

# Excitatory/Inhibitory Balance Emerges as a Key Factor for RBN Performance, Overriding Attractor Dynamics

Emmanuel Calvet,<sup>1\*</sup> Jean Rouat,<sup>1</sup> Bertrand Reulet<sup>2</sup>

<sup>1</sup>NECOTIS, Génie électrique, Université de Sherbrooke, Canada

<sup>2</sup>Institut Quantique, Département de Physique, Université de Sherbrooke, Canada

\*E-mail: emmanuel.calvet@usherbrooke.ca

**Reservoir computing provides a time and cost-efficient alternative to traditional learning methods. Critical regimes, known as the "edge of chaos," have been found to optimize computational performance in binary neural networks. However, little attention has been devoted to studying reservoir-to-reservoir variability when investigating the link between connectivity, dynamics, and performance. As physical reservoir computers become more prevalent, developing a systematic approach to network design is crucial. In this article, we examine Random Boolean Networks (RBNs) and demonstrate that specific distribution parameters can lead to diverse dynamics near critical points. We identify distinct dynamical attractors and quantify their statistics, revealing that most reservoirs possess a dominant attractor. We then evaluate performance in two challenging tasks, memorization and prediction, and find that a positive excitatory balance produces a critical point with higher memory performance. In comparison, a negative inhibitory balance delivers another critical point with better prediction performance. Interestingly, we show that the intrinsic attractor dynamics have little influence on performance in either case.**

# 1 Introduction

Reservoir Computing (RC) is a promising field for Machine Learning (ML), as the nonlinear reservoir requires no learning and the readout layer only needs linear regression [Maass et al., 2002, Jaeger and Haas, 2004], reducing time and computational cost [Schrauwen et al., 2007]. Furthermore, it has potential for real-world implementations as physical reservoirs and dedicated Neuromorphic chips do not always possess the ability to adapt [Benjamin et al., 2014, Merolla et al., 2014, Tanaka et al., 2019]. Around the same time in 2002, two models were developed: the Liquid State Machine (LSM) [Maass et al., 2002] and the Echo State Network (ESN) [Jaeger, 2001] (rectified version). These approaches differ in their neural models, with LSM using time-event-based neurons and ESNs using Artificial Neural Networks (ANN) with continuous activation functions [Jaeger, 2001]. The Random Boolean Network (RBN) [Bertschinger and Natschläger, 2004], with binary neurons, is a particularly promising model for LSM and allows for a direct relationship between the reservoir design and its performance in a task [Bertschinger and Natschläger, 2004, Natschläger et al., 2005, Snyder et al., 2012]. It is widely used to model and implement reservoirs [Rosin, 2015, Burkow and Tufte, 2016, Echlin et al., 2018, Komkov et al., 2021].

Studies on the RBN have demonstrated the existence of a phase transition in the dynamics of the reservoir for specific connectivity parameters. Close to the critical regime, an increase in performance in solving various tasks has been reported (boolean logic operations [Bertschinger and Natschläger, 2004], bit-parity check [Bertschinger and Natschläger, 2004], prediction of Mackey-Glass time series [Canaday et al., 2018]). As of now, most studies in the field of RC rely on phase diagrams to exhibit a statistical relationship between connectivity, dynamics, and performance [Bertschinger and Natschläger, 2004], [Büsing et al., 2010], [Snyder et al., 2012], [Krauss et al., 2019a], [Metzner and Krauss, 2022]. These results have been obtained by considering a limited number of reservoirs (from one [Metzner and Krauss, 2022], to 10 [Bertschinger and Natschläger, 2004], up to 100 [Krauss et al., 2019b]), and with a limited resolution in terms of the control parameter, due to the computational cost of these phase diagrams.

While phase diagrams are essential to comprehend the full range of the computational capabilities these systems can offer, one crucial point is rarely discussed. Since the reservoirs are randomly generated, there might be huge differences between them even though the statistics of their connectivity are the same. Indeed, close to the critical point, reservoir steady-state activities exhibit a wide range of dynamics as discussed by statistical

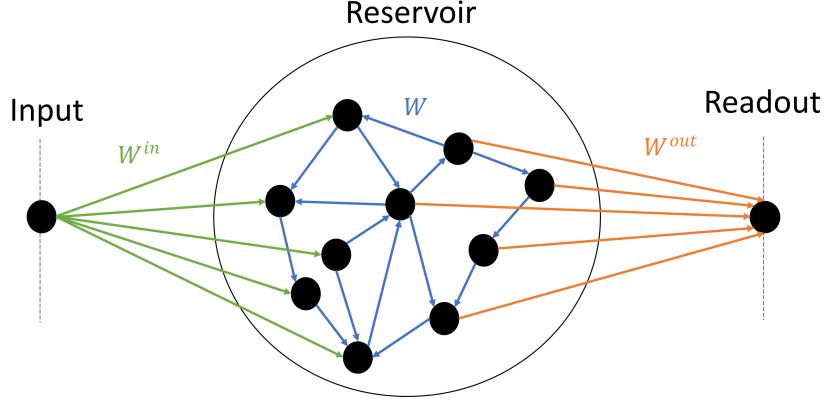
studies [Kinouchi and Copelli, 2006], [Del Papa et al., 2017], [Krauss et al., 2019b], and attractor classification [Seifert and Reggia, 2015], [Bianchi et al., 2016], [Krauss et al., 2019b], [Krauss et al., 2019a], [Metzner and Krauss, 2022].

This article aims at studying the variability of reservoir dynamics, performance, and their correlation. We consider randomly generated RBNs with a single control parameter related to the inhibitory/excitatory balance [Krauss et al., 2019a], tuned with high resolution to perform reliable statistical analysis. We study the excitatory/inhibitory balance, attractor dynamics, and performance, and show that the relationship between the three is more complex than previously thought. In line with the work of [Metzner and Krauss, 2022] on ESN, our research reveals that the RBN also possesses two critical points. Depending on whether the balance is in the majority excitatory or inhibitory, we show that reservoirs respectively exhibit optimal performance in either memory or prediction.

The article is organized as follows: in section 2, we describe the model and prove that it is controlled by the ratio of the standard deviation and mean of the weight distribution (noted  $\sigma^*$ ), which we use to perform all subsequent analyses. In section 3.1, we show that the sign of  $\sigma^*$  produces two critical regimes. In section 3.2, we classify the activity of free-running reservoirs into four classes according to their attractor dynamics for these two critical regimes. We show that each reservoir can be associated with its most dominant attractor. In section 4.1, we evaluate the relationship between connectivity, dominant attractor, and performance in memory and prediction tasks. We then investigate the relationship between the performances of the two tasks, critical regimes, and dominant attractors in section 4.2. This allows us to derive specific recommendations for simplifying the random generation process of reservoirs. Finally, we discuss our findings in section 5 and their implications for future works in section 6.

## 2 Model

The model consists of one input node, the reservoir itself, and an output node (Fig. 1). Half of the neurons inside the reservoirs are connected to the input, and the other half to the readout. Thus information between the input and the readout has to pass through the reservoir. The following subsections describe each component and how they are interconnected.



**FIG. 1** – Schematics of the model. The input node (left) randomly projects synaptic weights to half of the reservoir (center) (green); the reservoir is composed of random recurrent connections (blue); the readout (right) receives input from the other half of the reservoir (orange).

## 2.1 The reservoir

Phase transitions occur *stricto sensu* only in infinite systems, and critical phenomena are easier to observe in large systems [Lavis et al., 2021]. As such, we use an RBN model of size  $N = 10,000$  neurons, which is considerable compared to similar studies in the literature [Natschläger et al., 2005, Büsing et al., 2010, Metzner and Krauss, 2022]. The binary state  $x_i(t) \in \{0, 1\}$  of the neuron  $i$  at the time-step  $t$  (with  $t \in \mathbb{N}$ ), is given by:

$$x_i(t) = \theta \left( u_i(t) + \sum_{j=1}^N w_{ij} x_j(t-1) \right) \quad (1)$$

where  $\theta$  is the Heaviside step function:  $\theta(x) = 1$  if  $x > 0$  and  $\theta(x) = 0$  otherwise. Each neuron receives the same number of non-zero connections  $K = 16$ , in the range of values shown to display sharp phase transitions [Büsing et al., 2010]. The non-zero recurrent weights of the reservoir  $w_{ij}$  (blue arrows in Fig. 1) are i.i.d. and drawn from the Normal or Gaussian density probability function  $\mathcal{N}(\mu, \sigma)$ .  $u_i(t)$  is the external input of the neuron  $i$  at times  $t$ .

## 2.2 Input node

The input layer reduces to one node, receiving the time series  $u(t)$ . The input at times  $t$  of a given neuron  $i$  is:

$$u_i(t) = w_i^{in} u(t) \quad (2)$$

Where the input weight  $w_i^{in}$ , of neuron  $i$  (green arrows in Fig. 1) is drawn from a uniform distribution within  $[-0.5, 0.5]$ , and half of the weights are set to zero. According to Eq. 1, if the amplitude of the input far exceeds the total contribution of the recurrent weights, then the input mostly controls the dynamics. Our choice of parameters corresponds to an input of zero average and  $\sim 0.14$  the standard deviation, which is rather low compared to the recurrent weights. We show in part 4.1 that this choice makes the dynamics mostly controlled by the recurrent weights, which is the intended behaviour.

## 2.3 Readout

The adaptation mechanism is in the output layer only, which reduces here to one linear node with a sigmoid activation function  $f(x) = \frac{1}{1+e^{-x}}$ . As such, the output of the network is given by:

$$y(t) = f(W^{out} \vec{x} + c) \quad (3)$$

Since all experiments consist in reproducing a unidimensional time series, the output  $y$  is a scalar as well. The column vector  $\vec{x}$  represents the state of the reservoir neurons, while the output weights  $W^{out}$  (orange arrows in Fig. 1) are stored in a row vector of size  $N$ , with half of them set to zero. Lastly, the scalar  $c$  is the bias. The training is performed with a mean square error (MSE) loss function. Since we had a focus on collecting high-quality data regarding the link between connectivity and performance, we chose the ADAM optimizer [Kingma and Ba, 2015] over the more standard Ridge regression [Burkow and Tufte, 2016] often used in the literature. The implementation is made with the PyTorch library, and parameters  $\alpha = 0.001$ , and 4000 epochs (See Supplementary Materials 8.4 for more details).

## 2.4 Connectivity: the control parameter $\sigma^*$

To study the reservoir dynamics, one needs the proper definition of a control parameter. Previous work on the RBN often focuses on the average and variance of the recurrent weight matrix ([Bertschinger and Natschläger, 2004], [Natschläger et al., 2005]). In the following, we

demonstrate the existence of only one control parameter  $\sigma^*$  defined by:

$$\sigma^* = \sigma/\mu, \quad \mu \neq 0 \quad (4)$$

Where  $\mu$  is the mean of the weights and  $\sigma$  their standard deviation. Here we study the reservoir in the absence of external excitation,  $u_i(t) = 0$  in Eq. (1). Let us consider two reservoirs with the same architecture, the same initial state, and with respective weights matrices  $W$  and  $\lambda W$  with the scalar  $\lambda > 0$ . Since  $\lambda > 0$ , then  $\theta(\lambda x) = \theta(x), \forall x$ . Thus, according to Eq. (1) for  $u_i(t) = 0$ , the two networks are always in the same state. Thus  $(\lambda\mu, \lambda\sigma)$  gives rise to the same time evolution as  $(\mu, \sigma)$ . The two corresponding reservoirs are totally equivalent. We face two cases depending on  $\mu$ :

- When  $\mu = 0$ , choosing  $\lambda = 1/\sigma$  leads to the conclusion that all reservoirs  $(0, \sigma)$  are strictly equivalent to the reservoir  $(0, 1)$ . Hence reservoirs with  $\mu = 0$  are independent of  $\sigma$ .
- When  $\mu \neq 0$ , choosing  $\lambda = 1/|\mu|$  leads to the weights of the second reservoir distributed with a mean of  $\pm 1$  and a standard deviation  $\sigma/|\mu|$ . Hence we define the control parameter of the RBN as in Eq. (4).

