Neurocomputing 442 (2021): 269–280

DOI: https://doi.org/10.1016/j.neucom.2021.02.033

Ensemble of Diluted Attractor Networks with Optimized Topology for Fingerprint Retrieval

Mario González[a,*], Ángel Sánchez[b], David Dominguez[c], Francisco B. Rodríguez[c,*]

[a] *SI² Lab, Universidad de las Américas, Quito, Ecuador*
[b] *E.T.S. Ingeniería Informática, Universidad Rey Juan Carlos, 28933 Madrid, Spain*
[c] *Grupo de Neurocomputación Biológica, Departamento de Ingeniería Informática, Escuela Politécnica Superior, Universidad Autónoma de Madrid, 28049 Madrid, Spain*

Abstract

The present study analyzes the retrieval capacity of an Ensemble of diluted Attractor Neural Networks for real patterns (i.e., non-random ones), as it is the case of human fingerprints. We explore the optimal number of Attractor Neural Networks in the ensemble to achieve a maximum fingerprint storage capacity. The retrieval performance of the ensemble is measured in terms of the network connectivity structure, by comparing 1D ring to 2D cross grid topologies for the random shortcuts ratio. Given the nature of the network ensemble and the different characteristics of patterns, an optimization can be carried out considering how the pattern subsets are assigned to the ensemble modules. The ensemble specialization splitting into several modules of attractor networks is explored with respect to the activities of patterns and also in terms of correlations of the subsets of patterns assigned to each module in the ensemble network.

Keywords:
Attractor neural networks, Ensemble of diluted modules, Structured patterns, Module input optimization, Module specialization, Fingerprint retrieval, Information performance metrics

*Corresponding author
Email addresses: mario.gonzalez.rodriguez@udla.edu.ec (Mario González), angel.sanchez@urjc.es (Ángel Sánchez), david.dominguez@uam.es (David Dominguez), f.rodriguez@uam.es (Francisco B. Rodríguez)

1. Introduction

Amari-Hopfield type Attractor Neural Networks (ANNs) have been proposed to solve a large number of problems in different domains (Amit, 1992; Hopfield, 1982). This type of network is one of the computational models that have been intensively analyzed in different contexts (Agliari et al., 2012, 2014, 2015a,b). Patterns are dynamically stored throughout the learning process, using the generalized Hebb rule. In this type of network the patterns stored in it are considered attractors, so the dynamics of the neural system starting in a certain state converges to one of these stored attractors. Recurrent dynamics allow the attractor networks (similarly to the brain) to perform an interpretation process of pattern completion, and to construct coherent neural states from noisy and/or incomplete data. ANNs dynamics can be seen in many application domains like data clustering, pattern denoising or image-processing (e.g. super-resolution).

ANNs are related to several deep learning auto encoding models such as the Variational Autoencoders (Kingma and Welling, 2013) and denoising autoencoders (Vincent et al., 2008). These models can be applied recursively, however their convergence is not guaranteed nor the output quality over iterations. Recently, ANNs have been revised in light of deep learning methods and a convolutional bipartite architecture model has been proposed (Iuzzolino et al., 2019) with a new training loss, activation function, and connectivity constraints.

The communication between nodes (i.e. the information interaction) is another important factor in most network-based computational models (Du et al., 2018). Two aspects must be considered in the application of these models to optimization processes: network structure (i.e. topology) and information processing strategy, respectively. In general, a fully-connected structure is usually not a good choice and a sparse structure is preferable. Based on the defined network structure, the relevance of the network components (i.e. nodes or subnetworks) can be different in defined information processing strategies for optimization problems (Du et al., 2018). In the context of ANN, the maximum storage capacity is approximately $0.13 \times N$, where N is the number of neurons. This capability of storage is limited by the well-known crosstalk term (Hertz et al., 1991; Amit, 1992). Fully connected ANNs are very expensive computationally speaking, due to intensive updating of local neuron fields (each local field involves N nodes' activation). A possible alternative approach is the well-known diluted networks

(K connections per unit instead of N connections, with $K < N$), where it is shown that the computation is less intensive (each local field only involves a K neighborhood activation), and most importantly, the capacity of the attractor network per connection is also increased (Sompolinsky and Kanter, 1986; Derrida et al., 1987; Amit and Brunel, 1993; Arenzon and Lemke, 1994; Wemmenhove and Coolen, 2003).

Recently, it has been empirically demonstrated how the storage capacity of random patterns can be increased in an Ensemble of Attractor Neural Networks (EANN) using the concept of dilution of network connections in the different modules (Gonzalez et al., 2017; González et al., 2017), when we compare with a single attractor network with equal connectivity than all the modules (ensemble of attractors). With this new combination of attractors that exploits the concept of dilution in network connections, one wonders how this system will behave with real and non-random patterns (i.e. structured patterns). A first naive exploration of this question has been carried out in this regard in two congress works: for 2D Gesture Retrieval (Dávila et al., 2019), and for fingerprint patterns storage (González et al., 2019). In all these seminal works the connectivity of each module of the network ensemble is not optimized. In this work, we optimize the connectivity of the attractor sets using fingerprint images as real patterns and we perform an exhaustive analysis of how real patterns affect the storage characteristics of the attractor set. An initial analysis of how storage capacity varies depending on the dilution of connections for a single attractor network with random patterns was analyzed in (Dominguez et al., 2012), and later with fingerprint patterns in (González et al., 2014). Specifically in this work we looked for a topology that maximizes the storage capacity of fingerprint patterns for a set of attractor networks with a considerable dilution of connections in each of the modules of the ensemble. We compare the resulting capacity of optimized ensemble connectivity with the capacity of a single diluted attractor network with a number of connections equal to the sum of all the connections of the different modules in the ensemble.

Human fingerprints are one of the most extended biometric modalities to verify the identity and to recognize a person (Maltoni et al., 2009). This kind of pattern results from impression of the structure of ridges on the last joint of a person's finger. There are some properties which make a fingerprints widely used in forensic applications: (1) their uniqueness of the characteristic curved lines of ridges and (2) the stability of the papillary drawing over a person lifetime. Fingerprint retrieval problem consists in being able to recover as

accurately as possible the most similar fingerprint from a collection of learned patterns using a query fingerprint (which could be a noisy or incomplete pattern). Due to the increasingly large size of biometric databases, there is a need from computational methods to increase the efficiency of the fingerprint retrieval task (Maltoni et al., 2009).

In the present work we show that the maximum of the fingerprint storage capacity is obtained for an optimal number of modules of the EANN, and for a dilution connectivity in each module intermediate between the purely local connectivity and a completely random connectivity. Different types of local connectivities are also explored in the work, with 2D cross-type connectivity being the most optimal. Finally, due to the structured (non-random) nature of the patterns, the ensemble specialization of several modules of attractor networks is explored to optimize the capacity of the network. First, it is studied how to distribute the patterns by activities in the ensemble modules, thus specializing the modules for different activity degrees. It is also investigated how to distribute the correlated patterns among the different modules of the set of attractors (i.e., the ensemble modules) to optimize the storage capacity in the ensemble of attractors. It is shown that to optimize the capacity in this case it is necessary to distribute the fingerprints in the different modules so that each module has similar characteristics in terms of the correlation between the patterns learned by that module.

The rest of the paper is organized as follows. Section 2 describes different aspects of the EANN model (i.e., topology, learning, retrieval dynamics, and information measures). To better organize the results of the paper, the results are presented in two sections, 3 and 4. Section 3 presents the results related to topology optimization. In this section are shown the experimental performance results achieved by the EANN model with optimal topology versus a single ANN under equivalent conditions of complexity. Section 4 shows the results related to the specialization of the ensemble modules according to the structural characteristics of the patterns to be stored. A case example of pattern subset assignment and a case of module specialization are presented. Finally, Section 5 outlines the main conclusions of this experimental study and suggests future research.

2. EANN model

In this section the base module of the EANN model, the neural coding, the network topology, learning and retrieval dynamics, are described. Finally,

in this section, a schematic representation of the modularized EANN system is illustrated for the fingerprint retrieval and specialization, being compared to the single attractor network.

