Spatially Structured Information in Attractor Neural Networks using Metric Connectivity

Mario Salvador González Rodríguez

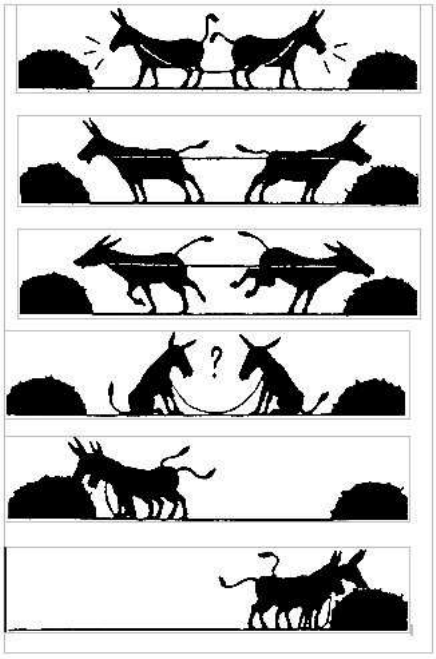
Escuela Politécnica Superior

Universidad Autónoma de Madrid

Doctoral Thesis

Advisor: David Dominguez Carreta

November 28, 2011



Acknowledgements

I am grateful to my advisor David Dominguez Carreta, for kindly sharing his knowledge to me. I would also like to thank to Fracisco B. Rodríguez, Eduardo Serrano, Ángel Sánchez, Rubem Erichsen Jr. and Walter K. Theumann for their collaboration during the research and publication of this work. Special thanks to Alberto Suárez at EPS-UAM for fruitful discussion. Thanks to my lab colleagues and GNB members at EPS-UAM. To the Artificial Intelligence and Computer Science Laboratory (LIACC) members at FEUP, Portugal. To EM ECW Lot 20 at Universidade do Porto for financial support.

Abstract

This thesis explores the emergence of spatially localized information in Attractor Neural Networks with metric connectivity. The spatial information examined is arranged in blocks of opposite correlations, that is, distributed in regions of pattern and anti-pattern. This block information is also characterized in relation with the retrieval of the spatial structure known as bumps, which is a distribution of localized regions that are correlated with the pattern along with uncorrelated regions.

In order to characterize the retrieval of spatially distributed information, the model of Attractor Neural Network on the small-world topology (local to random connectivity) is investigated. Different memory models are used. From a simple model of magnetic interactions to learning a sequence of sparse-coding correlated patterns. The models are rigorously defined and the results obtained by simulation are checked, and contrasted with a theoretical analysis when possible.

The transition between local and global information is evaluated for the network storage capacity and the network topological parameters. It is also explored the competition between the different types of localized information, that is, block and bump structures, according to the sparseness of the neural coding. As an extension of this spatial structures as attractor states of the neural network, and approaching real world applications with spatially organized data, it is characterized the learning and retrieval of patterns structured in regions (*objects*) with different levels of activity. Finally, it is proposed an application for storing and retrieving automotive traffic video that uses a coding structured in variable-activity objets.

Resumen

Esta tesis explora el surgimiento de información espacialmente localizada en Redes Neuronales Atractoras con conectividad métrica. La información espacial estudiada se organiza en bloques con correlación opuesta, es decir, distribuidas en regiones de patrón y anti-patrón. También es caracterizada la información en bloque respecto a la recuperación de otra estructura espacial conocida como "bumps", que es una distribución localizda de regiones que están correlacionadas con el patrón junto con regiones no correlacionadas.

Con el fin de caracterizar la recuperación de la información espacialmente distribuida, se investiga el modelo de Red Neuronal Atractora usando la topología "small-world", variando de la conectividad local a la aleatoria. Además, se utilizan diferentes modelos de memoria, desde un modelo simple de interacciones magnéticas hasta el aprendizaje de una secuencia de patrones correlacionados de baja actividad (sparse-coding patterns). Los modelos son rigurosamente definidos y los resultados obtenidos por simulación son contrastados con un análisis teórico cuando es posible.

Se investiga la recuperación de las estructuras espaciales mencionadas, y la transición entre la información local y global se evalúa en relación a la capacidad de almacenamiento y los parámetros topológicos de la red. También se explora la competencia entre los diferentes tipos de información localizada, es decir, las estructuras de bloques y de "bumps", de acuerdo al nivel de actividad del código neuronal (sparseness). Como una extensión de estas estructuras espaciales como estados atractores de la red neuronal, y acercándose a las aplicaciones del mundo real con información espacial organizada, se caracteriza

el aprendizaje y recuperación de patrones estructurados en regiones (objetos) con diferentes niveles de actividad. Finalmente, se propone una aplicación para el alamcenamiento y recuperación de videos de tráfico automovilístico que usa una codificación con objetos de actividad variable.

Contents

List of Figures

Chapter 1

Introduction

An Attractor Neural Network (ANN) is an arrangement of nodes (neurons), often recurrently connected, whose time dynamics relaxes to a stable pattern. For computational neuroscientists is important to determine which attractors are relevant in order to understand the principles behind the information processing in biological systems. Stable, persistent activity has been thought to be important for neural computation at least since Hebb (1949). In the work of Amit (1989) on attractors neural networks, was suggested that persistent neural activity in biological networks comes as results of dynamical attractors. Different kinds of attractors have been associated with different functions, such as memory (Wills et al., 2005), motor behavior (Stringer et al., 2003), and for recall and recognition (Ruppin and Yeshurun, 1991), to mention a few. Many works deal with different levels of abstraction with varying degrees of biological plausibility, for instances relying largely on spiking models (Izhikevich, 2003; Lago-Fernández et al., 2000). However, there is increasing evidence that many brain areas act as attractor networks as pointed by Wills et al. (2005). The neural assembling is also crucial in biological organisms, and one can argue that the comprehension of complex networks helps to understand the development, organization and emergence abilities of nervous systems (Bullmore and Sporns, 2009). In particular, the study of spatially localized *structures* using attractor networks has been the subject of this thesis. Sustaining or retrieving spatially localized information in the network is relevant for computational neuroscience conjectures as well as for

technical implementations. For example, localized profiles of activity are of interest in computational neuroscience when considering an explanation of working memory capacity (Edin et al., 2009). Also for real-world applications, where patterns are usually represented by spatially organized data, that is, they present an inner structure (González et al., 2011).

1.1 Attractor neural networks as computational memory models

The artificial neural network as a soft computing paradigm, is gaining importance. Soft computing, is an emerging approach to computing, which parallels the remarkable ability of the human brain to reason and learn in an flexible way, and provides answers to the practical situations where heuristic-type algorithms have proved to be valuable tools capable of providing solutions where exact algorithms are not able to. The influence of the neural network learning paradigm on Artificial Intelligence is profound. Computation by artificial neural networks is parallel and can deal with probabilistic or noisy data, is robust and fault tolerant and is capable of flexible learning.

An attractor neural network is a dynamic system. When a pattern is presented as a stimulus to the network, the system will settle to a memory state (attractor), that most closely resembles the input. If this initial state belongs to the attraction basin of an attractor, the network will converge to the corresponding stored pattern. The basin of attraction is the set of initial conditions that will evolve under the neural updating to the imprinted pattern (Eliasmith, 2007). ANN structures have been considered a plausible model for memory processing in the brain, specifically for auto-associative memory (Griniasty et al., 1993; Hopfield, 1982; Uezu et al., 2004).

From an implementation perspective and as pointed by Knoblauch et al. (2010), associative memories are computing architectures in which computation and data storage are not separated, in opposition to the classical von Neumann computing architecture, where there are of two separate modules: a central processing unit (CPU) for computation, and the random-access memory (RAM),

for data storage. Associative memory includes the random-access task as well as content-addressable memory tasks such as, pattern completion and denoising. ANNs have been effectively applied in a wide range of situations, for example, shape recognition (Amit and Mascaro, 1999), visual field data classification (Fink, 2004), automotive traffic video analysis (González et al., 2011).

Just as a remark, ANNs as associative memories, can be viewed for tractability, as an extension of the Ising model (Brush, 1967). In this sense neural networks have received lots of attention, equilibrium properties of fully connected Hopfield neural networks (Hopfield, 1982) have been well studied using spin-glass theory, especially the replica method (Amit et al., 1985a,b), as well as using signal-to-noise analysis for studying their dynamics (Amari and Maginu, 1988; Bollé et al., 2004; Okada, 1995).

1.1.1 Associative memory and Hebbian learning

Associations are stored in a set of synaptic weights between neurons using a local Hebbian learning rule. The basic associative memory problem is to store a set of independent patterns $\{\vec{\xi}^{\mu},\ \mu = 1,\dots,P\}$, and retrieve a pattern (say, $\vec{\xi} \equiv \vec{\xi}^{\mu}$) when the network starts from a neuron state $\vec{\sigma}^{0}$ which is close to it. The simplest way to achieve this in an ANN, is just to make the weights of the network a superposition of terms for each pattern: $W_{ij} = \sum_{\mu}^{P} \xi_i^{\mu}\xi_j^{\mu}$ (Amit, 1989; Hertz et al., 1991).

This is known as the Hebb learning rule. According to Hebb (1949): *The general idea is that any two neurons or systems of neurons that are repeatedly active at the same time will tend to become 'associated', so that activity in one facilitates activity in the other.* An associative memory model using the Hebb rule for all possible pairs ij, with binary units, is usually called a Hopfield model (Hertz et al., 1991). Associative networks are closely related to Hebbian cell assemblies and play an important role in neuroscience as models of neural computation for various brain structures, for example, neocortex, hippocampus, cerebellum, and mushroom body (Rolls and Treves, 1998).

Here one can remark, that the state of a Hebbian network evolves in two times scales. Slow time scale for the connection weights update and fast time scale for

the nodes activities update.

1.2 Global information in uniform networks

Associative networks often deal with global overlapping (similarity) between patterns and neural states, employing uniform, either full or random diluted connectivity to perform retrieval tasks. In fully connected and random diluted networks the connections between nodes are uniformly distributed with independence of the distance between nodes. The original Hopfield model (Hopfield, 1982), is a fully connected single layer associative network, with each neuron connected to every other neuron. Random dilution in the network connectivity implies that an edge or link can be present with probability p or not present with probability $1 - p$. In a random graph edges are chosen uniformly at random between each pair of nodes. A model for generating random graphs can be found in the classic article of Erdös and Rényi (1959). The uniform distributed connectivity allows the network to retrieve global information in an optimal and robust way when the patterns are uniformly distributed because pattern learning is resistant to damage of parts of the network (Amit, 1989; Hertz et al., 1991).

1.2.1 Fully connected neural networks

The fully connected ANN is deeply understood due to its basis in the Ising model of a ferromagnet. The network requires complete and symmetric connectivity, also no self-interactions (Amit, 1989; Hopfield, 1982). Given the fully connectivity, a spatial relationship between the units in the network can not be characterized. However, when ANNs are restricted only to near neighbors of neurons, the range over which the fluctuations in the states of one set of neurons are correlated with or affected by those of a set of neurons in another region is too small, and the system dynamics may be trapped in states that are not memories (spurious attractor) (Coolen and Sherrington, 1993). While full connectivity is not very realistic biologically, it simplifies the ANN model given the symmetry nature of the connections (Amit, 1989). It is also known to be the most efficient arrangement for (global) information storage and retrieval (Amit, 1989; Hertz et al., 1991). In

order to simulate a more *biological genuine* model rather than the fully connected networks, various random diluted models have been studied.

1.2.2 Random-diluted neural networks

In biological networks the synapses are known to be asymmetric and a neuron, in average, is connected only to a fraction of all neurons. Given the huge number of neurons, there is only a small number of interconnections in the human brain cortex ($\sim 10^{11}$ neurons and $\sim 10^{14}$ synapses). A considerable effort in the modeling of neural networks has been devoted to improving the storage and retrieval properties of the networks and to making them biologically plausible. Various random diluted models have been studied, weak diluted model (Zhang and Chen, 2007; Zillmer et al., 2006), finite connection model (Castillo and Skantzos, 2004; Wemmenhove and Coolen, 2003), and extremely diluted model (Derrida et al., 1987; Tamarit et al., 1991). It has also been shown that reliable simulations can be performed in strongly diluted systems that are comparable to analytical results, without using exponentially large networks (Arenzon and Lemke, 1994).

However, connectivity in neural systems is suggested to be far more complex than a random diluted or finite connected network (Ahn et al., 2008; Cherniak, 1994). Many numerical studies have focused on the performance of the Hopfield model using networks with spatially organized connectivity (Davey et al., 2004; Li and Chen, 2003; McGraw and Menzinger, 2003; Stauffer et al., 2003; Torres et al., 2004).

1.3 Structured information in metric networks

Metric is a function which defines a distance between elements in a set. In a uniform random network the connectivity does not depend on the distance between the nodes. Since any node can be connected by an edge to any other node in the network with equal probability, the connections have no spatial organization. On the contrary, in a regular network, for instance a ring, a k-nearest neighbor or a d-dimensional lattice, the connections are metrically organized. Furthermore, when a random network is embedded into a regular lattice *substrate* a spatial structured

can still be observed, which is called metric topology. Metric organized connectivity is relevant since this type of arrangement is more plausible from a biological point of view (Bullmore and Sporns, 2009), as well as, for optimizing technical applications. During the last few years it has become increasingly clear that understanding the behavior of many different systems passes through the comprehension of the dynamics of complex networks (Albert and Barabasi, 2002). Recent works have employed metric topologies such as small-world (Dominguez et al., 2007; Li and Chen, 2003; McGraw and Menzinger, 2003), scale-free (Stauffer et al., 2003; Torres et al., 2004), and modular networks (Johansson and Lansner, 2007; Levy et al., 1999; Roudi and Treves, 2008).

1.3.1 Small-world topology

Biologically plausible associative memories must have diluted and *spatially organized* connectivity, reflecting the situation in the cortex and hippocampus (Rolls and Treves, 1998). Small-world networks (Watts and Strogatz, 1998) provide a practical approach to this issue featuring *dense internal* connections and sparse *inter-modular* connections. As a result small-world networks have some areas (*sub-networks*) with highly connected nodes, with a few shortcuts between the different areas. This mimics a lot of networks found in biology including the brain, in which the majority of connections appear to occur between nearby regions and the pathways between different areas are to some degree diffuse (Damasio, 1994; Levy et al., 1999).

As described by Watts and Strogatz (1998) the small-world network starts in a regular lattice where each node is connected to its $K/2$ nearest neighbors on either side. A fraction ω of these connections are then re-wired to other randomly selected nodes. Self connections and repeated connections are not allowed. The result is a network that interpolates between a regular lattice and a completely random graph. A schematic representation of a small-world topology is presented in Fig. 1.1. For surprisingly low ω the network is in the *small-world regime* with primarily local connectivity and a few random long-range connections (shortcuts). The network remains highly clustered like a regular lattice, with a small characteristic path length, like a random graph. For any connected graph the

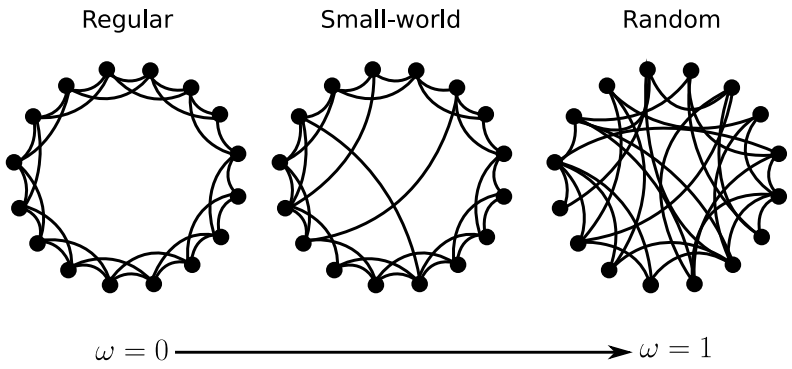


Figure 1.1: Schematic representation of a small-world topology (Watts-Strogatz model).

"characteristic path length" is the average distance between pairs of nodes. It is needed to remark that usually, the *small-world regime* is defined for the strong diluted model with static connections and nodes and mono valued as presented in Newman and Watts (1999). Defining the *small-world regime* for the dynamic nodes with strong feedback, highly connected nodes and variable weights in the connections, as used in the present work, are difficult to calculate and are out of the scope of this study. In this thesis the reference to *small-world* appeals to the topology employed to organize the network connectivity.

Attractor Neural Networks with metric connectivity have been studied recently, specially small-world topologies (Dominguez et al., 2007; Li and Chen, 2003; McGraw and Menzinger, 2003). Scale-free topologies have also gained increasing interest (McGraw and Menzinger, 2003; Stauffer et al., 2003; Torres et al., 2004). The retrieval ability of such networks is commonly measured by the overlap (m) between neuron states and memorized patterns and the load parameter α, which is expressed as the ratio between stored patterns and connections per node. Above a critical value of the load parameter α, no retrieval is possible and the overlap goes to zero (Hertz et al., 1991).

In addition, the mutual information between the stored patterns and the neural states has been proposed to compare the performance of different topologies in terms of memory retrieval (Dominguez and Bollé, 1998; Dominguez et al., 2007; Okada, 1996). Fig. 1.2 shows a comparison between the global overlap m and the mutual information i_m to measure the retrieval abilities for different levels of the

network randomness connectivity. Although the critical load increases monotonically with dilution and randomness, the information is a non-monotonic function of the load and reaches a maximum that corresponds to a nontrivial optimal topology (Dominguez et al., 2007). See Chapter 3 for a detailed examination of mutual information as a performance measure of the network retrieval abilities.

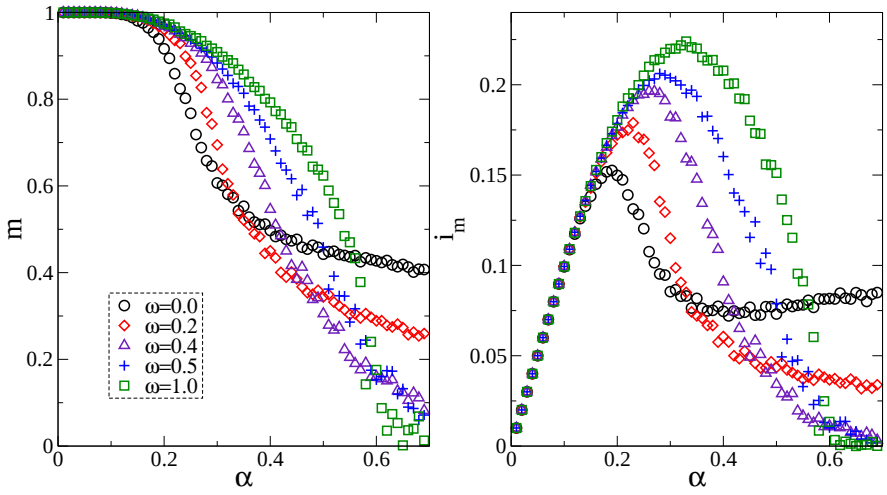


Figure 1.2: Overlap and mutual information as a measure of the network global retrieval performance for different values of the randomness parameter ω. Left: Overlap $\mathrm{m}(\alpha)$. Right: Information ratio $i(\alpha)$. Network with $N = 10^5$ neurons, and $K = 10^2$ neighbors.

Small-world networks with a moderate number of shortcuts can be almost as computationally efficient as a random network, retrieving global information while saving considerably on wiring costs (Bohland and Minai, 2001; McGraw and Menzinger, 2003; Morelli et al., 2004). This wiring cost constraint is an important factor in technical implementation, and is also a factor that affects the development and organization of biological neural networks such as of *C. elegans* (Ahn et al., 2008; Cherniak, 1994). This cost is also an important issue in the human brain, where the connectivity in the cortex and other brain regions is mainly local, with relatively sparse long-distances projections (Rolls and Treves, 1998).

Most works into associative memory networks with spatial structure (Hatchett et al., 2005a; Masuda and Aihara, 2004; McGraw and Menzinger, 2003; Morelli

et al., 2004; Nikoletopoulos et al., 2004) focus on the global retrieval of a pattern, without considering the possibility of spatially localized states. In general, these studies deal with the relation between the storage capacity and the degree of randomness in the network, considering only distributed information (overlap) along the network. However, for metric connectivity one can measure spatially structured information, distributed in local overlaps inside regions of contiguous neurons.

1.3.2 Bumps

The existence of spatially localized activity patterns, or bumps, has been investigated in a variety of neural network models with spatially distributed topology (Roudi and Treves, 2004, 2006; Rubin and Bose, 2004). In particular, sustained patterns of activity that are localized in space, have been recorded in several experimental settings. These patterns, often referred to as bumps of activity, have been correlated with neural circuits that encode directional or spatial information, e.g. working memory tasks (Brunel, 2003; Edin et al., 2009; Wang, 2001).

Bumps have been also measured as a localized retrieval state which has a region with high correlation (overlap) with the same region in the original stored pattern, and a region not correlated at all. In Koroutchev and Koroutcheva (2006) and Dominguez et al. (2012) have been shown that the sparsity of the code enhances the bump effects in memory networks.

Sparse coding is the representation of items by the activation of a relatively small set of neurons. Sparse-coding gives the model a biological plausibility since the brain suggests a general sparse-coding strategy, Experimental evidence for sparse coding has been found in several different sensory modalities in a variety of animals (Olshausen and Field, 2004). Sparse coding also improves the storage capacity of associative memory models in terms of the number of patterns that can be learned (Amit et al., 1987; Okada, 1996; Stroffek et al., 2007).

However, if information is to arise in the network using sparse coding representation, spatial information included; a reinforcement mechanism is required to sustain the levels of activity and to keep it localized (Dominguez et al., 2012;

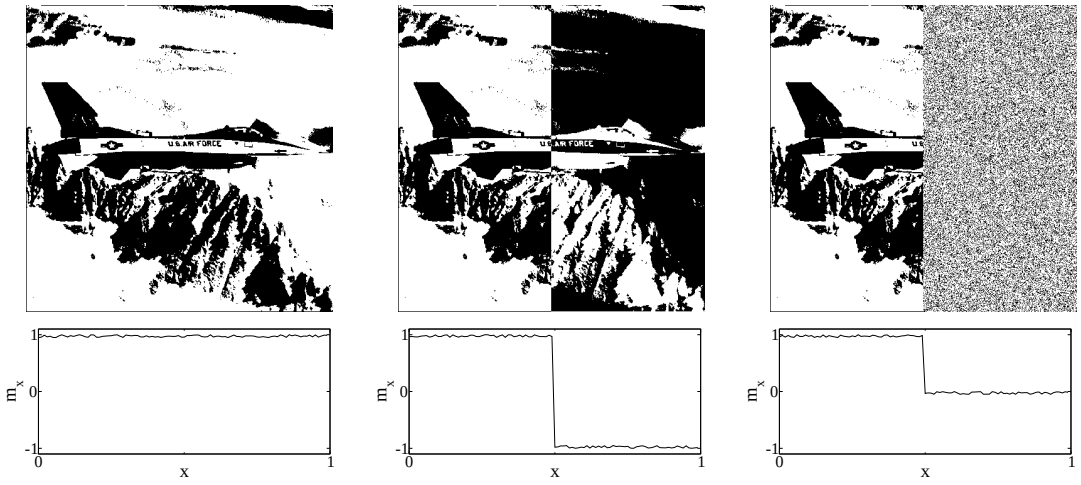


Figure 1.3: Top panels: Global original pattern (left), 2-blocks structure in a pattern/antipattern arrangement (middle), and bump structure arrangement (right). Bottom panels: Positive global overlap configuration $m_x \sim 1$ (left), Positive/negative overlaps configuration with $m_x \sim \pm 1$ respectively (middle), Bump configuration of overlaps with $m_x \sim 1$ for the positive correlated region and $m_x \sim 0$ for the uncorrelated region (right), $x \equiv i/N$.

Koroutchev and Koroutcheva, 2006). Some works use a rigid constraint so that bumps of activity can be sustained by the network (Roudi and Treves, 2008).

In Fig. 1.3 is presented a schematic representation of global (left panels) and structured information (middle and right panels). In the left panels a global pattern is depicted with a global positive overlap configuration (bottom-left). In the top-middle panel a 2-blocks structured pattern is presented, that is a configuration of pattern/antipattern with regions of positive/negative overlap respectively (bottom-middle panel). Finally, in the top-right panel the same pattern is presented in a bump configuration, with a region of positive overlap and an uncorrelated region with null overlap (bottom-right panel).

1.3.3 Blocks

Another type of spatial information is a block structure of opposite overlaps in localized regions of pattern/anti-pattern, as the one studied in Dominguez et al. (2009). Block regions of opposite activity have also also been reported

in González et al. (2009). Suppose a case where sequential blocks of pixels of a binary image are flipped, the overall pattern is probably still recognizable. For instance, Fig. 1.3 shows a binary image in black-and-white pixels in the top-left panel, which is considered as a global pattern. In the top-middle panel the same pattern is reassembled in two blocks of information. The configuration of such block states has vanishing global overlap with the original pattern. However, structured information can be measured, given that local regions carry some spatial order in blocks oscillating between negative and positive overlaps (bottom-middle panel). This is called block retrieval, to distinguish it from the usual global retrieval of the full pattern.

This negative overlap distribution could be thought as the Inverted Tuning Curves of activity, recorded in some cells in the Prefrontal Cortex of monkeys performing spatial working memory tasks (Wang et al., 2004). These neurons have a lower firing rate after some cue is presented than during spontaneous activity. This local order can emerge if the stimulus has some neighborhood structure and the network topology preserves some complexity, with stronger connectivity between nearest neurons than between neurons far from each other (Dominguez et al., 2012, 2009; González et al., 2009). This characteristic may allow the network to retrieve global information where local stimulus is relevant or only block information is at disposal for the network. It is also of interest to have networks capable to extract local information from structured stimuli allowing *block retrieval* of information.

1.4 Pattern correlation

Modeling neural systems with ANNs implies, from a physiological point of view, using metric topologies and sparse coding representation. The correlation between the input patterns has also to be considered. For example, the visual patterns on the retina are correlated across space and time (Liu et al., 2010). Also, in real-world applications, patterns usually present high correlation.

Correlations between the input patterns worsen the performance of the network since the storage of strongly correlated inputs may lead to ambiguity in memory retrieval due to interference. It can be argued that the brain has its

own mechanisms for the orthogonalization of the sensory inputs (Vinje and Gallant, 2000), thus using uncorrelated patterns is allowed for tractability reasons. However, for technical implementation purpose alternative solutions are needed. A well known approach is the pseudo-inverse learning rule to yield orthogonal patterns (Hertz et al., 1991) as well as expansion coding strategies (Trappenberg, 2002).

1.5 Contributions of this thesis

This thesis investigates the effects of the small-world topology upon the retrieval of spatially structured information. In Roudi and Treves (2008) is discussed the possibility of retrieving multiple patterns, in the form of multiple bumps of activity in different parts of the network. On the contrary this thesis considers the possibility of the coexistence of multiple local retrieval states of a unique pattern distributed in regions of positive/negative overlap along the network. While the local field arising from different patterns makes possible to retrieve all of them, the local field induced by a positive/negative pattern configuration cancels both out and the blocks are only stable for networks with local connectivity (Dominguez et al., 2009).

The existence of locally organized memories in blocks of positive/negative overlaps rises some questions that have been addressed in this work. Can stable block states emerge spontaneously in an associative network? How can be measured the information hidden in these blocks? Which neural architectures are able to convert this local information into global information? How is the block structure affected by the sparse representation of the coding? The proposal of a block-like structure could be closely related to biological brain systems, where different patterns of activity (blocks) arising from several cortical structures may be independently retrieved (Rolls and Treves, 1998). Blocks may also represent incomplete pieces of information which can be used to codify any kind of signals (images, voice, fingerprints, genetic code, etc). For network topologies with a majority of short range connections, the storage capacity of global information is severely disrupted (McGraw and Menzinger, 2003). However, local information emerges in the form of blocks of opposite activity/overlap. The local information

corresponds to configuration states which carry information in blocks of neighboring neurons, and are attractors of the network dynamics.

In Chapter 2, a modified Sherrington-Kirkpatrick model with a ferromagnetic bias plus random interactions is presented using the small-world topology. Blocks of opposite activity can be sustained in the network and the stability of the blocks depends on the network parameters of dilution and the number of random shortcuts. The block state is stable with a large basin of attraction. This stability depends on both the number of blocks and the network dilution (Dominguez et al., 2009; González et al., 2009).

In Chapter 3, a classic attractor network with Hebbian learning is considered. Nonuniform metric networks can retrieve fragments of patterns, in blocks of opposite overlaps without performing global retrieval, when a stimulus in such configuration is presented. A new measure related to the fluctuation of the local overlaps is introduced in order to characterize the block structured information, which proved to depend on the metric parameters of the network dilution and randomness connectivity.

In Chapter 4, an additional parameter is introduced and the structured information is characterized taking into account the sparseness of the coding. There is a competition of both types of structured information studied, Bumps and Blocks, according to the sparseness parameter. Whenever the connections are mainly local and the code is sparse (low level of neural activation), retrieving different blocks of pattern/anti-pattern implies sustaining different regions of activity along the network. Thus, a threshold strategy is introduced so that the network dynamics stabilizes a spatial configuration of blocks with opposite overlaps and different levels of activity. This threshold strategy depends mainly on the level of activity of the patterns and the global activity of the network, as well as the local activity of the neighborhood of each neuron. This translates into a self-control mechanism of the network dynamics (Dominguez and Bollé, 1998) according to the initial spatial stimulus (Dominguez et al., 2012). Sparse code benefits the bump effect and a competition arises with the block structure when the network level of activity increases. Both types of structured information are satisfied by the clustering behavior of the small-world topology, allied with the attractor properties of the network threshold dynamics.

In Chapters 2 to 4 the output of the network has been characterized according to the initial conditions of blocks and bumps spatial structures. Chapter 5, on the contrary, focuses on the input learned by the network, which is a pattern structured in localized *objects* with low and high activity. The learning and retrieval properties of the network with metric connectivity is characterized for this type of coding. Given the nature of the neural coding, an adequate threshold is needed in order to control the distinct activity of different regions in the pattern (Dominguez et al., 2012). One also has to deal with the pattern correlation due to the overlap between objects with same levels of activity. The results presented in this Chapter are preliminary and need to be contrasted with more extensive simulations and a theoretical analysis. However, it is worth studying this type of coding which is relevant for real world implementations, where information is represented by spatially organized data, distributed in regions with distinct levels of nodes activation as, illustrated in Chapter 6. Also, for theoretical neuroscience conjectures of information processing.

In Chapter 6, a real life application is addressed using patterns with some type of spatial structure: an automotive traffic video. A sparse-coding ANN with a small-world topology is used to learn/retrieve the video sequence. Given the nature of the data, the patterns are highly correlated and cyclic, a variant of the pseudo-inverse learning rule with a row-shifting schema is introduced to approach this issue (González et al., 2011). The proposed model avoids the segmentation and tracking of the involved targets and also some closely related difficulties. For instances from realistic traffic situations, e.g. public web-cams, where is not possible to control the variability of weather and/or illumination conditions, the feasibility of this technique for the storage and retrieval of traffic videos has proved to be valuable.