Eq. (4) characterizes the distribution of the weights: the mean is the sign of  $\sigma^*$ , and the standard deviation is its absolute value. Other distribution characterizations directly relate to  $\sigma^*$ . For instance, it is controlling the balance  $b$  between excitation and inhibition, defined by [Krauss et al., 2019a] as:

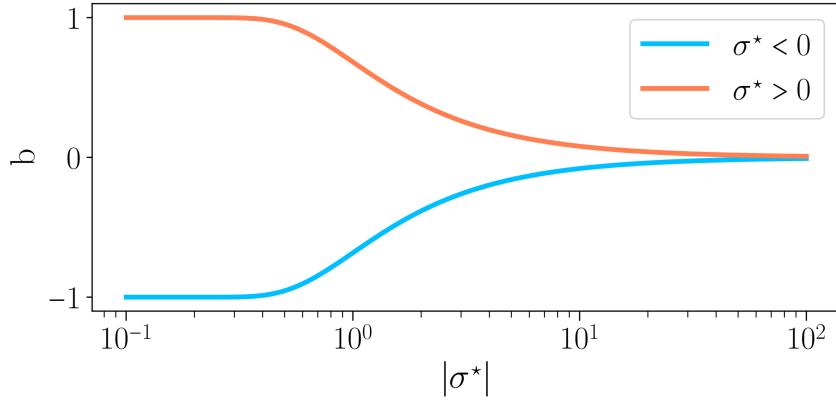
$$b = (S_+ - S_-)/S \quad (5)$$

$$S_{\pm} = \frac{S}{2}(1 \pm b) \quad (6)$$

With  $S = S_+ + S_- = KN$  the total number of synapses,  $S_-$  the number of inhibitory synapses ( $w_{ij} < 0$ ), and  $S_+$  the number of excitatory synapses ( $w_{ij} > 0$ ). By taking a normal weight distribution, the number of excitatory synapses is given by:

$$S_+ = S \int_0^{+\infty} \frac{1}{\sqrt{\pi}} e^{-\left(x - \frac{\mu}{\sqrt{2}\sigma}\right)^2} dx \quad (7)$$

By substituting Eq. (7) in Eq. (6), we find  $b = \text{Erf}[1/(\sqrt{2}\sigma^*)]$ , with Erf the error function. Thus, by controlling the weight distribution, our control parameter  $\sigma^*$  drives the excitatory to inhibitory balance and thus the reservoir dynamics, in line with [Krauss et al., 2019b] and



**FIG. 2** – Excitation/inhibition balance  $b$  as a function of the absolute value of the connectivity parameter  $\sigma^*$ , as defined in Eq. 4, for  $\sigma^* < 0$  (—) and  $\sigma^* > 0$  (—). When the average of weights is positive ( $\sigma^* > 0$ ), whereas the reverse is true when the average of weights is negative.

[Metzner and Krauss, 2022]. Fig. 2 shows the relationship between  $b$  and  $\sigma^*$ . The case  $\mu = 0$  corresponds to  $b = 0$  (perfect balance between excitation and inhibition) and  $\sigma^* \rightarrow \infty$ , for any value of  $\sigma$ . For  $\sigma^* < 0$ ,  $b \in [-1, 0]$ , i.e. there is a majority of inhibitory synapses while for  $\sigma^* > 0$ ,  $b \in [0, 1]$ , hence a majority of excitatory synapses.

To finish, the spectral radius  $\rho(W)$  of the weight matrix  $W$  is a particularly relevant quantity in the context of the ESN, a continuous version of reservoirs, where the critical point corresponds to a spectral radius of one. However, in the case of discontinuous activation functions, such as the one we have with the RBN, it has been shown that the Echo State Property (ESP) cannot be achieved. The spectral radius alone fails to characterize the dynamics and performance of these reservoirs [Oztuik et al., 2007], [Alexandre et al., 2009], [Tieck et al., 2018], [Balafrej et al., 2022]. In supplementary material 8.1, we explicitly discuss the link between  $\rho$  and the mean and variance of the weight matrix and show  $\rho$  is no particular interest in the study of the dynamics.

As a consequence, in the following, we will use  $\sigma^*$  as the unique control parameter (in the range of values displayed in supplementary material S8.2).

### 3 Statistics of dynamics at the critical points

The importance of neural networks dynamics in understanding their performances has been widely explored ([Bertschinger and Natschläger, 2004], [Büsing et al., 2010],

[Krauss et al., 2019a], [Metzner and Krauss, 2022]). The purpose of this part is to investigate the relationship between connectivity and dynamics through two statistical analyses:

1. **Activity statistics:** In section 3.1, we analyze the statistics of the neural activity as a function of the control parameter  $\sigma^*$ . We demonstrate the existence of two critical points and characterize them.
2. **Reservoirs attractors:** In section 3.2, we classify the steady state time evolution of the network activity into four distinct attractors and study the influence of the initial state and random weight generation. We show that reservoirs possess a dominant attractor independent of initial conditions.

### 3.1 Statistics of the activity of free-evolving reservoirs

#### 3.1.1 Methodology:

The first experiment is a free evolution of reservoirs in the absence of input, i.e.  $u_i = 0$ . We define the network activity as  $A(t) = \sum_i x_i(t)/N$ , with  $N$  the number of neurons in the reservoir, and  $A \in [0, 1]$ .  $A$  is also the proportion of excited neurons:  $A = 0$  if the network is extinguished ( $x_i = 0 \quad \forall i$ ),  $A = 1$  if the network is saturated ( $x_i = 1 \quad \forall i$ ). At the initial state, we randomly force 20% of neurons to an up state ( $x_i = 1$ ), i.e.  $A(t=0) = 0.2$ . After a transient regime of 1000 time steps, the reservoir reaches a steady state where we perform statistics. In the following,  $A$  will refer to the activity measured in that steady state (see supplementary material S9 for a more formal definition). For each value of  $\sigma^*$ , we perform statistics on 100 randomly generated reservoirs (see supplementary material S7.1 for more details on the experiment).

#### 3.1.2 Analysis:

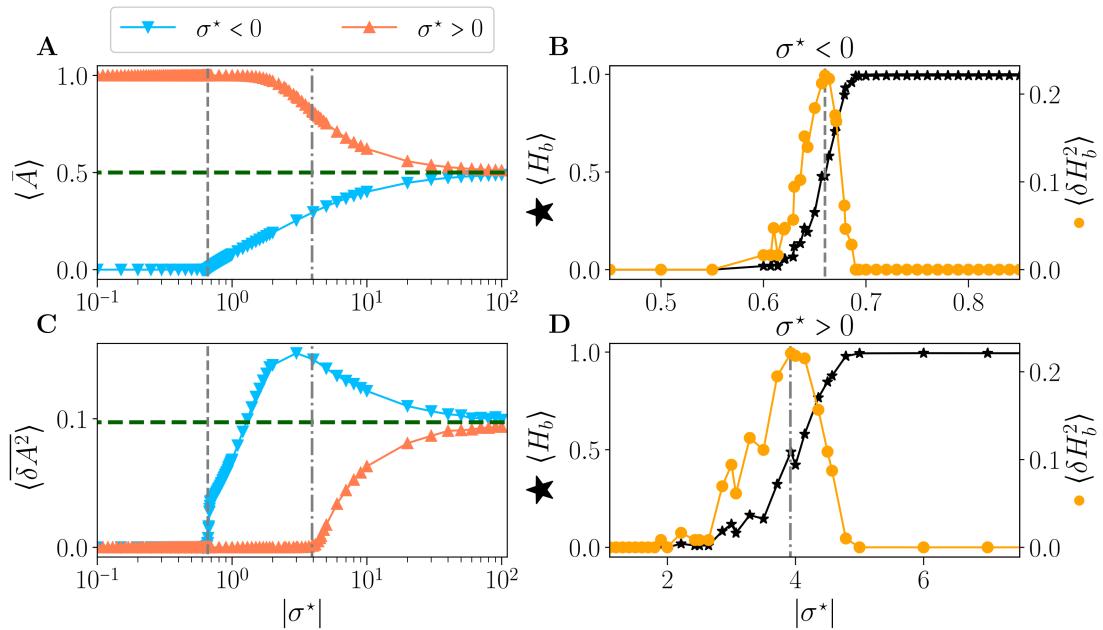
In the following, a bar over a variable  $\overline{(.)}$  represents an average over time for a given reservoir, while the brackets  $\langle . \rangle$  represents an average over different randomly generated reservoirs. In the first analysis, we calculate the time-average steady activity  $\bar{A}$  for a given reservoir and its time-variance  $\overline{\delta A^2}$ , where we define  $\delta A = A - \bar{A}$ . We average these quantities over the reservoirs to give  $\langle \bar{A} \rangle$  and  $\langle \overline{\delta A^2} \rangle$  for each value of  $\sigma^*$ . Next, we evaluate the average and variance over reservoirs of the binary entropy  $H_b$ , or BiEntropy [Croll, 2014] of the time-dependent activity. Compared to the Shanon entropy, the advantage of this metric is that it can discriminate ordered from disordered strings of binary digits. It has been used in machine learning [Mamun et al., 2016], [Zhou and Zeng, 2022], but to our knowledge, this is the first

time in reservoir computing. The binary entropy varies between 0 for fully ordered bit-streams and 1 for fully disordered ones. We compute the BiEntropy of the binarized time dependence of the steady activity for each reservoir (for the exact definition of all the metrics, see S10).

### 3.1.3 Results:

The time-average activity  $\langle \bar{A} \rangle$  as a function of  $\sigma^*$  is shown in Fig. 3A, for both signs of  $\sigma^*$  (Fig. 2). The green dashed line represents the value obtained for  $\mu = 0$ , i.e.  $\sigma \rightarrow \infty$ . The perfect balance in excitation ( $b = 0$ ) results in half of the neurons being activated  $\langle \bar{A} \rangle = 0.5$ . The variance  $\langle \delta \bar{A}^2 \rangle$  vs  $\sigma^*$  is shown in Figs. 3C. For the lowest values of  $|\sigma^*|$ , the reservoirs are frozen (zero variance) either extinguished (for  $\sigma^* < 0$ ) or saturated ( $\sigma^* > 0$ ). This corresponds to reservoirs being respectively purely inhibitory ( $b = -1$ ) or excitatory ( $b = 1$ ). Already at the level of the statistics of the activity, there is a clear difference between both signs of  $\sigma^*$ : for  $\sigma^* < 0$  there is a threshold in  $\sigma^*$  (vertical dashed line at  $\sigma^* \sim -0.7$ ) above which the average activity and its variance rise abruptly and simultaneously. In contrast, for  $\sigma^* > 0$ , there is a wide region where no dynamic is detected (zero variance), yet the network is not saturated but its activity decays continuously. The variance starts rising at  $\sigma^* > 4$  (vertical dash-dotted line).

The average BiEntropy  $\langle H_b \rangle$  vs.  $|\sigma^*|$  is plotted in Figs. 3B for  $\sigma^* < 0$  and 3D for  $\sigma^* > 0$  (left scale, black stars on both plots). These two plots zoom in the vicinity of the phase transition, as statistics are stationary elsewhere. There is a continuous transition between a fully ordered phase ( $\langle H_b \rangle = 0$ ) and a fully disordered one ( $\langle H_b \rangle = 1$ ). Since the BiEntropy is a measure of order, these results suggest that the transition we observe is related to the apparition of chaos in the reservoir above a critical value of  $\sigma^*$  ([Lewin and Bak, 1993], [Bertschinger and Natschläger, 2004], [Seifert and Reggia, 2015], [Kuśmierz et al., 2020]). The variance of the BiEntropy  $\langle \delta H_b^2 \rangle$  is shown in Figs. 3B for  $\sigma^* < 0$  and 3D for  $\sigma^* > 0$  (right scale, orange circles on both plots). It is zero when either in the ordered or disordered phase and spikes at the transition. Its maximum coincides with  $\langle H_b \rangle \simeq 0.5$ : the variance of BiEntropy captures the edge of chaos as a balance between order and disorder. More striking, it also coincides with the position at which the variance of the activity rises (vertical dashed lines). The peak of  $\langle \delta H_b^2 \rangle$  thus provides a clear definition of the position of two critical points:  $\sigma_c^* \simeq -0.66$  and  $\sigma_c^* \simeq 4.0$ , which correspond respectively to critical balances  $b_c \simeq -0.87$  (94% of inhibitory synapses) and  $b_c \simeq 0.19$  (60% of excitatory synapses). Moreover, the transition between order and disorder is much wider for  $\sigma^* > 0$ . This asymmetry between both signs of  $\sigma^*$  is a property of our model since  $\theta(-x) \neq \pm \theta(x)$  in Eq.(1). From now on, we will refer to the *critical points* as the point where the maximum of BiEntropy variance is obtained, and we will define the *critical regions* as the regions with  $\langle \delta H_b^2 \rangle \neq 0$ .