2.1. Modules coding, topology and dynamics

Let us define a particular ensemble module of the actual attractor neural network model. The definition is equivalent for the single network as well for each module in the ensemble. The state of a neuron by a set of N binary variables $\boldsymbol{\tau}^t = \{\tau_i \in 0,1; i = 1,\ldots,N\}$, where 1 and 0 represent, respectively, active and inactive states. The network stores a set of patterns, i.e. the fingerprints $\{\boldsymbol{\eta}^\mu,\ \mu = 1,\ldots,P\}$, through a learning process. The neurons are updated in time t and aim to retrieve the patterns, $\boldsymbol{\tau}^t \sim \boldsymbol{\eta}^\mu$, for large enough time t. The patterns are encoded as a set of binary variables $\boldsymbol{\eta}^\mu = \{\eta_i^\mu \in 0,1; i = 1,\ldots,N\}$, with pattern activity $a^\mu = \langle \eta^\mu \rangle \equiv \sum_i \eta_i^\mu / N$ (Dominguez et al., 2012; Doria et al., 2016). Similarly, the neural activity is $q \equiv \langle \tau \rangle$.

It is convenient to use the normalized variables, $\xi_i \equiv \frac{\eta_i - a}{\sqrt{A}}$, $\sigma_i \equiv \frac{\tau_i - q}{\sqrt{Q}}$, where $A \equiv Var(\eta) = a(1-a)$ and $Q \equiv Var(\tau) = q(1-q)$ are the pattern and neuronal variances, respectively. The neuron dynamics (Dominguez et al., 2009, 2012) can be written as

$$\tau_i^{t+1} = \Theta(h_i^t - \theta(a)),\ h_i^t \equiv \frac{1}{K}\sum_j J_{ij}\sigma_j^t,\ i = 1,\ldots,N, \tag{1}$$

where h_i^t represents the neuron's input field, and the step function is: $\Theta(x) = 0(1), x < (\geq)0$. A neural threshold defined as

$$\theta(a) = \frac{1-2a}{2\sqrt{A}}c, \tag{2}$$

where c is adjusted around $c \sim 1$, is necessary to keep the neural activity close to that of the learned patterns (Dominguez et al., 2012; González et al., 2015; Doria et al., 2016).

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

$$J_{ij} \equiv C_{ij}W_{ij}, \tag{3}$$

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 corresponds to a small-world network (Watts and Strogatz, 1998) with connectivity $K = K_l + K_r$, and random shortcuts ratio of $\omega = K_r/K$, where K_r corresponds to the long-range connections and K_l to the local connections. This definition is valid starting from a ring as well as other 1D or 2D topologies. The network is also characterized by the connectivity ratio $\gamma = K/N$, which represents the dilution degree of the ANN.

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

$$W_{ij}^{\mu} = W_{ij}^{\mu-1} + \xi_i^{\mu}\xi_j^{\mu}, \quad W_{ij}^{0} = 0. \tag{4}$$

After P learning steps, the weights reach the value $W_{ij} = \sum_{\mu}^{P} \xi_i^{\mu}\xi_j^{\mu}$.

We measure the retrieval ability of the network modules with the overlap,

$$m^{\mu} \equiv \frac{1}{N}\sum_{i}^{N} \xi_i^{\mu}\sigma_i, \tag{5}$$

which is the statistical correlation between the learned pattern ξ_i and the neural state σ_i. For $m = 1$ and $q \sim a$ one has a perfect retrieval of the pattern by the network, for $m = 0$ no retrieval is achieved, and for intermediate values the pattern is retrieved with noise. Hence, the value of m is a measure of the retrieval quality of the pattern performed by the network (Dominguez et al., 2009; Doria et al., 2016). Together with the overlap m, the load ratio $\alpha = P/K$ is useful to evaluate the network retrieval permanence. The initial state of the network is given by $\sigma_i^{t=0} = \xi_i^{\mu}$, which is a noiseless state loaded from a test pattern μ where $m_{t=0}^{\mu} = 1$.

2.2. Ensemble of Attractor Neural Networks (EANN)

A schematic representation of the single ANN is presented in Fig. 1-left. The connectivity ratio γ is diluted with $K < N$. A set of P fingerprints $\boldsymbol{\xi}$ is presented to the network in a learning phase, represented with the red dashed arrow. Then, this set of fingerprints is presented in a retrieval phase in order to test the recall abilities of the network in terms of the retrieved patterns load α, and the quality of the retrieval m. This is represented with the solid black arrow.

In Fig. 1-right, a schematic representation of an ensemble of ANN modules with a number of n components is presented. The connectivity in each ANN_b module b is highly diluted with $K_b \ll N$, $b \in \{1, ..., n\}$. The set of patterns is divided into disjoint subsets of uniform size $P_b = P/n$, and each pattern

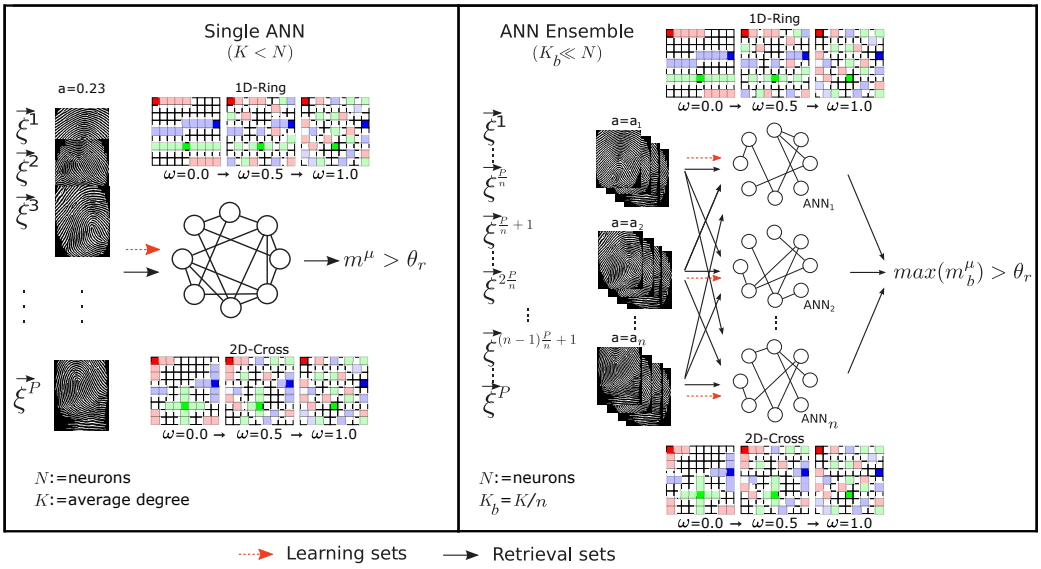

Figure 1: *Schematic representation of the single attractor vs. ensemble network model.* Left: single attractor network with entire fingerprint dataset. Right: ensemble of attractor networks with pattern subsets.

subset is learned by its corresponding ANN_b module as represented with the red dashed arrows. E.g. $\{\vec{\xi}^\mu,\ \mu = 1, \ldots, P/n\}$ for the first module $ANN_b, b = 1$, as shown in Fig. 1-right. The solid black arrows in Fig. 1-right, represent the retrieval stage, in which all the pattern subsets are presented to all ANN modules in order to test the discrimination among them. The target patterns are considered as retrieved by the ANN module with the higher overlap value over the retrieval threshold θ_r i.e. $max(m_b^\mu) > \theta_r$. For comparison purpose, the retrieval threshold is assumed to take the same value $\theta_r = 0.7$ for each component in the ensemble, as well as, for the single ANN system.

In both cases, single ANN and EANN systems, one looks for the optimal connectivity between the purely local and a completely random connectivity, to obtain a maximum recovery capacity. Two types of local connectivities, 1D-ring and 2D-cross, are considered in this study. In the EANN system one also looks for the optimal connectivity of the b modules to obtain a maximum recovery capacity with the two types of local connectivity that were used in the single attractor network. Thus the two systems schemes are compared with each having the same number of connections: the single attractor

network versus the EANN network with total number of connections being equal. The computational cost of the single ANN and the EANN are the same, as shown in (Gonzalez et al., 2017), one uses $K = K_b \times n$, the original dataset can be divided in subsets in order to increase the storage capacity when compared with single attractor.

As a use case, the EANN can perform specialization, where the fingerprints to be learned are partitioned into subsets according to the overlap between patterns and by assigning these subsets to different ensemble modules for reducing the overlap between patterns in each module (i.e, to reduce the cross-talk noise in the network). Three strategies are compared: random assignment, windowed assignment after sorting the datasets in terms of overlap between patterns, and alternate assignment of the sorted dataset. Next, we present another use case which consists in assigning subsets of patterns to modules according to similar dataset features, as it is the pattern activity. In this case an EANN with $n = 3, K_b = 100$ per module is used. Each module specializes its retrieval dynamics using a threshold adapted for the following pattern subsets activities: skeletonized fingerprints (activity $a \sim 0.0844$, and field threshold $\theta(a) = 1.4951$), normal binarized ($a \sim 0.2258,\ \theta(a) = 0.656$), and thickened fingerprints ($a \sim 0.4660,\ \theta(a) = 0.0681$). The same dataset have been used but differing the morphological operations applied over the binary images to achieve the respective activity as described in Fig. 2.

