of the mean clustering coefficient to the mean characteristic path length when compared to a random network with the same number of nodes and edges. The resulting structure is better than the one that would have emerged from a process pruning edges randomly. Moreover, the value of S for a fully-connected structure is 1 always, as the comparison with a random network with same number of nodes and edges yields exactly the same fully-connected structure. However, we observe that after pSTDP the fully-connected network becomes a network in which clearly $S > 1$ (see Fig. 3.1a), which implies the presence of the small-

world property. In other words, pSTDP shapes a functional small-world structure from a densely connected topology as it captures the patterns of avalanche activity occurring within the system.

In Sect. 2.3 we briefly described the model of Basalyga et al. (2011). Unlike ours, their model is not based on critical dynamics. They consider Erdös-Renyi networks of leaky integrate-and-fire units but only at a network size of $N = 100$ neurons, whereas we considered scale-free, random and fully-connected structures of sizes larger than 100 nodes. Since in their model only excitatory synapses are subject to pSTDP, the topological effects are biased; whereas in our model all synapses are plastic. Here, we reproduced their observation regarding the emergence of a small-world structure through STDP mechanisms. However, we achieve such result by considering critical dynamics.

Our analysis includes an inspection of the motif profile that resulted from pSTDP mechanisms. Song et al. (2005b) estimated the motif distribution of acute slices from the visual cortex of rats and observed that the motif profile of such networks differs from their random counterparts, in particular bidirectional connections were found to be more frequent than expected by chance. In contrast, we have obtained a decreased number of bidirectional motifs which is related to the particular form of the pSTDP rule used in our work due to the action of criticality within the system. As Song and colleagues claim, their counts are relative to random networks, whereas ours are absolute counts, that is, real counts which are not compared against a random network. In their work, they present a ratio of actual counts to that predicted by their null hypothesis. In this context, bidirectional connections show an overrepresentation, a situation which is not in clash with our observations. When the ratio of motifs containing bidirectional connections is compared to that of non-bidirectional connections in the same network profile, we observe that the latter is much larger than the former, an observation that is in agreement with ours. If we compare the resulting state of our model to the state before pSTDP was applied, we find that many bidirectional connections have been deleted. Moreover, it can be said that from a certain point of view, their observations are just snapshots of a dynamic process happening in the networks that they considered. Such a process would have another effect in the long run, in which bidirectional connections are found more rarely than before. Also, these authors do not mention anything related to the type of dynamic process that gave rise to the structures that they observed. Thus, the question regarding pSTDP as a mechanism that prunes two-way connections is not settled by their work.

Ex nihilo nihil fit. The brain is shaped not only by genetics but also by activity-dependent processes during development. It is known that several brain structures possess the features commonly associated to small-world networks (Sporns, 2010). In particular, the only neural network that has been mapped in its entirety, namely the nervous system of the nematode *C. elegans*, is known to possess the small-world property (Watts and Strogatz, 1998). In higher organisms, genetic processes cannot fully account for the existence of this particular network structure, as during the lifetime of an individual, modification of the network is present due to activity-dependent processes, e.g. in connection to learning and memory. Here, we claim that massively connected topologies combined with critical dynamics can give rise to a small-world structure already if a standard STDP rule such as pair-wise STDP is in place for adapting the network towards a *better-than-random* structure which is beneficial for information transmission across the system, and in which spikes of nodes are more successful at triggering subsequent spikes.

Nevertheless, more elaborate versions of pSTDP mechanisms are yet to be explored. In particular, we are interested in assessing the effects of triplets of spikes that occur between a pair of neurons, and whose action go beyond the simple pair-based rule explored in our model. In the next section we follow this direction.

5.3 tSTDP

In this section we extend the traditional model of spike-timing-dependent plasticity which considers only two spikes to estimate the modulation of synaptic efficacy between pre- and post-synaptic neurons (Pfister and Gerstner, 2005). Already in Sect. 2.2.4 we described the differences between pSTDP and tSTDP, and the motivation for their study.

In tSTDP, we consider three spikes for synaptic modulation: either two pre- and one post-synaptic spike, or one pre- and two post-synaptic spikes. In this way, tSTDP not only considers the temporal correlations between pre- and post-synaptic spikes but also the firing rate of the spiking units. The mechanism of operation can be summarised like this: potentiation is achieved when the pre-synaptic neuron fires in between two post-synaptic spikes, whereas depression occurs when the post-synaptic neuron fires in between two pre-synaptic spikes (see Fig. 2.4 in page 30).

In order to implement computationally the tSTDP learning rule, we add a couple of parameters to Eq. (5.1) presented in Sect. 5.2 which yields:

$$\Delta w_{ij}(\Delta T_1) = \begin{cases} a_p \exp\left\{\frac{-\Delta T_1}{T_p}\right\} \exp\left\{\frac{-\Delta T_2}{T_y}\right\} & \text{if } \Delta T_1 \geq 0 \\ -a_d \exp\left\{\frac{\Delta T_1}{T_d}\right\} \exp\left\{\frac{-\Delta T_2}{T_x}\right\} & \text{if } \Delta T_1 < 0 \end{cases}$$

where parameters a_p and T_p set the amount and duration of LTP, whereas a_d and T_d set the amount and duration of LTD, as with pSTDP; T_1 represents the difference between pre- and post-synaptic spikes, T_2 denotes the temporal difference between the two most immediate post-synaptic spikes (if $\Delta T_1 \geq 0$) or the temporal difference between the most immediate pre-synaptic spikes (if $\Delta T_1 < 0$); and T_x and T_y are two parameters that in a similar fashion as with parameters T_d and T_p set the amount of influence of immediate spikes for depression and potentiation, respectively.

In the following sections we present the results when applying tSTDP mechanisms to our model.

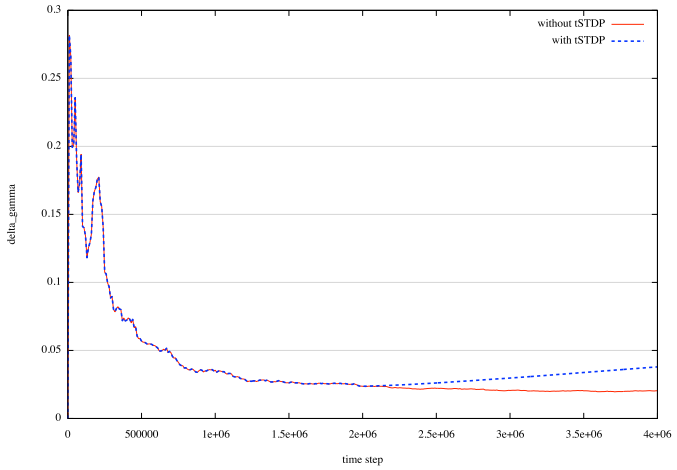
5.3.1 tSTDP impairs criticality

Similar to pSTDP the synaptic modulation triggered by tSTDP mechanisms alter the quality of the power-law approximation to the distribution of avalanche sizes. As expected, this in turn has effects over the error function and the exponent of the power-law fit. In Fig. 5.10 we present an example of this behaviour for a scale-free network of 256 nodes with low mean clustering coefficient. Here, we present the distribution of avalanche sizes, the evolution of the $\Delta\gamma$ curve representing the error in the power-law fit, along with the evolution of the corresponding exponent γ for two regimes: one in which no tSTDP was applied for the whole duration of the experiment, and another in which tSTDP mechanisms were applied after time step 2×10^6, that is, once the system reaches the critical state.

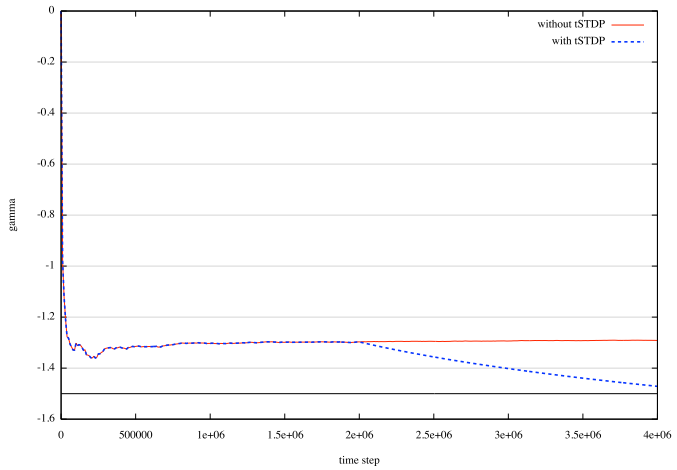
Moreover, in Table 6.1 in page 158 we present the value Λ, the largest eigenvalue associated to the weight matrices, that results from the application of tSTDP into our systems. The values found in column Λ_{tSTDP} in this table are less than unity, which is a signature of the deviation from criticality in the systems.

5.3.2 tSTDP modulates but does not prune connections

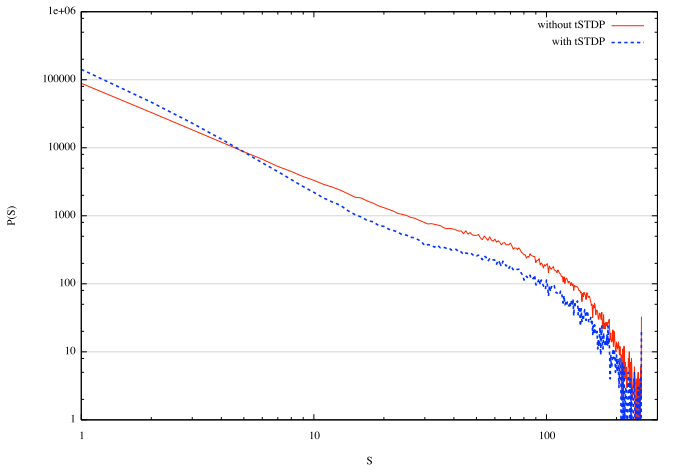
Although tSTDP has an effect on the modulation of the synaptic strength between two connected nodes, this type of plasticity is less effective than pSTDP on severing connections to the point that almost all edges survive after simulation time. Because of this, we do not observe any small-world structure emerging from fully-connected



(a) $\Delta\gamma$ curve.



(b) Exponent γ.



(c) Distribution of avalanche sizes.

Figure 5.10: Effects of tSTDP over the critical regime of scale-free network of size 256 with low mean clustering coefficient. We compare two different scenarios: no tSTDP during simulation time (red solid line), and tSTDP applied after time step 2×10^6 (blue dashed line). As with pSTDP, the synaptic modulation induced by tSTDP takes the system away from criticality. (a) Evolution of the $\Delta\gamma$ curve denoting the error in the power-law fit to our data, (b) Evolution of the exponent of the best-fit power-law. (c) Distribution of avalanche sizes in log-log plots.

networks as we did with pSTDP. At criticality, tSTDP acts as a mild modulator of synaptic weights based on the firing activity of pre- and post-synaptic nodes.

Figure 5.11 shows the number of edges of networks of 256 nodes where no STDP regime was applied (noSTDP), where pair-wise STDP was applied (pSTDP), and where triplet-wise STDP (tSTDP) was applied. Here, the systems ran for 4 million time steps. When a plasticity mechanism was applied, the system was allowed to develop without any STDP mechanism for the first half of the simulation time; afterwards, STDP was applied. We observe that pSTDP achieves more considerable edge pruning than tSTDP.

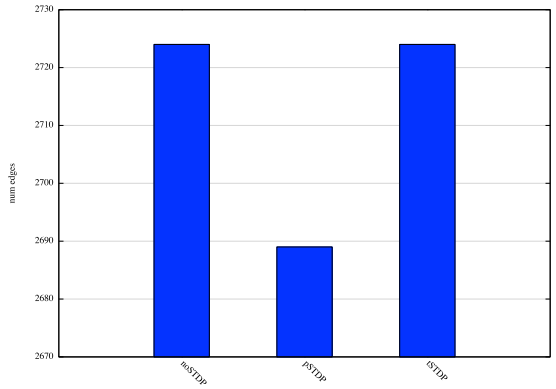
Why tSTDP does not achieve the same amount of pruning as pSTDP? Let node i send an edge to node j; and let j spike before i. For potentiation to take place, a second spike from j needs to take place. Let us refer to the first spike of j as j_1, and to its second spike as j_2, whereas the the only spike of i will be denoted as i_1. If node j's first spike and node i's only spike belong to the same avalanche we write $j_1, i_1 \in \mathrm{Av}_1$. As said above, for potentiation to occur, a second spike from j must take place. These leaves the following possibilities:

1. $j_1, i_1 \in \mathrm{Av}_1$ and $j_2 \in \mathrm{Av}_2$. Here, for the maximum potentiation to occur the inter-avalanche interval must be very small. Such a regime, in which firing rate is high, is associated with supercriticality and not with criticality.

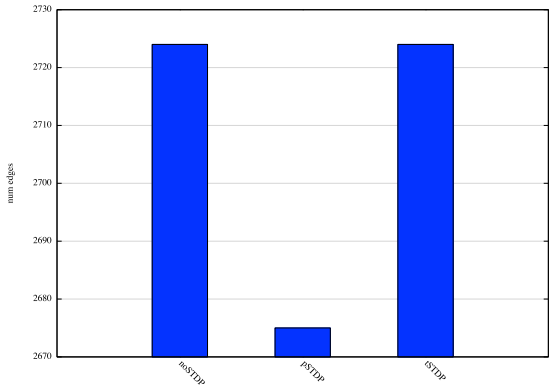
2. $j_1 \in \mathrm{Av}_1$ and $i_1, j_2 \in \mathrm{Av}_2$. Same as above. Here spikes are separated in two different avalanches that require to be as close as possible in order to potentiation be at its fullest.

3. $j_1, i_1, j_2 \in \mathrm{Av}_1$. Here what is required is that the three spikes occur in the same avalanche. For this to happen node j must be active twice in a single avalanche giving place to a non-Hamiltonian avalanche. As mentioned in Sect. 4.3.8 this occurs naturally in scale-free and random networks at criticality. Therefore we expect this to happen in this type of networks and to be non-existent in fully-connected structures.

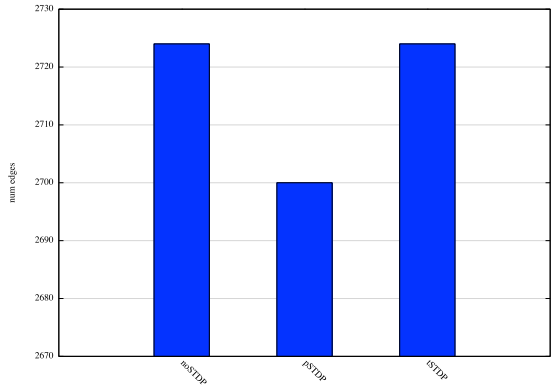
As we will show in the next section, the third case occurs in scale-free networks and this will result in weights reaching the largest allowed value of synaptic weight (w_{max}), whereas fully-connected networks fall in the above cases (1) and (2), in which the weights do not settle in the maximum synaptic weight and rather give rise to a bell-shaped distribution for the weights.



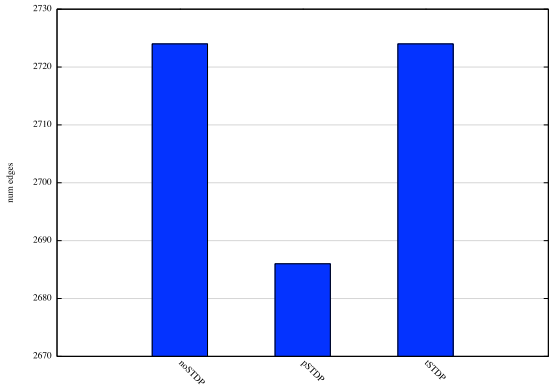
(a) Out-degree SF net with low mean CC.



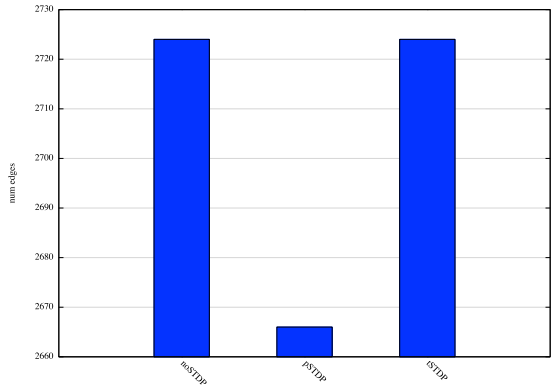
(b) In-degree SF net with low mean CC.



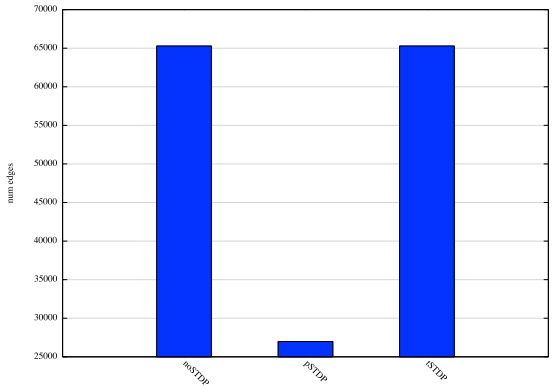
(c) Out-degree SF net with high mean CC.



(d) In-degree SF net with high mean CC.



(e) Random net.