Finally in Chapter 7.2, the concluding remarks from this thesis are presented, as well as, the open issues and future research that can be done as result of this work.

1.5.1 Thesis articles

This thesis is based on the following articles:

- Dominguez, D., González, M., Rodríguez, F. B., Serrano, E., Jr., R. E., and Theumann, W. (2012). Structured information in sparse-code metric neural networks. *Physica A: Statistical Mechanics and its Applications*, 391(3):799 – 808.

- González, M., Dominguez, D., and Ángel Sánchez (2011). Learning sequences of sparse correlated patterns using small-world attractor neural networks: An application to traffic videos. *Neurocomputing*, 74(14-15):2361 – 2367.

- González, M., Dominguez, D., and Rodríguez, F. B. (2009). Block attractor in spatially organized neural networks. *Neurocomputing*, 72:3795–3801.

- Dominguez, D., González, M., Serrano, E., and Rodríguez, F. B. (2009). Structured information in small-world neural networks. *Phys. Rev. E*, 79(2):021909.

- González, M., Dominguez, D., and Rodríguez, F. B. (2008a). Block activity in metric neural networks. *WASET Proceedings*, 27:56–59.

- González, M., Dominguez, D., and Rodríguez, F. B. (2008b). Learning block memories with metric networks. *WASET Proceedings*, 27:60–63.

Chapter 2

Block activity in a magnetic interaction model

2.1 Introduction

Attractor Neural Networks (ANNs) as models for associative memory usually use broad connectivity, either completely connected networks or diluted networks with random connectivity (Amit, 1989; Evans, 1989; Hertz et al., 1991; Treves, 1990). However from a neurobiological and an implementation perspective, it is logical to minimize the length of inter-node connections and consider networks whose connectivity is predominantly local. Biologically plausible associative memories must have sparse connectivity, reflecting the situation in the cortex and hippocampus (Johansson and Lansner, 2007; Levy et al., 1999; Rolls and Treves, 1998). ANNs with metric connectivity have been recently studied (Bohland and Minai, 2001; Dominguez et al., 2007; Johansson et al., 2006; Li and Chen, 2003; Lu et al., 2006; McGraw and Menzinger, 2003; Morelli et al., 2004; Torres et al., 2004), specially on the small-world topology (Watts and Strogatz, 1998). The underlying spatial organization of the network connectivity allows to retrieve interesting combinations of localized patterns of activity, and this raises the issue of their competitive interactions (Roudi and Treves, 2004).

ANNs usually deal with global correlation (overlap) between patterns and neural states, the response of a network to a given input stimulus leads to a particular configuration of the neural activity. However, a local order can emerge

if the stimulus has some neighborhood structure and the network topology preserves some complexity, with stronger connectivity between nearest neurons than between neurons far from each other (Koroutchev and Koroutcheva, 2006) allowing the *block retrieval* of information. For instance the image in Fig. 2.1-left, may be regarded as a global pattern, while the one in Fig. 2.1-center and right is the same pattern reassembled in two and four blocks of information respectively. These block states have null correlation with the original pattern, however they still carry information about the mandrill image (USC-SIPI, 2011).

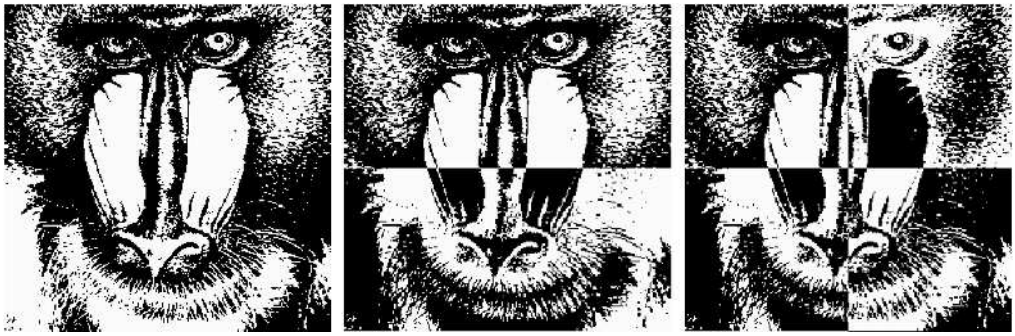


Figure 2.1: Left: the mandrill (original pattern). Center: the mandrill 2-blocks. Right: the mandrill 4-blocks.

For a spatially structured topology, local information inside blocks can be measured. A structured distribution of localized states, can carry some spatially ordered information, with blocks oscillating between negative and positive overlaps, corresponding to the pattern and anti-pattern, respectively. This is called block retrieval (B), to distinguish from the usual global retrieval (R).

In the present Chapter 2, a model of sparsely connected attractor neural network (Amit, 1989) on the small-world topology is presented. In this Chapter, a simplified network model, with random weights competing with a signal term, is used. First, an ordered macroscopic neural state is induced by a bias in the network random weight connections, and the network evolution into a global or block activity ordering is studied according to the initial conditions. The block-like attractor is explored to determine in which conditions (parameters), the network dynamics allows such a configuration. The dependence between the block structure and the network dilution, and the dynamical behavior of the

block and global attractors are characterized with a flow diagram. The transition between phases B and R is also evaluated.

2.2 The model

2.2.1 Topology and dynamics

A neuron i can be in one of two states, firing/quiescent, described by binary variables $\sigma_i \in \{\pm 1\}$. The state of a neuron σ_i is updated in time t according to the sigmoidal activation function:

$$\sigma_i^t = \text{sign}\left(h_i^{t-1} - \theta\right), \quad h_i^t \equiv \sum_j J_{ij}\sigma_j^t, \tag{2.1}$$

where h_i^t is the postsynaptic field arriving at neuron σ_i. Here sign is defined as: sign(z) $= 1$ if z ≥ 0, and sign(z) $= -1$ if z < 0. The variable θ is the firing threshold which is considered to be zero. A synchronous update is used.

The synaptic couplings between neurons i, j are $J_{ij} \equiv C_{ij}W_{ij}$, where $\mathbf{C} = \{C_{ij}\}$ is the topology matrix and $\mathbf{W} = \{W_{ij}\}$ are the synaptic weights. The topology matrix, usually referred to as adjacency matrix in the complex network literature, with $C_{ij} \in \{0, 1\}$ splits in local and random links. The local links connect each neuron to its K_L nearest neighbors, in a closed ring. The random links connect each neuron to K_R others uniformly distributed along the network (Hatchett et al., 2005b). Hence, the network degree is $K = K_L + K_R$. The network topology is then characterized by two parameters: the *connectivity* ratio, and the *randomness* ratio, defined respectively by:

$$\gamma = K/N, \ \omega = K_R/K. \tag{2.2}$$

Here ω plays the role of the rewiring probability in the slightly different original Watts-Strogatz *small-world* model (Watts and Strogatz, 1998). The connectivity and the weights are considered to be asymmetrical, $C_{ij} \neq C_{ji}$, and $W_{ij} \neq W_{ji}$.

2.2.2 Block activity

The weights W_{ij}, of the connections between neurons i and j, are composed of two terms (González et al., 2008a):

$$W_{ij} = cW_{ij}^r + (1-c)\overline{W}, \tag{2.3}$$

where W_{ij}^r are generated randomly to be either $+1$ or -1 with equal probability, representing either an excitatory or an inhibitory synapse, respectively. This term is multiplied by a parameter $c \in (0,1)$, and the bias term $\overline{W}$ is multiplied by $(1-c)$. $\overline{W} = 1$ is used for all synapses. According to the strength of c, the network is induced into a ferromagnetic state (Amit, 1989) (driven by the bias interaction), competing with a disordered state (driven by the random term).

The evolution of the network when initialized in b blocks (sets of neurons of type l_+ and l_-) is studied . The blocks are defined as the groups of neighbor neurons initialized as $\sigma_i = +1, i \in l_+$, and $\sigma_i = -1, i \in l_-$. A mesoscopic variable $a_l(t)$ is used to describe the neural activity of block l, with size $L = N/b$, as the fraction of neurons firing at time t,

$$a_l^t = \frac{1}{L}\sum_{i \in l} \sigma_i^t. \tag{2.4}$$

The macroscopic parameters are:

$$a = \langle a_l \rangle_b, \;\; v = \langle a_l^2 \rangle_b - a^2. \tag{2.5}$$

where a is the usual global activity (Amit, 1989; Paula et al., 2006), and $d \equiv \sqrt{v}$ stands for the block activity. The time index t is dropped for simplicity. One defines the activity of the positive and negative blocks (of type $l_\pm$) as a_+ and a_- respectively. For the case of uniform blocks, it can also be defined the global activity of the network as $a = (a_+ + a_-)/2$ and the block activity as $d = (a_+ - a_-)/2$.

The network can be in the following representative phases: global activity (R, with $a \neq 0, \; d = 0$), block activity (B, with $a = 0, \; d \neq 0$) and zero activity (Z, with $a = 0, \; d = 0$).

2.3 Simulation results

2.3.1 Block evolution

The dynamics in Eq. (2.1) was simulated for a network with $N = 10^5$, and connectivity $\gamma = 10^{-3}$. The network started in $b = 10$ contiguous blocks of positive/negative activity with $a_l^{t=0} \sim \pm 0.2$. For a fixed value of $c = 0.8$, the evolution in time of the block activities for different values of the randomness ratio is depicted in Fig. 2.2. In each panel the activities are smooth averaged over windows of $N_w = 2 \times 10^2$, and $x = i/N$ is defined.

In Fig. 2.2, left panels, the probability of random connections is $\omega = 0.1$. After $t = 10$ time steps, the blocks have been almost completed, and reach a stationary state in which the block structure is maintained, where $a_+ \sim 1$ and $a_- \sim -1$. In the right panels of Fig. 2.2 with $\omega = 0.3$, the block structure ($a_l^{t=10} \sim \pm 0.85$) is reached after $t \sim 10$ time steps. At $t = 60$ the active blocks were approximately filled with $a_+ \sim 1$, while the inactive blocks were destructed $a_- \sim 0$. In the next steps, the inactive blocks become attracted by the active ones, and the global phase is achieved, where an active ordering is accomplished.

In Fig. 2.3, both global and block activity order parameters are plotted against time evolution for the same values of the variables in Fig. 2.2. Fig. 2.3-left, shows that for a value of $\omega = 0.1$, a stable block activity is reached, $a \sim 0$, $d \sim 0.93$. It is seen that up to 10^6 time steps, the blocks don't change into global ordering. The behavior of the activity during the network evolution corresponds to the left panels in Fig. 2.2. In Fig. 2.3-right is seen that, for a larger value of the randomness parameter $\omega = 0.3$, after an initial retrieval of the full block ordering, the network almost suddenly (in a logarithmic time scale) switches from the block (B) phase to the global (R) phase. An active order is settled, $a \sim 1$, $d \sim 0$. This behavior corresponds to the right panels in Fig. 2.2.

2.3.2 Multi-blocks and dilution

Fig. 2.4 depicts the evolution of a network with $N = 10^4$, starting in $b = 10$ blocks chosen in a symmetric successive structure of large and small blocks, with $a_l^{t=0} \sim \pm 0.3$. The activities are smooth averaged over windows of $N_w = 50$. The

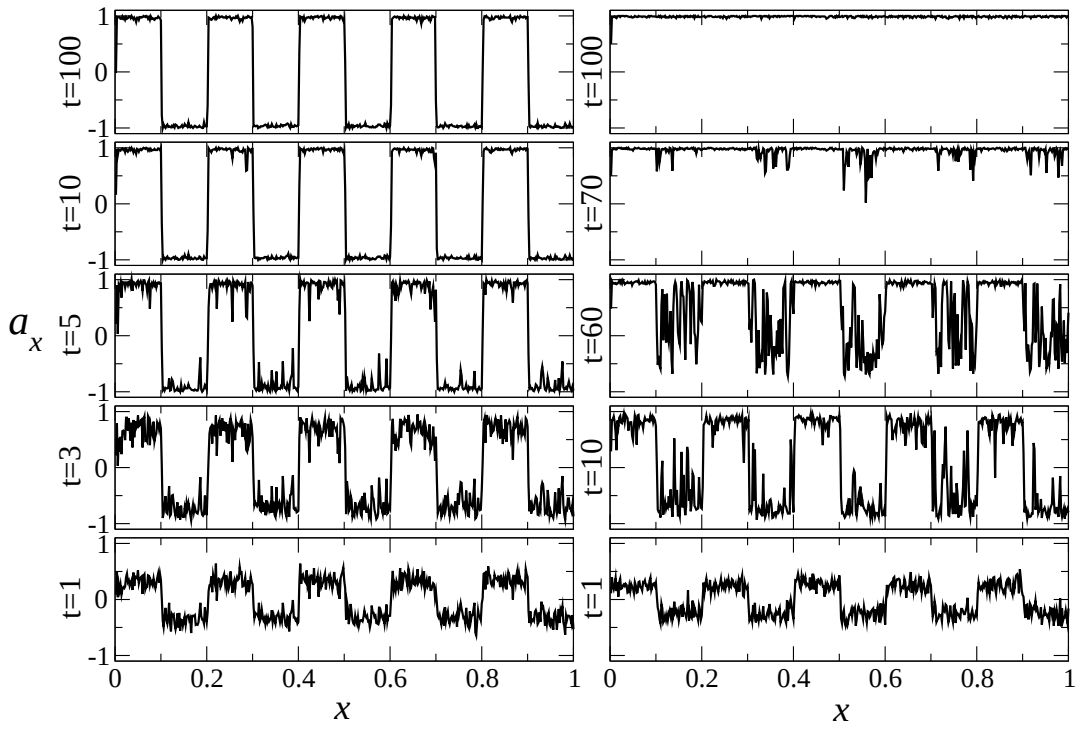


Figure 2.2: Network with N= 10^5, $\gamma = 10^{-3}$ and $c = 0.8$. Initial condition: $a_l^{t=0} \sim \pm 0.2$. Left panels: Evolution of the network into a block activity ordering, $\omega = 0.1$. Right panels: Evolution of the network into a global activity ordering, $\omega = 0.3$. $x \equiv i/N$.

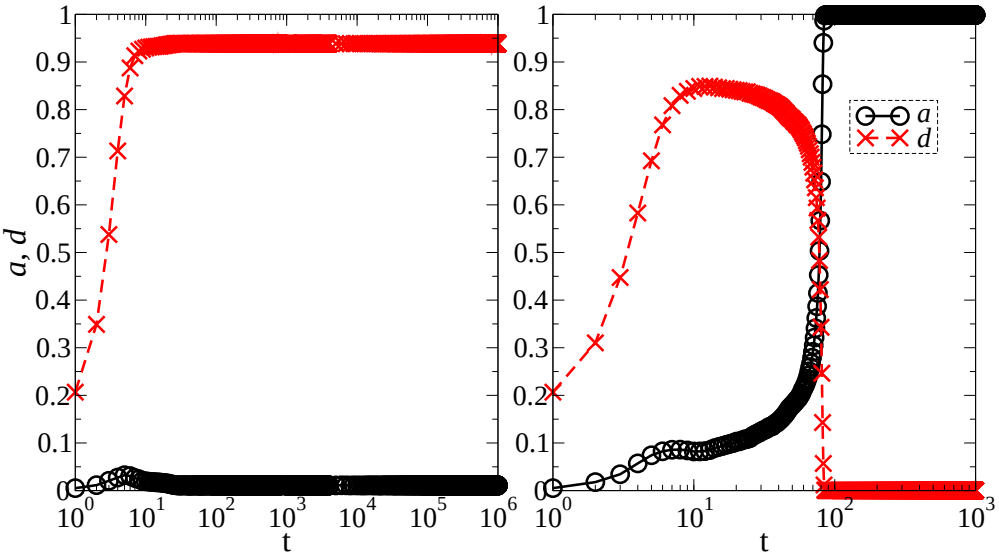


Figure 2.3: Network evolution in time, N= 10^5, $\gamma = 10^{-3}$ and $c = 0.8$. Left: Block phase, $\omega = 0.1$. Right: Global phase, $\omega = 0.3$. (Color on-line).

values of $\omega = 0.1, c = 0.4$ are fixed to achieve the block phase. The connectivity ratio γ is varied and its effect over the block structure is studied.

The block structure is sensitive to both, the number of blocks and the network connectivity γ. The block structure is stable when the network degree K is small, (γ decreases), as shown in the left panels of Fig. 2.4. Increasing K, some of the blocks may loose their stability and they are caught by blocks of larger size. Thus, a stable structure with a few blocks emerges, as seen in the middle panels. Further increasing the number of connections per node, will lead the network to the global state (right panels). In this case the border effect is strong, and the interaction between the positive and negative blocks is such that, their arrangement is no longer stable and one of them will finally attract the other to achieve a global state of positive or negative activity according to the winner blocks. A similar behavior would be observed if the number of blocks are varied for a fixed value of the network connectivity γ. Increasing the number of blocks will lead to a block phase with a narrow attractor basin.

Fig. 2.4 also shows that the block phase activity is robust for different spatial configurations of the initial block structure, and for a large amount of noise. This

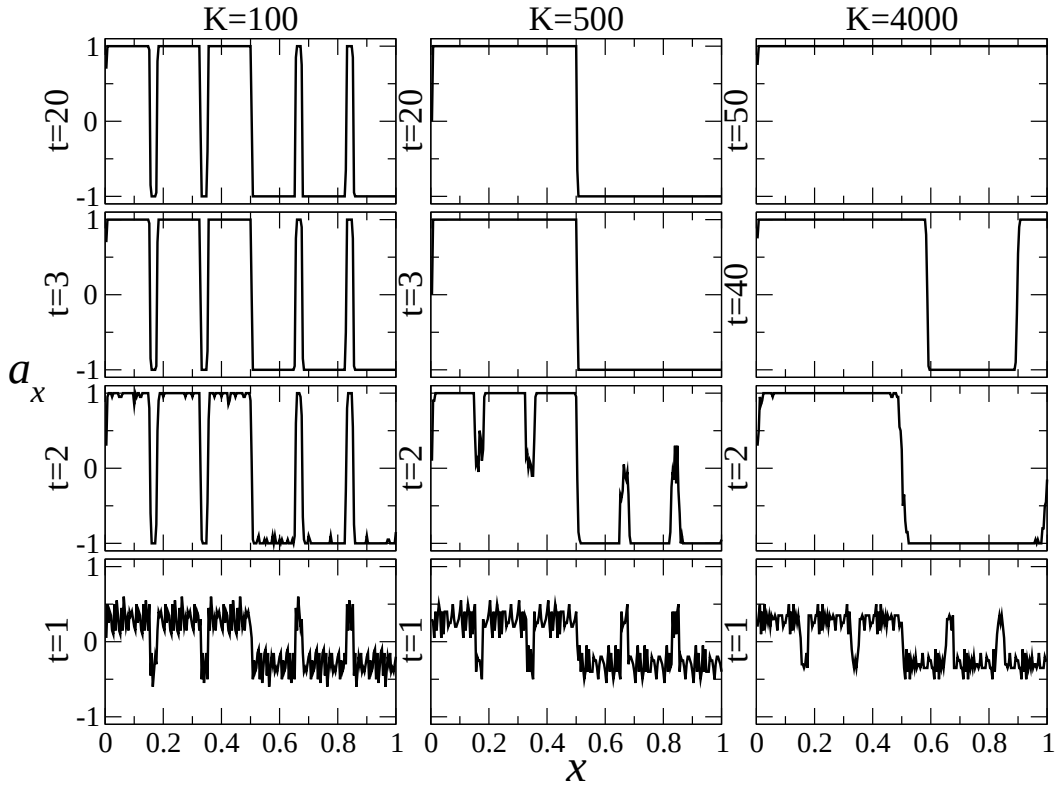


Figure 2.4: Evolution of the network into block or global phase according to the network connectivity $K = \gamma N$. The network starts in $b = 10$ non contiguous blocks of positive/negative activity ($a_l^{t=0} \sim \pm 0.3$). Network with $N = 10^4$, $\omega = 0.1,\ c = 0.4.\ x \equiv i/N$.

phase is also robust to an initial biased distribution with more positive blocks l_+ than negative blocks l_- or vice versa (Dominguez et al., 2009).

2.3.3 Phase diagram

The phase diagram for a network with $N = 10^5$, $K = 10^2$ is presented in Fig. 2.5. The diagram describes the regions of block (B_A, B_F) activity phase, with stationary states $a^* = 0$, $d^* > 0$, global (R) activity phase, with $a^* \neq 0$, $d^* = 0$, and zero (Z) activity phase, with $a^* = 0$, $d^* = 0$. It is useful to define these two regions for the block attractor according to the initial condition and to the basin of attraction, although they are the same phase. B_F ($d^{t=0} = 1$) is stable with a small attraction basin, B_A ($d^{t=0} = 0.2$) is stable with a large attraction basin.

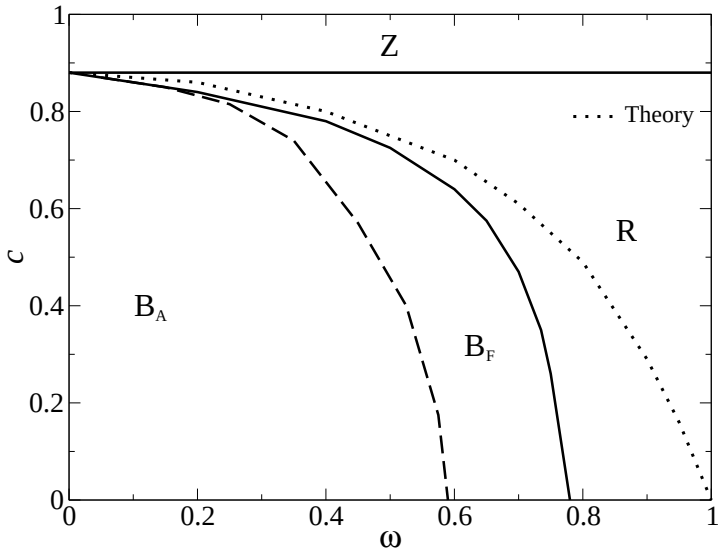


Figure 2.5: Phase diagram. B (B_A, B_F): block activity region. R: global activity region, Z: zero activity region. Network with N= 10^5 and $\gamma = 10^{-3}$.

It can be concluded from this figure that both, global and block activity appear for values of c less than approximately 0.88, which is the ordering saturation. The block activity B_F appears below the solid curve which is the block ordering saturation, and for ω not greater than approximately 0.78. For a larger ω, the stable phase is the R, for a smaller ω, B is the stable phase. A phase transition from the block to the global activity ordering is observed, when the randomness

(ω) or the c parameter of the network increases. The dotted curve corresponds to the B_F phase obtained analytically, according to the next section 2.4-Theory. The transition value of randomness with $c = 0$ slowly approaches $\omega_{BR} = 1$ when the network size increases. It was obtained $\omega_{BR} = 0.69, 0.78, 0.82, 0.86$ for $K = 50, 100, 200, 400$, respectively. Below the dashed curve the block phase B_A is stable.

2.3.4 Flow diagram

The flow diagram in Fig. 2.6 was constructed for two points in the regions of block ($\omega = 0.2$, $c = 0.7$) and global ($\omega = 0.5$, $c = 0.7$) phases presented in the Fig. 2.5. Different initial conditions were considered varying the values of $a^{t=0}, d^{t=0}$, and letting the system evolve. The stable fixed points are represented by filled circles, whereas the hollow circle corresponds to a metastable saddle point. Arrows indicate directions of the dynamic system flow.

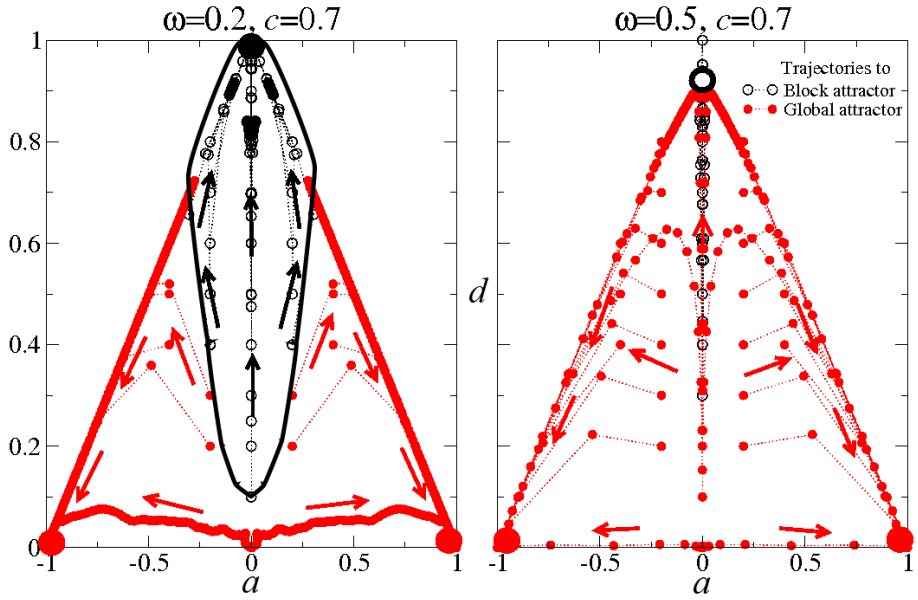


Figure 2.6: Flow diagram. Network with N= 10^5, $\gamma = 10^{-3}$, and $c = 0.7$. Left: $\omega = 0.2$, right: $\omega = 0.5$. Filled circles: stable fixed-point attractors. Hollow circle: saddle point. Red: global attractor, Black: block attractor. (Color on-line).

In the left panel the block activity phase is stable for a wide range of initial

conditions, except for very noisy block structures. The point $a = 0, d = 1$ can be considered a stable attractor and corresponds to the block phase. The block attractor basin is depicted by the region enclosed by the solid line, and the region outside represents the global attractor basin. The basin of attraction of the block attractor will shrink around the d-axis towards the point $a = 0, d \approx 1$ as one moves away from the block regions shown in the phase diagram of Fig. 2.5, being larger for B_A than for B_F. The points $a = \pm 1, d = 0$ are also stable attractors and correspond to the global phase.

In the right panel of Fig. 2.6, the global phase is stable for almost every initial condition of the system. The point $a = 0, d \approx 0.9$ is a saddle point. First, the network evolves to a block structure and then the trajectory describing the evolution of the system jumps to the global phase attractor, indicating that the neural stationary state corresponds to the global state with a positive or negative ordering.

Figure 2.7: Left: Noisy mandrill. Center: Block retrieval of the mandrill, $\omega = 0.3$, $c = 0.7$. Right: Global retrieval of the mandrill (negative), $\omega = 0.3$, $c = 0.8$.

2.3.5 An illustration

An application of the model to pattern retrieval is proposed, with a slight modification of the rule for the synaptic weights. A gauge transformation of the dynamics in Eq. (2.1) is applied, with $\sigma \to \sigma\xi$ and the Eq. (2.3) becomes

$$W_{ij} = cW^r_{ij} + (1 - c)\xi_i\xi_j. \tag{2.6}$$

where $\{\xi_i \in \pm 1\}_i^N$ represents a pattern with mean $\langle \xi \rangle \equiv \sum_i \xi_i / N = 0$. A natural image is presented to the network as a learning pattern (González et al., 2008b).

The random term W_{ij}^r is multiplied by a factor $c \in (0, 1)$ which accounts for all the previous synaptic processing, including both short-term and long-term memory of the network. The parameter c is called the *load ratio*, which resembles the amount of learned patterns by the network. The second term describes the Hebbian learning of the pattern multiplied by $1 - c$. For a given c, there is a competition between the learning pattern $\vec{\xi} \equiv \{\xi_i\}_i^N$ and the noise played by the random term.

Fig. 2.1-left shows the original pattern learned by the network which is the picture of the mandrill (USC-SIPI, 2011), a standard test image used in digital image processing. The picture is a 256×256 array of black/white pixels, properly formatted to be either ± 1. The stability and the attractor properties of the block state of the network are studied. The stimulus is characterized by a spatial partition of the pattern in a correct zone and an inverted zone. The correct zone has the neural states identical to the learned pattern, while the inverted zone has the neural states in opposition to the pattern: each neuron positioned at the same site as the pattern is switched over, i.e. the active neurons become inactive and the other way around. This is represented in Fig. 2.1-center. This state has vanishing overlap with the original pattern, however it still carries information about the picture of the mandrill. To study the stability of the block states, the network evolution should start precisely with the blocks, so that $a^{t=0} = 0$ but $d^{t=0} = 1$. Then, if the network stays at this block state, or if it goes close to it, it means that there is a block phase which is stable. However, this could be a marginal stability, with a very narrow attractor basin. To check for the size of the attractor basin, different initial conditions must be verified. Thus, the network ($N = 2^{16}$, $K = 2^6$) starts in a noisy 2-blocks mandrill ($a^{t=0} \sim 0$, $d^{t=0} \sim 0.2$), as presented in Fig. 2.7-left, and the neural dynamics leads to a stationary state. Two cases are depicted in Fig. 2.7. The attractor for equivalent topology ($\omega = 0.3$), with a load ratio of $c = 0.7$ is the block state and the recovered pattern correspond to Fig. 2.7-center. For a larger value of the load ratio $c = 0.8$, the attractor is the global state or the negative of it, as in Fig. 2.7-right.

2.4 Theory

The simulation results presented in the previous section can be supported by a straightforward theory based in a signal to noise ratio approximation (Amit, 1989; Hertz et al., 1991). The macroscopic equations for the bias model are obtained. A similar analysis can be done for the learning network.

2.4.1 Local field

Let the neurons be distributed within b blocks l, successively with positive and negative activities, $a_l = a_\pm$. Then, following Equations (2.5), the block activities can be written as

$$a_l = a + y_l d, \tag{2.7}$$

where $y_l \doteq \pm 1$ (according to the block) is a random variable.

The local field, Equation (2.1), with the Equation (2.3) for the synapses, can be separated in a signal and a noise terms,

$$h_i = (1-c)\overline{W}Ka_i + c\Omega_i \tag{2.8}$$

where $a_i \equiv \frac{1}{K}\sum_{j\in\{i\}}\sigma_j$, $\Omega_i \equiv \sum_{j\in\{i\}} W^r_{ij}\sigma_j$ are the activity restricted to the neighbors $\{i\}$, and the synaptic noise, respectively.

There are local and random neighbors for each neuron, hence the signal term itself splits in localized and randomized terms, namely

$$a_i = \frac{K_L}{K}a_i^L + \frac{K_R}{K}a_i^R, \tag{2.9}$$

with $a_i^{L,R} \equiv \frac{1}{K_{L,R}}\sum_{j\in L,R}\sigma_j$ for each of the fractions, where L and R are the local and random sets of neighbors, respectively, of the neuron σ_i.