2.3. EANN information measures

In order to evaluate the EANN performance, the retrieval efficiency R is defined as the number of learned patterns that are successfully retrieved $R = \frac{P_r}{P}$, where P_r is the overall number of retrieved patterns that satisfy $m^\mu > \theta_r$, and P is the overall number of patterns presented to the network during the learning phase. One has that $P \geq P_r$. Here, $\theta_r = 0.7$ is used as the retrieval threshold, unless stated otherwise. The mean retrieval overlap M is calculated over all patterns subset $\mu \in 1, 2, \ldots, P$, $M = \langle m \rangle_\mu = 1/P \sum_{\mu=1}^{P} m^\mu$. It is worth noting that in the case of the ANN ensemble, the retrieval pattern load is calculated as $\alpha_R = \frac{P_r}{K_b \times n}$ where n is the number of subnetworks. Thus, we use $K_b \times n = K$ constant for all network ensembles studied, where K is the connectivity of the single "dense" network. Also, it is of worth to define the pattern gain G of the ANN ensemble by taking the single ANN system retrieval performance in terms of recovered patterns (P_r^s) as baseline, and it is given by $G = P_r^e / P_r^s$. Here P_r^e stands for the number of total recovered patterns by the ANN ensemble and

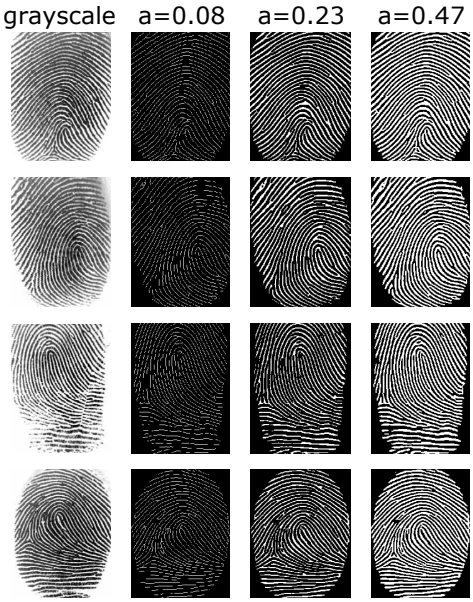

Figure 2: *Examples of fingerprints in the dataset and their associated transformations.* From left to right: gray-scale fingerprint, skeletonized fingerprint (activity $a \sim 0.08$), binarized fingerprint ($a \sim 0.23$), dilated (thickened) fingerprint ($a \sim 0.47$).

P_r^s stands for the patterns recovered by the single network at the maximum retrieval pattern load $max(\alpha_R)$.

2.4. Fingerprint datasets

Fig. 2 shows several representative examples of grayscale fingerprint images obtained from the sensor (left) followed by the binarized outputs: skeletonized, binarized with normal width, and dilated (thickened) fingerprints respectively. The fingerprint collections come from the International Fingerprint Verification Competition, editions FVC2000, FVC2002 and FVC2004 (Maio et al., 2004). Each edition has four different databases available, which are usually collected using the following sensors/technologies: optical, capacitive, thermal sweeping sensor, as well as synthetic generated fingerprints. The preprocessing details can be found in (González et al., 2014). These patterns are used as the training set of fingerprints that the network will learn. The mean activities, that is, the mean number of white pixels for each fingerprint dataset are: skeletonized: $a \sim 0.08$, normal width: $a \sim 0.23$, and dilated: $a \sim 0.47$.

3. Fingerprint recognition results: optimal topology

This section presents the performance results and comparison between the single and EANN systems. The EANN optimization in terms of the topology configuration (1D-ring vs 2D-cross) and for random connectivity shortcuts for different number of modules in the ensemble. A threshold analysis is also presented to understand how the performance of the ensemble system behaves in terms of the retrieval threshold.

The model in section 2 was implemented in C++, and Python 3 was used to process and visualize the output results. The code with a base example can be found at https://github.com/marsgr6/ann/tree/master/cpp_ann. All experiments, in sections 3 and 4, were carried out on a Linux cluster using Slurm cluster management and job scheduling system (Yoo et al., 2003).

3.1. Ensemble performance vs single network

In this section we compare through different examples the performances of a single attractor with an EANN network. These examples contain all the metrics used for comparison purposes, including the degree of randomness of the chosen network topology. Thus Fig. 3 depicts the retrieval performance for the normal activity: $a \sim 0.23$ dataset, where a proper dynamic threshold of $\theta(a) = 0.656$ is used. The performance measures depicted are the mean retrieval overlap M in blue lines with circle-markers, the retrieval efficiency $R = P_r/P$ in green lines with plus-markers, and retrieval capacity α_R, which is an information measure, in red lines with x-markers. The behavior shown in Fig. 3 shows the increasing of the number of patterns loaded in the network to be learned P in the x-axis. The y-axis depicts the the retrieval performance measures M, R, α_R. The curves for M and R remain with a value close to 1, indicating the retrieval of most of the patterns with high quality overlap m^μ. When the pattern load P surpasses a certain value, the retrieval quantity R and overall retrieval quality M goes continuously to 0. The critical value can be found using the information measure α_R, which measures the number of retrieval patterns in terms of the ensemble connectivity $\alpha_R = \frac{P_r}{K_b \times n}$. The curve for α_R shows a non monotonic behavior exhibiting a maximum value, which is used as a critical point and is depicted as a vertical red line in the top panels of Fig. 3.

In the same figure, top-left panel, is depicted the retrieval performance for the single network $N = 89420,\ K_b = 240,\ n = 1$ with $\omega = 1$ ratio of random shortcuts starting from a 1D-configuration ring network (equivalent to an

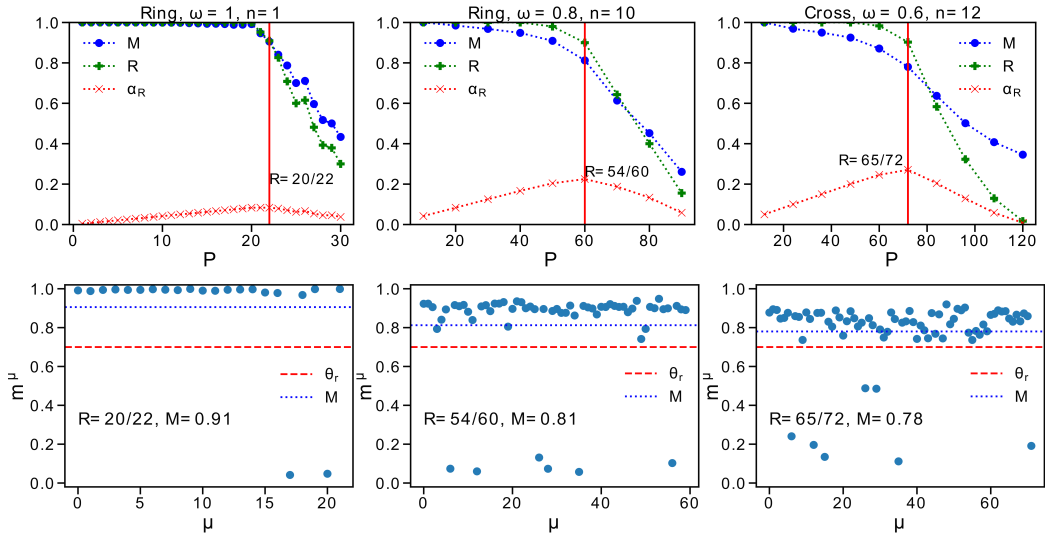

Figure 3: *Example of ensemble performance vs. single network for increasing number of patterns to be learned.* From left to right: random network $\omega = 1.0$ and $n = 1$ modules; 1D network with $\omega = 0.8$, $n = 10$ modules and initial structure of ring; 2D network with $\omega = 0.6$, $n = 12$ modules and initial structure of cross. Top panels: Ensemble performance curves M, R, α_R and maximum value cut-off for the pattern load P, indicated with the solid red vertical line, and the fraction of retrieved patterns $R = P_r/P$. Bottom panels: microscopic retrieval at the cut-off value, maximum α_R, vertical red line in top panels. Individual pattern overlaps m^μ (blue filled circles). Blue dotted line, mean retrieval overlap $M = \langle m^\mu \rangle$, red dashed line retrieval threshold θ_r.