(f) Fully-connected net.

Figure 5.11: Number of edges for networks of size $N = 256$ after no plasticity regime was applied (noSTDP), after pair-wise STDP was applied (pSTDP), and after triplet-wise STDP (tSTDP) was applied. Here, the systems ran for 4 million time steps. When STDP mechanisms were applied (either pSTDP or tSTDP), the system was allowed to develop for 2 million time steps with no STDP and afterwards these plasticity mechanisms were applied for the rest of the simulation. The only case in which networks were subject of synaptic pruning was when pair-wise STDP was used.

Let us consider the case where i fires before j. For depression to take place we require a second spike from node i. Similar to the case for potentiation, we have the following possibilities:

1. $i_1, j_1 \in \mathrm{Av}_1$ and $i_2 \in \mathrm{Av}_2$. As with the cases presented above, this situation requires a small inter-avalanche interval in order to get the maximum possible amount of synaptic depression.

2. $i_1 \in \mathrm{Av}_1$ and $j_1, i_2 \in \mathrm{Av}_2$. Same as above.

3. $i_1, j_1, i_2 \in \mathrm{Av}_1$. As with potentiation, this case requires non-Hamiltonian avalanches to be present in the system. As mentioned earlier, only scale-free and random estructures exhibit this kind of behaviour at criticality.

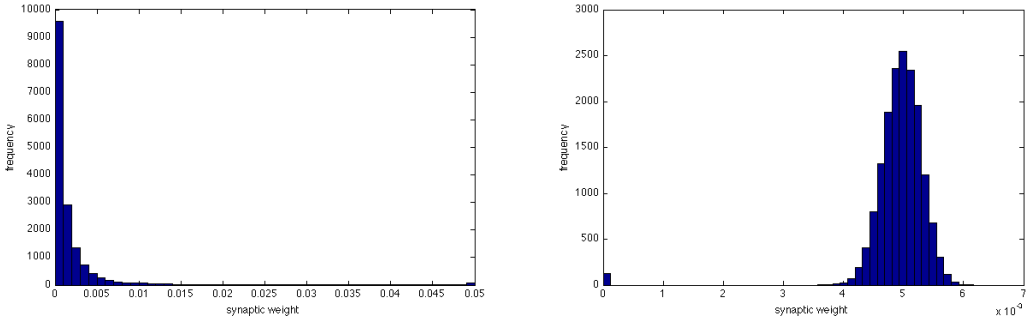
Unlike potentiation, we do not observe a large amount of synaptic depression occurring in our networks, and consequently no synaptic pruning either. This is due to the asymmetry of the STDP learning rule, in which potentiation is benefited over depression plus the fact that the rule requires either a small inter-avalanche interval in the system or non-Hamiltonian avalanches.

In the cases described above we have left out the case when each of the three spikes required for tSTDP belong to different avalanches; that is, $j_1 \in \mathrm{Av}_1$, $i_1 \in \mathrm{Av}_2$, and $j_2 \in \mathrm{Av}_3$. If the three spikes belonged to different avalanches, for either LTP or LTD to be at their maximum, the inter-avalanche interval should be small. This behaviour is compatible with supercriticality.

5.3.3 tSTDP gives rise to unimodal weight distributions in fully-connected networks

As mentioned in the last section, tSTDP gives rise to a bell-shaped weight distribution in fully-connected networks. Interestingly, the weight distribution for this type of structure under pSTDP mechanism takes a shape that resembles a power-law distribution with two modes: one close to 0 and another at w_{max}. Moreover, unlike pSTDP and other network structures, in tSTDP the weight distribution does not have w_{min} or w_{max} as minimum or maximum values.

In Fig. 5.12 we show a snapshot at time step 2×10^6 of the distribution of synaptic weights for a fully-connected network of 128 nodes for the plasticity regimes considered. Figure 5.12a shows the distribution for pair-wise STDP. Here, the shape of the



(a) pSTDP. (b) tSTDP.

Figure 5.12: Distribution of synaptic weights for a fully-connected network of 128 nodes under the two different plasticity regimes: (a) pair-wise STDP, and (b) triplet-wise STDP. Under pSTDP the distribution becomes bimodal at values w_{min} and w_{max}, whereas under tSTDP the distribution becomes bell-shaped with clear mean and variance. The distribution is far from w_{min} and w_{max}.

distribution resemble a power-law distribution. However, as simulation time increases the distribution becomes more and more bimodal, with one mode close to 0 and another at w_{max}. It is in this type of plasticity where we observe a small-world structure emerging. Figure 5.12b shows the weight distribution under tSTDP mechanisms. In this case, the shape resembles a Gaussian distribution with a clear mean value and variance. The distribution exhibits non-zero frequencies for zero weight, which corresponds to the diagonal of the adjacency matrix. This distribution does not have w_{min} nor w_{max} as minimum and maximum values.

5.3.4 Discussion

We started this section by wondering if by extending our plasticity model beyond simple spike pairs we would observe similar effects to the ones described in Sect. 5.2. For this purpose we considered triplets of spikes in the STDP model, which has been observed to fit better experimental data (see Sect. 2.2.4).

We observed that these mechanisms also take the system out of criticality. However, we do not observe a small-world structure emerging from fully-connected networks due to the practically non-existent synaptic pruning. The dynamics of our critical systems, characterised by metastable states interrupted by fast-relaxing dynamics (i.e. neuronal avalanches), does not allow for tSTDP to prune edges in the same proportion as with pSTDP. Our tSTDP rule acts only as a synaptic modulator driven by

the activity of the neurons.

Can critical dynamics coexist with STDP mechanisms? What we have observed in this chapter is that STDP modulates synaptic couplings based on activity in the system. Its long-term modulation has an overall effect that results in the impairment of criticality. To take the system back to the critical state we would require a mechanism to compensate for such modulation. The model developed by Levina et al. (2007) (Sect. 2.3.5) implements a compensatory mechanism of this type based on the dynamics of replenishing and exhausting synaptic resources. Thus, in this model critical dynamics can coexist with long-term plasticity mechanisms such as STDP. In the following chapter we present another compensatory mechanisms that allows the system to stay in the critical regime when STDP mechanisms are applied.

Chapter 6

Dynamic Synapses: Node-Success-Driven Plasticity

"In the economy, billions of individual processing units (people and firms) self-organize into productive networks. The mechanism that enables this self-organization is the transfer of money. Money flows in the opposite direction to goods and can be conceived of as a signal that indicates to a firm that its products are useful to others. By seeking to maximize their intake of money, producers are approximately maximizing their benefit to end-consumers."

—Lewis and Harris (2014)

As with the previous chapter, this chapter is situated in the context of dynamical synapses and it is devoted to a plasticity learning rule that we propose to compensate for the synaptic modulation induced by spike-timing-dependent plasticity (STDP). Moreover, although our initial interest was to propose a compensatory mechanism that would allow criticality to coexist with STDP, we soon were motivated by the idea of developing a local plasticity rule that would drive the system towards the critical state without any fine-tuning of the parameter α.

6.1 Introduction

Retro-synaptic signals have been defined as signals that travel across the synaptic cleft from the post-synaptic unit to the pre-synaptic one (Harris, 2008). Candidate molecules to be responsible of this type of retrograde communication between two neurons are neurotrophins (Zweifel et al., 2005), proteins (Kalinovsky and Scheiffele, 2004), cell-adhesion molecules (Dalva et al., 2007), lipids (Chevaleyre et al., 2006)

and even gases (Kishida and Klann, 2007). Previously, it has been reported that retrograde signaling is responsible for network development and apoptosis in the early stages brain formation (Purves, 1988). The mechanism of induced cell-death is related to neurotrophic factors travelling from post- to pre-synaptic cells in function of the relevance of their connection to the system. Additionally, this type of messaging has been proposed as a mechanism by which neurons compete with their neighbours to supply their targets with appropriate information in exchange for a "reward" that travels in the opposite direction of the action potential resembling the mechanism by which supply networks work in a free-market economy (Lewis and Harris, 2014). In this chapter we explore the idea that retro-synaptic signals might be also responsible for the emergence of critical behaviour in brain networks by studying their effects in a model of neuronal avalanches.

As the name suggests, neuronal avalanches are bursts of activity that resemble a domino effect triggered by spiking in groups of neurons. In the past decade, the observation of cascades of spontaneous neuronal activity, whose sizes and durations are distributed as a power-law, in brain tissue of rat cortex suggested a link between neural dynamics and phase transitions (Beggs and Plenz, 2003). Consequently, this triggered an avalanche of research in different experimental scenarios (Petermann et al., 2009; Klaus et al., 2011; Barbieri and Shimono, 2012; Friedman et al., 2012; Bellay et al., 2015; Shew et al., 2015), which confirmed the presence of neuronal avalanches *in vivo* as a previously unknown modality of brain operation.

It has been reported that neural networks whose dynamics are poised at the border of a phase transition, and whose activity is identified by power-law distribution of events, acquire several benefits for neural computation which include: optimal information transmission and maximum dynamic range (Kinouchi and Copelli, 2006), maximum information storage (Haldeman and Beggs, 2005), and stability of information transmission (Bertschinger and Natschläger, 2004). However, if brain networks operate at a critical point marking the border between regularity and randomness, how do they reach such dynamical regime?

Self-organised criticality (SOC) has been suggested as a concept to explain the emergence of critical dynamics in natural phenomena (Bak et al., 1988). It has been applied to describe the behaviour of phenomena as diverse as plate tectonics (Gutenberg and Richter, 1956), piles of granular matter (Frette et al., 1996), forest fires (Bak et al., 1990), mass extinctions (Bak, 1997), and crashes in the stock market (Johansen and Sornette, 1998), among others. This concept implies the existence of a critical

point that becomes an attractor in the collective dynamics of a system. Thus, the control parameter, which is an essential notion in the theory of critical phenomena and phase transitions, no longer needs to be fine-tuned by an external entity. Rather, the system implements some sort of feedback mechanism based on its internal dynamics that provides it with direct control over its control parameter.

In neural systems, self-organised critical models of neuronal avalanches have been put forward to explain the emergence of critical dynamics in brain tissue. These models can be divided in two categories, those that explain the emergence of critical behaviour through plasticity mechanisms such as spike-timing-dependent plasticity (Shin and Kim, 2006; Meisel and Gross, 2009), short-term plasticity (Levina et al., 2007), Hebbian-like plasticity processes (de Arcangelis et al., 2006; Pellegrini et al., 2007; de Arcangelis and Herrmann, 2010, 2012), and even non-Hebbian plasticity (Magnasco et al., 2009); and those that explain it through non-plastic mechanisms like axonal re-wiring (Bornholdt and Rohlf, 2000; Bornholdt and Röhl, 2003; Rybarsch and Bornholdt, 2014) or the balance between neuronal excitation and inhibition (Beggs and Plenz, 2003; de Arcangelis and Herrmann, 2010).

However, most models do not take into account the complex configuration of brain networks characterised by the presence of highly-connected units or hubs, as well as the presence of the small-world property which provides the system with an architecture efficient for information transmission across distant regions in the network (Sporns, 2010). Exceptions to this are the models by Pellegrini et al. (2007), de Arcangelis and Herrmann (2010), and de Arcangelis and Herrmann (2012), which introduce Apollonian networks, which capture both the notion of scale-invariant degree distributions and the small-world property.

Moreover, few of these models consider the relationship between learning and criticality, which could reveal the effects of this dynamical regime on cognition. Synaptic plasticity has been proposed as the neurobiological basis of learning and memory in brain networks. Some models of critical neuronal avalanches implement synaptic plasticity mechanisms either to explain the emergence of critical behaviour in the system (Meisel and Gross, 2009), or to provide the network with learning capabilities akin to artificial neural networks in machine learning (de Arcangelis and Herrmann, 2010). However, in the latter case it is not clear whether the critical regime survives the synaptic modulation induced by learning, whereas in the former case the implementation of spike-timing-dependent plasticity mechanisms is rather artificial as the system halts for predetermined periods of time in order to apply the synaptic modu-

lation based on the recent activity of the nodes in the system. The model developed by Levina et al. (2007) introduced depressing synapses which give rise to a form of short-term plasticity that results in the emergence of critical behaviour without tuning any control parameter. The resulting collective behaviour is robust against the synaptic modulation induced by spike-timing-dependent plasticity mechanisms. However, the model is not implemented in any pattern-learning scenario nor it takes into account complex network properties such as scale-invariant degree distributions.

In this chapter we put forward a novel model of self-organised critical neuronal avalanches in neural systems. Our approach differs from others by considering the introduction of retro-synaptic signals that inform the neuron about the behaviour at the level of its local surroundings. We propose a long-term synaptic plasticity mechanism based on this type of signals, which allows the system to reach the critical state autonomously, and such state is robust to the synaptic modulation induced by spike-timing-dependent plasticity mechanisms. Moreover, we take into account complex network features such as the small-world property and scale-invariant degree distributions which imply the presence of highly-connected nodes by considering scale-free structures with high clustering. By considering this type of topologies, our model considers the complex structure of brain networks which are neither regular nor random (Sporns, 2010).

6.2 Model

We present a basic model for neuronal avalanches (Eurich et al., 2002) that resembles the paradigmatic model of self-organised criticality: the sand-pile model of Bak and colleagues (Bak et al., 1988). The model consists of N non-leaky integrate-and-fire nodes. Each node j has a *membrane potential* h_j, which is a continuous variable that is updated in discrete time according to the equation:

$$h_j(t+1) = h_j(t) + \sum_{i=1}^{N} A_{ij} w_{ij} s_i(t) + I_{ext} \tag{6.1}$$

where A denotes the adjacency matrix with entries $A_{ij} = 1$ if node i sends and edge to node j, and $A_{ij} = 0$ otherwise, w_{ij} denotes the synaptic weight from node i to node j, $s_i(t) \in \{1,0\}$ represents the state of node i (active or quiescent, respectively) at time t, and I_{ext} denotes external input which is supplied to a node depending on the state of the system at time t. The mechanism of external driving works as follows: if there

is no activity at time t, then a node is chosen uniformly at random and its membrane potential is increased by a fixed amount through the variable I_{ext}. If $h_i(t)$ exceeds the threshold θ, then node i emits a spike, which changes the state of this node to active ($s_i(t) = 1$) and propagates its activity through its synaptic output. Afterwards, the node is reset, ie. $h_i(t+1) = 0$.

The propagation of activity might trigger subsequent activation of neurons. This results in cascading behaviour. Here, avalanches take the form of neuronal activity being propagated as a domino effect. While the system is relaxing, external drive stops. This guarantees that relaxation time occurs at a different time scale as the external driving. This corresponds to an infinite separation of time scales between external driving and relaxation dynamics that has been suggested as a necessary condition for critical behaviour (Jensen, 1998).

In the original model, the synaptic couplings w_{ij} are set by dividing a control parameter α by the mean connectivity of the network. The parameter α needs to be fine-tuned in order for critical dynamics to emerge. Here, we will set the synaptic weights randomly from an uniform distribution in the interval $(0,1)$ and let the system evolve according to a synaptic plasticity mechanism that we describe below.

We introduce a local measure of the performance of a node during simulation time that we call *node success*. The node success of node i at time t is the fraction of *out-neighbours* of this node that become active at time $t+1$ when node i spikes at time t, in other words:

$$\varphi_i(t) = \frac{\sum_{j=1}^{N} A_{ij} s_j(t+1)}{\sum_{j=1}^{N} A_{ij}} \tag{6.2}$$

where A is the adjacency matrix of the network, and $s_j(t+1)$ the state of node j at time $t+1$. The sum in the numerator runs for all out-neighbours j of node i.

Thus, node success measures the performance of an individual spike in terms of the subsequent spikes triggered by such initial activation, which occur within the out-neighbourhood of a given node. With this metric we construct a long-term plasticity rule that we call *node-success-driven plasticity* (NSDP). The intuition behind the NSDP learning rule is quite straightforward: if a node has very low success, then the node increases the synaptic weight *in all of its out-neighbourhood*, but only if the node is not spiking too often, that is, if the node possesses a low firing rate.

Formally, NSDP is defined by the following equations:

$$w_{ij}(t+1) = w_{ij}(t) + \Delta w_{ij}(\varphi_i(t), \Delta t_i) \tag{6.3}$$

where $\Delta t_i = t - t^i_{LS}$ denotes the difference between the spike of node i occurring at current time step t and its previous spike which occurred at t^i_{LS}. Whereas,

$$\Delta w_{ij}(\varphi_i(t), \Delta t_i) = A \exp\left\{\frac{-\varphi_i(t)}{B}\right\} - C \exp\left\{\frac{-\Delta t_i}{D}\right\} \tag{6.4}$$

with parameters A, B, C and D taken from $\mathbb{R}^+$. These parameters are phenomenological, that is, they are not meant to represent a particular physiological property of the living neuron. The action of the first term at the right hand of Eq. (6.4) *potentiates* the connection (i, j) according to the current node success of i, whereas the second term *depresses* it depending in the firing rate of i, which is succinctly captured by Δt_i.