**FIG. 3** – Statistics of the activity of free running reservoirs in the steady state as a function of  $|\sigma^*|$ . Each dot represents the statistics over 100 reservoirs ran once. Average over reservoirs of time average activity  $\langle \bar{A} \rangle$  (A), and average over reservoirs of time variance  $\langle \delta \bar{A}^2 \rangle$  (C), for  $\sigma^* < 0$  (▼) and  $\sigma^* > 0$  (▲). In all plots, the gray vertical lines represent the critical values of the control parameter for  $\sigma_c^* < 0$  (---) and  $\sigma_c^* > 0$  (-···). **B** and **D** zoom on the region of interest close to the critical points: average over reservoirs of BiEntropy  $\langle H_b \rangle$  (\*, left scale) and BiEntropy variance  $\langle \delta H_b^2 \rangle$  (●, right scale), for  $\sigma^* < 0$  (**B**) and  $\sigma^* > 0$  (**D**).

### 3.1.4 Discussion:

Similar to [Krauss et al., 2019b] and [Metzner and Krauss, 2022], the existence of phases separated by critical points as a parameter is varied is reminiscent of the phase diagrams drawn in thermodynamics. If we associate the state of a neuron, 0 or 1, to the state of an Ising spin, either down or up, then  $\langle \bar{A} \rangle$  corresponds to the average magnetization per spin of the network and  $\langle \delta \bar{A}^2 \rangle$  to the variance of its fluctuations, i.e. magnetization noise. At equilibrium, it is proportional to the magnetic susceptibility according to the fluctuation-dissipation theorem [Callen and Welton, 1951]. The total magnetization plays the role of an order parameter, and the transition order is obtained by considering discontinuities, as a function of temperature, of the order parameter and its derivatives with respect to the external field [?]. Here we observe that the average activity is always continuous as a function of  $\sigma^*$ . At the same time,  $\langle \delta \bar{A}^2 \rangle$  is continuous for  $\sigma^* > 0$  but shows a discontinuity at the critical point for  $\sigma^* < 0$ . This strongly suggests that the two "phase transitions" are of a different type.

## 3.2 Dominant attractor of reservoirs

In the previous section, we considered the *average* behaviour of reservoirs: for a given value of  $\sigma^*$  we averaged over many realizations of the distribution of synaptic weights. However, from a practical point of view, one wants to use one network to work with different inputs. This raises two questions: that of the reservoir-to-reservoir variability (do all reservoirs behave similarly?) and that of the sensitivity of a given reservoir to initial conditions. We address these questions in this section.

### 3.2.1 Methodology:

We submitted our reservoirs again to a free evolution without input ( $u_i(t) = 0$ ). For each value of  $\sigma^*$ , we created 100 reservoirs with randomly tossed weight matrices. Each reservoir is run 100 times, with a different random initial state, of activity  $A(t=0) = 0.2$  (for more details, see *Statistics of reservoirs* in S7).

### 3.2.2 Analysis:

We classify the attractor obtained in the steady-state activities, as proposed in [Krauss et al., 2019b]. We categorize the activity signals into one of the four types of attractors (see S11 for a grounded justification of each category):

- *Extinguished* activity: The steady-state activity  $A(t)$  is always zero. This means that the initial activity died during the transient phase and that the reservoir could not propagate it further in time. For simplicity, we will sometime refer to it as *dead* attractor.
- *Fixed* point attractor: The steady-state reservoir is active ( $A(t) \neq 0$ ), but the activity is independent of time ( $\delta A(t) = 0$ ). This includes the *saturated* states  $A(t) = 1$  of ref. [Seifter and Reggia, 2015].
- *Cyclic* attractor:  $A(t)$  is periodic with a periodicity larger than one time-step.
- *Irregular* attractor:  $A(t)$  is neither constant nor periodic within the duration of the simulation. Note that since the RBN is finite, discrete, and deterministic, given enough time, any sequence of states should eventually repeat, taking at most  $2^N$  time steps.

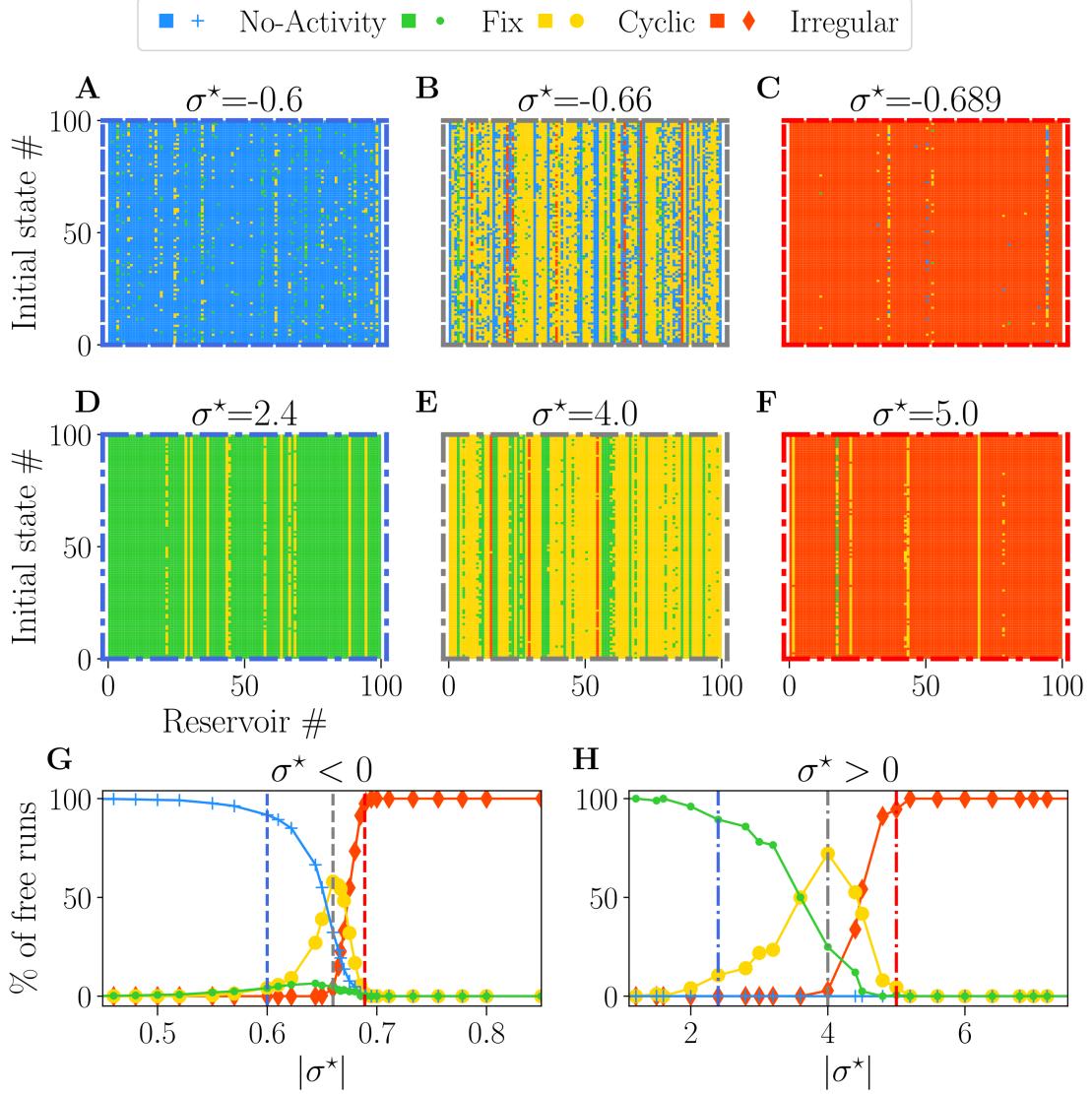
We determine the attractor obtained at the steady state for each reservoir and initial condition. We then compute the distribution of attractors for each value of  $\sigma^*$  obtained overall the initial conditions of all reservoirs. The statistics are thus computed on 10,000 steady activities for each  $\sigma^*$ .

### 3.2.3 Results:

Fig. 4A-F provide examples of attractors, encoded in the colours, for each reservoir (x-axis) and each initial condition (y-axis) for different values of  $\sigma^*$ . The left column (blue-bordered boxes) corresponds to values below the critical point (vertical blue lines on Figs. 4G,H), the center column (gray-bordered boxes) to values at the critical point (vertical gray lines), and the right column (red-bordered boxes) to values above the critical points (vertical red lines). The upper row displays negative  $\sigma^*$  values, while the lower row features positive  $\sigma^*$  values.

Away from the critical point, a dominant colour is observed, meaning that reservoirs exhibit a dominant attractor. Steady activities are predominantly extinguished and fixed on the left side of the critical point (A, D) and irregular at the right (C, F). Close to the critical points (B and E), there is an increase in the diversity of attractors, as previously observed [Karimipanah et al., 2017].

Fig. 4.G and H show the statistical distribution of all obtained attractors versus  $|\sigma^*|$ . As expected from the previous analysis, there is no attractor diversity on the far left and right of the plots, as we obtain one primary attractor. Dead (blue line) or fixed (green line) attractors are found for low values of  $|\sigma^*|$ , and their proportion decays slowly across the transition. Within the critical region coexist all attractors in various proportions. Chaotic attractors start to appear precisely at the transition (vertical gray lines), while the domain where cyclic attractors exist



**FIG. 4** – The attractor landscape of reservoirs: for  $\sigma^* < 0$  (**A**, **B**, **C**, **G**), and for  $\sigma^* > 0$  (**D**, **E**, **F**, **H**). The influence of initial conditions for specific values of  $\sigma^*$ : (**A**)  $\sigma^* = -0.6$ , (**B**)  $\sigma^* = -0.66$ , (**C**)  $\sigma^* = -0.689$ , (**D**)  $\sigma^* = 2.4$ , (**E**)  $\sigma^* = 4.0$ , (**F**)  $\sigma^* = 5.0$ . In each plot, the vertical axis represents different numbers (#) of initial random states of free-running reservoirs. The horizontal axis represents different numbers (#) of reservoirs with various initial weight tossing, randomly generated with distinct seeds. Pixels of colours represent the attractor obtained at the steady state, with the same colours as (**G**, **H**). **G**, **H**: percentage of steady-state activities belonging to each category of attractors: no-activity (+), fix (\*), cyclic (●), and irregular (◆). Each dot represents the statistics over 100 reservoirs ran 100 times, hence 1000 runs. Each (**G**)  $\sigma^* < 0$ , and (**H**)  $\sigma^* > 0$ . In each row (**A**, **B**, **C**) and (**E**, **F**, **G**), the coloured dashed boxes surrounding the plots correspond to the values of  $\sigma^*$ , indicated as vertical lines in plots **G** and **H**.

coincides with the critical region of nonzero BiEntropy variance (Fig. 3.B,D). The point at which cyclic attractors are most present is also precisely  $\sigma_c^*$ . These results corroborate what we inferred in the previous section: on the disordered phase  $|\sigma^*| > |\sigma_c^*|$ , attractors are irregular, while the ordered phase is characterized by fixed or dead attractors. We note an asymmetry between both sides of the transition: irregular attractors appear only in the disordered phase. From the point of view of the attractors, both signs of  $\sigma^*$  lead to similar behaviours, except again, that the transition region is much wider for  $\sigma^* > 0$ .

### 3.2.4 Discussion:

Our results suggest that the critical points enhance sensitivity to the initial states and configuration of the weights, explaining the reservoir-to-reservoir variance and increase in dynamic diversity. Reservoirs around the negative  $\sigma_c^*$  (Fig. 4.B) possessed a distribution of attractors with far more variety than the one with positive  $\sigma_c^*$  (Fig. 4.E), further reinforcing the idea that the sign of  $\sigma^*$  produces two distinct types of critical regimes. We quantified this in the supplementary material S12 by computing the entropy of reservoir attractor distributions plotted in Fig. 9. We interpret that result by suggesting that inhibition might be a key factor for enhancing dynamic diversity.

For the purpose of reservoir design, our findings suggest that with both critical points, most reservoirs possess an attractor obtained predominantly in most trials, independent of the initial state. The statistics of dominant reservoir attractors are presented in supplementary Fig. 9, and found to be similar to the one in Fig. 3. The presence of vertical colour lines in Fig. 4A-F means that, in most cases, the behaviour of the reservoirs does not depend on the initial state, even in the critical region (this is more thoroughly shown in the supplementary material S12). As a consequence, a dominant attractor can be associated with each reservoir, irrespective of the initial condition.