Erdös-Rényi random network (Löwe and Vermet, 2011)). The transition from retrieval $M \sim 1, R \sim 1$ to non-retrieval $M \sim 0, R \sim 0$ phases is more discontinuous for the single network than for the ensemble systems in top middle and right panels. The maximum α_R for the single network occurs at $P = 22$ patterns loaded, with a number of retrieved patterns over $P_r = 20$ occurs above the retrieval threshold $\theta_r = 0.7$, that is, $P_r = 20$ have individuals overlaps $m^\mu > \theta_r$. Thus, the retrieval efficiency achieved at the critical point for the single network systems is $R = P_r/P = 20/22 = 0.91$, with a mean retrieval overlap $M = \langle m \rangle_\mu = 0.91$. The individual pattern retrieval m^μ for the single network at the critical point, is depicted in Fig. 3 bottom-left panel, for $P = 22$, where $P_r = 20$ patterns have values of $m^\mu > \theta_r$, where $\theta_r = 0.7$ is depicted as red dashed horizontal line, two patterns are retrieved below such line corresponding to patterns $\mu \in \{18, 21\}$ with $m^\mu \approx 0$. The mean retrieval overlap $M = 0.91$ is depicted as a blue dotted

line. Note that the values of $M = 0.91$ and $R = 20/22$ correspond to the points intersecting the red vertical line at the maximum value of α_R, in the top-left panel. Fig. 3 top-middle panel depicts the retrieval performance for an ensemble with $N = 89420,\ K_b = 24,\ n = 10$ with random shortcut ratio $\omega = 0.8$, starting from a 1-D ring. The transition from retrieval to non-retrieval, maximum value of α_R, occurs at a number of loaded patterns $P = 60$, where $P_r = 54$ patterns are retrieved over $\theta_r = 0.7$, for a retrieval efficiency $R = 54/60 = 0.9$, with a mean retrieval overlap $M = 0.81$. Again, the corresponding individuals overlaps m^μ at the critical point are depicted in the bottom-middle panel. Six patterns are considered as not retrieved with individuals overlaps $m^\mu < 0.2$. Fig. 3 top-right panel shows the retrieval performance for an ensemble with $N = 89420,\ K_b = 20,\ n = 12$ with random shortcut ratio $\omega = 0.6$, starting from a 2-D cross grid. The transition from retrieval to non-retrieval, maximum value of α_R, occurs at a number of loaded patterns $P = 72$, where $P_r = 65$ patterns are retrieved over $\theta_r = 0.7$, for a retrieval efficiency $R = 65/72 = 0.9$, with a mean retrieval overlap $M = 0.78$. The corresponding individuals overlaps m^μ at the critical point are plotted in the bottom-middle panel. Seven patterns are considered as not retrieved with individuals overlaps $m^\mu < 0.55$. Note that the ensemble systems receive at each learning/retrieval step a number of patterns equivalent to the number of modules n, thus the curves for higher values of modules n have less points than curves lower modules.

The single system is used as a reference for calculating the gain $G = P_r^e/P_r^s$ achieved by the ensemble systems. For the single network $n = 1$ in Fig. 3 left panels, the number of retrieved patterns $P_r^s = 20$ can be used to calculate the gain for the ensembles, which for the middle panels corresponds to $G = 54/20 = 2.7$, that is the ensemble with $N = 89420,\ K_b = 24,\ n = 10,\ \omega = 0.8,\ 1D - ring$ manage to retrieve 2.7 more patterns than the single network. For the right panels the ensemble with $N = 89420,\ K_b = 20,\ n = 12,\ \omega = 0.6,\ 2D - cross$, the gain is $G = 65/20 = 3.25$ indicating that the ensemble manages to retrieve 3.25 more patterns than the single network.

From this introducing figure, one may conclude that the retrieval performance is optimized for an ensemble with modules starting in a 2D-cross configuration, for intermediate values of the randomness shortcut parameter $\omega = 0.6$. The optimization occurs for a given number of modules $n = 12$ with very diluted connectivity, however for extreme dilution, which is called as pathological dilution (larger number of modules), the performance worsens, given that for practical applications (experimental) such extreme dilutions

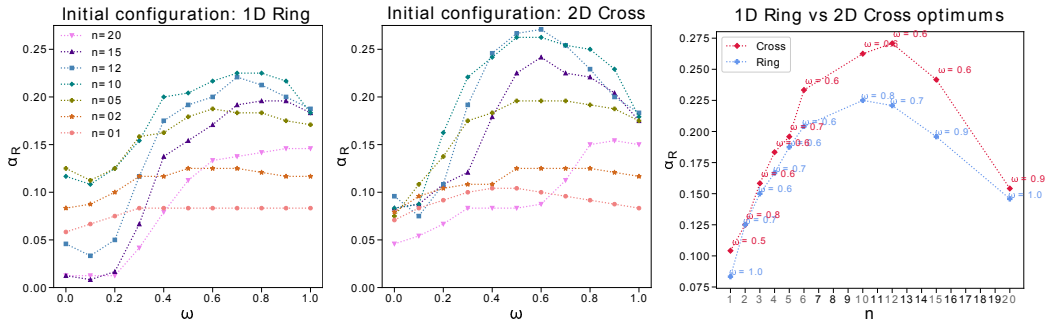

Figure 4: *Optimal topology parameter ω in the ANN ensemble for number of modules in 1D Ring and 2D Cross topologies.* Left and middle panels: Retrieval performance for topology parameter ω, initial network structure and number of ensemble modules for ring vs cross. Right panel: Comparison of best results in left and middle panels (i.e. optimums of retrieval performance for each topology, according to maximum α_R for the parameter ω and number of modules).

induce the network to non-retrieval states.

3.2. Ensemble optimal topology parameter ω for number of modules

Fig. 4 shows the extensive search analysis performed to identify the optimal parameters in terms of the random shortcuts ratio ω and the number of modules n in the ensemble. The analyzed parameters values are $\omega \in \{0.0, 0.1, 0.2, \ldots 1.0\}$. For $\omega = 0.0$ one gets a regular nearest neighbors 1D-ring or a 2D-cross network; and for $\omega = 1.0$, a random network. Intermediate values of the parameter ω represent the fraction of random shortcuts rewired from the initial regular configuration. The searched values for the number of modules are $n \in \{1, 2, 3, 4, 5, 6, 10, 12, 15, 20\}$.

In Fig. 4 left-panel the analysis performed is shown for the initial 1D-ring configuration, for the different values of the ω parameter. The retrieval capacity α_R is depicted in y-axis. Different values of the number of modules n are tested and depicted with different line colors and markers. The general behavior is that the retrieval capacity α_R increases with the fraction of random shortcuts ω starting from the regular topology 1D-Ring, and the retrieval capacity increases with the number of modules achieving a maximum capacity for $n = 10$ and then decreasing when the number of modules is larger implying very diluted modules where the pathological dilution translates to very low ensemble capacity. A similar analysis is performed for the 2D-cross initial configuration and is shown in Fig. 4 middle-panel. Again the

general behavior is that the retrieval capacity α_R increases with the value of ω, but now a maximum occurs for $n = 12$ modules with $\omega = 0.6$ and the capacity decreases for larger values, indicating an optimization of the network capacity occurs for intermediate values of the ω parameter. This behavior is more moderate for the 1D-ring in the left panel. The maximum capacity occurs for $n = 12$ modules decreasing for larger values due to the pathological dilution of each component in the ensemble. Fig. 4 right-panel summarizes the results for the 1D-ring and 2D-cross initial configurations in terms of the number of modules. As commented before the optimal capacity occurs for a given number of modules, showing a non monotonic behavior and maximum can be found for intermediates values of both number of modules n and ω parameters, occurring for $n = 10$ and $n = 12$ modules for the 1D-ring and 2D-cross with $\alpha_R \approx 0.22$ for $\omega = 0.8$ and $\alpha_R \approx 0.27$ for $\omega = 0.6$ respectively. The capacities achieved by the 2D-cross initial configuration are higher, as can be observed when comparing the left and middle panels, or the two curves in the summary plotted in the right panel. Thus the optimal ensemble parameter for learning an retrieving fingerprint patterns is the 2D-cross initial configuration with a random shortcut ratio of $\omega = 0.6$ and $n = 12$ modules. This is consistent with (González et al., 2014), where the 2D structure of the fingerprint pattern images was tested for different topology parameters and configurations in a single attractor network.