Network structure plays a role on the way that collective dynamics emerge and behave in a system. Moreover, complex network properties such as the presence of hubs and the small-world property might affect the system's behaviour in a way that can not be anticipated by studying simple heterogeneous structures such as random networks. Therefore, we consider six different network types in order to compare the onset of the critical regime in systems under the mechanism of NSDP:

i. fully connected,

ii. random,

iii. out-degree scale-free with *low* mean clustering coefficient (CC),

iv. out-degree scale-free with *high* mean CC,

v. in-degree scale-free with *low* mean CC, and

vi. in-degree scale-free with *high* mean CC.

All networks are directed, and for every node in the network the out- and the in-degree is larger than zero. This means that every node in the system is able to receive and transmit spikes. Out-degree scale-free networks exhibit a power-law in the out-degree distribution, whereas their in-degree distribution follows a Poisson distribution. A power-law out-degree distribution implies the presence of nodes with many outgoing connections that we call *broadcasting hubs*. On the other hand, an in-degree scale-free network exhibits a power-law in-degree distribution. This implies the existence of nodes with many incoming connections that we call *absorbing hubs*.

The purpose of tuning the overall clustering coefficient in a network results in different degrees of the small-world property on it (Holme and Kim, 2002). Therefore,

high and low levels of the mean clustering coefficient result in high and low degrees of the small-world property in our scale-free networks. Taken together, the small-world property and the presence of hubs will result in collective dynamics that differ from fully-connected and random structures.

6.3 Assessment of critical behaviour

Our first concern lies on how to determine that the collective dynamics of the system are in the critical state. Critical dynamics feature the presence of power-law distribution of events (e.g. size and duration of avalanches). Thus, a simple and straightforward way to look for criticality is to inspect the distribution of avalanche sizes and durations. If the distribution resembles a power-law distribution then we have reasons to suspect that the system is undergoing a phase transition.

We assess the quality of such a power law through the mean-squared deviation $\Delta\gamma$ from the best-matching power law with exponent γ obtained through regression in log-log scales. Our choice of using this method is due to its simplicity and justified by the asymptotic unbiasedness of the estimation. When this error function is at its minimum, that is, when the data is best approximated by a power-law distribution of avalanche sizes with exponent γ, is when the system is at the critical state.

A power-law fit of the distribution of avalanche sizes is by no means a sufficient condition for a system to be critical, but it is a necessary one. Other tests for criticality that we carry out are:

i. Analysis of the relationships between critical exponents of avalanche sizes and their lifetimes: From the theory of critical phenomena, we know that at criticality the distribution of several observables follow power-laws with mathematical expressions linking each other (Sornette, 2004). In particular, there is a power-law positive correlation between avalanche sizes and their lifetimes which only occurs at criticality (Beggs and Timme, 2012; Shew et al., 2015; Bellay et al., 2015).

ii. Analysis of the largest eigenvalue Λ associated to the matrix of synaptic weights: Larremore et al. (2011) studied the spectral decomposition of the weight matrix of a system at criticality, and concluded through analytical inspections along with numerical simulations that the largest eigenvalue of the weight matrix governs the network's dynamic range. When this value is unity, the system is in a critical state and its dynamic range is maximised.

iii. Analysis of data collapse: At criticality the dynamics of a system show no particular scale, thus resulting in a fractal geometry of its observables (Sornette, 2004; Kadanoff, 2010). A standard procedure to analyse the critical regime in a model of neuronal avalanches is to observe if avalanches exhibit a fractal structure. To this purpose, the average "shape" of avalanches is estimated by keeping track of the the lifetime of an avalanche and the number of nodes involved at each avalanche step. If the system is at criticality, then avalanches of different sizes would exhibit the same shape up to a scaling function, in such a way that data could be collapsed in order to observe how avalanche shapes resemble each other (Friedman et al., 2012).

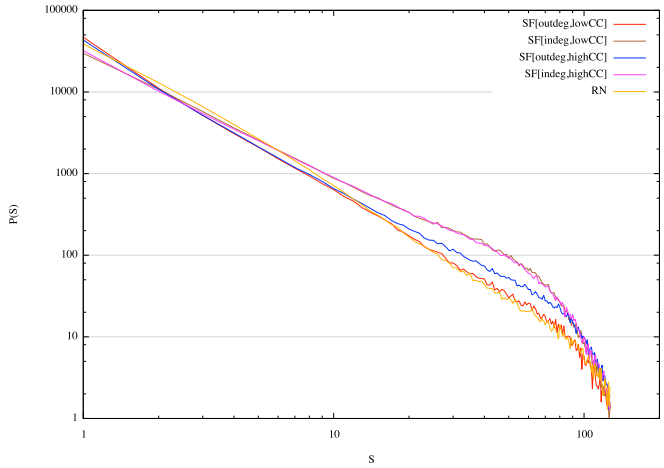
Our fist approach to analyse the critical behaviour of our systems will involve estimating the $\Delta\gamma$ function due to its ease to be be implemented in running time. When more formal analyses are required we will use the tests described above.

6.4 Results

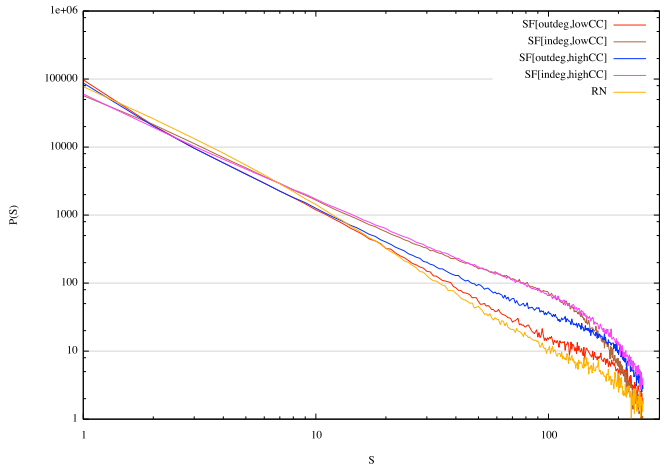
During our experiments, we consider three different system sizes: 128, 256, and 512 for each of the different network types considered (i.e. scale-free, random and fully-connected), and 20 different trials per network, which result from generating 20 different networks for each type. We fix the parameters B, C, and D in Eq. (6.4) to 0.1, 0.001, and 10, respectively; and we vary parameter A according to the system's topology and size. For heterogeneous topologies (scale-free and random) this value is set around 10^{-4}, 10^{-5}, and 10^{-6} for sizes 128, 256 and 512, respectively; whereas for fully-connected networks, this parameter was set around 2×10^{-6}, 5×10^{-7}, and 3×10^{-7}, for sizes 128, 256 and 512, respectively.

6.4.1 Distribution of avalanche sizes can be approximated by a power-law

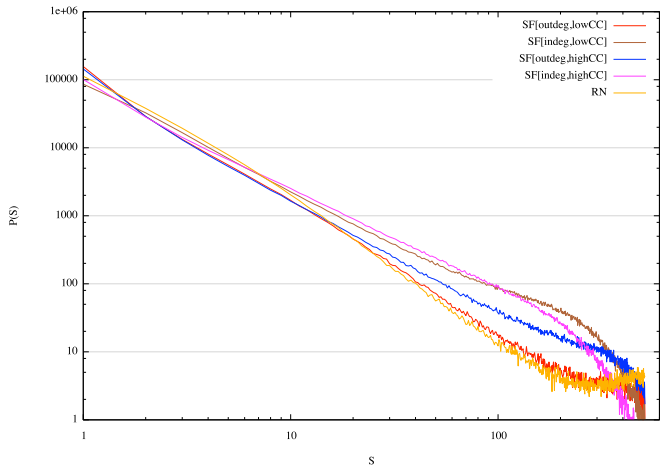
To start our examination of the critical state of our systems under NSDP we proceed by having a look at the distribution of avalanche sizes at the end of simulation time. We observe a power-law distribution of avalanche sizes that is identified by a straight line in a log-log plot of such a distribution until finite size effects take on. In Fig.6.1 we present only the mean values of these distributions for the sake of clarity.



(a) $N = 128$

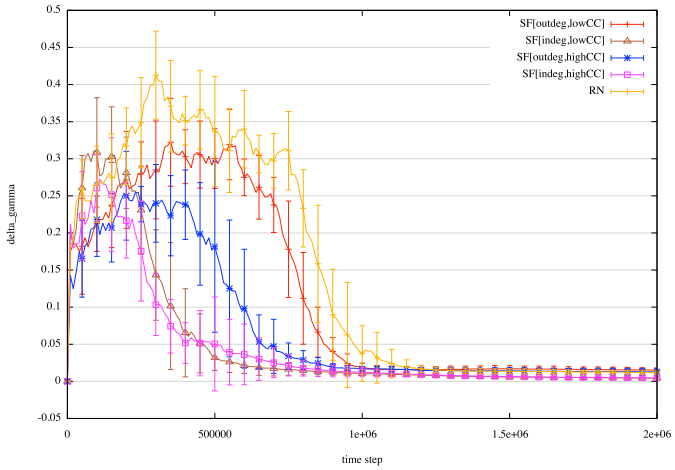


(b) $N = 256$

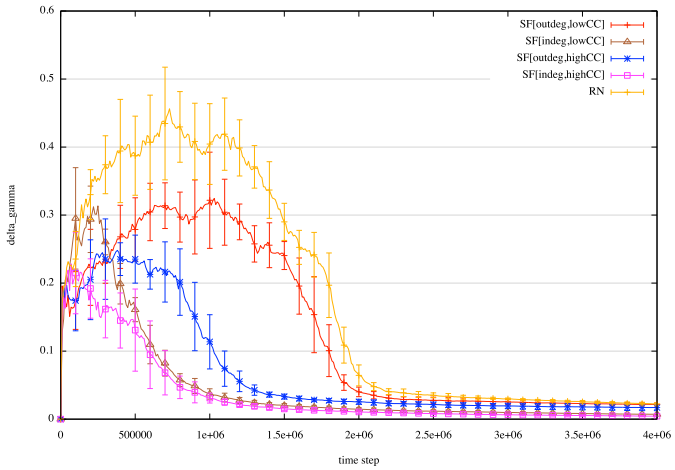


(c) $N = 512$

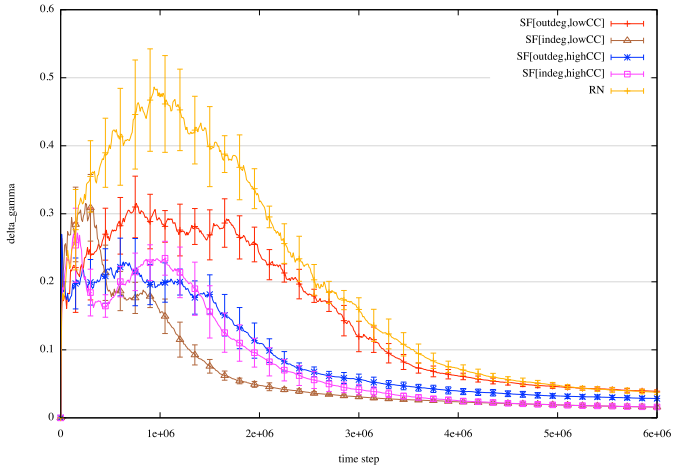
Figure 6.1: Double logarithmic plots of the distribution of avalanche sizes of scale-free and random nets under weight modulation by NSDP. The distribution follows a power-law for approx. two decades. Power-law distribution of avalanche sizes is a signature of criticality.



(a) $N = 128$

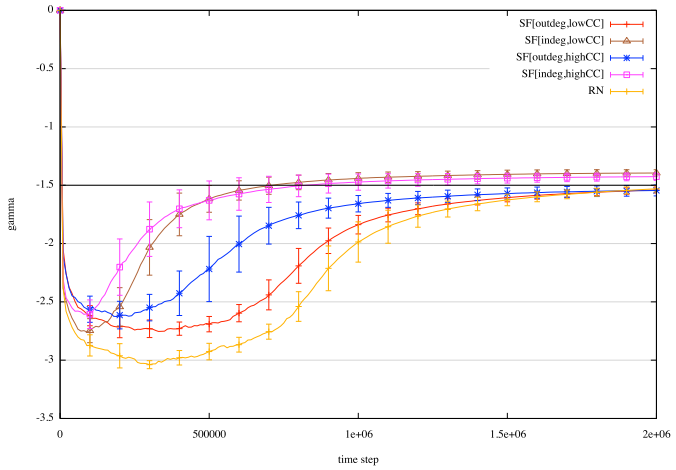


(b) $N = 256$

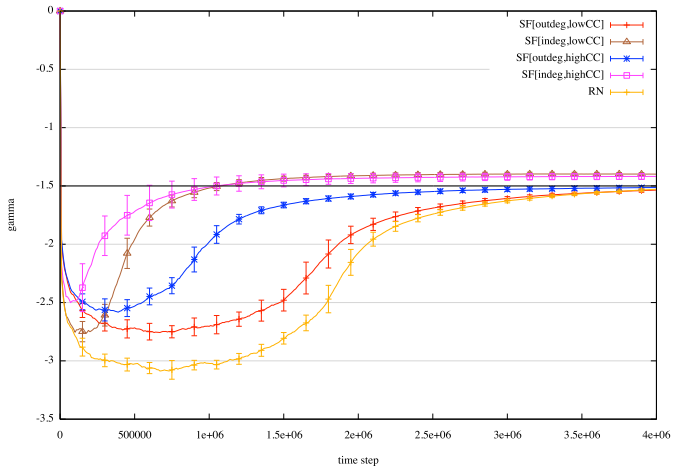


(c) $N = 512$

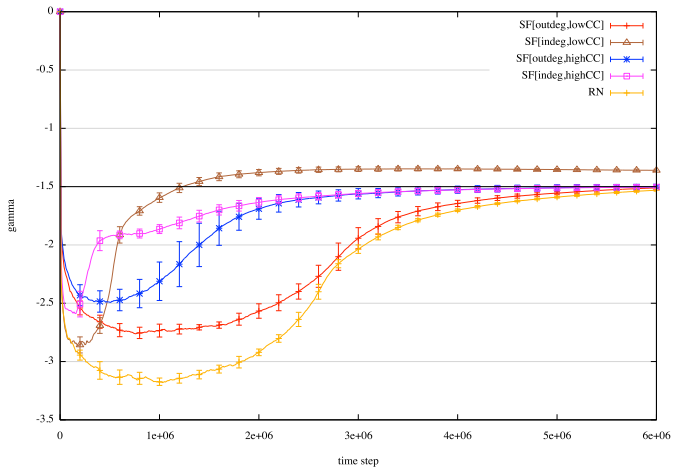
Figure 6.2: Evolution of $\Delta\gamma$ for systems under NSDP. Systems converge towards a low error (< 0.05) for the power-law approximation of avalanche size distribution, which implies that this distribution is well approximated by a power-law. (We present mean and standard deviations from our trials.)



(a) $N = 128$



(b) $N = 256$



(c) $N = 512$

Figure 6.3: Value of exponent γ of the power-law distribution of avalanche sizes of systems under NSDP. The value of the exponent lies close to -1.5 (black solid line) as it has been reported elsewhere for systems at criticality. The speed of convergence to this value depends on network topology. For instance, in-degree scale-free nets converge faster to $\gamma = -1.5$ due to the presence of absorbing hubs and clustering.

Next, we analyse the error of the power-law approximation to our data, that is, the $\Delta\gamma$ curve. We observe that these dynamical synapses achieve an error lower than 0.05 similar to the situation described for the static synapses in Sect. 4.3. This implies that the distribution of avalanche sizes are well approximated by a power-law distribution. We present the evolution of the fitting error in Fig. 6.2. In this figure it can be seen that the system converges to a low error after some time, that is, the error becomes low once the synaptic modulation induced by NSDP acts upon the random weights.

Now that we have seen that the distribution of avalanche sizes can be well approximated by a power-law, the remaining question is about the exponent of such power-law distribution. We observe that the value of such exponent lies close to -1.5 as it has been reported in several systems at criticality. We present the evolution of the value of the exponent γ in Fig. 6.3.

As it can be seen in Fig. 6.3, the speed of convergence to a low error regime and an exponent of $\gamma = -1.5$ depends on topology. In-degree scale-free networks converge faster than any other type of networks to a low error and to an exponent of -1.5. This can be explained by the presence of absorbing hubs and to some extent to the high amount of small-world-ness in these networks. On the other hand random networks and out-degree scale-free networks with low mean clustering coefficient exhibit the slowest convergence from the systems considered which can be attributed to low clustering within the network, and the absence of absorbing hubs.

6.4.2 Largest eigenvalue Λ close to unity

As we mentioned in Sect. 4.3.1, based on the work by Larremore et al. (2011) we analyse the value of the largest eigenvalue Λ associated to the matrix of synaptic weights. It has been shown that a system at criticality is identified by a value of $\Lambda = 1$. Table 6.1 in page 158 shows the value of the largest eigenvalue Λ for weight matrices under NSDP mechanisms (column Λ_{NSDP}). As it can be seen, the value of Λ is very close to unity for systems under this type of synaptic modulation. This implies that although initialised with random weights the system self-organises into a configuration of the weight matrix such that it gives rise to a largest eigenvalue Λ close to unity, which implies that the system is at criticality.