From Eq. (2.4), whenever the neighbors belong to a block, the localized field depends on its block activity, a_l. On the other hand, the randomized field is a sample of a global field, which does not depend on the block. Using the definitions in Eqs. (2.2) and (2.7), one gets $a_i = \omega a + (1-\omega)(a + y_l d)$ for the neurons in the bulk of a block. Finally one arrives to an approximation in the limit $K \to \infty$ for the local field of neurons in the block l,

$$h_l \equiv (1-c)\overline{W}K[\omega a + (1-\omega)(a + y_l d)(1-\gamma b)] + c\Omega \tag{2.10}$$

where the correction term $(1-\gamma b)$ accounts for the boundary effects between $a_{\pm}$ blocks.

The equation for the block-activity is then, according to Eq. (2.1)

$$a_l^{t+1} = \langle \text{sign}(h^t) \rangle_{\Omega}, \tag{2.11}$$

where the average in the angular brackets are over the noise Ω which is site independent. The average over Ω stands for the noise distribution.

This noise is Gaussian distributed, $\Omega \doteq N(0, \Delta^2)$ (Amit, 1989). Its variance is given by the sum of random and local terms, $\Delta^2 = Var(\Omega_i) = \omega\Delta_r^2 + (1-\omega)\Delta_l^2$. Neglecting the *feedback* terms, one has $\Delta^2 = K$. This approximation is valid in the limit of strongly diluted networks ($K \ll N$) (Dominguez et al., 2004). For local connections, extreme dilution eliminates the feedback, because the weights W_{ij} are uncorrelated.

2.4.2 Macrodynamics

According to the Eq. (2.11), the equations for the global and block activity, after averaging over the y are given by:

$$\begin{aligned} a^{t+1} &= \langle \text{sign}(h^t) \rangle_{y,\Omega} \\ d^{t+1} &= \langle y \ \text{sign}(h^t) \rangle_{y,\Omega}. \end{aligned} \tag{2.12}$$

The averages are over the noise distribution, $\Omega \sim N(0, K)$, and over the block distribution, y. The local field is:

$$h^t \equiv (1-c)\overline{W}K[\omega a^t + (1-\omega)(a^t + yd^t)(1-\gamma b)] + c\Omega. \tag{2.13}$$

The continuous transition from R to Z phase may be analyzed by taking first $d = 0$ in the Eqs. (2.12), which gives $a = \langle \text{sign}((1-c)\overline{W}Ka + c\Omega) \rangle$. Then expanding around $a \sim 0$, it gives the constant line: $c = [1 + \sqrt{\frac{\pi}{2K}}]^{-1}$, which coincides with border R-Z plotted in Fig. 2.5. The transition between B and R phases is not continuous, so no expansion around $d \sim 0$ is possible. However, assuming the distribution of blocks is binary, the equation: $d = \langle \text{sign}((1-c)\overline{W}K(1-\omega)d + c\Omega) \rangle$, is similar to the previous equation for a except that it depends on ω. The finite

solution $d > 0$ is stable only if $c < [(1-\omega) + \sqrt{\frac{\pi}{2K}}]^{-1}(1-\omega)$. There are quantitative differences between theory and simulations for the solid curve B-R in Fig. 2.5. In Eq. (2.10) $K \gg 1$ is used for theory, whereas for simulation $K = 100$. An additional noise to be added to Ω is neglected, due to the finiteness of the sum in the signal term a_i^R in Eq. (2.9). The finite size dependence of Δ needs more precise calculations, which is outside the scope of the present Chapter, but is clear that such noise increases with ω and with $1-c$. It has been checked with extensive simulations that the B region becomes wider as the connectivity K is increased.

2.5 Conclusions

The block activity/retrieval phase (B) was studied here with two simple models. When a bias in the synaptic weights is added to random weights, the network becomes ordered in a global activity phase (R), which resembles the ferromagnetic state in a spin system. The B phase, however, is spatially structured: within each (mesoscopic) block, the neurons are ordered. If the connections between each block are less relevant than inside the blocks, as is the case of small-world networks with few long-range shortcuts, the B phase is stable. If there are enough random long-range connections, the R phase attracts almost all space of configurations: even an initial condition close to a block structure leads to a final state where all neurons are ordered.

This physics model approach is also extensible to memory networks, as shown with an illustrative example, see Fig. 2.7. Also, in the next Chapter 3 a exhaustive analysis of a Hebbian memory network is presented. When the connections are preferentially local, there are spatial correlations between neurons which allow for states retrieving localized blocks of a pattern. It was observed that these block states are stable, with a large basin of attraction. The blocks' stability depends on both the number of blocks and the network dilution. The more blocks, less stable is the block structure. When the network degree K (γ) increases, the blocks lose stability emerging a structure with less blocks of larger size, eventually leading to a global state.

2. BLOCK ACTIVITY IN A MAGNETIC INTERACTION MODEL

The proposal of a block-like structure could be closely related to biological brain systems, on the one hand, where different sensory blocks of patterns (arising from several cortical structures) may be independently retrieved. On the other hand, there could be a relation with the cortical mechanism of binding that allows information previously stored in different regions to be shared between pathways to accomplish a structured and unified representation. Blocks may also represent incomplete pieces of information which can be used to codify images (as shown here) or other kind of signals (voice, fingerprints, genetic code, etc). Implementing ANN models to deal with this type of structured data is also of interest.

Chapter 3

Structured information in a Hebbian learning attractor network

3.1 Introduction

Attractor Neural Networks (ANNs) often deal with global overlapping between patterns and neural states, employing uniform connectivity to perform retrieval tasks. They are useful when the information involved is spatially distributed because pattern learning is resistant to damage of parts of the network (Amit, 1989). When these networks start from only local stimuli, however, no global information can be achieved. The retrieval ability of such networks are commonly measured by the overlap between neuron states and memorized patterns and the load parameter expressed as the ratio between stored patterns and links per node (synaptic connections per node) (Hertz et al., 1991). Indeed, above a critical value of the load parameter no retrieval is possible and the overlap goes to zero. In addition, the mutual information between the stored patterns and the neural states has been proposed to compare the performance of different topologies in terms of memory retrieval (Dominguez and Bollé, 1998; Dominguez et al., 2007; Okada, 1996). Although the critical load increases monotonically with dilution and randomness, the information is a non-monotonic function of the load and reaches a

maximum that corresponds to a nontrivial optimal topology (Dominguez et al., 2007).

Most studies into associative memory networks with spatial structure (Hatchett et al., 2005a; Masuda and Aihara, 2004; McGraw and Menzinger, 2003; Morelli et al., 2004; Nikoletopoulos et al., 2004) focus on the global retrieval of a pattern, without considering the possibility of spatially localized states. In general, these papers deal with the relation between the storage capacity and the degree of randomness in the network. The only source of information in a long-range connected ANN (either fully connected or random), is the standard uniformly distributed overlap along the network. However, for spatially structured topologies one might measure local overlaps inside blocks of contiguous neurons. Although the information about a pattern is invariant under the reverse transformation of the global overlap, it vanishes if only half of the neuron states are flipped. Suppose a case where sequential blocks of pixels of a binary image are flipped, the overall pattern is probably still recognizable. Similar to the example in the previous Chapter, Fig. 3.1 shows a natural image in black/white pixels in the left panel, while the central panel has 2-blocks of information and the right panel has 4-blocks. The configuration of such block states has vanishing average (global) overlap with the original pattern. However, the structured distribution of local overlaps carries some spatially ordered information, the blocks oscillating between negative and positive overlaps. In other words, is possible to recognize the image even if the information about the original pattern is zero. This is called a local retrieval, to distinguish it from the usual global retrieval of the full pattern.

Unlike previous studies about structured information (bumps) in ANNs (Koroutchev and Koroutcheva, 2006; Roudi and Treves, 2006), The simplest model of binary uniform neurons has been considered, with small-world connectivity and without any reinforcement mechanism. In this way one can single out the effect of topology on the structure of the retrieval attractor. Besides this issue, one has considered non-trivial block structures. Blocks of memories can appear spontaneously as a consequence of neural dynamics combined with the network topology. The topological conditions, for the synaptic connectivity, in which local overlaps are either stable or unstable are studied. In addition, a method to

Figure 3.1: Image with different spatial distributions of overlaps. Left: Original picture. Center: 2-Blocks. Right: 4-Blocks. Both spatial distributions, 2-Blocks and 4-Blocks, have null global overlap with the original picture.

measure local information and compare it with the global information has been proposed.

3.2 The model

3.2.1 Topology and dynamics

At any given time t, the network state is defined by a set of binary neurons $\vec{\sigma}^t = \{\sigma_i^t \in \pm 1, i = 1, \ldots, N\}$. The purpose of the network is to recover a set of independent patterns $\{\vec{\xi}^\mu,\ \mu = 1, \ldots, P\}$ that have been stored by a learning process. Each pattern, $\vec{\xi}^\mu = \{\xi_i^\mu \in \pm 1, i = 1, \ldots, N\}$, is a set of site-independent unbiased binary random variables, $p(\xi_i^\mu = \pm 1) = 1/2$.

The synaptic couplings between the neurons i and j are given by the adjacency matrix $J_{ij} \equiv C_{ij}W_{ij}$, where the topology matrix $\mathbf{C} = \{C_{ij}\}$ describes the connection structure of the neural network and in $\mathbf{W} = \{W_{ij}\}$ are the learning weights. The topology matrix is split into local and random links. The local links connect each neuron to its K_l nearest neighbors in a closed ring, while the random links connect each neuron to K_r others uniformly distributed in the network. Hence, the network degree is $K = K_l + K_r$. The network topology is then characterized by two parameters, the *connectivity* ratio and the *randomness* ratio, respectively

defined by:

$$\gamma = K/N,\ \omega = K_r/K, \tag{3.1}$$

where ω plays the role of a rewiring probability in the *small-world* model (Watts and Strogatz, 1998). An extremely diluted network is obtained as $\gamma \to 0$, and the storage cost of this network is $|\mathbf{J}| = N \times K$ if the matrix $\mathbf{J}$ is implemented as an adjacency list of K neighbors.

The task of the network is to retrieve a pattern (say, $\vec{\xi} \equiv \vec{\xi^{\mu}}$) starting from a neuron state $\vec{\sigma}^0$ which is close to it. This is achieved through the neuron dynamics

$$\sigma_i^{t+1} = \text{sign}(h_i^t), \tag{3.2}$$

$$h_i^t \equiv \frac{1}{K}\sum_j J_{ij}\sigma_j^t,\ i = 1, \ldots, N, \tag{3.3}$$

where h_i^t denotes the local field of neuron i at time t. A stochastic asynchronous updating is used in the present Chapter, except in section III-D and Fig. 3.6, where parallel dynamics were used to compare simulations with theory. Stochastic macro-dynamics take place due to the extensive learning of $P = \alpha K$ patterns, where α is the load ratio. The weight matrix $\mathbf{W}$ is updated according to the Hebb's rule,

$$W_{ij}^{\mu} = W_{ij}^{\mu-1} + \xi_i^{\mu}\xi_j^{\mu}. \tag{3.4}$$

Weights start at $W_{ij}^0 = 0$ and after P learning steps, they reach the value $W_{ij} = \sum_{\mu}^{P} \xi_i^{\mu}\xi_j^{\mu}$. The learning stage displays slow dynamics, being stationary within the time scale of the faster retrieval stage Eq. (3.2).

3.2.2 The information measures

Previous studies have only dealt with global measures of information, which are adequate to describe networks with no local connectivity. For small-world connectivity is useful to define blocks as the structured pieces of information that emerge in the network. If the contiguous neurons are distributed within b blocks, for simplicity each of size $L = N/b$, then the *block's overlap* between the neural states and one individual pattern restricted to the lth block ($l = 1, \ldots, b$) is:

$$m_l \equiv \frac{1}{L}\sum_{i \in l} \xi_i\sigma_i, \tag{3.5}$$

at an unspecified time step. One can consider m_l as a random variable and estimate the average of this variable across the blocks as: $\langle f_l \rangle_b \equiv \frac{1}{b}\sum_{l=1}^{b} f_l$.

The relevant order parameters measuring quality of retrieval are the mean (m) and the variance (v) of the block overlap distribution, defined as

$$m \equiv \langle m_l \rangle_b, \;\; v \equiv \langle m_l^2 \rangle_b - m^2. \tag{3.6}$$

Note that m is the usual *global* overlap, also written as $m \equiv \frac{1}{N}\sum_i \xi_i \sigma_i$. When the global overlap is zero and the size of the blocks is taken as $L = 1$, the network carries no macroscopic order. On the other hand, if there is only one block, $b = 1$, the variance is zero and no local information is carried. However, if the size is large but $1 \ll L < N$, the variance is finite and the blocks convey only local information. The standard deviation, which is named the *local* overlap, is $\delta = \sqrt{v}$. It is worth mentioning that the blocks are macroscopically scaled and hence the parameter δ is not related to a spin-glass, which is a microscopic ordering.

Together with the overlap, the load ratio $\alpha \equiv P/K$, that accounts for the storage capacity, is needed. As the number of stored patterns grows the network is not able to retrieve them and the overlap goes to zero. To fully describe the performance of a structured network with a unique measure, it is useful to apply tools from information theory (Dominguez et al., 2007). One first calculates the global mutual information, $\mathcal{M}$, a quantity used to measure the information that an observer can receive at the output of a channel. The recall process of stored patterns that is being considering here can be regarded as a channel, with the pattern as the input and the neuron states as the output. The mutual information can be defined in terms of the order parameter $m = \langle \sigma \xi \rangle_{\sigma,\xi}$, in the limits $K, N \to \infty$. The brackets represent an average over the joint distribution $p(\sigma, \xi)$ for a single neuron, which is understood as an ensemble distribution for the neuron states $\{\sigma_i\}$ and patterns $\{\xi_i^\mu\}$. This will give $\mathcal{M}[\sigma;\xi] = S[\sigma] - S[\sigma|\xi]$, with the entropy of the output channel $S[\sigma] = 1[bit]$ and the conditional entropy obeying (Dominguez and Bollé, 1998):

$$S[\sigma|\xi] = -\frac{1+m}{2}\log_2\frac{1+m}{2} - \frac{1-m}{2}\log_2\frac{1-m}{2}. \tag{3.7}$$

Hence one defines the global information ratio as

$$i_m(\alpha, m) \equiv \alpha \mathcal{M}[\sigma; \xi], \tag{3.8}$$

for independent neurons and uncorrelated patterns. The global information ratio is useful to evaluate the network performance when there is a global stimulus (Dominguez et al., 2007). Below, the local information due the distribution of blocks is estimated when the global overlap is $m = 0$.

One considers a sample of b independent blocks of pattern overlaps, their distribution being described by their mean m and variance v. The local information can be estimated from a Gaussian channel with the output state comprised of a signal term m_l with variance v and a noise term z, whose variance is assumed to be $v_z \sim 1$ (maximal for the signal). Thus, the mutual information satisfies $\mathcal{M}[\vec{\sigma}, \{\vec{\xi}^{\mu}\}] \sim S[m_l + z] - S[z]$ and the local information ratio is roughly (Cover and Thomas, 1991)

$$i_v(\alpha, m_l) \simeq \alpha \log_2(1 + v). \tag{3.9}$$

It should be noted that the underlying block distribution is unknown, except its first and second moments. The approximation in Eq. (3.9) supposes that the block distribution is Gaussian, and therefore it provides an upper bound estimation of the local information. The estimation for i_v is not as exact as the expression for the global information i_m. Nevertheless, it works well when $v = 0$ (all blocks have same overlap), such that $i_v = 0$ and there is only global information $i_m \geq 0$. Moreover, it also scales well with v in the case of perfect blocks ($v = 1$), since the information regarding the blocks corresponds to $i_v = \alpha$ bits. Note that if no spatial correlation emerges, m_l provides no information at all and it can be regarded as pure noise (both $i_m = 0$ and $i_v = 0$). Eq. (3.9) holds as long as the block is sufficiently large, $L \gg 1$. Both global and local information are not manipulated jointly in a unified formula, and therefore, they are analyzed separately.

The validity range for each expression in Eqs. (3.8,3.9), for the global and local information, are either the global ($\delta = 0$) or the local retrieval regime ($m = 0$), respectively. Finally, note that the global and local overlaps evolve with time according to the neuron dynamics Eq. (3.2), such that at each step they have a time index m^t and δ^t.

3.3 Results

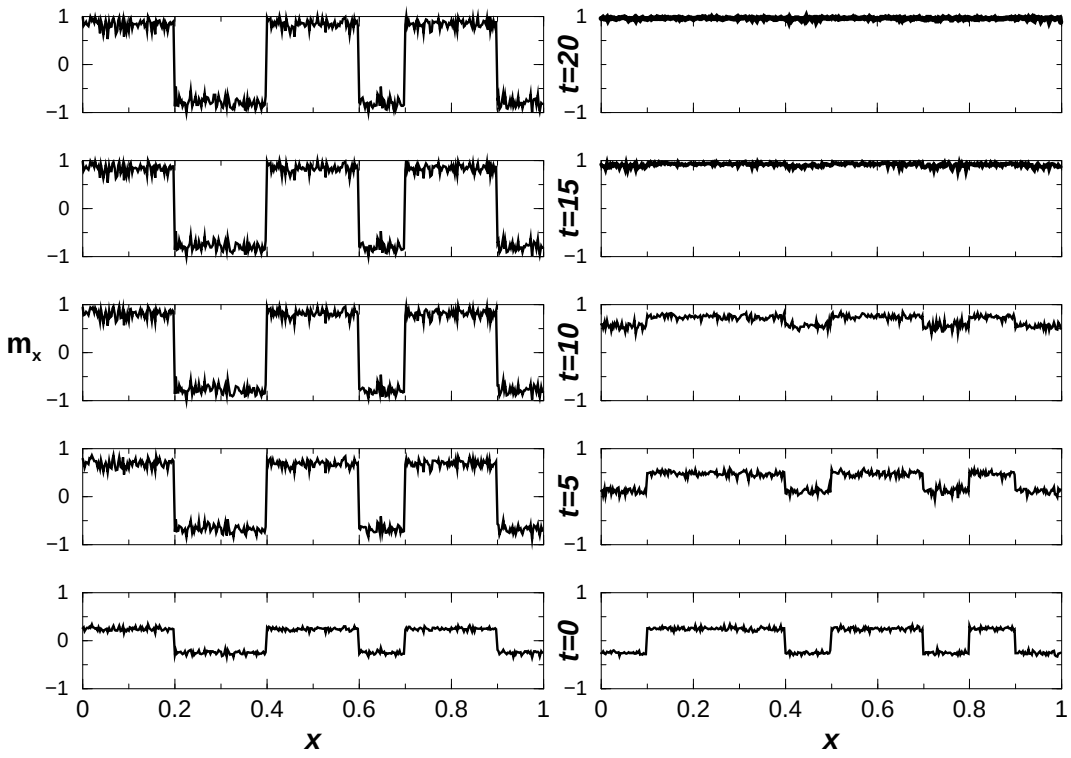


Figure 3.2: Evolution of blocks with $m_l^0 = \pm 0.3$, from $t = 0$ (bottom) to $t = 20$ (top), for $\gamma = 10^{-4}$, with $N = 10^6$. Left: $\omega = 0.1, \alpha = 0.05$. Right: $\omega = 0.5, \alpha = 0.20$. $x \equiv i/N$.

3.3.1 Simulations: the retrieval evolution

The dynamical neuron equations (3.2-3.4) are simulated with the topology parameters defined according the Eq. (3.1). The block overlaps time evolution is illustrated in Fig. 3.2 for a network of $N = 10^6$ neurons distributed in $b = 10$ blocks and with connectivity $\gamma = 10^{-4}$. The initial block overlaps were chosen at random from a discrete uniform distribution $m_l^{t=0} = \pm 0.3$. In the left panel $\alpha = 0.05$ and $\omega = 0.1$, in the right panel $\alpha = 0.20$ and $\omega = 0.5$. To improve the clarity of the figures, the neuron overlaps m_i have been smooth averaged across

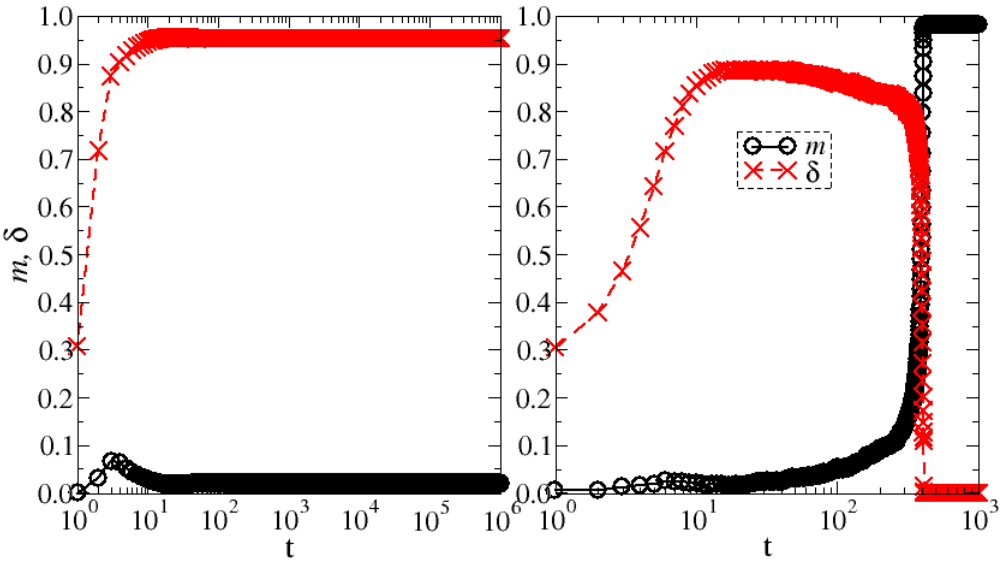


Figure 3.3: Network evolution, with $N = 10^6$, $\gamma = 10^{-4}$, $\omega = 0.3$. Left: B phase, $\alpha = 0.1$. Right: R phase, $\alpha = 0.2$. (Color on-line)

uniform windows (of size $\delta L = 10^3$) inside the blocks (of size $L = 10^5$). Hence the plotted curves, m_x, $x = i/N$, are smoother than the actual m_i, although some of the structure can still be appreciated in m_l. In these plots time evolves from $t = 0$ (bottom) to $t = 20$ (top), which is close to the stationary state. The left panels show that the network retains its initial block configuration: the blocks are retrieved as independent patterns increasing their overlaps to the fixed point $m_l^* \sim \pm 1.0$. In the right panel, the blocks lose their starting signals and the full pattern is completed, $m^* \sim 1.0$. While for local topology ($\omega = 0.1$) the block structure persists, for random topology ($\omega = 0.5$) the local information is converted into global information.

The time evolution of a network starting in blocks with $m_l^{t=0} \simeq \pm 0.3$ is depicted in Fig. 3.3. The randomness is $\omega = 0.3$. For a low load ratio, $\alpha = 0.1$, the network maintains its block configuration and the local overlap fluctuates slightly around the fixed point $\delta^* \sim 0.95$, as seen in the left panel. For a larger load, $\alpha = 0.2$, the network evolves to the global attractor in a more complex way, as seen in Fig. 3.3 right. Initially the network approaches the local state, improving the blocks' overlap ($m^t \sim 0, \delta^t \sim 0.9$). After $t = 400$ time steps, the

network quickly loses its block structure, and evolves into a global state with the completion of the learned pattern, $m \approx 1$. The stability in both cases has been checked, by running the simulation for up to $t = 10^6$ time steps. Although there may be microscopic cycles, the macroscopic state converges to either the local or the global retrieval phase (Paula et al., 2006). Just as a remark, the models of magnetic interaction and Hebbian learning correspond to an almost perfect "gauge" transformation from each other, perhaps with the exception that the Hebbian weights have triangular correlations, while the magnetic weights are independent. So, the dynamical results are approximately identical as depicted in Fig. 2.3, although the meaning of the parameters is not the same.

3.3.2 Simulations: the learning capacity

The stationary states of the network are studied as a function of the load ratio α for different values of the topological parameter ω. A sample of the simulation results is shown in Fig. 3.4. The stationary global and local overlaps, m^* an δ^* respectively, are plotted in the top panels for $\gamma = 10^{-3}$ and randomness ranging from $\omega = 0.0$ to $\omega = 1.0$. The network starts in a block configuration with $b = 10$, $m^0 = 0$ and $\delta^0 = 1$. It was observed that the global overlap remains close to $m^* \sim 1$ for random networks, while for local networks the local overlap increases to $\delta^* \sim 1$, up to the respective global and local critical capacities.

One reason for this behavior is that randomness decreases the mean-path-length between neurons, facilitating the propagation of the information around the network and yielding a global ordering. On the other hand, locality increases the clustering of neurons, slowing down the transmission of information across the network and stabilizing the formation of the blocks. It is worth noting that the network can retrieve the full pattern achieving $m^* \sim 1$ even starting at $m^0 = 0$, thanks to the role of the topology and the $\pm$ block overlaps,.

There is also a critical load for the global and local overlaps, α_R and α_B respectively. For instance, in the middle panel ($\omega = 0.5$) the local overlap decays sharply to $\delta^* = 0$ at $\alpha_B \sim 0.05$, coinciding with an increase of the global overlap to $m^* \sim 1$. Later in the learning process (increasing further on the load ratio),

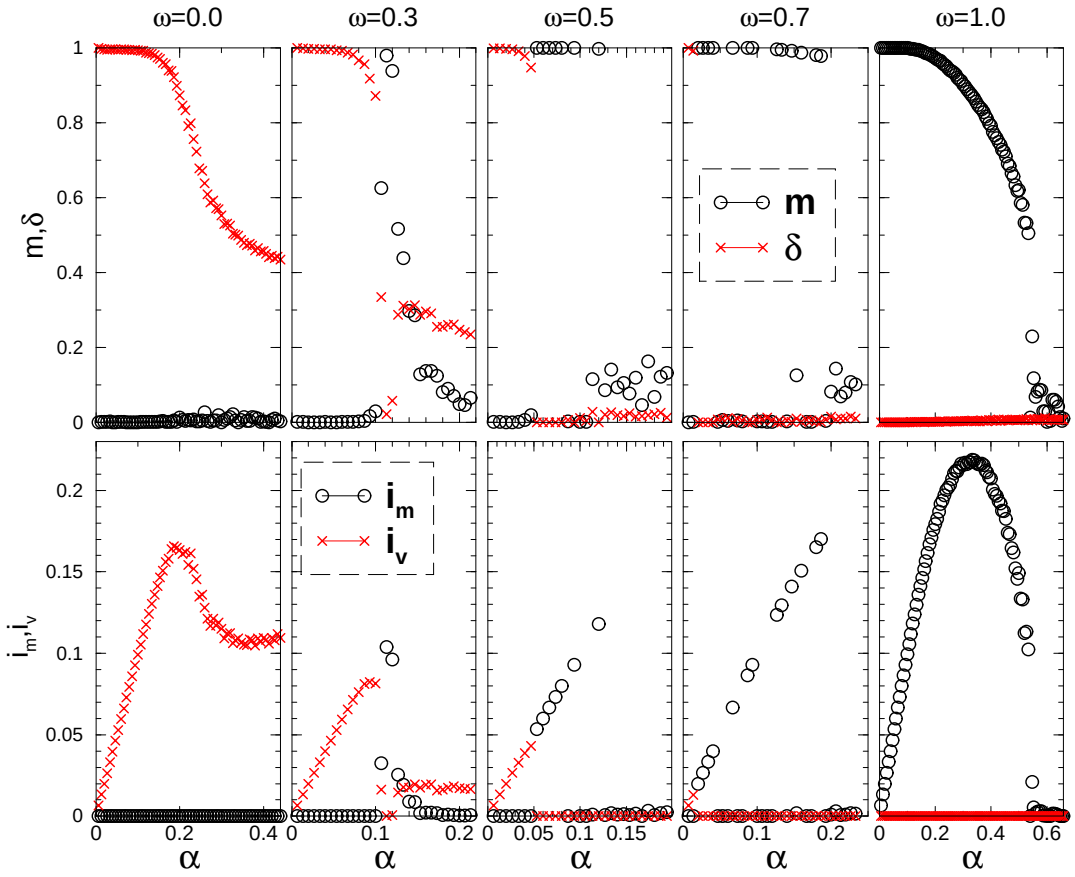


Figure 3.4: Global and local overlaps m, δ (top) and information i_m, i_v (bottom), vs α, for $b = 10$, $\gamma = 10^{-3}$, and ω from 0.0 (left) to 1.0 (right). Networks with $N = 3 \times 10^5$ and initial states $m^0 = 0, \delta^0 = 1$. (Color on-line)

at $\alpha_R \sim 0.11$, it establishes a second transition for the global overlap close to $m = 0$.

Fig. 3.4 (bottom) displays the global and the local information as function of the load ratio. The panels show that both local and global are a non-monotonic function of the load ratio. It can be seen that the maximal local (global) information i_v (i_m) decreases (increases) with the values of ω. Indeed, the maximum of the local information, $i_v \sim 0.17$ for a local topology, $\omega = 0.0$, is comparable to the maximum of the global information $i_m \sim 0.22$ for a random topology $\omega = 1.0$. Intermediate randomness (middle panels) led to competition between blocks and global structures. Indeed, above some load critical ratio α_B, the blocks lose stability and the neurons shift to the global information mode.

3.3.3 Phase diagram

According to the definitions of the order parameters, and following the results presented in the previous sections, the stationary states of this small-world ANN can be defined. The network may exhibit either a global *retrieval* (R) phase, with $m \neq 0, \delta = 0$, or a *local* retrieval (B) phase, with $m = 0, \delta > 0$, which carries non-vanishing information. Furthermore, there is a *zero* (Z) phase, with $m = 0$ and $\delta = 0$, without any information. When the network starts in the vicinity of a pattern, it will move closer to that pattern if the load ratio is lower than that of the global retrieval saturation, $\alpha_R(\omega)$. When the blocks of a network start successively near to a pattern or the inverse of that pattern, the block configuration will persist if the load is lower than the local retrieval saturation, $\alpha_B(\omega)$. For large ω, the stable phase is the R whereas for small ω, B is the stable phase.

In order to study these phases, extensive Monte Carlo simulations of the system were performed, with $K = 100$ and $N = 10^6$ neurons ($\gamma = 10^{-4}$). The phase diagram is shown in Fig. 3.5, for a network with $b = 2$ blocks. The initial condition for the B phase is $m^0 = 0$ and $\delta^0 = 1$. The separation between the R_1 and R_0 phases (the thin dashed line) is justified by their distinct dynamical character: while R_1 is the usual retrieval attractor for initial conditions $m^0 \neq$

$0, \delta^0 = 0$, the R_0 is the retrieval attractor when the network starts with no initial global overlap, $m^0 = 0, \delta^0 \neq 0$.