Fig. 5 depicts the retrieval performance in terms of M, R, α_R measures, as discussed in Fig. 3 for the critical values indicated with the red vertical line. The retrieval performance is presented for selected values of the ensemble parameters and the optimal topology 2D-cross initial configuration in the summary diagram in Fig. 4 right-panel. Fig. 5 top-left panel, shows that the single network system, $n = 1, K_b = 240$ modules, achieves to retrieve $P_r = 25$ patterns above the selected retrieval threshold $\theta_r = 0.7$. The number of retrieved patterns increases to $P_r = 56$ $n = 6, K_b = 40$ ensembles, in the top second panel; before reaching the maximum retrieval capacity of $P_r = 65$ for the $n = 12, K_b = 20$ ensemble in the top-third panel. For larger number of modules, $n = 15, K_b = 16$, the retrieval capacity starts to degrade, as shown in the top right panel of Fig. 5. Note that the smaller the number of modules in the ensemble, the more discontinuous is the transition between retrieval and non retrieval phases given that the connectivity is less diluted.

The individual retrieval is presented in Fig. 5 middle row panels for the critical value depicted with the vertical red line in the top panels. The larger the number of modules the larger the number of patterns that can be stored,

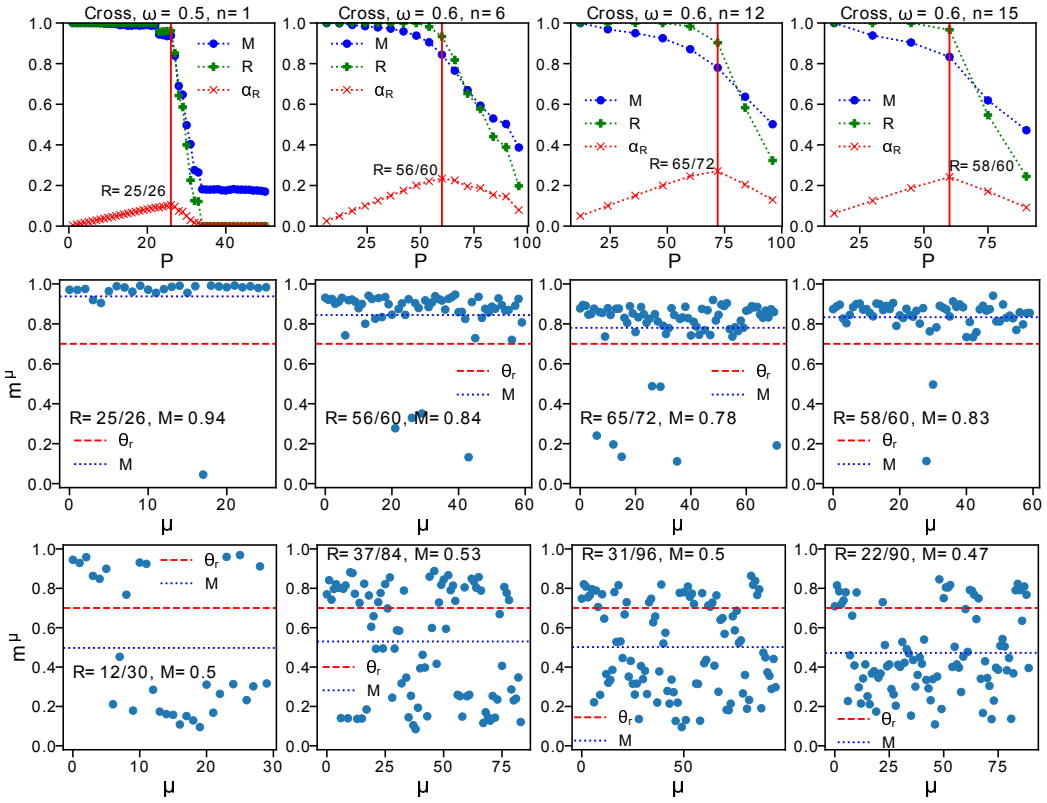

Figure 5: *Retrieval performance for several selected values of the ANN ensemble for optimal topology 2D-cross (see parameter ω for $n = \{1, 6, 12, 15\}$, in Fig. 4 right-panel).* Top panels: Ensemble performance curves M, R, α_R and maximum value cut-off for the pattern load P, indicated with the solid red vertical line, and the fraction of retrieved Patterns $R = P_r/P$. Middle and bottom panels: microscopic retrieval of the individual pattern overlaps m^μ. Blue dotted line, mean retrieval overlap $M = \langle m^\mu \rangle$, red dashed line retrieval threshold θ_r. Middle panels, microscopic retrieval at the cut-off value, maximum α_R, vertical red line in top panels. Bottom panels, microscopic retrieval to the right of the cut-off value. In each panel the value of $R = P_r/P$ indicates the load at each panel and the retrieved patterns.

although the quality of the retrieval is lower than the single network system with $M = 0.94$, the quality of the ensembles remain high with values around $M = 0.8$. Moving to the right of the critical point (right to the vertical red lone), the retrieval degrades for all ensemble systems as depicted in Fig. 5 bottom row panels.

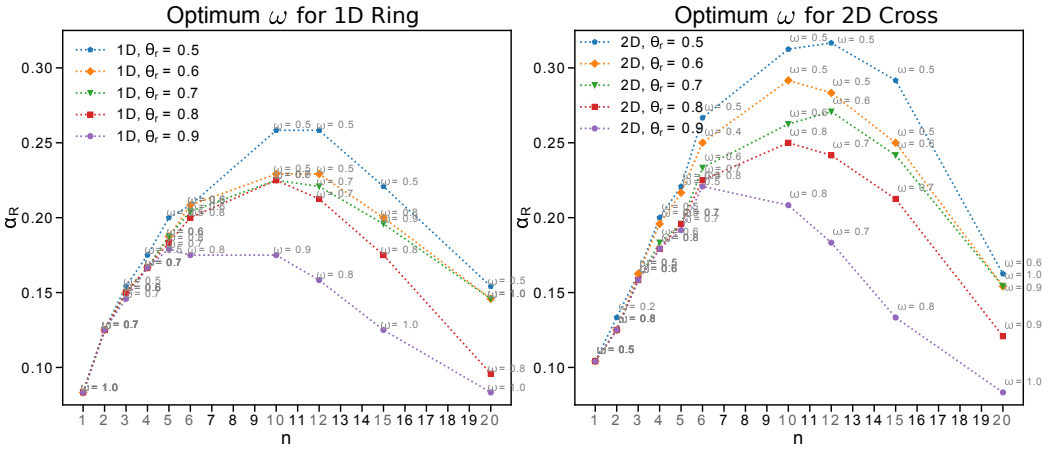

Figure 6: *Optimal retrieval performance for topology parameter ω, initial network structure and number of ensemble modules for different retrieval threshold.* Left: Ring initial topology structure. Right: Cross initial topology structure.

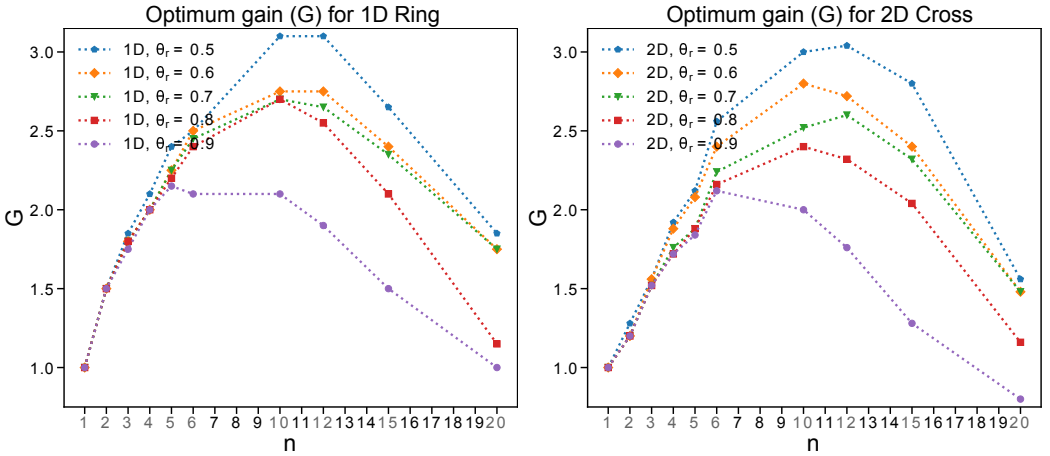

Figure 7: *Optimal gain G for topology parameter ω, initial network structure and number of ensemble modules for different retrieval threshold.* Left: Ring initial topology structure. Right: Cross initial topology structure.