Moreover, we present the evolution of the largest eigenvalue Λ for networks of size $N = 128$ in Fig. 6.4. Our systems start with random synaptic couplings which yield a largest eigenvalue $\Lambda \ll 1$, however as NSDP mechanisms start to modulate the synaptic

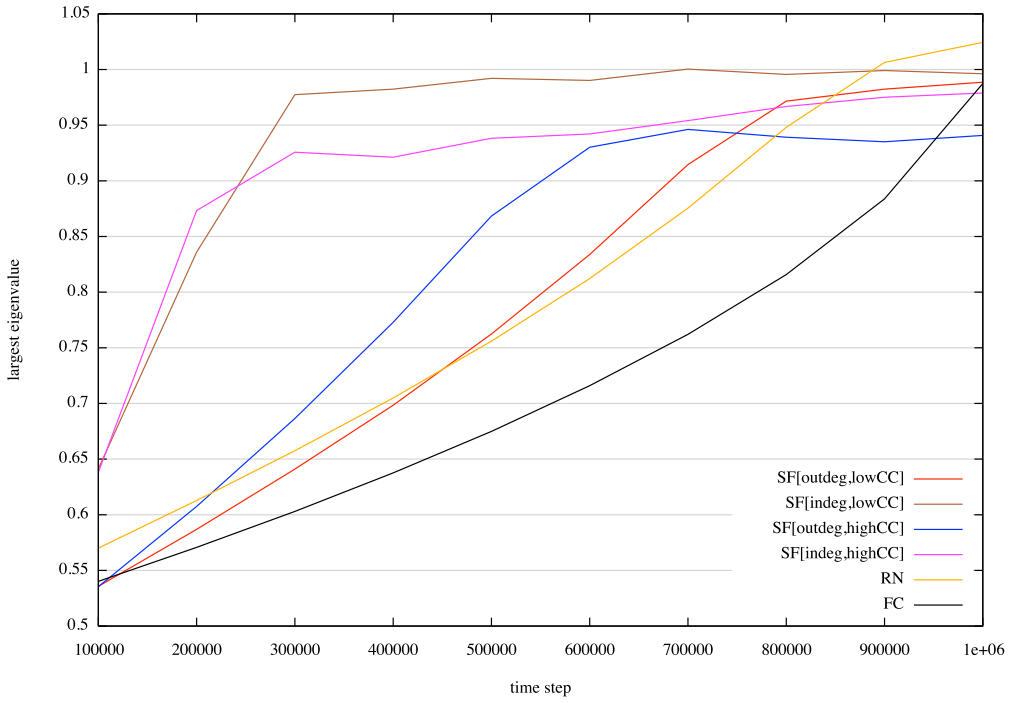


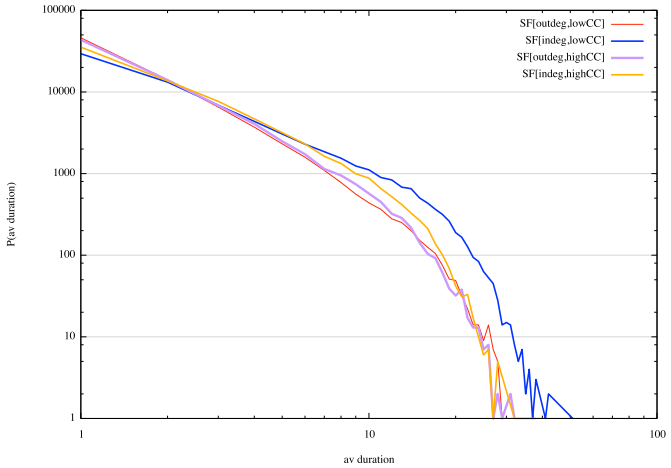
Figure 6.4: Evolution of the largest eigenvalue Λ associated to the matrices of synaptic weights for networks of 128 nodes. The value settles around unity once NSDP mechanisms take place. This implies that the system is at the critical state. (In this figure, we present the results of one trial, that is, one network per class.)

weights, the value of Λ increases and settles around unity, which identifies the critical state in the system.

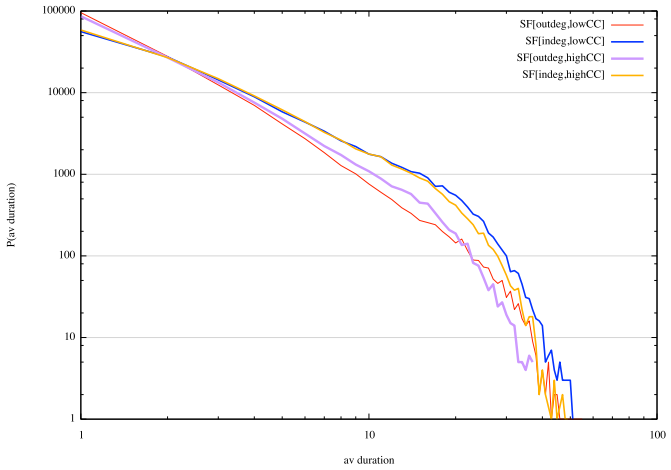
As in the last section, the convergence to a largest eigenvalue close to unity depends on topology. Here we observe a situation similar to that of the convergence to a low error regime and an exponent $\gamma = -1.5$. In-degree scale-free networks converge faster than any other type of network to $\Lambda = 1$. This can be explained by the presence of absorbing hubs in the network.

6.4.3 Distribution of avalanche lifetimes can be approximated by a power-law

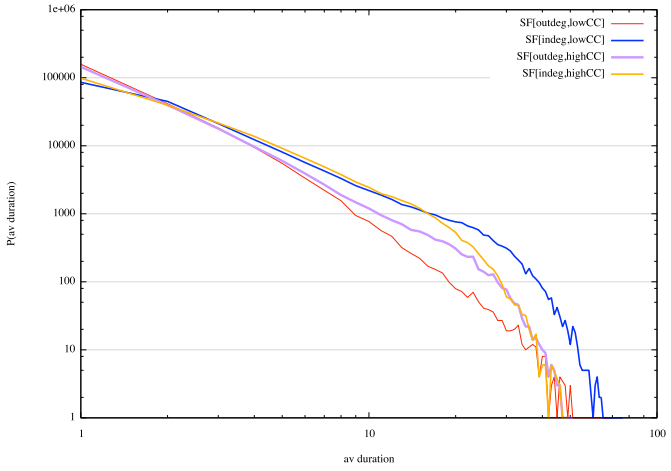
A system at criticality exhibits power-law distributions of several observables (Beggs and Timme, 2012). Similarly to what we presented for the case of static synapses in Sect. 4.3, our systems under NSDP exhibit distributions of avalanche durations that can be approximated by power-laws. This is shown in Fig. 6.5 for scale-free networks of all sizes considered.



(a) $N = 128$



(b) $N = 256$



(c) $N = 512$

Figure 6.5: Distribution of avalanche durations in log-log plots for scale-free nets of all sizes considered under the NSDP protocol. This distribution also exhibits a decreasing linear trend for a little more than a decade which implies that can be approximated by a power-law.

This fact is also related to the observation of a power-law correlation between avalanche sizes and their durations of the form $S \sim D^{\beta}$. This has been reported as a signature of criticality (Friedman et al., 2012; Bellay et al., 2015; Shew et al., 2015).

In Fig. 6.6 we shown this behaviour for all networks of size $N = 128$. Each blue dot in this plot is an avalanche denoted by its size and its lifetime, whereas the red straight line is the best-fit approximation to the data. This apparent linear trend in double logarithmic axes reveal a power-law relationship between avalanche sizes and lifetimes.

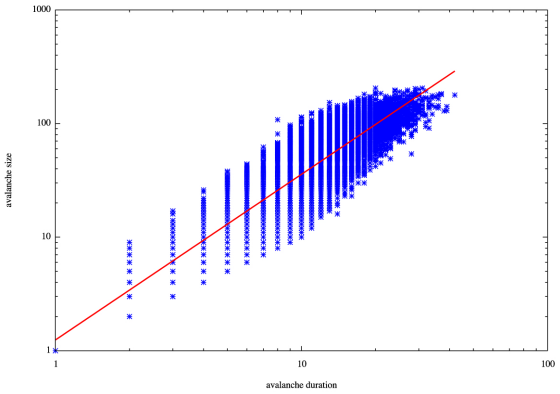
6.4.4 Avalanche shapes and data collapse

In this section we present the average shape of avalanches under the regime of NSDP. Recall that the avalanche shape is the curve that results by plotting the number of nodes involved at each step of the avalanche. In the current model every avalanche starts with one node becoming active. In the next step we expect to have one node on average becoming active as the result of the activation in a previous step.

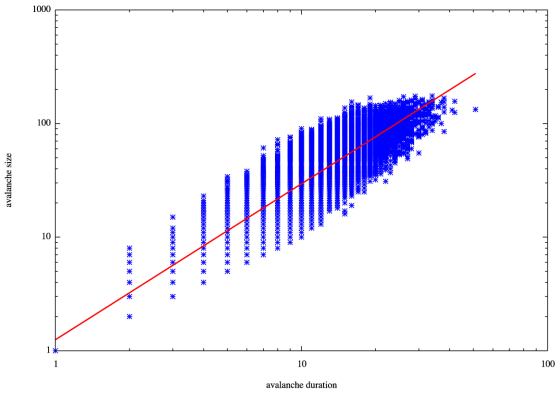
In Fig. 6.7 we present the average avalanche shapes for all types of network and all sizes considered in our protocol. As described in Sect. 4.3.4, the similarity of the avalanche shapes among different system sizes suggests a fractal structure underlying the avalanche dynamics. This is expected to occur in a system at criticality (Sornette, 2004; Friedman et al., 2012; Beggs and Timme, 2012).

In Fig. 6.8 we show an example of the data collapse from avalanche shapes of our in-degree scale-free networks with low mean clustering coefficient (Fig.6.7b). The fact that avalanche shapes from different system sizes can be re-scaled to match each other with certain accuracy not only suggests a fractal nature of the avalanche dynamics but also gives an insight of how avalanches might look in larger system sizes than the ones considered here.

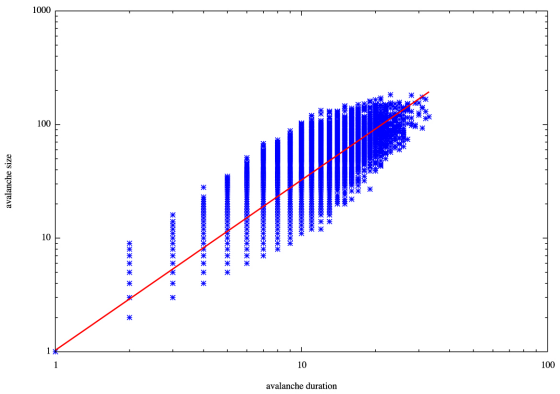
This concludes our examination of the critical state for our systems under the synaptic modulation of NSDP. As synaptic weights were initialised uniformly at random, and the system reached the critical state as a result of node activity based solely in the two rules that define our NSDP mechanism, we can safely say that this model exhibits *self-organised criticality*.



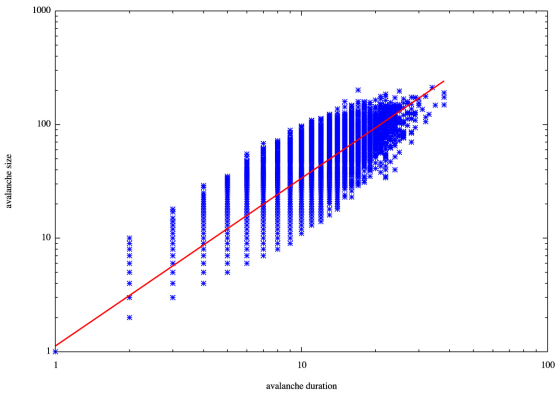
(a) Out-degree SF net with low mean CC. ($\beta = 1.45$)



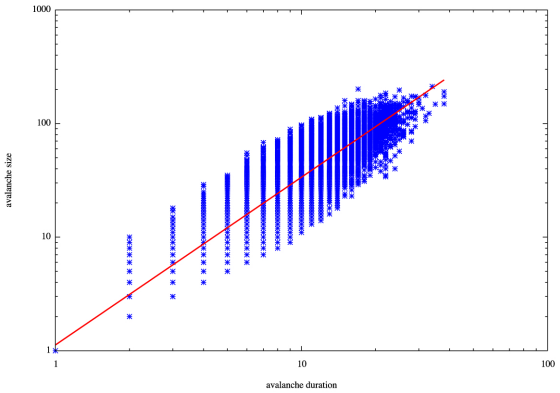
(b) In-degree SF net with low mean CC. ($\beta = 1.37$)



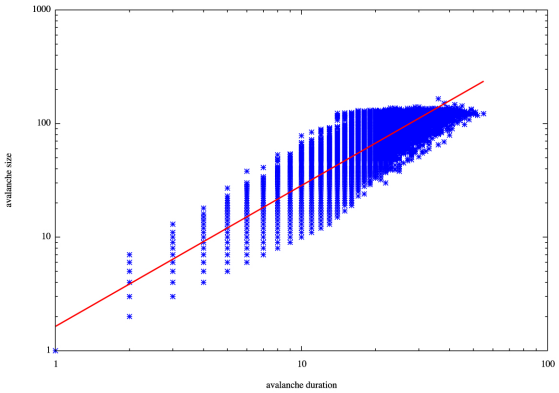
(c) Out-degree SF net with high mean CC. ($\beta = 1.49$)



(d) In-degree SF net with high mean CC. ($\beta = 1.47$)

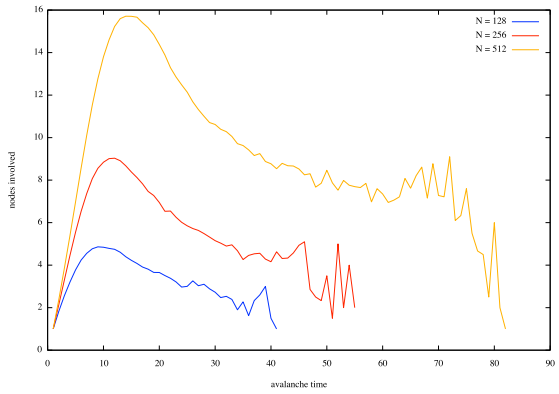


(e) Random net. ($\beta = 1.4$)

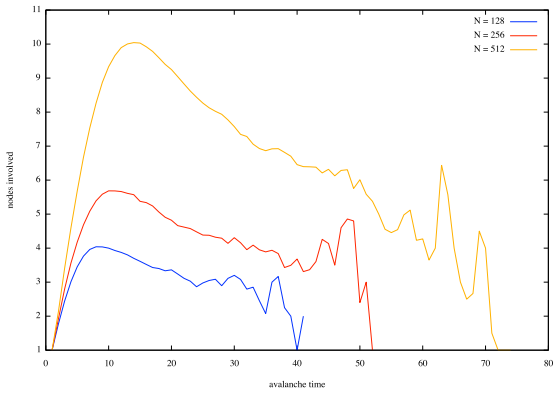


(f) Fully-connected net. ($\beta = 1.23$)

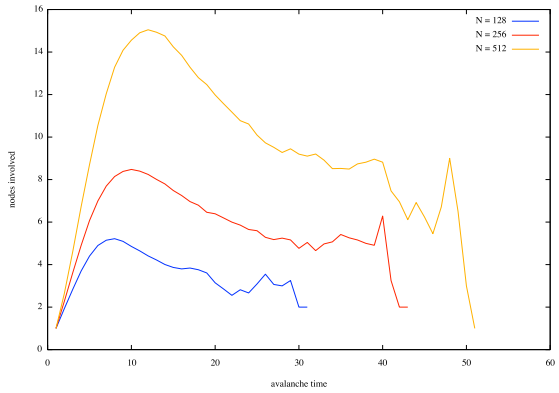
Figure 6.6: Linear relationship on logarithmic axes between avalanche sizes and durations for networks of 128 nodes. This behaviour implies a power-law correlation of the form $S \sim D^{\beta}$. In blue dots we present our data whereas in red we present the best-fit approximation. A relationship of this type is a signature of criticality.



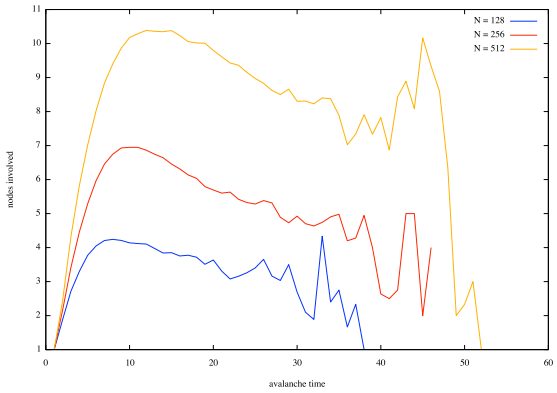
(a) Out-degree SF nets with low mean CC.



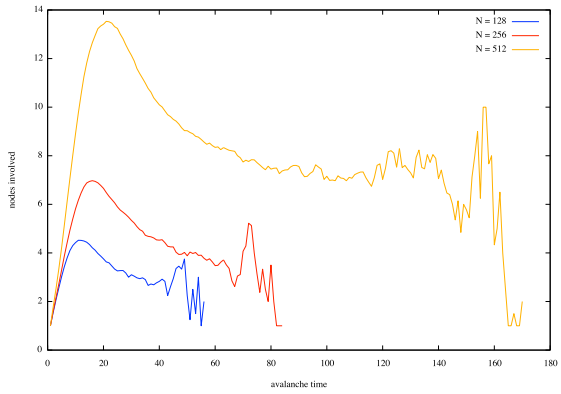
(b) In-degree SF nets with low mean CC.