There is a transition from the Z to R phases at a $\alpha_R(\omega)$, represented by the thick dashed line, below which the global retrieval information is always stable. The local information appears at the line $\alpha_B(\omega)$ (solid line), below which the B phase coexists with R. One could see that the B region holds steady at a larger α for local networks than for random networks. Note that mixed states, with both $m \neq 0, \delta \neq 0$, emerged from no pure R or B-like initial conditions, as seen in Fig. 3.2. It has been checked through extensive simulations that the local solution is robust for a wide range of numbers of blocks, and it was found that the transition $\alpha_B(\omega)$ in the phase diagram of Fig. 3.5 shrinks to zero for $b > 10^3$. The R region also collapses to the $\alpha = 0$ axis and only the Z phase survives.

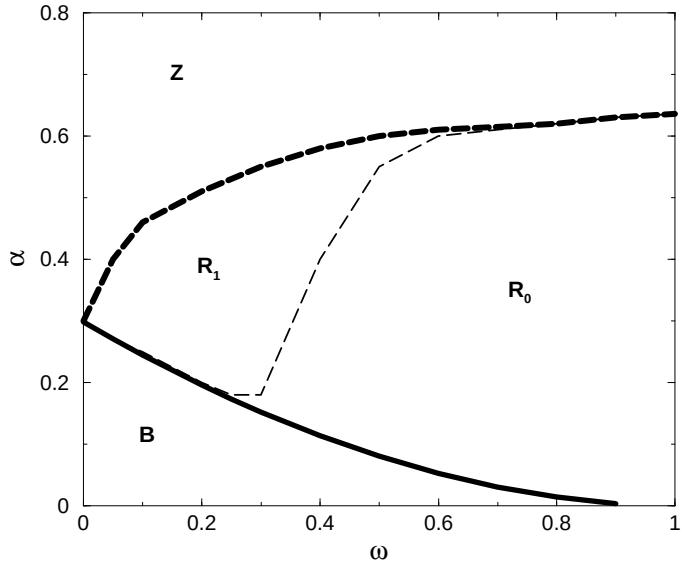


Figure 3.5: Phase diagram ($\omega \times \alpha$) for $N = 10^6$, $\gamma = 10^{-4}$ and $b = 2$. $B \equiv$ Local phase. $R_0(R_1) \equiv$ Retrieval phase with $m^0 = 0, \delta^0 = 1$ ($m^0 = 1, \delta^0 = 0$). $Z \equiv$ no information.

3.3.4 Theory

In this section a strongly-diluted network is considered and theoretical equations are proposed for the macroscopic order parameters. The sketch of a proof that

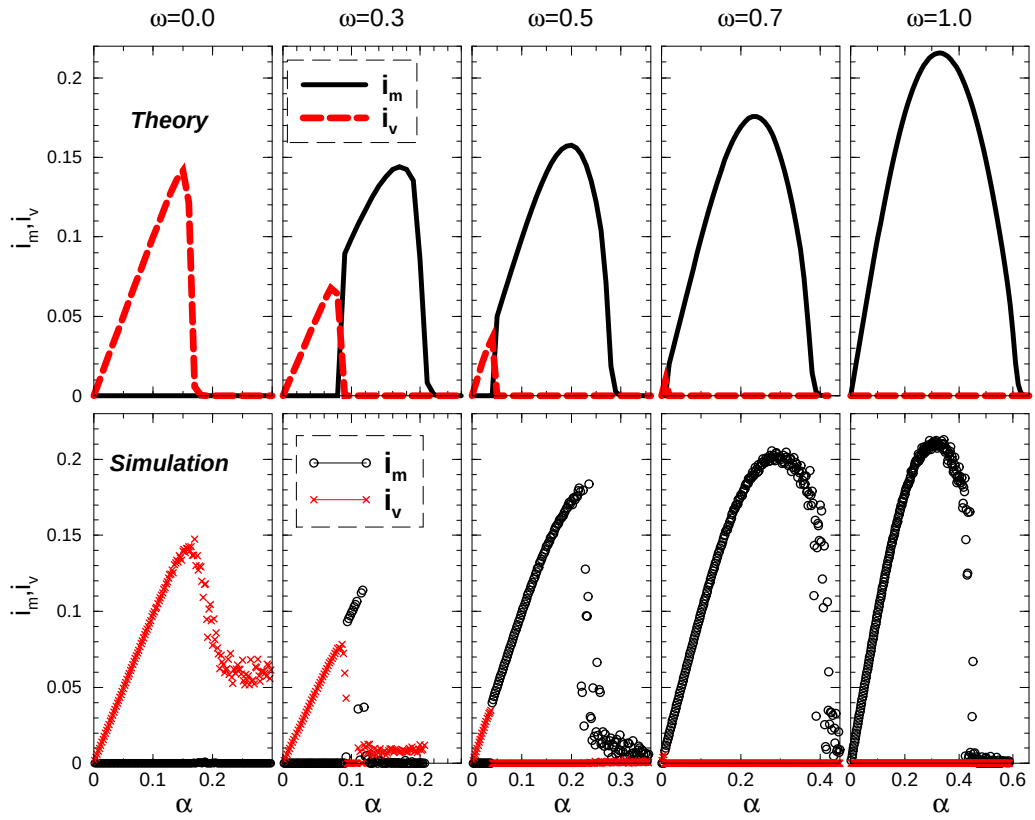


Figure 3.6: Global (solid lines and circles, i_m) and local (dashed lines and $\times$, i_v) information, vs α, for $b = 2$. Randomness runs from ω from 0 (left) to 1 (right), $\gamma = 10^{-2}$. Top: theory. Bottom: simulation with $N = 10^5, K = 10^3$ and parallel dynamics. Initial states: $m^0 = 0.04; \delta^0 = 1$. (Color on-line)

is valid for stationary equations can be found in the appendix. Here empirical dynamical equations whose fixed points coincide with the stationary solutions are added . This extrapolation is justified for parallel dynamics, for which all neurons are updated once at each time step. It is supposed that the neurons are distributed within b blocks, with positive and negative overlaps, $m_l \equiv m_\pm$. Hence the global overlap is $m = (m_+ + m_-)/2$ and the fluctuation between blocks is $\delta = (m_+ - m_-)/2$.

An approximation for the local field of neurons at block m_l at time step t gives

$$\xi h_l^t \equiv \omega m^t + (1-\omega)(m^t + y_l \delta^t)(1 - \gamma b) + \Omega^t \tag{3.10}$$

where ξ is the pattern being retrieved. The pattern-interference noise follows a Gaussian distribution, $\Omega \doteq N(0, \alpha r)$, where $r = \omega r_r + (1-\omega) r_l$ is the sum of the random and local *feedback* terms, $r = \omega r_r + (1-\omega) r_l$, with $r_r = 1$ and $r_l = (1-\chi)^{-2}$. The *susceptibility* χ arises from the local connections,

$$\chi^t = \frac{1}{\sqrt{\alpha r_l}} \langle z \, \mathrm{sign}(\xi h^t) \rangle_{y,z}. \tag{3.11}$$

With the field in Eq. (3.10), the macro-dynamics for the global and local overlap are

$$\begin{aligned} m^{t+1} &= \langle \mathrm{sign}(\xi h^t) \rangle_{y,z} \\ \delta^{t+1} &= \langle y \, \mathrm{sign}(\xi h^t) \rangle_{y,z}. \end{aligned} \tag{3.12}$$

The averages are over the binary distribution of $y \in \pm 1$ and a Gaussian $z \doteq N(0,1)$. There are two types of stationary states: (1) $m \neq 0, \delta = 0$, with $m = \mathrm{erf}(m/\sqrt{2\alpha r_l})$ and (2) $m = 0, \delta > 0$, with $\delta = \mathrm{erf}(\delta(1-\omega)/\sqrt{2\alpha r_l})$. The first is the usual Amit's solution (Amit, 1989), while the second is the local solution, which is stable if $(1-\omega) \geq \sqrt{\pi \alpha r/2}$. Adjusting the solid curve α_B in the Fig. 3.5 to $(1-\omega) \approx A_0 \alpha^{A_1}$, gives $A_1 = 0.51$, which fits well with this theoretical prediction if one assumes r is constant in the transition.

A comparison between the theoretical results (upper panels) and the simulation (bottom panels) for parallel dynamics is shown in Fig. 3.6. Qualitatively, the behaviors predicted by theory, of local retrieval with small ω and global retrieval

with large ω, agree quite well with that checked by simulations. The transition from local to global retrieval at a given $\alpha(\omega)$ (for intermediate randomness $0.3 \leq \omega \leq 0.7$), as well as the maximal local and global information, are also in agreement. Moreover, Eq. (3.10) explains why local retrieval fails for $L \sim K$, for which $\gamma b \sim 1$ and the boundary effects between blocks are relevant when compared to the bulk of the connected neurons. The degree of dilution plays the role of a resolution scale for the boundary effects. Therefore, only diluted networks are able to stabilize blocks.

Finally note that, both the theoretical results and the simulations start with $m^0 = 0.04$, because $m^* \neq 0$ would never be achieved if the initial global overlap is zero, as observed in the simulation for asynchronous update (see Fig. 3.4). The theoretical equations for the asynchronous dynamics are more involved than the parallel macro-dynamics and they lead to differential equations, the calculation of which is beyond the scope of this Chapter.

Except for this different dynamical behavior, both parallel and asynchronous updating produce similar attractors. It is worth comparing the simulation results of the lower panels of Fig. 3.6 with those of Fig. 3.4. First, it should be noted that they were calculated for different values of the network parameters (see the figure legends). For instance in Fig. 3.6 the connectivity is $K = 10^3$ while in Fig. 3.4 it is $K = 300$. The pattern storage capacity is related to K, hence the Fig. 3.6 displays more points for each learned pattern. In Fig. 3.4 the networks are also more diluted ($\gamma = 10^{-3}$) than in Fig. 3.6 ($\gamma = 10^{-2}$), which should lead to a larger storage capacity for the global overlap in Fig. 3.4 than in Fig. 3.6 (Dominguez et al., 2007). However, with a $\omega = 0.5$ and $\omega = 0.7$ in both panels, the more diluted network (Fig. 3.4) reaches a smaller critical load α_R. In fact, there are many holes in the curve that represent failures in the recovery of the pattern. The reason for this disagreement with the expected result is that the network has only $b = 2$ blocks in Fig. 3.6, in contrast with the $b = 10$ blocks in Fig. 3.4. The more blocks the network is initialized with, the more difficult it is to complete the pattern. This occurs because the borders between each block make them less stable.

3.4 Conclusions and discussion

In this Chapter was studied a new type of solution for an attractor neural network: the local retrieval phase (B) with overlaps structured in blocks. The dependence of the stability of the local B and the global retrieval (R) phases on the topological parameters of connectivity (γ) and randomness (ω) was analyzed. Although the storage capacity is severely disrupted when the long-range nature of the connections is lost, it was found that local information emerges when the network has a more local topology. The local information corresponds to configuration states which carry information in blocks of neighboring neurons, and are attractors of the network dynamics. A block structure might resemble the metastable mixed states studied in the seminal work of Amit (1989). These spurious states, where the network only recognizes mixture of patterns, may provide useful information, as might the B phase, because the blocks are spatially ordered. However, while the local field arising from different patterns can cooperate to retrieve both of them, the local field induced by negative/positive patterns cancels each of these out, and the blocks are unstable in networks without the type of topology studied here.

Both in biology and in hardware implementations of neural systems, mainly neighboring neurons are connected in networks. Such short-range architectures are much cheaper in terms of the wiring cost than long-range ones, but the downside is that the network loses most of its global retrieval capacity. In this Chapter was shown that such a structure induces another information retrieval capacity: that of local retrieval. Another novel situation that arises from this memory structure is that the information from blocks may be transferred to a global retrieval if the range of links is long enough, or if more patterns are stored. Hence, the topology complexity improves the retrieval attractor basin. It was found that the transition from $R \to B$ takes place for $\alpha \leq \alpha_B \approx (1-\omega)^2$, and a theory for strongly diluted networks was proposed, which fits well with the simulations. The information entropies were also estimated, for both R and B phases.

The blocks behave as independent pieces of information. Thus, instead of the small number P of patterns of size N that a diluted network can store, this phase is able to retrieve $b \times P$ patterns, each with size N/b. The existence of

an information phase with no global overlap may play a relevant role in natural neural networks, for instance, to manage a successful response to stimuli activating separated cortical areas (Roudi and Treves, 2004). Also in many applications of pattern classification, such as image recognition, by carrying local spatial information the overlaps may have opposite signals in separate blocks, although overall information might be generated. Minor changes in the topology **C**, for instance that suppress symmetry constraints, will lead to complex dynamics for the blocks, including cycles and chaos, which could model higher functions of the brain.

3.5 Appendix: Macro-dynamic equations

Let the neurons be randomly distributed within b blocks and for simplicity, each of size $L = N/b$, with positive and negative overlaps, $m_l = m_\pm$. The blocks $l = 1, \ldots, b$ are built as the sets $\lambda_l = \{i = (l-1)L + k; k = 1, \ldots, L\}$. Then, according to Eq. (3.6) the global overlap is $m = \sum_l^b m_l/b$ and the fluctuation between blocks is $\delta = \sqrt{v}$, with $v \equiv \sum_l^b m_l^2/b - m^2$. The block's overlap can be written as

$$m_l = m + y_l \delta, \tag{3.13}$$

where $y_l \doteq \pm 1$ (according to the block) is a random variable.

The local field of the neuron σ_i, Eq. (3.3), applying Eq. (3.4) for the weights, can be separated into a signal and a noise term, if a given pattern is being retrieved, say $\xi \equiv \xi^{\mu=1}$:

$$h_i \equiv \xi_i m_i + \Omega_i \tag{3.14}$$

where

$$m_i^\mu \equiv \frac{1}{K} \sum_{j \in C} \xi_j^\mu \sigma_j, \; \Omega_i \equiv \sum_{\mu > 1} \xi_i^\mu m_i^\mu \tag{3.15}$$

are the overlap restricted to the neighbors C of neuron σ_i, and the cross-talk noise, respectively.

There are local and random neighbors for each neuron, hence the signal term itself splits into localized and randomized terms, namely

$$m_i = \frac{K_r}{K} m_i^r + \frac{K_l}{K} m_i^l, \tag{3.16}$$

with $m_i^x \equiv \frac{1}{K_x}\sum_{j\in C_x}\xi_j\sigma_j$, $(x \equiv l, r)$ where K_l^i and K_r^i are the local and random sets of neighbors of the neuron σ_i, respectively.

From Eq. (3.5), whenever the neighbors belong to a block the local field depends on its block's overlap, m_l. On the other hand, the randomized field is a sample of a global field that does not depend on a block. Using the definition in Eq. (3.1), one arrives at an approximation for the local field of neurons in the block λ_l

$$\xi h_l \equiv \omega m + (1-\omega)(m + y_l\delta)(1-\gamma b) + \Omega \tag{3.17}$$

where ξ is the pattern being retrieved. The correction term $(1-\gamma b)$ accounts for the boundary effects between $m_\pm$ blocks.

The equation for the block's overlap is then $m_l = \langle \text{sign}(\xi h)\rangle_\Omega$, where the average in the angular brackets are over the noise Ω. But from Eq. (3.13) $m_l = m + y_l\delta$, and thus, after averaging over y_l one gets

$$\begin{aligned} m &= \langle m_l\rangle_y = \langle \text{sign}(\xi h)\rangle_{y,z} \\ \delta &= \langle y m_l\rangle_y = \langle y\ \text{sign}(\xi h)\rangle_{y,z}. \end{aligned} \tag{3.18}$$

The average over z stands for the noise distribution, Ω.

This noise is a large sum of almost-independent terms, which converges to a Gaussian distribution, $\Omega \doteq N(0,\Delta)$ (Hertz et al., 1991). Its variance $\Delta = \alpha r$ is given by the sum of random and local *feedback* terms, $r = Var(m_i^\mu) = \omega r_r + (1-\omega)r_l$. To deal with them, one can consider the residual overlaps ($\mu > 1$) as stochastic variables. If one expands the residual overlaps around $h_j^\mu \equiv h_j - \xi_i^\mu m_i^\mu$ it holds:

$$m_i^\mu \sim \frac{1}{K}\sum_{j\in C}\xi_j \text{sign}(h_j^\mu) + m_i^\mu\chi_i, \tag{3.19}$$

$$\chi_i \equiv \frac{1}{K}\sum_{j\in C}\frac{d}{dh_j^\mu}\text{sign}(h_j^\mu) \tag{3.20}$$

here χ_i is the *susceptibility*. The first term in the r.h.s of Eq. (3.19) is not correlated with the second term, and its variance is α. So, the stochastic equation reads $m_i^\mu(1-\chi) \simeq N(0,\alpha)$, and the feedback term is

$$r_x = (1-\chi_x)^{-2};\ x = l, r. \tag{3.21}$$

One now supposes only strongly diluted networks ($K \ll N$). For random connections, χ_r can be neglected since there is no feedback in the dynamics and $r_r = 1$ (Dominguez and Bollé, 1998). However, for local connections, even extreme dilution does not eliminate the feedback and thus, susceptibility can also be written as

$$\chi_l = \frac{1}{\sqrt{\alpha r_l}} \langle z \ \mathrm{sign}(\xi h) \rangle_{y,z}. \tag{3.22}$$

3. STRUCTURED INFORMATION IN A HEBBIAN LEARNING ATTRACTOR NETWORK

Chapter 4

Structured information in sparse coding networks

4.1 Introduction

Attractor Neural Networks (ANNs) with metric connections are being investigated through several approaches like pruning couplings (Zheng et al., 2010), scale-free networks (McGraw and Menzinger, 2003), small-world networks with binary neurons (Dominguez et al., 2007; Morelli et al., 2004; Nikoletopoulos et al., 2004), continuous neurons (Li and Chen, 2003) and integrate-and-fire neurons (Masuda and Aihara, 2004). Main attention has been attached to the global overlap which implies that the learning capacity, in general, is drastically disrupted for short-range neural networks. This is a counterpart for the considerable amount of wiring that can be saved by using metric architecture. Recently, attempts have been made to explore distinct types of spatially structured stable neural states, with localized overlaps and no meaning for networks with uniform couplings. Among them, one can distinguish two classes of structures: a bump (Koroutchev and Koroutcheva, 2006; Roudi and Treves, 2004, 2006) and a block (Dominguez et al., 2009; González et al., 2009) structure of overlaps. As an illustrative example, Fig. 4.1-left shows a natural image in black/white pixels which is a biased pattern with more active (black) than inactive (white) neurons. Global retrieval should recover this pattern in the same shape. In Fig. 4.1-center a 2-blocks structure of the same pattern is displayed: the left half has a positive overlap, while

the right half has a negative overlap with the original pattern. This negative overlap distribution could be thought as the Inverted Tuning Curves of activity, recorded in some cells in the Prefrontal Cortex of monkeys performing spatial working memory tasks (Wang et al., 2004). These neurons have a lower firing rate after some cue is presented than during spontaneous activity. In Fig. 4.1-right, a bump structure is presented. A bump is a localized retrieval state which has a region with high correlation with the original pattern (the left half), and another region which has no correlation at all (the right half). Spatially localized bumps of activity are relevant in attractor neural networks when trying to give a computational explanation of working memory capacity (Edin et al., 2009; Roudi and Treves, 2008).

Sparse coding is the representation of items by the activation of a relatively small set of neurons. Sparse-coding gives the model a biological plausibility since the brain suggests a general sparse-coding strategy (Foldiak and Endres, 2008). Sparse-coding is also favorable to increase the network capacity, because the cross-talk term between stored patterns decreases. However, to sustain a low rate of activity in associative memory networks is very difficult and a control mechanism becomes necessary (Dominguez and Bollé, 1998). In the present Chapter is used a dynamical threshold that depends on the global activity of the network, the local activity of the neuron's neighborhood, and the sparseness of the pattern coding. This threshold strategy is a reinforcement mechanism instead of a rigid constraint used in previous studies about bumps (Koroutchev and Koroutcheva, 2006; Roudi and Treves, 2006, 2008), allowing for the manifold of retrieval solutions for the network dynamics. Increasing the short-range nature of the connections, the network undergoes a transition from a global to a spatially structured retrieval phase. Furthermore, the different types of structures (block or bump) have been characterized, according to their dependence on the neural activity level. It is also shown which conditions favors the emergence of these localized states, and the regions in a phase diagram $\alpha \times a$ where they either coexist or dominate.

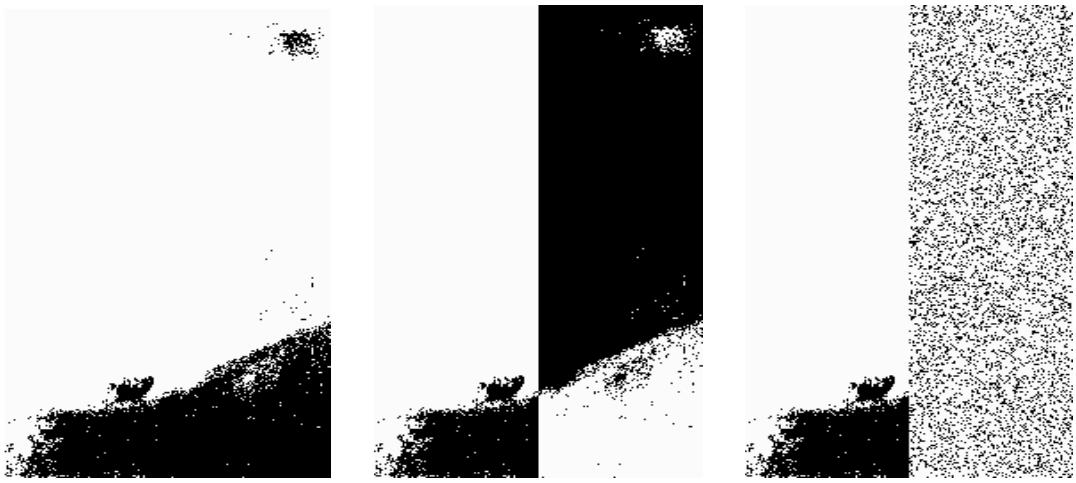


Figure 4.1: Left: Goya's Dog - original. Center: 2-blocks Dog. Right: Bump Dog, $a = 0.2$.

4.2 The model

4.2.1 Topology and dynamics

At any given discrete time t, the network state is defined by a set of N independent binary neuron variables $\boldsymbol{\tau}^t = \{\tau_i^t \in 0, 1; i = 1, \ldots, N\}$, where 1 and 0 represent, respectively, active and inactive states. The network aims to recover a set of independent patterns $\{\boldsymbol{\eta}^\mu,\ \mu = 1, \ldots, P\}$ that have been stored by a learning process. That means, a stable retrieval state satisfies $\boldsymbol{\tau}^t = \boldsymbol{\eta}^\mu$, for large enough time t. Each pattern, $\boldsymbol{\eta}^\mu = \{\eta_i^\mu \in 0, 1; i = 1, \ldots, N\}$, is a set of site-independent low-activity binary random variables distributed according the probability

$$p(\eta_i^\mu = 1) = a,\ p(\eta_i^\mu = 0) = 1 - a. \tag{4.1}$$

A uniform neural model is a particular case of the present low-activity model with $a = 1/2$ in Eq. (4.1). The later case was studied in the previous Chapter about block structured information.

The synaptic couplings between the neurons i and j are given by the adjacency matrix

$$J_{ij} \equiv C_{ij} W_{ij}, \tag{4.2}$$

where the topology matrix $\boldsymbol{C} = \{C_{ij}\}$ describes the connectivity structure of the neural network and $\boldsymbol{W} = \{W_{ij}\}$ is the matrix with the learning weights. The topology matrix is split into local and random links. Local links connect a given neuron to its K_l neighbors, with periodic boundary conditions. Random links connect each neuron to K_r other neurons, uniformly distributed in the network. Hence the total number of connections, per neuron, is $K = K_l + K_r$. Given neurons i and j, the corresponding matrix element $C_{ij} = 1$ if $|i - j| \leq K_l/2$. If $|i - j| > K_l/2$, $C_{ij} = 1$ with probability K_r/N, and $C_{ij} = 0$ otherwise. This way, the network topology is characterized by two parameters: the *connectivity* ratio and the *randomness* ratio, respectively defined by

$$\gamma = K/N\,, \quad \omega = K_r/K\,. \tag{4.3}$$

Here, K_l, K_r and N are assumed to be large. As in the *small-world* model (Watts and Strogatz, 1998), the parameter ω plays the role of a rewiring probability. The storage cost of this network is $|\boldsymbol{J}| = N \times K$ if the matrix $\boldsymbol{J}$ is implemented as an adjacency list of K neighbors, and one would like to find out the behavior of the network in terms of these two parameters. An extremely diluted network is obtained as $\gamma \to 0$ and one restricts to this case.

The retrieval of a pattern is achieved through the noiseless neuron dynamics

$$\tau_i^{t+1} = \Theta(h_i^t - \theta_i^t)\,, \quad i = 1, \ldots, N\,, \tag{4.4}$$

where

$$h_i^t \equiv \frac{1}{K} \sum_j J_{ij} \frac{\tau_j^t - q_j^t}{\sqrt{Q_j^t}} \tag{4.5}$$

denotes the local field at neuron i and time t and θ_i is its threshold of firing. In Eq. (4.5) is introduced the average activity of the neighborhood of neuron i, $q_i^t = \langle \tau^t \rangle_i$, and its corresponding variance, $Q_i^t = Var(\tau^t)_i = \langle (\tau^t)^2 \rangle_i - \langle \tau^t \rangle_i^2$. The neighborhood average is defined as $\langle f^t \rangle_i \equiv \sum_j C_{ij} f_j^t / K$. In Eq. (4.4) is used the step function:

$$\Theta(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0. \end{cases} \tag{4.6}$$

The uniform binary neural model is recovered when $a = 1/2$ (Dominguez et al., 2009). Although the normalization in Eq. (4.5) is not usual in the literature, it is

appropriate in order to consider the complete symmetry between the representation of a pattern and its anti-pattern, despite their opposite mean activity. Also, the local dependence in this normalization takes into account the structure in the neural states that are being studied in the present Chapter.

For convenience, in the sequel one uses the normalized variables, the site and time dependence being implicit:

$$\sigma \equiv \frac{\tau - q}{\sqrt{Q}},\ q \equiv \langle \tau \rangle\ ,\ Q \equiv Var(\tau) = q(1-q) \tag{4.7}$$

$$\xi \equiv \frac{\eta - a}{\sqrt{A}},\ a \equiv \langle \eta \rangle\ ,\ A \equiv Var(\eta) = a(1-a), \tag{4.8}$$

where a and q are the pattern and neural activities, respectively. The averages done in this Chapter run over different ensembles, and are indicated in each case. These variables can be directly translated to those used in most works found in the literature for uniform (non-biased) neurons, in the case of $a = 1/2$.

In terms of these normalized variables, the neuron dynamics reads

$$\sigma_i^{t+1} = g(h_i^t - \theta_i^t, q_i^t), \tag{4.9}$$

$$h_i^t \equiv \frac{1}{K} \sum_j J_{ij} \sigma_j^t,\ i = 1, \ldots, N, \tag{4.10}$$

where the gain function is given by

$$g(x, y) \equiv [\Theta(x) - y] / \sqrt{y(1-y)}. \tag{4.11}$$

The weight matrix $\boldsymbol{W}$ is updated according to the Hebb's rule,

$$W_{ij}^{\mu} = W_{ij}^{\mu-1} + \xi_i^{\mu} \xi_j^{\mu}. \tag{4.12}$$

Weights start at $W_{ij}^0 = 0$ and after P learning steps, they reach the value $W_{ij} = \sum_{\mu}^{P} \xi_i^{\mu} \xi_j^{\mu}$. The learning stage displays slow dynamics, being stationary within the time scale of the faster retrieval stage, Eq. (4.5). A stochastic macro-dynamics take place due to the extensive learning of $P = \alpha K$ patterns, where α is the load ratio.

4.2.2 The information measures

In order to characterize the retrieval ability, one needs to define appropriate information measures. Previous studies have usually dealt with global measures of information, which are adequate to describe networks with no local connectivity. In these cases the overlap and neural activity,

$$m \equiv \frac{1}{N}\sum_{i}^{N}\xi_i\sigma_i,\ q \equiv \frac{1}{N}\sum_{i}^{N}\tau_i, \tag{4.13}$$

which are the statistical correlation between the learned pattern ξ_i and the neural state σ_i, and the mean activity of the network, respectively, are enough to evaluate the network ability to retrieve a given pattern. When the pattern is successfully retrieved, $\tau_i = \eta_i, i = 1, ..., N$, one has $m \sim 1$ and $q \sim a$. For metric connectivity, it is useful to define blocks as the structured pieces of information that emerge in the network. If the contiguous neurons are distributed within b blocks, for simplicity each of size $L = N/b$, then one may define mesoscopic parameters restricted to the λth block ($\lambda = 1, \ldots, b$).

The *block's overlap* between the neural states and one individual pattern and the *block's activity* are

$$m_\lambda \equiv \frac{1}{L}\sum_{i\in\lambda}\xi_i\sigma_i \ \text{ and } \ q_\lambda \equiv \frac{1}{L}\sum_{i\in\lambda}\tau_i, \tag{4.14}$$

at an unspecified time step, respectively. One can consider m_λ as a random variable and estimate the average of this variable across the blocks as $\langle f_\lambda\rangle_b \equiv \frac{1}{b}\sum_{\lambda=1}^{b} f_\lambda$. For example, a network with $b = 2$ blocks can be characterized by the block parameters m_λ, q_λ, where $\lambda = \pm$ are labels for the two distinct regions. The following states are typical: global retrieval state (R, with $m_+ \sim m_- \sim 1, q_+ \sim q_- \sim a$), which carries non-vanishing uniform information; block retrieval state (B, with $m_+ \sim 1, m_- \sim -1, q_+ \sim a, q_- \sim 1 - a$), which carries structured information distributed in positive and negative overlaps with the target pattern; bump retrieval state (U, with $m_+ \sim 1, m_- \sim 0, q_+ \sim q_- \sim a$ which carries localized information distributed in a positive overlap region and a disordered region, respectively. Furthermore, there is a zero state (Z, with $m_+ \sim m_- \sim 0, q_+ \sim q_- \sim a$), without any information.