## 4 What drives performances

In this section, we examine whether there is a relationship between reservoir dynamics in the absence of input, as explored in the previous section, and its ability to perform two demanding tasks: memory and prediction. This is done in two steps:

- **Connectivity, attractor, and performance:** In section 4.1, we analyze the performance obtained separately in each task, depending on the control parameter, for each domi-

nant attractor category. We show that the key factor driving performance is the excitatory/inhibitory balance.

- **Attractor and cross-task performance:** In section 4.2, we analyze all reservoir performances independently of the control parameter. For each reservoir, we study the relationship between the performance obtained in each task and their dominant attractor. This allows us to deduce how to generate a reservoir for the best general purpose.

From now on, and for ease of notation, a reservoir with a dominant attractor obtained during free evolution (defined in previous section 3.2) will be referred to as either: a *extinguished*, *fix*, *cyclic*, or *irregular* reservoir (e.g., an *extinguished reservoir* refers to a reservoir with an extinguished dominant attractor).

## 4.1 Performance in memory and prediction tasks

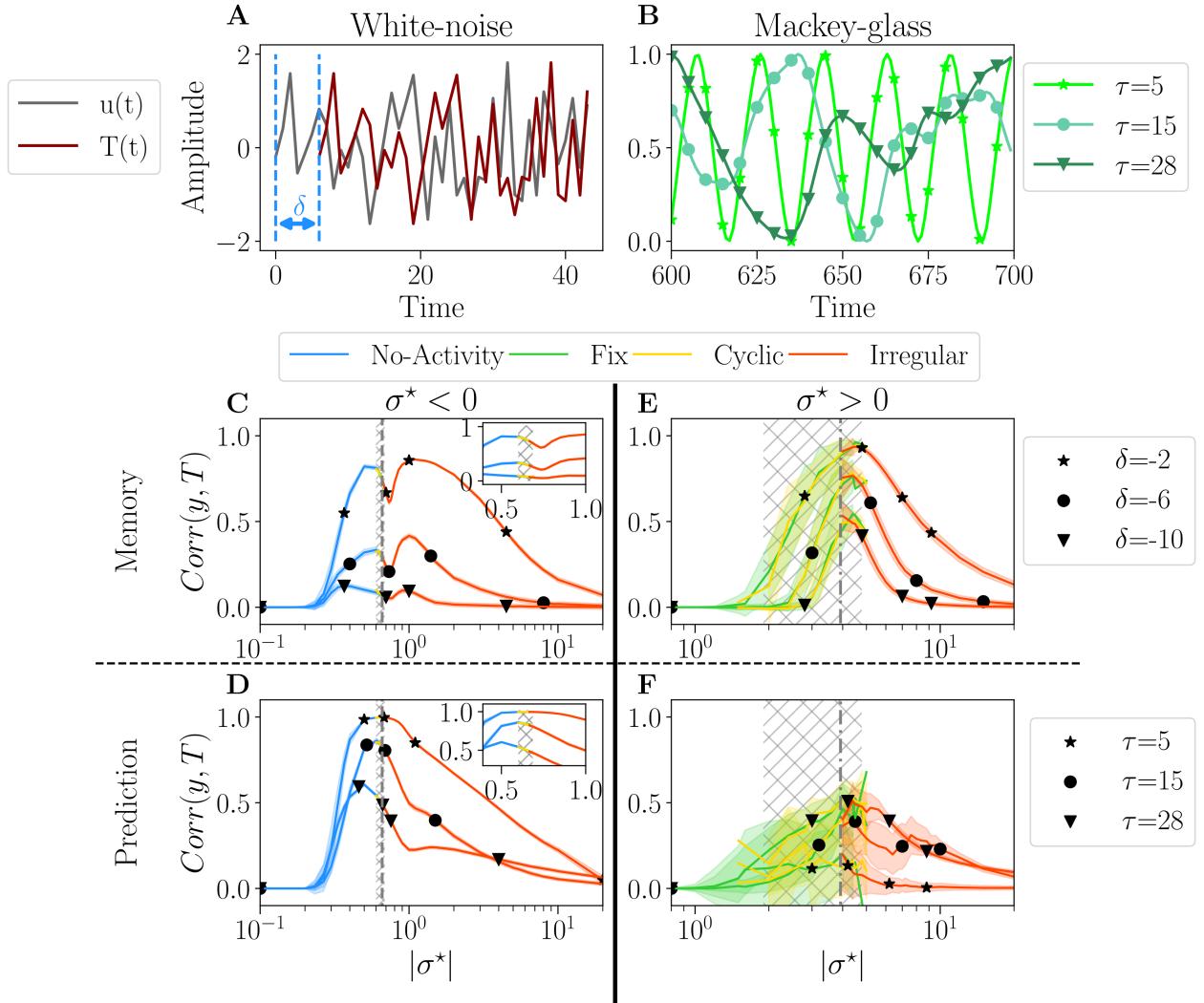
Close to the critical points, we obtained various dominant attractors for a single value of  $\sigma^*$ . This raises an important question regarding the relationship between the dominant attractor of a reservoir and its performance. Specifically, it is worth investigating whether the dominant attractor influences the reservoir’s performance. If this is the case, grouping attractor categories by discrete performance levels may be possible based on a single value of  $\sigma^*$ .

### 4.1.1 Methodology:

We evaluate the performance of the networks to execute two fundamental tasks: *memory* and *prediction*. Each reservoir receives an input  $u(t)$ , and the readout target is  $T(t) = u(t + \delta)$ , equal to the input shifted in time by  $\delta$  time steps.  $\delta < 0$  corresponds to a memory task, and  $\delta > 0$  to a prediction task. For each value of  $\sigma^*$ , we use 100 reservoirs, and each reservoir is run five times, with a different random tossing of the input weight matrix (more detail on the training procedure in S8.4).

The first task consists of memorizing a purely random signal (i.e., uncorrelated white noise), and since there is absolutely no correlation in the input, only memorization is involved. Fig. 5.A illustrates this task for one value of  $\delta = -6$ , with white noise as input  $u$ , and the target  $T$ .

For the second task, we explore the ability of the reservoir to predict a time series,  $\delta = 10$  time steps in the future. The input is the well-known Mackey-Glass time series, as it is a common benchmark of this type of task ([Hajnal and Lörincz, 2006], [Goudarzi et al., 2016], [Canaday et al., 2018], among others), notably testing the ability to infer non-linear dynamics.



**FIG. 5** – Performances for two tasks: white noise memory (**C, E**); and Mackey-glass prediction (**D, F**). **A** and **D**: examples of signals for each task with their respective parameters. **A**: White noise memory task, which consists in remembering the input (gray), to reproduce it in output (dark red) with a negative delay  $\delta$  (shown example corresponds to  $\delta = -6$ ). **B**: Mackey glass is controlled by the parameter  $\tau$  (see methodology 8.3 for more details), ranging from periodic to chaotic. **C, D, E, F**: the average performance  $Corr(y, T)$  between the output  $y$  and target  $T$ , plotted over  $|\sigma^*|$ , for each dominant attractor category: *no-activity*, *fix*, *cyclic* or *irregular*. For each value of  $\sigma^*$  we have 100 reservoirs. The solid line then represents the average over reservoirs belonging to the same attractor category; individual reservoir performances are averaged over 5 initial conditions. The shaded area represents one standard deviation. Higher correlations indicate better performance. The hatched gray area represents the critical regions, as defined in section 3.1. **C** and **D**: the performance in the white-noise memory task; three values of  $\delta$  are tested:  $-2$  ( $\star$ ),  $-6$  ( $\bullet$ ),  $-10$  ( $\blacktriangledown$ ). **E** and **F**: the performance of Mackey-Glass prediction ( $\delta = +10$ ); three values of  $\tau$  are tested:  $5$  ( $\star$ ),  $20$  ( $\bullet$ ),  $50$  ( $\blacktriangledown$ ). **C** and **D**: Performance for  $\sigma^* < 0$ , with inside each plot a zoom on the critical region. **E** and **F**: performance for  $\sigma^* > 0$ .

The signal regularity is controlled by the parameter  $\tau$ , see Fig. 5.B, ranging from periodic with  $\tau = 5$ , to chaotic for  $\tau = 28$  (more information on the experiments in S8.3).

#### 4.1.2 Analysis:

The performance of a reservoir is measured by computing the correlation product  $Corr(y, T)$  between the output  $y$  and the target  $T$ . A perfect match corresponds to a correlation of one, while a random output gives a zero correlation. An individual reservoir performance score is then obtained by averaging over the initial conditions. Each individual reservoir is associated with its dominant attractor, and the statistics of the performance of reservoirs are performed separately for each attractor.

#### 4.1.3 Results:

The average performance is plotted as a function of  $|\sigma^*|$  in Fig. 5.C,E for the memory task and in Fig. 5.D,F for the prediction task. The left column (**C** and **D**) corresponds to  $\sigma^* > 0$ , and the right column (**E** and **F**) to  $\sigma^* < 0$ . The colour of the lines corresponds to the attractor.

For  $\sigma^* < 0$  (Fig. 5.C and D), performance increases over a very wide range of  $\sigma^*$ , both for memory and prediction. This range includes the critical region (gray hatched area) but is vastly broader. Thus, being within the critical region is absolutely not mandatory to perform well. A shaded area in Figs. 5 indicates the spreading of the results. There is none in plots **C** and **D**, meaning that all reservoirs perform exactly the same for a given  $\sigma^*$ . Moreover, as  $\sigma^*$  is increased through the critical region, the dominant attractors change (see zooms in plot **C** and **D**), but surprisingly, there is no discontinuity in the performance. Indeed, even though the four attractor categories are present inside the gray area, their respective performance all align. This strongly suggests that the dynamics of the reservoir, as measured in the absence of input, is irrelevant for the performance. Only the value of  $\sigma^*$  matters. In both tasks, the average performance decreases monotonically with increasing difficulty via  $\tau$  and  $\delta$ . In the memory task, we register a dip in performance close to the critical point. This goes against the common assumption that the edge of chaos is optimal for memory [Natschläger et al., 2005]. In the prediction task, the peak of performance roughly coincides with the critical region, except for the greater difficulty, where the peak is slightly on the left.

The picture is very different for  $\sigma^* > 0$  (Figs. 5.E and F). First, the region in which some level of performance is observed is comparable to the critical region observed in free-running reservoirs. Second, there is substantial variability in performance across different reservoirs, as

indicated by the large shaded areas. Despite this variability, there is an overall dependence of performance on  $\sigma^*$ . The average performance of distinct dominant attractor categories is much noisier. However, despite being more noisy, the average performance of the distinct attractor categories aligns again, so there is still no evidence that the attractor category has any significant impact on the reservoir's performance. We observe that performance decreases as the difficulty of the memorization task increases, but interestingly, this trend appears to be inverted in the prediction task.

Once again, the two signs of  $\sigma^*$  give rise to different behaviour. In particular, networks with  $\sigma^* > 0$  memorize better and are less reliable than those with  $\sigma^* < 0$  but have poorer prediction capability. Yet, in all cases, attractors do not seem to be correlated to performance, as the top performance can be found in any of the four attractor categories.

#### 4.1.4 Discussion:

Our results somewhat challenge the common assumption that the edge of chaos is optimal for performance and suggest that this is true for reservoirs with a majority of excitation but not necessarily with a majority of inhibition. Reservoirs with negative  $\sigma^*$  exhibit very reliable performances with very low reservoir-to-reservoir variability over a range in  $\sigma^*$  much broader than the critical region. Since reservoirs behave the same, in practice, it is sufficient to generate one, with  $\sigma^*$  at the left of the critical region. However, if the goal is optimal memorization, it is wiser to choose  $\sigma \sim 0.4$  in the critical region and try different reservoirs until finding a good one, which requires training and testing.

## 4.2 Cross-task performance

Beyond studying the performance in memorization and prediction separately, as often done ([Bertschinger and Natschläger, 2004], [Büsing et al., 2010]), here we aim at answering the following question: are the reservoirs intrinsically good or bad, or does it depend on the task? In other words, are there general-purpose reservoirs and specialized ones?

#### 4.2.1 Analysis:

We analyze the absolute value of the performance of all reservoirs independently of the control parameter. For each reservoir, we study its performance in the memory task as a function of its performance in the prediction tasks (see section 4.1 *methodology*). For this, we fixed values of  $\delta$

(memory) and  $\tau$  (prediction), and we chose three levels of difficulty: simple ( $\tau = 5, \delta = -2$ ), average ( $\tau = 20, \delta = -6$ ) and difficult ( $\tau = 28, \delta = -10$ ).