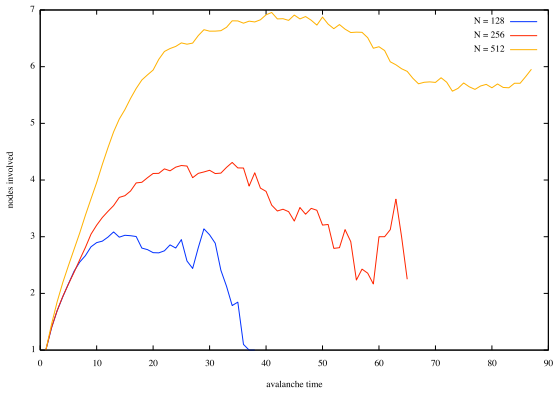
(c) Out-degree SF nets with high mean CC.



(d) In-degree SF nets with high mean CC.



(e) Random nets.



(f) Fully-connected nets.

Figure 6.7: Average avalanche shape for all networks and sizes considered under the regime of NSDP. The similarity of shapes among different sizes suggest the possibility of data collapsing in these systems. This fact implies that avalanches possess a fractal structure which is also a signature of criticality. (We present only mean values for the sake of clarity.)

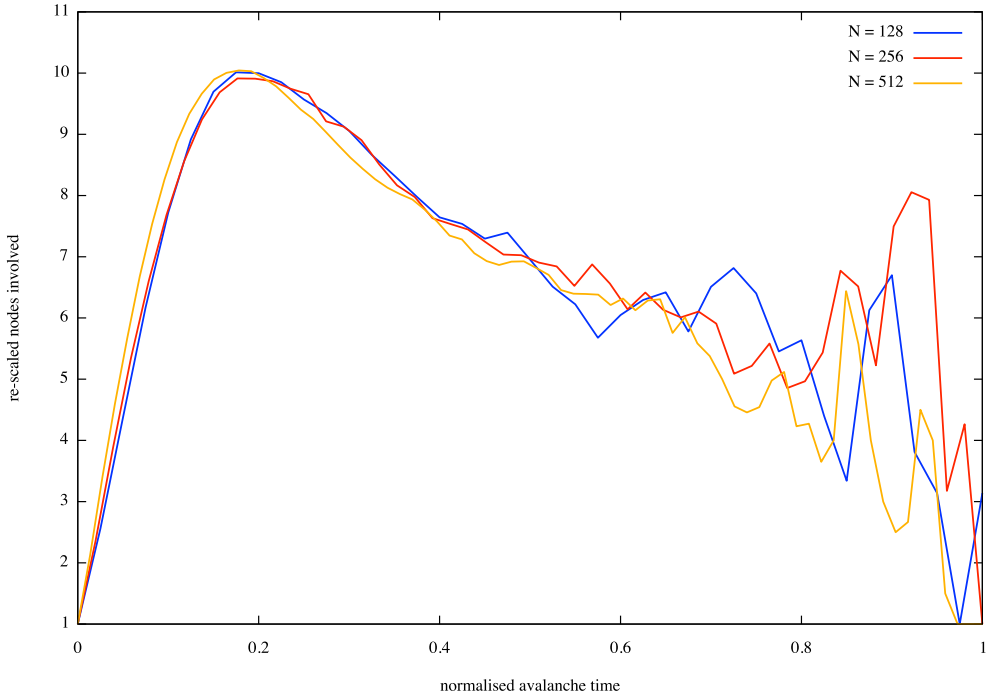


Figure 6.8: Data collapse for avalanche shapes from in-degree scale-free nets with low mean CC (Fig. 6.7b). Data collapsing reveals the fractal nature of avalanches in a system at criticality. (We present only mean values for the sake of clarity.)

6.5 NSDP & STDP

In Chapter 5 we showed how the critical state vanishes once STDP (either paired-wise or triplet-wise) take place in the dynamics of the system. This behaviour is captured by the value of the largest eigenvalue Λ shown in columns Λ_{pSTDP} and Λ_{tSTDP} for paired-wise STDP and triplet-wise STDP, respectively, in Table 6.1 in page 158. Thus, we wondered if by combining NSDP with STDP, we would obtain the synaptic modulation induced by STDP (which has been suggested as a basis for learning and memory) coexisting with criticality.

For this purpose we set up a protocol similar to that of Chap. 5, that is, we consider three system sizes (128, 256 and 512) and six different network types (see Sect. **??**). The synaptic weights in the coupling matrices associated to each of our networks are set according to the critical intervals found previously (see Sect. 4.2) by finding the values of the parameter α that allow the system to reach the critical state.

As with the protocol in Chapter 5, before applying STDP mechanisms into the system we allow criticality[1] to set in for a period of time in function of system size.

[1] In the form of large avalanches coexisting with small ones.

For networks of 128 nodes, we let the system reach the critical state for 10^6 time steps before introducing STDP mechanisms. For sizes 256 and 512 we do likewise for 2×10^6 and 3×10^6, respectively. Afterwards, we introduce STDP mechanisms for the same amount of time in which we drove the system to criticality without STDP. By the end of each simulation we end up with systems whose first half of the time were driven to criticality, whereas their second half were under STDP and NSDP regime.

STDP mechanisms were applied according to the equations presented in Sects. 5.2 and 5.3, for paired-wise STDP (pSTDP) and triplet-wise STDP (tSTDP), respectively. At the same time, we introduce NSDP mechanisms according to Eqs. (6.3) and (6.4) presented in this chapter. The value of the parameters A, B, C, and D in Eq. (6.4) are set depending on network type and size. The values of parameters B, C, and D are fixed to 0.1, 0.001, and 10, respectively; and we vary the value of A. For random and scale-free networks this parameter is set around 0.2, 0.1 and 0.02 for sizes 128, 256 and 512, respectively; whereas for fully-connected networks this parameter is set around 10^{-5}, 5×10^{-6}, and 3×10^{-6} for sizes 128, 256 and 512, respectively.

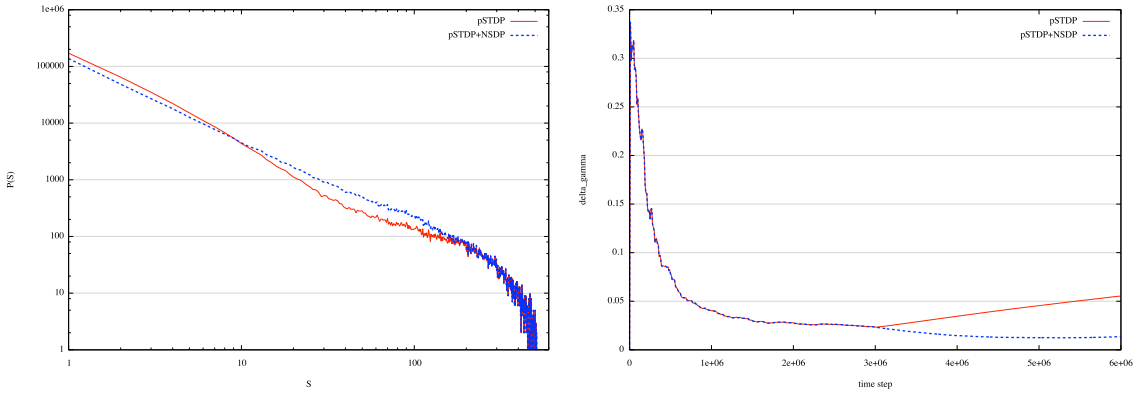
We report that the critical state was not lost once STDP mechanisms set in due to the synaptic modulation induced by the rules of NSDP. In the following sections we present our observations.

6.5.1 pSTDP

In Chapter 5 we described how fully-connected networks at the critical state develop a small-world structure when pSTDP mechanisms are introduced into the dynamics of the system. As well, we mentioned that the synaptic modulation induced by pSTDP impairs criticality. In this section we will briefly show how the critical state is not impaired by the action of pSTDP when NSDP is also introduced in the dynamics of the system. However, as NSDP does not allow synaptic weights to decrease too much[2] we do not observe synaptic pruning in the network, and therefore no small-world structure emerging.

Moreover, we can identify the status of the critical state by looking at the value of the largest eigenvalue Λ associated to the matrix of synaptic weights. Columns Λ_{pSTDP} and $\Lambda_{pSTDP+NSDP}$ in Table 6.1 show the value of the largest eigenvalue for systems under pSTDP mechanisms only and pSTDP mechanisms combined with NSDP mechanisms, respectively. The former shows values of Λ much less than unity, whereas the

[2]Due to the action of the first term at the right-hand side of Eq. (6.4), which tries to maximise the success of the node.



(a) Distribution of avalanche sizes.

(b) $\Delta\gamma$ curve.

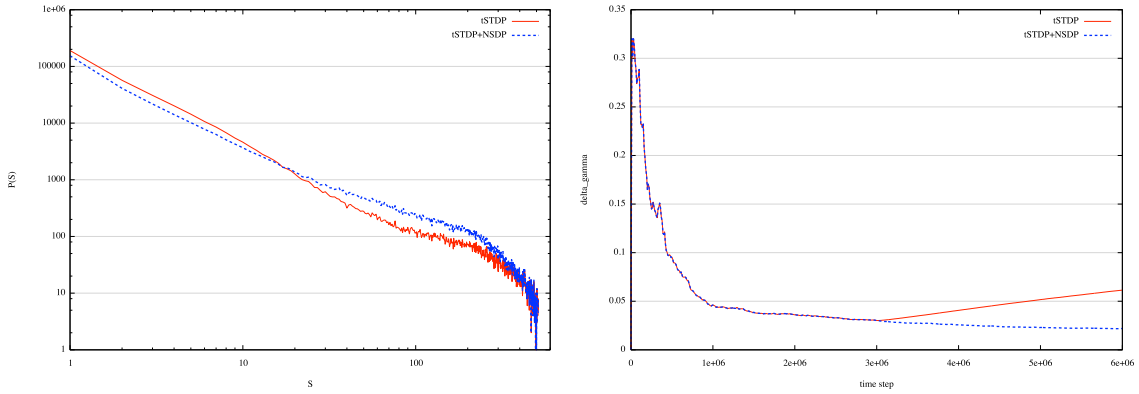
Figure 6.9: Distribution of avalanche sizes in logarithmic scales and deviation from best-fit power-law for in-degree scale-free networks with low mean CC of 512 nodes under two different regimes: pSTDP only (red, solid line), and pSTDP coexisting with NSDP mechanisms (blue, dashed line). (a) The shape of the distribution of avalanche sizes differs from a straight line when pSTDP sets in, which indicates a deviation from the critical regime. However, when NSDP is present in the system, the synaptic modulation induced allows the system to stay in the critical regime and exhibit a power-law distribution. (b) The error function $\Delta\gamma$ shows an increasing trend once pSTDP mechanisms take place in the system. However, if NSDP mechanisms are also present the system stays in the critical regime. A situation that is observed by the low error in the power-law fit.

latter shows values of Λ close to unity pointing out to the existence of critical behaviour in these systems.

Figure 6.9 shows an example of the action of NSDP on systems whose criticality has been impaired by pSTDP. When pSTDP sets in around time step 3×10^6 the synaptic modulation induced results in the vanishing of the critical state (see red solid line in Fig 6.9b). However, if pSTDP mechanisms are accompanied by NSDP mechanisms the critical regime is preserved (see blue dashed line in Fig 6.9b).

6.5.2 tSTDP

Now it is the turn of the triplet-spike version of STDP (tSTDP) to be combined with NSDP when the system is at criticality. As mentioned in Sect. 5.3, this type of plasticity exerts a milder synaptic modulation than pSTDP due to the nature of its mechanisms



(a) Distribution of avalanche sizes.

(b) $\Delta\gamma$ curve.

Figure 6.10: Distribution of avalanche sizes in logarithmic scales and deviation from best-fit power-law for in-degree scale-free networks with high mean CC of 512 nodes under two different regimes: tSTDP only (red, solid line), and tSTDP coexisting with NSDP mechanisms (blue, dashed line). (a) The shape of the distribution of avalanche sizes differs from a straight line when tSTDP sets in, which indicates a deviation from the critical regime. However, when NSDP is present in the system, the synaptic modulation induced allows the system to stay in the critical regime and exhibit a power-law distribution. (b) The error function $\Delta\gamma$ shows an increasing trend once tSTDP mechanisms take place in the system. However, if NSDP mechanisms are also present the system stays in the critical regime. A situation that is observed by the low error in the approximation.

and the avalanche dynamics of the system at the critical state. Nevertheless, the critical state vanishes once tSTDP sets in. This can be identified by the value of Λ in Table 6.1 under the column Λ_{tSTDP}. Here, the value of Λ is much less than unity which identifies the vanishing of criticality.

If we introduce NSDP mechanisms as specified by Eqs. (6.3) and (6.4) at the same time as tSTDP, the critical state is preserved. This is identified by the value of Λ in Table 6.1 under the column $\Lambda_{tSTDP+NSDP}$. Here, we can see that the value of the largest eigenvalue is close to unity, as expected from a system at criticality.

In Fig. 6.10 we show an example of an in-degree scale-free network with high mean clustering coefficient under two regimes: one in which only tSTDP is present (and sets in at time step 3×10^6) and another in which both tSTDP and NSDP mechanisms take place at the same time. It can be seen that tSTDP impairs the critical state (red solid lines in Fig. 6.10), but taken together tSTDP and NSDP preserve the critical regime

(blue dashed lines in the same figure).

6.6 Discussion

Node-success-driven plasticity is a mechanism that achieves self-organised criticality in the systems examined. The weight modulation induced through the two rules in the NSDP mechanism is sufficient to drive the system to the critical state. Interestingly, this modulation is activity-dependent and local. The nodes know nothing about the critical state nor about power-law distributions of events; the only thing they care about is to maximise their node success (in response of the activity of their out-neighbours) as long as they are not firing to often, that is, if their firing rate is not very high. If nodes follow these two rules, they will reach the critical state no matter how their synaptic weights were initialised in the first place.

Our presentation in Sect. 6.5 briefly described how node-success-driven plasticity allows the system to stay in the critical regime while spike-timing-dependent plasticity mechanisms occur in the system. Thus, this implies that learning and memory mechanisms could in principle occur during criticality in neural systems as long as a compensatory process (such as NSDP) is also present in the system. In such a regime, a neural system could acquire the benefits of these two worlds, namely, the capacity of learning through STDP modulation, and optimal information processing through critical dynamics.

This leads us to another important question regarding NSDP mechanisms. *How realistic is this type of synaptic plasticity in real brain networks?* Is there any evidence that this type of plasticity has any biological plausibility? In the next section we explore these questions.

6.6.1 Retro-synaptic signals

> *"The evidence we will present is circumstantial; however, the convergence of multiple lines of evidence, from molecular mechanisms to studies in behaving animals, suggests at least that this hypothesis deserves serious experimental investigation."*
>
> *—Harris (2008)*

The work presented in here suggests an hypothesis. The hypothesis that through signals that carry information regarding the success of a spike from a pre-synaptic unit

in terms of the subsequent spikes triggered at the post-synaptic neighbourhood a neural system is able to reach a critical state without external supervision or tuning.

A signal of this kind will have to travel in the opposite direction of an emitted spike, that is, from post- to pre-synaptic unit, without it being a spike by itself. Rather, this type of retro-synaptic signal will have to be some type of molecule emitted by a post-synaptic unit at the moment it becomes active and emits a spike. The released molecules would travel in the extracellular medium until they are captured by vesicles in the pre-synaptic unit, which in turn will modulate its outbound synaptic strength in function of the seized molecules from the post-synaptic neighbourhood.

Another mechanism of retro-synaptic signal might start with the release of molecules (other than neurotransmitters) from the pre-synaptic unit at the moment of a spike. This molecules could travel back to the pre-synaptic neuron carrying information about the success of the most recent spike emitted by this unit. For instance, if the post-synaptic unit seizes such molecules at the moment of spiking, then the absence of those molecules in the extracellular medium and the consequent inability of the pre-synaptic node to catch them back would inform this neuron that its most recent spike was successful in triggering subsequent spikes.

In the Introduction to this chapter, we mentioned some candidate mechanisms for retro-synaptic signaling in brain networks. In our particular context, we would require that this messaging also implements long-term synaptic modulation over post-synaptic units. To date, retrograde signals have been proposed as a mechanism by which neurons compete with their neighbours to supply their targets with appropriate information in exchange for a "reward" that travels in the opposite direction of the action potential. This mechanism has an analogy in the way supply networks work in a free-market economy; and it is also analogous to the mechanism of node-success maximisation described in this chapter. It has been suggested that this type of competition in neural systems might allow the brain to self-organise into functional networks giving rise to some sort of *Darwinian neurodynamics* (Lewis and Harris, 2014).

Thus, the idea of retro-synaptic signals modulating synaptic efficacies in function of successful spikes triggering subsequent activity is not an unsound one. Moreover, as we have shown in this chapter, this mechanism of plastic modulation could be the basic ingredient by which neural systems exhibit complex behaviour such as critical dynamics with all the benefits on information processing implied by it.