In the general case of b blocks, it is worth to define macroscopic parameters instead of the mesoscopic measures m_λ, q_λ. The relevant order parameters measuring the quality of retrieval are the mean (m) and the variance (v) of the block overlap distribution, defined as

$$m \equiv \langle m_\lambda \rangle_b \quad \text{and} \quad v \equiv \langle m_\lambda^2 \rangle_b - m^2. \tag{4.15}$$

Thus, m is the usual *global* overlap, given in Eq. (4.13). When the global overlap is zero and the size of the blocks is taken as $L = 1$, the network carries no macroscopic order. On the other hand, if there is only one block, $L = N$, the variance is zero and the network carries no local information. However, if the size is mesoscopic, $1 \ll L < N$, the variance is finite and the blocks convey only local information. The standard deviation, which is named the *block* overlap, is $\delta = \sqrt{v}$. It is worth mentioning that the blocks are macroscopically scaled and hence the parameter δ is not related to a spin-glass, which is a microscopic ordering (Amit, 1989; Hertz et al., 1991). Similarly, the mean and variance of the block's activity are defined as

$$q \equiv \langle q_\lambda \rangle_b \quad \text{and} \quad v_q \equiv \langle q_\lambda^2 \rangle_b - q^2. \tag{4.16}$$

Hence q is the global activity, and the standard deviation, $\delta_q = \sqrt{v_q}$ is the block neural activity.

Together with the overlap, one is interested in the load ratio $\alpha \equiv P/K$, that accounts for the storage capacity of the network. As the number of stored patterns grows the network is not able to retrieve them and the overlap goes to zero.

4.2.3 Threshold strategies

In order to retrieve patterns with low activity, it is necessary to use an adequate threshold for firing. If firing is not controlled, the neural activity could be higher (lower) than the pattern activity, whenever the threshold is too small (large).

The more sparse is the code, more sensible is the interval where the threshold can move on. On the one hand, one could use an optimal threshold where, for each learned pattern and initial condition, a threshold is chosen by hand, such that the retrieval is maximized (Roudi and Treves, 2008). This is not a realistic

strategy, since the neural network is not supposed to know the target during the retrieval process. Thus, it is convenient to use a fixed value for the threshold. Indeed, one can search for a threshold which is a function of the load rate, the pattern activity, the randomness and the connectivity parameters.

In particular, one could think about a dynamical threshold, which is updated each retrieval time step. A self-control threshold that works for very low activities was proposed in the sparse-coding literature(Dominguez and Bollé, 1998). An efficient strategy would be even more important in the case of block retrieval. For such a solution, the threshold should vary according to the neighborhood of the neurons. Here, the following simple strategy is proposed:

$$\theta_i^t = \begin{cases} \theta_0^t, \ q_i^t < 0.5 \\ -\theta_0^t, \ q_i^t > 0.5\,. \end{cases} \tag{4.17}$$

More sophisticated strategies can consider a soft function $\theta_i^t(q_i^t)$, as well as dependences on the noise and the other network and learning parameters, ω, γ, α and a.

The value of the θ^0 can be easily estimated as a function of the sparseness by taking a look at the local field. For the global solution, according to the Eqs. (4.10) and (4.12),

$$h_i^t = \frac{1}{K} \sum_{\mu} \xi_i^\mu \sum_{j \in C_i} \xi_j^\mu \sigma_j^t. \tag{4.18}$$

For simplicity, considering a single pattern $P = 1$ and a global retrieval state for a given time step, $\sigma_j^t = \xi_j^\mu$, the last expression becomes $h_i^t = \xi_i^\mu$.

On the other hand, the overlap is $m = \langle \xi\sigma \rangle_{\sigma,\xi}$, where the brackets represent an average over the joint distribution $p(\sigma, \xi)$ for a single neuron, which is understood as an ensemble distribution for the neuron states $\{\sigma_i\}$ and patterns $\{\xi_i^\mu\}$. Averaging over ξ using Eq. (4.1) and Eq. (4.8) one gets

$$m = a\xi_+\sigma_+ + (1-a)\xi_-\sigma_-, \quad where \quad \xi_+ \equiv \frac{1-a}{\sqrt{A}} \quad and \quad \xi_- \equiv \frac{-a}{\sqrt{A}}\,. \tag{4.19}$$

Here, $\sigma_\pm$ is the neuron state for the active and inactive pattern nodes, $\xi_\pm$ respectively. That means: $\sigma_\pm \equiv g(h^\pm - \theta_0)$, where $h^+ = (1-a)/\sqrt{A}$ and $h^- = -a/\sqrt{A}$.

So, in order that $\sigma_j^t = \xi_j^\mu$ (hence $m^t = 1$) persists for every time t, the threshold must satisfy: $h^- < \theta_0 < h^+$, one could choose

$$\theta_0(a) = \frac{1-2a}{2\sqrt{A}}. \tag{4.20}$$

This expression is used for the threshold as a function of the pattern activity. For instance, $\theta_0(a = 0.1) \sim 1.33$, which is the middle of the interval between the active signal and inactive one. Adding a dependency on the global activity of the network, q, one could obtain a useful threshold for the block and bump phases, as follows:

$$\theta_0^t = \begin{cases} \rho\theta_0, \ q^t > \dfrac{a+0.5}{2} \\ \dfrac{1}{\rho}\theta_0, \ q^t < \dfrac{a+0.5}{2}. \end{cases} \tag{4.21}$$

The empirical value of $\rho = 0.7$ is used. For $a = 0.1$, one has $\theta_0^t = 0.7 \times 1.33 \approx 0.93$ for the block solution ($q \sim 0.5$), and $\theta_0^t \approx 1.9$ for the bump and global solution ($q \sim a$).

4.3 Results

4.3.1 Simulations: the retrieval evolution

The neural dynamics, Eqs. (4.4-4.5), was simulated with the topology defined according Eq. (4.3) and the learning rule given by Eq. (4.12). The results reveal that the network can be found in the following representative stationary states: global retrieval phase (R, with $m \neq 0$, $\delta = 0$), which carries non-vanishing uniform information; block retrieval phase (B, with $m = 0$, $\delta > 0$) which carries structured information retrieving pattern and anti-pattern spatially distributed over the network; bump retrieval phase (U, with $m \sim 0.5$, $\delta \sim 0.5$) which carries spatially distributed information retrieving a correlated part of the pattern besides an uncorrelated one. Furthermore, there is a zero phase (Z, with $m = 0$, $\delta = 0$), without any information. Representing the different phases, the typical evolution of a network for $N = 10^5$, with topological parameter $\gamma = 10^{-2}$, randomness $\omega = 0.1$ and pattern activity $a = 0.1$ is shown in Fig. 4.2. In each panel the site overlaps $\xi_i\sigma_i$ are smooth averaged over windows of size $\Delta_N = 10^3$ for

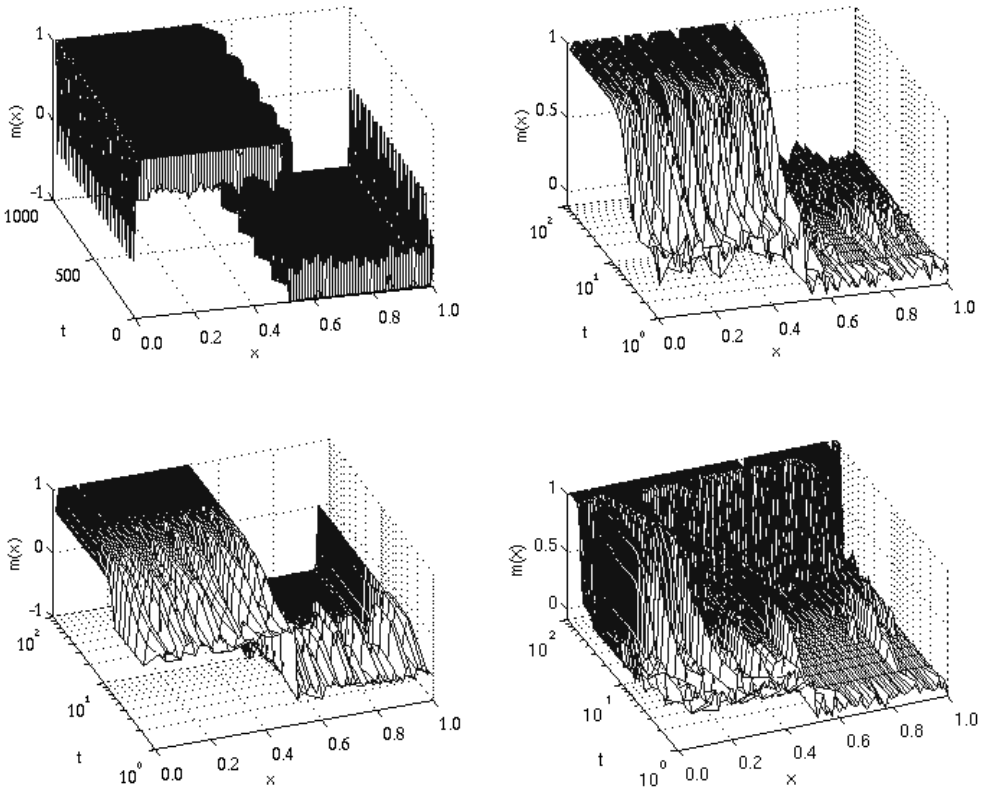


Figure 4.2: Time evolution of the microscopic overlaps. The parameters are $a = 0.1$, $N = 10^5, K = 10^3$ ($\gamma = 10^{-2}$), $\omega = 0.1$. Left panels: Block initial conditions with $\alpha = 0.01$; right panels: Bump initial conditions with $\alpha = 0.1$. Top panels: $\theta_0^t = 1.6$; bottom panels: $\theta_0^t = 1.0$.

every time t. Hence the plotted surfaces, $m^t(x)$, $x = i/N$, are softer than the actual site overlaps, although some structure can still be appreciated.

In the left panels, for $\alpha = 0.01$, the network starts in a noisy two-blocks spatial distribution ($m^0 \sim 0$, $\delta^0 = 0.4$). In the upper-left panel one sees that for a threshold value $\theta_0^t = 1.6$ the network evolves to a block structure that oscillates with period $\Delta t = 200$ at the borders of the positive and negative blocks. This is due to the fact that frontier effects on the local field of the neuron's neighborhood are amplified for large threshold values. Nevertheless, although the small fluctuations of the spatial structure for a large enough time, the block phase is a stationary stable state. In the lower-left panel, for a threshold value $\theta_0^t = 1.0$, It can be appreciated, in a logarithmic scale, how starting in a noisy two-blocks distribution the network evolves to an almost perfect block structure. That is, an optimized value of the threshold was translated into a more stable block phase. In this case, frontier effects are negligible, as compared to the upper-left panel.

In the right panels, the network starts in a noisy bump spatial distribution ($m^0, \delta^0 \sim 0.17$). The upper-right panel shows that, for a threshold value of $\theta = 1.6$, the network evolves to a bump phase. The part where the initial stimulus was correlated with a pattern, the pattern was completely retrieved (the errors were corrected), while the uncorrelated part remained quite stationary. This means that the bump phase is stationary in time. In the lower-right panel, for $\theta = 1.0$, the network evolves to a stationary global phase. On the contrary to the block structure seen in the left panels, a larger value of the threshold favors the stability of the bumps, because of the local field cannot reach large values for the uncorrelated region to surpass the threshold. For small threshold values, the signal part of the bump stimulus brings the whole network to a global phase, being unable to keep a spatially localized structure.

4.3.2 Simulations: the learning capacity

The stationary states of the network were studied in function of the load ratio α, for different values of the topological parameter ω. A sample of the simulation results is shown in Fig. 4.3, represented with symbols (circles and crosses). The

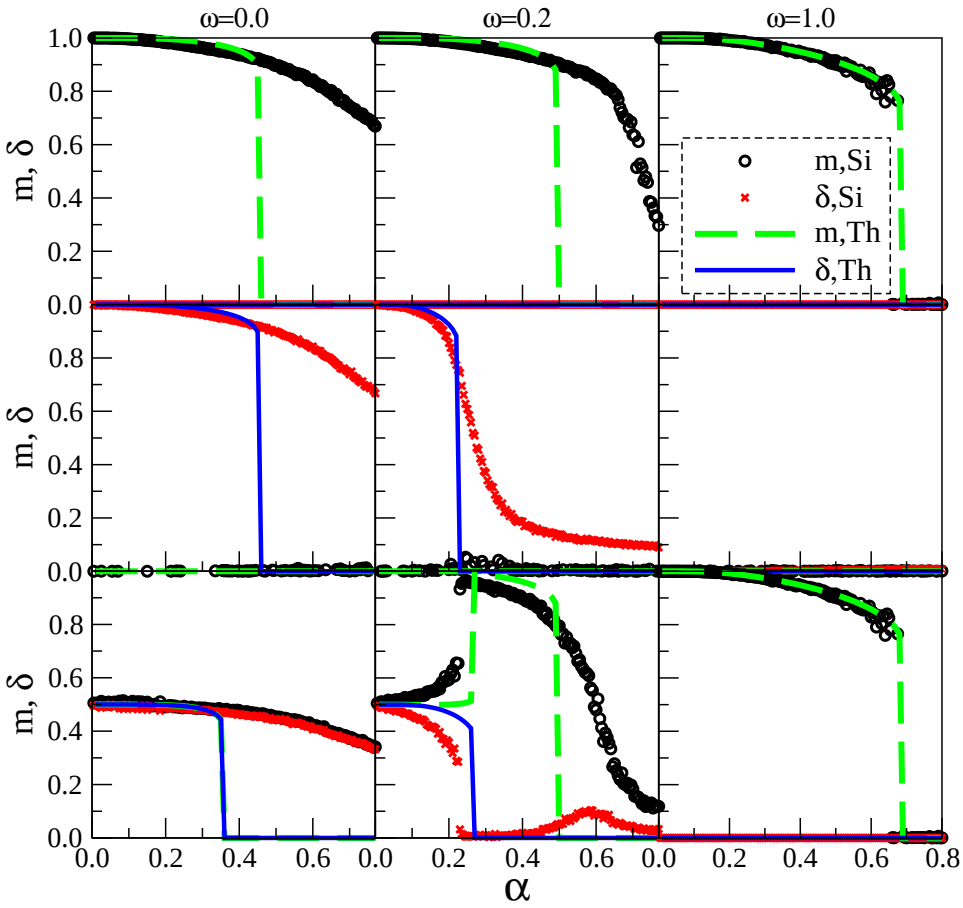


Figure 4.3: Global and Block overlaps, as a function of α for Theory and Simulation, with $a = 0.1$, $\omega = 0, 0.2, 1.0$, $\theta^0 = 1.3$. The network has $N = 10^6$, $K = 400$. Initial conditions: Global state (R, top panels), Block (B, middle panels), Bump (U, bottom panels). (Color on-line).

stationary global and local overlaps, m^* an δ^*, respectively, are plotted for $N = 10^6$ with connectivity $\gamma = 4 \cdot 10^{-4}$ and randomness values of $\omega = 0$, $\omega = 0.2$ and $\omega = 1.0$. The results from theory are plotted in the same figure, represented with lines (dashed and solid) and will be discussed later in section 4.3.4. For the top panels the network starts in a global configuration with $m^0 = 1$ and $\delta^0 = 0$. There is a clear distinction in which concerns the transitions from the global retrieval phase to the zero phase: they are smooth for the local networks ($\omega = 0$), while they are sharp for random networks ($\omega = 1.0$). The following criterion was used for the transition of the global to the zero phase: the network is in the zero phase either if $m < 0.8$ or the time of convergence exceeds $t = 200$ steps in the parallel neural updating, Eqs. (4.4-4.5). Similarly, the network is in the block phase if $\delta < 0.8$ and in the bump phase if $m \sim \delta < 0.4$.

It is observed that the critical load goes to $\alpha \sim 0.45$ for local networks. For an intermediate value $\omega = 0.2$, the critical load increases to $\alpha \sim 0.5$, while for random networks the critical load goes to a larger value of $\alpha \sim 0.7$. That is, when the information is spatially distributed (i.e. global phase) the optimal topology is a random network.

In the middle panels, the network starts in a two-blocks configuration with $m^0 = 0$ and $\delta^0 = 1$. For local networks ($\omega = 0$) the local overlap stays close to $\delta^* \sim 1$, up to the critical load $\alpha \sim 0.45$. For intermediate values of the random parameter, $\omega = 0.2$, is possible to sustain the blocks, although with a smaller critical load: $\alpha \sim 0.25$. Finally, for random networks, $\omega = 1.0$, is not possible to obtain a spatial configuration of blocks, and the network converges to a Z phase. Thus, one can see that the local character of the network allows the formation of spatially distributed structures, the shorter are the range of the connections, the more stable is the block phase.

In the bottom panels, the network starts in a single-bump configuration with $m^0 = 0.5$ and $\delta^0 = 0.5$. For local networks ($\omega = 0$) the bump is the only stable phase and the critical load goes to $\alpha \sim 0.35$. However, as the randomness network increases ($\omega = 0.2$), a re-entrance transition from bump to the more ordered global phase can be observed. Into the uncorrelated region, when the noise is small, the local field on each neuron is not sufficiently large to surpass the threshold value and, in consequence, the bump is stable. As the load increases, the noise becomes

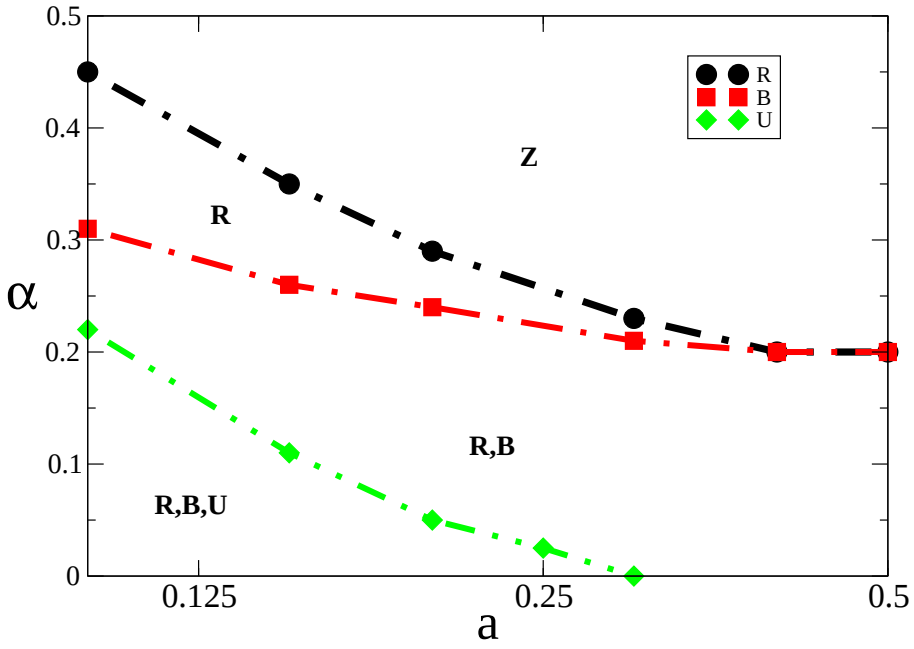


Figure 4.4: Phase diagram ($a \times \alpha$) for Simulation with $N = 2 \cdot 10^5$, $K = 2 \cdot 10^3$. The parameters are $\omega = 0.1$, and $b = 2$. $R \equiv$ Retrieval phase ($m \sim 1, \delta \sim 0$). $B \equiv$ Block phase ($m \sim 0, \delta \sim 1$). $U \equiv$ Bump phase ($m \approx 0.5, \delta \approx 0.5$). $Z \equiv$ Zero phase ($m \sim 0, \delta \sim 0$).

larger. The local field fluctuates between the pattern and the anti-pattern, but the only effect is to attract the uncorrelated region to a global order stable state. For random networks ($\omega = 1.0$), there is no sense for structured neural configurations because no spatial preference is generated by uniform connectivity.

4.3.3 Phase diagram

The phase diagram shown in Fig. 4.4 was obtained by simulation of a network with $N = 2{\cdot}10^5$, $K = 2{\cdot}10^3$ and randomness parameter $\omega = 0.1$. The results were obtained by computer simulations of the dynamics defined by Eqs. (4.4-4.5), with the topology defined in Eq. (4.3) and the learning rule given by Eq. (4.12). The dashed curves are simply guide for the eyes. The diagram displays the different regions of retrieval: Bump phase (U), Block phase (B) and Global phase (R). The non retrieval Zero phase (Z) is also shown. The regions are represented for

different values of the load ratio of learned patterns, α, against the sparseness, a. The criterion used to estimate the phase transitions are described in the previous section.

When the network starts in a bump configuration with $m^0 = 0.5$ and $\delta^0 = 0.5$, the bump phase survives for a load ratio that is lower than that of the bump retrieval saturation, $\alpha_U(\omega)$ curve. When the network starts in a block configuration with $m^0 = 0$ and $\delta^0 = 1$, this block phase survives if the load is lower than that of the block retrieval saturation, the $\alpha_B(\omega)$ curve. Finally, if the network starts in a global configuration with $m^0 = 1$ and $\delta^0 = 0$, it remains there if the load ratio is lower than that of the global retrieval saturation, the $\alpha_R(\omega)$ curve. Above the $\alpha_R(\omega)$ curve, no information is retrieved and the only achievable phase is Z, which carries no information. Below the diamond symbol curve $\alpha_U(\omega)$, and for sparseness values smaller than $a = 0.3$, the phases U, B, R coexist. What the phase will be effectively reached through the network dynamics depends on the initial conditions. Increasing the sparseness value, the Bump phase ceases, with the B and R phases coexisting below the squared symbol curve $\alpha_U(\omega)$. Increasing the value of the load ratio, the B phase is no longer stable above the curve $\alpha_B(\omega)$. Above this curve, the only phase supporting information is the global retrieval R. A further increase of the load ratio yields a transition to the zero information phase Z, at the circle symbol curve $\alpha_R(\omega)$.

The robustness of this behavior was checked by using different threshold strategies than that expressed in Eqs. (4.17,4.20,4.21). It was found that the corresponding phase diagrams are qualitatively similar to that shown in Fig. 4.4, whenever a reinforcement mechanism for the threshold depending on the local activity is employed.

4.3.4 Theory

In this section, a strongly-diluted network is considered and, consequently, theoretical equations for the macroscopic order parameters are proposed. The sketch of a proof that is valid for stationary equations can be found in the appendix.

In order to carry out the theoretical calculations, mesoscopic and macroscopic order parameters were defined. The mesoscopic order parameters are the block

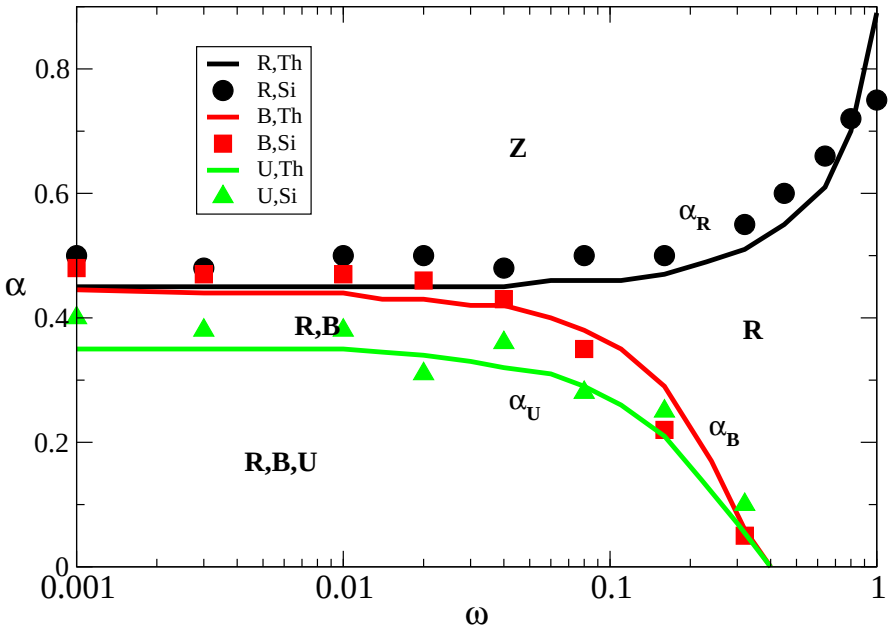


Figure 4.5: Phase diagram ($\omega \times \alpha$) for Theory (Lines) with $\gamma = 4 \cdot 10^{-4}$ and for Simulation (Symbols) with $N = 10^6$, $K = 400$. The parameters are $a = 0.1$, $\theta = 1.6$ and $b = 2$). $R \equiv$ Retrieval phase ($m \sim 1, \delta \sim 0$). $B \equiv$ Block phase ($m \sim 0, \delta \sim 1$). $U \equiv$ Bump phase ($m \approx 0.5, \delta \approx 0.5$). $Z \equiv$ Zero phase ($m \sim 0, \delta \sim 0$).

overlap m_λ and the block activity q_λ. They are described in Eq. (4.29), where the average is over the sites inside the blocks, according to Eq. (4.30). After successively separating in terms of signal/noise, local/random connections and global/block overlaps, one arrived at Eqs. (4.41-4.42) for the local field and the local activity, h_i, q_i respectively. The macroscopic parameters are the global overlap and the block overlap, m, δ respectively, defined by the ansatz in Eq. (4.31). Their recursive equations were obtained after averaging m_λ over y_λ, according Eqs. (4.43-4.44). Furthermore, one can average explicitly over the distribution of $y_\lambda \in \pm 1$ and $\xi \in \{(1-a)/\sqrt{A}, -a/\sqrt{A}\}$, obtaining

$$m = a\frac{1-a}{2\sqrt{A}}\langle \mathrm{g}(h_+^+ \quad -\theta, q_+) + \mathrm{g}(h_-^+ - \theta, q_-)\rangle_z \\ +(1-a)\frac{-a}{2\sqrt{A}}\langle \mathrm{g}(h_+^- - \theta, q_+) + \mathrm{g}(h_-^- - \theta, q_-)\rangle_z \quad (4.22)$$

and

$$\delta = a\frac{1-a}{2\sqrt{A}}\langle \mathrm{g}(h_+^+ \quad -\theta, q_+) - \mathrm{g}(h_-^+ - \theta, q_-)\rangle_z \\ +(1-a)\frac{-a}{2\sqrt{A}}\langle \mathrm{g}(h_+^- - \theta, q_+) - \mathrm{g}(h_-^- - \theta, q_-)\rangle_z \,. \quad (4.23)$$

Similarly, the equations for global and block activities defined in Eq. (4.32) are, according to Eqs. (4.45-4.46),

$$q = \frac{a}{2}\langle \mathrm{g}(h_+^+ - \theta, q_+) \quad +\mathrm{g}(h_-^+ - \theta, q_-)\rangle_z + \\ \frac{1-a}{2}\langle \mathrm{g}(h_+^- - \theta, q_+) + \mathrm{g}(h_-^- - \theta, q_-)\rangle_z \quad (4.24)$$

and

$$\delta_q = \frac{a}{2}\langle \mathrm{g}(h_+^+ - \theta, q_+) \quad -\mathrm{g}(h_-^+ - \theta, q_-)\rangle_z + \\ \frac{1-a}{2}\langle \mathrm{g}(h_+^- - \theta, q_+) - \mathrm{g}(h_-^- - \theta, q_-)\rangle_z \,. \quad (4.25)$$

The equations above use the definitions

$$h_\pm^+ \equiv \frac{1-a}{\sqrt{A}}[\omega m + (1-\omega)(m \pm \delta)(1-\gamma b)] + z\sqrt{\alpha r}\,, \\ h_\pm^- \equiv \frac{-a}{\sqrt{A}}[\omega m + (1-\omega)(m \pm \delta)(1-\gamma b)] + z\sqrt{\alpha r}\,. \quad (4.26)$$

The angular brackets indicate average over the noise variable z, that follows a Gaussian distribution: $\langle f(z)\rangle_z \equiv \int_{-\infty}^{\infty}(dz/\sqrt{2\pi})\exp(-z^2/2)f(z)$. See ref. Hertz et al. (1991). The noise variance αr is given by the sum of random and local *feedback* terms,

$$r = \omega r_r + (1-\omega)r_l \quad \text{and} \quad r_x = (1-\chi_x)^{-2}, \quad \text{with} \quad x = l, r\,. \tag{4.27}$$

The susceptibility for random connections is $\chi_r = 0$. For local connections, after averaging over both y and ξ, the susceptibility is given by

$$\chi_l = \frac{a}{2\sqrt{\alpha r_l}}\langle z[\mathrm{g}(h_+^+ \quad -\theta, q_+) + \mathrm{g}(h_-^+ - \theta, q_-)]\rangle_z + \frac{1-a}{2\sqrt{\alpha r_l}}\langle z[\mathrm{g}(h_+^- - \theta, q_+) + \mathrm{g}(h_-^- - \theta, q_-)]\rangle_z\,. \tag{4.28}$$

The numerical system of Eqs. (4.22)-(4.28), for the macroscopic parameters, is solved in function of the load α. The different solutions R, B and U are obtained through appropriate initial conditions. The results are plotted with solid (δ) and dashed (m) lines, in Fig. 4.3, together with simulation results represented by symbols, for randomness $\omega = 0$, $\omega = 0.2$ and $\omega = 1.0$. This procedure can be extended for several values of ω, in order to obtain the theoretical phase diagram $\alpha \times \omega$. The results were drawn in Fig. 4.5. There, the theoretical results (solid curves) can be compared to that obtained from numerical simulation (symbols). Fixed-point solutions which correspond to unstable stationary states were disregarded. A network with $N = 10^6$, $K = 400$ and sparseness value of $a = 0.1$ was used for the simulations. The criterion used to estimate the phase transitions with the simulation were described in the section 4.3.2.

For small values of ω, all the retrieval R, B and U phases are stable below $\alpha_U(\omega)$. Increasing the load ratio above $\alpha_U(\omega)$, the U phase becomes unstable, and the phases R and B phases coexist for small value of randomness parameter ω. For a further increase in ω, the B phase loses stability, and the only stable retrieval phase is R. Above the line $\alpha_R(\omega)$, no information is retrieved and Z is the only possible state.

As plotted in Figs. 4.3 and 4.5, simulation and theory are in qualitatively agreement for global, block and bump critical loads. There are some quantitative differences between the shape of the transitions, that will be discussed in the next section.

4.4 Conclusions and discussion

The main subject of this Chapter was the retrieval of spatially ordered information in an attractor neural network with sparse coding. The retrieval phases are, namely, Global retrieval (R) with uniform overlaps, Block retrieval (B) with structured overlaps, and Bump retrieval (U) with localized overlaps. The states R, B and U have a region of the network that is correlated with a given pattern; however, the other region is also correlated with the same pattern in the R phase, correlated with the anti-pattern in the B phase, and uncorrelated at all in the U phase. The dependence of the stability of these retrieval states was analyzed according the topological parameters connectivity (γ) and randomness (ω). The block and bump structures arise as stable solutions of the network dynamics for local connectivity ($\omega \ll 1$).

The dependence of these phases on the activity of the patterns (a) was also investigated. For large values of activity ($a > 0.3$), bumps are not possible, because the correlated region of the stimulus (whose neuron states at the initial time step overlaps with the pattern) attracts the whole network to a global overlap configuration. On the other hand, for patterns with low levels of activity, the correlated region is unable to rise the local fields of the uncorrelated region above the threshold value, and the bump is stable for a small load ratio, see Fig. 4.4. However, if the load ratio is large enough, the fluctuations in the local field increase and the neurons can achieve the correct sign (that is, correlated with the pattern), leading to the global phase R, as can be seen in the bottom-center panel of Fig. 4.3, for $\alpha > 0.25$.