#### 4.2.2 Results:

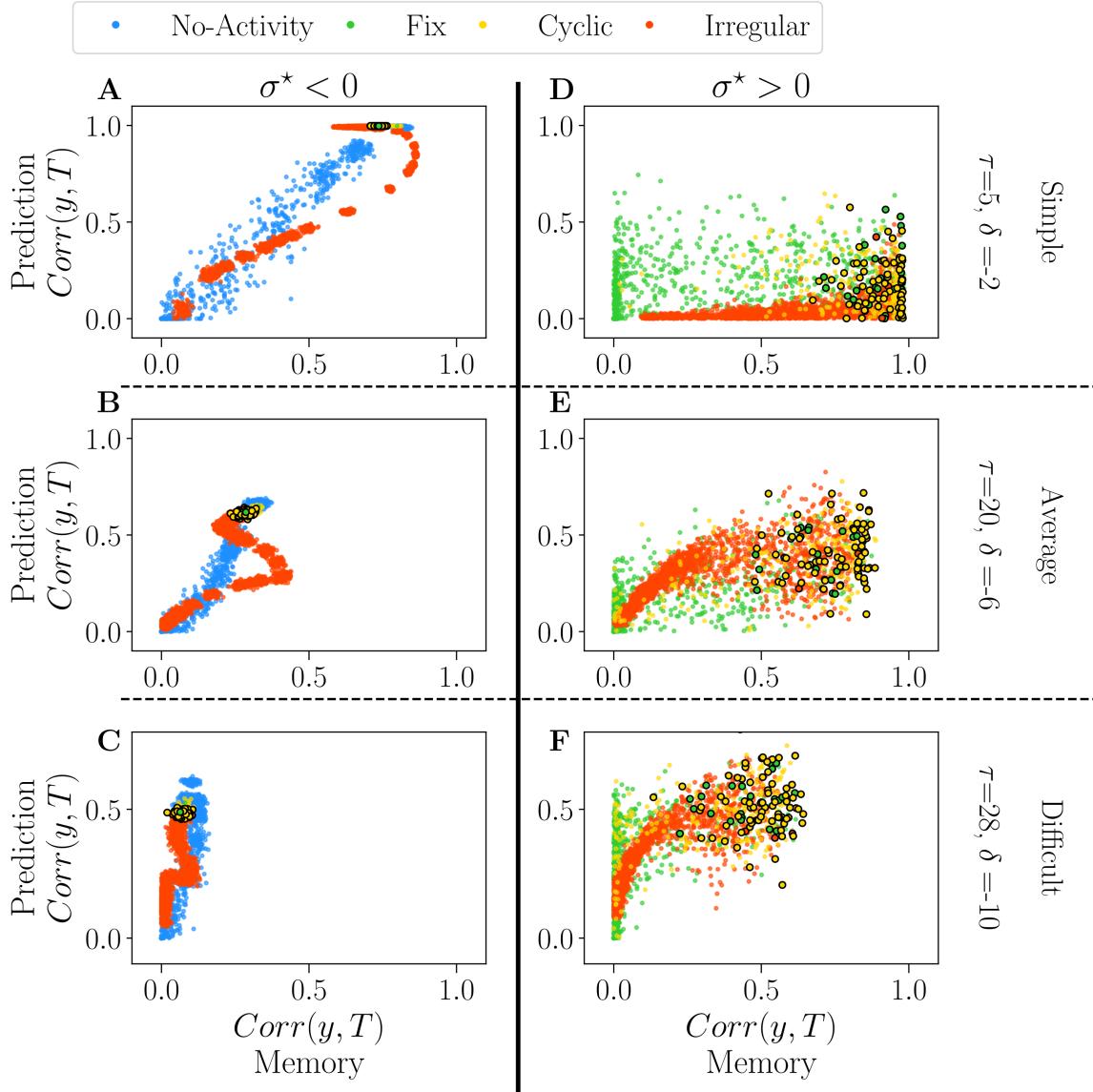
Fig. 6 displays the performance of reservoirs as coloured dots. As in the previous section, each colour corresponds to the dominant attractor of the reservoir. The vertical axis represents the performance in the prediction task (Mackey-Glass), and the horizontal axis that of the memory task (white noise). The columns correspond to the sign of  $\sigma^*$ , negative on the left, positive on the right. The rows correspond to the degree of difficulty, from simple (top) to difficult (bottom).

For  $\sigma^* < 0$  (left column), there is an apparent relationship in performance between the two tasks. The degree of difficulty roughly acts as a scaling factor on the curves, but the correlation is always clear. The reservoirs which are good at predicting do memorize well as well. There is, however, a reentrant region of irregular reservoirs (red) which are the best at memorizing but are not optimal for predicting, in particular for the intermediate degree of difficulty (see the red loop on Fig. 6B). This point corresponds to the maximum on the right of the dip observed in Fig 5C. Interestingly, in all difficulties, reservoirs at the critical point (encircled dots) create a narrow area with a good overall performance. This picture refines our previous analysis of the impact of attractors on performance, as it seems that extinguished reservoirs can perform slightly better at memory than the others, and the same for the irregular reservoirs at prediction. Nonetheless, the difference between attractors is rather small, and it could be argued that it is insignificant.

The picture is very different for  $\sigma^* > 0$  (right column). Performances are distributed as clouds of points. In Fig. 5, we observe a large reservoir variability in both tasks. Some reservoirs are suitable for one task and bad for the other one. For intermediate and difficult tasks, some reservoirs outperform the best ones with  $\sigma^* < 0$  in both memorization and prediction. Reservoirs at the critical point are found on the right of the plots, i.e. they promise good memorization but are nonetheless widely spread, especially in prediction (vertical axis). Overall, the distinct attractors occupy the space in overlapping and indistinguishable clouds; this confirms the previous analysis that attractors do not play a role in performance.

#### 4.2.3 Discussion:

Correlations in performances are very different for both signs of  $\sigma^*$ . Choosing a reservoir with  $\sigma^* \sim -0.66$  ensures a good, general-purpose reservoir but with suboptimal performance. In contrast, going into the positive side of  $\sigma^*$  may lead to the best reservoirs in a given task or even



**FIG. 6** – The performance of all reservoirs in the prediction tasks (Mackey-Glass) as a function of their performance in the memory tasks (white-noise), for  $\sigma^* < 0$  (**A, B, C**) and  $\sigma^* > 0$  (**D, E, F**). Reservoirs are again classified according to their dominant attractor (see supplementary material S13). Each dot represents an individual reservoir performance averaged over 5 initial conditions. In all plots, dots with black edges display reservoirs taken at the critical regimes. We chose three different pairs of values for the parameters of the tasks,  $\tau$  and  $\delta$ , each representing a different difficulty level: 1. Simple (**A** and **D**) the lowest difficulty in both tasks,  $\tau = 5$  and  $\delta = -2$ ; 2. Average (**B** and **E**) average difficulty for intermediary values of  $\tau = 20$  and  $\delta = -6$ ; 3. Difficult (**C** and **F**) difficult task for higher values of  $\tau = 28$  and  $\delta = -10$ .

better general purpose reservoirs, but this comes at a price: those gems cannot be found by the statistical analysis we have performed on their free running activity.

## 5 Conclusion

One of the main issues in the field of RC is the lack of principled methodology [Rodan and Tino, 2011] for reservoir design. This article aimed to quantify the impact of the random weight generation process to better understand the relationship between connectivity, dynamics, and performance. We demonstrated that the only control parameter is the ratio  $\sigma^* = \sigma/\mu$  through a Gaussian weight distribution, which indirectly regulates the excitatory/inhibitory balance. We found two critical points and observed that reservoirs typically possess a dominant attractor, regardless of their initial states.

We investigated the relationship between the performance, the control parameter, and the preferred attractor in memory and prediction tasks. Our results reveal that  $\sigma^*$ , hence the excitatory-inhibitory balance  $b$ , has a strong impact on performance in the two considered tasks while the attractor dynamics have none. We showed how to select a control parameter region that ensures good performance, thus providing a very efficient way to obtain high-performance reservoirs. This region corresponds to high attractor diversity. For  $\sigma^* < 0$ , the critical region is narrower and does not necessarily coincide with the top of performance, while for  $\sigma^* > 0$ , the critical region corresponds to the performing region.

For the tasks, we showed that negative  $\sigma^*$  values produced superior results in prediction, with reliable performance and low reservoir-to-reservoir variability. Therefore, it is sufficient to perform free-running and pick a single value of  $\sigma^*$ , preferably close to the critical point  $\sigma_c^*$ . In contrast, positive  $\sigma^*$  values were found to have higher performance in memory tasks but with greater volatility. Since a given  $\sigma^*$  value can lead to diverse performance outcomes, generating random reservoirs and testing them during training to select the best performers is still necessary. Given enough trials, however, our findings suggest that  $\sigma^* > 0$  can generate the bests general-purpose reservoirs.

## 6 Future work

We tested the impact of dynamics on performance in two types of tasks: memory and prediction, for various time series. It would be interesting to test if and how the balance and attractor dynamic impact other types of tasks and inputs, notably classification, as it is also a standard task

in machine learning. Moreover, extending the cross-task performance analysis to classification could potentially reveal interesting insights about its performance.

Still, one surprising result is the limited impact of the intrinsic attractor dynamics on performance. One could test the robustness of this result by refining the attractor category, and future work may reveal greater performance sensitivity to attractor dynamics. For example, the extinguished category included all reservoirs with activities dying before 1000 time steps. Refining the analysis could involve correlating performance with the average time before free-running reservoir activity dies out. Similarly, cyclic reservoirs could be refined by analyzing their period [Kinoshita et al., 2009], while some irregular activities may be considered cyclic when run for more extended periods. Moreover, it is possible that combining other types of analysis, such as correlation in space and time [Metzner and Krauss, 2022], avalanche distribution size [Siddiqui et al., 2018], basins of attraction [Del Giudice et al., 1998, Chinarov and Menzinger, 2000, Kinoshita et al., 2009], the number of attractors [Cabessa and Villa, 2018], and study of the reservoir topology [Kinoshita et al., 2009, Masulli and Villa, 2015], could provide better categorization of dynamics, with ultimately better predictive power of performance.

Finally, it would be of particular interest to see if our finding regarding the impact of the excitatory/inhibitory balance and dominant attractors also applies to other models, such as the quantum Ising spins, also used in the context of RC, which exhibit analogous phase transitions [Martínez-Peña et al., 2021], and improved memory and prediction of time series in its vicinity [Kutvonen et al., 2020].

## References

- [Alexandre et al., 2009] Alexandre, L. A., Embrechts, M. J., and Linton, J. (2009). Benchmarking reservoir computing on time-independent classification tasks. In *Proceedings of the International Joint Conference on Neural Networks*, pages 89–93.
- [Balafrej et al., 2022] Balafrej, I., Alibart, F., and Rouat, J. (2022). P-CRITICAL: a reservoir autoregulation plasticity rule for neuromorphic hardware. *Neuromorphic Computing and Engineering*, 2(2):024007.
- [Benjamin et al., 2014] Benjamin, B. V., Gao, P., McQuinn, E., Choudhary, S., Chandrasekaran, A. R., Bussat, J.-M., Alvarez-Icaza, R., Arthur, J. V., Merolla, P. A., and Boahen, K. (2014).

- Neurogrid: A Mixed-Analog-Digital Multichip System for Large-Scale Neural Simulations. *Proceedings of the IEEE*, 102(5):699–716.
- [Bertschinger and Natschläger, 2004] Bertschinger, N. and Natschläger, T. (2004). Real-Time Computation at the Edge of Chaos in Recurrent Neural Networks. *Neural Computation*, 16(7):1413–1436.
- [Bianchi et al., 2016] Bianchi, F. M., Livi, L., and Alippi, C. (2016). Investigating echo state networks dynamics by means of recurrence analysis. *IEEE transactions on neural networks and learning systems*, 29(2):427–439.
- [Blackledge and Mosola, 2020] Blackledge, J. M. and Mosola, N. (2020). A Statistically Significant Test to Evaluate the Order or Disorder of a Binary String. In *2020 31st Irish Signals and Systems Conference (ISSC)*, pages 1–6. IEEE.
- [Burkow and Tufte, 2016] Burkow, A. V. and Tufte, G. (2016). *Exploring Physical Reservoir Computing using Random Boolean Networks*. PhD thesis, Norwegian University of Science and Technology - NTNU.
- [Büsing et al., 2010] Büsing, L., Schrauwen, B., and Legenstein, R. (2010). Connectivity, dynamics, and memory in reservoir computing with binary and analog neurons. *Neural Computation*, 22(5):1272–1311.
- [Cabessa and Villa, 2018] Cabessa, J. and Villa, A. E. P. (2018). Attractor dynamics of a Boolean model of a brain circuit controlled by multiple parameters. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 28(10):106318.
- [Callen and Welton, 1951] Callen, H. B. and Welton, T. A. (1951). Irreversibility and generalized noise. *Physical Review*, 83(1):34–40.
- [Canaday et al., 2018] Canaday, D., Griffith, A., and Gauthier, D. J. (2018). Rapid time series prediction with a hardware-based reservoir computer. *Chaos*, 28(12):1–13.
- [Chinarov and Menzinger, 2000] Chinarov, V. and Menzinger, M. (2000). Computational dynamics of gradient bistable networks. *Biosystems*, 55(1-3):137–142.
- [Cohen et al., 2010] Cohen, O., Keselman, A., Moses, E., Rodríguez Martínez, M., Soriano, J., and Tlusty, T. (2010). Quorum percolation in living neural networks. *EPL (Europhysics Letters)*, 89(1):18008.