Retro-synaptic signaling might provide a biological interpretation of the first term of Eq. (6.4) for the NSDP mechanism. What about the second term which penalises

the increment of synaptic weight if a node's firing rate is high? How do single neurons know if they are spiking too often? An answer to this question can be given in terms of the depletion of synaptic resources of a single neuron. If a neuron is spiking often, its synaptic resources deplete, and a certain amount time is required in order to be replenished. The interplay between depletion and replenishment is the underlying mechanism behind the model of Levina et al. (2007) that has been shown to exhibit self-organised criticality.

6.6.2 The back-propagation algorithm

In 1957, the psychologist Frank Rosenblatt, a pioneer of artificial intelligence (AI), inspired by biology and theories of how brain cells work developed the *perceptron*: an artificial neural network (ANN) that was capable of identifying very simple geometric shapes. Some years later, in 1969 another couple of AI pioneers, Marvin Minsky and Seymour Papert, published a book called "Perceptrons", in which they showed how this ANN was unable to learn (or solve) simple logic functions, such as the *XOR* logic function. This shocked the community devoted to the advancement of ANNs, and subsequently led to a loss of interest in the field. Some researches continued to study ANNs and came to the conclusion that by adding more layers to the network, it would be able not only to solve the *XOR* problem but also to learn more complex functions. The problem, however, was that there was no optimum algorithm to train these multilayer ANNs. By the end of the 1980s, the research on ANNs would become mainstream once again after Yann LeCun and Geoffrey Hinton, developed independently a method to efficiently train the multilayer perceptron, namely, *the backpropagation algorithm* (Le Cun, 1986; Rumelhart et al., 1988). However, by the end of the 1990s ANNs would fall again in disuse when more efficient AI techniques appeared, such as *support vector machines*. This did not prevented some researchers to continue the research on ANNs, and inspired by the deep architecture of the brain cortex, Hinton and colleagues developed an algorithm (based on the backpropagation algorithm) to train deep ANNs (Hinton et al., 2006). This brings us to what today is known as *deep learning*.

As the name suggests, the backpropagation algorithm uses retrograde signals to communicate the fitting error back to the network while training an ANN in supervised learning scenarios[3]. The error is then used to estimate the weight updates at each

[3]Or unsupervised scenarios such as training autoencoders.

network connection, afterwards the error is re-estimated with the new weights and the process is repeated until convergence. Unlike the neurons that we have considered here, ANNs consist of non-spiking nodes, that is, non-threshold units. Therefore, there is no direct map between ANNs and networks of spiking nodes. Spiking nodes are all-or-none units that fire only when their membrane potential goes beyond a certain threshold. On the contrary, nodes in ANNs are continuous gateways whose activation function needs to be differentiable everywhere for the backpropagation algorithm to work. We believe that node-success-driven plasticity could in principle supply the basic requirements for the design of a backpropagation algorithm in spiking neurons. However, to this day this line of research is still unexplored.

There have been some efforts towards developing versions of the backpropagation algorithm for networks of spiking units, such as the *SpikeProp algorithm* and others (Bohte et al., 2002; Ponulak and Kasinski, 2010; Sporea and Grüning, 2013). Moreover, recently there has been a line of research regarding the plausibility of developing deep learning methods based on spiking networks (Nessler et al., 2013; Sheik et al., 2013; O'Connor et al., 2013). It has also been shown that deep learning architectures based on spiking neurons outperform those based on non-spiking neurons as the latter are expensive to implement on serial computers (O'Connor et al., 2013). Additionally, it has been shown that neuromorphic electronic architectures based on spiking units offer an efficient solution for developing compact sensory-motor neural processing systems for robotic applications (Sheik et al., 2013). When deep learning techniques are implemented in a VLSI architecture, the system benefits from the parallel and asynchronous nature of spiking units responding to signals coming from external sensors. Lastly, it has been suggested that Bayesian computation emerges from neural systems based on spiking nodes through the implementation of a version of STDP (Nessler et al., 2013). This links the concepts of spiking neural networks and STDP to that of the *Bayesian Brain Hypothesis* (BBH) (Knill and Pouget, 2004; Friston, 2010; Colombo and Seriès, 2012). The BBH states that the brain either at the level of circuits of neurons or at the level of concepts[4] performs some kind of Bayesian computation to update beliefs based on current sensory input, and this sensory information is represented probabilistically in the form of probability distributions. In this view, perception can be thought of sampling from a posterior distribution estimated through Bayes rule from priors that are updated via sensory stimuli.

Coming back to the backpropagation algorithm and ANNs, to the best of our

[4]Depending on the version of the BBH.

knowledge there has not been a study which relates deep learning methods based on spiking neurons and self-organised criticality[5]. We believe that NSDP could provide the link between these two concepts, and at the same time provide a learning algorithm based on the ideas behind backpropagation. This is an interesting direction of research due to the benefits on information processing that the critical regime entails.

6.6.3 To tune or not to tune

"To what extent is tuning needed? And, if it is *needed, does it then make any sense to speak of* self-organization*? I think that some degree of tuning is inevitable"*

—Jensen (1998)

Lastly, we conclude the exposition of this chapter by considering the question of fine-tuning. In this chapter, we have shown how systems self-organise into a critical state through NSDP. Thus, we became relieved from the burden of fine-tuning the control parameter α, but instead we acquire a new burden: that of estimating the appropriate values for parameters A, B, C and D in Eq.(6.4). Is there no way to be relieved from tuning *any* parameter in the system?

As the quote in the beginning of this section affirms, it seems that some amount of tuning is inevitable. Also, the issue of tuning or not tuning depends mainly on what we understand by control parameter. This notion, along with the concepts of order parameter and phase transitions are inherited from the theory of statistical mechanics. There, a control parameter can be thought of a knob or dial that when turned the system exhibits some quantifiable change. We say that the system self-organises if nobody turns that knob but the system itself. In order to achieve this, the elements comprising the system require some kind of feedback mechanism to be able to change their inner dynamics in response to their surroundings[6].

Denote by E a system in which we observe critical dynamics when the control knob is turned and left in a particular configuration. System E requires an external entity to turn the dial in order to make the system exhibit critical behaviour in some order parameters (e.g. critical exponents). In order to make E be able to turn the dial by itself we would have to mess with its internal configuration, and here is the *new*

[5] And possibly even the Bayesian Brain Hypothesis.

[6] For instance, in the model by Levina et al. (2007) this is achieved by the combined action of large cascading and the replenishing of synaptic resources; in our model this is achieved by the maximising of node success as long as the firing rate is not high.

need for tuning emerges. In the new internal configuration we would have to "plug" some new cables and make some new connections in a *very particular way* that at the end will result in system $E^\star$, which is now able to turn its own control knob. Some purists might affirm that we have just changed the place of fine-tuning from control dial to internal mechanisms. If that is the case, then the question about fine-tuning control parameters is in the realm of definitions and semantics. System E is the Eurich model, whereas $E^\star$ is this model under NSDP mechanisms. The latter does not require an external entity to turn the dial for the system to exhibit critical dynamics. However, its internal dynamics are configured in a *very particular way* in order to allow feedback mechanisms at the level of individual elements. Did we fine-tune their configuration? Yes. Otherwise, we would not have achieved what was desired, as nothing comes out of nothing. Did we change control parameter from α to A, B, C and D? No, the control parameter is still intact, and now it is in the hands of the system. However these agents, the nodes, know nothing about criticality; their only objective is to maximise their success as long as they are not spiking too much. Unlike them, when we -as external entities- turned the dial, we had the purpose of taking the system into (or out of) the critical regime. When we say that the control parameter is now in their hands we do not mean that their objective is to give rise to a particular collective behaviour. Rather, the control parameter (the dial) is changing as a result of their individual objectives.

Lastly and most importantly, the new configuration stresses the difference between global and local mechanisms. The control parameter α (the dial) is an external quantity that observes and governs the global (ie. the collective), whereas NSDP provides the system with local mechanism that have an effect over the collective. As mentioned in Sect.2.1.1, this is the main feature of a complex system.

Type	Subtype	Size	Λ_{static}	Λ_{pSTDP}	Λ_{tSTDP}	Λ_{NSDP}	$\Lambda_{pSTDP+NSDP}$	$\Lambda_{tSTDP+NSDP}$
Out-degree Scale-free	Low Mean CC	128	0.906 ± 0.029	0.27 ± 0.009	0.35 ± 0.01	0.99 ± 0.011	1.02 ± 0.04	1.08 ± 0.06
		256	0.9 ± 0.02	0.45 ± 0.007	0.52 ± 0.016	1.006 ± 0.004	1.026 ± 0.04	1.05 ± 0.02
		512	0.95 ± 0.01	0.7 ± 0.005	0.85 ± 0.002	1.01 ± 0.003	1.05 ± 0.02	1.11 ± 0.02
		1,024	0.97 ± 0.008					
	High Mean CC	128	0.89 ± 0.04	0.26 ± 0.01	0.33 ± 0.02	0.95 ± 0.011	1.049 ± 0.12	1.02 ± 0.09
		256	0.91 ± 0.03	0.43 ± 0.01	0.51 ± 0.01	0.96 ± 0.007	1.03 ± 0.04	1.06 ± 0.05
		512	0.91 ± 0.01	0.63 ± 0.01	0.73 ± 0.009	0.97 ± 0.009	1.03 ± 0.02	0.95 ± 0.01
		1,024	0.94 ± 0.01					
In-degree Scale-free	Low Mean CC	128	0.98 ± 0.02	0.21 ± 0.005	0.32 ± 0.009	0.99 ± 0.008	1.02 ± 0.05	1.13 ± 0.06
		256	0.99 ± 0.01	0.4 ± 0.006	0.49 ± 0.01	0.99 ± 0.003	1.08 ± 0.036	1.14 ± 0.02
		512	1.0006 ± 0.006	0.62 ± 0.006	0.68 ± 0.002	1.008 ± 0.004	1.008 ± 0.014	1.01 ± 0.02
		1,024	1.001 ± 0.004					
	High Mean CC	128	0.96 ± 0.02	0.23 ± 0.013	0.307 ± 0.02	0.96 ± 0.012	1.006 ± 0.102	1.09 ± 0.11
		256	0.99 ± 0.02	0.408 ± 0.016	0.47 ± 0.01	0.98 ± 0.012	1.07 ± 0.04	1.143 ± 0.06
		512	1.002 ± 0.014	0.61 ± 0.01	0.62 ± 0.007	0.95 ± 0.013	1.04 ± 0.04	1.004 ± 0.026
		1,024	1.01 ± 0.017					
Random		128	0.92 ± 0.012	0.28 ± 0.004	0.34 ± 0.003	1.02 ± 0.005	1.16 ± 0.02	1.04 ± 0.05
		256	0.99 ± 0.022	0.48 ± 0.003	0.53 ± 0.004	1.027 ± 0.007	0.98 ± 0.008	1.07 ± 0.015
		512	0.97 ± 0.001	0.77 ± 0.001	0.87 ± 0.001	1.03 ± 0.001	1.14 ± 0.001	0.98 ± 0.003
		1,024	0.98 ± 0.005					
Fully-connected		128	0.91 ± 0.034	0.098 ± 0.007	0.63 ± 0.0001	0.98 ± 0.001	0.704 ± 0.0001	0.901 ± 0.007
		256	0.93 ± 0.0006	0.23 ± 0.0062	0.78 ± 0.0002	0.99 ± 0.0008	0.72 ± 0.014	0.94 ± 0.002
		512	0.95 ± 0.001	0.33 ± 0.004	0.88 ± 0.0005	1.004 ± 0.001	0.75 ± 0.009	0.97 ± 0.0008
		1,024	0.98 ± 0.0005					

Table 6.1: Largest eigenvalue Λ of matrix W of synaptic weights.

Chapter 7

Conclusion

This concludes our presentation. In this chapter we briefly summarise our findings, and provide direction for future work.

7.1 Wrapping up

We started our exposition by considering the simplest case of networks, namely, static synapses. We began by extending the Eurich model of neuronal avalanches by considering not only fully-connected structures but also complex networks, whose features are a scale-invariant degree distribution and the presence of the small-world property. Here, our hypothesis was that these complex network properties exert an effect on critical behaviour that cannot be predicted by the standard Eurich model. Our desire to extend this model to complex networks is explained by the fact that many real-world networks, and in particular some brain structures, exhibit properties associated to this class of networks (Sect. 2.1.1).

In our simulations, we found that in-degree scale-free networks with high mean clustering coefficient exhibit a larger spike count than any other type of network considered with the same number of vertices and same number of edges at criticality. Thus, a scale-free structure with high degree of small-world-ness is a permutation of edges that maximises the activity in the system. Moreover, we introduced a local measure of performance that we called *node success* which measures the number of out-neighbours of a given node that become active after this node emits a spike. In other words, with the node success metric we assess the capacity of a node to trigger subsequent activations in its out-neighbourhood as a consequence of its own spiking. With this simple metric we were able to assess the "quality" of network structures in terms

of how successful nodes are when arranged in a particular topology. We observed that in-degree scale-free networks are structures that allow nodes to be more successful, whereas fully-connected networks are structures that exhibit the lowest spiking activity and the lowest mean node success. We showed that this is explained by a particular behaviour that we called *spike jamming*, which is a byproduct of the model dynamics and network topology.

We also showed how heterogeneous topologies give rise to a *plateau* in the $\Delta\gamma$ function that estimates the difference between data and the best power-law fit. This plateau arises when the function reaches a range of values in the control parameter for which the deviation is low. This suggests the presence of *Griffith phases*, which were already suggested as a mechanism by which neural systems stay close to the critical regime.

Next, we added spike-timing-dependent plasticity (STDP) mechanisms to our systems at the critical state. We observed that the critical state vanishes as a result of STDP modulation of the synaptic weights. This occurred in all network types and sizes considered. However, in fully-connected networks, STDP mechanisms combined with criticality carved a small-world structure out of the network, in which nodes reach a larger node success; thus, suggesting that critical dynamics combined with STDP mechanisms could account for the emergence of small-world structures from densely connected networks such as the one that some organisms have at the first stages of their lives. In this protocol, we employed simple pair-wise STDP mechanisms. Next, we implemented a version that considers triplets of spikes to achieve synaptic modulation. We discussed how this model fits better experimental data than simple pair-wise rules. We observed that as with pair-wise STDP, triplet-wise STDP affected the critical regime due to synaptic modulation. However, the modulation was not as strong as in the former case due to large inter-avalanche intervals which prevented strong changes in the synaptic weights. For this reason, no strong synaptic pruning was observed in the networks, which resulted in no emergence of small-world structures from fully-connected topologies.

Next, we wondered if there could be a plasticity mechanism that could compensate for the modulation induced by STDP in order to allow the system to stay at the critical regime. For this purpose we devised a plasticity mechanism based on the measure of node success previously introduced. Such a plasticity mechanism potentiate synaptic weights according to the success of a node, and depress them in function of its firing rate. We called this plasticity mechanism *node-success-driven plasticity* (NSDP). We

showed how NSDP allows the system to self-organise into the critical state without fine-tuning of the control parameter. That is, the system exhibits self-organised criticality under this mechanism. We presented some molecular candidates which could potentially account for this phenomenon in brain networks through retro-synaptic signaling. Afterwards, we introduced STDP mechanisms to our systems, which resulted in the two plasticity mechanisms coexisting in the system. Both STDP and NSDP possess different action mechanisms, but both result in synaptic modulation. We observed that the critical regime is preserved by the action of NSDP even when STDP mechanisms are present. This is valid for both pair-wise and triplet-wise versions of this type of plasticity. Therefore, NSDP is compensating for the modulation induced by STDP, which allows the system to stay in the critical regime, while at the same type providing the neural substrate required for learning and memory.

One way to observe the connectedness of our results is by following the next dialogue.

Q: I've read that some brain structures have the properties of scale-free networks (van den Heuvel and Sporns, 2011). Why would the brain prefer to have such a structure rather than a fully-connected architecture?

A: Neural tissue is expensive in terms of metabolic consumption, therefore the brain must do the best with a sparse architecture (Bullmore and Sporns, 2012). Scale-free networks provide robustness and fast communication among different parts of the network due to the presence of hubs (Albert et al., 2000; van den Heuvel et al., 2012). Moreover, criticality is a dynamical regime in which neural computation is optimised (Beggs, 2008). In our work, we showed how scale-free networks maximise neural activity and the success of spikes in terms of their capacity to trigger subsequent spikes. At the same time, we showed how all-to-all connectivity curtails neural activity. Additionally, we showed how scale-free networks might give rise to Griffith phases, thus extending the critical range in parameter space. These observations seem to imply that the brain benefits from scale-free architectures.

Q: It has also been found that some brain networks exhibit the small-world property (Sporns, 2010). How did it get there in the first place?

A: With our work, we suggest that critical dynamics combined with STDP can account for the emergence of such structure. The small-world property also imply benefits for the brain in terms of information transmission (Watts and Strogatz, 1998).

Q: But, if criticality is a requirement for small-world structures to emerge in brain networks, how do this type of dynamics emerge?