A threshold strategy, that benefits the retrieval of both global (R) and spatially structured (B and U) information states, allowing the network dynamics to stabilize such configurations according to the topological parameters, was introduced. This threshold strategy depends mainly on the level of activity of the patterns and the global activity of the network, as well as the local activity of the neighborhood of each neuron. This translates into a self-control mechanism of the network dynamics according to the initial stimulus (Dominguez and Bollé, 1998).

It worths to stress that the phase transitions, which are sharp for the theoretical results due to the thermodynamic limits $N \to \infty$, $Ka \to \infty$, $K \ll \ln(N)$, become smooth for the simulation results, because these limits can not be fulfilled, even with expensive computer memory resources. Simulation and theory results agree quite well in the regions where the R, B and U phases have a well defined values of the parameters m, δ. They disagree after the region of the transition defined for the theory, where the relaxation time is very large, and the network has not yet reached the convergence for the maximal times we have used in the simulation ($t < 400$ parallel steps). The finite size effects introduced in the simulations correspond, first, to the finiteness $N < \infty$. Next to the partial dilution $\gamma > 0$. Finally, all the finite size effects are amplified by the effective network dimension Ka which decreases with the low-activity. For instance, one used $N = 10^6, K = 400, a = 0.1$ in Fig. 4.3, which implies $Ka = 40$, i.e., far from the infinite limit. The apparent smooth transition is a reminiscent of a well known fact for small-world networks, close to the regular lattice behavior, with finite connectivity K and large dimension N (Newman and Watts, 1999). In this limit, the typical length scale diverges with vanishing short-cut density $\omega \to 0$, leading to a continuous transition.

4.5 Appendix: Macro-dynamic equations

In the formulation of the theory one restricts to the asymptotic limits $N \to \infty$, $Ka \to \infty$, with $K \ll \ln(N)$. Let the neurons be randomly distributed within b blocks and, for simplicity, each of size $L = N/b$, with positive and negative overlaps, $m_\lambda = m_\pm$. The blocks $\lambda = 1, \ldots, b$ are built as the sets $\mathbf{\Lambda} = \{i = (\lambda - 1)L + k; k = 1, \ldots, L\}$. The order parameters described in Eq. (4.14) are

$$m_\lambda \equiv \langle \xi\sigma \rangle_\lambda \quad \text{and} \quad q_\lambda \equiv \langle \tau \rangle_\lambda \,, \tag{4.29}$$

where the average is over the sites inside the blocks:

$$\langle f \rangle_\lambda = \frac{1}{L} \sum_{i \in \lambda} f_i. \tag{4.30}$$

Then, according to Eq. (4.15), the global overlap is $m = \sum_\lambda^b m_\lambda/b$ and the fluctuation between blocks is $\delta = \sqrt{v}$, with $v \equiv \sum_\lambda^b m_\lambda^2/b - m^2$. The block's overlap can be written as a function of the mean overlap m, and the block overlap δ, as

$$m_\lambda = m + y_\lambda \delta, \tag{4.31}$$

where $y_\lambda \doteq \pm 1$, (according to the block) is a random variable. By analogy, according to Eq. (4.16), the block's neural activity can be written as a function of the mean neural activity q and the block neural activity δ_q as

$$q_\lambda = q + y_\lambda \delta_q. \tag{4.32}$$

The dynamics of the order parameters is given by the updating of the neuron states, Eqs. (4.9)-(4.10). The local field of the neuron τ_i, applying Eq. (4.12) for the weights, reads

$$h_i^t = \frac{1}{K} \sum_\mu \xi_i^\mu \sum_{j \in C_i} \xi_j^\mu \sigma_j^t. \tag{4.33}$$

Now, one proceeds to split the local field in three steps: first the signal/noise separation; then the signal is separated in terms coming from the local and random connections; finally the local term is expressed as a function of the global and block overlaps. The noise term is considered at the end.

4.5.1 Signal/Noise

If a given pattern is being retrieved, say $\xi \equiv \xi^{\mu=1}$, the local field can be separated into a signal term and a noise term,

$$h_i \equiv \xi_i m_i + \Omega_i \,, \tag{4.34}$$

where

$$m_i^\mu = \langle \xi_j^\mu \sigma_j \rangle_C \,, \quad \text{with} \quad \langle f \rangle_C \equiv \frac{1}{K} \sum_{j \in C} f_j \,, \tag{4.35}$$

is the *graph-overlap*, which is the overlap restricted to the neighborhood C of neuron σ_i, and

$$\Omega_i \equiv \sum_{\mu > 1} \xi_i^\mu m_i^\mu \tag{4.36}$$

is the cross-talk noise due to the extensive number of learned patterns.

4.5.2 Local/Random

The graph-overlap for the retrieved pattern is the signal term which depends on the connectivity matrix C. There are local and random neighbors for each neuron. Hence, the signal term itself splits into localized and randomized terms, namely $m_i = \frac{K_r}{K} m_i^r + \frac{K_l}{K} m_i^l$, with

$$m_i^x \equiv \frac{1}{K_x} \sum_{j \in C_x} \xi_j \sigma_j, \ (x \equiv l, r) \tag{4.37}$$

where K_l^i and K_r^i are the local and random sets of neighbors of the neuron σ_i, respectively. Using the definitions of the topological parameters in Eq. (4.3), it can be written as

$$m_i = \omega m_i^r + (1 - \omega) m_i^l. \tag{4.38}$$

4.5.3 Global/Block

Taking a close look at the graph-overlap, Eq. (4.37), one rewrites it in function of the block overlaps. On the one hand, the random neighbors yields an average over neurons which are a uniform sample along the network. Thus it does not depend on a block, and results in a global mean $m_i^r = m$, up to fluctuations of order $1/\sqrt{N}$. Now, consider the local connections: whenever a neuron is in the bulk of a given block λ, its neighbors belong to the same block, $j \in \Lambda_\lambda$. So, for the local interactions, the graph-overlap is the same as the block's overlap, $m_i^l = m_\lambda$. Following Eq. (4.31),

$$m_i^l = (m + y_\lambda \delta)(1 - \gamma b), \tag{4.39}$$

up to fluctuations of order $1/\sqrt{Ka}$. The correction factor $(1 - \gamma b)$ accounts for the boundary effects between $m_\pm$ blocks. Finally, one obtain an expression for the graph-overlap:

$$m_i = \omega m + (1 - \omega)(m + y_\lambda \delta)(1 - \gamma b)\,. \tag{4.40}$$

Using these results, together with Eqs. (4.34) and (4.35), one arrives to an approximation for the local field of neurons in the block λ,

$$h_i \equiv \xi[\omega m + (1 - \omega)(m + y_\lambda \delta)(1 - \gamma b)] + \Omega\,, \tag{4.41}$$

where the noise Ω depends on the $(p-1)$ patterns which are not being retrieved. It also depends on the topology, and on the neural states. Similarly, one has an equation for the local neural activity

$$q_i \equiv \omega q + (1-\omega)(q - y_\lambda \delta_q). \tag{4.42}$$

Note that the sign of y_λ for the block's neural activity is opposite to the block's overlap sign, since the convention is that $y_\lambda = -1$ for the anti-pattern, which have the larger activity.

The equation for the block's overlap in Eq. (4.29) can be written as $m_\lambda = \langle \xi \mathrm{g}(h_i - \theta, q_i) \rangle_{\eta,z}$, where the average in the angular brackets are over the noise variable z and the pattern distribution η. But from Eq. (4.29) $m_\lambda = m + y_\lambda \delta$ and thus, after averaging over y_λ, one gets

$$m = \langle m_\lambda \rangle_y = \langle \langle \xi \mathrm{g}(h_i - \theta, q_i) \rangle_{\eta,y} \rangle_z \,, \tag{4.43}$$

$$\delta = \langle y m_\lambda \rangle_y = \langle \langle y \xi \mathrm{g}(h_i - \theta, q_i) \rangle_{\eta,y} \rangle_z \,. \tag{4.44}$$

Similarly, the block's activity can be written as $q_\lambda = \langle \mathrm{g}(h_i - \theta, q_i) \rangle_{\eta,\Omega}$, and the equations for the mean and block neural activity reads

$$q = \langle q_\lambda \rangle_y = \langle \langle \mathrm{g}(h_i - \theta, q_i) \rangle_{\eta,y} \rangle_z \,, \tag{4.45}$$

$$\delta_q = \langle y l_\lambda \rangle_y = \langle \langle y \mathrm{g}(h_i - \theta, q_i) \rangle_{\eta,y} \rangle_z . \tag{4.46}$$

4.5.4 Noise/Feedback

This noise is a large sum of almost-independent terms, which converges to a Gaussian distribution,

$$\Omega \sim z\sqrt{\Delta} \,, \quad z \doteq N(0,1) \,, \quad \Delta = \alpha r \,, \tag{4.47}$$

see Hertz et al. (1991). Its variance Δ is given by the sum of random and local *feedback* terms,

$$r = KVar(m_i^\mu) = \omega r_r + (1-\omega) r_l \,. \tag{4.48}$$

To deal with them, one can consider the residual overlaps ($\mu > 1$) as stochastic variables. For the local interactions the neighbors neurons are in the same block,

so one can replace $l^{\nu}_{j\in C} \sim l^{\nu}_{i}$. That means, all residual overlaps have the same probability distribution. If one expands the residual overlaps around $h^{\nu}_{j} \equiv h_j - \xi^{\nu}_{j} m^{\nu}_{j}$ it holds:

$$m^{\nu}_{i} \sim \frac{1}{K}\sum_{j\in C} \xi^{\nu}_{j} \mathrm{g}(h^{\nu}_{j} - \theta, q_j) + m^{\nu}_{i}\chi_i \,, \tag{4.49}$$

$$\chi_i \equiv \frac{1}{K}\sum_{j\in C} (\xi^{\nu}_{j})^2 \frac{d}{dh^{\nu}_{j}} \mathrm{g}(h^{\nu}_{j} - \theta, q_j) \,. \tag{4.50}$$

Here χ_i is the *susceptibility*. The first term in the r.h.s of Eq. (4.31) is not correlated with the second term, and its variance is

$$Var[\frac{1}{K}\sum_{j\in C} \xi^{\nu}_{j} \mathrm{g}(h^{\nu}_{j} - \theta, q_j)] = \alpha \,. \tag{4.51}$$

Thus, the stochastic equation reads $m^{\mu}_{i}(1-\chi) \simeq N[0,\alpha]$, and the feedback term is

$$r_x = (1-\chi_x)^{-2}; \; x = l, r. \tag{4.52}$$

Now supposing only strongly diluted networks ($K \ll N$). For random connections, χ_r can be neglected since there is no feedback in the dynamics and $r_r = 1$ (Dominguez and Bollé, 1998). However, for local connections, even extreme dilution does not eliminate the feedback, thus the susceptibility can also be written as

$$\chi_l = \frac{1}{\sqrt{\alpha r_l}} \langle z \; \mathrm{g}(h_\lambda - \theta, q_\lambda) \rangle_{y,z} \,, \tag{4.53}$$

where the averages are over the block distribution y and over the normalized Gaussian variable z.

Chapter 5

Structured patterns of variable-activity objects

5.1 Introduction

In the previous Chapters 2 to 4, the possibility of sustaining spatially distributed information in the network has been studied. That is, the output of the network was characterized according to the initial conditions of blocks and bumps structures. However the learned inputs were uniformly distributed. In the present Chapter one is interested in the input learned by the network, being that this input is spatially structured in regions of different levels of activity called *objects*. The learning and retrieval of patterns arranged in variable-activity objects is explored, and the impact of the metric connectivity on the performance of the Attractor Neural Network (ANN). In Fig. 5.1 is depicted a schematic representation of what one calls structured patterns of variable activity objects (left panels) in opposition to uniform coding (right panels) with no spatial structure, which is used in most ANN works. In Fig. 5.1 top-left panel is presented a structured of ten contiguous objects, one half with high local activity ($a_+ = 0.9$) and the other half with low local activity ($a_- = 0.1$). In the left-bottom panel the mean activity of each region is smooth averaged for sub-groups of neurons, thus some of the structure can be appreciated. In the top-right panel a uniform coding pattern is presented, where the global mean activity of the pattern is 0.5 as depicted in the bottom-right panel. Note that for the spatially structured pattern, averaging

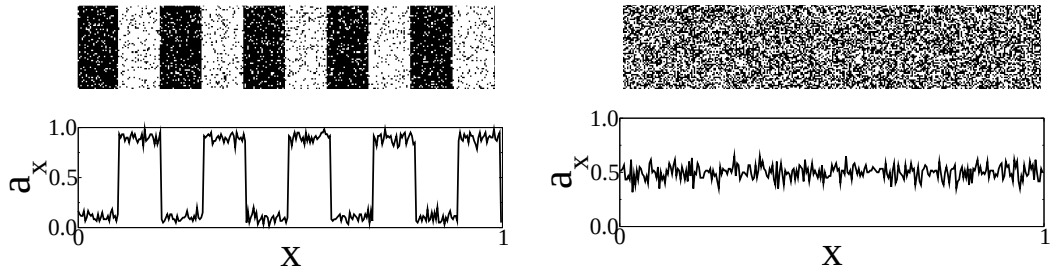


Figure 5.1: Left: Variable-activity structured pattern. Right: Uniform activity pattern.

over all the objects in the bottom-left panel, one obtains the same value of 0.5 for the global activity of the whole pattern.

A potential implication of this Chapter is the increment of the critical load α_c, that is, the number of patterns the network is able to store, given that the pattern structure resembles that of the sparse-coding pattern (antipattern) where the critical load is higher, compared with the uniform coding network (Amit, 1989; Hertz et al., 1991). However, considering the nature of the coding used, the correlation between patterns due to the potential overlap between objects of similar activity is a factor to take into account (Hertz et al., 1991). Also, an adequate threshold is needed in order to control the distinct levels of activity of the different objects along the network (Dominguez et al., 2012). Different threshold strategies are proposed, the correlation between the patterns analyzed, and the topological conditions that sustain the spatial structure of variable activity objects in the network are identified.

5.2 The Model

5.2.1 Neural coding

At any given discrete time t, the network state is defined by a set of N binary variables $\boldsymbol{\tau}^t = \{\tau_i^t \in 0, 1; i = 1, \ldots, N\}$, where 1 and 0 represent, respectively, active and inactive states. The network aims to recover a set of patterns $\{\boldsymbol{\eta}^\mu,\ \mu = 1, \ldots, P\}$ that have been stored by a learning process. That means, a stable

retrieval state satisfies $\boldsymbol{\tau}^t = \boldsymbol{\eta}^\mu$, for large enough time t. Each pattern, $\boldsymbol{\eta}^\mu = \{\eta_i^\mu \in 0, 1; i = 1, \ldots, N\}$, is a set of binary random variables distributed in objects (λ) of contiguous neurons (Dominguez et al., 2009; González et al., 2009) according the probability

$$\begin{aligned} p(\eta_i^\mu = 1) = a_-, i \in \lambda_- \\ p(\eta_i^\mu = 1) = a_+ \in \lambda_+. \end{aligned} \tag{5.1}$$

being $\lambda_\pm$ the labels for the low and high activity objects respectively, where $a_- \in (0, 0.5]$ and $a_+ = 1 - a_-$, stand for the corresponding average activity ratio of each object.

Let the contiguous neurons be distributed within β objects, for simplicity each of size $L = N/\beta$, where $\lambda = 1, \ldots, \beta$. If the number of high activity and λ_+ low activity λ_- objects are equal, the patterns are unbiased (i.e. the activity of the whole pattern is 1/2), however they are not uniform, but distributed in objects of distinct activity. The objects are distributed randomly helping to avoid the overlap between them. For $\beta = N$, that is, objects of size $L = 1$, the model presented in Chapter 3, with uniform and unbiased patterns with a global activity $a = 1/2$, is obtained. For $\beta = 1$, the model presented in Chapter 4 with uniform biased patterns with sparseness $a = a_-$, is obtained .

5.2.2 Topology and dynamics

In this Chapter the same model used in the previous Chapter 4 is employed, except for the coding of the input patterns learned which is presented in the previous section 5.2.1. The synaptic couplings between the neurons i and j are given by

$$J_{ij} \equiv C_{ij} W_{ij}, \tag{5.2}$$

where the topology matrix $\boldsymbol{C} = \{C_{ij}\}$ describes the connectivity structure of the neural network and $\boldsymbol{W} = \{W_{ij}\}$ is the matrix with the learning weights defined according to the Eq. (4.12). The network topology is characterized by the usual parameters: the *connectivity* ratio $\gamma = K/N$ and the *randomness* ratio $\omega = K_r/K$. The total number of connections $K = K_l + K_r$ splits between local connections to K_l neighbors and K_r random connections to other neurons

uniformly distributed in the network. The network dynamics, defined by Eqs. (4.4-4.11), is also valid in this case.

5.2.3 Information measures

The information measures employed in the present Chapter are the same presented in the previous Chapter in Eqs. (4.13-4.16). In the last Chapters the important measures were the mean (m) and the variance (v) of the block overlap distribution, defined in Eq. (4.15). In this Chapter, one is interested in the global retrieval of the structured input, then the important parameters to measure the retrieval of the network are the macroscopic global overlap m and global activity q defined in Eq. (5.6), as well as, the mesoscopic parameters for each object m_λ, q_λ in Eq. (4.14) that describe the inner structure sustained in the network. For convenience are presented here again.

One needs to define the following normalized variable, the site and time dependence being implicit:

$$\sigma \equiv \frac{\tau - q}{\sqrt{Q}},\ q \equiv \langle \tau \rangle\ ,\ Q \equiv Var(\tau) = q(1-q) \tag{5.3}$$

$$\xi \equiv \frac{\eta - a}{\sqrt{A}},\ a \equiv \langle \eta \rangle\ ,\ A \equiv Var(\eta) = a(1-a), \tag{5.4}$$

where a and q are the pattern and neural activities, respectively.

The mesoscopic order parameters are defined as

$$m_\lambda \equiv \frac{1}{L}\sum_{i\in\lambda} \xi_i \sigma_i \ \text{ and } \ q_\lambda \equiv \frac{1}{L}\sum_{i\in\lambda} \tau_i. \tag{5.5}$$

The macroscopic order parameters are defined as

$$m \equiv \langle m_\lambda \rangle_b \quad \text{and} \quad q \equiv \langle q_\lambda \rangle_b. \tag{5.6}$$

5.2.4 Pattern correlation

One can define the correlation g^μ between pattern μ and the other $P-1$ patterns as

$$g^{\mu} = \frac{1}{P-1} \sum_{\nu=1(\nu\neq\mu)}^{P} g^{\mu\nu}, \quad g^{\mu,\nu} = \frac{1}{N} \sum_{i}^{N} \xi_i^{\mu} \xi_i^{\nu}. \tag{5.7}$$

A set of $P = 100$ patterns of size $N = 10^5$ was generated, structured in objects with low activity $a_- = 0.1$ and high activity $a_+ = 1 - a_- = 0.9$. The mean correlation g^{μ} was measured between each pattern μ with the other $P-1$ patterns for different numbers of β objects. In Fig. 5.2 one can see that the correlation g^{μ} is approximately zero for random uniform patters $\beta = N$. As the number of objects β decreases, the correlation g^{μ} increases. Thus, one needs to take into account the correlation that occurs from the overlapped objects with same levels of activity between patterns, since this correlation worsens the performance of the network because the cross-talk term is high in this case.

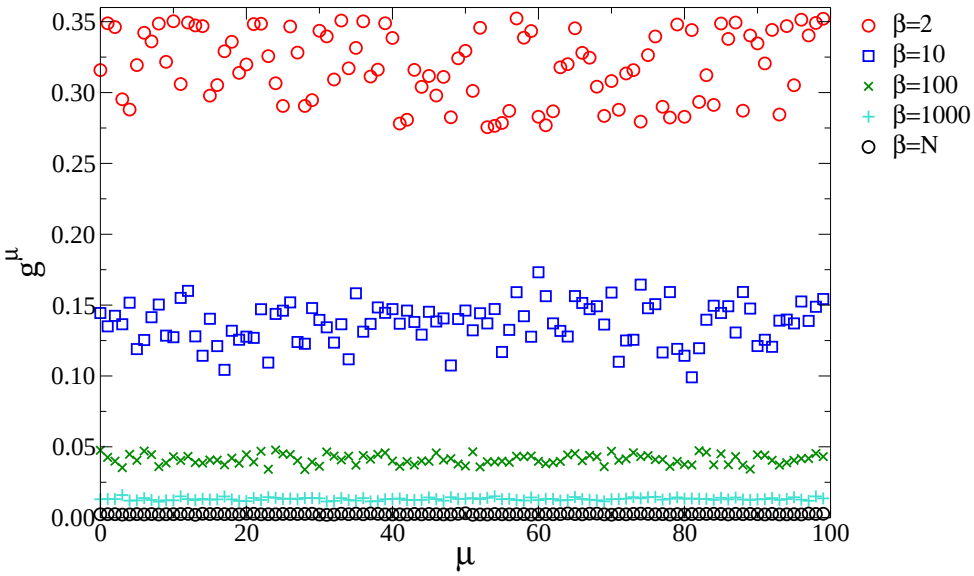


Figure 5.2: Mean correlation between patterns g^{μ} as the number of objects β increases. $P = 100$ patterns of size $N = 10^5$ with $a_- = 0.1$.

5.2.5 Threshold strategies

In order to retrieve patterns with objects of low and high activity, an adequate threshold for firing is necessary. One needs to define average activity q_i of the

neighborhood of neuron i, $q_i^t = \langle \tau^t \rangle_i$. The neighborhood average is defined as $\langle f^t \rangle_i \equiv \sum_j C_{ij} f_j^t / K$. Here are proposed the following threshold strategies.

- The *ρ-strategy* ($\theta^{t,r}$) as proposed in the last Chapter 4 under section 4.2.3, Eqs. (4.17-4.21) is tested in this Chapter.

- A simple trigonometric function strategy $\theta^{t,s}$:

$$\theta_i^{t,s} = \theta_0 sin(2\pi q_i). \tag{5.8}$$

- Square-cut function strategy $\theta^{t,c}$:

$$\theta_i^{t,c} = \begin{cases} \theta_0^t, \ q_i^t < 0.5 \\ -\theta_0^t, \ q_i^t > 0.5 \, , \end{cases} \tag{5.9}$$

  with $\theta_i^{t,c} = 0$ for the local activity values q_i below a_-, as well as for the values of q_i above $1 - a_-$.

- Linear-cut function strategy $\theta^{t,l}$:

$$\begin{aligned} \theta_i^{t,l} &= s(q_i - a_-) + \theta_0, \\ s &= \frac{2\theta_0}{2a_- - 1}. \end{aligned} \tag{5.10}$$

  As in the former strategy (Eq. 5.9) one uses $\theta_i^{t,l} = 0$ for the local activity values q_i below the a_-, as well as for the values of q_i above $1 - a_-$.

The following empirical values are used: $\theta_0(a_- = 0.1) = 0.3$, $\theta_0(a_- = 0.2) \sim 0.3$, $\theta_0(a_- = 0.3) \sim 0.2$, $\theta_0(a_- = 0.4) \sim 0.1$.

5.3 Results

5.3.1 Simulations: retrieval performance and threshold strategy

The dynamical neurons Eqs. (4.4)-(4.5) were simulated with the topology parameters defined according to Eq. (4.3). In Fig. 5.3 is shown the retrieval performance of the network for the different threshold strategies presented in Section 5.2.5. A

network with $N = 10^5, K = 10^2$ for a randomness parameter $\omega = 0.0$ was used, with $\beta = 100$ objects with $a_- = 0.1$ and a value of $\theta_0 = 0.3$. The performance is very similar for all threshold strategies, being slightly better for the trigonometric one $\theta^{t,c}$. The system's sensitivity to the threshold is better carried out using a smooth function.

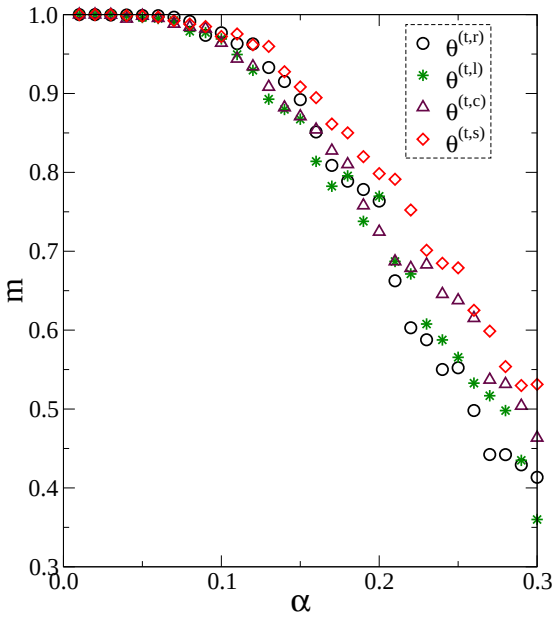


Figure 5.3: Retrieval overlap m as a function of the load ratio α for different threshold strategies. Network with $N = 10^5, K = 10^2, \omega = 0.0$ with $a_- = 0.1, \beta = 100$.

5.3.2 Simulations: retrieval performance and number of objects

A network with $N = 10^5, K = 10^2, \omega = 1.0$ was simulated, for different number of β objects with $a_- = 0.1$. the trigonometric threshold strategy $\theta^{t,s}$ with $\theta_0 = 0.3$ was used, see Eq. (5.8). In Fig. 5.4 is shown the retrieval performance as the number of objects increases. It can be appreciated that when increasing the number of β objects the performance of the network is improved because the correlation between patterns decreases. However the network is still far from its

optimal retrieval performance that corresponds to $\beta = N$ which is equivalent to the model presented in Chapter 3. Here the correlation between patterns is seriously affecting the storage capacity of the network. A usual solution is to use the pseudo-inverse approach (Hertz et al., 1991) to orthogonalize the correlated patterns as studied in the next Chapter 6, where automotive traffic video is stored and retrieved.

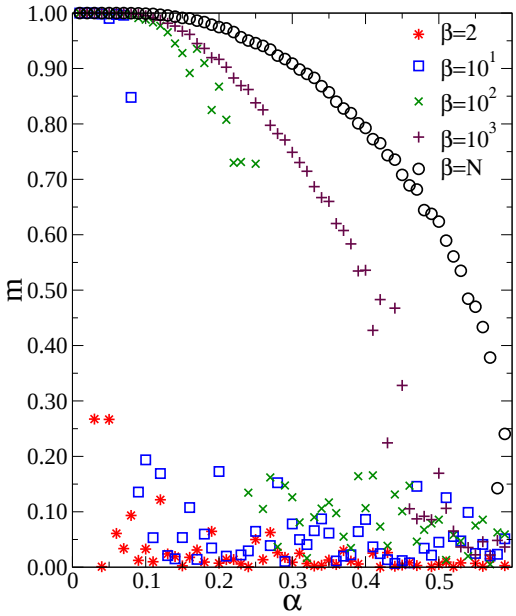


Figure 5.4: Retrieval overlap m as a function of the load ratio α as the number of objects β increases. Network with $N = 10^5, K = 10^2, \omega = 1.0$ with $a_- = 0.1$ and $\theta^{t,s}$ with $\theta_0 = 0.3$.

5.3.3 Retrieval performance and network randomness parameter

Using a network with $N = 10^5, K = 10^2$ and for $\beta = 100$ objects, with $a_- = 0.1$ and threshold strategy $\theta^{t,s}$ with $\theta_0 = 0.3$, the retrieval performance of the network is compared for different values of the randomness parameter. In Fig. 5.5 can be appreciated that the network performs better for high values of the parameter ω. Except for the pathological case of the regular network with $\omega = 0$ where the

results obtained are similar to the case of the totally random network $\omega = 1$. This can be due to the fact that in networks with only local connections, the nodes are highly influenced by their neighborhood activity, trapping the corresponding spatial structure so that the different objects are independently sustained. When adding some random shortcuts, that is for intermediate values of ω the network can not sustain the distinct levels of activity as the locality is lost, making impossible to sustain the different objects. Increasing the number of random shortcuts enough the network manage to give an overall response retrieving the global pattern while sustaining the different regions of activity increasing again its performance. This results need to be checked with more extensive simulations.

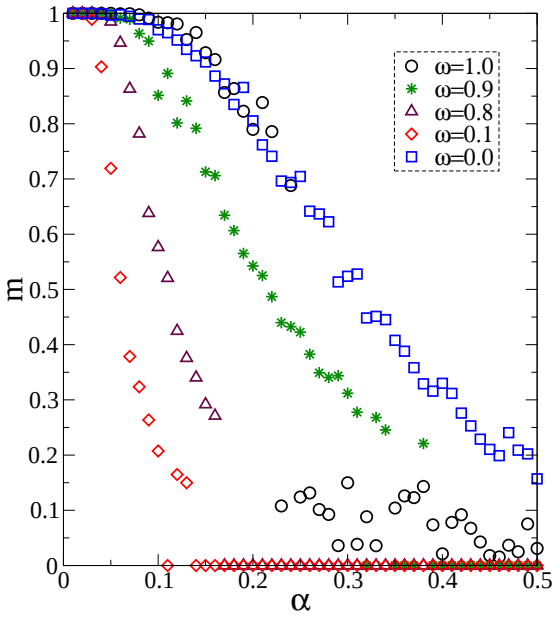


Figure 5.5: Retrieval overlap m as a function of the load ratio α for different values of the parameter ω. Network with $N = 10^5, K = 10^2$ with $\beta = 100, a_- = 0.1, \theta^{t,s}, \theta_0 = 0.3$.