3.3. Effect of the retrieval threshold over the ensemble capacity

Fig. 6 shows the retrieval capacity analysis for different values of the retrieval threshold $\theta_r \in \{0.5, 0.6, 0.7, 0.8, 0.9\}$. The retrieval threshold affects the modules performance in terms of quality of retrieval, however the retrieved pattern are always distinguished from the non retrieved patterns as was checked in (Gonzalez et al., 2017). The panels represents the retrieval

performance in terms of the information parameter α_R for each ensemble systems (module number) and their optimal ω parameter for the 1D-Ring in the top-left panel and the 2D-Cross in the top-right panel. The smaller the retrieval threshold θ_r the larger the retrieval capacity, this effect is more pronounced for the optimal region of the random shortcut parameter around $0.5 \leq \omega \leq 0.8$, and number of modules $n \in \{10, 12\}$. The retrieval capacity α_R is higher for the 2D-Cross initial configuration for all values of θ_R, achieving a maximum of $\alpha_R \sim 0.325$ versus the 1D-Ring configuration, which achieves $\alpha_R \sim 0.27$, for $\theta_R = 0.5$.

Fig. 7 depicts the performance in terms of the pattern gain, G, compared with each single system using as a reference each topology configuration, 1D-Ring single network ($n = 1$) in the left panel, and 2D-Cross single network ($n = 1$) in the right panel. Again the optimal ω parameter is depicted for each number of modules in the ensemble systems. In terms of pattern gain, G, the maximum is similar for both configurations (1D-Ring vs. 2D-Cross), $G \sim 3.1$, for $\theta_r = 0.5$. The gain is measured comparing the single system for for each configuration 1D-Ring vs. 2D-Cross with the respective ensemble for the number of modules n. This gain is similar to the one observed for random patterns, deserving further investigation to discover that it is an invariant measure.

4. Ensemble modules assignment and specialization results

This section presents the results of applying the ensemble of attractor networks to a case of modules specialization, where the assignment of patterns subsets to a module is carried out according the subset features, namely the patterns overlaps and activities.

4.1. Assignment of pattern subsets to ensemble modules

Given the ensemble system separation in modules and the pattern subsetting input, it originates an optimization assignment problem. The correlation between patterns increases the cross-talk noise term between patterns (Hertz et al., 1991; Amit, 1992), an application of the subset to module assignment is to feed the different modules in a way of reducing the overlap between patterns in the subset. Fig. 8 top-left panel, depicts the overlap of each pattern μ with the rest of patterns in the whole set. Each element of the overlap matrix $\mathbf{O}$, which is a $P \times P$ matrix, can be calculated as $O_{\mu\nu} = 1/N \sum_i^N \xi_i^\nu \xi_i^\mu$, indicating the overlap between patterns μ and ν. Once the overlap matrix is

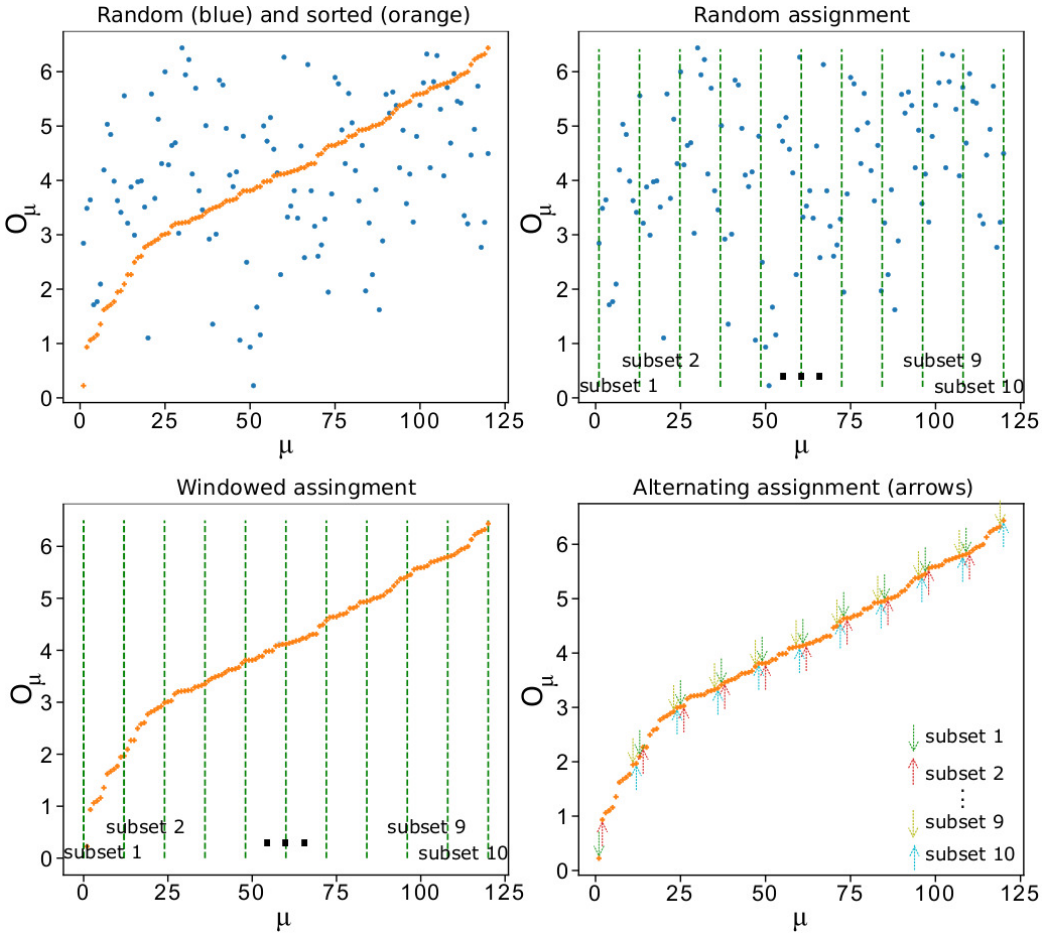

Figure 8: *Different strategies of assignment of pattern subsets to modules of the ensemble system.* O_μ defines the overlap between the pattern μ and the rest of $P-1$ patterns. Ensemble with $n = 10$ modules/subsets. Top-left: Blue points indexing is equivalent to random pattern assignment. The orange crosses represent the patterns ordered from least overlap to greatest overlap between one pattern and the rest. Top-right: Random assignment of subsets to modules, blue points between vertical lines (windows). Bottom-left: Windowed assignment of subsets to modules, orange crosses between vertical lines (windows) Bottom-right: Alternating assignment of subsets to modules, same color arrows.

calculated the sum of each row, subtracting the diagonal element, will give the total overlap of any pattern μ with all the pattern set: $O_\mu = \sum_{\nu:\forall\nu|\nu\neq\mu} O_{\mu\nu}$. This value is depicted, in Fig. 8 top-left panel, as blue circles. In the same panel is depicted the overlap of each pattern with the whole pattern set, which can be sorted in ascending ordered, as plotted with orange crosses,

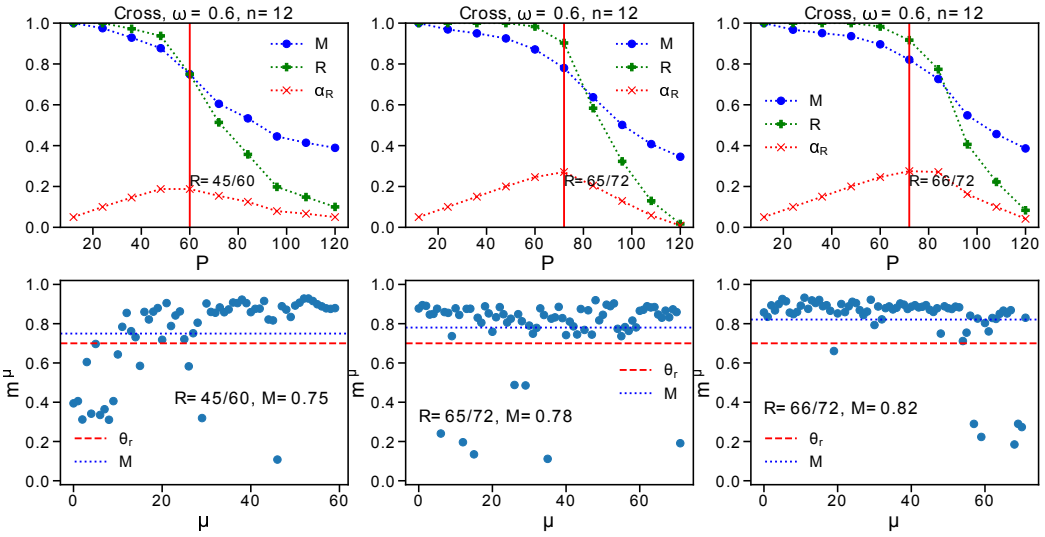

Figure 9: *Retrieval performance for different strategies of assignment of pattern subsets to modules of the ensemble.* Left panels: Windowed subset to module assignment. Center panels: Random assignment. Right panels: Alternating subset to module assignment. For each strategy top row shows the ensemble performance curves M, R, α_R, the maximum value cut-off for the pattern load P, is indicated with the solid red vertical line, and the fraction of retrieved patterns $R = P_r/P$. Finally, bottom row shows the microscopic retrieval for the critical value depicted with the vertical red line in the top panels.

this order gives the pattern that is less correlated with the rest, ascending to the one which is more correlated with the rest of patterns. This correlation value can be used as a criterion to assign the pattern sets to the modules in the ensemble. Three orders are tested, the first order consists on selecting the patterns subsets randomly, as depicted in Fig. 8 top-right panel. The second one, selecting a moving window of size corresponding to the number of modules, $n = 10$ subsets, windowed assignment, as depicted schematically Fig. 8 bottom-left panel. The third one, alternating patterns to build the $n = 10$ subsets, as depicted schematically in Fig. 8 bottom-right. One may conjecture that the third order must be better given that the modules are assigned with equivalent subsets of similarly uncorrelated patterns.