- [Coniglio et al., 1976] Coniglio, A., Nappi, C. R., Peruggi, F., and Russo, L. (1976). Percolation and phase transitions in the Ising model. *Communications in Mathematical Physics*, 51(3):315–323.
- [Croll, 2014] Croll, G. J. (2014). BiEntropy – the Measurement and Algebras of Order and Disorder in Finite Binary Strings. In John C Amson (University of St Andrews, U. and Louis H Kauffman (University of Illinois at Chicago, U., editors, *Scientific Essays in Honor of H Pierre Noyes on the Occasion of His 90th Birthday*, pages 48–64. Word Scientific, series on edition.
- [Del Giudice et al., 1998] Del Giudice, P., Fusi, S., Badoni, D., Dante, V., and Amit, D. J. (1998). Learning attractors in an asynchronous, stochastic electronic neural network. *Network: Computation in Neural Systems*, 9(2):183–205.
- [Del Papa et al., 2017] Del Papa, B., Priesemann, V., and Triesch, J. (2017). Criticality meets learning: Criticality signatures in a self-organizing recurrent neural network. *PLoS ONE*, 12(5):1–21.
- [Echlin et al., 2018] Echlin, M., Aguilar, B., Notarangelo, M., Gibbs, D., and Shmulevich, I. (2018). Flexibility of Boolean Network Reservoir Computers in Approximating Arbitrary Recursive and Non-Recursive Binary Filters. *Entropy*, 20(12):954.
- [Goudarzi et al., 2016] Goudarzi, A., Marzen, S., Banda, P., Feldman, G., Teuscher, C., and Stefanovic, D. (2016). Memory and Information Processing in Recurrent Neural Networks. 0(3):1–8.
- [Hajnal and Lörincz, 2006] Hajnal, M. A. and Lörincz, A. (2006). Critical echo state networks. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 4131 LNCS(September 2006):658–667.
- [Hämäläinen et al., 1993] Hämäläinen, M., Hari, R., Ilmoniemi, R. J., Knuutila, J., and Lounasmaa, O. V. (1993). Magnetoencephalography theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of Modern Physics*, 65(2).
- [Hopfield, 1982] 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).

- [Ising, 1925] Ising, E. (1925). Beitrag zur Theorie des Ferromagnetismus. *Zeitschrift für Physik*, 31(1).
- [Jaeger, 2001] Jaeger, H. (2001). The “echo state” approach to analysing and training recurrent neural networks-with an erratum note’.
- [Jaeger and Haas, 2004] Jaeger, H. and Haas, H. (2004). Harnessing Nonlinearity: Predicting Chaotic Systems and Saving Energy in Wireless Communication. *Science*, 304(5667).
- [Karimipanah et al., 2017] Karimipanah, Y., Ma, Z., and Wessel, R. (2017). Criticality predicts maximum irregularity in recurrent networks of excitatory nodes. *PLOS ONE*, 12(8):e0182501.
- [Kingma and Ba, 2015] Kingma, D. P. and Ba, J. (2015). Adam: A Method for Stochastic Optimization. In *3rd International Conference for Learning Representations, San Diego*.
- [Kinoshita et al., 2009] Kinoshita, S.-i., Iguchi, K., and Yamada, H. S. (2009). Intrinsic properties of Boolean dynamics in complex networks. *Journal of Theoretical Biology*, 256(3):351–369.
- [Kinouchi and Copelli, 2006] Kinouchi, O. and Copelli, M. (2006). Optimal dynamical range of excitable networks at criticality. *Nature Physics*, 2(5):348–351.
- [Komkov et al., 2021] Komkov, H., Pocher, L., Restelli, A., Hunt, B., and Lathrop, D. (2021). RF Signal Classification using Boolean Reservoir Computing on an FPGA. *Proceedings of the International Joint Conference on Neural Networks*, 2021-July.
- [Krauss et al., 2019a] Krauss, P., Schuster, M., Dietrich, V., Schilling, A., Schulze, H., and Metzner, C. (2019a). Weight statistics controls dynamics in recurrent neural networks. *PLOS ONE*, 14(4):e0214541.
- [Krauss et al., 2019b] Krauss, P., Zankl, A., Schilling, A., Schulze, H., and Metzner, C. (2019b). Analysis of structure and dynamics in three-neuron motifs. *Frontiers in Computational Neuroscience*, 13(February):1–6.
- [Kuśmierz et al., 2020] Kuśmierz, , Ogawa, S., and Toyoizumi, T. (2020). Edge of Chaos and Avalanches in Neural Networks with Heavy-Tailed Synaptic Weight Distribution. *Physical Review Letters*, 125(2):028101.
- [Kutvonen et al., 2020] Kutvonen, A., Fujii, K., and Sagawa, T. (2020). Optimizing a quantum reservoir computer for time series prediction. *Scientific Reports*, 10(1):1–8.

- [Lavis et al., 2021] Lavis, D. A., Kühn, R., and Frigg, R. (2021). Becoming Large, Becoming Infinite: The Anatomy of Thermal Physics and Phase Transitions in Finite Systems. *Foundations of Physics*, 51(5):1–69.
- [Lewin and Bak, 1993] Lewin, R. and Bak, P. (1993). Complexity: Life at the Edge of Chaos. *American Journal of Physics*, 61(8):764–765.
- [Maass et al., 2002] Maass, W., Natschläger, T., Markram, H., Maass, Natschlager, and Markram (2002). Real-time Computing Without Stable States: A New Framework for Neural Computation Based on Perturbations. *NEURCOMP: Neural Computation*, 14(11):2531–2560.
- [Mamun et al., 2016] Mamun, M. S. I., Ghorbani, A. A., and Stakhanova, N. (2016). An Entropy Based Encrypted Traffic Classifier. In *Information and Communications Security. ICICS 2015. Lecture Notes in Computer Science*, volume 9543, pages 282–294.
- [Martínez-Peña et al., 2021] Martínez-Peña, R., Giorgi, G. L., Nokkala, J., Soriano, M. C., and Zambrini, R. (2021). Dynamical phase transitions in quantum reservoir computing. *Physical Review Letters*, 127(10):1–14.
- [Masulli and Villa, 2015] Masulli, P. and Villa, A. E. P. (2015). The topology of the directed clique complex as a network invariant. *SpringerPlus*, 5(1):388.
- [Merolla et al., 2014] Merolla, P. A., Arthur, J. V., Alvarez-Icaza, R., Cassidy, A. S., Sawada, J., Akopyan, F., Jackson, B. L., Imam, N., Guo, C., Nakamura, Y., Brezzo, B., Vo, I., Esser, S. K., Appuswamy, R., Taba, B., Amir, A., Flickner, M. D., Risk, W. P., Manohar, R., and Modha, D. S. (2014). A million spiking-neuron integrated circuit with a scalable communication network and interface. *Science*, 345(6197):668–673.
- [Metzner and Krauss, 2022] Metzner, C. and Krauss, P. (2022). Dynamics and Information Import in Recurrent Neural Networks. *Frontiers in Computational Neuroscience*, 16(April):1–15.
- [Natschläger et al., 2005] Natschläger, T., Bertschinger, N., and Legenstein, R. (2005). At the edge of chaos: Real-time computations and self-organized criticality in recurrent neural networks. In *Advances in Neural Information Processing Systems 17*.
- [Oztuik et al., 2007] Oztuik, M. C., Xu, D., and Principe, J. C. (2007). Analysis and design of echo state networks. *Neural Computation*, 19(1):111–138.

- [Rodan and Tino, 2011] Rodan, A. and Tino, P. (2011). Minimum Complexity Echo State Network. *IEEE Transactions on Neural Networks*, 22(1):131–144.
- [Rosin, 2015] Rosin, D. P. (2015). *Dynamics of Complex Autonomous Boolean Networks*. Springer Theses. Springer International Publishing, Cham.
- [Schrauwen et al., 2007] Schrauwen, B., Verstraeten, D., and Van Campenhout, J. (2007). An overview of reservoir computing: Theory, applications and implementations. In *ESANN 2007 Proceedings - 15th European Symposium on Artificial Neural Networks*, pages 471–482.
- [Seifert and Reggia, 2015] Seifert, J. and Reggia, J. A. (2015). Lambda and the Edge of Chaos in Recurrent Neural Networks. *Artificial Life*, 21(1):55–71.
- [Sherrington and Kirkpatrick, 1975] Sherrington, D. and Kirkpatrick, S. (1975). Solvable model of a spin-glass. *Physical Review Letters*, 35(26).
- [Siddiqui et al., 2018] Siddiqui, M., Wedemann, R. S., and Jensen, H. J. (2018). Avalanches and generalized memory associativity in a network model for conscious and unconscious mental functioning. *Physica A: Statistical Mechanics and its Applications*, 490:127–138.
- [Snyder et al., 2012] Snyder, D., Goudarzi, A., and Teuscher, C. (2012). Computational Capabilities of Random Automata Networks for Reservoir Computing. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 87(4):1–8.
- [Tanaka et al., 2019] Tanaka, G., Yamane, T., Héroux, J. B., Nakane, R., Kanazawa, N., Takeda, S., Numata, H., Nakano, D., and Hirose, A. (2019). Recent advances in physical reservoir computing: A review. *Neural Networks*, 115:100–123.
- [Tieck et al., 2018] Tieck, J. C. V., Pogančić, M. V., Kaiser, J., Roennau, A., Gewaltig, M.-O., and Dillmann, R. (2018). Learning Continuous Muscle Control for a Multi-joint Arm by Extending Proximal Policy Optimization with a Liquid State Machine. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 11139 LNCS, pages 211–221.
- [Zhou and Zeng, 2022] Zhou, J. and Zeng, X. (2022). Physical Layer Secret Key Generation for Spatially Correlated Channels Based on Multi-Task Autoencoder. *2022 7th International Conference on Intelligent Computing and Signal Processing, ICSP 2022*, pages 144–150.

## Supplementary Material

## 7 The experiments

This article proposes two sets of experiments: in the first setting, one is interested in the free time evolution of the reservoir without input ( $u_i = 0$ ). The second set is devoted to the measurement of the capacity of the reservoir to execute two tasks.

### 7.1 Free running network

Both experiments in this section are based on the same simulation, and do not require the input and readout layers:

1. At time  $t = 0$  the network state is randomly initialized with provided *seed*, and a fixed 20% of active nodes ( $x_i = 1$ ).
2. Then the network is let run for a duration  $D = 2000$  time steps.

For each value of  $\sigma^*$ , we generate 100 reservoirs, i.e. reservoirs having all the same architecture but different weights. We consider two ways to perform statistics.

**Statistics over many reservoirs** (section 3.1).

Each **reservoir** is run once in a *free run simulation*. We report the result of 90 values of  $\sigma^* < 0$ , 110 values of  $\sigma^* > 0$ , and 80 for  $\mu = 0$ , for a total of 28,000 simulations.

**Statistics of reservoirs** (section 3.2).

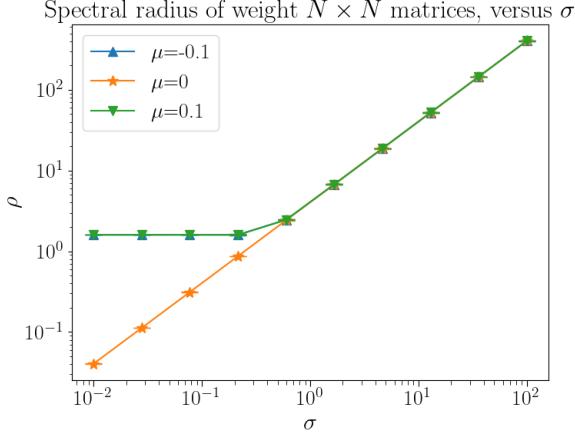
Each **reservoir** is run 100 times with different initial conditions. We report the result of 80 values of  $\sigma^*$  for each sign, for a total of 1,600,000 simulations.