A: Criticality can emerge in neural systems as a result of the balance of excitation and inhibition (Beggs and Plenz, 2003), or as a result of the interplay between replenishing and depletion of synaptic resources (Levina et al., 2007). Here we showed how another type of plasticity (NSDP) gives rise to self-organised critical behaviour.
Q: Can STDP and criticality coexist?
A: Yes, as in the model of Levina et al. (2007), and also in our model. This results in a system whose synaptic weight distribution represents the changes induced by learning, but at the same time results in modulation that allows the system to reach a critical state.

Lastly, further summarising, these are the main contributions of our research:

- Criticality occurs in complex networks, and complex topological features like the small-world property and the presence of hubs have an effect on the critical state. In particular, we showed how scale-free networks with absorbing hubs combined with the small-world property is a structure that maximises spiking activity as well as the success of spikes emitted. In contrast, we showed how fully-connected structures are in the other side of the spectre with low spiking activity and low node success due to the *spike jamming* effect.

- We presented a learning rule of pair-wise STDP, which combined with critical dynamics, gives to a small-world structure in fully-connected networks; thus suggesting this combination of processes as a way that neural systems achieve such a complex structure through activity-dependent mechanisms.

- Unlike simple spike pairs, spike triplets of STDP does not have a strong effect in network topology due to large inter-avalanche intervals. To the best of our knowledge, ours is the first study to implement triplet-wise STDP into critical neural systems.

- Self-organised criticality is achieved in a system whose elements modulate their synaptic efficacies in order to maximise the success of their spikes and get penalised if their firing rate is high. These are the mechanisms of node-success-driven plasticity.

- Node-success-driven plasticity regularises the synaptic modulation produced by STDP. This results in long-term plasticity that reflects the modulation induced by learning and memory (STDP), and at the same time provides the synaptic requirements for the network to self-organise towards the critical regime (NSDP).

7.2 Future Work

Naturally, there are questions that remained open in our presentation. We can think of several ways to extend the model we presented in Chapter 6. The first extensions to the model would be the introduction of inhibitory synapses and leakage. This would make our model more biologically plausible.

Perhaps most importantly it is to provide an answer to the question of how plausible is the existence of a plasticity rule whose mechanisms are similar to those of NSDP in brain networks? We believe that one direction of future research points to this issue. Currently, although the existence of retro-synaptic signals is well documented, its relationship to long-term plasticity has not been explored yet. With our work we would like to encourage experimental research towards this direction. We would like to stress the difference between not having yet any experimental evidence of such a mechanism, and having evidence of the non-occurrence of it.

7.2.1 Universality

After having studied the behaviour of critical systems in scale-free topologies, one could wonder how critical behaviour is related to the exponent of the power-law in this type of networks. Scale-free networks are characterised by a degree distribution that follows a power-law of the form $P(k) \sim k^{-\beta}$, where β is the exponent that parametrises the distribution. In our study we worked with $\beta = 1$. On the other hand, the critical regime is featured by the existence of power-law distribution of events; in particular, we have focused our exposition mainly in the distribution of avalanche sizes, which can be approximated by a distribution of the form $P(S) \sim S^{-\gamma}$. Wondering about the relationship between exponents β and γ comes natural. However, in this thesis, we left this question open. Nevertheless, we offer some insights regarding the possible outcome.

Our experience from working with different topologies and sizes makes us believe that no matter how different the structures are, when appropriately tuned (as in Chapter 4) or not tuned (as in Chapter 6) these systems will give rise to avalanche distributions whose exponent γ lies close to -1.5. We confirmed such observations in this thesis not only on static synapses, but also with dynamic synapses in which the control parameter α need not to be fine-tuned. However, how we reach such exponent ($\gamma = -1.5$) will be different in principle for each structure and size. That is, the critical intervals will differ in function of network type and size, but will result in an exponent that is close

to -1.5.

This insight is related to the notion of *universality*, in which systems from different contexts exhibit similarities across their critical exponents (Stanley, 1999). We suspect that the Eurich model and all its variations belong to a broad class of dynamical systems that can be identified by the presence of avalanche size distributions that follow a power-law with exponent $\gamma = -1.5$.

7.2.2 The Renormalisation Group

The Renormalisation Group (RG) is part of the standard toolbox of a statistical physicist. It provides the appropriate framework to study critical phenomena in order to obtain analytically values of observables such as critical exponents (Kadanoff, 2010). Before the RG formalism was developed, the common way to study critical phenomena was through mean-field theory. However, mean-field approximations disregard long-distance correlations, which at the critical point occur at every scale. The mean-field approach is based on averaging quantities close to a phase transition of the system. However, in the vicinity of a phase transition the system is dominated by large fluctuations that invalidate such averaging (Kadanoff, 2010). Nevertheless, mean-field theory yields good approximations regarding the location of critical points in parameter space, and the value of the critical exponents.

RG was put forward as an answer to the question that mean-field approximations have left open. The RG is based on the idea of coarse-graining the collective dynamics of a system in order to get a broader picture of its macroscopic behaviour through disregarding some details in the microscopic level. However, in order to apply the RG, the system must exhibit some scale-invariant features in its dynamics and in its structure. The critical point provides the first of these requirements. At this point the system exhibits fractal geometry identified by the scale-invariance of its dynamics. The RG has been used previously to study the *topological* phase transition from a random network to a small-world structure (Newman and Watts, 1999), and as well to study the self-organised critical (SOC) behaviour of the sandpile (Pietronero et al., 1994; Vespignani et al., 1995) and forest-fire models (Loreto et al., 1995). In all those situations, the underlying topology was a lattice, which is the simplest structure that exhibits self-similarity and scale-invariance. A study of SOC in complex networks under the formalism of the RG is still lacking.

Scale-free networks might offer the appropriate structural properties required to

coarse-grain effectively when applying the RG. Already the power-law degree distributions of these networks suggest the presence of scale-invariance in their structure. Moreover, the fractal geometry of complex networks has already been confirmed for different structures and network statistics (e.g. the presence of high clustering coefficients and low characteristic path lengths) (Song et al., 2005a; Goh et al., 2006; Kim et al., 2006; Daqing et al., 2011).

This implies that scale-free networks are an appropriate network structure in which to apply the RG to study SOC analytically. This provides a direction of future research. Moreover, a mapping between the RG and deep learning methods for machine learning has been found recently (Mehta and Schwab, 2014). This could bridge the concepts of SOC and deep learning in order to discover applications of the theory of critical phenomena to this technique of artificial intelligence; and on the opposite direction, offer some insights regarding learning in brain networks based on critical dynamics. Additionally, in Sect. 6.6 we presented some concepts that at first sight might seem disconnected, namely, deep networks of spiking neurons as an alternative to deep artificial neural networks (ANNs) to achieve complex learning mechanisms in machine learning, node-success-driven plasticity (NSDP) as a candidate mechanism to implement a back-propagation-like learning rule in networks of spiking neurons, SOC behaviour as a result of complex networks implementing NSDP mechanisms, and lastly, the RG as a framework to study SOC behaviour and learning in deep ANNs. Discovering the theoretical and practical mechanisms that bind all these concepts is perhaps the most interesting direction of future work that we might think of.

Lastly, consciousness has been portrayed as the "holy grail" of neuroscience (if not of all natural sciences). The concept lies in the intersection of disciplines of the weight of physics, biology, neuroscience, psychology, artificial intelligence, robotics, and -naturally- philosophy. A relationship among the notions of criticality, phase transitions, the RG, and consciousness has been bravely put forward (Werner, 2013). Although, the links seem rather informal and only at the conceptual level, the suggestion is very appealing. As expected there are more questions than answers in this respect. We would be satisfied if through the work presented in this thesis we provided some minimal insight in this direction.

Finally, for all questions that remained unanswered in this thesis, we can only appeal for the reader's consideration, and remind her that there is so much to do in so little time. *Ars longa, vita brevis.*

Appendix A

How to create a SF network

"The most productive use of power laws in the real world will therefore, we believe, come from recognizing their ubiquity (and perhaps exploiting them to simplify or even motivate subsequent analysis) rather than from imbuing them with a vague and mistakenly mystical sense of universality."

—Stumpf and Porter (2012)

As mentioned in Sect. 2.1.5, a scale-free network is a network whose degree distribution follows a power-law of the form

$$P(k) \sim k^{-\beta} \tag{A.1}$$

where k denotes node degree, and β the exponent of the power-law. In our work we set β to unity. Purported power-laws in empirical data have been claimed to follow a straight line in logarithmic axes in certain interval (x_{min}, x_{max}) of the frequency distribution of the data (Clauset et al., 2009; Stumpf and Porter, 2012).

In this section, we present the algorithm to generate scale-free networks used extensively in our research. The algorithm is based on ideas by Holme and Kim (2002) who extend the Barabasi-Albert algorithm (Barabási and Albert, 1999) by introducing a *triad-formation* step in order to maximise the over-all clustering coefficient in the network.

Unlike the Holme-Kim algorithm and the Barabasi-Albert algorithm, ours does not implement network growth nor preferential attachment. Rather, the user specifies network size N, power-law exponent β of the desired network and number of neighbouring nodes T to consider in the triad-formation step. The algorithm produces an adjacency matrix of a directed network whose *out-degree distribution* follows a power-law of the form of Eq. (A.1).

The algorithm in pseudo-code goes like this:

Inputs: N, β and T. We set $x_{min} = 1$, which means that our desired power-law starts at very beginning of the frequency distribution[1].

Let `currentDeg` $= x_{min}$ and `node_count` $= 0$. While `node_count` $< N$ and `currentDeg` $\leq N$ do:

- Estimate the number of nodes with out-degree `currentDeg` by computing the integer part of `(currentDeg)`$^{\beta}$ and store it on `numNodesCurrDeg`[2]
- Increase `node_count` by `numNodesCurrDeg`.
- While `numNodesCurrDeg` $\neq 0$ do:
  - Choose `nodeOrigin` at random from unused nodes, that is, nodes do not have out-going connections yet.
  - We must connect `nodeOrigin` with other `currentDeg` nodes. Keep track of number of nodes to connect current `node` by doing `nodes2connect`$=$ `currentDeg`.
  - While `nodes2connect` $\neq 0$ do:
    - Choose `nodeDest` at random from $N-1$ nodes[3].
    - Decrease `nodes2connect` by one.
    - *Triad-formation* step: If $T \neq 0$ and `nodes2connect` ≥ 0, choose another T nodes to connect `nodeOrigin` to. However, take these T nodes from the neighbours of `nodeDest`. When done, decrease `nodes2connect` by T[4].
  - Label `nodeOrigin` as used, and decrease `numNodesCurrDeg` by one.
- Increase `currentDeg` by one.

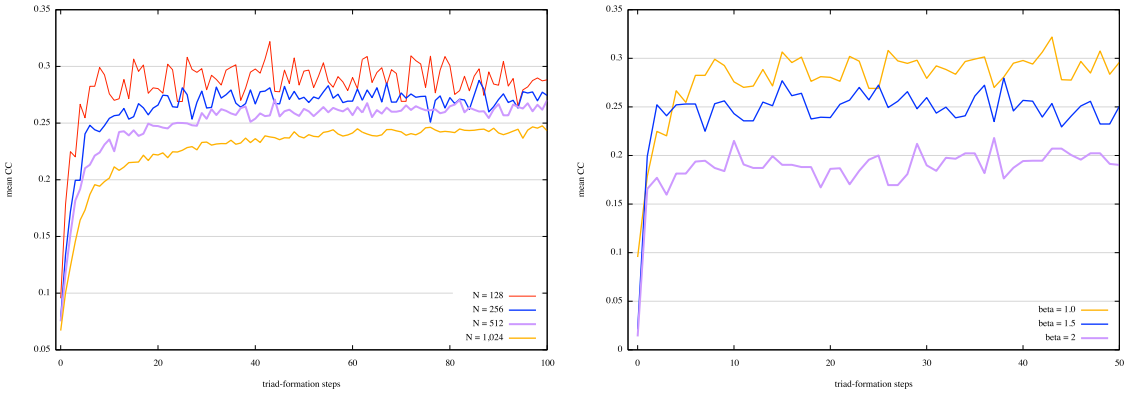
With some more details this algorithm will produce a directed network whose out-degree distribution follows a power-law. The algorithm produces an adjacency matrix

[1] We are not interested in the value of x_{max} as we are going to let the algorithm run until there are no nodes left to connect.

[2] Recall that a power-law degree distribution looks like Fig. 3.5b in page 47, where the smallest possible out-degree is 1 as every node in the network has at least one out-going connection; and the maximim possible out-degree is N. Each entry (x,y) in the histogram says: *there are y nodes of degree x*.

[3] That is, do not allow self-connections. Additionally, we check whether both nodes are already connected, in which case we choose another node

[4] Because in this step the algorithm does not add any extra edge, the resulting network will have exactly the same number of edges as if this step was never performed. This results in networks with more/less clustering with the same number of edges.



(a) Different network sizes for $\beta = 1.0$ (b) Different exponents β for network size $N = 128$

Figure A.1: Mean clustering coefficient per triad-formation step T for scale-free nets created by our algorithm. There is an upper bound for the maximum clustering coefficient in scale-free networks generated through our algorithm. (a) Smaller system sizes possess a larger upper bound than larger systems. (b) As well, smaller exponents in the power-law reach larger values of the mean clustering coefficient.

A. The process of reversing the direction of the edges in the network is equivalent to transposing A; therefore A^t is a directed network whose in-degree distribution follows a power-law. By this process we obtain out-degree and in-degree scale-free networks.

We experimented with the value of T in our algorithm, which specifies the number of triad-formation steps to be performed by the algorithm. We observed an upper bound in the value of the maximum mean clustering coefficient reachable by the network when varying the value of T. This implies that even if we increase the value of T, the value of the mean clustering coefficient will saturate. We show two examples of this behaviour in Fig. A.1.

The value of the upper bound is in function of system size N and exponent β of the power-law degree distribution. We observe that smaller systems sizes reach a larger upper bound (Fig. A.1a). Likewise, smaller exponents reach a larger mean clustering coefficient than larger ones (Fig. A.1b).

This points out a relationship between system size, power-law exponent and maximum clustering coefficient achievable, which we have not explored in our research. Do we observe this upper bound in real-world networks? Not really, as this upper bound might be the result of our requirements on network features. That is, we are asking the algorithm to generate a very particular *directed* scale-free network, one whose size is fixed to N, nodes share at least T neighbours, the power-law out-degree distribution

has an exponent of β, and all nodes have at least one out-going connection. These requirements might put a restriction on the maximum value achievable by the mean clustering coefficient. Real-world networks are free of these requirements. Moreover, they exhibit network growth which might explain the absence of this upper bound.

Lastly, we witness the effectiveness of the triad-formation step even when $T = 1$. The value of the mean clustering coefficient differs considerably from $T = 0$ to $T = 1$, that is, when we ask the algorithm to make a connection to a neighbouring node's neighbour.

The effects of the triad-formation step over the amount of the small-world property in our networks can be appreciated through the metric S given by Eq. (2.4) and whose value is presented in Table 3.1 in page 61 for the networks created using this algorithm. The more clustered the network is, the more amount of small-world-ness it has.

Bibliography

Albert, R. and Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74(1):47.

Albert, R., Jeong, H., and Barabási, A.-L. (2000). Error and attack tolerance of complex networks. *Nature*, 406(6794):378–382.

Amaral, L. A. and Ottino, J. M. (2004). Complex networks. *The European Physical Journal B-Condensed Matter and Complex Systems*, 38(2):147–162.

Bak, P. (1997). *How nature works*. Oxford University Press, Oxford.

Bak, P., Chen, K., and Tang, C. (1990). A forest-fire model and some thoughts on turbulence. *Physics Letters A*, 147(5):297–300.

Bak, P., Tang, C., and Wiesenfeld, K. (1988). Self-organized criticality. *Physical Review A*, 38(1):364.

Barabasi, A.-L. (2002). Linked: How everything is connected to everything else and what it means. *Plume Editors*.

Barabási, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439):509–512.

Barbieri, R. and Shimono, M. (2012). Criticality in large-scale brain fmri dynamics unveiled by a novel point process analysis. *Networking of Psychophysics, Psychology and Neurophysiology*, page 61.

Basalyga, G., Gleiser, P. M., and Wennekers, T. (2011). Emergence of small-world structure in networks of spiking neurons through stdp plasticity. In *From Brains to Systems*, pages 33–39. Springer.

Beggs, J. and Plenz, D. (2003). Neuronal avalanches in neocortical circuits. *the Journal of Neuroscience*, 23(35):11167.

Beggs, J. and Plenz, D. (2004). Neuronal avalanches are diverse and precise activity patterns that are stable for many hours in cortical slice cultures. *The Journal of Neuroscience*, 24(22):5216–5229.

Beggs, J. and Timme, N. (2012). Being critical of criticality in the brain. *Frontiers in Physiology*, 3.

Beggs, J. M. (2007). Neuronal avalanche. *Scholarpedia*, 2(1):1344.

Beggs, J. M. (2008). The criticality hypothesis: how local cortical networks might optimize information processing. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 366(1864):329–343.

Bellay, T., Klaus, A., Seshadri, S., and Plenz, D. (2015). Irregular spiking of pyramidal neurons organizes as scale-invariant neuronal avalanches in the awake state. *eLife*, 4:e07224.

Bertschinger, N. and Natschläger, T. (2004). Real-time computation at the edge of chaos in recurrent neural networks. *Neural Computation*, 16(7):1413–1436.