In Fig. 9 the retrieval performance for the different orders is presented. The EANN ensemble parameters are $n = 12, K_b = 20, \omega = 0.6$ corresponding to the optimum found in 4. Given such number of modules, the subsets are also $n = 12$. The worst performance corresponds to the moving window subsetting, presented in Fig. 9-left panel, where the maximum α_R occurs for

a load of $P = 60$, patterns, where the number of retrieved patterns over $\theta_R = 0.7$ is $P_r = 45$. The retrieval performance for the alternating subsetting by window, is similar than a random order, achieving $P_r/P = 65/72$ and $66/72$, respectively. Although, there is an overlap degree between fingerprints, a random order is enough to overcome this overlap as it has similar retrieval performance as the alternating order. The pattern to module assignment is more prominent for more correlated patterns, as in the gesture recognition presented in (Dávila et al., 2019), where a simple random order might not be enough. Further investigation is warranted.

The differences between the different pattern subsets to modules assignment can be observed in Fig. 10, using as performance evaluation an information measure defined as follows. An approximation from (Dominguez and Bollé, 1998; Dominguez et al., 2007, 2009) can be used to calculate the mutual information MI in terms of this mean retrieval overlap M. This will give

$$MI[M] = 1 - S[M], \quad S[M] = -\frac{1+M}{2}\log_2\frac{1+M}{2} - \frac{1-M}{2}\log_2\frac{1-M}{2}. \tag{6}$$

Hence the information ratio can be defined in terms of α_L and M as

$$i_M(\alpha_L, M) \equiv \alpha_L MI[M], \tag{7}$$

where $\alpha_L = \frac{P}{(K_b \times n)}$ is the measure of the pattern load P over the system connectivity.

As observed in Fig. 9 starting form a sorted dataset in terms of correlation with the rest of patterns, the windowed order, blue x-markers line, is the worst assignment given that correlated patterns are assigned to the same modules. Alternating the assignment gives better results, blue circle-markers line. This results, albeit, of being better than a random assignment, orange plus-markers line, the random assignment is almost equivalent, indicating that for the fingerprint dataset the subset to module assignment is not as critical as in more correlated patterns such as the 2D gestures explored in (Dávila et al., 2019).

4.2. Ensemble modules specialization

Fig. 11 shows the retrieval performance for specialized modules in an EANN system $N = 89420, K_b = 80, n = 3$. The specialization occurs according the activity of the pattern subsets as follows, module $b = 1$: skeletonized

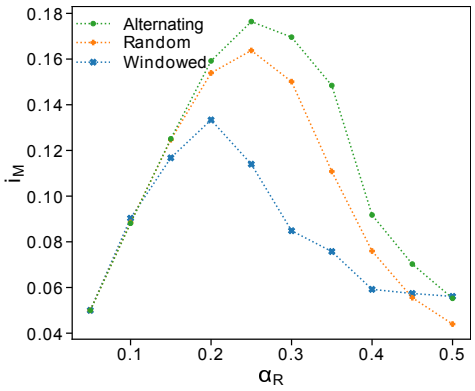

Figure 10: *Information ratio for the different assignment orders analyzed of Fig. 8.* The "alternating" assignment of pattern subsets to modules of the ensemble is the best strategy of the three, according the performance evaluation of the ANN ensemble in terms of information ratio.

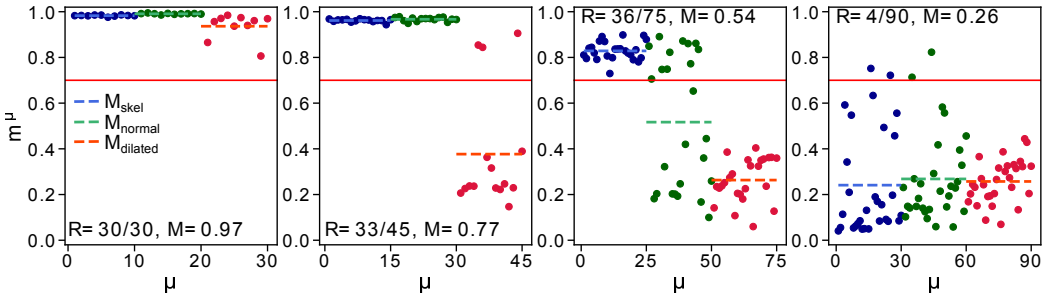

Figure 11: *Pattern specialization for different pattern activities and numbers of pattern loaded to the 3-module ensemble.* Skeletonized ($a \sim 0.0844$) in blue circles and M_{skel} in blue dashed line, normal binarized ($a \sim 0.2258$) in green circles and M_{normal} green dashed line, and dilated fingerprints ($a \sim 0.4660$) in red circles and $M_{dilated}$ in red dashed line. Solid red line corresponds to the retrieval threshold $\theta_r = 0.7$. From left to right increasing number of patterns loaded: $P = \{30, 45, 75, 100\}$ respectively.

fingerprints (activity and neural threshold $a \sim 0.0844, \theta(a) = 1.4951$); module $b = 2$: normal binarized ($a \sim 0.2258, \theta(a) = 0.6560$); and module $b = 3$: thickened fingerprints ($a \sim 0.4660, \theta = 0.0681$). The continuous red line means retrieval threshold overlap of the whole system ($\theta_r = 0.7$). The dashed lines represent the mean retrieval overlap of each module. That is, given the specialization, each module is evaluated for the subset it was assigned with. Note that each module gets a third of the whole pattern set ($n = 3$). In Fig. 11 can observed from left to right for increasing pattern

load $P = \{30, 45, 75, 90\}$ that the EANN system retrieval capacity worsens. However the degradation of the retrieval is not uniform for the modules. The first module assigned with the skeletonized fingerprints subset, $a = 0.0844$, is the one that achieves greater capacity when the number of patterns to learn is increased, i.e. see panel with $P = 75$ in Fig. 11; followed by the second module with normal binarized $a \sim 0.2258$ patterns. The worst is the third module (i.e., the one corresponding to dilated fingerprints) with near unbiased pattern $a \sim 0.4660$. This result is expected as the capacity of the attractor network increases with the sparseness of the pattern/network activity as studied in (Dominguez et al., 2012). One can think of an optimization of the modules connectivity, assigning more connections to the modules to improve the performance. That is, the connections, which are evenly distributed, in the EANN scheme in Fig. 1, can be redistributed assigning more connections to the modules where the performance is worst to increase the overall number of retrieved patterns. The modules that are in charge of patterns with low activity would need fewer connections than those in charge of higher activity patterns. Therefore, the distribution of connections would be made according to the activity of patterns that each module handles within the EANN.

5. Conclusions

The ensemble attractor model (EANN) was extended, in the present work, to deal with structured patterns, i.e. fingerprints. This work has also taken into account the study of i) the optimized number of diluted modules in EANN, and ii) the optimized intra-modular connectivity range in relation to storage capacity of 1D and 2D module topologies in the EANN. The EANN system proved to triple the capacity of a single attractor with the same connectivity and the network topology of modules was optimized. Such behavior is desirable for real patterns applications, such as the fingerprints tested in this work, where the EANN system was capable of retrieving three times the number of fingerprints of the single attractor system used as reference, with the same total connectivity. The performance retrieval was tested for two topology configurations 1D Ring vs 2D Cross, with the latter performing better given the 2D structure of the ridges in the fingerprints. For the topology parameter, random shortcut ratio ω, a non trivial topology, that is an intermediate value $\omega = 0.6$ optimizes the retrieval capacity of the EANN system. Similarly, the performance for increasing number of modules is non-

monotonic, reaching an optimum for $n = 12$ modules, where the connectivity is diluted, but has not reached the level of pathological dilution, where retrieval is hindered.