## 8 The control parameter

### 8.1 Link between $\sigma^*$ and $\rho$

The spectral radius  $\rho$  of a square matrix  $W$  is the largest of its eigenvalues in absolute value. We will consider the case where the size  $N$  of the network, here  $N = 10,000$  is large enough so that  $\rho$  is self-averaging: it does not depend on the exact coefficients of  $W$  but simply on the parameters  $(\mu, \sigma)$  of their distribution. It is showcased in Fig. 7 where the (inexistent) error bars represent the spread in the results for the calculation of  $\rho$  for different reservoirs of identical  $(\mu, \sigma)$ . One can thus consider  $\rho$  to be a function of  $\mu$  and  $\sigma$ .

The property  $\rho(\lambda W) = |\lambda|\rho(W)$  for any real  $\lambda$  leads to  $\rho(\lambda\mu, \lambda\sigma) = |\lambda|\rho(\mu, \sigma)$ . In the case  $\mu = 0$ , taking  $\lambda = 1/\sigma$  gives  $\rho(0, \sigma) = \alpha|\sigma|$  with  $\alpha = \rho(0, 1)$  a parameter independent of the



**FIG. 7** – The spectral radius  $\rho$ , of Gaussian weight matrices, computed for different values of the mean  $\mu(W)$ , and as a function of the standard deviation  $\sigma(W)$ . Each dot represents the average computed on 10 matrices, and errorbars represent two standard deviations.

weight matrix. We thus find that while the dynamics of the network is independent of  $\sigma$  for  $\mu = 0$  (see Fig. 3),  $\rho$  can take any value. In this case, there is no link between  $\rho$  and the dynamics of the network.

For  $\mu \neq 0$ , taking  $\lambda = 1/\mu$  gives  $\rho(\mu, \sigma) = |\mu|\beta(\sigma^*)$  with  $\beta(s) = \rho(1, s)$ . Thus for a given value of  $\sigma^*$ ,  $\rho$  can take any value. There is however a direct link between  $\beta$  and  $\sigma^*$ , shown in Fig. 7. For  $\sigma = 0$ , the weight matrix is  $W = \mu A$  with  $A$  a matrix with  $K$  ones and  $N - K$  zeroes in each row, thus obeying  $\rho(A) = \beta(0) = K$ , leading to  $\rho(\mu, 0) = |\mu|K$ , in agreement with our numerical simulation:  $\rho$  saturates at low  $\sigma$ . For the limit of large  $\sigma$  taking  $\lambda = 1/\sigma$  gives  $\rho(\mu, \sigma) = \sigma\rho(1/\sigma^*, 1) \simeq \sigma\rho(0, 1) = \alpha\sigma$ , which is the result for  $\mu = 0$ , as observed in Fig. 7.

While it may be tempting to replace  $\rho$  by its renormalized value  $\beta$ , both quantities suffer from the same weakness: they do not depend on the sign of  $\sigma^*$ , whereas the dynamics of the network strongly do.

## 8.2 Control parameter values

Below are the parameter values for the experiment of free running (see methodology in S7.1):

Start	End	Step
0.01	0.06	0.001
0.06	0.1	0.0007
0.07	0.1	0.001
0.1	0.2	0.01
0.2	1	0.1
1	11	1

**TAB. 1** – Values of  $\sigma(W)$  for the first experiment,  $\mu = -0.1$ .

Start	End	Step
0.01	0.07	0.001
0.07	0.2	0.01
0.2	0.5	0.02
0.5	1	0.1
1	11	1

**TAB. 2** – Values of  $\sigma(W)$  for the first experiment,  $\mu = 0.1$ .

### 8.3 Performance of tasks

All experiments in this section test the performance in a task, where learning is required for the readout layer. For all tasks, the following framework is applied. Assuming that  $F$  is the reservoir, the relationship between the input time series  $u(t)$  and the output layer  $y$  is  $y(t) = F(u(t))$ . The goal here is to learn the target  $T(t)$ , such that  $y(t) = T(t)$  and the target is set to:

$$T(t) = u(t + \delta) \quad (8)$$

Here,  $\delta$  is an integer which represents a number of time steps. This parameter serves the general purpose of setting the type of task:

- **Memory task** for  $\delta < 0$ , the reservoir output must reproduce the input received in the past.
- **Prediction task** for  $\delta > 0$ , the network output must produce an input not yet seen by the reservoir.

The (integer) parameter  $\delta$  also sets the difficulty of the task: the higher in absolute value, the more demanding it is.

**White noise memory** (section 4).  $u(t)$  is a zero mean, unit variance white i.i.d noise. Successive inputs are uncorrelated, so prediction is not involved, only memory. We report results for  $\delta \in \{-14, -10, -6, -2\}$  (see figure 4.A for an illustration).

**Prediction of Mackey-Glass series** (section 4).  $u(t)$  is the Mackey-Glass time series [Hajnal and Lörincz, 2006], which is common to benchmark reservoir computational capabilities. It is given by the following dynamical equation:

$$x_{t+1} = ax_t + \frac{bx_{t-\tau}}{c + x_{t-\tau}^d} \quad (9)$$

We choose  $a = 0.9$ ,  $b = 0.2$ ,  $c = 0.9$ ,  $d = 10$  and  $x_0 = 0.1$ . The dynamic of this equation can be controlled by varying  $\tau$ : as  $\tau$  increases, the time series evolves from periodic ( $\tau = 5$ ) to chaotic ( $\tau = 28$ ), with a continuous increase in complexity in between. Results are for a fixed  $\delta = 10$  and  $\tau = \{5, 15, 20, 28\}$  (see figure 4.B for an illustration).

## 8.4 Training

The two tasks follow the same protocol for training the readout:

- The network receives the input  $u$  for duration  $D = 2000$  time steps.
- The first 500 time steps are discarded and considered transients. This value is empirically obtained, adapted from the signal at hand, and coupled to a convergence test for permanent regime detection.
- The training is performed on the following 1500 time steps on the concatenated in-time reservoir outputs, using the optimization procedure described in 2.3.

Each experiment consists in 40 values of  $\sigma^*$ , 100 reservoirs per  $\sigma^*$  value, and each network is run 5 times with different randomly tossed inputs (i.e., 40,000 simulations). Each training is performed for 4000 epochs (with a total of 640,000,000 training epochs).

## 8.5 Performance

The metrics of performance are given by the *Pearson correlation coefficient* between the output and the target (each of length 1500). A perfect match corresponds to a correlation of 1 while 0 means an output of the network is not better than random.

# 9 Measure of networks dynamics

To evaluate the reservoir dynamics, we used the most straightforward measure one can imagine: the *activity*, which we defined as the sum of all spikes at each given time step, see supplementary material 10. We will take some time to make the reader appreciate the usefulness of such an approach. The sum of binary spikes is mathematically equivalent to how the magnetic field of spins in the ISING models [Ising, 1925] is computed. In ISING, the electron spins are modelled as miniature magnets that can take binary values  $[-1, 1]$ ), which flip with some probability

depending on temperature and the coupling with other spins. The ISING model is one of the closest physical models to neural networks. In fact, it has been widely used to design neural networks, for example, in the famous Hopfield network [Hopfield, 1982] or the Boltzmann machine [Sherrington and Kirkpatrick, 1975]. So one could argue that our activity signal is close to what Magnetoencephalography (MEG) is for brain activity [Hämäläinen et al., 1993].

Since this analysis does not depend on the micro-constituents, we make the case that the methodology performed in this article is easily transferable to other neural models and to other fields, such as neurosciences, physics, and the study of complex dynamical systems.

**Activity:** we define  $A(t)$ , the averaged activity of the network at time  $t$ , and for simplicity we will refer to it as *activity*. It is the normalized sum of all neural states  $x_i(t)$ :

$$A(t) = \frac{1}{N} \sum_{i=1}^N x_i(t)$$

**Steady state of reservoir:** In the main text (part 3.1), we defined the steady state as  $A$  for ease of notation, here for the sake of clarity, we define the steady activity  $A_s$  of free-running reservoirs during  $D = 2000$  time steps, as  $A_s = A(t \geq D/2)$ , the activity of the last  $D - D/2 = 1000$  time steps of free running simulations.

## 10 Metrics

### 10.1 Permanent regime statistics:

here are two statistical analyses we perform on the measure of the steady state activity  $A_s$  of a given reservoir, where  $\overline{(.)}$  represents the time-average over a quantity, and  $\delta A(t) = A(t) - \bar{A}$ :

- Permanent regime time-average:

$$\bar{A}_s = \frac{2}{D} \sum_{t=D/2}^D A(t)$$

- Permanent regime time-variance:

$$\overline{\delta A_s^2} = \frac{2}{D} \sum_{t=D/2}^D (A(t) - \bar{A}_s)^2$$

**Average over reservoirs of permanent regime statistics:** in the *first experiment* (see results in section 3.1), for each value  $\sigma^*$ , we compute the average (over reservoirs generated with 100 different seeds) of the permanent regime *average* and *variance* (overtime time step  $t$ ). Here  $A_s^r$  denotes the stationary activity of the reservoir of index  $r$  generated with  $seed = r \times 100$ , for a total number of reservoir  $R = 100$ . We compute the average over reservoirs of the permanent regime statistics, where the back  $\langle . \rangle$  denotes the average over reservoirs:

- Average over reservoirs of the permanent regime average:

$$\langle \bar{A}_s \rangle = \frac{1}{R} \sum_{r=1}^R \bar{A}_s^r$$

- Average over reservoirs of the permanent regime variance:

$$\langle \overline{\delta A_s^2} \rangle = \frac{1}{R} \sum_{r=1}^R \overline{(\delta A_s^r)^2}$$

**Statistics of the BiEntropy:** the BiEntropy solves one important limitation of the Shannon entropy, regarding the evaluation of binary words: as an example, let us take these two binary words "01010101" and "10110010". Although one is fully periodic and the other is somewhat random, they both have the same probability of occurrence of 1's and 0's; hence they have the same Shannon entropy. The Shannon entropy is thus insufficient to inform on the regularity versus the disorder of binary words. In contradiction, the binary entropy has been designed to discriminate patterns that the Shannon entropy could not [Blackledge and Mosola, 2020]. We argue that in our context this metric is perfectly fitted to evaluate the binary words of spike time patterns. Since phase transition in such systems are known to be between an ordered and disordered phase, while the critical regime, also known as the edge of chaos, is supposedly a mixture of both. One property of particular interest to us regarding that metric, is that it is bounded between 0 (for perfectly ordered words), and 1 (totally disordered).\*

- BiEntropy of permanent regime: the BiEntropy  $H_b$  is computed on binarized steady activity  $\delta A_s$ . First, we subtract the mean  $A_s - \bar{A}_s$ , and then all positive values are clamped to one and negative values to zero, resulting in a binary sequence. The  $H_b$  is then computed on this sequence after converting it to a string.

---

\*For example, the word "01010101" has  $H_b = 0.0078$ , and the word "10110010" has  $H_b = 0.7596$ . A convenient way to discriminate both words.

- Average over reservoirs of the permanent regime BiEntropy:

$$\langle H_b \rangle = \frac{1}{R} \sum_{r=1}^R H_b^r$$

- Variance over reservoirs of the permanent regime BiEntropy:

$$\langle \delta H_b^2 \rangle = \frac{1}{R} \sum_{r=1}^R (H_b^r - \langle H_b \rangle)^2$$

NB: given that neuron states are binary, one could wonder why employing a binarized version of a continuous variable like  $A$ . The reason is twofold: 1. the number of neurons in the network is  $N = 10000$ . Therefore, computing  $H_b$  on all neurons would be extremely costly. As such, reducing the number of neurons would be crucial, though it would necessitate a criterion for selection. This poses the risk of missing important information or introducing biases. 2. This would not be applicable in many real-world applications where access to the micro-constituents of the reservoir is difficult or even impossible. As such, we ensure our methodology is easily transferable to other areas comprising non-invasive studies of the brain.