Bi, G.-q. and Poo, M.-m. (1998). Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. *The Journal of Neuroscience*, 18(24):10464–10472.

Bi, G.-q. and Poo, M.-m. (2001). Synaptic modification by correlated activity: Hebb's postulate revisited. *Annual Review of Neuroscience*, 24(1):139–166.

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. *The Journal of Neuroscience*, 2(1):32–48.

Billings, G. and van Rossum, M. C. (2009). Memory retention and spike-timing-dependent plasticity. *Journal of Neurophysiology*, 101(6):2775–2788.

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., and Hwang, D.-U. (2006). Complex networks: Structure and dynamics. *Physics Reports*, 424(4):175–308.

Bohte, S. M., Kok, J. N., and La Poutre, H. (2002). Error-backpropagation in temporally encoded networks of spiking neurons. *Neurocomputing*, 48(1):17–37.

Bornholdt, S. and Röhl, T. (2003). Self-organized critical neural networks. *Physical Review E*, 67(6):066118.

Bornholdt, S. and Rohlf, T. (2000). Topological evolution of dynamical networks: Global criticality from local dynamics. *Physical Review Letters*, 84(26):6114.

Bullmore, E. and Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, 10(3):186–198.

Bullmore, E. and Sporns, O. (2012). The economy of brain network organization. *Nature Reviews Neuroscience*, 13(5):336–349.

Caravelli, F., Hamma, A., and Di Ventra, M. (2013). Scale-free networks as an epiphenomenon of memory. *arXiv preprint arXiv:1312.2289*.

Chevaleyre, V., Takahashi, K. A., and Castillo, P. E. (2006). Endocannabinoid-mediated synaptic plasticity in the cns. *Annu. Rev. Neurosci.*, 29:37–76.

Chialvo, D. R. (2006). Psychophysics: Are our senses critical? *Nature physics*, 2(5):301–302.

Clauset, A., Shalizi, C. R., and Newman, M. E. (2009). Power-law distributions in empirical data. *SIAM review*, 51(4):661–703.

Clopath, C. and Gerstner, W. (2010). Voltage and spike timing interact in stdp–a unified model. *Frontiers in synaptic neuroscience*, 2.

Colombo, M. and Seriès, P. (2012). Bayes in the brain: On bayesian modelling in neuroscience. *The British journal for the philosophy of science*, page axr043.

Dalva, M. B., McClelland, A. C., and Kayser, M. S. (2007). Cell adhesion molecules: signalling functions at the synapse. *Nature Reviews Neuroscience*, 8(3):206–220.

Daqing, L., Kosmidis, K., Bunde, A., and Havlin, S. (2011). Dimension of spatially embedded networks. *Nature Physics*, 7(6):481–484.

Dayan, P. and Abbott, L. F. (2001). *Theoretical neuroscience*. Cambridge, MA: MIT Press.

de Arcangelis, L. and Herrmann, H. J. (2010). Learning as a phenomenon occurring in a critical state. *Proceedings of the National Academy of Sciences*, 107(9):3977–3981.

de Arcangelis, L. and Herrmann, H. J. (2012). Activity-dependent neuronal model on complex networks. *Frontiers in physiology*, 3.

de Arcangelis, L., Perrone-Capano, C., and Herrmann, H. J. (2006). Self-organized criticality model for brain plasticity. *Physical review letters*, 96(2):028107.

Dorogovtsev, S. N., Goltsev, A. V., and Mendes, J. F. (2008). Critical phenomena in complex networks. *Reviews of Modern Physics*, 80(4):1275.

Eurich, C., Herrmann, J., and Ernst, U. (2002). Finite-size effects of avalanche dynamics. *Physical Review E*, 66(6):066137.

Frette, V., Christensen, K., Malthe-Sørenssen, A., Feder, J., Jøssang, T., and Meakin, P. (1996). Avalanche dynamics in a pile of rice. *Nature*, 379(6560):49–52.

Friedman, N., Ito, S., Brinkman, B. A., Shimono, M., DeVille, R. L., Dahmen, K. A., Beggs, J. M., and Butler, T. C. (2012). Universal critical dynamics in high resolution neuronal avalanche data. *Physical review letters*, 108(20):208102.

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, 11(2):127–138.

Gjorgjieva, J., Clopath, C., Audet, J., and Pfister, J.-P. (2011). A triplet spike-timing–dependent plasticity model generalizes the bienenstock-cooper-munro rule to higher-order spatiotemporal correlations. *Proceedings of the National Academy of Sciences*, 108(48):19383–19388.

Goh, K.-I., Salvi, G., Kahng, B., and Kim, D. (2006). Skeleton and fractal scaling in complex networks. *Physical review letters*, 96(1):018701.

Gross, T. and Blasius, B. (2008). Adaptive coevolutionary networks: a review. *Journal of the Royal Society Interface*, 5(20):259.

Gross, T. and Sayama, H. (2009). *Adaptive Networks: Theory, Models and Applications*. Springer Verlag.

Gutenberg, B. and Richter, C. F. (1956). Magnitude and energy of earthquakes. *Annals of Geophysics*, 9(1):1–15.

Haldeman, C. and Beggs, J. (2005). Critical branching captures activity in living neural networks and maximizes the number of metastable states. *Physical Review Letters*, 94(5):58101.

Harris, K. D. (2008). Stability of the fittest: organizing learning through retroaxonal signals. *Trends in neurosciences*, 31(3):130–136.

Herculano-Houzel, S. (2009). The human brain in numbers: a linearly scaled-up primate brain. *Frontiers in human neuroscience*, 3.

Hinton, G. E., Osindero, S., and Teh, Y.-W. (2006). A fast learning algorithm for deep belief nets. *Neural computation*, 18(7):1527–1554.

Holme, P. and Kim, B. J. (2002). Growing scale-free networks with tunable clustering. *Physical Review E*, 65(2):026107.

Humphries, M. D. and Gurney, K. (2008). Network small-world-ness: a quantitative method for determining canonical network equivalence. *PLoS One*, 3(4):e0002051.

Jensen, H. J. (1998). *Self-organized criticality: emergent complex behavior in physical and biological systems*, volume 10. Cambridge university press.

Johansen, A. and Sornette, D. (1998). Stock market crashes are outliers. *The European Physical Journal B-Condensed Matter and Complex Systems*, 1(2):141–143.

Jost, J. and Kolwankar, K. (2009). Evolution of network structure by temporal learning. *Physica A: Statistical Mechanics and its Applications*, 388(9):1959–1966.

Kadanoff, L. P. (2009). More is the same; phase transitions and mean field theories. *Journal of Statistical Physics*, 137(5-6):777–797.

Kadanoff, L. P. (2010). Theories of matter: infinities and renormalization. *arXiv preprint arXiv:1002.2985*.

Kalinovsky, A. and Scheiffele, P. (2004). Transcriptional control of synaptic differentiation by retrograde signals. *Current opinion in neurobiology*, 14(3):272–279.

Kim, J., Goh, K.-I., Salvi, G., Oh, E., Kahng, B., and Kim, D. (2006). Fractality in complex networks: critical and supercritical skeletons. *arXiv preprint cond-mat/0605324*.

Kinouchi, O. and Copelli, M. (2006). Optimal dynamical range of excitable networks at criticality. *Nature Physics*, 2(5):348–351.

Kishida, K. T. and Klann, E. (2007). Sources and targets of reactive oxygen species in synaptic plasticity and memory. *Antioxidants & redox signaling*, 9(2):233–244.

Klaus, A., Yu, S., and Plenz, D. (2011). Statistical analyses support power law distributions found in neuronal avalanches. *PLoS ONE*.

Knill, D. C. and Pouget, A. (2004). The bayesian brain: the role of uncertainty in neural coding and computation. *TRENDS in Neurosciences*, 27(12):712–719.

Larremore, D. B., Shew, W. L., and Restrepo, J. G. (2011). Predicting criticality and dynamic range in complex networks: effects of topology. *Physical Review Letters*, 106(5):058101.

Le Cun, Y. (1986). Learning process in an asymmetric threshold network. In *Disordered systems and biological organization*, pages 233–240. Springer.

Levina, A. (2008). A mathematical approach to self-organized criticality in neural networks. *Nieders. Staatsu. Universitätsbibliothek Göttingen. Dissertation (Ph. D. thesis), webdoc. sub. gwdg. de/diss/2008/levina/levina. pdf*.

Levina, A., Herrmann, J., and Geisel, T. (2007). Dynamical synapses causing self-organized criticality in neural networks. *Nature Physics*, 3(12):857–860.

Lewis, S. N. and Harris, K. D. (2014). The neural marketplace: I. general formalism and linear theory. *bioRxiv*, page 013185.

Linkenkaer-Hansen, K., Nikouline, V. V., Palva, J. M., and Ilmoniemi, R. J. (2001). Long-range temporal correlations and scaling behavior in human brain oscillations. *The Journal of neuroscience*, 21(4):1370–1377.

Loreto, V., Pietronero, L., Vespignani, A., and Zapperi, S. (1995). Renormalization group approach to the critical behavior of the forest-fire model. *Physical review letters*, 75(3):465.

Magnasco, M. O., Piro, O., and Cecchi, G. A. (2009). Self-tuned critical anti-hebbian networks. *Physical review letters*, 102(25):258102.

Massobrio, P., de Arcangelis, L., Pasquale, V., Jensen, H. J., and Plenz, D. (2015). Criticality as a signature of healthy neural systems. *Frontiers in systems neuroscience*, 9.

Mehta, P. and Schwab, D. J. (2014). An exact mapping between the variational renormalization group and deep learning. *arXiv preprint arXiv:1410.3831*.

Meisel, C. and Gross, T. (2009). Adaptive self-organization in a realistic neural network model. *Physical Review E*, 80(6):061917.

Milo, R., Itzkovitz, S., Kashtan, N., Levitt, R., Shen-Orr, S., Ayzenshtat, I., Sheffer, M., and Alon, U. (2004). Superfamilies of evolved and designed networks. *Science*, 303(5663):1538–1542.

Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., and Alon, U. (2002). Network motifs: simple building blocks of complex networks. *Science*, 298(5594):824–827.

Mitchell, M. (2006). Complex systems: Network thinking. *Artificial Intelligence*, 170(18):1194–1212.

Moretti, P. and Muñoz, M. A. (2013). Griffiths phases and the stretching of criticality in brain networks. *Nature communications*, 4.

Morrison, A., Diesmann, M., and Gerstner, W. (2008). Phenomenological models of synaptic plasticity based on spike timing. *Biological cybernetics*, 98(6):459–478.

Nessler, B., Pfeiffer, M., Buesing, L., and Maass, W. (2013). Bayesian computation emerges in generic cortical microcircuits through spike-timing-dependent plasticity. *PLoS Comput Biol*.

Newman, M. (2010). *Networks: an introduction*. Oxford University Press, Inc.

Newman, M. E. (2003). The structure and function of complex networks. *SIAM review*, 45(2):167–256.

Newman, M. E. and Watts, D. J. (1999). Renormalization group analysis of the small-world network model. *Physics Letters A*, 263(4):341–346.

O'Connor, P., Neil, D., Liu, S.-C., Delbruck, T., and Pfeiffer, M. (2013). Real-time classification and sensor fusion with a spiking deep belief network. *Frontiers in neuroscience*, 7.

Pearlmutter, B. A. and Houghton, C. J. (2009). A new hypothesis for sleep: tuning for criticality. *Neural computation*, 21(6):1622–1641.

Pellegrini, G. L., de Arcangelis, L., Herrmann, H. J., and Perrone-Capano, C. (2007). Activity-dependent neural network model on scale-free networks. *Physical Review E*, 76(1):016107.

Petermann, T., Thiagarajan, T., Lebedev, M., Nicolelis, M., Chialvo, D., and Plenz, D. (2009). Spontaneous cortical activity in awake monkeys composed of neuronal avalanches. *Proceedings of the National Academy of Sciences*, 106(37):15921–15926.

Pfister, J.-P. and Gerstner, W. (2005). Beyond pair-based stdp: A phenomenological rule for spike triplet and frequency effects. In *Advances in neural information processing systems*, pages 1081–1088.

Pfister, J.-P. and Gerstner, W. (2006). Triplets of spikes in a model of spike timing-dependent plasticity. *The Journal of neuroscience*, 26(38):9673–9682.

Pietronero, L., Vespignani, A., and Zapperi, S. (1994). Renormalization scheme for self-organized criticality in sandpile models. *Physical review letters*, 72(11):1690.

Ponulak, F. and Kasinski, A. (2010). Supervised learning in spiking neural networks with resume: sequence learning, classification, and spike shifting. *Neural Computation*, 22(2):467–510.

Priesemann, V., Wibral, M., Valderrama, M., Pröpper, R., Le Van Quyen, M., Geisel, T., Triesch, J., Nikolić, D., and Munk, M. H. (2014). Spike avalanches in vivo suggest a driven, slightly subcritical brain state. *Frontiers in systems neuroscience*, 8.

Purves, D. (1988). *Body and brain: a trophic theory of neural connections*. Harvard University Press.

Rubinov, M. and Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. *Neuroimage*, 52(3):1059–1069.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1988). Learning representations by back-propagating errors. *Cognitive modeling*, 5:3.

Rybarsch, M. and Bornholdt, S. (2014). Avalanches in self-organized critical neural networks: A minimal model for the neural soc universality class. *PloS one*, 9(4):e93090.

Sheik, S., Pfeiffer, M., Stefanini, F., and Indiveri, G. (2013). Spatio-temporal spike pattern classification in neuromorphic systems. In *Biomimetic and Biohybrid Systems*, pages 262–273. Springer.

Shew, W. L., Clawson, W. P., Pobst, J., Karimipanah, Y., Wright, N. C., and Wessel, R. (2015). Adaptation to sensory input tunes visual cortex to criticality. *Nature Physics*.

Shin, C. and Kim, S. (2006). Self-organized criticality and scale-free properties in emergent functional neural networks. *Physical Review E*, 74(4):045101.

Sjostrom, J. and Gerstner, W. (2010). Spike-timing dependent plasticity. *Scholarpedia*, 5(2):1362.

Song, C., Havlin, S., and Makse, H. A. (2005a). Self-similarity of complex networks. *Nature*, 433(7024):392–395.

Song, S., Sjöström, P. J., Reigl, M., Nelson, S., and Chklovskii, D. B. (2005b). Highly nonrandom features of synaptic connectivity in local cortical circuits. *PLoS Biology*, 3(3):e68.

Sornette, D. (2004). *Critical phenomena in natural sciences: chaos, fractals, selforganization and disorder: concepts and tools*. Springer Science & Business Media.

Sporea, I. and Grüning, A. (2013). Supervised learning in multilayer spiking neural networks. *Neural computation*, 25(2):473–509.

Sporns, O. (2010). *Networks of the Brain*. The MIT Press.

Sporns, O., Chialvo, D., Kaiser, M., and Hilgetag, C. (2004). Organization, development and function of complex brain networks. *Trends in Cognitive Sciences*, 8(9):418–425.

Stanley, H. E. (1999). Scaling, universality, and renormalization: Three pillars of modern critical phenomena. *Reviews of modern physics*, 71(2):S358.

Strata, P. and Harvey, R. (1999). Dale's principle. *Brain research bulletin*, 50(5):349–350.

Strogatz, S. H. (2001). Exploring complex networks. *Nature*, 410(6825):268–276.

Stumpf, M. P. and Porter, M. A. (2012). Critical truths about power laws. *Science*, 335(6069):665–666.

Tagliazucchi, E. and Chialvo, D. R. (2011). The collective brain is critical. *arXiv preprint arXiv:1103.2070*.

Touboul, J. and Destexhe, A. (2010). Can power-law scaling and neuronal avalanches arise from stochastic dynamics? *PloS one*, 5(2):e8982.

Uhlig, M., Levina, A., Geisel, T., and Herrmann, J. M. (2013). Critical dynamics in associative memory networks. *Frontiers in Computational Neuroscience*, 7.

van den Heuvel, M., Kahn, R., Goñi, J., and Sporns, O. (2012). High-cost, high-capacity backbone for global brain communication. *Proceedings of the National Academy of Sciences*, 109(28):11372–11377.

van den Heuvel, M. P. and Sporns, O. (2011). Rich-club organization of the human connectome. *The Journal of neuroscience*, 31(44):15775–15786.

Vespignani, A., Zapperi, S., and Pietronero, L. (1995). Renormalization approach to the self-organized critical behavior of sandpile models. *Physical Review E*, 51(3):1711.

Vitali, S., Glattfelder, J. B., and Battiston, S. (2011). The network of global corporate control. *PLoS One*, 6(10):e25995.

Watts, D. and Strogatz, S. (1998). Collective dynamics of small-world networks. *Nature*, 393(6684):440–442.

Werner, G. (2013). Consciousness viewed in the framework of brain phase space dynamics, criticality, and the renormalization group. *Chaos, Solitons & Fractals*, 55:3–12.

Yu, S., Huang, D., Singer, W., and Nikolić, D. (2008). A small world of neuronal synchrony. *Cerebral cortex*, 18(12):2891–2901.

Zweifel, L. S., Kuruvilla, R., and Ginty, D. D. (2005). Functions and mechanisms of retrograde neurotrophin signalling. *Nature Reviews Neuroscience*, 6(8):615–625.