Given the EANN modular structure and the fact that each module is assigned a subset of patterns, an assignment optimization problem arises. As a proof of concept, the patterns have been divided in subsets according to the overlap of each pattern with the rest of the pattern set. Two extreme cases were used, one that assign windowed subsets, where patterns with high overlap can be assigned to the same EANN module; and the other, reducing the overlap alternating the pattern assignment. The latter performed better given that the overlap between subsets was reduced. This proved to be a helpful characteristic of the EANN system, given that structured patterns, such as the fingerprints, are usually correlated. Taking advantage of the pattern subset to module assignment, another application arises, such as the module specialization according to intrinsic characteristics of the subset. An example for pattern subsets with different activities was presented for skeletonized fingerprints (low activity), normal binarized fingerprints (medium activity), and dilated fingerprints (high activity - unbiased). The module assigned with the skeletonized fingerprints outperformed the other two modules, this expected as the sparse pattern network have a larger storage capacity. With the modules specialization an optimization of the modules parameters, such as connectivity, topology parameters, dynamical parameters can be carried out. As future works, we propose the optimization of the module parameters according to the subsets intrinsic characteristics. Also, the ensemble will be explored for more complex biometric patterns, such as retina blood vessels, where the module specialization can be used in a multi-modal biometrics verification.

Acknowledgments

This work was funded by and UDLA-SIS.MGR.20.01. This research was also funded by the Spanish Ministry of Science and Innovation/FEDER, under the "RETOS" Programme, with grant numbers: TIN2017-84452-R and RTI2018-098019-B-I00; and by the CYTED Network "Ibero-American Thematic Network on ICT Applications for Smart Cities", grant number: 518RT0559.

References

Elena Agliari, Adriano Barra, Andrea Galluzzi, Francesco Guerra, and Francesco Moauro. Multitasking associative networks. *Physical review letters*, 109(26):268101, 2012.

Elena Agliari, Adriano Barra, Andrea Galluzzi, Francesco Guerra, Daniele Tantari, and Flavia Tavani. Metastable states in the hierarchical dyson model drive parallel processing in the hierarchical hopfield network. *Journal of Physics A: Mathematical and Theoretical*, 48(1):015001, 2014.

Elena Agliari, Adriano Barra, Andrea Galluzzi, Francesco Guerra, Daniele Tantari, and Flavia Tavani. Hierarchical neural networks perform both serial and parallel processing. *Neural Networks*, 66:22–35, 2015a.

Elena Agliari, Adriano Barra, Andrea Galluzzi, Francesco Guerra, Daniele Tantari, and Flavia Tavani. Hierarchical neural networks perform both serial and parallel processing. *Neural Networks*, 66:22–35, 2015b.

Daniel J Amit. *Modeling brain function: The world of attractor neural networks*. Cambridge university press, 1992.

Daniel J Amit and Nicolas Brunel. Adequate input for learning in attractor neural networks. *Network: Computation in Neural Systems*, 4(2):177–194, 1993.

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

Carlos Dávila, Mario González, Jorge-Luis Pérez-Medina, David Dominguez, Ángel Sánchez, and Francisco B Rodriguez. Ensemble of attractor networks for 2d gesture retrieval. In *International Work-Conference on Artificial Neural Networks*, pages 488–499. Springer, 2019.

Bernard Derrida, Elizabeth Gardner, and Anne Zippelius. An exactly solvable asymmetric neural network model. *EPL (Europhysics Letters)*, 4(2): 167, 1987.

David Dominguez, Kostadin Koroutchev, Eduardo Serrano, and Francisco B Rodríguez. Information and topology in attractor neural networks. *Neural computation*, 19(4):956–973, 2007.

David Dominguez, Mario González, Eduardo Serrano, and Francisco B Rodríguez. Structured information in small-world neural networks. *Physical Review E*, 79(2):021909, 2009.

David Dominguez, Mario González, Francisco B Rodríguez, Eduardo Serrano, R Erichsen Jr, and WK Theumann. Structured information in sparse-code metric neural networks. *Physica A: Statistical Mechanics and its Applications*, 391(3):799–808, 2012.

DRC Dominguez and Désiré Bollé. Self-control in sparsely coded networks. *Physical review letters*, 80(13):2961, 1998.

Felipe Doria, Rubem Erichsen Jr, Mario González, Francisco B Rodríguez, Ángel Sánchez, and David Dominguez. Structured patterns retrieval using a metric attractor network: Application to fingerprint recognition. *Physica A: Statistical Mechanics and its Applications*, 457:424–436, 2016.

Wenbo Du, Mingyuan Zhang, Wen Ying, Matjaž Perc, Ke Tang, Xianbin Cao, and Dapeng Wu. The networked evolutionary algorithm: A network science perspective. *Applied Mathematics and Computation*, 338: 33–43, 2018. ISSN 0096-3003. doi: https://doi.org/10.1016/j.amc.2018.06.002. URL `http://www.sciencedirect.com/science/article/pii/S0096300318304909`.

Mario González, David Dominguez, Francisco B Rodríguez, and Angel Sanchez. Retrieval of noisy fingerprint patterns using metric attractor networks. *International journal of neural systems*, 24(07):1450025, 2014.

Mario González, María del Mar Alonso-Almeida, Cassio Avila, and David Dominguez. Modeling sustainability report scoring sequences using an attractor network. *Neurocomputing*, 168:1181–1187, 2015.

Mario Gonzalez, David Dominguez, Angel Sanchez, and Francisco B. Rodriguez. Increase attractor capacity using an ensembled neural network. *Expert Systems with Applications*, 71:206 – 215, 2017. ISSN 0957-4174. doi: https://doi.org/10.1016/j.eswa.2016.11.035. URL `http://www.sciencedirect.com/science/article/pii/S0957417416306704`.

Mario González, David Dominguez, Ángel Sánchez, and Francisco B. Rodríguez. Capacity and retrieval of a modular set of diluted attractor networks with respect to the global number of neurons. In Ignacio

Rojas, Gonzalo Joya, and Andreu Catala, editors, *Advances in Computational Intelligence. IWANN 2017. Lecture Notes in Computer Science*, pages 497–506, Cham, 2017. Springer International Publishing. ISBN 978-3-319-59153-7.

Mario González, Carlos Dávila, David Dominguez, Ángel Sánchez, and Francisco B Rodriguez. Fingerprint retrieval using a specialized ensemble of attractor networks. In *International Work-Conference on Artificial Neural Networks*, pages 709–719. Springer, 2019.

J Hertz, J Krogh, and R Palmer. *Introduction to the theory of neural computation*. Santa Fe Institute Studies in the Sciences of Complexity, Vol. 1. Addison-Wesley, 1991.

John J Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the national academy of sciences*, 79(8):2554–2558, 1982.

M Iuzzolino, Y Singer, and M C Mozer. Convolutional bipartite attractor network. *arXiv preprint*, (arXiv:1906.03504v3), 2019.

D P Kingma and M Welling. Auto-encoding variational bayes. *arXiv preprint*, (arXiv:1312.6114), 2013.

Matthias Löwe and Franck Vermet. The hopfield model on a sparse erdös-renyi graph. *Journal of Statistical Physics*, 143(1):205–214, 2011.

Dario Maio, Davide Maltoni, Raffaele Cappelli, Jim L Wayman, and Anil K Jain. Fvc2004: Third fingerprint verification competition. In *International conference on biometric authentication*, pages 1–7. Springer, 2004.

Davide Maltoni, Dario Maio, Anil K Jain, and Salil Prabhakar. *Handbook of fingerprint recognition*. Springer Science & Business Media, 2009.

Haim Sompolinsky and I Kanter. Temporal association in asymmetric neural networks. *Physical review letters*, 57(22):2861, 1986.

P Vincent, H Larochelle, Y Bengio, and P-A Manzagol. Extracting and composing robust features with denoising autoencoders. In *International Conference on Machine Learning (ICML '08)*, pages 1096—-1103. ACM, 2008.

Duncan J Watts and Steven H Strogatz. Collective dynamics of 'small-world' networks. *nature*, 393(6684):440–442, 1998.

B Wemmenhove and ACC Coolen. Finite connectivity attractor neural networks. *Journal of Physics A: Mathematical and General*, 36(37):9617, 2003.

Andy B Yoo, Morris A Jette, and Mark Grondona. Slurm: Simple linux utility for resource management. In *Workshop on Job Scheduling Strategies for Parallel Processing*, pages 44–60. Springer, 2003.