## 11 Classification of attractors

In this section, we provide a more grounded explanation of the choice of attractor categories. In the main text (part 3.2), we developed a scheme to categorize activities by their respective attractor. Fig. 4 showed the histograms of attractors as a function of  $\sigma^*$ , and in this section, we provide a refined version of these statistics by adding two more categories of attractor:

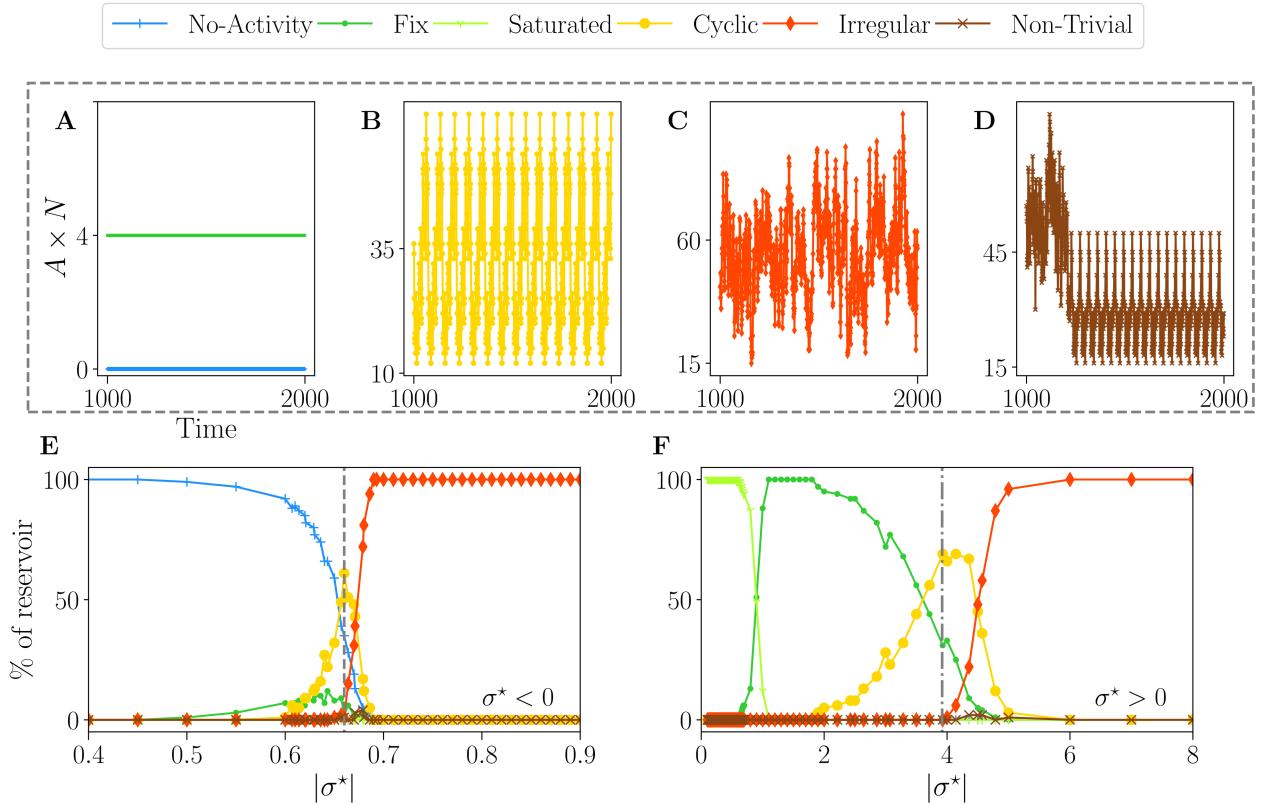
- *Saturated attractors*: The steady activity is saturated when all neurons are active at all times.
- *Non-trivial*: Any signal whose category of dynamics changes over time and that does not fit in previously mentioned types is considered non-trivial. This comprises cases where during the time window considered, activity suddenly changes from one type of dynamic to another.

We show in Fig. 8 examples of the various activities belonging to each category, *No-Activity* and *Fix* (**A**), *Cyclic* (**B**), *Irregular* (**C**), and lastly, *Non-Trivial* (**D**). In the following paragraphs,

we explain in more detail some specificities of the *fixed* point attractor category and the reason why did not treat the *Non-trivial* case separately from the irregular ones. First, we must mention that the *fixed* point attractor category could, in theory, encompass the *extinguished* and *saturated* cases as well. This is because they all fit inside the definition of a time derivative of zero. The reason why we separated inactive reservoirs from the two others comes from percolation theory [Coniglio et al., 1976]: formally, a reservoir has not percolated if the activity does not spread to infinity in space and/or time. While at the percolation threshold, activity will start to propagate indefinitely. We make the case that the percolation threshold, which does not coincide here with the critical points, constitutes another type of transition [Cohen et al., 2010], from inactive, to active, hence the distinction. This is visible in Fig. 8.E, as the fixed point attractors appear a bit before the *cyclic* ones, around  $\sigma^* \sim 0.5$ .

Second, in Fig. 8.F, we can see that the transition from *Saturated* to *Fix* attractor happens very early (in terms of  $\sigma^*$ ), compared to the rest of the attractors. So early, in fact, that it is not part of the phase transition. As you might recall, the analysis performed in part 3.1 revealed that this region has a zero variance of both activity and BiEntropy. We conclude that the transition from *Saturated* to *Fix* is not related to a change of dynamics but only to a change in amplitude. As a result, we have chosen not to differentiate the two.

Thirdly, as one can clearly see in both Fig. 8.E and F, the *Non-trivial* dynamics are very rare. As such, it is worth mentioning that in the analysis performed in S13, where we categorize reservoirs depending on their dominant attractors, not a single reservoir exhibits a dominance of *Non-Trivial* dynamics. This is important because it means this category of attractors is irrelevant for finding correlations between dominant attractors and performance.



**FIG. 8** – Examples of activities, classified into their respective category of attractor: (A) no-activity (+), fix (•); (B) cyclic (●), (C) irregular (◆), and (D) non-trivial (×). The percentage of trials (or activities) belonging to each category of attractors is plotted against the control parameter  $\sigma^*$ . (A) the phase transition for  $\sigma^* < 0$ , and (B) the phase transition obtained with  $\sigma^* > 0$ . In both cases, we zoom on the critical regions since the ordered, and irregular regimes have constant statistics. E and F: the statistics of attractors of free running simulations are computed for each value of  $\sigma^*$ , for which 100 reservoirs are run once (see methodology 7.1). (E)  $\sigma^* < 0$ , and (F)  $\sigma^* > 0$ .

## 12 Diversity of reservoir attractor distributions

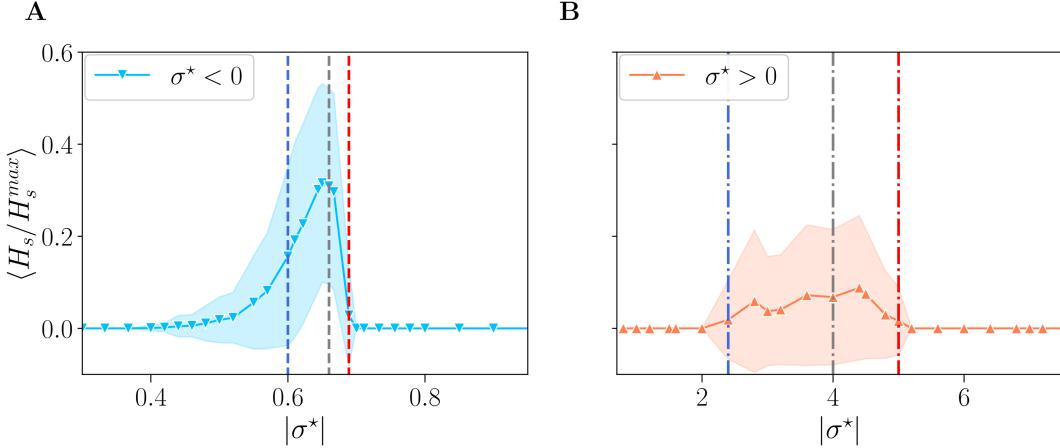
We compute the Shanon entropy  $H_s$ , with the goal of quantifying how varied are the attractor distributions of given reservoirs. Typically, if a reservoir activity always falls into one attractor, irrespective of the initial state, the Shanon entropy will give 0. On the other hand, the maximum entropy is obtained for a uniform distribution where each attractor is obtained 1/4 of the time, and  $H_s^{max} = -\log(1/4)$ . The Shanon Entropy is therefore normalized by its maximum value and averaged over the 100 reservoir of each value of  $\sigma^*$ .

To quantify reservoir dynamics diversity, we plotted the average normalized entropy of reservoir attractor distributions,  $\langle H_s/H_s^{max} \rangle$ , against the control parameter for  $\sigma^* < 0$  (**A**) and  $\sigma^* > 0$  (**B**). The shaded area represents one standard deviation. Lower values of this quantity indicate a stronger dominance of one attractor. Outside of the critical region, initial conditions are irrelevant, and reservoirs always converge to a single attractor. The region where  $\langle H_s \rangle$  increases correspond precisely to the region where the BiEntropy, shown in Fig. 9, is nonzero. In this region, different reservoirs may lead to different attractors. There is significantly more entropy, i.e. more competition between attractors, for  $\sigma^* < 0$ .

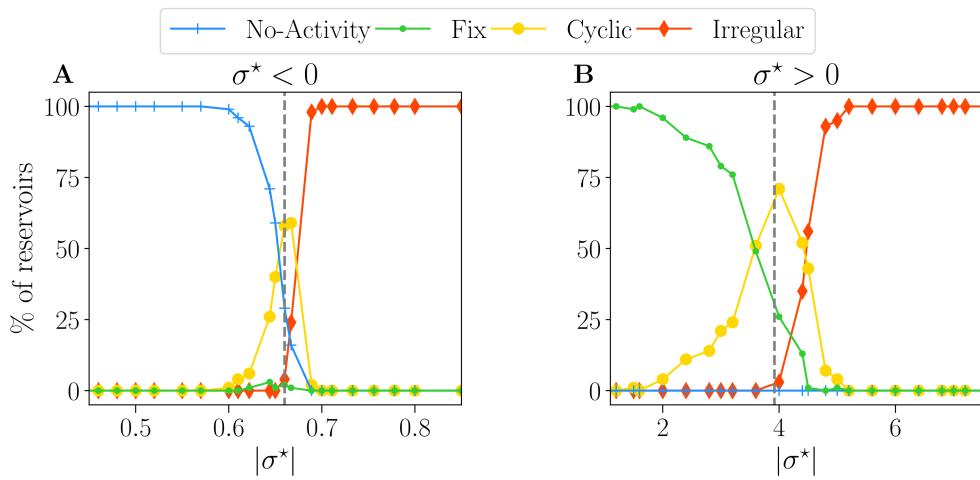
The higher variance in  $H_s$  also implies increased reservoir-to-reservoir variability. This variability is more pronounced for  $\sigma^* < 0$ , with the average entropy exceeding 0.25 near the critical point and the shaded area approaching 0.5. For  $\sigma^* > 0$ , reservoirs generally exhibit stronger dominant attractors, as the standard deviation never surpasses 0.25. Despite these differences, the normalized entropy remains consistently below 0.5 for all values of  $\sigma^*$ , irrespective of the sign. This value approximately corresponds to an attractor distribution where one attractor dominates 80% of the initial conditions, indicating that most reservoirs possess a peaked distribution with a predominant attractor.

## 13 Statistic of dominant attractors

Fig. 10 displays the statistics of the dominant attractors of each of the 100 reservoirs generated by  $\sigma^*$  values. As one can note, the statistics are also unchanged from Fig. 4. Taken together, these results indicate strong statistical robustness, as averaging over reservoirs is almost equivalent to averaging reservoirs themselves.



**FIG. 9** – The normalized Shanon entropy of individual reservoir attractor distributions. **A** and **B**: The average over reservoirs entropy of the entropy  $H_s$  versus the control parameter  $|\sigma^*|$ , for  $\sigma^* < 0$  (**A**) and  $\sigma^* > 0$  (**B**). The shaded area represents one standard deviation. The coloured hashed lines correspond to the boxes of Fig. 4: (blue)  $\sigma^* = -0.6$  (**A**) and  $\sigma^* = 2.4$  (**D**), (gray)  $\sigma^* = -0.66$  (**B**) and  $\sigma^* = 4.0$  (**E**), (red)  $\sigma^* = -0.689$  (**C**) and  $\sigma^* = 5.0$  (**F**).



**FIG. 10** – The statistics of reservoir dominant attractors. For each value of  $\sigma^*$ , 100 reservoirs are run 100 times with different initial conditions (see S7.1). The resulting activities are classified into the category of attractors: no-activity (+), fix (●), cyclic (○), and irregular (♦). The statistics of reservoirs with specified dominant attractor is plotted against the control parameter  $\sigma^*$ . (**A**)  $\sigma^* < 0$ , (**B**)  $\sigma^* > 0$ .