5.3.4 Simulations: the retrieval evolution

One lets a network with $N = 10^5, K = 10^2$, evolve as presented in Fig. 5.6, for a load ratio $\alpha = 0.1$. The trigonometric threshold strategy of $\theta^{t,s}$ with $\theta_0 = 0.3$ is used. In the top panels the site overlaps $\xi_i\sigma_i$ (top-left) and the site activities

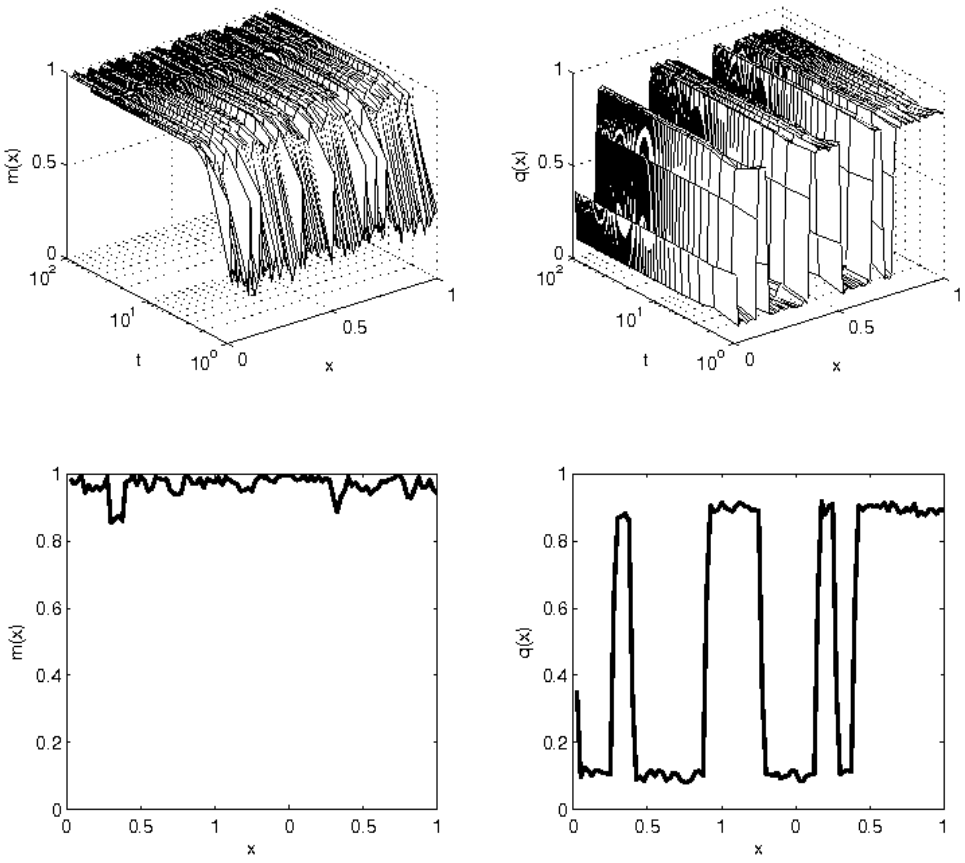


Figure 5.6: Top panels: evolution of the network microscopic overlaps of the objects. Network with $N = 10^5, K = 10^2, \omega = 1.0$ using $\beta = 20, a_- = 0.1$, initial condition $m^0 = 0.3$, $\alpha = 0.1$, threshold strategy $\theta^{t,s}$ with $\theta_0 = 0.3$, $x \equiv i/N$.

τ_i (top-right) are smooth averaged over windows of size $\Delta_N = 10^2$ for every time t. Hence the plotted surfaces, $m^t(x)$ and $q^t(x)$, $x = i/N$, are softer than the actual site overlaps and site activities respectively. However, some structure can still be appreciated. In the top-left panel, the network starts in a noisy spatial distribution of $\beta = 20$ objects with $a_- = 0.1$ and initial condition $m^0 = 0.3$. The network retrieve the target pattern with a value of $m^* = 0.97$ and the objects of low activity with q_- and high activity q_+ are sustained along the network, as seen in the top-right panel. It can be appreciated, in a logarithmic scale, how starting in a noisy configuration the network evolves to a global state where the target pattern is almost perfectly retrieved as shown in the bottom-left panel, for the smoothed site overlaps for the final time step $t = 100$. The bottom-right panel shows the smoothed site activities at the final time step, where the structure of the objects can be better appreciated.

5.4 Conclusions and discussion

In this Chapter is studied a pattern coding structured in objects of variable activity. It is explored how the topological parameters and a self-control mechanism presented as a dynamical threshold cooperate to sustain the structure of low and high activity objects. Further threshold strategies need to be developed, in order to improve the performance of the network for retrieving this type of variable activity coding. The threshold strategy has to be able to distinguish local from global activities in order to properly apply the control mechanism for the network dynamics. The correlation between patterns is relevant for this type of coding and needs to be addressed in order to improve the network performance. This has been considered in the next Chapter 6, where a variation of the pseudo-inverse approach is used to store and retrieve automotive traffic video using an attractor network.

The results presented in this Chapter are preliminary and need to be contrasted with more extensive simulations and a theoretical analysis. However, it is worth a further study of this type of structured coding for conjectures of information processing in biological neural systems, and also for real world implementations (González et al., 2011).

5. STRUCTURED PATTERNS OF VARIABLE-ACTIVITY OBJECTS

Chapter 6

An application for learning automotive video

6.1 Introduction

In the last Chapter 5 was suggested for the input patterns, a coding structured in objects with different levels of activity. In this Chapter an application of Attractor Neural Network for storing and retrieving automotive traffic video is suggested. The input patterns, the frames of the video, are spatially structured in objects e.g. cars, streets, sidewalks, etc., with different levels of activity. The information is localized within the frames in a similar way it is in the spatial structures studied so far in this thesis. In the present Chapter a real world implementation using structured patterns is presented for traffic surveillance. The application of video-based analysis to traffic surveillance (Kastrinaki et al., 2003) is an area of growing interest with the aim to detect both global events (i.e. number of vehicles in a road region) and local events (i.e. detection and tracking of a specific vehicle). As large amounts of video data are stored for analyzing the involved events on them, it becomes very important to develop efficient storage and retrieval techniques for these traffic videos. In general, these videos are sequences of frames where the involved patterns (i.e. moving vehicles) are highly correlated in time, specially in traffic congestion scenes. Most of existing works for this problem use an approach based on scene segmentation followed by vehicle tracking (Chan and Vasconcelos, 2005). In it, the vehicles are first detected in

the dynamic scene using adaptive-background techniques (Jung and Ho, 2001; Kastrinaki et al., 2003) and specific features like texture, color or shape (Chan and Vasconcelos, 2007), are extracted from the segmented targets for classification. Later, these vehicles are tracked using different techniques like optical flow (Li and Chellappa, 2002), Kalman filters (Yu et al., 2003) or particle filters (Ristic et al., 2004), among others. Segmentation and tracking tasks become more difficult on realistic traffic situations like possible vehicle congestions, variability of weather and/or illumination conditions. Moreover, the vehicle tracking results along time are highly dependent on a good segmentation of them. To avoid the need of segmentation and tracking individual vehicles, some holistic representations for the storage and retrieval of traffic videos have been proposed. Chan and Vasconcelos (2005) propose a dynamic texture representation to model the motion flow in the scene. They use the Kullback-Leibler divergence and the Martin distance to retrieve and classify traffic videos without need of segmentation. Xie et al. (2004) present another holistic method for traffic video retrieval using Hierarchical Self-Organizing Maps (HSOM). They extract the motion trajectories of the vehicles present in the video and these activity patterns are stored by the neural network, later this learned knowledge is combined with a semantic indexing stage to retrieve traffic sequences based on queries by keywords.

The aim of this Chapter is to learn and retrieve a sequence of patterns that are highly correlated over time, obtained from a traffic video sequence. A Hopfield-type of Attractor Neural Network (ANN) with a small-world connectivity distribution is used. In order to achieve this objective one must face some typical problems found in the literature on ANNs (Dominguez et al., 2007; Koroutchev and Koroutcheva, 2006). First, in real-world applications, such as video compression/retrieval, where patterns present high correlation, one has to deal with sparse coding patterns. Sparse-coding is the representation of items by the *strong activation* of a relatively small set of neurons (Olshausen and Field, 2004). This is a different subset of all available neurons when the patterns are uncorrelated. On the one hand, this sparse-coding gives the model a biological plausibility since the brain suggests a general sparse-coding strategy. Sparse-coding is also favorable to increase the network capacity, because the cross-talk term between stored patterns decreases. On the other hand, it is difficult to sustain a low rate

of activity in ANNs and a control mechanism must be used (Dominguez and Bollé, 1998). Second, the application suggested requires learning a sequence of time-correlated patterns. The noise induced by the overlap between patterns is much larger for correlated patterns than for random patterns (Hertz et al., 1991). This implies that the network capacity drops down to an asymptotically vanishing value. Correlations between the training patterns, as it happens for a video sequence, worsens the performance of the network since the cross-talk term can yield high values in this case (Trappenberg, 2002).

The contribution of this Chapter is twofold. First, a variant of the pseudo-inverse approach is introduced to learn/retrieve a sequence of correlated cyclic patterns (as is the case for a video sequence) using a sparse-coding ANN with a small-world topology. Second, to demonstrate the feasibility of this approach for the storage and retrieval of traffic videos where the frames are codified as variable-activity patterns as studied in the previous Chapter. The proposed model for this problem avoids the segmentation and tracking of the involved targets and also some closely related difficulties.

6.2 Proposed model

6.2.1 Neural coding

A network with N neurons and a fixed number of $K < N$ synaptic connections per neuron is considered. At any given discrete time t, the network state is defined by the set of N independent binary neurons $\vec{\tau}^t = \{\tau_i^t \in [0,1]\,; i = 0, \ldots, N-1\}$, each one active or inactive denoted respectively by the state 1 or 0. The aim of the network is to retrieve a sequence of correlated patterns (in this case, the consecutive frames of the video sequence) $\{\vec{\eta}^\mu,\ \mu = 1, \ldots, P\}$ that have been stored during a learning process. Each pattern $\vec{\eta}^\mu = \{\eta_i^\mu \in [0,1]\,; i = 1, \ldots, N\}$ is a set of biased binary variables with sparseness probability:

$$p(\eta_i^\mu = 1) = a^\mu, \qquad p(\eta_i^\mu = 0) = 1 - a^\mu. \tag{6.1}$$

The mean activity for each pattern μ is $a^\mu = \sum_i^N \eta_i^\mu / N \equiv \langle \eta^\mu \rangle$. The neural activity for any time t is given by the mean: $q^t = \sum_i^N \tau_i^t / N \equiv \langle \tau^t \rangle$.

6.2.2 Network topology

The synaptic couplings between the neurons i and j are given by the adjacency matrix $J_{ij} \equiv C_{ij}W_{ij}$, where the topology matrix $\mathbf{C} = \{C_{ij} \in [0,1]\}$ describes the connection structure of the neural network and $\mathbf{W} = \{W_{ij}\}$ is the matrix of learning weights. The topology matrix contains two types of links: the local and the random ones, respectively. The local links connect each neuron to its K_l nearest neighbors in a closed ring, while the random links connect each neuron to K_r others uniformly distributed in the network. Hence, the network degree is $K = K_l + K_r$. The network topology is then characterized by two parameters, the *connectivity ratio* γ and the *randomness ratio* ω, which are respectively defined by:

$$\gamma = K/N, \qquad \omega = K_r/K, \tag{6.2}$$

where ω plays the role of a rewiring probability in the *small-world* model (Watts and Strogatz, 1998).

The storage cost of this network is $|\mathbf{J}| = N \times K$ if the matrix $\mathbf{J}$ is implemented as an adjacency list, where all neurons have K neighbors.

6.2.3 Retrieval dynamics

The task of the network is to retrieve the whole learned sequence of patterns (i.e., the full video sequence) starting from an initial neuron state $\vec{\tau}^0$ which is a given seed frame or a state close to it. The retrieval is achieved through the noiseless neuron dynamics:

$$\tau_i^{t+1} = \Theta(h_i^t - \theta_i^t), \tag{6.3}$$

$$h_i^t \equiv \frac{1}{K}\sum_j J_{ij}\frac{\tau_j^t - q_j^t}{\sqrt{Q_j^t}}, \; i = 1, \ldots, N, \tag{6.4}$$

where h_i^t denotes the local field at neuron i and time t, and θ_i is its firing threshold. The local mean neural activity is $q_i^t = \langle \tau^t \rangle_i$, and its variance is $Q_i^t = Var(\tau^t)_i$. The local mean is given by spatial averaging: $\langle f^t \rangle_i \equiv \sum_j C_{ij} f_j^t / N = \sum_{k \in C_i} f_k^t / K$,

for any given function f of the neuron sites. Here the step function is used:

$$\Theta(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0. \end{cases} \tag{6.5}$$

For convenience, in the Chapter the normalized variables are used, where the site and time dependence are implicit:

$$\sigma \equiv \frac{\tau - q}{\sqrt{Q}}, \; q \equiv \langle \tau \rangle, \; Q \equiv Var(\tau) = q(1-q) \tag{6.6}$$

$$\xi \equiv \frac{\eta - a}{\sqrt{A}}, \; a \equiv \langle \eta \rangle, \; A \equiv Var(\eta) = a(1-a), \tag{6.7}$$

where a and q are the pattern and neural activities, respectively. The averages computed in this Chapter run over different ensembles, and are indicated in each case.

6.2.4 Learning dynamics

To state the proposed learning rule for storing cyclic patterns which are highly correlated, as is the case for a video sequence, one will recall the expression of the weights for the standard case (static and uncorrelated patterns), and then two straightforward extensions: static and correlated patterns, and cyclic and uncorrelated patterns. Cyclic patterns correspond to sequences of patterns of variable activities, with periodic conditions (C. Molter and Bersini, 2005), that means, the next to the last pattern is the first one, then $\xi^{\mu+P} = \xi^{\mu}$.

If the network learns a set $P = \alpha K$ of static and uncorrelated patterns, $\langle \xi^{\mu} \xi^{\nu} \rangle = 0$, these are stored by the network couplings W_{ij} using the classical Hebbian rule (Amit, 1989) for the Hopfield model:

$$W_{ij} = \frac{1}{N} \sum_{\mu=1}^{P} \xi_i^{\mu} \xi_j^{\mu}. \tag{6.8}$$

This rule for learning the weights can be generalized introducing a $P \times P$ matrix $A_{\mu\nu}$ in the following way:

$$W_{ij} = \frac{1}{N} \sum_{\mu,\nu=1}^{P} \xi_i^{\mu} A_{\mu\nu} \xi_j^{\nu}. \tag{6.9}$$

The standard case, given by Eq. (6.8), is obtained by using an identity matrix $A^I_{\mu\nu} = \delta_{\mu\nu}$.

For the situation of learning static and correlated patterns, the pseudo-inverse approach (Hertz et al., 1991) is a standard method to orthogonalize (i.e. to extract) the correlated patterns, and the matrix $A_{\mu\nu}$ is computed as follows:

$$A^C_{\mu\nu} = O^{-1}_{\mu\nu}, \quad O_{\mu\nu} \equiv \frac{1}{N}\sum_i^N \xi_i^\mu \xi_i^\nu, \tag{6.10}$$

where O is the $P \times P$ patterns overlap matrix.

For the case of learning cyclic (sequential with periodic conditions) and uncorrelated patterns the former Hebbian rule, Eq. (6.8), combined with a row-shifting schema of the identity matrix can be applied (Amit, 1989):

$$A^S_{\mu+1,\nu} = \delta_{\mu\nu}, \ \forall \mu \in [1,..,P-1], \quad A_{1,\nu} = \delta_{P,\nu} \ \forall \nu \in [1,..,P], \tag{6.11}$$

In the case of video sequences, one has cycles (or sequences of patterns) where there is a high temporal correlation between the successive frames. For this reason, a heuristic is proposed where the learning weights are computed by combining the pseudo-inverse approach with a row-shifting schema, as the one used for cyclic patterns. The proposed heuristic for this case (i.e. cycles and correlated patterns) has the following four steps:

1. Obtain the pattern overlap matrix O.
2. Compute its inverse matrix O^{-1}.
3. Rotate forwards cyclically the rows of O^{-1} to obtain a new matrix M.
4. Substitute matrix A by the new matrix M in Eq. (6.9) to compute the weights matrix W for the video sequence to be learned.

The previous stages are detailed next. First, the $P \times P$ overlap matrix O, describing the video sequence is computed by Eq. (6.10), and its inverse matrix O^{-1} is obtained next. This approach is thought to get fixed point solutions. However, if one is seeking a limit cycle solution (i.e. retrieving the whole sequence

of frames cyclically), then one must benefit from the interactions between one frame and the next one in the video. Therefore, the elements of the O^{-1} matrix are shifted as shown schematically in the following equations:

$$A^V_{\mu+1,\nu} = O^{-1}_{\mu\nu},\ \mu \in [1,..,P-1], \qquad A^V_{1,\nu} = O^{-1}_{P,\nu},\ \forall \nu \in [1,..,P], \tag{6.12}$$

obtaining the matrix A^V. The previous rule takes into account the dominant terms in the infra-diagonal positions of the matrix A^V. The sub-dominant terms account for the orthogonalization of the matrix O^{-1}. It is worth to note that the pseudo-inverse rule is a not local matrix, because the connections between every two neurons depend on the other neurons; it is also a non iterative rule, all patterns must be learned at the beginning of the retrieval process.

The learned weight matrix $\mathbf{W}$ is now calculated according to the rule in Eq. (6.9), where $A_{\mu\nu}$ is computed by applying the row-shifting schema given by Eq. (6.12). The learning stage displays slow dynamics, being stationary within the time scale of the faster retrieval stage, as shown by Eq. (6.3). A stochastic macro-dynamics takes place due to the extensive learning of $P = \alpha K$ patterns, where α is the load ratio.

6.2.5 Threshold strategies

In order to retrieve patterns with low activity, is necessary to use an adequate threshold of firing. If firing is not controlled, the neural activity could be higher (lower) than the pattern activity, whenever the threshold is too small (large).

The more sparse the code is, the more sensitive is the interval where the threshold can move (Dominguez and Bollé, 1998). On the one hand, one could use an optimal manually-chosen threshold, where for each learned pattern and initial condition, the retrieval is maximized. This is not a realistic strategy, since the neural network is not supposed to know the patterns during the retrieval process. Thus, a simple and convenient solution is to use a fixed value for the threshold. The value of $\theta_i = 1$ for the threshold was obtained experimentally for a sparseness ratio of $a \sim 0.1$, which is the mean sparseness of the frames in the analyzed videos.

6.2.6 The information measures

In order to evaluate the network retrieval performance, two measures are considered: the global overlap and the load ratio. The overlap is used as a temporal measure of information, which is adequate to describe instantaneously the network ability to retrieve each frame of the video. In this case, the overlap m_μ^t between the neural state σ^t at time t and the frame ξ^μ is:

$$m_\mu^t \equiv \frac{1}{N}\sum_i^N \xi_i^\mu \sigma_i^t, \tag{6.13}$$

which is the normalized statistical correlation between the learned frame η_i^μ and the neural state τ_i^t at a given iteration t in the sequence cycle. One lets the network evolve according to Eqs. (6.3) and (6.4), and measures the overlap between the network states and the video frames running over a whole sequence cycle of the learned video. The neural states $\{\vec{\tau}^t,\ t = 1, \ldots, P\}$ are compared cyclically with the learned frames $\{\vec{\eta}^\mu,\ \mu = 1, \ldots, P\}$. The network starts in an initial condition close to a given frame, say $\tau^{t=1} \sim \eta^{\mu=1}$, so that the time and frame label are synchronized, and the overlap for each frame at cycle $c = 0, 1, 2...$ is:

$$m_\mu^c \equiv \frac{1}{N}\sum_i^N \xi_i^\mu \sigma_i^{\mu+cP}. \tag{6.14}$$

The global overlap is defined as:

$$m^c = \langle m_\mu^c \rangle \equiv \frac{1}{P}\sum_{\mu=1}^P m_\mu^c \tag{6.15}$$

and it measures the network ability to retrieve the whole sequence of patterns. After a transient period of time, the network dynamics converges to a stationary regime where the global overlap m^c does not change in the next cycles. When this global overlap between the whole set of patterns (i.e. the video sequence) and their corresponding neural states is $m = 1$, the network has retrieved the complete sequence without noise. In this case, all the network states correspond perfectly to the frames of the video. When the global overlap m is zero, the network carries no macroscopic order. In this case, the video can not be retrieved. For intermediates values of m, where $0 < m < 1$, the video can be partially recovered

with a given level of noise (when m increases, a higher number of frames can be perfectly retrieved).

One is also interested in the load ratio $\alpha \equiv P/K$, that accounts for the storage capacity of the network. This ratio depends on the size of the video, which is $P \times N$ (i.e. the number of frames by their spatial resolution, where this resolution coincides with the number of neurons), and the amount of physical memory necessary to store the video, which is $K \times N$ representing the adjacency lists sizes (see the network topology subsection).

When the number of stored patterns increases, the noise due to interference between patterns also increases and the network is not able to retrieve them. Thus, the overlap m goes to zero. A good trade-off between a negligible noise (i.e when $1 - m \sim 0$) and a large video sequence (i.e. a high value of α) is desirable for any practical-purpose model.

6.3 Experimental evaluation

The learning times to store the traffic video sequences were very high for the network considered. In the experiments, this time was highly dependent on the parameter K, as well as the number of learned patterns P, and it varies between 100 min and near 2000 min depending on the network degree considered. In fact the learning time is of order $O(N \times K \times P^2)$, according to Eq. (6.12). Two video sequences for the experiments have been used: the first one, *Kiev*, corresponds to a densely transited crossroad zone in Kiev, Ukraine; and the second one, *roundabout*, corresponds to a roundabout area in a Spanish city. Different model parameter configurations were tested for both sequences to get more insight on how the network behaved during the learning and retrieval of correlated cyclic frames. The *Kiev* video sequence was captured by a live camera demo site from Axis company:
http://www.axis.com/es/solutions/video/gallery.htm.
It was recorded by an Axis Q1755 Network Camera as an AVI video and consisted of 1835 frames at 25 frames per second, that is 73.4 seconds of recording. The original *roundabout* video sequence consisted of about 15 min of video which was

recorded with a conventional camera at 30 fps with frames and only 650 frames have been used, that is 21.7 seconds of video for the experiments.

For the two analyzed sequences, the video pre-processing included three stages:

(1) The frames of the initial color video sequence were converted into binary patterns and stored as PNG images with dimension 384×356 black-and-white pixels for the *Kiev* sequence and 640×480 pixels for the *roundabout* sequence.

(2) The *Kiev* frames were resized to $96 \times 89 = 8,544$ pixels and the *roundabout* frames to 80×106 pixels, in order to get a reasonable network size for the simulations.

(3) A new subsequence of frames was created by uniformly sub-sampling the sequence obtained in the previous stage using a natural factor f, where $f \geq 1$ (i.e., the video subsequence is built with the original frames: 1, $1+f$, $1+2f$, ...). The goal is to ensure that the network is able to recover the whole stored sequence of frames. Consequently, one starts testing with f=1, then f=2, and so on, until the condition holds.

For the simulations a system with an Intel Core 2 Duo CPU E6750 at 2.66GHz and with 2GB of physical memory has been used. The Octave image package (Octave, 2011) was used for processing the image files into text files with the 0/1 binary format as the neuron states required. The network parameters used in the *Kiev* simulations were $N = 8544, K = 4250, \theta_i = 1.0$ for a sparseness $a = 0.10$. For this network size, the video sequence has been recovered each $f = 5$ frames, that is: $\frac{1835}{5} = 367$ frames. For the *roundabout* simulations a similar network were used with $N = 8480, K = 4240, \theta_i = 1.0$ for a sparseness $a = 0.07$, recovering the video sequence each $f = 5$ frames, that is: $\frac{650}{5} = 130$ frames. The video output comparing the original with the retrieved frames and the frames in text format can be found at: http://dl.dropbox.com/u/11890025/video5.zip for the *Kiev* sequence and at http://dl.dropbox.com/u/11890025/roundabout.zip for the *roundabout* sequence.

Fig. 6.1 and Fig. 6.2 show some sample post-processed frames of the stored and successfully retrieved video sequences for the *Kiev* and *roundabout* sequences,

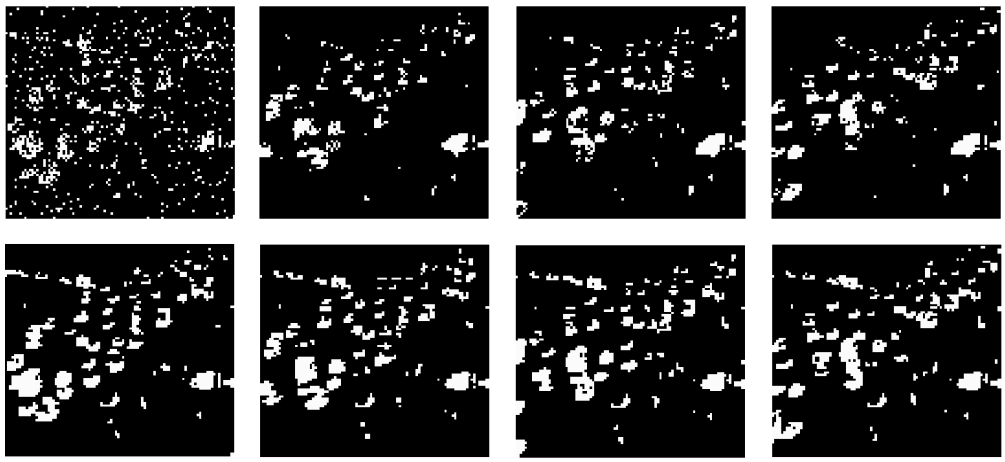


Figure 6.1: Some retrieved sample frames (from left to right, frame numbers 1, 21, 41 and 61) of the Kiev crossroad traffic video sequence for $f = 5$. Initial overlap $m^1 = 0.5$. Top panels: first cycle. Bottom panels: second cycle.

respectively. In Fig. 6.1 the seed used to start the retrieval was a noisy frame (top-left panel), with initial overlap $m_{\mu=1}^{c=0} = 0.5$. During the first cycle, the network is correcting the wrong pixels, (frame numbers 1, 21, 41 and 61 are presented in the top panels) $m^{c=0} \sim 0.93$, see Fig. 6.3-left. After a complete cycle the overlap reaches the stationary value of $m^{c=1} \sim 0.99$ (the same frames are shown in the bottom panels for the second cycle).

For the *roundabout* sequence in Fig. 6.2 the seed was a noisy frame (top-left panel), with initial overlap $m_{\mu=1}^{c=0} = 0.4$. The frame numbers 1, 11, 21 and 31 are presented in the top panels for $m^{c=0} \sim 0.97$, and bottom panels for the second cycle with a stationary value of $m^{c=1} \sim 0.98$, see Fig. 6.3-right.

6.3.1 Influence of the topology on the global overlap and the learning time

Using the previous network parameter setting (N,K,θ_i,a), Table 6.1 shows the dependence of global overlap and processing time on the random connections parameter ω at the learning stage.

As it can be observed, there is no significant difference between the processing time for learning the video with different values of ω and m parameters. This

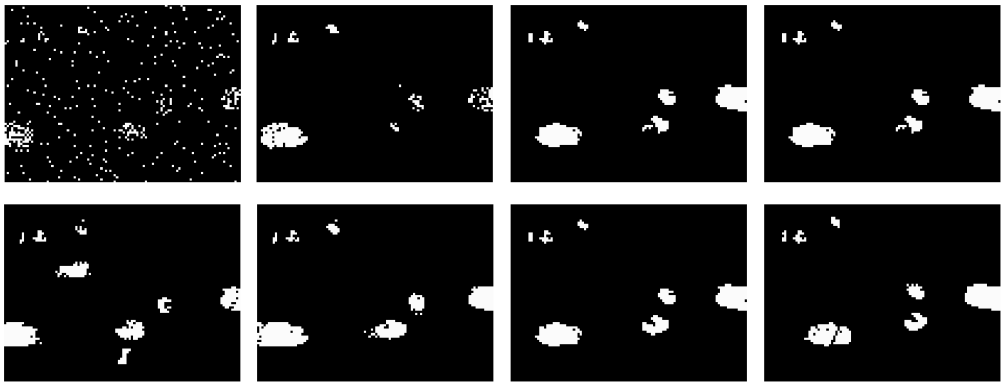


Figure 6.2: Some retrieved sample frames (from left to right, frame numbers 1, 11, 21 and 31) of the roundabout traffic video sequence for $f = 5$. Initial overlap $m^1 = 0.4$. Top panels: first cycle. Bottom panels: second cycle.

Kiev crossroad			Roundabout		
ω	m	learning time	ω	m	learning time
0.0	0.32117	499m39s	0.0	0.32060	100m19s
0.3	0.33677	500m08s	0.4	0.20104	101m01s
0.4	0.99751	500m25s	0.6	0.05548	101m43s
0.5	0.99767	501m30s	0.7	0.98344	102m50s
1.0	0.94742	504m27s	1.0	0.99732	104m51s

Table 6.1: Randomness ratio versus global overlap and learning time for the Kiev crossroad and roundabout video sequences.

slight difference is only due to the larger times to construct random networks than to construct local networks. The retrieval time for all cases was the same, around 5 minutes for the *Kiev* and 1 minute and a half for the *roundabout* sequences. In all cases, the respective memory usages for the learning and retrieval stages are about 14.3% and 10.4% of the whole computer memory, respectively.

One can conclude that, with a network with a randomness value of $\omega = 0.4$, the retrieval of the *Kiev* video sequence is possible and it saved considerably on wiring costs as the small-world topology suggests. It is also interesting to remark in Table 6.1 that the transition from confusion state (i.e. $m \sim 0$) to the retrieval state (i.e. $m \sim 1$) for *Kiev* traffic video happened around $\omega = 0.35$. This is related to an effective percolation of the information over all the network. Although the network is always connected, for smaller values of the randomness parameter, the synaptic strengths are not strong enough to percolate the information from some pixels to every region of the neuron states. For the *roundabout* video sequence, the randomness value for the transition from the confusion to the retrieval state, $\omega = 0.7$, is higher than in the *Kiev* video. This effect could be due to temporal correlation between frames which is smaller for the *roundabout* video.

It has also been experimented with a simpler "shifted-diagonal" Hebbian learning matrix (C. Molter and Bersini, 2005) replacing the pseudo-inverse rule(see Eqs. (6.9-6.11)). The maximal number of frames which could be retrieved for the *Kiev* video with $N = 8544, K = 4250, \omega = 0.5$ and with $m \sim 1$, was about $P = 16$. This choice is surely not appropriate for strongly correlated patterns and other learning rules like covariance rule (Dayan and Willshaw, 1991) or the Bayesian rule (Knoblauch, 2010) have been proposed to maximize the signal to noise ratio for a class of associative memories. A comparison with these models might be studied in a future work.

6.3.2 Robustness of the model with respect to the frame activity

The robustness of the model was tested (i.e. how overlapped the curves of average pattern and neural activities are along the frames of the video sequence) for a given network configuration: $N = 8544$, $K = 4250$ and $\omega = 0.4$. Fig. 6.3-left

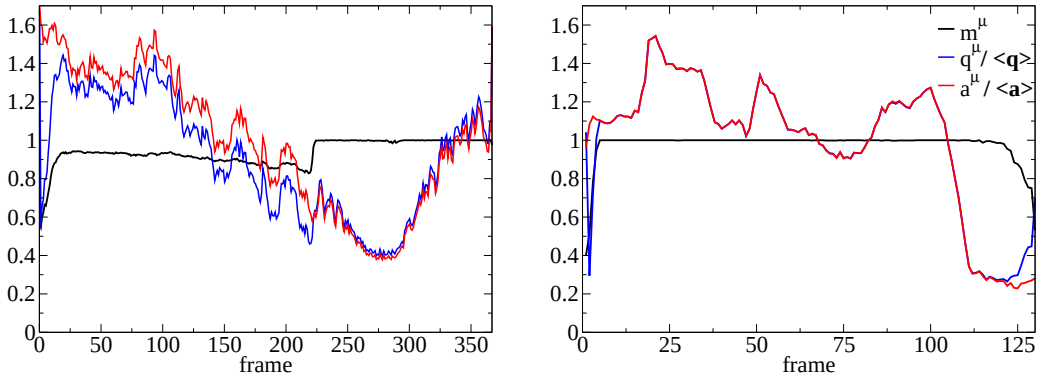


Figure 6.3: Left: Kiev crossroad sequence: Plot of overlapped pattern and neural activities against frames for $N = 8544$, $K = 4250$ and $\omega = 0.4$. Initial overlap $m_{\mu=1} = 0.5$. Right: Roundabout: Plot of overlapped pattern and neural activities against frames for $N = 8480$, $K = 4240$ and $\omega = 0.7$. Initial overlap $m_{\mu=1} = 0.4$. (Color on-line)

shows that the model is robust against a variable frame activity level, where the normalized activity (i.e. sparseness) of the frames a^μ/a varies in the range $0.4 < a^\mu/a < 1.6$). This graphic can be partitioned in three regions according the numbering of the frames. In a first region, where m (black line) varies from 0.55 to around 0.95 (from first frame to around frame 20), the average pattern (red line) and the neural (blue line) activities are uncorrelated and pattern activity is much higher than temporal neural activity. In a second region, where the value of m remains stable around 0.95 (from frame 21 to frame 225), the average pattern and neural activities are highly correlated but pattern activity is slightly larger than temporal neural activity. Finally, in the third region, where m equals to one from 226 to the end of the video, the pattern and neural activities are exactly coincident despite the significant changes in frame activity over time. The global overlap for the cycle is $m^c = 0.93$.

A similar curve for the roundabout sequence is presented in Fig. 6.3-right. for a network configuration: $N = 8480$, $K = 4240$ and $\omega = 0.4$. The overall behavior is similar to the *Kiev* sequence.

6.4 Conclusion

A Hopfield-type of Attractor Neural Network was used with a small-world connectivity distribution to learn and retrieve a sequence of highly correlated patterns. For this network model, a new weight learning heuristic which combines the pseudo-inverse approach with a row-shifting schema has been presented. The influence of the random connectivity ratio on the retrieval quality and learning time has been studied. This approach has been successfully tested for different combinations of the involved parameters on a complex traffic video sequence. Moreover, it was demonstrated to be robust with respect to highly-variable frame activity.

Another additional conclusion of this Chapter is that the more spatially correlated the frames are in average, the smaller is the range of the interaction (randomness parameter ω) which optimizes the retrieval of the video. The opposite also holds: the less spatially correlated the patterns are, the higher should be the value of ω. For instance, if there are large regions in the frames with high activity (i.e., a huge truck or bus in the corner) in a bulk of still background of the frame, then it is strongly spatially correlated. On the other hand, the threshold strategy used in the model is fundamental, since the dependency of θ with the neural activity (as well as with the pattern activity) is set in such a way that the network dynamics is self-controlled and it does not need from any human participation. For example, with the typical activity value used $a = 0.1$ in the traffic video, $\theta \sim 1$ was set in the neuron thresholds. For a uniform activity degree in the frames (i.e. $a = 1/2$), no threshold is needed ($\theta \sim 0$). Finally, for extremely sparse code (where $a \to 0$), the threshold increases to $\theta \sim 1/\sqrt{a}$.

Automatic video-based traffic monitoring systems are an alternative to loop detectors. Video-based systems provide updated global information on the analyzed traffic scene and also specific informations of the tracked vehicles. An interesting application of such systems is content-based traffic video retrieval, where using a query video, it is possible to retrieve another similar video from a database using some types of extracted features from the videos (i.e. textural information, motion trajectories of cars, etc). This can be useful for surveillance applications where one is interested in detecting certain events on the video

(i.e. accidents, congestions, etc). To achieve this goal, most approaches follow a feature-extraction approach which needs to segment the cars in the video and to track them individually. In a different way, using a holistic method like the proposed in this Chapter one can retrieve a complete video from a query frame, even if this frame represents a noisy scene of the video.

As the presented approach is holistic in the sense that no segmentation and feature extraction from the vehicles is required, one has to consider other holistic approaches applied to traffic videos for comparison purposes. The mentioned papers by Chan and Vasconcelos (2005) and Xie et al. (2004) do not segment the vehicles in the video, but they extract some global features from it (in particular, the complete motion information contained in the video), which are used for the retrieval task. They retrieve instances of traffic patterns using query videos; while in the proposed model the video can be retrieved using only a unique (possibly noisy) query frame. Moreover, the two compared papers do not quantitatively measure the video retrieval quality as one does using the global overlap. The proposed solution is suitable for the mentioned traffic application since it produces accurate retrieval results at reasonable time. However, the required learning times are still very large and the system needs improvement to be competitive with respect to those classical methods which segment the scene and track the moving targets. Moreover, the proposal can be now suited only to those traffic applications where the learning stage can be carried out off-line. Consequently, the use of complementary more-efficient strategies to compress the amount of memory required to store the patterns vectors like look-up tables (Knoblauch et al., 2010) or hashing techniques like LSH (Gionis et al., 1999) will be considered as future work.

Chapter 7

Conclusions and Summary

The model of Attractor Neural Network (ANN) on the small-world topology (local and random connectivity) has been investigated. The metric topology of network allows sustaining blocks of positive/negative activity in different regions of the network, and blocks of positive/negative overlaps for the case of memory networks. This block ordering is stable and compete with the global ordering according the topological parameters of the network. The proposal of a block-like structure could be closely related to biological brain systems, on the one hand, where different sensory blocks of patterns (arising from several cortical structures) may be independently retrieved. On the other hand, there could be a relation with the cortical mechanism of binding that allows information previously stored in different regions to be shared between pathways to accomplish a structured and unified representation.

In ANNs, the issue of retrieving mixed patterns randomly distributed on the network, has been discussed in the literature since the work of Amit (Amit, 1989), for fully connected networks. In the context of spatially organized connectivity, this problem has been addressed recently (Roudi and Treves, 2008), in a study which discuss the retrieval of multiple patterns at the same time, each on one part of the network, with graded response neurons. This thesis has considered the possibility of the coexistence of multiple local retrieval states of a unique pattern/antipattern configuration according to the geometric connectivity of the network. Different sectors of the pattern and anti-pattern can be retrieved in different portions of a network, whenever the connections are mainly local.

7.1 Summary of main results

The main results of this work can be summarized by chapter as follows:

Chapter 2

- Regions of opposite activity can be sustained in an Attractor Neural Network with metric connectivity.
- The change of stability from the block to the global attractors depends on the long-range character of the network connectivity.
- A larger number of blocks emerges with the network dilution.

Chapter 3

- The Attractor Neural Network with metric connectivity, using classic Hebbian learning, is able to retrieve patterns in blocks of opposite overlaps structured in a pattern/antipattern configuration.
- A new way to measure the local retrieval using a parameter that is related to the fluctuation of the block overlaps has been introduced.
- A transition from block retrieval to global retrieval occurs when the storage ratio increases and the topology becomes more random.

Chapter 4

- The sparse-coding network have retrieval abilities which are strongly dependent on the firing threshold of the neurons. A dynamical threshold that depends on the global activity of the network, the local activity of the neuron's neighborhood, and the sparseness of the pattern coding, has been used
- The structured information carried in the network has been characterized according to the fluctuations of the local overlaps, differentiating between bump and block phases.
- When the randomness in the network connectivity increases, the phase-diagram shows a transition from local to global retrieval. Furthermore, the local phase splits in a bump phase for low activity and block phase for high activity.

Chapter 5

- The storage and retrieval abilities of the ANN are studied for input patterns structured in *objects* with different levels of activity.
- For this type of patterns the correlation is relevant and needs to be addressed in order to improve the network performance.
- In order to properly sustain the different levels of activity of the *object* structures along the network, further threshold strategies have to be developed to enhance the performance of the network.

Chapter 6

- The feasibility of storing and retrieving automotive traffic videos using a sparse-coding ANN with a small-world topology demonstrated.
- A variant of the pseudo-inverse approach was introduced in order to learn/retrieve the sequence of correlated cyclic patterns of the video sequence.
- This approach has been successfully tested on a complex pattern, the traffic video sequences, for different combinations of the involved parameters. The suggested approach has also demonstrated to be robust with respect to highly-variable frame activity.

7.2 Future research

The robustness of this block phase for a whole set of topologies (for instance, scale-free networks) or neural models (for instance, integrate and fire neurons) worth to be checked. Also a systematic analysis of the probabilistic network dynamics, with a noise characterized by a temperature, remains to be done. These are subjects of future research. It is also worth to further improve the model presented in Chapter 5, for both theoretical neuroscience conjectures of information processing, as well as for implementation purposes, given that the variable-activity coding distribution can be found in real world data.

Conclusiones y sumario

El modelo de Red Neuronal Atractora (RNA) usando la topología "small-world" (conectividad local y aleatoria) ha sido investigado. La topología métrica de la red permite mantener bloques de actividad positiva/negativa en diferentes regiones de la red. Este ordenamiento en bloques de actividad es estable y compite con el ordenamiento global, según los parámetros topológicos de la red. La propuesta de una estructura de "bloques" podría estar estrechamente relacionada con los sistemas biológicos del cerebro, por un lado, cuando diferentes bloques sensoriales (provenientes de diferentes estructuras corticales) pueden ser independiente mantenidos. Por otro lado, podría haber una relación con el mecanismo de enlace (binding problem) que permite que la información previamente almacenada en diferentes regiones del cerebro sea compartida para lograr una representación unificada.

En RNAs, el tema de la recuperación de patrones mixtos distribuidos aleatoriamente en la red, se ha discutido en la literatura desde los trabajos de Amit (1989), con redes totalmente conectadas. En el contexto de redes con conectividad espacialmente organizada, este problema ha sido abordado recientemente por Roudi and Treves (2008) en un estudio que discute la recuperación de múltiples patrones al mismo tiempo, cada uno en una parte de la red, usando neuronas de respuesta graduada. En esta tesis se ha considerado la posibilidad de la coexistencia de varios estados de recuperación local de acuerdo a una configuración patrón/anti-patrón de acuerdo con la conectividad geométrica de la red. Diferentes sectores del patrón y anti-patrón pueden ser recuperados en diferentes partes de una red, siempre que las conexiones sean primordialmente locales. Además, se propuso una forma de medir la recuperación local utilizando un parámetro que está relacionado con la fluctuación de los solapamientos (overlap) de los bloques.

7.3 Recapitulación de los resultados

Los principales resultados de cada capítulo se resumen de la siguiente manera:

Capítulo 2

- Regiones de actividad opuesta pueden ser sostenidas en una Red Neuronal Atractora con conectividad métrica.
- El cambio de estabilidad del attractor en bloque al atractor global depende del carácter de largo alcance de conectividad de la red.
- Un mayor número de bloques surge con la dilución de la red.

Capítulo 3

- La Red Neuronal Atractora Hebbiana con conectividad métrica, es capaz de recuperar patrones en bloques con solapamientos opuestos.
- Se introdujo un nuevo método para medir la recuperación local utilizando un parámetro que está relacionado con la fluctuación del solapamiento de los bloques.
- La transición de la recuperación de bloque para la recuperación global ocurre cuando aumentan los patrones almacenados y la topología se vuelve aleatoria.

Capítulo 4

- La capacidad de recuperación de una red para patrones con codificación de baja actividad (sparse-coding), depende en gran medida del umbral de disparo de las neuronas. Se utilizó un umbral dinámico que depende de la actividad global de la red, la actividad local del vecindario de las neuronas, y de la actividad de los patrónes.
- El tipo de estructura de los estados neuronales se ha caracterizado utilizando un parámetro que está relacionado con las fluctuaciones de los solapamientos locales, diferenciando entre las fases de bloque y de "bumps".
- Cuando aumenta la aleatoriedad de la red, el diagrama de fases muestra una transición desde la recuperación local a la global. Por otra parte, la fase local se divide en una fase de "bumps" para baja actividad y fase de bloques para alta actividad.

Capítulo 5

- Se estudió la capacidad de almacenamiento y recuperación de la RNA para patrones de entrada estructurados en objetos con diferentes niveles de actividad.
- Para este tipo de patrones la correlación es relevante y debe ser tratada con el fin de mejorar el rendimiento de la red.
- Es necesario desarrollar nuevas estrategias de umbral a fin de mantener correctamente los diferentes niveles de actividad a lo largo de la red para los diferentes objetos estructurados.

Capítulo 6

- Se demostró la factibilidad de almacenar y recuperar videos de tráfico automovilístico utilizando una RNA con una topología "small-world" y codificación "sparse".
- Se introdujo una variante del enfoque pseudo-inverso para aprender/recuperar la secuencia de patrones cíclicos correlacionados de la secuencia del vídeo.
- Este enfoque ha sido probado con éxito en patrones complejos, como es el caso de las secuencias de vídeo de tráfico, para diferentes combinaciones de los parámetros involucrados. También ha demostrado ser robusto con respecto a la actividad altamente variable de los fotogramas del vídeo.

7.4 Trabajo futuro

Vale la pena investigar la robustez de esta fase de bloque para un conjunto de topologías (ej., redes libre de escala) o para diferentes modelos neuronales (ej., neuronas de integración y disparo). También queda por hacer un análisis sistemático de la dinámica de una red probabilística, caracterizada por una temperatura. Estos son temas de investigación futura. También vale la pena mejorar el modelo presentado en el Capítulo 5, tanto para conjeturar acerca del procesamiento neuronal de la información, así como para fines de implementacion, teniendo en cuenta que la codificación de actividad variable se pueden encontrar en la mayoría de las aplicaciones reales.

References

Ahn, Y.-Y., Kim, B. J., and Jeong, H. (2008). Wiring cost in the organization of a biological network. *arXiv:q-bio/0505009v1*. 5, 8

Albert, R. and Barabasi, A. (2002). Statistical mechanics of complex networks. *Rev. Mod. Phys.*, 74:47. 6

Amari, S.-I. and Maginu, K. (1988). Statistical neurodynamics of associative memory. *Neural Networks*, 1(1):63 – 73. 3

Amit, D. J. (1989). *Modeling brain function: the world of attractor neural networks*. Cambridge University Press, New York, NY, USA. 1, 3, 4, 17, 18, 20, 29, 30, 33, 46, 48, 59, 78, 93, 94, 105, 109

Amit, D. J., Gutfreund, H., and Sompolinsky, H. (1985a). Spin-glass models of neural networks. *Phys. Rev. A*, 32(2):1007–1018. 3

Amit, D. J., Gutfreund, H., and Sompolinsky, H. (1985b). Storing infinite numbers of patterns in a spin-glass model of neural networks. *Phys. Rev. Lett.*, 55(14):1530–1533. 3

Amit, D. J., Gutfreund, H., and Sompolinsky, H. (1987). Information storage in neural networks with low levels of activity. *Physical Review A*, 35:2293 – 2303. 9

Amit, Y. and Mascaro, M. (1999). Attractor networks for shape recognition. *Neural Computation*, 13:1415–1442. 3

REFERENCES

Arenzon, J. J. and Lemke, N. (1994). Simulating highly diluted neural networks. *Journal of Physics A: Mathematical and General*, 27(15):5161. 5

Bohland, J. and Minai, A. (2001). Efficient associative memory using small-world architecture. *Neurocomputing*, 38-40:489–496. 8, 17

Bollé, D., Blanco, J. B., and Verbeiren, T. (2004). The signal-to-noise analysis of the littlehopfield model revisited. *Journal of Physics A: Mathematical and General*, 37(6):1951. 3

Brunel, N. (2003). Dynamics and plasticity of stimulus-selective persistent activity in cortical network models. *Cereb. Cortex*, 13:1151–1161. 9

Brush, S. G. (1967). History of the lenz-ising model. *Rev. Mod. Phys.*, 39(4):883 – 893. 3

Bullmore, E. and Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, 10(3):186–98. 1, 6

C. Molter, U. S. and Bersini, H. (2005). Storing static and cyclic patterns in an hopfield neural network. 93, 101

Castillo, I. P. and Skantzos, N. S. (2004). The littlehopfield model on a sparse random graph. *Journal of Physics A: Mathematical and General*, 37(39):9087. 5

Chan, A. and Vasconcelos, N. (2005). Classification and retrieval of traffic video using auto-regressive stochastic processes. 89, 90, 104

Chan, A. and Vasconcelos, N. (2007). Road and trafic analysis from video. 90

Cherniak, C. (1994). Component placement optimization in the brain. *Journal of Neuroscience*, 14:2418–2427. 5, 8

Coolen, A. C. C. and Sherrington, D. (1993). Dynamics of fully connected attractor neural networks near saturation. *Phys. Rev. Lett.*, 71(23):3886–3889. 4

Cover, T. M. and Thomas, J. A. (1991). *Elements of Information Theory*. Wiley-Interscience. 38

Damasio, A. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain*. Grosset/Putnam, New York. 6

Davey, N., Christianson, B., and Adams, R. (2004). High capacity associative memories and small world networks. *International Joint Conference on Neural Networks*, pages 177–182. 5

Dayan, P. and Willshaw, D. (1991). Optimising synaptic learning rules in linear associative memories. *Biological Cybernetics*, 65:253 – 265. 10.1007/BF00206223. 101

Derrida, B., Gardner, E., and Zippelius, A. (1987). An exactly solvable asymmetric neural network model. *Europhys. Lett*, 4:167–173. 5

Dominguez, D. and Bollé, D. (1998). Self-control in sparsely coded networks. *Phys. Rev. Lett.*, 80:2961–2964. 7, 13, 33, 37, 51, 54, 60, 71, 76, 91, 95

Dominguez, D., González, M., Rodríguez, F. B., Serrano, E., Jr., R. E., and Theumann, W. (2012). Structured information in sparse-code metric neural networks. *Physica A: Statistical Mechanics and its Applications*, 391(3):799 – 808. 9, 11, 13, 14, 78

Dominguez, D., González, M., Serrano, E., and Rodríguez, F. B. (2009). Structured information in small-world neural networks. *Phys. Rev. E*, 79(2):021909. 10, 11, 12, 13, 25, 53, 56, 79

Dominguez, D., Koroutchev, K., Serrano, E., and Rodríguez, F. (2004). Mutual information and topology 1: Asymmetric network. In Yin, F., Wang, J., and Guo, C., editors, *Advances in Neural Networks - ISNN04*, volume 3173 of *LNCS*, pages 14–19, Berlin. 30

Dominguez, D., Koroutchev, K., Serrano, E., and Rodríguez, F. B. (2007). Information and topology in attractor neural networks. *Neural Comput.*, 19:956–973. 6, 7, 8, 17, 33, 34, 37, 38, 47, 53, 90

Edin, F., Klingberg, T., Johansson, P., McNab, F., Tegnér, J., and Compte, A. (2009). Mechanism for top-down control of working memory capacity. *Proceedings of the National Academy of Sciences of the United States of America*, 106(16):6802 – 6807. 2, 9, 54

Eliasmith, C. (2007). Attractor network. *Scholarpedia*, 2(10):1380. 2

Erdös, P. and Rényi, A. (1959). On random graphs, I. *Publicationes Mathematicae (Debrecen)*, 6:290–297. 4

Evans, M. R. (1989). Random dilution in a neural network for biased patterns. *Journal of Physics A*, 22:2103–2118. 17

Fink, W. (2004). Neural attractor network for application in visual field data classification. *Physics in Medicine and Biology*, 49(13):2799. 3

Foldiak, P. and Endres, D. (2008). Sparse coding. *Scholarpedia*, 3(1):2984. 54

Gionis, A., Indyk, P., and Motwani, R. (1999). Similarity search in high dimensions via hashing. In *Proceedings of the 25th International Conference on Very Large Data Bases*, VLDB '99, pages 518 – 529, San Francisco, CA, USA. Morgan Kaufmann Publishers Inc. 104

González, M., Dominguez, D., and Ángel Sánchez (2011). Learning sequences of sparse correlated patterns using small-world attractor neural networks: An application to traffic videos. *Neurocomputing*, 74(14-15):2361 – 2367. 2, 3, 14, 87

González, M., Dominguez, D., and Rodríguez, F. B. (2008a). Block activity in metric neural networks. *WASET Proceedings*, 27:56–59. 20

González, M., Dominguez, D., and Rodríguez, F. B. (2008b). Learning block memories with metric networks. *WASET Proceedings*, 27:60–63. 28

González, M., Dominguez, D., and Rodríguez, F. B. (2009). Block attractor in spatially organized neural networks. *Neurocomputing*, 72:3795–3801. 11, 13, 53, 79

Griniasty, M., Tsodyks, M. V., and Amit, D. J. (1993). Conversion of temporal correlations between stimuli to spatial correlations between attractors. *Neural Comput.*, 5:1–17. 2

Hatchett, J. P., Castillo, I. P., Coolen, A. C., and Skantzos, N. S. (2005a). Dynamical replica analysis of disordered ising spin systems on finitely connected random graphs. *Phys. Rev. Lett.*, 95:117204. 8, 34

Hatchett, J. P., Skantzos, N. S., and Nikoletopoulos, T. (2005b). Dynamic rewiring in small world networks. *Phys. Rev. E*, 72:066105. 19

Hebb, D. O. (1949). The organization of behavior. 1, 3

Hertz, J., Krogh, J., and Palmer, R. (1991). *Introduction to the Theory of Neural Computation*. Addison-Wesley, Boston. 3, 4, 7, 12, 17, 29, 33, 50, 59, 70, 75, 78, 84, 91, 94

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences of the United States of America*, 79(8):2554–2558. 2, 3, 4

Izhikevich, E. (2003). Simple model of spiking neurons. *Neural Networks, IEEE Transactions on*, 14(6):1569 – 1572. 1

Johansson, C. and Lansner, A. (2007). Imposing biological constraints onto an abstract neocortical attractor network model. *Neural Comput.*, 19:1871–1896. 6, 17

Johansson, C., Rehn, M., and Lansner, A. (2006). Attractor neural networks with patchy connectivity. *Neurocomputing*, 69:627–633. 17

Jung, Y.-K. and Ho, Y.-S. (2001). A feature-based vehicle tracking system in congested traffic video sequences. In *Proceedings of the Second IEEE Pacific Rim Conference on Multimedia: Advances in Multimedia Information Processing*, PCM '01, pages 190–197, London, UK, UK. Springer-Verlag. 90

Kastrinaki, V., Zervakis, M., and Kalaitzakis, K. (2003). A survey of video processing techniques for traffic applications. *Image and Vision Computing*, 21(4):359 – 381. 89, 90

Knoblauch, A. (2010). Optimal synaptic learning in non-linear associative memory. *Neural Networks*, pages 18 – 23. 101

Knoblauch, A., Palm, G., and Sommer, F. T. (2010). Memory capacities for synaptic and structural plasticity. *Neural Comput.*, 22:289–341. 2, 104

Koroutchev, K. and Koroutcheva, E. (2006). Bump formation in a binary attractor neural network. *Phys. Rev. E*, 73:026107. 9, 10, 18, 34, 53, 54, 90

Lago-Fernández, L. F., Huerta, R., Corbacho, F., and Sigüenza, J. A. (2000). Fast response and temporal coherent oscillations in small-world networks. *Physical Review Letters*, 84(12):2758–2761. 1

Levy, N., Horn, D., and Ruppin, E. (1999). Associative memory in a multimodular network. *Neural Comput.*, 11(7):1717–1737. 6, 17

Li, B. and Chellappa, R. (2002). A generic approach to simultaneous tracking and verification in video. *IEEE Transactions on Image Processing*, 11:530–544. 90

Li, C. and Chen, G. (2003). Stability of a neural network model with small-world connections. *Phys. Rev. E*, 68:052901–05. 5, 6, 7, 17, 53

Liu, W.-Z., Jing, W., Li, H., Gong, H.-Q., and Liang, P.-J. (2010). Spatial and temporal correlations of spike trains in frog retinal ganglion cells. *Journal of Computational Neuroscience*, pages 1–11. 10.1007/s10827-010-0277-9. 11

Lu, J., He, J., Cao, J., and Gao, Z. (2006). Topology influences performance in the associative memory neural networks. *Physics Letters A*, 354:335–343. 17

Masuda, N. and Aihara, K. (2004). Global and local synchrony of coupled neurons in small-world networks. *Biol. Cybernetics*, 90:302. 8, 34, 53

McGraw, P. N. and Menzinger, M. (2003). Topology and computational performance of attractor neural networks. *Phys. Rev. E*, 68:047102. 5, 6, 7, 8, 12, 17, 34, 53

Morelli, L., Abramson, G., and Kuperman, M. (2004). Associative memory on a small-world neural network. *Eur. Phys. J. B*, 38:495–500. 8, 17, 34, 53

Newman, M. E. J. and Watts, D. J. (1999). Scaling and percolation in the small-world network model. *PHYS.REV.E*, 60:7332. 7, 72

Nikoletopoulos, T., Coolen, A. C., Castillo, I. P., Skantzos, N. S., Hatchett, J. P., and Wemmenhove, B. (2004). Replicated transfer matrix analysis of ising spin models on small world lattices. *J. Phys. A: Math. Gen.*, 37:6455–64754. 9, 34, 53

Octave (2011). Octave gnu homepage, http://www.gnu.org/software/octave/. 98

Okada, M. (1995). A hierarchy of macrodynamical equations for associative memory. *Neural Networks*, 8(6):833 – 838. 3

Okada, M. (1996). Notions of associative memory and sparse coding. *Neural Network*, 8-9:1429. 7, 9, 33

Olshausen, B. and Field, D. (2004). Sparse coding of sensory inputs. 9, 90

Paula, D. R., Araujo, A. D., J. S. Andrade, J., Herrmann, H. J., and Gallas, J. A. C. (2006). Periodic neural activity induced by network complexity. *Phys. Rev. E*, 74(1):017102. 20, 41

Ristic, B., Arulampalam, S., and Gordon, N. (2004). Beyond the kalman filter: Particle filters for tracking applications. 90

Rolls, E. and Treves, A. (1998). *Neural Networks and Brain Function*. Oxford University Press, England. 3, 6, 8, 12, 17

Roudi, Y. and Treves, A. (2004). An associative network with spatially organized connectivity. *J. Stat. Mech.*, 2004(07):P07010. 9, 17, 49, 53

Roudi, Y. and Treves, A. (2006). Localized activity profiles and storage capacity of rate-based autoassociative networks. *Phys. Rev. E*, 73:061904. 9, 34, 53, 54

Roudi, Y. and Treves, A. (2008). Representing where along with what information in a model of a cortical patch. *PLoS Comput Biol*, 4(3):e1000012. 6, 10, 12, 54, 59, 105, 109

Rubin, J. and Bose, A. (2004). Localized activity patterns in excitatory neuronal networks. *Networks: Comput. Neural Syst*, 15:15. 9

Ruppin, E. and Yeshurun, Y. (1991). Recall and Recognition in an Attractor Neural Network Model of Memory Retrieval. *Connection Science*, 3:381 – 400. 1

Stauffer, D., Aharony, A., da Fontoura Costa, L., and Adler, J. (2003). Efficient hopfield pattern recognition on a scale-free neural network. *The European Physical Journal B - Condensed Matter and Complex Systems*, 32:395–399. 10.1140/epjb/e2003-00114-7. 5, 6, 7

Stringer, S. M., Rolls, E. T., Trappenberg, T. P., and de Araujo, I. E. T. (2003). Self-organizing continuous attractor networks and motor function. *Neural Networks*, 16:161–182. 1

Stroffek, J., Kuriscak, E., and Marsalek, P. (2007). Pattern storage in a sparsely coded neural network with cyclic activation. *Biosystems*, 89(1-3):257 – 263. Selected Papers presented at the 6th International Workshop on Neural Coding, 6th International Workshop on Neural Coding. 9

Tamarit, F. A., Stariolo, D. A., and Curado, E. M. F. (1991). Diluted-neural-network model with higher-order interactions. *Phys. Rev. A*, 43(12):7083–7086. 5

Torres, J. J., Munoz, M. A., Marro, J., and Garrido, P. L. (2004). Influence of topology on the performance of a neural network. *Neurocomputing*, 58-60:229–234. 5, 6, 7, 17

Trappenberg, T. (2002). *Fundamentals of Computational Neuroscience*. Oxford University Press, USA. 12, 91

Treves, A. (1990). Graded-response neurons and information encodings in autoassociative memories. *Phys. Rev. A*, 42(4):2418–2430. 17

Uezu, T., Hirano, A., and Okada, M. (2004). Retrieval properties of hopfield and correlated attractors in an associative memory model. *Journal of the Physical Society of Japan*, 73(4):867–874. 2

USC-SIPI (2011). The usc-sipi image database, http://sipi.usc.edu/database/. 18, 28

Vinje, W. E. and Gallant, J. L. (2000). Sparse coding and decorrelation in primary visual cortex during natural vision. *Science*, 287(5456):1273–1276. 12

Wang, X., Tegner, J., and Constantinidis, C. (2004). Division of labor among distinct subtypes of inhibitory neurons in a cortical microcircuit of working memory. *Proceedings of the National Academy of Sciences*, 101(5):1368 – 1373. 11, 54

Wang, X.-J. (2001). Synaptic reverberation underlying mnemonic persistent activity. *Trends in Neurosciences*, 24(8):455 – 463. 9

Watts, D. J. and Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393:440–422. 6, 17, 19, 36, 56, 92

Wemmenhove, B. and Coolen, A. C. C. (2003). Finite connectivity attractor neural networks. *Journal of Physics A Mathematical General*, 36:9617–9633. 5

Wills, T. J., Lever, C., Cacucci, F., Burgess, N., and OKeefe, J. (2005). Attractor dynamics in the hippocampal representation of the local environment. *Science*, 308(5723):873–876. 1

Xie, D., Hu, W., Tan, T., and Peng, J. (2004). Semantic-based traffic video retrieval using activity pattern analysis. In *ICIP*, pages 693–696. 90, 104

Yu, S.-H., Hsieh, J.-W., Chen, Y.-S., and Hu, W.-F. (2003). An automatic traffic surveillance system for vehicle tracking and classification. In *Proceedings of the 13th Scandinavian conference on Image analysis*, SCIA'03, pages 379–386, Berlin, Heidelberg. Springer-Verlag. 90

Zhang, P. and Chen, Y. (2007). Statistical neurodynamics for sequence processing neural networks with finite dilution. In *ISNN (1)*, pages 1144–1152. 5

Zheng, P., Tang, W., and Zhang, J. (2010). A simple method for designing efficient small-world neural networks. *Neural Networks*, 23(2):155 – 159. 53

Zillmer, R., Livi, R., Politi, A., and Torcini, A. (2006). Desynchronization in diluted neural networks. *Phys. Rev. E*, 74(3):036203. 5